

313
1/5/78
LA-7034-MS

Informal Report

16. 1724
UC-32 and UC-34c

Issued: November 1977

Data Analysis Facility at LAMPF

Dennis G. Perry
James F. Amann
Harold S. Butler
Cyrus J. Hoffman
Richard E. Mischke
E. Brooks Shera
Henry A. Thiessen



los alamos
scientific laboratory
of the University of California
LOS ALAMOS, NEW MEXICO 87545

An Affirmative Action/Equal Opportunity Employer

UNITED STATES
DEPARTMENT OF ENERGY
CONTRACT W-7405-ENG-36

DISTRIBUTION OF THIS DOCUMENT IS UNLIMITED

CONTENTS

ABSTRACT	1
I. PURPOSE OF THE STUDY AND PARTICIPANTS	1
II. SUMMARY OF RECOMMENDATIONS	2
III. INTRODUCTION	2
IV. HISTORY OF DATA ANALYSIS PROBLEM	2
A. Past Studies	3
B. Incompatibilities with CCF Operation	3
1. Access to the CCF	4
2. Security	4
3. Allocations	4
4. Magnetic Tapes	5
V. ANALYSIS OF REQUIREMENTS	5
A. Past Estimates	5
B. EPICS/HRS	6
C. General Experience	6
D. Magnetic Tape	7
E. Interactive Graphics	7
F. Memory Size	7
G. Projection of Experience	7
VI. SOLUTION TO REQUIREMENTS	7
A. Minicomputer Plus	8
1. Fast Array Processor (FAP)	9
2. FASTBUS Memory	9
3. Cache Memory	9
4. Fast Floating-Point Processor	9
B. Midicomputer	9
C. Operational Costs	9
1. Building Requirements	9
2. Support	11
3. Administration	11
VII. CONCLUSION	11

NOTICE

This report was prepared as an account of work sponsored by the United States Government. Neither the United States nor the United States Department of Energy, nor any of their employees, nor any of their contractors, subcontractors, or their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness or usefulness of any information, apparatus, product or process disclosed, or represents that its use would not infringe privately owned rights.

DATA ANALYSIS FACILITY AT LAMPF

by

Dennis G. Perry, James F. Amann, Harold S. Butler,
Cyrus J. Hoffman, Richard E. Mischke, E. Brooks Shera, and Henry A. Thiessen

ABSTRACT

This report documents the discussions and conclusions of a study held in July 1977 to develop the requirements for a data analysis facility to support the experimental program in medium-energy physics at the Clinton P. Anderson Meson Physics Facility (LAMPF).

I. PURPOSE OF THE STUDY AND PARTICIPANTS

This report documents the recommendations of a study group appointed to assess the current status of the data analysis facilities available for the direct support of the experimental program at the Clinton P. Anderson Meson Physics Facility (LAMPF), and as a consequence of its assessment, to recommend policies and a course of action for the future which will help make the most effective use of the resources of LAMPF and its user community.

The participants in this study fall into two groups. A larger group provided background information for a smaller data analysis committee which was responsible for the final recommendations. The people invited to this study were the following: *J. Amann, Carnegie-Mellon University; M. Barnes, LAMPF; P. Berardo, LAMPF; *P. Bevington, Case-Western Reserve University; J. Bradbury, LAMPF; *H. Butler, LAMPF; F. Cverna, Case-Western Reserve University; F. Dorr, LASL; K. Hanson, LAMPF; *C. Hoffman, LAMPF; C. Hwang, LAMPF; R. Jameson, LAMPF; S. Johnson, LAMPF; M. Kellogg, LAMPF; T. Kozlowski, Brookhaven National Laboratory; L. Mann, LASL; *R. Mischke, LAMPF; C. Nelson, LASL; *D. Perry, LAMPF; *E. Shera, LASL; J. Sicilian, LASL; D. Simmonds, LAMPF; N. Spencer, LAMPF; *H. Thiessen, LAMPF; and S. Zink, LAMPF. Those names marked with an asterisk constitute the data analysis committee.

Questions or comments regarding this report should generally be addressed to D. G. Perry or H. S. Butler.

