

Technical note

AD-A193 416

Host Computer System Capacity Management Procedures

Norman W. Watts, ACT-130
Paul Connolly, TSC
Robert Goettge, AST, Inc.

DTIC FILE COPY

2

DISTRIBUTION STATEMENT
Approved for Distribution

February 1988

DOT/FAA/CT-TN87/43

This document is available to the U.S. public
through the National Technical Information
Service, Springfield, Virginia 22161.

SDTIC
ELECTED
JUN 03 1988
E



U.S. Department of Transportation
Federal Aviation Administration

Technical Center
Atlantic City International Airport, N.J. 08405

88 6 2 021

Technical Report Documentation Page

1. Report No. DOT/FAA/CT-TN87/43	2. Government Accession No.	3. Recipient's Catalog No.	
4. Title and Subtitle Host Computer System Capacity Management Procedures		5. Report Date February 1988	6. Performing Organization Code ACT-130
7. Author(s) Norm Watts, Paul Connolly and Robert Goettge		8. Performing Organization Report No. DOT/FAA/CT-TN87/43	
9. Performing Organization Name and Address Federal Aviation Administration Technical Center Atlantic City International Airport, NJ 08405		10. Work Unit No. (TRAIS)	11. Contract or Grant No. T0503M
12. Sponsoring Agency Name and Address		13. Type of Report and Period Covered Technical Note 1987	
15. Supplementary Notes Other major contributors to this report were; Gary Morfitt, Robert Wiseman, Gary Wright, and Frank Yazek.		14. Sponsoring Agency Code	
16. Abstract The Federal Aviation Administration's Advanced Automation Program Office has recognized the need for monitoring and assessing the National Airspace System's operational performance and for long term planning during the life-cycle of the Host Computer System. The assessment of the operational performance involved the acquisition and analysis of field measurement data, while the long-term capacity planning entails execution of a Host Computer System analytical model using current and project traffic and other system loads. The procedures document defines the activities to be executed in: (1) measuring and monitoring operational performance, (2) measuring projecting system workloads, (3) predicting system performance using an analytical performance model, and (4) analyzing and reporting current and predicted future performance of the Host Computer System.			
17. Key Words Host Computer System Capacity Management Procedures Performance Modeling and Measurement National Airspace System		18. Distribution Statement This document is available to the public through the National Technical Information Service, Springfield, VA 22161	
19. Security Classif. (of this report) UNCLASSIFIED	20. Security Classif. (of this page) UNCLASSIFIED	21. No. of Pages 49	22. Price

TABLE OF CONTENTS

<u>Section</u>	<u>Page</u>
1. INTRODUCTION	1
1.1 PURPOSE	1
1.2 SCOPE	1
1.3 DOCUMENT ORGANIZATION	1
2. CAPACITY MANAGEMENT PLAN PROCEDURES OVERVIEW	1
2.1 CAPACITY MANAGEMENT OVERVIEW	1
2.1.1 Workload Measurement Activity	2
2.1.2 DR&A of Workload and Performance Measurements Activity	2
2.1.3 Workload and Performance Measurement Reporting	3
2.1.4 Performance Modeling Activity	3
2.1.5 Site Performance Prediction and Site Analysis Activity	3
2.1.6 Capacity Management Reporting Activity	4
3. WORKLOAD MEASUREMENT ACTIVITY	4
4. DR&A OF WORKLOAD AND PERFORMANCE MEASUREMENTS ACTIVITY	5
4.1 RAW SITE PERFORMANCE MEASUREMENT DATA INPUTS	5
4.2 CLASSES OF PROCESSED PERFORMANCE DATA	6
4.3 DR&A AND PERFORMANCE ANALYSIS PROGRAMS AND TOOLS	7
4.4 DATA SELECTION	10
4.5 DR&A PROGRAM AND TOOL SELECTION	11
4.6 EXECUTION OF DR&A TOOLS AND PROGRAMS	11
5. WORKLOAD AND PERFORMANCE MEASUREMENT REPORTING ACTIVITY	11
6. PERFORMANCE MODELING ACTIVITY	15
7. SITE PERFORMANCE PREDICTION AND SITE ANALYSIS ACTIVITY	16
8. CAPACITY MANAGEMENT REPORTING ACTIVITY	18
9. CAPACITY MANAGEMENT OVERVIEW	19
9.1 CAPACITY MANAGEMENT PROCEDURES	20
APPENDIX A - SAMPLE SITE CAPACITY PLAN	A-1

LIST OF ILLUSTRATIONS

<u>Figure</u>		<u>Page</u>
1-1	CAPACITY MANAGEMENT ACTIVITIES AND INTERRELATIONSHIPS	2
4-1	PERFORMANCE DATA REDUCTION FLOW	12
5-1	AST DATA FLOW	14
5-2	SPSS SCATTER PLOT DATA FLOW	14
9-1	TIME-ORDERED CAPACITY MANAGEMENT STEPS	19

LIST OF TABLES

<u>Table</u>		<u>Page</u>
4-1	PERFORMANCE DATA REDUCTION TOOL CAPABILITIES	12

Accession For	
NTIS GRA&I	<input checked="" type="checkbox"/>
DTIC TAB	<input type="checkbox"/>
Unannounced	<input type="checkbox"/>
Justification	
By _____	
Distribution/	
Availability Codes	
Dist	Avail and/or Special
A-1	



EXECUTIVE SUMMARY

The objective of the Host Computer system (HCS) Capacity Management Plan is to continuously monitor and project HCS operational performance to insure adequate capacity and response time requirements are satisfied during HCS's life cycle in the face of traffic load increases and functional enhancements to the operational software. This document describes the organizational planning processes and procedures necessary to assure this objective is satisfied.

Two critical tools required to satisfy this objective are the development of an accurate, high resolution performance measurement tool and a validated HCS performance predictive model. The first of these, the HCS Performance Measurement Tool, reduces Air Route Traffic Control Center (ARTCC) recorded HCS operational performance data to provide highly accurate performance measurement information in key areas such as device utilization, program element utilizations, and end-to-end response times. The second key tool in executing the HCS capacity Management Plan is a queuing network model of the en route air traffic control (ATC) system. This predictive model, the HCS Performance Prediction Tool, is basic to execution of the Capacity Management Plan because it facilitates projections of HCS performance as a function of future traffic loads, other system load factors, operational software enhancements, and site dependent allocation of functions to devices. These new tools form the nucleus of the HCS capacity planning.

The procedures contained in this document define a methodology for determining current HCS operational workload and anticipated workload growth. The workload parameters which drive the system are identified and the tools required to extract these parameters from real-time ARTCC measurement data are described and their use explained. Execution of this plan culminates in a capacity planning report for each ARTCC. The report will examine each center's peak-day workload (such as active tracks, active flight plans, and track life) in order to establish device utilizations and response times to controller entered messages. The report will either precede all new releases of operational ATC software or, at a minimum, will be made annually. It will graphically summarize each site's performance predictions and make recommendations on how to best allocate functions to devices should predetermined service levels not be met, e.g., rearrange data base, shift resource monitoring from disk to tape, and upgrade the CPU.

In summary, the Host capacity management procedures have been established at the Federal Aviation Administration Technical Center, when implemented, will provide a complete characterization of the present and future HCS performance at each en route center.

1. INTRODUCTION.

1.1 PURPOSE. The purpose of this document is to provide Air Traffic Service with information and procedures for the development, operation and continuing maintenance of a Host Computer System (HCS) Capacity Management Process. This document complements the HCS Capacity Management Plan Technical Note.

1.2 SCOPE. The procedures described in this document define the activities to be performed to ensure that the HCS installed at the Air Route Traffic Control Centers has sufficient capacity to meet specified performance requirements throughout the life of the system.

1.3 DOCUMENT ORGANIZATION. The procedures are addressed as activities which must be executed in the Capacity Management Process. The following sections have been identified for this process:

- a. Section 2: Capacity Management Plan Procedures Overview
- b. Section 3: Workload Measurement Activity
- c. Section 4: Data Reduction and Analyses (DR&A) of Workload and Performance Measurements Activity
- d. Section 5: Workload and Performance Measurement Reporting Activity
- e. Section 6: Performance Modeling Activity
- f. Section 7: Site Performance Prediction and Analysis Activity
- g. Section 8: Capacity Management Reporting Activity
- h. Section 9: Procedures Associated with Executing the Capacity Management Plan

2. CAPACITY MANAGEMENT PLAN PROCEDURES OVERVIEW.

This section provides an overview of the procedures to be executed in the activities associated with Capacity Management. Capacity Management is predicated on the existence of a calibrated analytical model capable of site dependent performance prediction. For this reason, this document includes a significant amount of information on both data selection and data reduction in accordance with modeling requirements.

2.1 CAPACITY MANAGEMENT OVERVIEW. An overview of the Capacity Management activities and their relationship is shown in Figure 1-1. The major activities are: Workload Measurement, Data Reduction and Analysis (DR&A), Performance Analysis, Workload and Performance Measurement Reporting, Performance Modeling, Site Performance

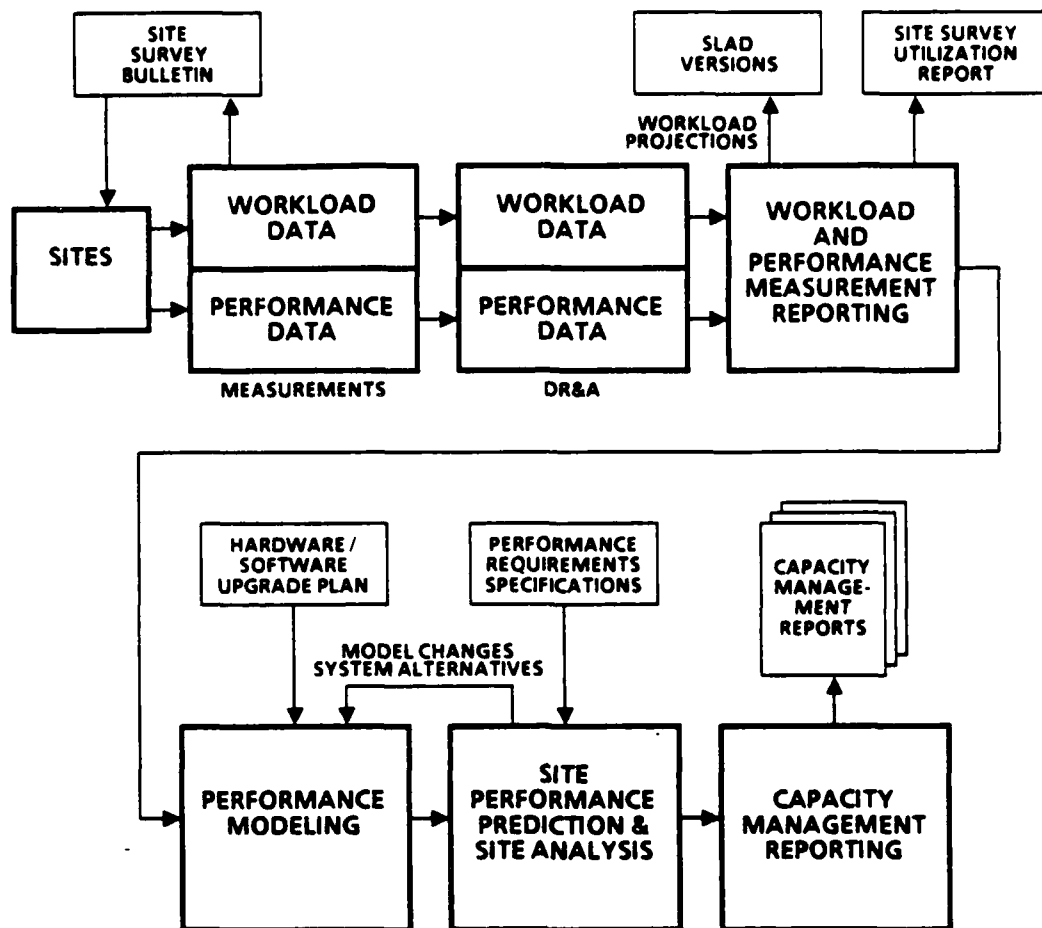


FIGURE 1-1. CAPACITY MANAGEMENT ACTIVITIES AND INTERRELATIONSHIPS

Prediction and Analysis, and Capacity Management Reporting.

