M.I.T. Laboratory for Computer Science
Computer Systems Research Groups

Request for Comments 260

June 27, 1984

CSC Report Draft, July, 1983--June, 1984

from J. H. Saltzer

Attached is the first complete draft of the CSC Annual Report for 1983-84. Please look it over, especially to help identify things our group did that slipped through the cracks.

COMPUTER SYSTEMS AND COMMUNICATION

Academic Staff

J.H. Saltzer, Group Leader
F.J. Corbato
M.V. Wilkes
D.P. Reed (leave of absence)

Graduate Students

D.L. Estrin
D.C. Feldmeier
K. Koile

Undergraduate Students

D.W. Gillies
F.S. Hsu
E. Jaeger
D.J. Karlson
M.L. Lambert
J.L. Romkey

Support Staff

N. Lyall
M.F. Webber

Visitors

R.K. Jain
B.G. Lindsay

COMPUTER SYSTEMS AND COMMUNICATIONS

1.1 INTRODUCTION

Rearrangements of Computer System group boundaries occurred in 1983–84, with the result that the area reported under this title is substantially smaller than in past years. The research projects of the group now fall in three categories: interorganization networks, local area network technology, and network services. The next sections describe these three areas in turn.

1.2 INTER-ORGANIZATION NETWORKS

During the past year Deborah Estrin continued her doctoral research on the interconnection of computer networks across organization boundaries (referred to as inter-organization networks or ION's). This interdisciplinary research encompasses both organization effects and new technical requirements. To date, the research results are of two kinds: hypotheses about organization effects of ION's, and design specifications for usage control mechanisms.

1.2.1 Hypotheses

Within an industry, ION's can alter the nature of competition among participants by introducing new barriers to market entry. Within an organization, interconnection can induce changes in the management of internal facilities and in the technical requirements on the underlying networks.

1.2.1.1 Industry Effects

The potential effects of ION's on businesses and industries include: smaller transaction overhead (e.g., less paperwork and fewer data-entry errors), improved inventory management, innovative product and service differentiation, increased interdependence among firms, and new barriers to market entry. In short, technology can affect the ways in which firms do business.

This research is on strategic effects that flow directly from technical choices. For example, one study investigates the use of non-standard protocols between customers and suppliers. In order to communicate via an ION, organizations must use the same protocol. Use of a standard protocol allows the equipment, training, and overall ION investment to be used for interchange with any organization that implements the standard. On the other hand, use of a non-standard protocol restricts customers to interchange with the small number of organizations that use that specialized protocol. Therefore, non-standard protocols act as market-entry barriers by encouraging customers to stay with the original ION-providing supplier, rather than make the investment required to switch to a new one, or forego the convenience of ION interchange.

In the longer term, three factors will erode the strategic advantage gained by using non-standard protocols: (1) Third parties will find it profitable to offer services that provide uniform interfaces to multiple suppliers; (2) Customers will favor those computer vendors that implement standardized protocols and interfaces in order to facilitate system-integration internally; (3) As desk-top computers replace limited-function terminals, customers will be able to deal with multiple standards more easily.

1.2.1.2 Organization Effects

Computer-to-computer communication can be qualitatively different from traditional forms of inter-organization communication (i.e., post and telephone). It allows a person or computer in one organization to cause something to happen in a second organization, without any intervening human decision from the second organization. Such possibilities can affect significantly the participants' management of internal resources.

Internal network facilities have as a primary goal the enhancement of internal communications and resource sharing. As a result, restriction of the communications possibilities is usually a secondary consideration or even undesirable. But, external connections force a new requirement— the need to control communications. At the very least, the communication system and all software must be able to distinguish internal from external users. In response to interconnection, a participant can implement technical and administrative mechanisms that discriminate against external users. But, if traditional protection mechanisms are used, internal users will be affected as well, perhaps inhibiting internal communication and resource sharing. Even if such internal changes are acceptable, it is difficult to move from an environment in which

resources are by default open to one in which they are by default closed. Alternatively, a participant may accept lack of control of external uses in order to maintain internal convenience. Or, lack of control may be a default that arises from insufficient concern or attention.

