LABORATORY FOR
COMPUTER SCIENCE

MASSACHUSETTS
INSTITUTE OF
TECHNOLOGY

MIT/LCS/TM-254

EMPIRICAL ANALYSIS OF A TOKEN RING NETWORK

DAVID C. FELDMEIER

JANUARY 1984

545 TECHNOLOGY SQUARE, CAMBRIDGE, MASSACHUSETTS 02139

Empirical Analysis of a Token Ring Network

David Charles Feldmeier

January 1984

This research was supported in part by the Defense Advanced Research Projects Agency of the Department of Defense and monitored by the Office of Naval Research under contract number N00014-75-C-0661

Massachusetts Institute of Technology
Laboratory for Computer Science
Cambridge, Massachusetts 02139

Empirical Analysis of a Token Ring Network

by

David Charles Feldmeier

Submitted to the
Department of Electrical Engineering and Computer Science
on January 26, 1984 in partial fulfillment of the requirements
for the Degree of Bachelor of Science

Abstract

The MIT Laboratory for Computer Science 10 Megabit token ring local area network was monitored. Over a one-week period 7 million packets and 1.3 billion bytes passed by the monitor. This thesis compares the MIT ring traffic with that observed on the Xerox Palo Alto Research Center experimental Ethernet by Shoch and Hupp.

- Packet length distribution on the ring was bimodal - packets were either short (less than 100 bytes) or long (between 530 and 576 bytes). This is consistent with observations on other local area networks. Most packets on the ring were small, but most of the bytes were transferred in large packets.

- 49% of the packets on the ring were internet packets, 46% were intranet packets, and 5% were transit packets.

- Ring hosts generated 4 times the load per node of Ethernet stations, suggesting that network applications are different in the two environments.

- Network utilization was a small fraction of its capacity - over the monitored week the busiest day had an average network load of 0.3%. Traffic occurs in bursts - the busiest hour had a 1.4% netload, the busiest minute had a 6% netload, and the busiest second had a 66% netload.

- An upper bound on the amount of ring down time was 1.4% of its total running time. This is the amount of time that no token existed, though the ring was not necessarily out of service that entire time.

- The probability of a bit error on the ring is estimated at less than 10^{-12}.

- Interpacket arrival time roughly followed a Poisson process.

- The network monitoring system is a useful diagnostic tool.

This thesis describes the monitoring tools and techniques, reports the measurements in detail, and compares the measurements with other published results.

Key Words: network measurement, network monitoring, rings, local networks

Acknowledgments

I would like to thank my thesis advisor Jerry Saltzer for providing his thoughts and encouragement throughout this thesis and the UROP work that preceded it. His advice was a great help during the design and building of the monitoring system and the writing of this thesis.

I would also like to thank Larry Allen, Dave Bridgham and John Romkey for their help. All three were constantly answering my questions about the PDP 11 computer, the C programming language and the Unix operating system. Without their help, the monitoring station would have taken much longer to build. They also helped me to explain puzzling observations of the ring network.

I would finally like to thank the people in the CSC and CSR research groups at the MIT Laboratory for Computer Science for making the lab an enjoyable place to work.

I dedicate this thesis to my parents, John and Carol, both of whom have done so much for me.

Table of Contents

Table of Figures

Table of Tables

Chapter One

Introduction

1.1 Purpose

The purpose of this project was to do an empirical study of the performance, reliability and traffic patterns of the Laboratory for Computer Science version two local area network. Performance data allows the comparison of the version two ring specifically and token ring networks in general with other local area network strategies, such as the Ethernet, to discover which type of local area network is decidedly better and why. Traffic patterns indicate how computer networks are used.

Performance analysis can be performed both theoretically and experimentally. Mathematical analysis of a local area network is complex and only roughly approximates the behavior of a network. Mathematical analysis is useful for determining which systems might be usable, but does not give enough information to choose among the different systems. Network data refines mathematical models for network evaluation. Improved mathematical models make evaluation and simulation of networks before construction more practical.

Empirical analysis determines the performance of an network. Not only can the network traffic be evaluated, but also network properties can be determined. Reported traffic on a local area network can improve mathematical models of traffic for theoretical evaluation and also suggest the traffic handling properties that a local area network and its hosts should have.

Contention bus networks are more popular in the United States because of the influence of the Ethernet, ring networks are more popular in Europe because of the Cambridge Ring. Although evaluation has been performed on the Ethernet, little has been done with token ring systems. This thesis will allow some comparison of the two major network control systems.

1.2 A Short History of Previous Work

The analysis of the token ring network was inspired by the paper *Measured Performance of an Ethernet Local Network* by John F. Shoch and Jon A. Hupp at Xerox Palo Alto Research Center. Their paper contains the result of network monitoring done on a 2.94 Megabit Ethernet at Xerox PARC. [8] The Ethernet report is the best work on local area networks that exists.

Robert V. Vieraitis Jr. wrote an undergraduate thesis about *A Performance Monitor for a Local Area Network.* [11] His thesis reports statistics gathered for the 1 Megabit version one token ring at the MIT Laboratory for Computer Science. Statistics for the version one network are valuable because work was done at the same lab with some of the same hosts on a token ring network.

This thesis is part of a planned comparison of the ring and the Ethernet proposed by Liba Svobodova in an MIT Laboratory for Computer Science internal working paper. Measurement of an operational network should answer some performance questions about local area networks.

1.3 A Brief Statement of Work Done

Before the work presented in this thesis, specialized hardware and software was constructed for network monitoring of the token ring. This *monitoring station* consists of a network interface and some custom hardware placed in a Digital PDP 11/10 computer. The monitoring station analyzes network data and runs a real-time display. Network data is compressed into large packets that are sent over the network to the *analysis machine.*

Work for this thesis was the programming of the analysis machine and the interpretation of the results gathered by the monitoring system. The analysis machine is a Digital VAX 11/750 running a server to receive and analyze compressed data packets from the monitoring station. Other programs put the analyzed data into a useful form for interpretation. The resulting data is presented in this thesis.

Chapter 2 introduces the version two token ring network and chapter 3 discusses network monitoring. Chapter 4 presents the data from network monitoring and chapter 5 compares these results with the data from other studies. Chapter 6 draws some conclusions about the research done on the version 2 ring network and chapter 7 provides some ideas for further study.

Chapter Two

Network Environment

2.1 Ring Network Characteristics

2.1.1 Physical Structure

The Version Two Laboratory for Computer Science Ringnet Local Area Network is a ten megabit star topology ring with token passing channel allocation and decentralized control. [5] Proteon, Inc. markets the version two ring under the name *proNET*. The ring at the Laboratory for Computer Science has about thirty-three stations on three different floors. The total length of the ring is about 1000 meters with 110 meters of fiber optic cable over one run, and an active repeater built into another cable run. The ring topology is a modified to be star shaped for ease of maintenance. [4] All ring nodes connect at a central point called the *wire center*. The wire center is passive and consists of relays and their associated driving circuitry. Nodes activate relays electronically to join the ring, otherwise the relay bypasses the node. If a node becomes disruptive to the ring, manual switches in the wire center allow any node to be disconnected from the ring from a central point.

The configuration of the ring at the MIT Laboratory for Computer Science is four stars, each with its own wire center, that connect together to form a continuous loop. Two stars are on the ninth floor, a star is on the fifth floor, and a star is on the second floor. The connection between floors nine and five is via a fiber-optic cable, and the connection between the fifth floor and the second floor is through an active repeater. An active repeater divides the length of cable from a fifth floor to the second floor and back to the fifth floor so the maximum wire distance between nodes is not exceeded if no host is connected on the second floor. An active repeater or fiber-optic link also automatically separates the ring into two operational pieces if the connection between wire centers is severed.

The ring accommodates 255 nodes, each with a unique address. Since the local address slot is a byte in length, the 256th address is used for the the *broadcast address*. Every node on the ring receives packets to the broadcast address regardless of the node's assigned address.

2.1.2 Ring Control

The ring control system is completely decentralized, which means that all nodes on the ring follow the same algorithm at all times. The algorithm depends on the fact that the nodes on the ring act in an asynchronous manner. The IBM experimental ring is not completely decentralized by this definition because at any given time there is a single station that monitors the network status and controls ring access, even though any station can be the monitor. [1]

The ring is controlled with various control characters. *Flags* comprise all control characters; a flag is a zero followed by seven ones. Bit stuffing distinguishes control characters from data. Bit stuffing is the automatic transmission of a zero after the transmission of six consecutive ones. The receiving node checks the seventh bit after the six ones. If it is a one, then a control character is being received; otherwise, the receiving station simply ignores the seventh bit and continues to receive as usual.

During normal operation, a circulating control character, the *token*, controls transmission. A station that receives a token and has a message to send may change the token into a *connector* (beginning of message) control character. The station appends the message to the connector, adds an End of Message (EOM) control character, a parity bit, a message not received bit and finally a new token. Ending a message with a token allows the next node with a message to transmit immediately after the previous message has passed. Each node is responsible for the removal of its own message from the ring. The receiving node simply copies the message as it passes by, changing at most the message not received bit at the end of the message. The message not received bit is examined on removal of the message from the ring to determine whether the destination node has received the message.

Nodes must determine whether normal operation of the ring has been disrupted and if so,

recover. Timers at each node trigger after 4 milliseconds with no control characters on the ring or 700 milliseconds with no token. If a node has a message to transmit and a watchdog timer has triggered, then transmission is begun immediately without waiting for a token. Each transmission ends with a token and this has the side effect of generating a new token on the ring. Contention is possible during recovery, but is rare. [6]

A complete description of the ring is contained in the operation manual for the proNET local area network. [3]

2.1.3 Computers on the Ring

The ring currently has about thirty-three hosts. Most of the computers on the ring are small timesharing systems, VAX 11/750s running Unix. Other hosts on the ring include a timesharing PDP 11/45, a PDP 11/40 and an IBM personal computer. The remaining machines on the ring are gateways to other networks. The only servers that are on the ring are Remote Virtual Disk servers. These servers are timesharing VAX 11/750s running Unix with large amounts of disk storage. Other computers can use virtual disks on these servers via the network. All other servers are on different networks. To determine the type of traffic that the gateways might contribute, the types of networks, the computers and services on those networks must be known. Page 17 illustrates the layout of the ring and other attached networks at the Laboratory for Computer Science.

2.2 Other MIT Networks

The ring is one of several networks that are on the MIT campus. This section describes how the ring net interacts with other networks on campus.

2.2.1 The ARPANET

The ring net is connected to ARPANET via two gateways, one of which is experimental. The ARPANET is a major path to the outside world and is also the path over which other machines on the ARPANET at MIT are accessible. This is the usual route to the

laboratory's Digital 20/60 research computer that is also a name server.

2.2.2 Three Megabit Ethernet

The ring is connected to a 2.94 Megabit Ethernet via a gateway called the *Bridge*. The 2.94 Megabit Ethernet has about 20 Altos (Xerox personal workstations), a file server for the Altos that is also a nameserver, a laserprinter (the main one for MIT), a spooler that translates from Internet Protocol (IP) to PARC Universal Packet (PUP) protocol to send files to the printer, and lastly a PDP 11 that is the front end for the Artificial Intelligence (AI) computer that also connects to the *Chaos* net and acts as a gateway between the 2.94 Megabit Ethernet and the Chaos net.

The Chaos net is a 4 Megabit virtual token bus topology network. Many of the machines on the MIT campus are attached to the Chaos net. Machines on the Chaos net speak the Chaos protocol. The Chaos net consists of several subnets; subnet 6 serves the Laboratory for Computer Science, and subnet 1 that serves much of campus. Both of these subnets connect to the Artificial Intelligence PDP 11 front end gateway.

2.2.3 The Ten Megabit Ethernet

Attached to the ring via yet another gateway is a 10 megabit Ethernet. This Ethernet is to eventually become a main network in the laboratory. It will consist of a spine that runs through the main computer rooms and down the center of the building. An Ethernet repeater attaches to the spine and repeats to an Ethernet for each floor. The system now covers three floors and serves 10 IBM personal computers and a Symbolics 3600 Lisp Machine.

2.2.4 The PC Gateway

The PC gateway allows personal computers to be attached to the ring via a serial line. This permits the attached personal computers to transmit and receive packets on the ring. The PC gateway can handle eight personal computers over serial lines, including dial up lines.

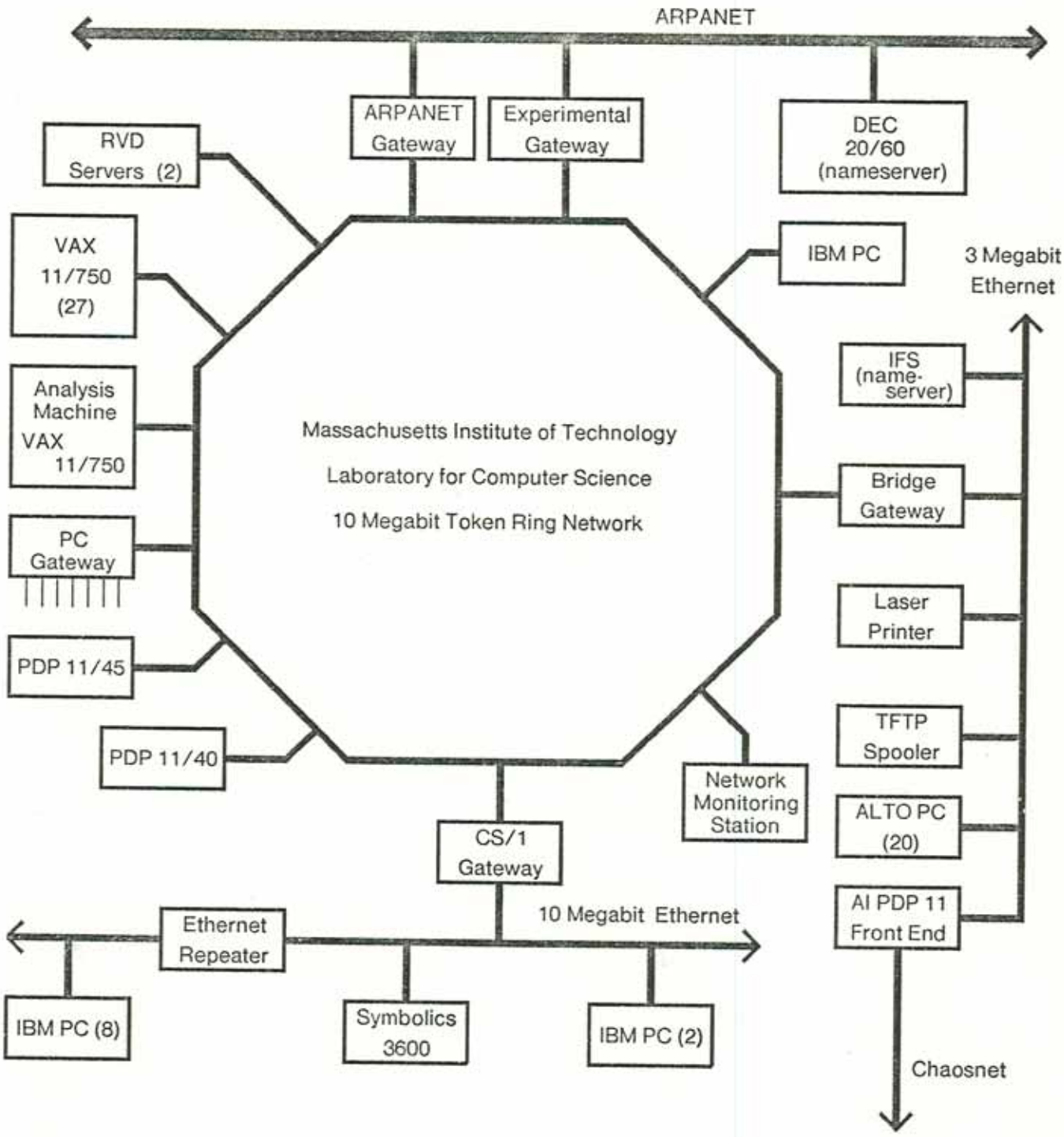


Figure 2-1: The MIT LCS 10 Megabit Ring Network

Chapter Three

Network Measurement

3.1 Types of Network Monitors

There are several ways to monitor a computer network, with various advantages and disadvantages to each. The possibilities considered below were discussed in a Laboratory for Computer Science internal working paper by Liba Svobodova.

