Computer Network-Based Scientific Collaboration in the Energy Research Community, 1973–1977: A Memoir

Gerald Estrin

In 1973, four national laboratory groups and three university energy research groups started an investigation of whether scientific collaboration could be improved with computer networks. Four years later, the seven groups were piggybacked onto the Arpanet, and more than 60 case histories of applications had been documented. This article reviews the early collaboration and tells how the energy research community was introduced to modern computer networking.

In 1973, some energy research groups started an investigation of whether scientific collaboration could be improved with computer networks. This Energy Research and Development Agency project lasted from 1973 to 1977 and culminated in a four-volume report, *General Purpose Computer Networks and Resource Sharing in ERDA*,[1] issued in 1977, bringing together the findings of investigators at Argonne National Laboratory, Brookhaven National Laboratory, Lawrence Berkeley Laboratory, Lawrence Livermore Laboratory, Massachusetts Institute of Technology's Laboratory for Nuclear Science, New York University's Courant Institute, and UCLA's Computer Science Department. (See the sidebar on software and also see the sidebar on organizations.) The project was supported by the Division of Physical Research through its Mathematics and Computer Science Research Program directed by Milton E. Rose (deceased 22 August 1993). The report was compiled by the ERDA Working Group on Computer Networking, of which I was Chair. There were 12 contributing authors, and this article draws liberally from the report (see Figure 1).

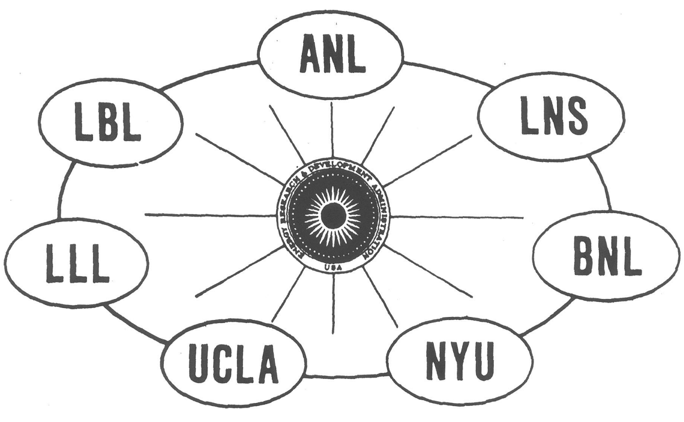


Figure 1. Logo of Working Group on Computer Networking. (From General Purpose Computer Networks and Resource Sharing in ERDA, vol. 1)

The investigation was of heterogeneous networks in which a variety of dissimilar computers and operating systems were interconnected nationwide. The investigation itself was an example of computational resource sharing through a heterogeneous network. More than 20 individuals, representing seven ERDA computing sites, coordinated, compiled, and formatted the data that made up the report. Volume 1 summarized the entire four-year investigation into computational resource sharing through heterogeneous computer networks and offered recommendations for the future. Volume 2 discussed technical and financial issues. Volume 3 analyzed the benefits and barriers encountered and provided case histories. Volume 4 provided the basic information needed to access and use the ERDA sites that were connected to the Arpanet.

This story begins with presentation of the first volume's summary and recommendations:

> The purpose and scope of the study is stated as:
>
> The ERDA research community encompasses the national research laboratories, universities, and private contractors undertaking energy related research, development, and demonstration projects. ERDA's mission requires the develop-

Software

BKY-Sesame: LBL Interactive System Monitor
FTP: file transfer protocol (used to transfer files between computers)
Linda: BNL modeling software
MACSYMA: MIT symbolic differential equation solver
MINPACK: ANL minimization software
MULTICS: MULTiplexed Information and Computing Service (a time-sharing operating system developed at MIT that served with the Xerox Graphic Printer for group generation, editing, and production of the four-part report on General Purpose Computer Networks and Resource Sharing in ERDA—see sidebar, Network Experience Report, UCLA-1).
NOS-TELEX: NYU's CDC 6600 time-sharing software
Planet: Institute for the Future's teleconferencing system
RATS: LLL gateway software
Runoff: text formatting program
SARA: UCLA System ARchitects Apprentice design and simulation software
Scope-Intercom BNL: CDC 6600 time-sharing software
SETL: NYU set logic software
Spice: LBL electronic chip design software
Telenet: an early commercial network service
TELNET: telecommunications network protocol
TSO: Time Shared Option (IBM)
WYLBUR: ANL time-sharing software

Organizations

Ames-ISU: Iowa State University at Ames, Iowa had extensive energy research support.
ANL: Argonne National Laboratory
BNL: Brookhaven National Laboratory
CCA: Computer Corporation of America, developer of the Datacomputer (CCA had received government support in the early 1970s to develop an experimental machine called the Datacomputer that would be capable of storing and managing a very large database. For example, it had the goal of managing a 10-year history of weather data that would be large enough to swamp even the largest ERDA sites. This provided a clear example of unique hardware and software that might be shared if accessible through the Arpanet. Note that Figure 5 shows CCA as a site on the Arpanet.)
ERDA HQ: The Energy Research and Development Agency Headquarters staff participated in our networking effort by facilitating our work and keeping the Washington office aware of our progress.
ERDA MCS: The Energy Research and Development Agency's Mathematical and Computer Sciences office.
LASL: Los Alamos Scientific Laboratory
LBL: Lawrence Berkeley Laboratory
LLL: Lawrence Livermore Laboratory
LNS: Laboratory for Nuclear Science, Massachusetts Institute of Technology
MFE: Magnetic Fusion Energy research program.
MFENET: Magnetic Fusion Energy NETwork; this included a CDC 7600 at LLL and would include a gateway and a CDC 6600 and other equipment.
NYU: Courant Institute, New York University
NASA-Ames: The Illiac IV was scheduled for delivery to NASA-Ames and was to be another example of a unique hardware resource that would benefit from widespread access through the Arpanet.
RHEL: Rutherford High Energy Laboratory, UK
SACNET: Secure Automatic Communications Network (available in the early 1970s but was not suitable for the kind of open resource sharing that we felt would provide great benefits to energy research).
SLAC: Stanford Linear Accelerator Center
U of I: University of Illinois
UCLA: University of California, Los Angeles (Computer Science Department)
UCLA CCN: University of California, Los Angeles (Campus Computer Network)

ment of large scale modeling systems, analysis programs and data bases. ERDA maintains an extremely large inventory of computing facilities which supply powerful computational capabilities for a wide spectrum of energy research and development activities.

