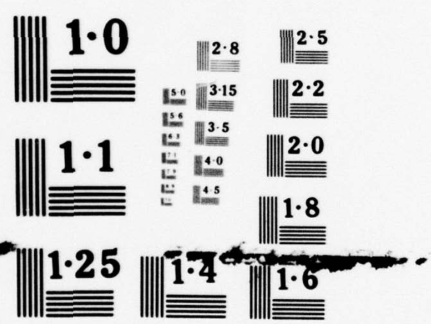
AD-A059 937 COMMAND AND CONTROL TECHNICAL CENTER WASHINGTON D C F/G 9/2
WORLDWIDE MILITARY COMMAND AND CONTROL SYSTEM (WWMCCS) H-6000 T--ETC(U)
FEB 78 B M WALLACK, G H GERO

UNCLASSIFIED CCTC-TM-180-78-VOL-1 SBIE-AD-E100 094 NL

1 OF 1
ADA
059937

END
DATE
FILMED
12-78
DDC



NATIONAL BUREAU OF STANDARDS
MICROCOPY RESOLUTION TEST CHART

AD-E100 094

(19)

ADA059937

DDC FILE COPY

CCTC

(18) SBIE

(12) 54p.

(9) TECHNICAL MEMORANDUM

(14) CCTC - TM-180-78 - VOL-1

VOLUME I

3 FEBRUARY 1978

(11)

(1)

LEVEL II

COMMAND & CONTROL TECHNICAL CENTER

DDC RECEIVED OCT 17 1978 B

(6)

WORLDWIDE MILITARY COMMAND AND CONTROL SYSTEM (WWMCCS)

H-6000 TUNING GUIDE,

VOLUME I. WWMCCS SYSTEM TUNING PROCESS.

(10) Barry M. /Wallace

George H. /Gero

DEFENSE COMMUNICATIONS AGENCY

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

AD-E100 094

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER CCTC-TM-180-78-VOL-1	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) Worldwide Military Command and Control System (WWMCCS). H-6000 Tuning Guide. Volume 1 - WWMCCS System Tuning Process.		5. TYPE OF REPORT & PERIOD COVERED Technical memorandum 6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) Barry M. Wallack George H. Gero		8. CONTRACT OR GRANT NUMBER(s)
9. PERFORMING ORGANIZATION NAME AND ADDRESS Command and Control Technical Center Attn: CCTC/CPE/C702 Washington, DC 20301		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS
11. CONTROLLING OFFICE NAME AND ADDRESS Same as above		12. REPORT DATE 3 February 1978 13. NUMBER OF PAGES 50
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office)		15. SECURITY CLASS. (of this report) Unclassified 15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Computers, Test and evlauation, Performance tests, Time sharing

Honeywell 6000 computers, Computer systems, Federal Computer Performance and Evaluation and Simulation Center

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

The Federal Computer Performance Evaluation and Simulation Center (FEDSIM) has developed a document for WWMCCS installations that can be used by site personnel to analyze the performance characteristics of their Honeywell 6000 (H-6000) computer systems. This document, called an H-6000 Tuning Guide, incorporates detailed analyses procedures that guide the analyst in applying specific techniques to improve system performance.

This volume describes the overall structure and application of the Tuning Guide.

SECURITY CLASSIFICATION OF THIS PAGE(When Data Entered)

Tt explains the approach, procedures, and process taken by the Tuning Guide to p
provide analyses of batch job turnaround time and GCOS Time Sharing System (TSS)
response time.

ACCESSION for		
NTIS	White Section	☑
DDC	Buff Section	☐
UNANNOUNCED		☐
JUSTIFICATION		
BY		
DISTRIBUTION/AVAILABILITY CODES		
Dist.	AVAIL and/or SPECIAL	
A		

SECURITY CLASSIFICATION OF THIS PAGE(When Data Entered)

COMMAND AND CONTROL TECHNICAL CENTER

Technical Memorandum TM 180-78

3 February 1978

H-6000 TUNING GUIDE

WWMCCS SYSTEM TUNING PROCESS

VOLUME I

PROJECT PERSONNEL

Barry M. Wallack
George H. Gero JR.
CCTC/CPE/C702
RM BE685, The Pentagon
Washington, D.C. 20301
AV - 692-2725
COMM - 6972725

APPROVED BY:

TIPTON P. MOTT-SMITH
Colonel, USAF
Deputy Director for
Computer Services

PREFACE

This report is based on detailed analysis of a large amount of technical information. The results address procedures for the analysis of batch turnaround time and GCOS Time Sharing System response time in World Wide Military Command and Control Systems (WWMCCS). Because of the complexity of the analysis procedures and their dependence on the WWMCCS workloads and operational environments, generalizing the procedures beyond the environment described or extracting conclusions without their respective qualifying conditions is not possible. Questions related to this report or to the possibility of extending the stated conclusions or recommendations should be addressed to the Computer Performance Evaluation Office, (C702), the Pentagon.

To gain a general understanding of the approach of the H-6000 Tuning Guide, Volume I Section II, Volume II Section II, and Volume III Section II should be read. One or more of the hypothesis tests (search procedures) in Volume II Sections IV-XII and Volume III Sections III-X should also be read. Not all these tests have to be read at the start of a tuning effort. Each should be read as it needs to be applied. To start a tuning effort, Volume I should be read and applied. The procedure for analysis of batch turnaround time begins in Volume II Section III. The procedure for analysis of Time Sharing response time begins in Volume III Section II.

The H-6000 Tuning Guide has never been tested by a novice in performance evaluation, although field tests have been conducted by FEDSIM personnel. For this reason, it remains a preliminary version.

TABLE OF CONTENTS

TABLE OF CONTENTS (Cont'd)

TABLES

FIGURES

ABSTRACT

The Federal Computer Performance Evaluation and Simulation Center (FEDSIM) has developed a document for WWMCCS installations that can be used by site personnel to analyze the performance characteristics of their Honeywell 6000 (H-6000) computer systems. This document, called an H-6000 Tuning Guide, incorporates detailed analysis procedures that guide the analyst in applying specific techniques to improve system performance.

The four volumes of the Tuning Guide present a precisely structured system of procedures for the analysis of the performance of WWMCCS computer services and systems:

Volume I — WWMCCS System Tuning Process. The first volume describes the overall structure and application of the Tuning Guide. It explains the approach, procedures, and processes taken by the Tuning Guide to provide analyses of batch job turnaround time and GCOS Time Sharing System (TSS) response time.

Volume II — Batch Turnaround Time Analysis Procedures. The second volume presents a set of procedures for analysis of batch job turnaround time. It first presents a model of the processes and queue points associated with batch job turnaround time and then describes nine tests that use the model to direct the analysis of turnaround time.

Volume III — TSS Response Time Analysis Procedures. The third volume serves the same general purpose and has the same general structure as Volume II. Volume III presents a complete set of procedures for investigating the response time of GCOS Time Sharing System (TSS) interactions. The volume first presents a model of the processes and queue points associated with TSS response time and then describes eight tests to direct an analysis of TSS response time.

Volume IV — H-6000 Tuning Guide Appendices. The fourth volume provides the appendices referenced by the other volumes of the Tuning Guide. The volume contains detailed descriptions of report formats and other referenced data.

ABSTRACT

The Federal Computer Performance Evaluation and Simulation Center (FEDSIM) has developed a document for WWMCCS installations that can be used by site personnel to analyze the performance characteristics of their Honeywell 6000 (H-6000) computer systems. This document, called an H-6000 Tuning Guide, incorporates detailed analysis procedures that guide the analyst in applying specific techniques to improve system performance.

The four volumes of the Tuning Guide present a precisely structured system of procedures for the analysis of the performance of WWMCCS computer services and systems:

Volume I — WWMCCS System Tuning Process. The first volume describes the overall structure and application of the Tuning Guide. It explains the approach, procedures, and processes taken by the Tuning Guide to provide analyses of batch job turnaround time and GCOS Time Sharing System (TSS) response time.

Volume II — Batch Turnaround Time Analysis Procedures. The second volume presents a set of procedures for analysis of batch job turnaround time. It first presents a model of the processes and queue points associated with batch job turnaround time and then describes nine tests that use the model to direct the analysis of turnaround time.

Volume III — TSS Response Time Analysis Procedures. The third volume serves the same general purpose and has the same general structure as Volume II. Volume III presents a complete set of procedures for investigating the response time of GCOS Time Sharing System (TSS) interactions. The volume first presents a model of the processes and queue points associated with TSS response time and then describes eight tests to direct an analysis of TSS response time.

Volume IV — H-6000 Tuning Guide Appendices. The fourth volume provides the appendices referenced by the other volumes of the Tuning Guide. The volume contains detailed descriptions of report formats and other referenced data.

I. INTRODUCTION

A. BACKGROUND

The Office of the Joint Chiefs of Staff (JCS) has directed that the Command and Control Technical Center (CCTC) develop a computer performance analysis capability to support the World Wide Military Command and Control System (WWMCCS).

