

**The ARPA Network**  
by Lawrence.G.Roberts  
Advanced Research Projects Agency  
Barry D. Wessler  
University of Utah

## **INTRODUCTION**

One of the most successful aspects of the experiments in the use of time-shared computer systems conducted during the past decade was the ability to share computing resources among all the users of the system. Controlled sharing of data and software as well as the sharing of the time-sharing system hardware has led to much higher programmer productivity and better overall utilization of the computing and user's resources. One of the prime reasons for the success of resource sharing is the on-line dialog capability between the man and the machine which has permitted the user to experiment more easily with new systems. This experimentation with existing resources has fostered greater acceptance and use of these resources. In addition, in more advanced systems where interprocess communications is permitted existing programs were often incorporated directly into larger software systems.

These techniques have been especially productive in the systems programming area. One reason for its limited success in the applications area has been the lack of a large enough community (critical mass phenomenon) in a single application area using the same time-sharing system. Resource sharing for large communities has been attempted by transferring the programs or data physically from one machine to another. This can be done only by imposing restrictive language standards and using identical hardware systems. Maintenance and updating problems and small variations of hardware and operating systems have dampened the success of program transferability experiments.

A viable alternative to program transferability, while permitting full resource sharing, is to provide a communications system which will permit users to access remote programs or data as if they were local users to that system. In addition, it should be possible for a user to create a program on his local machine which could make use of existing programs in the network as if it were available on his local machine. Rather than trying to move the programs from machine to machine, the network would allow the user or his program to communicate with a machine on which the program already executes. If enough machines can be connected into such a network then the total community in any particular application area may be sufficiently large enough to reach critical mass. We believe the computer would then become a more powerful tool for its users because of the resource sharing capability.

This chapter describes the design criteria and implementation of a communications system capable of satisfying resource sharing communication requirements. The communication system interconnects autonomous, independent computer systems nationwide into a Computer Network so as to permit interactive resource sharing between any pair of systems.

## **DESIGN OF A NETWORK COMMUNICATIONS SERVICE**

After initial network experiments conducted between MIT's Lincoln Lab and the System Development Corporation, it was clear that a completely new communications service was required in order to make an effective, useful resource-sharing computer network. The communication pipelines offered by the carriers would probably have to be a component of that service but were clearly inadequate by themselves. What was needed was a message service where any computer could submit a message destined for another computer and be sure it would be delivered promptly and correctly. Each interactive conversation or link between two computers would have messages flowing back and forth similar to the type of traffic between a user console and a computer. Message sizes of from one character to 1000 characters are characteristic of man-machine interactions and this should also be true for that network traffic where a man is the end consumer of the information being exchanged. Besides having a heavy bias toward short messages, network traffic will also be diverse. ' With twenty'

**The ARPA Network**  
by Lawrence.G.Roberts  
Advanced Research Projects Agency  
Barry D. Wessler  
University of Utah

computers, each with dozens of time-shared users, there might be, at peak times, one or more conversations between all 190 pairs of computers.

### **Reliability**

Communications systems, being designed to carry very redundant information for direct human consumption, have, for computers, unacceptably high down-time and an excessively high error rate. The line errors can 'easily' be fixed through error detection and retransmission; however, this does require the use of some computation and storage at both ends of each communication line. To protect against total line failures, there should be at least two physically separate paths to route each message. Otherwise the service will appear to be far too unreliable to count on and users will continue to duplicate remote resources rather than access them through the net.

### **Responsiveness**

In those cases where a user is making more or less direct use of a complete remote software system, the network must not cause the total round- trip delay to exceed the human short-term memory span of one to two seconds.

Since the time-sharing systems probably introduce at least a one-second delay, the network's end-to-end delay should be less than 112 second. The network response should also be comparable, if possible, to using a remote display console over a private voice grade line where a 50 character line of text (400 bits) can be sent in 200 ms. Further if inter- active graphics are to be available, the network should be able to send a complete new display page requiring about 20 kilobits of information within a second and permit interrupts (10-100 bits) to get through very quickly, hopefully within 30-90 ms. Where two programs are interacting without a human user being directly involved, the job will obviously get through sooner, the shorter the message delay. There is no clear critical point here, but if the communications system substantially slows up the job, the user will probably choose to duplicate the remote process or data at his site. For such cases, a reasonable measure by which to compare communications systems is the "effective bandwidth" [data block length for the job /end-to-end transmission delay].

