

# The PingER Project: Active Internet Performance Monitoring for the HENP Community

Warren Matthews and Les Cottrell, Stanford University

## ABSTRACT

The extraordinary network challenges presented by high energy nuclear and particle physics experiments has created a need for network monitoring both to understand present performance and to allocate resources to optimize performance between laboratories, and the universities and institutes collaborating on present and future experiments. The resulting Internet End-to-End Performance Monitoring project is called PingER. The monitoring infrastructure reflects the wide geographical spread of the collaborations, and involves a large number of research and commercial networks. The architecture of the data acquisition and methodology of the analysis have evolved over several years, and are described here in their present state. The strengths and weaknesses of the project are reviewed, and the derived metrics are discussed in terms of their diagnostic functions. The observed short-term effects and long-term trends are reviewed, and plans for future developments are described.

## INTRODUCTION

Modern high energy nuclear and particle (HENP) physics experiments at laboratories around the world present a significant challenge to wide-area networks. The *BaBar* collaboration at the Stanford Linear Accelerator Center (SLAC), the Relativistic Heavy Ion Collider (RHIC) groups at Brookhaven National Laboratory (BNL), and the Large Hadron Collider (LHC) projects under development at the European Center for Particle Physics (CERN) will all generate petabytes ( $10^{15}$  bytes) or exabytes ( $10^{18}$  bytes) of data during the lifetime of the experiment. Much of this data will be distributed via the Internet to the experiments' collaborators at universities and institutes throughout the world for analysis.

In order to assess the feasibility of the computing goals of these and future experiments, a number of committees and working groups founded projects to monitor performance. The combined efforts of these projects has resulted in a large end-to-end performance monitoring infrastructure being set in place with an active network probing system along with a set of tools for analyzing the

data. This architecture has become known as *PingER*, for ping end-to-end reporting.

In particular, the activity of two groups drives the monitoring activities and is the basis of the detailed *PingER* project review in this article. The first of these groups is the Network Monitoring Task Force (NMTF) of the Energy Sciences Network (ESnet), which takes particular interest in performance between laboratories funded by the U.S. Department of Energy (DOE) and the universities and institutes involved in research at these laboratories. The second group is the Standing Committee on Interregional Connectivity (SCIC) of the International Committee for Future Accelerators (ICFA), which addresses problem areas of international and especially transoceanic performance across multiple networks connecting laboratories and universities conducting HENP research.

In addition, a consortium of industry leaders working together under the Cross Industry Working Team (XIWT) and, in particular, the Internet Performance Working Group (IPERF) are also using the *PingER* framework to monitor performance on commercial networks. The International Atomic Energy Authority (IAEA) used *PingER* in a study of performance between their European offices and nuclear power stations in South America. The authors are also aware that several Internet service providers and network operation centers have trialed the tools; however, it appears that only the HENP analysis work has been made public.

## THE PINGER PROJECT

As its name indicates, the framework of the *PingER* project is based on the *ping* program familiar to network administrators and widely used for network troubleshooting. A *ping* involves sending an Internet Control Messages Protocol (ICMP) echo request [1] to a specified remote node which responds with an ICMP echo reply. It is also optional to send a data payload in the request which will be returned in the reply. The round-trip time (RTT) is reported; if multiple pings are dispatched, most implementations provide statistical summaries.

The performance of applications using TCP or UDP can be inferred from *PingER* metrics

The authors are with the Stanford Linear Accelerator Center (SLAC), a particle physics laboratory operated for the U.S. Department of Energy by Stanford University.

derived from the performance of ICMP packets because typically routers do not treat transit packets according to transport-level protocol and just send them on to the next hop. This assumption can be justified because studies [2] have shown a strong lower bound where a Hypertext Markup Protocol (HTTP) get is twice the minimum *ping* RTT. The factor of 2 can be understood because a minimal TCP transaction such as an HTTP get requires two round-trips, so the result indicates the packets experience similar conditions. A further source of validation is use of the ping packet loss and RTT in an equation to calculate the maximum TCP transfer rate [3] and compare the predicted value with measured TCP throughput using  $T_{TCP}$ . Initial studies show good agreement, and a detailed study is in progress to understand the agreement considering TCP congestion avoidance. The *PingER* methodology provides good agreement with protocols and applications important to end users. It also provides an advantage over monitoring performance of a specific application because the separate packet loss and RTT statistics provide details to understand network activity. Monitoring application performance, which may involve a complicated interaction of multiple packets with backoff and retransmit effects, only provides information on that particular application.

