

# Network Management and Performance Monitoring at SLAC\*

C.A. Logg and R.L.A. Cottrell

Stanford Linear Accelerator Center  
Stanford University  
Stanford, CA 94309

CAL@SLAC.Stanford.Edu  
COTTRELL@SLAC.Stanford.Edu

*Light  
All Thru  
ord  
Run  
4-13-95*

"The physical network plant and everything attached to it, including the software running on "computers" and other peripheral devices is the **system**."

## ABSTRACT

Subjectively, the ultimate measurers of **system** performance are the users and their perceptions of the performance of their networked applications. The performance of a **system** is affected by the physical network plant (routers, bridges, hubs, etc.) as well as by every "computer" and peripheral device that is attached to it, and the software running on the computers and devices. Performance monitoring of a network must therefore include computer systems and services monitoring and well as monitoring of the physical network plant. This paper will describe how this challenge has been tackled at SLAC, and how, via the World Wide Web<sup>1,2</sup>, this information is made available for quick perusal by concerned personnel and users.

## INTRODUCTION

The ultimate measure of **system** performance is the perception of the users. The users' perception is generally

based on the **response time** to various commands they enter at their keyboards. This maybe any command from a print command, the submission of a large compute intensive job, an interactive graphics session, an editing session, a directory list command, the sending and receiving of mail, literally any and all commands.

Response time can be affected by anything on the network, from physical network errors, network bandwidth availability, the reliability of software performing services, disk space (availability and configuration), cpu availability, the performance of network interfaces; that is, any activity on or the configuration of anything attached to the **system**.

Two UNIX<sup>R</sup> computers are used to perform network management and performance monitoring at SLAC. They are a Sun<sup>R</sup> (SunOS 4.1<sup>1</sup>) and an IBM<sup>R</sup> RS/6000 (AIX<sup>R</sup>) running Netview/6000<sup>R</sup>. An older system, a DECstation<sup>R</sup> 5000 running ULTRIX and DEC's discontinued MSU, is used for the Help Desk network map display. Software components used include: Perl<sup>3</sup>, REXX<sup>4</sup>, C, SAS<sup>R,5</sup>, the Tricklet SNMP routines<sup>6</sup>, and ORACLE<sup>R</sup>. Note that all these tools are not necessary. At a bare minimum, one

\* Work supported by Department of Energy contract DE-AC03-76SF00515

<sup>R</sup> indicates a registered trademark

needs a computer, SNMP<sup>7</sup> utilities, a programming language, and a graphing package.

The World Wide Web (W3 or WWW) is the interface used to make available the network management and network performance monitoring reports<sup>8</sup> at SLAC.

The WWW is a pool of network-accessible human knowledge. It is an initiative started at CERN, now with many participants. It has a body of software, and a set of protocols and conventions. WWW uses hypertext and multimedia techniques to make the WWW easy for anyone to roam, browse, and contribute to.

## OVERVIEW OF THE SLAC NETWORK

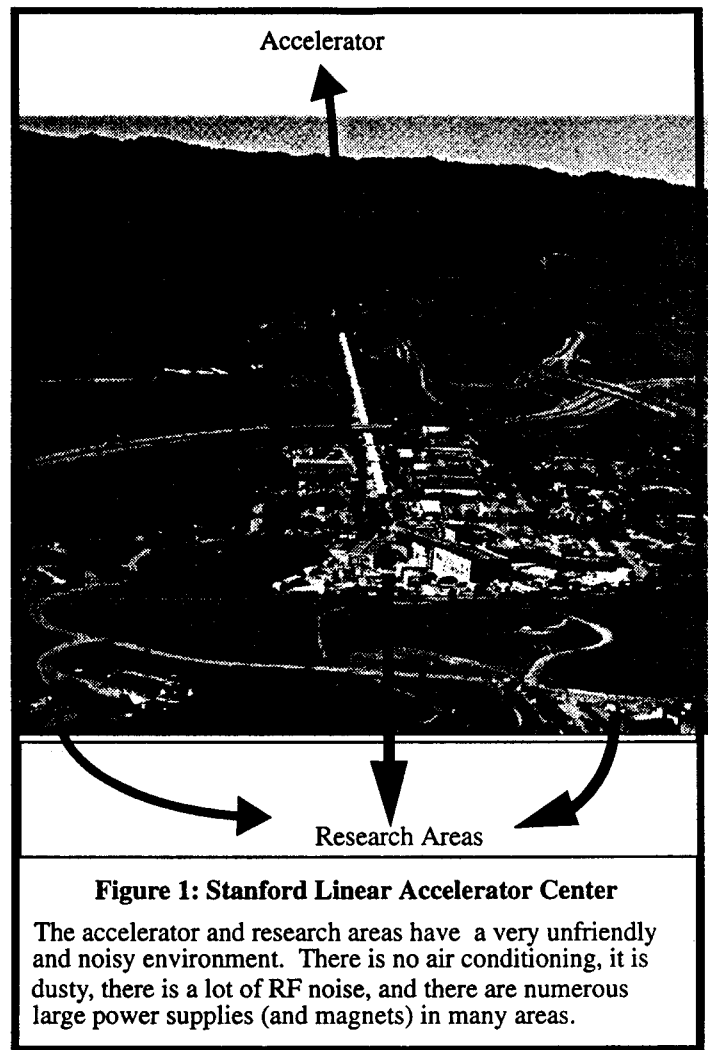
The Stanford Linear Accelerator Center (SLAC) is a national laboratory operated by Stanford University for the US Department of Energy. It is located on 426 acres of Stanford University land, about 40 miles south of San Francisco. SLAC has been in continuous use for over 25 years in a national research program that has made major contributions to the understanding of nature. The Center is one of a handful of laboratories worldwide that stands at the forefront of research into the basic constituents of matter and the forces that act between them. There are active programs in the development of accelerators and detectors for high energy physics research and of new sources and instrumentation for synchrotron radiation research.