3.1.1 Monitoring All Hosts

An obvious type of network monitoring is to record the network traffic on all the network hosts. Monitoring software will affect the performance of hosts on the network, altering the network load. On a network that cuts across administrative boundaries, permission to modify individual hosts may be difficult to obtain. Another problem is the correlation of data. To correlate the distributed data, an accurate clock that is synchronized across the various hosts is necessary. On a large network, the amount of data could be overwhelming. For these reasons, the monitoring of all hosts on the ring was abandoned as impractical.

3.1.2 Monitoring of Selected hosts

A simpler version of the above scheme is to monitor selected hosts on the network. The advantage over the monitoring of all hosts is that it is easier to monitor fewer hosts. Less data is combined, it is easier to get permission to work with a subset of hosts, and the network traffic is not affected as much by monitoring software overhead. The disadvantages are that not all data is recorded and if the monitored hosts are atypical, then the collected network data is skewed.

3.1.3 Probe Monitor

The probe monitor is an active monitor, unlike the other monitors discussed. A probe monitor injects traffic onto the network at intervals, an approach that corresponds to a random sampling. The probe transmissions discover the network state by their transmission delay. Probe monitors give little information about the traffic on the network during normal usage.

3.1.4 Spy Monitor

The spy monitor is a passive system that receives and analyzes all traffic passing through its network interface. A spy monitor perturbs the network by adding a network interface, but no other modifications to the network or its hosts are necessary. It collects statistics over a long period with little inconvenience and it collects more detailed statistics than a probe monitor can. A spy monitor must be fast and have a large amount of storage. It must be able to assimilate data at the maximum speed of the network over at least short intervals of time. On a network that has more than a million packets a day as does the ring, the storage requirements are in the tens of megabytes per day.

3.2 Monitoring the Ring

A passive spy monitoring system was chosen for simplicity and effectiveness. The original design of the monitoring station was a totally passive design in which a Digital PDP 11/20 computer was to do all the monitoring and analysis of the network data. The PDP 11/20 was not powerful enough, so a larger computer was needed. A larger computer could be used for little else besides network monitoring; servicing the network interface would use most of the processing power. As a compromise, a PDP 11/10 does all the receiving and pre-processing that is necessary, discarding packets with improper format and performing other low level functions. This pre-processed information is transmitted to the analysis machine for further analysis and long term storage of statistics. The monitoring system is passive, using network bandwidth for the transfer of data to the analysis machine. Other hosts on the ring are unaltered. Packets produced by the monitoring station cause 1.5% of

packets and 7% of the bytes on the ring.

3.2.1 The Network Monitoring Station

The *network monitoring station* is a spy type of monitor that uses a small amount of network bandwidth to report select statistics to an *analysis machine* on the ring. The monitoring station was built as a spy monitor to allow it to be a permanent addition to the ring for experimental, testing, and debugging purposes. The monitoring station is a Digital PDP 11/10 computer with specially constructed hardware that does data extraction and compression with some real time analysis.

The ring network interface hardware has two parts. The first is the control card (CTL). The CTL is the network interface that contains the modem and the low level transmit and receive mechanisms. The CTL converts data on the ring into 8 bit byte format that is transferred to the Host Specific Board (HSB). The HSB receives the information in byte format and stores it in a buffer for later DMA to the host.

To be effective, a monitoring station should miss as few packets as possible. Ring hardware must be reset after a packet and packets are not received while the hardware is resetting. The monitoring station CTL goes to special hardware that switches between two HSB cards. The first board receives as many packets as possible, but if the first board is unready to receive a packet, the special hardware switches the incoming packet to the second HSB board. The monitoring station's dual buffer strategy prevents packets from being missed. See figure 3-1.

The special hardware performs other functions such as timestamping arriving packets, finding the length of the incoming packet, and watching for token acquisition and loss on the ring. A frequency counter attached to the monitoring station observes signals on the CTL card. The counter measures ring frequency and inter-token time on the ring.

The monitoring station does data selection and compression. Eight bytes are stored for each received packet: ring destination, ring source, Internet Protocol, two bytes of packet length,

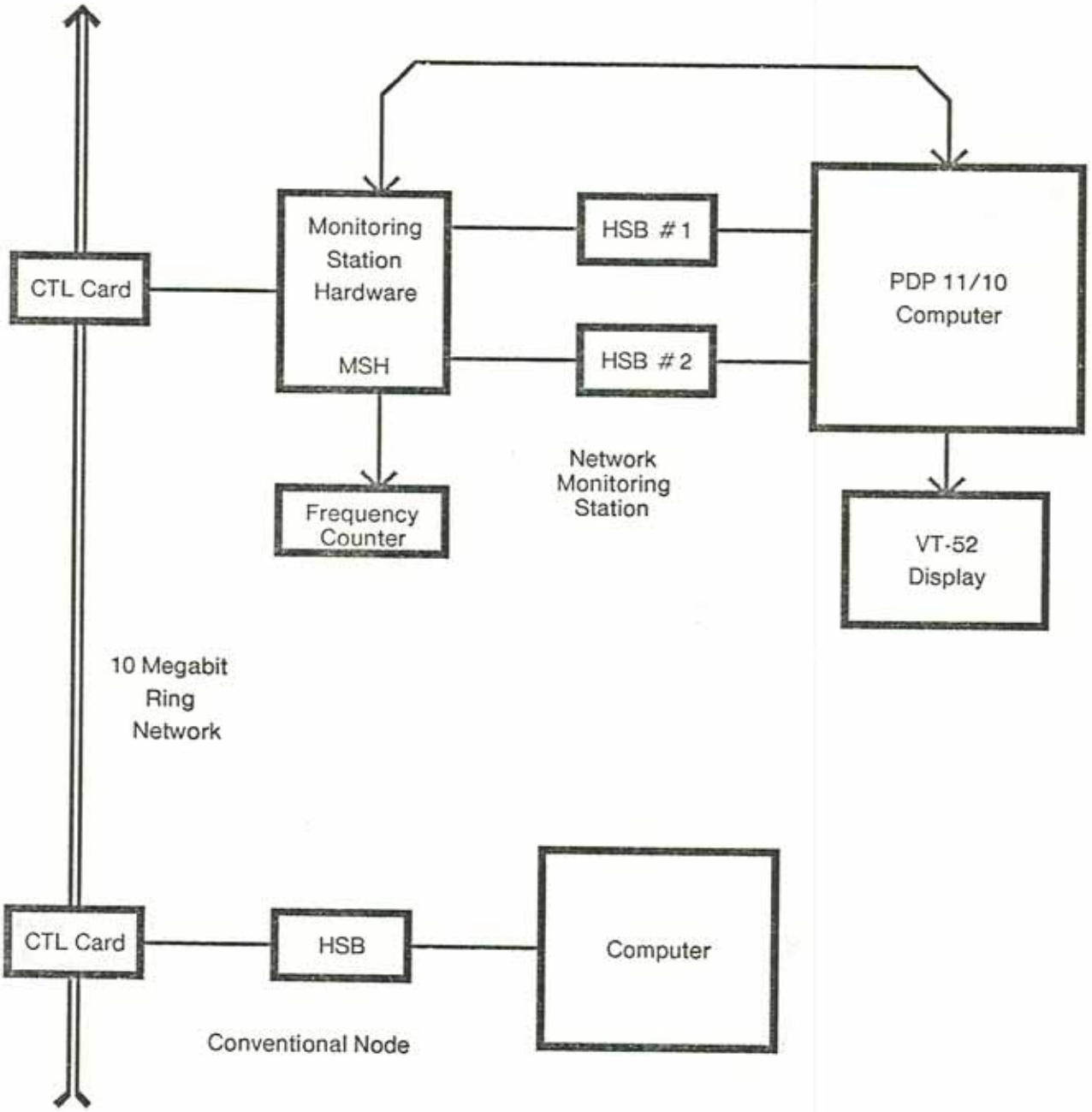


Figure 3-1: The Network Monitoring System

and three bytes of 25 microsecond resolution timestamp. Sixty-seven of these eight byte packet fields are sent to the analysis machine in each compressed data packet. Compressed data packets are sent unreliably to the analysis machine; the monitoring station follows a try-at-most-once transmission strategy to keep the netload contribution of the monitoring station low. About 5% of the data is lost because of unreliable transmission.

Appendix A contains a detailed description of the monitoring station.

3.2.2 The Analysis Machine

The *analysis machine* receives compressed data packets from the monitoring station via the ring. The analysis machine is a Digital VAX 11/750 timesharing computer that has a light network traffic load. The compressed data is analyzed to determine network characteristics and the results of analysis are stored. The analysis machine processes compressed data packets as they are received, statistics are collected and the original data is discarded.

The analysis program has a table of network usage during the day in 10 minute intervals.

Interpacket arrival time is computed from the timestamp of each packet. The interpacket arrival time histogram has 2000 slots of 1 millisecond resolution.

The packets are separated by protocol. For each protocol, there is a 256x256 host-table of source-destination pairs. A table of packet lengths is also kept for each protocol. Other programs interpret the data in the analysis tables to produce the data presented in chapter 4.

Appendix B contains a detailed description of the analysis machine.

3.2.3 Long Term Records

The monitoring station can generate about of fifteen megabytes of raw data on a reasonably busy day. It is impractical to keep all this data in storage, so it is analyzed immediately. To have some long term records, some "typical" days are to be recorded on a VAX that has a large amount of disk storage, with the data transferred to tape for long term storage and

later analysis. If something could not be determined from the analyzed data, information could be recovered from these tapes of raw received data.

Chapter Four

Experimental Results

The data in this section is mainly from a one week period beginning November 30th at 00:07 and ending December 6 at 23:54. The analysis machine logged 7,035,947 packets and 1,329,904,120 bytes over that week. The analysis machine failed to receive data about 374,798 packets (5%) from the monitoring station. The monitoring station missed less than 1% of the packets on the ring (estimated).

4.1 Network Traffic versus Time of Day

The number of packets per second received averaged over 10 minute intervals is shown in the histogram in figure 4-1. Network usage rises in the morning, is up sharply between 9:00 and 10:30 and begins to fluctuate around noon. Network usage is highest in the afternoon until it drops sharply at 6:00 in the evening. Usage remains high until 3:00 AM because of late workers, mainly students who have classes during the work day.

The Remote Virtual Disks cause the large spike that occurs at 5:10 AM. Spinning down the Remote Virtual Disks at 3:00 AM for system maintenance causes little traffic and traffic on the ring drops sharply by 3:10 AM. Machines spin the Remote Virtual Disk disks up at 5:00 AM and run file checks on the disks, causing many packets at this time.

4.2 Network Utilization

Network utilization is measured by the number of packets on a network or by the percentage of network resources being used. The network load is computed by counting the bytes in a packet, adding the local network encapsulation overhead and dividing by the network capacity.

Network utilization is low, with the largest recorded number of packets in a day being

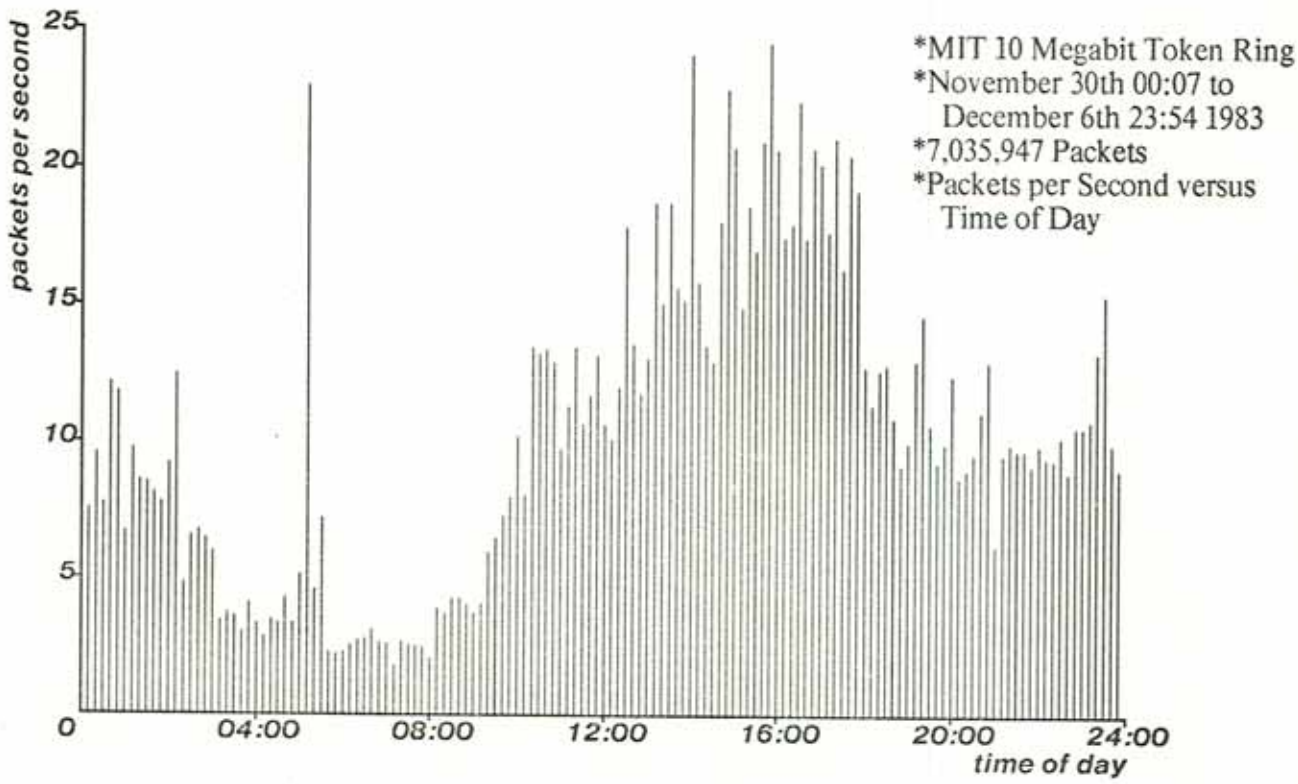


Figure 4-1: Packets per Second versus Time of Day

1,553,699, with an average network load of 0.30%. The highest load was 0.58% for a day with 1,401,827 packets. For the week of November 30 to December 6, the busiest day had 1,448,542 packets with a 0.26% netload. Utilization is higher over shorter periods. The busiest hour had 261,447 packets and a 1.4% netload; the busiest minute had 30,003 packets and a 5.6% netload; the busiest second had 1184 packets with a 66% netload. The 66% netload figure seems high. Occasionally when a node enters the ring, garbled data resembling a packet is received by the monitoring station, causing an apparent high network load until it is recognized as invalid. If this invalid data presents a 100% load and the no-token time-out is 700 milliseconds, the monitoring station would have seen a 70% netload; close to the 66% load observed.

