COMPUTER SYSTEMS DESIGN AND ANALYSIS THROUGH SIMULATION

George K. Hutchinson and John Norris Maguire
Lockheed Missiles & Space Company
Lockheed Aircraft Corporation, Sunnyvale, California

INTRODUCTION

In March 1965, the Lockheed Missiles & Space Company's Computation Center installed a UNIVAC 1107 with two FH880 drums and three UNIVAC 1004 print, card punch and read systems, two of which are remote. This configuration is illustrated in Fig. 1. The Central Processing Unit was scheduled to be upgraded to an 1108 in October 1965. The UNIVAC EXEC II monitor was used for the operating control of this equipment. Under this monitor, the typical job enters the system through a 1004 card reader, queues at a section of one drum designated the Input/Output buffer, is processed by the main frame using the drums for scratch storage and systems routines, occupies I/O drum buffer area for storage of output, and is printed on the originating 1004 system. The monitor "steals" cycles from the main frame to control the flow of data between the 1004's and the drum I/O buffer. The insert in Fig. 1 shows the original storage allocation of the two FH880 drums which allocated 262,000 36-bit words to the I/O buffer.

If the I/O buffer is full when the program residing in the CPU wishes to place output in the buffer, the main frame waits for the monitor to relieve the condition by sending output to the 1004's. During this period, the fast, expensive CPU logs "green lite" time while waiting for the peripheral data transmission, which takes place at a relatively slow rate. On the other hand, if an adequate input queue is not maintained on the I/O buffer the CPU must wait in a similar manner for a 1004 which is reading cards. The determination of the proper I/O buffer size and peripheral capacity immediately became a major concern to the systems group. Other 1107 users were contacted in an attempt to utilize their experience in actual usage or in the method of solution. The information gathered in this manner was useful in respect to preventing or solving many operating problems, but was not directly applicable to the buffer size/peripheral capacity decisions because of the difference in workload characteristics. The configuration illustrated in Fig. 1 was chosen as an interim one on the basis of the known requirements, estimated requirements, and a forecast of the total workload.

More accurate information on the workload characteristics and requirements became available after the initial system was installed and a more specific plan for the conversion of programs from other computers was laid out. At this point, the system designers turned to simulation as the tool to aid in the decisions for the 1108 system configuration scheduled for October 1965. Coupled with more accurate workload information was the recent
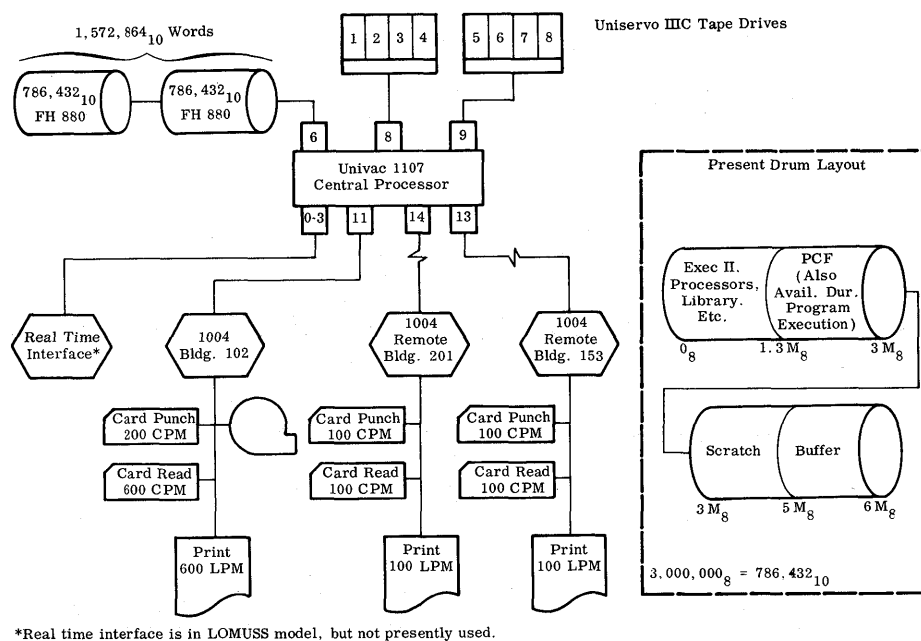
Figure 1. UNIVAC 1107 system configuration.

development of a generalized data processing model developed by Lockheed called the Lockheed Multiprocessor Simulation System (LOMUSS I). This system provided the vehicle for conducting dynamic simulations of proposed 1108 configurations.

