Internet End-to-end Performance Monitoring (IEPM) and the PingER project.

Les Cottrell and Warren Matthews
Stanford Linear Accelerator Center.

1. Introduction.

The ping end-to-end reporting (PingER) project began in early 1995. It was originally a system to monitor wide-area network performance between SLAC and collaborators of experiments. The ESnet Network Monitoring Task Force (NMTF) quickly become involved in the project and the work was extended to cover ESnet sites. Several DoE funded laboratories installed the PingER monitoring software which greatly improved our knowledge of the network.

In April 1997 the International Committee on Future Accelerators (ICFA) created the Network Task Force to cover the needs of the High Energy and Nuclear Physics (HENP) community. PingER became funded by DoE/MICS.

In 2001, the increasing need to understand the entire end-to-end user perception and throughputs achievable on high-speed paths led to the extension of the PingER concepts to use bandwidth measurement tools such as iperf and file transfer programs such as bbftp and GridFTP, popular amongst BaBar physicists and Grid developers/users. This new toolkit, called IEPM-BW to distinguish it from the light-weight IEPM-PingER, enabled the tie in to high-performance network science applications.

Currently the IEPM project is under the program led by Thomas Ndousse. Funding for PingER is due to expire this year. The monitoring continues to be useful and further funding is being sought. In addition new work inspired by the lessons learned from PingER and BW are being developed. In particular the MAGGIE proposal aims to continue the research in this area.

2. Evaluate the Progress.

PingER is a mature project and highly successful in terms of its original goals. It has been deployed throughout the High Energy and Nuclear Physics (HENP) community and has also been used by the European Data Grid (EDG) project, the International Atomic Energy Authority (IAEA), the Cross Industry Working Team (XIWT) and numerous governmental and commercial groups. The maintainers receive several enquiries each month regarding some aspect of the tools. Consequently it is believed the true installed base and research being conducted on the data is considerable.

PingER is based on the ubiquitous ping program. A single ping is sent to prime the cache. Next, by default (the parameters are configurable), 10 pings with 100 byte payloads are sent and finally 10 pings with 1000 byte payloads are sent. The reported packet loss and minimum, average, and maximum round trip time (RTT) are recorded. Furthermore

unreachability, quiescence, unpredictability, jitter and TCP throughput are derived. A metric for out-of-order packets using the IETF standard has been partially implemented.

The software consists of three parts; monitoring, analysis, and visualization. All the software is freely available from the SLAC and FNAL web sites. The tools currently require some manual configuration. However, it is intended to package the software using the gnu autoconf tools to provide quick and easy installation for the user.

Significant work has been applied to validate the results. The network performance has been compared to HTTP and FTP transfers. Such comparisons yield a strong correlation.

The PingER data goes back to January 1995 and are publicly available, as are the various analyses and web based visualization. PingER results have been used to identify sites requiring upgrades. It has also been used to compare performance between providers and sites in order to make decisions about where to allocate resources for experiments. It has been used to set routing policy and set expectations for applications such as Voice-over-IP (VOIP) and file transfer. Researchers consider it a rare resource in particular because of the length the project has run for and the availability of the data. Recently the maintainers learned of an economic policy report being prepared comparing economic indicators and network performance reported by PingER.

The DoE/MICS funded PingER project is currently deployed at 36 monitoring sites in 14 countries. More than 3600 end-to-end pairs are monitored and target hosts are at 473 sites in 79 countries. It is believed the DoE/MICS funded PingER is the most extensive end-to-end network monitoring project in the world. Reports of links of interest to specific groups such as BaBar are contained in special pages and searches can be conducted on specific geographic and affinity groups.

In reality, most network monitoring projects are actually monitoring end-to-end performance. This is because end-host effects are unavoidable. In fact end-host and local area network problems are the cause of most problems. Wide-area networks, especially well-engineered backbones such as ESnet and Abilene, are rarely the cause of performance issues.

In order to recognize the difference between network performance and the true user experience, PingER was extended to use throughput tools such as iperf and bbftp. Subsequently the code was re-written and dubbed IEPM-BW. Currently the pilot IEPM-BW project has been deployed at 10 sites in North America, Europe and Asia taking measurements to between 2 and 36 target sites each. The analysis engine has been configured to read the format and extract results from numerous throughput and file transfer tools. In addition traceroute monitoring, analysis and visualization was added and work has been conducted to use passive netflow data. Other work in progress includes studying the web100 variables.