4.3 Packet Length Distribution

Packet length distribution on the ring is bimodal - most packets are either small (less than 100 bytes in length) or large (between 530 and 576 bytes in length). The ring hardware receives packets up to 2048 bytes long, but software limits the size to 576 bytes (although some work is being done at 1082 bytes). The packet length distribution histograms on page 27 show lengths only up to 600 bytes because few packets are longer than 576 bytes.

4.3.1 Percentage of Packets versus Packet Length

Figure 4-2 shows a histogram of the percentage of packets at each length. Telnet (remote login) account for the two large spikes at 46 and 48 bytes. Remote Virtual Disk one block file transfers caused the smaller spike at 570 bytes. The monitoring station transmits 576 byte packets to the analysis machine and an incorrectly operating gateway sent the 578 byte packets.

4.3.2 Percentage of Bytes versus Packet Length

This histogram in figure 4-3 is similar to the previous one, except that each slot has been weighted by packet length. The largest spike is at 570 bytes, which suggests that most of the bytes on the network are sent during Remote Virtual Disk file transfers; the actual figure is 41% of all bytes.

4.4 Source - Destination Traffic Patterns

Tables 4-1 and 4-2 present the busiest hosts on the ring. The most active transmitters on the ring were the ARPANET gateway and the Remote Virtual Disk servers. The ARPANET gateway and one Remote Virtual Disk server each contributed 12% of the traffic on the ring. A timesharing system also contributed largely to the traffic. The timesharing machine is active because it is heavily used; most people login remotely over the network and use Remote Virtual Disk storage. The development machine is used to develop network code, which explains its many transmissions.

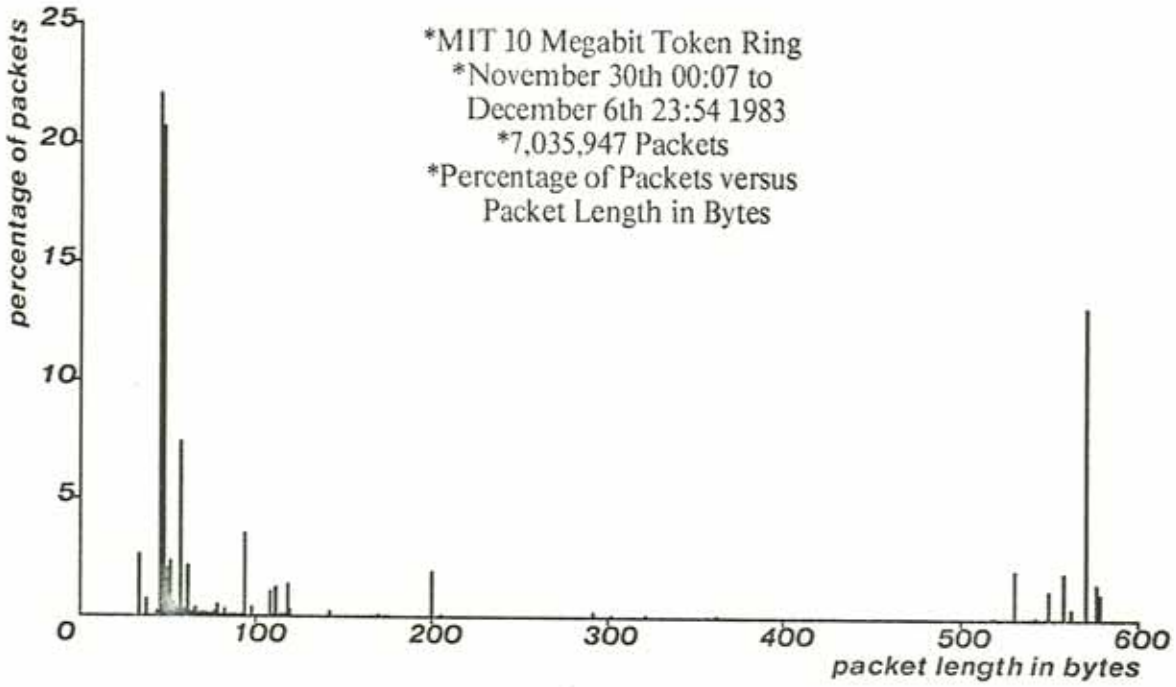


Figure 4-2:Percentage of Packets versus Packet Length

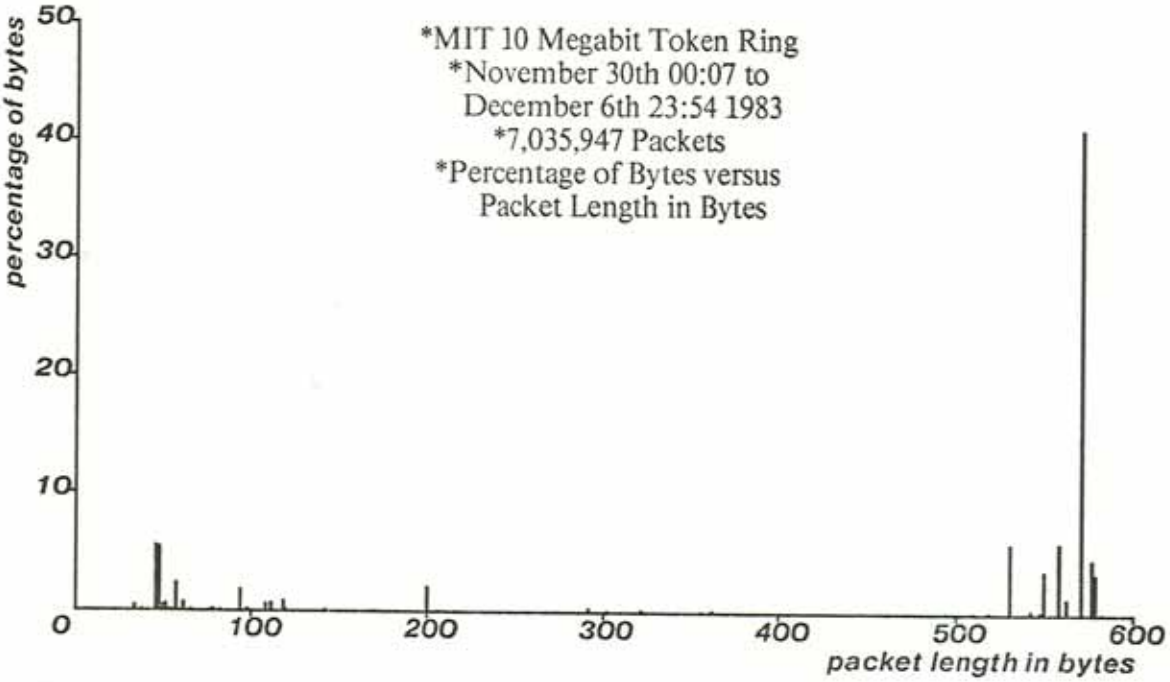


Figure 4-3:Percentage of Bytes versus Packet Length

Notice in figure 4-2 that almost 8% of the packets are destined for the unassigned address 0, which is an illegal address on the ring. Berkeley Unix 4.2 code being tested on two VAX 11/750s sent out a broadcast packet once a minute, but instead of sending to the ring broadcast address (address 255), packets were sent to address 0. Many incorrect packets were sent during the debugging of the Unix code.

Servers on the ring, specifically gateways and Remote Virtual Disk servers, transmitted 48% and received 46% of all packets on the ring.

Host Name	Ring Address	Primary Function	Packets Transmitted	Percentage of Packets
Gateway	4	gateway	850,955	12.09%
Milo	86	RVD server	840,496	11.95%
Borax	65	timeshare	730,176	10.38%
Bridge	5	gateway	541,607	7.70%
Opus	87	RVD server	530,599	7.54%
Dutch	67	development	391,976	5.57%
CLS	76	timeshare	383,502	5.45%
		total	4,269,311	60.68%

Table 4-1: Hosts that generate at least 5% of all packets

4.5 Interpacket Arrival Time

Packets received by the monitoring station are timestamped and a histogram of interpacket arrival time is produced by the analysis machine using the timestamps. The histogram consists of interpacket arrival times of 2 seconds or less with a resolution of 1 millisecond, although resolutions down to 25 microseconds are possible.

Host Name	Ring Address	Primary Function	Packets Received	Percentage of Packets
Gateway	4	gateway	1,000,909	14.23%
Opus	87	RVD server	679,677	9.66%
Borax	65	timeshare	575,897	8.19%
Unassigned	0	invalid	556,913	7.92%
Milo	86	RVD server	550,770	7.83%
Bridge	5	gateway	538,421	7.65%
		total	3,902,587	55.47%

Table 4-2: Hosts that were destinations of at least 5% of the packets

The interpacket arrival time histogram for 1 second is the figure on page 30. The histogram has the *percentage of packets* axis plotted on a logarithmic scale to determine whether interpacket arrival time is a Poisson process. The result is not strictly Poisson but can be reasonably approximated by the superposition of 2 or 3 Poisson functions. The short period Poisson is possibly because of the Remote Virtual Disk protocol. With this protocol, a single packet is sent to the server to request several disk blocks. Another cause is telnet remote login character echoes. The long period Poisson is likely because of telnet remote login characters sent by a remote terminal. Remote login and Remote Virtual Disk applications account for many packets on the ring.

Several distinct spikes occur on the curve above the noise. The largest spikes are at 1, 6, 16, 19, and 38 milliseconds. Smaller peaks exist between 200 and 300 milliseconds, around 700 milliseconds, and near 1 second. These spikes are likely the turn around time for transactions on the ring, such as a telnet (remote login) character echo. Another explanation is the time between consecutive blocks from a disk in a Remote Virtual Disk file transfer.

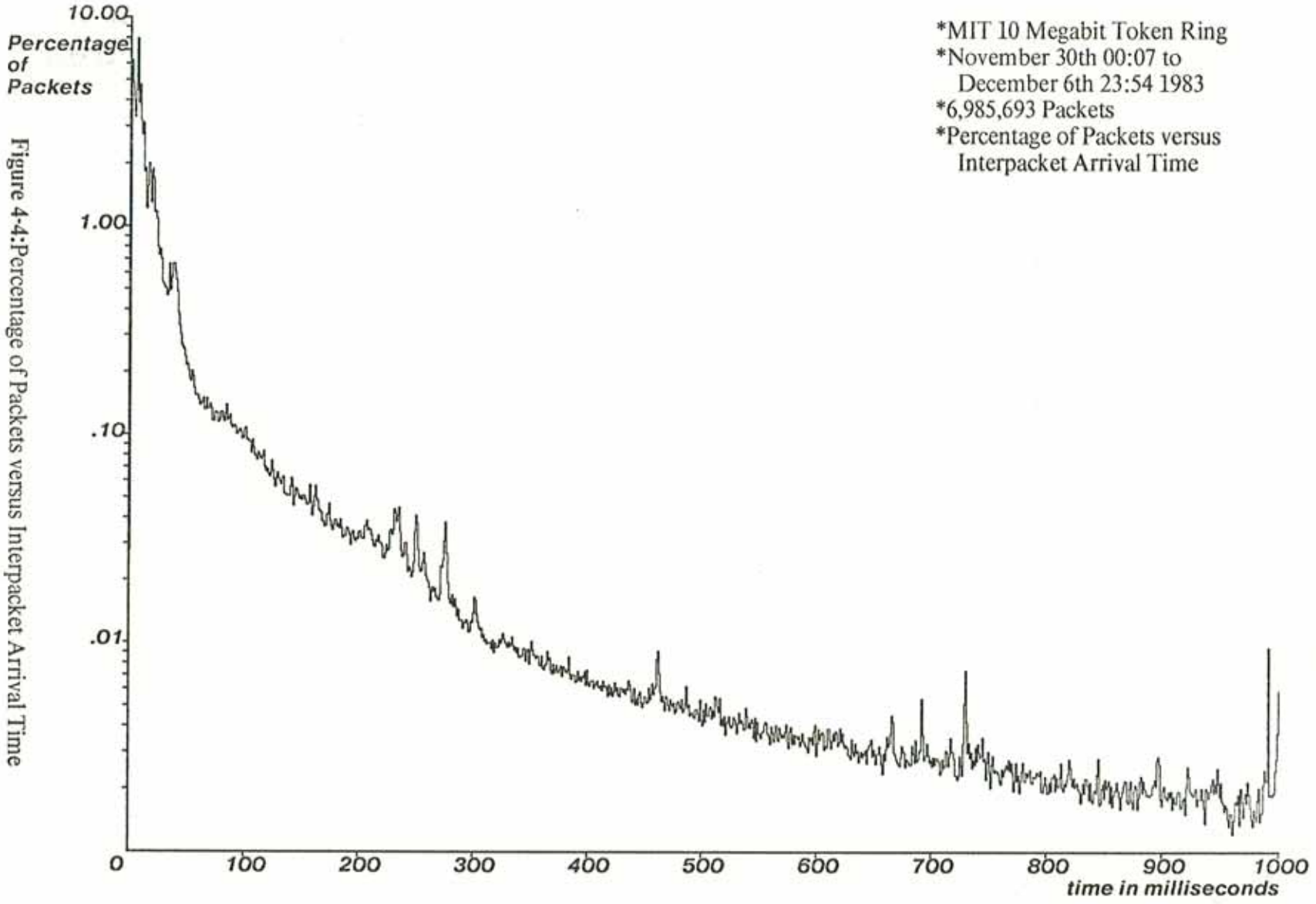


Figure 4-4:Percentage of Packets versus Interpacket Arrival Time

A histogram of the first 200 milliseconds plotted on a linear scale is shown in figure 4-5. The spikes near the beginning of the histogram are more distinct.