In December 1999 there were 20 *PingER* monitoring sites around the world. There are eight monitoring sites in the United States: the Stanford Linear Accelerator Center (SLAC) in California; the HEP Network Resource Center (HEPNRC) at the Fermi National Accelerator Laboratory near Chicago, Illinois; the Department of Energy (DOE) office in Washington, DC; the Brookhaven National Laboratory (BNL) in Upton, New York; the Atmospheric Radiation Measurement (ARM) facility at the Pacific Northwest National Laboratory in Richland, Washington; Carnegie-Mellon University (CMU) in Pittsburgh, Pennsylvania; Stanford University in Palo Alto, California; and the University of Maryland (UMD) at College Park, Maryland. There are two further North American sites in Canada: the National Meson facility (TRIUMF) near Vancouver, and at Carleton University in Ottawa. There are also seven sites in Europe: at CERN on the Swiss-French border near Geneva, Switzerland; the Deutsches Elektronen Synchrotron (DESY) near Hamburg, Germany; the Rutherford Laboratory (RAL) and Daresbury Laboratory, both in England; the Niels Bohr Institute (NBI) in Denmark; INFN's National Center for Telematics and Information (CNAF); and the Research Institute for Particle and Nuclear Physics (KFKI) in Budapest, Hungary. Finally, there are three monitoring sites in Asia: at the KEK facility in Tsukuba, Japan; the Riken Institute of Physical and Chemical Research (RIKEN) in Wako, Japan; and Academia Sinica (SINICA) in Taiwan, China.

The SLAC, HEPNRC, DOE, BNL, and ARM monitoring sites are connected to ESnet, which connects the DOE-funded laboratories in the United States. Fermilab (and hence HEPNRC) are also connected to the Metropolitan Research and Education Network (MREN), and BNL is also connected to the New York State Research

and Education network (NYSERnet). CMU and UMD are connected to the vBNS (Very-High-Performance Backbone Network Service). Stanford University is connected to CALREN2. UMD is also connected to the Abilene Gigapop in Washington, and CALREN2 is connected to the Abilene Gigapop in Sacramento. Each of these networks has a backbone typically running at OC12 (622 Mb/s), and most connections are OC3 (155 Mb/s). The TRIUMF and Carleton University monitoring sites are connected to provincial networks which are connected to the Canarie national research backbone (CA\*net2) operating at 2xOC3. Most of the European monitoring sites are connected to a national research network (NRN) and interconnected with the European TEN-155 network. Typically, in Western Europe, apart from Spain and Portugal, the TEN-155 backbone operates at 155 Mb/s. To Eastern Europe, Spain, and Portugal the connections are 10–45 Mb/s. Most NRNs have their own connection to North America. KEK is connected to the Japanese National Center for Science Information Systems (NACSIS) network, and also has a connection to ESnet. The RIKEN and SINICA monitoring sites have connections with commercial providers.

How these networks are interconnected and the routing policies practiced by the administrators are of critical importance to performance. Many networks have direct peering relationships, sometimes in several different locations. The STAR TAP in Chicago is a popular meeting point for research networks.

*PingER* sends 11 *pings* with a 100-byte payload, at 1 s intervals, followed by 10 *pings* with a 1000-byte payload, also at 1 s intervals, to each of a set of specified remote nodes listed in a configuration file. The first *ping* is discarded because it is assumed that it is slow due to priming caches. Studies using UDP echo packets have found that the first packet takes about 20 percent longer than subsequent packets [4]. The *ping* default timeout of 20 s is used. A study of poor links with long delays (in particular ones involving satellites) indicates that the number of *ping* packets returning after 20 s is small. For such links the number returning after 20 s but before 100 s is less than 0.1 percent. Small packets are sent to avoid fragmentation. Each set of 10 *pings* is called a *sample*, and each monitoring node-remote node combination a *pair*. In September 1999 there were 1977 pairs with 511 remote nodes at 355 sites in 54 countries on six continents.

Historically, each monitoring site has monitored remote nodes of interest to it, so monitoring sites at laboratories would *ping* the universities and institutes involved in collaborations at each laboratory, and monitoring sites at universities would *ping* the laboratories where research was conducted. Consequently, many countries are represented by only one node, but this node is usually at a research university, and the performance is assumed to be representative of networking in HENP rather than for the entire country. The United States and West Europe make up most of the nodes, and this is fair for a study of HENP internetworking because most of the international experiments and the universities that collaborate on them are

*The PingER methodology provides good agreement with protocols and applications important to end users. It also provides an advantage over monitoring performance of a specific application because the separate packet loss and RTT statistics provide details to understand network activity.*