Since the ERDA community is geographically distributed, it has been difficult to provide the complete spectrum of computational and data management facilities equally to all researchers. The consequence has been duplication of effort in some cases, and lack of computational facilities in other cases. General purpose computer networks offer the prospect of a more balanced distribution of computer resources throughout the ERDA community.

Since energy problems involve researchers in diverse fields residing at different institutions, more effective scientific collaboration has become necessary. If modern techniques such as electronic mail and computer-based teleconferencing were made accessible to scientists within and without ERDA that need might be fulfilled.

In the 1960s and 1970s, the national weapons laboratories (Los Alamos and Lawrence Livermore) played a major role in getting computer manufacturers and universities to propose new computer systems to solve problems that took too long to solve on exist-

WORKING GROUP ON COMPUTER NETWORKING

Mathematical and Computer Sciences (EE-03-04)
Office of Engineering,
Mathematical and Geo-sciences
Division of Basic Energy Sciences
ERDA

Gerald Estrin, Chairman

IMPLEMENTATION PANEL
Ed Franceschini. Chmn, NYU-Courant
Lawrence Amiot, ANL
Graham H. Campbell, BNL
James E. Donnelley, LLL
Robert Fink, LBL
Lyman R. Hazelton, Jr., LNS-MIT

INVESTIGATORS PANEL
Dennis E. Hall, Chmn, LBL
Ronald A. Bare, ANL
Graham H. Campbell, BNL
James E. Donnelley, LLL
Ed Franceschini, NYU-Courant
Terence E. Gray, UCLA
Lyman R. Hazelton, Jr., LNS-MIT
Muriel Kannel, LNS-MI1

OBJECTIVES PANEL
Gerald Estrin, Chmn, UCLA
Donald M. Austin, LBL
Charles N. Brashears, ERDA HQ
Elizabeth J. Campbell, LNS-MIT
Sidney Fernbach, LLL
Max Goldstein, NYU
Glen Ingram, ERDA MFE
Edward C. Jacques, ERDA HQ
Clair G. Maple, AMES-ISU
Robert B. Marr, BNL
James C. T. Pool, ERDA MCS
Milton E. Rose, ERDA MCS
Richard J. Royston, ANL
Ronald Schwartz, ERDA HQ

Figure 2. Implementation Panel, Investigators Panel, and Objectives Panel. (From Report on General Purpose Computer Networks and Resource Sharing I in ERDA)

It is often forgotten that computer based systems cannot function without the involvement of human beings. Many people contributed to this investigation. The current panel members are listed on the inside front cover. On this page we wish to acknowledge all others who made significant contributions. If anyone has been left out it was inadvertent.

Harvey Baumel, LNS
Jerome Becker, ANL
Herbert Bernstein, BNL
Robert Braden, UCLA-CCN
Bertram Bussell, UCLA
Irving Cohen, UCLA
Robert Oaley, MIT
Y. Feinroth, NYU
Anna Fisher, LNS
Michael Ford, LNS
Sheldon R. Furst, LBL
Robert L. Gardner, Jr., UCLA
Charles Gear, U of I
Robert Harvey, LBL
Robert Healy, LBL
Bradford Heckman, LBL
April Heiberger, ANL
S. Heller, BNL
Steve Kelem, UCLA
Roger B. Lazarus, LASL
P. Maclean, NYU
F. McGirt, LASL
Robert McMahon, ANL
Deane W. Merrill, LBL
A. Milonas, NYU
B. Miller, MFE
M. Nihalani, NYU
Doris Pahis, ANL
Beverly Robertson, UCLA
W. Russell, NYU
Kalman Saffran, LNS
Steven Scholbrock, ANL
M. H. Schwartz, ERDA
Marsha Joyce Shapiro, UCLA
Y. Shimamoto, BNL
L. Smith, SLAC
T. Stuart, NYU
T. Tsun, NYU
B. Wells, LASL
Joseph Wells, SLAC
Steven Wolfe, UCLA-CCN
Jack Worlton, LASL
L. Yeh, BNL

Figure 3. Significant contributors to working group effort. (From report on General Purpose Computer Networks and Resource Sharing in ERDA)

ing systems. The laboratories committed to purchase the first systems that came off the production line of key computer manufacturers (e.g., IBM and Control Data), even providing programmers to develop operating systems and application programs. Following traditions established at high-energy accelerator installations, researchers would go to the computational centers to do their experiments and then return to their home bases. There were specialized ERDA network developments to facilitate remote communication, but they were burdened with complex security requirements. We could ignore security issues when we dealt with the feasibility of using the insecure Arpanet. When we organized the resource-sharing study, we could draw on the experience of those who had worked with Arpanet. Furthermore, there was interest on the part of the Defense Advanced Research Projects Agency to bring powerful ERDA resources onto the network,

