NERSC

National Energy Research Scientific Computing Center

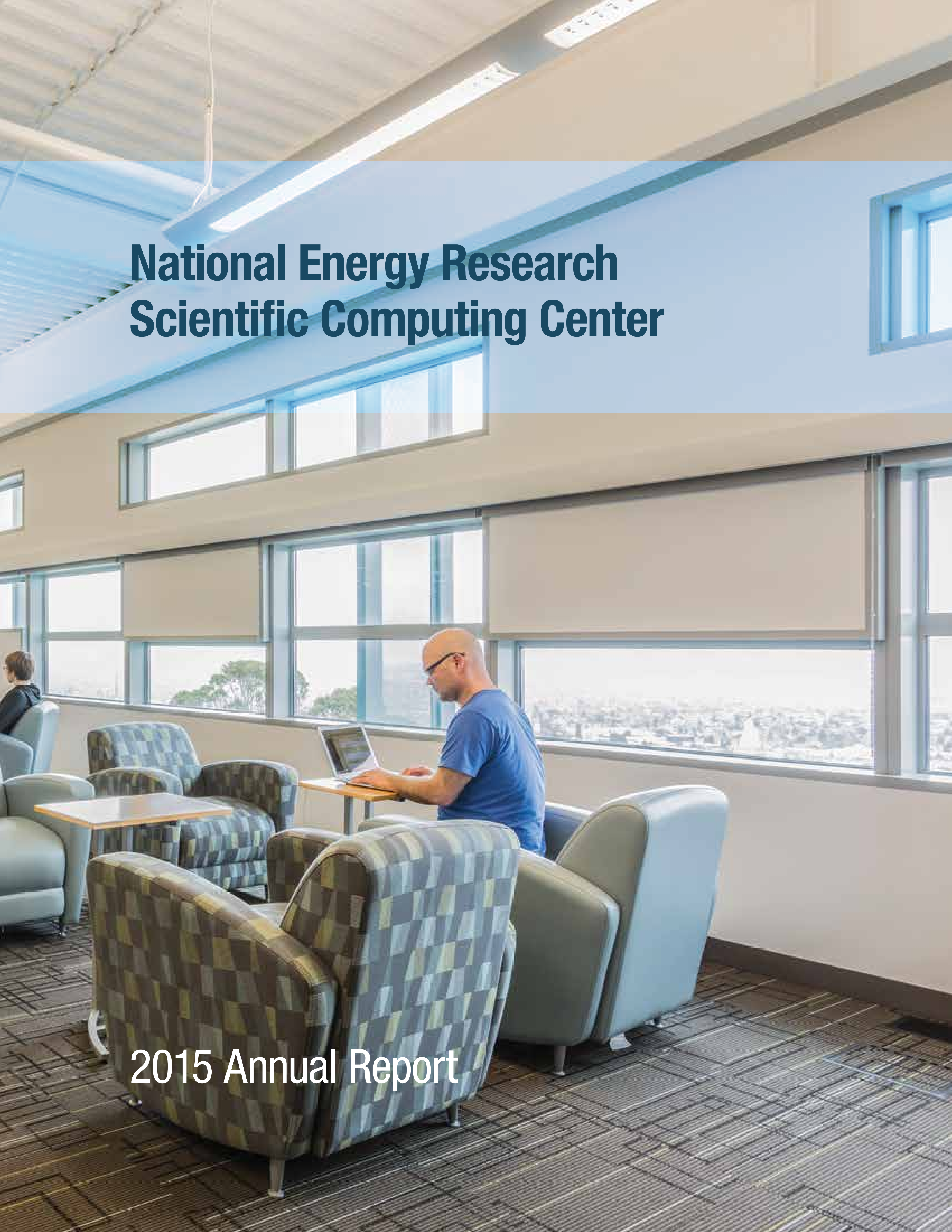
2015 Annual Report

Ernest Orlando Lawrence Berkeley National Laboratory
1 Cyclotron Road, Berkeley, CA 94720-8148

This work was supported by the Director, Office of Science, Office of Advanced Scientific Computing Research of the U.S. Department of Energy under Contract No. DE-AC02-05CH11231

Image Credits:
FRONT COVER: Roy Kaltschmidt, Lawrence Berkeley National Laboratory
BACK COVER (TOP TO BOTTOM): Cornford et al., *The Cryosphere*, 2015 (story p. 35); Scott French, Lawrence Berkeley National Laboratory (story p. 31); *Journal of Materials Research* (story p. 34); Ben Bowen, Lawrence Berkeley National Laboratory (story p. 44); Nicholas P. Brawand, University of Chicago (story p. 32)
THIS PAGE: © Emily Hagopian/Courtesy Perkins + Will

National Energy Research Scientific Computing Center



2015 Annual Report

Table of Contents

Director's Note: A Year of Transitions

2015 was a year of transition and change for NERSC on many levels. The most pronounced change was the center's move from the Oakland Scientific Facility in downtown Oakland to the new Shyh Wang Hall (also known internally as the Computational Research and Theory facility, or CRT) on the main Berkeley Lab campus. A significant amount of staff time was spent planning and implementing the move of computational systems, file systems and auxiliary systems while minimizing the impact on NERSC users and their ongoing projects.

Indeed, one of NERSC's greatest challenges in 2015 was running multiple systems across two different facilities with a constant number of staff. Despite having to physically move NERSC's largest system, Edison, and all file systems in November 2015, NERSC was able to deliver more than the committed 3 billion hours to its users—the same amount committed to in 2014 without the move. This was accomplished by keeping the Hopper system in production until Cori Phase 1 was stable.

Wang Hall itself brings many new changes. The building was designed to be extraordinarily energy efficient and expandable to meet the needs of next-generation HPC systems. It features 12.5 megawatts of power, upgradable to over 40 megawatts, a PUE (power usage effectiveness) of less than 1.1, ambient "free" cooling and the ability to reclaim heat from the computers to heat the building. The new machine room currently measures 20,000 square feet and features a novel seismic isolation floor to protect NERSC's many valuable resources.

Another big change in 2015 was the retirement of two of NERSC's legacy systems, Carver and Hopper. This was done in conjunction with the deployment of the first phase of Cori and the Edison move. As of early 2016, the transfer of systems into the new building was substantially complete and was accomplished with a minimum of downtime. The process was facilitated by a first-of-a-kind 400 Gbs link, provided by ESnet between Wang Hall and the Oakland Scientific Facility, that enabled moves of data to be effectively transparent to users.

Another major undertaking for NERSC in 2015 was the NERSC Exascale Science Applications Program (NESAP). We kicked off the project with 20 applications teams, plus an additional 20 teams participating in training and general optimization efforts. Through NESAP, Cray and Intel staff are able to work directly with code teams to optimize applications and learn about using Cray and Intel performance tools, which was unique because in most cases no such relationship previously existed.

2015 was also the year that NERSC decided there was enough momentum around HPC and data-intensive science that we should form a department focused on data. The new department has four groups, two of them new: a Data Science Engagement group that engages with users of experimental and observational facilities, and an Infrastructure Services group that focuses on infrastructure to support NERSC and our data initiative. The existing Storage Systems Group and Data and Analytics Group complete the new department.

In parallel with these changes, NERSC expanded its workforce in 2015, hiring 18 additional staff members. Our new staff have brought new perspectives to NERSC, and they are playing a critical role in our strategic initiatives. New positions became available due to a large number of retirements in 2014 and 2015, hiring of postdocs into NESAP and normal turnover and churn from competition with Bay Area technology companies.

Also in 2015, NERSC reinvigorated its Private Sector Partnership program, bringing on nine new industry partners in energy-specific fields such as semiconductor manufacturing and geothermal modeling. NERSC has been working with industry partners through the DOE's Small Business Innovation Research (SBIR) grant program for many years. Since 2011, industry researchers have been awarded 40 million hours at NERSC via the SBIR program, and NERSC currently has about 130 industry users from 50 companies.

All of these changes allowed us to continue to address our primary mission: to accelerate scientific discovery for the DOE Office of Science through HPC and data analysis. In 2015, NERSC users reported more than 2,000 refereed papers based on work performed at the center and NERSC staff contributed some 120 papers to scientific journals and conferences, showcasing the center's increasing involvement in technology and software development designed to enhance the utilization of HPC across a broad range of science and data-management applications.

Looking back on an exciting year, we are especially proud of what we were able to accomplish given the many transitions we experienced. NERSC is regarded as a well-run scientific computing facility providing some of the largest computing and storage systems available anywhere, but what really distinguishes the center is its success in creating an environment that makes these resources effective for scientific research and productivity. We couldn't achieve this without our dedicated staff, our great user community and the continued support of our DOE Office of Science sponsors.

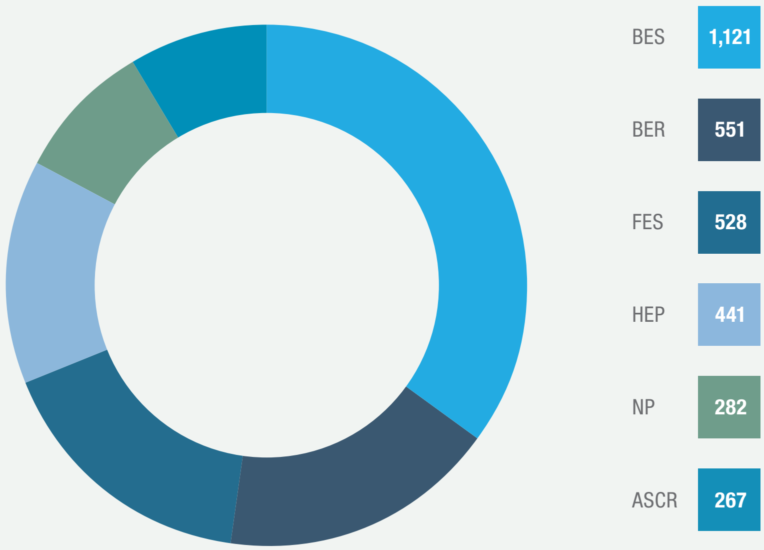
Sudip Dosanjh
NERSC Division Director

NERSC User/Usage Demographics

In 2015, the NERSC Program supported 6,000 active users from universities, national laboratories and industry who used approximately 3 billion supercomputing hours. Our users came from across the United States, with 48 states represented, as well as from around the globe, with 47 countries represented.

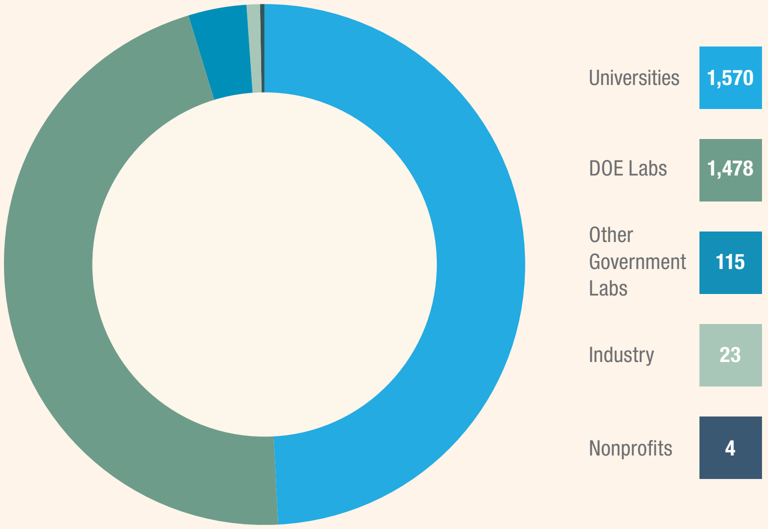
2015 NERSC Usage by DOE Program Office

(MPP HOURS IN MILLIONS)



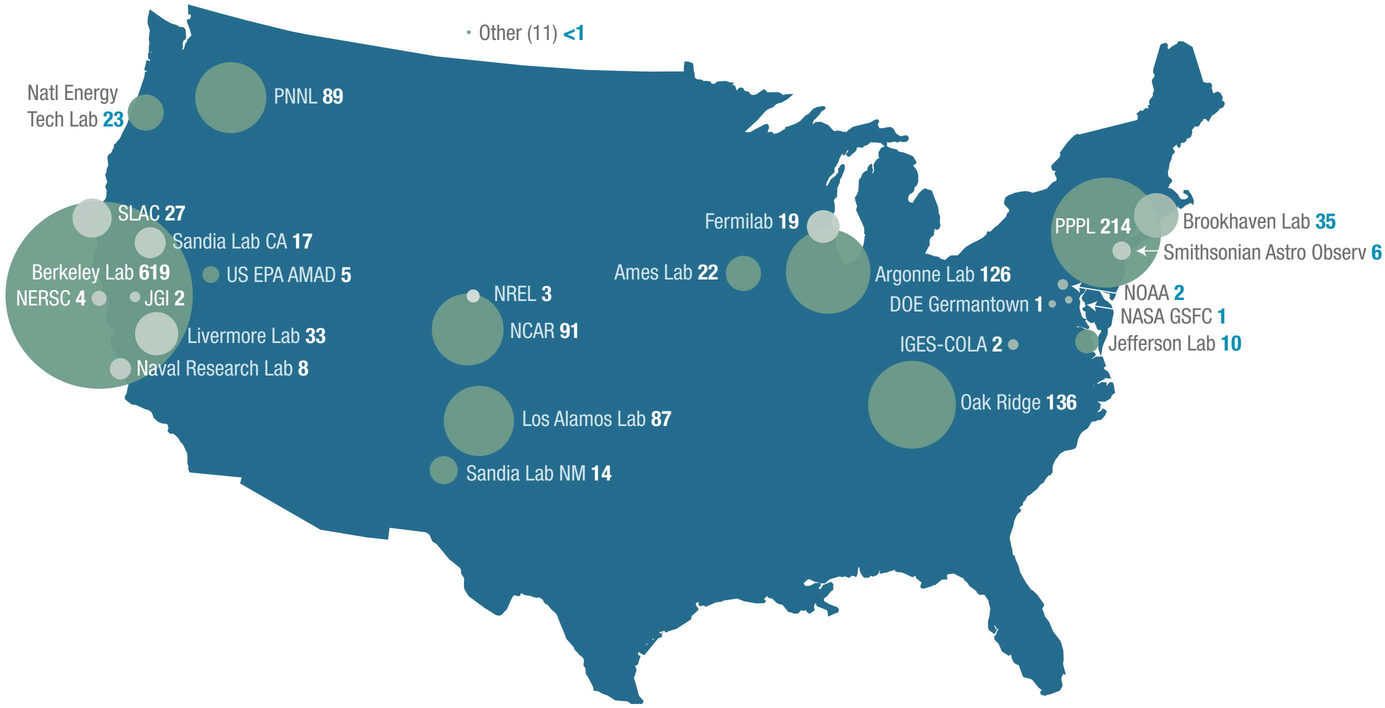
2015 NERSC Usage by Institution Type

(MPP HOURS IN MILLIONS)



2015 DOE & Other Lab Usage at NERSC

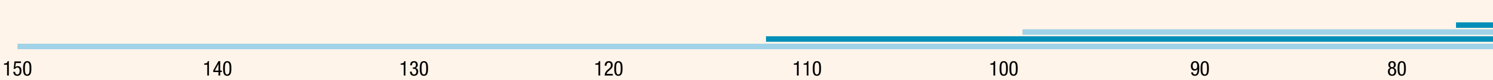
(MPP HOURS IN MILLIONS)



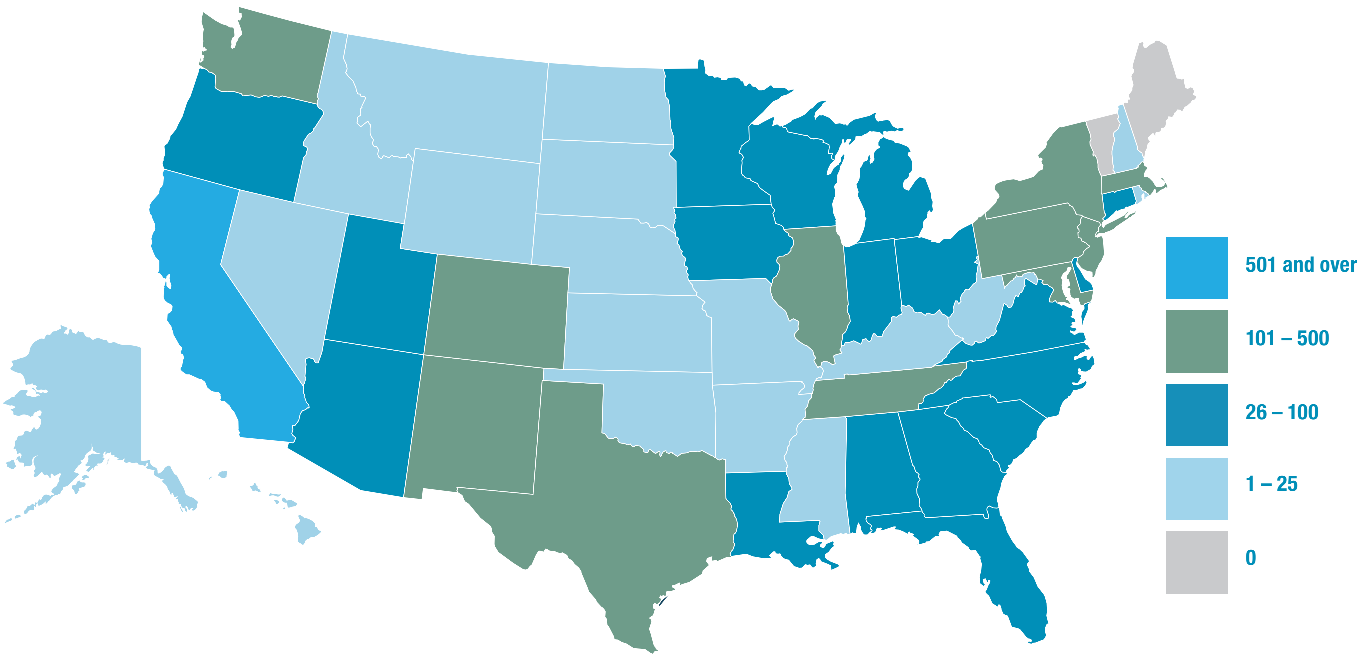
2015 Academic Usage at NERSC

(MPP HOURS IN MILLIONS)

MIT	150	Temple University	28	U. Missouri KC	11
William & Mary	112	Tulane Univ	28	Harvard Univ	10
U. Arizona	99	Northwestern Univ	25	U. Florida	8
UC San Diego	77	U. Penn	25	Penn State	8
UC Berkeley	75	Iowa State	23	Stanford Univ	8
U. Illinois U-C	72	U. Maryland	22	UMass Amherst	8
U. Washington	56	Colorado School Mines	20	Louisiana State	7
UC Irvine	54	U. Michigan	19	UC Santa Barbara	7
UCLA	53	U. Texas Austin	18	U. Tennessee	7
U. Colorado Boulder	50	U. Rochester	17	U. New Hampshire	7
Princeton Univ	40	U. Oklahoma	17	George Wash Univ	7
Caltech	40	U. Central Florida	16	U. Delaware	6
U. Chicago	37	Cornell Univ	15	Rice Univ	6
U. Kentucky	36	North Carolina State	13	UNC Chapel Hill	6
Columbia Univ	32	U. South Carolina	13	Dartmouth College	5
U. Wisc. Madison	32	Vanderbilt Univ	13	Clemson Univ	5

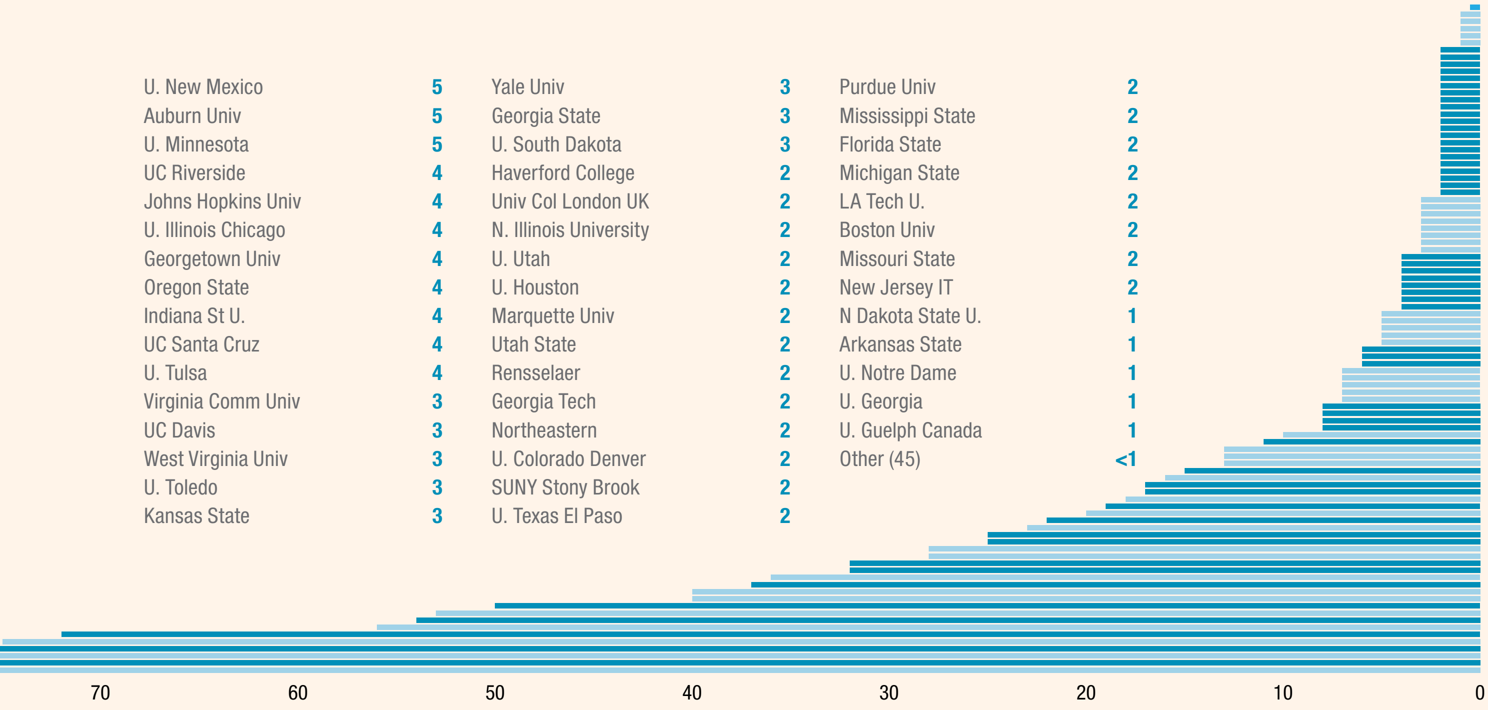


2015 NERSC Users by State



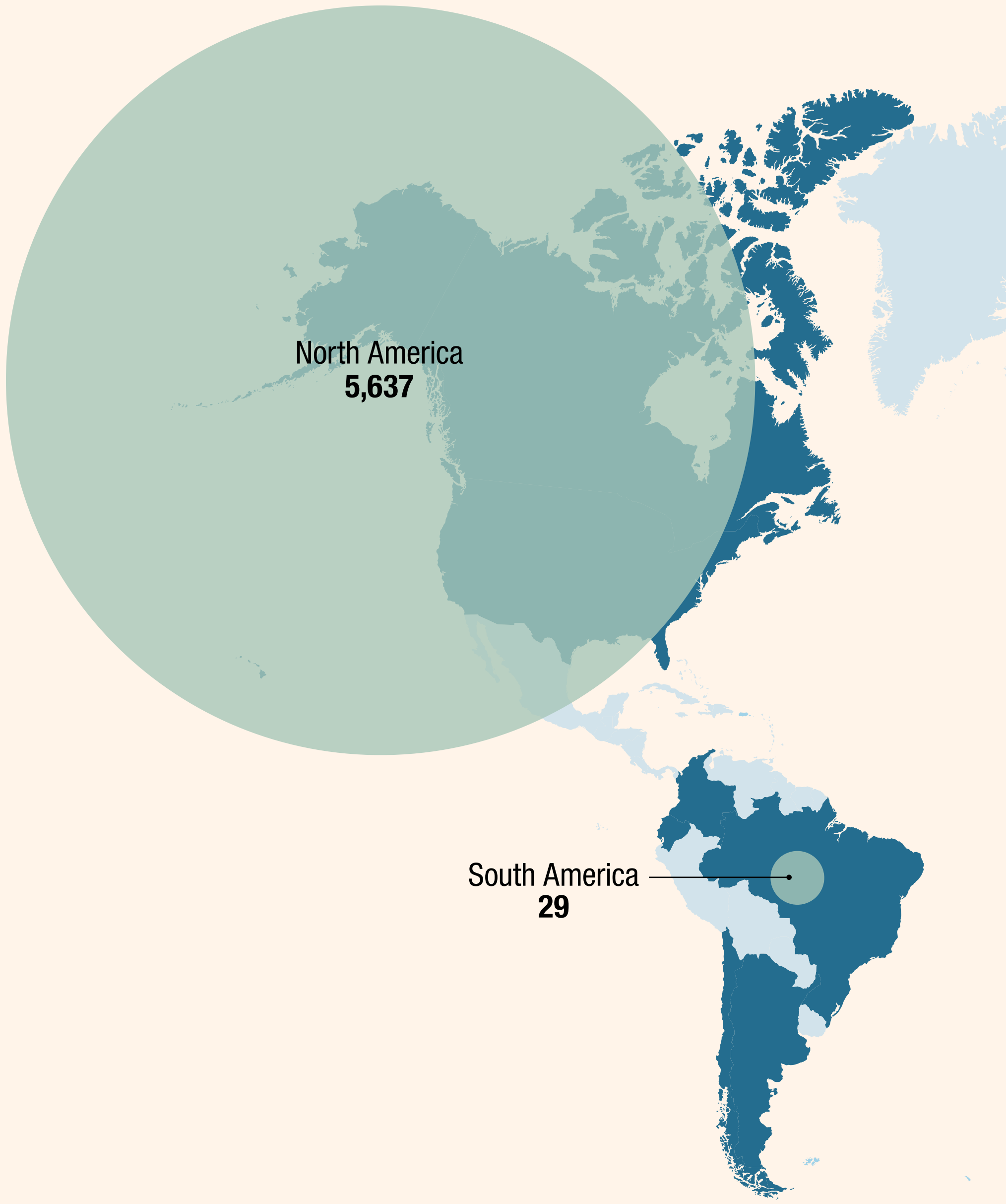
California	2,265	New Jersey	123	Iowa	64	Delaware	26	West Virginia	8
Illinois	368	Maryland	103	Ohio	43	Missouri	21	Mississippi	6
Tennessee	265	Michigan	89	Minnesota	41	South Dakota	18	Hawaii	5
New York	258	Connecticut	85	Utah	39	Kansas	16	Wyoming	5
Washington	231	North Carolina	83	District of Columbia	38	Kentucky	15	Nevada	4
Massachusetts	185	Indiana	78	Arizona	30	Oklahoma	12	Montana	3
Texas	169	Florida	72	Louisiana	30	New Hampshire	11	Alaska	2
Colorado	168	Georgia	69	South Carolina	29	North Dakota	10	Nebraska	2
Pennsylvania	151	Wisconsin	69	Alabama	28	Rhode Island	10	Idaho	1
New Mexico	144	Virginia	66	Oregon	28	Arkansas	8	Maine	0
								Vermont	0