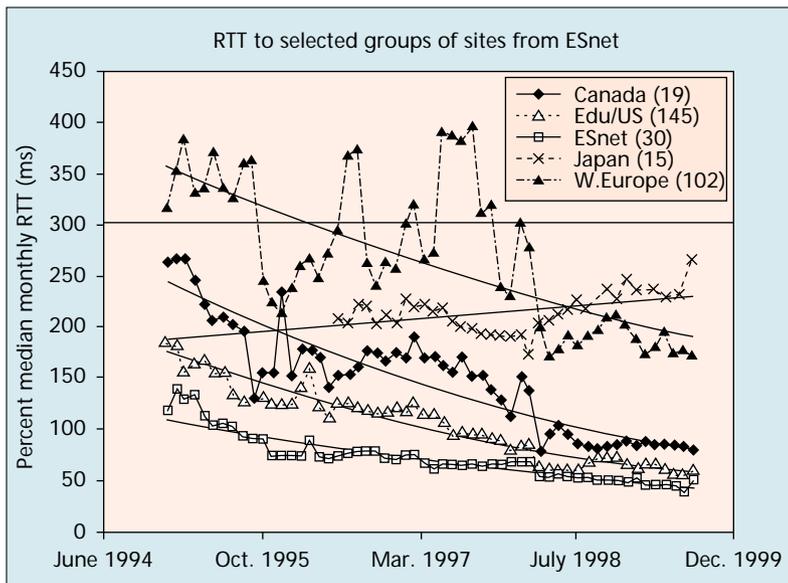


Figure 1. Round-trip time between ESnet and sites in several groups. The numbers in parentheses in the legend are the number of pairs aggregated.

in these regions. Around 70 percent of the remote nodes in the United States are educational or government sites, and 20 percent are connected to ESnet.

Recently the concept of *beacon sites* has been introduced, and all monitoring sites are requested to *ping* all the beacon sites. Beacon sites represent the various affinity groups monitored. All monitored sites, but especially beacon sites, should be reliable and available 24 hours a day, 7 days a week. The node should also be lightly loaded (or at least consistently loaded) and obviously responsive to *pings*. In order to achieve these goals and ensure they are kept, some beacon sites have created a CNAME. This enables the administrator to change the machine if it fails to meet any of the criteria. Other factors determining the selection of a beacon site include its physical location, backbone connectivity, and importance to HENP in general. Monitoring a selected subset of nodes from all monitoring sites gives better information for troubleshooting and understanding the network in general. Currently there are 53 beacon sites: 23 in the United States and Canada, 13 in West Europe, six in East Asia, three in South America, five in East Europe, two in Australia, and one in India.

Each monitoring site *pings* its set of remote nodes every half hour. The coarseness of the measurements was chosen so that their impact on the network is low (100 b/s/pair on average), but at the same time they provide reasonable trend information. *PingER* was never envisioned to do real-time measurements for network operation centers. For many links this rate provides adequate details; however, in some cases where packet loss is rare, it may be beneficial to increase the sampling rate and the number of packets in each sample. The statistical summary of RTT and packet loss from the output of the *ping* program are extracted; the median RTT is also calculated, and the data is written to a file. A common gateway interface (CGI) program running on a Web server at the monitoring site is used to display

<sup>1</sup> The situation is complicated by router algorithms such as Random Early Drop (RED), which discards packets even if the buffer is not full. However, such algorithms are not widely deployed in routers, and the above is sufficient to understand packet loss in general.

and retrieve the data from the file. All data is publicly available. Most monitoring sites provide a CGI program which displays the most recent results from that site on a Web page. Each day the archive site at HEPNRC retrieves the data and stores it in a simple attachment scheme (SAS) database. Currently, roughly 600 MBytes of data are stored each month.

The data is compiled into separate reports at the analysis site using code written in Perl and Java: a daily report where the data is aggregated to one value per hourly tick for each day; a monthly report where data is aggregated to one value per daily tick for each month; and a summary report containing monthly ticks. Each daily, monthly, and summary report is available with each monitoring site and remote node pair displayed individually, or nodes in the same domain at the same site can be aggregated. All reports are made available through a Web page as HTML tables and in tab-separated values (TSV) format for importing into spreadsheet packages. Graphs are generated at the archive site from the data using SAS and Perl.

### DATA ANALYSIS

Ideally, traffic should traverse the Internet at the maximum speed for the medium (e.g., the speed of light in glass for fiber). However, connections very rarely do. Packets must be received by and sent from the routers' network interfaces, and if there is significant congestion on the line, the delay caused by queuing and processing in routers may add a large delay, or the packet may even be dropped if the buffer is full.

The *PingER* analysis defines five metrics, designed to look for the effect of this queuing to estimate network performance. The five metrics are packet loss, RTT, unreachability, unpredictability, and quiescence.

**Packet Loss** — Packets must queue in buffers in order to be processed by a router, and if the queue is full the packet is discarded;<sup>1</sup> therefore, packet loss gives a good indication that at least part of the link is congested. Typically, performance of an application using TCP will deteriorate significantly after 3 percent packet loss due to retransmission of packets governed by the TCP algorithms, but the effect seen by the end user will vary according to the application. Different applications will vary in the extent to which packet loss affects their usability. Highly interactive applications such as videoconferencing will become unusable even with moderate packet loss, whereas noninteractive applications such as e-mail will work even across a network with high packet loss.

**Round-Trip Time** — The queuing in buffers described previously also affects the RTT. However, unlike packet loss, where it is possible to reduce losses to zero, it is never possible to reduce the RTT to less than the time taken for light to travel the distance along the fiber. The reported RTT is the minimum imposed by the laws of physics, the time taken for the packet to be accepted by the router interface, any delay caused by queuing, and the time taken for the packet to be transmitted from the interface. The minimum RTT indicates the length of the route

taken by the packets, the number of hops, and the line speeds. The distribution of RTT indicates the congestion. Changes in minimum RTT can be an indication of a route change.

