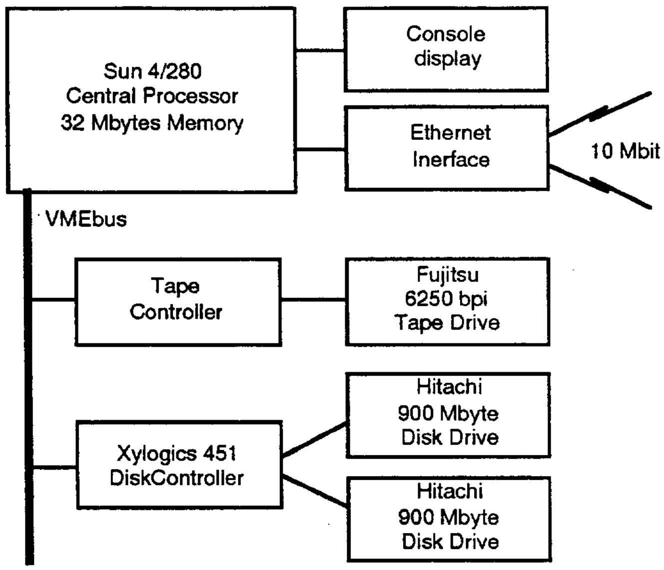

Figure 3. SUMEX-AIM SUN-4 Configuration

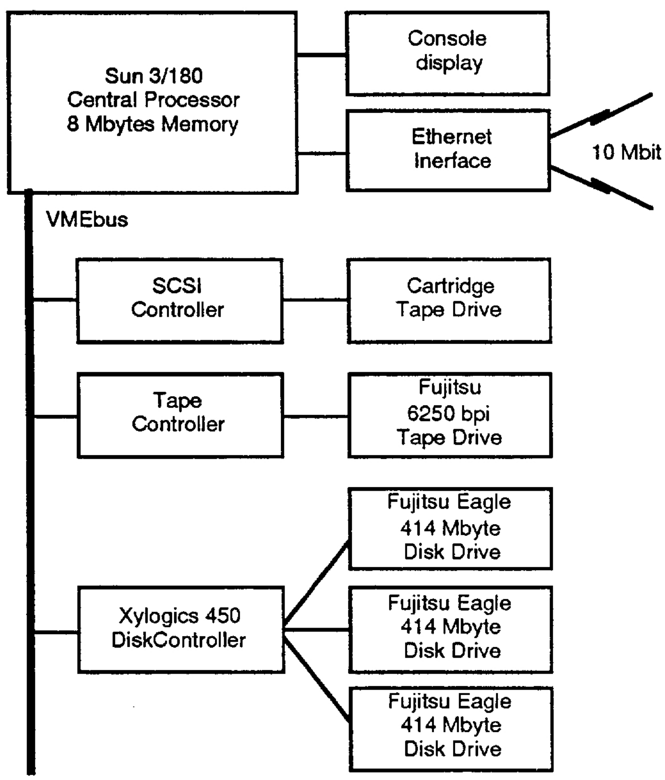

Figure 4. SUMEX-AIM SUN-3 File Server Configuration

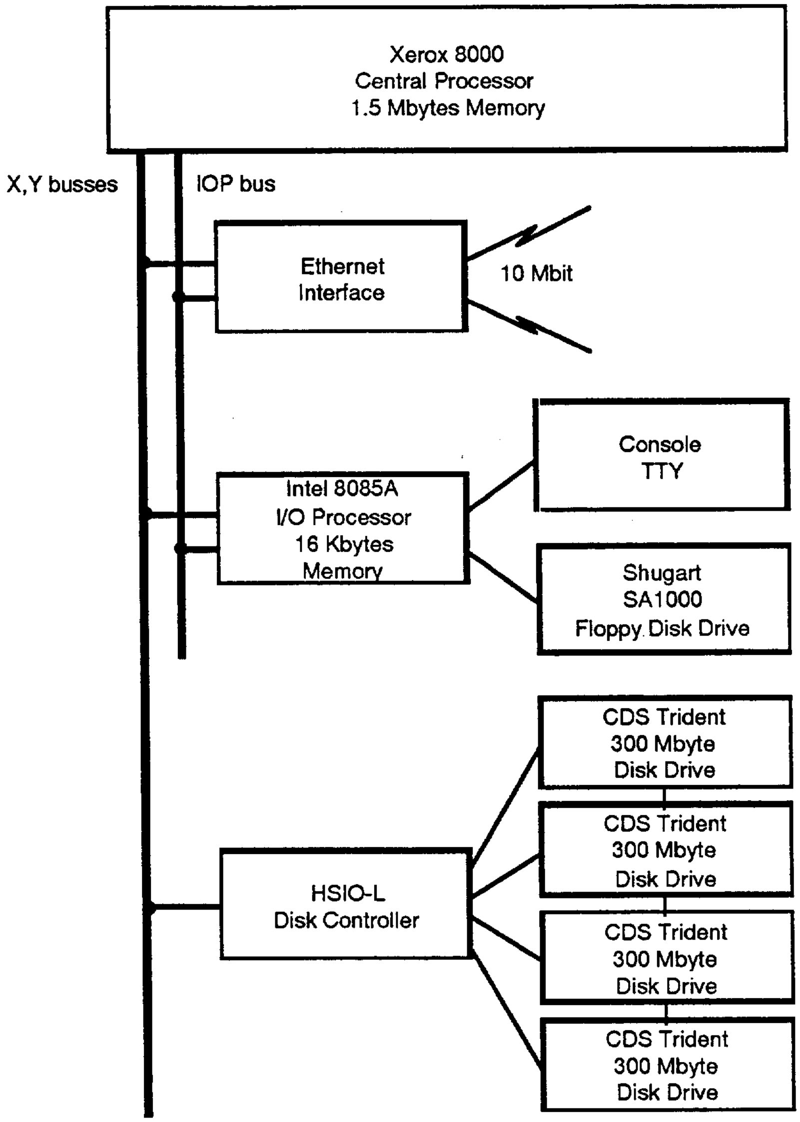

Figure 5. SUMEX-AIM Xerox File Server Configuration

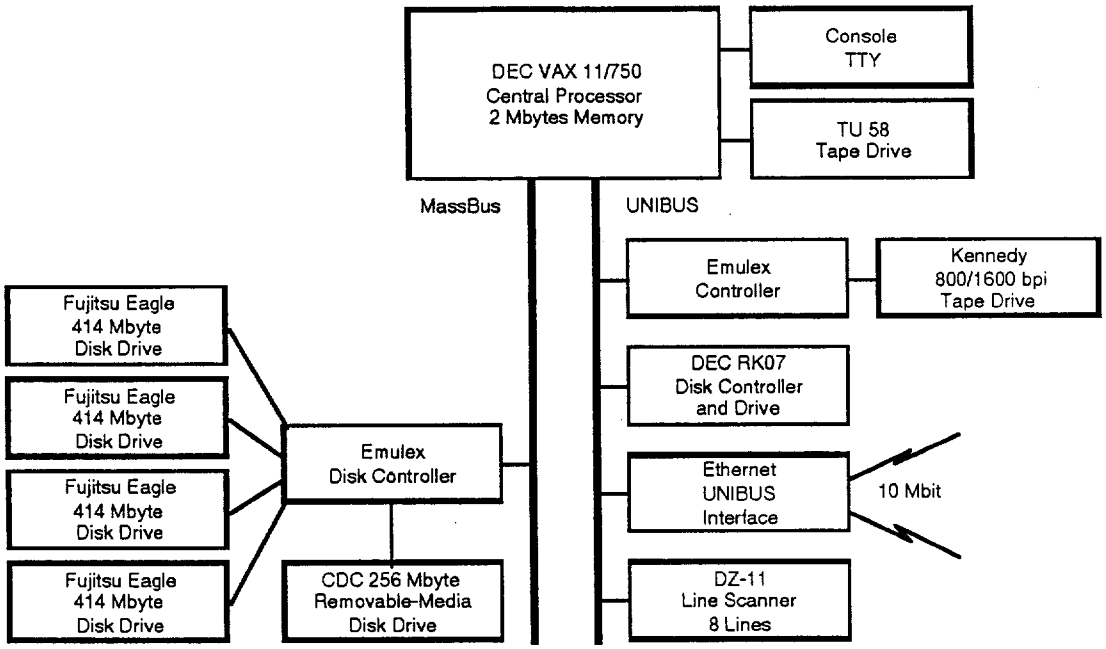

Figure 6. SUMEX-AIM VAX File Server Configuration

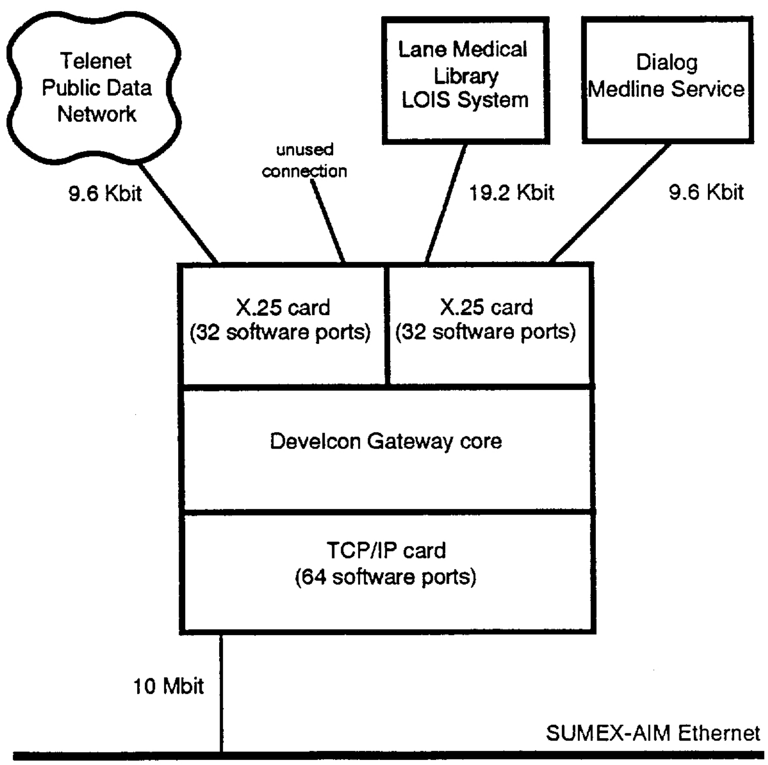

Figure 7. SUMEX-AIM Develcon X.25/TCP-IP Gateway Configuration

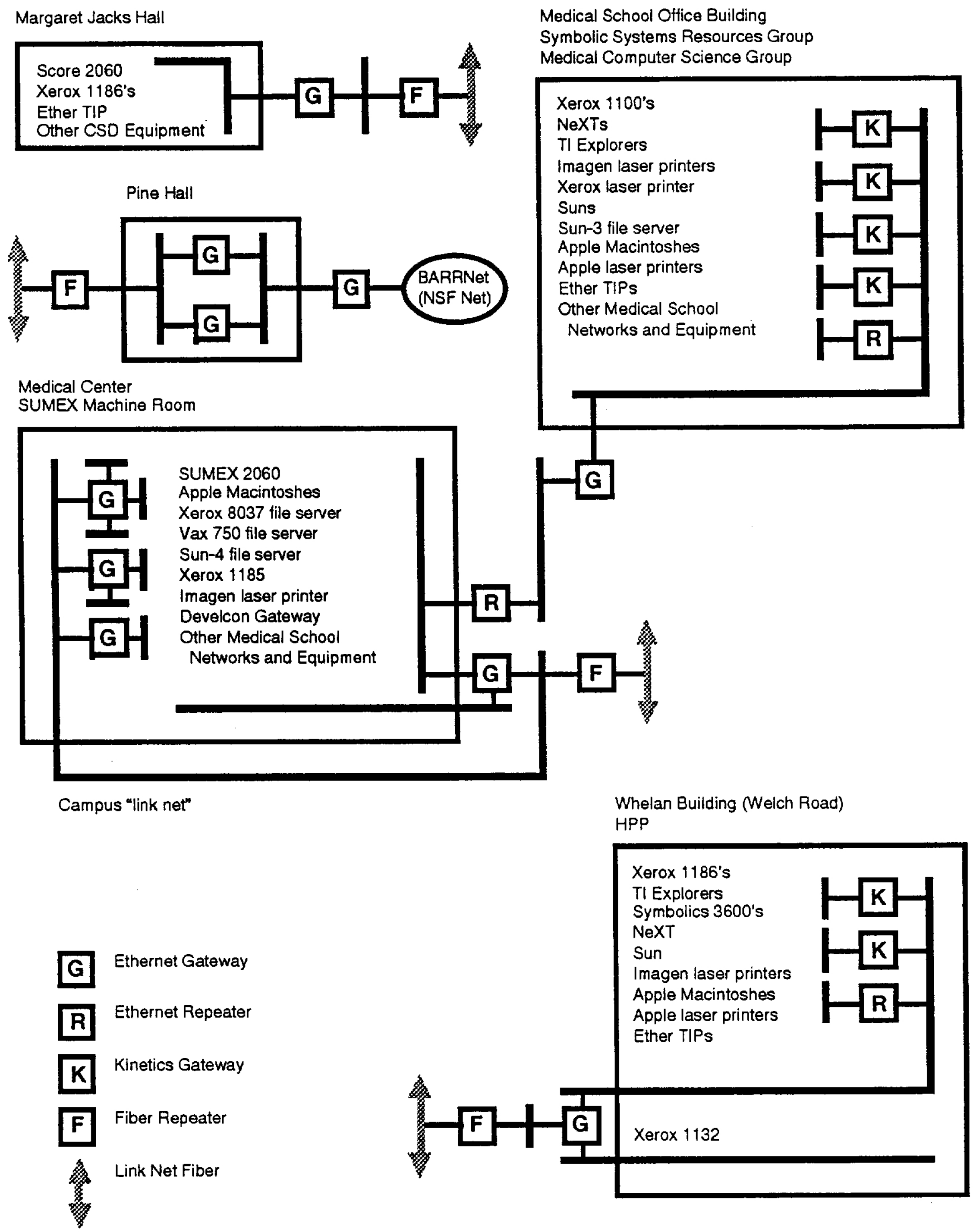

Figure 8. SUMEX-AIM Ethernet Configuration

III.A.2.7. Training Activities

The SUMEX resource exists to facilitate biomedical artificial intelligence applications. This user orientation on the part of the facility and staff has been a unique feature of our resource and is responsible in large part for our success in community building. The resource staff has spent significant effort in assisting users to gain access to the SUMEX-AIM resources at Stanford and use it effectively as well as in assisting AIM projects in designing their own local computing resources based on SUMEX experience. We have also spent substantial effort to develop, maintain, and facilitate access to documentation and interactive help facilities. The HELP and Bulletin Board subsystems have been important in this effort to help users get familiar with the computing environment.

We have regularly accepted a number of scientific visitors for periods of several months to a year, to work with us to learn the techniques of expert system definition and building and to collaborate with us on specific projects. Our ability to accommodate such visitors is severely limited by space, computing, and manpower resources to support them within the demands of our on-going research.

Finally, the training of graduate students is an essential part of the research and educational activities of the KSL. Based largely on the SUMEX-AIM community environment, we have had two unique, special academic degree programs at Stanford, the Medical Information Science program and the Masters of Science in AI, to increase the number of students we produce for teaching, research, and industry. A number of students are pursuing interdisciplinary programs and come from the Departments of Engineering, Mathematics, Education, and Medicine.

