grading this material, particularly as it relates to the inexperienced user.

In general we are attempting to up-grade the various DEC-originated subsystems to the newest versions to increase the chance of compatibility. We have recently done this with FORTRAN and MACRO and will bring the other programs along as soon as possible. The whole issue of compatibility is one which will receive attention. We will not be able to commit the necessary resources within available manpower to redo the TOPS-10 emulator correctly, but will keep chipping away.

Some requests to look into additional software subsystems have been received and we will consider mounting them if the community develops a definite need. Suggested augmentations include the simulation language SIMULA, the medical information system language MUMPS, BPL or APL, and subroutine libraries such as the biomedical statistics package. We will continue our efforts in the building of user subroutine libraries, concentrating initially on SAIL.

Informal Information Access:

One characteristic of the SUMEX community is the diversity of information, formal and informal, which flows around the system or is available from users. We want to work on ways to capture that information and direct it to other interested individuals. We have two major repositories for information, the <DOC> directory which contains the formal documentation of system facilities and procedures and the <BULLETINS> directory which is an accumulation of informal and dynamic information. We will be working on capabilities both to ease the entry and cataloging of information into these areas and to assist in guiding the user to that subset which is of interest to him at a given time. These user-oriented lookup protocols are, of course, strongly related to the problems of adaptive user interfaces to the system and each will benefit from the experience of the other.

Community Management:

We will continue to work with the management committees to recruit the additional high quality projects which can be accommodated and to evolve resource allocation policies which appropriately reflect assigned priorities and project needs. We hope to make more generally available information about the various projects both inside and outside of the community and thereby to promote the kinds of exchanges exemplified earlier and made possible by network facilities. The first workshop will provide much useful information about the strengths and weaknesses of the performance programs both in terms of criticisms from other AI projects and in terms of the needs off practicing medical people. We plan to use this experience to guide the community building aspects of SUMEX-AIM.

II.B SUMMARY OF RESOURCE USAGE

The following data give an overview of the resource usage from August 1974 (when we began keeping detailed statistics) through March 1975. There are four sub-sections containing data respectively for 1) resource usage by community (AIM, Stanford, and system), 2) resource usage by project, 3) Network usage data, and 4) measures of diurnal variations in system loading.

II.B.1 RELATIVE SYSTEM LOADING BY COMMUNITY

The SUMEX resource is divided, for administrative purposes, into 3 major communities: user projects based at the Stanford Medical School, user projects based outside of Stanford (national AIM projects), and systems development efforts. As defined in the resource management plan approved by BRB at the start of the project, the available resource will be divided between these communities as follows:

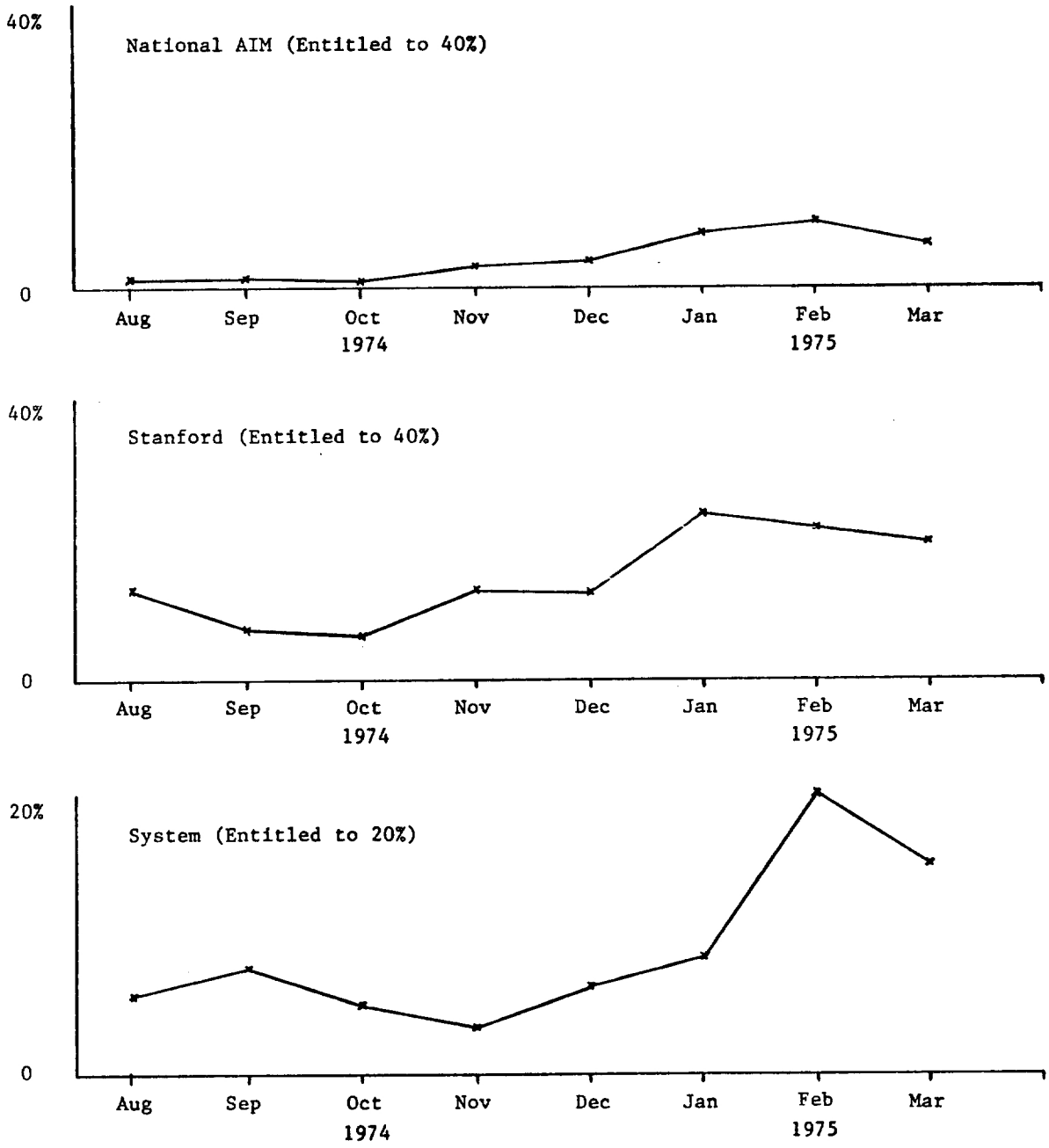
CPU Usage -	Stanford	40%
	AIM	40%
	System	20%
File Space -	Stanford	25,000 pages
	AIM	25,000 pages
	System	30,000 pages

We have recently brought an additional 40,000 pages of file space on-line (this happened since March and hence is not shown in these data) - this additional space will be equally divided, for the most part, between the Stanford and AIM user groups. The system file requirements include all of the subsystem files, documentation files, and other system-related files shared by all users in addition to staff directories. We expect the system allocation to grow somewhat more slowly than user space requirements, reflecting primarily the addition of new subsystems, documentation, and user information from time to time.

