



Caltech computer scientists develop FAST protocol to speed up Internet

[Related Links](#)

[FAST](#)

Caltech computer scientists have developed a new data transfer protocol for the Internet fast enough to download a full-length DVD movie in less than five seconds.

The protocol is called FAST, standing for Fast Active queue management Scalable Transmission Control Protocol (TCP). The researchers have achieved a speed of 8,609 megabits per second (Mbps) by using 10 simultaneous flows of data over routed paths, the largest aggregate throughput ever accomplished in such a configuration. More importantly, the FAST protocol sustained this speed using standard packet size, stably over an extended period on shared networks in the presence of background traffic, making it adaptable for deployment on the world's high-speed production networks.

The experiment was performed last November during the Supercomputing Conference in Baltimore, by a team from Caltech and the Stanford Linear Accelerator Center (SLAC), working in partnership with the European Organization for Nuclear Research (CERN), and the organizations DataTAG, StarLight, TeraGrid, Cisco, and Level(3).

The FAST protocol was developed in Caltech's Networking Lab, led by Steven Low, associate professor of computer science and electrical engineering. It is based on theoretical work done in collaboration with John Doyle, a professor of control and dynamical systems, electrical engineering, and bioengineering at Caltech, and Fernando Paganini, associate professor of electrical engineering at UCLA. It builds on work from a growing community of theoreticians interested in building a theoretical foundation of the Internet, an effort in which Caltech has been playing a leading role.

Harvey Newman, a professor of physics at Caltech, said the fast protocol "represents a milestone for science, for grid systems, and for the Internet."

"Rapid and reliable data transport, at speeds of one to 10 Gbps and 100 Gbps in the future, is a key enabler of the global collaborations in physics and other fields," Newman said. "The ability to extract, transport, analyze and share many Terabyte-scale data collections is at the heart of the process of search and discovery for new scientific knowledge. The FAST results show that the high degree of transparency and performance of networks, assumed implicitly by Grid systems, can be achieved in practice. In a broader context, the fact that 10 Gbps wavelengths can be used efficiently to transport data at maximum speed end to end will transform the future concepts of the Internet."

Les Cottrell of SLAC, added that progress in speeding up data transfers over long distance are critical to progress in various scientific endeavors. "These include sciences such as high-energy physics and nuclear physics, astronomy, global weather predictions, biology, seismology, and fusion; and industries such as aerospace, medicine, and media distribution.

"Today, these activities often are forced to share their data using literally truck or plane loads of data," Cottrell said. "Utilizing the network can dramatically reduce the delays and automate today's labor intensive procedures."

The ability to demonstrate efficient high performance throughput using commercial off the shelf hardware and applications, standard Internet packet sizes supported throughput today's networks, and requiring modifications to the ubiquitous TCP protocol only at the data sender, is an important achievement.

With Internet speeds doubling roughly annually, we can expect the performances demonstrated by this collaboration to become commonly available in the next few years, so the demonstration is important to set expectations, for planning, and to indicate how to utilize such speeds.

The testbed used in the Caltech/SLAC experiment was the culmination of a multi-year effort, led by Caltech physicist Harvey Newman's group on behalf of the international high energy and nuclear physics (HENP) community, together with CERN, SLAC, Caltech Center for Advanced Computing Research (CACR), and other organizations. It illustrates the difficulty, ingenuity and importance of organizing and implementing leading edge global experiments. HENP is one of the principal drivers and co-developers of global research networks. One unique aspect of the HENP

testbed is the close coupling between R&D and production, where the protocols and methods implemented in each R&D cycle are targeted, after a relatively short time delay, for widespread deployment across production networks to meet the demanding needs of data intensive science.

The congestion control algorithm of the current Internet was designed in 1988 when the Internet could barely carry a single uncompressed voice call. The problem today is that this algorithm cannot scale to anticipated future needs, when the networks will be compelled to carry millions of uncompressed voice calls on a single path or support major science experiments that require the on-demand rapid transport of gigabyte to terabyte data sets drawn from multi-petabyte data stores. This protocol problem has prompted several interim remedies, such as using nonstandard packet sizes or aggressive algorithms that can monopolize network resources to the detriment of other users. Despite years of effort, these measures have proved to be ineffective or difficult to deploy.