Similarly, the Internet2 end-to-end performance initiative (e2epi) has also recognized the importance of end-to-end monitoring and is developing a system for its users to be able to

obtain contact information for sites where measurements indicate the problem may lay. The PingER team work closely with the e2epi and associated PIPES development and is on the e2epi Technical Advisory Group.

3. Assess the value.

The PingER project has been hugely successful and continues to provide valuable insight into network performance. The many years of data taking provides a unique source for quantifying long-term trends. The effect of upgrades and vacations can be seen. PingER also provides evidence of the effect of policy changes. For example, the impact of peering arrangements and routing policy. PingER as a research project has provided the data for end users to understand what is going on in the network. PingER as a trouble-shooting tool continues to provide valuable evidence to direct administrators to address the problem. The accompanying presentation shows graphs selected from PingER monitoring that detail this value.

Recently PingER has once again proved useful. It has been employed by the electronic Journal Distribution service (eJDS) based at the Abdus Salaam International Center for Theoretical Physics (ICTP). The eJDS project distributes physics journals to its members worldwide. The project leaders are concerned about the ability of networks in developing countries and remote regions to perform the task adequately. Similarly the international nature of High Energy Physics has caused concern and ICFA created a special committee – the Standing Committee on Inter-regional Connectivity (SCIC) – to quantify the digital divide. It is estimated that 10-20% of HENP researchers are on the poorer side of the Digital Divide.

It is understood that PingER testing has limitations. The use of ICMP and in particular the practice of blocking or rate limiting ICMP packets has concerned the developers. Total filtering of the probes is easy to spot. However, rate limiting is potentially a cause of erroneous results. A study conducted in 2000 indicated that the occurrence of rate limiting between PingER sites was low. The developers work with local contacts and administrators to ensure the pings are not blocked, that they are suitably lightweight, and that the tests are as reasonably as possible an indication of the overall performance to that site.

Although largely unfunded, tremendous progress has been made with the IEPM-BW project. Numerous sites have volunteered resources to join the project and provided access to machines and we now have a valuable set of contacts/collaborators and accounts at over 40 sites. Future development of a funded successor to PingER/BW will have a sound basis of contacts and infrastructure on which to build.

4. Interactions with other projects.

The PingER team works closely with other monitoring projects. SLAC is home to a Surveyor box (defunct), a RIPE-TT box, AMP boxes (IPv4 and IPv6), a NIMI box and a box from the Self-Configuring Network Monitor (SCNM) project. In addition SLAC

gathers passive performance information from Cisco Netflow records recorded at the SLAC boundary.
SLAC is also part of the XIWT IPEX project and is evaluating certain commercial products such as NetPhysics.

During the lifetime of these projects, studies were conducted to correlate the results and understand the different methodologies. Typically results were closely correlated. In addition the use of multiple measurement tools provided extra insight into the performance of the network.

Other interaction with other project includes deriving throughput from packet loss and round trip time (RTT) using the equation of Mathis, Semke, Mahdavi and Ott. Comparison between derived throughput and high impact throughput tests using netperf and iperf indicate good agreement under reasonable conditions.

Work is underway to integrate the SCIDAC funded Available Bandwidth Estimation (ABwE) and pathload tools into the IEPM-BW framework.

The value of the IEPM-PingER/IEPM-BW data gathered and the need to make it available for scientific applications (e.g. for data steering) has prompted the developers to join the Global Grid Forum (GGF) Network Monitoring Working Group (NMWG). In particular the data is being made available by web services and OGSI-compliant grid services using the standardized schema under development by the group.

PingER was ported to IPv6 and provided some early measurements on the largely tunneled 6Bone network and the largely native 6REN network. Recently the PingER-6 monitoring has been revived.