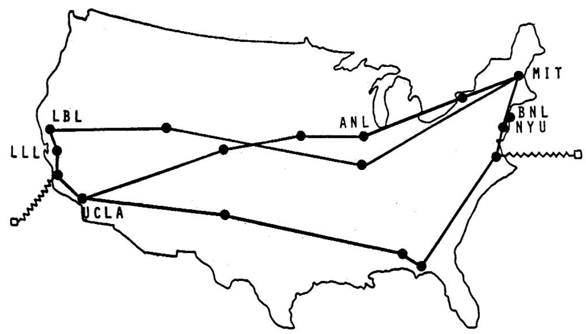


Figure 4. ERDA sites on the ARPA network. (From "Summary, Conclusions and Recommendations," *General Purpose Computer Networks and Resource Sharing in ERDA*, vol. 1, p. 14)

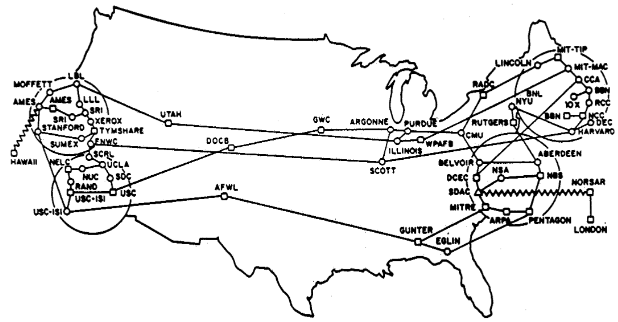


Figure 5. Arpanet geographic map, August 1976 (From "Summary, Conclusions and Recommendations," *General Purpose Computer Networks and Resource Sharing in ERDA*, vol. 1, p. 14)

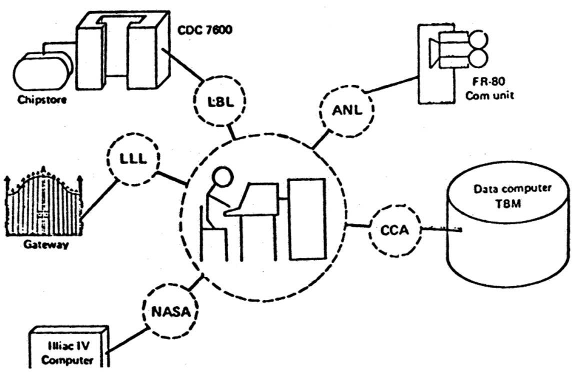


Figure 6. Accessing unique hardware: Reachable Through Arpanet (From "Implementation of Arpanet Facilities," *General Purpose Computer Networks and Resource Sharing in ERDA*, vol. 2, p. 5)

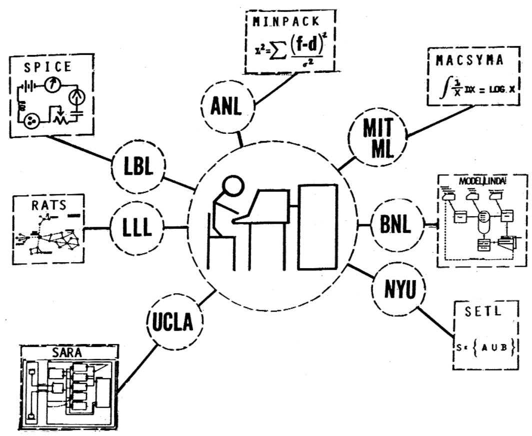


Figure 7. Accessing unique software: Reachable Through Arpanet (From "Implementation of Arpanet Facilities," *General Purpose Computer Networks and Resource Sharing in ERDA*, vol. 2, p. 7)

guaranteeing a cooperative attitude.

As of July 1977, seven ERDA research centers were connected and had direct on-site access to the full range of Arpanet facilities, including remote time-sharing systems, file transfer, and remote batch capabilities. During the previous two years, more than 60 network-based applications involving ongoing research projects had been initiated and analyzed. Some technical barriers to easy use of the network (such as inadequate consulting support, differences in command and programming languages, and saturation of host machines) were discovered and highlighted. Some important nontechnical barriers to network-based resource sharing were exposed; the most significant was potential loss of funding at local computing facilities because users could access lower-cost or more-powerful remote facilities.

Figure 2 identifies the working group and the study panels. Figure 3 lists people who made significant contributions to the project. Figure 4 shows the ERDA sites on the Arpanet. Figure 5 gives the complete Arpanet geographic map as of August 1976. Figure 6 shows some of the unique hardware. Figure 7 depicts some of the unique software made accessible by connection to the Arpanet.

(The software included the SARA design and simulation software at UCLA, the RATS gateway software at LLL, the Spice electronic chip design software at LBL, the MINPACK minimization software at ANL, the MACSYMA symbolic differential equation solver software at MIT, the Linda modeling software at BNL, and the SETL set logic software at NYU.)

Figure 6 and Figure 7 highlight the basic appeal of resource sharing through networks. If the solutions of problems at one installation were limited by the performance of local computers and if more-powerful computational resources were accessible at a remote site, then pressure would obviously grow to gain access to the remote resources. If unique software had been developed over many years at a site, then users would move data through the network and determine the applicability of that unique software. If successful, this remote use would avoid the costly and sometimes impracticable process of transporting or redeveloping software; at the very least, it would allow testing the effectiveness of the remote software as a prelude to developing local production software. The conclusions of the resource-sharing study were set down as follows:

> It has been demonstrated that general purpose computer networks can be used in the ERDA community to achieve two principal benefits of remote resource sharing: better utilization of resources and improved quality of research.
>
> The Arpanet is a smoothly functioning, general purpose computer communications network which supports the most versatile capability for heterogeneous, independent computer centers in existence. In particular, the large scale computational facilities available at ERDA centers can be adapted to support practical resource sharing on an operational basis.
>
> Unique or special computational resources will be effectively utilized in energy research and development programs if the resources are readily available to researchers.
>
> Electronic mail and computer teleconferencing increase scientific collaboration and management coordination.
>
> Duplication of effort in software development can be reduced by providing network facilities for accessing special programs at remote sites and for transporting software between cooperating sites.
>
> Large scale computations can be off-loaded to large centers, providing extra resources to a wider range of researchers.
>
> Interaction with non-ERDA research activities in the university and private sectors is greatly enhanced by computer network facilities.
>
> The initial costs for network connections are minimal compared to the long term benefits obtained.

