Carlson, and Trevor Pinch

n-Support Techniques and

bs: Toward a Theory of

ing Technology/Building Society

Dynamics of Technological Change

mation Management and
0

lge and Intelligent Machines
and the Politics of Discourse in

iscursive Politics of Genetic

ear Power and National Identity

isual Representations, Visual
eering
ation of Engineering Culture,
ce and America
Construction of the Landsat

orical Sociology of Nuclear

on Technical Change
ordinating Technology: Studies in
ations

Inventing the Internet

Janet Abbate
(1999)

The MIT Press
Cambridge, Massachusetts
London, England

3

"The Most Neglected Element": Users Transform the ARPANET

In light of the popularity of the Internet in the 1990s, we might expect that the ARPANET's first users would have quickly embraced the new technology. In practice, however, users did not move their research activities onto the network automatically or easily, and the results of such efforts were uneven. A number of diverse groups did make productive use of the ARPANET in the early 1970s, but other potential users were excluded or discouraged from using it, and many of ARPA's original predictions about how the network would benefit its users turned out to be wrong. The fact that the network became so successful is not something to be taken for granted, but rather something to be explained.

Historians have begun to call attention to the role of users in determining the features and ultimate success of a technology.[1] Typically, users are portrayed as consumers acting through the market, choosing one product or service over another. Occasionally, they are portrayed as concerned citizens pressing for regulations (e.g., safety standards). In any case, it is generally assumed that users become involved only after a technology has already been developed. But the ARPANET's ultimate "consumers"—the researchers who were to use it in their work—were directly involved in its development. During the ARPANET's first decade of operation, fundamental changes in hardware, software, configuration, and applications were initiated by users or were made in response to users' complaints or suggestions. It was, arguably, these activities that accounted for the perceived success of the system by ensuring that the ARPANET provided the types of services that users actually wanted. The ARPANET provides an instructive example of the variety of active roles users can play in shaping a new technology, and of the sometimes surprising results of their involvement.

"By No Means Complete or Perfect": The ARPANET as Experienced by Early Users

Many commentators on the popularized Internet of the 1990s have celebrated the advent of "cyberspace," the virtual realm in which people interact with computers and with other computer users.[2] Cyberspace provides an opportunity for individuals to create and explore imaginary environments, to experiment with different identities, and to establish new forms of community. Computer networks provide access to cyberspace, which appears as a welcoming, even playful environment in which newcomers receive instruction and encouragement from their fellow users.

The conditions encountered by the ARPANET users of the early 1970s stand in stark contrast to this rosy picture. Using the network and its host computers was difficult, the support systems were inadequate, and there was little opportunity to interact with other users. Michael Hart, one of the few early users from outside the field of computer science, later recalled that there was little on the net in the 1970s to attract users who weren't "computer geeks":

> You have to realize how FEW people were on the Net before the '80s. . . . There just weren't enough to support a conversation on any but the most geeky or the most general topics. . . . It was boring, unless you could "see" down the cables to the rest of the world . . . and into the future. (Michael S. Hart, email to author, 28 March 1997)

There *was* a sense of community among many of the ARPANET's users, but it predated the network and was based on their shared backgrounds, interests, and offline experiences. One challenge in making the ARPANET user friendly lay in translating activities that build community—sharing of information, support, recreation—to the network environment. In taking these steps for the first time, early users of the ARPANET laid the groundwork for future virtual communities.

The road to becoming an active ARPANET user was long and hard. The first challenge for any potential user was getting access to the network. In order for a site to get an ARPANET connection, someone there had to have a research contract with ARPA (or with another government agency approved by ARPA). A prospective network member who was not being funded by ARPA had to pay the cost of setting up their node, estimated in 1972 to be somewhere between $55,000 and $107,000 (RCA Service Company 1972, p. A-72).[3] Once a site was approved, ARPA had to order a new IMP or TIP from Bolt, Beranek and Newman, direct the Network Analysis Corporation to reconfigure the network to include the new node, and arrange with AT&T for a telephone link between the new node and the rest of the ARPANET (ibid., p. A-81). The new host site would be responsible for providing the hardware and software for the host-IMP interface and for implementing the host protocol, NCP, on its computer(s)—a task that might represent a year's work for a programmer. In short, adding a new site to the network was complicated and costly. A prospective site's access was limited by its ability to pay, by the need to belong to an ARPA-affiliated research group, and by the need to have expert programmers available to create and maintain the host software.

Once a site was connected to the ARPANET, though, access controls were much looser. In theory, access within each site was to be limited to individuals doing work for ARPA. In practice, few sites tried to enforce that policy. Once a university or a company had connected a computer to the network, anyone with an account on that computer (or access to a friend's account) could use network applications such as email and ftp simply by executing the proper commands. Often, sites even included these unofficial users in the listings they submitted to the NIC "white pages," an online directory of ARPANET users (McKenzie 1997). Few system administrators tried to add access restrictions to the network commands. According to BBN's ARPANET Completion Report, "despite a deeply ingrained government and Defense Department worry about unauthorized use of government facilities, it was possible to build the ARPANET without complex administrative control over access or complex login procedures or complex accounting of exactly who was using the net for what" (Heart et al. 1978, p. III-111). BBN argued that this relaxed access policy made the system simpler and thus contributed to its quick and successful completion.

Many members of the ARPANET community suspected that ARPA managers were aware that unsanctioned users were on the network and did not object. Unauthorized users who contributed improvements to the system may even have received tacit encouragement from ARPA. In the early years the ARPANET was underutilized, and ARPA had little reason to discourage users or activities that might make the network more popular. Increased use of the network would also make it easier for ARPA's computer scientists to evaluate the system's performance. In fact, a recreational mailing list for "science fiction lovers" was apparently allowed to operate over the ARPANET on the ground that it generated significant amounts of traffic and therefore provided an opportunity to observe the network's behavior under load (McKenzie 1997). Another unofficial but tolerated activity was Michael Hart's Project Gutenberg, an effort to make historically significant documents available over the network. Hart, who was not an ARPA researcher but who had acquired an account at the University of Illinois, began by posting the Declaration of Independence on his site's computer in December of 1971; Project Gutenberg was still in operation on the Internet 25 years later.

Once on the network, users theoretically had access to some of the most advanced computer systems in the United States; however, using those remote systems could be difficult, impractical, or unappealing.[4] For one thing, new sites were provided with only scattered and incomplete resources to get them started. A 1972 report by an outside consultant stated: "The network user, new and established, is probably the most neglected element within the present development atmosphere. The mechanisms for assisting and encouraging new members are relatively informal or nonexistent." (RCA Service Company 1972, p. 9) ARPA provided an initial briefing to prospective members. Each site was given some printed documentation, including protocol specifications, a Resource Notebook containing descriptions of resources available at various sites, and a directory of participants (ibid., p. A-79). Some additional information was available online at the Network Information Center, located at the Stanford Research Institute. Beyond that, new sites had to find help where they could. Typically, they turned to BBN or to more experienced host sites for advice. ARPA did not provide in-depth training, and there was no single source to which users could turn for help in setting up network operations and locating resources (ibid., p. 29).

Just finding out what was available on the ARPANET could be difficult. The network search tools that Internet and World Wide Web users would later take for granted did not yet exist. Many sites did not provide complete or up-to-date information for the Resource Notebook, nor did sites generally offer online consultation about their resources, so users had to contact the sites offline to find out what services might be available (RCA Service Company 1972, p. A-21). Throughout the ARPANET's existence, its managers struggled to get host administrators to provide adequate information about their computer resources, technical configurations, and users.[5] It is not clear whether this was the case because paperwork was a low priority for computer scientists, or because the information was difficult for them to pin down (since site configurations were continually changing), or because they did not want to make this information available (perhaps fearing a loss of local control). What is clear is that the difficulty of learning about host resources was a major obstacle for new users. And there was no guarantee that host machines or resources, once located, would continue to be available: a research site might amass data for a particular project and then remove it when the project was completed, or it might temporarily take its machines off the network without notifying remote users (ibid., p. A-76).

There seemed to be general agreement among users that the Network Information Center, which was supposed to provide network information and a means for users to interact, was not working as planned. The Network Information Center did have some successes. Notably, the Network Working Group's protocol developers used the NIC's text editing and bulletin board systems to prepare, distribute, and store Requests For Comments. RFCs proved to be a very effective way for a large group to participate in ongoing technical discussions—in large part because members of the NWG (in contrast with many host administrators) were highly motivated to make information available to their scattered collaborators.[6] But as a directory of network resources, the NIC fell short, both because sites failed to supply information on their resources and because many people found the software at the NIC difficult to use (RCA Service Company 1972, p. A-9). To help fill the void, in March of 1973 the Stanford Research Institute began publishing the *ARPANET News*, a newsletter that listed updated information on host resources. The *ARPANET News* reduced (but did not eliminate) the difficulty of locating resources (Hafner and Lyon 1996, p. 229).

Users who had managed to identify an attractive remote resource faced another obstacle in the lack of administrative mechanisms for arranging the remote use of computers. Many host administrators wanted to charge remote users for computer time, or at least to know who those users were. This meant that users had to contact someone at the remote site to set up a computer account, and if they were going to be charged for their usage they also had to obtain a purchase order from their local institution. These interactions almost always had to take place off line, since few sites were prepared to conduct such business over the network. In addition to the extra administrative burden on the user, it was difficult for many researchers to get approval to spend computing funds at other sites rather than at their institution's own computer center. And since all these arrangements had to be made before the researcher ever got access to the remote resource, a potential user had to weigh a definite cost against an unknown benefit.

Connections from user sites to the network were not always satisfactory. Users at Wright-Patterson Air Force Base, which used a TIP to access the ARPANET, found this arrangement inadequate, since they had to use noisy telephone links to get to the TIP and since phone calls going through the base's switchboard were limited to 5 minutes' duration (Lycos 1975, pp. 161–162). Other TIP users may have experienced similar connection problems. A 1975 report by the US Geological Service, which had set up a conferencing system using the ARPANET, noted that "a major drawback of the early system was the unreliability of the experimental computer network we were using. Access was limited, and frequent hardware failures made 'real' work all but impossible." (ibid., p. 59) As a result, the USGS switched from using the ARPANET to using the commercial networks Telenet and TYMNET (Turoff and Hiltz 1977, p. 60). That these sorts of difficulties were not necessarily the fault of the ARPANET itself must have been little comfort to users who found themselves unable to communicate over the network.

Another technical obstacle was incompatibility between computers at different sites. Many of the hosts were unique systems with their own command languages and data formats, and some required specialized hardware at the user end. And, except for the few sites that hoped to generate income from network users, there was no great incentive for host sites to adapt their systems for remote use by others; thus, users were often left to deal with incompatibilities as best they could. Compatibility problems proved much more difficult to resolve than anyone working on the ARPANET seems to have expected.[7]

A user who had surmounted all these obstacles still had to figure out how to operate the remote computer. The instructions provided for the ARPANET demonstration at the International Conference on Computer Communications in 1972 convey some idea of how complicated using the network and its computers could be. The instruction booklet, entitled Scenarios for Using the ARPANET at the ICCC (Anonymous 1972), began with a disclaimer that "the scenarios are by no means complete or perfect." It urged participants to "approach the ARPANET aggressively" and to "unhesitatingly call upon the ICCC Special Project People for the advice and encouragement you are sure to need." The detailed instructions for each scenario illustrated the many steps a user had to go through in order to connect a terminal to a TIP, tell the TIP which host to connect to, instruct the host computer how to handle the type of terminal being used, log in to the host computer, and run the desired application. Those who were already expert in various computer systems and ready to "approach the ARPANET aggressively" might find the network easy and exciting to use; for others, mastering the ARPANET must have been an uphill struggle.

Where could the novice user turn for help? Since most of the software available on the ARPANET had been developed as part of some local research project rather than as a commercial product, instruction and support tended to rely on informal local interactions. According to Alex McKenzie (email to author, 26 March 1997):

> I don't think that any of the hosts were all that easy to use if you weren't part of the computer community. Most of the hosts were operated by researchers and tended to change frequently. Although every site had substantial documentation, it tended to *not* be tutorial in nature. You learned to use a host by talking to the other users down the hall. With "the hall" extended to intercity distances, it wasn't even easy for computer scientists to learn how to use a remote system, much less for other communities to do so.

However, experience with the ARPANET caused managers of host systems to reevaluate how they provided user support. Because none of the sites had served remote users before joining the ARPANET, the typical site's modes of support—such as system updates posted on bulletin boards or face-to-face interactions between users and support staff—had implicitly relied on users' having physical access to its computer center. In response to requests for help from ARPANET users, some sites began supplementing or replacing older means of support with online documentation, system announcements sent to each user's terminal, the ability to query support staff by email, and/or methods for "linking" terminals so that the user and the remote system operator would see the same screen output and could work through a problem together (Heart et al. 1977, p. III-8). In addition, most sites provided telephone consultation (which was toll-free at the NIC), and ARPA asked each site to designate a system specialist who would be available to answer users' questions.

Host sites' attitudes toward adopting these new methods for helping network users varied widely, depending on whether they saw outside users as a source of income or as a drain on local users' resources. Computer users at UCLA complained that they were not welcomed at other host sites: "Computer operations managers at other nodes may feel that incoming traffic is disruptive, less important than their own needs, or that UCLA's use of the net should be shunted to slack hours." (Brinton 1971, p. 65) On the other hand, sites that wanted to attract remote customers so as to generate income for their computer centers were eager to adopt new support techniques. For remote users of these sites, the network might indeed seem to be a welcoming place. But a significant number of the ARPANET's users were not satisfied with the services they were offered, and they began to take matters into their own hands.

Improving the System: User Activism

The ARPANET created an environment of both frustration and opportunity for its users. Using the network could be difficult, but a person with skill and determination (and there were many of these in the ARPANET community) could devise new applications with few restrictions. Thus, users had both the incentive and the ability to experiment with the system to make it better meet their needs. In some cases users built new hardware or software for the network, or asked ARPA or BBN to do so. In other instances, users improvised new ways of using the existing infrastructure. Users also began to organize to press for more support from ARPA—an activity that exposed tensions between segments of the ARPANET community. Three aspects of the system that users' experiments affected noticeably were terminal interfaces, connection paths between hosts, and applications protocols.

Terminal Interface Systems

ARPANET users had mixed relations with Bolt, Beranek and Newman, the contractor responsible for building and operating the network's infrastructure. The BBN team took great pains to respond to trouble reports and keep the network in constant operation, and they made a number of improvements to the interface message processor. However, they were more reluctant to respond to demands for new features and services, especially when these threatened to increase their own management tasks.