CCTC, acting at the direction of the JCS, has specified that WWMCCS ADP managers are to apply various computer performance evaluation (CPE) tools and techniques to the systems now running at their sites. CCTC has also defined the need to instruct WWMCCS technical personnel in the selection and application of the CPE tools and techniques appropriate to individual WWMCCS ADP sites.

CCTC asked FEDSIM to plan and implement a document that could be employed by WWMCCS ADP personnel to diagnose problems and propose changes that would improve the performance of WWMCCS ADP systems.

B. PROJECT OBJECTIVE

The objective of the resulting FEDSIM project was to provide all WWMCCS installations with a document that could be used by staff personnel to analyze the performance characteristics of their ADP systems. This document, called an H-6000 Tuning Guide, was to contain sets of analysis procedures to improve system performance.

The product of the completed FEDSIM project is a four-volume H-6000 Tuning Guide (referred to hereafter as the Guide). The Guide volumes present a precisely structured system of procedures for the analysis of WWMCCS computer services and systems. The titles of the four volumes are: (1) WWMCCS System Tuning Process, (2) Batch Turnaround Time Analysis Procedures, (3) TSS Response Time Analysis Procedures, and (4) H-6000 Tuning Guide Appendices.

C. PURPOSE AND CONTENTS OF THE GUIDE

Computer jobs may be submitted by WWMCCS users to run as either batch jobs or as interactive jobs. Batch jobs, as processed by the WWMCCS systems, may be submitted by users at the site or may be initiated via a process called "job spawning" through the WWMCCS Time Sharing System. Interactive jobs addressed by the Guide are the subsystems that run under control of the WWMCCS Time Sharing System. The performance of both batch and interactive jobs can be measured and analyzed with reference to the amount of elapsed time that the system takes to process them. Batch job elapsed processing time is called batch turnaround time. Interactive job elapsed processing time is called TSS response time.

1. Volume I: WWMCCS System Tuning Process

The first volume describes the overall structure and application of the Guide. Volume I explains the approach, procedures, and processes taken by the Guide to provide analyses of batch turnaround time and TSS response time.

Volume I is organized into four sections. Section I (this section) is an introduction. Section II (COMPUTER SYSTEM PERFORMANCE TUNING) introduces system tuning concepts. Section III (PROBLEM DEFINITION PHASE) discusses pre-analysis and problem definition activities required of the Guide user. Section IV (PROBLEM ANALYSIS PHASE) provides general directions in the use of the other volumes.

2. Volume II: Batch Turnaround Time Analysis Procedures

Volume II presents a set of procedures for analysis of batch turnaround time. The volume first presents a model of the processes and queue points associated with batch job turnaround time and then describes nine tests that use the model to direct the analysis of turnaround time. Both the model and the tests are designed to be applied in the WWMCCS system environment.

Table I-1 lists the sections of Volume II and briefly identifies how each is used to conduct an analysis of batch job turnaround time.

3. Volume III: TSS Response Time Analysis Procedures

Volume III serves the same general purpose and has the same general structure as Volume II. Volume III presents a complete set of procedures for investigating the response time of TSS interactions. The volume first presents a model

SECTION NUMBER	TITLE	APPLICATION OF SECTION
I	Introduction	
II	Turnaround Time Model	1. Define Potential Delay Points 2. Hypothesize Causes for Delay
III	Turnaround Time Model Scan	1. Direct Analysis of Delay Points 2. Initiate Other Tests
IV	Seek Elongation Test	1. Confirm Disk Seek Delay 2. Identify Files to Move
V	Memory Constraint Test	1. Confirm High Memory Wait Time 2. Identify Source of Wait
VI	Device Errors Test	1. Determine Tape Handlers in Error 2. Determine Disk Units in Error
VII	Pathway Utilization Test	1. Validate I/O Service Queuing Delay 2. Isolate Libraries and Files to Move
VIII	Insufficient Device Test	1. Confirm Tape-Caused Delay 2. Confirm Disk Space-Caused Delay
IX	"Few Activities In System" Test	1. Substantiate Light Job Scheduling 2. Identify Scheduling Alternatives
X	IOS Delays Test	1. Confirm GCOS IOS Code As Delay 2. Propose Alternative Solutions
XI	Urgency Codes Test	1. Confirm Delay Caused by Urgency Codes 2. Determine if Initiated by Operator

CONTENTS OF TUNING GUIDE VOLUME II

TABLE I-1

of the processes and queue points associated with TSS response time and then describes eight tests to direct an analysis of TSS response time. Both the model and the tests are designed to be applied in the WWMCCS system environment.

Table I-2 lists the sections of Volume III and briefly identifies how each is to be used to conduct the analysis.

4. Volume IV: H-6000 Tuning Guide Appendices

Volume IV provides the appendices referenced by the other volumes. It contains detailed descriptions of report formats and other referenced data.

Table I-3 lists the sections of Volume IV and briefly identifies their contents.

SECTION NUMBER	TITLE	APPLICATION OF SECTION
I.	Introduction	
II.	Search Procedures	1. Define TSS Response Time Delay Points 2. Hypothesize Causes For Delay
III.	Subsystem CPU Time	1. Determine Subsystems With Unacceptable Response Time 2. Determine Dominant Subsystem CPU Time Category
IV.	TSS Executive CPU Time Search Procedure	1. Adjust Subsystem Dispatch Length 2. Conduct Wait Time Period Experiment 3. Conduct Cutoff Limits Experiment
V.	TSS Wait For CPU Time Search Procedure	1. Investigate TSS Priority 2. Examine TSS Dispatch Interval Length 3. Study Priority B Dispatch Blocking 4. Investigate I/O Queue Space 5. Examine Line Service Interval 6. Study SSA Module Usage
VI.	Disk I/O Time Search Procedure	1. Determine If I/O Queuing Is High 2. Investigate Relocation of TSS Work Files 3. Study Increasing I/O Queue Space
VII.	Memory Wait Time Search Procedure	1. Determine If Swapping Problem Exists 2. Investigate Increase Of TSS Core Size 3. Study Memory Priority Adjustment
VIII.	GWAKE Wait Time Search Procedure	1. Isolate Responsible Subsystems 2. Determine Reason for GWAKE
IX.	Output Wait Time Search Procedure	1. Investigate Memory Availability and User Errors 2. Study Use of I/O Queue Space 3. Investigate Possible DATANET Delays 4. Analyze Output Volume
X.	Non-TSS Service Wait Time Search Procedure	1. Determine Memory Availability 2. Study Urgency Codes 3. Examine I/O Pathway Utilization 4. Investigate CPU Execution Characteristics

CONTENTS OF TUNING GUIDE VOLUME III

TABLE I-2

APPENDIX	CONTENTS
A. Introduction	
B. Turnaround Time Analysis System	1. Concepts And Facilities 2. Defining Site Operating Processes 3. Report Formats And Data Elements
C. Mass Store Monitor	1. Concepts And Facilities 2. Report Formats And Data Elements
D. GCOS Reports	1. Startup Deck Listing 2. Startup File Map 3. SAVEMAST Listing
E. Memory Utilization Monitor	1. Concepts And Facilities 2. Report Formats And Data Elements
F. HEALS II	1. Concepts And Facilities 2. Report Formats And Data Elements
G. GESEP	1. Concepts And Facilities 2. Report Formats And Data Elements
H. Response Time Analysis	1. Concepts And Facilities 2. Report Formats And Data Elements

CONTENTS OF TUNING GUIDE VOLUME IV

TABLE I-3

II. COMPUTER SYSTEM PERFORMANCE TUNING

This section summarizes the process of computer system performance evaluation as implemented in the H-6000 Tuning Guide.

A. GENERAL GUIDE DEFINITIONS

1. Computer System Performance Evaluation

Computer system performance evaluation is a generic term applied to many techniques for determining the performance characteristics of both a computer system and its associated site processing activities. The performance characteristics may be compared to many criteria, including: (1) standards of economic operation, (2) technical norms, or (3) measures of service provided. The H-6000 Tuning Guide, by providing a structured set of analysis procedures and tuning steps, applies certain computer performance evaluation techniques to the computer system tuning process. This process aims to increase the technical capacity of the computer system to provide services to its users.

2. Computer System Performance Variables

In one way or another, the performance of a computer system is influenced by nearly every facet of the data processing function. The following examples illustrate the scope of the computer performance tuning process.

a. System Design. Computer application system design and development can be the starting point for performance degradation. Errors in original design with respect to I/O media selection, file structures, frequency of run, etc. may result in less than optimal performance for as long as an application is in existence.

b. Programming. A programmer's proficiency and the availability of program optimization tools, for example, will influence program design and coding, and affect system performance.

c. Hardware Configuration. Specific components of a computer system may be mismatched to the system as a whole, causing major subsystems (or the entire configuration) to operate at a reduced performance level. Even if the performance capabilities of the individual subsystems are reasonably well matched, the system may be poorly configured for the site's workload, resulting in poor performance.

d. System Software. The software supplied by the mainframe vendor may be inappropriately parameterized to fit the site workload, or may be a source of high overhead or bottlenecks to efficient workload processing.

e. Operations. An operations staff schedules the workload, provides job assembly (and library) services, and operates the system through the console. All of these functions are vital to the proper operation of the system. Mistakes, insufficient training, poor documentation, and a variety of other reasons may contribute to operational problems which substantially decrease system performance.

f. Communications Hardware and Software. A communications network, its interface to a central system, and the software used to control the on-line applications may have a significant impact on the system's overall performance.