2.1.1 Workload Measurement Activity. Workload Measurement is directed towards ensuring that adequate operational raw measurement data are obtained from each ARTCC for use in: writing a Site Survey Utilization Report, validating and updating a System Loads Analysis and Definition document and validating the HCS performance model. The starting point of this activity is writing a Site Survey Bulletin which defines the nature and extent of raw measurement data to be collected by each ARTCC. The site measurement data are forwarded to the DR&A of Workload and Performance Measurements Activity.

2.1.2 DR&A of Workload and Performance Measurements Activity. The DR&A of Workload and Performance Measurements Activity transforms raw Workload Measurement data into suitable forms for use in the Workload and Performance Measurement Reporting and Performance Modeling activities. The six classes of processed performance data are: Program Element (PE) Analysis (by calling PE), Response Time Summary, Response Time Analysis (by PE and device), Device Utilization, and System Overhead. The Performance Measurement DR&A tools are: Host Performance Measurement Tool (HPMT), HRT Reduction Tool (REDUC), Data

Analysis and Reduction Tool (DART), Response Time Tool (RTT), System Utilization Report Program (SURP), and the Timing Analysis Reduction Program (TARP).

2.1.3 Workload and Performance Measurement Reporting Activity. One function of the Workload and Performance Measurement Activity is directed towards generating HCS processed performance measurement data for use by the MITRE Corporation, the contractor responsible for providing the Federal Aviation Administration (FAA) with the current AAS System Loads Analysis and Definition (SLAD), dated August 1987 (see REF 1). MITRE will be involved in validating and updating the workload estimates and traffic forecasts contained in the SLAD using the Aircraft Management Program (AMP) data collected by the ARTCC sites in accordance with the Site Bulletin. The Workload and Performance Measurement Reporting Activity will forward the updated version of the AAS SLAD to the Performance Modeling Activity. Another important function of this activity is analysis of processed performance data for the purpose of generating the Site Survey Utilization Report. The Aggregate Statistics Tool (AST) and the SPSSTM statistics package, respectively, are used to generate composite statistics and scatter diagrams of HCS performance for use in this report.

2.1.4 Performance Modeling Activity. The Performance Modeling Activity (PMA) modifies the Host Performance Prediction Tool (HPPT) in support of model calibration, validation, and use by other activities. The model calibration process is initiated whenever a new release (hardware/software) of the HCS is tested at the FAATC or whenever an update to the SLAD is received. To calibrate the model, additions and changes to the model are made to reflect the changed hardware/software or workloads. The model is then calibrated so that its performance predictions are within certain tolerances of the performance of the test system as measured at the FAATC. The calibrated model is forwarded to the Site Performance Prediction Activity for use in predicting the performance of the sites and updating capacity plans.

Upon installation of the new system release at the sites, site workload and performance measurements are received by the PMA in support of model validation. Model performance predictions in processing the site workloads are compared to actual measurements. Differences between predicted and measured performance are normalized to produce a standard measure of significance. Differences which exceed a certain tolerance are forwarded to the Site Performance Prediction and Site Analysis Activity to assess their ramifications and determine resolutions.

The Performance Modeling Activity also modifies the model to reflect alternatives proposed by the Site Analysis Activity to alleviate any predicted capacity shortfalls.

2.1.5 Site Performance Prediction and Site Analysis Activity. The Site Performance Prediction and Site Analysis Activity uses the model,

interprets the model's results, and directs the model validation effort. Using the calibrated version of HPPT, this activity exercises the model to predict the performance of the HCS over time. The model parameters are incrementally changed to reflect the workloads as projected by the SLAD and any hardware/software changes forecast to occur over the planning horizon. Predicted response times and utilizations are compared to established performance criteria or guidelines and any shortfalls are identified and forwarded to the Capacity Management Reporting Activity (CMRA). Solutions for alleviating potential performance problems are proposed and through the assistance of the PMA, modifications to HPPT are made to reflect the proposed solutions so that their performance impact can be quantified. A ranked set of candidate solutions is then forwarded to the CMRA.

The Site Analysis function of this activity identifies the cause of the discrepancies in the expected performance (as predicted by HPPT) and the site measurement data. The ramifications of the observed differences are then assessed to determine if a modification to the model, site operations, or site procedures is warranted. If a change is made to HPPT to reflect the actual site's operations or procedures, the changed model is run to determine the long-term impact of the site's operation or procedures on performance. If the analysis determines that some aspect of the site's adaptation, operation, or procedures will significantly degrade long-term performance, then the degree of degradation and recommendations for change are forwarded to the CMRA.

2.1.6 Capacity Management Reporting Activity. The Capacity Management Process culminates in the Capacity Management Reporting Activity. This activity reports predicted site performance, potential operational problems or shortfalls in performance, and recommended solutions to performance shortfalls with their predicted performance impacts.

The primary output from this activity is a Site Capacity Plan which includes predictions of site workloads and performance over the planning horizon and includes any planned or recommended time-phased introduction of hardware or software upgrades to the site.

3. WORKLOAD MEASUREMENT ACTIVITY.

The starting point of Workload Measurement is to generate the Coordinated Site Survey Bulletin which defines the nature and extent of data to be collected at each ARTCC. The purpose of the Site Bulletin is to insure that collected site measurement data satisfies the requirements to support: SLAD validation and updates, measurement of current site dependent loads and levels of equipment utilizations, and validation of the performance model. Site data collection was coordinated to consolidate requests for ARTCC data into a single bulletin. The Workload Measurement Activity also involves obtaining raw measurement data at the FAA Technical Center whenever additional

testing using the new Host Load Tapes (HLT) with their three steady state traffic levels of 200, 400 and 600 tracks. (Note: the HLT tests are defined in REF 22-25.) These HLT tests permit operational type testing at Technical Center to determine the impact of HCS hardware and software enhancements on HCS performance prior to their field release.

The field, as well as the FAA Technical Center HLT testing, Workload Measurement Activity, culminates with the reception of the following HCS run time data:

- a. MMP data with bulletin designated options selected
- b. AMP data
- c. HRT data
- d. SAR data at bulletin designated SARC recording levels
- e. TAR data

4. DR&A OF WORKLOAD AND PERFORMANCE MEASUREMENTS ACTIVITY.

This section describes the procedures to be used in the reduction of raw performance measurement data. The objective of this DR&A activity is to transform the raw data collected from the HCS sites (ARTCC's) into suitable form for use in Workload and Performance Measurement Reporting and Performance Modeling Activities. The procedures defined in this section will only address: the input media to a DR&A or Performance Analysis program and tool; the selection of the best program and tool for generating a given class of processed performance data; and the output media from a DR&A program and tool to be passed on to next activity. It is important to note that many of the DR&A programs and tools used to support 9020 activities have limitations and inaccuracies which limits their usefulness to Capacity Management, in general, and Performance Analysis, in particular. This procedures document will only discuss the programs and tools available for use in processing the raw site measurement data for use in Workload Analysis and Performance Modeling. Information on how to use the DR&A tools is contained in the appropriate User's Manual.

4.1 RAW SITE PERFORMANCE MEASUREMENT DATA INPUTS. The raw performance measurement data received from the field sites are:

a. System Analysis Recordings (SAR) contain message data associated with each controller input action to HCS, data associated with each system generated output message, and data associated with information processed internally by the NAS software.

b. Timing Analysis Records (TAR) provide time traces of interrupts or system status changes during execution of the NAS operational software. TAR's are collected for external, service calls (SVC's), program, and input/output (I/O) interrupts plus exits from the NAS Monitor, program suspensions and normal program terminations. TAR's are output to the SAR files (tape/disk). TAR's data are often

used to generate pseudo-HRT data defined in the next paragraph.

c. High Resolution Timer (HRT) recordings provide time traces similar to TAR's but with a higher degree of resolution (1 microsecond for HRT vs 16 microseconds for TAR's) and for additional events during the execution of NAS operational software that are not covered by TAR's.

d. Monitor Minute Processor (MMP) recording provide data for reporting central processor utilization (CPU), channel utilization, storage utilization, PE utilization, and peripheral device utilization over selectable intervals of time (1 to 10 minutes). The MMP data uses the HRT files and recording media. Two factors that limit the usefulness of MMP data for application to Performance Analysis is the granularity of the data (averaged over the selected 1 to 10 minute time intervals) and, more significantly, the fact that all the data are rounded to the nearest percent, which for PE and device utilizations in Performance Analysis applications is far too coarse.

e. Aircraft Management Program (AMP) recording provides records of aircraft within an ARTCC control area for analysis of traffic loads by sector, fix, and airport, the results of which are incorporated into the SLAD.

4.2 CLASSES OF PROCESSED PERFORMANCE DATA. The raw performance measurement data collected from HCS sites are reduced into five classes of processed performance data:

a. Device Utilizations: Device utilization is defined to be the percentage of time that a device is busy during the measured time interval. The principal devices of interest for capacity management of the HCS are the primary CPU and the disks since these are the devices whose services are required to produce responses to controller input messages and process radar data for display on the PVDs.

b. PE Analysis: For the two major classes of device (CPU and disk), their utilizations are decomposed into percentages attributable to each of the Program Elements. Each PE contributing to the device busy time is analyzed for its individual contribution. Device utilization by a PE is dependent on the frequency of requests for service issued by the PE as well as the amount of service requested; both of these measures are reported by the measurement tools.

c. Response Times: Response times are defined to be the time interval between the arrival of an input event and the initiation of associated output events. The technical definition of response times used in HCS testing at the FAATC is defined in the HCS Engineering Requirement (and other documents referenced in the Requirement). It is important to understand the definition of the requirement in order to properly interpret the response time measurements produced by the various response time tools.

Response time tools measure HCS response times for each individual input and its associated outputs. The response times are measured over an interval of time and various statistics are computed from the response time samples. HCS requirements are specified for the mean and 90th percentile response time values. (99th percentile values are of some importance since such a requirement exists for the HCS component of ISSS.) Statistical measures are aggregated for both input/output message pairs and response time classes. An input/output pair is a unique combination of an input and output message type. Based on a set of assignment rules, these message pairs are further aggregated into one of five response time classes (classes 2-6). Radar data response times are assigned to class 2; controller messages are assigned to classes 3-6.

d. Response Time Analysis. For each device and PE, the component of the stimulus-response pair response time attributable to a PE for using or waiting to use the device.

e. System Overhead. The amount of CPU busy time attributable to such system overhead functions as PE dispatching and I/O and external interrupt processing.

4.3 DR&A AND PERFORMANCE ANALYSIS PROGRAMS AND TOOLS. There exist of variety of programs and tools to reduce and analyze HCS performance data. Several of these tools have overlapping capabilities yet the accuracy of the reports and suitability of the data for HPPT varies from tool to tool. The following section discusses some of the tools' limitations and their suitability for HPPT input and validation.

The programs and tools used to generate the five identified classes of performance data from the raw data are:

a. HPMT: This tool processes either HRT or pseudo-HRT data. HPMT produces system activity summaries from detailed event histories recorded on either tape. The histories yield virtually all measurement data required for site performance analysis and performance modeling, i.e., response times, device utilizations, detailed PE analysis and threads, branching frequency data, and input message arrival information. HPMT is the primary tool used to produce PE request frequency and service time data. For each PE, HPMT produces three statistical quantities -- execution rate, service time per execution, and utilization.

Execution Rate - This value is defined to be the total number of executions of a PE during the measured time interval divided by the length of the interval. HPMT counts PE executions differently than other tools (notably REDUC). Since response time is dependent on device utilizations and average service times, it is important that these parameters be accurately measured. To this end, HPMT counts only "useful" PE executions or executions in which the messages or data are processed (versus "null" executions in which the PE searches for work, but finds none). Clearly, the

average service time that results from counting only "useful" executions is larger than the average service time that results from counting all PE executions.

Service Times - For the CPU, service times are defined to be the amount of CPU time consumed by the PE per execution. For PEs which process multiple messages per execution, HPMT computes the service time per message processed in order to compute a message dependent execution time. HPMT reports the amount consumed by the PE plus any overhead consumed for HRT recording. In addition, to mean values, HPMT also reports the coefficient of variation of the CPU service time per execution.

Some PE service times depend upon the message type being processed. For these "message dependent" PE service times, HPMT will report a separate service time for each PE that calls the PE being analyzed. HPMT is limited in its ability to uniquely identify individual message processing intervals for several PEs, most notably HTI and CBC. Therefore, detailed service time data for these PEs must be carefully analyzed prior to using.

For disk service times, HPMT computes the number of disk accesses per PE execution and then calculates the average service time per disk access. The average disk service time per PE execution is therefore the product of these two values. Certain HCS configurations use multiple channels to access the disk devices. For these configurations, HPMT reports the disk utilization by channel and device. To determine the total disk utilization, the analyst must sum the device utilizations over all channels.