The BBN team made a number of changes from the original IMP specification as sites expressed their desire to use the IMPs in new ways (McKenzie 1997). The system was originally designed to have one host computer per IMP; however, users at UCLA—the very first site on the ARPANET—wanted to attach two computers to their IMP. To accommodate UCLA and other sites with multiple computers, Lawrence Roberts directed BBN to modify the IMP to handle more than one host. This paved the way for further innovations. Since a site's host computers were not always in the same location, multi-host IMPs tended to require longer connections to some of the hosts; since longer connections were more likely to suffer from line errors, BBN had to add error-checking procedures to the host-IMP interface. Users at UC Santa Barbara pushed the capabilities of the host-IMP interface to the limit when they requested a five-mile connection between a new host and their IMP so as to avoid having to install a second IMP. To accommodate this demand, BBN came up with a new interface, called VDH for "very distant host," that provided even more error checking.

The biggest innovation in the node design was to provide a way to connect terminals directly to the ARPANET, rather than expecting all terminals to be connected through host computers. Roberts decided in 1971 that such a terminal interface was needed in order to allow sites without host computers (such as ARPA itself) to connect to the network. This new type of node was called the terminal IMP (abbreviated "TIP").

The TIP greatly increased the numbers of sites and users that could access the ARPANET. However, many users were dissatisfied with the TIP interface, which represented a compromise between BBN's need for simplicity (in order to create a reliable system quickly) and the various requirements of terminal users. The most common complaint was that the TIP would handle only terminals for interactive computers ("asynchronous" terminals), while many people wanted to use terminals designed for batch processing computers ("synchronous" terminals). Some users also wanted to be able to program the TIPs to perform customized functions, such as reading files from magnetic tapes; they urged BBN to add new features to the TIP, or to allow users to modify their TIPs. Frank Heart, the head of the IMP group, commented: "Unfortunately, but perhaps not surprisingly, the limited goal and absolute restriction on user programming created considerable unhappiness in portions of the potential user community, and created considerable pressure for other 'better' terminal access techniques." (Heart et al. 1978, p. III-117)

From the perspective of Heart and the rest of the BBN team, these demands were not reasonable. The TIP was a small computer with no room for extra programs. The designers did not want to add additional disk storage, because this would make the TIP more liable to failure. Nor did the BBN team want to allow users to modify their TIPs: BBN was responsible for maintaining and upgrading the machines, a task that would be much harder if users were to make non-standard, possibly damaging changes. The BBN group felt that users simply did not understand how difficult it would be to provide user programming options, broad terminal support, and other special services.

Roberts had mandated some of the earlier changes in the IMP; however, perhaps swayed by Heart's arguments, he did not oblige BBN to make sweeping changes to the TIP. Instead, he gave financial support to users who wanted to develop alternative terminal interface systems.

Even before BBN had begun providing TIPs, some users had taken the initiative to build their own terminal interface machines. Such efforts continued after the TIP (with its perceived shortcomings) became available. In the first project of this kind, at the University of Illinois Center for Advanced Computation, W. J. Bouknight, G. R. Grossman, and D. M. Grothe designed a system they called the ARPA Network Terminal System (ANTS). The ANTS effort began in the summer of 1970, predating the introduction of the TIP. ANTS ran on a DEC PDP-11 minicomputer and was meant to provide an interface between an IMP and local terminals and modems. The system accommodated a wider range of terminals than the TIP, including the synchronous terminals typically used with batch machines. ANTS also allowed users to access other types of peripherals not supported by the TIP: users could read in data from punch cards, disks, or tapes for transmission across the network, and incoming data could be stored on these media or sent to local printers. Unlike the TIP, ANTS had disk storage, so users could keep files on their local ANTS system rather than on a distant host computer. ANTS also provided mechanisms for network access control and accounting.

The Illinois team put the ANTS system in place as soon as they received their IMP in April of 1971, and by September of 1971 ANTS was providing local terminal service. In the autumn of 1972 they added support for graphics terminals (Bouknight, Grossman, and Grothe [illegible]). The ANTS system gave the Illinois users [illegible]

complexity hindered transfer of the technology, and only a few sites ended up using it.

Another attempt to improve the ARPANET's terminal interface came from David Retz of the Speech Communications Research Lab at Santa Barbara. Retz had created a real-time data acquisition system for his ARPA-funded speech project, and he wanted to be able to send that data across the ARPANET for processing. The TIP, designed to handle commands typed from a terminal, was not equipped for such data transfers. Learning from the fate of ANTS, Retz and his colleagues decided to build a less ambitious terminal interface. Early in 1973 they began developing a system called ELF,[8] which also ran on a DEC PDP-11. Like ANTS, ELF allowed users to input and store files using local peripherals. ELF was more successful than ANTS, in part because it was simpler but also because its developers took advantage of the ARPANET to transfer the system to other sites. The ELF team would send the source code and binary files over the ARPANET to the target machine, perform remote debugging via the network, and keep in touch with remote ELF users through online release notes and bug reports (Retz and Schafer 1976, pp. 1012–1014). The initial ELF system was in experimental use at Santa Barbara in early 1974, and by 1976 about thirty ARPANET sites were using the ELF interface (ibid., p. 1007).

New Communications Paths

Members of the ARPANET community also found unexpected ways to use the network's communications links. The ARPANET was designed to connect distant computer centers, but users soon found a new application: sending data between computers at the same site. Local-area networks, which became ubiquitous in the 1980s, did not exist in the early 1970s; some manufacturers offered systems for networking their own line of computers, but there were no products for connecting different types of computers. Instead, users had to copy data onto tapes or other media and carry them from one local computer to the other. ARPA, concerned with the distribution of computing resources among different sites, had not focused on the networking of local computers, but users were very aware of the inconvenience of local data transfers.

When MIT's IMP was installed, in June of 1970, the ARPANET users there quickly realized that they could use the network to speed up communications between their local machines. No one had envisioned such a use of the ARPANET, and BBN's network monitors were puzzled when they started to notice that there was heavy traffic at the MIT IMP but not over MIT's outgoing lines. Eventually they realized that the MIT users had, in effect, turned their IMP into the hub of a local-area network (LAN). Soon other sites, among them Stanford University and the University of Southern California's Information Sciences Institute, began using the ARPANET as a LAN, and BBN itself began employing its IMP for local functions such as backups (McKenzie 1997). According to BBN's Frank Heart (1990), "the notion of using the IMP as a local connection was quite a surprise, to the extent that it became just common and had not been envisaged." By 1975 almost 30 percent of ARPANET traffic was intra-node (Heart et al. 1978, pp. III-77, III-91). A spontaneous innovation by users had contributed substantially to the use of the ARPANET and hence to its perceived value. Sites continued to use the ARPANET as a local-area network until "real" LANs based on Ethernet technology became available in the 1980s (McKenzie 1997).

Some users also created unusual or even illicit links between the ARPANET and other data communications systems. After the ARPANET was extended to England, physicists at the University of Illinois began using it to reach the Rutherford high-energy physics computer at Cambridge University. Rutherford had a separate connection to the Centre Européenne pour la Recherche Nucléaire (CERN) in Geneva through a European telecommunications carrier, but the carrier's regulations prohibited ARPA from using the Rutherford-CERN link to make connections from the United States to CERN. The Illinois physicists got around this restriction by using Rutherford as a dropoff point for files; by sending files or email between Illinois and CERN with a brief stop at Rutherford, they followed the letter of the law but were still able to create their own "virtual link" between the Illinois campus and the Geneva lab. Others used commercial networks to reach ARPANET sites. For example, John Day telecommuted from his home in Houston to his computer account at Illinois by setting up a connection from Houston to MIT through the commercial network Telenet, logging in to the MIT machine, and then going from MIT to Illinois through the ARPANET (John Day, telephone conversation with author, 11 April 1997). In both of these cases, the users needed to have accounts on machines at intermediate points (Rutherford, MIT) and needed to know how to use both networks, so making an

inter-network connection required skill, resources, and motivation. Such "virtual internets" provide another illustration of how resourceful users extended the capabilities of the ARPANET.

New Applications

Although many improvisations by users were encouraged or at least tolerated by ARPA, the agency did not always welcome users' attempts to steer the development of the system. This became painfully apparent when a group of user advocates tried to speed the development of upper-level protocols and applications for the ARPANET. In November of 1973, as enthusiasm for the network was beginning to grow among the ARPANET community, a group of systems developers who wanted to improve network services formed the Users Interest Working Group (USING). Members of this group began to critique the difficulty of using the network, and they lobbied ARPA to support the development of more and better applications. They also tried to create common tools for tasks such as accounting and editing; for example, they promoted a standard network editor called NetEd that was widely adopted.

Despite some initial support from ARPA, however, USING faced criticism when it tried to draw up a blueprint for the further development of the ARPANET's user services. Faced with organized action by users, the ARPA managers were evidently afraid that the network might slip out of their control. Members of USING were dissuaded from pushing their demands by ARPA program manager Craig Fields, who made it clear that the authority to make plans for the network lay with ARPA, not with USING. Early in 1974, Lawrence Roberts cut off funding for the development of upper-level protocols (Hafner and Lyon 1996, p. 230; John Day, telephone conversation with author, 11 April 1997).

The fate of USING revealed the limits of ARPA's generally nonhierarchical management approach. Individual users or research teams had tacit or explicit permission to add hardware and software to the system; ARPA even gave financial support for some of these experiments. However, users as a group had no say in the design decisions or funding priorities of the ARPANET project. The ARPANET experience is a reminder that the efforts of individuals to build virtual communities are constrained by the realities of money and power that support the infrastructure of cyberspace. ARPANET users continued to work for improved network applications, but after the demise of USING they focused on the more neutral activities of technical development and information sharing rather than on organized lobbying.

Rethinking the ARPANET's Purpose: Successes and Failures of Resource Sharing

When Lawrence Roberts described his original plan for the ARPANET, the goal he promoted was resource sharing: allowing individuals at different sites to share hardware, software, and data. The first published description of the ARPANET, co-authored by Roberts and his assistant Barry Wessler in 1970 and entitled "Computer Network Development to Achieve Resource Sharing," described the rationale for building the ARPANET as follows:

> Currently, each computer center in the country is forced to recreate all of the software and data files it wishes to utilize. In many cases this involves complete reprogramming of software or reformatting the data files. This duplication is extremely costly. . . . With a successful network, the core problem of sharing resources would be severely reduced. (Roberts and Wessler 1970, p. 543)

The resource-sharing ideal was similar to the vision of a "computer utility" that was popular at the time.[9] Both models assumed that users would be accessing large, centralized machines (analogous to the generating plants of an electric power utility), with the network acting as a distribution system for computing power. The goal of resource sharing was partially fulfilled by the ARPANET: some sites did provide remote services to a significant number of ARPANET users. But, as we have seen, there were many obstacles to providing computer services in a way that was convenient for distant users. Many of the sites that succeeded had administrators who were strongly motivated to sell networked services; others had financial backing from ARPA to build computer systems that were specially adapted for use over the network. Lacking these incentives, most sites did not invest the effort needed to make their computers easy to use from afar. Overall, therefore, the practice of resource sharing on the ARPANET fell far short of Roberts's expectations.

Roberts had envisaged that the ARPANET would be used mainly to access time sharing computers, and his design specifications were aimed at supporting interactive computer use. As it turned out, however, only a few time sharing systems seem to have had significant numbers of remote users. One of these was the MULTICS operating system, created by MIT's project MAC in the 1960s, which featured a popular mathematics program called MACSYMA. By 1976 an estimated 15–20 percent of the MULTICS computer's load came from remote ARPANET users (Day 1997; Heart et al. 1977, pp. III-24–III-25). This percentage, though significant, was less than expected, in part because the managers of the MULTICS system did not welcome outside users.[10]

Another popular time sharing system was TENEX, which had been developed at Bolt, Beranek and Newman for the DEC PDP-10. BBN and USC's Information Sciences Institute were the main providers of TENEX service to remote ARPANET users; BBN offered its service on a commercial basis to ARPANET users, while the ISI was subsidized by the US government and handled jobs from government agencies (Roberts 1974, p. 47; see also Heart et al. 1977, pp. III-18–III-19). TENEX machines were particularly popular for text processing and artificial intelligence programming in the LISP language. Howard Frank's group at the Network Analysis Company used a remote TENEX machine to replace their old batch processing programs with interactive programs. Frank (1990) reported that access to the TENEX system sped up the planning of changes to the ARPANET topology: "We were able to create the same design in a day that was taking us two weeks to get before that."

Though ARPA favored time sharing, the ARPANET also offered the services of batch processing computers. Two sites provided access to large IBM batch processing machines. One was UC Santa Barbara, which had an IBM 360/75 that was mainly used for image processing. The other was UCLA, which had an IBM 360/91; the largest and fastest general-purpose computer on the network for many years, it served as ARPA's main "number cruncher" (Heart et al. 1977, pp. III-9, III-17).

UCLA had been entrepreneurial in putting the 360/91 on the network. In 1969 it had approached ARPA with a request for funding to build an interface between its computer and the ARPANET (UCLA Campus Computing Network 1974, p. 4). For UCLA's computer managers, who were financially responsible for an expensive and underutilized machine, the network connection would provide a way to sell extra computing capacity; for ARPA, the arrangement would provide its researchers with access to a high-end computer at competitive rates. Since the UCLA managers were eager to sell computer time, they made special efforts to meet the needs of remote users by providing expanded telephone support, a consultant dedicated to helping remote users, a Users' Manual that was stored on disk and could be printed out at a remote site, modifications in IBM's operating system to help keep local operators aware of the status of remote users' programs, and a "status" command that allowed users to follow the progress of their jobs through the system. UCLA's commitment to service attracted remote users, and within a few years ARPANET users were providing 10–20 percent of the UCLA computer center's income (Roberts 1974; Greenberger et al. 1973, p. 30).

Meanwhile, ARPA was building its own number-crunching machine: ILLIAC IV, a supercomputer with 64 parallel processors and a large memory. The ILLIAC IV project began at the University of Illinois, but was eventually transferred to NASA's Ames Research Center at Moffett Field in California. Since it had been built as an experiment in computer design rather than to serve a particular computing need, ILLIAC IV represented a solution looking for a problem. ARPA managers hoped that putting ILLIAC IV on the network would encourage researchers to find applications for the new machine. Remote ARPANET users at Rand, NASA, and other sites did use ILLIAC IV to run large-scale computations needed for climate simulations, signal processing, seismic research, and physics calculations (Heart et al. 1977, p. III-28); however, many of these projects had been created to find employment for ILLIAC IV, not to meet the pre-existing needs of users.

One large resource that had been explicitly designed for the ARPANET was the Datacomputer, a database system located at the Computer Corporation of America in Cambridge, Massachusetts. The Datacomputer had been designed by a team led by Thomas Marill, who had worked on early network experiments with Lawrence Roberts. Marill was a staunch believer in the resource sharing model. He held that in a networked environment host sites would tend to become specialized to take advantage of economies of scale, and that the availability of these specialized resources would eventually make the general-purpose computer obsolete. The Datacomputer was meant to be an example of the specialized resources that would dominate the future of computing (Heart et al. 1977, pp. III-29–III-30). It consisted of a DEC PDP-10 computer with a 3-trillion-bit storage device and with programs for storing, organizing, and retrieving very large amounts of data. To accommodate the needs of network users, the system [illegible] facilities so that [illegible] transfer

data sets between different types of computers (Dorin and Eastlake 1976). The Datacomputer was heavily used by ARPA's seismic researchers, and the Argonne National Laboratory made it the repository of a climatological database. The Network Control Center stored statistical data on the performance of the IMP subnet on the Datacomputer, and MIT used it to store information it collected on ARPANET hosts (Heart et al. 1977, pp. III-35–III-36).