The *Medical Information Sciences* (MIS) program continues to be one of the most obvious signs of the local academic impact of the SUMEX-AIM resource[1]. The MIS program received University approval (in October 1982) as an innovative training program that offers MS and PhD degrees to individuals with a career commitment to applying computers and decision sciences in the field of medicine. In Spring 1987, a University-appointed review group unanimously recommended that the degree program be continued for another five years. The MIS training program is based in the School of Medicine, directed by Dr. Shortliffe, co-directed by Dr. Fagan, and overseen by a group of six University faculty that includes two faculty from the Knowledge Systems Laboratory. The specialized curriculum offered by the new program is intended to overcome the limitations of previous training options. It focuses on the development of a new generation of researchers with a commitment to developing new knowledge about optimal methods for developing practical computer-based solutions to biomedical needs.

---

1 Shortliffe, E. H., and Fagan, L. M. Research Training in Medical Informatics: The Stanford Experience. In Proceedings of *The International Symposium on Medical Informatics and Education,* Victoria, B.C., 1989.

The program accepted its first class of four trainees in the summer of 1983 and has now reached its steady-state size of approximately twenty-four graduate students. The program encourages applications from any of the following:

- medical students who wish to combine MD training with formal degree work and research experience in MIS;
- physicians who wish to obtain formal MIS training after their MD or their residency, perhaps in conjunction with a clinical fellowship at Stanford Medical Center;
- recent BA or BS graduates who have decided on a career applying computer science in the medical world;
- current Stanford undergraduates who wish to extend their Stanford training an extra year in order to obtain a "co-terminus" MS in the MIS program;
- recent PhD graduates who wish post-doctoral training, perhaps with the formal MS credential, to complement their primary field of training.