Utilization - HPMT reports device utilizations by PE as the percentage of time during the measurement interval that the device was busy servicing requests by the PE.

HPMT has a limitation on the accuracy of its I/O message pair aggregation into response time classes. Because HRT does not record the input source and output destination of the messages (and therefore cannot determine if an output is destined for an entering device), HPMT uses a default mix parameter to assign a few I/O pairs to classes. For an estimate of these mix parameters, DART can be used provided it has properly paired input and output messages.

Because of the scope, accuracy, and suitability for HPPT of the data produced by HPMT, this tool is the primary one used for the detailed performance analysis requirements of modeling. As a result, real and pseudo HRT data are of prime importance in the Capacity Management Process. HPMT should be used in accordance with its User's Manual.

b. **REDUC**: Uses either HRT or pseudo-HRT data to produce device utilization and PE behavior information. It presently is the only source of information on radar processing response times (class 2). Because there are severe timing accuracy limitations on much of the

information processed by REDUC, it is marginally useful for performance modeling applications. REDUC can be used to produce total CPU and disk utilizations although its CPU utilization will be slightly less than SURP's because it fails to record several minor CPU contributors. REDUC's disk utilization does not have this problem.

REDUC can also be used as a back-up for HPMT in producing PE CPU service times; REDUC does not compute disk service times by PE. However, REDUC and HPMT differ in how HRT overhead is attributed to service time. The PE service times are directly comparable only when REDUC and HPMT both use pseudo-HRT data as input. In addition, whenever HRT is ON, REDUC accumulates some CPU time into a fictitious PE called DBUF. In general, HPMT and REDUC will not produce the same average service times due to the difference in the way they count PE executions. REDUC should be used in accordance with FAA, REDUC User's Manual, NASP-9211-11H.

c. SURP. This DR&A program processes the MMP data. Its outputs (if selected) are:

- Processor Sample Report
- Processor Summary Report
- Storage Sample Report
- Storage Summary Report
- Channel Sample Report
- Channel Summary Report
- Peripheral Sample Report
- Peripheral Summary Report
- Interrupt Summary Report
- Program Analysis Report
- Data Address Report
- Instruction Address Report

SURP is the primary tool used to produce utilizations for all devices of interest in the capacity management. The real-time overhead to produce the data used by SURP is small which makes it a valuable tool for producing accurate measurements of utilizations and determining the overheads caused by other performance measurement functions. Typically, SURP reports utilization on a minute-by-minute basis with each value being the average utilization of the devices over the previous minute. Since device utilizations over a longer period are frequently needed, these minute-by-minute interval utilizations need to be averaged to compute utilizations over longer time periods. This process is currently done manually.

The SURP outputs used in site performance analysis and performance modeling are the one minute processor, channel and peripheral reports (sample or summary).

SURP also has an option to produce a magnetic tape output of MMP processed data that is compatible with the requirements of the BMDP™

statistical package. This option is used as an input source to the SPSS programs to generate scatter diagrams in the Site Survey Report (See section 5.1.b for the list of the scatter diagrams that can be generated from the BMDP tape). Use is in accordance with IBM, SURP, B093-SURP-UM.

d. DART: Used to produce response time reports. It accomplishes this by matching input and output messages obtained from SAR data files. Because of the half-second clock resolution and numerous message mismatching errors, DART response times have limited value to capacity management activities. Use is in accordance with FAA, DART User's Manual, NAP-9247-19H.

e. RTT: Operates in much the same manner as DART to analyze message response times from SAR data. Improved message pairing logic corrects many, but not all, of DART erroneous message pairings. The clock resolution of SAR data, again, limits RTT's value to capacity management related activities. Use is in accordance with RCA Corporation, RTT User's Manual, ND-H-ATC-M-4002.

f. TARP: Used to create a pseudo HRT tape from TARs data. Use is in accordance with FAA, TARP User's Manual, NAP-9227-15H, latest version.

4.4 DATA SELECTION. The criteria for selecting raw measurement data to be reduced and used in Workload Analysis and Performance Modeling activities is important because not all collected data are applicable and/or satisfactory for their use. The basic ground-rule is to find continuous intervals during which the HCS was in "steady-state." For capacity management purposes, steady-state refers to constant trackload and operational settings such as data recording levels and functional switches (conflict alert, Minimum Safe Altitude Warning (MSAW), etc.). Measurement during steady-state is important because this assumption is basic to data interpretation procedures as well as performance modeling predictions. In terms of trackload, steady-state is considered to be a variation in tracks of no more than five percent. Because HCS performance is significantly impacted by the operational conditions and recording levels, it important to know their status during the interval raw site measurement data are being reduced. Generally, this information can be checked by examining recordings of operator control messages. To ensure the proper setting of functional switches the activity of certain PEs should be checked (e.g., CAD for conflict alert, MRM for REMON).

Selection length of the time interval to be processed depends somewhat on the accuracy requirement of the processed data. Given the satisfaction of the steady-state restriction, the interval should be of sufficient length to "average out" statistical variations. For total device utilizations, several minutes of data are usually sufficient. However, if more detailed information on individual PE's or response time pairs is required, the length of the interval will have to take into account the arrival rates of the PE's or messages.

If one desires to compute mean values, a minimum of 30 samples is generally required; 90th percentile computations require a minimum of 100 samples.

Currently, HCS HRT and SAR recordings can have "data gaps." These data gaps appear to be associated with buffer overflow problems during HCS data acquisition. The TARP programs can be used to identify gaps in TAR's recording on the SAR tape while creating the pseudo-HRT tape. Data gaps on the real HRT tape can be detected using the Tape Summary mode of HPMT. Because HPMT does not have an option to discard data, it is sensitive to data gaps and, therefore, the selected time interval for HPMT processing must be gap free.

4.5 DR&A PROGRAM AND TOOL SELECTION. The performance data reduction flow is shown in Figure 4-1. This figure shows that the primary reduction tools SURP, HPMT, and REDUC use data recorded on the HRT tape (SURP use MMP data which is recorded on the HRT tape by HCS). Optionally, HPMT and REDUC can use a pseudo-HRT tape generated by TARP from TAR records on the SAR tape. Two response time tools use I/O recordings contained on the SAR tape; DART uses this information directly while RTT uses I/O LOG data produced by DART.

The capabilities of the tools (discussed in Section 4.3) for producing the five classes of performance data are summarized in Table 4-1. Because of both its coverage and accuracy for producing the data required for capacity management, HPMT is the primary tool used to reduce the raw performance data. Other tools which provide similar functionality to HPMT are used as back-ups whenever HPMT's required input data are unavailable.

4.6 EXECUTION OF DR&A TOOLS AND PROGRAMS. All of the DR&A tools/programs except RTT are executed on the HCS support processor. (RTT can only be run on the IBM 4341 computer system using SUPERWYLBUR.) Selection of data reduction options and the time intervals to be analyzed are defined in the User's Manual for the hosted DR&A programs/tools.

5. WORKLOAD AND PERFORMANCE MEASUREMENT REPORTING ACTIVITY.

The functions of the Workload Performance Reporting Activity are to:

- (1) provide measurement data (primary AMP data) to the MITRE Corporation which is under contract to the FAA to update and maintain versions of the System Loads Analysis and Definition (SLAD) document;
- (2) provide the updated SLAD to the Performance Modeling Activity for updating the model's workload parameter data base; and (3) generate the Site Survey Utilization Report using the Aggregate Statistics Tool (AST) and SPSS. Since the first two functions are straightforward and are primarily data passing functions, only the generation of the Site Survey Utilization Report is discussed below.

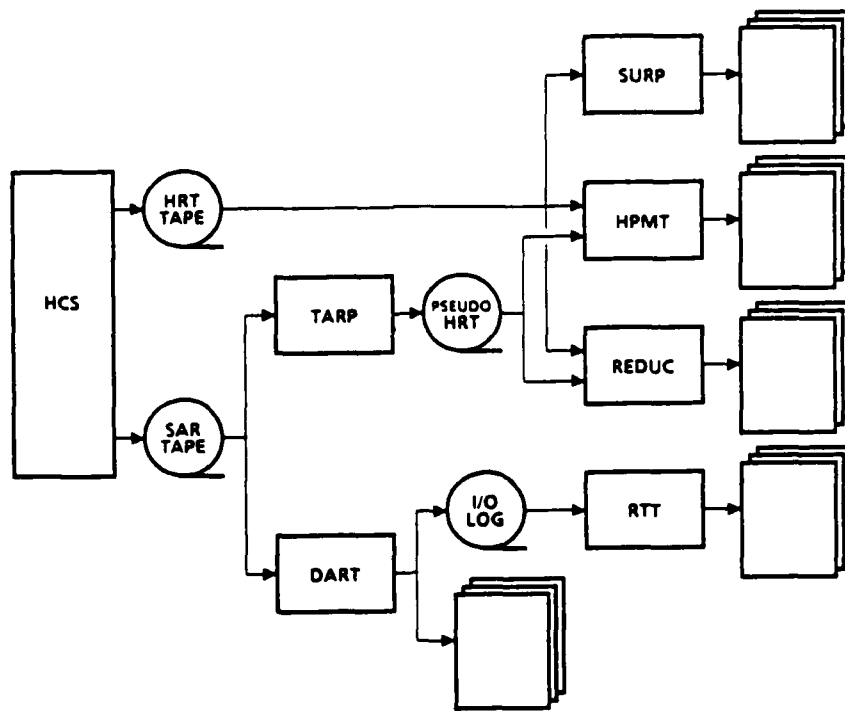


FIGURE 4-1. PERFORMANCE DATA REDUCTION FLOW

TABLE 4-1. PERFORMANCE DATA REDUCTION TOOL CAPABILITIES

TOOL	TOTAL DEVICE UTILIZATION	CAPACITY			RESPONSE TIME ANALYSIS	SYSTEM OVERHEAD
		PE ANALYSIS		RESPONSE TIME		
		CPU	DISK			
HPMT	X	X	X	X (CLASS 3-6)	X	
SURP	X					
REDUC	X	X		X (CLASS 2)		X
DART				X (CLASS 3-6)		
RTT				X (CLASS 3-6)		

The Site Survey Utilization Report reflects the current workload at each ARTCC site. This report is not part of the capacity report but will be generated when conditions warrant it or when specifically requested by AT or a site manager. The report includes:

a. Composite statistics of HCS performance which are generated by executing AST using processed performance data inputs in accordance with Figure 5-1. AST effectively creates a multi-site data base using appropriate processed measurement information. This multi-site data base is then input to the SPSS statistical package which generates composite statistics over all sites. This figure defines the source of the inputs and a flow of the steps involved in generating these inputs. The composite statistics generated by AST are:

1. SVC ratio statistics:

- a. Number of non-suspended SVC executions
- b. Number of executions per minute
- c. Percent of CPU utilization (hereafter called CPUu) per 1000 executions
- d. Percent of CPUu per 100 tracks
- e. Percent of CPUu per 100 active flight plans

2. PE report ratio statistics

- a. Number of full executions
- b. PE executions per minute
- c. Percent of CPUu per 1000 executions
- d. Percent of CPUu per 100 tracks
- e. Percent of CPUu per 100 active flight plans

3. CPU report ratio statistics

4. Percent of CPU PE processing

5. Composite DART log reports

- a. Input message report
- b. Output message report
- c. Summary message report

6. Composite DART response report

- a. Message response report

b. The scatter diagrams of single site performance measurement data are generated by the SPSS statistical package. The input to SPSS is the BMDP output tape option from the SURP DR&A