II. SUMMARY OF RECOMMENDATIONS

The recommendations of the study group are listed below and discussed in depth in the sections which follow. These recommendations are expected to satisfy the needs of experimental data analysis as presently understood and projected into the next three to five years. If circumstances change rapidly, then these recommendations should be reviewed again and amended appropriately.

- A. Sufficient computing power should be obtained to meet the estimated data analysis load of LAMPF. This may be accomplished with the following computer systems:
 - 1. Begin with the PDP-11/70 purchased in FY 77.
 - 2. In FY 78, purchase sufficient equipment to upgrade two computers--one to an enhanced PDP-11/70 for use at HRS/EPICS and another to a normal PDP-11/70. A third PDP-11/70, or a new 32-bit computer, should be ordered.
 - 3. In FY 79, the 32-bit computer should be enhanced with additional peripherals and a second, matching system should be ordered.
 - 4. A midcomputer or its equivalent in computing power and physical characteristics should be ordered as soon as possible, but arrival time should be no later than the first of FY 80.
- B. A building should be built to house the midcomputer and the other data analysis computer systems. This building should be located near the experimental area and should include work space for users, offices for computer support personnel, and storage space for magnetic tapes and supplies.
- C. The staff of personnel who support the data acquisition and analysis effort--programmers, consultants, operators, and data analysts--should be increased at a minimum by two additional people in FY 78, three additional people in FY 79, and by two additional people in FY 80.

III. INTRODUCTION

As LAMPF has developed, two study groups have been convened to analyze the problems associated with the acquisition and reduction of experimental data and to recommend guidelines for developing the needed facilities. These recommendations were made jointly by, and had the endorsement of, LAMPF users and staff.

Since the last study group met in August 1976, LAMPF has increased its output of experimental data considerably. HRS and EPICS have gone into operation and are producing large volumes of data. The other beam channels have become more prolific. The data analysis problem as foreseen only a year ago has changed because of unforeseen circumstances. Therefore, it was necessary to convene another study group to review the current status of data analysis facilities and to plan for their continued expansion. This study was held at LAMPF on July 18 and 19, 1977, with a concluding session on August 4.

IV. HISTORY OF DATA ANALYSIS PROBLEM

The analysis of experimental data is a preeminent problem at all of the major accelerator centers engaged in nuclear physics research. LAMPF is no exception.

A. Past Studies

The problem of data analysis at LAMPF was addressed formally for the first time in 1970. A study group recommended that LAMPF rely on the LASL CCF for its bulk computing needs, i.e., the computing capacity needed for data analysis.¹ The capacity required was projected to be equivalent to one CDC-6600 within one or two years of startup and equivalent to two CDC-6600's by the time all facilities were fully operational.

Those projections were made at a time when the CCF had an inventory of two CDC-7600's and three CDC-6600's. This computing capacity was expected to grow at an annual rate of 40%. Unfortunately, despite the best efforts of the C-Division management, the rate of growth was barely a third of that projected and allocations became a way of life.

In the face of decreasing allocations of computing time and a growing inventory of data tapes, a second study group was convened at LAMPF in 1976. This group recommended that LAMPF augment the CCF resources with a pair of PDP-11s configured for maximum data analysis capability.² This recommendation was based on tests that demonstrated the PDP-11 was a cost-effective vehicle for reducing data. The idea was to conserve the allocation of CCF time for the final stages of data analysis, which are computation intensive, and for calculations related to accelerator development. In anticipation of increased competition for CCF resources, i.e., the rumored departure of one CDC-7600, and the projected growth in the experimental programs at LAMPF, the group further recommended the acquisition of a medium-sized computer system in FY 79.

B. Incompatibilities with CCF Operation

In the year since the last study, the pressures from the data-analysis problem have continued to increase. HRS and EPICS, the two most prolific data producing facilities at LAMPF, have become operational. Unfortunately, no new computers have been added to the CCF to handle this increased load. On the contrary, the LAMPF allocation has been decreased. This fact is all the more significant because the experience gained last year established that the analysis of data from a certain class of experiments requires a larger word-length machine than the 16-bit PDP-11, thus making the CCF computers more attractive.