The choice between increasing internal controls or changing the inter-organization relationship is determined by characteristics of the participants. Organizations that manage resources conservatively and perceive strong proprietary interests will not accept greatly increased accessibility by outsiders. Instead, these organizations will adopt usage controls, even if internal practices must change as a result. In contrast, organizations with loose resource management policies and ill-defined proprietary concerns will accept accessibility by outsiders. In fact, such organizations may be more troubled by restrictions on internal operations, or compromise of function across the organization boundary, than by increased external accessibility.

1.2.2 Usage Control Mechanisms

Connection of computer networks across organization boundaries poses different design requirements from intra-organization connection. Designs for intra-organization networks emphasize connectivity, performance, and transparency with maximum function. In contrast, an ION participant may need to limit connectivity and function in order to restrict access by outsiders to internal facilities. Such restrictions should not compromise internal requirements for transparency and performance. This research aims to develop flexible usage control mechanisms that restrict external usage without increasing internal controls.

Traditional protection mechanisms expose all users to a single set of access controls. Such mechanisms are adequate for a computer that is dedicated to external functions. But, if a computer is used for internal functions as well, the organization may experience performance overheads, restricted information flows, and disincentives to resource sharing. If we think of a computer residing simultaneously in multiple logical networks whose usage control requirements differ (e.g., internal and external), then the computer must differentiate between members of the different networks if it is to implement more than one policy. Therefore, messages must be tagged so that the computer can appropriately and efficiently filter requests.

This research suggests that hiding network boundaries from applications that cross organization boundaries is inappropriate. However, recently-developed internetwork protocols present a *single* logical network to all applications in order to simplify the network interface. This conflict must be addressed in the design of both intra- and inter-organization network protocols.

1.2.3 Experimental Implementations

We have implemented two components of a mail-relay gateway suitable for inter-organization connections.

Don Gillies implemented a secure mail relay on an IBM Personal Computer as an undergraduate thesis project. The mail relay was inspired by the needs of the Laboratory for Computer Science Headquarters, which had a new office automation program for their UNIX system. They wanted to connect to the rest of the MIT computer network, but also wanted to protect their sensitive research accounting data. The same basic software design should be usable in the planned MIT-to-IBM mail relay connection.

The mail relay is built on top of the CSC group's PC/IP protocol implementation for the IBM PC. It contains a new a multiconnection TCP, user and server SMTP, and a reliable spooling program. It also has security mechanisms that leave audit trails, that detect unusually heavy mail traffic, that halt messages with suspicious ascii characters, and that record attempts to speak unauthorized protocols through the relay. The full design is explained in the thesis.

The second component (also implemented by Gillies) is mail-filtering software running on a VAX. This facility can be tailored to implement higher-level policies (e.g., no transit) for mail traffic that enters the network via the gateway.

By connecting the mail relay to an internal local area network on one side, and a telephone line on the other, we can implement a controlled, inter-organization, mail network. All messages that arrive over the telephone line can be sent to the VAX for higher-level filtering. In addition, we can use a link-level protocol that authenticates the origin of packets that arrive over the telephone line.

1.3 LOCAL AREA NETWORK TECHNOLOGY

1.3.1 Ring network monitoring results

The primary progress this year on the ring network research project was bringing into service of a ring network monitoring station, built by David Feldmeier, and the beginning of systematic collection of data on the proNET token ring network. In review, the proNET ring is a commercial version of a 10 Mbit/sec token ring local area network designed by this research group and reported in detail in previous reports. Its primary innovative feature is a star-shaped topology with a passive wiring center intended to make maintenance easy and availability very high. The L.C.S. installation currently covers three floors of the building using four interconnected wire centers. Although the installation is mostly done with twisted pair, there is one experimental fiber-optic section. There are 30 VAX 11/750 and 11/780 computers, five LSI-11's, a Bridge CS/1, and 3 IBM PC's connected to the ring. The Bridge CS/1 and several of the LSI-11's act as gateways to other local area networks, chiefly an experimental (3 Mbit/sec) Ethernet, a standard Ethernet, the ARPANET, and a serial-line network.