The following section describes the experiments in implementation of Arpanet facilities.

Implementation of Arpanet Facilities

An Implementation Panel, chaired by E. Franceschini at NYU-Courant Institute (see Figure 2), undertook the task of overseeing the connection of all sites to the Arpanet. The panel noted:

> The 60 or so nodes on the Arpanet offer a variety of hardware capabilities which have never been collected in a single site accessible to the general scientific community.... These include, for example, the Illiac IV parallel processor machine at NASA-Ames, the CDC 7600s at LBL and BNL, an IBM 370/195 at ANL and an IBM 370/195 at RHEL for large scale computation. There are terabit mass storage devices available at CCA (the Datacomputer), at LBL (the IBM Photodigital Chipstore), and at NASA-Ames (the Unicon device). Computer output to microfilm devices can be accessed at several sites ... providing high quality graphics output for network users.
>
> It is generally accepted that standardization in computing is slow in coming. Even a classical language like Fortran, because it permits machine dependencies to exist covertly, cannot always claim portability for its programs. Furthermore, really large scale software systems tend to take advantage of machine architectures and thus create portability problems. If they can be used over networks, there is no need for transporting.
>
> Access to unique large data bases presents a natural application of distributed processing. To transport a large data base to many computer systems is a difficult and wasteful task. While few people ever require the entire data base, they often need to sample freely from it. If one can store the data efficiently on one system, and provide easy and efficient access to all users, real savings in time and effort may accrue. The Datacomputer at CCA is a facility for implementing such a scheme. It provides a system and languages to permit the creation, maintenance, and access for data bases on a terabit storage device.
>
> Computer networks offer the possibility of enhancing research collaboration by scientists from several institutions. The field of high energy physics is particularly affected by this requirement, because there are only a few large particle accelerators available. In order to do experiments, groups from several institutions join their efforts in setting up the experiment, taking data, and analyzing the data. In addition to facilitating the sharing of data and software, computer networks can, through facilities such as network mail and the Planet teleconferencing system, enhance communications among the researchers.

Table 1 shows the implementation milestones that were met between 1973 and 1977.

Remote Resource-Sharing Experience and Findings

In October 1975, the technical problems associated with connecting the participating ERDA sites to the Arpanet were understood suf-

Table 1. Implementation Milestones: 1973–1977

1973	Q2	All facilities available at MIT-MULTICS and UCLA-CCN; IMP and HOST/IMP. Interface installed at LBL; Gateway to MFENET.
	Q3	Implementation goals defined at ANL, BNL, LBL, LNS, NYU, UCLA; Time sharing access to network from LLL
1974	Q1	ARPA approval for ANL, BNL, and NYU to join Arpanet.
	Q3	Varian V73's installed at ANL.
	Q4	PDP 11/40 installed at NYU.
1975	QI	IMP installed at ANL; HOST/IMP. Interface installed at NYU.
	Q2	Very distant host connection installed at BNL; IMP installed at NYU; Time sharing access to network from LBL; Time sharing access to CDC-6600 (BKY-SESAME) at LBL.
	Q3	Time sharing access to network from ANL and NYU; PDP 11 installed at BNL; File transfer and batch job submittal to CDC 6600/7600.
1976	Q2	Time sharing access to CDC-6600 (NOS-TELEX) at NYU.
	Q3	Batch job submittal to IBM 370/195 at ANL; PDP 11/34 upgrade installed at BNL; File transfer and batch job submittal to CDC 6600 at NYU.
	Q4	Time sharing access to CDC-6600 (Scope-Intercom) at BNL.
1977	Q1	Time sharing access to IBM 360/75 (TSO; WYLBUR) and file transfer at ANL.
	Q2	Batch job submittal to CDC 6600/7600 and file transfer at BRL.

Source: "Implementation of Arpanet Facilities," General Purpose Computer Networks and Resource Sharing, *vol. 2, p. 12.*

ficiently to face the critical measures of success for the resource-sharing project. An Investigators Panel was formed, chaired by Dennis Hall of Lawrence Berkeley Laboratory. The panel set out to gain firsthand experience in using the network for energy research.

The individual panel members explored use of the network for particular projects. At each ERDA site, network primers were prepared to ease access to remote ERDA resources. These primers were distributed to panel members and, with available consulting, made it reasonable for energy researchers to use the network; furthermore, guest accounts were set up to remove any administrative funding barrier. Finally, more than 60 case histories were documented at the seven ERDA sites, all based on active projects. Volume 3 of the ERDA report presents all of the case histories.

The Investigators Panel explained some technical problems encountered in remote resource sharing:

> ERDA scientists prescribe computational steps in a programming language such as Fortran, and these are made machine readable either by keypunching or by direct entry from an interactive terminal. For most applications it is also necessary to provide formatted data to the program. Most often these are supplied by automatic collection devices, or by very large online data bases.
>
> To cause the computation to be performed, the scientist must also specify the operating system services needed. Examples of these include input/output and data handling services, job and task control, security and protection, accounting information, and network access facilities. In short, these are all the operating system resources which are not selectable in the programming language itself. Specification is done in what has come to be called a "command language" or "job control language," and must also be made machine readable....
>
> The separation between programming languages and command languages is not a clean one ... what may be specified in a given Programming language at one host may have to be specified in the command language at a different host. This is a continual source of difficulty in the development of exportable software, and is encountered repeatedly in network applications.
>
> The effort required to learn enough of a command language to invoke a given task is called the "startup barrier" for that task.... there are no accepted command language standards.... In the Arpanet community alone there are 18 major operating system command languages. A similar diversity is found in the ERDA sites participating in this study.