U. New Mexico	5	Yale Univ	3	Purdue Univ	2
Auburn Univ	5	Georgia State	3	Mississippi State	2
U. Minnesota	5	U. South Dakota	3	Florida State	2
UC Riverside	4	Haverford College	2	Michigan State	2
Johns Hopkins Univ	4	Univ Col London UK	2	LA Tech U.	2
U. Illinois Chicago	4	N. Illinois University	2	Boston Univ	2
Georgetown Univ	4	U. Utah	2	Missouri State	2
Oregon State	4	U. Houston	2	New Jersey IT	2
Indiana St U.	4	Marquette Univ	2	N Dakota State U.	1
UC Santa Cruz	4	Utah State	2	Arkansas State	1
U. Tulsa	4	Rensselaer	2	U. Notre Dame	1
Virginia Comm Univ	3	Georgia Tech	2	U. Georgia	1
UC Davis	3	Northeastern	2	U. Guelph Canada	1
West Virginia Univ	3	U. Colorado Denver	2	Other (45)	<1
U. Toledo	3	SUNY Stony Brook	2		
Kansas State	3	U. Texas El Paso	2		



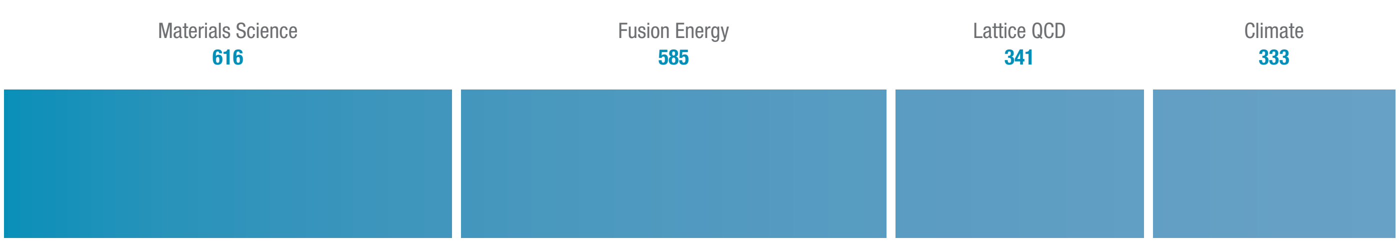
2015 NERSC Users by Country

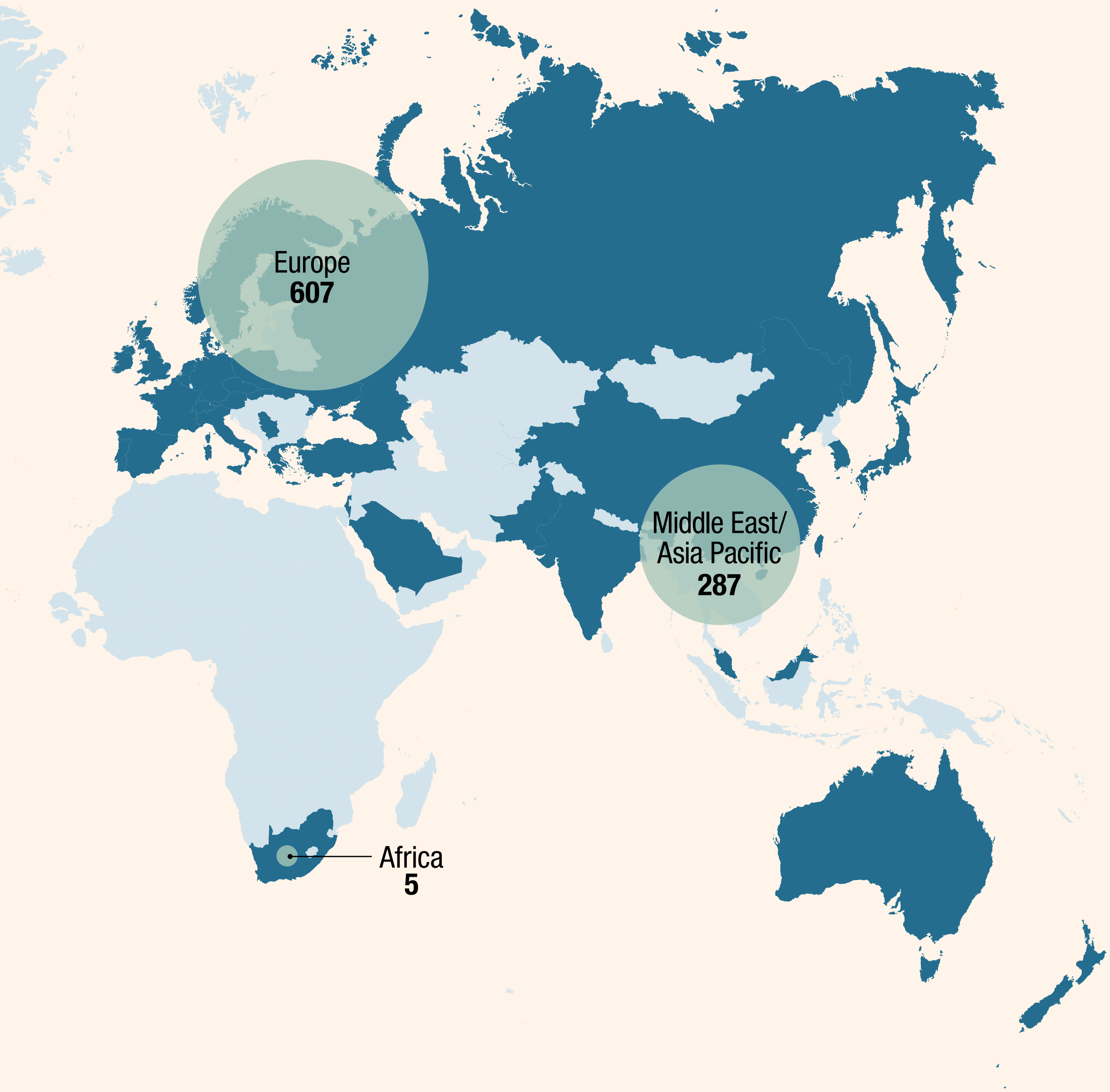
North America:	**5,637**
United States of America	5,594
Canada	38
Puerto Rico	5
South America:	**29**
Chile	15
Brazil	11
Argentina	1
Colombia	1
Ecuador	1
Africa:	**5**
South Africa	5
Middle East/ Asia Pacific:	**287**
People's Republic of China	142
Republic of Korea (South)	30
Japan	24
India	24
Israel	19
Australia	18
Taiwan, Province of China	17
Singapore	6
Saudi Arabia	4
Malaysia	1
New Zealand	1
Pakistan	1



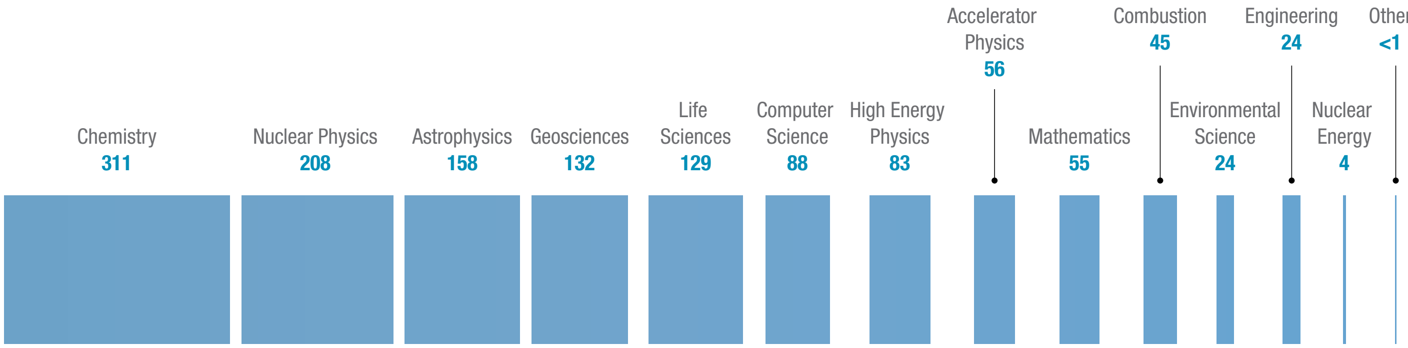
2015 NERSC Usage By Discipline

(MPP HOURS IN MILLIONS)





Europe:	607
United Kingdom	137
Germany	84
France	79
Italy	65
Switzerland	46
Spain	36
Denmark	34
Poland	23
Czech Republic	22
Russian Federation	14
Norway	12
Sweden	11
Netherlands	7
Finland	6
Greece	6
Portugal	6
Belgium	4
Cyprus	4
Ireland	3
Turkey	3
Slovenia	2
Austria	1
Serbia	1
Ukraine	1



Innovations

NERSC sets the bar high when it comes to providing its users with increased performance and additional services and operational efficiencies. The following pages highlight NERSC innovations from 2015 in our three strategic focus areas: exascale computing, data-intensive science and operational excellence.

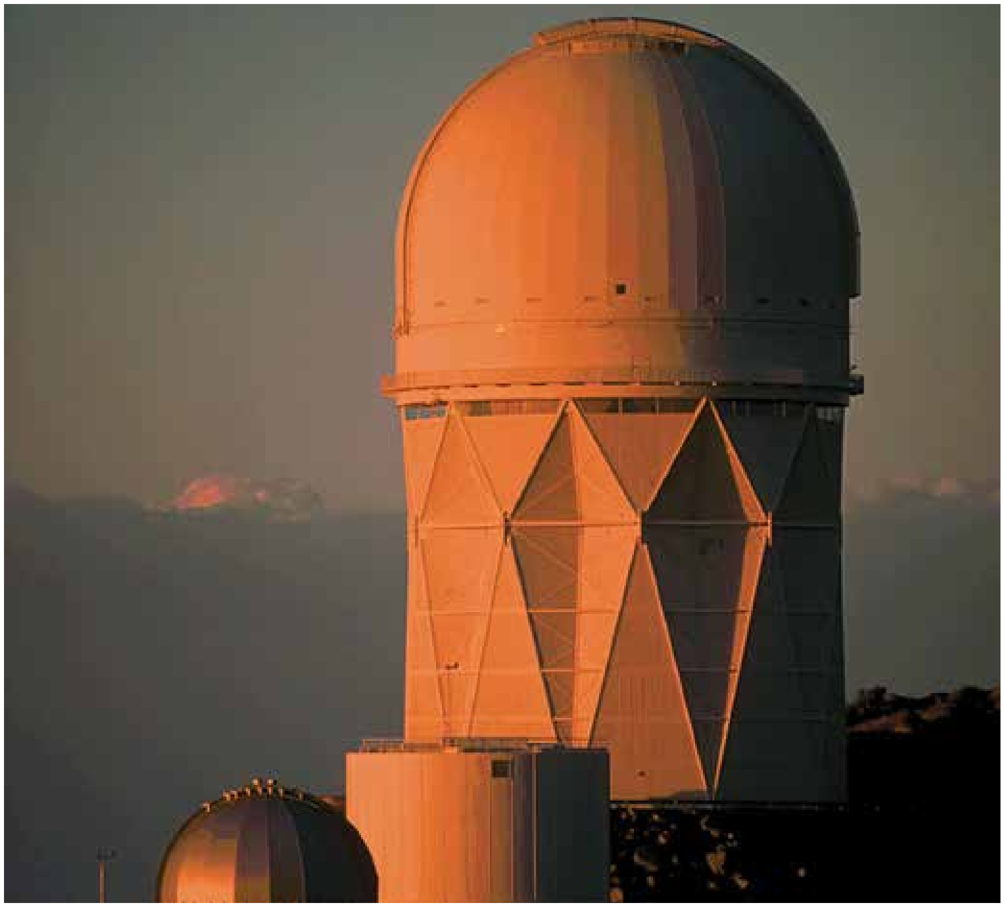
Paving the Way to Exascale

Powerful light sources (such as Berkeley Lab's Advanced Light Source, shown here), telescopes and accelerators are giving scientists greater insight into many phenomena, generating data at unprecedented speed and resolution. But getting the science out of the data requires powerful processing—such as exascale—to search, transform, analyze and interpret the data. *Image: Lawrence Berkeley National Laboratory*

Scientific discovery and translational R&D are increasingly data driven as faster, more powerful experimental facilities generate data at unprecedented speed and resolution. In turn, increasingly powerful processing tools are needed to search, transform, analyze and interpret this data. And it needs to be done quickly so that scientists can use the results to guide future experiments and make the most efficient use of facilities. Specific examples of these impending data challenges include:

- Genomics and metagenomics, where data volume will increase to ~100PB by 2021
- High energy physics, such as the Large Hadron Collider, with 50-100 exabytes of data starting in 2025
- Light sources, such as Berkeley Lab's Advanced Light Source, which will each generate about 1.5PB/month by 2020
- Climate, where annual data from simulations is expected to reach around 100PB by 2020

Toward this end, a principal goal of NERSC's exascale strategic initiative is to meet the ever-growing computing and data needs of the DOE Office of Science community by providing exascale computing and storage systems to the broad NERSC user community. Over the past year, NERSC made significant progress in developing the tools and techniques that will be used to assess the usability of exascale architectures in the context of the evolving NERSC workload.



The Dark Energy Spectroscopic Instrument (DESI) will measure the effect of dark energy on the expansion of the universe by obtaining optical spectra for tens of millions of galaxies and quasars. DESI will be conducted on the Mayall 4-meter telescope at Kitt Peak National Observatory in Arizona (shown here) starting in 2018. *Image: NOAO/AURA/NSF*

Extreme-Scale and Data-Intensive Platforms

An increasing number of scientific discoveries are being made based on the analysis of large volumes of data coming from observational science and large-scale simulation. These analyses are beginning to require resources only available on large-scale computing platforms, and data-intensive workloads are expected to become a significant component of the Office of Science's workload in the coming years.

For example, scientists are using telescopes on Earth—such as the Dark Energy Spectroscopic Instrument (DESI) managed by LBNL—and mounted on satellites to try to "see" the unseen forces, such as dark matter, that can explain the birth, evolution and fate of our universe. Detailed, large-scale simulations of how structures form in the universe will play a key role in advancing our understanding of cosmology and the origins of the universe.

To ensure that future systems will be able to meet the needs of extreme-scale and data-intensive computing in cosmology, particle physics, climate and other areas of science, NERSC has undertaken an extensive effort to characterize the data demands of the scientific workflows run by its users. In 2015 two important image-analysis workflows—astronomy and microtomography—were characterized to determine the feasibility of adapting these workflows to exascale architectures. Although both workflows were found to have poor thread scalability in general, we also determined that adapting these workflows to manycore and exascale node architectures will not be fundamentally difficult; coarse-grained threading and the use of large private buffers were the principal limits to scalability and optimization strategies to address these bottlenecks already exist.

The I/O requirements of these workflows were also characterized, and in general their I/O patterns are good candidates for acceleration on a flash-based storage subsystem. However, these workflows do not re-read data heavily, and it follows that transparent-caching architectures would not deliver the full benefit of the flash subsystem. Rather, these workflows would benefit most from the ability to explicitly stage data into the flash tier before they run.

Although this in-depth study revealed scalability issues in today's workflows, it did not identify any critical roadblocks to developing a single exascale system architecture that can meet the needs of both existing extreme-scale workloads and these data-intensive workflows. Work is ongoing, and the results of the broader study have been published in a whitepaper that is being provided as a part of the center's next-generation supercomputer (currently known as NERSC-9) technical requirements.

New Benchmarks for Data-Intensive Computing

The genome of bread wheat is five times larger than that of humans and contains many repeat sequences, making it challenging to obtain a reference sequence for this important crop. *De novo* assembly algorithms developed for such large genomes can dramatically speed data sequencing on supercomputers such as NERSC's Edison system and future exascale systems. *Image: Quer, https://commons.wikimedia.org/w/index.php?curid=15601542*

The increasing volumes of data that are driving the convergence of extreme-scale and data-intensive computing are accompanied by applications that stress aspects of system performance that have not been traditionally represented in benchmark suites. To ensure that these emerging applications will run well on future architectures, NERSC developed two new benchmark codes that are now a requirement in our next-generation supercomputer (currently known as NERSC-9) benchmarking suite.

In collaboration with LBNL's Computational Research Division, UC Berkeley and the Joint Genome Institute, NERSC developed Meraculous, a mini-app that performs *de novo* assembly of large genomes. Algorithmically, this mini-app performs parallel graph construction and traversal, and it relies on communication patterns that have not been stressed in traditional extreme-scale applications. However, these lightweight, single-sided communications are expected to be critical for communication performance between the throughput-optimized cores on exascale architectures.

Another bioinformatics application, BLAST, will become a significant source of the I/O load on the NERSC-9 system as it absorbs the high-throughput computing workflows currently running on a dedicated cluster for Joint Genome Institute users. This pattern-searching application matches a series of unknown DNA or protein sequences with known sequences by repeatedly scanning a database that is stored on disk. This process generates significant I/O loads that lie somewhere between the fully sequential and fully random I/O patterns traditionally benchmarked with the standard IOR benchmark. To address this, NERSC has extended the IOR benchmark to include a new mode that exercises aspects of I/O performance, including page cache and single-thread I/O, that are required by BLAST.

These new benchmarks represent the requirements of the broadening workload that is expected to require exascale-era resources. Optimizing for these data-intensive applications may reveal additional avenues for innovation in the I/O subsystem architectures being developed by vendors, and their inclusion in the procurement requirements for pre-exascale systems should provide a smoother on-ramp to exascale for data-intensive applications.

NERSC Introduces Power Measurement Library

Energy efficiency is a key driver in exascale architecture design considerations, and exascale nodes are expected to derive their performance from a large number of power-efficient, low-frequency compute cores. To better understand how application performance and power consumption are related, NERSC developed a profiling library to measure the power consumption of individual applications or subroutines. This library builds upon the low-level power-monitoring infrastructure provided by Cray by allowing application developers to measure the power consumed by a node while certain regions of code are being executed.

This library is currently being used by NERSC and LBNL's Computational Research Division to explore how compute and I/O performance change as CPU power efficiency increases. For example, the clock speed at which applications achieve peak energy efficiency was found to vary widely across applications due to the varying tradeoff between lower clock speeds and increased wall times. Those applications whose performance is limited by memory bandwidth achieve minimum energy consumption at lower frequencies, while codes that do not saturate the memory bandwidth achieve lower overall energy consumption at higher frequencies.

Similarly, applications that can saturate the I/O bandwidth of a node can achieve lower overall power consumption by reducing the CPU frequency during heavy I/O (for more on this, see https://sdm.lbl.gov/~sbyna/research/papers/201504-CUG-Power.pdf). This work found that the energy savings of running at each application's energy-optimal frequency were found to be small (less than 10 percent savings) relative to the increase in wall time (more than 10 percent longer) on the Cray XC30 platform. However, this library provides a powerful foundation on which the power-performance behavior of applications and kernels can be evaluated for emerging manycore architectures on the path to exascale.

Shifter is helping ATLAS, one of two general-purpose detectors at the Large Hadron Collider, run production level work more efficiently on NERSC's Edison system. *Image: CERN*

Data-Intensive Science

Shifter: Enabling Docker-Like Containers for HPC

The explosive growth in data coming out of experiments in cosmology, particle physics, bioinformatics and nuclear physics is pushing computational scientists to design novel software tools to help researchers better access, manage and analyze that data on current and next-generation HPC architectures.

One increasingly popular approach is container-based computing. In this context, a container is an isolated environment, used to run processes, that is constructed from an "image"—that is, a "snapshot" of an operating environment for an application or a service that typically contains the required operating system software, libraries and binaries in addition to anything specific to the application.

With the growing adoption of container-based computing, new tools to create and manage containers have emerged—notably Docker, an open-source, automated container deployment service. Docker containers wrap up a piece of software in a complete file system that houses everything it needs to run, including code, runtime, system tools and system libraries. This guarantees that the software will always operate the same, regardless of the environment in which it is running.

While Docker has taken the general computing world by storm in recent years, it has yet to be fully recognized in HPC. There are several reasons for this. First, HPC centers typically provide shared resources that are used by hundreds or, in the case of NERSC, thousands of users. Securely offering full Docker functionality is difficult in these shared environments. Second, Docker uses many advanced Linux kernel features that are not always available on large HPC systems. Finally, Docker is designed to run on systems that have local storage and direct access to the Internet, and HPC systems are typically not designed this way due to cost, performance and maintainability reasons.