The decisions associated with any or all of the above data processing functions can be considered variables that affect computer system performance.

3. Computer System Performance Tuning

The process of analyzing and appropriately adjusting computer system performance variables is known as computer system performance tuning.

4. Computer System Service Analysis Definition

a. WWMCCS Services. Within the context of the Guide, WWMCCS system services are: (1) batch job services and (2) GCOS Time Sharing System services. The service time for batch workload service is called batch turnaround time and the service time for TSS service is called TSS response time.

b. Turnaround Time. This is the total elapsed time taken by a job (or set of jobs) submitted to a WWMCCS site for batch processing. Batch turnaround time is comprised of computer system processing and physical input and output handling in the machine room, both before and after system processing. A job's turnaround time therefore includes all processing and waiting points through which the job must pass from submission until return to a user. The Guides's structured batch turnaround time analysis examines these processes and waiting points.

c. TSS Response Time. A time-related structure similar to batch processing turnaround time can be conceived for TSS service. The Guide's TSS response time analysis defines TSS Response Time as the elapsed time between an on-line users' request for service from the system and the system's request for further input from the user. The TSS analysis procedures of the Tuning Guide divide response time into processes and waiting points associated with: (1) CPU service, (2) disk I/O, (3) memory, (4) terminal I/O, and (5) special system processes.

d. Service/Resource Link. Guide analysis procedures use measures of system resources to analyze performance degradation. The Guide uses these resource measures (1) to isolate the processes or queue points that are major contributors to a particular elapsed service time, (2) to hypothesize causes of the elongation of these processes, and (3) to test the validity of the hypotheses, confirming the source of performance degradation. For example, the batch turnaround time analysis directs investigation to the GCOS process that is exhibiting the longest elapsed time of all GCOS system processes. This process (e.g., Core Allocation) requires system resources (i.e., core) to perform its service. A lack of these resources elongates the service (i.e., core allocation) to batch jobs.

e. Hypothesis Confirmation. The tests that the analyst is directed to conduct are used to confirm specific hypotheses as causes of the service elongation. These confirmations involve examining specific system software or performance data reports.

5. Computer System Tuning

Particular resource bottlenecks may be confirmed as elongating turnaround or response time. The Guide's analytical structure proposes specific solutions to correct or improve the degraded system performance. Several solutions can usually be applied to remove a particular bottleneck. In general, the Guide procedures provide up to four kinds of solutions to remove identified bottlenecks:

a. Scheduling Solutions. These solutions change the way that either batch or TSS workloads are scheduled for processing. They shift particular workloads to more evenly distribute system resources across the workload.

b. Parameter Solutions. These changes involve adjustments to system or subsystem functions. Examples include: (1) changes to the parameters of the GCOS Dispatcher or (2) a change in the placement of GCOS libraries. A solution may include specific changes to GCOS code, made through authorized software patch procedures.

c. Programming Solutions. These changes can involve modification of one or more application jobs running in the system. For example, Guide procedures are provided to investigate a program's execution characteristics in order to determine where it spends most of its execution time. This assists the programmer in examining the code. At a simpler level, Guide recommendations are made to speed application jobs by changing particular file locations discovered as delaying the job.

d. Sizing Solutions. These types of system change involve an increase (or a decrease) in the system's hardware configuration.

The Guide solutions are presented in a sequence that makes them easiest or least expensive to implement. Guide procedures direct the analyst to solutions that involve additions of new equipment only after other techniques have been tried.

B. COMPUTER SYSTEM TUNING USING THE GUIDE

Figure II-1 is a flow chart of the analysis process used throughout the Guide. The analysis process is comprised of two phases: (1) a Problem Definition Phase and (2) a Problem Analysis Phase. The activities of the Problem Definition Phase are directed toward determining whether a batch turnaround time or TSS response time problem actually exists. The activities of the Problem Analysis Phase are directed toward revealing causes of the identified turnaround time or response time problem.

1. Starting The Tuning Process

The tuning process may begin in one of three ways: (1) by direct request, (2) by an internal review of site-selected service elongation metrics, or (3) by user requests.

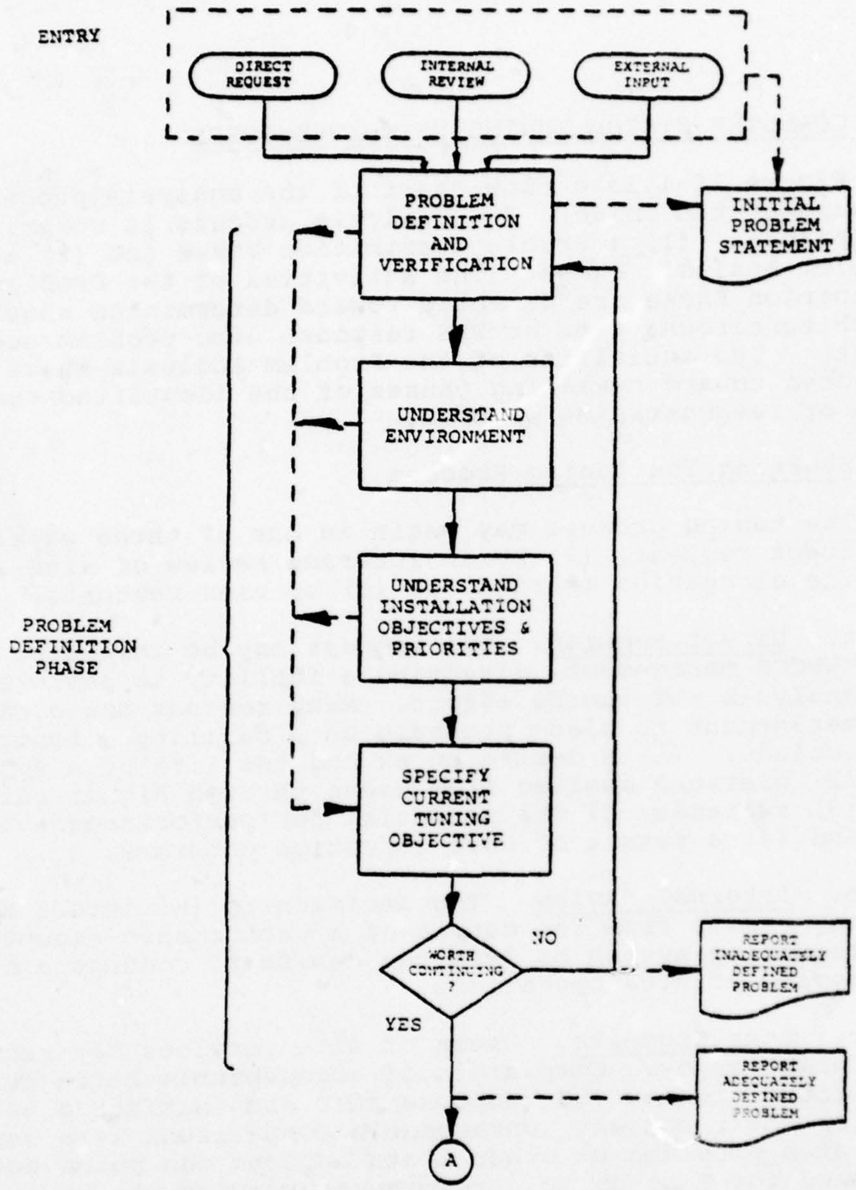
a. Direct Request. The request may be initiated by WWMCCS management, directing a facility to perform an analysis and tuning effort. Many reasons could cause management to place pressure on a facility. Examples include: (1) a desire to extend the life of a system, (2) pressure applied from users through higher authority, (3) awareness of the potential for performance evaluation, and (4) a result of cost reduction programs.

b. Internal Review. The decision to initiate a study may result from the output of a performance exception reporting system or from the desire to conduct a periodic review of site operations.

c. User Requests. Users of site services may request a tuning study. Complaints of unacceptable batch turnaround time or TSS response time can initiate a search for their causes. Unfavorable comparisons with service rates provided at other installations can point out the need for a study to determine if service can be improved.