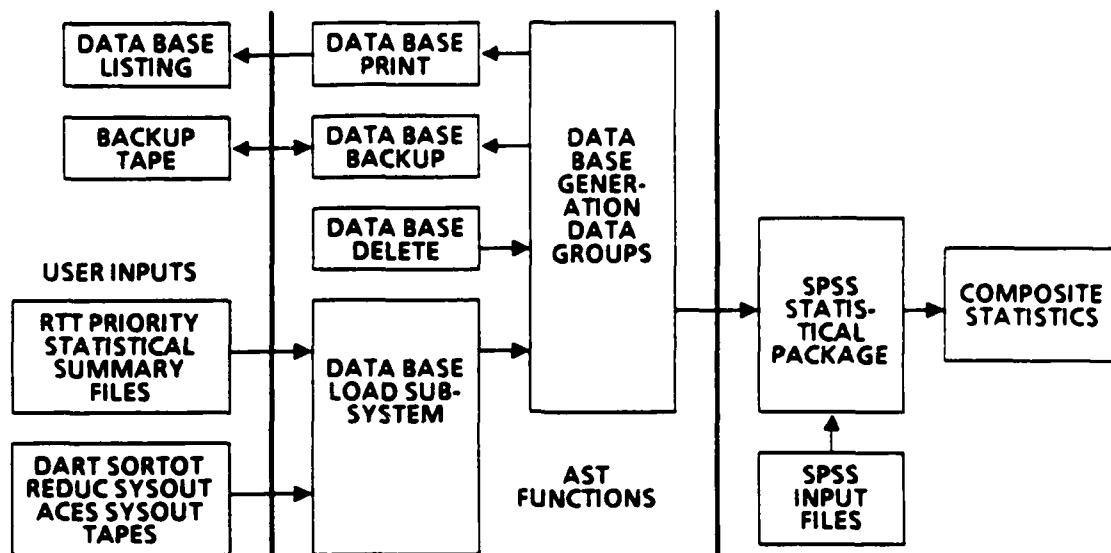


FIGURE 5-1. AST DATA FLOW

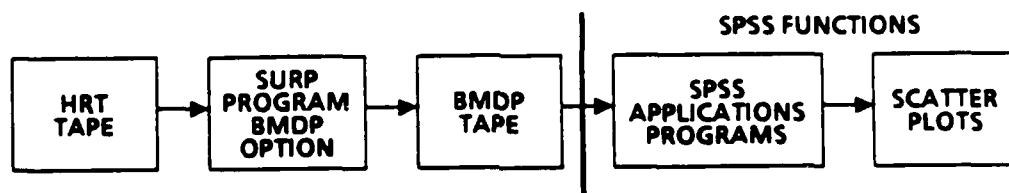


FIGURE 5-2. SPSS SCATTER PLOT DATA FLOW

program. Figure 5-2 defines the inputs and data flow associated with scatter diagram generation. The scatter diagrams are:

1. Regression equation
2. CPUu vs active flight plans
3. CPUu vs selector channel utilization
4. CPUu vs average disk/tape utilization
5. Track count vs average display channel utilization
6. Active flight plans vs time of day (TOD)
7. Track count vs TOD
8. CPUu vs TOD
9. Average selector channel vs TOD
10. CPUu maximum value vs active flight plans
11. CPUu maximum value vs tracks
12. CPUu maximum value vs channel utilization

All of the above composite statistics, regression equations, and scatter diagrams are the basis for the Site Survey Utilization Report.

This report reflects the current operational loading of the HCS and compares the loading of each site to all other sites. This report can provide information to MITRE relative to the SLAD's previously forecasted ARTCC loads versus actual measured ARTCC loads.

6. PERFORMANCE MODELING ACTIVITY.

The Performance Modeling Activity (PMA) is responsible for making all Host Performance Prediction Tool modifications in support of model calibration, validation, and use by other activities. To accomplish its mission, this activity performs three basic functions: (1) calibration of the model to measurements obtained primarily from the sites, or secondarily from FAA Technical Center tests using simulated (Host Load Tape) scenarios, (2) validation of the model to measurements obtained from actual site operations, and (3) use of the model to evaluate the performance impact of hardware/software/procedural changes proposed to alleviate predicted system bottlenecks.

The objective of the model calibration function is to develop a version of the model whose performance predictions are within the required tolerances of measurement data produced by the sites or by FAATC test HCS system processing HLT scenarios. One of the design goals in the development of the Host Load Tapes was to create a scenario that was representative of the site workloads in 1995. One of the design goals of HPPT was to represent the salient workload characteristics and operational procedural settings as model input parameters rather than to embed workload and procedural effects in the service time equations and branching frequencies. Assuming these goals have been achieved, a model whose service times and branching frequencies are within a specified tolerance of times and frequencies for a known workload (or set of workloads) can be used to make performance predictions for other workloads and operational settings. The process of ensuring that the model's performance predictions (e.g., service times and branching frequencies) are within the required tolerance is referred to as model calibration. Once a model has been calibrated, it needs to be recalibrated only if changes are made to the system (i.e., hardware and software) which would impact the calibrated service times and branching frequencies.

One additional model calibration function performed by the PMA is to update the model's data base of future workload parameters as defined in the SLAD. Anytime changes to the SLAD are made, the Workload and Performance Measurement Reporting Activity will forward these changes to the PMA for updating the workload data base. The model is then forwarded to the Site Performance Prediction and Site Analysis Activity to determine whether a new set of performance predictions using the updated workload is required.

In conjunction with the Site Analysis Activity (described in Section 7), the objective of the model validation function is to develop a version of the model whose performance predictions are also within the required tolerances of measurement data produced by the field sites.

To attain this objective, the model validation function first identifies discrepancies between the model's predictions and actual site measurements by running HPPT with workloads and operational settings identical to those of the site at the time the measurements were collected and then comparing the model's results to the measurements. Discrepancies whose normalized percent differences exceed a threshold are forwarded to the Site Analysis Activity to determine their cause. The magnitude of the threshold used to identify discrepancies requiring further analyses is currently based on engineering judgement. As experience with the Capacity Management Plan increases, this engineering judgement will be modified in accordance with the dictates of the increasing knowledge base.

Although testing at the FAATC with the HLT scenarios is primarily intended to test new functional enhancements to the Host (e.g., VFR intruder) under 1995 workload conditions, the tests can produce PE service time data which is useful for calibrating the model. Comparison of its measurements to field data can also be used to determine the representativeness of the SLAD's HCS workload, which has been established in the HLT scenarios. Of course, since the site's adaptation hardware, or operational procedures differ from those under test at the FAATC, then the calibration procedure can identify those differences.

Upon careful review of the identified model-measurement discrepancies, the Site Analysis Activity may recommend changes be made to the model. The Performance Modeling Activity is then responsible for implementing those changes and rerunning the model to ensure that its predictions are within specified tolerances. This validated model is then forwarded to the Site Performance Prediction Activity to re-forecast site performance.

The third type of activity performed by the PMA is to modify the model to represent alternative hardware, software, and operational settings recommended by the Site Performance Prediction and Site Analysis (SPPSA) to alleviate potential performance bottlenecks. These modified model versions are then forwarded to the SPPSA for their use in predicting the performance of these upgraded systems.

7. SITE PERFORMANCE PREDICTION AND SITE ANALYSIS ACTIVITY.

The Site Performance Prediction and Site Analysis Activity is responsible for exercising the model, interpreting the model's results, and directing the model validation effort. To accomplish its mission, this activity performs two primary functions: (1) predicting site performance using either a calibrated or validated version of HPPT and (2) analyzing site measurement data and site model results to explain differences in expected and measured performance.

The Site Performance Prediction (SPP) function involves exercising the models developed by the Performance Modeling Activity to predict site performance over time. Once HPPT has been calibrated to a

hardware/software configuration, the SPPSA exercises this model by setting the workload, operational settings, and other parameters to those characterizing the site. The model parameters are incrementally changed to reflect the workload and any hardware/software upgrades forecast to occur over the capacity planning horizon (currently 1995). A series of model experiments is then performed to produce performance predictions as they vary over the planning horizon.

Predicted performance metrics (e.g., response times and utilizations) are compared to performance criteria or guidelines (e.g., NAS-MD-318) and any shortfalls are identified and forwarded to the Capacity Management Reporting Activity (CMRA). The SPPSA attempts to find solutions for alleviating potential performance problems and enlists the aid of the PMA in modifying HPPT so that the performance impact of proposed solutions can be quantified. The search for candidate solutions for improving performance and alleviating bottlenecks looks first at changes to adaptation data, allocation of functions and devices, and operational procedures before looking at hardware upgrades. A ranked set of candidate solutions is forwarded to the CMRA.

The SPPSA also receives models whose workload data bases have been updated to reflect revisions to the SLAD. The SPPA determines (perhaps by running the model) whether the magnitude of the changes warrants rerunning the model to produce a new set of performance predictions. If a new set of predictions is deemed to be necessary, then the procedure defined above is followed.

The Site Analysis function of the SPPSAA is responsible for identifying the cause of the discrepancies between the expected performance (as predicted by HPPT) and the site measurement data. This activity receives the normalized percent differences that have been forwarded by the PMA and performs a detailed analysis of the hierarchical HPMT and HPPT performance reports to determine the reasons for the discrepancies. Potential candidates for explaining the differences in the modeling and measurement results include differences in the HCS workload scenario defined in the SLAD versus the site workload scenarios, differences in site adaptation (e.g., amount of memory allocated to the dynamic buffer), differences in the modeled versus measured operational procedures (e.g., writing data files to disk versus tape), and modeling inaccuracies or anomalies (e.g., incorrect PE service times).

The ramifications of the observed differences are then assessed to determine if a modification to the model, site operations, or site procedures is warranted. The changed model is then run to determine the long-term impact of the site's operation or procedures on performance. If the analysis determines that some aspect of the site's adaptation, operation, or procedures will significantly degrade long-term performance, then the degree of degradation and recommendations for change are forwarded to the CMRA. Those differences in the model versus measured results that are found to

have little or no impact on HPPT's performance predictions are documented and considered for incorporation into the model during the next calibration cycle.

Once changes to the model to resolve discrepancies have been identified, forwarded to the PMA, and the (re)validated model received from the PMA; the SPPSA produces a revised set of performance predictions using the updated model. As before, the problems are identified and solutions are searched and evaluated. The set of predictions, problems, proposed solutions and their predicted performance are forwarded to Capacity Management Reporting Activity.

8. CAPACITY MANAGEMENT PLANNING ACTIVITY.

The Capacity Management Reporting Activity is the primary source for reporting the results obtained from the other activities and therefore responsible for the formal reporting of predicted site performance, potential operational problems or shortfalls in performance, and recommended solutions to performance shortfalls with their predicted performance impacts (results produced by the SPPSAA).

In addition to the performance predictions reported by the CMRA, this activity also reports measured workload and performance trends from the data produced by the Workload Measurement and Performance Measurement Activities. The measurement reports produced by CMRA will highlight abnormal performance conditions such as excessive response times or utilization spikes and will identify the causes of these abnormalities (e.g., excessive queuing delay due to "batched" arrival of messages, utilization surges due to data being dumped from disk to tape).

The primary output from this activity is a Site Capacity Plan which includes predictions of site workloads and performance over the planning horizon and includes any planned or recommended time-phased introduction of the hardware or software upgrades to the site.

9. CAPACITY MANAGEMENT OVERVIEW

The Capacity Management Process is composed of three integrated phases as shown in Figure 9-1. The same general procedure is followed in each of the phases, and each phase can result in an updated Site Capacity Plan. The primary differences in the phases are the source and accuracy of the data that are input to the performance model to produce the Site Capacity Plans.