But the most important development from last year's experience was the recognition that there is a basic incompatibility associated with doing data analysis at the CCF. This observation is not a criticism of the CCF. The LASL CCF was created to fulfill a certain mission, and that mission dictated a complement of equipment and a philosophy of operation that are incompatible with the requirements for data analysis at LAMPF.

To be more specific, the equipment and operation of the CCF are optimized for large-scale calculations. In contrast, the core of the data analysis task at LAMPF is processing magnetic tapes. A large computing center can be optimized for this task as has been shown by the Fermi National Accelerator Laboratory. However, it is not appropriate to expect the LASL CCF to alter the parameters of its operation to satisfy the criteria for data analysis.

This fundamental incompatibility is compounded by four factors which are discussed in the following sections. No single factor by itself totally justifies the conclusion that the analysis of LAMPF data cannot be done reasonably at the CCF. However, all four factors interact in such a way as to create insurmountable difficulties.

1. Access to the CCF. The LASL CCF is located four miles west of LAMPF. This distance impedes access to the CCF because of the shortage of vehicles and limited parking. Yet many trips are required to transport the magnetic tapes to the CCF and return them to the users. In addition to the logistics of handling the magnetic tapes, there is a nonnegligible probability of losing tapes, thereby diluting the quality of an experiment. The move by the CCF to a time-sharing environment, while commendable in general, aggravates the problem of coordinating tapes with jobs. The net effect of a remote CCF is to reduce the user's efficiency over what it would be in a freely accessible, local facility.

2. Security. The LASL CCF operates on a "secure" environment to protect the classified information in the computers. As a consequence, the only personnel who may have indirect access to all of the LASL computers are those with a DOE "Q" clearance. Uncleared persons at LASL are restricted to a single CDC-6600 and may access it only through remote terminals. These security precautions not only hamper the uncleared user, they impose additional burdens on those who have clearances, so much so that essentially all users have given up trying to use the CCF for data analysis. And it appears that the security regulations will get worse before they get better; for example, the machine for unclassified work will probably not be integrated into the common file system, a decision which will further curtail the computing activities of uncleared users.

3. Allocations. The third factor which is curtailing the data analysis effort is the limited allocation of computing resources granted to LAMPF by the CCF. The current allocation is 33 minutes/day on one CDC-7600 and 31 minutes on another. These two machines operate under LTSS, a time-sharing operating system. On the third CDC-7600, LAMPF is allocated 2.06% of the weekly points--another measure of allocation. An additional hour/day of CDC-6600 time is available under the NOS operating system. Thus, the total CCF allocation is equivalent to one quarter of a CDC-6600 at a time when the needed capacity is four to eight times that allocation.

The CCF allocation to MP Division is supposed to support all of the computing activities at LAMPF, not just data analysis. A review of present usage disclosed that the current CCF allocation is almost totally invested in accelerator design and performance calculations, data base management, and theoretical nuclear physics calculations; very little is left over for data analysis. And the pressures for non-data-analysis calculations are growing.

This discussion of allocations is not meant to be critical of the CCF. The primary mission of the CCF is the support of programmatic weapons development work. As these programmatic requirements have grown, there has not been a corresponding growth in the CCF computing capacity because funds were not available. As a consequence, the CCF capacity was oversubscribed by 8% in FY 77 and is predicted to be oversubscribed by 20% in FY 78 according to C Division. Thus, there appears to be no relief in sight for the problem of allocations.

4. Magnetic Tapes. Data analysis at LAMPF involves the processing of 100 to 500 magnetic tapes per experiment. Depending on the type of analysis and the way in which the data were recorded, a single computer run may process one or more tapes, or multiple runs may be required for a single tape. In any case, the processing is highly interactive and result driven so that the order in which the tapes are processed is not necessarily predictable. This fact dictates that the tapes should be stored at the CCF if they are to be analyzed on the CCF computers. However, the CCF does not have the tape storage capacity for the 10 000 or more tapes that may be in various stages of analysis at any given time. The alternative is to transport the tapes to the CCF; however, that has been shown to be impractical, especially since the current security regulations require that a signed statement be attached to each magnetic tape assuring that the tape contains no classified data. Because of the lack of storage for outside tapes, locating a given tape for a computer run is often difficult. Moreover, the mounting of tapes requires additional effort on the part of already overworked operators and, consequently, jobs involving tapes are subject to long delays. Thus, the proper and speedy handling of magnetic tapes, which is crucial to an efficient data analysis facility, is incompatible with the present boundary conditions on the operation of the CCF.