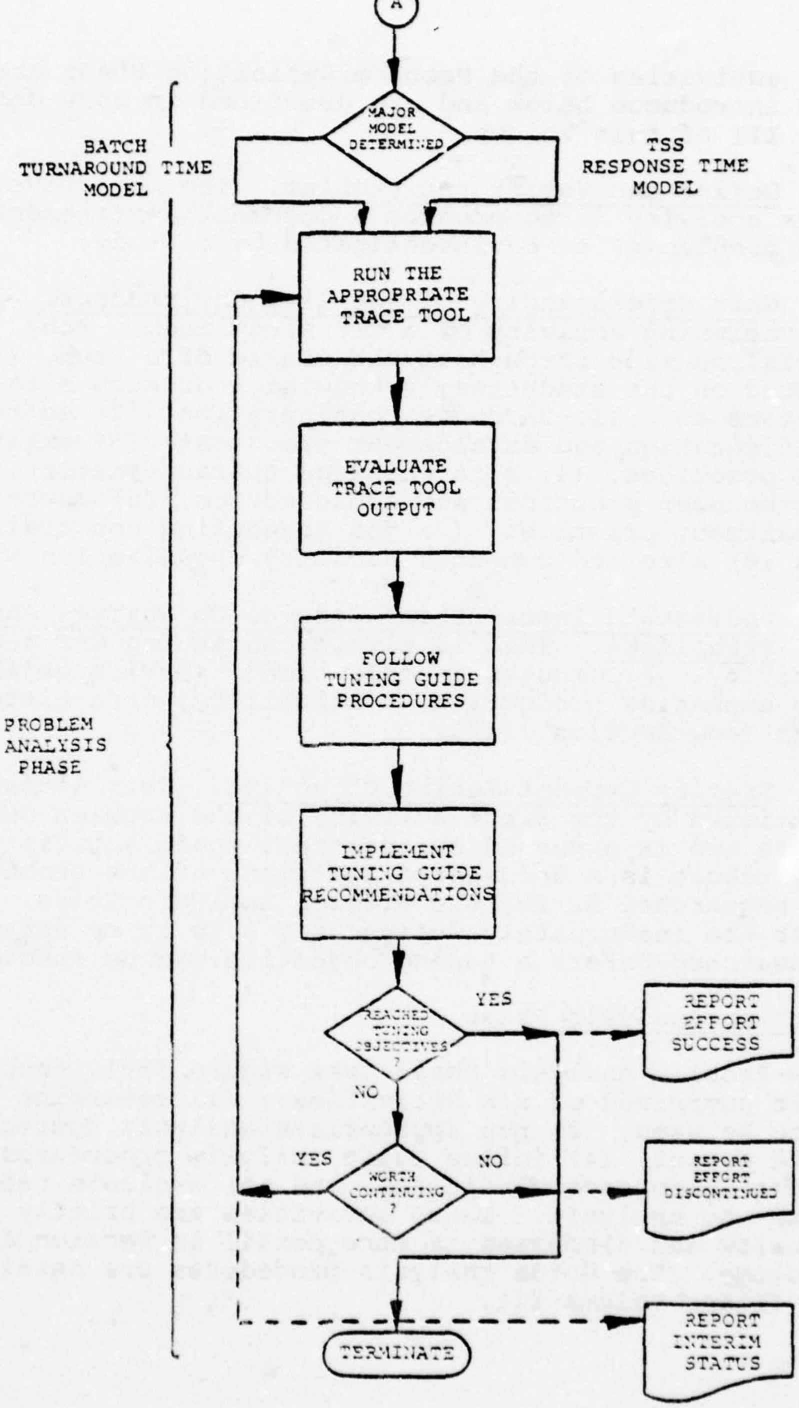
2. Problem Definition Phase

The Problem Definition Phase (see Figure II-1, first page) is comprised of four activities: (1) define and verify the problem, (2) gain understanding of facility environment, (3) understand installation service objectives and priorities, and (4) specify current tuning objectives. Some of these activities are initiated as a tuning study begins. Others are maintained as on-going activities.



FLOW CHART OF TUNING PROCESS

FIGURE II-1



FLOW CHART OF TUNING PROCESS

FIGURE II-1 (Cont'd)

The activities of the Problem Definition Phase are briefly introduced below and are described in more detail in Section III of this volume.

a. Define and Verify the Problem. The objective of this activity is to produce a documented statement of the problem(s) to be investigated by a study.

b. Gain Understanding of Facility Environment. This is a continuing activity of a CPE study team. Many decisions made throughout the course of a study will depend on the study team's knowledge of such site factors as: (1) hardware configuration, (2) software configuration and development practices, (3) existing CPE practices, (4) site workload characteristics, (5) system user practices and satisfaction, (6) operations department practices, (7) job scheduling constraints, and (8) site and computer facility organization structure.

c. Understand Installation Service Objectives And Priorities. This is also a continuing CPE team activity. A computer installation's service objectives may emphasize production, availability, or a mixture of both (see Section III.C).

d. Specify Current Tuning Objective. This activity is initiated by the first activity of the Problem Definition Phase and is expanded by the other phase activities. Its result is a documented statement of the problem to be researched during the Problem Analysis Phase. Problems that are inadequately defined may have to be further researched before a tuning objective can be specified.

3. Problem Analysis Phase

The Problem Analysis Phase (see Figure II-1, second page) is comprised of six activities: (1) determine major model to be used, (2) run appropriate analysis system, (3) evaluate output, (4) follow Guide analysis procedures, (5) implement Guide recommendations, and (6) evaluate need to continue the analysis. These activities are briefly introduced below and discussed in more detail in Section IV of this volume. The Guide analysis procedures are detailed in Volume II and Volume III.

a. Determine Major Model to be Used. The purpose of this activity is to choose between the Batch Turnaround Model and the TSS Response Time Model. The model selected depends on the tuning objectives specified in the Problem Definition Phase. Most tuning studies will be initiated as either batch turnaround time analyses or TSS response time analyses. Those that address neither of these two areas (e.g., an analysis of Transaction Processing System response times) may use the contents of this volume as a general guide; however, following the detailed procedures in Volumes II and III may be of little use to them.

b. Run Appropriate Analysis Tool. Both analyses require an initial collection of data to start an investigation. Special data collection tools have been developed for this purpose. The concepts, facilities and formats of the Batch Turnaround Time Analysis System are described in Appendix B. Similar information is provided in Appendix H for the TSS Response Time Analysis System.

c. Evaluate System Output. The Guide gives procedures for interpreting the initial output from the analysis systems. These procedures direct the CPE analyst toward certain hypotheses and further procedures which test these hypotheses.

d. Follow Guide Analysis Procedures. The Guide provides specific procedures to test hypotheses by examining the reports produced by standard WWMCCS performance tools. Specific system tuning steps can result from the tests.

e. Implement Guide Recommendations. The Guide recommends the following types of solutions to identified elongation causes: (1) scheduling changes, (2) parameter changes, (3) programming changes, and (4) configuration changes.

f. Evaluate Need to Continue the Analysis. This decision is made after the relevant recommendations have been implemented. If the current tuning objective (specified in the last activity of the Problem Definition Phase) has been met, the analysis is complete. If it has not been met, the Guide gives instructions for further investigation.

III. PROBLEM DEFINITION PHASE

This section describes the activities associated with the Problem Definition Phase of the WWMCCS System Tuning Process. These activities are: (1) define and verify the problem, (2) gain understanding of facility environment, (3) understand installation service objectives and priorities, and (4) specify current tuning objective.

A. DEFINE AND VERIFY THE PROBLEM

For whatever reason the tuning study was initiated, a study starting point or premise must be established. The problem must be described as well as it can at this point.

Problems may be stated at very gross levels. "Batch turnaround time is poor" or "TSS response time must be improved" are valid statements of problems at this time.

Site management must then verify that this is the problem they wish to pursue. The facility staff and users should verify that, in fact, this problem does exist and that they view it as a problem of importance. Note that a specific tuning objective is not yet defined; only the basic problem statement is verified.

B. GAIN UNDERSTANDING OF FACILITY ENVIRONMENT

This second step helps the CPE analyst relate the system environment to the problem statement. This activity may be time consuming if the analyst has little personal experience with the facility and the system.

Information collected during this activity will help determine the reason for the problem's importance; it may explain a reason for the problem's existence and help rank one problem against another. It probably will enable the selection of a particular tuning guide model. This activity will provide a means of understanding the significance of the Guide's tests. The activity forces an analyst to view the entire facility; this is important since many performance problems are caused by combinations of factors. The activity assists the analyst in understanding how to narrow, refine, modify, and improve the problem's definition in order to attempt a valid solution.

Each area described in the following paragraphs should be examined. It may help the analyst to examine them in the order they are given below. The associated lists (Tables III-1 through III-8) are intended to suggest areas of concern, and are not exhaustive. The analyst should:

1. Hardware Configuration

Determine the exact system configuration. Diagram the overall structure of the system. Collect (or begin collecting) reliability statistics on major configuration components and determine the history of hardware changes to the configuration. Table III-1 details several individual items of concern.

2. Software Configuration And Development Practices

Determine the exact operating system configuration. Note any local changes made to the system and any specialized major subsystems that are running. Note software monitors and other CPE measurement techniques used. Determine program optimization techniques used and note standards imposed on operations and programming staffs. Table III-2 details some individual items of concern.

3. Existing CPE Practices

Determine who has responsibility for computer performance evaluation at the site. Note sources of CPE data that are employed and determine how the data are used to evaluate system performance. Determine if personnel in all site functional areas (e.g., programming and operations) can relate to the performance data. If CPE studies have been conducted at the site in the past, how have they been documented? Table III-3 details some individual items of concern.