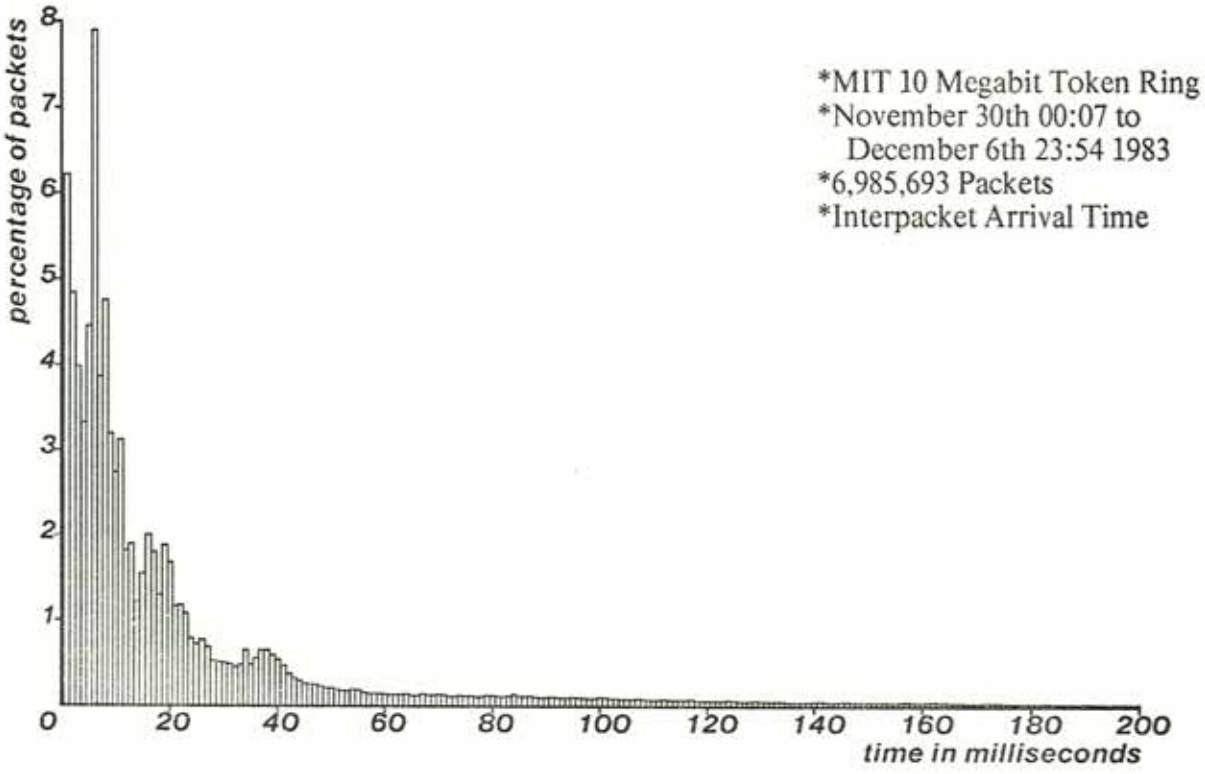


Figure 4-5: Percentage of Packets versus Interpacket Arrival Time

4.6 Protocol Usage

The protocol for the ring is the Internet Protocol, which is broken into several divisions. The Transmission Control Protocol (TCP) is a host-to-host protocol for reliable communication in internet environments. Almost 59% of the packets on the ring were TCP packets that are used for remote login and some file transfer. Most TCP packets are small, so TCP accounts for only 25% of the bytes on the ring.

Nearly 25% of the packets on the ring were Remote Virtual Disk (RVD) packets that are

used to request and deliver disk blocks of data from servers on the network. Most RVD packets are long, so RVD accounts for 58% of the bytes.

The User Datagram Protocol (UDP) is a user level protocol for transaction oriented applications. The major uses for UDP are name-request packets from the name server and file transfer, including mail. UDP packets accounted for about 15% of the packets and the bytes on the ring.

Internet Protocol	Number of of Packets	Percentage of Packets	Number of of Bytes	Percentage of Bytes
TCP	4,130,045	58.69%	336,484,225	25.30%
RVD	1,742,709	24.76%	773,863,891	58.19%
UDP	925,844	13.16%	207,365,219	15.59%
ICMP	235,690	3.35%	10,244,901	0.77%
others	3,219	0.04%	1,945,879	0.15%

Table 4-3: Percentage of Protocol Usage

4.7 Intranet, Internet, and Transit Packets

One might expect that most of the traffic on a local area network would be between two hosts on the same network, but this was not found to be the case with the ring. Packets fall into three categories: *intranet* (both source and destination on the ring), *internet* (source or destination, but not both on the ring) and *transit* packets (neither destination nor source on the ring). Internet traffic accounted for 49% of the traffic on the ring, with intranet comprising only 46% of the traffic. Transit packets accounted for about 5% of the traffic on the ring. The large quantity of internet traffic can be accounted for by remote login packets and mail packets.

Internet Protocol	Intranet Packets	Internet Packets	Transit Packets	Total Packets
TCP	33.93%	61.08%	4.99%	4,131,052
RVD	99.76%	0.16%	0.08%	1,742,586
UDP	15.26%	76.05%	8.69%	926,383
ICMP	0.36%	95.24%	4.40%	235,764
GGP	0.00%	100.00%	0.00%	8
total	46.15%	49.11%	4.74%	7,035,947

Table 4-4: Intranet, Internet, and Transit Traffic on the Ring

4.8 Network Reliability

Assessing the reliability of the ring is difficult. The monitoring station assumes the ring is down if there is no token, but even if the token is lost the ring will not reinitialize if no host has a packet to transmit. The major, if not only, cause of token loss is a host joining the ring. Joining the ring involves energizing a mechanical relay, which has a contact bounce time of a few milliseconds. This is more than enough to destroy any data on the ring. During the working day, some host joins or leaves the ring about every half hour.

The length of the ring changes with each host that enters the ring. The phase locked loops in each of the network interfaces must resynchronize on a common frequency, which takes about half a second. Since loss of control characters on the ring is noticed after 4 milliseconds, the ring loses and reacquires the token many times for a host entering the ring.

One might expect that the down times would be short (a few seconds or less) and that the up times for the ring would show a bimodal distribution - most of the up periods would be short, but most of the up time would be contained in a few long up periods. This is exactly what happened.

The measurement period was from January 5th 14:35 to January 7th 22:00 1984; the total monitoring time was 55:25. The monitoring station measures periods of no-token on the ring and places each no-token period in a slot. The time period covered by each slot increases by powers of two.

The no token time was between 23 and 46 minutes. Subtracting the maximum no token time from the total monitoring time gives a minimum token-present time of 54:39. The upper bound of ring down time is between 0.7% and 1.4% of its operating time.

The measurement of ring operation time is fuzzy both because of the resolution of the monitoring station statistics and because the token is reinitialized only when someone has data to send. If the token is lost but no station has data to transmit, the ring could be operational with no token for a long time. In the future, the monitoring station will actively send packets to find the up time more precisely by guaranteeing that a packet is sent soon after a token loss.

One quiet weekend the ring ran two and a half days without a token loss, allowing about 6 billion token circulations (and about 180 billion repetitions with 30 stations). This suggests that the error rate on the ring must be very low.

If the delay around the ring is 33 microseconds, the ring holds 330 bits at a nominal 10 Megahertz rate. The token is 10 bits long, and any single station sees a token every 33 microseconds or about 30,000 times per second. This is 300,000 bits per second that must be correct if the token remains intact. But there are 30 stations on the ring, so the number of bits passing all nodes is 9 million bits per second. This is 9 million bits per second correct over a 2.5 day period or 2×10^{12} correct bits. The node-to-node error rate is therefore less than 5×10^{-13}. The proNET manual states an observed error rate of better than 10^{-12}, so figures for the MIT ring are consistent.

The ring is shielded, twisted-pair cable averaging about 30 meters between nodes. The short distance between nodes is not near the limit of the proNET hardware, so the low error rate is expected. The low bit error rate suggests that nothing more complex than a parity bit is

required to detect errors on the ring; a Cyclic Redundancy Checksum (CRC) is not necessary because the network is so reliable. The network is reliable enough that it does not need separate error detection to improve performance. [7]

Chapter Five

Comparison of Results

This section compares the results presented in this thesis with other empirical measures of local networks. The best paper for comparison is the paper that was a model for this thesis: *Measured Performance of an Ethernet Local Network* by John Shoch and Jon Hupp. [8] The Ethernet is a 2.94 Megabit Carrier Sense Multiple Access with Collision Detection (CSMA/CD) bus network. Shoch and Hupp did their research at Xerox Palo Alto Research Center (PARC) where the Ethernet is 550 meters long and connects over 120 hosts, mostly personal workstations such as the Alto.

Robert V. Vieraitis Jr. presented a measurement of the version 1 ring done in this same lab in his Bachelor's thesis *A Performance Monitor for a Local Area Network.* [11] The version 1 ring was a 1 Megabit ring with 8 hosts (These hosts are now on the version 2 ring). The version 1 ring does not have relays to disconnect inactive nodes so network interfaces must always be powered on.

Apollo Computers, Inc. developed a ring network similar to the version 2 ring and a few measurements were presented in *The Architecture of an Integrated Local Network* by Paul Leach et al. [2] The Apollo ring is a 12 Megabit token ring that is star shaped with relays for maintainability as is the version 2 ring. The network at Apollo Computer has 124 nodes, mostly personal workstations.

5.1 Network Traffic versus Time of Day

The monitoring station accumulated packets on the ring over 10 minute periods. Packets were accumulated over 6 minute periods for both the Ethernet and the Apollo ring. A comparison among networks is possible because both 6 and 10 minutes are long enough that short term events cause little variation, but both periods are short enough that enough detail

exists to show how network traffic changes throughout the day. Packets were measured on the version 2 ring and the Apollo ring; network load is measured on the Ethernet. Network load is the number of data bits on a network divided by the capability of the network.

Xerox PARC and Apollo Computer are both different environments from the MIT Laboratory for Computer Science. Xerox and Apollo are businesses and tend to run on a 9:00 to 5:00 schedule that can be seen in their network traffic in figures 5-1 and 5-3. Notice that the Ethernet shows spikes in network load at noon and 6:00 PM, caused by people saving files before lunch and before leaving in the evening.

The Laboratory for Computer Science is in an academic environment. Although most staff and faculty work during the normal business day, there are many students who have classes during the business day. Students do computer work later in the evening, which accounts for the network load not dropping until 3:00 AM. See figure 5-2.

The Apollo ring had a peak over a 6 minute period of 230 packets per second, while the MIT ring with a quarter as many stations averaged 25 packets per second over 10 minutes. The Apollo ring had more than twice as many packets per node than did the ring because of the different computing environments. The Apollo system has nodes using a "shared" memory in which data can be on any node in the network, so many disk pages are retrieved from a remote disk rather than a local disk.

5.2 Network Utilization

Utilization of all four networks (Version 2 ring, Version 1 ring, Ethernet and Apollo ring) was a small fraction of network capacity. All but the Apollo ring have comparable utilization statistics. Comparing the utilizations directly is difficult because the speed and the hosts on the three networks are so different. A reasonable method of comparison is to find the number of bytes/second per node on the networks. Bytes per second on a network can be approximated by multiplying the speed of the network by its fraction of utilization and dividing by eight to convert from bits to bytes. Dividing this figure by the number of nodes on a network gives bytes/second per node. See table 5-1. Version 2 ring hosts