In addition, it was felt that conducting simulations of the existing 1107 system would be useful for validation purposes, determining the system capacity, aiding in evaluating different data transmission systems for the remote stations, and scheduling second and third shift operations.

STUDY OBJECTIVES

The primary objectives of the study were to determine which combination of I/O drum buffer size and peripheral devices and their locations would provide good service at a reasonable cost. The tradeoff between good service and system capacity is a management judgment, but simulations would indicate the relative changes in system performance measures (turnaround time, maximum queues, system cost, etc.) as a result of experimenting with different configurations. This problem was compli-

cated by an increasing workload as programs were generated for and converted to the 1107. At the start of this simulation study, the 1107 was running less than 16 hours per day and, as of the middle of May, the 1107 was on a 24-hour-per-day schedule.

The possibility of the main frame having to wait because of I/O buffer storage saturation could be reduced to any desired level by the addition of sufficient equipment, i.e., increased drum capacity, additional peripherals (1004's), or a faster data transmission capability. Thus, the basic problem was to balance the expected costs of the main frame waiting for space on the I/O drum buffer against the costs of additional drums and/or peripheral capacity. The large number of variables, their complex relationships, and the changing characteristics of the workload made simulation a logical choice for problem analysis.

DATA COLLECTION

The data collection phase of the study was undertaken with the study objectives as guidelines. Two catagories of data were collected, the data necessary

for the construction of the model of the 1107 configuration, and data to be used in establishing the validity of the model.

The 1107 monitor provided a major source of data as it summarized for each job processed the input volumes, main frame time, cards output, and the pages printed. The arrival patterns of the jobs into the system, together with the originating source was obtained from the log-in desk records. The routings that the jobs followed as they moved from processing step to processing step were identified and the equipment requirements at each step were determined. A typical routing is given below.

Frequently, the time at the various processing functions is collected and used as the basis for determining simulated processing times. The Lockheed Multiprocessor Simulation System (LOMUSS I) uses the volume of work to be processed and the processing rate of the specific equipment involved as the basis for processing time calculations during a simulation. Thus, if available, the number of cards input and the number of lines and cards output can be used rather than the time required to perform the function.

Summary data were collected for establishing the validity of the model on such system statistics as main frame utilization, average turnaround time (as a function of workload), average run time of jobs, system operating time, and daily beginning backlog of jobs. The use of these summary statistics in establishing the validity of the model will be discussed in a later section.

MODEL FORMULATION

The Lockheed Multiprocessor Simulation System (LOMUS I) was used in developing the 1107 model. LOMUSS I is a 7,500 SIMSCRIPT instruction, dynamic, generalized simulation system developed for the evaluation of data processing systems. It views a data processing system as an augmented job shop. Jobs are described by the functions to be performed and the volume of work to be done. The job routings may vary from a linear list to a complex PERT-like network of functions. A variable set of requirements may be associated with each function of each routing and the requirements may each be described with a variable level of detail. All of the above information is described for LOMUSS I through data cards. Thus the construction of the 1107 model was actually the translation of the in-

formation gathered in the data collection into LOMUSS I input formats.

The remaining information that was required for the 1107 model was the specification of the model output reports. LOMUSS I provides the analyst, at his option, with a wide variety of reports on such items as "snapshots" of queue lengths and machine status, job processing reports, critical paths of each job, turnaround time reports and machine utilization reports. Snapshot reports are provided at an increment set by the analyst. Summary reports are available at any time interval specified by the analyst. With this reporting power available, care must be exercised that (1) one is not inundated with reports and (2) that the essential information is provided. In the 1107 model, interest centered on the utilization of the main frame and the I/O buffer area, so report options were exercised to provide this information.

The development effort required was one man for approximately one month to achieve a working LOMUSS model of the basic 1107 system. Simulation of one day's operation of the 1107 system required about 10 minutes on an IBM 7094 computer.

MODEL VALIDATION

The following tests were applied to establish the validity of the model: a detailed trace of jobs through the system, a stability analysis, a comparison of summary statistics from the model with the same variables as collected in the data study and, finally, a test of the ability of the model to predict.

Initially, a small number of jobs were submitted to the model in order that a complete trace of their progress through the system could be printed out for detailed analysis. The following consistency checks were made (1) the job's path through the system with its routing, (2) the elapsed time at each function with the volume-rate specified time, (3) assignment of functions to equipment, (4) job processing order when queues developed, and (5) the matching of equipment capabilities with the function requirements.

