M.I.T. Laboratory for Computer Science — November 9, 1981
Computer Systems Groups — Request for Comments No. 211

DRAFT of Annual Report, July, 1980 - June. 1981

Cover note by M. Webber

Attached is a draft of the Annual Progress Report of the activities of the Computer Systems and Communications group. Please look it over and offer comments on

- omissions of significant activities
- details that are wrong
- anything else.

Check also the lists of papers, talks, committee memberships, etc., that appear at the end for mistakes and omissions.

The final version of this report will be included in the L.C.S. Annual Report, and we will also make copies for handout to visitors until the L.C.S. Annual Report is available. It is usually the case that the report gets distributed widely; many people follow our activities almost exclusively by this mechanism. Thus there is a significant payoff to making it good.

COMPUTER SYSTEMS AND COMMUNICATIONS

Academic Staff

J.H. Saltzer, Group Leader
D.D. Clark
F.J. Corbato
I. Grief

Research Staff

J.N. Chiappa
E.A. Martin

Graduate Students

R. Baldwin
G. Cooper
S. Curtis
D. Estrin
J. Frankel
S. Kent
L. Lopez
G. Simpson
V. Singh

Undergraduate Students

M. Agbleze
L. Allen
D. Bertko
J. Close
D. Feldmeier
T. Fitzgerald
D. Goldfarb
M. Greenwald
R. Houldin
L. Konopelski
C. Ludwig
S. Meadow
F. Meier zu Sieker
I. Mondori
M. Patton
M. Ross
R.Teal
J. Thomas
N. Webber

Support Staff

D. Fagin
N. Lyall
M. Webber

1. INTRODUCTION

At the beginning of 1981, the Computer Systems Research group split into two parts when David Reed founded the Computer Systems Structure group. This report mentions some joint work done prior to the first of the year, but in general the work prior to that time has been recorded in whichever group it most naturally fits. As suggested by its name, the Computer Systems and Communications group whose work is described here is interested in the impact of data communications on computer systems design. Most of the work of this group is experimental in nature, using the MIT local networks, the ARPANET, and the small and large computer systems at MIT as a laboratory. The activities of this group cover a wide range, from hardware design to system programming, and invoke many different levels of network implementation, from physical transmission level to operating system function.

2. RING NET HARDWARE

A major recent project at the physical transmission level has been the development of a 10 mbit/sec ring network, called the version 2 ring network (it replaces a 1 mbit/sec, version 1 ring). The research purpose of this development is to explore the hypothesis that a ring network is equal or superior to the Ethernet in day-to-day operation, maintenance, trouble isolation, and repair properties. At the same time, the development of this network is intended to provide a useful service in interconnecting the laboratory-developed NU personal computers and the several DEC PDP-11, LSI-11 and VAX-11 computers of the laboratory.

The Version 2 ring hardware is being marketed by Proteon Associates of Waltham, Mass. The ring interface consists of two parts, a net control/modem card and a host-specific interface to the PDP-11 UNIBUS. This set of two cards, at a price of $3200, is now in production, and has been shipped to several different customers. Other host-specific interface boards are under development in our laboratory for the NuBus and the S-100 bus, while Proteon Associates is developing one for the DEC Q-bus and exploring one for the Intel Multibus.

As part of this project, Glenn Simpson developed an alternative modem design that is completely digital, as opposed to the analog phase lock loop design that was produced for the modem by Proteon. The digital design has the advantages that it is more obviously reducable to VLSI, and it sidesteps some stability issues which had to be addressed in the closed loop analog modem.

As part of the ring interface development a large battery of test programs were developed by Liza Martin. These programs have now been released to other sites that are procuring the ring hardware.

A group of graduate students, as a class project, designed a VLSI chip containing a simplified, low-performance version of the ring network control circuitry. The chip will be fabricated and tested this summer, though it should be realized that this design was done as a feasability study rather than with serious intent of production use.

3. MIT INTERNET

Internetwork interconnection is a second major area of activity of this group, since there are now four high-speed local networks in the building at 545 Technology Square, as well as the ARPANET. Again in this area multiple purposes are being served. The research goal is to understand how to extend present interconnection techniques to a scale where perhaps 100 local networks and 5000 computers can communicate at high speeds. The anticipated data communication requirements of the MIT campus in 1990 serve as the focus for this interest. At the same time, service-producing results of the work are of immediate interest to other laboratory users.

A permanent high-speed connection was established between the LCS ringnet and the ARPANET, using a new C/30 IMP that was delivered for this purpose. There are now two net interconnection gateways in operation, one between the ARPANET and the 1 mbit/sec ring net, and the other connecting the 1 mbit/sec ring, the Ethernet and the Chaosnet. Substantial effort has been invested in stabilizing the service provided by these two machines, including an auto-restart facility which reboots the machines automatically over the network in case of software crash. The frequent outages which were initially observed with these interconnect machines now appear to be largely eliminated.