The staff is currently about 1400, of whom 150 are Ph.D. physicists. At any given time, there are typically 900 physicists from other institutions participating in the high energy physics and synchrotron radiation programs.

The SLAC network is physically spread over all the SLAC campus (Figure 1).

It is dynamic and continually growing as new research facilities are added at SLAC. The topology of the SLAC network is changing rapidly to accommodate the growth, and new networking technologies are continually incorporated. Both ethernet and FDDI are extensively employed.

The SLAC network has a large variety of computer hardware and operating systems, including: IBM<sup>R</sup> systems running VM<sup>R</sup> and AIX<sup>R</sup>, DEC VAX and Alpha systems<sup>R</sup> running VMS<sup>R</sup> and Ultrix<sup>R</sup>, Suns<sup>R</sup>, HP/UX<sup>R</sup>, SGI<sup>R</sup>, several flavors of PCs including IBM<sup>R</sup> compatibles, MACs<sup>R</sup>, Amigas<sup>R</sup>, running several flavors of operating systems and network software such as DOS<sup>R</sup>, OS/2<sup>R</sup>, Appletalk<sup>R</sup>, Novell<sup>R</sup>, MacOS<sup>R</sup>, TCP/IP, NextStep<sup>R</sup>, and DECNET<sup>R</sup>.



This all makes for a very heterogeneous network and all the problems associated with it, including interoperability problems.

## NETWORK MANAGEMENT AND PERFORMANCE MONITORING

Network Management is generally broken down into five management areas<sup>8</sup>: configuration, fault, performance, security, and accounting management. Only configuration, fault and performance management will be discussed here.

### 1.0 Configuration Management

SLAC is a very open environment with users and their computing equipment coming and going every day. This makes it difficult, with the current organization and tools

to do complete configuration management and know all the time what is attached to the system and where it is.

An Oracle database called CANDO<sup>9</sup>, developed at SLAC, is used to store information on the network cable plant, networking devices both passive and active, and computers and peripherals attached to cable plant. The lengths of wires, tap locations, etc. are contained in CANDO. Details of networking devices such as hubs, routers, ethermeters, and bridges are stored in CANDO. Information concerning domain name service (DNS) is contained in CANDO. Computers and peripherals, which are known about, are in CANDO with their names, IP addresses, ethernet addresses, physical locations, contact people, and operating systems. However, only devices which are known about can be put in CANDO.

To aid in configuration management, router Address Resolution Protocol (ARP) caches and bridge learn tables are searched to look for "new" devices on a regular basis. The log from Netview/6000 is also processed nightly for "new nodes" which it may have discovered during the day. This information is then processed for entry into CANDO.

Numerous flat files are created from the information in CANDO daily (or more frequently) via cron invoked SQL programs. These files are automatically made available for use by the Network Management Stations (NMS) and other network management and monitoring utilities (such as an Remote Network Monitoring (RMON) probe Graphical User Interface (GUI), network monitoring data collection, analysis and data reduction programs), which will be described later. Several of the files are also made available by WWW for SLAC general access.

## 2.0 Fault Management

**System** faults most frequently manifest themselves as: slow response, loss of connectivity between nodes, inaccessible file systems, hung processes on computing nodes, and dropped sessions.

There are several types of **system** faults which may be responsible for the above symptoms:

### 2.1 Physical Network Faults

Physical network faults such as CRC, alignment errors, broken and lost packets have a large effect on response time over a network. When a packet is damaged, it cannot be processed by the node for which it was destined. The sending node detects that a packet has reached its destination properly by receiving an acknowledgement from the receiver. Failure to receive an acknowledgement to a

packet causes the sender to timeout (can be many seconds) before resending the packet. These timeout delays can result in long delays in response for the users' applications.

Many network devices such as network traffic monitors (referred to herein as Ethermeters<sup>R</sup>), hubs, routers, bridges, and switches have a "program" that runs in them called an "SNMP agent". The SNMP agent provides access to its Management Information Base (MIB), and communication with the agent is performed via "The Simple Network Management Protocol" (SNMP). Utilizing SNMP, configuration and performance data can be read out of the devices with agents and analyzed. SLAC currently collects data from Ethermeters, routers, and bridges on an hourly basis, analyses that data, reduces it, and presents it for viewing via the WWW<sup>10</sup>.

The data collection is controlled via cron jobs and lists of nodes extracted on a regular basis from CANDO. The data collection is performed via perl programs using public domain SNMP access packages.