To overcome these challenges, NERSC developed Shifter, an open-source, scalable software package that leverages container computing to help supercomputer users run a wider range of software more easily and securely. Shifter enables Docker-like Linux containers and user-defined images to be deployed in an HPC environment—something that to date has not been possible. Shifter is specifically focused on the needs and requirements of the HPC workload, delivering the functionality NERSC and its users are looking for while still meeting the overall performance requirements and constraints that a large-scale HPC platform and shared resource providers impose. Shifter provides performance benefits for data-intensive codes that require many different dependencies because of the way it prepares and distributes the images. It also typically provides the fastest startup times and best scaling performance of any alternative we have measured.

Shifter is not a replacement for Docker; rather, it leverages the user interface that Docker makes available to people to create their software images but does not directly use Docker software internally. Through Shifter, users can supply NERSC with a Docker image that can then be converted to a form that is easier and more efficient to distribute to the compute nodes on a supercomputer. Shifter was initially prototyped and demonstrated on NERSC's Edison system and is now in production on Cori. It was fully open-sourced and released under a modified BSD license in 2015 and is being distributed via bitbucket (https://bitbucket.org/berkeleylab/shifter/).

NERSC collaborated closely with Cray during the development of Shifter and is working with other HPC centers to broaden its adoption outside of NERSC. At SC15, Cray announced plans to make Shifter available to all Cray XC series users; it will also be made available on the Cray CS400, Cray XE and Cray XK platforms in 2016. NERSC also worked closely with the Swiss National Computing Centre (CSCS), which has deployed and tested Shifter on its Cray CS-Storm cluster.

In 2015, two major science groups began working with Shifter: the LCLS experimental facility at SLAC and the high energy physics community at CERN. The LCLS has its own software environment, which has offered some challenges when migrating it to NERSC's HPC environment. Using Shifter, however, NERSC staff were able to create a Docker image in one day, demonstrating that this tool greatly reduces the staff effort needed to migrate applications into an HPC environment. In addition, the software loads much faster; in the case of the LCLS, before Shifter was implemented, when an image was ported to NERSC it could take up to half an hour for the software to start. With Shifter, it starts in 5-20 seconds.

Scientists from the Large Hadron Collider's (LHC) ALICE, ATLAS and CMS experiments at CERN have also been testing Shifter in conjunction with the CERN Virtual Machine File System (CVMFS) software package. CVMFS is a network file system that relies on http to deliver the full LHC software stack to local nodes. It requires fuse and root permissions to install, which has made installing it on Cray systems untenable. LHC users have been getting around this by untarring pre-made bundles of their software stack onto Cray compute nodes.

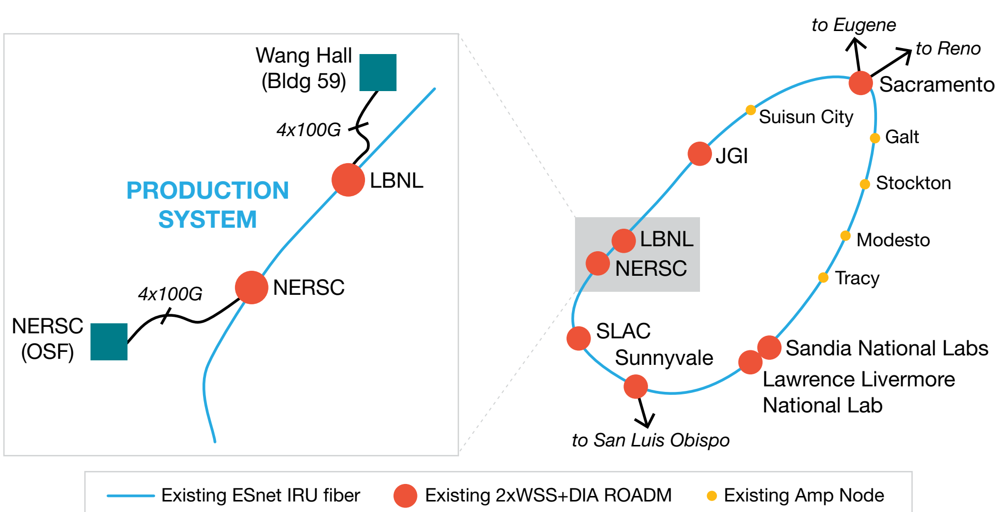
Building on work done by LHC scientists, members of NERSC's Data and Analytics Services group were able to build a file system image of the CVMFS repository. These images, which are uncharacteristically large for Shifter images, were successfully deployed using the Shifter framework. For example, the ATLAS image contains 300 GB (3.5 TB uncompressed) of data and 20 million files. Despite this, the Shifter framework has been used to run production level work on Edison out to scales of 1,000 nodes, with load times that equaled or surpassed the old tarball method (but with the full software framework instead of a small subset). For groups like this with highly specialized software stacks, Shifter allows them to run jobs as efficiently, if not more so, than they do in their current configurations and requires less effort by the user to port and optimize the codes to run on Cori and Edison.

But Shifter is not just for the largest HPC centers; it is also valuable for departmental and university clusters. Many universities and departments operate small- to medium-sized clusters used by a variety of researchers, and Shifter can enable even these smaller systems to be used more flexibly.

New 400Gbps Production Network Speeds Science

In 2015 NERSC collaborated with ESnet to build a 400 Gbps superchannel, the first-ever 400G production link to be deployed by a national research and education network. The connection, known as the BayExpress, provided critical support for NERSC's 6,000 users as the facility moved from its Oakland, Calif. location to the main LBNL campus.

To develop the 400 Gbps production connection as a step toward the next level of networking, ESnet and NERSC joined with Ciena, a global supplier of telecommunications networking equipment, software and services, and Level 3 Communications, a telecommunications and Internet service provider. The production link is deployed as a pair of 200 Gbps-per-wavelength connections and allows for a simple-to-deploy dual subcarrier 400 Gbps solution.



Map showing the 400G production link between Wang Hall at the main LBNL site and the Oakland Scientific Facility.

As part of this first-of-its-kind project, the team also set up a 400 Gbps research testbed for assessing new tools and technologies without interfering with production data traffic while allowing staff to gain experience operating a 400 Gbps optical infrastructure. For the testbed, the optical frequencies are more narrowly tuned to maximize use of the fiber spectrum, known as "gridless superchannels." The testbed uses a dark fiber link provided by Level 3 to ESnet and NERSC for six months. The team conducted field trials on the testbed during the latter part of 2015.

The 400 Gbps superchannel was further put to the test when it was used to relocate file systems from the Oakland Scientific Facility to the new Wang Hall at LBNL. Using features of IBM's Spectrum Scale file system (formerly GPFS), the NERSC Storage Systems Group successfully moved the 4.8 petabyte /project file system to Wang Hall in November 2015. Achieving transfer speeds up to 170 Gbps over the 400 Gbps link, NERSC engineers balanced the need to quickly complete the file system relocation while minimizing impact to users. Users continued to actively use the file system and noticed little impact from the move.

The last leg of the transfer of file system data to Wang Hall presented an additional challenge. The storage system network at Wang Hall is Infiniband, a high-bandwidth, low-latency switched network technology that is widely used in HPC centers, while data from the 400G superchannel uses the more common Ethernet networking technology. A method was needed to route between the Ethernet and Infiniband networks without degradation in transfer rates. To address this, NERSC deployed an array of 14 commodity compute systems running the open source VyOS network operating system. Unlike typical routers that make packet-forwarding decisions using custom-built hardware, VyOS uses software to route packets. With access to the software source code, NERSC modified the VyOS operating system to support Infiniband and provide failover for the connections on the Infiniband network. The 14 routers were deployed in pairs, with each member in the pair able to stand in for the other if a router failed. With tuning, a single router achieved an aggregate transfer rate of 37 Gbs of the 40 Gbs available.

During 2016, the link between the Oakland and LBNL facilities will remain in place until all systems have been moved to Wang Hall.

Cori Gets a Real-Time Queue

NERSC is increasingly being coupled with other user facilities to support analysis from instruments and detectors. In many cases, these use cases have a need for rapid turnaround since users may need to see the results of the analysis to determine the next set of experiments, configuration changes, etc.

To help support these workloads and others with similar needs, Cori now supports a real-time queue. Users can request a small number of on-demand nodes if their jobs have special needs that cannot be accommodated through the regular batch system. The real-time queue enables immediate access to a set of nodes for jobs that are under the real-time wallclock limit. It is not intended to be used to guarantee fast throughput for production runs. To support this queue, NERSC heavily leveraged features within the SLURM batch system.

Since the real-time queue could easily be misused, NERSC has given only a select set of projects access to this capability. To create this list, NERSC first sent a call for proposals to NERSC users asking if they had real-time computing needs with a justification during the allocation process.

The Palomar 48-inch telescope and time-lapse image of the night sky. The Palomar Transient Factory uses the real-time queue at NERSC to identify candidate transient objects for further analysis. *Image: Caltech*

NERSC then reviewed the list of projects and evaluated the justifications they provided. From this list, 15 were accepted, and as of January 2016 four projects are enabled and running in the real-time queue. Many of the submissions were associated with experimental facilities such as the ALS and LCLS or instruments such as telescopes and mass spectrometers.

One example of a project that is already making use of the real-time queue is the Palomar Transient Factory (PTF). The PTF survey uses a robotic telescope mounted on the 48-inch Samuel Oschin Telescope at the Palomar Observatory in Southern California to scan the sky nightly. As soon as the observations are taken, the data travels more than 400 miles to NERSC via the National Science Foundation's High Performance Wireless Research and Education Network and ESnet. At NERSC, computers running machine-learning algorithms in the Real-Time Transient Detection Pipeline scan through the data and identify events to follow up on. The real-time queues on Cori have allowed the project to go from shutter closed at the telescope to new transients classified in the database and distributed to the collaboration in less than 6 minutes for 95 percent of the images.

Other projects making early use of the real-time queue are the Metabolite Atlas and OpenMSI, which are using this capability to accelerate compound identification from mass spectrometry datasets. Identification is achieved by comparing measured spectra to all theoretically possible fragmentation paths for known molecular structures. To date, these projects have computed and stored complete fragmentation trees to a depth of five for greater than 11,000 compounds. These trees are then used to identify new molecules from raw experimental data using the Cori real-time queue, which allows Metabolite Atlas and OpenMSI users to search their raw spectra against these trees and obtain results in minutes. Without systems like Cori, these tasks would take months or would not be performed at all. The real-time queue enables users to make better use of their time, avoiding the variable wait times previously experienced using the standard batch queues.

iPython/JupyterHub Portal Debuts

The Jupyter Notebook is a web application that allows users to create and share documents that contain live code, equations, visualizations and explanatory text. JupyterHub is a multi-user version of the notebook designed for centralized deployments.

NERSC has deployed a JupyterHub service that allows users to spin up interactive iPython notebooks with access to their data and computational capabilities of the large-scale NERSC platforms. By integrating it with NERSC LDAP authentication, each user can start a private notebook on the NERSC JupyterHub server and run interactive computations through a Python shell. This can be used for interactive development and prototyping but is also useful for workflow orchestration and advanced visualization. For example, the OpenMSI project is using iPython notebooks to enable users to analyze and visualize mass spectrometry data.

NERSC's JupyterHub service also gives users the ability to customize their notebook environments and include their own packages and libraries. This makes it easy to include domain-specific or community-developed packages for specific forms of analysis or visualization. This powerful tool combines the expressiveness of a language like Python with the ability to perform interactive analyses through a web portal. One can interact with vast datasets stored on NERSC file systems and easily access Python tools for visualizing and analyzing the data.

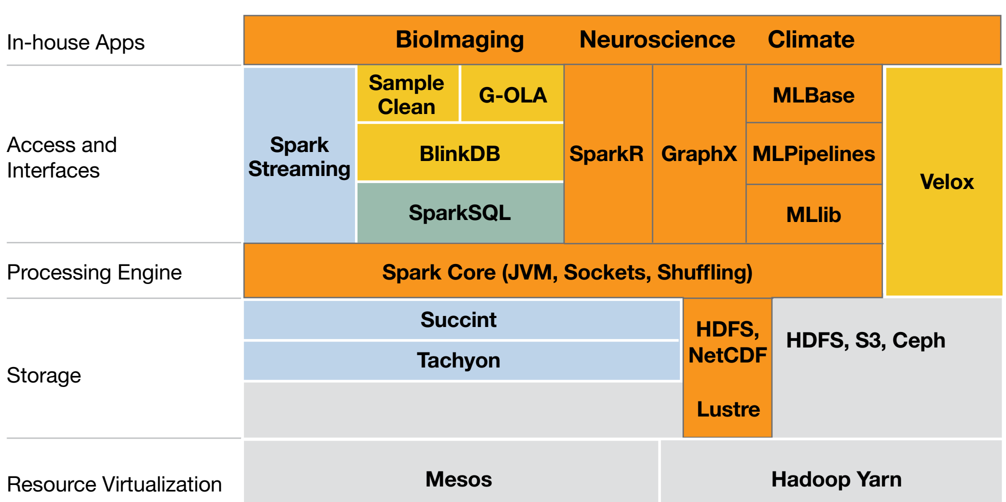
Despite being only in beta release mode during 2015 and not broadly advertised, the iPython/ JupyterHub deployment has already been used by over 100 users in domains spanning biology, cosmology, high-energy physics and climate. NERSC made this a full production service in early 2016 and is already exploring ways to more tightly integrate it with Cori to enable advanced analytics and workflow orchestration, including the ability to direct the execution of parallel jobs via iPython.

Cray-AMPLab-NERSC Partnership Targets Data Analytics

As data-centric workloads become increasingly common in scientific and industrial applications, a pressing concern is how to design large-scale data analytics stacks that simplify analysis of the resulting data. A collaboration between Cray, researchers at UC Berkeley's AMPLab and NERSC is working to address this issue.

The need to build and study increasingly detailed models of physical phenomena has benefited from advancement in HPC, but it has also resulted in an exponential increase in data from simulations as well as real-world experiments. This has fundamental implications for HPC systems design, such as the need for improved algorithmic methods and the ability to exploit deeper memory/storage hierarchies and efficient methods for data interchange and representation in a scientific workflow. The modern HPC platform has to be equally capable of handling both traditional HPC workloads and the emerging class of data-centric workloads and analytics motifs.

In the commercial sector, these challenges have fueled the development of frameworks such as Hadoop and Spark and a rapidly growing body of open-source software for common data analysis and machine learning problems. These technologies are typically designed for and implemented in



Orange highlights show portions of the Berkeley Data Analytics Stack that are being explored as part of the Cray/AMPLab/NERSC collaboration.

distributed data centers consisting of a large number of commodity processing nodes, with an emphasis on scalability, fault tolerance and productivity. In contrast, HPC environments are focused primarily on no-compromise performance of carefully optimized codes at extreme scale.

This scenario prompted the Cray, AMPLab and NERSC team to look at how to derive the greatest value from adapting productivity-oriented analytics tools such as Spark to HPC environments, and how to ensure that a framework like Spark can exploit supercomputing technologies like advanced interconnects and memory hierarchies to improve performance at scale without losing productivity benefits. The team is actively examining research and performance issues in getting Spark up and running on HPC environments, such as NERSC's Edison and Cori systems, with a long-term goal of bridging the gap between data analytics for commercial applications and high-performance scientific applications.

In addition, since linear algebra algorithms underlie many of NERSC's most pressing scientific data-analysis problems, the team is developing novel randomized linear algebra algorithms and implementing these algorithms within the AMPLab stack and on Edison and Cori (dubbed the Berkeley Data Analytics Stack). They are also applying these algorithms to some of NERSC's most pressing scientific data-analysis challenges, including problems in bioimaging, neuroscience and climate science.

Two of the main performance issues are file system I/O and node memory. The performance of Spark applications is dictated by the spark "shuffle," during which data is sent between nodes through a file system spill. Traditionally, commercial data centers have local disk, so shuffle performance is network-bound. In contrast, HPC systems like Edison and Cori have global file systems, so shuffle performance is dictated primarily by file system metadata latency. In addition, if a node runs out of memory during the shuffle, it will spill intermediately to the file system. HPC systems typically have less memory per node (and, thus, per core) compared to industry data centers. For large memory workloads, such as linear algebra algorithms on large matrices, HPC systems have to spill to the file system more often than data center workloads, which results in more of a hit on performance.

In order to see this difference in performance, we compared the performance of randomized linear algebra algorithm CX on a real scientific use case, a 1TB bioimaging dataset in a paper submitted to the 2016 IEEE International Parallel & Distributed Processing Symposium. To close some of the performance gaps observed in this use case, we tried a variety of techniques in collaboration with Cray and CRD to decrease the number of metadata operations in a Spark application. Testing has shown that these techniques can significantly improve performance, especially for workloads that impose a heavy shuffle load.

Deep Learning Comes to HPC

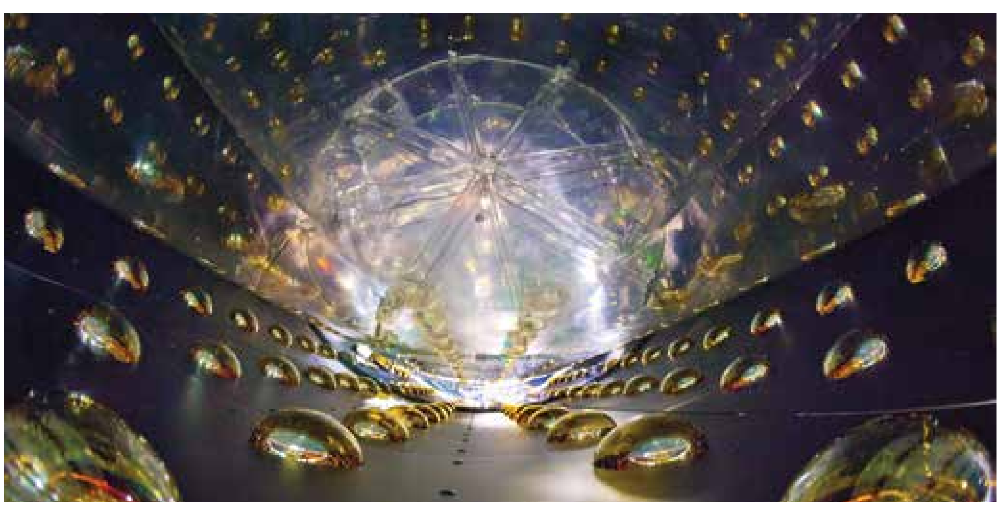
In 2015 researchers from NERSC and LBNL began applying deep learning software tools developed for HPC environments to a number of "grand challenge" science problems running computations at NERSC and other supercomputing facilities.

Deep learning, a subset of machine learning, is the latest iteration of neural network approaches to machine learning problems. Machine learning algorithms enable a computer to analyze pointed examples in a given dataset, find patterns in those datasets and make predictions about what other patterns it might find.

Deep learning algorithms are designed to learn hierarchical, nonlinear combinations of input data. They get around the typical machine learning requirements of designing custom features and produce state-of-the-art performance for classification, regression and sequence prediction tasks. While the core concepts were developed over three decades ago, the availability of large data, coupled with data center hardware resources and recent algorithmic innovations, has enabled companies like Google, Baidu and Facebook to make advances in image search and speech recognition problems.

Surprisingly, deep learning has so far not made similar inroads in scientific data analysis, largely because the algorithms have not been designed to work on high-end supercomputing systems such as those found at NERSC. But during 2015 NERSC began enabling the Daya Bay Neutrino Experiment to apply deep learning algorithms to enhance its data analysis efforts. Located in China, Daya Bay is an international neutrino-oscillation experiment designed to determine the last unknown neutrino mixing angle θ_{13} using anti-neutrinos produced by the Daya Bay and Ling Ao Nuclear Power Plant reactors. NERSC's Data and Analytics Services Group implemented a deep learning pipeline for Daya Bay at NERSC, using convolutional neural nets to automatically reduce and classify the features in the data. The work produced a supervised classifier with over 97 percent accuracy across different classes, as well as an unsupervised convolutional autoencoder capable of identifying patterns of physics interest without specialized input knowledge.

To our knowledge, this is the first time unsupervised deep learning has been performed on raw particle physics data. NERSC is now collaborating with physicists to investigate in detail the various clusters formed by the representation to determine what interesting physics is captured in them beyond the initial labeling and to incorporate such visualizations into the monitoring pipeline of the experiment.



NERSC's Data and Analytics Services Group has implemented a deep learning pipeline for the Daya Bay Neutrino Experiment Facility. *Image: Roy Kaltschmidt, Lawrence Berkeley National Laboratory*

Cori Phase I installation in the new Wang Hall computer room. *Image: Kelly Owen, Lawrence Berkeley National Laboratory*

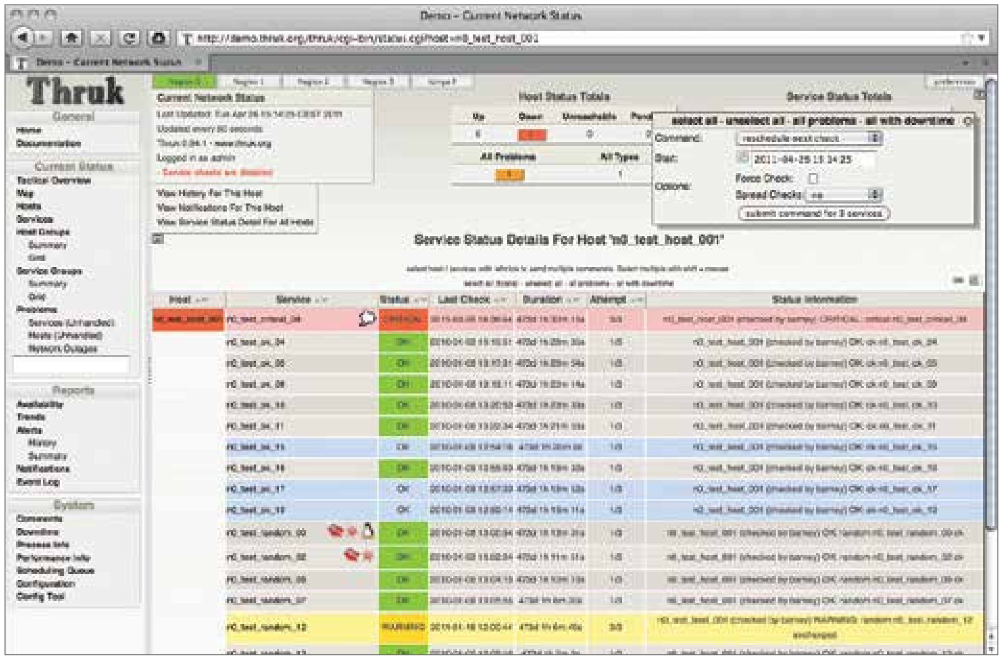
Operational Excellence

Thruk: A New Fault-Monitoring Web Interface

The move to Wang Hall prompted NERSC's Operational Technology Group (OTG)—the team that ensures site reliability of NERSC systems, storage, facility environment and ESnet—to implement several new fault-monitoring efficiencies. Foremost among them is Thruk, an enhancement to OTG's nagios fault-monitoring software that uses the Livestatus API to combine all instances of systems and storage into a single view.

Implementing Thruk provided OTG with an efficient method for correlating various alerts and determining if an issue affects systems throughout Wang Hall or just in certain parts of the facility. The most important advantage of this new tool is that it allows the OTG to monitor multiple aspects of systems in a single view instead of multiple tabs or multiple url interfaces. This has decreased the possibility of missing critical alerts. Other advantages include:

- A single location to view alerts to multiple systems
- A configuration interface that is flexible, allowing the OTG to easily add and remove systems
- The ability to send multiple commands at once without waiting for a response
- Multiple filtering schemes



Screen shot of the new Thruk interface.

Thruk has also been instrumental in supporting the OTG's efforts to improve workflow automation. In 2015 OTG staff implemented an automated workflow for timely reporting of user issues that uses program event handlers to manage multiple instances of issues and multiple action items. For example, alerts that called for a staff person to open a new ticket and assign it to a group can now be automated based on a filtering mechanism. The new interface saves the group's workflow approximately 17 hours per week, in addition to tickets now being opened accurately on a more consistent basis. In the past, the language and the information included on the ticket depended on which staff member opened the ticket.