Preliminary observations of the monitoring station appear in a Master's thesis by Feldmeier, also available as an L.C.S. Technical Memorandum. Some of the more interesting results are the following:

- Packet traffic is between 1 and 2 million packets/day. The ring is thus in production use, a vital link in the laboratory's resources.

- Load distribution is very similar to that reported by Shoch and Hupp at Xerox, with the difference that the individual nodes at M.I.T. seem to generate about 3 times as many bits/second/node. Two possible explanations of this more intense traffic are that the M.I.T. site has more remote paging, and that the VAX 11/750 computers are more powerful than the Xerox Alto, so they can make more frequent demands on the network.

- Internetwork traffic is some 40% of the total traffic on the ring. This large number may have important implications for gateway design and generally for internet plans, although it is hard to tell whether it would be similar in other environments with less diverse communications facilities.

- The distribution of packet interarrival times is decidedly non-Poisson: shorter interarrival intervals have much greater than Poisson probabilities. This observation has led Raj Jain to explore a network packet arrival model called "packet trains" reported below.

- "Network unavailable time," which accumulates whenever a token is continuously absent for more than one second, averages 5 to 10 seconds per 24-hour day. Since all reconfiguration and repair of the network at L.C.S. is done without turning off the network or the monitor, this small number suggests that the star configuration is extraordinarily effective in providing high availability.

1.3.2 Packet Trains for Network Workload Modeling

For network modeling and simulation, it is essential to know the right model for packet arrivals. The traffic measurements on the LCS token ring show that the packet arrivals do not follow the commonly used model of Poisson arrivals which assumes that packets arrive independently and that the arrival of a packet gives no clue about future arrivals. A more realistic model called *packet trains* has been proposed by Raj Jain, in which all packets of a file transfer form a train. The inter-packet interval between successive packets of a train has a distribution very different from inter-train interval. The former is a characteristic of the higher level protocols and system configurations while the later depends solely on user behavior. After the arrival of a locomotive (the first packet) of a train subsequent packet arrivals can be predicted with very little variance. Protocols that exploit this dependence of packet arrivals, e.g., reservation sharing, can be designed to use resources more efficiently than current ones which are optimized for random Poisson traffic.

1.3.3 High-bandwidth residential networks

A second network technology effort is just beginning this year: exploring the use of commercial cable television systems as a high-bandwidth local network technology that can reach the home. The primary progress on this topic has been to develop, in concept, techniques that can deal with the analog environment of the typical CATV system. It now appears that a promising approach is the use of spread-spectrum modulation techniques, for four reasons. First, spread-spectrum provides compatibility with miscellaneous services already in place on the same cable. Second, the anti-jamming properties of spread-spectrum modulation provide a counter-measure to deal with extreme interference from short-wave broadcasts in the frequency bands available for two-way data communications. A third property, that spread-spectrum signals are well hidden from non-spread spectrum receivers, may also be useful in allowing data signals to use radio frequencies on the cable that have been set aside to protect nearby aircraft communications. Finally, the spread-spectrum code-division multiple access technique may be an effective alternative to carrier-sense or token-passing as a channel access control technique. The theoretical properties of spread-spectrum modulation for this application all are sufficiently promising that it is time to begin a more detailed study, perhaps undertaking a modem design in the coming year.