While some sites specialized in serving remote users, other sites became consumers of network-based services. Roberts pointed out that several sites, such as the University of Illinois Center for Advanced Computation (CAC), were able to dispense with local time sharing machines altogether and to contract for basic computing services from remote sites. Because the CAC's projects required diverse computer resources, the ARPANET was used to access several different types of machines—in particular, a Burroughs 6700 at San Diego. Starting in August of 1972, the CAC got over 90 percent of its computing services over the ARPANET, at about 40 percent of the cost of the former local operation (Sher 1974, pp. 56–57). At first, it should be noted, the Illinois researchers had not been enthusiastic about switching from local to remote computers. Michael Sher, the associate director of the CAC, acknowledged that the Illinois users had had to adjust to the unfamiliar practice of remote computing:

> There was a great deal of skepticism among the center's programmers regarding their ability to develop systems and perform sophisticated applications programming over a network. Absolute control of computer resources, no matter what their quality, is normally not relinquished without significant reservations. (ibid., p. 57)

However, after being forced by economics to give up their local machine, the researchers found that networked computing had some unexpected advantages. Programmers could choose from diverse machines offering a range of services, including time sharing, fast calculation, and graphics routines. System developers were able to find remote users to test out experimental software (ibid., p. 58). In addition, with their newfound network expertise, the University of Illinois users were able to help their colleagues gain access to the network. As a result, Illinois researchers who were not funded by ARPA but who could suggest some defense-related application for their work were able to arrange to use the network for projects in economics, physics, and land use planning (John Day, telephone conversation with author, 11 April 1997).

Who Benefited?

Those who most readily benefited from the ARPANET were, not surprisingly, ARPA's computer scientists, who mainly used the network to trade files and information. Computer scientists had the expertise to use the system, and there were enough of them involved in the ARPANET project to form a community. And networking itself was a popular topic of research: one important ongoing activity was experimental research on topics such as switch design, protocols, and queuing theory. In addition to being a communications tool, then, the ARPANET was a source of empirical data and a test bed for new techniques.[11]

The ARPANET changed the way computer scientists worked and the types of projects that were feasible. Some collaborative projects, such as the development of the Common LISP programming language, would not have been possible without a means for extensive ongoing communication between many geographically separated groups (Sproull and Kiesler 1991, pp. 11, 32; Heart et al. 1977). In a 1986 *Science* article, several computer scientists noted: "The major lesson from the ARPANET experience is that information sharing is a key benefit of computer networking. Indeed it may be argued that many major advances in computer systems and artificial intelligence are the direct result of the enhanced collaboration made possible by ARPANET." (Jennings et al. 1986, p. 945) The National Institutes of Health sponsored a project at Stanford University to develop AI applications for medicine. This project supported AI studies at Rutgers, a medical diagnosis system at Pittsburgh, a database on eye disease at the University of Illinois, a distributed clinical database at the University of Hawaii, an effort at UCLA to model paranoid thought processes, and several activities at Stanford, including protein crystallography, an expert system for treating bacterial infections, and a program to aid chemists in determining molecular structures (Lycos 1975, pp. 193–197). Joshua Lederberg, a geneticist who had been one of the project's strongest advocates, described it as "one of the early 'collaboratories' enabled by the ARPANET" (message to Community Memory mailing list, 26 March 1997). In a 1978 article, Lederberg noted:

> Such a resource offers scientists both a significant economic advantage in sharing expensive instrumentation and a greater opportunity to share ideas about their research. This is especially timely in computer science, a field whose intellectual and technological complexity tends to nurture relatively isolated research groups.

Computer scientists also used the ARPANET to share software and other files. Most collaborative projects involved the transfer of files containing documents or programs. A procedure for anonymous file transfer, implemented early on, made it possible to leave files in a "guest" account for anyone who wished to retrieve them. This allowed files to be exchanged informally, even without the originator's knowledge. The Stanford computer scientist Les Earnest recalled:

> Another thing that happened a lot in the 1970s was benign theft of software. We didn't protect our files and found that both programs and data migrated around the net rather quickly, to the benefit of all. For example, I brought the first spelling checker into existence around 1966 but it wasn't picked up by anyone else, whereas the improved version (around 1971) quickly spread via ARPAnet throughout the world. (email to author, 28 March 1997)

In fact, in the 1970s a number of computer scientists had the impression that they and their colleagues were the *only* users of the ARPANET. David Farber, who was at Irvine, remembers only computer scientists being on the net (conversation with author, 22 February 1996). Les Earnest recalls that "nearly all ARPAnet participants in the early 1970s were computernicks. . . . I believe that there was very little academic or development activity outside of the realm of computer science" (email to author, 28 March 1997). A 1974 publication by UCLA describes the ARPANET as changing, between mid 1971 and mid 1974, "from a system programmer's experiment to an application programmer's tool" (UCLA Campus Computing Network 1974, p. 7)—hardly a move beyond the computer science community. According to these sources, computer experts dominated the network either because no one else was interested or because it was so difficult to use remote computers. Though there was some truth to this perception, use of the ARPANET did go beyond computing researchers.

Lawrence Roberts had announced in 1970 that the network would be used to support ARPA's researchers in behavioral science, climate dynamics, and seismology—researchers for whom computers were a tool, not a research focus (Roberts and Wessler 1970, p. 548). The man largely responsible for making this claim a reality was Stephen Lukasik, who in the early 1970s was both ARPA's director and the head of its seismology program. An early convert to the virtues of networking, Lukasik was (unlike Lawrence Roberts or Robert Taylor) involved in research areas besides computer science, and he sought out opportunities for ARPA's various contractors to work together. ARPA's use of the network for defense-oriented climate and seismic studies is a reminder that the ARPANET, though built by civilian research groups, was serving military needs from an early date.

ARPA's climate research program was one of the first to make serious use of the ARPANET. Predicting the weather has always been an important element in military planning; ideally, commanders would like to know seasonal weather conditions months in advance when planning attacks and invasions. Because of the chaotic nature of weather systems, however, such predictions are very difficult to make; to be feasible at all, they require very fast computers. Lukasik believed that climate modeling would be just the kind of data-intensive project that could provide a useful test of the ILLIAC IV supercomputer while also serving military needs. In the early 1970s he initiated a research program on global atmospheric circulation models that involved the Air Force, the Rand Corporation, the National Weather Bureau, the National Center for Atmospheric Research, Princeton University, and the Laboratory for Atmospheric Research at the University of Illinois.[12] In a typical use of the network, Illinois researchers would run large-scale hydrodynamic and meteorological simulations on the IBM 360/91 at UCLA, then send the output to the Information Sciences Institute (which had facilities for generating graphics), and finally send the graphics files back to Illinois, where the output would be displayed by local plotters. Climate researchers made extensive use of the ARPANET for remote job entry, file transfer, and interactive computing, and they reported that having access to the network allowed faster research and more efficient use of programmers' time (Sher 1974, p. 57; UCLA Campus Computing Network 1974, pp. 4–5; Heart et al. 1977, pp. III-17, III-54).

The seismology program—the other ARPA program that used the network in the early 1970s—was initially aimed at developing techniques for detecting tests of nuclear weapons in order to support a possible US-USSR treaty banning such tests. One sticking point to reaching agreement on such a treaty was the USSR's opposition to on-site inspection, the only known method of verification. Seismology seemed to offer a way out of this dilemma: if underground tests could be detected by seismic sensors located outside the USSR, it might be possible to provide verification of the treaty's terms without on-site inspection.

Detecting, locating, and identifying seismological events required large-scale data processing. To gather the raw data, hundreds of seismometers were arranged in arrays. ARPA had two of these arrays: the Large Aperture Seismic Array in Montana and the NORSAR array in Norway. Monitors at each array would collect data and send it to ARPA's Seismic Data Analysis Center in Alexandria, Virginia, where analysts would use signal processing to look for patterns that would indicate likely seismic events and to characterize these events in terms of time, location, depth, magnitude, and probable cause (earthquake or explosion).[13] Before the ARPANET, magnetic tapes containing the seismic data had to be mailed to the SDAC; that made it impossible to examine seismic events in real time. Lukasik arranged to connect the Montana and Norway sites to the ARPANET (the latter through a satellite link), which allowed the SDAC to begin analyzing data within hours—rather than weeks—of an event. Since files of seismic data could be very large, the SDAC was also an ideal test user of the Datacomputer. By 1976, ARPA's seismic analysts had stored 70 billion bits of seismic data in the Datacomputer (Heart et al. 1977, p. III-35).

In addition to ARPA's own researchers, other government-funded scientists experimented with using the ARPANET. Physicists at several universities used it to access powerful computers elsewhere, such as UCLA's IBM 360/91 (UCLA Campus Computing Network 1974, p. 6). Chemists were able to access a molecular mechanics system at UC San Diego and a computational chemistry project jointly run by Wright-Patterson Air Force Base and the University of Chicago, and geologists could participate in a conferencing system set up by the US Geological Service in 1973 (Lycos 1975, pp. 42, 156).

Some members of the armed services used the ARPANET to access ARPA's computer resources or participate in its research projects (Stephen Lukasik, telephone conversation with author, 1 May 1997). The Army used the network to collaborate with ARPA's ballistic missile program, and the Air Force participated in ARPA's online seismic research. The Navy used the network to access the ILLIAC IV for acoustic signal processing. Researchers at the Aeronautical Systems Division of Wright-Patterson Air Force Base used the network to collaborate with colleagues at the Argonne National Laboratory on mathematical and chemical research (Lycos 1975, pp. 156, 174).

But despite this scattering of applications and users, most of the ARPANET's capacity went unused in the early years. Computer researchers, who were supposed to be the network's primary beneficiaries, used only a few of its remote computers to any significant extent. Though ARPA's seismology and climatology projects made use of the network, most other non-computer-science groups used it only for small-scale experiments if at all. (There is no evidence that it was ever used by ARPA's behavioral scientists.) The hope that the ARPANET could substitute for local computer resources was, in most cases, not fulfilled.

The Decline of the Ideal of Resource Sharing

The ARPANET had been designed to give researchers access to computer resources that were presumed to be scarce. Ironically, however, many sites rich in computing resources seemed to be looking in vain for users. As the 1970s progressed, the demand for remote resources actually fell. ARPA managers went out of their way to find projects that had use for large computing resources such as ILLIAC IV and the Datacomputer. Despite a number of productive experiments using remote hardware and software, most members of the ARPANET community were not using the network the way it was originally intended: resource sharing, in the sense of running programs at remote sites, did not become the ARPANET's major purpose. In a 1990 interview, Leonard Kleinrock recalled:

> Originally, the network was supposed to provide resource sharing. . . . For example, you would log on to Utah to use their graphics capability there. At one time it was thought maybe you could import the software to your machine and run it locally. But the original concept was that you would do resource [sharing] through the network—that's not really what happened.

A number of factors may account for the demise of the ideal of resource sharing. Non-expert computer users wanting access to resources at other sites faced a daunting array of obstacles, as has already been noted. On the other hand, in an example of the "not invented here" syndrome, computer scientists who had created hardware or software at their own sites often were uninterested in using machines at other sites; and those who did want to use a remote program were more likely to copy it and run it on their local machine than to run it remotely on its "native" machine (David Farber, conversation with author, 22 February 1996).

Many had expected that the network would be used for "distributed computing." This meant dividing a computing task among two or more machines, each of which would run part of the program; the various parts of the program would interact over the network as necessary. Distributed computing was supposed to allow users to combine the capacities of various specialized machines. In practice, however, the ARPANET project saw little if any realization of distributed

computing.[14] Aside from the fact that there were incompatibilities between computers, the technical vision of distributed computing did not mesh with the administrative reality of using computer resources. The goal was to divide a task among various computers according to which machines were available and best suited for the job; ideally, the user would never need to know the details of the division of labor. In practice, however, because each site exercised administrative control over its computers, a user had to set up an account on each computer that might be used, and had to pay for whatever computer time was actually used. This made using multiple hosts more of a headache than a benefit to the average user (Heart et al. 1977, p. III-84). The problem was never solved, as Heart (1990) noted:

> When the network was originally built, Larry [Roberts] certainly had high in his set of goals the idea that different host sites would cooperatively use software at the other sites. There's a guy at host one, instead of having to reproduce the software on his computer, he could use the software over on somebody else's computer with the software in his computer. And that goal has, to this day, never been fully accomplished. . . . So that turned out not to be the main thing that was created by the ARPANET.[15]

In addition to these technical and administrative obstacles, the attitudes that underlay the goal of resource sharing—both those of ARPA personnel and those of computer professionals in general—were beginning to change as the 1970s progressed. As the computer industry matured and a wider range of high-performance computers became available, ARPA managers no longer felt it necessary for the agency to build its own large machines; this meant that finding ways to share large computer resources became less of a priority (McKenzie 1997). At the same time, an increasing number of scientists were turning to smaller computers to meet their research needs. Minicomputers became popular in laboratories because they were much less expensive than paying for time on a large computer, and because a scientist could have control over an entire computer and could adapt it to the specific requirements of his or her lab. Looking back in 1988, Lawrence Roberts concluded that resource sharing had made economic sense only in the days when most ARPA researchers were using mainframe computers. Since mainframes provided computing power in large, fixed amounts, it was sometimes difficult to match the size of local resources to the needs of local users, and thus it may have been more cost effective to obtain computer resources over a network. Once a site could give each researcher a mini- or microcomputer, however, a network that provided access to computer resources no longer offered an economic advantage (Roberts 1988, p. 158). These developments undermined the old "computer utility" view that the point of a network was to help users access large, centralized computing systems. If extensive use was to be made of the ARPANET, it would have to be for some other purpose.

Finding a "Smash Hit": Email

Had the ARPANET's only value been as a tool for resource sharing, the network might be remembered today as a minor failure rather than a spectacular success. But the network's users unexpectedly came up with a new focus for network activity: electronic mail.

Email (initially called "net notes" or simply "mail") made an inconspicuous entry onto the ARPANET scene. Since many time sharing systems provided ways for users to send messages to others on the same computer, personal electronic mail was already a familiar concept to many ARPANET users. By mid 1971, when most of the sites had their host protocols in place, several ARPANET sites had begun experimenting with ideas for simple programs that would transfer a message from one computer to another and place it in a designated "mailbox" file. At the Stanford Research Institute, for instance, Richard Watson proposed such a system in order to make it easier for the Network Information Center to collect and distribute information about ARPANET sites (Watson 1971).

