AD-A039 111 MITRE CORP BEDFORD MASS F/G 9/2
WWMCCS H6000 MULTIPROCESSOR PERFORMANCE EVALUATION. VOLUME I.(U)
FEB 77 G A NELSON F19628-77-C-0001
UNCLASSIFIED MTR-3350-VOL-1 ESD-TR-77-18-VOL-1 NL

1 OF 2
AD A039111



ADA039111

SYSTEMS TECHNOLOGY STUDY 3-77

ESD-TR-77-18

MTR 3350
REV. 1

WWMCCS H6000 MULTIPROCESSOR PERFORMANCE EVALUATION
VOLUME I

FEBRUARY 1977

DIRECTORATE OF SYSTEMS TECHNOLOGY
AIR FORCE DATA SYSTEMS DESIGN CENTER
AIR FORCE DATA AUTOMATION AGENCY
GUNTER AFS, AL 36114

DEPUTY FOR AFWWMCCS
ELECTRONIC SYSTEMS DIVISION
AIR FORCE SYSTEMS COMMAND
HANSCOM AFB, MA 01731

MITRE-BEDFORD
A DIVISION OF
THE MITRE CORPORATION
BEDFORD, MA 01730

DDC
RECEIVED
MAY 6 1977
A

AD No.
DDC FILE COPY

APPROVED FOR PUBLIC RELEASE; DISTRIBUTION UNLIMITED.

AIR FORCE (1) MARCH 29, 1977--405

When U.S. Government drawings, specifications, or other data are used for any purpose other than a definitely related government procurement operation, the government thereby incurs no responsibility nor any obligation whatsoever; and the fact that the government may have formulated, furnished, or in any way supplied the said drawings, specifications, or other data is not to be regarded by implication or otherwise, as in any manner licensing the holder or any other person or corporation, or conveying any rights or permission to manufacture, use, or sell any patented invention that may in any way be related thereto.

Do not return this copy. Retain or destroy.

PREFACE

Users should address questions related to the subject of this report or to the possibility of extending the stated conclusions or recommendations to the Chief, Operations Research Division, AFDSDC.

REVIEWED BY

James I. Clogston

JAMES I. CLOGSTON
Ch, Operations Research Div
Directorate of Systems Technology

APPROVED FOR RELEASE

Bruce L. Fowler

BRUCE L. FOWLER, Colonel, USAF
Director of Systems Technology

REVIEW AND APPROVAL

This technical report has been reviewed and approved for ESD publication.

David C. Peterson

DAVID C. PETERSON, Major, USAF
Project Officer

Walter W. Turgiss

WALTER W. TURGISS
Director of System Requirements
Deputy for AFWWMCCS

FOR THE COMMANDER

E. W. Milauckas

EDMUND W. MILAUCKAS, Colonel, USAF
Deputy for AFWWMCCS

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER: ESD-TR-77-18, Vol. I; AFDSDC-STS-3-77, Vol. I
2. GOVT ACCESSION NO.
3. RECIPIENT'S CATALOG NUMBER (9)

4. TITLE (and Subtitle): WWMCCS H6000 MULTIPROCESSOR PERFORMANCE EVALUATION. Volume I. (6)
5. TYPE OF REPORT & PERIOD COVERED: Final rept.
6. PERFORMING ORG. REPORT NUMBER: MTR-3359 Vol-1 (14)

7. AUTHOR(s): George A. Nelson (10)
8. CONTRACT OR GRANT NUMBER(s): F19628-77-C-0001 (15)

9. PERFORMING ORGANIZATION NAME AND ADDRESS
The MITRE Corporation
Box 208
Bedford, MA 01730

10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS
PE 63735F
Project 2188

11. CONTROLLING OFFICE NAME AND ADDRESS
Deputy for Air Force WWMCCS
Electronic Systems Division, AFSC
Hanscom Air Force Base, MA 01731 (11)

12. REPORT DATE: FEBRUARY 1977
13. NUMBER OF PAGES: 175

14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office): 177 p. (12) 2188 (16)

15. SECURITY CLASS. (of this report): UNCLASSIFIED
15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited.

(18) ESD, AFDSDC

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

(19) TR-77-18-Vol-1;

18. SUPPLEMENTARY NOTES

STS-3-77-Vol-1

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

H6000 CONFIGURATION ALTERNATIVES
HONEYWELL INFORMATION SYSTEMS H6000
MULTIPROCESSOR PERFORMANCE
PERFORMANCE EVALUATION
RELATIVE THROUGHPUT
WWMCCS ADP

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

This report presents an overall description of the WWMCCS Multiprocessor Performance Evaluation task sponsored by the Deputy for Air Force WWMCCS, ESD, during FY76. This task involved the collection and analysis of empirical data from controlled performance tests using synthetic workloads on WWMCCS H6000 computer systems comprised of from one to four central processing units. Volume I describes the rationale for initiating the task, the technical approach, the test results, and a

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

235 050

20. Abstract (continued)

summary of limitations, observations and applicability. Appendix I of Volume 1 documents an independent verification of the results performed by the AFDSDC. Volume II contains only detailed test data.

The goal of the task was to determine the relative throughput of several different H6000 multiprocessor configurations for different types of workloads. This information can be used by planners when trying to determine the best way to satisfy increasing workload requirements for existing WWMCCS computer systems.

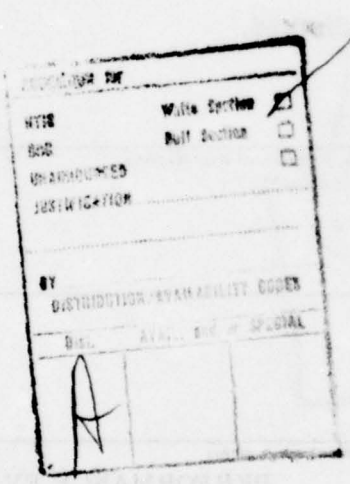


TABLE OF CONTENTS

	Page
VOLUME I	
LIST OF ILLUSTRATIONS	5
LIST OF TABLES	9
EXECUTIVE SUMMARY	11
SECTION I INTRODUCTION	17
1.1 Purpose of the Study	19
1.2 Scope of the Study	19
1.3 Initial Survey Results	20
1.4 Multiprocessors	21
1.5 H6000 Architecture	21
1.5.1 Central Processing Unit (CPU)	24
1.5.2 System Control Unit (SCU)	24
1.5.3 Memory Module (MM)	26
1.5.4 Input/Output Multiplexer (IOM)	27
1.5.5 Memory Interlace	28
SECTION II TECHNICAL APPROACH	31
2.1 Controlled Testing	31
2.1.1 System Environment	33
2.1.1.1 Hardware Configurations	33
2.1.1.2 System Software	36
2.1.1.3 System State	37
2.1.1.4 Operating Procedures	37
2.1.2 Workload Environment	38
2.1.2.1 Statistical Collection File	40
2.1.2.2 Variability of SCF Data	41
2.1.2.3 Site Accounting Data	45

TABLE OF CONTENTS (Continued)

	Page
2.1.2.4 Synthetic Workload Generator	45
2.1.2.5 Synthetic Workloads	50
2.1.2.6 Synthetic Program	53
2.1.2.7 Microscopic Analysis of Synthetic Program	53
2.1.3 Verification	55
2.2 Test Design	56
2.2.1 Test Measurements	56
2.2.1.1 Elapsed Time	56
2.2.1.2 Multiprogramming Depth (MPD)	56
2.2.1.3 IO/CP Ratio	59
2.2.1.4 Resource Utilization	60
2.2.2 Test Instrumentation	60
2.2.2.1 System Accounting Data	61
2.2.2.2 Software Monitors	61
2.2.2.3 Hardware Monitors	64
2.2.2.4 Console Log Sheets	64
2.2.3 Selected Configurations	65
2.2.4 Software Alterations	65
2.2.5 Test Procedures	66
SECTION III TEST RESULTS	71
3.1 Pilot Testing	71
3.1.1 Synthetic Workload Generator	71
3.1.2 Workload Generation and Testing	83
3.1.3 Software Monitors	85
3.1.4 Hardware Monitor	85
3.1.5 Data Reduction	86
3.1.6 Timing Runs	87
3.1.7 Test Procedures	89

TABLE OF CONTENTS (Continued)

	Page
3.2 Primary Testing	89
3.2.1 Test Configurations	89
3.2.2 Test Workloads	91
3.2.3 Test Conduct	103
3.2.4 Test Results	108
3.3 Secondary Testing	110
3.4 Data Analysis	111
3.4.1 Elapsed Time	112
3.4.2 Relative Throughput	142
SECTION IV SUMMARY	147
4.1 Test Conditions	147
4.1.1 Workload Limitations	147
4.1.1.1 Validation of Results	150
4.1.1.2 Workload Characterization Improvements	150
4.1.2 Hardware and Software Configuration Constraints	151
4.1.3 Operational Conditions	152
4.2 Observations	152
4.2.1 Relative Throughput	152
4.2.1.1 Observed Increases in Relative Throughput	153
4.2.1.2 Impact of Memory Interlacing	153
4.2.1.3 Other Observed Influences on Relative Throughput	154
4.2.2 I/O Generated Processor Time	154
4.2.3 Temporary File Purging Delays	155
4.3 Application to WWMCCS ADPE	155

TABLE OF CONTENTS (Concluded)

		Page
APPENDIX I	VERIFICATION STUDY	161
APPENDIX II	SURVEY RESULTS	167
REFERENCES		173
BIBLIOGRAPHY		175

LIST OF ILLUSTRATIONS

Figure Number		Page
VOLUME I		
1-1	Uncoupled Processors	22
1-2	Closely Coupled Processors	23
1-3	H6000 Basic Units	25
2-1	Technical Approach	32
2-2	Possible H6000 Configurations	34
2-3	Standard I/O Configuration	35
2-4	Steps in Analyzing SCF Data	52
2-5	Gantt Representation	57
2-6	Average Multiprogramming Depth	58
2-7	$\overline{\text{MPD}}$ Vs Average Throughput	59
2-8	Test Procedures	67
3-1	Synthetic Workload Multiprogramming Depth	73
3-2	Configuration A	92
3-3	Configuration B	93
3-4	Configuration C	94
3-5	Configuration D	95
3-6	Configuration E	96
3-7	Configuration F	97
3-8	Configuration G	98
3-9	Configuration H	99
3-10	WMPE Equipment Layout	106
3-11	Workload E, CPU Impact	114
3-12	Workload E, Memory Impact	114
3-13	Workload A, CPU Impact	115
3-14	Workload A, Memory Impact	115

LIST OF ILLUSTRATIONS (Continued)

Figure Number		Page
3-15	Workload B, CPU Impact	116
3-16	Workload B, Memory Impact	116
3-17	Workload 2, CPU Impact	117
3-18	Workload 2, Memory Impact	117
3-19	Workload C, CPU Impact	118
3-20	Workload C, Memory Impact	118
3-21	Workload 1, CPU Impact	119
3-22	Workload 1, Memory Impact	119
3-23	Workload D, CPU Impact	120
3-24	Workload D, Memory Impact	120
3-25	Workload 9, CPU Impact	121
3-26	Workload 9, Memory Impact	121
3-27	Workload F(2), CPU Impact	122
3-28	Workload F(2), Memory Impact	122
3-29	Workload F(1), CPU Impact	123
3-30	Workload F(1), Memory Impact	123
3-31	Workload 4, CPU Impact	124
3-32	Workload 4, Memory Impact	124
3-33	Workload 8, CPU Impact	125
3-34	Workload 8, Memory Impact	125
3-35	Workload 10, CPU Impact	126
3-36	Workload 10, Memory Impact	126
3-37	Workload 3, CPU Impact	127
3-38	Workload 3, Memory Impact	127
3-39	Workload E, CPU Impact	128
3-40	Workload E, Memory Impact	128

LIST OF ILLUSTRATIONS (Continued)

Figure Number		Page
3-41	Workload A, CPU Impact	129
3-42	Workload A, Memory Impact	129
3-43	Workload B, CPU Impact	130
3-44	Workload B, Memory Impact	130
3-45	Workload 2, CPU Impact	131
3-46	Workload 2, Memory Impact	131
3-47	Workload C, CPU Impact	132
3-48	Workload C, Memory Impact	132
3-49	Workload 1, CPU Impact	133
3-50	Workload 1, Memory Impact	133
3-51	Workload D, CPU Impact	134
3-52	Workload D, Memory Impact	134
3-53	Workload 9, CPU Impact	135
3-54	Workload 9, Memory Impact	135
3-55	Workload F(2), CPU Impact	136
3-56	Workload F(2), Memory Impact	136
3-57	Workload F(1), CPU Impact	137
3-58	Workload F(1), Memory Impact	137
3-59	Workload 4, CPU Impact	138
3-60	Workload 4, Memory Impact	138
3-61	Workload 8, CPU Impact	139
3-62	Workload 8, Memory Impact	139
3-63	Workload 10, CPU Impact	140
3-64	Workload 10, Memory Impact	140
3-65	Workload 3, CPU Impact	141
3-66	Workload 3, Memory Impact	141

LIST OF ILLUSTRATIONS (Concluded)

Figure Number		Page
4-1	Relative Throughput — H6060 Interleaving On	158
4-2	Relative Throughput — H6060 Interleaving Off	159
4-3	Relative Throughput — H6080 Interleaving On	160
I-1	Validation Data	164

LIST OF TABLES

Table Number		Page
VOLUME I		
1-I	AFWWMCCS ADP Configurations	18
1-II	Memory Interlace Examples	30
2-I	Variation in I/O Time	44
2-II	Site SCF Data	46
3-I	Test Results Calibration Run	74
3-II	Synthetic Workload Generator Calibration Calculations	77
3-III	Initial File Space Analysis	80
3-IV	Test Results Calibration Check	81
3-V	Comparison of Processor Time for Disk and Tape Assignment for File 04	82
3-VI	Test Results Final Calibration	84
3-VII	Configuration Identifiers	90
3-VIII	Workload Characteristics	102
3-IX	Testing Chronology	104
3-X	H6060 Relative Throughput	143
3-XI	H6060 Relative Throughput	144
3-XII	H6080 Relative Throughput	145
I-1	Test Data — Configuration A (Interlace Off)	162
I-2	Test Data — Configuration C (Interlace Off)	162
I-3	Linear Regression of Test Data Rel. Throughput (C to A) vs. IO/CP	165
I-4	Validation Workload Analysis	165

EXECUTIVE SUMMARY

Over the past several years the volume of Automatic Data Processing (ADP) support requirements at WWMCCS installations has been increasing. As a result, WWMCCS ADP planners are being forced to consider alternate ways of meeting these increased processing demands, while still staying within the ADP procurement limitations imposed by the Joint Chiefs of Staff (JCS) under the WWMCCS ADP program.

One possible way of providing increased processing capability within the limits of the WWMCCS ADP program is to add central processing units (CPU's) to the WWMCCS ADP configurations already in the field. However, quantitative information about the degree of improvement to be expected through the addition of CPU's has been unavailable to WWMCCS ADP planners, managers, and executives. Additional quantitative information is required to properly evaluate the relative performance of alternative computer system configurations to permit selection of cost effective configurations. Simple numerical values, such as "two central processing units when coupled together are equivalent to 1.8 central processing units," have been available in the past, but these single-valued estimators of increased throughput (work done per unit of time) overlooked the very important effects of variations between different sets of jobs and until now, included neither empirical data nor relative throughput for configurations with four central processing units.

A central processing unit in a WWMCCS (Honeywell H6000) computer system is a specific part of the hardware dedicated to doing the bulk of the calculations and comparisons required of the system. The inherent design of H6000 systems permits up to four central processing units to be connected in a single system to allow truly

PRECEDING PAGE BLANK-NOT FILMED

simultaneous processing of independent tasks. In some cases adding a central processing unit to a WWMCCS computer system may significantly decrease the total elapsed time needed to complete the processing of a set of jobs; however, in other cases, it may have only a minimal effect. This contrast is caused by the different characteristics of computer workloads. This is quite analogous to changes in miles per gallon of gasoline, where for the same vehicle, varying driving conditions and driver characteristics can have a significant impact on miles per gallon.

The purpose of the WWMCCS Multiprocessor Performance Evaluation task was to collect quantitative data about WWMCCS multiprocessor performance under various workload conditions. This report documents both the development of this performance data for combinations of various multiprocessor configurations and different classes of workloads and how this data can be used by WWMCCS ADP system managers in analyzing alternatives for system enhancement.

In order to provide quantitative information regarding multiprocessor throughput, the Deputy for Air Force WWMCCS, ESD, in response to HQ USAF tasking and with support from AFDSDC, undertook a task to measure throughput for various H6000 multiprocessor configurations.

At the beginning of this task, a survey of many organizations was conducted to determine if the required information existed or if similar efforts were already under way. The survey determined that a comprehensive study identifying throughput values for various H6000 multiprocessor configurations under varying workload conditions did not exist, nor was such an activity in progress.

A direct and pragmatic approach was selected for the WWMCCS Multiprocessor Performance Evaluation. The effort was based on strictly controlled live testing using real computer systems and a carefully selected range of test workloads.

Participants in the workload determination and/or testing were ESD, AFDSDC, CCTC, SAC, TAC, PACOM, ADCOM, and MAC.

Real workloads at WWMCCS sites vary from day-to-day, week-to-week, month-to-month, and site-to-site. The test workloads used in this task were carefully chosen to control certain characteristics, most notably the ratio of time used to do input and output operations for a workload versus the time used to do computation for the same workload. This ratio was varied over a range that covered real operational workloads of WWMCCS sites at MAC, PACOM, SAC, and TAC.

The test plan specified fourteen different test workloads, each having different characteristics and each executed on eight different configurations of H6060 and H6080 computer systems. The number of processors for both the H6060 and H6080 were varied from one to four (larger than any current WWMCCS installation). A total of 273 test cases was conducted and actual elapsed times collected during the tests were used to calculate relative throughput. The elapsed times were normalized to a single processor H6060 in the case of H6060 configurations and to a single processor H6080 in the H6080 test cases. Therefore, increases or decreases in throughput are presented relative to single processor configurations.

A summary of the results of all test cases is provided in Section III of this report. Simple summary graphs useful for estimation of throughput improvements are provided in Section III (Figures 3-10 through 3-66) and Section IV (Figures 4-1, 4-2, and 4-3). These summary graphs can be used to derive an estimate of potential gains in throughput as additional central processing units are added to a configuration. However, the analyst using these summary graphs must determine the characteristics of the current real and/or projected site workload prior to using the graphs. Normally this can be done from system produced information maintained at the site. Once the characteristics of the site workload are determined, the graph

can be used to interpolate between the data collected from the controlled test cases. Using this procedure, a simple estimation of potential throughput for the real workload can be derived. In order to effectively use the summary graphs this entire report must be studied to ensure complete understanding of the test conditions and the characteristics of the test workload.

The results of this effort show that the addition of H6000 central processing units can achieve a large or minimal increase in relative throughput depending on the characteristics of the workload. Observed ranges of relative throughput values, for the conditions tested, are presented in the following tables.

H6060

# of CPU's	Relative Throughput Observed		
	Lowest		Highest
1		1.00	
2	1.30	to	2.06
3	1.39	to	2.83
4	1.31	to	3.54

H6080

# of CPU's	Relative Throughput Observed		
	Lowest		Highest
1		1.00	
2	0.98	to	2.02
3	1.14	to	2.97
4	1.16	to	3.82

Single CPU H6060 and H6080 configurations were used as the baseline for comparison purposes. The ranges presented in the above table for multiple CPU configurations represent how the test workloads performed on multiple CPU configurations when compared to their performance on a single CPU configuration. The low number in a range typically relates to how well an I/O-intensive workload performed, and the high number relates to how well a CPU-intensive workload performed. For example, executing on a two CPU H6060 system, one of the I/O-intensive workloads ran about one-third faster with two CPU's than with one, and one of the CPU-intensive workloads ran about twice as fast. Simply stated, a performance improvement ranging from approximately 1.30 to 2.06 can be expected when upgrading from a single CPU H6060 to a two CPU H6060 depending on whether the workload is highly I/O or CPU intensive. The results of this task show the utilization of memory interlacing provided a relative throughput increase of approximately 5%.

The key factor in using the results of this report to evaluate various configuration possibilities at a specific site is a rigorous characterization of the workload.

This quantitiative information will be valuable to WWMCCS system analysts and system planners when contemplating system expansion to meet increased workload situations. Careful use of the results of this testing effort could prevent costly errors associated with attempts to solve increased workload problems.

Using workloads provided by TAC, a validation study of this technique was performed by AFDSDC. Based on the characteristics of the TAC workloads, a projection was made using the results of this task to predict relative throughput of TAC workloads on selected equipment configurations. Then, the workloads were actually executed on these configurations with resulting measured throughput values comparing very well with the predicted values. This is presented separately as Appendix I.

SECTION I

INTRODUCTION

In October 1971, the Honeywell Information Systems (HIS) Inc. Series 6000 computer system was selected as the standard automatic data processing (ADP) equipment for the World Wide Military Command and Control System (WWMCCS). The WWMCCS ADP is managed by the Joint Chiefs of Staff and includes 35 medium to large computer systems and their associated remote terminals operating in various DoD agencies. The configurations are geographically distributed over 18 time zones from Taiwan to Germany. Their purpose is to support planning for the employment of and command control of all of our armed forces by the President of the United States, the Secretary of Defense, the Joint Chiefs of Staff, and the major commanders in the field.

The hardware cost of the 35 systems installed and operating in various DoD agencies is in excess of $100 million. The Air Force has 17 systems installed at 10 different facilities with a hardware cost of approximately $50 million. The hardware configuration at each of these sites is briefly described in Table 1-I.

The WWMCCS ADP Contract (F19628-71-R-003) specified three system types as follows:

a. General Staff Support - Medium (GSS/M)

b. General Staff Support - Large (GSS/L)

c. Force Control (FC)

Within each system type, there are many possible configurations of central processor units (CPU's), memory and peripherals. All three system types allow up to four CPU's, but not more than two processors have been interconnected to date in an operational situation.

PRECEDING PAGE BLANK-NOT FILMED

Table 1-I

AFWWMCCS ADP Configurations*

DoD AGENCY	SYSTEM TYPE	NUMBER	CPU TYPE	CORE SIZE
1. ADCOM	FC	3 ea.	HIS 6080 (DUAL)	256K
	GSS/M	1 ea.	HIS 6060	----
2. AFDSC	GSS/M	1 ea.	HIS 6060	256K
3. TAC	GSS/M	1 ea.	HIS 6060	384K
4. AFSC	GSS/M	1 ea.	HIS 6060	320K
5. REDCOM	GSS/M	1 ea.	HIS 6060	256K
6. AU/AFDSDC	GSS/M	1 ea.	HIS 6060	192K
7. SAC	FC	2 ea.	HIS 6080 (DUAL)	384K
8. MAC	FC	1 ea.	HIS 6080 (DUAL)	512K
	FC/2	1 ea.	HIS 6080	192K
	GSS/M	2 ea.	HIS 6060	256K
9. PACAF	GSS/M	1 ea.	HIS 6060	192K
10. USAFE	GSS/M	1 ea.	HIS 6060	256K

*As of September 1975.

1.1 Purpose of the Study

The workload at many of the WWMCCS sites has been steadily increasing as mission requirements expand, and it is reasonable to expect that this trend will continue at least in the near future. It is expected that the WWMCCS sites will be equipped with HIS 6000's for the next five to eight years. The need to accommodate increased workloads has caused some organizations to identify alternate ways of obtaining more throughput (i.e., work done per unit time) from existing WWMCCS systems. One strong candidate is to increase the number of central processing units within a system. However, there were unanswered questions about how much more work dual-, triple-, and quadruple-processor systems could accomplish compared to a single-processor system. In order to provide quantitative information regarding multiprocessor throughput, the Deputy for Air Force WWMCCS (ESD/WW), in response to HQ USAF tasking and with the cooperation of the AFDSDC, undertook a task to measure throughput for various H6000 multiprocessor configurations.

The purpose of the study was to identify potential throughput improvements for selected H6000 multiprocessor configurations and to present this information in a simple format directly usable by ADP managers at WWMCCS sites. Throughput improvements could then be compared with relative cost (an H6060 central processing unit costs approximately $285,000 and an H6080 central processing unit costs approximately $325,000) and the result of the comparison used in the decision making process.

1.2 Scope of the Study

The scope of this study was purposely constrained in order to provide near-term results. The study deals with throughput -- one of several measures of effectiveness for any computer system; other measures such as response time and turn-around time are explitcly

excluded. Some measures such as processor utilization, memory utilization, and multiprogramming depth were collected as part of the test plan but were not the focus of the study.

Moreover, the scope of the study was constrained to the investigation of repeatable performance parameters collected via direct measurement from live tests. Only available tools were to be employed during the tests and developmental efforts were not to be required. Furthermore, the study drew upon previous work by other organizations to the maximum extent possible and involved operational WWMCCS commands such as MAC, PACOM, SAC, and TAC in the production of the test workloads. A separate verification of the study results was performed.

1.3 Initial Survey Results

Prior to beginning the test efforts, other activities related to WWMCCS multiprocessor performance evaluation were surveyed to determine if quantitative information on H6000 multiprocessor configurations was available. The conclusion reached was that some, but not all, of the needed information was available.

The following organizations were contacted during this survey:

a. System Development Corporation (SDC)

b. North American Air Defense Command (NORAD)

c. Pacific Command (PACOM)

d. DCA's Command and Control Technical Center (CCTC) formerly the Joint Technical Support Activity (JTSA)

e. Federal Computer Performance Evaluation and Simulation Center (FEDSIM)

f. Honeywell Information Systems Inc. (HIS)

g. Air Force Data Systems Design Center (AFDSDC)

h. Naval Command Systems Support Activity (NAVCOSSACT)

i. Navy Automatic Data Processing Evaluation Support Office (ADPESO)

j. Electronic Systems Division/Directorate of Computer Systems Engineering (ESD/MCI)

In the interest of completeness, the survey results are presented in Appendix II with the understanding that the comments or positions of the various agencies are not official and that changes in the status of some projects may have occurred since the survey period of August - September 1975.

1.4 Multiprocessors

Before attempting to discuss the relative throughput of WWMCCS H6000 multiprocessors (read multiple processors), a few words of background are appropriate. Multiprocessor, particularly as used throughout this report, is meant to describe a class of computer systems where two or more central processing units (CPU's) can operate with true simultaneity on two or more different programs (or two or more different parts of one program) and all CPU's have direct access to any portion of a common storage facility (e.g., core memory). Figures 1-1 and 1-2 illustrate this schematically. Figure 1-1 shows a pair of computer systems with completely uncoupled processors that can perform multiprogramming but cannot perform multiprocessing in the strict sense of the term. While the two systems shown could operate simultaneously on two different programs, they cannot have access to each other's memory without human intervention. Figure 1-2 shows this pair of computer systems reconnected as one multiprocessor system.

1.5 H6000 Architecture

This section provides a brief introduction to the basic units that comprise the H6000 series hardware architecture and establishes

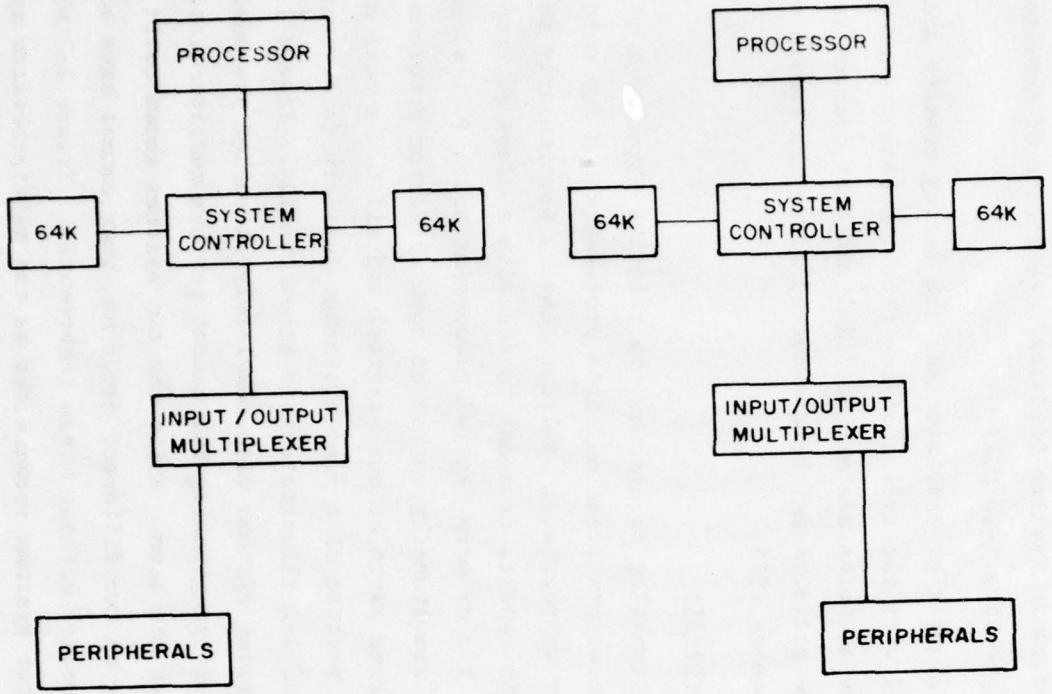


Figure 1-1. Uncoupled Processors

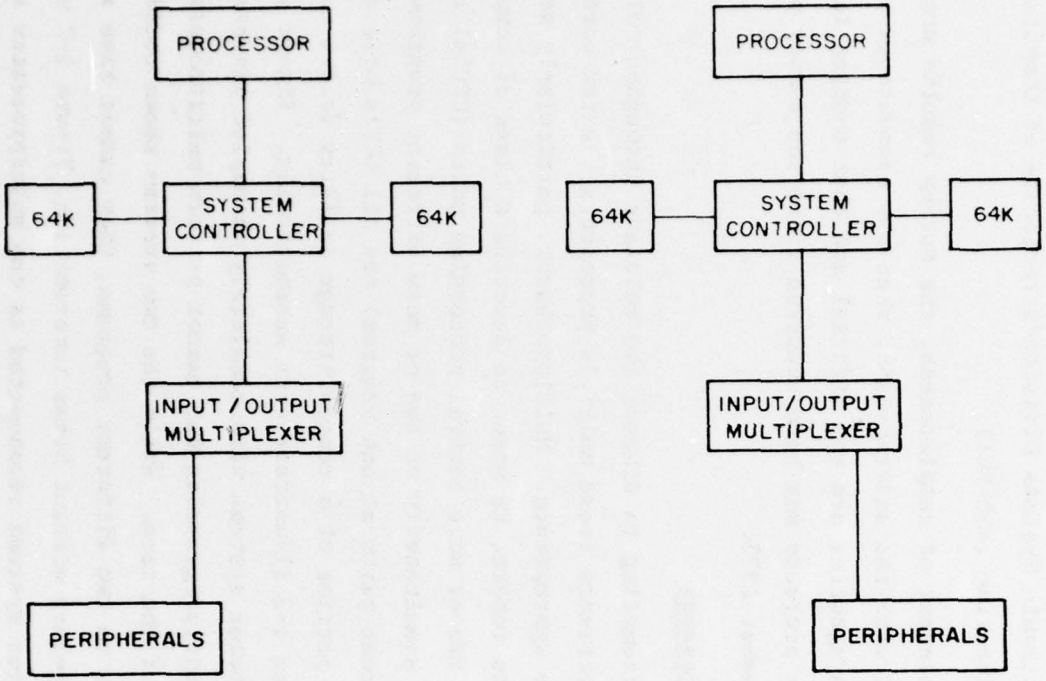


Figure 1-1. Uncoupled Processors

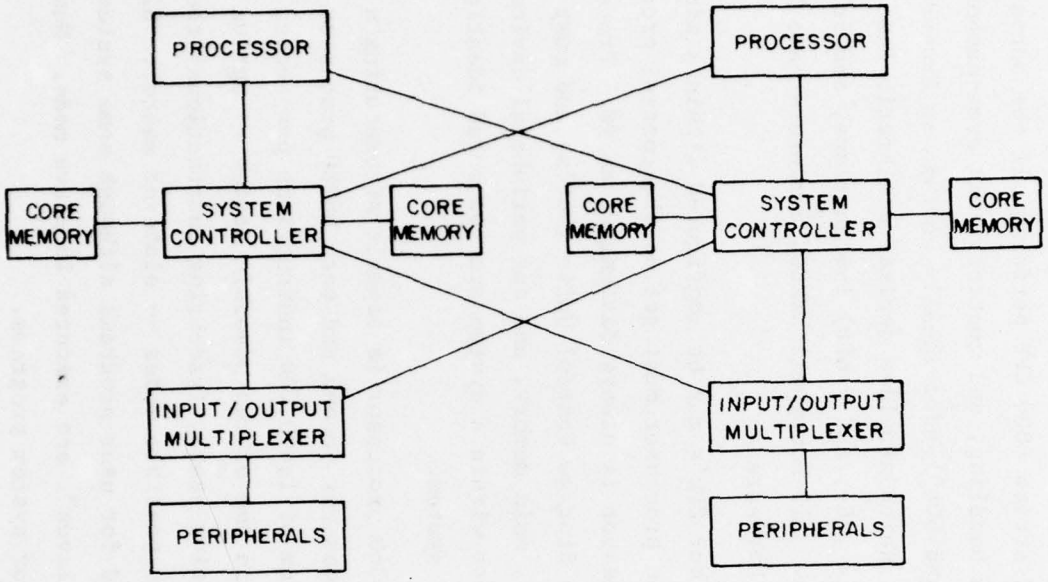


Figure 1-2. Closely Coupled Processors

a basis for later discussions. Figure 1-3 graphically portrays the basic units and their connectivity. More detailed explanations of the hardware components are provided in References [1] and [2].

1.5.1 Central Processing Unit (CPU)

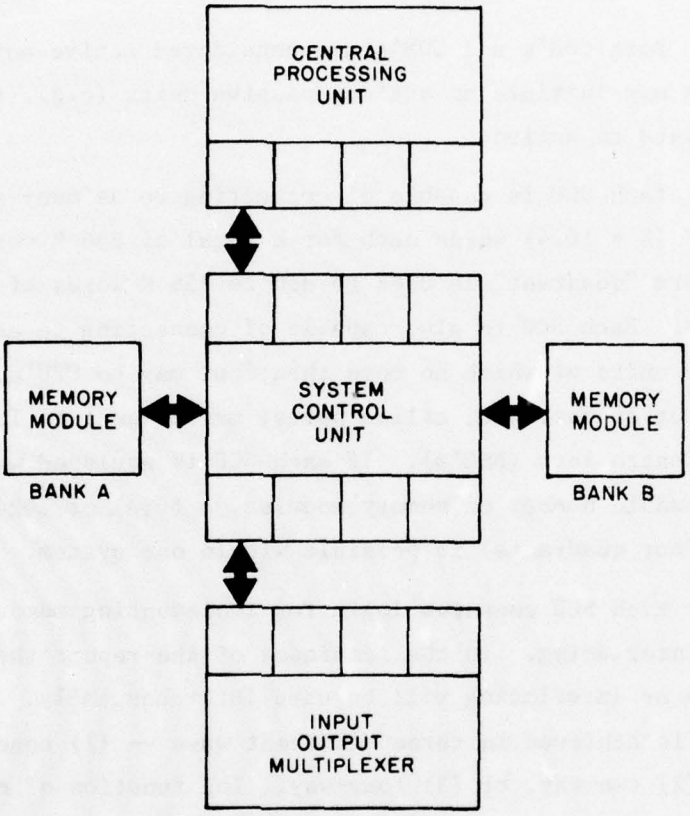
The HIS series 6000 CPU performs all the normal functions of computation, data handling, and control. All even-numbered systems (i.e., 6040, 6060, and 6080) additionally contain an Extended Instruction Set (EIS) unit, which provides decimal arithmetic instructions, character-oriented (4-, 6-, and 9-bit) instructions, and related control instructions. All Honeywell WWMCCS computers have been modified to include EIS hardware.

Up to four CPU's may be configured within a single H6060 or H6080 system. One processor must act as the control processor and the control processor is always designated as PØ. Processors can only be connected to Storage Control Units (SCU's) and must communicate with each other, main memory, and any peripheral device via the SCU's. All processors within a system are connected identically to all SCU's within the system.

Each H6060 processor is capable of executing a maximum of 550,000 instructions per second and each H6080 processor is capable of executing a maximum of 1,400,000 instructions per second. The processor is always in one of two possible states -- active or idle. When it is in the active state, (executing instructions) the processor must be in one of two possible modes -- slave or master. Slave mode is generally reserved for user programs although some system programs called "privileged slaves", are executed in slave mode. Master mode is always reserved for system programs.

1.5.2 System Control Unit (SCU)

The SCU controls access to all of main memory, whether ferrite core or metal-oxide semiconductor (MOS), from all active



IA-48,990

Figure 1-3. H6000 Basic Units

units within the system. Consequently, all memory modules (MM's), all CPU's, and all Input-Output Multiplexers (IOM's) must be connected to at least one SCU. In fact, all CPU's and all IOM's connect to all SCU's within a system.

Both CPU's and IOM's are considered active units. Active units may initiate an action; passive units (e.g., the SCU) cannot initiate an action.

Each SCU is capable of connecting to as many as four MM's of 64 K (K = 1024) words each for a total of 256 K words per SCU. The term "quadrant" is used to denote 256 K words of continguous memory. Each SCU is also capable of connecting to as many as eight active units of which no more than four may be CPU's. The remaining four connections, called ports, may connect to IOM's or to Bulk Store Controllers (BSC's). If each SCU is equipped with the maximum allowable number of memory modules, a total of 1024 K words of memory (four quadrants) is possible within one system.

Each SCU contains logic for implementing memory interleaving or interlacing. In the remainder of the report the term interleaving or interlacing will be used interchangeably. Memory interlacing is achieved in three different ways -- (1) none (interlace off), (2) two-way, or (3) four-way. The function of memory interlacing is sometimes split between SCU's and active units. The entire memory interlacing logic will be discussed in Section 1.5.5.

1.5.3 Memory Module (MM)

The memory modules provide high-speed storage of instructions and data for use by the CPU's and IOM's. Memory modules are connected directly to SCU's and can only be connected to SCU's. Therefore, all memory accesses are controlled by an SCU regardless of the source of the access request. Each MM consists of either 64 K or

128 K words of either ferrite core or MOS memory. In all cases, the memory is functionally equivalent except that for H6060's the memory cycle speed is 1.2 microseconds and for H6080's the memory cycle speed is 0.5 microsecond.*

There are two different memory module capacities available -- 64 K and 128 K words. Each memory module is composed of 32,768 (or 65,536) read (destroy)/write (restore) memory registers that are 74 bits wide. Two of these bits are parity bits; thus, each memory access provides two 36-bit computer words with consecutive addresses.

Any system configured with more than 256 K words of memory requires an extended memory option to increase the size of the address space. All WWMCCS H6000's have been equipped with the extended memory option.

1.5.4 Input/Output Multiplexer (IOM)

The IOM is an active unit and provides connections from all peripheral devices to the SCU and subsequently to the memory and CPU's. Because the IOM can operate without direct control by the CPU, a high degree of concurrency of CPU and I/O operations can be achieved.

As many as four IOM's may be configured in any one system. Each IOM consists of 24 useable data channels and can, therefore, control multiple I/O commands to a device and several simultaneous data transfers between memory and peripheral devices. The maximum data transfer rate of an IOM in an H6060 system is 3.75 million characters per second and 6 million characters per second in an H6080 system.

All IOM's in a system can be connected to all SCU's in a system (e.g., the IOM has four SCU connection ports). The IOM's are

---

*The new High-Performance MOS will operate at 0.75 microsecond in both machines.

always connected to the highest priority ports on the SCU (ports 0, 1, 2, and 3) since most peripheral devices must operate at fixed rates once the I/O operation has started.

1.5.5 Memory Interlace

Closely associated with the subject of memory interlacing is the subject of memory contention. Memory contention is inherent in any shared-memory multiprocessor system such as the H6000. Whenever more than one active unit (e.g., two processors, or one processor and one IOM) simultaneously require access to any of the memory locations within the same memory module, interference (memory contention) occurs. To resolve these conflicts, the eight SCU ports have inherent priorities relative to each other (unless disabled by switch setting). When configuring a system, IOM's are assigned top priority ports, due to the undesirability of causing loss of data once an I/O operation has begun. BSC's, if configured, are assigned next highest priority because they also transfer I/O data, but can tolerate delay. Processors are assigned the lowest priority because they are inherently asynchronous and can generally time-share memory with IOM's and BSC's. The higher-priority unit will be allowed to complete a double-word memory access before a lower-priority unit is allowed to start an access.