1.4 NETWORK SERVICES

1.4.1 Personal Computer Networking: PC/IP

Work continued on the development of network programs for the IBM PC. This year, several new programs were written, including one that displays all packets sent on the Ethernet, which is useful for debugging. Other programs placed in service include a file transfer server package, a remote printing user interface, and an interface to the ARPANET Network Information Center name directory. A major program which was finished was an implementation of RVD, the Remote Virtual Disk protocol, for the IBM PC. RVD was developed by LCS to permit shared (read only) disks and large private disks on VAXes that actually have only small ones. The PC implementation allows ten RVD drives on one machine, making it possible to have tremendous amounts of data accessible. In the current implementation, disk accesses are about as fast as a floppy disk. A drawback of the current implementation is that no network programs (most important, TFTP) can write to an RVD disk because of contention for the network interface. The problems are mostly caused by the unusual structure of the programs and awkward integration of the network with the PC DOS operating system. Another major addition this year was a driver for the proNET ring interface for the IBM PC. John Romkey worked on the programs and the RVD implementation.

A major release of the source language versions of the PC network programs was made, dated February 1. The release included most of the work done prior to RVD, binaries of all the programs, the User's Manual and the Programmer's Manual. It was 2.8Mbytes long. Copyright messages in the code grant permission to do anything except remove the copyright messages. Over 50 copies have been shipped at a $45 fee that covers the cost of a tape, duplication, and postage. Handling of the distribution was done largely by Muriel Webber.

1.4.2 On-line directory system

K. Koile completed a Master's thesis describing the design and implementation of an online directory assistance system called DIRSYS. The system, designed for use by members of the M.I.T. community, has an *incremental* interface that combines features of a paper phone book with those of a full-screen editor such as Emacs. Each directory entry is displayed in a compact one line per entry format, as are entries in a paper phone book. Since more information is available for each entry than in a paper phone book, a command is available for changing entries into a more expanded format in which additional information is displayed.

Users may "browse" through this electronic phone book by issuing commands similar to Emacs' cursor motion commands, or they may search for a specific name by typing the name. After each letter that the user types, DIRSYS moves the *highlight*, the means of emphasizing an entry, to the entry whose name string most closely matches what the user has typed so far. This incremental search mechanism is similar to that used in Emacs. The system provides a *help facility* with two levels: the first level reminds users of which commands are available; the second level describes the function of a specified command in detail. A *tutorial* is available for users who want a very detailed description of the system.

Finally, DIRSYS provides a facility for keeping the information in the directory database up-to-date. A user may submit *update requests*, which contain information about modifications to his directory entry, to the DIRSYS manager. When the update requests have been validated, the information contained in them is incorporated into the directory database by an *update daemon*, a program that runs every night to update the database.

A preliminary evaluation of DIRSYS indicates that the system can be used easily by both inexperienced and experienced computer users. Details of the evaluation were described in the thesis. The learning aids guide the novice without encumbering the experienced user. The commands are simple, easy to use, and consistently interpreted. The system provides prompts and polite, informative messages to the user. It also is robust in that it is very difficult to cause DIRSYS to fail to operate. Finally, individuals who used DIRSYS seemed to enjoy using the system, even though the system response was slow at times, and agreed that a directory system such as DIRSYS would provide a convenient service for use both inside and outside the M.I.T. community.

Doris Karlson, in an undergraduate thesis, explored the possibility of a search based on name sounds as an alternative to the incremental search of DIRSYS. A modified version of the Soundex

indexing system was tried on the M.I.T. directory database. Statistics from that database suggest two things: 1) A more sophisicated indexing system than Soundex seems necessary for reasonable human engineering, because the Soundex system tends to index too many names that do not sound alike in the same category; and 2) The frequency of similar-sounding names in a file of 20,000 entries is large enough that display of more than one name per line, 20 lines per display is required.

1.4.3 Analysis of Timeout Algorithms for Packet Retransmission

Almost all networking and distributed systems protocols have to cope with the problem of determining when a node should retransmit a packet that has not been acknowledged. A bad timeout algorithm may either flood the network with duplicate copies of packets and lead to unwanted congestion or it may take too long to retransmit a packet. An analysis of various timeout algorithms shows that during period of congestion (repeated packet loss), most timeout algorithms would either diverge to extremely high timeout intervals, or converge to a rather low timeout value. In either case, the throughput may be reduced to virtually zero or the circuit may be disconnected prematurely. Raj Jain did a simulation study of a variety of timeout algorithms and has compiled a list of guidelines for designing such algorithms and for setting parameter values.