The first working network mail program was created by Ray Tomlinson, a programmer at Bolt, Beranek and Newman. Tomlinson modified the mail program he had written for BBN's TENEX operating system to specify a host name as well as a user name in the mail address, and he modified another command so it would transfer mail files between machines. His programs were incorporated into subsequent versions of TENEX, so that other ARPANET sites with TENEX machines were able to take advantage of the email feature (Ray Tomlinson, email to author, 10 November 1997).

In 1972, the Network Working Group was working on the specification for the file transfer protocol, which would replace the use of telnet for file transfers. Several people suggested making an addition to the ftp standard that would support email transfer. This was done at a March 1973 meeting of the Network Working Group, and it became possible to send messages using ftp instead of Tomlinson's TENEX-specific command. The ftp-based method of mail transfer was used

until the early 1980s, when the NWG developed a separate mail protocol. (See Postel 1982.) Various members of the ARPANET community also wrote mail-reading programs that presented the contents of a user's mailbox in an organized way. One of the first mail readers was created by Lawrence Roberts; the first to become widely popular was MSG, written in 1975 by John Vittal at BBN.

Email quickly became the network's most popular and influential service, surpassing all expectations. The *ARPANET Completion Report* called its use by researchers for collaborative work the "largest single impact" of the ARPANET, noting that, along with the ability to easily share files, email had "changed significantly the 'feel' of collaborative research with remote groups" (Heart et al. 1978, p. III-110). Systems administrators began to use email for more mundane tasks, such as reporting hardware and software problems (Michael Hart, email to author, 26 March 1997). Inventive students participating in the early 1970s counterculture were rumored to use email for transcontinental drug deals (Les Earnest, email to author, 26 March 1997). ARPANET users came to rely on email in their day-to-day activities, and before long email had eclipsed all other network applications in volume of traffic. The *Completion Report* called electronic mail a "smashing success" and predicted that it would "sweep the country" (Heart et al. 1978, pp. III-113–III-115).

Email had several advantages over postal mail and the telephone. It was virtually instantaneous, and it did not require the sender and receiver to be available at the same time. Email programs were fairly simple to use even for computer novices, and the email addresses of registered ARPANET users could easily be found through the NIC. Copies of a message could be sent to several addresses at once, and widely dispersed groups could use email to coordinate their activities.[16] As email became popular, Roberts began funding "email hosts" at USC's Information Sciences Institute and at Bolt, Beranek and Newman; these provided email addresses for TIP users who did not have their own accounts on an ARPANET host (McKenzie 1997).[17] And ARPANET hosts that were connected to other networks (such as Telenet) often provided services for transferring mail between the two networks, further extending the community of email users.

Within the ARPA office and in the wider Department of Defense community, the use of email was vigorously promoted by Roberts and by ARPA director Stephen Lukasik. Roberts began using email to correspond with his contractors, thus giving Principal Investigators additional motivation to start using it themselves. Roberts found that email helped him overcome obstacles of time, space, and social distinctions in managing ARPA's many computer contracts. Alex McKenzie (email to author, 9 November 1997) recalled: "Email suited Larry's all-hours-of-the-night work habits and the far-flung set of projects he was responsible for. . . . He also liked the ability to use email to 'go around' the PIs and communicate directly with lower level employees." Lukasik also saw email as a convenient way to communicate with his managers and contractors. Roberts (1988, p. 168) recalled: "Steve Lukasik decided it was a great thing, and he made everybody in ARPA use it. So all these managers of ballistic missile technology, who didn't know what a computer was, had to start using electronic mail."[18]

At ARPA's headquarters, the appeal of the network had nothing to do with computers but everything to do with access to power. When Lukasik first got on the network, the only other people using email were computer scientists in the Information Processing Techniques Office. Program managers from ARPA's other offices began to notice that the IPTO contractors seemed to do better in the budget process because they were on closer terms with Lukasik. Wanting that same access, the other program managers began using email also. "It wasn't a technical issue," according to Lukasik (telephone conversation with author, 1 May 1997); "it was a management issue." "The way to communicate with me," Lukasik recalled, "was through electronic mail, and so almost all the offices then got on the net, and then the Strategic Office understood its utility and the Tactical Office understood its utility and my old Nuclear Monitoring Office understood its utility. . . . Of course, one can argue that even without me, everyone would be on networks because that's the way you work today, but in fact, you know, I really worked on it." From ARPA email began to spread to the rest of the military, and by 1974 "hundreds" of military groups were using the ARPANET for email (ibid.).

The popularity of email was not foreseen by the ARPANET's planners. Roberts had not included electronic mail in the original blueprint for the network. In fact, in 1967 he had called the ability to send messages between users "not an important motivation for a network of scientific computers" (Roberts 1967b, p. 1). In creating the network's host software, the Network Working Group had focused on protocols for remote login and file transfer, not electronic mail. Frank Heart (1990, p. 32) recalled: "When the mail was being developed, nobody thought at the beginning it was going to be the smash hit that it was. People liked it, they thought it was nice, but nobody imagined that it was going to be the explosion of excitement and interest that it

became." A draft of the *Completion Report* referred to email as "unplanned, unanticipated, and mostly unsupported" (Heart et al. 1977, p. III-67).

Yet the idea of electronic mail was not new. MIT's CTSS computer had had a message feature as early as 1965, and mail programs were common in the time sharing computers that followed (Heart et al. 1977, pp. III-70–III-71). According to BBN's John Day (telephone conversation with author, 11 April 1997), it was an obvious step for programmers to expand these mail services to the network: "The paradigm was, what do we have on the operating system and how can we provide that on the network?"

Why then was the popularity of email such a surprise? One answer is that it represented a radical shift in the ARPANET's identity and purpose. The rationale for building the network had focused on providing access to computers rather than to people. In justifying the need for a network, Roberts had compared the cost of using the network against the cost of sending computer data by other media, but he had not compared the cost of email against the costs of other means of communication. The paradigm of resource sharing may have blinded the ARPANET community to other potential uses of the network.[19]

It was not that ARPA's managers did not value computer-mediated interaction. Indeed, Taylor and Roberts had expressed their hopes that the ARPANET would help build a community of researchers. But Roberts and others had expected users to collaborate by sharing files and programs or by using the centralized bulletin boards at SRI's Network Information Center. To help the NIC fill this role, Douglas Engelbart had created NLS, an information resource that provided a sophisticated environment for creating databases and conducting online discussions. But NLS was unfamiliar and confusing to many people, especially since most remote users lacked the specialized interface hardware that the system was designed to use (Heart et al. 1977, pp. III-67–III-68). Consequently, ARPANET users did not rely much on NLS as a way to interact with other people (McKenzie 1997). Had NLS been easier to use, it might perhaps have become the preferred method of communication, rather than email.

Computer experts put the first email programs in place, but non-expert users also had a role in building this new capability. In particular, ARPA director Stephen Lukasik used his influence to make sure that new features were added to the mail readers, so that a network tool that had been designed for computer scientists would meet the needs of other users. In the first mail readers, invoking the mail command would produce a printout on the screen of all the messages in one's mailbox, with the most recent messages last. For computer scientists, this was an appropriate format. "Computer scientists used mail in an almost real-time situation," according to Lukasik; they would send messages back and forth in rapid succession, "almost like a conversation." They usually had no need to keep old messages, so they had to deal with only a few messages at a time. But, Lukasik notes, "as a *manager*, you need to keep email for some time." Lukasik often asked contractors and subordinates for numbers and details on their projects, and he needed to save the answers for future reference. For him, email was not a conversation but a way to gather information. Typically he would find himself with fifty or so old messages in his mailbox, and he would have to scroll through all of them to get to the new messages at the bottom (Lukasik, telephone conversation with author, 1 May 1997). On hearing Lukasik complain about this, Roberts wrote a new mail program that sorted incoming mail into folders and made it easy to selectively view or delete messages.

Email and mailing lists were crucial to creating and maintaining a feeling of community among ARPANET users. Mailing lists allowed users to send messages to a single list address (such as sf-lovers@hostname) at a host site where a list administrator maintained a database of list members. A program running on the host computer then automatically re-transmitted the message to each person in the database. This meant that an individual could communicate with a large group without having to send out numerous messages and without having to keep track of the addresses of all the members. Even more important, mailing lists allowed a virtual community to take on an identity that was more than the sum of the individuals who made it up. For example, a science fiction enthusiast did not need to be personally acquainted with others in order to join an online discussion of the latest popular story; the names of lists advertised the common interests of their members, and most lists were open to all who wished to participate. Mailing lists provided a way for people to "meet" and interact on the basis of shared interests, rather than relying on physical proximity or social networks.[20]

The ARPANET Remade

The ARPANET had been ushered into the public eye with a triumphant demonstration at the International Conference on Computer Communications in the autumn of 1972. But the ARPANET was not a finished product in 1972, nor was its success certain. The early years of the network were full of confusion, false starts, and frustration. Nonetheless, through many individual choices (and some top-down pressure from ARPA), people began to use the ARPANET and to discover how it could best serve their needs. These users were responsible for transforming the ARPANET from an experimental system with limited appeal to an operational service whose existence could be justified and even celebrated.

In the process of using the network, the ARPANET community developed a new conception of what networking meant. Since the original view of the network planners was that "resources" meant massive, expensive pieces of hardware or huge databases, they did not anticipate that people would turn out to be the network's most valued resources. Network users challenged the initial assumptions, voting with their packets by sending a huge volume of electronic mail but making relatively little use of remote hardware and software. Through grassroots innovations and thousands of individual choices, the old idea of resource sharing that had propelled the ARPANET project forward was gradually replaced by the idea of the network as a means for bringing people together.

Email laid the groundwork for creating virtual communities through the network. Increasingly, people within and outside the ARPA community would come to see the ARPANET not as a computing system but rather as a communications system. Succeeding generations of networks inspired by the ARPANET would be designed from the start to act as communications media. By embracing email, ARPANET users gave the network a new purpose and initiated a significant change in the theory and practice of networking.

4

From ARPANET to Internet

Over the course of a decade, the ARPANET (a single network that connected a few dozen sites) would be transformed into the Internet (a system of many interconnected networks, capable of almost indefinite expansion). The Internet would far surpass the ARPANET in size and influence and would introduce a new set of techniques to computer networking. For all its later importance, however, the Internet was not part of ARPA's initial networking plans. The Internet represented a new approach to networking, and its creation was prompted by a series of unforeseen events.

Even as the ARPANET was being developed, a small group of ARPA contractors were already working on the next generation of network technology for the military. In the course of trying to resolve some dilemmas they encountered in these other networking projects, ARPA researchers Robert Kahn and Vinton Cerf began to consider how to interconnect dissimilar networks.

The Internet architecture[1] that Cerf and Kahn proposed was eventually used not only to build the Internet itself but also as a model for other networks. One reason for the widespread acceptance of this new approach was that a range of interest groups participated in the Internet's design, including network researchers from outside the United States. Shaped both by ARPA's military concerns and by the opinions of an international community of network experts, the Internet would depart in significant ways from the design of the ARPANET, resulting in a system that was not simply bigger but also more flexible and decentralized.

Although the design of the Internet came from the international computer research community, the actual implementation was done under the auspices of the US military. Operational branches of the military began using the ARPANET during the 1970s, and

Department of Defense agencies other than ARPA became involved (somewhat reluctantly) in managing the network. The Department of Defense would play many roles in the emergence of the Internet: funding research and development, transferring technology to operational forces, using its financial resources to shape the commercial market for network products, and exercising management control over the ARPANET and its community of users.

New Directions in Packet Switching

The high-profile demonstration of the ARPANET at the 1972 International Conference on Computer Communications symbolically marked the completion of the original network project. That autumn also brought a change of personnel in ARPA's Information Processing Techniques Office. Lawrence Roberts left ARPA at the end of the year to head Telenet, BBN's commercial spinoff of the ARPANET. Around the same time, Robert Kahn, the BBN researcher who had been largely responsible for organizing the ARPANET demonstration, joined IPTO as a program manager.[2]

Kahn had no plans to pursue internetworking at this point; in fact, he did not expect to head a network project. On joining ARPA he had been told that he would manage a program in flexible manufacturing; when this was canceled soon after his arrival, Kahn was forced to find alternative projects for his office to fund. Since his main expertise lay in computer networking, he decided to look for defense-related projects in that area.

Kahn initiated an array of projects in network security and digital speech transmission. He also took up some investigations, begun by Roberts, of the application of packet switching techniques to two media that previously had not seemed suitable for data communications: land-based radio and satellite radio (hereafter referred to as "radio" and "satellite" respectively). Like the ARPANET, the experiments with radio and satellite combined fundamental research in computer science with potential for military applications. Packet radio seemed like an ideal medium for military field operations, since radio terminals (unlike the telephones of the 1970s) could be mobile. Satellite could provide worldwide communications and could support data-intensive defense applications, such as seismic monitoring of nuclear weapons tests; it also seemed suitable for use on Navy ships. Ironically, however, the biggest impact of the packet radio and satellite projects came

not from their use in military operations (which turned out to be limited, though both were used in the 1991 Gulf War) but from their unplanned contributions to local-area networking and internetworking.

Alohanet and Ethernet

Packet radio posed a new set of theoretical and engineering questions. Radio, as a broadcast medium, has capabilities and limitations quite different from those of the wired telephone network. In a broadcast system, each message transmitted is received by every host; the host to whom it is addressed accepts the message, while the others ignore it. There is no need to route messages between individual hosts, and a mobile radio host can move around within the area covered by the broadcast without disrupting its ability to receive messages.[3] It is also easy to send a message to many hosts at once. Before these benefits could be realized in a computer network, however, a basic theoretical question had to be answered. Two messages sent at the same time on the same broadcast channel will interfere with each other, so both will be reduced to gibberish. How could a radio network be designed to prevent or to recover from such collisions?

The first experiments that set out to answer this question began in 1970, when Lawrence Roberts was still head of IPTO. They were led by Norman Abramson of the University of Hawaii and were funded in part by the Navy and by ARPA. The Hawaii group wanted to explore packet switching as an alternative to costly dial-up telephone connections for accessing the university's computers. The ARPANET was using leased telephone lines for its links, but Hawaii's noisy telephone lines were ill suited for data transmission (Kahn 1975, p. 177; Kahn 1989, p. 13). Abramson's group decided to try packet radio as a potentially cheaper and better means of serving the university's computer users. In 1970, Roberts provided IPTO funding for Abramson's Alohanet, which would link radio terminals at the University of Hawaii's seven campuses and numerous research institutes to its main computer center near Honolulu (Abramson 1970, p. 281). Abramson used a computer interface modeled after the ARPANET's IMP, and his team received design support from ARPANET veterans at the Network Analysis Corporation and at UCLA.

The Alohanet design used two radio channels. On one channel, the computer interface (named Menehune[4]) would broadcast packets from the computer center to the user terminals. Since the Menehune would be the only machine transmitting on this channel, there would no interference. The other channel would carry all the traffic from the users' terminals to the computer center. How could all the users share a broadcast channel without interfering with one another?