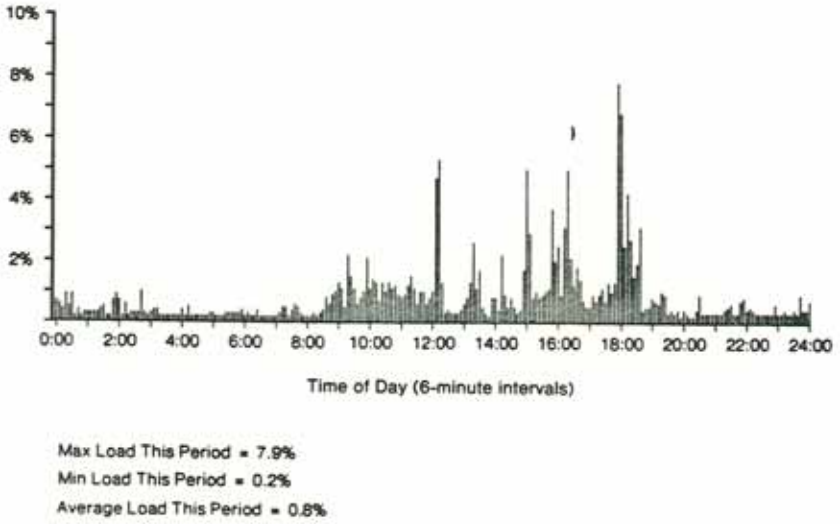


Figure 5-1: Ethernet Load versus Time of Day

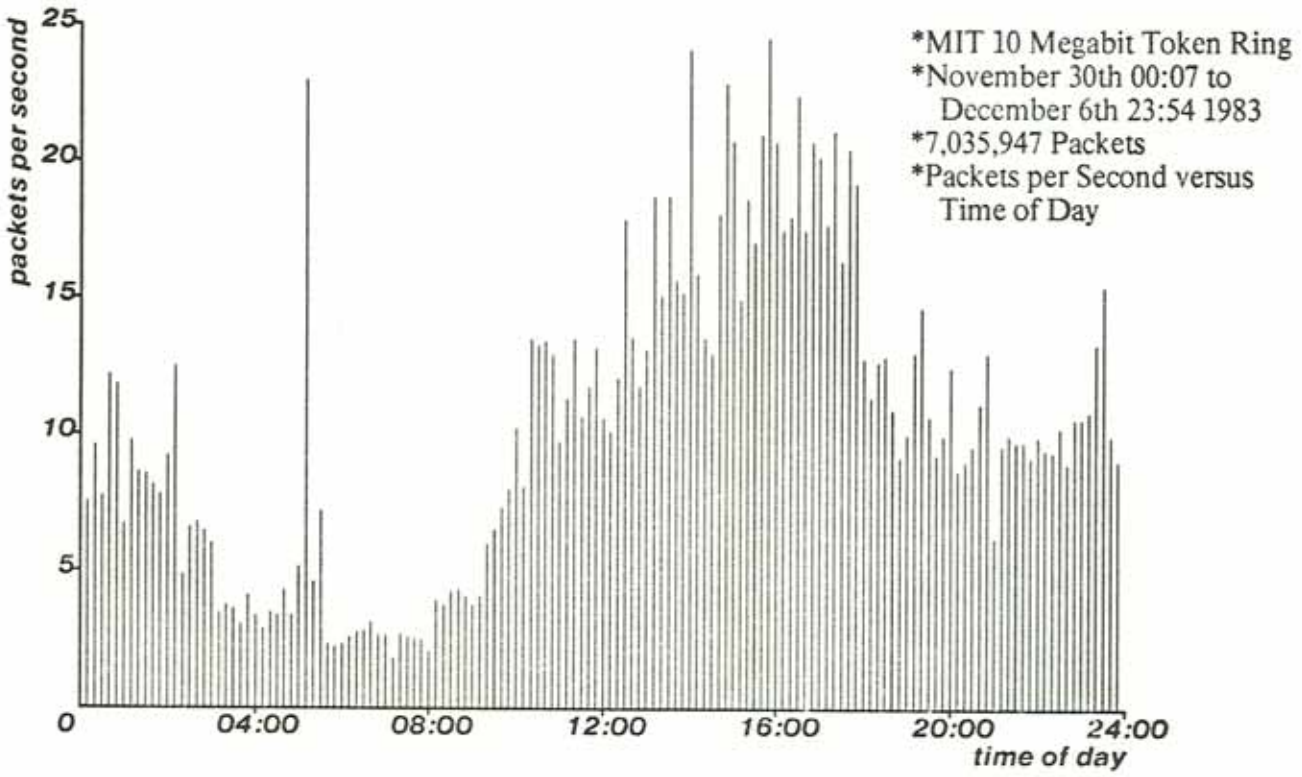


Figure 5-2: Version 2 Ring Packets versus Time of Day

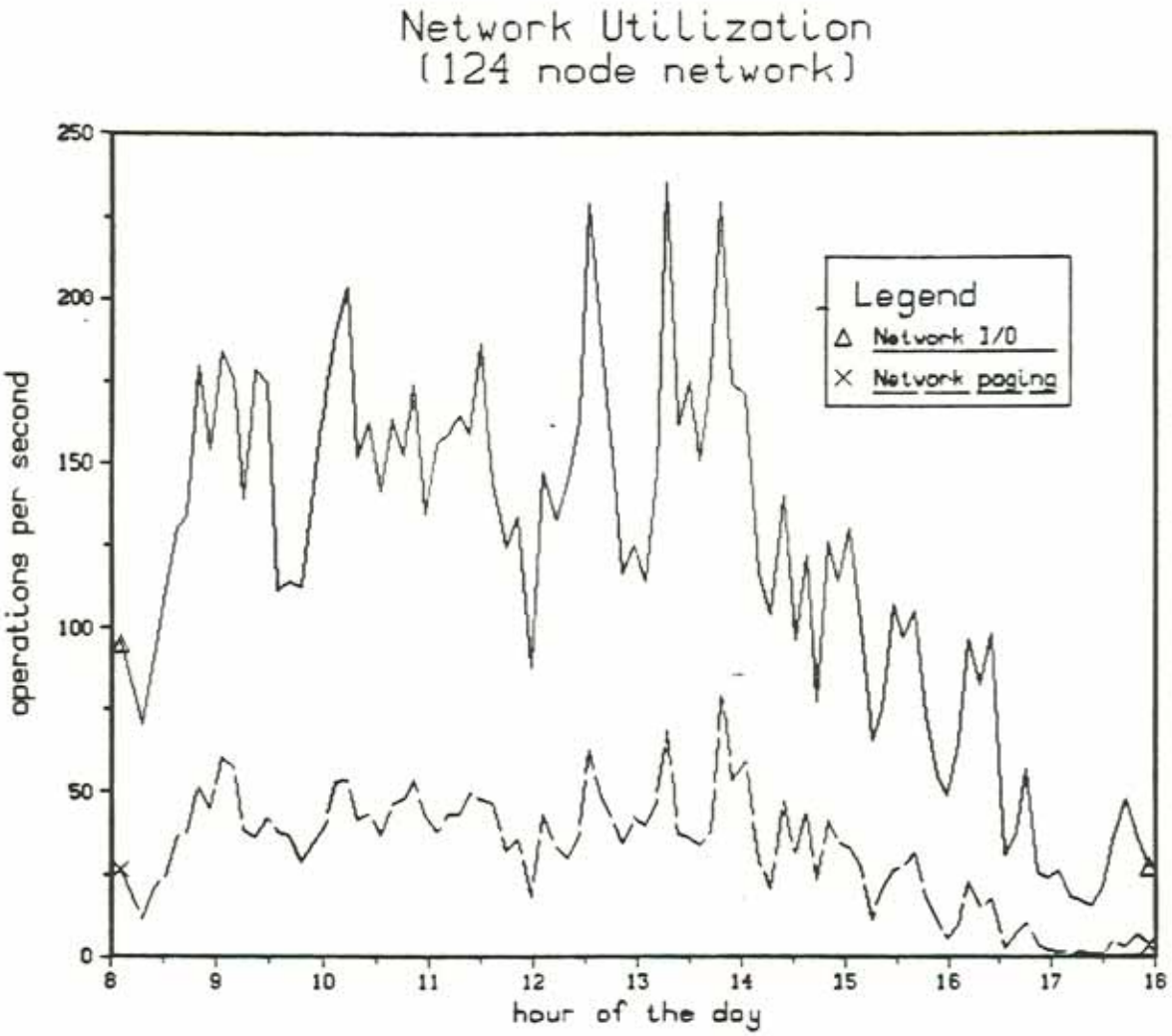


Figure 5-3: Apollo Ring Packets versus Time of Day

generated 4 times the load per node of Ethernet hosts, suggesting that the network applications are quite different in the two environments. The goal of the work at Xerox is office automation, the work at the Laboratory for Computer Science is in distributed computing.

The version 2 ring bytes/node per second for a one second period seems too high. This suspiciously large number could be because of the difficulty of measuring the peak load over a second on a ring that was referred to in section 4.2.

Measurement Period	Bytes/node V2 Ring	Bytes/node Ethernet	Bytes/node V1 Ring
Second	27,445	1133	1228
Minute	2312	521	491
Hour	562	110	22
Day	108	26	5

Table 5-1: Bytes/Second per Node for Three Networks

5.3 Packet Length Distribution

Packet length distribution on the ring was bimodal - packets were either short (under 100 bytes) or long (between 530 and 576 bytes). Shoch and Hupp reported similar findings in a paper on the Ethernet [9], as did Vieraitis in his paper on the version 1 ring. All show that most packets are small, but most bytes are sent in large packets. See figures 5-4 through 5-9.

The paper *Experience with Measuring Performance of Local Network Communications* by Terry and Andler makes the observation that packets are either small (under 128 bytes) or large (a 2K disk block). [10] On the ring, most large packets were 570 bytes - one disk block plus a header.

5.4 Source - Destination Traffic Patterns

On the ring, servers transmit 48% and receive 46% of the packets; on the Ethernet servers transmit 69% and receive 71% of the packets. The greater than 20% difference in both cases is really only a difference of interpretation. Servers on the ring refer to the Remote Virtual Disk servers and the gateways. Servers on the Ethernet refer to these types of servers and two timesharing machines. Personal workstations comprise most of the Ethernet hosts, and they communicate with the timesharing machines, not with each other. The ring is almost