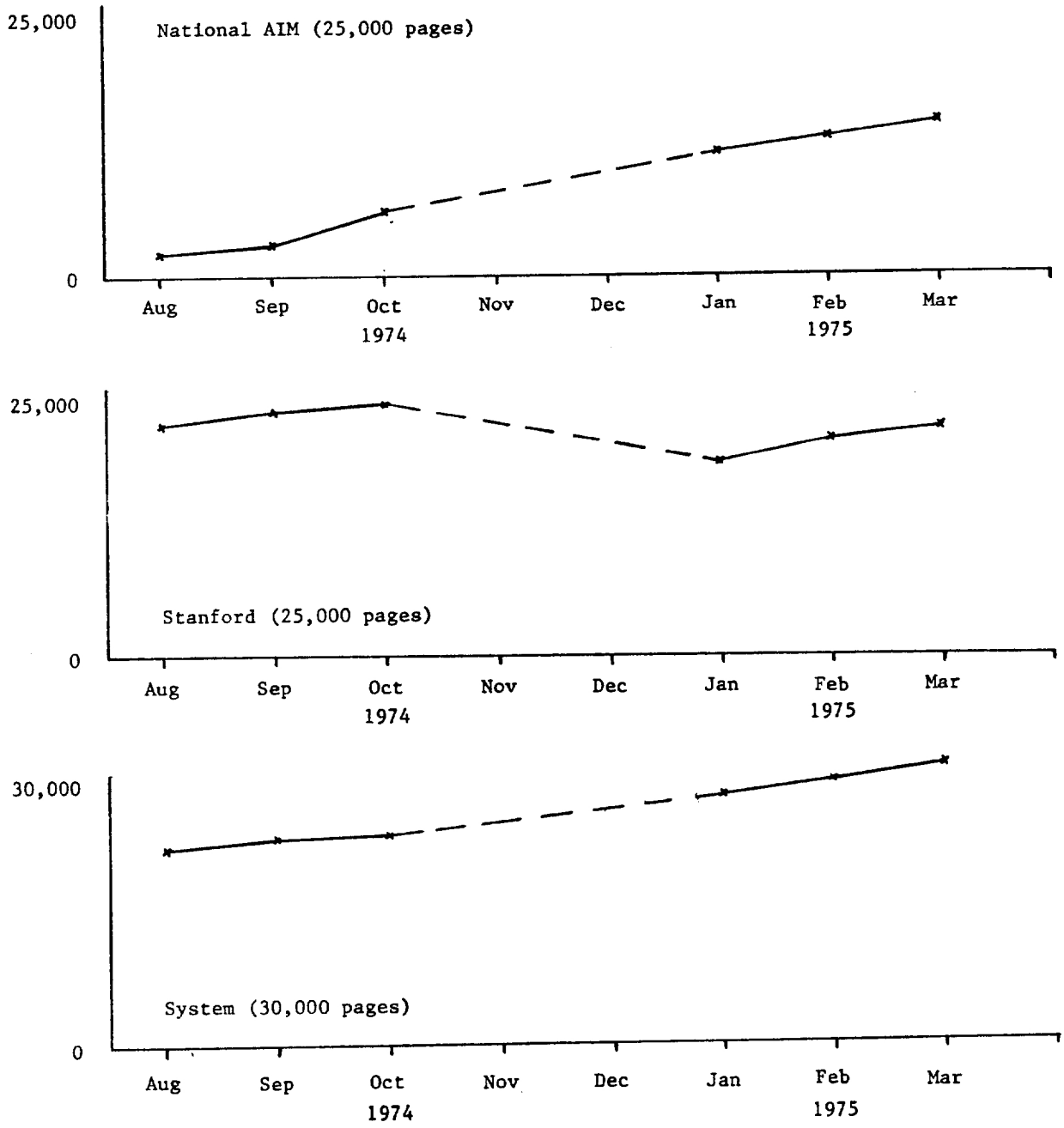
The following plots show monthly usage of CPU and file space resources for each of these three communities relative to their respective aliquots. Note that file space data are not available for November and December 1974 because we were making a transition between numeric and alphameric account designations.

As is evident in the data, many of the national AIM projects joined the SUMEX community during this year (as communication facilities became available) and their use of the facility began slowly. The rate of remote use is expected to increase substantially during the next few months.

CENTRAL PROCESSOR USE BY COMMUNITY



FILE SPACE USE BY COMMUNITY



II.B.2 INDIVIDUAL PROJECT AND COMMUNITY USAGE

The table following shows average resource usage by project in the past grant year. The data displayed include a description of the operational funding sources (outside of SUMEX-supplied computing resources) for currently active projects, average monthly CPU consumption by project (Hours/month), average monthly terminal connect time by project (Hours/month), and average file space in use by project (Pages/month, 1 page = 512 computer words). Averages were computed for each project for the months since August 1974 or since the project was admitted to the resource if after August 1974.

Many of the national AIM projects joined the SUMEX community during this year (as communication facilities became available) and their use of the facility began slowly. For this reason, many of the average usage figures for these projects are lower than is representative of their activity currently. We expect the remote usage to increase significantly in the next months.

RESOURCE USE BY INDIVIDUAL PROJECT

STANFORD COMMUNITY	CPU (Hrs/mo)	CONNECT (Hrs/mo)	FILE SPACE (Pages/mo)
1) DENDRAL PROJECT "Resource Related Research Computers and Chemistry" NIH RR-00612 (3 yr award) $323,933 this year	46.38	728.5	12325
2) MYCIN PROJECT "Computer-based Consult. in Clin. Therapeutics" HEW HSO-1544 (3 yr award) $124,000 this year	10.17	186.1	3264
3) PROTEIN STRUCT MODELING "Heuristic Comp. Applied to Prot. Crystallog." NSF DCR74-23461 (2 yrs.) $88,436 total	3.62	139.0	933
4) PILOT PROJECTS (see reports in Sec IV.B.1)	6.49	208.4	4691
COMMUNITY TOTALS	66.66	1262.0	21213

NATIONAL AIM COMMUNITY			
1) DIALOG PROJECT "Computer Model of Diagnostic Logic" HEW MB-00144 (3 yrs.) $125,027 this year	8.54	159.9	2063
2) Higher Mental Functions "Computer Models in Psychiatry and Psychother." NIH MH-06645 (3 yrs.) $170,000 this year NIH MH-27132 (2 yrs.) $130,000 this year	0.12	11.2	1021
3) MISL PROJECT "Medical Information Systems Laboratory" HEW MB-00114 (2 yrs.) $380,619 this year	0.68	53.1	242

4)	RUTGERS PROJECT "Computers in Biomedicine" NIH RR-00643 (3 yrs.) $285,240 this year	7.26	351.5	2659
5)	AIM Administration	1.93	66.2	2587
	COMMUNITY TOTALS	18.53	641.9	8572
SUMEX SYSTEM				
1)	Development	36.97	1441.5	8450
2)	Operations	6.04	618.9	17272
	COMMUNITY TOTALS	43.01	2060.4	25722
	RESOURCE TOTALS	128.20	3964.3	55507