In addition, a special one-year MS program is available for established academic medical researchers who may wish to augment their computing and statistical skills during a sabbatical break. As of spring of 1989, 46% of our enrolled trainees have previously received MD degrees and another 25% are medical students enrolled in joint degree programs. About 29% are candidates for the MS degree, while the remaining 71% are doctoral students. The program has 12 graduates to date, and another 4 students are expecting to complete degrees in June of 1989.

Except for the special one-year MS mentioned above, all students spend a minimum of two years at Stanford (four years for PhD students) and are expected to undertake significant research projects for either degree. Research opportunities abound, however, and they of course include the several Stanford AIM projects as well as research in psychological and formal statistical approaches to medical decision making, applied instrumentation, large medical databases, molecular biology, and a variety of other applications projects at the medical center and on the main campus. Several students are already contributing in major ways to the AIM projects and core research described elsewhere in this annual report.

We are pleased that the program already has an excellent reputation and is attracting superb candidates for training positions. The program's visibility and reputation is due to a number of factors:

- high quality students, many of whom publish their work in conference proceedings and refereed journals even before receiving their degrees; Stanford MIS students have won first prize in the student paper competition at the Symposium on Computer Applications in Medical Care (SCAMC) in 1985 and 1986, and have also received awards for their work at annual meetings of organizations such as the Society for Medical

Decision Making, the American Association for Medical Systems and Informatics (AAMSI), and the American Association for Artificial Intelligence (AAAI);

- a rigorous curriculum that includes newly-developed course offerings that are available to the University's medical students, undergraduates, and computer science students as well as to the program's trainees;
- excellent computing facilities combined with ample and diverse opportunities for medical computer science and medical decision science research;
- the program's great potential for a beneficial impact upon health care delivery in the highly technologic but cost-sensitive era that lies ahead.

