A 100-gigbit highway for science

April 30 2012

Climate researchers are producing some of the fastest growing datasets in science. Five years ago, the amount of information generated for the Nobel Prize-winning United Nations International Panel on Climate Change (IPCC) Fourth Assessment Report was 35 terabytes—equivalent to the amount of text in 35 million books, occupying a bookshelf 248 miles (399 km) long. By 2014, when the next IPCC report is published, experts predict that 2 petabytes of data will have been generated for it—that's a 580 percent increase in data production.

Because thousands of researchers around the world contribute to the generation and analysis of this data, a reliable, high-speed [network](#) is needed to transport the torrent of information. Fortunately, the Department of Energy's (DOE) ESnet (Energy Sciences Network) has laid the foundation for such a network—not just for climate research, but for all data-intensive science.

"There is a data revolution occurring in science," says Greg Bell, acting director of ESnet, which is managed by Lawrence Berkeley National Laboratory. "Over the last decade, the amount of scientific data transferred over our network has increased at a rate of about 72 percent per year, and we see that trend potentially accelerating."

In an effort to spur U.S. scientific competitiveness, as well as accelerate development and widespread deployment of 100-gigabit technology, the Advanced Networking Initiative (ANI) was created with $62 million in funding from the American Recovery and Reinvestment Act (ARRA) and implemented by ESnet. ANI was established to build a 100 Gbps

national prototype network and a wide-area network testbed.

To cost-effectively deploy ANI, ESnet partnered with Internet2—a consortium that provides high-performance network connections to universities across America—which also received a stimulus grant from the Department of Commerce's Broadband Technologies Opportunities Program.

Researchers Take a "Test Drive" on ANI

So far more than 25 groups have taken advantage of ESnet's wide-area testbed, which is open to researchers from government agencies and private industry to test new, potentially disruptive technologies without interfering with production science network traffic. The testbed currently connects three unclassified DOE supercomputing facilities: the National Energy Research Scientific Computing Center (NERSC) in Oakland, Calif., the Argonne Leadership Computing Facility (ALCF) in Argonne, Ill., and the Oak Ridge Leadership Computing Facility (OLCF) in Oak Ridge, Tenn.

"No other networking organization has a 100-gigabit network testbed that is available to researchers in this way," says Brian Tierney, who heads ESnet's Advanced Networking Technologies Group. "Our 100G testbed has been about 80 percent booked since it became available in January, which just goes to show that there are a lot of researchers hungry for a resource like this."

Climate 100

To ensure that researchers will use future 100-gigabit effectively, another ARRA-funded project called Climate 100 brought together middleware and network engineers to develop tools and techniques for

moving unprecedentedly massive amounts of climate data.

"Increasing network bandwidth is an important step toward tackling ever-growing scientific [datasets](#), but it is not sufficient by itself; next-generation high-bandwidth networks need to be evaluated carefully from the applications perspective as well," says Mehmet Balman of Berkeley Lab's Scientific Data Management group, a member of the Climate 100 collaboration.

According to Balman, climate simulation data consists of a mix of relatively small and large files with irregular file size distribution in each dataset. This requires advanced middleware tools to move data efficiently on long-distance high-bandwidth networks.

"The ANI testbed essentially allowed us to 'test drive' on a 100-gigabit network to determine what kind of middleware tools we needed to build to transport climate data," says Balman. "Once the development was done, we used the testbed to optimize and tune."

At the 2011 Supercomputing Conference in Seattle, Wash., the Climate 100 team used their tool and the ANI testbed to transport 35 terabytes of climate data from NERSC's data storage to compute nodes at ALCF and OLCF.

"It took us approximately 30 minutes to move 35 terabytes of climate data over a wide-area 100 Gbps network. This is a great accomplishment," says Balman. "On a 10 Gbps network, it would have taken five hours to move this **much data across the country.**"

Space Exploration

In 2024, the most powerful radio telescope ever constructed will go online. Comprising 3,000 satellite dishes spread over 250 acres, this

instrument will generate more data in a single day than the entire Internet carries today. Optical fibers will connect each of these 15-meter-wide (50 ft.) satellite dishes to a central high performance computing system, which will combine all of the signals to create a detailed "big picture."

"Given the immense sensor payload, optical fiber interconnects are critical both at the central site and from remote stations to a single correlation facility," says William Ivancic, a senior research engineer at NASA's Glenn Research Center. "Future radio astronomy networks need to incorporate next generation network technologies like 100 Gbps long-range Ethernet links, or better, into their designs."

In anticipation of these future networks, Ivancic and his colleagues are utilizing a popular high-speed transfer protocol, called Saratoga, to effectively carry data over 100-gigabit long-range Ethernet links. But because it was cost-prohibitive to upgrade their local network with 100-gigabit hardware, the team could not determine how their software would perform in a real-world scenario—that is, until they got access to the ANI testbed.

"Quite frankly, we would not be doing these speed tests without the ANI testbed," says David Stewart, an engineer at Verizon Federal Systems and Ivancic's colleague. "We are currently in the development and debugging phase, and have several implementations of our code. With the ANI testbed, we were able to optimize and scale our basic PERL implementation to far higher speeds than our NASA testbed."

End-to-End Delivery

Meanwhile, Dantong Yu, who leads the Computer Science Group at Brookhaven National Laboratory, used the ANI testbed to design an ultra-high-speed, end-to-end file transfer protocol tool to move science data at 100 gigabits per second across a national network.

"A network like ANI may be able to move data at 100 Gbps, but at each end of that connection there is a host server that either uploads or downloads data from the network," says Yu. "While the host servers may be capable of feeding data into the network and downloading it at 100 Gbps, the current software running on these systems is a bottleneck."

According to Yu, the bottlenecks are primarily caused by the number of times the current software forces the computer to make copies of the data before uploading it to the network.

"Initially I was testing this protocol at a very local lab level. In this scenario transfers happen in a split-second, which is far from reality," says Yu. "ANI allowed me to see how long it really takes to move data across the country, from East-to West Coast, with my software, which in turn helped me optimize the code."

The Next Steps

Within the next few months, the official ANI project will be coming to an end, but the community will continue to benefit for decades to come from its investments. The 100-gigabit prototype network will be converted into ESnet's fifth-generation production infrastructure, one that will be scale to 44 times its current. ESnet will also seek new sources of funding for the 100-gigabit testbed to ensure that it will be available to network researchers on a sustained basis.

"Since its inception, ESnet has delivered the advanced capabilities required by DOE science. Many of these capabilities are cost-prohibitive, or simply unavailable, on the commercial market," says Bell. "Because our network is optimized for the needs of DOE science, we're always looking for efficient ways to manage our large science flows. ESnet's new 100-Gigabit network will allow us to do that more flexibly and morecost-effectively than ever."

Provided by Lawrence Berkeley National Laboratory