Important economic barriers, alluded to earlier, were also highlighted:

> One important use of the network is to offload large scale computing tasks from smaller machines to the more powerful machines at the larger sites ... these tasks represent a major source of operating funds.... In light of this, the high degree of cooperation among the ERDA laboratories participating in this study is both remarkable and encouraging.

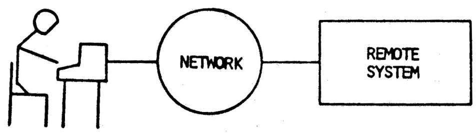


Figure 8. Using the network as a high-quality, high-speed telephone connection. (From *Remote Resource Sharing Experience and Findings*, vol. 3, p. 5)

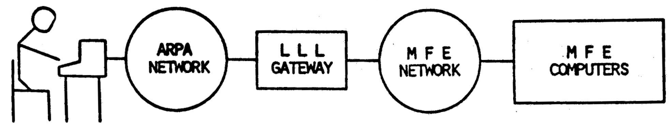


Figure 9. Using the LLL gateway to connect the ARPA net to the MFE net (From *Remote Resource Sharing Experience and Findings*, vol. 3, p. 5)

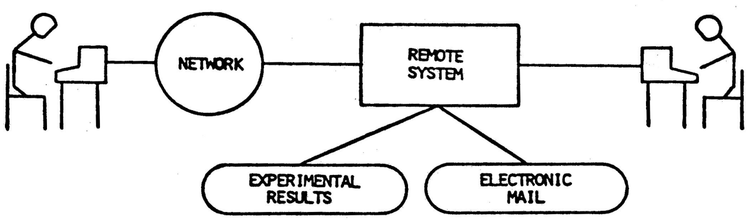


Figure 10. Using the network and electronic mail for scientific collaboration (From *Remote Resource Sharing Experience and Findings*, vol. 3, p. 5)

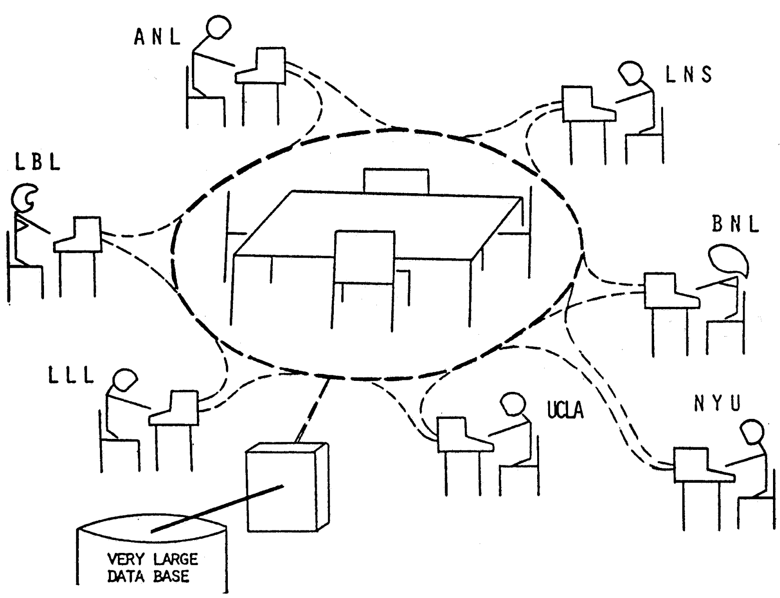


Figure 11. Teleconferencing (From *Remote Resource Sharing Experience and Findings*, vol. 3, p. 7)

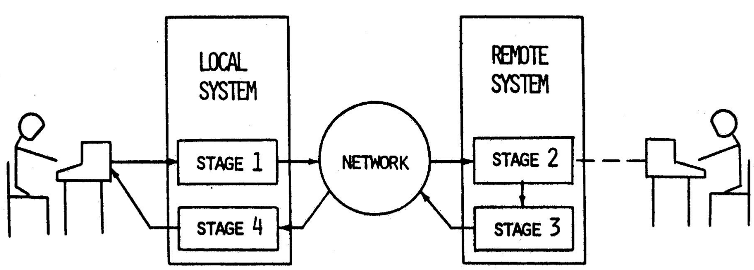


Figure 12. Network experience report UCLA-1 (From *Remote Resource Sharing Experience and Findings*, pp. 129–131, Appendix H)

The ERDA Investigators Panel characterized the case histories in Figure 8, Figure 9, Figure 10, Figure 11, and Figure 12.

At each site, the first use of the network was as a high-quality, high-speed telephone connection as exhibited in Figure 8. There was a negligible start-up barrier, because the connection was usually made to a familiar system; for example, a physicist who transferred to LBL from the Rutherford Laboratory in the United Kingdom needed to finish his Gamma Detector Efficiency Code that he had started at RHEL. After a few minutes of instruction, he was able to access the 8,000-mile-distant Rutherford computer complex from LBL as if it were local. Similarly, scientists were able to access the Magnetic Fusion Energy network developed at Lawrence Livermore Laboratory and thereby run programs on a Control Data 7600 as shown in Figure 9. Figure 10 exemplifies the kind of scientific collaboration that was carried out using electronic mail and maintaining a continuous transcript of every information exchange. A much more dramatic example of scientific collaboration arose in the distributed generation of the four-volume remote resource-sharing report using the facilities illustrated in Figure 11 and described in the *Network Experience Report UCLA-1* (Figure 12) reproduced from volume 3, *Remote Resource Sharing Experience and Findings*. A similarly dramatic example illustrating the use of software developed at different sites to solve a problem and avoid costly local software development is illustrated in the sidebar on Using Complementary Software to Reduce Duplicate Effort and described in the sidebar on Network Experience Report LBL-7.