The Alohanet solution was startling in its simplicity. The designers did not try to prevent collisions at all; they simply made sure the system could recover from collisions when they occurred. The method was called "random access" because access to the channel by different terminals was not scheduled or coordinated; each terminal transmitted on its own initiative whenever it had data to send. If two users happened to transmit packets at the same time, both packets would be garbled. The system relied on acknowledgments to keep data from being lost in such collisions. The Menehune acknowledged safe arrival of a packet. If the sending terminal did not receive such an acknowledgment (perhaps because the packet had been destroyed in a collision), the sender would re-transmit the packet. But if *two* terminals lost packets in a collision, what would prevent them from re-transmitting at the same moment and endlessly repeating the collision? The Aloha answer was to have the terminals wait before re-transmitting, each terminal choosing its waiting time at random from a specified range. In all probability, the two terminals would choose different times to re-transmit, and thus they would avoid a repeat collision (Binder et al. 1975, p. 206). Figure 4.1 illustrates packet transmission on a random-access channel.

The Aloha method, a significant advance in communications theory and practice, became a standard topic in computer science textbooks. It also provided inspiration to Robert Metcalfe, a graduate student at Harvard. Metcalfe had been drawn into the ARPANET development effort through a part-time job at MIT's Project MAC, and the experienced had inspired him to write his doctoral dissertation on packet switching networks. To satisfy his Harvard committee, he needed to find a theoretical aspect of networking on which to focus. Metcalfe happened to be friendly with Stephen Crocker, then an ARPA program manager. Crocker gave Metcalfe some papers describing the Aloha project. Intrigued, Metcalfe read the papers and came up with a mathematical model that would significantly improve the performance of the Aloha network. His key insight was that varying the re-transmission interval in response to traffic loads—waiting longer to re-transmit when traffic is heavy—could radically improve the throughput of such systems by cutting down the number of repeat collisions. Metcalfe

⑲

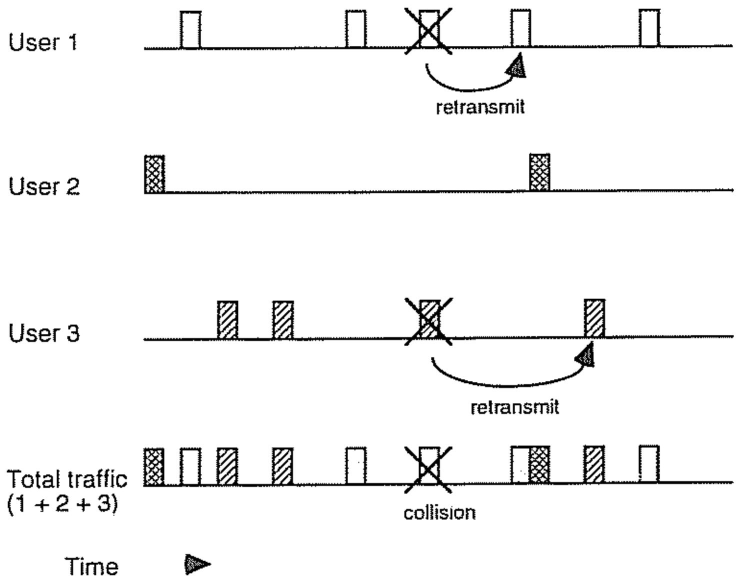

Figure 4.1
The Aloha technique. Packets from several users share a random-access channel. The different retransmit intervals keep the retransmitted packets from colliding a second time. Adapted from Abramson 1970.

made the analysis of this phenomenon, called "exponential back-off," the heart of his 1973 dissertation, titled Packet Communication.

In 1972, while still working on his Ph.D., Metcalfe took a job at the Xerox Palo Alto Research Center. PARC was a center of innovation; its staff of talented researchers included Robert Taylor, who had initiated the ARPANET project. Taylor had left ARPA to become associate director of the Computer Science Laboratory at PARC in 1970. In 1972 he was leading the development an innovative computer workstation called the Alto. A number of these workstations had been deployed around PARC, and Metcalfe was asked to design a system to connect them. Drawing on his dissertation work, Metcalfe created a random-access broadcast system that was initially known as the Alto Aloha network but was soon dubbed Ethernet (Thacker 1988, p. 274). Ethernet used a cable rather than a radio channel as the transmission medium. One advantage of using a cable was that it provided much more bandwidth. Alohanet had transmitted thousands of bits per second; Ethernet could carry millions per second. In combination with Metcalfe's improved re-transmission algorithm, the use of cable made

Ethernet a fast and efficient way to transmit packets over short distances.

Recognizing the commercial potential of his invention, Metcalfe left Xerox to found a company called 3Com. In 1981 3Com announced an Ethernet product for workstations, and in 1982 it introduced a version for personal computers. For the first time, owners of small computers had an affordable networking option, and Ethernet quickly became a standard technique for local area networking. By the late 1990s, millions of LANs around the world were using Ethernet (Metcalfe 1996, p. xix). By providing the technical foundation for Ethernet, ARPA's first investment in packet radio had the unanticipated dividend of helping to spawn a huge commercial market for LAN systems.

Packet Radio and Satellite

Robert Kahn decided that ARPA should follow up on the Alohanet project by building a packet radio network of its own in the San Francisco Bay area. That system, called PRNET, consisted of a control station, several broadcast nodes (called repeaters), and a multitude of radio sets that could be attached to computers or terminals. The radio units were built by the Collins Radio Group of Rockwell International, the control stations were supplied by Bolt, Beranek and Newman, and the Stanford Research Institute was in charge of system integration and testing; the Network Analysis Corporation and the University of California at Los Angeles also participated (Kunzelman 1978, pp. 157, 160). PRNET went into experimental operation in 1975 with a single control station and four repeaters, as illustrated in figure 4.2.

Since radio was already used for command and control in the field, ARPA's packet radio program seemed more directly applicable to military operations than the ARPANET had been. Kahn was especially interested in using packet radio as a way to transmit voice for command and control. Packet switching could make voice transmission more efficient and could correct for the errors introduced by noisy radio channels. It would also make it harder for an enemy to eavesdrop on a conversation, since a message that had been digitized and split into packets would be unintelligible until it was reassembled and decoded at the receiving end (US Congress 1974, p. 135). The design of PRNET reflected its intended use in combat situations. To make the equipment easy to set up and use, Kahn specified omnidirectional antennas so that users would not have to align them in the field, and he designed the system to keep track of the location and status of each

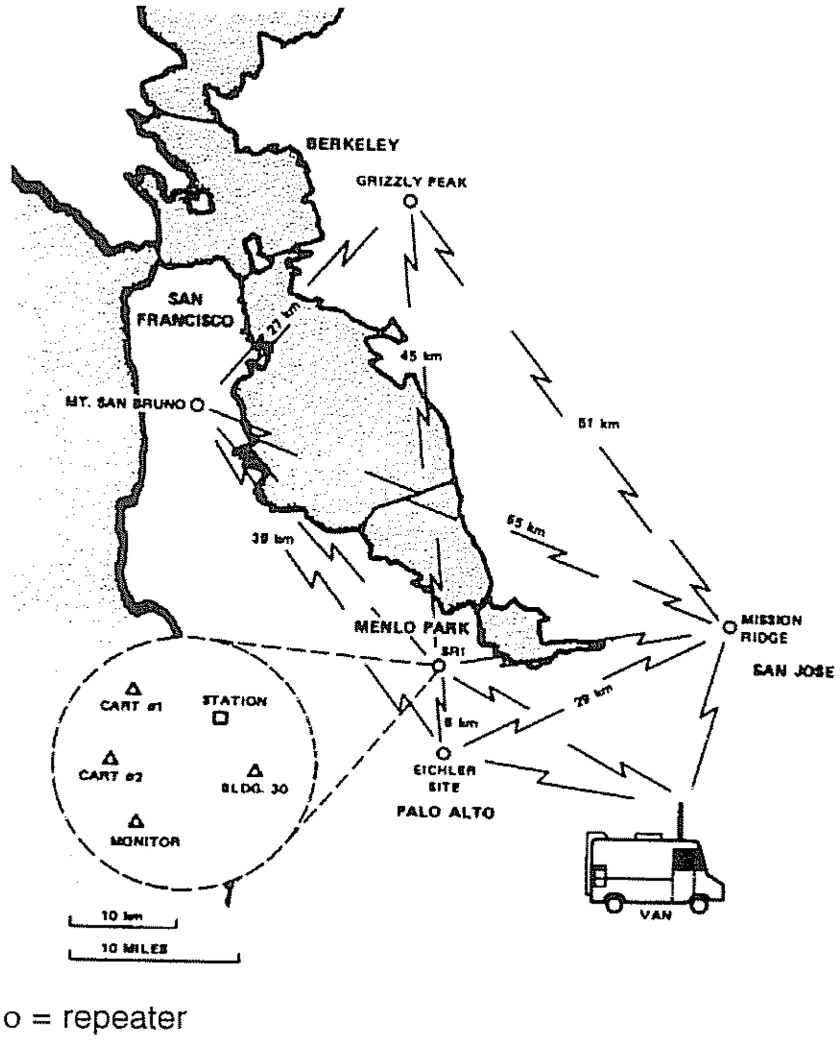

Figure 4.2
A map of the PRNET in 1977. The repeaters were located in elevated areas to increase their unobstructed transmission range. A radio-equipped van was used to test mobile communications. Source: Kahn et al. 1978.

component without human assistance (Kahn et al. 1978, p. 1478). Since the repeaters would operate unattended in remote outdoor areas, possibly in war zones, they were designed to be simple and rugged and to use minimal power. PRNET used various experimental data security techniques to prevent unauthorized access or tampering, and most control functions were located in the manned central station rather than in the more vulnerable repeaters. At the same time, PRNET made use of distributed control to survive an attack: the repeaters could take over routing without help from the control station (Kahn et al. 1978, p. 1417). Despite its careful design, however, PRNET was never developed to the point where it could be used in actual combat zones. ARPA managed to set up some test applications for the Army and the Air Force in which radio links were used to provide computer support for base operations, but the packet radio technology remained largely experimental (Cerf 1990).

ARPA ventured into another uncharted area with its packet satellite program. The development of packet switching networks in the 1960s and the 1970s had been paralleled by the development of satellite communications, which also had roots in the Cold War. In October of 1957, the USSR had launched the first artificial satellite, Sputnik I. Within a year, Sputnik had been joined in the sky by experimental satellites from the United States and Canada. In 1962, to encourage the use of satellites for peaceful purposes, President John F. Kennedy sent Congress a bill that became the Communications Satellite Act. This act supported the formation of a private corporation, Comsat, to provide commercial telecommunications service using satellites. At the same time, the United Nations was moving to create an organization to develop and operate a global communications satellite system that could be shared by all UN member countries. The International Telecommunications Satellite Organization (Intelsat) was established in August of 1964, with Comsat as the first US participant. Intelsat's first geosynchronous satellite was launched in April of 1965, and several generations of Intelsat satellites followed.

Satellites offered high bandwidth, and a single satellite could cover a wide area. Because of their high cost, however, satellite connections were rarely used for data transmission. Packet switching had the potential to make data communications via satellite economical. Kahn's immediate motivation for pursuing satellite network experiments was IPTO's seismic monitoring program (described in chapter 3 above), whose seismic sensors in Scandinavia generated voluminous streams

of data that had to be transferred to the United States for analysis. Kahn decided that a packet satellite system would provide the most efficient way to do this, and he persuaded ARPA's seismic monitoring research office to shift its data transfer operations to a satellite link.

IPTO began using Intelsat I for experimental satellite links in 1973, connecting first the University of Hawaii and then University College in London to the ARPANET. In the autumn of 1975, Kahn began work on the Atlantic Packet Satellite Network (SATNET) project. Jointly sponsored by ARPA, the British Post Office, and the Norwegian Telecommunications Authority, the SATNET project was intended to support both network research and the transmission of seismic data for defense purposes (Jacobs et al. 1978).[5] In its initial configuration, SATNET linked four sites: one in Maryland, one in West Virginia, one in England, and one in Norway. The Norwegian site, run by the Norwegian Defense Research Establishment, was associated with the seismic monitoring program; the British site was operated by Peter Kirstein, a researcher at University College London who had arranged to participate in ARPA's network program. SATNET was a broadcast system; the four stations used a single radio channel to communicate with the satellite. The ground stations were connected to packet switches that were similar to the ARPANET IMPs but had been specially adapted to handle the high bandwidths and long transmission delays involved in satellite communications (ibid., p. 1461).

By the mid 1970s, then, ARPA was operating three separate experimental networks: ARPANET, PRNET, and SATNET. All these networks used packet switching, but they used it in distinctly different ways that optimized the technique for each particular medium. Kahn began to think about bringing these three networks together while he was struggling to develop PRNET into something more than an experiment.

PRNET connected a single computer center at SRI to a set of mobile radio units. Portable terminals could be attached to the mobile units, but portable host computers did not yet exist. To make the network useful, a way would have to be found to reach additional host computers. Kahn (1990) later described the situation as follows:

> Partway through the first year of the program it became clear to me that we were going to have to have a plan for getting computer resources on the net. In 1973, mainframe computers were multi-million-dollar machines that required air-conditioned computer centers. You weren't going to connect them to a mobile, portable packet radio unit and carry it around. So my first

question was "How am I going to link this packet radio system to any computational resources of interest?" Well, my answer was, "Let's link it to the ARPANET."

But linking PRNET and ARPANET was no simple proposition. The networks were technically incompatible: PRNET used broadcast, ARPANET point-to-point transmission; ARPANET guaranteed reliable transmission and sequencing of packets, PRNET did not; and packet sizes and transmission speeds differed between the two networks (Norberg and O'Neill 1996, p. 182). Moreover, Kahn (1989, p. 15) realized that he would eventually want to link the ARPANET to additional networks, such as SATNET, that used still other techniques. No one in the field of computing had ever attempted to connect such dissimilar systems, and there were no models from which to work.

As Kahn began thinking about ways to address the general problem of interconnecting heterogeneous networks, he set in motion what would become the Internet program.[6]

The Internet Program

In the spring of 1973, Kahn approached Vinton Cerf (then at Stanford University) with the idea of developing a system for internetworking. Cerf and Kahn had worked together on testing the first ARPANET node at UCLA, and Cerf had been one of the original designers of the ARPANET host protocol, so Kahn felt that Cerf was the right person to turn to for help. "It just took one session," he recalled (1990), "before the two of us were on the same wavelength as to what we needed to do. And he and I just jointly worked it out from there." The two collaborated on the initial design of a system that would link ARPA's various networks to form what would become known as the ARPA Internet. In the summer of 1973 they wrote a paper outlining the basic Internet[7] architecture (Cerf and Kahn 1974). Cerf received an ARPA contract to work out the detailed specifications of the system, and in 1976 he joined Kahn at ARPA and took over as program manager for the agency's various network projects.[8]

Starting from Kahn's original problem (how to access host computers from the packet radio network), Cerf and Kahn raised two basic questions (Cerf 1990). First, if the packet radio network were to provide reliable connections with the host computers, it would need a host protocol that could compensate for its error-prone transmission medium. What would that host protocol look like? Second, what kind

of mechanism could provide an interface between two distinct networks such as PRNET and ARPANET? The answers they worked out would eventually become the basis for a set of internetworking techniques and for an experimental internet based on those techniques. But before they attempted to build such an internet, Cerf and Kahn sought out advice and opinions from the world's networking experts—a move that would significantly shape the resulting system.

