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

September 15, 1980
Request for Comments No. 198

Trip Report August 24 - September 7, 1980

by David P. Reed

My itinerary was: IFIP WG 6.4 Zurich Workshop on Local Area Computer Networks, 3 days; INRIA (French national computer research institute), 2 days; Cambridge University Computer Laboratory, 1 day. A planned visit to University of Newcastle-upon-Tyne was cancelled because too few of the scientists were available.

Also attending the Zurich workshop was Jerry Saltzer, whose trip report overlaps with this one.

1. IFIP WG 6.4 Workshop on Local Area Computer Networks --

Jerry Saltzer's report summarizes much of the information gathered here. As Jerry points out, there were a large number of attendees (93) from many countries. This led to a rewarding exchange of information.

I was particularly impressed by the number of locations and projects that are using high-speed local networks and by the number of network products available (or about to be avilable) commercially. Since many of these projects are well past the design stage, there were lots of people with significant experience in local network construction, maintenance, and application.

-- An observation made independently by several people is that the protocols designed for local networks must be concerned primarily with optimizing the host software and hardware, rather than with optimizing the network. Such problems as host interfaces that give too many interrupts, host interfaces that have to be "turned back on" after each packet received, host software difficulties in processing header formats, etc. all seemed to dominate discussions of network performance. In one sense, this is not surprising, since we have long contended that optimizing bandwidth on the local net by packing fields in packets, etc., is a bootless exercise.

-- CMU has begun a significant effort in distributed computing, motivated by the need for data sharing among PERQ's, VAXES, and other major workhouse computers and by their interest in distributed sensor networks. One of their projects is a central file server that will hold most of the shared files. They hope to use optical disks. Several of their ideas are similar to those of our SWALLOW project, particularly immutable, multi-version files.

-- IBM's Zurich laboratory seems to have jumped feet first into local networking. Their approach, however, is unusual -- they are using high-speed packet-switching nodes in a mesh-connected network, rather than simple rings or buses.

-- The use of local networks for real-time process control systems does not seem to be well understood, but there is strong interest in this use. There is a general concensus that environmental concerns are important, especially in the chemical plants and coal mines. However, questions about reliability and delay-variability requirements get widely varying answers. For example, some engineers want bit-error rates of 10^{-15} on the wire, while others seem nowhere near so criticial. Similarly, some engineers see the lack of fairness guarantees on the Ethernet as a serious problem, while others point out that statistical fairness at low Ethernet loads is adequate.

-- There seems to be general agreement that in most cases local area networks do not exist isolated from other communication networks. Thus most of the people at the conference were interested in using "internet" protocols of some sort for their local nets.

A highlight of the workshop was the discussion of the Xerox-DEC-Intel Ethernet standard and the developing IEEE local network protocol standards effort led by Gerry Clancy. J. Saltzer in his trip report summarizes the Xerox-DEC-Intel talk.

I found the IEEE standards proposals somewhat amazing. Clancy has led a large group of people in a crash effort to publish a standard at about the same time the Ethernet standards are published. The goals of Clancy's group seem much more ambitious, in some sense, than the Ethernet standard. They are very interested in a "technology-independent" interface and a speed independent interfae at very low levels, i.e., the "plug on the wall" to which you connect your computer port.

The success of both the IEEE standard and the Xerox-DEC-Intel effort may depend on the technical feasibility of the goals they have set themselves.

2. INRIA, Project Sirius --

Project Sirius is an effort to develop Sirius-Delta, a distributed database system aimed at real-time transaction processing applications. I spent both days discussing with them the structure of their system, and ideas about synchronization and recovery. Their system, being constructed on Realite 2000 minicomputers connected by a local network, is divided into three layers -- SER, a distributed executive that controls the communication and sequencing of steps in a single, multi-site transaction; SCORE, the concurrency control and synchronization system for managing distributed data; and SILOE, which analyzes a transaction and determines the steps to be executed, locating the data to be accessed. The head of the project is J. LeBihan, but I spent most of my time talking to Gerard LeLann and Simone Sedillot.

The SCORE system is quite interesting, since it embodies a complete lock-based synchronization and recovery scheme. One interesting feature is their variation of the two-phase commit protocol, which seems to be more reliable than that of Lampson and Sturgis. When actions at several sites are to be committed, the "prepare" phase message sent includes a list of the names of all sites involved. This information is then stored with the updated versions on stable storage. The commit message is then

broadcast by the coordinator. Once any one site receives the "commit" message, it commits the updates. If the coordinator is lost, then the prepared sites can inquire at each other prepared site to see if the transaction is committed. The idea is quite similar to the "voting" commit record idea in my dissertation.

3. **Cambridge University Computer Laboratory --**

At Cambridge I visited Roger Needham, who just assumed his new position as head of the Computer Science Department, replacing Maurice Wilkes who is now at DEC in Maynard.

My major interest was in learning about the Cambridge File Server. We had some useful discussions on this topic, and I discovered some interesting and useful things. They use an interesting strategy of multiple disk record sizes to optimize the space for small objects. They seem very happy with this decision.

They have an interesting strategy for garbage collection of the storage on the file server. This is controlled and managed by a remote node, and the interface is designed so that a failure at the remote node during garbage collection does not impact the file server.

Roger and I had a long discussion on the topic of the vulnerability of distributed systems to malicious or erroneous behavior. Our conclusion was that it is a serious problem, but not one that is unique to distributed systems. For example, even in systems like Multics, it is possible to disrupt service to others (e.g., the "plough" program). However, it is much easier to detect and locate such problems in a centralized time-shared system. It seems to be an important problem, yet there seems to be no good solution.