The program has been successful in raising financial and equipment support from industry and foundations. It is also recipient of a training grant from the National Library of Medicine. The latter grant was recently renewed for another five years with a study section review that praised both the training and the positive contribution of the SUMEX-AIM environment.

III.A.2.8. Resource Operations and Usage

(1) Operations and Support

The diverse computing environment that SUMEX-AIM provides requires a significant effort at operations and support to keep the resource responsive to community project needs. This includes the planning and management of physical facilities such as machine rooms and communications, system operations routine to backup and retrieve user files in a timely manner, and user support for communications, systems, and software advice. Maintaining the quality of these services has become increasingly difficult in the face of recent budget cuts and attendant staff cuts.

We spend significant time on new product review and evaluation such as Lisp workstations, terminals, communications equipment, network equipment, microprocessor systems, mainframe developments, and peripheral equipment. We also pay close attention to available video production and projection equipment, which has proved so useful in our dissemination efforts involving video tapes of our work.

We continue to operate the primary elements of our server equipment in a generally unattended manner. Operations costs are kept to a minimum by utilizing a student staff for routine tasks. Senior members of this staff provide improvements to the operations procedures in addition to training and supervising new students. This has provided SUMEX with a cost effective operations scheme, contributed to the education of the students, and assisted students in meeting their obligations in undergraduate financial aid programs.

While most of our equipment is concentrated in three computer equipment rooms, our move towards distributed computing has resulted in a substantial amount of equipment being installed in offices and student carrels. This physical distribution of the environment means that maintenance tasks (hardware and software) are more time-consuming because of the need to attend to the systems in remote locations.

(2) Resource Usage Details

The following data give the most cursory overview of various aspects of the SUMEX-AIM resource usage, based on the DEC 2060 accounting system in operation for part of the year — from May 1988 until the transition to the SUN-4 system in October and November of 1988. Measuring subsequent usage in a distributed environment is a much more difficult and ambiguous undertaking. It does not make sense, for example, to try to tally CPU or disk usage on individual workstations, anymore than we would try to measure a researcher's use of his pocket calculator. Similarly, much of the usage of the central servers is of a transient nature, copying a file here or there, archiving or retrieving a file, printing a document through the system spooler, reading mail and transmitting replies, etc. In the client/server model of our work environment, most often a user is not even logged in to a server in order to perform his tasks or if he is, most of the work is going on on his workstation with the central server providing background supporting services. Simply stating the server usage statistics (e.g., in CPU time consumed, connect time, or disk space used) does not begin to approach the measurement of effective computing resource usage in the pursuit of our research goals. To try to do so meaningfully would require a very large "instrumentation" effort that is not even contemplated in distributed workstation systems as they are distributed by vendors or third party system developers, that is not within the Council-approved research goals of the resource, and that would require an additional effort well beyond the resources that are available, with dubious return for the investment.

For example, since we moved our TELENET service to operate through the Develcon X.25/Ethernet gateway, connections from other parts of the country look *exactly* like connections from local workstations since they come into our servers over common-service network connection ports and are assigned randomly-selected TCP sockets for subsequent services. Since the Develcon gateway developer does not provide any accounting tools and the gateway itself does not know *who* the person is who is trying to establish a connection, we have no way to break down network usage between local and remote users, much less identify who is using how much of the various core services.

Thus, one of the consequences of our move to a distributed computing environment is that we will not be able to provide future meaningful data on resource usage as is possible with other central resources. For this year's report, we close out the record for the DEC 2060 usage with the following data:

- Overall resource loading data.
- Individual project and community usage.

(2.1) Overall Resource Loading Data

The following plot (see Figure 9) shows resource CPU usage over the entire history of the SUMEX-AIM project. This includes data from both the KI-TENEX system and the current DECsystem 2060. At the point where the SUMEX-AIM community switched over to the 2060 (February, 1983), you will notice a sharp change in the graphs. This is due to differences in scheduling, accounting, and processor speed calculations between the systems which we have not attempted to normalize.

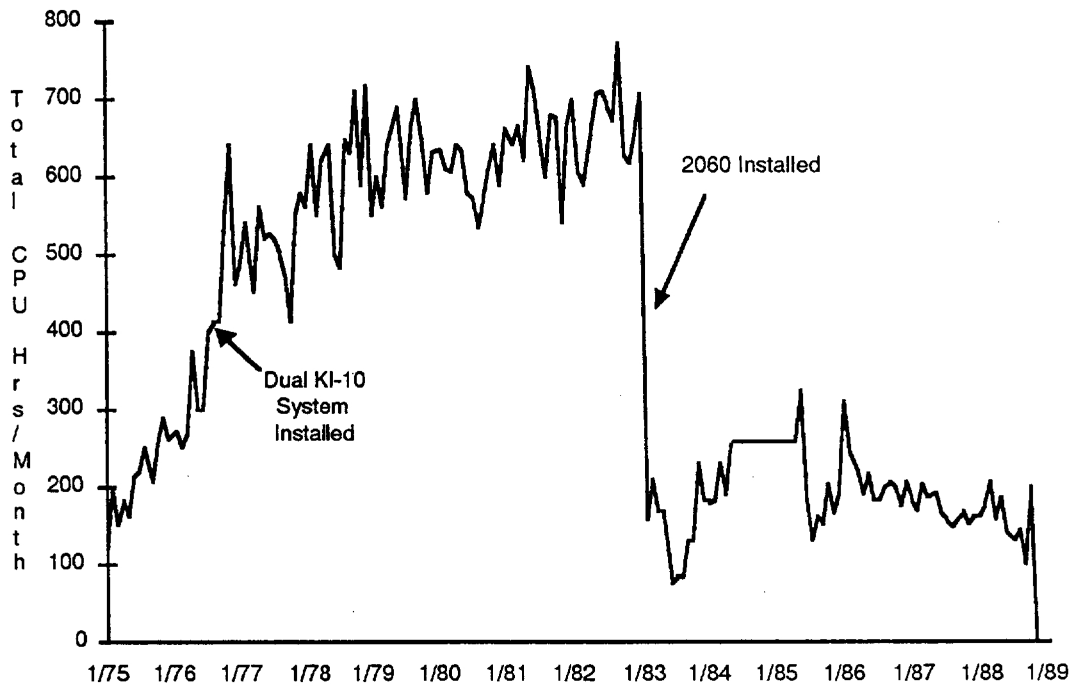

Figure 9. Total CPU Hours Consumed by Month

(2.2) Individual Project and Community Usage

The following histogram (Figure 10) shows the relative central resource usage during the past grant year by national AIM and Stanford collaborative projects and the core research groups before the DEC 2060 phase-out. The bars represent the fraction of total CPU time consumed by each project between May 1, 1988 and October 31, 1988, on the SUMEX-AIM DECsystem 2060 system.