Another advantage of the automation process involves opening tickets with vendors as a result of an alert. In the past, a staff person would either e-mail or call the vendor call center to report a problem. The phone call could take anywhere from 15 minutes to 45 minutes, depending on the vendor, and the time it took for the vendor tech to respond could be anywhere from one to four hours. In contrast, the automated process sends an e-mail to the vendor call center, which usually puts a priority on clearing out their inbox of requests. Thus a ticket sent to the vendor can be opened within 15 minutes, providing detailed information such as diagnostic logs as an attachment to the e-mail. This has resulted in improving vendor tech response times to under one hour. In addition, with the appropriate information provided with each alert, vendors spend less time on diagnosis and more time on problem resolution.

The event handlers can also create and assign work tickets to staff. For example, an alert that says a drive has failed and needs replacement can activate an event handler that opens a ticket and assigns it to a staff person who can replace that drive, work with vendors for parts under warranty and allow the staff to bring the node back into production sooner. One of the areas that allowed this process to be automated involves centralizing the location of the warrantied parts with their purchase date, serial number, vendor information and warranty expiration date. With the verification process handled automatically, it now takes only seconds to decide whether to exchange or acquire a part. This has resulted in compute nodes coming back to production in a shorter amount of time—in some instances as little as four to eight hours, where previously it could sometimes take twice that long.

Environmental Data Collection Projects

With the move to Wang Hall, NERSC is also taking the opportunity to collect environmental data from areas of the machine room where we had not previously done so at the Oakland Scientific Facility. We've installed sensors in different areas of the computational floor such as underneath the racks and even on the ceiling. We felt that getting a wide range of data on the environment will assist us in correlating between events that could affect the center and jobs running on the computing resources.

For example, we are collecting data from sensors at the substation, through the different levels of PDUs available in the building down to the node level if possible, including the UPS/generator setup. The data is collected from the subpanels and passed to the central data collection system via a Power Over Ethernet (PoE) network setup, since many of these switches and panels are located far away from the traditional network infrastructure. In one case, the substation might not include normal 110 volt power for the sensors and hardware, which is why PoE is an easy way to get both networking and power to these remote locations. We plan to eventually install more than 3,000 sensors, with each sensor collecting 10 data points every second.

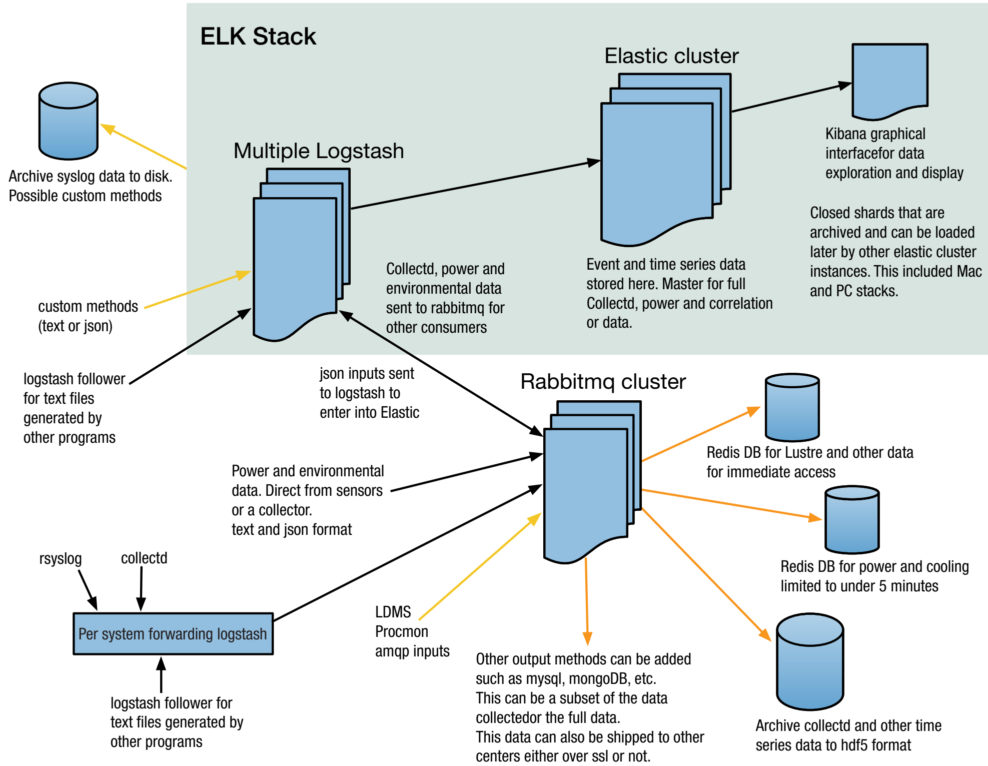
Because Wang Hall is so dependent on the environment, we can leverage all the environmental data to determine optimal times to run computation-intensive jobs or small jobs or potentially predict how the environment affects computation on a hot summer day.

Centralizing Sensor Data Collection

In addition to environmental sensors, NERSC's move to Wang Hall also prompted opportunities for the OTG to install other types of sensors in areas where this had not been done previously. These include temperature sensors from all parts of the computer floor and from all levels (floor to ceiling), power sensors from the substations to the PDUs in the rack, water flow and temperature sensors for the building and large systems and host and application measurement from all running systems. The data collected from these sensors gives us more ways to correlate causes to issues than was previously possible. We believe that this data should be able to answer questions such as power consumption of a job, power efficiency of systems for processing jobs or how we can leverage a schedule based on size of job to time of day or even the current weather.

However, the additional sensors generate a large amount of data that needs to be stored, queried, analyzed and displayed, and our existing infrastructure needed to be updated to meet this challenge. Thus, during 2015 the OTG created a novel approach to collecting system and environmental data and implemented a centralized infrastructure based on the open source ELK (Elasticsearch, Logstash and Kibana) stack. Elasticsearch is a scalable framework that allows for fast search queries of data; Logstash is a tool for managing events and parsing, storing and collecting logs; and Kibana is a data visualization plugin that we use on top of the indexed content collected by Elasticsearch.

Using the ELK stack allows us to store, analyze and query the sensor data, going beyond system administration and adding the ability to extend the analysis to improve the operational efficiency of an



The new centralized data collection infrastructure at NERSC.

entire facility. This process involves collecting information on hosts or sensors across the center. This data is then passed into the collection environment via RabbitMQ—an open-source message broker software—which sends and stores the data to the Elasticsearch framework. RabbitMQ is capable of passing data to many targets, and we leveraged this feature to send the same data to Redis (an open-source, in-memory data structure store) for real-time access to some data, and to Freeboard (an open-source dashboard) to display real-time data without any storage ability. We anticipate using this method to send data to multiple targets to address future collection issues as they arise.

All of the acquired sensor data is stored in a commonly accessible format—few changes need to be made when collecting new data or analyzing the data—that is centrally accessible to most staff without transferring the data from one storage method to another or re-collecting the data in a different configuration. This new method is already providing potential data analysis opportunities not previously possible.

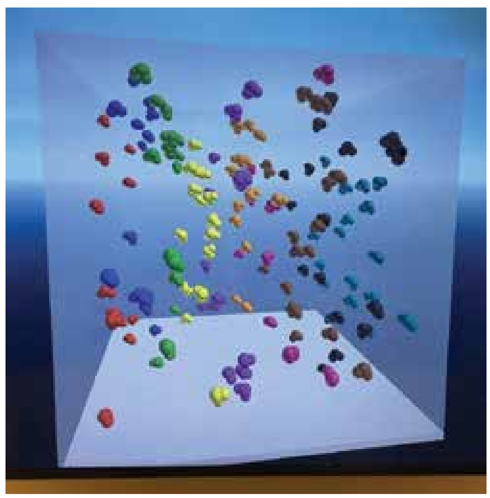
Compact Cori: An Educational Tool for HPC

In 2015 NERSC developed an HPC educational tool designed to inspire students of all ages, including middle and high-school levels, and demonstrate the value of HPC to the general public. A team of NERSC staff and student interns (two undergraduates and one high school student) created “Compact Cori,” a micro-HPC system based on Intel NUC (next unit of computing) hardware and a modern software stack based on python, MPI4PY, REST APIs, WebGL, Meteor and other modern web technologies.

The hardware design was inspired by the “Tiny Titan” micro-HPC developed at Oak Ridge National Laboratory. The hardware and software stack were chosen to be accessible and to represent modern programming techniques and paradigms with richer, more accessible educational value than traditional HPC languages like C and FORTRAN. The software stack allows students to explore several novel design points for HPC.

Despite being targeted for an educational mini-HPC, the software infrastructure represents a realistic model that can be applied to larger computing problems on production HPC resources. In particular, we developed a web-based, real-time, immersive 3D visualization technique for parallel simulations on Compact Cori. The user interface uses a web-based video game engine, a visualization concept new to HPC that allows for a rich human-computer interaction and realistic modeling of scientific data and is a compelling educational device. The web-based, real-time visualization framework represents a tangible advancement in the ability to engage users via science gateways as opposed to the traditional terminal/ssh type interaction.

Compact Cori is easy to transport and set up and carries a configurable price tag in the range of a few thousand dollars. Compact Cori is available for interactive demonstrations at NERSC.

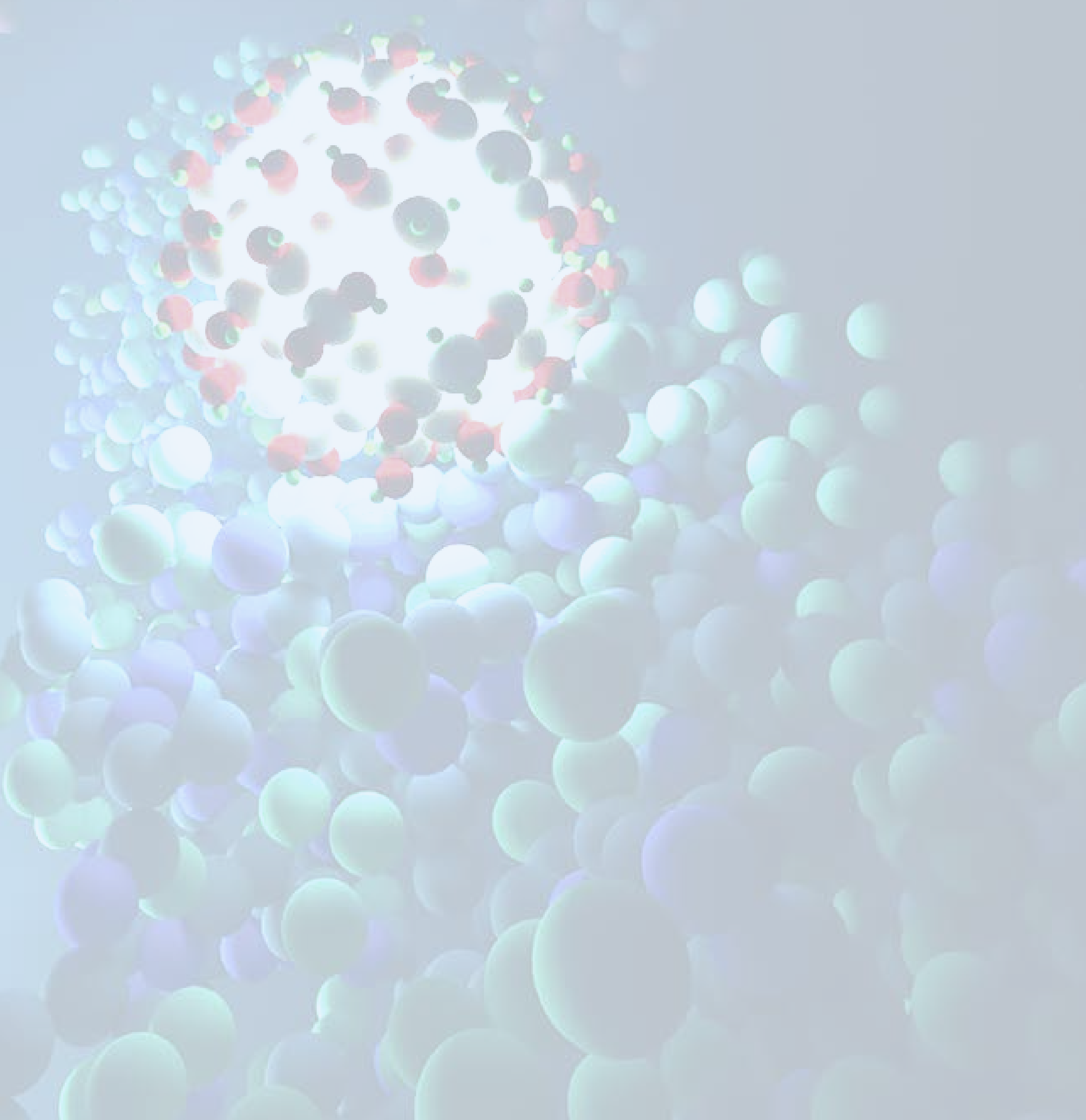


Left: Compact Cori uses 16 Intel NUC boards and a mobile Broadwell CPU. The NUC boards are linked via Gigabit Ethernet.

Right: An immersive 3D visualization technique demonstrated on Compact Cori.

Science Highlights

This section presents a selection of research highlights from 2015 based on computations and simulations run at NERSC, illustrating the breadth of science supported by the center for all six DOE program offices.



Simulations Link Mantle Plumes with Volcanic Hotspots

Imaging and Calibration of Mantle Structure at Global and Regional Scales Using Full-Waveform Seismic Tomography

BES—Geosciences

OBJECTIVE

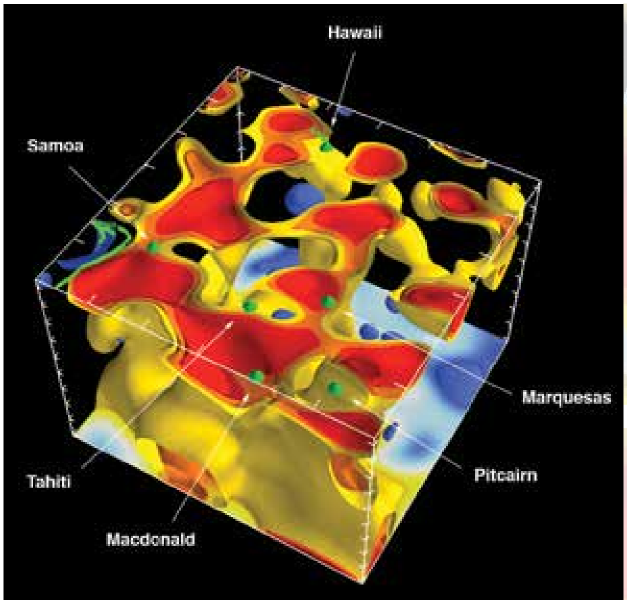
To demonstrate the relationship between plumes of hot rock rising through the Earth's mantle with surface hotspots that generate volcanic island chains.

Principal Investigator:
Barbara Romanowicz, UC Berkeley

FINDINGS/ACCOMPLISHMENTS

Numerical simulations run at NERSC helped UC Berkeley seismologists produce for the first time a 3D scan of Earth's interior that conclusively connects plumes of hot rock rising through the mantle with surface hotspots that generate volcanic island chains like Hawaii, Samoa and Iceland. Until this study, evidence for the plume and hotspot theory had been circumstantial, and some seismologists argued instead that hotspots are very shallow pools of hot rock feeding magma chambers under volcanoes.

Previous attempts to image mantle plumes have detected pockets of hot rock rising in areas where plumes have been proposed to exist. However, it was unclear whether they were connected to volcanic hotspots at the surface or at the roots of the plumes at the core mantle boundary 2,900 kilometers below the surface. The new, high-resolution map of the mantle showed the plumes were connected to many volcanic hotspots at the surface of the earth. It also revealed that below about 1,000 km under the surface of the Earth the plumes are between 600 and 1,000 km across, up to five times wider than geophysicists thought and at least 400 degrees Celsius hotter than surrounding rock.



A 3D rendering of shear wave speed anomalies beneath the Central Pacific between the core-mantle boundary (2891 km depth) and 400 km depth. Green cones and labels highlight locations of key hotspot volcanoes in this region. *Image: Scott French*

RESEARCH DETAILS

To create a high-resolution computed tomography image of Earth, the researchers used very accurate numerical simulations of how seismic waves travel through the mantle. They then compared their predictions to the ground motion measured by detectors around the globe. They mapped mantle plumes by analyzing the paths of seismic waves bouncing around Earth's interior after 273 strong earthquakes that shook the globe over the past 20 years.

Earlier attempts by other research groups often approximated the physics of wave propagation and focused mainly on the arrival times of only certain types of seismic waves, such as the P (pressure) and S (shear) waves, which travel at different speeds. In this case, the research team computed all components of the seismic waves, such as their scattering and diffraction, then tweaked the model repeatedly to fit recorded data using a method similar to statistical regression. The final computation required 3 million CPU hours on NERSC's Edison system.

Publication:
S.W. French, B. Romanowicz, "Broad plumes rooted at the base of the Earth's mantle beneath major hotspots," *Nature* 525, 95-99, September 3, 2015, doi:10.1038/nature14876

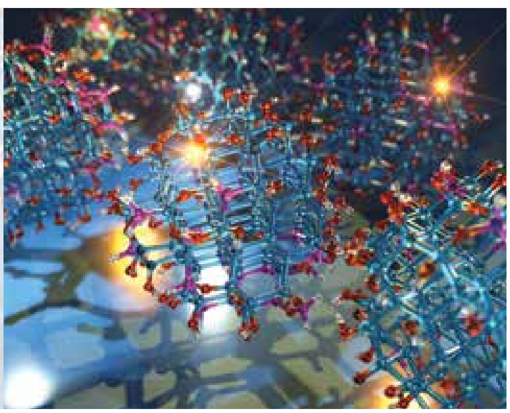
Full Story:
http://bit.ly/NERSCarMantlePlumes

What the Blank Makes Quantum Dots Blink?

Large-Scale Calculations on Nanostructured Heterogeneous Interfaces

BES—Materials Sciences

Principal Investigator:
Márton Vörös, University of Chicago



Silicon quantum dots are shown in various states of "blinking." The "on" crystals emit light (represented by a white dot) as an excited electron sheds excess energy as a photon. The "off" crystals are dark because their electrons (yellow) are trapped in surface defects and siphon off energy through other paths, such as heat or lattice vibrations. *Image: Peter Allen, University of Chicago*

Publication:
N.P. Brawand, M. Vörös, G. Galli, "Surface dangling bonds are a cause of B-type blinking in Si nanoparticles," *Nanoscale*, 2015, 7, 3737-3744, doi: 10.1039/C4NR06376G

Full Story:
http://bit.ly/NERSCarQuantumDots

OBJECTIVE

Quantum dots are nanoparticles of semiconductor that can be tuned to glow in a rainbow of colors. Since their discovery in the 1980s, these nanoparticles have held out tantalizing prospects for all kinds of new technologies, from paint-on lighting materials and solar cells to quantum computer chips, biological markers and even lasers and communications technologies.

Unfortunately, quantum dots often "blink," and this "fluorescence intermittency" has put a damper on many potential applications; for example, lasers and logic gates won't work very well with iffy light sources. Quantum dots can also absorb specific colors of light, but using them to harvest sunlight in photovoltaics is not yet very efficient, due in part to the mechanisms behind blinking.

FINDINGS/ACCOMPLISHMENTS

University of Chicago scientists computing at NERSC used *ab initio* simulations to probe the mysterious blinking process in silicon quantum dots. Their findings could help scientists fix the problem and pave the way for a number of new applications for this technology. The results are the first reported *ab initio* calculations showing that dangling bonds on the surface of oxidized silicon nanoparticles can act as efficient non-radiative recombination centers. These techniques could also be used to tackle the effects of trapping in solar cells, which can limit the efficiency of solar cells.

RESEARCH DETAILS

The team used simulated silicon (Si) nanoparticles configured with various defects and coated with silicon dioxide. Starting with three different possible defect states, they used the Hopper supercomputer to calculate the optical and electronic properties of the oxidized silicon nanoparticle with the scientific package called Quantum Espresso.

To perform their calculations, which took 100,000 processor hours on Hopper, the team first constructed virtual models. They computationally carved virtual holes out of a crystalline silicon oxide (SiO_2) matrix and inserted silicon quantum dots of various sizes, computing cycles of annealing and cooling to create a more realistic interface between the quantum dots and the SiO_2 matrix. Finally, dangling bond defects were introduced at the surface of quantum dots by removing a few selected atoms.

By computing the electronic properties and the rate at which electrons release energy, they found that trapped states do indeed cause quantum dot dimming. Dangling bonds on the surface of silicon nanoparticles trapped electrons where they recombined "non-radiatively" by releasing heat. That is, the electrons shed excess energy without radiating light. Dimming also depended on the overall charge of the entire quantum dot, the team found.

Stabilizing a Tokamak Plasma

3D Extended MHD Simulation of Fusion Plasmas

FES—Fusion SciDAC

OBJECTIVE

To use 3D simulations to gain new insights into the behavior of fusion plasmas and find improved methods for ensuring plasma stability in future tokamak reactors.

FINDINGS/ACCOMPLISHMENTS

A team of physicists from Princeton Plasma Physics Laboratory, General Atomics and the Max Planck Institute for Plasma Physics used NERSC's Edison computer to discover a mechanism that prevents the electrical current flowing through fusion plasma from repeatedly peaking and crashing. This behavior is known as a "sawtooth cycle" and can cause instabilities within the plasma's core.

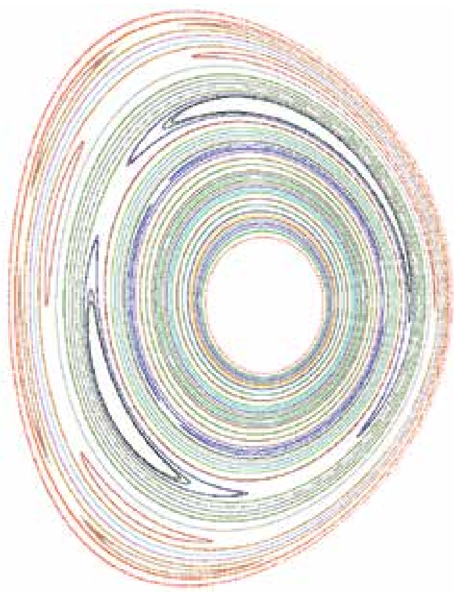
Learning how to model and study the behavior of fusion plasmas has important implications for ITER, the multinational fusion facility being constructed in France to demonstrate the practicality of fusion power. Instabilities in the plasma could destabilize and halt the fusion process. Preventing the destabilizing cycle from starting would therefore be highly beneficial for the ITER experiment and future tokamak reactors.

RESEARCH DETAILS

Running M3D-C1—a program that creates 3D simulations of fusion plasmas—on Edison, researchers found that under certain conditions a helix-shaped whirlpool of plasma forms around the center of the tokamak. The swirling plasma acts like a dynamo—a moving fluid that creates electric and magnetic fields. Together these fields prevent the current flowing through plasma from peaking and crashing.

During the simulations the scientists were able to virtually add new diagnostics, or probes, to the computer code, which measure the helical velocity fields, electric potential, and magnetic fields to clarify how the dynamo forms and persists. The persistence produces the voltage in the center of the discharge that keeps the plasma current from peaking.

Principal Investigator:
Stephen Jardin, Princeton Plasma Physics Laboratory



A cross-section of the virtual plasma showing where the magnetic field lines intersect the plane. The central section has field lines that rotate exactly once. *Image: Stephen Jardin, Princeton Plasma Physics Laboratory*

Publication:
S. C. Jardin, N. Ferraro, and I. Krebs, "Self-Organized Stationary States of Tokamaks," *Phys. Rev. Lett.* 115, 215001, November 17, 2015; DOI: http://dx.doi.org/10.1103/PhysRevLett.115.215001

Full Story:
http://bit.ly/NERSCarTokamakPlasma

Experiments + Simulations = Better Nuclear Power Research

Radiation Effects on the Mechanical Behavior of Ceramics

BES—Materials Science

Principal Investigator:
Ram Devanathan, Pacific Northwest National Laboratory

Publication:
C.A. Yablinsky, R. Devanathan, J. Pakarinen, J. Gan, D. Severin, C. Trautmann, T.R. Allen, "Characterization of swift heavy ion irradiation damage in ceria," *Journal of Materials Research* 30, 9, May 14, 2015. doi: 10.1557/jmr.2015.43

Full Story:
http://bit.ly/NERSCarNuclearPower

OBJECTIVE

An international collaboration of physicists is working to improve the safety and economics of nuclear power by studying how various cladding materials and fuels used in reactors respond to radiation damage.