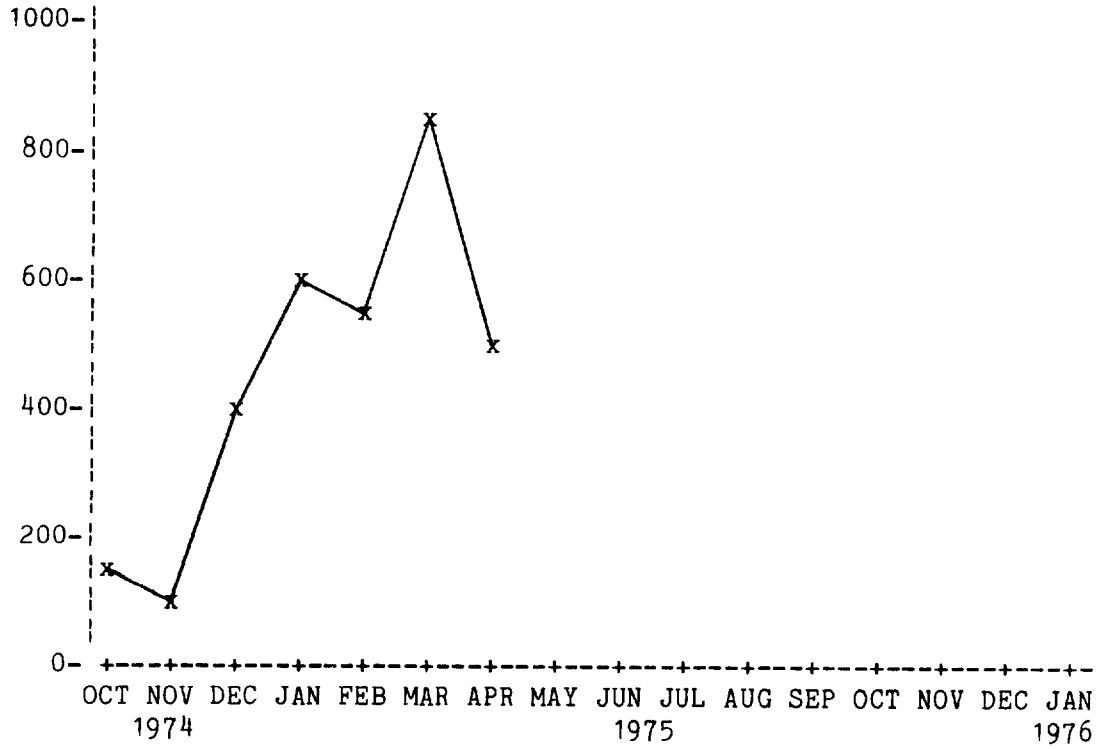
II.B.3 NETWORK USAGE STATISTICS

NETWORK USAGE PLOTS

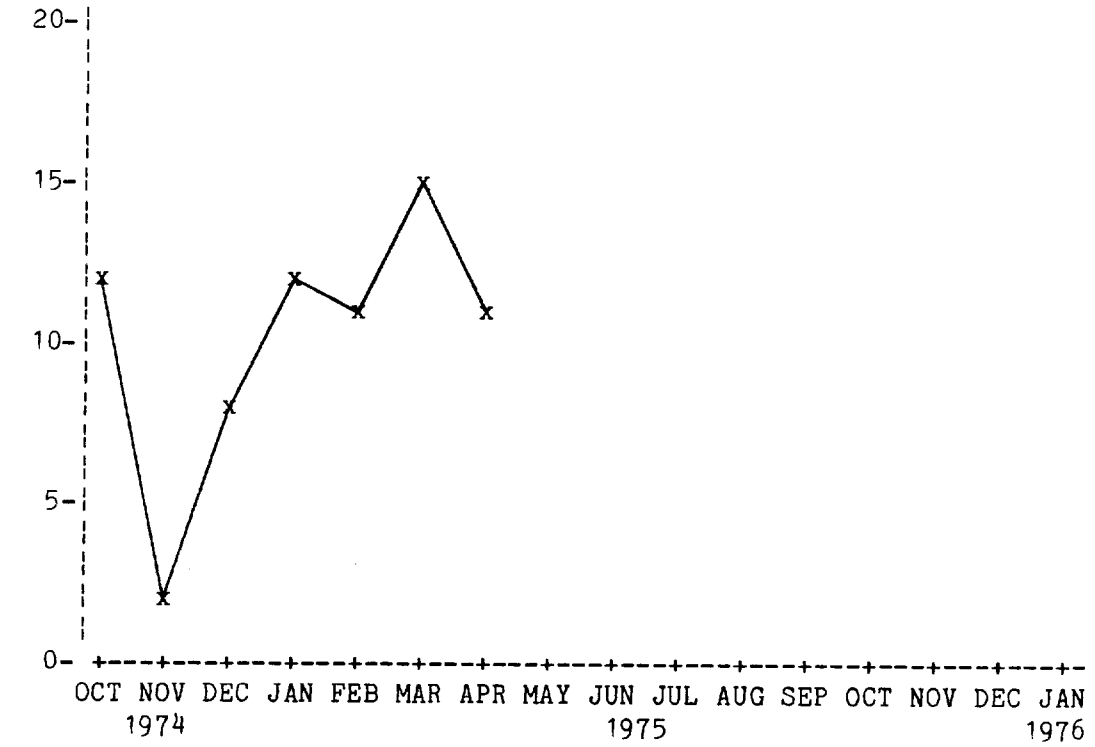
The following are plots of the major billing components for SUMEX-AIM TYMNET usage. These include the total connect time for terminals coming into SUMEX and the total number of characters transmitted over the net. The ratio of characters received at SUMEX to characters sent to the terminal is about 1:14 over the past couple of months. The plot on page 3 is for TOTAL character traffic.

Note that the high usage in October is because we were doing a great deal of testing of the net at that time. As of March, we announced that the IN-WATS lines would be terminated soon and that users should transfer to the TYMNET. Extrapolation of these data should be done very carefully because the apparent steep rise in February and March has apparently been tempered by the ARPANET connection becoming operational in late March. The Rutgers community, at least in part, is using the ARPANET in place of the TYMNET because of more convenient access through their TIP.

TYMNET CONNECT TIME (IN HOURS)



TYMNET CHARACTER TRANSMISSION (MILLIONS OF CHARS)



II.B.4 SYSTEM DIURNAL LOADING PROFILE

Since late February, we have been collecting more detailed information about time variations in system performance and loading. These data include user-oriented measurements such as load average, number of jobs, and percent of LISP use. In addition, we collect data characterizing internal system operating parameters such as paging rate, trap overhead, scheduler overhead, and drum use. Systematic measurements are taken every 20 minutes to give an overall picture of diurnal variations. Periodically we observe system dynamics on a much shorter time scale as well to observe more of the internal workings of the system - such data are so voluminous, however, that they cannot be collected continuously.

The following graphs give a feeling both for the type of data available and the loading characteristics they reveal. The plots are organized 2 per page with the quantity displayed labeled at the top of each graph along with average and extremum values of the data. The first 3 pages display composite data for the month of April 1975 - derived by averaging together all of the weekday data for that month as a function of time of day. The next 3 pages show somewhat different data for a single day recently in May (note the hash marks on the right side of these plots signify that the machine was down [for repair]).

The most striking feature from the monthly data is the expected peak loading during prime time and in fact a bi-modal distribution is apparent in some of the plots approximating the 3 hour shift between prime times on the east coast and west coast. Another striking feature is the dramatic difference between the monthly average data for April and the individual daily data. We feel there are at least two effects taking place to account for this difference in addition to simple day-to-day loading level variations. First, daily data is considerably more "noisy" than the average data with peak loading occurring in cycles spaced about 2-3 hours apart. We observe the phases of the various sub-peaks relative to time of day to be unpredictable other than gross effects leading to the bi-modal average data corresponding to heavy mid-afternoon use of the respective coasts. For example, whereas the peak in the monthly load average curve occurs around 1500 PDT, the peak for the daily plot is around 1600-1700 and the daily plot displays much higher "momentary" (20 minute average) excursions. In fact, over shorter intervals (1-2 minutes) load average peaks are even higher (load averages approaching 10 or more are not uncommon under present loading in the afternoon). Thus in taking the average for a month, the peak loading extremes which strongly influence the subjective "feel" of the system on a daily basis are blurred out and lost.

Secondly, we are apparently on a rather steeply rising curve in terms of system utilization. The composite load average data for March, for example, had a peak value of 2.7 as opposed to 3.7 for April. Of course, two points do not necessarily constitute a trend but this data corresponds with the subjective comments contributed by users in their project reports and other commentaries on system response. There is general agreement that loading has been increasing noticeably.