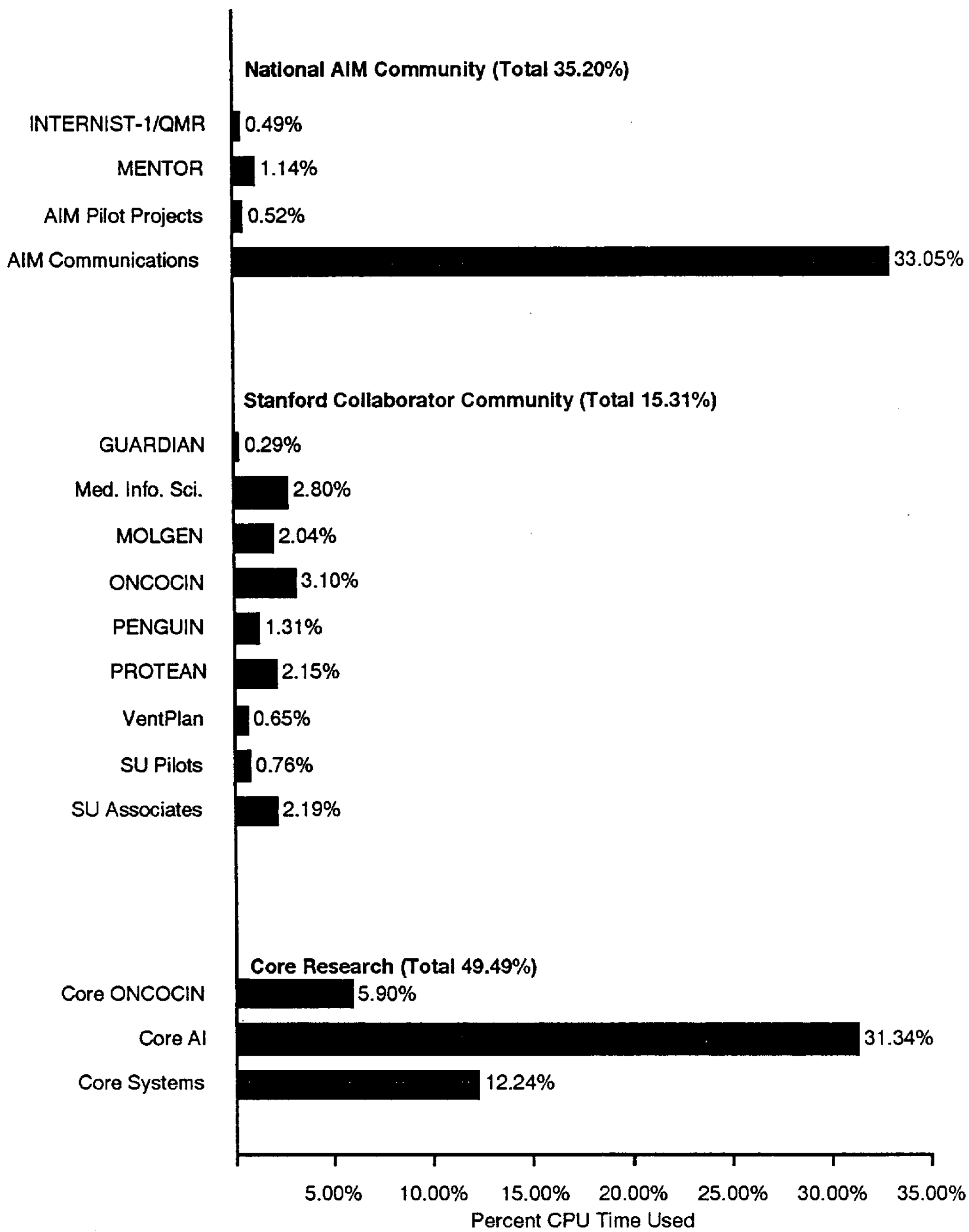

Figure 10. CPU Usage Histogram by Project and Community

Annualized Resource Use by Individual Project - 5/88 through 4/89

The following tabulated form of the data (Figure 12) shows total CPU consumption by project (Hours), total terminal connect time by project (Hours), and average file space in use by each project (Pages, 1 page = 512 computer words). These data were accumulated for each project for the months between May 1988 and October 1988 (when the users were transferred to the SUN-4 server), and were then annualized to the estimated project usage for an entire year.

National AIM Collaborator Community	**CPU (Hours)**	**Connect (Hours)**	**File Space (Pages)**
1) INTERNIST-I/QMR Project Jack D. Myers, M.D. Randolph A. Miller, M.D. University of Pittsburgh	4.20	158	743
2) MENTOR Project *Medical Evaluation of Therapeutic Orders* Stuart M. Speedie, Ph.D. University of Maryland Terrence F. Blaschke, M.D. Stanford University	9.80	6135	2057
3) AIM Pilot Projects PathFinder (Nathwani and Fagan)	4.44	341	1675
4) AIM Communications			
AIM Mail-Only Users	14.87	4837	8993
AAAI Management	6.29	3099	602
MCS Collaborators	6.76	795	984
MOLGEN Collaborators	1.08	82	1543
File/Information Access	253.22	14312	2
Guest and Other	0.86	45	3188
AIM Community Totals	301.52	29804	19787

Figure 11. Table of Resource Use by Project

Stanford Collaborator Community	CPU (Hours)	Connect (Hours)	File Space (Pages)
1) GUARDIAN Project Barbara Hayes-Roth, Ph.D. Department of Computer Science Adam Seiver, M.D. Department of Surgery Palo Alto VA Hospital	2.52	2141	298
2) Medical Information Sciences Edward H. Shortliffe, M.D., Ph.D. Lawrence M. Fagan, M.D., Ph.D. Department of Medicine	23.99	9436	3595
3) MOLGEN Project *Applications of Artificial Intelligence to Molecular Biology: Research in Theory Formation, Testing and Modification* Edward A. Feigenbaum, Ph.D. Peter Friedland, Ph.D. Charles Yanofsky, Ph.D. Depts. Computer Science/Biology	17.50	4129	4623
4) ONCOCIN Project *Knowledge Engineering for Medical Consultation* Edward H. Shortliffe, M.D., Ph.D. Lawrence M. Fagan, M.D., Ph.D. Department of Medicine	26.58	9267	7471
5) PENGUIN Project Gio C.M. Wiederhold, Ph.D. Depts. Computer Science and Medicine	11.24	2956	6315
6) PROTEAN Project Oleg Jardetzky School of Medicine	18.44	4009	2975