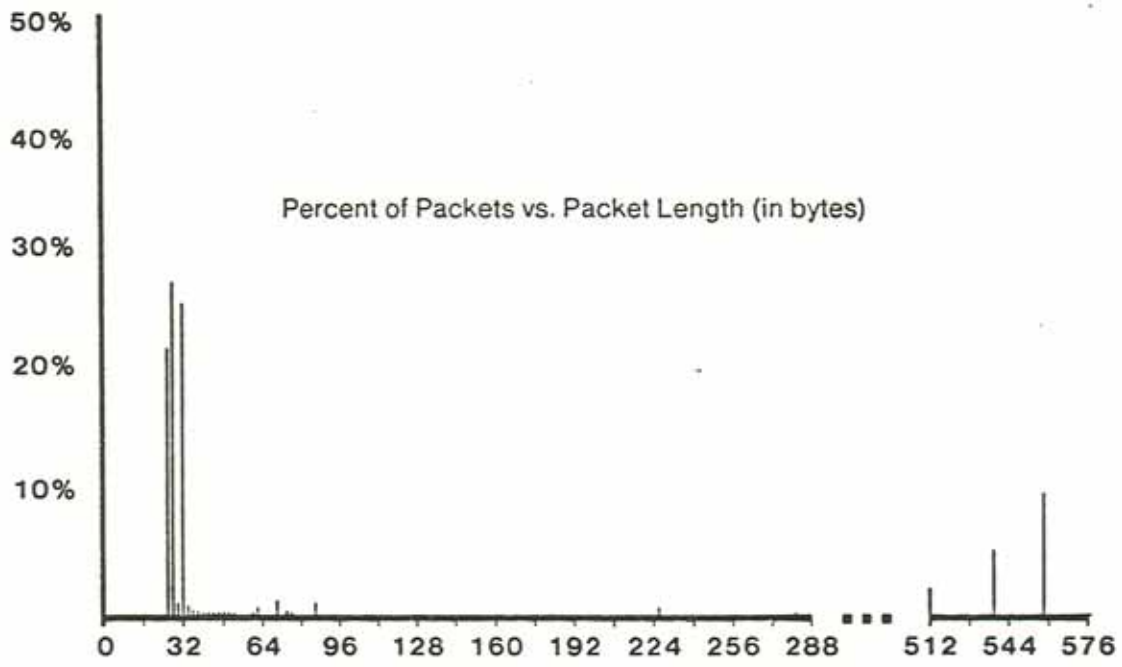


Figure 5-4:Ethernet Percentage of Packets versus Packet Length

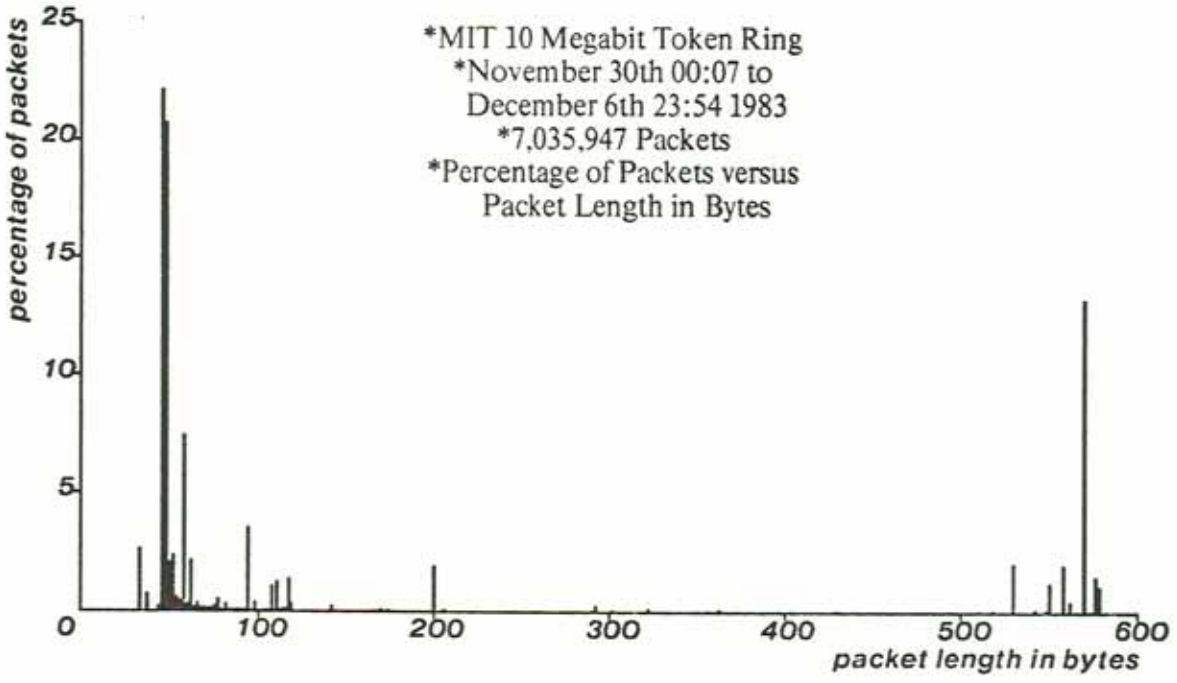


Figure 5-5:Version 2 Ring Percentage of Packets versus Packet Length

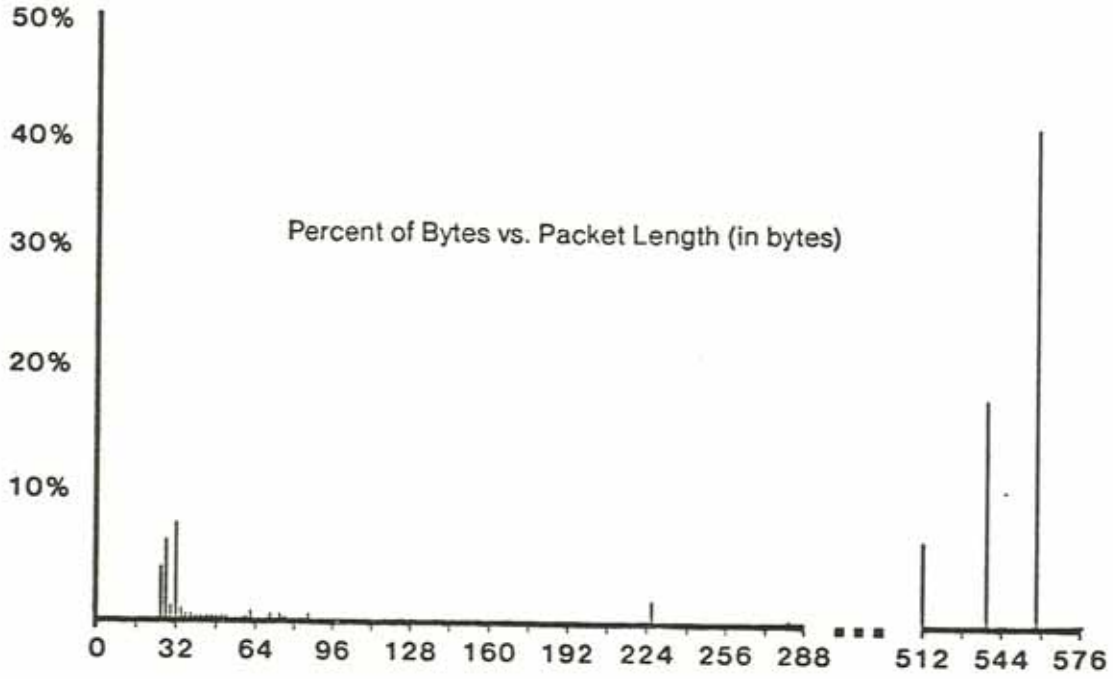


Figure 5-6: Ethernet Percentage of Bytes versus Packet Length

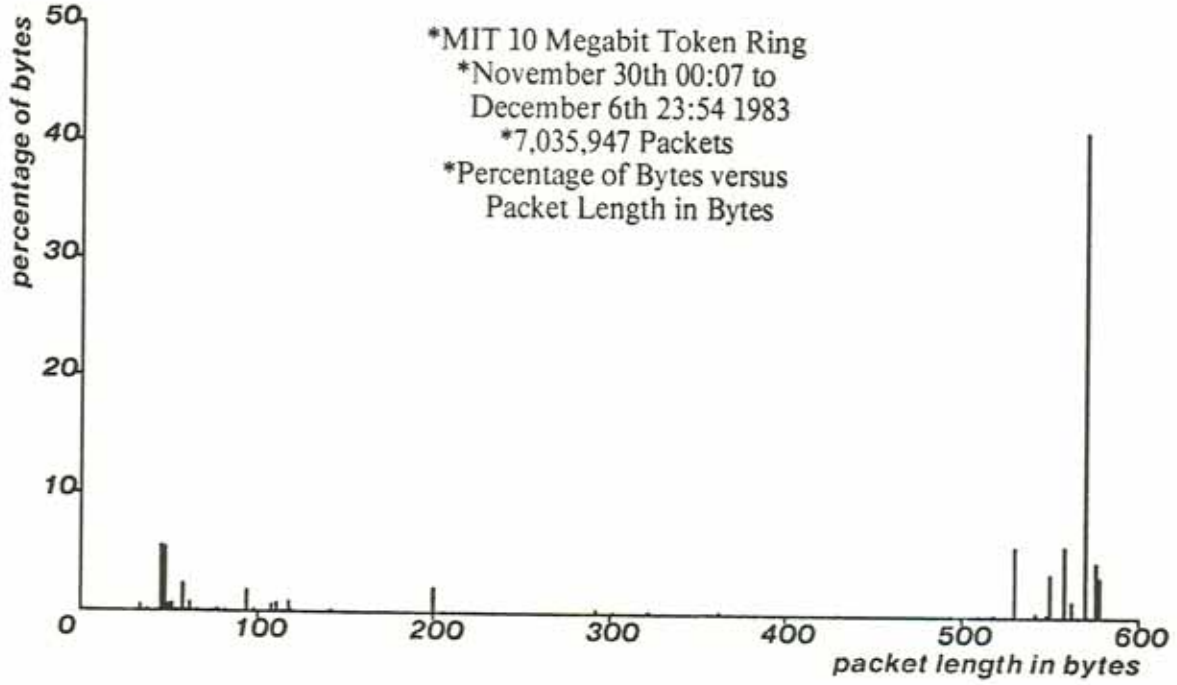


Figure 5-7: Version 2 Ring Percentage of Bytes versus Packet Length

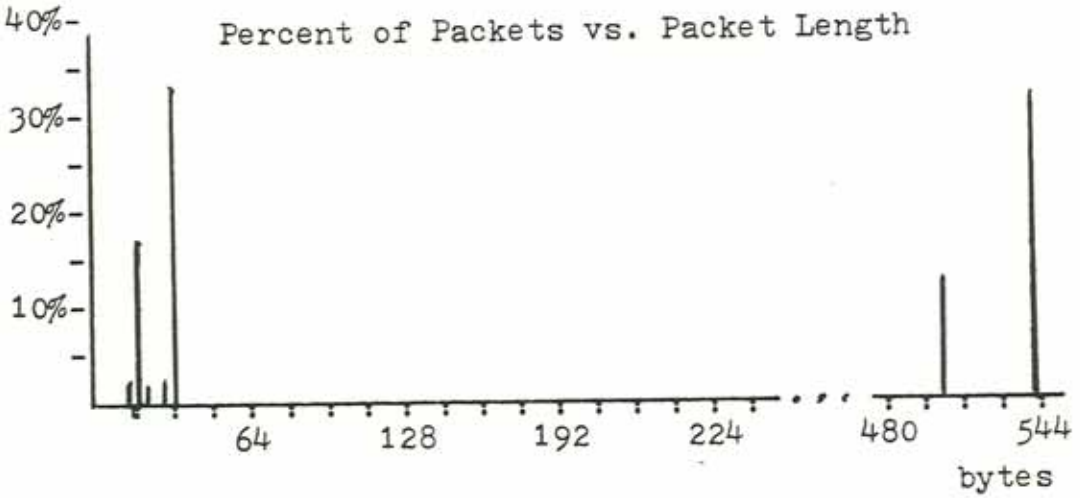


Figure 5-8: Version 1 Ring Percentage of Packets versus Packet Length

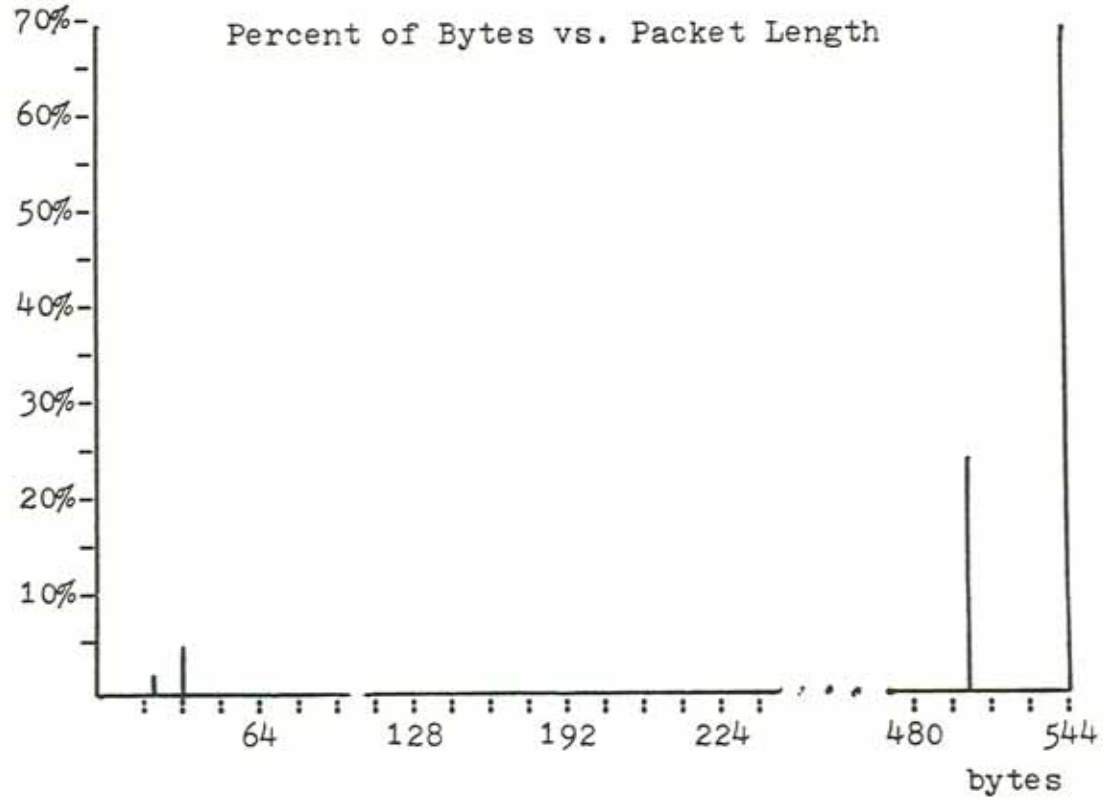


Figure 5-9: Version 1 Ring Percentage of Bytes versus Packet Length

entirely timesharing machines that communicate with one another and are not considered servers. Ring hosts are more autonomous, which means more duplication of data and less need of servers. The difference in server traffic is thus caused by both the different types of hosts and the definition of server on the two networks.

5.5 Interpacket Arrival Time

The version 2 ring interpacket arrival time histogram is presented in figure 5-11 for comparison with that obtained by Shoch and Hupp in figure 5-10. This histogram is linear on both axes and covers only 200 milliseconds; it is similar to the histogram obtained on the Ethernet.

Since the ring is 3.4 times faster than the Ethernet, it is to be expected that a significant number of packets would arrive within less than a millisecond of one another. Even full size (576 byte) packets take less than half a millisecond to transmit on the ring.

The interesting feature in figure 5-11 is the dip at 4 milliseconds and the spike at 6 milliseconds. The reason for both of these is unknown.

5.6 Intranet, Internet, and Transit Packets

The ring has 49% internet traffic, more than the 28% on the Ethernet. The Ethernet has almost no transit traffic, but the ring has 4% transit traffic. The ring is connected to the ARPANET, 2 Ethernets, and it has a PC gateway that transmits packets over serial lines. Differences in the network environment account for the larger number of internet packets on the ring.

Many computers at the Laboratory for Computer Science connect to the ring, but many of the smaller computers (such Altos and IBM PCs) that are used for remote login are on the two Ethernets in the building. This explains why many of the remote login packets (TCP) are internet packets.

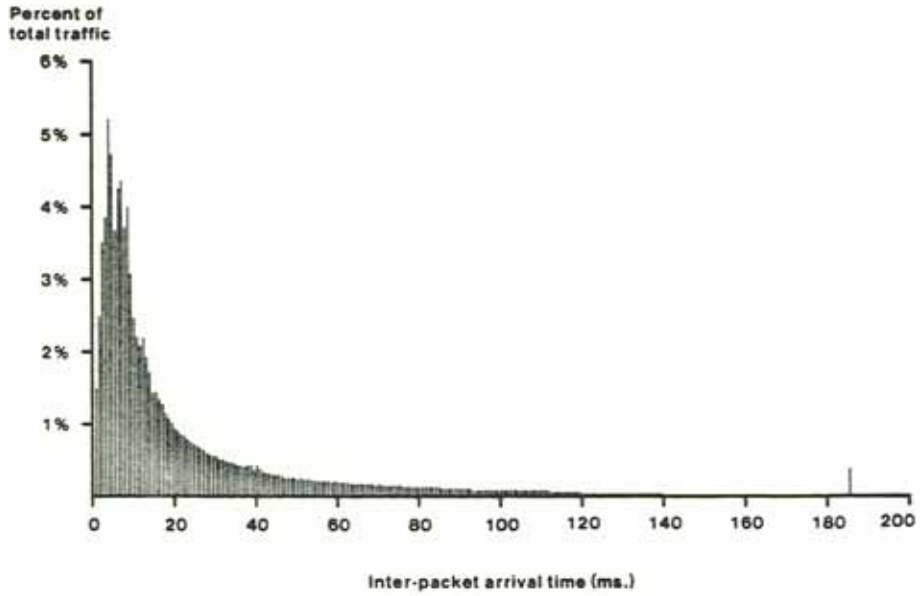


Figure 5-10: Ethernet Histogram of Interpacket Arrival Time

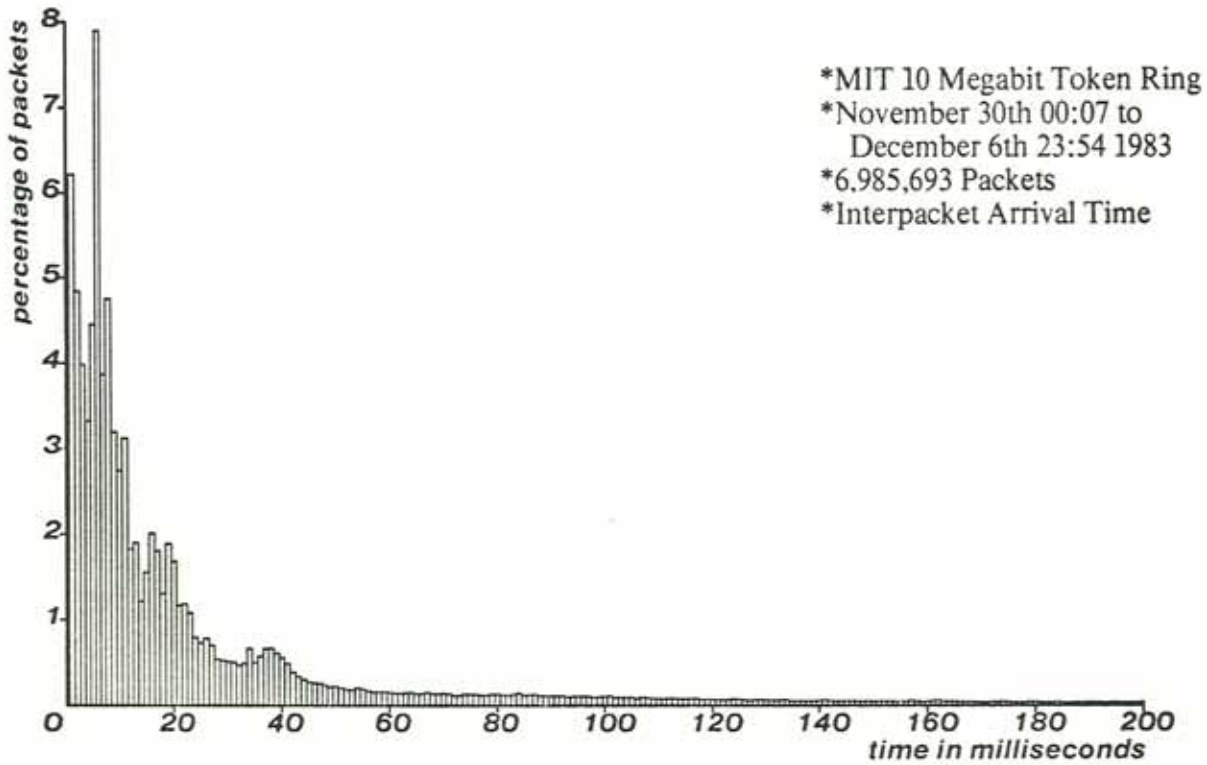


Figure 5-11: Version 2 Ring Histogram of Interpacket Arrival Time

Chapter Six

Conclusion

This chapter is a summary of the data collected by the monitoring station and the results of the project.

6.1 Network Traffic

Traffic on a local area network is bursty by nature. Although utilization over a day is low - less than 1%, high bandwidth is necessary to handle the bursts of traffic at high speed that occur during file transfers, for example. This confirms observations seen on the Ethernet.

Version 2 ring hosts generate four times the load per node of Ethernet hosts, suggesting different network application in the two environments. Despite the higher load per host on the ring, network utilization was a small fraction of network capacity. This observation confirms that the ring performs at a low utilization level as was intended in its design.