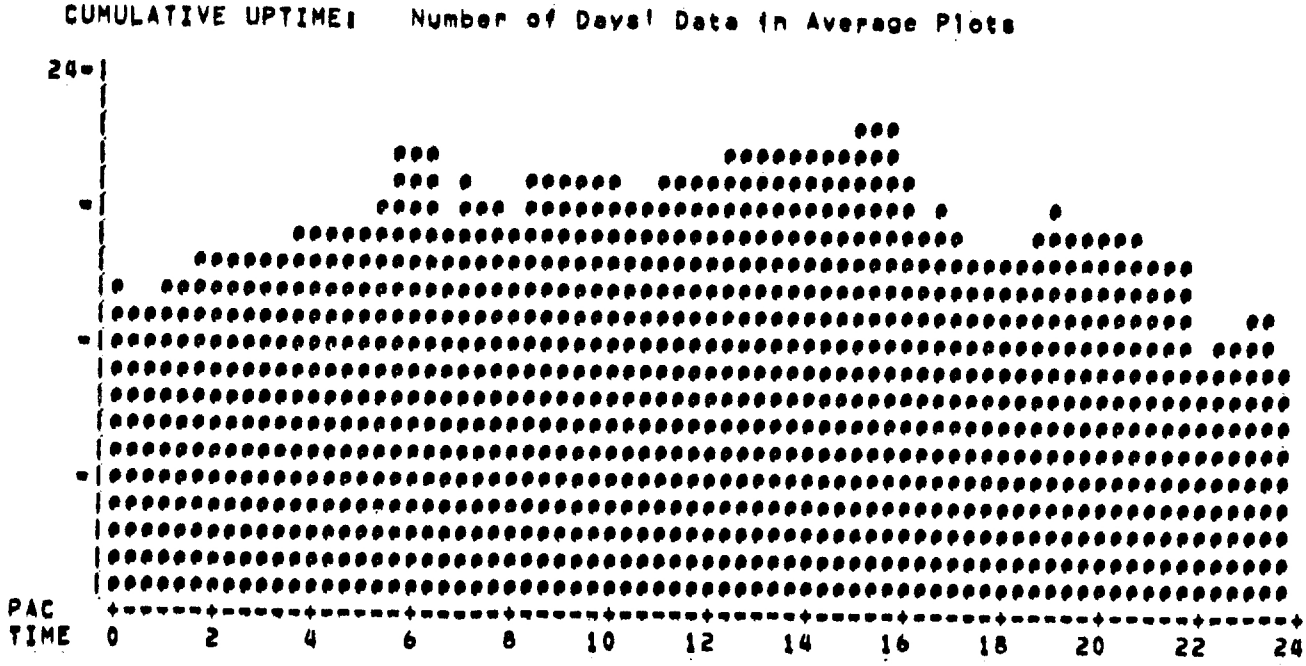
Several points mentioned in the report can be observed in the daily loading data. During peak loading, drum use somewhat exceeds the 3300 page fixed head capacity we now have on-line. We expect to mitigate this overflow through more intelligent management of swapping storage. The total number of jobs follows closely the average data plots but other measures of system use like load average and LISP job consumption tends to skew toward the evening hours. This is evidently a natural leveling process wherein the very large jobs have more difficulty contending during the day and the program developers have shifted their schedules to work later at night. Observe the very heavy (and efficient - 96%) use during the evening hours when the load average has dropped to about 3 or 4 and is arising to a large extent from LISP usage. It should also be noted that load average and overall percent usage are somewhat independent. Usage measures how many cycles are being consumed and load average measures between how many users these cycles are being divided (one CPU-bound job would show a usage of 100% but a load average of only 1.0).

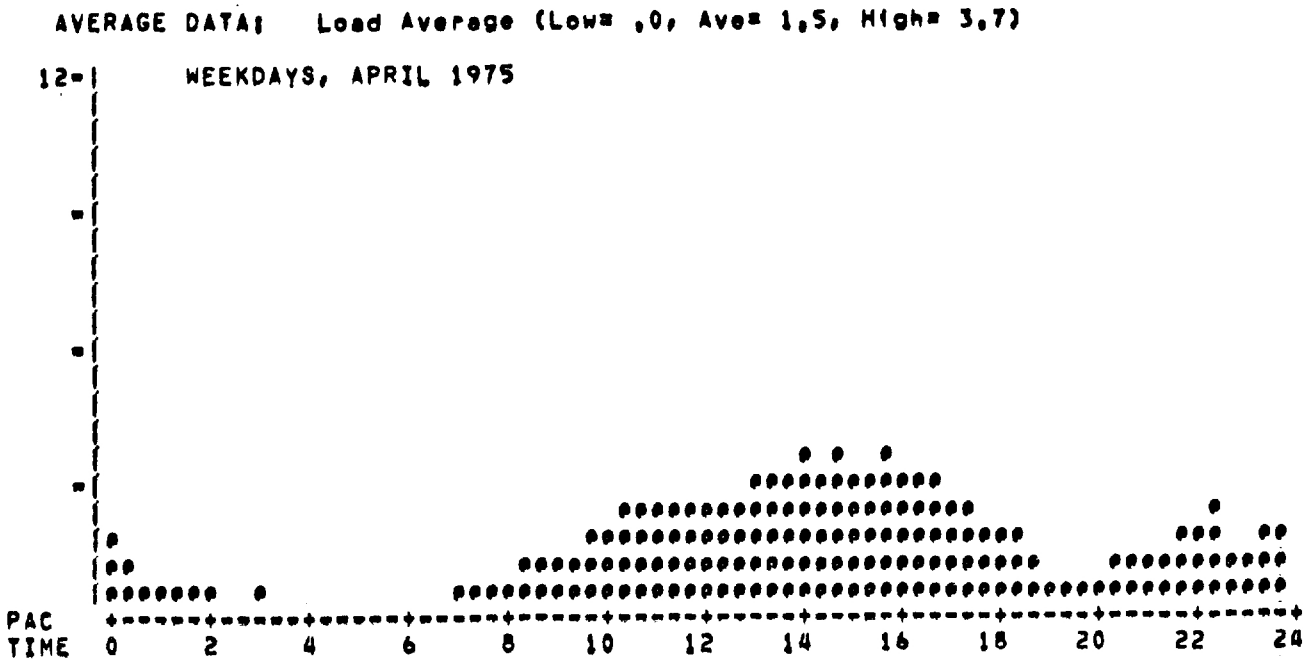
These data, while preliminary in the sense that we have only the trends of a couple of months to go on, are consistent with user comment and are the basis of our preliminary proposal to up-grade the SUMEX-AIM CPU (see page 4 AND Appendix C). While load-leveling (such as moving work on LISP programs to evening hours) may be acceptable (and desirable) for some program development work, it is unrealistic to ask physicians to adapt in a similar manner in trying out the AI programs. Our development and managerial incentives are directed toward making better use of the off-hours, freeing up prime time where possible. But, we must anticipate the need for more computing capacity during prime time, especially based on these measurements and the anticipated growth of the user community. We are continuing active work on the CPU up-grade plan discussed elsewhere in this report (see page 51).

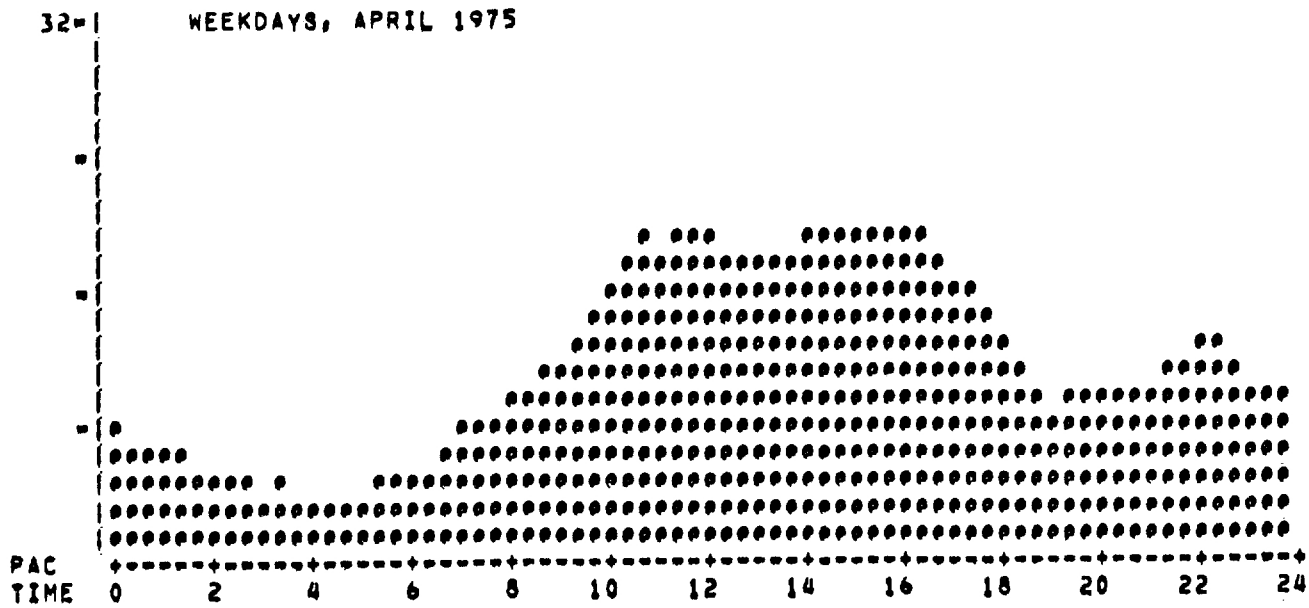
AVERAGE DATA PLOTS - WEEKDAYS, APRIL 1975

