

CA MICS® Resource Management

Capacity Planner Option Guide

Release 12.9



This Documentation, which includes embedded help systems and electronically distributed materials, (hereinafter referred to as the "Documentation") is for your informational purposes only and is subject to change or withdrawal by CA at any time. This Documentation is proprietary information of CA and may not be copied, transferred, reproduced, disclosed, modified or duplicated, in whole or in part, without the prior written consent of CA.

If you are a licensed user of the software product(s) addressed in the Documentation, you may print or otherwise make available a reasonable number of copies of the Documentation for internal use by you and your employees in connection with that software, provided that all CA copyright notices and legends are affixed to each reproduced copy.

The right to print or otherwise make available copies of the Documentation is limited to the period during which the applicable license for such software remains in full force and effect. Should the license terminate for any reason, it is your responsibility to certify in writing to CA that all copies and partial copies of the Documentation have been returned to CA or destroyed.

TO THE EXTENT PERMITTED BY APPLICABLE LAW, CA PROVIDES THIS DOCUMENTATION "AS IS" WITHOUT WARRANTY OF ANY KIND, INCLUDING WITHOUT LIMITATION, ANY IMPLIED WARRANTIES OF MERCHANTABILITY, FITNESS FOR A PARTICULAR PURPOSE, OR NONINFRINGEMENT. IN NO EVENT WILL CA BE LIABLE TO YOU OR ANY THIRD PARTY FOR ANY LOSS OR DAMAGE, DIRECT OR INDIRECT, FROM THE USE OF THIS DOCUMENTATION, INCLUDING WITHOUT LIMITATION, LOST PROFITS, LOST INVESTMENT, BUSINESS INTERRUPTION, GOODWILL, OR LOST DATA, EVEN IF CA IS EXPRESSLY ADVISED IN ADVANCE OF THE POSSIBILITY OF SUCH LOSS OR DAMAGE.

The use of any software product referenced in the Documentation is governed by the applicable license agreement and such license agreement is not modified in any way by the terms of this notice.

The manufacturer of this Documentation is CA.

Provided with "Restricted Rights." Use, duplication or disclosure by the United States Government is subject to the restrictions set forth in FAR Sections 12.212, 52.227-14, and 52.227-19(c)(1) - (2) and DFARS Section 252.227-7014(b)(3), as applicable, or their successors.

Copyright © 2014 CA. All rights reserved. All trademarks, trade names, service marks, and logos referenced herein belong to their respective companies.

Contact CA Technologies

Contact CA Support

For your convenience, CA Technologies provides one site where you can access the information that you need for your Home Office, Small Business, and Enterprise CA Technologies products. At <http://ca.com/support>, you can access the following resources:

- Online and telephone contact information for technical assistance and customer services
- Information about user communities and forums
- Product and documentation downloads
- CA Support policies and guidelines
- Other helpful resources appropriate for your product

Providing Feedback About Product Documentation

If you have comments or questions about CA Technologies product documentation, you can send a message to techpubs@ca.com.

To provide feedback about CA Technologies product documentation, complete our short customer survey which is available on the CA Support website at <http://ca.com/docs>.

Contents

Chapter 1: OVERVIEW	15
1.1 Primary Areas of Application.....	17
1.2 Capacity Planning Database	20
1.3 Analysis Capabilities	22
1.5 Online Facility Tutorial	25
1.5.1 Introduction to the Online Facility	25
1.5.2 Getting Started with the Online Facility	30
1.5.3 Capacity Planning Database File Definition Inquiries.....	34
1.5.4 Capacity Planning Analysis Inquiries	51
1.5.5 Replaying Cataloged Output	64
1.5.6 Capacity Planning Facilities Administration	64
1.5.7 Production Reporting Administration	75
1.5.8 Worksheet Facility.....	82
1.5.9 Selecting Files for MICF Analysis	83
1.5.10 Placing Analysis Inquiries in the SHARED Catalog	84
 Chapter 2: SAMPLE PLANNING APPLICATIONS	 87
2.1 z/OS Processor Planning.....	93
2.1.1 Usage Guidelines.....	94
2.1.2 Control Parameters	97
2.1.3 Standard Output	105
2.2 WLM Workload Planning.....	109
2.2.1 Usage Guidelines.....	110
2.2.2 Control Parameters	113
2.2.3 Standard Output	124
2.2.4 WLM Workload Planning Example	133
2.3 WLC Implementation Planning.....	138
2.3.1 Usage Guidelines.....	139
2.3.2 Control Parameters	140
2.3.3 Standard Output	148
2.4 IRD Implementation Planning	151
2.4.1 Usage Guidelines.....	152
2.4.2 Control Parameters	153
2.4.3 Standard Output	162
2.5 XCF Resource Planning	165
2.5.1 Usage Guidelines.....	166

2.5.2 Control Parameters	167
2.5.3 Standard Output	173
2.6 Coupling Facility Planning	174
2.6.1 Usage Guidelines	175
2.6.2 Control Parameters	176
2.6.3 Standard Output	182
2.7 z/VM Processor Planning	184
2.7.1 Usage Guidelines	184
2.7.2 Control Parameters	185
2.7.3 Standard Output	191
2.8 SMS Resource Planning	195
2.8.1 Usage Guidelines	195
2.8.2 Control Parameters	196
2.8.3 Standard Output	203
2.9 CICS Resource Planning	207
2.9.1 Usage Guidelines	207
2.9.2 Control Parameters	208
2.9.3 Standard Output	214
2.10 IMS Resource Planning	215
2.10.1 Usage Guidelines	216
2.10.2 Control Parameters	217
2.10.3 Standard Output	223
2.11 DB2 Resource Planning	224
2.11.1 Usage Guidelines	225
2.11.2 Control Parameters	226
2.11.3 Standard Output	232
2.12 WebSphere MQSeries Planning	236
2.12.1 Usage Guidelines	236
2.12.2 Control Parameters	237
2.12.3 Standard Output	243
2.13 WEB Server Resource Planning	245
2.13.1 Usage Guidelines	245
2.13.2 Control Parameters	246
2.13.3 Standard Output	252
2.14 TCP/IP Network Resource Planning	253
2.14.1 Usage Guidelines	253
2.14.2 Control Parameters	255
2.14.3 Standard Output	261
2.15 Tandem CPU Resource Planning	262
2.15.1 Usage Guidelines	262
2.15.2 Control Parameters	263
2.15.3 Standard Output	269

2.15.4 Data Elements Overview	273
2.16 VAX/VMS CPU Resource Planning	279
2.16.1 Usage Guidelines	279
2.16.2 Control Parameters	283
2.16.3 Standard Output	290
2.17 HP Openview UNIX Resource Planning	294
2.17.1 Usage Guidelines	294
2.17.2 Control Parameters	295
2.17.3 Standard Output	301
2.18 Tape Data Set Resource Planning	303
2.18.1 Usage Guidelines	303
2.18.2 Control Parameters	305
2.18.3 Standard Output	311

Chapter 3: CAPACITY PLANNING DATABASE 313

3.1 Structural Description	314
3.2 Usage Guidelines	317
3.3 Resource Element Files	321
3.3.1 Usage Guidelines	322
3.3.2 Screen Operation	330
3.3.3 Resource Element File Case Studies	367
3.4 Meta Files	446
3.4.1 Usage Guidelines	446
3.4.2 Screen Operation	448
3.4.3 Meta File Case Studies	478
3.5 Business Element Files	511
3.5.1 Usage Guidelines	511
3.5.2 Screen Operation	514

Chapter 4: PRESENTING FORECASTED DATA 525

4.1 Functional Description	525
4.2 Usage Guidelines	526
4.3 Standard Output	526
4.4 Presentation Technique Tutorial	529
4.5 Component Operation	531
4.5.1 Presentation Graphics Control Screen	532
4.5.2 Color Graphics Format Parameters	537
4.5.3 Y-Axis Reference Lines	538
4.6 Case Studies	540
4.6.1 Extended Example I/O Workload Analysis	540
4.6.2 Performance Group/Period Workload Tracking	546

4.6.3 System Workload and Performance Profile	549
4.6.4 Tracking Previous Forecasts of Resource Utilization	552
Chapter 5: PROFILE AND TRENDING ANALYSIS	557
5.1 Functional Description	557
5.2 Usage Guidelines	558
5.3 Standard Output	560
5.3.1 Historical Observations Report	560
5.3.2 Growth Rate Report	561
5.3.3 Estimates of Future Observations Report	563
5.4 Analytic Technique Tutorial	565
5.5 Component Operation	568
5.6 Case Studies	573
5.6.1 Extended Example I/O Workload Analysis	574
5.6.2 Performance Group/Period Workload Tracking	582
5.6.3 JES Spool	583
Chapter 6: SIMPLE REGRESSION ANALYSIS	587
6.1 Functional Description	588
6.2 Usage Guidelines	588
6.3 Standard Output	589
6.3.1 CSV Creation for CA MICS Q&R Reporting	592
6.4 Analytic Technique Tutorial	592
6.5 Component Operation	596
6.6 Case Studies	603
6.6.1 Extended Example I/O Workload Analysis	603
6.6.2 JES Spool	608
6.6.3 Application Tracking	610
Chapter 7: UNIVARIATE MODEL FORECASTING	613
7.1 Functional Description	613
7.2 Usage Guidelines	614
7.3 Standard Output	617
7.3.1 Model Analysis Report	618
7.3.2 Residual Analysis Report	620
7.3.3 R-Square Ranking Report	622
7.3.4 Updating the Capacity Planning Database	623
7.3.5 CSV Creation for CA MICS Q&R Reporting	624
7.4 Analytic Technique Tutorial	624
7.4.1 Linear, Quadratic, and Cubic Models	629

7.4.2 Processing Historical Observations	630
7.5 Component Operation	649
7.5.1 Univariate Model Forecasting Control Screen	650
7.5.2 Delete Selected Historical Data Observations Screen	658
7.5.3 Transform Historical Observations Screen	659
7.5.4 SAS Statements Screens	660
7.6 Case Studies	664
7.6.1 Extended Example Univariate Model Forecasting	665
7.6.2 Performance Group/Period Workload Tracking	673
7.6.3 System Workload and Performance Profile	676
7.6.4 Forecasting New Application Transaction Volumes	679

Chapter 8: WORKSHEET FACILITY 687

8.1 Functional Description	688
8.2 Usage Guidelines	688
8.3 Standard Output	688
8.3.1 Worksheet Report	689
8.3.2 Worksheet Effect Report	690
8.4 Forecasting Technique Tutorial	693
8.5 Product Operation	694
8.5.1 Worksheet Facility	694
8.5.2 Worksheet Effect Reports	719

Chapter 9: MULTIVARIATE REGRESSION FORECASTING 727

9.1 Functional Description	728
9.2 Usage Guidelines	729
9.2.1 Stepwise Multiple Regression Models	729
9.2.2 Choosing Model Parameters	731
9.2.3 Evaluating the Model	733
9.2.4 Forecasting Future Values from the Model	737
9.3 Standard Output	738
9.3.1 Update to the Forecast File	738
9.3.2 Standard Reports	739
9.4 Analytic Technique Tutorial	746
9.5 Component Operation	748
9.5.1 Multivariate Regression Forecasting Control Screen	749
9.5.2 Delete Historical Data Observations Screen	756
9.6 Case Studies	758
9.6.1 Capture Ratio Case Study	759
9.6.2 IMS Control Region Case Study	771

Chapter 10: BUSINESS ELEMENT FORECASTING

779

10.1 Functional Description	780
10.2 Usage Guidelines	781
10.3 Standard Output.....	786
10.3.1 Update to the Forecast File	786
10.3.2 Standard Reports	787
10.4 Analytic Technique Tutorial.....	794
10.4.1 Developing Business Element Models	795
10.4.2 Stepwise Multiple Regression Concepts	796
10.5 Component Operation	801
10.5.1 Business Element Forecasting Control Screen	802
10.5.2 Delete Selected Historical Observations Screen	808
10.6 Case Study	811
10.6.1 Problem Description	811
10.6.2 Creating a Business Element File.....	812
10.6.3 Developing a Business Element Model	817
10.6.4 Analyzing a Business Element Model	819
10.6.5 Automating the Collection of Business Elements	822

Chapter 11: WORKLOAD CHARACTERIZATION

823

11.1 Functional Description	824
11.2 Usage Guidelines	825
11.2.1 Historical Approaches	826
11.2.2 Basic Hypotheses	826
11.2.3 Application of Workload Characterization	828
11.2.4 Data Selection and Outliers.....	829
11.2.5 Suggested CA MICS Input Files	831
11.3 Standard Output.....	832
11.3.1 Descriptive Statistics Report	833
11.3.2 Cluster Centroids Report	835
11.3.3 Random Sample Cluster Assignment Report	836
11.3.4 Population Cluster Assignment Report	838
11.4 Analytic Technique Tutorial.....	840
11.4.1 Clustering Algorithms.....	840
11.4.2 Scaling	848
11.4.3 Algorithm Implementation.....	852
11.5 Component Operation	854
11.5.1 Workload Characterization Screen	855
11.5.2 Extended Options Menu	857
11.5.3 Global Data Selection Screen	858
11.5.4 Data Element Derivation Screen	860

11.5.5 Alternate Input File Screen	861
11.5.6 Save Selected Input Observations Screen	862
11.5.7 Save Cluster Output Observations Screen	863
11.5.8 Statistical Analysis Parameters Screen	864
11.5.9 Report Description Elements Screen	865
11.5.10 Inquiry Processing Screen	866
11.6 Case Studies	867
11.6.1 Batch Job Class	868
11.6.2 TSO Response	884
11.6.3 IMS Transaction	888

Chapter 12: INSTALLATION 891

12.1 Installation Procedure	892
12.2 Backups, Restores, and Compression	893
12.3 How to Enable Neugents technology in CA MICS	901
12.3.1 Requirements	901
12.3.2 Activate the CA MICS Root Directory	901
12.3.3 Enable Neugents technology	902

Chapter 13: RESOURCE COMPONENT ANALYSIS 903

13.1 Analysis using CA MICS files	903
13.1.1 Functional Description	904
13.1.2 Usage Guidelines	905
13.1.3 Standard Output	906
13.1.4 Analytic Technique Tutorial	907
13.1.5 Component Operation	907
13.1.6 Case Study	916
13.2 Analysis using CAP Database files	920
13.2.1 Functional Description	921
13.2.2 Usage Guidelines	922
13.2.3 Standard Output	923
13.2.4 Analytical Tutorial	924
13.2.5 Component Operation	924
13.2.6 Case Study	924

Chapter 14: WORKLOAD FORECASTING 931

14.1 Functional Description	931
14.2 Usage Guidelines	935
14.2.1 Neugents Technology Models	935
14.2.2 Choosing Model Parameters	936

14.2.3 Evaluating the Model	937
14.2.4 Forecasting Future Values with the Model	940
14.3 Standard Output.....	942
14.3.1 Update to the Forecast File	942
14.3.2 Standard Reports and Graphs	943
14.4 Analytic Technique Tutorial.....	958
14.5 Component Operation	958
14.5.1 Workload Forecasting Analysis Control Screen.....	958
14.5.2 Analysis File Data Elements Selection Screen	966
14.6 Case Studies	967
14.6.1 Processor Usage Study	967

Chapter 15: WORKLOAD CHARACTERIZATION 979

15.1 Functional Description	980
15.2 Usage Guidelines	981
15.2.1 Historical Approaches	982
15.2.2 Basic Hypotheses	982
15.2.3 Application of Workload Characterization.....	984
15.2.4 Data Selection and Outliers.....	985
15.2.5 Suggested CA MICS Input Files	987
15.3 Standard Output.....	988
15.3.1 Descriptive Statistics Report	990
15.3.2 Cluster Performance Summary Report	992
15.3.3 Cluster Population Summary Report.....	997
15.3.4 Data Value Exception Detail Report	1001
15.3.5 Sparse Cluster Population Report	1003
15.3.6 Cluster Index Analysis Graph.....	1005
15.3.7 Cluster Population Analysis Graph	1007
15.4 Analytic Technique Tutorial.....	1008
15.4.1 Clustering Algorithms.....	1008
15.4.2 Scaling	1016
15.4.3 Algorithm Implementation.....	1018
15.5 Component Operation	1019
15.5.1 Workload Characterization Screen	1020
15.5.2 Extended Options Menu	1022
15.5.3 Global Data Selection Screen	1023
15.5.4 Data Element Derivation Screen	1025
15.5.5 Alternate Input File Screen	1026
15.5.6 Save Selected Input Observations Screen	1027
15.5.7 Save Cluster Output Observations Screen	1028
15.5.8 Data Filtering Parameters Screen	1029

15.5.9 Clustering Execution Options Screen	1030
15.5.10 Report Description Elements Screen	1033
15.6 Case Studies	1034
15.6.1 Batch Job Class Study	1034

Appendix A: Messages **1061**

Messages.....	1061

Appendix B: QUICK START TUTORIAL **1113**

B.1 Quick Start Tutorial for CPU Planning	1114
B.1.1 Build the Capacity Database for Processor Planning	1115
B.1.2 Generate the CPU Forecast.....	1121
B.1.3 Create the CPU Worksheet	1125
B.1.4 Build Presentation Graphics.....	1128

Appendix C: INTERFACE WITH CA MICS Q&R WORKSTATION **1131**

Chapter 1: OVERVIEW

The CA MICS Capacity Planner addresses two key areas in capacity planning: workload characterization and workload forecasting. With the CA MICS Capacity Planner, you are able to perform key capacity planning studies and produce related reports using data that is maintained in the CA MICS database.

The CA MICS Capacity Planner Guide discusses analytic techniques and uses tutorials, reports, and case studies to explain workload characterization and workload forecasting.

Workload characterization helps you establish a workload baseline, based on current work activities, that will make future analyses more meaningful. Workload characterization uses data directly from the CA MICS database.

Workload forecasting helps you estimate growth in existing and new applications, and track and revise these estimates. Workload forecasting uses data from a special capacity planning database built using data elements that you select from the CA MICS database.

In addition to providing helpful forecasting techniques, the CA MICS Capacity Planner enables you to produce presentation graphics using the information you obtain. You can then develop a capacity planning strategy that is based on both past history and projections of future resource demand.

The CA MICS Capacity Planner enables you to maintain capacity history and projections information in both private and shared files.

The private files are accessible only to the planner who defines and maintains the data. This data are useful if you perform capacity planning functions alone, if you are developing ad hoc analyses, or if you are in the early phases of planning and are making frequent changes.

The shared files are useful if you share the capacity planning responsibilities among several planners and each may need to develop forecasts based on one set of shared files.

This section contains the following topics:

[1.1 Primary Areas of Application](#) (see page 17)

[1.2 Capacity Planning Database](#) (see page 20)

[1.3 Analysis Capabilities](#) (see page 22)

[1.5 Online Facility Tutorial](#) (see page 25)

1.1 Primary Areas of Application

The objective of any capacity planning study is to ensure that resources are available to provide the quality of service required by an organization. Capacity planning studies help to prevent resource shortages and save organizations from expensive slowdowns and down time.

The capacity planning process is divided into five functions: workload characterization, workload analysis and forecasting, performance prediction, financial analysis, and presenting the findings to management. A successful capacity planning process includes all of these functions:

- o Workload Characterization

Workload characterization is the ongoing analysis of the workload structure, resource consumption by each application, service levels, and system component utilization.

Your organization may have already characterized your workloads; if not, use the workload characterization functions of the CA MICS Capacity Planner as described in Chapter 11 of this guide.

- o Workload Analysis and Forecasting

Workload analysis and forecasting is predicting growth in existing applications and estimating the resource requirements of new applications, and tracking and revising these predictions, as appropriate.

The first step in analyzing and forecasting your resource consumption is to gather data about your system. You can use the facilities of the CA MICS Capacity Planner to build a data base of the elements you want to use as the basis of your studies. The process of defining files to build the database is described in Chapters 2 and 3 of this guide.

The Standard Applications files, described in Chapter 2, are predefined files that collect data contained in the CA MICS database and build a Capacity Planner database.

Resource Element, Meta, and Business Element files, described in Chapter 3, enable you to define the files to collect and/or augment the data contained in the Capacity Planner database.

Once you have defined your files and built your database, the next step is to analyze and forecast the data.

The analysis and forecasting tools of the CA MICS Capacity Planner fall into two categories: trend analysis and what-if analysis. The trend analysis tools, Profile and Trending and Simple Regression are described in Chapters 5 and 6. The what-if analysis tools, Univariate Modeling, Multivariate Regression and Business Element Forecasting are described in Chapters 7, 9, and 10.

o Financial Analysis

Financial analysis is reviewing different financial alternatives, once you've arrived at a set of configurations, to find those that provide the best return to the enterprise.

The CA MICS Capacity Planner can be used to analyze the data contained in CA MICS Accounting and Chargeback to assist in the financial analysis phase of capacity planning.

o Presentation Graphics

Presentation graphics enable you to present the set of configurations to management for further evaluation and, after the plan has been implemented, produce regular reports to compare the actual consumption levels against the plan.

With MICF access to CA MICS Capacity Planner files, you can produce high quality graphic reports as part of your capacity plan and present the results to management. This is described in Section 1.5.10 and in Chapter 4.

The first two functions, Workload Characterization and Workload Analysis and Forecasting, require processing vast quantities of data to identify and track current and new applications, monitor service and system component utilization levels, further refine the forecasting process, and build models of applications. (For further information see SAR79).

Although you may build models for predicting service and component utilization levels from small samples of data, you may have to analyze large quantities of data to ensure that the samples are representative of the population. The CA MICS database provides an ideal source for the quantities

of data required for the capacity planning process.

The CA MICS Capacity Planner provides a set of facilities that use the data in the CA MICS database and other business element data for capacity planning purposes. These facilities include: the private and shared capacity planning databases, analytic functions, and extensions to existing MICS functions.

Figure 1-1 shows the data process using the facilities of the CA MICS Capacity Planner.

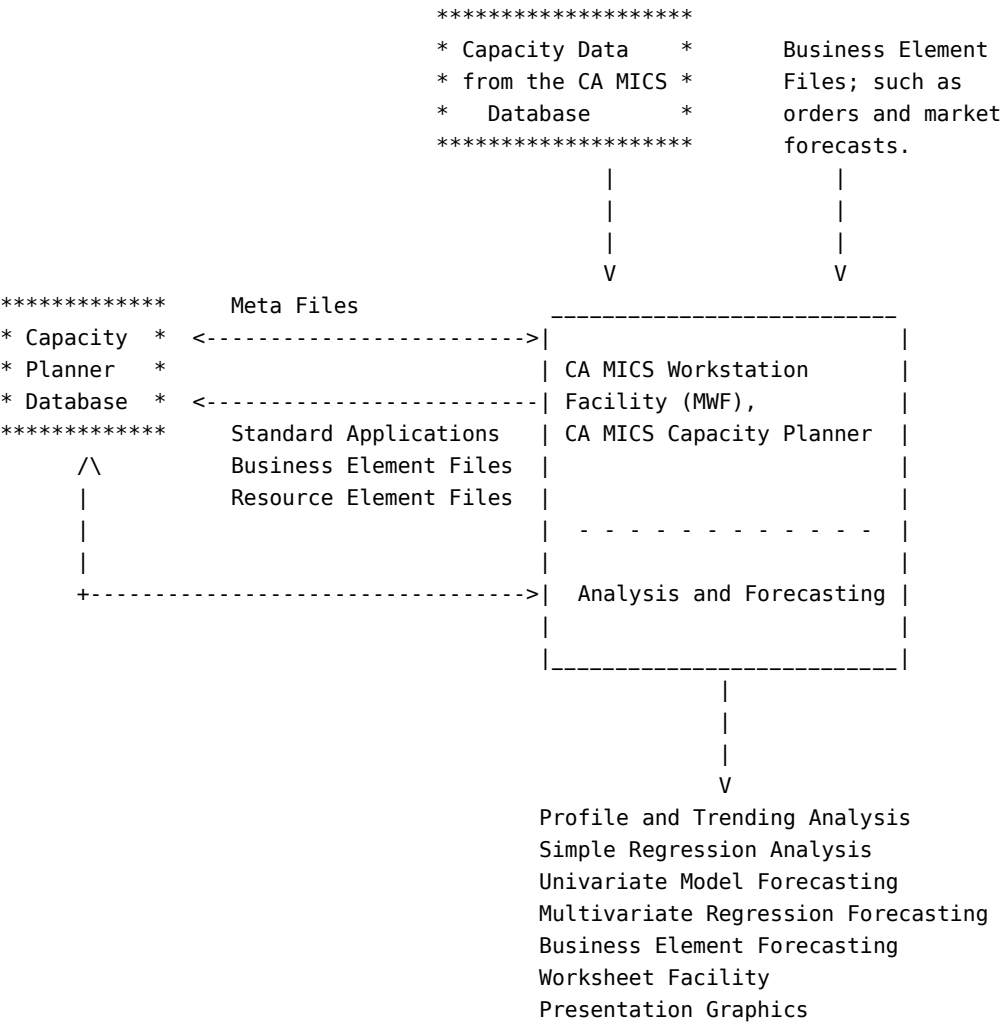


Figure 1-1. Data Process Through the CA MICS Capacity Planner

1.2 Capacity Planning Database

The CA MICS Capacity Planner provides facilities to maintain capacity planning information in databases that can be either shared (among all authorized users) or private (under your control). These SAS databases can be made up of four different types of files: Sample Application files, Resource Element files, Meta files, and Business Element files.

- o Sample Application Files are CA-defined views of data in the CA MICS database that can be used to do capacity planning in a number of different computing environments (for a complete discussion, see Chapter 2 of this guide). Sample Applications are designed to be a 'fast-path' approach to building a capacity planning database and do not require you to enter many parameters to start collecting CA MICS information in your capacity planning database.
- o Resource element files are user-defined files that extract information from the CA MICS database and store it in your capacity planning database. In contrast to the Sample Applications, resource element files let you choose the CA MICS files from which to extract data, decide the level to which the data is summarized, and create new data elements. The MICS facilities enable you to select files and data elements and provide a regular interface for you to use. Section 3.3 provides a full discussion on Resource Element files.
- o Meta files are user-defined files, similar to Resource Element files, however rather than using CA MICS files as input, Meta files use data that reside in either the Private (or Shared) Capacity Planner database. Meta files can summarize this data according to your specifications. Using Meta files, you can summarize the data from one Sample application in different ways to provide multiple views of the same information (for example, by department, by division, by company, and by the organization). See Section 3.4 for a full discussion of Meta files and their uses.

- o Business Element files are user-defined and contain information from data sources outside the CA MICS architecture. Business Element files are used to store business elements such as last year's total orders, or number of units sold, or future projected sales into the CA MICS Capacity Planner so you can perform Business Element Forecasting (BEF). Business Element Forecasting (also known as Natural Forecasting Units (NFU) or Natural Business Unit (NBU) Forecasting) is a workload forecasting technique that uses business metrics (for example, Units Sold, Claims Processed) to forecast future computing resource needs. Business Element files are discussed in Section 3.5 of this guide.

The CA MICS Capacity Planner uses the information contained in the Capacity Planner database as input to the workload forecasting facilities. You can use this to predict future consumption levels. The forecasting facilities of the CA MICS Capacity Planner can store your forecasts in the Capacity Planner database so you can use them in Presentation Graphics or as input to the presentation facilities of MICF.

Finally, with the MICF Production Reporting facility, you can automatically update the shared Capacity Planning database as the prerequisite CA MICS operational job completes.

1.3 Analysis Capabilities

The CA MICS Capacity Planner provides an extensive set of analysis functions to perform workload characterization, workload forecasting, and presentation graphics. The facilities use the information contained in the CA MICS database, and in the Capacity Planner database, as the input to the analytic functions provided. These functions include:

- o Profile and Trending (PFT): a facility that reads the data in the Capacity Planner database and generates compound growth rates for the data elements that you select. PFT interfaces with the Capacity Planner database definitions to provide an easy menu driven dialog that allows you to select the files and data elements that you would like to analyze. (See Chapter 5 for more information.)
- o Simple Regression (SMR): performs linear regression on a user-selected data element in the Capacity Planner database and displays results in either a color or printer graphic format. This facility can be very useful for quick analysis and understanding the growth of a particular data element. (See Chapter 6 for more information.)
- o Univariate Modeling (UVM): an easy-to-use facility that allows you to perform complex statistical analysis against information in your Capacity Planner database that results in forecasts that can be saved back into the Capacity Planner database. These forecasts can use any of the Capacity Planner file types (see Section 1.2). When a forecast is saved, a forecast file is built that can be used as input to either Presentation Graphics or to MICF's presentation facilities. (See Chapter 7 for more information.)
- o Multivariate Regression (MVR): a statistical analysis facility that generates savable forecasts. In contrast to Univariate Modeling, MVR can use the activity of data elements in different (or the same) Capacity Planner files to forecast the activity of another data element. These forecasts can be made available to Presentation Graphics or MICF. (See Chapter 9 for more information.)
- o Business Element Forecasting (BEF): uses the information provided by a user-defined Business Element File as the basis for forecasting the activity of a computer resource metric contained in the Capacity Planner database. BEF allows you to discover relationships that may exist between business activity and computer

resource consumption and use that as the basis of your workload forecasting studies. These forecasts can be saved for further analysis or as input to Presentation Graphics or MICF. (See Chapter 10 for more information.)

- o Workload Characterization: a method by which computer resource consumers can be grouped into like categories. This facility uses data contained in the CA MICS database to find relationships that are made apparent through the application of statistical clustering techniques. These clusters (or groupings) can then serve as the basis for workload forecasting studies so you can plan the capacity needs of your enterprise. (See Chapter 11 for more information.)
- o Presentation Graphics (PGR): a screen driven interface to a set of plotting routines that produce printer and color plots of information in the Capacity Planner database. These plots can include current and previous forecasts, confidence limits, and reference lines. These analyses can be tied into MICF Production Reporting so that they are generated as part of the update process for a capacity planning file. (See Chapter 4 for more information.)
- o Worksheet Facility: a method by which users can communicate their forecasted system resource needs to you. The Worksheets use files and data elements you select from the Capacity Planner database and format them into a report that you can distribute to your user base. The users then fill in their anticipated resource needs and return these to you. A data entry facility enables you to input this information and update the Capacity Planner database with the forecasts. The Worksheet Effect Report processes the information from the worksheets and combines it with actuals data and generated forecasts into a consolidated report that reflects anticipated and actual resource activity. (See Chapter 8 for more information.)

Note: SAS/STAT is a prerequisite for using SMR, UVM, MVR, BEF, and Workload Characterization. SAS/GRAPH is a prerequisite for using the graphics capabilities of PGR.

- o Resource Component Analysis using CA's U.S. patented Neugents Technology (NCA): a statistical analysis tool that permits the quantification of effects of a series of resource components against a specified analysis element. For example, processor busy is a function of the resources

used by a series of workloads. The Resource Component Analysis application can determine which of the workloads has the greatest impact on processor busy, and provide a weighted measure of the impact. This facility can be used to prepare for workload forecasting operations or as a general data analysis tool. (See Chapter 13 for more information.)

- o Workload Forecasting Analysis using Neugents Technology (NWF): a statistical analysis facility that generates savable forecasts. This facility works in a similar fashion to the Multivariate Forecasting facility mentioned above but uses a set of CA patented techniques to permit you to develop nonlinear models that relate different resource consumption data elements from one or more capacity planning database files. This forecasting capability provides for the investigation of possible relationships among different resource variables, and the forecasting of future requirements of a resource variable based on estimates of the future requirements of one or more different resource variables. (See Chapter 14 for more information.)
- o Workload Characterization using Neugents Technology (NWC): a method by which computer resource consumers can be grouped into like categories. This facility uses data contained in the CA MICS database to find relationships that are made apparent through the application of statistical clustering techniques. These clusters (or groupings) can then serve as the basis for workload forecasting studies or can be used as input for additional analyses. This facility is similar in functionality to the Workload Characterization mentioned above, but utilizes a set of CA patented techniques to perform the clustering operation. (See Chapter 15 for more information.)

Note: Neugents Technologies, provided at no additional cost, is a prerequisite for using NCA, NWF, or NWC.

All of these analytic facilities are implemented through a set of similar dialog interfaces, reducing the time to productive usage. They also have extensive tutorial support to assist you in their usage and the case studies found in this guide provide practical examples of using these facilities with actual data.

1.5 Online Facility Tutorial

The online facility provides the means for using the CA MICS Capacity Planner. This section introduces the online facility and explains how to use it. Additional online tutorials provide extensive guidance in using the features of the CA MICS Capacity Planner.

If you are familiar with the CA MICS Information Center Facility (MICF), you may notice many similarities between it and the CA MICS Capacity Planner's online facility. The online facility uses many of the MICF features. Any experience you have with MICF will benefit you as you begin using the CA MICS Capacity Planner's online facility.

The remainder of this guide contains detailed descriptions of each of the CA MICS Capacity Planner features and their operation using the online facility:

- 1 - Introduction to the Online Facility
- 2 - Getting Started with the Online Facility
- 3 - Capacity Planning Database File Definition Inquiries
- 4 - Capacity Planning Analysis Inquiries
- 5 - Replaying Cataloged Output
- 6 - Capacity Planning Facilities Administration

1.5.1 Introduction to the Online Facility

The following sections provide an overview of the purpose and major features of the online facility:

- 1 - Overview of the Online Facility
- 2 - Major Features of the Online Facility

1.5.1.1 Overview of the Online Facility

The online facility provides a convenient way to use the CA MICS Capacity Planner to rapidly obtain the results you need. With the online facility, you do not need an in-depth knowledge of CA MICS or SAS to effectively use the CA MICS Capacity Planner for building, maintaining, and analyzing capacity planning database files. The online facility's automatic validation procedures prevent you from making invalid requests, such as trying to use data elements not contained in the CA MICS files to build your capacity planning resource element files. Automatic validation also prevents syntax errors during subsequent analysis of capacity planning database files.

Following MICF's format, the online facility allows you to produce programs called file definition inquiries and analysis inquiries. You can design these inquiries either to collect capacity planning data, or to analyze that data and produce tabular and graphics reports.

The CA MICS Capacity Planner includes a number of standard capacity planning applications that you will find immediately useful through the online facility. You can easily generate file definition inquiries based on these standard applications and save them in your private MICF catalog, possibly further tailoring them to your particular requirements if necessary. Once a file definition inquiry has been created, it can be saved and reused simply by selecting it from the MICF catalog.

Online screens are provided for each of the data gathering, analytical, and reporting features of the CA MICS Capacity Planner. These screens and the chapters in this guide that discuss them will lead you through the use of each of these features.

A comprehensive, interactive tutorial facility assists you in learning and using the online facility. In addition, you have online access to all of the CA MICS documentation while you are using the online facility.

The following steps briefly describe how the online facility works:

- o The input, selection criteria, analysis, and report options are defined.
- o Transparent to you, the online facility dynamically builds

the file definition inquiry or analysis inquiry program from your specifications.

- o The inquiry is executed interactively or in batch, and the results either create or update a capacity planning file, or perform an analysis upon the data in your capacity planning database files. The reports and diagnostic output produced during the inquiry execution are displayed at your terminal or routed to a printer.

1.5.1.2 Major Features of the Online Facility

The online facility is a menu-based system that operates under IBM's Interactive Systems Productivity Facility (ISPF) and makes use of many of the features of the CA MICS Information Center Facility (MICF). If you are familiar with ISPF commands, you will find that the online facility menus behave like ISPF menus. The major features of the online facility include:

- o Inquiry composition
- o Inquiry modification
- o Execution-time specifications
- o Output replay
- o Facilities for advanced users

Each of these features is generally applicable to both file definition inquiries and analysis inquiries.

INQUIRY COMPOSITION

By completing one or a few online screens, you may create any number of special capacity planning inquiries. These inquiries allow you to do the following:

- o Create pre-designed capacity planning database files
- o Create custom capacity planning database files
- o Update and maintain these database files
- o Perform analyses and forecasting based upon the contents of these database files
- o Graphically present the results of forecasts

INQUIRY MODIFICATION

Once created, an inquiry may be stored in your private catalog, recalled, copied, deleted, or modified. You make changes to an inquiry using the same online facility screens that you originally used to create the inquiry. You do not, however, need to redisplay screens that did not change.

EXECUTION-TIME SPECIFICATIONS

You may define analysis inquiries so that key inquiry parameters can be specified at the time the analysis inquiry is executed. This facility allows you to report against varying data without changing the analysis inquiry. File definition inquiries do not support the execution-time specifications feature.

OUTPUT REPLAY

You can save the output produced when you execute a file definition or analysis inquiry, and replay it later. The output is saved by inquiry name, date, and time. Your CA MICS System Administrator can place inquiry outputs into a shared replay catalog that can be used throughout your installation.

When you replay an inquiry output, you can either view it on your terminal or send it to a hardcopy device. If you are replaying color graphics output, you do not need to use the type of device that was specified at the time the inquiry was originally executed. You can specify the color terminal, color printer, or color plotter type to use at the time you request the replay.

FACILITIES FOR ADVANCED USERS

If you are familiar with the CA MICS database and are accustomed to programming in SAS, a feature which you will find particularly helpful is the ability to include your own data selection and data manipulation SAS statements in your file definition inquiries. Through the exits provided, you can specify any valid SAS statement for inclusion in the SAS statements that are generated by the online facility to create and update capacity planning resource element files. This facility allows you to define sophisticated data selection logic and to generate new data elements to be included in your capacity planning resource element files.

1.5.2 Getting Started with the Online Facility

After you read this section, you will be able to start the online facility and know how to use its menus, commands, and help screens. In addition, you will be able to customize the online facility to better meet your needs.

The online facility contains extensive online tutorials (help screens) that you can request at any point in an online facility session. These screens provide detailed information about the operation you are specifying at the time you request help.

When you leave the tutorials, the online facility returns you to the place in your work session from which you requested help. Ask for help by typing HELP after the Command prompt (or the Option prompt on some screens) at the top of the screen. Your PF1 key is usually set to do this for you. When you request help in this manner, you are immediately shown a tutorial that is relevant to the screen on which you are working. To leave a tutorial, type END after the Command (or Option) prompt at the top of your screen. Your PF3 key is usually set to do this for you.

This remainder of this section discusses the following topics:

- 1 - Operating Conventions of the Online Facility
- 2 - Starting the Online Facility
- 3 - Customizing the Online Facility

For information on BACKUP and RESTORE, see Chapter 12.

1.5.2.1 Operating Conventions of the Online Facility

The CA MICS Capacity Planner online facility follows the same format and operating conventions as the CA MICS Information Center Facility (MICF).

Both the online facility and ISPF are very easy to use. Even if you have very little data processing experience, you will be able to learn the basics very quickly simply by using the terminal and signing on (though we recommend that you keep an experienced MICF or ISPF user handy for the first few minutes). If you would like more information on ISPF, you may want to consult IBM's manuals on ISPF called: Interactive System Productivity Facility (ISPF) Reference Summary, No. SC34-4445 and Interactive System Productivity Facility Getting Started, No. SC34-4440.

1.5.2.2 Starting the Online Facility

Start the online facility by selecting it from the CA MICS Workstation Facility (MWF) menu. The method of reaching the CA MICS Workstation Facility varies from installation to installation.

There are two common ways to start MWF. Your CA MICS System Administrator may have installed CA MICS so that you can reach MWF by selecting an option on your installation's ISPF primary option menu or another ISPF menu.

If you cannot start MWF from a standard ISPF menu, ask your CA MICS System Administrator for the name of the MWF CLIST.

If neither of these methods works, ask your CA MICS System Administrator to tell you how to start MWF.

More information on starting up and using MWF is contained in the PIOM.

Regardless of how you start MWF, once you select the options that bring you to the Capacity Planner Management Support Application, you will see the screen shown in Figure 1-2.

```
/----- Capacity Planner -----\
|Option ==>
|
|    0 - MICS Options
|        Review/modify MICS parameters and default options.
|    1 - Capacity Planning Data Base Files
|        Define, modify, or update capacity planning data base files.
|    2 - Capacity Planning Analysis
|        Define, modify, or update capacity planning analyses.
|    3 - Capacity Planner Worksheet Facility
|        Create, input, and produce capacity planning worksheets.
|    4 - Capacity Planning Facilities Administration
|        Maintain shared capacity planning data base files.
|    5 - Reports and Graphics
|        Review reports and graphics from inquiries and production jobs.
|    6 - Production Reporting Administration
|        Add, modify, execute, or delete MICS inquiry production jobs
|        or Capacity Planner production update jobs.
|    T - Tutorial
|    X - Exit
|
```




Figure 1-2. Capacity Planner Menu

For more detailed discussions on the Capacity Planner screen options, refer to the following sections of this guide or other CA MICS documentation:

- o For Option 0, MICS Options, refer to the CA MICS MICS User Guide and to sections 1.5.2.3 and 1.5.9 of this guide.
- o For Option 1, Capacity Planning Database Files, refer to Section 1.5.3 and Chapter 3 of this guide.
- o For Option 2, Capacity Planning Analysis, refer to Section 1.5.4 and Chapters 4 through 11 of this guide.
- o Option 3, Capacity Planner Worksheet Facility, refer to Section 1.5.8 of this guide.
- o Option 4, Capacity Planning Facilities Administration, refer to Section 1.5.8 and Chapter 8 of this guide.
- o For Option 5, Reports and Graphics, refer to Section 1.5.5 of this guide or to the CA MICS MICS Reference Guide.
- o For Option 6, Production Reporting Administration, refer to Section 1.5.7 of this guide or to the CA MICS MICS Reference Guide.

1.5.2.3 Customizing the Online Facility

This section provides a general description of how to customize the default parameters governing your online facility sessions. Parameters that are used by the CA MICS Capacity Planner online facility are shared with those used by MICF. Consequently, changes you make using Option 0 of the Capacity Planning menu are reflected when you use MICF, and vice-versa. You can find details on each parameter in the CA MICS MICF Reference Guide.

When you use the online facility, you must specify how to allocate data sets, what printer formats to use, what color graphics devices you are using, what ISPF profile you want to use with the online facility, and what defaults you would like to have when executing a capacity planning file definition or analysis inquiry. Defaults for each of the parameters in these categories were included with MICF when it was shipped to your installation. Your CA MICS System Administrator has further tailored these defaults for your site. In addition, the online facility often gives you the opportunity to revise parameters temporarily for a particular inquiry.

For these reasons, we recommend that you start using the online facility with the default options unchanged. As you become more familiar with the online facility, you may wish to reset some of these defaults. To do so, select Option 0, MICF Options, on the primary Capacity Planning menu.

1.5.3 Capacity Planning Database File Definition Inquiries

The purpose of this section is to introduce the online facilities that are available for creating, modifying, deleting, and executing file definition inquiries. Subsequent chapters of this guide provide direction on the specifications you must make to address various capacity planning requirements.

The following sections address these topics:

- 1 - Creating a File Definition Inquiry
- 2 - Modifying a File Definition Inquiry
- 3 - Repeating or Copying a File Definition Inquiry
- 4 - Deleting a File Definition Inquiry
- 5 - Executing a File Definition Inquiry

1.5.3.1 Creating a File Definition Inquiry

Creating a capacity planning database file is easy:

- o You tell the online facility that you are about to create a file definition inquiry.
- o You give the file definition inquiry a name and assign it to a catalog group. (A discussion of catalog groups follows.)
- o You specify the file definition process you plan to use.
- o You choose to save, execute, or cancel the file definition inquiry.

Note: If you decide to execute the inquiry, you must save it first.

As a user of the online facility, you will have your own private catalog of file definition inquiries. Before you create your first file definition inquiry, there is no PRIVATE capacity planning database for your user ID, but the first time you try to store data for a file definition inquiry, a PRIVATE capacity database is allocated automatically. This is illustrated below.

This database is a SAS Data Library with your user ID as the high-level prefix. The historical data and forecasts for the first and for all future inquiries will be stored in this PRIVATE capacity database. Section 12.2 of this guide provides backup and restore procedures that should be applied to your PRIVATE database.

To create a file definition inquiry, select Option 1, Capacity Planning Data Base Files, from the primary Capacity Planner menu (refer to Figure 1-2). The Capacity Planning Data Base Files screen shown in Figure 1-3 is displayed. You can use this screen to create, modify, and delete capacity planning database files.

```

/----- Capacity Planning Data Base Files -----\
|Command ==>                                     Scroll ==> PAGE |
|                                                                 |
|Catalog      ==> PRIVATE (PRIVATE/SHARED)  -----Line Commands-----|
|Catalog group ==> CAPACITY                  |A Add Notes B Browse S Select ||
|Select mode  ==> EXECUTE (EXECUTE/MODIFY)  |M Modify    E Execute I Insert ||

```

Name (subset) ==> ____		R Repeat	C Copy	D Delete

Cmd	Name	Title	Type	Status
-	-	-	-	-
_	DAS	DASD Sample Application	S	Updated 05/11/30
***** BOTTOM OF DATA *****				

Figure 1-3. Capacity Planning Data Base Files Screen - PRIVATE Catalog

When you enter the Capacity Planning Data Base Files screen, the default specification for Catalog is PRIVATE. The SHARED catalog facility permits you to copy inquiries that exist in the shared catalog into your private catalog.

Each catalog is divided into catalog groups. These groups contain the actual database file definition inquiries. When the Catalog group field contains no specification, the lower part of the screen displays the names of the various catalog groups that have been defined.

When you select a catalog group, either by typing an S (Select) in the Cmd column or by typing a catalog group name after the Catalog group prompt, the display portion of the screen shows the database file definition inquiries contained in that catalog group. To see the list of all file definition inquiries in your catalog, place an asterisk (*) after the Catalog group prompt. To see the list of all inquiries in all catalog groups starting with the same sequence of characters, code the sequence, followed by an asterisk, after the Catalog group prompt. For example,

Catalog Group $\implies S^*$

If you type `S*`, as shown in the example, you will see all

the definitions in all catalog groups starting with the letter S.

If the list of inquiries is long, you can scroll through it using the standard ISPF DOWN and UP commands. You can subset the list of inquiries by entering the beginning characters of the names you are interested in after the Name prompt. Only inquiries with names starting with these characters are shown.

The Select mode controls the action of the S (Select) line command and the SELECT primary command. The two possible values are EXECUTE or MODIFY, which default to the last value specified for this parameter:

- o Specify MODIFY to modify an existing file definition.
- o Specify EXECUTE to update a capacity planning database file (that is, execute the file definition inquiry).

The information about each file definition inquiry is displayed on a single line and includes the inquiry name, title, type of inquiry process (discussed below), and current status of the inquiry.

To create a new database file definition inquiry, type an I (Insert) on any line in the Cmd column. The Insert Capacity Planning Data Base File screen shown in Figure 1-4 is displayed.

```

/----- Insert Capacity Planning Data Base File -----\
|Option ==>|
|          |
|    File Name  ==>  ___|
|          Title  ==>  _____|
|          Author ==>  CA|
|          |
|Catalog Group Name ==> SAMPLES|
|          Title ==> Sample Application Files|
|          |
|          1 - Sample Applications|
|          Specify file definition parameters for one of the|
|          capacity planning sample applications.|
|          2 - Business Element File|
|          Specify file definition parameters for a file of user|
|          |

```

```
|      specified business element data.      |
|  3 - Resource Element File                |
|      Specify file definition parameters for a file of actuals |
|      from the CA MICS database.           |
|  4 - Meta File                          |
|      Specify file definition parameters for a file of actuals |
|      from the Capacity Planner database.  |
|-----/
```

Figure 1-4. Insert Capacity Planning Data Base File Screen

On the Insert Capacity Planning Data Base File screen, you can specify the inquiry name (one to three characters), title, and author. The catalog group name and title are copied from the information that applies to the line on which you entered the I. If you want to change either of them, you can do so here. Note that if you want to change any of the information on this screen later, you must first make a copy of the inquiry.

As shown in Figure 1-4, there are four different types of capacity planning processes that you can use to create a file definition inquiry: (Each of these is discussed in detail in Chapters 2 and 3 of this guide.)

- o Choosing Option 1 allows you to use one of the Capacity Planner sample applications for capacity planning.
- o Choosing Option 2 allows you to create capacity planning business element files that contain business activity related data elements defined by you.
- o Choosing Option 3 allows you to create a user-defined capacity planning database file using historical data from the CA MICS database. However, unlike the sample applications, you must specify ALL file definition parameters.
- o Choosing Option 4 allows you to create a user-defined meta capacity planning database file. Meta files are similar to Resource Element files; however, rather than using data from the CA MICS database, meta files use the shared or private Capacity Planner database.

Depending on the option you specify on this screen, pressing ENTER will give you a series of screens that allow you to complete the specifications for your file definition inquiry.

1.5.3.2 Modifying a File Definition Inquiry

Sometimes you may want to change an existing inquiry, thereby replacing the original. You may do this from the Capacity Planning Data Base Files screen (refer to Figure 1-4) either by typing an M line command next to the inquiry to be modified, or by specifying a Select mode of MODIFY, typing an S (Select) in the Cmd column on the line of the inquiry to be modified, and pressing ENTER. The screens you originally used to construct the existing inquiry are displayed. You may modify the parameters on these screens to conform to your new requirements. When you save the modified inquiry, it replaces the original inquiry in your catalog.

1.5.3.3 Repeating or Copying a File Definition Inquiry

This section describes how to use an existing file definition inquiry as a model for a new inquiry without changing the original inquiry. Figure 1-5 shows a typical display of the contents of a private catalog.

[illegible]

Figure 1-5. Repeating an Inquiry

To repeat a copy of a file definition inquiry in your private catalog, type an R (REPEAT) in the Cmd column on the line of the inquiry that you wish to copy and press ENTER. The Repeat File Definition screen shown in Figure 1-6 appears.

To copy an existing file definition from the shared catalog into your private catalog, you first need to view the shared catalog by changing PRIVATE to SHARED in the Catalog field on the screen shown above. Then type a C (COPY) in the Cmd column on the line of the inquiry that you wish to copy and

press ENTER. A Copy File Definition screen similar to the Repeat File Definition screen shown in Figure 1-6 displays.

```

/----- Repeat File Definition -----\
|Command ==>
|
|Catalog: PRIVATE
|
|      File name  ==> xxx
|      Title    ==> CPU Sample Application
|      Author   ==> CA
|      Type:    Sample Application
|
|Catalog group name ==> CAPACITY
|      Title ==> Capacity Planning
|
|Invoke SAS to copy history to your private catalog  ==> YES (YES/NO)
|Invoke SAS to copy forecasts to your private catalog ==> NO  (YES/NO)
|Convert Sample application to resource element file ==> NO  (YES/NO)
|
|Press ENTER to add the new file definition to your private catalog.
|or
|Press END to cancel.
|
|
\-----/

```

Figure 1-6. File Definition Inquiry Repeat Screen

Enter the name that this copied inquiry will use in the File name field. The values for the Title and Author fields are copied from the original inquiry. You may change them if you desire.

You are also given an opportunity to choose whether to copy any existing history or forecast SAS files at this time. The screen entries for these choices will default to YES only if such data exists. This copying will be done by invoking SAS in foreground.

If the file definition being repeated or copied is a Sample application, you can choose to convert the newly created copy to a resource element file. You would do this if you want to make subsequent modifications to the new definition by adding other input files or data elements, or by making alterations in some of its calculations.

When you are satisfied with the entries, press ENTER.

1.5.3.4 Deleting a File Definition Inquiry

Figure 1-7 shows a typical display of the contents of a private catalog.

[illegible]

Figure 1-7. Deleting an Inquiry

To delete one of these file definition inquiries in your private catalog, type a D for delete in the Cmd column on the line of the inquiry which you wish to delete and press ENTER. The Capacity Planning Database Files Confirm Delete screen shown in Figure 1-8 appears.

```

/----- Capacity Planning Data Base Files -----\
|Command ==>|
|
|*****|
|*                                           *|
|*           Confirm Delete                 *|
|*                                           *|

```

```
|      *      *
|      *   Press ENTER to DELETE or END to bypass   *
|      *      *
|      *****
|
|File name: CPU
|    Title: CPU Planning Application
|    Author: CA
|       Type: Sample Application
|    Status: Initial 98/11/01
|Update status: Batch initialization job submitted
|    Last update: 05/11/01:09:34
|    Catalog: PRIVATE
|Catalog group name: CAPACITY
|          Title: Capacity Planning
|
|Invoke SAS to delete the private catalog SAS files ==> YES (YES/NO)
|
|-----
```

Figure 1-8. Capacity Planning Database Files Confirm Delete Screen

Press ENTER to confirm the delete request or type END to cancel the delete request.

Specifying YES to delete SAS files on this confirmation screen will invoke SAS in the foreground to delete the files associated with the file definition. Specifying NO will delete the file definition without deleting the SAS files. If you delete the file definition but do not delete the associated SAS files, these files will no longer be accessible by the CA MICS Capacity Planner. If SAS files exist for this definition, this option will default to YES; otherwise, the option will default to NO. You would ordinarily allow this option to remain as it defaults, unless your installation does not permit even brief execution of SAS functions in foreground. If you change this option to NO, you should submit a batch job at some time in the future to delete these SAS files.

1.5.3.5 Executing a File Definition Inquiry

To execute a file definition inquiry, either enter an E line command next to the inquiry on the Capacity Planning Database Files screen (Figure 1-3), or specify a Select mode of EXECUTE, type an S (Select) in the Cmd column on the line of the inquiry you want to execute, and press ENTER. The Update Capacity Planning File screen (Figure 1-9) appears. You will be provided the opportunity to specify several execution options. The type of file definition inquiry you are executing will determine the screens that are displayed.

```

/----- Update Capacity Planning File -----\
|Command ==>|
|
|Executing CAP DB File def.:  XXC - CPU Standard Appl,03/10/93|
|
|Type of update           ==> INITIALIZE (INITIALIZE/UPDATE)|
|Replace existing file?   ==> N          (Y/N)|
|
|Input data source        ==> ARCHIVE    (ARCHIVE/ONLINE)|
|CA MICS file cycle(s)    ==> __ - __   (0 - 99)|
|Archive generation number ==> ___      (0 to -99)|
|
|Starting date of update   ==> _____ (ddmonyy)|
|Number of months to select ==> _____ (1 to 9999)|
|
|Mode of execution        ==> BATCH     (BATCH/BACKGROUND)|
|
|List updated file contents? ==> YES     (YES/NO)|
|
|
|
|Press END to invoke MICF inquiry execution to update the resource element file.|
|Enter CANCEL to terminate the resource element file update.|
\-----/

```

Figure 1-9. Update Capacity Planning File Screen

If you are executing a resource element file definition inquiry, the Update Capacity Planning File screen shown in Figure 1-9 displays. This screen and its control parameters are discussed in detail in Section 3.3.2.3. When you type END in the Command field and press ENTER, or when you press the END PF key (usually PF3), the CA MICS Database Inquiry Execution screen shown in Figure 1-11 displays.

```

/----- Business Element File Data Entry -----\
|Option ==>|
|
|Executing CAP DB File def.: BEF - Business Element File|
|
|1 - Initialize Business Element File|
|Prime business element file with null entries in preparation|
|for online data entry.|
|2 - Online Data Entry|
|Update the business element file using SAS/FSP. The file must|
|have been previously initialized.|
|3 - Batch Data Entry|
|Update the business element file using external card images.|
|The card image update can be executed in foreground or batch.|
|
|Starting date ==> _____ (ddmonyy) Number of months ==> ____ (1-9999)|
|Card image DSN ==> _____|
|
|-----\

```

Figure 1-10. Business Element File Data Entry Screen

If you are executing a business element file definition inquiry, the Business Element File Data Entry screen shown in Figure 1-10 displays. This screen and its control parameters are discussed in detail in Section 3.5.2.4. When you specify an option and press ENTER, the CA MICS Database Inquiry Execution screen shown in Figure 1-11 appears.

```

/----- MICS Execution -----\
|Command ==>|
|
|Executing CA MICS Inquiry: XXC - CPU Standard Application|
|
|Inquiry Execution Mode ==> BATCH (BATCH/BACKGROUND/DMS)|
|Override Execution Options ==> N (Y/N)|
|SAS/DMS Options: Resume ==> Y (Y/N) Output Replay ==> Y (Y/N)|
|
|Press ENTER to execute the inquiry, or press END to cancel execution.|
|
|-----\

```

```

|----- Batch Job Submission Parameters -----|
|Edit Generated Jobstream%====> Y (Y/N) Hold Inq. Output for Replay ====> N (Y/N)|
|
|SYSOUT Class   ====> A                               SYSOUT Form   %====> ____|
|Destination   %====> ____                           SYSOUT Writer%====> _____|
|SYSOUT Parms.%====> _____|
|
|Job Card Information:
|  ====> //jobname JOB (accounting information),
|  ====> // MSGLEVEL=(1,1),MSGCLASS=A,PRTY=5
|  ====> /*
|  ====> /*
|
|-----/

```

Figure 1-11. CA MICS Data Base Inquiry Execution Screen

The CA MICS Data Base Inquiry Execution screen shown in Figure 1-11 offers you the following execution options, which are discussed in the next three sections:

- 1 - Selecting the Execution Mode
- 2 - Overriding the Default Execution Options
- 3 - Overriding the Default Job Submission Parameters

1.5.3.5.1 Selecting the Execution Mode

Three possible execution modes exist: batch, foreground, and DMS (SAS Display Manager). In batch mode, the online facility creates and submits a batch job which executes the inquiry. In foreground mode, the online facility executes the inquiry interactively and lets you review the results immediately. DMS mode is like foreground except MICF executes your inquiry under the SAS Display Manager (SAS/DMS). With DMS mode execution, you get control after inquiry execution completes and can use SAS/DMS facilities to analyze intermediate work file contents, revise color graphics, etc.

To select the execution mode, type either BATCH (or B), FOREGROUND (or F), or DMS (or D) after the Inquiry Execution Mode prompt.

If you select foreground, the inquiry executes interactively and gives you the ability to browse your output online, print it, and save it. The disadvantages of foreground execution are that you cannot use your terminal for anything else while a file definition inquiry is executing and you cannot use archive tape input.

Figure 1-12 shows a comparison of the features available in the batch and foreground execution modes.

Feature	Batch	Foreground

Ability to use archive tape input	Yes	No
Ability to catalog output	Yes	Yes
Ability to replay color graphics	Yes	Yes
Ability to print reports	Yes	Yes
Ability to print color graphics	Yes	Yes
Terminal available for use during inquiry execution	Yes	No

Figure 1-12. Comparison of Batch and Foreground Execution

1.5.3.5.2 Overriding the Default Execution Options

You may not need to override the default execution options very often, but you should familiarize yourself with the available options by typing Y after the Override Execution Options prompt. The Execution Options screen appears and lets you specify the color graphics format set, the color graphics device, and the SAS execution parameters for the current inquiry.

Most of the parameters that you can set on this screen are self-explanatory; however, the Execution Options tutorial screens provide more detail if you should need it. You can reach these screens from the Execution Options screen by typing HELP at the Command prompt or (usually) by pressing your PF1 key.

There are two sets of execution parameters. One set is for batch job submissions. The other set is for inquiries run in foreground. These options are set originally by your CA MICS System Administrator, although you may personalize them by using Option 0, MICS Options, on the Capacity Planner menu shown in Figure 1-2. The online facility uses the set of options that corresponds to your choice of execution mode: batch or foreground.

1.5.3.5.3 Overriding the Default Job Submission Parameters

When you instruct the online facility to execute an inquiry as a batch job, it automatically uses your default job control specifications. These include your job statement and the SYSOUT specifications for the printed report and the SAS log.

You may override the default specifications for the current inquiry execution and modify batch job submission parameters and job card information from the CA MICS Database Inquiry Execution screen (Figure 1-11). If you want to edit the generated job stream and inquiry program before you submit the job, type Y at the Edit Generated Jobstream Before Submitting Batch Job prompt.

1.5.3.6 Adding and Browsing Notes

Once you have created a file in the CA MICS Capacity Planner, you can attach notes describing this file. The CA MICS Capacity Planner provides the capability of annotating your files with text comments that are automatically attached to the capacity file. You may want to add descriptive text, for example, "This file is from the New York data center, and contains information on System A and System B."

Once created, you can browse and print these notes. You can control whether or not notes are printed automatically when the capacity file update is executed.

1.5.3.6.1 Adding or Updating Notes

On the Capacity Planning Database Files screen, to add or update notes for a file, use the "A" line command next to the file name. A screen for entering text appears. You can input up to 80 characters across and as many lines as you need. The CA MICS Capacity Planner uses PDF/EDIT against a special member that is attached to the file, so you have all the capabilities of your PDF/EDIT environment. The note is actually attached when you "save" or "end" (PF3) from the screen.

1.5.3.6.2 Browsing Notes

On the Capacity Planning Database Files screen, to view notes, use the "B" line command next to the appropriate file name. A text screen containing the notes entered earlier appears. The CA MICS Capacity Planner uses PDF/BROWSE against the special annotation member that is attached to the file, so you can use all the capabilities of your PDF/BROWSE environment.

1.5.3.6.3 Printing Notes

To print the notes that are attached to a capacity file, specify "Yes" in the Enable: Notes report option of the file definition. For the standard applications, this option is located on the extended options screen for the each standard application. For Resource Element, Business Element, and Meta Files this option is on the main file definition screen for that type of file.

The notes appear in the SASLIST SYSOUT data set in the same format as you typed. To disable printing of the notes, specify "No" in the Enable: Notes report option.

1.5.4 Capacity Planning Analysis Inquiries

This section explains how to use the online facilities to create, modify, delete, and execute capacity planning analysis inquiries. Subsequent chapters of this guide provide direction on the parameters necessary to perform a particular type of analysis (for example, Simple Regression Analysis or Multivariate Regression Forecasting).

The remainder of this section addresses the following topics:

- 1 - Creating an Analysis Inquiry
- 2 - Modifying an Existing Analysis Inquiry
- 3 - Repeating or Copying an Existing Analysis Inquiry
- 4 - Deleting an Analysis Inquiry
- 5 - Executing an Analysis Inquiry

As with creating a capacity planning database file definition inquiry, the process of creating an analysis inquiry is straightforward. You tell the online facility that you are about to create an inquiry. Then you give the inquiry a name and assign it to a catalog group (a discussion on catalog groups follows). Next, you define the analysis technique you plan to use. At this point, you can choose to cancel, save, or execute the inquiry. If you decide to execute the inquiry before saving it, you will have an opportunity to cancel, save, or modify it after execution.

To create an analysis inquiry, select Option 2, Capacity Planning Analysis, from the primary Capacity Planning menu (refer to Figure 1-2). The Capacity Planning Analysis screen shown in Figure 1-13 displays. You can initiate all of the manipulations (that is, creating, modifying, and deleting) of analysis inquiries from this screen. You can create an inquiry by inserting an entirely new one, by repeating an existing one found in your private catalog, or by copying an existing one from the shared inquiry catalog. Line commands exist for each of these functions.

```

/------ Capacity Planning Analysis -----/
|Command ==>                                     Scroll ==> PAGE
|
|Catalog      ==> PRIVATE (PRIVATE/SHARED)   -----Line Commands -----
|Catalog group ==> CAPACITY                   |S Select  M Modify  E Execute |
|Select mode  ==> MODIFY (EXECUTE/MODIFY)    |I Insert  R Repeat  C Copy   |
|Name (subset) ==> _____                |D Delete                                |
|-----
|
|Cmd          Name                               Title                        Date       Time     Type
|- - - - -
|_ BEFM1      Business Element Forecast - BEF File    98/05/03   18:16  CAPBEF
|_ CLUC1      Workload Characterization               98/05/03   17:27  CAPCLU
|***** BOTTOM OF DATA *****
|
|
|
|
|
|
|
|
|
|
|
|

```


Figure 1-14. Insert Capacity Planning Analysis Screen

On the Insert Capacity Planning Analysis screen, you can specify the inquiry name (one to six characters), title, and author. The catalog group name and title are copied from the information that applies to the line on which you entered the I. If you want to change either of them, you can do so here. Note that if you want to change any of the information on this screen later, you must make a copy of the definition to do so.

As shown in Figure 1-14, there are two different types of capacity planning processes that you can use to create an analysis inquiry.

Choosing Option 1 allows you to perform workload characterization inquiries using the SAS FASTCLUS procedure. This option, and the subsequent screens and menus, is further discussed in Chapter 11 of this guide.

Choosing Option 2 displays the Workload Forecasting menu shown in Figure 1-15. This menu allows you to specify which of the workload forecasting techniques you plan to use. When you select an option from this menu and press ENTER, the control screens associated with each of the workload forecasting techniques available in the CA MICS Capacity Planner are displayed. These techniques and the associated control screens are discussed in Chapters 4 through 10 of this guide.

```
/----- Workload Forecasting -----\  
|Option ===>|  
|Composing CA MICS Inquiry: BEFMI Business Element Forecast - BEF File|  
|  
|1 - Profile and Trending Analysis|  
|   Report history, growth rates, and trends for multiple measures.|  
|2 - Simple Regression Analysis|  
|   Graph historical values with a linear regression forecast.|  
|3 - Multivariate Regression Forecasting|  
|   Forecast based on projected future values of other measures.|  
|4 - Univariate Model Forecasting|  
|   Forecast using time-related regression with smoothing.|
```

```
| 5 - Business Element Forecasting
|     Forecast based on the business plan.
| 6 - Presentation Graphics
|     Graph history and forecasts for presentation to management.
| 7 - Worksheet Effect Reports
|     Produce reports showing effect of user forecast from worksheets.
|
|
|-----
```

Figure 1-15. Workload Forecasting Menu

1.5.4.2 Modifying an Existing Analysis Inquiry

You may sometimes want to change an existing analysis inquiry, thereby replacing the original. You may do this from the Capacity Planning Analysis screen (see Figure 1-13) by typing M next to the inquiry name, or specifying MODIFY for Select mode, typing an S (Select) in the Cmd column corresponding to the inquiry you wish to modify, and pressing ENTER. The screens that were originally used to construct the existing inquiry are displayed. You can modify these screens as the need arises to change the analysis inquiry and re-execute it. When you save the modified analysis inquiry, it replaces the original inquiry in your catalog.

1.5.4.3 Repeating or Copying an Existing Analysis Inquiry

This section describes how to use an existing analysis inquiry as a model for a new inquiry without changing the original inquiry. Figure 1-16 shows a typical display of the contents of a private catalog.

```
/----- Capacity Planning Analysis -----\
|Command ==>                                     Scroll ==> PAGE |
|                                                                 |
|Catalog      ==> PRIVATE (PRIVATE/SHARED)   -----Line Commands -----|
|Catalog group ==> CAPACITY                   |S Select M Modify E Execute |
|Select mode   ==> MODIFY (EXECUTE/MODIFY)    |I Insert R Repeat C Copy   |
|Name (subset) ==> _____                |D Delete                    |
|                                                                 |
|-----\
|Cmd     Name           Title                                Date      Time    Type |
| - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - |
|R BEFM1 Business Element Forecast - BEF File          98/05/03  18:16 CAPBEF |
|- CLUC1 Workload Characterization                     98/05/03  17:27 CAPCLU |
|***** BOTTOM OF DATA *****|
|
|
|
|
|
```

Figure 1-16. Repeating an Analysis Inquiry

To repeat a copy of one of these definitions in your private catalog, type R (Repeat) in the Cmd column on the line of the inquiry you wish to repeat and press ENTER. The Inquiry Repeat/Copy screen shown in Figure 1-17 displays.

A similar procedure can be used to copy any existing capacity planning analysis inquiry from the shared catalog. Simply change the Catalog field from PRIVATE to SHARED on this same screen and type C (Copy) in the Cmd column on the line of the inquiry you want to copy, then press ENTER.


```

/----- Inquiry Repeat/Copy -----Repeat Inquiry\
|Command ==>
|
|Catalog: PRIVATE
|
|    Inquiry Name  ==> ____
|           Title  ==> Business Element Forecast - BEF File
|           Author ==> Charles Frost
|
|Catalog Group Name ==> CAPACITY
|           Title   ==> Capacity Planning
|
|
|Press ENTER to add the new file definition to your private catalog.
|or
|Press END to cancel.
|
|
|
|
|
\-----/

```

Figure 1-17. File Definition Repeat/Copy Screen

Enter the Inquiry Name by which this repeated or copied capacity planning analysis inquiry will be known after the Inquiry Name prompt. The values for Title, Author, Catalog Group Name, and Title are copied from the inquiry being repeated; however, you may change them if you desire. When you are satisfied with the entries, press ENTER. The Capacity Planning Analysis screen is redisplayed to allow you to select the repeated inquiry for modification, as discussed in Section 1.5.4.2.

1.5.4.4 Deleting an Analysis Inquiry

Figure 1-18 shows a typical display of the contents of a private catalog.

```

----- Capacity Planning Analysis -----
|Command ==>                                         Scroll ==> PAGE
|
|Catalog      ==> PRIVATE (PRIVATE/SHARED)   -----Line Commands -----
|Catalog group ==> CAPACITY                  |S Select  M Modify  E Execute |
|Select mode  ==> MODIFY (EXECUTE/MODIFY)    |I Insert  R Repeat  C Copy    |
|Name (subset) ==> _____                |D Delete                                |
|                                                -----
|
|Cmd          Name                               Title                                     Date       Time     Type
| - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - -
| D    BEFM1    Business Element Forecast - BEF File           98/05/03   18:16  CAPBEF
| _    CLUC1    Workload Characterization                      98/05/03   17:27  CAPCLU
|***** BOTTOM OF DATA *****
|
|
|
|
|
|
|

```

Figure 1-18. Using the Capacity Planning Analysis Files Screen to DELETE an Inquiry

To delete one of these inquiries from your private catalog, type D (Delete) in the Cmd column of the analysis inquiry to be deleted and press ENTER. The Database Inquiries Confirm Delete screen shown in Figure 1-19 displays. (See Section 1.5.6.3 for information on how to delete Shared inquiries.)

[illegible]

Press ENTER to confirm the delete request or type END to cancel the delete request.

1.5.4.5 Executing an Analysis Inquiry

To select an analysis inquiry for execution from the Capacity Planning Analysis screen shown in Figure 1-13, either type an E line command next to the inquiry, or type the EXECUTE primary command and an S line command, and press ENTER.

When you select an analysis inquiry for execution, the CA MICS Database Inquiry Execution screen shown in Figure 1-11 displays.

The CA MICS Database Inquiry Execution screen offers you the following execution options:

- 1 - Selecting the Execution Mode
- 2 - Overriding the Default Execution Options
- 3 - Overriding the Default Job Submission Parameters

After you specify the options you desire and press ENTER, you are provided with a fourth option:

- 4 - Selecting Execution-Time Data

Each of these options is discussed separately by section.

1.5.4.5.1 Selecting the Execution Mode

Three possible execution modes exist: batch, foreground, and DMS (SAS Display Manager). In batch mode, the online facility creates and submits a batch job which executes the inquiry. In foreground mode, the online facility executes the inquiry interactively and lets you review the results immediately. DMS mode is like foreground except MICF executes your inquiry under the SAS Display Manager (SAS/DMS). With DMS mode execution, you get control after inquiry execution completes and can use SAS/DMS facilities to analyze intermediate work file contents, revise color graphics, etc.

To select the execution mode, type either BATCH (or B), FOREGROUND (or F), or DMS (or D) after the Inquiry Execution Mode prompt.

If you select foreground, the inquiry executes interactively and gives you the ability to browse your output online, print it, and save it. The disadvantage of foreground execution is that you cannot use your terminal for anything else while a file definition inquiry is executing.

Figure 1-20 shows a comparison of the features available in the batch and foreground execution modes.

Feature	Batch	Foreground
Ability to catalog output	Yes	Yes
Ability to replay color graphics	Yes	Yes
Ability to print reports	Yes	Yes
Ability to print color graphics	Yes	Yes
Terminal available for use during inquiry execution	Yes	No

Figure 1-20. Comparison of Batch and Foreground Execution

1.5.4.5.2 Overriding the Default Execution Options

You may not need to override the default execution options very often, but you should familiarize yourself with the available options by entering Y after the Override Execution Options prompt. The Execution Options screen appears and lets you specify the color graphics format set, the color graphics device, and the SAS execution parameters for the current inquiry.

Most of the parameters that you can set on this screen are self-explanatory; however, the Execution Options tutorial screens provide more detail if you should need it. You can reach these screens from the Execution Options screen by typing HELP at the Command prompt or (usually) by pressing your PF1 key.

There are two sets of execution parameters. One set is for batch job submissions. The other set is for inquiries run in foreground. These options are set originally by your CA MICS System Administrator, although you may personalize them by using Option 0, MICS Options, on the Capacity Planning menu shown in Figure 1-2. The online facility uses the set of options that corresponds to your choice of execution mode: batch, foreground, or DMS.

1.5.4.5.3 Overriding the Default Job Submission Parameters

When you instruct the online facility to execute an inquiry as a batch job, it automatically uses your default job control specifications. These include your job statement and the SYSOUT specifications for the printed report and the SAS log.

You may override the default specifications for the current inquiry execution and modify batch job submission parameters and job card information from the CA MICS Database Inquiry Execution screen (see Figure 1-11). If you want to edit the generated job stream and inquiry program before you submit the job, type Y at the Edit Generated Jobstream Before Submitting Batch Job prompt.

```
/----- Execution-Time Parameters -----/
|Command ===>                                Scroll ===> CSR
|Press END to execute the inquiry. Enter CANCEL to cancel inquiry execution.
|Executing CA MICS Inquiry: CLUC1 - Workload Characterization
|
|-----
|Enter the Data Base ID. (1 character) for the Data Base to be processed.
|A                                     - _____
|
|Enter the range of CA MICS File Cycles (e.g., 01 - 04) to be processed.
|01                                 - 02
|
|
|
|
|
```

Chapter 1: OVERVIEW 63

1.5.5 Replaying Cataloged Output

The online facility uses the MICF facilities for replaying, printing, and deleting cataloged output. These are described fully in the CA MICS MICF User Guide and the CA MICS MICF Reference Guide.

1.5.6 Capacity Planning Facilities Administration

The Capacity Planning Facilities Administration process is used to maintain the shared capacity planning file catalog and its corresponding history files and forecast files. A private file catalog is maintained for each user, and a single shared file catalog is maintained for access by all users. This Capacity Planning Facilities Administration process is used to add capacity planning files to the shared catalog, thereby making them available to all users. Files may be copied to the shared catalog from any user's private catalog. Capacity Planning Facilities Administration can also be used to delete shared capacity planning files which are no longer required.

The Capacity Planning Facilities Administration process is also used to update shared capacity planning files. Shared files are updated through the same process as private files, except that access to this process must be done through Capacity Planning Facilities Administration.

In order to ensure control of any changes to this shared facility, only those who have been authorized are permitted to make changes to the shared catalog and its associated history files and forecast files. Changes requiring this authorization include additions to the shared catalog, deletions from it, and updates to files in the shared catalog. All such changes must be made using Capacity Planning Facilities Administration. The authorization procedure is described in Section 4.4.3 of the PIOM.

All users of the CA MICS Capacity Planner can make use of shared capacity files in their own analyses, since such use will not change the shared files. Any forecast saved as a result of such analyses will be saved in the user's private catalog and marked as a "Forecast only" file.

The following topics are addressed in this section:

- 1 - Using the Facilities Administration Screen
- 2 - Inserting a File Definition
- 3 - Deleting a File Definition
- 4 - Updating a Shared Capacity Planning File

1.5.6.1 Using the Facilities Administration Screen

To reach the Capacity Planning Facilities Administration screen, select option 4, Capacity Planning Facilities Administration, from the Capacity Planning menu (see Figure 1-2). The Capacity Planning Facilities Administration screen (shown in Figure 1-22) is used to copy file definitions and optionally the corresponding history files and forecast files from any user's private catalog (including the System Administrator's private file catalog) to the shared file catalog. The screen is also used to delete shared file definitions that are no longer required.

Note that file definitions are only composed in private file catalogs and then copied to the shared file catalog. When a file definition is copied from a private catalog to the shared catalog, the version in the private catalog remains in the private catalog and is not altered in any way.

```
/----- Capacity Planning Facilities Administration -----\
|Command ==>                                         Scroll ==> PAGE |
|
|Userid          ==> TS0USR1           (For insert)   -----Line Commands-----
|Catalog group ==> CAPACITY             |A Add Notes B Browse |
|Prefix         ==> _____ (Optional) |S Select    E Execute |
|Name (subset)  ==> __                 |I Insert     D Delete |
|                                                    -----
|
|Cmd      Name               Title                Type       Status
|- - - - -
|_ CPU        CPU Analysis Standard Application      S Updated 98/11/05
|S DSD        DASD Analysis Standard Application     S Updated 98/11/05
|_ SNA        SNA Network Analysis Standard Appl.   S Modified 98/11/05
|_ WKL        System Workload by Performance Group   F Forecast only
|***** BOTTOM OF DATA *****
|
|
|
|
```

Figure 1-22. Capacity Planning Facilities Administration Screen

The contents of the Capacity Planning Facilities Administration screen vary depending upon the specifications

entered in the data entry fields at the top of the screen. These control the subset of the shared capacity planning file catalog to be displayed. A catalog structure is provided to simplify retrieval. File definitions are stored in groups of related files. This screen can display the list of defined catalog groups or the list of file definitions cataloged under a specific catalog group.

There are five data entry fields on this screen:

- o Userid
- o Prefix
- o Catalog group
- o Name
- o Cmd

The data entry considerations for each of these are discussed separately below.

Userid: Required for insert processing. Defaults to the last value specified for this parameter. Enter a 1-7 character TS0 user identification (or logon-id). This is used to allocate the specified user's private capacity planning file catalog.

Prefix: Optional for insert processing. Defaults to the optional high level qualifier for private data sets defined by the CA MICS Administrator. If your installation uses high level qualifiers for private data sets other than TS0 userid, you should enter the qualifier for the specific user's catalog from which you intend to insert file definitions.

Catalog group: Optional. Defaults to the last value specified for this parameter. Enter a 1-8 character catalog group name or *. This is the catalog group to be displayed. Leave this parameter blank to display the list of defined catalog groups. Use the S (Select) line command or enter the group name in this field to select a catalog group and display a list of file definitions cataloged under that group. Specify * to list all file definitions.

Name: Optional. Defaults to the last value specified for this parameter. It can be one or two

characters and is used to subset the list of file definitions displayed on the screen. The Capacity Planning Facilities Administration screen will be limited to file definitions with names having the same leading characters as specified in this parameter. For example, specifying CP will limit the display to file definitions with names beginning with CP.

Cmd: Optional. This is the line command data entry field.

A to add notes to this file definition
B to browse the notes
S to select a file definition for update
E to execute a file definition update
I for insert a new definition
D to delete a file definition

When Catalog group is blank, each line of the display represents a catalog group and lists the name and title for the catalog group. When Catalog group is not blank, each line of the display represents a file definition and lists the name, title, type, and status of the file definition.

1.5.6.2 Inserting a File Definition

This section describes how to insert a file definition from a user's private file catalog into the shared file catalog. Figure 1-23 shows a typical display of the Capacity Planning Facilities Administration screen. When you insert a file definition into the shared file catalog, you will be given the opportunity to optionally copy the SAS files associated with this file definition from the user's private SAS capacity planning file to the shared capacity planning file.

The insert command also gives you the capability of inserting forecast results into an existing shared capacity planning file.

[illegible]

Figure 1-23. Inserting a File Definition

To insert a file definition into the shared file catalog, specify the Userid for the private capacity planning file from which the file definition will be copied.

Specify I in the Cmd column (can be specified on any line).
The Private File Definition Selection screen (Figure 1-24)

will be displayed for the specified Userid. Select the desired file definition and press ENTER.

This screen is similar to the Capacity Planning Facilities Administration screen, except that it displays the private file definitions for the specified Userid. When you select an entry on this screen, you are given the opportunity to change the file definition name, title, author, catalog group, and catalog group title, unless you select a "Forecast only" file.

The Private File Definition Selection screen includes an option to allow or disallow merging of such private forecast results into an existing shared file. This merging can only be done from a private "Forecast only" file which has the same name as an existing shared file. For files other than "Forecast only" files, you are also given the option of copying any corresponding private history or forecast data to the shared SAS file. An additional screen, not presented here, provides you with the capability of specifying these options.

To prevent the inadvertent destruction of critical planning data, you must make sure that any new file definition you insert in the shared file catalog does not have the same name as any existing shared file.

```
/----- Private File Definition Selection -----\
|Command ==>                                     Scroll ==> PAGE |
|                                                    |
|Userid: TS0USR1                                Catalog group ==> CAPACITY |
|Prefix:                                         Name (subset display) ==> __ |
|Merge private forecast file ==> N (Y/N)        |
|                                                    |
|Line Cmds: S Select                            |
|                                                    |
|Cmd      Name                                Title                Type      Status      |
| - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - |
| _ CPU      CPU Analysis Standard Application      S   Updated  98/10/15 |
| S DSD      DASD Analysis Standard Application      S   Updated  98/10/15 |
| _ SNA      SNA Network Analysis Standard Appl.    S   Modified 98/10/13 |
| _ WKL      System Workload by Performance Group    F   Forecast only |
|***** BOTTOM OF DATA *****|
|                                                    |
|                                                    |
|                                                    |
|                                                    |
|                                                    |
```



Figure 1-24. Private File Definition Selection Screen

1.5.6.3 Deleting a File Definition

This section describes how to delete a file definition from the shared file catalog. Figure 1-25 shows a typical display of the Capacity Planning Facilities Administration screen.

[illegible]

Figure 1-25. Deleting a File Definition

To delete a file definition, select the appropriate Catalog group and scroll the display until the desired file definition appears on the screen.

Specify D in the Cmd column for the desired file. The confirm delete screen will be displayed. When you press ENTER to confirm (Figure 1-26), the file is deleted. It CANNOT be recovered.

```

/----- Capacity Planning Facilities Administration -----\
|Command ==>|
|

```



```

|          *****          |
|          *                  *          |
|          *      Confirm Delete      *          |
|          *                  *          |
|          *      Press ENTER to DELETE or END to bypass      *          |
|          *                  *          |
|          *****          |
|      File name: DSD          |
|      Title: DASD Analysis Planning Application          |
|      Author: CA          |
|      Type: Sample Application          |
|      Status: Updated 98/11/05          |
|      Update status: Batch initialization job submitted          |
|      Last update: 05/11/30:09:34          |
|      Catalog: SHARED          |
|      Catalog group name: CAPACITY          |
|      Title: Capacity Planning          |
|      Invoke SAS to delete the shared catalog SAS files ==> YES (YES/NO)          |
|-----/

```

Figure 1-26. Confirming Deletion of a File Definition

Specifying YES to delete SAS files on this confirmation screen will invoke SAS in the foreground to delete the files associated with the file definition. Specifying NO will delete the file definition without deleting the SAS files. If you delete the file definition but do not delete the associated SAS files, these files will no longer be accessible by the CA MICS Capacity Planner. If SAS files exist for this definition, this option will default to YES; otherwise, it will default to NO.

You would ordinarily allow this option to remain as it defaults, unless your installation does not permit even brief execution of SAS functions in foreground. If you change this option to NO, you should submit a batch job at some time in the future to delete these SAS files.

1.5.6.4 Updating a Shared Capacity Planning File

This section describes how to Update (execute) a file in the shared file catalog. Figure 1-27 shows a typical display of the Capacity Planning Facilities Administration screen.

[illegible]

Figure 1-27. Updating a Shared Capacity Planning File

To update a shared capacity planning file, select the appropriate catalog group and scroll the display until the desired file definition appears on the screen.

When you specify S in the Cmd column for the desired file, the selected file update will be executed. At this time the Update Capacity Planning screen, which is described in Section 3.4.2.3, appears, giving you the opportunity to override or alter execution parameters.

1.5.7 Production Reporting Administration

Production Reporting Administration allows you to define and maintain production updating jobstreams of your shared Capacity Planning database files.

You might choose to define one jobstream to control updates for several month-based Capacity Planning files, and another to control updates for several week-based Capacity Planning files. These jobstreams could be scheduled to run after the relevant CA MICS updates have completed, thus consolidating your capacity file updates into convenient packets or jobstreams.

The following sections cover the use and input specifications of Production Reporting Administration for defining and maintaining jobstreams for updating your shared Capacity Planning database files:

- 1 - Using Production Reporting
- 2 - Inserting Capacity Planning File Definitions

1.5.7.1 Using Production Reporting

To reach Production Reporting Administration, select option 5, Production Reporting Administration, from the Capacity Planner main menu. Selecting this option will lead you through a series of MICF screens to build and maintain a reporting jobstream. This sequence of screens is more fully described in the CA MICS MICF User Guide.

You can access this same sequence of screens not only from the Capacity Planner main menu but also from the MICF Administration menu.

Production Reporting Administration provides you the ability to define and maintain production updating jobstreams of your shared Capacity Planning database files. It also permits you to define and maintain MICF inquiry production reporting. Consult the CA MICS MICF User Guide description of this capability and for a description of the overall function of Production Reporting Administration.

Before implementing the production updating of any Capacity Planner files, you need to categorize the files. Create a worksheet of files, timespan dependencies, and unit dependencies, as shown in the following example.

Capacity File	Timespan Dependencies			Unit Dependencies		
-----	-----			-----		
File Name	Daily	Weekly	Monthly	Primary P	CICS C	VCA V
CPD	X			X		
CPW		X		X		
CPM			X	X		
WKM			X	X		
SMM			X			X
TRM			X	X	X	

In this example, creation of five production updating job streams will manage updating the six capacity files.

The updates for CPD, CPW, and SMM can be defined in appropriate DAILY, WEEKLY, and MONTHLY production job streams and automatically submitted. When the primary (P) CA MICS unit completes its MONTHLY processing, capacity files CPM and WKM can be updated in a single production job stream. The update for the TRM capacity file, however, requires that MONTHLY processing be complete for both primary (P) and CICS

(C) units and cannot be scheduled automatically.

The job stream categories are summarized as follows:

Category	Files	Criteria for the Category	Automatically Submitted
Daily	CPD	Daily, Unit: P	Yes
Weekly	CPW	Weekly, Unit: P	Yes
Monthly	CPM, WKM	Monthly, Unit: P	Yes
Monthly	SMM	Monthly, Unit: V	Yes
Monthly	TRM	Monthly, Unit: P and C	No

Jobs can be scheduled to be automatically submitted during the DAILY, WEEKLY, or MONTHLY processing for the unit.

1.5.7.2 Inserting Capacity Planning File Definitions

The Reports and Graphics Jobstream screen in the sequence described in the CA MICS MICF Reference Guide has an option that states either "Specify MICF Inquiries for Reporting Jobstream" or "Specify Capacity Planning Files to Update". For jobstreams being newly created, this option always displays the first of these two messages. For jobstreams that have been previously defined, this option displays whichever of these two statements is appropriate for the contents of the jobstream. If, when you are initially building the jobstream, you answer Y to this option, you will next see the Reports and Graphics Jobstream screen shown in Figure 1-28.

```
/----- Reports and Graphics Jobstream -----  
|Command ===>  
|Select the type of reporting jobstream being defined or press END to cancel.  
  
|  
|      1 - Select Shared MICF Inquiries  
|          Select the shared MICF data base inquiries to be included in the  
|          reporting jobstream.  
|      2 - Select Shared Capacity Planning File Definitions  
|          Select the shared capacity planning files to be updated by the  
|          production jobstream.
```

Figure 1-28. Reports and Graphics Jobstream Screen

The second option on this screen, the one for selecting shared capacity planning file definitions, permits you to insert one or more shared capacity planning file processes into a production reporting jobstream. Once you choose one of the options on this screen, the entire jobstream you are defining will consist of either shared MICF inquiries or

shared capacity planning file definitions.

The following description shows you the sequence of actions you need to take to add a shared capacity file definition to a production reporting jobstream. Use a series of three screens (as shown in Figures 1-29, 1-30, and 1-31) to manage the content of these jobstreams:

```
/----- Capacity Planning File Definitions -----|  
|Command ==>                                         Scroll ==> CSR|  
|Reporting Jobstream: CAP001   - CAP UPDATE 001    |  
|                                                    |  
|                                                    Catalog Group ==> CAPACITY|  
|                                                    File Name      ==> __  |  
|Line Cmds: I Insert D Delete S Select M Move     |  
|Cmd        Name                    Title           Cycles   Periods|  
|-          -                      -              -       -|  
|- CPU         CPU STD APPL -- Weeks                |  
|- WK0         WKL STD APPL -- Weeks                |  
|***** BOTTOM OF DATA *****|
```

Figure 1-29. Capacity Planning File Definitions Screen

This screen shows the capacity planning database file updates you currently have included in the production reporting jobstream and permits you to delete any of these file updates from the jobstream, rearrange the sequence of these entries, or add new file update entries. When you run the jobstream containing these capacity file update processes shown on this screen, the files will be updated in the same sequence that they are listed on this screen.

Specify D in the Cmd column to delete a capacity file update from the jobstream.

Specify M in this same column to move a capacity file update to before or after another update in the list. Specify B or A to indicate the position before or after which you want to move the capacity file update.

Specify I in the Cmd column to insert another shared capacity file update in the jobstream. If you do this, you will see the next screen in the sequence, the Shared File Definition Selection screen shown in Figure 1-30.

[illegible]

Figure 1-30. Shared File Definition Selection Screen

This screen displays the shared file definitions available for insertion in the jobstream. You can select one or more of these shared capacity file definitions into the jobstream by entering an S in the Cmd column. If you do so, you will next see the Select Shared File Definition screen shown in Figure 1-31.

```

/----- Select Shared File Definition -----\
|Command ==>|
|Enter a ? in any data entry field for more information on valid values.|

```



```
|Reporting Jobstream: CAP001 - CAP UPDATE 001|
|
|      File Name:    CI0
|      Title:       CICS - Weeks
|      Author:      CA
|
|Catalog Group Name  ==> CAPACITY
|      Title        ==> Capacity Planning
|
|Default Cycle(s)    ==> __ - __ (00-99)
|      No. Weeks     ==> ___ (1-999)
|
|Press END to add the file definition to the reporting jobstream.
|or
|Enter CANCEL to return to the catalog display without adding the file.
|
|-----|
```

Figure 1-31. Select Shared File Definition Screen

The Select Shared File Definition screen is displayed in response to any S you enter on the preceding screen. This screen displays the shared file definition selected on the previous screen and allows you to specify a new catalog group and override the CA MICS cycles or number of weeks or months for the update process. The update process for the capacity planning file described by this screen will be entered into the production job stream.

1.5.8 Worksheet Facility

The Worksheet Facility allows you to create, modify, produce, enter, and apply Capacity Planning worksheets. These worksheets are fill-in-the-blank reports that are suitable for gathering user forecasts.

These user forecasts can then be used as part of the capacity planning process to better understand your environment and size your configurations. An analytic report, the Worksheet Effect Report, is part of the Analytic Functions of the CA MICS Capacity Planner. Presentation Graphics can display the effect of these user forecasts in conjunction with any statistical forecasts you may have produced.

For more information on the production and maintenance of worksheets and on the new Worksheet Effect Report, refer to Chapter 8. For information on Presentation Graphics, refer to Chapter 4.

1.5.9 Selecting Files for MICF Analysis

Once you have built some capacity planning files, you can use the CA MICS Information Center Facility (MICF) to generate other reports or graphics. Using MICF, you can create ad hoc reports, analyses, or graphs based on information stored in your capacity planning database or the shared capacity planning database.

To access MICF, select option 2, CA MICS Information Center Facility (MICF), from the CA MICS Workstation Facility (MWF). Select option 2, Database Inquiries on the MICF main menu.

To create a new inquiry, use the insert line command (that is, type an I in the Cmd column). Fill in the fields for inquiry name, title, author, catalog group name and title.

You are creating a structured inquiry, so select option 2, Structured Inquiry.

To select the capacity planning file you want to use as input to this analysis, select option 7, Data Manipulation, and then option 6, Save File Facilities. On the Saved File Facilities main screen, select option 3, Capacity Planning File Selection.

The Capacity Planning File Selection screen enables you to locate and retrieve information from the capacity planning database.

Type the name of the file you want to use as input to the inquiry in the Input File field (or use the scrollable, bottom half of the screen to locate the file name and select it.)

Type the File Type of the file (that is, actuals, forecast, or worksheet). The file type default is Actuals. (For a full description of the different file types, see Chapter 3 of this guide.)

Now that you've selected your file, MICF will automatically set up the necessary SAS and JCL to allocate and read that file into the inquiry you are building.

At this point, you can use any of the MICF functions to manipulate the data in this file. For more information on MICF, see the CA MICS MICF User Guide.

1.5.10 Placing Analysis Inquiries in the SHARED Catalog

The Capacity Planner Analysis Inquiries can be placed into the SHARED MICF catalog if you so desire. Once in the SHARED catalog all Capacity Planner users can then share these inquiries or copy them to their own private catalog for further modification.

Since these inquiries can be executed by anyone with authorization to use the Capacity Planner, they must reference files that exist in the SHARED Capacity Planner File Catalog and Database. During migration of these inquiries from the private to the SHARED catalogs, the SHARED catalog will be checked to ensure that the required files have been previously defined and migrated. If the required files are not present in the SHARED catalog, the migration process will be terminated and appropriate error messages issued.

By adhering to this standard you avoid possible conflicts with your security system, particularly if you wish to place these inquiries into Production Reporting. Any query that is to be executed by the Production Reporting facility must reside in the SHARED catalog.

The CA MICS Retrieval and Reporting Guide's Appendix C describes how to maintain the SHARED and private inquiry catalogs.

To promote user defined inquiries to the Capacity Planning SHARED catalog, you must first cause all required files to be defined in the CAP SHARED catalog. If the required file(s) were created by a Capacity Planning Standard application, you must promote that application, along with its files, or alternatively promote the application and then execute the application from the SHARED catalog to generate its files. If the required file(s) were created by a Capacity Planner Resource File definition, you must first promote the Resource Definition, and then execute the Resource Definition from the CAP SHARED catalog.

The basic rule regarding query promotion and execution is this: the file(s) must be defined in whatever catalog is used to contain and execute the query itself.

The Forecasting Analyses (Univariate Modeling, Multivariate Regression, and Business Element Forecasting) all have the capability to generate a forecast file that is then saved in the Capacity Planner database (for more information on the naming/structure of these files please see Chapter 3 of this

guide).

The default target of these forecast analysis operations is the same as the database used for maintaining the history files. If you want to share this data, you must use the Capacity Planner Administration Facility to migrate these file(s) to the SHARED catalog and database before attempting to migrate any database analysis inquiries that require these files.

You can override this default through the use of a new Execution-Time Option, Forecast Target, which enables you to specify that the Forecast File should be written to the SHARED Capacity Planner database directly. Note that you then must have the necessary authorization from your security administrator to update the SHARED Capacity Planner database.

When a Forecast Inquiry is placed into Production Reporting the target of the forecast will always be the SHARED Capacity Planner database. (For more information, see the CA MICS Retrieval and Reporting Guide.)

Chapter 2: SAMPLE PLANNING APPLICATIONS

The design and implementation of a capacity planning database can be time-consuming for a data center beginning a capacity planning methodology. The CA MICS Capacity Planner Option sample applications minimize the time required to build the capacity planning database.

Sample applications are predefined views of capacity planning information contained in the CA MICS database. For each sample application, most of the design decisions have already been made, including which CA MICS files will be accessed, which elements will be loaded into the database, and how the elements will be summarized. Only a few design decisions remain for the person implementing the sample application.

Figure 2-0.01 shows an overview of the sample application specification and execution process. In the first phase, you specify the site-dependent control parameters. In the process phase, you save and execute the control parameters, creating the data files and printed output. You can perform these tasks together, or store the specifications for later execution.

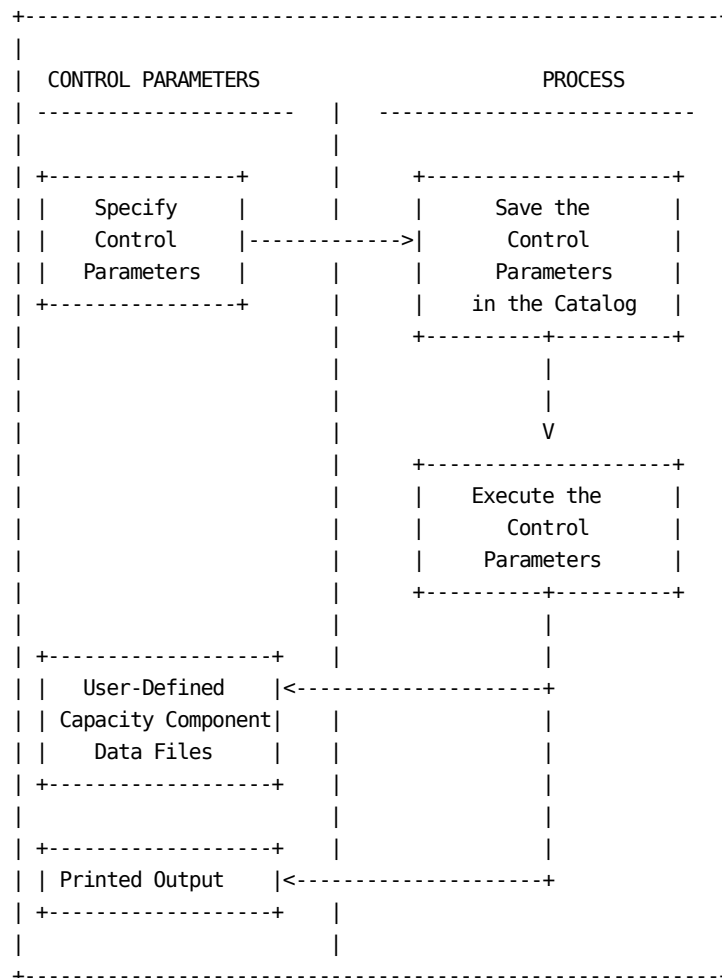


Figure 2-0.01. Sample Planning Application Overview

Although the sample applications are designed based on techniques and methodologies used by a wide spectrum of capacity planners, no single capacity planning methodology can satisfy the requirements of all data centers. For this reason, it is important for you to be aware of the following:

- o Using the sample applications is entirely optional. While they provide a productive and useful set of data files, you can also define your own user-variable files using other facilities of the CA MICS Capacity Planner Option.
- o A sample application that initially satisfies your capacity planning requirements may need future modifications to continue to meet your needs. Once you become familiar with the CA MICS Capacity Planner Option, you will be able to modify the generated sample application statements as needed.

The CA MICS Capacity Planner Option sample applications require that specific CA MICS files be active at either the DETAIL, DAYS, WEEKS, or MONTHS timespan. A matrix of the requirements is shown next:

The first group of applications is known as Mainframe Platform Planning.

CA MICS Files

		H		H		W		S		V		S		H		v		V	
Mainframe		A		A		L		C		M		C		A		C		C	
Platform		R		R		M		P		X		P		R		A		A	
Applications		C		L		S		P		S		X		C		V		D	
		P		P		E		A		G		C		F		O		A	
		U		C		C		G		P		M		P		A		A	
-----		---		---		---		---		---		---		---		---		---	
z/OS Processor Plan.		R				O													
-----		---		---		---		---		---		---		---		---		---	
WLM Workload Planning		R				R		R											
-----		---		---		---		---		---		---		---		---		---	
WLC Implementation		R		R															
-----		---		---		---		---		---		---		---		---		---	
IRD Implementation		R		R															
-----		---		---		---		---		---		---		---		---		---	
XCF Resource Mgmt.												R							
-----		---		---		---		---		---		---		---		---		---	
Coupling Facility Management														R					
-----		---		---		---		---		---		---		---		---		---	
z/VM Processor Fore.										R									
-----		---		---		---		---		---		---		---		---		---	
SMS Resource Fore.																R		R	
-----		---		---		---		---		---		---		---		---		---	

The next group of applications is known as Mainframe Platform Subsystems Planning.

	C I I S D D D M W T
	I M M N B B B Q E C
z/OS Subsystems	C S S T 2 2 2 R B P
Sample	C I I N D D D M W A
Applications	S S S C S S S M P P
	U U Y L U D Y A R I
-----	-----
CICS Resource Plan.	R
-----	-----
IMS Resource Planning	R R
-----	-----
DB2 Resource Planning	R R R
-----	-----
WebSphere MQSeries	R
-----	-----
Web Server Resource	R
-----	-----
TCP/IP Network Mgnt.	R R
-----	-----

The last group of applications is known as Distributed Platform Planning.

	T T D D A A T
Distributed	D D E E U U M
Systems	M M M M M M C
Sample	C S M S G A T
Applications	P Y O P P P M
	E P D R R P D
-----	-----
Tandem CPU Management	R O
-----	-----
VAX/VMS CPU Mgnt.	R R
-----	-----
HP Openview UNIX Mgt.	R R
-----	-----
Tape Data Set Mgt.	R
-----	-----

R-Required O-Optional

The sample applications use DETAIL, DAYS, WEEKS, or MONTHS as input. The applications are summarized to DAYS (DETAIL is summarized to DAYS), WEEKS, or MONTHS.

The following sections discuss the sample applications:

Mainframe Platform Planning:

- 2.1 - z/OS Processor Planning
- 2.2 - WLM Workload Planning
- 2.3 - WLC (Workload License Charging) Implementation Planning
- 2.4 - IRD (Intelligent Resource Director) Implementation Planning
- 2.5 - XCF Resource Planning
- 2.6 - Coupling Facility Planning
- 2.7 - z/VM Processor Planning
- 2.8 - SMS Resource Planning

Mainframe Platform Subsystem Planning:

- 2.9 - CICS Resource Planning
- 2.10 - IMS Resource Planning
- 2.11 - DB2 Resource Planning
- 2.12 - WebSphere MQSeries Planning
- 2.13 - Web Server Resource Planning
- 2.14 - TCP/IP Network Resource Planning
- 2.18 - Tape Data Set Resource Planning

Distributed Platform Planning:

- 2.15 - Tandem CPU Resource Planning
- 2.16 - VAX/VMS CPU Resource Planning
- 2.17 - HP Openview UNIX Resource Planning

This section contains the following topics:

- [2.1 z/OS Processor Planning](#) (see page 93)
- [2.2 WLM Workload Planning](#) (see page 109)
- [2.3 WLC Implementation Planning](#) (see page 138)
- [2.4 IRD Implementation Planning](#) (see page 151)
- [2.5 XCF Resource Planning](#) (see page 165)
- [2.6 Coupling Facility Planning](#) (see page 174)
- [2.7 z/VM Processor Planning](#) (see page 184)
- [2.8 SMS Resource Planning](#) (see page 195)
- [2.9 CICS Resource Planning](#) (see page 207)
- [2.10 IMS Resource Planning](#) (see page 215)
- [2.11 DB2 Resource Planning](#) (see page 224)
- [2.12 WebSphere MQSeries Planning](#) (see page 236)
- [2.13 WEB Server Resource Planning](#) (see page 245)
- [2.14 TCP/IP Network Resource Planning](#) (see page 253)
- [2.15 Tandem CPU Resource Planning](#) (see page 262)
- [2.16 VAX/VMS CPU Resource Planning](#) (see page 279)
- [2.17 HP Openview UNIX Resource Planning](#) (see page 294)
- [2.18 Tape Data Set Resource Planning](#) (see page 303)

2.1 z/OS Processor Planning

The z/OS Processor Planning Sample Application provides a capacity planning database file that contains the necessary elements for processor tracking and usage reporting.

The following sections cover the use, input specifications, and output of the sample application:

- 1 - Usage Guidelines
- 2 - Control Parameters
- 3 - Standard Output

2.1.1 Usage Guidelines

The z/OS Processor Planning Sample Application provides a capacity planning database file with the database elements you need for basic tracking, reporting, and forecasting of the processor workload. The information for this application is derived from either the DETAIL, DAYS, WEEKS, or MONTHS timespan of the CA MICS CPU Processor Activity (HARCPU) File.

Note: If you select the CPU Time Adjustment option, some additional data is used from the CA MICS Service Class Resource Consumption (WLMSEC) File.

Consequently, the z/OS Processor Planning Sample Application permits your data center to track and forecast processor usage at a high level of summarization (that is, at the SYSID and ZONE, HOUR, or DAYNAME level). Because the data is derived from the HARCPU File, this sample application provides information about system-wide processor use, rather than individual workload use.

In addition to providing basic information about processor utilization, the z/OS Processor Planning Sample Application also stores data concerning average batch, started task, and active TSO users in the system. You may find these elements useful in tracking and reporting. Also, you can sometimes make certain correlations between these elements and total processor utilization, thereby providing a technique for forecasting processor utilization on the basis of average batch, started task, and TSO user growth. See Chapter 14, Workload Forecasting, for more details on this technique.

NORMALIZATION

Another useful feature that the z/OS Processor Planning Sample Application provides is the ability to normalize processor utilization to a base CPU. Processor busy time is an indicator that is commonly used in capacity planning for reporting and forecasting. However, as systems are saturated and upgraded to more powerful processors, a graph of processor busy looks more like a "saw-tooth" function than a smoothly increasing function. This is because more powerful processors can process the same amount of work in less time than their slower predecessors could, and therefore spend less time processing. This contradicts the fact that the data center has really increased its processing capability with the upgrade.

You can circumvent this situation by normalizing all processor busy time to the equivalent of a base CPU. You can optionally request that the processor busy time and processor percent busy elements be adjusted to a base CPU of your specification. To use this feature, you must supply the following information:

- o The base CPU model type or name
- o The number of internal CPUs contained in the processor
- o The base CPU service unit to CPU second conversion factor

If the CPU Time Adjustment option is chosen, both the normalized and the unnormalized processor busy and percent utilization elements are still stored in the database and are available for reporting and analysis.

We recommend that you use this option to normalize workloads, since trending and forecasting CPU utilization without normalization is extremely difficult and often misleading.

NORMALIZATION EXAMPLE

Assume that a data center installed the z/OS Processor Planning Sample Application. They chose the optional CPU Time Adjustment and specified an IBM 2064-2C1 as the base CPU. The number of processors is 1 and the service unit coefficient is 14692.3.

These parameters, entered on the CPU Time Adjustment Parameters control screen, are then used to normalize CPU time values for any processor appearing in the CA MICS HARCPU File as the equivalent of an IBM 2064-2C1 processor.

Assume that one of the processors in the HARCPU File for this data center is an IBM 2084-302. During one week, this processor had a CPU busy time of 48 hours in 40 hours of operation. Since the 2084-302 has two internal processors, we can calculate the average percent busy for the 2084-302 for the week as:

$$100*(48/(2*40)) = 60\%.$$

The vendor has stated that the 2084-302 processor has a service unit to CPU second conversion factor of 20752.2, and we have chosen to use this factor for our normalization. Therefore, the normalized CPU busy time is:

$$48 * (20752.2 / 14692.3) = 67.79 \text{ hours}$$

In other words, this technique estimates that a workload requiring approximately 48 hours of CPU time on an IBM 2084-302 requires approximately 67.79 CPU hours on an IBM 2064-2C1. Since the 2064-2C1 has only one internal processor, the adjusted CPU percent utilization would be:

$$100 * (67.79 / (40 * 1)) = 169.47 \%$$

While it is impossible for a processor to run at 169% utilized, the calculation shows what the estimated utilization would be if the workload were running on the IBM 2064-2C1 processor. That is, we would need a minimum of two 2064-2C1 processors to process the work being done currently by the 2084-302.

Note that the use of an actual, commercially available processor as the base CPU is not required, and may not even be desirable for the data center. It would be perfectly acceptable to specify a hypothetical processor. For example:

```
Base CPU Model = MY_CPU
Number of Processors = 1
Service Unit Coefficient = 100
```

These parameters adjust all processor busy elements to the equivalent of a hypothetical processor having one internal CPU with a performance rating of 100 service units per CPU second. The fact that such a processor may not actually exist is not important, but rather that all processor time is being converted to a common unit that can have trends and be examined smoothly and consistently over time.

Also note that these calculations make no attempt to apply any analytic queuing formulas or benchmark estimates to compare one processor to another. The basic comparison of processors is calculated as the ratio of the service unit to CPU second conversion factors for the respective processors. These numbers can either be supplied by mainframe manufacturers or established by the data center through benchmark testing.

2.1.2 Control Parameters

You need relatively few parameters to implement the z/OS Processor Planning Sample Application. The control screens included in the online facility allow you to quickly enter the information required.

```

/----- z/OS Processor Planning -----\
|Command ==>|
|Enter a ? in any data entry field for more information on valid values.|
|Composing CAP DB File def.: CPU - z/OS Processor Sample Application|
|
|Summarization Level|
|      Primary      ==> MONTHS (DAYS/WEEKS/MONTHS)|
|      Secondary    ==> ZONE   (ZONE/HOUR/DAYNAME)|
|
|Input File Time-Span ==> DETAIL (DETAIL/DAYS/WEEKS/MONTHS)|
|
|Amount of data retained ==> 60 (1-9999 months)|
|Age of oldest forecast ==> 24 (1-9999 months)|
|History cycles ==> 3 (1-99 cycles)|
|Forecast cycles ==> 3 (1-99 cycles)|
|
|Data base ID (s) ==> _____|
|
|
|Specify extended options ==> Y (Y/N)|
\-----/

```

Figure 2-1.01. z/OS Processor Planning Control Screen

Figure 2-1.01 shows the z/OS Processor Planning screen, which is the main control screen for the z/OS Processor Planning Sample Application. It contains the following fields:

Summarization level:	The output summarization corresponds to the combination of the primary and secondary summarization levels. These control how the output file is summarized in your capacity planning database.
Input file timespan:	The timespan (either DETAIL, DAYS, WEEKS or MONTHS) of the HARCPU File that you want to use as input when creating the capacity planning database file. If you plan to select the CPU Time Adjustment option, the WLMSEC File must also be present at this timespan.
Amount of data retained:	The number of days, weeks, or months of historical data to be retained in the CAPxxx01 cycle of the capacity planning database file. Any historical observations older than what you specify for this parameter are dropped from the file during the file update.
Age of oldest forecast:	The number of days, weeks, or months that you want to retain a forecast after it is generated. Any forecast older than what you specify for this parameter is dropped when a new forecast is generated and saved.
History cycles:	The number of cycles or versions of the historical data that you want to retain. A new cycle is created when the file is updated and the existing cycles aged (for example, Cycle 01 is aged to Cycle 02, and Cycle 02 becomes Cycle 03).
Forecast cycles:	The number of cycles or versions of the forecast data that you want to retain. A new cycle is created when you specify forecast aging in one of the workload forecasting routines and the existing cycles are aged (for example, Cycle 01 is aged to Cycle 02, and Cycle 02 becomes Cycle 03).
Data base ID(s):	The database unit identifiers (up to fourteen units) that contain the HARCPU and WLMSEC Files to be used with this sample application. If you are uncertain in which database unit the HARCPU and WLMSEC Files reside, type a question

mark (?) in this field to display a message listing of the valid database unit identifiers for this sample application.

Specify If you want extended options, enter Y (YES), and
 Extended the Extended Options Menu shown in Figure 2-1.02
 Options: is displayed.

```

/----- Extended Options Menu -----\
|Command ==>|
|Inquiry Step: z/OS Processor Planning Sample Application|
|Combine all SYSIDs into a composite ==> N (Y/N)|
|  Composite SYSID ==> ____|
|Combine all zones into a composite ==> N (Y/N)|
|  Composite zone ==> _____|
|Specify:|
|  Common data selection parameters ==> Y (Y/N/R)|
|  Global data selection exit ==> Y (Y/N/R)|
|  CPU time adjustment parameters ==> Y (Y/N/R)|
|  Notes report enabled ==> YES (YES/NO)|
|  Contents reports enabled ==> YES (YES/NO)|
|-----\

```

Figure 2-1.02. z/OS Processor Planning Extended Options Menu

The extended options available with the z/OS Processor Planning Sample Application (Figure 2-1.02) are:

Combine all SYSIDs in the capacity planning
SYSIDs into database file. Enter a Y (YES) if you want data
a composite: for all of the SYSIDs to be summarized into a
single observation with the Composite SYSID
value.

Combine all zones, hours and daynames in the
zones/hours/ capacity planning database file. The prompt for
daynames this field changes to zones, hours, or daynames
into a depending on what you specify for the summary
composite: level parameter. Enter a Y (YES) if you want
all of the zones, hours, and daynames to be
summarized into a single observation with the
composite value.

Specify SYSIDs and zones, hours or daynames
common data to use for data selection in the CA MICS input
selection files. Enter Y (YES) to display the Common
parameters: Data Selection screen shown in Figure 2-1.03.
Enter R (RESET) to reset any specifications you
previously made to SYSID and zone, hour, or
dayname.

Specify You can use your own SAS statements to select
global data data from the CA MICS input files. Enter Y
selection (YES) to display the SAS System Statements
exit: screen shown in Figure 2-1.04. Enter R (RESET)
to reset any specifications you previously made
to the default of no user-specified data
selection.

Specify CPU Specify the parameters to use for CPU time
time adjustment (normalization), if you want this
adjustment optional feature. Enter Y (YES) to display the
parameters: CPU Time Adjustment Parameters screen shown in
Figure 2-1.05. Enter R (RESET) to reset any
specifications you previously made to CPU time
adjustment.

This option is not available if you specify the
composite SYSID option on the CPU Planning
screen shown in Figure 2-1.01. If you select
CPU time adjustment, you must also provide
database units that contain the WLMSEC File on
the CPU Planning screen.

The Common Data Selection screen, shown in Figure 2-1.03, allows you to define the following data selection criteria:

- SYSID:** The SYSIDs that you want to include in your capacity planning database file. You can enter a maximum of 15 one- to four-character SYSIDs. Leave this parameter blank to include data for all SYSIDs.
- Zone:** You can specify up to nine zones to be included in the capacity planning database file. Leave this parameter blank to include data for all zones.
- Hour:** Enter the five hours or hour ranges to be included in the capacity planning database file. Specification of this parameter is valid only if you specify the DETAIL, DAYS, or WEEKS timespan on the z/OS Processor Planning screen shown in Figure 2-1.01. Leave this parameter blank to include data for all hours.
- Dayname:** Enter up to five daynames to be included in the capacity planning database file. Leave blank to include all days. Input the first unique characters and the application will fill in the rest. Specification of this parameter is only valid when the input timespan is DETAIL or DAYS.

```

/----- SAS System Statements ----- COLUMNS 001 072--\
|Command ==>                               Scroll ==> PAGE |
|
|SAS Statements for the SELECT Macro          |
|-----|
|***** ***** TOP OF DATA *****|
|==MSG> SPECIFY SAS STATEMENTS TO SUBSET THE CAPACITY PLANNING DATA BASE FILE.|
|NOTE= %MACRO SELECT; <=== PLEASE DO NOT PLACE LINES BEFORE THIS LINE.      |
|000001                                     |
|000002                                     |
|000003                                     |
|000004                                     |
|000005                                     |
|000006                                     |
|000007                                     |
|000008                                     |
|000009                                     |
|000010                                     |
|NOTE= %MEND SELECT; <=== PLEASE DO NOT PLACE LINES AFTER THIS LINE.        |

```

```
***** ***** BOTTOM OF DATA *****
|
|
\-----/
```

Figure 2-1.04. SAS System Statements Screen-z/OS Planning Global Data Selection

The SAS System Statements screen shown in Figure 2-1.04 allows you to define your own data selection logic. You can code any valid SAS statements, as described in the SAS guides. The %MACRO SELECT and %MEND SELECT statements are already provided, and are field-protected from modification.

Note: This is a PDF/EDIT screen. All PDF/EDIT commands are available.

```
/----- CPU Time Adjustment Parameters -----\
|Command ==>
|
|Inquiry step: z/OS Processor Planning Sample Application
|
|Base CPU model                ==> 2094-722
|Number of processors in base CPU model ==> 22      (1-99)
|Base CPU service unit coefficient ==> 19777.5   (per processor)
|
|
|
|
|
|
|
|
|
|
|
```

Figure 2-1.05. CPU Time Adjustment Parameters Screen

The information on the CPU Time Adjustment Parameters screen shown in Figure 2-1.05 is required only if you choose CPU time normalization for the z/OS Processor Planning Sample Application.

Base CPU model: The 10-character identifier for the base processor. Any name is acceptable in this field. Do not use embedded blanks or equal signs (=) in this value. For example, if you want to normalize your actual CPU time values to the equivalent of a 2094-722, you would code 2094-722 in this field.

Number of processors in base CPU model: The number of internal CPUs in the base processor. This value must be an integer (that is, do not use a decimal point), and it must be greater than 0 and less than or equal to 99. For example, if you specify a 2094-722 as the base CPU model, then you would code a 22 in this field since a 2094-722 has twenty two internal CPUs.

Base CPU service unit coefficient: The service-units-to-CPU-seconds conversion factor for the base processor. The value must be a real number greater than 0. You can compute a relative performance rating by dividing this value by the corresponding service-unit-to-CPU-second factor for the processor represented by the SYSID (see Section 2.1.1 for an example of this calculation). You can use vendor-supplied values for the factor, or a factor based on your own evaluations. For example, if you specify a 2094-722 as the base CPU model, you would code 19777.5 in this field.

2.1.3 Standard Output

The capacity planning database file that is generated by the z/OS Processor Planning Sample Application contains the following information:

- o Data elements common to all CA MICS Capacity Planner database files:

SYSID
HOUR or ZONE or DAYNAME
YEAR
MNTHDAY or WEEK or MONTH
DATE

- o Data elements that are unique to the z/OS Processor Planning Sample Application database file:

- Character format data elements

CPUMODEL - CPU Model Type
MVSLEVEL - MVS Release
CPUMOD - Processor Model Family
CPCMODID - CPC Model Identifier

- Numeric format data elements

INTERVLS - Number of Recording Intervals
CPUNOCPU - Number of Internal Processors
CPUAVCPU - Avg Number of Assigned CP Processors
CPUTOONL - Total Online CP Processors
CPUAVONL - Avg Online CP Processors
CPUTOOZP - Total Online zAAP Processors
CPUAVOZP - Avg Online zAAP Processors
CPUTOOSU - Total Online zIIP Processors
CPUAVOSU - Avg Online zIIP Processors
CPUAVB - Avg Batch Users
CPUAVT - Avg TSO Users
CPUAVS - Avg Started Users
CPUAVR - Avg In & Ready Users
PRSMACP - Avg CP Processors in Complex
PRSMACDP - Avg CP Dedicated Processors
PRSMACVP - Avg CP Shared Logical Processors
CPUBTT - Total Batch Users
CPUTTT - Total TSO Users
CPUTTT - Total Started Users
CPURTT - Total In & Ready Users
CPURSMP - Total Number of Samples
PRSMTCP - Total CP Processors in Complex
PRSMTCDP - Total CP Dedicated Processors
PRSMTCVP - Total CP Shared Logical Processors
WLMSUCPU - SU:CPU-second Conversion Factor

SHR_CPS - Number of Shared CPs Defined in CEC

- Time format data elements

DURATION - Interval Duration
CPUTOOTM - Total CP Processors Online Time
CPUTOBTM - Total Time CP Processors Busy
CPUTODTM - Total CP Processors Dispatch Time
CPUTOVBT - Total CP Processors PR/SM MVS Busy Time
CPUZPBTM - Total Time zAAP Processors Busy
CPUZPDTM - Total Time zAAP Processors Dispatch Time
CPUZPOTM - Total zAAP Processors Online Time
CPUZPUTM - Total zAAP Processors Up Time
CPUSUBTM - Total Time zIIP Processors Busy
CPUSUDTM - Total Time zIIP Processors Dispatch Time
CPUSUOTM - Total zIIP Processors Online Time
CPUSUUTM - Total zIIP Processors Up Time
PRSMTEDT - Total Dispatch Time - Effective

- Percentage format data elements
 - CPUPCBSY - Pct CP Processors Busy (unadjusted)
 - CPUPCBZP - Pct zAAP Processors Busy (unadjusted)
 - CPUPCBSU - Pct zIIP Processors Busy (unadjusted)
- o Additional data elements if CPU time adjustment is specified:
 - Character format data element
 - BASECPU - Base Adjustment CPU Model Type
 - Numeric format data elements
 - ADJSERV - Base Adjustment Service Units/Second
 - ADJPRC - Base Adjustment Number of Processors
 - Time format data element
 - CPUADBTM - Base CPU Busy Time
 - Percentage format data elements
 - CPUADPCB - Base CPU Percent Busy

Many of the data elements used in the z/OS Processor Planning Sample Application database file have the same meaning and derivation as their identically named counterparts in the CA MICS HARCPU File. These data elements are CPUMODEL, MVSLEVEL, CPUNOCPU, CPUAVB, CPUAVT, CPUAVS, and DURATION.

The data elements that are derived from the CA MICS database and used by the z/OS Processor Planning Sample Application are calculated as follows:

CPUPCBSY: The percent busy for CP processors for the interval. It is calculated using the following formula:

```
IF CPUTOOTM THEN SELECT;
  WHEN (CPUTOBTM>0)
    CPUPCBSY=(CPUTOBTM*100)/CPUTOOTM;
  WHEN (CPUTOBTM=0) CPUPCBSY=0;
  OTHERWISE        CPUPCBSY=.;
END;
```

This element must have values between 0 and 100.

CPUPCBZP: The percent busy for zAAP processors for the interval. It is calculated using the following formula:

```
IF CPUZPOTM THEN SELECT;
  WHEN (CPUZPBTM>0)
    CPUPCBZP=(CPUZPBTM*100)/CPUZPOTM;
  WHEN (CPUZPBTM=0) CPUPCBZP=0;
  OTHERWISE        CPUPCBZP=. ;
END;
```

This element must have values between 0 and 100.

CPUPCBSU: The percent busy for zIIP processors for the interval. It is calculated using the following formula:

```
IF CPUZPOTM THEN SELECT;
  WHEN (CPUSUBTM>0)
    CPUPCBSU=(CPUSUBTM*100)/CPUSUOTM;
  WHEN (CPUSUBTM=0) CPUPCBSU=0;
  OTHERWISE        CPUPCBSU=. ;
END;
```

This element must have values between 0 and 100.

The additional data elements that are present if CPU time adjustment is specified are derived as follows:

BASECPU: The value that was specified for the base CPU model parameter on the CPU Time Adjustment Parameters screen shown in Figure 2-6.

CPUADBTM: The derivation of this value has two parts:

- o The SRM service unit to CPU time conversion factor WLMSUCPU is obtained from your WLMSEC File.
- o The value that you specify for the base CPU model parameter is used in the following formula:

$$\text{CPUADBTM} = \text{WLMSUCPU} * \text{CPUTOBTM} / \text{ADJSERV};$$

For example, if you specify a base CPU service model unit coefficient of 19777.5, then the following calculation is performed:

$$\text{CPUADBTM} = \text{WLMSUCPU} * \text{CPUTOBTM} / 19777.5;$$

CPUADPCB: This value is calculated using the value you specify for the number of processors in base CPU model parameter as follows:

$$\text{CPUADPCB} = 100 * (\text{CPUADBTM} * \text{CPUNOCPU}) / (\text{CPUTOOTM} * \text{ADJPRC});$$

For example, if you specify twenty two processors in your base CPU model, then the following calculation is performed:

$$\text{CPUADPCB} = 100 * (\text{CPUADBTM} * \text{CPUNOCPU}) / (\text{CPUTOOTM} * 22);$$

2.2 WLM Workload Planning

The WLM Workload Planning Sample Application provides a capacity planning database file that contains the necessary elements for workload tracking and reporting when utilizing WLM Goal mode.

The following sections cover the use, input specifications, and output of the sample application:

- 1 - Usage Guidelines
- 2 - Control Parameters
- 3 - Standard Output
- 4 - Goal Mode Planning Example

2.2.1 Usage Guidelines

The WLM Workload Planning Sample application provides a capacity planning database file with the database elements you need for tracking, reporting, and forecasting of the workloads on your system. The information for this application is primarily derived from either the DETAIL, DAYS, WEEKS, or MONTHS timespan of the CA MICS Service Class Resource Consumption (WLMSEC) file. The WLM Workload Planning Sample application also acquires additional planning elements from the CPU Activity (HARCPU) and the Paging Activity file (SCPPAG).

The WLM Workload Planning Sample application permits your installation to track and forecast resource consumption and transaction activity at a high level of summarization (that is, at the SYSID and ZONE, HOUR or DAYNAME level) for workloads. Because the data is derived from the WLMSEC file, this sample application provides information about individual workloads based on Service Classes.

In addition to providing basic information about workload processor utilization, the WLM Workload Planning Sample Application also stores data concerning application busy time, storage occupancy, I/O rates, and paging activity. You may find these elements useful in tracking and reporting. Also, you can sometimes make certain correlations between these elements and total processor utilization, thereby providing a technique for forecasting processor utilization on the basis of average workload growth. See Chapter 14, Workload Forecasting, for more details on this technique.

SERVICE CLASS WORKLOADS

The WLM Workload Planning Sample Application allows you to divide your service classes into groups, or workloads. For each defined workload, the application stores measures such as the number of ended transactions, CPU time consumed by the workload, total service units consumed by the workload, and DASD I/O statistics into the capacity planning database file.

You define Service Class Workloads by using one of two methods. The first method is to create a Workload Mapping Table in an 80-byte, fixed block data set or PDS member. The second method for creating Service Class Workloads is to write a CAPAPU exit using SAS code to specify the workloads to which each Service Class belongs.

Once defined, you can modify the workload definitions as often as necessary without having to respecify other Goal Mode Planning parameters.

NORMALIZATION

The WLM Workload Planning Sample Application provides the ability to normalize processor utilization to a base CPU. Processor busy time is an indicator commonly used in capacity planning for reporting and forecasting. However, as systems are saturated and upgraded to more powerful processors, a graph of processor busy time looks more like a saw-tooth function rather than a smooth, monotonically increasing function. This is because more powerful processors process the same amount of work in less time than their slower predecessors and therefore spend less time processing. This contradicts the fact that the installation has really increased its processing capability with the upgrade.

You can circumvent this situation by normalizing all processor busy time to the equivalent of a base CPU. You can optionally request that the Goal Mode busy time be adjusted to a base CPU of your specification. To use this feature, you must supply the following information:

- o The base CPU model type or name
- o The number of internal CPUs contained in the processor
- o The base CPU service unit to CPU second conversion factor (service unit coefficient)

If the CPU Time Adjustment option is chosen, both the normalized and the unnormalized Goal Mode busy elements are still stored in the database and are available for reporting and analysis.

It is recommended you use this option to normalize workloads, since trending and forecasting workload CPU utilization without normalization is extremely difficult and often misleading.

NORMALIZATION EXAMPLE

Assume an installation installed the WLM Workload Planning Sample Application. They chose the optional CPU time adjustment and specified an IBM 9121-742 as the base CPU. The number of processors is 4 and the service unit coefficient is 1351.68.

These parameters, entered on the CPU Time Adjustment Parameters control screen, are then used to normalize CPU time values for any processor appearing in the CA MICS WLMSEC file to the equivalent of an IBM 9121-742 processor. Assume that one of the processors in the WLMSEC file for this installation is an IBM 9672-RY5 (4 engines). During one week, one workload had a busy time of 50 hours in 100 hours of operation.

The vendor has stated that the 9672-RY5 (4 engines) processor has a service unit to CPU second conversion factor of 2902.76, and we have chosen to use this factor for our normalization. Therefore, the normalized Goal Mode busy time is:

$$50 * (2902.76 / 1351.68) = 107.38 \text{ hours}$$

In other words, this technique estimates that a workload requiring approximately 50 hours of CPU time on an IBM 9672-RY5 requires approximately 107.38 CPU hours on an IBM 9121-742.

Note that the use of an actual, commercially available processor as the base CPU is not required, and may not be desirable for the installation. It is acceptable to specify a hypothetical processor. For example:

```
Base CPU Model = MY_CPU
Number of Processors = 1
Service Unit Coefficient = 100
```

These parameters adjust all processor busy elements to the equivalent of a hypothetical processor having one internal CPU with a performance rating of 100 service units per CPU second. The fact that such a processor may not actually exist is not important, but rather that all processor time is being converted to a common unit that can be trended and examined smoothly and consistently over time.

Also note that these calculations make no attempt to apply any analytic queuing formulas or benchmark estimates to compare one processor to another. The basic comparison of processors is calculated as the ratio of the service unit to CPU second conversion factors for the respective processors. These numbers can be supplied by the mainframe manufacturers or can be established by the installation through benchmark testing.

2.2.2 Control Parameters

Relatively few parameters are needed to implement the Goal Mode Planning Sample Application. The control screens included in the online facility allow you to quickly enter the required information.

```

/----- WLM Workload Planning -----\
|Command ==>|
|Enter a ? in any data entry field for more information on valid values.|
|Composing CAP DB File def.: GM1 - WLM Workload Planning Sample Application|
|
|Summarization Level|
|    Primary      ==> MONTHS  (DAYS/WEEKS/MONTHS)|
|    Secondary    ==> ZONE    (ZONE/HOUR/DAYNAME)|
|
|Input File Time-Span  ==> DETAIL (DETAIL/DAYS/WEEKS/MONTHS)|
|
|Amount of data retained ==> 60   (1-9999 months)|
|Age of oldest forecast  ==> 24   (1-9999 months)|
|History cycles          ==> 3    (1-99 cycles)|
|Forecast cycles         ==> 3    (1-99 cycles)|
|
|Data base ID (s)       ==> P
|
|Specify extended options ==> Y (Y/N)
\-----/

```

Figure 2-2.01. WLM Workload Planning Screen

Figure 2-2.01 shows the WLM Workload Planning screen, which is the main control screen for the WLM Workload Planning Sample Application. It contains the following fields:

Summarization level:	The output summarization corresponds to the combination of the primary and secondary summarization levels. These control how the output file is summarized in your capacity planner database.
Input file timespan:	The timespan (either DETAIL, DAYS, WEEKS or MONTHS) of the HARCPU, WLMSEC and WLMSEM files that you wish to use as input when creating the Capacity Planner database file.
Amount of data retained:	The number of days, weeks or months of historical data to be retained in the CAPxxx01 cycle of the Capacity Planner database file. Any historical observations older than what you specify for this parameter are dropped from the file during the file update.
Age of oldest forecast:	The number of days, weeks or months that you want to retain a forecast after it is generated. Any forecast older than what you specify for this parameter is dropped when a new forecast is generated and saved.
History cycle:	The number of cycles or versions of the historical data you want to retain. A new cycle is created when the file is updated and the existing cycles aged (for example, Cycle 01 is aged to Cycle 02 and Cycle 02 becomes Cycle 03).
Forecast cycle:	The number of cycles or versions of the forecast data you want to retain. A new cycle is created when you specify forecast aging in one of the workload forecasting routines and the existing cycles are aged (for example, Cycle 01 is aged to Cycle 02 and Cycle 02 becomes Cycle 03).
Data base ID(s):	The database unit identifier(s) (up to fourteen units) that contain the WLMSEC and WLMSEM files to be used with this sample application. If you are uncertain in which database unit these files reside, type a question mark (?) in this field to display a message listing of the valid database unit identifiers for this sample application.

Specify extended options: If you want extended options, specify Y (YES), and the Extended Options Menu (Figure 2-2.02) is displayed.

```

/----- Extended Options Menu -----\
|Command ==>
|
|Inquiry Step: WLM Workload Planning Sample Application
|
|
|Combine all SYSIDs into a composite ==> N      (Y/N)
|  Composite SYSID                    ==> ----
|
|Combine all zones into a composite ==> N      (Y/N)
|  Composite zone                     ==> -----
|
|Specify:
|  Common data selection parameters ==> N      (Y/N/R)
|  Global data selection exit        ==> N      (Y/N/R)
|  Workload mapping definition       ==> Y      (Y/N/R)
|  Workload mapping (CAPAPU) exit    ==> N      (Y/N/R)
|  CPU time adjustment parameters   ==> N      (Y/N/R)
|  Notes report enabled              ==> YES    (YES/NO)
|  Contents reports enabled          ==> YES    (YES/NO)
|
|
|
|
|
|
|
|
|
|
|
\-----/

```

Figure 2-2.02. WLM Workload Planning Extended Options Menu

The extended options available with the WLM Workload Planning Sample Application (Figure 2-2.02) are:

Combine all SYSIDs in the capacity planning
SYSIDs into database file. Enter a Y (YES) if you want data
a composite: for all of the SYSIDs to be summarized into a
single observation with the Composite SYSID
value.

Combine all zones, hours, and daynames in
zones/hours/ the capacity planning database file. The prompt
daynames for this field changes to zones or hours or
into a daynames depending on what you specify for the
composite: summary level parameter. Enter a Y (YES) if you
want all of the zones, hours, and daynames to be
summarized into a single observation with the
composite value.

Specify SYSIDs and zones, and hours or
common data daynames to be used for data selection from the
selection CA MICS input file(s). Enter Y (YES) to display
parameters: the Common Data Selection screen (Figure
2-2.03.) Enter R (RESET) to reset any selections
you previously made to SYSID and zone, and hour
or daynames.

Specify You can use your own SAS statements to select
global data data from the CA MICS input file(s). Enter Y
selection (YES) to display the SAS System Statements
exit: screen (Figure 2-2.04). Enter R (RESET) to
reset any selections you previously made to the
default of no user- specified data selection.

Specify You can define a Workload mapping table. Enter
workload Y (YES) to display the Workload Mapping
mapping Definition screen (Figure 2-2.05), which
definition: provides you with the options for defining a
workload mapping table.

Specify You can use your own SAS statements to split
workload Service Classes into workload (CAPAPU) groups.
mapping Enter Y (YES) to display the SAS System
(CAPAPU) Statements screen.
exit:

Specify CPU Specify the parameters to be used CPU time
time for adjustment (normalization). Enter Y (YES) to
adjustment display the CPU Time Adjustment Parameters
parameters: screen (Figure 2-2.07). Enter R (RESET) to reset

any selections you previously made to CPU time adjustment.

This option is not available if you select the composite SYSID option on this screen.

You must select either the Workload mapping definition extended option or the Workload mapping (CAPAPU) exit to complete the Workload Planning Sample Application definition.

Notes	If you have entered notes associated with the
report	capacity planning database file definition, a
enabled:	Yes in this field prints the note as part of
	file updating.

Contents reports enabled:	A Yes in this field prints the contents of the capacity planning database file as part of file updating.

```

/----- Common Data Selection -----\
|Command ==>
|
|Inquiry Step:  WLM Workload Planning Sample Application
|
|SYSID ==>  _____
|           _____
|
|Zone  ==>  _ _ _ _ _ (1 - 9)
|
|Hour  ==>  __ - __ - __ - __ - __ - __ - __ (0 - 23)
|
|Dayname==>  _____
|
|
|
|
\-----\

```

Figure 2-2.03. Common Data Selection Screen - WLM Workload Planning

The Common Data Selection screen (Figure 2-2.03) allows you to define the following data selection criteria:

- SYSID:** The SYSIDs that you want to include in your capacity planning database file. You can enter a maximum of 15 one- to four-character SYSIDs. Leave this parameter blank to include data for all SYSIDs.
- Zone:** You can specify up to nine zones to be included in the capacity planning database file. Leave this parameter blank to include data for all zones.
- Hour:** Specify the five hours or hour ranges to be included in the capacity planning database file. Selection of this parameter is valid only if you enter the DETAIL, DAYS, or WEEKS timespan on the WLM Workload Planning screen (Figure 2-2.01). Leave this parameter blank to include data for all hours.
- Dayname:** You can enter up to five daynames to be included in the capacity planning database file. Leave blank to include all days. Input the first unique characters and the application will fill in the rest. Selection of this parameter is only valid when the input timespan is DETAIL or DAYS.

```

/----- SAS System Statements ----- COLUMNS 001 072--\
|Command ==>                               Scroll ==> PAGE |
|
|SAS Statements for the SELECT Macro          |
|-----|
|***** ***** TOP OF DATA *****|
|==MSG> SPECIFY SAS STATEMENTS TO SUBSET THE CAPACITY PLANNING DATA BASE FILE. |
|NOTE= %MACRO SELECT; <== PLEASE DO NOT PLACE LINES BEFORE THIS LINE.         |
|000001                                     |
|000002                                     |
|000003                                     |
|000004                                     |
|000005                                     |
|000006                                     |
|000007                                     |
|000008                                     |
|000009                                     |
|000010                                     |

```

```

|=NOTE= %MEND SELECT; <=== PLEASE DO NOT PLACE LINES AFTER THIS LINE.
|***** ***** BOTTOM OF DATA *****|
|
|-----/

```

Figure 2-2.04. SAS System Statements Screen - WLM Workload Planning

The SAS System Statements screen (Figure 2-2.04) lets you define your own data selection logic. You can code any valid SAS statements as described in the SAS manuals. The %MACRO SELECT and %MEND SELECT statements are already provided, and are field protected from modification.

Note: This is a PDF/EDIT screen. All PDF/EDIT commands are available.

```

/----- Workload Mapping Definition -----\
|Command ==>|
|
|Inquiry step: WLM Workload Planning Sample Application|
|
|Workload mapping table data set name|
|  ==> -----|
|
|Edit workload mapping table      ==> NO (YES/NO)|
|
|
|
|
|
|
|
|
|
|
|
|-----/

```

Figure 2-2.05. Workload Mapping Definition Screen

The Workload Mapping Definition screen (Figure 2-2.05) offers two options for defining a Workload Mapping Table: you can name an externally created data set containing valid workload mapping statements or you can define the workload mapping statements online. The following fields are included on this screen:

Workload mapping table data set name:	The name of an externally created data set that contains valid workload mapping statements. This parameter can be either a sequential data set or the member of a partitioned data set. Code this parameter as you would any other data set name. If the data is in a member of a partitioned data set, include the member name as part of this parameter. If the data set name is cataloged under your userid, do not enter your userid as part of the data set name. If you code a fully qualified data set name (that is, including a userid), enclose this data set name in single quotes.
Edit workload mapping table:	The specification that allows you to edit the workload mapping table. Enter Y (YES) if you want to invoke the online editor to modify workload mapping statements in the data set that is named on the workload mapping table data set name parameter.

WORKLOAD MAPPING TABLE

The Workload Mapping Table is specified using the ISPF/PDF Editor, which will be invoked if you specify Y (YES) for the Edit workload mapping table option. The table contains two types of entries, the first of which defines the workload group names and titles. The format of the entry is

WORKLOAD group-name group-title

where group-name is a one- to 12-character workload group name and group-title is a one- to 40-character name of the group. Group-title is used to build the DESCRIPT data element for each workload group.

The second entry type specifies which Service Classes are included in each group. The format is

srvclass group-name

where `srvc` is the one- to eight-byte service class name and `group-name` matches a group-name given in a `WORKLOAD` statement. Any service class that you do not specify on a workload mapping statement defaults to a workload group of `OTHER`.

An example of a very simple workload mapping table follows. Note how the descriptions are coded. There is ample use of non-text characters and embedded blanks. Do not use quotation marks to surround the descriptions.

```

WORKLOAD IMSPROD PRODUCTION IMS REGIONS
WORKLOAD BATCH ALL PRODUCTION BATCH
WORKLOAD TBATCH ALL TEST BATCH
WORKLOAD TSO ALL TSO
TSODEV  TSO
PRODCNTL TSO
IMSAPPL1 IMSPROD
IMSAPPL2 IMSPROD
HOTBATCH BATCH
REGBATCH BATCH
DEVBATCH TBATCH

```

Specifying this workload mapping would cause `WLMSEC` and `WLMSEM` data for both the `IMSAPPL1` and `IMSAPPL2` Service Classes to be combined into a workload entry named `IMSPROD`. Likewise, specifying this workload mapping would cause the `DEVBATCH` service class to be left isolated in its own workload called `TBATCH`.

```

/----- SAS System Statements ----- COLUMNS 001 072--\
|Command ==>                               Scroll ==> PAGE |
|
|SAS Statements for the CAPAURT Macro          |
|-----|
|***** ***** TOP OF DATA *****|
|==MSG> SPECIFY SAS STATEMENTS TO SUBSET THE CAPACITY PLANNING DATA BASE FILE. |
|NOTE= %MACRO CAPAURT; <=== PLEASE DO NOT PLACE LINES BEFORE THIS LINE.      |
|000001|
|000002 SELECT(WLMCLASS);                |
|000003   WHEN(HOTBATCH) CAPAPU = 'BATCH'; |
|000004   WHEN(SLOBATCH) CAPAPU = 'BATCH'; |
|000005   WHEN(TSO) D0;                   |
|000006       IF SYSID = 'PROD'           |
|000007       THEN CAPAPU = 'PROD TSO';   |
|000008       ELSE CAPAPU = 'TEST TSO';   |
|000009   END;                           |

```

```
|000010    OTHERWISE    CAPAPU = 'OTHERSTF';          |
|000011 END;                                              |
|000012                                              |
|000013                                              |
|000014                                              |
|000015                                              |
|000016                                              |
|000017                                              |
|=NOTE= %MEND CAPAURT; <=== PLEASE DO NOT PLACE LINES AFTER THIS LINE. |
|***** ***** BOTTOM OF DATA *****|
|
|-----/
```

Figure 2-2.06. Workload Mapping (CAPAPU) Exit Screen - WLM Workload Planning

The information on the CPU Time Adjustment Parameters screen (Figure 2-2.07) is required only if you choose CPU time normalization for the WLM Workload Planning Sample Application.

Base CPU model: The 10-character identifier for the base processor. Do not use embedded blanks or equal signs (=) in this value. For example, if you want to normalize your actual CPU time values to the equivalent of a 1000 ZZ, then code 1000ZZ in this field.

Number of processors in base CPU model: The number of internal CPUs in the base processor. This value must be an integer greater than 0 and less than or equal to 16.

Base CPU service unit coefficient: The service-units-to-CPU-seconds conversion factor for the base processor. The value must be a real number greater than 0. You can compute a relative performance rating by dividing this value by the corresponding service-unit-to-CPU-second factor for the processor represented by the SYSID (see Section 2.2.1 for an example of this calculation). You can use vendor-supplied values for the factor, or a factor based on your own evaluations.

2.2.3 Standard Output

The capacity planning database file that is generated by the WLM Workload Planning Sample Application contains the following data elements:

- A. Data elements common to all CA MICS Capacity Planner database files:

SYSID
ZONE or HOUR or DAYNAME
YEAR
MONTH or WEEK or MNTHDAY
DATE

B. Data elements unique to the WLM Workload Planning
Sample Application database file:

1. Character format data elements

CAPAPU - Workload Group
MVSLEVEL - MVS Software Level
RMFVER - RMF Version Number
CPUMOD - Processor Model Family
CPCMODID - CPC Model Identifier

2. Numeric format data elements

INTERVLS - Number of Recording Intervals
CPUNOCPU - Number of Internal Processors
CPUAVCPU - Avg Number of Assigned CP Processors
CPUTOONL - Total Online CP Processors
CPUAVONL - Avg Online CP Processors
CPUTOOZP - Total Online zAAP Processors
CPUAVOZP - Avg Online zAAP Processors
CPUTOOSU - Total Online zIIP Processors
CPUAVOSU - Avg Online zIIP Processors
CPUAVR - Avg In & Ready Users
PRSMACP - Avg CP Processors in Complex
PRSMACDP - Avg CP Dedicated Processors
PRSMACVP - Avg CP Shared Logical Processors
CPURTT - Total In & Ready Users
CPURSMP - Total Number of Samples
PRSMTCP - Total CP Processors in Complex
PRSMTCDP - Total CP Dedicated Processors
PRSMTCVP - Total CP Shared Logical Processors
PAGPIN - Non-VIO Non-Swap Page-Ins
RMFSYNCH - Intvls with RMF Synchronized to SMF
SECAVMPL - Avg Multiprogramming Level
SECAVPRS - Avg Storage Used - CS+ES
SECAVRTE - Avg Transaction Service Rate
SECAVSRS - Avg Storage Used - Shared Pages
SECAVTRN - Avg Transactions Active
SEDCDNT - DASD I/O Count
SECPSHPI - Page-Ins per Second - Hiperspace
SECPINS - SUs per Res Sec (Absorption Rate)
SECPSPIE - Page-Ins per Second - Expanded
SECPSPIN - Page-Ins per Second - Total
SECPSRCP - Transactions per Second
SECPSSPA - Shared Page-Ins from Auxiliary per Sec
SECPSSPE - Shared Page-Ins from ES per Sec
SECPSSRV - Service Units per Elapsed Second
SECTOARC - Transactions Completed Abnormal

SECTOBKA - Blocks Paged In - Auxiliary
SECTOBKE - Blocks Paged In - Expanded
SECTOBPE - Block Page-Ins - Exp
SECTOBPI - Block Page-Ins - Auxiliary
SECTOCPU - CPU Service Units
SECTOCRM - ESO Hiperspace Read Misses
SECTO CRS - Page Residency - CS
SECTOCUS - Samples - Using CPU
SECTOERS - Page Residency - ES
SECTOHPI - Page-Ins - Hiperspace
SECTOIOC - I/O Service Units
SECTOMSO - Storage Service Units
SECTONDD - Non-Dasd I/O Using or Delay
SECTOPIE - Page-Ins - Expanded
SECTOPIN - Page-Ins - Total
SECTOPRS - Page Residency - CS+ES
SECTORCP - Transactions Completed
SECTOSAM - Total Delay Samples
SECTOSPA - Shared Page-Ins from Auxiliary
SECTOSPE - Shared Page-Ins from ES
SECTOSRB - SRB CPU Service Units
SECTOSRS - Page Residency - Shared Pages
SECTOSRV - Total Service Units
SECTOSWC - Swaps
SECTOUSE - Total WLM Usings
SECTOZAP - CPU Service Units on zAAP
SECTZAPC - zAAP Eligible CPU Service Units on CP
SECTZAPCT - zAAP Eligible CPU Time on CP
SECTZAPTM - CPU Time on zAAP (Normalized)
SECTZAPUT - Un-normalized zAAP CPU Time
SECTOSUP - CPU Service Units on zIIP
SECTSUPC - zIIP Eligible CPU Service Units on CP
SECTSUPCT - zIIP Eligible CPU Time on CP
SECTSUPTM - CPU Time on zIIP (Normalized)
SECTSUPUT - Un-normalized zIIP CPU Time
WLMSCCPU - Service Coefficient - CPU
WLMSCSRB - Service Coefficient - SRB
WLMZAPNF - zAAP Service Time Normalization Factor
WLMSUPNF - zIIP Service Time Normalization Factor

SHR_CPS - Number of Shared CPs Defined in CEC

3. Time format data elements

ALLSRBTM - Total SRB CPU Time for MVS system
ALLTCBTM - Total TCB CPU Time for MVS system
ALLTCPTM - Total SRB & TCB Time for MVS system
CPUDUR - Recording Interval Time (HARCPU)

CPUT00TM - Total CP Processors Online Time
 CPUTOUTM - Total Processor Up Time
 CPUTOBTM - Total Time CP Processors Busy
 CPUTODTM - Total CP Processors Dispatch Time
 CPUTOVBT - Total CP Processors PR/SM MVS Busy Time
 CPUZPBTM - Total Time zAAP Processors Busy
 CPUZPDTM - Total Time zAAP Processors Dispatch Time
 CPUZPOTM - Total zAAP Processors Online Time
 CPUZPUTM - Total zAAP Processors Up Time
 CPUSBTM - Total Time zIIP Processors Busy
 CPUSUDTM - Total Time zIIP Processors Dispatch Time
 CPUSUOTM - Total zIIP Processors Online Time
 CPUSUUTM - Total zIIP Processors Up Time
 PRSMTEdT - Total Dispatch Time - Effective

DURATION - Recording Interval Time
 SECAVQUE - Avg Transaction Queue Time
 SECAVTET - Avg Transaction Elapse Time
 SECAVXET - Avg Transaction Execution Time
 SECCPUTM - TCB CPU Time
 SECDCTM - DASD I/O Connect Time
 SECDDTM - DASD I/O Disconnect Time
 SECDWTM - DASD I/O Wait Time
 SECSRBTM - SRB CPU Time
 SECTCPTM - Total TCB and SRB CPU Time
 SECTDSTM - Total DASD Service Time
 SECTOHST - Hiperspace Service Time
 SECTOIIT - I/O Interrupt Time
 SECTOQUE - Total Transaction Queue Time
 SECTORCT - Region Control Task Time
 SECTOTAT - Transaction Active Time
 SECTOTET - Transaction Elapsed Time
 SECTOTRR - In-Storage Residency Time
 SECTOXET - Transaction Execution Time
 SECZAPCT - zAAP Eligible CPU Time on CP
 SECZAPTM - CPU Time on zAAP (Normalized)
 SECZAPUT - Un-normalized zAAP CPU Time
 SECSUPCT - zIIP Eligible CPU Time on CP
 SECSUPTM - CPU Time on zIIP (Normalized)
 SECSUPUT - Un-normalized zIIP CPU Time

TOTCPUTM - Total CPU Time attributed to Workload

4. Percentage format data elements

CPUPCBSY - Pct CP Processors Busy (unadjusted)
 CPUPCBZP - Pct zAAP Processors Busy (unadjusted)
 CPUPCBSU - Pct zIIP Processors Busy (unadjusted)

PCATRCPU - Percent of CPU attributable to TCB and SRB time by all workloads, (i.e., the Capture Ratio)
PCTOTSRB - Percentage of total SRB time recorded by RMF, attributed to this workload
PCTOTTCB - Percentage of total TCB time recorded by RMF, attributed to this workload
PCTOTTCP - Percentage of total SRB & TCB time recorded by RMF, attributed to this workload

C. Additional data elements if CPU Time Adjustment is specified:

1. Character format data element

BASECPU - Base CPU Model

2. Numeric format data elements

ADJPRC - Base Adjustment Number of CPUs

ADJSERV - Base Adjustment Service Units/Second

WLMSUCPU - SU:CPU-Sec. Conv. Factor

3. Time format data element

WKLDASRB - Adjusted Workload SRB Time

WKLDATCB - Adjusted Workload TCB Time

WKLDACPU - Adjusted Workload SRB & TCB Time

WKLDATCP - Adjusted Total CPU Time

Many of the data elements used in the WLM Workload Planning Sample Application database file have the same meaning and derivation as their identically named counterparts in the CA MICS database. For more information on the derivation of these elements, see CA MICS Hardware and SCP Analyzer Guide.

The data elements that are derived from the CA MICS database and used by the WLM Workload Planning Sample Application are calculated as follows:

CAPAPU: The name of the workload. You specify this in either the Workload Mapping Table or the CAPAPU exit.

ALLSRBTM: The total SRB time consumed by all workloads on a system during the interval. This is the sum of SECSRBTM across the workloads.

ALLTCBTM: The total TCB time consumed by all workloads on a system during the interval. This is the sum of SECCPUTM across the workloads.

ALLTCPTM: The total SRB & TCB time consumed by all workloads on a system during the interval. This is the sum of SECTCPTM across the workloads.

PCATRCPU: The percent of SRB & TCB time attributable to workloads. This represents the sum of all SRB & TCB time measured by RMF for the represented workloads as a percentage of the hardware CPU busy time as measured by RMF. It is calculated using the formula:

$$\text{PCATRCPU} = 100 * \text{ALLTCPTM} / \text{CPUTOBTM}$$

This element should have values between 0 and 100.

PCTOTSRB: The percentage of total SRB time consumed by a workload. It is calculated using the formula:

$$\text{PCTOTSRB} = 100 * (\text{SECSRBTM} / \text{ALLSRBTM})$$

This element should have values between 0 and 100.

PCTOTTCB: The percentage of total TCB time consumed by a workload. It is calculated using the formula:

$$\text{PCTOTTCB} = 100 * (\text{SECCPUTM} / \text{ALLTCBTM})$$

This element should have values between 0 and 100.

PCTOTTCP: The percentage of total SRB & TCP time consumed by a workload. It is calculated using the formula:

$$\text{PCTOTTCP} = 100 * (\text{SECTCPTM} / \text{ALLTCPTM})$$

This element should have values between 0 and 100.

TOTCPU: This is the sum of all attributed CPU times as recorded by RMF. It is calculated using the formula:

$$\text{TOTCPU} = \text{SUM}(\text{SECSRB}, \text{SECCPUTM}, \text{SECTOIIT}, \text{SECTORCT}, \text{SECTOHST})$$

The additional data elements that are present if CPU time adjustment is specified are derived as follows:

BASECPU: The value that was specified for the Base CPU model parameter on the CPU Time Adjustment Parameters screen (Figure 2-2.07).

WKLDACPU: The adjusted workload SRB & TCB time. The derivation of this value has two parts:

- o The SRM service unit to CPU time conversion factor **WLMSCPU**, as recorded by RMF.
- o The value that you specify for the Base CPU service unit coefficient parameter is used in the formula:

$$\text{WKLDACPU} = \text{WLMSCPU} * \text{SECTCPTM} / \text{your value}$$

For example, if you specify a base CPU service unit coefficient of 1351.68, then the following calculation is performed:

$$\text{WKLDACPU} = \text{WLMSCPU} * \text{SECTCPTM} / 1351.68$$

- WKLDASRB: The adjusted workload SRB time. The derivation of this value has two parts:
- o The SRM service unit to CPU time conversion factor WLMSUCPU, as recorded by RMF.
 - o The value that you specify for the Base CPU service unit coefficient parameter is used in the formula:

$$\text{WKLDASRB} = \text{WLMSUCPU} * \text{SECSRBTM} / \text{your value}$$

For example, if you specify a base CPU service unit coefficient of 1351.68, then the following calculation is performed:

$$\text{WKLDASRB} = \text{WLMSUCPU} * \text{SECSRBTM} / 1351.68$$

- WKLDATCB: The adjusted workload TCB time. The derivation of this value has two parts:
- o The SRM service unit to CPU time conversion factor WLMSUCPU, as recorded by RMF.
 - o The value that you specify for the Base CPU service unit coefficient parameter is used in the formula:

$$\text{WKLDATCB} = \text{WLMSUCPU} * \text{SECCPUTM} / \text{your value}$$

For example, if you specify a base CPU service unit coefficient of 1351.68, then the following calculation is performed:

$$\text{WKLDATCB} = \text{WLMSUCPU} * \text{SECCPUTM} / 1351.68$$

WKLDATCP: The adjusted workload total CPU time. The derivation of this value has two parts:

- o The SRM service unit to CPU time conversion factor **WLMSUCPU**, as recorded by RMF.
- o The value that you specify for the Base CPU service unit coefficient parameter is used in the formula:

$$\text{WKLDATCP} = \text{WLMSUCPU} * \text{TOTCPU} / \text{your value}$$

For example, if you specify a base CPU service unit coefficient of 1351.68, then the following calculation is performed:

$$\text{WKLDATCP} = \text{WLMSUCPU} * \text{TOTCPU} / 1351.68$$

2.2.4 WLM Workload Planning Example

You can use the WLM Workload Planning Sample Application to build a Capacity Planning database file to track workload activity and resource consumption as reported by the CA MICS Hardware and SCP Analyzer.

Figures 2-2.08 through 2-2.12 show sample parameter specifications for a WLM Workload Planning Sample Application.

```

/----- WLM Workload Planning -----\
|Command ==>
|Enter a ? in any data entry field for more information on valid values.
|Composing CAP DB File def.:  GM1 - WLM Workload Planning Sample Application
|
|Summarization Level
|      Primary      ==> MONTHS  (DAYS/WEEKS/MONTHS)
|      Secondary    ==> ZONE    (ZONE/HOUR/DAYNAME)
|
|Input File Time-Span  ==> DAYS   (DETAIL/DAYS/WEEKS/MONTHS)
|
|Amount of data retained ==> 60   (1-9999 months)
|Age of oldest forecast  ==> 24   (1-9999 months)
|History cycles          ==> 3    (1-99 cycles)
|Forecast cycles         ==> 3    (1-99 cycles)
|
|Data base ID (s)       ==> P
|
|Specify extended options ==> Y (Y/N)
|
\-----/

```

Figure 2-2.08. WLM Workload Planning Screen - Example

In Figure 2-2.08, the primary summarization level parameter specifies that the application is to be summarized to using the MONTHS timespan. The Summary level parameter specifies that the data will be summarized in the ZONE level. The DETAIL timespan will be used as input. The P specified for the Database parameter indicates that the CA MICS WLMSEC and WLMSEM files are located in a database unit that carries an identifier of P.

In this example neither SYSIDs nor zones are aggregated into composites.

The information specified for the Maximum amount of data available for analysis and the Age of oldest forecast parameters indicates that historical data up to 60 months old and forecasts up to 24 months old are available for analysis. The History and Forecast parameters specify that three cycles of data will be retained.

The extended options should be requested to specify the Workload Mapping Table or CAPAPU exit at a minimum.

```

/----- Extended Options Menu -----\
|Command ==>
|
|Inquiry Step:  WLM Workload Planning Sample Application
|
|
|Combine all SYSIDs into a composite ==> N      (Y/N)
|  Composite SYSID                    ==> ____
|
|Combine all zones into a composite ==> N      (Y/N)
|  Composite zone                     ==> _____
|
|Specify:
|  Common data selection parameters ==> Y      (Y/N/R)
|  Global data selection exit        ==> N      (Y/N/R)
|  Workload mapping definition       ==> Y (Y/N/R)
|  Workload mapping (CAPAPU) exit    ==> N (Y/N/R)
|  CPU time adjustment parameters   ==> Y      (Y/N/R)
|  Notes report enabled              ==> YES    (YES/NO)
|  Contents reports enabled          ==> YES    (YES/NO)
|
|
|-----\

```

Figure 2-2.09. Extended Options Menu - Example

On the Extended Options Menu shown in Figure 2-2.09, the Specify common data selection, Specify CPU time adjustment, and Specify workload mapping definition options are selected.

```
/----- Common Data Selection -----\
|Command ==>
|
|Inquiry Step: WLM Workload Planning Sample Application
|
|SYSID ==> SYS1 SYS2 SYS3 ____ _ _ _ _ 
|          ____ _ _ _ _ 
|
|Zone   ==> 1 2 3 4 _ _ _ _ _             (1 - 9)
|
|Hour   ==> __ - __ - __ - __ - __ - __ - __           (0 - 23)
|
|Dayname==> _____
|
|_____
```

Figure 2-2.10. Common Data Selection Screen - Example

In Figure 2-2.10, the SYSIDs SYS1, SYS2 and SYS3 are selected for inclusion in the generated database files. Zones 1, 2, 3, and 4 will also be included. Data for all hours will be generated.

[illegible]

Figure 2-2.11. CPU Time Adjustment Parameters Screen - Example

Figure 2-2.11 shows that an IBM 3081-KX processor is chosen as the base CPU model for this example. The number of internal processors and the service-unit-to-CPU-second factor are determined from vendor information on the processor.

```
/----- Workload Mapping Definition -----\
|Command ==>
|
|Inquiry step: WLM Workload Planning Sample Application
|
|Workload mapping table data set name
|    ==> MYWKLMAP.DATA
|
|Edit workload mapping table          ==> NO (YES/NO)
|
|
|
|
|
|
|
|
|
|
```

Figure 2-2.12. WLM Workload Mapping Definition Screen - Example

In Figure 2-2.12 a workload mapping table is built externally using the ISPF Editor. The data set name of the volume mapping table is specified as MYWKLMAP.DATA.

2.3 WLC Implementation Planning

The WLC Implementation Planning Sample Application provides a capacity planning database file that contains the necessary elements for tracking and forecasting LPAR resource usage in MSUs. Both actual MSU usage and the 4 hour rolling average metrics are maintained along with changes in the computing environment. Additional elements associated to the usage and definition of defined capacity limit, capacity on demand, and Defined Group Capacity are also available with the sample application.

The following sections cover the use, input specifications, and output of the sample application:

- 1 - Usage Guidelines
- 2 - Control Parameters
- 3 - Standard Output

2.3.1 Usage Guidelines

The WLC Implementation Planning Sample Application provides a capacity planning database file with the database elements you need for basic tracking, reporting, and forecasting of LPAR resource usage.

The information for this application is primarily derived from either the DETAIL, DAYS, WEEKS, or MONTHS timespan of the PR/SM LPAR Config/Activity (HARLPC) File.

The CPCMODID and CPUMOD elements, which identify the processor type of the Central Processor Complex in use are obtained from the CA MICS CPU Activity (HARCPU) File.

Additional measurement metrics such as Physical CPU Adjustment Factor (CPUSUSEC), Service Units Available to MVS Image (CPUSUAVL), and Accumulated Long-Term Average of CPU MSUs (CPUTOLAC) are included as well.

In addition to providing basic information about LPAR resource utilization, the WLC Implementation Planning Sample Application provides measurement data concerning overall CPC utilization and capacity that can be tracked and forecasted as well.

With the introduction of Defined Group Capacity users will need the ability to summarize and forecast the MSU usage and the MSU 4-hr rolling average within a given Defined Capacity Group. The user can achieve this summarization by building a WLC planning file using the data selection feature to process only SYSIDs which belong to the Defined Capacity Group along with the Composite SYSID feature to summarize the groups MSU consumption and 4-hour rolling average.

Note: To accurately summarize a Capacity Group's MSU consumption and 4-hour rolling average all SYSID's belonging to the Capacity Group must be present within the input data.

2.3.2 Control Parameters

Relatively few parameters are needed to implement the WLC Implementation Planning Sample Application. The control screens included in the online facility allow you to quickly enter the required information.

```
/-----WLC Implementation Planner-----\  
|Command ==>  
|Enter a ? in any data entry field for more information on valid values.  
|Composing CAP DB File def.: WLC - WLC Implementation Planning  
|  
|Summarization Level  
|      Primary      ==> MONTHS  (DAYS/WEEKS/MONTHS)  
|      Secondary    ==> ZONE    (ZONE/HOUR/DAYNAME)  
|  
|Input File Time-Span ==> DETAIL  (DETAIL/DAYS/WEEKS/MONTHS)  
|  
|Amount of data retained ==> 60    (1-9999 months)  
|Age of oldest forecast  ==> 24    (1-9999 months)  
|History cycles          ==> 3     (1-99 cycles)  
|Forecast cycles         ==> 3     (1-99 cycles)  
|  
|Data base ID (s)       ==> _____  
|  
|  
|Specify extended options ==> Y (Y/N)  
\-----/
```

Figure 2-3.01. WLC Planning Control Screen

Figure 2-3.01 shows the WLC Planning screen, which is the main control screen for the WLC Implementation Planning Sample Application. It contains the following fields:

Summarization level:	The output summarization corresponds to the combination of the primary and secondary summarization levels. These control how the output file is summarized in your capacity planning database.
Input file timespan:	The timespan (either DETAIL, DAYS, WEEKS or MONTHS) of the HARCPU and HARLPU Files that you want to use as input when creating the capacity planning database file.
Amount of data retained:	The number of days, weeks or months of historical data to be retained in the CAPxxx01 cycle of the capacity planning database file. Any historical observations older than what you specify for this parameter are dropped from the file during the file update.
Age of oldest forecast:	The number of days, weeks or months that you want to retain a forecast after it is generated. Any forecast older than what you specify for this parameter is dropped when a new forecast is generated and saved.
History cycle:	Number of cycles or versions of the historical data you want to retain. A new cycle is created when the file is updated and the existing cycles aged (for example, Cycle 01 is aged to Cycle 02 and Cycle 02 becomes Cycle 03).
Forecast cycle:	Number of cycles or versions of the forecast data you want to retain. A new cycle is created when you specify forecast aging in one of the workload forecasting routines and the existing cycles are aged (for example, Cycle 01 is aged to Cycle 02 and Cycle 02 becomes Cycle 03).
Database ID(s):	The database unit identifier(s) (up to fourteen units) that contain the HARCPU and HARLPU Files to be used with this Sample application. If you are uncertain in which database unit the HARCPU and HARLPU Files reside, type a question mark (?) in this field to display a message listing of the valid database unit identifiers for this Sample application.

Specify extended options: If you want extended options, enter Y (YES), and the Extended Options Menu (Figure 2-3.02) is displayed.

```

/----- Extended Options Menu -----\
|Command ==>
|
|Inquiry Step: WLC Implementation Planning Sample Application
|
|
|Combine all SYSIDs into a composite ==> N      (Y/N)
|  Composite SYSID                    ==> ----
|
|Combine all zones into a composite ==> N      (Y/N)
|  Composite zone                     ==> -----
|
|Specify:
|  Common data selection parameters ==> N      (Y/N/R)
|  Global data selection exit        ==> N      (Y/N/R)
|  CPU time adjustment parameters   ==> N      (Y/N/R)
|  Notes report enabled              ==> YES    (YES/NO)
|  Contents reports enabled          ==> YES    (YES/NO)
|
|
|
|
\-----/

```

Figure 2-3.02. WLC Planning Extended Options Menu

The extended options available with the WLC Implementation Planning Sample Application (Figure 2-3.02) are:

Combine all SYSIDs into a composite:	You can combine SYSIDs in the capacity planning database file. Enter Y (YES) if you want data for all of the SYSIDs to be summarized into a single observation with the composite SYSID value.
Combine all zones/hours/ daynames into a composite:	You can combine zones, hours, and daynames in the capacity planning database file. The prompt for this field changes to zones or hours or daynames depending on what you specify for the summary level parameter. Enter Y (YES) if you want all of the zones, hours, and daynames to be summarized into a single observation with the composite value.
Specify common data selection parameters:	Specify the SYSIDs and zones, and data hours or daynames to be used for data selection from the CA MICS input file(s). Enter Y (YES) to display the Common Data Selection screen (Figure 2-3.03). Enter R (RESET) to reset any specifications you previously made to SYSID and zone, hour or dayname.
Specify global data selection exit:	You can use your own SAS statements to select data from CA MICS file(s). Enter Y (YES) to display the SAS System Statements screen (Figure 2-3.04). Enter R (RESET) to reset any selections you previously made to the default of no user-specified data selection.
Specify CPU time adjustment parameters:	Specify the parameters to use for CPU time adjustment (normalization). Enter Y (YES) to display the CPU Time Adjustment Parameters screen (Figure 2-3.03). Enter R (RESET) to reset any selections you previously made to CPU time adjustment.
	This option is not available if you select the composite SYSID option on this screen.
Notes reports enabled:	If you have entered notes associated with the capacity planning database file definition, a Yes in this field prints the note as part of file updating.
Contents	A Yes in this field prints the contents of the

```
reports      capacity planning file as part of file updating.
enabled:
```

[illegible]

Figure 2-3.03. Common Data Selection Screen - WLC Planning

The Common Data Selection screen (Figure 2-3.03) lets you to define the following data selection criteria:

- SYSID:** The SYSIDs that you want to include in your capacity planning database file. You can specify a maximum of 15 one- to four-character SYSIDs. Leave this parameter blank to include data for all SYSIDs.
- Zone:** You can specify up to nine zones to include in the capacity planning database file. Leave this parameter blank to include data for all zones.
- Hour:** Specify up to five hours or hour ranges to include in the capacity planning database file. Selection of this parameter is valid only if you enter the DETAIL, DAYS, or WEEKS timespan on the WLC Planning screen (Figure 2-3.03). Leave this parameter blank to include data for all hours. Singular hour values must be entered to the left of the dash.
- Dayname:** You can enter up to five daynames to include in the capacity planning database file. Leave blank to include all days. Input the first unique characters and the application will fill in the rest. Selection of this parameter is only valid when the input timespan is DETAIL or DAYS.

```

/----- SAS System Statements ----- COLUMNS 001 072--\
|Command ==>                               Scroll ==> PAGE |
|
|SAS Statements for the SELECT Macro          |
|-----|
|***** ***** TOP OF DATA *****|
|=MSG> SPECIFY SAS STATEMENTS TO SUBSET THE CAPACITY PLANNING DATA BASE FILE. |
|=NOTE= %MACRO SELECT; <=== PLEASE DO NOT PLACE LINES BEFORE THIS LINE.      |
|000001                                     |
|000002                                     |
|000003                                     |
|000004                                     |
|000005                                     |
|000006                                     |
|000007                                     |
|000008                                     |
|000009                                     |
|000010                                     |
|=NOTE= %MEND SELECT; <=== PLEASE DO NOT PLACE LINES AFTER THIS LINE.        |
|***** ***** BOTTOM OF DATA *****|

```

Figure 2-3.04. SAS System Statements Screen - WLC Planning Global Data Selection

The SAS System Statements screen (Figure 2-3.04) lets you define your own data selection logic. You can code any valid SAS statements as described in the SAS manuals. The %MACRO SELECT and %MEND SELECT statements are already provided, and are field protected from modification.

Note: This is a PDF/EDIT screen. All PDF/EDIT commands are available.

```
/----- CPU Time Adjustment Parameters -----  
|Command ===>  
|  
|Inquiry step: IRD Implementation Planning Sample Application  
|  
|Base CPU model ==> 3081-KX_____  
|Number of processors in base CPU model ==> 2_ (1-16)  
|Base CPU service unit coefficient ==> 396.4_ (per processor)  
|  
|  
|  
|  
|  
|  
|
```

Figure 2-3.05. CPU Time Adjustment Parameters Screen

The information on the CPU Time Adjustment Parameters screen, shown in Figure 2-3.05, is required only if you choose CPU time normalization for the IRD Implementation Planning Sample Application.

Base CPU model: The 10-character identifier for the base processor. Any name is acceptable in this field. Do not use embedded blanks or equal signs (=) in this value. For example, if you wish to normalize your actual CPU time values to the equivalent of a 3081-KX, then code 3081-KX in this field.

Number of processors in base CPU model: The number of internal CPUs in the base processor. This value must be an integer (that is, do not use a decimal point), and it must be greater than 0 and less than or equal to 16. For example, if you specify a 3081-K as the base CPU model, then code a 2 in this field since a 3081-K has two internal CPUs.

Base CPU service unit coefficient: The service-units-to-CPU-seconds conversion factor for the base processor. The value must be a real number greater than 0. You can compute a relative performance rating by dividing this value by the corresponding service-unit-to-CPU-second factor for the processor represented by the SYSID (see Section 2.1.1 for an example of this calculation). You may want to use vendor-supplied values for the factor, or a factor based on your own evaluations. For example, if you specify a 3081-K as the base CPU model, then code 396.4 in this field.

2.3.3 Standard Output

The capacity planning database file generated by the WLC Implementation Planning Sample Application contains the following data elements:

- A. Data elements common to all CA MICS Capacity Planner database files:

SYSID
ZONE or HOUR or DAYNAME
YEAR
MONTH or WEEK or MNTHDAY
DATE

- B. Data elements unique to the WLC Implementation Planning Sample Application database file:

1. Character format data elements

CAPAPU - LPAR Group Name (internal)
CPUMOD - Processor Model (from RMF)
CPCMODID - CPC Model Identifier
MVSLEVEL - MVS Release
PRSMCEC - PR/SM LPAR Central Complex
PRSMCGNM - Capacity Group Name
RMFVER - RMF Version Number

2. Numeric format data elements

INTERVLS - Number of Recording Intervals
CPUADJCH - Intvls with CPU Adjustment Factor Chgd
CPUSUACH - Intvls with Service Units Available Chgd
CPUSUAVL - Service Units Available to MVS Image
CPUSUSEC - Physical CPU Adjustment Factor
CPUTOLAC - Accum. Long-Term Average of CPU MSUs
CPUTOMSU - Actual Number of Consumed MSUs
LPCTOMSU - Actual Number of Consumed MSUs
LP_TOMSU - Total LPAR Processor Busy in MSUs
LPCBPS - Processor Relative Share
LPCCAPCH - Intvls with Defined Capacity Limit Chgd
LPCCAPDF - Intvls with Defined Capacity Limit
LPCDEFLM - Partition Defined Capacity Limit
LPCDIAGS - Number of Diagnose Samples

LPCNOCPU - Number of Logical Processors
LPCTONSW - Samples when WLM Capped Partition
PRSMPLNM - PR/SM Partition Name
PRSMLPNO - PR/SM Partition Number
PRSMCGLI - Capacity Group Limit
PRSMNOPP - PR/SM No. of Phys. Processors Assigned
PRSMTDP - PR/SM Dedicated Logical Processors
PRSMTLPA - PR/SM Active Logical Partitions
PRSMTLPS - PR/SM Total Defined Logical Partitions
PRSMTTP - PR/SM Total Processors in Complex
PRSMTVP - PR/SM Shared Logical Processors
PRSMUSCT - PR/SM Intvls Using PR/SM Feature
CPUCGJTS - Capacity Group Join Time Stamp
CPUCGJCT - Intvls with Capacity Group Changed
CPUMXLAC - Maximum Long-Term Average of CPU MSUs
CPUMXMSU - Maximum Number of Consumed MSUs
LPCMXMSU - Maximum Number of Consumed MSUs

3. Date and Time format data elements

CPCAVLTM - Total Available Processor Time in CPC
CPCUSETM - Total Processor Time Used in CPC
DATE - Interval End Date
DURATION - Recording Interval Time
LPCTODTM - Total Logical Processor Dispatch Time
PRSMTDTM - PR/SM Total Dispatch Time

4. Percentage format data elements

CPCPCLP - Percent Busy of CPC for this LPAR
CPCPCUSE - Percent Busy of Total CPC
LPCPCNSW - Partition WLM Capping Percentage

All of the CA MICS standard data elements used in the IBM License Manager Planning Sample Application database file have the same meaning and derivation as their identically named counterparts in the CA MICS HARLPC and HARCPU files.

The data elements that are derived from the CA MICS database and used by the WLC Implementation Planning Sample Application are calculated as follows:

CAPAPU: The CAPAPU element is developed internally by the application and consists of a logical concatenation of the PRSMENV & PRSMLPNO data values. This is done to cause separate observations to be created for each LPAR within each Central Processing Complex. This value will also ensure that historical observations are summarized correctly whenever a change is made within the computing environment. For example, adding a processor with the "Capacity on Demand" feature will be reflected within the PRSMENV value and subsequently cause a new observation to be created within the CAP Planning file.

CPCUSETM: The total processor time that the CPC was either dedicated or used in shared mode. It is calculated by the formula:

$$\text{CPCUSETM} = (\text{PRSMTPD} * \text{DURATION}) / \text{INTERVLS} + \text{PRSMDDTM}$$

where DURATION is the interval duration obtained from the HARLPC file.

CPCAVLTM: The total available processor time across the CPC including both dedicated and shared processors. It is calculated using the formula:

$$\text{CPCAVLTM} = (\text{PRSMTPP} * \text{DURATION}) / \text{INTERVLS}$$

where DURATION is the interval duration obtained from the HARLPC file.

CPCPCUSE: The percent of the total CPC consumed by all work. It is calculated using the following formula:

$$\text{CPCPCUSE} = 100 * (\text{CPCUSETM} / \text{CPCAVLTM})$$

This element must have values between 0 and 100.

CPCPCLPC: The percent of the total CPC consumed by this LPAR. It is calculated using the following formula:

$$\text{CPCPCLPC} = 100 * (\text{LPCTODTM} / \text{CPCAVLTM})$$

This element must have values between 0 and 100.

2.4 IRD Implementation Planning

The IRD (Intelligent Resource Director) Implementation Planning Sample Application provides a capacity planning database file that contains the necessary elements for tracking and forecasting processor resource usage at both the LPAR Cluster and LPAR level. Additionally, LPAR weighting and weight adjustments can be tracked at the LPAR level as well as changes in the computing environment, such as the defined capacity limit, defined group capacity, and the use of capacity on demand.

The following sections cover the use, input specifications, and output of the sample application:

- 1 - Usage Guidelines
- 2 - Control Parameters
- 3 - Standard Output

2.4.1 Usage Guidelines

The IRD Implementation Planning Sample Application provides a capacity planning database file with the database elements you need for basic tracking, reporting, and forecasting of processor resource usage at both the LPAR Cluster and LPAR levels.

The information for this application is primarily derived from either the DETAIL, DAYS, WEEKS, or MONTHS timespan of the PR/SM LPAR Config/Activity (HARLPC) File.

The CPCMODID and CPUMOD elements, which identify the processor type of the Central Processor Complex in use, are obtained from the CA MICS CPU Activity (HARCPU) File.

Additional measurement metrics, such as Physical CPU Adjustment Factor (CPUSUSEC), Service Units Available to MVS Image (CPUSUAVL), and Accumulated Long-Term Average of CPU MSUs (CPUTOLAC) are included as well.

In addition to providing basic information about processor resource utilization, the IRD Implementation Planning Sample Application provides measurement data concerning overall CPC utilization and capacity that can be tracked and forecasted as well.

2.4.2 Control Parameters

Relatively few parameters are needed to implement the IRD Implementation Planning Sample Application. The control screens included in the online facility allow you to quickly enter the required information.

```

/-----IRD Implementation Planning-----\
|Command ==>|
|Enter a ? in any data entry field for more information on valid values.|
|Composing CAP DB File def.:  IRD - IRD Implementation Planning|
|
|Summarization Level|
|      Primary      ==> MONTHS  (DAYS/WEEKS/MONTHS)|
|      Secondary    ==> ZONE    (ZONE/HOUR/DAYNAME)|
|
|Input File Time-Span ==> DETAIL (DETAIL/DAYS/WEEKS/MONTHS)|
|
|Amount of data retained ==> 60   (1-9999 months)|
|Age of oldest forecast  ==> 24   (1-9999 months)|
|History cycles          ==> 3    (1-99 cycles)|
|Forecast cycles         ==> 3    (1-99 cycles)|
|
|Data base ID (s)       ==> _____|
|
|
|Specify extended options ==> Y (Y/N)|
\-----/

```

Figure 2-4.01. IRD Implementation Planning Control Screen

Figure 2-4.01 shows the IRD Planning screen, which is the main control screen for the IRD Implementation Planning Sample Application. It contains the following fields:

Summarization level:	The output summarization corresponds to the combination of the primary and secondary summarization levels. These control how the output file is summarized in your capacity planning database.
Input file timespan:	The timespan (either DETAIL, DAYS, WEEKS or MONTHS) of the HARCPU and HARLPU Files that you wish to use as input when creating the capacity planning database file.
Amount of data retained:	The number of days, weeks or months of historical data to be retained in the CAPxxx01 cycle of the Capacity Planner database file. Any historical observations older than what you specify for this parameter are dropped from the file during the file update.
Age of oldest forecast:	The number of days, weeks or months that you want to retain a forecast after it is generated. Any forecast older than what you specify for this parameter is dropped when a new forecast is generated and saved.
History cycle:	The number of cycles or versions of the historical data you want to retain. A new cycle is created when the file is updated and the existing cycles aged (for example, Cycle 01 is aged to Cycle 02 and Cycle 02 becomes Cycle 03).
Forecast cycle:	The number of cycles or versions of the forecast data you want to retain. A new cycle is created when you specify forecast aging in one of the workload forecasting routines and the existing cycles are aged (for example, Cycle 01 is aged to Cycle 02 and Cycle 02 becomes Cycle 03).
Database ID(s):	The database unit identifier(s) (up to fourteen units) that contain the HARCPU and HARLPU Files to be used with this sample application. If you are uncertain in which database unit the HARCPU and HARLPU Files reside, type a question mark (?) in this field to display a message listing of the valid database unit identifiers for this sample application.

Specify
extended
options:

If you want extended options, enter Y (YES), and the Extended Options Menu (Figure 2-4.02) is displayed.

```

/----- Extended Options Menu -----\
|Command ==>
|
|Inquiry Step:  IRD Implementation Planning
|
|
|Combine all SYSIDs into a composite ==> N      (Y/N)
|  Composite SYSID                    ==>  ____
|
|Combine all zones into a composite ==> N      (Y/N)
|  Composite zone                     ==>  _____
|
|Specify:
|  Common data selection parameters ==> N      (Y/N/R)
|  Global data selection exit       ==> N      (Y/N/R)
|  CPU time adjustment parameters  ==> N      (Y/N/R)
|  Notes report enabled             ==> YES    (YES/NO)
|  Contents reports enabled         ==> YES    (YES/NO)
|
|
|
|
\

```

Figure 2-4.02. IRD Planning Extended Options Menu

The extended options available with the IRD Implementation Planning Sample Application (Figure 2-4.02) are:

Combine all SYSIDs into a composite: You can combine SYSIDs in the capacity planning database file. Enter a Y (YES) if you wish data for all of the SYSIDs to be summarized into a single observation with the Composite SYSID value.