FINDINGS/ACCOMPLISHMENTS

The research team—which includes representatives from Los Alamos, Pacific Northwest and Idaho national laboratories as well as institutions in Germany and Belgium—used molecular dynamics simulations run at NERSC to identify atomic-level details of early-stage damage production in cerium dioxide (CeO_2), a surrogate material used in nuclear research to understand the performance of uranium dioxide (UO_2).

RESEARCH DETAILS

The researchers coupled swift heavy ion experiments and electron microscopy analysis with parallel simulations—one of the few instances of combining physical experiments and simulations for this kind of research. The physical experiments involved implanting CeO_2 with 940 MeV gold ions and performing very careful high-resolution electron microscopy studies to understand the structure of the damaged material. Because the evolution of radiation damage is controlled by events that take place on the nanosecond scale, the simulations—run on NERSC's Hopper system—helped fill gaps in experimental understanding by simulating these transient events that take place on very small time and length scales.

While the molecular dynamics code (DL_POLY) the researchers used is fairly common, the simulations themselves were very computationally intensive; the team simulated a system of 11 million atoms for a fairly long time, something that couldn't have been done without a massively parallel computing resource.

.

BISICLES Ice Sheet Model Quantifies Climate Change

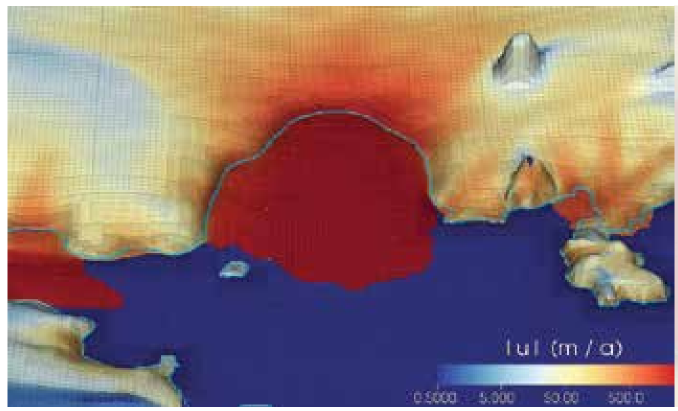
Predicting Ice Sheet and Climate Evolution at Extreme Scales

ASCR—Applied Mathematical Sciences

Principal Investigators:
Dan Martin, Esmond Ng, Lawrence Berkeley National Laboratory

OBJECTIVE

West Antarctica is one of the fastest warming regions on Earth, and its ice sheet has seen dramatic thinning in recent years. The ice sheet is losing significant amounts of ice to the ocean, with the losses not being offset by snowfall. The acceleration of West Antarctic ice streams in response to ocean warming could result in a major contribution to sea-level rise, but previous models were unable to quantify this.



Computer simulations run at NERSC show estimates of ice retreat in the Amundsen Sea Embayment by 2154. *Image: Stephen Cornford, University of Bristol*

FINDINGS/ACCOMPLISHMENTS

Using the high-resolution, large-scale computer model known as Berkeley-ISICLES (BISICLES), researchers were able to estimate how much ice the West Antarctic Ice Sheet could lose over the next two centuries and how that could impact global sea-level rise. These comprehensive, high-resolution simulations are a significant improvement from previous calculations (which were lower in resolution or scale) and will enable climate scientists to make more accurate predictions about West Antarctica's future.

RESEARCH DETAILS

The BISICLES adaptive mesh ice sheet model was used to carry out one, two, and three century simulations of the fast-flowing ice streams of the West Antarctic Ice Sheet, deploying sub-kilometer resolution around the grounding line since coarser resolution results in substantial underestimation of the response. The research team subjected the BISICLES model to a range of ocean and atmospheric change: no change at all, future changes projected by ocean and atmosphere models and extreme changes intended to study the upper reaches of future sea-level rise. The results of global climate models were fed into regional models of the Antarctic atmosphere and ocean, whose results were in turn used to force the ice-sheet model in this study. Some of the BISICLES simulations were run on NERSC's Hopper and Edison systems.

Publication:
S.L. Cornford, D.F. Martin, et al, "Century-scale simulations of the response of the West Antarctic Ice Sheet to a warming climate," *The Cryosphere,* 9, 1579-1600, 2015, doi: 10.5194/tc-9-1579-2015

Full Story:
http://bit.ly/NERSCarIceSheetModel

Tall Clouds from Tiny Raindrops Grow

Evolution in Cloud Population Statistics of the Madden-Julian Oscillation

BER—Climate and Environmental Sciences

Principal Investigator:
Samson Hagos, Pacific Northwest National Laboratory

PNNL scientists used real-world observations to simulate how small clouds are likely to stay shallow, while larger clouds grow deeper because they mix with less dry air. *Image: NASA Johnson Space Center*

OBJECTIVE

"Big clouds get bigger while small clouds shrink" may seem like a simple concept, but the mechanisms behind how clouds are born, grow and die are surprisingly complex. So researchers from Pacific Northwest National Laboratory (PNNL) used supercomputing resources at NERSC to simulate tropical clouds and their interaction with the warm ocean surface and compare the simulations to real-world observations. Scientists need to understand clouds to effectively predict weather patterns, including the potential for droughts and floods.

FINDINGS/ACCOMPLISHMENTS

By running simulations on NERSC's Hopper Cray XE6 system, the PNNL researchers found that factors as small as how sizes of raindrops were represented in a computer model made a big difference in the accuracy of the results. The simulations clearly showed that larger clouds tend to grow larger because they capture less dry air, while smaller clouds dwindle away.

RESEARCH DETAILS

The PNNL team collaborated with researchers from NASA Goddard Space Flight Center and the National Center for Atmospheric Research to run high-resolution regional models that simulated cloud lifecycles with different ranges of raindrop size. They compared those results to observational data collected from the Atmospheric Radiation Measurement Climate Research Facility's Madden-Julian Oscillation Investigation Experiment and the Earth Observing Laboratory's Dynamics of the Madden-Julian Oscillation research campaign, which collected data in the South Pacific from October 2011 to March 2012 to support studies on the birth, growth and evolution of certain types of clouds.

The scientists used satellite and ground-based radar measurements from the campaign to examine how well the Weather Research and Forecasting Model simulated tropical clouds. They looked specifically at four approaches to handling the physics of tiny raindrops and found that two key factors were rain rates at ground level and the modeling of "cold pools"—the cold, dry air flowing out from deep thunderstorms.

Publication:
S. Hagos, Z. Feng, C. Burleyson, K-S. Lim, C. Long, D. Wu, G. Thompson, "Evaluation of Convection-Permitting Model Simulations of Cloud Populations Associated with the Madden-Julian Oscillation using Data Collected during the AMIE/DYNAMO Field Campaign," *Journal of Geophysical Research: Atmospheres* 119(21):12,052-12,068. doi:10.1002/2014JD022143

Full Story:
http://bit.ly/NERSCarCloudScience

Supernova Hunting with Supercomputers

Palomar Transient Factory; Type Ia Supernovae

HEP—Cosmic Frontier

OBJECTIVE

Because the relative brightness of Type Ia supernovae can be measured so well no matter where they are located in the Universe, they make excellent distance markers. In fact, they were instrumental to measuring the accelerating expansion of the Universe in the 1990s—a finding that netted three scientists the 2011 Nobel Prize in Physics. Yet, astronomers still do not fully understand where they come from.

FINDINGS/ACCOMPLISHMENTS

Using a "roadmap" of theoretical calculations and supercomputer simulations performed at NERSC, astronomers observed for the first time a flash of light caused by a supernova slamming into a nearby star, allowing them to determine the stellar system—IC831—from which the supernova—iPTF14atg—was born. This finding confirms one of two competing theories about the birth of Type Ia supernovae: the single-degenerate model.

The intermediate Palomar Transient Factory (iPTF) analysis pipeline found the light signal from the supernova just hours after it ignited in a galaxy about 300 million light years away from Earth. iPTF depends on NERSC computational and global storage resources.

RESEARCH DETAILS

On May 3, 2014, the iPTF took images of IC831 and transmitted the data for analysis to computers at NERSC, where a machine-learning algorithm analyzed the images and prioritized real celestial objects over digital artifacts, leading to the discovery of the young supernova.

Principal Investigators:
Peter Nugent, Lawrence Berkeley National Laboratory; Stan Woosley, University of California, Santa Cruz

Simulation of the expanding debris from a supernova explosion (shown in red) running over and shredding a nearby star (shown in blue). *Image: Daniel Kasen, Lawrence Berkeley National Laboratory/ UC Berkeley*

Publication:
Y. Cao, S.R. Kulkarni, D. Andrew Howell, et al, "A strong ultraviolet pulse from a newborn type 1a supernova," *Nature* 421, 328-331, May 21, 2015, doi: 10.1038/nature14440

Full Story:
http://bit.ly/NERSCarSupernovaHunting

Petascale Pattern Recognition Enhances Climate Science

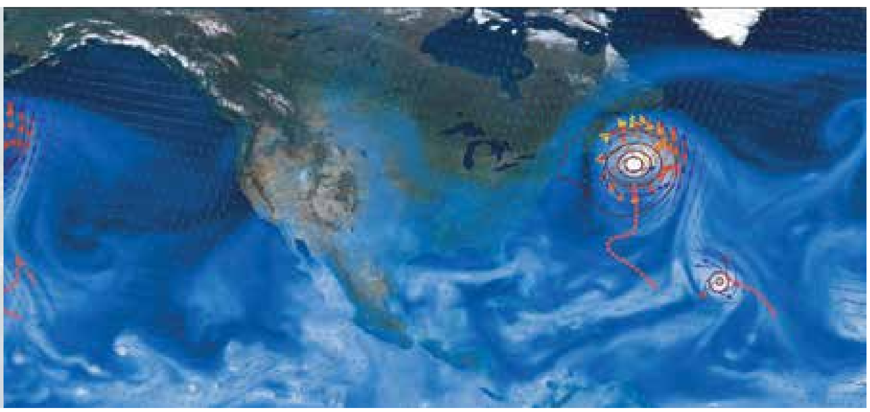
Calibrated and Systematic Characterization, Attribution and Detection of Extremes

BER—Climate and Environmental Sciences

Principal Investigator:
Travis O'Brien, Lawrence Berkeley National Laboratory

OBJECTIVE

To demonstrate how a new data analytics tool developed at Berkeley Lab uses pattern recognition to help climate researchers better detect extreme weather events, such as cyclones and atmospheric rivers, in large datasets.



TECA implements multivariate threshold conditions to detect and track extreme weather events in large climate datasets. This visualization depicts tropical cyclone tracks overlaid on atmospheric flow patterns. *Image: Prabhat, Lawrence Berkeley National Laboratory*

FINDINGS/ACCOMPLISHMENTS

Modern climate simulations produce massive amounts of data, requiring sophisticated pattern recognition algorithms to be run on terabyte- to petabyte-sized datasets. Using the new Toolkit for Extreme Climate Analysis (TECA), researchers from Berkeley Lab and Argonne National Laboratory demonstrated how TECA, running at full scale on NERSC's Hopper system and Argonne's Mira system, reduced the runtime for pattern detection tasks from years to hours.

"TECA: Petascale Pattern Recognition for Climate Science," a paper presented by the team at the 16th International Conference on Computer Analysis of Images and Patterns (CAIP) in September 2015, was awarded CAIP's Juelich Supercomputing Center prize for the best application of HPC technology in solving a pattern recognition problem.

RESEARCH DETAILS

For this project, the researchers downloaded 56TB of climate data from the fifth phase of the Coupled Model Intercomparison Project (CMIP5) to NERSC. Their goal was to access subsets of those datasets to identify three different classes of storms: tropical cyclones, atmospheric rivers and extra-tropical cyclones. All of the datasets were accessed through a portal created by the Earth Systems Grid Federation to facilitate the sharing of data; in this case, the data consisted of atmospheric model data sampled at six-hour increments running out to the year 2100. The data was stored at 21 sites around the world, including Norway, the United Kingdom, France, Japan and the U.S. NERSC's Hopper system was used to preprocess the data, which took about two weeks and resulted in a final 15TB dataset.

Publication:
"TECA: Petascale Pattern Recognition for Climate Science," 16th International Conference, CAIP 2015, Valletta, Malta, September 2-4, 2015

Full Story:
http://bit.ly/NERSCarPatternRecognition

Celeste: A New Model for Cataloging the Universe

MANTISSA: Massive Acceleration of New Techniques in Science with Scalable Algorithms

ASCR—Applied Math

OBJECTIVE

To demonstrate how Celeste—a new statistical analysis model developed by a collaboration of astrophysicists, statisticians and computer scientists from UC Berkeley, Harvard, Carnegie Mellon and Berkeley Lab—can automate digital sky surveys and make it possible to create a single catalog of the entire universe. Celeste is designed to be a better model for identifying the astrophysical sources in the sky and the calibration parameters of each telescope. It also has the potential to significantly reduce the time and effort that astronomers currently spend working with image data.

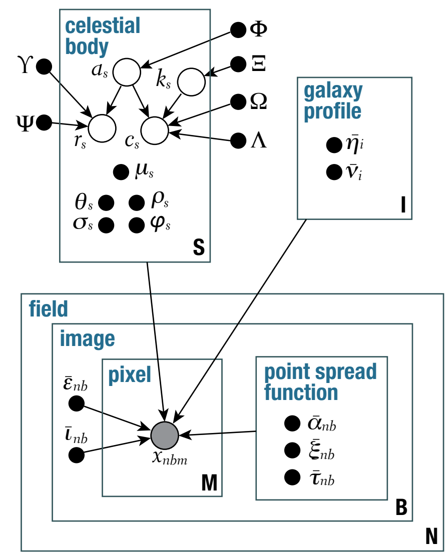
FINDINGS/ACCOMPLISHMENTS

The research team spent much of 2015 developing and refining Celeste, a hierarchical model designed to catalog stars, galaxies and other light sources in the universe visible through the next generation of telescopes. It also enables astronomers to identify promising galaxies for spectrograph targeting, define galaxies they may want to explore further and help them better understand dark energy and the geometry of the universe. Celeste uses analytical techniques common in machine learning and applied statistics but not so much in astronomy, implementing a statistical inference to build a fully generative model that can mathematically locate and characterize light sources in the sky.

RESEARCH DETAILS

The researchers have so far used Celeste to analyze pieces of Sloan Digital Sky Survey (SDSS) images, whole SDSS images and sets of SDSS images on NERSC's Edison supercomputer. These initial runs helped them refine and improve the model and validate its ability to exceed the performance of current state-of-the-art methods for locating celestial bodies and measuring their colors. The next milestone will be to run an analysis of the entire SDSS dataset all at once at NERSC. The researchers will then begin adding other datasets and building the catalog—which, like the SDSS data, will be housed on a science gateway at NERSC. In all, the Celeste team expects the catalog to collect and process some 500TB of data, or about 1 trillion pixels.

Principal Investigator:
Prabhat, Lawrence Berkeley National Laboratory



The Celeste graphical model. Shaded vertices represent observed random variables. Empty vertices represent latent random variables. Black dots represent constants.

Publication:
J. Regier, A. Miller, J. McAuliffe, R. Adams, M. Hoffman, D. Lang, D. Schlegel, Prabhat, "Celeste: Variational inference for a generative model of astronomical images," July 2015, 32nd International Conference on Machine Learning, Lille, France, *JMLR: W&CP* Volume 37.
J. Regier, J. McAuliffe, Prabhat, "A deep generative model for astronomical images of galaxies," *Neural Informational Processing Systems (NIPS) Workshop: Advances in Approximate Bayesian Inference.* 2015

Full Story:
http://bit.ly/NERSCarCeleste

Supercomputers Speed Search for New Subatomic Particles

Quantum Chromodynamics with Four Flavors of Dynamical Quarks

HEP—Lattice Gauge Theory

Principal Investigator:
Douglas Toussaint, University of Arizona

Artist's rendering of a rare B-meson "penguin" showing the quark-level decay process. *Image: Daping Du, Syracuse University*

OBJECTIVE

After being produced in a collision, subatomic particles spontaneously decay into other particles, following one of many possible decay paths. Out of 1 billion B mesons detected in a collider, only about 20 decay through this particular process. With the discovery of the Higgs boson, the last missing piece, the Standard Model of particle physics now accounts for all known subatomic particles and correctly describes their interactions. But scientists know that the Standard Model doesn't tell the whole story, and they are searching for evidence of physics beyond the Standard Model.

FINDINGS/ACCOMPLISHMENTS

A team of physicists in the Fermilab Lattice and MILC Collaboration has developed a new, high-precision calculation that could significantly advance the search for physics beyond the Standard Model.

RESEARCH DETAILS

The new calculation, which employs lattice quantum chromodynamics (QCD), applies to a particularly rare decay of the B meson (a subatomic particle), sometimes also called a "penguin decay" process. The team ran the calculation on several supercomputers, including NERSC's Edison and Hopper systems.

Generating the lattices and characterizing them is a time-intensive effort, noted Doug Toussaint, professor of physics at the University of Arizona and the PI who oversees this project's allocation at NERSC. The research team used Monte Carlo simulations to generate sample configurations of fields for QCD, then calculate numerous things by averaging the desired quantity over these sample configurations, or lattices, and use these stored lattices to calculate many different strong interaction masses and decay rates.

Publication:
J.A. Bailey, A. Bazavovo, C. Bernard, C.M. Bouchard, et al, "*B→πll* Form Factors for New Physics Searches from Lattice QCD," *Physics Review Letters* 115, 152002, October 7, 2015, doi: http://dx.doi.org/10.1103/PhysRevLett.115.152002

Full Story:
http://bit.ly/NERSCarPenguinDecay

What Causes Electron Heat Loss in Fusion Plasma?

Simulations of Field-Reversed Configuration and Other Compact Tori Plasmas

FES—Fusion Base Program

OBJECTIVE

Creating controlled fusion energy entails many challenges, but one of the most basic is heating plasma to extremely high temperatures and then maintaining those temperatures. Researchers at Princeton Plasma Physics Laboratory (PPPL) proposed an explanation for why the hot plasma within fusion facilities called tokamaks sometimes fails to reach the required temperature, even as beams of fast-moving neutral atoms are pumped into the plasma in an effort to make it hotter.

FINDINGS/ACCOMPLISHMENTS

3D simulations run at NERSC yielded new insights into why plasma in a tokamak fusion reactor sometimes fails to reach required temperatures. The findings enhance our understanding of electron energy transport in fusion experiments and could lead to improved temperature control in fusion devices such as ITER, the international fusion facility currently under construction in France.

RESEARCH DETAILS

The PPPL team focused on the puzzling tendency of electron heat to leak from the core of the plasma to the plasma's edge. While performing 3D simulations of past NSTX (National Spherical Tokamak Experiment) plasmas on computers at NERSC, they saw that two kinds of waves found in fusion plasmas appear to form a chain that transfers the neutral-beam energy from the core of the plasma to the edge, where the heat dissipates. While physicists have long known that the coupling between the two kinds of waves –known as compressional Alfvén waves and kinetic Alfvén waves–can lead to energy dissipation in plasmas, the PPPL team's findings were the first to demonstrate the process for beam-excited compressional Alfvén eigenmodes in tokamaks.

Principal Investigators:
Ronald Davidson, Elena Belova,
Princeton Plasma Physics Laboratory

NSTXU interior vacuum vessel. *Image: Elle Starkman/ PPPL Communications*

Publication:
E.V. Belova, N.N. Gorelenkov, E.D. Fredrickson, K. Tritz, N.A. Crocker, "Coupling of neutral-beam-driven compressional Alfven Eigenmodes to kinetic Alfven waves in NSTX tokamak and energy channeling," *Physics Review Letters* 115, 015001, June 29, 2015, doi: http://dx.doi.org/10.1103/PhysRevLett.115.015001

Full Story:
http://bit.ly/NERSCarFusionPlasma

Documenting CO_2's Increasing Greenhouse Effect at Earth's Surface

Center at LBNL for Integrative Modeling of the Earth System (CLIMES)

BER—Climate and Environmental Sciences

Principal Investigator:
William Collins, Lawrence Berkeley National Laboratory

Data for this study was collected using spectroscopic instruments at two sites operated by DOE's Atmospheric Radiation Measurement Climate Research Facility, including this research site on the North Slope of Alaska near the town of Barrow. *Image: Jonathan Gero*

OBJECTIVE

The influence of atmospheric CO_2 on the balance between incoming energy from the Sun and outgoing heat from the Earth is well established. But this effect had not been experimentally confirmed outside the laboratory.

FINDINGS/ACCOMPLISHMENTS

By measuring atmospheric CO_2's increasing capacity to absorb thermal radiation emitted from the Earth's surface over an 11-year period at two locations in North America, researchers for the first time observed an increase in CO_2's greenhouse effect at the Earth's surface. The team—which included researchers from Berkeley Lab, the University of Wisconsin-Madison and PNNL—attributed this upward trend to rising CO_2 levels from fossil fuel emissions. The measurements also enabled the scientists to detect, for the first time, the influence of photosynthesis on the balance of energy at the surface. They found that CO_2-attributed radiative forcing dipped in the spring as flourishing photosynthetic activity pulled more of the greenhouse gas from the air.

RESEARCH DETAILS

The research collaboration measured atmospheric CO_2's contribution to radiative forcing at two sites, one in Oklahoma and one on the North Slope of Alaska, from 2000 to the end of 2010. Using NERSC's Hopper and Edison systems and the Community Earth System Model, the team used about 5,000 CPU hours to perform tens of thousands of calculations to analyze the data. They ran a large number of radiative transfer calculations that turned out to be very parallel, so they then used the taskfarmer utility for some 500 calculations at a time, with each lasting about 10 minutes on a single core.

Publication:
D.R. Feldman, W.D. Collins, P.J. Gero, M.S. Torn, E.J. Mlawer, T.R. Shippert, "Observational determination of surface radiative forcing by CO_2 from 2000 to 2010," *Nature*, February 25, 2015, doi: 10.1038/nature14240

Full Story:
http://bit.ly/NERSCarGreenhouseEffect

What Ignites a Neutron Star?

Three-dimensional Studies of Convection in X-ray Bursts

NP—Astrophysics

OBJECTIVE

To use multi-dimensional simulations to gain new insights into the physics inside a neutron star and the behavior of dense nuclear matter by studying convection in Type I X-ray bursts—the thermonuclear "runaway" in a thin hydrogen/helium layer on the star's surface—thereby enhancing our understanding of how dense nuclear matter and neutron stars behave.

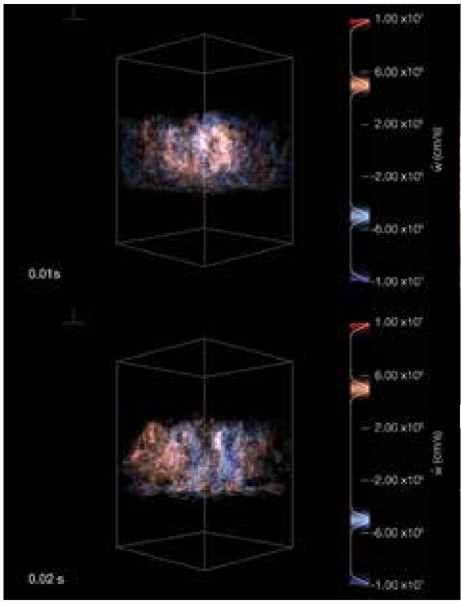
FINDINGS/ACCOMPLISHMENTS

Astrophysicists from Stony Brook University, Los Alamos National Laboratory and Berkeley Lab used supercomputing resources at NERSC to set the stage for the first detailed 3D simulations of convective burning in an X-ray burst. The research builds on the team's previous 2D X-ray burst studies at NERSC using MAESTRO, a hydrodynamics code developed at Berkeley Lab. One-dimensional hydrodynamic studies have been able to reproduce many of the observable features of X-ray bursts, such as burst energies and recurrence times. But multi-dimensional simulations are necessary to more fully understand the convection process in the extreme conditions found in a neutron star.