Includes dates:

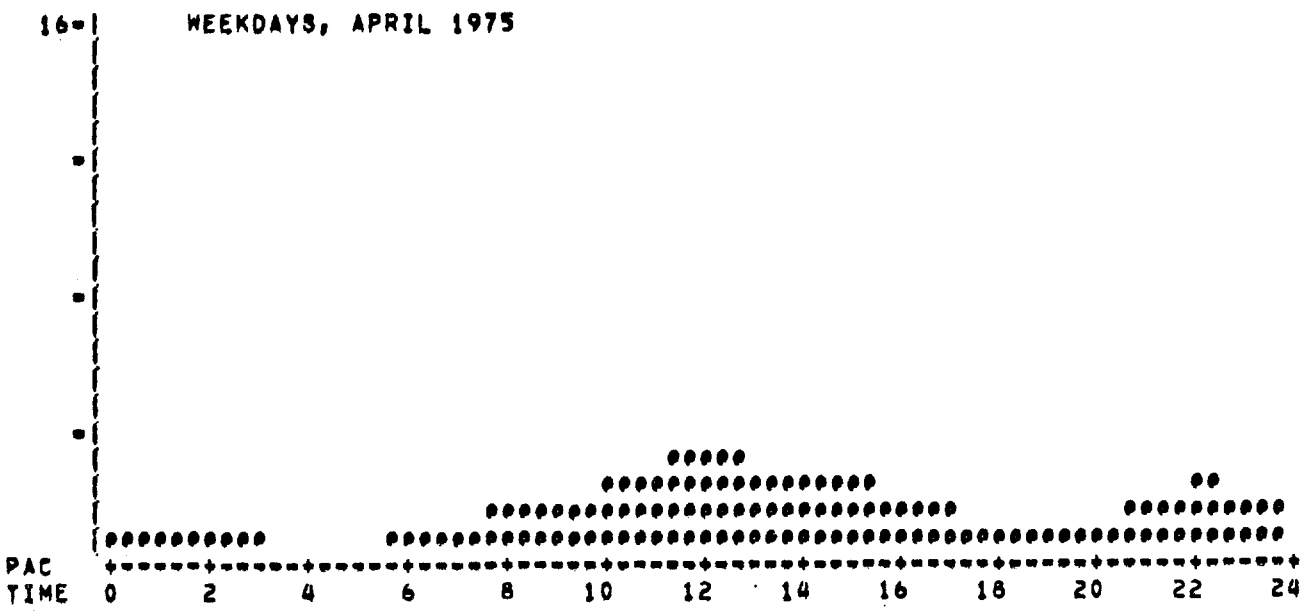
TUESDAY, 1-APR-75
WEDNESDAY, 2-APR-75
THURSDAY, 3-APR-75
FRIDAY, 4-APR-75
MONDAY, 7-APR-75
TUESDAY, 8-APR-75
WEDNESDAY, 9-APR-75
THURSDAY, 10-APR-75
FRIDAY, 11-APR-75
MONDAY, 14-APR-75
TUESDAY, 15-APR-75
WEDNESDAY, 16-APR-75
THURSDAY, 17-APR-75
FRIDAY, 18-APR-75
MONDAY, 21-APR-75
TUESDAY, 22-APR-75
WEDNESDAY, 23-APR-75
THURSDAY, 24-APR-75
FRIDAY, 25-APR-75
MONDAY, 28-APR-75
TUESDAY, 29-APR-75
WEDNESDAY, 30-APR-75



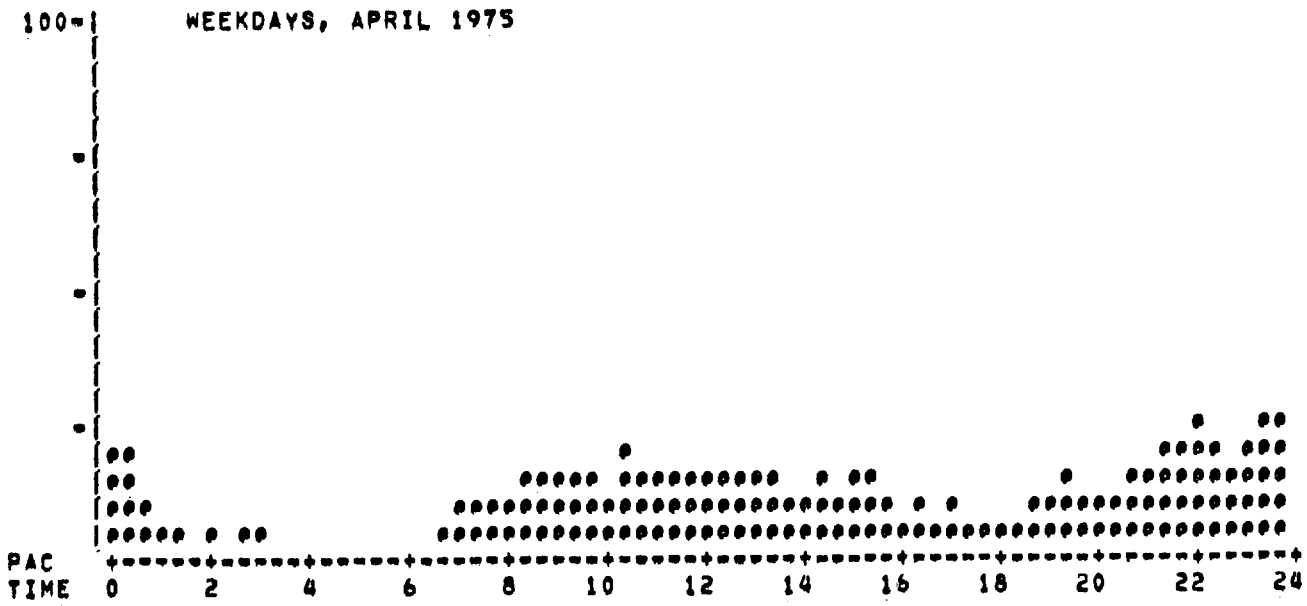

AVERAGE DATA: Load Average (Low= .0, Ave= 1.5, High= 3.7)
12-
WEEKDAYS, APRIL 1975
PAC TIME
0 2 4 6 8 10 12 14 16 18 20 22 24