Combine all zones/hours/daynames into a composite: You can combine zones, hours and daynames in the capacity planning database file. The prompt for this field changes to zones, hours, or daynames depending on what you specify for the summary level parameter. Enter Y (YES) if you want all of the zones, hours, and daynames to be summarized into a single observation with the composite value.

Specify common data selection parameters: Specify the SYSIDs and zones, data hours or daynames to be used for data selection from the CA MICS input file(s). Enter Y (YES) to display the Common Data Selection screen (Figure 2-4.03). Enter R (RESET) to reset any selections you previously made to SYSID and zone, hour or daynames.

Specify global data selection exit: You can use your own SAS statements to select data from CA MICS file(s). Enter Y (YES) to display the SAS System Statements screen (Figure 2-4.04). Enter R (RESET) to reset any selection you previously made to the default of no user-specified data selection.

Specify CPU time adjustment parameters: Specify parameters to use for CPU time adjustment (normalization). Enter Y (YES) to display the CPU Time Adjustment Parameters screen (Figure 2-4.05). Enter R (RESET) to reset any selection you previously made to CPU time adjustment.

This option is not available if you specify the composite SYSID option.

Notes report enabled: If you have entered notes associated with the capacity Planning file definition, a Yes in this field prints the note as part of file updating.

Contents reports A Yes in this field prints the contents of the capacity planner database file as part of file

```
enabled:      updating.
```

```

/------ Common Data Selection -----\
|Command ==>
|
|Inquiry Step:   IRD Implementation Planning
|
|SYSID ==>      _ _ _ _ _ 
|                _ _ _ _ _ 
|
|Zone    ==>     - - - - -                               (1 - 9)
|
|Hour     ==>    __ - __ - __ - __ - __ - __ - __        (0 - 23)
|
|Dayname==>     _____
|
|
|
|
|
|
|------\

```

Figure 2-4.03. Common Data Selection Screen - IRD Planning

The Common Data Selection screen (Figure 2-4.03) lets you define the following data selection criteria:

- SYSID:** The SYSIDs that you want to include in your capacity planning database file. You can specify a maximum of 15 one- to four-character SYSIDs. Leave this parameter blank to include data for all SYSIDs.
- Zone:** You can specify up to nine zones to include in the capacity planning database file. Leave this parameter blank to include data for all zones.
- Hour:** You can specify up to five hours or hour ranges to be included in the capacity planning database file. Selection of this parameter is valid only if you specify the DETAIL, DAYS, or WEEKS timespan on the IRD Implementation Planning screen (Figure 2-4.01). Leave this parameter blank to include data for all hours. Singular hour values must be entered to the left of the dash.
- Dayname:** You can specify up to five daynames to be included in the capacity planning database file. Leave blank to include all days. Input the first unique characters and this application will fill in the rest. Selection of this parameter is only valid when the input timespan is DETAIL or DAYS.

```

/----- SAS System Statements ----- COLUMNS 001 072--\
|Command ==>                               Scroll ==> PAGE |
|
|SAS Statements for the SELECT Macro          |
|-----|
|***** ***** TOP OF DATA *****|
|==MSG> SPECIFY SAS STATEMENTS TO SUBSET THE CAPACITY PLANNING DATA BASE FILE. |
|=NOTE= %MACRO SELECT; <== PLEASE DO NOT PLACE LINES BEFORE THIS LINE.      |
|000001                                     |
|000002                                     |
|000003                                     |
|000004                                     |
|000005                                     |
|000006                                     |
|000007                                     |
|000008                                     |
|000009                                     |

```

```
|000010|
|=NOTE= %MEND SELECT; <=== PLEASE DO NOT PLACE LINES AFTER THIS LINE.|
|***** ***** BOTTOM OF DATA *****|
|\-----|
```

Figure 2-4.04. SAS System Statements Screen - IRD Planning Global Data Selection

The SAS System Statements screen (Figure 2-4.04) lets you define your own data selection logic. You can code any valid SAS statements as described in the SAS manuals. The %MACRO SELECT and %MEND SELECT statements are already provided and field protected from modification.

Note: This is a PDF/EDIT screen. All PDF/EDIT commands are available.

```
/----- CPU Time Adjustment Parameters -----  
|Command ==>  
|  
|Inquiry step: IRD Implementation Planning Sample Application  
|  
|Base CPU model ==> 3081-KX_____  
|Number of processors in base CPU model ==> 2_ (1-16)  
|Base CPU service unit coefficient ==> 396.4_ (per processor)  
|  
|  
|  
|  
|  
|  
|
```

Figure 2-4.05. CPU Time Adjustment Parameters Screen

The information on the CPU Time Adjustment Parameters screen (Figure 2-4.05) is required only if you choose CPU time normalization for the IRD Implementation Planning Sample Application.

Base CPU model: The 10-character identifier for the base processor. Any name is acceptable in this field. Do not use embedded blanks or equal signs (=) in this value. For example, if you want to normalize your actual CPU time values to the equivalent of a 3081-KX, enter 3081-KX in this field.

Number of processors in base CPU model: The number of internal CPUs in the base processor. This value must be an integer (that is, do not use a decimal point), and it must be greater than 0 and less than or equal to 16. For example, if you specify a 3081-K as the base CPU model, enter the number 2 in this field since a 3081-K has two internal CPUs.

Base CPU service unit coefficient: The service-units-to-CPU-seconds conversion factor for the base processor. The value must be a real number greater than 0. You can compute a relative performance rating by dividing this value by the corresponding service-unit-to-CPU-second factor for the processor represented by the SYSID (see Section 2.1.1 for an example of this calculation). You can use vendor-supplied values for the factor, or a factor based on your own evaluations. For example, if you specify a 3081-K as the base CPU model, enter 396.4 in this field.

2.4.3 Standard Output

The capacity planning database file generated by the IRD Implementation Planning Sample Application contains the following data elements:

- A. Data elements common to all CA MICS Capacity Planner database files:

SYSID
ZONE or HOUR or DAYNAME
YEAR
MONTH or WEEK or MNTHDAY
DATE

- B. Data elements unique to the IRD Implementation Sample Application database file:

1. Character format data elements

CAPAPU - Resource Summarization Group (internal)
CPUMOD - Processor Model (from RMF)
CPCMODID - CPC Model Identifier
LPCCLUST - Cluster Name
PRSMCGNM - Capacity Group Name
MVSLEVEL - MVS Release
PRSMCEC - PR/SM LPAR Central Complex
RMFVER - RMF Version Number

2. Numeric format data elements

INTERVLS - Number of Recording Intervals
CPUADJCH - Intvls with CPU Adjustment Factor Chgd
CPUSUACH - Intvls with Service Units Available Chgd
CPUSUAVL - Service Units Available to MVS Image
CPUSUSEC - Physical CPU Adjustment Factor
CPUTOLAC - Accum. Long-Term Average of CPU MSUs
CPUGJCT - Intvls with Capacity Group Changed
LP_TOMSU - Total LPAR Processor Busy in MSUs
LPCACS - Processor Actual Weighting
LPCBPS - Processor Relative Share
LPCCAPCH - Intvls with Defined Capacity Limit Chgd
LPCCAPDF - Intvls with Defined Capacity Limit

LPCDEFLM - Partition Defined Capacity Limit
LPCDIAGS - Number of Diagnose Samples
LPCNOCPU - Number of Logical Processors
LPCTONSA - Samples at Maximum Weighting
LPCTONSW - Samples when WLM Capped Partition
PRSMCGLI - Capacity Group Limit
PRSMLPNM - PR/SM Partition Name
PRSMLPNO - PR/SM Partition Number
PRSMNOPP - PR/SM No. of Phys. Processors Assigned
PRSMTDP - PR/SM Dedicated Logical Processors
PRSMTLPA - PR/SM Active Logical Partitions
PRSMTLPS - PR/SM Total Defined Logical Partitions
PRSMTTP - PR/SM Total Processors in Complex
PRSMTVP - PR/SM Shared Logical Processors
PRSMUSCT - PR/SM Intvls Using PR/SM Feature

3. Date and Time format data elements

CPCAVLTM - Total Available Processor Time in CPC
CPCUSETM - Total Processor Time Used in CPC
CPUCGJTS - Capacity Group Join Time Stamp
DATE - Interval End Date
DURATION - Recording Interval Time
LPCTODTM - Total Logical Processor Dispatch Time
PRSMTDTM - PR/SM Total Dispatch Time

4. Percentage format data elements

CPCPCLPR - Percent Busy of CPC for this LPAR
CPCPCUSE - Percent Busy of Total CPC
LPCPCNSA - Maximum Weighting Percentage
LPCPCNSW - Partition WLM Capping Percentage

All of the CA MICS standard data elements used in the IRD Implementation Planning Sample Application database file have the same meaning and derivation as their identically named counterparts in the CA MICS HARLPC and HARCPU files.

The data elements that are derived from the CA MICS database and used by the IRD Implementation Planning Sample Application are calculated as follows:

CAPAPU: The CAPAPU element is developed internally by the application and consists of a logical concatenation of the PRSMENV & LPCCLUST or PRSMENV, LPCCLUST & PRSMLPNO data values, depending on the client specified option. This is done to cause separate observations to be created for each LPAR within each Central Processing Complex. This value will also ensure that historical observations are summarized correctly whenever a change is made within the computing environment. For example, adding a processor with the "Capacity on Demand" feature will be reflected within the PRSMENV value and subsequently cause a new observation to be created within the CAP Planning file.

CPCUSETM: The total processor time that the CPC was either dedicated or used in shared mode. It is calculated by the formula:

$$\text{CPCUSETM} = (\text{PRSMTP} * \text{DURATION}) / \text{INTERVLS} + \text{PRSMDDTM}$$

where DURATION is the interval duration obtained from the HARLPC file.

CPCAVLTM: The total available processor time across the CPC including both dedicated and shared processors. It is calculated using the formula:

$$\text{CPCAVLTM} = (\text{PRSMTP} * \text{DURATION}) / \text{INTERVLS}$$

where DURATION is the interval duration obtained from the HARLPC file.

CPCPCUSE: The percent of the total CPC consumed by all work. It is calculated using the following formula:

$$\text{CPCPCUSE} = 100 * (\text{CPCUSETM} / \text{CPCAVLTM})$$

This element must have values between 0 and 100.

CPCPCLPC: The percent of the total CPC consumed by this LPAR. It is calculated using the following formula:

$$\text{CPCPCLPC} = 100 * (\text{LPCTODTM} / \text{CPCAVLTM})$$

This element must have values between 0 and 100.

2.5 XCF Resource Planning

The XCF Resource Planning Sample Application provides a capacity planning database file that contains the necessary elements for the tracking and reporting of cross-system communication that uses the XCF facility of MVS.

The following sections cover the use, input specifications, and output of the sample application:

- 1 - Usage Guidelines
- 2 - Control Parameters
- 3 - Standard Output

2.5.1 Usage Guidelines

The XCF Resource Planning Sample application provides a capacity planning database file with the database elements you need for basic tracking, reporting, and forecasting of XCF usage on your MVS systems. The information for this application is derived from either the DETAIL, DAYS, WEEKS, or MONTHS timespan of the CA MICS Cross-System Coupling Facility (SCPXCM) file.

The XCF Resource Planning Sample application permits your installation to track and forecast XCF signalling between MVS images at a high level of summarization (that is, at the SYSID and ZONE, HOUR or DAYNAME level) for XCF members. Because the data is derived from the SCPXCM file, this sample application provides information about individual XCF member signalling.

XCF WORKLOADS

The XCF Resource Planning Sample Application uses five CAPAPU elements in order to keep the granularity of the capacity database at the XCF member level. These five CAPAPU elements are defined below. In addition to these five keys, the file is also summarized by the SYSID that did the RMF recording.

- CAPAPU - SYSPLEX Identifier (SYSPLEX)
- CAPAPU2 - XCF System Name (XCFSYSNM)
- CAPAPU3 - XCF Group Name (XCMGROUP)
- CAPAPU4 - XCF Member Name (XCMEMBER)
- CAPAPU5 - Member Residency (XCMMBRES)

The XCF Resource Planning Sample Application allows you to modify the setting of these CAPAPU fields by using the CAPAPU exit. Using this exit is optional, and only needed if you want to modify the level of granularity of the XCF capacity planning database.

2.5.2 Control Parameters

Relatively few parameters are needed to implement the XCF Resource Planning Sample Application. The control screens included in the online facility allow you to quickly enter the required information.

```
/----- XCF Resource Planning -----\
|Command ==>
|Enter a ? in any data entry field for more information on valid values.
|Composing CAP DB File def.: XCF - XCF Resource Planning Sample Application
|
|Summarization Level
|      Primary      ==> MONTHS (DAYS/WEEKS/MONTHS)
|      Secondary    ==> ZONE   (ZONE/HOUR/DAYNAME)
|
|Input File Time-Span ==> DETAIL (DETAIL/DAYS/WEEKS/MONTHS)
|
|Amount of data retained ==> 60 (1-9999 months)
|Age of oldest forecast ==> 24 (1-9999 months)
|History cycles ==> 3 (1-99 cycles)
|Forecast cycles ==> 3 (1-99 cycles)
|
|Data base ID (s) ==> _____
|
|
|Specify extended options ==> Y (Y/N)
\-----/
```

Figure 2-5.01. XCF Planning Control Screen

Figure 2-5.01 shows the XCF Planning screen, which is the main control screen for the XCF Resource Planning Sample Application. It contains the following fields:

Summariza- The output summarization corresponds to the
tion level: combination of the primary and secondary
summarization levels. These control how the
output file is summarized in your capacity
planning database.

Input file The timespan (either DETAIL, DAYS, WEEKS or
timespan: MONTHS) of the SCPXCM file that you want to use
as input when creating the capacity planning
database file.

Amount of data retained:	The number of days, weeks or months of historical data to be retained in the CAPxxx01 cycle of the capacity planning database file. Any historical observations older than what you specify for this parameter are dropped from the file during the file update.
Age of oldest forecast:	The number of days, weeks or months that you want to retain a forecast after it is generated. Any forecast older than what you specify for this parameter is dropped when a new forecast is generated and saved.
History cycle:	The number of cycles or versions of the historical data you want to retain. A new cycle is created when the file is updated and the existing cycles aged (for example, Cycle 01 is aged to Cycle 02 and Cycle 02 becomes Cycle 03).
Forecast cycle:	The number of cycles or versions of the forecast data you want to retain. A new cycle is created when you specify forecast aging in one of the workload forecasting routines and the existing cycles are aged (for example, Cycle 01 is aged to Cycle 02 and Cycle 02 becomes Cycle 03).
Database ID(s):	The database unit identifier(s) (up to fourteen units) that contain the SCPXCM file to be used with this sample application. If you are not sure where these files reside, type a question mark (?) in this field to display a listing of the valid database unit identifiers for this sample application.

Specify extended options:	If you want extended options, enter Y (YES), and the Extended Options Menu (Figure 2-5.02) is displayed.

```

----- Extended Options Menu -----
|Command ==>
|
|Inquiry Step: XCF Resource Planning Sample Application
|
|Combine all SYSIDs into a composite ==> N          (Y/N)
|   Composite SYSID                  ==> ____
|
|Combine all zones into a composite ==> N          (Y/N)
|   Composite zone                   ==> _____
|
|Specify:
|   Common data selection parameters ==> N          (Y/N/R)
|   Global data selection exit       ==> N          (Y/N/R)
|   CAPAPU derivation SAS statements ==> N          (Y/N/R)
|   Notes report enabled             ==> YES        (YES/NO)
|   Contents reports enabled         ==> YES        (YES/NO)
|
|
|
|
|

```

Figure 2-5.02. XCF Planning Extended Options Menu

The extended options available with the XCF Resource Planning Sample Application (Figure 2-5.02) are:

Combine all SYSIDs into a composite: You can combine SYSIDs in the capacity planning database file. Enter Y (YES) if you want data for all of the SYSIDs to be summarized into a single observation with the Composite SYSID value.

Combine all zones/hours/daynames into a composite:	You can combine zones, hours and daynames in the capacity planning database file. The prompt for this field changes to zones or hours or daynames depending on what you specify for the summary level parameter. Enter a Y (YES) if you want all of the zones, hours, and daynames to be summarized into a single observation with the Composite value.
Specify common data selection parameters:	Specify the SYSIDs and zones, hours or daynames to be used for data selection from the CA MICS input file(s). Enter Y (YES) to display the Common Data Selection screen (Figure 2-5.03.) Enter R (RESET) to reset any selections you previously made to SYSID and zone, hour or dayname.
Specify global data selection exit:	You can specify your own SAS statements to select data from the CA MICS input file(s). Enter Y (YES) to display the SAS System Statements screen (Figure 2-5.04). Enter R (RESET) to reset any selections you previously made to the default of no user-specified data selection.
CAPAPI derivation SAS statements:	You can use your own SAS statements to override the setting of the CAPAPI elements. Enter Y (YES) to display the SAS System Statements screen.
Notes report enabled:	If you have entered notes associated with the capacity planning file definition, a Yes in this field prints the note as part of file updating.

Contents reports enabled:	A Yes in this field prints the contents of the capacity planning database file as part of file updating.

```

|----- Common Data Selection -----|
|Command ==>|
|Inquiry Step: XCF Resource Planning Sample Application|
|SYSID ==> |-----|
|          |-----|
|Zone   ==> |-----| (1 - 9)
|Hour   ==> |_ _ ^ _ _ ^ _ _ ^ _ _ ^ _ _ ^ _ _| (0 - 23)
|Dayname==> |-----|

```

Figure 2-5.03. Common Data Selection Screen - XCF Planning

The Common Data Selection screen (Figure 2-5.03) allows you to define the following data selection criteria:

SYSID: The SYSIDs that you want to include in your capacity planning database file. You can specify a maximum of 15 one- to four-character SYSIDs. Leave this parameter blank to include data for all SYSIDs.

Zone: Select up to nine zones to include in the capacity planning database file. Leave this parameter blank to include data for all zones.

- Hour:** Select five hours, or hour ranges to be included in the capacity planning database file. Selection of this parameter is valid only if you specify the DETAIL, DAYS, or WEEKS timespan on the XCF Resource Planning screen (Figure 2-5.01). Leave this parameter blank to include data for all hours.
- Dayname:** Select up to five daynames to be included in the capacity planning database file. Leave blank to include all days. Input the first unique characters and this application will fill in the rest. Selection of this parameter is only valid when the input timespan is DETAIL or DAYS.

```

/----- SAS System Statements ----- COLUMNS 001 072--\
|Command ==>                               Scroll ==> PAGE |
|
|SAS Statements for the SELECT Macro
|-----|
|***** ***** TOP OF DATA *****|
|=MSG> SPECIFY SAS STATEMENTS TO SUBSET THE CAPACITY PLANNING DATA BASE FILE. |
|=NOTE= %MACRO SELECT; <=== PLEASE DO NOT PLACE LINES BEFORE THIS LINE.      |
|000001
|000002
|000003
|000004
|000005
|000006
|000007
|000008
|000009
|000010
|=NOTE= %MEND SELECT; <=== PLEASE DO NOT PLACE LINES AFTER THIS LINE.        |
|***** ***** BOTTOM OF DATA *****|
|
|-----/

```

Figure 2-5.04. SAS System Statements Screen - XCF Planning Global Data Selection

The SAS System Statements screen (Figure 2-5.04) lets you define your own data selection logic. You can code any valid SAS statements as described in the SAS manuals. The %MACRO SELECT and %MEND SELECT statements are already provided and field protected from modification.

Note: This is a PDF/EDIT screen. All PDF/EDIT commands are available.

2.5.3 Standard Output

The capacity planning database file that is generated by the XCF Resource Planning Sample Application contains the following data elements:

- A. Data elements common to all CA MICS Capacity Planner database files:

SYSID
ZONE or HOUR or DAYNAME
YEAR
MONTH or WEEK or MNTHDAY
DATE

- B. Data elements unique to the XCF Resource Planning Sample Application database file:

1. Character format data elements

CAPAPU - SYSPLEX Identifier
CAPAPU2 - XCF System Name
CAPAPU3 - XCF Group Name
CAPAPU4 - XCF Member Name
CAPAPU5 - XCF Member Residency

2. Numeric format data elements

XCMTORCV - Signals Received
XCMTOSNT - Signals Sent
TOTSGNLS - Total signals sent & received by that instance of the MEMBER
ALLSGNLS - Total signals sent & received for the SYSPLEX, as recorded by that RMF

3. Percentage format data elements

PCTSGNLS - Percentage of all signals that this instance of the MEMBER sent and received

Some of the data elements used in the XCF Resource Planning Sample Application database file have the same meaning as their identically named counterparts in the CA MICS database.

The data elements that are derived from the CA MICS database and used by the XCF Resource Planning Sample Application are calculated as follows:

TOTSGNLS: The total number of signals sent & recieved by this instance of the MEMBER. Members typically reside on multiple MVS images within the Enhanced SYSPLEX. Typically, there will be one instance of the MEMBER for each MVS image that is part of the Enhanced SYSPLEX. The formula used to derive this element is:

$$\text{TOTSGNLS} = \text{SUM}(\text{XCMTORCV}, \text{XCMTOSNT})$$

ALLSGNLS: The total number of signals sent & recieved by all members within the Enhanced SYSPLEX. This element is derived by summarizing TOTSGNLS by recording SYSID.

PCTSGNLS: The amount of signalling done by this instance of the member as a percentage of all messaging being done in the Enhanced SYSPLEX.

$$\text{PCTSGNLS} = 100 * \text{TOTSGNLS} / \text{ALLSGNLS}$$

2.6 Coupling Facility Planning

The Coupling Facility Planning Sample Application provides a capacity planning database file with the database elements necessary for tracking, reporting, and forecasting usage and activity of the Enhanced SYSPLEX Coupling Facility.

The following sections cover the use, input specifications, and output of the sample application:

- 1 - Usage Guidelines
- 2 - Control Parameters
- 3 - Standard Output

2.6.1 Usage Guidelines

The Coupling Facility Planning Sample Application provides a capacity planning database file with the database elements necessary for tracking, reporting, and forecasting the usage and activity associated with the Enhanced SYSPLEX Coupling Facility. The information for this sample application is derived from either the DETAIL, DAYS, WEEKS, or MONTHS timespan of the CA MICS CF/XES Configuration and Activity file (HARCFP).

The Coupling Facility Planning Sample Application permits your installation to track and forecast Coupling Facility usage and activity at a high level of summarization (that is, at the SYSID and ZONE, HOUR or DAYNAME level) for Coupling Facilities. Because the data is derived from the HARCFP file, this sample application provides information about each unique Coupling Facility from each recording RMF's point of view.

Coupling Facility Groups

Three CAPAPU elements, defined below, are used to keep the granularity of the capacity database at the Coupling Facility level. In addition to these three keys, the file is also summarized by the SYSID that did the RMF recording.

- CAPAPU - SYSPLEX Identifier (SYSPLEX)
- CAPAPU2 - Coupling Facility Name (CFYCFNM)
- CAPAPU3 - System using CF (CFYSYSNM)

The Coupling Facility Planning Sample Application allows you to modify the setting of these CAPAPU fields by using the CAPAPU exit. Using this exit is optional, and only needed if you want to modify the level of granularity of the Coupling Facility capacity planning database.

2.6.2 Control Parameters

Relatively few parameters are needed to implement the Coupling Facility Planning Sample Application. The control screens included in the online facility allow you to quickly enter the required information.

```
/----- Coupling Facility Planning -----\
|Command ==>
|Enter a ? in any data entry field for more information on valid values.
|Composing CAP DB File def.: CF1 - Coupling Facility Planning
|
|Summarization Level
|      Primary      ==> MONTHS  (DAYS/WEEKS/MONTHS)
|      Secondary    ==> ZONE    (ZONE/HOUR/DAYNAME)
|
|Input File Time-Span ==> DETAIL  (DETAIL/DAYS/WEEKS/MONTHS)
|
|Amount of data retained ==> 60    (1-9999 months)
|Age of oldest forecast ==> 24    (1-9999 months)
|History cycles ==> 3    (1-99 cycles)
|Forecast cycles ==> 3    (1-99 cycles)
|
|Data base ID (s) ==> _____
|
|
|Specify extended options ==> Y (Y/N)
\-----/
```

Figure 2-6.01. Coupling Facility Planning Screen

The Coupling Facility Planning screen (Figure 2-6.01), the main control screen for the Coupling Facility Sample Application, contains the following fields:

Summarization level:	The output summarization corresponds to the combination of the primary and secondary summarization levels. These control how the output file is summarized in your Capacity Planning database.
Input file timespan:	The timespan (either DETAIL, DAYS, WEEKS or MONTHS) of the HARCFP file that you want to use as input when creating the capacity planning database file.

Amount of data retained:	The number of days, weeks or months of historical data to be retained in the CAPxxx01 cycle of the capacity planning database file. Any historical observations older than what you specify for this parameter are dropped from the file during the file update.
Age of oldest forecast:	The number of days, weeks or months that you want to retain a forecast after it is generated. Any forecast older than what you specify for this parameter is dropped when a new forecast is generated and saved.
History cycle:	Number of cycles or versions of the historical data you want to retain. A new cycle is created when the file is updated and the existing cycles aged (for example, Cycle 01 is aged to Cycle 02 and Cycle 02 becomes Cycle 03).
Forecast cycle:	Number of cycles or versions of the forecast data you want to retain. A new cycle is created when you specify forecast aging in one of the workload forecasting routines and the existing cycles are aged (for example, Cycle 01 is aged to Cycle 02 and Cycle 02 becomes Cycle 03).
Database ID(s):	The database unit identifier(s) (up to fourteen units) that contain the HARCFP file to be used with this sample application. If you are not sure where these files reside, type a question mark (?) in this field to display a listing of the valid database unit identifiers for this sample application.

Specify extended options:	If you want extended options, enter Y (YES), and the Extended Options Menu (Figure 2-6.02) is displayed.

```

----- Extended Options Menu -----
|Command ==>
|
|Inquiry Step: Coupling Facility Planning
|
|Combine all SYSIDs into a composite ==> N          (Y/N)
|   Composite SYSID                  ==> ____
|
|Combine all zones into a composite ==> N          (Y/N)
|   Composite zone                   ==> _____
|
|Specify:
|   Common data selection parameters ==> N          (Y/N/R)
|   Global data selection exit       ==> N          (Y/N/R)
|   CAPAPU derivation SAS statements ==> N          (Y/N/R)
|   Notes report enabled             ==> YES        (YES/NO)
|   Contents reports enabled         ==> YES        (YES/NO)
|
|
|
|
|

```

Figure 2-6.02. Coupling Facility Planning Extended Options Menu

The extended options available with the Coupling Facility Planning Sample Application (Figure 2-6.02) are:

```
Combine all You can combine SYSIDs in the capacity planning
SYSIDs into database file. Enter Y (YES) if you want data
a composite: for all of the SYSIDs to be summarized into a
single observation with the Composite SYSID
value.
```

Combine all zones/hours/daynames into a composite:	You can combine zones, hours, and daynames in the capacity planning database file. The prompt for this field changes to zones, hours or daynames depending on what you specify for the summary level parameter. Enter Y (YES) if you want all of the zones, hours, and daynames to be summarized into a single observation with the composite value.
Specify common data selection parameters:	Specify the SYSIDs and zones, hours or daynames to be used for data selection from the CA MICS input file(s). Enter Y (YES) to display the Common Data Selection screen (Figure 2-6.03.) Enter R (RESET) to reset any selections you previously made to SYSID and zone, hour or daynames.
Specify global data selection exit:	You can use your own SAS statements to select data from the CA MICS input file(s). Enter Y (YES) to display the SAS System Statements screen (Figure 2-6.04). Enter R (RESET) to reset any selections you previously made to the default of no user-specified data selection.
CAPAPU derivation SAS statements:	You can use your own SAS statements to override the setting of the CAPAPU elements. Enter Y (YES) to display the SAS System Statements screen.
Notes report enabled:	If you have entered notes associated with the capacity planning database file definition, a Yes in this field prints the note as part of file updating.

Contents reports enabled: A Yes in this field prints the contents of the capacity planning file as part of file updating.

```

/----- Common Data Selection -----\
|Command ==>
|
|Inquiry Step:  Coupling Facility Planning
|
|SYSID ==>  _____
|           _____
|
|Zone  ==>  _ _ _ _ _ (1 - 9)
|
|Hour  ==>  _ _ - _ _ - _ _ - _ _ - _ _ - _ _ (0 - 23)
|
|Dayname==>  _____
|
|
|
|
\-----\

```

Figure 2-6.03. Common Data Selection Screen - Coupling Facility Planning

The Common Data Selection screen (Figure 2-6.03) allows you define the following data selection criteria:

SYSID: The SYSIDs that you want to include in your capacity planning database file. You can specify a maximum of 15 one- to four-character SYSIDs. Leave this parameter blank to include data for all SYSIDs.

Zone: You can specify up to nine zones to be included in the capacity planning database file. Leave this parameter blank to include data for all zones.

Hour: Specify the five hours or hour ranges to be included in the capacity planning database file. Selection of this parameter is valid only if you specify the DETAIL, DAYS, or WEEKS timespan on the Coupling Facility Planning screen (Figure 2-6.01). Leave this parameter blank to include data for all hours.

Dayname: You can enter up to five daynames to be included in the capacity planning database file. Leave blank to include all days. Input the first unique characters and this application will fill in the rest. Selection of this parameter is only valid when the input timespan is DETAIL or DAYS.

```

/----- SAS System Statements ----- COLUMNS 001 072--\
|Command ==>                               Scroll ==> PAGE |
|
|SAS Statements for the SELECT Macro                |
|-----|
|***** ***** TOP OF DATA *****|
|=MSG> SPECIFY SAS STATEMENTS TO SUBSET THE CAPACITY PLANNING DATA BASE FILE. |
|=NOTE= %MACRO SELECT; <=== PLEASE DO NOT PLACE LINES BEFORE THIS LINE.      |
|000001                                           |
|000002                                           |
|000003                                           |
|000004                                           |
|000005                                           |
|000006                                           |
|000007                                           |
|000008                                           |
|000009                                           |
|000010                                           |
|=NOTE= %MEND SELECT; <=== PLEASE DO NOT PLACE LINES AFTER THIS LINE.        |
|***** ***** BOTTOM OF DATA *****|
|
\-----/

```

Figure 2-6.04. SAS System Statements Screen - Coupling Facility Planning

The SAS System Statements screen (Figure 2-6.04) lets you define your own data selection logic. You can code any valid SAS statements as described in the SAS manuals. The %MACRO SELECT and %MEND SELECT statements are already provided and field protected from modification.

Note: This is a PDF/EDIT screen. All PDF/EDIT commands are available.

2.6.3 Standard Output

The capacity planning database file that is generated by the Coupling Facility Planning Sample Application contains the data following elements:

- A. Data elements common to all CA MICS Capacity Planner database files:

- SYSID
 - ZONE or HOUR or DAYNAME
 - YEAR
 - MONTH or WEEK or MNTHDAY
 - DATE

- B. Data elements unique to the Coupling Facility Planning Sample Application database file:

- 1. Character format data elements

- CAPAPU - SYSPLEX
 - CAPAPU2 - Coupling Facility Name
 - CAPAPU3 - System using Coupling Facility

2. Numeric format data elements

CFPTOCPU - Total Processors
CFPTOTSD - Total Storage Defined (KB)
CFPTOTSF - Total Storage Free (KB)
CFPTOCSD - Total Control Storage Defined (KB)
CFPTOCSE - Total Control Storage Free (KB)
CFPTODSA - Total Dump Storage Allocated (KB)
CFPTODSE - Total Dump Storage Free (KB)
CFPTOPTH - Total Paths Available
CFPTOREQ - Total Requests Made by this System
CFPTOURQ - Total Unsuccessful Requests
CFPTOSCG - Total Subchannels in IOGEN
CFPTOSCU - Total Subchannels in use
ALLTOREQ - Total Requests Made in SYSPLEX
ALLTOURQ - Total Unsuccessful Requests in SYSPLEX

3. Time format data elements

CFPTOBTM - Processor Busy Time
CFPTOWTM - Processor Wait Time

4. Percentage format data elements

PCTTOREQ - Percentage of total requests made by this system
PCTTOURQ - Percentage of total unsuccessful requests made by this system

Many of the data elements used in the Coupling Facility Planning Sample Application database file have the same meaning as their identically named counterparts in the CA MICS database.

The data elements that are derived from the CA MICS database and used by the Coupling Facility Planning Sample Application are calculated as follows:

ALLTOREQ: The total number of requests made in the Enhanced SYSPLEX.

ALLTOURQ: The total number of unsuccessful requests made in the Enhanced SYSPLEX.

PCTTOREQ: The percentage of signals made by this system in context of the entire Enhanced SYSPLEX.

$$\text{PCTTOREQ} = 100 * \text{CFPTOREQ} / \text{ALLTOREQ}$$

PCTTOURQ: The percentage of unsuccessful signals made by this system in context of the entire Enhanced SYSPLEX.

$$\text{PCTTOURQ} = 100 * \text{CFPTOURQ} / \text{ALLTOURQ}$$

2.7 z/VM Processor Planning

The z/VM Processor Planning Sample Application provides a capacity planning database file that contains the necessary elements for z/VM processor tracking and reporting.

The following sections cover the use, input specifications, and output of the sample application:

- 1 - Usage Guidelines
- 2 - Control Parameters
- 3 - Standard Output

2.7.1 Usage Guidelines

The z/VM Processor Usage Planning Standard Application provides a capacity planning database file with the database elements you need for basic tracking, reporting, and forecasting of z/VM activity. The information for this application is derived from either the DETAIL, DAYS, WEEKS, or MONTHS timespan of the z/VM System Global Data (VMXSGP) File.

The z/VM Processor Usage Planning Standard Application permits your installation to track and forecast z/VM activity at a high level of summarization (that is, at the SYSID and ZONE, HOUR or DAYNAME level). It should be noted that this standard application provides information about system-wide z/VM processor use, rather than individual workload use.

2.7.2 Control Parameters

Relatively few parameters are needed to implement the z/VM Processor Usage Planning Sample Application. The control screens included in the online facility allow you to quickly enter the required information.

```

/----- z/VM Processor Usage Planning -----\
|Command ==>|
|Enter a ? in any data entry field for more information on valid values.|
|Composing CAP DB File def.: ZVM - z/VM Processor Usage Plan. Sample Appl.|
|
|Summarization Level|
|      Primary      ==> MONTHS  (DAYS/WEEKS/MONTHS)|
|      Secondary    ==> ZONE    (ZONE/HOUR/DAYNAME)|
|
|Input File Time-Span ==> DETAIL  (DETAIL/DAYS/WEEKS/MONTHS)|
|
|Amount of data retained ==> 60    (1-9999 months)|
|Age of oldest forecast ==> 24    (1-9999 months)|
|History cycles ==> 3    (1-99 cycles)|
|Forecast cycles ==> 3    (1-99 cycles)|
|
|Data base ID (s) ==> _____|
|
|Specify extended options ==> Y (Y/N)|
\-----/

```

Figure 2-7.01. z/VM Planning Screen

Figure 2-7.01 shows the z/VM Processor Usage Planning screen, which is the main control screen for the z/VM Processor Usage Planning Sample Application. It contains the following fields:

Summarization level:	The output summarization corresponds to the combination of the primary and secondary summarization levels. These control how the output file is summarized in your capacity Planning database.
Input file timespan:	The timespan (either DETAIL, DAYS, WEEKS or MONTHS) of the VMXSGP File that you want to use as input when creating the capacity planning database file.

Amount of data retained:	The number of days, weeks or months of historical data to be retained in the CAPxxx01 cycle of the capacity planning database file. Any historical observations older than what you specify for this parameter are dropped from the file during the file update.
Age of oldest forecast:	The number of days, weeks or months that you want to retain a forecast after it is generated. Any forecast older than what you specify for this parameter is dropped when a new forecast is generated and saved.
History cycle:	Number of cycles or versions of the historical data you want to retain. A new cycle is created when the file is updated and the existing cycles aged (for example, Cycle 01 is aged to Cycle 02 and Cycle 02 becomes Cycle 03).
Forecast cycle:	Number of cycles or versions of the forecast data you want to retain. A new cycle is created when you specify forecast aging in one of the workload forecasting routines and the existing cycles are aged (for example, Cycle 01 is aged to Cycle 02 and Cycle 02 becomes Cycle 03).
database ID(s):	The database unit identifier(s) (up to fourteen units) that contain the VMXSGP File to be used with this sample application. If you are not sure where the VMXSGP files reside, type a question mark (?) in this field to display a listing of the valid database unit identifiers for this sample application.

Specify Extended Options:	If you want extended options, enter Y (YES), and the Extended Options Menu (Figure 2-7.02) is displayed.

```

/----- Extended Options Menu -----\
|Command ==>
|
|Inquiry Step:  z/VM Processor Usage Planning Sample Application
|
|Combine all SYSIDs into a composite ==> N          (Y/N)
|   Composite SYSID                  ==>  ____
|
|Combine all zones into a composite ==> N          (Y/N)
|   Composite zone                   ==>  _____
|
|Specify:
|   Common data selection parameters ==> Y          (Y/N/R)
|   Global data selection exit       ==> Y          (Y/N/R)
|   Notes report enabled             ==> YES        (YES/NO)
|   Contents reports enabled         ==> YES        (YES/NO)
|
|
|-----\

```

Figure 2-7.02. z/VM Planning Extended Options Menu

The extended options available with the z/VM Processor Usage Planning Sample Application (Figure 2-7.02) are:

Combine all SYSIDs into a composite: You can combine SYSIDs in the capacity planning database file. Enter Y (YES) if you want data for all of the SYSIDs to be summarized into a single observation with the Composite SYSID value.

Combine all zones/hours/daynames into a composite:	You can combine zones, hours, and daynames in the capacity planning database file. The prompt for this field changes to zones, hours or daynames depending on what you specify for the summary level parameter. Enter Y (YES) if you want all of the zones, hours, and daynames to be summarized into a single observation with the composite value.
Specify common data selection parameters:	Specify the SYSIDs and zones, hours or daynames to be used for data selection from the CA MICS input file(s). Enter Y (YES) to display the Common Data Selection screen (Figure 2-7.03). Enter R (RESET) to reset any selections you previously made to SYSID and zone, hour or dayname.
Specify global data selection exit:	You can use your own SAS statements to select data from the input input file(s). Enter Y (YES) to display the SAS System Statements screen (Figure 2-7.04). Enter R (RESET) to reset any selections you previously made to the default of no user- specified data selection.
Notes report enabled	If you have entered notes associated with the capacity planning file definition, a Yes in this field prints the note as part of file updating.

Contents reports: enabled:	A Yes in this field prints the contents of the capacity planner database file as part of file updating.

```

/----- Common Data Selection -----\
|Command ==>
|
|Inquiry Step: z/VM Processor Usage Planning Sample Application
|
|SYSID ==> _ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _
|          _ _ _ _ _ _ _ _ _ _
|
|Zone ==> _ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _      (1 - 9)
|
|Hour ==> __ - __ - __ - __ - __ - __ - __ - __      (0 - 23)
|
|Dayname==> _ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _
|
|
|
|
|
\-----\

```

Figure 2-7.03. Common Data Selection Screen - z/VM Planning

The Common Data Selection screen (Figure 2-7.03) allows you to define the following data selection criteria:

SYSID: The SYSIDs that you want to include in your capacity planning database file. You can specify a maximum of 15 one- to four-character SYSIDs. Leave this parameter blank to include data for all SYSIDs.

Zone: You can specify up to nine zones to be included in the capacity planning database file. Leave this parameter blank to include data for all zones.

- Hour:** You can specify up to five hours, or hour ranges to be included in the capacity planning database file. Selection of this parameter is valid only if you specify the DETAIL, DAYS, or WEEKS timespan on the z/VM Planning screen (Figure 2-7.01). Leave this parameter blank to include data for all hours. Singular hour values must be entered to the left of the dash.
- Dayname:** You can specify up to five daynames to be included in the capacity planning database file. Leave blank to include all days. Input the first unique characters and the application will fill in the rest. Specification of this parameter is only valid when the input timespan is DETAIL or DAYS.

```

/----- SAS System Statements ----- COLUMNS 001 072--\
|Command ==>                               Scroll ==> PAGE |
|
|SAS Statements for the SELECT Macro
|-----|
|***** ***** TOP OF DATA *****|
|==MSG> SPECIFY SAS STATEMENTS TO SUBSET THE CAPACITY Processor Usage Planning |
|NOTE= %MACRO SELECT; <=== PLEASE DO NOT PLACE LINES BEFORE THIS LINE. |
|000001
|000002
|000003
|000004
|000005
|000006
|000007
|000008
|000009
|000010
|NOTE= %MEND SELECT; <=== PLEASE DO NOT PLACE LINES AFTER THIS LINE. |
|***** ***** BOTTOM OF DATA *****|
|
\-----/

```

Figure 2-7.04. SAS System Statements Screen-z/VM Planning Global Data Selection

The SAS System Statements screen (Figure 2-7.04) lets you define your own data selection logic. You can code any valid SAS statements as described in the SAS manuals. The %MACRO SELECT and %MEND SELECT statements are already provided and field protected from modification.

Note: This is a PDF/EDIT screen. All PDF/EDIT commands are available.

2.7.3 Standard Output

The capacity planning database file generated by the z/VM Processor Usage Planning Sample Application contains the following:

- A. Data elements common to all CA MICS Capacity Planner database files:

SYSID
ZONE, HOUR or DAYNAME
YEAR
MONTH, WEEK or MNTHDAY
DATE

- B. Data elements unique to the z/VM Processor Usage Planning Sample Application database file:

- 1. Character format data elements

CPUMODEL - CPU Model Identification

- 2. Numeric format data elements

AVGACTU - Average Number of Active Users
AVGLOGU - Average Number of Logged Users
INTERVLS - Number of Recording Intervals
Q1EUSER - E1 Users in Eligible List
Q2EUSER - E2 Users in Eligible List
Q1LDUSR - Count of Loading Users in Q1
SGPXPGIN - Number of Fast Path Pugins Performed
SGPPGFCT - Count of Page Faults on this CPU
SGPPGRIO - Count of Read Operations for Paging
SGPTVRAT - Total to Virtual Ratio (System)
SGPTVRVM - Total to Virtual Ratio (Users)
SGPPGRTA - Page Rate to Auxiliary Storage
SGPPGRTE - Page Rate
SGPPGRTX - Page Rate to Expanded Storage
SGPSTGUT - Storage Utilization

SGPACTVF - Active Vector Facility Users
SGPVFLOD - Count of Vector Facility Loads
SGPVRFRE - System V=R Free Storage Area Size
SGPXTCCP - Expanded Storage Configured to CP
SGPXTSIZ - Expanded Storage Installed
SGPCDLDG - Count of Loading Users in Dispatch Lst
SGPCTXAV - Expanded Stor. Available but Not Used
SGPC1DLD - Count of Loading Users in Q1, Q2, & Q3
SGPC1DSP - Q1, Q2, and Q3 Users in Dispatch List
SGPC1ELD - Count of Loading Users in E1, E2, & E3
SGPC1ELG - E1, E2, and E3 Users in Eligible List
SGPC2DLD - Count of Loading Users in Q2, and Q3
SGPC2DSP - Q2 and Q3 Users in Dispatch List
SGPC2ELD - Count of Loading Users in E2, and E3
SGPC2ELG - E2, and E3 Users in Eligible List
SGPC3DLD - Count of Loading Users in Q3
SGPC3DSP - Q3 Users in Dispatch List
SGPC3ELD - Count of Loading Users in E3
SGPC3ELG - E3 Users in Eligible List
SGPPGOUT - Number of PAGOUTs Performed
SGPPGWIO - Count of Write Operations for Paging
SGPSPGIN - Number of Long Path PAGINs Performed
SGPCTXAV - Expanded Storage Available but Not Used
SGPCPPAR - CP Expanded Storage Partition Size
SGPNTCUP - Non-Trivial Trans. Run in Virtual UP
SGPNTCMP - Non-Trivial Trans. Run in Virtual MP
SGPFVRUD - Size of V=R Free Storage in Use
SGPFSYUD - Size of System Free Storage in Use
SGPFSTOR - Free Storage Available and In Use
SGPESWSS - Sum Of Estimated WSS for In Queue Users

3. Time format data elements

DURATION - Duration of the Interval Reported
NTRRSPUP - Non-Triv. Trans. Response Virtual UP
NTRRSPMP - Non-Triv. Trans. Response Virtual MP
SGPCPUTM - Total CPU Time Charged to Users
SGPPRBTM - Total CPU Time Spent in Emulation Mode
SGPSYSTM - Total CPU Time Charged to the System
SGPSUPTM - Total CP CPU Time Charged to Users
SGPTOTTM - Total CPU Time Consumed
SGPVFOTM - Vector CP Time Used to Switch Users
SGPVFVTM - Vec. User Time Used while in SIE Mode
SGPSYSTM - Total CPU Time Charged to the System

SGP002TM - Total Elapsed time for All 0/02 Records
 SGPNTRUP - Time for Non-Trivial Trans in Virtual UP
 SGPNTRMP - Time for Non-Trivial Trans in Virtual MP
 SGPWATTM - Total System Wait Time
 TRVRSPUP - Triv. Trans. Response Time Virtual UP
 TRVRSPMP - Triv. Trans. Response Time Virtual MP

4. Percentage format data elements

SGPPCFSU - CP Free Storage Utilization
 SGPPCPIX - Percent Page-Ins From Expanded Storage
 SGPPCPOX - Percent Page-Outs to Expanded Storage
 SGPPCPRB - Percent CPU Virtual Time
 SGPPCSUP - Percent CPU Supervisor Time
 SGPPCSYS - Percent CPU System Time
 SGPPCTOT - Percent CPU Total Time
 SGPPCVRP - V=R Free Storage Utilization
 SGPPCWAT - Percent CPU Wait Time
 SGPPCXU - Percent CP Expanded Storage in Use

Most of the previous data elements have the same meaning and derivation as their identically named counterparts in the CA MICS VMXSGP file.

AVGACTU: Average number of active users. It is calculated using the following formula:

$$\text{AVGACTU} = \text{SGPACTIV} / \text{INTERVLS}$$

where SGPACTIV is the accumulated number of active users during all recording intervals.

AVGLOGU: Average number of logged users. It is calculated using the following formula:

$$\text{AVGLOGU} = \text{SGPUSRLG} / \text{INTERVLS}$$

where SGPUSRLG is the accumulated number of logged on users during the recording intervals.

Q1EUSER: E1 users in eligible list. It is calculated using the following formula:

$$\text{Q1EUSER} = \text{SGPC1ELG} - \text{SGPC2ELG}$$

Q2EUSER: E2 users in eligible list. It is calculated using the following formula:

$$Q2EUSER = SGPC2ELG - SGPC3ELG$$

Q1LDUSR: Count of loading users in Q1. It is calculated using the following formula:

$$Q1LDUSR = SGPC1DLD - SGPC2DLD$$

NTRRSPUP: Non-Trivial transaction response time in a Virtual UP. It is calculated using the following formula:

$$NTRRSPUP = SGPNTUP / SGPNTCUP$$

where SGPNTUP is the total CPU time for non-trivial transactions in a Virtual UP, and SGPNTCUP is the count of non-trivial transactions executed in a Virtual UP.

NTRRSPMP: Non-Trivial transaction response time in a Virtual MP. It is calculated using the following formula:

$$NTRRSPMP = SGPNTMP / SGPNTCMP$$

where SGPNTMP is the total CPU time for non-trivial transactions in a Virtual MP, and SGPNTCMP is the count of non-trivial transactions executed in a Virtual MP.

TRVRSPUP: Trivial transaction response time Virtual UP. It is calculated using the following formula:

$$TRVRSPUP = SGPTRVUP / SGPTCTUP$$

where SGPTRVUP is the total CPU time for trivial transactions in a Virtual UP, and SGPTCTUP is the count of trivial transactions executed in a Virtual UP.

TRVRSPMP: Trivial transaction response time Virtual MP. It is calculated using the following formula:

$$\text{TRVRSPMP} = \text{SGPTRVMP} / \text{SGPTCTMP}$$

where SGPTRVMP is the total CPU time for trivial transactions is a Virtual MP, and SGPTCTMP is the count of trivial transactions executed in a Virtual MP.

2.8 SMS Resource Planning

The SMS Resource Planning Sample Application provides a capacity planning database file that contains the necessary elements for tracking and reporting storage in an SMS environment.

The following sections cover the use, input specifications, and output of the sample application:

- 1 - Usage Guidelines
- 2 - Control Parameters
- 3 - Standard Output

2.8.1 Usage Guidelines

The SMS Resource Planning Sample Application provides a Capacity Planning database file with the database elements you need for basic tracking, reporting, and forecasting of the SMS-managed storage usage on your system. The information for this application is derived from the CA MICS VOA Volume Allocation (VCAVOA) and Data Set Allocation (VCADAA) files.

STORAGE CATEGORIES

In the SMS Resource Planning Sample Application, the application workload group (CAPAPU) data element allows you to identify your individual SMS storage categories as workload groups for more meaningful tracking and reporting. For each SMS storage group, storage class, or management class, the application stores in the capacity planning database file measures such as megabytes allocated and used, and megabytes allocated, used, and available for the entire measured system.

The SMS Resource Planning Sample Application has four methods for specifying the manner in which the SMS-managed systems are grouped: 1) defining each SMS storage group as its own storage category (CAPAPU = STORNAME), 2) defining each SMS storage class as its own storage category (CAPAPU = STORSTGC), 3) defining each SMS management class as its own storage category (CAPAPU = STORMGTC), and 4) defining your own SAS statements to create the storage category groups and set the CAPAPU values accordingly.

2.8.2 Control Parameters

Relatively few parameters are needed to implement the SMS Planning Sample Application. The control screens included in the online facility allow you to enter the required information quickly.

```

/----- SMS Resource Planning -----\
|Command ==>
|Enter a ? in any data entry field for more information on valid values.
|Composing CAP DB File def.: SM1 - SMS Resource Planning Sample Application
|
|Summarization Level
|      Primary      ==> MONTHS  (DAYS/WEEKS/MONTHS)
|      Secondary    ==> ZONE    (ZONE/HOUR/DAYNAME)
|
|Input File Time-Span ==> DETAIL  (DETAIL/DAYS/WEEKS/MONTHS)
|
|Amount of data retained ==> 60    (1-9999 months)
|Age of oldest forecast  ==> 24    (1-9999 months)
|History cycles          ==> 3     (1-99 cycles)
|Forecast cycles         ==> 3     (1-99 cycles)
|
|Data base ID (s)       ==> _____
|
|SMS storage category (CAPAPU) ==> GROUP (GROUP/STGCLASS/MGTCLASS/EXIT)
|      Application unit title ==> Storage Group_____
|
|Specify extended options ==> Y (Y/N)
|
\-----/

```

Figure 2-8.01. SMS Planning Screen

Figure 2-8.01 shows the SMS Planning screen, which is the main control screen for the SMS Resource Planning Sample Application. It contains the following fields:

Summarization level:	The output summarization corresponds to the combination of the primary and secondary summarization levels. These control how the output file is summarized in your capacity planning database.
Input file timespan:	The timespan (either DETAIL, DAYS, WEEKS or MONTHS) of the VCAVOA and VCADAA Files that you want to use as input when creating the capacity Planning database file.
Amount of data retained:	The number of days, weeks or months of historical data to be retained in the CAPxxx01 cycle of the capacity planning database file. Any historical observations older than what you specify for this parameter are dropped from the file during the file update.
Age of oldest forecast:	The number of days, weeks or months that you want to retain a forecast after it is generated. Any forecast older than what you specify for this parameter is dropped when a new forecast is generated and saved.
History cycle:	The number of cycles or versions of the historical data you want to retain. A new cycle is created when the file is updated and the existing cycles aged (for example, Cycle 01 is aged to Cycle 02 and Cycle 02 becomes Cycle 03).
Forecast cycle:	The number of cycles or versions of the forecast data you want to retain. A new cycle is created when you specify forecast aging in one of the workload forecasting routines and the existing cycles are aged (for example, Cycle 01 is aged to Cycle 02 and Cycle 02 becomes Cycle 03).

Database ID(s):	The database unit identifier(s) (up to fourteen units) that contain the VCAVOA and VCADAA Files to be used with this sample application. If you are uncertain in which database unit the VCAVOA and VCADAA Files reside, type a question mark (?) in this field to display a message listing of the valid database unit identifiers for this sample application.
SMS storage category (CAPAPU):	The four options for defining workload groups (see Section 2.8.1, Usage Guidelines).
Application unit title:	The 1- to 40-character name for the CAPAPU data element. You can use this parameter to document how your workload groupings have been defined. This parameter is only available for entry if you specify EXIT for the SMS workload definition (CAPAPU) parameter. If you select any of the three SMS-related categorizations, this parameter will be specified for you as follows: GROUP -- Storage Group STGCLASS -- Storage Class MGTCLASS -- Management Class

Specify options:	If you want extended options, enter Y (YES), the the Extended Options Menu (Figure 2-8.02) is displayed.

```
/----- Extended Options Menu -----\n|Command ==>\n|\n|Inquiry Step: SMS Resource Planning Sample Application\n|\n|Combine all SYSIDs into a composite ==> N          (Y/N)\n|    Composite SYSID                ==> ____\n|\n|Combine all zones into a composite ==> N          (Y/N)\n|    Composite zone                  ==> _____\n|\nSpecify:\n|    Common data selection parameters ==> Y        (Y/N/R)\n|    Global data selection exit       ==> Y        (Y/N/R)\n|    CAPAPU derivation SAS statements ==> Y        (Y/N/R)\n|    Notes report enabled             ==> YES      (YES/NO)\n|    Contents reports enabled         ==> YES      (YES/NO)
```

Figure 2-8.02. SMS Planning Extended Options Menu

The extended options available with the SMS Resource Planning Sample Application are:

```
Combine all  You can combine SYSIDs in the capacity planning
SYSIDs into database file. Enter Y (YES) if you wish data
a composite: for all of the SYSIDs to be summarized into a
              single observation with the Composite SYSID
              value.
```

Combine all zones/hours/daynames into a composite:	You can combine zones, hours, and daynames in the capacity planning database file. The prompt for this field changes to zones or hours or daynames depending on what you specify for the summary level parameter. Enter Y (YES) if you want all of the zones, hours, and daynames to be summarized into a single observation with the composite value.
Specify common data selection parameters:	Specify the SYSIDs and zones, hours or daynames to be used for data selection from the CA MICS input file(s). Enter Y (YES) to display the Common Data Selection screen (Figure 2-8.03). Enter R (RESET) to reset any selections you previously made to SYSID and zone, hour or dayname.
Specify global data selection exit:	You can use your own SAS statements to select data from the CA MICS input file(s). Enter Y (YES) to display the SAS System Statements screen (Figure 2-8.04). Enter R (RESET) to display or reset any specifications you previously made to the default of no user-specified data selection.
Specify CAPAPI derivation SAS statements:	You can use your own SAS statements for defining your SMS storage categories. Enter Y (YES) to display the SAS System Statements screen. Enter R (RESET) to reset any SAS Statements you previously made to user-specified workload groupings. This option is available only if you specify EXIT for the SMS workload definition (CAPAPI) parameter on the SMS Planning screen.
Notes report enabled:	If you have entered notes associated with the capacity planning file definition, a Yes in this field prints the note as part of file updating.

Contents reports enabled:	A Yes in this field prints the contents of the capacity planning database file as part of file updating.

```

----- Common Data Selection -----
|Command ==>
|
|Inquiry Step: SMS Resource Planning Sample Application
|
|SYSID ==> _____
|          _____
|
|Zone   ==> _ _ _ _ _ (1 - 9)
|
|Hour   ==> _ _ ' _ _ ' _ _ ' _ _ ' _ _ ' _ _ (0 - 23)
|
|Dayname==> _____

```

Figure 2-8.03. Common Data Selection Screen - SMS Planning

The Common Data Selection screen (Figure 2-8.03) allows you to define the following data selection criteria:

SYSID: The SYSIDs you want to include in your capacity planning database file. You can specify a maximum of 15 one- to four-character SYSIDs. Leave this parameter blank to include data for all SYSIDs.

Zone: Specify up to nine zones to be included in the capacity planning database file. Leave this parameter blank to include data for all zones.

- Hour:** Specify the five hours, or hour ranges to be included in the capacity planning database file. Specification of this parameter is valid only if you specify the DETAIL, DAYS or WEEKS timespan on the SMS Planning screen (Figure 2-8.01). Leave this parameter blank to include data for all hours. You must enter singular hour values on the left side of the dash.
- Dayname:** Specify up to five daynames to be included in the capacity planning database file. Leave blank to include all days. Input the first unique characters and the application will fill in the rest. Selection of this parameter is only valid when the input timespan is DETAIL or DAYS.

```

/----- SAS System Statements ----- COLUMNS 001 072--\
|Command ==>                               Scroll ==> PAGE |
|
|SAS Statements for the SELECT Macro                               |
|-----|
|***** ***** TOP OF DATA *****|
|==MSG> SPECIFY SAS STATEMENTS TO SUBSET THE CAPACITY PLANNING DATA BASE FILE. |
|NOTE= %MACRO SELECT; <=== PLEASE DO NOT PLACE LINES BEFORE THIS LINE.      |
|000001                               |
|000002                               |
|000003                               |
|000004                               |
|000005                               |
|000006                               |
|000007                               |
|000008                               |
|000009                               |
|000010                               |
|NOTE= %MEND SELECT; <=== PLEASE DO NOT PLACE LINES AFTER THIS LINE.        |
|***** ***** BOTTOM OF DATA *****|
|
\-----/

```

Figure 2-8.04. SAS System Statements Screen - SMS Planning Global Data Selection

The SAS System Statements screen (Figure 2-8.04) lets you define your own data selection logic. You can code any valid SAS statements as described in the SAS manuals. The %MACRO SELECT and %MEND SELECT statements are already provided and protected from modification.

Note: This is a PDF/EDIT screen. All PDF/EDIT commands are available.

2.8.3 Standard Output

The capacity planning database file generated by the SMS Planning Sample Application contains the following data elements:

- A. Data elements common to all CA MICS Capacity Planner database files:

- SYSID
- ZONE (Always equal to "A") - see Note)
- YEAR
- MONTH or WEEK or MNTHDAY
- DATE

Note: The VCA Information Area database files do not use the ZONE variable. The SMS Resource Planning Sample Application automatically sets the ZONE variable to A. Code an A where a ZONE value is required when using the component's trending programs.

- B. Data elements unique to the SMS Resource Planning Sample Application database file:

- 1. Character format data elements

- CAPAPU - SMS Storage Category

2. Numeric format data elements

DAAALLSP - Amount of Space Allocated (Mbytes)
(for each Storage Category)
DAAUSESP - Amount of Space Used (Mbytes)
(for each Storage Category)
MBALOC - Megabytes Allocated for system
MBTOTL - Megabytes Total for measured system
MBUSED - Megabytes Used for system
PCATRAL - Percent of Total Allocation for CAPAPU
(for each Storage Category)
PCATRUS - Percent of Total Usage for CAPAPU
(for each Storage Category)
PCTALOC - Percent Allocated for System
PCTUSED - Percent Used for System
STGVOLNO - Total Volumes for Storage Group
(Present only if CAPAPU=GROUP)

DERIVATION OF THE SMS PLANNING DATA FILE ELEMENTS

The following paragraphs discuss the DASD Planning database file elements.

CAPAPU: The data element that contains the values found either in STORNAME, STORSTGC, STORMGTC, or values derived within the user's exit. In the capacity planning database, there is a separate observation for each value of CAPAPU for each MONTH, WEEK or DAY. In addition to the groups defined, there will be an observation for a CAPAPU value of "***OTHER**" that represents all the storage space having blank values for CAPAPU from STORNAME, STORSTGC, or STORMGTC. It is recommended, but not required, that, if the you write an exit, this exit provide a catchall category like "***OTHER**".

STGVOLNO: This is obtained from VCAVOA, if CAPAPU is set to storage group (STORNAME). Its creation is skipped if any CAPAPU other than storage group is selected.

In the calculation of the numeric data elements, the elements that are used from the CA MICS VCAVOA and VCADAA Files are the following:

DAAALLSP - Amount of Space Allocated (Mbytes)
DAAUSESP - Amount of Space Used (Mbytes)
VOACAPMB - Total Capacity of Volume in Megabytes

The data elements DAAALLSP, DAAUSESP, MBALOC, MBUSED, and MBTOTL all represent average usage statistics for the timespan being considered (that is, MONTH). Over the span of a month, actual use may vary considerably as SMS volumes are installed or discontinued, and data sets are allocated, extended, released, or scratched.

During the summarization process, VOAOLNO is used to weigh the individual observations from the VCAVOA File and DAASAMPL is used to weigh the individual observations from the VCADAA File. This allows the numeric data elements contained in the capacity planning database file to be calculated correctly, even if a volume resided on more than one device type during a month, or a particular category of data sets was present during only a part of a month.

The data elements from the VCAVOA and VCCDAA Files that are used to compute the DASD Planning file elements are themselves derived from the CA MICS VTOC Catalog Collector (VCC) Component. This product performs a periodic scan on all volumes that are defined to it. The elements defined next represent averages over all scans taken during the month.

DAAALLSP: The amount of space allocated for each CAPAPU. This value is computed as the sum of DAAALLSP for each unique value of CAPAPU.

DAAUSESP: The amount of space used for each CAPAPU. This value is computed as the sum of DAAUSESP for each unique value of CAPAPU.

MBALOC: The total allocated space in the measured system. This value is computed as the sum of DAAALLSP for the entire measured system.

MBUSED: The total used space in the measured system. This value is computed as the sum of DAAUSESP for the entire measured system. It represents productively used space since this space always contains some sort of data, either user or system defined. Because used space must be allocated, MBALOC is always greater than MBUSED. This space is allocated for the growth of existing data sets and therefore is not available for the addition of new data sets. The difference between MBALOC and MBUSED is the space that is allocated, but unused. A large amount of unused space indicates that a considerable amount of space has been reserved by current users, but is not being used.

MBTOTL: The total space in the measured system. This value is computed as the sum of VOACAPMB for each VOLSER measured.

Note that the above data element definitions have all been expressed in units of megabytes. For purposes of the SMS Planning Sample Application, a megabyte is defined as 1,000,000 bytes. This is consistent with published specifications for storage devices and with other SMS-related calculations performed for files in the VCA information area.

Percentage values have these definitions and derivations:

PCTUSED: Percent used for measured System. The percent of total space for the measured system which is used.

$$\text{PCTUSED} = 100 * (\text{MBUSED} / \text{MBTOTL})$$

PCTALOC: Percent allocated for measured system. The percent of total space for the measured system which is allocated.

$$\text{PCTALOC} = 100 * (\text{MBALOC} / \text{MBTOTL})$$

PCATRAL: Percent of total allocation for CAPAPU. The percent of total allocated space for the measured system which is allocated for a particular CAPAPU category.

$$\text{PCATRAL} = 100 * (\text{DAAALLSP} / \text{MBALOC})$$

PCATRUS: Percent of total usage for CAPAPU. The percent of total used space for the measured system which is used for a particular CAPAPU category.

$$\text{PCATRUS} = 100 * (\text{DAAALLSP} / \text{MBUSED})$$

2.9 CICS Resource Planning

The CICS Resource Planning Sample Application provides a capacity planning database file that contains the necessary elements for CICS resource tracking and reporting.

The following sections cover the usage, input specifications, and output of the sample application:

- 1 - Usage Guidelines
- 2 - Control Parameters
- 3 - Standard Output

2.9.1 Usage Guidelines

The CICS Resource Planning Sample Application provides a capacity planning database file with the data elements you need for basic tracking, reporting, and forecasting of the CICS workload.

In the CICS Resource Planning Sample Application, the application workload group (CAPAPU) data element allows you to group your individual CICS systems into logically similar workload groups for more meaningful tracking and reporting. Three methods are provided for specifying the manner in which the CICS systems are grouped: 1) defining each CICS system as its own workload group (CAPAPU = CICSID), 2) combining all CICS systems (no workload group), or 3) defining your own SAS statements to create the workload groupings and setting the CAPAPU values accordingly.

2.9.2 Control Parameters

You need relatively few parameters to implement the CICS Resource Planning Sample Application. The control screens included on the online facility allow you to quickly enter the information required.

```

/----- CICS Resource Planning -----\
|Command ==>
|Enter a ? in any data entry field for more information on valid values.
|Composing CAP DB File def.: CIC - CICS Resource Planning Sample Application
|
|Summarization Level
|      Primary      ==> MONTHS  (DAYS/WEEKS/MONTHS)
|      Secondary    ==> ZONE    (ZONE/HOUR/DAYNAME)
|
|Input File Time-Span ==> DETAIL  (DETAIL/DAYS/WEEKS/MONTHS)
|
|Amount of data retained ==> 60    (1-9999 months)
|Age of oldest forecast ==> 24    (1-9999 months)
|History cycles ==> 3    (1-99 cycles)
|Forecast cycles ==> 3    (1-99 cycles)
|
|Data base ID (s) ==> _____
|
|CICS workload definition (CAPAPU) ==> CICSID (CICSID/NONE/EXIT)
|      Application unit title ==> _____
|
|
|Specify extended options ==> N (Y/N)
\-----/

```

Figure 2-9.01. CICS Planning Screen

Figure 2-9.01 shows the CICS Planning screen, which is the main control screen for the CICS Resource Planning Sample Application. It contains the following fields:

Summarization level: The output summarization corresponds to the combination of the primary and secondary summarization levels. These control how the output file is summarized in your capacity planning database.

Input file timespan:	The timespan (either DETAIL, DAYS, WEEKS or MONTHS) of the CICC SU File that you want to use as input when creating the capacity planning database file.
Amount of Data retained:	The number of days, weeks or months of historical data to be retained in the CAPxxx01 cycle of the capacity planning database file. Any historical observations older than what you specify for this parameter are dropped from the file during the file update.
Age of oldest forecast:	The number of days, weeks or months that you want to retain a forecast after it is generated. Any forecast older than what you specify for this parameter is dropped when a new forecast is generated and saved.
History cycles:	The number of cycles or versions of the historical data that you want to retain. A new cycle is created when the file is updated and the existing cycles are aged (for example, Cycle 01 becomes Cycle 02 and Cycle 02 becomes Cycle 03).
Forecast cycles:	The number of cycles or versions of the forecast data that you want to retain. A new cycle is created when you specify forecast aging in one of the workload forecasting routines and the existing cycles are aged (for example, Cycle 01 becomes Cycle 02 and Cycle 02 becomes Cycle 03).
Data base ID(s):	The database unit identifier(s) (up to fourteen units) that contain the CICC SU File to be used with this sample application. If you are not sure where the CICC SU File resides, type a question mark (?) in this field to display a listing of the valid database unit identifiers for this sample application.
CICS workload definition (CAPAPU):	The three options for defining workload groups (see Section 2.9.1, Usage Guidelines).

Seven extended options are available with the CICS Resource Planning Sample Application, as shown in Figure 2-9.02. These options are described next.

Combine all SYSIDs into a composite:	You can combine SYSIDs in the capacity planning database file. Enter Y (YES) if you want data for all of the SYSIDs to be summarized into a single observation with the Composite SYSID value.
Combine all zones/hours/ daynames into a composite:	You can combine zones, hours and daynames in the capacity planning database file. The prompt for this field changes to zones, hours or daynames depending on what you specify for the summary level parameter. Enter Y (YES) if you want all of the zones, hours, and daynames to be summarized into a single observation with the composite value. Figure 2-9.03 illustrates specifying zones, hours, and daynames for inclusion in the database.
Specify common data selection parameters:	Specify the SYSIDs and zones, hours or daynames to be used for data selection from the CA MICS input file(s). Enter Y (YES) to display the Common Data Selection screen shown in Figure 2-9.03. Enter R (RESET) to reset any selections you previously made to SYSID and zone, hour or dayname.
Specify global data selection exit:	You can use your own SAS statements to select data from the CA MICS input file(s). Enter Y (YES) to display the SAS System Statements screen shown in Figure 2-9.04. Enter R (RESET) to reset any selections you previously made to the default of no user-specified data selection.
Specify CAPAPU derivation SAS statements:	You can use your own SAS statements for defining your CICS workloads. Enter Y (YES) to display the SAS System Statements screen. Enter R (RESET) to reset any SAS SYSTEM STATEMENTS you previously made to a user-specified workload grouping. This option is available only if you specify EXIT for the CICS workload definition (CAPAPU) parameter on the CICS Planning screen.
Notes report enabled:	If you have entered notes associated with the capacity planning file definition, a Yes in this field prints the note as part of file updating.

Hour: Specify the five hours or hour ranges to be included in the capacity planning database file. Selection of this parameter is valid only if you specify the DETAIL, DAYS, or WEEKS timespan on the CICS Planning screen shown in Figure 2-9.01. Leave this parameter blank to include data for all hours.

Dayname: Specify up to five daynames to be included in the capacity planning database file. Leave blank to include all days. Input the first unique characters and the capacity planner will fill in the rest. Specification of this parameter is only valid when the input timespan is detail or days.

```

/----- SAS System Statements ----- COLUMNS 001 072--\
|Command ==>                                         Scroll ==> PAGE |
|
|SAS Statements for the SELECT Macro                  |
|-----|
|***** ***** TOP OF DATA *****|
|=MSG> SPECIFY SAS STATEMENTS TO SUBSET THE CAPACITY PLANNING DATA BASE FILE. |
|=NOTE= %MACRO SELECT; <=== PLEASE DO NOT PLACE LINES BEFORE THIS LINE.      |
|000001                                           |
|000002                                           |
|000003                                           |
|000004                                           |
|000005                                           |
|000006                                           |
|000007                                           |
|000008                                           |
|000009                                           |
|000010                                           |
|=NOTE= %MEND SELECT; <=== PLEASE DO NOT PLACE LINES AFTER THIS LINE.        |
|***** ***** BOTTOM OF DATA *****|
|
\-----/

```

Figure 2-9.04. SAS System Statements Screen - CICS Planning Global Data Select on

The SAS System Statements screen shown in Figure 2-9.04 lets you define your own data selection logic. You can code any valid SAS statements, as described in the SAS manuals. The %MACRO SELECT and %MEND SELECT statements are already provided and are field protected from modification.

Note: This is a PDF/EDIT screen. All have PDF/EDIT commands are available.

2.9.3 Standard Output

The capacity planning database file that is generated by the CICS Resource Planning Sample Application contains the following information:

- A. Data elements common to all CA MICS Capacity Planner database files:

- SYSID
- ZONE or HOUR or DAYNAME
- YEAR
- MONTH or WEEK or MNTHDAY
- DATE

- B. Data elements that are unique to the CICS Resource Planning Sample Application database file:

- 1. Character format data elements

- CAPAPU - CICS Workload Group

- 2. Numeric format data elements

- CSUCTRN - Conversational Transactions Processed
 - CSULTRN - Long Transactions Processed
 - CSUMTRN - Medium Transactions Processed
 - CSUSTRN - Short Transactions Processed
 - CSUFCOPS - Number of CICS File Control Calls
 - CSUJCOPS - Number of CICS Journal Control Calls
 - CSUTDOPS - Number of CICS Transient Data Calls
 - CSUTSOPS - Number of CICS Temporary Storage Calls
 - CSUTRANS - Number of Ended Transactions
 - CSUSIOCT - Access Method Calls

3. Time format data elements

- * CSUAVCTM - Average Conversational Response Time
- * CSUAVLTM - Average Long Response Time
- * CSUAVMTM - Average Medium Response Time
- * CSUAVSTM - Average Short Response Time
- CSUAVTTM - Average Response Time for all transactions
- CSUCPUTM - Task CPU Time
- CSUFCWAT - Wait Time for CICS File Control
- CSUTSWAT - Wait Time for CICS Temporary Storage
- CSUJCWAT - Wait Time for CICS Journal Control
- DURATION - Interval Duration

* Note: These data elements are non-essential on the CA MICS database. These data elements contain missing values in your capacity planning database file unless you specify ESSENTIAL = ALL when the CGEN is run for your CA MICS CICS Analyzer.

Many of the data elements used in the CICS Resource Planning Sample Application database file have the same meaning and derivation as their identically named counterparts in the CA MICS CICCISU File.

2.10 IMS Resource Planning

The IMS Resource Planning Sample Application provides a capacity planning database file that contains the necessary elements for IMS resource tracking and reporting.

The following sections cover the use, input specifications, and output of the sample application:

- 1 - Usage Guidelines
- 2 - Control Parameters
- 3 - Standard Output

2.10.1 Usage Guidelines

The IMS Resource Planning Sample Application provides a capacity planning database file with the database elements you need for basic tracking, reporting, and forecasting of the IMS workloads on your system. The information for this application is derived from the CA MICS IMS User Activity - Full Function (IMISISU) and IMS System Activity (IMSISY) files.

WORKLOAD GROUPS

In the IMS Resource Planning Sample Application, the application workload group (CAPAPU) data element allows you to group your individual IMS systems into logically similar workload groups for more meaningful tracking and reporting. For each defined workload, the application stores in the capacity planning database file, measures such as transaction counts, CPU time, database calls, and various quantifications of response time.

Three methods are provided for specifying the manner in which the IMS systems are grouped: 1) defining each IMS system as its own workload group (CAPAPU = IMSID), 2) combining all IMS systems (CAPAPU = NONE, no workload group) and 3) defining your own SAS statements to create the workload group and set the CAPAPU values accordingly.

2.10.2 Control Parameters

Relatively few parameters are needed to implement the IMS Resource Planning Sample Application. The control screens included in the online facility allow you to quickly enter the required information.

```

/----- IMS Resource Planning -----\
|Command ==>
|Enter a ? in any data entry field for more information on valid values.
|Composing CAP DB File def.: IM1 - IMS Planning Standard Application
|
|Summarization Level
|      Primary      ==> MONTHS  (DAYS/WEEKS/MONTHS)
|      Secondary    ==> ZONE    (ZONE/HOUR/DAYNAME)
|
|Input File Time-Span ==> DETAIL  (DETAIL/DAYS/WEEKS/MONTHS)
|
|Amount of data retained ==> 60    (1-9999 months)
|Age of oldest forecast ==> 24    (1-9999 months)
|History cycles ==> 3    (1-99 cycles)
|Forecast cycles ==> 3    (1-99 cycles)
|
|Data base ID (s) ==> _____
|
|IMS workload definition (CAPAPU) ==> IMSID (IMSID/NONE/EXIT)
|      Application unit title ==> _____
|
|Specify extended options ==> Y (Y/N)
\-----/

```

Figure 2-10.01. IMS Resource Planning Screen

Figure 2-10.01 shows the IMS Resource Planning screen, which is the main control screen for the IMS Resource Planning Sample Application. It contains the following fields:

Summarization level: The output summarization corresponds to the combination of the primary and secondary summarization levels. These control how the output file is summarized in your capacity planning database.

Input file time-span: The timespan (either DETAIL, DAYS, WEEKS or MONTHS) of the IMSISU and IMSISY Files that you want to use as input when creating the capacity planning database file.

Amount of data retained:	The number of days, weeks or months of historical data to be retained in the CAPxxx01 cycle of the capacity planning database file. Any historical observations older than what you specify for this parameter are dropped from the file during the file update.
Age of oldest forecast:	The number of days, weeks or months that you want to retain a forecast after it is generated. Any forecast older than what you specify for this parameter is dropped when a new forecast is generated and saved.
History cycles:	The number of cycles or versions of the historical data you want to retain. A new cycle is created when the file is updated and the existing cycles aged (for example, Cycle 01 is aged to Cycle 02 and Cycle 02 becomes Cycle 03).
Forecast cycles:	The number of cycles or versions of the forecast data you want to retain. A new cycle is created when you specify forecast aging in one of the workload forecasting routines and the existing cycles are aged (for example, Cycle 01 is aged to Cycle 02 and Cycle 02 becomes Cycle 03).
Data base ID(s):	The database unit identifier(s) (up to fourteen units) that contain the IMSISU and IMSISY Files to be used with this standard application. If you are uncertain in which database unit the IMSISU and IMSISY Files reside, type a question mark (?) in this field to display a message listing of the valid database unit identifiers for this standard application.
IMS workload definition (CAPAPU):	The three options for defining workload groups (see Section 2.10.1, Usage Guidelines).
Application unit title:	The 1- to 40-character name for the CAPAPU data element. You can use this parameter to document how your workload groupings have been defined. This parameter is only available if you specify EXIT for the IMS workload definition (CAPAPU) parameter.

Combine all zones/hours/daynames into a composite:	You can combine zones, hours, and daynames in the capacity planning database file. The prompt for this field changes to zones or hours or daynames depending on what you specify for the summary level parameter. Enter Y (YES) if you want all of the zones, hours, and daynames to be summarized into a single observation with the composite value. Figure 2-10.02 illustrates specifying zones, hours, and daynames for inclusion in the database.
Specify common data selection parameters:	Specify the SYSIDs and zones, hours or daynames to be used for data selection from the CA MICS input file(s). Enter Y (YES) to display the Common Data Selection screen (Figure 2-10.03). Enter R (RESET) to reset any selections you previously made to SYSID and zone, hour or dayname.
Specify global data selection exit:	You can use your own SAS statements to select data from the CA MICS input file(s). Enter Y (YES) to display the SAS System Statements screen (Figure 2-10.04). Enter R (RESET) to reset any selections you previously made to the default of no user-specified data selection.
Specify CAPAPI derivation SAS statements:	You can use your own SAS statements for defining your IMS workloads. Enter Y (YES) to display the SAS System Statements screen. Enter R (RESET) to reset any SAS statements you previously made to user-specified workload groupings. This option is available only if you specify EXIT for the IMS workload definition (CAPAPI) parameter on the IMS Planning screen.
Notes report enabled:	If you have entered notes associated with the capacity planning file definition, a Yes in this field prints the note as part of file updating.

Contents reports enabled:	A Yes in this field prints the contents of the capacity planning database file as part of file updating.

[illegible]

Figure 2-10.03. Common Data Selection Screen - IMS Resource Planning

The Common Data Selection screen (Figure 2-10.03) allows you to define the following data selection criteria:

SYSID: The SYSIDs you want to include in your capacity planning database file. You can specify a maximum of 15 one- to four-character SYSIDs. Leave this parameter blank to include data for all SYSIDs.

Zone: Specify up to nine zones to be included in the capacity planning database file. Leave this parameter blank to include data for all zones.

Hour: Specify the five hours or hour ranges to be included in the capacity planning database file. Selection of this parameter is valid only if you specify the DETAIL, DAYS, or WEEKS timespan on the IMS Planning screen (Figure 2-10.01). Leave this parameter blank to include data for all hours. You must enter singular hour values on the left side of the dash.

Dayname: Specify up to five daynames to be included in the capacity planning database file. Leave blank to include all days. Input the first unique characters and the application will fill in the rest. Selection of this parameter is only valid when the input timespan is DETAIL or DAYS.

```

/----- SAS System Statements ----- COLUMNS 001 072--\
|Command ==>                                         Scroll ==> PAGE |
|
|SAS Statements for the SELECT Macro                  |
|-----|
|***** ***** TOP OF DATA *****|
|=MSG> SPECIFY SAS STATEMENTS TO SUBSET THE CAPACITY PLANNING DATA BASE FILE. |
|=NOTE= %MACRO SELECT; <=== PLEASE DO NOT PLACE LINES BEFORE THIS LINE.      |
|000001                                           |
|000002                                           |
|000003                                           |
|000004                                           |
|000005                                           |
|000006                                           |
|000007                                           |
|000008                                           |
|000009                                           |
|000010                                           |
|=NOTE= %MEND SELECT; <=== PLEASE DO NOT PLACE LINES AFTER THIS LINE.        |
|***** ***** BOTTOM OF DATA *****|
|
\-----/

```

Figure 2-10.04. SAS System Statements Screen - IMS Planning Global Data Select

The SAS System Statements screen (Figure 2-10.04) lets you define your own data selection logic. You can code any valid SAS statements as described in the SAS manuals. The %MACRO SELECT and %MEND SELECT statements are already provided and are field protected from modification.

Note: This is a PDF/EDIT screen. All PDF/EDIT commands are available.

2.10.3 Standard Output

The capacity planning database file generated by the IMS Resource Planning Sample Application contains the following data elements:

- A. Data elements common to all CA MICS Capacity Planner database files:

SYSID
ZONE or HOUR or DAYNAME
YEAR
MONTH or WEEK or MNTHDAY
DATE

- B. Data elements unique to the IMS Resource Planning Sample Application database file:

1. Character format data elements

CAPAPU - IMS Workload Group

2. Numeric format data elements

ISUCTRN - Conversational Transactions Processed
ISULTRN - Long Transactions Processed
ISUMTRN - Medium Transactions Processed
ISUSTRN - Short Transactions Processed
ISUETRNL - Excessive Transactions Processed
ISUDBDLT - Database Deletes
ISUDBGN - Database Get Nexts
ISUDBGU - Database Get Uniques
ISUDBIST - DataBase Inserts
ISUDBOTH - Other Database Calls (System Service)
ISUDBREP - Database Replaces
ISUDBTOT - Total Number of Database Calls
ISUTRANS - Number of Ended Transactions
ISUUOWCT - Total Units of Work Executed

3. Time format data elements

DURATION - Interval Duration
ISUMRSTM - Medium Transaction Response Time
ISUSRSTM - Short Transaction Response Time
ISUTRSTM - Transaction Response Time
ISUCRSTM - Conv. Transaction Response Time
ISUERSTM - Excessive Transaction Response Time
ISULRSTM - Long Transaction Response Time
ISUAVCTM - Average Conversational Response Time
ISUAVLTM - Average Long Response Time
ISUAVMTM - Average Medium Response Time
ISUAVSTM - Average Short Response Time
ISUAVETM - Average Excessive Response Time
ISUAVTTM - Average Response for All Transactions
ISUCPUTM - Task CPU Time
ISUTMCBF - Control Region Buffer CPU Time
ISUTMCDL - Control Region DL/I Time
ISUTMCRO - Control Region Overhead CPU Time
ISUTMDBF - Dependent Region Buffer Hdlr CPU Time
ISUTMDCP - Dependent Region Processing CPU Time
ISUTMDDL - Dependent Region DL/I Time
ISUTMDOC - Database Open and Close Time
ISUTMDRO - Dependent Region Overhead CPU Time
ISUTMSCH - Dependent Region Sched/Term CPU Time

2.11 DB2 Resource Planning

The DB2 Resource Planning Sample Application provides a capacity planning database file that contains the necessary elements for DB2 resource tracking and reporting.

The following sections cover the use, input specifications, and output of the sample application:

- 1 - Usage Guidelines
- 2 - Control Parameters
- 3 - Standard Output

2.11.1 Usage Guidelines

The DB2 Resource Planning Sample Application provides a capacity planning database file with the database elements you need for basic tracking, reporting, and forecasting of the DB2 workloads on your system. The information for this application is derived from either the DETAIL, DAYS, WEEKS, or MONTHS timespan of the following CA MICS files: the DB2 System Activity (DB2DSY) file, the DB2 Database Activity (DB2DSD) file, and the DB2 User Activity (DB2DSU) file.

The DB2 Resource Planning Sample Application permits your installation to track and forecast DB2 usage at a high level of summarization (that is, at the SYSID and ZONE, HOUR or DAYNAME level). Because the data is derived from the DB2DSY and DB2DSD files, this Sample application provides information about system-wide DB2 usage.

WORKLOAD GROUPS

The DB2 Resource Planning Sample Application allows you to divide your DB2 workloads into groups based on DB2 System Identifier (DB2ID) or to aggregate all DB2 systems (NONE). For each defined workload, the application stores measures such as CPU time, lock escalations, and synchronous I/O activity in the capacity planning database file.

A workload group, represented by the application unit (CAPAPU) data element, allows each DB2 system to be tracked as an entity for capacity planning purposes. If you choose not to utilize workload groups, all individual DB2 systems will have their information summarized into the capacity planning database file.

2.11.2 Control Parameters

Relatively few parameters are needed to implement the DB2 Planning Sample Application. The control screens included in the online facility allow you to quickly enter the required information.

```

/----- DB2 Resource Planning -----\
|Command ==>
|Enter a ? in any data entry field for more information on valid values.
|Composing CAP DB File def.: DB2 - DB2 Resource Planning Sample Application
|
|Summarization Level
|      Primary      ==> MONTHS  (DAYS/WEEKS/MONTHS)
|      Secondary    ==> ZONE    (ZONE/HOUR/DAYNAME)
|
|Input File Time-Span ==> DETAIL  (DETAIL/DAYS/WEEKS/MONTHS)
|
|Amount of data retained ==> 60    (1-9999 months)
|Age of oldest forecast ==> 24    (1-9999 months)
|History cycles ==> 3    (1-99 cycles)
|Forecast cycles ==> 3    (1-99 cycles)
|
|Data base ID (s) ==> _____
|
|DB2 workload definition (CAPAPU) ==> DB2ID (DB2ID/NONE)
|
|Specify extended options ==> Y (Y/N)
\-----/

```

Figure 2-11.01. DB2 Resource Planning Screen

Figure 2-11.01 shows the DB2 Resource Planning screen, which is the main control screen for the DB2 Resource Planning Sample Application. It contains the following fields:

Summarization level:	The output summarization corresponds to the combination of the primary and secondary summarization levels. These control how the output file is summarized in your capacity Planning database.
Input file timespan:	The timespan (either DETAIL, DAYS, WEEKS or MONTHS) of the DB2DSU, DB2DSD, and DB2DSY files that you want to use as input when creating the capacity Planning database file.

Amount of data retained	The number of days, weeks or months of historical data to be retained in the CAPxxx01 cycle of the capacity planning database file. Any historical observations older than what you specify for this parameter are dropped from the file during the file update.
Age of oldest forecast:	The number of days, weeks or months that you want to retain a forecast after it is generated. Any forecast older than what you specify for this parameter is dropped when a new forecast is generated and saved.
History cycle:	The number of cycles or versions of the historical data you want to retain. A new cycle is created when the file is updated and the existing cycles aged (for example, Cycle 01 is aged to Cycle 02 and Cycle 02 becomes Cycle 03).
Forecast cycle:	The number of cycles or versions of the forecast data you want to retain. A new cycle is created when you specify forecast aging in one of the workload forecasting routines and the existing cycles are aged (for example, Cycle 01 is aged to Cycle 02 and Cycle 02 becomes Cycle 03).
Database ID(s):	The database unit identifier(s) (up to fourteen units) that contain the DB2DSU, DB2DSD, and DB2DSY files to be used with this Sample application. If you are not sure where the DB2DSU, DB2DSD and DB2DSY files reside, type a question mark (?) in this field to display a listing of the valid database unit identifiers for this Sample application.
DB2 workload definition (CAPAPU):	The options for defining workload groups (see Section 2.11.1, Usage Guidelines).

Specify extended options: If you want extended options, enter Y (YES), and the Extended Options Menu (Figure 2-11.02) is displayed.

```

/----- Extended Options Menu -----\
|Command ==>
|
|Inquiry Step: DB2 Resource Planning Sample Application
|
|
|Combine all SYSIDs into a composite ==> N      (Y/N)
|  Composite SYSID                    ==> ----
|
|Combine all zones into a composite ==> N      (Y/N)
|  Composite zone                     ==> -----
|
|Specify:
|  Common data selection parameters ==> N      (Y/N/R)
|  Global data selection exit        ==> N      (Y/N/R)
|  Notes report enabled              ==> YES    (YES/NO)
|  Contents reports enabled          ==> YES    (YES/NO)
|
|
|
|
\-----/

```

Figure 2-11.02. DB2 Resource Planning Extended Options Menu

The extended options available with the DB2 Resource Planning Sample Application are as follows:

Combine all SYSIDs into a composite: You can combine SYSIDs in the capacity planning database file. Enter Y (YES) if you wish data for all of the SYSIDs to be summarized into a single observation with the Composite SYSID value.

Combine all zones/hours/daynames into a composite: You can combine zones, hours, and daynames in the capacity planning database file. The prompt for this field changes to zones or hours or daynames depending on what you specify for the summary level parameter. Enter Y (YES) if you want all of the zones, hours, and daynames to be summarized into a single observation with the composite value.

Specify common data selection parameters:	Specify the SYSIDs and zones, hours or daynames to select data from the CA MICS input file(s). Enter Y (YES) to display the Common Data Selection screen (Figure 2-11.03). Enter R (RESET) to reset selections you previously made to SYSID and zone, hour or dayname.
Specify global data selection exit:	You can use your own SAS statements to select data from the CA MICS input file(s). Enter Y (YES) to display the SAS System Statements screen (Figure 2-11.04). Enter R (RESET) to reset any selections you previously made to the default of no user-specified data selection.
Notes report enabled:	If you have entered notes associated with the capacity planning database file definition, a Yes in this field prints the note as part of file updating.

Contents reports enabled:	A Yes in this field prints the contents of the Capacity Planner database file as part of file updating.

```

----- Common Data Selection -----
|Command ==>
|
|Inquiry Step: DB2 Resource Planning Sample Application
|
|SYSID ==> _____
|          |_____
|
|Zone   ==> _ - - - - -                (1 - 9)
|
|Hour   ==> __ - __ - __ - __ - __    (0 - 23)
|
|Dayname==> _____
|
|
|
|
|

```

Figure 2-11.03. Common Data Selection Screen - DB2 Resource Planning

The Common Data Selection screen (Figure 2-11.03) allows you to define the following data selection criteria:

SYSID: The SYSIDs you want to include in your capacity Planning database file. You can specify a maximum of 15 one- to four-character SYSIDs. Leave this parameter blank to include data for all SYSIDs.

Zone: Specify up to nine zones to be included in the capacity planning database file. Leave this parameter blank to include data for all zones.

Hour: Specify the five hours or hour ranges to be included in the capacity planning database file. Selection of this parameter is valid only if you specify the DETAIL, DAYS, or WEEKS timespan on the DB2 Resource Planning screen (Figure 2-11.01). Leave this parameter blank to include data for all hours. You must enter the value to the left of the dash for singular hour entries.

Dayname: Specify up to five daynames to be included in the capacity planning database file. Leave blank to include all days. Input the first unique characters and the application will fill in the rest. Selection of this parameter is only valid when the input timespan is DETAIL or DAYS.

```

/----- SAS System Statements ----- COLUMNS 001 072--\
|Command ==>                               Scroll ==> PAGE |
|
|SAS Statements for the SELECT Macro
|-----|
|***** ***** TOP OF DATA *****|
|=MSG> SPECIFY SAS STATEMENTS TO SUBSET THE CAPACITY PlanNING DATA BASE FILE. |
|=NOTE= %MACRO SELECT; <=== PLEASE DO NOT PLACE LINES BEFORE THIS LINE.      |
|000001
|000002
|000003
|000004
|000005
|000006
|000007
|000008
|000009
|000010
|=NOTE= %MEND SELECT; <=== PLEASE DO NOT PLACE LINES AFTER THIS LINE.      |
|***** ***** BOTTOM OF DATA *****|
|
\-----/

```

Figure 2-11.04. SAS System Statements Screen - DB2 Resource Planning Global Data Selection

The SAS System Statements screen (Figure 2-11.04) lets you define your own data selection logic. You can code any valid SAS statements as described in the SAS manuals. The %MACRO SELECT and %MEND SELECT statements are already provided and field protected from modification.

Note: This is a PDF/EDIT screen. All PDF/EDIT commands are available.

2.11.3 Standard Output

The capacity planning database file generated by the DB2 Resource Planning Sample Application contains the following data elements:

- o Data elements common to all CA MICS Capacity Planner database files:

SYSID
ZONE or HOUR or DAYNAME
YEAR
MONTH or WEEK or MNTHDAY
DATE

- o Data elements unique to the DB2 Resource Planning Sample Application database file:

- Character format data elements

CAPAPU - DB2 Workload Group

- Numeric format data elements

ALLBFEXP - Total Getpage Caused Buffer Expansions
ALLGETPG - Total Getpage Requests
ALLPRFET - Total Prefetch Read Requests
ALLSYCIO - Total Synchronous Read I/Os
ALLWRKUP - Total Units of Work Updated
CPUOVRHD - Ratio of DB2 CPU Time to Attributable
DSDSABBP - Active Buffers In Buffer Pool
DSDMNRPG - Min Free Pages in Free Chain
DSDMXCTU - Max Pages Used for Cursor Table
DSDMXDBD - Max Pages Used for DBD
DSDMXDSO - Max Data Sets Open Concurrently
DSDMXPGE - Max Pages in EDM Pool
DSDMXRPG - Max Free Pages in Free Chain
DSDMXSKC - Max Pages Used for SKCT Table
DSDPAGE - No. Pages in EDM Pool
DSDSRFXS - Read Pagefixed W/O Real Storage
DSDSWFXS - Write Pagefixed W/O Real Storage
DSULEEM - No. Lock Escalations to Exclusive Mode
DSULESM - No. Lock Escalations to Shared Mode
DSUMXPL - Max. No. Page Locks Held
DSUSUSP - Suspend Count
DSUTIMO - Timeout Count
DSUWATTR - NO. Wait Trace Events Processed
TOTSTMTS - Total Statements Executed

- Time format data elements

CPUINDB2 - Total Attributable DB2 TCB
 DURATION - Recording Interval Time
 DSUALLTM - Lock & Latch Time
 DSUCPUTM - Accumulated CPU Time
 DSUELPTM - Elapsed Time
 DSUET2TM - Class 2 Accum. Elapsed Time in DB2
 DSUIOETM - DB2 I/O Time
 DSUTCBTM - TCB Time
 DSUTC2TM - Class 2 Accumulated TCB Time in DB2
 DSYSRTM0 - Accumulated SRB Time Addr0
 DSYSRTM1 - Accumulated SRB Time Addr1
 DSYSRTM2 - Accumulated SRB Time Addr2
 DSYTCTM0 - Accumulated TCB Time Addr0
 DSYTCTM1 - Accumulated TCB Time Addr1
 DSYTCTM2 - Accumulated TCB Time Addr2
 TOTCPSTM - Total DB2 System CPU Time
 TOTTCSTM - Total DB2 System TCB Time
 TOTSRSTM - Total DB2 System SRB Time

The following data elements used in the DB2 Planning Resource Sample Application database file have the same meaning and derivation as their identically named counterparts in the CA MICS files from which they are obtained.

DB2DSD: DSDSABBP, DSDMNRPG, DSDMXCTU, DSDMXDBD, DSDMXDSO,
 DSDMXPGE, DSDMXRPG, DSDMXSKC, DSDPAGE, DSDSRFXS
 DSDSWFXS

DB2DSU: DSUALLTM, DSUCPUTM, DSUELPTM, DSUET2TM, DSUIOETM,
 DSULEEM, DSULESM, DSUMXPL, DSUSUSP, DSUTCBTM,
 DSUTC2TM, DSUTIMO, DSUWATTR

DB2DSY: DSYSRTM0, DSYSRTM1, DSYSRTM2, DSYTCTM0, DSYTCTM1,
 DSYTCTM2

The following data elements are derived from the CA MICS database and used by the DB2 Resource Planning Sample Application:

ALLGETPG: The total number of GETPAGE requests issued by users to all buffer pools. It is calculated using the following formula:

$$\text{ALLGETPG} = \text{DSUSGPGR}$$

where DSUSGPGR is a summarized data element from the DB2DSU file quantifying the number of GETPAGE requests issued for all buffer pools.

ALLPRFET: The total number of Media Manager pre-fetch read requests to all buffer pools. It is calculated using the following formula:

$$\text{ALLPRFET} = \text{DSUSPRFH}$$

where DSUSPRFH is a summarized data element from the DB2DSU file quantifying the number of pre-fetch read requests issued for all buffer pools.

ALLSYCIO: The total number of DB2 Media Manager synchronous read requests performed for all users by all buffer pools. It is calculated using the following formula:

$$\text{ALLSYCIO} = \text{DSUSSRIO}$$

where DSUSSRIO is a summarized data element from the DB2DSU file quantifying the number of synchronous read requests performed by all buffer pools.

CPUINDB2: The total CPU time attributable to users of DB2. It is the Class 2 TCB time obtained from the DSUTC2TM data elements.

CPUOVRHD: The ratio of total DB2 system CPU time to the CPU time that is attributable to users of DB2. It is calculated using the following formula:

$$\text{CPUOVRHD} = \text{TOTCPSTM} / \text{CPUINDB2}$$

TOTCPSTM: The total DB2 system CPU time. It is the sum of the CPU time recorded for each address space: 0, 1, and 2, representing the Systems Service, Database Services, and the IMS Resource Lock Manager address spaces respectively. It is calculated using the following formula:

$$\text{TOTCPSTM} = \text{SUM}(\text{TOTTCSTM}, \text{TOTSRSTM})$$

TOTSRSTM: The total DB2 system SRB time. It is the sum of the SRB time recorded for each address space: 0, 1, and 2, representing the Systems Service, Database Services, and the IMS Resource Lock Manager address spaces respectively. It is calculated using the following formula:

$$\text{TOTSRSTM} = \text{SUM}(\text{DSYSRTM0}, \text{DSYSRTM1}, \text{DSYSRTM2})$$

TOTTCSTM: The total DB2 system TCB time. It is the sum of the TCB time recorded for each address space: 0, 1, and 2, representing the Systems Service, Database Services, and the IMS Resource Lock Manager address spaces respectively. It is calculated using the following formula:

$$\text{TOTTCSTM} = \text{SUM}(\text{DSYTCTM0}, \text{DSYTCTM1}, \text{DSYTCTM2})$$

TOTSTMTS: The total number of DB2 statements executed. It is the sum of the data elements in the DSU file containing statement counts (for example, DSUALTIX - Alter Index Statements, DSUALTST - Alter Stogroup Statements, and so on.) It is calculated using the following formula:

$$\begin{aligned} \text{TOTSTMTS} = \text{SUM}(\text{DSUINSRT}, \text{DSUUPDTE}, \text{DSUDELET}, \\ \text{DSUDESC}, \text{DSUPREP}, \text{DSUOPEN}, \\ \text{DSUCLOSE}, \text{DSUCRTAB}, \text{DSUCRINX}, \\ \text{DSUCTABS}, \text{DSUCRSYN}, \text{DSUCRDAB}, \\ \text{DSUCRSTG}, \text{DSUDEFVU}, \text{DSUDRPIX}, \\ \text{DSUDRPTA}, \text{DSUDRPTS}, \text{DSUDRPDB}, \\ \text{DSUDRPSY}, \text{DSUDRPST}, \text{DSUDRPVU}, \\ \text{DSUALTST}, \text{DSUALTTS}, \text{DSUALTTA}, \\ \text{DSUALTIX}, \text{DSUCMTON}, \text{DSULOCK}, \\ \text{DSUFETCH}, \text{DSULABEL}, \text{DSUSELCT}, \\ \text{DSUGRANT}, \text{DSUREVOK}); \end{aligned}$$

2.12 WebSphere MQSeries Planning

The WebSphere MQSeries Planning Sample Application provides a capacity planning database file that contains the necessary elements for MQSeries workload tracking and reporting.

The following sections cover the use, input specifications, and output of the sample application:

- 1 - Usage Guidelines
- 2 - Control Parameters
- 3 - Standard Output

2.12.1 Usage Guidelines

The WebSphere MQSeries Planning Sample Application provides a capacity planning database file with the database elements you need for basic tracking, reporting, and forecasting of the MQSeries workloads on your system. The information for this application is primarily derived from either the DETAIL, DAYS, WEEKS, or MONTHS timespan of the CA MICS WebSphere MQSeries Message Manager Accounting File (MQAMMA).

In the WebSphere MQSeries Planning Sample Application, the application workload group (CAPAPU) data element enables you to group your individual MQSeries systems into logically similar workload groups for more meaningful tracking and reporting. Three methods are provided for specifying the manner in which the MQSeries applications are grouped: 1) defining each MQSeries application as its own workload group (CAPAPU = MQMSSI), 2) combining all MQSeries applications (no workload group), 3) defining your own SAS statements to create the workload groupings by setting the CAPAPU values.

2.12.2 Control Parameters

Relatively few parameters are needed to implement the WebSphere MQSeries Planning Sample Application. The control screens included in the online facility allow you to quickly enter the required information.

```

/----- WebSphere MQSeries Planning -----\
|Command ==>
|Enter a ? in any data entry field for more information on valid values.
|Composing CAP DB File def.: MQ1 - WebSphere MQSeries Planning Sample Appl.
|
|Summarization Level
|      Primary          ==> MONTHS  (DAYS/WEEKS/MONTHS)
|      Secondary        ==> ZONE    (ZONE/HOUR/DAYNAME)
|
|Input File Time-Span    ==> MONTHS  (DETAIL/DAYS/WEEKS/MONTHS)
|
|Amount of data retained ==> 60      (1-9999 months)
|Age of oldest forecast  ==> 24      (1-9999 months)
|History cycles          ==> 3       (1-99 cycles)
|Forecast cycles         ==> 3       (1-99 cycles)
|
|database ID (s)         ==> _____
|
|MQSeries Workload Definition (CAPAPU) ==> MQMSSI (MQMSSI/NONE/EXIT)
|      Application unit title ==> _____
|
|Specify extended options ==> N      (Y/N)
\-----/

```

Figure 2-12.01. MQSeries Planning Screen

Figure 2-12.01 shows the MQSeries Planning screen, which is the main control screen for the WebSphere MQSeries Planning Sample Application. It contains the following fields:

Summarization level:	The output summarization corresponds to the combination of the primary and secondary summarization levels. These control how the output file is summarized in your capacity Planning database.
Input file timespan:	The timespan (either DETAIL, DAYS, WEEKS or MONTHS) of the MQAMMA Files that you want to use as input when creating the capacity planning database file.

Amount of data retained:	The number of days, weeks or months of historical data to be retained in the CAPxxx01 cycle of the capacity planning database file. Any historical observations older than what you specify for this parameter are dropped from the file during the file update.
Age of oldest forecast:	The number of days, weeks or months that you want to retain a forecast after it is generated. Any forecast older than what you specify for this parameter is dropped when a new forecast is generated and saved.
History cycle:	The number of cycles or versions of the historical data you want to retain. A new cycle is created when the file is updated and the existing cycles aged (for example, Cycle 01 is aged to Cycle 02 and Cycle 02 becomes Cycle 03).
Forecast cycle:	The number of cycles or versions of the forecast data you want to retain. A new cycle is created when you specify forecast aging in one of the workload forecasting routines and the existing cycles are aged (for example, Cycle 01 is aged to Cycle 02 and Cycle 02 becomes Cycle 03).
Database ID(s):	The database unit identifier(s) (up to fourteen units) that contain the MQAMMA file to be used with this Sample application. If you are not sure where the MQAMMA file resides, type a question mark (?) in this field to display a listing of the valid database units for this Sample application.
MQSeries Workload Definition (CAPAPU):	The three options for defining workload groups (see section 2.12.1, Usage Guidelines).
Application unit title:	The 1- to 40-character long name for the CAPAPU data element. Use this parameter to describe your workload groupings. This parameter is only available if you specify EXIT for the MQSeries workload definition (CAPAPU) parameter.

Specify If you want extended options, enter Y (YES), and
 extended the Extended Options Menu (Figure 2-12.02) is
 options: displayed.

```

/----- Extended Options Menu -----\
|Command ==>
|
|Inquiry Step: WebSphere MQSeries Planning Sample Application
|
|
|Combine all SYSIDs into a composite ==> N      (Y/N)
|  Composite SYSID                    ==> ----
|
|Combine all zones into a composite ==> N      (Y/N)
|  Composite zone                     ==> -----
|
|Specify:
|  Common data selection parameters ==> N      (Y/N/R)
|  Global data selection exit        ==> N      (Y/N/R)
|  CAPAPU derivation SAS statements ==> N      (Y/N/R)
|  Notes report enabled              ==> YES    (YES/NO)
|  Contents reports enabled          ==> YES    (YES/NO)
|
|
|
|
\-----/

```

Figure 2-12.02. MQSeries Planning Extended Options Menu

The extended options available with the WebSphere MQSeries Planning Sample Application (Figure 2-12.02) are:

Combine all You can combine SYSIDs in the capacity planning
 SYSIDs into database file. Enter Y (YES) if you want data
 a composite: for all of the SYSIDs to be summarized into a
 single observation with the Composite SYSID
 value.

Combine all You can combine zones, hours, and daynames in
 zones/hours/ the capacity planning database file. The prompt
 daynames for this field changes to zones or hours or
 into a daynames depending on what you specify for the
 composite: summary level parameter. Enter Y (YES) if you
 want all of the zones, hours, and daynames to be
 summarized into a single observation with the
 composite value.

Specify common data selection parameters:	Specify the SYSIDs and zones, hours or daynames to be used for data selection from the CA MICS input file(s). Enter Y (YES) to display the Common Data Selection screen (Figure 2-12.03). Enter R (RESET) to reset any selections you previously made to SYSID and zone, hour or dayname.
Specify global data selection	You can use your own SAS statements to select data from the CA MICS input file(s). Enter Y (YES) to display the SAS System Statements screen (Figure 2-12.04). Enter R (RESET) to reset any selections you previously made to the default of no user-specified data selection.
Specify CAPAPU derivation SAS statements:	You can use your own SAS statements for defining your MQSeries workloads. Enter Y (YES) to display the SAS System Statements screen. Enter R (RESET) to reset any SAS SYSTEM STATEMENTS you previously made to a user-specified workload grouping. This option is available only if you specify EXIT for the MQSeries Workload Definition (CAPAPU) parameter on the MQSeries Planning screen.
Notes report enabled:	A Yes in this field prints any notes associated with the capacity planning file definition as part of file updating.

- Hour:** Specify the five hours or hour ranges to be included in the capacity planning database file. Selection of this parameter is valid only if you specify the DETAIL, DAYS, or WEEKS timespan on the MQSeries Planning screen (Figure 2-12.01). Leave this parameter blank to include data for all hours.
- Dayname:** Specify up to five daynames to be included in the capacity planning database file. Leave blank to include all days. Input the first unique characters and the application will fill in the rest. Selection of this parameter is only valid when the input timespan is DETAIL or DAYS.

```

/----- SAS System Statements ----- COLUMNS 001 072--\
|Command ==>                               Scroll ==> PAGE |
|
|SAS Statements for the SELECT Macro          |
|-----|
|***** ***** TOP OF DATA *****|
|=MSG> SPECIFY SAS STATEMENTS TO SUBSET THE CAPACITY PLANNING DATA BASE FILE. |
|=NOTE= %MACRO SELECT; <=== PLEASE DO NOT PLACE LINES BEFORE THIS LINE.      |
|000001                                     |
|000002                                     |
|000003                                     |
|000004                                     |
|000005                                     |
|000006                                     |
|000007                                     |
|000008                                     |
|000009                                     |
|000010                                     |
|=NOTE= %MEND SELECT; <=== PLEASE DO NOT PLACE LINES AFTER THIS LINE.        |
|***** ***** BOTTOM OF DATA *****|
|
\-----/

```

Figure 2-12.04. SAS System Statements Screen - MQSeries Planning Global Data

The SAS System Statements screen (Figure 2-12.04) lets you define your own data selection logic. You can code any valid SAS statements as described in the SAS manuals. The %MACRO SELECT and %MEND SELECT statements are provided and are field protected from modification.

Note: This is a PDF/EDIT screen. All the PDF/EDIT commands are available.

2.12.3 Standard Output

The capacity planning database file that is generated by the WebSphere MQSeries Planning Sample Application contains the following data elements.

- A. Data elements common to all CA MICS Capacity Planner database files:

SYSID
 ZONE or HOUR or DAYNAME
 YEAR
 MONTH or WEEK or MNTHDAY
 DATE

- B. Data elements unique to the WebSphere MQSeries Planning Sample Application database file:

Note: For a detailed description of the data elements available for MQSeries platforms supported by CA MICS, see the CA MICS Analyzer Option for MQSeries Guide.

1. Character format data elements

CAPAPU - MQMSSI - Workload Group

2. Numeric format data elements

ALLGETA - Total MQGET Messages 0-99 Bytes
 ALLGETB - Total MQGET Messages 100-999 Bytes
 ALLGETC - Total MQGET Messages 1000-9999
 ALLGETD - Total MQGET For Messages ≥ 10000
 ALLPUTA - Total MQPUT For Messages 0-99 Bytes
 ALLPUTB - Total MQPUT For Messages 100-999 Bytes
 ALLPUTC - Total MQPUT For Messages 1000-9999
 ALLPUTD - Total MQPUT For Messages ≥ 10000
 MMACPUNI - Job instructions executed
 MMAGETA - MGET Req. for messages 0-99 bytes
 MMAGETB - MGET Req. for messages 100-999 bytes
 MMAGETC - MGET Req. for messages 1000-9999 bytes
 MMAGETD - MGET Req. for messages ≥ 10000 bytes
 MMAPUTA - MPUT Req. for messages 0-99 bytes
 MMAPUTB - MPUT Req. for messages 100-999 bytes
 MMAPUTC - MPUT Req. for messages 1000-9999 bytes
 MMAPUTD - MPUT Req. for messages ≥ 10000 bytes
 MMATMGET - Total number of MQGET messages
 MMATMPUT - Total number of MQPUT messages

3. Time format data elements

MMACPUTM - TCB/SRB CPU Time used

The data elements that are derived from the CA MICS database and used by the WebSphere MQSeries Planning Sample Application are calculated as follows:

CAPAPU:	The data element that contains the values found, which will be either MQMSSI, values derived within the user's exit, or no CAPAPU value.
ALLGETA:	Total MQGET Req. messages of size 0-99 bytes for this interval.
ALLGETB:	Total MQGET Req. messages of size 100-999 bytes for this interval.
ALLGETC:	Total MQGET Req. messages of size 1000-9999 bytes for this interval.
ALLGETD:	Total MQGET Req. messages of size >= 10000 bytes for this interval.
ALLMQCPU:	Total MQSeries CPU time (TCB/SRB) for this interval.
ALLPUTA:	Total MQPUT Req. messages of size 0-99 bytes for this interval.
ALLPUTB:	Total MQPUT Req. messages of size 100-999 bytes for this interval.
ALLPUTC:	Total MQPUT Req. messages of size 1000-9999 bytes for this interval.
ALLPUTD:	Total MQPUT Req. messages of size >= 10000 bytes for this interval.

2.13 WEB Server Resource Planning

The WEB Server Resource Planning Sample Application provides a capacity planning database file that contains the necessary elements for WEB Server workload tracking and reporting.

The following sections cover the use, input specifications, and output of the sample application:

- 1 - Usage Guidelines
- 2 - Control Parameters
- 3 - Standard Output

2.13.1 Usage Guidelines

The WEB Server Resource Planning Sample Application provides a capacity planning database file with the database elements you need for basic tracking, reporting, and forecasting of the WEB Server workloads on your system. The information for this application is primarily derived from either the DETAIL, DAYS, WEEKS, or MONTHS timespan of the CA MICS WebSphere Performance file (WEBWPR).

In the WEB Server Resource Planning Sample Application, the application workload group (CAPAPU) data element enables you to group your individual WEB Server systems into logically similar workload groups for more meaningful tracking and reporting. Three methods are provided for specifying the manner in which the WEB Server applications are grouped: 1) defining each WEB Server application as its own workload group (CAPAPU = WEBNAME), 2) combining all WEB Server applications (no workload group), 3) defining your own SAS statements to create the workload groupings by specifying CAPAPU values.

2.13.2 Control Parameters

Relatively few parameters are needed to implement the WEB Server Resource Planning Sample Application. The control screens included in the online facility allow you to quickly enter the required information.

```

/----- WEB Server Resource Planning -----\
|Command ==>|
|Enter a ? in any data entry field for more information on valid values.|
|Composing CAP DB File def.:WS1-WEB Server Resource Planning Sample Application|
|
|Summarization Level|
|    Primary        ==> MONTHS  (DAYS/WEEKS/MONTHS)|
|    Secondary      ==> ZONE    (ZONE/HOUR/DAYNAME)|
|
|Input File Time-Span    ==> MONTHS  (DETAIL/DAYS/WEEKS/MONTHS)|
|
|Amount of data retained    ==> 60    (1-9999 months)|
|Age of oldest forecast     ==> 24    (1-9999 months)|
|History cycles             ==> 3     (1-99 cycles)|
|Forecast cycles            ==> 3     (1-99 cycles)|
|
|database ID (s)           ==> _____|
|
|WEB Server Workload Definition (CAPAPU) ==> WEBNAME (WEBNAME/NONE/EXIT)|
|    Application unit title ==> _____|
|
|Specify extended options   ==> N      (Y/N)|
\-----/

```

Figure 2-13.01. WEB Server Planning Screen

Figure 2-13.01 shows the WEB Server Planning screen. This is the main control screen for the WEB Server Resource Planning Sample Application, and contains the following fields:

Summarization level:	The output summarization corresponds to the combination of the primary and secondary summarization levels. These control how the output file is summarized in your capacity planning database.
Input file timespan:	The timespan (either DETAIL, DAYS, WEEKS or MONTHS) of the WEBWPR Files that you want to use as input when creating the capacity planning database file.

Amount of data retained:	The number of days, weeks or months of historical data to retain in the CAPxxx01 cycle of the capacity planning database file. Any historical observations older than what you specify for this parameter are dropped from the file during the file update.
Age of oldest forecast:	The number of days, weeks or months that you want to retain a forecast after it is generated. Any forecast older than what you specify for this parameter is dropped when a new forecast is generated and saved.
History cycle:	The number of cycles or versions of the historical data you want to retain. A new cycle is created when the file is updated and the existing cycles aged (for example, Cycle 01 is aged to Cycle 02 and Cycle 02 becomes Cycle 03).
Forecast cycle:	The number of cycles or versions of the forecast data you want to retain. A new cycle is created when you specify forecast aging in one of the workload forecasting routines and the existing cycles are aged (for example, Cycle 01 is aged to Cycle 02 and Cycle 02 becomes Cycle 03).
Database ID(s):	The database unit identifier(s) (up to 14 units) that contain the WEBWPR file used with this sample application. If you are not sure where the WEBWPR file resides, type a question mark (?) in this field to display a listing of the valid database unit identifiers for this sample application.
WEB Server Workload Definition (CAPAPU):	The three options for defining workload groups (see section 2.13.1, Usage Guidelines).
Application unit title:	The 1- to 40-character name for the CAPAPU data element. You can use this parameter to describe your workload groupings. This parameter is only available if you specify EXIT for the WEB Server workload definition (CAPAPU) parameter.

Specify If you want extended options, specify Y (YES),
 extended and the Extended Options Menu (Figure 2-13.02)
 options: appears.

```

/----- Extended Options Menu -----\
|Command ==>
|
|Inquiry Step:  WEB Server Resource Planning Sample Application
|
|
|Combine all SYSIDs into a composite ==> N          (Y/N)
|  Composite SYSID                    ==>  ____
|
|Combine all zones into a composite ==> N          (Y/N)
|  Composite zone                     ==>  _____
|
|Specify:
|  Common data selection parameters ==> N          (Y/N/R)
|  Global data selection exit        ==> N          (Y/N/R)
|  CAPAPU derivation SAS statements ==> N          (Y/N/R)
|  Notes report enabled               ==> YES       (YES/NO)
|  Contents reports enabled          ==> YES       (YES/NO)
|
|
|
|
\-----/

```

Figure 2-13.02. WEB Server Planning Extended Options Menu

The extended options available with the WEB Server Resource Planning Sample Application (Figure 2-13.02) are:

Combine all SYSIDs into a composite: You can combine SYSIDs in the capacity planning database file. Enter Y (YES) if you want data for all of the SYSIDs to be summarized into a single observation with the Composite SYSID value.

Combine all zones/hours/daynames into a composite: You can combine zones, hours, and daynames in the capacity planning database file. The prompt for this field changes to zones, hours or daynames depending on what you specify for the summary level parameter. Enter Y (YES) if you want all of the zones, hours, and daynames summarized into a single observation with the composite value.

Specify common data selection parameters:	Specify the SYSIDs and zones, hours or daynames to use for data selection from the CA MICS input file(s). Enter Y (YES) to display the Common Data Selection screen (Figure 2-13.03). Enter R (RESET) to reset any selections you previously made to SYSID and zone, hour or dayname.
Specify global data selection	You can specify your own SAS statements to select data from the CA MICS input file(s). Enter Y (YES) to display the SAS System Statements screen (Figure 2-13.04). Enter R (RESET) to reset any selections you previously made to the default of no user-specified data selection.
Specify CAPAPU derivation SAS statements:	You can specify your own SAS statements for defining your WEB Server workloads. Enter Y (YES) to display the SAS System Statements screen. Enter R (RESET) to reset any SAS System Statements you previously made to a user-specified workload grouping. This option is available only if you specify EXIT for the WEB Server Workload Definition (CAPAPU) parameter on the WEB Server Planning screen.
Notes report enabled:	A Yes in this field prints any notes associated with the capacity planning database file definition as part of file updating.

Contents reports enabled:	A Yes in this field prints the contents of the capacity planning database file as part of file updating.

```

/----- Common Data Selection -----\
|Command ==>
|
|Inquiry Step:  WEB Server Resource Planning Sample Application
|
|SYSID ==>  _____
|           _____
|
|Zone ==>  _ _ _ _ _ (1 - 9)
|
|Hour ==>  __ - __ - __ - __ - __ - __ - __ - __ (0 - 23)
|
|Dayname==>  _____
|
|
|
|
\-----\

```

Figure 2-13.03. Common Data Selection Screen - WEB Server Planning

The Common Data Selection screen (Figure 2-13.03) lets you define the following data selection criteria:

SYSID: The SYSIDs that you want to include in your capacity planning database file. You can specify a maximum of 15 one- to four-character SYSIDs. Leave this parameter blank to include data for all SYSIDs.

Zone: Specify up to nine zones to be included in the capacity planning database file. Leave this parameter blank to include data for all zones.

Hour: Specify the five hours or hour ranges to be included in the capacity planning database file. Selection of this parameter is valid only if you specify the DETAIL, DAYS, or WEEKS timespan on the WEB Server Planning (Figure 2-13.03). Leave this parameter blank to include data for all hours.

Dayname: Specify up to five daynames to be included in the capacity planning database file. To include all days, leave this field blank. Input the first unique characters and the application will fill in the rest. Selection of this parameter is only valid when the input timespan is DETAIL or DAYS.

```

/----- SAS System Statements ----- COLUMNS 001 072--\
|Command ==>                               Scroll ==> PAGE |
|
|SAS Statements for the SELECT Macro          |
|-----|
|***** ***** TOP OF DATA *****|
|=MSG> SPECIFY SAS STATEMENTS TO SUBSET THE CAPACITY PLANNING DATA BASE FILE. |
|=NOTE= %MACRO SELECT; <=== PLEASE DO NOT PLACE LINES BEFORE THIS LINE.      |
|000001                                     |
|000002                                     |
|000003                                     |
|000004                                     |
|000005                                     |
|000006                                     |
|000007                                     |
|000008                                     |
|000009                                     |
|000010                                     |
|=NOTE= %MEND SELECT; <=== PLEASE DO NOT PLACE LINES AFTER THIS LINE.        |
|***** ***** BOTTOM OF DATA *****|
|
\-----/

```

Figure 2-13.04. SAS System Statements Screen - WEB Server Planning Global Data

The SAS System Statements screen (Figure 2-13.04) lets you define your own data selection logic. You can code any valid SAS statement as described in the SAS manuals. The %MACRO SELECT and %MEND SELECT statements are already provided and are field protected from modification.

Note: This is a PDF/EDIT screen. All the PDF/EDIT commands are available.

2.13.3 Standard Output

The capacity planning database file that is generated by the WEB Server Resource Planning Sample Application contains the following data elements.

- A. Data elements common to all CA MICS Capacity Planning database files:

SYSID
ZONE or HOUR or DAYNAME
YEAR
MONTH or WEEK or MNTHDAY
DATE

- B. Data elements unique to the WEB Server Resource Planning Sample Application database file:

Note: For a detailed description of the data elements available for WEB Server platforms supported by CA MICS, see the CA MICS Web Analyzer Option Guide.

- 1. Character format data elements

CAPAPU - WEBNAME - Workload Group

- 2. Numeric format data elements

WPRREQR - No. Requests Received
WPRRSPS - No. Responses Sent
WPRBYTR - No. Bytes Received
WPRBYTS - No. Bytes Sent

The data elements that are derived from the CA MICS database and used by the WEB Server Resource Planning Sample Application are calculated as follows:

CAPAPU: The data element that contains the values found will be either WEBNAME, values derived within the user's exit, or no CAPAPU value.

2.14 TCP/IP Network Resource Planning

The TCP/IP Network Resource Planning Sample Application provides a capacity planning database file that contains the necessary elements for TCP/IP resource tracking and reporting.

The following sections cover the use, input specifications, and output of the sample application:

- 1 - Usage Guidelines
- 2 - Control Parameters
- 3 - Standard Output

2.14.1 Usage Guidelines

The TCP/IP Network Resource Planning Standard Application provides a capacity planning database file with the database elements you need for basic tracking, reporting, and forecasting of the TCP/IP Network workloads on your system. The information for this application is primarily derived from either the DETAIL, DAYS, WEEKS, or MONTHS timespan of the CA MICS TCP/IP API Calls for Sockets File (TCPAPI).

In the TCP/IP Network Resource Planning Standard Application, the application workload group (CAPAPU) data element lets you group your individual TCP/IP systems into logically similar workload groups for more meaningful tracking and reporting. Three methods are provided for specifying the manner in which the TCP/IP applications are grouped: 1) defining each TCP/IP application as its own workload group (CAPAPU = NETNAME), 2) combining all TCP/IP applications (no workload group), 3) defining your own SAS statements to create the workload groupings by specifying CAPAPU values.

It may be important to identify specific remote addresses (TCPADDR) within your TCP/IP network. The TCP/IP remote address can be used to subset your TCP/IP volume to further delineate your TCP/IP volume statistics for planning purposes.

The TCP/IP TCPAPI file contains an element called NETWRKID. This element is populated from the SNTNIDRT exit contained within unit prefix.parms.

Consider the following CA MICS exit example:

```
/* */
/* Sample unit prefix.parms(SNTNIDRT) exit */
/* */
/* DCAHRND1 listed below can potentially have TCP/IP */
/* addresses from 141.202.195.0 to 141.202.195.255. */
/* */
/* TCPRADDR is stored as a character. */
/* */
IF TCPRADDR =: '141.202.195.' and ORGSYSID='PRD1'
  THEN
    NETWRKID='DCAHRND1'; /* Herndon Activity DCA1 */
  ELSE
    IF ORGSYSID='PRD1'
      THEN
        NETWRKID='DCAPROD1'; /* All other SYSID DCAPROD1 */
      ELSE
        IF TCPRADDR =: '141.202.255.' AND ORGSYSID='TST3'
          THEN
            NETWRKID='NYCNJ1'; /* New Jersey Activity */
          ELSE
            IF ORGSYSID='TST3'
              THEN
                NETWRKID='NYCPROD1'; /* All other SYSID NYCPROD1 */
              ELSE
                NETWRKID='*UNKNOWN';
            /* */
```

The above exit example subsets the TCP/IP volume sent to the remote addresses for New Jersey and Herndon. If necessary, the data can also be easily recombined during subsequent reporting for all of the SYSID by adding the statement: "If NETWRKID=: 'DCA' then NETWRKID='DCAPROD1' ".

The NETWRKID provides a method for identifying the volume of traffic within a given TCP/IP system.

2.14.2 Control Parameters

Relatively few parameters are needed to implement the TCP/IP Network Resource Planning Sample Application. The control screens included in the online facility allow you to quickly enter the required information.

```

/----- TCP/IP Network Resource Planning -----\
|Command ==>
|Enter a ? in any data entry field for more information on valid values.
|Composing CAP DB File def.: TP1 - TCP/IP Network Resource Plan. Sample Appl.
|
|Summarization Level
|      Primary          ==> MONTHS  (DAYS/WEEKS/MONTHS)
|      Secondary        ==> ZONE    (ZONE/HOUR/DAYNAME)
|
|Input File Time-Span    ==> MONTHS  (DETAIL/DAYS/WEEKS/MONTHS)
|
|Amount of data retained ==> 60      (1-9999 months)
|Age of oldest forecast  ==> 24      (1-9999 months)
|History cycles          ==> 3       (1-99 cycles)
|Forecast cycles         ==> 3       (1-99 cycles)
|
|database ID (s)         ==> _____
|
|TCP/IP Workload Definition (CAPAPU) ==> NETNAME (NETNAME/NONE/EXIT)
|      Application unit title ==> _____
|
|Specify extended options ==> N      (Y/N)
\-----/

```

Figure 2-14.01. TCP/IP Network Planning Screen

Figure 2-14.01 shows the TCP/IP Network Planning screen, which is the main control screen for the TCP/IP Network Planning Sample Application. It contains the following fields:

Summarization level: The output summarization corresponds to the combination of the primary and secondary summarization levels. These control how the output file is summarized in your capacity planning database.

Input file timespan: The timespan (either DETAIL, DAYS, WEEKS or MONTHS) of the TCPAPI Files that you want to use as input when creating the capacity planning database file.

Amount of data retained:	The number of days, weeks or months of historical data to retain in the CAPxxx01 cycle of the capacity planning database file. Any historical observations older than what you specify for this parameter are dropped from the file during the file update.
Age of oldest forecast:	The number of days, weeks or months that you want to retain a forecast after it is generated. Any forecast older than what you specify for this parameter is dropped when a new forecast is generated and saved.
History cycle:	The number of cycles or versions of the historical data you want to retain. A new cycle is created when the file is updated and the existing cycles aged (for example, Cycle 01 is aged to Cycle 02 and Cycle 02 becomes Cycle 03).
Forecast cycle:	The number of cycles or versions of the forecast data you want to retain. A new cycle is created when you specify forecast aging in one of the workload forecasting routines and the existing cycles are aged (for example, Cycle 01 is aged to Cycle 02 and Cycle 02 becomes Cycle 03).
Database ID(s):	The database unit identifier(s) (up to 14 units) that contain the TCPAPI file used with this Sample application. If you are not sure where the TCPAPI file resides, type a question mark (?) in this field to display a listing of the valid database unit identifiers for this sample application.
TCP/IP Workload Definition (CAPAPU):	The three options for defining workload groups (see section 2.14.1, Usage Guidelines).
Application unit title:	The 1- to 40-character name for the CAPAPU data element. You can use this parameter to describe your workload groupings. This parameter is only available if you specify EXIT for the TCP/IP Network workload definition (CAPAPU) parameter.


```

Specify      If you want extended options, specify Y (YES),
extended     and the Extended Options Menu (Figure 2-14.02)
options:     appears.

/----- Extended Options Menu -----\
|Command ==>
|
|Inquiry Step: TCP/IP Network Resource Planning Sample Application
|
|
|Combine all SYSIDs into a composite ==> N      (Y/N)
|  Composite SYSID                    ==> ----
|
|Combine all zones into a composite ==> N      (Y/N)
|  Composite zone                     ==> -----
|
|Specify:
|  Common data selection parameters ==> N      (Y/N/R)
|  Global data selection exit        ==> N      (Y/N/R)
|  CAPAPU derivation SAS statements ==> N      (Y/N/R)
|  Notes report enabled              ==> YES    (YES/NO)
|  Contents reports enabled          ==> YES    (YES/NO)
|
|
|
|
|
|-----\

```

Figure 2-14.02. TCP/IP Network Planning Extended Options Menu

The extended options available with the TCP/IP Network Resource Planning Sample Application (Figure 2-14.02) are:

Combine all SYSIDs into a composite: You can combine SYSIDs in the capacity planning database file. Enter Y (YES) if you wish data for all of the SYSIDs to be summarized into a single observation with the Composite SYSID value.

Combine all zones/hours/daynames into a composite: You can combine zones, hours, and daynames in the capacity planning database file. The prompt for this field changes to zones or hours or daynames depending on what you specify for the summary level parameter. Enter Y (YES) if you want all of the zones, hours, and daynames summarized into a single observation with the composite value.

Specify common data selection parameters:	Specify the SYSIDs and zones, hours or daynames to use for data selection from the CA MICS input file(s). Enter Y (YES) to display the Common Data Selection screen (Figure 2-14.03.) Enter R (RESET) to reset any selections you previously made to SYSID and zone, hour or dayname.
Specify global data selection	You can use your own SAS statements to select data from the CA MICS input file(s). Enter Y (YES) to display the SAS System Statements screen (Figure 2-14.04). Enter R (RESET) to reset any selections you previously made to the default of no user-specified data selection.
Specify CAPAPU derivation SAS statements:	You can use your own SAS statements for defining your TCP/IP workloads. Enter Y (YES) to display the SAS System Statements screen. Enter R (RESET) to reset any SAS System Statements you previously made to a user-specified workload grouping. This option is available only if you specify EXIT for the TCP/IP Workload Definition (CAPAPU) parameter on the TCP/IP Network Planning screen.
Notes report enabled:	A Yes in this field prints any notes associated with the capacity planning file definition as part of file updating.

Hour: Specify the five hours or hour ranges to be included in the capacity planning database file. Selection of this parameter is valid only if you specify the DETAIL, DAYS, or WEEKS timespan on the TCP/IP Network Planning (Figure 2-14.01). Leave this parameter blank to include data for all hours.

Dayname: Specify up to five daynames to be included in the capacity planning database file. To include all days, leave this field blank. Input the first unique characters and the application will fill in the rest. Selection of this parameter is only valid when the input timespan is DETAIL or DAYS.

```

/----- SAS System Statements ----- COLUMNS 001 072--\
|Command ==>                               Scroll ==> PAGE |
|
|SAS Statements for the SELECT Macro
|-----|
|***** ***** TOP OF DATA *****|
|=MSG> SPECIFY SAS STATEMENTS TO SUBSET THE CAPACITY PLANNING DATA BASE FILE. |
|=NOTE= %MACRO SELECT; <=== PLEASE DO NOT PLACE LINES BEFORE THIS LINE.      |
|000001
|000002
|000003
|000004
|000005
|000006
|000007
|000008
|000009
|000010
|=NOTE= %MEND SELECT; <=== PLEASE DO NOT PLACE LINES AFTER THIS LINE.        |
|***** ***** BOTTOM OF DATA *****|
|
|-----/

```

Figure 2-14.04. SAS System Statements Screen - TCP/IP Network Plan. Global Data

The SAS System Statements screen (Figure 2-14.04) lets you define your own data selection logic. You can code any valid SAS statement as described in the SAS manuals. The %MACRO SELECT and %MEND SELECT statements are already provided, and are field protected to prevent modification.

Note: This is a PDF/EDIT screen. All the PDF/EDIT commands are available.

2.14.3 Standard Output

The capacity planning database file that is generated by the TCP/IP Network Resource Planning Sample Application contains the following data elements.

- A. Data elements common to all CA MICS Capacity Planner database files:

SYSID
ZONE or HOUR or DAYNAME
YEAR
MONTH or WEEK or MNTHDAY
DATE

- B. Data elements unique to the TCP/IP Network Resource Planning Sample Application database file:

Note: For a detailed description of the data elements available for TCP/IP Network platforms supported by CA MICS, see the CA MICS SNA Analyzer Guide.

1. Character format data elements

CAPAPU - NETNAME - Workload Group
PRODUCT - SNA Network Product (Source) Name

2. Numeric format data elements

ALLBYTES - Total Bytes
APIBYTEI - Bytes In
APIBYTEO - Bytes Out

The data elements that are derived from the CA MICS database and used by the TCP/IP Network Resource Planning Sample Application are calculated as follows:

CAPAPU: The data element that contains the values found will be either NETNAME (NETWRKID), values derived within the user's exit, or no CAPAPU value.

ALLBYTES: Total of bytes in (APIBYTEI) and bytes out (APIBYTEO).

2.15 Tandem CPU Resource Planning

The Tandem CPU Resource Planning Sample Application provides a capacity planning database file that contains the necessary elements for tracking and reporting CPU utilization in a Tandem environment.

The following sections cover the use, input specifications, and output of the sample application:

- 1 - Usage Guidelines
- 2 - Control Parameters
- 3 - Standard Output
- 4 - Data Elements Overview

2.15.1 Usage Guidelines

The Tandem CPU Resource Planning Sample Application provides a capacity planning database file with the database elements you need for basic tracking, reporting, and forecasting of Tandem systems. The information for this application is derived from the CA MICS Tandem Analyzer's CPU Entity File (TDMCPE) and, optionally, the System Profile File (TDMSYP).

The Tandem CPU Resource Planning Sample Application allows you to plan by CPU number within the Tandem System by using the Capacity Planning Application Unit (CAPAPU) feature. There are three options available for CAPAPU value:

- 1) CPUNUM assigns each CPU Number to a CAPAPU value
- 2) NONE does not assign a CAPAPU value
- 3) EXIT assigns a CAPAPU value that you specify in the CAPAURT exit.

2.15.2 Control Parameters

Relatively few parameters are needed to implement the Tandem CPU Resource Planning Sample Application. The control screens included in the online facility allow you to enter the required information quickly.

```

/----- Tandem CPU Resource Planning -----\
|Command ==>
|Enter a ? in any data entry field for more information on valid values.
|Composing CAP DB File def.: TD1 - Tandem CPU Resource Planning Sample Appl.
|
|Summarization Level
|      Primary      ==> WEEKS   (DAYS/WEEKS/MONTHS)
|      Secondary    ==> ZONE    (ZONE/HOUR/DAYNAME)
|
|Input File Time-Span ==> DETAIL (DETAIL/DAYS/WEEKS/MONTHS)
|
|Amount of data retained ==> 60   (1-9999 months)
|Age of oldest forecast ==> 24   (1-9999 months)
|History cycles ==> 3   (1-99 cycles)
|Forecast cycles ==> 3   (1-99 cycles)
|
|Data base ID (s) ==> _____
|
|Tandem System Category (CAPAPU) ==> CPUNUM (CPUNUM/NONE/EXIT)
|
|      Application unit title ==> CPU Number _____
|
|Specify extended options ==> Y (Y/N)
|
\-----/

```

Figure 2-15.01. Tandem CPU Planning Screen

Figure 2-15.01 shows the Tandem CPU Planning screen, which is the main control screen for the Tandem CPU Resource Planning Sample Application. It contains the following fields:

Summarization level: The output summarization corresponds to the combination of the primary and secondary summarization levels. These control how the output file is summarized in your capacity planning database.

Input file timespan:	The timespan (either DETAIL, DAYS, WEEKS or MONTHS) of the TDMCPE and TDMSYP file that you want to use as input when creating the capacity planning database file.
Amount of data retained:	The number of days, weeks or months of historical data to be in the CAPxxx01 cycle of the capacity planning database file. Any historical observations older than what you specify for this parameter are dropped from the file during the file update.
Age of oldest forecast:	The number of days, weeks or months that you want to retain a forecast after it is generated. Any forecast older than what you specify for this parameter is dropped when a new forecast is generated and saved.
History cycle:	The number of cycles or versions of the historical data you want to retain. A new cycle is created when the file is updated and the existing cycles are aged (for example, Cycle 01 is aged to Cycle 02 and Cycle 02 becomes Cycle 03).
Forecast cycle:	The number of cycles or versions of the forecast data you want to retain. A new cycle is created when you specify forecast aging in one of the workload forecasting routines and the existing cycles are aged (for example, Cycle 01 is aged to Cycle 02 and Cycle 02 becomes Cycle 03).
Database ID(s):	The database unit identifier(s) (up to fourteen units) that contain the TDMCPE and TDMSYP files to be used with this Sample application. If you are not sure where the TDMCPE and TDMSYP files reside, type a question mark (?) in this field to display a listing of the valid database unit identifiers for this sample application.
Tandem System Category (CAPAPU):	The three options for defining Tandem System Category (see Section 2.15.1, Usage Guidelines).

Application unit title: The 1- to 40-character name for the CAPAPU data element. This parameter is only available for entry if you specify EXIT for the Tandem System Category. If you specify CPUNUM as the Tandem System Category, the title will be "CPU Number". You cannot specify a value for title if CAPAPU is NONE.

Specify extended options: If you want extended options, specify Y (YES), and the Extended Options Menu (Figure 2-15.02) is displayed.

```

/----- Extended Options Menu -----\
|Command ==>|
|
|Inquiry Step: Tandem CPU Resource Planning Sample Application|
|
|
|Combine all SYSIDs into a composite ==> N (Y/N)|
|Composite SYSID ==> ____|
|
|Combine all zones into a composite ==> N (Y/N)|
|Composite zone ==> _____|
|
|Specify:|
|Common data selection parameters ==> Y (Y/N/R)|
|Global data selection exit ==> Y (Y/N/R)|
|CAPAPU derivation SAS statements ==> N (Y/N/R)|
|Notes report enabled ==> YES (YES/NO)|
|Contents reports enabled ==> YES (YES/NO)|
|
|
|
|-----/

```

Figure 2-15.02. Tandem CPU Planning Extended Options Menu

The extended options available with the Tandem CPU Resource Planning Sample Application are:

Combine all SYSIDs into a composite: You can combine SYSIDs in the capacity planning database file. Enter Y (YES) if you want data for all of the SYSIDs to be summarized into a single observation with the Composite SYSID value.

Combine all zones/hours/daynames into a composite:	You can combine zones, hours, and daynames in the capacity planning database file. The prompt for this field changes to zones, hours or daynames depending on what you specify for the summary level parameter. Enter Y (YES) if you want all of the zones, hours, and daynames to be summarized into a single observation with the composite value.
Specify common data selection parameters:	Specify the SYSIDs and zones, hours or daynames to be used for data selection from the CA MICS input file(s). Enter Y (YES) to display the Common Data Selection screen (Figure 2-15.03). Enter R (RESET) to reset any selections you previously made to SYSID and zone, hour or dayname. The field defaults to N (NO), specifying no common data selection parameters.
Specify global data selection exit:	You can use your own SAS statements to select data from the CA MICS input file(s). Enter Y (YES) to display the SAS System Statements screen (Figure 2-15.04). Enter R (RESET) to reset any selections you previously made to the default of no user-specified data selection. The field defaults to N (NO) specifying that you are not coding a data selection exit.
Specify CAPAPU derivation SAS statements:	You can assign a CAPAPU value when Tandem System Category is set to EXIT on the Tandem CPU Planning Screen (Figure 2-15.01). Enter Y (Yes) to display the SAS System Statements screen. Enter R (RESET) to reset any SAS Statements you previously entered for CAPAPU assignment. This option is available only if you specify EXIT for CAPAPU parameter on the Tandem CPU Planning screen. The field defaults to N (NO) specifying that Tandem System Category is not set to EXIT.
Notes report enabled:	If you have entered notes associated with the capacity planning file definition, a Yes in this field prints the note as part of file updating.

Contents reports enabled:	A Yes in this field prints the contents of the capacity planning file as part of file updating.

----- Common Data Selection -----

Command ==>

Inquiry Step: Tandem CPU Resource Planning Sample Application

SYSID ==> -----

Zone ==> ----- (1 - 9)

Hour ==> -- - -- - -- - -- - -- (0 - 23)

Dayname==> -----

Figure 2-15.03. Common Data Selection Screen - Tandem CPU Planning

The Common Data Selection screen (Figure 2-15.03) allows you to define the following data selection criteria:

- SYSID:

The SYSIDs you want to include in your capacity planning database file. You can specify a maximum of 15 one- to four-character SYSIDs. Leave this parameter blank to include data for all SYSIDs.
- Zone:

Specify up to nine zones to be included in the capacity planning database file. Leave this parameter blank to include data for all zones.

- Hour:** Specify five hours, or range of hours, to include in the capacity planning database file. Selection of this parameter is valid only if you specify the DETAIL, DAYS, or WEEKS timespan on the Tandem CPU Planning screen (Figure 2-15.01). Leave this parameter blank to include data for all hours. You must enter singular hour values on the left side of the dash.
- Dayname:** Specify up to five daynames to include in the capacity planning database file. Leave this parameter blank to include all days. Input the first unique characters and the application will fill in the rest. Selection of this parameter is only valid when the input timespan is DETAIL or DAYS.

```

/----- SAS System Statements ----- COLUMNS 001 072--\
|Command ==>                               Scroll ==> PAGE |
|
|SAS Statements for the SELECT Macro
|-----|
|***** ***** TOP OF DATA *****|
|=MSG> SPECIFY SAS STATEMENTS TO SUBSET THE CAPACITY PLANNING DATA BASE FILE. |
|=NOTE= %MACRO SELECT; <=== PLEASE DO NOT PLACE LINES BEFORE THIS LINE.      |
|000001
|000002
|000003
|000004
|000005
|000006
|000007
|000008
|000009
|000010
|=NOTE= %MEND SELECT; <=== PLEASE DO NOT PLACE LINES AFTER THIS LINE.        |
|***** ***** BOTTOM OF DATA *****|
|
\-----/

```

Figure 2-15.04. SAS System Statement-Tandem CPU Planning Global Data Selection

The SAS System Statements screen (Figure 2-15.04) lets you define your own data selection logic. You can code any valid SAS Statement as described in the SAS manuals. The %MACRO SELECT and %MEND SELECT statements are already provided and field protected to prevent modification.

Note: This is a PDF/EDIT screen. All the PDF/EDIT commands are available.

2.15.3 Standard Output

The capacity planning database file generated by the Tandem CPU Resource Planning Sample Application contains the following data elements:

- A. Data elements common to all CA MICS Capacity Planner database files:

SYSID
ZONE, HOUR or DAYNAME
YEAR
WEEK, MONTH or MNTHDAY
DATE

- B. Data elements unique to the Tandem CPU Resource Planning Sample Application database file:

1. Character format data elements

CAPAPU - Tandem System Category
CPETYPE - Tandem CPU type

2. Numeric format data elements

From the CPU Entity File (TDMCPE):

CPECAHIT - Number of Cache Hits
CPECBSTM - CPU Busy Time
CPEDIOS - Disk I/Os
CPEDISP - Number of Process Dispatches
CPEIBSTM - Interrupt Busy Time
CPEOVHTM - Process Overhead Time
CPEPCBSY - Percent CPU Busy
CPEPCIBS - Percent CPU Interrupt Busy
CPEPCOBS - Percent CPU Process Overhead Busy
CPEPCSBS - Percent CPU Send Busy
CPEPSCAH - Cache Hits Per Second

CPEPSDIO - Disk I/Os Per Second
CPEPSDSP - Dispatches Per Second
CPEPSSWP - Swaps Per Second
CPEPSSTM - Send Busy Time
CPESWAPS - CPU Swaps
DURATION - Recording Interval

If Tandem System Category (CAPAPU) is NONE or EXIT, the following numeric format elements from the System Profile File (TDMSYP) are also kept on the capacity planning database file.

SYNOCPU - Maximum number of CPU processors
SYPTOUTM - Total system uptime

DERIVATION OF THE TANDEM CPU PLANNING DATA FILE ELEMENTS

The following paragraphs discuss the Tandem CPU Resource Planning data base file derived elements.

CAPAPU: The data element that contains the values found; either TDMCPUNO, values derived within the user's exit; or no CAPAPU value. In the capacity planning database, there is a separate observation for each value of CAPAPU for each WEEK, MONTH or MNTHDAY.

When the Tandem System Category is the default value, CPUNUM, the CA MICS Tandem CPU Entity File (TDMCPE), is the only file read and CAPAPU is the only derived element. The utilization and rate elements in the final capacity database are calculated using the standard CA MICS derivation macros. The duration, or uptime, is based on the time during the interval that a processor is online. If a CPU processor is available for half the week, the utilizations and rates are based on the time that the processor was available, and not the total available hours in a week.

When the Tandem System Category is NONE or EXIT, the CA MICS Tandem Analyzer System Profile File (TDMSYP) is read in addition to the Tandem Analyzer CPU Entity File (TDMCPE). Data elements from the TDMSYP file are used to determine the uptime across CPUs.

The following derived elements are created:

CPUCOUNT - Number of CPU processors found in the TDMCPE file. CPUCOUNT is assigned a value of 1 for each observation that is read from the TDMCPE file.

SYSDUR - The system duration, or system uptime, for the timespan being processed.

$$\text{SYSDUR} = \text{DURATION}$$

One new element, SYPTOUTM, is derived and the utilization and rates are recalculated using SYPTOUTM for the uptime of the summarized data.

For example, if CAPAPU is NONE, the timespan is weeks, summarization is hour, and zone selection is zone 1 (08:00 - 17:00, Monday-Friday), SYSDUR would be 45 hours. This is the available time, per processor.

The available time for the system (CAPAPU=NONE), or other system category (CAPAPU=EXIT), is SYSDUR multiplied by CPUCOUNT.

The time and count data elements are summarized to the specified summarization level and the percent and rate per second calculations are calculated by dividing the total times and counts by the NEW duration value.

The derived elements and recalculated elements are:

SYPUPTM: The total up time for the Tandem system (CAPAPU = NONE) or the subset of the Tandem system (CAPAPU=EXIT).

$$\text{SYPUPTM} = \text{SYSDUR} * \text{CPUCOUNT}$$

CPEPCBSY: Percent CPU busy time. This is the percent of total CPU busy time.

$$\text{CPEPCBSY} = (\text{CPECBSTM} / \text{SYPTOUTM}) * 100$$

CPEPCIBS: Percent interrupt busy time. This is the percent of time the CPU was busy handling interrupts.

$$\text{CPEPCIBS} = (\text{CPEIBSTM} / \text{SYPTOUTM}) * 100$$

CPEPCOBS: Percent overhead busy time. This is the percent of time reported by the CPU MEASURE Entity File as system overhead.

$$\text{CPEPCOBS} = (\text{CPEOVHTM}/\text{SYPTOUTM}) * 100$$

CPEPCSBS: Percent send busy time. This is the percent of time the CPU was busy sending data to other CPUs.

$$\text{CPEPCSBS} = (\text{CPESBSTM}/\text{SYPTOUTM}) * 100$$

CPEPSCAH: Cache hit rate per second. This is the number of cache hits per second.

$$\text{CPEPSCAH} = (\text{CPECAHIT}/\text{SYPTOUTM})$$

CPEPSDIO: Disk I/Os per second. This is the number of disk I/Os per second.

$$\text{CPEPSDIO} = (\text{CPEDIOS}/\text{SYPTOUTM})$$

CPEPSDSP: Dispatch rate per second. This is the number of PROCESS dispatches per second.

$$\text{CPEPSDSP} = (\text{CPEDISP}/\text{SYPTOUTM})$$

CPEPSSWP: Swaps per second. This is the number of swaps per second.

$$\text{CPEPSSWP} = (\text{CPESWAPS}/\text{SYPTOUTM})$$

You will probably only use the CPEPCBSY element for forecasting, however, other utilization and rates are included in the capacity planning database file for your convenience should you have a need to use these values for forecasting.

The utilizations and rates may also be used for performance trending reports. The capacity planning database may be accessed through MICF for creating long term printer and graphics reports. See Section 3.1 of this guide for more information on the structure of the capacity planning database file.

2.15.4 Data Elements Overview

The intent of this section is to provide the capacity planner with an understanding of the information chosen for the Tandem CPU Resource Sample Application and how these variables can be applied to capacity planning situations in a Tandem environment.

CPETYPE - Tandem CPU Type

The CPU type of the Tandem processor on which MEASURE data was recorded. It is alphanumeric. Although not directly useful for capacity planning, CPETYPE provides an audit trail of changes that may have occurred in your Tandem systems.

CPECBSTM - CPU Busy Time

The amount of time that the CPU was executing instructions. CPECBSTM is maintained as a TIME14.2 SAS variable. Unlike Percent CPU Busy (CPEPCBSY), this is a direct measurement of the amount of time that the CPU was active. Because CPU is the most critical resource, and CPECBSTM determines the amount being used, this is one of the most important variables in capacity planning, regardless of the operating environment.

The capacity planner should watch this value closely and be aware of any unusual fluctuations in its size (either up or down). This variable can be used as a forecast element in univariate modeling, however, be wary of changes in Tandem CPU Type (CPETYPE) that may affect the rate at which instructions are executed.

CPEPCBSY - Percent CPU Busy

A numeric variable that reflects the percentage of uptime that the CPU was busy executing instructions. The value ranges from 0 to 100% (see Section 2.15.3) for the derivation) and is a relative indicator of how much CPU resource is consumed (or remaining). Like CPU Busy Time (CPECBSTM), this value can be used to compare the activity of one system (or CPU) to another.

CPEPCBSY provides a simple indicator of processor resources and can also be used for forecasting purposes. Although it provides the same information as CPECBSTM, it is more helpful because you do not need to examine uptime to see how much CPU time is available. It is also useful for presenting a capacity plan to management.

CPEPCBSY is the most commonly used value for capacity planning. Other elements from the CPU file are included in the capacity database file because they may be of some interest to the capacity planner, however, CPEPCBSY is the recommended element for forecasting CPU capacity needs.

CPEDIOS - Disk I/Os

The number of disk I/Os processed in the CPU. It is an accumulated numeric variable and is limited only by practical bounds. CPEDIOS is a relative indicator of I/O activity on the system.

I/O is second only to CPU utilization when evaluating the capacity of a computer system. Increases in I/O indicate growth in workloads on the system. Use CPEDIOS to forecast when the I/O counts may cause system saturation.

CPEPSDIO - Disk I/Os Per Second

The rate of disk I/Os per second of uptime for the processor. See Section 2.15.3 for its derivation. It is numeric and limited only by practical bounds. Like CPEDIOS, CPEPSDIO is a relative indicator of I/O activity on the system.

Because CPEPSDIO is in rate per second rather than total count, it is more suitable for trending and forecasting than CPEDIOS. Monitor the trend of I/Os per second to determine when your system will reach its I/O processing saturation limit.

CPEDISP - Number of Process Dispatches

The number of processes dispatched by the CPU. It is an accumulated numeric variable and is limited only by practical bounds. The number of dispatches is an indicator of the amount of work being completed on the system.

Temporary increases in CPEDISP indicate periods of heavy work on the system, while a constant increase in this value indicates steady growth in workloads.

CPEIBSTM - Interrupt Busy Time

The amount of time the CPU was busy handling interrupts. CPEIBSTM is maintained as a time variable. The interrupts are usually for processing of messages and I/O operations. Monitor the changes CPEIBSTM in comparison to I/O rates and messages sent or received for indications of performance problems.

Disproportional increases in CPEIBSTM to the total CPU Busy Time (CPECBSTM) may indicate problems with I/O or message processes. You may want to report CPEIBSTM as a percentage of CPECBSTM.

CPEOVHTM - Process Overhead Time

The CPU time spent on overhead that cannot be identified with a particular process. It is maintained as a time variable. It is limited only by practical bounds, however, it is usually a small percentage of the total CPU Busy Time (CPECBSTM). Like CPEIBSTM, you may want to report CPEOVHTM as a percentage of CPECBSTM.

CPEPCOBS - Percent CPU Process Overhead Busy

The percentage of uptime that the CPU was busy with overhead processing. See Section 2.15.3 for information on its derivation. The value of CPEPCOBS ranges from 0 to 100%, however, it is usually only a fraction of a percent of the total processor uptime.

CPEPCOBS is most useful when shown as a percentage of the total CPU Busy Time (CPECBSTM). Low overhead busy time indicates efficient use of the processor. You may want to use this element to measure the trend of overhead busy time over an extended period of time. Disproportional increases in the element should be cause for concern.

CPEPSSTM - Send Busy Time

The amount of time the CPU was busy sending data to another system. CPEPSSTM is maintained as a time variable and is limited only by practical bounds. There may be a slight overlap in CPEPSSTM and Interrupt Busy Time (CPEIBSTM).

Like CPEIBSTM, CPEPSSTM is most useful when shown as a percentage of Total CPU Busy (CPECBSTM). You may want to look at this element for trends over a long period of time.

CPEPCSBS - Percent CPU Send Busy

The percentage of uptime the processor is busy sending data to another system. It is maintained as a numeric variable and ranges from 0 to 100%. See Section 2.15.3 for the derivation. Like other percent variables, you can determine how much of the available time is being used without examining other variables.

CPEPCSBS is most meaningful when compared to the total CPU Busy Time (CPECBSTM). Monitor the trend of this element to see what percentage of processor time is used for interaction between systems.

CPEPCIBS - Percent CPU Interrupt Busy

The percentage of uptime the processor is busy handling system interrupts. CPEPCIBS is maintained as a numeric element and ranges from 0 to 100%. See Section 2.15.3 for the derivation. Like other percent variables, you can determine how much of the available time is being used without examining other variables.

CPEPCIBS is most meaningful when compared to the total CPU Busy Time (CPECBSTM). Monitor the trend of this element to see what percentage of processor time is used for handling interrupts.

CPECAHIT - Number of Cache Hits

The number of times the required data was found in cache and resulted in a reduced disk I/O count. It is numeric and limited only by practical bounds.

The capacity planner should monitor this value. Decreased cache hits may indicate a need for more cache.

CPEPSCAH - Cache Hits Per Second

CPEPSCAH measures the rate of cache hits per second and is numeric. See Section 2.15.3 for the derivation. Higher rates indicate better performance and a sufficient capacity of cache resources.

The capacity planner should watch for decreases in this rate as an indication the system is approaching its capacity limit. Decreases in the cache hits per second will result in increased disk I/Os.

CPEPSDSP - Dispatches Per Second

The rate of dispatches per second for the processor. It is numeric and is limited only by practical bounds. See Section 2.15.3 of this guide for the derivation. Because CPEPSDSP factors in the uptime, this number is more useful for forecasting dispatches.

The capacity planner will use this number to determine the throughput rate of the processor. An increase in the dispatch rate may indicate an increase in the work being done or a change in the mix of work. When the dispatch rate is a constant value, review system use carefully to determine whether the rate is constant because all necessary work is being completed or because the system is saturated.

CPEPSSWP - Swaps Per Second

The swap rate per second for the processor. It is a numeric value and is limited only by practical bounds. See Section 2.15.3 for the derivation.

CPEPSSWP is an indicator of memory usage. Low rates indicate adequate memory, while higher rates indicate contention for the available memory on the system. Monitor this value to forecast memory requirements.

CPESWAPS - CPU Swaps

CPESWAPS is an accumulated numeric counter of swaps in and out of the system by the memory manager. It is limited only by practical bounds. The swap count is a count of the number of swaps, not the number of pages swapped.

The capacity planner will monitor this value to determine when more memory is required. Because this is a counter and not a rate, be aware of the interval duration when using this variable.

DURATION - Recording Interval

DURATION is a time variable indicating the duration for the interval used for this observation in the capacity database file. The duration value is taken from the TDMCPE file and may be less than the system duration, SYSDUR, if the processor was not active for the full interval. DURATION is used in rate and utilization calculations when the CAPAPU value is CPUNUM.

This value is included on the capacity database file for informational purposes. It may serve as an indicator of missing processor data.

Note: If Tandem System Category (CAPAPU) is NONE or EXIT, the System Profile File (TDMSYP) is merged with the CPU Entity File (TDMCPE) to determine the duration of the intervals for the complete system. Only in this circumstance are the remaining elements kept on the capacity database file.

SYPNOCPU - Maximum Number of CPU Processors

The maximum number of CPU processors in use for the entire system during this interval. It is numeric from 1 to 16.

SYPNOCPU is kept on the capacity database file for informational purposes and is not recommended as a forecast element.

SYPTOUTM - Total System Uptime

The total time the system is active. SYPTOUTM is stored as a time element. It is derived, as shown in Section 2.15.3, by the Tandem CPU Resource Sample Application.

SYPTOUTM is used as the denominator for calculating rates and percent utilization when data is summarized across CPU processors.

SYPTOUTM is kept in the capacity database file because it is used to derive other elements. It is not recommended as a forecast element.

CPUCOUNT - Number of CPU Processors Found in the TDMCPE File

The number of CPU processors in the TDMCPE file. It is numeric and accumulated. Because data is summarized across CPU processors when NONE or EXIT is specified, you need to know how many CPU processors are included in the system or subsystem represented in the capacity planning database file. CPUCOUNT is used in the derivation of SYSTOUTM.

CPUCOUNT is included in the capacity database file because it is used to derive other values. It is not recommended as a forecasting element.

SYS DUR - System Duration or System Uptime

The duration or uptime of the entire system for an interval. It is stored as a time element. The derivation is shown in Section 2.15.3. SYS DUR is used when CAPAPU is EXIT or NONE, rather than using the duration from the individual processors in the TDMCPE file, because it gives the duration of the entire system. For example, for the WEEKS timespan, for hour 10, for the prime shift (Monday through Friday), SYS DUR would be 5 hours.

2.16 VAX/VMS CPU Resource Planning

The VAX/VMS CPU Resource Planning Sample Application provides a capacity planning database file that contains the necessary elements for CPU resource tracking and reporting in a VAX/VMS environment.

The following sections cover the use, input specifications, and output of the sample application:

- 1 - Usage Guidelines
- 2 - Control Parameters
- 3 - Standard Output

2.16.1 Usage Guidelines

The VAX/VMS CPU Resource Planning Sample Application provides a capacity planning database file with the database elements you need for basic tracking, reporting, and forecasting of the VAX/VMS CPU workload. The information for this application is derived from either the DETAIL, DAYS, WEEKS, or MONTHS timespan of the CA MICS VMS MONITOR Modes (DEMMOD) file and the CA MICS VMS System Profile (DEMSPR) file. This sample application provides information about system-wide processor use. Data for individual CPU IDs in the DEMMOD file are summarized to the SYSID (system node) level. Consequently, the VAX/VMS CPU Resource Planning Sample Application permits your installation to track and forecast CPU use at a high level of summarization (that is, at the SYSID and ZONE, HOUR or DAYNAME level).

In addition to providing basic information about processor utilization, the VAX/VMS CPU Resource Planning Sample Application also stores data concerning direct and buffered I/O, page faulting, and Inswaps. You may find these elements useful in tracking and reporting. For more information about the data contained in the VAX/VMS files see the VAX/VMS Analyzer Guide.

NORMALIZATION

Another useful feature that the VAX/VMS CPU Resource Planning Sample Application provides is the ability to normalize processor utilization to a base CPU. Processor busy time is an indicator that is commonly used in capacity planning for reporting and forecasting. However, as systems are saturated and upgraded to more powerful systems, a graph of processor busy looks more like a "saw-tooth" function than a smooth, monotonically increasing function. This is because more powerful systems can process the same amount of work in less time than their slower predecessors could, and therefore spend less time processing. This hides the fact that the installation has really increased its processing capability with the upgrade.

You can circumvent this situation by normalizing all processor busy time to the equivalent of a base CPU. You may optionally request that the processor busy time and processor percent busy elements be adjusted to a base CPU of your specification. To use this feature, you must supply the following information:

- o The base CPU model type, or name
- o The base CPU Vax Units of Performance (VUP) rating

If the CPU Time Adjustment option is chosen, both the normalized and the unnormalized CPU busy and percent utilization elements are still stored in the database and are available for reporting and analysis.

We recommend that you use this option to normalize workloads, since trending and forecasting CPU utilization without normalization is extremely difficult and often misleading.

IMPORTANT NOTE

You must also be certain that the VUPs ratings for your VAX/VMS systems are stated correctly in the CA MICS SYSID parameter. These values are used to rate the CPU information as it is input into the CA MICS database by the CA MICS VAX/VMS Analyzer Option.

NORMALIZATION EXAMPLE

The concept of normalization can be made clearer through an example. Suppose that an installation installed the VAX/VMS CPU Resource Planning Sample Application. They chose the optional CPU time adjustment and specified a VAX 8350 as the base CPU, with a relative performance rating of 2.3 (VUP).

These parameters, entered on the CPU Time Adjustment Parameters control screen, are then used to normalize CPU time values for any processor appearing in the CA MICS VMS MONITOR Modes (DEMOD) file to the equivalent of a VAX 8350 processor. Assume that one of the CPUs in the DEMOD file for this installation is a VAX 11/780, with a relative performance rating of 1 (VUP). During one week, this CPU had an average processor busy time of 35 hours in 40 hours of operation. We can calculate the average percent busy for the VAX 11/780 for the week as:

$$100 * (35 / 40) = 87.5\%.$$

Given that the VAX 11/780 has a VUP rating of 1, by convention, we have chosen to use this factor for our normalization. Therefore, the normalized average CPU busy time is:

$$35 * (1 / 2.3) = 15.2 \text{ hours}$$

In other words, this technique estimates that a workload requiring approximately 35 hours of CPU time on a VAX 11/780 requires approximately 15.2 CPU hours on a VAX 8350. The adjusted CPU percent utilization would be:

$$100 * (15.2 / 40) = 38.0\%$$

While normalization can calculate that a CPU runs at over 100% utilized, this type of calculation shows what the estimated utilization would be if the workload were running on a different type of VAX System. That is, we would need multiple systems of a given type if the workload exceeds the capacity of the system.

Note first that the use of an actual, commercially available VAX system as the base CPU is not required, and may not even be desirable for the installation. It would be perfectly acceptable to specify a hypothetical system. For example:

```
Base CPU Model   = MY_CPU
CPU VUP rating   = 10
```

These parameters adjust all processor busy elements to the equivalent of a hypothetical CPU having a relative performance rating of 10 VUPs. The fact that such a CPU may not actually exist is not important, but rather that all CPU time is being converted to a common unit that can be trended and examined smoothly and consistently over time.

These parameters are not dependent on the number of processors in the VAX system under study, since both the VUP rating and the summarized data represent the VAX system as a whole rather than individual processors in the CPU.

Also note that these calculations make no attempt to apply any analytic queuing formulas or benchmark estimates to compare one processor to another. The basic comparison of systems is calculated as the ratio of relative performance ratings for the respective processors. These numbers can be supplied by the manufacturer or can be established by the installation through benchmark testing.

2.16.2 Control Parameters

You need relatively few parameters to implement the VAX/VMS CPU Resource Planning Sample Application. The control screens included in the online facility allow you to quickly enter the information required.

```

/----- VAX/VMS CPU Resource Planning -----\
|Command ==>
|Enter a ? in any data entry field for more information on valid values.
|Composing CAP DB File def.: VAX - VAX/VMS Planning Sample Application
|
|Summarization Level
|      Primary      ==> MONTHS  (DAYS/WEEKS/MONTHS)
|      Secondary    ==> ZONE    (ZONE/HOUR/DAYNAME)
|
|Input File Time-Span ==> DETAIL  (DETAIL/DAYS/WEEKS/MONTHS)
|
|Amount of data retained ==> 60    (1-9999 months)
|Age of oldest forecast ==> 24    (1-9999 months)
|History cycles ==> 3    (1-99 cycles)
|Forecast cycles ==> 3    (1-99 cycles)
|
|Data base ID (s) ==> _____
|
|Specify extended options ==> Y (Y/N)
\-----/

```

Figure 2-16.01 VAX/VMS CPU Planning Control Screen

Figure 2-16.01 shows the VAX/VMS CPU Planning screen, which is the main control screen for the VAX/VMS CPU Resource Planning Sample Application. It contains the following fields:

Summarization level:	The output summarization corresponds to the combination of the primary and secondary summarization levels. These control how the output file is summarized in your capacity planning database.
Input file timespan:	The timespan (either DETAIL, DAYS, WEEKS or MONTHS) of the DEMMOD and DEMSPR files used as input to the capacity planning database file.

Amount of data retained	The number of days, weeks or months of historical data to be retained in the CAPxxx01 cycle of the capacity planning database file. Any historical observations older than what you specify for this parameter are dropped from the file during the file update.
Age of oldest forecast:	The number of days, weeks or months that you want to retain a forecast after it is generated. Any forecast older than what you specify for this parameter is dropped when a new forecast is generated and saved.
History cycles:	Number of cycles or versions of the historical data you want to retain. A new cycle is created created when the file is updated and the existing cycles aged (for example, Cycle 01 is aged to Cycle 02 and Cycle 02 becomes Cycle 03).
Forecast cycles:	Number of cycles or versions of the forecast data you wish to retain. A new cycle is created when you specify forecast aging in one of the workload forecasting routines and the existing cycles are aged (for example, Cycle 01 is aged to Cycle 02 and Cycle 02 becomes Cycle 03).
Database ID(s):	The database unit identifier(s) (up to 14 units) that contain the DEMMOD and DEMSPR files to be used with this sample application. If you are not sure where the DEMMOD and DEMSPR files reside, type a question mark (?) in this field to display a listing of the valid database unit identifiers for this sample application.

Specify common data selection parameters:	Specify the SYSIDs and zones, hours or daynames to be used for data selection from the CA MICS input file(s). Enter Y (YES) to display the Common Data Selection screen shown in Figure 2-16.03. Enter R (RESET) to reset any selections you previously made to SYSID and zone, hour or dayname.
Specify global data selection exit:	You can use your own SAS statements to select data from the CA MICS input file(s). Enter Y (YES) to display SAS System Statements screen shown in Figure 2-16.04. Enter R (RESET) to reset any selections you previously made to the default of no user-specified data selection.
Specify CPU time adjustment parameters:	Specify the parameters to use for CPU time adjustment (normalization), if you desire this optional feature. Enter Y (YES) to display the CPU Time Adjustment Parameters screen shown in Figure 2-16.05. Enter R (RESET) to reset any selections you previously made to CPU time adjustment. This option is not available if you specify the composite SYSID option on this screen shown in Figure 2-16.02.
Notes report enabled:	If you have entered notes associated with the capacity planning file definition, a Yes in this field prints the note as part of file updating.

Contents reports enabled:	A Yes in this field prints the contents of the capacity planning file as part of file updating.

```

/------ Common Data Selection -----\
|Command ==>
|
|Inquiry Step: VAX/VMS Planning Sample Application
|
|SYSID ==>      _ _ _ _ _ 
|              _ _ _ _ _ 
|
|Zone   ==>     _ - - - -         (1 - 9)
|
|Hour   ==>    __ - __ - __ - __ - __ - __ - __ - __       (0 - 23)
|
|Dayname==>    _____
|
|
|
|
|\-----\

```

Figure 2-16.03. Common Data Selection Screen - VAX/VMS CPU Planning

The Common Data Selection screen shown in Figure 2-16.03 allows you to define the following data selection criteria:

SYSID: The SYSIDs that you want to include in your capacity planning database file. You can specify a maximum of 15 one- to four-character SYSIDs. Leave this parameter blank to include data for all SYSIDs.

Zone: Specify up to nine zones to include in the capacity Planning database file. Leave this parameter blank to include data for all zones.

Hour: Specify the five hours or hour ranges to include in the capacity planning database file. This parameter is valid only if you specify the DETAIL, DAYS, or WEEKS timespan on the VAX/VMS CPU Planning screen shown in Figure 2-16.01. Leave this parameter blank to include data for all hours.

Dayname: Specify up to five daynames to include in the capacity planning database file. Leave this parameter blank to include all days. Input the first unique characters and the application will fill in the rest. This parameter is only valid when the input timespan is DETAIL or DAYS.

```

/----- SAS System Statements ----- COLUMNS 001 072--\
|Command ==>                               Scroll ==> PAGE |
|
|SAS Statements for the SELECT Macro
|-----|
|***** ***** TOP OF DATA *****|
|=MSG> SPECIFY SAS STATEMENTS TO SUBSET THE CAPACITY PLANNING DATA BASE FILE. |
|=NOTE= %MACRO SELECT; <=== PLEASE DO NOT PLACE LINES BEFORE THIS LINE.      |
|000001
|000002
|000003
|000004
|000005
|000006
|000007
|000008
|000009
|000010
|=NOTE= %MEND SELECT; <=== PLEASE DO NOT PLACE LINES AFTER THIS LINE.        |
|***** ***** BOTTOM OF DATA *****|
|
\-----/

```

Figure 2-16.04. SAS System Statements Screen - VAX/VMS CPU Planning Global Data Selection

The SAS System Statements screen shown in Figure 2-16.04 lets you define your own data selection logic. You can code any valid SAS statements, as described in the SAS manuals. The %MACRO SELECT and %MEND SELECT statements are already provided and are field protected to prevent modification.

Note: This is a PDF/EDIT screen. All the PDF/EDIT commands are available.

```

/------ CPU Time Adjustment Parameters -----\
|Command ==>
|
|Inquiry step: VAX/VMS CPU Resource Planning Sample Application
|
|Base CPU model              ==> VAX-8350____
|Base CPU VUP rating         ==> 2.3___
|
|
|
|
|
|
|
|
|
|

```

Figure 2-16.05. CPU Time Adjustment Parameters Screen

The information on the CPU Time Adjustment Parameters screen shown in Figure 2-16.05 is required only if you choose CPU time normalization for the VAX/VMS CPU Resource Planning Sample Application.

Base CPU model:	The 10-character identifier for the base VAX system. Any name is acceptable in this field. Do not use embedded blanks or equal signs (=) in this value. For example, if you want to normalize your actual CPU time values to the equivalent of a VAX 8350, then type VAX-8350 in this field.

Base CPU VUP rating: The relative performance (VUP) rating for the base processor. The value must be a real number greater than 0 (zero). You can use vendor-supplied values for the factor, or use a factor based on your own evaluations. For example, if you specify a VAX-8350 as the base CPU model, then type 2.3 in this field.

2.16.3 Standard Output

The capacity planning database file that is generated by the VAX/VMS CPU Resource Planning Sample Application contains the following information:

- A. Data elements common to all CA MICS Capacity Planner database files:

SYSID
ZONE or HOUR or DAYNAME
YEAR
MONTH or WEEK or MNTHDAY
DATE

- B. Data elements that are unique to the VAX/VMS CPU Resource Planning Sample Application database file:

1. Character format data elements

none

2. Numeric format data elements

INTERVLS - Number of Recording Intervals (MODES)
SPRBIOS - Buffered I/Os
SPRPCUNI - Number of Instructions Executed
SPRDIOS - Direct I/Os
SPRFPGFL - Free List Page Faults
SPRINSWP - Inswaps
SPRMPGFL - Modified List Page Faults
SPRPGFLT - Page Faults
SPRPSBIO - Buffered I/O Rate
SPRPSDIO - Direct I/O Rate
SPRPSFLT - Page Fault Rate
SPRPSFPL - Free List Page Fault Rate

SPRPSINS - Inswap Rate
 SPRPSMPL - Modified List Page Fault Rate
 SPRPSRIO - Read I/O Rate
 SPRPSSYS - System Page Fault Rate
 SPRRDIOS - Page Read I/Os
 SPRSYSFL - System Page Faults
 SPRSYTCT - SYT SYSTEM Record Count (Type 17)
 VUPS - VAX Units of Performance

3. Time format data elements

DURATION - Recording Interval Time (MODES sum)
 MODCOMTM - Compatibility Mode Time
 MODCPUTM - Total CPU Time
 MODEXETM - Executive Mode Time
 MODIDLTM - Idle Time
 MODINTTM - Interrupt Stack Time
 MODKERTM - Kernel Mode Time
 MODMPSTM - Multi-Processor Sync Time
 MODSUPTM - Supervisor Mode Time
 MODUSRTM - User Mode Time
 MODVECTM - CPU Time of Vector Consumers
 SPRCPUTM - System CPU Busy Time
 SPRIOVDR - IOV IO Record Duration (Type 4)
 SPRPGEDR - PGE PAGE Record Duration (Type 3)
 SPRSYTDR - SYT SYSTEM Record Duration (Type 17)

4. Percentage format data elements

MODPCCOM - Compatibility Mode Percent
 MODPCCPU - Total CPU Percent
 MODPCEXE - Executive Mode Percent
 MODPCIDL - Idle Percent
 MODPCINT - Interrupt Stack Percent
 MODPCKER - Kernel Mode Percent
 MODPCMPS - Multi-Processor Sync Percent
 MODPCSUP - Supervisor Mode Percent
 MODPCUSR - User Mode Percent
 USRPCTOT - User Mode as Percent of Total CPU

C. Additional data elements if CPU time adjustment is specified:

1. Character format data element

BASECPU - Base CPU Model

2. Numeric format data elements

VUPSA - VAX Units of Performance Adjusted
VUPS - VAX Units of Power

3. Time format data element

ADJCOMTM - Adjusted Compatibility Mode Time
ADJCPUTM - Adjusted Total CPU Time
ADJEXETM - Adjusted Executive Mode Time
ADJINTTM - Adjusted Interrupt Stack Time
ADJKERTM - Adjusted Kernel Mode Time
ADJMPSTM - Adjusted Multi-Processor Sync Time
ADJSUPTM - Adjusted Supervisor Mode Time
ADJUSRTM - Adjusted User Mode Time

4. Percentage format data elements

ADJPCCOM - Adjusted Compatibility Mode Percent
ADJPCCPU - Adjusted Total CPU Percent
ADJPCEXE - Adjusted Executive Mode Percent
ADJPCINT - Adjusted Interrupt Stack Percent
ADJPCKER - Adjusted Kernel Mode Percent
ADJPCMPS - Adjusted Multi-Processor Sync Percent
ADJPCSUP - Adjusted Supervisor Mode Percent
ADJPCUSR - Adjusted User Mode Percent

Many of the data elements used in the VAX/VMS CPU Resource Planning Sample Application database file have the same meaning and derivation as their identically named counterparts in the CA MICS DEMMOD and DEMSPR Files. These data elements are named MODnnnnn and SPRnnnnn respectively, plus the DURATION and INTERVLS elements.

The data elements that are derived from the CA MICS database and used by the VAX/VMS CPU Resource Planning Sample Application are calculated as follows:

USRPCTOT: The percentage of CPU busy time spent in USER Mode. It is calculated by the formula:

```
IF MODCPUTM GT 0 THEN  
USRPCTOT=MODUSRTM*100/MODCPUTM;
```

VUPS: The relative performance (VUP) rating for the CPU, before adjustment, as specified to the CA MICS SYSID parameter for each SYSID (node). It is calculated by the formula:

```
IF (SPRCPUTM GT 0) THEN
VUPS=SPRCPUNI/SPRCPUTM;
```

This formula extracts the VUPS value component from the SPRCPUNI element on the CA MICS database.

The additional data elements that are present if CPU time adjustment is specified are derived as follows:

BASECPU: The value that was specified for the Base CPU model parameter on the CPU Time Adjustment Parameters screen shown in Figure 2-16.05.

VUPSA: The value that was specified for the Base CPU VUP rating on the CPU Time Adjustment Parameters screen shown in Figure 2-16.05.

ADJnnnnn The adjusted value data elements are named ADJnnnnn and where the nnnnn corresponds to the name of the unadjusted element from the DEMMOD file. The adjustment is performed as follows:

```
IF (SPRCPUTM GT 0) AND (VUPSA GT 0) THEN DO;
```

```
ADJCOMTM=MODCOMTM*(VUPS/VUPSA); Compatibility Mode Time
ADJPCCOM=MODPCCOM*(VUPS/VUPSA); Compatibility Mode Pct
```

```
ADJCPUTM=MODCPUTM*(VUPS/VUPSA); Total CPU Time
ADJPCCPU=MODPCCPU*(VUPS/VUPSA); Total CPU Pct
```

```
ADJEXETM=MODEXETM*(VUPS/VUPSA); Executive Mode Time
ADJPCEXE=MODPCEXE*(VUPS/VUPSA); Executive Mode Pct
```

```
ADJINTTM=MODINTTM*(VUPS/VUPSA); Interrupt Stack Time
ADJPCINT=MODPCINT*(VUPS/VUPSA); Interrupt Stack Pct
```

```
ADJKERTM=MODKERTM*(VUPS/VUPSA); Kernel Mode Time
ADJPCKER=MODPCKER*(VUPS/VUPSA); Kernel Mode Pct
```

```
ADJMPSTM=MODMPSTM*(VUPS/VUPSA); Multi-Processor Sync Time
ADJPCMPS=MODPCMPS*(VUPS/VUPSA); Multi-Processor Sync Pct
```

```
ADJSUPTM=MODSUPTM*(VUPS/VUPSA); Supervisor Mode Time
ADJPCSUP=MODPCSUP*(VUPS/VUPSA); Supervisor Mode Pct

ADJUSRTM=MODUSRTM*(VUPS/VUPSA); User Mode Time
ADJPCUSR=MODPCUSR*(VUPS/VUPSA); User Mode Pct

END;
```

2.17 HP Openview UNIX Resource Planning

The HP Openview UNIX Resource Planning Sample Application provides a capacity planning database file that contains the necessary elements for HP UNIX resource tracking and reporting.

The following sections cover the use, input specifications, and output of the sample application:

- 1 - Usage Guidelines
- 2 - Control Parameters
- 3 - Standard Output

2.17.1 Usage Guidelines

The HP Openview UNIX Resource Planning Sample Application provides a capacity planning database file with the database elements you need for basic tracking, reporting, and forecasting of the UNIX workloads on your system. The information for this application is primarily derived from either the DETAIL, DAYS, WEEKS, or MONTHS timespan of the CA MICS Global Performance File (AUMGPR) and the CA MICS Application Performance File (AUMAPP).

In the HP Openview UNIX Resource Planning Sample Application, the application workload group (CAPAPU) data element enables you to group your individual UNIX systems into logically similar workload groups for more meaningful tracking and reporting. Three methods are provided for specifying the manner in which the UNIX applications are grouped: 1) defining each UNIX application as its own workload group (CAPAPU = AUMAPU), 2) combining all UNIX applications (no workload group), 3) defining your own SAS statements to create the workload groupings by setting the CAPAPU values.

2.17.2 Control Parameters

Relatively few parameters are needed to implement the HP Openview UNIX Resource Planning Sample Application. The control screens included in the online facility allow you to quickly enter the required information.

```
/----- HP Openview UNIX Resource Planning -----\
|Command ==>
|Enter a ? in any data entry field for more information on valid values.
|Composing CAP DB File def.: MW1 - HP Openview UNIX Resource Plan. Sample Appl.
|
|Summarization Level
|      Primary          ==> MONTHS  (DAYS/WEEKS/MONTHS)
|      Secondary        ==> ZONE    (ZONE/HOUR/DAYNAME)
|
|Input File Time-Span    ==> MONTHS  (DETAIL/DAYS/WEEKS/MONTHS)
|
|Amount of data retained ==> 60      (1-9999 months)
|Age of oldest forecast  ==> 24      (1-9999 months)
|History cycles          ==> 3       (1-99 cycles)
|Forecast cycles         ==> 3       (1-99 cycles)
|
|Data base ID (s)        ==> _____
|
|UNIX Workload Definition (CAPAPU) ==> AUMAPU  (APPLNM/NONE/EXIT)
|      Application unit title ==> _____
|
|Specify extended options ==> Y       (Y/N)
\-----/
```

Figure 2-17.01. HP Openview UNIX Planning Screen

Figure 2-17.01 shows the HP Openview UNIX Planning screen, which is the main control screen for the HP Openview UNIX Resource Planning Sample Application. It contains the following fields:

- Summarization level: The output summarization corresponds to the combination of the primary and secondary summarization levels. These control how the output file is summarized in your capacity planning database.
- Input file timespan: The timespan (either DETAIL, DAYS, WEEKS or MONTHS) of the AUMGPR and AUMAPP Files that you want to use as input when creating the capacity planning database file.

