Edison supercomputer electrifies scientific computing

January 30 2014

Named in honor of American inventor Thomas Alva Edison, Edison is a Cray XC30 supercomputer that will serve 5,000 researchers at the National Energy Research Scientific Computing Center, the Department of Energy's most productive scientific computing facility. Edison is unique among scientific supercomputers for its ability to handle data and simulation with equal ease and a design and deployment centered on scientific productivity. Credit: Lawrence Berkeley National Laboratory

The National Energy Research Scientific Computing (NERSC) Center recently accepted "Edison," a new flagship supercomputer designed for scientific productivity. Named in honor of American inventor Thomas Alva Edison, the Cray XC30 will be dedicated in a ceremony held at the Department of Energy's Lawrence Berkeley National Laboratory (Berkeley Lab) on Feb. 5, and scientists are already reporting results.

About 5,000 researchers working on 700 projects and running 600 different codes compute at NERSC, which is operated by Berkeley Lab. They produce an average of 1,700 peer-reviewed publications every year, making NERSC the most productive scientific computing center serving the Department of Energy's Office of Science.

"We support a very broad range of science, from basic energy research to climate science, from biosciences to discovering new materials, exploring high energy physics and even uncovering the very origins of the universe," said NERSC Director Sudip Dosanjh.

Edison can execute nearly 2.4 quadrillion floating-point operations per second (petaflop/s) at peak theoretical speeds. While theoretical speeds are impressive, "NERSC's longstanding approach is to evaluate proposed systems by how well they meet the needs of our diverse community of researchers, so we focus on sustained performance on real applications," said NERSC Division Deputy for Operations Jeff Broughton, who led the Edison procurement team.

"For us, what's really important is the scientific productivity of our users," Dosanjh said. That's why Edison was configured to handle two kinds of computing equally well: data analysis and simulation and modeling.

Traditionally, scientific supercomputers are configured to simulate and model complex phenomena, such as nanomaterials converting electricity into photons of light, climate changing over decades or centuries, or interstellar gases forming into stars and galaxies. Simulations require a lot of processors running in unison, but not necessarily a lot of memory for each processor.

Data analysis, such as genome sequencing or molecular screening programs that search for promising new materials or drugs, often involves high throughput computing—running large numbers of loosely coupled simulations simultaneously. Such "ensemble computing" requires more memory per node and has typically been relegated to separate computer clusters. As instruments and experiments deliver more and more data however, scientists need more computing power to crunch it; so smaller clusters no longer suffice.

Jeff Broughton, NERSC deputy for operations and leader of the Edison procurement; Kathy Yelick, associate laboratory director for computing sciences at Berkeley Lab; and Sudip Dosanjh, NERSC division director, pose with Edison, NERSC's new flagship supercomputer.

"Facilities throughout the Department of Energy are being inundated with data that researchers don't have the ability to understand, process or analyze sufficiently," said Dosanjh. Historically, NERSC was an exporter of data as scientists ran large-scale simulations and then moved that data to other sites. But with the growth of experimental data coming from other sites, NERSC is now a net importer, taking in a petabyte of data in fields such as biosciences, climate and high-energy physics each month.

Both types of computing rely heavily on moving data, said Dosanjh. "So Edison has been optimized for that: It has a really high-speed interconnect, it has lots of memory bandwidth, lots of memory per node, and it has very high input/output speeds to the file system and disk system."

"If you have a computing resource like Edison, one with the flexibility to run different classes of problems, then you can apply the full capacity of your system to the problem at hand, whether that be high-throughput genome sequencing or highly parallel climate simulations," said Broughton.

Because Edison does not employ accelerators, such as graphics processing units (GPUs), scientists have been able to move their codes from NERSC's previous flagship system (a Cray XE6 named for computer scientist Grace Hopper) to Edison with little or no changes, another consideration meant to keep scientists doing science instead of rewriting code.

"We were able to open Edison to all our users shortly after installation for testing, and the system was immediately full," said Broughton. By the time Edison was accepted and placed into production, scientists had logged millions of processor hours of research into areas as varied as carbon sequestration, nanomaterials, cosmology, and combustion.

And while researchers may not see or appreciate Edison's advances in energy efficiency, it will impact their ability to do science. "In coming years, performance will be more limited by power than anything else, so energy efficiency is critical," said Dosanjh.

In preparation for its 2015 move into a custom-built data center (the Computational Research and Theory facility), Edison is the first supercomputer at NERSC to rely solely on outside air for cooling, a technique known as "free cooling." Edison is cooled without mechanical chillers. Instead water is circulated through outdoor cooling towers and back into the system's internal radiators, which cool air rather than heat it. Fans located between each pair of cabinets in a row pull air in one end; circulate it through a radiator, over the hot components and on to the next set of cabinets before it exits at the row's end. This side-to-side airflow, or transverse cooling, is more energy efficient than the typical front-to-back flow of most systems.

Edison will be dedicated as part of the annual NERSC Users Group being held February 3-6 at Berkeley Lab (See http://www.nersc.gov/NUG14 for more information.)

"As we celebrate NERSC's 40th anniversary, it's quite fitting we start the year by dedicating Edison, a system that embodies our guiding principle over the last four decades: computing in the service of science," said NERSC director Dosanjh."

Edison by the Numbers

- 124,608 processing cores
- 332 terabytes memory
- 2.39 petaflop/second peak performance
- 462 terabytes/second global memory bandwidth
- 11 terabytes/second network bisection bandwidth
- 7.56 petabytes disk storage
- 163 gigabytes/second I/O bandwidth

Provided by Lawrence Berkeley National Laboratory

Citation: Edison supercomputer electrifies scientific computing (2014, January 30) retrieved 19 April 2024 from https://phys.org/news/2014-01-edison-supercomputer-electrifies-scientific.html

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.