

# NERSC, Cray move forward with next-generation scientific computing

22 April 2015



A new Cray XC40 will be first supercomputer in Berkeley Lab's new Computational Research and Theory facility. Credit: Photo by Roy Kaltschmidt, Berkeley Lab

The U.S. Department of Energy's (DOE) National Energy Research Scientific Computing (NERSC) Center and Cray Inc. announced today that they have finalized a new contract for a Cray XC40 supercomputer that will be the first NERSC system installed in the newly built Computational Research and Theory facility at Lawrence Berkeley National Laboratory.

This [supercomputer](#) will be used as Phase 1 of NERSC's next-generation system named "Cori" in honor of bio-chemist and Nobel Laureate Gerty Cori. Expected to be delivered this summer, the Cray XC40 supercomputer will feature the Intel Haswell processor. The second phase, the previously announced Cori system, will be delivered in mid-2016 and will feature the next-generation Intel Xeon Phi processor "Knights Landing," a self-hosted, manycore processor with on-package high bandwidth memory that offers more than 3 teraflop/s of double-precision peak performance per single socket node.

NERSC serves as the primary high performance computing facility for the Department of Energy's Office of Science, supporting some 6,000 scientists annually on more than 700 projects. This latest contract represents the Office of Science's ongoing commitment to supporting computing to address challenges such as developing new energy sources, improving energy efficiency, understanding climate change and analyzing massive data sets from observations and experimental facilities around the world.

"This is an exciting year for NERSC and for NERSC users," said Sudip Dosanjh, director of NERSC. "We are unveiling a brand new, state-of-the-art computing center and our next-generation supercomputer, designed to help our users begin the transition to exascale computing. Cori will allow our users to take their science to a level beyond what our current systems can do."

"NERSC and Cray share a common vision around the convergence of supercomputing and big data, and Cori will embody that overarching technical direction with a number of unique, new technologies," said Peter Ungaro, president and CEO of Cray. "We are honored that the first supercomputer in NERSC's new center will be our flagship Cray XC40 system, and we are also proud to be continuing and expanding our longstanding partnership with NERSC and the U.S. Department of Energy as we chart our course to exascale computing."

## Support for Data-Intensive Science

A key goal of the Cori Phase 1 system is to support the increasingly data-intensive computing needs of NERSC users. Toward this end, Phase 1 of Cori will feature more than 1,400 Intel Haswell compute nodes, each with 128 gigabytes of memory per node. The system will provide about the same sustained application performance as NERSC's Hopper system, which will be retired later this year.

The Cori interconnect will have a dragonfly topology based on the Aries interconnect, identical to NERSC's Edison system.

However, Cori Phase 1 will have twice as much memory per node than NERSC's current Edison supercomputer (a Cray XC30 system) and will include a number of advanced features designed to accelerate data-intensive applications:

- Large number of login/interactive nodes to support applications with advanced workflows
- Immediate access queues for jobs requiring real-time data ingestion or analysis
- High-throughput and serial queues can handle a large number of jobs for screening, uncertainty qualification, genomic data processing, image processing and similar parallel analysis
- Network connectivity that allows compute nodes to interact with external databases and workflow controllers
- The first half of an approximately 1.5 terabytes/sec NVRAM-based Burst Buffer for high bandwidth low-latency I/O
- A Cray Lustre-based file system with over 28 petabytes of capacity and 700 gigabytes/second I/O bandwidth

In addition, NERSC is collaborating with Cray on two ongoing R&D efforts to maximize Cori's data potential by enabling higher bandwidth transfers in and out of the compute node, high-transaction rate data base access, and Linux container virtualization functionality on Cray compute nodes to allow custom software stack deployment.

"The goal is to give users as familiar a system as possible, while also allowing them the flexibility to explore new workflows and paths to computation," said Jay Srinivasan, the Computational Systems Group lead. "The Phase 1 system is designed to enable users to start running their workload on Cori immediately, while giving data-intensive workloads from other NERSC systems the ability to run on a Cray platform."

### **Burst Buffer Enhances I/O**

A key element of Cori Phase 1 is Cray's new DataWarp technology, which accelerates application I/O and addresses the growing performance gap between compute resources and disk-based storage. This capability, often referred to as a "Burst Buffer," is a layer of NVRAM designed to move data more quickly between processor and disk and allow users to make the most efficient use of the system. Cori Phase 1 will feature approximately 750 terabytes of capacity and approximately 750 gigabytes/second of I/O bandwidth. NERSC, Sandia and Los Alamos national laboratories and Cray are collaborating to define use cases and test early software that will provide the following capabilities:

- Improve application reliability (checkpoint-restart)
- Accelerate application I/O performance for small blocksize I/O and analysis files
- Enhance quality of service by providing dedicated I/O acceleration resources
- Provide fast temporary storage for out-of-core applications
- Serve as a staging area for jobs requiring large input files or persistent fast storage between coupled simulations
- Support post-processing analysis of large simulation data as well as in situ and in transit visualization and analysis using the Burst Buffer nodes

### **Combining Extreme Scale Data Analysis and HPC on the Road to Exascale**

As previously announced, Phase 2 of Cori will be delivered in mid-2016 and will be combined with Phase 1 on the same high speed network, providing a unique resource. When fully deployed, Cori will contain more than 9,300 Knights Landing compute nodes and more than 1,900 Haswell nodes, along with the file system and a 2X increase in the applications I/O acceleration.

"In the scientific computing community, the line between large scale data analysis and simulation and modeling is really very blurred," said Katie Antypas, head of NERSC's Scientific Computing and Data Services Department. "The combined Cori system is the first system to be specifically

designed to handle the full spectrum of computational needs of DOE researchers, as well as emerging needs in which data- and compute-intensive work are part of a single workflow. For example, a scientist will be able to run a simulation on the highly parallel Knights Landing nodes while simultaneously performing data analysis using the Burst Buffer on the Haswell nodes. This is a model that we expect to be important on exascale-era machines."

Provided by Lawrence Berkeley National Laboratory

APA citation: NERSC, Cray move forward with next-generation scientific computing (2015, April 22) retrieved 24 May 2019 from <https://phys.org/news/2015-04-nersc-cray-next-generation-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.*