New software for these machines, which incorporates more sophisticated routing algorithms and better maintainability and modifiability, has been coded in C, and is being prepared for installation now.

We have worked with a committee being chaired by F.J. Corbato to plan for the networking of MIT. The primary output of this committee has been a memorandum outlining the scope of the data communication problem at MIT and specifying the general requirements a network must meet.

4. NEW PROTOCOL DEVELOPMENT

Dave Clark continues to participate in the ARPA working group developing Internet and Transmission Control Protocol (TCP). These protocols have now been adopted as DOD standards for internetting, and are beginning to be placed in service here at MIT as well as elsewhere in the ARPA community. A number of projects are underway to explore new protocols. In particular, for page or record level file access across a network, we have developed a protocol named Simple File Access Protocol (SFAP), which might be suitable for support of remote files for personal computers. Geoff Cooper, as a master's thesis, is exploring possible implemention techniques for SFAP. In particular, he is investigating the question of whether layering SFAP into a reliable datagram protocol and file transfer superstructure is of any benefit, or whether the layering is in fact an artificial structure which produces inefficiencies and bulkiness in the code.

A related area of protocol, the control of routing by specifying the complete route when a packet enters the network, so as to simplify and speed up forwarding nodes, is the subject of a master's thesis in progress by Vineet Singh. In this thesis, Singh is developing algorithms by which administratively distinct regions of a network can maintain local routing information and cooperate with other regional routing services to materialize a complete route for any given connection.

5. PROTOCOL PERFORMANCE

One of our principal research goals for the year is to determine how TCP will perform given the wide range of networks over which it is intended to operate. In particular, we are interested in how the performance characteristics of the protocol impact on the implementation done inside the host machine. In order to explore this, we have extensively metered the implementation of TCP on Multics, and we have produced a number of implementations of TCP for the Xerox Alto desktop computer, which implementations have been used for performance experiments as well as for service. The conclusions of our preliminary study are presented in a document now in preparation, but the general results are that: 1) the details of TCP are not a material contributor to overhead observed. The actual protocol processing is a very small part of the observed cost. The most important single host-related cost of TCP is that the packet is checksummed in software, a somewhat expensive operation whose benefit has been clearly demonstrated. 2) We have shown that the classical flow control mechanism, windows, does not necessarily work well under all circumstances in which TCP is expected to operate. When the total delay in the net is substantially more than the buffer space available in the receiving host, it is difficult to make windows work in such a way that high transmission rate is achieved. Mismanagement of windows, which we have identified by the term Silly Window Syndrome, is a major contributor to the abysmal throughput sometimes observed in the internet. 3) We have demonstrated that operating system overhead is a principal contributor to poor performance of host protocol implementations. Especially in classical operating systems, which were not intended to be used as communications processors, process switching is sufficiently costly that dealing with individual incoming packets causes a very large cost in process management. We are currently exploring what might be appropriate structures for operating systems that should operate well in a network environment.

6. MULTI-PROTOCOL COMPUTER MAIL

The ARPANET mail protocols do not permit delivery of mail to computers connected to nets other than the ARPANET. The ARPA community has proposed a new protocol, called Mail Transfer Protocol (MTP) which operates on top of TCP to provide an internet mail service. A protocol converting mail forwarded is now in operation on the Multics system, which makes it possible to send and receive mail to systems such as the RTS 11/70 which is connected only to the local net. Our group has also produced a mail package for UNIX, which is capable of using the Multics mail forwarder. This makes it possible to transfer mail between UNIX and systems on the ARPANET not yet upgraded to the new mail protocol.

7. HAND-HELD TERMINAL

This year, our group has started a project to produce a portable terminal small enough to fit in a pocket. Our goal is to exploit the technology and packaging that is now available and being used in various sorts of pocket computers and large calculators. The fundamental idea of our research is that a pocket terminal will only be really useable if the application program is specifically programmed to deal with this class of terminal. Our first test application has been in receiving and sending mail. We have programmed an Alto display to simulate the eventual terminal display, a single line of 36 LCD characters, and have put together various packages to see how one might read text through such a display. We are also developing a first version of a hardware prototype, which is not packaged as a small unit, but which has the correct sort of display, so that we can experiment with the actual technology.

8. OTHER ACTIVITIES

A new project was initiated by Jerry Saltzer to explore the possibility of a two-way community cable TV distribution system as a high-bandwidth data communication path to private homes. Discussions have been held with Continental Cablevision of Massachusetts concerning the possibilities of setting up an experimental data service on the CATV system that is about to be installed in the city of Newton. Deborah Estrin has completed an initial study of the technical feasibility of using Ethernet-like CSMA/CD for control of access in such a system, with the general conclusion that CSMA/CD will work well unless most traffic is small packets such as terminal input. (This same conclusion applies to the 10Mb/sec Xerox Ethernet.) Discussions are underway with potential modem manufacturers with the goal of initiating a joint project with one of them.