The major effect of poor response time is on interactive sessions such as telnet,<sup>2</sup> or packetized video or voice, where even fairly moderate delay can cause severe disruption. Applications that do not require such a level of interactivity (e.g., e-mail) may appear to perform well even with high delay.

**Unreachability** — If no reply is received from all 10 *ping* packets sent to a remote node, the remote node is considered unreachable. It is important for accurate network performance analysis that the cause of this unreachability be network performance. However, computers crash and become unresponsive, or there is some reason other than network performance that no replies were received, and it is extremely difficult to program analysis code to tell the difference. In addition, statistical fluctuation means that pairs involving links that suffer high packet loss will sometimes be reported as unreachable when in fact the node is reachable but the packet loss exceeds 90 percent. In practice it is left to the analyst's judgment of whether the node is truly unreachable due to network problems. Often a number of nodes will become unreachable at the same time, indicating that a common cause has affected them all. High-performance research networks are considered as good as it gets, but they too experience glitches. Typically, less than 1 percent of the *PingER* samples between nodes on these networks are completely lost; hence, unreachability less than 1 percent is classified as good. However, it is somewhat subjective to claim that a user would consider any amount of unreachability acceptable.

**Quiescence** — If all 10 *ping* packets sent to a remote node receive a reply, the network is considered quiescent or nonbusy. The frequency of this zero packet loss may be an indication of the use of the network. For example, a network that is busy eight work hours per weekday and quiescent at other times would have a quiescent percentage of about 75 percent. If the network is busy all during the workday, it is considered "poor" and probably needs upgrading.

**Unpredictability** — Unpredictability is derived from a calculation based on the variability of packet loss and round trip time.<sup>3</sup> The *ping* success is the proportion of replies received from the number of packets sent, and the *ping* rate is twice the ratio of the *ping* payload to the average RTT. In any time period, the ratio of the average and maximum *ping* success, *s*, and that of the average and maximum *ping* rate, *r*, is combined to create the unpredictability, *u*, where

$$u = \frac{1}{\sqrt{2}} \sqrt{(1-r)^2 + (1-s)^2}.$$

The derived value is a measure of the stability and consistency of a link; a high-performing link with low packet loss and low RTT will be ranked good, but a poorly performing link with consistently high packet loss and high RTT will

also be ranked good as long as the packet loss is consistent. However, links where packet loss and RTT can vary dramatically will rank poorly because the end user will be unsure just how the link will perform.

**LIMITATIONS OF THE PINGER METHODOLOGY**  
There are two main issues with the *PingER* methodology: periodic sampling and the use of ICMP packets.

Periodic sampling using cron to send *pings* at regular intervals has proved to be a powerful tool in understanding network performance; however, periodic network events that occur at times other than when the samples are scheduled will never be observed, and periodic network events that do occur regularly just as the samples are scheduled may make the network performance appear to be poorer than it is. The more effective method, and in accordance with RFC 2330 [5], would be to observe the network at random times using a Poisson distribution of samples.

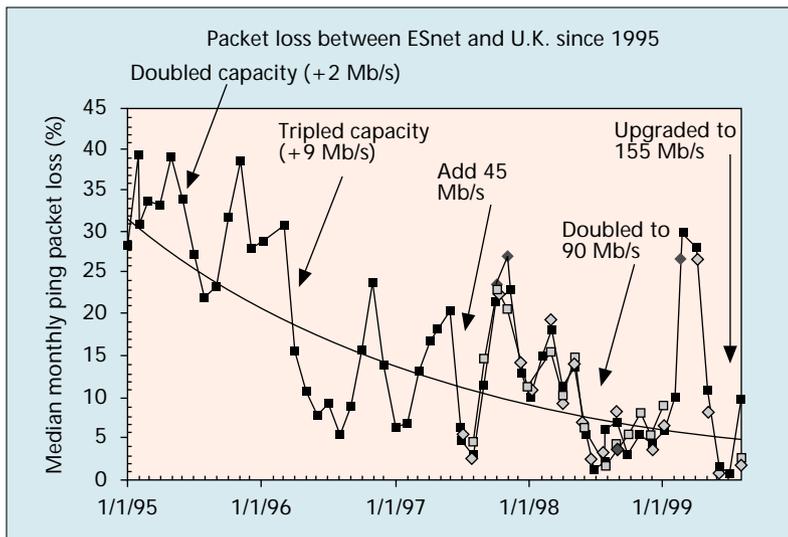
The use of ICMP packets is the source of some controversy because quality of service (QoS) techniques deliberately hinder the progress of some types of packets in order to increase the performance for other types, and ICMP packets may be given low priority in order for higher priority to be given to TCP and UDP packets. Furthermore, *ping* packets can be used in certain types of security attacks such as smurf attacks. Some networks give low priority to *ping* in order to reduce the effect of such attacks, so *ping* monitoring makes the network appear to have poorer performance than it actually has. In some cases rate limiting is also observed. This is usually implemented by networks with low-bandwidth connections and restricts the amount of ICMP traffic allowed to flow through the router. Once the limit has been reached, further ICMP (including *ping* packets) will be dropped. It may be possible to use a TCP- or UDP-based echo program, but these may involve their own security issues and are often blocked too. Other monitoring projects such as Surveyor [6] used dedicated machines to exchange one-way UDP packets.