### **Capacity**

The capacity required is proportional to the number and variety of services available from the network. As the number of nodes increases, the traffic is expected to increase more than linearly, until new nodes merely duplicate available network resources. The number of nodes in the experimental network was chosen to:

1. involve as many computer researchers as possible to develop network protocol and operation procedures,
2. involve special facilities, such as the ILLIAC, to distribute its resources to a wider community,
3. involve as many disciplines of science as possible to measure the effect of the network on those disciplines, and
4. involve many different kinds of computers and systems to prove the generality of the techniques developed. The nodes of the network were generally limited to:
  - those centers for which the network would truly provide a cost benefit,

**The ARPA Network**  
 by Lawrence.G.Roberts  
 Advanced Research Projects Agency  
 Barry D. Wessler  
 University of Utah

- government-funded projects because of the use of special rate government-furnished communications, and
- ARPA-funded projects where the problems of intercomputer accounting could be deferred until the network was in stable operation.

The size of the experimental network was chosen to be approximately 15 nodes nationwide. It was felt that this would be large and diverse enough to be a useful utility and to provide enough traffic to adequately test the network communication system.

Although it is extremely difficult to estimate potential traffic levels in a resource-sharing environment one approach is to examine the traffic created by hardware service sharing; probably the first activity within the network. Hardware service sharing is where each node obtains a portion of its computer capability from each of several large service centers instead of operating its own less efficient general purpose system. The traffic generated by this type of activity can be estimated by examining the total input-output bandwidth to a user of a computer center, the traffic which would have to be transmitted via the network to remote users. This; includes traffic to line printers, user consoles, user tape units and input from the card reader, consoles, etc. Since line printers alone usually require 2-30 KB, the total I/O traffic is likely to be 2-50 KB.

**Cost**

To be a useful utility, it was felt that Communications costs for the network should be less than 20% of the computing costs of the systems connected through the network. . This is in contrast to the rising costs of remote access communications which often cost as much as the computing equipment.

If we examine why communications usually cost so much we find that it is not the communications channel per se, but our inefficient use of them, the switching costs, or the operations cost. To obtain a perspective on the price we commonly pay for communications let us evaluate a few methods. As an example, let us use a distance of 1400 miles since that is the average distance between pairs of nodes in the projected ARFA Net- work. A useful measure of communications cost is the cost to move one million bits of information, cents /megabit. In the table below this is computed for each media. It is assumed for-leased equipment and data set rental that the usage is eight hours per working day.

**Table 1. Cost per megabit for Various Communication Media 1400 Mile Distance Media**

Telegram	\$3300.00	For 100 words at 30 bits/wd, daytime
Night Letter	565.00	For 100 words at 30 bits/wd, overnight delivery
Computer Console	374.00	18 baud avg. use2, 300 baud DDD service line & data sets only
TELEX	204.00	50 baud teletype service

**The ARPA Network**  
 by Lawrence.G.Roberts  
 Advanced Research Projects Agency  
 Barry D. Wessler  
 University of Utah

DDD (103A)	22.50	300 baud data sets, DDD daytime service
AUTODIN	8.20	2400 baud message service, full use during working hours
DDD (202)	3.45	2000 baud data sets
Letter	3.30	Airmail, 4 pages, 250 wds/pg, 30 bits/wd
W.U. Broadband	2.03	2400 baud service, full duplex
WATS	1.54	200 baud, used 8 hrs/working day
Leased Line (201)	.57	2000 baud, commercial, full duplex
Data 50	.47	50 KB dial service, utilized full duplex
Leased Line (303)	.23	50 KB, commercial, full duplex
Mail DEC Tape	.20	2.5 megabit tape, airmail
Mail IBM Tape	.034	100 megabit tape, airmail

Special care has also been taken to minimize the cost of the multiplexor or switch. Previous store and forward systems like DoD's AUTODIN system have had such complex, expensive switches that over 950/q of the total communications service cost was for the switches. Other switch services adding to the system's cost, deemed superfluous in a computer network, were: long term message storage, multi-address messages and individual message accounting.

The final cost criteria was to minimize the communications software development cost required at each node site. If the network software could be generated centrally, not only would the cost be significantly reduced, but also the reliability would be significantly enhanced.

**THE ARPA NETWORK**

Three classes of communications systems were investigated as candidates for the ARPA Network: fully interconnected point to point leased lines, line switched (dial-up) service, and message switched (store and forward) service. For the kind of service required, it was decided and later verified that the message switched service provided the greater flexibility, higher effective bandwidth, and lower cost than the other two systems.

**The ARPA Network**  
by Lawrence.G.Roberts  
Advanced Research Projects Agency  
Barry D. Wessler  
University of Utah

The standard message switched service uses a large central switch with all the nodes connected to the switch via communication lines; this configuration is generally referred to as a Star. Star systems perform satisfactorily for large blocks of traffic (greater than 100 kilobits per message), but the central switch saturates very quickly for small message sizes. This phenomenon adds significant delay to the delivery of the message. Also, a Star design has inherently poor reliability since a single line failure can isolate a node and the failure of the central switch is catastrophic.

