#209

CHALLENGING PROBLEMS IN THE DESIGN OF COMPUTER-COMMUNICATION NETWORKS

by

Leonard Kleinrock

Computer Science Department
School of Engineering and Applied Science
University of California, Los Angeles 90024
Phone: (213) 825-2543

ABSTRACT

In the past few years we have seen the emergence of numerous computer-communication networks of various types. The structure and sophistication of these networks varies over a considerable range from the highly specialized networks designed to handle specific tasks in a carefully controlled environment to the more generalized networks which handle a variety of tasks in a highly unstructured environment. The successes and failures of these networks in terms of economy, service, response time, throughput, coverage, reliability, use and convenience is certainly not uniform from network to network. In this paper we examine some of the general principles of network design which have emerged, some of the experience we have gained with respect to a particular generalized network (the ARPA experimental computer network) and then lastly we discuss some of the open problems and challenges which as yet remain in the design of data networks. The problems addressed are not only with regard to the technical design of networks but also with regard to some of the human factors questions which become so very important in successful network operation.

CHALLENGING PROBLEMS IN THE DESIGN OF COMPUTER-COMMUNICATION NETWORKS*

by

Leonard Kleinrock
Computer Science Department
University of California, Los Angeles

1. INTRODUCTION

Whenever a set of computer systems are interconnected in a way which permits them to communicate with each other in order to share such resources as hardware, software, or data bases, then we may think of that collection as a computer-communication network. Just as the 1960's represented the decade of time-sharing, so the 1970's represent the era of computer networks.

The interest in computer networks has grown enormously in the last few years, and the complexity of the problems one faces in creating such networks is staggering. The considerations range from the highly technical, mathematical and engineering design questions (many of which are "within reach") to the extremely frustrating political, legal, social, management and ethical questions which penetrate the fabric of our society today. This growth in the need for data communications comes from a large number of varied application areas. For example, the finance industry, including banking and insurance firms has a growing need for remote data processing (such as electronic funds transfer, etc.). In the field of medicine and health there is a need for large information banks with remote access. Educational computing needs currently emphasize interactive use as opposed to data entry, retrieval and acquisition. Large government agencies have vast data exchange requirements both military and non-military. Point of sales terminals by retail organizations is a fast-growing applications field. Information retrieval is of great importance in the transportation field currently and control of traffic load is a fast growing area. Large corporations now are exchanging data among their many central and regional offices. Other industries have a natural need for computer networks (for example, airline reservations systems, travel services, etc.). A vast use is foreseen for access to information processing directly from the consumer's home (shopping, voting, the use of electronic

---

*This work was supported by the Advanced Research Projects Agency of the Department of Defense under Contract No. DAHC 15-69-C-0285.

will take place, we identify a second "incompatibility." Remote terminals come in a *variety* of types ranging from many low-speed, inexpensive teletypewriters to a few higher-speed, intelligent computer display terminals. Typically, these terminals are located over some *distributed* geographical region, perhaps in clusters. They tend to operate *asynchronously* in the sense that the characters they generate are spaced nonuniformly in time. They tend to have a very *low duty-cycle* and also tend to generate data in *bursts* (a nasty combination). Among this multitude of terminal types there is also a multitude of *incompatible* coded versions of the common alphanumeric symbols. Most of the terminals are relatively *cheap* and *unsophisticated*. They operate relatively *independently* of each other and are unaware of other terminal behavior. On the other hand, the main computer complex typically consists of *one*, or at most a few, *large* central processors. These machines are large-scale, *high-speed* computers which operate *synchronously* with a *high duty-cycle*. Usually, a single *standard* representation for alphamuneric symbols is used. They are very *expensive*, highly *sophisticated* and, if there is more than one at a given complex, then considerable energy may be expended to see that they *cooperate* in an efficient, symbiotic fashion.

Another difficulty concerns the telephone network which was originally designed to carry voice traffic, an analog signal whose highly redundant nature combats the various effects of noise present in the telephone network. In order to transmit data traffic over this network, the digital signals generated by computers and terminals must be converted to the analog signals which the telephone network carries on its voice grade lines; the devices which carry out this conversion are referred to as "modems" (which is a contraction of the word modulator-demodulator). Since this network is designed to handle voice signals, then one anticipates that adapting such a system for data transmission (whose statistics vary considerably from that of voice signals) is a formidable task. The difference between data and voice traffic lies at the source of many of the difficulties in the design of data communication networks; the principal characteristic is that unpredictable demands arise at unpredictable times leading to bursty low duty-cycle traffic. This requires buffering, smoothing, multiplexing, concentrating, etc.