Using Complementary Software to Reduce Duplicate Effort

(From *Remote Resource Sharing Experience and Findings*, vol. 3, p. 7)

Network Experience Report UCLA-1
ERDA Program: BES–M&CS
Title: Distributed Report Generation
Cooperating Sites: ANL, BNL, LBL, LLL, LNS, NYU, UCLA
UCLA Contact: G. Estrin
Resources Accessed: ANL, BBN.CCN, CCA, LBL, MIT-Multics, LLL, BNL, NYU, MIT-MC, MITXGP

Motivation

Producing this report required the participation of ERDA sites that were widely distributed geographically. Each site was involved in determining what information should be in the report, who should be responsible for which parts, and how they should be organized. Each site was to prepare their own set of appendices, which would eventually be merged into one set. In addition, individuals were charged with coordinating certain sections and writing portions of the main body. From the first discussion of the report, it was clear that computer teleconferencing could be an extremely valuable communication and coordination tool. Later it became even more clear that a great deal of time could be saved during the production phase of the report if network resources were fully utilized. In addition, the use of text formatting and editing programs, and a Xerox Graphic Printer, promised the possibility of a better report than could have been produced using conventional manual methods.

Perhaps the most important motivation was the prospect of distributing portions of the report for review as soon as they were written and at electronic speeds. Distributing changes was equally important.

Procedure

It was agreed that the report would be produced using text formatting programs at MIT-MULTICS and MIT's Xerox Graphic Printer. A new directory was established in the MULTICS Storage System for files associated with this report. In some cases, portions of the report were generated at an author's local host and then transmitted, using a File Transfer Program, to MULTICS. In other cases, MULTICS was accessed remotely using a network connection and text files created there using one of the MULTICS text editors. Since MULTICS is connected to both Arpanet and Telenet, the report authors (and staff) could establish a connection by dialing a local Telenet port, using an Arpanet TIP or going through their local host and establishing a telnet connection. It should be mentioned that there are a wide variety of text editing programs in existence, and the network gives authors the freedom to choose one they like.

Frequently, a remote host was accessed purely because that host had a "better" text editor than the one available on the local host. All participants were given access to the files being accumulated at MULTICS so everyone could review and comment, as well as see visible signs of progress. The Planet teleconferencing system at BBN was used extensively during the report writing period to facilitate dissemination of procedures, guidelines, status, etc. and as a message board for participants to solicit help in accomplishing a particular task, such as moving a file from one host to another. Other remote resources were used in preparing this report as well. A facility at MULTICS for storing all of the user input and system output from an interactive terminal session on a file was used in preparing some of the scenarios in Appendix F, which show how to use some of the resources at each of the participating laboratory sites. A procedure was established to aid in communicating suggested changes to a portion of the report. Once the author had placed the section draft online, he would send a message to whomever he wanted to review the section, stating the name and location of the file, as well as any necessary access control information. Then the reviewer(s) would print the section locally and annotate changes on their own copy. Then they would make a copy of the file and, using a text editor, incorporate their suggested changes in this copy of the file. The convention was established that all changes would be made in uppercase, so that they would easily stand out when the author was perusing the modified file. Deletions were indicated by a series of *X*s. After the suggested modifications were made, the author—and whoever else was involved in that section—was given the name and location of the modified file. The modified file was then printed locally by the interested parties for further review. Once the initial versions of each section were available at MULTICS, they were grouped into larger sections, and additional format control information was inserted. The text was then processed by a text formatting program called Runoff and the output directed to a Xerox Graphic Printer [XGP]. At this point, a meeting was held to review the initial version of the report in its printed form, make additional revisions, and merge graphical information. After all revisions had been made to the text, the entire report was processed once again by the Runoff program and printed on the XGP. During the period when files were being accumulated at MULTICS, procedures were established for periodically making backup copies of all files and storing them on the CCA Datacomputer. Also, since it was important for all of the participants to have access to all the files, procedures were established to make sure that one person making changes to a file did not wipe out the changes made by someone else.

continued on p. 50

continued from p. 49

Observed Benefits

A much higher quality report was produced than could have been possible using conventional methods (i.e., typewriters and the Post Office). In addition, the work was accomplished in a much shorter time than would have been possible otherwise. Only three face-to-face meetings were required from inception to completion (a period of about six months). This was largely due to the effectiveness of the Planet teleconferencing system as a coordination tool. Planet also minimized the number of lengthy phone calls between participants. From the time the decision was made to use network facilities to produce the report to the printing of the first version of the completed report was less than three weeks.