5. Future Work - Elements that should be added.

Experience has brought to light the need to extend IEPM-BW and a new proposal is being prepared. The Measurement and Analysis for the Global Grid and Internet End-to-end performance (MAGGIE) proposal will provide a secure framework with which to test multiple file transfer and throughput testing tools. In particular the NIMI security model utilizing the Akenti software will be developed and integrated with the measurement engine. This framework will allow measurements to be made on demand. Co-ordinated measurements, perhaps using the Network Weather Service (NWS) concept of cliques, will also be enabled. Predictions will also be added allowing physicists to request an estimate of the time it will take to transfer a large file to a selection of remote sites. Results will be published on the web, and also via Web/Grid Services. The ability for separate projects to share results and third-party tools to obtain the data on demand are core concepts. The availability of network performance data via web or grid services will allow steering of grid applications and enable resource brokers to make decisions based on network performance. Currently there is little interest from the developers of resource brokers to make use of this service. In addition algorithms to cross reference different data sets and pinpoint the location of a problem will be developed.

6. Summary

The results described here and in the accompanying presentation reflect the success of the IEPM-PingER project in addressing its original goal of monitoring network performance. This groundbreaking project has had a major impact on the understanding and ability to troubleshoot networks used by the High Energy Physics community. The success of PingER and the extension to a throughput monitoring system has led to a great deal of interaction across projects and communities. A myriad of network performance projects have emerged during the lifetime of PingER. The opportunity to integrate tools and projects within a follow-up project is irresistible.

Moreover, the goals of the larger National Collaboratory and Network Research begs the continuation of the IEPM work. In the summary of the review, it was stated

> All of the National Collaboratory and Network Research projects have specific goals and objectives, but all of you involved in those projects are also part of a much larger, longer term effort, namely creating an infrastructure that will enable geographically separated scientists to effectively work together as a team and that will facilitate remote access to both facilities and data.

Similar sentiments have been expressed in the recent ESnet roadmap workshop and the Large Scale Networking (LSN) meeting. Creation of a cross-domain monitoring infrastructure is an essential ingredient to the success of data grids.

> **"Unfortunately, network management research has historically been very under-funded, because it is difficult to get funding bodies to recognize this as legitimate networking research."**
>
> Sally Floyd
> IAB Concerns & Recommendations Regarding Internet Research & Evolution.
> http://www.ietf.org/internet-drafts/draft-iab-research-funding-00.txt

7. Acknowledgements.

The authors would like to thank the many developers, past and present who have worked so hard to build this project.

> Connie Logg, Jerrod Williams, Jiri Navratil, David Martin, Frank Nagy, Al Thomas, Maxim Grigoriev and Fabrizio Coccetti.

Furthermore we gratefully acknowledge the advisors and experts who so kindly donate their hard won knowledge.

> Brian Tierney, Eric Boyd, Jeff Boote, Matt Zekauskas, Matt Mathis, Russ Hobby, Vern Paxson, Andy Adams, Kc Claffy and Iosif Legrand.

A special thank you to all the administrators and engineers who keep the network running, and especially those who volunteer their spare time to assist with monitoring projects such as IEPM-PingER and IEPM-BW.

Finally a very big thank you for the financial support from DoE/MICS.

8. References.

This document, the accompanying presentation and other IEPM related papers and talks are one line at http://www-iepm.slac.stanford.edu/paperwork

[IEPM] The Internet End-to-end Performance Monitoring homepage is http://www-iepm.slac.stanford.edu

[PingER] The PingER monitoring page is http://www-iepm.slac.stanford.edu/pinger

[IEPM-BW] The BW monitoring page is http://www-iepm.slac.stanford.edu/bw

[MAGGIE] Details of the MAGGIE proposal can be found at http://www-iepm.slac.stanford.edu/maggie

[ICFA-SCIC] The ICFA SCIC page is http://icfa-scic.web.cern.ch/ICFA-SCIC

[XIWT] The Cross-Industry Working Team home page is http://www.xiwt.org

[AMP] Details of the Active Monitoring Project (AMP) can be found at http://moat.nlanr.net

[NIMI] Details of the NIMI project can be found at http://www.ncne.nlanr.net/nimi

[E2E PI] Details of the Internet2 End-to-end Performance Initiative can be found at http://e2epi.internet2.edu

[GGF NMWG] The Global Grid Forums Network Monitoring Workding Group home page is at http://www-didc.lbl.gov/NMWG