Memory interlacing allows for the spreading of consecutive memory address-pairs across more than one physical memory module. Both two- and four-way interleaving are provided. When there is no interlacing (i.e., all interlace switches turned off), then all memory locations are sequentially addressed throughout the total complement of memory modules attached to the system. Since programs generally reference memory in a pattern of consecutive addresses and since one memory module of 64 K words is generally more than one program would utilize, contention or queuing for access to the memory module is likely.

When two-way interlacing is selected, switches on each of the SCU's are set to enable the SCU interlace logic to distribute consecutive address-pairs to alternate memory banks. Each SCU is capable of controlling two memory banks (A&B) and each memory bank may consist of as many as 128 K words (2 memory modules of the 64 K words variety). Since the logic is active both when the program is initially loaded and also during all subsequent references, unless the system is restarted with the interlace switches off, the address translation (alteration) mapping is consistent and the proper reference is accomplished. Two-way interlacing of this form requires that the address space of both memory banks attached to one SCU be identical.

Another form of two-way interlacing is possible even if the address space for each memory bank is not the same. To achieve this form of two-way memory interlace, the interlace switches on the SCUs must be off and there must be an even number of SCU's attached to the system. This form utilizes memory interlace switches on each of the active units and enables logic which directs two consecutive address-pairs (4 words) to alternate SCU's within the system.

Four-way interlacing is simply a combination of both types of two-way interlacing. Table 1-II provides an example of address mapping for all four interlace possibilities. The example used is for a program or group of programs consecutively referencing sixteen contiguous logical addresses in a two-SCU configuration with two banks of 64 K words on each SCU. The SCU's are denoted as 0 and 1 and the memory banks are denoted as A and B.

Table 1-II

Memory Interlace Examples

INTERLACE "OFF":

LOGICAL ADDRESS 0	SCU-0, MEMORY BANK A, PHYSICAL ADDRESS 0
" " 1	" 0, " " A, " " 1
" " 2	" 0, " " A, " " 2
" " 3	" 0, " " A, " " 3
" " 4	" 0, " " A, " " 4
" " 5	" 0, " " A, " " 5
" " 6	" 0, " " A, " " 6
" " 7	" 0, " " A, " " 7
" " 8	" 0, " " A, " " 8
" " 9	" 0, " " A, " " 9
" " 10	" 0, " " A, " " 10
" " 11	" 0, " " A, " " 11
" " 12	" 0, " " A, " " 12
" " 13	" 0, " " A, " " 13
" " 14	" 0, " " A, " " 14
" " 15	" 0 " " A " " 15

SCU INTERLACE "ON" ONLY (2-WAY)

LOGICAL ADDRESS 0	SCU 0, MEMORY BANK A, PHYSICAL ADDRESS 0
" " 1	" 0, " A, " " 1
" " 2	" 0, " B, " " 0
" " 3	" 0, " B, " " 1
" " 4	" 0, " A, " " 2
" " 5	" 0, " A, " " 3
" " 6	" 0, " B, " " 2
" " 7	" 0, " B, " " 3
" " 8	" 0, " A, " " 4
" " 9	" 0, " A, " " 5
" " 10	" 0, " B, " " 4
" " 11	" 0, " B, " " 5
" " 12	" 0, " A, " " 6
" " 13	" 0, " A, " " 7
" " 14	" 0, " B, " " 6
" " 15	" 0 " B " " 7

ACTIVE INTERLACE "ON" ONLY (2-WAY):

LOGICAL ADDRESS 0	SCU 0, MEMORY BANK A, PHYSICAL ADDRESS 0
" " 1	" 0, " " A, " " 1
" " 2	" 0, " " A, " " 2
" " 3	" 0, " " A, " " 3
" " 4	" 1, " " A, " " 0
" " 5	" 1, " " A, " " 1
" " 6	" 1, " " A, " " 2
" " 7	" 1, " " A, " " 3
" " 8	" 0, " " A, " " 4
" " 9	" 0, " " A, " " 5
" " 10	" 0, " " A, " " 6
" " 11	" 0, " " A, " " 7
" " 12	" 1, " " A, " " 4
" " 13	" 1, " " A, " " 5
" " 14	" 1, " " A, " " 6
" " 15	" 1 " " A " " 7

ALL INTERLACE "ON" (4-WAY):

LOGICAL ADDRESS 0	SCU 0, MEMORY BANK A, PHYSICAL ADDRESS 0
" " 1	" 0, " " A, " " 1
" " 2	" 0, " " B, " " 0
" " 3	" 0, " " B, " " 1
" " 4	" 1, " " A, " " 0
" " 5	" 1, " " A, " " 1
" " 6	" 1, " " B, " " 0
" " 7	" 1, " " B, " " 1
" " 8	" 0, " " A, " " 2
" " 9	" 0, " " A, " " 3
" " 10	" 0, " " B, " " 2
" " 11	" 0, " " B, " " 3
" " 12	" 1, " " A, " " 2
" " 13	" 1, " " A, " " 3
" " 14	" 1, " " B, " " 2
" " 15	" 1 " " B " " 3

SECTION II

TECHNICAL APPROACH

The technical approach selected for the WWMCCS Multiprocessor Performance Evaluation study was empirical testing in a strictly controlled environment. The organization of the technical approach, as presented in this section, is depicted in Figure 2-1. As in all computer performance evaluation studies, the environment consists of a system or systems to be tested and a workload or workloads to be executed by the system for measurement purposes.

The first phase of the technical approach involved planning a controlled workload set and selecting a controlled system set. Each of these areas is further described in the following paragraphs.

2.1 Controlled Testing

In order to determine the relative throughput ranges of HIS 6000 multiprocessor configurations and the interdependencies between system and workload parameters, a series of controlled tests was designed.

Prior work [3,4] applying the methods of controlled testing (or design of experiments) to computer performance evaluation studies has helped isolate factors influencing performance. Such techniques constrain exogenous interactions which might impact the experimental data collected during a test. This enables the systematic examination of variations in the computer system parameters (e.g., resource utilizations) as a reaction to a set of workload demands.

In accordance with the preliminary experiment design, a system environment and a workload environment were defined as the two elements of the controlled testing.

BEST AVAILABLE COPY

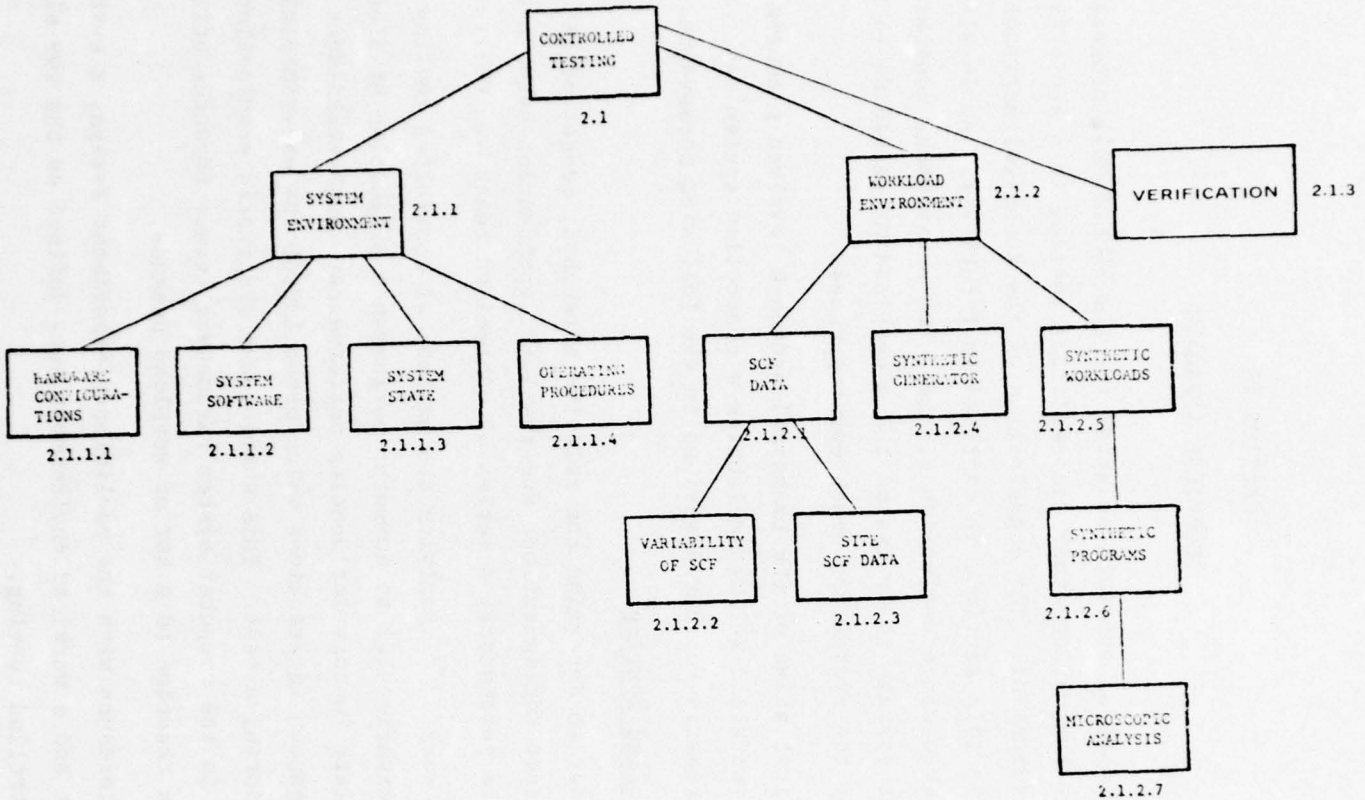


Figure 2-1. Technical Approach

2.1.1 System Environment

The system environment was defined as four distinct segments -- 1) the hardware configurations, 2) the system software, 3) the system state, and 4) the operating procedures for the experimental testing. The control mechanisms for each of the segments are discussed below.

2.1.1.1 Hardware Configurations

Hardware configurations to be used during the tests focused on the major hardware components, namely the number of central processing units (CPU's), system control units (SCU's) and memory modules (MM's). A total of 144 combinations of CPU's, SCU's, and MM's is possible within the design constraints of the H6000 for both H6060's and H6080's. Many of the 144 combinations were deemed unrealistic and were eliminated from further consideration. For example, 1 CPU with 1024K words of core memory and 4 CPU's with 128K words of core memory are both combinations that were not selected because of system resource imbalance. Figure 2-2 illustrates the full range of selection alternatives and the eight combinations chosen for testing. The configurations selected provide a full range of CPU's (1 to 4), a full range of SCU's (1 to 4), and a range of memory modules (4, 6, 8, 12, 16). It was established that the eight configurations chosen would be tested both as H6060 and H6080 systems -- for a total of 16 test configurations. It was also decided to test each configuration with memory interlace and without memory interlace.

A standard I/O subsystem was selected for use in all test cases. As shown in Figure 2-3, this subsystem consisted of two IOM's (each with 2 physical channels and 4 logical channels) and six DSS-190 disk units crossbarred such that any of the six disk units could be accessed from either of two MPC's. Also, tape drives, a printer, and a card reader were included for operation of the system.

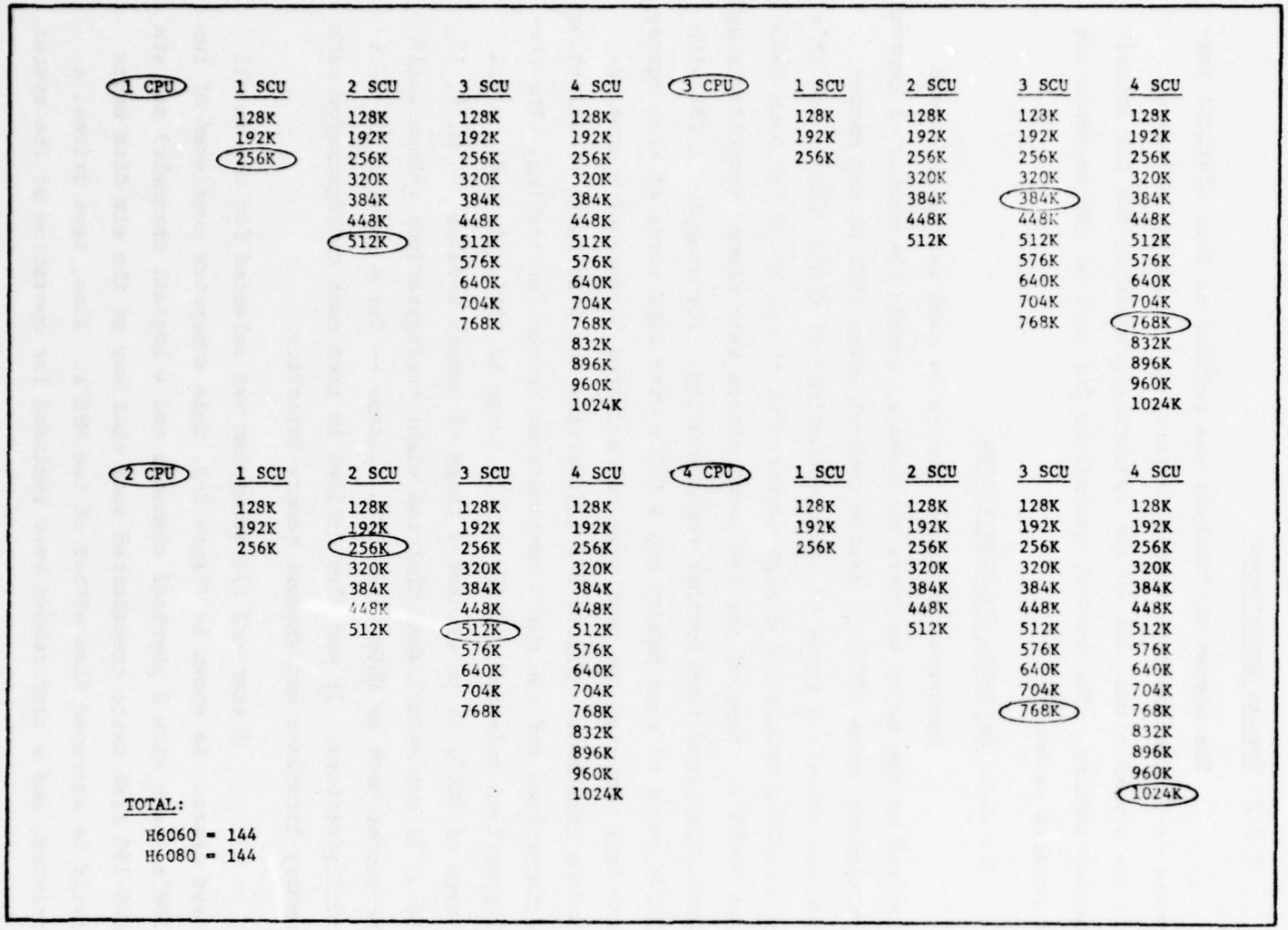


Figure 2-2. Possible H6000 Configurations

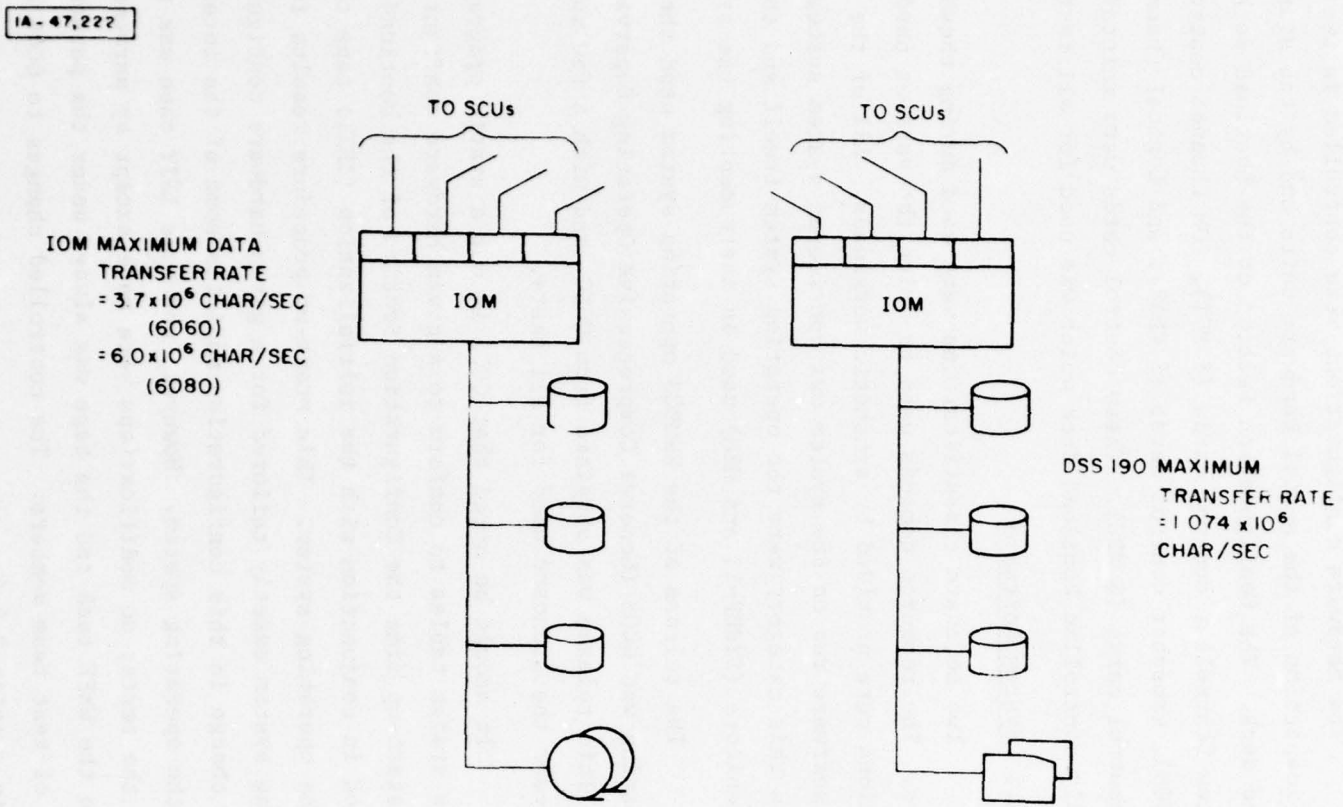


Figure 2-3. Standard I/O Configuration

The hardware configurations were controlled in two ways; by visual inspection of the actual hardware units and by the start-up or bootload deck. The Configuration Section of the bootload deck contains System Controller control cards ($ MCT), IOM Channel control cards ($ IOM), Crossbar control cards ($ XBAR), and Logical Channel Grouping control cards ($ MPC). These control cards were maintained in a rigidly controlled bootload deck which was used for all testing.

2.1.1.2 System Software

Two software classifications were used during these experiments. The resource demands used to drive the various hardware configurations were provided by synthetic workloads. All of the remaining software run on the system was considered system software. Included in this category were the operating system itself and the software monitors (SYRUP-II and MSM) used in instrumenting the system.

The release of the WWMCCS operating system used throughout the testing was GCOS (General Comprehensive Operating Supervisor) WW 6.2.1. This release was obtained from CCTC, and with a few modifications became the standard used for all tests.

It should be noted that GCOS is not a static program, but changes system tables to conform to a given hardware configuration. At start-up time the Configuration Section of the bootload deck is used in conjunction with the initialization (INIT) tape to generate the operating system. This start-up procedure results in an operating system exactly tailored for a given hardware configuration. Any change in this configuration modifies some of the internal tables of the operating system. However, the same INIT tape was used throughout the tests; no modifications were made except by personnel assigned to the WMPE task and the tape was always under the personal cognizance of test team members. The controlled changes to GCOS are discussed in section 2.2.4.

The two software monitors used in the tests were programs specifically designed to collect user and system oriented data. They are explained in sections 2.2.2.2 and 2.2.2.3 respectively of this report. The software monitors were controlled in a manner similar to the operating system. The collector portions were stored on the system in a permanent file. A permanent file save tape (PERMSAVE) was used to restore the software monitors whenever the system was loaded from scratch. The PERMSAVE tape was controlled in the same way as the INIT tape. Control cards that enable the monitors to become part of the operating system were included in the Patch Section of the bootload deck which was controlled as described above.

2.1.1.3 System State

Describing the current condition of a computer system can be accomplished using the concept of system state. The various components of a computer system, during the system's operation, experience changes in their status. Some of these changes may be relatively long, such as an I/O operation; or emphemeral, such as the change in the program counter.

The initial system state for the tests described in this report established the beginning system environment and the point at which system loading began. By rigidly controlling this initial system state, i.e., having the system completely quiescent, the system was isolated from other than specific test demands.

2.1.1.4 Operating Procedures

Because operator interactions are key factors in the environment of a computer system, carefully defined operating procedures were established and followed. For the HIS 6000 system such procedures are given in the operator's manual[5] which describes the System Startup and Initialization procedures as well as other miscellaneous operating techniques.

During the tests only standard Honeywell and special test procedures were used. Section 2.2.5 describes the special test procedures designed specifically for this task.

2.1.2 Workload Environment

A workload is a collection of all the jobs processed by a computer system in a finite period of time. An actual workload is the collection of jobs processed by an existing computer system during its normal operation. A subset of the actual workload is generally used in conducting controlled experiments with a computer system and is referred to as the test or benchmark workload.

There are several methods of constructing the test workload. In one method (the benchmark method), the performance analyst and the installation manager select a set of user jobs that appear to represent the user workload. The criteria for selection are usually the frequency of occurrence and the magnitude of demands on the system resources. The advantage of this method is that it uses real jobs and consequently is viewed as a real workload even though only a subset of the total real workload is used. The disadvantages of the benchmark method are: (a) the test workload is not flexible as it is constructed from jobs with fixed characteristics, (b) duplication of large amounts of data on auxiliary storage is expensive, and (c) security and privacy considerations may prevent the use of some jobs.

In another method (the synthetic workload method), the test workload consists of a set of synthetic programs constructed to place known demands on the major resources of the computer system. The synthetic workload is a collection of individual user jobs. Synthetic jobs may be conveniently employed to construct test workloads that are flexible and reproducible. A synthetic job is essentially a program with control parameters built into it. By varying the

values of these parameters, the magnitude of the demands placed on system resources may be varied. The parameters may be selected manually after analysis of the characteristics of the real workload or may be generated automatically based on the system-collected accounting data. The advantage of constructing test workloads based on resource demands is that the magnitude of the resource demands is readily available from the accounting package integral to the operating system for billing purposes. The disadvantages of this method are that (a) any inaccuracies in the accounting data are reflected in the generation of the parameters for the synthetic job and (b) the accounting data is machine-dependent.

A synthetic test workload can also be constructed from application programs selected using functional units as the criterion. In this machine-independent description, the workload is classified based on the nature of processing; for example, compilation, sort-merge, matrix-inversion. The disadvantage of this method, called the kernel approach, is the difficulty in obtaining information about the nature of processing. The advantage of this approach is that is is machine-independent and therefore may be employed for competitive benchmarking.

The choice of the method is influenced by the specific application at hand, the availability and the level of detail of the accounting data, and the aims of the performance study. The success of the method selected is assessed by the following factors: the ease with which the accounting data is reduced to determine the test workload characteristics, the flexibility incorporated into the test workload, and the extent to which the experimental results obtained from the test workload are applicable to real life situations.

For the WWMCCS Multiprocessor Performance Evaluation task, a workload environment consisting of synthetic workloads based on accounting data from WWMCCS operational sites was selected as the

most reasonable method of generating a number of repeatable and controlled workloads for throughput analysis.

2.1.2.1 Statistical Collection File

The Statistical Collection File (SCF) maintains the accounting data for GCOS in the HIS 6000 series computer systems. There are 18 different types of SCF records that collect information for both batch and interactive jobs. The details of these several types of records appear in the Honeywell manual on Summary Edit Program[6]. In this study only batch jobs were selected for analysis and hence only Type 1 records were used to obtain information about the workload characteristics. The quantities of interest for this study were:

1. Processor Time
2. I/O time
3. Core size
4. Job start time
5. Job finish time

The first three quantities may be used in characterizing user jobs. The last two quantities are useful in calculating the elapsed time for either a job or a workload; they are also useful in determining the degree of multiprogramming and the construction of Gantt charts (see section 2.2.1.2 for further discussion).

SCF data is used as input to the GCOS Summary Edit Program (GSEP) that prints out a summary report of resource utilization. The summary can be specified by record types or can be for all types of records collected during a session. The GSEP summary report may be used to obtain the following quantities for each test workload:

1. Total processor time
2. Total I/O time
3. Ratio of I/O time to processor time (IO/CP)
4. Elapsed time for the workload

Based on the SCF data, test workloads can be constructed for use in controlled experiments. However, the use of SCF data has at least two known limitations. First, SCF data can be considered static by nature because the variation of resource demands with time is not recorded. The information is collected only after the fact. The dynamic interaction among the active jobs and their collective interaction with the system is not captured by SCF. The variations of resource utilizations with time are considered dynamic data; this data can be collected by software and hardware monitors and can shed some light on the nature of interaction among the active jobs and their collective interaction with the system. Second, the SCF data has variance; the same job executed on the same system does not usually produce identical SCF data.

2.1.2.2 Variability of SCF Data

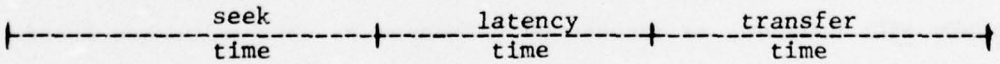
In this study, SCF data was the source used for constructing the test workloads; hence, it is relevant to investigate the accuracy of data recorded by SCF. In the WWMCCS community it is felt that the accuracy of the data recorded by SCF is approximately $\pm$5%; however, no documented reference to this value was found.

In this study, variability refers to the amount of variation in the processor and I/O time recorded by SCF for the same job when the job is run several times on the same dedicated system. While this would appear to provide identical conditions which should produce identical SCF data, the conditions actually vary within the computer system from run-to-run due, primarily, to variations in I/O access which, in turn, influences other scheduling algorithms. The

variation may be attributed to two sources; the first is the accuracy with which SCF collects the data and the second is an actual change in the quantity measured. In this section the term variability refers to the sum of these two quantities.

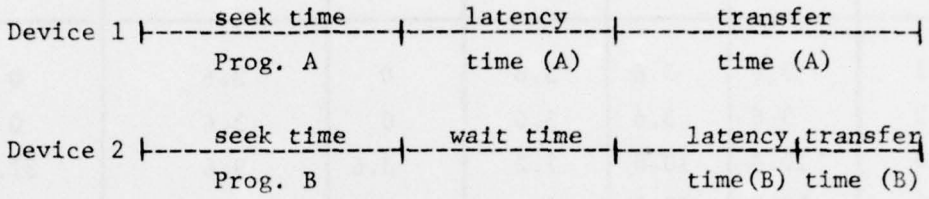
I/O time for an I/O initiation is the sum of seek time, latency time, data transfer time, and wait time. The data transfer time is not subject to large variation because it depends only on the size of the buffer and the speed characteristics of the channel and device. The wait time depends on the number of jobs in the mix (the multiprogramming depth), the type of processing being performed by the individual jobs in the mix, and the number of available channels and devices. The seek time depends on such quantities as the number of files allocated to the physical device, the size of the files, and the number of jobs in contention for the particular device. The latency time is the rotational delay caused when the read/write head is waiting for the beginning of the record. The wait and seek time are largely responsible for the variation in the I/O time.

The following examples illustrate the process described above. The first case involves one program performing I/O to one physical device. The timing diagram (not to scale) is shown below.



This is a simple case of uniprogramming, and hence, no wait is involved. The time per I/O initiation is equal to the time the device and channel are effectively busy.

The second example involves two programs performing I/O simultaneously to two devices connected to a single channel. The timing diagram (not to scale) can be represented as follows:



In this case, program B has to wait for the availability of the channel before it can transfer data. The I/O time for program B includes the wait time and, as a consequence, the measured value of I/O time will not be equal to the time the device and channel are effectively busy. As the number of programs increases the contention for the channel increases with corresponding increases in wait time. When two or more programs are accessing the same device, there is a further increase in wait time due to device contention; as a result, their seeks cannot be overlapped since they are to the same device.

The I/O time recorded by SCF is the cumulative sum of times required to perform all the I/O initiations requested by a job. As briefly described above, the I/O time is made up of statistical quantities and varies every time there is a change in the environment in which the job is running. Such a variation was observed during our experiments at CCTC, Reston, Va. The experimental results are presented in Table 2-I and show that the range in variation is from 0% to 37.5%. No variation occurred in 15 of the 26 values, probably because of the small values of I/O time involved.

A study was conducted at the Air Force Data Systems Design Center (AFDSDC) to quantify the variability of key SCF

Table 2-I

Variation of I/O Time

SNUMB	I/O TIME (T) in Seconds			$\Delta = T_{MAX} - T_{MIN}$	Ave. of $T = \overline{T}$	Percent Variation $100\ \Delta / \overline{T}$
	Run 1	Run 2	Run 3			
1	3.6	3.6	3.6	0	3.6	0
2	3.6	3.6	3.6	0	3.6	0
3	10.8	10.8	7.2	3.6	9.6	37.5
4	18.0	18.0	14.4	3.6	16.8	21.4
5	21.6	25.2	21.6	3.6	22.8	15.8
6	28.8	39.6	36.0	10.8	34.8	31.0
7	32.4	39.6	43.2	10.8	38.4	28.1
8	43.2	50.4	57.6	14.4	50.4	28.6
9	133.2	154.8	169.2	36.0	152.4	23.6
10	298.8	331.2	324.0	32.4	318.0	10.2
11	550.8	576.0	579.6	28.8	568.8	5.1
12	802.8	853.2	748.8	104.4	801.6	13.0
13	1393.2	1411.2	1746.0	352.8	1516.8	23.3
14	3.6	3.6	3.6	0	3.6	0
15	3.6	3.6	3.6	0	3.6	0
16	3.6	3.6	3.6	0	3.6	0
17	3.6	3.6	3.6	0	3.6	0
18	3.6	3.6	3.6	0	3.6	0
19	3.6	3.6	3.6	0	3.6	0
20	3.6	3.6	3.6	0	3.6	0
21	3.6	3.6	3.6	0	3.6	0
22	3.6	3.6	3.6	0	3.6	0
23	3.6	3.6	3.6	0	3.6	0
24	3.6	3.6	3.6	0	3.6	0
25	3.6	3.6	3.6	0	3.6	0
26	3.6	3.6	3.6	0	3.6	0

values[7]. In this study a test workload consisting of many synthetic programs was run several times and the results show that the processor time measured fairly constant from run-to-run as indicated by a standard deviation of 1% to 2% of the mean processor time. A larger variation in I/O time was observed as indicated by a standard deviation of 5% to 15% of the mean value. Complete details of the AFDSDC study are given in Reference 7.

2.1.2.3 Site Accounting Data

The use of SCF data from operational Air Force WWMCCS sites was selected as the most expeditious method of generating synthetic test workloads that (1) reflect to some degree the system resource demands for CPU time, I/O time, and core memory that actually occur at operational sites and (2) allow use of available synthetic workload generation tools. Accounting data in the form of SCF data was provided by participating Air Force WWMCCS sites for use in generating synthetic test workloads. The sites participating in the study included MAC, PACOM, SAC, and TAC.

Ten SCF data tapes were selected for synthetic workload generation. The characteristics of the data on each of the ten accounting tapes is summarized in Table 2-II. This data was used as input to the automatic synthetic workload generator described in section 2.1.2.4 of this report to produce the synthetic test workloads described in section 2.1.2.5 of this report.

2.1.2.4 Synthetic Workload Generator

The synthetic workload generator (provided by Horan and Venese) was chosen for this study to construct test workloads. The generator builds, in three steps, a test workload using SCF accounting data. The first step consists of analyzing the input accounting data, which includes the processor time (X1), I/O time (X2) and core size

Table 2-II

Site SCF Data

SCF IDENTIFIER	NUMBER OF JOBS	CPU TIME (Seconds)			I/O TIME (Seconds)			CORE USED (K Words)		
		MIN.	MAX.	TOTAL	MIN.	MAX.	TOTAL	MIN.	MAX.	TOTAL
1	34	0.31	9799.27	39277.27	0.07	4472.02	16308.85	5	57	953
2	251	0.07	5548.81	25739.90	0.00	1712.16	11015.35	4	62	5725
3	1185	0.03	1423.86	51360.82	0.00	7829.23	156190.03	5	62	22435
4	1169	0.08	4635.49	59665.88	0.00	12385.74	85849.19	4	71	28967
5	31789	0.04	33397.60	1492489.05	0.00	719896.32	5992574.00	2	125	712892
6	254	0.12	7439.68	27061.02	0.00	3861.69	33741.93	2	45	2933
7	253	0.11	9029.36	21316.83	0.00	5789.87	37759.27	4	40	3686
8	180	0.14	3912.98	11863.40	0.01	8492.67	16925.84	5	50	3998
9	2886	0.12	2910.70	106979.18	0.00	9429.31	187445.34	5	61	68867
10	3726	0.10	5872.41	177510.97	0.00	16081.63	285321.08	5	55	86761

(X3) for every individual job in the actual workload. Every set of values (X1, X2, X3) is examined to determine the respective maximum and minimum values, from which a three-dimensional grid with the following cell dimensions is constructed.

$$X1 = \frac{X1_{max} - X1_{min}}{100} \quad \text{along X1 axis}$$

$$X2 = \frac{X2_{max} - X2_{min}}{60} \quad \text{along X2 axis}$$

$$X3 = \frac{X3_{max} - X3_{min}}{10} \quad \text{along X3 axis}$$

There are 60,000 equal cells in the three-dimensional matrix. The input data is scanned again and the jobs are placed in the appropriate cells depending on the individual values of X1, X2, and X3. The individual cell population (i.e., the number of jobs in each cell) is divided by the total number of jobs in the actual workload to obtain the joint probability density function and the cell coordinates are given by the values of X1, X2, and X3 at the midpoint of the cell. This process identifies the most frequently occurring jobs in the actual workload. The generator has editing features that allow the maximum value of each of the three variables to be discarded and the next highest value to be selected as the new maximum value.

The synthetic workload generator was designed to produce a test workload using the accounting data with a minimum of manual intervention. Because of this, the workload characterization part of the generator has certain limitations. The division of the total

workload into cells to approximate the real workload characteristics is arbitrarily performed by dividing the range between the maximum and minimum values of the real workload into equal parts, without inspection of the actual distribution of the jobs over the above range. In follow-on studies, the actual distribution of the jobs (for example, histograms for the selected variables) should be considered before deciding on the cell division.

During the second step, the test workload characteristics are determined. Either the number of jobs in the test workload or the total processor time of the test workload may be specified. For these tests the number of jobs in each test workload was specified. The number of jobs specified is multiplied by the joint-probability density for each cell to determine the number of jobs to be generated with the characteristics of each cell and rounded off to the next higher integer, thus completing the determination of the test workload.

During the third step, the test workload characteristics are converted into synthetic program parameters by means of the calibration equations built into the workload characterization portion of the synthetic workload generator. The processor time (for the user program) and the I/O time (for the user program) were selected as the major characteristics to be duplicated in constructing the test workload. The values of these two variables were obtained from the Statistical Collection File (SCF data). The synthetic program used consisted of two parameters, NIO (number of I/O loops) and NCPU (number of CPU loops), which determine the number of records written to or read from the four files and the number of executions of the CPU kernel, respectively. Calibration consists of relating the CPU time and I/O time for a synthetic job with the corresponding values of NCPU and NIO. The workload characterization program must be calibrated for the system under test. This is due to the fact that, simplistically, the synthetic program parameters (NIO and NCPU) are

determined by dividing the I/O time (and/or CPU time) to be synthetized by the time used to complete one I/O loop (and/or CPU loop). In reality, the calibration equations are more complex. In order to perform an I/O operation; instructions must be executed in the CPU, which results in accumulation of CPU time, when performing I/O loops. Hence, the calibration equations contain I/O-to-CPU coupling terms. The coupling effect, which limits the range over which the synthetic program can be used, may be lessened by reducing the value of the processor time per I/O loop. This was accomplished by using GMAP subroutines to perform the I/O instead of FORTRAN statements. The reduction in the value of processor time per I/O loop was accompanied by some decrease in the I/O time per I/O loop, but still yielded a higher I/O-to-CPU ratio.

Another problem that exists in the calibraiton needs to be mentioned. The value of NIO is determined by the I/O time of the individual job, which is not reproducible, i.e., the same job, run under identical conditions, may yield different values of I/O time (see section 2.1.2.2 for details). The I/O time of a job is subject to variation and the amount of variation depends on several mix-dependent parameters; for example, the degree of multiprogramming and the nature of processing performed by the individual jobs. The values of the synthetic program parameters, NIO and NCPU, are calculated based on the values of processor time and I/O time determined by the real workload. Because of the coupling effects of the synthetic program and the non-reproducibility of I/O time per job, the synthetic program when executed may not yield the values of the processor and I/O time of the real workload.

Time for one loop can vary drastically when major changes to the system are made. In the case of this study, recalibration was needed because the release of GCOS being used was several releases beyond the release used for the previous calibration. The

actual calibration was accomplished during the pilot testing and is described in section 3.1.1 of this report.

Minor changes were made to the synthetic workload generator. The calibration constants were changed as described in section 3.1.1. The initial size of the four files that each synthetic program requested was changed from 20 to 200 links. The minimum core size of the synthetic program was changed from 11K to 12K words. Because of the compiler restrictions (in version GCOS SR WW6.2.1), the minimum amount of core required to execute the synthetic program was 12K and, as a consequence, user jobs requesting less than 12K were represented by the minimum value of 12K words. The user programs that requested a core size greater than 12K were represented by synthetic programs with core sizes equal to the requested size. This method of core representation is useful in testing the system core allocation policies and the resulting influence on the multiprogramming depth. It should be noted, however, that only 12K words of all the jobs in the test workload participated actively during processing. The values of NIO and NCPU are determined by using the calibration equations discussed earlier. The value of the core size is set to the required value via the LIMITS card in the job stream. The test workload is then punched out in the form of a series of synthetic program control card decks.

2.1.2.5 Synthetic Workloads

Construction of a test workload is similar to model building. The only way to capture all the details of the actual workload is to duplicate the actual workload at great expense of money and time. The test workload should be representative of the actual workload in order that valid conclusions may be drawn from its use. The modeling of an actual application workload requires a

compromise between the details to be represented and the time needed to process the test workload.

A representative test workload may be defined as a test workload having reasonable fidelity to the actual workload that satisfies several practical constraints. The variables chosen for representation depend on the goals of the evaluation study, the system being studied, and the availability of performance monitors.

Construction of a representative test workload consists of two steps: characterization of the actual workload and implementation of the synthetic workload. The first step involves (1) selection of a method for characterizing the workload and choice of the characteristic variables, (2) selection of a method for representing the workload, and choice of criteria to evaluate the "representativeness" of the test workload, and (3) determination of the test workload characteristics. The second step involves (1) design of the synthetic program(s) to be used, (2) calibration of the synthetic program, and (3) translation of test workload characteristics into the number of and parameters for, the synthetic programs. The experimental evaluation of multiprocessor performance reported in this document was primarily for the purpose of determining values of relative throughput when the number of processors is increased. Two sets of test workloads were used in the experiments. The first set of test workloads (7) was constructed using the accounting data provided by the participating sites; the second set (7) was artificially constructed so as to cover a range of values for I/O time/CPU time (e.g., IO/CP ratio).