References

Publications

1. Comer, M.H., "Loose Consistency in a Personal Computer Mail System," S.M. thesis, MIT, Department of Electrical Engineering and Computer Science, Cambridge, MA. June, 1983. Also MIT Laboratory for Computer Science Technical Report, MIT/LCS/TR-316, June, 1984.

2. Feldmeier, D.C., "Empirical Analysis of a Token Ring Network," S.B. thesis, MIT, Department of Electrical Engineering and Computer Science, Cambridge, MA. January, 1984. Also MIT Laboratory for Computer Science Technical Memorandum, MIT/LCS/TM-254, January, 1984.

3. Koile, K., "The Design and Implementation of an Online Directory Assistance System," S.M. thesis, MIT, Department of Electrical Engineering and Computer Science, Cambridge, MA. December, 1983. Also MIT Laboratory for Computer Science Technical Report, MIT/LCS/TR-313, December, 1983.

4. Estrin, D., and Sirbu, M., "Cable Television Networks as an Alternative to the Local Loop," to appear in Journal of Telecommunications Networks, 1984.

5. Saltzer, J.H., Reed, D.P., and Clark, D.D., "End-To-End Arguments in System Design," to appear in ACM Transactions on Computer Systems, November, 1984.

Theses Completed

1. Comer, M.H., "Loose Consistency in a Personal Computer Mail System," S.M. thesis, MIT, Department of Electrical Engineering and Computer Science, Cambridge, MA. June, 1983.

2. Feldmeier, D.C., "Empirical Analysis of a Token Ring Network," S.B. thesis, MIT, Department of Electrical Engineering and Computer Science, Cambridge, MA. June, 1984.

3. Frankston, C., "The Amber Operating System," S.B. thesis, MIT, Department of Electrical Engineering and Computer Science, Cambridge, MA. May, 1984.

4. Gillies, D.W., "Improved Network Security with a Trusted Mail Relay," S.B. thesis, MIT, Department of Electrical Engineering and Computer Science, Cambridge, MA. June, 1984.

5. Karlson, D.J., "Soundex Searching in an Online Directory Assistance System," S.B. thesis, MIT, Department of Electrical Engineering and Computer Science, Cambridge, MA. June, 1984.

6. Koile, K., "The Design and Implementation of an Online Directory Assistance System," S.M. thesis, MIT, Department of Electrical Engineering and Computer Science, Cambridge, MA. December, 1983.

Theses in Progress

1. Estrin, D.L., "Inter-organization Networks," Ph.D. thesis, MIT, Department of Electrical Engineering and Computer Science, Cambridge, MA, expected date of completion, September, 1985.

2. Jaeger, E., "Third Party Access Control and Accounting Schemes," S.B. thesis, MIT, Department of Electrical Engineering and Computer Science, Cambridge, MA, expected data of completion, December, 1984.

Conference Participation

1. Estrin, D.L., Technology for Meaningful Work, Institute of Policy Studies, Allentown, PA. April, 1984.

2. Estrin, D.L., IBM–MIT Workshop on Security, Boston, MA. May, 1984.

3. Saltzer, J.H., Chairman, ACM 9th Symposium on Operating Systems Principles, Bretton Woods, NH. October, 1983.

Talks

1. Corbato, F.J., "System Issues in Project Athena," Panel discussion. Ninth ACM Symposium on Operating Systems Principles, Bretton Woods, NH. October, 1983.

2. Saltzer, J.H., "Interorganization Links and Community Networks," presented at ILP Symposium on "Networked Computer Systems," M.I.T., Cambridge, MA, May 14, 1984.

Committees

1. Corbato, F.H., National BUreau of Standards: Evaluation Panel for Scientific Computing

2. Saltzer, J.H., CSNET Executive Committee

3. Saltzer, J.H., Cable Communication Advisory Board, City of Newton, MA.