V. ANALYSIS OF REQUIREMENTS

Before proposing solutions to the data analysis problem at LAMPF, it is necessary to document the requirements that must be met.

A. Past Estimates

The summer study of 1976 surveyed 30 experiments typical to LAMPF to estimate the magnitude of future requirements for data analysis. They projected a total of 10 000 to 15 000 magnetic tapes of data per year from LAMPF. Since each tape holds a few $\times 10^8$ bits of data, LAMPF must be prepared to reduce a few $\times 10^{12}$ bits of data per year given the present array of beam ports. The experience of other accelerator centers indicates that a CDC-6600 can analyze about 10^5 bits /s. Given the number of effective seconds of CDC-6600 time in a year, LAMPF was projected to need an in-house computing capacity equivalent to one CDC-6600.

The study group then addressed the alternatives for acquiring this capacity. By analyzing the same data tapes on a CDC-6600 and a PDP-11/45, it was determined that one CDC-6600 was equal to five to six PDP-11/45's. On the other hand, when actual costs for reducing data on the two types of machines were compared, it was determined that tapes could be analyzed for about half the CDC-6600 cost on a PDP-11/45 because of proportionately lower operating and maintenance costs. Consequently, the group recommended that LAMPF acquire, as soon as possible, two PDP-11/70 class computers and a permanent building to house a data analysis facility. (The PDP-11/70 was specified because it was the highest performance model of the PDP-11 line.)

Moreover, the group recommended that \$300K be budgeted in FY 79 for a major enhancement of the data analysis facility in recognition of the anticipated growth in computing needs as the intensity of the accelerator approached 1 mA.

Unfortunately, these recommendations came after the FY 77 budget allocations had been fixed. The best that could be done was to order one PDP-11/70 for delivery in 1977. No funds were available for a building, so the computer will have to be put in the Terminal Building. A second PDP-11/70 was budgeted for FY 78.

Recently, it has been recognized that while the PDP-11/70 is suitable for the type of analysis used for the estimates last summer, there is another typical class of data analysis codes that needs a larger and faster main frame.

B. EPICS/HRS

EPICS and HRS are two of the largest producers of data at LAMPF. EPICS presently can produce data up to a rate of 100 events/s. The PDP-11/45 at EPICS can analyze about 30 events/s when it is available for data analysis. As utilization of EPICS increases, the data acquisition computer will no longer be available for data analysis. It is anticipated that the data event rate for data acquisition will drop somewhat to about 20 events/s because of the lower counting rates typical of pion experiments.

HRS presently produces data up to a rate of 100 events/s. This rate is expected to increase to about 400 events/s with buffered TDC's and ADC's. The HRS PDP-11/45 can analyze about 40 events/s when available.

EPICS and HRS combined need a data analysis facility capable of analyzing data at the rate of 500 events/s or greater. (It is assumed that a 50% data-taking efficiency due to machine and experiment downtime is cancelled by a need to replay some tapes more than once.) The data analysis facility also requires large amounts of memory (a minimum of 256-K bytes), disk space (a minimum of 50-M bytes), interactive graphics, and the other usual assortment of peripherals required to effectively use a computer.

In addition to the type of analysis described above, i.e., replay, a large amount of time must be spent combining runs, finding areas of peaks, subtracting backgrounds, and normalizing results. Typically, this work takes as many man hours as the replay but requires far less CPU time. This is best done in a time-shared mode and is a prime candidate for moving to a larger time-shared machine.