#### 2.1.1 The Ethermeter Data Collection, Analysis, Reduction, and Presentation

Ethermeters (which reside on almost every individual segment at SLAC) are probed hourly for # of bytes seen, # of packets seen, # of crc errors, # of alignment errors, # of short packets, # of collisions, and the # of broadcast and multicast packets. Information on peak values for the previous hour for those MIB variables is also obtained. Other information including packet size distribution, the top-10 segment talkers, and protocol distribution information is obtained.

Once a day, in the early morning, an analysis of this collected data for the previous day, is done by a SAS program. The SAS program generates plots of the segment utilization information, error information, and packet size distributions for the previous day ("yesterday"), the previous fourteen days ("fortnightly"), and previous 28 days ("monthly"), and the previous six months. It also creates an extensive tabular report which summarizes the utilization and errors for the previous day.

This tabular report is then reduced to a report which shows a summary of only the data which may indicate that there are problems on network. This is done by applying thresholds to the data in the tabular report. For example, error rates over 1 in 10000 packets are highlighted. In addition, the report generated by the data reduction program is created in HyperText Markup Language (HTML) for display by the World Wide Web. Any values over the thresholds are linked via hypertext to a graph of the data and other information such as top-10 segment talkers and protocol

NCSA Mosaic: Document View

File Options Navigate Annotate Help

Document Title: [REDACTED]

Document URL: [REDACTED]

### Quick Guide to SLAC's Networking Reports and Data - Overview

SLAC Oct 1994

Clickhouse Summary Data Summary Report Problems

- For a QUICK REVIEW of yesterday's happenings
- ETHERMETERS
- BRIDGES
- ROUTERS
- FDDI RING ANALYSES - Under Development
- PROTOCOL ANALYSES
- PING ANALYSIS

Authors

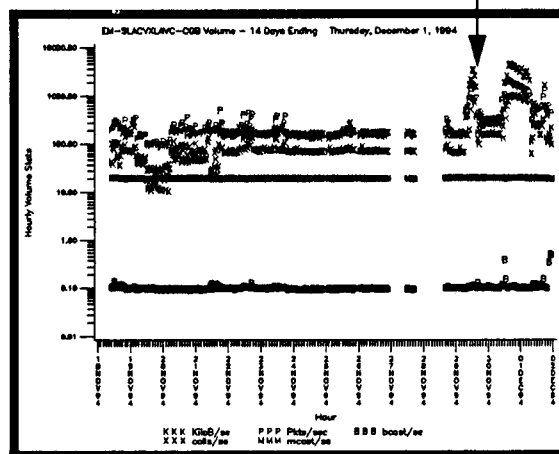
There are various thresholds applied to the daily data analysis and reduction for the extraction of "alertable" situations.

For a QUICK REVIEW of yesterday's happenings and data analysis and reduction see:

- The Ethermeter and Bridge Alert Summary for yesterday
- The Router Alert Summary for yesterday
- The Network Systems Summary Report for yesterday
- The Network Services Summary Report for yesterday and so far today
- It is often a good idea to search thru the ProbTrak database or review the last 40 ProbTrak entries to see what problems have been reported recently.
- The Change Log also provides a log of incidents and changes.
- A list of newnodes discovered by Netview/6000 yesterday is available as well as logs by month.
- One can also review the data for segments of the network by subnet.

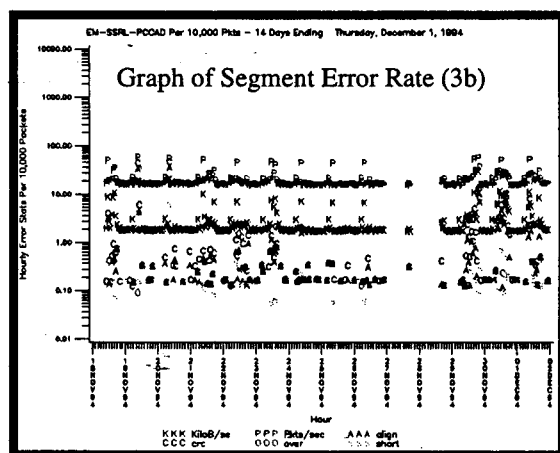
Back Forward Home Reload Open... Save As... Clone New Window Close Window

Note Increase in Packets and Kilobytes



Hypertext Link to Graph of Data Volume (3a)

Hypertext Link to Ethermeter and Bridge Summary Report (2)



NCSA Mosaic: Document View

File Options Navigate Annotate Help

Document Title: [REDACTED]

Document URL: [REDACTED]

### Ethermeter Alert Summary

(The full ethermeter daily summary is also available)

utilization & contention Analysis: Thursday, December 1, 1994

Node	util	100%sec/g	hr	util
EM-SLACVAVC-COB	22.38	9.47	18	22.38
EM-SLACVAVC-COB	22.38	17.35	1	22.38

Traffic Volume & Error Analysis Over the Entire Day: Thursday, December 1, 1994

Node	pkts	kiloB	sec	coll	short	err	align	over
EM-SLACVAVC-COB	828	449	0.2	19.8	150.41	0.0	0.0	0.0
EM-SLACVAVC-COB	21	3	0.8	13.4	0.0	0.4	1.4	0.0