Figure 11. Table of Resource Use by Project (Continued)

Stanford Collaborator Community (Continued)	CPU (Hours)	Connect (Hours)	File Space (Pages)
7) VentPlan Project *ICU Management* Lawrence Fagan, M.D., Ph.D. Department of Medicine Adam Seiver, M.D. Department of Surgery Palo Alto VA Hospital Lewis Sheiner, M.D. Dept of Laboratory Medicine UC San Francisco	5.54	2980	861
8) Stanford Pilot Projects			
REFEREE Project (Fagan and Buchanan)	6.52	2256	475
9) Stanford Associates	18.76	10120	3563
Stanford Community Totals	131.09	47295	30176

Core ONCOCIN Research	CPU (Hours)	Connect (Hours)	File Space (Pages)
1) Core ONCOCIN and Medical Information Sciences Edward H. Shortliffe, M.D., Ph.D. Lawrence M. Fagan, M.D., Ph.D. Department of Medicine	50.57	18703	11066
Core ONCOCIN Research Totals	50.57	18703	11066

Figure 11. Table of Resource Use by Project (Continued)

	Core AI Research	CPU (Hours)	Connect (Hours)	File Space (Pages)
1)	Advanced Architectures Edward A. Feigenbaum, Ph.D. Computer Science Department	85.50	25844	7893
2)	Blackboard Architectures Barbara Hayes-Roth, Ph.D. Computer Science Department	46.78	10427	5405
3)	Large Multi-Use Knowledge Bases Edward A. Feigenbaum, Ph.D. Richard Keller, Ph.D. Yumi Iwasaki, Ph.D. Computer Science Department	26.62	12428	2019
4)	Software Design Project H. Penny Nii Computer Science Department	8.28	1426	1467
5)	MS:AI Student Projects Edward A. Feigenbaum, Ph.D. Computer Science Department	5.18	1975	468
6)	Machine Learning Studies Bruce G. Buchanan, Ph.D. University of Pittsburgh	20.43	4230	3922
7)	SOAR Project Paul R. Rosenbloom, Ph.D. Information Sciences Institute University of Southern California	8.46	6668	563
8)	HPP Administration	49.18	15416	9839
9)	HPP Associates	18.04	3119	3825
	Core AI Research Totals	268.47	81534	35401

Figure 11. Table of Resource Use by Project (Continued)

Core Systems Research	**CPU (Hours)**	**Connect (Hours)**	**File Space (Pages)**
1) SUMEX Staff R & D Thomas C. Rindfleisch Departments of Medicine and Computer Science	102.94	29234	17118
2) Systems Associates	1.92	337	1188
Core Systems Research Totals	104.86	29572	18306

System Operations	**CPU (Hours)**	**Connect (Hours)**	**File Space (Pages)**
1) System Operations	724.76	85231	2709
2) Other Opns & Mgmnt Services	100.22	247	1779
System Operations Totals	824.98	85478	4488
Resource Grand Totals	1681.49	292385	119224

Figure 11. Table of Resource Use by Project (Concluded)

III.B. Research Highlights

In this section we describe several research highlights from the past year's activities. These include notes on existing projects that have passed important milestones, new pilot projects that have shown progress in their initial stages, and other core research and special activities that reflect the progress, impact, and influence the SUMEX-AIM resource has had in the scientific and educational communities.

III.B.1. INTERNIST-I/QMR

The INTERNIST-1/QMR project, under Drs. Jack Myers and Randy Miller at the University of Pittsburgh, has been developing a high-level computer diagnostic program in the broad field of internal medicine as an aid in the solution of complex medical problems. This system has won recognition both as a project in artificial intelligence, encoding one of the largest knowledge bases ever attempted, covering almost 600 different diseases and 4500 possible patient findings and as an able problem-solver in internal medicine.

Over the past decade, this program has been able to analyze many hundreds of difficult diagnostic problems, including cases published in medical journals (particularly Case Records of the Massachusetts General Hospital, in the New England Journal of Medicine), CPCs, and unusual problems of patients in the Pittsburgh Medical Center. In most instances, although not all, INTERNIST-I has performed at the level of the skilled internist. The INTERNIST-I program has also been used in recent years to develop patient management problems for the American College of Physician's Medical Knowledge Self-assessment Program.

The QUICK MEDICAL REFERENCE (QMR) program, developed under the leadership of Dr. Miller, incorporates most of the INTERNIST-I knowledge base and diagnostic consultative program, together with much more powerful user interface and query tools than INTERNIST-I ever had. And QMR runs on an IBM PC-AT workstation. QMR has served as a means to distribute the knowledge base to over twenty other academic medical institutions where it can be used as an "electronic textbook" in medical education at all levels — by medical students, residents and fellows, and faculty and staff physicians. This distribution is continuing to expand.

The Pittsburgh group, including researchers, residents in internal medicine, and fellows in medical informatics, is continuing to expand the knowledge base and to polish the QMR diagnostic consultant program. The medical knowledge base continues to grow both in the incorporation of new diseases and in the modification of diseases already profiled so as to include recent advances in medical knowledge.

Limited field trials of QMR were begun in 1987 on the clinical services in internal medicine at the Health Center of the University of Pittsburgh. A "computer-based diagnostic consultation service" was made available to attending physicians and house staff of our two main teaching hospitals. Institutional Review Board (IRB) approval was granted to the service before it was initiated. In the near term, these clinical field trials will be extended to other university health centers which have expressed interest in working with the system.

III.B.2. PathFinder

One of the most difficult areas in surgical pathology is the microscopic interpretation of lymph node biopsies. Most pathologists have difficulty in accurately classifying lymphomas. Several cooperative oncology group studies have documented that while experts show agreement with one another, the diagnosis rendered by a "local" pathologist may have to be changed by expert lymph node pathologists (expert hematopathologists) in as many as 50% of the cases.