In addition to the above issues, there are also the pathologies of out-of-order responses and duplicate responses. We record out-of-order packets. Between December 1998 and September 1999, for the SLAC measurements, we recorded less than 0.1 percent samples with out-of-order responses, and 2/3 of these came from one site in China. In all cases the out-of-order packet was due to an extraordinarily long response time (up to 68 times as long as a normal response). Occasionally a ping echo request will result in greater than one echo response. For example, we may send out six pings (sequence numbers 0-5), and the sequence numbers of the responses are 1, 1, 2, 2, 4, 4; that is, the pings with sequence numbers 0 and 3 are lost, but the system stops listening after six responses are received, so the response from ping 5 is ignored. Between December 1998 and September 1999, for the SLAC measurements, less than 0.01 percent of the samples have had duplicate response, and all were involved with a host in Russia. Almost all of these are packets which experienced long delays (up to 94 s), and

*The use of ICMP packets is the source of controversy because QoS techniques deliberately hinder the progress of some types of packets in order to increase the performance for other types, and ICMP packets may be given low priority in order for higher priority to be given to TCP and UDP packets.*

<sup>2</sup> Of course, everyone uses ssh these days.

<sup>3</sup> Developed by Hans-Werner Braun.



■ **Figure 2.** Packet loss between U.S. sites connected to ESnet and sites at universities in the U.K. connected to the JANet NRN. Packet loss decreases significantly as the capacity of the link is upgraded, and during university holidays.

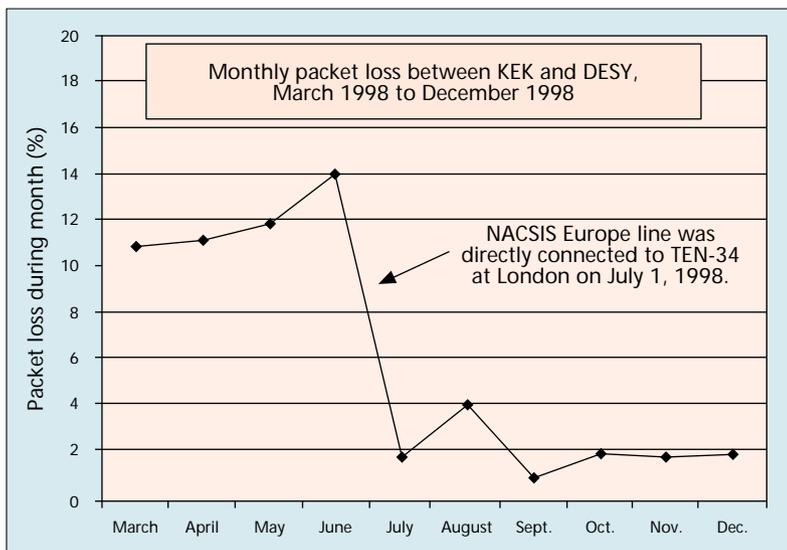
the ping client incorrectly associated the response with the wrong echo request.

## RESULTS AND CONCLUSIONS FROM PINGER MONITORING

*PingER* provides insight into a multitude of network activity. To summarize the results, these are divided into measurements of long-term trends brought about by planning and resource allocation; and the short-term glitches and changes such as outages, and the immediate outcome of upgrades and reconfiguration.

### LONG-TERM TRENDS

In the United States, performance on high-speed research networks such as ESnet, vBNS, and Abilene is good, and performance between these networks is also good.



■ **Figure 3.** Packet loss between KEK in Japan and DESY in Germany. Packet loss decreased from unusable to acceptable after the NACSIS line from Japan to Europe was relocated to London and connected directly to the TEN-34 network.

During normal operation, packet losses are negligible and RTTs are minimum. A major cause of this high performance is that the links are far from saturated; in fact, at times the utilization is less than 5 percent [7]. Connections between the networks avoid congested Internet exchange points. Future HENP experiments and other next-generation Internet (NGI) projects will involve significantly more utilization, but the long-term trend for most connections monitored by *PingER* has so far been toward steady improvement. Figure 1 shows the RTT between sites on ESnet, and sites in several regions and groups. Exponential fits to the data are included to show the overall improvement even though the monitoring has grown over time to include more remote regions. In most cases advances in infrastructure have stayed at least one step ahead of the demands of bandwidth-hungry applications.