Figure 2-4 describes schematically the various steps in reducing the site accounting data into test workloads. The reduction procedure may be described as follows. Raw SCF data collected on several days or several sessions is processed by the merge program, a COBOL program that essentially recasts the raw data into system standard format and resequences any out-of-sequence

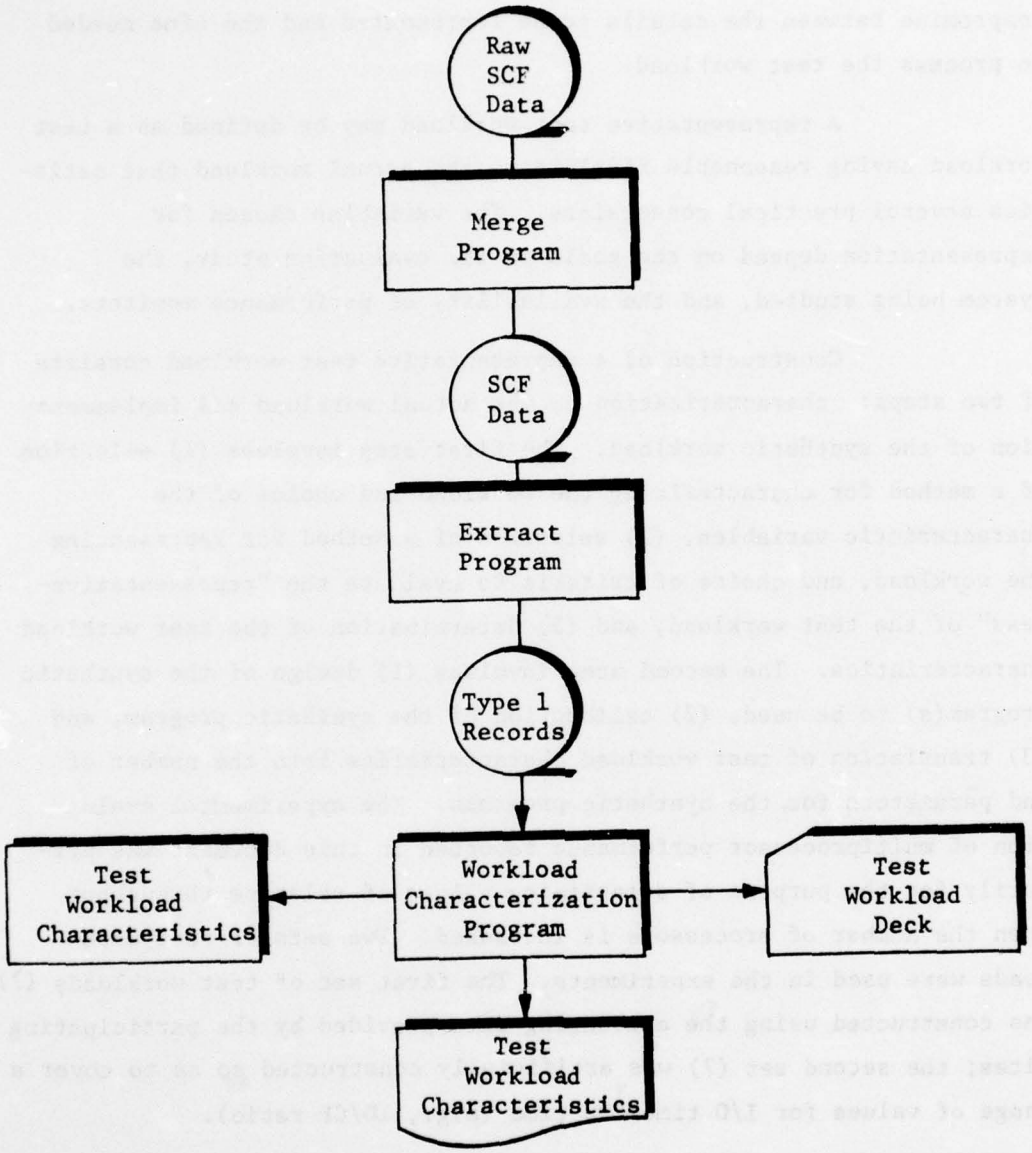


Figure 2-4. Steps in Analyzing SCF Data

data blocks. The resequencing is necessary because every time the system is reloaded the block sequence number is reset to the initial value. The output of the merge program is then fed to the extract program that extracts all the Type 1 records, which contain information about the magnitude of demands placed by the batch jobs on the system resources. The output of the extract program is then fed to the synthetic workload generator.

Table 3-VIII shows the characteristics of the 14 test workloads generated for the experiments. The table lists the number of jobs in each test workload, the values of synthetic program parameters, the processor time, the I/O time and the ratio of processor time to I/O time. Workloads 1 through 4 and 8 through 10 were produced using the accounting data provided by operational WWMCCS sites. However, they cannot be considered as completely representative of the site workloads primarily because the interactive workload of the site was not explicitly represented.

2.1.2.6 Synthetic Program

The synthetic program (SYN) used in this study simulates file activity on four files. The load on the I/O subsystem (measured in terms of I/O time) is varied by changing the number of records written to or read from the four files. The file manipulation activity is interspersed with execution of a compute kernel and the number of times the kernel is executed determines the load on the processor (measured in terms of CPU time). The calibration of the synthetic workload generator is a function of the actual time needed to execute one I/O loop and/or one CPU loop.

2.1.2.7 Microscopic Analysis of Synthetic Program

The synthetic program used in the construction of the test workload has been described in terms of its global behavior in the previous section. The global behavior, in this context, refers to the total processor and I/O time required by the job. Other job

characteristics that influence the performance of a workload are the memory reference patterns, the time variation of the resource demand, and the degree of overlap between I/O and CPU usage.

Originally, it was planned to study the memory referance patterns using a hardware monitor, but this effort was unsuccessful. Software monitors, SYRUP-II and Mass Store Monitor (MSM), were used to study the microscopic interaction. A synthetic program was specially constructed for this purpose with the following parameter settings: NCPU = 1,000,000 and NIO = 100,000. This job was run single-thread and the elapsed time under these conditions for the job was 35.22 minutes. During this interval the state of the system resources was sampled by SYRUP approximately every second. The quantities sampled were: (1) percent of processor time used by user programs, (2) percent of IOM time used by user programs, (3) number of user programs executing, (4) number of user programs swapped (5) number of user programs queued, (6) amount of memory used by the user programs, (7) percent of time used by the synthetic program, and (8) percent of IOM time used by the synthetic program. Simultaneous with the running of SYRUP, the mass store monitor (MSM) was used to measure the number of connects to each physical device. The number of connects refers to the number of I/O initiations.

The SYRUP results indicate a minimum value of 70% and a maximum value of 98% for the processor utilization. It should be pointed out that SYRUP was monitoring both the synthetic program and MSM. The SYRUP results could not be used to obtain any quantitative insight into the behavior of the synthetic program; however they show qualitatively how the utilization of the processor varies with time.

The synthetic program is a single-activity program written in FORTRAN and uses GMAP procedures for performing the reads and writes. Three parameters influence the execution of the synthetic program: core size, the CPU-loop parameter (NCPU), and

the I/O-loop parameter (NIO). The core size is specified in the $LIMITS card and the specified core is allocated to the synthetic program by the core allocator. The synthetic program uses four files and the requested primary space is allocated by the peripheral allocator. The synthetic program imitates a general file update program and initially sets up a table of numbers and then executes the CPU kernel NRPT times where NRPT = NCPU/3. After this burst of processor activity, NIO records are written out to files 1 and 2. This I/O activity is followed by another burst of processor activity during which the CPU kernel is executed NRPT times. At the end of the second burst of processor activity, NIO records are read back from files 1 and 2. After the reading, the CPU kernel is executed NRPT times followed by the writing of NIO records to files 3 and 4. The synthetic program thus consists of three sets of processor and I/O activity. The values of NIO and NCPU control the duration of these processor and I/O activities. Writing and reading of records are accomplished by transferring control to the GMAP routines at the appropriate time.

The CPU kernel consists of adding numbers in a table initialized in the beginning. The table has 1000 serial numbers starting from 100. A single execution of the CPU loop consists of adding 10 numbers from the table and checking the sum independently by using the formula for the sum of numbers. The successive numbers to be summed are the first, the eighth, the twenty-seventh, etc., value in the table, that is, the Kth value in the summation is the K^3th entry in the table.

2.1.3 Verification

The conclusions of this study are intended to be applicable to other systems besides those used to develop the test workloads. AFDSDC was provided an early draft of this report and was asked to verify by direct application the applicability of the study. Their report is reproduced herein as Appendix I.

2.2 Test Design

The design of the tests used for the WWMCCS Multiprocessor Performance Evaluation task concentrated on placing known resource demands generated by synthetic workloads on instrumented systems. By carefully controlling the system environment to prevent extraneous interactions, measurements could be gathered to reflect the system's reaction to the workload. This was done in an attempt to identify those system parameters that influence throughput and, hence, relative throughput.

2.2.1 Test Measurements

A broad range of performance measurements was planned for collection during the testing to provide an indication of the system's reaction to the range of synthetic workloads in terms of elapsed time and resource utilizations.

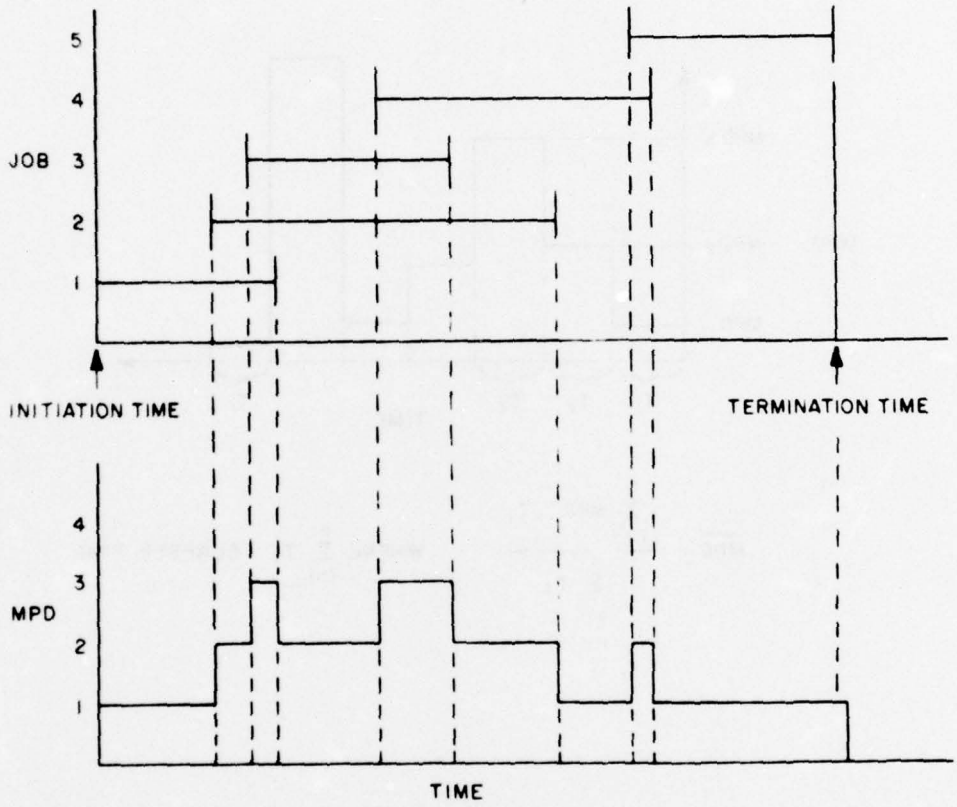
2.2.1.1 Elapsed Time

Elapsed time is the difference between the start and finish times of a particular workload. Two sources of this measurement were planned for the tests. The GSEP and console log listings both give the start and stop time of each job within the workload, from which the elapsed time for the entire workload can be determined.

2.2.1.2 Multiprogramming Depth (MPD)

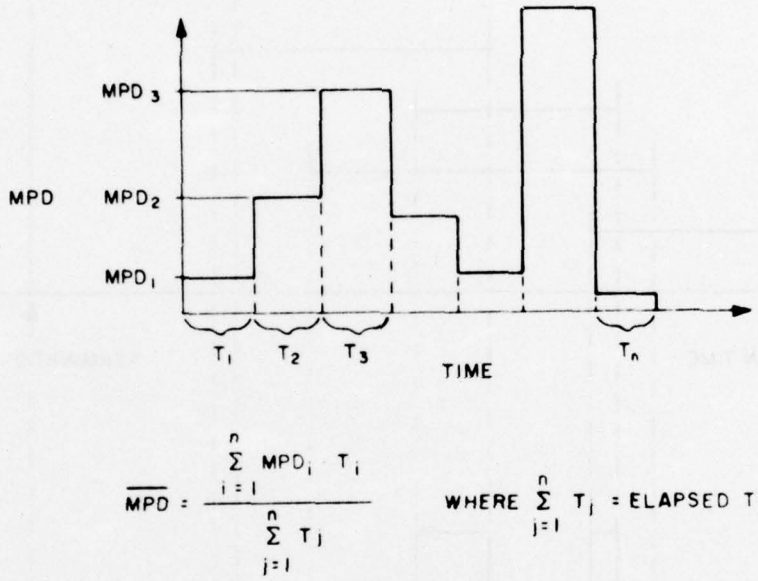
The number of user programs or jobs in concurrent operation is called the multiprogramming depth. The average multiprogramming depth (MPD) can be determined using the SCF data and a Gantt Chart representation as shown in Figures 2-5 and 2-6.

The impact of MPD on average throughput is shown in Figure 2-7. Initially, a rapid increase in throughput is experienced due mainly to the IO-CPU overlap at the lower degrees of multiprogramming. However, as the number of executing programs increases, the overhead within the operating system needed to handle interrupt and scheduling processing begins to degrade the throughput. The



IA-47,223

Figure 2-5. Gantt Representation



$$\overline{MPD} = \frac{\sum_{i=1}^{n} MPD_i \cdot T_i}{\sum_{j=1}^{n} T_j} \qquad \text{WHERE } \sum_{j=1}^{n} T_j = \text{ELAPSED TIME}$$

IA-47,221

Figure 2-6. Average Multiprogramming Depth

IA-47,543

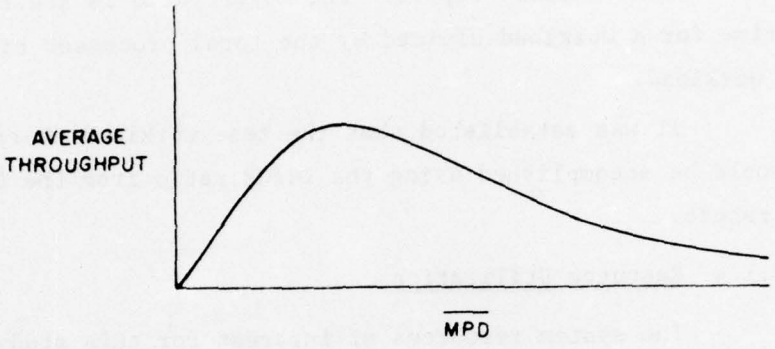


Figure 2-7. $\overline{MPD}$ vs Average Throughput

maximum multiprogramming depth to be allowed was planned at a value of 35. SYRUP-II data, as well as the console log listings, was planned to monitor the variation in multiprogramming depth during the tests.

2.2.1.3 IO/CP Ratio

Jobs or workloads can be characterized qualitatively as I/O-bound, balanced, or CPU-bound. The IO/CP ratio of a workload derives from the amount of time a workload spends doing computational and I/O operations. The IO/CP ratio can be used to characterize a

job or workload quantitatively. This ratio is provided automatically as part of the GSEP summary report. The IO/CP ratio is the total channel time for a workload divided by the total processor time for the same workload.

It was established that the test workload characterization would be accomplished using the IO/CP ratio from the GSEP summary report.

2.2.1.4 Resource Utilization

The system resources of interest for this study were the CPU's, SCU's, core memory, and the disk subsystem. Resource utilization measurements for these hardware components, or the percentage of time a given component was busy during the elapsed time, can be recorded by software monitors. These utilizations, in a controlled environment, result from the demand placed on the system resources by the scheduling and allocation policies of GCOS on behalf of the test workloads.

Several software monitors were investigated, as part of the test design, in order to provide a mechanism for collecting system resource utilizations during the execution of the test workloads on the selected configurations. The software monitors selected are discussed in section 2.2.2.2 of this report.

2.2.2 Test Instrumentation

To collect a variety of performance information on the system's response to the synthetic workload demands, a number of measurement tools were planned for use during the testing. By monitoring the resource utilizations of a configuration's components

(i.e., CPU's, SCU's, memory and I/O subsystem), a corroboration of the system throughput variations can be achieved.

The types of performance information to be collected are dependent on the availability of measurement tools. The Honeywell accounting package (GSEP, GCOS Summary Edit Program) and two software monitors (SYRUP-II and MSM) were planned for use in all tests. A hardware monitor (TESDATA 1185) was also planned for use but was not successful.

2.2.2.1 System Accounting Data

The HIS 6000 system software provides for the collection of user-oriented accounting information (i.e., SCF File) and a reduction program to process it. This accounting data is placed on the SCF tape as the system terminates the processing of the individual jobs in the synthetic workload.

An edit program (GSEP), using the SCF data as input, produces a series of reports for use in resource accounting. These reports include data on CPU time used by a job or activity, the start and stop times of a job or activity, file accesses made by a program, the number of records read or written and to what units, and other information not germane to this study.

2.2.2.2 Software Monitors

Software monitors are special-purpose programs which reside in the memory of the measured system. These monitors can be independent programs executed with high priority or can be implemented by modifying the system software itself. Although software monitors provide information not readily obtainable from other sources, they do impose a cost in system overhead because of the resources (i.e., CPU, memory and I/O) they themselves consume.

Generally, software monitors are classified as either event-driven (i.e., initiated as the result of some event), or sampling (i.e., collect data at fixed, but possibly adjustable time intervals). For the sampling type software monitor, the interval of time between activations is very important. Too large an interval means a system change might be missed; too small an interval causes increased system overhead to collect the information. An event-driven software monitor is specifically designed to record information on particular occurrences (e.g., the reassignment of a processor, I/O operation of a given device type). Because of their specificity, event-driven monitors have limited application and varying degrees of overhead.

SYRUP-II is a sampling type of software monitor. It was developed at Hq. SAC specifically for the HIS 6000 computer systems. Version II is a revision of an earlier release of SYRUP that incorporates changes to decrease the amount of memory needed and decrease CPU overhead. The data collected by SYRUP-II is written to the SCF tape, but a dedicated tape can be assigned via an operator console command. An operator interface is provided to dynamically adjust the sampling rate, determine system status, and start or stop the collection of particular information. A sampling rate of 10 seconds was chosen for all testing described in this report.

The information selected to be collected by SYRUP-II included average processor overhead and active times, average percentage processor time for system and user programs, IOM counts for each I/O multiplexer over time, and average counts for programs swapped and currently executing.

Along with the SYRUP-II monitor, a reduction program is available. This program has the ability to merge several SCF tapes, selectively process information delimited by date and time, and produce a graphical output of collected information. The data

can be averaged by second, minute, hour, or day. For the WMPE tests, averaging by day was used to ensure that all the information for a synthetic workload would be processed.

The Mass Store Monitor (MSM) is an event-driven software monitor. It was developed by Honeywell, under contract from CCTC, to capture data on disk subsystem activity. MSM modifies the GCOS Dispatcher (.MDISP) module and is initiated by an I/O request for disk access. A dedicated tape unit is required to operate MSM. The MSM program must reside in memory at all times during its use. It is, therefore, given a privileged status and once in memory cannot be swapped or relocated. It is appropriate to start MSM before any other software monitors or synthetic programs are initiated, preventing the fragmentation of memory possible if MSM were started at some other time.

Like SYRUP-II, MSM has a reduction program to process the information collected on the tape. The types of data and statistics provided by the MSM reduction package included histograms for device and file space accesses, device and IOM utilizations, seek movement distances and frequencies.

The Memory Utilization Monitor (MUM) is also an event-driven software monitor. It was developed by Honeywell, under contract to CCTC, to capture data on the use of main memory during a given period of time. The MUM package consists of the capture program, which places a hook into the GCOS Dispatcher (.MDISP) module, and a data reduction program which generates reports from the data collected.

MUM reports include a graphic display of memory demands versus availability for a given period of time, various percentage distributions for memory requests by size, memory allocator queue size distributions and memory demands with processor activity.

2.2.2.3 Hardware Monitors

Although a TESDATA 1185 hardware monitor was planned for use during the Phoenix tests and was connected to the CCTC system for familiarization, no data was collected.

However, some comment should be made about the use of hardware monitors in throughput studies. A hardware monitor is a device used to detect electrical impulses at selected points in the circuitry of a computer. Some monitors, like the TESDATA 1185, have a minicomputer as part of their structure to process information and produce statistics as it records the data. The electrical impulses are sensed by high-impedance probes attached to the system. Since a hardware monitor does not use any system resources, it is transparent to the system software and does not impact the operation.

Precise measurements of overall resource utilizations and concurrent usage can be achieved using a hardware monitor. Some measurements, (e.g., memory access interference) can only be obtained by using a hardware monitor and, conversely, most job related information and queue length information is unavailable to a hardware monitor.

2.2.2.4 Console Log Sheets

The operator's console is a hard-copy, typewriter-like device. The console log, like the accounting data, records start and stop times of jobs and shows any interactions between the operator and the system, such as the request to mount a tape or the occurrence of a memory parity error. The console log also indicates the number of programs in concurrent execution. The plan for the WWMCCS Multiprocessor Performance Evaluation task included the console logs as an integral part of the instrumentation.

2.2.3 Selected Configurations

Eight hardware configurations were chosen for use in the testing. Although these represent a small subset of the number of H6000 configurations possible, they do include a good sampling of WWMCCS installations.

Control cards for the Configuration Section of the bootload deck were prepared for each of the eight configurations. Schematic diagrams of the eight configurations were prepared and are included in section 3.2.1 of this report.

2.2.4 Software Alterations

Four changes were required to the GCOS SR WW6.2.1 release obtained from CCTC due to the use of special software monitors and the delays in job initiation caused by the interaction between the temporary files used by the synthetic workloads and the WWMCCS security module. The following changes were made using octal patches to the $ PATCH Section of the bootload deck and the addition of an object module to the $ LOAD Section:

1. A patch was made to GCOS modules .SWAP and .MDISP to provide system status information to SYRUP-II.
2. An object module for .MXSA5 was added to allow console operator interaction with SYRUP-II.
3. A patch was made to the GCOS dispatcher (.MDISP) to allow disk activity to be recorded to the MSM software monitor.
4. A patch was made to .MALC5 and .MALC6 to disable the WWMCCS security module (FS49).

2.2.5 Test Procedures

An initial set of test procedures, using a prearranged sequence of steps, was developed to ensure the proper initial system state and to control the submission of test workloads to the system. Figure 2-8 shows the initial set of instructions developed for use during the conduct of the tests. During the pilot testing at CCTC, these procedures were refined for use during the primary testing at Phoenix.

I. SYSTEM STATE

A. Configuration Deck
- Verify that the Configuration Section specifies the actual hardware connected for the test.
- Have the operator confirm the set-up of the Configuration Section, check the interleaving switch settings on the SCU's, CPU's, and IOM's and note which of the CPU's, SCU's, and IOM's are being used; i.e., sketch a hardware configuration diagram.

B. Disk Packs
- The six removable disk packs should be mounted and powered up.
- They may have been used as scratch packs for processing: check this - if they have been, they must be restored.
- RESTORE PACKS

  Give the operator the latest SAVE Tape.
- BEGIN BOOT PROCESS

  Give the operator the INIT Tape.

C. SCF & SYRUP-II
- Have the operator mount a scratch accounting tape and label it appropriately.

D. Boot System
- Using the Bootload Deck have the operator boot the system.
- Check that the Bootload Deck does not cause any extraneous error messages on the operator's console. If any occur, investigate the reason(s).

Figure 2-8. Test Procedures

	E. System Information	- Have the operator enter the following console commands to verify the system configuration and file definition: LISTCA, LISTCN, and TYPEFG. - Have the operator use the card-to-printer utility to list the Bootload Deck. - Be sure that all test generated output goes to one dedicated printer. - Do not separate the printed output while the test is being run. - Have the operator enter the console command LIMIT NONE. - Have the operator enter the console command PURGE ALL (mount scratch tape on request).
II.	START MONITORS	
	A. Spawn MSM	- Have the operator spawn MSM00. It will call for a scratch tape to be mounted. Label it appropriately.
	B. Start SYRUP	- Run the job to start SYRUP (small deck with SNUMB SYRUP. It will ask for PRIVITY and the operator should allow it to be run. The following table will be displayed on the console listing: H6000 SYSTEM BEWARE --SYRUP- II IS ON THE AIR-- MAINFRAME PERIPHERALS

Figure 2-8. Test Procedures (Continued)

C.	Check SYRUP Status	- Be sure to type each SYRUP command exactly as shown!!! - Use operator interface (Request and End-of-Mesaage). In response to ??? type SYRUP STATS. The resultant table will be printed: **SYRUP STATUS** RECORD IN-USE/ACCNT/ON-LINE MAIN X X PERIPH X X COMM NO ON-LINE OVERLAY IN USE
D.	START HARDWARE MONITOR (if available)	- Have hardware monitor team begin their data collection
E.	Space Console Log	- Space the paper in the operator's console to a new page to separate the start-up processing and the actual workload start- and end-of job statistics. - Note on the console log the workload and hardware configuration used in this experiment.
F.	Set up sign	- Place sign on console "DO NOT ENTER OR REQUEST INFORMATION DURING THE TEST."
III.	START TEST	
A.	Read Workload	- Have the operator read in experiment workload from the appropriate card reader.
B.	Check Progress	- Do not use the console command LSTAL indiscriminately. Watch for the EOJ's and when it looks like all jobs have ended, then use LSTAL; MSMOO and SYRUP should be the only jobs remaining.

Figure 2-8. Test Procedures (Continued)

	C. Run Until End	- Run the entire workload.
IV.	FINISH DATA COLLECTION	
	A. Terminate MSM	- Have the operator terminate MSM00 by using the TERM console command. Save the data tape for data reduction.
	B. Terminate SYRUP	- In response to the operator interface, i.e., ???, type SYRUP STOP MAIN. Do the same again with SYRUP STOP PERIPH. SYRUP may now be terminated using the console command TERM. Save the accounting tape for data reduction.
	C. Terminate Hardware Monitor	- Stop the hardware monitor. Save the output tape.
V.	DATA REDUCTION	
	A. Accounting Tape	- Run GSEP on the accounting tape with option 1. This is the SCF & SYRUP data tape.
	B. SYRUP Reduction	- Run the SYRUP data reduction package, making the appropriate changes in the tape label card and parameter cards (the tape label card should have a 5-digit number matching the label on the accounting tape, and the parameters are given in the SYRUP documentation).
	C. MSM Reduction	- Run the MSM reduction package against the MSM data tape.
VI.	COMPLETE TEST	
	A. Console Log	- Space the console paper to a new page and mark end-of-experiment.
	B. Printer Output	- Separate the printer output from the test and label it appropriately

Figure 2-8. Test Procedures (Concluded)

SECTION III

TEST RESULTS

Testing activity for the WWMCCS Multiprocessor Performance Evaluation task took place during the period of January - July 1976. The tests were conducted in accordance with the test plan and in the controlled environment described in Section II. However, not all of the planned test cases were conducted; of the planned 336 test cases, a total of 273 test cases (81%) were actually conducted. The overall test effort included three major phases. The three phases, called pilot testing, primary testing, and secondary testing, are described in this section.

3.1 Pilot Testing

The pilot testing was accomplished using dedicated H6000 time at the CCTC Computer Facility in Reston, Va. during the period of January - April 1976. The pilot tests involved eleven test sessions during which the synthetic workload generator was recalibrated, synthetic test workloads were built and tested, software monitors were installed and checked out, a hardware monitor was connected to the test system, data reduction and archiving procedures were developed, timing runs for some test workloads were accomplished, and test procedures were verified.

3.1.1 Synthetic Workload Generator

The first effort under the pilot testing was the recalibration of the synthetic workload generator. As discussed in section 2.1.2.5, the generator had to be recalibrated to reflect changes in system characteristics. The recalibration for the tests described in this report was done in three stages.

Initially, a five-hour synthetic test workload was executed on an H6060 single processor system to collect timing data for recalibration of the equations in the workload characterization program which is part of the synthetic workload generator. The jobstream consisted of 50 jobs with various settings for NIO (number of I/O loops) and NCPU (number of CPU loops). The actual values used for each job (SNUMB) are shown in Table 3-I with the measured values for processor time of each job in the calibration run. The execution sequence of the workload is illustrated in Figure 3-1 which also provides a graphical representation of the multiprogramming depth for the workload. In one set of 13 jobs, the number of CPU loops (NCPU) executed was zero and the number of I/O loops (NIO) executed varied from zero to 100,000. In a second set of 13 jobs, NIO was zero and NCPU varied from 10 to 100,000. The remaining jobs used six combinations of non-zero parameter settings. The test results presented in Table 3-I were used to reformulate the calibration equations for the synthetic workload generator.

A set of 7 jobs each used identical parameter settings (NCPU = 100,000 and NIO = 10,000). For these jobs the processor time varied from 186.5 seconds to 199.1 seconds, whereas the I/O time varied from 237.6 seconds to 363.6 seconds. The processor time variation is approximately 6% and as expected the I/O times varied as much as 35%. This trend was confirmed for different values of parameter settings. A set of six jobs each used identical parameter settings (NCPU = 10,000 and NIO = 15,000). For these jobs the I/O time varied from 345.6 seconds to 547.2 seconds (i.e., 37%).

The measured values of processor time and I/O time were used to determine the coefficients of the calibration equations.

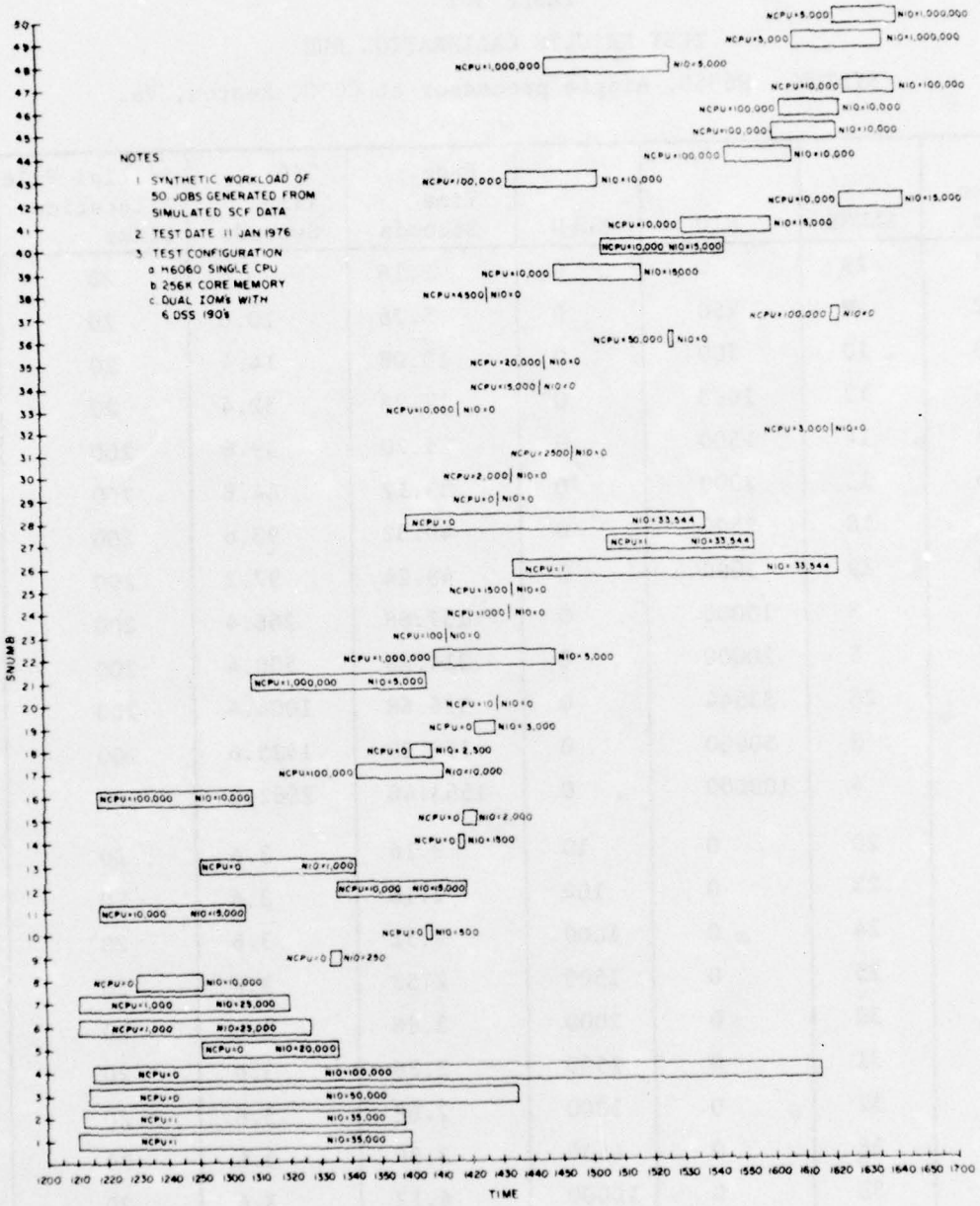


Figure 3-1. Synthetic Workload Multiprogramming Depth

TABLE 3-I

TEST RESULTS CALIBRATION RUN

SYSTEM - H6060, single processor at CCTC, Reston, Va.

Run No.	SNUMB	NIO	NCPU	Proc. Time Seconds	I/O Time Seconds	Initial File Allocation Links
1	29	0	0	2.16	3.6	20
2	9	250	0	5.76	10.8	20
3	10	500	0	10.08	14.4	20
4	13	1000	0	17.28	32.4	20
5	14	1500	0	25.20	39.6	200
6	15	2000	0	33.12	64.8	200
7	18	2500	0	40.32	93.6	200
8	19	3000	0	48.24	97.2	200
9	8	10000	0	157.68	266.4	200
10	5	20000	0	314.28	500.4	200
11	28	33544	0	526.68	1004.4	200
12	3	50000	0	782.28	1425.6	200
13	4	100000	0	1553.40	2862.0	200
14	20	0	10	2.16	3.6	20
15	23	0	100	2.16	3.6	20
16	24	0	1000	2.52	3.6	20
17	25	0	1500	2.52	3.6	20
18	30	0	2000	2.88	3.6	20
19	31	0	2500	2.88	3.6	20
20	32	0	3000	2.88	3.6	20
21	38	0	4500	3.96	3.6	20
22	33	0	10000	6.12	3.6	20
23	34	0	15000	7.92	3.6	20
24	35	0	20000	9.72	3.6	20
25	36	0	50000	21.96	3.6	20

TABLE 3-I

TEST RESULTS CALIBRATION RUN (Concluded)

Run No.	SNUMB	NIO	NCPU	Proc. Time Seconds	I/O Time Seconds	Initial File Allocation Links
26	37	0	100000	40.68	3.6	20
27	46	10000	100000	186.48	237.6	20
28	47	10000	100000	186.48	223.2	20
29	45	10000	100000	189.36	252.0	20
30	44	10000	100000	198.72	295.2	20
31	43	10000	100000	199.08	241.8	200
32	16	10000	100000	199.10	280.8	20
33	17	10000	100000	192.24	363.6	200
34	42	15000	10000	217.10	345.6	200
35	41	15000	10000	246.40	414.0	20
36	40	15000	10000	244.10	374.4	200
37	39	15000	10000	231.50	547.2	200
38	11	15000	10000	246.60	414.0	20
39	12	15000	10000	235.80	543.6	200
40	21	5000	1000000	479.50	187.2	20
41	22	5000	1000000	473.40	187.2	200
42	48	5000	1000000	476.30	144.0	20
43	49	5000	1000000	461.20	133.2	200
44	50	5000	1000000	462.60	122.4	20
45	27	33544	1	532.40	867.6	20
46	26	33544	1	501.50	1000.8	200
47	6	25000	1000	401.04	680.4	20
48	7	25000	1000	389.52	727.2	200
49	1	35000	1	559.80	1004.4	20
50	2	35000	1	545.00	946.8	200

$$NIO = \frac{[(X2 - C2) * 1000]}{T2} \quad (1)$$

$$NCPU = \frac{[(X1 - C1) * 1000] - (T3 * NIO)}{T1} \quad (2)$$

where

NIO = Number of I/O loops needed in synthetic program

NCPU = Number of CPU loops needed in synthetic program

X1 = CPU time (seconds) to be synthesized

X2 = I/O time (seconds) to be synthesized

C1 = CPU offset (seconds)

C2 = I/O offset (seconds)

From the calculations shown in Table 3-II, values for the coefficients were determined to be:

C1 = 2.16 seconds

C2 = 3.60 seconds

T1 = 0.36 milliseconds

T2 = 27.57 milliseconds

T3 = 15.40 milliseconds

where

T1 = Average CPU time for one CPU loop (milliseconds)

T2 = Average I/O time for one I/O loop (milliseconds)

T3 = Average CPU time for one I/O loop (milliseconds)

and

T4 = (N1 - T3 * NIO) * 1000

T4 = Average CPU time for m CPU loops and n I/O loops (milliseconds)

N1 = X1 - C1

N2 = X2 - C2

TABLE 3-II

SYNTHETIC WORKLOAD GENERATOR
CALIBRATION CALCULATIONS

SNUMB	NCPU	NIO	LINKS	CORE	CPU SEC	I/O SEC	N1	N2	T1	T2	T3	T4
1	1	35,000	20	15K	559.80	1004.4	557.64	1000.8	-	28.59	-	18.64
2	1	35,000	200	20K	545.04	946.8	542.88	943.2	-	26.95	-	3.88
3	0	50,000	200	30K	782.28	1425.6	780.12	1422.0	-	28.44	15.60	-
4	0	100,000	200	40K	1553.40	2862.0	1551.24	2858.4	-	28.58	15.51	-
5	0	20,000	200	50K	314.28	500.4	312.12	496.8	-	24.84	15.61	-
6	1,000	25,000	20	15K	401.04	680.4	398.88	676.8	-	27.07	-	13.88
7	1,000	25,000	200	20K	389.52	727.2	387.36	723.6	-	28.94	-	2.36
8	0	10,000	200	30K	157.68	266.4	155.52	262.8	-	26.28	15.55	-
9	0	250	20	40K	5.76	10.8	3.60	7.2	-	28.80	14.40	-
10	0	500	20	50K	10.08	14.4	7.92	10.8	-	21.60	15.85	-
11	10,000	15,000	20	15K	246.60	414.0	244.44	410.4	-	27.36	-	13.44
12	10,000	15,000	200	20K	235.80	543.6	233.64	540.0	-	36.00	-	2.64
13	0	1,000	20	30K	17.28	32.4	15.12	28.8	-	28.80	15.12	-
14	0	1,500	200	40K	25.20	39.6	23.04	36.0	-	24.00	15.36	-
15	0	2,000	200	50K	33.12	64.8	30.96	61.2	-	30.60	15.48	-
16	100,000	10,000	20	15K	199.08	280.8	196.92	277.2	-	27.72	-	42.92
17	100,000	10,000	200	20K	192.24	363.6	190.08	360.0	-	36.00	-	206.00
18	0	2,500	200	30K	40.34	93.6	38.18	90.0	-	36.00	15.27	-
19	0	3,000	200	40K	48.24	97.2	46.08	93.6	-	31.20	15.36	-
20	10	0	20	50K	2.16	3.6	0	0	-	0	-	-
21	1,000,000	5,000	20	15K	497.52	187.2	477.36	183.6	-	36.72	-	400.36
22	1,000,000	5,000	200	20K	473.40	187.2	471.24	183.6	-	36.72	-	394.24
23	100	0	20	30K	2.16	3.6	0	0	-	0	-	-

TABLE 3-II (Concluded)

SNUMB	NCPU	NIO	LINKS	CORE	CPU SEC	I/O SEC	N1	N2	T1	T2	T3	T4
24	1,000	0	20	40K	2.52	3.6	0.36	0	0.360	–	–	0.36
25	1,500	0	20	50K	2.52	3.6	0.36	0	0.240	–	–	0.36
26	1	33,544	200	50K	501.48	1000.8	499.32	997.2	–	29.73	–	-17.26
27	1	33,544	20	50K	532.44	867.6	530.28	864.0	–	25.76	–	13.70
28	0	33,544	200	15K	526.68	1004.4	524.52	1000.8	–	29.84	15.64	7.94
29	0	0	20	20K	2.16	3.6	0	0	–	–	–	–
30	2,000	0	20	30K	2.88	3.6	0.72	0	0.360	–	–	0.72
31	2,500	0	20	40K	2.88	3.6	0.72	0	0.288	–	–	0.72
32	3,000	0	20	50K	2.88	3.6	0.72	0	0.240	–	–	0.72
33	10,000	0	20	15K	6.12	3.6	3.96	0	0.369	–	–	3.96
34	15,000	0	20	20K	7.92	3.6	5.76	0	0.384	–	–	5.76
35	20,000	0	20	30K	9.72	3.6	7.56	0	0.378	–	–	7.56
36	50,000	0	20	40K	21.96	3.6	19.80	0	0.396	–	–	19.80
37	100,000	0	20	50K	40.68	3.6	38.52	0	0.385	–	–	38.52
38	4,500	0	20	15K	3.96	3.6	1.80	0	0.400	–	–	1.80
39	10,000	15,000	200	20K	231.48	547.2	229.32	543.60	–	36.24	–	1.68
40	10,000	15,000	200	30K	244.08	374.4	241.92	370.80	–	24.72	–	10.92
41	10,000	15,000	20	40K	246.24	414.0	244.08	410.40	–	27.36	–	13.08
42	10,000	15,000	200	50K	217.08	345.6	214.92	342.00	–	22.80	–	-16.08
43	100,000	10,000	200	15K	199.08	241.2	196.92	193.32	–	19.33	–	–
44	100,000	10,000	20	20K	198.72	295.2	196.56	192.96	–	19.30	–	–
45	100,000	10,000	20	30K	189.36	252.0	187.20	183.60	–	18.36	–	–
46	100,000	10,000	20	40K	186.48	237.6	184.32	180.72	–	18.07	–	–
47	100,000	10,000	20	50K	186.48	223.2	184.32	219.60	–	21.96	–	30.32
48	1,000,000	5,000	20	15K	476.28	144.0	474.12	140.40	–	28.08	–	397.12
49	1,000,000	5,000	200	20K	461.16	133.2	459.00	129.60	–	25.92	–	382.00
50	1,000,000	5,000	20	30K	462.60	122.4	460.44	118.80	–	23.76	–	383.44
TOTAL									3.20	992.44	184.75	
AVERAGE									0.36	27.57	15.40	

N1 = CPU time for execution of m CPU loops (seconds)

N2 = I/O time for execution of n I/O loops (seconds)

m = NCPU

n = NIO

The amount of CPU time accumulated from I/O loops (T3) was judged to be too large to allow generation of I/O intensive workloads; so an attempt was made to lower the value of T3 before proceeding with the workload generation.