An alternative to the star, suggested by the Rand Study "On Distributed Communications" is a fully distributed message switched system. Such a system has a switch or store and toward center at every node in the network. Each node has a few transmission lines to other nodes; messages are therefore routed from node to node until reaching their destination. Each transmission line thereby multiplexes messages from a large number of source-destination pairs of nodes. The distributed store and forward system was chosen, after careful study, as the ARFA Network communications system. The properties of such a communication system are described below and compared with other systems.

A more complete description of the implementation, optimization and initial use of the network can be found in a series of five papers presented at the 1970 Spring Joint Computer Conference. The first paper, by Roberts and Wessler, from which this chapter was extracted in part presents an overview of the need for a computer network, communication system requirements, properties of the ARPA Network, and the overall cost factors. The second paper by Heart, et al<sup>5</sup> describes the design, implementation and performance characteristics of the message switch. The third paper by Kleinrock<sup>6</sup> derives procedures for optimizing the capacity of the transmission facility in order to minimize cost and average message delay. The fourth paper by Frank, et al<sup>7</sup> describes the procedure for finding optimized network topologies under various constraints. The last paper by Carr, et al<sup>8</sup> is concerned with the system software required to allow the network computers to talk to one another. This final paper describes a first attempt at Intercomputer protocol, which is expected to grow and mature as we gain experience in computer networking.

### **Network Properties**

The switching centers use small general purpose computers called Interface Message Processors (IMPs) to route messages, to error check the transmission lines and to provide asynchronous digital interface to the main (HOST) computer. The (IMPs) are connected together via 50 Kbps data transmission facilities using common carrier (ATT) point to point leased lines. The topology of the network transmission lines was selected to minimize cost, maximize growth potential, and yet satisfy all the design criteria.

### **Reliability**

The network specification requires that the delivered message error rates be matched with computer characteristics, and that "the down-time of the communications system be extremely small. Three steps have been taken to insure these reliability characteristics:

- at least two trans - mission paths exist between any two nodes,
- a 24 bit cyclic check sum is provided for each 1000 bit block of data and
- the IMP is ruggedized against external environmental conditions and its operation is independent of any electromechanical devices (except fans).

**The ARPA Network**  
by **Lawrence.G.Roberts**  
**Advanced Research Projects Agency**  
**Barry D. Wessler**  
**University of Utah**

So far, the down-time of the transmission facility has averaged 2.3% outage for each line, however, the duplication of paths reduce the average down-time between any pair of nodes, due to transmission failure, to approximately 0.4%. The cyclic check sum was chosen based on the performance: characteristics of the transmission facility; it is designed to detect long burst errors. The code is used for error detection only, with retransmission on an error. This check reduces the undetected bit error rate to one in 10<sup>12</sup> or about one undetected error per year in the entire network.

The ruggedized IMP was chosen to provide as high a reliability as possible; however, solid reliability data is not available yet. The elimination of mass storage devices from the IMP results in lower cost' less down-time; and greater throughput performance of the WP, but implies no long term message storage and no message accounting by the IMP. If these functions are later needed they can be added by establishing a special node in the net- work. This node would accept accounting information from all the IMPs and also could be routed all the traffic destined for HOSTs which are down. We do not believe these functions are necessary, but the network design is capable of providing them.

### **Responsiveness**

The target goal for responsiveness was .5 seconds transit time from any node to any other, for a 1000 bit (or less) block of information. Actual response averages .1 sec for 1000 bit blocks and .3 sec for 8000 bit messages for all traffic levels less than saturation. After saturation the transit time rises quickly because of excessive queuing delays. However, saturation is avoided by the net acting to choke off the inputs for short periods of time, reducing the buffer queues while not significantly increasing the delay.

### **Capacity**

The capacity of the network is the throughput rate at which saturation Occurs. The saturation level is a function of the topology and capacity of the transmission lines and the average size of the blocks sent over the transmission lines. The analysis of capacity was performed by Network Analysis Corporation during the optimization of the network topology. As that analysis shows, the network has the ability to flexibly increase its capacity by adding additional transmission lines. The use of 230.4 KB communication services, where appropriate, considerably improves the cost-performance of the network.