In Europe, where NRNs provide connectivity for research entities within each country, performance within the NRN is usually good. Performance across the TEN-155 backbone is also good, but connectivity to Eastern Europe is often poor.

International connectivity, particularly transoceanic performance, has improved dramatically since the *PingER* project began. Figure 2 shows the packet loss between ESnet sites in the United States and universities in the United Kingdom since January 1995. An exponential fit to the data is marked to show the overall trend toward better performance. Initially, the only monitored link was SLAC to RAL, and the 2 Mb/s connection was suffering heavy packet loss. In June 1997 three more ESnet monitoring sites began to monitor links to different U.K. sites. It can be seen that the measured packet loss tracks from pair to pair extremely well. This observation was the basis for the decision to select beacon sites rather than attempt full mesh ping. Packet loss decreases when the capacity of the trans-Atlantic link is upgraded, but increases again as the extra headroom is quickly taken up. Packet loss also decreases significantly during University holidays. The trans-Atlantic link, as of July 5, 1999, consists of 2 x OC3 connections (155 Mb/s each), an increase in bandwidth of about 150 times in 4.5 years.

### SHORT-TERM GLITCHES AND CHANGES

Over shorter timescales, all connections, even the high-performance research networks, suffer glitches. This section describes a few examples to illustrate typical observations.

The connection between the KEK facility in Japan and the DESY laboratory in Germany, shown in Fig. 3, suffered high packet loss. On 1 July 1998, the NACSIS Europe line was directly connected to TEN-34 in London, rather than the traffic transiting across a third-party network creating a more direct connection.

Colorado State University is connected to the Abilene Network. ESnet peers with Abilene in several places, and packets from CSU to SLAC were being routed via Chicago. Figure 4 shows the effect of changing the route to send packets bound for SLAC to the ESnet-Abilene peering point in Sacramento. In this case the route change was planned, and it can be seen that the

RTT became much better. Similar changes are often observed and are indicative of routing changes, but these are frequently not planned, and more often things get worse, usually for small periods of time due to a router outage.

The connectivity between sites in Scandinavian countries connected to the NORDUnet and sites in North America ranges from good to poor. In December 1998, ping monitoring reported that these links suddenly became unusable. Figure 5 shows the packet loss between NBI and FNAL. However, the apparent deterioration in performance was due to the installation of Smurf filters on NORDUnet's U.S. connection. These filters are designed to defeat a security attack by giving very low priority to *ping* packets, which results in many packets being dropped. TCP and UDP traffic is unaffected by these filters.

Figure 6 shows the zero packet loss frequency (quiescence) between RAL and the National Institute for Nuclear and High Energy Physics (NIKHEF) in the Netherlands. Typically, the link was busy during the workday and suffered only minor packet loss at other times until the TEN-155 backbone became operational on 11 December, providing additional bandwidth such that the connectivity during the workday became better and packet loss was reduced.

The Institute for High Energy Physics (IHEP) in Beijing, China, has a direct connection to the KEK facility, which provides connectivity to the HENP community. However, the HENP community often found nodes at IHEP unreachable; the connection was completely saturated. When the link was upgraded from 64 to 128 kb/s on 30 October 1998, the unreachability decreased from up to 25 percent to less than 5 percent. Further improvement is required to be able to utilize the Internet for research at IHEP, but even a modest upgrade can make a vast improvement.

## CONCLUSIONS REGARDING THE PINGER METHODOLOGY

The *Pinger* methodology has been highly effective. The results discussed earlier illustrate how a simple tool can provide insight into the cause and effect of short-term glitches and long-term trends. New technologies certainly provide a challenge for the future development of the *Pinger* tools, and the creativity and imagination of the analyst trying to understand performance. In some cases, which will certainly increase, *Pinger* will not be able to provide accurate monitoring, but these cases would probably defeat many sophisticated monitoring tools too.

## FUTURE WORK

Particle physics experiments keep getting more powerful and more complex. They probe deeper into the subnuclear world, and require larger collaborations of physicists to run them, analyze the huge quantities of data, and spread the cost. Physicists are also an increasingly cosmopolitan group, and require access to the experimental data from their desktops at their home institutions. The capacity of research networks will grow and new technologies be developed. Fur-