An Inclusive Collaboration

Though Cerf and Kahn were the main architects of the Internet, they had a number of collaborators both from within the ARPANET group and from a growing international networking community. Among the members of the ARPA research community who were involved in designing the Internet were Yogan Dalal, Richard Karp, and Carl Sunshine (graduate students of Cerf's at Stanford); Stephen Crocker (who had worked on NCP as a graduate student and who was now at IPTO); Jon Postel of the Information Sciences Institute at USC; Robert Metcalfe of Xerox PARC; and Peter Kirstein of University College London (Cerf 1990, pp. 29, 33–34).

A number of computer researchers from outside the United States became involved in the project through the International Network Working Group, which had been formed at the 1972 International Conference on Computer Communications. The INWG brought together representatives from the world's major packet switching projects—the ARPANET, the British NPL network, and a French research network called Cyclades—and from various national telecommunications carriers who were planning packet switching networks of their own. The group soon affiliated itself with the International Federation for Information Processing (an association of technical societies that exchanged information and cooperated on the development of new technologies), thus adding to the INWG's visibility and legitimacy within the international computer science community.

Though the INWG had no formal authority to create international standards for computing, its members hoped to reach an informal agreement on internet standards so as to be able to interconnect their various systems (Curran and Cerf 1975, p. 8). Among the most active members from the United States were Franklin Kuo (who had worked on the Alohanet), Alex McKenzie of BBN, and Vint Cerf, who chaired the INWG from 1972 through 1976 (Cerf 1990). Cerf's involvement in the INWG allowed him to draw on the combined experience and expertise of this international networking community; it also encouraged him to expand the focus of the ARPA's internet program so that the proposed system would accommodate the various types of packet networks being built in Europe as well as ARPA's own networks.[9]

The ultimate conception and design of the Internet system would be shaped by the agendas of all these participants. In addition, the ARPA managers, who ultimately had to justify the program in military terms, wanted the system to support the complex requirements of the armed services. Writing in 1978, Cerf noted that computers were becoming ubiquitous in military equipment:

> A fundamental premise of all current Command, Control and Communications (C3) research is that digital technology and computing systems will play a central role in the future. It is already apparent that computers are being employed in tactical as well as strategic military equipment. . . . To make this collection of computers, sensors, and databases useful, it is crucial that the components be able to intercommunicate. (Cerf 1979, p. 288)

For military purposes it was important that the Internet accommodate different types of networks, since it was expected that military communications systems would be optimized for a variety of service environments:

> Ethernet ideas might serve well in garrison or aboard a ship. Packet radio concepts are crucial for local area mobile communication (e.g., land mobile, ground-air, ship-ship). ARPANET technology is appropriate for fixed installations such as in CONUS [the continental United States] or Europe. Finally, packet satellite supports wide geographic coverage while permitting efficient and dynamic allocation of transmission capacity as needed. The conclusion is that many different transmission technologies are needed for military operations and therefore, a sensible C3 system must incorporate a strategy for the interoperation of dissimilar computer communication networks. (Cerf 1979, pp. 288–289)

ARPA had started the packet radio and satellite programs to meet the military's perceived need for these different types of systems. The design of the Internet would also support this objective.

The members of the INWG were motivated by a common desire to enlarge the scope of their networks through interconnection, but they had divergent views on internet design principles. The most active French members came from the Cyclades project. Named after a group of islands in the Aegean Sea (since it connected isolated "islands" of computing), Cyclades was an experimental network project begun in 1972 with funding from the French government. Its architects, Louis Pouzin and Hubert Zimmerman, had very definite ideas about internetworking. In fact, Cyclades, unlike ARPANET, had been explicitly designed to facilitate internetworking; it could, for instance, handle varying address formats and varying levels of service (Pouzin 1975b, p. 416).

Cyclades was based on a very simple packet switching system. Rather than having the network maintain an ongoing connection between a pair of hosts, as the ARPANET did, Cyclades simply delivered individual packets (known as "datagrams"). Pouzin and Zimmerman argued that keeping network operations simple made it easier to build an internet. "The more sophisticated a network," according to Pouzin (1975b, p. 429), "the less likely it is going to interface properly with another. In particular, any function except sending packets is probably just specific enough not to work in conjunction with a neighbor." To keep the network's functions to a minimum, the French researchers argued, it was necessary for the host protocol to take on the primary responsibility for maintaining reliable connections. This went contrary to both the way BBN had designed the ARPANET and the way telecommunications carriers in France and elsewhere were planning to design their public data networks.[10] Perhaps anticipating opposition to their unconventional approach, the members of the Cyclades group were extremely vigorous in advocating their internetworking philosophy. Pouzin and Zimmerman were active in INWG. Another member of the Cyclades team, Gerard Lelann, worked in Cerf's lab at Stanford, where he was able to participate directly in the design of ARPA's internet system. According to Cerf (1990), the Cyclades group "had a lot to do with the early discussions of what the [host protocol] would look like."

England's National Physical Laboratory, which had pioneered packet switching techniques, was also involved in internetworking research. In 1971, a science and technology study group of the European Economic Community (precursor of the European Union) recommended the building of a multinational computer research network. The proposed European Informatics Network would help member countries share computer resources, would promote computer science research, and would provide a European test bed for networking techniques. Work began in 1973, and by 1976 the EIN was providing network service to ten countries.[11] Its British node was located at the National Physical Laboratory, and Derek Barber, who had worked on the original NPL network, led the development of EIN.

As part of the EIN experiment, researchers at the NPL set up a connection between their network and the EIN. They also made a trial connection between their network and the Experimental Packet Switching Service being offered by the British Post Office. In the course of these experiments, the NPL team confronted what they called the "basic dilemma" of internetworking: in order to get the most reliable and efficient service, it would be necessary to implement common host protocols on all the networks, but this would also require a substantial restructuring of existing network systems (Laws and Hathway 1978, p. 280). The NPL tried two approaches: for the EIN connection they translated between two different host protocols, while for the EPSS connection they used a common host protocol in both networks. Their experience confirmed that the translation approach was awkward and inefficient, and that establishing a standard host protocol would be the preferable way to build an internet (Laws and Hathway 1978, pp. 282–283).

Corporate researchers at Xerox PARC also played a significant part in designing the Internet. While Vinton Cerf and Robert Kahn were initiating the Internet program, Robert Metcalfe, David Boggs, and others at PARC were developing both the Ethernet local area network technology and a proprietary internet system called PARC Universal Packet ("Pup").[12] The initial Pup system was designed to connect several wide-area networks used by PARC (the ARPANET, PRNET, and the company's own leased-line network) and a number of LANs that used Ethernet (Boggs et al. 1979, p. 1). Drawing on ideas Metcalfe had presented in his 1973 dissertation, the Pup system provided a simple datagram service at the network level and relied on the hosts to provide reliable connections (ibid., pp. 3, 9). This approach to internetworking was similar to the Cyclades philosophy, but it arose from a local concern: the technical constraints of Ethernet. An Ethernet system has no "intelligence" inside the network; there is only a piece of cable connecting the computers, rather than a set of packet switching minicomputers as in the ARPANET. In an Ethernet system, therefore, the hosts must take most of the responsibility for running the network. This design was replicated in Pup. "An important feature of the Pup internet model," Boggs noted (ibid., p. 2), "is that the hosts *are* the internet." "Most hosts," he continued, "connect directly to a local network, rather than connecting to a network switch such as an IMP, so subtracting all the hosts would leave little more than wire." It is not surprising, therefore, that Metcalfe, when he joined the discussion on how to design the ARPA Internet, argued that the system should be based on the Pup approach of having simple network requirements and strong host protocols (McKenzie 1997).

Beyond Xerox PARC, which was oriented toward research rather than toward commercial production, there seems to have been no corporate participation in the design of the Internet. Few people outside the computer science community had even heard of the ARPANET in the early 1970s, and fewer still could have recognized that the Internet would someday become an important public and commercial technology. Like its predecessor, the Internet was designed, informally and with little fanfare, by a self-selected group of experts.

Designing the Internet

In June of 1973, Cerf organized a seminar at Stanford University to discuss the design of the proposed Internet and its host protocol, called the Transmission Control Protocol (TCP). All the interest groups mentioned above were represented at this meeting. As Cerf remarked in 1990, "TCP turned out to be the open protocol that everybody had a finger in at one time or another."

The seminar addressed the two questions originally raised by Cerf and by Robert Kahn: What was the best design for a universal host protocol that would work on unreliable networks such as the PRNET and not only on reliable ones such as the ARPANET? And how should the networks be attached to one another? Though these questions generated some debate, the participants were able to find enough common ground to define an approach on which most of them could agree. There was nothing inevitable about this agreement—as we will see in the next chapter, network design issues would become a source of international conflict by the mid 1970s. However, there seems to have been an emerging consensus among the computer researchers enlisted by ARPA on some basic principles.[13]

In answer to the first question, the group decided that TCP should have the responsibility for providing an orderly, error-free flow of data from host to host. Vinton Cerf, Gerard Lelann, and Robert Metcalfe collaborated closely on the specifications for TCP (Cerf 1990),[14] and thus the protocol reflected the design philosophies of Cyclades and Ethernet while deviating significantly from the approach that had been taken with the ARPANET. The ARPANET subnet was a very reliable communications system; the original host protocol, NCP, counted on this and did not have any error-recovery mechanisms. Cerf later pointed out that "even though the ARPANET was considered kind of a datagram-like system—because you put a label on the front [of each individual packet] and say 'here, deliver this'—underneath, inside the IMPs . . . things were delivered in sequence. And if they weren't in sequence there was something wrong." PRNET, on the other hand, might lose packets or deliver them out of sequence. "So we really needed a complete rethinking of the protocol suite," Kahn (1989, p. 19) recalled. TCP did much more than just set up a connection between two hosts: it verified the safe arrival of packets using acknowledgments, compensated for errors by re-transmitting lost or damaged packets, and controlled the rate of data flow between the hosts by limiting the number of packets in transit. All this made it feasible to provide reliable communications over a network as unreliable as PRNET. Cerf and Kahn planned for TCP to replace NCP as the ARPANET's host protocol and be the standard host protocol on every subsequent network built by ARPA.

Establishing a single universal host protocol was not the only possible approach to building an internet. One obvious alternative would have been to continue using different host protocols in different networks and create some mechanism for translating between them. This would have avoided the necessity of replacing existing host protocols, but Cerf and Kahn knew that such a design would not scale up gracefully: if the number of networks being connected were to grow large, the translation requirements would become unworkable. For Cerf and Kahn, the efficiency and flexibility of having a common protocol were worth the effort of converting the older system. Perhaps as important, the common protocol would create a particular type of experience for Internet users. According to Cerf (1990): "We wanted to have a common protocol and a common address space so that you couldn't tell, to first order, that you were actually talking through all these different kinds of nets. That was the principal target of the Internet protocols." Having to translate between different protocols would have emphasized the boundaries between networks, and the Internet's designers wanted the system to appear seamless. Indeed, they were so successful that today's Internet users probably do not even realize that their messages traverse more than one network.

To connect the networks physically, Cerf and Kahn proposed the creation of special host computers called "gateways." A gateway would be connected to two or more networks and would pass packets between them; all inter-network traffic would flow through these gateways. The gateways would maintain routing tables indicating how to send packets to each member network. Besides connecting networks, they would also help to accommodate differences between network systems by translating between the different local packet formats (Cerf 1979, p. 292).[15] Gateways buffered the local networks from having to know about the overall topology of the network. This made the system easier to scale up, because the local networks did not have to keep track of changes in the rest of the Internet; if a new network were added to the system, only the gateways would need to know.

This division of responsibility between local networks (which handled their own internal operations) and gateways (which routed messages between networks) was exemplified by the way the Internet handled addresses. ARPANET hosts had not needed addresses: packets were sent to a particular IMP, and it was assumed that a single host was connected there. If, instead, an entire network had been connected to that IMP, there would have been no way to specify which host on that network was supposed to receive the packet (Kahn 1989, p. 18).[16] The designers of the Internet had to devise a system of host addresses that would enable packets to be directed to a particular host on a particular network. They chose to create a hierarchical address system: one part of the address would specify the name of a network, while another part would give the name of an individual host within that network. The hierarchical address scheme facilitated the division of labor between gateways and local networks. Within each local network, the nodes would not have to know anything about non-local addressing or routing; they would simply send all packets with a non-local address to a gateway. Gateways would know how to route packets to any network, but they would not need to know the locations of host computers within that network (Cerf 1979, p. 297). This scheme kept the complexity of each part of the system manageable.

The system worked out by Cerf, Kahn, and their colleagues addressed the project's original requirements: it provided a protocol that would work over unreliable networks, and it solved the basic internetworking problems of routing and translating packet formats between networks. But some in the Internet group were critical of this initial design. Since the gateways used TCP, they ended up performing reliability functions (sequencing, error control, flow control) that were already being handled by the hosts; this made the gateways unnecessarily complex. The Xerox group compared this to "the complicated measures required to avoid deadlock conditions in the Arpanet—

conditions which are a direct consequence of attempting to provide reliable delivery of every packet" (Boggs et al. 1979, p. 8). They argued that the gateways should provide only a simple datagram service.

At an internetworking meeting at the University of Southern California in January of 1978, Vint Cerf, Jon Postel, and Danny Cohen discussed this issue and came up with a solution. They proposed splitting the TCP protocol into two separate parts: a host-to-host protocol (TCP) and an internetwork protocol (IP). The pair of protocols became known collectively as TCP/IP. IP would simply pass individual packets between machines (from host to packet switch, or between packet switches); TCP would be responsible for ordering these packets into reliable connections between pairs of hosts (Cohen 1978, p. 179). The idea of having a separate internet protocol was modeled in part on Xerox's Pup network, which was being developed around the same time (Cerf 1980, p. 11; Boggs et al. 1979, pp. 4–6).

With the new version of the Internet protocols, gateways could be simpler: they would run only IP, and they would no longer have to duplicate the host functions (now confined to TCP). The minimal functions required of the Internet Protocol also put fewer demands on member networks.[17] Introducing the new set of protocols, Cerf (1980, p. 10) wrote: "The Internet Protocol (IP) has been designed around the premise that few assumptions can be made about the type of service available from any given network." He also argued that the stripped-down functionality of IP would make military networks more robust and therefore more likely to "meet the requirements of operation under hostile conditions" (ibid., p. 11) He noted that, by accommodating diverse networks, the design would allow the armed forces to create specialized networks and also to introduce new technologies over time without major disruption of the system. "Thus," he noted elsewhere, "the problems of dealing with dissimilar tactical and strategic networks and with evolving computer communication network technology can be solved in a single stroke." (Cerf 1979, p. 289) But the version of TCP/IP that became standard in 1980 was more than a military product; it also reflected the ideas and interests of an international community of network researchers.