4. Site Workload Characteristics

Determine how the existing workload uses system resources. Note patterns of resource use by selected jobs. Determine any obvious bottlenecks to the handling of jobs within the installation. Table III-4 lists some areas of concern.

5. System Users

Expand upon the workload analysis by investigating the practices of the major users in the installation. Determine how special priorities are assigned to these users and whether the users can directly control system resources. Note chargeback schemes used (if any) to level the workload across the operating day and night. Determine the levels of user satisfaction (or dissatisfaction) exhibited. Table III-5 details these and other items of concern.

AREA	ITEM
Central Processing Unit	Model(s) Options Special Features
Memory	Size(s) Speed(s) Interleaving
I/O Data Pathways	Types of Channels Peripheral Device Subsystems Connected Physical/Logical Data Paths Manual/Automatic Switching Capabilities for Devices Rated I/O Speeds (Channels/ Controllers/Devices) Prioritization for Memory Access Contention with CPU for Memory Access
Peripherals	Types and Quantities Capacities Addressing Conventions I/O Path Priorities Data Transfer Modes (i.e., Burst, Word, etc.) Maximum Number Per Controller/ Channel Maximum Number Active Under a Given Controller Overlapped Activity Permitted Between Devices Operations; Special Features Restrictions Imposed on I/O Data Paths by Type of Device

HARDWARE CONFIGURATION

TABLE III-1

AREA	ITEM
Communications Network	Number of Lines Baud Rates of Lines Ownership of Lines Terminal Types Synchronous or Asynchronous Line Types Buffer Size Requirements on Host Computer Physical Location of Users and Terminals Polling Discipline
Configuration	Diagram Showing All Physical/ Logical Interconnections
Reliability Statistics	Mean Time Between Failure for Major Components Mean Time To Repair for Major Components
Hardware History	Growth of Hardware Inventory Over Time Reasons for Hardware Upgrades/ Changes

HARDWARE CONFIGURATION

TABLE III-1 (Cont'd)

AREA	ITEM
Operating System	Release Version Site Options Enabled Additional Options Permitted
Local Modification	Changes Modules Obtained from Users Group
Major Utility Systems	Vendor-Standard Packages In-House Developed Packages Software House Packages Packages Received from User Groups
Special Capability Systems	Text Editors On-Line Retrieval Systems Special Timesharing Interface Subsystems Special RJE Systems Special Operator-Assist Systems Library Aid Systems Program Flow Chart Systems Program Evaluator Systems (Static)
System Accounting File Data	Types of Data Quantity of Data Storage Media
System Accounting Data Reduction System	Commercially Available Developed In-House Frequency of Summarization Purpose for Which Data Are Accumulated
Software Monitors	Types of Data Collected Quantities of Data Captured Normal Monitoring Schedule

SOFTWARE CONFIGURATION AND DEVELOPMENT PRACTICES

TABLE III-2

AREA	ITEM
Language Processors	Percentage of Programs Coded in Each Language Special Versions (e.g., Optimizers) Available
Program Optimization Systems	Module Reorganizers Instruction Frequency Analyzers (Dynamic)
Programming Standards Imposed	I/O Data Block Sizes Syntax Standards Specific Practices Outlawed Instruction in Efficient Programming Techniques Provided

SOFTWARE CONFIGURATION AND DEVELOPMENT PRACTICES

TABLE III-2 (Cont'd)

AREA	ITEM
Responsibility for CPE	Identify Staff Members Other Duties Power to Influence Policies and Procedures
Sources of CPE Data	Manual Shift Logs System Console Messages Standard System Utilities (e.g., SAVEMAST Listings) System Accounting Data Files Software Monitors Hardware Monitors
CPE Data Usage	To What Management Level Are CPE Reports Presented? What Periodic Reports Are Generated? What Is the Intent of Reporting CPE Data?
Completeness of Data Analysis	What Data Elements Are Reported by Exception? What Are Exception Threshold Parameters? Who Sets Threshold Limits? How Are Thresholds Determined?
Acceptance of Recommendations	Ignored by Other Technical Personnel Ignored by Management Treated as Invasion into Realm of Responsibility by Certain Managers Accepted Unwillingly Accepted and Generally Implemented Enthusiastically Supported by High Enough Level of Management to Implement Changes
CPE Documentation	In What Format Are Studies Documented? Who Receives Documentation? What Changes Have Been Documented?
CPE Follow-Up	History of Change Follow-Up Formal Reevaluation Procedures

EXISTING CPE PRACTICES

TABLE III-3

AREA	ITEM
Workload Characteristics	Proportion of Workload by Distinct User Function Predominance of Workload toward Batch/On-Line/Certain Peripherals I/O- or CPU-Dominance of Workload Patterns of Workload Through Day: - Volume Changes - Time-Dependent Output - Cyclic Patterns of Workload Across Month/Year
Historical Resource Usage by Selected Jobs	Distributions of Memory Demand Distributions of CPU Usage Distributions of Elapsed Times Distributions of Response Times/ Turnaround Times Distributions of Activities Per Job Distributions of Time Sharing Commands and Resource Demands of Individual Commands Distributions of Peripherals Used (i.e., Disk, Tape) I/O Loads by Device Type Distributions of I/O Loads by Job/Activity
Job Breakdown	New Program Development New Program Test/Debug Modifications to Existing Programs Reruns - Distribution by Function/ Department/Major System - Analysis by Reason for Failure
Workload Handling	Backlog? Response Times Met? Time-Dependent Completions Met? Stand-Alone Execution for Critical Jobs?
Workload Security	Impact on Processing of Other Workload Priority of Secured Workload Secure Running Periods

SITE WORKLOAD CHARACTERISTICS

TABLE III-4

AREA	ITEM
Workload Distribution by User/Function	Number of Jobs/Activities CPU Usage I/O Usage Specific Peripheral Usage Type of Processing (Batch, On-Line, RJE)
Special Priorities Established by User/ Function	Initiation Priorities (i.e., Job Scheduler Classes)
Ability of Users to Control System Resources	Dedicated Disk Spindles Limitations on Use of Temporary On-Line Storage Limitations on Number of Concurrent Jobs Backup Files Maintained in Library Memory Limitations Time-Of-Day Restrictions for Certain Processing
Charge-Back Schemes	Charge-Back Algorithms Used (if any): - Resource Use - Priority Execution - Time-Of-Day Variations - Variations Arising from Manual Intervention and Support Requirements Are Budget Estimates for Computer Use Required?: - By Department - By Function/Major System
Naming Conventions	Are Job/Program Naming Conventions Imposed on Users That Might Aid in Defining the Workload? Are There Similar Conventions Imposed on Disk/Tape Files to Identify Storage Media Use?

SYSTEM USERS

TABLE III-5

AREA	ITEM
System Scheduler Conventions	Do the Conventions Vary at Operator Discretion or User Demand to Affect Availability of the System Throughout the Day? Are the Conventions Static?
User Satisfaction	How Are User Complaints Handled? - Formal Procedure - Log Maintained - Periodic Management Review Current Level of User Satisfaction for Turnaround Time/Response Time/ Services Formal Procedure to Determine User Satisfaction

SYSTEM USERS

TABLE III-5 (Cont'd)

6. Operations Practices

Determine the operating shift schedule for the site. Note pre-scheduled non-production time and unscheduled non-production time. Examine training provided for operators and systems programmers. Study site library maintenance procedures. Examine the production control points established by operations. Determine how formalized logs are maintained to track work as it passes through the installation. Table III-6 lists areas of concern.

7. Batch Job Scheduling

Determine how batch work is accepted for processing at the site. Note the techniques used to schedule the batch workload that are either automated or are manually implemented. Table III-7 outlines areas of concern.

8. Site Operating Processes

Prepare for the batch turnaround time analysis. Construct site Pre-Processing and Post Processing models for the Batch Turnaround Time Analysis System as directed in Guide Appendix B.

9. Site And Computer Facility Organization

Determine the organization structure and reporting authority at the site. This includes the organization structure of the installation as well as the organization of the site itself. Determine the extent to which system sizing activities are organizationally separated from operations and applications systems development. Table III-8 outlines areas of concern.

C. UNDERSTAND INSTALLATION SERVICE OBJECTIVES AND PRIORITIES

This third step determines site processing objectives and priorities to be applied to decisions made during both the Problem Definition and Problem Analysis Phases.

1. Installation Service Objectives

To be effective, a computer performance evaluation (CPE) effort must consider the hardware, software, personnel, and service objectives of the computer installation being analyzed.