Data collected on the version 2 ring, the Ethernet, the version 1 ring and by Terry and Andler shows that packets are either large (a disk block) or small (less than 128 bytes). Network interface design, buffer allocation in networked computers, and protocols could be improved by taking advantage of this fact.

Internet traffic on the ring accounts for 49% of the packets. Experience at the Laboratory for Computer Science is that networks tend to get awkward as they grow large. As local area networks become more common, the trend in the future will be to have several small local area networks connected together by gateways or repeaters.

Interpacket arrival time is close to Poisson and could be modeled that way. This behavior was seen on both the version 2 ring and the Ethernet. Peaks in the interpacket arrival time histogram correspond to turn-around time for some network hosts.

An upper bound on ring down time is the time that the ring had no circulating token - 1.4% of the running time. Because of the initialization strategy used on the ring, lack of a token does not necessarily indicate a non-functional ring.

The probability of a bit error on the ring is estimated at less than 10^{-12} based on a two and a half day period with no token loss.

All of the above show that a token ring is a usable architecture for a local area network.

6.2 Network Monitoring

Many of the results in this thesis could not have even been guessed without a network monitoring system, such as the large quantity of internet traffic on the ring.

The monitoring station is used to check whether proper packets are being transmitted by experimental software. The network monitoring station is the first diagnostic tool used for checking operation of the ring whenever things go wrong. The monitoring station has great potential for network debugging, software optimization, and maintenance.

Chapter Seven

Suggestions for Future Work

This chapter discusses future directions that measurement of local area networks can take.

7.1 Continuation of Monitoring at LCS

The network monitoring system has been operating for two months. This allows a snapshot of the network, but gives little indication how traffic will change with time. Continued monitoring the ring for a length of time would give more complete data. A project that should happen soon is the storage of some typical days of network traffic on magnetic tape for future analysis, which would be like snapshots in a photo album of the ring in operation.

Rajendra K. Jain, a visiting scientist from Digital Equipment Corporation, will be studying the ring with the current monitoring equipment.

Improvements could be made to the monitoring station. A hardwired line could replace the ring as a means of moving data from the monitoring station to the analysis machine. The advantage is that the monitoring station would not contribute to the traffic on the ring, which would reduce the monitoring artifact (statistics indicate that a modest speed serial line of 64 kbits/sec would work).

The reliability figures for the ring are fuzzy because the monitoring station measures the time that no token exists on the ring. Because of the ring's initialization strategy, even an operational ring may not have a circulating token. The monitoring station could transmit a packet after a token loss, restoring the token at the earliest possible time to get a better estimate of ring availability.

7.2 Monitoring Other Locations

MIT is not the only location with a proNET ring. Other universities include Berkeley, UCLA, and the University of Wisconsin at Madison. Commercial firms include Realshare and Sohio. Monitoring should be done at some of these locations to determine whether the MIT proNET is a typical installation and whether the type of load is different at non-university sites.

7.3 Monitoring of Other Token Rings

Almost no other literature exists about measurements done on rings. Some preliminary work was done on the Apollo ring, but little information about the ring itself was discussed. The Apollo ring is similar in design to the proNET, so data about the Apollo ring would be good for comparison.

The IBM experimental ring is another token ring that should be monitored. The IBM experimental ring, though similar in concept to the version 2 ring, has several special design features intended to contribute to performance.

7.4 More Detailed Analysis of Token Resource Allocation System

This thesis does little to compare the Ethernet CSMA/CD bus and the version 2 token ring as strategies for design of local area networks. It would be interesting to see more study of token passing versus CSMA/CD allocation.

References

[1]

Bux, W., et al.
A Local-Area Communication Network Based on a Reliable Token-Ring System.
In *Proc. IFIP TC. 6 International In-Depth Symposium on Local Computer Networks*, pages 69-82. (Florence, Italy, April, 1982).

[2]

Leach, P.J., et al.
The Architecture of an Integrated Local Network.
Accepted for publication in *IEEE Journal on Selected Areas in Communications*, November, 1983.

[3]

Operation and Maintenance Manual for the proNET model p1000 Unibus Local Network System.
9th edition, Proteon, Inc., 1983.

[4]

Saltzer, J. H. and Pogran, K. T.
A Star-Shaped Ring Network with High Maintainability.
In *Proc. MITRE-NBS Local Area Communication Network Symposium*, pages 179-189. (Boston, Mass., May, 1979).
Reprinted in *Computer Networks* 5(4):239-244, (Oct./Nov., 1980).

[5]

Saltzer, J. H.; Pogran, K. T.; and Clark D. D.
Why a Ring?
In *Proc. IEEE Seventh Data Communication Symposium*, pages 211-217. (Mexico City, Mexico, October, 1981).

[6]

Saltzer, J. H.
Communication Ring Initialization Without Central Control.
Technical Memorandum TM-202, MIT Laboratory for Computer Science, (December, 1981).

[7]

Saltzer, J. H.; Reed, D. P.; and Clark D. D.
End-to-End Arguments in System Design.
In *Proc. 2nd International Conference on Distributed Computing Systems*, pages 509-512. (Paris, France, April, 1981).

[8]

Shoch, John F. and Hupp, Jon A.
Measured Performance of an Ethernet Local Network.
Communications of the ACM 23(12):711-721, (December, 1980).

[9]

Shoch, John F. and Hupp, Jon A.
Performance of the Ethernet Local Area Network--A Preliminary Report.
In *Proc. IEEE COMPCON80*, pages 318-322. February, 1980).

[10]

Terry, D. and Andler, S.
Experience with Measuring Performance of Local Network Communications.
In *Digest of Papers Twenty-Sixth IEEE Computer Society International Conference*, pages 203-205. (San Francisco, Ca., April, 1983).

[11]

Vieraitis, Robert V. Jr.
A Performance Monitor for a Local Area Network.
Undergraduate Thesis, Electrical Engineering and Computer Science Dept., Massachusetts Institute of Technology, May, 1980.

Appendix A

Monitoring Station Operation

A.1 Monitoring Station Hardware

Two different computers comprise the monitoring system. The *monitoring station* is a special purpose computer that does selective data collection and compression with some real-time analysis. The monitoring station consists of a Digital PDP-11/10 computer with 32k words of MOS semiconductor memory, a DL-11 asynchronous line interface, a proNET ring interface, two proNET Unibus Host Specific Boards, and some specially constructed hardware. The monitoring station runs a real-time display on a VT-52 terminal, which is useful for checking the current state of the ring. Compressed data is sent via the network to the *analysis machine*, a VAX 11/750 computer running 4.1 Berkeley Unix. The analysis machine receives the data from the monitoring station in compressed form for storage of network data and long term analysis of network parameters.

A.2 proNET Network Hardware

The network hardware consists of two parts. The first part is a generalized interface between the network and a fifty-wire ribbon-cable interface. This *Ring Control Board* or CTL has the modem and the low level receive and transmit mechanisms. The CTL performs the necessary lower level functions for ring management. The CTL recognizes three different control characters on the network. These control characters are the *token* (free to transmit), *connector* (Beginning of Message or BOM), and End of Message (EOM). The first byte after the connector is the packet's destination address. If the destination address matches the address of the CTL or the CTL is in *match-all* mode, the CTL begins to receive the message. The CTL takes bits serially from the network and transfers the data to the Host Specific Board (HSB) one byte at a time across the fifty-wire ribbon-cable . As the HSB receives these bytes of data, it stores them in an on-board packet buffer with 2046

bytes of storage. When the HSB receives the message complete signal, it begins a DMA to the memory of the host (16 bit word transfers into the PDP-11/10). When the DMA is complete, the HSB interrupts the processor. This interrupt allows the host to re-enable the HSB to receive another packet.

Between the time that a packet is received and the time that the HSB is re-enabled, the network interface is deaf to the net and does not receive any packets. For a network monitoring station to be effective, it should miss as few packets as possible. One method of missing few packets is to reduce the re-enable time of the network interface by receiving only the data that interests you. For the Internet Protocol (used exclusively on the ring), the first seventeen bytes of the packet contain the relevant information. These first seventeen bytes include ring destination address, ring source address, the ring protocol and the particular internet protocol. Since the network interface receives only the first seventeen bytes of a packet, it has the receiving time for the rest of the packet to recover for the next packet.

A second method to miss few packets is to have two network interfaces. One interface receives a packet from the ring while the other resets for the next packet. The original monitoring station used the first method, but the hardware did not work well enough. The version of the monitoring station used in the measurements reported here effectively had two network interfaces.

A.3 Special Hardware

The special hardware for the monitoring station switches the incoming bytes of data from the CTL card to one of two HSB cards in the PDP 11/10. The CTL card fifty wire ribbon cable runs to the special hardware. The special hardware then determines if the first HSB is ready to receive a packet. If so, then the Monitoring Station Hardware (MSH) allows the first HSB to receive data from the network. If the HSB card is unready, then the second HSB card receives the network data. A counter in the MSH increments if neither card is ready to receive. The MSH switches the receiving HSB by selecting four of the fifty wires in

the ribbon cable, those wires that are control lines for the HSB. Ground, data lines, and transmit control lines are connected to both HSB cards at all times.

A 10 Megahertz clock on the MSH provides a timebase. Dividing the clock down provides a 25 microsecond resolution clock. A 32 bit counter run by this clock provides all timestamps. The counter output is accessible over the computer bus and it also runs to latches that are triggered by various events. A loss of token on the ring triggers a latch and the acquisition of token triggers a second latch. A third latch triggers when the first HSB receives a packet and the fourth latch is triggered when the second HSB receives a packet. The outputs of these latches are also accessible from the bus.

Packets on the ring may be any length because they are delimited by connectors (BOM) and EOM control characters. No field in the ring header contains the length of the packet, so the length must be determined by counting bytes. During normal operation, packet length is determined by checking the DMA counter on the HSB to see how many bytes were transferred into memory. To speed up the monitoring station, only the first 17 bytes are DMAed into memory. Because no other packet length counter is built into the HSB, counters in the MSH determine the number of bytes received by each HSB for an arriving packet.

The last component of the MSH is a vectored interrupt system. The current time is stored in latches whenever the ring acquires or loses a token, and an interrupt is triggered. This allows the PDP 11/10 to queue the times of ring disruptions for transmission to the analysis machine.

Other hardware on the monitoring station includes provisions for monitoring signals from the CTL. Signals on the CTL run from the CTL through buffers to BNC sockets on the front panel of the PDP 11/10. The signals observable are token, flag (the basic unit of BOM, EOM or token), network clock (10 MHz), raw clock for differential Manchester encoding (20 MHz), digitally recovered input data, and digitally encoded output data. A frequency counter mounted on the monitoring station monitors the system clock or the raw clock. The proNET ring network has a distributed clock algorithm in which the PLLs on all

the modems must adjust their frequency to a mutually agreeable value. In effect, the bit time adjusts so that the ring holds an integral number of bits. The operating frequency of the ring changes with the number of nodes, the length of wire, and the particular nodes that are in operation. The frequency counter also measures the circulation period of the token around the ring. The token completes a ring circulation without delay if no packet is transmitted, which is common with a low network load. This is the minimum possible time of token circulation and is the amount of delay around the ring. Each node on the ring has ten bit times of delay of a hundred nanoseconds per bit, therefore the delay in microseconds around the ring is equal to the number of nodes in the ring plus the propagation delay through the wire. The wires in the ring at the MIT add about 15% to the station delays.

A.4 Monitoring Station Software

The monitoring station software was written with two thoughts in mind. The first was fast servicing of interrupts. The interrupt routines written in assembly language take about 30 microseconds to execute. The rest of the software had to be efficient, but for convenience of programming it had to be a higher level language. The C programming language was chosen primarily on the basis of availability. The interrupt routines placed data in circular buffers and the main software would analyze the data in the circular buffers as a first priority. The system priorities are: 1) retrieval of network data (interrupt level), 2) preprocessing of network data, 3) transmission of pre-processed data, 4) incrementing the clock, 5) display of data, 6) character fetch from the keyboard.

A.5 Real-time Analysis

The Network Monitoring Station has a VT-52 display for real time network analysis. The screen has four different windows: a help display, a packets display, a netload display, and an error display. The help window displays the various window options available as in figure 1. The packets window displays the number of packets seen over various time intervals. Along the bottom of the screen are four clocks displaying current time, time since last ring reinitialization, starting time and running time. See figure 2. The netload display is

similar, except that the percentage of total ring resources in use replaces the number of packets observed. The netload display is illustrated in figure 3. The error display is mostly for checking on the number of errors that the monitoring station has made. It displays the number of packets missed by the monitoring station because of slow interrupts, the number of bad format packets that have occurred, the number of token losses on the ring, and the number of transmission errors on packets transmitted to the analysis machine. See figure 4.