Traffic Volume & Error Analysis for Peak Hours: Thursday, December 1, 1994

Node	hr>	pkts	hr>	kiloB	hr>	sec	hr>	coll	hr>	short	hr>	err	hr>	align	hr>	over
EM-SLACVAVC-COB	18	150.41	18	478	22	0.5	3	20.6	18	0.28	1	0.32	21	0.32	0	0.00
EM-SLACVAVC-COB	0	150.41	1	478	22	0.5	3	20.6	18	0.12	1	0.01	13	0.01	0	0.00
EM-SLACVAVC-COB	0	150.41	1	478	22	0.5	3	20.6	18	2.61	9	2.18	9	2.18	0	0.00

### Bridge Alert Summary

(The full bridge daily summary is also available)

Traffic Volume & Error Analysis Over the Entire Day: Thursday, December 1, 1994

Node	pkts	sec	coll	short	err	align	over
EM-SLACVAVC-COB	1	1	0	0	0	0	0
EM-SLACVAVC-COB	10	16	14	0	0	0	0
EM-SLACVAVC-COB	4	3	0	0	0	0	0

Traffic Volume & Error Analysis for Peak Hours: Thursday, December 1, 1994

Node	hr>	pkts	hr>	sec	hr>	coll	hr>	short	hr>	err	hr>	align	hr>	over
EM-SLACVAVC-COB	18	7	16	7.15	0	0	0	0.22	11	11	0	0	0	0
EM-SLACVAVC-COB	0	7	16	7.15	0	0	0	0.22	11	11	0	0	0	0
EM-SLACVAVC-COB	0	7	16	7.15	0	0	0	0.22	11	11	0	0	0	0

Back Forward Home Reload Open... Save As... Clone New Window Close Window

Top 10 Talkers (by hour) Table (3c)

Top 10 Talkers per hour for 94/12/01 00:00:00 to 94/12/01 23:59:59 for /usr/local/netdata/top10/data/EM-SLACVAVC-COB.docx																									
Numbers represent the percent of the total packets a pair represents for those hours that they were in the top10. Top 10 Total % is the percentage of total packets that were identified as being in the top10.																									
From node	To node	HOUR																							
		0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23
SLACCL	SLD01	2	2	3	2	3	2	3	2	3	2	3	2	3	2	3	2	3	2	3	2	3	2	3	2
SLACCL	SLD02																								
SLACCL	SLD03				2	7	7	7					3												
SLD01	SLACCL																								
UW0001	SLD03							3																	
UW0002	SLACCL																								
SC0	SLD01							3		2	3					4	9	6	4	5	2	2	2	2	2
SC0	SLD02											2	5	5	5	3	2	4	2	1	1	3	2	2	1
SC0	SLD03										2	3	6	11	4	9	2	9	5	1	3	2	2	2	1
SC0	SLD04							3	2	9	6	1	4	4	11	6	3	3	1	4	5	5	1	14	6
SC0	SLD05										4	2	13	3	3	2	3	3	3	2	1	3	3	2	2
SC0	SLD06																								
SC0	SLD02																								
SLACCL	SLD01																								
SLACCL	SLD02																								
SLACCL	SLD03																								
SLACCL	SLD04																								
SLACCL	SLD05																								
SLACCL	SLD06																								
SLD02	SLD07	10	10	11	15	10	9	10	16																

distribution, which may be useful in the investigation of a problem.

Figure 2, component (1) shows the Quick Guide to SLAC's Networking Reports and Data WWW page, its link to the Ethermeters and Bridge Alerts (Figure 2, component (2)) and some examples of some of the hypertext links for the reduced data to graphs and tables of relevant information. Figure 2 components (3a), (3b), and (3c) result from activating the hypertext links connected with alerts for utilization and errors. They are respectively, (3a) = utilization information, (3b) = error information, (3c) = top-10 talker information. Also available through other hypertext links, contained in Figure 2 component (2), but not demonstrated in figure 2, are protocol distribution information, a graph for all segments of (shorts+collisions)/good\_packets by hour, and a link to another HTML page (Figure 3) which provides access to all the reports and graphs for the subnet that the device is on.

Contents	
• Today's Reports	
• Yesterday's Reports	
• Last 14 Day Reports	
• Last 28 Day Reports	
• Last 180 Day Reports	
Today's Reports	
RT-PUB1-CH1	Today's Ethermeter Volume graphs
RT-PUB1-CH2	Today's Ethermeter Volume graphs
RT-PUB1-CH3	Today's Ethermeter Error graphs
RT-PUB1-CH4	Today's Ethermeter Error graphs
RT-PUB1-CH5	Today's Packet Size Distribution
RT-PUB1-CH6	Today's Packet Size Distribution
RT-PUB1-CH7	Today's Ethermeter Protocol data
RT-PUB1-CH8	Today's Ethermeter Protocol data
RT-PUB1-CH9	Today's Ethermeter Top 10 data
RT-PUB1-CH10	Today's Ethermeter Top 10 data
RT-PUB1-CH11	Today's NAT Bridge plots
RT-PUB1-CH12	Today's NAT Bridge plots
RT-CG-PUB1	Today's Router Volume graphs for physical router= ROUTER6.
RT-CG-PA10-PUB1	Today's Router Volume graphs for physical router= ROUTER6.
RT-CG-PUB1	Today's Router Error graphs for physical router= ROUTER6.
RT-CG-PA10-PUB1	Today's Router Error graphs for physical router= ROUTER6.
RT-CG-PUB1	Today's Router Miscellaneous graphs for physical router= ROUTER6.
RT-CG-PA10-PUB1	Today's Router Miscellaneous graphs for physical router= ROUTER6.
Yesterday Reports	
RT-PUB1-CH1	Yesterday's Ethermeter Volume graphs
RT-PUB1-CH2	Yesterday's Ethermeter Volume graphs
RT-PUB1-CH3	Yesterday's Ethermeter Error graphs
RT-PUB1-CH4	Yesterday's Ethermeter Error graphs
RT-PUB1-CH5	Yesterday's Packet Size Distribution
RT-PUB1-CH6	Yesterday's Packet Size Distribution