C. General Experience

An analysis of users' experience with data acquisition has shown that acquisition event rates may vary from 1 to 400 events/s. Each event may consist of 1 to 200 words of 12-bit data. The number of tapes generated per experiment may vary from 10 to 500. The time required to analyze these tapes may vary from 1 to 5 h/tape on a PDP-11/45. From this, one can determine that the time required to analyze a typical experiment may vary from 200 to 1000 h with a nominal value being 500 h/experiment. One can summarize the requirements for these experiments as requiring a moderately fast computer capable of handling large amounts of data on magnetic tape.

When the nominal EPICS/HRS requirements are added to the estimates from all other channels that will be analyzed at LAMPF, the total requires the approximate

equivalent computing power of two CDC-6600's to be available today with an anticipated growth by another factor of 2.

D. Magnetic Tape

One of the major problems of data analysis is the handling and storing of large volumes of magnetic tapes. It is estimated that each experiment will produce from 100 to 500 magnetic tapes. Any data analysis facility must be able to handle tapes from about 60 experiments per year, or anywhere from 5 000 to 30 000 magnetic tapes per year, with 15 000 being a nominal number.

E. Interactive Graphics

Another prime requirement for data analysis is the ability to display and plot results. The ability to interactively display small portions of data while adjusting selection criteria is invaluable in saving data processing time. After final selection criteria, production runs may proceed with plots needed of the reduced data.

F. Memory Size

Many analysis programs require more than the 32-K word limit imposed by the architecture of the PDP-11 computer. In some cases, this may be circumvented by clever programming and segmentation, but in many cases it proves to be a real limit. These cases require memory sizes which are more typical of larger computers, e.g., 256-K bytes per program task.

G. Projection of Experience

Our experience with making projections has indicated that it is difficult to project requirements when data acquisition technology is changing so rapidly that the amount of data that can be collected changes by an order of magnitude in a few years. It is also difficult to project how these needs can be met when their fulfillment depends on plans one cannot influence. A greater degree of control may be had with one's own facility.

Our understanding of the data analysis problem indicates that a solution may be obtained with a modest amount of capital expenditure phased over several years. We also anticipate that a greater volume of data will be acquired as future fast-buffered systems are realized. This increase is projected to be close to a factor of 2 beyond what is required today.

VI. SOLUTION TO REQUIREMENTS

A comparison of several representative computers and a measure of their performance is given in Table I. One measure of performance is derived from the Whetstone Benchmarks.³ In the Whetstone system, the program is first translated into an intermediate instruction code, which is then executed by an interpretive program. The results of analysis of the Benchmark output are given in kilo-Whetstone instructions per second (kWI/s). Another measure of performance is the Normalized Service Unit (NSU) used by DOE. The NSU is based on the available clock time of a CDC-7600. The Whetstone Benchmark is perhaps more

applicable to our situation and mix of jobs since it more closely represents our situation. To satisfy the data analysis needs of LAMPF requires a facility with at least 3600 kWI/s by FY 78 and a computer with certain physical characteristics to be available in FY 80.

TABLE I
COMPUTER COMPARISON

Computer	NSU	kWI/s	Normalized k WI/s	Estimated Capital Cost (K\$)	Cost NSU (K\$)	Cost NKWI/s (K\$)
CDC-7600	1.0	8000	1.0	7000	7000	7000
CDC-6600	0.22	1800	0.23	2000	9091	8696
DEC 10	0.16	1000	0.13	1000	6250	7692
DEC 20	-	1000	0.13	500	-	3846
PDP-11/70+	-	1000	0.13	200	-	1538
PDP-11/70	0.16	700	0.09	150	938	1667
PDP-11/60	-	600	0.08	100	-	1250
PDP-11/45	0.07	245	0.03	100	1429	3333
PDP-11/34	-	45	0.02	75	-	3750

Since some fraction of data analysis (such as theoretical curve fitting) involves the running of very large codes, it would be useful to have a computer that would run these codes. Certain other types of data analysis also require large codes (such as ray tracing). All data analysis codes appear to be compute-bound, and the speed of these machines is an advantage for compute-bound jobs. A possible solution is to choose a CDC-7600 or some multiple number of CDC-6600 computers. The acquisition of such a large computer, however, has several disadvantages. One is the very high purchase and maintenance costs in comparison to their performance. Another disadvantage is that large machines require carefully controlled environments, which means substantial costs for a building and its maintenance. It is also estimated that its operating costs would be very high; at least six to eight full-time personnel would be required to operate and manage such a facility. In addition, much of the software used for data acquisition would have to be rewritten for data analysis. Also, the lead time for procurement of a major computer system is prohibitive.