Amount of data retained:	The number of days, weeks or months of historical data to be retained in the CAPxxx01 cycle of the capacity planning database file. Any historical observations older than what you specify for this parameter are dropped from the file during the file update.
Age of oldest forecast:	The number of days, weeks or months that you want to retain a forecast after it is generated. Any forecast older than what you specify for this parameter is dropped when a new forecast is generated and saved.
History cycle:	Number of cycles or versions of the historical data you want to retain. A new cycle is created when the file is updated and the existing cycles aged (for example, Cycle 01 is aged to Cycle 02 and Cycle 02 becomes Cycle 03).
Forecast cycle:	Number of cycles or versions of the forecast data you want to retain. A new cycle is created when you specify forecast aging in one of the workload forecasting routines and the existing cycles are aged (for example, Cycle 01 is aged to Cycle 02 and Cycle 02 becomes Cycle 03).
Database ID(s):	The database unit identifier(s) (up to 14 units) that contain the AUMGPR and AUMAPP Files to be used with this sample application. If you are not sure where the AUMGPR and AUMAPP Files reside, type a question mark (?) in this field to display a listing of the valid database unit identifiers for this sample application.
UNIX Workload Definition (CAPAPU):	The three options for defining workload groups (see section 2.16.1, Usage Guidelines).
Application unit title:	The 1- to 40-character long name for the CAPAPU data element. You can use this parameter to define your workload groupings. This parameter is only available if you specify EXIT for the UNIX workload definition (CAPAPU) parameter.

Specify extended options:	If you want extended options, specify Y (YES), and the Extended Options Menu (Figure 2-17.02) is displayed.

```
/----- Extended Options Menu -----\
|Command ==>
|
|Inquiry Step: HP Openview UNIX Resource Planning Sample Application
|
|Combine all SYSIDs into a composite ==> N            (Y/N)
|    Composite SYSID                  ==> ____
|
|Combine all zones into a composite ==> N            (Y/N)
|    Composite zone                   ==> _____
|
|Specify:
|    Common data selection parameters ==> N          (Y/N/R)
|    Global data selection exit       ==> N          (Y/N/R)
|    CAPAPU derivation SAS statements ==> N          (Y/N/R)
|    Notes report enabled             ==> YES        (YES/NO)
|    Contents reports enabled         ==> YES        (YES/NO)
```

Figure 2-17.02. HP Openview UNIX Planning Extended Options Menu

The extended options available with the HP Openview UNIX Resource Planning Sample Application (Figure 2-17.02) are:

Combine all SYSIDs into a composite: You can combine SYSIDs in the capacity planning database file. Enter Y (YES) if you want data for all of the SYSIDs to be summarized into a single observation with the Composite SYSID value.

Combine all You can combine zones, hours, and daynames in zones/hours/ the capacity planning database file. The prompt daynames for this field changes to zones, hours or into a daynames depending on what you specify for the composite: summary level parameter. Enter Y (YES) if you want all of the zones, hours, and daynames to be summarized into a single observation with the composite value.

Specify common data selection parameters:	Specify the SYSIDs and zones, hours or daynames to be used for data selection from the CA MICS input file(s). Enter Y (YES) to display the Common Data Selection screen (Figure 2-17.03). Enter R (RESET) to reset any selections you previously made to SYSID and zone, hour or dayname.
Specify global data selection exit:	You can use your own SAS statements to select data from the CA MICS input file(s). Enter Y (YES) to display the SAS System Statements screen (Figure 2-17.04). Enter R (RESET) to reset any selections you previously made to the default of no user-specified data selection.
Specify CAPAPU derivation SAS statements:	You can use your own SAS statements for defining your UNIX workloads. Enter Y (YES) to display the SAS System Statements screen. Enter R (RESET) to reset any SAS System Statements you previously made to a user-specified workload grouping. This option is available only if you specify EXIT for the UNIX Workload Definition (CAPAPU) parameter on the HP Openview UNIX Planning screen.
Notes report enabled:	A Yes in this field prints any notes associated with the capacity planning file definition as part of file updating.

Contents reports enabled:	A Yes in this field prints the contents of the capacity planning file as part of file updating.

Common Data Selection

Command ==>

Inquiry Step: Workload Planning Sample Application

SYSID ==> _____

Zone ==> _____ (1 - 9)

Hour ==> ____ - ____ - ____ - ____ - ____ - ____ (0 - 23)

Dayname==> _____

Figure 2-17.03. Common Data Selection Screen - HP Openview UNIX Planning

The Common Data Selection screen (Figure 2-17.03) allows you to define the following data selection criteria:

- SYSID:

The SYSIDs that you want to include in your capacity planning database file. You can specify a maximum of 15 one- to four-character SYSIDs. Leave this parameter blank to include data for all SYSIDs.
- Zone:

Specify up to nine zones to include in the capacity planning database file. Leave this parameter blank to include data for all zones.

Hour: Specify the five hours or hour ranges to include in the capacity planning database file. Selection of this parameter is valid only if you specify the DETAIL, DAYS, or WEEKS timespan on the HP Openview UNIX Resource screen (Figure 2-17.01). Leave this parameter blank to include data for all hours.

Dayname: Specify up to five daynames to be included in the capacity planning database file. Leave this parameter blank to include all days. Input the first unique characters and the application will fill in the rest. This parameter is only valid when the input timespan is DETAIL or DAYS.

```

/----- SAS System Statements ----- COLUMNS 001 072--\
|Command ==>                               Scroll ==> PAGE |
|
|SAS Statements for the SELECT Macro          |
|-----|
|***** ***** TOP OF DATA *****|
|=MSG> SPECIFY SAS STATEMENTS TO SUBSET THE CAPACITY PLANNING DATA BASE FILE. |
|=NOTE= %MACRO SELECT; <=== PLEASE DO NOT PLACE LINES BEFORE THIS LINE.      |
|000001                                     |
|000002                                     |
|000003                                     |
|000004                                     |
|000005                                     |
|000006                                     |
|000007                                     |
|000008                                     |
|000009                                     |
|000010                                     |
|=NOTE= %MEND SELECT; <=== PLEASE DO NOT PLACE LINES AFTER THIS LINE.        |
|***** ***** BOTTOM OF DATA *****|
|
\-----/

```

Figure 2-17.04. SAS System Statements Screen - HP Openview UNIX Planning Global Data

The SAS System Statements screen (Figure 2-17.04) lets you define your own data selection logic. You can code any valid SAS statements as described in the SAS manuals. The %MACRO SELECT and %MEND SELECT statements are already provided and are field protected to prevent modification.

Note: This is a PDF/EDIT screen. All the PDF/EDIT commands are available.

2.17.3 Standard Output

The capacity planning database file that is generated by the HP Openview UNIX Resource Planning Sample Application contains the following data elements.

- A. Data elements common to all CA MICS Capacity Planner database files:

SYSID
ZONE or HOUR or DAYNAME
YEAR
MONTH or WEEK or MNTHDAY
DATE

- B. Data elements unique to the HP Openview UNIX Resource Planning Sample Application database file:

NOTE: The availability of some of these data elements depends on your operating environment. For a description of the data elements available for various UNIX platforms supported by CA MICS, see the CA MICS Analyzer Option for MeasureWare Guide.

1. Character format data elements

CAPAPU - CAPAPU9 - Workload Group

2. Numeric format data elements

AUMSAMPL - Number of data samples
AUMCPMDL - System model identification
GPRAPRCT - Sum of average processes active
GPRLCOLS - LAN collisions
GPRDSKIO - Physical disk I/Os
GPRPFLCT - Number of memory page faults during the interval
GPRSWPCT - Number of process swaps during the interval
GPRLDKCT - Number of logical disk I/Os during the interval
APPDSKIO - Number of physical disk I/Os

3. Time format data elements

DURATION - Time in the measurement interval
GPRCPUTM - CPU use during the interval
GPRSYSTM - System CPU use during the interval
GPRUSRMTM - User CPU use during the interval
APPCPUTM - Application's CPU use during the interval
APPUSRMTM - User CPU use during the interval
APPSYSTMTM - CPU use for system processing

The data elements that are derived from the CA MICS database and used by the HP Openview UNIX Resource Planning Sample Application are calculated as follows:

CAPAPU: The data element that contains the values found which will be either AUMAPU, values derived within the user's exit, or no CAPAPU value.

TOTCPUTM: Total Workload application's CPU time consumed. The total CPU time consumed by all workload applications on the system. This is the sum of APPCPUTM across the workloads.

TOTDSKIO: Total Workload disk transfers. The total number of disk transfers by all workloads on the system. This is the sum of APPDSKIO across the workloads.

PCTOTCPU: Workload percentage of total application CPU. The percentage of total application CPU time consumed by the workload. It is calculated using the following formula:

$$\text{PCTOTCPU} = 100 * (\text{APPCPUTM} / \text{TOTCPUTM})$$

PCTOTDSK: Workload percentage of total disk transfers. The percentage of total disk transfers consumed by the workload. It is calculated using the following formula:

$$\text{PCTOTDSK} = 100 * (\text{APPDSKIO} / \text{TOTDSKIO})$$

2.18 Tape Data Set Resource Planning

The Tape Data Set Resource Planning Sample Application provides a capacity planning database file that contains the necessary elements for tracking, forecasting, and reporting on tape data sets.

The following sections cover the use, input specifications, and output of the sample application:

- 1 - Usage Guidelines
- 2 - Control Parameters
- 3 - Standard Output

2.18.1 Usage Guidelines

DATA SOURCE

The Tape Data Set Resource Planning Sample Application processes data from the CA MICS Tape Analyzer Option. The TLM component provides IT management information for tape data set ownership and usage. The TMCTMD file provides a daily snapshot of tape data set characteristics and is used as input into the sample application. This file contains detailed information about each tape data set managed by the tape management system. At summarized timespans, the data set name and volser are dropped, but the summarization key provides enough granularity to track and forecast tape data sets at the Tape Library Management (TAPELIB) and Volume Group (VOLGROUP) levels.

The capacity planner sample application creates a capacity database containing a subset of CA MICS elements that reside in the TMCTLM file. With this database, an analyst can track, forecast, and report on tape-related characteristics like these:

- o Tape Data Set Size
- o Tape Data Set Age
- o Last Used in Days
- o Encryption
- o Space Charges

CAPAPUs

The application workload group (CAPAPU) data elements let you group your individual tape library systems or volume groups into logically similar workloads for meaningful tracking and forecasting. By default, CAPAPU is set to TAPELIB, and CAPAPU2 is set to VOLGROUP. You have the ability to redefine all CAPAPUs using the CAPAPU Derivation SAS Statements option on the Extended Options Menu.

COMPOSITE HOUR

Raw data for the TLM component is collected by the TMCC collector. The best practice is to run a TMCC raw data collection once daily to provide a daily snapshot of the tape data sets managed by the library. Because of this, the Composite Zone/hour/day option defaults to Y. This summarizes all zones, hours, or day names into a single observation, masking the time at which the TMCC collector ran.

2.18.2 Control Parameters

Relatively few parameters are needed to implement the Tape Data Set Resource Planning Sample Application. The control screens included in the online facility let you quickly enter the required information.

```

/----- Tape Data Set Planning -----\
|Command ==>|
|Composing CAP DB File def.: TLM - Physical Tape Data Set vs Virtual|
|Summarization Level|
|      Primary      ==> MONTHS  (DAYS/WEEKS/MONTHS)|
|      Secondary    ==> ZONE    (ZONE/HOUR/DAYNAME)|
|Input File Timespan      ==> MONTHS  (DETAIL/DAYS/WEEKS/MONTHS)|
|Amount of data retained  ==> 60      (1-9999 months)|
|Age of oldest forecast   ==> 24      (1-9999 months)|
|History cycles           ==> 3       (1-99 cycles)|
|Forecast cycles          ==> 3       (1-99 cycles)|
|Database ID (s)          ==> T
|
|      First Application Unit (CAPAPU ) ==> TAPELIB  ( Use Extended Options )
|      Second Application Unit (CAPAPU2) ==> VOLGROUP ( to re-define CAPAPUs )
|Specify extended options      ==> N      (Y/N)  ** Specified **
\-----/

```

Figure 2-18.01. Tape Data Set Planning Screen

Figure 2-18.01 shows the Tape Data Set Planning screen, which is the main control screen for this sample application. It contains the following fields:

Summarization level: The output summarization corresponds to the combination of the primary and secondary summarization levels. These control how the output file is summarized in your capacity planning database.

Input file timespan: The timespan (either DETAIL, DAYS, WEEKS or MONTHS) of the AUMGPR and AUMAPP files that you want to use as input when creating the capacity planning database file.

Amount of data retained:	The number of days, weeks, or months of historical data to be retained in the CAPxxx01 cycle of the capacity planning database file. Any historical observations older than what you specify for this parameter are dropped from the file during the file update.
Age of oldest forecast:	The number of days, weeks, or months that you want to retain a forecast after it is generated. Any forecast older than what you specify for this parameter is dropped when a new forecast is generated and saved.
History cycle:	Number of cycles or versions of the historical data you want to retain. A new cycle is created when the file is updated and the existing cycles aged (for example, Cycle 01 is aged to Cycle 02 and Cycle 02 becomes Cycle 03).
Forecast cycle:	Number of cycles or versions of the forecast data you want to retain. A new cycle is created when you specify forecast aging in one of the workload forecasting routines and the existing cycles are aged (for example, Cycle 01 is aged to Cycle 02 and Cycle 02 becomes Cycle 03).
Database ID(s):	The database unit identifiers (up to 14 units) that contain the TMCTMD file to be used with this sample application. If you are not sure where the TMCTMD file resides, type a question mark (?) in this field to display a listing of the valid database unit identifiers for this sample application.

```

Specify      If you want extended options, specify Y (YES),
extended    and the Extended Options Menu (Figure 2-17.02)
options:    is displayed.

/----- Extended Options Menu -----\
|Command ==>
|
|Inquiry Step:  Tape Data Set Planning
|
|
|Combine all zones into a composite ==> Y      (Y/N)
|  Composite zone                ==> A
|
|Specify:
|  Common data selection parameters ==> N      (Y/N/R)
|  Global data selection exit       ==> N      (Y/N/R)
|  CAPAPU derivation SAS statements ==> N      (Y/N/R)
|  Notes report enabled             ==> YES    (YES/NO)
|  Contents reports enabled         ==> YES    (YES/NO)
|
\-----/

```

Figure 2-18.02. Tape Data Set Planning Extended Options Menu

The extended options available with the Tape Data Set Resource Planning Sample Application (Figure 2-18.02) are the following:

Combine all zones/hours/daynames into a composite: You can combine zones, hours, and days in the capacity planning database file. The prompt for this field changes to zones, hours, or days, depending on what you specify for the summary-level parameter. Enter Y (YES) if you want all of the zones, hours, and days to be summarized into a single observation with the composite value. By default this option is set to Y.

Specify common data selection parameters: Specify the SYSIDs and zones, hours or days to be used for data selection from the CA MICS input files. Enter Y (YES) to display the Common Data Selection screen (Figure 2-17.03). Enter R (RESET) to reset any selections you previously made to SYSID and ZONE, HOUR, or DAYNAME.

Specify global data selection exit:	You can use your own SAS statements to select data from the CA MICS input files. Enter Y (YES) to display the SAS System Statements screen (Figure 2-17.04). Enter R (RESET) to reset any selections you previously made to the default of no user-specified data selection.
Specify CAPAPU derivation SAS statements:	You can use your own SAS statements for defining your application units. Enter Y (YES) to display the SAS System Statements screen. Enter R (RESET) to reset any SAS System Statements you previously made to a user-specified workload grouping.
Notes report enabled:	A YES in this field prints any notes associated with the capacity planning file definition as part of file updating.
Contents reports enabled:	A YES in this field prints the contents of the capacity planning file as part of file updating.

```

/----- Common Data Selection -----\
|Command ==>
|
|Inquiry Step:  Workload Planning Sample Application
|
|SYSID ==>  _ _ _ _ _
|           _ _ _ _ _
|
|Zone  ==>  _ _ _ _ _ (1 - 9)
|
|Hour  ==>  _ _ - _ _ - _ _ - _ _ - _ _ (0 - 23)
|
|Dayname==>  _ _ _ _ _
|
|
|
|
\-----/

```

Figure 2-18.03. Common Data Selection Screen - Tape Data Set Planning

The Common Data Selection screen (Figure 2-17.03) lets you define the following data selection criteria:

- SYSID:** The SYSIDs that you want to include in your capacity planning database file. You can specify a maximum of 15 one to four-character SYSIDs. Leave this parameter blank to include data for all SYSIDs.
- Zone:** Specify up to nine zones to include in the capacity planning database file. Leave this parameter blank to include data for all zones.
- Hour:** Specify the five hours or hour ranges to include in the capacity planning database file. Selection of this parameter is valid only if you specify the DETAIL, DAYS, or WEEKS timespan on the Tape Data Set Resource screen (Figure 2-18.01). Leave this parameter blank to include data for all hours.

Dayname: Specify up to five days to be included in the capacity planning database file. Leave this parameter blank to include all days. Input the first unique characters and the application will fill in the rest. This parameter is only valid when the input timespan is DETAIL or DAYS.

```

/----- SAS System Statements ----- COLUMNS 001 072--\
|Command ==>                               Scroll ==> PAGE |
|
|SAS Statements for the SELECT Macro          |
|-----|
|***** ***** TOP OF DATA *****|
|=MSG> SPECIFY SAS STATEMENTS TO SUBSET THE CAPACITY PLANNING DATA BASE FILE. |
|=NOTE= %MACRO SELECT; <=== PLEASE DO NOT PLACE LINES BEFORE THIS LINE.      |
|000001                                     |
|000002                                     |
|000003                                     |
|000004                                     |
|000005                                     |
|000006                                     |
|000007                                     |
|000008                                     |
|000009                                     |
|000010                                     |
|=NOTE= %MEND SELECT; <=== PLEASE DO NOT PLACE LINES AFTER THIS LINE.        |
|***** ***** BOTTOM OF DATA *****|
|
\-----/

```

Figure 2-18.04. SAS System Statements Screen - Tape Data Set Planning Global Data

The SAS System Statements screen (Figure 2-17.04) lets you define your own data selection logic. You can code any valid SAS statements as described in the SAS manuals. The %MACRO SELECT and %MEND SELECT statements are already provided and are field-protected to prevent modification.

Note: This is a PDF/EDIT screen. All the PDF/EDIT commands are available.

2.18.3 Standard Output

The capacity planning database file that is generated by the Tape Data Set Resource Planning Sample Application contains the following data elements.

- o Data elements common to all CA MICS Capacity Planner Option database files:

SYSID
 ZONE or HOUR or DAYNAME
 YEAR
 MONTH or WEEK or MNTHDAY
 DATE

- o Data elements unique to the Tape Data Set Resource Planning Sample Application database file:

- Character format data elements

CAPAPU - CAPAPU9 - Tape Workload Group

- Numeric format data elements

TMDURSP - Amount of Space-Time MB per Hour
 TMDSSZM - Approx. Data Set Size MB
 TMDAVSZM - Avg. Approx. Data Set Size
 TMDBLKCT - Block Count
 TMDDSNCT - Data Set Count
 TMDSEGS - Data Set Segments
 TMDAVSEG - Avg. Data Set Segments
 TMDCOST - Space Charges
 TMDAGE - Data Set Age (Days)
 TMDMXAGE - Max Data Set Age (Days)
 TMDAVAGE - Avg. Data Set Age (Days)
 TMDLUSE - Days Since Last Used
 TMDMXLUS - Max Days Since Last Used
 TMDAVLUS - Avg. Days Since Last Used
 TMDNCRYP - Encrypted Data Set Count

The data elements that are derived from the CA MICS database and used by the Tape Data Set Resource Planning Sample Application are calculated as follows:

CAPAPU: The data element that contains the values found, which will be either TAPELIB or values derived within the user's exit.

CAPAPU2: The data element that contains the values found, which will be either VOLGROUP or values derived within the user's exit.

Chapter 3: CAPACITY PLANNING DATABASE

The CA MICS Capacity Planner enables analysts to create a capacity planning database that is derived from, but independent of, the CA MICS database. This database satisfies the long-term historical requirements for forecasting workloads and resource utilization without retaining unnecessary data.

The most fundamental technique employed by capacity planners is the analysis of historical data describing workload resource consumption, system service levels, and system component utilizations. Although the CA MICS database contains a wealth of these data elements, the collection of such an historical series from multiple cycles of the CA MICS database (while operationally trivial) requires significant CPU resources. Moreover, it may be infeasible to retain a significant number of cycles of the DETAIL, DAYS, WEEKS, and/or MONTHS timespans for capacity planning because of the DASD space required. To minimize these resource requirements, the concept of a specialized capacity planning database was developed.

We supply interactive screens with the CA MICS Capacity Planner that allow you to define any number of capacity planning files within the capacity planning database. These files contain the various data elements that the capacity planner finds useful and has selected and/or computed from a variety of files in the CA MICS database. These files are then used as input to Presentation Graphics, Profile and Trending Analysis, Simple Regression Analysis, Univariate Model Forecasting, Capacity Planning Worksheet, Multivariate Regression Forecasting, and Business Element Forecasting facilities of the CA MICS Capacity Planner. These facilities are discussed in Chapters 4 through 10.

The following sections describe the operation and reporting facilities of the CA MICS Capacity Planner Products's database system.

This section contains the following topics:

- [3.1 Structural Description](#) (see page 314)
- [3.2 Usage Guidelines](#) (see page 317)
- [3.3 Resource Element Files](#) (see page 321)
- [3.4 Meta Files](#) (see page 446)
- [3.5 Business Element Files](#) (see page 511)

3.1 Structural Description

Typically, Capacity Planning requires you to retain far fewer data elements than what CA MICS database maintains; however, it requires many more cycles of the data elements than CA MICS normally maintains. The CA MICS Capacity Planner lets you develop composite files, known as resource element files and business element files, that merge the necessary data elements from many CA MICS files into a single Capacity Planning file. A resource element file is a SAS data set that contains data elements you select from CA MICS files or from user-defined data elements that are computed from CA MICS data elements. A business element file is a SAS data set that contains user-specified and supplied data elements from sources other than CA MICS files.

The CA MICS Capacity Planner provides a bridge between the capacity planner and the CA MICS database. The files you produce require a minimum of DASD space for retaining a great number of historical observations. Moreover, the resources required to process these files are minimal.

With the application of CAP6480 first delivered with r12 SP3 (PSP0804), the Capacity Planner began implementing SAS compression on newly created historical Capacity Planner database cycles.

An internal SAS macro variable CAPCMPRS set to YES has been added to shared.prefix.SOURCE(\$CAPMSTR). With this variable set to yes, any capacity database cycle created either during an initialization of a new database, or during an update to an existing database, is created with SAS compression active. This method will compress the new 01 cycle of the historical database. During subsequent database updates the database aging routine will copy the compressed 01 cycle to the 02 cycles deleting the highest cycle in retention, gradually over time compressing all files of the capacity database.

In some cases, a compressed SAS file may take more space than if it were not compressed. To determine your compression savings for each newly created database file, review your SASLOG from either the initial capacity database create or the subsequent database update jobs.

In your review, consider that capacity database compression is active for all historical files in the entire capacity database, so the best way to determine if compressing your capacity database saves you space is to compress the entire capacity database. Performing this function will provide compression statistics for the entire capacity file.

If you prefer to compress all previously created private capacity database history files use the JCL provided in the Backup, Restore and Compression section of this guide.

When you create a resource element or business element file, you supply a three-character file identifier as one of the parameters. This identifier generates unique SAS data set names for the SAS data sets that you need to support a particular study (using the analysis programs documented in Chapters 4 through 10). Moreover, the presence of this identifier in the SAS data set names lets you conduct multiple studies simultaneously.

When you create resource element, meta, or business element files, three groups of data sets are generated in the CA MICS Capacity Planning database:

- o A file or set of files containing the historical observations called CAPxxxnn, where xxx is the file identifier and nn denotes a cycle number. For business element files, the CAPxxxnn files will contain all of the observations that you provide.
- o A file or set of files containing the forecasted observations called CPFxxxnn, where xxx is the file identifier and nn denotes a cycle number. These observations are generated by the Univariate Modeling, Business Element Forecasting, and Multivariate Regression Forecasting routines of the CA MICS Capacity Planner. These files do not exist until you run one of these routines and specify that forecasts are to be saved and/or aged.
- o A file or set of files containing worksheet observations called CPWxxxnn, where xxx is the file identifier and nn denotes a cycle number. These observations are generated by the Worksheet Facility that is part of the CA MICS Capacity Planner. These files do not exist until you run a Worksheet Apply routine.

Figure 3-1 provides a schematic diagram of the data sets generated by the CA MICS Capacity Planner. You specify the number of cycles (nn) when the files are defined.

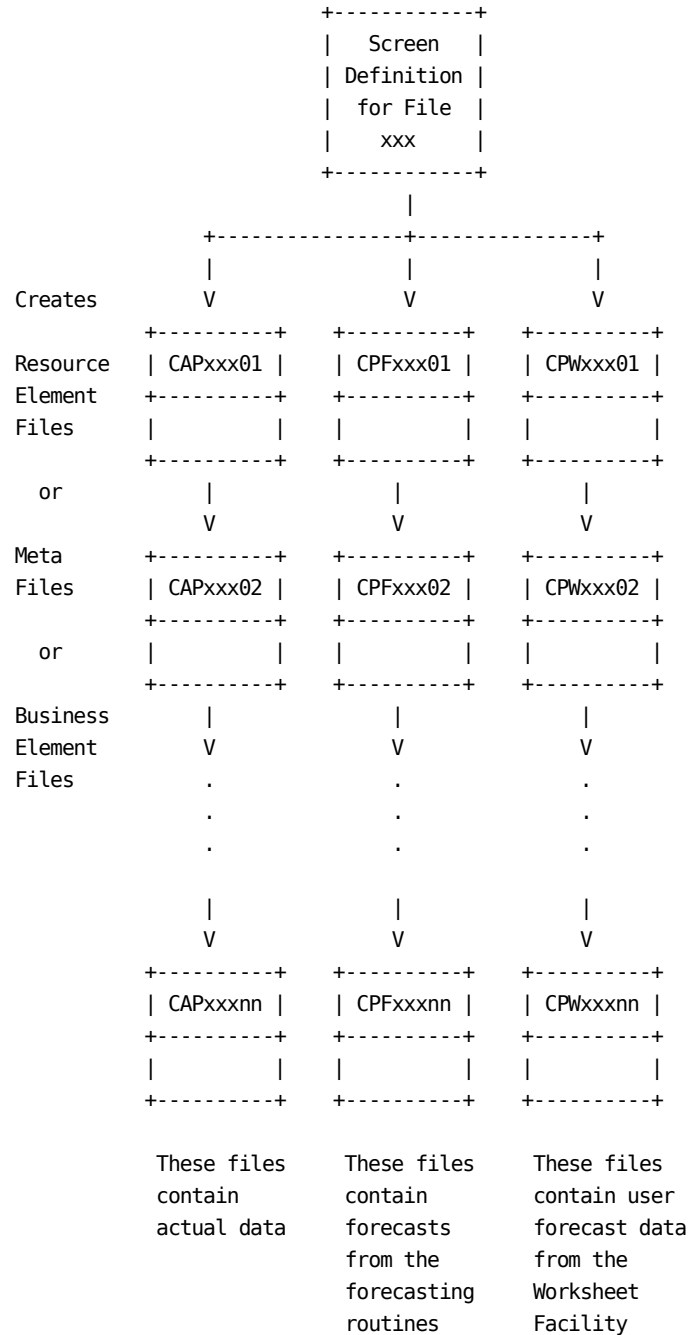


Figure 3-1. Resource, Meta, and Business Element File Creation

3.2 Usage Guidelines

Resource element files contain data elements you select from CA MICS files or user-defined data elements computed from CA MICS data elements. For example, you might construct a hardware utilization profile of the system that contains CPU, LCU, expanded storage, and device activity data. Since an important issue in workload forecasting is the tracking and verification of previous forecasts, you could obtain a similar set of observations some weeks or months later to use in evaluating the previous forecast.

Meta files contain data elements you select from already defined Capacity Planner files or user-defined data elements computed from other data elements. For example, you might create a meta file that combines information from two different standard application files, containing information from the MVS/ESA Standard Application in conjunction with data from the VAX/VMS CPU Planning Standard Application.

Business element files contain user-specified and supplied data elements from sources other than CA MICS files. For example, you might construct a workload profile of the system containing the number of general ledger accounts, the number of accounts updated, the number of reports generated, etc. You can directly specify forecasts of future workloads or they can be generated by the CA MICS Capacity Planner's forecasting tools.

Since different versions of each of these files can be created over a period of time, the product creates a file definition from the parameters you supply on the control screens. This definition is stored as a file definition inquiry in a standard MICF catalog. You can create additional copies or variations of the file by repeating, modifying, and executing the cataloged inquiry. This eliminates the need to retype all of the control parameters each time you create another file.

Figure 3-2 presents a schematic representation of the creation of a capacity planning database resource element file. You can specify any number of standard CA MICS files to use as input in creating and updating the capacity planning database through parameters you specify on the online screens. You can identify the data elements that you wish to extract from each file, along with data selection criteria and the level of summarization of the data.

Each capacity planning database file can have, as part of the summarization sequence, up to nine, user-defined data elements, CAPAPUs, which are the application workload units. You can use these data elements to provide installation-specific summarization of the data. The selected data elements from each of the specified CA MICS files are then merged together. You can also have user-defined data elements which are computed from the CA MICS data elements. The capacity planning database files serve as input to the various forecasting and analysis programs which are a part of this product.

	Primary	Secondary
Input File	Summarization	Summarization
TimeSpan	Level	Level
DETAIL	DAY	ENDTS/DAYNAME/HOUR/ZONE
DETAIL	WEEK	HOUR/ZONE
DETAIL	MONTH	HOUR/ZONE
DAYS	DAY	DAYNAME/HOUR/ZONE
DAYS	WEEK	HOUR/ZONE
DAYS	MONTH	HOUR/ZONE
WEEKS	WEEK	HOUR/ZONE
MONTH	MONTH	HOUR (if available)/ZONE

Note: When the secondary summarization level is END TIME STAMP, the files in which the data elements reside cannot be merged. For example, if you select DETAIL data with a summarization of ENDTS, you can only select one CA MICS file as input. Also note, ENDTS is a required data element.

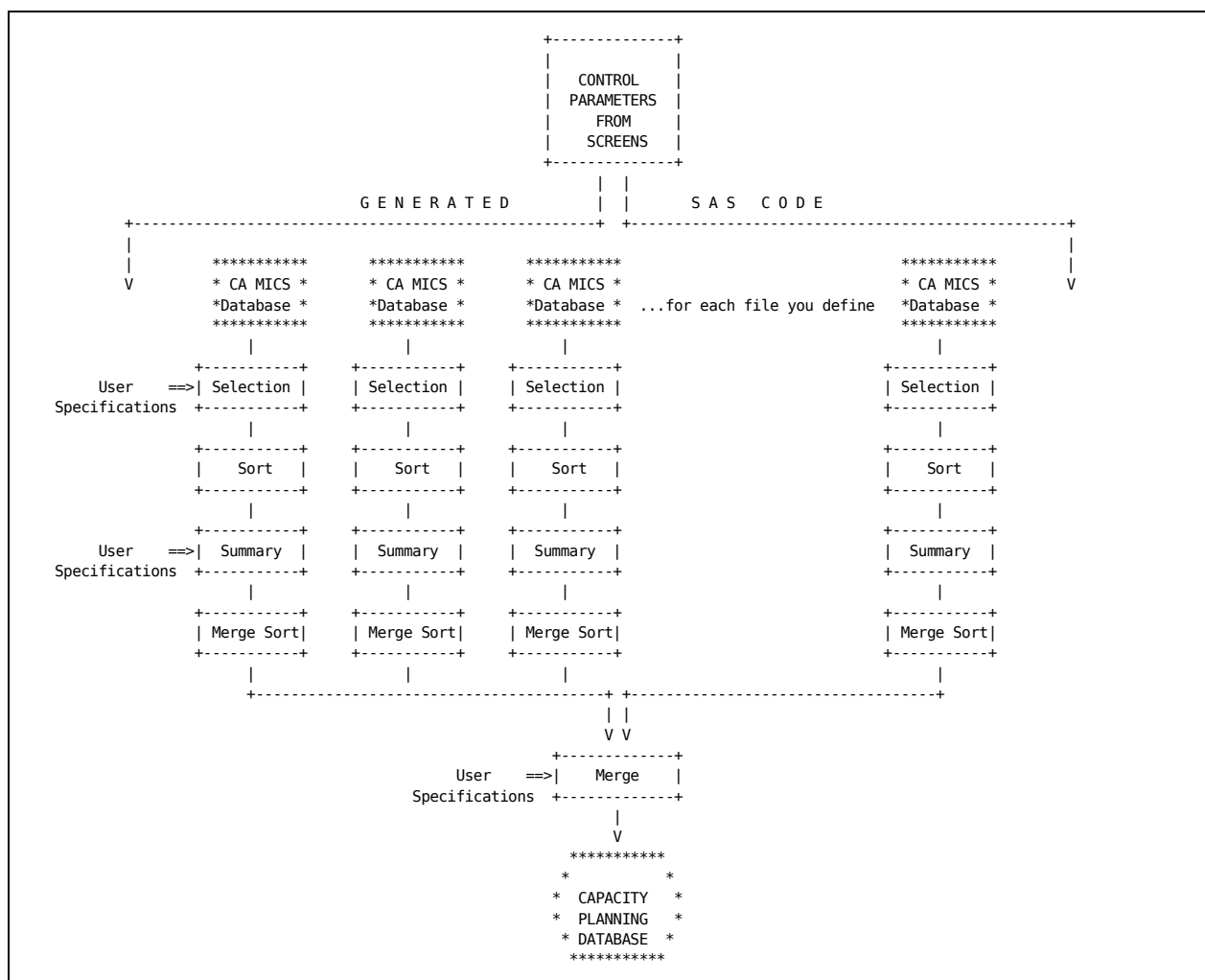


Figure 3-2. Capacity Planning Database Creation Flow Chart

The CA MICS Capacity Planner's online facility introduces the concepts of application workload unit and forecast age. These are important concepts to understand if you wish to take full advantage of the product's capabilities.

APPLICATION WORKLOAD UNIT (CAPAPU)

In many capacity planning applications you may find it useful to group observations from the input CA MICS files into logically similar application or workload groups. For instance, in tracking DASD use at your installation you might find it more meaningful to track the use of groups of volumes (for example, SYSDA volumes or TSO online volumes) rather than to track the use of each separate volume. You can perform this task through the application workload unit feature.

You can implement the application workload unit feature by adding up to nine, optional, 64-byte character data elements (CAPAPUs) to the capacity planning database historical and forecast files. This feature is implemented in your individual capacity planning database files only if you specify it. If you specify CAPAPUs, you are responsible for specifying the SAS statements to perform the application workload groupings and for defining the corresponding values of the CAPAPU data elements.

FORECAST AGE

For any numeric data element in your capacity planning database files, it is possible to generate and track multiple forecasts that are each based on a different historical range of data. The method that the capacity planning forecast files uses for identifying the separate forecasts is the inclusion of two data elements: FORYEAR, and FORMNDY, FORMONTH, or FORWEEK (depending on the primary summarization level of the historical data used to generate the forecast). These data elements identify the most recent historical observation used to generate the forecast. Each observation forecasted during a single execution of a workload forecasting routine will contain the same values for these two data elements.

For example, if, when using the DASD Planning Standard Application, you generate a six-month forecast for the BDPTOTL data element (total space used) that is based on historical data ranging from January 1997 through June 1998, the value of FORYEAR would be 98 and the value of FORMONTH would be 6 for each of the six forecasted observations.

If you then produce a second set of six-month forecasts based on historical data ranging from January 1998 through January 1999, the value of `FORYEAR` would be 99 and the value of `FORMONTH` would be 1 for each of these six forecasted observations. Both sets of forecasted observations are available in the `CPFxxx01` cycle for tracking and reporting through the Presentation Graphics routines of the CA MICS Capacity Planner.

Whenever you save new forecasts, the `CPFxxx01` cycle of your forecast files is scanned to determine if any forecasts existing in the file are older than the forecast retention value that you specified when creating the capacity planning database. If so, they are dropped from the file. If you specified that forecast aging was to occur, then these forecasted observations still exist in your older cycles of the forecast files. Aging of the 01 cycle into the 02 cycle occurs prior to this scanning and only the 01 cycle of your forecast files is scanned for forecast deletion.

3.3 Resource Element Files

The Resource Element Files section describes how to operate the CA MICS Capacity Planner with user-defined resource element files. The user-defined database files differ from the standard applications in that you are responsible for selecting the data elements that you wish to extract from the CA MICS database and also for supplying the SAS statements that are necessary to derive any non-CA MICS data elements that you wish to include in the file.

The following sections describe the screens that control this facility:

- 1 - Usage Guidelines
- 2 - Screen Operation
- 3 - Resource Element File Case Studies

3.3.1 Usage Guidelines

The resource element files hold subsets of CA MICS file data element values extracted from either the DETAIL, DAYS, WEEKS or MONTHS timespan. You select the data elements for your resource element file. Like business element files, there can be any number of resource element files. A three-character user-supplied name identifies each of the resource element file definitions.

Refer to Figure 3-1 to see the relationship between resource element and business element files. The number of cycles (or copies) of the files maintained is represented by nn. These cycles have the following values:

- Cycle 01: Data from the last file load or update plus any data from the previous Cycle 01 that is not older than allowed by the parameter Amount of data retained on the Resource Element file Definition screen shown in Figure 3-7). The CAPxxx01 file contains user-supplied observations while the CPFxxx01 file contains the forecasted observations from the workload forecasting routines. The CPFxxx01 file is not created until you generate and save a forecast for the first time.
- Cycle 02: Data that is contained in the Cycle 01 file when it is aged. Aging occurs before the Cycle 01 file is loaded or updated. The Cycle 02 file acts as a backup copy of the Cycle 01 file as it existed before the last file update. This cycle and all succeeding cycles may contain data older than allowed in the Cycle 01 file because of the Amount of data retained parameter. Aging of the CPFxxxnn files (forecast data files) does not occur until you explicitly specify forecast file aging in one of the workload forecasting routines.
- Cycle 03 through Cycle nn: Data from the Cycle xx-1 file just before it was aged. This occurs each time the Cycle 01 file is updated.

If, in September 1997, another load of four months of CA MICS data is requested, the sequence of execution would be as follows: First, the data in the Cycle 01 file is aged (copied) into the Cycle 02 file. Next, the existing Cycle 01 file is updated with the additional four months of data and any observations in the new Cycle 01 file that are older than the maximum data retention value you specified are dropped. Figure 3-3 shows the status of the files.

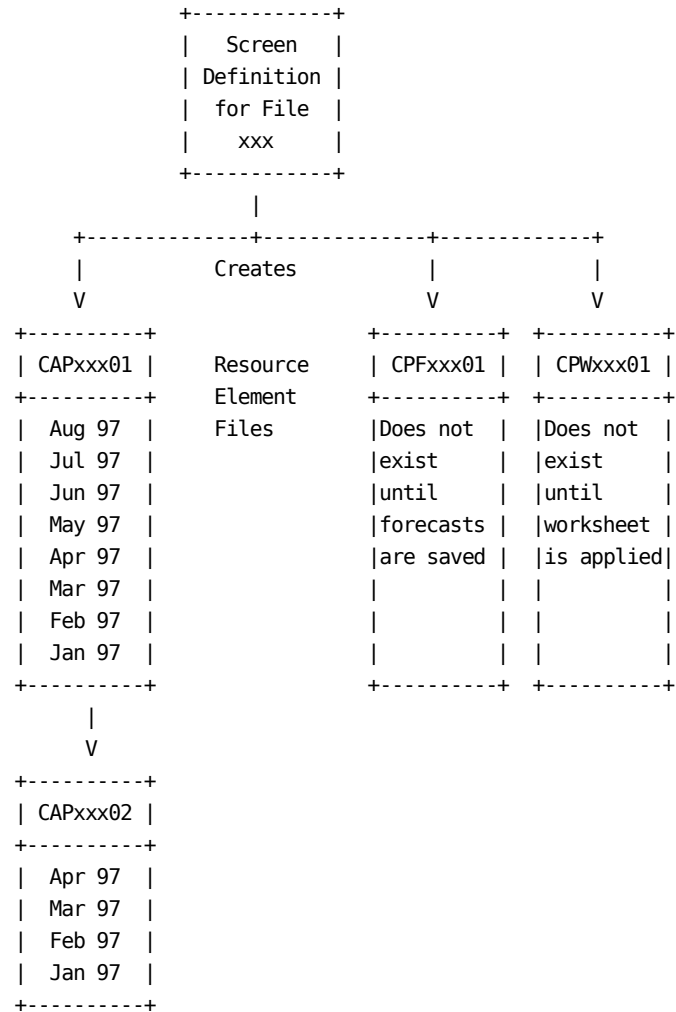


Figure 3-4. Files after the Second Loading of Data


```
+-----+  
| CAPxxx03 |  
+-----+  
| Apr 97 |  
| Mar 97 |  
| Feb 97 |  
| Jan 97 |  
+-----+
```

Figure 3-5. Files after the Third Loading of Data

If you load data a fourth time in May 1998, the old Cycle 03 file is dropped from the Capacity Planning database file because of the user-specified limit. Section 3.3.3.1.1 describes how to specify this limit.

+-----+			
Screen			
Definition			
for File xxx			
+-----+			
V			
+-----+			
V	Creates	V	V
+-----+			
CAPxxx01	Resource	CPFxxx01	CPWxxx01
+-----+			
Apr 98	Element	Does not	Does not
Mar 98	Files	exist	exist
Feb 98		until	until
Jan 98		forecasts	worksheet
Dec 97		are saved	is applied
Nov 97			
Oct 97			
Sep 97			
Aug 97			
+-----+			
V			
+-----+			
CAPxxx02			
+-----+			
Dec 97			
Nov 97			
Oct 97			
Sep 97			
Aug 97			
Jul 97			
Jun 97			
May 97			
Apr 97			
Mar 97			
Feb 97			
Jan 97			
+-----+			
V			
+-----+			
CAPxxx03			
+-----+			
Aug 97			
Jul 97			
Jun 97			
May 97			
Apr 97			
Mar 97			
+-----+			


```

| Feb 97 |
| Jan 97 |
+-----+

```

Figure 3-6. Files after the Fourth Loading of Data

Note throughout this example that forecast files do not exist until you save forecasts. Forecast data is created and aged in the same way as actual data. Note that the historical files (CAP files) and the forecast files (CPF files) are not aged together. The CAP files are aged whenever you perform a file update. You can age the CPF files only when you request aging through a parameter on the workload forecast routine control screens.

Throughout this example that worksheet files do not exist until you applied one. Worksheet data is created and aged in the same way as actual data. Note that the historical files (CAP files) and the worksheet files (CPW files) are not aged together. The CAP files are aged whenever you perform a file update. You can age the CPW files when you apply the worksheet to the database.

In addition to the data elements defined for the resource element files containing the derived from CA MICS files (the CAPxxxxnn files), the forecast files (CPFxxxxnn files) contain three additional data elements:

- FILE_NO:** The types of forecast data from the workload forecasting routines. If FILE_NO is 11, the data values are point estimates. If FILE_NO is 12, the data values are confidence limits.
- FORYEAR:** The year value of the most recent historical observation used to derive the forecast. For example, if you specify to Univariate Model Forecasting that a forecast is to be generated based on data from June 1997 through October 1998, the value of FORYEAR would be 98. This data element containing this value would be present on each of the observations that are forecasted during that execution of Univariate Model Forecasting.

FORMONTH/	The month, week, or day value of the most recent
FORWEEK/	historical observation used to derive the
FORMNDY:	forecast. For example, if you specify to
	Univariate Model Forecasting that a monthly
	forecast is to be generated based on data from
	June 1997 through October 1998, the value of
	FORMONTH would be 10. This data element
	containing this value would be present on each
	of the observations that are forecasted during
	that execution of Univariate Model Forecasting.

3.3.2 Screen Operation

The various online screens that the CA MICS Capacity Planner uses to create, modify, and load user-defined resource element files are described in the following sections:

- 1 - Creating a Database Specification
- 2 - Modifying a Database Specification
- 3 - Executing a Database Specification

3.3.2.1 Creating a Database Specification

The online screens that the CA MICS Capacity Planner uses to create user-defined resource element files are described in the following sections:

- 1 - File Definition
- 2 - Element Selection
- 3 - Specification Processing

3.3.2.1.1 File Definition

Figure 3-7 shows the first step for creating a user-defined resource element file.

```

/----- Resource Element File Definition -----\
|Command ==>|
|Enter a ? in any data entry field for more information on valid values.|
|CAP DB File def.: SYS - System Activity Tracking File|
|
|Summarization Level|
|Primary ==> MONTHS (DAYS/WEEKS/MONTHS)|
|Secondary ==> ZONE (ZONE/HOUR/ENDTS/DAYNAME)|
|
|Amount of data retained ==> 60 (1-9999 months)|
|Age of oldest forecast ==> 24 (1-9999 months)|
|History cycles ==> 3 (1-99 cycles)|
|Forecast cycles ==> 3 (1-99 cycles)|
|
|Enable:|
|Notes report ==> N (YES/NO)|
|Contents reports ==> N (YES/NO)|
|
|Specify extended options ==> Y (Y/N)|
\-----/

```

Figure 3-7. Resource Element File Definition Screen

The Resource Element File Definition screen shown in Figure 3-7 is the main control screen for defining resource element files. It contains the following fields:

Summarization level: The output summarization corresponds to the combination of the primary and secondary summarization levels. These control how the output file is summarized in your Capacity Planner database.

Note: When you choose a secondary summarization level of ENDTS, you can only select one CA MICS file as input. This is due to the difficulty in merging ENDTS values across multiple files.

Amount of data retained:	Number of days, weeks, or months of historical data to be retained in the CAPxxx01 cycle of the Capacity Planning database file. Any historical observations older than what you specify for this parameter are dropped from the file during the file update.
Age of oldest forecast:	Number of days, weeks or months you wish to retain a forecast after it is generated. Any forecast older than what you specify for this parameter is dropped when a new forecast is generated and saved.
History cycles:	The number of cycles or versions of the historical data that you wish to retain. A new cycle is created when the file is updated and the existing cycles are aged (that is, Cycle 01 becomes Cycle 02 and Cycle 02 becomes Cycle 03).
Forecast cycles:	The number of cycles or versions of the forecast data that you wish to retain. A new cycle is created when you specify forecast aging in one of the workload forecasting routines and the existing cycles are aged (that is, Cycle 01 becomes Cycle 02 and Cycle 02 becomes Cycle 03).
Notes report enabled:	If you have entered notes associated with the file definition, a Yes in this field prints the note as part of file updating.
Contents reports enabled:	A Yes in this field prints the contents of the capacity planner file as part of file updating.

Specify The specification for selecting extended
 Extended options. If you specify Y (YES), the Extended
 Options: Options Menu shown in Figure 3-8 displays.

```

/----- Extended Options Menu -----\
|Command ==>
|
|Inquiry Step:
|
|Combine SYSIDs into a composite ==> N      (Y/N)
|      Composite SYSID ==> _____
|Combine zones into a composite ==> N      (Y/N)
|      Composite zone ==> _____
|
|Number of CAPAPUS ==> 0      (0 through 9)
|CAPAPU title (First APU) ==> _____
|CAPAPU2 title ==> _____
|CAPAPU3 title ==> _____
|CAPAPU4 title ==> _____
|CAPAPU5 title ==> _____
|CAPAPU6 title ==> _____
|CAPAPU7 title ==> _____
|CAPAPU8 title ==> _____
|CAPAPU9 title ==> _____
|
|Specify
|  SAS Statements to set CAPAPUs ==> N      (Y/N/R)
|
\-----/

```

Figure 3-8. Resource Element Extended Options Menu

Combine all The specification for combining SYSIDs in the
 SYSIDs into Capacity Planning database file. Specify a Y
 a composite: (YES) if you wish data for all of the SYSIDs to
 be summarized into a single observation with the
 Composite SYSID value. Figure 3-28 illustrates
 specifying SYSIDs for inclusion in the database.

Combine all zones/hours/daynames into a composite:	The specification for combining zones/hours in the Capacity Planning database file. The prompt for this field changes to zones, hours or daynames depending on what you specify for the summary level parameter. Specify a Y (YES) if you want all of the zones/hours/daynames to be summarized into a single observation with the Composite zone/hour/dayname value. Figure 3-28 illustrates specifying zones/hours for inclusion in the database.
Number of CAPAPUs:	Specify the number (0 to 9) of application units (CAPAPUs) you are defining for this file. This activates the optional CAPAPU feature so that these data elements are present in the Capacity Planning database file, and are included in the summarization sequence during file summarization and merging. If you specify that CAPAPUs are to be active, you will be prompted to specify the SAS statements needed to derive the CAPAPU groupings and corresponding CAPAPU values.
CAPAPU titles:	Specify the 1-40 character title for each CAPAPU you are defining for this data element.

Specify SAS Statements to set CAPAPUs: The specification that allows you to use your own SAS Statements for defining your resource element file. Specify Y (YES) to display the SAS Statements screen shown in Figure 3-9. This option is only available if you specify CAPAPUs.

```

/----- SAS System Statements -----\
|Command ==>                               Scroll ==> CSR |
|
|SAS Statements for the CAPAURT Macro
|-----|
|***** ***** TOP OF DATA *****|
|==MSG> SPECIFY SAS STATEMENTS TO SET CAPAPU VALUES.
|=NOTE= %MACRO CAPAURT; <=== PLEASE DO NOT PLACE LINES BEFORE THIS LINE.
|000001
|000002
|000003
|000004
|000005
|000006
|000007
|000008
|000009
|000010
|=NOTE= %MEND CAPAURT; <=== PLEASE DO NOT PLACE LINES AFTER THIS LINE.
|***** ***** BOTTOM OF DATA *****|
\-----/

```

Figure 3-9. SAS Statements Screen - Defining CAPAPU

The SAS Statements screen shown in Figure 3-9 allows you to define the SAS statements required to group CA MICS input observations into application workload groups that are appropriate for your installation. All statements that you specify must conform with the syntax specified in the SAS manuals. The CAPAPU(s) value is a maximum of 64 characters.

Note that the %MACRO CAPAURT and %MEND CAPAURT statements are provided for you and that they are field-protected from modification.

A special data element, CAPIFILE, is available to you for use in defining your CAPAPU values. This three-character data element contains the CA MICS file suffix that corresponds to the CA MICS file that you choose for processing after you select the Insert CA MICS File and Data Element Selection option on the Resource Element File menu (see Section 3.3.2.1.2). For example, if you specify the Insert File and Data Element Selection option twice, once for the SCPPGA File and again for the SCPPPA File, the value of CAPIFILE will be PGA during the processing of the SCPPGA File and PPA during the processing of the SCPPPA File. The following example illustrates the use of the CAPIFILE data element in defining CAPAPU values.

You can rewrite the Performance Group/Period Workload Tracking example in Section 3.3.3.2 to use the CAPAPU feature by defining the following statements in the CAPAPU exit:

```
%MACRO CAPAURT;
  IF CAPIFILE EQ 'PPA' AND PERFGRP EQ 2 THEN DO;
    IF PERFPER EQ 1 THEN CAPAPU='WORKLOAD1';
    ELSE IF PERFPER EQ 2 THEN CAPAPU='WORKLOAD2';
    ELSE IF PERFPER EQ 3 THEN CAPAPU='WORKLOAD3';
    ELSE IF PERFPER EQ 4 THEN CAPAPU='WORKLOAD4';
    ELSE IF PERFPER EQ 5 THEN CAPAPU='WORKLOAD5';
  END;
  ELSE IF CAPIFILE EQ 'PGA' THEN DO;
    IF PERFGRP EQ 7 THEN CAPAPU='WORKLOAD6';
    ELSE IF PERFGRP EQ 62 THEN CAPAPU='WORKLOAD7';
    ELSE IF PERFGRP NE 2 THEN CAPAPU='WORKLOAD8';
  END;
%MEND CAPAURT;
```


3.3.2.1.2 Element Selection

Figure 3-10, Resource Element File menu, shows the next step in creating a user-defined resource element file definition.

```
/----- Resource Element File -----  
|Option ===>  
|  
|Composing CAP DB file def.: SYS - System Activity Tracking File  
|  
|    1 - Insert Common Data Selection  
|        Specify data selection criteria to apply to ALL files.  
|    2 - Insert CA MICS File and Data Element Selection  
|        Select a CA MICS file and data elements.  
|    3 - Insert Cross-File Data Element Definition  
|        Specify cross-file data element derivations.
```

Figure 3-10. Resource Element File Menu

After you complete specifying the resource element file definition control parameters, you are ready to define the steps to build your Capacity Planning database file. (Refer to Figure 3-2 to see how these steps correspond to user specifications for the selection, summary, and merge phases of data base creation.) The Resource Element File menu provides you with the opportunity to define selection criteria that you wish to apply to each CA MICS file that you define as an input data source for your Capacity Planning database file.

You may choose from the following options on the Resource Element File menu:

Insert Common Data Selection: Specify Option 1, Insert Common Data Selection, to specify the data selection criteria that you wish to apply to every CA MICS file that provides input data in building your Capacity Planning database file. The Common Data Selection screen shown in Figure 3-11 is displayed.

You may specify only one common data selection step for each database definition inquiry. You must complete this before proceeding with the succeeding steps.

Insert New MICS File & Data Element Selection: Specify Option 2, Insert CA MICS File and Data Element Selection, to define a CA MICS file and data element selection step for each CA MICS file that provides input data for your Capacity Planning database file. The CA MICS File and data Element Selection screen shown in Figure 3-13 is displayed. (Both the file-unique selection and summary phases shown in Figure 3-16 occur during this step.)

Using this screen and its associated exit screens, you can define both the CA MICS and the non-CA MICS data elements that will comprise your Capacity Planning database file, along with any file-unique data selection and/or summarization logic.

If you wish to define any non-CA MICS data elements that are derived from CA MICS data elements from a single CA MICS file, you can usually specify the data element derivation SAS statements through one of the exit screens associated with this screen. If you are defining non-CA MICS data elements that are derived from CA MICS data elements from multiple CA MICS files, you can specify the data element derivation SAS statements in the cross-file data element definition step.

Insert Cross-File Data Element Definition: Specify Option 3, Insert Cross-File Data Element Definition, to define any non-CA MICS data elements that you wish to derive from CA MICS data elements from multiple CA MICS files. The Cross-File Data Element Definition screen shown in Figure 3-21 is displayed. (This step corresponds to the merge phase shown in Figure 3-2.) You must also specify the data element derivation SAS statements that will be used to derive these non-CA MICS data elements.

Since the specifications associated with this step apply to the merging of data from multiple CA MICS files, you may specify a cross-file data element definition step only if you specify multiple CA MICS file and data element selection steps. In this case, the cross-file data element definition step is required to create a valid file definition inquiry.

Each time you complete the specifications for an individual step and type END at the Command prompt, then press ENTER or (usually) the PF3 key. The Resource Element File Menu shown in Figure 3-10 is redisplayed to allow you to insert another step in your file definition inquiry. Repeat this process until you define all of the steps necessary to build your Capacity Planning database file.

The three options are further discussed in the following sections:

- 1 - Insert Common Data Selection
- 2 - Insert CA MICS File and Data Element Selection
- 3 - Insert Cross-File Data Element Definition

3.3.2.1.2.1 Insert Common Data Selection

If you select Option 1, Insert Common Data Selection, on the Resource Element File menu, the Common Data Selection screen shown in Figure 3-11 is displayed.

```

----- Common Data Selection -----
|Command ==>
|Enter a ? in any data entry field for more information on valid values.
|Composing CAP DB File def.: SYS - System Activity Tracking File
|
|
|SYSID   ==> _____
|         ____ _
|Zone    ==> _ _ _ _ _ (1 - 9)
|
|Hour     ==> __ - __ - __ - __ - __ - __ - __ - __ (0 - 23)
|
|Dayname  ==> _____
|
|Specify common data selection SAS statements ==> N (Y/N/R)
-----
```

Figure 3-11. Common Data Selection Screen

The data selection criteria that you specify using this screen is applied to each CA MICS file that provides input data in creating and updating your Capacity Planning database file.

The parameters for the Common Data Selection screen are defined below:

SYSID: The SYSIDs that you wish to include in your Capacity Planning database file. You may specify a maximum of 15 one- to four-character SYSIDs. Leave this parameter blank to include data for all SYSIDs. The standard method is to create one record for each SYSID that is created for each day, week, or month processed.

Alternatively, you may wish to summarize multiple SYSIDs into a single record. You can do this by using the Composite SYSID parameter on the Resource Element File Definition screen shown in Figure 3-7.

- Zone:** The zones that you wish to include in your Capacity Planning database file. You may specify a maximum of nine one-character zones. Leave this parameter blank to include data for all zones. The standard method is to create one record for each zone that is created for each week or month processed. Alternatively, you may wish to summarize multiple zones into a single record. You can do this by using the Composite zone parameters on the Resource Element File Definition screen shown in Figure 3-7.
- Hour:** The hours that you wish to include in your Capacity Planning database file. You may specify a maximum of five hours or hour ranges. Leave this parameter blank to include data for all hours.
- Dayname:** The specification of up to five daynames to be included in the Capacity Planner database file. Leave blank to include all days. Input the first unique characters and the CA MICS Capacity Planner will fill in the rest. Specification of this parameter is only valid when the input timespan is DETAIL or DAYS.

Specify common data selection SAS statements:	Put a Y in this field to specify your own SAS statements to subset the CA MICS input files. The SAS Statements screen shown in Figure 3-12 displays. Specify R (Reset) to reset any SAS statements you previously made.


```

/----- SAS System Statements -----\
|Command ==>                               Scroll ==> CSR |
|
|Inquiry Step:  Common Data Selection          |
|
|SAS Statements for the SELECT  Macro          |
|-----|
|***** ***** TOP OF DATA *****|
|==MSG> SPECIFY SAS STATEMENTS TO SUBSET THE CAPACITY PLANNING DATA BASE FILE. |
|=NOTE= %MACRO SELECT; <=== PLEASE DO NOT PLACE LINES BEFORE THIS LINE.      |
|000001                                         |
|000002                                         |
|000003                                         |
|000004                                         |
|000005                                         |
|000006                                         |
|000007                                         |
|000008                                         |
|000009                                         |
|000010                                         |
|=NOTE= %MEND SELECT; <=== PLEASE DO NOT PLACE LINES AFTER THIS LINE.        |
|***** ***** BOTTOM OF DATA *****|
|
\-----/

```

Figure 3-12. SAS Statements Screen - Selecting Common Data

Using the SAS Statements screen shown above, you can specify your own data selection logic to apply to each CA MICS file that provides input data to your Capacity Planning database file.

Note that the %MACRO SELECT and %MEND SELECT statements are provided for you and that they are field-protected from modification.

3.3.2.1.2.2 Insert CA MICS File and Data Element Selection

If you select Option 2, Insert CA MICS File and Data Element Selection, on the Resource Element File menu, the CA MICS File and Data Element Selection screen shown in Figure 3-13 is displayed.

```

/----- CA MICS File and Data Element Selection -----\
|Command ==>                                         Scroll ==> CSR |
|Enter a ? in any data entry field for more information on valid values. |
|Modifying CAP DB File def.: R03 - sample resource element file          |
|CA MICS file      ==> CPU      (fff) - CPU ACTIVITY FILE                |
|Input Time-Span   ==> MONTHS (DETAIL/DAYS/WEEKS/MONTHS)                 |
|Data base (s)     ==> A                                                |
|Include CAPAPUs   ==> N (Y/N)                                           |
|Summarize file by ==> _____ | I Insert C Copy D Delete | |
|Specify file exits ==> N (Y/N)   | R Repeat E Calculations  | |
|                                   | S Depend Clauses & Label  | |
|                                   ----- | |
|                                   Sum Data | |
|Cmd  Name  Label (long name)              D Type Type Order Len Output-Format |
|-----|-----|-----|-----|-----|-----|-----|
|_  DATE    Interval End Date              RET  D   000   7   DATE9.          |
|_  MONTH   Month of Year                  RET  N   000   2   _____          |
|_  SYSID   System Identifier              RET  A   000   4   _____          |
|_  TEST    test                          COMP N   000   5   _____          |
|_  YEAR    Year of Century                RET  N   000   2   _____          |
|_  ZONE    Time Zone                      RET  A   000   1   _____          |
|***** BOTTOM OF DATA *****|
|
\-----/

```

Figure 3-13. CA MICS File and Data Element Selection Screen

The parameters for the CA MICS File and Data Element Selection screen are defined below:

CA MICS file	The CA MICS file that contains the data elements that you wish to include in your resource element file or that you wish to use in deriving non-CA MICS data elements. If you are uncertain of the file name, type a question mark (?) in this field to obtain a complete selection list for your installation.
Input File Time-Span	The timespan (DETAIL, DAYS, WEEKS, or MONTHS) of the file you are using as input.
Data base:	The database unit identifier(s) (up to fourteen units) that contain the specified CA MICS file. If you are uncertain of which database unit the CA MICS file resides in, type a question mark (?) in this field to obtain a complete selection list for your installation.
Include CAPAPUs:	The specification for whether the Application unit (CAPAPU) data elements are to be included in the summarization sequence for this file. This parameter defaults to Y (YES) if CAPAPUs are defined for this resource element file. The CAPAPU derivation exit is invoked for each input file observation. If you do not want to include CAPAPUs in the summarization sequence, type an N (NO) in this field.
Summarize file by:	The specification that allows you to add a data element to the intermediate file summarization sequence of this resource element file. This may be either a CA MICS or a non-CA MICS data element. Any data element that you specify as a summarization data element is dropped from the database file following the summarization phase. If you want to include a data element in the summarization sequence and retain it on the file, use the CAPAPU feature to hold the values of the data element.

Specify file exits: The specification for CA MICS file processing exits. You must specify a Y (YES) if you intend to code CA MICS file-unique data selection, summary, or data derivation logic. If you specify a Y, the CA MICS File Processing Exits Menu shown in Figure 3-16 is displayed.

The display portion of this screen contains a scrollable list of the CA MICS and non-CA MICS data elements that you can either include in your capacity planning database file or use to derive other data elements. It contains the following fields:

Cmd: The line commands are:

- I (insert) to insert a new data element.
- R (repeat) to repeat a data element.
- S (depend clauses and label) to specify independent data elements upon which this data element depends. See Figure 3-14.
- C (copy) to copy a data element.
- E (expression) to create SAS code to calculate the value of this specific data element. See Figure 3-15.
- D (delete) to delete a data element.

Name: The name of the data element. This may be either the name of a CA MICS data element that exists in the input CA MICS file or any valid SAS name, as described in the SAS manuals. If you wish to define non-CA MICS data elements, choose variable names that are descriptive of the values they represent.

Note that when you first see this screen, it displays only the control data elements that are always present in your capacity planning database file. If you attempt to modify these data elements, an error will occur. Define an additional row for each data element that you wish to include in your capacity planning database file by typing an I (INSERT) in the Cmd column of any existing row.

If you are uncertain of the correct spelling of a CA MICS data element or if you would like to see a complete list of the CA MICS data elements available in the CA MICS input file, type a question mark (?) in the name field to display the screen shown in Figure 3-20.

- Data Element Label (long name):** The data element label that is displayed on the CAP File and Data Element Selection screen and on many reports. This label is supplied automatically for input data elements.
- D:** The data element drop flag. Typing a D in this field makes the corresponding data element available for use in calculations through the cross-file data element definition step of the capacity planning database file update, but drops the data element from the resulting capacity planning database file. Leave the field blank to keep the corresponding data element on the capacity planning database file.
- Sum Type:** The summarization type of the data element: RET (retain), ACCM (accumulate), MIN (minimum), MAX (maximum), and COMP (compute).
- Data type:** The type of data element. This is used to validate data values and usage for the data element. Valid types are Alphanumeric (A), Numeric (N), Money (M), Time (T), Hex (H), Date (D), Time-of-day (TD), and Date-Time (DT). This value is supplied automatically for CA MICS data elements.
- Order:** Define the execution order of the data elements you defined by entering ordinal numeric values greater than 100.
- Note:** This is required if you are going to use this Resource Element file as input to a Meta file.

The order value is automatically defined for CA MICS data elements and Capacity Planner Control elements. Data elements that require other data elements are automatically included in the Capacity Planner database file. In particular, two variables could appear that will not show up in the output file: @@FIRST and @@LAST.

Len: The length of the data element. This value is supplied automatically for CA MICS data elements.

Output Format: The SAS format that is used to print data element values. This value is supplied automatically for CA MICS data elements.

```

/----- Resource Element File Definition -----\
|Command ==>
|
|Data Element
|   Name:      MYVAR   From File: CPU  CPU ACTIVITY FILE
|   Label ==> MY VARIABLE
|   Type ==> COMPUTE
|
|-----
|DEPEND 01==> CPUPCASY _____
|DEPEND 02==> _____
|DEPEND 03==> _____
|DEPEND 04==> _____
|DEPEND 05==> _____
|DEPEND 06==> _____
|DEPEND 07==> _____
|DEPEND 08==> _____
|DEPEND 09==> _____
|DEPEND 10==> _____
|
|
\-----/

```

Figure 3-14. Defining Resource Element File Dependencies and Labels

The Depend Clause and Label screen is used to define those variables in the input file that are required by the SAS statements (either in the CAPSMRY or in Expression Statements) to calculate the value of the data element being selected. For example, if you had a data element MYVAR that required the presence of CPUPCBSY to calculate its value, you would place CPUPCBSY on this screen, as shown in Figure 3-14. Once specified, the CA MICS Capacity Planner will carry forward these independent elements as required to recalculate the value of MYVAR.

```
/----- Expression Statements -----\  
|Command ==>|  
|  
|Expression statements for variable: MYVAR|  
|-----|  
|***** TOP OF DATA *****|  
|=NOTE= SET THE EXPRESSION STATEMENTS FOR THIS DATA ELEMENT.|  
|000001 IF CPUPCBSY LT 70.0|  
|000002 THEN MYVAR = 'OK';|  
|000003 ELSE IF CPUPCBSY GT 70.0|  
|000004     THEN MYVAR = 'WARN';|  
|000005     ELSE IF CPUPCBSY GT 85.0|  
|000006         THEN MYVAR = 'INVESTIGATE';|  
|000007         ELSE IF CPUPCBSY GT 95.0|  
|000008             THEN MYVAR = 'CONSTRAINED';|  
|000009|  
|000010|  
|000011|  
|***** BOTTOM OF DATA *****|  
\-----/
```

Figure 3-15. Setting Expression Statements

You use this screen (Figure 3-15) to set the Expression Statements for a particular variable name (in this case MYVAR). The CA MICS Capacity Planner will then use these SAS Statements to calculate the value for MYVAR. If you have already coded statements in a MACRO (either CAPSMRY or CAPMERG) you can use the Converting Macros to Expression Statements to perform this association.

[illegible]

Figure 3-16. CA MICS File Processing Exits Menu

Figure 3-16 shows the CA MICS File Processing Exits Menu that is displayed if you specify a Y (YES) for the Specify CA MICS file processing exits parameter on the CA MICS File and Data Element Selection screen shown in Figure 3-13.

The following options are available:

Specify file unique data selection SAS stmts:	The specification for assigning SAS statements to select file-unique data. Specify a Y (YES) if you wish to supply the SAS statements necessary to implement a special observation selection method.

Figures 3-17 and 3-18 show how you can use the SAS Statements screen to enter these SAS statements.

Figure 3-17. SAS Statements Screen - Selecting File Unique Data

Use the SAS Statements screen shown in Figure 3-17 to code valid SAS statements that conform to the syntax described in the SAS manuals to define CA MICS file-unique data selection logic.

```

/----- SAS System Statements -----\
|Command ==>                                Scroll ==> CSR|
|
|Inquiry Step:   File & Data Element Sel.: JOB - Batch User Job Activity File
|
|SAS Statements for the CAPSMRY Macro
|-----\
|***** ***** TOP OF DATA *****\
|==MSG> SPECIFY SAS STATEMENTS FOR TRANSPOSING THE FILE OR TO SET VARIABLES.
|NOTE= %MACRO CAPSMRY; <== PLEASE DO NOT PLACE LINES BEFORE THIS LINE.
|000001
|000002
|000003
|000004
|000005
|000006
|000007
|000008
|000009
|000010
|NOTE= %MEND CAPSMRY; <== PLEASE DO NOT PLACE LINES AFTER THIS LINE.
|***** ***** BOTTOM OF DATA *****\
|
|-----\

```

Figure 3-18. SAS Statements Screen - Deriving New Data Elements

Use the SAS Statements screen shown in Figure 3-18 to code valid SAS statements that conform to the syntax described in the SAS manuals to derive non-CA MICS data elements or summarization logic.

The CAPSMRY macro is used for several different purposes by the CA MICS Capacity Planner and it is very important to understand how this relates to the overall creation of a Capacity Planner file and, particularly, the creation of your own data elements.

You can specify user-created data elements and the type of data element you are creating. The data element types available are: Maximum, Minimum, Accumulated, Retained, and Computed. This type designation defaults to Retained for your data elements unless you override the setting, which may be any of these values. The data element type is used by the CA MICS Capacity Planner to determine how to process a particular user data element. To determine how to use the CAPSMRY for a particular data element please review the following criteria and take the appropriate action accordingly.

If you wish to use the Meta File Facility:

For data elements that have a data element type of computed you do not need to use the CAPSMRY exit. You should use Expression Statements to contain the SAS code needed to calculate the value of the data element. If you currently have files that were created with a release of the CA MICS Capacity Planner prior to CAP6100 and wish to convert a data element from the default of RETained to COMputed, there is an option on the extended options menu provided for this migration. Make sure that you change the Data Element Type for these variables (from RETained) because Expression Statements are only valid for COMputed variables.

For data elements that are not COMputed, you would specify the necessary SAS statements in the CAPSMRY Macro to set the initial value of the variable. If you are transposing the file, you should specify the SAS statements to perform this transformation here in the CAPSMRY. The data element type takes affect for non-computed variables only in the Meta File Facility.

If you do not wish to use the Meta File Facility:

If you do not wish to use the Meta File Facility, the data data element type is less important (the default of RETained is sufficient) and all variables should be set via the CAPSMRY exit. This is true whether or not you are transposing the file.

The following Data Element Type table describes the relationship between the CAPSMRY macro and expression statements.

Data Element Type	Used in Meta File?	Need SAS Code in CAPSMRY?	Need Expression Statements?
Maximum	No	Yes, to set the value	No
Maximum	Yes	Yes, to set initial value	No, handled in Meta Files automatically
Minimum	No	Yes, to set the value	No
Minimum	Yes	Yes, to set initial value	No, handled in Meta files automatically
Accumulated	No	Yes, to set the value	No
Accumulated	Yes	Yes, to set initial value	No, handled in Meta files automatically
Retained	No	Yes, to set the value	No
Retained	Yes	Yes, to set initial value	No, handled in Meta files automatically
Computed	No	Yes, to set the value	No
Computed	Yes	No	Yes, Expression statements used by both Resource Element File creation and in Meta files

```
/----- Converting Macros to Expression Statements -----\  
|Command ==>                                         Scroll ==> CSR |  
|  
|Enter the variable names that are associated with the SAS statements shown |  
|below.                                             |  
|  
|Variable  SAS Statements in Macro                  |  
|-----  
|MYVAR2__ IF CPUAVB GT 10  /* IF THE AVERAGE NUMBER OF BATCH JOBS */  
|MYVAR2__ THEN MYVAR2 = 'CHECK'; /* IS GREATER THAN 10 WE SHOULD  */  
|MYVAR2__ ELSE MYVAR2 = 'OK'; /* CHECK IT OUT */  
|***** BOTTOM OF DATA *****  
|-----/
```

Figure 3-19. Converting Macros to Expression Statements

You use this screen (see Figure 3-19) to convert existing MACRO statements into Expression Statements that calculate the value of a particular variable. In the example shown in Figure 3-19, the variable MYVAR2 is a user-created data element that is used to warn the user of this file whenever CPUAVB is larger than 10. This variable was originally created using the CAPSMRY MACRO. By placing "MYVAR2" in the Variable Column of this screen, the CA MICS Capacity Planner will automatically associate these SAS Statements with MYVAR2. When you End from this display, the CA MICS Capacity Planner will remove the statements from the MACRO and move them into the Expression Statement table for the variable MYVAR2.

```

/----- Data Element Selection -----\
|Command ==>                               Scroll ==> CSR |
|
|Inquiry Step: Common Data Selection          |
|Input File:  MONTHS JOB - BATCH USER JOB ACTIVITY FILE |
|
|Line Cmds: S Select  H Data Dictionary Display |
|
|          Select Data Elements
|Cmd  Element      Data Element Label (long name)
| -  - - - - - - - - - - - - - - - - - - - -
| _  ACCTN01  DIVISION
| _  ACCTN02  PROJECT
| _  ACCTN03  USER
| _  ENDTS    End Time Stamp
| _  JOBACOMM Communication Data Set Allocations
| _  JOBACTTM Job Active Time
| _  JOBADASD DASD Data Set Allocations
| _  JOBAGRAF Graphics Data Set Allocations
| _  JOBALCTM Job Allocation Time
| _  JOBAMSS  MSS Data Set Allocations
| _  JOBATAPE Tape Data Set Allocations
| _  JOBAUREC Unit Record Data Set Allocations
| _  JOBAVIO  VIO Data Set Allocations
|
\-----/

```

Figure 3-20. Data Element Selection Screen

If you enter a question mark (?) in the "name" field of the CA MICS File and Data Element Selection screen shown in Figure 3-13, the Data Element Selection screen (Figure 3-20), which is the final screen in this series, is displayed.

This screen presents a full list of data elements within the CA MICS file that you specified. You may select data elements to include in the resource element file by placing an S (SELECT) in the Cmd column that corresponds to the data element you want to include in your capacity planning database file.

3.3.2.1.2.3 Insert Cross-File Data Element Definition

If you specify data selection from more than one CA MICS file through Option 2 on the Resource Element File menu shown in Figure 3-10, you must also insert a cross-file data element definition step in your file definition. Select Option 3, Cross-File Data Element Definition, to define data elements that are calculated from CA MICS or user-defined data elements from more than one file. This screen displays all the data elements, including the control data elements, that you selected from all CA MICS files.

```

/----- Cross-File Data Element Definition -----\
|Command ==>                                     Scroll ==> CSR |
|Enter a ? in any data entry field for more information on valid values. |
|Modifying CAP DB File def.: R03 - sample resource element file          |
|                                                                           |
|Specify:                                     -----Line Commands-----|
| SAS exit stmts.   ==> N (Y/N/R)           | I Insert C Copy D Delete | |
|                                                                           | |
|                                     | R Repeat E Calculations  | |
| Expression stmts. ==> N (Y/N/R)           | S Depend Clauses & Label | |
|                                                                           | |
|                                     -----|
|                               Sum  Data
|Cmd  Name    Label (long name)    D Type Type Order Len Output Format|
|-----|
| _  MONTH    Month of Year                RET  N    000  2  _____|
|                                     |
| _  YEAR     Year of Century              RET  N    000  2  _____|
|                                     |
| _  SYSID    System Identifier            RET  A    000  4  _____|
|                                     |
| _  CPUPCBSY Pct Processor Bus           COMP N    050  5  MAPCT.   |
|                                     |
| _  DATE     Interval End Date            RET  D    000  7  DATE9.   |
|                                     |
| ***** BOTTOM OF DATA *****|
|
\-----/

```

Figure 3-21. Cross-File Data Element Definition Screen

The parameters for the Cross-File Data Element Definition screen are defined below:

Specify SAS exit stmts:	The specification for assigning SAS statements to derive cross-file elements. Specify a Y (YES) if you wish to define cross-file data elements. (See Figure 3-22.)
Specify Expression stmts:	The specification for defining Expression statements. Specify a Y (YES) if you wish to define expression statements.
Cmd:	The line commands are: - I (insert) to insert a new data element. - R (repeat) to repeat a data element. - S (depend clauses and label) to specify independent data elements upon which this data element depends. See Figure 3-14. - C (copy) to copy a data element. - E (expression) to create SAS code to calculate the value of this specific data element. See Figure 3-15. - D (delete) to delete a data element.
Name:	The name of the data element. This may be either the name of a CA MICS data element that exists in the input CA MICS file or any valid SAS name, as described in the SAS manuals. If you wish to define non-CA MICS data elements, choose variable names that are descriptive of the values they represent.

Note that when you first see this screen, it displays only the control data elements that are always present in your Capacity Planning database file. If you attempt to modify these data elements, an error will occur. Define an additional row for each data element that you wish to include in your Capacity Planning database file by typing an I (INSERT) in the Cmd column of any existing row.

If you are uncertain of the correct spelling of a CA MICS data element or if you would like to see a complete list of the CA MICS data elements available in the CA MICS input file, type a question mark (?) in the name field to display the screen shown in Figure 3-20.

Data Element Label (long name): The data element label that is displayed on the CAP File and Data Element Selection screen and on many reports. This label is supplied automatically for CA MICS data elements.

D: The data element drop flag. Typing a D in this field makes the corresponding data element available for use in calculations through the cross-file data element definition step of the Capacity Planning database file update, but drops the data element from the resulting Capacity Planning database file. Leave the field blank to keep the corresponding data element on the Capacity Planning database file.

Sum Type: The summarization type of the data element. Summarization types are RET (retain), ACCM (accumulate), MIN (minimum), MAX (maximum), and COMP (compute).

Data type: The type of data element. This is used to validate data values and usage for the data element. Valid types are Alphanumeric (A), Numeric (N), Money (M), Time (T), Hex (H), Date (D), Time-of-day (TD), and Date-Time (DT). This value is supplied automatically for CA MICS data elements.

Order: Define the execution order of the data elements you defined by entering ordinal numeric values greater than 300.

Note: This is required if you are going to use this Resource Element file as input to a Meta file.

The order value is automatically defined for CA MICS data elements and Capacity Planner control elements. Data elements that require other data elements are automatically included in the Capacity Planner database file. In particular, two variables could appear that will not show up in the output file: @@FIRST and @@LAST.

Len: The length of the data element. This value is supplied automatically for CA MICS data elements.

Output Format: The SAS format that is used to print data element values. This value is supplied automatically for CA MICS data elements.

```

/----- SAS System Statements -----\
|Command ==>                               Scroll ==> CSR |
| |                                         |
|Inquiry Step:  Cross-file data element derivation |
| |                                         |
|SAS Statements for the CAPMERG Macro |
|-----|
|***** ***** TOP OF DATA *****|
|==MSG> SPECIFY SAS STATEMENTS TO SET VARIABLES IN THE CROSS-FILE MERGE. |
|=NOTE= %MACRO CAPMERG; <=== PLEASE DO NOT PLACE LINES BEFORE THIS LINE. |
|000001 |
|000002 |
|000003 |
|000004 |
|000005 |
|000006 |
|000007 |
|000008 |
|000009 |
|000010 |
|=NOTE= %MEND CAPMERG; <=== PLEASE DO NOT PLACE LINES AFTER THIS LINE. |
|***** ***** BOTTOM OF DATA *****|
| |
\-----/

```

Figure 3-22. SAS Statements Screen - Deriving Cross-file Data Elements

Using the SAS Statements screen shown in Figure 3-22, you can code any valid SAS statements that conform to the syntax described in the SAS manuals to derive non-CA MICS data elements.

Note that the %MACRO CAPMERG and %MEND CAPMERG statements are provided for you and that they are field protected from modification.

```
/----- Converting Macros to Expression Statements -----\  
|Command ==>                                         Scroll ==> CSR |  
|                                                     |  
|Enter the variable names that are associated with the SAS statements shown |  
|below.                                              |  
|                                                     |  
|Variable  SAS Statements in Macro                  |  
|-----|-----|  
|RATIO___  IF CPUTOBTM GT 0                          |  
|RATIO___  THEN RATIO = PGACPUTM / CPUTOBTM;          |  
|RATIO___  ELSE RATIO = .;                          |  
|-----|-----|  
|***** BOTTOM OF DATA *****|  
\-----/
```

Figure 3-23. Converting Macros to Expression Statements

You use this screen (see Figure 3-23) to convert existing MACRO statements into Expression Statements that calculate the value of a particular variable. In the example shown in Figure 3-23, the variable RATIO is a user-created data element that corresponds to the ratio of Performance Group Activity (PGA) file CPU time consumed and the total Busy time of the processor. By placing "RATIO" in the Variable Column of this screen the CA MICS Capacity Planner will automatically associate these SAS Statements with RATIO. When you End from this display the CA MICS Capacity Planner will remove the statements from the MACRO and move them into the Expression Statement table for the variable RATIO.

3.3.2.1.3 Specification Processing

After you specify all of the parameters that are necessary to build your Capacity Planning database file, type END at the Command prompt of the Resource Element File menu shown in Figure 3-10, and press ENTER or (usually) press the PF3 key to display the Resource Element File screen shown in Figure 3-24. This screen lists the file definition steps that you defined. You can change specifications for any step by typing an S on the Cmd column of the step you wish to change. In the same way, you can delete or repeat a step by typing D or R, respectively.

```
/-----Validation Required.-\
|Command ===>                                     Scroll ==> CSR |
|Composing CAP DB File def.: SYS - System Activity Tracking File
|File definition must be validated. Enter VALIDATE to validate the file def.
|Update the capacity planning data base file ==> N (Y/N)
|Line Cmds: S Select I Insert D Delete R Repeat
|Cmd          Data Base File Definition Step
|- -----
|_ Resource Element File Definition
|_   File & Data Element Sel.: JOB - BATCH USER JOB ACTIVITY FILE
|_   File & Data Element Sel.: CPU - CPU PROCESSOR ACTIVITY FILE
|_ Cross-File Data Element Definition
|***** BOTTOM OF DATA *****
```

Figure 3-24. Resource Element File Screen

Before you execute your resource element file definition inquiry, you must first validate it to ensure that you have defined all of the necessary steps. For instance, if you specify more than one file and data element selection step, then the online facility ensures that you also specify a cross-file data element definition step to cause the data from each of the file and data element selection steps to be merged.

You may request validation of your resource element file definition inquiry by typing a V (or VALIDATE) at the Command prompt of the Resource Element File screen and pressing ENTER. Diagnostic messages inform you that the inquiry has been validated or why it cannot be validated so you can correct the inquiry step specifications and request validation again.

If you are unable to validate your inquiry and want to defer correcting the problems until a later time, type END at the Command prompt and press ENTER or (usually) press the PF3 key to save your unvalidated inquiry. The Capacity Planning Data Base Files screen, which displays the resource element file definition inquiries in your catalog, is displayed.

After you validate your resource element file definition inquiry, you have an opportunity to execute the inquiry and update the Capacity Planning database file by typing a Y at the Update Capacity Planning database file prompt and pressing ENTER. The Update Capacity Planning File screen shown in Figure 3-26 is displayed. After execution processing is completed, the Resource Element File screen is redisplayed. After you are satisfied with all the specifications, type END at the Command prompt to save your resource element file definition. The Capacity Planning Database File screen is then redisplayed. REST

3.3.2.2 Modifying a Database Specification

After you define a resource element file definition, you can then modify it. The Resource Element File screen shown in Figure 3-25 shows all of the specifications that you can use to manipulate resource element file definitions. You can select a definition by typing an S (SELECT) in the Cmd column of the desired definition. You can also add a new definition by typing an I (INSERT) in the Cmd column of any existing row or by specifying an R (REPEAT) in the Cmd column of a specific row, causing that definition to be duplicated. You may then select the new definition and modify it to conform to your requirements.

```

/----- Resource Element File -----
|Command ===>                                Scroll ===> CSR
|
|Composing CAP DB File def.: SYS - System Activity Tracking File
|
|Update the capacity planning data base file ===> N (Y/N)
|
|Line Cmds: S Select  I Insert  D Delete  R Repeat
|
|Cmd                Data Base File Definition Step
|  - -----
|  - Resource Element File Definition
|  - File & Data Element Sel.: JOB - BATCH USER JOB ACTIVITY FILE
|  - File & Data Element Sel.: CPU - CPU PROCESSOR ACTIVITY FILE
|  - Cross-File Data Element Definition
|***** BOTTOM OF DATA *****
|
|
|-----

```

Figure 3-25. Resource Element File Screen

3.3.2.3 Executing a Database Specification

When you select a resource element file definition, including standard applications, for execution, the Update Capacity Planning screen shown in Figure 3-26 is displayed.

```

/----- Update Capacity Planning File -----\
|Command ==>                                     |
|                                                    |
|Composing CAP DB File def.: R03 - sample resource element file |
|                                                    |
|Type of update                ==> INITIALIZE (INITIALIZE/UPDATE) |
|Replace existing file?        ==> N          (Y/N)              |
|                                                    |
|Input data source              ==> ARCHIVE    (ARCHIVE/ONLINE)   |
|CA MICS file cycle(s)          ==> __ - __   (0 - 99)           |
|Archive generation number      ==> 0         (0 to -99)          |
|                                                    |
|Starting date of update        ==> _____ (ddmonyy)         |
|Number of months to select     ==> _____ (1 to 9999)       |
|                                                    |
|Mode of execution              ==> BATCH      (BATCH/BACKGROUND) |
|                                                    |
|List updated file contents?    ==> YES        (YES/NO)           |
|                                                    |
|Press END to invoke MICF inquiry execution to update the resource element file. |
|Enter CANCEL to terminate the resource element file update.      |
|                                                    |
\-----/

```

Figure 3-26. Update Capacity Planning File Screen

The parameters for the Update Capacity Planning File screen are described below:

Type of update:	The specification for initializing or updating the Capacity Planning database file. If the resource element file does not exist or if you plan to overwrite the current file, specify INITIALIZE. If you wish to add to an existing file, specify UPDATE.

Note that if you specify INITIALIZE and the file already exists, all historical and forecasted observations that you previously placed in the file will be lost.

Replace existing file? The specification for deleting an old file and replacing it with a new file. If you specify INITIALIZE and the file already exists, you must set this parameter to Y (YES) to delete the old version and confirm your intent to initialize and replace. Otherwise, specify N (NO).

Input data source: The specification for extracting online or archived data. Specify ONLINE if the data that you wish to extract from the CA MICS files is in the online portion of the CA MICS database. Specify ARCHIVE if you wish the data to be read from a CA MICS history archive tape.

Note: If any of the input files have an input timespan of DETAIL or DAYS, you may only use ONLINE as your input data source.

CA MICS file cycle(s): The CA MICS file cycles to be used as input to loading the resource element file if you specify ONLINE for Data source.

Archive generation number: The archive generations that you wish to use as input if you specify ARCHIVE for data source.

Update starting date: The start date for date range selection. If Start date is specified, only historical observations at least as recent as this value are included in the analysis. The discussion of the following control parameter describes the options for controlling date range selection.

Number of days/weeks/months to select: The maximum number of days, weeks, or months of historical observations to be included in the analysis. Note that once you specify the input file, the field name changes to reflect the timespan of the input file. When used in conjunction with the Start date parameter, several options are available for date range selection control:

- o If you specify only Update start date, all historical observations at least as recent as the value specified for Update start date is selected.

- o If you specify both Update start date and Number of days/weeks/months, the date range for data selection begins with the specified value for Start date and extends for the specified number of days, weeks, or months.
- o If you specify only Number of days/weeks/months, the date range for data selection includes that number of days, weeks, or months prior to the submit date.
- o If you specify neither parameter, then all historical observations are selected when Update type is INITIALIZE. When Update type is UPDATE, the default of 1 for Number of days/weeks/months updates the Capacity Planning database file with data for the last day/week/month, based on the CA MICS calendar.

Mode of execution: The specification for executing the resource element file update. Specify BATCH to submit the resource element file update as a batch job. Specify FOREGROUND to execute it immediately as an online SAS program.

List updated file contents? The specification to execute a PROC CONTENTS and a PROC PRINT of the resource element file to be executed after the database is updated. To do so, specify a Y (YES).

The CA MICS Database Inquiry Execution screen appears showing the execution parameters that are available before the dataBASE update job is submitted for processing. This screen and its control parameters and options are discussed in detail in Section 1.5.3.5, Executing a File Definition Inquiry.

3.3.3 Resource Element File Case Studies

The Resource Element File case studies, which are also continued as examples in Chapters 4 through 10, illustrate the use of the Capacity Planning resource element files.

These five case studies are presented in the sections that follow:

- 1 - Extended Example I/O Workload Analysis
- 2 - Performance Group/Period Workload Tracking
- 3 - JES Spool Activity
- 4 - Tracking Applications
- 5 - System Workload and Performance Profile

3.3.3.1 Extended Example I/O Workload Analysis

This extended example of I/O workload analysis illustrates how to apply the capacity planning resource element files in analyzing the I/O workload of an IBM 4381-2. You can also apply this study to track channel loadings on any MVS/370 system. The I/O subsystem of an XA processor does not uniquely identify a single channel path for a single I/O, so this type of study does not apply to an MVS/XA system.

The following topics are addressed:

- 1 - Problem Description
- 2 - Control Parameters
- 3 - Output Report

3.3.3.1.1 Problem Description

You can use the capacity planning resource element file to collect monthly historical data to analyze the growth in the utilization of the base group of selector and block multiplexor channels on an IBM 9021-711. The scope of this study is to determine if load growth requires the acquisition of the optional channel group.

The base channel group of a 9021-711 consists of one-byte multiplexor channel (channel 0) and five block multiplexor channels (channels 1 through 5). Therefore, data is collected for channels 1 through 5 of the system. Two values characterize the workload of each of these channels:

- o The number of successful start I/Os (SIOs) issued by the channel. This value is provided by the CA MICS data element PCASIO in the HARPCA File.
- o The utilization (that is, percent busy) of the channel. This value is provided by the CA MICS data element PCAPCSY in the HARPCA File.

The following is a discussion of the design of the control stream that generates these values.

CONTROL STREAM DESIGN

When you select a CA MICS file, such as the HARPCA file, as input to the resource element files, always consult the file description in the CA MICS Database Structure and Content Guide to ensure that the required data elements are not lost when the file is sorted and summarized by the default sequence that is used by the capacity planning database generator. This sequence is

SYSID ZONE YEAR timespan

For the HARPCA File, the summary sequence for the MONTHS timespan is:

SYSID CHANNEL CPUID YEAR MONTH ZONE

If the HARPCA file is summarized by the default sequence, the data for the individual channels is summarized into one meaningless observation unless you add CHANNEL to the summary sequence of the file. Since the system is a uniprocessor, the CPUID is unimportant.

The program initially selects the data elements that you specify from the CA MICS file. In the data selection step, the program invokes the user-defined selection exits. These exit macros perform two functions:

- o Deleting unnecessary observations to reduce the processing requirements of the program
- o Calculating user-defined data elements

For this extended example, the selection exit is used to delete the observations for the channels that are not involved in the study. While in this example deleting the unnecessary channel observations represents only a small potential savings, the exit macros are introduced to remind you of the potential resource savings when files like the BATJOB or BATPGM files are processed.

After the HARPCA data is summarized, one record for each unique SYSID ZONE YEAR MONTH CHANNEL sequence exists in the file. This data set is processed by a data step that invokes the user-supplied summary logic that is coded in the summary exit macro. The user-supplied summary logic is used to assign the PCASIO and PCAPBSY values to the user-defined data elements for the individual channels.

For this example, the user-defined data elements are the following:

- CHAN1BS & CHAN1CNT - the utilization and SIO values for channel 1
- CHAN2BS & CHAN2CNT - the utilization and SIO values for channel 2
- CHAN3BS & CHAN3CNT - the utilization and SIO values for channel 3
- CHAN4BS & CHAN4CNT - the utilization and SIO values for channel 4
- CHAN5BS & CHAN5CNT - the utilization and SIO values for channel 5

These values are assigned by testing the value of the CHANNEL data element and then assigning the PCASIO and PCAPCSY values to the user-defined data elements.

After the data step in which the user summary exit is invoked, the observations are sorted into the default summary sequence for the final merge step. This summary sequence is

SYSID ZONE YEAR timespan

The records are then merged with the default sequence key. After the records are merged, the program drops the summary and drop variables, resulting in a single record produced for each month that contains the five channel utilizations (CHAN1BS through CHAN5BS) and the five SIO counts (CHAN1CNT through CHAN5CNT).

3.3.3.1.2 Control Parameters

Figures 3-27 through 3-38 show the control screens that were used to create and execute this resource element file definition example.

```

/----- Resource Element File Definition -----\
|Command ==>
|Enter a ? in any data entry field for more information on valid values.
|      CAP DB File def.:  SYS - System Activity Tracking File
|
|Summarization Level
|      Primary   ==> MONTHS  (DAYS/WEEKS/MONTHS)
|      Secondary ==> ZONE    (ZONE/HOUR/ENDTS/DAYNAME)
|
|Amount of data retained ==> 60      (1-9999 months)
|Age of oldest forecast  ==> 24      (1-9999 months)
|History cycles          ==> 3       (1-99 cycles)
|Forecast cycles         ==> 3       (1-99 cycles)
|
|Enable:
|  Notes report          ==> N       (YES/NO)
|  Contents reports      ==> N       (YES/NO)
|
|Specify extended options ==> Y      (Y/N)
\-----/

```

Figure 3-27. Specifying Parameters for the Resource Element File

Option 1, Insert Common Data Selection, is chosen on the Resource Element File screen shown in Figure 3-27.

```

/----- Common Data Selection -----\
|Command ==>                               |
|Enter a ? in any data entry field for more information on valid values.  |
|Composing CAP DB File def.:  CHN - I/O Channel Example                    |
|                                                                            |
|SYSID  ==> TS01 TS02 _____          |
|              _____          |
|Zone    ==> 1 2 _ _ _ _ _              (1 - 9) |
|Hour    ==> _ - _ _ - _ _ - _ _ - _ _ (0 - 23) |
|Dayname ==> _____          |
|                                                                            |
|Specify common data selection SAS statements ==> N (Y/N/R)                |
\-----/

```

Figure 3-29. Common Data Selection Screen

The Common Data Selection screen shown in Figure 3-29 specifies that observations with SYSIDs that are equal to TS01 or TS02, and with ZONEs 1 or 2 are to be included in the resource element data file. No user-defined common data selection statements are to be specified.

[illegible]

Figure 3-30. Selecting CA MICS Files and Data Elements

Option 2, Insert CA MICS File and Data Selection, is chosen on the Resource Element File screen shown in Figure 3-30.

The CA MICS File and Data Element Selection screen shown in Figures 3-31 and 3-32 specifies that the CA MICS PCA file in database unit P is to be used. The data elements to be included in the resource element file are listed. A Y specified for the Specify CA MICS file processing exits parameter indicates that user-defined logic for selecting specific channels is to be used.

```

/----- CA MICS File and Data Element Selection -----\
|Command ==>                                                    Scroll ==> CSR|
|Enter a ? in any data entry field for more information on valid values.|
|Composing CAP DB File def.:  CHN - I/O Channel Example|
|CA MICS file      ==> PCA      (fff) -  PHYSICAL CHANNEL ACTIVITY FILE|
|Input Time-Span   ==> MONTHS (DETAIL/DAYS/WEEKS/MONTHS)|
|Data base (s)     ==> P|
|Include CAPAPUs   ==> N (Y/N)|
|Summarize file by ==> CHANNEL_|
|Specify file exits ==> N (Y/N)|
|
|-----Line Commands-----|
| I Insert C Copy D Delete |
| R Repeat E Calculations  |
| S Depend Clauses & Label |
|-----|
|
|                               Sum Data
|Cmd  Name   Label (long name)  D Type Type Order Len Output-Format|
|-----|
|_ CHANNEL Physical Channel
|
|                               D RET  A    050   2  _____|
|_ CHAN1BS Channel 1 Busy
|
|                               _ COMP N    050   7  PERCENT_____|
|_ CHAN1CNT Channel 1 Start
|
|                               _ ACCM N    050   7  _____|
|_ CHAN2BS Channel 2 Busy
|
|                               _ COMP N    050   7  PERCENT_____|
|_ CHAN2CNT Channel 2 Start
|
|                               _ ACCM N    050   7  _____|
|_ CHAN3BS Channel 3 Busy
|
|                               _ COMP N    050   7  PERCENT_____|
|_ CHAN3CNT Channel 3 Start
|
|                               _ ACCM N    050   7  _____|
|***** BOTTOM OF DATA *****|
|-----|

```

Figure 3-31. CA MICS File and Data Element Selection Screen - Screen 1

Figure 3-31. CA MICS File and Data Element Selection Screen - Screen 2

```

/----- CA MICS File and Data Element Selection -----\
|Command ==>                                         Scroll ==> CSR |
|Enter a ? in any data entry field for more information on valid values. |
|Modifying CAP DB File def.: CHN - I/O Channel Example |
|CA MICS file      ==> PCA    (fff) - PHYSICAL CHANNEL ACTIVITY FILE   |
|Input Time-Span   ==> MONTHS (DETAIL/DAYS/WEEKS/MONTHS)              |
|Data base (s)     ==> P                                              |
|Include CAPAPUS   ==> N (Y/N)          -----Line Commands----- |
|Summarize file by ==> CHANNEL_         | I Insert C Copy D Delete |
|Specify file exits==> Y (Y/N)         | R Repeat E Calculations  |
|                                     | S Depend Clauses & Label  |
|                                     ----- |
|                                     |
|                                     |
|                                     | Sum Data
|Cmd Name Label (long name) D Type Type Order Len Output-Format |
|-----|-----|-----|-----|-----|-----|-----|
| _ PCASIO Successful I/Os D ACCM N 050 13 _____ |
| _ SYSID System Identifier _ RET A 000 4 _____ |
| _ YEAR Year of Cent _ RET N 000 2 _____ |
| _ ZONE Time Zone _ RET A 000 1 _____ |
|***** BOTTOM OF DATA *****|

```

Figure 3-32. CA MICS File and Data Element Selection Screen - Screen 3

The CA MICS File Processing Exits Menu shown in Figure 3-33 indicates that both data selection and data element derivation logic are to be defined by the user.

```
/----- SAS System Statements -----\  
|Command ==>                               Scroll ==> CSR |  
|  
|Inquiry Step:  File & Data Element Sel.: PCA - PHYSICAL CHANNEL ACTIVITY FILE|  
|  
|SAS Statements for the CAPSLCT Macro      |  
|-----|  
|***** ***** TOP OF DATA *****|  
|==MSG> SPECIFY SAS STATEMENTS TO SUBSET THIS PARTICULAR FILE.      |  
|=NOTE= %MACRO CAPSLCT; <== PLEASE DO NOT PLACE LINES BEFORE THIS LINE. |  
|000001 IF CHANNEL = '1' OR CHANNEL = '2' OR CHANNEL = '3' OR      |  
|000002   CHANNEL = '4' OR CHANNEL = '5';                          |  
|=NOTE= %MEND CAPSLCT; <== PLEASE DO NOT PLACE LINES AFTER THIS LINE. |  
|***** ***** BOTTOM OF DATA *****|  
|  
\-----/
```

Figure 3-34. Selecting File-Unique Data

The SAS Statements screen in Figure 3-35 shows the SAS statements that are to be used to define the new data elements. CHANxBS is the Channel Busy percentage for channel x while CHANxCNT is the SIO count for the same channel.

```

/----- Resource Element File ----- Inquiry Validated.-/
|Command ===>                               Scroll ==> CSR |
|
|Composing CAP DB File def.: CHN - I/O Channel Example
|
|Update the capacity planning data base file ===> Y (Y/N)
|
|Line Cmds: S Select  I Insert  D Delete  R Repeat
|
|Cmd                Data Base File Definition Step
|  - -----
|  _ Resource Element File Definition
|  _ Common Data Selection
|  _ File & Data Element Sel.: PCA - PHYSICAL CHANNEL ACTIVITY FILE
|***** BOTTOM OF DATA *****
|
|
|
|
|
|
|
|
|
|
|-----/

```

Figure 3-36. Validating the Inquiry

After the resource element file definition parameters are specified, the Resource Element File screen is used to validate the inquiry. Figure 3-36 shows this screen after VALIDATE is typed at the Command prompt.

```
/----- Update Capacity Planning File ----- Confirm Save.-\
|Command ==>|
|Modifying CAP DB File def.: CHN - I/O Channel Example|
|This file definition has been changed and must be saved before updating the|
|file.|
|Do you wish to save the file definition and update the file? ==> YES (YES/NO)|
|Instructions:|
|  Press ENTER key to save the file definition and update the file.|
|  Reply NO and press enter key to cancel the file update request.|
|-----/
```

Figure 3-37. Saving the Inquiry

The Update Capacity Planning File screen shown in Figure 3-37 is used to save the inquiry so that it can be executed.

```

/----- Update Capacity Planning File -----\
|Command ==>                                     |
|                                                    |
|Composing CAP DB File def.: CHN - I/O Channel Example |
|                                                    |
|Type of update                ==> INITIALIZE (INITIALIZE/UPDATE) |
|Replace existing file?        ==> N          (Y/N)              |
|                                                    |
|Input data source              ==> ONLINE    (ARCHIVE/ONLINE)   |
|CA MICS file cycle(s)         ==> 01 - 12    (0 - 99)            |
|Archive generation number     ==> ___        (0 to -99)          |
|                                                    |
|Starting date of update       ==> _____ (ddmonyy)         |
|Number of months to select    ==> 12__      (1 to 9999)         |
|                                                    |
|Mode of execution              ==> BATCH     (BATCH/BACKGROUND) |
|                                                    |
|List updated file contents?   ==> YES       (YES/NO)            |
|                                                    |
|Press END to invoke MICF inquiry execution to update the resource element file. |
|Enter CANCEL to terminate the resource element file update. |
|                                                    |
\-----/

```

Figure 3-38. Executing the Inquiry

Cycles 1 through 12 are specified as input to the database loading process on the Update Capacity Planning File screen shown in Figure 3-38. The updated resource element file will be printed after the update is complete.

3.3.3.1.3 Output Report

Figure 3-39 shows the output that is created by the PROC PRINT of the updated data set.

CA MICS Capacity Planner														
OBS	DATE	MONTH	SYSID	YEAR	ZONE	CHAN1BS	CHAN1CNT	CHAN2BS	CHAN2CNT	CHAN3BS	CHAN3CNT	CHAN4BS	CHAN4CNT	CHAN5CNT
1	30SEP97	9	TS01	97	1	5.3569	10283160	18.9219	12621388	18.8242	16971048	7.2551	3640411	15.7917
2	30OCT97	10	TS01	97	1	5.5168	11306966	19.4883	13861055	19.6445	18385386	7.3911	3822387	14.7969
3	30NOV97	11	TS01	97	1	6.3586	11414859	19.7500	11717108	20.2109	16663774	9.9683	4338143	15.9124
4	31DEC97	12	TS01	97	1	5.3074	9976870	18.0391	12327566	18.5586	16800745	8.8547	5247391	14.6099
5	29JAN98	1	TS01	98	1	6.0647	10611907	18.9922	11156555	19.3750	16185623	8.9697	4257223	15.6548
6	26FEB98	2	TS01	98	1	6.3376	10960937	23.0977	12830799	18.8438	15257229	11.1433	5641456	16.2344
7	31MAR98	3	TS01	98	1	5.4783	10025924	19.6797	12843967	16.9063	13669835	9.5486	5318293	14.8582
8	30APR98	4	TS01	98	1	4.7642	9870709	17.9570	13093325	15.2307	14904987	7.6726	5033416	14.3140
9	28MAY98	5	TS01	98	1	4.5576	8851487	17.7852	11399322	15.6868	13645679	8.2495	4494512	14.5493
10	30JUN98	6	TS01	98	1	4.9968	10545954	17.5703	12099310	15.7957	14443173	9.2905	5437664	15.1023
11	30JUL98	7	TS01	98	1	6.9377	14139601	21.0313	13436249	18.6172	15858031	11.7161	6582489	17.7891
12	31AUG98	8	TS01	98	1	8.5986	14220067	30.9336	17448773	15.0393	14690555	10.7354	6879567	15.3848
13	30SEP97	9	TS01	97	2	3.6626	3844805	21.0391	9137237	25.5938	12868504	9.1514	4359236	16.0000
14	30OCT97	10	TS01	97	2	3.6033	4058408	21.1641	9761257	23.7773	12025920	8.2043	3752326	14.9243
15	30NOV97	11	TS01	97	2	3.5081	3559998	21.0078	8160196	26.0352	11506972	9.4600	3786755	15.9133
16	31DEC97	12	TS01	97	2	2.6174	3018221	22.1406	10373391	30.1250	12866071	11.2361	4973300	16.9727
17	29JAN98	1	TS01	98	2	2.0652	2197819	18.7461	8134965	24.0313	10484711	8.0525	3572788	15.1262
18	26FEB98	2	TS01	98	2	2.4753	2669361	21.7070	8721177	24.9844	11027921	10.0513	4014299	16.1367
19	31MAR98	3	TS01	98	2	1.9980	2196140	18.5313	7999537	22.3164	11310782	9.1003	3852128	13.9209
20	30APR98	4	TS01	98	2	1.5474	1753226	17.7422	7390090	20.3477	9696053	7.9121	3435097	14.7136
21	28MAY98	5	TS01	98	2	1.5369	1670242	16.2227	6574310	18.8555	9463561	8.3237	3304728	13.3713
22	30JUN98	6	TS01	98	2	2.0928	2596361	17.6797	7756252	21.4180	11278067	9.4841	3893271	15.0049
23	30JUL98	7	TS01	98	2	3.1802	3551786	20.6289	7575368	26.1992	13215648	11.2686	3863744	14.9900
24	31AUG98	8	TS01	98	2	3.3550	3475517	30.0859	10478833	23.7578	10705401	11.3088	4154695	13.9648
25	30SEP97	9	TS02	97	1	5.8586	4117146	19.7070	19625621	20.8047	22304391	12.7476	9302529	10.8057
26	30OCT97	10	TS02	97	1	14.7346	2722390	19.8320	20338573	19.9375	22521512	15.4590	11686631	10.7395
27	30NOV97	11	TS02	97	1	13.7773	3809002	14.9670	15289886	18.5938	19254275	12.3804	8388867	8.9019
28	31DEC97	12	TS02	97	1	10.0374	4721964	13.4104	13712714	16.6367	17817953	15.4211	10188718	9.0366
29	29JAN98	1	TS02	98	1	12.7666	5548874	16.8281	13845993	17.6797	18586811	12.6563	8258431	9.9780
30	26FEB98	2	TS02	98	1	10.8706	7771787	16.3789	13405987	15.9539	17237412	13.2949	10653306	10.2646
31	31MAR98	3	TS02	98	1	13.4224	9735884	15.5269	14151218	15.9382	17819024	14.6516	10932092	9.5662
32	30APR98	4	TS02	98	1	11.6868	5365007	13.5408	13405430	13.8350	17208020	13.9941	10965281	8.9309
33	28MAY98	5	TS02	98	1	8.6855	6873986	14.3179	13105377	13.7134	15188405	13.4080	9084066	10.0325
34	30JUN98	6	TS02	98	1	10.0339	7030206	15.9690	14749902	14.5276	16348721	16.5117	10999226	10.6196
35	30JUL98	7	TS02	98	1	8.2434	7093482	14.4727	13043652	13.4624	14116174	15.6594	10289280	8.7729
36	31AUG98	8	TS02	98	1	13.1682	9970887	24.8711	17365170	13.7458	13949379	12.1421	9414241	8.8535
37	30SEP97	9	TS02	97	2	2.7336	880241	18.6016	9049081	25.0039	10550887	10.5205	3584787	11.0811
38	30OCT97	10	TS02	97	2	7.3264	679033	21.8281	9497450	24.7461	10452369	14.0229	4612623	10.7666
39	30NOV97	11	TS02	97	2	6.6284	784134	17.5781	7030350	21.5352	8291186	8.3538	2943352	11.3120
40	31DEC97	12	TS02	97	2	4.1448	858595	13.8445	5741095	16.3867	6745006	9.0964	3310562	8.4487
41	29JAN98	1	TS02	98	2	4.4250	956579	18.9531	7723973	21.4648	8545945	10.4185	3550842	11.7129
42	26FEB98	2	TS02	98	2	3.6133	1217462	20.1680	7515701	20.2422	7863046	13.5088	4195880	13.0649
43	31MAR98	3	TS02	98	2	6.2932	2400223	23.8320	9675267	25.9297	11585347	15.1511	5334095	14.5879
44	30APR98	4	TS02	98	2	4.4189	1045310	22.6758	8853916	22.8203	10761397	15.0293	4769231	12.6448
45	28MAY98	5	TS02	98	2	3.0962	1307297	20.6953	7138379	22.4414	9365947	14.7139	3981306	13.8040
46	30JUN98	6	TS02	98	2	3.2813	1239208	22.0625	8164637	20.7148	9408193	14.6201	4561230	12.8960
47	30JUL98	7	TS02	98	2	2.0996	896128	18.2734	5934381	16.2227	6571253	14.1196	3805317	8.7349
48	31AUG98	8	TS02	98	2	3.4624	1362194	26.0430	7640659	17.9375	5866009	14.2581	3940249	9.8955

Figure 3-39. Resource Element File Contents Report

3.3.3.2 Performance Group/Period Workload Tracking

In this performance group/period workload tracking example, an analytic model was developed using the CA MICS MVS Model Generator to represent the peak morning and afternoon hours workload (that is, 8 AM to 9AM and 2PM to 3PM) of a 3083 CPU, SYSID TS01.

The following topics are discussed by section:

- 1 - Problem Description
- 2 - Control Parameters
- 3 - Output Report

3.3.3.2.1 Problem Description

The following chart represents the workloads that were defined to the model:

WORKLOAD	PERFORMANCE GROUPS AND/OR PERIODS
1 - Period 1 TS0	Performance Group 2 Period 1
2 - Period 2 TS0	Performance Group 2 Period 2
3 - Period 3 TS0	Performance Group 2 Period 3
4 - Period 4 TS0	Performance Group 2 Period 4
5 - Period 5 TS0	Performance Group 2 Period 5
6 - Development TS0	Performance Group 7
7 - JES	Performance Group 62
8 - Overhead	All Other Performance Groups

Although the analytic model was validated for the period selected, there are still two other questions that must be answered:

- o Is the selected day representative of the same time periods on other days?
- o How can the workload transaction arrival rates be expected to grow in the future?

A resource element file was used to collect 12 weeks of data from ZONE 1 (Monday through Friday prime shift) for the workload. You can compare these values to the transaction counts reported for the modeled interval to see if the modeled interval is similar to the historical data. (Later in this document, we show how you can use the workload forecasting program to estimate the future values of the workload transaction arrival rates.)

3.3.3.2.2 Control Parameters

Figures 3-40 through 3-55 show the control screens that we used to create and execute this resource element file example.

```

/----- Resource Element File Definition -----\
|Command ==>|
|Enter a ? in any data entry field for more information on valid values.|
|CAP DB File def.: PRF - Performance Group Example|
|
|Summarization Level|
|Primary ==> WEEKS (DAYS/WEEKS/MONTHS)|
|Secondary ==> ZONE (ZONE/HOUR/ENDTS/DAYNAME)|
|
|Amount of data retained ==> 60 (1-9999 weeks)|
|Age of oldest forecast ==> 24 (1-9999 weeks)|
|History cycles ==> 3 (1-99 cycles)|
|Forecast cycles ==> 3 (1-99 cycles)|
|
|Enable:|
|Notes report ==> N (YES/NO)|
|Contents reports ==> N (YES/NO)|
|
|Specify extended options ==> N (Y/N)|
\-----/

```

Figure 3-40. Specifying Parameters for the Resource Element File

Option 1, Insert Common Data Selection, is chosen on the Resource Element File screen shown in Figure 3-41.

```

/----- Common Data Selection -----\
|Command ==>                               |
|Enter a ? in any data entry field for more information on valid values. |
|Composing CAP DB File def.: PRF - Performance Group Example |
|                               |
|SYSID  ==> TS01  _____ |
|                               |
|Zone   ==> 1  _ _ _ _ _ (1 - 9) |
|Hour   ==> 08 - 09  14 - 15  __ - __  __ - __  __ - __ (0 - 23) |
|Dayname ==> _____ |
|                               |
|                               |
|Specify common data selection SAS statements ==> N (Y/N/R) |
\-----/

```

Figure 3-42. Common Data Selection Screen

The Common Data Selection screen shown in Figure 3-42 specifies that observations with SYSIDs that are equal to TS01, ZONE 1, and HOURS 8 and 14 are to be included in the resource element file.

```
/----- Resource Element File -----  
|Option ==> 2  
|  
|Composing CAP DB file def.: PRF - Performance Group Example  
  
|    1 - Insert Common Data Selection  
|        Specify data selection criteria to apply to ALL files.  
|    2 - Insert CA MICS File and Data Element Selection  
|        Select a CA MICS file and data elements.  
|    3 - Insert Cross-File Data Element Definition  
|        Specify cross-file data element derivations.
```

Figure 3-43. Selecting CA MICS Files and Data Elements

Option 2, Insert CA MICS File and Data Selection, is chosen on the Resource Element File screen shown in Figure 3-43.

The CA MICS File and Data Element Selection screen shown in Figure 3-44 specifies that the CA MICS PPA file in database unit P is to be used. The data elements to be included in the resource element file or to be used to derive non-CA MICS data elements are listed.

```

/----- CA MICS File and Data Element Selection -----\
|Command ==>                                         Scroll ==> CSR |
|Enter a ? in any data entry field for more information on valid values. |
|Composing CAP DB File def.: PRF - Performance Group Example |
|CA MICS file      ==> PPA      (fff) - MVS PERFORMANCE PERIOD ACTIVITY FILE |
|Input Time-Span   ==> WEEKS   (DETAIL/DAYS/WEEKS/MONTHS) |
|Data base (s)     ==> P |
|Include CAPAPUs   ==> N (Y/N) |
|Summarize file by ==> PERFPER_ |
|Specify file exits ==> Y (Y/N) |
| |
| |
|-----Line Commands-----|
| I Insert C Copy D Delete |
| R Repeat E Calculations |
| S Depend Clauses & Label |
|-----|
| |
| Sum Data |
| Cmd Name Label (long name) D Type Type Order Len Output-Format |
|-----|
| _ PERFPER Perf Group Period D RET N 050 5 _____ |
| _ DATE Interval End Date _ RET D 000 7 DATE7. _____ |
| _ PPATTX Trans Terminated _ ACCM N 050 13 _____ |
| _ SYSID System Identifier _ RET A 000 4 _____ |
| _ WEEK Week of Year _ RET N 000 2 _____ |
| _ WKL01 Period 1 TS0 _ ACCM N 120 13 _____ |
| _ WKL02 Period 2 TS0 _ ACCM N 120 13 _____ |
|***** BOTTOM OF DATA *****|
| |
\-----/

```

Figure 3-44. CA MICS File and Data Element Selection Screen - Screen 1

The CA MICS File Processing Exits Menu shown in Figure 3-45 indicates that both data selection and data element derivation logic are to be defined by the user.

```

/----- SAS System Statements ----- Scroll ==> CSR
|
|
|Inquiry Step:   File & Data Element Sel.: PPA - MVS PERFORMANCE PERIOD
|
|SAS Statements for the CAPSLCT Macro
|-----
|***** ***** TOP OF DATA *****
|=MSG> SPECIFY SAS STATEMENTS TO SUBSET THIS PARTICULAR FILE.
|=NOTE= %MACRO CAPSLCT; <=== PLEASE DO NOT PLACE LINES BEFORE THIS LINE.
|000001 IF PERFGRP = 2;
|=NOTE= %MEND CAPSLCT; <=== PLEASE DO NOT PLACE LINES AFTER THIS LINE.
|***** ***** BOTTOM OF DATA *****
|
|-----
\

```

Figure 3-46. Selecting File-Unique Data

The SAS Statements screen in Figure 3-46 shows the SAS statements that are used to subset the input data to observations for performance group 2 only.

```

/----- SAS System Statements -----\
|Command ==>                               Scroll ==> CSR
|
|Inquiry Step:   File & Data Element Sel.: PPA - MVS PERFORMANCE PERIOD
|
|SAS Statements for the CAPSMRY Macro
|-----
|***** ***** TOP OF DATA *****
|=MSG> SPECIFY SAS STATEMENTS FOR TRANSPOSING THE FILE OR TO SET VARIABLES.
|=NOTE= %MACRO CAPSMRY; <=== PLEASE DO NOT PLACE LINES BEFORE THIS LINE.
|000001 IF PERFPER = 1 THEN WKL01 + PPATTX ;
|000002 IF PERFPER = 2 THEN WKL02 + PPATTX ;
|000003 IF PERFPER = 3 THEN WKL03 + PPATTX ;
|000004 IF PERFPER = 4 THEN WKL04 + PPATTX ;
|000005 IF PERFPER = 5 THEN WKL05 + PPATTX ;
|=NOTE= %MEND CAPSMRY; <=== PLEASE DO NOT PLACE LINES AFTER THIS LINE.
|***** ***** BOTTOM OF DATA *****
|-----
\

```

Figure 3-47. Deriving New Data Elements

The SAS Statements screen in Figure 3-47 shows the SAS statements that are used to define the new data elements. WRK0x is the transaction count for workload x.

The CA MICS File and Data Element Selection screen shown in Figure 3-48 specifies that the CA MICS PGA file in database unit P is to be used. The data elements to be included in the resource element file or to be used to derive non-CA MICS data elements are listed.

```

/----- CA MICS File and Data Element Selection -----\
|Command ==>                                         Scroll ==> CSR |
|Enter a ? in any data entry field for more information on valid values. |
|Composing CAP DB File def.: PRF - Performance Group Example |
|CA MICS file      ==> PGA      (fff) - MVS PERFORMANCE GROUP ACTIVITY FILE |
|Input Time-Span   ==> WEEKS   (DETAIL/DAYS/WEEKS/MONTHS) |
|Data base (s)     ==> P |
|Include CAPAPUs   ==> N (Y/N) |
|Summarize file by ==> PERFGRP_ |
|Specify file exits ==> Y (Y/N) |
| |
| |
|-----Line Commands-----|
| I Insert C Copy D Delete |
| R Repeat E Calculations  |
| S Depend Clauses & Label |
| |
| |
| Cmd  Name  Label (long name)  Sum  Data |
|-----|-----|-----|-----|-----|-----|-----|-----|
| _  PERFGRP  Perf Group Number  D RET  N   050   5  -----|
| _  DATE     Interval End Date  _ RET  D   000   7  DATE7.-----|
| _  PPATTX   Trans Terminated  D ACCM N   050  13  -----|
| _  SYSID    System Identifier  _ RET  A   000   4  -----|
| _  WEEK     Week of Year       _ RET  N   000   2  -----|
| _  WKL06    Development TS0    _ ACCM N   120   7  -----|
| _  WKL07    JES                _ ACCM N   120   7  -----|
| ***** BOTTOM OF DATA *****|
| |
\-----/

```

Figure 3-48. CA MICS File and Data Element Selection Screen - Screen 1

The CA MICS File Processing Exits Menu shown in Figure 3-49 indicates that the data element derivation logic is to be defined by the user.

```
/----- SAS System Statements-----\  
|Command ==>                               Scroll ==> CSR |  
|  
|Inquiry Step:  File & Data Element Sel.: PPA - MVS PERFORMANCE PERIOD |  
|  
|SAS Statements for the CAPSMRY Macro |  
|-----  
|***** ***** TOP OF DATA *****  
|==MSG> SPECIFY SAS STATEMENTS FOR TRANSPOSING THE FILE OR TO SET VARIABLES. |  
|=NOTE= %MACRO CAPSMRY; <=== PLEASE DO NOT PLACE LINES BEFORE THIS LINE. |  
|000001 IF PERFGRP EQ 7 THEN WKL06 + PGATTX ; |  
|000002 ELSE IF PERFGRP EQ 62 THEN WKL07 + PGATTX ; |  
|000003 ELSE IF PERFGRP NE 2 THEN WKL08 + PGATTX ; |  
|=NOTE= %MEND CAPSMRY; <=== PLEASE DO NOT PLACE LINES AFTER THIS LINE. |  
|***** ***** BOTTOM OF DATA *****  
|  
|-----\  
/
```

Figure 3-50. Deriving New Data Elements

The SAS Statements screen in Figure 3-50 shows the SAS statements that are used to define the new data elements. WRK0x is the transaction count for workload x.

```
/----- Resource Element File -----  
|Option ==> 3  
|Composing CAP DB file def.: PRF - Performance Group Example  
  
|    1 - Insert Common Data Selection  
|        Specify data selection criteria to apply to ALL files.  
|    2 - Insert CA MICS File and Data Element Selection  
|        Select a CA MICS file and data elements.  
|    3 - Insert Cross-File Data Element Definition  
|        Specify cross-file data element derivations.
```

Figure 3-51. Specifying a Cross-File Data Element Definition

Option 3, Insert Cross-File Data Element Definition, is chosen on the Resource Element File screen shown in Figure 3-51.

```
/----- SAS System Statements -----\
|Command ==>                               Scroll ==> CSR |
| |
|Inquiry Step:  File & Data Element Sel.: PPA - MVS PERFORMANCE PERIOD |
| |
|SAS Statements for the CAPMERG Macro |
|-----|
|***** ***** TOP OF DATA *****|
|==MSG> SPECIFY SAS STATEMENTS TO SET VARIABLES IN THE CROSS-FILE MERGE. |
|=NOTE= %MACRO CAPMERG; <=== PLEASE DO NOT PLACE LINES BEFORE THIS LINE. |
|000001 WKL01 = WKL01 / 10 ; |
|000002 WKL02 = WKL02 / 10 ; |
|000003 WKL03 = WKL03 / 10 ; |
|000004 WKL04 = WKL04 / 10 ; |
|000005 WKL05 = WKL05 / 10 ; |
|000006 WKL06 = WKL06 / 10 ; |
|000007 WKL07 = WKL07 / 10 ; |
|000008 WKL08 = WKL08 / 10 ; |
|=NOTE= %MEND CAPMERG; <=== PLEASE DO NOT PLACE LINES AFTER THIS LINE. |
|***** ***** BOTTOM OF DATA *****|
| |
\-----/
```

Figure 3-52. Deriving Cross-File Data Elements

After the resource element file definition parameters are specified, the Resource Element File screen is used to validate the inquiry. Figure 3-53 shows this screen after VALIDATE is typed at the Command prompt.

```
/----- Update Capacity Planning File ----- Confirm Save.-\
|Command ==>
|
|Modifying CAP DB File def.: PRF - Performance Group Example
|
|
|This file definition has been changed and must be saved before updating the
|file.
|
|
|Do you wish to save the file definition and update the file? ==> YES (YES/NO)|
|
|
|Instructions:
|
|    Press ENTER key to save the file definition and update the file.
|
|    Reply NO and press enter key to cancel the file update request.
|
|
|-----/
```

Figure 3-54. Saving the Inquiry

The Update Capacity Planning File shown in Figure 3-54 is used to save the inquiry so that it can be executed.

```

/----- Update Capacity Planning File -----\
|Command ==>                                     |
|                                                 |
|Composing CAP DB File def.:  PRF - Performance Group Example |
|                                                 |
|Type of update                ==> INITIALIZE (INITIALIZE/UPDATE) |
|Replace existing file?        ==> N          (Y/N)               |
|                                                 |
|Input data source              ==> ONLINE    (ARCHIVE/ONLINE)    |
|CA MICS file cycle(s)         ==> 01 - 09    (0 - 99)             |
|Archive generation number     ==> ___        (0 to -99)           |
|                                                 |
|Starting date of update       ==> _____ (ddmonyy)          |
|Number of months to select    ==> 9___      (1 to 9999)          |
|                                                 |
|Mode of execution             ==> BATCH      (BATCH/BACKGROUND)  |
|                                                 |
|List updated file contents?   ==> YES        (YES/NO)             |
|                                                 |
|Press END to invoke MICF inquiry execution to update the resource element file. |
|Enter CANCEL to terminate the resource element file update.      |
|-----\

```

Figure 3-55. Executing the Inquiry

Cycles 1 through 9 are specified as input to the database loading process on the Update Capacity Planning File screen shown in Figure 3-55. The updated resource element file will be printed after the update is complete.

3.3.3.2.3 Output Report

Figure 3-56 shows the output that is created by the PROC PRINT of the updated data set.

CA MICS Capacity Planner													
OBS	DATE	SYSID	WEEK	YEAR	ZONE	WKL06	WKL07	WKL08	WKL01	WKL02	WKL03	WKL04	WKL05
1	25JUL98	TS01	30	98	1	2565.9	0.0	1301.3	11102.5	1270.9	45.0	0.4	1.4
2	31JUL98	TS01	31	98	1	2155.1	0.0	1352.6	11026.7	1251.8	49.3	0.6	0.8
3	07AUG98	TS01	32	98	1	2148.7	0.0	1387.4	11232.2	1247.4	47.3	0.6	1.6
4	14AUG98	TS01	33	98	1	2297.2	0.2	753.6	10096.8	1201.9	45.5	0.3	1.1
5	21AUG98	TS01	34	98	1	2254.0	0.0	615.7	10604.8	1361.2	66.9	0.1	0.6
6	28AUG98	TS01	35	98	1	2169.1	0.0	573.3	10192.3	1226.0	41.5	0.4	0.8
7	04SEP98	TS01	36	98	1	2960.0	0.0	760.1	11505.6	1230.3	63.8	0.1	1.8
8	11SEP98	TS01	37	98	1	1743.2	0.0	431.1	8881.0	1263.0	50.0	0.4	3.6
9	18SEP98	TS01	38	98	1	2283.9	0.0	484.5	11927.8	1123.4	38.8	0.2	1.7

Figure 3-56. Resource Element File Contents Report

3.3.3.3 JES Spool Activity

The JES spool activity case study illustrates how to use the CA MICS Capacity Planner to create a resource element file for tracking the print load of remote printers.

The following topics are discussed by section:

- 1 - Problem Description
- 2 - Control Parameters
- 3 - Output Report

3.3.3.3.1 Problem Description

For this case study, a resource element file was created to track the print load of seven remote printers for a shared spool JES2 installation.

The remote printers used in this study are listed below:

Remote Number	Location
=====	=====
0	Computer Center, JES LOCAL
12	Engineering
34	Systems Programming
80	Tech Services
91	Production Control
93	Refinery
96	Application Development

The case study addresses the following areas:

- o Summarizing multiple ZONEs and SYSIDs
- o Using a selection macro to limit the volume of data processed by the program. Rather than summarizing the data for all the remote printers, a selection macro excludes observations for remote printers other than the seven being considered in the study.

3.3.3.3.2 Control Parameters

Figures 3-57 through 3-68 show the control screens that we used to create and execute this resource element file example.

```
/----- Resource Element File Definition -----\  
|Command ==>|  
|Enter a ? in any data entry field for more information on valid values.|  
|CAP DB File def.: JES - JES Spool Example|  
|  
|Summarization Level|  
|Primary ==> MONTHS (DAYS/WEEKS/MONTHS)|  
|Secondary ==> ZONE (ZONE/HOUR/ENDTS/DAYNAME)|  
|  
|Amount of data retained ==> 60 (1-9999 months)|  
|Age of oldest forecast ==> 24 (1-9999 months)|  
|History cycles ==> 3 (1-99 cycles)|  
|Forecast cycles ==> 3 (1-99 cycles)|  
|  
|Enable:|  
|Notes report ==> N (YES/NO)|  
|Contents reports ==> N (YES/NO)|  
|  
|Specify extended options ==> N (Y/N)|  
\-----/
```

Figure 3-57. Specifying Parameters for the Resource Element File

Option 1, Insert Common Data Selection, is chosen on the Resource Element File shown in Figure 3-58.

```

/----- Common Data Selection -----\
|Command ==>|
|Enter a ? in any data entry field for more information on valid values.|
|Composing CAP DB File def.: JES - JES Spool Example|
|
|
|SYSID ==> TS01 TS02 _____|
|          _____|
|
|Zone ==> _____ (1 - 9)|
|
|Hour ==> ____ - ____ - ____ - ____ - ____ - ____ - ____ - ____ (0 - 23)|
|
|Dayname ==> _____|
|
|
|
|
|Specify common data selection SAS statements ==> Y (Y/N/R)|
\-----/

```

Figure 3-59. Common Data Selection Screen

The Common Data Selection screen shown in Figure 3-59 specifies that observations with SYSIDs that are equal to TS01 or TS02 are to be included in the resource element file. SAS statements are specified to further control data selection.

```

/----- SAS System Statements -----\
|Command ==>                               Scroll ==> CSR |
|
|Inquiry Step:  File & Data Element Sel.: JES - JES SP00L |
|
|SAS Statements for the SELECT  Macro |
|-----|
|***** ***** TOP OF DATA *****|
|==MSG> SPECIFY SAS STATEMENTS TO SUBSET THE CAPACITY PLANNING DATA BASE FILE. |
|=NOTE= %MACRO SELECT; <=== PLEASE DO NOT PLACE LINES BEFORE THIS LINE. |
|000001 IF ROUTE = 000 OR ROUTE = 012 OR ROUTE = 034 OR |
|000002     ROUTE = 080 OR ROUTE = 091 OR ROUTE = 093 OR |
|000003     ROUTE = 096 ; |
|=NOTE= %MEND SELECT; <=== PLEASE DO NOT PLACE LINES AFTER THIS LINE. |
|***** ***** BOTTOM OF DATA *****|
|
\-----/

```

Figure 3-60. Selecting Common Data


```

/----- CA MICS File and Data Element Selection -----/
|Command ==>                                     Scroll ==> CSR|
|Enter a ? in any data entry field for more information on valid values.|
|Modifying CAP DB File def.: JES - JES Spool Example|
|CA MICS file      ==> SPL      (fff) - BATCH USER SPOOL ACTIVITY FILE|
|Input Time-Span   ==> MONTHS (DETAIL/DAYS/WEEKS/MONTHS)|
|Data base (s)     ==> P|
|Include CAPAPUs   ==> N (Y/N)|
|Summarize file by ==> ROUTE____|
|Specify file exits ==> Y (Y/N)|
|-----Line Commands-----|
| I Insert C Copy D Delete |
| R Repeat E Calculations  |
| S Depend Clauses & Label |
|-----|
|                               Sum Data|
|Cmd  Name   Label (long name)  D Type Type Order Len Output-Format|
|-----|
|_ RMT080LR Remote 080 Lines|
|_ ACCM N    120  7  _____|
|_ RMT091LR Remote 091 Lines|
|_ ACCM N    120  7  _____|
|_ RMT093LR Remote 093 Lines|
|_ ACCM N    120  7  _____|
|_ RMT096LR Remote 096 Lines|
|_ ACCM N    120  7  _____|
|_ SPLNLR   Logical Writer|
|_ D ACCM N    050 13  _____|
|_ SYSID    System Identifier|
|_ RET  A     000  4  _____|
|-----|

```

Figure 3-62. CA MICS File and Data Element Selection Screen - Screen 2

The CA MICS File Processing Exits Menu shown in Figure 3-64 indicates that the data element derivation logic is to be defined by the user.

```

/----- SAS System Statements -----\
|Command ==>                               Scroll ==> CSR |
|
|Inquiry Step:  File & Data Element Sel.: JES - JES SP00L |
|
|SAS Statements for the CAPSMRY Macro |
|-----|
|***** ***** TOP OF DATA *****|
|==MSG> SPECIFY SAS STATEMENTS FOR TRANSPOSING THE FILE OR TO SET VARIABLES. |
|=NOTE= %MACRO CAPSMRY; <=== PLEASE DO NOT PLACE LINES BEFORE THIS LINE. |
|000001 IF ROUTE = 000 THEN RMTLOCLR + SPNLR ; |
|000002 IF ROUTE = 012 THEN RMT012LR + SPNLR ; |
|000003 IF ROUTE = 034 THEN RMT034LR + SPNLR ; |
|000004 IF ROUTE = 080 THEN RMT080LR + SPNLR ; |
|000005 IF ROUTE = 091 THEN RMT091LR + SPNLR ; |
|000006 IF ROUTE = 093 THEN RMT093LR + SPNLR ; |
|000007 IF ROUTE = 096 THEN RMT096LR + SPNLR ; |
|=NOTE= %MEND CAPSMRY; <=== PLEASE DO NOT PLACE LINES AFTER THIS LINE. |
|***** ***** BOTTOM OF DATA *****|
|
\-----/

```

Figure 3-65. Selecting File-Unique Data

The SAS Statements screen in Figure 3-65 shows the SAS statements that are used to subset the input data to ROUTEs under consideration only, and to derive new data elements to hold the values for the number of print lines for each of the remote printers of interest.

[illegible]

Figure 3-66. Validating the Inquiry

After the resource element file definition parameters are specified, the Resource Element File is used to validate the inquiry. Figure 3-66 shows this screen after VALIDATE is typed at the Command prompt.

```
/----- Update Capacity Planning File ----- Confirm Save.-\
|Command ==>|
|          |
|Modifying CAP DB File def.: JES - JES Spool Example|
|          |
|          |
|This file definition has been changed and must be saved before updating the|
|file.     |
|          |
|Do you wish to save the file definition and update the file? ==> YES (YES/NO)|
|          |
|          |
|Instructions:|
|  Press ENTER key to save the file definition and update the file.         |
|  Reply NO and press enter key to cancel the file update request.         |
|          |
|          |
\-----/
```

Figure 3-67. Saving the Inquiry

The Update Capacity Planning File screen shown in Figure 3-67 is used to save the inquiry so that it can be executed.

```
/----- Update Capacity Planning File -----\  
|Command ==>  
|  
|Composing CAP DB File def.: JES - JES Spool Example  
|  
|Type of update           ==> INITIALIZE (INITIALIZE/UPDATE)  
|Replace existing file?   ==> N          (Y/N)  
|  
|Input data source        ==> ONLINE    (ARCHIVE/ONLINE)  
|CA MICS file cycle(s)    ==> 01 - 05    (0 - 99)  
|Archive generation number ==> ___      (0 to -99)  
|  
|Starting date of update   ==> _____ (ddmonyy)  
|Number of months to select ==> 5___     (1 to 9999)  
|  
|Mode of execution        ==> BATCH     (BATCH/BACKGROUND)  
|  
|List updated file contents? ==> YES      (YES/NO)  
|  
|  
|Press END to invoke MICF inquiry execution to update the resource element file.  
|Enter CANCEL to terminate the resource element file update.  
|  
\-----/
```

Figure 3-68. Executing the Inquiry

Cycles 1 through 5 are specified as input to the database loading process on the Update Capacity Planning File screen shown in Figure 3-68. The updated resource element file will be printed after the update is complete.

3.3.3.3.3 Output Report

Figure 3-69 shows the output that is created by the PROC PRINT of the updated data set.

CA MICS Capacity Planner												
OBS	DATE	MONTH	SYSID	YEAR	ZONE	RMTLOCLR	RMT012LR	RMT034LR	RMT080LR	RMT091LR	RMT093LR	RMT096LR
1	31MAY98	5	SPOL	98	A	170892378	1878785	257370	120771	17080144	9607193	3464558
2	30JUN98	6	SPOL	98	A	181624570	1803782	440657	126215	13941631	9232379	2415886
3	31JUL98	7	SPOL	98	A	178294112	1637214	418027	101766	15656930	10687353	1799042
4	31AUG98	8	SPOL	98	A	183922396	1828908	541960	117295	7790766	8053527	392477
5	30SEP98	9	SPOL	98	A	174622520	502184	62965	3468	2071539	1723890	109273

Figure 3-69. Resource Element File Contents Report

3.3.3.4 Tracking Applications

The tracking applications case study illustrates how to create a resource element file that describes the resource requirements of a payroll application from the MONTHS timespan of the BATJOB File.

The following topics are discussed by section:

- 1 - Problem Description
- 2 - Control Parameters
- 3 - Output Report

3.3.3.4.1 Problem Description

The tracking applications case study addresses the following areas:

- o Summarizing multiple SYSIDs and ZONES into a composite SYSID and ZONE. Although you may normally expect an application to run during a specific zone or on a specific processor, summarizing data from all zones and systems prevents the loss of data when the application runs in an unexpected manner.
- o Using a selection exit to limit the volume of data that is processed by the program. Rather than summarizing the data for all jobs, the selection exit retains data for the application being studied and excludes data for all of the other jobs.

Note that identifying applications in the MONTHS timespan of the BATJOB File depends on the way that you define the BATJOB File accounting fields. In the test installation, applications are identified by two three-character account codes (ACCTN01 and ACCTN02). The payroll application is identified by the account codes P39 and 028. We chose the following data elements to characterize the application for this case study:

Data Element	Description
=====	=====
JOBTCBTM	Total TCB CPU time used by the application
JOBSRBTM	Total SRB CPU time used by the application
JOBNLR	Total lines printed by the application
JOBMXWSS	The maximum working set size of the application
JOBCOUNT	The total number of jobs executed during the month
JOBSTEPS	The total number of job steps executed during the month
JOBSYSAB	The total number of system ABENDs experienced by the application during the month
JOBUSRAB	The total number of user ABENDs experienced by the application during the month

Although these data elements do not provide a complete characterization of all resources used by the application, they are examples of key indicators that capacity planners would find interesting. The CA MICS BATJOB File contains many other data elements which you could select to augment this capacity database file, based on your own interest.

3.3.3.4.2 Control Parameters

Figures 3-70 through 3-78 show the control screens that we used to create and execute this resource element file example.

```

/----- Resource Element File Definition -----\
|Command ==>|
|Enter a ? in any data entry field for more information on valid values.|
|CAP DB File def.: PAY - Payroll Application Example|
|
|Summarization Level|
|Primary ==> MONTHS (DAYS/WEEKS/MONTHS)|
|Secondary ==> ZONE (ZONE/HOUR/ENDTS/DAYNAME)|
|
|Amount of data retained ==> 60 (1-9999 months)|
|Age of oldest forecast ==> 24 (1-9999 months)|
|History cycles ==> 3 (1-99 cycles)|
|Forecast cycles ==> 3 (1-99 cycles)|
|
|Enable:|
|Notes report ==> N (YES/NO)|
|Contents reports ==> N (YES/NO)|
|
|Specify extended options ==> N (Y/N)|
\-----/

```

Figure 3-70. Resource Element File Definition Screen

Option 1, Insert Common Data Selection, is chosen on the Resource Element File screen shown in Figure 3-71.

```

/----- Common Data Selection -----\
|Command ==>                               |
|Enter a ? in any data entry field for more information on valid values.         |
|Composing CAP DB File def.:  PAY - Payroll Application Example                   |
|                                                                                  |
|SYSID   ==> TS01 TS02  ____  ____  ____  ____  ____  ____  ____  ____  ____  ____|
|          ____  ____  ____  ____  ____  ____  ____  ____  ____  ____|
|Zone    ==>  _ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _|
|Hour    ==>  _ _ - _ _ _ - _ _ - _ _ - _ _ - _ _ - _ _ - _ _ _ _ _ _ _ _ _ _ _ _|
|Dayname ==>  _ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _|
|                                                                                  |
|Specify common data selection SAS statements ==> Y (Y/N/R)                       |
\-----/

```

Figure 3-72. Common Data Selection Screen

The Common Data Selection screen shown in Figure 3-72 specifies that observations with SYSIDs that are equal to TS01 or TS02 are to be included in the resource element data file. SAS statements are specified to further control data selection.

```
/----- SAS System Statements -----\  
|Command ==>                               Scroll ==> CSR |  
|  
|Inquiry Step:  File & Data Element Sel.: PAY - PAYROLL APPLICATION |  
|  
|SAS Statements for the SELECT  Macro |  
|-----|  
|***** ***** TOP OF DATA *****|  
|==MSG> SPECIFY SAS STATEMENTS TO SUBSET THE CAPACITY PLANNING DATA BASE FILE. |  
|=NOTE= %MACRO SELECT; <=== PLEASE DO NOT PLACE LINES BEFORE THIS LINE. |  
|000001 IF ACCTN01 = 'P39' AND ACCTN02 = '028' ; |  
|=NOTE= %MEND SELECT; <=== PLEASE DO NOT PLACE LINES AFTER THIS LINE. |  
|***** ***** BOTTOM OF DATA *****|  
|  
|-----|  
\-----/
```

Figure 3-73. Common Data Selection Screen

Option 2, Insert CA MICS File and Data Element Selection, is chosen on the Resource Element File screen shown in Figure 3-74.

The CA MICS File and Data Element Selection screen shown in Figure 3-75 specifies that the CA MICS JOB file in database complex P is to be used. The data elements to be included in the resource element file are listed.

```

/----- CA MICS File and Data Element Selection -----\
|Command ==>                                         Scroll ==> CSR |
|Enter a ? in any data entry field for more information on valid values. |
|Composing CAP DB File def.: PAY - Payroll Application Example |
|CA MICS file      ==> JOB      (fff) - BATCH USER JOB ACTIVITY FILE |
|Input Time-Span   ==> MONTHS (DETAIL/DAYS/WEEKS/MONTHS) |
|Data base (s)     ==> P |
|Include CAPAPUs   ==> N (Y/N) |
|Summarize file by ==> _____ |
|Specify file exits ==> N (Y/N) |
| |
| |
|-----Line Commands-----|
| I Insert C Copy D Delete |
| R Repeat E Calculations  |
| S Depend Clauses & Label |
|-----|
| |
| Sum Data |
| Cmd Name Label (long name) D Type Type Order Len Output-Format |
|-----|
| _ DATE Interval End Date _ RET D 000 7 DATE7_____ |
| _ JOBCOUNT Job Executions _ ACCM N 050 5 _____ |
| _ JOBMXWSS Max Step Work Set _ MAX N 050 5 _____ |
| _ JOBNLR Total Log. Writer _ ACCM N 050 13 _____ |
| _ JOBSRBTM Job SRB CPU Time _ ACCM T 050 10 TIME14.2_____ |
| _ JOBSTEPS Number of Steps _ ACCM N 050 10 _____ |
| _ JOBSYSAB Job System Abends _ ACCM N 050 5 _____ |
|***** BOTTOM OF DATA *****|
| |
\-----/

```

Figure 3-75. CA MICS File and Data Element Selection Screen - Screen 1

After the resource element file definition parameters are specified, the Resource Element File screen is used to validate the inquiry. Figure 3-76 shows this screen after VALIDATE is typed at the Command prompt.

```
/----- Update Capacity Planning File ----- Confirm Save.-\
|Command ==>|
|Modifying CAP DB File def.: PAY - Payroll Application Example|
|This file definition has been changed and must be saved before updating the|
|file.|
|Do you wish to save the file definition and update the file? ==> YES (YES/NO)|
|Instructions:|
|  Press ENTER key to save the file definition and update the file.|
|  Reply NO and press enter key to cancel the file update request.|
|-----/
```

Figure 3-77. Saving the Inquiry

The Update Capacity Planning file screen shown in Figure 3-77 is used to save the inquiry so that it can be executed.

```
/----- Update Capacity Planning File -----\  
|Command ==>  
|  
|Composing CAP DB File def.:  PAY - Payroll Application Example  
|  
|Type of update                ==> INITIALIZE (INITIALIZE/UPDATE)  
|Replace existing file?        ==> N          (Y/N)  
|  
|Input data source              ==> ONLINE    (ARCHIVE/ONLINE)  
|CA MICS file cycle(s)         ==> 01 - 05    (0 - 99)  
|Archive generation number     ==> ___        (0 to -99)  
|  
|Starting date of update       ==> _____ (ddmonyy)  
|Number of months to select    ==> 5___      (1 to 9999)  
|  
|Mode of execution             ==> BATCH      (BATCH/FOREGROUND)  
|  
|List updated file contents?   ==> YES        (YES/NO)  
|  
|  
|Press END to invoke MICS inquiry execution to update the resource element file.  
|Enter CANCEL to terminate the resource element file update.  
|  
\-----/
```

Figure 3-78. Executing the Inquiry

Cycles 1 through 5 are specified as input to the database loading process on the Update Capacity Planning File screen shown in Figure 3-78. The updated resource element file will be printed after the update is complete.

3.3.3.4.3 Output Report

Figure 3-79 shows the output that is created by the PROC PRINT of the updated data set.

CA MICS Capacity Planner													
OBS	DATE	JOB COUNT	JOB NL	JOB SRBTM	JOB STEPS	JOB SYSAB	JOB TCBTM	JOB USRAB	JOB MXWSS	MONTH	SYSID	YEAR	ZONE
1	31MAY98	470	69572	0:49:06.09	1725	10	16:56:39.10	9	280	5	TS0A	98	A
2	30JUN98	682	44740	1:09:51.00	2190	46	21:30:41.10	20	491	6	TS0A	98	A
3	31JUL98	556	78848	1:04:02.63	1439	33	19:55:22.87	13	724	7	TS0A	98	A
4	31AUG98	718	153084	1:04:57.95	1618	57	20:01:14.77	26	608	8	TS0A	98	A
5	30SEP98	559	57062	0:56:43.29	1495	24	17:10:10.31	32	1599	9	TS0A	98	A

Figure 3-79. Resource Element File Contents Report

3.3.3.5 System Workload and Performance Profile

The system workload and performance profile case study illustrates how to combine data elements from multiple CA MICS files into a single capacity planning database file to allow you to track system workload and performance levels.

The following topics are discussed by section:

- 1 - Problem Description
- 2 - Control Parameters
- 3 - Output Report

3.3.3.5.1 Problem Description

For this example, we selected the following seven data elements from three CA MICS files:

Data Element	CA MICS File	Comments
CPUMPBS	HARCPU	Average CPU multi-processor busy. This user defined data element is computed from CPUPC0BS and CPUPC1BS, which are the percent busy values for CPUs 0 and 1, respectively.
CPUMXT	HARCPU	The maximum number of TSO users.
PAGPSSPG	SCPPAG	Average pages per second for the system.
TS0N0LOG	TS0TS0	Total number of TSO logons.
TS0UPTM	TS0TS0	TSO up time.
TS0AVSTM	TS0TS0	TSO average short response time.
TS0AVTTM	TS0TS0	Average response time for all TSO commands.

Although these data elements do not provide a complete performance and workload report for the system, they provide an overview of the system's activities and illustrate how you can use the capacity planning database generator to assemble data elements from a variety of CA MICS files.

3.3.3.5.2 Control Parameters

Figures 3-80 through 3-91 show the control screens used to create and execute this resource element file example.

```

/----- Resource Element File Definition -----\
|Command ==>|
|Enter a ? in any data entry field for more information on valid values.|
|CAP DB File def.: WKL - System Workload Example|
|
|Summarization Level|
|Primary ==> WEEKS (DAYS/WEEKS/MONTHS)|
|Secondary ==> ZONE (ZONE/HOUR/ENDTS/DAYNAME)|
|
|Amount of data retained ==> 60 (1-9999 weeks)|
|Age of oldest forecast ==> 24 (1-9999 weeks)|
|History cycles ==> 3 (1-99 cycles)|
|Forecast cycles ==> 3 (1-99 cycles)|
|
|Enable:|
|Notes report ==> N (YES/NO)|
|Contents reports ==> N (YES/NO)|
|
|Specify extended options ==> N (Y/N)|
\-----/

```

Figure 3-80. Resource Element File Definition Screen

Option 1, Insert Common Data Selection, is chosen on the Resource Element File screen shown in Figure 3-81.

```

/----- Common Data Selection -----\
|Command ==>                               |
|Enter a ? in any data entry field for more information on valid values.  |
|Composing CAP DB File def.: WKL - System Workload Example                |
|                                                                           |
|SYSID   ==> TS01  _ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _ |
|               _ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _ |
|                                                                           |
|Zone    ==> 1  _ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _ |
|                                                                           |
|Hour    ==> _ _ - _ _ - _ _ - _ _ - _ _ - _ _ - _ _ - _ _ _ _ _ _ _ _ |
|                                                                           |
|Dayname ==> _ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _ |
|                                                                           |
|                                                                           |
|                                                                           |
|Specify common data selection SAS statements ==> N (Y/N/R)                |
\-----/

```

Figure 3-82. Common Data Selection Screen

The Common Data Selection screen shown in Figure 3-82 specifies that observations with SYSIDs equal to TS01 and with ZONE 1 are to be included in the resource element data file.

[illegible]

Figure 3-83. Specifying CA MICS Files and Data Elements


```

/----- CA MICS File and Data Element Selection -----
|Command ==>                                         Scroll ==> CSR |
|Enter a ? in any data entry field for more information on valid values. |
|Modifying CAP DB File def.: WKL - System Workload Example |
|CA MICS file      ==> CPU      (fff) - CPU PROCESSOR ACTIVITY FILE |
|Input Time-Span   ==> WEEKS   (DETAIL/DAYS/WEEKS/MONTHS) |
|Data base (s)     ==> P |
|Include CAPAPUs   ==> N (Y/N) |-----Line Commands-----|
|Summarize file by ==> ----- | I Insert C Copy D Delete |
|Specify file exits ==> N (Y/N) | R Repeat E Calculations |
|                                     | S Depend Clauses & Label |
|                                     -----|
|                                     Sum Data |
|Cmd  Name      Label (long name)          D Type Type Order Len Output-Format |
|-----|-----|-----|-----|-----|-----|-----|-----|
|_  YEAR      Year of Century              _ RET  N    000   2  -----|
|_  ZONE      Time Zone                    _ RET  A    000   1  -----|
|-----|-----|-----|-----|-----|-----|-----|

```

Figure 3-84. CA MICS File and Data Element Selection Screen - Screen 2

The CA MICS File Processing Exits Menu shown in Figure 3-85 indicates that the data element derivation logic is to be defined by the user.

```
/----- SAS System Statements -----\  
|Command ==>                               Scroll ==> CSR |  
|  
|Inquiry Step:  File & Data Element Sel.: CPU - CPU PROCESSOR ACTIVITY FILE |  
|  
|SAS Statements for the CAPSMRY Macro |  
|-----|  
|***** ***** TOP OF DATA *****|  
|==MSG> SPECIFY SAS STATEMENTS FOR TRANSPOSING THE FILE OR TO SET VARIABLES. |  
|=NOTE= %MACRO CAPSMRY; <=== PLEASE DO NOT PLACE LINES BEFORE THIS LINE. |  
|000001 CPUMPBS = (CPUPCOBS + CPUPC1BS) /2 ; |  
|=NOTE= %MEND CAPSMRY; <=== PLEASE DO NOT PLACE LINES AFTER THIS LINE. |  
|***** ***** BOTTOM OF DATA *****|  
|  
|-----/
```

Figure 3-86. Creating a User-defined Data Element

The CA MICS File and Data Element Selection screen shown in Figure 3-88 specifies that the CA MICS T50 file in database unit P is to be used. The data elements to be included in the resource element file are listed.

/----- CA MICS File and Data Element Selection -----							
Command ==>				Scroll ==> CSR			
Enter a ? in any data entry field for more information on valid values.							
Modifying CAP DB File def.: WKL - System Workload Example							
CA MICS file ==> TSO (fff) - TSO SYSTEM ACTIVITY FILE							
Input Time-Span ==> WEEKS (DETAIL/DAYS/WEEKS/MONTHS)							
Data base (s) ==> P							
Include CAPAPUs ==> N (Y/N)				-----Line Commands-----			
Summarize file by ==> _____				I Insert C Copy D Delete			
Specify file exits ==> N (Y/N)				R Repeat E Calculations			
				S Depend Clauses & Label			

Sum Data							
Cmd	Name	Label (long name)	D	Type	Type	Order	Len Output-Format

_	DATE	Interval End Date					
			_	RET	D	000	7 DATE7._____
_	MONTH	Month of Year					
			_	RET	N	000	2 _____
_	SYSID	System Identifier					
			_	RET	A	000	4 _____
_	TSOAVSTM	Avg Short Response					
			_	COMP	N	050	7 TIME._____
_	TSOAVTTM	Avg Response All					
			_	COMP	N	050	7 TIME._____
_	TSOONOLOG	User Logons					
			_	ACCM	N	050	7 _____
_	TSOUPTM	TSO Avail Time					
			_	ACCM	N	050	7 TIME._____

Figure 3-88. CA MICS File and Data Element Selection Screen - Screen 1


```

/----- CA MICS File and Data Element Selection -----/
|Command ==>                                         Scroll ==> CSR |
|Enter a ? in any data entry field for more information on valid values. |
|Modifying CAP DB File def.: WKL - System Workload Example |
|CA MICS file      ==> TSO      (fff) - TSO SYSTEM ACTIVITY FILE |
|Input Time-Span   ==> WEEKS   (DETAIL/DAYS/WEEKS/MONTHS) |
|Data base (s)     ==> P |
|Include CAPAPUs   ==> N (Y/N) |-----Line Commands-----|
|Summarize file by ==> _____ | I Insert C Copy D Delete ||
|Specify file exits ==> N (Y/N) | R Repeat E Calculations ||
|                               | S Depend Clauses & Label ||
|                               |-----|
|                               |
|                               | Sum Data |
|Cmd  Name  Label (long name)  D Type Type Order Len Output-Format |
|-----|-----|-----|-----|-----|-----|-----|
|_  YEAR    Year of Century    _ RET  N    000  2  _____ |
|                               |
|_  ZONE    Time Zone          _ RET  A    000  1  _____ |
|                               |
\-----/

```

Figure 3-88. CA MICS File and Data Element Selection Screen - Screen 2

```

/----- Resource Element File -----Inquiry Validated.-\
|Command ==>                               Scroll ==> CSR |
|
|Composing CAP DB File def.: WKL - System Workload Example |
|
|Update the capacity planning data base file ==> Y (Y/N)    |
|
|Line Cmds: S Select  I Insert  D Delete  R Repeat          |
|
|Cmd              Data Base File Definition Step           |
| - -----|
| - Resource Element File Definition                       |
| - Common Data Selection                                  |
| - File & Data Element Sel.: CPU - CPU PROCESSOR ACTIVITY FILE |
| - File & Data Element Sel.: PAG - MVS PAGING ACTIVITY FILE  |
| - File & Data Element Sel.: TSO - TSO SYSTEM ACTIVITY FILE  |
| - Cross-File Data Element Definition                     |
|***** BOTTOM OF DATA *****|
|
|
|
|
|-----/

```

Figure 3-89. Validating the Inquiry

After the resource element file definition parameters are specified, the Resource Element File screen is used to validate the inquiry. Figure 3-89 shows this screen after VALIDATE is typed at the Command prompt.

```
/----- Update Capacity Planning File ----- Confirm Save.-\
|Command ==>|
|          |
|Modifying CAP DB File def.: WKL - System Workload Example|
|          |
|          |
|This file definition has been changed and must be saved before updating the|
|file.      |
|          |
|          |
|Do you wish to save the file definition and update the file? ==> YES (YES/NO)|
|          |
|          |
|Instructions:|
|          |
|  Press ENTER key to save the file definition and update the file.        |
|          |
|  Reply NO and press enter key to cancel the file update request.         |
|          |
|          |
\-----/
```

Figure 3-90. Saving the Inquiry

The Update Capacity Planning File screen shown in Figure 3-90 is used to save the inquiry so that it can be executed.

```

/----- Update Capacity Planning File -----\
|Command ==>|
|
|Composing CAP DB File def.: WKL - System Workload Example|
|
|Type of update          ==> INITIALIZE (INITIALIZE/UPDATE)|
|Replace existing file?  ==> N          (Y/N)|
|
|Input data source       ==> ONLINE    (ARCHIVE/ONLINE)|
|CA MICS file cycle(s)   ==> 01 - 26   (0 - 99)|
|Archive generation number ==> ___      (0 to -99)|
|
|Starting date of update ==> _____ (ddmonyy)|
|Number of months to select ==> 26__   (1 to 9999)|
|
|Mode of execution       ==> BATCH     (BATCH/FOREGROUND)|
|
|List updated file contents? ==> YES    (YES/NO)|
|
|Press END to invoke MICF inquiry execution to update the resource element file.|
|Enter CANCEL to terminate the resource element file update.|
|
\-----/

```

Figure 3-91. Executing the Inquiry

Cycles 1 through 26 are specified as input to the database loading process on the Update Capacity Planning File screen shown in Figure 3-91. The updated resource element file will be printed after the update is complete.

3.3.3.5.3 Output Report

Figure 3-92 shows the output that is created by the PROC PRINT of the updated data set.

CA MICS Capacity Planner												
OBS	DATE	CPUMXT	SYSID	WEEK	YEAR	ZONE	CPUMPBS	PAGPSSPG	TSOONOLOG	TSOUPSTM	TSOAVSTM	TSOAVTTM
1	23APR98	129	TS01	17	98	1	61.4238	33.2198	14592	246:41:13.63	0:00:00.37	0:00:00.79
2	30APR98	167	TS01	18	98	1	58.6719	32.8897	14406	247:26:18.19	0:00:00.56	0:00:00.89
3	07MAY98	138	TS01	19	98	1	64.3906	35.1352	15830	245:41:18.13	0:00:00.38	0:00:00.78
4	14MAY98	146	TS01	20	98	1	63.8496	34.2255	15348	247:06:19.88	0:00:00.44	0:00:00.86
5	21MAY98	148	TS01	21	98	1	63.3184	36.5478	15486	247:51:21.15	0:00:00.36	0:00:00.85
6	28MAY98	133	TS01	22	98	1	59.0957	31.8148	14422	247:56:19.52	0:00:00.36	0:00:00.73
7	04JUN98	125	TS01	23	98	1	59.9043	30.9136	11020	197:55:55.96	0:00:00.39	0:00:00.79
8	11JUN98	148	TS01	24	98	1	57.8789	34.2631	15092	247:51:23.49	0:00:00.39	0:00:00.76
9	18JUN98	208	TS01	25	98	1	59.2773	36.3417	15040	244:21:18.72	0:00:00.43	0:00:00.88
10	25JUN98	159	TS01	26	98	1	59.1797	33.9071	14786	238:22:05.21	0:00:00.41	0:00:00.78
11	02JUL98	139	TS01	27	98	1	66.5547	47.0149	15744	242:06:34.54	0:00:00.55	0:00:00.98
12	09JUL98	136	TS01	28	98	1	53.3340	33.6379	11890	236:46:54.05	0:00:00.50	0:00:00.89
13	16JUL98	182	TS01	29	98	1	69.8828	63.7973	16778	247:42:56.80	0:00:00.78	0:00:01.26
14	23JUL98	149	TS01	30	98	1	67.2285	53.0309	15504	246:19:56.86	0:00:00.66	0:00:01.13
15	30JUL98	196	TS01	31	98	1	71.9414	64.2385	14598	246:56:42.68	0:00:00.74	0:00:01.22
16	06AUG98	133	TS01	32	98	1	64.1016	42.4233	14620	247:51:18.95	0:00:00.40	0:00:00.76
17	13AUG98	242	TS01	33	98	1	63.3496	46.0162	14476	242:50:34.61	0:00:00.50	0:00:01.07
18	20AUG98	132	TS01	34	98	1	64.4785	45.5301	14484	242:11:34.16	0:00:00.47	0:00:00.96
19	27AUG98	115	TS01	35	98	1	63.3281	46.4081	14020	247:36:16.22	0:00:00.44	0:00:00.88
20	03SEP98	150	TS01	36	98	1	65.8555	46.0491	14532	245:41:21.55	0:00:00.47	0:00:00.92
21	08SEP98	122	TS01	37	98	1	76.2949	47.3450	11130	187:51:25.56	0:00:01.01	0:00:01.86
22	16SEP98	123	TS01	38	98	1	60.0039	47.9458	12872	230:56:22.15	0:00:00.65	0:00:01.46
23	20SEP98	136	TS01	39	98	1	66.7480	48.3453	14008	246:01:32.37	0:00:00.80	0:00:01.45
24	28SEP98	119	TS01	40	98	1	67.2891	45.1462	11718	194:41:15.07	0:00:00.80	0:00:01.42
25	04OCT98	136	TS01	41	98	1	69.3281	49.2254	15534	249:46:46.15	0:00:00.72	0:00:01.47
26	12OCT98	123	TS01	42	98	1	64.3730	44.5164	11774	200:31:19.81	0:00:00.66	0:00:01.33

Figure 3-92. Resource Element File Contents Report

3.4 Meta Files

This section describes how to operate the CA MICS Capacity Planner to create user-defined meta capacity planning files. Meta files are similar to resource element files in format, however rather than using the data in the CA MICS database, meta files use either a user's private Capacity Planning database or the shared Capacity Planning database for its input. Extensions have been added to the Meta file facility to resummairize CA MICS data elements. Furthermore, the user can specify the SAS statements to derive new elements based on the file being read or on the relationship among several files.

Through the use of the Meta File Facility of the CA MICS Capacity Planner you can create new capacity planning files that can summarize existing files in the capacity planning database or consolidate capacity planner files into unified consistent database files.

The following sections describe the screens that control this facility:

- 1 - Usage Guidelines
- 2 - Screen Operation
- 3 - Meta File Case Studies

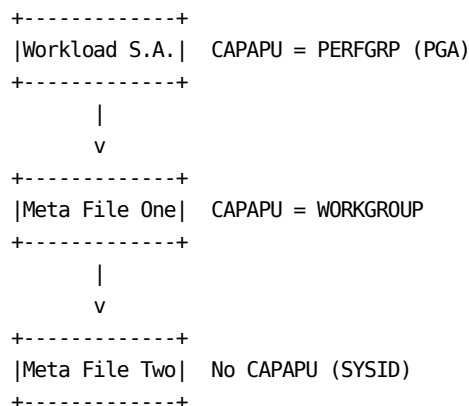
3.4.1 Usage Guidelines

Meta files hold subsets of CA MICS Capacity Planner data element values at either the DAYS, WEEKS or MONTHS timespan. You select the data elements for your meta file. Like business element files, there can be any number of meta files. A three-character user-supplied name identifies each of the meta file definitions.

The structure and format of the meta files is the same as that found for resource element files (refer to Section 3.3.1 for a full discussion), however, there is one major difference between a resource element file and a meta file. Meta files can be chained together so that the output of one meta file update can be used as input to another meta file. With this, the user can create summarized files that only depend on one extract from the CA MICS database. A brief example will illustrate this summarization feature (see the meta file case studies for a more complete example):

Given a capacity planner file based on the Workload Standard Application and the Workload Mapping Definition has a one-to-one mapping between the performance group number and the CAPAPU value. The resulting capacity planner file has one observation for each unique performance group in CA MICS. If there was a need to see this information summarized at the workgroup level (say a group of performance groups) and also summarized for the entire system, two meta files would address this need. Before the introduction of the Meta File facility, the user would have had to create two more standard application files with the necessary Mapping definitions to provide the desired summarization.

By using the Meta File facility, the user could create a meta file with a CAPAPU that corresponded to the workgroup level and then create another meta file that used the previous file as input and summarized up to the SYSID level. Graphically this looks like this:



3.4.2 Screen Operation

The various online screens that the CA MICS Capacity Planner uses to create, modify, and execute user-defined meta files are described in the following sections:

- 1 - Creating a Meta File Specification
- 2 - Modifying a Meta File Specification
- 3 - Executing a Meta File Specification

3.4.2.1 Creating a Database Specification

The online screens that the CA MICS Capacity Planner uses to create user-defined meta files are described in the following sections:

- 1 - File Definition
- 2 - Element Selection
- 3 - Specification Processing

3.4.2.1.1 File Definition

Figure 3-93 shows the first step for creating a user-defined meta file.

```

/----- Meta File Definition -----\
|Command ==>|
|Enter a ? in any data entry field for more information on valid values.|
|Composing CAP DB File def.: SYS - System Activity Tracking File|
|
|Summarization Level|
|          Primary  ==> MONTHS   (DETAIL/DAYS/WEEKS/MONTHS)|
|          Secondary ==> ZONE    (ZONE/HOUR/ENDTS/DAYNAME)|
|
|Capacity Planner Data Base ==> PRIVATE (Private/Shared)|
|
|Amount of data retained  ==> 60   (1-9999 months)|
|Age of oldest forecast   ==> 24   (1-9999 months)|
|History cycles           ==> 3    (1-99 cycles)|
|Forecast cycles          ==> 3    (1-99 cycles)|
|
|Enable:|
|  Notes report           ==> YES   (YES/NO)|
|  Contents reports       ==> YES   (YES/NO)|
|
|Specify extended options ==> N     (Y/N/R)|
\-----/

```

Figure 3-93. Meta File Definition Screen

The Meta File Definition screen shown in Figure 3-93 is the main control screen for defining meta files. It contains the following fields:

Summarization level: The output summarization corresponds to the combination of the primary and secondary summarization levels. These control how the output file is summarized in your Capacity Planner database.

Note: When you choose a secondary summarization level of ENDTS, you can only select one CA MICS file as input. This is due to the difficulty in merging ENDTS values across multiple files.

Capacity Planner Database: The specification that tells where ALL input files that make up this Meta file reside. Choose PRIVATE (the default) if you wish the meta file to be made up of information in your own Capacity Planning database. Choose SHARED if you wish to process information in the Shared Capacity Planning database.

Note: If this meta file will ever be executed using the Production Updating Feature of CA MICS, this field must be set to SHARED, because you will not be able to copy the file definition from your private catalog to the shared catalog.

Amount of data retained: The number of days, weeks, or months of historical data to be retained in the CAPxxx01 cycle of the Capacity Planning database file. Any historical observations older than what you specify for this parameter are dropped from the file during the file update.

Age of oldest forecast: The number of days, weeks, or months that you wish to retain a forecast after it is generated. Any forecast older than what you specify for this parameter is dropped when a new forecast is generated and saved.

History cycles: The number of cycles or versions of the historical data that you wish to retain. A new cycle is created when the file is updated and the existing cycles are aged (that is, Cycle 01 becomes Cycle 02 and Cycle 02 becomes Cycle 03).

Forecast cycles: The number of cycles or versions of the forecast data that you wish to retain. A new cycle is created when you specify forecast aging in one of the workload forecasting routines and the existing cycles are aged (that is, Cycle 01 becomes Cycle 02 and Cycle 02 becomes Cycle 03).

Specify The specification for selecting extended
 Extended options. If you specify Y (YES), the Extended
 Options: Options Menu shown in Figure 3-94 displays.

```

/----- Extended Options Menu -----\
|Command ==>
|
|Inquiry Step:
|
|Combine SYSIDs into a composite ==> N      (Y/N)
|      Composite SYSID ==> _____
|Combine zones into a composite ==> N      (Y/N)
|      Composite zone ==> _____
|
|Number of CAPAPUS ==> 0      (0 through 9)
|CAPAPU title (First APU) ==> _____
|CAPAPU2 title ==> _____
|CAPAPU3 title ==> _____
|CAPAPU4 title ==> _____
|CAPAPU5 title ==> _____
|CAPAPU6 title ==> _____
|CAPAPU7 title ==> _____
|CAPAPU8 title ==> _____
|CAPAPU9 title ==> _____
|
|Specify
|  SAS Statements to set CAPAPUs ==> N      (Y/N/R)
|
\-----/

```

Figure 3-94. Meta File Extended Options Menu

Combine all The specification for combining SYSIDs in the
 SYSIDs into Capacity Planning database file. Specify a Y
 a composite: (YES) if you wish data for all of the SYSIDs to
 be summarized into a single observation with
 the Composite SYSID value.

Combine all The specification for combining zones/hours in
 zones/hours/ the Capacity Planning database file. The
 daynames prompt for this field changes to zones, hours or
 into a daynames depending on what you specify for the
 composite: summary level parameter. Specify a Y (YES) if
 you want all of the zones/hours/daynames to be
 summarized into a single observation with the
 Composite zone/hour/dayname value.

Number of CAPAPUs:	Specify the number (0 to 9) of application units (CAPAPUs) you are defining for this element. This activates the optional CAPAPU feature so that these data elements are present in the Capacity Planning database file, and are included in the summarization sequence during file summarization and merging. If you specify that CAPAPUs are to be active, you will be prompted to specify the SAS statements needed to derive the CAPAPU groupings and corresponding CAPAPU values.
CAPAPU titles:	Specify the 1-40 character title for each CAPAPU you are defining for this data element.
Specify SAS Statements to set CAPAPUs:	The specification that allows you to use your own SAS Statements for defining your resource element file. Specify Y (YES) to display the SAS Statements screen shown in Figure 3-95. This option is only available if you specify CAPAPUs.

```

/----- SAS System Statements -----\
|Command ==>                               Scroll ==> CSR |
|
|SAS Statements for the CAPAURT Macro          |
|-----|
|***** ***** TOP OF DATA *****|
|=MSG> SPECIFY SAS STATEMENTS TO SET CAPAPU VALUES. |
|=NOTE= %MACRO CAPAURT; <=== PLEASE DO NOT PLACE LINES BEFORE THIS LINE. |
|000001                                         |
|000002                                         |
|000003                                         |
|000004                                         |
|000005                                         |
|000006                                         |
|000007                                         |
|000008                                         |
|000009                                         |
|000010                                         |
|=NOTE= %MEND CAPAURT; <=== PLEASE DO NOT PLACE LINES AFTER THIS LINE. |
|***** ***** BOTTOM OF DATA *****|
\-----/

```

Figure 3-95. SAS Statements Screen - Defining CAPAPU

The SAS Statements screen shown in Figure 3-95 allows you to define the SAS statements required to group CAP input observations into application workload groups that are appropriate for your installation. All statements that you specify must conform with the syntax specified in the SAS manuals.

Two special data elements are available to you for use in defining your CAPAPU values. The first, CAPIFILE, is a three-character data element that contains the CAP file name that corresponds to the CAP file that you chose for processing after you selected the Insert CAP File and Data Element Selection option on the Meta File menu (see Section 3.4.2.1.2). The second, CAPIAPU (including CAPIAPU2 to CAPIAPU9), corresponds to the CAPAPU value found for the current observation being read from the input CAP file that you chose. For example, if you specify the Insert CAP File and Data Element Selection twice, once for CAP file: CPG, which has one observation for each Performance Group and again for CAP file: CPA, which has one observation for each Performance Group/Performance Period combination. CAPIFILE would be 'CPG' when reading the CPG file and 'CPA' when reading the CPA file. The following example illustrates the use of the CAPIFILE and CAPIAPU data elements in defining new CAPAPU values for the output meta file:

```
%MACRO CAPAURT;
  IF CAPIFILE EQ 'CPA' AND SUBSTR(CAPIAPU,1,1) EQ 2 THEN DO;
    IF SUBSTR(CAPIAPU,2,1) EQ 1 THEN CAPAPU='WORKLOAD1';
    ELSE IF SUBSTR(CAPIAPU,2,1) EQ 2 THEN CAPAPU='WORKLOAD2';
    ELSE IF SUBSTR(CAPIAPU,2,1) EQ 3 THEN CAPAPU='WORKLOAD3';
    ELSE IF SUBSTR(CAPIAPU,2,1) EQ 4 THEN CAPAPU='WORKLOAD4';
    ELSE IF SUBSTR(CAPIAPU,2,1) EQ 5 THEN CAPAPU='WORKLOAD5';
  END;
  ELSE IF CAPIFILE EQ 'CPG' THEN DO;
    IF SUBSTR(CAPIAPU,1,1) EQ 7 THEN CAPAPU='WORKLOAD6';
    ELSE IF SUBSTR(CAPIAPU,1,2) EQ 62 THEN CAPAPU='WORKLOAD7';
    ELSE IF SUBSTR(CAPIAPU,1,1) NE 2 THEN CAPAPU='WORKLOAD8';
  END;
%MEND CAPAURT;
```

3.4.2.1.2 Element Selection

Figure 3-96, Meta File menu, shows the next step in creating a user-defined meta file definition.

```

----- Meta File -----
|Option ==>
|
|Composing CAP DB file def.: SYS - System Activity Tracking File
|
|    1 - Insert Common Data Selection
|        Specify data selection criteria to apply to ALL files.
|    2 - Insert CAP File and Data Element Selection
|        Select a Capacity Planner file and data elements.
|    3 - Insert Cross-File Data Element Definition
|        Specify cross-file data element derivations.

```

Figure 3-96. Meta File Menu

After you complete specifying the meta file definition control parameters, you are ready to define the steps to build your Capacity Planning database file. (Refer to Figure 3-2 to see how these steps correspond to user specifications for the selection, summary, and merge phases of database creation.) The Meta File menu provides you with the opportunity to define selection criteria that you wish to apply to each CAP file that you define as an input data source for your meta file.

You may choose from the following options on the Meta File menu:

Insert Common Data Selection: Specify Option 1, Insert Common Data Selection, to specify the data selection criteria that you wish to apply to every CAP file that provides input data in building your meta file. The Common Data Selection screen shown in Figure 3-97 is displayed.

You may specify only one common data selection step for each meta file. You must complete this before proceeding with the succeeding steps.

Insert CAP File and Data Element Selection: Specify Option 2, Insert CAP File and Data Element Selection, to define a CAP file and data element selection step for each CAP file that provides input data for your meta file. The CAP File and Data Element Selection screen shown in Figure 3-99 is displayed. (Both the file-unique selection and summary phases shown in Figure 3-100 occur during this step.) You may select this option more than once.

Using this screen and its associated exit screens, you can define the CAP data elements that will comprise your meta file, along with any file-unique data selection and/or summarization logic.

If you wish to define any data elements that are derived from data elements from a single CAP file, you can usually specify the data element derivation SAS statements through one of the exit screens associated with this screen. If you are defining data elements that are derived from data elements from multiple CAP files, you can specify the data element derivation SAS statements in the cross-file data element definition step.

Insert Cross-File Data Element Definition: Specify Option 3, Insert Cross-File Data Element Definition, to define any data elements that you wish to derive from data elements from multiple CAP files. The Cross-File Data Element Definition screen shown in Figure 3-102 is displayed. (This step corresponds to the merge phase shown in Figure 3-2.) You must also specify the data element derivation SAS statements that will be used to derive these data elements.

Since the specifications associated with this step apply to the merging of data from multiple CAP files, you may specify a cross-file data element definition step only if you specify multiple CAP file and data element selection steps. In this case, the cross-file data element definition step is required to create a valid file definition inquiry.

Each time you complete the specifications for an individual step, END from the screen to redisplay the Meta File Menu shown in Figure 3-104. You can then insert another step in your meta file definition. Repeat this process until you define all of the steps necessary to build your meta file.

The three options are further discussed in the following sections:

- 1 - Insert Common Data Selection
- 2 - Insert CAP File and Data Element Selection
- 3 - Insert Cross-File Data Element Definition

3.4.2.1.2.1 Insert Common Data Selection

If you select Option 1, Insert Common Data Selection, on the Meta File menu, the Common Data Selection screen shown in Figure 3-97 is displayed.

```

/----- Common Data Selection -----
|Command ==>
|Enter a ? in any data entry field for more information on valid values.
|Composing CAP DB File def.:  SYS - System Activity Tracking File
|
|SYSID   ==>  _ _ _ _ _
|           _ _ _ _ _
|
|Zone    ==>  _ _ _ _ _ (1 - 9)
|
|Hour    ==>  _ _ - _ _ - _ _ - _ _ - _ _ - _ _ (0 - 23)
|
|Dayname ==>  _ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _
|
|
|Specify common data selection SAS statements ==> N (Y/N/R)
\-----

```

Figure 3-97. Common Data Selection Screen

The data selection criteria that you specify using this screen is applied to each CAP file that provides input data in creating and updating your Capacity Planning database file.

The parameters for the Common Data Selection screen are defined below:

- SYSID:** The SYSIDs that you wish to include in your Capacity Planning database file. You may specify a maximum of 15 one- to four-character SYSIDs. Leave this parameter blank to include data for all SYSIDs. The standard method is to create one record for each SYSID that is created for each day, week, or month processed. Alternatively, you may wish to summarize multiple SYSIDs into a single record.
- Zone:** The zones that you wish to include in your Capacity Planning database file. You may specify a maximum of nine one-character zones. Leave this parameter blank to include data for all zones. The standard method is to create one record for each zone that is created for each week or month processed. Alternatively, you may wish to summarize multiple zones into a single record.
- Hour:** The hours that you wish to include in your Capacity Planning database file. You may specify a maximum of five hours or hour ranges. Leave this parameter blank to include data for all hours.
- Dayname:** The specification of up to five daynames to be included in the Capacity Planner database file. Leave blank to include all days. Input the first unique characters and the Capacity Planner will fill in the rest. Specification of this parameter is only valid when the input timespan is DETAIL or DAYS.

Specify common data selection SAS statements:	The specification for the SYSIDs and zones or hours to be used for data selection from the CAP input file(s). Specify Y (YES) to display the SAS Statements screen shown in Figure 3-98. Specify R (Reset) to reset any specifications you previously made to SYSID and zone or hour.


```

/----- SAS System Statements -----\
|Command ==>                               Scroll ==> CSR |
|SAS Statements for the SELECT Macro          |
|-----|
|***** TOP OF DATA *****|
|==MSG> ENTER THE SAS STATEMENTS TO SET CAPAPU VALUES. |
|=NOTE= %MACRO SELECT; <=== PLEASE DO NOT PLACE LINES BEFORE THIS LINE. |
|000001|
|000002|
|000003|
|000004|
|000005|
|000006|
|000007|
|000008|
|000009|
|000010|
|=NOTE= %MEND SELECT; <=== PLEASE DO NOT PLACE LINES AFTER THIS LINE. |
|***** BOTTOM OF DATA *****|
| |
\-----/

```

Figure 3-98. SAS Statements Screen - Selecting Common Data

Using the SAS Statements screen shown above, you can specify your own data selection logic to apply to each CAP file that provides input data to your Capacity Planning database file. The special data element, CAPIFILE, is available to the SELECT MACRO and can be used in subsetting the input file.

Note that the %MACRO SELECT and %MEND SELECT statements are provided for you and that they are field-protected from modification.

3.4.2.1.2.2 Insert Capacity File and Data Element Selection

If you select Option 2, Insert CAP File and Data Element Selection, on the Meta File menu, the CAP File and Data Element Selection screen shown in Figure 3-99 is displayed.

```

/----- Capacity Planner File and Data Element Selection -----\
|Command ==>                                     Scroll ==> CSR |
|Enter a ? in any data entry field for more information on valid values. |
|Composing CAP DB File def.:  SYS - System Activity Tracking File      |
|                               |                                       |
|CAP file           ==> WK1 (fff) - Workload Standard Application      |
|                               |                                       |
|                               |-----Line Commands-----|
|Include CAPAPUs    ==> N (Y/N)      | I Insert C Copy D Delete |
|Specify file exits ==> N (Y/N)      | R Repeat E Calculations  |
|                               | S Depend Clauses & Label   |
|                               |-----|
|                               |                               |
|                               | Sum  Data                    |
|Cmd  Name    Label (long name)  D Type Type Order Len Output Format|
|-----|-----|-----|-----|-----|-----|-----|
|_  DATE      Interval End Date  _ RET  D      _____ 7  DATE7.  |
|_  MONTH     Month of Year      _ RET  N      _____ 2  _____|
|_  SYSID     System Identifier   _ RET  A      _____ 4  _____|
|_  YEAR      Year of Century     _ RET  N      _____ 2  _____|
|_  ZONE      Time Zone           _ RET  A      _____ 1  _____|
|***** BOTTOM OF DATA *****|
|                               |
\-----/

```

Figure 3-99. CAP File and Data Element Selection Screen

The parameters for the CAP File and Data Element Selection screen are defined as follow:

CAP file: The Capacity Planner file that contains the data elements that you wish to include in your meta file or that you wish to use in deriving data elements. If you are uncertain of the file name, type a question mark (?) in this field to obtain a complete selection list for your site.

- Include CAPAPUs: The specification for whether the application unit (CAPAPU) data element is to be included in the summarization sequence for this file. This parameter defaults to Y (YES) if CAPAPU is defined for this meta file. The CAPAPU derivation exit is invoked for each input file observation. If you do not want to include CAPAPU in the summarization sequence, type an N (NO) in this field.
- Specify file exits: The specification for CAP file processing exits. You must specify a Y (YES) if you intend to code CAP file-unique data selection, or data derivation logic. If you specify a Y, the CAP File Processing Exits Menu shown in Figure 3-100 is displayed.

The display portion of this screen contains a scrollable list of the data elements that you can either include in your Capacity Planning database file or use to derive other data elements. It contains the following fields:

Cmd: The line commands are:

- I (insert) to insert a new data element.
- R (repeat) to repeat a data element.
- S (depend clauses and label) to specify independent data elements upon which this data element depends. See Figure 3-14.
- C (copy) to copy a data element.
- E (expression) to create SAS code to calculate the value of this specific data element. See Figure 3-15.
- D (delete) to delete a data element.