Figure 3: Index to All Graphs and Reports Related to the "PUB1" subnet

### 2.1.2 The Bridge Data Collection, Analysis, Reduction, and Presentation

SNMP accessible bridges on the SLAC network are probed once an hour for information on # of packets received and forwarded, # of buffer and controller overflows, # of multicast packets seen, # of packets blocked and passed, the number of collisions and the number of CRC and alignment errors seen by the bridge.

Just like the Ethermeter data, the bridge data is analyzed, graphed and summarized in a tabular report. The tabular report is then analyzed, and only those values exceeding certain thresholds are put into an HTML'ized report for access via WWW. The threshold for errors is 1 in 10000 packets.

Figure 2, component 2 shows the Bridge Alert Summary report, and one of its hypertext links to the relevant graph of the bridge interface data (Figure 2, component (4)).

### 2.1.3 The Router Data Collection, Analysis, Reduction, and Presentation

Once an hour data is extracted via SNMP from all the routers. Data collected for each interface includes: # of incoming and outgoing packets and octets, the # of incoming errors and the number of each type of error (CRCs, framing, runts, giants, ignored packets, unrecognizable protocol packets), the # of incoming and outgoing packets discarded, the number of incoming and outgoing queue drops, the # of collisions on output.

Just like the bridge and ethermeter data, the router data for each interface on each router is analyzed and reduced into an HTML'ized file for quick perusal via WWW. The thresholds for errors are the same, that is 1 error in 10000 packets.

An example of the Router Alert Summary page and its HTML links is not included, but it works and looks like the ethermeter and bridge summary reports.

### 2.1.4 Connectivity Monitoring

Connectivity monitoring of the **system** involves checking that everything is reachable from everywhere. This would be an onerous task at the present, so at this point only connectivity from the network data collection server and Netview/6000 to critical servers and physical network equipment is performed. This is done three different ways.

#### 2.1.4.1 Monitoring the NMS logs:

One way involves processing the Netview/6000 network event log file ("trapd.log") and noting when interfaces were discovered "up" or "down". This information only appears for nodes that are "managed" by Netview/6000. In general, only "critical" nodes and physical network equipment are managed. Critical nodes include major cpu servers, file servers, machines running network information services (NIS), mail, WWW, DNS, news service, print service, and network management and monitoring services. It should be noted that this method says nothing about the actual state of the critical nodes, it only indicates their physical connectivity to the Netview/6000.

In addition, any new nodes which were discovered by Netview/6000 appear in the log file. These are also extracted daily and put into a file. The list of new nodes which have been added to the network can be very useful in tracking down physical network problems which "suddenly" pop up.

#### 2.1.4.2 Probing for System Up Time:

This method is used primarily for issuing pages and alerts for critical systems. It was implemented several years ago, and is to be redesigned in the near future. An SNMP request requesting the system up time is issued every 5 minutes. If a node fails to respond, an attempt is made to ping it. Depending upon the system, an alert (pop-up x-window message) or page can be issued if either the node does not respond to the SNMP request, or if it doesn't respond to the ping. Ping and SNMP probe different layers of the system. Ping is at a lower level and may respond even when SNMP doesn't. Because this SNMP query requires the execution of code in the system being probed, it may reveal that internal system congestion is present in the node being probed when it times out.

#### 2.1.4.3 Pinging Critical Nodes and Network Equipment

SLAC critical nodes, router interfaces, bridges, ethermeters, and other network equipment with IP addresses are pinged periodically, although this is done primarily for performance monitoring purposes. In addition various collaborators' nodes are probed via ping. This process will be described in detail under Performance Monitoring. Figure 4 shows an example of how the daily ping data is displayed for connectivity purposes and response time.