Thus we have a complex interaction of partially incompatible systems, and our task is to provide a "message service" which is "invisible" in the sense computers which are in communication but rather this communication involves a sequence of message transmissions sharing the communication lines with other messages in transit. The maximum message size is 8064 bits. Thus a pair of HOSTS will typically communicate over the net via a sequence of transmitted messages. To obtain delays of a few tenths of a second for such messages and to lower the required IMP buffer storage, the IMP program partitions each message into one or more packets, each containing at most 1008 bits. Each packet of a message is transmitted independently to the destination where the message is reassembled by the IMP before shipment to that destination HOST. Thus we have the concept of a packet-switching network [1, 14].

The network was designed so as to achieve both a rapid delivery for the short interactive messages as well as a high bandwidth for the long data files. The first of these goals has been achieved; in a lightly loaded network the response time for short messages varies between one tenth and two tenths of a second even over many hops, and in a more heavily loaded network this response time is usually less than one half a second. As regards the second goal, we are able to provide a high bandwidth for long messages under light and moderate traffic loads; however under heavy traffic loads the throughput falls somewhat. From the time that the network came to life in September 1969 (when the first IMP was connected at UCLA), considerable effort has gone into measuring and evaluating the performance of this network. As a result changes in the IMP program have been made from time to time, the most recent occurring in the spring of 1972; these changes eliminated some important problems with regard to deadlock conditions in the network [12]. Currently the operating procedure is such that no multipacket message (a message consisting of more than one packet) is allowed to enter the network until storage for that message has been reserved at the destination IMP. The HOST computers are essentially unaware of the notion of packets and deal exclusively with messages; it is therefore required that multipacket messages be reassembled prior to the delivery of that message to its destination HOST. The buffers (eight of which are set aside for each multipacket message) turn out to be a critical resource in the ARPA Network. The high bandwidth for long sequences of messages is obtained by requiring only the first message in such a sequence to go through this reservation procedure; from that point on the reservation is held for the message sequence, subject to a timeout. At the same time we obtain rapid response for short messages (single packet messages) by transmitting them major difficulty with these lines is that occasionally there are extended periods (hours or days) of line outages. The Network Control Center monitors both IMP and line outages continually and the following table summarizes the network operation over a ten month period [11].

Month	Average Line Outage	Average IMP Down	# of Nodes	Average HOST Inter-site Output (packets/day)
September 1971	.59%	3.27%	18	51,386
October	1.66%	1.77%	18	95,930
November	1.65%	5.50%	18	116,515
December	3.21%	3.95%	19	107,896
January 1972	1.02%	1.92%	19	172,037
February	1.23%	2.73%	19	224,668
March	1.36%	4.00%	23	240,144
April	.88%	2.86%	25	362,064
May	1.11%	2.57%	25	505,639
June	.41%	.97%	29	807,164

This table shows that the phone lines and the IMPs have roughly equivalent reliabilities. It is fair to conclude that with these error detection precautions, the telephone communication circuits do not create a problem in the performance or growth of these networks. When a line or IMP goes down, the network routing procedure will automatically adapt to the new condition thereby preventing congestion.

The IMP

The original IMP [2] was constructed using a Honeywell 516 computer; this is a 16 bit machine with a memory cycle time of 0.96 microseconds. Configured as an IMP the cost is approximately $100,000. Recently the Honeywell 316 computer (cycle time of 1.6 microseconds) has been configured as an IMP costing approximately $50,000; the new IMPs are of this type. An IMP is provided with 12 K of core.* This piece of equipment is responsible for all the processing of packets, which includes: decomposition of HOST messages into packets; routing; relaying and receiving store-and-forward packets; acknowledging accepted packets and retransmitting unacknowledged packets; reassembling packets into messages at the destination HOST; generating control messages, etc. In addition the IMP program is responsible for gathering statistics, performing on-line testing, and monitoring the network status. The shortest packet

---

*Recently, it was decided to increase the core storage of an IMP to 16K to provide more buffers. Similarly, the TIP (see below) is now to be provided with 28K. These additions are currently in the process of being installed.

differing types may be connected to a given TIP and up to three modem and/or HOST interfaces may be connected. That which distinguishes a TIP from an IMP (aside from the additional 8 K of core) is a device known as a multiline controller (MLC) which allows terminal connections to the IMP. The terminals are handled on a per-character basis with start and stop bits (even on synchronous lines). Data rates and character bit length may be set for each terminal line by the TIP program itself. For each input and output terminal line, two full characters are buffered - the one currently being assembled or disassembled and one further character to account for memory accessing delays. The MLC contains 256 integrated circuits (MSI and LSI) and is approximately the same complexity as the basic Honeywell 316 computer itself. Each line interface unit contains an additional 31 ICs. A TIP costs approximately $100,000. The additional 8 K memory is required for the special TIP code, tables, buffer storage for terminal messages, etc. The per character processing time is about 75 microseconds and the overhead per message can be extremely large (a factor of 10 or 20 in bandwidth) when single characters are sent one at a time. Approximately 5% of the TIP will be lost in performing as an IMP, even in the absence of IMP traffic. The TIP bandwidth is approximately 500 kilobits in the absence of terminal traffic (for full size messages). The TIP average per machine down rate is approximately 2%.