Previously, it had been suspected that the large value of processor time per I/O loop might be due to frequent execution of MME (master mode entry) GEMORE to handle the growth in file size. To check this point, tests were run in which the initial file space allocated was changed from 20 links to 200 links. The results in Table 3-III indicate no significant trend in the processor time for the two cases. However, small reductions were noted in most cases; therefore, the synthetic job was changed to request 200 links instead of 20 links in order to help lower the undesired CPU time resulting from I/O loops.

During the next test period, a 26 job synthetic workload was executed to check the effects of the recalibration. The results of this test are presented in Table 3-IV. Also, a test was run to determine the effect of changing the I/O loop of the synthetic job from disk to tape. Five jobs were run separately with one of the four files (File 04) assigned to tape. The results are shown in Table 3-V. The table also shows a comparison of the processor time for these jobs when File 04 was assigned to disk. The results show no significant change in the processor time for the two cases.

TABLE 3-III

INITIAL FILE SPACE ANALYSIS

SNUMB	LINKS	NCPU	NIO	CPU SEC	I/O SEC	Δ CPU	Δ I/O
21	20	1,000,000	5,000	479.52	187.2	——	——
22	200	1,000,000	5,000	473,40	187.2	- 6.12	0.0
48	20	1,000,000	5,000	476.28	144.0	——	——
49	200	1,000,000	5,000	461.16	133.2	-15.12	- 10.8
16	20	100,000	10,000	199.08	280.8	——	——
17	200	100,000	10,000	192.24	363.6	- 6.84	+ 82.8
44	20	100,000	10,000	198.72	259.2	——	——
43	200	100,000	10,000	199.08	241.2	+ 0.36	- 18.0
11	20	10,000	15,000	246.60	414.0	——	——
12	200	10,000	15,000	235.80	543.6	-10.80	+129.6
41	20	10,000	15,000	246.24	414.0	——	——
40	200	10,000	15,000	244.08	374.4	- 2.16	- 39.6
6	20	1,000	25,000	401.04	680.4	——	——
7	200	1,000	25,000	389.52	727.2	-11.52	+ 46.8
27	20	1	33,544	532.44	867.6	——	——
26	200	1	33,544	501.48	1000.8	-30.96	+133.2
1	20	1	35,000	559.80	1004.4	——	——
2	200	1	35,000	545.04	946.8	-14.76	- 57.6

TABLE 3-IV

TEST RESULTS CALIBRATION CHECK

SNUMB	NIO	NCPU	Proc. Time Seconds	I/O Time Seconds
1	0	0	1.80	3.6
2	250	0	5.76	7.2
3	500	0	9.36	14.4
4	1000	0	16.92	28.8
5	1500	0	24.48	32.4
6	2000	0	32.04	57.6
7	2500	0	39.96	75.6
8	3000	0	46.80	97.2
9	10000	0	151.56	295.2
10	20000	0	300.60	648.0
11	33544	0	512.28	950.4
12	50000	0	760.32	1195.2
13	100000	0	1524.96	2718.0
14	0	10	1.80	3.6
15	0	100	2.16	3.6
16	0	1000	2.52	3.6
17	0	1500	2.52	3.6
18	0	2000	2.88	3.6
19	0	2500	2.88	3.6
20	0	3000	3.24	3.6
21	0	4500	3.60	3.6
22	0	10000	5.76	3.6
23	0	15000	7.92	3.6
24	0	20000	9.72	3.6
25	0	50000	21.24	3.6
26	0	100000	41.40	3.6

TABLE 3-V

COMPARISON OF PROCESSOR TIME FOR
DISK AND TAPE ASSIGNMENT FOR FILE 04

SNUMB	NIO	NCPU	Proc. Time Seconds	Assignment of File 04
8	3000	0	46.80	Disk
			46.08	Tape
9	10000	0	151.56	Disk
			149.78	Tape
10	20000	0	300.60	Disk
			303.06	Tape
11	33544	0	512.28	Disk
			507.35	Tape
12	50000	0	760.32	Disk
			747.85	Tape

Next, the FORTRAN I/O loop of the synthetic program was replaced by a GMAP I/O loop in an attempt to lower the amount of CPU time needed to execute an I/O loop. The results of this modification are presented in Table 3-VI and were considered generally favorable in reducing the CPU time per I/O loop. These test results were used to recalculate the coefficients of the calibration equations. The resulting coefficients, listed below, were used throughout the synthetic test workload generation.

C1 = 1.80 seconds
C2 = 3.60 seconds
T1 = 0.34 milliseconds
T2 = 13.82 milliseconds
T3 = 4.41 milliseconds

3.1.2 Workload Generation and Testing

Using the synthetic workload generator, a series of synthetic test workloads were produced and tested at CCTC and AFDSDC. The test workloads included twelve synthetic workloads generated from site SCF data and five synthetic workloads constructed by selecting a series of synthetic programs with values of CPU time and I/O time which would produce desired IO/CP ratios for the workload.

Four synthetic test workloads generated by Hq. SAC and three synthetic test workloads generated by PACOM were executed to ensure proper operation. Three synthetic test workloads were generated at CCTC from SCF tapes supplied by Hq. MAC and were executed to ensure proper operation, but were not used in subsequent testing because the synthetic workloads were significantly different than the real workloads. This difference stemmed from the failure of the synthetic batch jobs to properly represent the transaction processing workload on the MAC machines.

TABLE 3-VI

TEST RESULTS FINAL CALIBRATION

SNUMB	NIO	NCPU	Proc. Time Seconds	I/O Time Seconds
1	0	0	1.80	3.6
2	250	0	2.88	3.6
3	500	0	3.96	10.8
4	1000	0	6.12	18.0
5	1500	0	8.28	21.6
6	2000	0	10.44	28.8
7	2500	0	12.60	32.4
8	3000	0	15.12	43.2
9	10000	0	46.44	133.2
10	20000	0	92.52	298.8
11	33544	0	151.20	550.8
12	50000	0	226.44	802.8
13	100000	0	472.68	1393.2
14	0	10	1.80	3.6
15	0	100	1.80	3.6
16	0	1000	2.16	3.6
17	0	1500	2.16	3.6
18	0	2000	2.52	3.6
19	0	2500	2.52	3.6
20	0	3000	2.88	3.6
21	0	4500	3.24	3.6
22	0	10000	5.40	3.6
23	0	15000	7.20	3.6
24	0	20000	9.36	3.6
25	0	50000	20.52	3.6
26	0	100000	39.24	3.6

3.1.3 Software Monitors

During the pilot testing, three software monitors were installed on the test system for possible use during primary and secondary testing. The monitors are described in section 2.2.2. Since the Mass Store Monitor (MSM) and the Memory Utilization Monitor (MUM) are mutually exclusive, it was decided to use MSM instead of MUM for the relative throughput tests.

The SYRUP-II monitor obtained from SAC was installed and several test runs were made to determine which sampling rate to use. Sampling rates of 0, 5, 10, and 15 seconds were all investigated. A sampling of 10 seconds was selected for use during the remainder of the testing. The SYRUP-II collector was run concurrently with both MSM and MUM to check for incompatibilities and none was discovered.

The collector portions of all three monitors appeared to collect the expected data once the initial installation problems were resolved. All three collectors stored the data on magnetic tape for subsequent data reduction.

3.1.4 Hardware Monitor

During the pilot testing at CCTC, an attempt was made to further instrument the test system with a hardware monitor. The hardware monitor available for use on this task was a Tesdata 1185. The purpose of the hardware monitor was to collect data, which could not be obtained with software monitors, on memory contention and memory reference patterns for the synthetic workloads.

The monitor was connected to the test system at CCTC with the assistance of HIS Field Engineering. The monitor was connected to probe points obtained from the HIS Probe Point Manual. Several attempts were made to collect information regarding SCU port contention during the execution of a synthetic workload. None of the attempts was successful.

During the installation of the hardware monitor, two special null (NOP) instructions were installed in the synthetic program. These instructions are only intended to be used for hardware monitoring and do not appear in the normal programming manuals. One NOP was placed as the first instruction of the synthetic program to be executed and the other NOP was placed as the last instruction to be executed. The placement of the NOP's was intended to allow correlation of software monitor and hardware monitor outputs.

After determining that the available monitor did not have sufficient capacity to measure the number of probe points necessary, with enough resolution to be of value, the hardware monitoring effort was terminated.

3.1.5 Data Reduction

The data reduction portion of each of the three software monitors was exercised during the pilot testing. The primary purpose was to become familiar with the format and use of the controls in the data reduction packages and with the format and content of the generated reports. This effort was intended to refine the list of data to be extracted by the data reduction programs as well as the data to be extracted from the GSEP reports. The final determination of data to be reduced is presented in section 3.2.4.

During the pilot testing is was decided that the value (at least for this task) of the SYRUP-II output reports would be significantly increased if the average value of each item were printed in addition to the instantaneous value at each sampling period. This enhancement to SYRUP-II was obtained from Hq. SAC.

Also, during the pilot testing a procedure for archiving raw data was developed. After each of the magnetic tapes produced by the collector portion of the software monitors was processed by the appropriate data reduction program, the tapes were consolidated

onto full reels of like-data at 1600 bits-per-inch. A full reel was made up of varying numbers of files which could be selectively retrieved at a later time. As the tapes were consolidated, a file directory log was maintained in order to allow subsequent location of specific test cases.

3.1.6 Timing Runs

Several timing runs for various test workloads on various configurations were accomplished during the pilot testing. The timing runs uncovered a potential testing problem with the system and workload interaction related to the use of temporary files and the security module. Also, the timing runs suggested a possible configuration alternative, related to crossbarring of the IOM's and MPC's, that could lead to improved throughput for certain types of workloads.

Test workload 7 was run on configuration A (1 CPU, 1 SCU, 256K, standard I/O) with an elapsed time of 55.32 minutes. The same test workload executed on a 1 CPU,2 SCU, 384K, standard I/O, configuration had an elapsed time of 55.44 minutes. And, the same test workload executed on a 2 CPU, 2 SCU, 384K, standard I/O, configuration had an elapsed time of 54.24 minutes. On a 3 CPU, 3 SCU, 384K, standard I/O, configuration, test workload 7 had an elapsed time of 56.34 minutes. A subsequent run of this same workload during the tests in Phoenix had an elapsed time of 34.44 minutes, on a 1 CPU, 1 SCU, 256K, standard I/O, configuration. The elapsed times indicate no substantial improvement with larger configurations during the pilot testing, but approximately 38% improvement during the Phoenix tests. This improvement is the result of two changes in the test environment. First, in Phoenix the security module was disabled. Second, in Phoenix the IOM's and MPC's were crossbarred. These conditions were not the case during the pilot testing, but the test workload was identical in all cases.

Based on the elapsed times from the pilot testing it was obvious that something was wrong. After consultation with Honeywell, it was determined that the interaction of the security module and the temporary files used by the synthetic program was causing job initiation to be too slow to maintain an adequate MPD. As a result, it was decided to disable the security module for the remainder of the throughput tests. More discussion of this subject is included in section 3.2.2.

The crossbarring of the IOM's and MPC's was not physically possible during the pilot testing at CCTC, but was possible in Phoenix because the equipment was being specially set up for the throughput tests. It was decided that the crossbarring would have a favorable effect on the elapsed times of all test cases, thereby allowing more test runs. Therefore, during the Phoenix tests, the IOM's and MPC's were crossbarred.

Under the same test environments, test workload 8 indicated similar responses as shown below:

CPU	SCU	CORE	ELAPSED TIME	LOCATION
1	1	256K	68.46 minutes	Reston, Virginia
3	3	384K	68.52 minutes	Reston, Virginia
1	1	256K	51.60 minutes	Phoenix, Arizona
3	3	394K	35.46 minutes	Phoenix, Arizona

Because this task was not specifically investigating the throughput improvements related to the two differences in the test environment, the test cases were not structured in a manner to allow separation of the improvement between the crossbarring and the security module.

3.1.7 Test Procedures

The test procedures listed in section 3.2.3 were developed and refined during the pilot testing. It was determined that the SYSOUT from the synthetic jobs contained no information of value that was not also available through GSEP. Hence, there was no need to print SYSOUT from every synthetic job. Consequently, the PURGE ALL command was added to the test procedures to inhibit printing of any output from the synthetic jobs.

All of the procedures in 3.2.3 were checked and verified prior to the primary testing in Phoenix.

3.2 Primary Testing

The majority of the test cases were successfully conducted at the Honeywell Information Systems Inc. facility on Black Canyon Drive in Phoenix, Arizona during the period of 10 May - 21 May 1976. Using twenty-five different hardware configurations and fourteen different synthetic workloads, a total of 245 test cases was run. The test cases using H6060 configurations used a total of 106 hours of computer time and test cases using H6080 configurations used a total of 63 hours of computer time.

This section reports on the conditions under which the tests were conducted and the data collected during the tests. The test configurations are presented in section 3.2.1, the test workloads in section 3.2.2, and the test conduct in section 3.2.3. Finally, the preliminary test results are included in section 3.2.4.

3.2.1 Test Configurations

A total of twenty-five different H6000 hardware configurations was used during the Phoenix testing. The test plan called for 32 different configurations, but seven configurations were not run because of time limitations. Of the planned 32 test configurations,

16 were H6060 configurations and 16 were H6080 configurations. In each case, there were eight configurations with memory interlace turned off and the same eight configurations with memory interlace turned on. The twenty-five test configurations completed were as follows:

a. H6060 memory interlace off - 6
b. H6060 memory interlace on - 7,
c. H6080 memory interlace off - 4,
d. H6080 memory interlace on - 8.

The eight configurations ranged from one CPU to four CPU's and from 256K to 1024K words of core memory. All configurations had exactly the same complement of I/O devices and channel assignments.

Each test case was identified by a two-character code which denoted the type of CPU (H6060 and H6080) and the specific configuration. The coding used is shown below and is used throughout this paper to identify test data and test results.

Table 3-VII
Configuration Identifiers

H6060	H6080	# CPUs	# SCUs	# WORDS OF CORE MEMORY
6A	8A	1	1	256K
6B	8B	1	2	512K
6C	8C	2	2	256K
6D	8D	2	3	512K
6E	8E	3	3	384K
6F	8F	3	4	768K
6G	8G	4	3	768K
6H	8H	4	4	1024K

The schematic diagrams of the eight configurations are shown in Figures 3-2 through 3-9. In all cases, the IOM's and MPC's were crossbarred and the MPC's and disk drives were crossbarred in order to provide the maximum number of independent I/O transfer paths between SCU's and the disk drives.

The system software was essentially the same for all cases. The operating system was GCOS system release WW 6.2.1 with an octal patch to disable the WWMCCS security module (FS 49) and a patch to enable the operator interface to one of the software monitors (SYRUP-II). The security module was disabled in order to eliminate unrealistic delays in job initiation resulting from the total use of temporary files by the synthetic workloads. This point will be discussed further in section 3.2.2.

Two software monitors were used during all test cases in addition to the normal GCOS Statistical Collection File (SCF). The primary software monitor was the System Resource Utilization Package (SYRUP-II release 2.2) provided by Hq. SAC and the secondary monitor was the Mass Store Monitor (MSM) provided by CCTC/WAD. System resources utilization data collected by the software monitors is delineated in section 3.2.4.

3.2.2 Test Workloads

During the WMPE tests, a series of fourteen different synthetic workloads was used to measure elapsed time and pertinent resource utilization. Each synthetic workload was composed of approximately fifty synthetic programs. Each synthetic program was either a unique SNUMB within the jobstream or a unique ACTIVITY within a job (SNUMB).

The synthetic programs which made up the workloads were all constructed using the Buchholz[8] yardstick-program concepts.

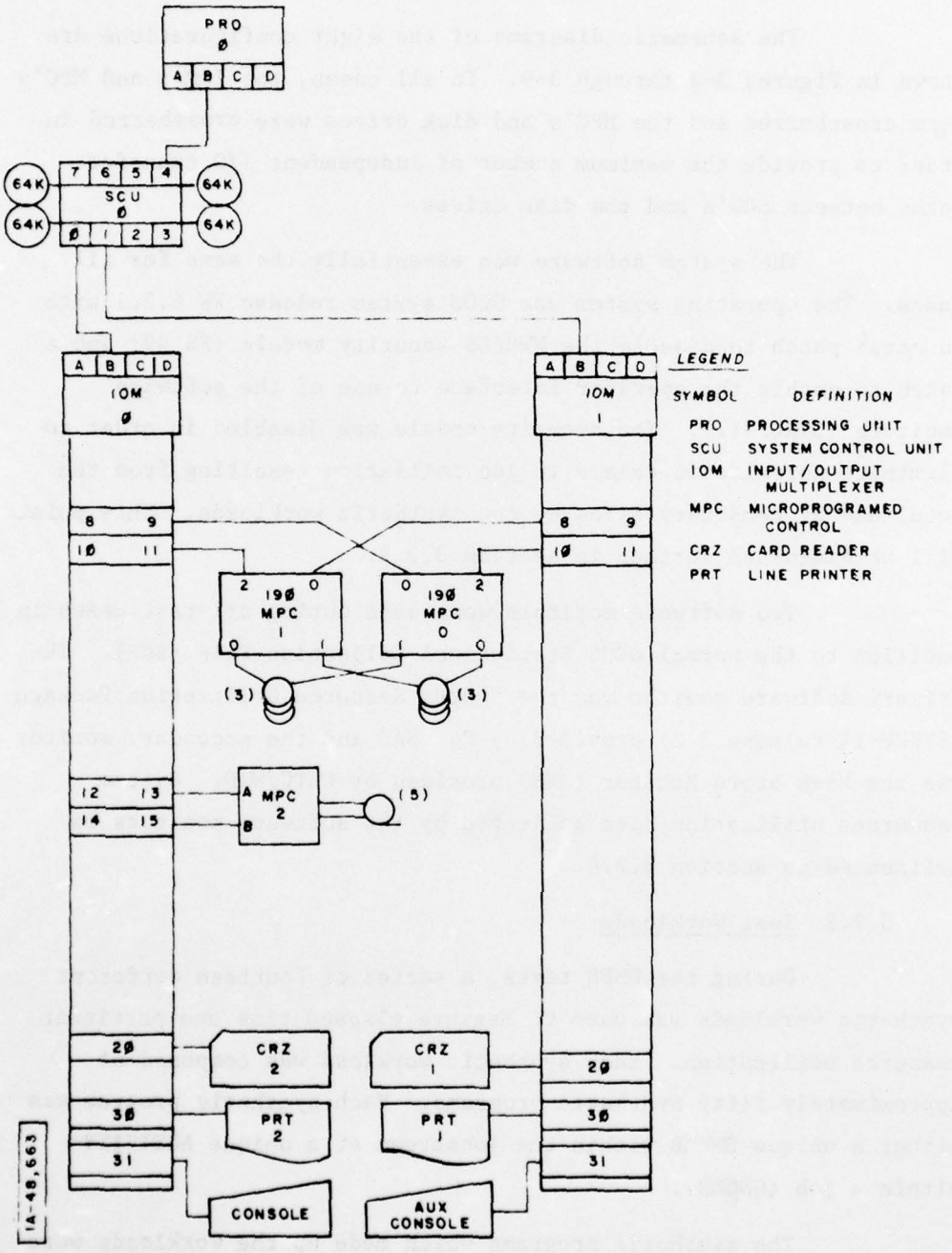


Figure 3-2. Configuration A

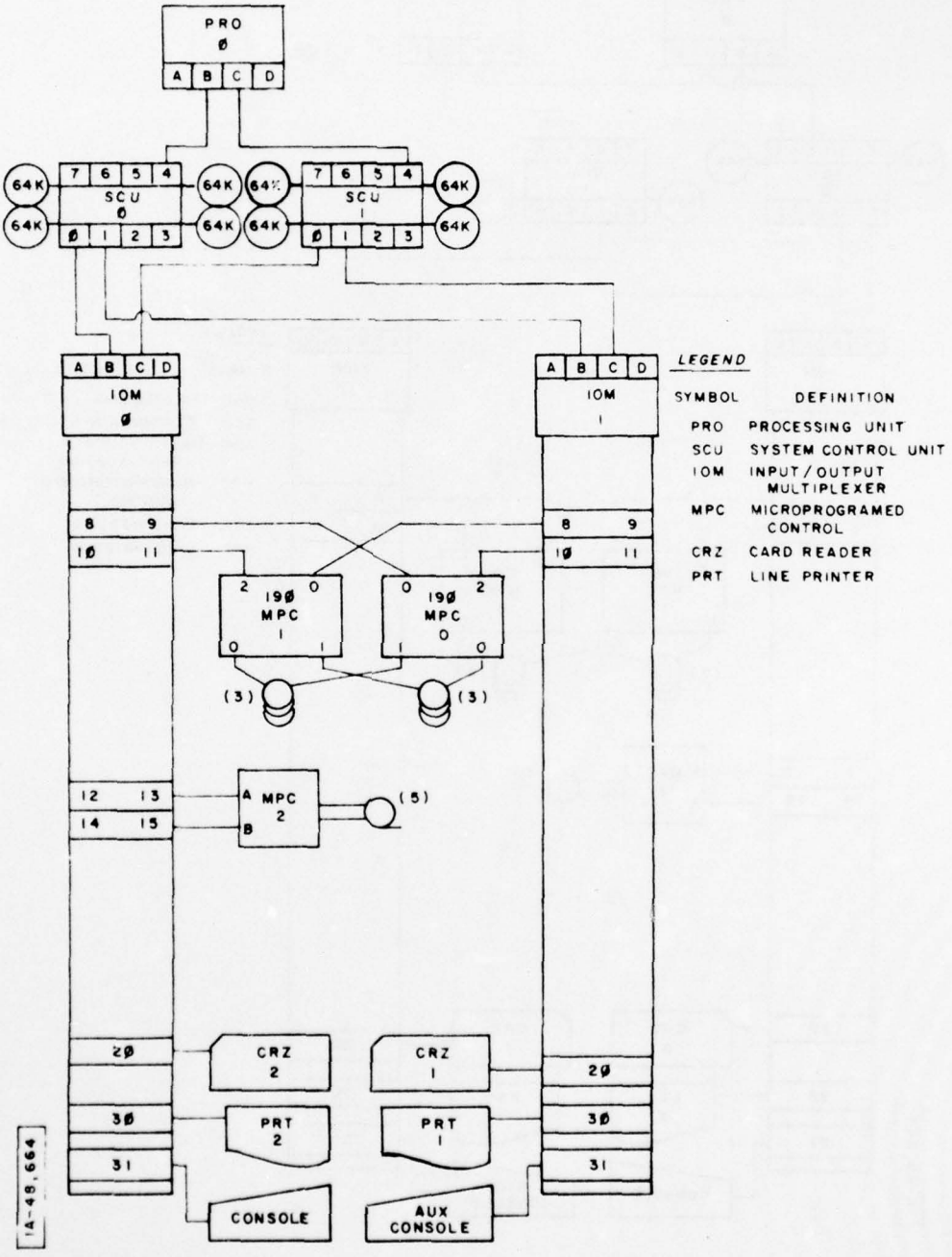


Figure 3-3. Configuration B

AD-A039 111 MITRE CORP BEDFORD MASS F/G 9/2
WWMCCS H6000 MULTIPROCESSOR PERFORMANCE EVALUATION. VOLUME I.(U)
FEB 77 G A NELSON F19628-77-C-0001
UNCLASSIFIED MTR-3350-VOL-1 ESD-TR-77-18-VOL-1 NL

2 OF 2
AD A039111



END

DATE FILMED
5-77

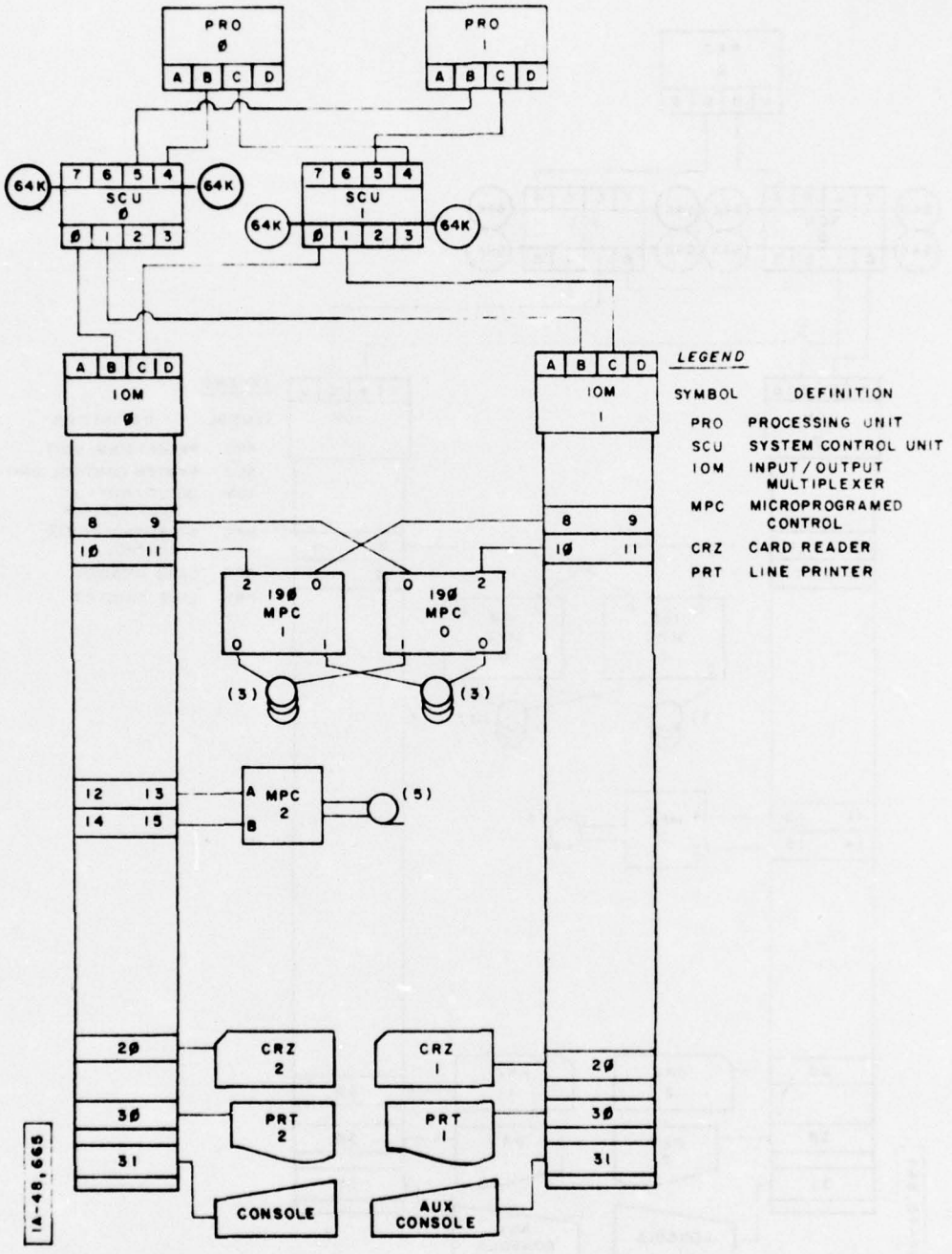


Figure 3-4. Configuration C

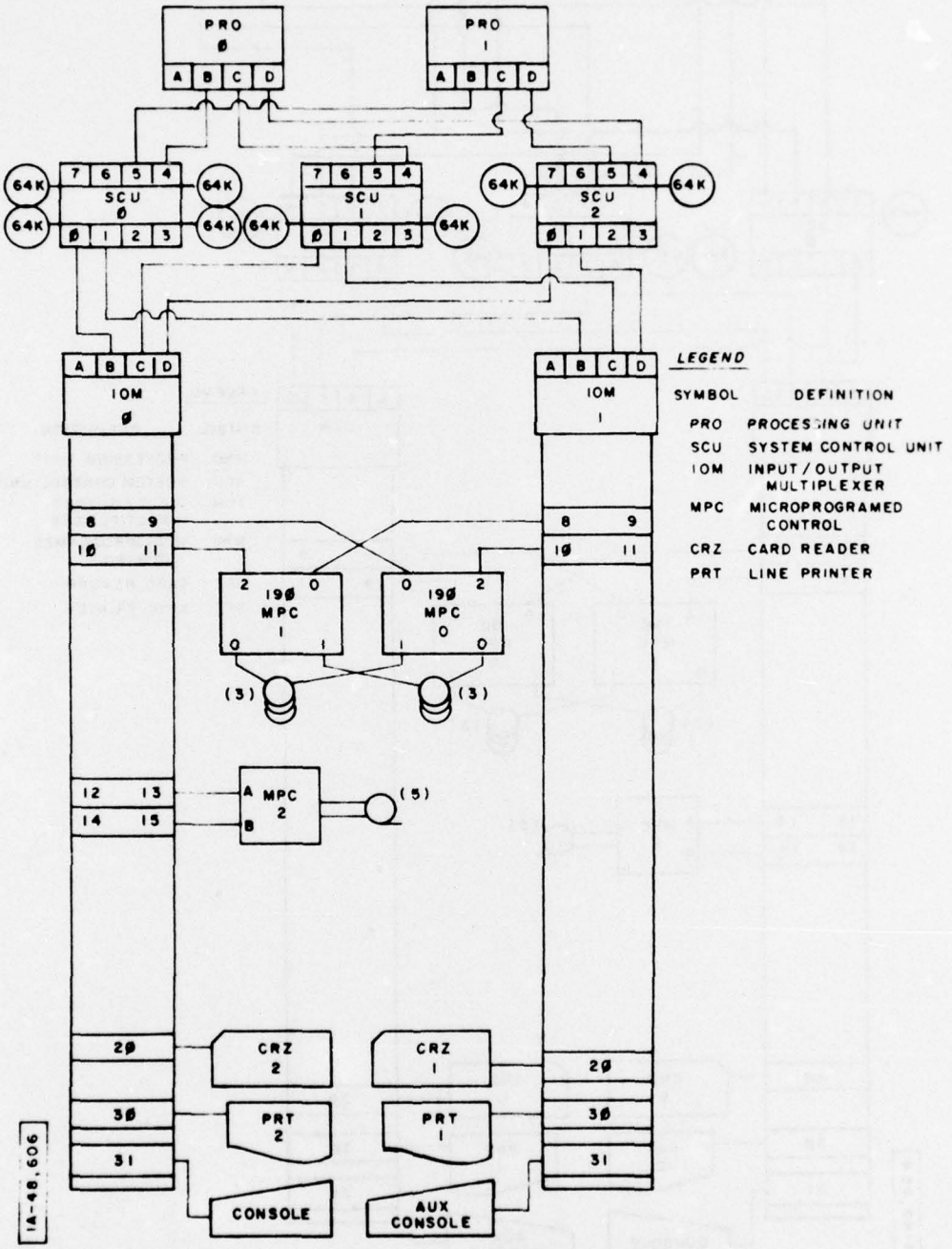


Figure 3-5. Configuration D

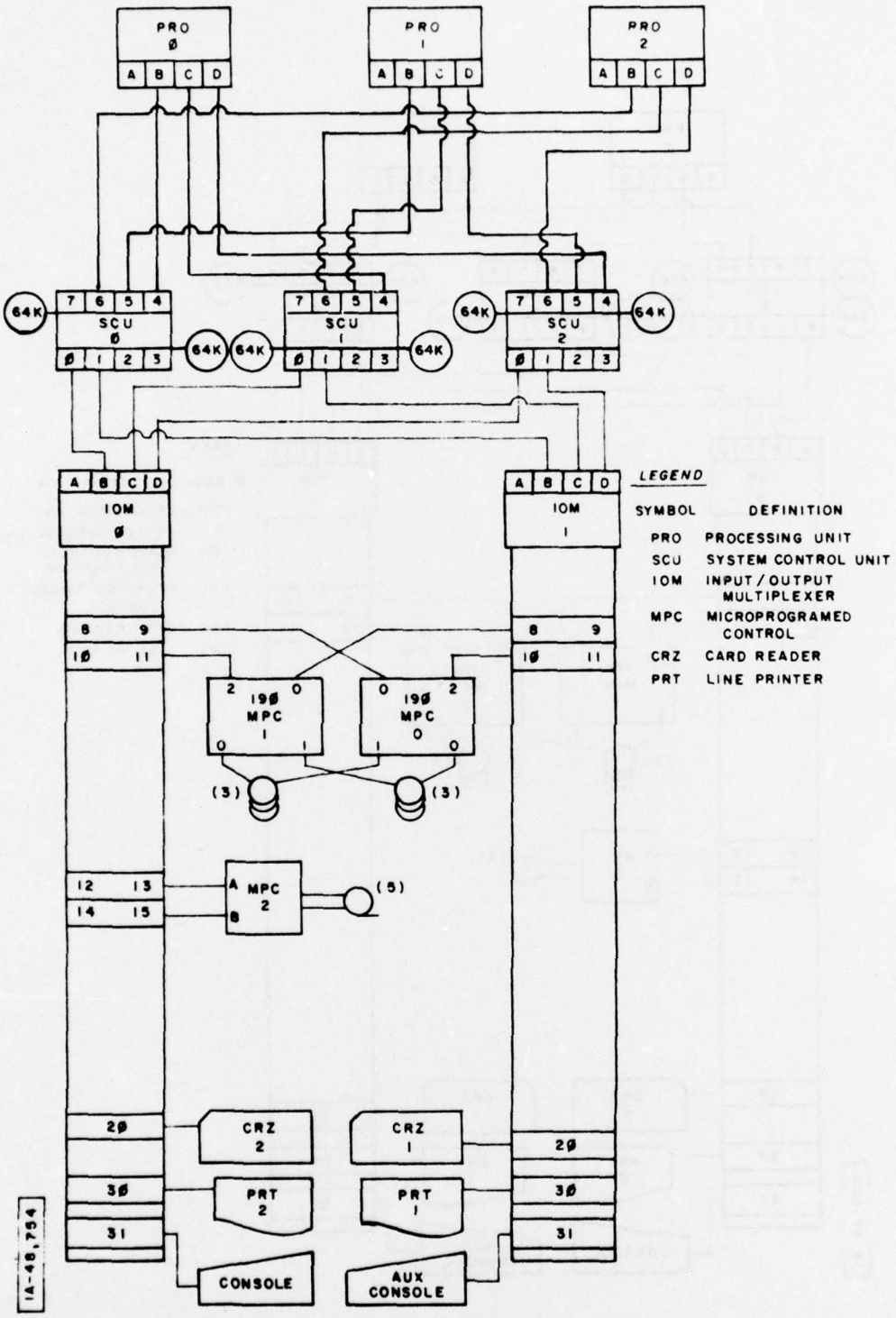
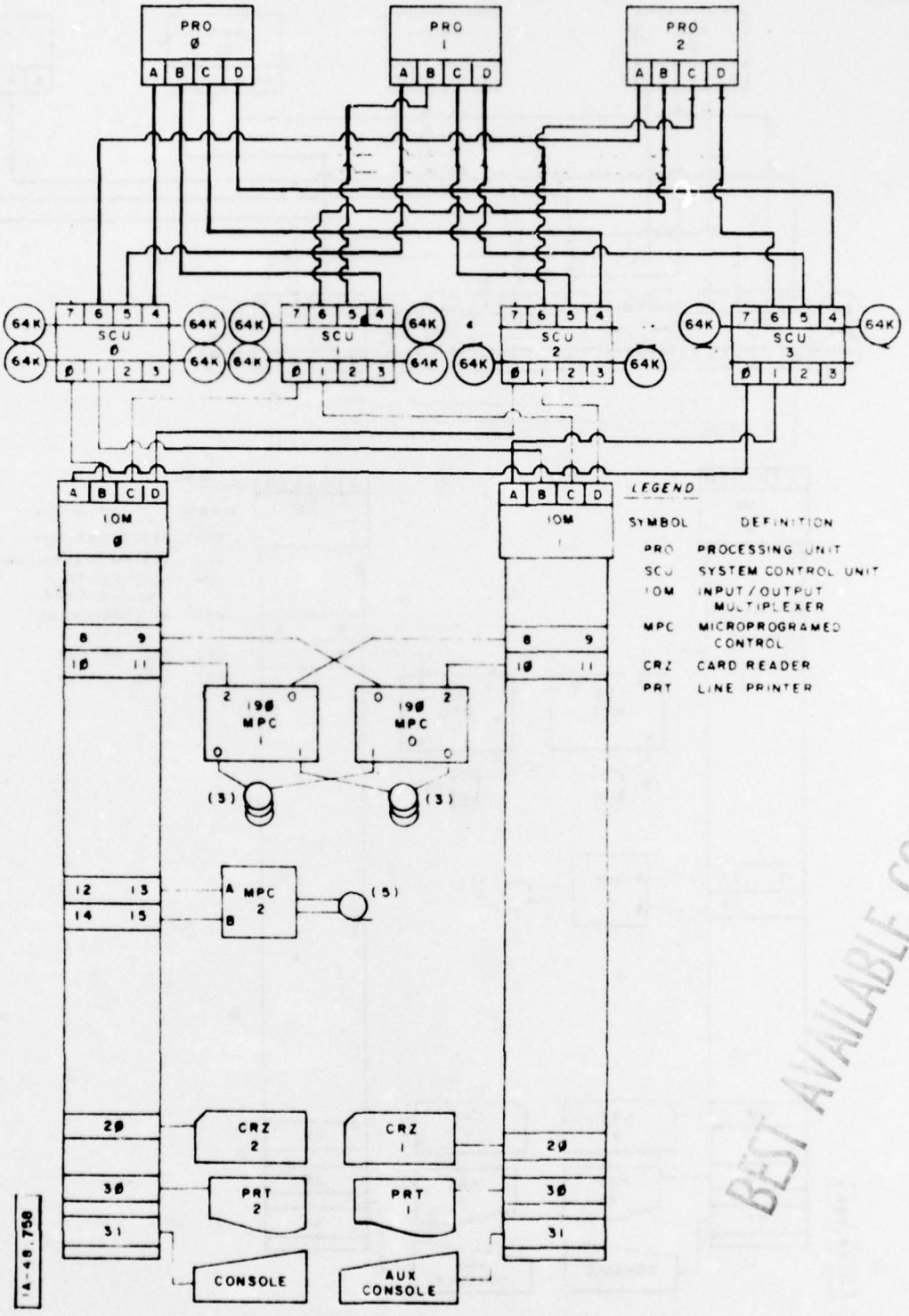