With this in mind, it is apparent from Table I that there exist more cost-effective solutions to the data analysis problem at LAMPF. One that this section will discuss is based on a medium-sized computer coupled with several minicomputers with special hardware to enhance their performance.

A. Minicomputer Plus

It is estimated that most of the EPICS and HRS data analysis can be done on a minicomputer complex with the equivalent computing power of 1600 to 2000 kWI/s. It is proposed that this need be met with an enhanced PDP-11/70 computer that would provide about 1000 kWI/s. This machine (to be delivered in FY 78) in conjunction with a PDP-11/70 (~700 kWI/s) delivered in FY 77 and another PDP-11/70 or 32-bit computer to be delivered in FY 78 would give a computing power of about

2400 kWI/s, or more than half the total required for LAMPF. As buffered fast acquisition becomes a reality, the needs of EPICS/HRS could double.

Ways of enhancing a PDP-11/70 or a PDP-11/45 to increase its performance are discussed below. Any or all of these ideas may need to be implemented. It should be noted that the speed factors cannot be multiplied together to gain even more speed.

1. Fast Array Processor (FAP). Fast array processors are commercially available, which would be used to speed up certain processes within the data-analysis program. It is estimated that one could expect a factor of 2 increase in speed on a PDP-11/70 for certain applications. The estimated cost of the increase in speed is \$30 to \$50K.

2. FASTBUS Memory. The PDP-11/45 and 11/55 computers have an additional bus besides the UNIBUS. This bus, called the FASTBUS, can be used to add very fast memory to the system. This fast memory can be expected to increase the speed of a PDP-11/45 or a PDP-11/55 by a factor of 2. The cost for this increase is about \$10K/32-K words of memory, or \$40K for a 124-K word system.

3. Cache Memory. If one does not want to add fast memory to the FASTBUS, cache memory may be added to the 11/45 or 11/55 to speed up access to memory on the UNIBUS. Depending on the design of the cache and the program running on the machine, a speedup of about a factor of 1.5 may be expected. The cost for cache memory is a total of \$5K per machine.

4. Fast Floating-Point Processor. Another alternative to speeding up FORTRAN programs is a fast floating-point processor. These may be added to 11/70, 11/60, 11/55, and 11/45 computers. The expected speedup is about a factor of 2, depending on how much floating-point calculation is done in the FORTRAN program. The incremental cost is about \$5K per machine.

B. Midicomputer

The requirement to run large codes required for certain classes of data analysis makes the choice of a medium-sized computer attractive. In fact, this requirement overcomes the disadvantages of the higher operating cost-performance ratio. Since the minicomputers only provide about 2400 kWI/s, an additional 1000 to 1200 kWI/s are required to meet the anticipated data analysis needs of LAMPF.

These additional 1000 kWI/s may be obtained by the use of a class of machine characterized by a DEC System 20 with a price of about \$825K. A configuration that would meet the needs of data analysis is proposed in Table II.

C. Operational Costs

Unfortunately, the costs of doing data analysis do not end with the purchase of the hardware required to do the job. Additional costs incurred are building requirements, facilities, administration, and operational support.