Initial Experiments

Designing the protocols was only the first step toward building the Internet. Putting the design into practice took several years; even in simplified form, the network protocols performed a complex set of

functions and were difficult to implement correctly in software. The initial version of TCP was specified in 1974. Bolt, Beranek and Newman had an implementation of TCP for the TENEX operating system completed in February of 1975, though it was not reported to be debugged until November. BBN also built its first experimental gateway in 1975, connecting an in-house research network to the ARPANET. Stanford University implemented TCP during 1975, and in November the BBN and Stanford groups set up an experimental TCP connection between their sites. The early tests revealed a number of deficiencies in the design, forcing the Internet group to revise the TCP specification (McKenzie 1991a).

The BBN group proposed testing TCP over the satellite network, and they began installing experimental gateways at BBN (in 1975), at University College London (in November of 1976), and at the Norwegian Defense Research Establishment (in June of 1977) (Travers 1991). As they conducted tests over these links, the Stanford and University College London researchers discovered that badly programmed implementations of TCP could drastically degrade the network's performance (Bennett and Hinchley 1978, p. 406). The protocol specification was only a "blueprint"; it was up to the host system's programmers to make a working version of the protocol—a lesson that would become painfully clear when the entire ARPANET community tried to adopt TCP. By late 1977, however, the test sites were ready to try out the new protocols, and ARPA demonstrated its first multinetwork connection. Experimenters sent packets from a van on a California freeway through PRNET to an ARPANET gateway, then through the ARPANET to a SATNET gateway on the East Coast, over SATNET to Europe, and finally back through the ARPANET to California (figure 4.3).

For the computer scientists, the 1977 demonstration confirmed the feasibility of the Internet scheme. For ARPA, it was also a way of highlighting the military potential of the new technology. Cerf (1990) emphasized that "all of the demonstrations that we did had military counterparts," suggesting how connections between radio, satellite, and telephone networks could be used during wartime:

> What we were simulating was a situation where somebody was in a mobile unit in the field, let's say in Europe, in the middle of some kind of action trying to communicate through a satellite network to the United States, and then going across the US to get to some strategic computing asset. . . . There were a number of such simulations or demonstrations like that, some of which

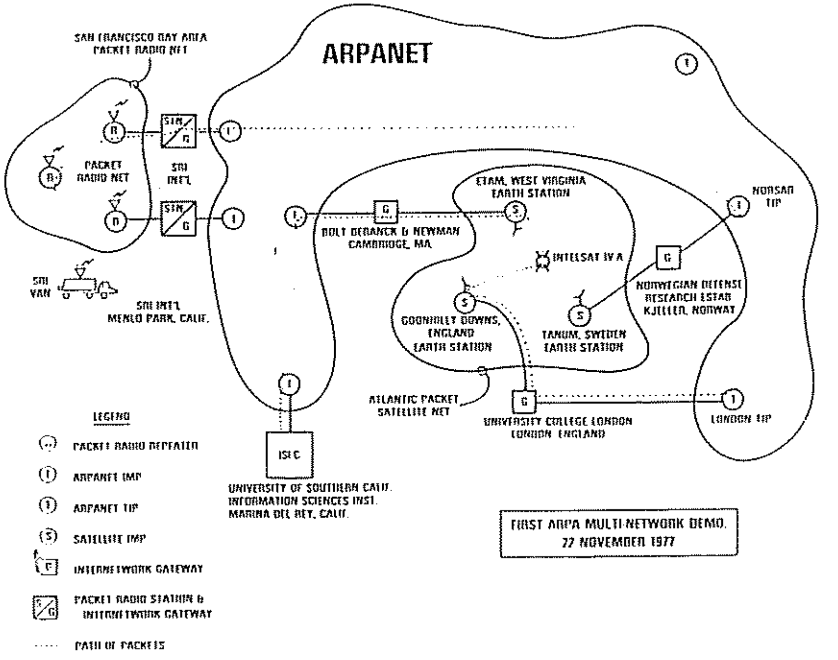

Figure 4.3
Diagram of 1977 Internet demonstration.

were extremely ambitious. They involved the Strategic Air Command at one point, where we put airborne packet radios in the field communicating with each other and to the ground, using the airborne systems to sew together fragments of Internet that had been segregated by a simulated nuclear attack.

The successful three-way interconnection of the ARPANET, PRNET, and SATNET represented the beginning of the Internet as an operational system. The design of the Internet made it possible for the networks to operate independently but still communicate, which benefited ARPA's experimental network projects. For instance, SATNET researchers could use the ARPANET to coordinate project personnel, monitor SATNET equipment, and generate test traffic; at the same time, SATNET remained a separate system from the ARPANET, which gave researchers the freedom to conduct possibly disruptive experiments on SATNET without disturbing ARPANET users (Jacobs et al. 1978, pp. 1462–1464). The ability of local networks to maintain their autonomy while participating in the Internet also made it easier to include networks from outside ARPA. After the demonstration, a number of new defense and research networks joined ARPA's evolving Internet, including the Defense Communications Agency's Experimental Data Network, the Army's Fort Bragg packet radio network, various Ethernet LANs at Xerox PARC, an experimental packet radio network at BBN, the network of MIT's Laboratory of Computer Science, and the British Post Office's Experimental Packet Switching System (Cerf and Kirstein 1978, p. 302).

To encourage sites to adopt TCP, ARPA began funding implementations of it for various operating systems. In 1977, ARPA funded BBN to incorporate TCP/IP into the popular Unix operating system, and one of the system's creators, Bill Joy, added TCP/IP to the Berkeley version of Unix.[18] ARPA also funded implementations for IBM machines, for the DEC TOPS-20 system, and for other operating systems (McKenzie 1997).

The ARPANET, however, did not adopt TCP/IP immediately. ARPA managers encouraged host sites to implement the new protocol, but did not force them to do so. Most sites chose to continue using NCP: the old protocol was providing perfectly adequate service within the ARPANET, and researchers who were not actively involved in internetworking experiments had no immediate motivation to switch protocols. Implementing TCP was difficult; to make matters worse, the specification kept changing as the Internet team adopted new ideas and as experimental use revealed shortcomings in the design. It was not ARPA's research community, therefore, that pushed for the transition from the ARPANET to the Internet.

Military Involvement in the Internet

The impetus for adopting TCP/IP came from the operational branches of the military (the armed forces and the agencies that support their day-to-day operations). Not all commanders were eager to adopt ARPA's new networking techniques, and there was often a clash of cultures between the ARPANET's research and military communities. But a combination of circumstances caused the Defense Communications Agency, which provided communications services for the armed forces, to view the ARPANET as an important part of its own system-building plans. As the DCA began to depend on the ARPANET, its managers took an active role in guiding the system's technical evolution and eventually championed full adoption of the Internet protocols.

The operational defense agencies first became interested in the ARPANET as a model for replacing their existing networks with more advanced technology. The National Security Agency commissioned Bolt, Beranek and Newman to create two smaller versions of the ARPANET for the intelligence community,[19] and the Defense Communications Agency experimented with ARPANET technology as part of its plans to upgrade the WorldWide Military Command and Control Systems.

In 1962, during the Cuban Missile Crisis, President Kennedy had discovered that the US military did not have an effective worldwide communications system for command and control. To remedy this, the Department of Defense had initiated the WorldWide Military Command and Control Systems, which built on an earlier system devised for the Strategic Air Command.[20] WWMCCS consisted of a hierarchy of command and control centers around the world that were equipped with computer systems to gather data on the status of forces and to store war plans (Zraket 1990). But the initial communications system for WWMCCS, which used leased lines to connect the computers, was far from satisfactory. In fact, since the modems used were so slow, the personnel at the centers often found it quicker to put data on a tape and transport it on an airplane than to transfer it over the phone lines (Eric Elsam, telephone conversation with author, 22 July 1997).

The Defense Communications Agency was eager to try more advanced technology for its new WWMCCS network, called WIN. In 1972 the agency contracted with Bolt, Beranek and Newman for a three-IMP network called PWIN that was used to develop software and test operations for WIN (US Congress 1972, p. 822). After a successful demonstration of PWIN, the DCA built the operational WIN network. In addition to transferring techniques, hardware, and skilled personnel from the ARPANET into a new military project, the WIN project convinced a number of people at the DCA that packet switching represented the future of data communications.[21]

Soon after beginning the WIN network, the DCA took on a new and unexpected role as the ARPANET's operator. ARPA, as a small research agency, was not well suited to provide routine data communications services. Once the ARPANET had passed the experimental stage, ARPA began looking for a new operator. In 1972, ARPA and BBN began to consider transferring the ARPANET to another government agency or a commercial carrier, with the hope that it would grow into a nationwide public service (Ornstein et al. 1972, p. 253; McQuillan et al. 1972, p. 752).[22] After discussing the matter with the Federal Communications Commission and other agencies, ARPA's managers decided to find a commercial operator who would buy the network hardware from ARPA, receive an FCC license as a specialized common carrier, and supply communications services to the government and other customers (US Congress 1972, p. 822). A 1974 report that ARPA commissioned from Paul Baran concurred that moving network operations to competitive commercial suppliers would stimulate the US networking industry and make it easier for military and civilian users to share the use of the ARPANET (Kuo 1975, p. 13).[23]

Many members of the ARPANET community, including Robert Kahn, Lawrence Roberts, and Howard Frank, took part in the effort to find a new operator. AT&T, the largest telecommunications carrier in the United States, seemed the most likely candidate. Roberts and Frank met with AT&T managers to explain how the network could be scaled up for commercial use (Roberts 1978, p. 49; Frank 1990, pp. 26–27; Kleinrock 1990, p. 36). AT&T declined, perhaps because the packet switching business was too small and too different from conventional telecommunications to seem worth its while. In 1975, after lengthy discussions among Roberts, ARPA director George Heilmeier, and other DoD personnel, the ARPA managers decided to temporarily transfer operational responsibility for the ARPANET to the Defense Communications Agency. ARPA would continue to provide funding and technical direction, and access would be open to DoD users and to government contractors approved by the DCA. The agreement left the fate of the network after three years unresolved, since ARPA still hoped to find a home for the ARPANET outside the government; nonetheless, the DCA ended up operating the network well beyond the initial three-year period.

It may seem unusual that the operational branches of the military took so little heed of the ARPANET before 1975. But ARPA played an unusual role in the Department of Defense: ARPA's whole purpose was to pursue research projects that were far ahead of the contemporary state of the art and were not tied too closely to specific applications (such as weapons systems). This role freed ARPA to look beyond the immediate concerns of the armed forces, but it also meant that ARPA sometimes had to work hard to get other military agencies interested in its innovations. It was only after the DCA took over operation of the ARPANET that the network began to be used by the armed forces in any extensive way.

The ARPANET as a Defense System

After officially assuming control of the ARPANET on 1 July 1975, the Defense Communications Agency began to reorient the network away from its research origins and toward routine military operations. Military users made increasing use of the network now that they could arrange connections through the DCA. "It was their normal way," Kahn (1990, p. 40) observed; "they didn't have to deal with a research agency." This effort to transfer ARPA's network technology to the various commands was aided by a personnel change within the Information Processing Techniques Office. Army Colonel David Russell became director of IPTO a few months after the DCA took over the ARPANET, and he helped to accelerate contacts with the armed forces and to promote the ARPANET as a test bed for new computerized command and control systems (Kahn 1990, p. 38; Klass 1976, p. 63).[24] By 1976, the Air Force, the Navy, and the Army were all using the ARPANET to experiment with such systems.[25]

The DCA imposed its own style of management on the ARPANET. The DCA's ARPANET manager, Major Joseph Haughney, commented:

> When the network was small, a decentralized management approach was established due to the nature of the network and the small community of users. This promoted flexibility and synergy in network use. Now that the network has grown to over 66 nodes and an estimated four to five thousand users, flexibility must be tempered with management control to prevent waste and misuse. (Haughney 1980a)

Under the new regime, prospective sites had to go through a more involved process to get access to the ARPANET; Bolt, Beranek and Newman and other contractors had to work through bureaucratic rather than informal channels; and there were more rules for what could or could not be done with the network (McKenzie 1997). The DCA was more serious than ARPA had been about preventing use of the network for "frivolous" activities, even if these activities did not disrupt network operations. For instance, in March of 1982 the DCA's new ARPANET manager, Major Glynn Parker, complained about an "email chain letter" that had been circulating on the network and threatened to cut off hosts whose users forwarded the letter (Parker 1982a). The DCA also wanted to cut down on the computer scientists' common practice of copying files across the network without their owners' explicit permission, which had become an accepted way for users to share the latest improvements in network software. The DCA expressed concern that government information might be inappropriately released to the general public or sold to industry, and it instituted a new policy that a file's owner had to give explicit consent before any copying could be done (Haughney 1981b).

As military use of the ARPANET grew, the DCA also tried to enforce the network's access policy, which many researchers felt had been more honored in the breach. In theory, the benign neglect of access controls that had allowed system administrators to turn a blind eye to unauthorized users would no longer be tolerated. Frequent reminders in the DCA's online newsletter that "all unauthorized use of the ARPANET is prohibited"[26] suggest that local administrators were not quick to enforce this policy, however. Some host administrators did not even know who all their ARPANET users were, since their computers were not set up to control which users could access the network. Haughney (1981b) warned these administrators that they would have to start monitoring or restricting access to their machines:

> If unauthorized users are found on the net because of a weak or nonexistent host access control mechanism, we will review the host's access mechanisms and request improvements. If the host refuses a review or refuses to make the suggested improvements, we will take action to terminate its network access. This is a club of last resort, but we will use it to protect other network users who have invested time and money to bring their controls up to par.

Haughney presented these measures as necessary to protect the military computer systems from malicious infiltration, stressing that the aim of the new access controls was "to ensure that we can verify proper resource utilization and prevent unauthorized penetrations" (1980a).

The DCA's heightened concern with network security was a response to wider trends in computing in the 1970s. In January of 1975, only a few months before the DCA assumed control of the ARPANET, the world's first personal computer was introduced in the United States. The Altair 8800 was made by a small company called Micro Instrumentation Telemetry Systems and advertised in the magazine *Popular Electronics*. It was primitive, and it was sold as a kit, but its price was astonishingly low: $379. The Altair 8800 was an instant hit with amateur computer enthusiasts, who place thousands of orders during the first few months it was advertised. Suddenly, a technology that had been restricted to authority figures in academia, business, and government was in the hands of teenage hobbyists. Members of a new "hacker" subculture quickly made improvements to the Altair and began devising more user-friendly machines, and by the late 1970s there was a thriving market for personal computers.