A distributed PASCAL compiler is being implemented by James Frankel as part of a Harvard University Ph.D. thesis to explore the problems of distributing a large, predictable computation over several network-connected computers. A completely dynamic resource allocation scheme is being used, in which the compiler explores the network to find out how many free computers are around, then it assigns parts of the compilation computation task accordingly.

Publications

1. Saltzer, J., Reed, D., and Clark, D. "Source Routing for Campus-Wide Internet Transport," IFIP Working Group 6.4 Workshop, Workshop on Local Area Networks, Zurich, Switzerland, August, 1980.

2. Saltzer, J., Reed, D., and Clark, D., "End-to-End Arguments in Systems Design," Proceedings Second International Conference on Distributed Computing Systems, Paris, France, pp. 509-512.

3. Kent, S., "Protecting Externally Supplied Software in Small Computers," MIT/LCS/TR-255, MIT, Laboratory for Computer Science, Cambridge, Ma., September 1980.

Theses Completed

1. Fitzgerald, T., "The Implementation of a File Transfer Protocol on Multics," S.B. thesis, MIT, Department of Electrical Engineering and Computer Science, Cambridge, Ma., May 1981.

2. Kent, S., "Protected Externally Supplied Software in Small Computers," Ph.D. dissertation, MIT, Department of Electrical Engineering and Computer Science, Cambridge, Ma., September 1980.

3. Thomas, J., "A Multi-Protocol Network Mail Transport Facility for Multics," S.B. thesis, MIT, Department of Electrical Engineering and Computer Science, Cambridge, Ma., September 1980.

Theses in Progress

1. Baldwin, R., "An Evaluation of the Recursive Machine Architecture," S.M. thesis, MIT, Department of Electrical Engineering and Computer Science, Cambridge, Ma., expected February 1982. (Also S.B.)

2. Houldin, R., "Formats and Controls for a One-Line Computer Terminal Display," S.B. thesis, MIT, Department of Electrical Engineering and Computer Science, Cambridge, Ma., expected August 1981.

3. Lopez, L., "Gateway Congestion Control," S.M. thesis, MIT, Department of Electrical Engineering and Computer Science, Cambridge, Ma., expected January 1982.

4. Ludwig, C., "An Implementation of a Portable Personal Terminal," S.B. thesis, MIT, Department of Electrical Engineering and Computer Science, Cambridge, Ma., expected January 1982.

5. Martinez, D., "A Central Switcher for Line Switched Message Oriented Computer Input/Output," S.M. thesis, MIT, Department of Electrical Engineering and Computer Science, Cambridge, Ma., expected December 1981. (Also S.B.).

6. Meier zu Sieker, F., "A Telex Gateway for the Internet," S.B. thesis, MIT, Department of Electrical Engineering and Computer Science, Cambridge, Ma., expected December 1981.

7. Simpson, G., "NEMO- a monitoring station for a local area network," S.M. thesis, MIT, Department of Electrical Engineering and Computer Science, Cambridge, Ma., expected February 1982. (also S.B.).

8. Singh, V., "A Routing Service for Campus-Wide Internet Transport," S.M. thesis, MIT, Department of Electrical Engineering and Computer Science, Cambridge, Ma., expected September 1981.

Talks

1. Clark, D., "On the Limits of Distributed Atomic Update,"

   Workshop on Fundamental Issues in Distributed Computing,
   Pala Mesa, Ca., December 1980;
   5th Berkeley Workshop, Berkeley, Ca., February 1981,
   also Program Committee and Session Chairman.

2. Saltzer, J., "Why A Ring?" Xerox Palo Alto Research Center, December 1980.

3. Saltzer, J., "End-to-End Arguments in System Design,"

   University of Cambridge, August 1981;
   Workshop on Fundamental Issues in Distributed Computing,
   Pala Mesa, Ca., December 1980;
   IBM San Jose Research Laboratory, December 1980;
   INRIA Workshop on Distributed Computing, April 1981.

4. Saltzer, J., "Rings, Ethernets and Broadband, the Underpinnings of Local Networks," New York Academy of Sciences, May 1980.

5. Saltzer, J., "Evolution of Distributed Computer Systems," Computer Science Conference Address, National Security Agency, May, 1980.

Committee Memberships

Clark, D., DARPA IPTO Internet TCP Working Group

Chiappa, N., DARPA IPTO Internet TCP Working Group

Martin, E., DARPA IPTO Internet TCP Working Group

Saltzer, J., DoD/DDRE Security Working Group Member