In October 1972, a demonstration of the ARPA Network was conducted in conjunction with the International Conference on Computer Communications (ICCC) in Washington, D.C. Approximately 30 terminals from various manufacturers were connected to a TIP at the conference site. Instruction booklets were made available to the conference attendees which described methods for accessing various resources on the ARPA Network through these terminals. The procedure which one goes through in reaching a remote computer facility is as follows. First he sits down, powers up the terminal, and then initiates a simple dialogue with the TIP. Then he requests the TIP to make a connection to a remote HOST and when this is accomplished he ignores the TIP and proceeds to login to the remote HOST. Following this, as has always been the case, the user then ignores the operating system of that HOST and communicates directly with the user process he has now been put in contact with. During that ICCC demonstration, the true power of the ARPA Network became apparent not only to the uninitiated users of the network, but also to the sophisticated and experienced users as they observed peak traffic rates of 60,000 packets per hour

4. CHALLENGING PROBLEMS

Our opening comments express concern over the difficulty of wedding computers and communications. This apprehension was based on certain apparent incompatibilities between the computer and communication industries and between terminal and computer behavior. We have discussed one example of successful computer network operation and this among others provides a strong basis for confidence in these networks.

There remain several open questions in network design. For example, what structure should a high bandwidth IMP have? How can efficient use be made of a variety of high bandwidth circuits? The entire question of large networks poses numerous challenging questions; for example, how should these large networks be partitioned for effective design and what operational procedures should they follow? The introduction of satellite links to overseas nodes brings up an entirely new set of questions with regard to simultaneous access to a wide band channel; the rub is that these nodes must communicate with each other in order to control the use of that channel and this communication must take place over the channel which they are attempting to control. The creation of the ARPA Network has stimulated considerable research into how programs and operating systems should communicate with each other; this question is of interest even independent of the network operation.

It is interesting to note that the present ARPA Network is expanding rather quickly both in size and in traffic. Many outside groups are initiating efforts to gain access to the network. Currently ARPA is considering the transfer of the network from under its own research and development control to some other operational agency or specialized carrier. The selection of such an agency is by no means trivial and has already raised some difficult questions. Moreover as other non-ARPA users gain access to the network we must resolve the very nasty questions regarding charging mechanisms, privacy and security guarantees, guaranteed access and service, etc.

It is perhaps fair to say that whereas a variety of significant technical problems face us with regard to the growth of telecommunications and remote data processing, it is clear that these will not be the significant problem areas of the future. It does not take much thought to realize that the major problems are social, economical, political and even ecological in nature. Moreover, one must carefully examine the real goals that the customer of remote data processing may be measuring a proposal against. For example, the

8. Crocker, S.D., J.F. Heafner, R.M. Metcalfe and J.B. Postel, "Function-Oriented Protocols for the ARPA Computer Network," *AFIPS Conference Proceedings*, 40:271-279, SJCC, 1972.

9. Thomas, R.H. and D.A. Henderson, "McROSS--A Multi-Computer Programming System," *AFIPS Conference Proceedings*, 40:281-293, SJCC, 1972.

10. Roberts, L.G., "Extensions of Packet Communication Technology to a Hand Held Personal Terminal," *AFIPS Conference Proceedings*, 40:295-298, SJCC, 1972.

11. McKenzie, A.A., B.P. Cossell, J.M. McQuillan, M.J. Thrope, "The Network Control Center for the ARPA Network," *ICCC Proceedings*, pp 185-191, October 1972.

12. McQuillan, J.M., W.R. Crowther, B.P. Cossell, D.C. Walden and F.E. Heart, "Improvements in the Design and Performance of the ARPA Network," *AFIPS Conference Proceedings*, 41:714-754, FJCC, 1972.

13. Fultz, G.L. and L. Kleinrock, "Adaptive Routing Techniques for Store-and-Forward Computer-Communication Networks," *Proceedings of the International Conference on Communications*, pp 39-1--39-8, 1971.

14. Davies, D.W., "The Principals of a Data Communication Network for Computers and Remote Peripherals," *Proceedings of IFIPS Congress*, Edinburgh, August 1968.

15. Gerla, M., "The Design of Store-and-Forward Networks for Computer Communications," Ph.D. dissertation, School of Engineering and Applied Science, University of California, Los Angeles, 1973.