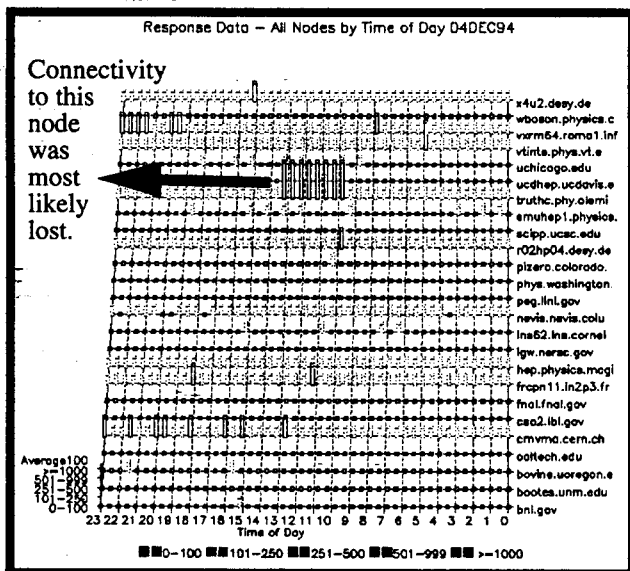


Figure 4: Ping Response Times for Off-site Nodes

## 2.2 Logical Network Faults

Logical network faults such as mis-routing and IP address conflicts (two nodes using the same IP address simultaneously) can cause poor response symptoms. If a packet is mis-routed, it may never get to its destination, or may be delayed. If two devices are using the same IP address, the packets received by the computers may confuse them and result in broken sessions and lengthy delays in response.

Comparison of ARP caches may lead to the detection of duplicate IP addresses if an IP address appears with two different ethernet addresses in the ARP caches. This detection is included in the processing of the ARP caches and an alert is generated when a duplicate IP address is detected.

Incorrect routing is often detected during the data collection process, and by examining the Netview 6000 log files for redirect messages issues by the routers.

## 2.3 Network Adapter Faults

Network adapter faults can bring the physical network down or cause it to be inoperative for intermittent periods of time. They certainly cause an interruption of service from the server experiencing them.

These can be difficult to detect. Currently a command is issued periodically to each computer system to read the system log. The network adapter and other network errors are then logged to a file which is made available by the Quick Guide entry "The Networked Systems Summary Report for Yesterday" (in Figure 2 component (1)). SeF-Figure 5 for an example of the WWW information. Note it also lists connectivity outages for the previous day.

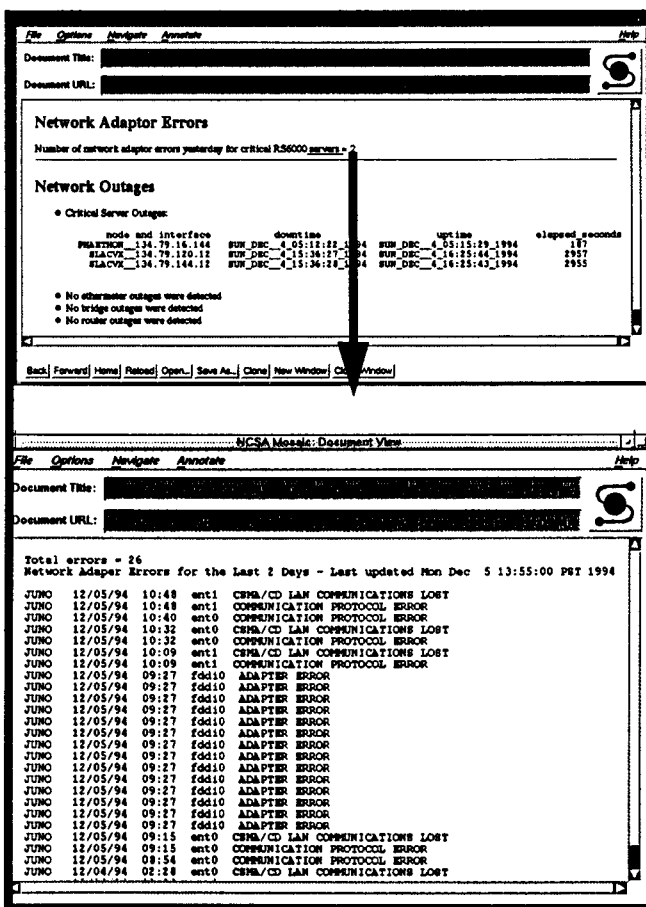


Figure 5: Network Adapter Errors and Network Outages

## 2.4 Computer Driver Faults

Computer driver faults such as producing incomplete/illegal packets on a network or not correctly following the arbitration protocol can cause serious problems by disrupting good packets.

The physical network fault monitoring often will indicate CRC and alignment errors which are actually the result of computer's network interface driver. When the alerts indicate a high level of errors on a segment, RMON is often used to capture the packets and analyze them. This can sometimes point to the computer interface issuing the defective packets.

## 2.5 Network Services Faults

Network services are provided by various computing nodes and programs running on them. If the printing daemon stops working, users cannot get their printed output. If the mail daemon stops working, mail may be lost, bounced, or not delivered in a timely fashion. These are discussed under Performance Management.

## 2.6 Computing Operating System Faults and Incorrect Configuration

Computer operating system faults and incorrect operating system configurations can result in the use of the wrong network interface, inefficient use of a network interface, and incorrect routing of packets.

The monitoring of physical network load often provides information in this area. Some segments may have unexpectedly heavy traffic for example. Upon examining the top-10 data, the segment traffic volume, and the protocol distributions, the offending nodes can usually be identified. The increase in traffic exhibited in Figure 2 component (3a) was the result of an FDDI interface not being used after a reconfiguration of its node. The traffic was all going over the ethernet whose Ethermeter graph is displayed.