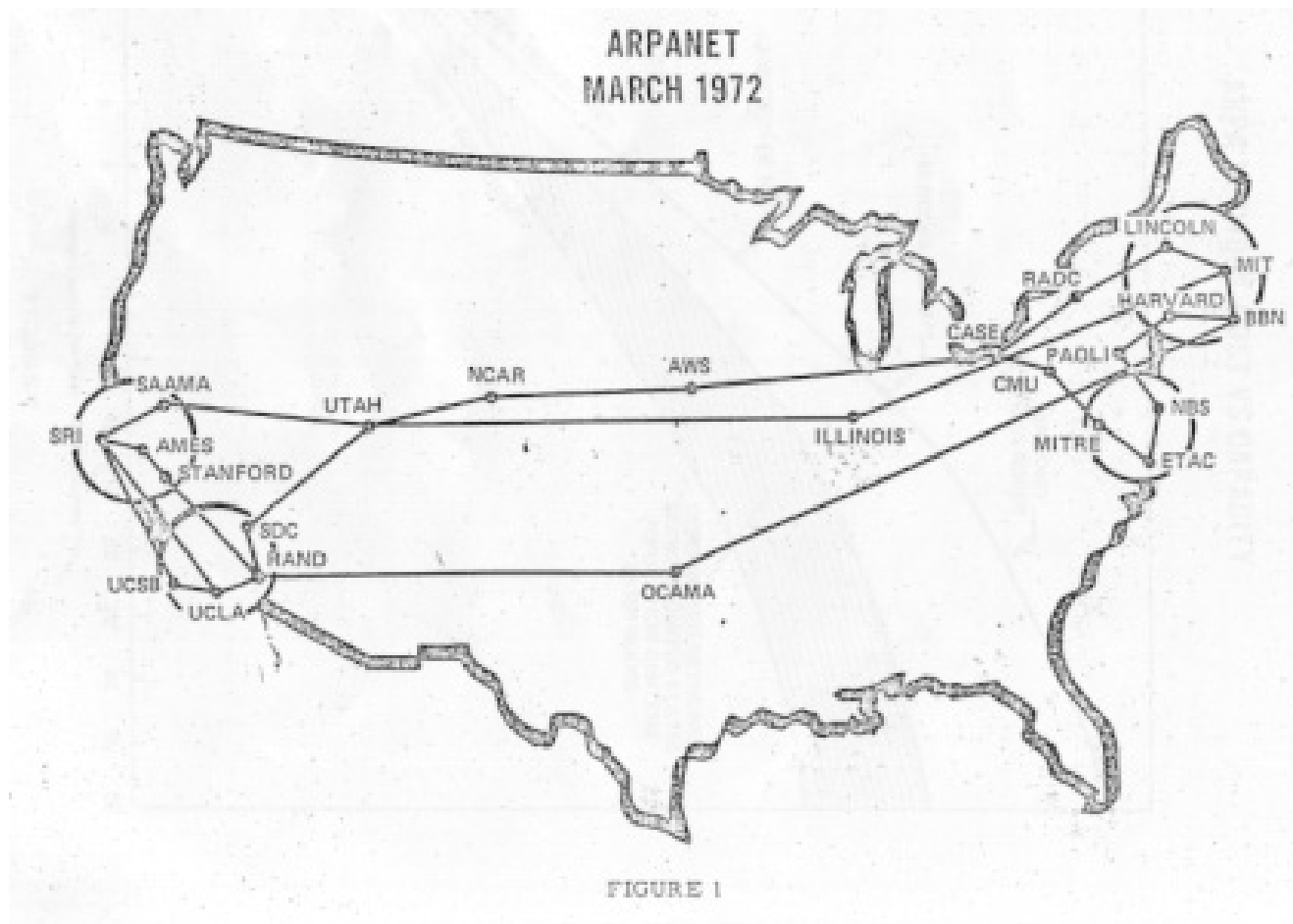
### **Configuration**

The flexibility in topology of communication lines permitted in the ARPA Network allows controlled growth in capacity as the nodes increase their use of the Net. The topology can vary from a simple loop connecting all nodes once to a densely connected system such as the network shown in Figure 1. This fact suggests that Network users can be charged on the basis of use, since a heavy user will cause the network manager to put in more lines increasing the overall cost of the network. An equitable basis for charging has been found for the currently planned 20 nodes net. Although current users are not yet being charged, the charging policy is useful for planning purposes and will be implemented later, as the network size and use expands. The basis of the charging scheme is shown in Figure 2 where the cost per node is plotted against the average capacity per node. Data for this graph was obtained from Dr. H. Frank<sup>7</sup>, Network Analysis Corporation, and is based on dozens of pos Bible optimal network topologies produced by computer analysis. The slope of the ARPANET line represents the incremental cost for increasing the Net's capacity. Dr. Frank also determined that the traffic need not be uniformly distributed among the nodes in order to achieve the total traffic capacity capability. Large variations in the distribution of average traffic among the nodes will result in