The National Cancer Institute recognized this problem in 1968 and created the Lymphoma Task Force which is now identified as the Repository Center and the Pathology Panel for Lymphoma Clinical Studies. The main function of this expert panel of pathologists is to confirm the diagnosis of the "local" pathologists and to ensure that the pathologic diagnosis is made uniform from one center to another. But this approach is only a partial answer to the problem as the panel annually reviews only 1,000 cases whereas more than 30,000 new cases of lymphomas are reported each year.

The PathFinder project, under Dr. Bharat Nathwani of the University of Southern California and M.D./Ph.D. students David Heckerman and Eric Horvitz of the Stanford University Medical Computer Science Group, has been exploring the use of a computer-based diagnostic program that provides advice on over 70 common benign and malignant diseases of the lymph node based on over 100 histologic features. The design of the program was influenced by the architecture of the INTERNIST-1 program, also developed on the SUMEX resource.

Project computer science research is studying formal techniques for decision making under uncertainty, including the assessment and representation of important dependencies among morphologic features and diseases, reasoning about the costs and benefits of alternative information acquisition strategies, the acquisition and use of expert knowledge bases from multiple experts, the customization of the system's reasoning and explanation behaviors to reflect the expertise of the user, and controlling the naturalness of complex formal reasoning techniques. A group of expert pathologists from several centers in the U.S. have showed interest in the program and helped to provide the structure of the knowledge base for the Pathfinder system.

Originally written in Lisp on the SUMEX 2060, PathFinder was converted two years ago to MPW Object Pascal on the Macintosh II. Much of the recent testing and refinement of the knowledge base has been carried out within the Macintosh II environment. The group conducted a study to compare the performance of the system with that of the domain expert. In the evaluation, a community pathologist used the Pathfinder system to analyze a set of difficult cases. Fifty-three cases were were selected in sequence from a large library of referrals. The work showed a close correspondence between the behavior of the system and expert decision making.

III.B.3. The Distributed SUMEX-AIM Community

SUMEX-AIM has undergone a major transition this past year, with the movement of users from the tried and true DEC 2060 resource to a new UNIX-based SUN-4. During the inexorable development and maturation of inexpensive and powerful workstations that are now on most researcher's desks, the 2060 has provided us with a link to the past and a stable continuation of essential research, communications, and other network services for the AIM community. But, under recent budget pressures, we were unable to continue a reliable, well-maintained 2060 service and so, in October 1988, we brought a SUN-4 UNIX computer on-line to replace the old machine for most SUMEX-AIM community functions. We have continued to operate the 2060 in background mode under a much lower cost and less responsive maintenance arrangement but even this level of access will be ended and the machine will be shut down completely by the end of July. We are making every effort to ensure continued access, through the SUN-4, to all the files that were archived under TENEX (1975 - 1983) and TOPS-20 (1983 - present). We will also provide continued access to the annual full file system dumps we have kept for each of the past 14 years.

Of course, after we shut the 2060 down, none of the early seminal programs developed by the SUMEX-AIM community will be accessible — including DENDRAL, MYCIN/EMYCIN, INTERNIST-1, SECS, MOLGEN/UNITS, AGE, PARRY, GUIDON, RADIX, CRYSALIS, PUFF-VM, and on and on... As we turn this next page in SUMEX-AIM history, we realize that most other 2060's that could run these programs are rapidly disappearing as well. so an era is indeed passing.

But even as we watch these changes nostalgically, the future is bright. The Apple Macintosh II workstations we chose are performing extremely well as a general computing environment for researchers and staff, the TI Explorer Lisp machines (including the microExplorer Macintosh coprocessor) are a sound base as the near-term high-performance Lisp research environment, and, after considerable effort, the SUN-4 has taken over most of the 2060 functions as the central system network server (network services, file services, printing services, etc.). Initial user response to the introduction of these systems has been overwhelmingly enthusiastic, even though there have been many "rough edges" to be smoothed out along the way toward full systems integration. Our core development work is coming along well for providing remote access between workstations and servers, integrating a solid support of the TCP-IP network protocols, and building a powerful distributed electronic mail system. The new Mac II mail system will be introduced this summer and we believe that it will be a significant quantum improvement over the serial TTY-based systems of the past.

III.B.4. ONCOCIN

ONCOCIN, developed under Drs. Ted Shortliffe and Larry Fagan at Stanford University, is an expert system for clinical oncology, designed for use in managing chemotherapy after a diagnosis has been reached. Because anticancer agents tend to be highly toxic, and because their tumor-killing effects are routinely accompanied by damage to normal cells, the rules for monitoring and adjusting treatment in response to a given patient's course over time are complex and difficult to memorize. ONCOCIN integrates a temporal record of a patient's treatment with an underlying knowledge base of treatment protocols and rules for adjusting dosage, delaying treatment, aborting cycles, ordering special tests, and similar management details. The program uses this knowledge to help physicians with decisions regarding the management of specific patients.

Oncologists use ONCOCIN routinely for recording and reviewing patient data, replacing the conventional recording of data on a paper flowsheet. With its knowledge of the patient's chemotherapy protocol, ONCOCIN then provides assistance by suggesting appropriate therapy at the time that the day's treatment is to be recorded on the flowsheet. Physicians maintain control of the decision and can override the computer's recommendation if they wish. ONCOCIN also indicates the appropriate interval until the patient's next treatment and reminds the physician of radiologic and laboratory studies required by the treatment protocol.

The ONCOCIN Project started in July 1979 and the first version ran on DECSystem-10/20 mainframes using character-oriented terminals. Limitations in terminal capabilities to adapt to user interface needs and the high cost of central machines necessitated a reimplementation of the system on workstations in 1984. In 1986, we placed the Xerox Lisp machine version of ONCOCIN in the Stanford Oncology Day Care clinic. This version is a completely different program from the mainframe in that it uses graphical user interfaces extensively, new protocols are entered through a powerful graphical data entry interface (OPAL), and it has a revised knowledge representation and reasoning component. In 1987, we began to explore the use of continuous speech recognition as an alternate entry method for communicating with ONCOCIN.