## 3.0 Performance Management

Performance management entails looking at the overall performance of the **system's** components. In addition to looking at physical network response, this includes network services, as well as the performance of servers in providing disk and cpu resources. One of the characteristics of users' complaints is that they often think it is the physical network giving rise to problems of slow response. In reality, the situation is often much more complex. The network can be working just fine, physically, but if the servers are poorly configured or overloaded, the server performance itself may be bad.

### 3.0.1 Examining Physical Network Performance

The data collection and analysis described under Fault Management is also used for Performance Management. It is actually used to examine the physical performance and loading of the network cable plant itself.

The percentage of collisions also has a threshold applied to it, and if anything exceeds that threshold it is alerted. Collisions themselves are not bad, they are normal. But if there is a high level of collisions it is an indication of contention for a segment, and transactions can timeout before they are able to be placed on the net.

Other network overload conditions are indicated by the number of ignored, discarded packets (in and out), queue drops (in and out queues) in the case of devices such as routers or bridges. Any of these will result in a failed network transaction, and slowed response due to the need for a retry of the transaction.

### 3.0.2 Monitoring Lowest Level Network Communication Performance via Ping

The pinging of critical nodes and network equipment mentioned briefly in the Fault Management section provides extensive information on node to node network timing and inter-segment transmission of packets.

This data collection and analysis works as follows:

- A single ping is done to satisfy resolving the name.
- Next 5 pings with a 100 byte packet and 5 pings with a 1000 byte packet are executed. The minimum, maximum, and average round trip times are calculated, as well as the number of packets that did not make the roundtrip. This data is logged to a flat file.
- Once a day an analysis of the ping data for all nodes is done. Graphs of the frequency of response times (Figure 5) are made, and a report which lists packet loss is created. Any packet loss over 1% is alerted in a tabular report.

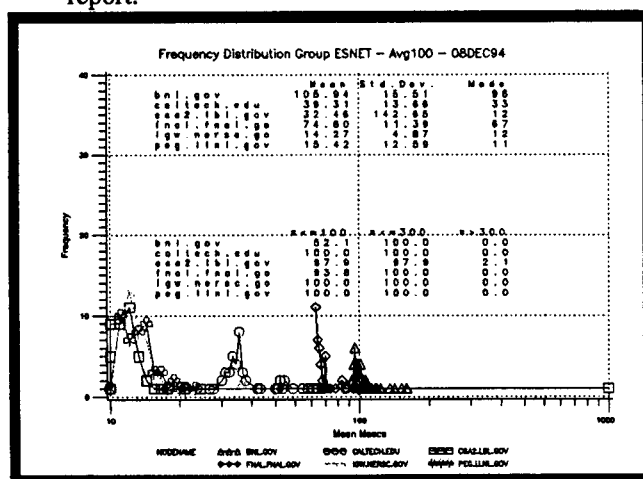


Figure 5: Frequency Distribution of Average Ping Response time for ESNET Nodes

In addition, several collaborators nodes (Figure 4) around the world are treated to this process. This has been helpful in tracking down WAN communications problems that were being blamed on the SLAC internal network.

### 3.0.3 Monitoring Network Services Performance

The performance of network services such as Network File Service (NFS), mail, nameservice, and WWW greatly affects overall response time of the system. We are currently in the process of developing mechanisms for the performance monitoring of these services.

The process entails timing the execution of a command such as a nameservice lookup (to each nameserver), probing a port for WWW or Simple Mail Transfer Protocol (SMTP), sending mail to various mail servers and timing how long it takes to get to its destination, an SNMP command, or accessing files on different hosts and timing how long those commands take.

Currently the analysis of the service timing data entails creating a tabular histogram with 4 bins. For example, SMTP service is probed and the response is histogrammed in 4 bins of <1.0 seconds, <5.0 seconds, less than 10.0 seconds, and greater than 10.0 seconds. In addition, mail is sent through various mail servers, and its delivery timed.

Figure 6 shows an example of the tables currently generated for monitoring network mail services.

Response Time Service Summary for Local Internet Mail Delivery on Dec 5, '94									
Node	ave	< 300.0sec	< 600.0sec	< 900.0sec	> 900.0sec	maxtime	maxsec		
SCSU1	255.15	69.6 (32)	89.1 (41)	97.8 (45)	2.2 (1)	09:45	1434.0		
SCSU6	5.48	100.0 (44)	100.0 (44)	100.0 (44)	0.0 (0)	11:44	83.0		
SERV02	1010.94	62.5 (20)	71.9 (23)	71.9 (23)	28.1 (9)	10:43	6254.0		
SERV03	43.67	100.0 (21)	100.0 (21)	100.0 (21)	0.0 (0)	08:29	137.0		

Response Time Service Summary for SMTP Probe to Port 25 on Dec 5, '94									
Node	ave	< 1.0sec	< 5.0sec	< 10.0sec	> 10.0sec	maxtime	maxsec		
SCSU1	9.36	0.0 (0)	42.4 (25)	78.0 (46)	22.0 (13)	07:35	94.7		
SCSU6	3.47	0.0 (0)	98.3 (58)	100.0 (59)	0.0 (0)	14:20	5.9		
SERV02	6.82	0.0 (0)	59.3 (35)	84.7 (50)	15.3 (9)	11:05	23.1		
SERV03	4.54	0.0 (0)	86.4 (51)	94.9 (56)	5.1 (3)	08:50	14.6		

Figure 6: Example of Network Services Timing Report

Ultimately this data will be further analyzed and graphed to establish baselines and alert when performance of a service starts to exceed the baseline value. This would indicate that there is a deterioration in service and that further analysis and action is necessary.