RESEARCH DETAILS

The research team used 5 million CPU hours on NERSC's Edison system as an initial testbed to explore multiple models and parameter space. They then used this information to run 3D simulations of the X-ray bursts at the Oak Ridge Leadership Computing Facility, comparing the simulations to previously created 2D simulations. Using the MAESTRO code, they modeled a small box on the surface of a neutron star, not the entire star; specifically, they modeled tens of meters on a side and 10 meters of depth on the surface of the star, which was enough to resolve features only a few centimeters in length, to see the burning and to see that it drives the convection. As a follow-on, the team is running a new class of simple model problems at NERSC that uses a more realistic initial model and reaction network.

Principal Investigator:
Michael Zingale, Stony Brook University



Volume renderings of the vertical velocity field at t=0.01 s (top) and 0.02 s (bottom) for the wide calculation. Upward moving fluid is in red and downward moving is blue. *Image: Michael Zingale, Stony Brook University*

Publication:
M. Zingale, C.M. Malone, A. Nonaka, A.S. Almgren, J.B. Bell, "Comparisons of two- and three-dimensional convection in Type I x-ray bursts," *Astrophysical Journal* 807:60, July 1, 2015, doi: 10.1088/0004-637X/807/1/60

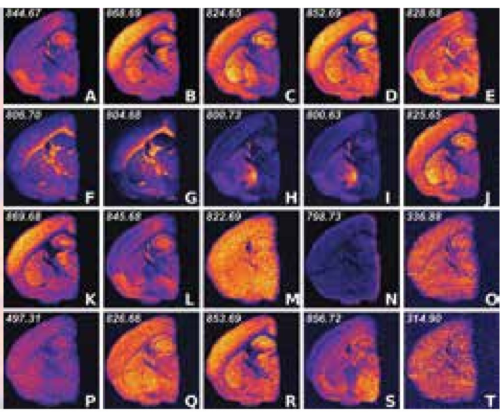
Full Story:
http://bit.ly/NERSCarXrayBursts

'Data Deluge' Pushes MSI Analysis to New Heights

MANTISSA: Massive Acceleration of New Techniques in Science with Scalable Algorithms

ASCR—Applied Math, BER—Biological Systems Science

Principal Investigators:
Prabhat, Lawrence Berkeley National Laboratory; Ben Bowen, University of California, Berkeley



Ion-intensity visualization of the 20 most important ions in a mouse brain segment selected by the CX/CUR algorithm. Of the 20 ions, little redundancy is present, pointing to the effectiveness of the CX approach for information prioritization. *Image: Ben Bowen, Lawrence Berkeley National Laboratory*

OBJECTIVE

Researchers are combining mathematical theory and scalable algorithms with computational resources at NERSC to address growing data-management challenges in climate research, high-energy physics and the life sciences.

FINDINGS/ACCOMPLISHMENTS

To develop better methods for analyzing large datasets, in 2014 Berkeley Lab launched a project dubbed MANTISSA (Massive Acceleration of New Technologies in Science with Scalable Algorithms). MANTISSA's first order of business was to improve data analysis in mass spectrometry imaging (MSI). Datasets from MSI provide unprecedented characterization of the amount and spatial distribution of molecules in a sample, leading to detailed investigations of metabolic and microbial processes at subcellular to centimeter resolutions. But the raw datasets range from gigabytes to terabytes in size, which can be a challenge for even large computing systems to wade through efficiently.

Studies applying MANTISSA's algorithms to MSI datasets have demonstrated how a mathematical theory such as randomized linear algebra (RLA)—commonly used in engineering, computational science and machine learning for image processing and data mining—can be combined with one of the most sophisticated imaging modalities in life sciences to enhance scientific progress.

RESEARCH DETAILS

Researchers from UC Berkeley and Berkeley Lab ran computations involving two RLA algorithms—CX and CUR—on NERSC's Edison system using two MSI datasets: mammalian brain and lung sections. They found that using these algorithms streamlined the data analysis process and yielded results that were easier to interpret because they prioritized specific elements in the data.

Publication:
J. Yang, O. Rubel, Prabhat, M. Mahoney, B.P. Bowen, "Identifying important ions and positions in mass spectrometry imaging data using CUR matrix decompositions," *Analytical Chemistry,* 87 (9), 4658-4666, March 31, 2015, doi: 10.1021/ac5040264

Full Story:
http://bit.ly/NERSCarMSIDatasets

Could Material Defects Actually Improve Solar Cells?

Doping bottleneck, van der Waals Interaction and Novel Ordered Alloy

BES—Materials Science

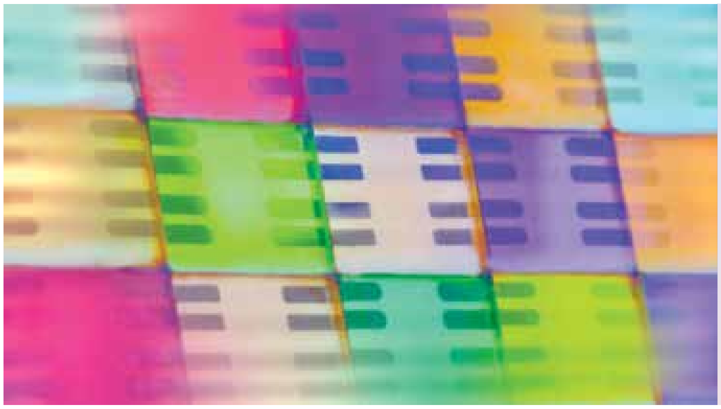
OBJECTIVE

Scientists at the National Renewable Energy Laboratory (NREL) are using supercomputers to study what may seem paradoxical: that certain defects in silicon solar cells may actually improve their performance.

Principal Investigator:
Suhuai Wei, National Renewable Energy Laboratory

FINDINGS/ACCOMPLISHMENTS

Deep-level defects frequently hamper the efficiency of solar cells, but this theoretical research suggests that defects with properly engineered energy levels can improve carrier collection out of the cell or improve surface passivation of the absorber layer. For solar cells and photoanodes, engineered defects could possibly allow thicker, more robust carrier-selective tunneling transport layers or corrosion protection layers that might be easier to fabricate.



Deep-level defects frequently hamper the efficiency of solar cells, but NREL research suggests that defects with properly engineered energy levels can improve carrier collection out of the cell. *Image: Roy Kaltschmidt, Lawrence Berkeley National Laboratory*

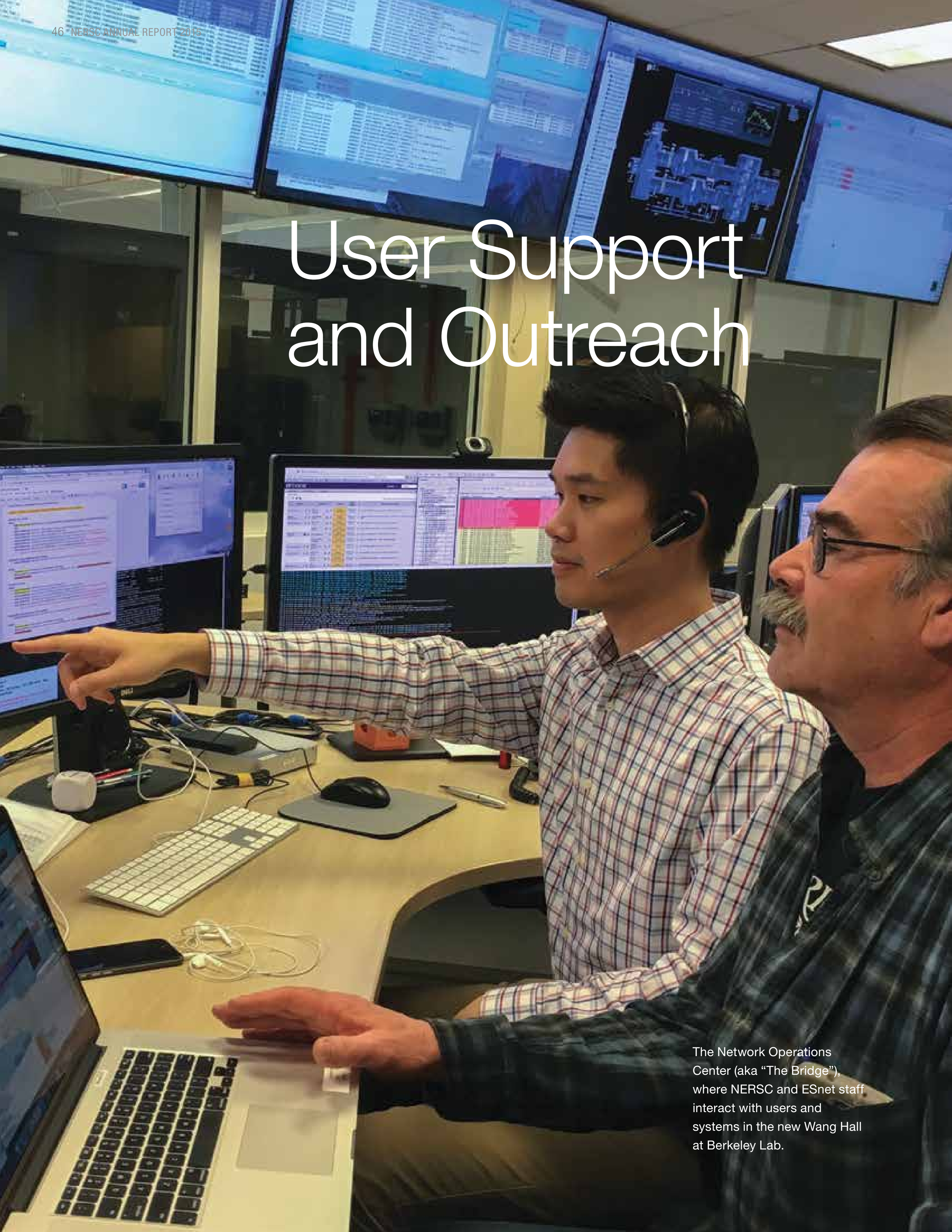
RESEARCH DETAILS

Researchers at NREL ran simulations to add impurities to layers adjacent to the silicon wafer in a solar cell. Namely, they introduced defects within a thin tunneling silicon dioxide layer that forms part of the "passivated contact" for carrier collection and within the aluminum oxide surface passivation layer next to the silicon (Si) cell wafer. In both cases, specific defects were identified to be beneficial. NERSC's Hopper system was used to calculate various defect levels; the researchers ran a total of 100 calculations on Hopper, with each calculation taking approximately eight hours on 192 cores.

Publication:
Y. Liu, P. Stradins, H. Deng, J. Luo, S. Wei, "Suppress carrier recombination by introducing defects: The case of Si solar cell," *Applied Physics Letters* 108, 022101, January 11, 2016.

Full Story:
http://bit.ly/NERSCarMaterialDefects

User Support and Outreach



The Network Operations Center (aka "The Bridge"), where NERSC and ESnet staff interact with users and systems in the new Wang Hall at Berkeley Lab.

NERSC has long been well regarded for the depth and breadth of its user support, and 2015 was no exception: in our annual user survey, the Overall Satisfaction with NERSC score was tied with our highest ever score in this category. One reason for this is that NERSC consultants and account support staff are available to users around the clock via email and an online web interface. Users worldwide can access basic account support (password resets, resetting login failures) online or via the NERSC operations staff 24 x 7, 365 days a year. In 2015, NERSC users hailed from 48 U.S. states and 47 countries.

With the center's reorganization in 2015, a cross-group set of consultants now provide the first line of support between NERSC and its active user community, which numbers about 6,000 and includes representatives from universities, national laboratories and industry. These staff are responsible for problem management and consulting, helping with user code optimization and debugging, strategic project support, web documentation and training, third-party applications and library support, running Office of Science-wide requirements reviews and coordinating NERSC's active users group.

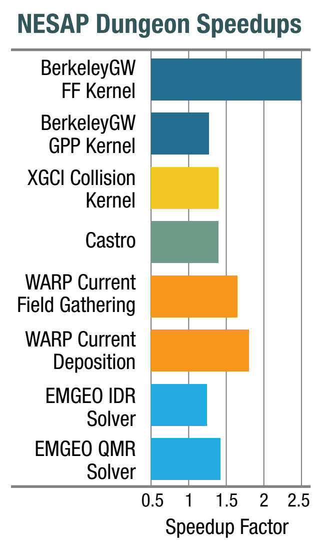
NESAP Begins to Bloom

Throughout 2015, NERSC expanded its efforts to help users prepare for the Cori Phase 2 system coming online in summer 2016. NERSC established the NERSC Exascale Scientific Applications Program (NESAP), a collaborative effort designed to give code teams and library and tool developers a unique opportunity to prepare for Cori's manycore architecture. NERSC selected 20 projects to collaborate with NERSC, Cray and Intel and access early hardware, special training and preparation sessions with Intel and Cray staff. In addition, another 24 projects, as well as library and tool developers, are participating in NESAP via NERSC training sessions and early access to prototype and production hardware.

One of the main goals of NESAP is to facilitate a venue for DOE Office of Science application teams to work productively with Intel and Cray engineers who specialize in hardware architecture, compiler and tool design. Through the NERSC-8 Cray Center of Excellence, Cray engineers are paired with NESAP code teams to profile and address a large number of application performance issues requiring code refactoring to improve data locality (such as tiling and cache blocking), thread scaling and vectorization.

After working closely with Cray engineers for months, the NESAP code teams are invited to attend a 2.5-day optimization "dungeon session" at Intel. These sessions generally focus on deep optimization efforts related specifically to the Intel Xeon Phi hardware. Hardware, compiler and tools experts from Intel are present to work closely with the code teams. In 2015 we completed four intensive optimization dungeon sessions at Intel with an average of three code teams at each; most teams achieved optimization of 1.5x to 2.5x in code speed-up. NESAP members have already gained a tremendous amount of experience optimizing codes for the Xeon Phi architecture as well as using Cray and Intel performance tools. NERSC has passed this knowledge onto NERSC users through regular training sessions and special events such as the Xeon Phi "hack-a-thons."

NERSC is in the process of hiring up to eight postdocs to assist in the NESAP project. To date, six postdocs have been hired and placed with NESAP teams. The first three postdocs arrived in mid 2015; the other three arrived in early 2016.



Code speedups observed during four NESAP "dungeon sessions" in 2015.

NERSC's NESAP Post-Docs, as of March 2016

NAME	NESAP PROJECT	START DATE
Brian Friesen	BoxLib	May, 2015
Andrey Ovsyannikov	Chombo Crunch	July 2015
Taylor Barnes	Quantum Espresso	July 2015
Tuomas Koskela	GTC	January 2016
Tareq Malas	EMGeo	January 2016
Mathieu Lobet	WARP	February 2016

NERSC Sponsors Trainings, Workshops, Hackathons

NERSC holds training events and workshops aimed at improving the skills of users of NERSC resources. Events are aimed at users of all skill levels, from novice to advanced.

In 2015 NERSC held two targeted trainings for users new to NERSC resources, one in conjunction with the annual NERSC User Group meeting, the other coinciding with the beginning of the school year. We also offered training for NERSC users in particular application areas. Users of the materials science applications VASP and Quantum Espresso had a training focused on those two applications in June, while users of BerkeleyGW participated in a 2.5-day workshop in November designed to help them use that code more efficiently.

In addition, NERSC held workshops to train users on specific code-development tools. A training specialist from Allinea presented a session on using the DDT debugger and the MAP profiler, a presenter from Rogue Wave software taught users to use the TotalView debugger and Intel taught classes on VTune and other tools.

The center also offered opportunities for hands-on development, beginning in February with the NUG Hackathon, where users were able to work with real tools under the guidance of NERSC and vendor staff. Working in teams, attendees learned how to add OpenMP to an application and how to vectorize code. Inspired by the success of this event, in July NERSC held a similar event at the Computational Science Graduate Fellowship Annual Review Workshop and at the IXPUG conference in September, drawing more than 30 participants to the latter event.

During the hackathons, NERSC staff were on hand to provide optimization strategies, help attendees use coding tools and answer questions. Attendees did the work of refactoring code to make it run faster, and remote viewers were able to follow along and pose questions via a chat interface. For example, many attendees had their first look at Intel's VTune Amplifier, a performance analysis tool targeted for users developing serial and multithreaded applications. One attendee used VTune to identify a hot loop in his code and also to determine that he is not memory-bandwidth bound. The code contained a large number of flops and instructions but was not being vectorized due to a loop dependence. Experts from NERSC and Intel helped him restructure the code so that the flop-intensive loop no longer contained the variable dependence. This change enabled the flop-heavy loop to vectorize and sped up his entire application run by 30 percent.

Other NERSC training events in 2015 covered programming concepts of particular utility for Cori. Ruud van der Pas, one of the leading OpenMP experts, taught a course on OpenMP, and Cray's Nathan Wichmann, an expert in application performance, provided guidance on optimizing for the Knights Landing architecture in a webinar. In total, NERSC hosted 14 training sessions during the year, some of which were full-day events.

NERSC also hosted a series of developer and user community workshops as part of our NESAP program. For example, in conjunction with IXPUG, we hosted a "Density Functional Theory (DFT) for Exascale" day targeted at the developer community. The meeting attracted speakers and attendees representing the top DFT codes in the world, including VASP, Quantum ESPRESSO, SIESTA, QBOX, PETOT, PARATEC, BerkeleyGW, WEST, Yambo and DG-DFT. Attendees came from all over the U.S. and Europe to attend the workshop and discuss advancements in DFT methods targeting exascale architectures, areas of commonality for collaboration and the development and use of optimized shared libraries.

In parallel with the DFT developer community meeting, NERSC hosted an accelerator modeling code developer meeting targeted at porting particle accelerator codes to Xeon Phi. The workshop had 32 attendees from LBNL, LLNL, SLAC, Fermilab, UCLA and industry.

NERSC also hosted the third annual BerkeleyGW workshop, bringing developers and approximately 50 users together for a 2.5-day intensive training session on advanced materials science simulation using BerkeleyGW and NERSC HPC resources.

IXPUG Annual Meeting a Hit

NERSC's newest supercomputer, Cori, is a Cray XC system based on Intel's Xeon Phi Knights Landing processor. Most applications currently being used at NERSC are expected to require additional code development to fully take advantage of the computational power of the new manycore Knights Landing processor. Cori Phase 2 is expected to arrive in the summer of 2016, and other DOE national laboratories, including Argonne, Los Alamos and Sandia, will also be installing Knights Landing-based systems in the next year or so.

Because of the challenges and potential of the Xeon Phi processors, representatives from research institutions that will be deploying systems using the processor formed the Intel Xeon Phi Users Group (IXPUG), an independent users group that provides a forum for exchanging ideas and information.

NERSC hosted the IXPUG annual meeting September 28 - October 2, 2015. The event drew more than 100 members of the HPC community from around the world, who spent four days preparing for the Knights Landing architecture.

The first day consisted of workshops in which there were hands-on opportunities to optimize a code kernel or use advanced MPI and OpenMP. On the second day, there were talks about the Xeon Phi roadmap, the preparations for manycore architectures being taken by many HPC facilities and

More than 100 people attended the first IXPUG meeting, held September 28 – October 2, 2015 in Berkeley Lab's Wang Hall.

experiences with developing for the Xeon Phi. The third day was a workshop on tools and techniques for optimizing for the Xeon Phi, and the fourth day featured workshops on libraries and programming models for the Xeon Phi.

Application Portability Efforts Progress

Having an extremely broad set of applications running on our systems, NERSC is very aware that our users, as well as the applications that run on our systems, also run on systems at other large-scale computing facilities, such as the Leadership Computing Facilities at Oak Ridge or Argonne, NCAR, National Science Foundation computing facilities such as NCSA and TACC, as well as at international computing centers. Thus it is important that applications run well across multiple systems and architectures. The present challenge is that there are two architectural paths on the exascale trajectory, which, while converging slowly, nevertheless present portability challenges. For example, NERSC and the ALCF are both procuring Cray-Intel-based manycore architectures while the OLCF is procuring an IBM-NVIDIA based system.

The Office of Science ASCR facilities have been collaborating and discussing best practices for porting applications to advanced architectures since March 2014. The first meeting included ALCF, OLCF and NERSC exchanging information on each other's systems and tools. By the second meeting, participants had expanded to include the NNSA labs and vendors providing briefings on their toolsets. Finally, in the latter half of 2015, two larger workshops were held: the HPC Operational Review (with the topic of application performance portability) and an HPC Portability Workshop at the SC15 conference.

Going forward, NERSC is partnering with ALCF and OLCF to examine two real applications to assess how easy or difficult it is to achieve application performance portability across architectures. The goal will be to use the same code base on two different architectures. We expect this effort to expand in 2016 and 2017 as prototype hardware becomes available.

Preliminary Findings from the HPC Operational Review

Current Practices	Emerging Practices	Opportunities
Applications truly running on both GPU and CPU architectures today create a code structure that allows a "plug-and-play" of architecture-specific modules	Use frameworks like RAJA, KOKKOS or TIDA to enable portability	Train early career staff on multi-disciplinary computational science areas (performance, applications, architecture and algorithms)
No applications currently use the exact code base and run with performance across GPU and CPU systems	Consider using domain-specific language	Use libraries to encapsulate architecture specific optimizations
Libraries that have been optimized for a particular architecture should be used when possible	Use LLVM as a backend code generator/compiler/translator	Use OpenMP 4.x as a standard that is portable across architectures
Memory management should be abstracted in code so it can easily be replaced	Use autotuning to find best parameters for different architectures	Adopt OpenACC features into OpenMP and focus investments into one standard
Write modular, layered code without vendor-specific compiler optimizations	Architecting code in a tasking model	Use DSLs, compilers, translators and code generators to achieve architecture tuned performance
Pay careful attention to data structure layout(s), which improves data locality and can improve portability on many architectures		

Data and Analytics Support

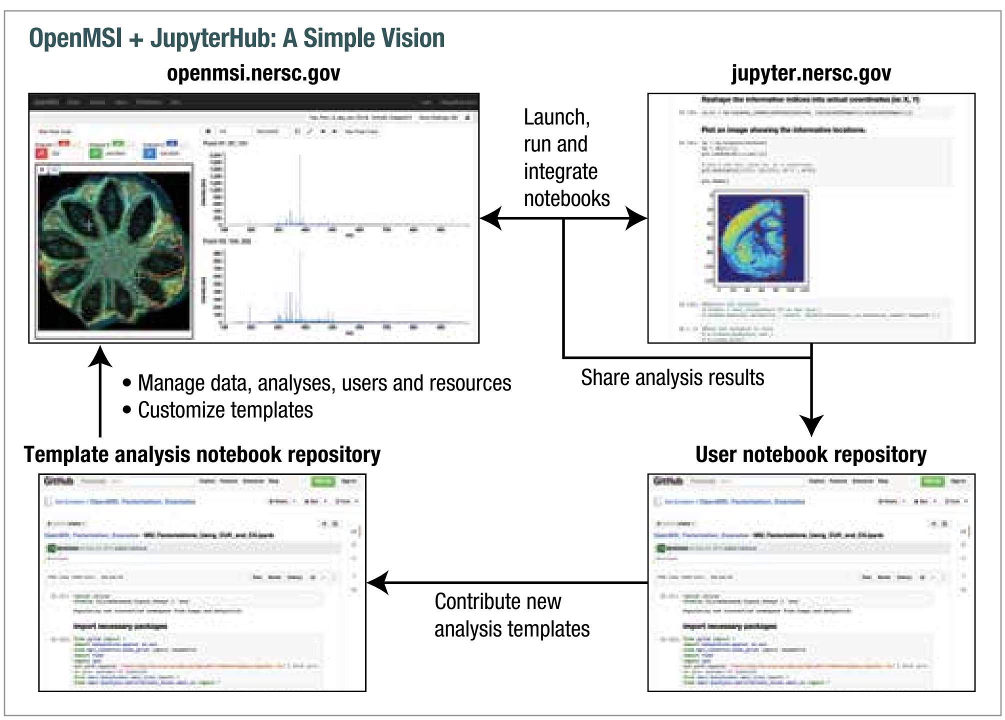
OPENMSI + IPYTHON/JUPYTER = INTERACTIVE COMPUTING

In 2015 NERSC provided user support to a number of project teams focused on data-intensive science. Some involve tools accessible to the broad NERSC community, while others were tight collaborations with particular projects.