When the above validity checks were completed, the 1107 model was loaded with the number of jobs (160) that the actual system was running as established by the data collection. The model indicated that this level load could be handled on a 2-shift (16-hour) basis, which corresponded with the actual operation. The load was then increased in

steps, observing the effects on the system at each step. The plots in Fig. 2 were generated at this time. Illustrated in Fig. 2 is the arrival rate pattern of jobs into the system and number of jobs in the I/O drums queue plotted as a function of (1) the time of day and (2) the aggregate workload rate. A

detailed analysis of the simulation output provided information on the "mix" (input, waiting to be punched or printed) of the I/O queue which was manually translated to "words of drum storage required." The model predicted that the quarter million words of I/O drum storage would be saturated
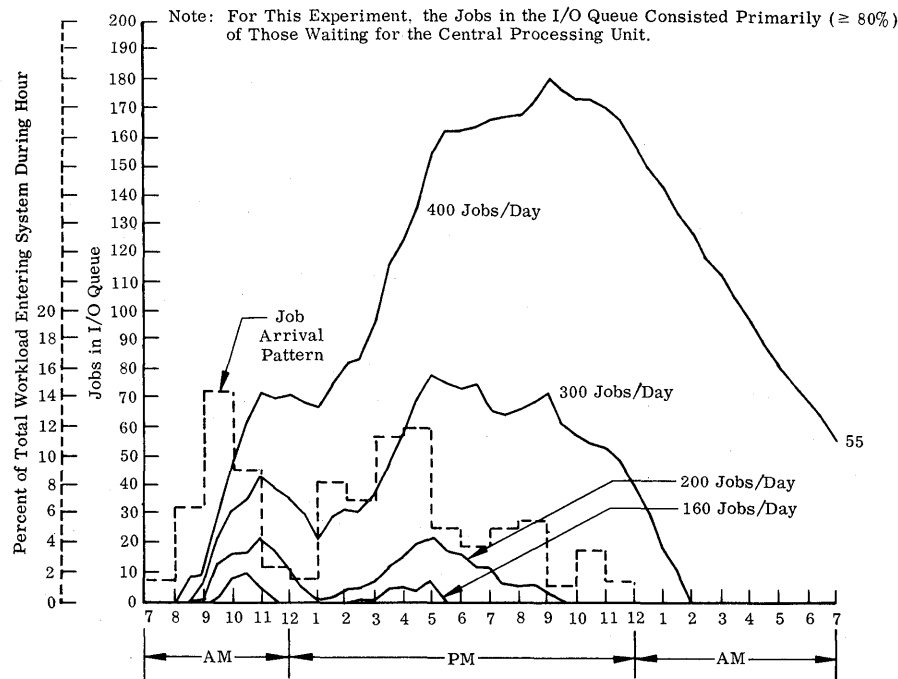

Figure 2. Analysis of UNIVAC 1107 workload rates and drum input/output queues.

in the late morning and late afternoon when the workload rate exceeded 200 jobs per day. This created surprise among some systems and operations personnel who had felt that the I/O buffer was adequate for the 1107 configuration but confirmed the intuitive judgment of others. During the month of April the workload gradually built up until April 8th when a record 215 jobs were processed. On that day, the I/O drum buffer reached a saturation point for the first time. From a scientific point of view, this was only one of many factors in the validation procedure; but from a practical point of view this accurate drum saturation prediction proved to be very significant for purposes of implementing model results in the actual system, which was the responsibility of operations and systems personnel. More system status data (words being utilized on the I/O drum queue, jobs queued at card readers, etc.) were sampled frequently during the following two weeks. This information further

validated the model and lent credence to results obtained with higher workloads rates fed into the model through parameter variation.