AREA	ITEM
Shift Information	Days of Week System Available Hours of Day for Shift Definitions Shifts per Day System Available
Prescheduled Non-production Time	Preventive Maintenance Operator Instruction System Programmers' Dedicated Time Shift Change/IPL Periods Normally Slack Shifts Special Test Time Special Dedicated Mode Operation
Unscheduled Non-production Time	History of Hardware Down Time History of Software-Related Down Time History of Repair Delays IPL Counts/Reasons
Shop Access Policy	Closed Shop to All but Operations Personnel Open Shop with Limited or Unrestricted Access
Training for Operators and Systems Programmers	Hiring Experience Levels Formal Schools On-The-Job Training In-House Classes Contracted On-Site Classes Periodic Training Updates or Advanced Training Job Descriptions Manning Requirements for Different Levels of Proficiency Number of Supervisors and Workers

OPERATIONS PRACTICES

TABLE III-6

AREA	ITEM
Library Maintenance	Proximity to Computer Room Are Tapes and Disk Packs Dispatched Prior to Job Submission? Availability of System Displays for Mount Messages Automatic Library Maintenance Procedures Migration Policies for On-Line Data Sets to Other Media Retention Cycle for Tape - Generation or Time Based?
Production Control Points	User Problem Resolution Points (e.g., Lost Listings, JCL Correction) Work Entrance Points User Special Service Points (e.g., SYSOUT Bursting, Decollation) Periodic Meetings with Prime Users to Resolve Problems, Estimate Future Requirements for Support, Reduce Friction)
Formal Logs Maintained	Shift Supervisor Hardware/Software Trouble Maintenance Records User Complaints Reruns Operator Errors Special SYSOUT Forms SNUMB Registers for Batch Work
Operator Specialization	Separate Console Operators Separate Peripheral Operators Separate Library Personnel Availability of Console Messages for Each Operating Station

OPERATIONS PRACTICES

TABLE III-6 (Cont'd)

AREA	ITEM
Work Acceptance	Proportion Scheduled As Received Proportion RJE Scheduled As Received Proportion Prescheduled, Repetitive, Known in Advance Proportion Interactive, Unknown in Advance, Unscheduled
Scheduling Constraints	Completion Date/Time Dependent Scheduling Resource-Based Scheduling Evaluation of Compatible Job Mix Performed Use of Software Scheduling Packages Establishment of Priorities Handling of Classified or Sensitive Jobs

BATCH JOB SCHEDULING

TABLE III-7

AREA	ITEM
Computer Facility As Seen by Users	Viewed As a Service Facility with Little Power Over the Applications Run Are Users an Integral Part of the Same Organization As the Computer Facility? Are the Users Within the Same Organizational Authority Even if Not Under the Same Executive?
Compare Level of Facility to Other Similar Site Departments	Programming Systems Programming Production Control Scheduling Plans and Requirements Functions Users
Separation of Functions	Hardware Sizing and Operation Hardware Sizing and Applications Development

SITE AND COMPUTER FACILITY ORGANIZATION

TABLE III-8

Although the first three areas are readily identifiable, the site's service objectives are often misunderstood. An installation's objectives can emphasize production or availability, or a mixture of both. A mixture is generally dominated by either a production or availability objective. A production objective attempts to take full advantage of the capabilities of the system in terms of throughput. With a customer-oriented availability objective, however, the interests of the users take priority over the most efficient use of system resources. In either case, the computer installation's management objective is return on investment. For the production objective, the most important investment is the computer; for the availability objective, the most important investment is the people or system using the computer. The difference between these two objectives is reflected in how the performance parameters are evaluated.

a. Production Objective. Traditional data processing is production-oriented. The computer installation is viewed as a large investment for capacity to do routine work. Much of that capacity depends on a high degree of simultaneous use of many system components. Management's production objective is to use as many of the components as much of the time as possible by scheduling compatible jobs. Under these conditions, low batch turnaround and TSS response times are secondary to high machine utilization. Management's success is generally defined in terms of a relatively high number of jobs contending for computer resources (a high multiprogramming level), high central processor unit (CPU) activity, nearly full memory utilization, and highly active channels. Schedule-driven processing uses internal and external priority allocations to sustain efficiency. Programs must be written to share available resources. This can be done with program segmentation techniques, use of a minimum number of devices, the indirect use of input/output through spooling, and adherence to rigid standards - all at the expense of the individual job. Data must be so distributed that device activity is economically justified with the attainment of a low system wait, low unit cost, and high simultaneity.

b. Availability Objective. Service-oriented systems are more concerned with return on investment for users (managers, scientists, or programmers) than with the computer itself. Low TSS response and batch turnaround times are critical. Their increase delays user operations and such applications as on-line command and control systems, real-time update systems, scientific support services, and program development systems. Demand-driven processing varies in activity levels, both by users and type of work. Minimum emphasis is placed on scheduling. The system must therefore have available capacity ready to respond to demand. Such a system must have utilization rates well below the 100% limit in order to accommodate the variations in demand. Success is measured in terms of user satisfaction and little emphasis is placed on reporting high utilization figures.

c. Mixed Objectives. Growing numbers of computer installations are encountering (1) demands for highly responsive services and (2) pressure from management for high production rates. The two objectives are not mutually exclusive. Predominance of either objective can be identified within operational time periods, and management must evaluate whether the satisfaction of one objective might deleteriously affect the satisfaction of another.

Computer performance evaluation can serve the management of either type of computer installation and can also serve mixtures that might have different objectives, depending on the time of day or day of the week. However, it is important that managers, auditors, and executives recognize the implications of the different objectives when they compare the performance of one installation to another.

2. Installation Priorities

The installation's priorities derive from the predominance of either the production objective or the availability objective at a particular site. They also, by implication, determine the sequence in which system tuning solutions are applied.

a. Service Priorities. A site may feel that low TSS response time is more desirable than low batch turnaround time for a certain period of the operating day. The analysis procedures presented in the Guide are not generally directed toward determining which of the two is "more important" than the other. However, the sequence of examining either of the two service areas may be affected by this priority.

b. Tuning Solution Priorities. Tuning solutions are presented in Volume II and Volume III to correct system bottleneck conditions. In nearly all cases, more than one solution is specified to correct a problem. The solutions are generally proposed in a sequence that recommends the more quickly (or easily) applied solution be implemented before others. Installation requirements may change this sequence.

D. SPECIFY CURRENT TUNING OBJECTIVE

This fourth step is used to determine whether the originally specified problem can actually be investigated with Guide analysis procedures.

1. Tuning Objective

A tuning objective is a stated (i.e., documented) goal of an analysis. A tuning objective must be stated in specific and quantified terms. The tuning objective statement is used to determine when a particular Guide analysis has been completed.

Examples of well-stated tuning objectives are: (1) "reduce mean response time for (stated) non-trivial TSS commands to five seconds" and (2) "reduce the mean batch turnaround time for (stated) jobs to 1.5 hours." A well-stated tuning objective includes: (1) a definition of the workload category, (2) a description of the process to be investigated, and (3) a service metric value.

Examples of badly-stated tuning objectives are "reduce TSS response time" and "improve turnaround time." Note the missing components of these tuning objectives.

2. Tuning Objective Decision

Determine whether the tuning objective as documented is realistic, attainable, and a cost effective target.

a. Attainable Tuning Objectives. Tuning objectives must be attainable. It might be possible to reduce a program's elapsed time to certain limits, but not to a desired limit, simply because of the quantity of I/O and computation that occurs within the program. If this is the case, re-evaluate the tuning objective.

b. Realistic Tuning Objectives. It might be possible to reduce a program's elapsed time to the desired amount, however, particular site constraints may prevent its being a realistic goal. These constraints might have their source outside of the tuning effort, but directly affect the internal performance of the program. These might include, for example, the requirement to give certain other jobs higher priority. If this is the case, re-evaluate the tuning objective.

c. Cost-Effective Tuning Objectives. The additional effort, cost, and time required to achieve an attainable and realistic tuning objective might not be worth it. A particular increment of performance improvement might simply not be cost effective. Determine the amount of improvement likely with additional effort. Decide whether the effort might be better off abandoned.

3. Determine If Worth Continuing

The final activity of the Problem Definition Phase is to determine whether the tuning process itself is worth continuing. There may be potential performance problems of greater magnitude or immediate importance that may have been uncovered during the Problem Definition Phase. Document the decision and the tuning objective and obtain management approval of them.

4. Begin Problem Analysis

Having obtained concurrence from management, begin the second half of the performance evaluation process (see Figure II-1, second page): the Problem Analysis Phase. The activities of this phase (see Section IV) involve the selection of a particular tuning model, analysis of the system processes defined by the model, and tuning of system components.

IV. PROBLEM ANALYSIS PHASE

This section describes the activities of the Guide's Problem Analysis Phase (see Figure II-1, second page). These activities include: (1) determine major model, (2) run appropriate analysis system, (3) evaluate analyzer output, (4) follow analysis procedures, (5) implement tuning recommendations, and (6) evaluate the requirement to continue the analysis.