Figure 3-6. Configuration E



BEST AVAILABLE COPY

Figure 3-7. Configuration F

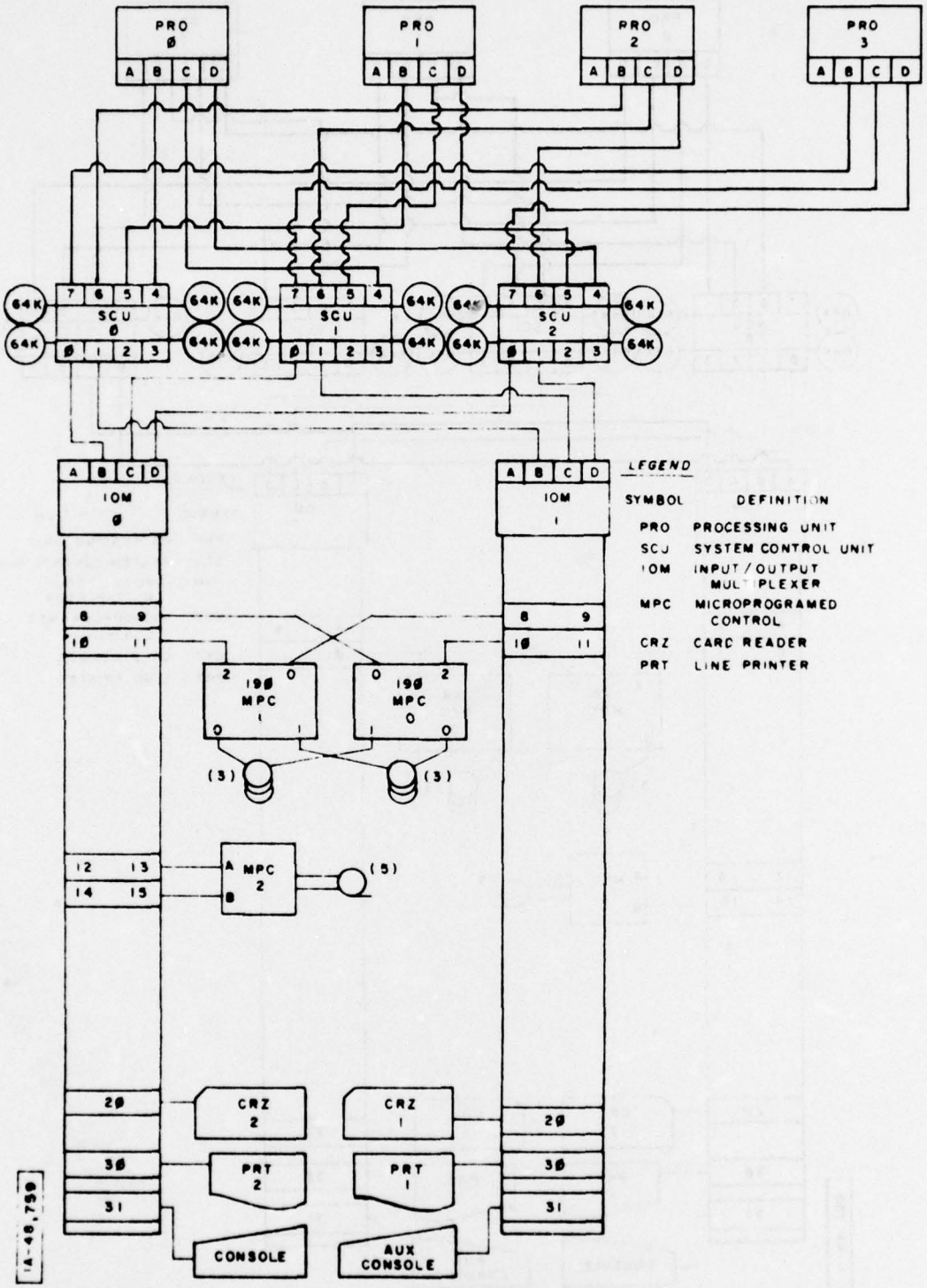


Figure 3-8. Configuration G

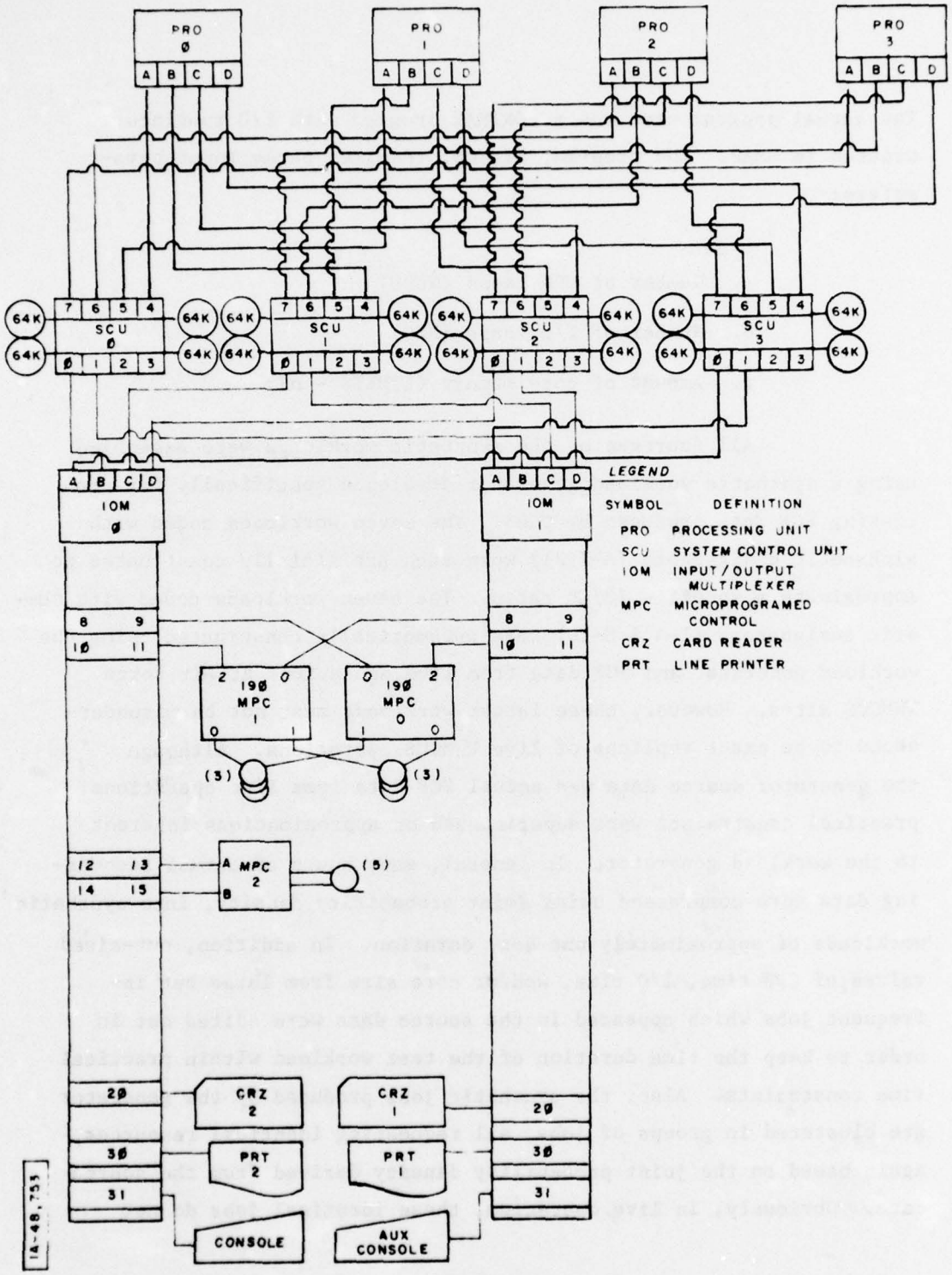


Figure 3-9. Configuration H

The actual program used was a FORTRAN program with I/O routines written in GMAP. The program, SYNJOB, required three input parameters:

a. Number of CPU loops (NCPU)

b. Number of I/O loops (NIO)

c. Amount of core memory (LIMITS - nK)

All fourteen of the synthetic workloads were assembled using a synthetic workload generator developed specifically for processing SCF data produced by GCOS. The seven workloads coded with alphabetic designators (A-F(2)) were each artificially constructed to approximate a specific IO/CP ratio. The seven workloads coded with numeric designators (1-4 & 8-10) were automatically constructed using the workload generator and SCF data from live operations at Air Force WWMCCS sites. However, these latter workloads must not be misunderstood to be exact replicas of live WWMCCS operations. Although the generator source data was actual SCF data from live operations, practical constraints were superimposed on approximations inherent to the workload generator. In general, many hours of actual accounting data were compressed using joint probability density, into synthetic workloads of approximately one-hour duration. In addition, out-sized values of CPU time, I/O time, and/or core size from large but infrequent jobs which appeared in the source data were edited out in order to keep the time duration of the test workload within practical time constraints. Also, the synthetic jobs produced by the generator are clustered in groups of jobs, all requesting identical resources, again based on the joint probability density derived from the source data. Obviously, in live operation, these identical jobs do not

necessarily cluster themselves. However, the seven workloads produced from real SCF data, although not exact replicas of site workloads, are more indicative of site workloads than the seven desk-designed artificial workloads.

During the pilot testing of the workloads, operating system and software monitors at CCTC, it was observed that job initiation was very slow. Consequently, jobs initiated early would complete before very many other jobs were started and the multiprogramming depth (MPD) never reached realistic values. Subsequently, it was determined that the slow initiation was caused by the temporary files that each job allocates in order to do I/O operations. The temporary files interacting with the WWMCCS security module, which obliterates all temporary files before and after use, were causing jobs to be delayed in initiation. Replacing the temporary files with permanent files was considered and rejected because of the enormous disk storage requirement. It was decided, instead, to offset the use of so many temporary files by disabling the routine that performs the cleaning action (i.e., FS 49). Implicit in this decision was the assumption that, overall, the disabled security module and the unrealistically large number of temporary files would roughly balance out any effect on the test results.

The characteristics of each workload are listed in Table 3-VIII. As the table shows, the workloads requested CPU and I/O resources from the system in varying amounts and were intended to provide performance data on a range of workloads from highly CPU-bound to moderately I/O-bound. The IO/CP ratio is the total channel time accumulated for a workload divided by the total processor time accumulated for the same workload.

All jobs within a workload were executed at equal priority. All fourteen workloads were run on each of sixteen different hardware configurations at least once.

TABLE 3-VIII

WORKLOAD CHARACTERISTICS

WORKLOAD IDENTIFIER	NO. OF JOBS	ACTIVITIES	NO. OF CPU LOOPS	PROC. TIME*	NO. OF I/O LOOPS	CHAN. TIME*	ELAPSED TIME*	IO/CP RATIO*
E	40	40	12,000,000	82.829	120	0.897	87.293	0.01
A	50	50	6,495,300	50.013	43,420	12.184	57.057	0.24
B	50	50	5,676,770	45.205	52,100	14.876	52.741	0.32
2	49	49	3,909,091	31.637	46,039	14.529	38.161	0.45
C	50	50	4,454,420	38.115	65,200	18.983	46.030	0.49
1	24	24	7,161,466	45.760	177,230	51.132	61.547	1.11
D	50	50	2,906,032	43.679	150,000	55.067	57.630	1.26
9	48	48	1,758,858	36.038	253,812	86.352	57.799	2.39
F(2)	18	48		9.998	(101,786)est.	29.635	19.350	2.96
F(1)	24	49		13.024	(135,342)est.	39.405	26.943	3.02
4	49	49	453,102	35.996	353,143	118.413	64.959	3.28
8	50	50	296,293	27.431	271,250	94.085	51.569	3.42
10	49	49	262,742	43.995	462,364	152.726	80.686	3.47
3	49	49	41,181	21.996	237,582	81.887	42.609	3.72

* H6060, 1 CPU, 256K, INTERLACE OFF

3.2.3 Test Conduct

The WMPE tests were conducted by a team of AFDSDC, ESD, HIS, MITRE, and SAC personnel at the Honeywell Information Systems, Inc. plant in Phoenix, Arizona. The chronology of testing is listed in Table 3-IX. The tests took place during the period of 10 May - 21 May 1976 on equipment specifically assembled in Phoenix for these tests. The test period began with a visual inspection of the equipment. The equipment layout was as shown in Figure 3-10.

The controlled copy of the operating system (GCOS SR WW 6.2.1) was loaded from a combination of the initialization (INIT) tape, a PERMFILE SAVE tape, and the start-up deck. Each time the system was restarted after a configuration change, the same materials were used.

A special calibration workload was run initially to determine the effect of disabling the FS 49 security module. This workload was run first with FS 49 enabled and then rerun with the FS 49 disabling patch applied. A patch was also made to SYRUP to allow it to operate without a Datanet 355 attached to the system.

Between each configuration change, including turning memory interlace on or off, the system was rebooted. The system was booted with the same start-up deck each time except for the CONFIG section, which had changes made for the MCT cards each time the number of CPU's or memory modules were changed. Appendix I in Volume II contains a listing of the start-up deck. After each reboot, the following procedures were used.

a. Spawn LSTCN & LSTCA from the console to obtain a listing of the system master catalog.

b. Enter TYPFG from the console to obtain a listing of the configuration.

Table 3-IX

Testing Chronology

DAY	DATE	DURATION	TIME	CONFIG	MEMORY INTERLACE	WORKLOADS
MON	5/10/76	16 Hrs.	0800-2400	6A	ON	A, B, C, D
TUES	5/11/76	24 Hrs.	0000-2400	6A	ON	E, F(1), F(2)
				6A	OFF	ALL
WED	5/12/76	24 Hrs.	0000-2400	6B	OFF	ALL
				6G	OFF	A, B, C, D, E, F(1), F(2)
THUR	5/13/76	22 Hrs.	0000-2200	6G	OFF	1, 2, 3, 4, 8, 9, 10
				6G	ON	A, B, C, D, E, F(1), F(2)
				6H	OFF	ALL
				6H	ON	A, B, C, D, E, F(1), F(2)
FRI	5/14/76	14 Hrs.	0000-2200	6E	OFF	ALL
				6E	ON	A, B, C, D, E, F(1), F(2)
SAT	5/15/76	12 Hrs.	0800-2200	6F	OFF	ALL
				6F	ON	A, B, C, D, E, F(1), F(2)
SUN	5/16/76	8 Hrs.	0800-1600	6D	ON	A, B, C, D, E, F(1), F(2)
				6C	ON	A, B, C, D, E, F(1), F(2)
MON	5/17/76	11 Hrs.	1300-2400	8H	ON	A, B, C, D, E, F(1), F(2)
				8H	OFF	ALL
TUES	5/18/76	24 Hrs.	0000-2400	8G	ON	A, B, C, D, E, F(1), F(2)
				8G	OFF	ALL
				8F	ON	A, B, C, D, E, F(1), F(2)
				8F	OFF	ALL

Table 3-IX

Testing Chronology (Concluded)

DAY	DATE	DURATION	TIME	CONFIG	MEMORY INTERLACE	WORKLOAD
WED	5/19/76	24 Hrs.	0000-2400	8E	ON	A, B, C, D, E, F(1), F(2)
				8E	OFF	ALL
				8D	ON	A, B, C, D, E, F(1), F(2)
				8C	ON	A, B, C, D, E, F(1), F(2)
				8B	ON	A, B, C, D, E, F(1), F(2)
THUR	5/20/76	4 Hrs.	0000-0400	8A	ON	A, B, C, D, E, F(1), F(2)

BEST AVAILABLE COPY

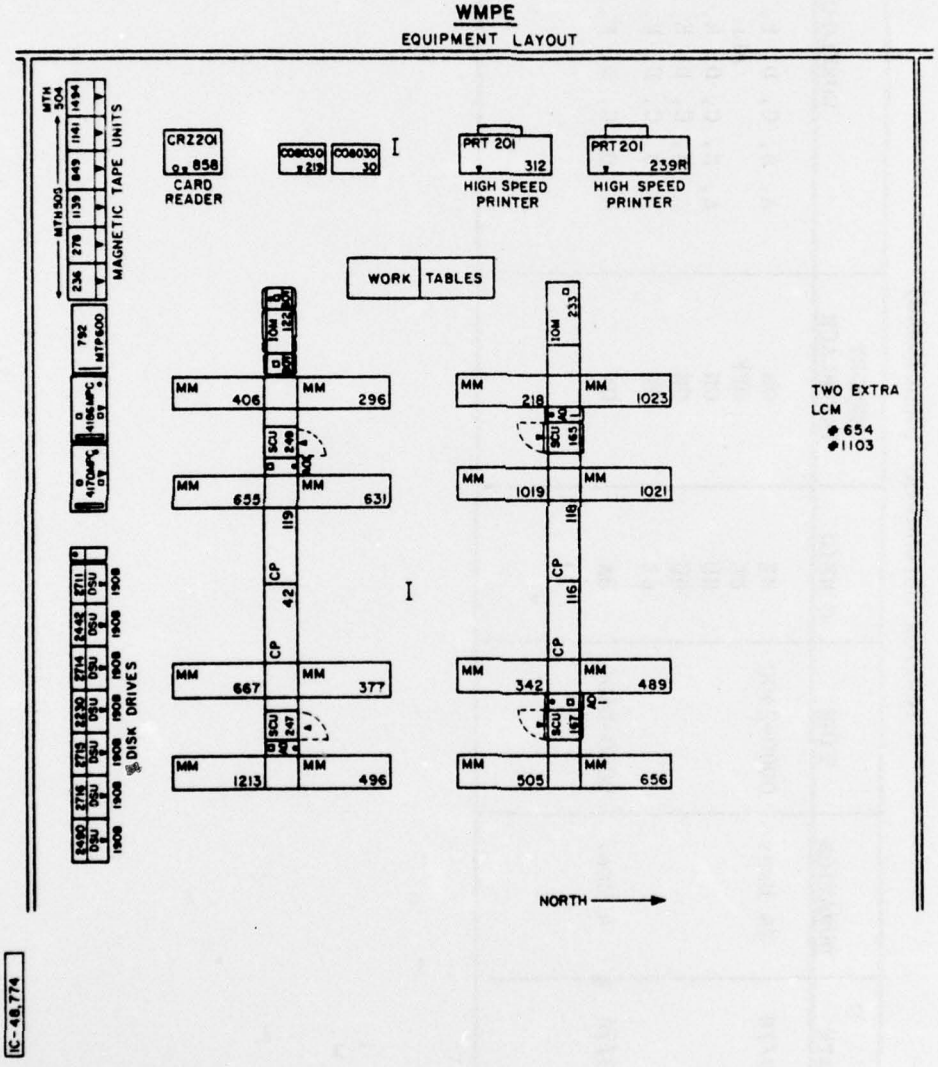


Figure 3-10. WMPE Equipment Layout

c. Enter LIMIT NONE from the console to disable any sieve limits.

d. Enter PURGE ALL from the console to inhibit printing of any output from the synthetic jobs.

Each workload was executed in sequence using the following procedures:

a. Spawn MSMØØ from the console to start MSM data collection.

b. Load SYRUP card deck into card reader.

c. Enter RUN SYRUP from the console to start SYRUP data collection. The SYRUP sampling interval was set for 10 seconds in all cases.

d. Load workload card deck into card reader to start jobs running without any further operator intervention. (Workloads F(1) and F(2) were run from IMCV tapes instead of card decks.)

e. No operator intervention from the console allowed unless required by a malfunction until workload had completed execution.

f. Enter LSTAL from console to verify that only MSM and SYRUP are still executing.

g. Enter TERM MSMØØ from the console to stop MSM data collection.

h. Enter SYRUP STOP MAIN and SYRUP STOP PERIPH from the console to end SYRUP data collection.

i. Enter TERM SYRUP from the console to terminate SYRUP execution.

j. Enter ACCNT from the console to terminate SCF data collection.

After each workload was completed, data-reduction programs were run for GESEP, MSM, and SYRUP. The raw data tapes were then consolidated onto master tapes for archival storage. A complete file directory of all archival tapes is available from the author.

3.2.4 Test Results

The data reduction reports from GESEP, MSM, and SYRUP-II were used to extract pertinent system data. The data selected for analysis are listed below:

a. GESEP

1. Total elapsed time for workload
2. Total processor time for workload
3. Total channel time for workload
4. Channel vs processor ratio for workload

b. SYRUP-II

1. Multiprogramming depth (maximum)
2. Multiprogramming depth (average)
3. Average percentage of processor utilization for system programs for each processor
4. Average percentage of processor utilization for user programs for each processor
5. Average percentage of IOM utilization for system programs for each IOM

6. Average percentage of IOM utilization for user programs for each IOM

7. Average percentage of processor time active for each processor

8. Average percentage of processor overhead time for each processor

9. Average memory used for each quadrant of memory

10. Average count of connects on each channel of each IOM

c. MSM

1. Total count of connects for workload only

All of the extracted data are presented in Appendix II of Volume II, Tables 1 through 64. The tables are organized with all workloads and all configurations within a group shown in matrix form on each table. There are four groups of tables divided as follows:

a. H6060, MEMORY INTERLACE OFF - Tables 1 through 17
b. H6060, MEMORY INTERLACE ON - Tables 18 through 32
c. H6080, MEMORY INTERLACE OFF - Tables 33 through 48
d. H6080, MEMORY INTERLACE ON - Tables 49 through 64

The conventions used in preparing the tables were as follows:

a. Data for this entry was either not applicable or the test case was not planned to be conducted -- denoted by three dashes.

b. Data for this entry was planned and was collected but lost during data reduction -- denoted by two asterisks.

c. Data for this entry was planned but the test case was not conducted because of time limitations -- denoted by blanks.

3.3 Secondary Testing

The secondary testing at CCTC, Reston, Va., was an unplanned outcome of the primary testing. Not all of the tests planned to be conducted in Phoenix were accomplished. A total of 245, out of a planned 336, test cases was completed in Phoenix. This resulted from an extremely ambitious test plan for the primary testing. Consequently, an attempt was made to complete the planned series of test cases after the primary testing.

While conducting the secondary testing, a problem was encountered with the setting of the system scheduler file. During the later part of the pilot testing, during all of the primary testing, and during the secondary testing, the .NORM queue for the system scheduler was set at a value of 35 jobs maximum via the configuration section of the bootload deck. However, during the secondary testing the data reduction program indicated that the setting was 20, rather than 35 as requested in the bootload deck. Once discovered, the erroneous setting was overriden using an available console command and subsequent tests were conducted with .NORM = 35. However, since the operating system continued to set .NORM = 20 during start-up and console action was required to reset .NORM = 35, it was concluded that the control copy of GCOS had become contaminated.

A total of 28 test cases was run during the secondary testing. Test cases for configuration 6C (interlace off) with all 14 test workloads were completed after resetting .NORM = 35. Test cases for configuration 6D (interlace off) with test workloads 8, 9, 10, A, B, C, D, E, F(1), and F(2) were completed after resetting .NORM = 35. Test cases for configuration 6D (interlace off) with test workloads 1, 2, 3, and 4 were completed with .NORM = 20 and were not rerun

During the secondary testing, the configurations were as shown in Figures 3-4 and 3-5 except that IOM's and MPC's were not crossbarred as shown. Instead, all four logical channels from IOM-0 were connected to MPC-0 and all four logical channels of IOM-1 were connected to MPC-1.

After concluding that any test results from the secondary testing would be questionable due to the uncontrolled behavior of the system scheduler and the difference in I/O configuration, the secondary testing effort was terminated

The results from all secondary testing are included in the overall test results; because of the anomaly with the system scheduler, results from these test cases are not considered to be reliable.

3.4 Data Analysis

Since the focus of this task was an attempt to determine the relative throughput for various WWMCCS H6000 system configurations, the analysis of the collected data was concentrated on the elapsed time, and changes in elapsed time, for the various test workloads for each test configuration. The baseline configuration in all cases was defined as a single H6000 CPU with 256K of core memory connected through one SCU, no memory interlacing, and with the previously defined standard I/O configuration. The choice of the baseline configuration was an arbitrary choice to make all relative throughput values positive and, hopefully, greater than unity.

Data extracted from the data reduction reports and contained in Tables 1 - 64 of Volume II, Appendix II were analyzed for two specific information types -- elapsed time, and relative throughput. Each of the areas is discussed below.

3.4.1 Elapsed Time

The elapsed time data contained in Volume II, Appendix II, Tables 1, 17, 33, and 49 were plotted in histogram form for graphical analysis. The resulting graphs are included in this section as Figures 3-11 through 3-66. The following three areas of interest were analyzed.

a. Effect on elapsed time as CPU's were added to the configuration.

b. Effect on elapsed time as core memory was added to the configuration.

c. Effect on elapsed time resulting from memory interlacing.

Figures 3-11 through 3-38 represent elapsed times for all fourteen test workloads on all eight H6060 configurations. The odd-numbered figures in this group show the impact of adding CPU's to the configuration and the effect of memory interlacing. The even-numbered figures show the impact of adding core memory to the configuration and the effect of memory interlacing. Qualitatively, the figures indicate sizable decreases in elapsed time for CPU-intensive workloads as the number of CPU's in the configuration is increased and small decreases in elapsed time for I/O-intensive workloads as the number of CPU's in the configuration is increased. These results are as expected. Additionally, the data tables in Volume II can be used to determine, quantitatively, the percentage improvement as resources are added to the configuration. For example, from Figure 3-16 and Table 1 of Volume II, Appendix II, Test Workload 2 shows an elapsed

time decrease from 38 minutes to 21 minutes, or 42% improvement, when a second CPU is added. The same workload exhibits a 33% improvement over the 2 CPU configuration when the third CPU is added and a 14% improvement over the 3 CPU configuration when the fourth CPU is added. Test Workload 2 is considered to be CPU-intensive and shows sizable improvement as more CPU's are added.

Test Workload 3, on the other hand, is considered to be I/O-intensive and shows only slight improvement as more CPU's are added. Improvements of 21%, 13%, and 9% were measured for Workload 3 on 2 CPU, 3 CPU, and 4 CPU configurations respectively.

Figures 3-39 through 3-66 represent elapsed times for all fourteen test workloads on all eight H6080 configurations. The arrangement of the figures is the same as the the H6060 figures discussed earlier. Also, the qualitative and quantitative results follow the same trend as the H6060 data. For both H6060 and H6080, most of the test cases indicate a 5% to 15% decrease in elapsed time with memory interlacing turned on. Some workload/configuration combinations show only 1% to 2% improvement and one combination shows a 28% increase in elapsed time (see Figure 3-58). The 28% increase in elapsed time seems to be anomolous; however, the data has been checked several times and appears to be correct. As shown in Figures 3-11 through 3-66, the impact of adding core memory to a configuration was generally small (i.e., less than 12% improvement).