### 3.0.4 Network Server Capacity Monitoring

There are several aspects to this area including cpu utilization, utilization patterns, load distribution, trends and capacity; and, disk utilization, utilization patterns, load distribution, trends and capacity. Information on these activities can be obtained by seeing reference 11.

### 3.0.5 Networked Server Realtime Performance Monitoring

Realtime "computer" performance monitoring has not yet been tackled. In 1995, SLAC will explore this area.

As indicated in this section, Performance Monitoring is a very complex problem. Not only must the physical network performance be monitored, but the network services and the servers themselves.

## FUTURE PLANS

The way network monitoring is done today at SLAC is not likely to scale very well over the next few years. Currently it is centralized on two machines, an RS/6000 running Netview/6000 and SAS analysis, and a SUN which does the data collection and non-SAS analysis. The disks used for the data and analysis output are in NFS and very heavily accessed. Although the traffic for data collection on the individual subnets is small, all the small loads must traverse the busy main server net, adding to that load.

Proposed design changes include adding more data collection and analysis machines and distributing them in major sections of the network. The machine in a given area would then only collect data and analyze it for that local area, although the data and analysis output would be accessible from anywhere at SLAC. In addition, the data collection and analysis machines would have their own local disk space which would cut down dramatically on the network bandwidth used for file access. This would also help reduce the number of points of failure for the data collection and analysis.

This distribution of network monitoring servers would also facilitate better connectivity monitoring, as connectivity would be monitored from several points in the system, not just one point.

Many new physical network plant technologies are going to be added in the coming months. Monitoring code will have to be developed and incorporated in the existing system for these new technologies.

## SUMMARY AND CONCLUSIONS

Network monitoring and performance analysis of today's **systems** is a very complex matter. It involves not only monitoring the capacity, configuration, and performance of the physical network plant, but also the capacity, configuration, and performance of everything attached to the physical network plant. There are many intertwining factors which have an effect on it. To get a handle on it it is necessary to collect and analyze voluminous amounts of data, and then reduce the data to a form that it can be effi-



ciently and effectively reviewed. One way to approach it has been shown here.

All of this has also been very valuable for setting objectives and expectations for what users should expect from the **system** and for seeing how well we are meeting the objectives. Current SLAC network service level objectives include:

- keeping network layer response time (as measured by ping) to less than 10msec for better than 95% of the samples,
- maintain sub-second response times for the majority of trivial network service requests (such as nameservice),
- for 90% of the on-site mail to be delivered in less than 30 minutes for trivial mail items,
- and provide for network reachability of 99.5%.

## REFERENCES

1. Krol, Ed; *The Whole Internet Catalog*, O'Reilly & Associates, 2nd edition, Sebastopol, CA, 1994.
2. Berners-Lee, T, R. Cailiau, A. Luotonen, H. Nielsen and A. Secret, "The World-Wide Web", Comm. ACM, Vol. 37, No. 8, p 77, August 1994.
3. Wall, Larry & Schwartz, Randal; *Programming perl*, O'Reilly & Associates, Sebastopol, CA, 1990.
4. *Uni-REXX*, The Workstation Group, Rosemont, IL, 1994.
5. SAS Institute Inc.; *SAS Language Reference, Version 6, First Edition*, Cary, NC.
6. van Oorschot, Ir. Jan; *Tricklet*, Data Network Performance Analysis Project, CARDIT, Delft University of Technology, 2600 GA Delft, The Netherlands. Email: J.P.M.vOorschot@et.tudelft.nl.
7. Rose, Marshall, *The Simple Book, An Introduction to Management of TCP/IP-based Networks*, Prentice Hall, Englewood Cliffs, New Jersey, 1991.
8. Leinward, Allan and K. Fong, *Network Management, A Practical Perspective*, Addison-Wesley Publishing Company, 1993.
9. Downey, Teresa, *CANDO Reference Guide*, URL: <http://www.slac.stanford.edu/cando/cando.ps>. February 1993.
10. URL to access the SLAC Network Performance Reports is: <http://www.slac.stanford.edu/netdoc/perf-rep.html>.
11. White, Lois, URL=<http://www.slac.stanford.edu/usr/local/doc/systems/usage-report-overview>.

## ACKNOWLEDGEMENTS

The authors would like to thank their many colleagues at SLAC who have contributed to the success of this system:

- Lois White (LMWHITE@SLAC.Stanford.Edu) has been responsible for all the SAS programming, and the individual computer system cpu and disk space accounting,
- John Halperin (JXH@SLAC.STANFORD.Edu) has helped with the development of tools to monitor the performance of network services,
- Teresa Downey (TERESA@SLAC.Stanford.Edu) developed and maintains the CANDO database.,
- Don Pelton (DEP@SLAC.Stanford.Edu) who installed the Tricklet SNMP routines,

and, the other members of the Network Development and Network Operations groups who have used the system in the heat of the battle and provided invaluable feedback on how to make it more useful to them.