Name: The name of the data element. This may be either the name of a data element that exists in the input CAP file or any valid SAS name, as described in the SAS manuals. If you wish to define new data elements, choose variable names that are descriptive of the values they represent.

Note that when you first see this screen, it displays only the control data elements that are always present in your Capacity Planning database file. If you attempt to modify these data elements, an error will occur. Define an additional row for each data element that you wish to include in your Capacity Planning database file by typing an I (INSERT) in the Cmd column of any existing row.

If you are uncertain of the correct spelling of a CAP data element or if you would like to see a complete list of the CAP data elements available in the Capacity Planner input file, type a question mark (?) in the name field to display the screen shown in Figure 3-101.

Data Element Label (long name): The data element label that is displayed on the CAP File and Data Element Selection screen and on many reports. This label is supplied automatically for input data elements.

D: The data element drop flag. Typing a D in this

field makes the corresponding data element available for use in calculations through the cross-file data element definition step of the Capacity Planning database file update, but drops the data element from the resulting Capacity Planning database file. Leave the field blank to keep the corresponding data element on the Capacity Planning database file.

Sum Type: The summarization type of the data element. Summarization types are RET (retain), ACCM (accumulate), MIN (minimum), MAX (maximum), and COMP (compute).

Date type: The type of data element. This is used to validate data values and usage for the data element. Valid types are Alphanumeric (A), Numeric (N), Money (M), Time (T), Hex (H), Date (D), Time-of-day (TD), and Date-Time (DT). This value is supplied automatically for CA MICS data elements.

Order: Define the execution order of the data elements you defined by entering ordinal numeric values greater than 1000.

Note: This is required if you are going to use this Resource Element file as input to a Meta file.

The order value is automatically defined for CA MICS data elements and Capacity Planner Control elements. Data elements that require other data elements are automatically included in the Capacity Planner database file. In particular, two variables that will not show up in the output file: @@FIRST and @@LAST.

Len: The length of the data element. This value is supplied automatically for CA MICS data elements.

Output Format: The SAS format that is used to print data element values. This value is supplied automatically for CA MICS data elements.

```
/----- Capacity Planner File Processing Exits Menu -----\
|Command ==>|
|           |
```

```
|Inquiry Step:  File & Data Element Sel.: WK1 -  Workload Standard Application |
|
|Specify file unique data selection SAS stmts.  ==> N (Y/N/R)
|  (file level SELECT exit)
|
|Specify new data element derivation SAS stmts. ==> N (Y/N/R)
|  (file transpose and summary exit)
|
|Set expression statements for data elements.  ==> N (Y/N)
|  (convert SAS stmts to Expression stmts)
|
|
|
|-----/
```

Figure 3-100. CAP File Processing Exits Menu

Figure 3-100 shows the CAP File Processing Exits Menu that is displayed if you specify a Y (YES) for the Specify CAP file processing exits parameter on the CAP File and Data Element Selection screen shown in Figure 3-99.

The following options are available:

Specify
file unique
data Specify a Y (YES) to supply SAS statements in
selection the CAPSLCT macro, used to implement a special
SAS stmts: observation selection method.

Specify Specify a Y (YES) to supply SAS statements in
new data the CAPSMRY macro, used to define new data
element elements for the meta file and to calculate
derivation values for CA MICS derived data elements.
SAS stmts:

Set
expression
statements Set to Y (YES) to convert existing Meta file
for data SAS statements to expression statements or
elements: calculations.

The CAPSMRY exit should only be used for the setting the initial value of non-Computed variables or to transpose the CAP file. The SAS Statements necessary to set the value for a Computed variable should be placed in Expression Statements.

The following illustration describes the CAPSMRY exit.

Data Element Type	Need SAS Code in CAPSMRY?	Need Expression Statement?
Maximum	Yes, to set initial value	No, handled automatically in Meta files
Minimum	Yes, to set initial value	No, handled automatically in Meta files
Accumulated	Yes, to set initial value	No, handled automatically in Meta files
Retained	Yes, to set	No, handled

	initial value	automatically in Meta files
Computed	No	Yes, Expression statements used in Meta files when variable is created and in subsequent Meta Files.

```

/----- Converting Macros to Expression Statements -----\
|Command ==>                                         Scroll ==> CSR |
| |
|Enter the variable names that are associated with the SAS statements shown |
|below. |
| |
|Variable SAS Statements in Macro |
|-----|
|MYVAR2__ IF CPUAVB GT 10 /* IF THE AVERAGE NUMBER OF BATCH JOBS */ |
|MYVAR2__ THEN MYVAR2 = 'CHECK'; /* IS GREATER THAN 10 WE SHOULD */ |
|MYVAR2__ ELSE MYVAR2 = 'OK'; /* CHECK IT OUT */ |
|***** BOTTOM OF DATA *****|
\-----/

```

Figure 3-101. Converting Macros to Expression Statements

You use this screen (see Figure 3-101) to convert existing MACRO statements into Expression Statements that calculate the value of a particular variable. In the example shown in Figure 3-101, the variable MYVAR2 is a user-created data element that is used to warn the user of this file whenever CPUAVB is larger than 10. This variable was originally created using the CAPSMRY MACRO. By placing "MYVAR2" in the Variable Column of this screen the Capacity Planner will automatically associate these SAS Statements with MYVAR2. When you End from this display the Capacity Planner will remove the statements from the MACRO and move them into the Expression Statement table for the variable MYVAR2.

```

/----- Data Element Selection -----\
|Command ==>                               Scroll ==> CSR |
|
|Inquiry Step: Common Data Selection          |
|Input File:  WK1 - Workload Standard Application |
|
|Line Cmds: S Select  H Data Dictionary Display |
|
|          Select Data Elements               |
|Cmd  Element      Data Element Label (long name) |
| -  - - - - - - - - - - - - - - - - - - - - |
| _  CAPAPU        Workload Group               |
| _  CPUDUR        Recording Interval Time       |
| _  CPUTOBTM      Total Time Processor Busy     |
| _  CPUTOUTM      Total Processor Up Time       |
| _  DATE          Interval End Date             |
| _  DESCRIPT      Workload Group Long Name      |
| _  HOUR          Hour of Day                   |
| _  MVSLEVEL      MVS Software Level            |
| _  PAGPIN        Non-VIO Non-Swap Page-Ins     |
| _  PCATRCPU      Percent of CPU Attributable to Workload |
| _  PCTOTCPU      Percent Total CPU Time Consumed |
|
\-----/

```

Figure 3-102. Data Element Selection Screen

If you enter a question mark (?) in the Data Element Label field of the CAP File and Data Element Selection screen shown in Figure 3-99, the Data Element Selection screen (Figure 3-102), which is the final screen in this series, is displayed. This screen lists the data elements in the CAP file that you specified. To include data elements in the meta file, place an S (SELECT) in the Cmd column to the left of the data element you want to include in your meta file.

3.4.2.1.2.3 Insert Cross-File Data Element Definition

If you specify data selection from more than one CAP file (using Option 2 on the Meta File menu shown in Figure 3-96), you must also insert a cross-file data element definition step in your meta file definition. Select Option 3, Insert Cross-File Data Element Definition from the Meta File Menu, to define data elements that are calculated from data elements contained in more more than one file. This screen displays all the data elements, including the control data elements, that you selected from all CAP files.

```

/----- Cross-File Data Element Definition -----\
|Command ==>                                     Scroll ==> CSR |
|Enter a ? in any data entry field for more information on valid values. |
|Composing CAP DB File def.:  SYS - System Activity Tracking File      |
|                                                                    |
|Specify:                                     -----Line Commands-----|
| SAS exit stmts.   ==> N (Y/N/R)           | I Insert C Copy D Delete | |
|                                                                    | |
| Expression stmts. ==> N (Y/N)             | R Repeat E Calculations  | |
|                                                                    | |
|                                                                    | S Depend Clauses & Label | |
|                                                                    | |
|                                                                    | |
|                                                                    | |
|Cmd  Name      Label (long name)           Sum  Data |
|D Type Type Order Len Output Format |
|-----|-----|-----|-----|-----|-----|
|_ CPUAVT      Workload Group                _ RET  N    000   5   _____|
|_ DATE        Workload Group 2              _ RET  D    000   7   DATE7._____|
|_ JOBCPUTM    Workload Group 3              _ RET  T    000  10   TIME14.2_____|
|_ MONTH       Workload Group 4              _ RET  N    000   2   _____|
|_ SYSID       Workload Group 5              _ RET  A    000   4   _____|
|_ YEAR        Workload Group 6              _ RET  N    000   2   _____|
|_ ZONE        Workload Group 7              _ RET  A    000   1   _____|
|***** BOTTOM OF DATA *****|
\-----/

```

Figure 3-103. Cross-File Data Element Definition Screen

The parameters for the Cross-File Data Element Definition screen are defined below:

Specify SAS exit stmts:	The specification for assigning SAS statements to derive cross-file elements. Specify a Y (YES) if you wish to define cross-file data elements. (See Figure 3-103.)
Specify Expression stmts:	The specification for defining Expression statements. Specify a Y (YES) if you wish to define expression statements.
Cmd:	The line commands are: - I (insert) to insert a new data element. - R (repeat) to repeat a data element. - S (depend clauses and label) to specify independent data elements upon which this data element depends. See Figure 3-14. - C (copy) to copy a data element. - E (expression) to create SAS code to calculate the value of this specific data element. See Figure 3-15. - D (delete) to delete a data element.
Name:	The name of the data element. This may be either the name of a CA MICS data element that exists in the input CAP file or any valid SAS name, as described in the SAS manuals. If you wish to define new data elements, choose variable names that are descriptive of the values they represent. Note that the screen displays the CA MICS control data elements that are always present in your meta file. If you attempt to modify these data elements, an error will occur. Define an additional row for each data element that you wish to include in your meta file by typing an I (INSERT) in the Cmd column of any existing row.
Data Element Label (long name):	The data element label that is displayed on the CAP File and Data Element Selection screen and on many reports. This label is supplied automatically for CA MICS data elements.
D:	The data element drop flag. Typing a D in this field makes the corresponding data element available for use in calculations through the

cross-file data element definition step of the Capacity Planning database file update, but drops the data element from the resulting Capacity Planning database file. Leave the field blank to keep the corresponding data element on the Capacity Planning database file.

Sum Type: The summarization type of the data element:
RET (retain), ACCM (accumulate), MIN (minimum), MAX (maximum), and COMP (compute).

Date type: The type of data element. This is used to validate data values and usage for the data element. Valid types are Alphanumeric (A), Numeric (N), Money (M), Time (T), Hex (H), Date (D), Time-of-day (TD), and Date-Time (DT). This value is supplied automatically for CA MICS data elements.

Order: Define the execution order of the data elements you defined by entering ordinal numeric values greater than 3000.

Note: This is required if you are going to use this Resource Element file as input to a Meta file.

The order value is automatically defined for CA MICS data elements and Capacity Planner Control elements. Data elements that require other data elements are automatically included in the Capacity Planner database file. In particular, two variables could appear that will not show up in the output file: @@FIRST and @@LAST.

Len: The length of the data element. This value is supplied automatically for CA MICS data elements.

Output Format: The SAS format that is used to print data element values. This value is supplied automatically for CA MICS data elements.

```
/----- SAS System Statements -----\  
|Command ==>                               Scroll ==> CSR |  
|  
|Inquiry Step: Cross-file data element derivation          |  
|                                                           |
```

```

|SAS Statements for the CAPMERG Macro
|-----|
|***** ***** TOP OF DATA *****|
|==MSG> ENTER THE SAS STATEMENTS TO SET VARIABLES IN THE CROSS-FILE MERGE.|
|=NOTE= %MACRO CAPMERG; <=== PLEASE DO NOT PLACE LINES BEFORE THIS LINE.|
|000001|
|000002|
|000003|
|000004|
|000005|
|000006|
|000007|
|000008|
|000009|
|000010|
|=NOTE= %MEND CAPMERG; <=== PLEASE DO NOT PLACE LINES AFTER THIS LINE.|
|***** ***** BOTTOM OF DATA *****|
|
|-----|
\-----/

```

Figure 3-104. SAS Statements Screen - Deriving Cross-file Data Elements

Using the SAS Statements screen shown in Figure 3-104 above, you can code any valid SAS statements that conform to the syntax described in the SAS manuals to derive new data elements.

Note that the %MACRO CAPMERG and %MEND CAPMERG statements are provided for you and that they are field protected from modification.

```

/-----\
|-----\
|Command ==>                               Scroll ==> CSR |
|
|Enter the variable names that are associated with the SAS statements shown |
|below. |
|
|Variable SAS Statements in Macro |
|-----|
|RATIO___ IF CPUTOBTM GT 0 |
|RATIO___ THEN RATIO = PGACPUTM / CPUTOBTM; |
|RATIO___ ELSE RATIO = .; |
|_____ |
|***** ***** BOTTOM OF DATA *****|
|-----|
\-----/

```

Figure 3-105. Converting Macros to Expression Statements

You use this screen (see Figure 3-105) to convert existing MACRO statements into Expression Statements that calculate the value of a particular variable. In the example shown in Figure 3-105, the variable `RATIO` is a user-created data element that corresponds to the ratio of Performance Group Activity (PGA) file CPU time consumed and the total Busy time of the processor. By placing "RATIO" in the Variable Column of this screen the Capacity Planner will automatically associate these SAS Statements with `RATIO`. When you End from this display the Capacity Planner will remove the statements from the MACRO and move them into the Expression Statement table for the variable `RATIO`.

This screen lists the file definition steps that you defined. You can change specifications for any step by typing an S on the Cmd column of the step you wish to change. In the same way, you can delete, insert, or repeat a step by typing D, I, or R, respectively.

Figure 3-106. Meta File Screen

Before you execute your meta file definition inquiry, you must first validate it to ensure that you have defined all of the necessary steps. For instance, if you specify more than one file and data element selection step, then the online facility ensures that you also specify a cross-file data element definition step to cause the data from each of the file and data element selection steps to be merged.

You may request validation of your meta file definition inquiry by typing a V (or VALIDATE) at the Command prompt of the Meta File screen and pressing ENTER. Diagnostic messages inform you that the inquiry has been validated or why it cannot be validated so you can correct the inquiry step specifications and request validation again.

If you are unable to validate your inquiry and want to defer correcting the problems until a later time, END from this screen to save your unvalidated inquiry. The Capacity Planning Database Files screen, which displays the meta file definition inquiries in your catalog, is displayed.

After you validate your meta file definition inquiry, you have an opportunity to execute the inquiry and update the Capacity Planning database file by typing a Y at the Update Capacity Planning database file prompt and pressing ENTER. The Update Capacity Planning File screen shown in Figure 3-101 is displayed. After execution processing is completed, the Meta File screen is redisplayed. After you are satisfied with all the specifications, type END at the Command prompt to save your meta file definition. The Capacity Planning database File screen is then redisplayed.

3.4.2.2 Modifying a Database Specification

After you define a meta file definition, you can then modify it. The Meta File screen shown in Figure 3-107 shows all of the specifications that you can use to manipulate meta file definitions. You can select a definition by typing an S (SELECT) in the Cmd column of the desired definition. You can also add a new definition by typing an I (INSERT) in the Cmd column of any existing row or by specifying an R (REPEAT) in the Cmd column of a specific row, causing that definition to be duplicated. You may then select the new definition and modify it to conform to your requirements.

```

/----- Meta File -----
|Command ==>                               Scroll ==> CSR
|
|Composing CAP DB File def.: SYS - System Activity Tracking File
|
|Update the capacity planning data base file ==> N (Y/N)
|
|Line Cmds: S Select  I Insert  D Delete  R Repeat
|
|Cmd                Data Base File Definition Step
| - -----
| - Meta File Definition
| - File & Data Element Sel.: VM1 - VM/ESA Standard Application
| - File & Data Element Sel.: ES1 - MVS/ESA Standard Application
| - Cross-File Data Element Definition
|***** BOTTOM OF DATA *****
|
|
|
|-----

```

Figure 3-107. Meta File Screen

3.4.2.3 Executing a Database Specification

When you select a meta file definition a special Update Capacity Planning Meta File screen is displayed (an example is shown in Figure 3-108).

```
/----- Update Capacity Planning Meta File -----\  
|Command ==>|  
|  
|Composing CAP DB File def.:  SYS - System Activity Tracking File|  
|  
|Update type                ==> INITIALIZE (INITIALIZE/UPDATE)|  
|Replace existing           ==> N (Y/N)|  
|Private Capacity Planner data base cycle ==> 01|  
|Update start date          ==> _____ (ddmonyy)|  
|Number of months           ==> _____ (1-9999)|  
|Execution mode             ==> BATCH      (BATCH/BACKGROUND)|  
|  
|List updated CAP DB file contents ==> YES (YES/NO)|  
|  
|  
|  
|  
|  
|Press END to invoke MICF inquiry execution to update the meta file.|  
|Enter CANCEL to terminate the meta file update.|  
\-----/
```

Figure 3-108. Update Capacity Planning Meta File Screen

The parameters for the Update Capacity Planning File screen are described below:

Update type: The specification for initializing or updating the Capacity Planning database file. If the meta file does not exist or if you plan to overwrite the current file, specify INITIALIZE. If you wish to add to an existing file, specify UPDATE.

Note that if you specify INITIALIZE and the file already exists, all historical and forecasted observations that you previously placed in the file will be lost.

Replace existing: The specification for deleting an old file and replacing it with a new file. If you specify INITIALIZE and the file already exists, you must set this parameter to Y (YES) to delete the old version and confirm your intent to initialize and replace. Otherwise, specify N (NO).

Capacity Planner Database Cycle: The Private (or Shared) Capacity Planner database file cycle you wish to use as input into the meta file update run.

Update start date: The start date for date range selection. If Start date is specified, only historical observations at least as recent as this value are included in the analysis. The discussion of the following control parameter describes the options for controlling date range selection.

Number of days/ weeks/ months: The maximum number of days, weeks, or months of historical observations to be included in the analysis. Note that once you specify the input file, the field name changes to reflect the time-span of the input file. When used in conjunction with the Start date parameter, several options are available for date range selection control:

- o If you specify only Update start date, all historical observations at least as recent as the value specified for Update start date is selected.

- o If you specify both Update start date and Number days/weeks/months, the date range for data selection begins with the specified value for Start date and extends for the specified number of days, weeks, or months.
- o If you specify only Number of days/weeks/months, the date range for data selection includes that number of days, weeks or months prior to the submit date.
- o If you specify neither parameter, then all historical observations are selected when Update type is INITIALIZE. When Update type is UPDATE, the default is 1 for Number of days/weeks/months, based on the CA MICS calendar.

Execution mode: The specification for executing the meta file update. Specify BATCH to submit the meta file update as a batch job. Specify FOREGROUND to execute it immediately as an online SAS program.

List updated CAP DB file contents: The specification to execute a PROC CONTENTS and a PROC PRINT of the meta file to be executed after the database is updated. To do so, specify a Y (YES).

The CA MICS Database Inquiry Execution screen shows the execution parameters that are available before the database update job is submitted for processing. This screen and its control parameters and options are discussed in detail in Section 1.5.3.5, Executing a File Definition Inquiry.

3.4.3 Meta File Case Studies

The Meta File case studies illustrate the use of Capacity Planning meta files.

The case studies presented in the sections that follow are:

- 1 - Combining Standard Applications
- 2 - Multi-Level Application Summarization

3.4.3.1 Combining Standard Applications

The combining standard applications case study illustrates using the CA MICS Capacity Planner to create a meta file that combines the VAX/VMS Standard Application with the Tandem Standard Application into a distributed systems tracking file.

The following topics are addressed:

- 1 - Problem Description
- 2 - Control Parameters
- 3 - Output Report

3.4.3.1.1 Problem Description

For this case study, a meta file was created to track the activity of two distributed systems, one a Tandem system used for ATM processing, the other, a VAX cluster used for a stock trading system.

The systems used in this study are as follows:

SYSID	Description
=====	=====
TDM1	Tandem System Supporting ATM Activity
STK1	Stock Trading System Running on a VAX 8350

This case study addresses the following areas:

- o Combining SYSIDs
- o How to use the Meta File Facility to combine two standard application files.

3.4.3.1.2 Control Parameters

Figures 3-109 through 3-115 show the control screens that were used to create and execute this meta file definition example.

```
/----- Meta File Definition -----\  
|Command ==>  
|Enter a ? in any data entry field for more information on valid values.  
|Composing CAP DB File def.: DST - Distributed System Ex.  
|  
|Summarization Level  
|          Primary  ==> WEEKS    (DETAIL/DAYS/WEEKS/MONTHS)  
|          Secondary ==> ZONE     (ZONE/HOUR/ENDTS/DAYNAME)  
|  
|Capacity Planner Data Base ==> PRIVATE (Private/Shared)  
|  
|Amount of data retained  ==> 60    (1-9999 weeks)  
|Age of oldest forecast   ==> 24    (1-9999 weeks)  
|History cycles           ==> 3     (1-99 cycles)  
|Forecast cycles          ==> 3     (1-99 cycles)  
|  
|Enable:  
|  Notes report           ==> YES    (YES/NO)  
|  Contents reports       ==> YES    (YES/NO)  
|  
|Specify extended options ==> N      (Y/N/R)  
\-----/
```

Figure 3-109. Specifying Parameters for the Meta File

c. Set the primary summarization level to WEEKS

- o Set the primary summarization level to WEEKS
- o Set the secondary summarization level to ZONE
- o Indicate that the input standard applications are in the private Capacity Planner database
- o Indicate that extended options are not to be included in the meta file

```
/----- Meta File -----  
|Option ==> 2  
|  
|Composing CAP DB file def.: DST - Distributed System Ex.  
|  
|    1 - Insert Common Data Selection  
|        Specify data selection criteria to apply to ALL files.  
|    2 - Insert CAP File and Data Element Selection  
|        Select a Capacity Planner file and data elements.  
|    3 - Insert Cross-File Data Element Definition  
|        Specify cross-file data element derivations.
```

Figure 3-110. Selecting CAP Files and Data Elements

Option 2, Insert CAP File and Data Element Selection, is chosen on the Meta File screen shown in Figure 3-110.

The CAP File and Data Element Selection screens shown in Figure 3-111 specify that the CAP File VX1 is being used. The data elements to be included in the meta file are listed. Note that since there will no summarization of this file, you do not have to specify a processing exit for derived data elements.

```

/----- Capacity Planner File and Data Element Selection -----\
|Command ==>                               Scroll ==> CSR |
|Enter a ? in any data entry field for more information on valid values |
|
|Composing CAP DB File def.: DST - Distributed System Ex. |
|
|CAP file ==> VX1 (fff) - VAX/VMS SA - Stock Quote System |
|
|                                     -----Line Commands----- |
|Include CAPAPUs ==> N (Y/N)          | I Insert C Copy D Delete | |
|Specify file exits ==> N (Y/N)       | R Repeat E Calculations | |
|                                     | S Depend Clauses & Label | |
|                                     ----- |
|                                     Sum Data |
|Cmd  Name      Label (long name)          D Type Type Order Len Output Format |
|-----|-----|-----|-----|-----|-----|-----|-----|
|_  ADJCOMTM Adjusted Compatibility Mode Time | COMP  T 120  13  TIME14.2 |
|_  ADJCPUTM Adjusted Total CPU Time          | COMP  T 120  13  TIME14.2 |
|_  ADJEXETM Adjusted Executive Mode Time     | COMP  T 120  13  TIME14.2 |
|_  ADJINTTM Adjusted Interrupt Stack Time    | COMP  T 120  13  TIME14.2 |
|_  ADJKERTM Adjusted Kernel Mode Time        | COMP  T 120  13  TIME14.2 |
|_  ADJMPSTM Adjusted Multi-Processor Sync Time | COMP  T 120  13  TIME14.2 |
|_  ADJPCCOM Adjusted Compatibility Mode Percent | COMP  N 120  5  MAPCT. |
|_  ADJPCCPU Adjusted Total CPU Percent       | COMP  N 120  5  MAPCT. |
|-----|-----|-----|-----|-----|-----|-----|
\-----/

```

Figure 3-111. CAP File and Data Element Selection Screen--Screen 1 (Part 1 of 7)

```

/----- Capacity Planner File and Data Element Selection -----\
|Command ==>                               Scroll ==> CSR |
|Enter a ? in any data entry field for more information on valid values |
|
|Composing CAP DB File def.: DST - Distributed System Ex. |
|
|CAP file ==> VX1 (fff) - VAX/VMS SA - Stock Quote System |
|
|                                     -----Line Commands----- |
|Include CAPAPUs ==> N (Y/N)          | I Insert C Copy D Delete |
|Specify file exits ==> N (Y/N)       | R Repeat E Calculations |
|                                     | S Depend Clauses & Label |
|                                     ----- |
|                                     Sum Data |
|Cmd Name Label (long name) D Type Type Order Len Output Format |
|-----|-----|-----|-----|-----|-----|-----|
|_ ADJPCXE Adjusted Executive Mode Percent |
|                                     COMP N 120 5 MAPCT. |
|_ ADJPCINT Adjusted Interrupt Stack Percent |
|                                     COMP N 120 5 MAPCT. |
|_ ADJPCKEK Adjusted Kernel Mode Percent |
|                                     COMP N 120 5 MAPCT. |
|_ ADJPCMPS Adjusted Multi-Processor Sync Percent |
|                                     COMP N 120 5 MAPCT. |
|_ ADJPCSUP Adjusted Supervisor Mode Percent |
|                                     COMP N 120 5 MAPCT. |
|_ ADJPCUSR Adjusted User Mode Percent |
|                                     COMP N 120 5 MAPCT. |
|_ ADJSUPTM Adjusted Supervisor Mode Time |
|                                     COMP T 120 13 TIME14.2 |
|_ ADJUSRTM Adjusted User Mode Time |
|                                     COMP T 120 13 TIME14.2 |
|_ BASECPU Base CPU |
|                                     RET A 000 10 ----- |
|_ DATE Interval End Date |
|                                     RET D 000 7 DATE7. |
|_ DURATION Recording Interval Time (MODES sum) |
|                                     ACCM T 050 13 TIME15.2 |
\-----/

```

Figure 3-111. CAP File and Data Element Selection Screen--Screen 2 (Part 2 of 7)

```

/----- Capacity Planner File and Data Element Selection -----\
|Command ==>                               Scroll ==> CSR |
|Enter a ? in any data entry field for more information on valid values |
|
|Composing CAP DB File def.: DST - Distributed System Ex. |
|
|CAP file ==> VX1 (fff) - VAX/VMS SA - Stock Quote System |
|
|                                     -----Line Commands----- |
|Include CAPAPUs ==> N (Y/N)          | I Insert C Copy D Delete | |
|Specify file exits ==> N (Y/N)       | R Repeat E Calculations | |
|                                     | S Depend Clauses & Label | |
|                                     ----- |
|                                     Sum Data |
|Cmd  Name      Label (long name)          D Type Type Order Len Output Format |
|-----|-----|-----|-----|-----|-----|-----|
|_ INTERVLS Number of Recording Intervals (MODES) |
|          ACCM  N  050    8  ----- |
|_ MODCOMTM Compatibility Mode Time |
|          ACCM  T  050   13  TIME14.2 |
|_ MODCPU TM Total CPU Time |
|          ACCM  T  050   13  TIME14.2 |
|_ MODEXETM Executive Mode Time |
|          ACCM  T  050   13  TIME14.2 |
|_ MODIDLTM Idle Time |
|          ACCM  T  050   13  TIME14.2 |
|_ MODINTTM Interrupt Stack Time |
|          ACCM  T  050   13  TIME14.2 |
|_ MODKERTM Kernel Mode Time |
|          ACCM  T  050   13  TIME14.2 |
|_ MODMPSTM Multi-Processor Sync Time |
|          ACCM  T  050   13  TIME14.2 |
|_ MODPCCOM Compatibility Mode Percent |
|          COMP  N  050    5  MAPCT. |
|_ MODPCCPU Total CPU Percent |
|          COMP  N  050    5  MAPCT. |
|_ MODPCEXE Executive Mode Percent |
|          COMP  N  050    5  MAPCT. |
|-----|-----|-----|-----|-----|-----|
\-----/

```

Figure 3-111. CAP File and Data Element Selection Screen--Screen 3 (Part 3 of 7)

```

/----- Capacity Planner File and Data Element Selection -----\
|Command ==>                               Scroll ==> CSR |
|Enter a ? in any data entry field for more information on valid values |
|Composing CAP DB File def.: DST - Distributed System Ex. |
|CAP file ==> VX1 (fff) - VAX/VMS SA - Stock Quote System |
|
|                                     -----Line Commands-----|
|Include CAPAPUs ==> N (Y/N)         | I Insert C Copy D Delete |
|Specify file exits ==> N (Y/N)      | R Repeat E Calculations |
|                                     | S Depend Clauses & Label |
|                                     -----|
|                                     Sum Data |
|Cmd  Name      Label (long name)          D Type Type Order Len Output Format|
|-----|-----|-----|-----|-----|-----|-----|
|_ MODPCIDL Idle Percent                    COMP N 050 5 MAPCT. |
|_ MODPCINT Interrupt Stack Percent         COMP N 050 5 MAPCT. |
|_ MODPCKER Kernel Mode Percent             COMP N 050 5 MAPCT. |
|_ MODPCMPS Multi-Processor Sync Percent    COMP N 050 5 MAPCT. |
|_ MODPCSUP Supervisor Mode Percent         COMP N 050 5 MAPCT. |
|_ MODPCUSR User Mode Percent               COMP N 050 5 MAPCT. |
|_ MODSUPTM Supervisor Mode Time            ACCM T 050 13 TIME14.2 |
|_ MODUSRTM User Mode Time                 ACCM T 050 13 TIME14.2 |
|-----|-----|-----|-----|-----|-----|
\-----/

```

Figure 3-111. CAP File and Data Element Selection Screen--Screen 4 (Part 4 of 7)

```

/----- Capacity Planner File and Data Element Selection -----\
|Command ==>                               Scroll ==> CSR |
|Enter a ? in any data entry field for more information on valid values |
|Composing CAP DB File def.: DST - Distributed System Ex. |
|CAP file ==> VX1 (fff) - VAX/VMS SA - Stock Quote System |
|
|                                     -----Line Commands-----|
|Include CAPAPUs ==> N (Y/N)         | I Insert C Copy D Delete |
|Specify file exits ==> N (Y/N)      | R Repeat E Calculations |
|                                     | S Depend Clauses & Label |
|                                     -----|
|                                     Sum Data |
|Cmd  Name      Label (long name)          D Type Type Order Len Output Format|
|-----|-----|-----|-----|-----|-----|-----|
|_ MODVECTM CPU Time of Vector Consumers      ACCM T   050   13  TIME14.2 |
|_ SPRBIOS Buffered I/Os                     ACCM N   050   13  |
|_ SPRCPUNI Number of Instructions Executed   ACCM N   050   13  |
|_ SPRCPUTM System CPU Busy Time             ACCM T   050   13  TIME14.2 |
|_ SPRDIOS Direct I/Os                      ACCM N   050   13  |
|_ SPRFPGFL Free List Page Faults           ACCM N   050   13  |
|_ SPRINSWP Inswaps                         ACCM N   050   13  |
|_ SPRIOVDR IOV IO Record Duration (Type 4)  ACCM T   050   13  TIME15.2 |
|_ SPRMPGFL Modified List Page Faults       ACCM N   050   13  |
|_ SPRPGEDR PGE PAGE Record Duration (Type 3) ACCM T   050   13  TIME15.2 |
|-----|-----|-----|-----|-----|-----|-----|
\-----/

```

Figure 3-111. CAP File and Data Element Selection Screen--Screen 5 (Part 5 of 7)

```

/----- Capacity Planner File and Data Element Selection -----\
|Command ==>                               Scroll ==> CSR |
|Enter a ? in any data entry field for more information on valid values |
|Composing CAP DB File def.: DST - Distributed System Ex. |
|CAP file ==> VX1 (fff) - VAX/VMS SA - Stock Quote System |
|
|-----Line Commands-----|
|Include CAPAPUs ==> N (Y/N) | I Insert C Copy D Delete |
|Specify file exits ==> N (Y/N) | R Repeat E Calculations |
|                               | S Depend Clauses & Label |
|-----|
|                               Sum Data |
|Cmd Name Label (long name) D Type Type Order Len Output Format |
|-----|
|_ SPRPGFLT Page Faults ACCM N 050 13 |
|_ SPRPSBIO Buffered I/O Rate COMP N 050 13 |
|_ SPRPSDIO Direct I/O Rate COMP N 050 13 |
|_ SPRPSFLT Page Fault Rate COMP N 050 13 |
|_ SPRPSFPL Free List Page Fault Rate COMP N 050 13 |
|_ SPRPSINS Inswap Rate COMP N 050 13 |
|_ SPRPSMPL Modified List Page Fault Rate COMP N 050 13 |
|_ SPRPSRIO Read I/O Rate COMP N 050 13 |
|-----|
\-----/

```

Figure 3-111. CAP File and Data Element Selection Screen--Screen 6 (Part 6 of 7)

```

/----- Capacity Planner File and Data Element Selection -----\
|Command ==>                               Scroll ==> CSR |
|Enter a ? in any data entry field for more information on valid values |
|Composing CAP DB File def.: DST - Distributed System Ex. |
|CAP file ==> VX1 (fff) - VAX/VMS SA - Stock Quote System |
|
|-----Line Commands-----|
|Include CAPAPUs ==> N (Y/N) | I Insert C Copy D Delete |
|Specify file exits ==> N (Y/N) | R Repeat E Calculations |
|                               | S Depend Clauses & Label |
|-----|
|Cmd  Name      Label (long name)      Sum  Data
|D Type Type Order Len Output Format|
|-----|
|_ SPRPSSYS System Page Fault Rate      COMP  N   050   13
|                                     ACCM  N   050   13
|_ SPRRDIOS Page Read I/Os              ACCM  N   050   13
|_ SPRSYSFL System Page Faults           ACCM  N   050   13
|_ SPRSYTCT SYT SYSTEM Record Count (Type 17) ACCM  N   050    8
|_ SPRSYTDR SYT SYSTEM Record Duration (Type 17) ACCM  T   050   13  TIME15.2
|_ SYSID      System Identifier           RET   A   050    4
|_ USRPCTOT User Mode as Percent of Total CPU COMP  N   120    5  MAPCT.
|_ VUPS       VAX Units of Power          RET   N   050   13
|_ VUPSA      VAX Units of Power Adjusted RET   N   120   13
|_ WEEK       Week of Year                RET   N   000    2
|_ YEAR       Year of Century              RET   N   000    2
|_ ZONE       Time Zone                   RET   A   000    1
|***** BOTTOM OF DATA *****|
\-----/

```

Figure 3-111. CAP File and Data Element Selection Screen--Screen 7 (Part 7 of 7)

The CAP File and Data Element Selection screens shown in Figure 3-112 specify that the CAP File TD1 is being used. The data elements to be included in the meta file are listed.

```

/----- Capacity Planner File and Data Element Selection -----\
|Command ==>                               Scroll ==> CSR |
|Enter a ? in any data entry field for more information on valid values |
|Composing CAP DB File def.: DST - Distributed System Ex. |
|CAP file ==> TD1 (fff) - Tandem SA - ATM System |
|
|-----Line Commands-----|
|Include CAPAPUs ==> N (Y/N) | I Insert C Copy D Delete |
|Specify file exits ==> N (Y/N) | R Repeat E Calculations |
|                               | S Depend Clauses & Label |
|-----|
|Cmd  Name      Label (long name)      Sum  Data |
|-----|-----|-----|-----|-----|-----|-----|
|_ TDMCPUNO CPU Number                  D RET  N    050  3 |
|_ CAPAPU   CPU Number                  RET  A    000  12 |
|_ CPECAHIT Cache Hits                  ACCM N    050  13 |
|_ CPECBSTM CPU Busy Time               ACCM T    050  13  TIME14.2 |
|_ CPEDIOS  Disk IOs                   ACCM N    050  13 |
|_ CPEDISP  Process Dispatches          ACCM N    050  13 |
|_ CPEIBSTM Interrupt Busy Time         ACCM T    050  13  TIME14.2 |
|_ CPEOVHTM Process Overhead Time       ACCM T    050  13  TIME14.2 |
|-----|

```

Figure 3-112. CAP File and Data Element Selection Screen--Screen 1 (Part 1 of 3)

```

/----- Capacity Planner File and Data Element Selection -----\
|Command ==>                               Scroll ==> CSR |
|Enter a ? in any data entry field for more information on valid values |
|
|Composing CAP DB File def.: DST - Distributed System Ex. |
|
|CAP file ==> TD1 (fff) - Tandem SA - ATM System |
|
|-----Line Commands-----|
|Include CAPAPUs ==> N (Y/N) | I Insert C Copy D Delete |
|Specify file exits ==> N (Y/N) | R Repeat E Calculations |
| | S Depend Clauses & Label |
|-----|
|
|          Sum Data
|Cmd  Name      Label (long name)      D Type Type Order Len Output Format
|-----|
|_ CPEPCBSY Percent CPU Busy              COMP N    050  5  MAPCT.
|_ CPEPCIBS Percent CPU Interrupt Busy      COMP N    050  5  MAPCT.
|_ CPEPCOBS Percent CPU Process Overhead Busy COMP N    050  5  MAPCT.
|_ CPEPCSBS Percent CPU Send Busy           COMP N    050  5  MAPCT.
|_ CPEPSCAH Cache Hits Per Second          COMP N    050  5
|_ CPEPSDIO Disk IO Per Second             COMP N    050  5
|_ CPEPSDSP Dispatches per Second          COMP N    050  5
|_ CPEPSSWP Swap Rate Per Second           COMP N    050  5
|
|***** BOTTOM OF DATA *****|
\-----/

```

Figure 3-112. CAP File and Data Element Selection Screen--Screen 2 (Part 2 of 3)

```

/----- Capacity Planner File and Data Element Selection -----\
|Command ==>                               Scroll ==> CSR |
|Enter a ? in any data entry field for more information on valid values |
|Composing CAP DB File def.: DST - Distributed System Ex. |
|CAP file ==> TD1 (fff) - Tandem SA - ATM System |
|
|-----Line Commands-----|
|Include CAPAPUs ==> N (Y/N) | I Insert C Copy D Delete |
|Specify file exits ==> N (Y/N) | R Repeat E Calculations |
|                               | S Depend Clauses & Label |
|-----|
|                               Sum Data |
|Cmd Name Label (long name) D Type Type Order Len Output Format |
|-----|
|_ CPESBSTM Send Busy Time ACCM T 050 13 TIME14.2 |
|_ CPESWAPS CPU Swaps ACCM N 050 13 |
|_ CPETYPE CPU Type RET A 050 8 |
|_ CPUCOUNT CPU Count ACCM N 050 2 |
|_ DATE Interval End Date RET D 000 7 DATE7. |
|_ DURATION Recording Interval Time ACCM T 050 13 TIME15.2 |
|_ ZONE Zone RET N 000 2 |
|_ SYSID System Identifier RET A 000 4 |
|_ WEEK Week of Year RET N 000 2 |
|_ YEAR Year of Century RET N 000 2 |
|***** BOTTOM OF DATA *****|
\-----/

```

Figure 3-112. CAP File and Data Element Selection Screen--Screen 3 (Part 3 of 3)

```
/----- Meta File -----Inquiry Validated.\
|Command ==>                               Scroll ==> CSR |
|
|Composing CAP DB File def.: DST - Distributed System Ex.  |
|
|Update the capacity planning data base file ==> Y (Y/N)    |
|
|Line Cmds: S Select  I Insert  D Delete  R Repeat         |
|
|Cmd              Data Base File Definition Step          |
| - -----
| - Meta File Definition
| - CAP File & Data Element Sel.: VX1 - VAX/VMS SA - Stock Quote System
| - CAP File & Data Element Sel.: TD1 - Tandem SA - ATM System
| - Cross-File Data Element Definition
|***** BOTTOM OF DATA *****
|
|
|
|
|
|
|
|
|
|
|-----/
```

Figure 3-113. Validating the Inquiry

After the meta file definition parameters are specified, the meta file is used to validate the inquiry. Figure 3-113 shows this screen after VALIDATE is typed at the Command prompt.

```
/----- Update Capacity Planning File ----- Confirm Save.\
|Command ==>
|
|Modifying CAP DB File def.: DST - Distributed System Ex.
|
|
|This file definition has been changed and must be saved before updating the
|file.
|
|
|Do you wish to save the file definition and update the file? ==> YES (YES/NO)
|
|
|
|Instructions:
|
|    Press ENTER key to save the file definition and update the file.
|
|    Reply NO and press enter key to cancel the file update request.
|
|
|-----/
```

Figure 3-114. Saving the Inquiry

The Update Capacity Planning File screen shown in Figure 3-114 is used to save the inquiry so that it can be executed.

```
/----- Update Capacity Planning File ----- DST Saved.-\
|Command ==>|
|
|Composing CAP DB File def.: DST - Distributed System Ex.|
|
|Update type ==> INITIALIZE (INITIALIZE/UPDATE)|
|Replace existing ==> N (Y/N)|
|
|PRIVATE Capacity Planner data base cycle(s) ==> 01 - __|
|Update start date ==> _____ (ddmonyy)|
|Number of weeks ==> 5__ (1-9999)|
|
|Execution mode ==> BATCH (BATCH/FOREGROUND)|
|
|List updated CAP DB file contents ==> YES (YES/NO)|
|
|
|
|
|Press END to invoke MICF inquiry execution to update the meta file.|
|Enter CANCEL to terminate the meta file update.|
\-----/
```

Figure 3-115. Executing the Inquiry

Cycle 1 is specified as input to the database loading process on the Update Capacity Planning File screen shown in Figure 3-115. Since we are doing an INITIALIZE all data in the PRIVATE Capacity Planning database for these two standard applications will be copied to our new meta file. The updated meta file will be printed after the update is complete.

3.4.3.1.3 Output Report

Figure 3-116 shows the output that is created by the PROC PRINT of the updated data set.

M I C S C A P A C I T Y P L A N N E R P R O D U C T									
OBS	ZONE	SYSID	WEEK	YEAR	INTERVLS	DURATION	MODCOMTM	MODEXETM	MODIDLTM
1	1	DST1	24	91	240	60:06:15.56	0:00:00.00	0:23:02.77	52:22:32.05
2	1	DST1	25	91	360	90:06:15.82	0:00:00.00	0:03:58.46	89:18:05.41
3	1	DST1	26	91	360	90:06:35.40	0:00:00.00	0:02:54.75	89:17:53.64
4	2	DST1	24	91	416	104:08:28.62	0:00:00.00	0:11:21.97	99:56:42.70
5	2	DST1	25	91	600	150:10:01.50	0:00:00.00	0:05:07.69	149:05:14.73
6	2	DST1	26	91	598	149:40:09.96	0:00:00.00	0:04:02.55	148:27:32.36
7	2	DST1	27	91	2	0:30:01.70	0:00:00.00	0:00:18.73	0:26:50.88
8	3	DST1	24	91	190	47:32:42.96	0:00:00.00	0:00:41.46	47:21:25.12
9	3	DST1	25	91	384	96:05:40.28	0:00:00.00	0:01:36.01	95:39:40.21
MODINTTM	MODKERTM	MODMPSTM	MODSUPTM	MODUSRTM	MODVECTM				
0:54:40.71	3:48:33.88	0:15:02.08	0:06:19.10	2:16:04.96	0:00:00.00				
0:05:59.19	0:26:52.85	0:00:13.60	0:01:57.88	0:09:08.43	0:00:00.00				
0:07:11.56	0:27:49.06	0:00:20.06	0:00:18.59	0:10:07.74	0:00:00.00				
0:31:13.75	2:03:51.33	0:07:30.48	0:01:50.44	1:15:57.88	0:00:00.00				
0:07:22.48	0:36:04.73	0:00:08.14	0:03:28.75	0:12:34.85	0:00:00.00				
0:05:58.46	0:45:44.24	0:00:35.04	0:01:50.55	0:14:26.70	0:00:00.00				
0:00:08.58	0:01:07.94	0:00:00.19	0:00:20.59	0:01:14.79	0:00:00.00				
0:01:33.54	0:08:15.51	0:00:02.06	0:00:00.53	0:00:44.71	0:00:00.00				
0:03:21.33	0:17:41.91	0:00:04.66	0:00:22.10	0:02:54.03	0:00:00.00				
MODPCCOM	MODPCCPU	MODPCEXE	MODPCIDL	MODPCINT	MODPCKER	MODPCMPS	MODPCSUP	MODPCUSR	
0.00 %	12.85 %	0.63 %	87.14 %	1.51 %	6.33 %	0.41 %	0.17 %	3.77 %	
0.00 %	0.89 %	0.07 %	99.10 %	0.11 %	0.49 %	0.00 %	0.03 %	0.16 %	
0.00 %	0.90 %	0.05 %	99.09 %	0.13 %	0.51 %	0.00 %	0.00 %	0.18 %	
0.00 %	4.02 %	0.18 %	95.96 %	0.49 %	1.98 %	0.12 %	0.02 %	1.21 %	
0.00 %	0.71 %	0.05 %	99.27 %	0.08 %	0.40 %	0.00 %	0.03 %	0.13 %	
0.00 %	0.80 %	0.04 %	99.18 %	0.06 %	0.50 %	0.00 %	0.02 %	0.16 %	
0.00 %	10.59 %	1.03 %	89.40 %	0.47 %	3.77 %	0.01 %	1.14 %	4.15 %	
0.00 %	0.39 %	0.02 %	99.60 %	0.05 %	0.28 %	0.00 %	0.00 %	0.02 %	
0.00 %	0.45 %	0.02 %	99.54 %	0.05 %	0.30 %	0.00 %	0.00 %	0.05 %	

Figure 3-116. Meta File Contents Report (Part 1 of 3)

MODCPUTM	SPRSYTCT	SPRBIOS	SPRCPUNI	SPRCPUTM	SPRDIOS	SPRFGFL	SPRINSWP	SPRIOUDR
7:43:43.50	120	1211327	61211.7	7:43:43.50	27400	282873	0	30:03:07.78
0:48:10.41	180	72554	6358.9	0:48:10.41	25472	109592	0	45:03:07.91
0:48:41.76	180	97965	6427.9	0:48:41.76	20357	105844	0	45:03:17.70
4:11:45.85	208	659178	33232.9	4:11:45.85	30640	116817	0	52:04:14.31
1:04:46.64	300	85869	8550.6	1:04:46.64	36046	109779	0	75:05:00.75
1:12:37.54	299	81629	9586.6	1:12:37.54	34027	206687	0	74:50:04.98
0:03:10.82	1	2140	419.8	0:03:10.82	2958	2064	0	0:15:00.85
0:11:17.81	95	18829	1491.2	0:11:17.81	5063	25714	0	23:46:21.48
0:26:00.04	192	40497	3432.1	0:26:00.04	13289	54246	0	48:02:50.14
SPRMPGFL	SPRPGEDR	SPRPGFLT	SPRRDIOS	SPRSYSFL	SPRSYTDR	SPRPSBIO	SPRPSDIO	SPRPSFLT
209018	30:03:07.78	1360478	69048	41	30:03:07.78	11.1965	0.25326	12.5752
102132	45:03:07.91	320588	3975	6	45:03:07.91	0.4473	0.15705	1.9766
86172	45:03:17.70	295780	4066	9	45:03:17.70	0.6040	0.12551	1.8236
137667	52:04:14.31	687385	30054	17	52:04:14.31	3.5165	0.16345	3.6669
158687	75:05:00.75	389956	4123	7	75:05:00.75	0.3177	0.13336	1.4427
133803	74:50:04.98	586023	4694	3	74:50:04.98	0.3030	0.12630	2.1752
1921	0:15:00.85	9990	331	0	0:15:00.85	2.3755	3.28357	11.0895
46036	23:46:21.48	87601	0	0	23:46:21.48	0.2200	0.05916	1.0236
94412	48:02:50.14	188272	398	0	48:02:50.14	0.2341	0.07683	1.0885
SPRPSFPL	SPRPSINS	SPRPSMPL	SPRPSRIO	SPRPSYS	BASECPU	ADJCOMTM	ADJPCUTM	
2.61465	0	1.93199	0.638224	0.000378971	TESTING	0:00:00.00	3:24:02.34	
0.67571	0	0.62971	0.024509	0.000036994	TESTING	0:00:00.00	0:21:11.78	
0.65256	0	0.53128	0.025068	0.000055488	TESTING	0:00:00.00	0:21:25.57	
0.62318	0	0.73440	0.160327	0.000090689	TESTING	0:00:00.00	1:50:46.57	
0.40614	0	0.58708	0.015253	0.000025897	TESTING	0:00:00.00	0:28:30.12	
0.76720	0	0.49666	0.017424	0.000011136	TESTING	0:00:00.00	0:31:57.32	
2.29117	0	2.13243	0.367431	0.000000000	TESTING	0:00:00.00	0:01:23.96	
0.30046	0	0.53792	0.000000	0.000000000	TESTING	0:00:00.00	0:04:58.24	
0.31361	0	0.54583	0.002301	0.000000000	TESTING	0:00:00.00	0:11:26.42	
ADJEXETM	ADJINTTM	ADJKERTM	ADJMPSTM	ADJPCCOM	ADJPCCPU	ADJPCEXE	ADJPCINT	
0:10:08.42	0:24:03.51	1:40:34.11	0:06:36.92	0.00 %	5.65 %	0.28 %	0.66 %	
0:01:44.92	0:02:38.04	0:11:49.65	0:00:05.98	0.00 %	0.39 %	0.03 %	0.04 %	
0:01:16.89	0:03:09.89	0:12:14.39	0:00:08.83	0.00 %	0.39 %	0.02 %	0.05 %	
0:05:00.07	0:13:44.45	0:54:29.79	0:03:18.21	0.00 %	1.77 %	0.08 %	0.21 %	
0:02:15.38	0:03:14.69	0:15:52.48	0:00:03.58	0.00 %	0.31 %	0.02 %	0.03 %	
0:01:46.72	0:02:37.72	0:20:07.47	0:00:15.42	0.00 %	0.35 %	0.01 %	0.02 %	
0:00:08.24	0:00:03.78	0:00:29.89	0:00:00.08	0.00 %	4.66 %	0.45 %	0.20 %	
0:00:18.24	0:00:41.16	0:03:38.02	0:00:00.91	0.00 %	0.17 %	0.01 %	0.02 %	
0:00:42.24	0:01:28.59	0:07:47.24	0:00:02.05	0.00 %	0.19 %	0.01 %	0.02 %	

Figure 3-116. Meta File Contents Report (Part 2 of 3)

ADJPCCKR	ADJPCMP5	ADJPCSUP	ADJPCUSR	ADJSUPTM	ADJUSRTM	USRPCTOT	VUPS	VUPSA	DATE	
2.78 %	0.18 %	0.07 %	1.66 %	0:02:46.80	0:59:52.58	29.34 %	2.2	5	15JUN91	
0.21 %	0.00 %	0.01 %	0.07 %	0:00:51.87	0:04:01.31	18.97 %	2.2	5	22JUN91	
0.22 %	0.00 %	0.00 %	0.08 %	0:00:08.18	0:04:27.41	20.80 %	2.2	5	29JUN91	
0.87 %	0.05 %	0.01 %	0.53 %	0:00:48.59	0:33:25.47	30.17 %	2.2	5	15JUN91	
0.17 %	0.00 %	0.01 %	0.06 %	0:01:31.85	0:05:32.13	19.42 %	2.2	5	22JUN91	
0.22 %	0.00 %	0.00 %	0.07 %	0:00:48.64	0:06:21.35	19.88 %	2.2	5	29JUN91	
1.65 %	0.00 %	0.50 %	1.82 %	0:00:09.06	0:00:32.91	39.19 %	2.2	5	06JUL91	
0.12 %	0.00 %	0.00 %	0.01 %	0:00:00.23	0:00:19.67	6.59 %	2.2	5	15JUN91	
0.13 %	0.00 %	0.00 %	0.02 %	0:00:09.72	0:01:16.57	11.15 %	2.2	5	22JUN91	
CPETYPE	CPECBSTM	CPEIBSTM	CPEOVHTM	CPESBSTM	CPESWAPS	CPECAHIT	CPEDIOS	CPEDISP	CPEPCBSY	CPEPCIBS
TXP	4:24:27.20	0:45:34.09	0:00:06.30	0:04:21.72	2471	737887	685026	7652570	48.97 %	8.43 %
TXP	5:44:08.37	1:06:16.71	0:00:05.58	0:08:31.50	1504	978760	701681	10539838	38.28 %	7.37 %
TXP	3:29:59.26	0:47:24.84	0:00:04.11	0:15:07.79	1767	0	0	4666824	38.88 %	8.78 %
TXP	3:44:32.00	0:50:50.98	0:00:14.26	0:12:29.07	4743	0	0	5546717	24.97 %	5.65 %
TXP	4:10:52.66	0:53:38.76	0:00:11.03	0:13:04.64	4982	436988	568179	5845211	46.45 %	9.93 %
TXP	4:48:29.62	1:04:14.21	0:00:10.64	0:12:05.04	3305	1082554	1032595	7391847	32.08 %	7.14 %
TXP	3:30:01.63	0:45:58.43	0:00:01.67	0:08:06.85	243	173195	456849	5926072	38.89 %	8.51 %
TXP	2:59:54.19	0:39:15.37	0:00:03.64	0:08:11.61	541	200435	402456	4629368	20.01 %	4.36 %
TXP	3:08:39.80	0:37:32.29	0:00:01.44	0:03:48.30	7200	320532	359570	4753180	34.93 %	6.95 %
CPEPCOBS	CPEPCBS	CPEPSCAH	CPEPSDIO	CPEPSDSP	CPESSWP	CPUCOUNT				
0.01 %	0.80 %	22.7734	21.1406	236.188	0.07626	1				
0.01 %	0.94 %	18.1445	13.0088	195.402	0.02788	1				
0.01 %	2.80 %	0.0000	0.0000	144.035	0.05454	1				
0.02 %	1.38 %	0.0000	0.0000	102.832	0.08792	1				
0.03 %	2.42 %	13.4871	17.5352	180.406	0.15376	1				
0.01 %	1.34 %	20.0664	19.1406	137.039	0.06127	1				
0.00 %	1.50 %	5.3455	14.1001	182.902	0.00750	1				
0.00 %	0.91 %	3.7158	7.4612	85.824	0.01003	1				
0.00 %	0.70 %	9.8928	11.0977	146.699	0.22221	1				

Figure 3-116. Meta File Contents Report (Part 3 of 3)

3.4.3.2 Multilevel Application Summarization

The multilevel application summarization case study shows how to use the CA MICS Capacity Planner to create a meta file that takes information from the Workload Standard Application and summarizes it into broader categories (for example, Department).

The following topics are addressed:

- 1 - Problem Description
- 2 - Control Parameters
- 3 - Output Report

3.4.3.2.1 Problem Description

For this case study, a single meta file was created to summarize the activity from the workload standard application into workgroups.

The systems used in this study are as follows:

SYSID	Description
=====	=====
SYSA	Production System W/Claims & Insurance Adjustors

This case study addresses the following areas:

- o Recalculating Data Elements
- o The use of the CAPIAPU internal variable
- o How to use the Meta File Facility to summarize an input Capacity Planner file to a higher level of summarization

3.4.3.2.2 Control Parameters

Figures 3-117 through 3-123 show the control screens that were used to create and execute this meta file definition example.

```

/----- Meta File Definition -----\
|Command ==>|
|Enter a ? in any data entry field for more information on valid values.|
|Composing CAP DB File def.: MT1 - Multi-Level Meta Ex.|
|
|Summarization Level|
|      Primary  ==> WEEKS      (DETAIL/DAYS/WEEKS/MONTHS)|
|      Secondary ==> HOUR      (ZONE/HOUR/ENDTS/DAYNAME)|
|
|Capacity Planner Data Base ==> PRIVATE (Private/Shared)|
|
|Amount of data retained  ==> 60      (1-9999 weeks)|
|Age of oldest forecast   ==> 24      (1-9999 weeks)|
|History cycles           ==> 3       (1-99 cycles)|
|Forecast cycles          ==> 3       (1-99 cycles)|
|
|Enable:|
|  Notes report           ==> YES      (YES/NO)|
|  Contents reports       ==> YES      (YES/NO)|
|
|Specify extended options ==> Y        (Y/N/R)|
\-----/

```

Figure 3-117. Specifying Parameters for the Meta File

The specifications on the Meta File Definition screen shown in Figure 3-117 do the following:

- o Set the timespan to WEEKS.
- o Set the summary level to HOUR.
- o Indicate that the input standard applications are in the Private Capacity Planner database.
- o Indicate that the optional data element CAPAPU is to be included in the meta file using the extended options (shown in Figure 3-94).
- o Indicate extended options (shown in Figure 3-94) are to be included in the meta file.

```

/----- SAS System Statements -----\
|Command ==>                               Scroll ==> CSR |
| |                                           |
|Inquiry Step:  File & Data Element Sel.: CPU - CPU PROCESSOR ACTIVITY FILE |
| |                                           |
|SAS Statements for the CAPAURT Macro |
|-----|
|***** TOP OF DATA *****|
|==MSG> SPECIFY SAS STATEMENTS TO SET CAPAPU VALUES. |
|NOTE= %MACRO CAPAURT; <=== PLEASE DO NOT PLACE LINES BEFORE THIS LINE. |
|000001 IF CAPIAPU EQ 'TS01' THEN CAPAPU = 'CLAIMS'; |
|000002 ELSE IF CAPIAPU EQ 'TS02' THEN CAPAPU = 'ADJUST'; |
|000003 ELSE IF CAPIAPU EQ 'PROD1' THEN CAPAPU = 'CLAIMS'; |
|000004 ELSE IF CAPIAPU EQ 'PROD2' THEN CAPAPU = 'ADJUST'; |
|000005 ELSE IF CAPIAPU EQ 'SYSTEMS' THEN CAPAPU = 'SYSTEMS'; |
|000006 ELSE CAPAPU = 'OTHER'; |
|NOTE= %MEND CAPAURT; <=== PLEASE DO NOT PLACE LINES AFTER THIS LINE. |
|***** BOTTOM OF DATA *****|
| |                                           |
| |                                           |
| |                                           |
\-----/

```

Figure 3-118. Setting the CAPAPU

Option 2, Insert CAP File and Data Element Selection, is chosen on the Meta File screen shown in Figure 3-113.

The CAP File and Data Element Selection screens shown in Figure 3-120 specify that the CAP File WK1 is being used. The data elements to be included in the meta file are listed. Note that since you are doing summarization, you need to set an exit for those data elements that need to be recalculated.

```

/----- Capacity Planner File and Data Element Selection -----\
|Command ==>                               Scroll ==> CSR |
|Enter a ? in any data entry field for more information on valid values |
|
|Composing CAP DB File def.: MT1 - Multi-Level Meta Ex. |
|
|CAP file ==> WK1 (fff) - Workload SA - Prod. Workloads |
|
|-----Line Commands-----|
|Include CAPAPUs ==> Y (Y/N) | I Insert C Copy D Delete |
|Specify file exits ==> Y (Y/N) | R Repeat E Calculations |
|                               | S Depend Clauses & Label |
|-----|
|                               Sum Data |
|Cmd  Name      Label (long name)      D Type Type Order Len Output Format|
|-----|
|_ CAPAPU      Workload Group              RET  A   000  12 |
|_ CPUDUR      Recording Interval Time      ACCM T   120  13  TIME13.2 |
|_ CPUTOBTM    Total Time Processor Busy     ACCM T   050  13  TIME12.2 |
|_ CPUTOUTM    Total Processor Up Time       ACCM T   050  13  TIME12.2 |
|_ DATE        Interval End Date            RET  D   000  7   DATE7. |
|_ DESCRIPT   Workload Group Long Name      RET  A   120  40 |
|_ HOUR        Hour of Day                  RET  N   000  2 |
|_ MVSLEVEL    MVS Software Level           RET  A   050  8 |
|-----|
\-----/

```

Figure 3-120. CAP File and Data Element Selection Screen--Screen 1 (Part 1 of 5)

```

/----- Capacity Planner File and Data Element Selection -----\
|Command ==>                               Scroll ==> CSR |
|Enter a ? in any data entry field for more information on valid values |
|Composing CAP DB File def.: MT1 - Multi-Level Meta Ex. |
|CAP file ==> WK1 (fff) - Workload SA - Prod. Workloads |
|-----Line Commands-----|
|Include CAPAPUs ==> Y (Y/N) | I Insert C Copy D Delete |
|Specify file exits ==> Y (Y/N) | R Repeat E Calculations |
| | S Depend Clauses & Label |
|-----|
|Cmd  Name      Label (long name)      Sum  Data |
|D Type Type Order Len Output Format |
|-----|
|_ PAGPIN      Non-VIO Non-Swap Page-Ins |
|          ACCM N      050  13 |
|_ PCATRCPU    Percent of CPU Attributable to Workload |
|          COMP N      120  3  MAPCT. |
|_ PCT0TCPU    Percent Total CPU Time Consumed |
|          COMP N      120  3  MAPCT. |
|_ PCT0TSRB    Percent Total SRB Time Consumed |
|          COMP N      120  3  MAPCT. |
|_ PCT0TTCB    Percent Total TCB Time Consumed |
|          COMP N      120  3  MAPCT. |
|_ PCTT0TPG    Percent Total Pagein Rate Consumed |
|          COMP N      120  3  MAPCT. |
|_ PGAAVAFT    Average Real+Expanded Frames |
|          COMP N      050  10 |
|_ PGAAVETM    Avg Time of Ended Transactions |
|          COMP T      050  10  TIME9.3 |
|_ PGAAVSTM    Avg Transaction Service Time |
|          COMP T      050  10  TIME9.3 |
|-----|

```

Figure 3-120. CAP File and Data Element Selection Screen--Screen 2 (Part 2 of 5)

```

/----- Capacity Planner File and Data Element Selection -----\
|Command ==>                               Scroll ==> CSR |
|Enter a ? in any data entry field for more information on valid values |
|
|Composing CAP DB File def.: MT1 - Multi-Level Meta Ex. |
|
|CAP file ==> WK1 (fff) - Workload SA - Prod. Workloads |
|
|-----Line Commands-----|
|Include CAPAPUs ==> Y (Y/N) | I Insert C Copy D Delete |
|Specify file exits ==> Y (Y/N) | R Repeat E Calculations |
| | S Depend Clauses & Label |
|-----|
|
|Sum Data
|Cmd Name Label (long name) D Type Type Order Len Output Format|
|-----|
|_ PGAAVWTM Avg Transaction Wait for Service Time |
| | COMP T 050 10 TIME9.3 |
|_ PGACPUTM Total CPU Time |
| | ACCM T 050 13 TIME12.2 |
|_ PGAHIN Hiperspace Page-Ins |
| | ACCM N 050 13 |
|_ PGAHMS ESO Hiperspace Read Misses |
| | ACCM N 050 13 |
|_ PGAITS I/O Service Units |
| | ACCM N 050 13 |
|_ PGAMPL Multi-Programming Level |
| | RET N 050 5 |
|_ PGAMTS Main Storage Service Units |
| | ACCM N 050 13 |
|-----|

```

Figure 3-120. CAP File and Data Element Selection Screen--Screen 3 (Part 3 of 5)


```

/----- Capacity Planner File and Data Element Selection -----\
|Command ==>                                     Scroll ==> CSR |
|Enter a ? in any data entry field for more information on valid values |
|
|Composing CAP DB File def.: MT1 - Multi-Level Meta Ex.
|
|CAP file ==> WK1 (fff) - Workload SA - Prod. Workloads
|
|-----Line Commands-----|
|Include CAPAPUs ==> Y (Y/N) | I Insert C Copy D Delete |
|Specify file exits ==> Y (Y/N) | R Repeat E Calculations |
|                               | S Depend Clauses & Label |
|-----|
|                               Sum  Data
|Cmd  Name      Label (long name)  D Type Type Order Len Output Format|
|-----|
|_ PGAPIN      Number of Page-Ins
|                               ACCM N   050  13
|_ PGAPSHIN    Hiperspace Page-Ins per Second
|                               COMP N   050  5
|_ PGAPSHMS    ESO Hiperspace Read Misses per Second
|                               COMP N   050  5
|_ PGAPSINS    Interval Service Rate
|                               COMP N   050  5
|_ PGAPSPIN    Page-Ins per Second
|                               COMP N   050  5
|_ PGASER      Total Service Units
|                               ACCM N   050  13
|_ PGASPP      Swap Sequences
|                               ACCM N   050  13
|-----|

```

Figure 3-120. CAP File and Data Element Selection Screen--Screen 4 (Part 4 of 5)

```

/----- Capacity Planner File and Data Element Selection -----\
|Command ==>                               Scroll ==> CSR |
|Enter a ? in any data entry field for more information on valid values |
|
|Composing CAP DB File def.: MT1 - Multi-Level Meta Ex.
|
|CAP file ==> WK1 (fff) - Workload SA - Prod. Workloads
|
|
|-----Line Commands-----|
|Include CAPAPUs ==> Y (Y/N) | I Insert C Copy D Delete |
|Specify file exits ==> Y (Y/N) | R Repeat E Calculations |
|                               | S Depend Clauses & Label |
|-----|
|
|Cmd  Name      Label (long name)      Sum  Data
|D Type Type Order Len Output Format|
|-----|
|_ PGASRBTM SRB CPU Time
|                               ACCM T    050  13  TIME12.2
|_ PGATAT  Residency Time
|                               ACCM T    050  13  TIME9.3
|_ PGATCBTM TCB CPU Time
|                               ACCM T    050  13  TIME12.2
|_ PGATTX  Transactions Terminated
|                               ACCM N    050  13
|_ SYSID   System Identifier
|                               RET  A    050  4
|_ TOTSRBTM Total SRB Time Consumed This System
|                               ACCM T    120  13  TIME12.2
|_ TOTTCBTM Total TCB Time Consumed This System
|                               ACCM T    120  13  TIME12.2
|_ WEEK    Week of Year
|                               RET  N    000  2
|_ YEAR    Year of Century
|                               RET  N    000  2
|-----|

```

Figure 3-120. CAP File and Data Element Selection Screen--Screen 5 (Part 5 of 5)

```
/----- Meta File -----Inquiry Validated.\n|Command ==> Scroll ==> CSR |\n|\n|Composing CAP DB File def.: MT1 - Multi-Level Meta Ex.\n|\n|Update the capacity planning data base file ==> Y (Y/N)\n|\n|Line Cmds: S Select I Insert D Delete R Repeat\n|\n|Cmd Data Base File Definition Step\n|- -----\n|_ Meta File Definition\n|_ CAP File & Data Element Sel.: WK1 - Workload SA - Prod. Workloads\n|***** BOTTOM OF DATA *****\n|\n|\n|\n|\n|\n|\n|\n|\n|
```

Figure 3-121. Validating the Inquiry

After the meta file definition parameters are specified, the meta file is used to validate the inquiry. Figure 3-121 shows this screen after VALIDATE is typed at the Command prompt.

```
/----- Update Capacity Planning File ----- Confirm Save.\
|Command ==>
|
|Composing CAP DB File def.: MT1 - Multi-Level Meta Ex.
|
|
|This file definition has been changed and must be saved before updating the
|file.
|
|
|Do you wish to save the file definition and update the file? ==> YES (YES/NO)
|
|
|
|Instructions:
|
|    Press ENTER key to save the file definition and update the file.
|
|    Reply NO and press enter key to cancel the file update request.
|
|
|-----/
```

Figure 3-122. Saving the Inquiry

The Update Capacity Planning File screen shown in Figure 3-122 is used to save the inquiry so that it can be executed.

```

/----- Update Capacity Planning File ----- DST Saved.-\
|Command ==>|
|
|Composing CAP DB File def.: MT1 - Multilevel Meta Ex.|
|
|Update type                ==> INITIALIZE (INITIALIZE/UPDATE)|
|Replace existing           ==> N (Y/N)|
|
|PRIVATE Capacity Planner data base cycle(s) ==> 01 - __|
|
|Update start date          ==> _____ (ddmonyy)|
|Number of weeks            ==> _____ (1-9999)|
|
|Execution mode              ==> BATCH      (BATCH/BACKGROUND)|
|
|List updated CAP DB file contents ==> YES (YES/NO)|
|
|
|
|Press END to invoke MICF inquiry execution to update the meta file.
|Enter CANCEL to terminate the meta file update.
\-----/

```

Figure 3-123. Executing the Inquiry

Cycle 1 is specified as input to the database loading process on the Update Capacity Planning File screen shown in Figure 3-123. Since you are doing an INITIALIZE, all data in the PRIVATE Capacity Planning database for this standard application will be copied to the new meta file. The updated meta file will be printed after the update is complete.

3.4.3.2.3 Output Report

Figure 3-124 shows the output that is created by the PROC PRINT of the updated data set.

OBS	ZONE	SYSID	WEEK	YEAR	TOTSRBTM	TOTTCBTM	PAGPIN	MVSLEVEL	CPUTOBTM	CPUTOUTM	CPUDUR	PGACPUTM	PGASRBTM	
1	1	S008	22	98	3:59:35.36	27:30:38.31	906574	SP3.1.3	67:29:12.21	72:00:00.46	36:00:00.23	0:55:19.74	0:02:22.93	
2	1	S008	23	98	5:10:17.51	33:56:44.00	1019275	SP3.1.3	79:56:41.09	85:56:58.52	42:58:29.26	1:35:43.93	0:02:09.00	
3	1	S008	24	98	4:57:04.90	35:55:17.02	1240731	SP3.1.3	86:40:29.48	89:59:58.83	44:59:59.42	1:00:20.99	0:02:10.33	
4	1	S008	25	98	5:07:26.89	34:19:17.26	1011178	SP3.1.3	79:36:11.11	87:45:45.03	43:52:52.51	1:10:19.13	0:03:49.15	
5	1	S008	26	98	5:00:42.75	31:27:56.51	1032329	SP3.1.3	77:34:18.14	87:10:11.38	43:35:05.69	1:41:37.26	0:02:25.11	
6	1	S008	27	98	4:20:26.15	23:51:01.77	685754	SP3.1.3	54:49:03.94	68:55:44.56	34:27:52.28	0:40:42.06	0:02:08.89	
7	1	S008	28	98	5:26:01.29	38:24:19.89	804749	SP3.1.3	65:58:42.67	93:43:22.54	34:04:26.55	0:36:46.46	0:01:46.11	
8	1	S008	29	98	6:57:12.86	50:57:28.49	1242727	SP3.1.3	90:05:32.84	128:41:02.73	42:53:40.91	0:42:25.66	0:02:14.69	
PGATAT	PGATCBTM	PGAHIN	PGAHMS	PGAITS	PGAMTS	PGAPIN	PGASER	PGASPP	PGATTX	PGAMPL	PGAPSHIN	PGAPSHMS	PGAPSINS	PGAPSPIN
12:16:08.	0:52:56.81	0	0	1750522	55223196	35154	93633125	44620	43187	0.31660	0.00000	0	671.2	0.3
12:02:43.	1:33:34.93	0	0	3130124	69569898	26425	136129461	23333	22501	0.28067	0.00000	0	881.1	0.2
13:36:36.	0:58:10.67	0	0	2554919	39096333	31459	81637362	19068	18274	0.26604	0.00000	0	443.3	0.2
18:54:52.	1:06:29.98	0	0	2416498	80720366	46920	129728096	57476	53864	0.43361	0.00000	0	826.1	0.3
17:13:26.	1:39:12.15	5776	0	1982346	131697964	41152	201011453	34966	31419	0.40919	0.03812	0	1326.5	0.3
7:00:01.4	0:38:33.18	0	0	2183431	59773809	17443	87455660	12763	12363	0.23659	0.00000	0	805.7	0.2
11:33:07.	0:35:00.35	0	0	1307131	35971094	18012	58719809	17742	17719	0.33751	0.00000	0	476.5	0.1
7:13:15.2	0:40:10.98	385	0	1746471	44277734	24456	70615541	20919	21716	0.16478	0.00244	0	447.6	0.2
PGAAVAFT	PGAAVETM	PGAAVSTM	PGAAVTRN	PGAAVWTM	CAPAPU	DESCRIPT	PCTATRCPU	PCTOTTCPU	PCTOTSRB	PCTOTTCB	PCTTOTPG	DATE		
480.71	0:00:01.3	0:33:43.4	0.3317	0:00:00.0	ADJUST	ALL Adjustors	46.68 %	1.36 %	0.99 %	3.20 %	3.87 %	29MAY98		
453.99	0:00:02.5	0:49:08.8	0.2988	0:00:00.1	ADJUST	ALL Adjustors	48.93 %	1.99 %	0.69 %	4.59 %	2.59 %	05JUN98		
388.16	0:00:03.4	0:27:01.7	0.2733	0:00:00.0	ADJUST	ALL Adjustors	47.15 %	1.16 %	0.73 %	2.69 %	2.53 %	12JUN98		
445.13	0:00:01.8	0:28:21.1	0.4856	0:00:00.1	ADJUST	ALL Adjustors	49.55 %	1.47 %	1.24 %	3.22 %	4.64 %	19JUN98		
480.59	0:00:03.3	0:48:56.4	0.4517	0:00:00.3	ADJUST	ALL Adjustors	47.02 %	2.18 %	0.80 %	5.25 %	3.98 %	26JUN98		
514.21	0:00:03.5	0:51:19.1	0.2617	0:00:00.3	ADJUST	ALL Adjustors	51.42 %	1.23 %	0.82 %	2.69 %	2.54 %	03JUL98		
370.03	0:00:02.7	0:20:02.7	0.3962	0:00:00.4	ADJUST	ALL Adjustors	66.44 %	0.92 %	0.54 %	1.51 %	2.23 %	10JUL98		
388.88	0:00:01.6	0:42:22.2	0.1761	0:00:00.1	ADJUST	ALL Adjustors	64.28 %	0.78 %	0.53 %	1.31 %	1.96 %	24JUL98		

3-124. Meta File Contents Report

3.5 Business Element Files

Historically, one of the most difficult problems that the Capacity Planner encounters is the forecasting of the resource requirements of specific applications. A popular approach has been to use business element data (sometimes called natural forecast units or key volume indicators) to develop relationships between the volume of business activity in the sector using the application and the amount of resource consumption.

The CA MICS Capacity Planner supports both the collection and maintenance of business element data and the analysis of business activity and resource consumption relationships between these indicators.

This section describes how to use the CA MICS Capacity Planner to create user-defined business element files. You are responsible for defining the data elements that you wish to include in your database file and for executing the data base load job whenever appropriate.

The following sections describe the screens that control this facility:

- 1 - Usage Guidelines
- 2 - Screen Operation

3.5.1 Usage Guidelines

The term "business element" is used to represent a measurement unit that characterizes activity in any specified division of an organization. The objective of a business element forecast is to investigate possible statistical relationships between such measurements and the demand for computer resources.

The business element files represent business element values for your installation at either the DAYS, WEEKS or MONTHS timespan. Because interest in specific business elements varies from installation to installation, all but one of the data elements contained in a business element file must be user-defined.

The DATE data element is the exception. This data element must always be present in a business element file. It is the end date of the interval (that is, day, week, or month) represented by the record. The business element file database generation routines automatically calculate and maintains two additional data elements for you: YEAR and MNTHDAY/WEEK/MONTH (depending on the timespan that you define). You can define the remainder of the data elements in the business element file using the Business Element File Definition screen (see Figure 3-122).

As with the resource element files, you can define any number of business element files. You supply a three-character that identifies each business element file definition. This name must be unique.

The Capacity Planning database file contains two types of business element files. The first type contains observed values of historical data. The second type stores forecasts generated by the univariate or multivariate software. Each type of file is also defined using multiple cycles.

Cycle 01: Data from the last file load or update plus any data from the existing 01 cycle that is not older than allowed by the Maximum data available for analysis parameter on the Business Element File Definition screen shown in Figure 3-122. The CAPxxx01 File contains user-supplied observations, while the CPFxxx01 File contains forecasted observations from the workload forecasting routines. As with the resource element files, the CPFxxxnn File is not created until you generate and save a forecast for the first time.

Cycle 02: Data that is contained in the Cycle 01 File when it is aged. Aging occurs before the Cycle 01 File is loaded or updated. The Cycle 02 File acts as a backup copy of the Cycle 01 File as it existed before the last file update.

Cycle 03 A copy of the data from the Cycle xx-1 File just through before it is aged. This occurs each time the
Cycle nn: Cycle 01 File is updated.

The Cycle 01 File is the input file for the analysis and forecasting programs. It is created the first time you execute the file definition inquiry. This process reads the business elements that you define either from an externally created data set or from a SAS/FSP session.

Note that the historical files (CAP files) and the forecast files (CPF files) are not aged together. The CAP files are aged whenever you perform a file update. You can age the CPF files only when you request aging through a parameter on the workload forecast control screen.

In addition to the data elements defined for the business element files containing the observations that you provide (the CAPxxxxnn files), the forecast files (CPFxxxxnn files) contain three additional data elements:

- FILE_NO: The types of forecast data from the workload forecasting routines. If FILE_NO is 11, the data values are point estimates. If FILE_NO is 12, the data values are confidence limits.
- FORYEAR: The year value of the most recent historical observation used to derive the forecast. For example, if you specify to Univariate Model Forecasting that a forecast is to be generated based on data from June 1998 through October 1999, the value of FORYEAR would be 99. This data element containing this value would be present on each of the observations that are forecasted during that execution of Univariate Model Forecasting.
- FORMNDY/
FORWEEK/
FORMONTH: The day, week, or month value of the most recent historical observation used to derive the forecast. For example, if you specify to Univariate Model Forecasting that a monthly forecast is to be generated based on data from June 1998 through October 1999, the value of FORMONTH would be 10. This data element containing this value would be present on each of the observations that are forecasted during that execution of Univariate Model Forecasting.

Two methods are supported for entering your business element observations into the business element file. These methods are identified on the Business Element File Data Entry screen shown in Figure 3-125.

With the first method, you enter the observations directly into the business element file via a SAS/FSP session invoked for you by the business element file database generation routines. If you select this method, you first need to initialize the business element file by selecting Option 1 on the Business Element File Data Entry screen.

The initialization process places a number of entries into the file that contains values for the DATE and YEAR and MNTHDAY/WEEK/MONTH data elements, and leaves your remaining business data elements with missing values. You specify the number of initialized entries and the starting date for the first initialized entry. Once you initialize the file, you can enter data values as often as you want into the initialized entries by selecting Option 2 on the Business Element File Data Entry screen. Should you eventually need more entries than you first initialized, you can add them using the normal SAS/FSP commands that are available for this function.

The second method of entering your business element observations is to build an external file (not a SAS file) containing the values for the business elements. An easy method of doing this is to invoke the ISPF Editor and enter the data values directly into a sequential file or into a member of a partitioned data set. Then select Option 3 on the Business Element File Data Entry screen to read the data directly into your business element file from the input file you specify. If you use this method, you do not have to initialize the file.

3.5.2 Screen Operation

The following sections describe the online screens that the CA MICS Capacity Planner uses to define business element files:

- 1 - File Definition
- 2 - File Data Elements
- 3 - Inquiry Processing
- 4 - Business Element File Data Entry

3.5.2.1 File Definition

The first step in creating a business element file definition is to identify the timespan of the data that will be maintained in the file and to specify the data retention control parameters. You may specify these parameters on the Business Element File Definition screen shown in Figure 3-125.

```

/----- Business Element File Definition -----\
|Command ==>|
|Enter a ? in any data entry field for more information on valid values.|
|Composing CAP DB File def.: BEF - Business Element File|
|
|
|Time-span      ==> MONTHS (DAYS/WEEKS/MONTHS)|
|Starting date   ==> _____ (ddmmyy)|
|Number of months ==> ____ (for pre-formatting records)|
|
|
|Maximum amount of data available for analysis ==> 60 (1-9999 months)|
|Age of oldest forecast                        ==> 24 (1-9999 months)|
|Number of cycles retained:      History file ==> 3 (1-99 cycles)|
|                                Forecast file ==> 3 (1-99 cycles)|
|
|
|
|
|Specify business element file data elements ==> N (Y/N)|
\-----/

```

Figure 3-125. Business Element File Definition Screen

The Business Element File Definition screen shown in Figure 3-125 is the main control screen for defining business element files. It contains the following fields:

Timespan:	The timespan (either DAYS, WEEKS, or MONTHS) of the CA MICS file(s) that you wish to use as input for processing when the business element file is loaded. This timespan must match that of the resource element or standard application files against which business-related measures are analyzed.
Starting date:	An optional parameter that provides the starting date for the initialized entries in your business element file if file initialization is to be performed. Section 3.5.1 discusses the initialization process.
Number of months:	The number of initialized observations that are placed in your business element file if file initialization is performed.
Maximum amount of data available for analysis:	The number of months or weeks of historical data to be retained in the database. Any observations older than what you specify for this parameter are dropped from the file during update processing.
Age of oldest forecast:	The number of days, weeks, or months that you wish to retain a forecast after it is generated. Any forecast older than what you specify for this parameter is dropped whenever a new forecast is generated. The values of the FORYEAR and FORMNDY/FORWEEK/FORMONTH are used to determine the age of the forecast.
Number of history cycles retained:	The number of cycles or versions of the historical data that you wish to retain. A new cycle is created when the file is updated and the existing cycles are aged (that is, Cycle 01 becomes Cycle 02 and Cycle 02 becomes Cycle 03).
Number of forecast cycles retained:	The number of cycles or versions of the forecast data that you wish to retain. A new cycle is created when the file aging is specified on one of the workload forecasting routine control screens and the existing cycles are aged (that is, Cycle 01 becomes Cycle 02 and Cycle 02 becomes Cycle 03).

Specify business element data elements:	The specification for the business element file data elements. If you specify a Y (YES), the Business Element File Data Elements screen shown in Figure 3-126 is displayed.

3.5.2.2 File Data Elements

The second step in creating a business element file definition is specifying the business element that you wish to include in the file using the Business Element File Data Elements screen shown in Figure 3-126.

[illegible]

Figure 3-126. Business Element File Data Elements Screen

The parameters for the Business Element File Data Elements screen are defined below:

Name: The name of a business element. This may be any valid SAS name as described in the SAS manuals. The data elements YEAR, MONTH, WEEK, and MNTHDAY are exceptions. Do not specify them since these are automatically supplied for you based on the data value that you supply for the DATE data element. Choose variable names that are descriptive of the values they represent.

Data Element Label (long name): The data element label that is displayed on the data element selection screen and on many reports.

Tp: The type of data element. This is used to validate the data values and usage for this data element. Valid types are Alphanumeric (A), Numeric (N), Money (M), Time (T), Hex (H), Date (D), Time-of-day (TD), and Date-Time (DT).

Length: The length of the data element.

Output Format: The SAS format that is used to print the data element values.

3.5.2.3 Inquiry Processing

After you complete the previous steps, you may process the business element file definition inquiry that you specified using the Inquiry Processing screen shown in Figure 3-127.

```

/----- Inquiry Processing -----\
|Command ==>
|
|
|Active MICF Inquiry: BEF      - Business Element File
|           Author:          Charles Frost
|
|
|Specify Inquiry processing ==> EXECUTE (EXECUTE/MODIFY/SAVE/CANCEL)
|EXECUTE - Execute the inquiry.
|MODIFY  - Redisplay the inquiry for modification.
|SAVE    - Exit from inquiry composition and save the inquiry.
|CANCEL  - Exit from inquiry composition without saving the inquiry.
|
|
|Instructions:
|
|Specify desired option and press ENTER key.
|
\-----\

```

Figure 3-127. Inquiry Processing Screen

The Inquiry Processing screen displays the name and author of the business element file definition inquiry and allows you to specify the type of processing you desire. You may execute the inquiry immediately, redisplay it for modification (saved as defined), or cancel it.

3.5.2.4 Business Element File Data Entry

If you select the EXECUTE option on the Inquiry Processing screen, the Business Element File Data Entry screen shown in Figure 3-128 appears.

```

/----- Business Element File Data Entry -----\
|Option ==>                                     |
|Composing CAP DB File def.: BEF - Business Element File |
|
|      1 - Initialize Business Element File           |
|      Prime business element file with null entries in preparation |
|      for online data entry.                         |
|      2 - Online Data Entry                         |
|      Update the business element file using SAS/FSP. The file must |
|      have been previously initialized.               |
|      3 - Batch Data Entry                         |
|      Update the business element file using external card images.  |
|      The card image update can be executed in foreground or batch. |
|
|      Starting date ==> _____ (ddmonyy)      Number of months ==> ____ (1-9999) |
|      Card image DSN ==> _____              |
|
|-----\

```

Figure 3-128. Business Element File Data Entry Screen

The Business Element File Data Entry screen provides you with the options for initializing (only required if you intend to use the SAS/FSP method of data entry), and for loading business element data into your business element file.

Choosing Option 1, Initialize Business Element File, causes any previous data that you entered in your business element file to be discarded. A new file is created with initialized entries based on the following parameters:

Starting date:	The execution-time override for the Starting date parameter on the Business Element File Definition screen shown in Figure 3-125. Specify this parameter to define the starting date for file initialization.
Number of months:	The execution-time override for the Number of months parameter on the Business Element File Definition screen shown in Figure 3-125. Specify this parameter to initialize the number of days, weeks, or months in your business element file.

Choosing Option 2, Online Data Entry, provides direct, interactive access to your business element file for data entry. You can use the SAS Full Screen Product (SAS/FSP), which is a prerequisite, to implement this option. When you select this option, SAS/FSP is invoked in the foreground through MICF. To use this method, you must first initialize the file by selecting Option 1. If you use all of the previously initialized entries, DO NOT reinitialize the file unless you want to discard all of your previously entered data. Instead, use the SAS/FSP commands for adding observations to a SAS data set.

Choosing Option 3, Batch Data Entry, invokes a batch job that reads your business element observations from the data set that you identified in the following parameter:

Card image DSN:	The data set name of the card image. This parameter points either to a sequential data set or a member of a partitioned data set that contains business element data that you create external to the CA MICS Capacity Planner.

Code this parameter as you would any other data set name. If the data is a member of a partitioned data set, include the member name as part of this parameter. If the data set name is

cataloged under your userid, then do not enter your userid as part of the data set name. If you code a fully qualified data set name (that is, including a user ID), then enclose this data set name in single quotes.

Figure 3-129 shows a typical example of a SAS/FSP data entry screen for a business element file.

1...5...10...15...20...25...30...35...40...45...50...55...60...65...70...75...80		
1	/-----\	
2	Command ==>	Screen 1
3		Obs 1
4	-----	
5		
6	YEAR:	98
7		
8	MONTH:	1
9		
0	DATE:	31JAN98
1		
2	INVOICES:	-----
3		
4	UPDATES:	-----
5		
6	OPN_ITEM:	-----
7		
8		
9		
0		
1		
2		
3		
\		

Figure 3-129. Business Element File Data Entry Screen

Chapter 4: PRESENTING FORECASTED DATA

The Presentation Graphics software produces plots that depict the forecasts generated using Univariate Modeling, Multivariate Regression, and Business Element Forecasting, and the results of applying any worksheets studies, if appropriate. This facility is designed to address the following requirements:

- o High quality graphic presentations of the historical data observations and forecasts for output either to a line printer or a color graphics device
- o Tracking of previous forecasts to determine if the statistical models developed during the forecasting process are reliable

This section contains the following topics:

[4.1 Functional Description](#) (see page 525)

[4.2 Usage Guidelines](#) (see page 526)

[4.3 Standard Output](#) (see page 526)

[4.4 Presentation Technique Tutorial](#) (see page 529)

[4.5 Component Operation](#) (see page 531)

[4.6 Case Studies](#) (see page 540)

4.1 Functional Description

Presentation Graphics allows you to produce color or line printer graphics suitable for presenting analysis results and forecasts generated with Univariate Modeling, Multivariate Regression, and Business Element Forecasting, and the results of applying any worksheet studies, if appropriate. You can also produce plots of the historical observations contained in any of the resource element or business element capacity planning database files.

Optional parameters allow you to control colors and print fonts. You can also add capacity lines to the graph. Line printer graphics are always available; however, SAS/GRAPH software is a prerequisite for the color graphics option.

4.2 Usage Guidelines

Forecasting for capacity planning is typically a quarterly, semiannual, or annual activity. Regularly produced management reports track the accuracy of these forecasts against actual values. The reports display historical data, the most recent forecast, and a previous forecast, if it is available. For example, you have a capacity planning database file that is updated during a regular monthly execution and you routinely prepare three-month projections. You want to produce a quarterly report for management showing the results of the last forecast and next quarter's projection. You can produce this management report with Presentation Graphics.

4.3 Standard Output

Presentation Graphics produce plots that depict the forecasts generated using Univariate Modeling, Business Element Forecasting, Multivariate Regression, and worksheets. These plots show the historical data series, previous forecasts, current forecasts, and confidence limits on each forecasted point.

Figure 4-1 shows a sample printer graphics report for a TSO response study. The title line of the report lists the SYSID, ZONE, and report title that you specify on the Presentation Graphics screen.

The body of the report displays the plot of the specified data element. The Y-axis label shown in the plot is specified for the Y-axis label control parameter. The values on the Y-axis are scaled so that the numerical values shown are less than 1,000. If the Y-axis values are scaled by the program, then a scale block is generated at the base of the Y-axis.

The scaling of the Y-Axis is not performed for variables that represent time, such as the response variable, in this plot. Time variables are recognized through the CA MICS convention of ending the variable name with the characters TM. For example, PGACPUTM represents the amount of MVS TCB time accumulated for a performance group. These variables are carried as SAS time values in the CA MICS database in units of seconds.

Time variables are printed using a TIMEw.d format (HH:MM:SS.DDD) for hours, minutes, seconds, and a decimal portion of seconds. Scaling of time values presents a readability problem, since hours and minutes do not accumulate on a decimal basis. To aid in readability of time values, the precision of the output format is designed to float with the magnitude of the time value to be printed, from a maximum of HHHHH:MM:SS. down to H:MM:SS.DDDD. This allows printing of time values from 100 microseconds up to 10,000 hours with no loss of precision.

For printer plots, seven characters are used to denote the points in the body of the plot:

- * - historical observations
- + - current forecasted values
- W - worksheet applied forecasts
- P - previously forecasted values
- U - the upper 95 percentile confidence limit
- L - the lower 95 percentile confidence limit
- = - a user-specified limit
- | - line marking where history ends and future begins

A legend of these characters appears in the lower left corner of the report. The example shown in Figure 4-1 includes both a previous forecast and user-specified limits.

For color graphics plots, you can specify the symbols, line types, and colors to use to plot the points in the body of the plot. Figure 4-12 shows the Color Graphics Format Parameters control screen that provides you with these options and others.

The X-axis values correspond to the dates of the observations. Depending on the timespan of the historical observations and your CA MICS calendar, you will see one of three X-axis formats: days, weeks, or months. For printer or color graphics plots, all days, weeks, or months are shown on a single plot.

Presentation Graphics automatically selects the proper report format.

Presentation Graphics also enables you to specify up to six 60-character lines of information that are printed as footnotes on both the printer and the color graphics plots.

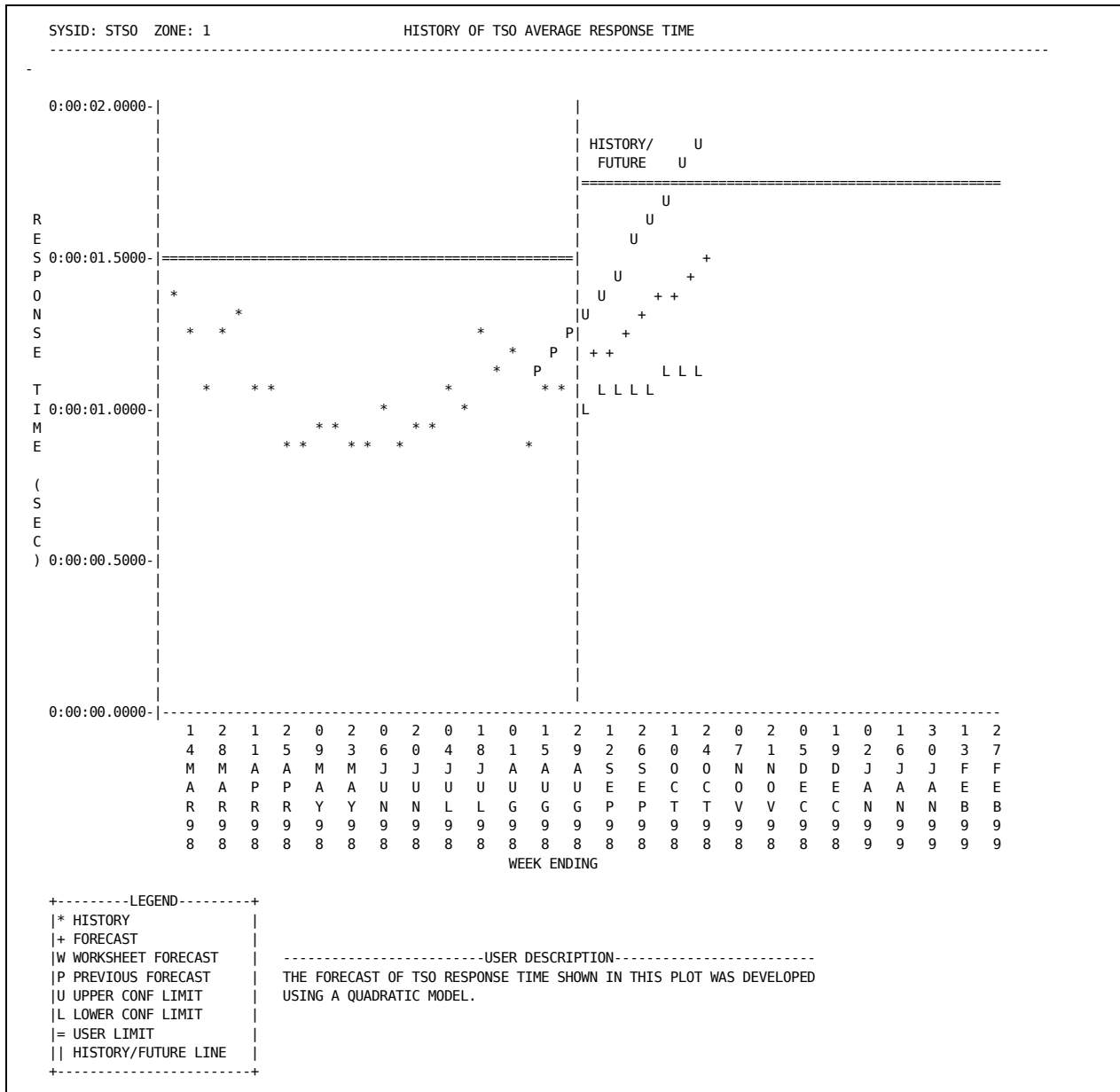


Figure 4-1. Presentation Graphics Report

4.4 Presentation Technique Tutorial

Presentation Graphics uses the historical observations and forecasts that reside in the Capacity Planning database files. (The creation and use of these data sets is discussed in Chapters 2 and 3.) Figure 4-2 shows an example of a set of these files.

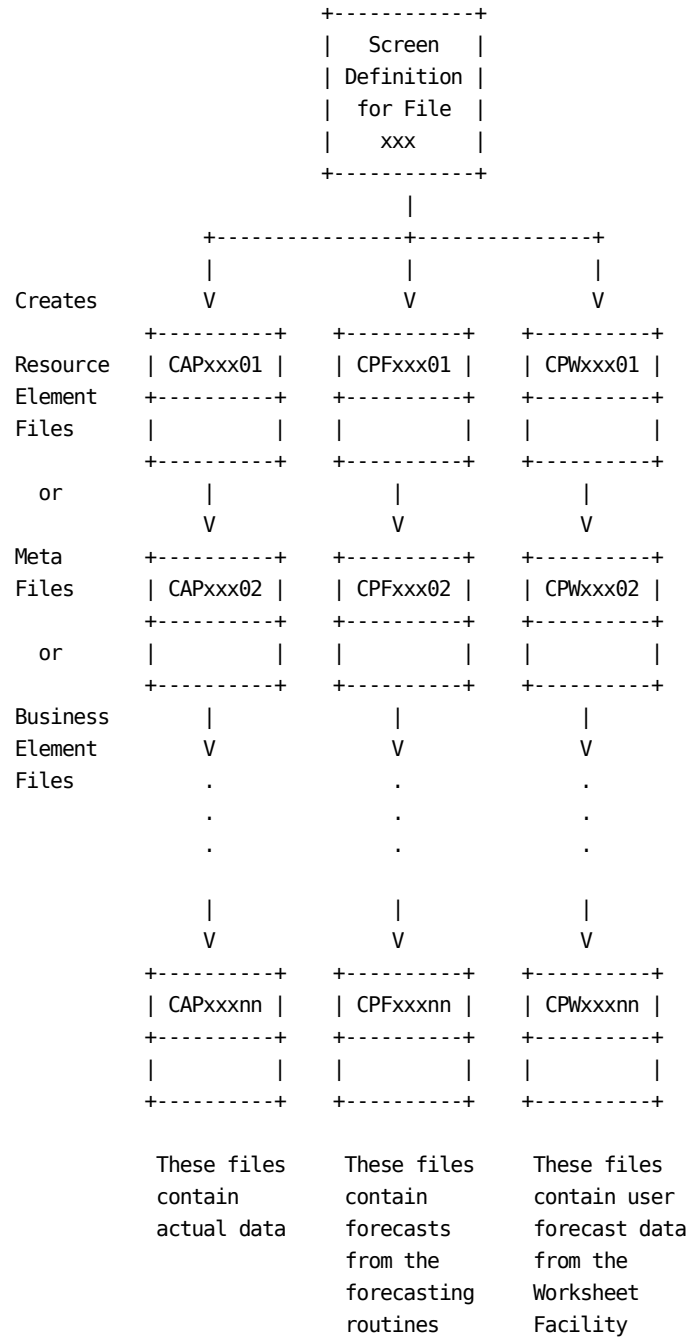


Figure 4-2. Typical Capacity Planning Database Files

If you would like a more complete explanation of this file structure, see Section 3.1.

It is important to remember that Capacity Planning forecast files may contain more than one forecast for a single data element. This typically occurs when forecasts are produced on a regular basis. Hence, after the second time a forecast is produced, the forecast file contains the forecasted observations for both the current (most recent) forecast and the previous forecast. Presentation Graphics automatically plots the observations for both the current and the previous forecast. If more than one previous forecast exists in the forecast file, then you may optionally specify which of these previous forecasts is to be plotted.

4.5 Component Operation

The Component Operation Section describes the control parameters that you can use to operate Presentation Graphics.

The following topics are addressed:

- 1 - Presentation Graphics Control Screen
- 2 - Color Graphics Format Parameters
- 3 - Y-Axis Reference Lines

4.5.1 Presentation Graphics Control Screen

To get to the Presentation Graphics Screen (shown in Figure 4-3) on your terminal, you must follow a short set of menus which are more fully described in Section 1.5 of this guide. The Presentation Graphics Screen appears after you indicate that you want to graph history and forecasts for presentation to management. Two fields on this screen are filled in on the fourth line when you first see it: the inquiry name and the inquiry description, separated by a dash. You entered these two fields on the Insert Capacity Planning Analysis Screen when you requested that a new inquiry be added.

```

----- Presentation Graphics -----
|Command ==>
|Enter a ? in any data entry field for more information on valid values.
|Composing CA MICS Inquiry:  _____ - _____
|
|Report title ==> _____
|
|Selection criteria
|      Input file ==> _____
|      Start date ==> _____ (ddmmyyyy)
|      SYSID ==> _____
|      Zone ==> _____
|      No. months ==> _____ (1-9999)
|      Specify CAPAPU Values ==> N (Y/N)
|
|Data element      ==> CHAN5CNT
|Previous forecast date ==> _____ (ddmmyyyy)
|Y-Axis label      ==> _____
|Plot type         ==> PRINTER (COLOR/PRINTER)
|Include Worksheets ==> _ (Y/N)
|
|Specify extended options ==> N (Y/N)
|
-----

```

Figure 4-3. Presentation Graphics Control Screen

Some of the parameters on the Presentation Graphics Screen shown in Figure 4-3 are required, while others are optional. The following is a discussion of the various parameters as they appear in this figure.

Report Title: The 60-character title field for the Presentation Graphics plot. No double quotes (") or semicolons (;) may appear within the body of the title.

- Input file:** The resource element, business element, or meta Capacity Planning database file containing the historical observations to be used as input to the analysis. If you are uncertain of the files or file names available in your Capacity Planning database, type a question mark (?) in the field and a selection list will be displayed.
- Start date:** The start date for date range selection. If the start date is specified, only historical observations at least as recent as this value are included in the analysis. The discussion under No. months describes the options for controlling the date range selection.
- SYSID:** The SYSID for which the report is produced. If the Capacity Planning database file was defined as having a composite SYSID, this value is automatically supplied on the screen.
- Zone:** The Zone, Hour, or Dayname to be included in the analysis, depending on whether you defined the Capacity Planning database file at the Zone, Hour, or Dayname secondary summarization level. If you defined the file as having a composite Zone, Hour, or Dayname this value is supplied automatically on the screen. Note that the field name will be changed automatically to either Zone, Hour, or Dayname once you specify the input file to reflect the input file summarization level.
- No. months:** The maximum number of days, weeks, or months of historical observations that will be included in the analysis. Note that once you specify the input file, the field name will change to reflect the time-span of the Input file. When used in conjunction with the Start date parameter, several options are available for date range selection control:
- o If you specify only Start date, all historical observations at least as recent as the value specified for Start date will be selected.

- o If you specify both Start date and No. months, the date range for data selection begins with the value specified for Start date and extends for the specified number of days, weeks, or months.
- o If you specify only No. months, the date range for data selection will include only this number of the most recent days, weeks, or months.
- o If you specify neither parameter, then all historical observations will be selected.

Specify CAPAPU values:	Put a Y in this field to specify the CAPAPU value(s) to be included in the analysis. The CAPAPU value is specified on the CAPAPU Selection Values screen, shown in Figure 4-4, that appears if you enter a Y in this field. You may specify it only if the Capacity Planning database file is defined as containing the CAPAPU data element.
Data element:	The input file data element that is to be plotted. If you are uncertain as to the correct name of the data elements on the input file, type a question mark (?) in this field and a selection list of available data elements will be presented.
Previous forecast date:	The specific previous forecast to be plotted. Note that in addition to the previous forecast, the current forecast (the one corresponding to the most recent historical observation to be plotted) is plotted automatically.
Y-Axis label:	The 30-character label for the Y-axis for the Presentation Graphics plot. No double quotes (") or semicolons (;) may appear within the body of the label.
Plot type:	The specification for the type of plot. You may specify COLOR for a color graphics plot or PRINTER for a line printer plot. Printer graphics are always available, however SAS/GRAPH is a prerequisite for color graphics.

Include Worksheets: Include any user forecasts generated for this data element through the use of the Capacity Planner Worksheet Facility. If present, this will produce an additional line on the output generated by Presentation Graphics.

Specify extended options: A Y in this field displays the Extended Options Menu screen, shown in Figure 4-5.

If you entered a Y in the Specify CAPAPU values field on the Presentation Graphics Control Screen, the CAPAPU Selection Values screen displays, as shown in Figure 4-4. On the CAPAPU Selection Values screen, you may specify the CAPAPU values (up to nine) to be included in the analysis. You may specify CAPAPUs only if the Capacity Planning database file is defined as containing the CAPAPU data elements. (Note that Presentation Graphics does not support wild cards.)

```
----- CAPAPU Selection Values -----
|Command ==>|
|
|Composing CA MICS Inquiry: _____|
|
|      Rule ----+----1----+----2----+----3----+----4----+----5----+----6----|
|CAPAPU  ==>|-----|
|CAPAPU 2 ==>|-----|
|CAPAPU 3 ==>|-----|
|CAPAPU 4 ==>|-----|
|CAPAPU 5 ==>|-----|
|CAPAPU 6 ==>|-----|
|CAPAPU 7 ==>|-----|
|CAPAPU 8 ==>|-----|
|CAPAPU 9 ==>|-----|
|
|
|-----|
```

Figure 4-4. CAPAPU Selection Values Screen

```

----- Extended Options Menu -----
|Command ==>
|
|Inquiry Step:  Presentation Graphics
|
|Description text:
|  ==> -----
|  ==> -----
|  ==> -----
|  ==> -----
|  ==> -----
|  ==> -----
|
|Override color graphics format parameters      ==> N (Y/N/R)
|          Color graphics format set           ==> PENPLOT
|
|Specify Y-axis reference (capacity) lines      ==> N (Y/N/R)
|
-----

```

Figure 4-5. Presentation Graphics Extended Options Menu

Some of the parameters on the Extended Options Menu shown in Figure 4-5 are required, while others are optional. The following is a discussion of the various parameters as they appear in this figure.

Description text: Up to six 60-character lines of descriptive text to appear as footnotes to the plot.

Override color graphics format parameters: Specify Y for yes, N for no, or R to reset and return to previous value.

Color graphics format set: Enter the format type that you specified in MICF.

Specify Y-axis reference (capacity) lines: Specify Y for yes, N for no, or R to reset and return to previous value.

4.5.2 Color Graphics Format Parameters

You can use the Color Graphics Format Parameters extended option to override parameter defaults which control color graphics format and appearance. These default values are specified through the MICF Color Graphics Format Parameters process. You can customize these defaults to match your most common reporting requirements; you need to override color graphics format defaults only for exceptional reporting requirements.

Figure 4-6 illustrates the Color Graphics Format Parameters screen.

```
/----- Color Graphics Format Parameters -----\
|Command ==>                                     Format set: PENPLOT |
|
|Inquiry step: Presentation Graphics                |
|
|Title print font      ==> XSWISSBE   Title color      ==> BLUE   |
|Draw border around plot ==> N   (Y/N) Background color ==> -----|
|Axis color            ==> BLACK     Area inside axis color ==> -----|
|Text color            ==> BLACK     |
|
|History:              Sym ==> STAR   Line ==> SOLID   Color ==> RED  |
|Forecast:              Sym ==> PLUS  Line ==> DOTTED  Color ==> BLUE |
|Previous forecast: Sym ==> PLUS  Line ==> DASHED   Color ==> GREEN  |
|
|Upper confidence: Sym ==> TRIANGLE Line ==> SOLID   Color ==> YELLOW|
|Lower confidence: Sym ==> TRIANGLE Line ==> SOLID   Color ==> YELLOW|
|Reference limits: Sym ==> NONE   Line ==> SOLID   Color ==> BLACK  |
|
|
|Symbol (NONE,PLUS,STAR,TRIANGLE,DIAMOND,HASH,SQUARE,A-Z,0-9,or SAS/GRAPH value)|
|Line   (SOLID,DASHED,DOTTED,1-32) |
|Color  (May be any color available on your graphics device.) |
\-----/
```

Figure 4-6. Color Graphics Format Parameters Screen

The parameters for the Color Graphics Format Parameters shown in Figure 4-6 screen are defined below:

Title print font: The character format (font) for the title line(s). Character font names are one to nine characters in length. Refer to the SAS/GRAPH User's Guide for a list of valid character fonts.

Title color:	The color of the title line(s).
Draw border around plot:	The specification to frame the graph with a border (that is, draw a line around the outside of the display).
Background color:	The specification to generate the graph on a colored background (that is, fill the entire graph with a color).
Axis color:	The color of the x-axis and y-axis.
Area inside axis color:	The specification to draw the plot lines on a colored background (that is, only the area inside the axis is filled).
Text color:	The color of all text words except the titles and labels.
Symbol:	The symbol for plotting the data values. See the SAS/GRAPH User's Guide for the available symbols.
Line type:	The line type for plotting the data values. See the SAS/GRAPH User's Guide for the available line types.
Color:	The color for plotting the data values. See the SAS/GRAPH User's Guide for the available colors.

4.5.3 Y-Axis Reference Lines

You may specify horizontal reference lines on the plots produced by Presentation Graphics to indicate control limits or objective values for the data element being plotted. For example, if a two-hour objective has been set for Class B job turnaround times, you might want to show this value as a limit or objective in a report produced for the data element. You may specify a maximum of 99 limits per plot.

Figure 4-7 shows a sample Y-Axis Reference Lines screen.

```

/----- Y-Axis Reference (Capacity) Lines -----\
|Command ===>                                Scroll ===> CSR|
|
|Inquiry Step: Presentation Graphics
|
|Line Cmds: I Insert D Delete
|
|Cmd      Date          Value
| -  ----  -  ----
| -  ----  -  ----
| -  ----  -  ----
| -  ----  -  ----
| -  ----  -  ----
| -  ----  -  ----
|***** BOTTOM OF DATA *****|
|
|
|
|
|
|
|
|
|
|
|
|
|

```

Figure 4-7. Y-Axis Reference Lines Screen

The parameters for the Y-Axis Reference Lines screen are defined below:

Date: The date on which this limit should begin to appear in the output plot. The dates need not be week-ending or month-ending dates since Presentation Graphics uses greater than or equal logic to select the proper limits.

Value: The numeric value of the limit. Note that SAS time formats are not supported for input and must be specified as numeric fields. For example, to place a reference line at the two-hour vertical point on the Y-axis, specify 7200 for the Value parameter.

4.6 Case Studies

The Case Studies Section presents three studies that illustrate the use of Presentation Graphics with historical and forecasted data.

The first case study, Extended Example I/O Workload Analysis, uses the 9021-711 I/O workload study introduced in Section 3.3.3.1 and forecasted in Section 6.6.1. This study shows a simple case of plotting historical data and the forecast generated by Univariate Model Forecasting.

The next case study, Performance Group/Period Workload Tracking, continues the performance group workload tracking study introduced in Section 3.3.3.2 and continued in Section 7.6.2. This study shows you how to use Presentation Graphics to track growth in TSO transaction rates. Following this is a similar study, System Workload and Performance Profile, using the system profile resource element file from the Capacity Planning database file originally developed in Section 3.3.3.5 and continued in Section 6.6.3.

The final case study, Tracking Previous Forecasts of Resource Utilization, shows you how to prepare a quarterly reporting system. This study shows how you can use multiple forecasts maintained in the same Capacity Planning database file in a single report, allowing you to project future trends and track the effectiveness of your earlier forecasts at the same time.

The following case studies are presented by section:

- 1 - Extended Example I/O Workload Analysis
- 2 - Performance Group/Period Workload Tracking
- 3 - System Workload and Performance Profile
- 4 - Tracking Previous Forecasts of Resource Utilization

4.6.1 Extended Example I/O Workload Analysis

This extended example applies Presentation Graphics to the univariate modeling forecast that was produced for the channel 5 SIO count in Section 7.6.1. In particular, the following topics are addressed:

- 1 - Problem Description
- 2 - Control Parameters
- 3 - Output Report

4.6.1.1 Problem Description

The purpose of this study is to prepare a Presentation Graphics printer plot report for the forecast of channel 5 SIO counts that we prepared in Section 7.6.1. In that extended example, we developed a linear model to forecast the SIO counts for the channel based on a geometrically smoothed historical data series.

4.6.1.2 Control Parameters

Figure 4-8 shows the Presentation Graphics screen that contains the control parameters used in this case study.

```

----- Presentation Graphics -----
|Command ==>
|Enter a ? in any data entry field for more information on valid values.
|Composing CA MICS Inquiry: CHNPLT - Channel 5 Workload Study
|
|Report title ==> HISTORY OF CHANNEL 5 SIO COUNT
|
|Selection criteria
|      Input file ==> CHN - I/O Channel Example,03/10/93
|      Start date ==> _____ (ddmnyyyy)
|      SYSID ==> TS01
|      Zone ==> 1
|      No. months ==> _____ (1-9999)
|      Specify CAPAPU Values ==> N          (Y/N)
|
|Data element      ==> CHAN5CNT
|Previous forecast date ==> _____ (ddmnyyyy)
|Y-Axis label      ==> SIO COUNT
|Plot type         ==> PRINTER  (COLOR/PRINTER)
|Include Worksheets ==> N          (Y/N)
|
|Specify extended options ==> Y          (Y/N)
|
-----

```

Figure 4-8. SIO Count Study Control Parameters

The parameters for the Presentation Graphics screen shown in Figure 4-8 are defined below:

Report title:	The title that appears on the Presentation Graphics reports.

Input file:	The CHN resource element file that was originally defined in the extended example in Section 3.3.3.1 of this guide.
Start date:	The start date for date range selection. Since neither start date nor number of months is specified for this analysis, no date range selection is applied to the historical observations from the CHN resource element file.
SYSID:	The historical observations for SYSID TS01 only.
Zone:	The Zone, Hour, or Dayname to be included in the analysis.
No. months	The maximum number of months that will be included in the analysis.
Specify CAPAPU value:	The specification for CAPAPU. The CHN resource element file was not defined as containing the CAPAPU data element, so no CAPAPU value may be specified on the CAPAPU Selection Values screen.
Data element:	The name of the data element (CHAN5CNT) in the capacity planning resource file which contains SIO counts for channel 5.
Previous forecast date:	The date of any previous forecast. In this case, there are no previous forecasts, so the entry is left blank.
Y-Axis label:	The label for the Y-axis. For this example, the label is SIO COUNT.
Plot type:	The specification for the type of plot. PRINTER is selected for the line printer plot.
Include Worksheets:	The specification to include information produced through the use of the Capacity Planner Worksheet Facility.

Specify extended options:	Type a Y in this field to enter the extended options menu, shown in Figure 4-9.


```

----- Extended Options Menu -----
|Command ==>
|
|Inquiry Step: Presentation Graphics
|
|Description text:
| ==> LINEAR MODEL OF CHANNEL 5 SIO COUNT DEVELOPED WITH_____
| ==> GMA SMOOTHED DATA_____
| ==> _____
| ==> _____
| ==> _____
| ==> _____
|
|Override color graphics format parameters      ==> N (Y/N/R)
|          Color graphics format set           ==> PENPLOT
|
|Specify Y-axis reference (capacity) lines      ==> N (Y/N/R)
|
-----

```

Figure 4-9. SIO Count Study Extended Options

The parameters for the Presentation Graphics Extended Options Menu shown in Figure 4-9 are defined below:

Description text: The descriptive text to appear as footnotes to the plot.

Override color graphics format parameters: Specify Y for yes, N for no, or R to reset and return to previous value.

Color graphics format set: Enter the format type that you specified in MICF.

Specify Y-axis reference (capacity) lines: Specify Y for yes, N for no, or R to reset and return to previous value.

4.6.1.3 Output Report

Figure 4-10 provides the Presentation Graphics printer plot for the extended example. The X-axis of this plot is scaled by months, with two years shown. This scale is selected based on the definition of the resource element file.

Note that the Y-axis is scaled so that the maximum Y-axis label is less than 1,000. In this case, each value on the Y-axis is scaled by 100,000, as indicated in the scale block to the lower left of the origin on the plot. Because we did not specify user-defined references, and no previous forecasts were available, neither appear in the printer plot.

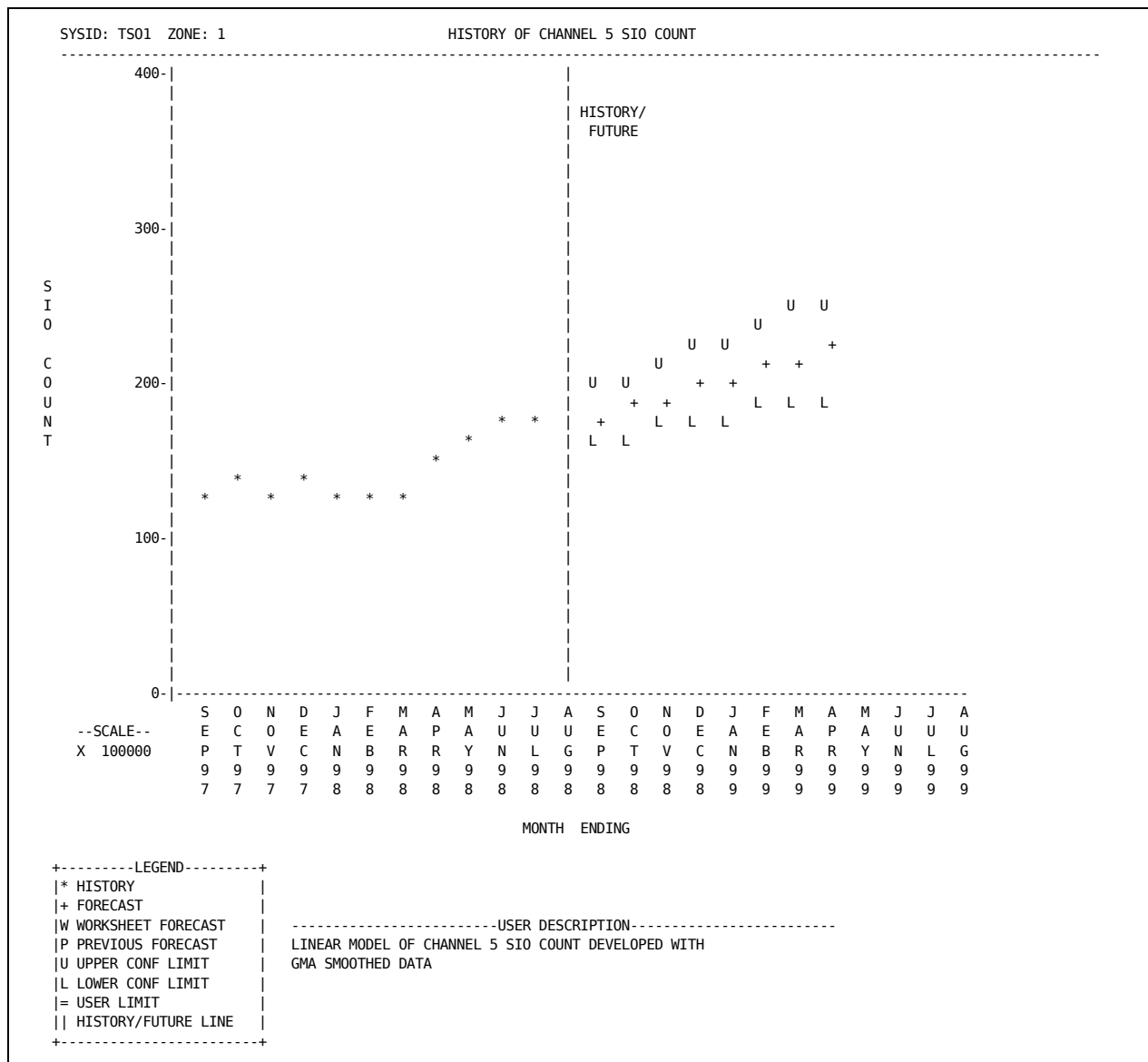


Figure 4-10. Extended Example Presentation Graphics Report

4.6.2 Performance Group/Period Workload Tracking

Continuing the Performance Group/Workload Tracking study introduced in Section 3.3.3.2 and continued in Section 6.6.2, Figures 4-11 and 4-12 illustrates the parameters that were entered on the Presentation Graphics control screen. Figure 4-13 shows the printer plot produced by Presentation Graphics.

```
----- Presentation Graphics -----
|Command ==>
|Enter a ? in any data entry field for more information on valid values.
|Composing CA MICS Inquiry: PRFPLT - TSO Period 1 Transactions
|
|Report title ==> HISTORY OF TSO PERIOD 1 TRANSACTIONS
|
|Selection criteria
|      Input file ==> PRF - TSO Period 1 Transactions,09/11/98
|      Start date ==> _____ (ddmmyyyy)
|      SYSID ==> TS01
|      Zone ==> 1
|      No. months ==> _____ (1-9999)
|      Specify CAPAPU Values ==> N (Y/N)
|
|Data element      ==> WKL01
|Previous forecast date ==> _____ (ddmmyyyy)
|Y-Axis label      ==> TRAN PER HOUR
|Plot type         ==> PRINTER (COLOR/PRINTER)
|Include Worksheets ==> N (Y/N)
|
|Specify extended options ==> Y (Y/N)
|
-----
```

Figure 4-11. Performance Group Study Control Parameters

```
----- Extended Options Menu -----
|Command ==>
|
|Inquiry Step: Presentation Graphics
|
|Description text:
| ==> UNIVARIATE MODEL DEVELOPED IN SECTION 7.6.2_____
| ==> _____
| ==> _____
| ==> _____
| ==> _____
| ==> _____
|
|Override color graphics format parameters    ==> N (Y/N/R)
|          Color graphics format set         ==> PENPLOT
|
|Specify Y-axis reference (capacity) lines    ==> N (Y/N/R)
|-----
```

Figure 4-12. Performance Group Study Extended Options

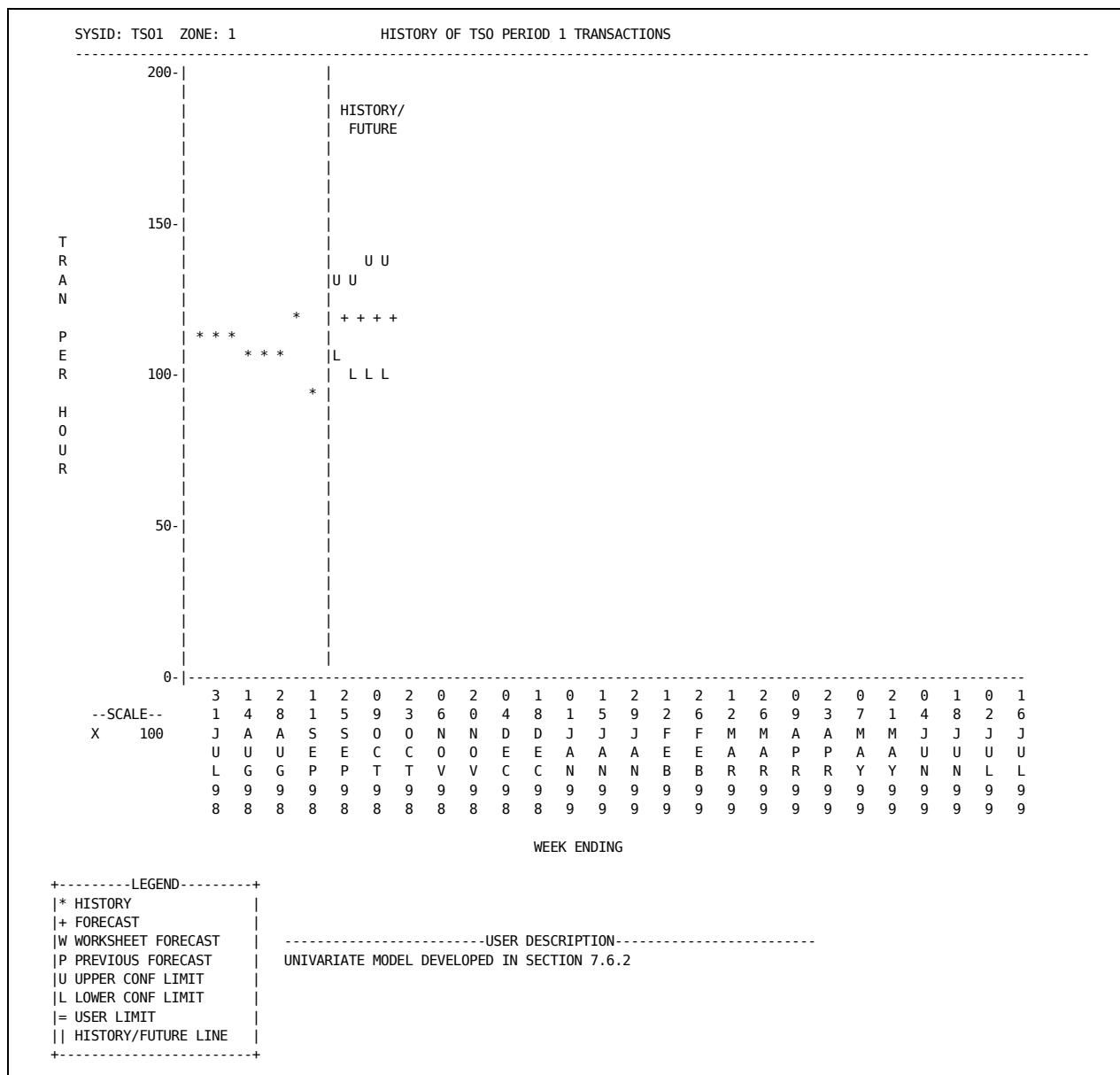


Figure 4-13. Presentation Graphics Report for the Performance Group Study

4.6.3 System Workload and Performance Profile

Continuing the System Workload and Performance Profile case study introduced in Section 3.3.3.5 and continued in Section 6.6.3, Presentation Graphics was used to produce a printer plot of the forecast of multiprocessor CPU busy that was developed. Figures 4-14 and 4-15 show the control parameters used to produce the report. Figure 4-16 shows the printer plot produced by Presentation Graphics.

```

----- Presentation Graphics -----
|Command ==>
|Enter a ? in any data entry field for more information on valid values.
|Composing CA MICS Inquiry: WKLPLT - CPU Utilization Plot
|
|Report title ==> HISTORY OF MULTIPROCESSOR CPU UTILIZATION FOR SYSID TS01
|
|Selection criteria
|      Input file ==> WKL - System Workload Example
|      Start date ==> _____ (ddmonyyyy)
|      SYSID ==> TS01
|      Zone ==> 1
|      No. months ==> _____ (1-9999)
|      Specify CAPAPU Values ==> N (Y/N)
|
|Data element      ==> CPUMPBS
|Previous forecast date ==> _____ (ddmonyyyy)
|Y-Axis label      ==> PCT CPU BUSY
|Plot type         ==> PRINTER (COLOR/PRINTER)
|Include Worksheets ==> N (Y/N)
|
|Specify extended options ==> Y (Y/N)
|
-----

```

Figure 4-14. System Workload Study Control Parameters

```
----- Extended Options Menu -----
|Command ==>
|
|Inquiry Step: Presentation Graphics
|
|Description text:
| ==> PLOT OF MODEL DEVELOPED IN SECTION 7.6.3_____
| ==> _____
| ==> _____
| ==> _____
| ==> _____
| ==> _____
|
|Override color graphics format parameters    ==> N (Y/N/R)
|           Color graphics format set        ==> PENPLOT
|
|Specify Y-axis reference (capacity) lines    ==> N (Y/N/R)
|-----
```

Figure 4-15. System Workload Study Extended Options

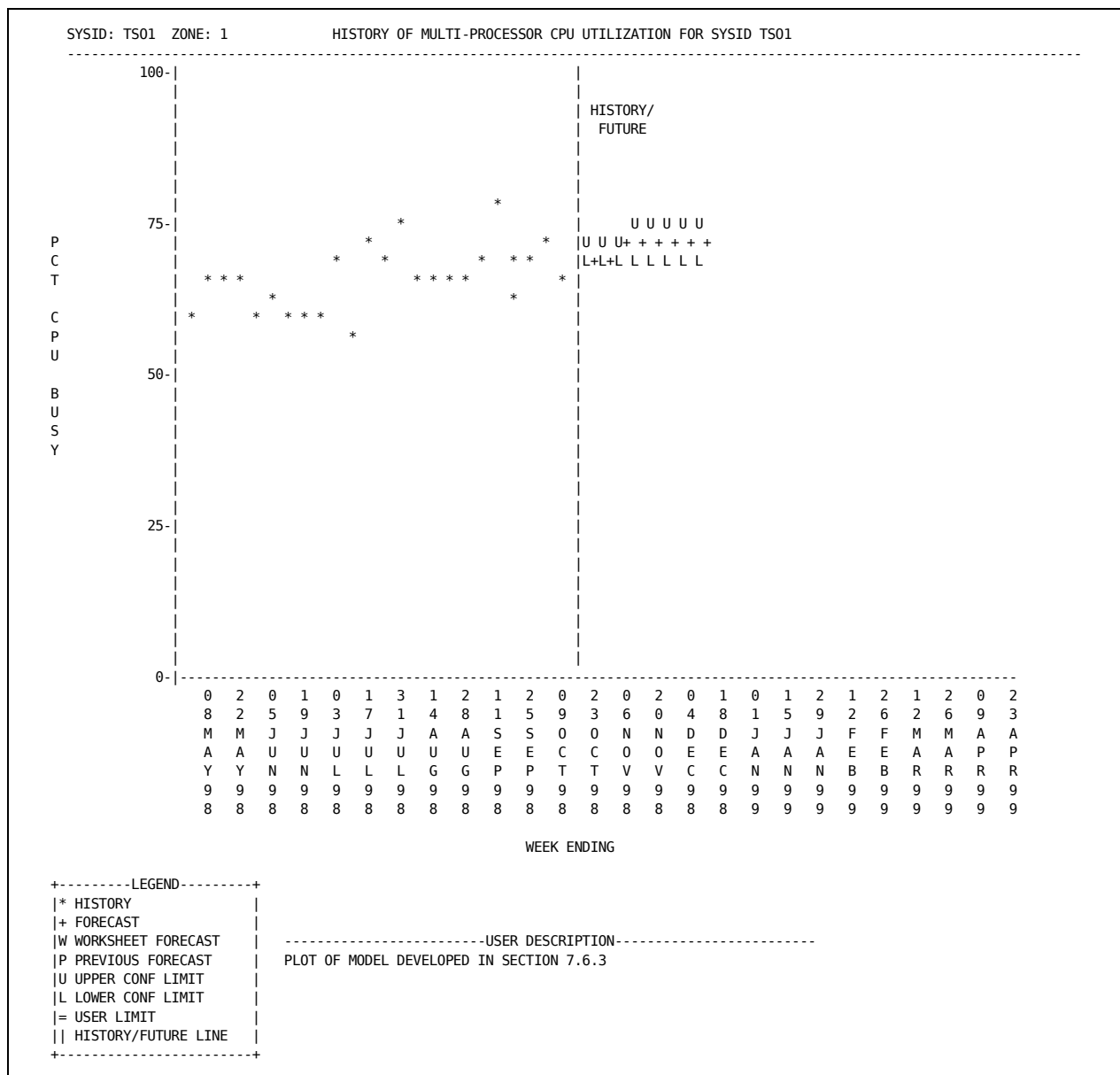


Figure 4-16. Presentation Graphics Report for the System Workload Case Study

4.6.4 Tracking Previous Forecasts of Resource Utilization

In this example, you use channel 5 SIO count data from the 9021-711 I/O study in Section 6.6.1 to produce the required quarterly management report. The following sections discuss this report:

- 1 - Problem Description
- 2 - Control Parameters
- 3 - Quarterly Report

4.6.4.1 Problem Description

Data for channel 5 SIO counts was collected monthly, beginning in September 1985. As part of a regular capacity management program, projections are made quarterly, beginning in March 1986. Each projection is based on the previous six months' historical values. By the time we are ready to produce the June 1986 quarterly report, the capacity planning forecast file contains multiple forecasts (one for each quarter since March 1985), each representing a forecast based on the previous six months of data.

The object of the June 1986 quarterly report is to show the current projection for channel 5 SIO counts and to track the projection from the previous quarter against the more recent actual observations.

4.6.4.2 Control Parameters

Figures 4-17 and 4-18 show Presentation Graphics screens that contain the control parameters that we used to produce the quarterly management report.

Because multiple forecasts, each based on different data, are maintained in the Capacity Planning database files, it is easy to plot one of the previous forecasts using the forecast date as the Previous forecast date parameter. If you do not specify a value for this parameter, Presentations Graphics automatically selects the most recent prior forecast for plotting.

In this example, the current forecast is based on data through June 1998 (and is thus the June 1998 forecast). The most recent prior forecast is the March 1998 forecast (the one based on data through March 1998). Leaving the previous forecast parameter blank has the same result as the specification shown in Figure 4-19 (that is, the specified previous forecast, March 1998, is also the default previous forecast).

```

----- Presentation Graphics -----
|Command ==>
|Enter a ? in any data entry field for more information on valid values.
|Composing CA MICS Inquiry: CHNQPL - Quarterly Channel 5 Plot
|
|Report title ==> HISTORY OF CHANNEL 5 SIO COUNT
|
|Selection criteria
|      Input file ==> CHN - I/O Channel Example,03/10/98
|      Start date ==> _____ (ddmnyyyy)
|      SYSID ==> TS01
|      Zone ==> 1
|      No. months ==> _____ (1-9999)
|      Specify CAPAPU Values ==> N (Y/N)
|
|Data element      ==> CHAN5CNT
|Previous forecast date ==> _____ (ddmnyyyy)
|Y-Axis label      ==> SIO COUNT
|Plot type         ==> PRINTER (COLOR/PRINTER)
|Include Worksheets ==> N (Y/N)
|
|Specify extended options ==> Y (Y/N)
|
-----

```

Figure 4-17. I/O Channel Study Control Parameters

```
----- Extended Options Menu -----
|Command ==>
|
|Inquiry Step:  Presentation Graphics
|
|Description text:
| ==> LINEAR MODEL DEVELOPED WITH GMA SMOOTHED DATA_____
| ==> _____
| ==> _____
| ==> _____
| ==> _____
| ==> _____
|
|Override color graphics format parameters      ==> N (Y/N/R)
|          Color graphics format set           ==> PENPLOT
|
|Specify Y-axis reference (capacity) lines      ==> N (Y/N/R)
|
-----
```

Figure 4-18. I/O Channel Study Extended Options

4.6.4.3 Quarterly Reports

Figure 4-19 shows the quarterly management report produced by the parameters entered on the Presentation Graphics screen shown in Figures 4-17 and 4-18.

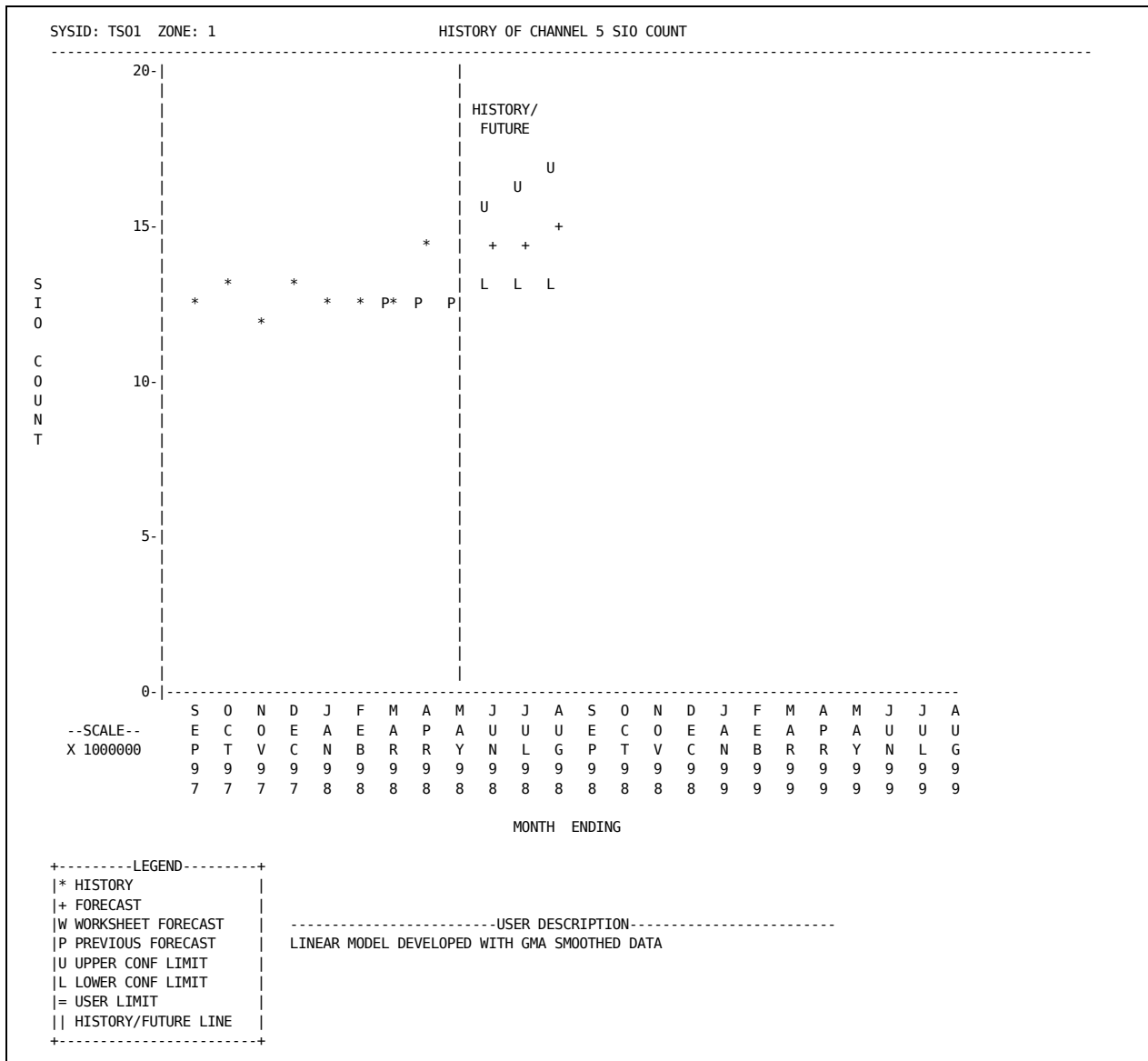


Figure 4-19. Extended Example Quarterly Management Report

Chapter 5: PROFILE AND TRENDING ANALYSIS

The Profile and Trending Analysis software provides a convenient method for presenting historical data, calculating growth rates, and estimating future values of the data elements contained in your Capacity Planning database. The facility is designed to address two requirements:

- o Tabular reporting of historical data observations
- o Provision of simple "sense of direction" estimates of future values.

The reports can be produced at either the WEEKS, MONTHS, or DAYS timespan.

The following sections describe Profile and Trending Analysis and its reports.

This section contains the following topics:

- [5.1 Functional Description](#) (see page 557)
- [5.2 Usage Guidelines](#) (see page 558)
- [5.3 Standard Output](#) (see page 560)
- [5.4 Analytic Technique Tutorial](#) (see page 565)
- [5.5 Component Operation](#) (see page 568)
- [5.6 Case Studies](#) (see page 573)

5.1 Functional Description

Profile and Trending Analysis allows you to tailor and generate reports that list historical data observations and estimate future observations using a compound rate of growth methodology. You may apply the facility to any of the files in your Capacity Planning database.

Profile and Trending Analysis calculates the compound growth rate between the first and the last observation (the long-term rate of growth) and between the last two observations (the short-term rate of growth). Based on the control parameters you provide, the program estimates a specified number of future observations for each data element, using either the long- or short-term rate of growth. The program produces three primary reports:

- o Historical Observations Report. This report lists the historical values for each of the data elements you specify.
- o Growth Rate Report. This report details the short- and long-term growth rates calculated for each of the data elements you specify.
- o Estimates of Future Observations Report. This report details the future estimates for each of the data elements you specify. These estimates are calculated using the specified growth rate.

5.2 Usage Guidelines

The principal items to consider when you use Profile and Trending Analysis are the number of future intervals you wish to estimate and whether you want the future estimates based on the long- or short-term rate of change.

The number of intervals that can be forecast based on a historical data series is a function of two variables:

- o The length (that is, number of observations) of the historical data series
- o The behavior of the historical data series

As a general rule, the number of forecast intervals that you specify should never exceed the length of the historical data series. Limiting the length of the forecast to a maximum of one-third or one-half the length of the historical data series is often a far more judicious choice.

You must also consider the qualitative behavior of the historical data series when you select the number of intervals to be forecast. If the historical data series appears to represent a smooth trend, then you can specify a forecast period one-half to two-thirds the length of the data. If the historical data series exhibits high variability or apparent randomness, then you should specify a very short period.

Quantitative techniques for determining the behavior of the historical data series and the length of forecast are included in Chapter 7, Univariate Model Forecasting.

You must also specify whether the future estimates will be based on the short- or long-term growth rates. Each of these two options have both advantages and disadvantages.

- o If you specify the short-term growth rate, the future estimates are based on the rate of change exhibited between the final two observations. You can use this option, which is extremely sensitive to the system's most recent behavior, to identify changes in long-term behavior patterns (that is, when the signs of the short- and long-term growth rates are different).
- o The short-term growth rate can produce deceptive results for data elements that you report as totals, such as total TSO LOGONS, when the MONTH level of detail is specified. This problem is caused by the varying number of days that is represented by any month. Whenever possible, convert such variables to percents or rates to avoid the problems that are introduced by dissimilar measurement intervals.
- o The long-term growth rate produces much more conservative estimates, since the significant variations that can occur between any two time periods are dampened when more intermediate time periods are included. This dampening results since the long-term rate of growth is the n -th root (see Equation 1 in Section 5.4) of the growth rate between the two observations.

For extended forecasts, the long-term growth rate often represents a better choice, since the likelihood of significant growth rates is small. However, be careful to avoid the long-term growth rate if there was some significant change to the hardware or software that produced one or more steps in the historical data series.

5.3 Standard Output

Profile and Trending Analysis produces three primary reports which describe the historical data, show the compounded growth rates, and give estimates of future growth. These reports are described in the following sections:

- 1 - Historical Observations Report
- 2 - Growth Rate Report
- 3 - Estimates of Future Observations Report

5.3.1 Historical Observations Report

The Historical Observations Report provides a summary of the historical values contained in the Capacity Planning database file. You define the structure of this report using the Profile and Trending Analysis control screen shown in Figure 5-6.

While the principal application of these reports is intended to be the presentation of historical data, you can also use them to examine the behavior of the data before applying the more detailed statistical analysis procedures provided in the remaining chapters.

Figure 5-1 shows a sample Historical Observations Report. The following information is shown:

SYSID: The SYSID specified by the SYSID control parameter.

ZONE: The ZONE specified by the LVL control parameter.

For each historical observation, the following data elements appear in the report columns:

MONTH: The MONTH and YEAR of the historical observation. If you specify the WEEKS timespan when the Capacity Planning database file is created, then the title for this column is WEEK ENDING and the week ending dates appear in the report. If you specify the DAYS timespan when the Capacity Planning database file is created, then the title for this column is DAY ENDING and the day ending dates appear in the report.

You specify the remaining columns of the report through parameters on the Profile and Trending Analysis control screen. For example, the first column of the report shown in Figure 5-1 is:

CHANNEL 1 The historical observations for the average
 BUSY (%): monthly channel 1 busy values computed for ZONE
 1.

+-----+-----+-----+-----+-----+-----+-----+-----+							
SYSID: TS01							
ZONE: 1		EXTENDED EXAMPLE --- MONTHLY HISTORICAL OBSERVATIONS					
+-----+-----+-----+-----+-----+-----+-----+-----+							
MONTH	CHANNEL 1	CHANNEL 2	CHANNEL 3	CHANNEL 4	CHANNEL 5		
	BUSY (%)	BUSY (%)	BUSY (%)	BUSY (%)	BUSY (%)		
+-----+-----+-----+-----+-----+-----+-----+-----+							
SEP 97	5.4	18.9	18.8	7.3	15.8		
OCT 97	5.5	19.5	19.6	7.4	14.8		
NOV 97	6.4	19.8	20.2	10.0	15.9		
DEC 97	5.3	18.0	18.6	8.9	14.6		
JAN 98	6.1	19.0	19.4	9.0	15.7		
FEB 98	6.3	23.1	18.8	11.1	16.2		
MAR 98	5.5	19.7	16.9	9.5	14.9		
APR 98	4.8	18.0	15.2	7.7	14.3		
MAY 98	4.6	17.8	15.7	8.2	14.5		
JUN 98	5.0	17.6	15.8	9.3	15.1		
JUL 98	6.9	21.0	18.6	11.7	17.8		
AUG 98	8.6	30.9	15.0	10.7	15.4		
+-----+-----+-----+-----+-----+-----+-----+-----+							

Figure 5-1. Historical Observations Report

5.3.2 Growth Rate Report

The Growth Rate Report provides the values of the long- and short-term compound growth rates for the variables that you specify on the Profile and Trending Analysis control screen (see Figure 5-6). Figure 5-2 shows a sample Growth Rate Report.

+-----+-----+-----+-----+-----+-----+-----+-----+							
SYSID: TS01							
ZONE: 1		EXTENDED EXAMPLE --- GROWTH RATE PER MONTH					
+-----+-----+-----+-----+-----+-----+-----+-----+							
INTERVAL	CHANNEL 1	CHANNEL 2	CHANNEL 3	CHANNEL 4	CHANNEL 5		
	BUSY (%)	BUSY (%)	BUSY (%)	BUSY (%)	BUSY (%)		
+-----+-----+-----+-----+-----+-----+-----+-----+							
11 MOS	4.40 %	4.57 %	-2.02 %	3.63 %	-0.24 %		
LAST MO	23.94 %	47.08 %	-19.22 %	-8.37 %	-13.52 %		
+-----+-----+-----+-----+-----+-----+-----+-----+							

Figure 5-2. Growth Rate Report

The following information is shown:

SYSID: The SYSID specified by the SYSID control parameter.

ZONE: The ZONE specified by the LVL control parameter.

INTERVAL: Two interval values are shown in the report. The first interval identifies the number of intervals used to compute the long-term rate of growth. The second interval specifies either LAST WK, LAST MO, or LAST DY depending on the time-span you specified for the Capacity Planning database file.

You can specify the remaining columns of the report using the control parameters on the Profile and Trending Analysis control screen. For example, the first column of the report shown in Figure 5-2 contains the following information:

CHANNEL 1 BUSY (%): The two rates of growth that are shown for channel 1 utilization. The long-term growth rate (that is, for the last 11 months) for utilization is 4.40%. Simply stated, the August 1998 value of 8.6% busy represents a 4.40% compound increase per month from the value recorded for September 1997. The short-term rate of growth (that is, 23.94%) is the percent difference between the values for July and August 1998.

These growth rates provide a sense of direction for the variables you specify. In general, the long-term growth rate is a much more reliable estimator, since it is less susceptible to step function changes that occur when a new workload is introduced. The short-term growth rate value is useful for determining when a long-term trend has been reversed.

For example, the channel utilization values for channel 4 (see Figure 5-2) were growing at a long-term rate of growth of 3.63%. However, the differences between July and August 1998 result in a short-term rate of growth of -8.37%. The sign difference between the long- and short-term rates represents a trend reversal.

A positive long-term growth rate does not imply that each of the observations in the historical series has been steadily increasing, but only that the long-term trend between the first and last observation is positive (see Figure 5-1 for channel 1).

5.3.3 Estimates of Future Observations Report

The Estimates of Future Observations Report provides estimates of future observations for specified data elements. As with the Historical Observation and the Growth Rate Reports, you define the structure of this report using the Profile and Trending Analysis control screen (see Figure 5-6).

Figure 5-3 shows a sample Estimates of Future Observations Report.

+-----+-----+-----+-----+-----+-----+-----+-----+							
SYSID: TS01							
ZONE: 1		EXTENDED EXAMPLE --- ESTIMATES OF FUTURE MONTHLY OBSERVATIONS					
+-----+-----+-----+-----+-----+-----+-----+-----+							
MONTH	CHANNEL 1 BUSY (%)	CHANNEL 2 BUSY (%)	CHANNEL 3 BUSY (%)	CHANNEL 4 BUSY (%)	CHANNEL 5 BUSY (%)		
+-----+-----+-----+-----+-----+-----+-----+-----+							
SEP 98	9.0	32.3	14.7	11.1	15.3		
OCT 98	9.4	33.8	14.4	11.5	15.3		
NOV 98	9.8	35.4	14.1	11.9	15.3		
DEC 98	10.2	37.0	13.9	12.4	15.2		
JAN 99	10.7	38.7	13.6	12.8	15.2		
FEB 99	11.1	40.4	13.3	13.3	15.2		
+-----+-----+-----+-----+-----+-----+-----+-----+							

Figure 5-3. Estimates of Future Observations Report

The following information is shown:

SYSID: The SYSID specified by the SYSID control parameter.

ZONE: The ZONE specified by the LVL control parameter.

The number of estimated future observations is determined by a control parameter you specify on the Profile and Trending Analysis control screen. For each estimated future observation, the following data elements appear in the report columns:

MONTH: The MONTH and YEAR of the historical observation. If you specify the WEEKS time-span when the Capacity Planning database file is created, the title of this column is WEEK ENDING and the week ending dates appear in the report. If you specify the DAYS timespan when the Capacity Planning database file is created, the title of this column is DAY ENDING and the day ending dates appear in the report.

NOTE: If you specified DAYNAME as the secondary summarization level then these dates will be seven days apart. Otherwise, the Capacity Planner will generate one observation per day.

You can specify the remaining columns of the report through the control parameters on the Profile and Trending Analysis control screen. For example, the first column of the report shown in Figure 5-3 contains the following information:

**CHANNEL 1
BUSY (%):** The future estimates of the average monthly values for channel 1 utilization. To calculate these values, repetitively apply the growth rate you specified to the final observation for the variable. For channel 1 busy, the estimated value of 9.0% for September 1998 corresponds to a 4.4% growth of the August 1998 value, 8.6% (taken from the historical report shown in Figure 5-1).

These reports are intended to provide simple sense of direction estimates for future values. These estimates are not statistically based forecasts, but provide estimates based on historical growth rates.

5.4 Analytic Technique Tutorial

Profile and Trending Analysis analyzes historical series of observations of data elements contained in the Capacity Planning database. Each historical data series is analyzed to determine two compound growth rates:

- o Long-term Growth Rate: The compound growth rate which, if applied to the first observation n times (where n is the number of time intervals between the first and last historical observations), yields the final observation in the series.
- o Short-term Growth Rate: The compound growth rate which, if applied to the next to last observation n times (where n is the number of time intervals between the last two historical observations), yields the final observation in the series. This value is defined as the short-term growth rate. These growth rates provide a sense of direction for estimating future resource requirements (MOR81).

The formula for the compound rate of growth is shown in Equation 1:

$$g = \left[1 + \frac{X(n) - X(m)}{X(m)} \right]^{1/(n-m)} - 1 \quad (\text{Eqn 1})$$

where: $X(i)$ is the i -th observation of user-specified variable X .

g is the compound rate of growth.

n is the number of the last time interval being used for the rate of growth calculations. (The time intervals are numbered starting with the oldest observations.)

m is the number of the previous time interval being used for the calculations.

The formula calculates the j-th root of the percent change between any two observations of X. Therefore, if the growth rate is multiplied by the first observation j times, the result is equal to the final observation.

You can then develop a forecast on the next k observations using either the long-term or short-term growth rate. The number of intervals to be forecast (k), and which growth rate should be used for the calculations are specified via control parameters on the Profile and Trending Analysis control screen (see Figure 5-4). The rationale for these selections is discussed in Section 5.2, Usage Guidelines.

The formula used to calculate the future estimates is shown in Equation 2:

$$X(j+1) = X(j) * (1 + g) \quad (\text{Eqn 2})$$

where: X(i) is the i-th observation of user-specified variable X.

g is the compound rate of growth.

j is the last known or estimated time interval.

j+1 is the time interval being estimated from the previous time interval using the compound rate of growth g.

The formula calculates each new observation by multiplying the growth rate by the previous observation. Although the growth rate for any variable will not remain a constant, this technique is valuable for determining a sense of direction for the variables you specify.

The following example illustrates the application of the compound growth rate methodology. Typical data for monthly CPU utilization values are shown below:

MONTH	INTERVAL NUMBER	CPU UTILIZATION
JAN 97	1	54.3
FEB 97	2	56.7
MAR 97	3	55.8
APR 97	4	58.0
MAY 97	5	59.1
JUN 97	6	60.9

Using Equation 1, The values for the first and last time intervals (January and June 1997) were substituted to get the following:

$$g = \left| 1 + \frac{(60.9 - 54.3)}{54.3} \right|^{** (1/(6-1))} - 1$$

$$g = 1.12^{** 0.2} - 1.0 = 1.02 - 1.0 = 0.02$$

Solving for the short-term rate of growth, you get the following:

$$g = \left| 1 + \frac{(60.9 - 59.1)}{59.1} \right|^{** (1/(6-5))} - 1$$

$$g = 1.03^{** 1.0} - 1.0 = 1.03 - 1.0 = 0.03$$

Assuming that you want to estimate the future observation using the long-term rate of growth, use Equation 2 to estimate the value for July 1997:

$$X(7) = 60.9 * (1 + 0.02) = 62.1$$

To estimate successive observations based on the previous estimate and the rate of growth, reapply the equation.

5.5 Component Operation

The Component Operation Section describes the screen and parameters that are used to control Profile and Trending Analysis. The control parameters allow you to define the Capacity Planning database file to be used as input, the time periods, and the variables to be used in the forecasting process. In addition, you may tailor the contents of the reports produced.

To get to the Profile and Trending Analysis screen (shown in Figure 5-4) on your terminal, you must follow a short set of menus which are more fully described in Section 1.5 of this guide. The Profile and Trending Analysis control screen appears after you indicate that you want to insert a new Profile and Trending Analysis inquiry into your catalog of analysis inquiries. Two fields on this screen are filled when you first see it: the inquiry name and the inquiry description, separated by a dash. You entered these two fields on the Insert Capacity Planning Analysis screen when you requested that a new inquiry be added.

```

/----- Profile and Trending Analysis ----- ROW 1 OF 63 \
|Command ==>                               Scroll ==> CSR |
|Enter a ? in any data entry field for more information on valid values. |
|Composing CA MICS Inquiry:      -      |
|                                |
|Report title ==> Adhoc Inquiry - Will Not Be Saved |
|                                |
|Selection criteria: |
|      Input file ==> ___ - _____ |
|      Start date ==> _____ (ddmonyyyy) |
|      SYSID ==> _____ |
|      Zone ==> _____ |
|      No. months ==> _____ (1-9999) |
|Specify CAPAPU Values ==> N      (Y/N) |
|                                |
|Forecast length (weeks) ==> _____ (1-999) |
|Forecast type ==> _____ (LONG/SHORT) |
|                                |
|      Name   Page  Col.  Col. Heading 1  Col. Heading 2  Output Format |
|      ----- -   -   -----  -----  -----  ----- |
|      _____ 1   1   _____  _____  _____ |
|      _____ 1   2   _____  _____  _____ |
|      _____ 1   3   _____  _____  _____ |
|      _____ 1   4   _____  _____  _____ |
|      _____ 1   5   _____  _____  _____ |
|      _____ 1   6   _____  _____  _____ |
|      _____ 1   7   _____  _____  _____ |
|      _____ 2   1   _____  _____  _____ |
|      _____ 2   2   _____  _____  _____ |
|                                |
\-----/

```

Figure 5-4. Profile and Trending Analysis Control Screen

Some of the parameters in Figure 5-4 are required, while others are optional. The following is a discussion of the various parameters as they appear in this figure.

Input file: The resource element or business element Capacity Planning database file that contains the historical observations to be used as input to the analysis. If you are uncertain of the files or file names available in your Capacity Planning database, type a question mark (?) in the field and a selection list will be displayed.

Start date: The start date for date range selection. If Start date is specified, only historical observations at least as recent as this value are included in the analysis.

SYSID: The SYSID for which the report is to be produced. If the Capacity Planning database file was defined as having a composite SYSID, this value is automatically supplied on the screen.

Zone: The Zone, Hour, or Dayname to be included in the analysis, depending on whether you defined the Capacity Planning database file at the Zone, Hour, or Dayname summarization level. If you defined the file as having a composite Zone, Hour, or Dayname, this value is supplied automatically on the screen. Note that the field name will be changed automatically to Zone, Hour, or Dayname once you specify the input file to reflect the input file summarization level.

No. months: The maximum number of days, weeks, or months of historical observations that will be included in the analysis. Note that once you specify the input file, the field name will change to reflect the timespan of the input file. When used in conjunction with the Start date parameter, several options are available for date range selection control:

- o If you specify only Update start date, all historical observations at least as recent as the value specified for Update start date are selected.

- o If you specify both Update start date and Number months/weeks/days, the date range for data selection begins with the specified value for Start date and extends for the specified number of days, weeks, or months.
- o If you specify only Number of months/weeks/days, the date range for data selection includes only the number of the most recent months, weeks, or days.
- o If you specify neither parameter, then all historical observations are selected.

Note: If the file uses the secondary summarization level of DAYNAME, then these dates will be seven days apart; otherwise you will receive one observation per day.

Specify CAPAPI values:	Put a Y in this field to specify the CAPAPI values to be included in the analysis. The CAPAPI value is specified on the CAPAPI Selection Values screen, shown in Figure 5-5. You may specify it only if the Capacity Planning database file is defined as containing at least one CAPAPI.
Report title:	The 60-character title field for the profile and trending reports. No double quotes (") or semicolons (;) may appear within the body of the title.
Forecast length dy/wk/mn.:	The length of forecast period in days, weeks, or months. Note that the field name changes once you specify the input file to reflect the timespan of the input file.
Forecast type:	The growth rate (long-term or short-term) used to estimate future values.

The lower portion of the control screen contains a scrollable data entry display in which you may specify the data elements to be included in the analysis and on the reports. One entry is provided for each of seven columns on nine pages. Blank entries represent blank columns. Trailing blank pages are not printed. The following fields are available for each row in the display area:

Name:	The name of the data element to be included in the analysis. If you are uncertain as to the correct name of the data elements on the input file, type a question mark (?) in this field and a selection list of available data elements will be presented.
Page:	The page of the report on which the data element appears. You cannot change this field.
Column:	The column in which the data element appears. You cannot change this field.
Col. Heading 1 and Col. Heading 2:	The first and second lines of the column heading. The maximum permissible length for the column headings is 14 characters. No double quotes (") or semicolons (;) may appear within the body of the headings.
Output format:	The output format that is used to print the values for this data element in the reports. The report field is 14 columns wide, so any value that you enter here should specify a field width of 14, or the values will not be right-justified. If you specify a format value on the CA MICS File and Data Element Selection screen when the Input file is defined, this value is supplied automatically as a default. If you do not specify any value here, then the Profile and Trending Analysis uses one of two default output formats. For any data element name that ends in the characters TM, a SAS format of TIME14.2 is used. All other data elements default to a SAS format of 14.2.

If you entered a Y in the Specify CAPAPU values field on the Profile and Trending Analysis screen, the CAPAPU Selection Values screen displays, as shown in Figure 5-5. On the CAPAPU Selection Values screen, you may specify the CAPAPU values (up to nine) to be included in the analysis. You may specify CAPAPUs only if the Capacity Planning database file is defined as containing the CAPAPU data elements. (Note that Profile and Trending Analysis does not support wild cards.)

```

/----- CAPAPU Selection Values -----\
|Command ==>                               |
|                                           |
|Composing CA MICS Inquiry: _____    |
|                                           |
|      Rule  ---+---1---+---2---+---3---+---4---+---5---+---6---|
|CAPAPU   ==> _____|
|CAPAPU 2 ==> _____|
|CAPAPU 3 ==> _____|
|CAPAPU 4 ==> _____|
|CAPAPU 5 ==> _____|
|CAPAPU 6 ==> _____|
|CAPAPU 7 ==> _____|
|CAPAPU 8 ==> _____|
|CAPAPU 9 ==> _____|
|                                           |
|                                           |
\-----/

```

Figure 5-5. CAPAPU Selection Values Screen

5.6 Case Studies

This section presents three case studies that show you how to apply Profile and Trending Analysis to solve capacity planning problems that many data processing installations commonly encounter.

The first case study, Extended Example I/O Workload Analysis, is an extended example that shows a profile and trending analysis of I/O growth on a small MVS system. This study is typical of studies performed to determine if a machine upgrade is required to support the future demands of the current workload. A complete set of reports, along with interpretative commentary, is presented with this study.

The second study, Performance Group/Period Workload Tracking, is an analysis of the number of transactions processed by a TSO performance group, divided by periods. This study is used to determine if the character of the TSO workload is changing over time.

The third study, JES Spool, views the loading of various printers on a JES2 system. This study contrasts the load on the local printer in the data center with six different departmental remote printers.

The case studies are presented in the following sections:

- 1 - Extended Example I/O Workload Analysis
- 2 - Performance Group/Period Workload Tracking
- 3 - JES Spool

5.6.1 Extended Example I/O Workload Analysis

This I/O workload analysis extended example details the application of Profile and Trending Analysis to the I/O workload data collected in the extended example introduced in Section 3.3.3.1.

To illustrate the use of the reports, monthly data was collected for the selector and block multiplexor channels on the base channel group of a 9021-711 to study the growth of the system's I/O workload. Each channel is characterized by average percent busy and the number of SIOs completed during the month.

The following topics are addressed:

- 1 - Problem Description
- 2 - Control Parameters
- 3 - Output Reports

5.6.1.1 Problem Description

The purpose of this study is to develop simple estimates of the future I/O load that would be experienced by the selector and multiplexor channels on a 9021-711. The example shows 12 months of historical data collected for channels 1 through 5 for the system SYSID TS01 for ZONE 1, and a report designed to present the historical data and estimates.

```

/----- Profile and Trending Analysis ----- ROW 1 OF 63--\
|Command ==>                               Scroll ==> CSR |
|Enter a ? in any data entry field for more information on valid values. |
|Composing CA MICS Inquiry: CHNPFT - Channel SIO Profile and Trending |
| |
|Report title ==> Extended Example |
| |
|Selection criteria: |
|   Input file ==> CHN - I/O Channel Example |
|   Start date ==> _____ (ddmmyyyy) |
|   SYSID ==> TS01 |
|   Zone ==> 1_____ |
|   No. months ==> _____ (1-9999) |
|   Specify CAPAPU Values ==> N (Y/N) |
| |
|Forecast length (months) ==> 6 (1-999) |
|Forecast type ==> LONG (LONG/SHORT) |
| |
|   Name   Page  Col.   Col. Heading 1  Col. Heading 2  Output Format |
|   -----  -   -   -----  -----  ----- |
|   CHAN1BS   1   1   CHANNEL 1     BUSY (%)     14.1 |
|   CHAN2BS   1   2   CHANNEL 2     BUSY (%)     14.1 |
|   CHAN3BS   1   3   CHANNEL 3     BUSY (%)     14.1 |
|   CHAN4BS   1   4   CHANNEL 4     BUSY (%)     14.1 |
|   CHAN5BS   1   5   CHANNEL 5     BUSY (%)     14.1 |
|   -----  1   6   -----  -----  ----- |
|   -----  1   7   -----  -----  ----- |
|   CHAN1CNT   2   1   CHANNEL 1     SIO COUNT    14.1 |
|   CHAN2CNT   2   2   CHANNEL 2     SIO COUNT    14.1 |
\-----/

```

Figure 5-6. Extended Example Control Parameters - Screen 1

```

/----- Profile and Trending Analysis ----- ROW 1 OF 63 \
|Command ==>                               Scroll ==> CSR |
|Enter a ? in any data entry field for more information on valid values. |
|Composing CA MICS Inquiry: CHNPFT - Channel SIO Profile and Trending |
| |
|Report title ==> Extended Example |
| |
|Selection criteria: |
|      Input file ==> CHN - I/O Channel Example |
|      Start date ==> _____ (ddmonyyyy) |
|      SYSID ==> TS01 |
|      Zone ==> 1_____ |
|      No. months ==> ____ (1-9999) |
|Specify CAPAPU Values ==> N (Y/N) |
| |
|Forecast length (weeks) ==> 6 (1-999) |
|Forecast type ==> LONG (LONG/SHORT) |
| |
|  Name   Page  Col.   Col. Heading 1  Col. Heading 2  Output Format |
|  ----- -  -  -----  -----  ----- |
|  CHAN3CNT  2   3   CHANNEL 3      SIO COUNT      14.1 |
|  CHAN4CNT  2   4   CHANNEL 4      SIO COUNT      14.1 |
|  CHAN5CNT  2   5   CHANNEL 5      SIO COUNT      14.1 |
|  ----- 2   6   -----  -----  ----- |
|  ----- 2   7   -----  -----  ----- |
|  ----- 3   1   -----  -----  ----- |
|  ----- 3   2   -----  -----  ----- |
|  ----- 3   3   -----  -----  ----- |
|  ----- 3   4   -----  -----  ----- |
\-----/

```

Figure 5-6. Extended Example Control Parameters - Screen 2

5.6.1.2 Control Parameters

The Profile and Trending Analysis screen shown in Figure 5-6 shows the control parameters that are provided for Profile and Trending Analysis.

Report title: The title that appears on the reports.

Input file: The CHN resource element file that was originally defined in the extended example in Section 3.3.3.1 of this guide.

Start date	The start date for date range selection. Since neither a start date nor a number of days, weeks, or months are specified for this analysis, no date range selection is applied to the historical observations from the CHN resource element file.
SYSID:	Historical observations only for SYSID TS01.
Zone:	The summary level selection criteria for the analysis. Since the CHN resource element file was defined in Section 3.4.3.1 at the ZONE summary level, this parameter specifies that only historical observations for ZONE 1 are included in the analysis.
No. months:	The number of days, weeks, or months that will be included in the analysis.
Specify CAPAPU Values:	Specify Y or N. The CHN resource element file was not defined as containing the CAPAPU data element, therefore no CAPAPU value may be specified on this screen.
Forecast length:	The length of the forecast observation. Six is selected as a conservative number of estimates to be made from the 12 historical observations that are contained in the CHN resource element file.
Forecast type:	The type of forecast observation. The six future observations to be forecast are calculated based on the long-term rate of growth.
Name, Page, Col., Heading, Output Format:	These parameters define the contents of the Column reports. These parameters specify which and variables are tracked and provide column headings, report formats, and the page and column sequence of variables on the reports. Note that the five channel percent busy data elements are specified to appear on the first page of the report and that the five channel I/O count data elements are specified to appear on the second page.

5.6.1.3 Output Reports

Executing Profile and Trending Analysis using the control parameters on the Profile and Trending Analysis screen discussed in Section 5.6.1.2 produces three reports, which are discussed in the following sections:

- 1 - Channel Utilization Reports
- 2 - SIO Count Reports
- 3 - Estimates of Future Observations Report

5.6.1.3.1 Channel Utilization Reports

The Historical Observations and Growth Rate Reports for channel utilization are shown in Figure 5-7. By studying the detail of the historical observations and looking at the growth rates, you can conclude that the I/O load on the system is increasing slightly in the long term. The short-term growth rates show a very high degree of change.

The high variability in the short term, where channel 2 shows an increase of 47.08% in a single month, may indicate some instability in the workload. Looking at actual data, you can see that the load on channel 2 generally runs in the high teens, with an occasional jump into the low 20s, until the final observation, which shows a sudden increase to 30.9%.

A sudden, unexplained change is cause for concern when you are trying to forecast from this data since many forecasting methodologies place heavy emphasis on the most recent value of the observed data when estimating the forecast. If the additional load on channel 2 is due to a transient condition, the forecast values will be much too high. You should investigate the reason for the change in channel 2 utilization before you undertake a more detailed study.

These reports present historical data with the growth rates, allowing you to examine the behavior of the data before proceeding with the more detailed statistical analysis procedures provided in the remaining chapters. This gives you a chance to judge whether the data provides a good basis for further study.

+-----+-----+-----+-----+-----+-----+-----+-----+							
SYSID: TS01							
ZONE: 1		EXTENDED EXAMPLE --- MONTHLY HISTORICAL OBSERVATIONS					
+-----+-----+-----+-----+-----+-----+-----+-----+							
MONTH	CHANNEL 1	CHANNEL 2	CHANNEL 3	CHANNEL 4	CHANNEL 5		
	BUSY (%)	BUSY (%)	BUSY (%)	BUSY (%)	BUSY (%)		
+-----+-----+-----+-----+-----+-----+-----+-----+							
SEP 97	5.4	18.9	18.8	7.3	15.8		
OCT 97	5.5	19.5	19.6	7.4	14.8		
NOV 97	6.4	19.8	20.2	10.0	15.9		
DEC 97	5.3	18.0	18.6	8.9	14.6		
JAN 98	6.1	19.0	19.4	9.0	15.7		
FEB 98	6.3	23.1	18.8	11.1	16.2		
MAR 98	5.5	19.7	16.9	9.5	14.9		
APR 98	4.8	18.0	15.2	7.7	14.3		
MAY 98	4.6	17.8	15.7	8.2	14.5		
JUN 98	5.0	17.6	15.8	9.3	15.1		
JUL 98	6.9	21.0	18.6	11.7	17.8		
AUG 98	8.6	30.9	15.0	10.7	15.4		
+-----+-----+-----+-----+-----+-----+-----+-----+							
+-----+-----+-----+-----+-----+-----+-----+-----+							
SYSID: TS01							
ZONE: 1		EXTENDED EXAMPLE --- GROWTH RATE PER MONTH					
+-----+-----+-----+-----+-----+-----+-----+-----+							
	CHANNEL 1	CHANNEL 2	CHANNEL 3	CHANNEL 4	CHANNEL 5		
	BUSY (%)	BUSY (%)	BUSY (%)	BUSY (%)	BUSY (%)		
+-----+-----+-----+-----+-----+-----+-----+-----+							
INTERVAL							
+-----+-----+-----+-----+-----+-----+-----+-----+							
11 MOS	4.40 %	4.57 %	-2.02 %	3.63 %	-0.24 %		
LAST MO	23.94 %	47.08 %	-19.22 %	-8.37 %	-13.52 %		
+-----+-----+-----+-----+-----+-----+-----+-----+							

Figure 5-7. Extended Example Historical Observations and Growth Rate Reports: Channel Utilizations

5.6.1.3.2 SIO Count Reports

The Historical Observations and Growth Rate Reports for SIO counts are shown in Figure 5-8. Like the Channel Utilization Reports, these show that the I/O load is increasing slightly over the long term, but with a large amount of recent variability.

By comparing long- and short-term trends between related variables, you can often see changes in the fundamental behavior of the workload. The growth rates shown in the SIO Count Report are generally smaller than the corresponding growth rates in channel utilization. This may mean that the average amount of data transferred by an I/O is increasing over time. Does this indicate that larger block sizes are being used? A change in the character of the workload certainly affects the interpretation of the forecasts.

These reports are intended to provide simple sense of direction estimates for future values. These estimates are in no way statistically based forecasts, but rather provide estimates based on historical growth rates.

+-----+-----+-----+-----+-----+-----+-----+-----+							
SYSID: TS01							
ZONE: 1		EXTENDED EXAMPLE --- ESTIMATES OF FUTURE MONTHLY OBSERVATIONS					
+-----+-----+-----+-----+-----+-----+-----+-----+							
MONTH	CHANNEL 1 BUSY (%)	CHANNEL 2 BUSY (%)	CHANNEL 3 BUSY (%)	CHANNEL 4 BUSY (%)	CHANNEL 5 BUSY (%)		
+-----+-----+-----+-----+-----+-----+-----+-----+							
SEP 98	9.0	32.3	14.7	11.1	15.3		
OCT 88	9.4	33.8	14.4	11.5	15.3		
NOV 98	9.8	35.4	14.1	11.9	15.3		
DEC 98	10.2	37.0	13.9	12.4	15.2		
JAN 99	10.7	38.7	13.6	12.8	15.2		
FEB 99	11.1	40.4	13.3	13.3	15.2		
+-----+-----+-----+-----+-----+-----+-----+-----+							
+-----+-----+-----+-----+-----+-----+-----+-----+							
SYSID: TS01							
ZONE: 1		EXTENDED EXAMPLE --- ESTIMATES OF FUTURE MONTHLY OBSERVATIONS					
+-----+-----+-----+-----+-----+-----+-----+-----+							
MONTH	CHANNEL 1 SIO COUNT	CHANNEL 2 SIO COUNT	CHANNEL 3 SIO COUNT	CHANNEL 4 SIO COUNT	CHANNEL 5 SIO COUNT		
+-----+-----+-----+-----+-----+-----+-----+-----+							
SEP 98	14645337.2	17970160.8	14499094.8	7289359.2	18741264.3		
OCT 98	15083325.6	18507128.2	14310129.9	7723561.2	19466597.9		
NOV 98	15534412.7	19060140.8	14123627.7	8183627.2	20220003.6		
DEC 98	15998990.1	19629678.0	13939556.2	8671097.7	21002568.1		
JAN 99	16477461.3	20216233.5	13757883.6	9187605.2	21815419.7		
FEB 99	16970241.9	20820315.9	13578578.8	9734879.3	22659730.7		
+-----+-----+-----+-----+-----+-----+-----+-----+							

Figure 5-9. Extended Example Estimates of Future Observations Report

5.6.2 Performance Group/Period Workload Tracking

Profile and Trending Analysis was used to track the performance group/period workload data introduced in Section 3.3.3.2. Using the long-term rate of change, a three-month estimate of the arrival rates for each of the workloads is developed. Figure 5-10 shows the control parameters that are used to define the study. Figure 5-11 provides the Historical Observations, Growth Rate, and Future Estimates Reports.

```

/----- Profile and Trending Analysis ----- ROW 1 OF 63 \
|Command ==>                                     Scroll ==> CSR |
|Enter a ? in any data entry field for more information on valid values. |
|Composing CA MICS Inquiry: PRFPFT - Performance Group Profile and Trending |
| |
|Report title ==> PERFORMANCE GROUP/PERIOD WORKLOAD TRACKING |
| |
|Selection criteria: |
|   Input file ==> PRF - Performance Group Example |
|   Start date ==> _____ (ddmmyyyy) |
|   SYSID ==> TS01 |
|   Zone ==> 1_____ |
|   No. months ==> _____ (1-9999) |
|   Specify CAPAPU Values ==> N (Y/N) |
| |
|Forecast length (weeks) ==> 3 (1-999) |
|Forecast type ==> LONG (LONG/SHORT) |
| |
|   Name   Page  Col.   Col. Heading 1  Col. Heading 2  Output Format |
|-----|-----|-----|-----|-----|-----|
| WKL01    1    1    PERIOD 1          TS0          14.1 |
| WKL02    1    2    PERIOD 2          TS0          14.1 |
| WKL03    1    3    PERIOD 3          TS0          14.1 |
| WKL04    1    4    PERIOD 4          TS0          14.1 |
| WKL05    1    5    PERIOD 5          TS0          14.1 |
| WKL06    1    6    DEVELOPMENT       TS0          14.1 |
| WKL08    1    7    _____        OVERHEAD      14.1 |
| _____ 2    1    _____        _____        _____ |
| _____ 2    2    _____        _____        _____ |
\-----/

```

Figure 5-10. Performance Group/Period Workload Tracking

+-----+ SYSID: TS01 ZONE: 1 +-----+ PERFORMANCE GROUP/PERIOD WORKLOAD TRACKING --- WEEKLY HISTORICAL OBSERVATIONS +-----+							
WEEK ENDING	PERIOD 1 TSO	PERIOD 2 TSO	PERIOD 3 TSO	PERIOD 4 TSO	PERIOD 5 TSO	DEVELOPMENT TSO	OVERHEAD
24JUL1998	11102.5	1270.9	45.0	0.4	1.4	2565.9	1301.3
31JUL1998	11026.7	1251.8	49.3	0.6	0.8	2155.1	1352.6
07AUG1998	11232.2	1247.4	47.3	0.6	1.6	2148.7	1387.4
14AUG1998	10096.8	1201.9	45.5	0.3	1.1	2297.2	753.6
21AUG1998	10604.8	1361.2	66.9	0.1	0.6	2254.0	615.7
28AUG1998	10192.3	1226.0	41.5	0.4	0.8	2169.1	573.3
04SEP1998	11505.6	1230.3	63.8	0.1	1.8	2960.0	760.1
11SEP1998	8881.0	1263.0	50.0	0.4	3.6	1743.2	431.1
18SEP1998	11927.8	1123.4	38.8	0.2	1.7	2283.9	484.5
+-----+							
+-----+ SYSID: TS01 ZONE: 1 +-----+ PERFORMANCE GROUP/PERIOD WORKLOAD TRACKING --- GROWTH RATE PER WEEK +-----+							
INTERVAL	PERIOD 1 TSO	PERIOD 2 TSO	PERIOD 3 TSO	PERIOD 4 TSO	PERIOD 5 TSO	DEVELOPMENT TSO	OVERHEAD
07 WKS	1.03 %	-1.75 %	-2.10 %	-9.43 %	2.81 %	-1.65 %	-13.16 %
LAST WK	34.31 %	-11.05 %	-22.40 %	-50.00 %	-52.78 %	31.02 %	12.39 %
+-----+							
+-----+ SYSID: TS01 ZONE: 1 +-----+ PERFORMANCE GROUP/PERIOD WORKLOAD TRACKING --- ESTIMATES OF FUTURE WEEKLY OBSERVATIONS +-----+							
WEEK ENDING	PERIOD 1 TSO	PERIOD 2 TSO	PERIOD 3 TSO	PERIOD 4 TSO	PERIOD 5 TSO	DEVELOPMENT TSO	OVERHEAD
26SEP1998	12050.6	1103.8	38.0	0.2	1.7	2246.2	420.7
03OCT1998	12174.7	1084.5	37.2	0.2	1.8	2209.2	365.3
10OCT1998	12300.0	1065.5	36.4	0.1	1.8	2172.7	317.2
+-----+							

Figure 5-11. Performance Group/Period Workload Tracking Report

5.6.3 JES Spool

Profile and Trending Analysis was used to estimate three future monthly observations of the print load for the JES work stations introduced in Section 3.3.3.3. The long-term rate of change is used to develop these estimates. Figure 5-12 shows the control parameters used to define the study and Figure 5-13 provides the Historical Observations, Growth Rate, and Future Estimates Reports.

