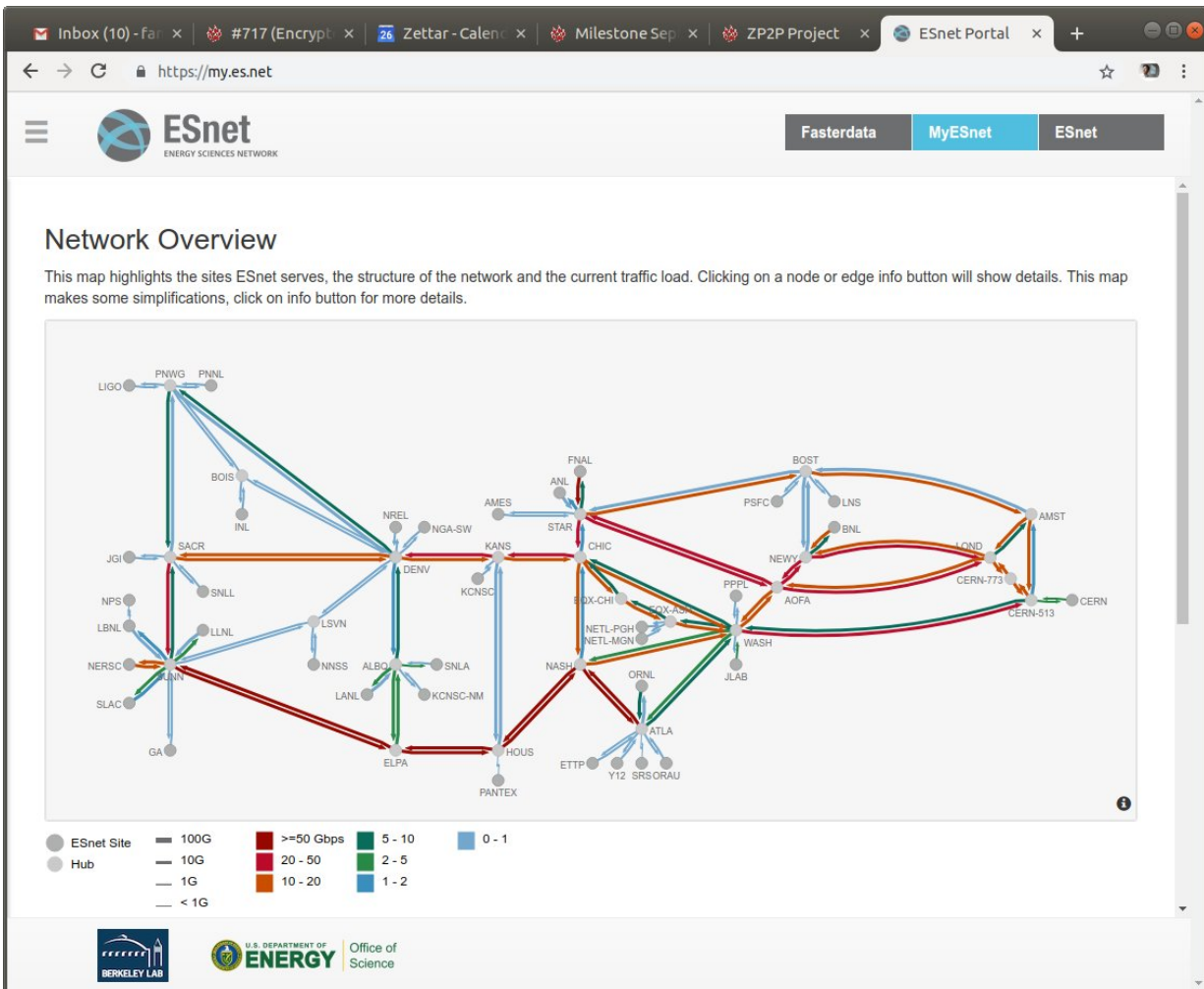


Record-setting transfer of 1 petabyte of data

October 23 2018, by Jon Bashor



This screen shot from MyESnet shows the southern links of the network dark red, indicating data transfers exceeding 50 Gbps during the 5,000-mile transfer conducted by SLAC and Zettar Inc.