A. DETERMINE MAJOR MODEL

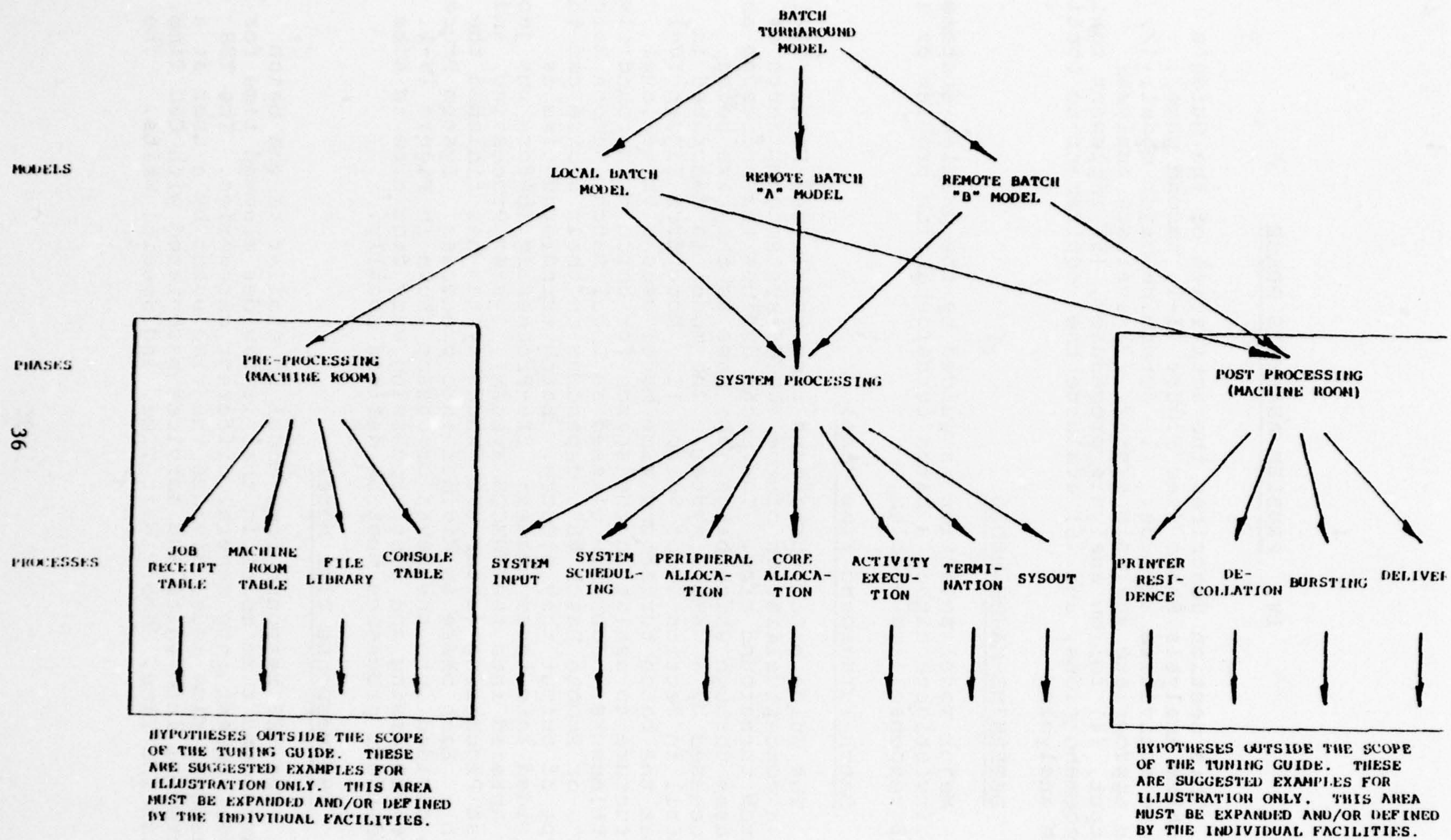
Major model selection is guided by the problem statement: to investigate either a batch turnaround time problem or a TSS response time problem.

1. Batch Turnaround Time Model

The Guide batch turnaround time model (see Figure IV-1) is a conceptualization of the WWMCCS system components of batch turnaround time. The model defines the processes and phases through which batch jobs pass as they are being processed by a WWMCCS system. The model is described in detail in Section II of Volume II. Note from Figure IV-1 that the batch turnaround time model uses a three-level structure to assist in the search for batch turnaround time bottlenecks. Jobs are classed as Local Batch, Remote Batch "A", or Remote Batch "B", depending on their source and the type of output they produce. Batch turnaround time is divided into three phases: Pre-Processing (before the job is entered into the WWMCCS system), System Processing, and Post Processing (after the WWMCCS system has finished the job). Each phase is divided into processes. System Processing is divided into the seven processes shown in Figure IV-1. Pre-Processing and Post Processing vary from site to site and their processes must be defined locally.

2. TSS Response Time Model

The TSS response time model is similar to the batch turnaround time model in that it divides elapsed time for a user request into several different categories. The TSS response time model divides the time spent by a user at a terminal into waits and services associated with CPU Time, Disk I/O Time, Memory Wait Time, and Special Waits. The



SUMMARY OF BATCH TURNAROUND TIME MODEL STRUCTURE

FIGURE IV-1

model further subdivides each of these categories into two or more subcategories. The amount of response time associated with each category and subcategory guides further investigation into improving response time. Again, choice of a particular category for further investigation depends on the amount of response time associated with that category.

Figure IV-2 provides a graphic description of the TSS Response Time Model, which is defined in detail in Volume III.

B. RUN THE APPROPRIATE ANALYSIS SYSTEM

The two Guide models employ special software to determine where a user spends the largest amount of elapsed time in their respective processes.

1. Batch Turnaround Time Analysis System

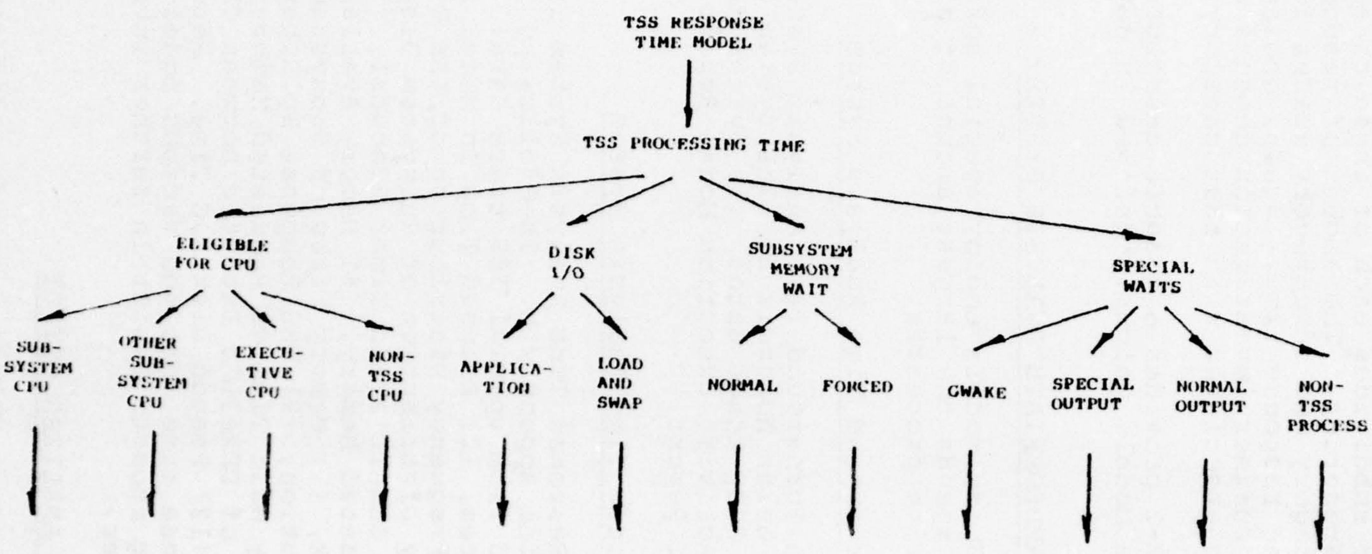
The Batch Turnaround Time Analysis System is described in detail in Guide Appendix B. The system generates four reports from GCOS trace data: (1) System Report, (2) Model Report, (3) Activity Execution Model Report, and (4) Disk Space Refusal Report.

2. TSS Response Time Analysis System

The TSS Response Time Analysis System is described in detail in Guide Appendix H. The following twelve reports are generated from special TSS trace data: (1) Elapsed Time In Model States, (2) Elapsed Time In Model States By Subsystem, (3) Frequency Distribution of TSS Dispatch Lengths, (4) Frequency Distribution of Subsystem Dispatch Lengths (Joined Over DERAIL's And Timer Runouts), (5) Memory Wait Time vs Requested Memory, (6) Memory Available For Subsystems--Largest Block, (7) Memory Sizes of Subsystems, (8) TSS Core Size Distribution, (9) TSS Response By Subsystem, (10) Urgent Memory Wait Time vs Requested Memory, (11) Frequency Distribution of DERAIL's Executed Between CPU Eligibility Losses, and (12) Pseudo Disk I/O Time. Reports one and two divide response time into the various model categories, and the remaining reports assist in further investigations into the categories.