The objective of Phase 1, Site Forecast, is to predict the performance impact of planned software/hardware upgrades prior to their development, test, and installation at the FAATC. During this phase, various upgrade alternatives can be evaluated to determine their long-term capacity impacts. HPPT is used to evaluate the upgrade

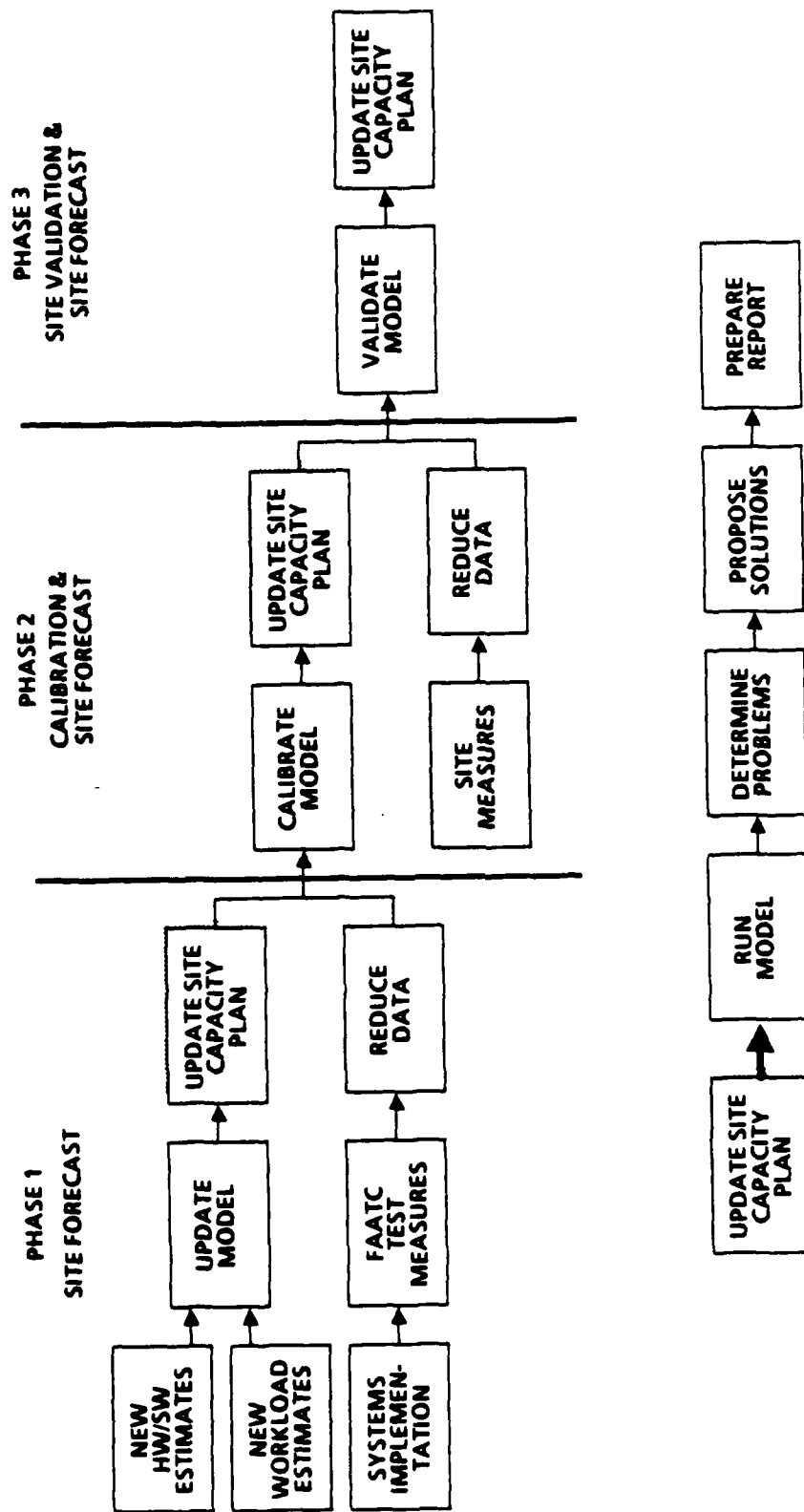


FIGURE 9-1. TIME-ORDERED CAPACITY MANAGEMENT STEPS

alternatives based on engineering estimates for the amount of service requested by the new/modified program elements and vendor claims or benchmark results for the performance characteristics of the new hardware. HPPT is updated to reflect the new/modified hardware and software. The SLAD workloads forecast for the time frame of the system upgrade are then modeled in order to predict the future performance of the HCS. The purpose of this step is to determine well in advance if any performance shortfalls will occur, and to propose alternatives that will alleviate the shortfalls. The standard procedure for producing site capacity plans is followed: performance shortfalls are identified, capacity upgrades are modeled, and recommendations are proposed.

The objective of Phase 2, Calibration and Site Forecast, is to predict the performance impact of planned software/hardware upgrades prior to their test and installation at the sites. Engineering estimates for amount of service requested by the PEs and vendor claims for hardware performance are replaced in HPPT by measurement data obtained from the FAATC test system while processing HLT workloads. HPPT's performance predictions are then compared to HPMT's measurement reports, and discrepancies are resolved. The model calibration process is explained in much greater detail in the HPPT Model documentation. Once the model is calibrated, the procedure for updating site capacity plans may be invoked if the calibrated model differs significantly from that used during Phase 1.

The objective of Phase 3, Validation and Site Forecast, is to resolve any discrepancies between expected and measured performance, that is, to validate the performance predictions made by HPPT. Estimates of site workloads and operational procedures and settings are replaced in HPPT by measurement data obtained from the sites. HPPT's performance predictions are then compared to HPMT's measurement reports, and discrepancies are resolved. The model validation process is explained in much greater detail in the HPPT Model documentation. Once the model is validated, the procedure for updating site capacity plans may be invoked if the validated model differs significantly from that used during Phase 2.

9.1 CAPACITY MANAGEMENT PROCEDURES. This section defines the steps to be performed to implement the various phases of the Capacity Management Process and their associated HCS Capacity Management Activities. The validity of the Site Capacity Plans is strongly dependent on the data used by HPPT. During Phase 1, only estimates of service times are used, and it is not until Phase 2 that measured times are available. Consequently, although major capacity shortfalls can be identified during Phase 1, it is not until the model has been calibrated that highly accurate capacity plans can be generated.

Since Site Forecast is really a subphase of each of the three major phases, and Calibration is covered in considerable detail in the HPPT model documentation, only Phase 3 (Validation and Site Forecast) is described by the following concise series of steps.

Data Reductions:

1. The data outlined in site bulletin is captured by site
2. The site sends several cassette tapes containing HRT data to FAATC
3. FAATC ACT 130 logs cassette tapes into library
4. Using the SURP tool on the HOST Support Processor find peak steady state time during which HRT recording was on and other operational settings were constant
5. Convert HRT cassette tapes to round tapes at 1600 Bpi using FAATC 3083 job shop for the time determined in step 4. In the future cassette tape will be used directly on the Host Support Processor
6. Based on the performance measures (e.g., means, percentiles, class or pair response times) to be reported, determine the length of the period to be analyzed by HPMT
7. Run HPMT to scan the entire tape and report on gaps in the data. Find a gap free period of the appropriate length that falls within the steady state interval identified in Step 4. If no such period exists, then find a gap free period in another steady state (but not peak) interval. If a gap free period still cannot be found, rely on HPMT back-up data sources.
8. Using HPMT tool select the time determined in step 4 and generate the reports
9. Run REDUC on the same tape and for the same time frame generating the PE Statistics Summary Report.

Analysis:

1. Determine average track load during HRT measurement period from the SURP minute-by-minute output
2. Establish the site environment during the HRT recording period by examining the HPMT Program Element Service Time Summary Report for the following:
 - a) Determine what disk (if any) HRT is recorded on by investigating PE MRC for a disk service time.
 - b) To determine how and where On-line Recording is routed, check PE MOC as in step 2a.

c) Conflict Alert, DYSIM, MSAW and REMON functions are determined by noting whether the corresponding PEs, CAD, KSM, RVD, MRM have been executed.

3. Examine the REDUC PE Statistics Summary Report and extract the CPU utilization required to perform interrupt processing.
4. Develop the site operational performance envelope by executing the HPPT for two scenarios. First run with HRT-on, TARS-on, MMP-high to establish maximum utilization. Then run with HRT-off, TARS-off, and MMP-low to get the minimum. Each scenario is run up to the year and/or trackload that an anticipated change will be fielded in either hardware or software. At this point the model (HPPT) is changed to reflect the changes in the real system and the two scenarios run up to the year of that change. This is repeated for the expected life of the system. Figures 3.1-1, 3.3-1 and 3.3-2 in the appendix are representative of the output. The area between the two stepwise discontinuous curves represents the expected life cycle operating range for this site.
5. Compute the measured total CPU utilization by adding:
 - a) the total CPU utilization for PEs reported by HPMT Program Element Service Time Summary Report.
 - b) the CPU required to perform interrupt processing as reported by REDUC in step 3.

The number should fall within the operating envelope. For example, the measured performance reported is shown in Figure 3.1-1 in the appendix. If the number falls outside this range, detailed analysis at the PE level is required (See step 8).

6. The measured total disk utilization is reported by the HPMT Program Element Service Time Summary Report. Again this number should fall within the operating envelope for that particular disk drive.
7. Detailed analysis of a sites resources are examined by investigating the basic unit of resource consumption the Program Element (PE). Using the site environment and trackload determined in steps 1 and 2 above, run HPPT to obtain expected utilizations and response times
8. From the HPPT output select those PE which exhibit the highest percentages of CPU usage. Compare these with the highest users determined by HPMT. They should be the same PEs and the percentage differences normalized to the measured utilizations should be less than 10%. Failure to meet either condition requires the analyst to investigate the assumptions pertaining

to workload, functionally and service demands. Figure 3.4-1 in the appendix is an example of this comparison. HPPT model documentation contains information and procedures for reconciling these differences.

9. Disk usage is analyzed by performing the same activities as stated in the previous step. Figure 3.4-2 in the appendix is an example of this output.
10. Controller message response times are measured by HPMT and predicted by HPPT. The response times predicted in step 7 are based on the average PE service times for all threads, but the PE service time in a particular thread can be significantly different from the average. For example, PCE can be used by QZ and DM messages. Because QZ has fewer input parameters to process its service demand is only 5 milliseconds. DM on the other hand requires 10 milliseconds. When the sample size of an individual I/O pair is small, the HPPT projected response time can be different from the HPMT measurement because of this difference. The overall average I/O pairs response times, not the individual I/O pair response time should be within 30% of one another in comparing model and measured response time.

Report Generation:

1. The indicated outputs from the measurement tools and the model generated in the proceeding steps are input to the Apple microcomputer.
2. The graphics developed for the report are generated using LISAGRAPH and LISADRAW. LISADRAW is used to develop performance envelopes as seen in Figure 3.3-1 and 3.3-2 in the appendix. These include the development of shaded areas between two curves for better graphical presentation. LISAGRAPH uses independent variables on the x-axis, dependent variable on the y-axis and generates either box charts or line graphs. Examples of this are shown in Figures 3.4-1 and 3.4-2 in the appendix. Both the LISAGRAPH and the LISADRAW data base need to be developed and maintained in order to provide figures for future Site Capacity Plans.
3. The text portion of the report is written by a capacity manager. The text explains any inconsistency between the expected values and those measured at this time. Also, the manager checks service levels for compliance with requirements and recommends courses of action to alleviate any bottleneck or shortfall.

APPENDIX
SAMPLE SITE CAPACITY PLAN

1.0 OVERVIEW

This is the first in a series of Capacity Management Reports which will be generated periodically for this site. Reports will be generated several months prior to the installation of a new NAS software version or a hardware upgrade to the site. The purpose of the report is to determine if any changes to the hardware/software or operational procedures need to be made in order to process the projected workloads while meeting the performance requirements.

2.0 SITE MEASUREMENT DATA RECORDINGS

The Seattle system build number used for this analysis was SEAH162. This system is the 4e0.0 Host version. The following data tapes were used in the analysis and are available at the FAA Technical Center.

FAATC

<u>VLIB#</u>	<u>DATE</u>	<u>TYPE</u>	<u>START-STOP TIMES</u>
AC8948	6/25/87	HRT	MMP DATA-00:00:00-18:02:00 HRT DATA-17:00:00.5-18:00:00.5
AC8962	6/25/87	SAR1	15:50:00-18:15:00 TARS 7FF
AC8963		SAR2	18:15:00-21:32:00

All of the analysis is confined to data taken from these recording periods.

2.1 HOST PREDICTED PERFORMANCE TOOL (HPPT) PARAMETERS

The HPPT model, calibrated to version 4e0.0 of the software, was run with the following parameters set for all scenarios for which data was generated.

Workload Parameters

Active Flight Plan/Trackload Ratio	1.34
Track Life	39 Minutes
Year	1987

Environment Switches

DYSIM	OFF
Conflict Alert	ON
E-MSAW Alert	ON
Near Term Enhancement	OFF
SARC Level	4
HRT Recording (MRC)	DISK2
On-Line Recording (MOC)	DISK2

In addition, the following variable conditions were used for the four scenarios:

Track Counts between 100-600 track in 50 track increments

Scenario 1 HRT-ON; TARS-ON; MMP-HIGH

Scenario 2 HRT-OFF; TARS-ON; MMP-HIGH

Scenario 3 HRT-OFF; TARS-OFF; MMP-HIGH

Scenario 4 HRT-OFF; TARS-OFF; MMP-LOW

where MMP-HIGH is CPUU, CHNU, STRU, PERU, PEMN, INTR - ON

MMP-LOW is CPUU - ON

3.0 PERFORMANCE ANALYSIS

The analysis conducted on the Seattle site data was from SAR tapes with TARS 7FFF and an HRT tape with MMP data and HRT recordings. The data from time period 17:00:00 to 17:30:00 was selected for detailed analysis since it represented a time period of high traffic for which all available data, both SAR, HRT, and MMP was available. Additional analysis was conducted on data from 00:00:00 through 18:02:00 using the System Utilization Report Program (SURP). The analysis showed that the track count varied for the measurement period (1700-1730) between 179 tracks and 206 tracks with total average track load of 196. Total CPU averaged 12.5% over this period. Over the entire period of observation (0000-1802), track count and CPU utilization varied from a low of 19 tracks and 4% CPU utilization to 209 tracks and 14% CPU utilization.