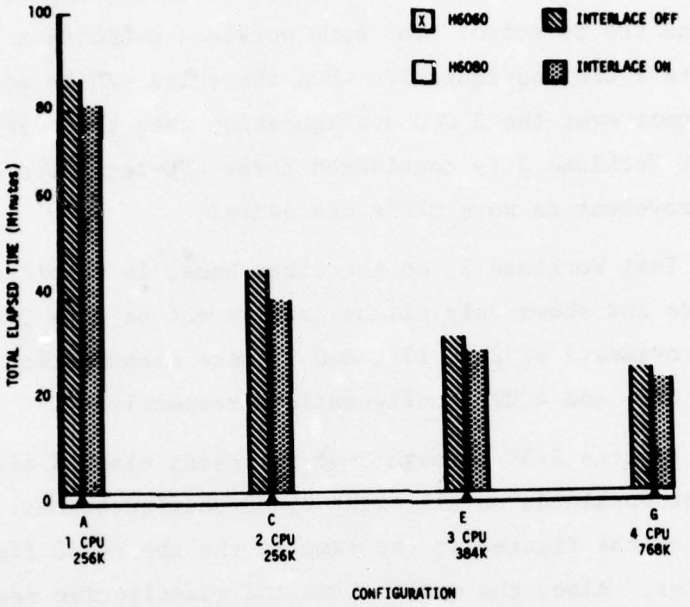


Figure 3-11. Workload E, CPU Impact

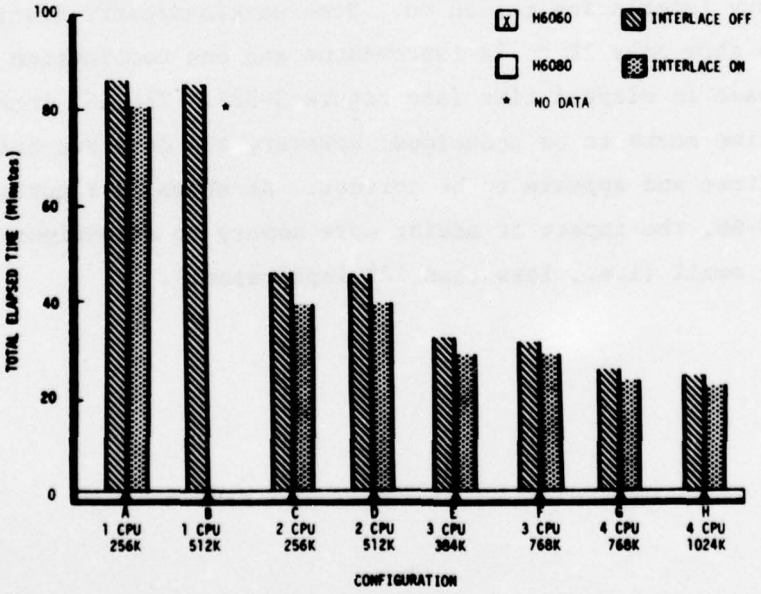


Figure 3-12. Workload E, Memory Impact

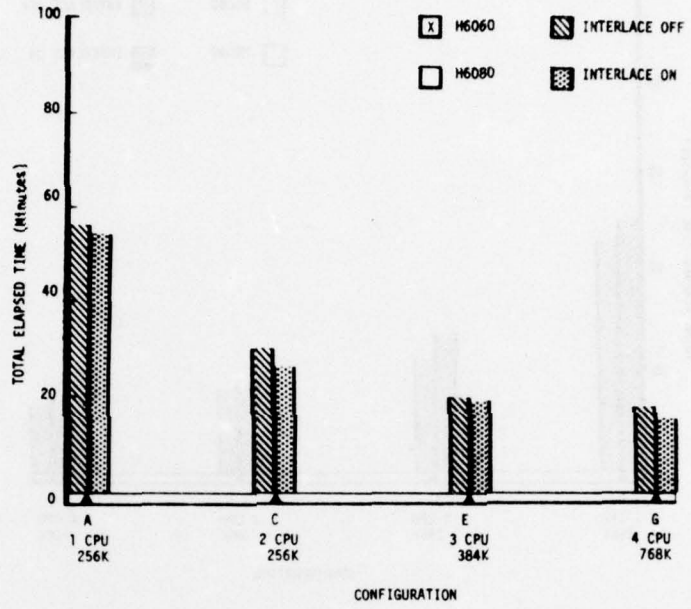


Figure 3-13. Workload A, CPU Impact

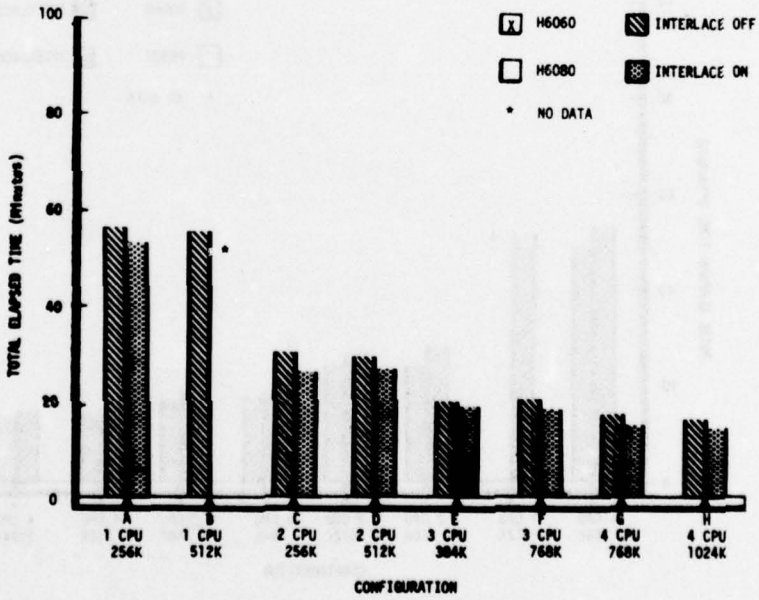


Figure 3-14. Workload A, Memory Impact

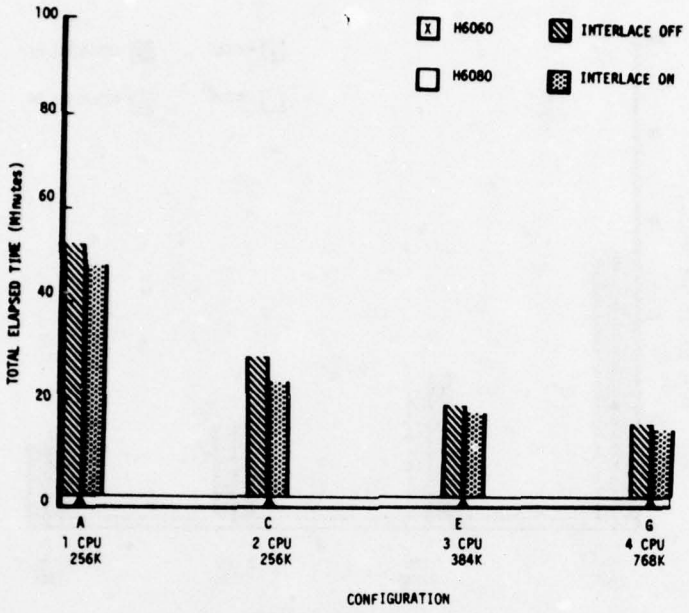


Figure 3-15. Workload B, CPU Impact

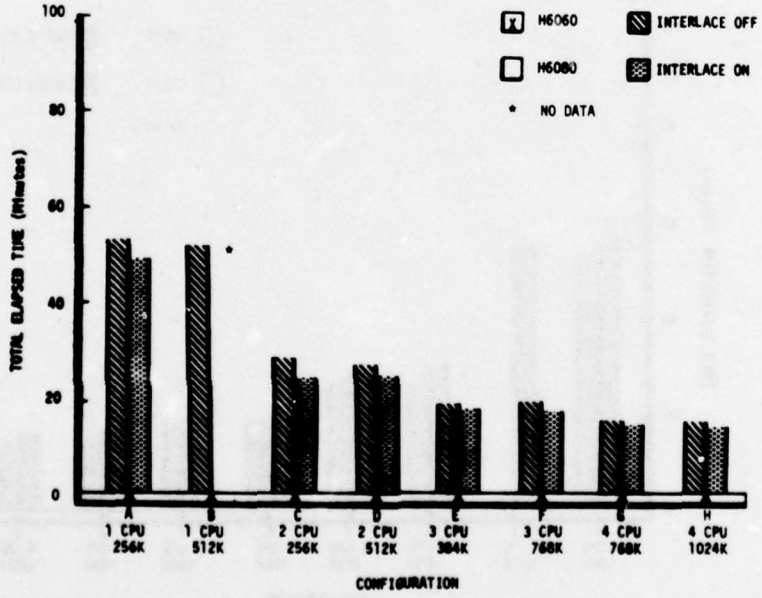


Figure 3-16. Workload B, Memory Impact

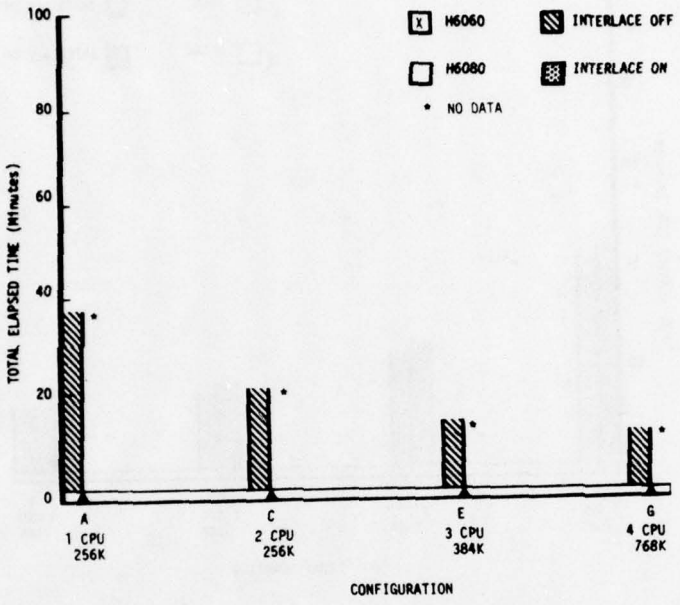


Figure 3-17. Workload 2, CPU Impact

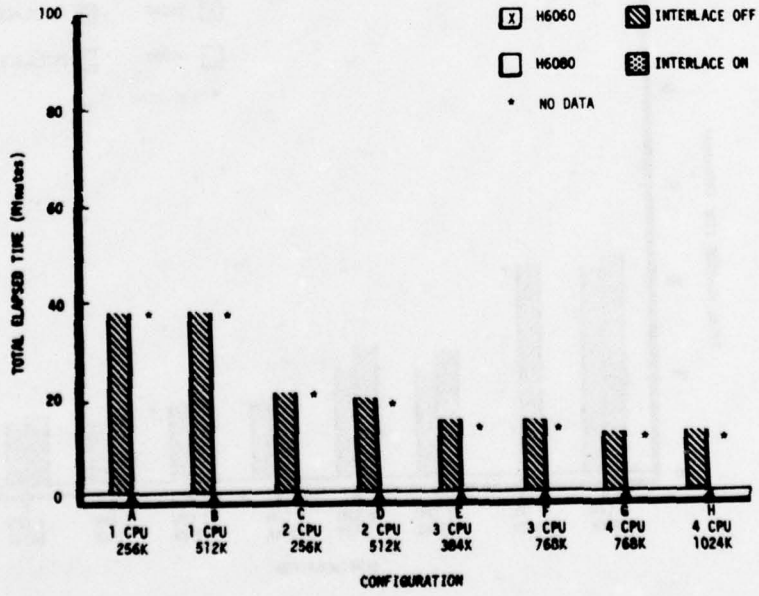


Figure 3-18. Workload 2, Memory Impact

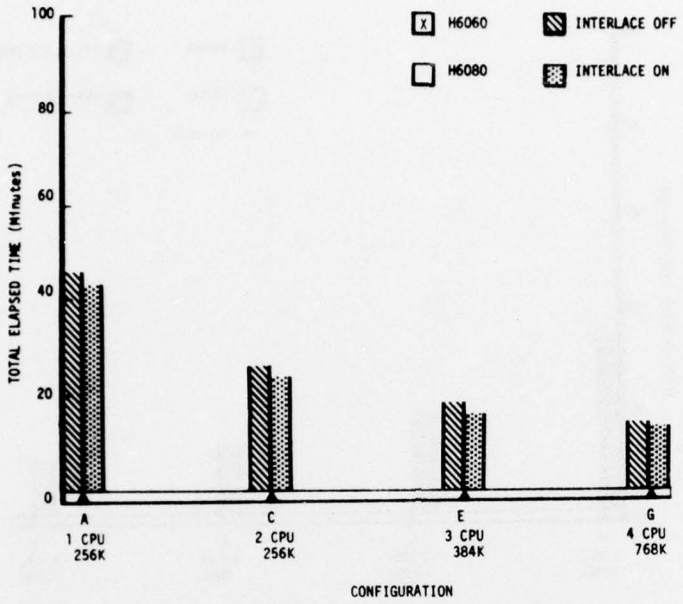


Figure 3-19. Workload C, CPU Impact

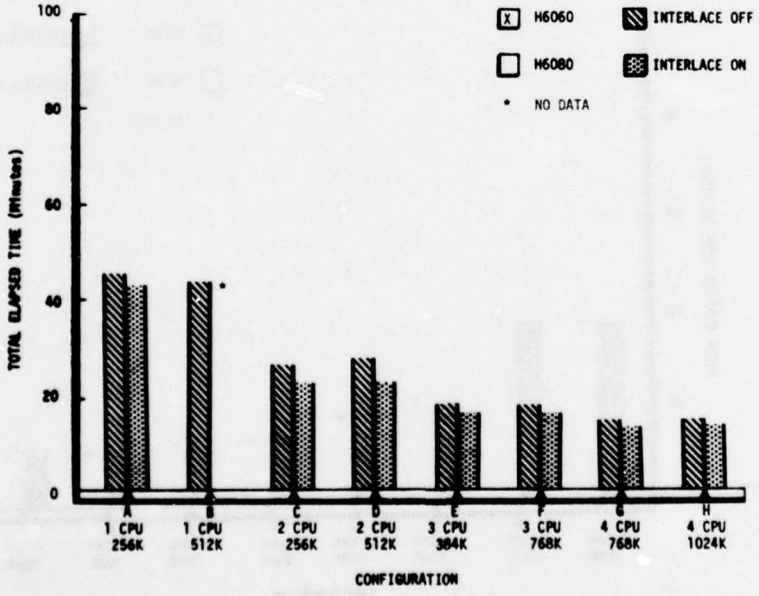


Figure 3-20. Workload C, Memory Impact

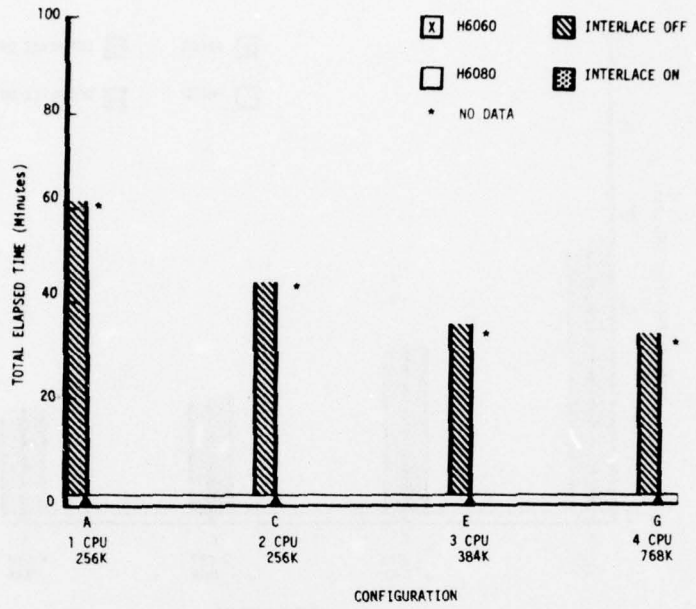


Figure 3-21. Workload 1, CPU Impact

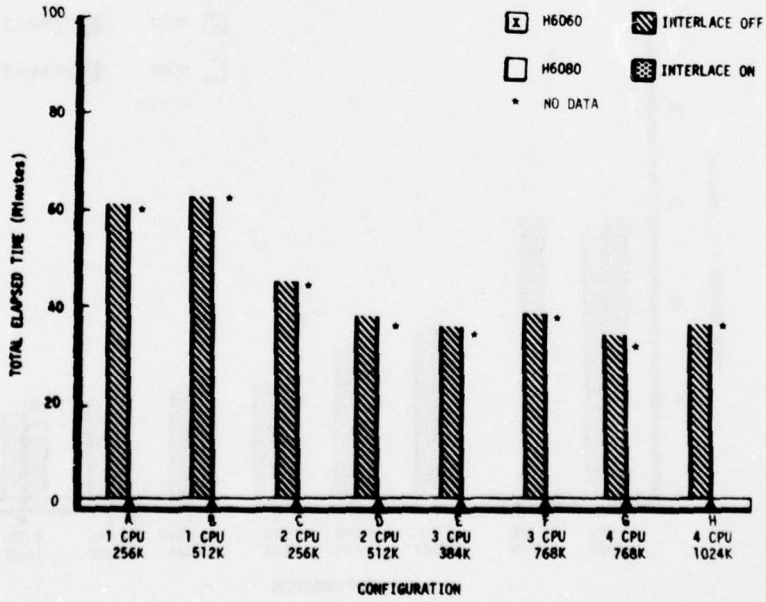


Figure 3-22. Workload 1, Memory Impact

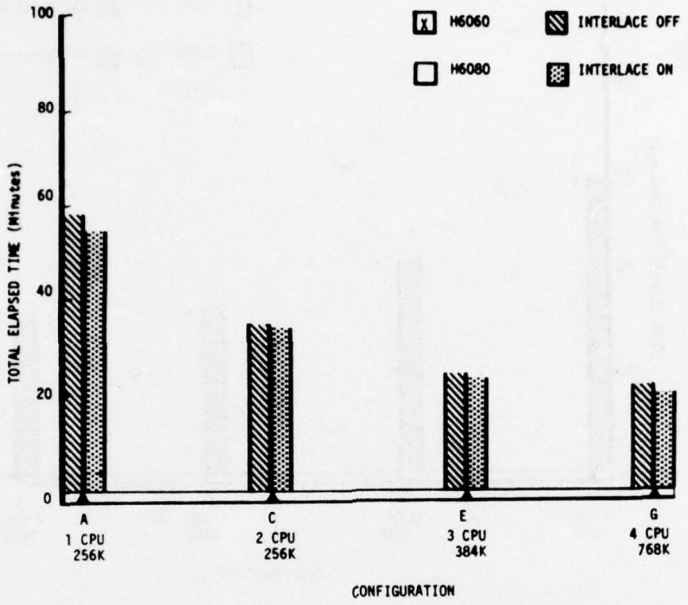


Figure 3-23. Workload D, CPU Impact

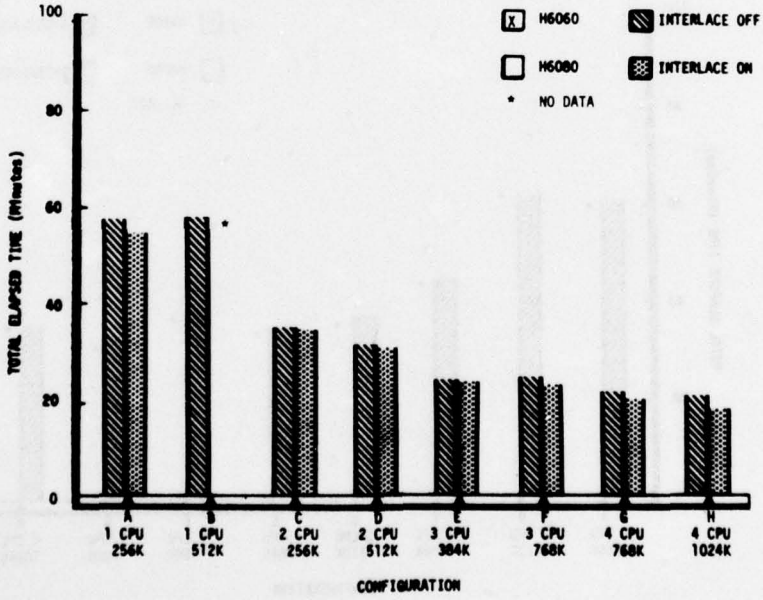


Figure 3-24. Workload D, Memory Impact

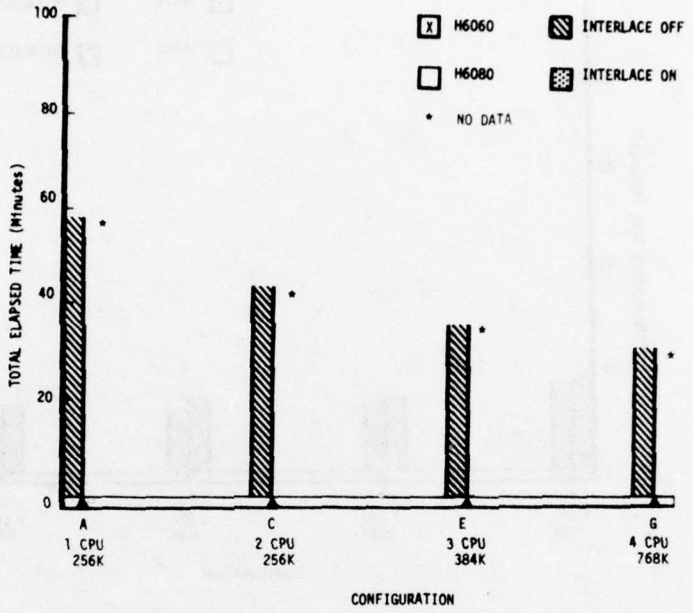


Figure 3-25. Workload 9, CPU Impact

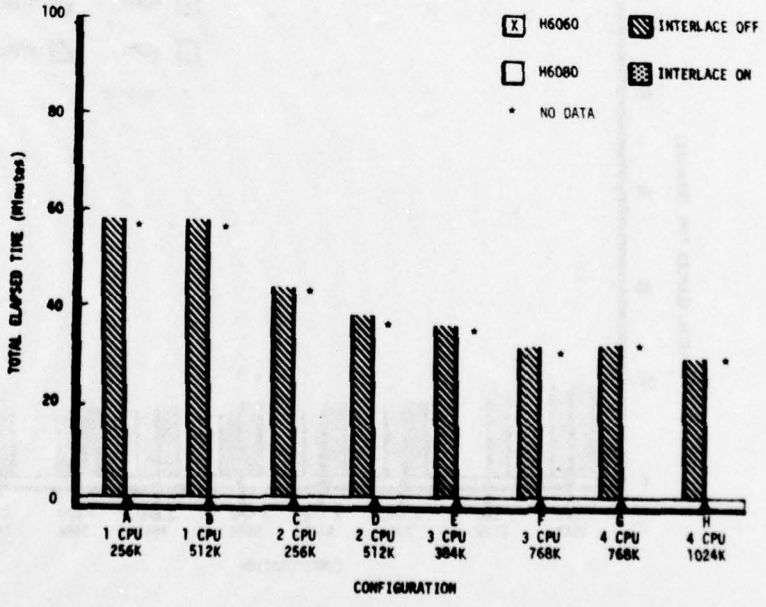


Figure 3-26. Workload 9, Memory Impact

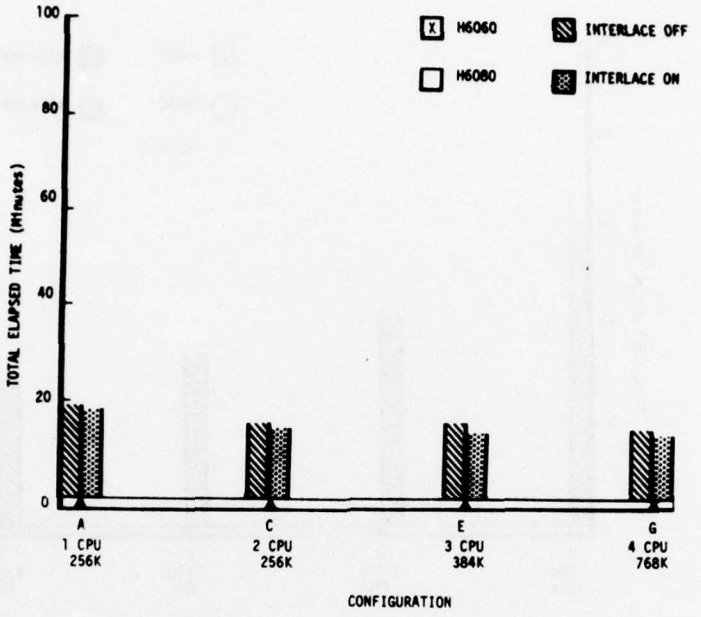


Figure 3-27. Workload F(2), CPU Impact

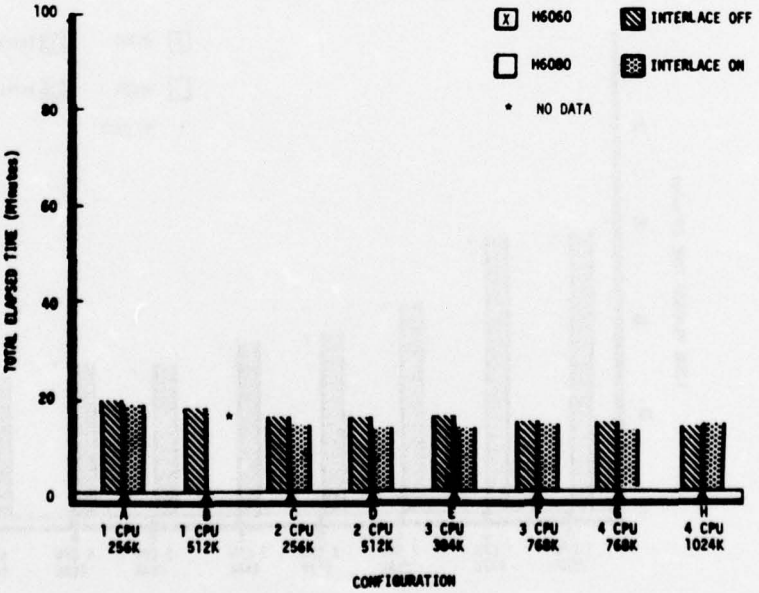


Figure 3-28. Workload F(2), Memory Impact

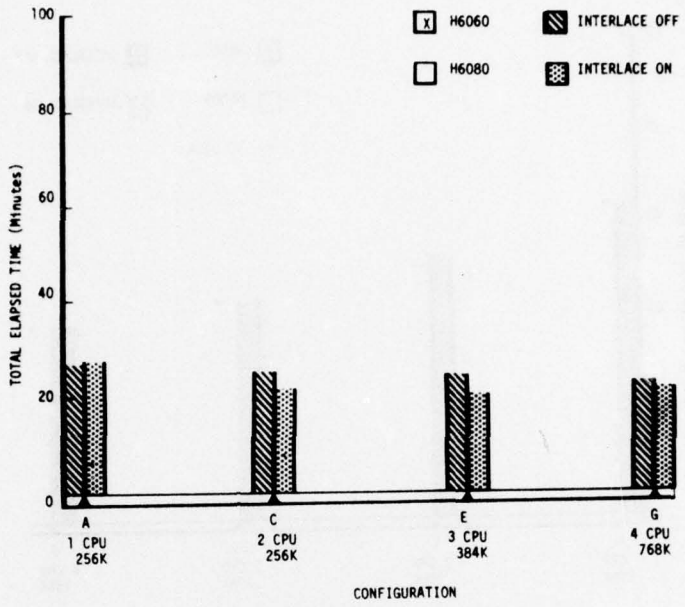


Figure 3-29. Workload F(1), CPU Impact

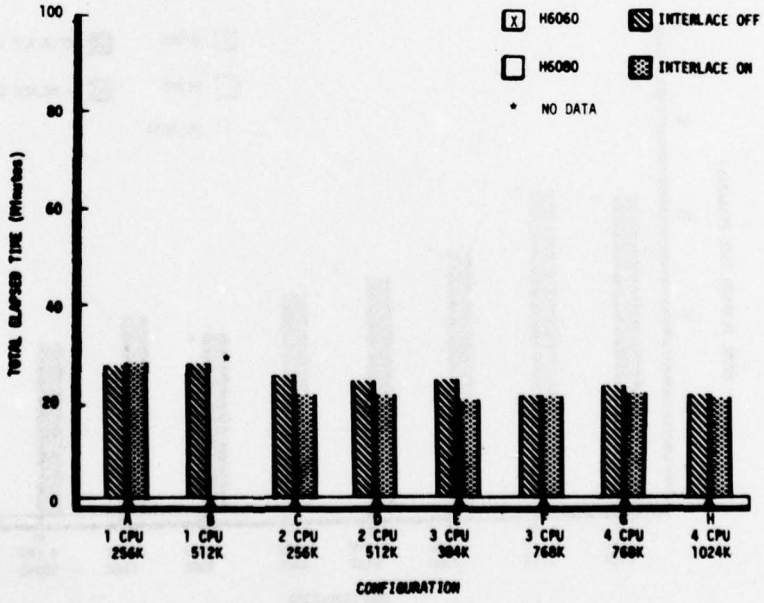


Figure 3-30. Workload F(1), Memory Impact

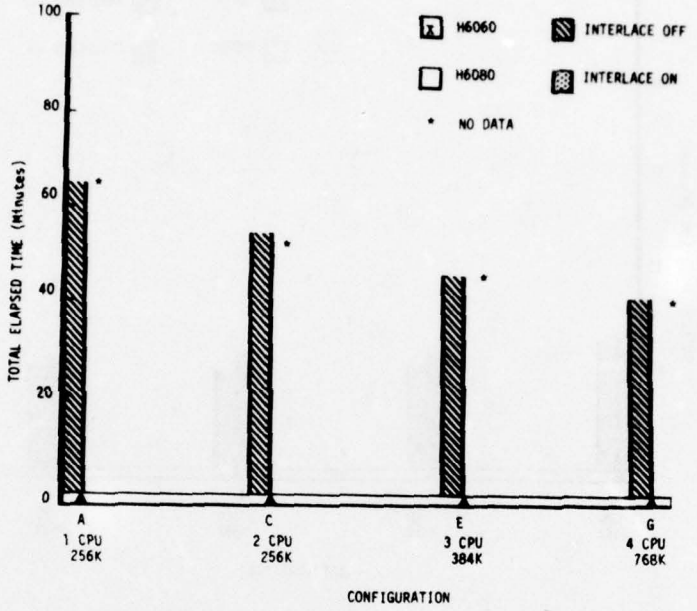


Figure 3-31. Workload 4, CPU Impact

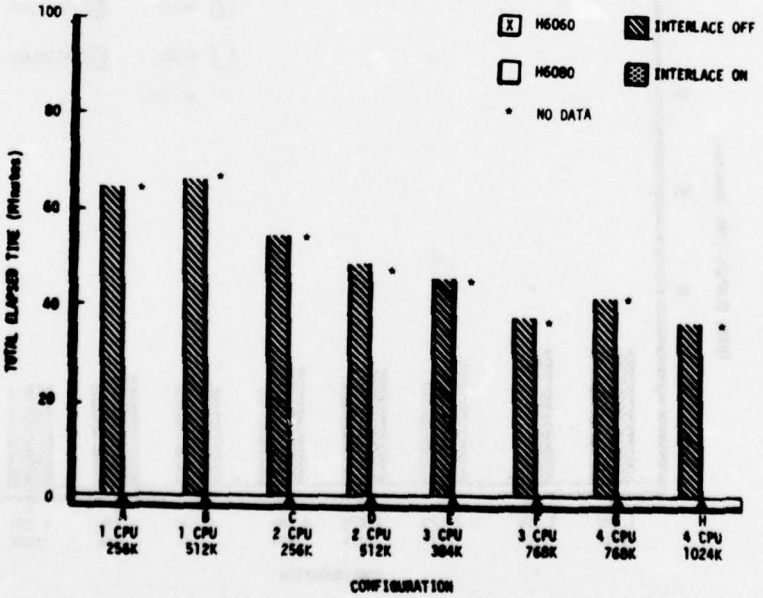


Figure 3-32. Workload 4, Memory Impact

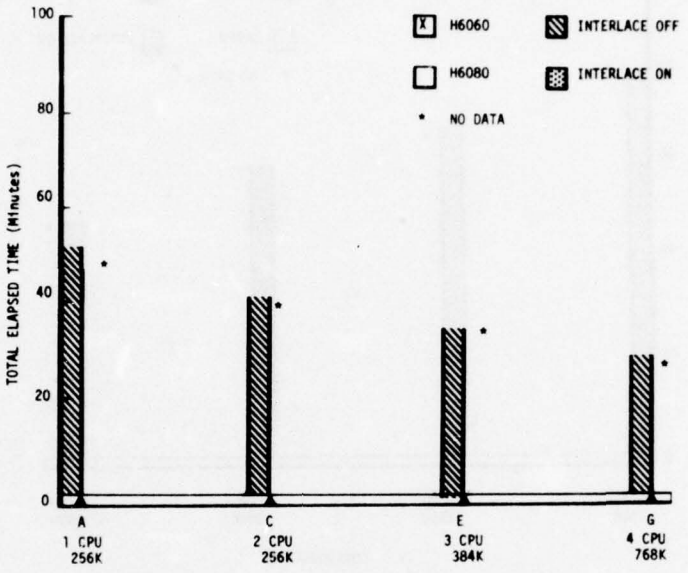


Figure 3-33. Workload 8, CPU Impact

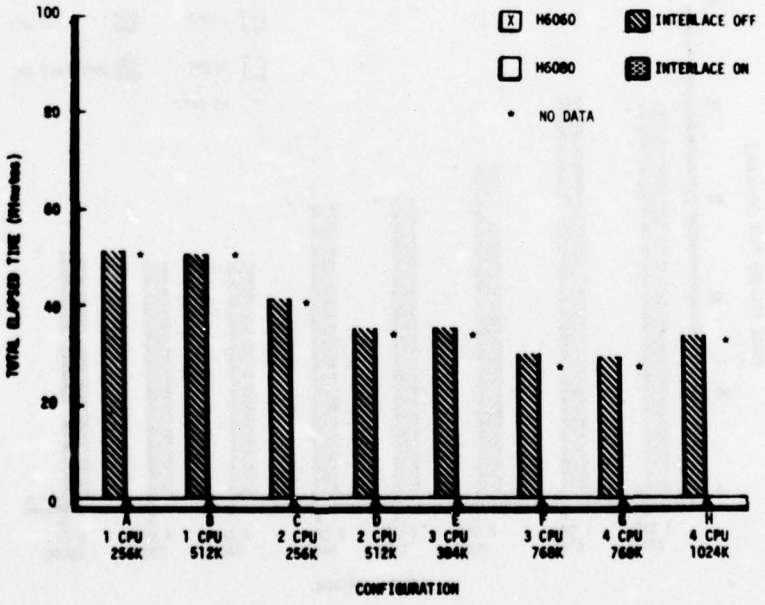


Figure 3-34. Workload 8, Memory Impact

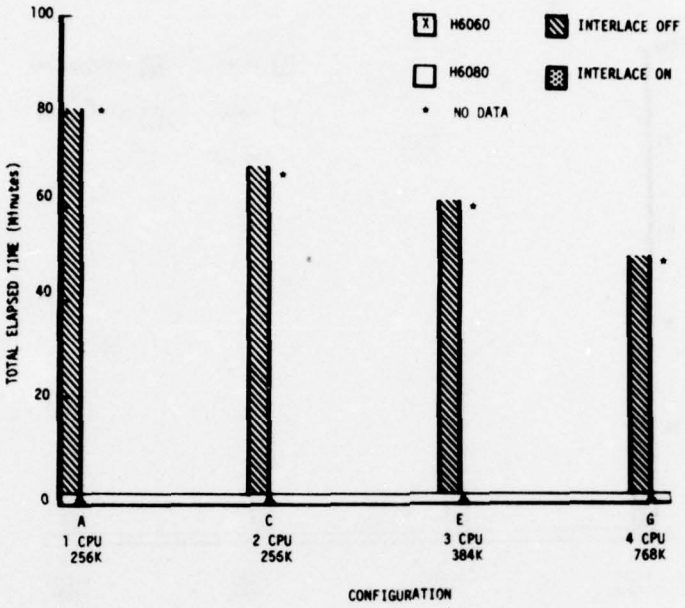


Figure 3-35. Workload 10, CPU Impact

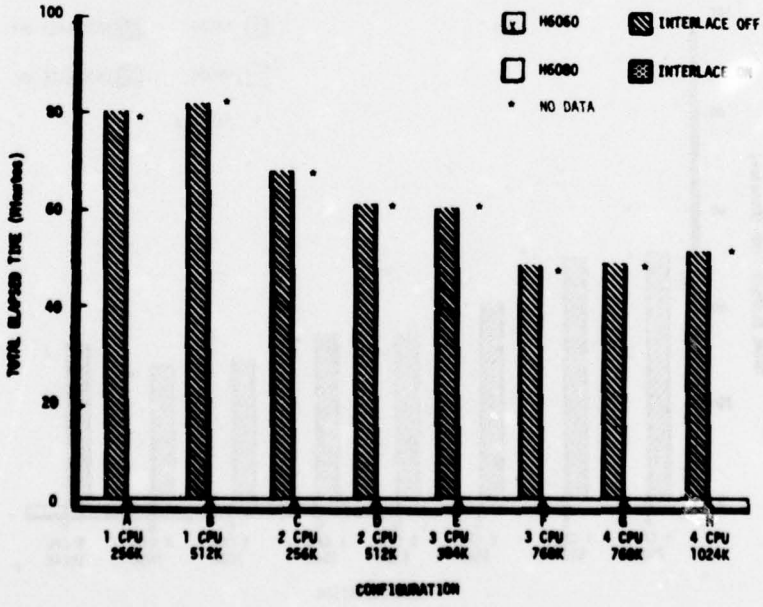


Figure 3-36. Workload 10, Memory Impact

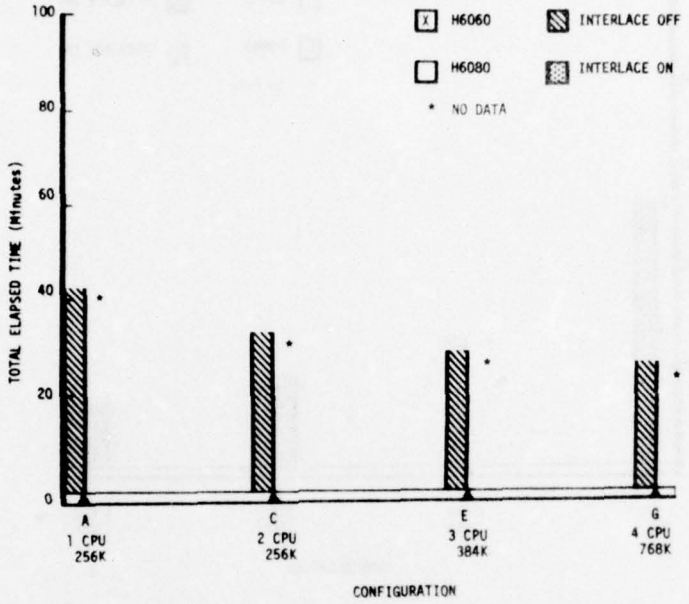


Figure 3-37. Workload 3, CPU Impact

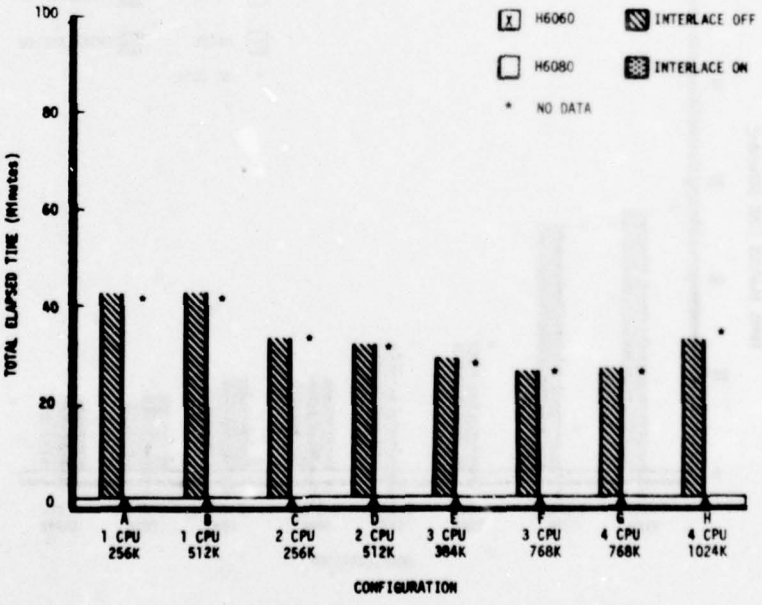


Figure 3-38. Workload 3, Memory Impact

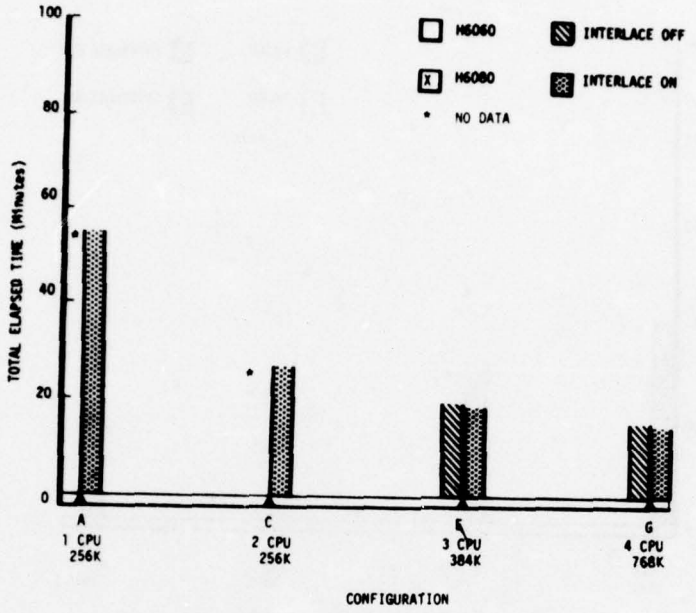


Figure 3-39. Workload E, CPU Impact

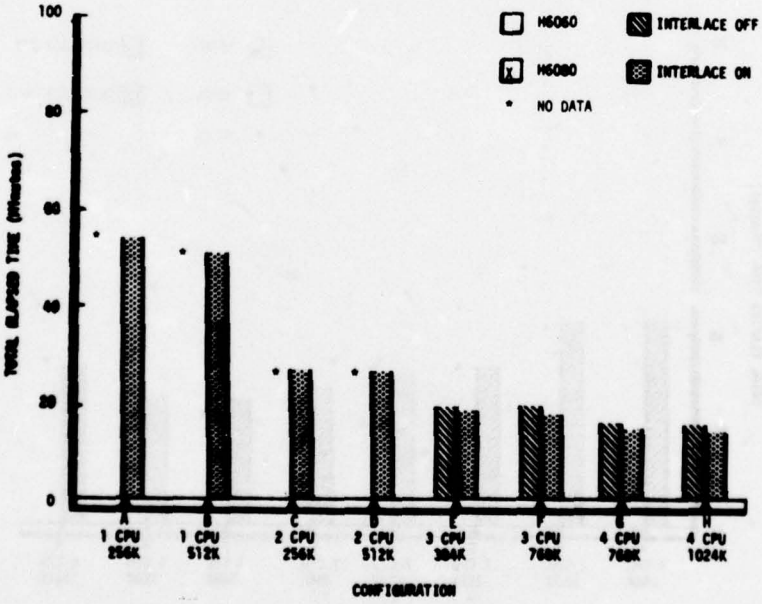


Figure 3-40. Workload E, Memory Impact

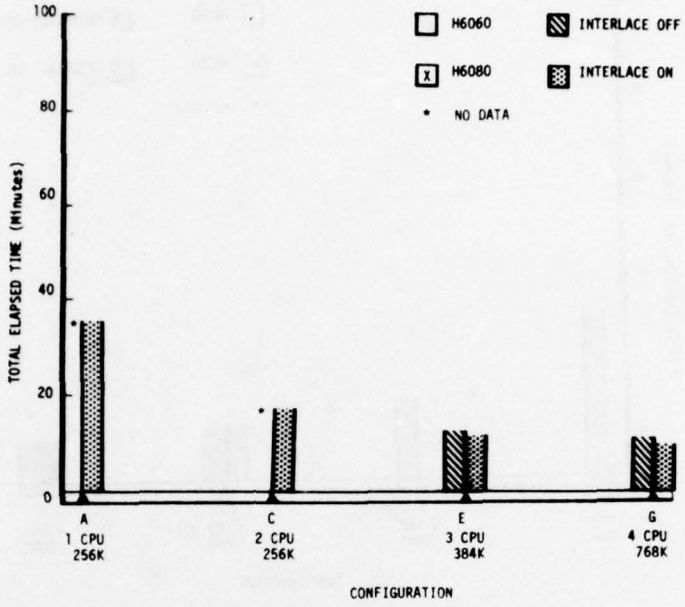


Figure 3-41. Workload A, CPU Impact

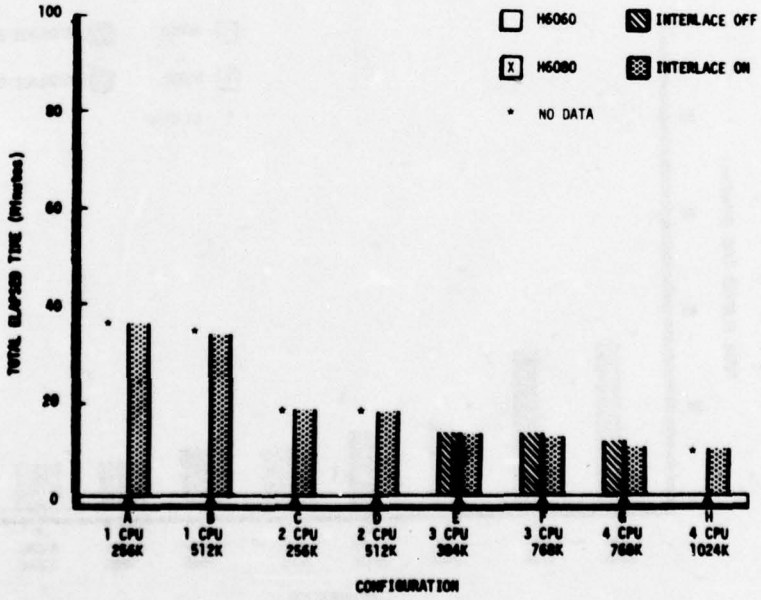


Figure 3-42. Workload A, Memory Impact

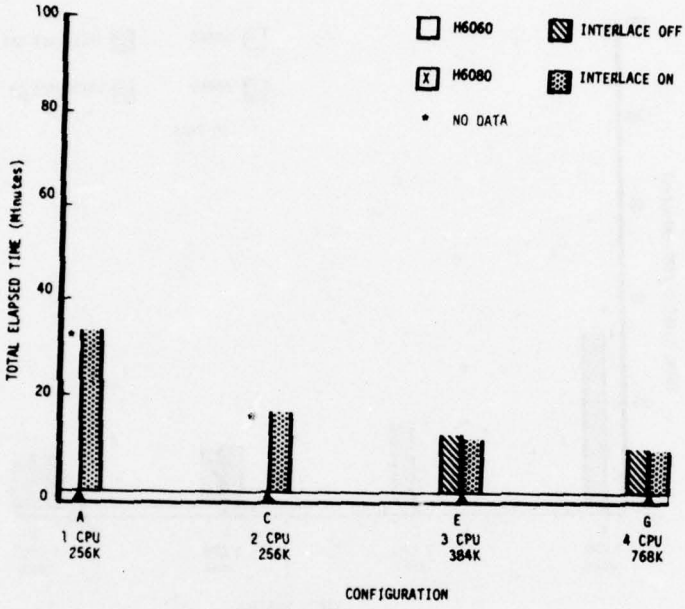


Figure 3-43. Workload B, CPU Impact

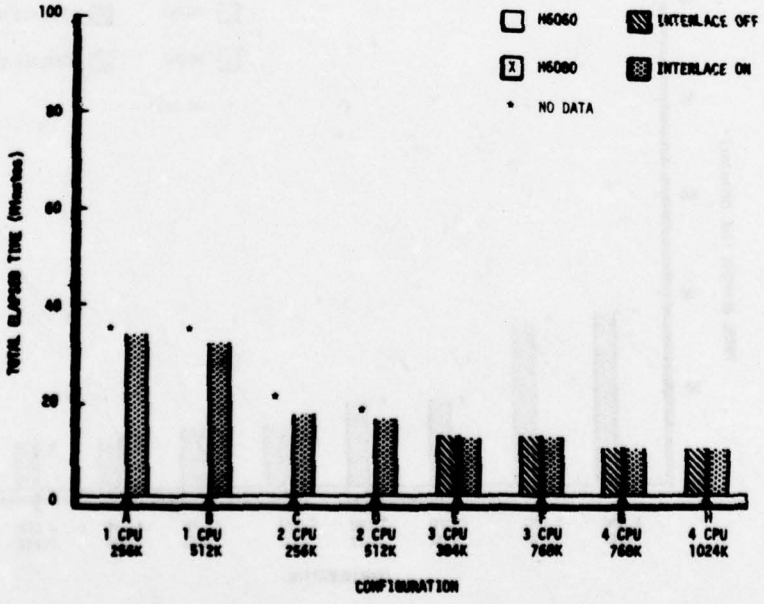


Figure 3-44. Workload B, Memory Impact

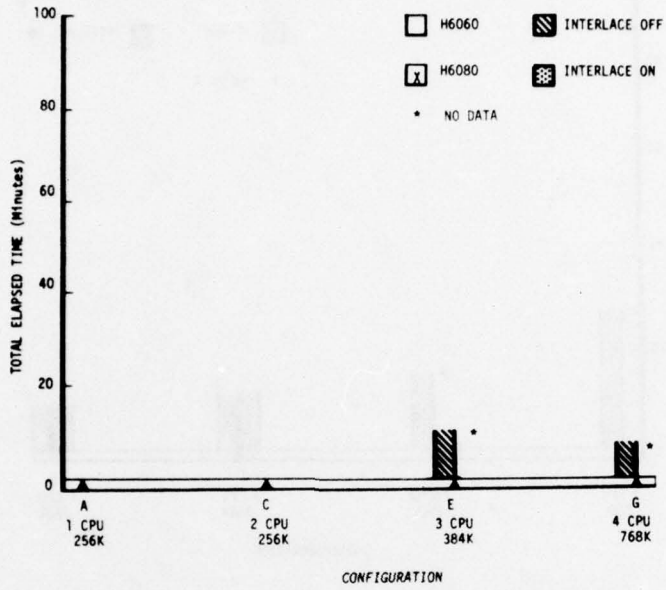


Figure 3-45. Workload 2, CPU Impact

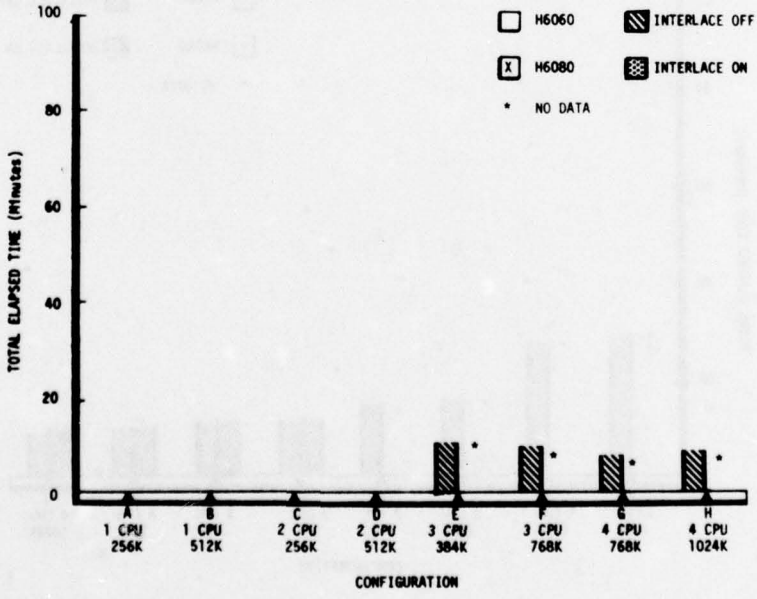


Figure 3-46. Workload 2, Memory Impact

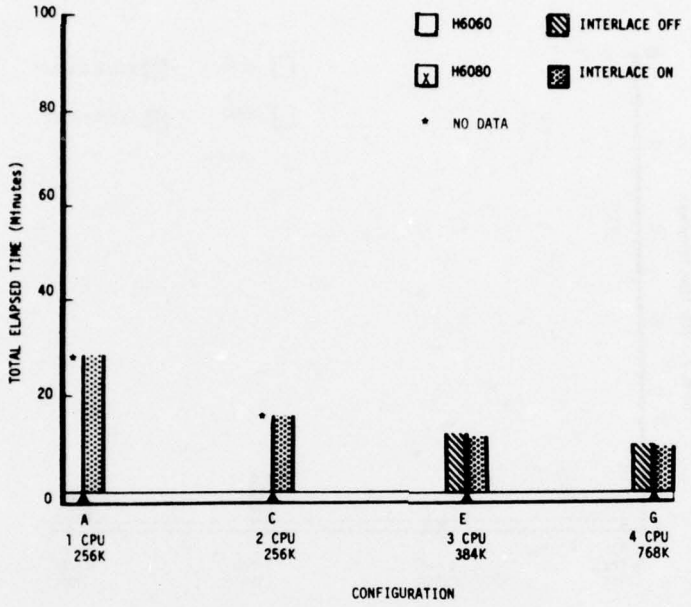


Figure 3-47. Workload C, CPU Impact

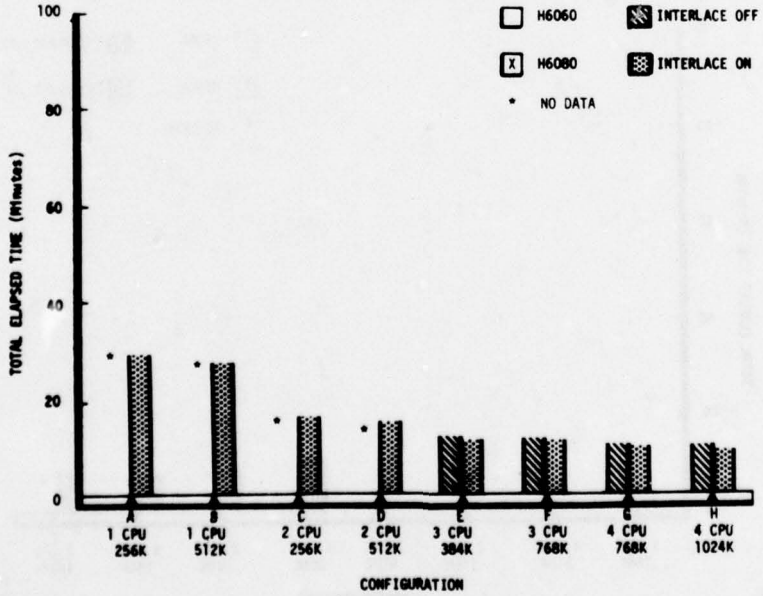


Figure 3-48. Workload C, Memory Impact

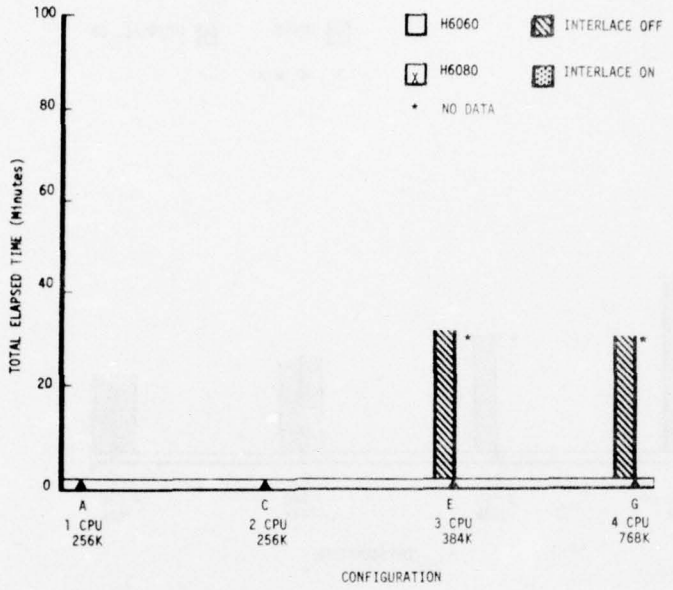


Figure 3-49. Workload 1, CPU Impact

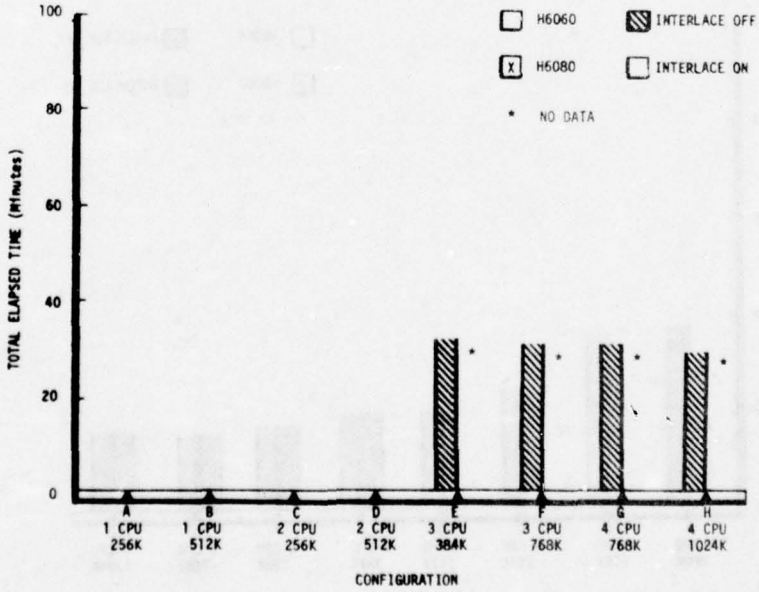


Figure 3-50. Workload 1, Memory Impact

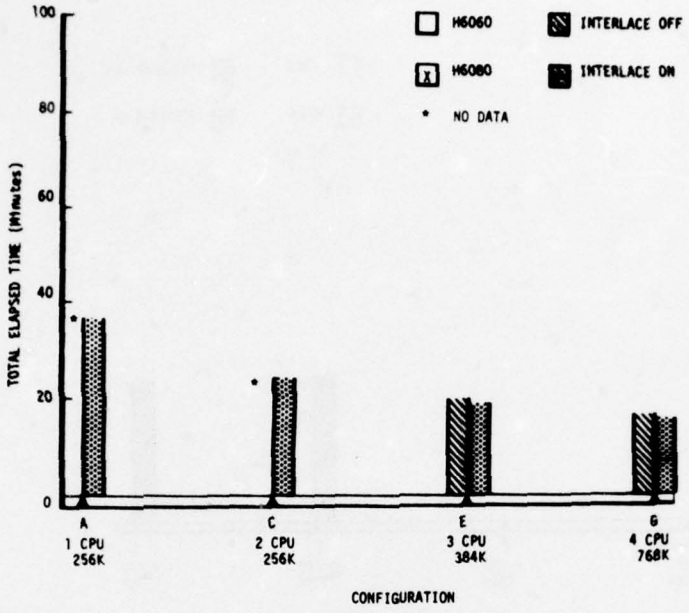


Figure 3-51. Workload D, CPU Impact

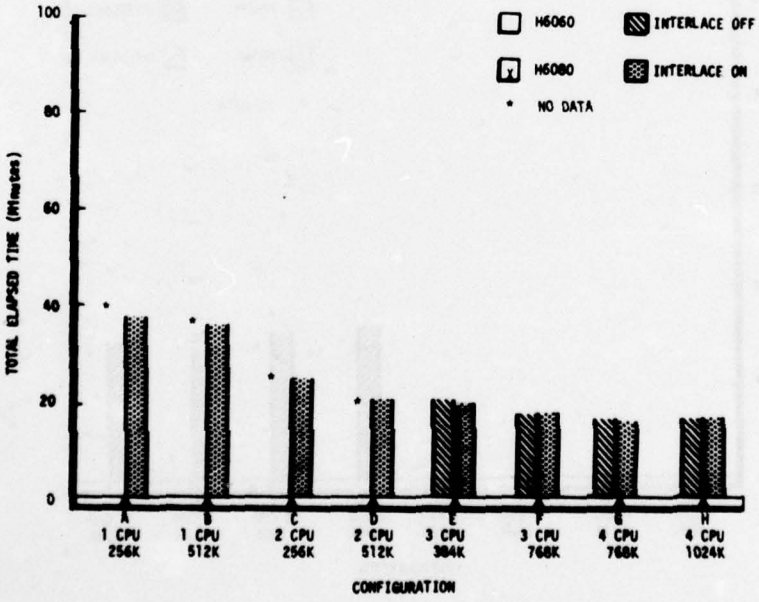


Figure 3-52. Workload D, Memory Impact

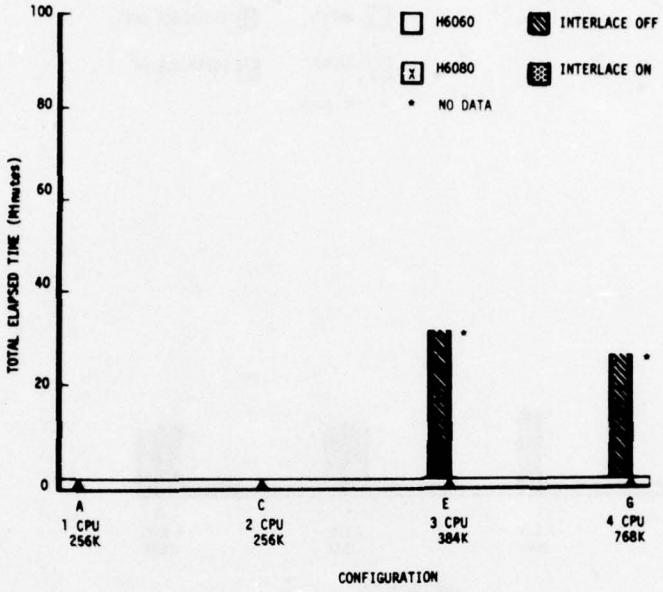


Figure 3-53. Workload 9, CPU Impact

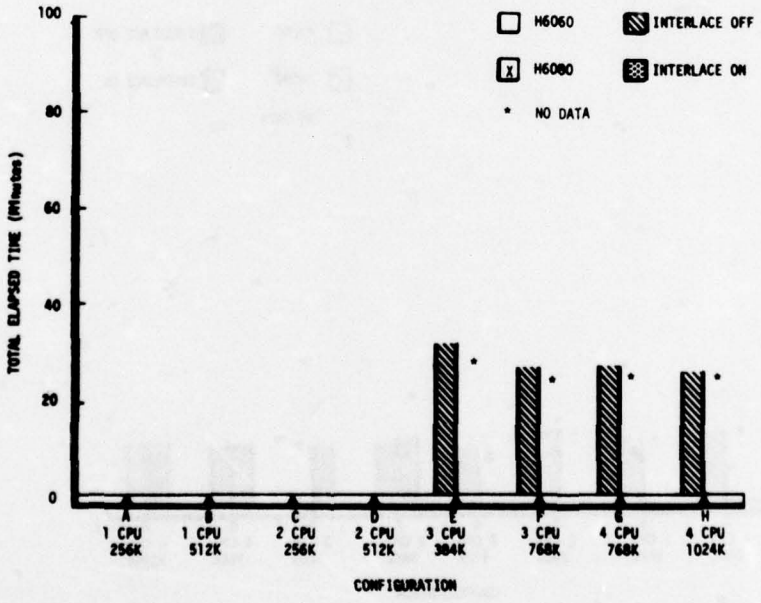


Figure 3-54. Workload 9, Memory Impact

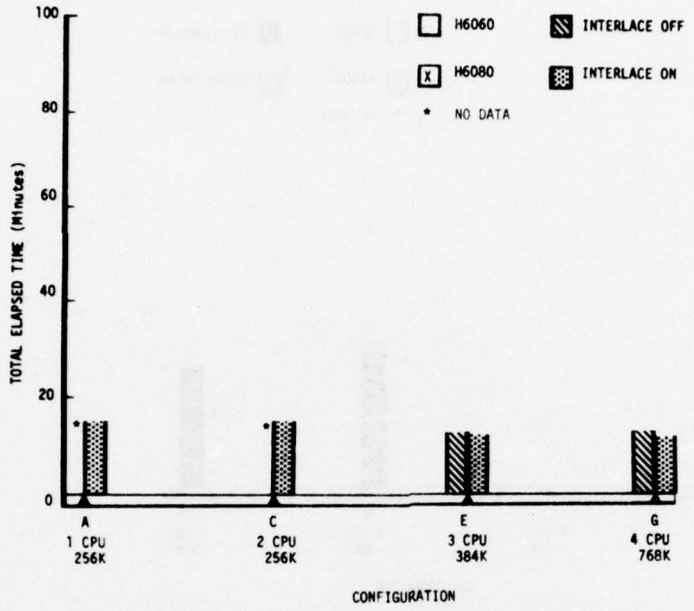


Figure 3-55. Workload F(2), CPU Impact

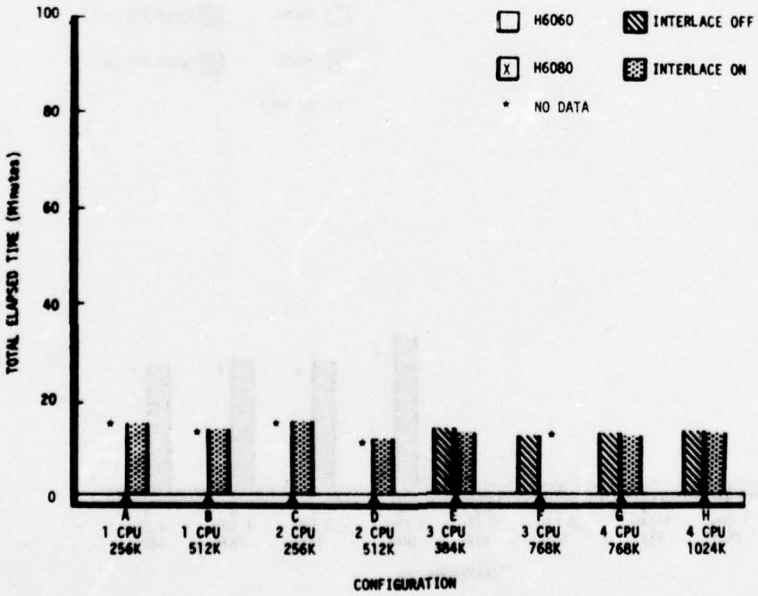


Figure 3-56. Workload F(2), Memory Impact

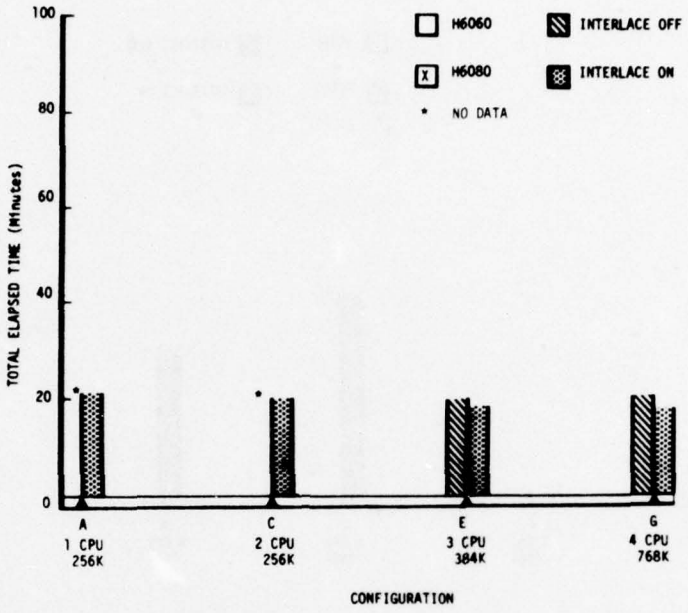


Figure 3-57. Workload F(1), CPU Impact

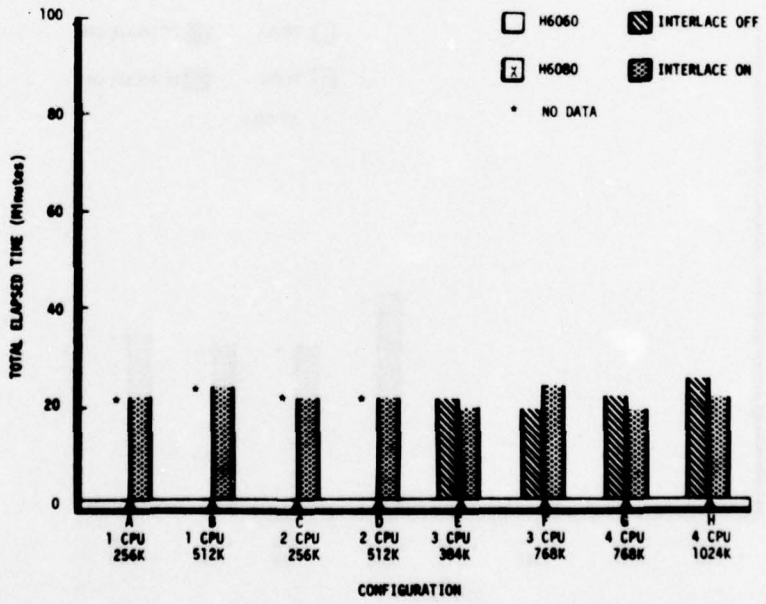


Figure 3-58. Workload F(1), Memory Impact

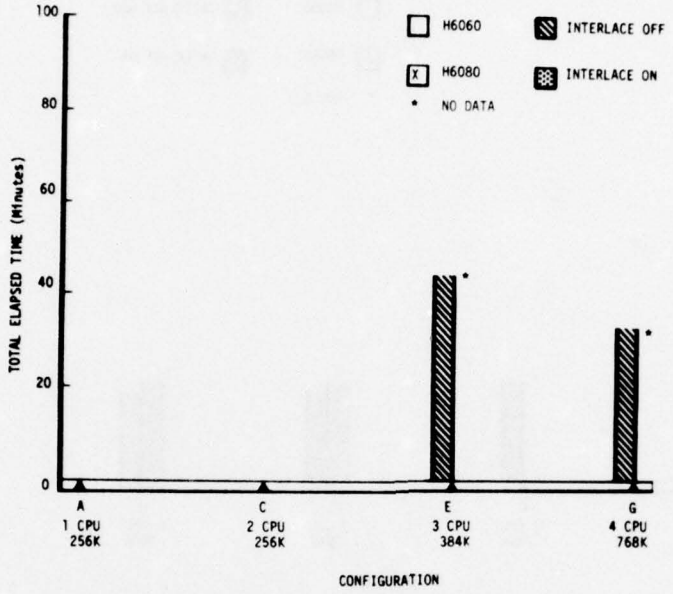


Figure 3-59. Workload 4, CPU Impact

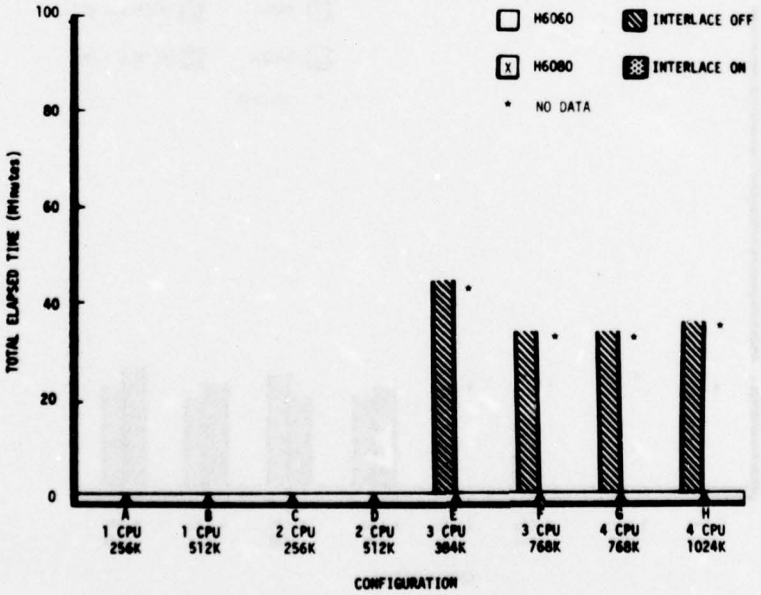


Figure 3-60. Workload 4, Memory Impact

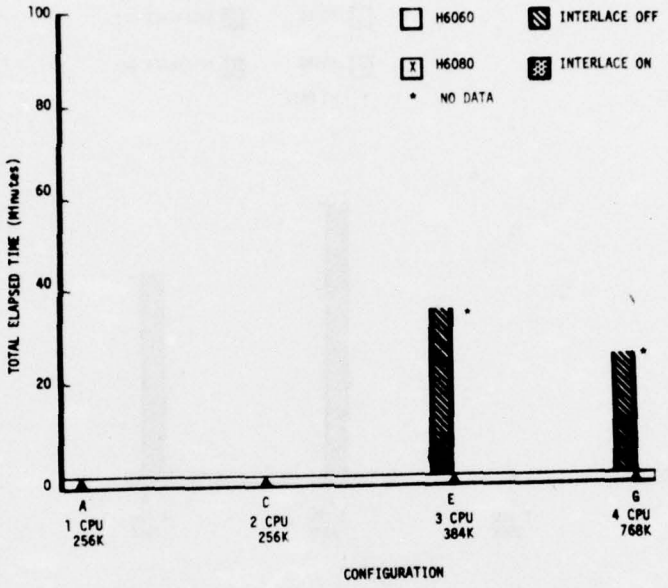


Figure 3-61. Workload 8, CPU Impact

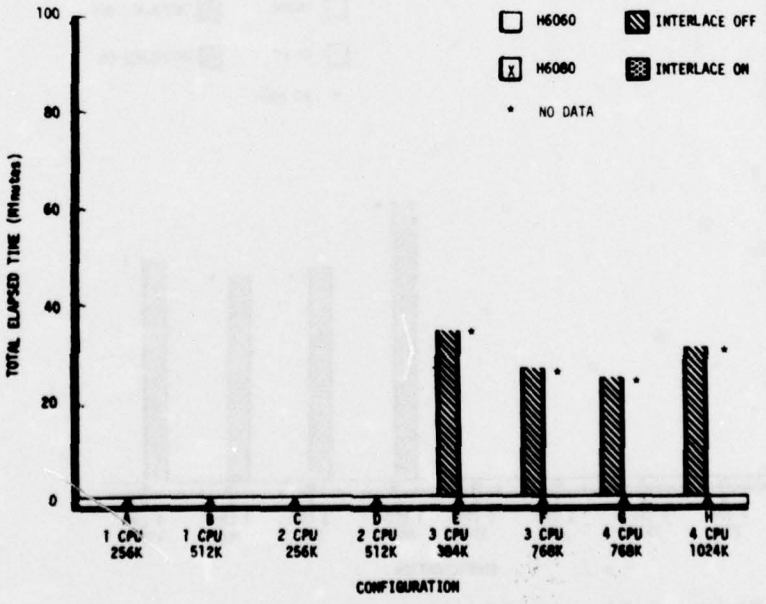


Figure 3-62. Workload 8, Memory Impact

FIGURE 62. WORKLOAD 10, CPU IMPACT

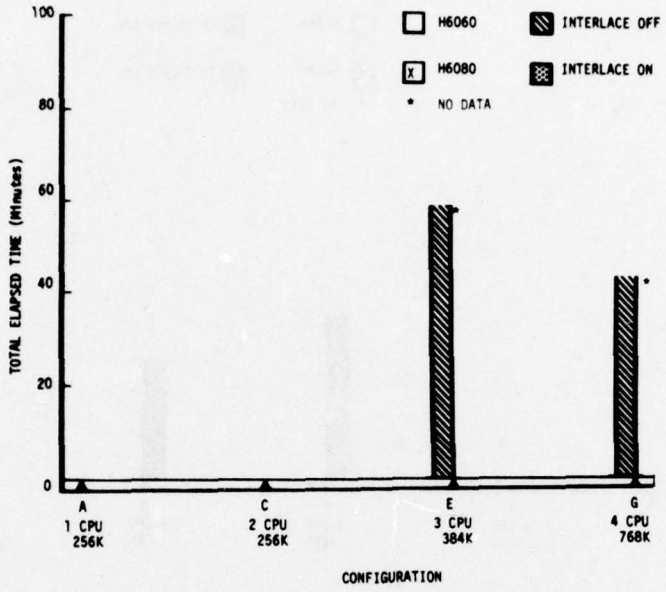


Figure 3-63. Workload 10, CPU Impact

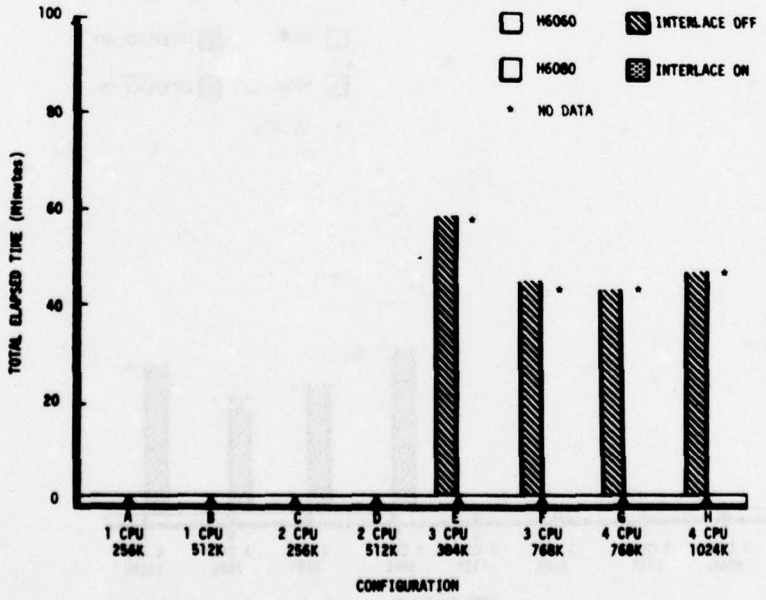


Figure 3-64. Workload 10, Memory Impact

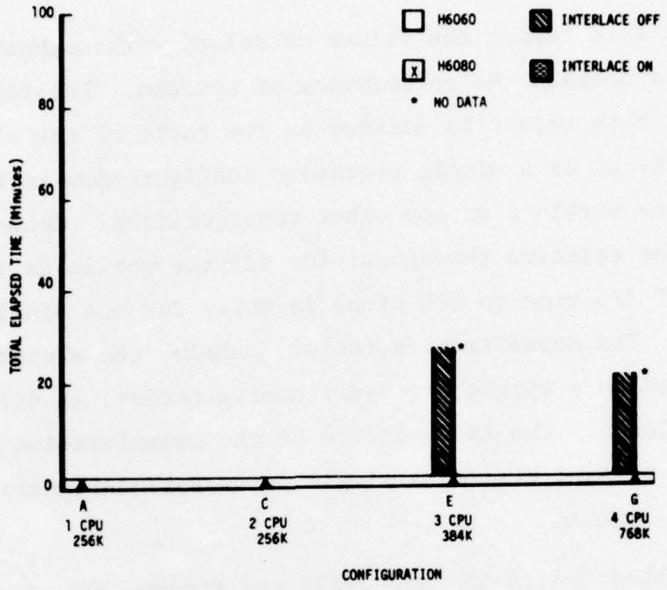


Figure 3-65. Workload 3, CPU Impact

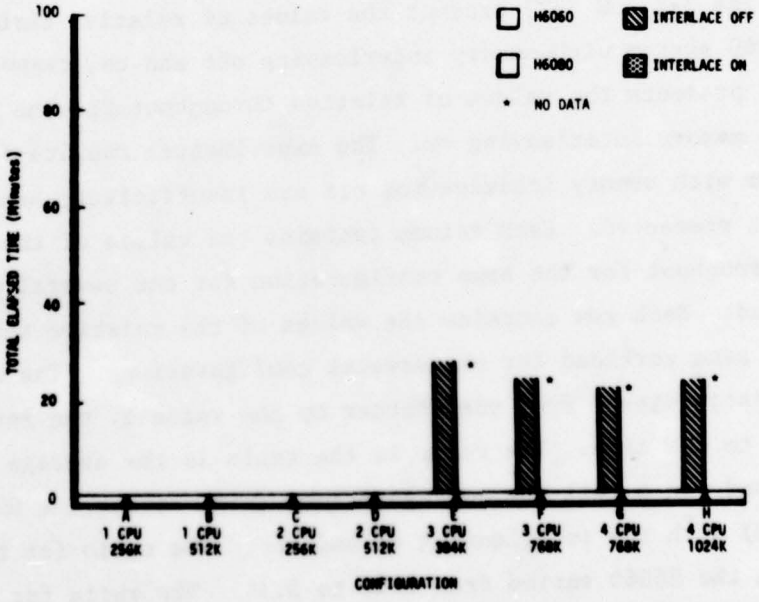


Figure 3-66. Workload 3, Memory Impact

3.4.2 Relative Throughput

In this report the values of relative throughput are used to describe and compare the performance of systems. Relative throughput as used in this report is defined as the ratio of the elapsed time of a given workload on a single processor configuration to the elapsed time of the same workload on any other configuration. Consequently, the value of the relative throughput for all the workloads (regardless of the ratio of I/O time to CPU time) is unity for the single processor configuration. The normalization factor, namely, the elapsed time of a given workload on a single processor configuration, is different for different workloads. The above choice of the normalization factor was arbitrary; the elapsed time on a single processor configuration was chosen for convenience.

Tables 3-X, 3-XI, and 3-XII and Figures 4-1, 4-2, and 4-3 show the variation of relative throughput with increases in the number of processors for the several workloads used in the experiments. Tables 3-X and 3-XI present the values of relative throughput for the H6060 system with memory interleaving off and on, respectively. Table 3-XII presents the values of relative throughput for the H6080 system with memory interleaving on. The experimental results for the H6080 system with memory interleaving off are insufficient; hence they are not presented. Each column contains the values of the relative throughput for the same configuration for the several workloads studied. Each row contains the values of the relative throughput for the same workload for the several configurations. The workloads are distinguished from one another by the value of the ratio of I/O time to CPU time. The ratio in the table is the average value for a workload run on all the configurations (with either the H6060 or the H6080) with the interleaving on and off. The ratio for the workloads on the H6060 varied from 0.01 to 3.34. The ratio for these workloads on the H6080 varied from 0.01 to 5.07. The ratio for the

TABLE 3-X. H6060 RELATIVE THROUGHPUT

CONFIGURATION	Code	6A	6B	6C	6D	6E	6F	6G	6H
	CPUs	1	1	2	2	3	3	4	4
	SCUs	1	2	2	3	3	4	3	4
	IOMs	2	2	2	2	2	2	2	2
	Core	256K	512K	256K	512K	384K	768K	768K	1024K
WORKLOADS - IO/CPU RATIO	E/0.01	1.000	1.010	1.889	1.908	2.756	2.801	3.490	3.607
	A/0.28	1.000	1.019	1.870	1.941	2.800	2.812	3.317	3.594
	B/0.40	1.000	1.023	1.843	1.926	2.772	2.759	3.439	3.541
	2/0.48	1.000	1.014	1.792	*	2.689	2.602	3.139	3.079
	C/0.62	1.000	1.036	1.752	1.677	2.545	2.621	3.271	3.332
	1/1.12	1.000	0.983	1.379	*	1.715	1.600	1.813	1.730
	D/1.23	1.000	1.001	1.631	1.829	2.371	2.323	2.602	2.795
	9/2.26	1.000	0.994	1.327	1.508	1.627	1.809	1.859	1.986
	F(2)/2.78	1.000	1.042	1.252	1.267	1.312	1.354	1.402	1.502
	F(1)/3.07	1.000	0.958	1.055	1.104	1.109	1.268	1.169	1.242
	4/3.06	1.000	0.973	1.179	*	1.419	1.706	1.560	1.773
	8/3.13	1.000	1.012	1.218	1.454	1.456	1.715	1.751	1.538
	10/3.25	1.000	0.977	1.168	1.325	1.323	1.649	1.630	1.566
	3/3.34	1.000	0.994	1.274	*	1.473	1.635	1.616	1.312

DATA SOURCE = GESEP HOST MACHINE = H6060 MEMORY INTERLACE = OFF

TABLE 3-XI. H6060 RELATIVE THROUGHPUT

CONFIGURATION	Code	6A	6B	6C	6D	6E	6F	6G	6H
	CPUs	1	1	2	2	3	3	4	4
	SCUs	1	2	2	3	3	4	3	4
	IOUs	2	2	2	2	2	2	2	2
	Core	256K	512K	256K	512K	384K	768K	768K	1024K
WORKLOADS - IO/CPU RATIO	E/0.01	1.000	-	2.054	2.031	2.827	2.926	3.507	3.734
	A/0.28	1.000	-	2.053	1.996	2.816	2.921	3.535	3.740
	B/0.40	1.000	-	2.056	2.005	2.735	2.845	3.427	3.668
	2/0.48	-	-	-	-	-	-	-	-
	C/0.62	1.000	-	1.960	1.933	2.697	2.758	3.318	3.329
	1/1.12	-	-	-	-	-	-	-	-
	D/1.23	1.000	-	1.578	1.832	2.320	2.378	2.735	3.041
	9/2.26	-	-	-	-	-	-	-	-
	F(2)/2.78	1.000	-	1.343	1.415	1.460	1.388	1.470	1.387
	F(1)/3.07	1.000	-	1.296	1.316	1.393	1.353	1.310	1.326
	4/3.06	-	-	-	-	-	-	-	-
	8/3.13	-	-	-	-	-	-	-	-
	10/3.25	-	-	-	-	-	-	-	-
	3/3.34	-	-	-	-	-	-	-	-

HOST MACHINE - H6060 MEMORY INTERLACE - ON

TABLE 3-XII. H6080 RELATIVE THROUGHPUT

CONFIGURATION	Code	8A	8B	8C	8D	8E	8F	8G	8H
	CPUs	1	1	2	2	3	3	4	4
	SCUs	1	2	2	3	3	4	3	4
	IOMs	2	2	2	2	2	2	2	2
	Core	256K	512K	256K	512K	384K	768K	768K	1024K
WORKLOADS - IO/CPU RATIO	E/0.01	1.000	1.066	2.021	2.084	2.974	3.097	3.817	4.017
	A/0.46	1.000	1.064	1.990	2.066	2.931	2.979	3.647	3.691
	B/0.64	1.000	1.060	1.956	2.075	2.821	2.836	3.557	3.608
	2	-	-	-	-	-	-	-	-