This case study shows how significantly the final monthly, weekly or daily value may influence growth rate estimates. As shown in Figure 5-13, the final monthly observations for all of the remote stations are significantly lower in September than in any previous month. This resulted in calculations of significant negative long- and short-term growth rates for each of these remotes.

```

/----- Profile and Trending Analysis ----- ROW 1 OF 63 \
|Command ==>                               Scroll ==> CSR |
|Enter a ? in any data entry field for more information on valid values. |
|Composing CA MICS Inquiry: JES PFT - JES Spool Profile and Trending |
| |
|Report title ==> JES WORKLOAD STUDY |
| |
|Selection criteria: |
|      Input file ==> JES - JES Spool Example |
|      Start date ==> _____ (ddmmyyyy) |
|      SYSID ==> SPOL |
|      Zone ==> A_____ |
|      No. months ==> _____ (1-9999) |
|      Specify CAPAPU Values ==> N (Y/N) |
| |
|Forecast length (weeks) ==> 3 (1-999) |
|Forecast type ==> LONG (LONG/SHORT) |
| |
|      Name      Page  Col.   Col. Heading 1  Col. Heading 2  Output Format |
|      -----  -    -    - |
|      RMTLOCLR__ 1     1     LOCAL             PRINT LINES     14._____|
|      RMT012LR__ 1     2     REMOTE 012         PRINT LINES     14._____|
|      RMT034LR__ 1     3     REMOTE 034         PRINT LINES     14._____|
|      RMT080LR__ 1     4     REMOTE 080         PRINT LINES     14._____|
|      RMT091LR__ 1     5     REMOTE 091         PRINT LINES     14._____|
|      RMT093LR__ 1     6     REMOTE 093         PRINT LINES     14._____|
|      RMT096LR__ 1     7     REMOTE 096         PRINT LINES     14._____|
|      _____ 2     1     _____         _____         _____|
|      _____ 2     2     _____         _____         _____|
\-----/

```

Figure 5-12. JES Workload Profile and Trending

+-----+ SYSID: SPOL ZONE: A -----+ JES WORKLOAD STUDY --- MONTHLY HISTORICAL OBSERVATIONS -----+ MONTH LOCAL REMOTE 012 REMOTE 034 REMOTE 080 REMOTE 091 REMOTE 093 REMOTE 096 PRINT LINES PRINT LINES PRINT LINES PRINT LINES PRINT LINES PRINT LINES PRINT LINES PRINT LINES -----+ MAY 98 170892378 1878785 257370 120771 17080144 9607193 3464558 JUN 98 181624570 1803782 440657 126215 13941631 9232379 2415886 JUL 98 178294112 1637214 418027 101766 15656930 10687353 1799042 AUG 98 183922396 1828908 541960 117295 7790766 8053527 392477 SEP 98 174622520 502184 62965 3468 2071539 1723890 109273 -----+ +-----+							
+-----+ SYSID: SPOL ZONE: A -----+ JES WORKLOAD STUDY --- GROWTH RATE PER MONTH -----+ INTERVAL LOCAL REMOTE 012 REMOTE 034 REMOTE 080 REMOTE 091 REMOTE 093 REMOTE 096 PRINT LINES PRINT LINES PRINT LINES PRINT LINES PRINT LINES PRINT LINES PRINT LINES PRINT LINES -----+ 04 MOS 0.54 % -28.10 % -29.67 % -58.83 % -40.99 % -34.92 % -57.86 % LAST MO -5.06 % -72.54 % -88.38 % -97.04 % -73.41 % -78.59 % -72.16 % -----+ +-----+							
+-----+ SYSID: SPOL ZONE: A -----+ JES WORKLOAD STUDY --- ESTIMATES OF FUTURE MONTHLY OBSERVATIONS -----+ MONTH LOCAL REMOTE 012 REMOTE 034 REMOTE 080 REMOTE 091 REMOTE 093 REMOTE 096 PRINT LINES PRINT LINES PRINT LINES PRINT LINES PRINT LINES PRINT LINES PRINT LINES PRINT LINES -----+ OCT 98 175567709 361085 44283 1428 1222485 1121987 46050 NOV 98 176518013 259631 31144 588 721430 730241 19406 DEC 98 177473461 186682 21903 242 425740 475274 8178 -----+ +-----+							

Figure 5-13. JES Workload Profile and Trending Report

Chapter 6: SIMPLE REGRESSION ANALYSIS

The Simple Regression Analysis software provides a convenient method for producing linear regression-based forecasts of CA MICS or user-computed data elements from capacity planning database files. This facility is designed to address two requirements:

- o Graphic presentations of historical data
- o Simple regression-based forecasts of future values
- o Optional Q&R formatted CSV files of forecast data

The plots can be produced either at the WEEKS, MONTHS, or DAYS timespan, depending on the specifications that you made in the File Definition Inquiry. The plots can either be formatted for a printer or for a color graphics device. Printer graphics are always available, however, SAS/GRAPH is a prerequisite for the color graphics option.

The Q&R formatted CSV files contain both the historical (actual) and forecasted values of the data element chosen for analysis. This data can be subsequently downloaded to a PC workstation and used for additional analysis and presentation using the CA MICS Q&R product.

Note: SAS/STAT is a prerequisite for using Simple Regression Analysis.

The following sections describe Simple Regression Analysis operation and its reports.

This section contains the following topics:

- [6.1 Functional Description](#) (see page 588)
- [6.2 Usage Guidelines](#) (see page 588)
- [6.3 Standard Output](#) (see page 589)
- [6.4 Analytic Technique Tutorial](#) (see page 592)
- [6.5 Component Operation](#) (see page 596)
- [6.6 Case Studies](#) (see page 603)

6.1 Functional Description

Simple Regression Analysis generates simple linear regression-based forecasts of the data elements that are contained in the Capacity Planning database files. Simple Regression Analysis produces plots, either for a line printer or for a color graphics device, that include historical data observations and the forecasts of future values. These forecasts are based on least squares linear regression models.

Simple Regression Analysis uses the SAS REG procedure to fit a least squares trend line to the historical data observations. You may then use the model thus developed to estimate future observations.

6.2 Usage Guidelines

The main factor you must assess for Simple Regression Analysis is the number of future intervals that you wish to estimate using the linear regression model. As with Profile and Trending Analysis, discussed in Chapter 5, the number of intervals that you forecast based on an historical data series is dependent on two characteristics of the historical data series:

- o The length (that is, the number of observations) of the historical data series
- o The behavior of the historical data series

The number of forecast intervals that you specify should never exceed the length of the historical data series used to produce the model. In fact, the probable error associated with each future value increases as a function of the number of the future values estimated. Quantitative techniques for determining these confidence limits for regression-based forecasts are discussed in Chapter 7, Univariate Model Forecasting. Since the probability of error increases with forecast length, we recommend a maximum forecast length of one-third to one-half the length of the historical data series.

You must also consider the qualitative behavior of the historical data series when you select the number of intervals to be forecast. If the differences between the regression line and the historical data series are small, then you can specify a longer forecast. If the differences are large (that is, greater than 10 to 15 percent of the actual value), then you should specify a very short period.

6.3 Standard Output

Simple Regression Analysis produces plots that track both the historical data and the least squares trend line over time. Forecasted future values of the trend line are represented. Using Simple Regression Analysis screens, you can control the number of forecast intervals, the axis titles, and the plot title.

Additionally, Q&R formatted CSV files can be optionally produced for later use by the CA MICS Q&R product. The optional CSV file creation is discussed in chapter 7.

Figure 6-1 is a sample simple regression forecast plot that shows the historical data, trend, and forecast for the SIO count on channel 1. The data is taken from the 4381- 2 I/O workload analysis case study, which was originally developed in Section 3.3.3.1. The contents of this graph are defined using the Simple Regression Analysis screen (see Figure 6-3).

The following is a description of the title information:

- o A user-specified title. In this example, EXTENDED EXAMPLE.
- o A second user-specified title. In this example, CHANNEL 1 SIO COUNT VS MONTH.
- o A title line generated by the program. This line contains the SYSID and ZONE that you selected for inclusion in the analysis.
- o A blank line.

- o The next three lines of the title are generated by the SAS PROC PLOT for printer plots. They are not included on color graphics plots. These lines are:
 - Designation of the historical observations
 - Designation of the regression values that correspond to the historical observation
 - Designation of the forecasted values
- o Y-axis title line. The Y-axis of the plot shows the range of values associated with the historical observations and the predicted values. You may specify the label for this axis.
- o X-axis title line. The X-axis of the plot shows the range of dates that correspond to the historical data and the predicted values. The program generates the label for this axis.

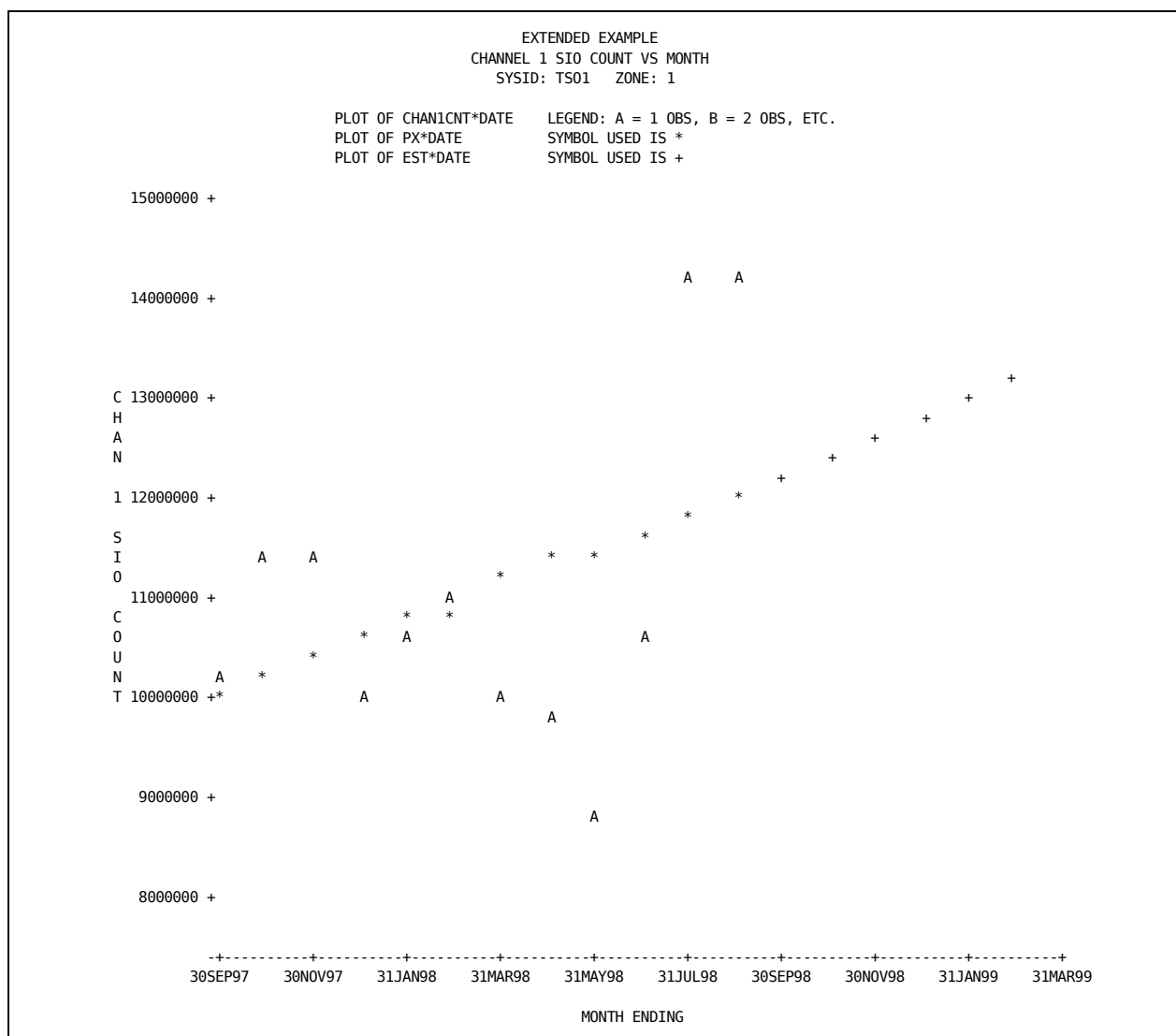


Figure 6-1. Simple Regression Forecast Plot

6.3.1 CSV Creation for CA MICS Q&R Reporting

A CSV (comma separated value) can be created for use by the CA MICS Q&R product by specifying "Y" on the Create CSV Output option of the query panel. This file can be downloaded and used by the CA MICS Q&R product to generate high-resolution graphs of the capacity planning data used in Simple Regression Analysis. The file will contain all of the Actual and Forecasted model data values used in the analysis. This provides the analyst a complete set of data to use for subsequent analysis and presentation with the CA MICS Q&R product.

6.4 Analytic Technique Tutorial

Historically, linear regression has been one of the most commonly applied techniques for forecasting computer requirements (ART79). Simple Regression Analysis uses linear regression to analyze a series of observations of user-selected variables from any of the files in the CA MICS database or user-defined variables that are computed from data elements contained in the CA MICS Database. This series of historical observations can be developed from either the WEEKS, MONTHS, or DAYS timespans. The historical series is analyzed using SAS PROC REG to develop a linear model that relates the user-specified variable to the independent variable time. This section provides an introduction to the regression analysis technique that the facility uses.

The fundamental concept of linear regression (MIL65) is that the curve of some dependent variable y is linear as a function of x . That is, for any given x , the mean distribution of y can be computed as:

$$y = b + m * x \quad (\text{Eqn 1})$$

where b is a constant that defines the mean value of y when x is zero

m is the slope of the line

The estimated y values exhibit differences from the actual y values that are used to produce the linear model. These differences are denoted by an error term, e . Therefore, we must expand Equation 1 to include an error term as shown in Equation 2:

$$y = b + m * x + e \quad (\text{Eqn 2})$$

The error term represents small random variations about the expected mean value of y . The value of e for any given observation depends on possible measurement errors and variables other than x which may influence y .

To give an example of linear regression, consider the following data which describes CPU utilization (% CPU) as a function of the number of active address spaces (ASIDs):

obs	1	2	3	4	5	6	7
ASIDs	10	15	20	25	30	35	40
% CPU	50	59	62	69	75	78	83

As the SAS plot in Figure 6-2 indicates, a linear relationship exists between percent CPU busy and the number of active address units for the data presented in the previous table.

Thus, if you hypothesize that a linear relationship exists between the number of active address spaces and percent CPU busy, you are faced with the problem of how to estimate the parameters m and b that are introduced in Equations 1 and 2. The statistical technique that we used to estimate these parameters is called the method of least squares. The method of least squares produces estimates of the parameters m and b such that the error term, e , for each of the observations is as small as possible.

Since SAS PROC REG is used to estimate the values of the parameters m and b , you need not understand the details of the least squares technique to use the software. However, a brief description of the least squares method is presented below to provide you with an overview of the technique.

The technique for estimating the values of m and b initially requires the computation of five parameters:

SX The sum of the x observations

SY The sum of the y observations

SSX The sum of the squares of the x observations

SXY The sum of x times y for all of the observations

n The number of observations

These parameters are used in the following equations to determine the values of m and b . The equations are called the "normal equations."

$$SY = b * n + m * SX \quad (\text{Eqn 3})$$

$$SXY = b * SX + m * SSX \quad (\text{Eqn 4})$$

The normal equations are a set of two linear equations which can be solved simultaneously to determine the values of the parameters m and b .

For the sample problem previously discussed, we determined the values of the SX , SY , SSX , SXY , and n to be:

$$\begin{array}{rclcl} SX & = & 175 & & SXY & = & 12,650 \\ SY & = & 476 & & n & = & 7 \\ SSX & = & 5,075 & & & & \end{array}$$

Substituting these values into the normal equations, Equations 3 and 4, we produce the following:

$$476 = 7b + 175m$$

$$12,650 = 175b + 5,075m$$

In solving these equations, you will find that $b = 41.21$ and $m = 1.07$. Using these values, you can expand the table of data that presents the actual CPU utilization and ASID count observations to include the estimated value of percent CPU busy (denoted by y_{est}) and the value of the error term, e .

obs	1	2	3	4	5	6	7
ASIDs	10	15	20	25	30	35	40
% CPU	50	59	62	69	75	78	83
y_{est}	52.0	57.4	62.7	68.1	73.4	78.8	84.1
e	-2.0	+1.6	-0.7	+0.9	+1.6	-0.8	-1.1

Although the regression of percent CPU busy as a function of number of active address spaces achieved good results, the same data analysis repeated with data from a different system could achieve much more suspect results. We alluded to the reason for this potential variation when we introduced the concept of the error term in an earlier paragraph. The error term represents potential measurement errors and/or influences on the y variable from independent variables other than x .

In the case of percent CPU busy, this analysis is also influenced by the CPU absorption rate of each of the address spaces. For the example presented above, the rates are apparently similar. However, for systems processing a wide variety of workload types, differences in CPU absorption rates for the individual address spaces could result in significant differences between the actual and estimated observations. Statistical estimators for determining the goodness of fit of regression models is discussed in Chapter 7, Univariate Model Forecasting.

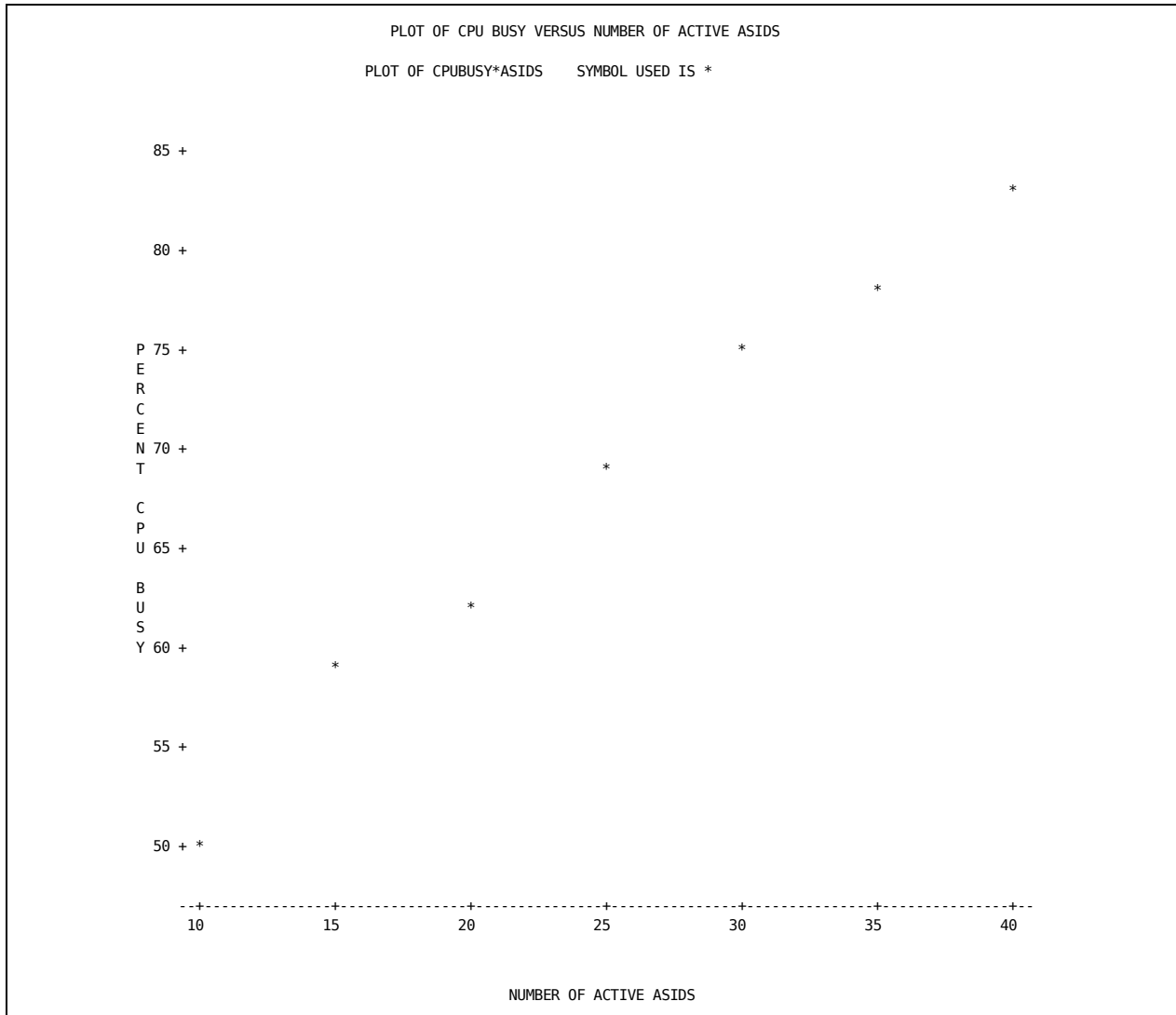


Figure 6-2. Percent CPU Busy vs. ASIDs

6.5 Component Operation

This section describes the control parameters for Simple Regression Analysis inquiries. These parameters allow you to define analyses based on simple linear regression, controlling titles, labels, and the presentation media. You can present your analysis on a line printer or on a color graphics device, however, SAS/GRAPH is a prerequisite for the color graphics option.

To get to the Simple Regression screen (shown in Figure 6-3), you must follow a short set of menus that are more fully described in Section 1.5 of this guide. The Simple Regression screen appears after you indicate that you want to insert a new Simple Regression Analysis inquiry into your catalog of analysis inquiries. Two fields on this screen are filled in on the fourth line when you first see it: inquiry name and the inquiry description, separated by a dash. You entered data in these fields on the Insert Capacity Planning Analysis screen when you requested that a new inquiry be added.

```

/----- Simple Regression Analysis -----\
| Command ==>                                     |
|                                                    |
| Composing CA MICS Inquiry: _____ - _____ |
|                                                    |
| Report title ==> _____ |
|                                                    |
| Selection criteria:                               |
|           Input file ==> _____ |
|           Start date ==> _____ (ddmnyyyy) |
|           SYSID ==> _____ |
|           Zone ==> _____ |
|           No. months ==> _____ (1-9999) |
|           Specify CAPAPU Values ==> _ (Y/N) |
|                                                    |
| Plot title ==> _____ |
| Forecast data element ==> _____ |
| Axis label ==> _____ |
| Forecast length (weeks) ==> _____ (1-9999) |
|                                                    |
| Plot type ==> PRINTER (COLOR/PRINTER) |
| Specify color graphics parms==> N (Y/N/R) |
| Color graphics format set ==> PENPLOT |
|                                                    |
|           Create CSV Output ==> N (Y/N) |
\-----/

```

Figure 6-3. Simple Regression Control Screen

Some of the parameters on the Simple Regression control screen shown in Figure 6-3 are required, while others are optional. A description of the parameters as they appear in this figure follows.

Report title: The 60-character title field for the Simple Regression Analysis reports. No double quotes or semicolons (;) may appear within the body of the title.

Input file: The capacity planning database file containing the historical observations to use as input to the analysis. If you are uncertain of the files or file names available in your capacity planning database, type a question mark (?) in the field and a selection list will appear.

Start date: The start date for date range selection. If Start date is specified, only historical observations at least as recent as this value are included in the analysis. The discussion of the following control parameter describes the options for controlling date range selection.

SYSID: The SYSID for which the report will be produced. If the Capacity Planning database file was defined as having a composite SYSID, this value is automatically supplied.

No. months: The maximum number of days, weeks, or months of historical observations that will be included in the analysis. Note that once you specify the input file, the field name will change to reflect the timespan of the input file. When used in conjunction with the Start date parameter, several options are available for date range selection control:

- o If you specify only Start date, all historical observations at least as recent as the value specified for Start date will be selected.

- o If you specify both Start date and No. months, the date range for data selection begins with the value specified for Start date and extends for the specified number of days, weeks or months.
- o If you specify only No. months, the date range for data selection will include only the number of the most recent days, weeks or months.
- o If you specify neither parameter, then all historical observations are selected.

Zone: The Zone, Hour, or Dayname to be included in the analysis, depending on whether you defined the capacity planning database file at the Zone, Hour, or Dayname summarization level. If you defined the file as having a composite Zone, Hour or Dayname, this value is supplied automatically on the screen. Note that the field name will be changed automatically to either Zone, Hour or Dayname once you specify the input file to reflect the input file summarization level.

Specify CAPAPU values: Put a Y in this field to specify the CAPAPU value(s) to be included in the analysis. The CAPAPU value is specified on the CAPAPU Selection Values screen, shown in Figure 6-4, that appears if you enter a Y in this field. You may specify it only if the Capacity Planning database file is defined as containing CAPAPU(s).

Plot Title: An optional second title line for the reports if needed. This is a 40-character field and no double quotes (") or semicolons (;) can appear within the body of the title.

Forecast data element: The data element will be the subject of the Simple Regression Analysis. If you are uncertain of the correct data element names in the input title, type a question mark (?) in this field and a selection list of available data elements will be presented.

Axis label:	The label for the vertical axis on the simple regression plot. You may specify up to 40 characters for the label and no double quotes (") or semicolons(;) can appear within the body of the label. If you do not specify a label, this parameter is automatically filled with the forecast data element label that you specified when the input file was created.
Forecast length (weeks):	The length of forecast period in days, weeks, or months. Note that the field name changes once you specify the input file to reflect the timespan of the Input file. If the file is a days level file, the Capacity Planner generates a forecast observation every seven days if the file uses DAYNAME summarization; otherwise an observation is generated for each day.
Plot type:	The specification for the type of plot. You can specify COLOR for a color graphics plot or PRINTER for a line printer plot. Printer graphics are always available, however, SAS/GRAPH is a prerequisite for color graphics. The default is a non-graphics printer. If you indicate that the results are to be presented on a color graphics device, you can also indicate whether you want to override the default color graphics format parameters. If you specify that you do, an additional screen appears to give you an opportunity to do this.
Create CSV Output:	Default value is "N" for NO. Specify "Y" for YES to create a Q&R formatted CSV that contains those data elements used in the analysis. Specify "N" for NO to bypass CSV creation.

If you entered a Y in the Specify CAPAPU values field on the Profile and Trending Analysis screen, the CAPAPU Selection Values screen appears, as shown in Figure 6-4. On the CAPAPU Selection Values screen, you can specify the CAPAPU values (up to nine) to be included in the analysis. You may specify CAPAPUs only if the Capacity Planning database file is defined as containing the CAPAPU data elements. (Note that Simple Regression Analysis does not support wild cards.)


```

/----- CAPAPU Selection Values -----\
|Command ==>
|
|Composing CA MICS Inquiry:  _____
|
|      Rule  ---+---1---+---2---+---3---+---4---+---5---+---6---
|CAPAPU  ==>  _____
|CAPAPU 2 ==>  _____
|CAPAPU 3 ==>  _____
|CAPAPU 4 ==>  _____
|CAPAPU 5 ==>  _____
|CAPAPU 6 ==>  _____
|CAPAPU 7 ==>  _____
|CAPAPU 8 ==>  _____
|CAPAPU 9 ==>  _____
|
|
\-----/

```

Figure 6-4. CAPAPU Selection Values Screen

```

/----- Color Graphics Format Parameters -----\
|Command ==>                                     Format set: PENPLOT
|
|Inquiry step: Simple Regression Analysis
|
|Title print font      ==> SWISSB      Title color      ==> BLUE
|Draw border around plot ==> N (Y/N) Background color  ==> _____
|Axis color           ==> BLACK      Area inside axis color ==> _____
|Text color           ==> BLACK
|
|Actual values:      Symbol ==> STAR      Line type ==> SOLID
|                   Color  ==> BLUE
|
|Regression values:  Symbol ==> PLUS      Line type ==> DASHED
|                   Color  ==> GREEN
|
|Forecast values:    Symbol ==> DIAMOND   Line type ==> DOTTED
|                   Color  ==> RED
|
|
|Symbol (NONE,PLUS,STAR,TRIANGLE,DIAMOND,HASH,SQUARE,A-Z,0-9,or SAS/GRAPH value)
|Line (SOLID,DASHED,DOTTED,1-32)
|Color (May be any color available on your graphics device.)
\-----/

```

Figure 6-5. Color Graphics Format Parameters Screen

The screen shown in Figure 6-5 allows you to define the color graphics parameters that will determine the appearance of your Simple Regression Analysis color plot. The following parameters are available:

Title print font: The character format (font) for the title line(s). Character font names are one to nine characters in length. Refer to the SAS/GRAPH Software Volume 1 for a list of valid character fonts.

Title color: The color of the title line(s).

Draw border around plot: The specification to frame the graph with a border (that is, draw a line around the outside of the display).

Background color: The specification to generate the graph on a colored background (that is, fill the entire graph with a color).

Axis color: The color of the x-axis and y-axis.

Area inside axis color: The specification to draw the plot lines on a colored background (only the area inside the axis is filled).

Text color: The color of all text words except the titles and labels.

Symbol: The symbol for plotting the data values. See the SAS/GRAPH Software Volume 1 for the available symbols.

Line type: The line type for plotting the data values. See the SAS/GRAPH Software Volume 1 for the available line types.

Color: The color for plotting the data values. See the SAS/GRAPH Software Volume 1 for the available colors.

6.6 Case Studies

This section presents three case studies that illustrate how to use the simple regression plots for resource forecasting. The first study, Extended Example I/O Workload Analysis, applies the technique to the analysis of the 9021-711 I/O workload that we originally introduced in Section 3.3.3.1. The second study, JES Spool, looks at trends in JES remote printer use. The final study, Application Tracking, shows an example of the simple regression technique used to track the CPU consumption of a critical application system.

The case studies are presented in the following sections:

- 1 - Extended Example I/O Workload Analysis
- 2 - JES Spool
- 3 - Application Tracking

6.6.1 Extended Example I/O Workload Analysis

To illustrate the use of the simple regression forecast plot, monthly data was collected for the channels on a 9021-711 to study the growth of the system's I/O workload. Each of the channels was characterized by average percent busy and the number of SIOs completed during the month.

The generation of the resource element file used in this study is discussed in Section 3.3.3.1. In this example, the following topics are addressed:

- 1 - Problem Description
- 2 - Control Parameters
- 3 - Output Report

6.6.1.1 Problem Description

The purpose of this study is to produce linear regression based forecasts of the future I/O load that would be experienced by the channels on a 9021-711. Twelve months of historical data was collected for channels 1 through 5 for the system SYSIDs TS01 and TS02 for ZONES 1 and 2.

6.6.1.2 Control Parameters

Figure 6-6 shows the screen containing the control parameters that are provided to Simple Regression Analysis.

```

/----- Simple Regression Analysis -----\
| Command ==>                                     |
|                                                    |
| Composing CA MICS Inquiry: CHNSMR - Simple Regression of Channel Util. |
|                                                    |
| Report title ==> EXTENDED EXAMPLE                 |
|                                                    |
| Selection criteria:                               |
|           Input file ==> CHN                     |
|           Start date ==> _____ (ddmonyyyy) |
|           SYSID ==> TS01                         |
|           Zone ==> _____                   |
|           No. months ==> _____ (1-9999)      |
|           Specify CAPAPU Values ==> _ (Y/N)       |
|                                                    |
| Plot title ==> CHANNEL 1SIO COUNT VS MONTH_____ |
| Forecast data element ==> CHAN1CNT               |
| Axis label ==> CHANNEL 1 SIO COUNT               |
| Forecast length (weeks) ==> 6 (1-9999)           |
|                                                    |
| Plot type ==> PRINTER (COLOR/PRINTER)            |
| Specify color graphics parms==> N (Y/N/R)        |
| Color graphics format set ==> PENPLOT            |
|                                                    |
\-----/

```

Figure 6-6. Extended Example Control Parameters

Report title: The title that appears on the resulting graph.

Input file: The CHN resource element file that was originally defined in the extended example in Section 3.3.3.1 of this guide.

Start date: The start date for date range selection. Since neither a start date nor a number of days, weeks or months are specified for this analysis, no date range selection is applied to the historical observations from the CHN resource element file.

SYSID: Historical observations only for SYSID TS01.

- Zone: The summary level selection criteria for the analysis. Since the CHN resource element file was defined in Section 3.3.3.1 at the ZONE summary level, this parameter specifies that only historical observations for ZONE 1 are included in the analysis.
- No. months: The maximum number of days, weeks, or months that will be included in the analysis.
- Specify CAPAPU values: Put a Y in this field to specify the CAPAPU value to be included in the analysis. The CAPAPU value is specified on the CAPAPU Selection Values screen, shown in Figure 6-4, that appears if you enter a Y in this field. You may specify it only if the capacity planning database file is defined as containing the CAPAPU data element.
- Plot Title: An optional second title line for the reports if needed. This is a 40-character field and no double quotes (") may appear within the body of the title.
- Forecast data element: The name of the data element (CHAN1CNT) in the capacity planning resource file which contains SIO counts for channel 1.
- Axis label: The label for the X-axis. Since the axis label was left blank, the online facility automatically filled it with the label of the data element, CHANNEL 1 SIO COUNT.
- Forecast length: The length of the forecast observation. Six is selected as a conservative number of estimates to be made from the 12 historical observations that are contained in the CHN resource element file.
- Plot type: The specification for the type of plot to be used. For this example, PRINTER is selected for a line printer plot.

Executing the analysis produces the plot report for the Channel 1 SIO counts shown in Section 6.6.1.3.

By modifying screen entries, then executing each modification, it is possible to quickly request forecasts of SIO counts and channel busy for each of the channels. These forecasts produce reports similar in format to the one shown, so they will not be presented here.

6.6.1.3 Output Report

Simple Regression Analysis produces the plot shown in Figure 6-7. This plot illustrates some of the problems you may encounter with historical data. (These types of data behavior problems are discussed in Section 6.2.) Although the historical observations from September 1997 through April 1998 are reasonably well behaved, the six monthly values recorded from May 1998 through August 1998 illustrate a great deal of variability. If the program had executed with a historical data series ending in May or June 1998, the resulting trend would have been negative. Therefore, you should be hesitant to base an extended forecast on this data series.

Investigation of the historical data might reveal some reasons for the variations in their values. For example, a week of data might have been lost in May or a new workload might have been added to the system starting in June. Univariate Model Forecasting, which is discussed in Chapter 7, allows you to apply the results of this type of investigation by deleting or adjusting historical data points.

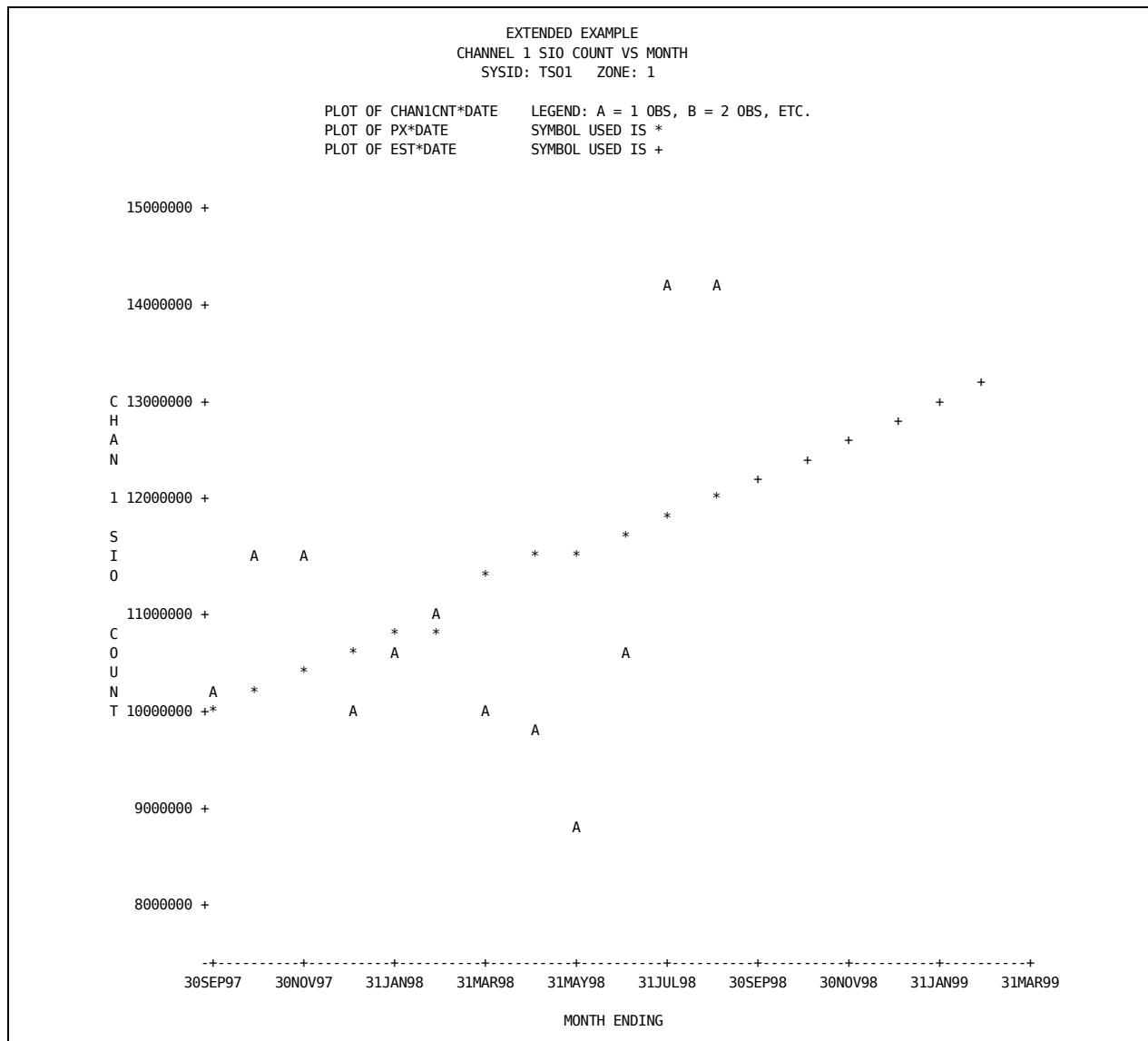


Figure 6-7. Extended Example Simple Regression Forecast Plot

6.6.2 JES Spool

Simple Regression Analysis was used to develop a four-month projection of the number of lines printed by local JES printers. Figure 6-8 shows the control parameters used for the study while Figure 6-9 shows the plot that was produced for the specified model.

```

/----- Simple Regression Analysis -----\
| Command ==>
|
| Composing CA MICS Inquiry: JESSMR - Simple Regression of Local Print Load
|
| Report title ==> LOCAL PRINT WORKLOAD
|
| Selection criteria:
|       Input file ==> JES
|       Start date ==> _____ (ddmonyyyy)
|       SYSID ==> SPOL
|       Zone ==> _____
|       No. weeks ==> _____ (1-9999)
|       Specify CAPPU Values ==> _ (Y/N)
|
| Plot title ==> _____
| Forecast data element ==> RMTLOCLR
| Axis label ==> PRINT LINES
| Forecast length (weeks) ==> 4 (1-9999)
|
| Plot type ==> PRINTER (COLOR/PRINTER)
| Specify color graphics parms==> N (Y/N/R)
| Color graphics format set ==> PENPLOT
|
\-----/

```

Figure 6-8. JES Workload Control Parameters

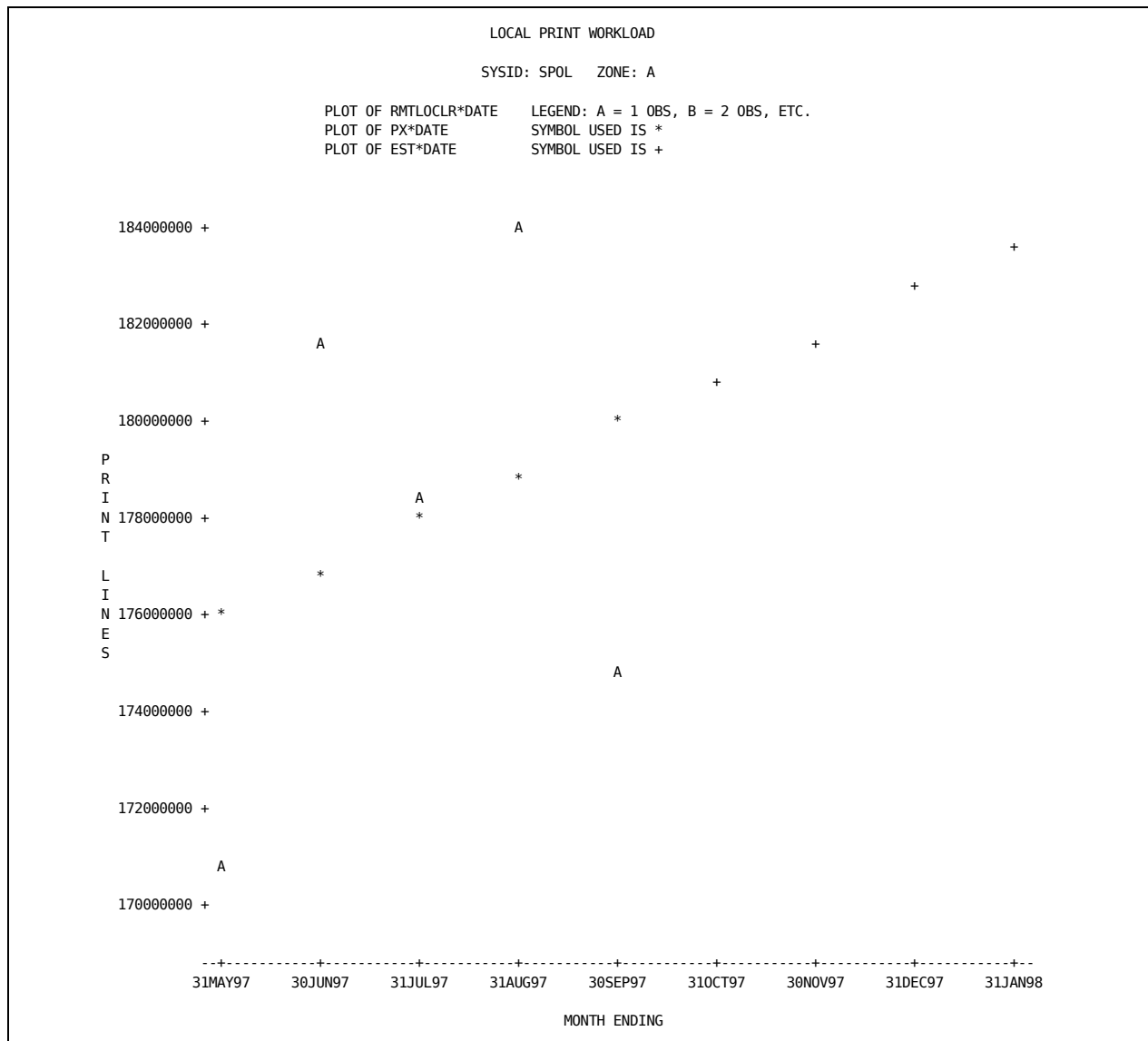


Figure 6-9. Simple Regression Model of Local Print Lines

6.6.3 Application Tracking

Simple Regression Analysis was used to develop a four-month, time-based linear regression model of the CPU time consumed by the payroll application. Figure 6-10 provides the control parameters that we used for the study and Figure 6-11 shows the plot that was produced for the specified model.

```

/----- Simple Regression Analysis -----\
| Command ==>
|
| Composing CA MICS Inquiry: PAYSMR - Simple Regression of Payroll Appl.
|
| Report title ==> Payroll CPU requirements
|
| Selection criteria:
|       Input file ==> PAY
|       Start date ==> _____ (ddmonyyyy)
|       SYSID ==> TSOA
|       Zone ==> _____
|       No. weeks ==> _____ (1-9999)
|       Specify CAPAPU Values ==> _ (Y/N)
|
| Plot title ==> _____
| Forecast data element ==> JOBTCBTM
| Axis label ==> CPU TIME
| Forecast length (weeks) ==> 4 (1-9999)
|
| Plot type ==> PRINTER (COLOR/PRINTER)
| Specify color graphics parms==> N (Y/N/R)
| Color graphics format set ==> PENPLOT
|
\-----/

```

Figure 6-10. Application Tracking Control Parameters

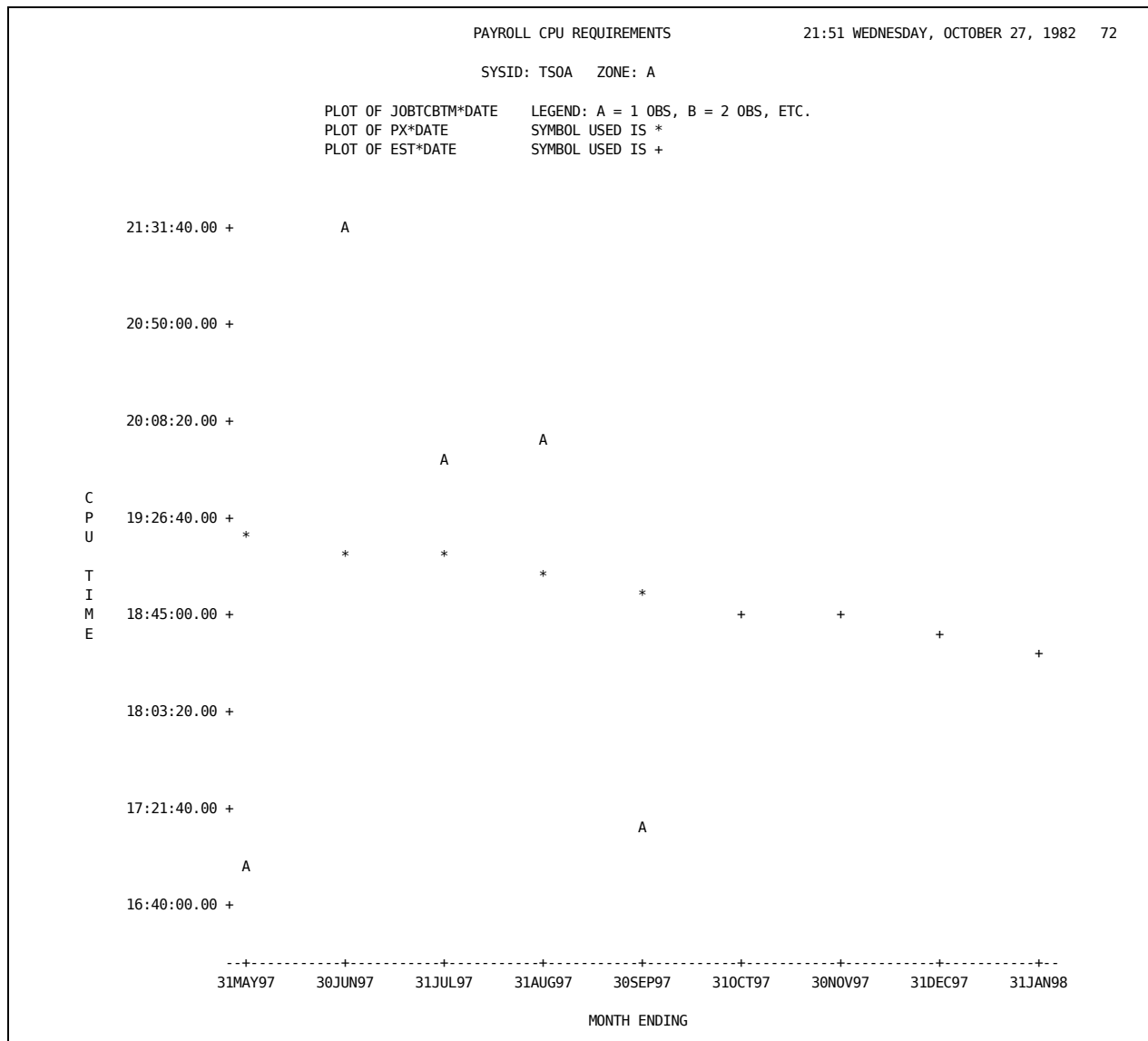


Figure 6-11. Simple Regression Payroll CPU Time Forecast

Chapter 7: UNIVARIATE MODEL FORECASTING

The Univariate Model Forecasting software allows you to investigate the behavior of historical data series and forecast future observations with a univariate model. This facility is designed to address three requirements:

- o Fitting linear, quadratic, and cubic models to historical data series.
- o Smoothing historical data series to minimize problems introduced by random variations in the historical data observations.
- o Allowing you to modify and/or delete historical data points to deal with errant data.

Note: SAS/STAT is a prerequisite for using Univariate Model Forecasting.

This section contains the following topics:

[7.1 Functional Description](#) (see page 613)

[7.2 Usage Guidelines](#) (see page 614)

[7.3 Standard Output](#) (see page 617)

[7.4 Analytic Technique Tutorial](#) (see page 624)

[7.5 Component Operation](#) (see page 649)

[7.6 Case Studies](#) (see page 664)

7.1 Functional Description

Univariate Model Forecasting allows you to develop forecasts based on regression using linear, quadratic, or cubic models of data elements contained in the Capacity Planning database files. These files can be either resource element files based on data from one or more CA MICS files, meta files or business element files based on data obtained from any other source (see Chapter 3).

These forecasts are based on "least squares" regression models. Univariate modeling uses the SAS REG procedure to fit a line of least squares to the historical data. The model developed for the historical data is then used to estimate future observations for the number of periods which you specify. Optionally, it can store the forecast values in special Capacity Planning forecast files.

Univariate Model Forecasting offers several additional features to assist you in fitting a least squares line to the data. Computer system data often exhibits a high degree of variability, and is sometimes a non-linear function. Both high variability and functional form make fitting a least squares line very difficult.

To overcome problems with variability and functional form, Univariate Model Forecasting offers you smoothing functions based on geometric moving averages, logarithmic transformation, and user-defined transformations. You can use geometric moving averages to dampen a series which exhibits a high degree of variability (both positive and negative) around a central trend. Logarithmic transformations allow you to analyze exponential relationships using linear models. You can also code a user exit to define any other type of transformation which may apply to your installation. An example of a user-defined transformation is the normalization of CPU time to the speed of a base CPU.

7.2 Usage Guidelines

This section discusses three methods for evaluating the fit of regression models:

- o The coefficient of determination. This value, which is commonly called "r-squared," indicates the percent of the variability in the historical data that is explained by the model.
- o The F value for testing that all of the parameters of the model are zero except for the intercept.

- o The development of confidence limits for the model. Confidence limits which are based on the standard error for the model allow you to determine confidence bounds (that is, upper and lower limits) for the forecasts that are produced with a model.

Each of these methods is based on statistical tests of certainty using the regression model. A tutorial on regression models is presented in Section 7.4. If you are not familiar with regression models, we recommend that you review the tutorial in conjunction with this section on usage guidelines.

COEFFICIENT OF DETERMINATION

Perhaps the best technique for evaluating the fit of a proposed model is the value of r-squared. This parameter, called the coefficient of determination, is calculated using the following equation:

$$r^2 = \frac{b * SY + m * SXY - n * \bar{Y} ** 2}{SXY - n * \bar{Y} ** 2} \quad (\text{Eqn 1})$$

Basically, r-squared indicates the percentage of the variability in the historical data that is explained by the model (SAS82). If the model is evaluated and the resulting r-squared is 0.98, the proposed model explains 98% of the variability of the historical data. For most forecasting applications, an r-squared value of 0.70 or greater is more than suitable for evaluating models (SAR79).

NULL HYPOTHESIS

The second method for evaluating models is to test the "null hypothesis" using the F value for the model. The null hypothesis states that the historical data is as well represented by the mean value (that is, the Y intercept) as by the proposed model. The F value is computed using the ratio of the sum of the squares values calculated for the model. (Details of this calculation are beyond the scope of this tutorial.) You may compare the F value to standard statistical tables to determine the probability that the proposed model would have arisen under the null hypothesis. This probability is given by the value of p.

A "good model" is indicated by a large F value and a very small p probability. Although most statistics texts recommend testing the null hypothesis to a level of 0.001, a value of 0.01, investigators studying short series of computer measurement data (SAR79) suggest a value of 0.01. An F statistic can be misleading when there are too few degrees of freedom (that is, less than 15 or 20). Therefore, the F statistic should not be your primary consideration in evaluating models developed for short series of historical observations.

CALCULATION OF CONFIDENCE LIMITS

The third method for evaluating regression models is the calculation of confidence limits. If the proposed linear model does not fit the historical data perfectly, you can expect that errors will also be present in the values predicted by the model. You can bound these errors by confidence limits. You can determine the average confidence limit for a regression model using the standard error value calculated for the model. Equation 2 shows the equation for calculating the standard error.

$$s_e = \sqrt{\frac{SSY - b * SY - m * SXY}{n-2}} \quad ** \ 1/2 \quad (\text{Eqn 2})$$

By solving the equation, you can determine the standard error, and using this value, you can estimate the confidence limits for the forecast. For example, if a 99% confidence limit is calculated for a forecast, then you can be 99% sure that all future values will fall between the upper and lower confidence limits.

For large samples (that is, n greater than 30), a 68% confidence limit is approximated by plus or minus one times the standard error; a 95.5% confidence limit is approximated by two times the standard error; and a 99.7% confidence limit (that is, 99.7% of actual future values should fall within plus or minus three standard errors of their predicted value) is approximated by three times the standard error. For smaller samples, the number of standard errors required to approximate the confidence limits is determined using a "t statistic". (The details of this calculation are beyond the scope of this tutorial.) Note that the width of the confidence interval (that is, the difference between the exact upper and lower confidence limits) increases with the length of the forecast. That is, the width of the confidence limit increases monotonically with the length of the forecast.

The SAS REG procedure calculates an exact 95% confidence limit for each point in the historical data series and for each point estimated by the model (SAS81). These calculations are based on the standard error and are similar in nature to the equations presented in this section. (The details of these calculations are beyond the scope of this tutorial.) Confidence limits allow you to evaluate the range of potential errors that may exist in the forecasts produced by the program.

7.3 Standard Output

Univariate Model Forecasting produces analysis reports and, optionally, an update to a Capacity Planning forecast file.

The Model Analysis Report gives a summary of the regression analysis. This report includes key regression statistics, such as the r -squared, F , and p . It also includes estimates for the coefficients of the least squares line.

The Residual Analysis Report analyzes the residuals which occur when the model is fit to the historical data, and shows the future predicted values of the series. This report includes a printed graph of the residuals to allow for quick visual judgment of the fit of the model.

Optionally, Q&R formatted CSV files can be created that contain both the historical (actual) and forecasted values of the data element chosen for analysis. This data can be subsequently downloaded to a PC workstation and used for additional analysis and presentation using the CA MICS Q&R product.

After the regression model is executed, you can optionally store the predicted values of the model's dependent variable in a capacity planning forecast file. This allows you to use the presentation graphics programs described in Chapter 4 to develop a graph of both the historical and forecast data.

The reports are presented in the following order:

- 1 - Model Analysis Report
- 2 - Residual Analysis Report
- 3 - R-Square Ranking Report
- 4 - Updating the Capacity Planning Database
- 5 - CSV Creation for CA MICS Q&R Reporting

7.3.1 Model Analysis Report

The Model Analysis Report provides a summary of the regression analysis that was conducted for the particular model that you specified. Figure 7-1 illustrates the Model Analysis Report for a study that relates channel SIO counts to the passage of time. You can use this report to evaluate whether or not the model produced by Univariate Model Forecasting acceptably represents the historical data series. This decision is based on the r-squared and F values discussed in Section 7.2.

CA MICS Capacity Planner							
ANALYSIS OF UNIVARIATE FORECASTING MODEL							
MODEL OF: CHAN5CNT							
BASED ON: LINEAR							
				-----TIME ELEMENTS-----			
R**2	F	P	INTERCEPT	NAME	COEFFICIENT	T	P
0.75	30.84	.0004	9736797	LINEAR	587938	5.55	.0004

Figure 7-1. Model Analysis Report

The Model Analysis Report shown in Figure 7-1 presents the following information:

MODEL OF: The data element name to be forecast.

BASED ON: The time elements (linear, quadratic and/or cubic) that you used to develop the model.

The first four columns of the report are discussed below:

R**2: The r-squared value for the model. We recommend R-squared values of 0.70 and above for accepting models that are produced using Univariate Model Forecasting.

F: The F statistic for the model. As we discussed in Section 7.2, the F statistic is calculated from the ratio of sums of squares of the model. Basically, larger F values indicate a more reliable model. The F statistics may be unreliable for models developed with a small number of points (that is, less than 30).

P: The probability that the model proposed is significant (that is, more reliable than simply forecasting the average value of the historical observations). Although most statistics texts recommend testing the null hypothesis to a 0.001 level, 0.01 has been recommended for models based on computer measurement data (SAR79).

INTERCEPT: The mean value of y when x is zero.

The next four columns are produced for each time variable that is used in the model. (For example, if a model specified that CPUHRS is to be forecast based upon linear, quadratic and cubic, time elements, then three lines would appear under the heading TIME ELEMENTS.) These columns contain the following information:

NAME: The name of the time element to be analyzed. This name can be either LINEAR, QUADRATIC, or CUBIC.

COEFFICIENT: The coefficient for the time element in the regression equation.

T: The T statistic. The T statistic is similar in nature to the F statistic. Larger values indicate "goodness." Note that the value may be misleading for models that are based on small samples.

P: The probability that the time element you specified makes a significant contribution to the model to be developed. Once again, we recommend a 0.01 acceptance value.

7.3.2 Residual Analysis Report

The Residual Analysis Report shows how the model proposed by the analyst fits the historical data series. Figure 7-2 shows an example of a Residual Analysis Report for the channel SIO count.

CA MICS Capacity Planner								
UNIVARIATE MODELING RESIDUAL ANALYSIS								
UNIVARIATE FORECAST OF: CHAN5CNT								
BASED ON TIME ELEMENTS: LINEAR								
DATE	CHAN5CNT	TRANSFORMED* CHAN5CNT	PREDICTED CHAN5CNT	RESIDUAL CHAN5CNT	-----PLOT OF RESIDUALS-----			CONFIDENCE
					-195E4	0.0	+195E4	LIMITS +/-
30SEP97	11882418	.	10324734	1557683.74		0+++++		1624240.54
31OCT97	12860633	12371525.50	10912672	1947961.05		0+++++		1416789.90
28NOV97	11743581	12302107.00	11500610	242971.35		0++		1221119.32
31DEC97	12510012	12126796.50	12088547	421464.66		0++++		1043874.36
30JAN98	12071368	12290690.00	12676485	-605117.03		-----0		896056.61
27FEB98	12076524	12073946.00	13264423	-1187899		-----0		794269.29
31MAR98	12186976	12131750.00	13852360	-1665384		-----0		757306.06
30APR98	14355185	13271080.50	14440298	-85113.10		0		794269.29
29MAY98	15011392	14683288.50	15028236	-16843.79		0		896056.61
30JUN98	17176448	16093920.00	15616173	1560274.52		0+++++		1043874.36
31JUL98	17421158	17298803.00	16204111	1217046.83		0+++++		1221119.32
31AUG98	18042957	17732057.50	16792049	1250908.14		0+++++		1416789.90
31SEP98	.	.	17379987	.		0		1624240.54
31OCT98	.	.	17967924	.		0		1839490.02
31NOV98	.	.	18555862	.		0		2060095.20
31DEC98	.	.	19143800	.		0		2284505.07
31JAN99	.	.	19731737	.		0		2511700.04
31FEB99	.	.	20319675	.		0		2740987.65
31MAR99	.	.	20907613	.		0		2971883.57
31APR99	.	.	21495550	.		0		3204040.12

* - DATA SMOOTHED USING A GEOMETRIC MOVING AVERAGE. ALPHA = 0.50

Figure 7-2. Residual Analysis Report

The Residual Analysis Report shown in Figure 7-2 presents the following information:

UNIVARIATE

FORECAST OF: The name of the data element to be forecast.

BASED ON

TIME The time elements (LINEAR, QUADRATIC, and/or
ELEMENTS: CUBIC) that you are using to develop the model.

The columns of the report are discussed below:

DATE: The ending date of the week or month that is
 represented by the observation.

CHAN5CNT: The actual historical observations for the data
 element being modeled (in this case, CHAN5CNT).
 If the last two characters of the the data
 element name are TM, the facility assumes that
 the value is in seconds and uses a SAS TIME10.1
 format for the report.

TRANSFORMED The transformed value. If you do not specify
(CHAN5CNT): that historical observations are to be
 transformed by one of the options available on
 the Transform Historical Observations screen
 shown in Figure 7-12, the TRANSFORMED value will
 be identical to the historical value. If you do
 specify that historical observations are to be
 transformed, then a footnote is generated to
 indicate the type of processing that is
 performed on the historical data series. The
 presence of a footnote is indicated by an
 asterisk (*) following the word TRANSFORMED. If
 the last two characters of the the data element
 name are TM, the facility assumes that the value
 is in seconds and uses a SAS TIME10.1 format for
 the report.

PREDICTED The y value predicted for the interval using the
(CHAN5CNT): regression equation fit to the historical data.
 If the last two characters of the data element
 name are TM, the facility assumes that the value
 is in seconds and uses a SAS TIME10.1 format for
 the report.

RESIDUAL (CHAN5CNT):	The error term, e, in the regression equation. The residual is equal to the difference between the actual historical observation and the predicted value. The residual value for a point deleted from the modeling process is of no concern, since the model was not developed to account for the observation. If the last two characters of the data element name are TM, the facility assumes that the value is in seconds and uses a SAS TIME10.1 format for the report.
PLOT OF RESIDUALS:	The residual values that are plotted over the range from minus to plus the maximum residual value.
CONFIDENCE LIMITS:	The 95% confidence limits for the forecast. As discussed in Section 7.2, the confidence limits bound the potential error that may exist in the forecasts produced by the model. If the last two characters of the the data element name are TM, the facility assumes that the value is in seconds and uses a SAS TIME10.1 format for the report.

7.3.3 R-Square Ranking Report

The R-Square Ranking report provides you with an overview of the results of all of the forecasts performed during this forecasting run. The report is presented in ascending R-Square sequence and gives the r-square, f-statistic, and p-values for the statistical model generated.

If either the r-square or p-value for the run exceeds the thresholds set for the report, the Warning Flag column will be set to show the which threshold was exceeded. If the flag is set to RS-WARN then the r-square for this particular forecast was below the threshold; if the flag is set to P-WARN then one of the p-values for one of the variables in the model was exceeded. If both thresholds are exceeded then the RS-WARN takes precedence.

The other fields on this report correspond to the equivalent values found on the other reports associated with the forecasting process.

```

CA MICS Capacity Planner
R-SQUARE RANKING REPORT
Testing MVR

MODEL OF: CHAN5CNT
BASED ON: LINEAR

WARNING  R**2    P      SYSID
FLAG
-----  -
0.75    .0004   S008

R-SQUARE WARNING VALUE : 0.50  P-VALUE WARNING THRESHOLD : 0.05

```

Chapter 7: UNIVARIATE MODEL FORECASTING 623

In the example shown in Figure 7-2, data for September 1997 through August 1998 is available in the historical file. After performing Univariate Model Forecasting, the corresponding forecast file contains observations from September 1998 through April 1999.

You can use Presentation Graphics, which is described in Chapter 4, to present both the historical input to Univariate Model Forecasting and the results of forecasting studies such as these.

7.3.5 CSV Creation for CA MICS Q&R Reporting

A CSV (comma separated value) can be created for use by the CA MICS Q&R product by specifying "Y" on the Create CSV Output option of the query panel. This file can be downloaded and used by the CA MICS Q&R product to generate high-resolution graphs of the capacity planning data used in Simple Regression Analysis. The file will contain all of the Actual and Forecasted model data values used in the analysis. This provides the analyst a complete set of data to use for subsequent analysis and presentation with the CA MICS Q&R product.

7.4 Analytic Technique Tutorial

Regression is one of the most popular techniques used to forecast computer requirements (ART79). Univariate Model Forecasting uses regression to analyze historical observations from either of two sources:

- o Capacity planning resource element files based on data from any of the files in the CA MICS database
- o Capacity planning business element files

You can develop a series of historical observations in a resource element file either from the DAYS, WEEKS, or MONTHS timespans in the CA MICS database. You can extract each data element in such a resource file directly from a CA MICS file or calculate it from one or more CA MICS data elements obtained from one or more CA MICS files.

The historical series in a business element file can represent various volume indicators related to your organization's business or any other non-CA MICS data source.

Either type of historical series is analyzed during Univariate Model Forecasting by using SAS PROC REG (SAS81) to develop models (linear, quadratic, or cubic) that relate user-specified variables to time as the independent variable. Thus, Univariate Model Forecasting is based upon the hypothesis that there is a trend in the historical series of data that can aid in predicting the future.

SIMPLE REGRESSION

To review the concepts of simple regression (MIL65), hypothesize that the curve of some dependent variable y is linear as a function of x . That is, for any given x , the mean distribution of the y s can be computed as

$$y = b + m * x \quad (\text{Eqn 3})$$

where b is a constant that defines the mean value of y
when x is zero

and m is the slope of the line

The estimated y values exhibit differences from the actual y values that are used to produce the linear model. These differences are denoted by an error term, e . Therefore, we can expand Equation 3 to include an error term as shown in Equation 4:

$$y = b + m * x + e \quad (\text{Eqn 4})$$

Basically, the error term represents small random variations about the expected mean value of y . The value of e for any given observation depends on possible measurement errors and variables other than x that may influence y .

To see an example of linear regression, consider the following data, which describes CPU utilization (% CPU) as a function of the number of active address spaces (ASIDs):

obs		1	2	3	4	5	6	7
-----		-----	-----	-----	-----	-----	-----	-----
ASIDs		10	15	20	25	30	35	40
-----		-----	-----	-----	-----	-----	-----	-----
% CPU		50	59	62	69	75	78	83

As the SAS plot shown in Figure 6-2 indicates, a linear relationship exists between percent CPU busy and the number of active address spaces for the data presented in the previous table.

METHOD OF LEAST SQUARES

Thus, if you hypothesize that a linear relationship exists between the number of active address spaces and % CPU busy, you have the problem of how to estimate the parameters m and b , which we introduced previously in Equations 3 and 4. The statistical technique that we used to estimate these parameters is called the method of least squares. The method of least squares produces estimates of the parameters m and b such that the error term, e , for each of the observations is as small as possible.

Since you can use SAS PROC REG to estimate the values of the parameters m and b , you need not understand the details of the least squares technique to use the facility. However, a brief description of the least squares method is presented below to provide you with an overview of the technique. The technique for estimating the value of m and b initially requires computing five parameters:

- SX: The sum of the x observations
- SY: The sum of the y observations
- SSX: The sum of the squares of the x observations
- SXY: The sum of x times y for all of the observations
- n : The number of observations

The next two parameters are not required to calculate the values of m and b , however they are used to calculate two other values that will be used later to evaluate the fit of the model. These values are the following:

SSY The sum of the squares of the y observations

$\bar{Y}$ The average of the y observations

The first set of parameters are used in the following equations to determine the values of m and b . These equations are called the "normal equations":

$$SY = b * n + m * SX \quad (\text{Eqn 5})$$

$$SXY = b * SX + m * SSX \quad (\text{Eqn 6})$$

The normal equations are a set of two linear equations that can be solved simultaneously to determine the values of the parameters m and b .

For the sample problem previously discussed, we determined the values of SX , SY , SSX , SXY , and n to be

$$SX = 175 \quad SXY = 12,650$$

$$SY = 476 \quad \bar{Y} = 68$$

$$SSX = 5,075 \quad n = 7$$

$$SSY = 33,184$$

Substituting these values into the normal equations (Equations 5 and 6) produces the following:

$$476 = 7b + 175m$$

$$12,650 = 175b + 5,075m$$

In solving these equations, you find that $b = 41.21$ and that $m = 1.07$. Using these values, you can expand the table of data that presented the actual CPU utilization and ASID count observations to include the estimated value of percent CPU busy (denoted by y_{est}) and the value of the error term, e .

obs	1	2	3	4	5	6	7
ASIDs	10	15	20	25	30	35	40
% CPU	50	59	62	69	75	78	83
y_{est}	52.0	57.4	62.7	68.1	73.4	78.8	84.1
e	-2.0	+1.6	-0.7	+0.9	+1.6	-0.8	-1.1

Although this regression of percent CPU busy as a function of the number of active address spaces achieves good results, the same data analysis repeated with data from a different system could achieve much more suspect results. We alluded to the reason for this potential variation when we introduced the concept of the error term in an earlier paragraph. The error term represents potential measurement errors and/or influences on the y variable from other independent variables than x .

In the case of percent CPU busy, this analysis is also influenced by the CPU absorption rate of each of the address spaces. For the example presented above, the rates are similar among the various address spaces. However, for systems processing a wide variety of workload types, differences in CPU absorption rates for the individual address spaces could result in significant differences between the actual and estimated observations.

The following sections discuss statistical estimators for evaluating the fit of regression models:

- 1 - Linear, Quadratic, and Cubic Models
- 2 - Processing Historical Observations

7.4.1 Linear, Quadratic, and Cubic Models

Univariate Model Forecasting allows you to evaluate models that involve linear, quadratic, and cubic functions of time. The characteristics of these models are as follows:

- o A linear model proposes that the data element can be modeled based on the interval number (that is, time period). Linear models are considered to be the most reliable for the short series of historical observations that you normally encounter. The general form of a linear model is shown in the following equation:

$$y = m * x + b \quad (\text{Eqn 7})$$

where y is the data element, x is the interval number, m is the coefficient of x (that is, the slope), and b is the intercept.

- o A quadratic model proposes that the data element can be modeled based on the interval number squared. The general form of the quadratic model is shown in the following equation.

$$y = m1 * x^2 + m2 * x + b \quad (\text{Eqn 8})$$

where y is the data element, x is the interval number, m1 and m2 are the coefficients of the interval numbers, and b is the intercept.

- o A cubic model proposes that the data element can be modeled based on the interval number cubed. The general form of the cubic model is shown in the following equation.

$$y = m1 * x^3 + m2 * x^2 + m3 * x + b \quad (\text{Eqn 9})$$

where y is the data element, x is the interval number, m1, m2, and m3 are the coefficients of the interval numbers, and b is the intercept.

You should avoid higher order models (that is, quadratic and cubic) for historical data series that include less than 30 historical observations because these models are often unreliable when they are based on a small number of observations. With shorter series, these models have a propensity to select coefficients for their equations that almost ideally represent the historical data points but rapidly diverge beyond the range of the historical observations.

7.4.2 Processing Historical Observations

Often the series of historical observations present in a capacity planning database file poses a number of problems to the analyst attempting to develop forecasting models.

Among these problems are the following:

- o The presence of errant observations in the historical data series. These observations result from holidays, missing data, and other anomalies.
- o Apparent random variations around a central trend. This type of variation is common in random arrival workloads, such as testing or in a service bureau's workload.
- o Exponential trends associated with service level measurements. Rather than being linear, the trends associated with service measurements are often exponential since they result from an underlying queuing relationship.
- o Step functions and other changes introduced by the addition of new workloads to the system.

The following sections discuss approaches to these problems:

- 1 - Deleting Errant Observations
- 2 - Data Smoothing Using a Geometric Moving Average
- 3 - Data Smoothing Using a LOG10 Function
- 4 - User-Defined Data Transformations

7.4.2.1 Deleting Errant Observations

The following table shows average CPU utilizations collected for a 3090-200J processor, and illustrates the problems that often occur with historical observations:

Week Ending	Observation Number	% CPU BUSY
=====	=====	=====
31OCT97	1	71.0
07NOV97	2	72.0
14NOV97	3	72.2
21NOV97	4	73.8
28NOV97	5	62.5
05DEC97	6	74.0
12DEC97	7	75.2
19DEC97	8	75.0
26DEC97	9	53.7
02JAN98	10	61.0
09JAN98	11	76.4
16JAN98	12	78.0

Figure 7-4 shows a scatter plot of the data. A linear regression model developed for this historical CPU utilization data has the following parameters:

$n = 12$, the number of historical observations

$b = 62.70$, the y intercept

$m = 1.17$, the slope of the line

$r^2 = 0.25$, the coefficient of determination

$F = 0.02$, the F value

$p = 0.90$, the probability that we should reject the hypothesis

$s_e = 6.68$, the standard error

The predicted and residual values for the historical data series are shown in the following table:

Week Ending	Observation Number	% CPU BUSY	Est % CPU	Residual (error)
=====	=====	=====	=====	=====
31OCT97	1	71.0	63.9	7.1
07NOV97	2	72.0	65.1	9.9
14NOV97	3	72.2	66.2	6.0
21NOV97	4	73.8	67.4	6.4
28NOV97	5	62.5	68.6	-6.1
05DEC97	6	74.0	69.8	4.2
12DEC97	7	75.2	70.9	4.3
19DEC97	8	75.0	72.1	2.9
26DEC97	9	53.7	73.3	-19.6
02JAN98	10	61.0	74.5	-13.5
09JAN98	11	76.4	75.6	0.8
16JAN98	12	78.0	76.8	1.2

As you can see in the model parameters and residual values in this table, the proposed model fits the historical data very poorly. In many cases, these problems are introduced by poorly behaved historical data rather than by the type of model selected by the analyst. In this example, three observations in the historical data (28NOV97, 26DEC97, and 02JAN98) are significantly different from the remainder of the historical data points. Investigation reveals that these three weeks represent holidays, presenting two alternatives:

- o Compensating the historical data points. For example, you could attempt to compensate for the missing data by multiplying by some constant. Unfortunately, such constants are guesses made by the analyst. Therefore, we do not recommend that you compensate historical data.
- o Deleting the errant historical data points. Although this reduces the number of points available for developing the model, it does not introduce any of the analyst's biases into the modeling process and is statistically defensible, since these weeks really do represent a different category of work for the processor.

Deleting the historical observations for the holiday weeks results in a substantially better model, giving significantly improved parameters. The parameters of the model are shown below:

$n = 9$, the number of observations

$b = 70.78$, the y intercept

$m = 0.57$, the slope of the line

$r^2 = 0.93$, the coefficient of determination

$F = 162$, the F value

$p = 0.0001$, the probability that we should reject the hypothesis

$s_e = 0.64$, the standard error

The predicted and residual values for the model that is developed from the historical series with the three holiday weeks deleted are shown in the following table.

Week Ending	Observation Number	% CPU BUSY	Est % CPU	Residual (error)
31OCT97	1	71.0	71.4	-0.4
07NOV97	2	72.0	71.9	-0.1
14NOV97	3	72.2	72.5	0.3
21NOV97	4	73.8	73.0	-0.8
28NOV97	5	.	73.6	.
05DEC97	6	74.0	74.2	-0.2
12DEC97	7	75.2	74.7	0.5
19DEC97	8	75.0	75.3	-0.3
26DEC97	9	.	75.9	.
02JAN98	10	.	76.4	.
09JAN98	11	76.4	77.0	0.6
16JAN98	12	78.0	77.6	-0.4

The model developed from the historical data series after the three holiday weeks were deleted is significantly better than the model developed before this deletion. This example shows the value of deleting errant historical data points. Note that the WEEKS timespan is probably more attractive for building models since there are often too few monthly observations for deletion to be an attractive alternative if the MONTHS timespan is used.

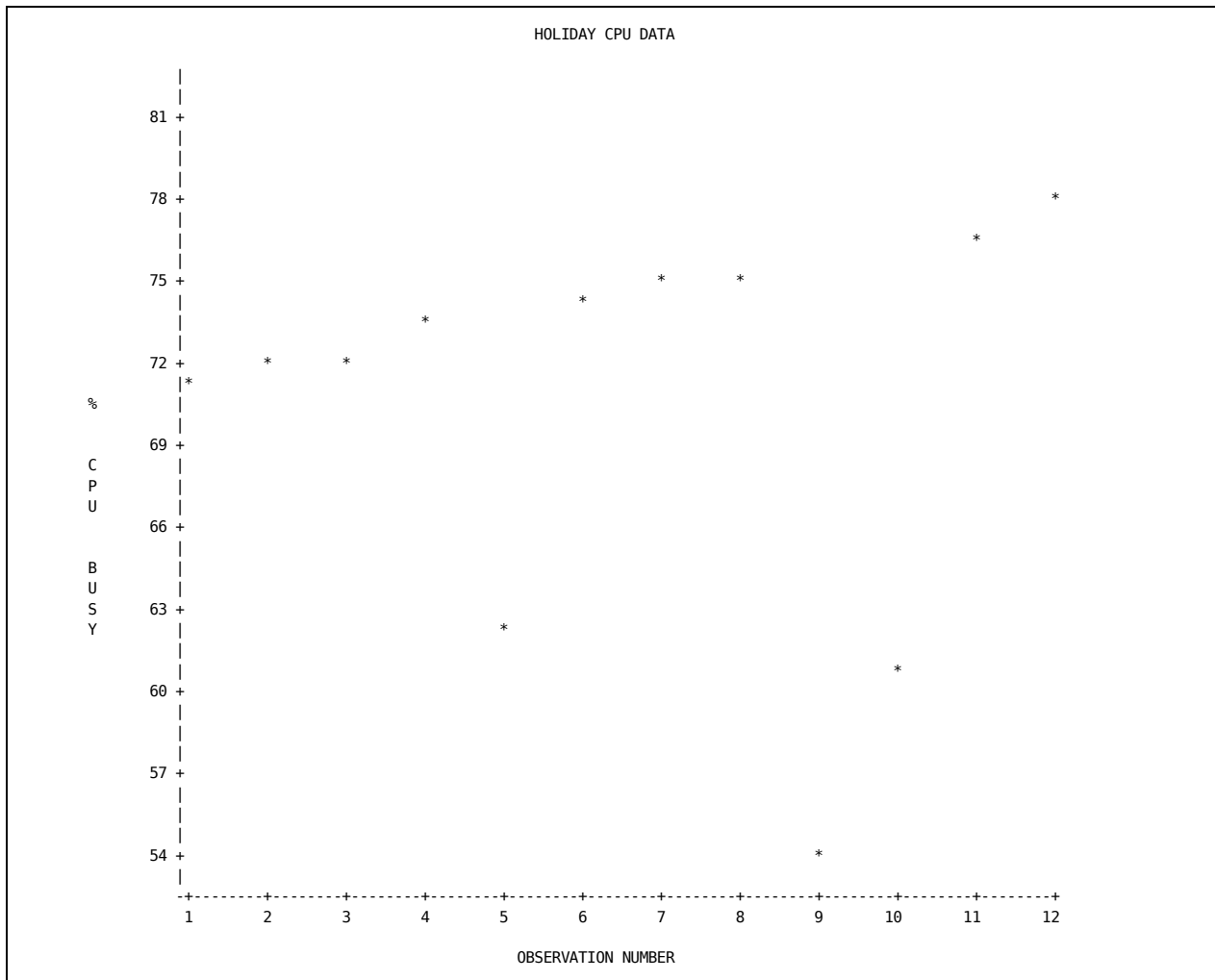


Figure 7-4. Weekly CPU Utilizations

7.4.2.2 Data Smoothing Using a Geometric Moving Average

One problem that analysts often encounter when attempting to forecast computer requirements is the high degree of variability shown by the observations in an historical data series. These variations are typically caused by random elements (for example, test jobs) that often exist in a system's workload. The following table shows the monthly observations of test jobs that are processed by a moderately sized MVS system, illustrates this problem.

Month	Observation Number	Test Jobs
=====	=====	=====
JAN98	1	2900
FEB98	2	3070
MAR98	3	2950
APR98	4	3080
MAY98	5	3200
JUN98	6	3150

Figure 7-5 shows a scatter plot of the data. A linear regression model developed for this historical data series has the following parameters:

$n = 6$, the number of historical observations

$b = 2881$, the y intercept

$m = 50.6$, the slope of the line

$r^2 = 0.68$, the coefficient of determination

$F = 8.47$, the F value

$p = 0.04$, the probability that we should reject the hypothesis

$s_e = 72.6$, the standard error

The predicted and residual values for the historical data series are shown in the following table:

Month	Observation Number	Test Jobs	Est. Jobs	Residual (error)
=====	=====	=====	=====	=====
JAN98	1	2900	2932	-32
FEB98	2	3070	2982	88
MAR98	3	2950	3033	-88
APR98	4	3080	3084	-4
MAY98	5	3200	3134	66
JUN98	6	3150	3185	35

Although one could argue that the model produced is marginally acceptable, only 68% of the variability in the historical data is accounted for by the model. You can treat this type of apparent randomness in historical data through the use of data smoothing techniques. Although a wide variety of techniques are available, perhaps the simplest is the geometric moving average (BAR77). A geometric moving average (GMA) is attractive in that you only need to sacrifice one historical data observation to calculate the smoothed series. (Other techniques require you to sacrifice many more historical data observations. For example, a five-point moving average requires you to sacrifice the first five observations.) You can calculate a geometric moving average using the following equation.

$$x(j) = \alpha * x(j) + \beta * x(j-1), \\ \text{for all } j \geq 2 \text{ (Eqn 10)}$$

$$\text{where } \alpha + \beta = 1.0$$

Thus, you can use the first and second observation to calculate a new value for the second observation, and so forth. Another feature of the geometric moving average is that you can select the degree of smoothing. If you select a large value for alpha (that is, $0.5 \leq \alpha < 1.0$), the smoothed series is less sensitive to variations between observations. Conversely, if you select a small value for alpha (that is, $0.0 < \alpha < 0.5$), the smoothed series is more responsive to variations in the historical data series.

For example, you can apply a geometric moving average to the historical data observations shown in the previous table. For this example, alpha equals 0.5. Hence, beta is equal to 0.5. The observation for January would be lost and the observation for February would be computed as

$$\text{FEB98} = 0.5 * 2900 + 0.5 * 3070 = 2985$$

The value for March would be computed based on the smoothed observation for February and the actual value for March. This procedure would be continued for the remainder of the observations, resulting in the following table:

Month	Observation Number	GMA Test Jobs
=====	=====	=====
JAN98	1	.
FEB98	2	2985
MAR98	3	2968
APR98	4	3024
MAY98	5	3112
JUN98	6	3131

Using the smoothed observations a second linear model was developed. The parameters for this model are shown below:

$n = 5$, the number of historical observations

$b = 2869$, the y intercept

$m = 43.6$, the slope of the line

$r^2 = 0.88$, the coefficient of determination

$F = 23.36$, the F value

$p = 0.02$, the probability that we should reject the hypothesis

$s_e = 29.6$, the standard error

The following table shows the predicted and residual values developed using this model:

Month	Obs #	GMA Test Jobs	Test Jobs	Est. Jobs	Residual (error)
=====	===	=====	=====	=====	=====
JAN98	1	.	2900	2914	.
FEB98	2	2985	3070	2957	28
MAR98	3	2968	2950	3000	-32
APR98	4	3024	3080	3043	-19
MAR98	5	3112	3200	3087	25
JUN98	6	3131	3150	3131	0

The forecast developed using the smoothed data series is much better behaved than the first forecast that was based on the untreated historical data series. Data smoothing is a powerful technique that you can use to minimize the effects of apparently random variations in historical data.

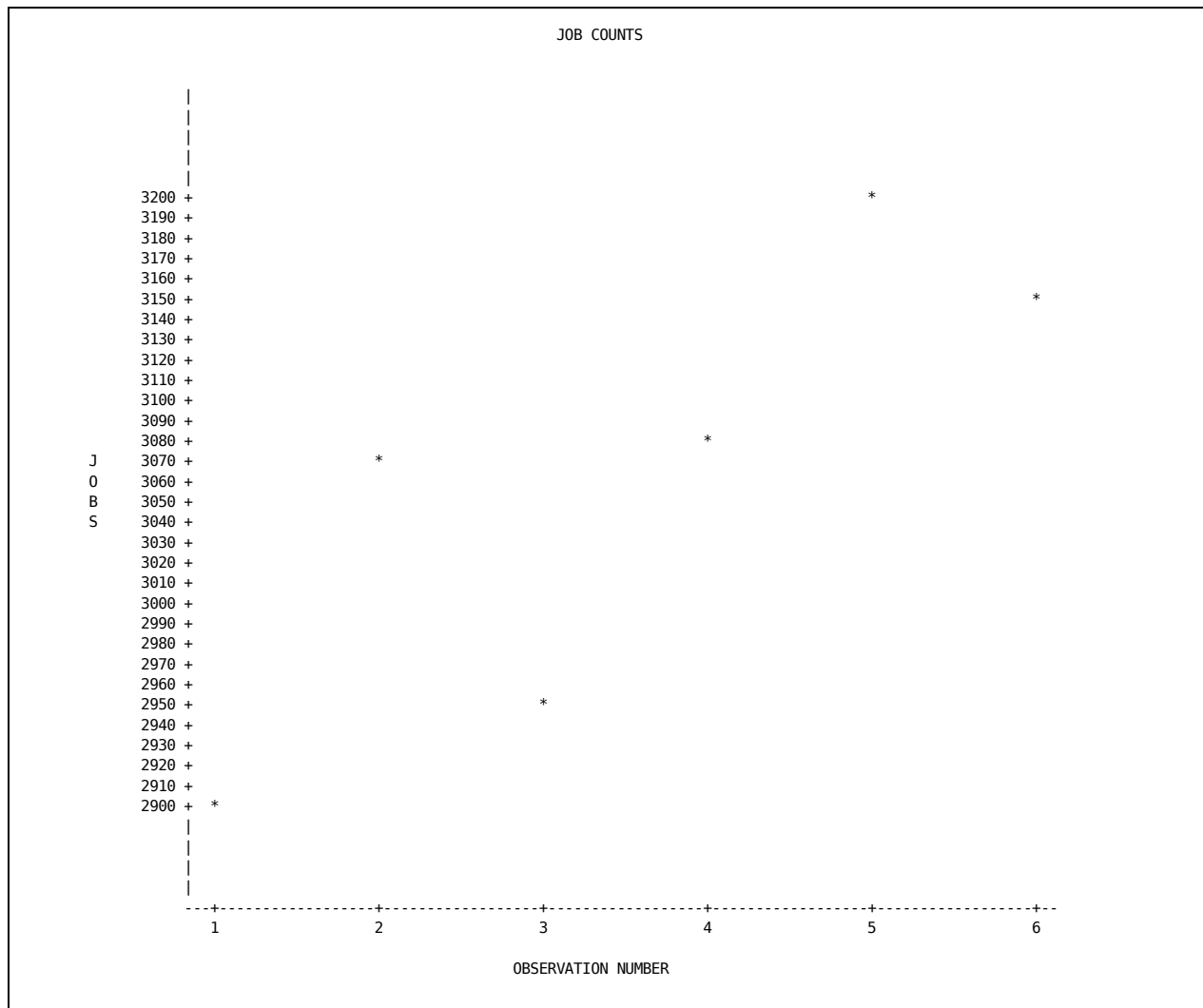


Figure 7-5. Monthly Job Counts

7.4.2.3 Data Smoothing Using a LOG10 Function

Fitting trend lines to service-oriented measurements is another problem that you may encounter. While trend-based estimates of future service levels may be inferior to the results that you can obtain using analytic queuing models, you can often fit regression models to service-oriented measurements to obtain a sense of direction for the values. Consider the following table of total pages per second collected for a 3090-200J processor:

Month	Observation Number	Page /Sec
=====	=====	=====
JAN98	1	75
FEB98	2	77
MAR98	3	80
APR98	4	86
MAY98	5	97
JUN98	6	113

Figure 7-6 shows a scatter plot of the data. A linear regression model that was developed for this historical data series has the following parameters:

n = 6, the number of historical observations

b = 83.1, the y intercept

m = 1.41, the slope of the line

r^2
r = 0.17, the coefficient of determination

F = 29.3, the F value

p = 0.01, the probability that we should reject the hypothesis

s
e = 14.8, the standard error

The predicted and residual values for the historical data series are shown in the following table:

Month	Observation Number	Page /Sec	Est. Pages	Residual (error)
=====	=====	=====	=====	=====
JAN98	1	75	84.5	9.5
FEB98	2	77	85.9	8.9
MAR98	3	80	87.3	7.3
APR98	4	86	88.7	2.7
MAY98	5	97	90.1	-6.9
JUN98	6	113	91.5	-21.5

The model poorly represents the historical data collected for paging. This is indicated by the small value of r-squared. To understand why this model fails, you must consider the underlying mechanism that produced the observations. Paging is an example of a service-oriented measurement that results from an underlying queuing relationship. That is, rather than expecting a linear increase in paging to a linear increase in load, you can expect an exponential increase. A standard technique (KEL74) for modeling exponential data is to use a logarithmic function to transform the data into a linear form. Univariate Model Forecasting employs a log base 10 function, as shown in the following equation:

$$x(j) = \text{LOG}_{10}(x(j)+1.0), \text{ for all } j \quad (\text{Eqn } 11)$$

Note that 1.0 is added to the historical observation before the transformation to avoid the undefined result of the LOG₁₀ function at zero. To interpret the forecasted values, we must reverse the transformation performed in Equation 11. The following equation shows this reverse transformation:

$$x(j) = 10^{x(j)} - 1.0, \text{ for all } j \quad (\text{Eqn } 12)$$

The 1.0 that is subtracted in Equation 12 corresponds to the 1.0 that was added in Equation 11. A LOG10 function was applied to the historical paging data presented in the previous table to "linearize" the observations. The result of this transformation is shown in the following table:

Month	Observation Number	LOG10 Pages
=====	=====	=====
JAN98	1	1.88
FEB98	2	1.89
MAR98	3	1.90
APR98	4	1.93
MAY98	5	1.99
JUN98	6	2.05

Using the smoothed observations a second linear model was developed. The parameters of this model are shown below:

n = 6, the number of historical observations

b = 1.84, the y intercept

m = 0.03, the slope of the line

r^2
r = 0.96, the coefficient of determination

F = 31.82, the F value

p = .005, the probability that you should reject the hypothesis

s = -0.06, the standard error. (Note that you must
e transform this value using Equation 11 to
obtain the actual standard error,
 $10^{*-0.06} = 0.88$.)

The future observations predicted by the model must be transformed back from the logarithmic scale in the same manner as the standard error value (described above). The predicted and residual values developed from model are shown in the following table:

Month	Obs #	Page /Sec	LOG10 Page	Est	Trans Est	Residual (error)
				LOG10 Pages		
JAN98	1	75	1.88	1.87	74.1	0.9
FEB98	2	77	1.89	1.90	79.4	-2.4
MAR98	3	80	1.90	1.93	86.1	-6.1
APR98	4	86	1.93	1.96	91.2	-5.2
MAY98	5	97	1.99	1.99	97.7	-0.7
JUN98	6	113	2.05	2.02	104.7	8.3

Forecasts developed using the transformed paging observations are superior to the forecast based on the original values. You should consider logarithmic transformation any time you are modeling service-related values like device utilizations, turnaround times, or response times.

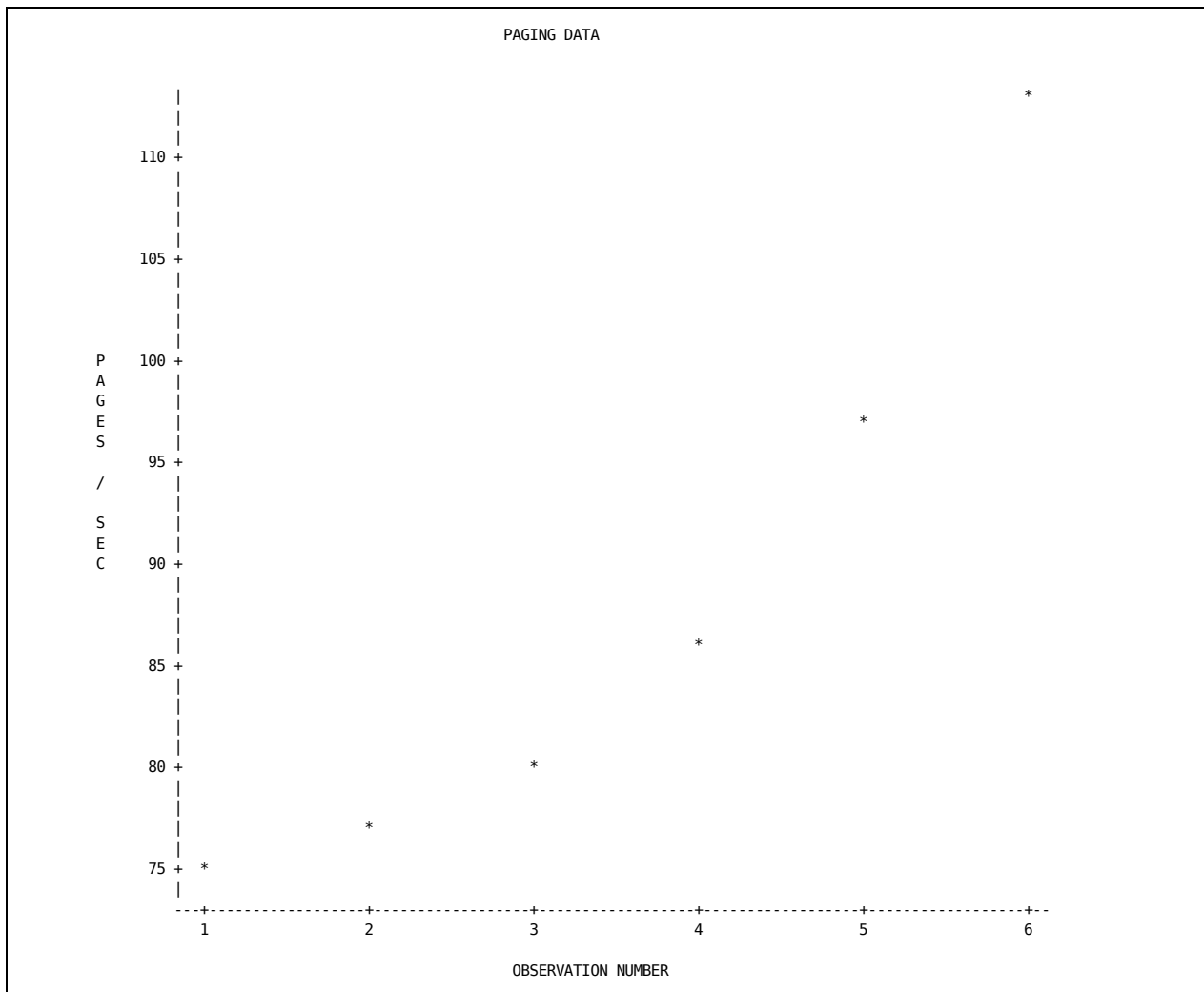


Figure 7-6. Monthly Page/Sec Values

7.4.2.4 User-Defined Data Transformations

Step function changes in the workload or other known changes can often introduce problems into linear regression forecasts. The following table illustrates this type of problem by presenting monthly observations of the average number of jobs per hour submitted by an engineering group. In April, three more engineers were added to the group, resulting in a step function increase of approximately 10 jobs per hour.

Month	Observation Number	Jobs Hour
JAN98	1	19
FEB98	2	25
MAR98	3	23
APR98	4	34
MAY98	5	37
JUN98	6	40

Figure 7-7 shows a scatter plot of the data. A linear regression model developed for this historical data series has the following parameters:

$n = 6$, the number of historical observations

$b = 14.5$, the y intercept

$m = 4.3$, the slope of the line

$r^2 = 0.92$, the coefficient of determination

$F = 5.1$, the F value

$p = .003$, the probability that we should reject the hypothesis

$s_e = 2.7$, the standard error

The following table shows the predicted and residual values for the historical data series:

Month	Observation Number	Jobs Hour	Est. Jobs	Residual (Error)
=====	=====	=====	=====	=====
JAN98	1	19	18.8	0.2
FEB98	2	25	23.2	1.8
MAR98	3	23	27.5	4.5
APR98	4	34	31.8	2.7
MAY98	5	37	36.2	-6.9
JUN98	6	40	40.5	-21.5

Although the r-squared of 0.92 for the model is very good, the results are somewhat suspicious due to the step function increase that occurred in the workload. You might want to modify the historical series to account for the step function increase of 10 jobs per hour in April. There are two ways to approach the problem:

- o Subtracting 10 from the April through June observations.
- o Adding 10 to the January through March observations.

You would probably want to modify the January through March observations since you would expect the future observations to be more like the April through June sales. Therefore, you would transform the historical observations using the following equation:

$$x(j) = x(j)+10, \text{ for } j \leq 3 \quad (\text{Eqn } 13)$$

$$x(j) = x(j), \quad \text{for } j > 3$$

Univariate Model Forecasting allows you to specify SAS statements to accomplish these transformations. The following table shows the result of this transformation:

Month	Observation Number	Transformed Jobs
=====	=====	=====
JAN98	1	29
FEB98	2	35
MAR98	3	33
APR98	4	34
MAY98	5	37
JUN98	6	40

Using the transformed observations, a second linear model was developed. The parameters of this model are shown below:

$n = 6$, the number of historical observations

$b = 28.5$, the y intercept

$m = 1.77$, the slope of the line

$r^2 = 0.79$, the coefficient of determination

$F = 15.23$, the F value

$p = 0.01$, the probability that the hypothesis should be rejected

$s_e = 1.92$, the standard error

The following table shows the predicted and residual values developed using this model:

	Obs	Jobs	Trans.	Est	
Month	#	Hour	Jobs	Jobs	Residual
=====	==	=====	=====	=====	=====
JAN98	1	19	29	30.2	-1.2
FEB98	2	25	35	32.0	3.0
MAR98	3	23	33	33.8	-0.8
APR98	4	34	34	35.6	-1.6
MAY98	5	37	37	37.3	-0.3
JUN98	6	40	40	39.1	0.9

Note that the r-squared and F values for the resulting model, while acceptable, are poorer than the values calculated for the original historical data series. However, the statistical model proposed is much more reasonable since it accounts for a significant step that occurred in the data. You should always examine each model to determine whether the proposed model is reasonable for the data being modeled.

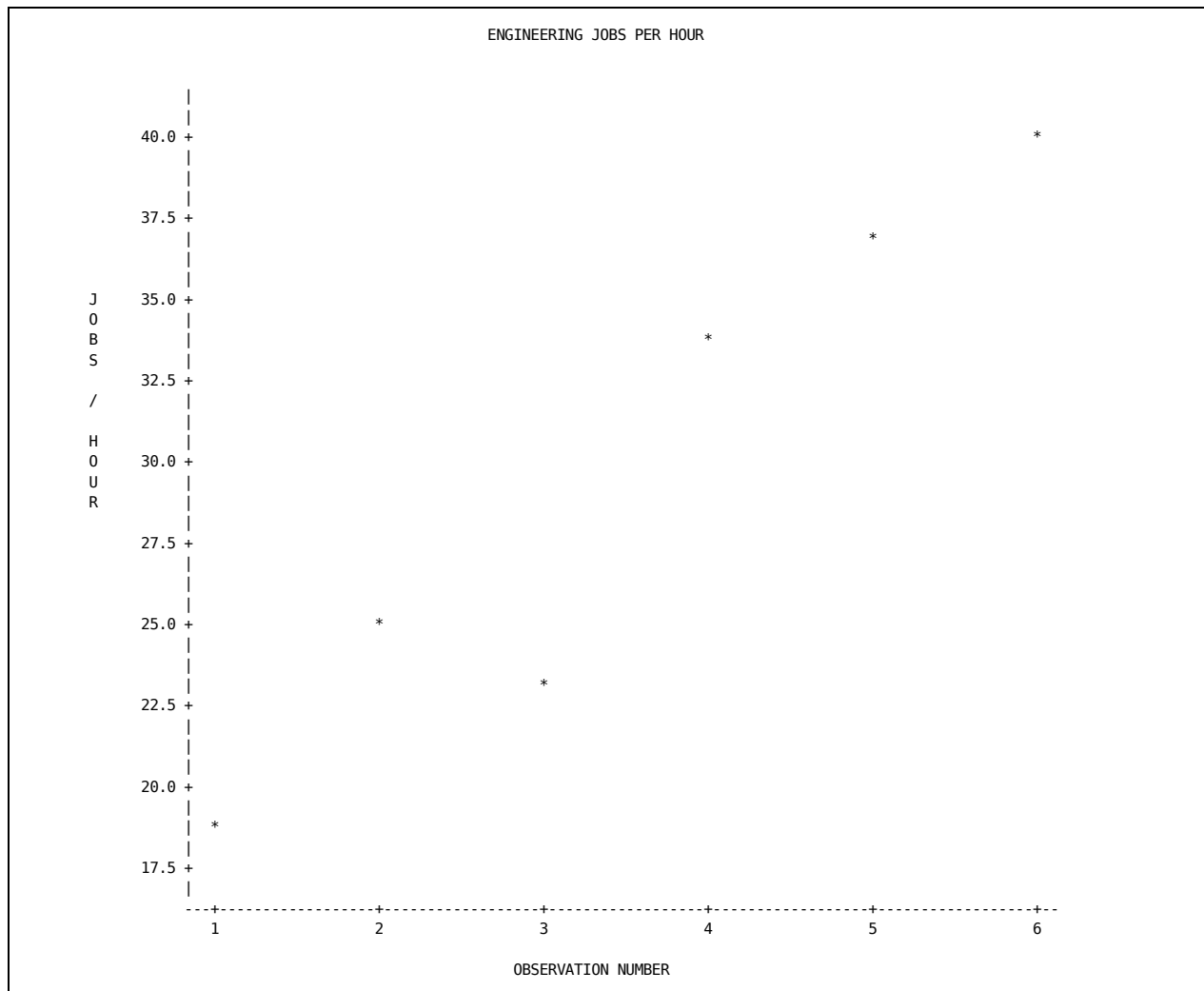


Figure 7-7. Engineering Jobs Per Hour

7.5 Component Operation

The Component Operation Section describes the control parameters that we used to build Univariate Model Forecasting inquiries. These parameters allow you to define forecasts based on linear, quadratic, and cubic regression, making use of a variety of techniques for data smoothing.

Because of the variety of optional features associated with Univariate Model Forecasting, you may need to use several screens to develop a forecast. The following sections discuss the use of each of these screens.

- 1 - Univariate Model Forecasting Control Screen
- 2 - Delete Selected Historical Observations Screen
- 3 - Transform Historical Data Observations Screen
- 4 - SAS Statements Screens

7.5.1 Univariate Model Forecasting Control Screen

Univariate Model Forecasting control parameters allow you to develop regression-based forecast models. You can save the forecasts produced by the program in a forecast file that corresponds to the resource or business element file you specify as input to the model.

Prior to seeing the Univariate Model Forecasting screen shown in Figure 7-8, you must follow a short set of menus which are more fully described in Section 1.5 of this guide. The Univariate Model Forecasting control screen is displayed after you indicate that you want to insert a new Univariate Model Forecast inquiry into your catalog of inquiries. Two fields are filled in on the fourth line when you first see this screen: inquiry name and inquiry description, separated by a dash. You entered these two fields on the Insert Capacity Planning Analysis screen when you requested that a new inquiry be added.

```

/----- Univariate Model Forecasting -----\
|Command ==>
|
|Composing CA MICS Inquiry: TSTUVM - Univariate Modeling Inquiry
|
|Report title ==> Univariate Modeling Inquiry
|
|Selection criteria:
|
|           Input file ==> R02 - RESOURCE ELEM FILE WITH CAPAPU
|           Start date ==> _____ (ddmonyyyy)
|           SYSID ==> _____
|           Hour ==> _____
|           No. weeks ==> _____ (1-9999)
|           Save forecast ==> YES (YES/NO/AGE)
|           Specify CAPAPU Values ==> N (Y/N)
|Forecast data element ==> PPATCBTM
|Forecast length (weeks) ==> 10 (1-9999)
|Confidence limits (percent) ==> 70 (70/90/95)
|
|Time element terms: Linear ==> YES (YES/NO)
|                       Quadratic ==> YES (YES/NO)
|                       Cubic ==> YES (YES/NO)
|           Create CSV Output ==> N (Y/N)
|Specify extended options ==> N (Y/N)
\-----/

```

Figure 7-8. Univariate Model Forecasting Control Parameters

Some of the parameters in the Univariate Model Forecasting screen shown in Figure 7-8 are required and some are optional. The following is a discussion of the various parameters:

Report title:	The 60-character title field for the Univariate Model Forecasting reports. No double quotes (") or semicolons (;) may appear within the body of the title.
Input file:	The resource element or business element Capacity Planning database file containing the historical observations to be used as input to the analysis. If you are uncertain of the files or file names available in your Capacity Planning database, type a question mark (?) in the field and a selection list will be displayed.

- Start date:** The start date for date range selection. If Start date is specified, only historical observations at least as recent as this value are included in the analysis. The discussion of the following control parameter describes the options for controlling date range selection.
- SYSID:** The SYSID for which the report is to be produced. If the Capacity Planning database file was defined as having a composite SYSID, this value is automatically supplied on the screen.
- Hour:** The Zone, Hour, or Dayname to be included in the analysis, depending on whether you defined the Capacity Planning database file at the Zone, Hour, or Dayname summarization level. If you defined the file as having a composite Zone, Hour, or Dayname, this value is supplied automatically on the screen. Note that the field name will be changed automatically to either Zone, Hour, or Dayname once you specify the input file to reflect the input file summarization level.
- No. days/
weeks/
months:** The maximum number of days, weeks, or months of historical observations that will be included in the analysis. Note that once you specify the input file, the field name will change to reflect the timespan of the Input file. When used in conjunction with the Start date parameter, several options are available for date range selection control:
- o If you specify only Start date, all historical observations at least as recent as the value specified for Start date will be selected.
 - o If you specify both Start date and No. dy/wk/mn, the date range for data selection begins with the value specified for Start date and extends for the specified number of days, weeks, or months.

	- o If you specify only No. dy/wk/mn, the date range for data selection will include only this number of the most recent days, weeks, or months. - o If you specify neither parameter, then all historical observations will be selected.
Save Forecast:	Three options for saving the forecasted observations: - o Specifying YES causes the forecasted observations to be saved in the Capacity Planning database forecast file (CPFxxx01) without aging the file, as described in Chapter 3. - o Specifying NO results in the forecasted observations not being saved in the forecast file. If you select this option, the forecast file will not be aged. - o Specifying AGE causes the forecasted observations to be saved in the forecast file and causes the forecast files to be aged, as described in Chapter 3.
Specify CAPAPU Values	Put a Y in this field to specify the CAPAPU value(s) to be included in the analysis. The CAPAPU value is specified on the CAPAPU Selection Values screen, shown in Figure 7-9, that appears if you enter a Y in this field. You may specify CAPAPUs only if the Capacity Planning database file is defined as containing CAPAPU(s).
Forecast data element:	The data element that is to be the subject of Univariate Model Forecasting. If you are uncertain of the correct name of the data elements in the input file, type a question mark (?) in this field and a selection list of available data elements will be presented.

Forecast length day/week/month:	The length of forecast period in days, weeks, or months. Note that the field name changes, once you have specified the Input file, to reflect the timespan of the Input file. (Generally, you should not specify a forecast length greater than the length of your historical data series. See Section 6.2 for guidelines on choosing this value.)\
Confidence limits (percent):	The degree of confidence to be used in calculating the upper and lower confidence limits for the forecasted values of the forecast data element. The default is 95%.
Time element terms:	The time elements to be used in the regression process. You must specify at least one of the following terms: LINEAR, QUADRATIC, or CUBIC. You may specify models of the forms shown in Equations 7, 8, and 9 (see Section 7.4.1). For example, you could develop a model of CPUHRS based on linear and quadratic time terms by specifying LINEAR and QUADRATIC. The model you can develop using these terms corresponds to Equation 8. You could also develop a model of CPUHR based on a cubic time term, corresponding to a special case of Equation 9, where m1 and m2 are zero.
Create CSV Output:	Default value is "N" for NO. Specify "Y" for YES to cause a Q&R formatted CSV to be created that contains those data elements used in the analysis. Specify "N" for NO to bypass CSV creation.
Specify extended options:	The specification for extended options. Specify a Y for the opportunity to specify the options on the Extended Options Menu (Figure 7-10).

If you entered a Y in the Specify CAPAPU Values field on the Univariate Model Forecasting screen, the CAPAPU Selection Values screen displays, as shown in Figure 7-9. On the CAPAPU Selection Values screen, you may specify the CAPAPU values (up to nine) to be included in the analysis. You may specify CAPAPUs only if the Capacity Planning database file is defined as containing the CAPAPU data elements. You may use wild cards (?) and (*) when specifying CAPAPUs. A question mark (?) replaces one character; an asterisk (*) replaces all the characters after that point.

```

/----- CAPAPU Selection Values -----\
|Command ==>                               |
|                                           |
|Composing CA MICS Inquiry: TSTUVM Univariate Modeling Inquiry|
|                                           |
|      Rule  ----+----1----+----2----+----3----+----4----+----5----+----6----|
|CAPAPU  ==>  -----|
|CAPAPU 2 ==>  -----|
|CAPAPU 3 ==>  -----|
|CAPAPU 4 ==>  -----|
|CAPAPU 5 ==>  -----|
|CAPAPU 6 ==>  -----|
|CAPAPU 7 ==>  -----|
|CAPAPU 8 ==>  -----|
|CAPAPU 9 ==>  -----|
|                                           |
\-----/

```

Figure 7-9. CAPAPU Selection Values

If you entered a Y in the Specify Extended Options field on the Univariate Model Forecasting screen, the Extended Options Menu appears, as shown in Figure 7-10.

```
/----- Extended Options Menu -----\  
|Command ==>|  
|  
|Inquiry Step: Univariate Model Forecasting|  
|  
|Print Reports          ==> YES  (YES/NO)|  
|Maximum R-Square for report ==> 0.55 (0.00-1.00)|  
|P-Value flag          ==> 0.05 (0.000-1.000)|  
|  
|Delete historical data observations ==> N (Y/N/R)|  
|Transform historical observations ==> N (Y/N/R)|  
|  
\-----/
```

Figure 7-10. Extended Options Menu

Print Reports:	The Print Reports parameter controls the printing of the Analysis of the Model and Residual Analysis reports. If you type YES, these reports are produced. If you type NO and the forecasted file is NOT a Business Element File, these reports are suppressed. These reports will always be produced for Business Element Files; however, the R-Square Ranking Report is not produced.
Maximum R-Square for report:	The Maximum R-Square for report parameter is used during the production of the R-Square Ranking Report to control the flagging of particular forecasts. If the R-Square of a generated model is less than this value, it will be flagged with an RS-WARN flag. A value of 1.0 in this field will disable the production of the RS-WARN flag in the R-Square Ranking Report.

P-Value flag: The P-Value Flag is also used in the production of the R-Square Ranking Report to determine if any of the P-values of the model variables should be flagged with the P-WARN flag. If any P-value in the model is greater than this value, it will be flagged in the report. A value of 1.0 in this field will disable the production of the P-WARN flag in the R-Square Ranking Report.

Note: The RS-WARN flag takes precedence over the P-WARN flag. This means that if a model exceeds both of these warning criteria, the RS_WARN flag will be printed.

The other options shown on the Extended Options Menu are discussed in the following sections:

- 1 - Delete Selected Historical Data Observations
- 2 - Transform Historical Observations

7.5.2 Delete Selected Historical Data Observations Screen

If you specify YES for extended options on the Univariate Model Forecasting screen, you have the option to delete selected historical observations from the modeling process. If you select this option, the Delete Historical Data Observations screen shown in Figure 7-11 is displayed.

```
/----- Delete Historical Data Observations -----\  
|Command ==>                                     Scroll ==> CSR |  
|  
|Inquiry Step:  Univariate Model Forecasting      |  
|  
|Line Cmds:  I Insert  D Delete  R Repeat  M Move  C Copy  |  
|  
|Cmd  -----  Dates to be dropped (specify in ddmonyy format)  -----  
| -  -----  -----  -----  -----  -----  -----  -----  
| -  -----  -----  -----  -----  -----  -----  -----  
| -  -----  -----  -----  -----  -----  -----  -----  
| -  -----  -----  -----  -----  -----  -----  -----  
| -  -----  -----  -----  -----  -----  -----  -----  
| -  -----  -----  -----  -----  -----  -----  -----  
| -  -----  -----  -----  -----  -----  -----  -----  
| -  -----  -----  -----  -----  -----  -----  -----  
| -  -----  -----  -----  -----  -----  -----  -----  
| -  -----  -----  -----  -----  -----  -----  -----  
| -  -----  -----  -----  -----  -----  -----  -----  
| -  -----  -----  -----  -----  -----  -----  -----  
| -  -----  -----  -----  -----  -----  -----  -----  
| -  -----  -----  -----  -----  -----  -----  -----  
| -  -----  -----  -----  -----  -----  -----  -----  
|  
|-----|
```

Figure 7-11. Delete Selected Historical Observations Screen

Note that the dates you specify on this screen must exactly match the DATE value of the historical observations to be excluded from the analysis. Although deleted observations still appear in the reports, they will not be used to develop the regression model.

7.5.3 Transform Historical Observations Screen

If you specify YES for extended options on the Univariate Model Forecasting screen, you will have the opportunity to specify the transformation of historical observations. Using this option, you can specify either GMA smoothing, a LOG10 transformation, or a user-defined transformation. Specify NONE if you do not want smoothing or transformation. NONE is the default.

If you specify USER for this option, then you are required to provide the SAS statements to perform the transformation and/or the reverse transformation by specifying a Y for at least one of the last two options on this screen.

```

/----- Transform Historical Observations -----\
|Command ==>                                     |
|                                                |
|Inquiry Step:  Univariate Model Forecasting    |
|                                                |
|Smoothing technique ==> NONE  (GMA/LOG10/USER/NONE) |
|                                                |
|                                                |
|Geometric moving average data smoothing parameters: |
|  Equation alpha value  ==> ____  (.01 - .99)      |
|  Initial value         ==> _____             |
|                                                |
|User defined data transformation SAS statements:  |
|  Statements to transform historical data  ==> N (Y/N/R) |
|  Stmts. to reverse history transformations ==> N (Y/N/R) |
|                                                |
|                                                |
|                                                |
\-----/

```

Figure 7-12. Transform Historical Observations Screen

The Transform Historical Observations screen allows you to modify the historical data series prior to the modeling process. (The processing of historical data observations is discussed in Section 7.4.2.) You can choose from three different types of optional processing:

- o Data smoothing using a geometric moving average

- o A log base 10 transformation of the historical observations
- o A user-specified exit to process the historical observations and possibly a corresponding exit to reverse the transformation process after the forecasting process

If you specify that the data is to be smoothed using a geometric moving average, then one other parameter is required and one more is optional. You must specify the value of alpha ($0 < \alpha < 1$) to be used in the geometric moving average equation (Equation 10). We recommend that you initially try a value of 0.5.

In addition, you may optionally specify an alternative value for the initial historical observation. This value is used to replace the first value when you do not believe that the first point of the historical data series is representative. If you do not wish to have an alternative value, leave this parameter blank.

If you specify that the historical observations are to be transformed using a LOG10 function, no other parameters are required or allowed on this screen.

If you specify that the historical observations are to be transformed using your own SAS code, you need to indicate on this screen which type of transformation you wish to include:

- o Statements to transform historical data
- o Statements to reverse the transformation after the analysis

If you select either of these transformations, the appropriate screen is displayed to allow you to enter your own SAS code. These screens are discussed in Section 7.5.4.

7.5.4 SAS Statements Screens

You can use the first SAS Statements screen, shown in Figure 7-13, for coding SAS statements to transform historical data prior to the regression analysis performed by Univariate Model Forecasting.

You can use the second screen, shown in Figure 7-14, to reverse this transformation process after the regression is done, so that the reports show the untransformed values.

```

/----- SAS System Statements -----\
|Command ==>                                Scroll ==> PAGE|
|
|SAS Statements for the ASGN_TRN Macro
|-----\
|***** ***** TOP OF DATA *****|
|==MSG> ENTER SAS STATEMENTS TO PERFORM A USER TRANSFORMATION.|
|==NOTE= %MACRO ASGN_TRN; <=== PLEASE DO NOT PLACE LINES BEFORE THIS LINE.|
|000001|
|000002|
|000003|
|000004|
|000005|
|000006|
|000007|
|000008|
|000009|
|000010|
|==NOTE= %MEND ASGN_TRN; <=== PLEASE DO NOT PLACE LINES AFTER THIS LINE.|
|***** ***** BOTTOM OF DATA *****|
|
|-----\

```

Figure 7-13. SAS Statements Screen - ASGN_TRN Macro

```

/----- SAS System Statements -----\
|Command ==>                               Scroll ==> PAGE |
|
|SAS Statements for the RET_TRN Macro
|-----|
|***** ***** TOP OF DATA *****|
|==MSG> SPECIFY SAS STATEMENTS TO REVERSE A USER TRANSFORMATION. |
|=NOTE= %MACRO RET_TRN; <=== PLEASE DO NOT PLACE LINES BEFORE THIS LINE. |
|000001
|000002
|000003
|000004
|000005
|000006
|000007
|000008
|000009
|000010
|=NOTE= %MEND RET_TRN; <=== PLEASE DO NOT PLACE LINES AFTER THIS LINE. |
|***** ***** BOTTOM OF DATA *****|
|
\-----/

```

Figure 7-14. SAS Statements Screen - RET_TRN Macro

The remainder of this section describes considerations for coding these macros.

Note that the %MACRO statements appearing on each of these screens are already provided for you and are field-protected from modification.

Univariate Model Forecasting does not work directly with the variable names on the input SAS file, but rather uses SAS macro variables to pass the names taken from the control parameters into the program's code. This allows Univariate Model Forecasting to be sufficiently generalized to handle all incoming SAS files without imposing naming conventions on you. If you are writing SAS statements to transform the original historical data using the first screen shown in Figure 7-13, refer to the following symbolic names for variables:

&VNAME: The original dependent variable

_TRAN_V_: The transformed dependent variable

In some instances, you may want to code a reverse transformation macro that returns the predicted variables to the original dependent variable units. When the reverse transformation macro is invoked, the SAS variable names that contain the critical data are as follows:

- ___ESTS: The predicted value of the transformed dependent variable (three leading underscores)
- _____UCL: The upper confidence limit of the transformed predicted variable (five leading underscores)
- _____LCL: The lower confidence limit of the transformed predicted variable (five leading underscores)

To illustrate this concept, assume there is a resource element file that contains CPU utilization data. Further, assume that on March 8, 1998, a new CPU was installed that is approximately 30% more powerful than its predecessor. A CPU utilization trend over time would probably show a sudden decline in CPU utilization on or around March 8, 1998. This discontinuity in the trend would result in a very poor fit by a linear, quadratic, or cubic equation.

To provide a set of historical points that can result in a better fit to a standard equation, you can transform all the historical data after March 8, 1998, by multiplying it by 1.3 (representing the 30% increase provided by the new processor). Following the curve fitting, the projected trend is transformed back into units of the current processor by dividing the forecast values by 1.3. (This adjustment for relative processor speeds can be done for you automatically if you use the CPU Planning Standard Application described in Chapter 3. The process described above is for illustrative purposes only.)

On the first screen, shown in Figure 7-13, you would enter the following SAS statements:

```
IF DATE GE '08MAR98'D THEN _TRAN_V_ = 1.3 * &VNAME;
ELSE _TRAN_V_ = &VNAME;
```

On the second screen, you would enter the following SAS statements:

```
IF DATE GE '08MAR98'D THEN DO;
  ___ESTS = ___ESTS / 1.3;
  _____LCL = _____LCL / 1.3;
  _____UCL = _____UCL / 1.3;
END;
```

7.6 Case Studies

To illustrate the use of the univariate modeling techniques, four case studies are presented in this section.

The first case study, Extended Example Univariate Model Forecasting, is an extended example that applies Univariate Model Forecasting to the study of the I/O workload of a 9021-711 and illustrates how to use a geometric moving average. This study is a continuation of the study started in Section 3.3.3.1.

The second study, Performance Group/Period Workload Tracking, shows how you can use Univariate Model Forecasting to track the growth of a workload that is defined by a performance group. This example shows how to use the data deletion facility to discard an observation that is unusually low due to an unexpected event, such as a system outage. This case study is a continuation of the study started began in Section 3.3.3.2.

The third study, System Workload and Performance Profile, looks at trends in CPU utilization and uses the geometric moving average to dampen the variability in the original data. This case study is a continuation of the study started in Section 3.3.3.5.

The final study, Forecasting New Application Transaction Values, uses a user-defined transformation to forecast the new volume of an online workload that you expect to add work to. This type of analysis is useful for estimating future transaction loads, which you can then enter into an analytic model to measure whether sufficient capacity is available to meet the demand.

The case studies are presented in the following order:

- 1 - Extended Example Univariate Model Forecasting
- 2 - Performance Group/Period Workload Tracking
- 3 - System Workload and Performance Profile
- 4 - Forecasting New Application Transaction Values

7.6.1 Extended Example Univariate Model Forecasting

This extended example shows how to apply Univariate Model Forecasting to the I/O workload data collected in the extended example that we introduced in Section 3.3.3.1. The following topics are addressed:

- 1 - Problem Description
- 2 - Control Parameters
- 3 - Output Reports

7.6.1.1 Problem Description

The purpose of this study is to develop a model to forecast the start I/O counts and the utilization loads that would be experienced by the channels on a 9021-711. The collecting historical data to support this example was discussed in the extended example in Section 3.3.3.1. Twelve months of historical data was collected for channels 1 through 5 for the system SYSIDs TS01 and TS02 for ZONES 1 and 2.

This extended example concentrates on the start I/O counts for channel 5 on SYSID TS01 during ZONE 1. Although the development of the remaining forecasts (utilizations and the SIO counts for the other channels) are not shown in this example, they are similar to the forecast discussed in the following sections.

The first step in the modeling process is to obtain a listing of the historical data observations. To do this, specify YES for the List updated CAP DB file contents parameter on the Update Capacity Planning File screen shown in Chapter 3 when a database update inquiry is executed. Figure 7-15 shows a sample report from the update of the CHN Capacity Planning database file. Note that while the CHN database file contains observations for TS01 and TS02 for Zones 1 and 2, this extended example focuses only on the observations for TS01 for Zone 1. A cursory investigation of the channel 5 SIO counts (data element CHAN5CNT) reveals an increasing trend with a significant amount of random variations. (These random variations are not surprising since the system being analyzed processes a random arrival batch and TS0 workload in a service bureau.)

Therefore, the initial approach to developing a model for this system was to fit a linear trend to the historical values after they were smoothed using a geometric moving average.

CA MICS Capacity Planner															
OBS	DATE	MONTH	SYSID	YEAR	ZONE	CHAN1BS	CHAN1CNT	CHAN2BS	CHAN2CNT	CHAN3BS	CHAN3CNT	CHAN4BS	CHAN4CNT	CHAN5BS	CHAN5CNT
1	30SEP97	9	TS01	97	1	5.3569	10283160	18.9219	12621388	18.8242	16971048	7.2551	3640411	15.7917	11882418
2	30OCT97	10	TS01	97	1	5.5168	11306966	19.4883	13861055	19.6445	18385386	7.3911	3822387	14.7969	12860633
3	30NOV97	11	TS01	97	1	6.3586	11414859	19.7500	11717108	20.2109	16663774	9.9683	4338143	15.9124	11743581
4	31DEC97	12	TS01	97	1	5.3074	9976870	18.0391	12327566	18.5586	16800745	8.8547	5247391	14.6099	12510012
5	29JAN98	1	TS01	98	1	6.0647	10611907	18.9922	11156555	19.3750	16185623	8.9697	4257223	15.6548	12071368
6	26FEB98	2	TS01	98	1	6.3376	10960937	23.0977	12830799	18.8438	15257229	11.1433	5641456	16.2344	12076524
7	31MAR98	3	TS01	98	1	5.4783	10025924	19.6797	12843967	16.9063	13669835	9.5486	5318293	14.8582	12186976
8	30APR98	4	TS01	98	1	4.7642	9870709	17.9570	13093325	15.2307	14904987	7.6726	5033416	14.3140	14355185
9	28MAY98	5	TS01	98	1	4.5576	8851487	17.7852	11399322	15.6868	13645679	8.2495	4494512	14.5493	15011392
10	30JUN98	6	TS01	98	1	4.9968	10545954	17.5703	12099310	15.7957	14443173	9.2905	5437664	15.1023	17176448
11	30JUL98	7	TS01	98	1	6.9377	14139601	21.0313	13436249	18.6172	15858031	11.7161	6582489	17.7891	17421158
12	31AUG98	8	TS01	98	1	8.5986	14220067	30.9336	17448773	15.0393	14690555	10.7354	6879567	15.3848	18042957
13	30SEP97	9	TS01	97	2	3.6626	3844805	21.0391	9137237	25.5938	12868504	9.1514	4359236	16.0000	10247086
14	30OCT97	10	TS01	97	2	3.6033	4058408	21.1641	9761257	23.7773	12025920	8.2043	3752326	14.9243	10098182
15	30NOV97	11	TS01	97	2	3.5081	3559998	21.0078	8160196	26.0352	11506972	9.4600	3786755	15.9133	9260268
16	31DEC97	12	TS01	97	2	2.6174	3018221	22.1406	10373391	30.1250	12866071	11.2361	4973300	16.9727	11058613
17	29JAN98	1	TS01	98	2	2.0652	2197819	18.7461	8134965	24.0313	10484711	8.0525	3572788	15.1262	8586867
18	26FEB98	2	TS01	98	2	2.4753	2669361	21.7070	8721177	24.9844	11027921	10.0513	4014299	16.1367	9080303
19	31MAR98	3	TS01	98	2	1.9980	2196140	18.5313	7999537	22.3164	11310782	9.1003	3852128	13.9209	8682421
20	30APR98	4	TS01	98	2	1.5474	1753226	17.7422	7390090	20.3477	9696053	7.9121	3435097	14.7136	9351902
21	28MAY98	5	TS01	98	2	1.5369	1670242	16.2227	6574310	18.8555	9463561	8.3237	3304728	13.3713	8148197
22	30JUN98	6	TS01	98	2	2.0928	2596361	17.6797	7756252	21.4180	11278067	9.4841	3893271	15.0049	9578109
23	30JUL98	7	TS01	98	2	3.1802	3551786	20.6289	7575368	26.1992	13215648	11.2686	3863744	14.9900	8710056
24	31AUG98	8	TS01	98	2	3.3550	3475517	30.0859	10478833	23.7578	10705401	11.3088	4154695	13.9648	8859281
25	30SEP97	9	TS02	97	1	5.8586	4117146	19.7070	19625621	20.8047	22304391	12.7476	9302529	10.8057	12295717
26	30OCT97	10	TS02	97	1	14.7346	2722390	19.8320	20338573	19.9375	22521512	15.4590	11686631	10.7395	12889200
27	30NOV97	11	TS02	97	1	13.7773	3809002	14.9670	15289886	18.5938	19254275	12.3804	8388867	8.9019	10402395
28	31DEC97	12	TS02	97	1	10.0374	4721964	13.4104	13712714	16.6367	17817953	15.4211	10188718	9.0366	12847178
29	29JAN98	1	TS02	98	1	12.7666	5548874	16.8281	13845993	17.6797	18586811	12.6563	8258431	9.9780	15816183
30	26FEB98	2	TS02	98	1	10.8706	7771787	16.3789	13405987	15.9539	17237412	13.2949	10653306	10.2646	17803695
31	31MAR98	3	TS02	98	1	13.4224	9735884	15.5269	14151218	15.9382	17819024	14.6516	10932092	9.5662	17208264
32	30APR98	4	TS02	98	1	11.6868	5365007	13.5408	13405430	13.8350	17208020	13.9941	10965281	8.9309	18366040
33	28MAY98	5	TS02	98	1	8.6855	6873986	14.3179	13105377	13.7134	15188405	13.4080	9084066	10.0325	18941408
34	30JUN98	6	TS02	98	1	10.0339	7030206	15.9690	14749902	14.5276	16348721	16.5117	10999226	10.6196	20710472
35	30JUL98	7	TS02	98	1	8.2434	7093482	14.4727	13043652	13.4624	14116174	15.6594	10289280	8.7729	17823803
36	31AUG98	8	TS02	98	1	13.1682	9970887	24.8711	17365170	13.7458	13949379	12.1421	9414241	8.8535	15638684
37	30SEP97	9	TS02	97	2	2.7336	880241	18.6016	9049081	25.0039	10550887	10.5205	3584787	11.0811	6455642
38	30OCT97	10	TS02	97	2	7.3264	679033	21.8281	9497450	24.7461	10452369	14.0229	4612623	10.7666	6382688
39	30NOV97	11	TS02	97	2	6.6284	784134	17.5781	7030350	21.5352	8291186	8.3538	2943352	11.3120	5859400
40	31DEC97	12	TS02	97	2	4.1448	858595	13.8445	5741095	16.3867	6745006	9.0964	3310562	8.4487	5035398
41	29JAN97	1	TS02	98	2	4.4250	956579	18.9531	7723973	21.4648	8545945	10.4185	3550842	11.7129	6831805
42	26FEB98	2	TS02	98	2	3.6133	1217462	20.1680	7515701	20.2422	7863046	13.5088	4195880	13.0649	7158157
43	31MAR98	3	TS02	98	2	6.2932	2400223	23.8320	9675267	25.9297	11585347	15.1511	5334095	14.5879	8925649
44	30APR98	4	TS02	98	2	4.4189	1045310	22.6758	8853916	22.8203	10761397	15.0293	4769231	12.6448	7748332
45	28MAY98	5	TS02	98	2	3.0962	1307297	20.6953	7138379	22.4414	9365947	14.7139	3981306	13.8040	7103022
46	30JUN98	6	TS02	98	2	3.2813	1239208	22.0625	8164637	20.7148	9408193	14.6201	4561230	12.8960	7367164
47	30JUL98	7	TS02	98	2	2.0996	896128	18.2734	5934381	16.2227	6571253	14.1196	3805317	8.7349	5075665
48	31AUG98	8	TS02	98	2	3.4624	1362194	26.0430	7640659	17.9375	5866009	14.2581	3940249	9.8955	4596676

Figure 7-15. CHN Resource Element File Contents Report

7.6.1.2 Control Parameters

Figure 7-16 shows the Univariate Model Forecasting, Extended Options, and Transform Historical Observations screens that contain the control parameters used to perform univariate modeling on the channel 5 SIO count.

```

/----- Univariate Model Forecasting -----\
|Command ==>
|
|Composing CA MICS Inquiry: CHNUVM - Univariate Modeling Inquiry
|
|Report title ==> Univariate Modeling Inquiry
|
|Selection criteria:
|           Input file ==> CHN - I/O Channel Example
|           Start date ==> _____ (ddmmyyyy)
|           SYSID ==> TS01
|           Zone ==> 1_____
|           No. months ==> _____ (1-9999)
|           Save forecast ==> YES      (YES/NO/AGE)
|           Specify CAPAPU Values ==> N      (Y/N)
|Forecast data element      ==> CHAN5CNT
|Forecast length (months)   ==> 8          (1-9999)
|Confidence limits (percent) ==> 95        (70/90/95)
|
|Time element terms: Linear  ==> YES      (YES/NO)
|           Quadratic ==> ____      (YES/NO)
|           Cubic      ==> ____      (YES/NO)
|Specify extended options   ==> Y        (Y/N)
\-----/

```

Figure 7-16. Extended Example Control Parameters (Screen 1 of 3)

```

/----- Extended Options Menu -----\
|Command ==>
|
|Inquiry Step: Univariate Model Forecasting
|
|Print Reports                ==> YES    (YES/NO)
|Maximum R-Square for report   ==> 0.55  (0.00-1.00)
|P-Value flag                  ==> 0.05  (0.000-1.000)
|
|Delete historical data observations ==> N (Y/N/R)
|Transform historical observations  ==> N (Y/N/R)
|
\-----/

```

Figure 7-16. Extended Example Control Parameters (Screen 2 of 3)

```

/----- Transform Historical Observations -----\
|Command ==>                                     |
|Inquiry Step:  Univariate Model Forecasting      |
|Smoothing technique          ==> GMA   (GMA/LOG10/USER/NONE) |
|                                                    |
|Geometric moving average data smoothing parameters: |
|    Equation alpha value ==> .5       (.01 - .99) |
|    Initial value       ==> _____ |
|                                                    |
|User defined data transformation SAS statements:  |
|  Statements to transform historical data ==> N (Y/N/R) |
|  Stmts. to reverse history transformations ==> N (Y/N/R) |
|                                                    |
\-----/

```

Figure 7-16. Extended Example Control Parameters (Screen 3 of 3)

The Univariate Model Forecasting screen shown in Figure 7-16 contains the following parameters.

Report title: The title that appears on the reports.

Input file: The CHN resource element file that was originally defined in the extended example in Section 3.3.3.1 of this guide.

Start date and No. months: The start date for date range selection and the maximum number of days, weeks, or months included in the analysis.

Since neither Start date nor No. months is specified for this analysis, no date range selection is applied to the historical observations from the CHN resource element file.

SYSID: Historical observations only for SYSID TS01.

Zone:	The summary level selection criteria for the analysis. Since the CHN resource element file was defined in Section 3.3.3.1 at the ZONE summary level, this parameter specifies that only historical observations for ZONE 1 are included in the analysis.
Save Forecast:	The specification to save the results of the analysis in the forecast file. For this example, YES is specified so that the forecast file will not be aged prior to saving (that is, prior values are overwritten).
Specify CAPAPU Values:	The specification for CAPAPU. The CHN resource element file was not defined as containing the CAPAPU data element, therefore no CAPAPU value may be specified on this screen.
Forecast Data Limit:	The name of the data element (CHAN5CNT) in the Capacity Planning resource file that contains SIO counts for channel 5.
Forecast length:	The length of the forecast observation. Eight is selected as a conservative number of estimates to be made from the 12 historical observations that are contained in the CHN resource element file. (You should never specify a forecast longer than the historical data series.)
Confidence limits (percent):	The degree of confidence to be used in calculating the upper and lower confidence limits for the forecasted values of the forecast data element. The default is 95 %.
Time element terms:	The time elements to be used in the regression process. LINEAR is specified for this example.
Specify extended options:	The specification for selecting extended options. Specifying extended options causes the Transform Historical Observations screen to display. In this example, GMA (geometric moving average) is chosen for the Smoothing technique parameter, with 0.5 for the alpha value.

7.6.1.3 Output Reports

We used the control parameters that we discussed in Section 7.6.1.2 to perform Univariate Model Forecasting. The following reports are discussed by section:

- 1 - Model Analysis Report
- 2 - Residual Analysis Report

7.6.1.3.1 Model Analysis Report

Figure 7-17 shows the Model Analysis Report for the channel 5 I/O data. Note that the value of r-squared is a reasonable 0.75. The F value of 30.84 appears to be large and the p-value associated with this F is 0.0004, which is very small. You can interpret this p-value as the probability that you would have estimated this model, given that the data was purely random, and that no linear relationship truly exists. The p-value says that there are no more than four chances in 10,000 that these estimates would arise due to random error.

The model that is shown in the report is in the following form:

$$\text{CHAN5CNT} = 9736797 + 587938 * T1$$

In this model, T1 is the number of the month and CHAN5CNT is the monthly total I/O count for channel 5. This model indicates that you can expect I/O loads to rise by about 588,000 I/Os each month.

CA MICS Capacity Planner							
ANALYSIS OF UNIVARIATE FORECASTING MODEL							
MODEL OF: CHAN5CNT							
BASED ON: LINEAR							
R**2	F	P	INTERCEPT	NAME	COEFFICIENT	T	P
0.75	30.84	.0004	9736797	LINEAR	587938	5.55	.0004

Figure 7-17. Extended Example Model Analysis Report

7.6.1.3.2 Residual Analysis Report

Figure 7-18 shows a sample Residual Analysis Report for the channel I/O study. The regression model that was developed using the smoothed data has a maximum of approximately 10 to 12% difference from the actual historical observations. You must expect this magnitude of residuals, while non-trivial, from a model that uses smoothing to eliminate apparent random variations from the historical data series.

When viewing a Residual Analysis Report, look for obvious patterns in the unexplained data. For example, in Figure 7-18, you can see large positive residuals in the oldest and newest historical observations and large negative residuals in the middle observations. This pattern indicates that the data forms an arc around the regression line. The data is above the line initially, slides down below the line in the winter months, then rises again in April.

If a pattern is identified in the residuals, you should reconsider your choice of model. In this example, a quadratic model may provide more predictive power. When more data is available, try the quadratic model and compare it to the linear model by contrasting the r-squared, F, and p values, as illustrated in Sections 7.4.2.2 through 7.4.2.4, to determine which model is better for a limited number of data points.

A quadratic model may fit the data quite well but may fail to provide useful predictions for future elements, especially if relatively few historical observations are available. Therefore, it is not recommended that you use quadratic models when less than 30 historical observations are available.

CA MICS Capacity Planner								
UNIVARIATE MODELING RESIDUAL ANALYSIS								
UNIVARIATE FORECAST OF: CHAN5CNT								
BASED ON TIME ELEMENTS: LINEAR								
DATE	CHAN5CNT	TRANSFORMED* CHAN5CNT	PREDICTED CHAN5CNT	RESIDUAL CHAN5CNT	-----PLOT OF RESIDUALS-----			CONFIDENCE LIMITS +/-
					-195E4	0.0	+195E4	
30SEP97	11882418	.	10324734	1557683.74		0+++++		1624240.54
30OCT97	12860633	12371525.50	10912672	1947961.05		0+++++		1416789.90
30NOV97	11743581	12302107.00	11500610	242971.35		0++		1221119.32
31DEC97	12510012	12126796.50	12088547	421464.66		0++++		1043874.36
29JAN98	12071368	12290690.00	12676485	-605117.03		-----0		896056.61
26FEB98	12076524	12073946.00	13264423	-1187899		-----0		794269.29
31MAR98	12186976	12131750.00	13852360	-1665384		-----0		757306.06
30APR96	14355185	13271080.50	14440298	-85113.10		0		794269.29
28MAY98	15011392	14683288.50	15028236	-16843.79		0		896056.61
30JUN98	17176448	16093920.00	15616173	1560274.52		0+++++		1043874.36
30JUL98	17421158	17298803.00	16204111	1217046.83		0+++++		1221119.32
31AUG98	18042957	17732057.50	16792049	1250908.14		0+++++		1416789.90
31SEP98	.	.	17379987	.		0		1624240.54
31OCT98	.	.	17967924	.		0		1839490.02
31NOV98	.	.	18555862	.		0		2060095.20
31DEC98	.	.	19143800	.		0		2284505.07
31JAN99	.	.	19731737	.		0		2511700.04
31FEB99	.	.	20319675	.		0		2740987.65
31MAR99	.	.	20907613	.		0		2971883.57
31APR99	.	.	21495550	.		0		3204040.12

* - DATA SMOOTHED USING A GEOMETRIC MOVING AVERAGE. ALPHA = 0.50

Figure 7-18. Extended Example Residual Analysis Report

7.6.2 Performance Group/Period Workload Tracking

Univariate Model Forecasting was used to develop a forecast of transaction arrival rates by performance groups/periods that were identified in the case study discussed in Section 3.3.3.2. Refer to Figure 3-56 to consider the historical observations for Workload 01 (denoted by WKL01).

There is a significant decrease in the transaction arrival rate for the workload in the week of September 11th. The week contains a one-day holiday, thus you can reasonably expect a decrease in the volume of the workloads for the week. Therefore, a linear model was proposed for the workload and deleted the observation for the week of September 11th. Figure 7-19 shows the control parameters that were supplied to produce the model.

```

/----- Univariate Model Forecasting -----\
|Command ==>
|
|Composing CA MICS Inquiry: PRFUVM - Univariate Modeling Inquiry
|
|Report title ==> Univariate Modeling Inquiry
|
|Selection criteria:
|
|           Input file ==> PRF - Performance Group Example
|           Start date ==> _____ (ddmonyyy)
|           SYSID ==> TS01
|           Zone ==> 1_____
|           No. months ==> _____ (1-9999)
|           Save forecast ==> YES      (YES/NO/AGE)
|           Specify CAPAPU Values ==> N      (Y/N)
|Forecast data element      ==> WKL01
|Forecast length (months)   ==> 4      (1-9999)
|Confidence limits (percent) ==> 95     (70/90/95)
|
|Time element terms: Linear  ==> YES     (YES/NO)
|           Quadratic ==> ____ (YES/NO)
|           Cubic      ==> ____ (YES/NO)
|Specify extended options   ==> Y      (Y/N)
\-----/

```

Figure 7-19. Univariate Model Forecasting Control Parameters (Screen 1 of 2)

CA MICS Capacity Planner								
UNIVARIATE MODELING RESIDUAL ANALYSIS								
UNIVARIATE FORECAST OF: WKL01								
BASED ON TIME ELEMENTS: LINEAR								
DATE	WKL01	TRANSFORMED WKL01	PREDICTED WKL01	RESIDUAL WKL01	-----PLOT OF RESIDUALS-----			CONFIDENCE LIMITS +/-
					-2334	0.0	+2334	
24JUL98	11102.50	11102.50	10688.77	413.73		0+++		986.87
31JUL98	11026.70	11026.70	10763.89	262.81		0++		812.25
07AUG98	11232.20	11232.20	10839.02	393.18		0+++		667.92
14AUG98	10096.80	10096.80	10914.14	-817.34		-----0		577.08
21AUG98	10604.80	10604.80	10989.26	-384.46		---0		566.08
28AUG98	10192.30	10192.30	11064.38	-872.08		-----0		639.08
04SEP98	11505.60	11505.60	11139.50	366.10		0+++		772.61
11SEP98	8881.00	.	11214.62	.		0		941.25
18SEP98	11927.80	11927.80	11289.74	638.06		0+++++		1129.39
26SEP98	.	.	11364.86	.		0		1328.77
03OCT98	.	.	11439.98	.		0		1535.01
10OCT98	.	.	11515.11	.		0		1745.69
17OCT98	.	.	11590.23	.		0		1959.37

Figure 7-20. Univariate Forecast of Performance Group/Period Workloads

7.6.3 System Workload and Performance Profile

Univariate Model Forecasting was used to develop a forecast of the multiprocessor CPU busy value that was collected for system TS01 in the system workload and performance profile case study introduced in Section 3.3.3.5. Refer to the Resource Element File Contents Report in Figure 3-92 to consider the historical observations for the CPUMPBS data element. Since the system being studied processes a random arrival testing workload, significant week to week variations around the central trend represented in the data were observed. Therefore, a geometric moving average to smooth the historical data was used so that a linear model for the observations could be developed. Figure 7-21 shows the control parameters used to produce the model.

```

/----- Univariate Model Forecasting -----\
|Command ==>
|
|Composing CA MICS Inquiry: WKLUV - Univariate Modeling Inquiry
|
|Report title ==> Univariate Modeling Inquiry
|
|Selection criteria:
|
|           Input file ==> WKL - System Workload Example
|           Start date ==> _____ (ddmmyyyy)
|           SYSID ==> TS01
|           Zone ==> 1_____
|           No. weeks ==> _____ (1-9999)
|           Save forecast ==> YES      (YES/NO/AGE)
|           Specify CAPAPU Values ==> N      (Y/N)
|Forecast data element ==> CPUMPBS
|Forecast length (weeks) ==> 8      (1-9999)
|Confidence limits (percent) ==> __      (70/90/95)
|
|Time element terms: Linear ==> YES      (YES/NO)
|           Quadratic ==> ____      (YES/NO)
|           Cubic ==> ____      (YES/NO)
|Specify extended options ==> Y      (Y/N)
\-----/

```

Figure 7-21. System Univariate Model Forecasting (Screen 1 of 3)

```

/----- Extended Options Menu -----\
|Command ==>
|
|Inquiry Step: Univariate Model Forecasting
|
|Print Reports ==> YES (YES/NO)
|Maximum R-Square for report ==> 0.55 (0.00-1.00)
|P-Value flag ==> 0.05 (0.000-1.000)
|
|Delete historical data observations ==> N (Y/N/R)
|Transform historical observations ==> N (Y/N/R)
|
\-----/

```

Figure 7-21. System Univariate Model Forecasting (Screen 2 of 3)

```

/----- Transform Historical Observations -----\
|Command ==>|
|
|Inquiry Step: Univariate Model Forecasting
|
|Smoothing technique ==> GMA (GMA/LOG10/USER/NONE)
|
|
|
|Geometric moving average data smoothing parameters:
|      Equation alpha value ==> .5 (.01 - .99)
|      Initial value ==> _____
|
|
|
|User defined data transformation SAS statements:
|  Statements to transform historical data ==> N (Y/N/R)
|  Stmts. to reverse history transformations ==> N (Y/N/R)
|
|
|
|-----\

```

Figure 7-21. System Univariate Model Forecasting (Screen 3 of 3)

Figure 7-22 shows the Residual Analysis Report for the model. As you can see, all of the residual values are within 10 percent of the observed values except for the weeks of July 4th and September 5th. Since these two weeks represent holidays, you can reasonably expect to encounter significant residual values for these weeks.

UNIVARIATE MODELING RESIDUAL ANALYSIS							
UNIVARIATE FORECAST OF: CPUMPBS BASED ON TIME ELEMENTS: LINEAR							
DATE	CPUMPBS	TRANSFORMED* CPUMPBS	PREDICTED CPUMPBS	RESIDUAL CPUMPBS	-----PLOT OF RESIDUALS-----		
					-10	0.0	+10
							CONFIDENCE LIMITS +/-
24APR98	58.67	.	59.92	-1.25		--0	2.44
02MAY98	64.39	61.53	60.26	4.13		0++++++	2.30
09MAY98	63.85	64.12	60.59	3.26		0+++++	2.15
16MAY98	63.32	63.58	60.93	2.39		0++++	2.01
23MAY98	59.10	61.21	61.26	-2.17		---0	1.88
30MAY98	59.90	59.50	61.60	-1.69		---0	1.75
06JUN98	57.88	58.89	61.93	-4.06		-----0	1.63
13JUN98	59.28	58.58	62.27	-2.99		---0	1.52
20JUN98	59.18	59.23	62.61	-3.43		-----0	1.42
27JUN98	66.55	62.87	62.94	3.61		0++++++	1.33
04JUL98	53.33	59.94	63.28	-9.94		-----0	1.26
11JUL98	69.88	61.61	63.61	6.27		0+++++++	1.20
18JUL98	67.23	68.56	63.95	3.28		0+++++	1.17
25JUL98	71.94	69.58	64.28	7.66		0+++++++	1.16
01AUG98	64.10	68.02	64.62	-0.52		-0	1.18
08AUG98	63.35	63.73	64.95	-1.60		---0	1.22
15AUG98	64.48	63.91	65.29	-0.81		-0	1.28
22AUG98	63.33	63.90	65.62	-2.30		---0	1.36
29AUG98	65.86	64.59	65.96	-0.10		0	1.45
05SEP98	76.29	71.08	66.30	10.00		0+++++++	1.56
12SEP98	60.00	68.15	66.63	-6.63		-----0	1.68
19SEP98	66.75	63.38	66.63	0.12		0	1.68
26SEP98	67.29	67.02	66.97	0.32		0	1.80
03OCT98	69.33	68.31	67.30	2.03		0++++	1.93
10OCT98	64.37	66.85	67.64	-3.26		-----0	2.07
17OCT98	70.13	67.25	67.97	2.15		0++++	2.21
24OCT98	.	.	68.31	.		0	2.35
31OCT98	.	.	68.64	.		0	2.50
07NOV98	.	.	68.98	.		0	2.65
14NOV98	.	.	69.31	.		0	2.80
21NOV98	.	.	69.65	.		0	2.96
28NOV98	.	.	69.98	.		0	3.11
05DEC98	.	.	70.32	.		0	3.27
12DEC98	.	.	70.66	.		0	3.42

* - DATA SMOOTHED USING A GEOMETRIC MOVING AVERAGE. ALPHA = 0.50

Figure 7-22. Univariate Model of System CPU Utilization

7.6.4 Forecasting New Application Transaction Volumes

Capacity planners are frequently asked to forecast the potential impact of applications that are still in development. System designers and targeted end-users may supply estimated workload characteristics or transaction volumes based on comparisons to existing systems. You can develop early projections based upon estimates of this type using Univariate Model Forecasting.

In this case study, an existing, inquiry-only online application is to be replaced by a system with both inquiry and update capabilities. It was estimated that during peak hour, inquiry transactions would account for 60% of the new system's transaction volume. It was further estimated that inquiry transactions would increase 15% with installation of the new system, but that overall transaction volume growth would not change. The purpose of this study is to provide an initial projection of the transaction volumes to be expected for the new application during the first three months following its installation.

The estimates indicate that transaction growth should be modeled as a step function, with the step occurring in January 1999. Note that the magnitude of the step is estimated as a function of the transaction volume expected in December 1998. Since no data are available beyond June 1998, two models were developed.

The first model was developed to project transaction volumes for the existing system through December 1998. Figure 7-23 shows the contents of the resource element file that contains monthly peak hour transaction volumes for the existing application. A linear model, which uses all 18 months of historical data was proposed. Figure 7-24 shows the control parameters that we used to produce the model. Figure 7-25 shows the resulting Residual Analysis Report.

CA MICS Capacity Planner						
OBS	DATE	MONTH	CISTRANS	SYSID	ZONE	YEAR
1	31JAN97	1	4043.0	IMS1	1	97
2	28FEB97	2	3890.0	IMS1	1	97
3	31MAR97	3	4037.0	IMS1	1	97
4	30APR97	4	4128.0	IMS1	1	97
5	31MAY97	5	4037.0	IMS1	1	97
6	30JUN97	6	4314.0	IMS1	1	97
7	31JUL97	7	4373.0	IMS1	1	97
8	31AUG97	8	4740.0	IMS1	1	97
9	30SEP97	9	4759.0	IMS1	1	97
10	31OCT97	10	4856.0	IMS1	1	97
11	30NOV97	11	4741.0	IMS1	1	97
12	31DEC98	12	5099.0	IMS1	1	98
13	31JAN98	1	5149.0	IMS1	1	98
14	28FEB98	2	4857.0	IMS1	1	98
15	31MAR98	3	5506.0	IMS1	1	98
16	30APR98	4	5196.0	IMS1	1	98
17	31MAY98	5	5194.0	IMS1	1	98
18	30JUN98	6	5477.0	IMS1	1	98

Figure 7-23. Contents of Transaction Volume Resource Element File


```

/----- Univariate Model Forecasting -----\
|Command ==>
|
|Composing CA MICS Inquiry: UVMW2 - Univariate Modeling Inquiry
|
|Report title ==> Univariate Modeling Inquiry - CICS
|
|Selection criteria:
|
|           Input file ==> CIS - CPU Analysis w/Composite SYSID/ZONE
|           Start date ==> _____ (ddmmyyyy)
|           SYSID ==> IMS1
|           Zone ==> 1_____
|           No. weeks ==> _____ (1-9999)
|           Save forecast ==> YES      (YES/NO/AGE)
|           Specify CAPAPU Values ==> N      (Y/N)
|Forecast data element ==> CISTRANS
|Forecast length (weeks) ==> 6      (1-9999)
|Confidence limits (percent) ==> __      (70/90/95)
|
|Time element terms: Linear ==> YES      (YES/NO)
|           Quadratic ==> ____      (YES/NO)
|           Cubic ==> ____      (YES/NO)
|Specify extended options ==> N      (Y/N)
\-----/

```

Figure 7-24. Control Parameters for Existing System Model

C A M I C S C A P A C I T Y P L A N N E R							
UNIVARIATE MODELING RESIDUAL ANALYSIS							
UNIVARIATE FORECAST OF: CISTRANS							
BASED ON TIME ELEMENTS: LINEAR							
DATE	CISTRANS	TRANSFORMED CISTRANS	PREDICTED CISTRANS	RESIDUAL CISTRANS	-----PLOT OF RESIDUALS-----		
					-301	0.0	+301
							95% CONFIDENCE LIMITS +/-
31JAN97	4043.00	4043.00	3889.39	153.61		0+++++++	156.05
28FEB97	3890.00	3890.00	3983.42	-93.42		-----0	142.91
31MAR97	4037.00	4037.00	4077.45	-40.45		--0	130.33
30APR97	4128.00	4128.00	4171.49	-43.49		--0	118.48
31MAY97	4037.00	4037.00	4265.52	-228.52		-----0	107.62
30JUN97	4314.00	4314.00	4359.55	-45.55		--0	98.07
31JUL97	4373.00	4373.00	4453.58	-80.58		-----0	90.25
31AUG97	4740.00	4740.00	4547.62	192.38		0+++++++	84.63
30SEP97	4759.00	4759.00	4641.65	117.35		0++++++	81.68
31OCT97	4856.00	4856.00	4735.68	120.32		0++++++	81.68
30NOV97	4741.00	4741.00	4829.72	-88.72		-----0	84.63
31DEC97	5099.00	5099.00	4923.75	175.25		0+++++++	90.25
31JAN98	5149.00	5149.00	5017.78	131.22		0+++++++	98.07
28FEB98	4857.00	4857.00	5111.82	-254.82		-----0	107.62
31MAR98	5506.00	5506.00	5205.85	300.15		0+++++++	118.48
30APR98	5196.00	5196.00	5299.88	-103.88		-----0	130.33
31MAY98	5194.00	5194.00	5393.91	-199.91		-----0	142.91
30JUN98	5477.00	5477.00	5487.95	-10.95		0	156.05
31JUL98	.	.	5581.98	.		0	169.62
31AUG98	.	.	5676.01	.		0	183.53
30SEP98	.	.	5770.05	.		0	197.70
31OCT98	.	.	5864.08	.		0	212.08
30NOV98	.	.	5958.11	.		0	226.64
31DEC98	.	.	6052.15	.		0	241.33

Figure 7-25. Univariate Model of Existing Application

The second model was developed to project transaction volumes for the new application. Again, a linear model was proposed, with a user transformation defined to account for the anticipated step function. The magnitude of the step was calculated from the percentage estimates provided and the predicted transaction volume for December 1998 was obtained from the first model. Figure 7-26 shows the control parameters used to produce the model.

```

/----- Univariate Model Forecasting -----\
|Command ==>
|
|Composing CA MICS Inquiry: UVMW3 - Univariate Modeling Inquiry - CICS2
|
|Report title ==> Univariate Modeling Inquiry
|
|Selection criteria:
|
|           Input file ==> CIS - CPU Analysis w/Composite SYSID/ZONE
|           Start date ==> _____ (ddmmyyyy)
|           SYSID ==> IMS1
|           Zone ==> 1_____
|           No. months ==> _____ (1-9999)
|           Save forecast ==> YES      (YES/NO/AGE)
|           Specify CAPAPU Values ==> N      (Y/N)
|Forecast data element      ==> CISTRANS
|Forecast length (weeks)    ==> 9        (1-9999)
|Confidence limits (percent) ==> __      (70/90/95)
|
|Time element terms: Linear  ==> YES      (YES/NO)
|           Quadratic ==> ____      (YES/NO)
|           Cubic      ==> ____      (YES/NO)
|Specify extended options   ==> Y        (Y/N)
\-----/

```

Figure 7-26. Forecasting for New Application Model (Screen 1 of 4)

```

/----- Extended Options Menu -----\
|Command ==>
|
|Inquiry Step: Univariate Model Forecasting
|
|Print Reports                ==> YES      (YES/NO)
|Maximum R-Square for report  ==> 0.55    (0.00-1.00)
|P-Value flag                 ==> 0.05    (0.000-1.000)
|
|Delete historical data observations ==> N (Y/N/R)
|Transform historical observations  ==> N (Y/N/R)
|
\-----/

```

Figure 7-26. Forecasting for New Application Model (Screen 2 of 4)

```

/----- Transform Historical Observations -----\
|Command ==>                                     |
|                                                    |
|Inquiry Step:  Univariate Model Forecasting        |
|                                                    |
|Smoothing technique                               ==> USER  (GMA/LOG10/USER/NONE) |
|                                                    |
|                                                    |
|Geometric moving average data smoothing parameters: |
|      Equation alpha value ==> ___      (.01 - .99) |
|      Initial value       ==> _____ |
|                                                    |
|                                                    |
|User defined data transformation SAS statements:    |
|  Statements to transform historical data   ==> Y (Y/N/R) |
|  Stmts. to reverse history transformations ==> N (Y/N/R) |
|                                                    |
|                                                    |
\-----/

```

Figure 7-26. Forecasting for New Application Model (Screen 3 of 4)

```

/----- SAS System Statements -----\
|Command ==>                                     Scroll ==> PAGE |
|                                                    |
|SAS Statements for the ASGN_TRN Macro              |
|-----|
|***** ***** TOP OF DATA *****|
|==MSG> SPECIFY SAS STATEMENTS TO PERFORM A USER TRANSFORMATION. |
|NOTE= %MACRO ASGN_TRN; <=== PLEASE DO NOT PLACE LINES BEFORE THIS LINE. |
|000001 _TRAN_V_=&VNAME+5569; |
|NOTE= %MEND ASGN_TRN; <=== PLEASE DO NOT PLACE LINES AFTER THIS LINE. |
|***** ***** BOTTOM OF DATA *****|
| |
\-----/

```

Figure 7-26. Forecasting for New Application Model (Screen 4 of 4)

The special variable name `_TRAN_V_` in the SAS statement on the Univariate Model Forecasting screen shown in Figure 7-26 refers to the transformed historical value of the dependent variable, CISTRANS, and the special variable name, `&VNAME`, refers to the value of CISTRANS prior to transformation. Section 7.5.4 describes these special variable names in more detail.

Figure 7-27 shows the Residual Analysis Report for the model. Note that the residuals shown are identical to those in Figure 7-25. This is to be expected, since all historical data is transformed by adding the same constant value. Figure 7-28 shows a plot of the resulting composite forecast.

C A M I C S C A P A C I T Y P L A N N E R							
UNIVARIATE MODELING RESIDUAL ANALYSIS							
UNIVARIATE FORECAST OF: CISTRANS							
BASED ON TIME ELEMENTS: LINEAR							
DATE	CISTRANS	TRANSFORMED CISTRANS	PREDICTED CISTRANS	RESIDUAL CISTRANS	-----PLOT OF RESIDUALS-----		95% CONFIDENCE LIMITS +/-
					-301	0.0	+301
31JAN97	4043.00	9612.00	9458.39	153.61		0+++++++	156.05
28FEB97	3890.00	9459.00	9552.42	-93.42		-----0	142.91
31MAR97	4037.00	9606.00	9646.45	-40.45		--0	130.33
30APR97	4128.00	9697.00	9740.49	-43.49		--0	118.48
31MAY97	4037.00	9606.00	9834.52	-228.52		-----0	107.62
30JUN97	4314.00	9883.00	9928.55	-45.55		---0	98.07
31JUL97	4373.00	9942.00	10022.58	-80.58		-----0	90.25
31AUG97	4740.00	10309.00	10116.62	192.38		0+++++++	84.63
30SEP97	4759.00	10328.00	10210.65	117.35		0+++++	81.68
31OCT97	4856.00	10425.00	10304.68	120.32		0+++++	81.68
30NOV97	4741.00	10310.00	10398.72	-88.72		-----0	84.63
31DEC97	5099.00	10668.00	10492.75	175.25		0+++++++	90.25
31JAN98	5149.00	10718.00	10586.78	131.22		0+++++	98.07
28FEB98	4857.00	10426.00	10680.82	-254.82		-----0	107.62
31MAR98	5506.00	11075.00	10774.85	300.15		0+++++++	118.48
30APR98	5196.00	10765.00	10868.88	-103.88		-----0	130.33
31MAY98	5194.00	10763.00	10962.91	-199.91		-----0	142.91
30JUN98	5477.00	11046.00	11056.95	-10.95		0	156.05
31JUL98	.	.	11150.98	.		0	169.62
31AUG98	.	.	11245.01	.		0	183.53
30SEP98	.	.	11339.05	.		0	197.70
31OCT98	.	.	11433.08	.		0	212.08
30NOV98	.	.	11527.11	.		0	226.64
31DEC98	.	.	11621.15	.		0	241.33
31JAN99	.	.	11715.18	.		0	256.14
28FEB99	.	.	11809.21	.		0	271.05
31MAR99	.	.	11903.24	.		0	286.03

* - DATA SMOOTHED USING USER DEFINED MACRO

Figure 7-27. Univariate Model of New Application

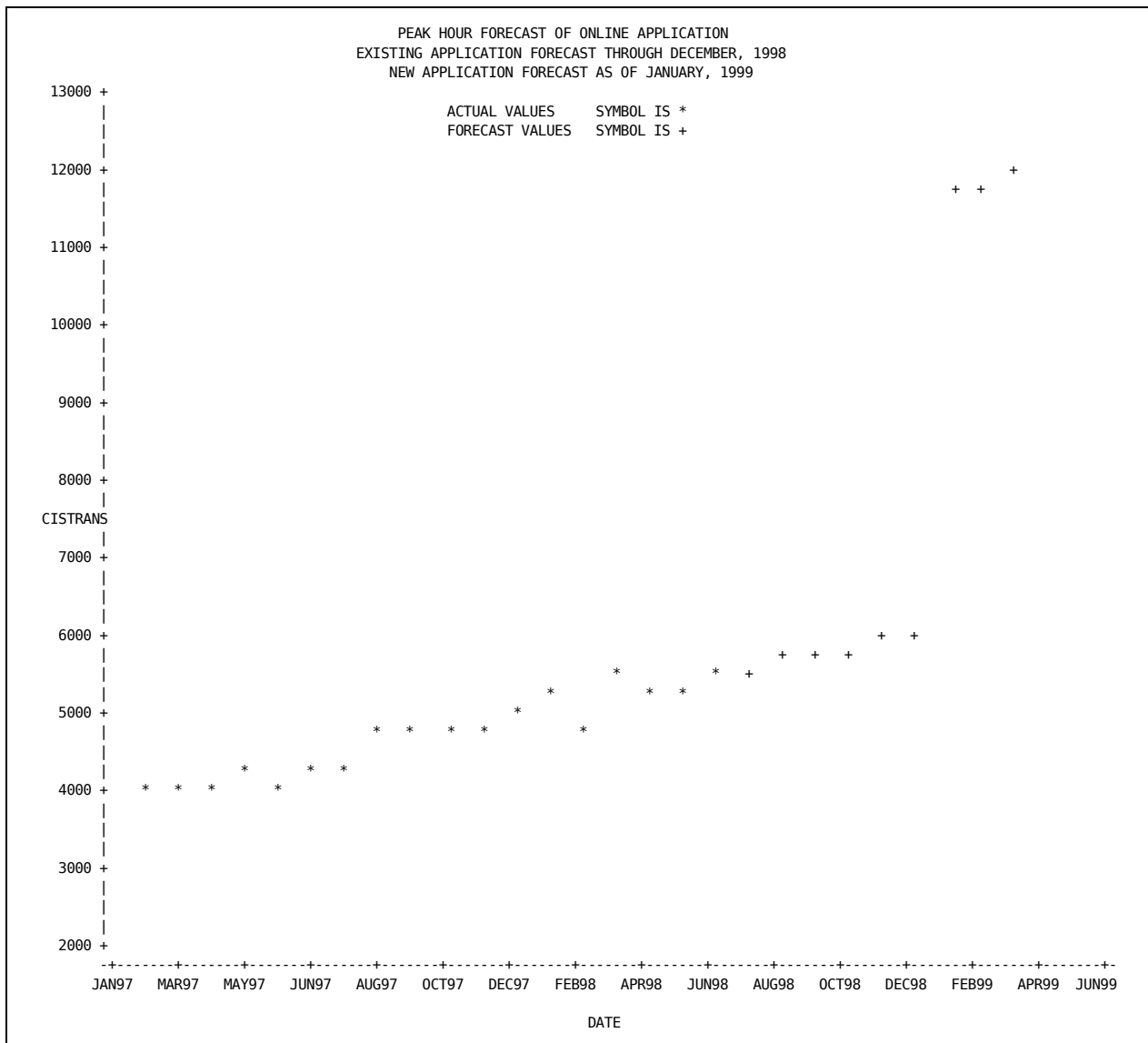


Figure 7-28. Transaction Volume Forecast

Chapter 8: WORKSHEET FACILITY

The Worksheet Facility allows you to generate reports suitable for distribution to your users so they can fill in their projections of future consumption levels. Once these reports have been returned (or filled in by you) you can consolidate the forecasts by entering the data into the Worksheet Entry screen and apply them to your Capacity Planning database. Once applied, facilities are provided to show the effects of these projections on your Capacity Planning actuals data.

The Worksheet Facility enables you to perform these forecasting functions using the following Capacity Planner options.

- o Menu option (3) on the Capacity Planner Main Menu takes you into Worksheet Creation, Modification, Production (Execution), Worksheet Enter, and Database Apply.
- o An extension to the Presentation Graphics Facility places the results of the user forecast on generated graphics.
- o An analytic report, Worksheet Effect Report, shows the effect of applying the Worksheet values to the Capacity Planning database.

The following sections describe the operation and reporting facilities of the Capacity Planner Worksheet Facility.

For information regarding Presentation Graphics, see Chapter 4 of this guide.

This section contains the following topics:

- [8.1 Functional Description](#) (see page 688)
- [8.2 Usage Guidelines](#) (see page 688)
- [8.3 Standard Output](#) (see page 688)
- [8.4 Forecasting Technique Tutorial](#) (see page 693)
- [8.5 Product Operation](#) (see page 694)

8.1 Functional Description

The Worksheet Facility allows you to tailor and generate reports that list historical data on a Worksheet suitable for distribution to users for their input. These worksheets can optionally reflect forecasts that have been created using other analysis facilities of the CA MICS Capacity Planner. Once these worksheets have been filled in, you can enter the information into the Capacity Planner and then apply these forecasts to the Capacity Planning database. You may create worksheets based on any standard application, resource element, meta, or business element file in your Capacity Planning database.

8.2 Usage Guidelines

Often when doing a capacity plan you find yourself having to gather information from your user base on their plans. These user forecasts are then aggregated into a whole which is used to formulate the consumption numbers upon which the capacity plan is based. These consumption values are then compared to current levels to ascertain the need for an upgrade. The Worksheet Facility provides an automated method of generating the reports distributed to your user base (or used for your own work) and then showing the effect of these projections on your capacity planning actuals data in the Capacity Planner database. The Worksheet Effect Reports can be placed into the Production Reporting Facility of MICF to produce these reports on a regular basis so you can track how well the actual values match the projected values.

8.3 Standard Output

The Worksheet Facility produces two different output reports, the worksheet itself and the Worksheet Effect Report. These reports are described in the following sections:

- 1 - Worksheet Report
- 2 - Worksheet Effect Report

8.3.1 Worksheet Report

The Worksheet Report displays information from a CA MICS Capacity Planner file in a suitable format for use as a fill-in-the-blank worksheet and can optionally display forecast information generated by the Forecasting Facilities of the Capacity Planner. The content of this report is defined using the Worksheet Definition Screen shown in Figure 8-5.

Figure 8-1 shows a sample Worksheet generated from the Workload Standard Application. The title lines of the report lists the SYSID, ZONE, CAPAPU(s), and the report title you specify on the Worksheet Definition screen.

The body of the report consists of columns that correspond to the time periods you select with the rows showing the data elements that you selected. Each data element row consists of the actual value selected, the optionally selected forecast lines, and a blank line for filling in a forecast value. A footer section has an Instructions block for any text you wish to place there and a User Comments block for your user to fill in, if they so desire.

Capacity Planning Worksheet Workload Consumption by Customer Area Payroll Department									
	DEC97	JAN98	FEB98	MAR98	APR98	MAY98	JUN98	JUL98	AUG98
Batch Jobs Actuals	7895	6665	34664	.	.	.	.	.	.
Forecast Batch Jobs	7666	8909	9886	10200	11849	14004	15794	18865	21786
User Projection Batch Jobs									
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----									
CICS Transactions Address Updates									
Forecasted Address Updates									
User Projection Address Updates									
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----									
CICS Transactions Employee Adds									
User Projection Employee Adds									
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----									
CICS Transactions Employee Change of Status									
Forecasted Change of Status Trans									
User Projection Change Of Status									
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----									
INSTRUCTIONS: Please update these forecasted values with your anticipated usage values. If you wish to specify a percentage growth rate, place in the first blank space, if you wish to modify each month, place modification in corresponding column of report. If you have any questions, call the Capacity Planning Department at X5000. If you have any comments, please use the space at the left. Thank you. USER COMMENTS: _____ _____ _____ _____ _____ _____									

Figure 8-1. Worksheet Report

8.3.2 Worksheet Effect Report

The Worksheet Effect Report displays information from CA MICS Capacity Planner Files and Forecasts and shows the results of entered worksheet information.

Figure 8-2 shows a sample Worksheet Effect Report generated from the Workload Standard Application. The title lines of the report lists the SYSID, ZONE, CAPAPU(s), and the report title you specify on the worksheet screen.

The body of the report consists of columns that correspond to the time periods you select with the rows showing the data elements that you selected. Each data element row consists of the actual value selected, the optionally selected forecast lines, and a line corresponding to the worksheet values you enter (if present). A footer section has a Comment block for any text you wish to place there.

Capacity Planning Worksheet Effect Report Workload Consumption by Customer Area									
Payroll Department									
	DEC97	JAN98	FEB98	MAR98	APR98	MAY98	JUN98	JUL98	AUG98
Batch Jobs Actuals	7895	6665	34664	.	.	.	.	.	.
Forecast Batch Jobs	7666	8909	9886	10200	11849	14004	15794	18865	21786
User Projection Batch Jobs	9000	9300	10000	11000	13000	15500	17000	23000	25000
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----									
CICS Transactions Address Updates									
Forecasted Address Updates									
User Projection Address Updates									
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----									
CICS Transactions Employee Adds									
User Projection Employee Adds									
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----									
CICS Transactions Employee Change of Status									
Forecasted Change of Status Trans									
User Projection Change Of Status									
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----									
INSTRUCTIONS: Please update these forecasted values with your anticipated usage values. If you wish to specify a percentage growth rate, place in the first blank space, if you wish to modify each month, place modification in corresponding column of report. If you have any questions, call the Capacity Planning Department at X5000. If you have any comments, please use the space at the left. Thank you. USER COMMENTS: _____ _____ _____ _____ _____ _____									

Figure 8-2. Worksheet Effect Report

8.4 Forecasting Technique Tutorial

The Worksheet Facility provides a method of generating documents that can be used as part of the overall information gathering process that leads to the creation of the Capacity Plan. Used in conjunction with the other forecasting facilities of the Capacity Planner, worksheets provide additional insight into the plans of your users and assist in understanding how workloads behave on your systems.

The Worksheet Forecasting process follows a series of steps described as follows:

- o Step one, Worksheet Creation, involves the defining a worksheet based upon a file you choose in the CA MICS Capacity Planning database. This file could be a resource element file, a standard application file, a meta file, or a business element file. Once you have selected a file for input, you can choose the data elements you would like to appear on the report. You can specify any instructions you would like to appear on the report and build subsetting logic if desired.
- o Step two involves producing (executing) a worksheet. Production of the worksheet generates the Worksheet Report (see Figure 8-1) and creates/updates a member in the CA MICS Capacity Planner data set that contains the data you selected (see step four).
- o Step three takes place outside the CA MICS Capacity Planner and involves filling in the worksheets with the information you desire. This process can be done by yourself, through user interviews, or by directly distributing the worksheets to your users. Once the worksheets are completed, you can proceed to step four.
- o Step four involves using the ENTER function of the Worksheet Facility to enter the values from a filled in worksheet into the Capacity Planning data set. Once entered, you can proceed to step five.

- o Once the values have been entered into the Capacity Planning Data Set, you need to APPLY this information to the Capacity Planner database file you specified in the step one. The APPLY function is part of the Worksheet Facility and invokes SAS to take the information in the Capacity Planner Data Set and create the Capacity Planner Worksheet Forecast File (see Chapter 3 for more information on the format of the Capacity Planner database).
- o After the worksheet information has been applied, you can see the effect of your worksheet through the creation and execution of a Worksheet Effect Report. The Effect Report is similar in layout and creation to the worksheet itself, however, you will use the Capacity Planning Analysis option of the Capacity Planner to create/modify a Worksheet Effect Report. This new report is option 7 on the Workload Forecasting main menu.

8.5 Product Operation

This section describes the screens and parameters that are used to control the Worksheet Facility. It is divided into two sections that describe the creation of worksheets and the creation of Worksheet Effect Reports:

- 1 - Worksheet Facility
- 2 - Worksheet Effect Reports

8.5.1 Worksheet Facility

This section describes how to create new worksheets, modify worksheets, produce a worksheet, enter the filled-in worksheets, and apply these values to the Capacity Planning database. These operations are described in the following sections:

- 1 - Creating a Worksheet
- 2 - Modifying a Worksheet
- 3 - Producing a Worksheet
- 4 - Entering a Worksheet
- 5 - Applying a Worksheet

8.5.1.1 Creating a Worksheet

```

/----- Capacity Planner Worksheet Functions -----\
|Command ==>                                     Scroll ==> CSR |
|
| Commands: Apply, Catalog, Create, End, Enter, Help, Modify, Produce, Return, |
|           Select                                                                    |
| Command  Num  Short  Function                                                         |
| Name      Name  Name  Name                                                            |
|-----|-----|-----|-----|
| 1      WRK  WORKSHEET                                                                    |
|          Create a new capacity planning worksheet                                    |
|
| 2      ENT  ENTER                                                                    |
|          Enter values from a worksheet into the apply data set                      |
|
| 3      APP  APPLY                                                                    |
|          Apply entered values to capacity planning actuals                        |
|
| 4      MDW  MODWORK                                                                    |
|          Modify an existing capacity planning worksheet                            |
|
| 5      PRW  PROWORK                                                                    |
|          (Re)Produce an existing capacity planning worksheet                      |
|
|***** BOTTOM OF DATA *****|
\-----/

```

Figure 8-3. Worksheet Function Menu

The Capacity Planner Worksheet Function Menu, shown in Figure 8-3, is displayed the first time that you enter the Worksheet Facility of the Capacity Planner or whenever you issue the MENU Primary Command. This scrollable display allows you to perform any of the functions described in Section 8.4. The command list displayed is the same whether you use the Primary Command Field (the top of the screen) or the Line Command Fields (in front of each row of the scrollable display).

A short description of each option available and the alias follows. Commands available from the Primary Command Line are shown in uppercase (for example, CREATE); Line Commands are shown with an initial capital (for example, Insert):

- o CREATE (1, WRK, WORKSHEET, Insert; or Select on line 1)--This command allows you to create a new Capacity Planner Worksheet and takes to you the screen shown in Figure 8-4. You may specify the name of the worksheet when entering the CREATE command on the Primary Command Line (for example, CREATE TEST01).
- o ENTER (2, ENT, ENTER; or Select on line 2)--Once a worksheet has been filled out, you need to enter the information back into the Capacity Planner. After which, you apply this data to the Capacity Planner database file. See Section 8.5.1.4 for more information on ENTERing worksheet values.
- o APPLY (3, APP, APPLY; or Select on line 3)--After the values have been ENTERed from a worksheet you need to APPLY this information to the Capacity Planner database for further analysis and for the creation of Worksheet Effect Reports.
- o MODIFY (4, MDW, MODIFY; or Select on line 4)--If you wish to make changes to an existing worksheet, you can use this command to modify an existing worksheet. For more information on the MODIFY process, see Section 8.5.1.2.
- o PRODUCE (5, PRO, PROWORK; or Select on line 5)--This option takes you into worksheet production. Worksheet production will execute a SAS job that will use the worksheet definition you defined to PRODUCE the Worksheet Report (see Section 8.3.1 for a description of the Worksheet Report). For more information on worksheet production, see Section 8.5.1.3.

Besides the options listed previously there are several other commands that you can use on this screen. They are:

- o HELP--The HELP command will bring up the tutorial screens that describe this screen. Help as a Line Command will provide a Help screen describing the individual option.
- o END--The END command returns you to the Capacity Planner Main Menu.

- o RETURN--The RETURN command also returns you to the Capacity Planner Main Menu.
- o CATALOG--If you have created at least one worksheet, this command will display the Worksheet Catalog screen. For an example of this screen, see Figure 8-9 in Section 8.5.1.2.

[illegible]

Figure 8-4. Creating A New Worksheet

If, on the previous screen, you chose option 1 (or one of its aliases), you would see the screen in Figure 8-4. This is where you specify the name of the worksheet, the title displayed on the Worksheet Catalog Display, and the author's name. Once you have filled out all the fields to your satisfaction, enter END to continue with the Worksheet Definition (see Figure 8-5). You can enter CANCEL to cancel this worksheet definition.

```

----- Worksheet Definition -----
|Command ==>                               Scroll ==> PAGE
|
|  Creating Worksheet Inquiry:  TEST01 - Testing Worksheets
|
|Report Title ==> _____
|Selection criteria:
|
|          Input file ==> ____ (fff) - RESOURC ELEM FILE WITH CA
|          SYSID ==> ____
|          HOUR ==> __
|          Specify CAPAPU values ==> _ (Y/N)
|
|Number of previous weeks      ==> ____ (1-9999)
|Number of future weeks        ==> ____ (1-9999)
|Specify Extended Options      ==> N (Y/N)
|
|
|      Name      Label Line One  Label Line Two  Output Format  Include Forecast
|      -----
|      _____
|      _____
|      _____
|      _____
|      _____
|      _____
|      _____
|
|-----

```

Figure 8-5. Worksheet Definition

This is the main definition screen for a worksheet. Here you enter such items as the input file, selection criteria, the amount of data to be displayed, the report title, any extended options, and the data elements you would like to appear on the Worksheet Report.

The following is a description of the parameters on this screen and how they can be used to tailor the worksheet to your requirements:

- o Report Title--The title that will appear on the worksheet.
- o Input File--Required. This is the Capacity Planner file to be used as input to your worksheet. It can be any file type in the CA MICS Capacity Planner. If you do not know the name of the file you wish to use, enter a '?' mark in this field to get a list of files to choose from.

- o SYSID--Required if the input is not a business element file. This is the SYSID that you would like on the worksheet.
- o Hour--The Zone, Hour or Dayname from the input file you would like on the worksheet. This is a required field for all input files except for a business element file, in which case it is not used.
- o Specify CAPAPU values--Put a Y in this field to specify the CAPAPU value(s) to be included in the analysis. The CAPAPU value is specified on the CAPAPU Selection Values screen, shown in Figure 8-6, that appears if you enter a Y in this field. You may specify it only if the Capacity Planning database file is defined as containing CAPAPU(s).
- o Number of previous weeks--The number of days, weeks, or months (depending on the specification of the input file) of actuals data you would like on the Worksheet.
- o Number of future weeks--The number of days, weeks, or months (depending on the specification of the input file) of user forecast fields you would like on the worksheet.
- o Specify Extended Options--A Y in this field will take you to the Extended Options Menu (see Figure 8-7 for an example of this screen). The Extended Options allow you to optionally add a list of instructions to the worksheet and/or write your own SAS code to further subset the input file.

The lower portion of the control screen contains a scrollable data entry display in which you may specify the data elements to be included in the worksheet. One line represents one line on the report, blank lines will not appear on the report. The following fields are available for each row in the display area:

- o Name--The name of the data element from the input file. You may specify a question mark '?' in this field to get a scrollable list of data elements on the file. You can then select the data elements you wish to have on the worksheet.
- o Label Line One--The first line of label information to appear on the worksheet (see Section 8.3.1). This defaults to the data element name. No double quotes (") or semicolons (;) may appear in this field.

- o Label Line Two--The second line of label information to appear on the worksheet (see Section 8.3.1). This defaults to blank. No double quotes (") or semicolons (;) may appear in this field.
- o Output Format--The output format that is used to print the values for this data element on the worksheet. The report field is 12 columns wide, so any value that you enter here should specify a field width of 12, or the values will not be right-justified. If you did specify a format value for the data element when the file was created, this value is supplied as the default. If you do not specify a value, then the Worksheet Facility will default to a value of 12. for numeric values, and TIME12.2 for time values. Character values are not allowed.
- o Include Forecast--If you have used the analytic facilities of the Capacity Planner on this file (Univariate Modeling, Multivariate Regression, and Business Element Forecasting) and would like those forecasted values to appear on this report, specify Yes to include them in the worksheet. Otherwise use No, the default, to block these values from the worksheet.

Once you have defined all of the fields to your satisfaction, you can END from this screen and the Worksheet Definition will be saved in your Worksheet Catalog. If you would like to quit Worksheet Definition without actually creating a worksheet, the CANCEL Primary Command will return you to either the Worksheet Catalog or the Worksheet Function Menu. There you may PRODUCE the newly created worksheet, CREATE a new worksheet, or leave the Worksheet Facility.

If you specified Y on the Worksheet Definition screen indicating that you want to specify CAPAPU values, the CAPAPU Selection Value screen appears as shown in Figure 8-6. You may specify it only if the Capacity Planning database file is defined as containing the CAPAPU data element. This value can be made up of the following when your input file contains a CAPAPU:

- * - Select all values of CAPAPU for this ZONE/HOUR/DAYNAME. May be the only character specified, but must be at the end of any other values specified.
- ? - Select all characters that fall in this position of the CAPAPU.

a - An explicit character other than ? and *.

```

/----- CAPAPU Selection Values -----\
|Command ==>
|
|
|Composing CA MICS Inquiry: _____
|
|      Rule  ---+---1---+---2---+---3---+---4---+---5---+---6--
|CAPAPU    ==> -----
|CAPAPU 2  ==> -----
|CAPAPU 3  ==> -----
|CAPAPU 4  ==> -----
|CAPAPU 5  ==> -----
|CAPAPU 6  ==> -----
|CAPAPU 7  ==> -----
|CAPAPU 8  ==> -----
|CAPAPU 9  ==> -----
|
|
\-----/

```

Figure 8-6. CAPAPU Selection Values Screen

```
/----- CAP File Processing Extended Options Menu -----\
|Command ===>
|
|Worksheet Name: WORKSHEET CREATION
|
|Specify instructions for the worksheets.           ==> N (Y/N/R)
|
|Specify file selection criteria.                   ==> N (Y/N/R)
|    (file subsetting exit)
|
|
|
|
|
|
|
|
```

Figure 8-7. CAP File Processing Extended Options

Figure 8-7 shows the CAP File Processing Extended Options Menu that is displayed if you specify a Y (YES) for the Specify Extended Options parameter on the Worksheet Definition screen shown in Figure 8-5.