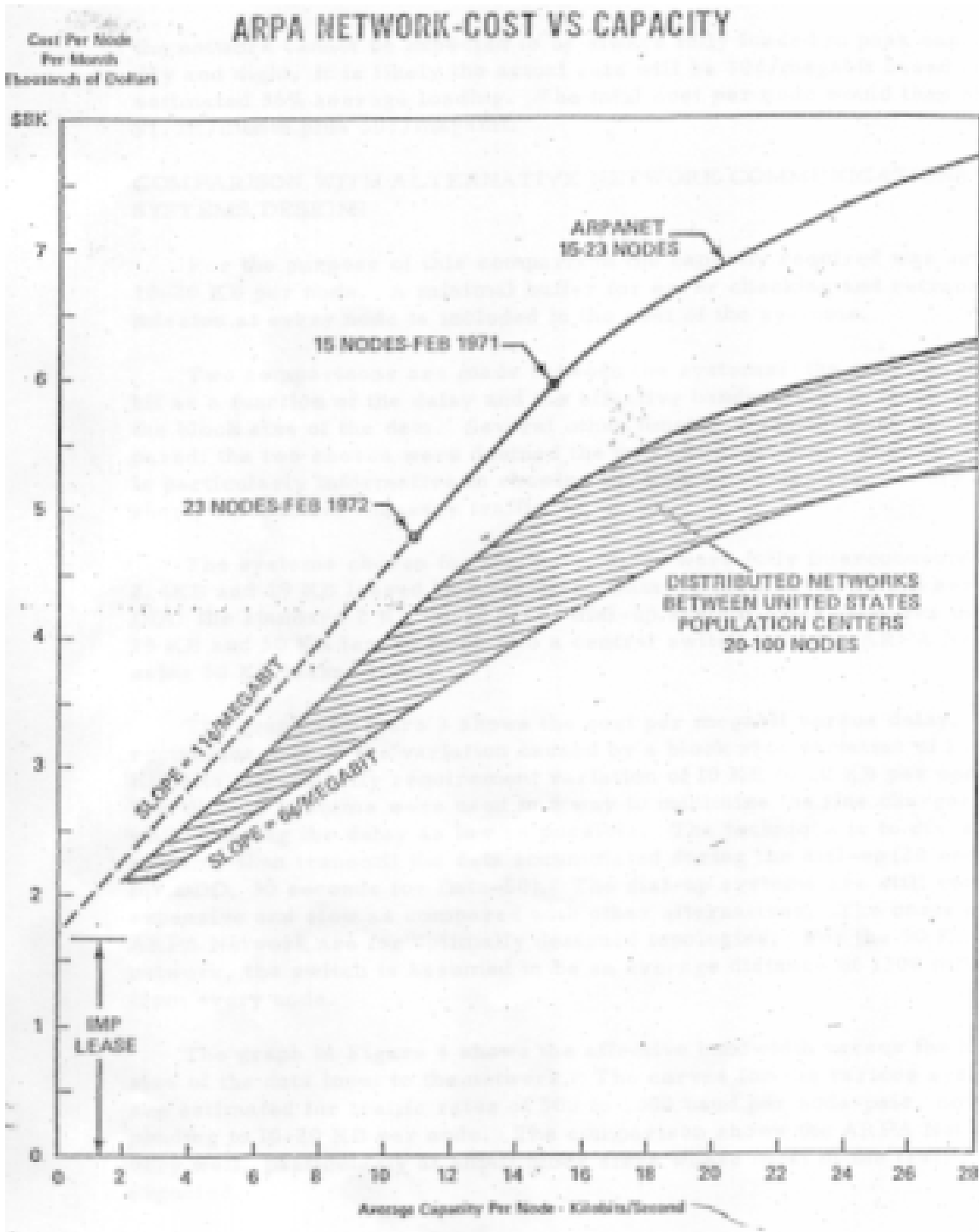
**The ARPA Network**  
by Lawrence.G.Roberts  
Advanced Research Projects Agency  
Barry D. Wessler  
University of Utah

saturation occurring at approximately the same total traffic for the same topology as if the traffic were uniformly distributed. The effect of one node reducing or increasing its traffic can therefore be found by reducing or increasing the total network traffic by that amount.

From Figure 2 it can be seen that the cost of increasing the capacity results in an equivalent transmission cost of 11c/megabit. However, since the network cannot be expected to be always fully loaded to peak capacity, day and night, it is likely the actual rate will be 30f /megabit based on an estimated 36% average loading. The total cost per node would then be \$1.7K/month plus 30f/megabit.



**The ARPA Network**  
 by Lawrence.G.Roberts  
 Advanced Research Projects Agency  
 Barry D. Wessler  
 University of Utah



**COMPARISON WITH ALTERNATIVE NETWORK COMMUNICATIONS SYSTEMS DESIGNS**

For the purpose of this comparison the capacity required was set at 10-20 KB per node. A minimal buffer for error checking and retransmission at every node is included in the cost of the systems.