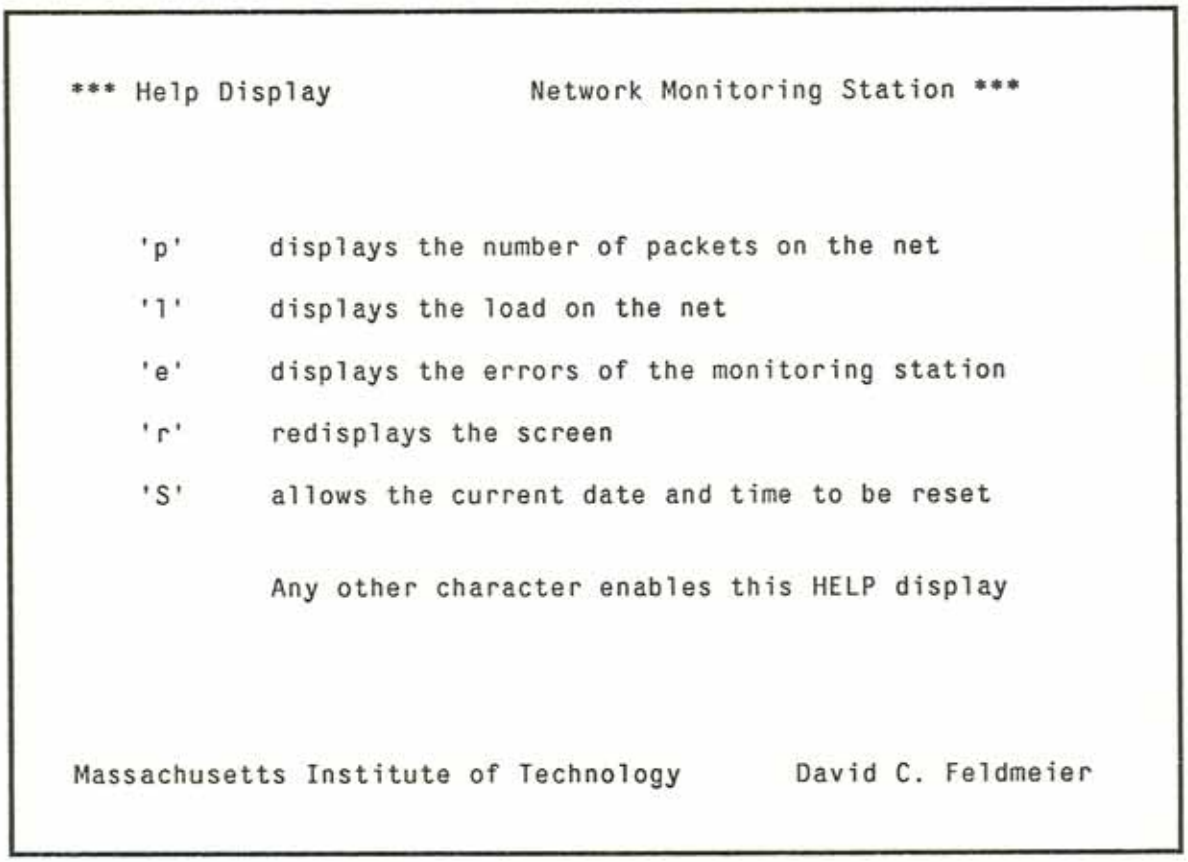
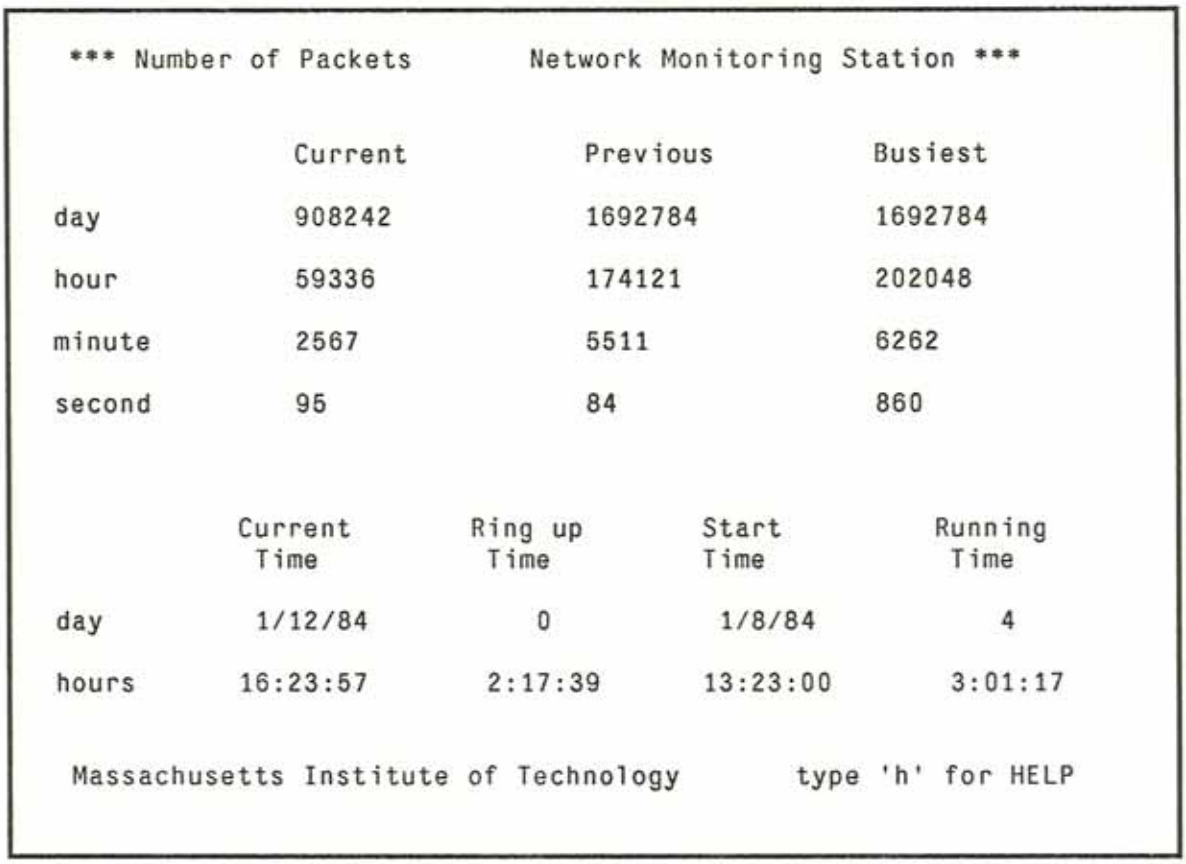


Figure 1:HELP window on the monitoring station

A.6 Transmission of Information to Analysis Machine

The monitoring station does data compression of network traffic and sends this compressed data to the analysis machine for further processing. The monitoring station eliminates



*** Number of Packets Network Monitoring Station ***

	Current	Previous	Busiest
day	908242	1692784	1692784
hour	59336	174121	202048
minute	2567	5511	6262
second	95	84	860

	Current Time	Ring up Time	Start Time	Running Time
day	1/12/84	0	1/8/84	4
hours	16:23:57	2:17:39	13:23:00	3:01:17

Massachusetts Institute of Technology type 'h' for HELP

Figure 2: Typical *packets display* on the monitoring station

faulty packets (those that have improper format or are less than the minimum Internet Protocol packet size) and compresses the remainder into large packets that are sent to the analysis machine. These packets are composed of a local network header, an Internet Protocol header, a User Datagram Protocol header, a clock field (six bytes) and 67 packet fields (each field is eight bytes) for a total of 576 bytes.

The header fields are simply the standard headers used on the ring. The clock field is used for several purposes. One byte of the clock field is the sequence number of the packet. A second byte contains six bits of time (the current minute) and two bits that specify what the remaining four bytes contain. The four possibilities are: current time of the 32 bit clock in the monitoring station, the 32 bit time that the token was lost, time that the token was

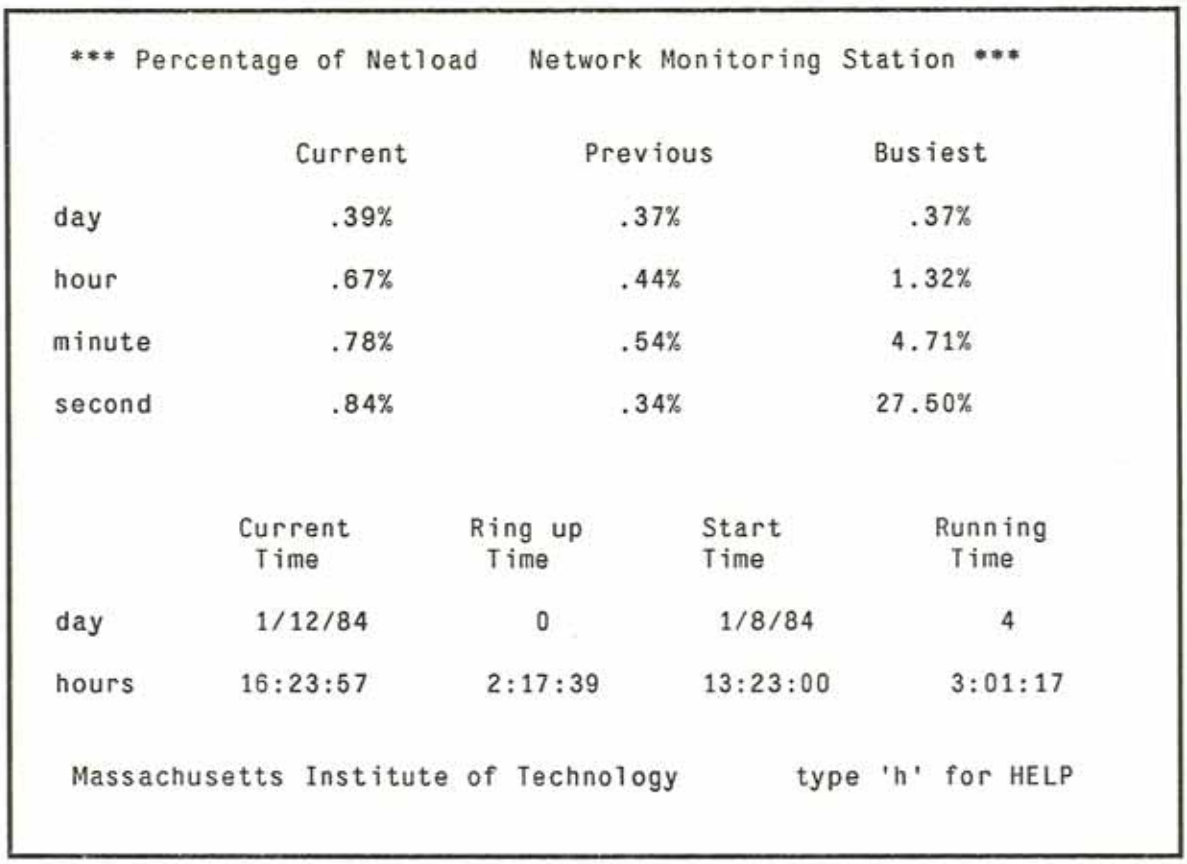


Figure 3: Typical *percentage display* on the monitoring station

acquired, or time that the ring glitched (token loss and reacquisition too quickly to measure). All token acquisition/loss data is sent via the clock field. The monitoring station buffers this data (which tends to occur in bursts) and transmits it as quickly as possible. Each packet field contains eight bytes: ring destination address, ring source address, internet protocol, two bytes of packet length, and three bytes of arrival time to the nearest 25 microseconds.

Compressed data packet transmission to the analysis machine is not reliable. To guarantee that network load caused by the monitoring station remains low, the monitoring station follows a try-at-most-once transmission strategy which also eliminates the need for acknowledgment packets from the analysis machine. If packets lost on the way to the

```
*** Error Display          Network Monitoring Station ***

    Number of Packets Missed by Both HSBs is          6212
    Number of Packets Received by Second HSB is       15483
    Number of Bad Format Packets Received is          176482
    Number of Packets Smaller than IP Header is       5741

            Receive Status HSB #1 is          107
            Receive Status HSB #2 is          107

    Transmit Status for HSB #1 is   200
            errors  878     refused 421     bad format 471
            timeout 0       overrun 0       nxm 0

    Number of Ring Initializations  9333
    Number of Token Losses          9334
    Instant Token Loss/Ring Initialization  2


Massachusetts Institute of Technology       type 'h' for HELP
```

Figure 4: Typical *error display* on the monitoring station

analysis machine are at random, no biasing of data will occur. Transmission of the clock field is more reliable, but not guaranteed. If the monitoring station realizes that the packet to the analysis machine was lost (bits in the refused packet indicate that the packet was not received or the packet was damaged), the clock field data is retransmitted in the next compressed data packet. The clock field can still be lost if the network software on the analysis machine runs out of network received packet buffers, but this loss occurs rarely.

A compressed data packet is sent to the analysis machine for every 67 packets (not including the compressed data packet) monitored on the ring. Because the monitoring station transmits at a rate proportional to the traffic on the ring, compressed data packets cause the peak traffic load to be 7% higher than it really is. The monitoring station defers packet transmission under high load conditions, but the packet still must be sent.

The monitoring station would produce less monitoring artifact if it did not use the network to report statistics. An alternative method of data transmission is a hardwired line to the analysis machine. Packets sent by the monitoring station account for 7% of the bytes on the ring. If the monitoring station can buffer information for a minute, then the line to the analysis machine needs to carry 7% of the data bits on the ring. If the busiest minute is 10% of the ring capacity, then 10 Megabit x 10% x 7% = 70 kbit/sec over the hardwired line. A hardwire line may be run in the future to improve the quality of statistics from the network monitoring system.

Appendix B

Long-term Analysis - the Analysis Machine

The program for long term analysis runs on a Digital VAX 11/750 timesharing computer with 4.1 Berkeley Unix. The specific VAX for the analysis machine was chosen because of its low network traffic load, enabling it to receive the most packets from the monitoring station. The analysis machine receives data from the network monitoring station over the ring network. Packets sent by the monitoring station account for about 7% of the bytes on the net. The compressed data packets received by the analysis machine are processed immediately. Long term storage is impractical, because on a busy day over 15 megabytes of storage would be necessary.

B.1 Data Analysis

The analysis machine analyzes the packets from the monitoring station as they are received. When the analysis program is first started, it looks for files of network data, creating any that do not exist. In its errorfile, it enters the time and date at which it was started. It also logs the date and time of the first packet received from the monitoring station. Every 10 minutes, the analysis program writes out errors into an errorfile and writes into a file the number of compressed data packets received over the last 10 minutes.

The analysis program sleeps until a packet arrives. Packets are transferred from the network buffers into a circular queue. If packets are in the queue and no packets in the netbuffers, processing begins. The sequence number of each packet is checked; duplicates are discarded and missing sequence numbers are noted. Each compressed data packet has its clock field checked for a token loss or token acquisition. The time of the last acquisition or loss is stored so the ring up time or the ring down time can be computed. The result is entered into a log table.

Interpacket arrival time is computed using the 25 microsecond resolution timestamp in each packet data section. The difference between the time of this packet and the last packet is computed and the correct bucket is incremented in a 1 millisecond resolution, 2000 slot histogram. The first packet after a ring recovery or the first packet in a compressed data packet that has no predecessor is considered invalid and not included in the interpacket arrival time histogram. The compressed data packet is also not in the histogram.

Packets are then separated by protocol. Six of the eight protocols recognized are the Internet Protocol (IP) subdivisions: Exterior Gateway Protocol (EGP), Gateway to Gateway Protocol (GGP), Internet Control Message Protocol (ICMP), Remote Virtual Disk (RVD), Transmission Control Protocol (TCP), and User Datagram Protocol (UDP). The other two divisions are all other IP protocol subdivisions and all non-IP protocols.

For each protocol, there is an 256x256 host-table that contains the number of packets between every source-destination pair on the ring. The number of packets of each length from 26 bytes (the length of the IP protocol) to 2046 bytes (the hardware limit of the ring) are stored for each protocol.

B.2 Display Programs

The analysis program does as little as possible to process incoming packets to increase its speed. The computationally intensive work is done by the display programs. Display programs worked with data placed in tables by the analysis program.

A host activity program worked with the 256x256x8 table of source, destination and protocol. The user specified a host and protocol and the program displayed all those hosts that communicated with the given host. The number of packets sent in each direction were displayed for the given protocol. Another version of this program went through all the host and ordered them by total number of packets transmitted and total number of packets received. A third version of this program found the number of packets to a gateway to a ring host, from a gateway to a ring host, and from gateway to gateway. Other programs simply turned the tables into histograms and formatted the data for a plotting program.

Tables in the thesis were computed using various display programs.