Although we have successfully moved ONCOCIN into a stable and useful system on the Xerox Lisp workstations, it is now clear that this environment will not provide the means for dissemination we need for economical and technical reasons. Based on the success of this earlier work, strong interest has developed, from such diverse quarters as the National Cancer Institute and the Stanford Hospital, for developing a fully operational version of ONCOCIN that can be broadly used in oncology clinics outside our research laboratory. This past year, the Stanford Hospital started a program to assist in the transfer of innovative medical technology out of the laboratory to patient care and ONCOCIN was selected as one of 10 projects to be funded from a large group of competing proposals. So it is once again time to consider a redesign of the system. The dilemma for the project that is still unresolved is how to maintain a cohesiveness between ongoing research work to extend ONCOCIN and generalize it for applications to other domains and the operational needs of a widely disseminated practical system. Much thought has gone into this problem this past year, including issues such as which of the modern workstation alternatives to select (Lisp machine, IBM PC, Apple Macintosh, SUN or NeXT UNIX workstation, ...), what language to pick (C, Lisp, ...), and whether the research and operational systems can really be consistent versions of a single system?

In order to understand the scope and practical issues involved, we have begun an experiment to port ONCOCIN to a TI microExplorer running inside of a Mac II during the last six months. We have completed the translation of the Ozone object-oriented system, the temporal network and most of the reasoner. We will next approach the design of the user interface, which must be rewritten anew, since the current interface depends heavily on the graphical capabilities of the Xerox workstations. We are also starting a study of the overall design and specification of an "integrated" oncologist's workstation, under NCI sponsorship, that will lead to an attempt to coordinate federal, academic, and industrial efforts to implement such a system.

III.C. Administrative Changes

There have been few administrative changes within the project this past reporting year. As we reported last year, Professor Buchanan, who had been one of the leaders of the core AI research effort, left in July 1988 to take a faculty post at the University of Pittsburgh. At about the same time, two Research Associates joined the core AI research effort, Dr. Yumi Iwasaki, who finished her PhD at Carnegie Mellon University, and Dr. Thomas Gruber, who did his PhD at the University of Massachusetts. Overall, these changes have caused a shift in emphasis of the core AI research work away from the machine learning work that Prof. Buchanan headed and toward the large, multiuse knowledge base work that has been getting underway.

Effective August 1, 1988, we discontinued the fee-for-service cost recovery system we have been using during the last two years to collect from Stanford users the resource operating costs not covered by NIH support. The operating cost shortfall is now being paid entirely by KSL core research projects (i.e., none of the Stanford or national collaborator projects pays anything for their increasingly communication-oriented use of SUMEX-AIM) — under the previous cost center model, *all* Stanford users paid fees. These direct payments avoid the large accounting overhead of a cost center to collect relatively small bills each month and are now made on the basis of a "bulk purchase" agreement, reached between resource management and the KSL Principal Investigators. Our reasons for this move are described in more detail in section III.D.2 (Cost Center Management).

III.D. Resource Management and Allocation

III.D.1. Overall Management Plan

Early in the design of the SUMEX-AIM resource, an effective management plan was worked out with the Biotechnology Resources Program (now Biomedical Research Technology Program) at NIH to assure fair administration of the resource for both Stanford and national users and to provide a framework for recruitment and development of a scientifically meritorious community of application projects. This structure has been described in some detail in earlier reports and is documented in our recent renewal application. It has continued to function effectively as summarized below.

The AIM Executive Committee meets periodically by teleconference to advise on new user applications, discuss resource management policies, plan workshop activities, and conduct other community business. The Advisory Group meets as needed to review project applications. (See Appendix C for a current listing of AIM committee membership).

We actively recruit new application projects and disseminate information about AI in biomedicine. With the development of more decentralized computing resources within the AIM community outside of Stanford, the use of SUMEX resources by AIM members has shifted more and more toward communication with colleagues and access to information.

With the advice of the Executive Committee, we have opened SUMEX-AIM resources widely to biomedical users desiring electronic communications facilities.

We have carefully reviewed on-going projects with our management committees to maintain a high scientific quality and relevance to our biomedical AI goals.

We continue to provide active support for the AIM workshops. The most recent one was held in the spring of 1988 at Stanford University, under the auspices of the American Association for Artificial Intelligence (AAAI). Planning is underway for an AIM workshop in the spring of 1990.

We have continued to provide systems advice to users attempting to set up computing resources at their own sites, based on the expertise developed in the SUMEX resource environment.

We have tailored resource policies to aid users whenever possible within our research mandate and available facilities.

III.D.2. Cost Center

Our plan for the term of the current grant had been a resolute but responsible transition of the SUMEX-AIM resource to a distributed community model of operation. While there has continued to be a group of national and local users — particularly young projects needing seed support prior to obtaining major funding — that depend on a central shared resource like the SUMEX mainframe, powerful and widely available workstation equipment has rapidly become accessible at a cost that most projects can afford, even young ones. Thus, the period of critical dependence on the DEC 2060 for raw computing cycles is largely past and its role in supporting routine computing and communication services was soon to be replaced by other more cost effective equipment. We were in the process of implementing the phase-out of the SUMEX 2060 machine over this year when the 11% budget cut forced us to shut down the machine to users in a much more precipitous manner (see the report of Core Systems Development in Section III for more details). In the course of this much more hasty change-over from the DEC 2060-based to a SUN-4-based resource, it became clear that the cost center fee-for-service approach to recovering unsubsidized operating costs was counterproductive to our efforts. There were three main reasons for this assessment:

- In the face of such a large budget cut, the administrative overhead costs associated with running the cost center became prohibitive. Since they would have had to be recovered through the rates charged, this would have raised rates relatively much faster than previously planned, and our user projects were unprepared for this magnitude of increase.
- The rapidly rising costs of fee-for-service computing were having a strongly negative effect on the research we were striving to support — particularly student computing. Students were receiving increasing pressure from their mentors and sponsors to contain costs and so were confronted with unworkable decisions about how to avoid extra computer runs to polish papers for publication or to continue development or running test cases on their research programs. This was exactly the kind of counterproductive bureaucracy we had struggled to avoid in the early phases of the SUMEX-AIM resource, but were increasingly forced into because of pressures from BRTP.
- As we moved to a distributed model of computing, it was becoming increasingly difficult to assess realistically a measure of resource usage on which to base fee-for-service charges. Since many of the services used in such an environment do not have any accounting associated with them, because vendors do not provide the tools or because such measurements would add an unreasonable overhead to the basic service mechanism, we simply cannot allocate costs on the basis of usage because we do not have a valid measure of usage.

Thus, effective August 1, 1988, we discontinued the fee-for-service cost recovery system used during the precious two grant years. The operating cost shortfall is now being paid entirely by KSL core research projects (i.e., none of the Stanford or national collaborator projects pays anything for their increasingly communication-oriented use of SUMEX-AIM). The payments are negotiated between resource management and the KSL Principal