During the initial model runs with the actual 1107 workload, the 1107 utilization reported by the model was 54 percent. The actual operating data collected showed 56 percent utilization. The average turnaround time of jobs in the model as a function of workload level was compared with turnaround times obtained from the actual operating system. These data are illustrated in Fig. 3. The differences between the actual and model turnaround times are relatively constant and represent the manual handling times exterior to the hardware of the actual system and the domain of the model. The time between clocking in a job and actually inserting it into the card reader could have been accounted for in the model, but the actual times were found to be about constant and not significant to the problems for which the model was being utilized.
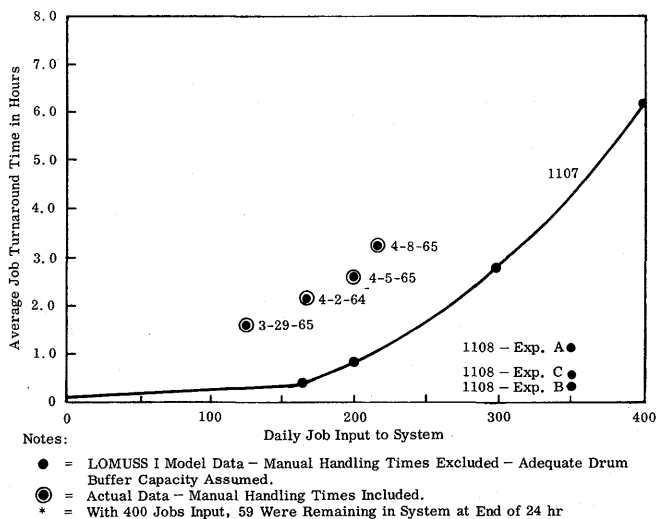
There were some manual procedures that were immediately implemented to minimize the drum saturation problem, namely, when the operator detected a saturation condition he could stop the input flow by stopping one or more card readers and/or he could dump some of the drum output queue onto tape for later processing during a low workload period (which would increase the turnaround time). The model made a contribution by pointing up the importance of the console operator to system efficiency and by so doing, helped accelerate work on improved manual procedures and automatic system status indicators (for example, modification of the executive routine to give an online console printer message when the I/O drum buffer became saturated).

Another pertinent piece of information was the 350-job-per-day 1107 system capacity predicted by the model as illustrated in Fig. 2. This provided a useful guideline to establishing system programmer support levels and program conversion (from other systems to the UNIVAC system) schedules prior to installation of the faster 1108 CPU system.
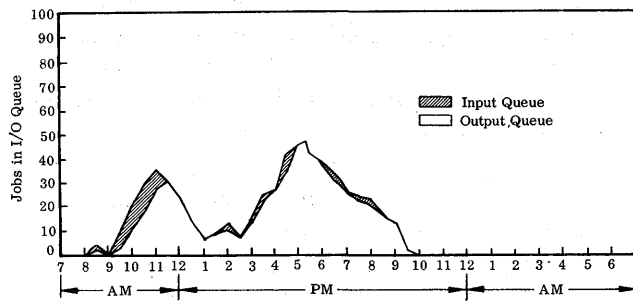
The next step in designing and conducting the simulation experiments was to incorporate some tentative system hardware plans into the model. These plans are summarized as follows:

- Installation of a high-speed data link between the central system and a high-use remote station.
- Installation of an additional on-site UNIVAC 1004 print, card read and punch system.
- Replacing the 1107 CPU with an 1108 CPU.

Three experiments were conducted which recognized the above system changes combined with a forecasted change in the characteristics and level of the 1108 workload. The output from these experiments was the primary information used for determining the 1108 system configuration for October 1965.

Some output from experiment A is illustrated in Fig. 4, where the original 1107 model configuration was run with (1) a 350-job-per-day input which was the maximum forecasted level of work through 1965 and (2) the faster 1108 CPU in place of the 1107. Three system performance measures were significantly affected in experiment A:

- The average job turnaround time for 350 jobs dropped from over 4 hours to about 1 hour.

Figure 3. Analysis of UNIVAC 1107/1108 turnaround time as function of workload (manual handling times exterior to hardware system excluded for LOMUSS I simulation).

SYSTEMS ANALYSIS

The simulations were conducted primarily for determining an 1108 system configuration for October 1965, but information obtained from the 1107 simulations in April was useful to the then operating system.

With the workload characteristics remaining approximately constant, the model showed that a 3-shift operation would be required when the workload rate reached approximately 250 jobs per day. This can be seen in Fig. 2 when a linear interpolation is made between the I/O queue plots of 200 and 300 jobs per day. That is, a queue of jobs will still be on the drum after midnight, which is the end of the second shift, if the workload rate exceeds 250 jobs per day.
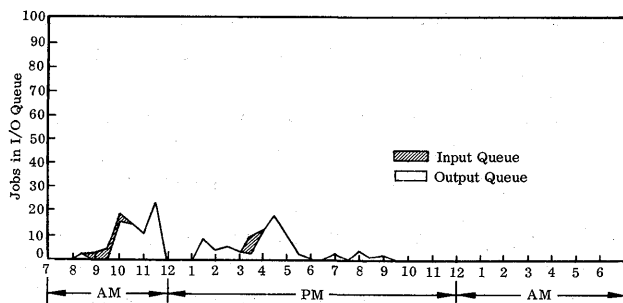
Notes:

1. Average Turnaround Time = 1.02 Hours (Manual Handling Times Exterior to Hardware System Excluded.)
2. 1108 CPU Utilization = 49% (Two Shift Operation).
3. Maximum Requirement for I/O Drum Buffer Storage Occurs at 5:10 PM and Equals 1,860,000 Words.