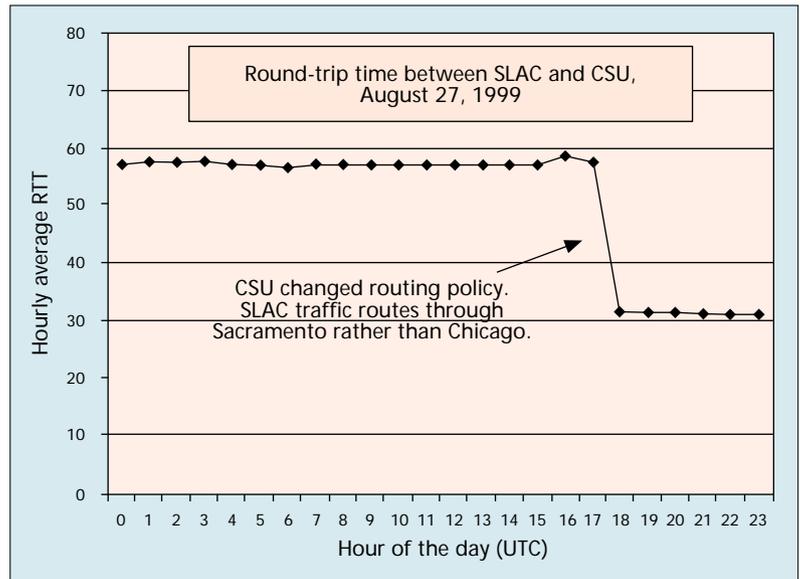


Figure 4. Average round-trip time between SLAC and Colorado State University dropped after CSU changed routing policy so that packets destined for SLAC were sent to Sacramento rather than the ESnet hub in Chicago.

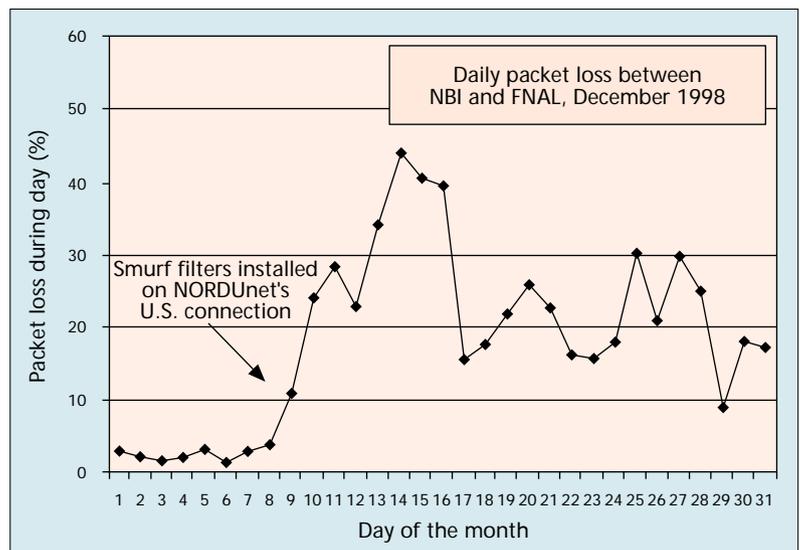
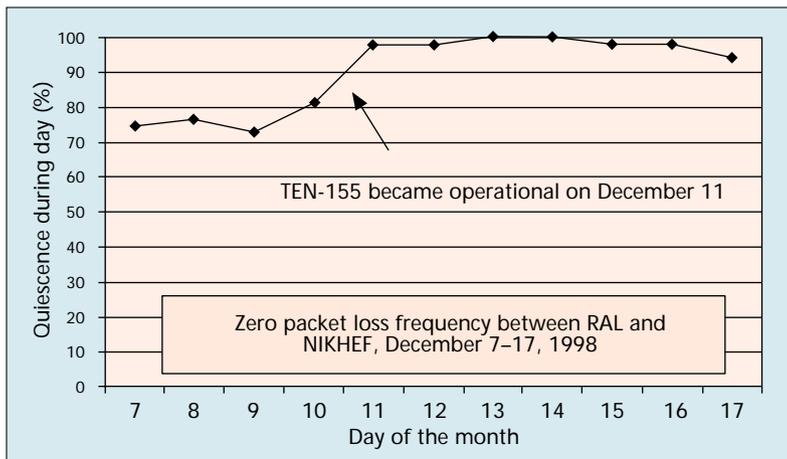


Figure 5. Packet loss between NBI and FNAL appeared to increase wildly in December 1998. However, this is an effect of a filter used to prevent security attacks and is not a true indication of network performance.

ther development of the *Pinger* framework is necessary to keep pace and provide accurate performance monitoring.

A number of projects are actively being worked on. Several new metrics that can be derived from the *Pinger* data for use in determining network performance are to be included in the analysis and will be reported in the future. The variation between the individual RTT in each sample, what may be called packet delay variation or jitter, is being studied on a testbed that has been set up between SLAC, Lawrence Berkeley National Laboratory (LBNL), and Sandia National Laboratory. This test bed is being used for a detailed study of voice over IP (VoIP) and QoS. New tools and techniques may be developed to understand how network performance correlates to new applications. An exten-



■ **Figure 6.** Frequency of zero packet loss (quiescence) between RAL and NIKHEF improved when the TEN-155 network became operational in December 1998.

sion to the PingER framework to include traceroute is being developed. A metric based on minimum RTT will also be included, and correlations with route changes studied. Further comparison between the TCP throughput, measured with *TCP* and other tools, and predicted values of the maximum TCP transfer rate [3], using the *PingER* packet loss and RTT, are underway. Also, we are looking at extracting signatures to identify ICMP rate limiting.