Using a 5,000-mile network loop operated by ESnet, researchers at the Department of Energy's SLAC National Accelerator Laboratory (SLAC) and Zettar Inc. recently transferred 1 petabyte in 29 hours, with encryption and checksumming, beating last year's record by 5 hours, an almost 15 percent improvement.

The project is aimed at achieving the high data [transfer](#) rates needed to accommodate the amount of data to be generated by the Linac Coherent Light Source II (LCLS II), which is expected to come online in 2020. The LCLS is the world's first hard X-ray free-electron laser (XFEL) and its strobe-like pulses are just a few millionths of a billionth of a second long, and a billion times brighter than previous X-ray sources. LCLS II will provide a major jump in capability – moving from 120 pulses per second to 1 million pulses per second. Scientists use LCLS to take crisp pictures of atomic motions, watch chemical reactions unfold, probe the properties of materials and explore fundamental processes in living things.

The increased capability is expected to generate data transfers of multiple terabits per second— as the experimental results are sent from SLAC to Department of Energy's (DOE) supercomputing facilities at Lawrence Berkeley and Oak Ridge national laboratories for analysis. As the DOE's dedicated network user facility for scientific research, ESnet carries data between universities and DOE's national laboratories and national user facilities along a multi-100 gigabits-per-second (Gbps) backbone network.

According to Stanford Emeritus Les Cottrell, the current demonstration of [high speed data](#) transport is part of a program to develop and evaluate tools to address the upcoming decade's critical data transfer requirements for the next-generation XFELs at SLAC.



Graph shows that the petabyte transfer trial accounted for one-third of the total network traffic on ESnet during the 29-hour run. Credit: US Department of Energy

Zettar, a National Science Foundation-funded software firm in Palo Alto, develops hyperscale data distribution software solution capable of multi-100+Gbps transfer rates and collaborates with ESnet and DOE national labs. For the trial, the team used ESnet's On-Demand Secure Circuits and Advance Reservation System (OSCARS) to reserve the bandwidth for the run.

"The innovative and versatile nature of ESnet's OSCARS enabled us to design a unique 5000-mile 100 Gbps loop, which our ESnet colleagues quickly helped establish in 2015," said Zettar Founder and CEO Chin Fang. "The design of ESnet enables the use of a separate dedicated channel for the testing without impact to or interference from production traffic. We were thus able to have the data storage clusters and the data transfer nodes co-located in the same rack although they

were at both ends of the loop. Physically they were separated by inches, but digitally they are separated by 5000 miles. Further the clusters themselves have a self-imposed cap of 80 Gbps."

Fang said the five-hour improvement in transferring the petabyte of data was accomplished thanks to improvements in both software and hardware. This year, Zettar significantly improved the parallelism of its software to speed up data transfers for both encrypted and non-encrypted data. Intel also provided the team with faster processors and solid-state memory products for moving data in and out of main storage systems faster. In this case, the project used simulated datasets with file size distributions provided by the LCLS Data Management Group.

"The flexible traffic engineering design of ESnet's network ensures that TCP-based transfer protocols can be used effectively for high-throughput data movement," said ESnet Director Inder Monga. "ESnet's guiding strategy is to anticipate and stay ahead of the [data transfer](#) requirements of our scientific users. Ongoing trials like the ones we are supporting for SLAC are important for us in that we see the directions that the research teams are heading and we can enhance our support and infrastructure to be ready when facilities like LCLS II come online."

Provided by US Department of Energy

Citation: Record-setting transfer of 1 petabyte of data (2018, October 23) retrieved 27 April 2024 from <https://phys.org/news/2018-10-record-setting-petabyte.html>

<p>This document is subject to copyright. Apart from any fair dealing for the purpose of private study or research, no part may be reproduced without the written permission. The content is provided for information purposes only.</p>
--