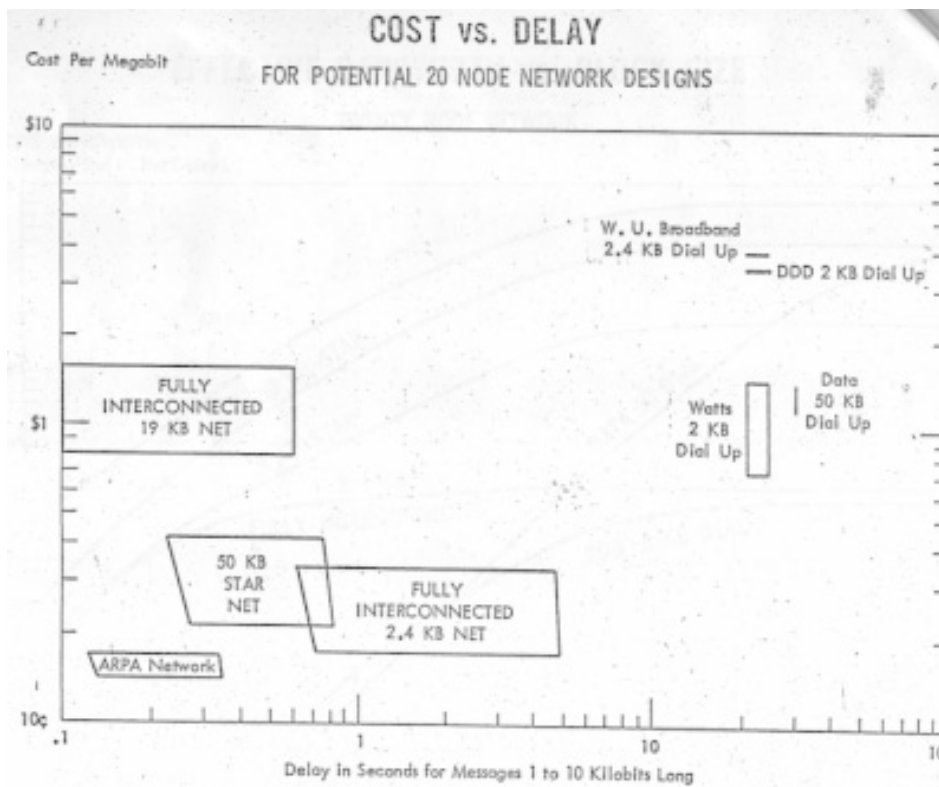
**The ARPA Network**  
 by Lawrence.G.Roberts  
 Advanced Research Projects Agency  
 Barry D. Wessler  
 University of Utah

Two comparisons are made between the systems: the cost per mega-bit as a function of the delay and the effective bandwidth as a function of the block size of the data. Several other functions were plotted and compared; the two chosen were deemed the most informative. The latter graph is particularly informative in showing the effect of using the network for short, interactive message traffic.

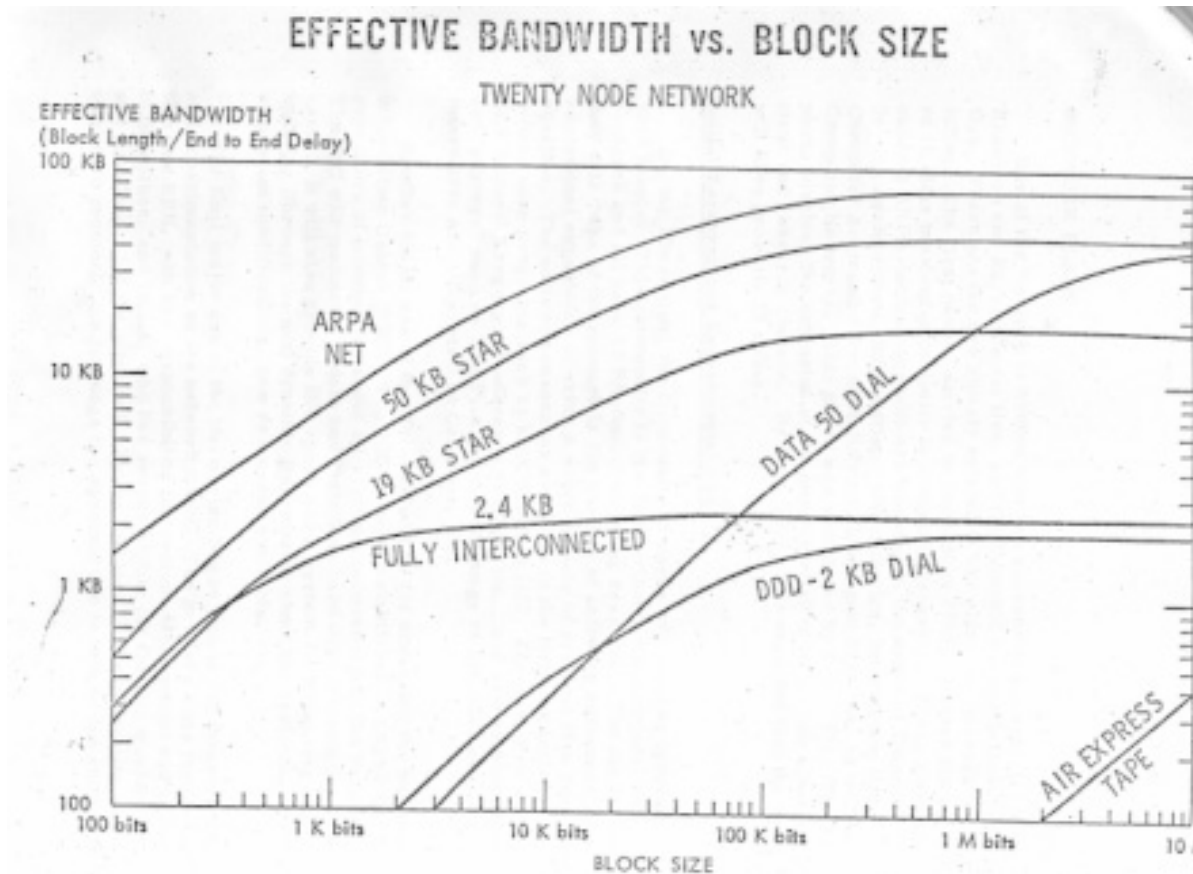
The systems chosen for the comparison were fully interconnected 2.4KB and 19 KB leased line systems, Data-50 the dial-up 50 KB service, DDD the standard 2 KB voice grade dial-up system, Star networks using 19 KB and 50 KB leased lines into a central switch and the ARFA Network using 50 KB leased lines.