They are, however, critical steps in our evolution toward ultrascale networks. Sustaining high performance on a global network is extremely challenging and requires concerted advances in both hardware and protocols. Experiments that achieve high throughput either in isolated environments or using interim remedies that by-pass protocol instability, idealized or fragile as they may be, push the state of the art in hardware and demonstrates its performance limit. Development of robust and practical protocols will then allow us to make effective use of the most advanced hardware to achieve ideal performance in realistic environments.

The FAST team addresses the protocol issues head-on to develop a variant of TCP that can scale to a multi-gigabit-per-second regime in practical network conditions. The integrated approach that combines theory, implementation, and experiment is what makes their research unique and fundamental progress possible.

Using standard packet size that is supported throughout today's networks, the current TCP typically achieves an average throughput of 266 Mbps, averaged over an hour, with a single TCP/IP flow between Sunnyvale near SLAC and CERN in Geneva, over a distance of 10,037 kilometers. This represents an efficiency of just 27 percent. The FAST TCP sustained an average throughput of 925 Mbps and an efficiency of 95 percent, a 3.5-times improvement, under the same experimental condition. With 10 concurrent TCP/IP flows, FAST achieved an unprecedented speed of 8,609 Mbps, at 88 percent efficiency, that is 153,000 times that of today's modem and close to 6,000 times that of the common standard for ADSL (Asymmetric Digital Subscriber Line) connections.

The 10-flow experiment sets another first in addition to the highest aggregate speed over routed paths. It is the combination of high capacity and large distance that causes performance problems. Different TCP algorithms can be compared using the product of achieved throughput and the distance of transfer, measured in bit-meter-per-second, or bmpps. The world record for the current TCP is 10 peta (1 followed by 16 zeros) bmpps, using a nonstandard packet size. The Caltech/SLAC experiment transferred 21 terabytes over six hours between Baltimore and Sunnyvale using standard packet size, achieving 34 peta bmpps. Moreover, data was transferred over shared research networks in the presence of background traffic, suggesting that FAST can be backward compatible with the current protocol. The FAST team has started to work with various groups around the world to explore testing and deploying FAST TCP in communities that need multi-Gbps networking urgently.

The demonstrations used a 10 Gbps link donated by Level(3) between StarLight (Chicago) and Sunnyvale, as well as the DataTAG 2.5 Gbps link between StarLight and CERN, the Abilene backbone of Internet2, and the TeraGrid facility. The network routers and switches at StarLight and CERN were used together with a GSR 12406 router loaned by Cisco at Sunnyvale, additional Cisco modules loaned at StarLight, and sets of dual Pentium 4 servers each with dual Gigabit Ethernet connections at StarLight, Sunnyvale, CERN, and the SC2002 show floor provided by Caltech, SLAC, and CERN. The project is funded by the National Science Foundation, the Department of Energy, the European Commission, and the Caltech Lee Center for Advanced Networking.

One of the drivers of these developments has been the HENP community, whose explorations at the high-energy frontier are breaking new ground in our understanding of the fundamental interactions, structures and symmetries that govern the nature of matter and space-time in our universe. The largest HENP projects each encompasses 2,000 physicists from 150 universities and laboratories in more than 30 countries.

Rapid and reliable data transport, at speeds of 1 to 10 Gbps and 100 Gbps in the future, is a key enabler of the global collaborations in physics and other fields. The ability to analyze and share many terabyte-scale data collections, accessed and transported in minutes, on the fly, rather than over hours or days as is the current practice, is at the heart of the process of search and discovery for new scientific knowledge. Caltech's FAST protocol shows that the high degree of transparency and performance of networks, assumed implicitly by Grid systems, can be achieved in practice.

This will drive scientific discovery and utilize the world's growing bandwidth capacity much more efficiently than has been possible until now.

Navigating Tools



Search for: in

Email us at prmedia@caltech.edu with comments, questions or suggestions.