A version of *PingER* using Poisson sampling has been developed, but deployment has been delayed due to stability problems. Extra effort will be allocated to resolving the issues and installing the code at the monitoring sites. More configurability will be built into the framework. Variability in payload sizes will be studied to understand if this is a more friendly way to monitor low-bandwidth connections. Variability in the number of *ping* packets in a sample and the sampling rate will also be studied on very well performing links.

Several *PingER* monitoring sites are also home to other network monitoring projects such as *Surveyor* [6], *NIMI* [8], *AMP* [9], and *RIPE* [10]. Studies looking at correlations between network performance determined by these projects have begun at SLAC. In addition, passive monitoring using various sniffing tools is being deployed, and agreements and disagreements will be studied.

A major challenge is the visualization of results. The data and reports will be made available in other formats for importing into statistical analysis and histogramming packages. Techniques developed by other groups will be examined.

#### ACKNOWLEDGMENTS

The authors would like to acknowledge the suggestions of many people, in particular the insightful comments of David Williams and Harvey Newman, and the members of the network monitoring mailing list. Thanks to John Halperin and Charles Granieri for their work on the network monitoring project, and to Bill Lidinsky, Shiqi He, and especially David Martin for main-

taining the archive site and developing the data gathering tools. Special thanks go to Rebecca Moss for her hard work figuring out all the political, research, and network affiliations of all the sites monitored.

This work would not have been possible without the efforts of the maintainers of the monitoring sites: Wen-Shui Chen, Ricardo G. Patará, Darren Curtis, Mike O'Connor, Olivier Tarter, Wade Hong, Michael Procario, Robin Maskin, Michael Ernst, Michael C. Weaver, Cristina Vistoli, Jae-young Lee, Fukuko Yuasa, Bjorn S. Nilsson, Takashi Ichihara, Piroska Giese, Andrew Daviel, and Drew Baden; and the support and funding of the Mathematical, Information, and Computational Sciences Division of the U.S. Department of Energy (DOE/MICS).

We would also like to thank the members of the ICFA-SCIC and ESnet committees for their support of this work.

#### REFERENCES

- [1] J. Postel, "Internet Control Message Protocol," RFC 792, <ftp://ftp.isi.edu/in-notes/rfc792.txt>, Sept. 1981.
- [2] R. L. A. Cottrell and J. Halperin, "Effects of Internet Performance on Web Response Times," <http://www.slac.stanford.edu/comp/net/wan-mon/ping/correlation.html>, Dec. 1996.
- [3] M. Mathis *et al.*, "The Macroscopic Behavior of the TCP Congestion Avoidance Algorithm," *Comp. Commun. Rev.*, vol. 27, no. 3, July 1997.
- [4] M. Horneffer, IPPM mailing list, <http://www.advanced.org/IPPM/archive/0246.html>, Jan. 1997.
- [5] V. Paxson *et al.*, "Framework for IP Performance Metrics," RFC 2330, <ftp://ftp.isi.edu/in-notes/rfc2330.txt>, May 1998.
- [6] The Surveyor Project Advanced Networks, <http://www.advanced.org/surveyor>
- [7] Abilene Network Operations Center Web site, <http://www.abilene.iu.edu>
- [8] The National Internet Measurement Infrastructure (NIMI) Project, <http://www.psc.edu/networking/nimi>
- [9] The Active Measurement Project (AMP), <http://amp.nlanr.net/AMP>
- [10] The RIPE Test Traffic Project, <http://www.ripe.net/test-traffic>

#### BIOGRAPHIES

LES COTTRELL ([cottrell@slac.stanford.edu](mailto:cottrell@slac.stanford.edu)) left the University of Manchester, England, in 1967 with a Ph.D. in nuclear physics to pursue fame and fortune on the Left Coast of the United States. He joined SLAC as a research physicist in high energy physics, focusing on real-time data acquisition and analysis in the Nobel prize winning group that discovered the quark. In 1973–1974, he spent a year's leave of absence as a visiting scientist at CERN in Geneva, Switzerland, and he spent 1979–1980 at the IBM U.K. Laboratories at Hursley, England, where he obtained United States Patent 4,688,181 for a dynamic graphical cursor. He is currently the assistant director of the SLAC Computing Services group and leads the computer networking and telecommunications areas. He is also a member of the Energy Sciences Network Site Coordinating Committee (ESCC) and chair of the ESnet Network Monitoring Task Force. He was a leader of the effort that, in 1994, resulted in the first Internet connection to mainland China. He is also the leader of the DOE-sponsored Internet End-to-end Performance Monitoring (IEPM) effort.

WARREN MATTHEWS ([warrenm@slac.stanford.edu](mailto:warrenm@slac.stanford.edu)) obtained a Ph.D. in particle physics; then, realizing packets were more fun than particles, he went to pursue packets across networks for an Internet service provider. He joined SLAC as a network specialist in 1997 as part of the DOE Internet End-to-end Performance Monitoring (IEPM) group.