The graph in Figure 3 shows the cost per megabit versus delay. The rectangles outline the variation caused by a block size variation of 1 to 10 Kilobits and capacity requirement variation of 10 KB to 20 KB per node. The dial-up systems were used in a way to minimize the line charges while keeping the delay as low as possible. The technique is to dial a system, then transmit the data accumulated during the dial-up (20 seconds for DDD, 30 seconds for Data-50). The dial-up systems are still very expensive and slow as compared with other alternatives. The costs of the ARPA Network are for optimally designed topologies. For the 50 KB Star network, the switch is assumed to be an average distance of 1300 miles from every node.

The graph in Figure 4 shows the effective bandwidth versus the block size of the data input to the network. The curves for the various systems are estimated for traffic rates of 500 to 1000 baud per node-pair, corresponding to 10-20 KB per node. The comparison shows the ARPA Net does very well, particularly at small block sizes where most of the traffic is expected.



**The ARPA Network**  
by Lawrence.G.Roberts  
Advanced Research Projects Agency  
Barry D. Wessler  
University of Utah



### NETWORK PLANS

Use of the Network is broken into two successive phases:

- Initial Research and Experimental Use, and
- External Research Community Use.

These phases are closely related to our plans for Network implementation. The first phase, started in September 1969, involves the connection of 15 sites involved principally in computer research. These sites are current ARPA contractors who are working in the areas of Computer System Architecture, Information System Design, Information Handling, Computer Augmented Problem Solving, Intelligent Systems, as well as Computer Networks. This phase was completed by Feb. 1971. The second phase involves the extension of the number of sites to include other applications and research centers. By Feb. 1972, it is expected that the network will have grown to 23 nodes.

### Initial Research and Experimental Use

During Phase One, the community of users will number approximately 2000 people. This community is involved primarily in computer Science research and all have ARPA-funded on-going research. The

**The ARPA Network**  
by **Lawrence.G.Roberts**  
**Advanced Research Projects Agency**  
**Barry D. Wessler**  
**University of Utah**

major use they will make of the network is the sharing of software resources and the educational experience of using a wider variety of systems than previously possible. The software resources available to the Network include: advanced user programs such as MA THLAB at MIT, Theorem Proves at SRI, Natural Language Processors at BBN, etc., and new system software and languages such as LEAP, a graphic language at Lincoln Lab. LC2, an interactive ALGOL system at Carnegie. etc. Another major use of the Network will be for accessing the Network Information Center (NIC). The NIC is being established at SRI as the repository of information about all systems connected into the Network. The NIC will maintain, update and distribute hard copy information to all users. It will also provide file space and a system for accessing and updating (through the net) dynamic information about the systems, such as system modifications, new resources available, etc.

The final major use of the Net during Phase One is for measurement and experimentation on the network itself. The primary sites involved in this are BBN, who has responsibility for system development and system maintenance, and UCLA, who has responsibility for the Net measurement and modeling. All the sites will also be involved in the generation of intercomputer protocol, the language the systems use to talk to one another.

### **External Research Community Use**

During the time period after August 1971, additional nodes will be installed to take advantage of the Network in other ARPA-funded research disciplines: Behavioral Science, Climate Dynamics and Seismology. The use of the Network at these nodes will be oriented more toward the distribution and sharing of stored data, and in the latter two fields the use of the ILLIAC IV.