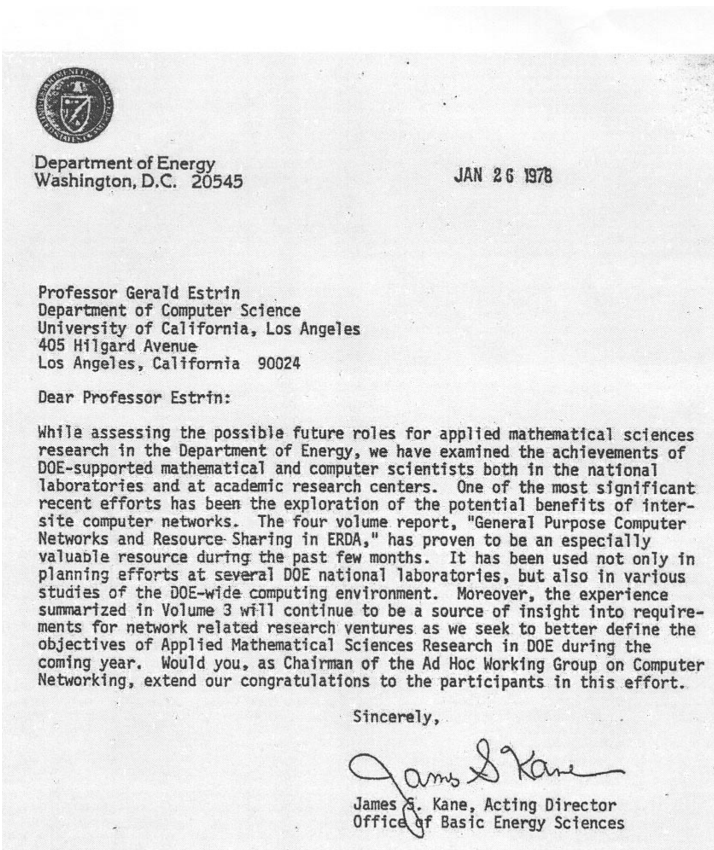
Observed Barriers

Becoming familiar with the various network resources, such as text editors, formatters, file transfer programs, etc., required additional learning for many of the participants. While not an insurmountable barrier, this cost can be minimized only through greater standardization of key network resources (e.g., command languages) and better educational tools, including consultants, self-teaching programs, and improved documentation—both online and conventional.

Additional problem areas related to differences in writing style and format of the several coauthors and the problems associated with controlling font changes on the XGP printer. The former problem provided strong evidence that time spent in establishing standards for style and format—before anyone starts writing—is more than worthwhile. In fact, the development of text editor macros that define a standard format for, say, experimental results, is an excellent idea. The difficulties associated with font changes on the XGP were due to limitations of the text formatting program. Escape sequences were defined that would later be translated to XGP font change commands. These escape sequences proved to be confusing and cumbersome.

Evaluation

Successful.



Department of Energy
Washington, D.C. 20545

JAN 26 1978

Professor Gerald Estrin
Department of Computer Science
University of California, Los Angeles
405 Hilgard Avenue
Los Angeles, California 90024

Dear Professor Estrin:

While assessing the possible future roles for applied mathematical sciences research in the Department of Energy, we have examined the achievements of DOE-supported mathematical and computer scientists both in the national laboratories and at academic research centers. One of the most significant recent efforts has been the exploration of the potential benefits of inter-site computer networks. The four volume report, "General Purpose Computer Networks and Resource Sharing in ERDA," has proven to be an especially valuable resource during the past few months. It has been used not only in planning efforts at several DOE national laboratories, but also in various studies of the DOE-wide computing environment. Moreover, the experience summarized in Volume 3 will continue to be a source of insight into requirements for network related research ventures as we seek to better define the objectives of Applied Mathematical Sciences Research in DOE during the coming year. Would you, as Chairman of the Ad Hoc Working Group on Computer Networking, extend our congratulations to the participants in this effort.

Sincerely,

James S. Kane

James S. Kane, Acting Director
Office of Basic Energy Sciences

Figure 13. Department of Energy letter. (Personal letter, 16 January 1978)

There was one event, characteristic of efforts to use newly developing technologies, that arose during the UCLA-1 experience. We were fighting deadlines to put together the first draft of the four-volume report, and I recall one night when Terry Gray and I were working into the wee hours to enter some of our material. Our data entry person suddenly said that something was wrong with the system, because material that she was sure she had entered was not there. Gray began some detective work. He discovered that someone else on the network had become so enamored with a new spell-check program that he had collected several report sections, put them through the spell-check, and then put them back. The only problem was that, meanwhile, there had been new entries made, and the correctly spelled material overwrote the more recently edited work. We found it necessary to awaken the well-meaning culprit at home and make him undo the havoc that he had created. This explains why the case history refers to explicit protocols for group editing. Another dramatic scene that will stay forever in my memory occurred the last night of editing, when participants at every ERDA site were reviewing the report document. A small group of us had accepted the task of making final changes. It was a characteristic of the ERDA support structure that everyone's funding came from the same source, and each task force member became particularly tuned to making sure that the report said good things about his or her laboratory and, in some cases, revealed a relative weakness in a

Network Experience Report LBL-7 ERDA Program: High-Energy Physics (EC)

Title: PBAR-P Partial Wave Analysis
Cooperating Sites: LBL, RHEL
LBL Contact: D. Hall (415) 843-2740 x6053 or FTS 451-6053

Resources Accessed:
Experimental data—RHEL, CERN, BNL.
Partial wave generator code—RHEL
Electronic mail, MUGWUMP—RHEL
Cluster analysis code—LBL
Continuity fitting code—LBL
Pole extraction code—RHEL
Graphics display code—RHEL

Motivation

Physicists using the network as a mechanism for maintaining communication discovered that complementary software existed at their respective sites which would allow a highly sophisticated processing chain to be constructed for analysis of proton antiproton annihilations. Transporting either set of software to the other site would have been considerably more effort than merely transferring output at each stage of the process.

Procedure

Data from previous experiments from many groups was formatted and transmitted to RHEL (currently, CERN uses an RJE link and BNL uses air mail, but the potential for data acquisition over the network is obvious). Next, the partial wave generation code at RHEL was used to produce many sets of partial wave coefficients as possible solutions. These coefficient sets were transmitted to LBL (using FTP) for further analysis.

The cluster analysis program at LBL was invoked to separate the possible solutions into several groups. This is an iterative process which requires judgment on the part of the physicist. The separation process is currently underway with physicists from both sites communicating by electronic mail.

When cluster analysis is completed, the results will be passed to a continuity fitting code also at LBL, which will further eliminate unlikely solutions. The results of the continuity fit will then be transferred to RHEL (again by FTP), which will further eliminate unlikely solutions. The results of the continuity fit will then be fed to the RHEL pole-extraction program which will use them as weak constraints to limit the number of possible solutions still further.