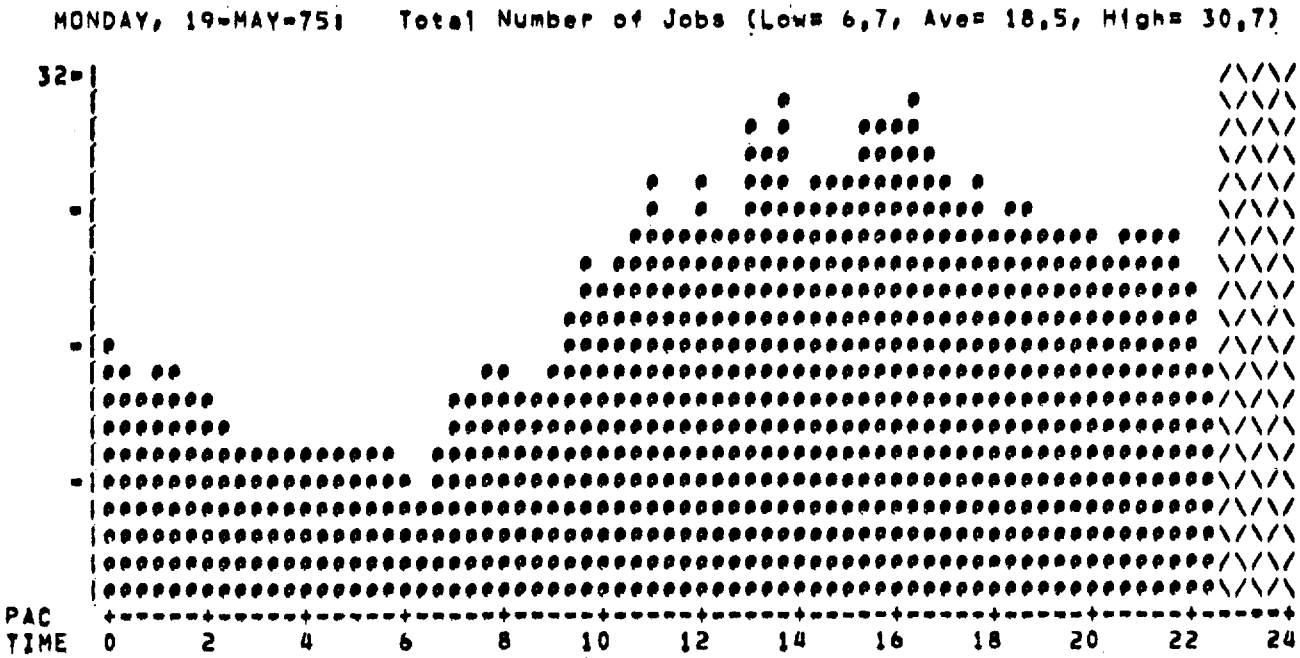

AVERAGE DATA: Total Number of Jobs (Low= .0, Ave= 11.2, High= 19.8)
32-
WEEKDAYS, APRIL 1975
PAC TIME
0 2 4 6 8 10 12 14 16 18 20 22 24

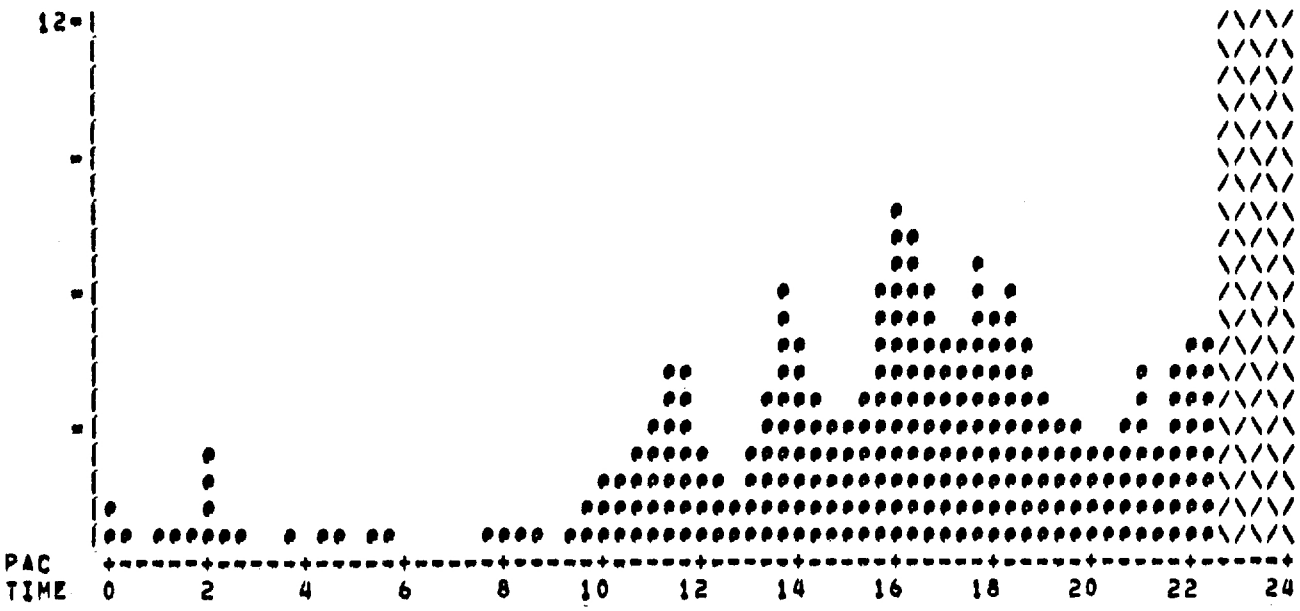
AVERAGE DATA: Number of LISP Jobs (Low= ,0, Ave= 1,4, High= 3,2)