C. EVALUATE ANALYZER OUTPUT

Both procedure Volumes provide instructions for evaluating initial analyzer outputs. These instructions guide



SUMMARY OF TSS RESPONSE TIME MODEL STRUCTURE

FIGURE IV-2

the analyst to specific test procedures. For both analyses, the instructions are implemented through a form that is filled in by the analyst. This form directs the type and sequence of tests to be conducted during the search.

D. FOLLOW GUIDE TEST PROCEDURES

The fourth activity of the Problem Analysis Phase is to follow the specific analysis procedures documented in Volume II and Volume III.

1. Test Procedure Format

A Guide test procedure contains a series of analytical steps that are directed toward a diagnostic objective. For example, one of the procedures under the CPU Execution Characteristics Test in the batch turnaround time analysis has the following objective: "...to determine the dominant CPU user." The steps of these procedures involve: (1) examining specific reports to obtain performance data, (2) entering selected metric values on a form, (3) calculating ratios or percentages from these entry values, and (4) making certain tuning decisions (and subsequent recommendations) from the calculated values.

In some Guide test procedures, several CPE tools may be used to gather data. If this is the case, it is the analyst's responsibility to provide synchronized data (i.e., reports produced from multiple sources that are measuring the systems over a common elapsed time).

Each test procedure uses a form on which the significant measurement values are entered. The form helps clarify the procedures; use of the form documents the tuning effort for later management and analyst reference.

A ratio analysis technique is used to compute values which are quantities compared with tuning decision criteria. As an example, one of the procedure tests results in the calculation of an Activity CPU Ratio and a GCOS CPU Ratio. These ratios are then employed in a decision step to further direct the test. All computations required in each test are made directly on the appropriate test form under the direction of a procedure step.

The bracketed numbers or expressions in the text of volumes II and III (for example, [0.3] or [the next to the largest]) are an attempt to clarify what should be considered as "high" or "low" for a particular decision. The suggested values are based on estimates of "typical" statistics and workloads; they will not be correct in every case. If a situation does not seem to "fit" the assumptions or if a decision value is nearly equal to the bracketed expression, three actions may be indicated: (1) change the bracketed value based on experience, (2) explore both logical paths leading from the decision in question, and/or (3) seek outside help.

2. Batch Turnaround Time Tests

Nine tests are included in the Turnaround Time Analysis Procedures (Volume II, Sections IV - XII). The Turnaround Time Model Scan section gives instructions for evaluating the initial Batch Turnaround Time Analysis System output and directs the analyst to one or more of the following tests.

a. Seek Elongation Test. This test confirms excessive seek lengths by calculating Weighted Average Seek Length for all configuration disk units. The test then identifies the GCOS and User files that are contributing to the seek elongation on the active disk units.

b. Memory Constraint Test. This test identifies bottlenecks due to lack of memory using the following memory-related metrics: (1) Average Elapsed Memory Wait Time, (2) Processor Idle Time, and (3) Number of I/O Entries Queued.

c. Device Errors Test. This test uses direct measurement of device errors to isolate faulty devices and media. Procedures include: (1) Determine Tape Handler Units in Error and (2) Determine Disk Units in Error.

d. Pathway Utilization Test. This test identifies disk channel bottlenecks by direct measurement of the following disk I/O queue conditions: (1) device queues, (2) channel queues, and (3) queues for both channel and device.

e. CPU Execution Characteristics Test. This test analyzes CPU bottlenecking. It first determines the CPU's utilization level and then determines if the CPU was dominated by GCOS or user program code execution. Two test procedures are then executed to identify and investigate particular programs.

f. Insufficient Devices Test. This test determines whether there are sufficient tape drives and/or disk space.

g. "Few Activities In System Test". This test uses an analysis of memory demand to determine whether additional jobs could have been scheduled in the system.

h. IOS Delays Test. This test isolates excessive IOS overhead by direct measurement of IOS delay time using GCOS trace entry data.

i. Urgency Codes Test. This test confirms job elongation due to urgency code problems.

3. TSS Response Time Tests

The TSS response time tests are documented in a format similar to those of the batch turnaround time tests. The details of these tests are contained in Volume III.

a. Subsystem CPU Time Search Procedure. This procedure deals with a system where a large part of response time is spent using the CPU. A system with unacceptable TSS response time may achieve acceptable response for most types of commands, but achieve unacceptable response for a few. This procedure determines whether adjusting the relative CPU priorities of subsystems will result in acceptable response for all types of commands.

b. TSS Executive CPU Time Search Procedure. This procedure deals with a system where a large part of response time is spent waiting for executive functions. The amount of TSS Executive CPU time is controlled by several factors: the types of demands made by other subsystems, various waits between periodic executive functions, and a few cutoff periods that limit how long TSS will continue to request some action from GCOS. This procedure investigates these sources of TSS Executive CPU Time.

c. TSS Wait For CPU Time Search Procedure. This procedure investigates means of manipulating the dispatching priorities for TSS and emphasizes granting TSS priority over other jobs.

d. Disk I/O Time Search Procedure. High disk I/O times may be caused by a general disk I/O problem, by wrong usage or placement of work files, by insufficient I/O queue space, and by excessive use of disk files. General disk I/O problems are referred to the Pathway Utilization Test of the batch turnaround time analysis.

e. Memory Wait Time Search Procedure. Long memory wait times are caused by lack of TSS swap core in which to load subsystems. This may be caused by three types of problems: (1) a swap file problem that prevents swapping of subsystems, (2) too little swap core, and (3) excessive memory demand. These, and other memory wait time related hypotheses, are investigated in this procedure.

f. GWAKE Wait Time Search Procedure. The procedure for analyzing GWAKE time directs the analyst to the subsystems involved.

g. Output Wait Time Search Procedure. This procedure includes (1) investigate memory availability and user errors, (2) investigate I/O queue space, (3) investigate possible DATANET delays, and (4) investigate output volume.

h. Non-TSS Service Wait Time Search Procedure. The Non-TSS process state is used by the TSS Executive to describe subsystems that are waiting for certain types of service to be performed outside of TSS. While the factors causing long Non-TSS Service Waits are outside TSS and they are not diagnosed by the TSS Response Time Analysis System, suggestions are offered which may help reduce these waits.

E. IMPLEMENT GUIDE TUNING RECOMMENDATIONS

The analyst is directed during or at the end of each test to a series of tuning recommendations. These recommendations incorporate generally accepted system tuning practices as applied directly to the WWMCCS system.

1. Types of Guide Tuning Recommendations

The Turnaround Time tuning recommendations, called "tuning solutions," include: (1) scheduling solutions, (2) parameter solutions, (3) programming solutions, and (4) sizing solutions. One or more solutions may be proposed for an identified bottleneck.

a. Scheduling Solutions. These recommendations propose changes to the job schedule. Particular scheduling steps depend upon the problem uncovered.

b. Parameter Solutions. These involve a change to some GCOS or subsystem function. The change may involve a parameter in GCOS code, system file placement or any process affected by a change to system software.

c. Programming Solutions. These are proposed to modify programs, files, or job control cards. Some changes can be implemented quickly; others will take some time.

d. Sizing Solutions. These are proposed to modify a hardware configuration in a specific manner. Both increases and decreases in configuration components can be recommended.

The TSS Response Time tuning recommendations are integrated into the procedures.

2. General Selection Criteria

The tests in both models provide specific tuning recommendations. As a general rule, these recommendations are proposed in the following sequence: (1) scheduling changes, (2) parameter changes, (3) programming changes, and (4) sizing changes. Sizing solutions should be specified only after all other approaches have been attempted. It is expected that each site will modify its own selection criteria as it develops experience with the Guide.

3. Complex Selection Criteria

There are times when specific tuning recommendations cannot be made. In these cases, the analyst is directed toward a general solution to the problem.

As an example, an indication that CPU power should be increased cannot specify how many CPU's should be added to a configuration. It is not possible for a simple analytical procedure to define the specific number of processors to be added to a configuration. The analyst is required to investigate this type of tuning recommendation by employing such tests as benchmarks. A test of this complex nature is beyond the scope of the Guide.

4. Verification of Change

Measurement should be continued for a representative period of time after changes have been made to the system. Measurement is required to verify that changes achieve their desired result.

F. EVALUATE REQUIREMENT TO CONTINUE

Analysts should determine whether tuning objectives have been met after each implementation of a tuning step. Even if a tuning objective is not met at any particular iteration of tests and tuning step, an analyst may observe an improvement in turnaround time or TSS response time. This improvement will act as a checkpoint, indicating that the overall analysis is proceeding in the right direction. If it appears difficult to reach a tuning objective after several attempts, it is possible that the original tuning objective may have been over stated. In this case the analyst may want to reformulate the tuning objectives.

The Guide cannot cover every possible situation. Some cases may require specialist attention. The Guide (Volumes II and III) tries to anticipate this in many cases by calling to the analyst's attention combinations of data values which the Guide does not cover and which call for special attention. In these cases and others where the Guide appears not to analyze the situation properly, experienced computer measurement help should be obtained.