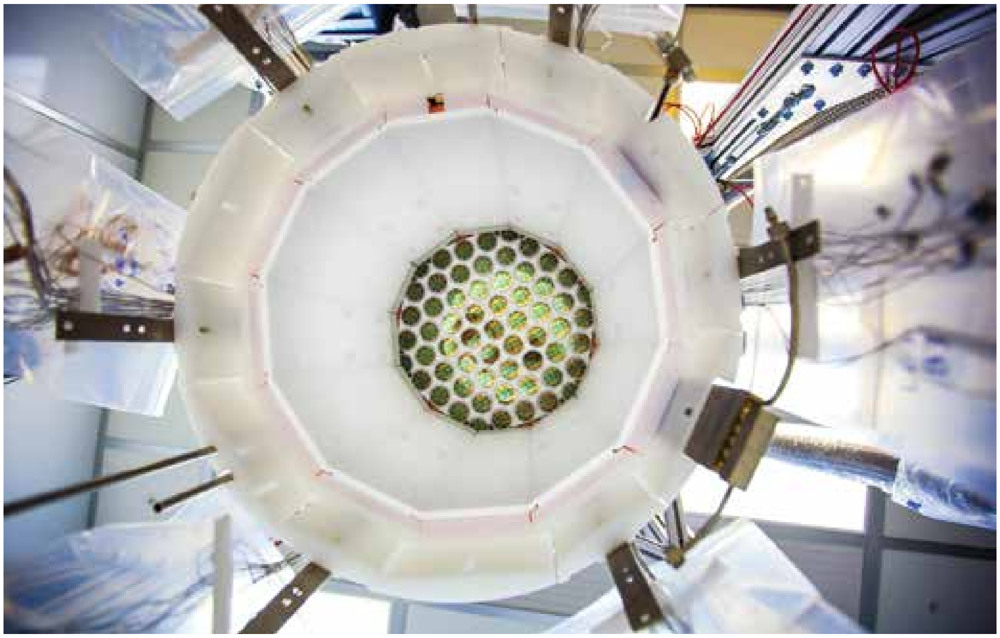
Toward this end, members of NERSC's Data and Analytics Support (DAS) group worked with Berkeley Lab's OpenMSI team to use the Jupyter/iPython notebook interface to enable interactive computing for the OpenMSI project. The DAS team helped create functionality to enable OpenMSI users to automatically open iPython notebooks on the NERSC JupyterHub server (https://jupyter.nersc.gov) with runnable code and analyses customized by the user from the OpenMSI portal. DAS also enabled users to bring their own Python libraries and environments into the notebook via a pluggable mechanism.

Other features of the Jupyter/iPython notebook interface include:

- Customizable template notebooks
- Programmatic and graphical user interface to easily customize notebooks—for example, change analysis parameters via a query string or user interface fields—without requiring users to write specific code to do this
- Ability to customize and integrate external JupyterHub service with a science gateway
- Ability to save notebooks directly to external repositories (such as github) without requiring users to know git
- Integration of user-level data access controls with Jupyter
- Shared notebooks: Sharing code and methods between expert programmers and beginners alike



The Jupyter/iPython Notebook Interface.



A view inside the LUX detector. Photomultiplier tubes can pick up the tiniest bursts of lights when a particle interacts with xenon atoms. *Image: Matthew Kapust, Sanford Underground Research Facility*

Supporting LUX with SciDB

The Large Underground Xenon (LUX) dark matter experiment operates nearly a mile underground at the Sanford Underground Research Facility in South Dakota. The LUX instrument was built to directly detect the faint interactions from galactic dark matter in the form of weakly interacting massive particles (WIMPS). In its initial data run, LUX collected 83 million events, containing 600 million pulses, and wrote 32TB of raw data. Of this data, only 160 events survived basic analysis cuts for potential WIMP candidates. The data rate for the new dark matter run, which started in 2014, exceeds 250TB/year.

As a case study for high energy physics, NERSC's DAS group worked with the LUX team to load the raw data collected from the instrument's inaugural run in 2013 into NERSC's SciDB testbed—an open source database system designed to store and analyze extremely large array-structured data. In this case, analysis of this kind of data typically requires a researcher to search through about 1 million 10MB files within a 10TB dataset. If a WIMP candidate is spotted, the researcher would save the candidate to another file—a process that is slow and cumbersome, taking at least a day and involving analysis steps that are difficult to share. But the DAS group found that by using the SciDB testbed, the same search took just 1.5 minutes from start to finish. Since the SciDB testbed was implemented at NERSC in 2013, there has been so much interest in the use cases that the center deployed a dedicated set of nodes for this kind of analysis. These nodes will be one of the main analysis platforms for the LUX experiment.

Burst Buffer Early User Program Moves Forward

In August 2015, NERSC put out a Burst Buffer Early User Program call for proposals, asking NERSC's nearly 6,000 users to describe use cases and workflows that could benefit from accelerated I/O. NERSC received over 30 responses from the user community. These proposals spanned the DOE Office of Science and represented a breadth of use cases. NERSC staff evaluated each proposal using the following criteria:

- Representation among all six DOE Office of Science programs
- Ability for application to produce scientific advancements

- Ability for application to benefit significantly from the burst buffer
- Resources available from the application team to match NERSC/vendor resources

NERSC originally intended to select five teams to partner with on early burst buffer use cases. However, because of the large response, NERSC chose to support 13 application teams, meaning that each application team would have a point of contact for support from NERSC's staff. In addition, NERSC chose to give an additional 16 teams early access to the burst buffer hardware, but without dedicated support from a NERSC staff member.

Initially, the burst buffer was conceived for the checkpoint and restart use case; however, with NERSC's broad workload, the use cases turned out to be much more diverse, ranging from coupling together complex workflows and staging intermediate files to database applications requiring a large number of I/O operations. Ongoing bi-weekly meetings are held to update users on the latest burst buffer software patches, to offer example use cases and best practices and to allow users to share their experiences.

The Burst Buffer Early User Program Early Access Projects

PROJECT	DOE OFFICE	BB DATA FEATURES
Nyx/Boxlib cosmology simulations (Ann Almgren, LBNL)	HEP	I/O bandwidth with BB; checkpointing; workflow application coupling; *in situ* analysis
Phoenix: 3D atmosphere simulator for supernovae (Eddie Baron, U. Oklahoma)	HEP	I/O bandwidth with BB; staging intermediate files; workflow application coupling; checkpointing
Chombo-Crunch + VisIt for carbon sequestration (David Trebotich, LBNL)	BES	I/O bandwidth with BB; *in situ* analysis/visualization using BB; workflow application coupling
Sigma/UniFam/Sipros Bioinformatics codes (Chongle Pan, ORNL)	BER	Staging intermediate files; high IOPs; checkpointing; fast reads
XGC1 for plasma simulation (Scott Klasky, ORNL)	FES	I/O bandwidth with BB; intermediate file I/O; checkpointing
PSANA for LCLS (Amadeo Perazzo, SLAC)	BES/BER	Staging data with BB; workflow management; in-transit analysis
ALICE data analysis (Jeff Porter, LBNL)	NP	I/O bandwidth with BB; read-intensive I/O
Tractor: cosmological data analysis (DESI) (Peter Nugent, LBNL)	HEP	Intermediate file I/O using BB; high IOPs
VPIC-IO performance (Suren Byna, LBNL)	HEP/ACSR	I/O bandwidth with BB; *in situ* data analysis; BB to stage data
YODA: Geant4 sims for ATLAS detector (Vakhtang Tsulaia, LBNL)	HEP	BB for high IOPs; stage small intermediate files
ALS SPOT Suite (Craig Tull, LBNL)	BES/BER	BB as fast cache; workflow management; visualization
TomoPy for ALS image reconstruction (Craig Tull, LBNL)	BES/BER	I/O throughput with BB; workflow management; read-intensive I/O
kitware: VPIC/Catalyst/ParaView (Berk Geveci, kitware)	ASCR	*in situ* analysis/visualization with BB; multi-stage workflow

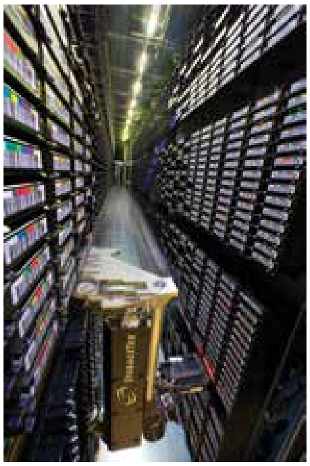
Additional Burst Buffer Early Access Projects

- Image processing in cryo-microscopy/structural biology, Sam Li, UCSF (BER)
- Htslib for bioinformatics, Joel Martin, LBNL (BER)
- Falcon genome assembler, William Andreopoulos, LBNL (BER)
- Ray/HipMer genome assembly, Rob Egan, LBNL (BER)
- HipMer, Steven Hofmeyr, LBNL (BER/ASCR)
- CESM Earth System model, John Dennis, UCAR (BER)
- ACME/UV-CDAT for climate, Dean N. Williams, LLNL (BER)
- Global View Resilience with AMR, neutron transport, molecular dynamics, Andrew Chien, University of Chicago (ASCR/BES/BER)
- XRootD for Open Science Grid, Frank Wuerthwein, UCSD (HEP/ASCR)
- OpenSpeedShop/component-based tool framework, Jim Galarowicz, Krell Institute (ASCR)
- DL-POLY for material science, Eva Zarkadoula, ORNL (BES)
- CP2K for geoscience/physical chemistry, Chris Mundy, PNNL (BES)
- ATLAS simulation of ITK with Geant4, Swagato Banerjee, University of Louisville (HEP)
- ATLAS data analysis, Steve Farrell, LBNL (HEP)
- Spark, Costin Iancu, LBNL (ASCR)

NERSC has identified a broad range of applications and workflows with a variety of use cases for the burst buffer, beyond the checkpoint-restart use case. NERSC burst buffer early users are now enabled on the Cori system and beginning to test their workflows on the burst buffer. In 2016 we will be examining application performance and comparisons to the file system.

Use Cases for 12 Burst Buffer Early User Program Projects

APPLICATION	I/O BANDWIDTH: READS	I/O BANDWIDTH: WRITES (CHECK-POINTING)	HIGH IOPS	WORKFLOW COUPLING	*IN SITU*/ IN-TRANSIT ANALYSIS AND VISUALIZATION	STAGING INTERMEDIATE FILES/PRE-LOADING DATA
Nyx/Boxlib		X		X	X	
Phoenix 3D		X		X		X
Chomo/ Crunch + Visit		X		X	X	
Sigma/ UniFam/Sipros	X	X	X			X
XGC1	X	X				X
PSANA				X	X	X
ALICE	X					
Tractor			X	X		X
VPIC/IO					X	X
YODA			X			X
ALS SPOT/ TomoPy	X			X	X	X
kitware				X	X	



The High Performance Storage System, which has been used for archival storage at NERSC since 1998, received a major upgrade in 2015. *Image: Roy Kaltschmidt, Lawrence Berkeley National Laboratory*

Operational Excellence

HPSS DISK CACHE GETS AN UPGRADE

NERSC users today are benefiting from a business decision made three years ago to upgrade the High Performance Storage System (HPSS) disk cache. For archiving purposes, NERSC uses HPSS, which implements hierarchical storage management, comprising a disk farm front end (referred to as a disk cache) for short-term storage and a tape back end for long-term storage. When new files land on the disk cache they are immediately copied to tape so that they exist in both places. When the disk cache starts to get full, an algorithm selects and removes files that have already been migrated to tape.

An I/O analysis conducted by NERSC's Storage Systems Group in 2013 changed the way HPSS would be architected going forward. The group found that the vast majority of the data that comes into NERSC is accessed within the first 90 days. Previously five days of peak I/O was NERSC's metric for sizing its disk cache.

With the new insights into when users most often access their data, NERSC chose to go in a new direction with its hardware choices for the cache upgrade, which was rolled out in early 2015. Rather than buying SAS drives or fiber channel drives designed to support high performance bandwidth needs, they opted to go with a more capacity-oriented technology involving Nearline-SAS drives, providing 2PB of disk cache for the HPSS Archive system.

With the new arrays, the HPSS disk cache can now retain 30 days' worth of data—roughly 1.5PB—thus reducing I/O bottlenecks and improving users' productivity. Another key benefit is improved performance. There has also been a decrease in consult tickets since deploying the new disk cache.

STREAMLINING DATA AND ANALYTICS SOFTWARE SUPPORT

In 2015, staff from NERSC, in collaboration with LBNL's Computational Research Division, organized a series of working groups to streamline the data and analytics portfolio supported by NERSC. NERSC has been deploying a range of data capabilities, and there has been a strong need to gain a better understanding of the rapidly evolving data technology landscape. The objective of these working groups was to identify, evaluate and recommend technologies for official adoption and deployment on NERSC systems.

Working groups were convened in the following areas:

- Workflows
- Data management (databases and storage formats)
- Data analytics

The sessions consisted of an initial inventory-gathering phase that included a broad survey of relevant technologies in each specific area. Based on the initial survey, the working groups organized deep-dive sessions, with presentations from multiple vendors, software developers and users of a given technology. Finally, the group members did some additional evaluation and offered further recommendations for the data technology stack to be deployed and supported at NERSC.

The conclusions of these working groups, along with existing user needs, resulted in an officially supported set of data technologies for science—NERSC's Big Data Software Portfolio. These technologies provide a deployment vehicle for several research projects and a path for productive engagement with data science collaborators.

NERSC's Big Data Software Portfolio

CAPABILITIES	TECHNOLOGY AREAS	TOOLS, LIBRARIES
Data Transfer + Access	Globus, Grid Stack, Authentication	Globus Online, Grid FTP
	Portals, Gateways, RESTful APIs	Drupal/Django, NEWT
Data Processing	Workflows	Swift, Fireworks
Data Management	Formats, Models Databases	HDF5, NetCDF, ROOT MongoDB, SciDB, PostgreSQL, MySQL
	Storage, Archiving	Lustre/GPFS, HPSS, SRM
Data Analytics	Statistics, Machine Learning	python, R, ROOT BDAS/Spark
	Imaging	OMERO, Fiji, MATLAB
Data Visualization	SCiVis InfoVis	VisIt, Paraview Tableau

THE BIG MOVE: MINIMIZING DISRUPTIONS, IMPROVING OPERATIONS

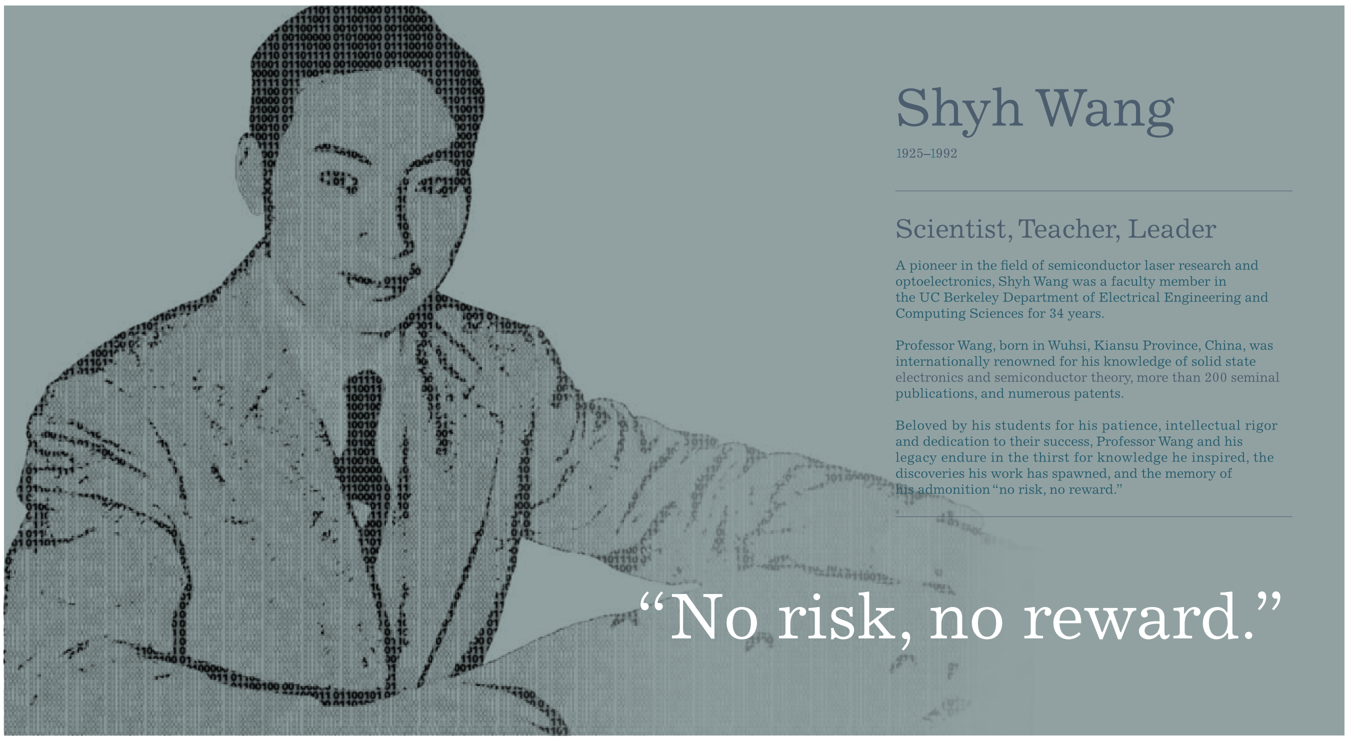
NERSC's move to Wang Hall was a complicated process, and part of the planning involved ensuring that there would be minimal disruption to users. One example of this was staggering the shutdown and move of the Cray systems, ensuring that the Cori Phase 1 system was in place and available for computation prior to shutting down Hopper and Edison.

In terms of Hopper, while computation on it was stopped in December, the login nodes and data remained available to users for several weeks after decommissioning the system. Edison's move from the Oakland Scientific Facility to Wang Hall began on November 30, 2015; the system came back to production on January 4, 2016, almost a week earlier than anticipated. NERSC staff worked over the holiday break to bring the system back up for users. The NERSC global file systems suffered no downtime during the move as data were synced over a 400G network between the Oakland and Berkeley facilities. Users only experienced lower performing file systems for a few days.

NERSC's move to the new building also provided opportunities for operational improvements. For example, NERSC's Operations Technology Group includes matrixed staff funded by ESnet to support the Network Operations Center (NOC) for ESnet. With the transition of the NOC to Wang Hall, the closer proximity of NERSC and ESnet staff has created new opportunities for technology transfer and collaboration between the two divisions. One example is a new NOC Lecture Series, a 90-minute session held once a month on wide-area networking topics that provides a venue for ESnet and NERSC engineers to update operational knowledge on current implementations, new software and new configurations. The series has provided a systematic method for operations staff to improve procedures and diagnostic capabilities and to learn how to use new tools. These sessions also give NERSC staff expanded opportunities to learn about work in another facility.

Center News

Shyh Wang Hall features a number of conference rooms and "just in time" meeting rooms for private discussions and gatherings–many with spectacular views of UC Berkeley and the San Francisco Bay.



NERSC's New Home: Shyh Wang Hall

One of the biggest things to happen at NERSC in 2015 was the center's move back to the main Berkeley Lab campus and into Shyh Wang Hall, a brand new, state-of-the-art facility for computational science. Wang Hall is named in honor of Shyh Wang, a professor at UC Berkeley for 34 years who was known for his research in semiconductors, magnetic resonances and semiconductor lasers, which laid the foundation for optoelectronics.

Wang Hall is named in honor of Shyh Wang, a professor at UC Berkeley for 34 years who was known for his research in semiconductors, magnetic resonances and semiconductor lasers.

The $143 million structure, financed by the University of California, provides an open, collaborative environment bringing together nearly 300 staff members from NERSC, ESnet and the Computational Research Division, plus colleagues from nearby UC Berkeley, to encourage new ideas and approaches to solving some of the nation's biggest scientific challenges.

The 149,000 square foot facility, built on a hillside overlooking the UC Berkeley campus and San Francisco Bay, houses one of the most energy-efficient computing centers anywhere, tapping into the region's mild climate to cool NERSC's supercomputers and eliminating the need for mechanical cooling. The building features large, open bays on the lowest level, facing west toward the Pacific Ocean, that draw in natural air conditioning for the computing systems. In addition, heat captured from those systems is recirculated to heat the building.

The new building was also designed to be flexible to meet the needs of next-generation HPC systems. The building has 12.5 megawatts of power (upgradable to over 40 megawatts) and is

The symbolic connection of the 400Gbps optical network at the Wang Hall dedication ceremony in November 2015. Left to right: DOE's Barbara Helland, UC President Janet Napolitano, former Berkeley Lab director Paul Alivisatos, Dila Wang (wife of Shyh Wang), Berkeley Lab Associate Director Kathy Yelick, UC Berkeley Chancellor Nicholas Dirks, Ipitek CEO Michael Salour and Berkeley Lab Deputy Director Horst Simon. *Image: Kelly Owen, Lawrence Berkeley National Laboratory*

designed to operate at a PUE (power usage effectiveness) of less than 1.1. The machine room measures 20,000 square feet (expandable to 30,000) and features a novel seismic isolation floor to protect the center's many resources in the event of an earthquake.

Wang Hall's innovate design can be seen throughout the building. For example, in addition to the views, the west-facing windows were carefully placed to ensure both optimum light and energy efficiency. The architects also made sure that areas near the windows on all sides of the building offer open space and casual drop-in seating so that all staff can enjoy the space and be inspired by the views.

The building's design is also intended to promote and encourage staff interaction across the three Computing Sciences divisions that now call Wang Hall home: NERSC, ESnet and CRD. Toward this end, there's a mix of space in the top two floors of the building: offices, cubicles, open seating areas, "bullpens," just-in-time meeting spaces and conference rooms of all sizes. The top two floors also feature breakrooms with floor-to-ceiling windows looking out at San Francisco and the Bay Area, an open kitchen and comfortable seating.

Because of the dramatic views and open environment, Wang Hall has already become a popular venue for meetings and conferences for groups within and outside of Computing Sciences, with additional opportunities for computer-room tours and "photo ops" with NERSC's newest supercomputer, Cori.

Changing of the Guard: Hopper and Carver Retire, Cori Moves In

Another big change in 2015 was the retirement of two of NERSC's legacy systems, Carver and Hopper.

Carver, which was decommissioned on September 30, was named in honor of American scientist George Washington Carver. The IBM iDataPlex cluster featured nodes with relatively high memory per core, a batch system that supported long running jobs and a rich set of third-party software applications. Carver also served as the interactive host for various experimental testbeds, such the Dirac GPU cluster and Hadoop.

Hopper, a Cray XE6 system, was retired on December 15. Named in honor of computing pioneer Admiral Grace Hopper, who developed the first compiler and originated the concept of machine-independent programming languages, Hopper was NERSC's first petaflop system, with a peak performance of 1.28 Petaflops/sec, 153,216 compute cores, 212TB of memory and 2PB of online disk storage. In fact, Hopper was the fifth fastest supercomputer in the world in 2010, according to the TOP500 list of the world's supercomputers.