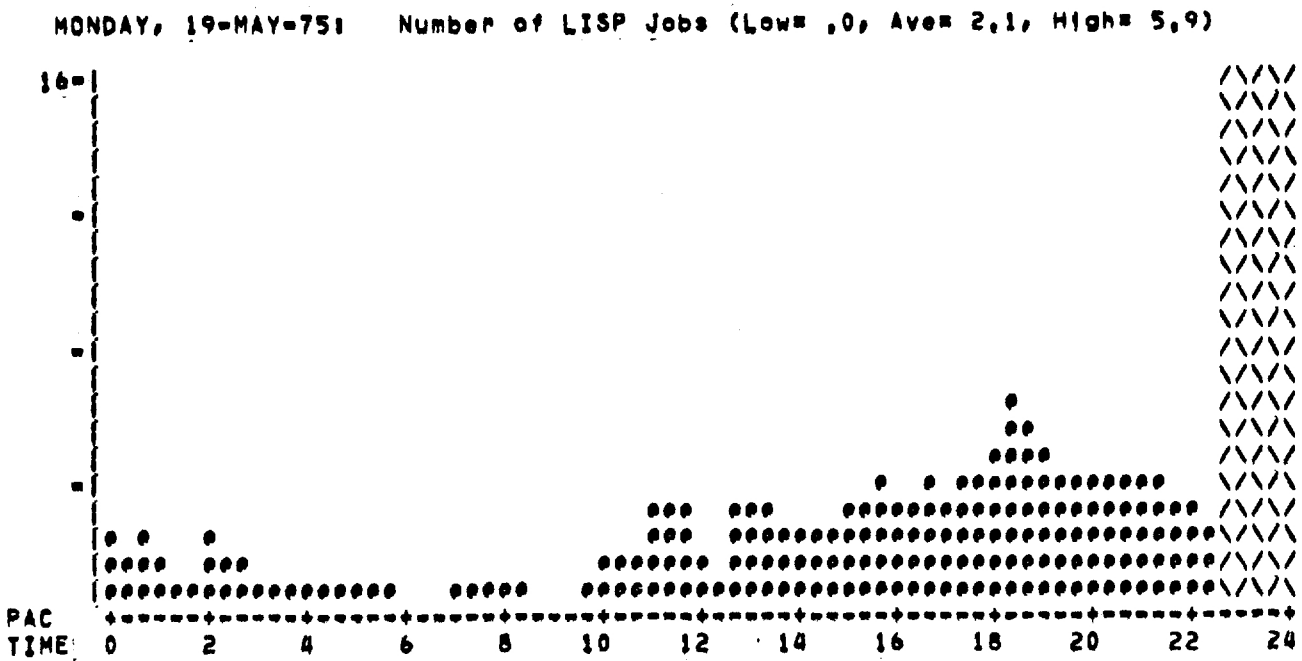


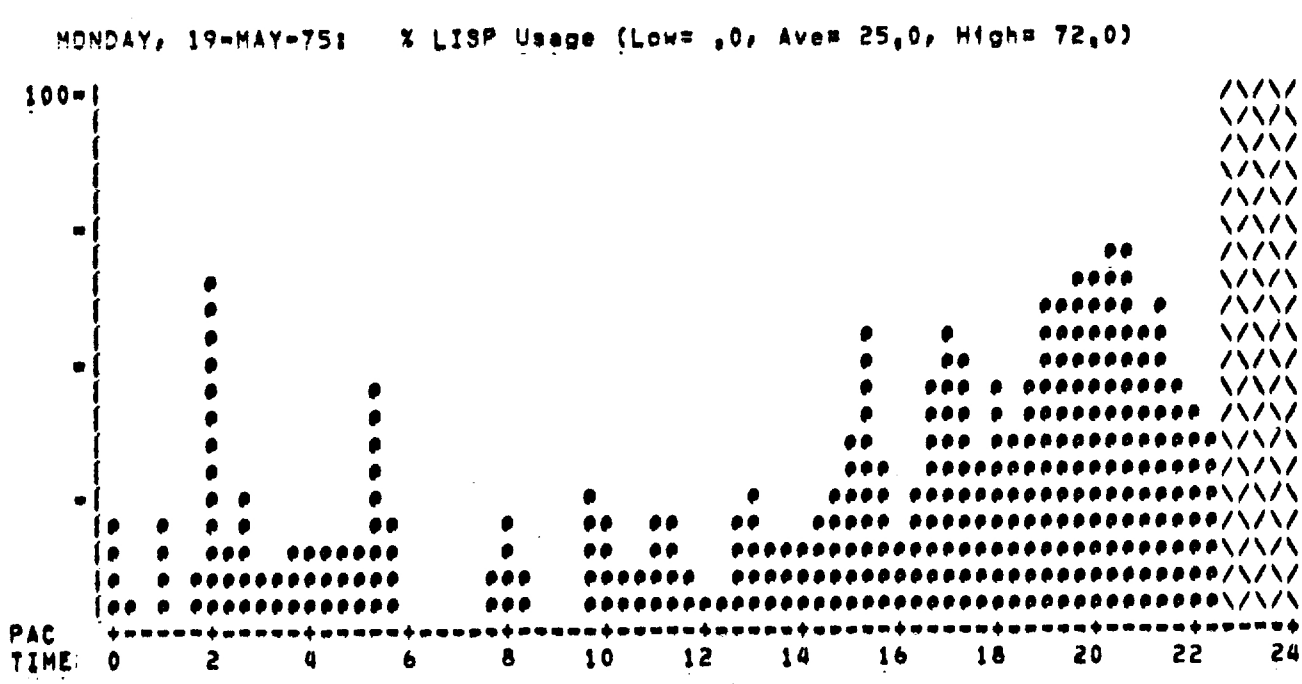
AVERAGE DATA: % LISP Usage (Low= ,0, Ave= 10,5, High= 24,0)

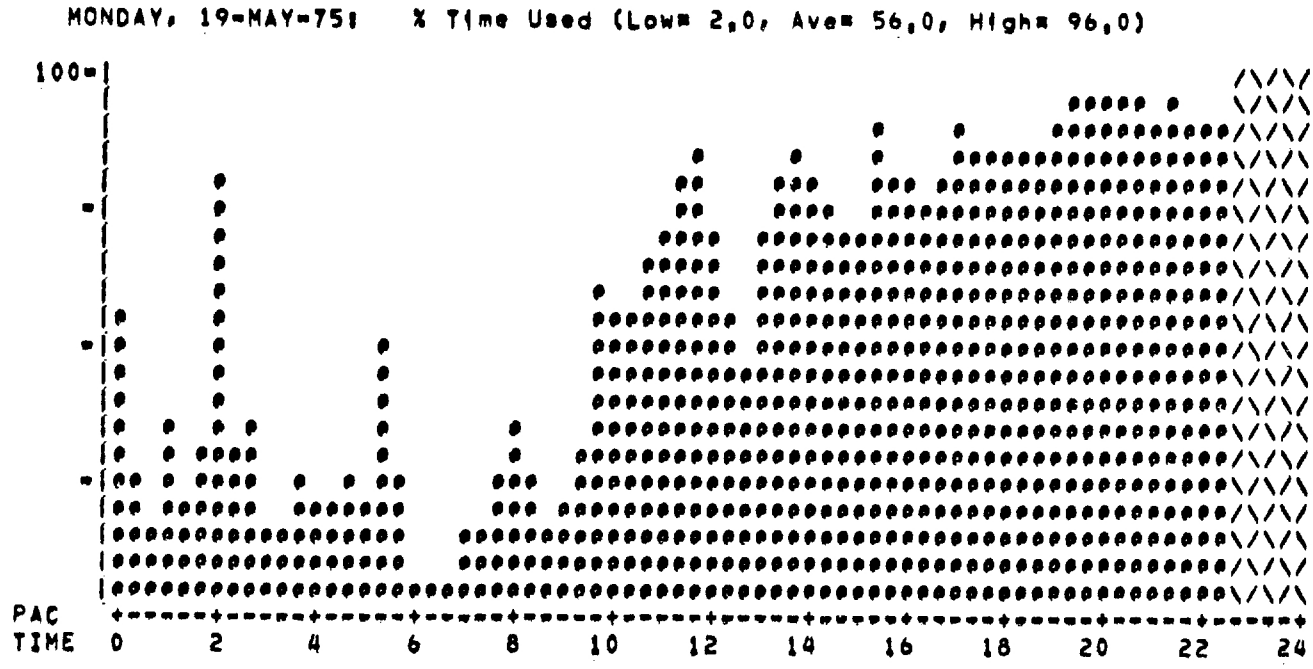


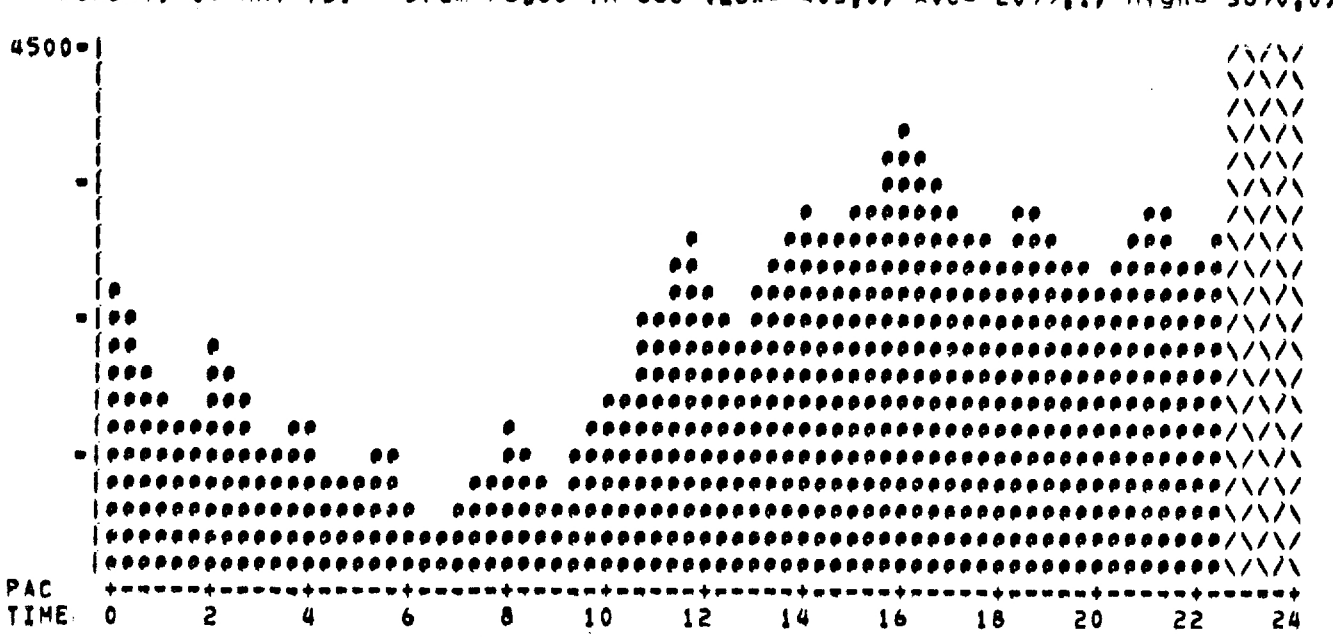

MONDAY, 19-MAY-75: Total Number of Jobs (Low= 6,7, Ave= 18,5, High= 30,7)
32
PAC
TIME
0
2
4
6
8
10
12
14
16
18
20
22
24