Figure 4. Experiment A — present 1107 system with 1108 CPU and a 350-job-per-day workload.

- System throughput increased substantially as the 350 jobs were processed in only 2 shifts instead of 3.
- The mix of the I/O queue changed from primarily input to primarily output and the peak I/O drum buffer requirement rose to almost 2 million words.

Similar output from experiment B is illustrated in Fig. 5 where experiment A was repeated except that a faster data transmission facility for a high-

Notes:

1. Average Turnaround Time = 0.36 Hours (Manual Handling Times Exterior to Hardware System Excluded.)
2. Maximum Requirement for I/O Drum Buffer Storage Occurs at 11:30 AM and Equals 900,000 Words.

Figure 5. Experiment B — 1108 CPU with building-153 COAX and additional 1004 system in building 102.

use remote station was added to the model along with an additional on-line, on-site UNIVAC 1004 Read, Punch, and Print System.

Two system performance measures were significantly affected in experiment B:

- The average job turnaround time dropped from 1 hour to 0.36 hour.

- The peak I/O drum buffer requirement dropped from almost 2 million to about 900,000 words.

The output from experiment C is illustrated in Fig. 6 where experiment B was repeated except that the mean run time of the jobs was raised to 6 min-
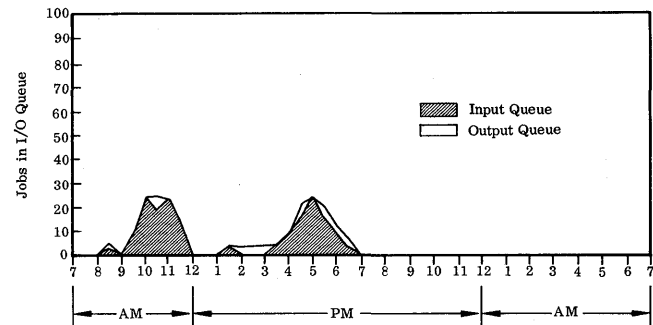
Notes:

1. Average Turnaround Time = 0.45 Hours (Manual Handling Times Exterior to Hardware System Excluded.)
2. Maximum Requirement for I/O Drum Buffer Storage Occurs at 10:30 AM and Equals 354,000 Words.

Figure 6. Experiment C — same as Experiment B but with an average CPU time of 6 minutes instead of 2.54 minutes (as measured on 1107).

utes to more than account for an expected increase in the average run time characteristic of the workload.

The significant result from experiment C was:

- The peak I/O drum buffer requirement was reduced from 900,000 to 354,000 words.

CONCLUSION

The quantitative information provided by experiments A, B, and C was used to revise the tentative 1108 system hardware plan. The significant difference between the old and new plan for the 1108 configuration was the change from all rapid access, expensive UNIVAC 432 drums to a reduced 432 drum capacity and an increased (from ¼ to ¾ million words) I/O drum buffer area using the slower and less expensive 880 drums. The simulations showed that the forecasted 1108 workload would initially be primarily constrained by peripheral data transfer rates and therefore the faster 432 drums would not alleviate this problem. In addition, the peak I/O queues indicated that a larger I/O drum buffer area would provide improved system performance (e.g., reduced turnaround time) and minimize manual intervention.

The LOMUSS model of the Lockheed UNIVAC

on-line, time-sharing, remote terminal system simulated two critical periods (spring and fall of 1965) and provided information upon which the design of the 108 system configuration was based. An effort is continuing which will monitor the level and characteristics of the workload, equipment utilization, turnaround time, etc., for further model validation. The information derived from this continuing effort, coupled with up to date workload forecasts, will determine whether further model experimentation will be conducted for future system decisions.

Lockheed has been constructing general and specific models of computer systems for five years using a variety of simulation languages. As illustrated by the implementation of the above simulation study, the development of LOMUSS has significantly assisted in establishing simulation as a design and analysis tool of the Lockheed Missiles & Space Company.