3.1 ANALYSIS OF CPU UTILIZATION

Figure 3.1-1 presents the Seattle predicted performance envelope. It was generated using the HPPT for track counts of 100 through 600 tracks. Four different scenarios, corresponding to the most likely and extreme field situations were run, each generating a predicted performance curve. The scenarios are:

- a. HRT ON, TARS ON, MMP high*

This is the least likely configuration to be run at an ARTCC since HRT is only used for performance measurements.

- b. HRT OFF, TARS, ON, MMP high*

A possible normal configuration.

- c. HRT OFF, TARS OFF, MMP high*

A possible normal configuration.

SEATTLE 4e0.0 System CPU PREDICTED PERFORMANCE ENVELOPE CPU Verses Track Count
 (SARC 4:C/A, E-MSAW ON; DYSIM, REMON OFF)

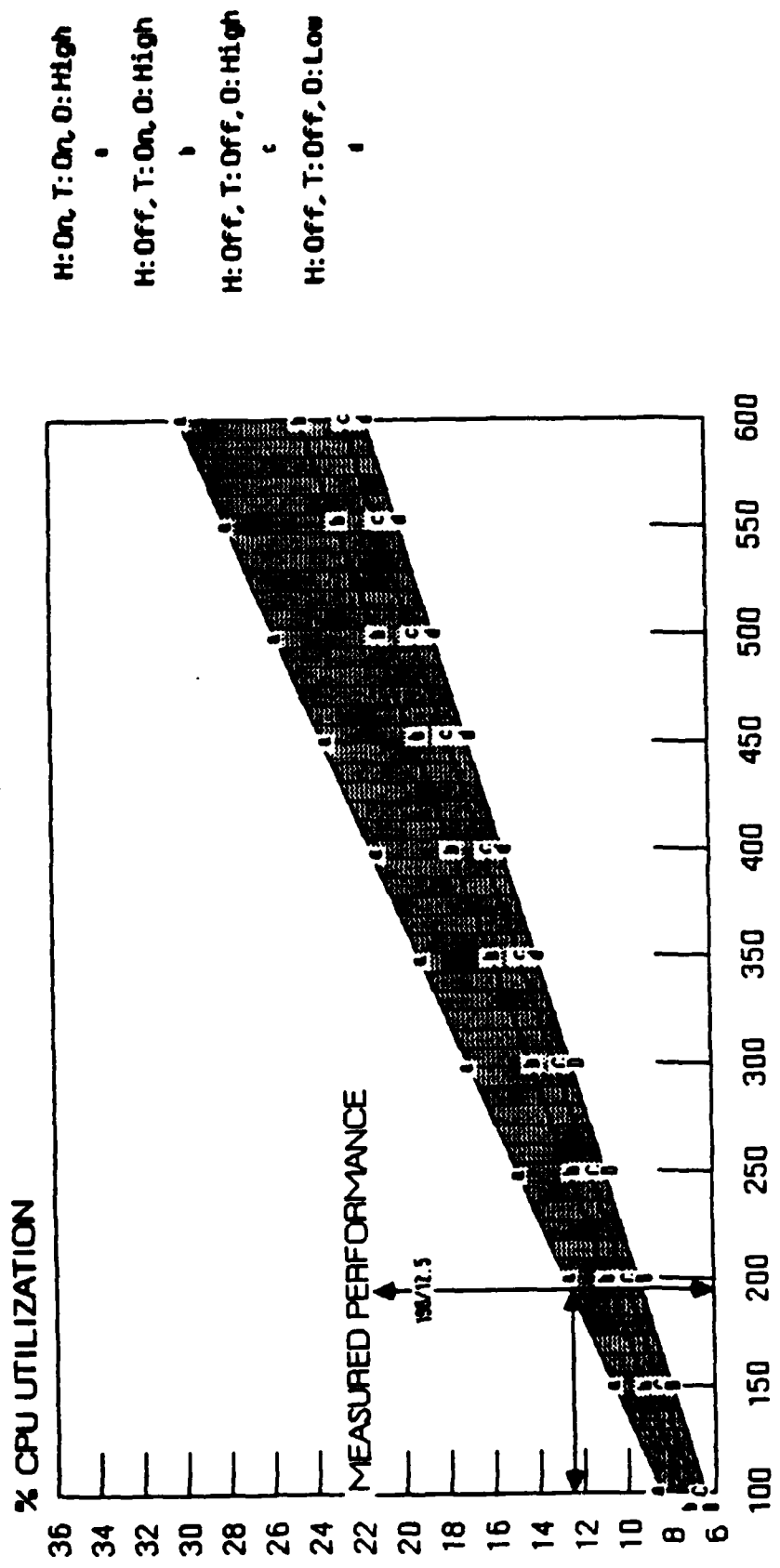


FIGURE 3.1-1. TRACKLOAD (H:HRT, T:TARS, O:Online Monitor (MMP))

d. HRT OFF, TARS FFFF OFF, MMP low*

A most likely normal configuration.

*See Section 2.1 for definitions of MMP low and high.

In all scenarios, E-MSAW and Conflict Alert were on and SARC 4 recording level was in effect. The performances envelope presents the range of CPU utilization for track counts between 100 and 600 tracks. Also, superimposed on Figure 3.1-1 is the highly precise data measured at the site and for which detailed analysis has been performed. The figure clearly shows the system is performing within the expected operating range.

Figure 3.1-2 compares (for equivalent configurations and track counts) the measured Seattle CPU utilization and the expected utilization as predicted by HPPT. It can be seen that the current measured CPU utilization at Seattle is close to the expected utilization.

SEATTLE 4e0.0 System PREDICTED CPU VS ACTUAL CPU PERFORMANCE

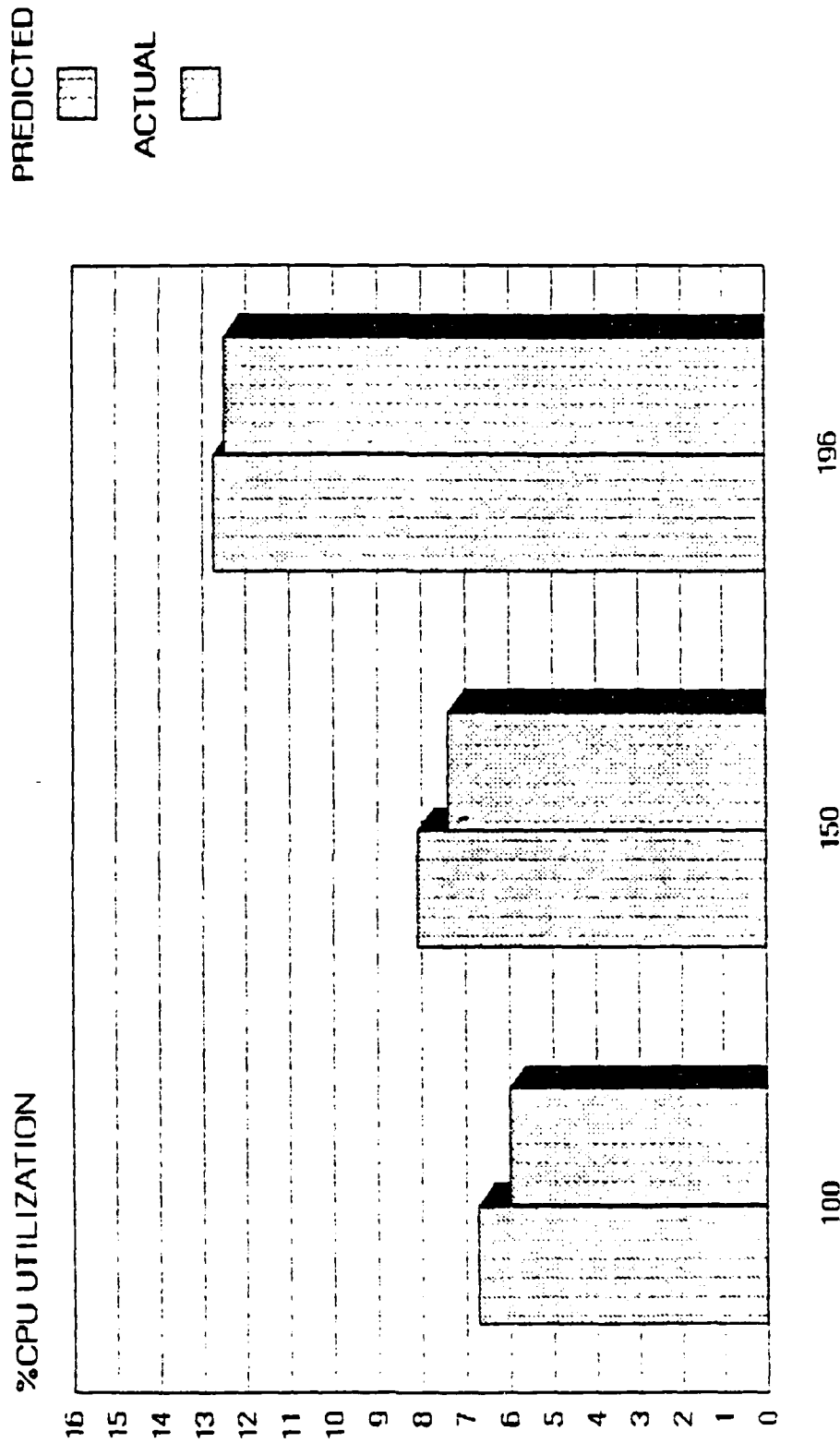


FIGURE 3.1-2. TRACKLOAD (100 & 150) (HRT-OFF, TARS-OFF, MMP-HIGH) TRACKLOAD (196) (HRT-ON, TARS-ON, MMP-HIGH)

3.2 RESPONSE TIME ANALYSIS

Figure 3.2-1 compares, for equivalent configurations and trackload, the measured Seattle mean response times from HPMT and the expected mean response times as predicted by HPPT. The measured values cover the time interval 17:10-17:20 when the average trackload was 202 tracks. The figure shows that actual response times are close to the expected response time performance for this trackload. In addition, the actual response times (on the order of tens of milliseconds) are very small compared to response time performance guidelines which require mean response times on the order of seconds.

Table 3.2-1 presents individual input/output message pair response times for the interval 17:00-17:30 when the average trackload was 196. The table provides additional verification of satisfactory response time performance in this Seattle data sample. All 90th percentile response time measurements were under 0.2 seconds, and only a single maximum response time exceeded 1.0 seconds.