The following options are available:

- o Specify instructions for the worksheets--The specification for placing instruction lines at the bottom of the worksheet reports. See Figure 8-8 to see the screen that is displayed. Specify 'Y' to place instructions on the worksheet; specify 'R' to reset (clear) any instructions that are already defined.
- o Specify file selection criteria--The specification for building a SAS Macro to subset the input Capacity Planner file. See Figure 8-9 to see the screen that is displayed. Specify a 'Y' to create/update the macro. Specify an 'R' to reset (clear) any macro that may be there.

When you are finished with this screen, END to return to the Worksheet Definition screen (see Figure 8-5).

```

/----- Worksheet Instructions -----\
|Command ==>                               Scroll ==> CSR |
|
|Specify instructions that will appear at the bottom of the worksheet page. |
|
|Line Instructions
| - -----
| 1 -----
| 2 -----
| 3 -----
| 4 -----
| 5 -----
| 6 -----
| 7 -----
|
|
|
|
|
|
|
|
|
|
\-----\

```

Figure 8-8. Worksheet Instructions

Figure 8-8 shows the Worksheet Instructions screen. You can specify up to seven lines of instructions for the worksheet. To save your instructions without leaving this screen, enter the SAVE command. To void any changes made to this screen, enter the CANCEL command. To leave this screen and save any changes you made, enter the END command.

```

/----- SAS System Statements ----- COLUMNS 001 072--\
|Command ==>                               Scroll ==> PAGE |
|
|SAS Statements for the SELECT Macro          |
|***** TOP OF DATA *****|
|=MSG> SPECIFY SAS STATEMENTS TO SUBSET THE CAPACITY PLANNING DATA BASE FILE. |
|=NOTE= %MACRO SELECT; <=== PLEASE DO NOT PLACE LINES BEFORE THIS LINE.      |
|000001                                     |
|000002                                     |
|000003                                     |
|000004                                     |
|000005                                     |
|000006                                     |
|000007                                     |
|000008                                     |
|000009                                     |
|000010                                     |
|000011                                     |
|000012                                     |
|000013                                     |
|000014                                     |
|=NOTE= %MEND SELECT; <=== PLEASE DO NOT PLACE LINES AFTER THIS LINE.        |
|***** BOTTOM OF DATA *****|
\-----/\

```

Figure 8-9. Worksheet Subsetting SAS Statements

Figure 8-9 shows the SAS System Statements screen that you use to code a SELECT macro that will be used to subset the input Capacity Planner file. This is a PDF Edit screen, so you can use any of the standard ISPF commands. To leave this screen and save any changes you made, enter the END command.

8.5.1.2 Modifying a Worksheet

This section of the Product Operation section describes how to use the Worksheet Catalog and modify an existing Worksheet definition.

```

/----- Capacity Planner Worksheet Catalog -----\
|Command ==>                                     Scroll ==> CSR |
|                                                    |
|Commands: Apply, Create, End, Enter, Help, Menu, Modify, Produce, Return, |
|           Select                                   |
|                                                    |
|Command  Name                Title                File   Status   |
|-----|-----|-----|-----|-----|-----|
|_____ BEF001 Business Element File - Invoice File   BF1   Produced 08AUG97 |
|MODIFY  TEST01 Testing Worksheets - CPU Stnd. Appl.   C10   Modified 25JUL97 |
|_____ WRK001 Worksheet Built From WKL Stnd. Appl.   Wk1   Entered 19SEP97 |
|***** BOTTOM OF DATA *****|
|                                                    |
|                                                    |
|                                                    |
|                                                    |
|                                                    |
|                                                    |
|                                                    |
|                                                    |
|                                                    |
|                                                    |
|_____|
\-----/

```

Figure 8-10. Capacity Planner Worksheet Catalog

The Capacity Planner Worksheet Catalog appears after you have created worksheet definition and any time you enter the CATALOG Primary Command when displaying the Worksheet Function screen (see Figure 8-3). In this case, you are going to modify the worksheet definition for TEST01.

The Worksheet Catalog is a scrollable display showing the names of created worksheets, the title of the worksheet definition, the input Capacity Planner file, and worksheet status. Similar to the Worksheet Function Menu many commands can be entered either on the Primary Command Line or as line commands.

The commands available on this screen have some characteristics in common: when used on the Primary Command Line you can specify the worksheet name at the same time, for example, CREATE TEST02; all commands entered as line commands, except Insert, operate on the worksheet catalog line the cursor is on. A short description of the worksheet commands follows:

- o MODIFY (or Modify)--Used to update an existing worksheet.
- o CREATE--Used to create a new worksheet definition.
- o PRODUCE (or Produce)--Used to produce a worksheet (that is, run a job to create the Worksheet Report).
- o ENTER--Used to enter the values on a worksheet into the Capacity Planner in anticipation of applying them to the Capacity Planner database.
- o APPLY--Takes the information you entered for the worksheet and creates a SAS file that will be placed in your Capacity Planner database (see Section 3.1 for a description of this file).
- o INSERT--An alias for CREATE.
- o DELETE--Used to delete a worksheet definition from the Worksheet Catalog.

The following commands are also available on this screen:

- o HELP--Used to see the HELP screen associated with this screen.
- o MENU--Used to display the Worksheet Function Menu. See Section 8.5.1.1 for a description of this menu.
- o END--Used to return to the Capacity Planner Main Menu.
- o RETURN--An alias for END.

```

/----- Worksheet Definition -----\
|Command ==>                               Scroll ==> PAGE|
|
| Modifying Worksheet Inquiry:  TEST01 - Testing Worksheets - CPU Stnd. Appl. |
|
|Report Title ==> Testing the Worksheet Facility - CPU Activity Tracking|
|Selection criteria:|
|      Input file ==> C10          (fff) - RESOURCE ELEM FILE WITH CA|
|      SYSID ==> SYS1|
|      ZONE ==> 1|
|      Specify CAPAPU values ==> N      (Y/N)|
|
|Number of previous weeks    ==> 12      (1-9999)|
|Number of future weeks     ==> 24      (1-9999)|
|Specify Extended Options    ==> Y      (Y/N)|
|
|  Name      Label Line One  Label Line Two  Output Format  Include Forecast|
|-----|-----|-----|-----|-----|
| CPUPCBSY   CPU Percent     Busy           MAPCT.         NO|
| CPUADJPC   CPU Adjusted    PCBUSY    600S    MAPCT.         NO|
|-----|-----|-----|-----|-----|
|-----|-----|-----|-----|-----|
|-----|-----|-----|-----|-----|
|-----|-----|-----|-----|-----|
|-----|-----|-----|-----|-----|
\-----\

```

Figure 8-11. Modifying a Worksheet Definition

Figure 8-11 shows the Worksheet Definition screen with the values filled in from the created worksheet. Once you make the desired changes, you can enter END to save the worksheet or, if you make a mistake, you can enter CANCEL to lose the modifications.

```
|----- Capacity Planner Worksheet Catalog -----|
|Command ==>                                         Scroll ==> CSR|
|
|Commands: Apply, Create, End, Enter, Help, Menu, Modify, Produce, Return,
|          Select
|
|Command      Name              Title                  File       Status
|-----|-----|-----|-----|-----|
|_____ BEF001 Business Element File - Invoice File   BF1    Produced 10AUG91
|_____ TEST01 Testing Worksheets - CPU Stnd. Appl.   C10    Modified 27JUL91
|_____ WRK001 Worksheet Built From WKL Stnd. Appl.   Wk1    Entered 21SEP91
|***** BOTTOM OF DATA *****|
```

Figure 8-12. Capacity Planner Worksheet Catalog after Modification

After the modifications are complete and END is entered on the Worksheet Definition screen, you again see the Worksheet Catalog with worksheet TEST01 showing a new modified date.

8.5.1.3 Producing a Worksheet

This section describes how to produce a worksheet. The PRODUCE function will invoke MICF execution facilities to generate an inquiry that can be executed. This process creates the Worksheet Report and member in the Capacity Planner data set that contains the data you requested. Once the worksheets have been filled in, you can use the ENTER function of the Worksheet Facility to enter these values into the Capacity Planner data set for later APPLY processing.

```
/----- Capacity Planner Worksheet Catalog -----\
|Command ==>                                         Scroll ==> CSR|
|
|Commands: Apply, Create, End, Enter, Help, Menu, Modify, Produce, Return,
|          Select
|
|Command   Name                               Title                               File    Status
|-----|
|_____ BEF001 Business Element File - Invoice File      BF1 Produced 08AUG97
|PRODUCE TEST01 Testing Worksheets - CPU Std. Appl.       C10 Modified 25JUL97
|_____ WRK001 Worksheet Built From WKL Std. Appl.        Wk1 Entered 19SEP97
|***** BOTTOM OF DATA *****|
|
|
|
|
|-----|
```

Figure 8-13. Worksheet Catalog.

Figure 8-13 shows the Worksheet Catalog with 'PRODUCE' entered in front of the Worksheet Definition TEST01. Entering 'PRODUCE TEST01' on the Command Line has the same effect. The Capacity Planner verifies that the worksheet definition can be produced (that is, is executable) and then invokes the MICF Execution Facility to produce the worksheets and update the Capacity Planner Data Set.

```

/----- MICF Execution -----\
|Command ==>|
|
|Executing CA MICS Inquiry: TEST01 - Testing Worksheets - CPUStd. Appl.|
|
|Inquiry Execution Mode    ==> BATCH      (BATCH/FOREGROUND)|
|Override Execution Options ==> N        (Y/N)|
|SAS/DMS Options: Resume  ==> Y          (Y/N)   Output Replay ==> Y (Y/N)|
|
|Press ENTER to execute the inquiry, or press END to cancel execution.|
|
|----- Batch Job Submission Parameters -----|
|Edit Generated Jobstream%==> N (Y/N)  Hold Inq. Output for Replay ==> N (Y/N)|
|
|SYSOUT Class  ==> A                               SYSOUT Form %==> ____|
|Destination  %==> ____                           SYSOUT Writer%==> ____|
|SYSOUT Parms.%==> _____|
|
|Job Card Information:|
|  ==> //jobname JOB (accounting information),|
|  ==> // MSGLEVEL=(1,1),MSGCLASS=A,PRTY=5|
|  ==> /*|
|  ==> /*|
|
\-----/

```

Figure 8-14. CA MICS Worksheet Inquiry Execution Screen

For more information about the CA MICS Worksheet Inquiry Execution Screen shown in Figure 8-14, see Section 1.5 of this guide.

```

/----- Capacity Planner Worksheet Catalog -----
|Command ==>                                     Scroll ==> CSR |
|
|Commands:  Apply, Create, End, Enter, Help, Menu, Modify, Produce, Return,
|           Select
|
|Command    Name                               Title                               File    Status
|-----
|_____ BEF001 Business Element File - Invoice File    BF1  Produced 08AUG97
|_____ TEST01 Testing Worksheets - CPU Std. Appl.    C10  Produced 25JUL97
|_____ WRK001 Worksheet Built From WKL Std. Appl.    Wk1  Entered 19SEP97
|***** BOTTOM OF DATA *****
|
|
|
|
|
|
|-----

```

Figure 8-15. Worksheet Catalog. After Worksheet Production.

Once the worksheet is produced, you are returned to the Worksheet Catalog and the Status Field is updated to reflect the change.

Figure 8-16 shows the Capacity Planner Worksheet Catalog screen, displaying all the available worksheets.

[illegible]

Figure 8-17. Selecting Nine CAPAPU Values

Figures 8-17 and 8-18 shows the screen sequence when you are selecting nine CAPAPUs.

```

/-----\
|      Entering Worksheet Inquiry : WK1R01 - WORKSHEET WITH 9 CAPAPUS      |
|Command ==>                               Scroll ==> PAGE                |
|                                                                              |
|Capacity Planner File: R01 - RESOURCE ELEM FILE WITH CAPAPU              |
|                                                                              |
|Selected                                                                    |
|  CAPAPU   : 0                                                             |
|  CAPAPU2  : 1                                                             |
|  CAPAPU3  : PG000000                                                      |
|  CAPAPU4  : CAPAPU4                                                       |
|  CAPAPU5  : CAPAPU5                                                       |
|  CAPAPU6  : CAPAPU6                                                       |
|  CAPAPU7  : CAPAPU7                                                       |
|  CAPAPU8  : CAPAPU8                                                       |
|  CAPAPU9  : CAPAPU9                                                       |
|                                                                              |
|                                                                              |
|      SYSID   HOUR      DATE      VARIABLE   ACTUAL   WORKSHEET   |
|      -----  -----  -----  -----  -----  -----  |
|      S008    10        06FEB1998 PPATCBTM   0:16:29.68  -----  |
|      S008    10        13FEB1998 PPATCBTM   0:14:15.31  -----  |
|      S008    10        20FEB1998 PPATCBTM   0:14:05.62  -----  |
|      S008    10        27FEB1998 PPATCBTM   0:15:57.92  -----  |
|      S008    10        06MAR1998 PPATCBTM   0:12:45.77  -----  |
|                                                                              |
\-----/

```

Figure 8-18. Worksheet Inquiry With Nine CAPAPUs

```
/----- Select CAPAPU Values for Entry -----\  
|Command ==>                                     Scroll ==> PAGE |  
|  
|Worksheet Name: WK1R02 - WORKSHEET WITH 8 CAPAPUS |  
|  
|  
|CMD Select the CAPAPU(S) to which you wish to enter worksheets   Entered  APU|  
|-----  
| S 0                                                                1 |  
|  1                                                                2 |  
|  PG000000                                                                3 |  
|  CAPAPU4                                                                4 |  
|  CAPAPU5                                                                5 |  
|  CAPAPU6                                                                6 |  
|  CAPAPU7                                                                7 |  
|  CAPAPU8                                                                8 |  
|-----  
\  
/
```

Figure 8-19. Selecting Eight CAPAPU Values

Figures 8-19 and 8-20 show the screen sequence when you are selecting 8 CAPAPUs.

```

/-----\
|      Entering Worksheet Inquiry : WK1R02 - WORKSHEET WITH 8 CAPAPUS      |
|Command ==>                                                                    Scroll ==> PAGE|
|
|Capacity Planner File: R02 - RESOURCE ELEM FILE WITH CAPAPU
|
|Selected
|  CAPAPU   : 0
|  CAPAPU2  : 1
|  CAPAPU3  : PG000000
|  CAPAPU4  : CAPAPU4
|  CAPAPU5  : CAPAPU5
|  CAPAPU6  : CAPAPU6
|  CAPAPU7  : CAPAPU7
|  CAPAPU8  : CAPAPU8
|
|
|      SYSID   HOUR      DATE      VARIABLE   ACTUAL   WORKSHEET
|      -----
|      S008    10        13FEB1998 PPATCBTM   0:14:15.31  -----
|      S008    10        20FEB1998 PPATCBTM   0:14:05.62  -----
|      S008    10        27FEB1998 PPATCBTM   0:15:57.92  -----
|      S008    10        06MAR1998 PPATCBTM   0:12:45.77  -----
|      S008    10        13MAR1998 PPATCBTM   0:14:22.87  -----
|      S008    10        20MAR1998 PPATCBTM   0:13:23.37  -----
|
\-----/

```

Figure 8-20. Worksheet Inquiry With Eight CAPAPUs

```

/----- Worksheet Entry -----\
|Command ==>                               Scroll ==> CSR |
|
|  Entering Worksheet Inquiry:  TEST02 - Further Batch Activity Projections
|
|  Capacity Planner File: BA2 - Batch Activity File - System Level
|
|
|      VARIABLE      ACTUAL      WORKSHEET
|  Cmd SYSID  HOUR   DATE      NAME      VALUE      VALUE
|  -  - - - -  - - -  - - - - -  - - - - -  - - - - -
|  -  S008    10    31JAN1998  AVRAGBAT    2360    -----
|  -  S008    10    28FEB1998  AVRAGBAT    3465    -----
|  -  S008    10    31MAR1998  AVRAGBAT    5665    -----
|  -  S008    10    30APR1998  AVRAGBAT    7571    -----
|  -  S008    10    31MAY1998  AVRAGBAT    8774    -----
|  -  S008    10    30JUN1998  AVRAGBAT    -----
|  -  S008    10    31JUL1998  AVRAGBAT    -----
|  -  S008    10    31AUG1998  AVRAGBAT    -----
|  -  S008    10    30SEP1998  AVRAGBAT    -----
|  -  S008    10    31OCT1998  AVRAGBAT    -----
|  -  S008    10    30NOV1998  AVRAGBAT    -----
|  -  S008    10    31DEC1998  AVRAGBAT    -----
|  -  S008    10    31JAN1999  AVRAGBAT    -----
|  -  S008    10    28FEB1999  AVRAGBAT    -----
|***** BOTTOM OF DATA *****|
\-----/

```

Figure 8-21. Entering Worksheet Values - Resource Element File

Figure 8-21 shows the Worksheet Entry screen for Resource Element files.

```

/----- Worksheet Entry -----\
|Command ==>                               Scroll ==> CSR |
|
|  Entering Worksheet Inquiry:  TEST03 - Process INV BEF File
|
|  Business Element File: INV - Invoice Files
|
|      VARIABLE      ACTUAL      WORKSHEET
|Cmd  DATE          NAME        VALUE      VALUE
|-----
|  _  31JAN1998  ORDERS          859839  -----
|  _  28FEB1998  ORDERS         1087935  -----
|  _  31MAR1998  ORDERS         1349796  -----
|  _  30APR1998  ORDERS         1589798  -----
|  _  31MAY1998  ORDERS         1770987  -----
|  _  30JUN1998  ORDERS          -----
|  _  31JUL1998  ORDERS          -----
|  _  31AUG1998  ORDERS          -----
|  _  30SEP1998  ORDERS          -----
|  _  31OCT1998  ORDERS          -----
|  _  30NOV1998  ORDERS          -----
|  _  31DEC1998  ORDERS          -----
|  _  31JAN1999  ORDERS          -----
|  _  28FEB1999  ORDERS          -----
|***** BOTTOM OF DATA *****|
\-----/

```

Figure 8-22. Entering Worksheet Values - Business Element File

Figure 8-22 shows the Worksheet Entry screen for Business Element files.

8.5.1.5 Applying a Worksheet

This section describes how to apply the entered worksheet values to the Capacity Planner database.

[illegible]

Figure 8-23. Applying Worksheet Values Catalog Display.

[illegible]

Figure 8-24. Applying Worksheet Values Foreground SAS Invocation.

8.5.2 Worksheet Effect Reports

This section describes the screen and parameters that are used to control the production of Worksheet Effect Reports. The control parameters allow you to define the Capacity Planning database file to be used as input, the time periods, and the variables to be placed on the report. In addition, you may subset the input data or tailor the contents of the reports produced.

To get to the Worksheet Effect Report screen (shown in Figure 8-25) on your terminal, you must follow a short set of menus which are more fully described in Section 1.5 of this guide. The Worksheet Effect Report control screen appears after you indicate that you want to insert a new Worksheet Effect Report inquiry into your catalog of analysis inquiries. Two fields are filled on the fourth line when you first see it: the inquiry name and the inquiry description, separated by a dash. You enter these two fields on the Insert Capacity Planning Analysis screen when you request that a new inquiry be added.

```

/----- Worksheet Report ----- Row 1 of 63\
|Command ==>                               Scroll ==> CSR |
|Enter a ? in any data entry field for more information on valid values. |
|Composing CA MICS Inquiry: WRKREP - Worksheet Testing - Effect Report |
| |
|Creating Worksheet Inquiry: WK1R01 - WORKSHEET WITH 9 CAPAPUS |
| |
|Report Title ==> Worksheet Testing - Effect Report |
|Selection criteria: |
| |
|           Input file ==> ___ (fff) - |
|           SYSID ==> ____ |
|           HOUR ==> __ |
|           Specify CAPAPU values ==> _ (Y/N) |
| |
|Number of previous weeks ==> ____ (1-9999) |
|Number of future weeks ==> ____ (1-9999) |
|Specify Extended Options ==> N (Y/N) |
| |
|   Name   Label Line One Label Line Two Output Format Include Forecast |
|-----|-----|-----|-----|-----|
| PPASRBTM PPASRBTM _____ TIME12.2 NO |
| PPATCBTM PPATCBTM _____ TIME12.2 NO |
|-----|-----|-----|-----|-----|
|-----|-----|-----|-----|-----|
|-----|-----|-----|-----|-----|
|-----|-----|-----|-----|-----|
|-----|-----|-----|-----|-----|
\-----/

```

Figure 8-25. Worksheet Effect Report Control Screen

Some of the parameters in Figure 8-25 are required, while others are optional. The following is a discussion of the various parameters as they appear in this figure.

Report title: The 60-character title field for the Worksheet Effect Report. No double quotes (") or semicolons (;) may appear within the body of the title.

Input file: The resource element or business element Capacity Planning database file that contains the historical observations to be used as input to the report. If you are uncertain of the files or file names available in your Capacity Planning database, type a question mark (?) in the field and a selection list will be displayed.

SYSID:	The SYSID for which the report is to be produced. If the Capacity Planning database file was defined as having a composite SYSID, this value is automatically supplied on the screen.
Hour:	The Zone, Hour or Dayname from the input file you would like on the worksheet. This is a required field for all input files except for a business element file, in which case it is not used.
Specify CAPAPU values:	Put a Y in this field to specify the CAPAPU value(s) to be included in the analysis. The CAPAPU value is specified on the CAPAPU Selection Values screen, shown in Figure 8-26, that appears if you enter a Y in this field. You may specify it only if the Capacity Planning database file is defined as containing the CAPAPU data element.
Number of previous weeks:	The number of days, weeks, or months (depending on the specification of the input file) of actual data you would like on the Worksheet.
Number of future weeks:	The number of days, weeks, or months (depending on the specification of the input file) of user forecast fields you would like on the worksheet.
Specify Extended Options:	A Y in this field will take you to the Extended Options Menu (see Figure 8-27 for an example of this screen). The Extended Options Menu allows you to optionally add a list of comments to the generated report and/or write your own SAS code to subset the input file.

The lower portion of the control screen contains a scrollable data entry display in which you may specify the data elements to be included on the report. Blank entries represent blank columns. The following fields are available for each row in the display area:

Name:	The name of the data element to be included on the report. If you are uncertain as to the correct name of the data elements on the input file, type a question mark (?) in this field and a selection list of available data elements will be presented.

Col. The first and second lines of the column heading.

Heading 1 The maximum permissible length for the column and Col. headings is 12 characters. No double quotes (") or semicolons (;) may appear within the body of the headings.

Heading 2:

Output format: The output format that is used to print the values for this data element in the reports. The report field is 12 columns wide, so any value that you enter here should specify a field width of 12, or the values will not be right-justified. If you specify a format value on the CA MICS File and Data Element Selection screen when the Input file is defined, this value is supplied automatically as a default. If you do not specify any value here, then the Profile and Trending Analysis uses one of two default output formats. For any data element name that ends in the characters TM, a SAS format of TIME12.2 is used. All other data elements default to a SAS format of 12.2.

```

/----- CAPAPU Selection Values -----\
|Command ==>|
|Composing CA MICS Inquiry: _____|
|
|      Rule  ---+---1---+---2---+---3---+---4---+---5---+---6---|
|CAPAPU  ==>|-----|
|CAPAPU 2 ==>|-----|
|CAPAPU 3 ==>|-----|
|CAPAPU 4 ==>|-----|
|CAPAPU 5 ==>|-----|
|CAPAPU 6 ==>|-----|
|CAPAPU 7 ==>|-----|
|CAPAPU 8 ==>|-----|
|CAPAPU 9 ==>|-----|
|
|
\-----/

```

Figure 8-26. CAPAPU Selection Values Screen

[illegible]

Figure 8-27. CAP File Processing Extended Options Menu

Figure 8-27 shows the CAP File Processing Extended Options Menu that is displayed if you specify a Y (YES) for the Specify Extended Options parameter on the Worksheet Effect Report screen shown in Figure 8-25.

The following options are available:

- o Specify file selection criteria -The specification for building a SAS Macro to subset the input Capacity Planner File. See Figure 8-29 to see the screen that is displayed. Specify a 'Y' to create/update the macro. Specify an 'R' to reset (clear) any macro that may be there.
- o Specify footnotes for the worksheets--The specification for placing footnotes at the bottom of the effect reports. See Figure 8-28 to see the screen that is displayed. Specify 'Y' to place footnotes on the worksheet; specify 'R' to reset (clear) any footnotes that are already defined.

When you are finished with this screen, END to return to the Worksheet Effect Reports Control screen (see Figure 8-25).

```

/----- SAS System Statements -----\
|Command ==>                               Scroll ==> CSR |
|
|Inquiry Step: WRKREP - Worksheet Testing - Effect Report |
|***** TOP OF DATA *****|
|==MSG> SPECIFY SAS STATEMENTS TO SUBSET THE CAPACITY PLANNING DATA BASE FILE. |
|=NOTE= %MACRO SELECT; <=== PLEASE DO NOT PLACE LINES BEFORE THIS LINE. |
|000001 |
|000002 |
|000003 |
|000004 |
|000005 |
|000006 |
|000007 |
|000008 |
|000009 |
|000010 |
|000011 |
|000012 |
|000013 |
|000014 |
|=NOTE= %MEND SELECT; <=== PLEASE DO NOT PLACE LINES AFTER THIS LINE. |
|***** BOTTOM OF DATA *****|
\-----/

```

Figure 8-28. SAS Statements to Subset the Input File

Figure 8-28 shows the SAS System Statements screen that you use to code a SELECT macro that will be used to subset the input Capacity Planner file. This is a table display screen. See Section 1.5 for more information on how to use this type of screen. Save any changes you made. Enter the END command.

```

/----- Worksheet Effect Report Footnotes -----\
|Command ==>                                     Scroll ==> CSR |
|SAS statements for the specified exit (SELECT macro) have been deleted. |
|Specify footnotes that will appear at the bottom of the report page.   |
|                                                                           |
|Line Comments                                                             |
| - ----- |
| 1 ----- |
| 2 ----- |
| 3 ----- |
| 4 ----- |
| 5 ----- |
| 6 ----- |
| 7 ----- |
|                                                                           |
\-----/

```

Figure 8-29. Worksheet Effect Report Footnotes Screen.

Figure 8-29 shows the Worksheet Effect Report Footnotes screen. You can specify up to seven lines of footnotes for the report. To save your footnotes without leaving this screen, enter the SAVE command. To void any changes made to this screen, enter the CANCEL command. To leave this screen and save any changes you made, enter the END command.

Chapter 9: MULTIVARIATE REGRESSION FORECASTING

The Multivariate Regression Forecasting software allows you to develop regression models that relate different resource consumption data elements or business-related data elements from one or more capacity planning database files. This forecasting capability addresses the following requirements:

- o The investigation of possible relationships among different resource variables
- o The investigation of possible relationships among different business element variables
- o The investigation of possible relationships between a resource variable and a group of business-related variables
- o The forecasting of future requirements of a resource variable based on estimates of the future requirements of one or more different resource variables or business-related variables

Note: SAS/STAT is a prerequisite for using Multivariate Regression Forecasting.

The following sections describe this forecasting procedure's operation and reports.

This section contains the following topics:

- [9.1 Functional Description](#) (see page 728)
- [9.2 Usage Guidelines](#) (see page 729)
- [9.3 Standard Output](#) (see page 738)
- [9.4 Analytic Technique Tutorial](#) (see page 746)
- [9.5 Component Operation](#) (see page 748)
- [9.6 Case Studies](#) (see page 758)

9.1 Functional Description

Multivariate Regression Forecasting produces regression models that relate different resource consumption data elements or business-related data elements from one or more capacity planning database files. These forecasts are based on stepwise multiple regression models using any of the resource element or business element files produced by the CA MICS Capacity Planner. You can use the same or different files as the source for dependent and independent variables in the model.

The process implements the multivariate forecasting methodology in a two-step procedure. The first step uses the SAS REG procedure with the stepwise option to investigate the relationships among the historical observations of the variables. This step determines which of several historical variables have the most predictive power on the variable to be forecast. In the second step, the SAS REG procedure builds a predictive model using the variables selected in the previous step, estimates future values of the forecast variable, and generates confidence intervals on the forecasts.

The operation and usage of Multivariate Regression Forecasting is nearly identical to that of the Business Element Forecasting, described in Chapter 10. In Business Element Forecasting, the control screens ensure you are forecasting a resource consumption variable based on its possible relationship to one or more business-based variables.

Multivariate Regression Forecasting allows you to investigate relationships among variables from any two capacity planning database files, or among variables from a single file. Any of these files can be resource element files or business element files. Thus, although Multivariate Regression Forecasting and Business Element Forecasting use the same statistical techniques, the first type of forecasting is potentially much more general.

To illustrate the use of Multivariate Regression Forecasting, we have developed a multivariate regression model that estimates long-range capture ratios. This model relates the historical use of specified processor workload categories to overall processor use. A complete description and interpretation of the capture ratio model is presented in Section 9.6.1.

9.2 Usage Guidelines

Multivariate Regression Forecasting is a technique for investigating the mathematical relationships between one or more resource variables contained in capacity planning data base files. The mathematical relationships which may exist between variables with a high degree of correlation can be expressed as models. You can use these models to forecast future values of the resource variables.

There are three broad steps in the development of a multivariate regression model:

- o Choosing model parameters
- o Evaluating the model
- o Forecasting future values from the model

These topics are discussed after a brief review of stepwise multiple regression models. This introduction presents some of the basic terminology used in the evaluation of regression models. A more detailed development of the multiple regression technique is presented in Section 9.4.

The following topics are presented by section:

- 1 - Stepwise Multiple Regression Models
- 2 - Choosing Model Parameters
- 3 - Evaluating the Model
- 4 - Forecasting Future Values from the Model

9.2.1 Stepwise Multiple Regression Models

Within the context of the CA MICS Capacity Planning component, the term "multivariate regression" refers to multivariate linear regression.

A multivariate regression model is a linear model which explains one variable, known as the dependent variable, in terms of a linear combination of several other variables, known as independent variables. The general form of the model can be written as follows:

$$y = b + m1 * x1 + m2 * x2 + m3 * x3 + \dots mj * xj \text{ (Eqn 1)}$$

In this form the dependent variable, denoted by y , is shown as a combination of a constant term, b , added to a group of j independent variables, denoted by $x_1, x_2, \dots x_j$.

Each of the x variables is multiplied by an m coefficient which represents the directional slope of a regression surface in a j -dimensional space. When only one independent variable is used, the equation for the model is:

$$y = b + m_1 * x_1 \quad (\text{Eqn 2})$$

This is the equation for a univariate regression model, which is described in Section 7.4. The m coefficient is the slope of a line in a two-dimensional space, defined by a y -axis and an x -axis.

If the model is extended to two independent variables, it expands to:

$$y = b + m_1 * x_1 + m_2 * x_2 \quad (\text{Eqn 3})$$

This is a three-dimensional space, defined by a y -axis, an x_1 -axis, and an x_2 -axis. The m_1 and m_2 coefficients represent the slopes of a plane, or flat surface, within the space. It becomes very difficult to visualize the physical appearance of the regression surface for models with more than two independent variables.

In the case study presented in Section 9.6.1, we construct a multivariate regression model which relates total CPU time to the TCB CPU time used by key system workloads. For this model the y and x_j terms are:

	x_1 - TSOCPUTM	x_2 - BATCPUTM
y - TOTCPUTM	x_3 - IMSCPUTM	x_4 - CICCPUTM

To use Multivariate Regression Forecasting, you must supply observed values of the dependent and independent variables. The procedure generates estimates of $m_1, m_2, \dots m_j$, and various test statistics, such as r -squared, F , and p , which are the multivariate analogs to the statistics for univariate modeling described in Section 7.4.

In a general multivariate model, you determine the number of independent variables, assuming that you know what the independent variables should be and how many there are. At the start of an exploratory analysis, however, you do not know which variables are related, and how many variables should be analyzed.

To solve this problem, Multivariate Regression Forecasting uses what is known as a stepwise process to determine which independent variables have the most predictive power on the dependent variable. This process, which is based on the SAS REG procedure with the stepwise option, attempts to find which subset of a group of variables forms the best predictive model. For example, you may supply ten variable names. The process chooses the best predictor, then the second best, then the third best, and so on, by trying out each of the ten variables individually and retaining the best. Once the first is known, the process tries this one with each of the remaining nine to select the best two-element model. The process continues the testing, finding the best three-element model, four-element model, and so on. It stops when it finds no more variables to test or no greater statistical quality in the models.

9.2.2 Choosing Model Parameters

The first step in the development of a multivariate regression model is to determine the appropriate model parameters (that is, the dependent and independent variables to be evaluated).

Not all possible combinations of independent and dependent resource variables within the capacity planning database provide usable or even sensible models in the capacity planning process. Exploring every possibility in searching for the right models would be an enormously time-consuming and time-wasting enterprise. And even though the technique of stepwise multiple regression is useful in eliminating certain models, there are still too many combinations to evaluate. Throughout the modeling and forecasting process, your own knowledge and experience are the best possible guidelines for finding the correct applications for Multivariate Regression Forecasting and for discarding inappropriate ones.

Certain applications will never provide suitable models for Multivariate Regression Forecasting because they do not fit such a model on the basis of purely theoretical reasons. For example, a model that consists of the response time as the dependent variable, and the transaction rate and transaction service time (and possibly other variables) as the independent variables do not usually provide a suitable multivariate regression model. This is because analytic queuing theory demonstrates that the most appropriate models for these variables are non-linear.

On the other hand, certain processor overhead functions can provide reasonable multivariate regression models. Even though the hypothetically correct model may not necessarily be linear, it may be close enough to allow very good approximation by a linear model.

The following are some examples of situations where multivariate regression can provide reasonable models:

- o Uncaptured processor utilization. This is the amount of processor busy time used for unrecorded overhead functions. This example is discussed in more detail in Section 9.6.1.
- o IMS Control Region overhead analysis. You can correlate the amount of time used by an IMS control region to the amount of processor time used by its dependent regions. This example is discussed in more detail in Section 9.6.2.
- o System address spaces. Certain system address spaces and tasks provide services for all users of the system. Examples of these services include the master scheduler, JES, VTAM, TCAS, monitors, global enqueue propagators, and others. You can frequently correlate use of these systems to processor use of the other non-system address spaces, such as batch and TSO address spaces. Such a model could be helpful, since future resource consumption estimates for the batch and/or TSO address spaces might then be used to produce future estimates of resource consumption by this group of system address spaces.

- o Recorded processor utilization. You can generally relate processor use to the average number of batch, TSO, and started tasks that are running on the processor. This application of Multivariate Regression Forecasting enables you to predict future processor use on the basis of the anticipated number of users.

When you choose the applications for Multivariate Regression Forecasting analysis, ensure that the desired resource variables are present in capacity planning database files. All independent variables in the model must exist within a single capacity planning database file. The dependent variable may exist in the same or a different capacity planning database file.

In addition, you must ensure that the capacity planning data base file containing the dependent variable and the file containing the independent variables are of the same time-span (either DAYS, WEEKS, or MONTHS), and that they both cover a concurrent period of historical data.

9.2.3 Evaluating the Model

The second step in developing a Multivariate Regression Forecasting model is evaluating how well the model fits the data. For this purpose, you can use the same basic statistics that you use in univariate linear regression with multivariate regression analysis. The most commonly used statistics are the r-squared, F, and p statistics. The statistics for Multivariate Regression Forecasting and univariate linear regression are similar, although the formula for calculating multivariate regression statistics is more involved.

As we discussed in Chapter 7, r-squared, F, and p are interpreted as follows:

- o r-squared is the coefficient of determination (also referred to as the coefficient of multiple correlation). This value indicates what percent of the variability in the historical data is explained by the proposed model. You would typically be seeking r-squared values greater than or equal to 0.7.

- o F is a parameter calculated from the ratio of sums of squares terms for the models. In general, larger F values indicate more reliable models. However, the F value has very little significance when less than 30 historical data points are used to develop the model.
- o p is the probability that the F value is significant. A value less than or equal to 0.01 has been recommended for computer measurement applications (SAR79). You can discount the p value if you use a small number of historical observations to build the model.

There are countless numbers of tests and statistics that aid the sophisticated statistician in the evaluation of models. However, there are three basic tests that you can employ. Used in concert, these tests should enable you to evaluate how well the models fit the data relatively quickly and with relatively little prerequisite mathematical knowledge.

TEST 1 - Examine the r-squared, F, and p statistics.

We recommend an r-squared value greater than 0.7. The p statistic should be small (we recommend a value of 0.01 or less). Like most rules of thumb, these rules cannot be considered inviolate.

The F value should also be large, although the size is relative and depends upon, among other things, the number of independent variables. The F-value is used to calculate the p statistic, so it is probably better to rely upon the p statistic until you gain familiarity with models and their statistics.

The r-squared, F, and p values are printed on the Model Analysis Report.

TEST 2 - Visually examine the data.

Visual examination of the actual and predicted data is very important and can sometimes show that a model whose r-square, F, and p statistics look bad may be significantly improved by the exclusion of a small number of outliers from the data.

The Multivariate Residual Analysis Report can help you visually examine the actual, predicted, and residual values. Ideally the residuals should appear totally random. If you notice a pattern in the residuals, you may want to reconsider your choice of independent variables.

TEST 3 - Apply the common sense test.

The common sense test asks, "Does this model make any sense?" To illustrate this point, consider the following example. You decide to analyze a model whose dependent variable is CPU utilization, measured in CPU hours per zone per week. The three independent variables are the average number of started tasks (STC), the average number of batch jobs (BAT), and the average number of TSO users (TSO) during the week.

Suppose that a stepwise regression shows that the best r-squared model is

$$\text{CPU\_HRS} = (1.5 \times \text{TSO}) + (.89 \times \text{BAT}) + (-.23 \times \text{STC}) + 2.45$$

For this model, the r-squared value is 0.82, the F-value is 323.26, and the p statistic is .0012.

Based on these values, the model passes TEST 1 with flying colors. Furthermore, since visual analysis of the actual and predicted data indicates that there appear to be no serious outliers, you can consider TEST 2 a success.

However, this model fails TEST 3 because the coefficient for the STC variable is negative. If this model were truly valid, then you would want to run as many started tasks as possible on this processor, for each additional running started task reduces CPU utilization by 0.23 hours of CPU time per week. Such a notion violates common sense, so we must discard this model.

Generally, situations like the above (in which TESTs 1 and 2 pass and TEST 3 fails miserably) most often occur when there is insufficient historical data included in the evaluation of the model. This situation can also occur if there is duplication among the variables you use in the analysis. For example, predicting total CPU utilization based on batch job, TSO session, and IMS transaction CPU times can yield negative coefficients if the IMS message region batch jobs are not excluded from the historical data. In summary, always apply TEST 3.

ATYPICAL BEHAVIOR

When TEST 3 fails, look for atypical behavior on the system during the measurement interval. For example, this model shows a negative relationship between started tasks and overall CPU time. Perhaps the system programmers were testing some new online systems, and they were allowed to test only during non-peak periods. Because of this, they started several online systems whenever the load on the system dropped off. In this case, you could isolate the test online systems as a separate independent variable, and repeat the analysis.

Not all problems with data are the result of atypical workloads. A more serious problem which causes TEST 3 to fail is a violation of the basic assumptions of regression. In the theoretical formation of a regression model, we assume that the independent variables are truly independent of each other. This assumption is often violated by computer system data. In our example we use the number of TSO users, batch jobs, and started tasks. If TSO is used for development and test, you would expect that increased TSO use would lead to increased batch usage as the programmers submit more jobs.

This violation of the regression assumptions, where the independent variables are correlated with each other, is known as multicollinearity. The only way to recover from data which represents multicollinearity is to change the data to prevent the problem. If the TSO and batch activity are correlated, you must remove the correlation by creating a new variable which represents batch jobs not due to TSO. This variable is independent of TSO, so you may repeat this analysis with your basic independence assumption satisfied.

9.2.4 Forecasting Future Values from the Model

The final step in developing the multivariate regression model is to use the model to predict future values of the dependent resource variable. This is easily accomplished using Multivariate Regression Forecasting.

To predict future values for the dependent resource variable, you must assume that forecast values for all independent resource variables used in the model are stored in the forecast file, which corresponds to the resource file containing these independent resource variables. You can supply these predicted values for the independent resource variables by using any of the forecasting methods available with the CA MICS Capacity Planning Component: Univariate Model Forecasting, Business Element Forecasting, or by using a different model generated by Multivariate Regression Forecasting.

It is possible to keep multiple forecasts for different time periods in a forecast file. The future data observations of the independent variables must come from a forecast period that corresponds to the period for which you are forecasting the dependent variable via Multivariate Regression Forecasting. For example, assume that a resource file named ABC contains data representing 24 months of resource use extracted from the CA MICS database. You intend to use variables from this file as independent variables for Multivariate Regression Forecasting. You want to forecast ahead 12 months. You have another resource file named XYZ containing data for the same time period. You intend to use one of its variables as the dependent variable in this forecast process:

ABC file: historical data for Jan97-Dec98
XYZ file: historical data for Jan97-Dec98

The first step is to use some forecasting process, perhaps Univariate Model Forecasting, to produce a forecast for Jan99-Dec99 for the independent variables in the ABC file and save this forecast in its corresponding ABC forecast file:

ABC forecast file: forecast data for Jan99-Dec99

Next, use Multivariate Regression Forecasting to produce a forecast of the dependent variable in the XYZ historical file, based on your forecasts of the independent variables in the ABC file. You might direct Multivariate Regression Forecasting to use all the data from the ABC file and the XYZ file to calculate the parameters for the model. These parameters and the data from the ABC forecast file are used to produce a forecast for the dependent variable, optionally saving it in an XYZ forecast file:

XYZ forecast file: forecast data for Jan99-Dec99

Make one test upon the data in the ABC forecast file before using it to produce the forecast for the XYZ file. Any forecast file can contain multiple forecasts. Thus, you might also have saved a forecast in the ABC file several months ago, perhaps based only on actual historical data through July, 1998. Each forecast observation is stamped with the year and month or week of the latest historical data used to produce it. Multivariate Regression Forecasting guards against using any forecast data as input except data which has a forecast stamp equal to the latest historical data used during the multivariate regression analysis. This ensures that you do not use forecast data which does not correspond to the historical data used in the multivariate regression analysis.

9.3 Standard Output

Output from Multivariate Regression Forecasting consists of printed reports and, optionally, data values stored in a forecast file. These are covered in more detail in the following sections:

- 1 - Update to the Forecast File
- 2 - Standard Reports

9.3.1 Update to the Forecast File

You can use the Save forecast option on the Multivariate Regression Forecasting control screen shown in Figure 9-10 to save forecasted values of the dependent variable in the forecast file that corresponds to the file containing the dependent variable.

Typically, as a user of Multivariate Regression Forecasting, you can develop a model using the three steps outlined in Section 9.2. This is almost always an iterative process, for the first few choices of model parameters may not be the best choices. Formulate the model and then review the results of that model using the three tests described in Section 9.2.3. Should the tests fail, make corrections and adjustments to the proposed model and rerun the forecast.

Finally, you develop a model that satisfies your requirements. At this point, you can use the model to generate future values for the dependent variable. To do this, future values of ALL the independent variables must exist. By rerunning the same model with the Save forecast option mentioned above, the forecasting process generates the future values of the dependent variable and stores them in the appropriate file.

In the forecast file, future values are stored in a special way that is used by other functions of the CA MICS Capacity Planner to produce comparative graphs and charts. The FILE_NO variable is used to distinguish between the various types of observations within forecast files. The Multivariate Regression, Univariate Model Forecasting, and Business Element Forecasting facilities store a future value with a FILE_NO of 11, and a confidence range with a FILE_NO value of 12. In addition, each observation of a forecast contains the variables FORYEAR and FORMNDY, FORWEEK or FORMONTH, which contain the date for the most recent historical data used in making the forecast. These concepts are discussed in more detail in Section 3.2.

9.3.2 Standard Reports

Multivariate Regression Forecasting produces three reports, which are discussed in the following sections:

- 1 - Model Analysis Report
- 2 - Residual Analysis Report
- 3 - R-Square Ranking Report

9.3.2.1 Model Analysis Report

Figure 9-1 presents a sample Model Analysis Report for Multivariate Regression Forecasting. The Model Analysis Report provides a summary of the regression analysis that was conducted for the specified multivariate regression model.

CA MICS Capacity Planner							
ANALYSIS OF MULTIVARIATE REGRESSION MODELS							
MODEL OF: TOTCPUTM							
BASED ON: TSOCPUTM IMSCPUTM CICCPUTM							
R**2	F	P	INTERCEPT	NAME	COEFFICIENT	F	P
0.97	1127.97	.0001	10:57:58.3	TSOCPUTM	1.77499	71.77	.0001
				IMSCPUTM	1.14392	77.90	.0001
				CICCPUTM	1.13065	31.88	.0001

Figure 9-1. Model Analysis Report

The Model Analysis Report contains the following information:

MODEL OF: The dependent data element that you wish to forecast.

BASED ON: The independent data elements on which the analysis is based.

The first four columns of the report are discussed below.

R2:** The r-squared value for the model. Although the SAS REG procedure attempts to use every one of the independent elements suggested by the analyst, the procedure excludes any term that does not result in a minimum improvement in the r-squared value as specified on the Multivariate Regression Forecasting screen. We recommend an r-squared value of 0.70 and above for the acceptance of models produced with Multivariate Modeling Forecasting.

F: The F statistic for the model. As discussed in Section 9.2.3, the F statistic is calculated from the ratio of sums of squares of the model. Larger F values indicate a more reliable model. The F statistics may be unreliable for models developed with a small number of points (that is, less than 30). An F value is provided for each step made by the regression procedure.

P: The probability that the model proposed is significant (that is, more reliable than forecasting the average value of the historical observations). Although most statistics texts recommend testing the null hypothesis to a 0.001 level, we recommend 0.01 for models based on computer measurement data. A p value is provided for each step made by the regression procedure.

INTERCEPT: The b value for the regression equation.

The next four columns are produced for each independent element that you specify. These columns contain the following information:

NAME: The name of the independent element being analyzed.

COEFF: The coefficient for the independent element in the regression equation. The coefficient is the M value mentioned in Sections 9.2.1 and 9.4. Negative coefficients indicate that you can expect resource consumption to decrease with increases in the independent element value. This situation is usually not normal in resource consumption models.

F: The F statistics calculated for the business element. This F value indicates the significance of the independent element in the model that was developed. Once again, larger values indicate a more reliable model and the value may be misleading for models based on small samples.

P: The probability that the specified business element makes a significant contribution to the model developed. We recommend a 0.01 acceptance value.

This report is used to evaluate whether the model that is produced adequately models an application. Although the values of r-squared, F, and p indicate the statistical significance of the proposed model, remember that a proper evaluation includes all three tests enumerated in Section 9.2.3: statistical interpretation, visual inspection, and the common sense test.

9.3.2.2 Residual Analysis Report

The Residual Analysis Report shows how the multivariate regression model fits the historical data series. Figure 9-2 shows a sample Residual Analysis Report.

CA MICS Capacity Planner									
MULTIVARIATE REGRESSION FORECAST AND RESIDUAL ANALYSIS									
MULTIVARIATE REGRESSION BASED FORECAST OF: TOTCPUTM									
MODEL BASED ON INDEPENDENT ELEMENTS : TSOCPUTM BATCPUTM IMSCPUTM CICCPUTM									
DATE	TOTCPUTM	INDEPENDENT-ELEMENT NAME	VALUE	PREDICTED TOTCPUTM	RESIDUAL TOTCPUTM	-----PLOT OF RESIDUALS-----			
						-18723	0.0	□	95% CONFIDENCE LIMITS
04JAN97	95:34:40.7	TSOCPUTM	47321	91:47:15.0	3:47:25.7		0.0		1:26:03.7
		BATCPUTM	37461				0		
		IMSCPUTM	118575				0		
		CICCPUTM	60716				0		
11JAN97	91:11:54.9	TSOCPUTM	47779	90:36:08.0	0:35:46.9		0		1:20:01.3
		BATCPUTM	37663				0		
		IMSCPUTM	115851				0		
		CICCPUTM	59529				0		
18JAN97	93:57:10.4	TSOCPUTM	49732	92:14:06.2	1:43:04.2		0.0		1:04:26.9
		BATCPUTM	36463				0		
		IMSCPUTM	119521				0		
		CICCPUTM	59195				0		
25JAN97	92:25:59.9	TSOCPUTM	52382	93:46:39.6	-1:20:39.6		-----0		1:02:14.1
		BATCPUTM	35767				0		
		IMSCPUTM	121372				0		
		CICCPUTM	59275				0		
01FEB97	94:15:30.1	TSOCPUTM	52535	94:58:53.1	-0:43:22.9		--0		1:03:10.5
		BATCPUTM	36859				0		
		IMSCPUTM	122859				0		
		CICCPUTM	59992				0		
08FEB97	96:10:44.4	TSOCPUTM	54022	97:40:15.2	-1:29:30.8		-----0		1:08:11.3
		BATCPUTM	36984				0		
		IMSCPUTM	126560				0		
		CICCPUTM	61944				0		

Figure 9-2. Residual Analysis Report

The Residual Analysis Report includes the following information:

MULTIVARIATE
REGRESSION

BASED The dependent data element name that you
FORECAST OF: specify on the entry screen.

MODEL The independent elements that you use to
BASED ON develop the model. This list may differ from
INDEPENDENT the list of independent elements specified in
ELEMENTS: the Model Analysis Report shown in Figure 9-1.
 This difference results from deleting proposed
 independent elements based on the minimum
 r-squared improvement criterion.

The columns of the report are discussed below:

DATE: The date ending value for the week or month
 represented by the observation.

Independent The actual historical observations for the data
Element element being modeled. Note that if the last
Name two characters of the data element name are TM,
(TOTCPUTM): the program assumes that the value is in
 seconds and uses a SAS TIME10.1 format for the
 report.

Two columns appear under the heading INDEPENDENT-ELEMENT for each independent element used in the model.

NAME: The name of the independent element.

VALUE: The value of the independent element for the
 interval.

The remaining columns of the report are described as follows:

PREDICTED The predicted value for the dependent variable
(TOTCPUTM): for the interval using the regression equation
 fit to the historical data. Note that if the
 last two characters of the data element name
 are TM, the program assumes that the value is
 in seconds and uses a SAS TIME10.1 format for
 the report.

RESIDUAL (TOTCPUTM):	The error term in the regression equation. The residual is equal to the difference between the actual historical observation and the predicted value. The residual value for a point deleted from the modeling process is of no concern since the model was not developed to account for the observation. Note that if the last two characters of the data element name are TM, then the program assumes that the value is in seconds and uses a SAS TIME10.1 format for the report.
PLOT OF RESIDUALS:	The residual values plotted over the range from minus to plus the maximum residual value.
CONFIDENCE LIMITS:	The confidence limits for the forecast. The confidence limits bound the potential error that may exist in the forecasts produced by the model. The confidence limits are a function of the independent element estimates provided, rather than monotonically increased as was the case with models developed by the Univariate Model Forecasting. If the last two characters of the data element name are TM, the program assumes that the value is in seconds and uses a SAS TIME10.1 format for the report.

9.3.2.3 R-Square Ranking Report

The R-Square Ranking report provides you with an overview of the results of all of the forecasts performed during this forecasting run. The report is presented in ascending R-Square sequence and gives the r-square, f-statistic, and p-values for the statistical model generated. The r-squares and p-values for the individual variables in the best model are also provided.

If either the r-square or p-value for the run exceeds the thresholds set for the report, the Warning Flag column will be set to show the which threshold was exceeded. If the flag is set to RS-WARN then the r-square for this particular forecast was below the threshold; if the flag is set to P-WARN then one of the p-values for one of the variables in the model was exceeded. If both thresholds are exceeded then the RS-WARN takes precedence.

The other fields on this report correspond to the equivalent values found on the other reports associated with the forecasting process.

You can use the R-Square Ranking report to review the results of all your forecasts and to indicate whether you need to take further action. For example, you may wish to review the Model Analysis report or the Residual Analysis report to see what may be affecting these results. We recommend that for critical workloads you review these other reports regardless of whether or not they are flagged on this report.

CA MICS Capacity Planner					
R-SQUARE RANKING REPORT					
Testing MVR					
MODEL OF: MBALOC					
BASED ON: MBTOTL STGVOLNO DATE					
WARNING FLAG	R**2	P	SYSID	VARIABLE	P
-----	-----	-----	-----	-----	-----
RS-WARN	0.431	.0001	S008	MBTOTL	0.004
				STGVOLNO	0.003
				DATE	0.048
P-WARN	0.71	.0031	S008	MBTOTL	0.003
				STGVOLNO	0.040
				DATE	0.741
	0.71	.0001	S008	MBTOTL	0.043
				STGVOLNO	0.023
				DATE	0.044
R-SQUARE WARNING VALUE : 0.50 P-VALUE WARNING THRESHOLD : 0.05					
* - CAPACITY PLANNING INQUIRY: MVRSMS					
- THIS FORECAST HAS BEEN SAVED					

Figure 9-3. R-Square Ranking Report

9.4 Analytic Technique Tutorial

Within the context of the CA MICS Capacity Planner, the term multivariate regression refers to multivariate linear regression. Non-linear techniques of multivariate regression exist, but a discussion of these techniques goes beyond the scope of this text. In data processing, non-linear models are primarily useful in formulating analytical queuing models and studying response time as the dependent variable. If necessary, you can evaluate non-linear models by changing the functional form of the independent variables, using a multivariate analog to the quadratic and cubic regression models described in Section 7.4.1.

The technique of multivariate regression is identical to that of univariate linear regression, except that more variables are involved than the x and y variables (a tutorial on univariate linear regression is presented in Section 7.4). In multivariate regression, the hypothesis is that the dependent variable y is a linear function of several independent variables: $x_1, x_2, \dots, x_n$.

If you recall the introduction to linear regression in Section 7.4, you know that the equation for a simple linear regression model is:

$$y = b + m * x \quad (\text{Eqn 4})$$

where b is a constant that defines the mean value of y when x is zero

m is the slope of the line that relates y and x

y is the value to be predicted (dependent variable)

x is the value on which the prediction is based (independent variable)

You can expand this equation to be a multiple regression equation by replacing the m and x terms in the previous equation by a set of terms as shown in the following equation:

$$y = b + \sum_{j=1}^n m_j * x_j \quad (\text{Eqn 5})$$

where b is a constant that defines the mean value of y when all x_j are zero

m_j is the slope of the line that relates y and x_j

x_j is the j -th independent variable

Recall the example shown in Figure 9-1 in which the y and x_j terms of Equation 5 are:

$$\begin{array}{ll} x_1 - \text{TSOCPUTM} & x_2 - \text{BATCPUTM} \\ y - \text{TOTCPUTM} & \\ x_3 - \text{IMSCPUTM} & x_4 - \text{CICCPUTM} \end{array}$$

You may solve multiple regression equations for any number of x terms, but you must select some subset of these terms to build models for each of the resource consumption values. While you can make these selections arbitrarily, stepwise multiple regression provides an ideal technique for investigating potential relationships between resource variables.

Although numerous stepwise multiple regression algorithms are available, this program uses Goodnight's MAXR algorithm, which is implemented in the SAS REG procedure using the stepwise option.

For example, consider the development of a statistical model that attempts to relate the CPU utilization of a processor to three potential resource elements named A, B, and C. In this example, A, B, and C are the independent (or x) resource variables, while CPU utilization is the dependent (y) resource variable. The name stepwise multiple regression describes the function of the algorithm. For our small example, the algorithm attempts to solve the problem using three steps:

- o The algorithm evaluates three linear one-term models (that is, one based on A, one on B, and one on C) to determine which of the three potential elements had the greatest correlation to the resource being modeled. This correlation is judged based on the r-squared value for each of the models. The algorithm attempts to maximize r-squared. For the purposes of this example, assume that the B element has the highest correlation.
- o The algorithm would then evaluate two linear two-term models based on B and the remaining two elements (that is, B and A, and B and C). Once again, the algorithm chooses the best model based on the maximum r-squared. For the purposes of this example, assume that the pair of elements that resulted in the maximum r-squared value is B and C.
- o Finally, the algorithm would evaluate one linear model based on A, B, and C.

The algorithm stops at any step where the addition of a term to the model does not result in an increase of the r-squared value by a certain amount. Thus, you can employ stepwise multiple regression to investigate the relationship between several resource consumption variables, letting the algorithm choose the most likely combinations and eliminate the least likely.

9.5 Component Operation

The Component Operation Section describes the control parameters which are used to build Multivariate Regression Forecasting inquiries. These parameters allow you to define forecasts based on stepwise multivariate regression, optionally making use of selective data deletion.

Because of the optional feature of selective data deletion, you may need to use two screens to develop a multivariate regression forecast. The following sections discuss the use of each of these screens.

- 1 - Multivariate Regression Forecasting Control Screen
- 2 - Delete Historical Data Observations Screen

9.5.1 Multivariate Regression Forecasting Control Screen

The Multivariate Regression Forecasting control parameters allow you to develop forecast models based on stepwise multivariate regression. You can save the forecasts produced by the program in a forecast file corresponding to the file you specify as the source of the dependent variable for the model.

Using a short set of menus that are described in Section 1.5 of this guide, you will arrive at the Multivariate Regression Forecasting screen shown in Figure 9-4. The Multivariate Regression Forecasting control screen appears after you indicate that you want to insert a new Multivariate Regression Forecasting analysis inquiry into your catalog of inquiries. Two fields on this screen are filled in on the fourth line when you first see it: the inquiry name and the inquiry description, separated by a dash. You entered these two fields on the Insert Capacity Planning Analysis screen when you requested that a new inquiry be added.

```

/----- Multivariate Regression Forecasting -----\
|Command ==>
|Enter a ? in any data entry field for more information on valid values.
|Composing CA MICS Inquiry: ADHOC - Adhoc Inquiry - Will Not Be Saved
|
|Report title ==> Adhoc Inquiry - Will Not Be Saved
|
|Selection criteria:
|    Dependent file ==> ___ -
|    Dependent element ==> _____
|    Start date ==> _____ (ddmonyyyy)
|    SYSID ==> _____
|    LVL ==> _____
|    No. wk/mn ==> _____ (1-9999)
|    Specify CAPAPU Values ==> N (Y/N)
|Save forecast ==> ___ (YES/NO/AGE)
|Independent file ==> ___ -
|Independent elements ==> _____
|
|Confidence limits (percent) ==> __ (70/90/95)
|Min R-square improvement ==> _____ (0.000-0.100)
|Delete observations ==> N (Y/N/R)
|
\-----/

```

Figure 9-4. Multivariate Regression Forecasting Control Screen

Some of the parameters on the Multivariate Regression Forecasting screen shown in Figure 9-4 are required, while others are optional. The following is a discussion of the various parameters as they appear in this figure.

Report title:	The 60-character title for Multivariate Regression Forecasting reports. No double quotes (") or semicolons (;) may appear within the body of the title.
Dependent file:	The resource element or business element capacity planning database file containing the historical observations for the dependent variable of the regression model. If you are uncertain of the files or file names available in your capacity planning database, type a question mark (?) in the field and a selection list will be displayed.

Dependent element:	The name of the dependent variable in the regression model. If you are uncertain of the correct name of the data elements in the input file, type a question mark (?) in this field and a selection list of available data elements will be presented.
Start date:	The start date for date range selection. If Start date is specified, only historical observations at least as recent as this value will be included in the analysis. The discussion of the following control parameter describes the options for controlling date range selection.
SYSID:	The SYSID for which the report is to be produced. If the Dependent file is a resource element file and is defined as having a composite SYSID, this value is supplied automatically on the screen. Note that if the Independent file is a resource element file, it must also contain historical observations for this SYSID. If the Independent file is a business element file, this requirement does not apply, since business element files are not defined as containing the SYSID data element.
LVL:	The Zone, Hour or Dayname to be included in the analysis, depending on whether you define the Dependent file at the ZONE, HOUR, or DAYNAME summarization level. If the Dependent File is a resource element file and you define it as having a composite ZONE, HOUR, or DAYNAME this value is automatically supplied on the screen. Note that if the Independent file is a resource element file, it must also contain historical observations for this ZONE, HOUR or DAYNAME. If the Independent file is a business element file, this requirement does not apply, since business element files are not defined as containing the ZONE, HOUR or DAYNAME data element. Note that the field name will automatically be changed from LVL to either ZONE, HOUR, or DAYNAME to reflect the Input file summarization level once you specify the Input file.

No. wk/mn: The maximum number of days, weeks, or months of historical observations that will be included in the analysis. Note that the field name will change once you specify the Dependent file to reflect the time-span of the Dependent file. When used in conjunction with the Start date parameter, several options are available for date range selection control:

- o If you specify only Start date, all historical observations at least as recent as the value specified for Start date will be selected.
- o If you specify both Start date and No. wk/mn, the date range for data selection will begin with the value specified for Start date and will extend for the specified number of days, weeks, or months.
- o If you specify only No. wk/mn, the date range for data selection will include only this number of the most recent days, weeks, or months.
- o If you specify neither parameter, then all historical observations will be selected.

Specify CAPAPU Values: Put a Y in this field to specify the CAPAPU value(s) (up to nine) to be included in the analysis. The CAPAPU values are specified on the CAPAPU Selection Values screen shown in Figure 9-5. You may specify CAPAPUs only if the Dependent file is a resource element file and is defined as containing the CAPAPU data elements. Note that if the Independent file is a resource element file, it must also contain historical observations for the specified CAPAPU values. If the Independent file is a business element file, this requirement does not apply, since business element files are not defined as containing the CAPAPU data element.

- Save forecast:
- Three options for saving the forecasted observations in the forecast file corresponding to the Dependent file.
- o Specifying YES causes the forecasted observations to be saved in the capacity planning database forecast file (CPFxxx01) without aging the file as described in Chapter 3.
 - o Specifying NO results in the forecasted observations not being saved in the forecast file. If you select this option, the forecast file will not be aged.
 - o Specifying AGE causes the forecasted observations to be saved in the forecast file and the forecast files to be aged as described in Chapter 3.
- Independent file:
- The resource element or business element capacity planning database file containing the historical observations for the independent variable of the regression model. If you are uncertain of the files or file names available in your capacity planning database file, type a question mark (?) in the field and a selection list will be displayed.

Independent elements: The names of up to ten independent variables for the regression model. If you are uncertain of the correct name of the data elements in the Input file, type a question mark (?) in this field and a selection list of available data elements will be displayed. The number of Independent elements that you may specify is limited by the number of historical observations in the capacity planning database files. This restriction is best explained by the following logical relationship:

$$n - ie - 1 \geq 3$$

Where n is the number of historical observations and ie is the number of Independent elements that you specify. For example, you must have 14 historical observations in order to specify 10 independent elements $(14 - 10 - 1) \geq 3$. This restriction is enforced to provide a minimum of three degrees of freedom for the statistical models developed by the program.

Confidence limits (percent): The degree of confidence to be used in calculating the upper and lower confidence limits for the forecasted values of the Dependent element. The default is 95%.

Minimum R-squared improvement: The r-squared improvement criterion for the regression model. This value excludes any independent element from the regression model whose inclusion does not increase the overall r-squared value by at least this amount. The default is 0.01.

Delete observations: Enter a Y to obtain the screen shown in Figure 9-7. Enter an R to reset (nullify) any previous date deletion specifications.

If you entered a Y in the Specify CAPAPU values filed on the Multivariate Regression Forecasting screen, the CAPAPU Selection Values screen displays, as shown in Figure 9-5. On the CAPAPU Selection Values screen, you may specify the CAPAPU values (up to nine) to be included in the analysis. You may specify CAPAPUs only if the capacity planning data base file is defined as containing the CAPAPU data elements. You may use wild cards (?) and (*) when specifying CAPAPUs. A question mark (?) replaces one character; an asterisk (*) replaces all the characters after that point.

```
/----- CAPAPU Selection Values -----\
|Command ==>|
|
|Composing CA MICS Inquiry: PGRR01 PRESENTATION GRAPHICS FOR CAPAPU|
|
|      Rule ----+----1----+----2----+----3----+----4----+----5----+----6---|
|CAPAPU  ==> -----|
|CAPAPU 2 ==> -----|
|CAPAPU 3 ==> -----|
|CAPAPU 4 ==> -----|
|CAPAPU 5 ==> -----|
|CAPAPU 6 ==> -----|
|CAPAPU 7 ==> -----|
|CAPAPU 8 ==> -----|
|CAPAPU 9 ==> -----|
|
|\-----/\
```

Figure 9-5. CAPAPU Selection Values Screen

9.5.2 Delete Historical Data Observations Screen

If you entered a Y in the Delete observations field on the Multivariate Regression Forecasting screen, the Extended Options Menu appears, as shown in Figure 9-6.

```

/----- Extended Options Menu -----\
|Command ==>                               |
|                                           |
|Inquiry Step:  Multivariate Regression Analysis |
|                                           |
|Print Reports                               ==> YES  (YES/NO) |
|Maximum R-Square for report                 ==> 0.55  (0.00-1.00) |
|P-Value flag                               ==> 0.05  (0.000-1.000) |
|                                           |
|Delete historical data observations ==> N (Y/N/R) |
|                                           |
\-----/

```

Figure 9-6. Extended Options Menu

Print Reports:	The Print Reports parameter controls the printing of the Analysis of the Model and Residual Analysis reports. If you type YES, these reports are produced. If you type NO and the forecasted file is NOT a Business Element File, these reports are suppressed. These reports will always be produced for Business Element Files; however, the R-Square Ranking Report is not produced.
Maximum R-Square for report:	The Maximum R-Square for report parameter is used during the production of the R-Square Ranking Report to control the flagging of particular forecasts. If the R-Square of a generated model is less than this value, it will be flagged with an RS-WARN flag. A value of 1.0 in this field will disable the production of the RS-WARN flag in the R-Square Ranking Report.

P-Value flag: The P-Value Flag is also used in the production of the R-Square Ranking Report to determine if any of the P-values of the model variables should be flagged with the P-WARN flag. If any P-value in the model is greater than this value, it will be flagged in the report. A value of 1.0 in this field will disable the production of the P-WARN flag in the R-Square Ranking Report.

Note: The RS-WARN flag takes precedence over the P-WARN flag. This means that if a model exceeds both of these warning criteria, the RS_WARN flag will be printed.

9.6.1 Capture Ratio Case Study

This case study presents one of several methods for examining CPU capture ratios. The method presented here probably does not produce exactly the same results as another method, largely because capture ratio determination is an inexact science.

In performance analysis and capacity planning, the basis of any study is the examination of measured data. In this discussion, assume that the system being measured is a processor running the MVS operating system (although the underlying theory is independent of the type of operating system). In an MVS system, basic measurement data is usually derived from SMF and RMF data, which you store in the CA MICS Batch and Operation Analyzer (SMF) and Hardware and SCP Analyzer (RMF) areas respectively. For this case study we will narrow our discussion to RMF data and consider the differences between data stored in the MVS Performance Group Activity (SCPPGA) File and the CPU Processor Activity (HARCPU) File.

DEVELOPMENT OF CONCEPTUAL MODEL

If you sum all the processor busy time elements (as stored in the CPUxBTM variables of the CA MICS CPU Processor Activity (HARCPU) File) for a specific processor over a specific time interval, you would arrive at a value of CPU seconds which you could call T - for total CPU time. And if you add all the CPU time (as stored in the CA MICS data element PGACPUTM in the CA MICS MVS Performance Group Activity (SCPPGA) File) for all performance groups over the same interval on the same processor, this amount would add up to a second number of CPU seconds which you could call A - for accountable CPU time.

In capacity planning and performance analysis, the value T is desirable, for it represents every second of the interval during which the processor is busy performing some function. Therefore, this number represents the utilized capacity of the processor. Unfortunately, the number T is a very gross value in that it simply measures how often the processor is busy, but not what or whose work is being processed. This is a function of the input data for the HARCPU File (the RMF Type 70 record), which contains no detailed analysis of the workload.

On the other hand, the value A, which is derived from the SCPPGA File, contains a substantial amount of detail concerning what workload is being processed and how much resource is required to process it. This detail is provided by performance group, which is one of the principle methods for subdividing workloads in MVS systems, and which is provided as input to the CA MICS system through the RMF Type 72 records. The value A is also very desirable in capacity planning and performance analysis studies, for this data can pinpoint who is doing the work, not just how much work is being done.

If the values of A and T are equal over the exact same time intervals, there would be no problem or need for further analysis. The numbers are, in fact, not equal. This inequality is the basis for all the mystery surrounding the analysis of capture ratios. It is always the case that A is less than T, and usually by a substantial margin. You can formulate this situation using the equation below:

$$T = A + U \quad (\text{Eqn 6})$$

In Equation 6, you can think of the variable U as the uncaptured CPU time. That is, this is the number of CPU seconds during the interval that were unaccounted for in any performance group that is active on the system. Nevertheless, these CPU seconds are expended doing services for active tasks. Examples of such unaccounted services would be certain paging and swapping operations, a certain amount of SVC processing, a certain amount of I/O operations, etc. See the data element \$SMFCPU for a more detailed discussion of accountable CPU time.

APPLICATION OF CONCEPTUAL MODEL

The analysis is more complicated than Equation 6 leads you to believe. The accountable time (A) is not viewed as a single number, but as separate accountable CPU times for a certain number of workloads that subdivide the entire processor workload. Consider the following example.

The system being examined is SYS1 (this is the CA MICS SYSID value). This system processes a varied workload, which can be divided by performance group. The workloads are a TSO subsystem, an IMS subsystem, a CICS subsystem, and a BATCH subsystem. There are also system-related tasks that must execute to provide an array of common services, such as GRS, JES2, VTAM, TCAS, and RMF.

The following performance group assignments are made in the IEAIPS01 member of SYS1.PARMLIB:

TSO: 2, 40

BATCH: 1, 3, 16, 23

CICS: 9

IMS: 6, 11, 14, 18

SYSTEM TASKS: 4, 5, 7, 8, 10, 12

(The fact that these performance groups do not define a contiguous group is of no concern. The performance groups listed above constitute ALL performance groups currently running on SYS1--not including performance group 0.)

Previously, the capacity planner established two resource element files to contain CPU time information concerning these groups and processor use as a whole. These files are called:

TOT - Total Usage File

WKL - Workload Usage File

In our example, the total CPU time (which we called T above) is represented in the TOT file by the data element TOTCPU. The accountable CPU time is broken into several subsystem CPU times. These are represented in the WKL file UWWKLC01 by the elements TSOCPUTM, BATCPUTM, IMSCPUTM, CICCPUTM, and SYSCPUTM. The sum of these five elements is the accountable CPU time (which we called A above).

The formula is not the simple one of Equation 6, but the slightly more complex one of Equation 7:

$$\begin{aligned} \text{TOTCPU} &= \text{TSOCPU} + \text{BATCPU} + \text{IMSCPU} + \text{CICCPU} + \\ &\quad \text{SYSCPU} + U \end{aligned} \quad (\text{Eqn 7})$$

IMPACT OF UNCAPTURED SYSTEM TIME

As we noted previously, U , the uncaptured CPU time, is due to system functions such as paging, swapping, SVC processing, etc. Such functions are not used in equal proportions by all MVS subsystems. For example, TSO tends to use paging/swapping services more heavily than do batch or IMS. And TSO, IMS, and CICS tend to be heavier SVC users than batch. On the other hand, batch tends to use I/O more heavily than online systems.

Consequently, the proportion of the value U that is contributed by the TSO subsystem may be much different from that contributed by the IMS system and this is similarly true of all the subsystems. Furthermore, the amounts of U contributed by each subsystem are also completely dependent upon the configuration of the individual MVS system. For example, subsystems that are heavily storage-fenced contribute much less to the amount U than unfenced subsystems do; processors with expanded storage contribute different amounts to the U value than the same subsystem on a comparable processor using DASD paging/swapping systems; and the MVS release can cause differences in the apportionment.

Thus, even Equation 7 does not adequately describe the situation, and so we need to formulate Equation 8:

$$\begin{aligned} \text{TOTCPU} &= m1 * \text{TSOCPU} + m2 * \text{BATCPU} + m3 * \text{IMSCPU} \\ &\quad + m4 * \text{CICCPU} + m5 * \text{SYSCPU} + b \end{aligned} \quad (\text{Eqn 8})$$

The values $m_1 \dots m_5$ are multipliers that are applied to the various CPU times. The difference (for example, $(m_1 * TSOCPUTM) - TSOCPUTM$) is the amount of uncaptured CPU time that has been contributed by the subsystem. One can see that this capture ratio problem has now been stated in the form of a multilinear equation, which will therefore lend itself to solution by multilinear regression.

The values m_1 through m_5 are sometimes called the multipliers. Their inverses are called capture ratios, that is, the value $1/m_1$ is the TSO capture ratio for the TSO system, while $1/m_4$ is the CICS capture ratio, and so on. The reason that $1/m_1$ (for example) is referred to as the capture ratio is that it can be interpreted as the proportion of total TSO CPU utilization accounted for in the CA MICS MVS Performance Group Activity (SCPPGA) File.

From the above discussion, you can draw some conclusions about the values $m_1, \dots, m_5$ and b . First, the values $m_1, \dots, m_5$ should all be positive values. Furthermore, from our discussion in the previous two paragraphs, you can see that the values $m_1, \dots, m_5$ must be greater than 1. These are important observations, for if a model that we are evaluating results in m values that are less than 1, this model will fail our third test - the common sense test.

INTERPRETATION OF MODEL

Sections 9.6.1.1 through 9.6.1.3 explain each of the reports in our case study. These reports represent the final model that is evaluated after several preliminary attempts are discarded. In the preliminary models, the variable SYSCPUTM is retained as an independent variable in the model, which always results in an m value for SYSCPUTM less than 1. Finally, SYSCPUTM was eliminated as an independent variable, which resolves that problem.

Why did this situation happen? SYSCPUTM represents the CPU utilization of all the common system services on the processor. As it turns out, the values of SYSCPUTM are related to the values of the other independent variables in a linear fashion. This situation, in which one or more of the independent variables are related to each other by a basically linear relationship, is termed colinearity, and invariably has a bad effect on multilinear regression models. It usually reveals itself as a model that fails test 3 - the common sense test. You must be cautious of hidden linear relationships between the independent variables. When you delete such relationships, either combine related workloads into a single "lump" workload, or drop one of the workloads altogether, which is the method we use in our example.

PROBLEMS OF INTERPRETATION

Think of value b as a "noise" factor (that is, the amount of idle CPU time within the interval in the operating system). The situation when the value of b is rather large for certain operating system releases is sometimes referred to as the low utilization effect. This means that at low CPU utilization rates, the operating system tends to consume a large amount of CPU cycles, although its task is only to search for more work to do.

Because the value of b represents the amount of CPU busy time consumed in a system with no load upon it, this value can sometimes be unreliable when determined by multilinear regression techniques. This is because most production systems rarely, if ever, experience periods with no load upon them (systems with no load for substantial periods of time are usually halted for economic reasons). Therefore, it is unlikely that there will be many, if any, historical data points representing a no-load period. At low utilization, total CPU busy time tends to be a non-linear function of the subsystem CPU times (for example, BATCPUTM, TSOCPUTM), and therefore a linear model may badly estimate the value b while providing very good predicted values for total CPU use at higher utilization levels. In short, do not blindly trust the value b determined in a multilinear regression model without more investigation.

Figure 9-8 illustrates a typical case of low utilization effects. The asterisks (*) illustrate the placement of historical data -- mostly in the upper utilization areas -- while the plus signs (+) illustrate the points of the linear regression model. The point labeled *b* shows where the regression predicts the intercept to be, while the point labeled *true b* shows where the real utilization curve intersects the Y axis. The true utilization curve departs from the predicted linear model at around 20% utilization in our illustration, and runs along the curve depicted with minus signs (-).

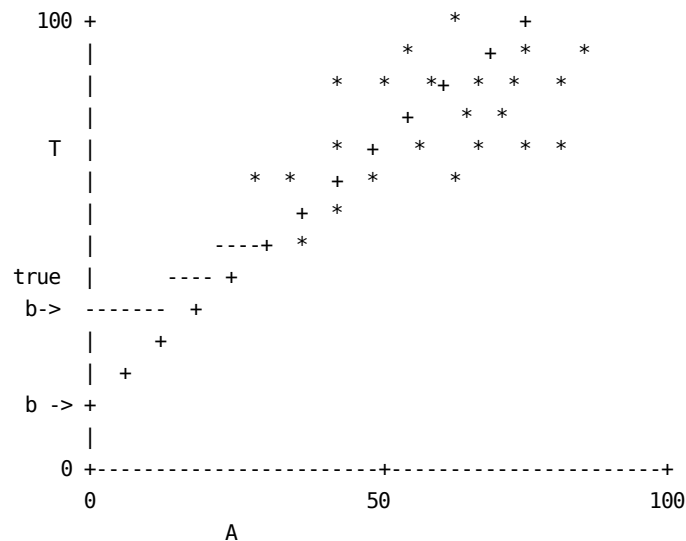


Figure 9-8. The Low Utilization Effect

PREDICTIONS OUTSIDE THE RANGE OF MEASUREMENT DATA

The problem introduced by low utilization effects is really a specific instance of a more widespread problem in multivariate regression analysis. This is the problem of predicting values of the dependent variable outside the range of the independent variables. The most basic underlying assumption in the use of multilinear regression analysis is that the relationship exhibited between the dependent and independent variables is linear. While this may be essentially true in the range of the independent variables, it may not be true outside of that range. The technique of multivariate regression has no way of predicting, or even warning, of potential non-linearity outside of the range of the variables. It is your own experience and knowledge that must provide the guidelines in these situations.

Now that we have expanded on some of the theory of capture ratios, and have defined the case study, the next sections explore the results of the reports that are generated from the multilinear regression model we developed for our capture ratio study:

- 1 - Control Parameters
- 2 - Model Analysis Report
- 3 - Residual Analysis Report

9.6.1.1 Control Parameters

Using a short set of menus that are described in Section 1.5 of this guide, you will arrive at the Multivariate Regression Forecasting screen shown in Figure 9-9.

Figure 9-9 shows the control parameters that we used to generate the final multivariate regression model in this case study.

```

/----- Multivariate Regression Forecasting -----\
|Command ==>
|Enter a ? in any data entry field for more information on valid values.
|Composing CA MICS Inquiry: CAPRAT - Capture ratios on test workload
|
|Report title ==> Capture Ratios on test workload
|
|Selection criteria:
|    Dependent file ==> TOT -
|    Dependent element ==> TOTCPUTM
|    Start date ==> _____ (ddmnyyyy)
|    SYSID ==> SYS1
|    Zone ==> 1_____
|    No. weeks ==> _____ (1-9999)
|    Specify CAPAPU Values ==> N (Y/N)
|Save forecast ==> ____ (YES/NO/AGE)
|Independent file ==> WKL - WORKLOAD USAGE FILE
|Independent elements ==> BATCPUTM CICCPUTM IMSCPUTM TSOCPUTM _____
|
|-----
|Confidence limits (percent) ==> __ (70/90/95)
|Min R-square improvement ==> 0.001 (0.000-0.100)
|Delete observations ==> N (Y/N/R)
|
\-----/

```

Figure 9-9. Capture Ratio Case Study Control Parameters

Note that in this case study the analyst previously generated two resource element files. The first, TOT, contains a data element, TOTCPUTM, which represents the total systemwide CPU time. This file and this data element are specified on the control screen as the Dependent file and the Dependent element. The second file, WKL, contains the data elements that represent the CPU time consumed by each of the workloads running on the system. This file and the corresponding data elements are specified as the Independent file and Independent elements for this case study. Note also that the analyst previously generated and saved forecasts for each of the independent data elements, using one of the forecasting routines of this component. These forecasts are required for Multivariate Regression Forecasting to produce forecasts of the dependent variable.

The analyst also specified that the study would focus on SYSID SYS1 and Zone 1. Because no value is specified for the Confidence limits parameter, we will use the default value of 95%.

In this case study, a Minimum R-squared improvement value of .001 is used for the final model. Sometimes in the evaluation of multilinear models, you may feel that the default value of .05 does not quite provide a fine enough distinction between models to produce the best multilinear regression model. Normally, a value of .01 should be sufficient to make this distinction.

In this particular case study, the first evaluation of the model using the default Minimum R-squared improvement value resulted in values for m of Equation 8 in Section 9.6.1 that appeared to be somewhat high (for example, greater than 2). The analyst therefore attempted the same regression model, but with a Minimum R-squared improvement value of 0.001. This time the stepwise regression process produced a different model that provided m values which were closer to the true values. The regression process may provide different models with the same list of potential dependent variables, and the same input data, but with different Minimum R-squared improvement values. This is due to there being a model of marginally higher r -squared value that fits the data but which the regression process discarded as being irrelevant due to the default Minimum R-squared improvement value, but which it will not discard with a lower value.

There can be a significant amount of subjective judgment involved in the evaluation of regression models. It is highly unlikely that two analysts working independently with the same set of data would produce the same capture ratio coefficients. This does not mean that either of them is necessarily incorrect, since this method of capture ratio determination claims only to be an estimation technique. There is no generally accepted technique for determining what the exact capture ratios are on a specific processor. Since the uncaptured processor time is by definition not distributed to the various workloads by the operating system, there can be no method of independently checking the results. Thus, the results of any capture ratio analysis are only an estimate.

9.6.1.2 Model Analysis Report

Figure 9-10 shows the Model Analysis Report for this case study. The model that was chosen by the stepwise regression process is the last one listed. It has an r -squared value of .97 and acceptable F and p values, so that the first test is passed.

The only variable that gives some concern is the BATCPUTM independent variable, which has a somewhat high value for the p statistic. You can interpret the p statistic for each of the independent variables as the probability that the coefficient (in this case 1.11713) is 0. While a 3.45% probability is not really very large, it is considerably higher than its neighboring independent variables. In this case, the cause of this somewhat higher p statistic can be explained by the fact that the batch workload is a much smaller workload than the others, and therefore variances in the workload may be harder to account for in a linear model. In any event, you should not be very concerned with this p statistic, since the capture ratio for batch that results in this model seems very reasonable.

Recall from the discussion in Section 9.6.1 that the capture ratio is equal to the inverse of the regression coefficient, m . The capture ratio values resulting from this model are:

$$\text{TSO} : 1 / 1.42295 = 0.70$$

$$\text{BATCH} : 1 / 1.11713 = 0.90$$

$$\text{IMS} : 1 / 1.32701 = 0.75$$

$$\text{CICS} : 1 / 1.28791 = 0.77$$

Finally, note that the b value for this model is negative, violating the common sense test. So why do we consider this final model an acceptable one?

Recall the discussion of the introduction in Section 9.6.1 concerning low utilization effects. Examination of the input historical data series (see Figure 9-11) reveals that all of the historical data points occur at moderately high to very high CPU utilization rates. Consequently, the predicted value for b is meaningless, since there are no historical data points to guide the regression when the load is light. Therefore, you can determine that the value of b in this case, even though it is negative, is unimportant and uninteresting.

CA MICS Capacity Planner							
ANALYSIS OF MULTIVARIATE REGRESSION MODELS							
MODEL OF: TOTCPUTM							
BASED ON: TSOCPUTM BATCPUTM IMSCPUTM CICCPUTM							
R**2	F	P	INTERCEPT	NAME	COEFFICIENT	F	P
0.94	1684.11	.0001	44:16:57.0	CICCPUTM	2.96864	1684.11	.0001
0.95	977.72	.0001	29:12:42.3	IMSCPUTM	0.594775	16.58	.0001
				CICCPUTM	2.72044	895.89	.0001
0.96	1286.48	.0001	-2:19:38.6	TSOCPUTM	2.88379	1195.40	.0001
				IMSCPUTM	1.63519	222.39	.0001
0.97	1127.97	.0001	10:57:58.3	TSOCPUTM	1.77499	71.77	.0001
				IMSCPUTM	1.14392	77.90	.0001
				CICCPUTM	1.13065	31.88	.0001
0.97	877.25	.0001	-3:58:16.7	TSOCPUTM	1.42295	29.20	.0001
				BATCPUTM	1.11713	4.60	.0345
				IMSCPUTM	1.32701	74.89	.0001
				CICCPUTM	1.28791	37.62	.0001

Figure 9-10. Capture Ratio Case Study Model Analysis Report

9.6.1.3 Residual Analysis Report

The first page of the Residual Analysis Report for this case study is shown in Figure 9-11. Examination of this report confirms that TEST 2 (visual inspection of the predicted versus the actual data) passed acceptably. Comparison of TOTCPUPTM and PREDICTED TOTCPUPTM shows little difference between the actual and predicted values.

CA MICS Capacity Planner									
MULTIVARIATE REGRESSION FORECAST AND RESIDUAL ANALYSIS									
MULTIVARIATE REGRESSION BASED FORECAST OF: TOTCPUPTM									
MODEL BASED ON INDEPENDENT ELEMENTS : TSOCPUTM BATCPUPTM IMSCPUTM CICCPUTM									
DATE	TOTCPUPTM	INDEPENDENT-ELEMENT NAME	VALUE	PREDICTED TOTCPUPTM	RESIDUAL TOTCPUPTM	-----PLOT OF RESIDUALS-----			
						-18723	0.0		95% CONFIDENCE LIMITS
04JAN97	95:34:40.7	TSOCPUTM	47321	91:47:15.0	3:47:25.7		0.0		1:26:03.7
		BATCPUPTM	37461				0		
		IMSCPUTM	118575				0		
		CICCPUTM	60716				0		
11JAN97	91:11:54.9	TSOCPUTM	47779	90:36:08.0	0:35:46.9		0		1:20:01.3
		BATCPUPTM	37663				0		
		IMSCPUTM	115851				0		
		CICCPUTM	59529				0		
18JAN97	93:57:10.4	TSOCPUTM	49732	92:14:06.2	1:43:04.2		0.0		1:04:26.9
		BATCPUPTM	36463				0		
		IMSCPUTM	119521				0		
		CICCPUTM	59195				0		
25JAN97	92:25:59.9	TSOCPUTM	52382	93:46:39.6	-1:20:39.6		-----0		1:02:14.1
		BATCPUPTM	35767				0		
		IMSCPUTM	121372				0		
		CICCPUTM	59275				0		
01FEB97	94:15:30.1	TSOCPUTM	52535	94:58:53.1	-0:43:22.9		--0		1:03:10.5
		BATCPUPTM	36859				0		
		IMSCPUTM	122859				0		
		CICCPUTM	59992				0		
08FEB97	96:10:44.4	TSOCPUTM	54022	97:40:15.2	-1:29:30.8		-----0		1:08:11.3
		BATCPUPTM	36984				0		
		IMSCPUTM	126560				0		
		CICCPUTM	61944				0		

Figure 9-11. Capture Ratio Case Study Residual Analysis Report

9.6.2 IMS Control Region Case Study

The IMS control region case study examines the IMS subsystem, but the application is very similar in aim and execution to the capture ratio case study.

OVERVIEW OF IMS

Think of an IMS system as a collection of address spaces (at least two) that function together to process certain types of batch and/or transaction-oriented workloads. The address spaces fall mainly into two categories: a control region and dependent regions. There is only one control region for an IMS system (although with IMS Version 2, this distinction becomes somewhat murky, since the DL/I and DBRC functions previously contained within the control region have been moved outboard to separate "dependent" address spaces. We can ignore this exception for the moment and assume that all IMS control functions are contained in one place).

The control region provides the control functions necessary to keep the IMS system running as a unified database/data communications system: integrity, recovery, data sharing, dispatching, queuing, etc. The dependent regions are the address spaces that process the user-generated workload, both batch job and transaction. Dependent regions execute under the control and guidance of the IMS control region in much the same way that all executing tasks within a processor run under the control and guidance of the operating system.

For our case study, we are looking at two types of dependent regions: the BMP (Batch Message Processing region), and the MPP (or Message Processing Program). The BMP can be described as a batch job that runs under the guidance and control of the IMS control region, while an MPP is a task (job or started task) that processes transactions normally originating at terminals and returns output to terminals.

Just as all executing tasks within a processor incur overhead due to the control functions exercised on their behalf by the operating system, so do all IMS BMP and IMS transactions incur control region overhead. The more transactions processed, or BMPs run, the more CPU time that is consumed by the control region on their behalf. In large IMS systems, the amount of CPU time used by the control region alone can be quite significant (30 to 50% of a processor), and this utilization must be taken into account in the capacity planning process.

If, for example, the capacity planner of the XYZ Corporation knows that the IMS transaction load will increase by a certain amount, and the IMS BMP workload will increase by a certain amount, then the capacity planner still must determine how much increase will occur in control region utilization before adequately predicting the growth requirements for the IMS system.

To do this, the capacity planner guesses that the problem of determining the control region utilization from the MPP and BMP region utilizations could probably be estimated by a multilinear regression model. Assume that the control region has been established in its own performance group in the IEAIPS01 member of SYS1.PARMLIB. Furthermore, assume that the MPP regions have been established in one (or more) separate performance groups, and that the BMPs also execute in their own performance group(s). The key point is that the control, MPP, and BMP regions' utilization may all be isolated from one another by performance group assignments.

The capacity planner can then establish a single capacity planning database resource element file to archive utilization statistics from the CA MICS SCPPGA File. This resource element file is called IMS. The CPU time variables in this file are called CTLCPUTM, MPPCPUTM, and BMPCPUTM (denoting aggregated CPU time for all control regions, message processing regions, and batch message regions, respectively).

The capacity planner reasons that if MPP processor use increases, there is a strong likelihood that control region use will increase by a proportional amount. The same is also true of BMP use. Therefore, a model of control region use would be:

$$\text{CTLCPUTM} = m1 * \text{MPPCPUTM} + m2 * \text{BMPCPUTM} + b \quad (\text{Eqn } 9)$$

You could interpret this formula by saying that each additional CPU second of MPP processor use incurs $m1$ CPU seconds of control region use, while each additional second of BMP use incurs an additional $m2$ seconds of control region use.

The interpretation of the value *b* is similar to that of the capture ratio case study (Section 9.6.1). This is a "noise" factor, or the amount of processor busy time used by the control region during a no-load period. In other words, the control region uses this amount of CPU time just searching for work to do. However, as in the case of the capture ratio example, you must take care in evaluating the value of *b*. See Section 9.6.1 for a discussion on the caution needed when predicting outside the range of historical data.

The next sections explore the results of this model:

- 1 - Control Parameters
- 2 - Model Analysis Report
- 3 - Residual Analysis Report

9.6.2.1 Control Parameters

Figure 9-12 shows the control parameters used to generate the final multilinear regression model of this case study.

```

/----- Multivariate Regression Forecasting -----\
|Command ==>
|Enter a ? in any data entry field for more information on valid values.
|Composing CA MICS Inquiry: IMSOVH - IMS Control Region Overhead Study
|
|Report title ==> Capture ratios on test workload
|
|Selection criteria:
|    Dependent file ==> IMS - IMS Analysis File
|    Dependent element ==> CTLCPUTM
|    Start date ==> _____ (ddmnyyyy)
|    SYSID ==> SYS2
|    Zone ==> 1_____
|    No. weeks ==> _____ (1-9999)
|    Specify CAPAPU Values ==> N (Y/N)
|Save forecast ==> _____ (YES/NO/AGE)
|Independent file ==> IMS - IMS Analysis File
|Independent elements ==> MPPCPUTM BMPCPUTM _____
|
|-----
|Confidence limits (percent) ==> __ (70/90/95)
|Min R-square improvement ==> 0.001 (0.000-0.100)
|Delete observations ==> N (Y/N/R)
|
\-----/

```

Figure 9-12. IMS Control Region Case Study Control Screen

Note that the same resource file (IMS) was used for both the dependent variable (CTLCPUTM) and the two independent variables (MPPCPUTM and BMPCPUTM). Note also that the analyst previously generated and saved forecasts for each of the independent variables, using one of the forecasting routines of this component. This is required for Multivariate Regression Forecasting to produce forecasts of the dependent variable.

In this case study, a Minimum R-squared improvement value of .001 is used for the final model. Sometimes in the evaluation of multilinear models you may feel that the default value of .05 does not quite provide a fine enough distinction between models to produce the best multilinear regression model. Normally, a value of .01 should be sufficient to make this distinction.

In this study, the first execution of the program uses the default value for Minimum R-squared improvement. However, this value is not fine enough to distinguish between the two models listed in the Model Analysis Report, shown in Figure 9-13. When this type of situation occurs, the algorithm chooses the first model, because the second model failed to provide the Minimum R-squared improvement you specified. The model that the stepwise regression process chooses is therefore the one which uses only MPPCPUTM as the independent variable. In the example, the analyst decided to see the results of a model that chose both MPPCPUTM and BMPCPUTM as independent variables before committing to one or the other. Therefore the program was rerun with a lower Minimum R-squared improvement value to force the choice of the second model.

9.6.2.2 Model Analysis Report

The Model Analysis Report in Figure 9-13 shows two models for evaluation. The best one-variable model includes only MPPCPUTM as the independent variable, while there is only one two-variable model to evaluate.

CA MICS Capacity Planner							
ANALYSIS OF MULTIVARIATE REGRESSION MODELS							
MODEL OF: CTLCPUTM							
BASED ON: MPPCPUTM BMPCPUTM							
R**2	F	P	INTERCEPT	NAME	COEFFICIENT	F	P
0.98	5806.11	.0001	3:06:54.3	MPPCPUTM	0.293982	5806.11	.0001
0.98	3286.01	.0001	1:14:09.0	MPPCPUTM	0.277912	2763.89	.0001
				BMPCPUTM	0.170892	17.50	.0001

Figure 9-13. IMS Control Region Case Study Model Analysis Report

As we noted previously, on the first execution of the model the analyst used the default Minimum R-squared improvement value of 0.05, which resulted in a choice of the first model (the one-variable model). The re-execution of the analysis with a Minimum R-squared improvement value of 0.001 resulted in the two-variable model. Both models show excellent values of r-squared, F, and p, with the r-squared values for both models less than 1% apart.

In a situation like this, it is hard to defend the choice of one model over the other on the basis of statistical evidence alone. The analyst decided to use the two-variable model since it takes into account both types of dependent regions and therefore (in the analyst's opinion) passes the common sense test with a marginally better score than does the one-variable model.

The two-variable model indicates that the value of m1 is 0.277912 and the value of m2 is 0.170892. This means that each additional CPU second of MPP processor utilization will generate an additional 0.277912 CPU seconds of control region overhead. Similarly, each additional CPU second of BMP processor workload will generate an additional 0.17 CPU seconds of control region overhead.

9.6.2.3 Residual Analysis Report

The first page of the Residual Analysis Report for this case study is shown in Figure 9-14. Examining this report confirms that TEST 2 (visual inspection of the predicted versus the actual data) passed acceptably. Comparison of CTLCPUTM and PREDICTED CTLCPUTM shows a very close agreement between the actual and predicted values.

The independent variable names for this model are listed under the column INDEPENDENT-ELEMENT NAME, which indicates that this model is the two-variable model.

CA MICS Capacity Planner									
MULTIVARIATE REGRESSION FORECAST AND RESIDUAL ANALYSIS									
MULTIVARIATE REGRESSION BASED FORECAST OF: CTLCPUTM									
MODEL BASED ON INDEPENDENT ELEMENTS : MPPCPUTM BMPCPUTM									
DATE	CTLCPUTM	INDEPENDENT-ELEMENT NAME	VALUE	PREDICTED CTLCPUTM	RESIDUAL CTLCPUTM	-----PLOT OF RESIDUALS-----			95% CONFIDENCE LIMITS
						-8111	0.0		
03JAN98	11:05:54.2	MPPCPUTM	94201	11:13:28.4	-0:07:34.2		-0		0:19:20.9
		BMPCPUTM	57228				0		
10JAN98	11:56:38.0	MPPCPUTM	90148	11:03:29.5	0:53:08.4		0		0:22:46.0
		BMPCPUTM	60315				0		
17JAN98	10:25:08.5	MPPCPUTM	91410	11:10:32.4	-0:45:23.9		0		0:22:59.2
		BMPCPUTM	60737				0		
24JAN98	10:48:24.3	MPPCPUTM	83932	10:43:06.5	0:05:17.8		0		0:26:46.6
		BMPCPUTM	63266				0		
31JAN98	11:55:52.8	MPPCPUTM	90346	11:04:00.0	0:51:52.8		0		0:22:35.8
		BMPCPUTM	60171				0		
07FEB98	11:07:44.0	MPPCPUTM	87145	10:45:01.8	0:22:42.2		0		0:21:43.3
		BMPCPUTM	58716				0		
14FEB98	11:44:51.8	MPPCPUTM	95966	11:20:54.8	0:23:57.1		0		0:18:52.5
		BMPCPUTM	56970				0		
21FEB98	10:57:17.1	MPPCPUTM	94796	11:17:09.4	-0:19:52.3		0		0:19:31.7
		BMPCPUTM	57553				0		
28FEB98	11:42:45.9	MPPCPUTM	94842	11:24:42.0	0:18:03.8		0		0:21:51.3
		BMPCPUTM	60127				0		
07MAR98	11:26:49.1	MPPCPUTM	97073	11:37:16.9	-0:10:27.8		0		0:22:17.1
		BMPCPUTM	60916				0		

Figure 9-14. IMS Control Region Case Study Residual Analysis Report

Chapter 10: BUSINESS ELEMENT FORECASTING

The Business Element Forecasting software employs a methodology that uses business factors derived from sources outside of the data processing installation to forecast capacity requirements. This differs from the other forecasting methods implemented in the CA MICS Capacity Planner because the future requirements are not estimated from historical patterns of computer resource consumption. Instead, you assume that the demand for computer services is directly related to the amount of activity occurring in other departments or divisions of the organization. The methodology provides tools for determining if such a relationship exists, and for forecasting from this relationship.

Chapter 3 discusses the structure of business element history files and the structure of resource element history files. It also presents the structure and content of the forecast files which contain forecasts resulting from Business Element Forecasting.

Note: SAS/STAT is a prerequisite for using Business Element Forecasting.

This section contains the following topics:

- [10.1 Functional Description](#) (see page 780)
- [10.2 Usage Guidelines](#) (see page 781)
- [10.3 Standard Output](#) (see page 786)
- [10.4 Analytic Technique Tutorial](#) (see page 794)
- [10.5 Component Operation](#) (see page 801)
- [10.6 Case Study](#) (see page 811)

10.1 Functional Description

The term "business element" refers to a unit of measurement used to characterize activity in some division of the organization. The objective of Business Element Forecasting is to investigate possible statistical relationships between such measurements and the demand for computer resources. If you can establish such a relationship, you can build a statistical model which relates the level of one or more business elements to the amount of computer resources needed to service that level of business.

Many types of organizations use the values of business elements as a part of business planning. For example, a company which underwrites life insurance has an actuarial department to study demographic factors such as age, health, occupation, and geographic location to determine the expected size of future insurance claims. In addition, most organizations attempt to forecast future sales volumes and production costs as part of their budgetary planning.

Business Element Forecasting allows you to develop forecasts that relate application resource consumption to business elements that you have identified in your installation. These forecasts are based on stepwise multiple regression models. You may apply the program to any of the resource element history files that are produced by the capacity planning database generator (see Chapter 3).

Once you establish a relationship between business elements and capacity requirements, you can use the forecast data for the business elements to drive the capacity planning process. This allows you to base the data processing capacity plan on the expected future workload of the organization as a whole, rather than attempt to estimate future use only from the perspective of historical use of the data processing resource.

The program uses the SAS REG with the stepwise option to fit a multiple regression model to the historical resource use and historical business element values, perhaps suggested by the users of the application. The model is then used to estimate the future resource requirements of the application, using estimates of the business elements provided by the application's users. The SAS REG procedure is used to determine the confidence limits for the forecasts.

10.2 Usage Guidelines

In addition to analyzing the values of r -squared, F , and p statistics of the regression model (discussed in Section 10.4), there are three other techniques for evaluating business element models:

- o Determining when business factors can be excluded from the model
- o Calculating confidence limits for the business element based forecasts
- o Inspecting the model

EXCLUDING BUSINESS ELEMENTS

In the development of multiple regression models, each of the independent variables used to create the model represents a source of potential error. Because of this, one of the objectives of most multiple regression studies is to attempt to identify the smallest set of independent variables that produces a model that accurately represents the historical data.

For example, consider the development of a model that relates the disk EXCPs consumed by a life insurance application to the number of policies in effect, new policies issued, claims processed, and the number of cash value loans in effect. When we used stepwise multiple regression to evaluate these potential business elements, we produced four models. (In the example in the Section 10.4, Analytic Technique Tutorial, we show every potential set of combinations to illustrate how the algorithm decides the order in which the business elements are added to the model. In the following table, we only show the model that results from each step.) The four steps of the regression modeling process are shown as follows:

# of Variables/ Step	Business Elements	R**2
=====	=====	=====
1	New Policies	0.72
2	New Policies Claims	0.81
3	New Policies Claims Policies in Effect	0.87
4	New Policies Claims Policies in Effect Cash Value Loans	0.89

Since each business element introduces another potential source of error into the model, you must weigh the contribution each new term makes to the accuracy of the model relative to the potential error that it introduces. Previous investigators have recommended that, for a new business element to be included in the model, the addition of the term should result in an improvement of r-squared of 0.05 (ART78, SAR79). Some stepwise regression algorithms allow you to specify a minimum r-squared improvement parameter. In these algorithms, the regression stops when adding another term to the model does not result in the specified minimum improvement.

Since such a parameter is not available in the current version of the SAS REG stepwise option, stepwise attempts to use every business element you propose as long as a non-zero increase in the r-squared value results. However, the program post processes the output from the stepwise option to provide you with the ability to specify a minimum r-squared improvement for each business element added to the model. The default for this parameter is 0.05.

In the example shown in the table, each of the first three steps of the regression result in an improvement of r-squared of at least 0.05. However, the addition of the cash value loan term to the model only improves the r-squared value by 0.02. Because of this, Business Element Forecasting would develop a forecast based only on new policies, claims, and policies in effect, unless the parameter was overridden.

CALCULATING CONFIDENCE LIMITS

You can calculate confidence limits for multiple regression models to measure the statistical significance of the forecasted values. The width of the confidence limits is influenced by three factors:

- o The amount of measurement data collected
- o The variability seen in the measurement data
- o The distance between forecasted values and the measurement data

Since a smaller confidence interval implies a better, more reliable forecast, it is desirable to control these factors as much as possible. The amount of data processed is a function of the cost of data collection, so you should collect as much as you can afford. You cannot usually influence the amount of variability in the measurement data. The third factor, the distance between the forecast and the measurement data, provides the best leverage in a business element based forecast.

In other sections of this guide you worked with forecasts based on historical data. The forecast observations estimate future resource demand. As more observations are forecast, the forecasts move progressively farther away from the measured data. With time-based forecasting this is inevitable, since we can never measure the future.

In business element forecasting, this is not necessarily the case. The width of the confidence limits depends on the estimated values of the business elements. As long as the estimates of the future business elements fall within the ranges of the volumes used to build the forecasting model, you can expect the predictions made by the model to have relatively small confidence limits. However, the farther the estimates of the future business elements are from these ranges, the wider you can expect the resulting confidence interval to be.

Thus, if an application is currently being tested with relatively small volumes of input data, forecasts based on the relationship of recorded business elements to actual resource consumption are probably not applicable for predicting resource consumption when the application is in production, processing significantly larger volumes (that is, higher business element values) of input data.

The use of confidence limits assumes that the business element estimates being used are error-free. If you are not able to obtain accurate estimates of the future business, it may not be worthwhile to attempt to develop a model.

Bowie reported this type of problem in her GUIDE presentation in 1980 (BOW80). Bowie developed a multiple regression model of the resources consumed by a CICS-based automobile reservation system. However, the company's business forecasting group did not attempt to develop firm estimates of future rentals but, instead, continually adjusted the size of the rental car fleet to meet demand. Because of this, Bowie reported that her business element based forecasts of future resource consumption for the application were no better than simple regression models developed with smoothed historical resource consumption data.

INSPECTING THE MODEL

Perhaps the most important step in evaluating business element-based forecasting is the inspection of the model. Regardless of how well the proposed model fits the historical data, the most important question that can be asked is: "Does this model seem reasonable?"

Viewed from a statistical level, many business element models may look like an excellent representation of the application, but these models may also fail to satisfy your intuitive understanding of the system. In the case study in Section 10.6, we present an example of this type of model. Our example relates the CPU hours consumed by the accounts receivable application to invoices, updates, and open items.

The statistical tests, using the r-squared and the F statistic, show that the model is statistically valid. However, inspection of the results reveals that two of the three model coefficients are negative and that the intercept value is greater than any of the historical observations. Therefore, the model predicts that if no input data is supplied to the application during a month, 544 CPU minutes would be consumed. Moreover, as the input volumes increase, you can expect the resources consumed by the application to decrease. This model does not make sense as a forecasting tool.

Although you might initially expect the intercept value to be zero, you should expect a positive value, since there are a number of functions (like database backups, control card processing, etc.) that consume resources every time the application runs, regardless of the input volumes. There are several reasons why quirks may sometimes occur when you apply the business element methodology. A few common examples are listed below:

- o Duplicate or erroneous SMF data might be used to determine the monthly resource consumption values. You can expect the CA MICS data validation process to eliminate such errors from models built using the CA MICS database.
- o During low volume weeks there might be extensive reruns of the application. Since CA MICS carries ABEND counts, you can identify these types of problems.
- o The application may be in a state of change. If a more efficient program is implemented during the period in which the data is being collected, you might expect to see a drop in resources, even though volumes are increased.

- o The business elements that are used to develop the model may not be related to computer resource use, even though they are directly related to the volume of work processed by the application. This problem occurs when there are other business elements (not considered in the model) which have a greater effect on the application.

There are any number of other factors that may influence the modeling process. By inspecting the model, you can identify such problems before errant forecasts are produced.

10.3 Standard Output

Output from Business Element Forecasting consists of printed reports and, optionally, data values stored in a forecast file. These are covered in more detail in the following sections:

- 1 - Update to the Forecast File
- 2 - Standard Reports

10.3.1 Update to the Forecast File

You can use the Save forecast optional parameter on the Business Element Forecasting control screen to save forecasted values of the dependent variable in the forecast file that corresponds to the file containing the dependent variable.

Typically, as a user of Business Element Forecasting, you can develop a model using the steps outlined in Section 10.2. This is almost always an iterative process, for the first few choices of model parameters may not be the best choices. Formulate the model and then review the results of that model using the tests described in Sections 10.2 and 10.4. Should the tests fail, you must make corrections and adjustments to the proposed model and rerun the forecast.

Finally, you develop a model that satisfies your requirements. At this point, you can use the model to generate future values for the dependent variable. To do this, future values of ALL the independent variables must exist. By rerunning the same model using the Save forecast option mentioned above, the forecasting process generates the future values of the dependent variable and stores them in the appropriate file. You may obtain the future values of the independent business elements in either of two ways: (1) obtaining estimates from the users of the applications, or, (2) generating forecasts by using one of the forecasting routines. If both types of future value estimates are present, Business Element Forecasting uses those provided by you or the users of the applications.

In the forecast file, future values are stored in a special way that is used by other functions of the CA MICS Capacity Planner to produce comparative graphs and charts. The FILE_NO variable is used to distinguish between the various types of observations within forecast files. The Multivariate Regression, Univariate Model, and Business Element Forecasting facilities store a future value with a FILE_NO of 11, and a confidence range with a FILE_NO value of 12. In addition, each observation of a forecast contains the variables FORYEAR and FORMNDY, FORWEEK, or FORMONTH, which contain the date for the most recent historical data used in making the forecast. These concepts are discussed in more detail in Section 3.2.

10.3.2 Standard Reports

Business Element Forecasting produces two reports, which are discussed in the following sections:

- 1 - Model Analysis Report
- 2 - Residual Analysis Report
- 3 - R-Square Ranking Report

10.3.2.1 Model Analysis Report

The Model Analysis Report provides a summary of the business element regression analysis. This report is used to evaluate whether the model acceptably represents the application. Although the values of r-squared, F, and p indicate the statistical significance of the proposed model, you should inspect the intercept value and coefficients to see if the model is reasonable.

A sample Model Analysis Report is shown in Figure 10-1. The report includes the following information:

MODEL OF: Dependent data element named as the element on the Business Element Forecasting control screen (shown in Figure 10-3).

BASED ON: Business elements proposed by the analyst for the model.

The first four columns of the report are discussed below.

R**2: The r-squared value for the model. Although the SAS REG stepwise option procedure attempts to use every one of the business elements that you suggest, Business Element Forecasting excludes any term that does not result in a minimum improvement in the r-squared value (default 0.05, which may be altered by the DELRSQ control statement). For this example, each subsequent r-squared represents at least an increase of 0.05. Hence, all of the terms are included in the model developed by the program. We recommend an r-squared value of 0.70 and above for the acceptance of models produced using this software.

F: The F statistic for the model. The F statistic is calculated from the ratio of sums of squares of the model. Larger F values indicate a more reliable model. Note that the F statistics may be unreliable for models developed with less than 30 historical data points. An F value is provided for each step made by the regression procedure.

P: The probability that the model proposed is significant, that is, more reliable than forecasting the average value of the historical observations. Although most statistics texts recommend testing the null hypothesis to a 0.001 level, 0.01 is recommended for models based on computer measurement data. A p value is provided for each step made by the regression procedure.

INTERCEPT: The y intercept (b value) for the regression equation. The intercept value should be smaller than the actual historical data observations. Otherwise, a model has been developed that indicates that resource consumption decreases with increases to input.

The next four columns are produced for each business element you specify as independent elements on the Business Element Forecasting control screen (Figure 10-3). In this example, the business elements are invoices, updates, and open items. (These four columns appear under the heading BUSINESS ELEMENTS.) These columns contain the following:

NAME: The name of the business element being analyzed.

COEFFICIENT: The coefficient for the business element in the regression equation. The coefficient is the m value in the regression equation. Negative coefficients indicate that you can expect resource consumption to decrease with increases in the business element value.

F: The F statistics calculated for the business element. This F value indicates the significance of the business element in the model that was developed. Once again, larger values indicate a more reliable model and the value may be misleading for models based on small samples.

P: The probability that the specified business element makes a significant contribution to the model developed. We recommend a 0.01 acceptance value.

CA MICS Capacity Planner							
ANALYSIS OF BUSINESS ELEMENT MODELS							
MODEL OF: BILEXCPS							
BASED ON: INVOICES UPDATES OPN_ITEM							
R**2	F	P	INTERCEPT	-----BUSINESS-ELEMENTS-----			
				NAME	COEFFICIENT	F	P
0.78	17.26	.0089	7768.21	INVOICES	-0.159782	17.26	.0089
0.91	20.86	.0077	9091.08	INVOICES	-0.14591	28.35	.0060
				OPN_ITEM	-.0328758	6.27	.0665
0.98	44.36	.0055	7938.98	INVOICES	-0.122026	47.07	.0063
				UPDATES	0.196209	8.91	.0584
				OPN_ITEM	-.0309463	16.42	.0271

Figure 10-1. Model Analysis Report

10.3.2.2 Residual Analysis Report

The Residual Analysis Report details how the business element based model you propose fits the historical data series. A Residual Analysis Report for the BILEXCPS data element is shown in Figure 10-2. The following information is shown:

BUSINESS

ELEMENT The data element that is named as the dependent
BASED element on the Business Element Forecasting
FORECAST OF: control screen (Figure 10-3).

MODEL BASED The business elements that are used to develop
ON BUSINESS the model. This list may differ from the list
ELEMENTS: of business elements specified in the Model
 Analysis Report. This difference results from
 the deletion of proposed business elements based
 on the minimum r-squared improvement criterion.

The columns of the report are discussed below:

DATE: The date ending value for the week or month
 represented by the observation.

Resource The actual historical observations for the
Usage data element (in this case, BILEXCPS) being
Element modeled. Note that if the last two characters
(BILEXCPS): of the data element name are TM, the program
 assumes that the value is in seconds and uses a
 SAS TIME10.1 format for the report.

Two columns appear under the heading BUSINESS ELEMENTS for each independent element used in the model.

NAME: The name of the business element.

VALUE: The user-supplied business element value.

The remaining columns of the report are as follows:

PREDICTED The y value predicted for the interval using the
(BILEXCPS): regression equation fit to the historical data.
Note that if the last two characters of the data
element name are TM, the program assumes that
the value is in seconds and uses a SAS TIME10.1
format for the report.

RESIDUAL The error term, e, in the regression equation.
(BILEXCPS): The residual is equal to the difference between
the actual historical observation and the
predicted value. The residual value for a point
deleted from the modeling process is of no
concern since the model was not developed to
account for the observation. A SAS TIME10.1
format is used if the variable name ends in
TM.

PLOT OF The residual values plotted over the range from
RESIDUALS: minus to plus the maximum residual value.

CONFIDENCE The confidence limits for the forecast. The
LIMITS: confidence limits bound the potential error (e)
that may exist in the forecasts produced by the
model.

CA MICS Capacity Planner									
BUSINESS ELEMENT FORECAST AND RESIDUAL ANALYSIS									
BUSINESS ELEMENT BASED FORECAST OF: BILEXCPS									
MODEL BASED ON BUSINESS ELEMENTS : INVOICES UPDATES OPN_ITEM									
DATE	BILEXCPS	-BUSINESS-ELEMENTS- NAME	VALUE	PREDICTED BILEXCPS	RESIDUAL BILEXCPS	-----PLOT OF RESIDUALS-----			CONFIDENCE LIMITS +/-
						-58	0.0	+58	
31JAN98	4205.00	INVOICES	24017	4147.62	57.38		0+++++		172.66
		UPDATES	3145				0		
		OPN_ITEM	47752				0		
28FEB98	4268.00	INVOICES	21570	4292.32	-24.32		-----0		145.01
		UPDATES	2960				0		
		OPN_ITEM	51552				0		
31MAR98	3970.00	INVOICES	23411	3932.92	37.08		0+++++		133.57
		UPDATES	2012				0		
		OPN_ITEM	49896				0		
30APR98	4384.00	INVOICES	21644	4391.34	-7.34		--0		171.87
		UPDATES	2500				0		
		OPN_ITEM	45144				0		
31MAY98	3908.00	INVOICES	23476	3961.48	-53.48		-----0		115.22
		UPDATES	2709				0		
		OPN_ITEM	53136				0		
30JUN98	3595.00	INVOICES	26311	3644.67	-49.67		-----0		171.37
		UPDATES	2108				0		
		OPN_ITEM	48384				0		
31JUL98	3651.00	INVOICES	24774	3610.66	40.34		0+++++		157.48
		UPDATES	2103				0		
		OPN_ITEM	55512				0		
31AUG98	.	INVOICES	25000	3764.39	.		0		133.61
		UPDATES	2000				0		
		OPN_ITEM	49000				0		
30SEP98	.	INVOICES	27000	3483.33	.		0		224.90
		UPDATES	2600				0		
		OPN_ITEM	54000				0		
31OCT98	.	INVOICES	23000	3893.67	.		0		292.69
		UPDATES	3150				0		
		OPN_ITEM	60000				0		

Figure 10-2. Extended Example Residual Analysis Report

10.3.2.3 R-Square Ranking

The R-Square Ranking report provides you with an overview of the results of all of the forecasts performed during this forecasting run. The report is presented in ascending r-square sequence and gives the r-square, f-statistic, and p-values for the statistical model generated. The r-squares and p-values for the individual variables in the best model are also provided.

If either the r-square or p-value for the run exceeds the thresholds set for the report, the Warning Flag column will be set to show the which threshold was exceeded. If the flag is set to RS-WARN then the R-square for this particular forecast was below the threshold; if the flag is set to P-WARN then one of the p-values for one of the variables in the model was exceeded. If both thresholds are exceeded then the RS-WARN takes precedence.

The other fields on this report correspond to the equivalent values found on the other reports associated with the forecasting process.

You can use the R-Square Ranking report to review the results of all your forecasts and to indicate whether you need to take further action. For example, you may wish to review the Model Analysis report or the Residual Analysis report to see what may be affecting these results. We recommend that for critical workloads you review these other reports regardless of whether or not they are flagged on this report.

CA MICS Capacity Planner				
R-SQUARE RANKING REPORT				
Testing MVR				
MODEL OF: BILEXCPS				
BASED ON: INVOICES UPDATES OPN_ITEM				
WARNING FLAG	R**2	P	VARIABLE	P
-----	-----	-----	-----	-----
	0.98	.0055	INVOICES	.063
			UPDATES	.0584
			OPN_ITEM	.0271
R-SQUARE WARNING VALUE : 0.50 P-VALUE WARNING THRESHOLD : 0.05				

Figure 10-3. R-Square Ranking Report

10.4 Analytic Technique Tutorial

Historically, one of the most difficult problems encountered by the analyst was the forecasting of the resource requirements of specific applications (ART79). Perhaps the most common approach to this problem is the use of periodic user surveys to elicit estimates of future resource requirements like EXCPs or CPU seconds. Unfortunately, the results of such surveys have been subject to significant error since application users typically have little feel for the relationship of their business activities to resource consumption. Therefore, investigators have proposed the use of multiple regression techniques to build statistical models that relate business activities to resource consumption (ART79, SAR79, GOL79, BOW80).

For example, rather than asking the users of the payroll application to estimate CPU consumption or other resource requirements, you could attempt to build statistical models based on factors like total employees, pay checks produced, and/or other breakdowns of total employees such as salary/hourly or weekly/bi-weekly/monthly pay periods. These factors have been given names such as natural forecasting units, key volume indicators, and natural business units (KOL76, SAR79, BRO78). The CA MICS Capacity Planning Component calls these factors business elements.

This section is an overview of the steps required to establish a Business Element Forecasting model using CA MICS. The primary analytical technique used in this process is stepwise multiple regression. The overview includes the basic statistical concepts and terminology used in the construction of the forecast.

The following topics are presented:

- 1 - Developing Business Element Models
- 2 - Stepwise Multiple Regression Concepts

10.4.1 Developing Business Element Models

You can develop forecasting models based on business elements using a seven-step process. Because of the significant amount of time required to interview the users and developers of each application that you wish to model, we recommend that you reserve business element-based forecasting for significant applications where Univariate Model Forecasting has proved unsuccessful (that is, one or more forecasts failed to track the actual values). The seven development steps are:

1. Interview the developers of the application to determine what they believe are the most significant influences on the execution characteristics of the application (that is, data volume associated with various input record types or input volume of various files).

While this interview probably will not identify actual business elements, you must have a good understanding of the application's characteristics before you interview the application's users.

2. Interview the users of the application to determine the principal volume characteristics of their function. For example, the principal volume characteristics of a department in an insurance company might be new policies, renewals, and claims. You can use the knowledge of the application that you obtained from the developers to make initial guesses as to which of these business elements has the greatest correlation to the resource consumption of the application you are modeling.
3. Obtain historical observations for the business elements identified in the interview. In the insurance example discussed above, you would want to obtain monthly or weekly historical values (depending on the timespan for which you wanted to model the application) for the number of new policies, renewals, and claims. You would also obtain future estimates of these business volumes from the users.
4. Create a business element history file in the capacity planning database that contains the historical and future estimates of the business elements provided by the users.

5. Identify a resource element history file that tracks the computer resources used by the application. Create such a file if it does not already exist. Depending upon the way the installation defines performance groups, reporting performance groups, job names, and program names, you might choose to build this file from the CA MICS SCPPGA, SCP_RG, BATJOB, and/or BATPGM Files.
6. Use Business Element Forecasting to select business elements that have a high correlation to the resource consumption of the application. Use the future estimates of the selected business elements to develop forecasts of the future resource consumption values. Note that the future estimates may be provided by you or generated by one of the forecasting routines of this component.
7. Continue to track the forecast of the application's resource consumption and the estimates of the business elements provided by the application's users. Be aware that accuracy of this methodology is dependent not only on the characteristics of the statistical model proposed, but the accuracy of the business element estimates provided by the application's users.

10.4.2 Stepwise Multiple Regression Concepts

The stepwise multiple regression procedures used by Business Element Forecasting are an extension of the linear regression models presented in Chapters 6 and 7. In the introduction to linear regression provided in Section 6.4.1, we presented the equation for a simple linear regression model as:

$$y = b + m * x \quad (\text{Eqn 1})$$

where b is a constant that defines the mean value of y when x is zero

m is the slope of the line that relates y and x

y is the value to be predicted

x is the value on which the prediction is based

You can expand this simple equation to be a multiple regression equation by replacing the m and x terms in the previous equation by a set of terms as shown in the following equation:

$$y = b + \sum_{j=1}^n m_j * x_j \quad (\text{Eqn 2})$$

where b is a constant that defines the mean value of y when x is zero

m_j is the slope of the line that relates y and x_j

x_j is the j -th business element

Although you can easily solve multiple regression equations of this type for any set of x terms (that is, business elements), you are often presented with numerous potential business elements by your application's users. Therefore, you must select some subset of these business factors to build models for each of your application's resource consumption values. While you could make these selections arbitrarily, stepwise multiple regression provides an ideal technique for investigating potential relationships between business elements and resource consumption.

Although numerous stepwise multiple regression algorithms are available, this program uses Goodnight's MAXR algorithm, which is implemented in the SAS REG procedure with the stepwise option. For example, consider the development of a statistical model that attempts to relate the CPU consumption of an application to three potential business elements called A, B, and C. The name, stepwise multiple regression, describes the function of the algorithm. For our example, the algorithm would attempt to solve the problem using three steps:

1. The algorithm would evaluate three linear one-term models (that is, one based on A, one on B, and one on C) to determine which of the three business elements had the greatest correlation to the resource being modeled. This correlation is judged based on the r-squared value for each of the models. The algorithm attempts to maximize r-squared. For the purposes of this example, assume that the B business element had the highest correlation.
2. The algorithm would evaluate two linear two-term models based on B and the remaining two business elements (that is, B and A, and B and C). Once again, the algorithm chooses the best model based on the maximum r-squared. For the purposes of this example, assume that the pair of business elements that resulted in the maximum r-squared value was B and C.
3. The algorithm would evaluate one linear model based on A, B, and C.

The algorithm stops at the step in which it finds the best r-square value. After this point, adding a term to the model does not result in an increase of the r-squared value. Thus, you can employ stepwise multiple regression to investigate the relationship between application resource consumption and a potential business element. Although a discussion of Goodnight's MAXR algorithm is far beyond the scope of this document, we can illustrate the function of the algorithm by considering the following example.

This procedure was used to predict the CPU and EXCP resource requirements of an accounts receivable application (SAR79). The application is characterized by monthly volumes of invoices, updates, and open items which were processed by the installation. In this example the model to be investigated is:

$$y - \text{BILCPUTM} \quad x - \text{UPDATES} \\ 2 \\ x - \text{INVOICES} \quad x - \text{OPN\_ITEM} \\ 1 \quad 3$$

A stepwise multiple regression is used to evaluate a model of CPU time consumption. First, the software evaluates three models of CPU time consumption based on invoices, updates, and open items. The results of these models are shown in the following table:

Model Number	Business Elements	Coef.	Intercept CPU Mins	R**2	F	p
1	Invoices	-0.011	520.2	0.72	12.7	0.02
2	Updates	0.030	180.2	0.36	2.9	0.15
3	Open Items	-0.003	419.0	0.26	1.8	0.24

The coefficient and intercept terms shown in the table correspond to the m and b values discussed in Equation 1. The values of r-squared, F, and p are interpreted below:

- o r-squared is the coefficient of determination. This value indicates what percent of the variability in the historical data is explained by the proposed model. Typically, we are seeking r-squared values greater than or equal to 0.7.
- o F is a measure of statistical quality calculated from the ratio of sums of squares terms for the models. In general, larger F values indicate more reliable models. Note that the F value has very little significance when less than 30 historical data points are used to develop the model.

- o p is the probability that the F is significant. A value less than or equal to 0.01 is recommended for computer measurement applications (SAR79). Note that you can discount the p value if you use only a few historical observations to build the model.

Since the maximum value of r-squared resulted from a model based on invoices, you would select it as the best one-term model that could be developed from the proposed business elements.

The second step of the process is evaluating potential two-term models that you develop from the one-term model developed above. The possible models in this case are based on invoices and updates, and invoices and open items. The results for these models are shown in the following table:

Model Number	Business Elements	Intercept		R**2	F	p
		Coef.	CPU Mins			
1	Invoices Updates	-0.010 0.013	448.5	0.77	6.8	0.05
2	Invoices Open Items	-0.010 -0.002	611.7	0.84	10.6	0.03

Since the maximum value of r-squared resulted from a model based on invoices and open items, it would be selected as the best two-term model that could be selected from the proposed business elements.

The final step of the process is evaluating a three-term model based on all of the business elements supplied by the user. This model is shown in the following table.

Model Number	Business Elements	Intercept		R**2	F	p
		Coef.	CPU Mins			
1	Invoices Updates Open Items	-0.009 0.012 -0.002	544.1	0.88	7.6	0.07

Since the maximum value of r-squared resulted from a model based on invoices, updates, and open items, you would select the best model that you could develop from the proposed business elements.

10.5 Component Operation

The Component Operation Section describes control parameters that you can use to build the Business Element Forecasting analysis inquiries. These parameters allow you to create regression models that relate the resource consumption of an application to user-supplied business elements using stepwise multivariate regression, optionally making use of selective data deletion.

You can produce forecasts with these models using estimates of future business element volumes which you have obtained either from internal business sources or by using one of the forecasting routines of this component. Estimates you enter in the business element history file are recognized by Business Element Forecasting as future estimates by the fact that the DATE data element contains date values greater than the current date when you execute the analysis.

Estimates produced by one of the component forecasting routines are stored in the forecast files (CPFxxx01) and are recognized by the FILE_NO data element, which is 11 if the observation is a future estimate and 12 if the observation is confidence range. Business Element Forecasting always uses the future estimates you provide rather than those produced by the forecasting routines of this component if both are present.

You can optionally store the forecasts produced by Business Element Forecasting in the resource element forecast file corresponding to the dependent file.

Chapter 3 contains detailed descriptions of the screens and procedures necessary to build a business element file, along with descriptions of the methods available for building and updating resource element files.

Because of the optional feature of selective data deletion, you may need to use two screens to develop a business element forecast. The following sections discuss the use of each of these screens.

- 1 - Business Element Forecasting Control Screen
- 2 - Delete Historical Data Observations Screen

10.5.1 Business Element Forecasting Control Screen

Prior to arriving at the Business Element Forecasting screen (shown in Figure 10-4) on your terminal, you will follow a short set of menus which are more fully described in Section 1.5 of this guide. You will see this Business Element Forecasting screen after you indicate that you want to insert a new Business Element Forecast analysis inquiry into your catalog of inquiries. Two fields on this screen will be filled in on the fourth line when you first see it: the inquiry name and the inquiry description, separated by a dash. You entered these two fields on the Insert Capacity Planning Analysis screen when you requested that a new inquiry be added.

```

/----- Business Element Forecasting -----\
| Command ==>
|Enter a ? in any data entry field for more information on valid values.
|Composing CA MICS Inquiry: ADHOC - Adhoc Inquiry - Will Not Be Saved
|
|Report title ==> Adhoc Inquiry - Will Not Be Saved
|
|Selection criteria:
|    Dependent file ==> ___ -
|    Dependent element ==> _____
|    Start date ==> _____ (ddmnyyyy)
|    SYSID ==> _____
|    LVL ==> _____
|    No. wk/mn ==> _____ (1-9999)
|    Save forecast ==> _____ (YES/NO/AGE)
|    Specify CAPAPU Values ==> N (Y/N)
|
|Independent Bus. file ==> ___ -
|Independent elements ==> _____
|
|Confidence limits (percent) ==> __ (70/90/95)
|Min R-squared improvement ==> _____ (0.000-0.100)
|Delete observations ==> N (Y/N/R)
\-----/

```

Figure 10-4. Business Element Forecasting Control Screen

The Business Element Forecasting screen shown in Figure 10-4 presents the control parameters that allow you to develop forecast models based on stepwise multivariate regression. The forecasts produced by the program can be saved in a forecast file corresponding to the file you specify as the source of the dependent variable for the model.

Some of these parameters in Figure 10-4 are required, while others are optional. The following is a discussion of the various parameters as they appear in this figure.

Report title:	The 60-character title field for the Business Element Forecasting reports. No double quotes (") or semicolons (;) may appear within the body of the title.
Dependent file:	The resource element or business element capacity planning database file containing the historical observations to be used for the dependent variable of the regression model. If you are uncertain of the files or file names available in your capacity planning database, type a question mark (?) in the field and a selection list will be displayed.
Dependent element:	The name of the dependent variable in the regression model. If you are uncertain of the correct name of the data elements in the Input file, type a question mark (?) in this field and a selection list will be displayed.
Start date:	The start date for date range selection. If Start date is specified, only historical observations at least as recent as this value are included in the analysis. The discussion of the following control parameter describes the options for controlling date range selection.
SYSID:	The SYSID for which the report is to be produced. If the capacity planning database file is defined as having a composite SYSID, this value is automatically supplied on the screen.

LVL: The ZONE, HOUR, or DAYNAME to be included in the analysis, depending on whether you defined the Dependent file at the ZONE, HOUR, or DAYNAME summarization level. If the Dependent file is a resource element file and you defined it as having a composite ZONE, HOUR, or DAYNAME this value is supplied automatically on the screen. Note that if the Independent file is a resource element file, this requirement does not apply, since business element files are not defined containing the ZONE, HOUR, or DAYNAME data element. Note that the field name will automatically be changed from LVL to either ZONE, HOUR, or DAYNAME to reflect the input file summarization level once you specify the input file.

No. wk/mn: The maximum number of days, weeks, or months of historical observations that will be included in the analysis. Note that once you specify the input file, the field name will change to reflect the timespan of the Input file. When used in conjunction with the Start date parameter, several options are available for date range selection control:

- o If you specify only Start date, all historical observations at least as recent as the value specified for Start date will be selected.
- o If you specify both Start date and No. wk/mn, the date range for data selection begins with the value specified for Start date and extends for the specified number of days, weeks, or months.
- o If you specify only No. wk/mn, the date range for data selection will include only this number of the most recent days, weeks, or months.
- o If you specify neither parameter, then all historical observations will be selected.

Save forecast:	Offers you three options for saving the forecasted observations in the forecast file corresponding to the Dependent file. - o Specifying YES causes the forecasted observations to be saved in the capacity planning database forecast file (CPFxxx01), without aging the file as described in Chapter 3. - o Specifying NO will result in the forecasted observations not being saved in the forecast file. If this option is selected, the forecast file will not be aged. - o AGE will cause the forecasted observations to be saved in the forecast file, and will cause the forecast files to be aged as described in Chapter 3.
Specify CAPAPU Values:	Put a Y in this field to specify the CAPAPU value(s) to be included in the analysis. The CAPAPU value is specified on the CAPAPU Selection Values screen, shown in Figure 10-5 (which appears if you put a Y in this field). You may specify up to nine CAPAPUs on this screen only if the capacity planning database file is defined as containing the CAPAPU data elements.
Independent file:	The resource element or business element capacity planning database file containing the historical observations for the independent variable of the regression model. If you are uncertain of the files or file names available in your capacity planning database file, type a question mark (?) in the field and a selection list will be displayed.

Independent elements: The names of up to ten independent variables for the regression model. If you are uncertain of the correct name of the data elements in the Independent file, type a question mark (?) in this field and a selection list of available data elements will be displayed. The number of Independent elements that you may specify is limited by the number of historical observations in the capacity planning database files. This restriction is best explained by the following logical relationship:

$$n - ie - 1 \geq 3$$

Where n is the number of historical observations and ie is the number of Independent elements that you specify. For example, you must have 14 historical observations in order to specify 10 independent elements $(14-10-1) \geq 3$. This restriction is enforced to provide a minimum of three degrees of freedom for the statistical models developed by the program.

Confidence limits (percent): The degree of confidence to be used in calculating the upper and lower confidence confidence limits for the forecasted values of the Dependent element. The default is 95%.

Minimum R-squared improvement: The r-squared improvement criterion for the regression model. This value excludes any independent element from the regression model whose inclusion does not increase the overall r-squared value by at least this amount. If you do not enter any value for this parameter, Business Element Forecasting uses a value of 0.01.

Specify historical data observations to delete: The dates of specific observations to be excluded from the analysis. Enter a Y for this parameter to obtain the screen shown in Figure 10-7. Enter an R to reset (nullify) any previous date deletion specifications.

If you entered a Y in the Specify CAPAPU values field on the Business Element Forecasting Control Screen, the CAPAPU Selection Values screen displays, as shown in Figure 10-5. On the CAPAPU Selection Values screen, you may specify up to nine CAPAPU values to be included in the analysis. You may specify CAPAPUs only if the capacity planning database file is defined as containing the CAPAPU data elements. You may use wild cards (?) and (*) when specifying CAPAPUs. A question mark (?) replaces one character; an asterisk (*) replaces all the characters after that point.

```

/----- CAPAPU Selection Values -----\
|Command ==>|
|
|Composing CA MICS Inquiry: BEFR01 - FORECASTING USING BUSINESS DATA|
|
|      Rule ----+----1----+----2----+----3----+----4----+----5----+----6----|
|CAPAPU  ==> -----|
|CAPAPU 2 ==> -----|
|CAPAPU 3 ==> -----|
|CAPAPU 4 ==> -----|
|CAPAPU 5 ==> -----|
|CAPAPU 6 ==> -----|
|CAPAPU 7 ==> -----|
|CAPAPU 8 ==> -----|
|CAPAPU 9 ==> -----|
\-----/

```

Figure 10-5. CAPAPU Selection Values Screen

10.5.2 Delete Selected Historical Observations Screen

If you entered a Y in the Specify historical data observations to delete on the Business Element Forecasting screen, the Extended Options Menu appears, as shown in Figure 10-6.

```
/----- Extended Options Menu -----\
|Command ==>                               |
|                                           |
|Inquiry Step:  Business Element Forecasting |
|                                           |
|Print Reports                ==> YES  (YES/NO) |
|Maximum R-Square for report   ==> 0.55  (0.00-1.00) |
|P-Value flag                 ==> 0.05  (0.000-1.000) |
|                                           |
|Delete historical data observations ==> N (Y/N/R) |
|Transform historical observations ==> N (Y/N/R) |
|                                           |
\-----/
```

Figure 10-6. Extended Options Menu

Print Reports:	The Print Reports parameter controls the printing of the Analysis of the Model and Residual Analysis reports. If you type YES, these reports are produced. If you type NO and the forecasted file is NOT a Business Element File, these reports are suppressed. These reports will always be produced for Business Element Files; however, the R-Square Ranking Report is not produced.

Maximum R-Square for report:	The Maximum R-Square for report parameter is used during the production of the R-Square Ranking Report to control the flagging of particular forecasts. If the R-Square of a generated model is less than this value, it will be flagged with an RS-WARN flag. A value of 1.0 in this field will disable the production of the RS-WARN flag in the R-Square Ranking Report.

P-Value flag: The P-Value Flag is also used in the production of the R-Square Ranking Report to determine if any of the P-values of the model variables should be flagged with the P-WARN flag. If any P-value in the model is greater than this value, it will be flagged in the report. A value of 1.0 in this field will disable the production of the P-WARN flag in the R-Square Ranking Report.

Note: The RS-WARN flag takes precedence over the P-WARN flag. This means that if a model exceeds both of these warning criteria, the RS_WARN flag will be printed.

If you specify historical data observations for deletion on the Business Element Forecasting screen, you may specify the deletion of selected historical observations from the modeling process. Although deleted observations still appear in the reports, they are not used to develop the regression model. Use the Delete Historical Data Observations screen shown in Figure 10-7 to delete observations.

[illegible]

Figure 10-7. Delete Selected Historical Observations Screen

Note that the date values you specify on this screen must exactly match the DATE value of the historical observations in the Dependent and Independent files to be excluded from the analysis.

10.6 Case Study

The Case Studies Section provides a case study using Business Element Forecasting, the control parameters used to conduct the analysis, and the resulting analysis reports. The development and maintenance of business element history files and forecast files is discussed in detail in Chapter 3. However, we will discuss the creation of the particular business element database files that we used in the case study analyses in some detail to help you better understand the overall process of Business Element Forecasting.

The following topics are addressed by section:

- 1 - Problem Description
- 2 - Creating a Business Element File
- 3 - Developing a Business Element Model
- 4 - Analyzing a Business Element Model
- 5 - Automating the Collection of Business Element Data

10.6.1 Problem Description

In this case study, assume that the system is an online application that is processing transactions against a central database used for accounts receivable. The transactions perform three basic functions: processing invoices, logging credit updates, and making inquiries against open items. The organization tracks the volumes of each of these business elements for management reporting in several divisions.

The objective of this case study is to investigate whether a relationship exists between the three business elements and the number of EXCPs required by the application. If such a relationship can be established, a business element model can be generated and used to forecast future application resource requirements.

This case study is based on a paper published by Sarna which described an early effort at Business Element Forecasting (SAR79).

10.6.2 Creating a Business Element File

There are three primary business elements to be used in the study: invoices, updates, and open items. Assume that this data is available from sources within the organization and has been collected from these sources. The identification and collection of this data comprises the first three steps of the methodology described in Section 10.4.1.

The following table shows the data from the study that we are using for our example:

Month	Invoices	Updates	Open Items	CPU Mins	EXCPs (1000)
1	24017	3145	49752	273	4205
2	21570	2960	51552	268	4268
3	23411	2012	49896	262	3970
4	21664	2500	45144	280	4384
5	23476	2709	53156	250	3908
6	26311	2108	48384	221	3595
7	24744	2013	55512	230	3651

Following Step 4 in the methodology, you need to create a business element history file in the capacity planning database. You may use one of two methods: a batch update or an online process using the SAS/FSP product. In this case, since the data is already tabulated, the batch update is more appropriate. This allows creation of the business element file containing all the historical data in one process. If the forecasting model is successful, you can use the SAS/FSP update to add new business element data to the file on a monthly basis to continue the forecast tracking process.

The historical observation for invoices, updates, open items, CPU minutes, and EXCPs are shown in the following table. Since dates are required by the software, an arbitrary date was assigned to each observation. To illustrate the use of forecasting, estimates of the business element values for August through October are also included in the table.

Date	Invoices	Updates	Open Items	CPU Mins	EXCPs (1000)
=====	=====	=====	=====	=====	=====
31JAN98	24017	3145	49752	273	4205
28FEB98	21570	2960	51552	268	4268
31MAR98	23411	2012	49896	262	3970
30APR98	21664	2500	45144	280	4384
31MAY98	23476	2709	53156	250	3908
30JUN98	26311	2108	48384	221	3595
31JUL98	24744	2013	55512	230	3651
31AUG98	25000	2000	49000		
30SEP98	27000	2600	54000		
31OCT98	23000	3150	60000		

The Business Element File definition screen shown in Figure 10-8 is used to define the business element file.

```

/----- Business Element File Definition -----\
|Command ==>                                     |
|Modifying CAP DB File def.: BEF - Business Element File |
|Time-span      ==> MONTHS (DAYS/WEEKS/MONTHS)          |
|Starting date   ==> _____ (ddmonyy)                |
|Number of months ==> 24 (for pre-formatting records)    |
|Maximum amount of data available for analysis ==> 60 (1-9999 months) |
|Age of oldest forecast ==> 24 (1-9999 months)           |
|Number of cycles retained: History file ==> 3 (1-99 cycles) |
|Forecast file ==> 3 (1-99 cycles)                       |
|Specify business element file data elements ==> Y (Y/N) |
\-----/

```


Use the Business Element File Data Elements screen shown in Figure 10-9 to define the data elements INVOICES, OPN_ITEM, and UPDATES. The data element DATE is provided automatically.

```

/----- Business Element File Data Entry -----\
| Option ==> 3                                     |
| Modifying CAP DB File def.: BEF - Business Element File |
|                                                     |
|   1 - Initialize Business Element File               |
|     Prime business element file with null entries in preparation |
|     for online data entry.                           |
|   2 - Online Data Entry                             |
|     Update the business element file using SAS/FSP. The file must |
|     have been previously initialized.                 |
|   3 - Batch Data Entry                             |
|     Update the business element file using external card images.  |
|     The card image update can be executed in foreground or batch. |
|                                                     |
| Starting date ==> _____ (ddmonyy)   Number of months ==> 24 (1-9999) |
| Card image DSN ==> USER.CNTL(BEFDATA)              |
|                                                     |
\-----/

```

Figure 10-10. Business Element File Data Entry

Use the Business Element File Data Entry screen shown in Figure 10-10 to read data into the business element file. You select batch data entry instead of preparing the file for online entry or performing the actual online data entry. We also specify the name of the data set, USER.CNTL(BEFDATA) and indicate the name of the file from which the business element data is to be read.

On the CA MICS database Inquiry Execution screen, specify that you plan to execute this file update inquiry in foreground.

A generated file update program reads the business elements from the specified card image data set. Each record must contain a date of the form DDMONYY followed by the business element values in the order that they are specified as shown in the Business Element File Data Elements screen (Figure 10-9). For this example, the following SAS input statement is generated to read the business element data:

```
INPUT DATE DATE7. INVOICES UPDATES OPN_ITEM;
Hence, the following records are supplied in the
USER.CNTL(BEFPARM) file:
```

31JAN98	24017	3145	49752
28FEB98	21570	2960	51552
31MAR98	23411	2012	49896
30APR98	21664	2500	45144
31MAY98	23476	2709	53156
30JUN98	26311	2108	48384
31JUL98	24744	2013	55512
31AUG98	25000	2000	49000
30SEP98	27000	2600	54000
31OCT98	23000	3150	60000

Figure 10-11. Business Element Data Input File

The business element file update includes a report of the contents of the business element file so that successful job completion can be verified. The Business Element File Contents Report shown in Figure 10-12 is generated for this example.

CA MICS Capacity Planner						
OBS	YEAR	MONTH	DATE	INVOICES	UPDATES	OPN_ITEM
1	98	1	31JAN98	24017.00	3145.00	47752.00
2	98	2	28FEB98	21570.00	2960.00	51552.00
3	98	3	31MAR98	23411.00	2012.00	49896.00
4	98	4	30APR98	21644.00	2500.00	45144.00
5	98	5	31MAY98	23476.00	2709.00	53136.00
6	98	6	30JUN98	26311.00	2108.00	48384.00
7	98	7	31JUL98	24774.00	2103.00	55512.00
8	98	8	31AUG98	25000.00	2000.00	49000.00
9	98	9	30SEP98	27000.00	2600.00	54000.00
10	98	10	31OCT98	23000.00	3150.00	60000.00

Figure 10-12. Business Element File Contents Report

10.6.3 Developing a Business Element Model

Once the business element file is created, perform step 5 in the procedure listed in Section 10.4.1 to identify or create a resource element file in the capacity planning database file. You would create such a file using the capacity planning database generator, described in Chapter 3. For this case study, assume that this file has already been created, and that it contains data for the time period under study.

At this point, all the data that you need to perform the study is available in the capacity planning database file, either in a business element file or in a resource element file. The next step in the methodology is to run a stepwise multiple regression to identify the best model for the number of EXCPs performed by the application. Since the business element file contains estimates of the values of business elements in the future as well as the corresponding historical values, this analysis also generates forecasts for the application EXCPs. Figure 10-13 shows the Business Element Forecasting screen with the parameters that we entered to invoke the stepwise procedures.

```

/----- Business Element Forecasting -----\
|Command ==>
|Enter a ? in any data entry field for more information on valid values.
|Composing CA MICS Inquiry: BEFM1 - Business Element Forecasting - BEF FILE
|
|Report title ==> Business Element Forecasting - BEF File
|
|Selection criteria:
|    Dependent file ==> SYS - System Planning - MONTH/ZONE/SYSID/APU
|    Dependent element ==> BILEXCPS
|        Start date ==> _____ (ddmonyyyy)
|        SYSID ==> CMPX
|        Zone ==> A_____
|        No. months ==> ____ (1-9999)
|        Save forecast ==> ____ (YES/NO/AGE)
|        Specify CAPAPU Values ==> N (Y/N)
|
|Independent Bus. file ==> BEF - Business Element File
|Independent elements ==> INVOICES UPDATES OPN_ITEM _____
|
|Confidence limits (percent) ==> ____ (70/90/95)
|Min R-squared improvement ==> _____ (0.000-0.100)
|Delete observations ==> N (Y/N/R)
\-----/

```

Figure 10-13. Business Element Forecasting

On the Business Element Forecasting screen shown in Figure 10-13, we specify the three-character names of the resource element file and the business element file for Business Element Forecasting. For business element-based forecasting studies, you may want to analyze a resource element file that summarizes all of the SYSIDs and ZONEs. Although you may normally schedule the application on a particular SYSID during some particular ZONE, you would not want to exclude the resources consumed by the application when it is late or when it executes on some other SYSID due to hardware outages.

The screen indicates that a business element-based model of the data element BILEXCPS will be developed based on invoices, updates, and open items. Since we provide no minimum r-squared improvement criterion, the incremental r-squared improvement value for this study defaults to 0.05. Since we specify no confidence limits, 95 % confidence limits is the default. Since this execution of Business Element Forecasting generates forecasts of the application EXCPs, we specify on this screen that we want to save the forecast. The forecast will be saved in the resource element forecast file corresponding to the SYS resource element file (the dependent file).

10.6.4 Analyzing a Business Element Model

The Model Analysis Report corresponding to this case study is presented in Figure 10-14. Section 10.3.2.1 contains a detailed discussion of each of the fields appearing on this report.

Several points should be noted in this case study. At first glance the values of r-squared, F, and p all suggest that this is a solid model of the relationship between EXCPs and the specified business elements. In fact, the r-squared value indicates that the proposed model explains 98% of the variability in the historical data.

However, there are negative coefficients associated with the business elements included in the model. If the model is truly valid, these would suggest that, as the business element volumes increase, the expected resource consumption of the application decreases. This does not make sense and you should reject the proposed model for this reason. You should investigate the historical data to determine if reruns or other factors caused anomalies in the historical data.

CA MICS Capacity Planner							
ANALYSIS OF BUSINESS ELEMENT MODELS							
MODEL OF: BILEXCPS							
BASED ON: INVOICES UPDATES OPN_ITEM							
R**2	F	P	INTERCEPT	NAME	COEFFICIENT	F	P
0.78	17.26	.0089	7768.21	INVOICES	-0.159782	17.26	.0089
0.91	20.86	.0077	9091.08	INVOICES	-0.14591	28.35	.0060
				OPN_ITEM	-.0328758	6.27	.0665
0.98	44.36	.0055	7938.98	INVOICES	-0.122026	47.07	.0063
				UPDATES	0.196209	8.91	.0584
				OPN_ITEM	-.0309463	16.42	.0271

Figure 10-14. Model Analysis Report

The Residual Analysis Report details how the business element-based model proposed by the analyst fits the historical data series. The Residual Analysis Report corresponding to the case study is presented in Figure 10-15. Section 10.3.2.2 contains a detailed discussion of each of the fields appearing on the standard report.

Generally, a good model has no pattern in the plot of the residual values. The presence of a pattern in the residuals often suggests that there is some systematic pattern to the historical data that is not being accounted for by the model. For this example, the residuals are very small (the residuals are plotted on a scale from -58 to +58, while the observed values range from 3595 to 4268), so the errors in the predicted values are on the order of 1%. Also, there seems to be no systematic pattern to the residuals. These observations support the conclusion drawn from the model analysis report that, on a purely statistical basis, the model appears valid.

Confidence limits are reported on the right side of the Residual Analysis Report. For a business element model, these limits are a function of the estimates for future business elements rather than simply increasing monotonically as was the case with models developed by Univariate Model Forecasting. Note that the confidence limits for the August forecast are small since the business element estimates provided to the model were in the minimum to maximum range observed in the historical data. However, the confidence limits for September and October are larger since the business element estimates are outside of the range of observed historical values.

C A M I C S C A P A C I T Y P L A N N E R									
BUSINESS ELEMENT FORECAST AND RESIDUAL ANALYSIS									
BUSINESS ELEMENT BASED FORECAST OF: BILEXCPS									
MODEL BASED ON BUSINESS ELEMENTS : INVOICES UPDATES OPN_ITEM									
DATE	BILEXCPS	-BUSINESS-ELEMENTS- NAME	VALUE	PREDICTED BILEXCPS	RESIDUAL BILEXCPS	-----PLOT OF RESIDUALS-----			CONFIDENCE LIMITS +/-
						-58	0.0	+58	
31JAN98	4205.00	INVOICES	24017	4147.62	57.38		0+++++		172.66
		UPDATES	3145				0		
		OPN_ITEM	47752				0		
28FEB98	4268.00	INVOICES	21570	4292.32	-24.32		-----0		145.01
		UPDATES	2960				0		
		OPN_ITEM	51552				0		
31MAR98	3970.00	INVOICES	23411	3932.92	37.08		0+++++		133.57
		UPDATES	2012				0		
		OPN_ITEM	49896				0		
30APR98	4384.00	INVOICES	21644	4391.34	-7.34		--0		171.87
		UPDATES	2500				0		
		OPN_ITEM	45144				0		
31MAY98	3908.00	INVOICES	23476	3961.48	-53.48		-----0		115.22
		UPDATES	2709				0		
		OPN_ITEM	53136				0		
30JUN98	3595.00	INVOICES	26311	3644.67	-49.67		-----0		171.37
		UPDATES	2108				0		
		OPN_ITEM	48384				0		
31JUL98	3651.00	INVOICES	24774	3610.66	40.34		0+++++		157.48
		UPDATES	2103				0		
		OPN_ITEM	55512				0		
31AUG98	.	INVOICES	25000	3764.39	.		0		133.61
		UPDATES	2000				0		
		OPN_ITEM	49000				0		
30SEP98	.	INVOICES	27000	3483.33	.		0		224.90
		UPDATES	2600				0		
		OPN_ITEM	54000				0		
31OCT98	.	INVOICES	23000	3893.67	.		0		292.69
		UPDATES	3150				0		
		OPN_ITEM	60000				0		

Figure 10-15. Residual Analysis Report

10.6.5 Automating the Collection of Business Elements

The business element data was already tabulated at the beginning of this case study, so we chose a batch update to create the business element file. This type of update is appropriate when the business element data is readily available and the study is of short duration.

Once you establish a business element model, you will want to set up procedures for continuing to use the model to track forecasts and predict future capacity requirements. To do so, you must collect business element data from other departments and add it to the capacity planning database on a weekly or monthly basis.

For installations that have the SAS/FSP Product, we have provided a technique to help you collect this data in a regular manner. The technique creates a business element file within the capacity planning database file, which contains only the dates for which planning will be done. All values of the business elements are set to a missing value. This provides a formatted file which you can access using the SAS/FSP interface provided by the Business Element File Data Entry screen. This interface is discussed in Section 10.6.2.4.

The same screens shown in Section 10.6.2 are used for defining this file. When the Business Element File Data Entry screen was displayed, we chose to initialize the business element file instead of selecting batch data entry. This primed the file with null entries in preparation for online entry.

After the file was primed, we once more chose the execute option for the file, but selected the option for online data entry, which updated the file using SAS/FSP.

Chapter 11: WORKLOAD CHARACTERIZATION

The Workload Characterization software, provided in the CA MICS Capacity Planner, enables you to apply a clustering methodology to the analysis of your installation's workload. Workload characterization is an attractive approach to capacity planning and performance management problems because it allows the number of workload elements that need be considered in a study to be reduced from tens of thousands to only a few.

Clustering methods are a statistical extension of scatter plots to identify similarities and differences between workloads. Scatter plots are often difficult to prepare and depend heavily on visual interpretation of the data. The need for visual interpretation limits the use of scatter plots to two or perhaps three axes. Clustering overcomes these disadvantages through its ability to recognize patterns in multiple dimensions. Using the CA MICS database as an input data source for clustering simplifies and extends the application of the technique.

Note: SAS/STAT is a prerequisite for using Workload Characterization.

This section contains the following topics:

- [11.1 Functional Description](#) (see page 824)
- [11.2 Usage Guidelines](#) (see page 825)
- [11.3 Standard Output](#) (see page 832)
- [11.4 Analytic Technique Tutorial](#) (see page 840)
- [11.5 Component Operation](#) (see page 854)
- [11.6 Case Studies](#) (see page 867)

11.1 Functional Description

During the early 1970s, Domenico Ferrari introduced the concept of workload characterization (FER72). The basic premise of the concept was that studies of computer systems could be facilitated by the identification of a number of "resource patterns" that could be used to statistically represent a system's workload. Since capacity planning and performance management studies often encompass tens of thousands of jobs or hundreds of thousands of transactions, the ability to represent the workload by a limited number of resource patterns can significantly simplify a study. (Typical studies result in the identification of from 10 to 30 patterns.)

Ferrari identified these resource patterns, which he called clusters, by visual inspection of two-dimensional scatter plots. For example, Ferrari reported studies where he examined scatter plots of print lines and CPU seconds representing the resource usage patterns of student jobs. Unfortunately, the dependence on visual inspection limited the technique to two or perhaps three dimensions, since graphic representations become extremely complex for higher dimensions. Also, the technique depended upon the visual skill of the interpreter, since some analysts could "see" patterns better than others. These problems limited the applications of what was otherwise a valid concept.

In 1976, however, two papers were published that detailed the application of statistical pattern recognition (clustering) techniques to the workload characterization problem. They allowed problems of higher dimensions to be solved and eliminated the dependence on the analyst's interpretive skill and judgment. Agrawala and Mohr reported on their initial studies of Univac workloads at the University of Maryland (AGR76) and Artis reported on studies of IBM workloads at Bell Laboratories (ART76). Since that time, these and a number of other authors have published numerous papers on the subject in various publications, such as the Computer Measurement Group (CMG) proceedings.

Workload Characterization allows you to apply cluster-based workload characterization techniques to the CA MICS database. The software can be applied to any of the unsummarized files that are supported in the DETAIL timespan. The program identifies clusters (that is, similar patterns of activity) in the workload using the SAS FASTCLUS procedure.

The use of clustering techniques introduces other statistical issues. Among these are scaling and the treatment of outliers. Improper scaling and failure to account for the influence of abnormally large observations are the most common reasons for invalid workload characterization studies. These issues are discussed in detail in Section 11.2.4.

11.2 Usage Guidelines

Before the introduction of cluster-based workload characterization, a number of other workload characterization techniques were utilized. These approaches provide valuable insight into potential errors in the workload characterization process.

The equivalent job approach fails to provide an adequate characterization, since it lacks the important property of robustness. Relatively small changes in the workloads can cause large changes in the definition of an equivalent job. This means that the results of the equivalent job approach are too closely associated with the data used in the analysis to provide general insight into the workloads on the system. Cluster-based workload characterization can replace commonly practiced ad hoc techniques.

The most important issues regarding implementation concern the type of data elements that are used to characterize the clusters and the treatment of unusual points which cannot be represented by clusters. These points, known as outliers, can have a profound effect on the clustering study.

These topics and the use of certain CA MICS database files for clustering analysis are discussed in the following sections:

- 1 - Historical Approaches
- 2 - Basic Hypotheses
- 3 - Application of Workload Characterization
- 4 - Data Selection and Outliers
- 5 - Suggested CA MICS Input Files

11.2.1 Historical Approaches

One early approach presented to capacity planning was the equivalent job methodology presented by Jenkins in 1977 (JEN77). Most simply described, the resources consumed by all of the workload elements (for example, job, steps, transactions) over some number of weeks or months are divided by the number of workload elements that were executed during the period. The result of this division is called an equivalent workload element.

An equivalent workload concept initially appears to be very attractive since you can estimate future resource requirements by simply estimating the number of equivalent workload elements. This value is multiplied by the equivalent workload description to calculate the future resource requirements.

11.2.2 Basic Hypotheses

The problems associated with the equivalent workload element approach can best be introduced by stating the two hypotheses that were introduced by Artis for testing proposed workload characterizations (ART76). These hypotheses are the "representative hypothesis" and the "stationary hypothesis."

The representative hypothesis tests how well the proposed characterization represents each element of the workload. Rather than testing the proximity of the mean value of the workload to the proposed characterization, you must evaluate the differences between the proposed characterization and each element of the workload.

For example, consider the pathological example of a system that runs 100 testing jobs using less than 5 CPU seconds and 100 production jobs using about 10 CPU minutes each day. (Of course, real workloads are much more complex.) This would result in an equivalent job that required about 5 CPU minutes. Unfortunately, no job in the workload has resource requirements similar to this equivalent job. Therefore, rather than simplifying the problem by reducing the number of workload elements to be considered, we have changed the nature of the problem.

The characteristics of a system that runs 200 equivalent (that is, average) jobs per day would be much different from those of the hypothetical system discussed above. It is interesting to note that the equivalent workload element can only hope to pass the representative hypothesis for systems where all of the workload elements are approximately equal. To be valid, any proposed workload characterization must pass the representative hypothesis test.

The stationary hypothesis evaluates the likelihood that the proposed workload characterization is valid for future workloads. To test this hypothesis, you must wait and test the characterization's applicability over a number of weeks or months. (Alternatively, you may produce the characterization on older historical data and compare it to more recent history that is not being used to do the characterization.)

You should never consider a workload characterization as a tool for forecasting unless this hypothesis has been tested. Using the hypothetical system discussed above, the smallest change in the ratio of test jobs to production jobs would significantly change the workload characterization. Moreover, even small changes in the average resource consumptions for testing or production jobs would have the same effect. Therefore, it is unlikely that any equivalent workload element approach would ever pass the stationary hypothesis test.

In general, very few proposed workload characterizations will ever pass the stationary hypothesis test. This does not mean that most workload characterizations are invalid; it does mean that most are unsuitable for use as forecasting units.

11.2.3 Application of Workload Characterization

Perhaps one of the most common workload characterization studies is the determination of job classes. Typically, the analyst guesses at a solution and then evaluates how well it fits the system's workload. This is known as the ad hoc approach. For example, consider a hypothetical system where the following job class structure has been proposed.

JOB CLASS	CPU MINS	PRINT LINES
=====	=====	=====
A	2	5,000
B	5	20,000
C	unlim	unlim

If 60% of the jobs are assigned to class A, 25% to class B, and 15% to class C, the analyst might assume that the classes provide a good representation of the workload.

Applying the representative hypothesis test could reveal problems. For example, if 30% of the jobs use less than 5 CPU seconds and print less than 2,000 lines, then they would be poorly represented by the class A limits. Moreover, their service characteristics could be much poorer when they are randomly mixed with the larger jobs in class A than they could be if they are assigned to a class by themselves.

Another problem that can result from arbitrarily established job classes is the specification of a job class limit that bisects a natural structure in the data. For example, if 15% of the jobs normally require between 110 and 130 CPU seconds, imposing a two-minute limit could force the users of jobs into class B to ensure that they are not canceled at the two-minute limit. This problem can also be identified by the representative hypothesis test since the approximate two CPU minute jobs would be poorly represented by the class B limits.

The characteristics of an installation's job mix are continually evolving. As a result, it is probably unreasonable to apply the stationary hypothesis test to a job class structure. However, you should examine the job class structure on a quarterly or semiannual basis.

Both the equivalent job and ad hoc batch approaches to workload characterization that we have discussed show the problems that result from using mean values of arbitrarily selected limits. To meet the requirements of the first hypothesis, we must exploit the natural structure of the workload. Statistical pattern recognition techniques (clustering) offer a powerful tool for determining whether a natural structure exists in an installation's workload data.

11.2.4 Data Selection and Outliers

Selecting data elements for the clustering procedure is one of the most important steps in the workload characterization process. In clustering terminology, these data elements are known as features. The result of the clustering process is a ranking of similar types of workloads, based on the mean values of these features.

The objective is to select the minimum number of features that can be used to accurately describe the characteristics of the workload to be analyzed. A small number of features is desirable, since it greatly reduces the number sources of variability in the characterization, thus making the results more robust and easier to understand. Note that each time an additional feature is added, splits may occur in as many as half of the existing clusters.

Another consideration in the feature selection process is the representation of outliers. Outliers are data elements which cannot be generalized as similar to other data elements in any of the chosen features. One example of an outlier is a batch job which falls into a CPU-bound loop due to a programming error. Another example is a periodic process, such as a monthly close of a financial application, which is run rarely but requires a large number of resources. Outliers have the highest probability of being poorly represented by the developed workload characterization.

There are two classes of features that are normally used in workload characterization studies:

- o Rate-based features, such as EXCPs per CPU second
- o Total features, such as CPU minutes

Rate-based features are desirable because they allow you to control the content of the outliers. For example, consider the use of CPU minutes and EXCPs as total features for characterizing the resource consumption of a workload. Because the largest jobs would be outliers, the resulting characterization would poorly represent them. Unfortunately, these jobs might represent a significant percentage of the resources that are consumed.

To avoid this problem, you could use rate-based features. When you divide EXCPs by CPU seconds, you change the content of the outliers. Rather than representing jobs that consume significant resources, they represent jobs with exceptional EXCP/CPU second rates. The probability that these jobs would consume significant resources is very small, since the ratio of elapsed time to CPU resource consumption for I/O-bound jobs is very high.

Such rate-based features have been called instantaneous workload descriptions, since they characterize the activities of a workload element for an instant of CPU time (ART76, ART78).

Although using total features introduces concerns about representing outliers, they have many applications to capacity planning and performance problems. For example, using total features for job and transaction class structures treats the occasional massive workload element as an exception. By handling the occasional element as an exception, the system can concentrate on exploiting the characteristics of the majority of the workload elements to simplify scheduling. You must, however, always be aware of the content of the outliers and pay particular attention to the marginally assigned resource vectors to verify that the objectives of the study are not compromised.

11.2.5 Suggested CA MICS Input Files

Although you may analyze any of the unsummarized CA MICS files in the DETAIL timespan, a number of files may particularly interest you. In the Batch Information Area, these are the following:

BATJOB	Contains one record for each job
BATPGM	Contains one record for each job step, that is, program
BAT_TS	Contains one record for each TSO session
BAT_ST	Contains one record for each started task

In the TSO Information Area, the following files may interest you:

TS0TSU	Contains one record for each CA TSO/MON PM interval in which a user was active
TS0TSI	Contains one record for each command or program executed interactively

In the CICS Information Area, you can create files in the DETAIL timespan that contain one record for each transaction using the _USRXfff macros. These macros are discussed in Chapter 10 of the CA MICS CICS Analyzer Guide. In the event that you use this option, the following files may interest you:

CICCSU	Contains one record for each transaction sorted by user
CICCAU	Contains one record for each transaction sorted by application

You can also create IMS transaction files in the DETAIL time-span that contain one record for each transaction using the `_USRXfff` macros. These macros are discussed in Chapter 10 of the CA MICS IMS Analyzer Guide. In the event that you choose this option, the following files may interest you:

IMISISU	Contains one record for each transaction sorted by user
IMSIAU	Contains one record for each transaction sorted by application

Although you can analyze other CA MICS information areas and files using Workload Characterization, no results have been published that detail the application of clustering algorithms to other than workload element data.

11.3 Standard Output

Workload Characterization produces reports that describe the original data, show the clusters which are identified from this data, and analyze the degree of fit with which these clusters characterize the data.

The four primary reports are the following:

- o Descriptive Statistics Report. This report details the characteristics of the variables that are specified as criteria for the clustering analysis. Scaling is an ancillary issue introduced by numerical clustering algorithms. Scaling considerations will be discussed under Analytic Technique Tutorial in Section 11.4 of this document.
- o Cluster Centroids Report. This report details the characteristics of the patterns that were identified in the workload. Each of these patterns statistically represents a subset of the installation's workload.
- o Cluster Assignment Report. This report provides the resource consumption characteristics of each of the clusters. For example, you can use this report to identify resource consumption patterns that may be resulting in performance or capacity problems. Like the Cluster Centroids Report, this report is based on a statistical sample of your installation's workload.

- o Population Cluster Assignment Report. This report provides the same information as the Cluster Assignment Report, except that the entire input data set, rather than the statistical subset, is included. This report allows you to judge how well the characterization fits the entire workload.

The sections which follow describe the fields, data elements, and format of each report.

- 1 - Descriptive Statistics Report
- 2 - Cluster Centroids Report
- 3 - Random Sample Cluster Assignment Report
- 4 - Population Cluster Assignment Report

11.3.1 Descriptive Statistics Report

The Descriptive Statistics Report provides a summary of the descriptive statistics that are calculated for the cluster features. This report is used to evaluate the effects of trimming 2.5% (or the specified value) off the right-hand tails of the distributions on the sample averages and standard deviations. Figure 11-1 illustrates a sample Descriptive Statistics Report.

DESCRIPTIVE STATISTICS								
NUMBER OF OBSERVATIONS: 2000								
	-----SAMPLE-----					-----TRIMMED 97.50%-----		
FEATURE	MIN	MAX	AVERAGE	STD DEV	CV	AVERAGE	STD DEV	CV
JOBMXNTA	0	6	0.2135	0.54687	2.6	0.154951	0.391913	2.5
JOBTCBTM	0.01	7115.71	16.6509	175.139	10.5	5.80658	11.0867	1.9
JOBNLRL	0	334275	2327.42	12277.9	5.3	1019.77	2339.31	2.3

Figure 11-1. Descriptive Statistics Report

The Descriptive Statistics Report lists the following fields:

OBSERVATIONS: The number of observations used for the scaling procedure. The default value is 2,000. The value shown in this report differs from 2,000 in only two cases. The first case results when the population you select from the CA MICS database contains less than 2,000 observations. The second case occurs when you override the default value on the Workload Characterization screen.

FEATURE: The feature name. The feature names that are listed in the report correspond to the features that are specified on the Workload Characterization screen.

MIN: The minimum value observed for the feature in the sample.

MAX: The maximum value observed for the feature in the sample.

For the randomly selected **SAMPLE**, the following values are calculated:

AVERAGE: The average calculated for the feature observations in the sample.

STD DEV: The standard deviation calculated for the feature observations in the sample.

CV: The coefficient of variation. The CV is calculated by dividing the standard deviation by the average.

For the trimmed **SAMPLE**, the percentage of the distribution used for calculating the trimmed mean statistics is shown. In Figure 11-1, 97.5% of the observations were included in the calculation of the trimmed mean statistics.

AVERAGE: The mean calculated for the feature observations in the trimmed sample.

STD DEV: The standard deviation calculated for the feature observations in the trimmed sample.

CV: The coefficient of variation. The CV is calculated by dividing the standard deviation by the average.

11.3.2 Cluster Centroids Report

The Cluster Centroids Report provides a summary of the patterns that are identified by the clustering procedure. This report allows you to determine the characteristics of the patterns that are identified. Figure 11-2 illustrates a sample Cluster Centroids Report.

CLUSTER CENTROIDS						
CLUSTER	JOBMXNTA	JOBTCBTM	JOBNLR	RADIUS	NO ASSIGN	PCT
1	0.0	0:00:05	3450.7	2.00	145	7
2	1.0	0:00:03	277.8	1.00	163	8
3	0.0	0:00:57	90.9	3.00	68	3
4	1.0	0:00:23	64.4	2.50	56	3
5	0.0	0:00:14	181.8	1.00	156	8
6	2.0	0:00:03	209.3	3.00	58	3
7	0.0	0:00:01	267.7	1.00	1172	59
8	0.0	0:00:07	8165.7	3.00	54	3
9	1.0	0:00:03	3950.7	2.50	20	1
10	0.0	0:00:11	12211.6	3.00	81	4
11	0.0	0:00:36	261.4	2.00	27	1

Figure 11-2. Cluster Assignment - Randomly Selected Sample Report

The Cluster Centroids Report lists the following fields:

CLUSTER: The cluster number. Cluster numbers are assigned sequentially to the patterns that are identified by the algorithm. Note that the order in which the clusters are identified is not an indicator of merit.

A column is included in the report for each of the features that are specified using the Workload Characterization screen (see Section 11.5). In the report shown in Figure 11-8, the columns that are generated are JOBMXNTA, JOBTCBTM, and JOBNLR.

JOBMXNTA:	The average number of tapes allocated by jobs assigned to the cluster. Note that this variable is included in the CA MICS database only if the DVCT option is specified to the CA MICS Batch and Operations Analyzer. In the distributed CA MICS system, this option is not activated.
JOBTCBTM:	The average CPU time consumed by jobs assigned to the cluster.
JOBNLR:	The average number of lines printed by jobs assigned to the cluster.
RADIUS:	The radius of the cluster in standard deviations. Cluster radii vary from a minimum of 1.0 to a maximum of 3.0 standard deviations.
NO ASSIGN:	The number of resource vectors that were assigned to the cluster. As a rule, the radius of a cluster is inversely proportional to the number of resource vectors assigned. More simply stated, clusters that represent the highest percentages of the resource vectors are most likely to be those with the smallest radii.
PCT:	The percent of the resource vectors in the sample that are represented by the cluster.

11.3.3 Random Sample Cluster Assignment Report

The Cluster Assignment Report for the randomly selected sample indicates how well the clusters represent the resource vectors in the sample. The report also provides an analysis of the resources that are consumed by the clusters. The resources detailed in this report are specified on the Workload Characterization screen (see Section 11.5).

A sample Cluster Assignment Report for the randomly selected sample is shown in Figure 11-3. It lists the following fields:

TOTAL OBSERVATIONS:	The total number of resource vectors summarized by the report.
OBS WITHIN RADIUS:	The number of resource vectors that are assigned within the radius of one of the clusters. These observations are well represented by the clusters.
MARGINAL ASSIGNMENTS:	The number and percent of resource vectors that are not within the radius of any of the clusters. The marginal assignments should represent no more than 5 to 7% of the sample.

The columns of the report are the following:

CLUS:	The cluster number.
WITHIN:	The number of resource vectors assigned to the cluster that are within the radius of the cluster.
MARGIN:	The number of resource vectors assigned to the cluster that are outside of the cluster's radius. A resource vector is marginally assigned to a cluster when it is farther away from the cluster's centroid than the radius of the cluster.

The last four report columns are the resource names that are specified on the Workload Characterization screen (see Section 11.5). The fields are as follows:

JOBTCBTM:	The total TCB CPU time consumed by the job assigned to the cluster.
JOBEDASD:	The total number of DASD EXCPs issued by jobs assigned to the cluster.
INITMIN:	The total number of initiator minutes consumed by jobs assigned to the cluster.

JOBNLNR: The total print lines generated by jobs assigned to the cluster.

The second line for each cluster provides the percent of the specified resource consumed by the cluster.

RESOURCE CONSUMPTION BY CLUSTER -- RANDOMLY SELECTED SAMPLE						
TOTAL OBSERVATIONS :		2000				
OBS WITHIN RADIUS :		1892				
MARGINAL ASSIGNMENTS:		108	5.4%			
CLUS	WITHIN	MARGIN	JOBTCBTM	JOBEDASD	INITMIN	JOBNLNR
1	145	0	0:11:19	188270	321.142	461175
	7.25%		2.04%	3.15%	2.12%	9.91%
2	163	0	0:08:51	184617	1329.24	46110
	8.15%		1.59%	3.09%	8.78%	0.99%
3	25	43	5:55:41	2646216	5432.45	157324
	3.40%		64.09%	44.26%	35.86%	3.38%
4	51	5	0:25:14	356752	1644.46	27583
	2.80%		4.55%	5.97%	10.86%	0.59%
5	156	0	0:35:38	212340	848.633	27388
	7.80%		6.42%	3.55%	5.60%	0.59%
6	52	6	0:07:58	182933	1081.27	43144
	2.90%		1.43%	3.06%	7.14%	0.93%
7	1172	0	0:22:04	390123	1383.77	298355
	58.60%		3.97%	6.52%	9.14%	6.41%
8	52	2	0:09:56	146459	208.562	417324
	2.70%		1.79%	2.45%	1.38%	8.97%
9	20	0	0:02:03	54634	176.678	110967
	1.00%		0.37%	0.91%	1.17%	2.38%
10	30	51	1:01:05	1221149	1433.75	3037632
	4.05%		11.01%	20.42%	9.47%	65.26%
11	26	1	0:15:13	395459	1287.72	27829
	1.35%		2.74%	6.61%	8.50%	0.60%
TOTALS			9:15:02	5978952	15147.7	4654831

Figure 11-3. Resource Consumption by Cluster - Randomly Selected Sample

11.3.4 Population Cluster Assignment Report

The Population Cluster Assignment Report indicates how well the clusters represent all of the resource vectors in the population. The report also provides an analysis of the resources that are consumed by the clusters. The resources that are detailed in this report are specified on the Workload Characterization screen (see Section 11.5).

Perhaps the best way to discuss the application of this report is to review the representative hypothesis that was introduced in Section 11.2.2. The representative hypothesis tests how close the resource vectors are to the characterization that was proposed for the workload. You can test this hypothesis by assigning the resource vectors for the full population to the clusters developed for the sample.

Test this hypothesis by comparing the marginal assignment percentage for the sample of resource vectors, shown previously in Figure 11-3, to the marginal assignment percentage for the full population shown in Figure 11-4. If these two values are not significantly different (+ or - 2%), then the conditions of the representative hypothesis have been met. If the test is not met, you should repeat the procedure with a larger sample size.

RESOURCE CONSUMPTION BY CLUSTER -- TOTAL POPULATION						
TOTAL OBSERVATIONS :		18633				
OBS WITHIN RADIUS :		17576				
MARGINAL ASSIGNMENTS:		1057	5.7%			
CLUS	WITHIN	MARGIN	JOBTCBTM	JOBEDASD	INITMIN	JOBNLR
1	1296	1	1:31:42	1667621	2830.38	4247705
	6.96%		1.99%	3.15%	1.99%	11.14%
2	1556	2	1:14:19	1578031	13008.4	396291
	8.36%		1.61%	2.98%	9.13%	1.04%
3	280	463	48:34:27	23590139	60822.1	1696906
	3.99%		63.26%	44.60%	42.68%	4.45%
4	432	80	4:06:36	5403274	15489.5	291784
	2.75%		5.35%	10.22%	10.87%	0.77%
5	1517	11	5:53:01	3952170	8446.99	323429
	8.20%		7.66%	7.47%	5.93%	0.85%
6	476	86	1:55:07	2155617	10764.1	620991
	3.02%		2.50%	4.08%	7.55%	1.63%
7	11064	0	3:29:46	3489342	11599.6	2825604
	59.38%		4.55%	6.60%	8.14%	7.41%
8	411	6	1:08:52	1178249	1611.2	3173915
	2.24%		1.49%	2.23%	1.13%	8.33%
9	131	13	0:15:07	319366	1888.09	795981
	0.77%		0.33%	0.60%	1.32%	2.09%
10	186	372	6:18:10	6998863	9600.94	23412139
	2.99%		8.21%	13.23%	6.74%	61.42%
11	227	23	2:19:40	2554131	6444.38	335516
	1.34%		3.03%	4.83%	4.52%	0.88%
TOTALS			76:46:46	52886803	142506	38120261

Figure 11-4. Cluster Assignment - Total Population Report

11.4 Analytic Technique Tutorial

This section discusses Workload Characterization methodology. Although we do not attempt to provide a complete discussion of the topic of workload characterization, we do provide an introduction to the statistical techniques employed by the CA MICS Capacity Planning. The tutorials presented in this section concentrate on the following topics:

- 1 - Clustering Algorithms
- 2 - Scaling
- 3 - Algorithm Implementation

11.4.1 Clustering Algorithms

Although there are a wide variety of statistical algorithms that you can use to develop clusters, one of the simplest and most popular algorithms is the K-means algorithm (MAC67). A variation of the K-means algorithm is employed in the SAS FASTCLUS procedure used by the CA MICS Capacity Planner (SAS82b). To understand how clustering algorithms work, you must understand some additional definitions:

Resource Vector:	An n-tuple of measurements selected to describe a workload element. For example, you might choose to describe jobs in terms of the three distinct measurements: CPU time, print lines, and tape mounts. Each job would then be represented by a 3-tuple (CPU time, print lines, and tape mounts).
Feature:	A component of a resource vector. In the above example, CPU time, print lines, and tape mounts are the three features of resource vectors describing the workload element, job.
Space:	The n-dimensional coordinate system in which resource vectors are defined.
Feature Mean:	The mean of all values of a given feature taken over a set of resources.

Feature Variance:	The variance (that is, the standard deviations squared) associated with the distribution of values of a given feature, taken over a set of resource vectors. Note that the feature with the greatest variance defines the longest coordinate axis of the space.
Centroid:	The geometric center of the set of all resource vectors contained in a space. The location of the centroid is defined by an n-feature vector. Its component values are the feature means of the resource vectors in the space.
Cluster:	A set of resource vectors in a given space plus their centroid. Note that the centroid of a set of vectors need not be one of the original vectors in the space.
Geometric Distance:	The distance between any two points (that is, resource vectors) in a given space. It is calculated by squaring the difference between the points (that is, the feature values) for each feature, summing the squares over all features, and taking the square root of the resulting sum. The formula for a two-dimension geometric distance is better known as the Pythagorean theorem. A formula for calculating geometric distance is presented in Equation 1 below.
Radius:	The greatest geometric distance between a centroid and any of its associated resource vectors.
Objective Radius:	An externally defined parameter which specifies a maximum desired radius for clusters identified by a clustering algorithm.
Splitting:	The partitioning of a space into two smaller spaces. A space is split in half along the axis of greatest variance. A clustering algorithm will continue to split resource vector spaces as long as the radius of any cluster in the partition is greater than the specified objective radius.

Consider the following two-dimensional clustering example. Fifteen resource vectors were selected for the example. They are shown in the following table:

VECTOR NUM.	X	Y
=====	=====	=====
1	1	1
2	2	1
3	3	1
4	3	8
5	4	7
6	4	8
7	4	9
8	5	8
9	8	4
10	9	3
11	9	4
12	9	5
13	10	4
14	10	8
15	11	8

Note: The range of values selected for the two features in this example are similar to avoid the problems that are introduced by scaling considerations. Scaling is discussed in Section 11.4.2. The objective radius for this example is 1.

A plot of the resource vectors is shown in Figure 11-5. The resource vectors are represented by asterisks (*) and the centroid of the space is represented by a pound sign (#).

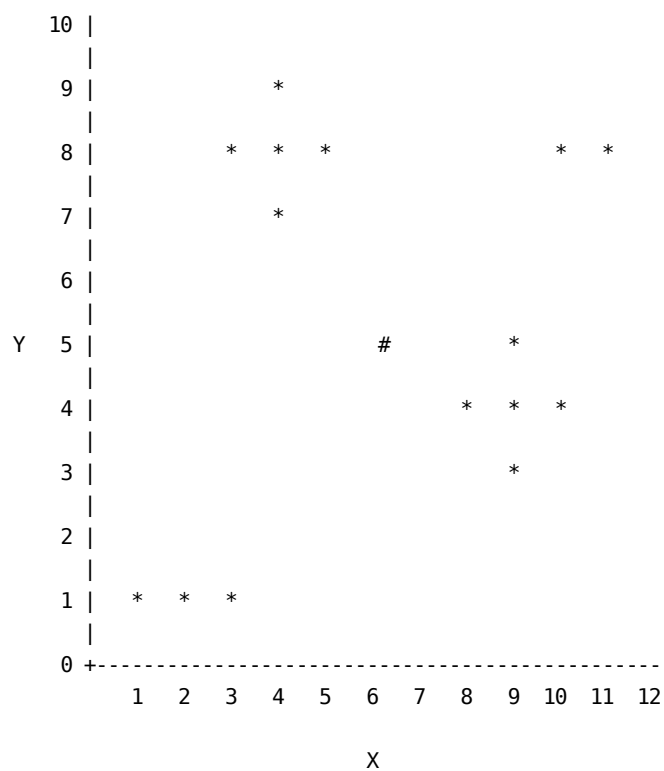


Figure 11-5. Initial Space

The first step in the solution process is to calculate the feature means, variances, and radius for the space. (The centroid is defined by the feature means.) The calculated values are as follows:

SPACE	FEATURE	CENTROID (MEAN)	VARIANCE	RADIUS
all	X	6.13	11.41	6.67
	Y	5.27	8.50	

As mentioned in the definitions above, the radius of the space is calculated using the geometric distance equation. This equation states that the distance, D , between the centroid of the k -th cluster, $C(k)$, and the i -th resource vector, $R(i)$, is equal to the square root of the sum of the squares of their feature differences. The general form of the geometric distance equation is shown in Equation 1, where j is the feature number index.

$$D = \sqrt{\sum_{j=1}^n (C(k,j) - R(i,j))^2}^{1/2} \quad (\text{Eqn 1})$$

For example, the farthest resource vector from the centroid of the space shown in Figure 11-5 is the resource vector at (1,1). The calculation of the distance between the resource vector and the centroid is shown in Equation 2.

$$D = \sqrt{(6.13-1.0)^2 + (5.27-1.0)^2}^{1/2} \quad (\text{Eqn 2})$$

$$D = \sqrt{26.32 + 18.23}^{1/2} = 6.67$$

As shown in the previous table, the variance in the X direction is 11.4 and the variance in the Y direction is 8.50. Since the radius of the space is greater than the objective radius, the space is split at the centroid along the axis of greatest variance. This results in two spaces, A and B, as is shown in Figure 11-6.