The spread of computer expertise to a much wider segment of the American population increased the risk that hackers would be able to break into restricted military systems on the ARPANET. Computer administrators had only to look to the telephone system for an example of the type of "unauthorized penetrations" Haughney was worried about. The 1970s saw the widespread use of "blue boxes"—devices that mimicked the control tones used by the telephone system—to fraudulently obtain free phone service (AT&T Bell Laboratories 1982, pp. 430, 432). A number of highly publicized incidents dramatized how pranksters known as "phone phreaks" used blue boxes to make free calls all over the world, often just for the challenge of mastering the telephone system. Phone phreaks came from the same world of young, undisciplined technophiles as computer hackers; for instance, before Steve Jobs and Stephen Wozniak started Apple, they had been in the business of making and selling blue boxes (Campbell-Kelly and Aspray 1996, p. 244). Haughney (1981a) warned ARPANET host managers that "the advent of lowcost, home computer systems has subjected the ARPANET to increased probing by computer freaks."

DCA managers were particularly concerned about the TIPs—the network nodes that allowed users to reach the ARPANET by dialing up from a terminal rather than having to go through a host computer. Initially, anyone with a terminal and the telephone number of the local TIP could use the ARPANET. To increase security, the DCA instituted a new system of logins and passwords to ensure that only authorized TIP users would have access to the network.

Another unforeseen set of circumstances spurred the DCA to become involved in the Internet effort. In 1976, DCA managers decided to procure from Western Union an upgraded data network to replace the outdated AUTODIN (Automatic Digital Network), a message switching network that the DCA had built for the military in the early 1960s. The new network, called AUTODIN II, was meant to replace AUTODIN, WIN, and the military sites on the ARPANET. AUTODIN II was slated to go into operation late in 1979 and would connect some 160 host computers and 1300 terminals (Kuo 1978, p. 309).[27] The DCA considered dismantling the ARPANET once AUTODIN II had been constructed, but the agency eventually concluded that there was still a role in the Department of Defense for a research-oriented network. Disruptive experiments would clearly be out of place on an operational military network such as AUTODIN II, and they would be just as unwelcome on a commercial data network, which would be the main alternative for researchers if the ARPANET were dismantled. Therefore, the DCA planned to leave the research portion of the ARPANET intact and to set up gateways to connect it to AUTODIN II (Kuo 1978, p. 312).

The DCA's decision to create an internetwork link between the ARPANET and AUTODIN II meant that the agency suddenly had a need for ARPA's new Internet protocols. Kahn and Cerf had been actively promoting TCP/IP as a potential standard for DoD networks, and in 1980 the Office of the Secretary of Defense formally adopted the ARPA protocols—which were still somewhat experimental—as military standards (Kahn 1990; Parker and Cerf 1982; Cerf 1980, p. 11). As Kahn (1990) explained, TCP/IP was the only system available that could meet the DoD's needs:

> We needed to switch over to the internet protocol because connections between multiple nets needed an internet protocol. . . . The sweep of events at the time was such that DoD really had to decide what guidance to give people who were connecting their computers to the net as newer sites came in. "What do we tell them?" So they finally decided to standardize [TCP/IP], because it was really the only game in town at that point.

It was these pragmatic considerations, rather than any demand from the research community, that drove the DoD to take the first decisive step toward making TCP the standard for ARPANET hosts.

By 1981 the armed services—which were being asked to pay for and use the new system—were complaining that AUTODIN II was too expensive and technically deficient. Don Latham, the Assistant Secretary of Defense for "C3I" (Command, Control, Communications, and Intelligence), asked the DCA to come up with an alternative, but the agency was not able to do so. Latham then appointed Colonel Heidi B. Heiden, who had been the Army's planning director for computer communications, to join the DCA and put together a team to come up with an alternative network design. The DoD did not immediately abandon the AUTODIN II effort; rather, it gave Western Union and Heiden's team six months to prepare their systems for evaluation by a DoD review board, which would choose one (Heidi Heiden, telephone conversation with author, 30 July 1997).

Heiden did not want to build new hardware, as the AUTODIN II group was doing. His plan was to use the DoD's existing packet switching networks—the ARPANET, WIN, and MINET (a version of the ARPANET used in Europe)—as the basis for a new Defense Data Network (Heiden and Duffield 1982; Harris et al. 1982). He wanted to use commercial technology wherever possible, to cut development costs and to give the DoD competing sources for its components. He believed that ARPA's Internet protocols would provide the best service, and he defended them against rival standards, such as the protocols that were then being developed by the International Organization for Standardization (Heiden, telephone conversation, 30 July 1997). In April of 1982 the review board chose Heiden's Defense Data Network plan over AUTODIN II, putting the ARPANET back at the center of the DoD's networking plans (Parker 1982b).

The Transition to TCP/IP

Since the Internet protocols were to serve as the common language for the new Defense Data Network, it became imperative that the ARPANET sites adopt TCP/IP and retire the older and more widely used NCP. After the Internet protocols had been successfully tested on the ARPANET, they would be introduced on the other participating defense networks. In March of 1981, Major Joseph Haughney announced that all ARPANET hosts would be required to implement TCP/IP in place of NCP by January of 1983 (Haughney 1980a, 1980c, 1981a). His successor, Major Glynn Parker, commented on this decision: "Just as it did a decade ago, the ARPANET community is leading the way into a new networking territory of great importance to the future of US military command and control systems." (Parker and Cerf 1982)

The reality beneath Parker's inspiring words was that the DCA and ARPA were forcing a traumatic upheaval in the ARPANET community. Most host sites were still relying on NCP, and converting to the new Internet protocols proved to be an enormous effort. "The transition from NCP to TCP was done in a great rush," one participant recalled, "occupying virtually everyone's time 100% in the year 1982. *Nobody* was ready. . . . It was a major painful ordeal." (Crispin 1991) Dan Lynch, a computer systems manager, recalled: "Dozens of us system managers found ourselves on a New Year's Eve trying to pull off this massive cutover. We had been working on it for over a year. There were hundreds of programs at hundreds of sites that had to be developed and debugged." (Lynch 1991) Lynch made up buttons that read "I Survived the TCP Transition" and passed them out to his colleagues (ibid.). Alex McKenzie of BBN agreed that there had been a "mad rush at the end of 1982" to make the deadline (McKenzie 1991b). Clearly the transition to the Internet protocols would not have occurred so quickly—perhaps not at all at many sites—without considerable pressure from the military managers.

Most host system managers had no compelling interest in converting to the Internet protocols, and the transition required a number of steps that would cost the host sites time and money. Haughney warned the ARPANET sites in July of 1980: "Unless you have already begun development of the protocols, you may want to start budgeting for the protocol software development for your host" (Haughney 1980a). The first transition occurred in January of 1981, when the new Internet packet format, with 96-bit rather than 32-bit headers, came into use. Hosts had to make sure that all their network applications produced packets with the new headers; if not, they would be unable to use the ARPANET as of January 1981. The next step would be writing TCP software for each type of host computer—which, as the earlier efforts to implement TCP had shown, was no easy task. Hosts would also have to adopt updated versions of the applications protocols ftp and telnet, a new mail standard called Simple Mail Transfer Protocol (SMTP), and a new addressing scheme for mail (Feinler 1982b).[28] At the same time, the sites had to replace their old IMPs or TIPs with new versions designed by BBN to run the Internet protocols (Haughney 1980a; Parker 1982b).[29]

To support inter-network routing, the Internet needed a name server—a database of host names that, when queried with the name of a host, would supply the host's network address. The name server was created by a large group of ARPANET members and went into service at the Network Information Center at SRI in July of 1982.[30] Since NCP and TCP were incompatible, some sites ran both protocols and acted as translators between TCP and NCP hosts during the transition period (McKenzie 1997). Until 1 January 1983 both protocols would also be accepted by the IMPs, but after that date BBN would set the IMPs to reject packets that used the NCP format.

When the cutoff date arrived, only about half the sites had actually implemented a working version of TCP/IP. IPTO director Robert Kahn recalled:

> The biggest problem was just getting people to believe that it was real. . . . We sent messages to everybody, alerting them to the timing and yet one week

> before we were still getting messages, "Is this really going to happen next week?" or "Let us know if you decide to really go ahead with this." (Kahn 1990)

Those who had not created the necessary software for their computers were unpleasantly surprised when BBN upheld the ARPA-DCA policy and cut them off from the network. In addition, many sites that had tried to convert to TCP discovered errors in their implementations and were forced to revert to NCP (Heiden 1983a). Kahn (1990) recalled that it took a long time for all the sites to get their new TCPs working properly:

> Managing it was traumatic for a while. I mean, the phone was ringing off the hook every few minutes. Every day someone new would complain, "I used to be able to do this, and now I can't." Shaking it all down was also a problem. Even the places that thought they were going to convert properly suddenly found that while theirs worked with the three or four places that they thought it would, or had tried it out with, it didn't work with some others.

To keep the network running, host sites with nonfunctional TCPs were temporarily allowed to run NCP while they worked on the problem. Any site that had not converted to TCP/IP by the cutoff date was required to submit a request for an exception, justify its failure to be ready, and set a schedule for converting (Heiden 1982). By March of 1983, when the next deadline arrived, about half of the remaining sites still did not have the new protocols running, and the routine was repeated. By June every host was running TCP/IP. A major milestone in the evolution of the Internet had been passed (Heiden, telephone conversation with author, 30 July 1997).

Steps toward a Civilian Internet

After converting the ARPANET to TCP/IP, the DCA and ARPA took two more steps that would help set the stage for the development of a large-scale civilian Internet.

One step was to segregate the ARPANET's military users and its academic researchers, who had been coexisting somewhat uneasily since the DCA's takeover of the ARPANET in 1975. The DCA and its military users were concerned that the academic sites could not or would not enforce strict access controls. One BBN manager put it this way: "The research people like open access because it promotes the sharing of ideas. . . . But the down side is that somebody can also launch an attack." (Broad 1983, p. 13) The DCA warned in 1982 that the ARPANET was increasingly vulnerable to "intrusion by unauthorized, possibly malicious, users . . . as the availability of inexpensive computers and modems have made the network fair game for countless computer hobbyists" (Harris et al. 1982, p. 78). To protect the military sites from this perceived threat, Heiden decided to split the ARPANET into two separate networks: a defense research network (still called ARPANET) and an operational military network (MILNET). The ARPANET would continue to be used to develop and test new networking technologies, while MILNET sites would be equipped with encryption devices and other security measures to support their military functions (Lukasik 1997). The decision to split the network was announced on 4 October 1982, and the MILNET was officially established on 4 April 1983 (Heiden 1983b). The actual physical separation of the two networks took a bit longer. Each host and IMP had to be assigned to either MILNET or ARPANET, and telephone links had to be rearranged so that only IMPs from the same network would be interconnected. A few hosts were attached to both networks to provide internetwork communications (Heiden 1983b). The new arrangement meant that the ARPANET was once again a research-oriented network dominated by universities. This would make it much easier to imagine transferring the network to civilian control.

The second step to was to commercialize the Internet technology. Heiden was eager to have commercial sources for Internet products. ARPA had already funded various contractors to write TCP implementations, most notably for the Unix operating system. Heiden stepped up this effort at technology transfer, setting up a $20 million fund to finance computer manufacturers to implement TCP/IP on their machines (Heiden, telephone conversation with author, 30 July 1997). All the major computer companies took advantage of this opportunity, and by 1990 TCP/IP was available for virtually every computer on the American market. This gave a tremendous momentum to the spread of the ARPA protocols, helping to ensure that they would become a de facto standard for networking.

Ushering in the Internet Era

In the period 1973–1983, ARPA created a new generation of technologies for packet radio, packet satellite, and internetworking. The ARPANET went through a number of transformations: the entire network community switched to TCP/IP, the military users were split off to their own network, and the ARPANET became part of a larger system—the

Internet—that encompassed a number of military and experimental networks. Owing in large part to ARPA's influence, the field of computer networking underwent a conceptual transformation: it was no longer enough to think about how a set of *computers* could be connected; network builders now also had to consider how different *networks* could interact. The dominant model for internetworking would be the system worked out by Vinton Cerf and Robert Kahn.

In the years since the Internet was transferred to civilian control, its military roots have been downplayed. However, it should not be forgotten that ARPA's new networking techniques were shaped in many ways by military priorities and concerns. Like the original ARPANET project, the radio, satellite, and Internet programs followed a philosophy of promoting heterogeneity and decentralization in network systems that mirrored the US military's diverse and scattered operations. The use of new communications media was meant to make it easier to tailor command and control systems to specific military environments, such as jeeps, ships, or airplanes. The idea that network protocols should be simple and adaptable derived in part from the military's continued concern with survivability. Even civilian developments were shaped by the military: Robert Metcalfe drew on the ARPA-funded Alohanet work in developing Ethernet, and Heidi Heiden funded the commercialization of TCP/IP. Finally, it was the determined efforts of DCA managers to get TCP/IP running throughout the ARPANET that set the stage for the emergence of a worldwide, publicly accessible Internet in the late 1980s.[31]

But military shaping is only part of the story. The Internet approach would not have been so influential had it not served the needs and interests of a diverse networking community. The Department of Defense could require the US armed forces to use the TCP/IP protocols, but it could not force others to adopt them. The ARPA system was not the only option available: by the mid 1970s, both computer manufacturers and telecommunications carriers were beginning to offer their own internetworking systems, which might have served as the basis for a worldwide network system. Cerf and Kahn's collaborative approach to system design helped ensure that TCP/IP would become the technology of choice. And, as Cerf (1990) observed, the Internet's ruggedness made it appealing for civilian as well as military applications:

> There were all kinds of challenges for this technology to overcome that were military in nature, that were problems that were caused by very hostile environments. Now as it has turned out, the robustness in the system has been helpful in the civilian sector, too. They may not be as dramatic, but a cable cut through an optical fiber line is just as devastating as nuking some central office somewhere, as far as communications is concerned.

By coordinating defense and civilian interests, the Internet's designers were able to create a system that would appeal to a broad spectrum of potential network builders.

The story of the Internet's origins departs from explanations of technical innovation that center on individual inventors or on the pull of markets. Cerf and Kahn were neither captains of industry nor "two guys tinkering in a garage." The Internet was not built in response to popular demand, real or imagined; its subsequent mass appeal had no part in the decisions made in 1973. Rather, the project reflected the command economy of military procurement, where specialized performance is everything and money is no object, and the research ethos of the university, where experimental interest and technical elegance take precedence over commercial application. This was surely an unlikely context for the creation of what would become a popular and profitable service. Perhaps the key to the Internet's later commercial success was that the project internalized the competitive forces of the market by bringing representatives of diverse interest groups together and allowing them to argue through design issues. Ironically, this unconventional approach produced a system that proved to have more appeal for potential "customers"—people building networks—than did the overtly commercial alternatives that appeared soon after.