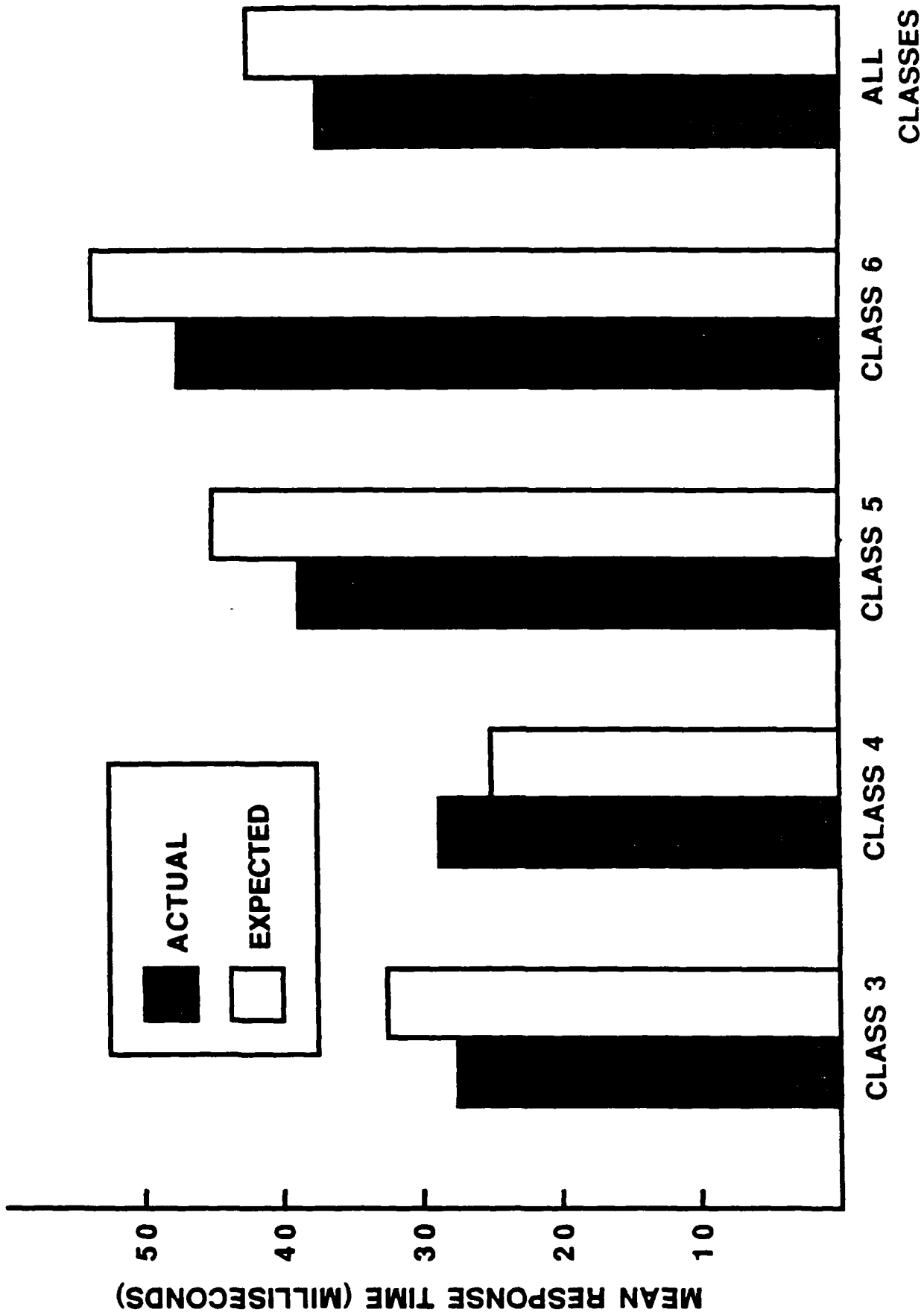


FIGURE 3.2-1. SEATTLE 4e0.0 SYSTEM ACTUAL RESPONSE TIME VS. EXPECTED PERFORMANCE

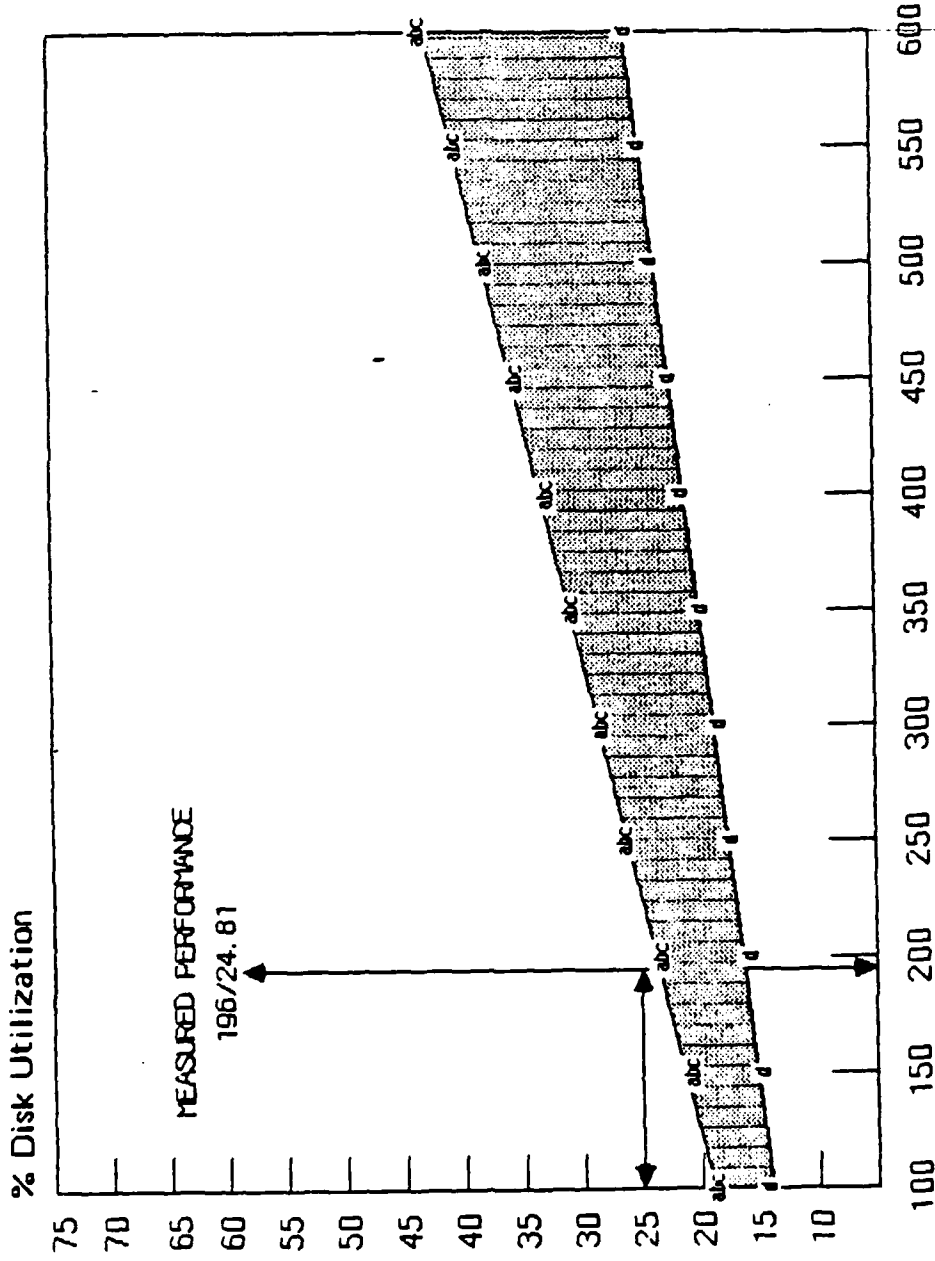
3.3 DISK UTILIZATION

Figures 3.3-1 and 3.3-2 present the expected disk utilization for the system. The same scenario outlined for the CPU was used in generating these figures. The total number of disk I/O accesses on the two disk devices is about the same but the utilization is 24.82% and 53.89%, respectively. Disk 1 is considered to be within the operating envelope but the large discrepancy between measurement and projection for Disk 2 warrants deeper investigation. Since HRT data is recorded on Disk 2 and is indeed a significant amount of data, perhaps the data layout on the disk may be very inefficient.

The first figure shows that measured Disk 1 utilization is within the expected range. However, the second figure shows that Disk 2 utilization is significantly higher than expected. This difference between expected and measured utilization is analyzed in detail in Section 3.4.2.

SEATTLE PREDICTED DISK UTILIZATION

DISK 1 (H=HRT, T=TARS, M=MMP)



H-ON,T-ON,M-HIGH

a

H-OFF,T-ON,M-HIGH

b

H-OFF,T-OFF,M-HIGH

c

H-OFF,T-OFF,M-LOW

d

FIGURE 3.3-1. DISK 1 SEATTLE 4e0.0 SYSTEM PREDICTED PERFORMANCE ENVELOPE

SEATTLE PREDICTED DISK UTILIZATION

DISK 2 (H=HRT, T=TARS, M=MIMP)

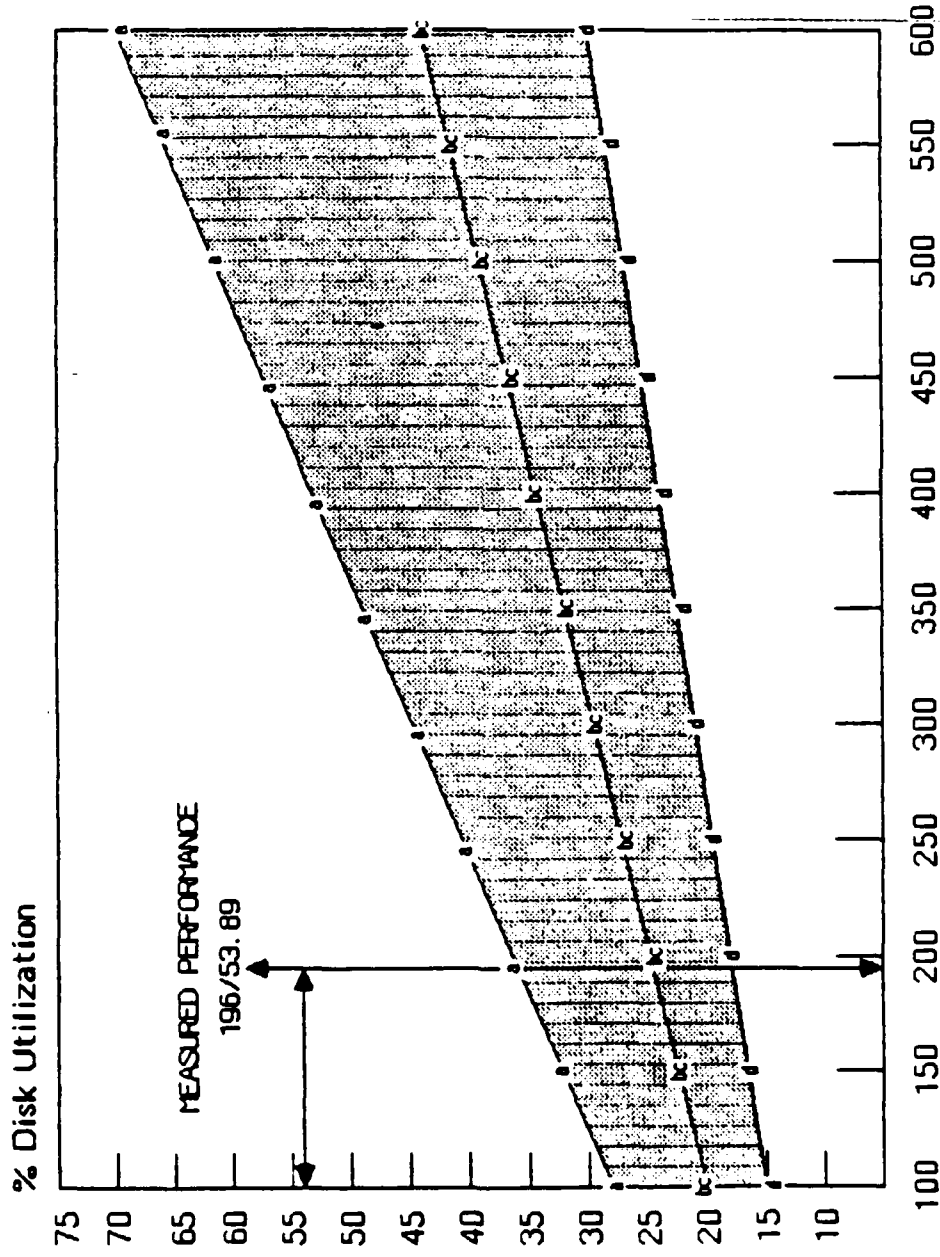


FIGURE 3.3-2. DISK 2 SEATTLE 4e0.0 SYSTEM PREDICTED PERFORMANCE ENVELOPE

3.4 PE ANALYSIS

The basic unit of resource consumption is the Program Element (PE) and is the smallest individually schedulable unit of software in the Host system. The following sections examine the CPU and Disk utilization for the heaviest PE users.

3.4.1 TOP CPU USERS

Figure 3.4-1 shows the absolute percentage utilization that each PE contributes to the total CPU utilization. The measured and expected values are within the tolerances as one would expect having viewed Figure 3.1-1. The behavior of the CPU is normal and is what was anticipated for this workload.