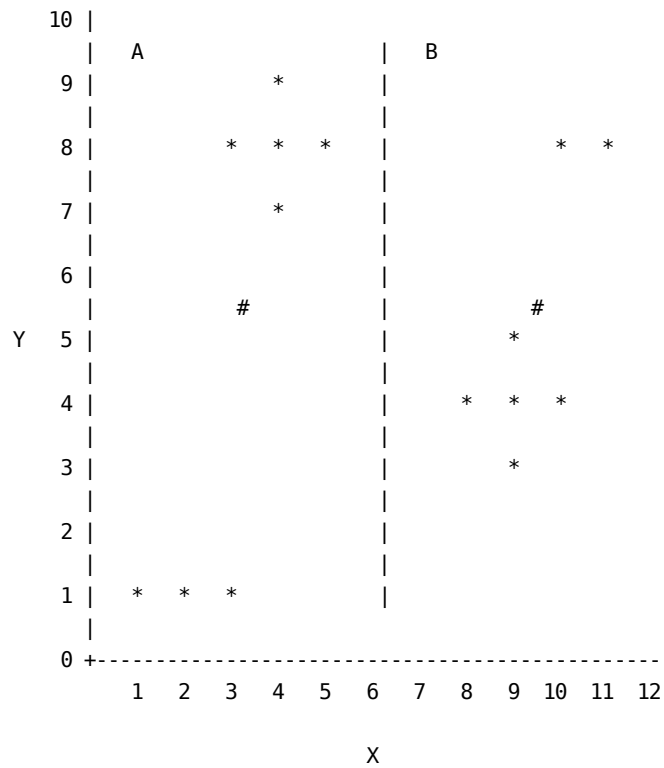


Figure 11-6. First Split

The characteristics of the two spaces resulting from the split are shown in the table below:

SPACE	FEATURE	CENTROID (MEAN)	VARIANCE	RADIUS
A	X	3.25	2.70	4.92
	Y	5.38	13.41	
B	X	9.33	1.07	3.31
	Y	5.14	4.14	

Once again, the radii of the spaces are greater than the objective radius. Hence, each of the spaces is split along the axis of greatest variance. This results in four spaces, A, B, C, and D, as shown in Figure 11-7. Note that in spaces A, C, and D the centroids happen to coincide with a resource vector.

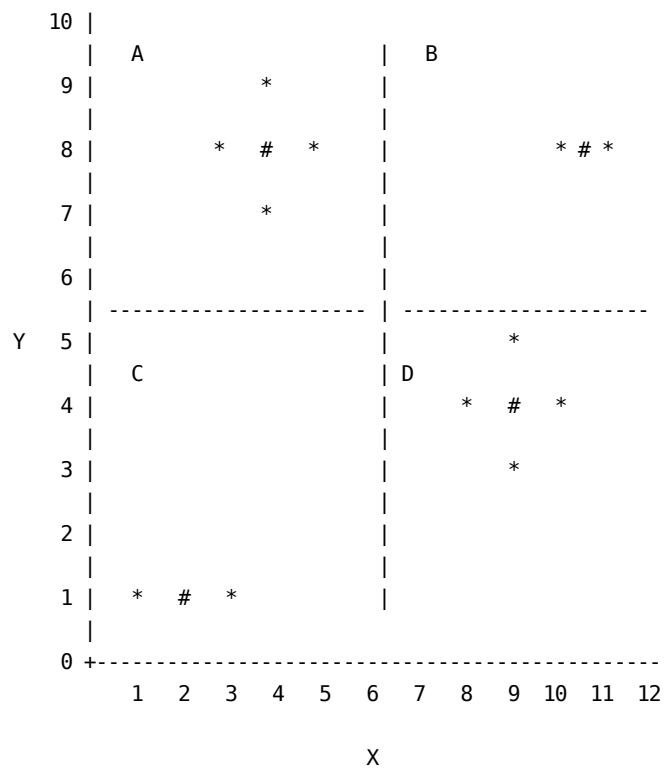


Figure 11-7. Second Split

The characteristics of the four spaces resulting from the split are shown in the table below:

SPACE	FEATURE	CENTROID (MEAN)	VARIANCE	RADIUS
A	X	4.0	0.50	1.0
	Y	8.0	0.50	
B	X	10.5	0.50	0.5
	Y	8.0	0.00	
C	X	2.0	1.00	1.0
	Y	1.0	0.00	
D	X	9.0	0.50	1.0
	Y	4.0	0.50	

Since all of the radii are less than or equal to the objective radii, a solution for the space introduced in Figure 11-7 has been developed. Although this problem is very simplistic, it illustrates the concepts of the K-mean algorithm.

Thus, four clusters have been identified that represent the 15 resource vectors in the space. Although the actual implementation of the clustering algorithm is somewhat more complex, the same principles apply when thousands of workload elements are analyzed with the program.

11.4.2 Scaling

One problem that is introduced by the use of statistical clustering algorithms is scaling. In the clustering example discussed in the previous section, we were careful to select a similar numerical range for the X and Y variables. Unfortunately, actual workload data presents a much wider range of values.

Consider the problem of developing job classes. The job classes in this example are based on CPU minutes and print lines. The maximum value observed for CPU minutes would be far less than 1,000 minutes, while the likelihood of printing more than 1,000 lines is very high. If the differences in these features are squared in the geometric distance equation, the number of CPU minutes used by a job would appear to be insignificant. Therefore, the variables must be scaled prior to the clustering process so that relative differences between the features have a similar influence on the geometric distance calculation.

A number of different scaling techniques can be employed. Two of the most popular are unit scaling and Z statistic scaling.

UNIT SCALING

Unit scaling is a simple technique for solving the problems introduced by the different range of values associated with each of the resource vector's features. Most simply stated, unit scaling maps the range of values associated with each feature into the range 0 to 1. Equation 3 details this mapping:

$$s_i = \frac{x_i - x_{\min}}{x_{\max} - x_{\min}} \quad (\text{Eqn 3})$$

where: s_i is the scaled value of x_i

x_i is i-th observation of the x feature

$x_{\min}$ is the minimum of all x_i

$x_{\max}$ is the maximum of all x_i

Although this transformation solves the problems introduced by the different ranges of values associated with the features, it does not provide any assistance in estimating a value for the objective radius.

Z STATISTIC SCALING

Z statistic scaling is another technique that has been employed in workload characterization studies.

Equation 4 details Z statistic scaling:

$$z_i = \frac{x_i - \bar{x}}{std} \quad (\text{Eqn 4})$$

where: z_i is the Z statistic for x_i .

x_i is i-th observation of the x feature.

$\bar{x}$ is the mean value of the x distribution.

std is the standard deviation of the x distribution.

The Z statistic transforms the range of values to the range -3.5 to +3.5. In other words, the Z statistic transforms each x observation into the number of standard deviations that it lies from the mean value. Since a standard deviation can be considered as a measure of a "significant difference" with a distribution, you can establish the value for the objective radius in terms of standard deviations. Since the Z statistic solves both the problem of scaling the features and selecting a meaningful value for the objective radius, it was selected for use in the workload characterization approach implemented in the CA MICS Capacity Planner.

One other complication related to scaling is the presence of outliers in the feature distributions. With workload measurement data, it is not uncommon to find that one or two observations out of every 10,000 are an order of magnitude different than any of the other observations in the distribution. Although the presence of these few observations changes the mean value by only a few percentage points, they have a profound effect on the standard deviation. If such an exaggerated standard deviation value is used in the scaling process, then the body of the data would be crunched so that the outliers could be properly scaled.

To correct the problem introduced by outliers, it is necessary to exclude the outliers from the calculation of the population statistics, that is, the mean and standard deviation. Previous investigators have trimmed from 1 to 5 percent of the observations from the right-hand tail of the distribution. By default, the workload characterization program trims 2.5 percent of the observations from the right-hand side of the distribution prior to the calculation of the population statistics. It is important to note that these observations are not excluded from the workload characterization study, but are simply excluded from the calculation of the population statistics.

When the feature observations are scaled using the trimmed mean and standard deviation values, 97.5 percent of the values are scaled between -3.5 and +3.5. The values that are excluded from the calculation of the population statistics are scaled to values greater than +3.5. It is very likely that if any of the resource vectors are poorly represented by the characterization of the workload developed, it will be one of the resource vectors that contains a feature that is excluded from the calculation of the population statistics.

The potentially poor representation of outliers is one of the primary concerns in the selection of features for the clustering process. Criteria for selecting features for clustering are discussed in Section 11.2.4.

11.4.3 Algorithm Implementation

The workload characterization program is comprised of these major steps:

- 1 - Selecting data elements (features) from the CA MICS file based on user input.
- 2 - Randomly selecting 2,000 resource vectors from the population selected from the CA MICS database.
- 3 - Calculating the trimmed mean and standard deviation values for the 2,000 resource vector sample.
- 4 - Scaling the resource vector sample using the trimmed population statistics.
- 5 - Clustering the sample of scaled resource vectors.
- 6 - Scaling the population of resource vectors using the trimmed population statistics.
- 7 - Assigning the full population to the clusters developed for the sample of resource vectors. This final step evaluates how well the population is represented by the clusters developed for the randomly selected sample. If the population fits the proposed characterization, then the requirements of the first hypothesis have been satisfied.

The first step in the program is the selection of the user-specified elements from a CA MICS database file. The CA MICS file and database elements (features) are specified with parameters on the Workload Characterization screen. The control parameters are discussed in Section 11.5.

The second step is the random selection of 2,000 resource vectors from the population selected from the CA MICS database. Experience has shown that a sample size of 2,000 is suitable for analyzing populations as large as 500,000 workload elements. The selection process involves sorting the resource vectors into order by arrival time. The samples are then selected from the time-ordered population using a uniform sampling rate. You may specify alternative sample sizes if the default sample size proves unsuitable.

The third step in the process is the calculation of the trimmed mean and standard deviations for each of the features. This involves the sorting of the observations for each feature into order so that the tail of the distribution can be excluded. Although 2.5% is the default value for trimming the observations, you can override this value.

In the fourth step, the sample resource vectors are scaled using the trimmed mean statistics.

In the fifth step, the SAS FASTCLUS procedure is used to identify clusters in the sample resource vectors. One problem introduced by FASTCLUS is the way it treats outliers. When FASTCLUS encounters an outlier, it assigns the point to a cluster by itself to meet the objective radius criteria. Other clustering algorithms, like ISODATA (HAL73), selectively expand the radii of some clusters to avoid forming clusters that represent less than a specified percentage of the sample.

To implement a similar feature in the workload characterization program, the results of the FASTCLUS procedure are post processed to remove clusters that do not represent a minimum percent (0.5%) of the sample. Ball and Hall introduced the concept of a minimum cluster size in the ISODATA points algorithm in 1965 (BAL65). Within limits, selected cluster radii are expanded to include the majority of the points that were represented by the deleted clusters. The remaining resource vectors that do not fit within the radii of any of the expanded clusters are "marginally assigned" to the nearest cluster centroid.

Marginal assignment percentages of one or two are common. The marginal assignment percentage, the percent of the sample observations represented by each cluster, and the percent of the resources represented by each cluster are also calculated and reported.

In the sixth step, all of the resource vectors in the population are scaled using the trimmed mean statistics.

In the final step, the marginal assignment percentage, the percentage of the observations represented by each cluster, and the percent of the resources represented by each of the clusters are calculated and reported for the population. If the marginal assignment rate and the percent of the resource vectors represented by each cluster agree with the values calculated for the sample, then the first hypothesis has been successfully tested. You can request that the data set that is created by the population assignment DATA step be copied to a user-defined SAS file for later analysis.

11.5 Component Operation

The Component Operation Section describes the screens and parameters you can use to control workload characterization. The control parameters allow you to define the CA MICS file you wish to use as input, the time periods you wish to specify, and the variables you wish to use in the clustering process. Each of the control screens is discussed the following sections.

- 1 - Workload Characterization Screen
- 2 - Extended Options Menu
- 3 - Global Data Selection Screen
- 4 - Data Element Derivation Screen
- 5 - Alternate Input File Screen
- 6 - Save Selected Input Observations Screen
- 7 - Save Cluster Output Observations Screen
- 8 - Statistical Analysis Parameters Screen
- 9 - Report Description Elements Screen
- 10 - Inquiry Processing Screen

11.5.1 Workload Characterization Screen

Figure 11-8 illustrates the Workload Characterization screen.

```

/----- Workload Characterization -----\
| Command ==>
|
| Composing CA MICS Inquiry: CLUCIO - Workload Characterization
|
| CA MICS file ==> -
| Time-span:    DETAIL                      Cycle(s) ==> __ - __
| Data base ==> _
| Allow Run-Time Override of Data Base ==> Y (Y/N) and/or Cycle(s) ==> Y (Y/N)
|
| Analysis based on      ==> _____
| Report for each cluster ==> _____
|
| Data Selection:                      Allow run-time override
| Date      ==> _____ ==> N (Y/N)
| Hour      ==> _____ ==> N (Y/N)
| Zone      ==> _____ ==> N (Y/N)
| SYSID     ==> _____ ==> N (Y/N)
|
| Specify extended options ==> N (Y/N)
\-----/

```

Figure 11-8. Workload Characterization Screen

The Workload Characterization screen lists the following fields:

CA MICS file: The CA MICS file that contains the data elements that you wish to include in your workload characterization analysis. If you are uncertain of the correct file name, type a question mark (?) in this field to obtain a complete selection list for your installation.

Cycle(s): The cycle or cycles of the CA MICS input file that are used in the analysis. For most installations, there is one cycle of DETAIL CA MICS data for each day. You may specify a maximum of 21 cycles.

Database:	The database unit identifier(s) (up to five units) that contain the specified CA MICS file. An eight-character database unit name is displayed after you specify this identifier. If you are uncertain of which database unit the CA MICS file resides in, type a question mark (?) in this field to obtain a complete selection list for your installation.
Analysis based on:	The features (data elements) that are analyzed by the clustering procedure. You may define a maximum of ten features. Cluster variables must be numeric. These data elements may either be CA MICS data elements or non-CA MICS data elements. If you specify non-CA MICS data elements, you must provide the SAS code to derive them via the exit available on the Extended Options Menu, Figure 11-9.
Report for each cluster:	The data elements that will be detailed in the Resource Consumption by Cluster Reports (see Section 11.3.3). At a minimum, you will generally want to specify all of the data elements specified for the Analysis based on parameter. In addition, you may specify other numeric data elements that you might find informative in the reports. For example, clusters for a job class study might be developed using CPU time, tape mounts, and print lines as features. You might also wish to report on tape EXCPs, tape drive occupied hours, disk EXCPs, and average elapsed time of the jobs assigned to the clusters. A maximum of ten report variable names may be specified. Report variables must be numeric. These data elements may be either CA MICS data elements or non-CA MICS data elements. If non-CA MICS data elements are specified, you must provide the SAS code to derive them via the exit available on the Extended Options Menu, shown in Figure 11-9.
Date selection:	The selection criteria for Date, Hour, Zone, and SYSID to subset the data you are examining. You must enter date values in the ddmonyy format.

Specify extended options:	The specification for selecting extended options. Select this option if you wish to specify additional parameters.

11.5.2 Extended Options Menu

Typing a Y at the Specify extended options prompt on the Workload Characterization screen displays the Extended Options Menu shown in Figure 11-9.

```
/----- Extended Options Menu -----  
|Command ==>  
|  
|Inquiry Step: Workload Characterization  
|  
|Specify global data selection exit          ==> N (Y/N/R)  
|Specify non-CA MICS data element derivation exit==> N (Y/N/R)  
|  
|Specify alternate input file                ==> N (Y/N/R)  
|Save selected input observations            ==> N (Y/N/R)  
|Save cluster output observations           ==> N (Y/N/R)  
|  
|Specify statistical analysis parameters     ==> N (Y/N/R)  
|Define report description element names    ==> N (Y/N/R)  
|  
|  
|  
|  
|  
|  
|
```

Figure 11-9. Extended Options Menu

You may select any of the options available on this screen by typing a Y (YES) at any of the prompts. Typing an R (RESET) resets the option and nullifies any specifications you made for that option.

11.5.3 Global Data Selection Screen

Figure 11-10 illustrates the Global Data Selection Macro Definition screen.

```

/----- SAS Statements -----\
|Command ==>                               Scroll ==> CSR |
|
|Inquiry Step:  Workload Characterization
|
|Line Cmds: I Insert  D Delete  R Repeat  M Move  C Copy
|
|Cmd  SAS Statements for:  Specifying global data selection
| - -----
| _ - %MACRO SELEOF;
|
| - -----
| - -----
| - -----
| - -----
| - -----
| - -----
| - -----
| - -----
| - -----
| - -----
| - -----
| _ - %MEND SELEOF;
| ***** BOTTOM OF DATA *****
|
\-----/

```

Figure 11-10. Global Data Selection Macro Definition Screen

Workload characterization provides a user-exit that you can use to control data selection from the CA MICS database. To invoke this exit, code SAS statements in the SELEOF macro. This macro is bracketed by the %MACRO and %MEND statements which are provided for you and are field protected to prevent accidental modification.

The default for the SELEOF macro is null. This macro is invoked when the workload observations are read from the specified CA MICS file. You may use it to delete observations from the data as it is selected. For example, in a job class study, you might wish to exclude production job classes from the analysis. If production jobs are designated by job class P, then you would code the following:

```
%MACRO SELEOF;  
  IF JOBCLASS='P' THEN GO TO T_EOF;  
%MEND SELEOF;
```

Note that you do not have to code the %MACRO and %MEND statements, as they are already provided for you.

The IF statement should specify branching to T_EOF rather than acting as a subsetting IF or specifying DELETE. A subsetting IF or a DELETE statement may cause the program to function incorrectly.

You must code each specified SAS statement in accordance with the syntax of the SAS Macro Language, as described in the SAS manuals.

11.5.4 Data Element Derivation Screen

If you require user-defined data elements, use the Data Element Derivation Macro Definition screen shown in Figure 11-11 to specify their derivation.

```

/----- SAS Statements -----\
|Command ===>                               Scroll ===> CSR|
|
|Inquiry Step: Workload Characterization
|
|Line Cmds: I Insert  D Delete  R Repeat  M Move  C Copy
|
|Cmd  SAS Statements for:  Specifying data element derivation
|  -  -----
|  -  %MACRO  DRVMAC;
|  -  -----
|  -  -----
|  -  -----
|  -  -----
|  -  -----
|  -  -----
|  -  -----
|  -  -----
|  -  -----
|  -  -----
|  -  -----
|  -  -----
|  -  -----
|  -  -----
|  -  %MEND  DRVMAC;
|
|***** BOTTOM OF DATA *****|
\-----\

```

Figure 11-11. Data Element Derivation Macro Definition Screen

Workload characterization provides a user-exit for defining the procedures for computing data elements that are not part of the CA MICS database. This exit is invoked by coding SAS statements in the DRVMAC macro. This macro is bracketed by the %MACRO and %MEND statements which are provided for you and are field protected from modification. The default for the DRVMAC macro is null.

You must code each data element derivation SAS statement in accordance with the syntax of the SAS Macro Language, as described into the SAS manuals.

11.5.5 Alternate Input File Screen

Figure 11-12 illustrates the Alternate Input File screen.

```
/----- Alternate Input File -----\
|Command ==>
|
|Inquiry Step: Workload Characterization
|
|SAS file name(s) ==> _____
|
|
|SAS Data set allocation instructions:
|   Data set name   ==> _____
|   Disposition     ==> SHR  (SHR/OLD)
|
|
|
|
|
```

Figure 11-12. Alternate Input File Screen

Workload characterization has the ability to read data either from CA MICS files or from non-CA MICS external SAS files. The Alternate Input File screen allows you to specify that the input data for the workload characterization is to be a non- CA MICS SAS file or files, and allows you to specify the SAS data set name and up to five SAS files that are to be used as input to the analysis. Note that the SAS data must already be created.

11.5.6 Save Selected Input Observations Screen

Figure 11-13 illustrates the Save Selected Input Observations screen.

[illegible]

Figure 11-13. Save Selected Input Observations Screen

Workload characterization has the ability to write the observations that your workload characterization analysis inquiry selected to an external SAS file for further processing by independent SAS programs or for future processing by the workload characterization software. The Save Selected Input Observations screen allows you to specify that the selected observations are to be written to an external SAS file, and to specify the SAS data set name and the name of the SAS file to which they are to be written. Note that the SAS data set must already be created.

11.5.7 Save Cluster Output Observations Screen

Figure 11-14 illustrates the Save Cluster Output Observations screen.

[illegible]

Figure 11-14. Save Cluster Output Observations Screen

Workload Characterization has the ability to write the SAS cluster observations from the SAS clustering program to external files for additional processing. The Save Cluster Output Observations screen allows you to specify the SAS data set and the name of the SAS file to which the cluster observations are to be written. Note that the SAS data set must already be created.

11.5.8 Statistical Analysis Parameters Screen

Figure 11-15 illustrates the Statistical Analysis Parameters screen.

```

/----- Statistical Analysis Parameters ----- ROW 1 OF 5 \
|Command ==>                                     Scroll ==> CSR |
|
|Inquiry Step:  Workload Characterization
|
|Sample size           ==> 2000   (nnnnnn)
|Observation trimming (percent) ==> 97.5   (nn.n)
|Objective radius      ==> 1.5    (1.0-2.0)
|
|Delete observations if the following elements contain SAS missing values or 0:
|
|Cmd Element      Data element label (long name)          Warning
| - - - - -      - - - - -
| - - - - -
| - - - - -
| - - - - -
| - - - - -
| - - - - -
|***** BOTTOM OF DATA *****
|
|
|
\-----/

```

Figure 11-15. Statistical Analysis Parameters Screen

The Statistical Analysis Parameters screen allows you to specify the statistical parameters that will be used in the workload characterization analysis. The available parameters are described below.

Sample size: The sample size to be used for the clustering procedure. The default sample size, which is 2,000, should prove suitable for populations as large as 500,000.

Observation trimming (percent): The percent of the distribution to be included when the trimmed mean and standard deviation are calculated. The default value is 97.5%.

Objective radius:	The objective radius for the clusters to be created. The default value is 1.5 standard deviations.

This screen also allows you to specify deletion for any input observation that has a zero or missing value for a specified data element name when it is read from the CA MICS database. For example, in a job class study you might wish to delete the observations for any jobs whose TCB CPU time is zero or missing. To do so, you would code JOBTCBTM in the Element field of the display table at the bottom of the screen. You can specify only numeric variables. To obtain additional lines, enter an I in the Cmd column of any of the rows in the display portion of the screen.

11.5.9 Report Description Elements Screen

Figure 11-16 illustrates the Report Description Elements screen.

[illegible]

Figure 11-16. Report Description Elements Screen

The Report Description Elements screen allows you to specify additional descriptive data elements to be included in the optional SAS file that can be created by workload characterization (see Section 11.5.7). Unlike the cluster, zero deletion, and report variables, the descriptive variables need not be numeric. For example, in a job class study you might wish to include additional data elements like JOBCCLASS (job class), JOB (job name), and RDRTS (reader time stamp) for use in your SAS programs that process the optional SAS File.

11.5.10 Inquiry Processing Screen

Figure 11-17 illustrates the Inquiry Processing screen.

```
/----- Inquiry Processing -----\  
|Command ==>  
|  
|   Active MICF Inquiry:  CLUCI  - Workload Characterization  
|           Author:      Charles Frost  
|  
|   Specify Inquiry processing ==> EXECUTE (EXECUTE/MODIFY/SAVE/CANCEL)  
|   EXECUTE - Execute the inquiry.  
|   MODIFY  - Redisplay the inquiry for modification.  
|   SAVE    - Exit from inquiry composition and save the inquiry.  
|   CANCEL  - Exit from inquiry composition without saving the inquiry.  
|  
|Instructions:  
|   Specify desired option and press ENTER key.  
|  
|-----/
```

Figure 11-17. Inquiry Processing Screen

Typing END at the Command prompt or pressing the END PF key (usually PF3) while on the Workload Characterization screen causes the Inquiry Processing screen to be displayed. Your available options are described in the text of the screen. Note that you may save a partially composed inquiry for later modification (completion) and execution.

11.6 Case Studies

The Case Studies Section presents case studies to illustrate how you can apply the clustering methodology to solve capacity planning and performance management problems that many data processing installations commonly encounter.

The batch job classes case study shows how you can use clustering to investigate patterns in the batch jobs being processed. You can use information to develop workload classes which can then form a reasonable basis for service objectives. The statistical uniformity of the classes enables you to assess the degree to which the service objective can be met. You can use the same information to define a job class structure for establishing an initiator scheme that is consistent with the service objectives. A full set of reports and an explanation for each report is provided with this study.

The TSO response study shows how you can apply clustering to a TSO workload to establish a characterization, which you can use to develop a benchmark system. The basic objectives of the study are presented through Descriptive Statistics, Cluster Centroids, and Cluster Assignment Reports.

The IMS transaction study looks at an IMS workload to determine a transaction class strategy. The study includes a description of the input data and the centroid report.

The case studies are shown in the following sections:

- 1 - Batch Job Class
- 2 - TSO Response
- 3 - IMS Transaction

11.6.1 Batch Job Class

The Batch Job Class case study applies the workload characterization program to identify job classes for an installation. The following topics are addressed by section:

- 1 - Problem Description
- 2 - Control Parameters
- 3 - Report Interpretation

11.6.1.1 Problem Description

This problem asks you to determine a new set of resource-limited job classes for an installation. Job class structures are generally established to meet two objectives (FER78, ART79a):

- o Provide operational controls
- o Facilitate scheduling and implement service policies

You usually establish job classes that you define to meet the first objective for reasons other than resource requirements. For example, online systems like IMS and CICS are often run in unique job classes to avoid interference with other workloads. Also, many computer centers use one or two job classes to segregate production jobs from the remainder of their workload.

Job classes that you establish to meet the second objective provide your users with alternative classes to which they can assign their jobs based on resource requirements. Typically, the service objective of a job class decreases (that is, jobs have longer turnaround times) as its resource requirements increase. Using a job class structure that focuses on scheduling and service has a number of potential benefits:

- o Many jobs can be completed in a short time (assuming demand) since the job classes representing smaller jobs may be expedited.
- o The availability of better turnaround time for the smaller classes may encourage users to conserve resources.
- o Job scheduling is simplified since each job class represents a homogeneous set of jobs.

Clustering is an ideal approach to this problem since we want to exploit the natural characteristics of the workload. If we arbitrarily select job classes, we might select classes that split normal resource consumption patterns or choose some that are too large or small to be useful.

For this sample installation, we investigated job classes based on CPU time, print lines, and tape drives. We selected these resources to quantify the system, output, and operator requirements presented by the jobs. Although a number of other features could have been selected, only three were chosen to keep the resulting job class structure as simple as possible.

11.6.1.2 Control Parameters

Figure 11-18 illustrates the Workload Characterization screen.

```

/----- Workload Characterization -----\
|Command ==>|
|
|Composing CA MICS Inquiry: CLASS - Batch Job Class Case Study|
|
|CA MICS file ==> JOB -|
|Time-span:      DETAIL|
|Data base ==> A|
|Allow Run-Time Override of Data Base ==> Y (Y/N) and/or Cycle(s) ==> Y (Y/N)|
|
|Analysis based on ==> JOBMXNTA JOBTCBTM JOBNLR__|
|
|Report for each cluster ==> JOBTCBTM JOBEDASD INITMIN JOBNLR|
|
|Data Selection:|
|Date ==> - - - - -|
|Hour ==> - - - - -|
|Zone ==> - - - - -|
|SYSID ==> - - - - -|
|
|Specify extended options ==> Y (Y/N)|
\-----/

```

Figure 11-18. Workload Characterization Screen Control Parameters

The Workload Characterization screen lists the following fields:

CA MICS file: The CA MICS file to be used for workload characterization. (For this example, we specified the CA MICS BATJOB File for a job class study.)

Cycles: The number of cycles to be specified. (For example, we specified seven cycles to analyze a full week of data. In an actual study, analyze as much data as possible before you initiate a new job class structure.)

Database: The single-character database identifier for this file at your site.

Analysis based on: The CA MICS data element feature names (that is, tape drives, CPU time, and print lines) to be used in the clustering procedure. (As discussed in Section 11.2.4, you should specify as few features as possible.)

CPU time, DASD EXCPs, initiator-occupied minutes (that is, elapsed time), and print lines are specified for reporting purposes. (Because initiator-occupied minutes is not a standard element in the CA MICS database, it is calculated using the user-defined macro DRVMAC that is specified on another screen).

You may select additional option specifications on the Extended Options Menu. Figure 11-19 shows an example in which we specified four of the extended options.

```

/----- Report Description Elements ----- ROW 1 OF 2 \
|Command ==>                               Scroll ==> CSR |
|
|Inquiry Step: Workload Characterization
|
|Line cmds: I Insert D Delete M Move
|
|Cmd Element      Data element label (long name)          Warning
| - - - - -      - - - - -
| _ JOB
| _ JOBCCLASS
|***** BOTTOM OF DATA *****
|
|
|
|
|
|
|
|
|
|
|
\-----

```

Figure 11-20. Report Description Elements Screen

We excluded jobs that used no CPU time from the study, since they did not execute due to a JCL error, or some other abnormal condition. To exclude jobs with no CPU usage, specify the data element name, `JOBTCBTM`, on the Statistical Analysis Parameters screen shown in Figure 11-22.

[illegible]

Figure 11-23. External Report Output File Definition

In this case study we wanted to create the SAS data set name, ALLOBS, that contains the cluster analysis output. To do this, we specified ALLOBS on the External Report Output File Definition screen shown in Figure 11-23.

```

/----- Inquiry Processing -----
|Command ===>
|
|
|
|Active MICF Inquiry:  CLASS  - Batch Job Class Case Study
|                        Author:      Charles Frost
|
|
|Specify Inquiry processing ===> EXECUTE (EXECUTE/MODIFY/SAVE/CANCEL)
|EXECUTE - Execute the inquiry.
|MODIFY  - Redisplay the inquiry for modification.
|SAVE    - Exit from inquiry composition and save the inquiry.
|CANCEL  - Exit from inquiry composition without saving the inquiry.
|
|
|Instructions:
|
|Specify desired option and press ENTER key.
|
\-----

```

Figure 11-24. Inquiry Processing

After a workload characterization analysis inquiry is created, the Inquiry Processing screen shown in Figure 11-24 is displayed. From this screen, you can execute the newly created inquiry, modify it, save it without executing it, or cancel the parameter specifications you just entered.

11.6.1.3 Report Interpretation

The example we discussed in Section 11.6.1.2 showed you how to use the control screens to execute Workload Characterization. The program produces five reports, which are discussed in the following sections:

- 1 - Descriptive Statistics Report
- 2 - Cluster Centroids Report
- 3 - Random Sample Cluster Assignment Report
- 4 - Population Cluster Assignment Report
- 5 - Analysis of the ALL_OBS SAS Data Set

11.6.1.3.1 Descriptive Statistics Report

The Descriptive Statistics Report provides a summary of the descriptive statistics calculated for the cluster features. This report is used to evaluate the effects of trimming 2.5% (or the specified value) off the right-hand tails of the distributions on the sample averages and standard deviations. A sample report is shown in Figure 11-25.

In Section 11.4.3 we discussed trimming a small percentage of the right-hand tail from the distribution to minimize the estimate of the standard deviation that was made for the population. For example, if the average value for CPU time in the full sample is 16.6 with a standard deviation of 175.1 seconds, then the coefficient of variation (that is, the standard deviation divided by the mean) is 10.55. As a general rule, a high coefficient of variation (that is, greater than 2.5) indicates that the outliers are still present in the trimmed distribution.

If, for example, 2.5% of the observations are excluded from the calculation of the sample statistics, the average and the standard deviation would be 5.8 and 11.1 seconds, respectively. These values result in a coefficient of variation of 1.9. In this batch job class study, the statistical behavior of the observations would be significantly improved by excluding 2.5% of the observations.

If the coefficient of variation for one of the features exceeds the recommended range, you should decrease the percent of the distribution to be included in the trimmed statistics by adjusting Observation trimming (percent) on the Statistical Analysis Parameters screen shown in Figure 11-22.

DESCRIPTIVE STATISTICS								
NUMBER OF OBSERVATIONS: 2000								
-----SAMPLE-----						-----TRIMMED 97.50%-----		
FEATURE	MIN	MAX	AVERAGE	STD DEV	CV	AVERAGE	STD DEV	CV
JOBMXNTA	0	6	0.2135	0.54687	2.6	0.154951	0.391913	2.5
JOBTCBTM	0.01	7115.71	16.6509	175.139	10.5	5.80658	11.0867	1.9
JOBNLR	0	334275	2327.42	12277.9	5.3	1019.77	2339.31	2.3

Figure 11-25. Descriptive Statistics Report

11.6.1.3.2 Cluster Centroids Report

The Cluster Centroids Report provides a summary of the patterns that are identified by the clustering procedure. A sample Cluster Centroids Report is shown in Figure 11-26. This report allows you to determine the characteristics of the patterns that are identified.

CLUSTER CENTROIDS						
CLUSTER	JOBMXNTA	JOBTCBTM	JOBNLR	RADIUS	NO ASSIGN	PCT
1	0.0	0:00:05	3450.7	2.00	145	7
2	1.0	0:00:03	277.8	1.00	163	8
3	0.0	0:00:57	90.9	3.00	68	3
4	1.0	0:00:23	64.4	2.50	56	3
5	0.0	0:00:14	181.8	1.00	156	8
6	2.0	0:00:03	209.3	3.00	58	3
7	0.0	0:00:01	267.7	1.00	1172	59
8	0.0	0:00:07	8165.7	3.00	54	3
9	1.0	0:00:03	3950.7	2.50	20	1
10	0.0	0:00:11	12211.6	3.00	81	4
11	0.0	0:00:36	261.4	2.00	27	1

Figure 11-26. Cluster Centroids Report

On the sample, consider the values shown for CLUSTER in the report. Cluster 7 represents approximately 59% of all jobs that executed during the seven-day period. The cluster centroid represents a job that uses no tape drives, uses one second of CPU time, and prints approximately 270 lines. The radius of the cluster is 1.0 standard deviations. Using the trimmed standard deviation values from the example in Figure 11-15, we can determine the bounds of the jobs assigned to the cluster. Consider the following table for cluster 7:

FEATURE	CENTROID	STD DEV	LOWER BOUND	UPPER BOUND
=====	=====	=====	=====	=====
JOBTCBTM	0:01	0:11	-0:10	+0:12
JOBMXNTA	0	0.39	-0.39	+0.39
JOBNLR	267	2339	-2072	+2606

By multiplying the positive and negative values of the cluster's radius times by the standard deviation for each of the features, you can determine the possible bounds of the cluster. Note that you may allocate tapes only in integer values and that jobs cannot consume negative resource. Also, note that the minimum value of CPU time observed in the distribution was .01 second. This simplification yields the following table:

FEATURE	CENTROID	STD DEV	LOWER BOUND	UPPER BOUND
=====	=====	=====	=====	=====
JOBTCBTM	0:01	0:11	0:01	+0:12
JOBMXNTA	0	0.39	0	0
JOBNLR	267	2339	0	+2606

Although the second table still indicates that the resulting cluster has some positive variability, the closeness of the cluster's mean values to the minimum values (from the Descriptive Statistics Report, Figure 11-27) for the distribution indicates that the number of extreme values is small.

11.6.1.3.3 Random Sample Cluster Assignment Report

The Cluster Assignment Report for the randomly selected sample indicates how well the clusters represent the resource vectors in the sample. The report also provides an analysis of the resources that are consumed by the clusters. A sample Cluster Assignment Report for the randomly selected sample is shown in Figure 11-27. The resources detailed in this report are specified on the Workload Characterization screen (see Figure 11-18).

This report demonstrates the potential benefits of job classes that you developed using the clusters. For example, consider cluster 7. This cluster represents approximately 59% of all the jobs processed by the system. These jobs use less than 10% of the resources consumed by all jobs that execute on the system. By exploiting the characteristics of cluster 7, you could establish a job class and an initiator structure that would isolate jobs in this cluster. By prioritizing the initiators which process these jobs, you can guarantee excellent service to more than half of the jobs while minimally impacting resources on the system as a whole.

RESOURCE CONSUMPTION BY CLUSTER -- RANDOMLY SELECTED SAMPLE						
TOTAL OBSERVATIONS : 2000						
OBS WITHIN RADIUS : 1892						
MARGINAL ASSIGNMENTS: 108 5.4%						
CLUS	WITHIN	MARGIN	JOBTCBTM	JOBEDASD	INITMIN	JOBNLR
1	145	0	0:11:19	188270	321.142	461175
	7.25%		2.04%	3.15%	2.12%	9.91%
2	163	0	0:08:51	184617	1329.24	46110
	8.15%		1.59%	3.09%	8.78%	0.99%
3	25	43	5:55:41	2646216	5432.45	157324
	3.40%		64.09%	44.26%	35.86%	3.38%
4	51	5	0:25:14	356752	1644.46	27583
	2.80%		4.55%	5.97%	10.86%	0.59%
5	156	0	0:35:38	212340	848.633	27388
	7.80%		6.42%	3.55%	5.60%	0.59%
6	52	6	0:07:58	182933	1081.27	43144
	2.90%		1.43%	3.06%	7.14%	0.93%
7	1172	0	0:22:04	390123	1383.77	298355
	58.60%		3.97%	6.52%	9.14%	6.41%
8	52	2	0:09:56	146459	208.562	417324
	2.70%		1.79%	2.45%	1.38%	8.97%
9	20	0	0:02:03	54634	176.678	110967
	1.00%		0.37%	0.91%	1.17%	2.38%
10	30	51	1:01:05	1221149	1433.75	3037632
	4.05%		11.01%	20.42%	9.47%	65.26%
11	26	1	0:15:13	395459	1287.72	27829
	1.35%		2.74%	6.61%	8.50%	0.60%
TOTALS			9:15:02	5978952	15147.7	4654831

Figure 11-27. Random Sample Cluster Assignment Report

11.6.1.3.4 Population Cluster Assignment Report

The Cluster Assignment Report for the population indicates how well the clusters represent all of the resource vectors. The report also provides an analysis of the resources that are consumed by the clusters. The resources detailed in this report are specified on the Workload Characterization screen (see Figure 11-18).

The representative hypothesis that was introduced in Section 11.2.2 tests how close the resource vectors are to the characterization proposed for the workload. We propose to test this hypothesis by assigning the resource vectors for the full population to the clusters developed for the sample.

You can test this hypothesis by comparing the marginal assignment percentage for the sample of resource vectors, shown previously in Figure 11-17, to the marginal assignment percentage for the full population shown in Figure 11-28. Since these two values are nearly the same (5.4% vs. 5.7%), you can consider the conditions of the representative hypothesis met. If the test had not been met, you would have repeated the procedure with a larger sample size.

RESOURCE CONSUMPTION BY CLUSTER -- TOTAL POPULATION						
TOTAL OBSERVATIONS :		18633				
OBS WITHIN RADIUS :		17576				
MARGINAL ASSIGNMENTS:		1057	5.7%			
CLUS	WITHIN	MARGIN	JOBTCBTM	JOBEDASD	INITMIN	JOBNLR
1	1296	1	1:31:42	1667621	2830.38	4247705
	6.96%		1.99%	3.15%	1.99%	11.14%
2	1556	2	1:14:19	1578031	13008.4	396291
	8.36%		1.61%	2.98%	9.13%	1.04%
3	280	463	48:34:27	23590139	60822.1	1696906
	3.99%		63.26%	44.60%	42.68%	4.45%
4	432	80	4:06:36	5403274	15489.5	291784
	2.75%		5.35%	10.22%	10.87%	0.77%
5	1517	11	5:53:01	3952170	8446.99	323429
	8.20%		7.66%	7.47%	5.93%	0.85%
6	476	86	1:55:07	2155617	10764.1	620991
	3.02%		2.50%	4.08%	7.55%	1.63%
7	11064	0	3:29:46	3489342	11599.6	2825604
	59.38%		4.55%	6.60%	8.14%	7.41%
8	411	6	1:08:52	1178249	1611.2	3173915
	2.24%		1.49%	2.23%	1.13%	8.33%
9	131	13	0:15:07	319366	1888.09	795981
	0.77%		0.33%	0.60%	1.32%	2.09%
10	186	372	6:18:10	6998863	9600.94	23412139
	2.99%		8.21%	13.23%	6.74%	61.42%
11	227	23	2:19:40	2554131	6444.38	335516
	1.34%		3.03%	4.83%	4.52%	0.88%
TOTALS			76:46:46	52886803	142506	38120261

Figure 11-28. Population Cluster Assignment Report

11.6.1.3.5 Analysis of the ALL_OBS SAS Data Set

The final step of the job class identification process involves some decisions based on the cluster descriptions and the analyst's experience. Although the clusters developed in the example provide an excellent representation of the system's workload, there are too many clusters to define a job class structure with a class for each structure. Providing a minimum number of alternative job classes would simplify the job class selection process. Therefore, after analyzing the data from the cluster centroids and the population statistics reports (see Figures 11-26 and 11-28), we proposed four job classes. These classes are shown in the following table.

CLASS	JOBMXNTA	JOBTCBTM	JOBNLR	CLUSTERS REPRESENTED
=====	=====	=====	=====	=====
1	0	5	2000	7
2	0	60	5000	2,4,9,6
3	>=1, <=2	60	20000	1,3,5,8,10,11
4	unlim	unlim	unlim	outliers

A SAS program was written to evaluate the proposed classes. The program attempts to assign each job to the classes (tested in ascending order of size) until a match is found. The results from the execution of this program are shown in Figure 11-29.

As you can see in the figure, 10,722, 4,155, 1,489, and 2,267 jobs fell into the four classes. These values correspond to 58, 22, 8, and 12%, respectively, of the system's workload. To test the sensitivity of these job class limits, we should repeat the analysis varying the job class limits by + or -5%. If the results of these analyses do not significantly change the percentage of the workload represented by each of the classes (that is, more than 2 or 3% of the jobs change classes), then you can be confident that none of the job classes bisect a resource consumption pattern in the workload.

```

1      DATA _NULL_;
2      SET JOBFIL.ALL_OBS END=EOF_SW;
3      RETAIN C1 C2 C3 C4 0;
4
5      IF JOBMXNTA<=0 & JOBTCBTM<=05 & JOBNLR<=02000 THEN DO;
6          C1+1;
7          GO TO T_EOF;
8      END;
9      IF JOBMXNTA<=0 & JOBTCBTM<=60 & JOBNLR<=05000 THEN DO;
10         C2+1;
11         GO TO T_EOF;
12     END;
13     IF JOBMXNTA>=1 & JOBMXTAP<=2 & JOBTCBTM<=60 & JOBNLR<=20000 THEN DO;
14         C3+1;
15         GO TO T_EOF;
16     END;
17     C4+1;
18     T_EOF:
19     IF EOF_SW THEN DO;
20         PUT C1= / C2= / C3= / C4=;
21         P1=C1/(C1+C2+C3+C4)*100;
22         PUT P1=;
23         P2=C2/(C1+C2+C3+C4)*100;
24         PUT P2=;
25         P3=C3/(C1+C2+C3+C4)*100;
26         PUT P3=;
27         P4=C4/(C1+C2+C3+C4)*100;
28         PUT P4=;
29     END;
30
31     C1=10722
32     C2=4155
33     C3=1489
34     C4=2267
35     P1=57.45
36     P2=22.30
37     P3=7.99
38     P4=12.17
39
40     NOTE: THE DATA STATEMENT USED 1.76 SECONDS AND 384K.
41     RUN;
42     NOTE: SAS USED 384K MEMORY.
43
44     NOTE: SAS INSTITUTE INC.
45           SAS CIRCLE
46           BOX 8000
47           CARY, N.C. 27511-8000

```

Figure 11-29. Sample ALL_OBS Analysis Program

11.6.2 TSO Response

In this example, as part of a benchmark study, an installation wants to prepare TPNS scripts to represent TSO user activity. In particular, they wish to identify patterns of think time and short, medium, and long TSO response events to specify the TPNS scripts.

To identify the TSO activity patterns, the TSOTSU File was analyzed. Four features were specified on the Workload Characterization screen:

AVGTHINK - Average think time. This value is computed by dividing the total think time (TSUTHKTM) by the number of think time events (TSUTHCNT).

SREVPCT - The ratio of short response events to total response events for the user during the CA TSO/MON PM measurement interval. This value is calculated by dividing the short response event count (TSUSRESC) by the total response event count (TSUTRESC).

MREVPCT - The ratio of medium response events to total response events for the user during the CA TSO/MON PM measurement interval. This value is calculated using the TSUMRESC and TSUTRESC values.

LREVPCT - The ratio of long response events to total response events for the user during the CA TSO/MON PM measurement interval. This value is calculated using the TSULRESC and TSUTRESC values.

You may be interested in the reason we used ratios rather than counts to characterize the response event. Although we could have calculated the arrival rate using the total response event values, we could also have calculated the arrival rate based on the average think time. Therefore, you need only the ratio of the response event sizes rather than response event counts, since you can use the average think time to calculate the arrival rate. As we discussed in Section 11.2.4, such rated features offer significantly fewer problems in the control of outliers.

We selected five data elements as reporting variables:

TSUTGETS - The number of terminal TGETS issued

TSUTPUTS - The number of terminal TPUTS issued

TSUTCBTM - The TCB CPU time consumed by the commands

TSUSRBTM - The SRB CPU time consumed by the commands

TSUSWPSZ - The number of 1K blocks swapped for the user

Sample statistics, cluster centroids, and cluster assignments for the total population are shown in Figures 11-30, 11-31, and 11-32.

Figure 11-30 provides the sample statistics for the TSO data. The maximum and minimum values for the ratios of short, medium, and long commands are all 0 and 1. The 0 values correspond to intervals when none of a user's response events were of a specified duration. Conversely, the 1 values correspond to intervals during which all of a user's response events were of a specified duration.

Since all of the ratios were bounded by 0 and 1, outliers were not a consideration. The minimal differences between the sample and trimmed statistics demonstrate this fact. However, examining the statistics for average think time reveals some unexpected values. At first, the maximum value (14,627 seconds) of the average think time appears to be invalid. However, these long think times correspond to sessions where the terminal was simply left logged on for several hours without any activity. Therefore, the outliers represent only inactive terminals. Because of this, the representation of the outliers is of little or no concern for this study.

The centroids of the 15 clusters that were developed are shown in Figure 11-31, TSO Cluster Centroids Report. Consider the characteristics shown for cluster 6. Cluster 6 represents approximately 21% of all of the TSO user measurement interval records. The users represented by this cluster have an average think time of 21.2 seconds and experience short and medium response events in a ratio of 9 to 1. These characteristics are the typical input requirements for system drivers like TPNS.

The final report for the TSO study, shown in Figure 11-32, is the Cluster Assignment Report. In a benchmark study, you would use this report to verify that the percent of the resources that are consumed in the benchmark by the synthetic workloads correspond to the distributions for the actual system.

DESCRIPTIVE STATISTICS								
NUMBER OF OBSERVATIONS: 2000								
FEATURE	MIN	MAX	-----SAMPLE-----			-----TRIMMED 95.00%-----		
			AVERAGE	STD DEV	CV	AVERAGE	STD DEV	CV
AVGTHINK	0.87	14627	107.463	555.687	5.2	42.0253	63.6419	1.5
SREVPCT	0	1	0.38403	0.388045	1.0	0.351269	0.370578	1.0
MREVPCT	0	1	0.358206	0.375055	1.0	0.324072	0.353648	1.1
LREVPCT	0	1	0.257764	0.403596	1.6	0.218288	0.375076	1.7

Figure 11-30. TSO Descriptive Statistics Report

CLUSTER CENTROIDS							
CLUSTER	AVGTHINK	SREVPCT	MREVPCT	LREVPCT	RADIUS	NO ASSIGN	PCT
1	19.7	0.0	0.0	1.0	1.50	339	17
2	25.0	0.1	0.9	0.0	1.50	347	17
3	127.9	0.5	0.5	0.0	2.00	33	2
4	147.8	0.0	1.0	0.0	1.50	49	2
5	20.0	0.5	0.5	0.0	1.50	395	20
6	21.2	0.9	0.1	0.0	1.00	411	21
7	19.6	0.5	0.1	0.4	1.50	55	3
8	233.1	0.0	0.0	1.0	3.00	28	1
9	427.0	0.0	1.0	0.0	2.50	18	1
10	20.8	0.1	0.4	0.5	2.00	85	4
11	540.5	0.0	1.0	0.0	2.00	81	4
12	267.1	1.0	0.0	0.0	3.00	34	2
13	296.9	0.0	1.0	0.0	2.00	21	1
14	103.6	1.0	0.0	0.0	1.50	61	3
15	122.0	0.0	0.0	1.0	2.00	43	2

Figure 11-31. TSO Cluster Centroids Report

CLUSTER ASSIGNMENT -- TOTAL POPULATION							
TOTAL OBSERVATIONS :		15104					
OBS WITHIN RADIUS :		14568					
MARGINAL ASSIGNMENTS:		536 3.5%					
CLUS	WITHIN	MARGIN	TSUTGETS	TSUTPUTS	TSUTCBTM	TSUSRBTM	TSUSWPSZ
1	2652	0	42742	179645	1:30:01	0:06:54	2534180
	17.56%		22.42%	24.99%	27.99%	19.76%	22.84%
2	2715	0	27731	133719	0:54:28	0:06:07	1844684
	17.98%		14.54%	18.60%	16.94%	17.52%	16.62%
3	247	0	1327	5877	0:02:38	0:00:20	110092
	1.64%		0.70%	0.82%	0.82%	0.95%	0.99%
4	365	4	1188	9135	0:02:24	0:00:16	122952
	2.44%		0.62%	1.27%	0.75%	0.77%	1.11%
5	2888	0	45130	162621	1:11:16	0:08:11	2440348
	19.12%		23.67%	22.62%	22.16%	23.48%	21.99%
6	3114	0	53320	125302	0:41:40	0:07:25	2508048
	20.62%		27.96%	17.43%	12.96%	21.28%	22.60%
7	367	3	5450	26948	0:15:30	0:01:30	315648
	2.45%		2.86%	3.75%	4.82%	4.30%	2.84%
8	179	16	578	3790	0:01:55	0:00:13	75240
	1.29%		0.30%	0.53%	0.59%	0.61%	0.68%
9	98	4	203	1176	0:00:26	0:00:03	20740
	0.68%		0.11%	0.16%	0.14%	0.14%	0.19%
.							
.							
15	293	0	1269	8928	0:05:42	0:00:31	166184
	1.94%		0.67%	1.24%	1.77%	1.49%	1.50%
TOTALS			190681	718918	5:21:38	0:34:53	11097664

Figure 11-32. TSO Cluster Assignment Report

11.6.3 IMS Transaction

To investigate potential IMS transaction classes, we used Workload Characterization to analyze DETAIL IMS transaction data collected with the optional _USRXISU exit in the IMS Component. In particular, we analyzed the IMS transaction data to determine patterns based on CPU time, message queue calls, and database calls.

To identify potential IMS transaction classes, we analyzed the DETAIL timespan IMSISU File. (The IMSISU File is not normally supported in the DETAIL timespan, but was included through the use of an optional exit.)

We specified three data elements on the Workload Characterization screen:

ISUCPUTM - The CPU time consumed by the transaction

ISUDBTOT - The total number of database calls issued by the transaction

ISUMSTOT - The total number of message queue calls issued by the transaction

We selected four data elements as reporting variables:

ISUDBGU - The number of database GET UNIQUES issued by the transaction

ISUDBGN - The number of database GET NEXTs issued by the transaction

ISUMSGET - The number of message queue GET UNIQUE and GET NEXTs issued by the transaction. This value is computed from ISUMSGGU and ISUMSGGN.

ISUCPUTM - The CPU time consumed by the transaction

The centroids of the six clusters that we developed are shown in Figure 11-33. Cluster 2 represents 71% of the transactions that were processed. We could potentially use the transactions in this cluster to define an IMS transaction class. However, we would have to perform further analysis on the ALL_OBS SAS data set to verify that the DBD requirements of the transactions do not conflict (ART78, LEE83).

CLUSTER CENTROIDS						
CLUSTER	ISUCPUTM	ISUDBTOT	ISUMSTOT	RADIUS	NO ASSIGN	PCT
1	0.13	11.9	1.23	1.00	219	11
2	0.21	9.8	2.99	1.00	1423	71
3	0.17	17.8	0.5	1.00	77	4
4	1.13	150.3	1.0	2.00	41	2
5	0.85	1.3	61.2	1.50	137	7
6	0.53	0.9	0.1	1.00	103	5

Figure 11-33. IMS Cluster Centroids Report

Chapter 12: INSTALLATION

The Installation Chapter describes installation, backup, restore, and compression procedures for the CA MICS Capacity Planner and how to enable Neugents technology.

This section contains the following topics:

[12.1 Installation Procedure](#) (see page 892)

[12.2 Backups, Restores, and Compression](#) (see page 893)

[12.3 How to Enable Neugents technology in CA MICS](#) (see page 901)

12.1 Installation Procedure

Instructions for loading the CA MICS Capacity Planner were included with the CA MICS product distribution tape. In the second step of these instructions you ran the product load job to copy all required product modules and members into your CA MICS libraries. Use the following checklist to complete the installation of the CA MICS Capacity Planner. All references to the PIOM are to sections of the CA MICS Planning, Installation, Operation, and Maintenance Guide.

- ___ 1. Using the JCL shown in Figure 12-1, allocate a shared capacity planning database for your installation. The MICSNDBx cataloged procedure is used to execute the program, where x is the database identifier of the primary CA MICS unit database. If you are unfamiliar with your CA MICS database identifiers, consult your CA MICS System Administrator.

```
//ALLOCATE EXEC MICSNDBx
//*
//* SHARED CAPACITY PLANNING DATABASE ALLOCATION
//* SPECIFY VOLSER AND UNIT TYPE
//*
//CAPACITY DD DSN=sharedprefix.MICS.CAPACITY,
//          DISP=(NEW,CATLG,DELETE),
//          DCB=(DSORG=PS,RECFM=FS),
//          SPACE=(TRK,(300,50)),UNIT=____,VOL=SER=_____
//SYSIN DD *
DATA CAPACITY.DUMMY; DUMMY=1; RUN;
//
```

Figure 12-1. Capacity Planning Database Allocation JCL

Note that the 300 tracks specified is an estimate of initial requirements. You should carefully monitor the space utilization of the capacity planning database to optimize DASD utilization.

- ___ 2. Update the product statement in JCLDEF in the primary unit to include the character string 'CAP' to indicate the product is installed.

```
+-----+
| Reference(s): PIOM, Section 2.3.3.2.1.1 |
+-----+
```

- ___ 3. Generate jobs GDGSGEN, BACKUP, BACKUPM, and MONTHLY using the JCLGENU job.

```
+-----+
| Reference(s): PIOM, Section 3.5.3.4 |
+-----+
```

- ___ 4. Run the GDGSGEN job.

```
+-----+
| Reference(s): PIOM, Section 3.5.6.3 |
+-----+
```

- ___ 5. Run a CGEN for all the products you have installed in your complex. The CGEN jobs exist in your sharedprefix.MICS.CNTL library.

```
+-----+
| Reference(s): PIOM, Section 3.4.3 |
+-----+
```

Note that if you do not run a CGEN for all of your installed products, you will experience errors when using the Capacity Planner Online Facilities.

You have now completed the product installation.

12.2 Backups, Restores, and Compression

PRIVATE CAPACITY PLANNING DATABASE BACKUP

Your private capacity planning database is NOT backed up as part of the CA MICS primary unit database backup. For this reason you need to run your own backup job periodically to back up both your private capacity planning database SAS data set and the ISPF tables that are essential to the operation of the CA MICS Capacity Planner. The frequency with which you run the backup job depends on how often you update your capacity planning database files and how often you create and save analysis inquiries. In general, if you update your capacity planning database files weekly, you should run your backup job weekly. If you run your updates monthly, you should run your backup job monthly.

A number of methods are possible for backing up your private capacity planning database SAS data set and the corresponding ISPF tables. The checklist and sample JCL below provide one method for accomplishing the backup to a generation data set.

- ___ 1. You need to know the data set name of your private capacity planning database file. If you already know the data set name, proceed to Step 7 of this checklist. Otherwise, access the main CA MICS Workstation Facility (MWF) screen as described in Chapter 1 of this guide. Select Option 4, Management Support Applications (APPL)
- ___ 2. From the Management Support Applications screen, select Option 3, Capacity Planning.
- ___ 3. From the Capacity Planning screen, select Option 0, MICF Options.
- ___ 4. From the MICF Options screen, select Option 6, Data Set Allocation Parameters.
- ___ 5. From the Data Set Allocation screen, select the row identified as CAPACITY by placing an S in the Cmd column of this row. The resulting screen displays the data set name and the allocation parameters that you will use to allocate your private capacity planning database file.
- ___ 6. Record the data set name for use in the JCL shown in Figures 12-2 and 12-3.

- ___ 7. If you have not already defined the generation data sets for backing up your private capacity planning SAS data set and the corresponding ISPF tables, you need to do so now. You only need to execute this step the first time you back up your capacity planning database file. Figure 12-2 shows one method of doing so for a VSAM catalog structure.

```
//DEFINGDG EXEC PGM=IDCAMS
//*
/* DEFINE VSAM BACKUP GDG FOR PRIVATE CAPACITY FILE
/*
//SYSPRINT DD SYSOUT=*
//SYSIN DD *
DEFINE GDG -
  (NAME(userid.MICS.BACKUP.CAPACITY) -
   SCRATCH -
   LIMIT(3) )
DEFINE GDG -
  (NAME(userid.MICS.BACKUP.TABLES) -
   SCRATCH -
   LIMIT(3) )
```

Figure 12-2. JCL to Define Backup GDGs for Private Files

- ___ 8. You are now ready to execute the JCL to back up your private capacity planning database file and the corresponding ISPF tables. You may use the JCL shown in Figure 12-3 for this purpose. Note that this is a two-step job and that you can make several assumptions.

Refer to the first step, BKUPSAS, in Figure 12-3. If the name of the SAS PROC at your installation is not SAS, change the JCL to reflect the correct name. Also, if the name of the private capacity planning SAS data set that you identified in Step 6 of this checklist does not match the format shown below for the CAPACITY DD statement, then change the JCL to reflect the correct name. The data set named on the BKUPCAP DD statement refers to a generation data set that you either defined previously or that you defined by executing the JCL in Step 7 of this checklist. As with the CAPACITY DD statement, ensure that the data set name on the BKUPCAP DD statement reflects the correct name at your installation.

Refer to the second step, BKUPTBL, in Figure 12-3. If the name of your private ISPF table library does not follow the format shown on the SYSUT1 DD statement, change the JCL to reflect the correct name. The data set name on the SYSUT2 DD statement refers to a generation data set that you either defined previously or you defined by executing the JCL in Step 7. As with the SYSUT1 DD statement, ensure that the data set name on the SYSUT2 DD statement reflects the correct name at your installation.

```
//BKUPSAS    EXEC MICSNDBx
/*
/* PRIVATE CAPACITY PLANNING DATABASE SAS DATA SET BACKUP
/*
//CAPACITY  DD DSN=userid.MICS.USER.CAPACITY,DISP=SHR
/*
//BKUPCAP   DD DSN=userid.MICS.BACKUP.CAPACITY(+1),
//           DISP=(NEW,CATLG,DELETE),UNIT=unit,
//           DCB=(userid.MICS.USER.CAPACITY),
//           LABEL=(1,RETPD=360)
//SYSIN     DD *
PROC COPY IN=CAPACITY OUT=BKUPCAP; RUN;
/*
//BKUPTBL   EXEC PGM=IEBCOPY
/*
/* PRIVATE CAPACITY PLANNING ISPF TABLE BACKUP
/*
//SYSPRINT  DD SYSOUT=*
//SYSUT1    DD DISP=SHR,DSN=userid.MICS.USER.TABLES
//SYSUT2    DD DSN=userid.MICS.BACKUP.TABLES(+1),
//           DISP=(NEW,CATLG,DELETE),UNIT=unit,
//           DCB=(userid.MICS.USER.TABLES),
//           LABEL=(1,RETPD=360)
//SYSUT3    DD UNIT=SYSDA,SPACE=(TRK,(3))
//SYSUT4    DD UNIT=SYSDA,SPACE=(TRK,(3))
//SYSIN     DD *
            COPY INDD=SYSUT1,OUTDD=SYSUT2
```

Figure 12-3. Private Capacity Planning Backup JCL

- ___ 9. You have now backed up your private capacity planning database file SAS data set and the corresponding ISPF tables.

PRIVATE CAPACITY PLANNING DATABASE RESTORE

If you find it necessary to restore your private capacity planning database file, you may do so by executing the sample JCL shown in Figure 12-4. As with the backup job described above, you can make a number of assumptions about data set and PROC names at your installation. Review the discussion above and ensure that the PROC and data set names at your installation are correctly reflected in the following JCL before submitting the job.

```
//RSTRSAS    EXEC MICSNDBx
/*
/**PRIVATE CAPACITY PLANNING DATABASE SAS DATA SET RESTORE
/*
//CAPACITY   DD DSN=userid.MICS.USER.CAPACITY,DISP=OLD
/*
//BKUPCAP    DD DSN=userid.MICS.BACKUP.CAPACITY(0),
//           DISP=OLD
//SYSIN      DD *
PROC COPY IN=BKUPCAP OUT=CAPACITY; RUN;
/*
//RSTRTBL    EXEC PGM=IEBCOPY
/*
/**PRIVATE CAPACITY PLANNING ISPF TABLE RESTORE
/*
//SYSPRINT   DD SYSOUT=*
//SYSUT1     DD DSN=userid.MICS.BACKUP.TABLES(0),
//           DISP=OLD
//SYSUT2     DD DISP=SHR,DSN=userid.MICS.USER.TABLES
//SYSUT3     DD UNIT=SYSDA,SPACE=(TRK,(3))
//SYSUT4     DD UNIT=SYSDA,SPACE=(TRK,(3))
//SYSIN      DD *
              COPY INDD=SYSUT1,OUTDD=SYSUT2
```

Figure 12-4. Private Capacity Planning Restore JCL

PRIVATE CAPACITY PLANNING DATABASE COMPRESSION

If you find it necessary to compress your private capacity database files containing the historical observations (CAPxxxnn), where xxx is the file identifier and nn denotes a cycle number, you may do so by executing sample JCL shown in Figure 12-5.

This JCL first creates a temporary SAS data set. The first PROC COPY will copy all historical files from your private capacity database into the temporary data set. The second PROC COPY will copy and compress all files back into your private capacity database.

Review your SASLOG to determine your compression savings for the entire capacity database. In some cases, a compressed SAS file may take more space than if it were not compressed. If you determine that compression is not a good option for your private capacity database, rerun the compression JCL, setting the COMPRESS option to COMPRESS=NO. Running this job will uncompress all your private capacity historical database files.

As with the backup & restore jobs described above, you can make a number of assumptions about data set and PROC names at your installation. Review the discussion above and ensure that the PROC and data set names at your installation are correctly reflected in the following JCL before submitting the job.

```
//CMRSPR    EXEC MICSNDBx
/*
/**PRIVATE CAPACITY PLANNING DATABASE COMPRESSION
/*
//CAPACITY  DD DSN=userid.MICS.USER.CAPACITY,DISP=OLD
/*
//TEMPCAP   DD DSN=userid.MICS.TEMP.CAPACITY,
//           DISP=(NEW,DELETE,DELETE),UNIT=unit,
//           DCB=(userid.MICS.USER.CAPACITY),
//           SPACE=(CYL,(5,2))
//SYSIN     DD *
PROC COPY IN=CAPACITY OUT=TEMPCAP;
  SELECT CAP: ; RUN;
OPTIONS COMPRESS=YES;
PROC COPY IN=TEMPCAP OUT=CAPACITY NOCLONE;
  SELECT CAP: ; RUN;
/*
```

Figure 12-5. Private Capacity Planning Compression JCL

NOTE:

To turn off compression during a capacity database initialization or subsequent database update, modify the SAS macro variable CAPCMPRS=YES in shared.prefix.SOURCE(\$CAPMSTR) to CAPCMPRS=NO;

SHARED CAPACITY PLANNING DATABASE BACKUP AND RESTORE

The CA MICS shared capacity planning database and its associated tables are backed up whenever the CA MICS primary unit database backup job is executed. The JCL required to restore the shared capacity planning database and the tables is shown in Figure 12-6.

```
*****
* WARNING: This restore process will also restore tables *
* which are used in controlling and maintaining the entire *
* CA MICS system.                                         *
*                                                         *
* YOU SHOULD NOT PERFORM THIS YOURSELF, BUT SHOULD ASK THE *
* CA MICS SYSTEM ADMINISTRATOR TO PERFORM IT FOR YOU AFTER *
* CONFIRMING THAT THIS RESTORE PROCESS WILL NOT DESTROY ANY *
* RECENT MAINTENANCE, SHARED MICF INQUIRIES, OR SHARED MICF *
* CATALOGED REPORTS THAT THE ADMINISTRATOR HAS APPLIED TO *
* THE SYSTEM.                                             *
*                                                         *
* If you need to restore only a few tables describing *
* shared files or shared inquiries, you should call *
* CA MICS Product Support                                *
*****
```

Use the MICSNDBx cataloged procedure to execute the program, where x is the database identifier of the primary CA MICS unit database.

```
//RSTRSAS    EXEC MICSNDBx
//*
//* SHARED CAPACITY PLANNING DATABASE RESTORE
//*
//CAPACITY   DD DSN=sharedprefix.MICS.CAPACITY,DISP=OLD
//*
//BKUPCAP    DD DSN=tapeprefix.MICS.BACKUP.CAPACITY(0),
//            DISP=OLD
//SYSIN       DD *
PROC COPY OUT=CAPACITY IN=BKUPCAP; RUN;
/*
//RSTRTLIB   EXEC PGM=IEBCOPY
//*
//*SHARED CAPACITY PLANNING ISPF TABLE RESTORE
//*
//SYSPRINT   DD SYSOUT=*
//SYSUT1     DD DSN=tapeprefix.MICS.BACKUP.ISPTLIB(0),
//            DISP=OLD
//SYSUT2     DD DISP=OLD,DSN=sharedprefix.MICS.ISPTLIB
//SYSUT3     DD UNIT=SYSDA,SPACE=(TRK,(3))
//SYSUT4     DD UNIT=SYSDA,SPACE=(TRK,(3))
//SYSIN      DD *
              COPY INDD=SYSUT1,OUTDD=SYSUT2
//
```

Figure 12-6. Shared Capacity Planning Restore JCL

SHARED CAPACITY PLANNING DATABASE COMPRESSION

12.3 How to Enable Neugents technology in CA MICS

The CA MICS Performance Manager, Capacity Planner, and StorageMate products include features that use CA's U.S. patented Neugents(R) technology.

Before you use Neugents, be sure your environment is ready. See these subsequent sections for more information:

- 1 - Requirements
- 2 - Activate the CA MICS Root Directory
- 3 - Enable Neugents Technology

12.3.1 Requirements

CA MICS delivers mainframe-based Neugents technology that requires C/C++ runtime features available in the language environment delivered with z/OS 1.3 and above. If you plan to use CA MICS product features that rely on Neugents technology, then CA MICS must be installed on a z/OS system at level 1.3 or above.

12.3.2 Activate the CA MICS Root Directory

To activate the CA MICS root directory, review Section 5.7 in the PIOM Guide and complete the activation checklist detailed in Section 5.7.3.

NOTE: If the CA MICS root directory has been enabled during the installation of another CA MICS product, you do not have to do it again.

12.3.3 Enable Neugents technology

Once you have activated the CA MICS root directory on USS, you can enable Neugents technology. By following the steps in this procedure, you will enable Neugents technology by delivering the Neugents technology DLLs to the correct directory under the CA MICS root directory.

To enable Neugents technology, complete the following steps:

- 1.____ Edit `sharedprefix.MICS.PARMS(JCLGENUC)` so that it contains the following:

```
MQRHFSU
```

Submit the job in `sharedprefix.MICS.CNTL(JCLGENUC)`. Ensure that there are no error messages in `MICSLOG` or `SYSTSPRT`, that the `MICSLOG` contains the normal termination message, `BAS10999I`, and that the job completes with a condition code of zero.

- 2.____ Unload Neugents technology HFS files to the CA MICS Root Directory.

Submit the following job:

```
'sharedprefix.MICS.CNTL(MQRHFSU)'
```

Ensure that there are no error messages in `MICSLOG` or `SYSTSPRT`, and that the job completes with a condition

code of zero.

Chapter 13: RESOURCE COMPONENT ANALYSIS

A new facility has been added to the CA MICS Capacity Planner that will analyze the effects of a series of numeric data values upon a client specified Analysis or dependent element, and then quantify those effects. This is an exploratory data analysis tool and can be used in selecting data elements for later use in resource forecasting operations.

This new facility can be executed in two different modes: one using any CA MICS file as input, and the other, using a CAP database as input. Although there are differences in how the facility is invoked between each mode, the internal processing is the same in both cases.

This section contains the following topics:

[13.1 Analysis using CA MICS files](#) (see page 903)

[13.2 Analysis using CAP Database files](#) (see page 920)

13.1 Analysis using CA MICS files

A new facility, Resource Component Analysis, using CA MICS files has been added to the CA MICS MVS Capacity Planner that will analyze the effects of a series of numeric data values (independent elements) upon a client specified dependent element, and then quantify those effects. This exploratory data analysis tool can be used in selecting appropriate data elements for use in populating future capacity planning databases.

The sections that follow describe this facility.

- 1 - Functional Description
- 2 - Usage Guidelines
- 3 - Standard Output
- 4 - Analytic Technique Tutorial
- 5 - Component Operation
- 6 - Case Studies

13.1.1 Functional Description

The Resource Component Analysis application uses the Relative Importance facility (CA's U.S. patented Neugents(R) technology) and utilizes proprietary analytical methods to understand and quantify the effect of a list of data values (independent elements) against a client specified dependent element. While a similar analysis can be performed using standard statistical means such as correspondence matrix, the Relative Importance facility can not only determine linear relationships among data values but non-linear relationships as well. This is important when exploring data that may not be well understood and can result in a more accurate understanding of the relationships among the data values. Any files to be analyzed using this option of the Resource Component Analysis application must reside in the CA MICS database.

The process itself is straightforward. First, determine the dependent data element to use for analysis. The CA MICS Capacity Planner will then accept a list of quantitative independent data elements defined on the input screen. The client can enter a '?' in the input area of the scrollable portion of the input panel and obtain a list of all data elements for the file specified as the input file. Any numerical quantitative data element(s) can be selected, and multiple elements can be chosen concurrently. Note that a minimum of two elements must be defined or an error message will be displayed.

The application will validate all chosen independent elements for "missing" values. These values are not valid for performing Relative Importance and Neugents will ignore them during processing. Input observations containing "missing" values will be bypassed and listed on an error report.

Additionally, all independent values will be analyzed to determine if they are totally co-related with the dependent value. For example, in the HARCPU file, CPUTOBTM and CPUPCBSY would be totally co-related and as such, CPUPCBSY values would bias the Relative Importance process and cause the Importance Index (Weight) values to be distorted. The application will compute co-relation weights for all independent elements and will cause any totally co-related values to be dropped from analysis and flagged on the Descriptive Statistics report.

During execution, the Relative Importance facility will analyze the values of each selected independent element and develop a weight that quantifies the effect of the data value against the dependent element. The sum of the weights for all selected elements will be 100 and will quantify the impact of each selected independent data element against the dependent variable. Those data values with a greater weight have a more significant impact on the dependent element than those with a lesser weight, and will likely be of more interest in any future analysis.

The CA MICS Capacity Planner Option include features that utilize CA's U.S. Patented Neugents(R) technology.

To exploit this technology, the CA MICS Root Directory under Unix System Services needs to be initialized and the Neugents technology infrastructure enabled.

For detail instructions on how to initialize the CA MICS root directory and enable Neugents technology review, see Chapter 12 of this guide.

13.1.2 Usage Guidelines

Use of the Resource Component Analysis application is easy and the operation is straightforward. You can normally execute a complete analysis with a single panel and a few key strokes. Once the analysis has completed, you can examine the output report produced, quickly identify those independent elements that have minimal effect on the dependent element values, remove those elements, and reexecute the analysis query again. This process can be executed as often as necessary until the best set of independent elements remains. Once an optimum set of independent elements has been determined, they can be used in future studies.

13.1.3 Standard Output

The Resource Component Analysis facility currently produces a single printed output: The Relative Importance Analysis Report. This report contains a set of simple descriptive statistics for the dependent elements and their associated values, as well the relative importance weight assigned to each independent individual element. A sample of this report is shown next.

Resource Component Analysis Report							
Report For: 01/27/04							
Feature	Description	Minimum	Maximum	Average	Std. Dev.	CV	Rel. Imp.
CPUTOBTM	Total Time Processor Busy	0:08:15.73	0:27:00.70	0:19:23.79	0:03:23.89	17.52	
CPUAVB	Avg Batch Users	3.16	10.23	5.16	1.83	35.50	0.2265
CPUAVE	Avg OMVS Users	12.00	12.10	12.03	0.02	0.15	0.2834
CPUAVS	Avg Started Users	196.47	221.82	206.65	7.42	3.59	0.2333
CPUAVT	Avg TSO Users	2.04	16.72	9.96	4.15	41.71	0.2568
CPUTODTM	Total Dispatch Time	0:08:15.73	0:27:00.70	0:19:23.79	0:03:23.89	17.52	0.0000***
Total Relative Importance Weights:							1.0000
*** - This Field(s) is 100% Correlated to the CPUTOBTM Value and was Dropped from Analysis.							

The Relative Importance Analysis report contains the following information:

ELEMENT:	The name of the independent and dependent elements being analyzed.
DESCRIPTION:	The SAS label of the selected data element, from either the CA MICS GENLIB definition or supplied by the user.
MIN. VALUE:	The minimum value of the selected data element observed in the input data.
MAX. VALUE:	The maximum value of the selected data element observed in the input data.
MEAN VALUE:	The mean of the values of the selected data element observed in the input data.
STD:	The standard deviation of the values of the selected data element observed in the input data.

CV:	The coefficient of variance of the values of the selected data element observed in the input data. This is a standard statistic and is a ratio of the mean and standard deviation measurements.
REL. IMP.:	The relative importance of this particular independent element as a function of the dependent element value. Neugents technology computes this value internally using a mixture of both linear and nonlinear techniques and expresses this as a ratio or weight factor, relative to 100.

13.1.4 Analytic Technique Tutorial

This section is not available at this time.

13.1.5 Component Operation

This section describes the control parameters that are used in the execution of the Resource Component Analysis inquiry. The control parameters allow you to define the CA MICS file you want to use as input, the time periods you want to analyze, the dependent elements, and the associated independent elements you want to use in the analysis process.

The following sections discuss each control panel in detail:

- 1 - Resource Component Analysis Control Panel
- 2 - Extended Options Panel
- 3 - Global Data Selection Panel
- 4 - Data Element Derivation Panel
- 5 - Alternate Input File Panel
- 6 - Save Selected Input Observations Panel
- 7 - Inquiry Processing Panel

13.1.5.1 Resource Component Analysis Control Panel

Resource Component Analysis is one of the capacity planning application options available through the CA MICS Workstation Facility (MWF). Chapter one of this guide explains how to use MWF to reach the Capacity Planner options and subsequently the Relative Importance control panel shown in Figure 13-1.01.

```

/----- Resource Component Analysis -----\
|Command ==>                               Scroll ==> CSR      |
|                                           |
|Composing CA MICS Inquiry:   IOISC   - Adhoc Inquiry -      |
|CA MICS file ==> CPU - CPU ACTIVITY FILE           Time-span DETAIL |
|Data base ==> P                Allow Run-time Override of DB(s) => Y (Y/N) |
|Cycle(s) ==> 01 - 02          Allow Run-time Override of Cycles => Y (Y/N) |
|                                           |
|Data Selection:                               Allow runtime override |
|  Date           ==>          -              -              ==> N (Y/N) |
|  Hour           ==>          -              -              ==> N (Y/N) |
|  Zone           ==>                               ==> N (Y/N) |
|  SYSID          ==>                               ==> N (Y/N) |
|                                           |
|Specify extended options ==>   (Y/N) |
|                                           |
|Specify Dependent Element ==> CPUTOMSU ( Use a ? to select from list ) |
|                                           |
|Line Cmds: I Insert  D Delete |
|                                           |
|Specify Two or More Independent Elements Below ( Use a ? to select from list ) |
|                                           |
| X Element                Description |
| - ----- |
| CPUAVB   Average Batch Users |
| CPUAVT   Average TSO Users |
| CPUAVS   Average Started Tasks |
| CPUAVE   Average OMVS Users |
|                                           |
\-----/

```

Figure 13-1.01. Resource Component Analysis

Some of the parameters on this panel are required, while others are optional. Each parameter is discussed next.

CA MICS file:	The CA MICS database file containing the historical observations for the Analysis file used in the Relative Importance Analysis application. If you are uncertain of the files or file names available in your CA MICS database, type a question mark (?) in the field and a selection list will be displayed.
Data base:	The single-character database identifier for this file at your site.
Cycle(s):	The number of cycles to be specified. For example, we specified two cycles to analyze two days worth of data.
Data Selection:	The selection criteria for Date, Hour, Zone, and SYSID to subset the data you are examining. You must enter date values in the ddmonyy format.
Specify extended options:	Select this option if you want to specify additional parameters. Enter Y (YES) to display the Extended Options panel.
Specify Dependent Element:	The name of the dependent element used in the Relative Importance Analysis application. If you are uncertain of the elements available in your CA MICS file, type a question mark (?) in the field and a selection list will be displayed.
Specify TWO or More Independent Elements:	The independent elements that you select for the analysis. If you are uncertain of the elements available in your CA MICS file, type a question mark (?) in the field and a selection list will be displayed.

13.1.5.2 Extended Options Panel

Entering Y (YES) at the Specify extended options prompt on the Relative Importance panel displays the Extended Options Panel shown in Figure 13-1.02.

```

/----- Extended Options Menu -----\
|Command ==>
|
|Inquiry Step:  Relative Importance
|
|
|Specify global data selection exit          ==> N (Y/N/R)
|Specify non-CA MICS data element derivation==> N (Y/N/R)
|
|Specify alternate input file                ==> N (Y/N/R)
|Save selected input observations            ==> N (Y/N/R)
|
|
|
|
|
\-----\

```

Figure 13-1.02. Extended Options Panel

You can select any of the options available on this panel by entering Y (YES) at any of the prompts. Entering R (RESET) nullifies any specifications you made for that option.

13.1.5.3 Global Data Selection Panel

Figure 13-1.03 displays the Global Data Selection Macro Definition panel.

[illegible]

Figure 13-1.03. Global Data Selection Macro Definition Panel

Relative Importance provides a user-exit that you can use to control data selection from the CA MICS database. To invoke this exit, code SAS statements in the SELEOF macro. This macro is bracketed by the %MACRO and %MEND statements that are field protected to prevent accidental modification.

The default for the SELEOF macro is null. This macro is invoked when the Relative Importance observations are read from the specified CA MICS file. You can use it to delete observations from the data as it is selected. For example, in a job class study, you might want to exclude production job classes from the analysis. If production jobs are designated by job class P, code the following:

```
%MACRO SELEOF;  
  IF JOBCLASS='P' THEN GO TO T_EOF;  
%MEND SELEOF;
```

Note that you do not have to code the %MACRO and %MEND statements, as they are already provided for you.

The IF statement should specify branching to T_EOF rather than acting as a subsetting IF or specifying DELETE. A subsetting IF or a DELETE statement may cause the program to function incorrectly.

You must code each specified SAS statement in accordance with the syntax of the SAS Macro Language, as described in the SAS Language Reference: Dictionary (Version 8).

13.1.5.5 Alternate Input File Panel

Figure 13-1.05 displays the Alternate Input File panel.

```
/----- Alternate Input File -----\
|Command ==>
|
|Inquiry Step: Relative Importance
|
|SAS file name(s) ==> _____
|
|
|SAS Data set allocation instructions:
|   Data set name    ==> _____
|   Disposition      ==> SHR  (SHR/OLD)
|
|
|
|
|
```

Figure 13-1.05. Alternate Input File Panel

Resource Component Analysis has the ability to read data from CA MICS files, or from non-CA MICS external SAS files. The Alternate Input File panel lets you specify non-CA MICS SAS files as input data for the analysis. You can specify the SAS data set name and up to five SAS files to be used as input. Note that the SAS data must already be created.

13.1.5.6 Save Selected Input Observations Screen

Figure 13-1.06 illustrates the Save Selected Input Observations screen.

[illegible]

Figure 13-1.06. Save Selected Input Observations Screen

Resource Component Analysis has as the ability to save the observations that your analysis inquiry selected to an external SAS file for further processing by independent SAS programs, or for future processing by CA MICS Capacity Planner. The Save Selected Input Observations screen allows you to save the selected observations to an external SAS file and specify the SAS data set name and the name of the SAS file in which they should be saved. Note that the SAS data set must already be created.

13.1.5.7 Inquiry Processing Panel

Figure 13-1.07 illustrates the Inquiry Processing panel.

```

/----- Inquiry Processing -----\
|Command ===>
|
|
|Active MICF Inquiry:  RICPU - Relative Importance
|                Author:      Your Name
|
|
|Specify Inquiry processing ==> EXECUTE (EXECUTE/MODIFY/SAVE/CANCEL)
|EXECUTE - Execute the inquiry.
|MODIFY  - Redisplay the inquiry for modification.
|SAVE    - Exit from inquiry composition and save the inquiry.
|CANCEL  - Exit from inquiry composition without saving the inquiry.
|
|
|Instructions:
|
|Specify desired option and press ENTER key.
|
\-----\

```

Figure 13-1.07. Inquiry Processing Panel

Entering END at the Command prompt or pressing the END PF key (usually PF3) while on the Relative Importance Analysis panel displays the Inquiry Processing panel. The available options are described in the text of the panel. Note that you can save a partially composed inquiry for later modification and execution.

13.1.6 Case Study

The following case study involves the use of Relative Importance Analysis used in preparation of a processor resource forecast.

1 - Processor Resource Study

13.1.6.1 Processor Resource Usage

This case study illustrates how you can perform a preliminary analysis of processor usage data and better understand which general workloads appear to influence the overall system usage. This is a simple exploratory data step that can help direct other, more detailed analyses. Since capacity planning is often done in a very individualized fashion and the actuals metrics vary among sites, it is important to have tools that are flexible and permit you to choose among a variety of metrics, measurements and methods to achieve the best outcome.

DEVELOPMENT OF CONCEPTUAL STUDY

Knowledge gained from this and similar limited analyses can help you choose which metrics to employ in more detailed studies. For example, you might decide that simple usage data as depicted here is too limited to be reliable and that data from the WLMSECxx or related files would be required. But for quick studies and simple "back of the envelope" analyses, the method employed here may be sufficient.

APPLICATION OF CONCEPTUAL STUDY

The systems being examined are SYS1 and SYS2 (the CA MICS SYSID values). These systems process varied workloads, which can be loosely grouped together as BATCH, TSO, STC and USS. They are also used for both system development and non-scheduled production, and serve a number of remotely located groups. The hardware platforms have changed over time and both LPARs reside on a z-900 based CEC executing z-OS version 1.4.

The DETAIL HARCPU data from a typical day was chosen and the analyses performed. The analysis report was reviewed and the decision made to drop the 'average' data elements (CPUAVx) in favor of the 'maximum' values (CPUMXx). The query was then reexecuted and the results noted.

The next sections explore the results of the report that is generated from the Relative Importance Analysis we developed for the workload usage study:

- 1 - Control Parameters
- 2 - Relative Importance Analysis Report

13.1.6.1.1 Control Parameters

Using the short set of menus that are described in Section 1.5 of this guide, you will arrive at the Resource Component Analysis screen shown in Figure 13-1.08. This figure shows the control parameters that were used to generate the Resource Component Analysis query in this case study.

```

/----- Resource Component Analysis -----\
|Command ==>                               Scroll ==>|
|
|Modifying CA MICS Inquiry: CPUBSY - Analyze Processor Busy by Workload|
|CA MICS file ==> CPU - CPU ACTIVITY FILE           Time-span: DETAIL|
|Data base ==> P                Allow Run-Time Override of DB(s) => N (Y/N)|
|Cycle(s) ==> 01 - 01          Allow Run-Time Override of Cycles => N (Y/N)|
|
|Data Selection:                               Allow run-time override|
| Date      ==> _____ - _____ - _____ ==> N (Y/N)|
| Hour      ==> __ - __ - __ - __ - __ - __ ==> N (Y/N)|
| Zone      ==> 1 2 3 _ _ _ _ _ ==> N (Y/N)|
| SYSID     ==> SYS1 SYS2 _ _ _ _ _ ==> N (Y/N)|
|
|Specify extended options ==> N (Y/N)|
|
|Specify Dependent Element ==> CPUTOBTM ( Use a ? to select from list )|
|
|Line Cmds: I Insert  D Delete|
|
|Specify Two or More Independent Elements Below ( Use a ? to select from list )|
|
| X Element          Description|
| - - - - - - - - - - - - - - -|
| __ CPUAVB      Avg Batch Users|
| __ CPUAVE      Avg OMVS Users|
| __ CPUAVS      Avg Started Users|
| __ CPUAVT      Avg TSO Users|
| __ CPUMXB      Max Batch Users|
| __ CPUMXE      Max OMVS Users|
| __ CPUMXS      Max Started Users|
| __ CPUMXT      Max TSO Users|
|
\-----/

```

Figure 13-1.08. Resource Component Analysis Case Study Control Parameters

Note that this sample study is conducted using a "typical" days activity and the entire days usage will be selected as input. Two systems, SYS1 and SYS2, are incorporated in the study; together they are used for a variety of development and production purposes.

The data elements used to analyze the CPUTOBTM value are specified in the "Specify Two or More Independent Elements Below" list. For the first execution, we have included both the 'average' and 'maximum' values for each of the basic workloads: batch; TSO, STC and USS. This decision was made to determine if there were significant differences among these values and which set would be better used for the final analysis.

13.1.6.1.2 Relative Importance Analysis Report

The figures below show the Relative Importance Analysis Reports for this case study.

As this report shows, there is no significant difference between the 'average' and 'maximum' data element values and in both groups, BATCH (CPUxxB) is the least relevant. The other workloads, are about even in terms of affect on the CPUTOBTM value. The BATCH workload value was retained, and the 'maximum' data values were used in the second execution.

Resource Component Analysis Report							
Report For: 01/28/04							
Feature	Description	Minimum	Maximum	Average	Std. Dev.	CV	Rel. Imp.
CPUTOBTM	Total Time Processor Busy	0:08:15.73	0:27:00.70	0:19:23.79	0:03:23.89	17.52	
CPUAVB	Avg Batch Users	3.16	10.23	5.16	1.83	35.50	0.1465
CPUAVE	Avg OMVS Users	12.00	12.10	12.03	0.02	0.15	0.2229
CPUAVS	Avg Started Users	196.47	221.82	206.65	7.42	3.59	0.1482
CPUAVT	Avg TSO Users	2.04	16.72	9.96	4.15	41.71	0.1537
CPUMXB	Max Batch Users	5.00	17.00	8.76	2.85	32.50	0.0846
CPUMXE	Max OMVS Users	13.00	13.00	13.00	0.00	0.00	0.0000***
CPUMXS	Max Started Users	199.00	224.00	208.87	7.73	3.70	0.1126
CPUMXT	Max TSO Users	3.00	19.00	11.73	4.55	38.78	0.1315
Total Relative Importance Weights:							1.0000
*** - This Field(s) is 100% Correlated to the CPUTOBTM Value and was Dropped from Analysis.							

Figure 13-1.09. Relative Importance Analysis Report - first execution

For the second execution, only the 'average' (CPUAVx) data values was used. Not surprisingly, the approximate ratios of the Relative Importance Index value didn't change much between executions. In a real-life study, the actual processor times for the workloads would need to be considered as well as the simple number of users, and a detailed workload characterization study would be required before using the metrics for any forecasting operation.

Resource Component Analysis Report							
Report For: 01/27/04							
Feature	Description	Minimum	Maximum	Average	Std. Dev.	CV	Rel. Imp.
CPUTOBTM	Total Time Processor Busy	0:08:15.73	0:27:00.70	0:19:23.79	0:03:23.89	17.52	
CPUAVB	Avg Batch Users	3.16	10.23	5.16	1.83	35.50	0.2265
CPUAVE	Avg OMVS Users	12.00	12.10	12.03	0.02	0.15	0.2834
CPUAVS	Avg Started Users	196.47	221.82	206.65	7.42	3.59	0.2333
CPUAVT	Avg TSO Users	2.04	16.72	9.96	4.15	41.71	0.2568
Total Relative Importance Weights:							1.0000

Figure 13-1.10. Relative Importance Analysis Report - second execution

13.2 Analysis using CAP Database files

A new facility has been added to the CA MICS Capacity Planner that will analyze the effects of a series of numeric data values upon a client specified analysis element, and then quantify those effects. This is an exploratory data analysis tool and can be used in selecting data elements for later use in resource forecasting operations.

The sections that follow describe this resource forecasting facility.

- 1 - Functional Description
- 2 - Usage Guidelines
- 3 - Standard Output
- 4 - Analytic Technique Tutorial
- 5 - Component Operation
- 6 - Case Studies

13.2.1 Functional Description

Resource Component Analysis utilizes CA's U.S. patented Neugents technology facility known as Relative Importance. The Relative Importance facility utilizes proprietary analytical technology to understand and quantify the effect of a list of data values against the client specified analysis data element. While a similar analysis can be performed using standard statistical means such as correspondence matrix, the Relative Importance facility can not only determine linear relationships among data values, but non-linear relationships as well. This is important when exploring data that may not be well understood and can result in a more accurate understanding of the relationships among the data values.

As with other CA MICS Capacity Planner facilities, any files to be analyzed using the Resource Component Analyzer must reside in the capacity planning database. The analysis can be performed using either one or two database files, but both files must be summarized in the same timespan and maintain the same key structure.

Determine the data element to use for analysis and the CA MICS Capacity Planner will produce a default list of all quantitative data elements contained in the file defined under "Select Elements from." Edit this list as appropriate by either selecting or unselecting data elements before executing the application.

The application will validate all chosen independent elements for 'missing' values. Such values are not valid for performing Relative Importance and Neugents will ignore such values during processing. Input observations containing 'missing' values will be bypassed and listed on an error report.

Additionally, all training values will be analyzed to determine if they are totally co-related with the forecast value. For example, in a database built from HARCPU data, CPUTOBTM and CPUPCBSY would be totally co-related and as such, CPUPCBSY values would bias the Relative Importance process and cause the Importance Index (Weight) values to be distorted. The application will compute co-relation weights for all training elements and will cause any totally co-related values to be dropped from analysis and flagged on the Descriptive Statistics report.

During execution, the Relative Importance facility will analyze the values of each selected element from the "Select Elements from" list and develop a weight that quantifies the effect of the data value against the Analysis element. The sum of the weights for all selected elements will be 100 and will quantify the impact of each selected data element against the analysis variable. Those data values with a greater weight have a more significant impact against the analysis element than those with a lesser weight, and will likely be of more interest in any future forecasting operations.

13.2.2 Usage Guidelines

Usage of the Resource Component Analysis facility is easy and the component operation is straightforward. Since the facility creates a default list of data elements for later analysis, you can quickly execute a complete analysis with a single panel and a few key strokes. Once the analysis has completed, you can examine the output report produced, quickly determine those components that have minimal effect on the analysis element values, unselect those elements and reexecute the component analysis again. This is frequently an iterative process and can be executed as often as necessary until the best list of component elements remains. Once an optimum set of component elements has been determined, selected elements can be used as input to the workload forecasting process, discussed in Chapter 14.

13.2.3 Standard Output

The Resource Component Analyzer currently produces a single printed output: The Resource Component Analysis Report. This report contains a set of simple descriptive statistics for the Analysis element and each of the component definition values, as well the relative importance weight assigned to the individual components. A sample of this report is shown next.

Resource Component Analysis Report							
Report For: 01/27/04							
Feature	Description	Minimum	Maximum	Average	Std. Dev.	CV	Rel. Imp.
CPUADBTM	Adjusted Processor Busy Time	19:48:03.06	177:30:19.63	150:24:18.49	28:23:26.31	18.88	
BATCH	Batch Processor Usage Time	1:34:22.97	23:54:35.96	12:30:32.64	4:29:03.99	35.85	0.1261
DATABASE	Database Processor Usage Time	6:32:51.97	76:14:35.50	57:31:46.57	13:15:40.67	23.05	0.4012
ONLINE	Online Systems Processor Usage Time	4:03:04.13	39:19:11.42	33:07:25.75	6:37:24.04	20.00	0.3231
TSO	TSO Processor Usage Time	0:02:28.15	0:54:55.72	0:17:59.79	0:07:52.44	43.75	0.1496
Total Relative Importance Weights:							1.0000

Figure 13-2.01. Resource Component Analysis Report

The Resource Component Analysis report contains the following information:

- ELEMENT:** The name of the Analysis or Training element being analyzed.
- DESCRIPTION:** The SAS label of the selected data element, from either the CA MICS GENLIB definition or supplied by the user.
- MIN. VALUE:** The minimum value of the selected data element observed in the input data.
- MAX. VALUE:** The maximum value of the selected data element observed in the input data.
- MEAN VALUE:** The mean of the values of the selected data element observed in the input data.
- STD:** The standard deviation of the values of the selected data element observed in the input data.

CV:	The coefficient of variance of the values of the selected data element observed in the input data. This is a standard statistic and is a ratio of the mean and standard deviation measurements.
REL. IMP.:	The Relative Importance of this particular component as a function of the Analysis element value. Neugents technology computes this value internally using a mixture of both linear and nonlinear techniques and expresses this as a ratio or weight factor, relative to 100.

13.2.4 Analytical Tutorial

This section is not available at this time.

13.2.5 Component Operation

The Component Operation Section describes the control parameters that are used to execute the Resource Component Analysis inquiry.

These parameters allow you to specify which files and elements to analyze, making use of selective data deletion.

The Resource Component Analysis query is invoked entirely within a single screen, which is described in the next section.

13.2.6 Case Study

The following case study involves the use of Resource Component Analysis used in preparation of a processor resource forecast.

1 - Processor Resource Study

13.2.6.1 Processor Resource Usage

This case study shows how you can query a set of workloads and determine which workloads most significantly affect the processor resource usage.

In preparing for a capacity planning forecast, it is important for you to understand which workloads have the greatest impact on processor utilization. While you might want to simply look at workload CPU time, there may well be relationships among workloads that are not readily apparent. A better approach is to use Neugents technology Relative Importance, which is a nonlinear process and is therefore more aware of such relationships.

This case study assumes that two files exist in the capacity planning database, one built with the CPU Planning Standard application using the CPU Processor Activity (HARCPU) data as input, and another built as a Resource Element file. The Resource Element is derived from the Service Class Resource Consumption (WLMSEC) data. Both files in this sample reside in the CA MICS Hardware and CSP Analyzer (RMF) area and are summarized at the "Weeks" timespan. The WLM workload usage data has been mapped to a series of client defined workload names and the file contains one set of workload data for each observation.

DEVELOPMENT OF CONCEPTUAL STUDY

This study is being conducted as a prelude to performing a capacity study of a single system using 53 weeks of input data. The workload measurement data will be analyzed with respect to the processor resource usage in an effort to identify those workloads that are most relevant to future forecasting efforts. This study will be performed initially using all of the workloads, and then executed again after removing (unselecting) those workloads that appear to be less relevant.

APPLICATION OF CONCEPTUAL STUDY

The system being examined is SYS1 (this is the CA MICS SYSID value). This system processes a varied workload, which can be divided by WLM Service Class and summarized by Workload Name. The workloads are a TSO subsystem, a Database subsystem, an Online subsystem, a BATCH subsystem and a small USS based workload. There are also system-related tasks that must execute to provide an array of common services, such as GRS, JES2, VTAM, TCAS, and RMF.

Two files contain CPU time information concerning these workloads and processor use as a whole. These files are:

- APU - Processor Usage File
- WKL - Workload Usage File

In our example, the processor usage for each workload is analyzed against the total processor usage and the effect quantified into a weight factor.

The next sections explore the results of the report that is generated from the Resource Component Analysis we developed for workload usage study:

- 1 - Control Parameters
- 2 - Resource Component Analysis Report

13.2.6.1.1 Control Parameters

Using a short set of menus that are described in Section 1.5 of this guide, you will arrive at the Resource Component Analysis screen shown in Figure 13-2.02.

This figure shows the control parameters that we used to generate the Resource Component Analysis query in this case study.

```

/----- Resource Component Analysis Using Neugents -----\
|Command ==>|
|Enter a ? in any data entry field for more information on valid values.|
|      Analysis Inquiry:  APU  - Analysis CPU Usage|
|
|      Primary CAP file ==> APU - Adjusted Processor Usage File|
|      Analysis element ==> CPUADBTM|
|      Start date ==> _____ (ddmnyyyy)|
|      SYSID ==> SYS1|
|      Zone ==> 1|
|      No. wk/mn ==> 53      (1-9999)|
|      Specify CAPAPU Values ==> N      (Y/N)|
|      Select Elements from ==> WKL - Workload Usage File|
|
|Show ==> ALL      (Selected/All)  Line Cnds: S Select  U Unselect|
|
|Cmd  Name      Element Description      Selected ?|
|-----|
|Batch      Batch Workload Usage|
|Database    Database Regions Usage|
|OMVS        USS Application Usage|
|Online      Online Applications Usage|
|Other        Misc. Applications Usage|
|STC          All Started tasks Usage|
|SYSTEM      System task Usage|
|TSO          TSO Workload Usage|
|-----|
\-----/

```

Figure 13-2.02. Resource Component Analysis Case Study Control Parameters

Note that in this case study the analyst previously generated two capacity planning files. The first, APU, contains a data element, CPUADBTM, which represents the total systemwide CPU time, normalized to an IBM 9672-R95. This file and data element are specified on the control screen as the Primary CAP file and Analysis element respectively. The second file, WKL, contains the data elements that represent the CPU time consumed by each of the workloads running on the system. This file and the corresponding data elements are specified in the "Select Elements from" list and contain the elements for this case study. Note that the CPU times contained in the Workload definition file have been normalized as well.

The focus of this study is on a SYSID of SYS1 and a Zone of 1.

For the first execution of this case study, all of the elements (workloads) listed under "Select Elements from" will be selected by the analyst. Subsequent executions will be made with just those workloads that demonstrate a significant Relative Importance score.

13.2.6.1.2 Resource Component Analysis Report

The figures below show the Resource Component Analysis Reports for this case study.

As this report shows, the DATABASE and ONLINE workloads are the largest contributors to the overall processor usage. The next most relevant workloads are BATCH and TSO. Of these, TSO is interesting in that while it ranks fourth in terms of relative importance, its average CPU usage value is rather small, less than 18 minutes per week. Other workloads that are much higher in CPU usage rank less, perhaps indicating a relationship between the TSO workload and others, the DATABASE workload being the likely candidate. The analyst should do additional research regarding this result to better understand the overall affect TSO may have on the capacity planning effort.

Resource Component Analysis Report							
Report For: 01/27/04							
Feature	Description	Minimum	Maximum	Average	Std. Dev.	CV	Rel. Imp.
CPUADBTM	Adjusted Processor Busy Time	19:48:03.06	177:30:19.63	150:24:18.49	28:23:26.31	18.88	
BATCH	Batch Processor Usage Time	1:34:22.97	23:54:35.96	12:30:32.64	4:29:03.99	35.85	0.0965
DATABASE	Database Processor Usage Time	6:32:51.97	76:14:35.50	57:31:46.57	13:15:40.67	23.05	0.2836
OMVS	USS Processor Usage Time	0:03:04.02	2:15:20.94	0:42:05.17	0:35:42.67	84.85	0.0588
ONLINE	Online Systems Processor Usage Time	4:03:04.13	39:19:11.42	33:07:25.75	6:37:24.04	20.00	0.2648
OTHER	Misc. Processor Usage Time	0:00:00.37	0:08:29.81	0:00:45.17	0:01:35.03	210.39	0.0417
STC	Started Tasks Processor Usage Time	0:17:04.28	16:20:51.37	2:41:52.34	3:36:16.00	133.60	0.0732
SYSTEM	System Tasks Processor Usage Time	0:51:09.92	15:39:31.50	5:39:10.84	2:54:14.42	51.37	0.0867
TSO	TSO Processor Usage Time	0:02:28.15	0:54:55.72	0:17:59.79	0:07:52.44	43.75	0.0947
Total Relative Importance Weights:							1.0000

Figure 13-2.03 Resource Component Analysis Report - first execution

For a second execution, only the DATABASE, ONLINE, BATCH and TSO workloads are used. The report produced is shown next. Notice that the Importance weight values have changed as one would expect, but that the weight for the TSO workload has increased, indicating a stronger relationship with the other workloads. Again, additional analysis may be indicated before making any final conclusions regarding the workloads and their makeup.

Resource Component Analysis Report							
Report For: 01/27/04							
Feature	Description	Minimum	Maximum	Average	Std. Dev.	CV	Rel. Imp.
CPUADBTM	Adjusted Processor Busy Time	19:48:03.06	177:30:19.63	150:24:18.49	28:23:26.31	18.88	
BATCH	Batch Processor Usage Time	1:34:22.97	23:54:35.96	12:30:32.64	4:29:03.99	35.85	0.1261
DATABASE	Database Processor Usage Time	6:32:51.97	76:14:35.50	57:31:46.57	13:15:40.67	23.05	0.4012
ONLINE	Online Systems Processor Usage Time	4:03:04.13	39:19:11.42	33:07:25.75	6:37:24.04	20.00	0.3231
TSO	TSO Processor Usage Time	0:02:28.15	0:54:55.72	0:17:59.79	0:07:52.44	43.75	0.1496
Total Relative Importance Weights:							1.0000

Figure 13-2.04. Resource Component Analysis Report - second execution

Chapter 14: WORKLOAD FORECASTING

Workload Forecasting using CA's U.S. patented Neugents technology software allows you to develop nonlinear models that relate different resource consumption data elements from one or more capacity planning database files. This forecasting capability addresses the following requirements:

- o The investigation of possible relationships among different resource variables
- o The forecasting of future requirements of a resource variable based on estimates of the future requirements of one or more different resource variables.

The following sections describe how Workload Forecasting operates and the reports it produces.

This section contains the following topics:

- [14.1 Functional Description](#) (see page 931)
- [14.2 Usage Guidelines](#) (see page 935)
- [14.3 Standard Output](#) (see page 942)
- [14.4 Analytic Technique Tutorial](#) (see page 958)
- [14.5 Component Operation](#) (see page 958)
- [14.6 Case Studies](#) (see page 967)

14.1 Functional Description

Workload Forecasting using Neugents technology produces nonlinear models that relate different resource consumption data elements from one or more capacity planning database files. These forecasts are based on models generated using Neugents technology, and may use any of the resource element files produced by CA MICS Capacity Planner. You can use the same or different files as the source for the Forecast and Analysis variables in the model. The Forecast element is the data value being modeled and the Analysis elements are used to "train" the model before actual forecasting begins.

Upon execution, CA MICS Capacity Planner will read up the Primary CAP file, select those observations that meet the filtering criteria specified on panel CAPG700, and sort the data in preparation for modeling. Likewise, the Analysis file elements are first selected using panel CAPG710, observations are selected using the same filtering criteria used for the Primary CAP file, and validated to ensure that no missing or zero values exist. If invalid values are detected, a report is produced and execution is terminated. If the input data has been determined to be valid, a set of descriptive statistics of the input values is created and a report produced accordingly.

The modeling process uses a two-fold approach. First, a set of data elements is selected from the file specified under Select Elements from, and the training phase is executed. The training process enables the model to understand the relationship(s) between the Analysis element values and both the actual and predicted values of the Forecast element. Once the model is trained, a report is generated that contains statistics used to evaluate the output of the training phase. These include: the ANOVA R^2 value for the model, produced by Neugents technology itself; the Model Error value, which is the average value of the differences between the predicted and actual values within the forecast data and the Model RMS Error value, which is the average of the Model Error values, squared. Examination of these values will help the analyst determine how effective the model has been in predicting the Forecast element values.

Once the model has been trained, the forecasting phase is then executed. In Neugents technology parlance, forecasting is done by a process called "consulting." Basically, the trained model is "consulted" internally so that the relationships established between the Analysis (training) elements and the Forecast element can be used to determine future values of the Forecast element. The forecasting phase is actually executed in two parts. First, the values of the Analysis elements are extrapolated for the specified number of cycles. The extrapolation can use either a MANUAL or AUTOMATIC adjustment option specified by the analyst. If the MANUAL option is chosen, the analyst supplies an adjustment percentage for each selected Analysis element. If the AUTOMATIC option is chosen, CA MICS Capacity Planner examines the historical data values for each selected Analysis element and determines an appropriate adjustment factor. In either case, the adjustment factors are used to generate extrapolated Analysis element values, and in turn, these are used to predict new Forecast element values during modeling. The forecasted values can be optionally saved in a new file, as can be done with other forecasting methods.

The following reports and printer graphs are automatically generated by the Workload Forecast Analysis application:

- Model Input Descriptive Statistics
- Model Training Phase Summary
- Model Forecasting Phase Summary report
- Model Forecasting Phase Analysis graph
- Forecast Check report

An optionally requested report called the Model Forecasting Phase Analysis report can also be produced and used for a more detailed analysis of the model processing. The graphic output is produced as simple "printer graphics" for convenience, and can be produced as either high-resolution color graphics, or generated as a MQR compatible Tab delimited file suitable for downloading to a PC workstation and used for additional analysis and presentation.

Support is available for setting a threshold value alert and raising a condition if the threshold is exceeded. This information is shown in the Forecast Check report. The threshold percentage value you set is first applied against the baseline forecasted values of the data element chosen to generate the model. This then sets the upper and lower threshold limit values. Each observation selected for processing is compared against these limits. Those observations found to exceed the limits are flagged as exceptions and an alert is raised. The alerting process causes attention to those observations that may impact the validity or accuracy of the forecast so that the analyst can take appropriate action. The alerts are summarized on the MICSLOG and a report is produced detailing all selected observations and their alert status.

If the ServicePlus Service Desk product is installed and the MICS global Service Desk option is activated, you may additionally create a service desk alert. For more information on using the ServicePlus Service Desk alert facility, refer to section 4.9 of the CA MICS Resource Management Planning, Installation, Operation and Maintenance Guide.

To illustrate the use of Workload Forecasting using Neugents technology, we have developed a Workload forecasting model that estimates future processor usage based on workload consumption. This model relates the historical use of specified processor workload categories to overall processor use. A complete description and interpretation of the model is presented in Section 14.2.

The CA MICS Capacity Planner Option includes features that utilize CA's U.S. Patented Neugents(R) technology.

In order to exploit this technology the CA MICS Root Directory under Unix System Services needs to be initialized and the Neugents technology infrastructure enabled.

For detailed instructions on how to initialize the CA MICS root directory and enable Neugents technology, review and complete the steps presented in Chapter 12 of this guide.

14.2 Usage Guidelines

Workload Forecasting using Neugents technology is a technique for investigating the mathematical relationships between one or more resource variables contained in capacity planning database files. The mathematical relationships that may exist between variables with a high degree of relative importance can be expressed as models. You can then use these models to forecast future values of the resource variables.

There are three major steps in the development of a Neugents technology model:

- o Choosing model parameters
- o Evaluating the Training phase
- o Forecasting future values using the model

These topics are discussed after a brief review of Neugents technology models. This introduction presents some of the basic terminology used in the evaluation of Neugents technology models.

The following topics are presented by section:

- 1 - Neugents technology Models
- 2 - Choosing Model Parameters
- 3 - Evaluating the Model
- 4 - Forecasting Future Values with the Model

14.2.1 Neugents Technology Models

Neugents technology models utilize a nonlinear methodology to first train using a set of client specified data elements and then predict future values of the specified Forecast element. The techniques employed are internally complex, but presented in a simple, straight-forward manner, eliminating any knowledge requirement of those processes. This is in contrast to normal statistical methods, which often deliver satisfactory results, but generally require a considerable amount of knowledge on the part of the client regarding specific methods, measurements and output information. Neugents technology provides accurate forecasting while permitting the analyst to focus on the problem at hand, not the underlying methodologies.

Aside from simplified use, Neugents technology are nonlinear in nature, are generally less sensitive to "outliers" and other data variability issues, and will normally produce more accurate predictions using the types of data commonly found in capacity planning applications.

The Resource Component Analysis application can be used to evaluate the data elements being considered for training the model. Training element values that show a Relative Importance ranking (weight) value of less than 0.05 should probably be dropped from training usage. Of course, the analyst should always make such decisions based on his personal knowledge of the resources and workloads, as well as the business objectives of the data center.

14.2.2 Choosing Model Parameters

Choosing the model parameters for the Workload Forecasting application using Neugents technology is reasonably simple. Aside from the Primary CAP file and Analysis element specifications, the CA MICS Capacity Planner software provides a set of default values that will permit a valid model to be obtained with minimal effort.

Building a Neugents technology model can be and often is an iterative process. The model is always trained first and a forecast produced according to the Forecast Length specification. Descriptive statistics of the input data, and model performance statistics from both the training and forecasting phases are produced. Additionally, a graph depicting the results of the training and forecasting phases is generated.

Using the default Forecasting Method of AUTOMATIC will cause a model to be based on the historical growth of the values of the Training elements, and should provide a good starting point for capacity planning. Once better information concerning growth rates is available, then the Forecasting Method can be set to MANUAL and specific adjustment factors used for each selected workload.

When specifying a Forecast Length value, use something reasonable that relates to the business requirements. Neugents technology is quite adept at forecasting, but any forecast begins to lose validity when the period is exaggerated. Depending on the quantity and granularity of the input data, a forecast period of 90-180 days is probably a good place to start. Save your forecasts and compare them to the actual data when it becomes available. As with any forecasting technique, the validity of the forecast is a function of length of time covered by the input history data. For example, we would never suggest using six months of monthly summarized data and then attempting to forecast 18 months into the future. The Neugents technology model would produce a forecast, but it would be of no practical value.

14.2.3 Evaluating the Model

The second step in developing a Workload Forecasting Analysis model is evaluating how well the data fits the model. For this purpose, you can use the model summary report produced after the training phase. The statistics are the ANOVA R-squared, Average Model Error and the Average Model RMS Error.

The statistics produced by the model are interpreted as follows:

- o ANOVA R-squared is the coefficient of determination and termed the "goodness of fit". This value indicates what percent of the variability in the historical (training) data is explained by the proposed model. You would typically be seeking r-squared values greater than or equal to 0.70.
- o Average Model Error is the mean value of the differences between the predicted and actual data values occurring within the Forecast data and observed by the model, expressed as a ratio of the Forecast data values. A value of less than .10 should be considered acceptable.
- o Average RMS Error is the mean value of squared differences between the predicted and actual data values occurring within the Forecast data and observed by the model. This value is expressed in the same units as the Forecast element value.

As with statistical modeling methods, the analyst should review the Neugents technology modeling output statistics to determine how effective the model trained and how accurately it represents the historical data values before proceeding with any forecasting operations.

The following steps should be taken to ensure model accuracy and provide confidence in modeled results.

Step 1 - Examine the ANOVA R-squared, Model Error and RMS Model Error statistics.

We recommend an ANOVA R-squared value greater than 0.70. This value is an estimate of how many of the actual Forecast element values were accurately predicted during the training phase. Thus, an ANOVA value of .70 indicates that 70 percent of the Forecast element values were accurately predicted by the model. A value less than .70 would indicate that the input data for the training phase did not properly define the forecasting problem and the modeling data needs to be revised. A well defined model should produce an ANOVA R-squared value of at least .85, with higher values typically observed.

The Model Error calculated by the Workload Forecasting application is expressed as a ratio, and as such should be less than .10. Many models can produce model error measurements of less than .05, so an error value of more than .10 would likely indicate that improper choices were made when selecting training elements and that the model requires revision.

The RMS Error is the square root of the mean of the squared actual model errors. While no specific recommendation is currently available, the value here should be "small" with respect to the Forecast element values and should track the average model error.

The ANOVA R-squared, Model Error and RMS Error values are printed on the Model Error Summary Report.

Step 2 - Visually examine the data.

Visual examination of the actual and predicted data is very important and can sometimes show that a model whose ANOVA R-square, or other statistics look bad may be significantly improved by the exclusion of a small number of outliers from the data. While models built using Neugents technology are generally less affected by "outliers" and other "noise" in the input data, they are not completely immune, and the Workload Forecasting Analysis application generates both descriptive statistics as well as reports and graphs to help evaluate model performance and to identify problem data values.

Step 3 - Apply the common sense test.

The common sense test asks, "Does this model make any sense?" To execute and depend on various forecasting models you should understand the data well. For example, you can load up a year's worth of resource usage data, construct and train a model, check the statistics and decide the model is valid and can be used for forecasting, when actually it is not valid. If you are using weekly data for planning purposes and you are a bank that does a lot of credit card processing, examining the data visualization graphs and reports might erroneously cause you to decide the fourth week of November is an outlier. In fact, it is probably the single most important week of the year. The reason is that Thanksgiving week and the Friday after Thanksgiving is typically the single largest shopping day of the year, with credit cards melting down all over the country. The obvious motto is: know your business and know your data.

Any forecasting technique needs validation, and Neugents technology is no different. Execute your models and save the forecasts. Once the new planning data is available, compare your previous forecasts to the actual planning data. Are you tracking well? If not, investigate and find out why. You may need to include more workloads into your model, or switch to the MANUAL option and "grow" your workloads more aggressively. Or you may need to get creative and define a special workload that represents planned new business and use this "dummy" workload to document the need for additional capacity. A good deal of an analysts work is in understanding the business and ensuring that his plans properly address these needs. Tools such as Neugents technology can help create better forecasts, but no tool can substitute for the knowledge the analyst possesses.

14.2.4 Forecasting Future Values with the Model

The final step in developing the Workload Forecasting Analysis model is to use the model to predict future values of the Forecast resource variable. This is easily accomplished using the Workload Forecasting Analysis model along with Neugents technology.

Predicting future values of the Forecast element value requires the analyst to first extrapolate the values of the selected Analysis element values and then to use those values as input to the model prediction process. Workload Forecasting Analysis provides assistance in this process with internal options for extrapolating the Analysis data values, so that you do not need to rely on external means. You may choose to adjust the Analysis data values using either the MANUAL option, where an adjustment value (as a percentage) is applied against each selected Analysis element (workload), or the AUTOMATIC option where the application examines the Analysis element values to determine the growth rate of each workload, and then assigns each respective growth rate internally. Once the growth rate has been assigned, the Workload Forecasting Analysis application extrapolates the data values for each selected Analysis element for the number of cycles specified in the Forecast Length. Neugents technology is then invoked to predict future values of the Forecast variable using the newly extrapolated data values as input. The output of the forecasting execution is a file containing the original input values, including the Forecast and selected Analysis element values, and all output values. This file is used to generate a variety of reports and graphs.

One consideration in any forecasting operation is "where to start". That is to say, what point in time is appropriate to use as a starting data value for the forecasting operation. Conventional wisdom says to use the last known set of Analysis element values (workloads) to begin forecasting, but this can result in misleading results. Suppose, for example that the most recent set of data points occurred in a time period of reduced usage. Starting forecasting from that point might well result in a forecast that is too low in magnitude. The opposite could certainly occur if the starting set of data points were higher than what normally occurred in operation of the system. To address that issue, the Workload Forecasting Analysis application first reviews the historical data values for each of the Analysis elements independently, determines an adjustment factor for each value and creates a starting set of data points from the adjusted set. Note that while this set of data points may well mimic a real point in time, it is only a base line estimate but will generally provide better results when the input history data is not well behaved.

The newly predicted values of the Forecast element, along with the extrapolated values of the Analysis elements can be optionally saved in a CA MICS Capacity Planner Forecast file for additional analysis or presentation purposes.

14.3 Standard Output

Output from Workload Forecasting Analysis consists of printed reports and/or graphs, color graphic output, an MQR compatible Tab delimited file suitable for downloading to a PC workstation, and, optionally, data values stored in a CA MICS Capacity Planner Forecast file. These are covered in more detail in the following sections:

- 1 - Update to the Forecast File
- 2 - Standard Reports and Graphs

14.3.1 Update to the Forecast File

You can use the Save forecast option on the Workload Forecasting Analysis control screen shown below to save predicted values of the Analysis variable in the Capacity Planner Forecast file.

In the Forecast file, future values are stored in a special way that is used by other functions of the CA MICS Capacity Planner to produce comparative graphs and charts. The FILE_NO variable is used to distinguish between the various types of observations within Forecast files. The Workload Forecasting Analysis, Multivariate Regression, Univariate Model Forecasting, and Business Element Forecasting facilities store a future value with a FILE_NO of 11, and a confidence range with a FILE_NO value of 12. In addition, each observation of a forecast contains the variables FORYEAR and FORMNDY, FORWEEK or FORMONTH, which contain the dates for the most recent historical data used in making the forecast. These concepts are discussed in more detail in Section 3.2.

```

/----- Workload Forecasting Analysis -----\
|Command ==>
|Enter a ? in any data entry field for more information on valid values.
|      CA MICS Inquiry: CPW - Sample Workload Forecasting Query
|Report title ==> Processor Usage Forecast
|
|Model Data Selection:
|Primary CAP file ==> APU          Forecast length ==> 13
|Forecast element ==> CPUADBTM   Forecasting Method ==> AUTOMATIC (Manual/Auto
|      Start date ==> 01JUN200   Adjustment Option ==> INTERVAL (Int./Annual
|      SYSID ==> SYSB           Specify elements ==> YES      (YES/NO)
|      ZONE ==> 1               Select Elements from ==> APU   (CAP file)
|      No. Weeks ==> 53          Save forecast ==> YES      (YES/NO/AGE/CHK)
|      CAPAPU Values ==> NO      Alerting Threshold ==> NO   (00-99 Percent)
|      Activate Alerting ==> NO  (YES/NO)
|
|-----\
|Output Generation Options:
|Forecasting Report      ==> YES      Color ==> NO   CSV ==> NO
|Regression Residuals Report ==> NO      Color ==> NO   CSV ==> NO
|      Color graphics format set ==> TERMINAL
|Output Formatting Options:
|Format for Predicted Values ==> HOUR
|HAXIS Interval Values      ==> QTR      Format ==> MONYY
|-----\

```

14.3.2 Standard Reports and Graphs

Workload Forecasting Analysis produces a series of printer reports and graphs, color graphics, as well as CA MICS Q&R WorkStation compatible Tab delimited files which are discussed in the following sections:

- 1 - Descriptive Statistics
- 2 - Training Phase Summary
- 3 - Forecasting Phase Report
- 4 - Forecasting Phase Graph
- 5 - Regression Model Output Graph
- 6 - Regression Model Residual Graph
- 7 - Forecast Check Report

14.3.2.1 Model Input Descriptive Statistics Report

Figure 14-1 shows a sample of the Input Descriptive Statistics Report. These statistics can be used to gauge the validity and robustness of the input data values.

Model Input Descriptive Statistics						
Report For: 01/28/04						
Feature	Description	Minimum	Maximum	Average	Std. Dev.	CV
CPUTOBTM	Total Time Processor Busy	7:38:18.77	56:58:16.02	42:54:27.52	9:15:55.30	21.59
CPUAVB	Avg Batch Users	4.84	43.81	25.30	13.77	54.42
CPUAVS	Avg Started Users	119.27	199.50	152.56	29.11	19.08
CPUAVT	Avg Tso Users	9.76	16.20	12.81	1.61	12.56

Figure 14-1. Model Input Descriptive Statistics Report

The Model Input Descriptive Statistics report contains the following information:

- ELEMENT:** The name of the Forecast or training element being analyzed.
- DESCRIPTION:** The SAS label of the selected data element, from either the CA MICS GENLIB definition or supplied by the user.
- DESCRIPTION:** The SAS label of the selected data element, from either the CA MICS GENLIB definition or supplied by the user.
- MIN. VALUE:** The minimum value of the selected data element observed in the input data.
- MAX. VALUE:** The maximum value of the selected data element observed in the input data.
- MEAN VALUE:** The mean of the values of the selected data element observed in the input data.
- STD:** The standard deviation of the values of the selected data element observed in the input data.
- CV:** The coefficient of variation. The CV is calculated by dividing the standard deviation by the average.

14.3.2.2 Training Phase Statistics Summary Report

This report displays the model statistics from the training phase and is used to evaluate the success and performance of the model training.

SUMMARY OF NEUGENT'S TRAINING PHASE
NUMBER OF CASES USED FOR TRAINING: 000040
AVERAGE MODEL ERROR MEASURED: 0.0092
AVERAGE MODEL RMS ERROR MEASURED: 0.0001
'GOODNESS OF FIT' FOR MODEL: 0.9960

Figure 14-2. Training Phase Statistics Summary Report

The Model Training Phase Summary report contains the following information:

NUMBER OF CASES USED FOR TRAINING:	The number of observations (cases) actually selected by the model for training purposes.
AVERAGE MODEL ERROR:	The mean of all error values detected during the Training phase, expressed as a percentage of the Analysis element values. Error values are the difference between the predicted and original Analysis element values.
AVERAGE MODEL RMS:	The Root-Mean-Square value of the total model error. This is an estimate of the standard deviation of model errors and serves as an estimate of the distribution of error values.
GOODNESS OF FIT:	The ANOVA R**2 value produced by Neugents Technology that represents the percentage of predicted values addressed by the model. This is a general measure of how effective the model is performing.

14.3.2.3 Forecasting Phase Analysis Report

This report displays all of the observations from the Forecasting file, including the training, the actual and predicted Forecast data values, along with individual delta values and model errors. Additionally, the predicted future values of the Forecast element is displayed. The model statistical summary is displayed on each page of the report for reference. This report can be used to document the model for a better understanding of the various data values.

NEUGENTS MODEL FORECASTING PHASE ANALYSIS								
REPORT OF FORECASTED CPUADBTM VALUES								
'GOODNESS OF FIT' FOR MODEL: 0.9960								
DATE	BATCH	DATABASE	ONLINE	TSO	ORIGINAL INPUT VALUE	PREDICTED VALUE	DELTA VALUE	MODEL ERROR
15SEP2001	2.67	45.54	29.12	0.23	144.27	142.87	1.40	0.0097
22SEP2001	2.75	41.29	31.21	0.35	136.17	137.78	1.62	0.0119
29SEP2001	3.06	53.18	34.63	0.39	162.61	162.11	0.49	0.0030
06OCT2001	7.93	63.50	31.33	0.38	159.36	158.31	1.04	0.0065
13OCT2001	13.99	71.30	32.69	0.29	165.47	162.30	3.17	0.0191
20OCT2001	12.99	66.66	34.37	0.32	160.39	165.71	5.32	0.0331
27OCT2001	13.88	52.05	29.12	0.32	133.68	134.29	0.61	0.0045
03NOV2001	14.19	70.01	35.56	0.30	167.47	169.48	2.02	0.0120
10NOV2001	14.69	50.32	32.00	0.29	141.49	142.92	1.42	0.0101
17NOV2001	13.91	51.62	35.40	0.34	147.46	146.03	1.43	0.0097
24NOV2001	15.50	38.81	25.59	0.23	115.67	115.47	0.20	0.0017
01DEC2001	17.20	58.84	37.84	0.30	163.94	163.48	0.46	0.0028
08DEC2001	13.75	60.88	36.40	0.24	159.86	158.93	0.94	0.0059
15DEC2001	15.15	60.09	37.31	0.31	163.05	161.10	1.95	0.0120
22DEC2001	17.44	56.54	33.22	0.18	154.92	154.17	0.75	0.0048
29DEC2001	10.33	31.49	16.82	0.14	86.25	86.11	0.14	0.0016
31DEC2001	1.57	6.55	4.05	0.04	19.80	19.94	0.14	0.0069
05JAN2002	11.88	34.93	21.15	0.17	99.09	99.41	0.32	0.0032
12JAN2002	16.31	60.84	37.65	0.38	167.67	166.42	1.25	0.0075
19JAN2002	15.24	61.11	38.68	0.20	166.61	166.93	0.32	0.0019
26JAN2002	15.85	55.85	37.19	0.29	157.22	159.74	2.52	0.0161
02FEB2002	14.77	63.80	39.32	0.36	169.00	172.13	3.13	0.0185
09FEB2002	11.24	60.06	34.22	0.35	152.11	154.39	2.28	0.0150
16FEB2002	12.15	59.74	36.37	0.23	158.76	160.74	1.98	0.0125
23FEB2002	10.47	53.77	33.49	0.20	142.05	143.95	1.91	0.0134
02MAR2002	9.15	57.73	36.66	0.31	150.36	148.77	1.59	0.0106
09MAR2002	12.67	61.95	39.14	0.34	165.46	164.99	0.47	0.0028
16MAR2002	11.95	67.60	39.30	0.38	173.09	170.94	2.15	0.0124
23MAR2002	9.11	59.23	37.02	0.55	154.08	155.29	1.21	0.0078
30MAR2002	13.53	61.93	37.49	0.31	163.67	160.67	3.00	0.0183
06APR2002	19.47	75.42	34.47	0.21	177.51	178.54	1.03	0.0058
13APR2002	9.74	76.24	34.57	0.23	164.34	165.44	1.10	0.0067
20APR2002	12.89	75.06	36.22	0.25	168.64	165.82	2.82	0.0167
27APR2002	10.31	74.04	37.12	0.25	166.66	167.80	1.14	0.0068
04MAY2002	14.57	55.43	29.77	0.25	142.77	141.59	1.17	0.0082
11MAY2002	12.67	64.10	31.22	0.36	155.23	154.30	0.93	0.0060
18MAY2002	15.78	55.85	35.58	0.28	157.35	157.38	0.03	0.0002
25MAY2002	11.24	55.38	33.01	0.27	145.13	142.69	2.44	0.0168
01JUN2002	14.46	66.30	33.47	0.24	160.80	161.15	0.35	0.0022
08JUN2002	23.91	66.17	35.16	0.92	176.78	176.05	0.72	0.0041
15JUN2002	15.35	65.61	35.57	0.34		167.29		
22JUN2002	15.49	66.01	35.69	0.34		168.18		
29JUN2002	15.63	66.41	35.82	0.34		168.87		
06JUL2002	15.78	66.81	35.94	0.35		169.34		
13JUL2002	15.92	67.21	36.06	0.35		169.55		
20JUL2002	16.06	67.62	36.18	0.35		169.51		
27JUL2002	16.20	68.02	36.30	0.35		169.22		
03AUG2002	16.34	68.42	36.43	0.35		168.67		
10AUG2002	16.48	68.82	36.55	0.36		167.88		
LEGEND: ACTUAL VALUES (CPUADBTM) - 'A'								
PREDICTED VALUES - 'P'								
NEUGENTS MODEL FORECASTING PHASE ANALYSIS								
REPORT OF FORECASTED CPUADBTM VALUES								
'GOODNESS OF FIT' FOR MODEL: 0.9960								
DATE	BATCH	DATABASE	ONLINE	TSO	ORIGINAL INPUT VALUE	PREDICTED VALUE	DELTA VALUE	MODEL ERROR
7AUG2002	16.62	69.22	36.67	0.36		166.88		
4AUG2002	16.77	69.63	36.79	0.36		165.69		
1AUG2002	16.91	70.03	36.91	0.36		164.36		

7SEP2002	17.05	70.43	37.04	0.36	162.94

Figure 14-5. Forecasting Phase Analysis Report

The Model Forecasting Phase Analysis Report contains the following information:

GOODNESS OF FIT:	The ANOVA R**2 value produced by Neugents technology that represents the percentage of predicted values addressed by the model. This is a general measure of how effective the model is performing.
DATE:	The ending date of the observations. For example, if the capacity planning file used for forecasting is summarized at the WEEKS timespan, then the DATE value will be the date at the end of the week's data being processed.
Training Elements:	The next "n" fields are the training elements specified by the client.
ORIGINAL INPUT VALUE:	The value of the Forecast element for the observation being reported.
PREDICTED VALUE:	The value of the Forecast element as predicted by the model.
DELTA VALUE:	The difference between the model predicted and original (Forecast) element values.
MODEL ERROR:	The absolute DELTA value expressed as a percentage of the original Forecast element value.

14.3.2.4 Forecasting Phase Analysis Graph

This graph displays all of the observations from the Forecasting file, including the actual and predicted values of the Forecast element. The model "goodness of fit" statistic is displayed on each page of the report for reference. This graph can be used to evaluate the success of the model forecasting phase and to demonstrate the forecasted values of the Forecast data element.

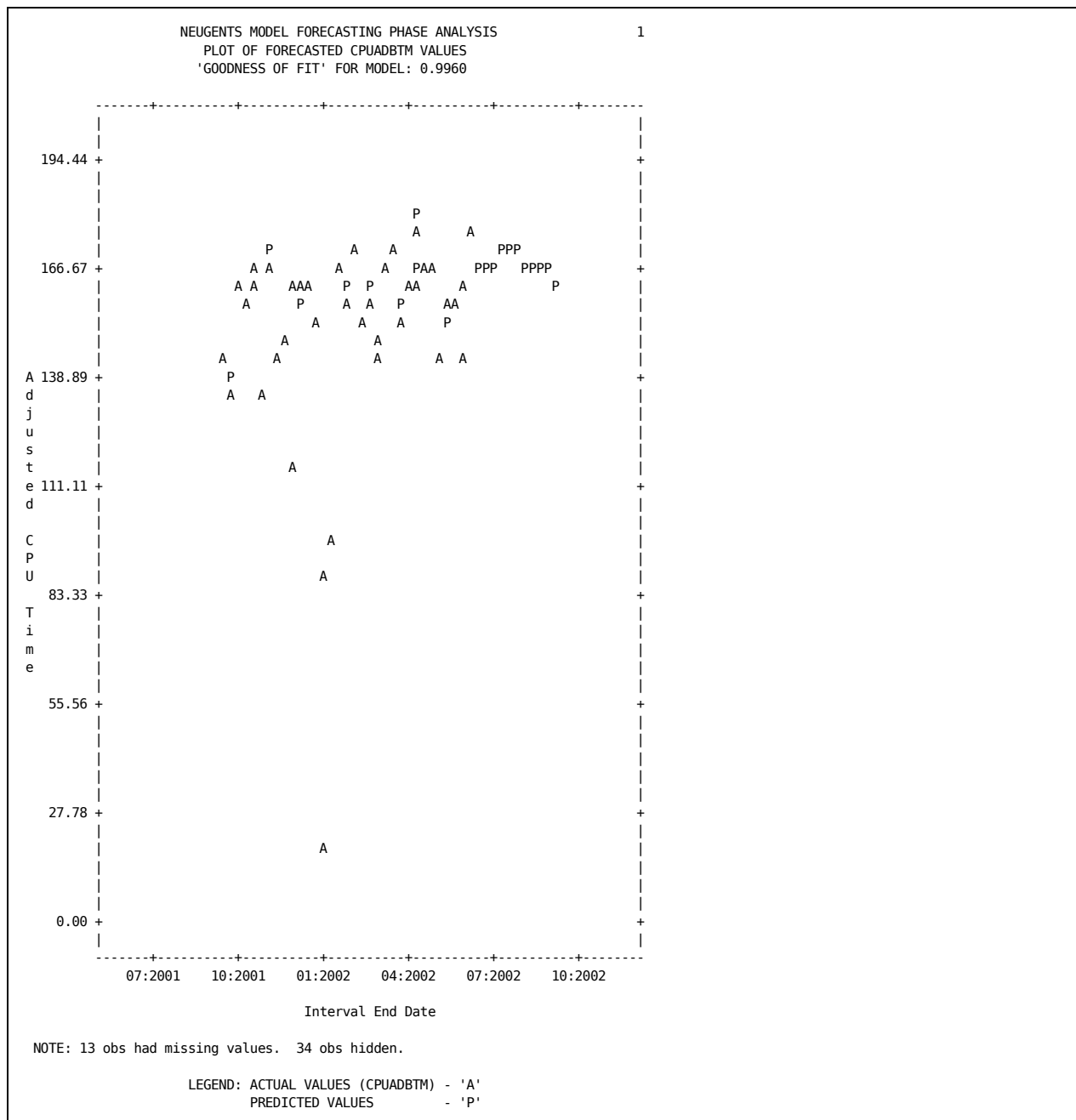


Figure 14-6. Forecasting Phase Analysis Graph

This graph provides an easy visual evaluation of model effectiveness by displaying the actual and predicted values of the Forecast element. This display, along with the model "goodness of fit" statistic repeated here, allow the analyst to make a quick and accurate assessment of model performance during Forecasting and determine whether the model should be considered as trustworthy. Future Forecast element values predicted by the model actually can increase and decrease over time, owing to the non-linear methodology employed by Neugents technology. This is a strength of this particular approach and provides for more accurate results.

14.3.2.5 Regression Model Output Graph

This graph displays all of the observations from the Forecast file, including the actual and predicted values of the Forecast element. The linear regression statistics are displayed on each page of the report for reference. This graph can be used to evaluate the success of the model Forecast phase and to demonstrate the predicted values of the Forecast data element.

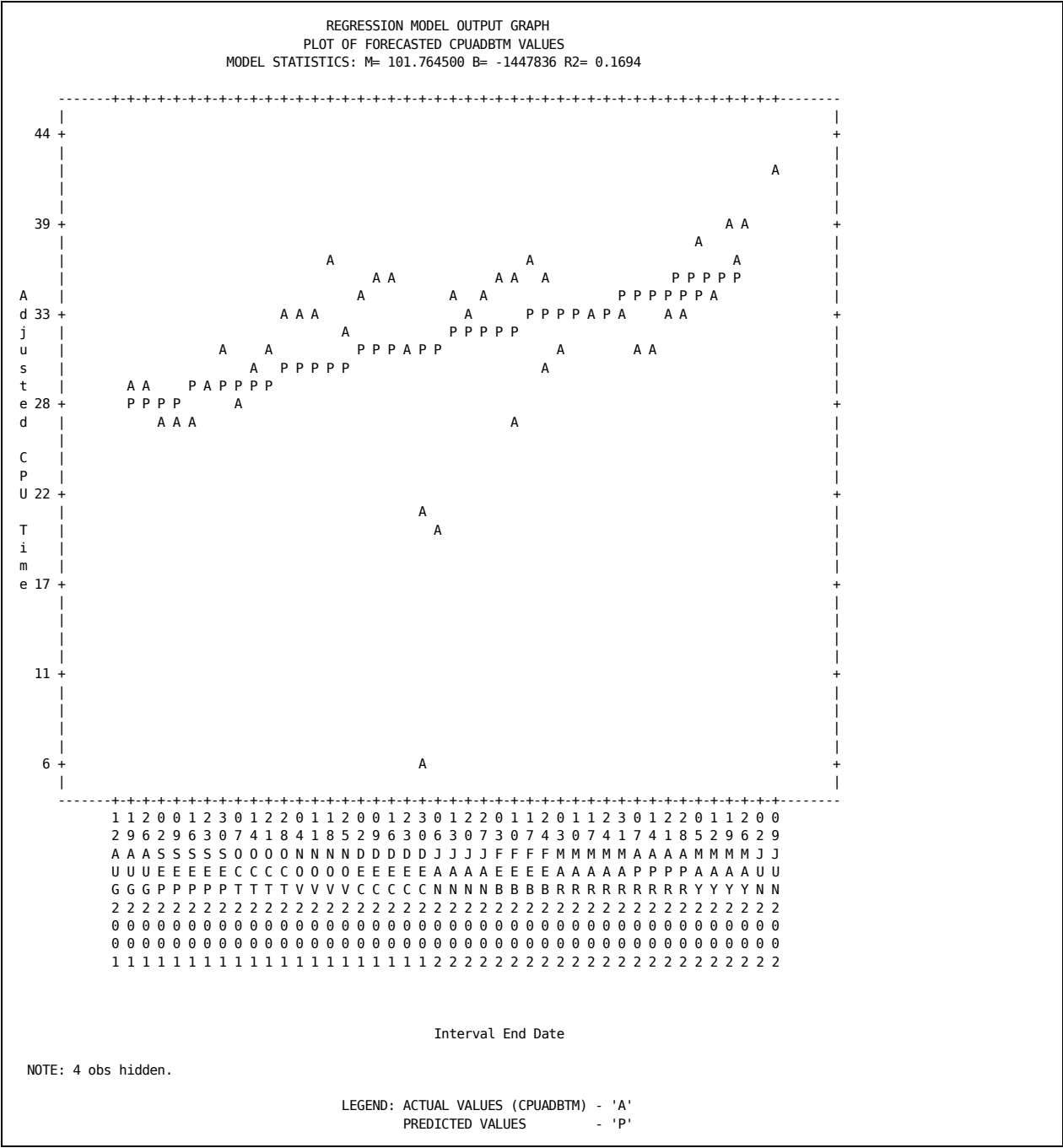


Figure 14-7. Regression Model Output Graph

This graph provides an easy visual evaluation of model effectiveness by displaying the actual and predicted values of the Forecast element. This display, along with the regression model statistics repeated here allow the analyst to make a quick and accurate assessment of model performance during prediction and to then determine whether the model is representative of the input data. Simple linear regression techniques are used primarily as a data exploration tool and have limited capability in capacity planning applications, as evidenced by the extremely poor R^2 value in the above example.

14.3.2.6 Regression Model Residuals Graph

This graph displays residual values from the regression model operation. Residual values are differences observed between the actual and predicted values of the dependent (Forecast) variable. The linear regression statistics are displayed on each page of the report for reference. This graph can be used to help evaluate the success of the model by ensuring that the model is functioning properly.

REGRESSION MODEL RESIDUALS GRAPH PLOT RESIDUAL VS ACTUAL VALUES	1

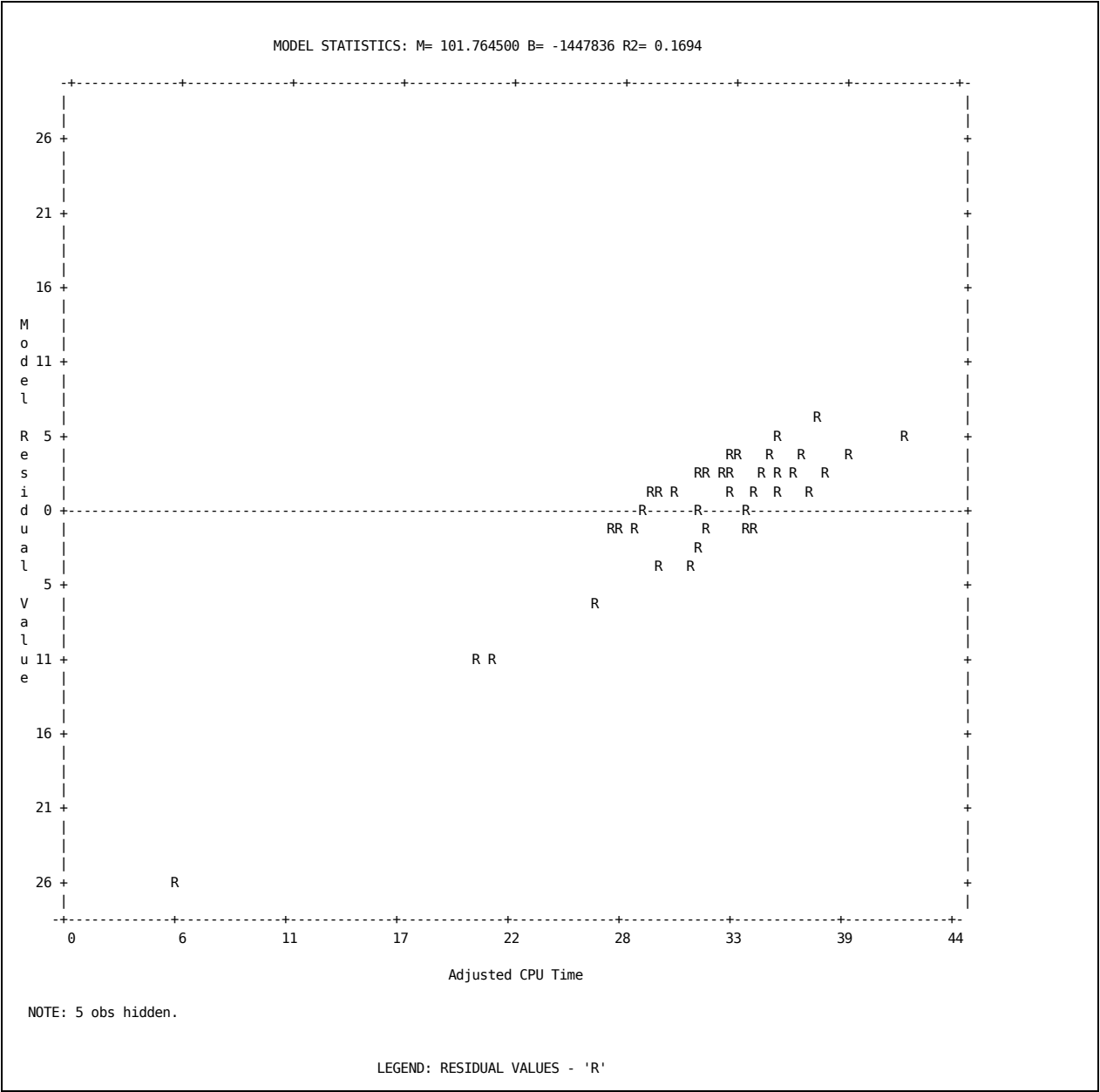


Figure 14-8. Regression Model Residuals Graph

Anytime a linear regression model is used, the residual values should be plotted against the actual values of the dependent (Forecast) variable and examined for statistical abnormalities. The residual values should cluster around zero and should be present in a somewhat "random" pattern. Residuals that track the original data values or form other specific patterns can indicate problems within the model. Consult any basic statistics book for additional details.

14.3.2.7 Forecasting Check Report

This report is produced only when the Forecasting Alerting option is chosen. The report displays those observations that occur in common between the Baseline Forecast (original) file, the updated Actual (observed values) file, and the Current (temporary) Forecast file, as ordered by the Sysid and time span values. Additionally, the upper and lower alerting limits used to generate alerts and the percentage of variation observed from the Forecast Baseline Values.

This report is used to understand what alerts were raised during the forecast operation, if any, and which limits were exceeded. Appropriate MICSLOG messages will be issued to summarize the alerting operation as well.

Analysis of Neugents Forecasting Model: Forecast Check
 Forecast Data Verification for: CPUADBTM - Adjusted CPU Time
 Actual File Input Date Range: 04DEC2004 - 10SEP2005
 Baseline Forecast Date Range: 13AUG2005 - 15OCT2005
 Forecast Level: WEEK Forecast Length: 10 Forecast Threshold: 10 %
 Observations Selected from Sysid: PROD for HOUR 22

Date	Actual Observed Values	Baseline Forecasted Values	Current Forecasted Values	Lower Alert Limit	Upper Alert Limit	Percent From Baseline	ALERT Flag
13AUG2005	70:21:05.92	96:04:01.54		86:27:37.39	105:40:25.69	-26.77	<====
20AUG2005	101:49:03.12	97:30:39.02		87:45:35.12	107:15:42.93	4.42	
27AUG2005	123:22:45.59	98:57:35.50		89:03:49.95	108:51:21.05	24.68	<====
03SEP2005	107:35:46.14	100:24:29.04		90:22:02.14	110:26:55.95	7.16	
10SEP2005	85:20:15.03	101:50:55.80		91:39:50.22	112:02:01.38	-16.21	<====
17SEP2005		103:16:30.34	72:20:35.55	92:56:51.31	113:36:09.37	-29.95	<====
24SEP2005		104:40:46.12	103:21:33.33	94:12:41.50	115:08:50.73	-1.26	
01OCT2005		106:03:15.96	134:24:46.87	95:26:56.36	116:39:35.55	26.74	<====
08OCT2005		107:23:32.64	105:30:18.81	96:39:11.38	118:07:53.91	-1.76	
15OCT2005		108:41:09.54	106:38:11.24	97:49:02.59	119:33:16.50	-1.89	

Figure 14-8A Forecasting Check Report

The Forecast Check Report contains the following information:

DATE:	The ending date of the observations. For example, if the capacity planning file used for forecasting is summarized at the WEEKS timespan, then the DATE value will be the date at the end of the week's data being processed.
Actual Observed Values:	The values of the data element used for forecasting and found in the CAP Planning file used as input for the Neugents Forecasting Model query.
Baseline Forecasted Values:	The values of the data element used for forecasting and found in the Forecast file created when during the Neugents Forecasting Model query. For the Forecast Check operation to complete successfully, a standard forecast query must have been previously executed and the resultant Forecast file saved.
Current Forecasted Values:	The values of the data element used for forecasting and found in the temporary Forecast file created during the Neugents Forecasting Check query.
Lower Alert Limit:	The lower threshold limit used for alert processing. This value is derived from the Baseline Forecasted Values, minus the Forecast Threshold percentage. Any Actual Observed or Current Forecasted values that are less than this value raise an alert condition.
Upper Alert Limit:	The upper threshold limit used for alert processing. This value is derived from the Baseline Forecasted Values, plus the Forecast Threshold percentage. Any Actual Observed or Current Forecasted values that are greater than this value raise an alert condition.
Percent From Baseline:	The variation of either then Actual Observed Values or the Current Forecasted Values from the Baseline Forecasted Values, expressed as a percentage.
ALERT Flag:	An indicator, if present which flags a given observation as having a data value that caused an alert condition to be raised.

14.4 Analytic Technique Tutorial

Not available at this time.

14.5 Component Operation

This section describes the control parameters that are used to execute the Workload Forecasting Analysis inquiry.

These parameters allow you to specify the files and elements to be analyzed.

The Workload Forecast Analysis Query is invoked using either one or two screens, depending on the options selected. The screens are described in the following sections:

- 1 - Workload Forecasting Analysis Control Screen
- 2 - Analysis File Data Elements Selection Screen

14.5.1 Workload Forecasting Analysis Control Screen

The Workload Forecasting Analysis control parameters allow you to develop forecast models based on Neugents technology. You can optionally save the forecasts produced by the program in a Capacity Planning Forecast file for additional analysis or presentation.

Using a short set of menus that are described in Section 1.5 of this guide, takes you to the Workload Forecasting Analysis Control Screen shown in Figure 14-9. This screen appears after you indicate that you want to insert a new Workload Forecasting Analysis Query. Two fields on this screen are filled in on the third line when you first see it: the inquiry name and the inquiry description, separated by a dash. You entered these two fields on the Insert Capacity Planning Analysis screen when you requested that a new inquiry be added.

```

/----- Workload Forecasting Analysis -----\
|Command ==>
|Enter a ? in any data entry field for more information on valid values.
|      CA MICS Inquiry: CPW - Sample Workload Forecasting Query
|Report title ==> Processor Usage Forecast
|
|Model Data Selection:
|Primary CAP file ==> APU      Forecast length ==> 13
|Forecast element ==> CPUADBTM Forecasting Method ==> AUTOMATIC (Manual/Auto
|      Start date ==> 01JUN200 Adjustment Option ==> INTERVAL (Int./Annual
|      SYSID ==> SYSB      Specify elements ==> YES      (YES/NO)
|      ZONE ==> 1      Select Elements from ==> APU      (CAP file)
|      No. Weeks ==> 53      Save forecast ==> CHK      (YES/NO/AGE/CHK)
|      CAPAPU Values ==> NO      Alerting Threshold ==> NO      (00-99 Percent)
|      Activate Alerting ==> NO      (YES/NO)
|
|-----\
|Output Generation Options:
|Forecasting Report      ==> YES      Color ==> NO      CSV ==> NO
|Regression Residuals Report ==> NO      Color ==> NO      CSV ==> NO
|      Color graphics format set ==> TERMINAL
|Output Formatting Options:
|Format for Predicted Values ==> HOUR
|HAXIS Interval Values      ==> QTR      Format ==> MONYY
|-----\

```

Figure 14-9. Workload Forecasting Analysis Control Screen

Some of the parameters on the Workload Forecasting Analysis screen shown in Figure 14-9 are required, while others are optional. The following is a discussion of the various parameters as they appear in this figure.

PARAMETER:	DESCRIPTION / USAGE:
Report title:	The 60-character title for Workload Forecasting Analysis using Neugents technology reports. No double quotes (") or semicolons (;) may appear within the body of the title.

Primary CAP file:	The capacity planning database file containing the historical observations for the forecast variable of the model. This file may be either a Sample Application or user defined Resource Element file, but must reside in the Capacity Planning database. If you are not sure of the files or file names available in your capacity planning database, type a question mark (?) in the field and a selection list will be displayed.
Forecast element:	The name of the Forecast variable in the model. If you are not sure of the correct name of the data elements in the input file, type a question mark (?) in this field and a selection list of available data elements will be
Start date:	The start date for date range selection. If Start date is specified, only historical observations at least as recent as this value will be included in the analysis.
SYSID:	The SYSID for which the report is to be produced. Note that the Secondary CAP file (specified under "Select Elements from" on the screen) must also contain historical observations for this SYSID.
LVL:	The LVL field was reset to Zone because the data was summarized by Zone. The Zone, Hour or Dayname to be included in the analysis, depending on whether you define the Primary CAP file at the ZONE, HOUR, or DAYNAME summarization level. The Secondary CAP file must also contain historical observations for this ZONE, HOUR or DAYNAME. The field name will automatically be changed from LVL to either ZONE, HOUR, or DAYNAME to reflect the Input file summarization level once you specify the Input file.

No. wk/mn: The maximum number of days, weeks, or months of historical observations included in the analysis. Note that the field name will change once you specify the file to reflect the timespan of the Primary CAP File. When used in conjunction with the Start date parameter, several options are available for date range selection control:

- o If you specify only Start date, all historical observations at least as recent as the value specified for Start date will be selected.
- o If you specify both Start date and No. wk/mn, the date range for data selection will begin with the value specified for Start date and will extend for the specified number of days, weeks, or months.
- o If you specify only No. wk/mn, the date range for data selection will include only this number of the most recent days, weeks, or months.
- o If you do not specify a parameter, then all historical observations will be selected.

Specify CAPAPU Values: Put a Y in this field to specify the CAPAPU value(s) (up to nine) to be included in the analysis. The CAPAPU values are specified on the CAPAPU Selection Values screen shown in Figure 14-10. You may specify CAPAPUs only if the Primary CAP file is a resource element file and is defined as containing the CAPAPU data elements. Note that the Secondary CAP file must also contain historical observations for the specified CAPAPU values.

Forecast the Length: The length of a forecast period, in terms of timespan assumed by the Primary CAP file. In other words, a Forecast Length of 13 when the Primary CAP file is summarized at the WEEKS timespan would be 13 WEEKS. A non-zero value is required; the current default value is 1.

Forecasting Method:	AUTOMATIC (default) indicates that the Workload Forecasting application will determine the adjustment values for each selected training element. MANUAL indicates that the analyst will specify the adjustment values on panel CAPG710. Changing this value from AUTOMATIC to MANUAL forces the Specify Elements value to be set to YES.
Adjustment Option:	INTERVAL (default) causes the Adjustment value to be applied equally across the Forecast Length period. ANNUAL causes the Adjustment value to be annualized before application. For example, an Adjustment value of 20 percent for a particular workload (training element) would be applied evenly across a 13 week Forecast Length if MANUAL is specified, resulting in a 20 percent increase in workload processor time by the end of the forecast period. Specifying ANNUAL would cause the 20 percent increase to be taken across one year, and the resulting increase then applied across the forecasting period. In this case, the workload processor time would be increased by 5 percent by the end of the forecasting period.
Specify Elements:	YES (default) forces the display of panel CAPG710 when panel CAPG700 is "ended". Panel CAPG710 is used to select a set of elements to be used for training, as well as to specify an adjustment percentage when using MANUAL adjustments.
Select Elements from:	The file that contains the data elements used for the training phase. This will default to the same file specified for the Primary CAP file, but may be any Sample Application, user defined Resource Element or Business Element Forecasting file that resides in the Capacity Planning database. If you are not sure of the files or file names available in your capacity planning database, type a question mark (?) in the field and a selection list will be displayed.

Save Forecast:	YES/NO/AGE/CHK determines if a file containing the forecasted future values of the Analysis element should be saved by itself (YES), saved and aged across three cycles (AGE), used only for ALERT detection and processing (CHK) or ignored (NO, the default) value. - o Specifying YES causes the forecasted observations to be saved in the Capacity Planning database Forecast file (CPFxxx01) without aging the file as described in Chapter 3. - o Specifying AGE causes the forecasted observations to be saved in the Primary CAP file and be aged as described in Chapter 3. - o Specifying NO results in the forecasted observations not being saved in the Forecast file. If you select this option, the Forecast file will not be aged. - o Specifying CHK results in the forecasted observations being used to detect and process ALERTS. If you select this option, no new Forecast file will be created but a Forecast file must exist.
Alerting Threshold:	A value from 01 to 99 percent, or "NO," if the the Save Forecast option does not specify 'CHK'. This value is used to determine the threshold values used for evaluating whether an Alert should be raised. The default value is 10 percent. Using the default for example, an Alert would be raised if the current Actual or Forecasted value deviated from the Baseline Forecast value by 10 percent.
Activate Alerting:	YES / NO specifies whether the Service Desk facility is available and used for Alert processing (YES) or that Alert processing is handled internally only (NO).

OUTPUT GENERATION OPTIONS:

Forecasting Report: YES / NO specifies whether a printer report of the training and forecasted values, along with the model evaluation statistics should be printed (YES) or ignored (NO).

Regression Residuals Report: YES / NO specifies whether a printed report of linear regression residual values should be printed (YES) or ignored (NO). This is a special case where the client wants to perform an exploratory data analysis on the Forecast element values. A linear regression procedure is executed against Forecast element values. The regression residual values are presented so that their values can be observed and verified.

Note: This report is not currently available, but a graph of residuals is produced automatically.

Color: YES / NO (Default is NO). Specifies whether a high resolution graphics should be generated.

CSV: YES / NO (Default is NO). Specifies whether a CA MICS Q&R WorkStation compatible comma separated value (CSV) file should be generated (YES) or ignored (NO). This file can be downloaded to a PC and used to generate graphs using the CA MICS Q&R WorkStation.

OUTPUT FORMATTING OPTIONS:

Format for Predicted Values: TIME / HOUR / MINUTE specifies how predicted values will be formatted in output operations. The default value of null causes the predicted values to assume the same format as the original Analysis element. A value of TIME causes a SAS format of TIME10.0 to be used. A value of HOUR causes a SAS format of HOUR6.4 to be used. A value of MINUTE causes time values to be first converted to minutes and then displayed using a SAS format of 5.0.

HAXIS Interval Values: DAY / WEEK / MONTH / QTR specifies the units used for HAXIS values for printer graphs. The default value of null causes the HAXIS values to be assigned using the timespan of the Analysis file.

HAXIS Interval Format: DATE / DDMMYY / MMDDYY / YYMMDD / MMY / MONYY causes various SAS formats to be used for the HAXIS values for printer graphs. The SAS formats used are as follows:

DATE	DATE9.
DDMMYY	DDMMYY10.
MMDDYY	MMDDYY10.
YYMMDD	YYMMDD10.
MMY	MMYYC7.
MONYY	MONYY7.

The default value of null causes a SAS format of DATE9.

If you entered a "Y" in the CAPAPU Values field on the Workload Forecasting Analysis screen, the CAPAPU Selection Values screen displays, as shown in Figure 14-10. On this screen, you can specify the CAPAPU values (up to nine) to be included in the analysis, only if the capacity planning database file is defined as containing the CAPAPU data elements.

```

/----- CAPAPU Selection Values -----\
|Command ==>|
|
|Composing CA MICS Inquiry: Neugents Workload Forecasting with CAPAPU|
|
|      Rule  +---+1---+2---+3---+4---+5---+6---|
|CAPAPU  ==>|-----|
|CAPAPU 2 ==>|-----|
|CAPAPU 3 ==>|-----|
|CAPAPU 4 ==>|-----|
|CAPAPU 5 ==>|-----|
|CAPAPU 6 ==>|-----|
|CAPAPU 7 ==>|-----|
|CAPAPU 8 ==>|-----|
|CAPAPU 9 ==>|-----|
|
|-----/

```

Figure 14-10. CAPAPU Selection Values Screen

14.5.2 Analysis File Data Elements Selection Screen

If you entered "YES" for the Specify Elements option, panel CAPG710 depicted in figure 14-11 is automatically displayed after exiting the Workload Forecasting Analysis main panel, CAPG700. This panel is used to select those workloads chosen for analysis. In this sample case, this panel would have been displayed and filled out by the analyst. The BATCH, ONLINE, DATABASE and TSO workloads mentioned initially were selected for training.

```

/-----Analysis File Data Elements-----\
|Command ==>                               Scroll ==> CSR |
|
|      CAP DB File def.: WLM - Workload Resource Usage Data
|
|      Forecasting Method: AUTOMATIC
|      Adjustment Option: INTERVAL
|
|      Show ==> SELECTED (Selected/All)
|
|      Line Cmds: S Select  U Unselect
|
|      Cmd   Name           Description           Adj.  Selected?
|      _    BATCH         Batch Workload Usage     N/A   YES
|      _    ONLINE        Online Workload Usage     N/A   YES
|      _    DATABASE      Database Workload Usage    N/A   YES
|      _    TSO           TSO Workload Usage         N/A   YES
|
\-----/

```

Figure 14-11. Analysis File Element Selection

Note that because the SHOW option was set to "SELECTED", only those workloads that had been previously selected are displayed. This helps avoid panel "clutter" and focuses more on the problem being addressed.

Once data entry on this panel has been completed, exiting the panel will result in the model being executed.

14.6 Case Studies

The case study that follows is an example involving the use of Workload Forecasting Analysis.

1 - Processor Usage Planning

14.6.1 Processor Usage Study

This case study presents an analysis of processor busy time using a set of primary workloads to drive the study. It assumes that a preliminary analysis using the Resource Component Analysis application has been performed using the same Forecast and Analysis files as are input to this study, and that a subset of workloads that best represent the processor usage has been identified.

In performance analysis and capacity planning, the basis of any study is the examination of measurement data. In this discussion, assume that the system being measured is a processor running the MVS operating system (although the underlying theory is independent of the type of operating system). In an MVS system, basic measurement data is usually derived from SMF and RMF data, which you store in the CA MICS Batch and Operation Analyzer (SMF) and Hardware and SCP Analyzer (RMF) areas respectively. For this case study we will narrow our discussion to RMF data and consider the relationships between data stored in the WLM Service Class Resource Consumption (WLMSEC) File and the CPU Processor Activity (HARCPU) File.

DEVELOPMENT OF CONCEPTUAL MODEL

Although a typical system executes many transactions across a number of workloads, generally speaking, there will be a subset of workloads. For example, some small number of workloads that tend to drive the processor busy. Some systems may even be devoted primarily to a single workload such as a dedicated IMS or CICS system, but that is not the normal situation. Although you could perform capacity planning using all workloads, such an approach requires more effort on the part of the analyst and would likely not yield more accurate results. Generally, it is better to plan on the primary workloads and ignore smaller components.

In our case study, four workloads have been determined to be the most "important", statistically speaking: BATCH, ONLINE, DATABASE and TSO. There were actually eight workloads represented in the input Analysis file, but four were discarded for the purposes of this case study. One interesting point was noted while preparing this case study. The TSO workload was ultimately included even though it was a minor user of processor resources, considerably less than some of the workloads that were discarded. The preliminary analysis using the Resource Component Analysis application assigned a higher Importance value to this workload than the other minor players. While the exact reason for this observation is unknown, Neugents technology employed by this application uses a nonlinear process to analyze data and therefore can discover relationships among data that are not as apparent when using other statistical methods of analysis. It is quite possible that the small TSO workload is employed mainly to "drive" one of the other larger workloads, such as the DATABASE workload. This demonstrates that there may be an indirect relationship between the TSO and DATABASE workloads, despite the rather low resource usage.

APPLICATION OF CONCEPTUAL MODEL

The model was developed and the training phase executed. The Training Phase Summary statistics showed that out of 53 weeks of input history data, the application selected 40 cases for training, yielding a model "goodness of fit" of 0.9960, meaning that the model prediction could account for over 99 percent of the forecasted values. Additionally, the average model error was less than 1 percent, indicating that the model was performing well.

A period of 13 weeks was chosen for the Forecast Length and the model output again demonstrated some interesting characteristics. In particular, the model predicted a nearly "flat" growth rate, not especially surprising given the nature of the input data. But the predicted values actually first curve upwards slightly and then curve downwards towards the end of the forecast. This clearly shows the nonlinear nature of the processor data and ability of Neugents technology to establish relationships among such data.

The rest of this section discusses the following topics relating to this case study:

- 1 - Control Parameters
- 2 - Input Data Descriptive Statistics
- 3 - Training Phase Summary
- 4 - Forecasting Phase Analysis Report
- 5 - Forecasting Phase Analysis Graph

14.6.1.1 Control Parameters

Using a short set of menus that are described in Section 1.5 of this guide, you will arrive at the Workload Forecasting Analysis screen shown in Figure 14-12.

Figure 14-12 shows the control parameters that were used to generate the final Workload Forecasting Analysis model in this case study.

```

/----- Workload Forecasting Analysis -----\
|Command ==>
|Enter a ? in any data entry field for more information on valid values.
|      CA MICS Inquiry: CPW - Sample Workload Forecasting Query
|Report title ==> Processor Usage Forecast
|
|Model Data Selection:
|Primary CAP file ==> APU      Forecast length ==> 13
|Forecast element ==> CPUADBTM Forecasting Method ==> AUTOMATIC (Manual/Auto
|      Start date ==> 01JUN200 Adjustment Option ==> INTERVAL (Int./Annual
|      SYSID ==> SYS1      Specify elements ==> YES      (YES/NO)
|      ZONE ==> 1      Select Elements from ==> APU      (CAP file)
|      No. Weeks ==> 53      Save forecast ==> YES      (YES/NO/AGE/CHK)
|      CAPAPU Values ==> NO      Alerting Threshold ==> NO      (00-99 Percent)
|      Activate Alerting ==> NO      (YES/NO)
|
|-----\
|Output Generation Options:
|Forecasting Report      ==> YES      Color ==> NO      CSV ==> NO
|Regression Residuals Report ==> N/A      Color ==> NO      CSV ==> NO
|      Color graphics format set ==> TERMINAL
|Output Formatting Options:
|Format for Predicted Values ==> HOUR
|HAXIS Interval Values      ==> QTR      Format ==> MONYY
\-----/

```

Figure 14-12. Processor Usage Study Control Parameters

Note that in this case study the analyst previously generated two resource element files. The first, APU, contains a data element, CPUADBTM, which represents the processor resource usage time, adjusted to an IBM 9672-r95 system. This adjustment was done to accommodate the fact that the LPAR being analyzed has been moved over time to different hardware platforms. This file and data element are specified on the control screen as the Primary CAP file and the Forecast element respectively. The second file, WLM, contains the data elements that represent the processor resources consumed by each of the workloads running on the system, and is used to train the model. This file is specified as the "Select Elements from" file and the corresponding data elements are specified on panel CAPG710.

The analyst also specified that the study would focus on SYSID SYS1 and Zone 1, and will initially use 53 weeks of planning data taken beginning in June of 2001.

For this study, we selected the Forecast report, in addition to the normally produced reports and graphs for output. If many observations are used in a given study, it might be advantageous to exclude the Forecasting report to reduce the volume of printed output.

14.6.1.2 Model Input Descriptive Statistics Report

Figure 14-13 shows the Input Descriptive Statistics Report for this case study. These statistics can be used to gauge the validity and robustness of the input data values.

MODEL INPUT DESCRIPTIVE STATISTICS					
REPORT FOR: 09/17/02					
ELEMENT	MIN. VALUE	MAX. VALUE	MEAN VALUE	STD	CV
CPUADBTM	19.80	177.51	150.41	28.39	18.88
BATCH	1.57	23.91	12.51	4.48	35.85
DATABASE	6.55	76.24	57.53	13.26	23.05
ONLINE	4.05	39.32	33.12	6.62	20.00
TSO	0.04	0.92	0.30	0.13	43.75

Figure 14-13. Processor Usage Case Study Input Descriptive Statistics Report

14.6.1.3 Training Phase Statistics Summary Report

This report displays the model statistics from the training phase and is used to evaluate the success and performance of the model training.

```
SUMMARY OF NEUGENT'S TRAINING PHASE
NUMBER OF CASES USED FOR TRAINING: 000040
    AVERAGE MODEL ERROR MEASURED: 0.0092
    AVERAGE MODEL RMS ERROR MEASURED: 0.0001
    'GOODNESS OF FIT' FOR MODEL: 0.9960
```

Figure 14-14. Training Phase Statistics Summary

For this case study, 40 cases were actually selected by the Workload Forecasting Analysis application for the training phase. The Average Model Error value of 0.0092 indicates that a model error rate of less than 1 percent was observed and the "Goodness of Fit" value of 0.9960 indicates that the model predicted accurate values of the Forecast element 99.6 percent of the time during training.

14.6.1.4 Forecasting Phase Analysis Report

This report displays all of the observations from the Forecasting file, including the Analysis elements used for training, the actual and predicted Forecast data values, along with individual delta values and model errors. The predicted future values of the Forecast element are also displayed. The model statistical summary is displayed on each page of the report for reference. This report can be used to document the model for a better understanding of the various data values.

NEUGENTS MODEL FORECASTING PHASE ANALYSIS								
REPORT OF FORECASTED CPUADBTM VALUES								
'GOODNESS OF FIT' FOR MODEL: 0.9960								
DATE	BATCH	DATABASE	ONLINE	TSO	ORIGINAL INPUT VALUE	PREDICTED VALUE	DELTA VALUE	MODEL ERROR
15SEP2001	2.67	45.54	29.12	0.23	144.27	142.87	1.40	0.0097
22SEP2001	2.75	41.29	31.21	0.35	136.17	137.78	1.62	0.0119
29SEP2001	3.06	53.18	34.63	0.39	162.61	162.11	0.49	0.0030
06OCT2001	7.93	63.50	31.33	0.38	159.36	158.31	1.04	0.0065
13OCT2001	13.99	71.30	32.69	0.29	165.47	162.30	3.17	0.0191
20OCT2001	12.99	66.66	34.37	0.32	160.39	165.71	5.32	0.0331
27OCT2001	13.88	52.05	29.12	0.32	133.68	134.29	0.61	0.0045
03NOV2001	14.19	70.01	35.56	0.30	167.47	169.48	2.02	0.0120
10NOV2001	14.69	50.32	32.00	0.29	141.49	142.92	1.42	0.0101
17NOV2001	13.91	51.62	35.40	0.34	147.46	146.03	1.43	0.0097
24NOV2001	15.50	38.81	25.59	0.23	115.67	115.47	0.20	0.0017
01DEC2001	17.20	58.84	37.84	0.30	163.94	163.48	0.46	0.0028
08DEC2001	13.75	60.88	36.40	0.24	159.86	158.93	0.94	0.0059
15DEC2001	15.15	60.09	37.31	0.31	163.05	161.10	1.95	0.0120
22DEC2001	17.44	56.54	33.22	0.18	154.92	154.17	0.75	0.0048
29DEC2001	10.33	31.49	16.82	0.14	86.25	86.11	0.14	0.0016
31DEC2001	1.57	6.55	4.05	0.04	19.80	19.94	0.14	0.0069
05JAN2002	11.88	34.93	21.15	0.17	99.09	99.41	0.32	0.0032
12JAN2002	16.31	60.84	37.65	0.38	167.67	166.42	1.25	0.0075
19JAN2002	15.24	61.11	38.68	0.20	166.61	166.93	0.32	0.0019
26JAN2002	15.85	55.85	37.19	0.29	157.22	159.74	2.52	0.0161
02FEB2002	14.77	63.80	39.32	0.36	169.00	172.13	3.13	0.0185
09FEB2002	11.24	60.06	34.22	0.35	152.11	154.39	2.28	0.0150
16FEB2002	12.15	59.74	36.37	0.23	158.76	160.74	1.98	0.0125
23FEB2002	10.47	53.77	33.49	0.20	142.05	143.95	1.91	0.0134
02MAR2002	9.15	57.73	36.66	0.31	150.36	148.77	1.59	0.0106
09MAR2002	12.67	61.95	39.14	0.34	165.46	164.99	0.47	0.0028
16MAR2002	11.95	67.60	39.30	0.38	173.09	170.94	2.15	0.0124
23MAR2002	9.11	59.23	37.02	0.55	154.08	155.29	1.21	0.0078
30MAR2002	13.53	61.93	37.49	0.31	163.67	160.67	3.00	0.0183
06APR2002	19.47	75.42	34.47	0.21	177.51	178.54	1.03	0.0058
13APR2002	9.74	76.24	34.57	0.23	164.34	165.44	1.10	0.0067
20APR2002	12.89	75.06	36.22	0.25	168.64	165.82	2.82	0.0167
27APR2002	10.31	74.04	37.12	0.25	166.66	167.80	1.14	0.0068
04MAY2002	14.57	55.43	29.77	0.25	142.77	141.59	1.17	0.0082
11MAY2002	12.67	64.10	31.22	0.36	155.23	154.30	0.93	0.0060
18MAY2002	15.78	55.85	35.58	0.28	157.35	157.38	0.03	0.0002
25MAY2002	11.24	55.38	33.01	0.27	145.13	142.69	2.44	0.0168
01JUN2002	14.46	66.30	33.47	0.24	160.80	161.15	0.35	0.0022
08JUN2002	23.91	66.17	35.16	0.92	176.78	176.05	0.72	0.0041
15JUN2002	15.35	65.61	35.57	0.34		167.29		
22JUN2002	15.49	66.01	35.69	0.34		168.18		
29JUN2002	15.63	66.41	35.82	0.34		168.87		
06JUL2002	15.78	66.81	35.94	0.35		169.34		
13JUL2002	15.92	67.21	36.06	0.35		169.55		
20JUL2002	16.06	67.62	36.18	0.35		169.51		
27JUL2002	16.20	68.02	36.30	0.35		169.22		
03AUG2002	16.34	68.42	36.43	0.35		168.67		
10AUG2002	16.48	68.82	36.55	0.36		167.88		
LEGEND: ACTUAL VALUES (CPUADBTM) - 'A'								
PREDICTED VALUES - 'P'								
NEUGENTS MODEL FORECASTING PHASE ANALYSIS								
REPORT OF FORECASTED CPUADBTM VALUES								
'GOODNESS OF FIT' FOR MODEL: 0.9960								
DATE	BATCH	DATABASE	ONLINE	TSO	ORIGINAL INPUT VALUE	PREDICTED VALUE	DELTA VALUE	MODEL ERROR
17AUG2002	16.62	69.22	36.67	0.36		166.88		
24AUG2002	16.77	69.63	36.79	0.36		165.69		

31AUG2002	16.91	70.03	36.91	0.36	164.36
07SEP2002	17.05	70.43	37.04	0.36	162.94

Figure 14-17. Forecasting Phase Analysis Report

This report displays one line per observation period; weekly in this case study. This sample report uses the "HOURS" option for formatting the resource element values and shows a sample of the modeled input. Since this information is also available in a more concise form as a printer graph, a typical model execution will probably generate the graph(s) and reexecute, generating the reports if more detail is required.

For this case study, the entire report has been made available in order to demonstrate the prediction of future values of the Forecast element, CPUADBTM.

14.6.1.5 Forecasting Phase Analysis Graph

This graph displays all of the observations from the Primary CAP file, including the actual and predicted values of the Analysis element. The model "goodness of fit" statistic is displayed on each page of the report for reference. This graph can be used to evaluate the success of the model forecasting phase.

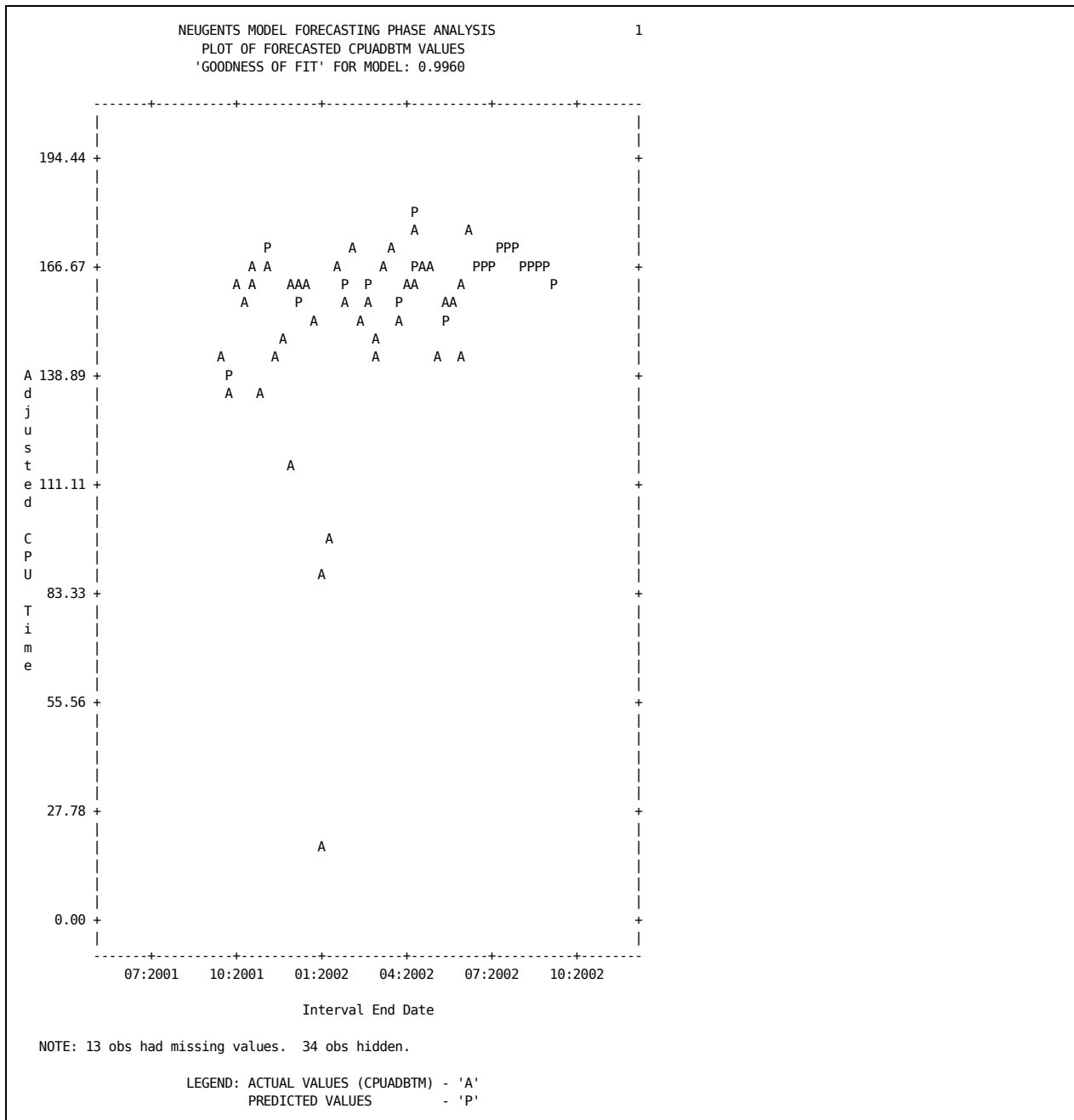


Figure 14-18. Forecasting Phase Analysis Graph

This graph provides an easy visual evaluation of model effectiveness by displaying the actual and predicted values of the Forecast element. This display, along with the model "goodness of fit" statistic repeated here allow the analyst to make a quick and accurate assessment of model performance during forecasting and determine whether the model should be considered as trustworthy. In this case study, in spite of the odd pattern of the usage data, the model has performed very well in predicting the Forecast element values and is considered to be reliable. Note that the future values predicted by the model actually increase initially and then decrease, owing to the non-linear methodology employed by Neugents technology. This is a strength of this particular approach and provides for more accurate results.

Chapter 15: WORKLOAD CHARACTERIZATION

Workload Characterization using CA's U.S. patented Neugents technology, provided in the CA MICS Capacity Planner, enables you to apply a clustering methodology to the analysis of your installation's workload. Workload characterization is an attractive approach to capacity planning and performance management problems because it allows the number of workload elements that need be considered in a study to be reduced from tens of thousands to only a few.

Clustering methods are a statistical extension of scatter plots to identify similarities and differences between workloads. Scatter plots are often difficult to prepare and depend heavily on visual interpretation of the data. The need for visual interpretation limits the use of scatter plots to two or perhaps three axes. Clustering overcomes these disadvantages through its ability to recognize patterns in multiple dimensions. Using the CA MICS database as an input data source for clustering simplifies and extends the application of the technique.

Note: Neugents technology is a prerequisite for using Workload Characterization.

The following sections describe the analytic techniques, operational considerations, and reports used in the application.

This section contains the following topics:

- [15.1 Functional Description](#) (see page 980)
- [15.2 Usage Guidelines](#) (see page 981)
- [15.3 Standard Output](#) (see page 988)
- [15.4 Analytic Technique Tutorial](#) (see page 1008)
- [15.5 Component Operation](#) (see page 1019)
- [15.6 Case Studies](#) (see page 1034)

15.1 Functional Description

During the early 1970s, Domenico Ferrari introduced the concept of workload characterization (FER72). The basic premise of the concept was that studies of computer systems could be facilitated by the identification of a number of "resource patterns" that could be used to statistically represent a system's workload. Since capacity planning and performance management studies often encompass tens of thousands of jobs or hundreds of thousands of transactions, the ability to represent the workload by a limited number of resource patterns can significantly simplify a study. (Typical studies result in the identification of from 10 to 30 patterns.)

Ferrari identified these resource patterns, which he called clusters, by visual inspection of two-dimensional scatter plots. For example, Ferrari reported studies where he examined scatter plots of print lines and CPU seconds representing the resource usage patterns of student jobs. Unfortunately, the dependence on visual inspection limited the technique to two or perhaps three dimensions, since graphic representations become extremely complex for higher dimensions. Also, the technique depended upon the visual skill of the interpreter, since some analysts could "see" patterns better than others. These problems limited the applications of what was otherwise a valid concept.