	C/0.98	1.000	1.059	1.859	2.021	2.592	2.687	3.141	3.332
	1	-	-	-	-	-	-	-	-
	D/1.97	1.000	1.045	1.529	1.815	1.898	2.208	2.449	2.301
	9	-	-	-	-	-	-	-	-
	F(2)/4.62	1.000	1.093	0.980	1.261	1.147	-	1.207	1.174
	F(1)/5.07	1.000	0.912	1.024	1.015	1.136	0.902	1.157	1.003
	4	-	-	-	-	-	-	-	-
	8	-	-	-	-	-	-	-	-
	10	-	-	-	-	-	-	-	-
	3	-	-	-	-	-	-	-	-

HOST MACHINE - H6080 MEMORY INTERLACE - ON

same workload is higher on the H6060 than on the H6060 because of the increased processor speed (instructions/sec) reduces the processor time while the I/O time remains roughly constant.

As seen from Tables 3-X and 3-XI, for ratios in the range of 0.01 to 0.4, the increase in relative throughput is almost linear with increases in the number of processors (for H6060). For example, for the workload E (ratio of I/O time to CPU time of 0.01) the relative throughput increases to 1.908 when the number of processors is increased to 2, increases to 2.801 when the number of processors is 3, and increases to 3.607 when the number of processors is 4. For I/O bound loads, the values of relative throughput exhibit an asymptotic behavior. For example, for the workload F(2) (ratio of I/O time to CPU time of 3.07), the relative throughput increases to 1.252 when the number of processors is increased to 2, increases to 1.312 when the number of processors is 3, and increases to 1.402 when the number of processors is 4. For this workload there was a 40% increase in relative throughput when the number of processors was increased from one to four. For the CPU-bound workload E, there was a 260% increase in the relative throughput when the number of processors was increased from one to four.

As discussed above, the values of relative throughput are useful in comparing the performance of a single workload on several configurations. However, they should not be used in comparing the performance of several workloads on the same configuration. For example, from Table 3-X the relative throughput for the configuration 6C is 1.87 for workload A and 1.055 for workload F(1). But no valid conclusions can be drawn by comparing the two values of relative throughput.

SECTION IV

SUMMARY

In order to provide quantitative information regarding multiprocessor throughput, the Deputy for Air Force WWMCCS, ESD, in response to Hq USAF tasking and with support from AFDSDC, undertook a task to measure throughput for various H6060 and H6080 configurations. Performance data was obtained by running synthetic workloads on these equipment configurations. This section summarizes both the performance information which has been extracted from the data and the test conditions under which the data were obtained. Section 4.1 reviews the workloads, hardware and software configurations, and operational conditions under which the tests were conducted. Section 4.2 contains observations about the results of the tests and Section 4.3 explains the application of the test results to operational WWMCCS sites.

4.1 Test Conditions

The workloads, hardware and software configurations, and operational conditions under which the tests were run are summarized in this section.

4.1.1 Workload Limitations

Since a test workload is only a model of a real workload, test results obtained using any test workload must be applied with caution in the analysis and prediction of the performance of systems processing real workloads. The test workloads used in these tests consisted of 14 synthetic workloads, 7 of which were based on SCF data from SAC and PACOM, and 7 of which were arbitrarily defined to extend the range of IO/CP ratios in the test workload set. The characteristics of the synthetic workloads used in the tests must be considered when applying the test results to actual workloads.

The remainder of this section identifies workload related factors that effect the representativeness of the test workloads derived from SCF data and also identifies other synthetic workload characteristics that may be different from those of actual WWMCCS workloads. These differences should be noted, but must not be misconstrued as invalidating results of this report. The validation study performed by AFDSDC (Appendix I) clearly indicates the utility of the data and results presented herein.

a. The synthetic workloads used in the tests consisted of batch jobs. As a result, the system overhead resource utilization and the resource demand patterns of on-line activities which are a portion of the workload at some WWMCCS sites are not reflected in the synthetic workloads.

b. Workloads differ both in the amount of resources used and in the ways in which those resources are demanded. The design of the synthetic workload generator determines the way in which CPU, I/O, and memory utilization data obtained from the SCF tape are transformed into resource utilization specifications for the synthetic workload. The following properties of the synthetic workload generator which affect that transformation must be considered:

(1) In order to characterize a real workload, SCF job data is sorted into a three-dimensional array having 100 equal intervals for CPU time, 60 equal intervals for I/O time, and 10 equal intervals for memory size. The size of the CPU interval is determined by dividing the difference between the maximum and minimum CPU utilization values found in the workload by 100. The sizes for the I/O and memory intervals are determined in a similar way. The synthetic workload generator treats the jobs placed into each cell of the array as having CPU, I/O, and memory utilization values equal to the midpoint values for each cell.

(2) The job having the largest value for each dimension of the array (i.e., the CPU dimension, I/O dimension, and memory dimension) was discarded for some test workloads.

(3) The subset of jobs selected to represent the total set of jobs (so that tests can be run in a reasonable amount of time) may have resource utilizations which are not exactly proportional to those of the total set of the real workload.

c. The synthetic workload generator design also determines both the way and how well the resource utilization requirement specifications, derived from the SCF data, are met by the synthetic jobs constituting the synthetic test workload. The following characteristics of the synthetic programs are noted:

(1) All of the synthetic programs have the same structure and therefore have identical resource utilization patterns. That is to say, all of the I/O time in all of the programs is generated by iterations of the same I/O loop and all of the CPU time is generated either by the I/O or by iterations of the same CPU loop.

(2) There is high locality* of code in the synthetic programs.

(3) All of the I/O operations are to disk files.

(4) I/O activity consisting of sequential file processing only, results in a very small amount of I/O time utilized per I/O transaction. The smaller the amount of I/O time utilized per transaction the greater the number of transactions required to accumulate a given amount of I/O time (i.e., the amount of I/O time specified). Therefore, the synthetic program I/O activity may generate more I/O dependent CPU time and system overhead time than would normally be expected for the same amount of I/O time in real workloads.

---

* All of the instructions executed by the synthetic program have core memory addresses that are restricted to a very small portion of the total memory space.

4.1.1.1 Validation of Results

The ratio of I/O time to CPU time has been identified during this study as one of the parameters that influence relative throughput. Experiments with the synthetic test workloads during the WWMCCS Multiprocessor Performance Evaluation task have shown that this ratio determines the amount of increase in the relative throughput for a given workload.

A validation of the results contained in this report conducted by AFDSDC is contained in Appendix I. AFDSDC used a linear regression model of the form $Y = 1.847 - 0.201\ X$ for the primary test data, where Y = IO/CP ratio and X = relative throughput for configuration C. When the validation test workloads were executed and the results compared with the earlier results, the validation data fell within the 90% confidence interval for the regression.

It is recommended that several test workloads be constructed from application programs currently operational at WWMCCS sites. The ratio of I/O time to CPU time should be a criterion in the construction of these test workloads and the ratio should be within the range of values studied during the tests reported in this document. The workloads should be used on controlled tests with several HIS 6000 multiprocessor configurations to determine relative throughput. These experimental values may be compared with those predicted by using the WMPE results. A favorable comparison will further increase the confidence in the results reported herein.