The data sharing between data management systems or data retrieval systems will begin an important phase in the use of the Network. The concept of distributed data bases and distributed-access to the data is one of the most powerful and useful applications of the network for the general data processing community. As described above, if the Network is responsive in the human time frame, data bases- can be stored and maintained at a remote location rather than duplicating them at each site the data is needed. Not only can the data be accessed as if the user were local, but also as a Network user he can write programs on his own machine to collect data from a number of locations for comparison, merging or further analysis.

Because of widespread use of the ILLIAC IV, it will undoubtedly be the single most demanding node in the Network. Users will not only be sending requests for service but will also send very large quantities of input and output data, e. g., a 10 b1.t weather map, over the Net. Projected uses of the ILLIAC include weather and climate modeling, picture processing, linear programming, matrix manipulations, and extensive work in other areas of simulation and modeling.

The ILLIAC installation will also include a trillion bit mass store. An experiment is being planned to use 10% of the/storage (100 billion bits) as archival storage for all the nodes on the Net. This kind of capability may help reduce the number of tape drives and/or data cells in the Network.

### **FUTURE**

There are many applications of computers for which current communication technology is not adequate. One such application is the specialized customer service computer systems in existence or envisioned for the future; these services provide the customer with information or computational capability. If no commercial computer network service is developed, the future may be as follows:

**The ARPA Network**  
by **Lawrence.G.Roberts**  
**Advanced Research Projects Agency**  
**Barry D. Wessler**  
**University of Utah**

One can envision a corporate officer in the future having many different consoles in his office: one to the stock exchange to monitor his own company's and competitor's activities, one to the commodities market to monitor the demand for his product or raw materials, one to his own company's data management to system monitor inventory, sales, payroll, cash flow, etc., and one to a scientific computer used for modeling and simulation to help plan for the future. There are probably many people within that same organization who need some of the same services and potentially many other services. Also, though the data exists in digital form on other computers, it will probably have to be keypunched into the company's modeling and simulation system in order to perform analyses. Though these services are desirable, the executive would be faced with learning to use several different consoles and protocols.

The organization providing the service has a hard time, too. In addition to collecting and maintaining the data, the service must have field offices to maintain the consoles and the communications multiplexors adding significantly to their cost. A large fraction of that cost is for communications and consoles, rather than the service itself. Thus, the services which can be justified are very limited.

Let us now paint another picture given a nationwide network for computer-to-computer communication. The service organization need only connect its computer into the net. It probably would not have any consoles other than for data input, maintenance, and system development. In fact, some of the services data input may come from another service over the Net. Users could choose the service they desired based on reliability, cleanliness of data, and ease of use, rather than proximity or sole source. Large companies would connect their computers into the net and contract with service organizations for the use of those services they desired. The executive would then have one console, connected to his company's machine. He would have one standard way of requesting the service he desires with a far greater number of services available to him.

For the small company, a master service organization might develop, similar to today's time-sharing service, to offer console service to people who cannot afford their own computer. The master service organization would be wholesalers of the services and might even be used by the large companies in order to avoid contracting with all the individual service organizations.

The kinds of services that will be available and the cost and ultimate capacity required for such service is difficult to predict. It is clear, however, that the network philosophy is adopted and if it is made widely available through a common carrier, that the communications system will not be the limiting factor in the development of these services as it is now.

## **REFERENCES**

T. MARILL and L. ROBERTS.  
Toward a Cooperative Network of Time-Shared Computers  
AFIPS Conference Proceedings Nov 1966

P. E. JACKSON and C.D. STUBBS  
A Study of Multi-Access Computer Communications  
AFIPS Conference Proceedings Vol 34 p 491 1969

PAUL BARAN, et al  
On Distributed Communications  
RAND Series Reports Aug 1964

**The ARPA Network**  
by **Lawrence.G.Roberts**  
**Advanced Research Projects Agency**  
**Barry D. Wessler**  
**University of Utah**

L. G. ROBERTS and B.D. WESSLER  
Computer Network Development to Achieve Resource Sharing  
AFIPS Conference Proceedings May 1970

F.E. HEARTS, R.E. KAHN, S.M. ORNSTEIN, W.R.CROWTHER and D. C. WALDEN  
The Interface Message Processor for the ARPA Network  
AFIPS Conference Proceedings May 1970

L. KLEINROCK  
Analytic and Simulation Methods in Computer Network Design  
AFIPS Conference Proceedings May 1970

H. FRANK, IT. FRISCH and W. CHOU  
Topological Considerations in the Design of the ARPA Computer Network  
AFIPS Conference Proceedings May 1970

S. CARR, S. CROCKER and V. CERF  
HOST-HOST Communication Protocol in the ARPA Network  
AFIPS Conference Proceedings May 1970