In 1976, however, two papers were published that detailed the application of statistical pattern recognition (clustering) techniques to the workload characterization problem. They allowed problems of higher dimensions to be solved, and eliminated the dependence on the analyst's interpretive skill and judgment. Agrawala and Mohr reported on their initial studies of Univac workloads at the University of Maryland (AGR76) and Artis reported on studies of IBM workloads at Bell Laboratories (ART76). Since that time, these and a number of other authors have published numerous papers on the subject in various publications, such as the Computer Measurement Group (CMG) proceedings.

Workload Characterization using Neugents technology allows you to apply cluster-based workload characterization techniques to the CA MICS database. The software can be applied to any of the unsummarized files that are supported in the DETAIL timespan. The program identifies clusters (that is, similar patterns of activity) in the workload using a Neugents technology based clustering procedure.

The use of clustering techniques introduces other statistical issues. Among these are scaling and the treatment of outliers. Improper scaling and failure to account for the influence of abnormally large observations are the most common reasons for invalid workload characterization studies. These issues are discussed in detail in Section 15.2.4.

The CA MICS Capacity Planner Option includes features that utilize CA's U.S. Patented Neugents(R) technology.

In order to exploit this technology the CA MICS Root Directory under Unix System Services needs to be initialized and the Neugents technology infrastructure enabled.

For detailed instructions on how to initialize the CA MICS root directory and enable Neugents technology, review and complete the steps presented in Chapter 12 of this guide.

15.2 Usage Guidelines

Before the introduction of cluster-based workload characterization, a number of other techniques were utilized. These approaches provide valuable insight into potential errors in the workload characterization process.

The equivalent job approach fails to provide an adequate characterization, since it lacks the important property of robustness. Relatively small changes in the workloads can cause large changes in the definition of an equivalent job. This means that the results of the equivalent job approach are too closely associated with the data used in the analysis to provide general insight into the workloads on the system. Cluster-based workload characterization can replace commonly practiced ad hoc techniques.

The most important issues regarding implementation concern the type of data elements that are used to characterize the clusters, and the treatment of unusual points that cannot be represented by clusters. These points, known as outliers, can have a profound effect on the clustering study.

These topics and the use of certain CA MICS database files for clustering analysis are discussed in the following sections:

- 1 - Historical Approaches
- 2 - Basic Hypotheses
- 3 - Application of Workload Characterization
- 4 - Data Selection and Outliers
- 5 - Suggested CA MICS Input Files

15.2.1 Historical Approaches

One early approach presented to capacity planning was the equivalent job methodology described by Jenkins in 1977 (JEN77). Most simply described, the resources consumed by all of the workload elements (for example, job, steps, transactions) over some number of weeks or months are divided by the number of workload elements that were executed during the period. The result of this division is called an equivalent workload element.

An equivalent workload concept initially appears to be very attractive since you can estimate future resource requirements by simply estimating the number of equivalent workload elements. This value is multiplied by the equivalent workload description to calculate the future resource requirements.

15.2.2 Basic Hypotheses

The problems associated with the equivalent workload element approach can best be introduced by stating the two hypotheses that were introduced by Artis for testing proposed workload characterizations (ART76). These hypotheses are the "representative hypothesis" and the "stationary hypothesis."

The representative hypothesis tests how well the proposed characterization represents each element of the workload. Rather than testing the proximity of the mean value of the workload to the proposed characterization, you must evaluate the differences between the proposed characterization and each element of the workload.

For example, consider the example of a system that runs 100 testing jobs using less than 5 CPU seconds and 100 production jobs using about 10 CPU minutes each day. (Of course, real workloads are much more complex.) This would result in an equivalent job that required about 5 CPU minutes. Unfortunately, no job in the workload has resource requirements similar to this equivalent job. Therefore, rather than simplifying the problem by reducing the number of workload elements to be considered, we have changed the nature of the problem.

The characteristics of a system that runs 200 equivalent (that is, average) jobs per day would be much different from those of the hypothetical system discussed above. It is interesting to note that the equivalent workload element can only hope to pass the representative hypothesis for systems where all of the workload elements are approximately equal. To be valid, any proposed workload characterization must pass the representative hypothesis test.

The stationary hypothesis evaluates the likelihood that the proposed workload characterization is valid for future workloads. To test this hypothesis, you must wait and test the characterization's applicability over a number of weeks or months. (Alternatively, you can produce the characterization on older historical data and compare it to more recent history that is not being used to do the characterization.)

You should never consider a workload characterization as a tool for forecasting unless this hypothesis has been tested. Using the hypothetical system discussed above, the smallest change in the ratio of test jobs to production jobs would significantly change the workload characterization. Moreover, even small changes in the average resource consumptions for testing or production jobs would have the same effect. Therefore, it is unlikely that any equivalent workload element approach would ever pass the stationary hypothesis test.

In general, very few proposed workload characterizations will ever pass the stationary hypothesis test. This does not mean that most workload characterizations are invalid; it does mean that most are unsuitable for use as forecasting units.

15.2.3 Application of Workload Characterization

Perhaps one of the most common workload characterization studies is the determination of job classes. Typically, the analyst guesses at a solution and then evaluates how well it fits the system's workload. This is known as the ad hoc approach. For example, consider a hypothetical system where the following job class structure has been proposed.

JOB CLASS	CPU MINS	PRINT LINES
=====	=====	=====
A	2	5,000
B	5	20,000
C	unlim	unlim

If 60 percent of the jobs are assigned to class A, 25 percent to class B, and 15 percent to class C, the analyst might assume that the classes provide a good representation of the workload.

Applying the representative hypothesis test could reveal problems. For example, if 30 percent of the jobs use less than 5 CPU seconds and print less than 2,000 lines, then they would be poorly represented by the class A limits. Moreover, their service characteristics will be much poorer when they are randomly mixed with the larger jobs in class A than they will be if they are assigned to a class by themselves.

Another problem resulting from arbitrarily established job classes is the specification of a job class limit that bisects a natural structure in the data. For example, if 15 percent of the jobs normally require between 110 and 130 CPU seconds, imposing a two-minute limit could force the users of jobs into class B to ensure that they are not canceled at the two-minute limit. This problem is also identified by the representative hypothesis test since the approximate two CPU minute jobs are poorly represented by the class B limits.

The characteristics of an installation's job mix are continually evolving. As a result, it is probably unreasonable to apply the stationary hypothesis test to a job class structure. However, you should examine the job class structure on a quarterly or semiannual basis.

Both the equivalent job and ad hoc batch approaches to workload characterization discussed here show the problems that result from using mean values of arbitrarily selected limits. To meet the requirements of the first hypothesis, we must exploit the natural structure of the workload. Statistical pattern recognition techniques (clustering) offer a powerful tool for determining whether a natural structure exists in an installation's workload data.

15.2.4 Data Selection and Outliers

Selecting data elements for the clustering procedure is one of the most important steps in the workload characterization process. In clustering terminology, these data elements are known as features. The result of the clustering process is a ranking of similar types of workloads, based on the mean values of these features.

The objective is to select the minimum number of features that can be used to accurately describe the characteristics of the workload to be analyzed. A small number of features is desirable, since it greatly reduces the number of sources of variability in the characterization, making the results more robust and easier to understand. Note that each time an additional feature is added, splits may occur in as many as half of the existing clusters.

Another consideration in the feature selection process is the representation of outliers. Outliers are data elements that cannot be generalized as similar to other data elements in any of the chosen features. One example of an outlier is a batch job which falls into a CPU-bound loop due to a programming error. Another example is a periodic process, such as a monthly close of a financial application, which is run rarely but requires a large number of resources. Outliers have the highest probability of being poorly represented by the developed workload characterization.

There are two classes of features that are normally used in workload characterization studies:

- o Rate-based features, such as EXCPs per CPU second
- o Total features, such as CPU minutes

Rate-based features are desirable because they allow you to control the content of the outliers. For example, consider the use of CPU minutes and EXCPs as total features for characterizing the resource consumption of a workload. Because the largest jobs would be outliers, the resulting characterization would poorly represent them. Unfortunately, these jobs might represent a significant percentage of the resources that are consumed.

To avoid this problem, you could use rate-based features. When you divide EXCPs by CPU seconds, you change the content of the outliers. Rather than representing jobs that consume significant resources, they represent jobs with exceptional EXCP/CPU second rates. The probability that these jobs would consume significant resources is very small, since the ratio of elapsed time to CPU resource consumption for I/O-bound jobs is very high.

Such rate-based features have been called instantaneous workload descriptions, since they characterize the activities of a workload element for an instant of CPU time (ART76, ART78).

Although using total features introduces concerns about representing outliers, they have many applications to capacity planning and performance problems. For example, using total features for job and transaction class structures treats the occasional massive workload element as an exception. By handling the occasional element as an exception, the system can concentrate on exploiting the characteristics of the majority of the workload elements to simplify scheduling. You must, however, always be aware of the content of the outliers and pay particular attention to the marginally assigned resource vectors to verify that the objectives of the study are not compromised.

15.2.5 Suggested CA MICS Input Files

Although you may analyze any of the unsummarized CA MICS files in the DETAIL timespan, a number of files may particularly interest you. In the Batch Information Area, they are as follows:

- BATJOB Contains one record for each job
- BATPGM Contains one record for each job step, that is, program
- BAT_TS Contains one record for each TSO session
- BAT_ST Contains one record for each started task

In the TSO Information Area, the following files may interest you:

- TS0TSU Contains one record for each CA TSO/MON PM interval in which a user was active
- TS0TSI Contains one record for each command or program executed interactively

In the CICS Information Area, you can create files in the DETAIL timespan that contain one record for each transaction using the _USRXfff macros. These macros are discussed in Chapter 10 of the CA MICS Analyzer Option for CICS Guide. In the event that you use this option, the following files may interest you:

- CICCSU Contains one record for each transaction sorted by user
- CICCAU Contains one record for each transaction sorted by application

You can also create IMS transaction files in the DETAIL timespan that contain one record for each transaction using the _USRXfff macros. These macros are discussed in Chapter 10 of the CA MICS Analyzer Option for IMS Guide. In the event that you choose this option, the following files may interest you:

IMISISU	Contains one record for each transaction sorted by user
IMSIAU	Contains one record for each transaction sorted by application

Although you can analyze other CA MICS information areas and files using Workload Characterization, no results have been published that detail the application of clustering algorithms to other than workload element data.

15.3 Standard Output

Workload Characterization using Neugents technology produces reports and simple graphs that describe the original data, show the clusters that are identified from this data, and analyze the degree of fit with which these clusters characterize the data.

The five reports and two graphs are:

- o Descriptive Statistics Report. This report details the characteristics of the variables that are specified as criteria for the clustering analysis.
- o Cluster Performance Summary. This report consists of three sections: Clustering Executions, a summary of the options selected by the analyst for the study; Cluster Feature Contents, a list of each of the elements or features used for clustering, and the boundaries selected for determining normalcy of the data values; and Cluster Population Summary, which lists each cluster, its size and population, and an index used to indicate relative performance of the cluster.

- o Cluster Population Summary. This report provides the resource consumption characteristics of each of the clusters. For each cluster, its size and population is summarized and all clustering and reporting elements are listed, along with their respective usage values. This report permits you to identify resource consumption patterns that may represent performance or capacity problems.
- o Data Value Exception Detail. This report provides a detail listing of any observation that has one or more clustering elements that are identified as outliers. The report is sequenced in the same order as is the original CA MICS files used as input, and can be used to further investigate why certain cases have been identified as outliers and whether or not they actually represent abnormal situations.
- o Sparse Cluster Population. This report provides a listing of all clusters that have been identified as 'sparse' and their population contents. A sparse cluster is defined within this application as one having less than 0.5% of the overall population of the study. For example, if a sample of 2,000 cases were selected for analysis, then a sparse cluster would contain ten or fewer observations.
- o Cluster Index Analysis. This graph depicts a measurement of relative performance for each cluster, by presenting an index calculated by Neugents technology in terms of the root mean square of the distances of the resource vectors or feature values from the cluster centroid. The cluster index of each cluster is displayed in either a Block or a horizontal bar chart, depending on the number of clusters being presented. This chart permits you to quickly access the overall performance of the workload characterization effort.
- o Cluster Population Analysis. This graph depicts the population of the top "n" clusters in a PIE chart, which can be used to understand pictorially how the study data actually populated the clusters.

The following sections describe the fields, data elements, and format of each report.

- 1 - Descriptive Statistics Report
- 2 - Cluster Performance Summary
- 3 - Cluster Population Summary
- 4 - Data Value Exception Detail
- 5 - Sparse Cluster Population report
- 6 - Cluster Index Analysis graph
- 7 - Cluster Population Analysis graph

15.3.1 Descriptive Statistics Report

The Descriptive Statistics Report provides a summary of the statistics that are calculated for the cluster features. Figure 15-1 illustrates a sample Descriptive Statistics Report.

Workload Characterization Analysis Clustering Input Descriptive Statistics For: Thursday, June 19, 2003						
Summary for Total Sample:						
Number of Sample Observations:		2,000				
Feature	Description	Minimum	Maximum	Average	Std. Dev.	CV
JOBTCTBM	Job TCB CPU Time	0:00:00.01	0:40:00.28	0:00:13.14	0:01:42.89	7.83
JOBEDASD	Total EXCPs	1.00	1,154,523.00	5,992.30	34,582.08	5.77
Summary for Trimmed Sample:						
Number of Sample Observations:		1,919				
Feature	Description	Minimum	Maximum	Average	Std. Dev.	CV
JOBTCTBM	Job TCB CPU Time	0:00:00.01	0:01:00.81	0:00:03.74	0:00:07.71	2.06
JOBEDASD	Total EXCPs	1.00	38,096.00	2,505.00	4,508.48	1.80

Figure 15-1. Descriptive Statistics Report

Note that the report has two sections: one for the original sample randomly taken from the input data; and the other, the sample data after trimming, using the Sample Trim Limit value specified by the analyst on the Execution Options panel.

The Descriptive Statistics Report contains the following fields:

OBSERVATIONS: The number of observations selected for processing.

FEATURE:	The feature name. The feature names that are listed in the report correspond to the features that are specified on the Workload Characterization panel. In Neugents technology, this is also called a "pattern."
DESCRIPTION:	The SAS label of the selected data element, from either the CA MICS GENLIB definition or supplied by the user.
MINIMUM:	The minimum value observed for the feature in the sample.
MAXIMUM:	The maximum value observed for the feature in the sample.
AVERAGE:	The average calculated for the feature observations in the sample.
STD DEV:	The standard deviation calculated for the feature observations in the sample.
CV:	The coefficient of variation. The CV is calculated by dividing the standard deviation by the average.

15.3.2 Cluster Performance Summary Report

```

Workload Characterization Analysis                               1
Cluster Performance Summary
For: Monday, June 23, 2003

Clustering Execution Options

Cluster Input Data: SAMPLE
Clustering Method: RADIUS
Maximum Cluster Radius: 3.0

Training Sample Size: 2000
Sample Trim Limit: 97.5 Pct.
Sampling Percentage : 2.0 Pct.

Include Outliers in Clusters: YES
Include Sparse Clusters: YES
Sparse Cluster Limit: 0.5 Pct.

Report Cluster Population: YES
Report Using Account Codes: YES
Report Sparse Clusters : YES
Report Outliers Separately: YES
Create Cluster Index Graph: YES
Create Population Graph: YES

Input CA MICS Files: BATJOB
Input Dataset Name: 'COGDA01.AUDIT.BATJOB'

Clustering Variables: JOBTCBTM JOBEXCPS
Reporting Variables: JOBNLR
Delete obs. if zero: JOBTCBTM
Save Input Observations as: SAVEOBS
In User Dataset: 'COGDA01.02.CAPACITY'

Save Cluster Observations as: CLUSOUT
In User Dataset: 'COGDA01.02.CAPACITY'
SYSID's Selected: XAL1
Zones Selected: 1 2 3
Hours Selected: 00 23
Date Ranges Selected: 01JAN01 31DEC03

Note: Global Data Selection exit in force.
      Selection of input data may be affected.

Note: Non-CA MICS data element derivation exit in use.

```

Cluster Feature Contents				
Feature Name	Description	-----Outlier Limits-----		
		Lower	Upper	Cases
JOBTCBTM	Job TCB CPU Time	0:00:00.00	0:0 1:05.70	50
JOBEDASD	DASD EXCPS	0.00	38,933.00	50
				=====
				100

Note: A given case may be classified as an outlier more than once if multiple analysis elements contain abnormal values. Therefore, the totals in this report will not necessarily agree with those of the Population section.

Workload Characterization Analysis						2
Cluster Performance Summary						
For: Monday, June 23, 2003						
Cluster Population Summary						
Cluster	Radius	Normal Obs.	Outlying Obs.	Total Obs.	% of Population	Clustering Index
1	0.00	0	14	14	0.70	0.00
2*	0.82	0	2	2	0.10	0.82
3*	0.82	0	6	6	0.30	0.51
4*	1.17	0	2	2	0.10	1.17
5*	1.20	0	2	2	0.10	1.20
6*	1.39	4	0	4	0.20	1.01
7*	1.43	0	3	3	0.15	1.02
8*	1.50	0	2	2	0.10	1.50
9*	1.55	0	5	5	0.25	1.22
10	1.61	1,541	0	1,541	77.05	0.34
11	1.71	15	0	15	0.75	0.77
12*	1.86	0	9	9	0.45	1.26
13*	2.01	0	2	2	0.10	2.01
14	2.04	4	8	12	0.60	1.19
15	2.07	129	0	129	6.45	0.95
16	2.22	12	0	12	0.60	0.80
17	2.25	0	11	11	0.55	1.69
18*	2.29	0	2	2	0.10	2.29
19*	2.45	0	2	2	0.10	2.45
20*	2.54	0	3	3	0.15	1.99
21*	2.70	0	2	2	0.10	2.70
22*	2.71	2	6	8	0.40	1.68
23	2.72	39	0	39	1.95	1.70
24	2.83	173	0	173	8.65	0.98
		=====	=====	=====	=====	
		1,919	81	2,000	100.00	
'*' Indicates that the cluster is sparsely populated. Sparse clusters are defined as having a population that is less than 0.5% of the total population. In this study, the sparse cluster population limit is 10 cases.						

Figure 15-2. Cluster Performance Summary report

The Cluster Performance Summary report is presented in three sections as described below:

The first section explains the execution options you have selected for the study being performed, and documents which reports or graphs will be produced.

For an explanation of the execution options and their respective usage, see Section 15.5 in this guide.

The second section of the report presents the feature contents of the clusters, including the limits used to determine outliers, and the number of times (cases) observations were declared as outliers because of a given feature value.

FEATURE NAME:	The name of the data element chosen for clustering. This is normally a CA MICS data element but can be a computed (user defined) element if required.
DESCRIPTION:	The SAS label of the selected data element, from either the CA MICS GENLIB definition or supplied by the user.
OUTLIER LIMITS:	Statistical bounds to determine if a given data value is considered "normal" or represents an abnormal condition. These values are computed based on the Sample Trim Limit value specified on the Execution Options panel.
LOWER:	Cases (observations) whose value for this feature are less than this defined value are considered outliers. Given that most performance measurement data is positive in scope, the current implementation uses a value of zero for this boundary.
HIGHER:	Cases (observations) whose value for this feature are greater than this defined value are considered outliers.
CASES:	The number of cases (observations) that were declared as outliers because of this feature value. Note that a given case may be flagged for multiple times and therefore the number of cases presented in this section may exceed the total number of outlying observations in the next report section.

A footnote to this affect is printed at the bottom of this report section.

The last report section presents a summary of the size, population and general performance of each cluster.

CLUSTER NUMBER:	The cluster number. Cluster numbers are assigned sequentially to the patterns that are identified by the algorithm. Note that the order in which the clusters are identified is not an indicator of merit.
RADIUS:	The geometric distance from the cluster center centroid to the outer boundary of the cluster, expressed in terms of Standard Deviations. The outer limit of this value is defined as the Maximum Cluster Radius value on the Executions Options panel.
NORMAL OBS:	The count of observations within this cluster where all feature values were found to be "normal" in a statistical sense. In this implementation, normal feature values are those that are less than the value determined by the Sample Trim Limit. For example, if the Sample Trim Limit is 97.5%, then all feature values of "normal" clusters would reside below the 97.5 percentile of the sample.
OUTLYING OBS:	The count of observations within this cluster where one or more feature values were found to be "outliers" in a statistical sense. In this implementation, outlying feature values are those that are greater than the value determined by the Sample Trim Limit. For example, if the Sample Trim Limit is 97.5%, then at least one feature value of "outlying" clusters would reside above the 97.5 percentile of the sample.
TOTAL OBS:	The sum of NORMAL and OUTLYING observations.
% OF POPULATION:	The percent of the total sample populations represented by this cluster.

CLUSTERING INDEX: Formally called the Performance Index, this metric was renamed in this implementation to avoid confusion with similar terms in the z/OS Workload Manager. It is the Root Mean Square of the distances of all cases (observations) within the cluster, and serves as a simple measure of clustering effectiveness. The lower this value, the tighter the fit of the cluster data. This is not to say that outliers are not present; only that they are not distorting the cluster shape by their presence.

Note the footnote at the bottom of this report section. It refers to "sparse" clusters and presents the definition and limits used for determining which clusters are considered "sparse". Sparse clusters are generally populated by outliers and are often dropped from further analysis after their contents are reviewed.

15.3.3 Cluster Population Summary Report

A sample Cluster Population Summary report is shown below in Figure 15-3.

Workload Characterization Analysis							1
Cluster Population Summary							
For: Tuesday, June 24, 2003							
Cluster: 1							
Radius: 0.00 Clustering Index: 0.00 Maximum Cluster Radius: 3.0 STDs							
Normal Obs.:	0	0.00%	Observations <= 97.5 Pct.				
Outlying Obs.:	14	0.70%	Observations >= 97.5 Pct.				
Total Obs.:	14	0.70%					
Cluster Feature Resources							
Feature	Description	Minimum Value	Average Value	Standard Deviation	Maximum Value	Total Value % of Total	
JOBTCBTM	Job TCB CPU Time	25	656	687	2,400	9,186 34.96	
JOBEDASD	DASD EXCPS	5,103	236,581	298,042	1,154,523	3,312,131 27.64	
Cluster Report Resources							
Feature	Description					Total Value % of Total	
JOBNLR	Total Logical Writer Records					0 0.00	
Cluster: 10							
Radius: 1.61 Clustering Index: 0.34 Maximum Cluster Radius: 3.0 STDs							
Normal Obs.:	1,541	77.05%	Observations <= 97.5 Pct.				
Outlying Obs.:	0	0.00%	Observations >= 97.5 Pct.				
Total Obs.:	1,541	77.05%					
Cluster Feature Resources							
Feature	Description	Minimum Value	Average Value	Standard Deviation	Maximum Value	Total Value % of Total	
JOBTCBTM	Job TCB CPU Time	0	1	2	14	1,873 7.13	
JOBEDASD	DASD EXCPS	1	914	1,155	6,523	1,408,821 11.76	
Cluster Report Resources							
Feature	Description					Total Value % of Total	
JOBNLR	Total Logical Writer Records					19,968 99.56	
Sparse Clusters Have Been Excluded							

Figure 15-3. Cluster Population Summary report

Workload Characterization Analysis Cluster Population Summary	
For: Monday, June 23, 2003	
The following clusters were determined to be 'sparse' in their population and have been excluded from processing.	
Cluster	Population
2	2
3	6
4	2
5	2
6	4
7	3
8	2
9	5
12	9
13	2
18	2
19	2
20	3
21	2
22	8

CLUSTER NUMBER:	Cluster numbers are assigned sequentially to the patterns that are identified by the algorithm. Note that the order in which the clusters are identified is not an indicator of merit.
RADIUS:	The geometric distance from the cluster center centroid to the outer boundary of the cluster, expressed in terms of Standard Deviations. The outer limit of this value is defined by the analyst as the Maximum Cluster Radius value on the Execution Options panel.
CLUSTERING INDEX:	Formally called the Performance Index, this metric was renamed in this implementation to avoid confusion with similar terms in the z/OS Workload Manager. It is the Root Mean Square of the distances of all cases (observations) within the cluster and serves as a simple measure of clustering effectiveness. The lower this value, the tighter the fit of the cluster data. This is not to say that outliers are not present; only that they are not distorting the cluster shape by their presence.

MAXIMUM CLUSTER RADIUS: The maximum size of any cluster in the study under consideration, expressed in terms of standard deviations. This value is specified by the user on the Execution Options panel and is presented here for documenting the cluster definitions.

NORMAL OBS: The count of observations within this cluster where all feature values were found to be normal in a statistical sense. In this implementation, normal feature values are those that are less than the value determined by the Sample Trim Limit. For example, if the Sample Trim Limit is 97.5%, then all feature values of "normal" clusters will reside below the 97.5 percentile of the sample.

OUTLYING OBS: The count of observations within this cluster where one or more feature values were found to be "outliers" in a statistical sense. In this implementation, outlying feature values are those that are greater than the value determined by the Sample Trim Limit. For example, if the Sample Trim Limit is 97.5%, then at least one feature value of "outlying" clusters will reside above the 97.5 percentile of the sample.

TOTAL OBS: The sum of NORMAL and OUTLYING observations.

CLUSTER FEATURE RESOURCES:

For each feature (clustering element) defined on the Workload Characterization panel, a separate line is generated here and contains the following elements:

FEATURE: The name of the data element chosen for clustering. This is normally a CA MICS data element, but can be a computed (user defined) element if required.

DESCRIPTION: The SAS label of the selected data element, from either the CA MICS GENLIB definition or supplied by the user.

MINIMUM VALUE: The minimum of all values for this feature within this cluster.

AVERAGE VALUE:	The average of all values for this feature within this cluster. This value approximates the centroid value for this feature and is also referred to as the feature mean.
STANDARD DEVIATION:	The standard deviation of all values for this feature within this cluster.
MAXIMUM VALUE:	The maximum of all values for this feature within this cluster. This value approximates the outer boundary value for this feature.
TOTAL VALUE:	The sum of all values for this feature within this cluster.
% OF TOTAL:	The percentage of the sum of this feature's values for this cluster compared to the sum for the entire sample population. For example, in the report above, the JOBTCBTM represented by cluster 1 is nearly 35% of the JOBTCBTM for the entire sample.

CLUSTER REPORT RESOURCES:

For each feature (reporting element) defined on the Workload Characterization panel, a separate line is generated here and contains the following elements:

FEATURE:	The name of the data element chosen for clustering. This is normally a CA MICS data element, but can be a computed (user defined) element if required.
DESCRIPTION:	The SAS label of the selected data element, from either the CA MICS GENLIB definition or supplied by the user.
TOTAL VALUE:	The sum of all values for this feature within this cluster. This value represents the total of this reporting element for this cluster.
% OF TOTAL:	The percentage of the sum of this feature's values for this cluster compared to the sum for the entire sample population. For example, in the report above, the JOBNLR represented by cluster 1 is 0% of the JOBNLR for the entire sample.

If the INCLUDE SPARSE CLUSTERS options on panel CAPG910U has been set to "NO" by the analyst, the page of the report will contain a listing of all clusters excluded, and their populations.

The following fields are presented in this section:

CLUSTER Cluster numbers are assigned sequentially
NUMBER: to the patterns that are identified by the
 algorithm. Note that the order in which the
 clusters are identified is not an indicator of
 merit.

POPULATION: The population of each excluded cluster.

15.3.4 Data Value Exception Detail Report

This is a simple listing of any observation that contains one or more outlying data values.

The application will determine the key sequence structure of the input file, modify the key structure to remove date and time key values except for the End Timestamp, and present the listing using the modified key sequence. You can optionally exclude the CA MICS account codes from the reports by specifying a value of "NO" to the REPORT USING ACCOUNT FIELDS option on the Executions Options panel. This report can be used to verify that the outliers are truly exceptional, and not merely unusual cases.

Note that each outlying clustering data value is flagged with an "*". This permits you to identify outliers in the data and understand which elements(s) contain abnormal data values.

Workload Characterization Analysis									
Data Value Exception Detail									
For: Monday, June 23, 2003									
Listing of Observations With Outlier Values									
System	Job	Job	Job	TCB	CPU				
Identifier	DIVISION	DEPARTMENT	RESERVED	Group	Identification	End TimeStamp	Time	DASD	EXCPS
XAL1	BNK	01622	*****	7	FCSBK150	07NOV02:10:08:53.48	0:01:49.84*	65,053*	
	DEV	01056	*****	7	LTSIPCSX	14NOV02:17:54:24.33	0:40:00.28*	354,840*	
		01060	*****	5	CYGIPCS	12NOV02:10:46:57.98	0:01:41.88*	178,301*	
		01062	*****	0	AAAAA05M	07NOV02:13:30:34.99	0:00:06.64	90,943*	
				5	BBBBB02D	06NOV02:13:50:11.44	0:02:10.87*	34,038	
					CCCCC04B	12NOV02:15:43:33.43	0:03:32.57*	2,879	
						12NOV02:15:55:19.91	0:01:10.59*	2,702	
						12NOV02:16:10:50.76	0:01:14.38*	2,661	
				7	CCCCC043	14NOV02:12:03:07.41	0:00:24.55	126,664*	
					DDDDD02D	15NOV02:10:46:48.50	0:01:17.18*	15,968	
						15NOV02:13:30:36.48	0:01:14.50*	15,781	
					JSIEXTR	11NOV02:14:33:57.06	0:00:03.56	48,436*	
					EEEE01D	11NOV02:15:27:19.09	0:02:15.77*	25,010	
						12NOV02:10:24:31.11	0:01:23.27*	26,935	
	01064	*****	3	FFFFF06L	04NOV02:14:41:52.90	0:00:16.10	46,024*		
					FFFFF06M	04NOV02:14:42:19.03	0:00:16.12	46,024*	
	01076	*****	0	GGGGG01L	11NOV02:15:31:18.24	0:00:02.36	42,283*		
						11NOV02:15:34:53.13	0:00:02.16	39,770*	
					GGGGG01M	11NOV02:17:08:46.41	0:00:02.49	44,844*	
						11NOV02:17:16:02.77	0:00:02.85	51,817*	
						11NOV02:17:25:56.11	0:00:02.38	44,865*	
						11NOV02:17:46:16.77	0:00:02.12	42,335*	
						11NOV02:17:49:34.96	0:00:04.81	101,604*	
						11NOV02:17:51:30.57	0:00:02.46	48,545*	
						11NOV02:17:53:40.22	0:00:04.31	89,564*	
						11NOV02:17:55:53.45	0:00:02.67	53,444*	
						12NOV02:12:58:13.85	0:00:02.91	50,464*	
				7	HHHHH01Y	07NOV02:15:09:18.12	0:00:31.65	43,576*	
						07NOV02:15:34:08.89	0:00:59.88	79,414*	
						15NOV02:11:34:19.60	0:00:33.91	45,279*	
	01266	*****	7	IIIII04Q	15NOV02:11:00:52.65	0:00:05.13	95,660*		
GIS	00045	*****	0	A02ICST3	12NOV02:11:11:59.69	0:01:31.50*	9,099		
				MPIRSADT	08NOV02:17:22:57.44	0:01:37.44*	5,808		
				TCWFAST\$	11NOV02:11:23:34.38	0:00:17.46	99,448*		
						11NOV02:11:50:11.47	0:00:19.64	156,990*	
				TESTXCBK	14NOV02:11:47:28.60	0:03:38.29*	10,763		
				TOPCDEVL	06NOV02:17:18:56.79	0:07:28.57*	47,494*		
				TOPCDEV2	08NOV02:09:51:53.06	0:05:40.66*	148,789*		
						08NOV02:11:42:22.55	0:01:40.63*	7,267	
						08NOV02:13:41:37.00	0:01:38.81*	7,179	
						08NOV02:13:49:04.83	0:01:33.66*	7,094	
				TOPCQAQA	12NOV02:10:05:58.65	0:03:00.82*	14,794		
						14NOV02:18:24:00.78	0:01:44.87*	25,197	
				TOPDDEV2	13NOV02:10:25:47.61	0:20:52.34*	538,224*		
				TOPDQAQA	06NOV02:11:35:54.43	0:17:50.60*	1,154,523*		
				WKBKVM82	04NOV02:12:20:36.27	0:21:08.57*	225,024*		
				WKBKVM83	04NOV02:18:11:15.80	0:23:50.50*	230,790*		
				WKBKVM84	11NOV02:10:52:19.66	0:20:15.46*	220,572*		
				WKBKVM85	11NOV02:12:50:10.33	0:25:27.51*	220,572*		
Note: '*' Indicates that the data value is an outlier									

Figure 15-4. Data Value Exception Detail report

The Data Value Exception report is dynamically built and does not have a fixed format. The first 'n' columns of the report contain the sequence fields of the CA MICS file used for analysis. These fields are documented in the CA MICS Data Dictionary. The remaining elements are those specified by the user on the Workload Characterization panel as clustering elements. If these are CA MICS data elements, which is the typical case, then their definitions and usage data can be located in the CA MICS Data Dictionary as well.

15.3.5 Sparse Cluster Population Report

The Sparse Cluster Population report is a listing of all clusters that have been defined as "sparse". Sparse clusters are by definition those clusters having less than 0.5% of the population of the sample. Generally, these clusters contain outliers, but may also contain normal cases.

The application will determine the key sequence structure of the input file, modify the key structure to remove date and time key values, except for the End Timestamp, and present the listing using the modified key sequence. You can optionally exclude the CA MICS account codes from the reports by specifying a value of "NO" on the REPORT USING ACCOUNT FIELDS option on the Execution Options panel. This report can be used to verify that the sparse clusters are truly exceptional cases, not merely unusual cases.

Note that each outlying clustering data value is flagged with an "*". This permits you to not only identify outliers in the data, but to understand which elements contain abnormal data values.

Sparse Cluster Population Report									
For: Monday, June 23, 2003									
Listing of Observations Within Cluster									
Cluster Number	System Identifier	DIVISION	DEPARTMENT	RESERVED	Job Group	Job Identification	End Time Stamp	Job TCB CPU Time	DASD EXCPS
2	XAL1	GIS	00045	*****	5	MPGLAUDT	07NOV02:11:39:02.05	0:02:06.60*	7,340
	XAL1	GIS	02046	*****	0	T1PR2NEW	13NOV02:17:37:26.57	0:02:09.01*	108
3	XAL1	GIS	00045	*****	0	A02ICST3	12NOV02:11:11:59.69	0:01:31.50*	9,099
	XAL1	GIS	00045	*****	0	MPIRSADT	08NOV02:17:22:57.44	0:01:37.44*	5,808
	XAL1	GIS	00045	*****	0	TOPCDEV2	08NOV02:11:42:22.55	0:01:40.63*	7,267
	XAL1	GIS	00045	*****	0	TOPCDEV2	08NOV02:13:41:37.00	0:01:38.81*	7,179
	XAL1	GIS	00045	*****	0	TOPCDEV2	08NOV02:13:49:04.83	0:01:33.66*	7,094
	XAL1	UNK	UNK	*****	0	CACMPQA	14NOV02:10:09:47.07	0:01:41.14*	8,383
4	XAL1	GIS	00045	*****	5	MPMINAS	14NOV02:16:45:15.65	0:00:35.54	70,439*
	XAL1	GIS	00045	*****	5	MPMIPACE	14NOV02:10:30:56.85	0:00:40.26	80,646*
5	XAL1	DEV	01076	*****	7	FEESE01Y	07NOV02:15:34:08.89	0:00:59.88	79,414*
	XAL1	GIS	02045	*****	0	BEHMAMDB	11NOV02:15:08:03.50	0:01:18.45*	79,294*
6	XAL1	DEV	01062	*****	5	MCCTI03H	07NOV02:14:32:14.63	0:00:34.83	23,429
	XAL1	DEV	01062	*****	7	DIXPA02M	15NOV02:15:35:10.77	0:00:42.85	26,409
	XAL1	GIS	02045	*****	0	BEHMAMDB	11NOV02:15:31:53.21	0:00:27.57	27,384
	XAL1	GIS	02045	*****	0	BEHMAMEN	11NOV02:14:52:19.61	0:00:24.20	23,585
7	XAL1	GIS	00045	*****	0	TCWFAST\$	11NOV02:11:50:11.47	0:00:19.64	156,990*
	XAL1	GIS	00045	*****	5	STARLOAD	13NOV02:15:58:32.57	0:00:28.49	148,507*
	XAL1	GIS	00045	*****	5	STARLOAD	13NOV02:16:26:38.42	0:00:26.74	148,503*
8	XAL1	GIS	00045	*****	0	WKBKVM85	11NOV02:12:50:10.33	0:25:27.51*	220,572*
	XAL1	GIS	02046	*****	7	MELRI01B	07NOV02:18:12:06.64	0:25:34.91*	233,415*
9	XAL1	DEV	01062	*****	0	BEAMA05M	07NOV02:13:30:34.99	0:00:06.64	90,943*
	XAL1	DEV	01076	*****	0	JERMI01M	11NOV02:17:49:34.96	0:00:04.81	101,604*
	XAL1	DEV	01076	*****	0	JERMI01M	11NOV02:17:53:40.22	0:00:04.31	89,564*
	XAL1	DEV	01266	*****	7	HENR004Q	15NOV02:11:00:52.65	0:00:05.13	95,660*
	XAL1	GIS	00045	*****	0	TCWFAST\$	11NOV02:11:23:34.38	0:00:17.46	99,448*
12	XAL1	DEV	01062	*****	5	JONCA04B	12NOV02:15:55:19.91	0:01:10.59*	2,702
	XAL1	DEV	01062	*****	5	JONCA04B	12NOV02:16:10:50.76	0:01:14.38*	2,661
	XAL1	DEV	01062	*****	7	DIXPA02D	15NOV02:10:46:48.50	0:01:17.18*	15,968
	XAL1	DEV	01062	*****	7	DIXPA02D	15NOV02:13:30:36.48	0:01:14.50*	15,781
	XAL1	GIS	00045	*****	5	VRPDIV	08NOV02:17:27:14.42	0:01:23.85*	9,966
	XAL1	GIS	02046	*****	0	FAIRSCAN	07NOV02:11:39:42.42	0:01:14.27*	13,231
	XAL1	GIS	02046	*****	0	FAIRSCAN	08NOV02:14:40:26.10	0:01:13.71*	13,234
	XAL1	UNK	UNK	*****	0	DBTNGFR2	14NOV02:09:26:28.85	0:01:12.29*	5,774
	XAL1	UNK	UNK	*****	0	T1CAPS2	11NOV02:09:07:14.75	0:01:21.77*	15,217
	XAL1	UNK	UNK	*****	0	T1CAPS2	11NOV02:09:07:14.75	0:01:21.77*	15,217
13	XAL1	DEV	01060	*****	5	CYGIPCS	12NOV02:10:46:57.98	0:01:41.88*	178,301*
	XAL1	SPT	04733	*****	5	BEEST02U	08NOV02:11:58:26.55	0:02:11.93*	173,919*
Note: '*' Indicates that the data value is an outlier									

Figure 15-4A. Sparse Cluster Population report

The Sparse Cluster Population report is dynamically built, and does not have a fixed format. The first 'n' columns of the report contain the sequence fields of the CA MICS file used for analysis. These fields are documented in the CA MICS Data Dictionary. The remaining elements are those specified by the analyst on the Workload Characterization panel as clustering elements. If these are CA MICS data elements, which is the typical case, then their definitions and usage data can be located in the CA MICS Data Dictionary as well.

15.3.6 Cluster Index Analysis Graph

The Cluster Index Analysis graph can produce two different charts, depending on data volumes and the print line size specified in the MICF options:

- o BLOCK chart
- o Horizontal bar (HBAR) chart

In either case, one entry is made for each cluster selected and the clustering index value is displayed.

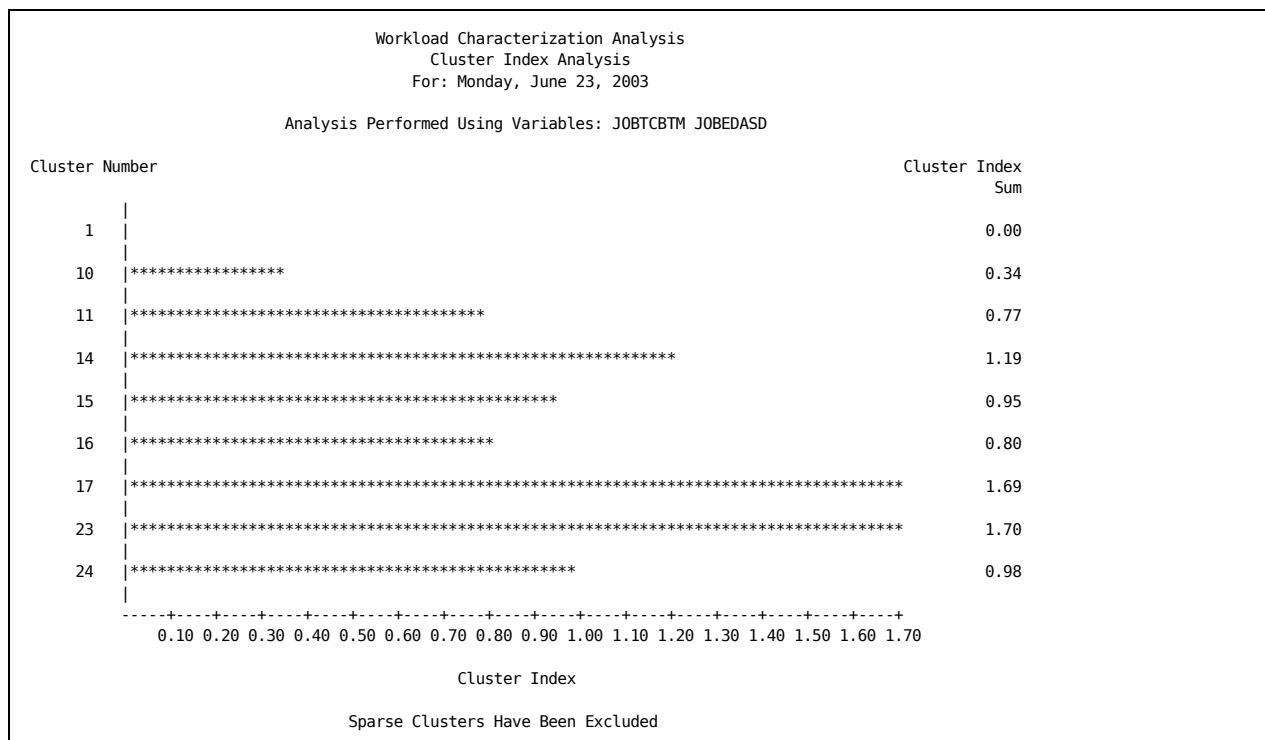


Figure 15-4B. Cluster Index Analysis graph

The clustering index value is the Root Mean Square (RMS) of the distances of all feature values within the cluster. As stated previously, this is a relative measurement of cluster performance and "goodness of fit" of the data, with respect to the cluster centroid values. A lower value represents a better performing cluster, indicating more tightly formed clusters and fewer outlying values.

This graph is a simple, yet effective way to quickly evaluate the overall performance of the clustering operation, and can be used to communicate this performance to non-technical personnel.

15.3.7 Cluster Population Analysis Graph

Cluster Population Analysis is a PIE chart depicting the populations of the top 'n' clusters, where 'n' depends on the printer 'linesize' and 'pagesize' specifications, and the distribution of data values within the sample. For example, the number of 'slices' of the PIE chart may be reduced if one slice is large with respect to the remaining slices.

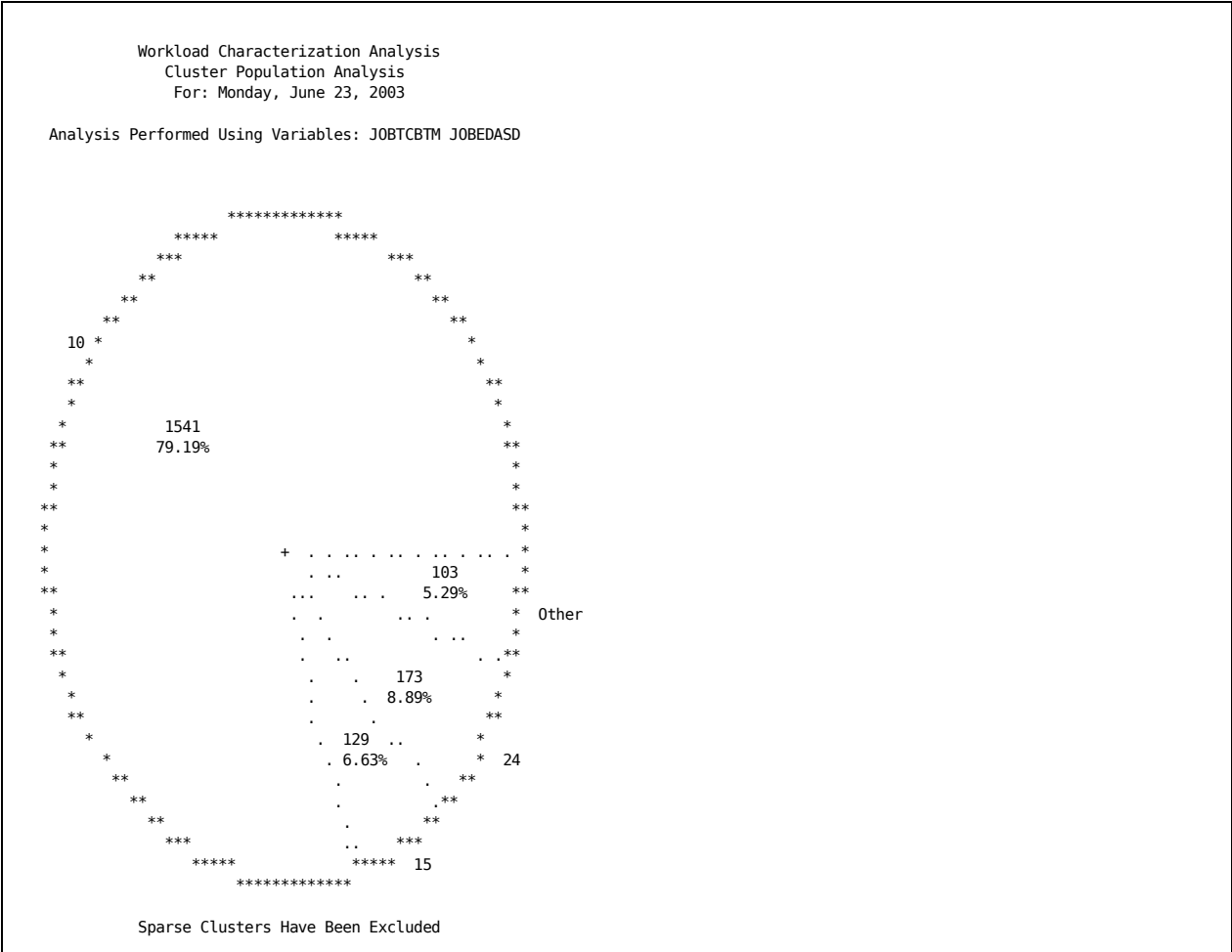


Figure 15-4C. Cluster Population graph

This chart provides a simple, yet effective way of presenting the population distribution of the study data in a high-level summary.

15.4 Analytic Technique Tutorial

This section discusses general Workload Characterization methodology. While this is not a complete discussion of the topic, it does provide an introduction to the statistical techniques employed by the Neugents technology in this process. The tutorials presented in this section concentrate on the following topics:

- 1 - Clustering Algorithms
- 2 - Scaling
- 3 - Algorithm Implementation

15.4.1 Clustering Algorithms

Although there are a wide variety of statistical algorithms that you can use to develop clusters, one of the simplest and most popular algorithms is the K-means algorithm (MAC67). A hybrid of both the hierarchical and the K-means algorithm is employed within Neugents technology for clustering operations.

To understand how clustering algorithms work, you must understand the following definitions.

Resource Vector:	An n-tuple of measurements selected to describe a workload element. For example, you might choose to describe jobs in terms of the three distinct measurements: CPU time, print lines, and tape mounts. Each job would then be represented by a 3-tuple (CPU time, print lines, and tape mounts). This is also sometimes referred to as a usage vector, or simply a vector. Neugents technology uses the term "pattern" when describing such a set of related measurements.
Feature:	A component of a resource vector. In the above example, CPU time, print lines, and tape mounts are the three features of resource vectors describing the workload element, job.
Space:	The n-dimensional coordinate system in which resource vectors are defined.
Feature Mean:	The mean of all values of a given feature taken over a set of resources.

Feature Variance:	The variance (that is, the standard deviations squared) associated with the distribution of values of a given feature, taken over a set of resource vectors. Note that the feature with the greatest variance defines the longest coordinate axis of the space.
Centroid:	The geometric center of the set of all resource vectors contained in a space. The location of the centroid is defined by an n-feature vector. Its component values are the feature means of the resource vectors in the space.
Cluster:	A set of resource vectors in a given space plus their centroid. Note that the centroid of a set of vectors need not be one of the original vectors in the space.
Geometric Distance:	The distance between any two points (that is, resource vectors) in a given space. It is calculated by squaring the difference between the points (that is, the feature values) for each feature, summing the squares over all features, and taking the square root of the resulting sum. The formula for a two-dimension geometric distance is better known as the Pythagorean theorem. A formula for calculating geometric distance is presented in the following first equation.
Radius:	The greatest geometric distance between a centroid and any of its associated resource vectors.
Relative Position:	The ratio of the geometric distance between a given resource vector or pattern and the radius of the cluster, expressed as a value from 0.00 to 1.00. The lower the value of the relative position, the closer the particular resource vector or pattern is to the cluster center.
Objective Radius:	An externally defined parameter that specifies a maximum desired radius for clusters identified by a clustering algorithm.

Splitting: The partitioning of a space into two smaller spaces. A space is split in half along the axis of greatest variance. A clustering algorithm will continue to split resource vector spaces as long as the radius of any cluster in the partition is greater than the specified objective radius.

Consider the following two-dimensional clustering example. Fifteen resource vectors were selected for the example. They are shown in the following table:

VECTOR NUM.	X	Y
=====	=====	=====
1	1	1
2	2	1
3	3	1
4	3	8
5	4	7
6	4	8
7	4	9
8	5	8
9	8	4
10	9	3
11	9	4
12	9	5
13	10	4
14	10	8
15	11	8

Note: The range of values selected for the two features in this example are similar. This avoids the problems that are introduced by scaling considerations. Scaling is discussed in Section 15.4.2. The objective radius for this example is 1.

A plot of the resource vectors is shown in Figure 15-5. The resource vectors are represented by asterisks (*) and the centroid of the space is represented by a pound sign (#).

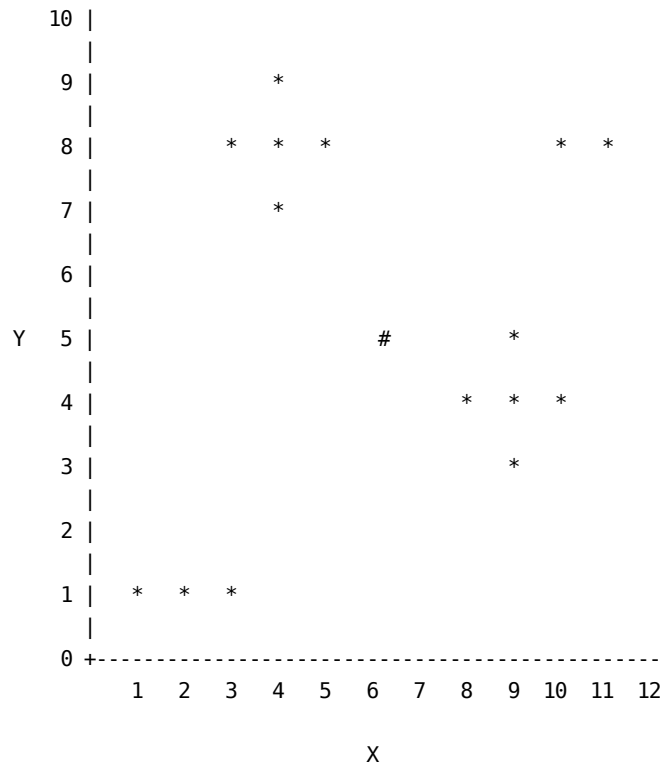


Figure 15-5. Initial Space

The first step in the solution process is to calculate the feature means, variances, and radius for the space. (The centroid is defined by the feature means.) The calculated values are as follows:

SPACE	FEATURE	CENTROID (MEAN)	VARIANCE	RADIUS
all	X	6.13	11.41	6.67
	Y	5.27	8.50	

As mentioned previously, the radius of the space is calculated using the geometric distance equation. This equation states that the distance, D , between the centroid of the k -th cluster, $C(k)$, and the i -th resource vector, $R(i)$, is equal to the square root of the sum of the squares of their feature differences. The general form of the geometric distance equation is shown in Equation 1, where j is the feature number index.

$$D = \sqrt[n]{\sum_{j=1}^n (C(k,j) - R(i,j))^2}^{1/2} \quad (\text{Eqn 1})$$

For example, the farthest resource vector from the centroid of the space shown in Figure 15-5 is the resource vector at (1,1). The calculation of the distance between the resource vector and the centroid is shown in Equation 2.

$$D = \sqrt{(6.13-1.0)^2 + (5.27-1.0)^2}^{1/2} \quad (\text{Eqn 2})$$

$$D = \sqrt{26.32 + 18.23}^{1/2} = 6.67$$

As shown in the previous table, the variance in the X direction is 11.4 and the variance in the Y direction is 8.50. Since the radius of the space is greater than the objective radius, the space is split at the centroid along the axis of greatest variance. This results in two spaces, A and B, as is shown in Figure 15-6.

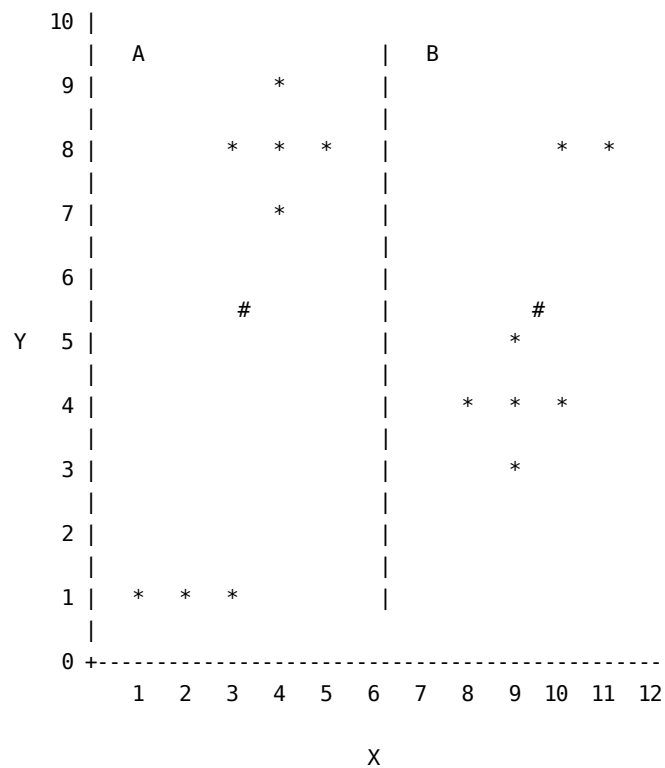


Figure 15-6. First Split

The characteristics of the two spaces resulting from the split are shown in the following table.

SPACE	FEATURE	CENTROID (MEAN)	VARIANCE	RADIUS
A	X	3.25	2.70	4.92
	Y	5.38	13.41	
B	X	9.33	1.07	3.31
	Y	5.14	4.14	

Once again, the radii of the spaces are greater than the objective radius. Hence, each of the spaces is split along the axis of greatest variance. This results in four spaces, A, B, C, and D, as shown in Figure 15-7. Note that in spaces A, C, and D the centroids happen to coincide with a resource vector.

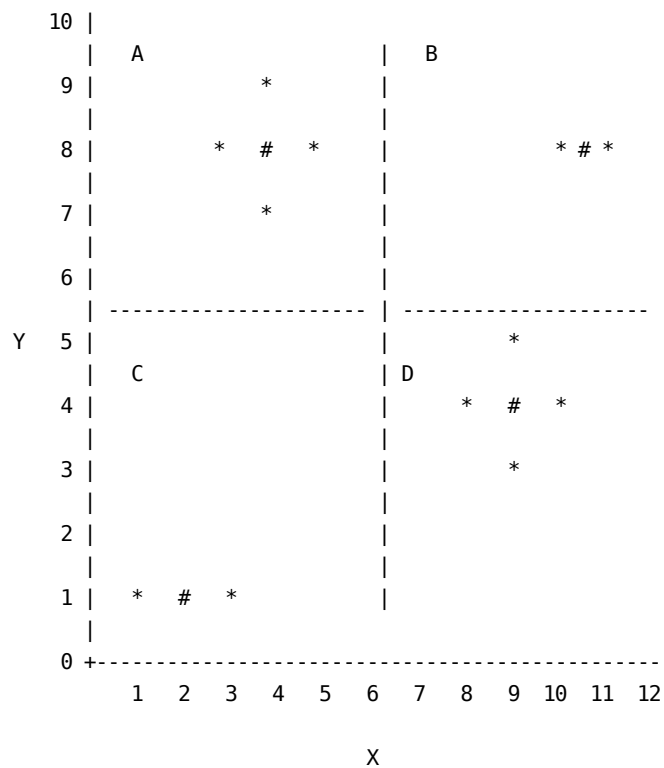


Figure 15-7. Second Split

The characteristics of the four spaces resulting from the split are shown in the following table.

SPACE	FEATURE	CENTROID (MEAN)	VARIANCE	RADIUS
A	X	4.0	0.50	1.0
	Y	8.0	0.50	
B	X	10.5	0.50	0.5
	Y	8.0	0.00	
C	X	2.0	1.00	1.0
	Y	1.0	0.00	
D	X	9.0	0.50	1.0
	Y	4.0	0.50	

Since all of the radii are less than or equal to the objective radii, a solution for the space introduced in Figure 15-7 has been developed. Although this problem is very simplistic, it illustrates the concepts of the K-mean algorithm.

Thus, four clusters have been identified that represent the 15 resource vectors in the space. Although the actual implementation of the clustering algorithm is somewhat more complex, the same principles apply when thousands of workload elements are analyzed with the program.

15.4.2 Scaling

One problem that is introduced by the use of statistical clustering algorithms is scaling. In the clustering example discussed in the previous section, a similar numerical range for the X and Y variables was selected. Unfortunately, actual workload data presents a much wider range of values.

Consider the problem of developing job classes. The job classes in this example are based on CPU minutes and print lines. The maximum value observed for CPU minutes would be far less than 1,000 minutes, while the likelihood of printing more than 1,000 lines is very high. If the differences in these features are squared in the geometric distance equation, the number of CPU minutes used by a job would appear to be insignificant. Therefore, the variables must be scaled prior to the clustering process so that relative differences between the features have a similar influence on the geometric distance calculation.

Neugents technology for clustering operations uses a simple unit scaling technique described next.

UNIT SCALING

Unit scaling is a simple technique for solving the problems introduced by the different range of values associated with each of the resource vector's features. Most simply stated, unit scaling maps the range of values associated with each feature into the range 0 to 1 (AGR76). Equation 3 details this mapping.

$$s_i = \frac{x_i - x_{\min}}{x_{\max} - x_{\min}} \quad (\text{Eqn 3})$$

where: s_i is the scaled value of x_i

x_i is i-th observation of the x feature

$x_{\min}$ is the minimum of all x_i

$x_{\max}$ is the maximum of all x_i

A modified version of UNIT scaling is currently used within Neugents technology whereby data values are scaled internally to a range of -1 to +1.

The potentially poor representation of outliers is one of the primary concerns in the selection of features for the clustering process. Criteria for selecting features for clustering are discussed in Section 15.2.4.

15.4.3 Algorithm Implementation

The workload characterization program is comprised of these major steps:

- 1 - Select data elements (features) from the CA MICS file based on user input.
- 2 - Execute Neugents technology training phase, which examines the input data and builds a set of clusters using its own proprietary techniques.
- 3 - Neugents technology then executes the "consulting" phase, where all observations (patterns) are consulted against the trained model and each observation or pattern is assigned to its final cluster. Outliers are marginally assigned during this process.
- 4 - Clustered data, along with model performance statistics are returned to the user.

The first step in the program is the selection of the user-specified elements from a CA MICS database file. The file and database elements (features) are specified with parameters on the Workload Characterization screen. The control parameters are discussed in Section 15.5.

The second step is to perform the training phase, in a manner not unlike that done in the Workload Forecasting. A model of clusters is produced from the input data using an iterative process involving a combination of hierarchical and "k-means" methods. The actual process involved is proprietary, but permits accurate and rapid cluster creation. A Relative Position value is calculated for each observation or pattern in the manner previously described for cluster evaluation purposes, and the model is then ready for the final phase.

During the next phase, the input data is reprocessed and "consulted" against the trained model. Each observation is now assigned to a permanent cluster and any outliers are marginally assigned, meaning they are put with a cluster that best describes the vector data contents.

Lastly, the final clustered data, along with model performance statistics is made available to Workload Characterization application.

15.5 Component Operation

This section describes the screens and parameters you can use to control workload characterization. The control parameters allow you to define the CA MICS file you wish to use as input, the time periods you wish to specify, and the variables you wish to use in the clustering process. Each of the control screens is discussed in the following sections.

- 1 - Workload Characterization Screen
- 2 - Extended Options Menu
- 3 - Global Data Selection Screen
- 4 - Data Element Derivation Screen
- 5 - Alternate Input File Screen
- 6 - Save Selected Input Observations Screen
- 7 - Save Cluster Output Observations Screen
- 8 - Data Filtering Parameters Screen
- 9 - Clustering Execution Parameters Screen
- 10 - Report Description Elements Screen

15.5.1 Workload Characterization Screen

Figure 15-8 illustrates the Workload Characterization screen.

```

/----- Workload Characterization using Neugents Technology -----\
| Command ==> |
| |
| Composing CA MICS Inquiry: CLUCIO - Workload Characterization |
| |
| CA MICS file ==> - |
| Time-span: DETAIL | Cycle(s) ==> __ - __ |
| Data base ==> _ |
| Allow Run-Time Override of Data Base ==> Y (Y/N) and/or Cycle(s) ==> Y (Y/N) |
| |
| Analysis based on ==> _____ |
| |
| Report for each cluster ==> _____ |
| |
| |
| Data Selection: | Allow run-time override |
| Date ==> _____ - _____ - _____ ==> N (Y/N) |
| Hour ==> __ - __ - __ - __ ==> N (Y/N) |
| Zone ==> _____ ==> N (Y/N) |
| SYSID ==> _____ ==> N (Y/N) |
| |
| Specify extended options ==> N (Y/N) |
\-----/

```

Figure 15-8. Workload Characterization Screen

The Workload Characterization screen contains the following fields:

CA MICS file: The CA MICS file that contains the data elements that you wish to include in your workload characterization analysis. If you are uncertain of the correct file name, type a question mark (?) in this field to obtain a complete selection list for your installation.

Cycle(s): The cycle or cycles of the input file that are used in the analysis. For most installations, there is one cycle of DETAIL data for each day. You may specify a maximum of 21 cycles.

Database:	The database unit identifier(s)(up to five units) that contain the specified CA MICS file. An eight-character database unit name is displayed after you specify this identifier. If you are uncertain of which database unit the file resides in, type a question mark (?) in this field to obtain a complete selection list for your installation.
Analysis based on:	The features (data elements) that are analyzed by the clustering procedure. You can define a maximum of ten features. Cluster variables must be numeric. These data elements may either be CA MICS data elements or non-CA MICS data elements. If you specify non-CA MICS data elements, you must provide the SAS code to derive them via the exit available on the Extended Options Menu, Figure 15-9.
Report for each cluster:	The data elements that will be detailed in the Cluster Population Summary report (see Section 15.3.3). You can specify any numeric data elements that you find informative in the reports. For example, clusters for a job class study might be developed using CPU time, tape mounts, and print lines as features. You might also wish to report on tape EXCPs, tape drive occupied hours, disk EXCPs, and average elapsed time of the jobs assigned to the clusters. A maximum of ten report variable names can be specified. Report variables must be numeric. These data elements may be either CA MICS data elements or non-CA MICS data elements. If non-CA MICS data elements are specified, you must provide the SAS code to derive them via the exit available on the Extended Options Menu, shown in Figure 15-9.
Date selection:	The selection criteria for Date, Hour, Zone, and SYSID to subset the data you are examining. You must enter date values in the ddmonyy format.
Specify extended options:	The specification for selecting extended options. Select this option if you wish to specify additional parameters.

Typing a Y at the "Specify extended options" prompt on the Workload Characterization screen displays the Extended Options Menu shown in Figure 15-9.

```

/----- Extended Options Menu -----\
|Command ==>
|
|Inquiry Step:  Workload Characterization
|
|
|Specify global data selection exit          ==> N (Y/N/R)
|Specify non-CA MICS data element derivation ==> N (Y/N/R)
|
|Specify alternate input file                ==> N (Y/N/R)
|Save selected input observations             ==> N (Y/N/R)
|Save cluster output observations            ==> N (Y/N/R)
|
|Specify data filtering parameters           ==> N (Y/N/R)
|Specify clustering execution options        ==> N (Y/N/R)
|Define report description element names     ==> N (Y/N/R)
|
|
|
|
\-----/

```

Figure 15-9. Extended Options Menu

You may select any of the options available on this screen by typing a Y (YES) at any of the prompts. Typing an R (RESET) resets the option and nullifies any specifications you made for that option.

The default for the SELEOF macro is null. This macro is invoked when the workload observations are read from the specified CA MICS file. You may use it to delete observations from the data as it is selected. For example, in a job class study, you might wish to exclude production job classes from the analysis. If production jobs are designated by job class P, then you would code the following:

```
%MACRO SELEOF;  
  IF JOBCLASS='P' THEN GO TO T_EOF;  
%MEND SELEOF;
```

Note that you do not have to code the %MACRO and %MEND statements, as they are already provided for you.

The IF statement should specify branching to T_EOF rather than acting as a subsetting IF or specifying DELETE. A subsetting IF or a DELETE statement may cause the program to function incorrectly.

You must code each specified SAS statement in accordance with the syntax of the SAS Macro Language, as described in the SAS manuals.

15.5.5 Alternate Input File Screen

Figure 15-12 illustrates the Alternate Input File screen.

```
/----- Alternate Input File -----\
|Command ==>
|
|Inquiry Step: Workload Characterization
|
|SAS file name(s) ==> _____
|
|
|SAS Data set allocation instructions:
|   Data set name    ==> _____
|   Disposition      ==> SHR  (SHR/OLD)
|
|
|
|
```

Figure 15-12. Alternate Input File Screen

Workload characterization has the ability to read data from either CA MICS files or external SAS files. The Alternate Input File screen allows you to specify that the input data for the workload characterization is a SAS file or files, and input the SAS data set name and up to five SAS files to use as input to the analysis. Note that the SAS data must already be created.

15.5.6 Save Selected Input Observations Screen

Figure 15-13 illustrates the Save Selected Input Observations screen.

[illegible]

Figure 15-13. Save Selected Input Observations Screen

Workload characterization has the ability to save the observations that your workload characterization analysis inquiry selected to an external SAS file for further processing by independent SAS programs, or for future processing by the workload characterization software. The Save Selected Input Observations screen allows you to specify that the selected observations are to be written to an external SAS file, and to specify the SAS data set name and the name of the SAS file to which they are to be written. Note that the SAS data set must already be created.

15.5.7 Save Cluster Output Observations Screen

Figure 15-14 illustrates the Save Cluster Output Observations screen.

[illegible]

Figure 15-14. Save Cluster Output Observations Screen

Workload Characterization has the ability to save the SAS cluster observations from the SAS clustering program to external files for additional processing. The Save Cluster Output Observations screen allows you to specify the SAS data set and the name of the SAS file to which the cluster observations are to be written. Note that the SAS data set must already be created.

15.5.9 Clustering Execution Options Screen

Figure 15-15A illustrates the Clustering Execution Options screen.

```

/----- Execution Options for Neugents Technology ----- ROW 1 OF 5 \
|
|Inquiry Step:  Workload Characterization Using Neugents Technology
|
|      Clustering Input Data ==> SAMPLE      (SAMPLE/POPULATION)
|      Clustering Method   ==> RADIUS      (RADIUS/NUMBER OF CLUSTERS)
|      Maximum Cluster Radius ==> 3.0      (1.0-3.5) If RADIUS Specified
|      Target Number of Clusters ==>      (03-64)  If CLUSTER Specified
|
|      Training Sample Size ==> 2000      (nnnnnn/ALL)
|      Sample Trim Limit   ==> 02.5      (0.00-99.9)  Default: 2.5 Pct.
|
| Include Outliers in Clusters ==> YES      (YES/NO)  Trim 02.5 Pct.
| Include Sparse Clusters ==> YES      (YES/NO)  LT 0.5 Pct. of Obs.
|
|      Report Cluster Population ==> NO      (YES/NO)
|      Report Using Account Fields ==> YES      (YES/NO)
|      Report Outliers Separately ==> NO      (YES/NO)
|      Report Sparse Clusters ==> NO      (YES/NO)
|      Create Cluster Index Graph ==> NO      (YES/NO)
|      Create Population Graph ==> NO      (YES/NO)
|
\-----/

```

Figure 15-15A. Clustering Execution Options Screen

This screen is used to enter the execution options for the clustering process and contains the following fields:

CA MICS file:	The CA MICS file that contains the data elements that you wish to include in your workload characterization analysis. If you are uncertain of the correct file name, type a question mark (?) in this field to obtain a complete selection list for your installation.
Clustering Input Data:	Either SAMPLE or POPULATION, default is SAMPLE. This parameter informs the application to select either a SAMPLE of the input data, or to use the entire POPULATION for processing.

Clustering Method:	Either RADIUS or CLUSTER, default value is RADIUS. The clustering operation will be controlled by this option. If RADIUS is specified, then clustering will be limited to a certain cluster size or RADIUS. If CLUSTER is chosen, then clustering will produce only a certain number of clusters, but the clusters will vary more in size and can be quite large in scope.
Maximum Cluster Radius:	A range from 1.0 to 3.5, with a default value of 3.0. The Maximum size of the largest cluster to be formed, is specified in terms of standard deviations. Clustering will stop once any cluster exceeds this value. Use of this option will cause more clusters to be formed but will generally result in a better 'fit' of the data.
Target Number of Clusters:	If CLUSTER is specified for the Clustering Method, then a number from 3 to 64 must be entered here, with the default value being 8. This is the number of clusters that will be produced in a particular execution. This options is useful when using clustering as a data analysis tool where you might want to simply partition the data into heterogeneous groups, but are not concerned about the size or exact content of the clusters.
Training Sample Size:	A numeric value up to 999,999 or the word 'ALL', if ALL data is chosen for the sample. The default value is 2,000 observations. The actual sample size is computed to be 2% of the sample or 2,000 cases, whichever is greater.
Sample Trim Limit:	A percentage of the sample to be removed, or 'trimmed' to improve the normality of the data and allow for better descriptive statistics to be obtained. The default value is 2.5% and will cause those observations that contain the top 2.5% of values for each cluster feature to be removed. This should effectively remove observations that contain very large values which tend to distort descriptive statistics and should improve the shape and fit of the data within the clusters.

Include Outliers in Clusters:	YES/NO, with a default value of YES. This option indicates if observations are considered outliers, when they contain one or more feature values that exceed the trim limit. Including these observations improves the quality of the study by ensuring that all sampled data will be accounted. Outliers may be excluded if they have been reviewed and a determination is made that they are not relevant.
Include Sparse Clusters:	YES/NO, with a default value of NO. This option controls whether sparse clusters are included for subsequent reporting or other post-processing operations. Sparse clusters are defined as those clusters that contain less than 0.5% of the sample population. Note, if you select the Default value of NO, sparse clusters are still created and the data remains available, but is merely excluded from various reports and graphs.
Report Cluster Population:	YES/NO, with the default being NO. This option lets you decide if you want to generate this report. The report contains detailed usage statistics for each included cluster, with both cluster features and optional reporting elements being reported.
Report Using Account Fields:	YES/NO, with the default value being YES. Controls whether user defined CA MICS account fields are included in detail reports. Inclusion of these fields improves the content and usability of detailed reports, but can result in a print line overflow and distortion of the report format.
Report Outliers Separately:	YES/NO, with the default value being NO. Controls whether to generate a detail report of all observations that contain one or more outlier cluster feature values. This report can be used to review and investigate outliers and their possible effect on the clustering operation.

Report	YES/NO, with the default value being NO.
Sparse	Controls whether to generate a detail report of
Clusters:	all sparse clusters and their observations. This report can be used to better understand the sparse clusters and to determine whether their content should be considered for additional processing.
Create	YES/NO, with the default value being NO.
Cluster	Controls whether to generate a printer graph of
Index	the cluster index values. This graph provides
Graph:	a high-level pictorial overview of cluster performance for the study.
Create	YES/NO, with the default value being NO.
Cluster	Controls whether to generate a printer
Population:	graph of the cluster population. This graph provides a high-level pictorial overview of cluster population for the study.

15.5.10 Report Description Elements Screen

Figure 15-16 illustrates the Report Description Elements screen.

```

/----- Report Description Elements ----- ROW 1 OF 5\
|Command ==>                               Scroll ==> CSR |
|
|Inquiry Step: Workload Characterization
|
|Line cmds: I Insert  D Delete  M Move
|
|Cmd Element   Data element label (long name)           Warning
| - - - - -   - - - - - - - - - - - - - - - - - - - - -
| - - - - -
| - - - - -
| - - - - -
| - - - - -
| - - - - -
|
\-----/

```

Figure 15-16. Report Description Elements Screen

The Report Description Elements screen allows you to specify additional descriptive data elements to be included in the optional SAS file that can be created by workload characterization (see Section 15.5.7). Unlike the cluster, zero deletion, and report variables, the descriptive variables need not be numeric. For example, in a job class study you might wish to include additional data elements like JOBCCLASS (job class), JOB (job name), and RDRTS (reader time stamp) for use in your SAS programs that process the optional SAS File.

15.6 Case Studies

The Case Studies Section presents case studies to illustrate how you can apply the clustering methodology to solve capacity planning and performance management problems that many data processing installations commonly encounter.

The batch job classes case study shows how you can use clustering to investigate patterns in the batch jobs being processed. You can use information to develop workload classes which can then form a reasonable basis for service objectives. The statistical uniformity of the classes enables you to assess the degree to which the service objective can be met. You can use the same information to define a job class structure for establishing an initiator scheme that is consistent with the service objectives. A full set of reports and an explanation for each report is provided with this study.

The case study is shown in the following sections:

- 1 - Batch Job Class Study

15.6.1 Batch Job Class Study

The Batch Job Class case study applies the workload characterization program to identify job classes for an installation. The following topics are addressed by section:

- 1 - Problem Description
- 2 - Control Parameters
- 3 - Report Interpretation

15.6.1.1 Problem Description

This problem asks you to determine a new set of resource-limited job classes for an installation. Job class structures are generally established to meet two objectives (FER78, ART79a):

- o Provide operational controls
- o Facilitate scheduling and implement service policies

You usually define job classes to meet the first objective for reasons other than resource requirements. For example, online systems like IMS and CICS are often run in unique job classes to avoid interference with other workloads. Also, many computer centers use one or two job classes to segregate production jobs from the remainder of their workload.

Job classes established to meet the second objective provide users with alternative classes for job assignments based on resource requirements. Typically, the service objective of a job class decreases (that is, jobs have longer turnaround times) as its resource requirements increase. Using a job class structure that focuses on scheduling and service has a number of potential benefits:

- o Many jobs can be completed in a short time (assuming demand) since the job classes representing smaller jobs may be expedited.
- o The availability of better turnaround time for the smaller classes may encourage users to conserve resources.
- o Job scheduling is simplified since each job class represents a homogeneous set of jobs.

Clustering is an ideal approach to this problem since we want to exploit the natural characteristics of the workload. If we arbitrarily select job classes, we might select classes that split normal resource consumption patterns or choose some that are too large or small to be useful.

For this sample installation, we investigated job classes based on CPU time and tape drives usage. We selected these resources to quantify the system, output, and operator requirements presented by the jobs. Additionally, several other reporting metrics were included to quantify the resources represented by each cluster. Although a number of other features could have been selected for analysis, only these were chosen to keep the resulting job class structure as simple as possible.

15.6.1.2 Control Parameters

Figure 15-18 illustrates the Workload Characterization screen.

```

/----- Workload Characterization Using Neugents Technology -----\
|Command ==>|
|
|Composing CA MICS Inquiry: CLASS - Batch Job Class Case Study|
|
|CA MICS file ==> JOB -|
|Time-span:      DETAIL                      Cycle(s) ==> 01 - 10|
|Data base ==> A|
|Allow Run-time Override of Data Base ==> Y (Y/N) and/or Cycle(s) ==> Y (Y/N)|
|
|Analysis based on      ==> JOBTCBTM  TAPEUSE_  _____  _____  _____|
|
|Report for each cluster ==> JOBEDASD  JOBETAPE  JOBTOHRS  JOBNLR  _____|
|
|
|Data Selection:                      Allow run-time override|
|Date      ==> _____ - _____ - _____ ==> N (Y/N)|
|Hour      ==> __ - __ - __ - __ - __ ==> N (Y/N)|
|Zone      ==> 1 2 3 4 _ _ _ _ _ ==> N (Y/N)|
|SYSID     ==> SYSA _____ ==> N (Y/N)|
|
|Specify extended options ==> Y (Y/N)|
\-----/

```

Figure 15-18. Workload Characterization Screen Control Parameters

The Workload Characterization screen contains the following fields:

CA MICS file:	The CA MICS file to be used for workload characterization. (For this example, we specified the CA MICS BATJOB File for a job class study.)
Cycles:	The number of cycles to be specified. (For example, we specified ten cycles. In an actual study, analyze as much data as possible before you initiate a new job class structure.)
Database:	The single-character database identifier for this file at your site.
Analysis based on:	The CA MICS data element feature names (CPU Time and Tape Drive Usage) to be used in the clustering procedure. Because TAPEUSE is a user supplied element that is used to summarize tape device usage, it is defined on the Non-CA MICS data derivation exit panel and uses the DRVMAC macro to populate the data values. (As discussed in Section 15.2.4, you should specify as few features as possible.) DASD EXCPs, Tape EXCPs, Tape Drive Occupancy Time, and print lines are specified for reporting purposes.

```

/----- Extended Options Menu -----\
|Command ==>
|
|Inquiry Step:  Workload Characterization
|
|
|Specify global data selection exit          ==> N (Y/N/R)
|Specify non-CA MICS data element derivation exit==> Y (Y/N/R)
|
|Specify alternate input file                ==> N (Y/N/R)
|Save selected input observations            ==> N (Y/N/R)
|Save cluster output observations           ==> Y (Y/N/R)
|
|Specify data filtering parameters          ==> Y (Y/N/R)
|Specify clustering execution options       ==> Y (Y/N/R)
|Define report description element names    ==> N (Y/N/R)
|
|
|
\-----

```

Figure 15-19. Workload Characterization Extended Options Menu

You may select additional option specifications on the Extended Options Menu. Figure 15-19 shows an example with four of the extended options specified.

```

/----- SAS Statements -----\
|Command ==>                               Scroll ==> CSR |
|
|Inquiry Step:  Workload Characterization
|
|Line Cmds: I Insert  D Delete  R Repeat  M Move  C Copy
|
|Cmd  SAS Statements for:  Specifying data element derivation
|  - -----
|  - %MACRO DRVMAC;
|    IF JOBMXNTA EQ 0 THEN TAPEUSE = 0;
|    ELSE IF 01 LE JOBMXNTA LE 3 THEN TAPEUSE = 1;
|    ELSE TAPEUSE = 2;
|    LABEL TAPEUSE = "Tape Device Usage Indicator";
|  - %MEND DRVMAC;
|***** BOTTOM OF DATA *****
|
|
|
\-----/

```

Figure 15-20. User-Defined Data Element Derivation Macro

The SAS statements that compute the user-defined data elements are placed in the DRVMAC macro shown in Figure 15-20.

```

/----- Save Cluster Output Observations -----\
|Command ==>
|
|Inquiry Step:  Workload Characterization
|
|SAS file name      ==> ALLOBS
|
|SAS data set name  ==> 'your.dataset.name'
|
|
|
\-----/

```

Figure 15-21. External Report Output File Definition

In this case study we wanted to create the SAS data set name, ALLOBS, that contains the cluster analysis output. To do this, we specified ALLOBS as the SAS file name on the External Report Output File Definition screen shown in Figure 15-21.

```

/---      Data Filtering Options For Neugents Technology      ----  ROW 1 OF 1\
|Command ==>                                         Scroll ==> CSR |
|
|Inquiry Step:  Workload Characterization
|
|  Please enter the names of elements that should be checked for zero
|  values. Observations with elements specified below that contain zero
|  values will be subsequently discarded from further processing.
|
|Line Cmds:%I+Insert %D+Delete %M+Move
|
|Cmd Element      Data element label (long name)          Warning
| -  - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - -
| _  JOBTCBTM
| -
| -
|***** BOTTOM OF DATA *****
|
|
|-----/

```

Figure 15-22. Data Filtering Options

We excluded jobs that used no CPU time from the study, since they did not execute due to a JCL error, or some other abnormal condition. To exclude jobs with no CPU usage, specify the data element name, JOBTCBTM, on the Data Filtering Options for Neugents Technology screen shown in Figure 15-22.


```

/----- Execution Parameters for Neugents Technology ----- ROW 1 OF 5\
|
|Inquiry Step: Workload Characterization Using Neugents Technology|
|
|      Clustering Input Data ==> SAMPLE      (SAMPLE/POPULATION)|
|      Clustering Method   ==> RADIUS      (RADIUS/NUMBER OF CLUSTERS)|
|      Maximum Cluster Radius ==> 3.0      (1.0-3.5) If RADIUS Specified|
|      Target Number of Clusters ==>        (03-64) If CLUSTER Specified|
|
|      Training Sample Size ==> 2000      (nnnnnn/ALL)|
|      Sample Trim Limit   ==> 02.5      (0.00-99.9) Default: 2.5 Pct.|
|
| Include Outliers in Clusters ==> YES      (YES/NO) Trim 02.5 Pct.|
| Include Sparse Clusters ==> YES      (YES/NO) LT 0.5 Pct. of Obs.|
|
| Report Cluster Population ==> YES      (YES/NO)|
| Report Using Account Fields ==> YES      (YES/NO)|
| Report Outliers Separately ==> YES      (YES/NO)|
| Report Sparse Clusters ==> NO      (YES/NO)|
| Create Cluster Index Graph ==> NO      (YES/NO)|
| Create Population Graph ==> NO      (YES/NO)|
|
|-----/

```

Figure 15-23. Execution Parameters Screen

Using the default execution parameter values for the study will normally provide good results, especially for a first time execution. For this case study, we will generate the Cluster Population Summary and Data Value Exception Detail reports, and include both outliers and sparse clusters in both the analysis and in all reports. This should provide a good overall picture of the data under study and help determine how to apply the clusters in the job class structure. All of these options are presented in Figure 15-23 above.

```

/----- Report Description Elements ----- ROW 1 OF 2\
|Command ==>                               Scroll ==> CSR |
|
|Inquiry Step: Workload Characterization                |
|
|  Note: The application will automatically insert all key sequence
|         elements relevant to the CA MICS file being analyzed into
|         the clusters for use in optional reporting.  There is no need
|         to include any key sequence elements.
|
|Line cmds: I Insert  D Delete  M Move
|
|Cmd Element      Data element label (long name)        Warning
| - - - - -      - - - - - - - - - - - - - - - - - - -
| _  JOBCCLASS
| -
|***** BOTTOM OF DATA *****
|
\-----/

```

Figure 15-24. Report Description Elements Screen

Since we wished to create an optional external report data set during the study, we also specified two descriptive variables. We selected the variable JOBCLASS (job class), so that the results of the cluster analysis could be compared with the current job class structure and the assignment of particular jobs could be determined. These data elements are specified on the Report Description Elements screen shown in Figure 15-24.

```

/----- Inquiry Processing -----\
|Command ==>|
|           |
|           |
|   Active MICF Inquiry:  CLASS  - Batch Job Class Case Study   |
|           Author:      Charles Frost                         |
|           |
|   Specify Inquiry processing ==> EXECUTE (EXECUTE/MODIFY/SAVE/CANCEL)|
|   EXECUTE - Execute the inquiry.                               |
|   MODIFY  - Redisplay the inquiry for modification.           |
|   SAVE    - Exit from inquiry composition and save the inquiry. |
|   CANCEL  - Exit from inquiry composition without saving the inquiry. |
|           |
|Instructions:|
|           |
|   Specify desired option and press ENTER key.                 |
|           |
|           |
\-----/

```

Figure 15-24A. Inquiry Processing

After a workload characterization analysis inquiry is created, the Inquiry Processing screen shown here is displayed. From this screen, you can execute the newly created inquiry, modify it, save it without executing it, or cancel the parameter specifications you just entered.

15.6.1.3 Report Interpretation

The example we discussed in Section 15.6.1.2 showed you how to use the control screens to execute Workload Characterization. The program produces five reports and two graphs, of which four will be used in this study.

- 1 - Descriptive Statistics Report
- 2 - Cluster Performance Summary
- 3 - Cluster Population Summary
- 4 - Data Value Exception Detail report

Additionally, a separate section will explain how you can further analyze the output data produced during clustering, and explore possible job class scenarios derived from the case study.

15.6.1.3.1 Descriptive Statistics Report

The Descriptive Statistics Report provides a summary of the statistics that are calculated for the cluster features. Figure 15-25 illustrates a sample Descriptive Statistics Report.

Workload Characterization Analysis						
Clustering Input Descriptive Statistics						
For: Thursday, June 19, 2003						
Summary for Total Sample:						
Number of Sample Observations:		2,000				
Feature	Description	Minimum	Maximum	Average	Std. Dev.	CV
JOBTCBTM	Job TCB CPU Time	0:00:00.01	0:40:00.28	0:00:13.14	0:01:42.89	7.83
JOBEDASD	Total EXCPs	1.00	1,154,523.00	5,992.30	34,582.08	5.77
Summary for Trimmed Sample:						
Number of Sample Observations:		1,919				
Feature	Description	Minimum	Maximum	Average	Std. Dev.	CV
JOBTCBTM	Job TCB CPU Time	0:00:00.01	0:01:00.81	0:00:03.74	0:00:07.71	2.06
JOBEDASD	Total EXCPs	1.00	38,096.00	2,505.00	4,508.48	1.80

Figure 15-25. Descriptive Statistics Report

Note that the report has two sections: one for the original sample randomly taken from the input data; and the other, the sample data after trimming, using the Sample Trim Limit value specified by the user on the Execution Options panel.

The Descriptive Statistics Report contains the following fields:

OBSERVATIONS: The number of observations selected for processing.

FEATURE: The feature name. The feature names that are listed in the report correspond to the features that are specified on the Workload Characterization screen. In Neugents technology, this is also called a "pattern".

DESCRIPTION: The SAS label of the selected data element, from either the CA MICS GENLIB definition or supplied by the user.

MINIMUM: The minimum value observed for the feature in the sample.

MAXIMUM: The maximum value observed for the feature in the sample.

AVERAGE: The average calculated for the feature observations in the sample.

STD DEV: The standard deviation calculated for the feature observations in the sample.

CV: The coefficient of variation. The CV is calculated by dividing the standard deviation by the average.

15.6.1.3.2 Clustering Performance Summary Report

This report is the anchor of the Workload Characterization application and provides a complete overview of the results of the clustering process. You can use this report to make a quick determination of how well the operation performed and to better understand how each cluster feature contributed to the overall performance of the process. Additionally, the execution options and operating parameters used in the process are included to provide a record of the complete operation.

Workload Characterization Analysis	1
Cluster Performance Summary	
For: Monday, July 1, 2003	
Clustering Execution Options	
Cluster Input Data: SAMPLE	
Clustering Method: RADIUS	
Maximum Cluster Radius: 3.0	
Training Sample Size: 2000	
Sample Trim Limit: 97.5 Pct.	
Sampling Percentage : 2.0 Pct.	
Include Outliers in Clusters: YES	
Include Sparse Clusters: NO	
Sparse Cluster Limit: 0.5 Pct.	
Report Cluster Population: YES	
Report Using Account Codes: YES	
Report Sparse Clusters : NO	
Report Outliers Separately: NO	
Create Cluster Index Graph: NO	
Create Population Graph: NO	
Input CA MICS Files: BATJOB	
Input Dataset Name: 'COGDA01.AUDIT.BATJOB'	

Cluster Feature Contents

Feature Name	Description	-----Outlier Limits-----		Cases
		Lower	Upper	
JOBTCBTM	Job TCB CPU Time	0:00:00.00	0:05:40.49	50
TAPEUSE	Tape Device Usage Indicator	0.00	1.00	1

=====

Note: A given case may be classified as an outlier more than once if multiple analysis elements contain abnormal values. Therefore, the totals in this report section will not necessarily agree with those of the Population section.

Workload Characterization Analysis

2

Cluster Performance Summary

For: Monday, July 1, 2003

Cluster Population Summary

Cluster	Radius	Normal Obs.	Outlying Obs.	Total Obs.	% of Population	Clustering Index
1	0.00	1	13	14	0.70	0.00
2*	0.00	0	2	2	0.10	0.00
3*	0.17	0	2	2	0.10	0.17
4*	0.20	0	2	2	0.10	0.20
5*	0.36	5	0	5	0.25	0.21
6*	0.65	0	2	2	0.10	0.65
7*	0.69	0	2	2	0.10	0.69
8*	1.06	0	6	6	0.30	0.82
9	1.19	1,840	0	1,840	92.00	0.19
10*	1.20	6	0	6	0.30	0.82
11	1.44	31	0	31	1.55	0.79
12*	1.44	1	6	7	0.35	1.12
13	1.71	12	0	12	0.60	0.66
14	2.02	46	0	46	2.30	0.72
15*	2.44	0	3	3	0.15	1.74
16*	2.88	0	7	7	0.35	1.58
17*	2.91	0	6	6	0.30	1.81
18*	2.99	7	0	7	0.35	1.57
		=====	=====	=====	=====	
		1,949	51	2,000	100.00	

'*' Indicates that the cluster is sparsely populated.

Sparse clusters are defined as having a population that is less than 0.5% of the total population. In this study, the sparse cluster population limit is 10 cases.

Figure 15-26. Clustering Performance Summary report

In the above report, cluster 9 accounts for nearly 92 percent of all jobs. The jobs represented by this cluster are the bulk of the work performed by this LPAR, and any new job class structure should ensure that these jobs receive priority in terms of the number of initiators. Additionally, clusters 11 & 14 contain significant numbers of jobs and will need to be considered in any new job class scheme as well.

15.6.1.3.3 Clustering Population Summary Report

This report presents information regarding the content of each cluster, including the number of both normal and outlying observations, the statistical summary of each cluster feature, and the usage summary of each of the reporting element.

The report is normally used to better understand how the sample data is distributed among the clusters, how each cluster is actually functioning, and what data values are represented by each cluster.

If Sparse Clusters are not included, then a separate report is produced at the bottom depicting those clusters that were excluded and their respective populations.

Workload Characterization Analysis							1
Cluster Population Summary							
For: Tuesday, July 1, 2003							
Cluster: 1							

Radius: 0.00 Clustering Index: 0.00 Maximum Cluster Radius: 3.0 STDs							
Normal Obs.:	1	0.05%	Observations << 97.5 Pct.				
Outlying Obs.:	13	0.65%	Observations >> 97.5 Pct.				
Total Obs.:	14	0.70%					
Cluster Feature Resources							
Feature	Description	Minimum Value	Average Value	Standard Deviation	Maximum Value	Total Value	% of Total
JOBTCBTM	Job TCB CPU Time	165	6,232	7,031	18,090	87,253	61.43
TAPEUSE	Tape Device Usage Indicator	0	0	1	2	3	3.13
Cluster Report Resources							
Feature	Description					Total Value	% of Total
JOBEDASD	DASD EXCPS					45,230,910	64.98
JOBETAPE	Tape EXCPS					1,354,117	2.47
JOBTOTHS	Tape Occupancy Hours					4	2.86
JOBNLR	Total Logical Writer Records					0	0.00

Cluster: 9							

Radius: 1.19 Clustering Index: 0.19 Maximum Cluster Radius: 3.0 STDs							
Normal Obs.:	1,840	92.00%	Observations << 97.5 Pct.				
Outlying Obs.:	0	0.00%	Observations >> 97.5 Pct.				
Total Obs.:	1,840	92.00%					
Cluster Feature Resources							
Feature	Description	Minimum Value	Average Value	Standard Deviation	Maximum Value	Total Value	% of Total
JOBTCBTM	Job TCB CPU Time	0	3	6	41	5,753	4.05
TAPEUSE	Tape Device Usage Indicator	0	0	0	0	0	0.00
Cluster Report Resources							
Feature	Description					Total Value	% of Total
JOBEDASD	DASD EXCPS					7,715,579	11.08
JOBETAPE	Tape EXCPS					0	0.00
JOBTOTHS	Tape Occupancy Hours					0	0.00
JOBNLR	Total Logical Writer Records					100,736	99.91

Cluster: 11							

Radius: 1.44 Clustering Index: 0.79 Maximum Cluster Radius: 3.0 STDs							
Normal Obs.:	31	1.55%	Observations << 97.5 Pct.				
Outlying Obs.:	0	0.00%	Observations >> 97.5 Pct.				
Total Obs.:	31	1.55%					
Cluster Feature Resources							
Feature	Description	Minimum Value	Average Value	Standard Deviation	Maximum Value	Total Value	% of Total
JOBTCBTM	Job TCB CPU Time	46	84	26	130	2,593	1.83
TAPEUSE	Tape Device Usage Indicator	0	0	0	0	0	0.00
Sparse Clusters Have Been Excluded							

Workload Characterization Analysis
Cluster Population Summary
For: Tuesday, July 1, 2003

2

Cluster: 11

Radius: 1.44 Clustering Index: 0.79 Maximum Cluster Radius: 3.0 STDs
Normal Obs.: 31 1.55% Observations <= 97.5 Pct.
Outlying Obs.: 0 0.00% Observations >= 97.5 Pct.
Total Obs.: 31 1.55%

Cluster Report Resources

Feature	Description	Total Value	% of Total
JOBEDASD	DASD EXCPS	563,173	0.81
JOBETAPE	Tape EXCPS	0	0.00
JOBTOHRS	Tape Occupancy Hours	0	0.00
JOBNLRL	Total Logical Writer Records	0	0.00

Cluster: 13

Radius: 1.71 Clustering Index: 0.66 Maximum Cluster Radius: 3.0 STDs
Normal Obs.: 12 0.60% Observations <= 97.5 Pct.
Outlying Obs.: 0 0.00% Observations >= 97.5 Pct.
Total Obs.: 12 0.60%

Cluster Feature Resources

Feature	Description	Minimum Value	Average Value	Standard Deviation	Maximum Value	Total Value	% of Total
JOBTCBTM	Job TCB CPU Time	105	160	22	180	1,921	1.35
TAPEUSE	Tape Device Usage Indicator	1	1	0	1	12	12.50

Cluster Report Resources

Feature	Description	Total Value	% of Total
JOBEDASD	DASD EXCPS	530,916	0.76
JOBETAPE	Tape EXCPS	3,504,873	6.39
JOBTOHRS	Tape Occupancy Hours	8	5.66
JOBNLRL	Total Logical Writer Records	0	0.00

Cluster: 14

Radius: 2.02 Clustering Index: 0.72 Maximum Cluster Radius: 3.0 STDs
Normal Obs.: 46 2.30% Observations <= 97.5 Pct.
Outlying Obs.: 0 0.00% Observations >= 97.5 Pct.
Total Obs.: 46 2.30%

Cluster Feature Resources

Feature	Description	Minimum Value	Average Value	Standard Deviation	Maximum Value	Total Value	% of Total
JOBTCBTM	Job TCB CPU Time	0	19	23	84	884	0.62
TAPEUSE	Tape Device Usage Indicator	1	1	0	1	46	47.92

Cluster Report Resources

Feature	Description	Total Value	% of Total
JOBEDASD	DASD EXCPS	1,984,232	2.85
JOBETAPE	Tape EXCPS	1,053,286	1.92
JOBTOHRS	Tape Occupancy Hours	5	3.29
JOBNLRL	Total Logical Writer Records	88	0.09

Sparse Clusters Have Been Excluded

Figure 15-27. Clustering Population Summary report
There are several points to be made concerning the Population Summary.

Cluster 1 contains 14 cases within a radius of 0. This is a special case that can occur when a small number of nearly identical observations reside together. Notice that while these 14 cases represent only 0.7% of the sample, they account for over 61% of TCB time and nearly 65% of all DASD EXCPs. Clearly, while these are exceptional cases, they must be appropriately addressed for any job class structure to be considered successful.

Cluster 9 contains the majority of the batch work under study, and represents 92% of the cases in the sample. The average TCB time for jobs represented in the cluster is 3 seconds, with a maximum TCB time of 41 seconds being observed, and no tape devices utilized. Clearly, an installation could define a single job class with several initiators and ensure excellent service for such jobs.

Cluster 11 represents jobs again requiring no tape devices, but using more TCB time. Such jobs could be treated in a manner similar to those represented in cluster 9, but fewer initiators would be required. An alternative might be to use the same job class that was used for jobs contained in cluster 9, but adjust the WLM Service class definitions to reduce service for these jobs. A review the service requirements should be made before making a final decision.

The remaining clusters, 13 & 14 represent tape jobs that have heavier processor requirements. The classes and number of initiators assigned depend on tape drive availability and service requirements.

Workload Characterization Analysis		1
Cluster Population Summary		
For: Tuesday, July 1, 2003		
The following clusters were determined to be 'sparse' in their population and have been excluded from processing.		
Cluster:	Population:	
2	2	
3	2	
4	2	
5	5	
6	2	
7	2	
8	6	
10	6	
12	7	
15	3	
16	7	
17	6	
18	7	

Figure 15-27A. Clustering Population Summary: Sparse Clusters

The above report section simply presents those clusters that were "sparse" in population. An optional report on these clusters can be generated, but was not in this particular study.

15.6.1.3.4 Data Value Exception Detail Report

This report provides a simple listing of all observations that are considered outliers, including a flag indicating which cluster feature(s) value caused the outlier status. This report can be used to help verify the criteria used for clustering, to identify those cases that should be considered as mainstream, despite the outlier status, and to determine how such cases should be considered in any post-processing efforts.

Workload Characterization Analysis
Data Value Exception Detail
For: Tuesday, July 1, 2003

1

Listing of Observations With Outlier Values

			Job		Job		Job TCB CPU		Tape Device	
System	DIVISION	DEPARTMENT	RESERVED	Group	Identification	End Time Stamp	Time	Usage	Indicator	
XAL1	DEV	01056	*****	7	LTSIPCSX	14NOV02:17:54:24.33	0:40:00.28*		0	
		01063	*****	5	HEMDA023	09NOV02:03:50:34.20	0:06:00.34*		0	
	GIS	00045	*****	0	A02DDSTT	07NOV02:21:15:40.57	4:56:12.13*		0	
						09NOV02:20:05:43.32	0:32:20.13*		0	
						12NOV02:21:58:37.99	1:24:08.19*		0	
					A02ICSTT	07NOV02:21:15:41.47	0:14:03.47*		0	
					A02SGATE	09NOV02:20:03:54.73	0:11:27.99*		0	
					DISKIB03	12NOV02:02:52:05.00	0:17:34.86*		1	
						14NOV02:23:43:56.22	0:08:37.51*		1	
					DISKIB04	14NOV02:22:42:28.48	0:07:55.69*		1	
					DISKIB08	05NOV02:07:26:09.92	0:25:25.81*		1	
					MICSDDAP	04NOV02:04:17:41.38	0:22:59.33*		1	
					MICSDDAV	04NOV02:02:18:04.06	0:03:46.72		2*	
					TOPCDEV2	08NOV02:09:51:53.06	0:05:40.66*		0	
					TOPDDEVL	06NOV02:17:20:43.25	0:28:48.32*		0	
						09NOV02:20:06:02.62	0:12:11.98*		0	
					TOPDDEV2	12NOV02:21:58:38.00	0:53:26.01*		0	
						13NOV02:10:25:47.61	0:20:52.34*		0	
					TOPDQAQA	06NOV02:11:35:54.43	0:17:50.60*		0	
						12NOV02:21:58:37.99	0:07:03.85*		0	
					TOPSPCDB	05NOV02:20:03:27.63	4:20:09.74*		0	
						09NOV02:20:04:14.55	1:12:34.37*		0	
						10NOV02:20:03:10.09	4:13:26.19*		0	
						14NOV02:20:03:18.98	5:01:29.65*		0	
						WKBKU2	09NOV02:16:32:04.49	0:16:44.47*		1
						WKBKU24	09NOV02:22:42:12.68	0:20:37.46*		1
						WKBKU27	10NOV02:01:31:32.87	0:24:41.93*		1
						WKBKU3	10NOV02:05:39:32.77	0:11:02.99*		1
						WKBKU30	10NOV02:06:14:29.78	0:21:42.96*		1
						WKBKU40	10NOV02:08:37:18.76	0:20:39.41*		1
						WKBKU61	10NOV02:14:51:03.13	0:16:30.64*		1
						WKBKU63	10NOV02:16:19:18.56	0:16:41.95*		1
						WKBKU67	10NOV02:19:43:18.32	0:19:06.31*		1
						WKBKU69	10NOV02:22:06:54.12	0:17:34.56*		1
						WKBKU80	09NOV02:08:03:40.48	0:13:40.25*		1
						WKBKU81	11NOV02:00:10:54.65	0:13:22.97*		1
						WKBKU86	04NOV02:02:23:29.11	0:15:29.06*		1
			11NOV02:04:51:03.56	0:17:21.82*		1				
		WKBKVM81	04NOV02:14:41:56.87	0:19:28.42*		1				
			11NOV02:08:57:18.42	0:19:25.72*		1				
		WKBKVM84	11NOV02:10:52:19.66	0:20:15.46*		1				
		WKBKVM85	05NOV02:03:42:10.23	0:25:11.81*		1				
			11NOV02:12:50:10.33	0:25:27.51*		1				
		WKBKVM86	05NOV02:04:04:08.44	0:23:29.65*		1				
		02046	*****	0	T1PR2NEW	14NOV02:12:56:45.81	0:07:12.03*		0	
			7	MELRI01A	09NOV02:18:24:56.07	0:20:52.65*		0		
				MELRI01B	07NOV02:18:12:06.64	0:25:34.91*		0		
	UNK	UNK	*****	0	CACTAPG	07NOV02:21:17:12.89	0:18:03.35*		0	

Note: '*' Indicates that the data value is an outlier

Workload Characterization Analysis								2
Data Value Exception Detail								
For: Tuesday, July 1, 2003								
Listing of Observations With Outlier Values								
System	Job	Job	Job	Tape	Device			
Identifier	DIVISION	DEPARTMENT	RESERVED	Group	Identification	End Time Stamp	Job TCB CPU Time	Usage Indicator
XAL1	UNK	UNK	*****	0	CACTUSA	12NOV02:21:58:38.00	0:06:32.16*	0
					TNGWE2	09NOV02:20:08:39.48	0:06:52.27*	0
					TOPDEM	12NOV02:21:58:38.00	0:25:24.12*	0

Figure 15-28. Cluster Performance Index graph

Note that in all but one case, the observations are outliers because of abnormal TCB Time usage. Such jobs can be handled either by a special exception job class, or by establishing WLM Service class definitions that manage such jobs in a lower priority scheme, such as making them Discretionary.

15.6.1.3.5 Analysis of the Clustering Output Data Set

The final step of the job class identification process involves some decisions based on the cluster descriptions and the analyst's experience. Although the clusters developed in the example provide a reasonable representation of the system's workload, additional work is needed to define an appropriate job class structure.

After reviewing the Cluster Performance and Population Summary reports, a simple analysis involving selecting, sorting and printing was executed against the output statistical data from the clustering operation. The output of the analysis is presented next.

Listing of Selected Workload Characterization Cases Presented by Clustering Feature Values												1
Cluster Type	Minimum TAPEUSE Value	Mean TAPEUSE Value	Maximum TAPEUSE Value	Minimum JOBTCBTM Value	Mean JOBTCBTM Value	Maximum JOBTCBTM Value	Cluster Number	Cluster Radius	Total Normal Observations	Total Outlier Observations	Total Observations	
Normal	0	0.00000	0	0:00:00.01	0:00:03.13	0:00:41.26	9	1.190520	1,840	0	1,840	
	0	0.00000	0	0:00:46.15	0:01:23.66	0:02:09.66	11	1.436100	31	0	31	
	0	0.00000	0	0:03:04.29	0:04:40.17	0:05:32.62	18	2.993240	7	0	7	
	1	1.00000	1	0:00:00.03	0:00:19.22	0:01:23.85	14	2.017900	46	0	46	
	1	1.00000	1	0:01:45.20	0:02:40.05	0:03:00.43	13	1.712520	12	0	12	
	1	1.00000	1	0:03:31.30	0:03:38.57	0:03:50.19	5	0.362838	5	0	5	
	1	1.00000	1	0:04:15.47	0:04:49.99	0:05:28.36	10	1.197910	6	0	6	
-----									1,947	0	1,947	
Normal												

Listing of Selected Workload Characterization Cases Presented by Clustering Feature Values												2
Cluster Type	Minimum TAPEUSE Value	Mean TAPEUSE Value	Maximum TAPEUSE Value	Minimum JOBTCBTM Value	Mean JOBTCBTM Value	Maximum JOBTCBTM Value	Cluster Number	Cluster Radius	Total Normal Observations	Total Outlier Observations	Total Observations	
Other	0	0.00000	0	0:05:40.33	0:06:25.95	0:07:12.03	12	1.438660	1	6	7	
	0	0.00000	0	0:11:27.99	0:11:49.98	0:12:11.98	7	0.686676	0	2	2	
	0	0.00000	0	0:17:50.60	0:17:56.97	0:18:03.35	4	0.199029	0	2	2	
	0	0.00000	0	0:20:52.34	0:20:52.49	0:20:52.65	2	0.004837	0	2	2	
	0	0.00000	0	0:25:24.12	0:25:29.51	0:25:34.91	3	0.168434	0	2	2	
	1	1.00000	1	0:07:55.69	0:08:16.60	0:08:37.51	6	0.652809	0	2	2	
	1	1.00000	1	0:13:22.97	0:14:10.76	0:15:29.06	15	2.444520	0	3	3	
	1	1.00000	1	0:16:30.64	0:17:04.72	0:17:34.86	8	1.063880	0	6	6	
	1	1.00000	1	0:19:06.31	0:20:10.82	0:21:42.96	16	2.876600	0	7	7	
	1	1.00000	1	0:22:59.33	0:24:32.67	0:25:27.51	17	2.914180	0	6	6	
	0	0.21429	2	0:02:44.98	1:43:52.37	5:01:29.65	1	0.000000	1	13	14	
	-----									2	51	53
Other												

The code used for this analysis divided clusters into two groups: one containing most of the normal data and the other representing the outliers. This division was done mainly to demonstrate that the outliers will normally be considered as exceptions to the job classification process. Usually only the regular (non-sparse) clusters will be processed and the sparse clusters and outliers will be treated as exceptions. In our case study, three of the sparse clusters, clusters 5, 10, and 18 were included in the "normal" clusters because they contained all normal observations. One regular cluster, cluster 1, was placed in the "Other" category because it contained almost all outliers. These decisions require installation knowledge and careful planning before implementation.

A SAS program was written to evaluate the proposed classes. The program utilizes the clustering criteria from the above report and attempts to assign each job (case) to the proposed classes (tested in ascending order of size) until a match is found. The results from the execution of this program are shown in Figure 15-30.

```
DATA JOBCLASS;
SET TOTAL_SAMPLE_CLUSTERS;

TITLE;
FOOTNOTE;

SELECT (TAPEUSE);
  WHEN (0) DO;
    IF JOBTCBTM LE 40 THEN CLASS = "A";
    ELSE IF JOBTCBTM LE 300 THEN CLASS = "B";
    ELSE CLASS = "X";
    OUTPUT;
  END;

  WHEN (1) DO;
    IF JOBTCBTM LE 300 THEN CLASS = "C";
    ELSE CLASS = "X";
    OUTPUT;
  END;

  OTHERWISE DO;
    CLASS = "X";
    OUTPUT;
  END;
```

```
END;  
RUN;  
  
PROC FREQ DATA=JOBCLASS;  
  TABLES CLASS;  
  FORMAT CLASS $CLASS.;  
  TITLE1 "Analysis of Proposed Job Classes";  
RUN;
```

The output from the job class structure evaluation is presented next.

Analysis of Proposed Job Classes					
The FREQ Procedure					
Proposed Job Class					
Proposed Job Class	Frequency	Percent	Cumulative Frequency	Cumulative Percent	
A: 0 Tapes; LE 40 secs. TCB Time	1837	91.85	1837	91.85	
B: 0 Tapes; LE 300 secs. TCB Time	39	1.95	1876	93.80	
C: 1-3 Tapes; LE 300 secs. TCB Time	66	3.30	1942	97.10	
X: <===== All Other Jobs =====>	58	2.90	2000	100.00	

The code above can be changed to suite the installation service and resource requirements. The proposed job class structure does a reasonably good job representing the resource requirements for the jobs being studied. The exception class ("X") contains only 58 jobs and represents less than 3.0 percent of the total number of jobs being analyzed. Changing the classification criteria in the code above might place more jobs in the exception class. This is not necessarily bad, but keep in mind that in most installations such jobs often require manual scheduling or similar support. Obviously the criteria used was selected mainly for illustrative purposes and might not be realistic given an installation's requirements. You can use the reports and substitute a variety of criteria, execute the simulation code and evaluate the results.

Appendix A: Messages

Messages

This appendix lists all messages generated by the CA MICS Capacity Planner. Most messages are caused by various conditions in the data. The messages are listed in ascending numerical sequence and include the full text of the message, the type, the reason for the message, the action you should take, and applicable document references.

Four type codes are used to categorize the messages:

1. Information: A note which documents a CA MICS Capacity Planner option or potentially important feature in the data.
2. Warning: A condition of either the data or the control statements which does not affect the CA MICS Capacity Planner's operation, but which may lead to unexpected results.
3. Error: An indication that a problem has been encountered with a control statement which will prevent a successful run of the component. Execution is stopped after all control statements are processed.
4. Termination: An indication that a critical problem has been encountered with the data which requires the component to terminate processing immediately.

The message texts often include references to information contained in a control statement or values of variables from temporary CA MICS Capacity Planner work files. In these cases, the message text will contain a string in double quotes, indicating that a value will be substituted into the text at execution time. For example, message CAP00031 reads:

```
NO DATA FOR "ELEMENT". SYSID="SYSID" ZONE="ZONE"
```

CAP00008

```
+-----+  
| C A P 0 0 0 0 8 |  
+-----+
```

TEXT: NO PREVIOUS FORECAST DATA FOR "CPFxxx01",
SYSID="SYSID",ZONE="ZONE"

TYPE: Information

REASON: Possible data error in the current analysis
software. No forecast data was found in the
CPFxxx01 File for the given SYSID and ZONE.
This is normal for the first execution of a plot
request.

ACTION: None required. Check the data source to ensure
that it contains valid data for the given SYSID
and ZONE combination.

REFERENCES: Section 4.6, Presenting Forecasted Data

CAP00026

```
+-----+  
| C A P 0 0 0 2 6 |  
+-----+
```

TEXT: INVALID OBSERVATION: WEEK, YEAR, OR DATE
MISSING.

TYPE: Error

REASON: An observation read from the capacity planning
database file specified by the current data
suffix is invalid and has been deleted.

ACTION: Check the data source to ensure that only valid
observations are present, then resubmit.

REFERENCES: Section 7.6, Univariate Model Forecasting
Section 9.6, Multivariate Model Forecasting
Section 10.6, Business Element Forecasting

CAP00031

```
+-----+  
| C A P 0 0 0 3 1 |  
+-----+
```

TEXT: NO DATA FOR "ELEMENT". SYSID="SYSID"
ZONE="ZONE"

TYPE: Information

REASON: No observations were found for the data element
listed for the SYSID and ZONE specified.

ACTION: None required.

REFERENCES: Section 6.4, Simple Regression Analysis
Section 7.4, Univariate Model Forecasting
Section 9.4, Multivariate Model Forecasting
Section 10.4, Business Element Forecasting

CAP00032

```
+-----+  
| C A P 0 0 0 3 2 |  
+-----+
```

TEXT: LESS THAN THREE DEGREES OF FREEDOM FOR THE MODEL
OF: "ELEMENT", REQUEST DENIED.

TYPE: Information

REASON: Not enough data is available for the statistical
procedures used by the modeling software. There
must be at least three more observations than
the number of business elements specified on the
MODEL control statement available for the
requested SYSID and ZONE in the capacity
planning database file.

ACTION: None required. Defer use of this forecasting
technique until more data has been added to the
capacity planning database file.

REFERENCES: Section 9.4, Multivariate Model Forecasting
Section 10.4, Business Element Forecasting

CAP00078

```
+-----+  
| C A P 0 0 0 7 8 |  
+-----+
```

TEXT: LESS THAN FOUR OBSERVATIONS FOR "ELEMENT" FOR
SYSID "SYSID" - ZONE "ZONE".

TYPE: Information

REASON: There is not enough data in the capacity
planning database file to perform the trend
analysis.

ACTION: Defer this analysis until more data has been
added to the capacity planning database file.

REFERENCES: Section 5.6, Profile and Trending Analysis

CAP00086

```
+-----+  
| C A P 0 0 0 8 6 |  
+-----+
```

TEXT: LESS THAN THREE OBSERVATIONS FOR "ELEMENT" FOR
SYSID "SYSID" - ZONE "ZONE".

TYPE: Information

REASON: There is not enough data in the capacity
planning database file to perform the simple
regression analysis.

ACTION: Defer this analysis until more data has been
added to the capacity planning database file.

REFERENCES: Section 6.6, Simple Regression Analysis

CAP00094

```
+-----+  
| C A P 0 0 0 9 4 |  
+-----+
```

TEXT: TOO FEW HISTORICAL DATA POINTS TO MODEL
"ELEMENT" FOR "SYSID" - ZONE "ZONE".

TYPE: Error

REASON: There is not enough data in the capacity
planning database file to perform the simple
regression analysis.

ACTION: Defer this analysis until more data has been
added to the capacity planning database file.

REFERENCES: Section 7.6, Univariate Model Forecasting

CAP00109

```
+-----+  
| C A P 0 0 1 0 9 |  
+-----+
```

TEXT: EXECUTION IS TERMINATED DUE TO PREVIOUS ERRORS.

TYPE: Information

REASON: Previously flagged errors in the control stream
resulted in the termination of this run.

ACTION: Correct all flagged errors and resubmit.

REFERENCES: Section 3.2, Capacity Planning Database

CAP00201

```
+-----+  
| C A P 0 0 2 0 1 |  
+-----+
```

TEXT: UNABLE TO ALLOCATE VOLTABLE. ERROR "ERROR".

TYPE: Termination

REASON: The VOLTABLE data set specified on the DASD
Planning Standard Application control screen
could not be allocated.

ACTION: Check that you entered the correct data set name
of an existing DASD VOLTABLE and that it is not
currently allocated to some other job or session
with a disposition of old, then resubmit.

REFERENCES: Section 3.3.2.2, Capacity Planning Database

CAP00202

```
+-----+  
| C A P 0 0 2 0 2 |  
+-----+
```

TEXT: GROUPNAME "GROUPNAME" HAS ALREADY BEEN
SPECIFIED.

TYPE: Termination

REASON: The DASD volume group "GROUPNAME" has been
specified more than once in the DASD VOLTABLE.

ACTION: Review the entries in your DASD VOLTABLE and
ensure that each GROUP named on a GROUP
statement is only specified once, then resubmit.

REFERENCES: Section 3.3.2.2, Capacity Planning Database

CAP00203

```
+-----+
| C A P 0 0 2 0 3 |
+-----+
```

TEXT: VOLSER "VOLSER" HAS ALREADY BEEN SPECIFIED.

TYPE: Termination

REASON: The DASD volume "VOLSER" has been specified either explicitly or implicitly more than once in the DASD VOLTABLE.

ACTION: Review the entries in your DASD VOLTABLE and ensure that no volume is named more than once, then resubmit. Remember that by using the colon notation you may have inadvertently referenced a volume more than once.

REFERENCES: Section 3.3.2.2, Capacity Planning Database

CAP00204

```
+-----+
| C A P 0 0 2 0 4 |
+-----+
```

TEXT: VOLSER "VOLSER" IS WITHIN THE PREVIOUSLY SPECIFIED VOLSER RANGE OF "VOLRANGE".

TYPE: Termination

REASON: The identified VOLSER is within the range of VOLSERs specified via the colon notation for volume entries in the DASD VOLTABLE.

ACTION: Either remove the explicit reference to the identified VOLSER or change the specification implied by the colon notation, then resubmit.

REFERENCES: Section 3.3.2.2, Capacity Planning Database

CAP00205

```
+-----+  
| C A P 0 0 2 0 5 |  
+-----+
```

TEXT: UNABLE TO ALLOCATE WKLTABLE. ERROR="HEX ERROR"

TYPE: Warning

REASON: The CA MICS Capacity Planner attempted to
dynamically allocate the workload data set
table, but allocation failed.

ACTION: Ensure that the data set still exists and has
not been migrated, and that no users or jobs
have the data set allocated, then resubmit.

CAP00206

```
+-----+  
| C A P 0 0 2 0 6 |  
+-----+
```

TEXT: PERFGRP "%PERFGRP" HAS ALREADY BEEN SPECIFIED.

TYPE: Warning

REASON: This performance group value has already been
specified. A duplicate entry exists in the
table. Remove one of the entries for the value
displayed in the error message.

ACTION: Review the table and remove any duplicate
entries encountered, then resubmit the job.

CAP00207

```
+-----+  
| C A P 0 0 2 0 7 |  
+-----+
```

TEXT: WLMCLASS "%WLMCLASS" HAS ALREADY BEEN SPECIFIED.

TYPE: Termination

REASON: This workload class value has already been specified. A duplicate entry exists in the table. Remove one of the entries for the value displayed in the error message.

ACTION: Review the table and remove any duplicate entries encountered, then resubmit the job.

CAP00210

```
+-----+  
| C A P 0 0 2 1 0 |  
+-----+
```

TEXT: THE REQUIRED PARAMETER "PARM" IS MISSING.

TYPE: Termination

REASON: The required parameter for the Capacity Planning Migration Utility has not been specified.

ACTION: Enter the required parameter and resubmit.

REFERENCES: None

CAP00211

```
+-----+  
| C A P 0 0 2 1 1 |  
+-----+
```

TEXT: "PARM" MUST BE x OR LESS CHARACTERS.

TYPE: Termination

REASON: The specified parameter exceeded the maximum length.

ACTION: Correct the entry and resubmit.

REFERENCES: None

CAP00212

```
+-----+  
| C A P 0 0 2 1 2 |  
+-----+
```

TEXT: "parm" MUST BE x CHARACTERS IN LENGTH.

TYPE: Termination

REASON: The specified parameter was not the correct length.

ACTION: Correct the entry and resubmit.

REFERENCES: None

CAP00213

```
+-----+  
| C A P 0 0 2 1 3 |  
+-----+
```

TEXT: FILETYPE MUST BE A, C, OR D.

TYPE: Termination

REASON: The only valid values for the FILETYPE parameter are A, C, and D. Your entry for this parameter specified some other value.

ACTION: Correct the entry and resubmit.

REFERENCES: None

CAP00214

```
+-----+  
| C A P 0 0 2 1 4 |  
+-----+
```

TEXT: NO VOLTABLE HAS BEEN SPECIFIED.

TYPE: Termination

REASON: When migrating a DASD Standard Application, you must specify the data set name of the DASD VOLTABLE. None was specified.

ACTION: Provide a valid VOLTABLE parameter and resubmit.

REFERENCES: None

CAP00215

```
+-----+  
| C A P 0 0 2 1 5 |  
+-----+
```

TEXT: THE CAPACITY PLANNING DATABASE MIGRATION
UTILITY HAS COMPLETED SUCCESSFULLY.

TYPE: Information

REASON: Your request to migrate an existing capacity
planning database file has executed
successfully.

ACTION: None required.

REFERENCES: None

CAP00216

```
+-----+  
| C A P 0 0 2 1 6 |  
+-----+
```

TEXT: THE ISPF TABLE LIBRARY HAS EITHER NOT BEEN
ALLOCATED OR IS IN USE BY ANOTHER USER.

TYPE: Termination

REASON: The ISPF table data set specified on the ISPTABL
DD statement cannot be allocated by the Capacity
Planning Migration Utility.

ACTION: Check that you coded the correct data set name
on the ISPTABL and ISPTLIB DD statements and
that no users or jobs have the data set
allocated, then resubmit.

REFERENCES: None

CAP00217

```
+-----+  
| C A P 0 0 2 1 7 |  
+-----+
```

TEXT: THE VALUE FOR "parm" MUST BE BETWEEN x AND y.

TYPE: Termination

REASON: The specified parameter was not within the required minimum and maximum lengths.

ACTION: Check the value that you specified for the parameter to ensure that it is within the required lengths, then resubmit.

REFERENCES: None

CAP00218

```
+-----+  
| C A P 0 0 2 1 8 |  
+-----+
```

TEXT: INQUIRY "xxx" ALREADY EXISTS IN THE CAPACITY PLANNING CATALOG.

TYPE: Termination

REASON: You requested the migration of an existing capacity planning database file, but specified the migrated file with the name of an inquiry that already exists in your new capacity planning catalog.

ACTION: Check your CA MICS Capacity Planner's catalog of capacity planning database file inquiries. Either delete the existing inquiry, or specify another value for the "NEWFILE" Migration Utility parameter, then resubmit.

REFERENCES: None

CAP00219

```
+-----+  
| C A P 0 0 2 1 9 |  
+-----+
```

TEXT: NO "xxx" WAS FOUND TO CONVERT. ENSURE THAT YOU HAVE SPECIFIED THE CORRECT SUFFIX FOR THE CONVERSION FILE.

TYPE: Termination

REASON: The file that you specified with the OLDFILE parameter for the Migration Utility could not be found.

ACTION: Check that you coded the correct data set name for the OLDFILE parameter, and that the file actually exists in the SAS data set named on the CAP4100 DD statement, then resubmit.

REFERENCES: None

CAP01149

```
+-----+  
| C A P 0 1 1 4 9 |  
+-----+
```

TEXT: "NUMBER" OBSERVATIONS WERE DELETED DUE TO MISSING VALUES.

TYPE: Information

REASON: The identified number of observations were deleted from the analysis due to the specifications from the control screen.

ACTION: None required.

REFERENCES: Section 11.6, Workload Characterization

CAP01150

```
+-----+  
| C A P 0 1 1 5 0 |  
+-----+
```

TEXT: NO OBSERVATIONS MET SELECTION CRITERIA.
EXECUTION TERMINATES.

TYPE: Termination

REASON: The selection criteria defined via the control
screen resulted in 0 observations being selected
for analysis.

ACTION: Examine the selection criteria specified on the
control screen, as well as the input data.
Modify the selection criteria or select from a
different CA MICS data source, and resubmit.

REFERENCES: Section 11.6, Workload Characterization
Sections 3.4 & 3.5, Capacity Planning Database

CAP01151

```
+-----+  
| C A P 0 1 1 5 1 |  
+-----+
```

TEXT: NUMBER OF CLUSTERS GENERATED EXCEEDS 99 - USE
FEWER VARIABLES FOR CLUSTERING.

TYPE: Error

REASON: The variables specified for the Analysis based
on parameter on the control screen resulted in
more than the maximum of 99 clusters being
generated.

ACTION: Specify a fewer number of variables, and
resubmit.

REFERENCES: Section 11.6, Workload Characterization

CAP01153

```
+-----+  
| C A P 0 1 1 5 3 |  
+-----+
```

TEXT: FILE 'xxxx' IS MISSING OR CONTAINS NO DATA
AFTER FILTERING.

TYPE: Warning

REASON: The file being processed has no data after
filtering has been applied.

ACTION: Verify that the selection criteria you
specified is correct.

REFERENCES: Chapters 4 and 8

CAP01200

```
+-----+  
| C A P 0 1 2 0 0 |  
+-----+
```

TEXT: NOTES MEMBER EMPTY. PRINT OF NOTES BYPASSED.

TYPE: Warning

REASON: User specified that they wished to have notes
printed, but no notes are contained in the
notes member.

ACTION: Check that you have attached notes to the file.

REFERENCES: Section 1.5.3.6 - Adding and Browsing Notes

CAP01201

```
+-----+  
| C A P 0 1 2 0 1 |  
+-----+
```

TEXT: UNABLE TO ALLOCATE NOTES MEMBER. ERROR='xxxx'
HEX_ERR.

TYPE: Termination

REASON: CA MICS Capacity Planner detected an error while trying to allocate the Notes member. The user had asked for notes to be printed, however there was an error in allocation. Printing of the notes was bypassed.

ACTION: The CA MICS Capacity Planner could not allocate the notes member. Check the JES log for indications of an allocation error.

REFERENCES: Section 1.5.3.6 - Adding and Browsing Notes

CAP01202

```
+-----+  
| C A P 0 1 2 0 2 |  
+-----+
```

TEXT: UNABLE TO ALLOCATE NOTES DSN. ERROR='xxxx'
HEX_ERR.

TYPE: Termination

REASON: CA MICS Capacity Planner detected an error while trying to allocate the Notes member. The user had asked for notes to be printed, however there was an error in allocation. Printing of the notes was bypassed.

ACTION: The CA MICS Capacity Planner could not allocate the notes data set name. Check the JES log for indications of an allocation error.

REFERENCES: Section 1.5.3.6 - Adding and Browsing Notes

CAP01203

```
+-----+  
| C A P 0 1 2 0 3 |  
+-----+
```

TEXT: USER NOTES PRINTED FOR CAPACITY FILE: 'xxxx'
CAPWORK.

TYPE: Informational

REASON: This message indicates that the notes have
been printed as requested.

ACTION: none.

REFERENCES: Section 1.5.3.6 - Adding and Browsing Notes

CAP01204

```
+-----+  
| C A P 0 1 2 0 4 |  
+-----+
```

TEXT: UNABLE TO PRINT NOTES. DATA SET NAME IS BLANK.

TYPE: Warning

REASON: The CA MICS data set name containing the Notes
members that you have created is empty. The
printing of the notes is bypassed.

ACTION: The CA MICS Capacity Planner could not allocate
the notes data set name. Check the JES log for
indications of an allocation error.

REFERENCES: Section 1.5.3.6 - Adding and Browsing Notes

CAP01205

```
+-----+  
| C A P 0 1 2 0 5 |  
+-----+
```

TEXT: FORECAST GENERATED FOR SYSID: 'xxxx'. R-SQUARE
VALUE IS 'yyyy'. F-STATISTIC IS 'zzzz'. P-VALUE
IS 'aaaa'.

TYPE: Informational

REASON: The user is running a forecast for a file and
for SYSID 'xxxx' with results of R-Square of
'xxxx', an F-statistic of 'zzzz' and a p-value
of 'aaaa'.

ACTION: none.

REFERENCES: See subsection 2 of the Chapter associated with
the type of forecast being performed. Chapter
7. for UVM, Chapter 9 for MVR, and Chapter 10
for BEF.

CAP01206

```
+-----+  
| C A P 0 1 2 0 6 |  
+-----+
```

TEXT: FORECAST GENERATED FOR SYSID: 'xxxx'.
WORKLOAD NAME 'yyyy'.

TYPE: Informational

REASON: The CA MICS Capacity Planner generated a
forecast for SYSID 'xxxx' and CAPAPU 'yyyy'.

ACTION: none.

REFERENCES: none.

CAP01207

```
+-----+  
| C A P 0 1 2 0 7 |  
+-----+
```

TEXT: FORECAST GENERATED FOR SYSID: 'xxxx'SYSID. ,
WORKLOAD NAME 'xxxx'CAPAPU. ,
WORKLOAD NAME2 'xxxx'CAPAPU2.

TYPE: Informational

REASON: The CA MICS Capacity Planner generated a
forecast for SYSID 'xxxx' and CAPAPU 'yyyy'.

ACTION: none.

REFERENCES: none.

CAP01208

```
+-----+  
| C A P 0 1 2 0 8 |  
+-----+
```

TEXT: FORECAST GENERATED FOR SYSID: 'xxxx'SYSID. ,
WORKLOAD NAME 'xxxx'CAPAPU.
WORKLOAD NAME2 'xxxx'CAPAPU2.
WORKLOAD NAME3 'xxxx'CAPAPU3.

TYPE: Informational

REASON: The CA MICS Capacity Planner generated a
forecast for SYSID 'xxxx' and CAPAPU 'yyyy'.

ACTION: none.

REFERENCES: none.

CAP01209

```
+-----+  
| C A P 0 1 2 0 9 |  
+-----+
```

TEXT: FORECAST GENERATED FOR SYSID: 'xxxx'SYSID. ,
 WORKLOAD NAME 'xxxx'CAPAPU. ,
 WORKLOAD NAME2 'xxxx'CAPAPU2. ,
 WORKLOAD NAME3 'xxxx'CAPAPU3. ,
 WORKLOAD NAME4 'xxxx'CAPAPU4.

TYPE: Informational

REASON: The CA MICS Capacity Planner generated a
 forecast for SYSID 'xxxx' and CAPAPU 'yyyy'.

ACTION: none.

REFERENCES: none.

CAP01210

```
+-----+  
| C A P 0 1 2 1 0 |  
+-----+
```

TEXT: FORECAST GENERATED FOR SYSID: 'xxxx'SYSID. ,
 WORKLOAD NAME 'xxxx'CAPAPU. ,
 WORKLOAD NAME2 'xxxx'CAPAPU2. ,
 WORKLOAD NAME3 'xxxx'CAPAPU3. ,
 WORKLOAD NAME4 'xxxx'CAPAPU4. ,
 WORKLOAD NAME5 'xxxx'CAPAPU5.

TYPE: Informational

REASON: The CA MICS Capacity Planner generated a
 forecast for SYSID 'xxxx' and CAPAPU 'yyyy'.

ACTION: none.

REFERENCES: none.

CAP01211

```
+-----+  
| C A P 0 1 2 1 1 |  
+-----+
```

TEXT: FORECAST GENERATED FOR SYSID: 'xxxx'SYSID. ,
WORKLOAD NAME 'xxxx'CAPAPU. ,
WORKLOAD NAME2 'xxxx'CAPAPU2. ,
WORKLOAD NAME3 'xxxx'CAPAPU3. ,
WORKLOAD NAME4 'xxxx'CAPAPU4. ,
WORKLOAD NAME5 'xxxx'CAPAPU5. ,
WORKLOAD NAME6 'xxxx'CAPAPU6.

TYPE: Informational

REASON: The CA MICS Capacity Planner generated a
forecast for SYSID 'xxxx' and CAPAPU 'yyyy'.

ACTION: none.

REFERENCES: none.

CAP01212

```
+-----+  
| C A P 0 1 2 1 2 |  
+-----+
```

TEXT: FORECAST GENERATED FOR SYSID: 'xxxx'SYSID. ,
WORKLOAD NAME 'xxxx'CAPAPU. ,
WORKLOAD NAME2 'xxxx'CAPAPU2. ,
WORKLOAD NAME3 'xxxx'CAPAPU3. ,
WORKLOAD NAME4 'xxxx'CAPAPU4. ,
WORKLOAD NAME5 'xxxx'CAPAPU5. ,
WORKLOAD NAME6 'xxxx'CAPAPU6. ,
WORKLOAD NAME7 'xxxx'CAPAPU7.

TYPE: Informational

REASON: The CA MICS Capacity Planner generated a
forecast for SYSID 'xxxx' and CAPAPU 'yyyy'.

ACTION: none.

REFERENCES: none.

CAP01213

```
+-----+  
| C A P 0 1 2 1 3 |  
+-----+
```

TEXT: FORECAST GENERATED FOR SYSID: 'xxxx'SYSID. ,
 WORKLOAD NAME 'xxxx'CAPAPU. ,
 WORKLOAD NAME2 'xxxx'CAPAPU2. ,
 WORKLOAD NAME3 'xxxx'CAPAPU3. ,
 WORKLOAD NAME4 'xxxx'CAPAPU4. ,
 WORKLOAD NAME5 'xxxx'CAPAPU5. ,
 WORKLOAD NAME6 'xxxx'CAPAPU6. ,
 WORKLOAD NAME7 'xxxx'CAPAPU7. ,
 WORKLOAD NAME8 'xxxx'CAPAPU8.

TYPE: Informational

REASON: The CA MICS Capacity Planner generated a
 forecast for SYSID 'xxxx' and CAPAPU 'yyyy'.

ACTION: none.

REFERENCES: none.

CAP01214

```
+-----+
| C A P 0 1 2 1 4 |
+-----+
```

TEXT: FORECAST GENERATED FOR SYSID: 'xxxx'SYSID. ,
WORKLOAD NAME 'xxxx'CAPAPU. ,
WORKLOAD NAME2 'xxxx'CAPAPU2. ,
WORKLOAD NAME3 'xxxx'CAPAPU3. ,
WORKLOAD NAME4 'xxxx'CAPAPU4. ,
WORKLOAD NAME5 'xxxx'CAPAPU5. ,
WORKLOAD NAME6 'xxxx'CAPAPU6. ,
WORKLOAD NAME7 'xxxx'CAPAPU7. ,
WORKLOAD NAME8 'xxxx'CAPAPU8. ,
WORKLOAD NAME9 'xxxx'CAPAPU9.

TYPE: Informational

REASON: The CA MICS Capacity Planner generated a
forecast for SYSID 'xxxx' and CAPAPU 'yyyy'.

ACTION: none.

REFERENCES: none.

CAP01215

```
+-----+
| C A P 0 1 2 1 5 |
+-----+
```

TEXT: VALUE IS 'yyyy'. F-STATISTIC IS 'zzzz'. P-VALUE
IS 'aaaa'.

TYPE: Informational

REASON: The user is running a forecast for a file and
for SYSID 'xxxx' with results of R-Square of
'xxxx', an F-statistic of 'zzzz' and a p-value
of 'aaaa'.

ACTION: None.

REFERENCES: See subsection 2 of the Chapter associated with
the type of forecast being performed. Chapter
7. for UVM, Chapter 9 for MVR, and Chapter 10
for BEF.

CAP01216

```
+-----+  
| C A P 0 1 2 1 6 |  
+-----+
```

TEXT: WORKLOAD NAME 'xxx'CAPAPU.

TYPE: Informational

REASON: This is an informational message that comes out after message number CAPI0026, CAPI0032, or CAPI0094 that gives the CAPAPU associated with the error message.

ACTION: Verify that your selection criteria is correct and that there is sufficient data to perform this forecast.

REFERENCES: See the description for messages CAPI0026, CAPI0032, and CAPI0094.

CAP01217

```
+-----+  
| C A P 0 1 2 1 7 |  
+-----+
```

TEXT: WORKLOAD NAME 'xxx'CAPAPU. ,
WORKLOAD NAME2 'xxx'CAPAPU2.

TYPE: Informational

REASON: This is an informational message that comes out after message number CAPI0026, CAPI0032, or CAPI0094 that gives the CAPAPU associated with the error message.

ACTION: Verify that your selection criteria is correct and that there is sufficient data to perform this forecast.

REFERENCES: See the description for messages CAPI0026, CAPI0032, and CAPI0094.

CAP01218

```
+-----+  
| C A P 0 1 2 1 8 |  
+-----+
```

TEXT: WORKLOAD NAME 'xxx'CAPAPU. ,
 WORKLOAD NAME2 'xxx'CAPAPU2. ,
 WORKLOAD NAME3 'xxx'CAPAPU3.

TYPE: Informational

REASON: This is an informational message that comes out
 after message number CAPI0026, CAPI0032, or
 CAPI0094 that gives the CAPAPU associated with
 the error message.

ACTION: Verify that your selection criteria is correct
 and that there is sufficient data to perform
 this forecast.

REFERENCES: See the description for messages CAPI0026,
 CAPI0032, and CAPI0094.

CAP01219

```
+-----+  
| C A P 0 1 2 1 9 |  
+-----+
```

TEXT: WORKLOAD NAME 'xxx'CAPAPU. ,
 WORKLOAD NAME2 'xxx'CAPAPU2. ,
 WORKLOAD NAME3 'xxx'CAPAPU3. ,
 WORKLOAD NAME4 'xxx'CAPAPU4.

TYPE: Informational

REASON: This is an informational message that comes out
 after message number CAPI0026, CAPI0032, or
 CAPI0094 that gives the CAPAPU associated with
 the error message.

ACTION: Verify that your selection criteria is correct
 and that there is sufficient data to perform
 this forecast.

REFERENCES: See the description for messages CAPI0026,
 CAPI0032, and CAPI0094.

CAP01220

```
+-----+  
| C A P 0 1 2 2 0 |  
+-----+
```

TEXT: WORKLOAD NAME 'xxxx'CAPAPU. ,
 WORKLOAD NAME2 'xxxx'CAPAPU2. ,
 WORKLOAD NAME3 'xxxx'CAPAPU3. ,
 WORKLOAD NAME4 'xxxx'CAPAPU4. ,
 WORKLOAD NAME5 'xxxx'CAPAPU5.

TYPE: Informational

REASON: This is an informational message that comes out
 after message number CAPI0026, CAPI0032, or
 CAPI0094 that gives the CAPAPU associated with
 the error message.

ACTION: Verify that your selection criteria is correct
 and that there is sufficient data to perform
 this forecast.

REFERENCES: See the description for messages CAPI0026,
 CAPI0032, and CAPI0094.

CAP01221

```
+-----+  
| C A P 0 1 2 2 1 |  
+-----+
```

TEXT: WORKLOAD NAME 'xxxx'CAPAPU. ,
 WORKLOAD NAME2 'xxxx'CAPAPU2. ,
 WORKLOAD NAME3 'xxxx'CAPAPU3. ,
 WORKLOAD NAME4 'xxxx'CAPAPU4. ,
 WORKLOAD NAME5 'xxxx'CAPAPU5. ,
 WORKLOAD NAME6 'xxxx'CAPAPU6.

TYPE: Informational

REASON: This is an informational message that comes out
 after message number CAPI0026, CAPI0032, or
 CAPI0094 that gives the CAPAPU associated with
 the error message.

ACTION: Verify that your selection criteria is correct
 and that there is sufficient data to perform
 this forecast.

REFERENCES: See the description for messages CAPI0026,
 CAPI0032, and CAPI0094.

CAP01222

```
+-----+  
| C A P 0 1 2 2 2 |  
+-----+
```

TEXT: WORKLOAD NAME 'xxxx'CAPAPU. ,
 WORKLOAD NAME2 'xxxx'CAPAPU2. ,
 WORKLOAD NAME3 'xxxx'CAPAPU3. ,
 WORKLOAD NAME4 'xxxx'CAPAPU4. ,
 WORKLOAD NAME5 'xxxx'CAPAPU5. ,
 WORKLOAD NAME6 'xxxx'CAPAPU6. ,
 WORKLOAD NAME7 'xxxx'CAPAPU7.

TYPE: Informational

REASON: This is an informational message that comes out
 after message number CAPI0026, CAPI0032, or
 CAPI0094 that gives the CAPAPU associated with
 the error message.

ACTION: Verify that your selection criteria is correct
 and that there is sufficient data to perform
 this forecast.

REFERENCES: See the description for messages CAPI0026,
 CAPI0032, and CAPI0094.

CAP01223

```
+-----+  
| C A P 0 1 2 2 3 |  
+-----+
```

TEXT: WORKLOAD NAME 'xxxx'CAPAPU. ,
 WORKLOAD NAME2 'xxxx'CAPAPU2. ,
 WORKLOAD NAME3 'xxxx'CAPAPU3. ,
 WORKLOAD NAME4 'xxxx'CAPAPU4. ,
 WORKLOAD NAME5 'xxxx'CAPAPU5. ,
 WORKLOAD NAME6 'xxxx'CAPAPU6. ,
 WORKLOAD NAME7 'xxxx'CAPAPU7. ,
 WORKLOAD NAME8 'xxxx'CAPAPU8.

TYPE: Informational

REASON: This is an informational message that comes out
 after message number CAPI0026, CAPI0032, or
 CAPI0094 that gives the CAPAPU associated with
 the error message.

ACTION: Verify that your selection criteria is correct
 and that there is sufficient data to perform
 this forecast.

REFERENCES: See the description for messages CAPI0026,
 CAPI0032, and CAPI0094.

CAP01224

```
+-----+
| C A P 0 1 2 2 4 |
+-----+
```

TEXT: WORKLOAD NAME 'xxxx'CAPAPU. ,
 WORKLOAD NAME2 'xxxx'CAPAPU2. ,
 WORKLOAD NAME3 'xxxx'CAPAPU3. ,
 WORKLOAD NAME4 'xxxx'CAPAPU4. ,
 WORKLOAD NAME5 'xxxx'CAPAPU5. ,
 WORKLOAD NAME6 'xxxx'CAPAPU6. ,
 WORKLOAD NAME7 'xxxx'CAPAPU7. ,
 WORKLOAD NAME8 'xxxx'CAPAPU8. ,
 WORKLOAD NAME9 'xxxx'CAPAPU9.

TYPE: Informational

REASON: This is an informational message that comes out
 after message number CAPI0026, CAPI0032, or
 CAPI0094 that gives the CAPAPU associated with
 the error message.

ACTION: Verify that your selection criteria is correct
 and that there is sufficient data to perform
 this forecast.

REFERENCES: See the description for messages CAPI0026,
 CAPI0032, and CAPI0094.

CAP01225

```
+-----+
| C A P 0 1 2 2 5 |
+-----+
```

TEXT: UNIVARIATE MODELING FORECAST BEING PERFORMED
 FOR VARIABLE NAME:'xxxx'

TYPE: Informational

REASON: The user is running a UVM forecast for the
 variable name 'xxxx'.

ACTION: none.

REFERENCES: See Section 7.5.1

CAP01226

```
+-----+  
| C A P 0 1 2 2 6 |  
+-----+
```

TEXT: BUSINESS ELEMENT FORECAST BEING PERFORMED FOR
VARIABLE NAME: 'xxx'.

TYPE: Informational

REASON: The user is running a BEF forecast for the
variable name 'xxx'.

ACTION: none.

REFERENCES: See Section 10.5.1

CAP01227

```
+-----+  
| C A P 0 1 2 2 7 |  
+-----+
```

TEXT: MULTIVARIATE REGRESSION FORECAST BEING
PERFORMED FOR VARIABLE NAME: 'xxx'.

TYPE: Informational

REASON: The user is running a MVR forecast for the
variable name 'xxx'.

ACTION: none.

REFERENCES: See Section 9.5.1

***** new *****

CAP02000

```
+-----+  
| C A P 0 2 0 0 0 |  
+-----+
```

TEXT: "CAPMSG"

TYPE: Informational

REASON: Multi-use generic message

ACTION: none.

REFERENCES: none.

CAP02010

```
+-----+  
| C A P 0 2 0 1 0 |  
+-----+
```

TEXT: "CAPMSG"

TYPE: Informational

REASON: Multi-use generic message; future use

ACTION: none.

REFERENCES: none.

CAP02050

```
+-----+  
| C A P 0 2 0 5 0 |  
+-----+
```

TEXT: "PHASE": "_PGMTTL" BEGINNING EXECUTION...

TYPE: Informational

REASON: Processing phase _PGMTTL is starting

ACTION: none.

REFERENCES: none.

CAP02060

```
+-----+  
| C A P 0 2 0 6 0 |  
+-----+
```

TEXT: NEUGENTS MODEL "_FUNCT" PHASE BEGINNING...

TYPE: Informational

REASON: Neugents internal phase _FUNCT starting

ACTION: none.

REFERENCES: None.

CAP02070

```
+-----+  
| C A P 0 2 0 7 0 |  
+-----+
```

TEXT: "PHASE": "_PGMTTL" ENDING EXECUTION...

TYPE: Informational

REASON: Processing phase _PGMTTL ending.

ACTION: none.

REFERENCES: None.

CAP02075

```
+-----+  
| C A P 0 2 0 7 5 |  
+-----+
```

TEXT: "PHASE": "_PGMTTL" ABNORMALLY TERMINATING

TYPE: Error.

REASON: Processing _PGMTTL abnormally terminated.

ACTION: Neugents processing job will end.

REFERENCES: None.

CAP02080

```
+-----+  
| C A P 0 2 0 8 0 |  
+-----+
```

TEXT: "V1" FILE CONTAINS "N" RECORDS SELECTED FOR PROCESSING.

TYPE: Informational

REASON: N records have been selected for processing from file V1.

ACTION: none.

REFERENCES: None.

CAP02085

```
+-----+  
| C A P 0 2 0 8 5 |  
+-----+
```

TEXT: DATA FROM FILE "V1" WILL BE EXCLUDED, YOUR RESULTS MAY BE NEGATIVELY AFFECTED.

TYPE: Warning.

REASON: Some observations from file VI were excluded, either due data value errors or unmatched observations among files. Modeling results may be impacted and should be examined in light of this message.

ACTION: none.

REFERENCES: None.

CAP02090

```
+-----+  
| C A P 0 2 0 9 0 |  
+-----+
```

TEXT: NO FORECAST FILE ENTRY FOUND FOR "V1" ENTRY:,

TYPE: Warning.

REASON: An observation in the Primary CAP file was not
matched in the alternate input file.

ACTION: The original observation is discarded from
further processing. See message CAP02095 for
additional details.

REFERENCES: None.

CAP02095

```
+-----+  
| C A P 0 2 0 9 5 |  
+-----+
```

TEXT: %FNAME %FDATA

TYPE: Warning.

REASON: One of the key values of the unmatched
observation is displayed. FNAME is the name of
the key variable; FDATA is the data value of
the key variable.

ACTION: None.

REFERENCES: None.

CAP02100

```
+-----+  
| C A P 0 2 1 0 0 |  
+-----+
```

TEXT: REPORT: "V1" BEING CURRENTLY GENERATED ON
DDNAME //"V1"

TYPE: Informational

REASON: Generation of report V1 has started.

ACTION: The report will be written out on DDNAME V1.
This message will only be issued if the MICF
environment is not being used. Message CAP02105
will normally be issued instead.

REFERENCES: None.

CAP02105

```
+-----+  
| C A P 0 2 1 0 5 |  
+-----+
```

TEXT: REPORT: "V1" BEING CURRENTLY GENERATED.

TYPE: Informational

REASON: Generation of report V1 has started.

ACTION: Reports generated within the MICF environment
are normally written out to the SASLIST DD.

REFERENCES: None.

CAP02110

```
+-----+  
| C A P 0 2 1 1 0 |  
+-----+
```

TEXT: REPORT: "V1" BEING CURRENTLY GENERATED VIA CSV FILE.

TYPE: Informational

REASON: Generation of report V1 has started and will be created using a CSV file.

ACTION: This option is used to create MQR compatible CSV files.

REFERENCES: None.

CAP02120

```
+-----+  
| C A P 0 2 1 2 0 |  
+-----+
```

TEXT: GRAPH: "V1" BEING CURRENTLY GENERATED ON DDNAME // "V1"

TYPE: Informational

REASON: Generation of graph V1 has started.

ACTION: The graph will be written out on DDNAME V1.
This message will only be issued if the MICF environment is not being used. Message CAP02125 will normally be issued instead.

REFERENCES: None.

CAP02125

```
+-----+  
| C A P 0 2 1 2 5 |  
+-----+
```

TEXT: GRAPH: "V1" BEING CURRENTLY GENERATED.

TYPE: Informational

REASON: Generation of graph V1 has started.

ACTION: Graphs generated within the MICF environment
are normally written out to the SASLIST DD.

REFERENCES: None.

CAP02130

```
+-----+  
| C A P 0 2 1 3 0 |  
+-----+
```

TEXT: GRAPH: "V1" BEING CURRENTLY GENERATED VIA CSV
FILE.

TYPE: Informational

REASON: Generation of graph V1 has started and will be
created using a CSV file.

ACTION: This option is used to create MQR compatible
CSV files.

REFERENCES: None.

CAP02150

```
+-----+  
| C A P 0 2 1 5 0 |  
+-----+
```

TEXT: INVALID DATA VALUES DETECTED IN "V1" FILE,
"_PGMTTL" CANNOT PROCESS SUCCESSFULLY, WHEN
DATA VALUES ARE MISSING OR INVALID. SEE
BADTRAIN REPORT FOR DETAILS.

TYPE: Warning

REASON: Missing or zero values were detected in data
elements in one or more observations.

ACTION: The observations are discarded from further
processing. Report "BADTRAIN" will contain a
list of observations that were dropped.

REFERENCES: None.

CAP02160

```
+-----+  
| C A P 0 2 1 6 0 |  
+-----+
```

TEXT: MODEL SUMMARY STATISTICS HAVE BEEN CREATED ON
DDNAME //MODLSTAT

TYPE: Informational

REASON: The Model Summary Statistics report is being
created.

ACTION: The report will be written out on DDNAME
MODLSTAT. This message will only be issued if
the MICF environment is not being used. Message
CAP02165 will normally be issued instead.

REFERENCES: None.

CAP02165

```
+-----+  
| C A P 0 2 1 6 5 |  
+-----+
```

TEXT: MODEL SUMMARY STATISTICS HAVE BEEN CREATED.

TYPE: Informational

REASON: The Model Summary Statistics report is being created.

ACTION: Reports generated within the MICF environment are normally written out to the SASLIST DD.

REFERENCES: None.

CAP02170

```
+-----+  
| C A P 0 2 1 7 0 |  
+-----+
```

TEXT: "_FORTYP" FORECASTING OPTION SELECTED -, "V1"
"V2" OF PREDICTED DATA WILL BE GENERATED.
BEGINNING WITH "X_SDATE".

TYPE: Informational.

REASON: For forecasting option _FORTYP, V1
months/weeks/days of data were selected for
processing, beginning with X_DATE. This
message defined the time frame of the modeling
study.

ACTION: None.

REFERENCES: None.

CAP02210

```
+-----+  
| C A P 0 2 2 1 0 |  
+-----+
```

TEXT: "REPORTING LEVELS EXCEEDED MAXIMUM LIMIT, QUERY
EXECUTION WILL BE TERMINATED"

TYPE: Error.

REASON: Currently, a maximum of nine levels can be
specified for Q&R charting and reporting.

ACTION: Program CAPNVP22 detected this condition and
will terminate with abend u998.

REFERENCES: None.

CAP02300

```
+-----+  
| C A P 0 2 3 0 0 |  
+-----+
```

TEXT: MACRO "PHASE" BEGINNING EXECUTION...

TYPE: Informational

REASON: Macro PHASE is starting execution. This macro
is used to provide linkage to Neugents
technology.

ACTION: None.

REFERENCES: None.

CAP02310

```
+-----+  
| C A P 0 2 3 1 0 |  
+-----+
```

TEXT: MACRO "PHASE" FAILED WITH RETURN CODE "RC"

TYPE: Error.

REASON: A non-zero return code value was detected upon
 return from Neugents.

ACTION: Neugents processing will be terminated.

REFERENCES: None.

CAP02320

```
+-----+  
| C A P 0 2 3 2 0 |  
+-----+
```

TEXT: MACRO GLOBALS MISSING REQUIRED OPERAND
 "MODNAME", EXECUTION TERMINATING.

TYPE: Error.

REASON: An internal error has occurred. This message
 should never occur except during product
 development.

ACTION: Processing terminates.

REFERENCES: None.

CAP02325

```
+-----+  
| C A P 0 2 3 2 5 |  
+-----+
```

TEXT: MACRO GLOBALS MISSING REQUIRED OPERAND
 "MODTITLE", EXECUTION TERMINATING.

TYPE: Error

REASON: An internal error has occurred. This message
 should never occur except during product
 development.

ACTION: Processing terminates.

REFERENCES: None.

CAP02390

```
+-----+  
| C A P 0 2 3 9 0 |  
+-----+
```

TEXT: MACRO "PHASE" TERMINATING, ABORT REQUESTED.

TYPE: Error.

REASON: A terminal error condition was detected in
 macro PHASE. This message is preceded by
 either message CAP02310 or CAP02590.

ACTION: Neugents processing terminates.

REFERENCES: None.

CAP02395

```
+-----+  
| C A P 0 2 3 9 5 |  
+-----+
```

TEXT: MACRO "PHASE" ENDED WITH RETURN CODE "RC"

TYPE: Informational

REASON: The macro PHASE ended execution, with return code RC.

ACTION: None.

REFERENCES: None.

CAP02410

```
+-----+  
| C A P 0 2 4 1 0 |  
+-----+
```

TEXT: MODEL TRAINING PHASE COMPLETED SUCCESSFULLY.
TRAINING STATISTICS ARE SUMMARIZED ON DDNAME
//DELTASTS, AND ARE INCLUDED BELOW:
NUMBER OF CASES USED FOR TRAINING: "N",
AVERAGE MODEL ERROR MEASURED: "V1",
AVERAGE MODEL RMS ERROR MEASURED: "V2",
(GOODNESS OF FIT) FOR MODEL: "V3"

TYPE: Informational

REASON: Neugents Value Prediction models first execute a training phase, and produce several statistics used to evaluate the model. These statistics are displayed in the message.

ACTION: The client should review the statistics before deciding on the validity of the model.

REFERENCES: None.

CAP02420

```
+-----+  
| C A P 0 2 4 2 0 |  
+-----+
```

TEXT: MODEL FORECASTING PHASE COMPLETED SUCCESSFULLY.
 FORECASTING STATISTICS ARE SUMMARIZED ON DDNAME
 //CONSFCS, AND ARE INCLUDED BELOW:
 NUMBER OF CASES USED FOR FORECASTING: "N",
 (GOODNESS OF FIT) FOR MODEL: "V1"

TYPE: Informational

REASON: Neugents Value Prediction models secondly
 execute a consulting phase, and produce
 additional statistics used to evaluate the
 model. These statistics are displayed in the
 message.

ACTION: The client should review the statistics before
 deciding on the validity of the model.

REFERENCES: None.

CAP02430

```
+-----+  
| C A P 0 2 4 3 0 |  
+-----+
```

TEXT: LINEAR REGRESSION PHASE COMPLETED SUCCESSFULLY.
MODEL STATISTICS ARE SUMMARIZED ON DDNAME
//MODLSTAT, AND ARE INCLUDED BELOW: NUMBER OF
CASES USED FOR MODELING: "V1",
SLOPE OF REGRESSION LINE: "V2",
INTERCEPT VALUE FOR IND. ELEMENT: "V3",
(GOODNESS OF FIT) FOR MODEL: "V4"

TYPE: Informational

REASON: Workload Forecasting will execute a simple
linear regression operation when a single file
is specified and only a forecasting element is
provided. Regression models produce various
statistics used to evaluate the model. These
statistics are displayed in the message.

ACTION: The client should review the statistics before
deciding on the validity of the model. The
report will be written out on DDNAME MODLSTAT.
This message will only be issued if the MICF
environment is not being used. Message CAP02435
will normally be issued instead.

REFERENCES: None.

CAP02435

```
+-----+  
| C A P 0 2 4 3 5 |  
+-----+
```

TEXT: LINEAR REGRESSION PHASE COMPLETED SUCCESSFULLY.
 MODEL STATISTICS ARE SUMMARIZED ON DDNAME
 //SASLIST AND ARE INCLUDED BELOW: NUMBER OF
 CASES USED FOR MODELING: "V1",
 SLOPE OF REGRESSION LINE: "V2",
 INTERCEPT VALUE FOR IND. ELEMENT: "V3",
 (GOODNESS OF FIT) FOR MODEL: "V4"

TYPE: Informational

REASON: Workload Forecasting will execute a simple
 linear regression operation when a single file
 is specified and only a forecasting element is
 provided. Regression models produce various
 statistics used to evaluate the model. These
 statistics are displayed in the message.

ACTION: The client should review the statistics before
 deciding on the validity of the model. Reports
 generated within the MICF environment are
 normally written out to the SASLIST DD.

REFERENCES: None.

CAP02500

```
+-----+  
| C A P 0 2 5 0 0 |  
+-----+
```

TEXT: OPEN ISSUED FOR "STD" FILE. RETURN CODE IS:
"RC"

TYPE: Informational

REASON: An open was issued for file with DDNAME "STD".
Such files are used by Neugents processing for
passing data between Neugents and the Capacity
Planning applications, or are required by USS
for internal use. The return code from open
processing is displayed as "RC".

ACTION: None.

REFERENCES: None.

CAP02510

```
+-----+  
| C A P 0 2 5 1 0 |  
+-----+
```

TEXT: OPEN FAILED FOR "STD" FILE. SYSTEM ERROR
MESSAGE IS "SYSMSG" EXECUTION TERMINATED.

TYPE: Error

REASON: An open was issued for file with DDNAME "STD".
Such files are used by Neugents processing for
passing data between Neugents and the Capacity
Planning applications, or are required by USS
for internal use. The return code from open
processing is displayed as "RC". Generally, a
non-zero return code value will cause
processing to be terminated.

ACTION: Review the specific system message for action.

REFERENCES: None.

CAP02590

```
+-----+  
| C A P 0 2 5 9 0 |  
+-----+
```

TEXT: ERROR TEXT FROM NEUGENTS PROCESSING
"ERRMSG" EXECUTION TERMINATED.

TYPE: Error

REASON: Any error text returned by USS in the STDERR
dataset will be displayed herein.

ACTION: Review specific error message for action.

REFERENCES: None.

CAP02610

```
+-----+  
| C A P 0 2 6 1 0 |  
+-----+
```

TEXT: %V1 %V2 Values Have Exceeded the %V3 Threshold
During the Reporting Period.

TYPE: Informational

REASON: The number of data values for the Forecast
element that exceeded the specified threshold
value is displayed. Each value that exceeds the
threshold will result in an alert condition
being raised.

ACTION: None.

REFERENCES: None.

CAP02620

```
+-----+  
| C A P 0 2 6 2 0 |  
+-----+
```

TEXT: Forecast Results Should be Reviewed for Accuracy

TYPE: Informational

REASON: This is displayed at the end of the query operation if any data values have exceeded the threshold and resulted in Alerts being raised.

ACTION: The Forecast reports should be reviewed as the Alerts indicated a departure of the current forecasted results from previous expectations.

REFERENCES: None.

CAP02690

```
+-----+  
| C A P 0 2 6 9 0 |  
+-----+
```

TEXT: ALERT Processing Requested but no Forecast File was Found in the CAP Data Base

TYPE: Warning

REASON: Threshold Alert processing uses the Forecast file from a prior execution to determine if an exception has occurred. No such file was found, so Threshold Alert processing was ignored.

ACTION: None.

REFERENCES: None.

Appendix B: QUICK START TUTORIAL

This appendix provides a step-by-step tutorial to guide you through a typical setup of a CPU capacity plan from building and collecting data for the capacity database to generating and presenting the forecast in presentation graphics.

The tutorial uses the sample z/OS Processor Planning application, one of the many predefined views of resource consumption delivered by the CA MICS Capacity Planner, to assist you in collecting capacity planning data. However, your main capacity issue may require a different sample application, such as the one for DB2, CICS, or goal mode. In any case, this tutorial uses the CPU Planning application as an example to show you how to work with a typical sample Application.

Before you begin, it is recommended that you review sections 2.0 and 2.1 of this guide. These sections describe the sample z/OS Processor Planning application and the types of data it collects for processor usage planning. Section 2.1 will help you understand the example that is used in this tutorial. However, if you need to use a different sample application, you must read another section of chapter 2. For example, you will need to read section 2.17 if you need to use the Goal Mode sample Application.

While not required, a basic knowledge of the MWF (CA MICS Workstation Facility) and the MICF (CA MICS Information Control Facility) is helpful. This tutorial guides you through the Capacity Planner's online facility using MWF and discusses reporting interfaces in MICF.

The z/OS Processor Planning sample Application that is described in this tutorial uses the MONTHS timespan of the HARCPU and WLMSEC files in the CA MICS database and archive tapes. You can substitute a different timespan in the actual planning process based on your requirements and database configuration.

Assumptions

- o You have a basic knowledge of MICF.
- o You have reviewed sections 2.0 and 2.1 of this guide.

- o If you are going to use the z/OS Processor Planning sample application, the HARCPU and WLMSEC files must be available online or on CA MICS archive tapes. The MONTHS timespan is preferred for this example.
- o If you need to use another sample application, you can check the Usage Guidelines section for that application to find out what CA MICS files are required. Each sample Application that is described in Chapter 2 of this guide contains a section called Usage Guidelines that tells what CA MICS files are required.

Introduction to the Exercise

This z/OS Processor sample application exercise provides CPU utilization statistics for performance analysis and can serve as input for future capacity planning forecasts. This procedure does not discuss option parameters such as Common Data Selection (Figure 2-3). These parameters can also be specified to subset the data further.

The tutorial consists of four sections that are designed to help you:

- o Build the Capacity Database for CPU planning
- o Generate the CPU Forecast
- o Create the CPU Worksheet
- o Build Presentation Graphics

This section contains the following topics:

[B.1 Quick Start Tutorial for CPU Planning](#) (see page 1114)

B.1 Quick Start Tutorial for CPU Planning

Assumptions:

- o You have a basic knowledge of MICF
- o You have reviewed sections 2.0 and 2.1 of the CA MICS Capacity Planner.
- o Multiple months of HARCPU and WLMSEC. You must omit CPU adjustment if the WLMSEC file is not available. Data must be available online or on CA MICS archive tapes.

Introduction:

This z/OS Processor Planning sample exercise provides processor utilization statistics for performance analysis and can serve as input for future forecasts for capacity planning. This procedure does not discuss option parameters such as Common Data Selection (Figure 2-3). These parameters can also be specified to further subset the data.

B.1.1 Build the Capacity Database for Processor Planning

The CA MICS capacity database for an application will be stored in a SAS Data Library as members that contain the historical data and the forecasts that are based on the historical data. As a user, you will store your capacity database files on a SAS data library that has your TSO userid as the high level prefix. This user SAS data library is called a Private capacity database. Once you are satisfied that the files for a capacity application are correct, they can be copied to a common SAS Data Library called the Shared Capacity Database, which has the CA MICS sharedprefix as its high level qualifier.

If you have never created an application under the CA MICS Capacity Planner, the SAS Data Library for your private capacity database will not yet exist. This SAS Data Library will be allocated automatically as you follow a procedure such as the one described in this section. If you have already built an application sometime in the past, then the SAS Data Library for your private capacity database should exist already. It might be archived on your system or it might have been deleted. If it was deleted and you have followed the backup procedure described in section 12.2 of this guide, then you could restore your private capacity database by using the "Private Capacity Planning Database Restore" procedure which is also described in section 12.2.

When you decide to build a capacity database, you must determine what data to collect and the timespan at which the data should be summarized.

This exercise uses the MONTHS timespan for input and summarization since exceptions or unique system behavior will not be a concern at this higher timespan. However, if there are certain days or certain hours of a day that usually have high CPU utilization, the ZONE element could be defined so that these peak periods can be isolated. In building the Capacity Database, peak periods could be eliminated from the capacity application.

To start the exercise:

1. If this is the first time that you are using the Capacity Planner, the SAS Data Library that will contain your capacity database has not yet been allocated. Thus you should go into the MWF 4;3;0;6 panel ("Data Set Allocation Parameters"). Under the "Data Set" column are entries for CAPACITY and for CAPPDS. You will see something similar to these lines:

```
_ CAPACITY  MICS6.USER.CAPACITY      TRK 20  10      3390
_ CAPPDS    MICS6.USER.CAPPDS        TRK 20  10  20  3390
```

The names under "Data Set Name Suffix" will be concatenated after TSO userid to form the names of these datasets when they are allocated. Thus your private capacity database will reside in

userid.MICS6.USER.CAPACITY

Note that the names here have 'MICS6.USER' as part of the suffix. This is purely arbitrary and could be different in your CA MICS complex.

If you need to change either of these names, you can do so on this panel. Also, you should change the allocation space allocation parameters and the "DASD unit" if other values are appropriate.

2. Proceed to MWF 4;3;1 and INSERT a new Capacity Query (Documentation: CAP Section 1.5.3).

Code values that you want for the "File Name" and "Title".
Record the File Name here: _____.

Select 1 for "Sample Applications" and press ENTER.

3. A list of available Sample Applications will appear. For our example here, we will select 1 for "z/OS Processor Planning" (Documentation: CAP Section 2.1) and press ENTER.

4. Keep the Summarization Level defaults where MONTHS is the Primary and ZONE is the secondary.

Specify at least one unit data base ID. If you are doing a Processor Planning application, then the data base ID must be for a RMF unit.

You may code Y after "Specify extended options" and press ENTER.

5. Go to the Extended Options Menu (Section 2.1.1 and 2.1.2).

Specify CPU time adjustment parameters as described in Figure B-1.

NOTE: WLMSEC is required if the CPU time adjustment is specified.

```
/----- CPU Time Adjustment Parameters -----\
|Command ==>
|
|Inquiry step: z/OS Processor Planning Application
|
|Base CPU model                ==> MYCPU_____
|Number of processors in base CPU model ==> 1_      (1-16)
|Base CPU service unit coefficient    ==> 100___ (per processor)
|
|
|
|
|
|
|
|
|
|
```

Figure B-1. CPU Time Adjustment Parameters Screen

6. After completing the entries in these definition screens, exit out of them until you arrive at the "Inquiry Processing" panel. Code EXECUTE after "Specify Inquiry Processing" and press ENTER

The "Update Capacity Planning File" screen will appear. See Figure B-2.

```

/----- Update Capacity Planning File -----\
|Command ==>|
|
|Executing CAP DB File def.: QC1 - Quick Start CPU Example|
|
|Type of update          ==> INITIALIZE (INITIALIZE/UPDATE)|
|Replace existing file?  ==> N          (Y/N)|
|
|Input data source       ==> ARCHIVE    (ARCHIVE/ONLINE)|
|CA MICS file cycle(s)   ==> __ - __   (0 - 99)|
|Archive generation number ==> -6_     (0 to -99)|
|
|Starting date of update ==> _____ (ddmonyy)|
|Number of months to select ==> ____   (1 to 9999)|
|
|Mode of execution       ==> BATCH     (BATCH/FOREGROUND)|
|
|List updated file contents? ==> YES    (YES/NO)|
|
|
|
|Press END to invoke MICF inquiry execution to update the resource element file.|
|Enter CANCEL to terminate the resource element file update.|
\-----/

```

Figure B-2. Update Capacity Planning File Screen

- Complete the entries on this panel. Since this is the first time the capacity database will be updated for this application, the "Type of update" should be INITIALIZE and "Replace existing file" should be N.

The "Input data source" can be ARCHIVE or ONLINE. If ARCHIVE is chosen, then the "Archive generation number" must also be chosen. On the other hand, if ONLINE is chosen, then the input will come directly from an existing CA MICS timespan. In this case, the "CA MICS file cycle(s)" must be specified and the "Archive generation number" must be blank.

The Starting date and the number of months to select should be coded.

You may select BATCH or FOREGROUND as the "Mode of execution". You might find the BATCH option to be better as a learning tool, since you can study a listing of the batch job before it is submitted for execution.

Note: It is best to bring in as much data as possible. It is recommended that you specify ARCHIVE (the relative GDG can be specified on the execute, such as 0, -1, -2, -3, etc.) to use the monthly archive tapes. Rerun a second time using the online files in order to have as much history as possible for analysis. Remember, you have to change the "Number of Months to select" to bring in more than one month's data.

8. Press END and follow the panels until you reach the MICF Execution screen. Code Y for "Override Execution Options" and for "Edit Generated Jobstream". Press ENTER.

Review the "Execution Options" and press END to generate the batch update job.

Note that in the JCL portion of this job is a DD statement for CAPCOMDS. The DSN should refer to the SAS Data Library that will contain the capacity database members for this job. By the time this batch job has been generated, this SAS Data Library should have been allocated.

9. Submit the job and track its progress. The output should have a PROC CONTENTS listing of all the variables in the member that has been created. It should also contain a listing of the CA MICS performance data that were input to the job.
10. You will need to run the query you just built on a regular basis. The execution interval depends on the timespan selected; ie., days, weeks, months. This may be accomplished by one of the following methods:
 - a. Production Reporting. The private query will need to be copied to the SHARED inquiry catalog (MWF 4;3;4).
 - b. Run the JCL from the inquiry within a production scheduling system such as CA-7.
 - c. Run the PRIVATE inquiry each day (MWF 4;3;1).

11. For this exercise, copy this query to the SHARED capacity database.

This enables the scheduling of this query and makes it available for all three of the scheduling options listed in Step 10.

- a. Go to MWF 4;3;4 to copy the PRIVATE query to the SHARED capacity database.
 - b. Go to MWF 4;3;1 and delete the original Private query (you can always copy the SHARED copy back if the need arises) to eliminate duplicate data.
12. You should now have one copy of capacity history in your sharedprefix.CAPACITY This copy will be backed up by your DAILY or MONTHLY CA MICS backup job for your primary unit.

Note that you will need to set up your own backup and restore jobs for your private capacity database by following the directions in section 12.2 of this guide.

B.1.2 Generate the CPU Forecast

You can use the following procedure to create a forecast using the CA MICS Capacity Planner database as input.

Assumptions:

- o The input data are stored within the shareprefix.CAPACITY database.
 - o Linear Univariate Modeling will be used in this exercise.
 - o The forecast may fail if there are not enough historical data.
1. Go to Capacity Planning Analysis or Option 2 on the Capacity Planner main screen to create a forecast.
 2. Insert a new inquiry. Code an Inquiry Name and Title. Write the Inquiry Name here: _____.

Choose Workload Forecasting (option 2) and press ENTER.

3. Choose 4 for Univariate Model Forecasting and press ENTER.

4. Enter the parameters necessary for Univariate modeling.
See Figure B-3.

Use the same database that you used in the Section C.1.1.
(Build Capacity Database). You recorded information in
Step 2 in the Build Capacity Database section.

```
/----- Univariate Model Forecasting -----\  
|Command ==>  
|  
|Composing CA MICS Inquiry: QC1F1 - Quick Start Forecast for CPU  
|  
|Report title ==> Quick Start Forecast for CPU  
|  
|Selection criteria:  
|      Input file ==> QC1 - Quick Start Exercise  
|      Start date ==> _____ (ddmmyyyy)  
|      SYSID ==> XAD1  
|      Zone ==> 1  
|      No. months ==> 8          (1-9999)  
|      Save forecast ==> AGE      (YES/NO/AGE)  
|      Specify CAPAPU Values ==> N (Y/N)  
|Forecast data element ==> CPUPCBSY  
|Forecast length (months) ==> 18 (1-9999)  
|Confidence limits (percent) ==> 90 (70/90/95)  
|  
|Time element terms: Linear ==> YES (YES/NO)  
|      Quadratic ==> NO (YES/NO)  
|      Cubic ==> NO (YES/NO)  
|Specify extended options ==> N (Y/N)  
|-----/
```

Figure B-3. Univariate Model Forecasting Control Parameters

If you code Y after "Specify extended options", then an "Extended Options Menu" will appear. The last line of this menu has the option "Transform historical observations". If you code Y for this, another panel will appear. This panel provides the following options for the user to enter SAS code that will modify the historical data:

- Statements to transform historical data
- Statements to reverse history transformations

fit changes in your hardware environment as in these examples:

When the first option is chosen, the original data could be transformed to show a 30% faster CPU processor on June 9.

```
IF DATE GE '09JUN02'D
  THEN _TRAN_ = 1.3 * &VNAME;
ELSE _TRAN_V_ = &VNAME;
```

Note that the symbolic variable names used here are

&VNAME = the original dependent variable

_TRAN_V_ = the transformed dependent variable

When the second option is chosen, the transformation from the first option could be reversed by these statements:

```
IF DATE GE '09JUN02'D THEN DO:
  _ESTS      = _ESTS / 1.3;
  _LCL       = _LCL  / 1.3;
  _UCL       = _UCL  / 1.3;
END;
```

You can END out of the "Transform Historical Observations" screen to go back to the "Extended Options Menu".

If there are specific data points that you would like to delete from the historical data, you can choose the "Delete historical data observations" screen. This screen allows you to enter the dates for those observations that need to be deleted. Press END twice to go back to the "Univariate Model Forecasting" screen.

5. Press END again to bring up the "Inquiry Processing" screen. Choose EXECUTE and press ENTER.
6. Go to MWF 2;4;2 to copy this PRIVATE query to SHARED MICF query catalog.

NOTE: You may need to ask the CA MICS Administrator to copy this query to the CA MICS complex SHARED MICF catalog. Only inquiries in the SHARED MICF catalog can update data within sharedprefix.CAPACITY.

7. Go to MF 4;3;2 and execute this query (as recorded in Step 2 of this section). Ensure that the query you are running is from the SHARED MICF catalog. On the execution, specify update SHARED database forecast:

```

/----- Execution-Time Parameters -----\
|Command ==>                               Scroll ==> CSR |
|Press END to execute the inquiry. Enter CANCEL to cancel inquiry execution. |
|Executing CA MICS Inquiry: QC1F1 - Quick Start Forecast for CPU |
| |
|-----\
| Update forecast file (YES/NO/AGE)? Save forecast to Private or Shared DB. |
|AGE                                - SHARED _____ |
| |
| |
| |
| |
| |
| |
| |
| |
| |
| |
|-----\

```

Figure B-4. Execution-Time Parameters Screen

The shareprefix.CAPACITY should now contain a FORECAST and HISTORY.

B.1.3 Create the CPU Worksheet

Use the following procedure to create a worksheet using forecast and history from the CA MICS Capacity Planner database as input.

1. Choose Option 3 or the Capacity Planner Worksheet Facility from the Capacity Planner main screen.

The history and the forecast data created in previous steps in this appendix will be used as input.

2. Enter CREATE on the command line to create a new worksheet.
3. Enter the following:
 - a. Use the file recorded in Section C.1.1, Step 2 as the Input File.
 - b. Specify Zone 1 (or the Zone that you selected).
 - c. Specify your site dependent SYSID.
 - d. Specify CPUPCBSY and Y to include Forecast.
4. Press ENTER and then PF3 or END
5. Type PRODUCE to generate the Worksheet creation job.
6. After the job is complete, TYPE the word ENTER on the same line where you typed PRODUCE.
7. Enter worksheet forecast in the worksheet forecast column. Press PF3 or END when complete.
8. Type the word APPLY on the same line to generate a job that applies the worksheet forecast.
9. The worksheet file is now stored in your Private capacity file.

You can verify the worksheet's existence using MWF 4;3;1 (see figure B-5).

10. Go to MWF 4;3;4 to copy your WorkSheet to sharedprefix.CAPACITY database.

Specify Merge Forecast = Y

```

/------ Private File Definition Selection -----/
|Command ==>                                Scroll ==> PAGE |
|
|Userid: RUEDA02                            Catalog group      ==> QUICKS
|Prefix:                                    Name (subset display) ==> __
|Merge private forecast file ==> Y (Y/N)
|
|Line Cmds: S Select
|
|Cmd      Name                        Title                      Type        Status
|- - - - -|-----|-----|-----|-----|-----|-----|-----|
|_   QC1       Quick Start Monthly CPU Data                    F    Forecast Only
|S   QC2       Quick Start (QC1)                               F    Forecst/Worksheet
|***** BOTTOM OF DATA *****|
|
|
|
|
|
|
|
|
|
|
|
|
|

```

Figure B-6. Private File Definition Selection Screen

11. A SAS PROC CONTENTS executed on your sharedprefix.CAPACITY will report the existence of history, forecast, and worksheet files. (See naming convention as described in Section 4.4 of this guide).

B.1.4 Build Presentation Graphics

The following procedure documents the required steps for creating charts from CA MICS Capacity Planner database history, forecast and worksheets.

1. Choose Option 2 or Capacity Planning Analysis from the Capacity Planner main screen.
2. Insert a new inquiry and choose Workload Forecasting. Enter the name of the analysis here:_____.

3. Choose 6 for Presentation Graphics.
4. Specify the following:
 - a. Specify the Original file used in the previous sections of this Quick Start Tutorial.
 - b. Specify Y for Worksheets.
 - c. Use Color only if you have host graphic support available on your system.
5. Execute the inquiry (See Figure B-7).

```

----- Presentation Graphics -----
|Command ==>
|Enter a ? in any data entry field for more information on valid values.
|Composing CA MICS Inquiry: CPUM1 - Actual vs. Forecast CPU
|
|Report title ==> CPU: Actual vs. Forecast (by Month)
|
|Selection criteria
|      Input file ==> QC1 - Quick Start
|      Start date ==> _____ (ddmonyyyy)
|      SYSID ==> XAD1
|      Zone ==> 1
|      No. months ==> 12 (1-9999)
|      Specify CAPAPU Values ==> N (Y/N)
|
|Data element      ==> CPUPCBSY
|Previous forecast date ==> _____ (ddmonyyyy)
|Y-Axis label      ==> Pct Processor Busy
|Plot type         ==> COLOR (COLOR/PRINTER)
|Include Worksheets ==> Y (Y/N)
|
|Specify extended options ==> N (Y/N)
|
-----

```

Figure B-7. Presentation Graphics Control Screen

Appendix C: INTERFACE WITH CA MICS Q&R WORKSTATION

The CA MICS Capacity Planner provides an interface to the CA MICS Q&R WorkStation, a windows-based product that provides advanced graphics and reporting facilities on your PC.

Once you enable the interface, you can use the CA MICS Q&R WorkStation product to generate presentation graphics and perform further analysis of the data using various tools and facilities provided by the product, including:

- o Build queries to extract data from the shared capacity database to your PC.
- o Build presentation graphics in Microsoft Excel - the CA MICS Q&R WorkStation delivers a reporting facility that automatically generates charts using CA MICS-provided Excel macros. For advanced Excel users, the reporting facility allows you to further analyze and enhance the presentation of the data using the extensive features and functions in Excel.
- o Publish your reports on the intranet for fast, one-time delivery to your users across the enterprise. CA MICS Q&R WorkStation enables you to view the data in graphical and tabular format via a web browser.

Implementation

Perform the following steps to implement the interface between the CA MICS Capacity Planner and CA MICS Q&R WorkStation.

Note the interface allows you to access data in the shared capacity database only. Therefore, you must move data from the private database to the shared database if you wish to access the private files.

- ____ 1. Locate the member CAPMDB in sharedprefix.MICS.CNTL. If the job exists, proceed to Step 3. Otherwise, perform Step 2 to generate the job.

- _____ 2. Edit `sharedprefix.MICS.PARMS(JCLGENUC)` so that it contains the following:

`CAPMDB`

Submit the job in `sharedprefix.MICS.CNTL(JCLGENUC)`. Ensure that there are no error messages in `MICSLOG` or `SYSTSPRT`, that the `MICSLOG` contains the normal termination message, `BAS10999I`, and that the job completes with a condition code of zero.

- _____ 3. Submit the job in `sharedprefix.MICS.CNTL(CAPMDB)`. Ensure that there are no error messages in `MICSLOG` or `ISPLOG` and that the job completes with a condition code of zero.

The `CAPMDB` job builds a meta file that describes the content of the shared capacity database. As such, the `CAPMDB` job must be rerun whenever the content of the shared capacity database has been altered. Otherwise, the information in CA MICS Q&R WorkStation will be inaccurate.

The meta file built by `CAPMDB` is stored in `sharedprefix.MICS.CAPACITY`. It is used as input to the meta database build process in CA MICS Q&R WorkStation.