ENVIRONMENT: 200 Tracks HTR=on, IARS=on, HMP=high

SEATTLE ARTCC

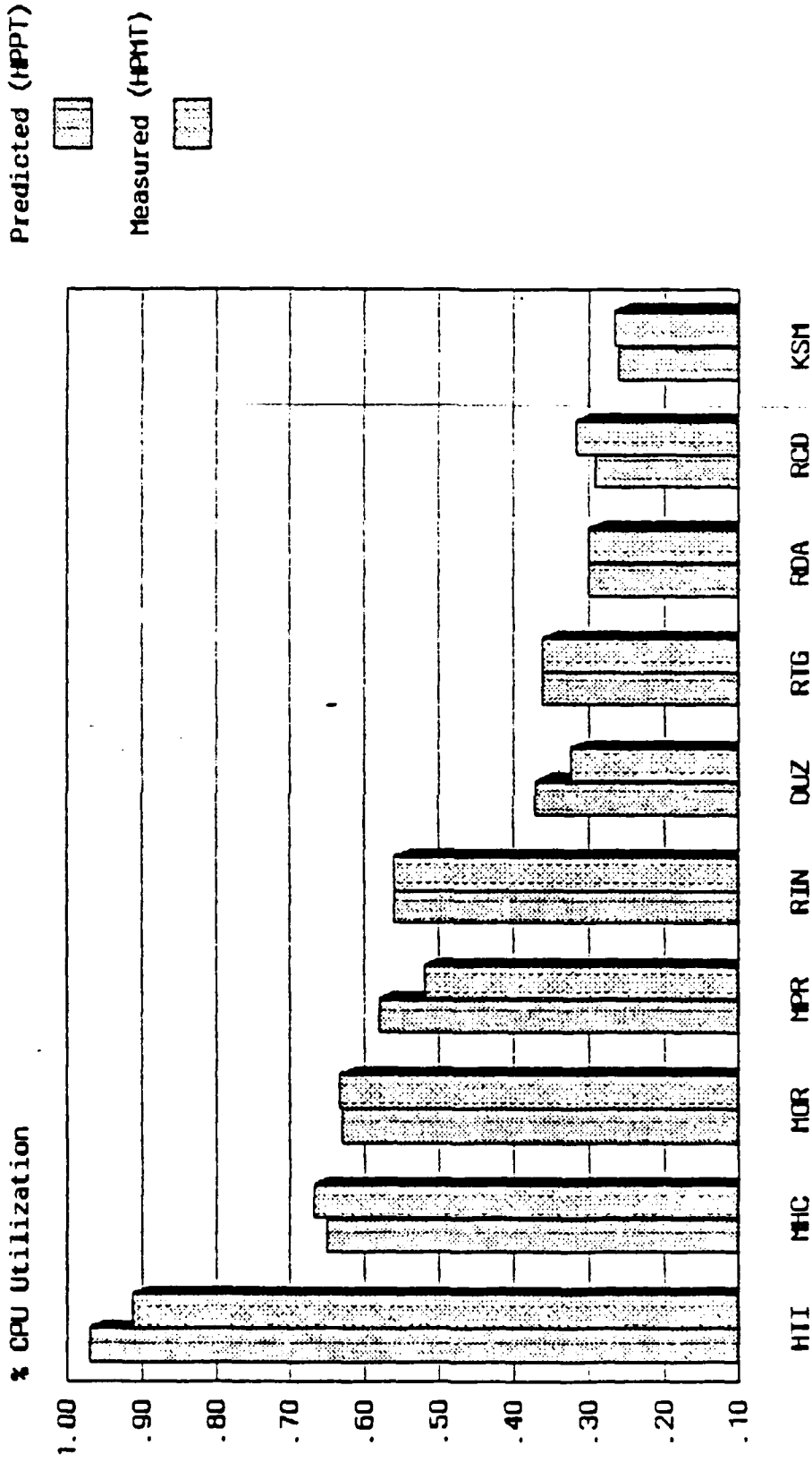


FIGURE 3.4-1. A COMPARISON OF THE TEN LARGEST USERS OF CPU AS MEASURED BY HPMT AND PREDICTED BY HPPT

3.4.2 TOP DISK USERS

Disk I/O activity on a PE basis is displayed in Figure 3.4-2. The reason for the large discrepancy pointed out in Figure 3.3-2 is now obvious. PE MHC and MPR are out of line especially with regards to Disk 2. Detailed analysis of MHC using the HPMT tool has revealed that the measured service demand of Disk 2 is more than three times larger than on Disk 1 even though the total number of disk I/O accesses on each of the two disks are about the same. On the other hand, the service demand for MPR seems reasonable but the disk access rate differs significantly. Some further investigation is needed to understand the causes of these differences.

SEATTLE ARTDC

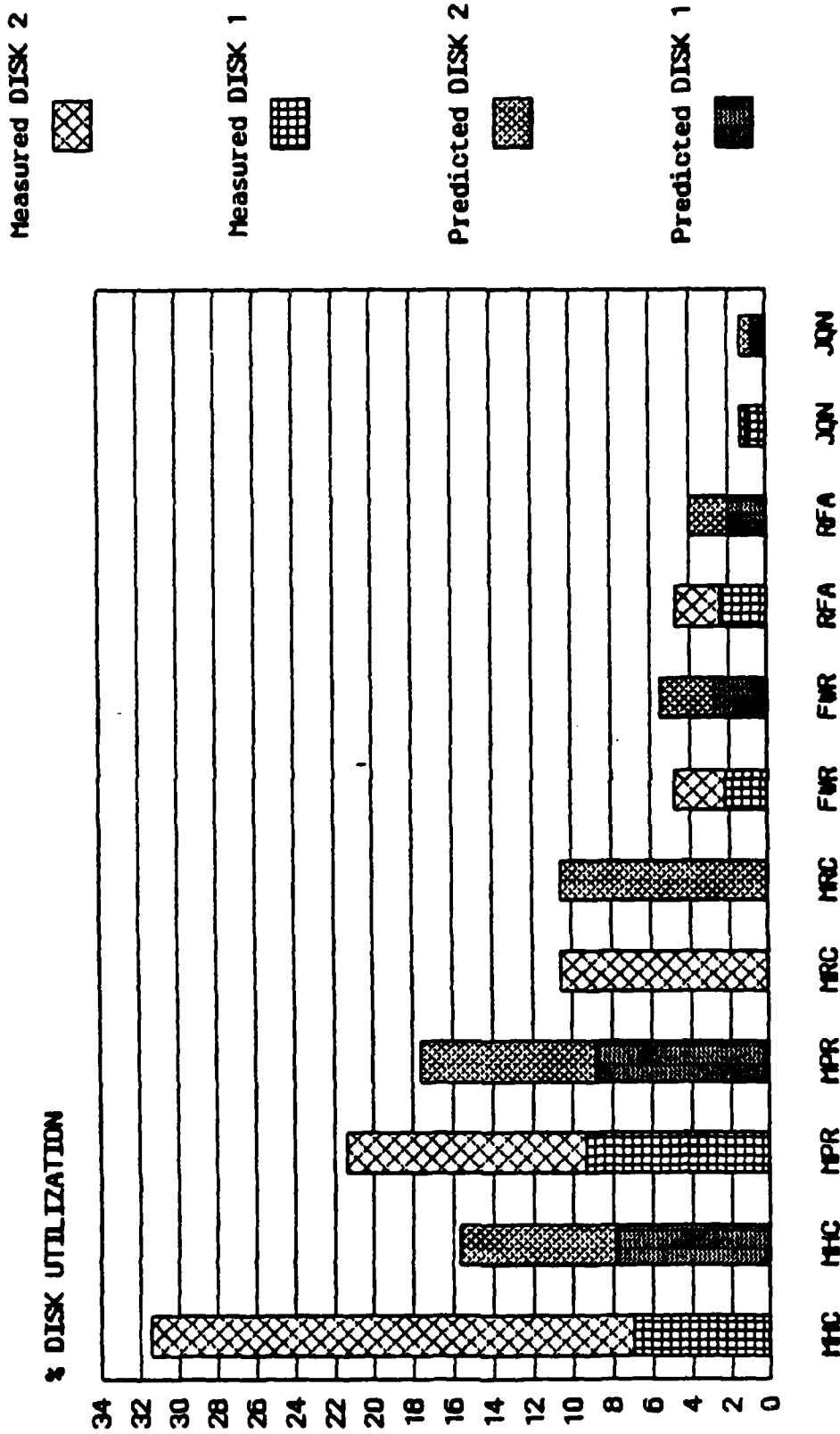


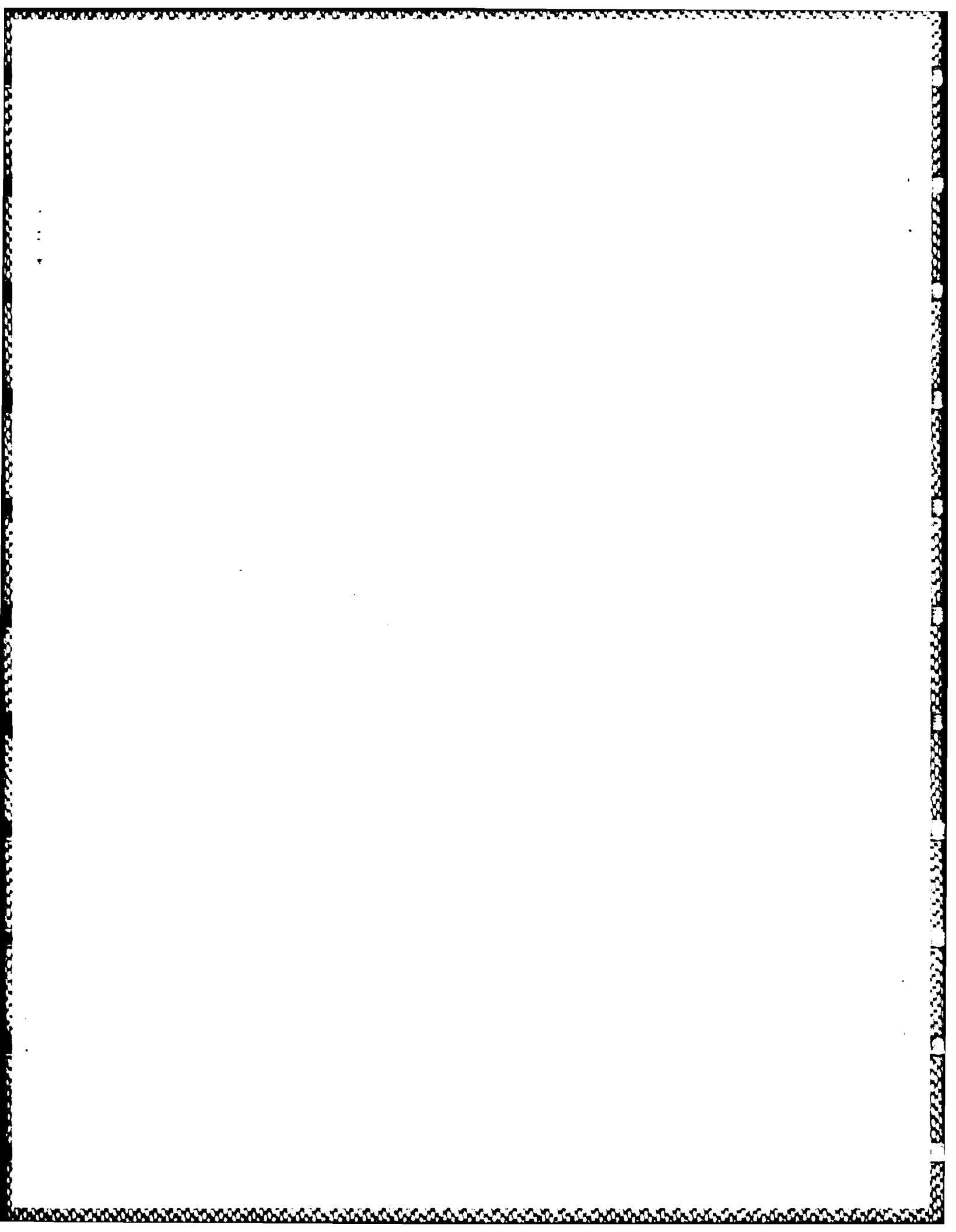
FIGURE 3.4-2. PE'S WHICH ARE THE BIGGEST USERS OF DISKS 1 & 2

4.0 CONCLUSIONS AND RECOMMENDATIONS

The data and analysis presented in this report show that the performance of the Seattle HCS under operational conditions is generally within the expected range. One exception is an unexpectedly large disk utilization. Even with this higher than expected utilization, response time performance is well within the nominal range.

The disks are more highly utilized than the processor and are therefore more likely to become the bottleneck resource if resource demands continue to increase proportionally to the number of tracks. The system's monitor and resource monitoring are by far the major disk users. Eventually this high disk utilization will have an effect on response time.

It is recommended that the site's data systems specialist carefully examine the file structure on Disk 2 and further investigate the PE MHC. Also considerable thought should be given to recording resource monitoring data to tape which will significantly lower disk utilization.



END

DATE

FILMED

DTIC

JULY 88