These results will then be fed to a graphics display code (MUGWUMP) at RHEL for final checking and publication.

Observed Benefits

Energy-related scientific research and development

Scientific collaboration and communication
Reduction of duplicate software development

Observed Barriers

Insufficient Network Hardware Support

File transfer speed between RHEL and the network is limited by the 1,200-baud telephone link between from the London TIP to RHEL. For this application, file transfers typically require two hours to complete. It is significant that the physicists are sufficiently interested in the results to be willing to come in nights and weekends to perform the transfers when network traffic and the LBL file system are relatively quiescent.

Evaluation

Successful and continuing.

This impressive procedure was initiated by the physicists themselves, which is typical in high-payoff situations. Familiarity with their local command languages and programming languages meant that neither group had to learn the other's system, beyond what had already been learned in establishing initial contact. The fact that slow file transfer speeds were deemed acceptable is further evidence of the importance attached to the application. In general, however, 1,200 baud is too slow for most resource-sharing applications.

competing laboratory. As the deadline approached, there developed a stream of proposed changes that flooded our lines. Somehow the clock ticked away the remaining time, and we were uniformly happy with the result.

Conclusion

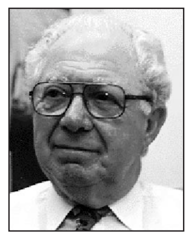
The four-year project launched the energy research community into the exploding world of computer network technology. The researchers were given links to a much larger research community. There had been earlier exploration of resource sharing through the MFENET (which attempted to interconnect computers dedicated to a special application domain) and through SACNET (which was concerned with secure communications). However, piggybacking on the emerging Arpanet allowed testing of new methods that might not have occurred for years and probably spurred the growth of Arpanet by introducing a well-recognized and supported application domain to demonstrate resource-sharing effectiveness. Figure 13 shows the January 1978 letter from the Department of Energy testifying to the significance of this project.

There are at least two other ways in which this early work had a major impact on computing. The energy research community consisted of professionals who faced challenges at the cutting edge of computational capabilities. Defense funds were available to buy the most-advanced computer systems as soon as they were produced; generally, funds permitted buying one at LLL and participating in development of the software systems that put the hardware to work. There was little collaboration between the national laboratories; in fact, there was a great deal of competition to acquire the next computer, because that helped the laboratory to attract funds and researchers. The practice among high-energy physicists was for them to travel to a particle accelerator center and take turns using the equipment. The computer network brought them into a new era, in which they could do much more work remotely, share resources, and then gather face-to-face as needed to discuss their models and experiments. Furthermore, the collaboration demonstrated in the planning, execution, and reporting on experiments in these reports left a legacy for future network-based research. The resource sharing that was documented made it much more likely that an expenditure on new technology would be utilized more fully within the relatively short lifetime of a computer system and that there might be less waste in software development efforts.

This project was a clear example of thinking "outside the box." Rose was head of the Mathematics and Computer Science Research program in ERDA in 1973. I was, then, Principal Investigator for a research program ERDA supported at UCLA. I had also been Principal Investigator of Defense Advanced Research Projects Agency's pre-Arpanet contract at UCLA; therefore, several of the graduate students who worked with me had been and continued to be deeply involved in the creation and growth of the Arpanet. During a discussion with Rose, I suggested that computer scientists and applied mathematicians might significantly increase their impact on the energy research program. If he would authorize a project to explore resource sharing through general-purpose computer networking, I was willing to lead the working group in evaluating its potential. He stuck his neck out, and I think he never regretted doing so. Without his official support and the impact of Rose's small but precious budget, this project would not have succeeded. It suffered from the same kind of problem that caused great difficulty in initial development of the Arpanet. There was no incentive for computer center directors at ERDA national laboratories or center directors at Defense Advanced Research Projects Agency-supported sites to share their facilities with outsiders. In both cases, I believe, the combination of incentive funds and threat of loss of support bought the interest and cooperation of center directors.

Reference

1. *General Purpose Computer Networks and Resource Sharing in ERDA*, 15 July 1977, Working Group on Computer Networking, Mathematical and Computer Sciences (EE-03-04), Office of Engineering, Mathematical and Geo-sciences, Division of Basic Energy Sciences, ERDA. Vol. 1, *Summary, Conclusions and Recommendations*, 26 pp. Vol. 2, *Implementation of Arpanet Facilities*, 55 pp. Vol. 3, *Remote Resource Sharing Experience and Findings*, 141 pp. Vol. 4, *Arpanet Resource Handbook for ERDA Sites*, 247 pp.



Gerald Estrin received the PhD in electrical engineering at the University of Wisconsin, Madison in January 1951. Estrin was a member of John von Neumann's Electronic Computer Group at the Institute for Advanced Study in Princeton, New Jersey, from 1950 to 1956. After completion of the IAS machine, he took leave to direct the building of the WEIZAC computer at the Weizmann Institute of Science in Rehovot, Israel.

In 1956, Estrin became a faculty member at the University of California, Los Angeles, where he built the academic program in computer engineering and led research on reconfigurable computer architecture and methods and tools to support design of computer systems. Estrin was involved in ARPA networking projects starting in 1965 and sparked interest of the energy research community in resource sharing through networks. He served as Chair of the UCLA Computer Science Department from 1979 to 1982 and from 1985 to 1988. Estrin has been recognized as AAAS Fellow, Guggenheim Fellow, IEEE Fellow, and IEEE Computer Society Pioneer.

Direct questions and comments to Gerald Estrin, Professor Emeritus, Computer Science Department, 4731C Boelter Hall, University of California, Los Angeles, CA 90095; estrin@cs.ucla.edu.