4.1.1.2 Workload Characterization Improvements

A more comprehensive, validated WWMCCS workload characterization package is needed. The synthetic workloads used in these tests do not completely represent site workloads; on-line activity is represented as resource demands by synthetic batch jobs, memory reference pattern parameters cannot be varied, and job inter-arrival intervals are not variable. SCF data, used as input to the workload generator, and the design of the generator do not permit a totally representative workload to be produced.

Additional work is needed to identify key workload parameters that accurately determine and reflect the characteristics of a H6000 workload. Then, further work is needed to modify the workload generation process so that the desired representativeness is achieved.

Other possible extensions are the use of software and hardware monitor data as additional inputs to a synthetic workload generator. Augmenting the generator with such information would enable the inclusion of on-line activity and memory reference patterns as workload variables.

Extending the workload variables to include memory reference patterns would allow the test of cache memory systems. Cache systems have performance characteristics strongly influenced by the memory reference patterns of executing programs. Synthetic workloads used in the testing of cache memory systems must incorporate reference pattern information reflecting actual site program coding practices.

4.1.2 Hardware and Software Configuration Constraints

a. The primary purpose of the tests was to obtain information about the changes in performance attributable to changes in the number of processors in a system. The number of processors in a system cannot reasonably be varied independently of some other system resources, such as SCU's and memory. Of the many processor/SCU/memory configurations possible, eight were used in the tests. These eight were chosen to reflect the existing WWMCCS HIS 6000 configurations and the large multiprocessor configurations into which they might reasonably grow.

b. The I/O subsystem consisted of DSS 190 disks only and was held constant over all configurations.

c. GCOS System Release WW 6.2.1 was used with the security module intentionally disabled. As a result of disabling

the security module, normal purging of temporary files and the non-trivial amounts of time involved were eliminated.

d. Two software monitors, SYRUP-II and MSM, were run continuously during the test. Their execution created an unknown, but assumed constant, amount of additional system overhead.

4.1.3 Operational Conditions

a. During the conduct of the testing (e.g., the actual execution of the test workloads), operator intervention was strictly curtailed. No console entries were allowed unless absolutely necessary for the continuation of the test. In the few cases where significant intervention became necessary, the test workload was rerun.

b. The synthetic jobs of a synthetic workload were fed to the system all at once (i.e., all jobs were put into the card reader at once) in the order in which they were produced by the synthetic workload generator. This resulted in jobs with identical resource utilization characteristics being grouped together in the input deck, the small jobs first, and all jobs were assigned the same priority.

4.2 Observations

The first two observations below are related to the performance of the configurations tested. The last two are not related directly to the test goals but are based on information obtained from the test activities.

4.2.1 Relative Throughput

As defined in Section 3.4.2, relative throughput is the ratio of elapsed times for a given workload run on a uni- and multiprocessor configuration. The elapsed time for the uniprocessor configuration was selected to be the normalization factor used to compare the change in elapsed time of a given workload run on several configurations.

4.2.1.1 Observed Increases in Relative Throughput

In all experiments where additional processors were added to a test configuration, increases in relative throughput were observed. The range of relative throughput improvement varied in an almost linear manner for some workloads and asymptotically for others.

In Section 3.4.2 a group of tables and figures depict the relative throughput changes. Tables 3-X and 3-XI show the results for H6060's with and without memory interleaving activated. Table 3-XII shows the results for H6080's with memory interleaving on. For the H6060's with memory interleaving off, a maximum relative throughput change of 260.7% was observed for the CPU-bound workload E (IO/CP ratio of 0.01) and a change of 31.2% was observed for I/O-bound workload 3 (IO/CP ratio of 3.34). With memory interlace on, the maximum relative throughput change was 273.4% for workload E and 32.6% for workload F(1) (IO/CP ratio of 3.07). The H6080 tests with memory interlace on showed a maximum relative throughput change of 301.7% for workload E and a maximum change of 0.3% for workload F(1) (IO/CP ratio of 5.07).

The different IO/CP ratios for the same workloads run on H6060's and H6080's (e.g., workload F(1)) are due to the difference in the speed of the processors, the I/O time remains relatively fixed and a change in IO/CP ratio results.

4.2.1.2 Impact of Memory Interlacing

As mentioned above, the test workloads were run on configurations using the memory interlacing feature both on and off. The maximum degree of interlacing allowed by the symmetry of the hardware was always used.

For all cases (except the one instance noted in Section 3.4.1) where interlacing was activated, an increase in relative throughput was experienced over the non-interlaced configuration running the same workload. Relative throughput improvement due to interlacing had an observed average value of approximately 5% while the maximum improvement was nearly 18%.

Because definitive data is not available on the memory reference patterns of the test synthetic workloads, the observed impact of memory interlacing may be biased. Therefore, the interlacing effectiveness and its influence on relative throughput cannot be rigorously quantified.

4.2.1.3 Other Observed Influences on Relative Throughput

Small increases in relative throughput were observed in cases where the number of processors remained fixed, but the number of SCU's and memory size changed. *Recognizing that other factors* affect the observed relative throughput, the analysis concentrated only on the effects of varying processors. Other factors were considered and observed to be ancillary; no detailed analysis of their effects was accomplished.

4.2.2 I/O Generated Processor Time

As part of the calibration process for the synthetic workload generator, the amount of processor time per synthetic program I/O loop on the HIS 6060 was determined to be 4.41 ms. (See Section 3.1.1.) Because of the way in which the loop was coded, this figure can be considered to be the minimum amount of CPU time required to start an I/O operation on the H6060.

The total processor time required for an I/O operation includes the interrupt handling and redispatch times as well as the start time. These other times are not directly available from the test data and hence have not been determined. Knowledge of any of

these times, however, is useful in estimating resource utilization and elapsed time characteristics of applications programs and in the development of analytic or simulation models.

4.2.3 Temporary File Purging Delays

During the preparations for the tests, it was determined that the WWMCCS security module (FS49), which purges temporary disk areas before and after use, substantially delays job or activity initiation and termination. (See Sections 3.1.6 and 3.2.2.) For example, in processing one synthetic workload the average elapsed time between jobs starts, under a condition of queued jobs and available memory, was 1 minute 15 seconds with the security package enabled and only 2 seconds with the package disabled.

The difference in time indicates a large amount of resource usage by the purge package. For a facility that runs a large number of short jobs, the total throughput capability may be seriously reduced by the operation of the security package.

4.3 Application to WWMCCS ADPE

The relative throughput results of the tests cover a broad range of HIS 6000 hardware configurations that reflect WWMCCS installations, and the test workloads cover a broad range of possible IO/CP ratios. The results are applicable to WWMCCS sites within the context of the limitations outlined in Sections 3.4 and 4.1.

The results of the relative throughput experiments are presented in graphical form using the number of processors and interleaving status as the hardware configuration identifiers. Test workloads are distinguished by the IO/CP ratios used to plot the relative throughput curves. The ranges of the IO/CP ratio are 0.01 to 3.07 for the H6060's; and 0.01 to 5.07 for the H6080's.

Figures 4-1 and 4-2 present the variation of the relative throughput with the increase in the number of processors for the H6060 system with memory interleaving on and off, respectively. Figure 4-3 presents similar variation for the H6080 system with memory interleaving on. In these figures the configurations are distinguished from one another by the number of processors. In some cases, the curves are not for individual workloads. Because of the coalescing of the data points, several workloads have been grouped together and the trend has been indicated by smooth curves. For these cases, the ratios (of I/O time to CPU time) grouped are indicated on the right-hand side. In those cases for which the curves are for individual workloads, the data points for the two memory sizes of each configuration are averaged and the trend indicated by a smooth curve with the ratio for the corresponding workload appearing on the right-hand side.

These summary graphs can be used to derive a preliminary estimate of potential gains in throughput as additional central processing units are added to a configuration. However, the analyst must determine the characteristics of the current real and/or projected site workload prior to using these summary graphs (normally this can be done from system produced information maintained at the site). Once the characteristics of the site workload are determined, the graphs can be used to interpolate between the data collected from the controlled test cases. Using this procedure, a simple estimation of potential throughput for the real workload can be derived.

Refer to Figure 4-2 and consider a WWMCCS site using a dual H6060, not using interleaving, and having an average IO/CP ratio of 2.5. By drawing a curve of the same general form as that given for the IO/CP ratios of 2.26 and 2.78 to approximate a ratio of 2.5, the relative throughput that can be gained by adding more processors

may be estimated. If such a curve is drawn, the IO/CP ratio of 2.5 is seen to have a relative throughput value of 1.4 with two processors and with the addition of a third processor, the value becomes 1.6. The percentage improvement in relative throughput expected would be approximately 14%. Example cases for more CPU-bound workloads would have a greater percentage of improvement.

Figures 4-1 and 4-3 can be used in exactly the same manner to yield estimates of relative throughput as a function of workload characteristics.

The quantitative data collected during this study effort will be valuable to WWMCCS system analysts and planners when investigating alternative ways to increase system capacity to meet expanding workloads. Careful use of the results of this testing effort, when combined with space-availability and other site considerations, will permit analyzing costly configuration decisions and in assessing the relative merit of adding CPU's to solve the expanding workload problem.

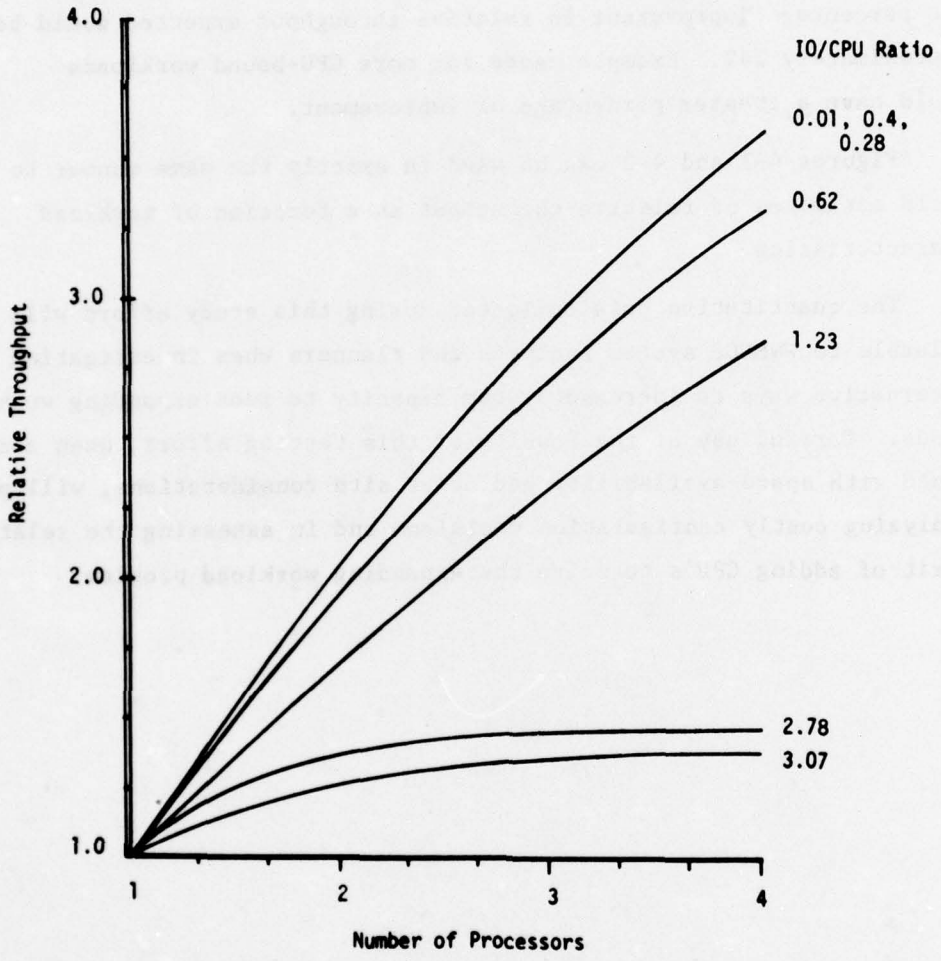


Figure 4-1. Relative Throughput - H6060 Interleaving On

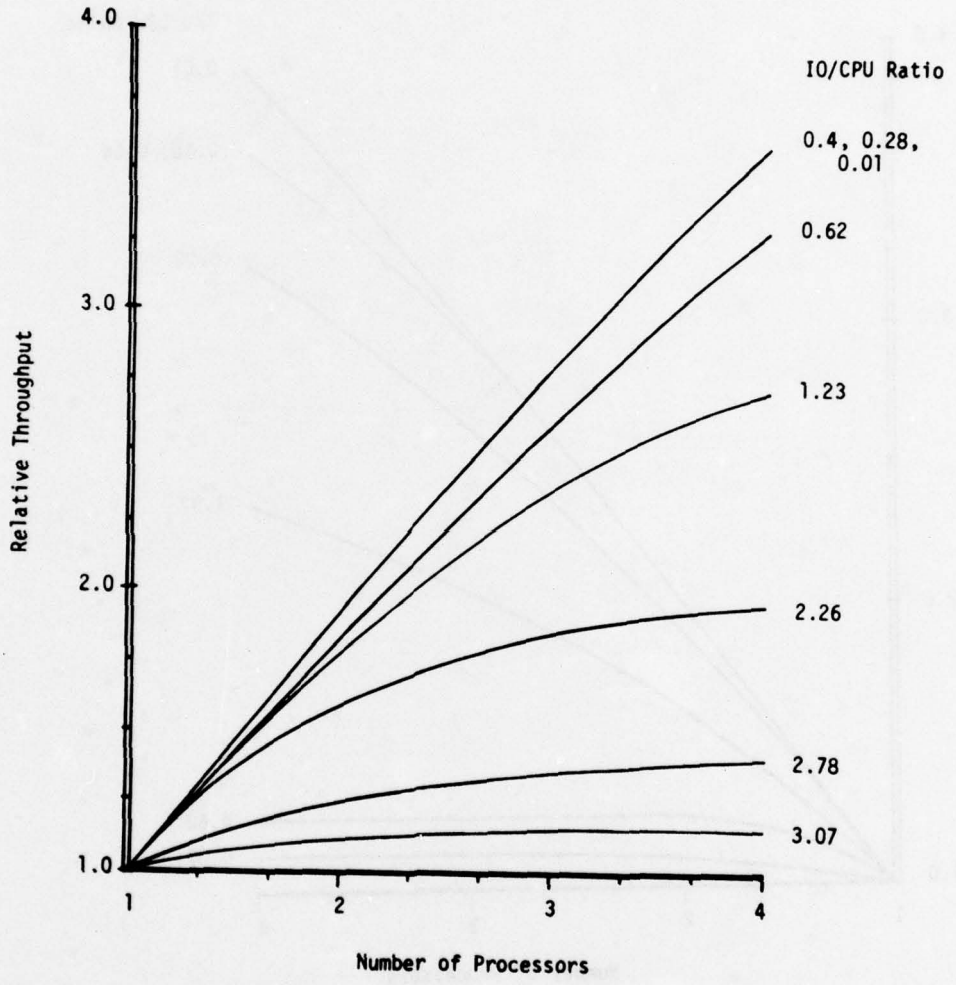


Figure 4-2. Relative Throughput - H6060 Interleaving Off

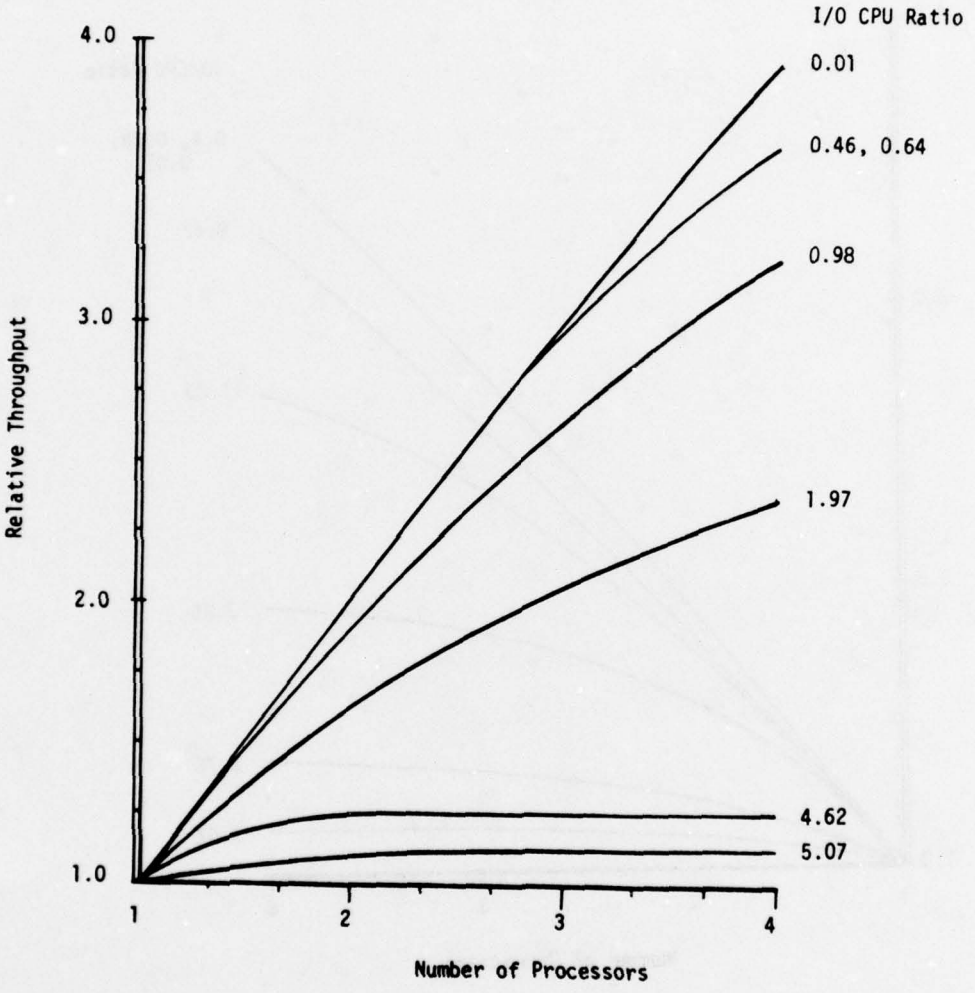


Figure 4-3. Relative Throughput - H6080 Interleaving On

APPENDIX I

VERIFICATION STUDY

Following the formalization of test results and conclusions presented in this report, a validation effort was undertaken by personnel who were not members of the project team. Conceptually, the validation would consist of the following steps: generation of synthetic workloads from source workloads not previously used in the test; benchmarking these workloads on several of the configurations used in the test; and comparing the relative throughput obtained in these benchmarks with that predicted in the test results and conclusions. It was implicit that an extensive validation could not be accomplished, as this would require a magnitude of effort approaching that of the original test. For the validation, two synthetic workloads were generated from two monthly SCF tapes from TAC. These will be referred to as workloads T1 and T2. These workloads would be run on H6060 configurations A, C, and E, representing single, dual, and triple CPU configurations, respectively.

The benchmarks were run at the CCTC computer facility at Reston, Virginia, on 3 December 1976. Because of hardware difficulties, only configurations A and C were available.

The measured IO/CP ratios on configuration A were 2.51 for workload T1 and 2.58 for workload T2. The test workloads closest in IO/CP ratio to these values were workload 9 (2.39) and workload F2 (2.96). The benchmark results for these four workloads (T1, T2, 9, and F2) are presented in Tables I-1 (configuration A) and I-2 (configuration C). Examination of multiprogramming depth (MPD), Memory Utilization, and Processor Utilization (especially on configuration C) reveals qualitative differences between the processing of workloads T1 and T2 and that of workloads 9 and F2. Basically, it

Table I-1

Test Data - Configuration A (Interlace Off)

Wkld	Elp Time (min)	CPU Time (min)	I/O Time (min)	IO/CP	MPD Max	MPD Avg	% CPU PØ	% CPU P1	Average Mem Util.
9	57.8	36.0	86.4	2.39	16	9.3	95.9	-	222.9
T1	58.2	30.7	77.1	2.51	17	10.8	88.8	-	225.3
T2	64.6	34.2	88.4	2.58	16	11.1	95.9	-	237.5
F2	19.4	10.0	29.6	2.96	13	8.0	85.6	-	206.2

Table I-2

Test Data - Configuration C (Interlace Off)

Wkld	Elp Time (min)	CPU Time (min)	I/O Time (min)	IO/CP	MPD Max	MPD Avg	% CPU PØ	% CPU P1	Average Mem Util.
9	43.6	36.9	84.5	2.29	15	9.2	67.4	69.7	225.1
T1	38.6	32.5	77.5	2.38	17	10.8	81.6	77.6	231.1
T2	15.8	37.2	92.0	2.47	17	10.6	79.4	75.2	237.5
F2	15.5	10.7	28.4	2.66	11	7.1	60.1	59.9	192.9

appears that jobs were generally smaller for T1 and T2, leading to higher multiprogramming depth (MPD), and hence higher CPU utilizations, especially on the dual processor configuration C. As this consideration is independent of the primary workload characterization parameter (IO/CP ratio), a comparison of the relative throughput measured for workloads T1 and T2 with that for workloads 9 and F2 alone is not meaningful. Because of this consideration, relative throughput data from all test workloads was examined to determine whether the throughput measured for T1 and T2 was consistent with this data.

Relative throughput (configuration C compared to configuration A) is plotted as a function of IO/CP ratio in Figure I-1 for all test workloads and for workloads T1 and T2. A linear regression model was developed for the data points representing the test workloads (i.e., excluding workloads T1 and T2). The results of this analysis are depicted in Table I-3. Good correlation is indicated, and the F statistic is significant at the 99% level. As indicated in Table I-4, the observed values for workloads T1 and T2 lie within the 90% confidence interval for this regression. Therefore, it is concluded that the validation data (workloads T1 and T2) is consistent with the test data used in the study.

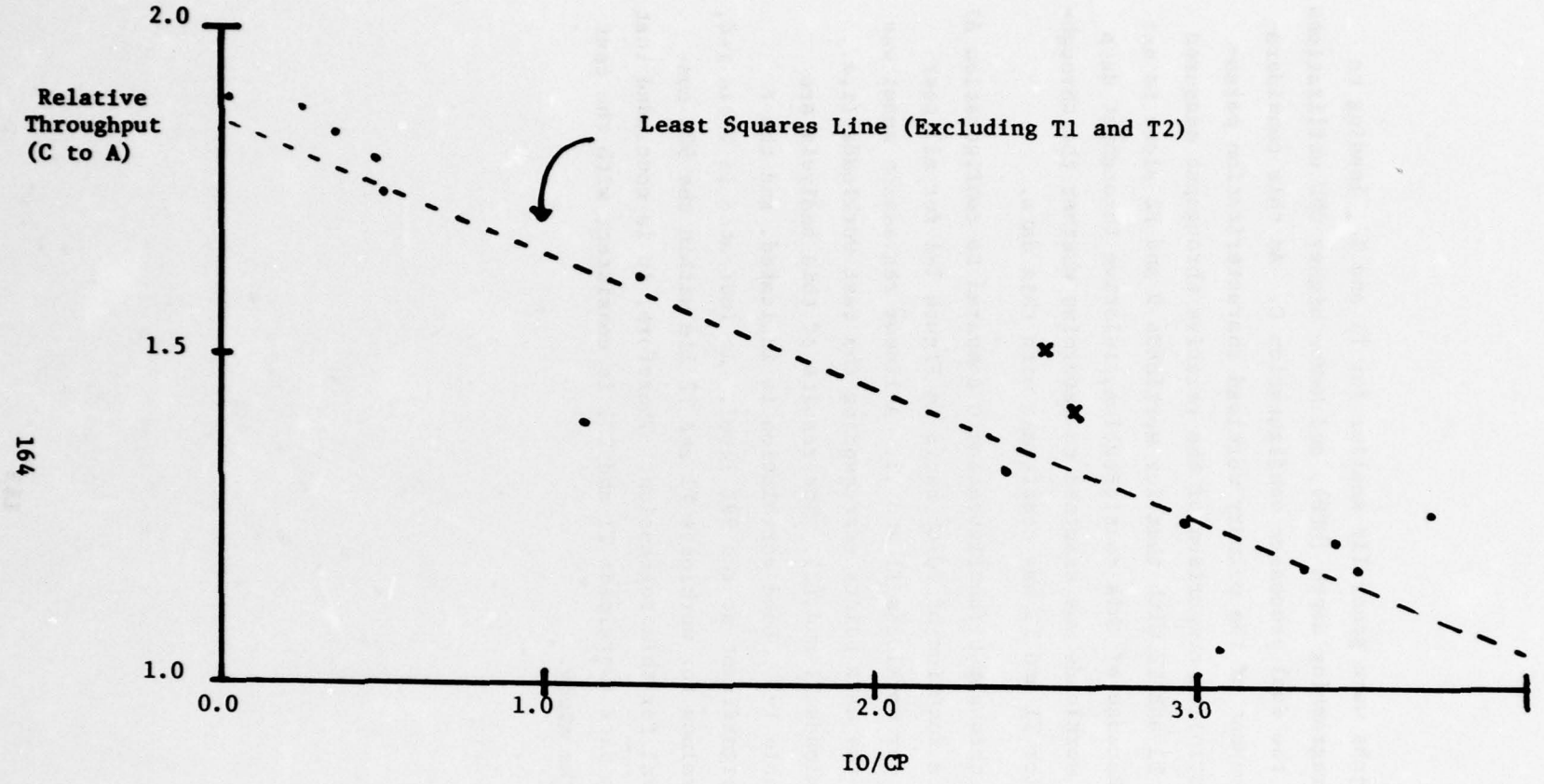


Figure I-1. Validation Data

Table I-3

Linear Regression of Test Data
Rel. Throughput (C to A) Vs. IO/CP

$$Y = 1.847 - .201\ X$$

Coefficient of Correlation = - 0.939

Coefficient of Determination = 0.882

	Degree Free	Sum Squares	Variance Estimate
Regression	1	1.071	1.071
Remainder	12	0.143	0.072
Total	13	1.215	

$$F_{1,12} = 89.737$$

Table I-4

Validation Workload Analysis

Workload	x	y	Regression $\hat{y}$	90% Limits on y Lower	Upper
T1	2.51	1.51	1.34	1.16	1.53
T2	2.58	1.41	1.33	1.14	1.52

APPENDIX II

SURVEY RESULTS

The initial survey of other activities related to the WWMCCS Multiprocessor Performance Evaluation task was conducted in August and September 1975. The information obtained during the survey is briefly discussed below:

1. System Development Corporation (SDC)

A study by SDC, published as TN-CN-1040/003/00, was reviewed. This study used a system simulation to investigate the effect of increased memory capacity for a fixed workload on a H6080 dual-processor configuration. The study report addressed only the dual-processor configuration and asserted, but did not substantiate a 28% GCOS overhead requirement.

2. North American Air Defense Command (NORAD)

NORAD is presently involved in a facility upgrade to the Cheyenne Mountain Complex. In addition to the upgrade, but not directly related, there have been two informal studies conducted that have WWMCCS ADP performance connotations. The SDC study (just described) was one and a conceptual plan for a quadruple processor to handle the combined SCC and NORAD Command System (NCS) workloads was considered briefly. Neither of the above efforts provided conclusive multiprocessor throughput information.

3. Pacific Command (PACOM)

PACOM investigated the pros and cons of a regional data processing center. This effort was related to five WWMCCS computers installed in Hawaii plus an additional central processing unit (CPU) to be acquired by NAVY FLEET OPS. All five systems were HIS 6000

PRECEDING PAGE BLANK-NOT FILMED

single processor machines. There was a concern at PACOM that not enough core memory was available to application programs because of heavy core requirements for system software. A perceived advantage of the proposed consolidation effort was the amount of installed core memory that could be available to application programs if fewer copies of the system software were needed. Thus, PACOM was concerned about the throughput of various multiprocessor configurations.

4. Command and Control Technical Center/WWMCCS ADP Directorate (CCTC/WAD)*

Several efforts related to the WWMCCS multiprocessor performance evaluation task were sponsored by CCTC. CCTC was funding at a very low level, a Navy Post Graduate School (NPGS) study to define a set of queueing equations that would represent the WWMCCS ADP system. Some of the equations had been tested on the CCTC computer. The equations were for a dual-processor configuration with infinite channels and equal priority for all jobs.

FEDSIM was engaged in the development of a WWMCCS Tuning Guide under CCTC sponsorship. This work is described below. In the past, CCTC had contracted with FEDSIM for an H6000 sensor probe-point document.

Although these CCTC efforts did not provide all of the information required by the AFWWMCCS sites regarding multiprocessor performance, they did provide pieces of the needed information. CCTC had available hardware resources to connect single-, dual-, and triple-processor H6060 configurations and agreed to support the pilot testing for the AFWWMCCS multiprocessor study.

* In January 1976, JTSA was made part of the Command and Control Technical Center, primarily the WWMCCS ADP Directorate.

5. Federal Computer Performance Evaluation and Simulation Center (FEDSIM)

As noted above, FEDSIM was developing a WWMCCS H6000 Tuning Guide under a work statement from CCTC. This work is directed toward the areas of: a) tuning techniques applicable to the WWMCCS standard system; b) system variables available for tuning; and c) effect of changing performance variables on system efficiency and throughput. The end product is to be an H6000/WWMCCS tuning document to be used by a WWMCCS performance analyst and will describe such performance tuning procedures and disciplines as will enable the WWMCCS performance analyst to conduct a successful tuning diagnosis.

This work will not provide multiprocessor performance data, but should provide specific system tuning guidelines. The guide will address dual-processor configurations but not triple- or quadruple-processor systems. The WWMCCS Tuning Guide has not been released as of November 1976.

6. Honeywell Information Systems (HIS)

HIS has been involved for some time in H6000 multiprocessor performance evaluation. A memory interference and interlacing study was conducted about two years ago by Ken Norlund (HIS) and Dr. Stan Siegal (DCA). The study was conducted using GCOS 2.1, a CPU-bound jobstream, and single-, dual-, and triple-CPU configurations. The study results were not published, but draft documents contain the following single-valued results:

CPU's	Performance Ratio
1	1.0
2	1.8
3	2.4

7. Air Force Data Systems Design Center (AFDSDC)

AFDSDC is supporting a Data Project Directive (DPD) which tasks AFDSDC with overall performance evaluation responsibility for AFWWMCCS. As part of the AFDSDC support to AFWWMCCS, a performance evaluation effort for HIS Series 6000 was initiated. This effort is part of a larger task entitled AFWWMCCS Configuration Control and was in the embryonic stage at the time of the survey.

In August 1975, it was determined that AFDSDC could not take the lead role in the WWMCCS Multiprocessor Performance Evaluation task. However, there were several related activities at AFDSDC which could support the WMPE task. In support of AFWWMCCS Configuration Control, AFDSDC receives Statistical Collection File (SCF) tapes from AFWWMCCS sites. Some of this data was used for the WMPE task. Another AFDSDC activity, involving H6000 simulation using the *System and Computer Evaluation and Review Technique* (SCERT) software, was directed primarily to H6000 I/O operations and algorithms but there was interest in expanding the effort to encompass all of GCOS. A third effort was related to System Accounting and Resource Analysis (SARA) implementation on the HIS Series 6000. SARA is an analyzer and formatter for SCF data.

AFDSDC has demonstrated interest in performance evaluation but had not addressed the multiprocessor relative performance issues. (AFDSDC supported ESD in accomplishing the multiprocessor evaluation task described herein.)

8. Naval Command Systems Support Activity (NAVCOSSACT)

There were several performance evaluation efforts in progress at NAVCOSSACT during the survey period, but none was specifically related to multiprocessor performance. Four separate, but coordinated efforts were identified at NAVCOSSACT. The first was a Performance

Evaluation Group concerned with both UNIVAC 1108 and HIS 6060 performance measurement. This group had done HIS 6000 hardware monitoring to determine the correlation between SCF data and hardware monitor data. The report is available as NAVCOSSACT TN-17. They also conducted a performance evaluation of the HIS 6080 dual processor computer system at Naval Command Systems Support Activity. The report is available as NAVCOSSACT TN-20. This group developed a software package called Trace Reduction and Analysis Package (TRAP). TRAP is a collection of software routines designed to reduce data collected from cyclic dumping of the GCOS trace table. This package was not completely operational at the time of the survey but was being used to determine operating conditions in the Message Input Processing (MIP) software and to locate congestion causes in MIP. Some work has been done to instrument, with software, the transaction processing system in order to determine transaction processing times and delays.

The Support for WWMCCS Accounting Programs (SWAP) was developed at NAVCOSSACT. This system is designed to support operating statistics based on System Collection File (SCF) data.

A third effort was specifically aimed at improving the performance of applications programs. This effort was focused on UNIVAC 1108 applications. Improvements in applications program performance have been very noticeable and this effort was considered very successful at NAVCOSSACT. Efforts to improve H6000 application programs have been started.

The fourth effort at NAVCOSSACT was related to system simulation using a Computer Aided System Evaluation (CASE) simulation of the H6000 with a workload derived from SCF data. A log analyzer program has been written to produce graphical reports from SCF data.

No relative performance data for multiprocessor configurations was found at NAVCOSSACT; however, reports on most of the above efforts were obtained and working level relationships were established.

9. Navy Automatic Data Processing Evaluation Support Office (ADPESO)

ADPESO is primarily responsible for computer system selection for the Navy. Because of the nature of its mission, the majority of ADPESO work is not oriented towards synthesis of resource utilization from actual workloads but instead towards benchmark programs made up of functional kernels.

An informal study of multiprocessor performance was conducted at ADPESO for HIS 6000 multiprocessor configurations of up to four processors. The study also included IBM and UNIVAC multiprocessors. The report of the study was not published because of difficulties maintaining rigorous configuration control during the test. During the study, a set of queueing equations were developed which produce essentially the same performance ratios as the test results.

At the time of this survey, there was no multiprocessor performance evaluation activity at ADPESO.

10. Electronic Systems Division/Directorate of Computer Systems Engineering (ESD/MCI)

This agency has been involved in Air Force computer evaluation studies for several years, recently completing a study of the SAC PACER system which utilizes an HIS 6000 computer. However, ESD/MCI has not conducted studies to determine the relative throughput of H6000 multiprocessor configurations.

REFERENCES*

1. "Series 6000 Summary Description," Honeywell Information Systems, Incorporated, DA48 Rev. 1, 1972.

2. J.S. Glenn, "The WWMCCS Computers: Part 1, System 6000 Overview," unpublished lecture aids, April 1974.

3. A.C. Shetler and T.E. Bell, "Computer Performance Analysis: Controlled Testing," R-1436-DCA, Defense Communications Agency, Joint Technical Support Activity, April 1974.

4. T.E. Bell, B.W. Boehm, and R.A. Watson, "Computer Performance Analysis: Framework and Initial Phases for a Performance Improvement Effort," R-549-1-PR, RAND Corporation, November 1972.

5. "System Startup," Honeywell Information Systems, Incorporated, DD33 Rev. 1, May 1975.

6. "Summary Edit Program," Honeywell Information Systems, Incorporated, DD24A Rev. 0, January 1975.

7. "Variability of Key SCF Statistics," Air Force Data Systems Design Center, A08-76, March 1976.

8. W. Buchholz, "A Synthetic Job for Measuring System Performance," IBM Systems Journal, Volume 8, Number 4, 1969.

*Other sources that are not in the public domain and hence cannot be included in this compilation have been used in preparing this report.

BIBLIOGRAPHY

Air Force Data Systems Design Center, "Data Project Plan," August 1975.

Crothers, C.G., "Workload Determination and Representation for On-Line Computer Systems," ESD-TR-74-54, Electronic Systems Division, AFSC, Hanscom Air Force Base, Massachusetts, January 1974.

Esposito, J.E. and K. Sreenivasan, "A Semi-Empirical Approach for Evaluating Computer System Performance," ESD-TR-73-297, Electronics Systems Division, AFSC, Hanscom Air Force Base, Massachusetts, December 1973.

Esposito, J.E., "Statistical Analysis to Determine Digital Computer Workload Characteristics," ESD-TR-74-175, Electronics Systems Division, AFSC, Hanscom Air Force Base, Massachusetts, June 1974.

Federal Computer Performance Evaluation and Simulation Center, "PEM Tool Evaluation," Volume I, April 1974.

Federal Computer Performance Evaluation and Simulation Center, "PEM Tool Evaluation," Volume II, April 1974.

Federal Computer Performance Evaluation and Simulation Center, "PEM Tool Evaluation," Volume III, April 1974.

Federal Computer Performance Evaluation and Simulation Center, "PEM Tool Evaluation," Volume IV, April 1974.

Feldmeyer, B.A., et al., "Preliminary Computer Performance Evaluation Guidelines," MTR-2805, The MITRE Corporation, Bedford, Massachusetts, April 1974.

Grodsky, Dr. H.R., "Performance Measurement Within the HIS Series 6000," ESD-TR-73-119, Electronics Systems Division, AFSC, Hanscom Air Force Base, Massachusetts, April 1973.

Honeywell Information Systems, Inc., "Dispatcher and Peripheral Allocation SMD," BR20, January 1972.

Honeywell Information Systems, Inc., "Rollcall, Core Allocation, and Operator Interface SMD," BR21, April 1972.

Honeywell Information Systems, Inc., "I/O Supervision SMD," BR23, May 1972.

PRECEDING PAGE BLANK-NOT FILMED

Honeywell Information Systems, Inc., "I/O Supervision SMD, Addendum A," BR23A, November 1972.

Honeywell Information Systems, Inc., "I/O Supervision SMD, Addendum B," BR23B, January 1973.

Honeywell Information Systems, Inc., "I/O Supervision SMD, Addendum C," BR23C, August 1973.

Honeywell Information Systems, Inc., "Comprehensive Index for Software Reference Manuals," BR28, December 1972.

Honeywell Information Systems, Inc., "Time-Sharing Executive SDM," BR29, January 1974.

Honeywell Information Systems, Inc., "Series 6000 Summary Description," DA48, October 1975.

Information Systems Technology Applications Office, "Computer Accounting Study (CASS)," March 1974.

Information Systems Technology Applications Office, "Data Project Plan for Computer Performance Evaluation," March 1974.

Information Systems Technology Applications Office, "Honeywell Information Systems Computer Performance Evaluation Sensor Point Reference Manual" (MCI-75-5), March 1975.

Information Systems Technology Applications Office, "Computer Performance Evaluation (CPE) Results for MAC," 7 March 1975.

Joint Technical Support Activity, "GCOS Documentation — Job Processing System Schedules Queue Servicing," Task 301, Subtask 2, March 1975.

Joint Technical Support Activity, "GCOS Documentation — Job Processing Transfer of Control Macros and Module Loading Routines," Task 301, Subtask 2, March 1975.

Joint Technical Support Activity, "Work Statement for WWMCCS Tuning Guide," 7 March 1975.

Joint Technical Support Activity, "GCOS Documentation — Static State Hierarchical Structure," Task 301, Subtask 1. April 1975.

BIBLIOGRAPHY (Concluded)

Joint Technical Support Activity, "GCOS Documentation — I/O Supervision," Task 505, Subtask 1, July 1975.

Navy Automatic Data Processing Equipment Selection Office, "An Experiment in the Use of Synthetic Programs for System Benchmarking," NCC 1974.

Navy Automatic Data Processing Equipment Selection Office, "System for Efficient Program Portability," NCC 1974.

Navy Automatic Data Processing Equipment Selection Office, "Synthetic Benchmark Programs Library List," 10 July 1975.

Oliver, Paul, "Queuing Equations for Multiprocessors," Navy Automatic Data Processing Equipment Selection Office, no date.

Pacific Air Command, "Regional Processing Center Study — PACOM ADP Workload Data," 17 July 1975.

Sreenivasan, K. and A. Kleinman, "On the Construction of Representative Synthetic Workloads," MTP-143, The MITRE Corporation, Bedford, Massachusetts, March 1973.

Sreenivasan, K. and A.J. Kleinman, "Construction and Applications of Representative Synthetic Workloads," ESD-TR-73-212, Electronics Systems Division, AFSC, Hanscom Air Force Base, Massachusetts, August 1973.

Sreenivasan, K., "On the Use of Accounting Data in the Evaluation of Computer System Performance," MTP-149, The MITRE Corporation, Bedford, Massachusetts, January 1974.

Sreenivasan, K., "On a Comparison Between DYNAPROBE Output and LOGGER Data," MTP-159, The MITRE Corporation, Bedford, Massachusetts, May 1975.

System Development Corporation, "GRTS 355 Timing Analysis," 21 April 1975.

☆U. S. GOVERNMENT PRINTING OFFICE: 1977--700-002--69