MONDAY, 19-MAY-75: Load Average (Low= ,0, Ave= 2,4, High= 7,7)
12
PAC
TIME
0
2
4
6
8
10
12
14
16
18
20
22
24


MONDAY, 19-MAY-75: Number of LISP Jobs (Low= ,0, Ave= 2,1, High= 5,9)
16-
PAC
TIME: 0 2 4 6 8 10 12 14 16 18 20 22 24


MONDAY, 19-MAY-75: % LISP Usage (Low= ,0, Ave= 25,0, High= 72,0)
100-
PAC
TIME: 0 2 4 6 8 10 12 14 16 18 20 22 24


MONDAY, 19-MAY-75: % Time Used (Low= 2,0, Ave= 56,0, High= 96,0)
100=
PAC
TIME 0 2 4 6 8 10 12 14 16 18 20 22 24


MONDAY, 19-MAY-75: Drum Pages In Use (Low= 405,0, Ave= 2099,1, High= 3690,0)
4500=
PAC
TIME 0 2 4 6 8 10 12 14 16 18 20 22 24

II.C RESOURCE EQUIPMENT SUMMARY

The following table gives a list of the items of equipment purchased to date for the SUMEX resource along with details on vendor, description, price, and date.

ITEM	QUANTITY	DESCRIPTION	MANUFACTURER	MODEL NUMBER	DATE INSTALLED	DATE ACCEPTED	PURCHASE PRICE	SOURCE FUNDS
KI-10 CPU	1	Central processor, including console	Digital Equipment Corporation	KI-10	3/1/74	4/24/74	$178,500	NIH
Memory	3	Core memory (64K words including 4 MC-10 memory ports	Digital Equipment Corporation	MF-10G	3/1/74	4/24/74	$224,910	NIH
	1	Core memory (64K words including 4 MC-10 memory ports	Digital Equipment Corporation	MF-10G	11/74	12/74	$ 63,754	NIH
	1	Memory port multiplexer	Digital Equipment Corporation	MX-10	8/74	9/74	$ 4,770	NIH
Clock	1	Programmable clock	Digital Equipment Corporation	DK-10	3/1/74	4/24/74	$ 2,678	NIH
Disk System	1	Single double density disk controller	Digital Equipment Corporation	RP-10C	3/1/74	5/1/74		
	1	Memory data channel	Digital Equipment Corporation	DF-10	3/1/74	4/24/74		
	4	Double density disk drives and disk packs	Digital Equipment Corporation	RP-03	3/1/74	4/24/74	$108,153	NIH
	3	Double density disk drives and disk packs	Digital Equipment Corporation	RP-03R	2/75	3/75	$ 44,636	NIH

ITEM	QUANTITY	DESCRIPTION	MANUFACTURER	MODEL NUMBER	DATE INSTALLED	DATE ACCEPTED	PURCHASE PRICE	SOURCE FUNDS
Swapping Storage	2	Fixed head disk with 1.7M word capacity and 4 track parallel access	Digital Development Corporation	A-7312-D-8	1/75	3/75	$ 37,206	NIH
	1	Special systems controller for DDC disks	Digital Equipment Corporation	RES-10	10/74	11/74	$ 81,090	NIH
DEC Tapes (TU-56)	1	DEC tape control	Digital Equipment Corporation	TD-10	3/1/74	4/24/74	↓	↓
	1	Dual DEC tape drive	Digital Equipment Corporation	TU-56	3/1/74	4/24/74	$ 17,850	NIH
Magnetic Tapes (2 x TU-30)	1	Magnetic tape controller	Digital Equipment Corporation	TM-10A	3/1/74	4/24/74	↓	↓
	2	Tape transports	Digital Equipment Corporation	TU-30	3/1/74	4/24/74	$ 31,238	NIH
Line Printer	1	Special systems line printer control for Data Products 2410	Digital Equipment Corporation	Special	6/74	7/74	$ 7,208	NIH
	1	Line printer with 96 character drum, vertical format control, parity check	Data Products	2410	6/74	7/74	$ 18,963	NIH

ITEM	QUANTITY	DESCRIPTION	MANUFACTURER	MODEL NUMBER	DATE INSTALLED	DATE ACCEPTED	PURCHASE PRICE	SOURCE FUNDS
GT-40	1	Graphics terminal	Digital Equipment Corporation	GT-40	3/1/74	4/24/74	$ 11,156	NIH
Line Scanner	1	Data line scanner	Digital Equipment Corporation	DC-10A	3/1/74	4/24/74	↓	↓
	1	8-line unit	Digital Equipment Corporation	DC-10B	3/1/74	4/24/74	$ 16,275	NIH
TYMNET Interface	1	PDP-10 TYMNET communications controller	TYMSHARE		8/74	10/74	$ 50,774	NIH
ARPANET Interface	1	BB&N ARPANET/KI-10 interface	Bolt, Beranek & Newman		1/75	2/75	$ 21,200	NIH
PDP-11/10	1	Communications processor	Digital Equipment Corporation	PDP-11/10	2/75	3/75	$ 13,445	NIH

ITEM	QUANTITY	DESCRIPTION	MANUFACTURER	MODEL NUMBER	DATE INSTALLED	DATE ACCEPTED	PURCHASE PRICE	SOURCE FUNDS
Terminals	1	Terminal	Data Terminals Communications	DTC-300	3/18/74	6/74	$ 4,597	NIH
	2	Terminals - Execuport portable with carry case	Computer Transceiver Systems, Inc.	311-3	3/18/74	6/74	$ 6,402	NIH
	6	Terminals - elite CRT with edit capabilities	Datamedia	2500	9-10/74	11/74	$ 13,250	NIH
Keyboards	3	Keyboards, special, for leased Datamedia elite 2500 CRT terminals at - NIH Johns Hopkins Univ. Washington Univ.	Datamedia	70DVK7019	11/74	12/74	$ 1,156	NIH

ITEM	QUANTITY	DESCRIPTION	MANUFACTURER	MODEL NUMBER	DATE INSTALLED	DATE ACCEPTED	PURCHASE PRICE	SOURCE FUNDS
Modems	16	Auto answer modems	Prentice Electronics	P-113B	5/6/74	5/6/74		
	5	Auto answer modems	Prentice Electronics	P-1200/150	5/6/74	5/6/74		
	5	Originate modems	Prentice Electronics	P-1200/150	5/6/74	5/6/74		
	4	Modem enclosure with loopback switch and cables	Prentice Electronics	P-100	5/6/74	5/6/74		
	4	Modem enclosures for 8 modems with cables, power supply, digital loopback, line loopback, indicator lights	Prentice Electronics	P-850	5/6/74	5/6/74		
	2	Acoustic coupler modems	Prentice Electronics	DC-22	3/74	3/74		
	3	Modem enclosure with live loopback switch to house P-103F modems	Prentice Electronics	P-100	3/74	3/74	$ 9,415	NIH
Oscilloscope	1	Oscilloscope	Tektronix, Inc.	475DM43	1/75		$ 3,476	NIH

II.D PUBLICATIONS

Publications for the SUMEX staff have included papers describing the SUMEX-AIM resource coincident with its dedication last November (SIGART Newsletter, Sept. 1974; and ACM SIGBIO meeting, San Diego, November 1974 (oral presentation)). In addition, a substantial effort has gone into developing, upgrading, and extending documentation about the SUMEX-AIM resource, the SUMEX-TENEX system, and the many subsystems available to users. These efforts include a number of major documents (such as SOS, PUB, and TENEX-SAIL manuals) as well as a much larger number of document upgrades, user information and introductory notes, and policy guidelines (see Appendix E, Appendix H, and Appendix I). Publications for individual user projects are summarized in the respective reports and Appendix F, page 167.

III RESOURCE FINANCES

III.A REFERENCE TO BUDGETARY DETAILS

The budgetary materials for the SUMEX project covering past actual costs, current performance, and estimates for the next grant year are submitted in separate document to the NIH. Only that section referenced earlier and describing preliminary plans for SUMEX-AIM CPU up-grading are included here