1. Building Requirements. The PDP-11/70 to be delivered in FY 77 can be housed in the Terminal Building although with a large degree of discomfort. It is anticipated that the air conditioning of this building will be inadequate and additional cooling will be required. Also, some work must be done to reduce

TABLE II
COMPUTER CONFIGURATION

	<u>Cost</u>	<u>Maintenance (\$/month)</u>
DEC 20/50-C	503 000	2 773
256-K words memory		
LA 36 console		
RP06 controller plus two drives		
TU45 controller plus one drive		
16 communication lines		
One additional RP06 drive	36 650	215
Two additional TU45 drives	28 000	312
Line printer (1200 l/min)	46 200	294
Card reader (300 c/min)	6 170	75
Terminals		
Four 4010 @ \$5K	20 000	300
Twelve VT52 @ \$1800	21 600	240
Network communications package	35 000	
PDP-11/34 front end	34 640	235
One additional RK05 drive	5 600	39
Large disk system for 11/34	25 000	290
KS90 for 729 7-track drives for 11/34	7 000	55
Three IBM 729 drives	-	225
Plotter, xy-36-in. drum for DEC 20 plus software	31 500	~150
Networking hardware for 11/34	<u>1 400</u>	25
Sub total	801 760	
Software		
RSX-11D license	4 200	
F4P license	3 000	
DEC 20 FORTRAN license	8 250	
DECNET license DEC 20	5 000	
DECNET license 11/34	<u>2 750</u>	
Total cost	824 960	5 228
Lease purchase option on 60-month payback		
2.5% of purchase price/month - \$20 624		

the noise level of this building as it is currently too high for comfortable working conditions. These conditions must be rectified in FY 78. The proposed PDP-11/70 for FY 78 can be housed in the HRS counting room with little additional cost. If an additional PDP-11/70 or 32-bit computer is acquired in FY 78 and FY 79, additional space will be required.

The acquisition of a midcomputer in FY 80 or sooner will also require additional space. It is recommended that this space be located near the experimental area and provide space for all the data analysis computers. Space should also be provided for consultants and experimenters along with permanent magnetic tape storage.

2. Support. The support of the data analysis facility is a large operational cost. This support involves the maintenance of the computers, software support, consulting, and operation of the computer.

The number of additional support personnel is estimated to be two to four, depending on the support level required. An operator/consultant is required for each two PDP-11/70 systems. In addition, the midcomputer would require about three people. This is based on the experience of CTR-6, which runs a large PDP-10 installation. They have one operator, one programmer consultant, and one systems programmer. CTR-6 feels that this is the minimum to run their system efficiently and would like to add an additional support operator/technician.

3. Administration. With a distributed data analysis facility, it is important that close administrative control be exercised. This involves usage control, perhaps through a recharge system, so that most efficient use of the facility may be made by all groups. Also, the direction and supervision of operators, consultants, and programmers required to support the facility is best administered from a single point. Though the location of various computers might be distributed and people assigned to meet the needs of special interests, effective control of the facility can only be accomplished by a single administrator.

Another aspect of administrative control is the assigning of priorities to certain types of computing. It is presently felt that the performance of the midcomputer should be optimized for time-sharing small jobs during the prime shift and reoptimized for production runs in the two after-hours shifts.

VII. CONCLUSION

Although it is hoped that the study has arrived at conclusions that will satisfy the requirements for LAMPF's data analysis problem, one should be aware that changing circumstances may alter those requirements. This report addresses only the problem of data analysis, but resources are also needed for continuing support of the data acquisition systems with upgrading of computers and development of new data acquisition hardware.

This summer study of 1977 recognizes these requirements and recommends that the data analysis facility development not be accomplished at the expense of, but in cooperation with, these other vital developments.

REFERENCES

1. H. L. Anderson, P. R. Bevington, J. A. Buchanan, H. S. Butler, R. E. Chrien, F. A. Kirsten, J. N. Snyder, and R. F. Thomas, "LAMPF Data-Acquisition System," Los Alamos Scientific Laboratory report LA-4504-MS (August 1970).
2. R. F. Thomas, J. F. Amann, H. S. Butler, J. M. Gallup, R. H. Jeppesen, S. S. Johnson, W. K. McFarlane, M. W. McNaughton, R. E. Mischke, D. G. Perry, M. M. Freedom, R. J. Ridge, E. B. Shera, and H. A. Thiessen, "Computing in Support of Experiments at LAMPF," Los Alamos Scientific Laboratory report LA-6543-MS (October 1976).
3. Harry J. Curnow, in Benchmarking, Nicholas Benwell, Ed. (Hemisphere Publishing Corp., Washington, 1975), pp. 99-114.