The phasing out of Hopper and Carver was done in conjunction with the deployment of the first phase of Cori in Wang Hall and the move of Edison from the Oakland Scientific Facility to Berkeley Lab. Cori Phase 1 is being configured as a data partition and will be joined with the larger Cori Phase 2 system arriving in the fall of 2016. The Cori data partition has a number of unique features designed to support data-intensive workloads, including a real-time queue; a burst buffer that provides high bandwidth, low latency I/O; a large amount of memory per node (128GB/node); several login/interactive nodes to support applications with advanced workflows; and high throughput and serial queues.

A key element of Cori Phase 1 is Cray's new DataWarp burst buffer technology, which accelerates application I/O and addresses the growing performance gap between compute resources and disk-based storage. The burst buffer is a layer of NVRAM designed to move data more quickly between processor memory and disk and allow users to make the most efficient use of the system. To ensure that they can take full advantage of DataWarp's capabilities, NERSC also launched the Burst Buffer Early User Program, which enables us to work closely with a set of users who have diverse use cases for the burst buffer.

APEX: A Joint Effort for Designing Next-Generation HPC Systems

2015 also saw the formation of the Alliance for Application Performance at Extreme Scale (APEX), a collaboration between Lawrence Berkeley, Los Alamos and Sandia national laboratories that is focused on the design, acquisition and deployment of future advanced technology high performance computing systems.

NERSC and ACES first began collaborating in 2010 with the procurement and deployment of two Cray XE-6 systems: Hopper at NERSC and Cielo at LANL. The partnership was formalized and strengthened with the Trinity/NERSC-8 project for the procurement of the Trinity system at LANL and the Cori system at NERSC.

APEX is designed to build on this partnership. Over the years, the labs have independently deployed world-leading supercomputers to support their respective scientific missions. In joining together, they aim to work even more closely with vendors to shape the future of supercomputer design to deliver ever-more capable systems to solve problems of national importance.

Initially, APEX will focus on future computing technologies via two new advanced computing systems: Crossroads for the New Mexico Alliance for Computing at Extreme Scale (ACES, a Los Alamos and Sandia partnership), and NERSC-9, NERSC's next-generation system. In addition to providing more computing capability to NERSC users, the NERSC-9 system will advance new technologies on the path to exascale and demonstrate unique features to support data-intensive workloads.

APEX will also allow the three labs to pursue areas of collaboration with vendors, providing substantial value to the Crossroads and NERSC-9 systems in the areas of increased application performance, increased workflow efficiency and enhanced reliability of the systems. The APEX partners have released a workflows document that describes the way in which users perform their science end-to-end, including their data motion needs and accesses. A workload analysis and the workflows document can be found at http://www.nersc.gov/research-and-development/apex/apex-benchmarks-and-workflows/.

Industry Partnership Program Expands

NERSC has been working with industry partners through the DOE's Small Business Innovation Research (SBIR) grant program for many years. Since 2011, industry researchers have been awarded 40 million hours at NERSC via the SBIR program, and NERSC currently has about 130 industry users from 50 companies. In 2015, through its Private Sector Partnership program, the center added nine new projects in energy-specific fields ranging from semiconductor manufacturing to geothermal modeling:

COMPANY	R&D COMMUNIITY	AWARD TYPE
Chemical Semantics Inc.	Private sector materials science	SBIR
Civil Maps	Monitoring of energy infrastructure	DDR*
Cymer Inc.	Improving EUV Lithography Sources with HPC	CRADA
Exabyte	Web based HPC for materials science	BES
Global Foundries	Semiconductor device design, chip fab	DDR
Kinectrics	Metal alloys for pipe networks	DDR
Melior Innovations	Geothermal reservoir modeling	DDR
Nano Precision Products	Optical interconnect engineering	SBIR
Quanturmscape	Solid-state batteries for electric cars	DDR

*Directors Discretionary Reserve

Among the nine new partners is Melior Innovations, which is developing and optimizing software for accurate modeling of geothermal power plants. Subsurface networks containing thousands of fractures present complex dynamic flow and diffusion problems that must be solved for thousands of time steps. Melior Innovations has been collaborating with the open source FEniCS Project to develop simulation capabilities for fluid flow problems in large geophysical discrete fracture networks. Using NERSC'S Edison system, Melior has already improved the scalability of a solver code for geothermal energy so problems can be studied in greater detail.

Here's a look at how some of NERSC's other industry partners are using the center's supercomputing resources:

- Electric utilities are keenly interested in evaluating the effects of climate change and extreme weather on wind energy. Toward this end, Vertum Partners—a company specializing in applying atmospheric simulations to renewable energies, operational weather forecasts and long-term climate dataset construction—has been working with NERSC to create 4D next-generation mesoscale numerical weather simulations that model complex atmospheric data resolved to local wind farm scales. Using WRF, a code designed to solve 4D models of the atmosphere, on NERSC supercomputers, Vertum was able to evaluate the sensitivity of the model results to different land surface types, atmospheric datasets and other forcings both internal and external to the model. This will allow for optimal turbine site selection and evaluating the true climate change effects to wind energy.

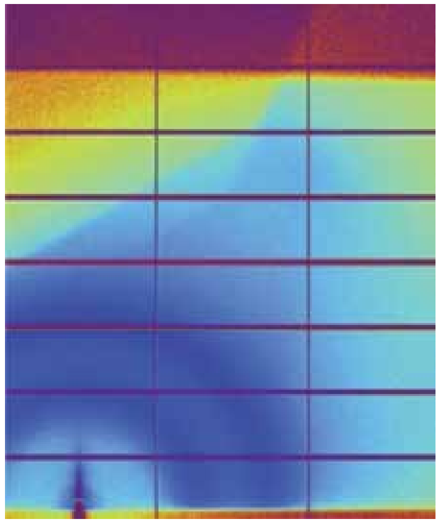
- DynaFlow, a provider of research and development services and products in fluid dynamics and material sciences, used a NERSC allocation to run a numerical model that can accurately capture flow dynamics of microbubbles that can help and/or hinder industrial machinery. The findings, published in the *Journal of Fluids Engineering* (April 2015) and *Chemical Engineering Science* (May 2015), verified the importance of "bubbly flow" in many industrial applications; for example, it can cause cavitation leading to efficiency loss in turbomachinery, but it can also be exploited to enhance cutting, drilling, cleaning and chemical manufacturing. In particular, the simulations run at NERSC showed that bubble cloud characteristics and initial pressure could be "tuned" to generate very high cavitation forces that could be used for beneficial purposes.

- GlobalFoundries, one of the largest semiconductor foundries in the world, provides a unique combination of advanced technology and manufacturing to more than 250 customers. In 2015, the company began using NERSC supercomputers to improve product development time and costs and facilitate the discovery of new materials for next-generation low-power, high-performance semiconductors through atomic scale simulations. GlobalFoundries is using these computational findings to guide development of its new 22-nm FDSOI chips.

- Civil Maps, which uses artificial intelligence and HPC to create comprehensive 3D maps, is using LIDAR monitoring data to help utility companies achieve more reliable power grids. LIDAR monitoring of power grid assets is crucial to the reliably engineered operation of distributed energy assets. For example, Pacific Gas & Electric Co. in California is using a combination of aerial surveys and mobile LIDAR mapping to monitor the growth of trees near power lines, eliminating the need for tedious land surveying using handheld GPS receivers. Geographic information specialists can then use visualizations of the data to explore the infrastructure on a computer. However, the raw, unorganized 3D LIDAR data gathered from the mobile sensors is very large in size. Civil Maps is tapping into NERSC's resources to test and optimize algorithms and process the enormous data sets generated by LiDAR scans.

On a broader scale, NERSC is participating in DOE's new High Performance Computing for Manufacturing Program (HPC4Mfg), announced in September 2015. HPC4Mfg couples U.S. manufacturers with world-class computational research and development expertise at Lawrence Berkeley, Lawrence Livermore and Oak Ridge national laboratories to fund and foster public-private R&D projects that enhance U.S. competitiveness in clean energy manufacturing and address key challenges in U.S. manufacturing.

In February 2016, DOE announced $3 million for the first 10 HPC4Mfg projects. Each of the projects will receive approximately $300,000 to fund the national labs to partner closely with each chosen company to provide expertise and access to HPC systems aimed at high-impact challenges. GlobalFoundries was one of the projects selected; the company will collaborate with Berkeley Lab and NERSC to optimize the design of transistors under a project entitled "Computation Design and Optimization of Ultra-Low Power Device Architectures."



A GISAXS pattern from a printed organic photovoltaic sample as it dried in the beamline, part of a "superfacility" experiment based at Berkeley Lab. *Image: Craig Tull, Lawrence Berkeley National Laboratory*

'Superfacilities' Concept Showcases Interfacility Collaborations

Throughout 2015 NERSC saw increased interest in the coupled operation of experimental and computational facilities throughout the DOE Office of Science complex—the so-called "superfacility."

A superfacility is two or more facilities interconnected by means of a high performance network specifically engineered for science (such as ESnet) and linked together using workflow and data management software such that the scientific output of the connected facilities is greater than it otherwise could be. This represents a new experimental paradigm that allows HPC resources to be brought to bear on experimental work being done at, for example, a light source beamline within the context of the experimental run itself.

Adding HPC to facility workflows is a growing trend, and NERSC is leveraging its experience with the Joint Genome Institute, the Advanced Light Source, the Large Hadron Collider, the Palomar Transient Factory, AmeriFlux and other experimental facilities to shape the interfaces, schedules and resourcing models that best fit what our facility users need. Treating facilities as users of HPC requires long-term data planning, real-time resource scheduling and pipelining of workflows.

NERSC's superfacility strategy focuses on scalable data analysis services that can be prototyped and reused as a new interfacility collaborations emerge. For example, in 2015 a collaborative effort linking the Advanced Light Source (ALS) at Berkeley Lab with supercomputing resources at NERSC and Oak Ridge National Laboratory via ESnet yielded exciting results in organic photovoltaics research. In a series of experiments, the operators of a specialized printer of organic photovoltaics used HPC to observe the structure of the material as it dried. Organic photovoltaics show promise as less expensive, more flexible materials for converting sunlight to electricity.

During the experiments, scientists at the ALS simultaneously fabricated organic photovoltaics using a miniature slot-die printer and acquired beamline data of the samples as they were made. This data was transferred to NERSC via ESnet, where it was put into the SPOT Suite data portal and underwent analysis to translate the image pixels into the correct reciprocal space. Using the Globus data management tool developed by the University of Chicago and Argonne National Laboratory, the output from those jobs was then sent to the Oak Ridge Leadership Computing Facility for analysis, where the analysis code ran on the Titan supercomputer for six hours nonstop on 8,000 GPUs.

A unique feature of this experiment was co-scheduling beamline time at the ALS with time and resources at NERSC and Oak Ridge, plus utilizing ESnet to link it all together. This was the first time a beamline was able to simultaneously take data and perform large-scale analysis on a supercomputer.

NERSC and the 2015 Nobel Prize in Physics

In 2015 NERSC was once again fortunate to be part of the story behind a Nobel Prize. This time it was the 2015 Nobel Prize in Physics, awarded to the leaders of two large neutrino experiments: Takaaki Kajita of Tokyo University, who led the Super-Kamiokande experiment, and Arthur B. McDonald of Queen's University in Kingston, Ontario, Canada, head of the Sudbury Neutrino Observatory (SNO). According to the Nobel Prize committee, Kajita and McDonald were honored "for the discovery of neutrino oscillations, which shows that neutrinos have mass."

NERSC's connection to this Nobel Prize is through SNO, a huge neutrino detector based in Sudbury, Ontario. The observatory is located in an old copper mine 2 km underground to shield it as much as possible from the noise background of particle interactions that take place on Earth's surface.

The SNO detector was turned on in 1999, and from the earliest days SNO data was transferred to NERSC, where the center's PDSF cluster was used for what became known as the "West Coast Analysis." When the discovery of neutrino flavor mixing in solar neutrinos was published in 2001 in *Physical Review Letters*, NERSC's role was well established and recognized by scientists working on the project; in fact, they presented NERSC's PDSF team with a signed and framed copy of the journal article.

The SNO experiment wound down in 2006, but the data continues to prove invaluable. To assure its integrity and safekeeping, NERSC was chosen to house the data archive. Archiving and moving all the data to NERSC required a close collaboration between SNO, NERSC and ESnet.

When the Nobel Prize for Physics was announced in October 2015, one of the leads on the SNO project, Alan Poon of Berkeley Lab, asked that the following email be distributed to the entire NERSC staff:

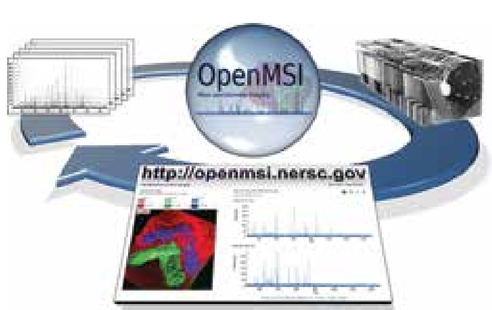
> *SNO has been blessed by the top-notch support and facility at NERSC. Without NERSC's support, SNO would not have been able to process and reconstruct the data, simulate the data and run massive jobs for the physics fits so smoothly and successfully.*
>
> *We appreciate your continual support of our data archive at HPSS. As you can imagine, we want to preserve this precious data set, and once again, NERSC has come to our assistance. Many, many thanks from all of us in SNO.*
>
> *Best,*
> *Alan*

To which Dr. McDonald added in a subsequent email:

> *May I add my thanks to those from Alan. We greatly appreciate your support.*
>
> *Best regards,*
> *Art McDonald*

Art McDonald at Berkeley Lab in November 2015, where he outlined NERSC's and Berkeley Lab's contributions to the SNO Collaboration.

OpenMSI Wins 2015 R&D 100 Award



OpenMSI, the most advanced computational tool for analyzing and visualizing mass spectrometry instruments (MSI) data, was one of seven Berkeley Lab winners of a 2015 R&D 100 award. Oliver Rübel of the Computational Research Division and Ben Bowen of the Environmental Genomics and Systems Biology Division led the development of OpenMSI, with collaboration from NERSC.

MSI technology enables scientists to study tissues, cell cultures and bacterial colonies in unprecedented detail at the molecular level. As the mass and spatial resolution of MS instruments increase, so do the number of pixels in MS images and data size. Nowadays, MSI datasets range from tens of gigabytes to several terabytes. Thus, basic tasks like opening a file or plotting spectra and ion images become insurmountable challenges.

OpenMSI overcomes these obstacles by making highly optimized computing technologies available via a user-friendly interface. Because OpenMSI leverages NERSC's resources to process, analyze, store and serve massive MSI datasets, users can now work on their data at full resolution and in real time without any special hardware or software. They can also access their data on any device with an Internet connection.

Presented by *R&D Magazine*, the annual R&D 100 Awards recognize the top 100 technology products from industry, academia and government-sponsored research, ranging from chemistry and materials to biomedical and information technology breakthroughs.

NERSC Staff Honored with Director's Awards for Exceptional Achievement

Brent Draney (left) and Eli Dart (center) received their Director's Award from Berkeley Lab Deputy Director Horst Simon.

Two long-time NERSC staff members—Brent Draney and Lynne Rippe—were among five Berkeley Lab Computing Sciences employees to be honored with 2015 Director's Awards for Exceptional Achievement.

Draney, head of NERSC's Networking, Security and Servers Group, and Eli Dart of ESnet's Science Engagement Team, were recognized for their work in developing the Science DMZ, a network architecture that allows science data to securely bypass institutional firewalls. The Science DMZ has been endorsed by the National Science Foundation, which has funded Science DMZs at more than 100 universities across the U.S. Draney and Dart, who was working at NERSC when the Science DMZ concept was born, were honored in the area of operations for "achievement in operational effectiveness, process re-engineering or improvement, resource management and efficiency or partnerships across organizational/departmental boundaries."

Lynne Rippe

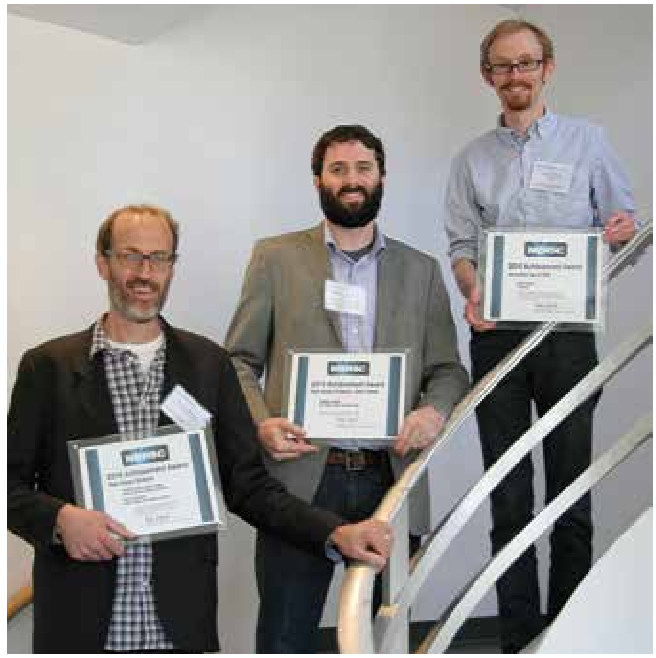
Rippe was awarded a Berkeley Lab Citation for her longtime procurement work in support of NERSC. A lead subcontracts administrator with the Lab's Office of the Chief Financial Officer, Rippe has been the procurement department lead on every major NERSC system dating back to the center's days at Lawrence Livermore National Laboratory. She was also responsible for the development of the Best Value Source Selection process for large-scale computer procurements, which has been widely adopted within DOE laboratories. The Berkeley Lab Citation recognizes the highest level of service to Berkeley Lab and the network of DOE laboratories.

4th Annual HPC Achievement Awards Highlight Innovative Research

NERSC announced the winners of the fourth annual High Performance Computing Achievement Awards in March 2016 during the annual NERSC Users Group meeting at Berkeley Lab. The awards recognize NERSC users who have either demonstrated an innovative use of HPC resources to solve a scientific problem or whose work has had an exceptional impact on scientific understanding or society. To encourage younger scientists who are using HPC in their research, NERSC also presents two early career awards.

NERSC 2016 AWARD FOR HIGH IMPACT SCIENCE: OPEN

Charles Koven and William Riley of Berkeley Lab's Climate and Ecosystem Sciences Division and David Lawrence of the National Center for Atmospheric Research were honored for their role in using an Earth system model to demonstrate the atmospheric effect of emissions released from carbon sequestered in melting permafrost soil. Running simulations on NERSC's Hopper system, the team demonstrated that thawing permafrost releases enormous amounts of long-frozen carbon into the atmosphere—more than carbon taken by plants.



The 2016 NERSC HPC Achievement Award winners (left to right): Charles Koven, Nathan Howard, Scott French. Not pictured: William Riley, David Lawrence, Min Si. Image: *Margie Wylie, Lawrence Berkeley National Laboratory*

The model is the first to represent permafrost processes as well as the dynamics of carbon and nitrogen in the soil. According to the simulations, by the year 2300—if climate change continues unchecked—the net loss of carbon to the atmosphere from Arctic permafrost would range from between 21 petagrams and 164 petagrams. That's equivalent to between two years and 16 years of human-induced CO_2 emissions.

"We found the rate of permafrost thaw and its effect on the decomposition of deep carbon will have a much bigger impact on the carbon cycle than the availability of deep nitrogen and its ability to spark plant growth," Koven said. Their findings were reported in the *Proceedings of the National Academy of Sciences* in March 2015.

NERSC 2016 AWARD FOR HIGH IMPACT SCIENCE: EARLY CAREER

Nathan Howard, a research scientist at MIT's Plasma Science and Fusion Center, is being honored in this category for his pioneering computational work in plasma turbulence simulations. In particular, Howard—who was a postdoc at the University of California, San Diego until joining MIT in 2015—carried out the most physically comprehensive simulations of tokamak plasma microturbulence to date, according to Christopher Holland, associate research scientist in the Center for Energy Research at UCSD, who nominated Howard for this award.

One roadblock in the quest for fusion is that, to date, computer models have often been unable to predict exactly how turbulence will behave inside the reactor. In fact, there have long been differences between predictions and experimental results in fusion experiments when studying how turbulence contributes to heat loss in the confined plasma. But Howard discovered a solution to this disparity: By performing high-resolution multi-scale simulations, the team was able to simultaneously resolve multiple turbulence instabilities that have previously been treated in separate simulations. A series of these multi-scale simulations run on NERSC's Edison system found that interactions between turbulence at the tiniest scale (that of electrons) and turbulence at a scale 60 times larger (that of ions) can account for the mysterious mismatch between theoretical predictions and experimental observations of the heat loss.

These groundbreaking simulations used well over 100 million core-hours of computing time, and Howard was the top user of NERSC computer time in 2014 and in the top five for 2015. The results resolved a long-standing model-experiment discrepancy and led to the most significant advance in our understanding of tokamak turbulence in over a decade, according to Holland.

NERSC 2016 AWARD FOR INNOVATIVE USE OF HPC: OPEN

Former Berkeley Lab computational scientist Scott French was honored in this category for his role in helping seismologists create a unique 3D scan of the Earth's interior that resolved some long-standing questions about mantle plumes and volcanic hotspots. (See science highlight on p. 31.)

Working with Barbara Romanowicz, a professor of earth and planetary science at UC Berkeley, French—who now works at Google—ran a number of simulations at NERSC, producing for the first time a computed tomography scan that conclusively connects plumes of hot rock rising through the mantle with surface hotspots that generate volcanic island chains like Hawaii, Samoa and Iceland. Until this study, evidence for the plume and hotspot theory had been circumstantial, and some seismologists argued instead that hotspots are very shallow pools of hot rock feeding magma chambers under volcanoes.

To create the high-resolution 3D image of Earth, French ran numerical simulations of how seismic waves travel through the mantle and compared their predictions to ground motion actually measured by detectors around the globe. They mapped mantle plumes by analyzing the paths of seismic waves bouncing around Earth's interior after 273 strong earthquakes that shook the globe over the past 20 years. The simulations, run on NERSC's Edison system, computed all components of the seismic waves. French tweaked the model repeatedly to fit recorded data using a method similar to statistical regression. The final computation required 3 million CPU hours on Edison; parallel computing shrank this to a couple of weeks.

NERSC 2016 AWARD FOR INNOVATIVE USE OF HPC: EARLY CAREER

Min Si, a graduate student from the University of Tokyo who is working at Argonne National Laboratory, received this award for her pioneering work in developing novel system software in the context of MPI3 one-sided communication. This software has had a transformative impact on the field of computational chemistry by completely eliminating the need for custom ports of global arrays, according to Jeff Hammond, a research scientist in the Parallel Computing Lab at Intel Labs, who nominated Si for the award.

The tool Si created is Casper, a library that sits between an MPI3 application and any MPI implementation, including proprietary ones such as Cray MPI and Intel MPI. When Casper is used, the application sees an ideal implementation of MPI3: it makes excellent asynchronous progress for all operations, including noncontiguous and accumulating one-sided operations that are rarely, if ever, supported in hardware. This is achieved without the usual overheads of asynchronous progress, which include mutual exclusion and oversubscription resulting from communication helper threads. Casper is a boon to MPI3 users and implementers because it resolves the longstanding tension over asynchronous progress, which torments both groups because it previously was not possible to achieve it without tradeoffs.

Using Casper, NWChem runs as fast using MPI3 as it does using DMAPP on NERSC Cray XC systems without compromising all of the productive features of MPI, such as profiling and debugging tool support, noted Hammond, who previously worked as a computational scientist at Argonne with Si.

Organizational Changes Reflect Evolving HPC Data Environment

In February 2015, NERSC announced several organizational changes to help the Center's 6,000 users manage their data-intensive research more productively.

Notably, NERSC determined that there was enough momentum around HPC for data-intensive science that it warranted a new department focused on data: the Scientific Computing & Data Services Department, which is being led by Katie Antypas, who was also named Division Deputy for Data Science. The department comprises four groups, two of them new: Data Science Engagement (led by Kjiersten Fagnan), which engages with users of experimental and observational facilities, and Infrastructure Services, which focuses on infrastructure to support NERSC and the center's data initiative. The existing Storage Systems Group and Data and Analytics Group complete the new department.

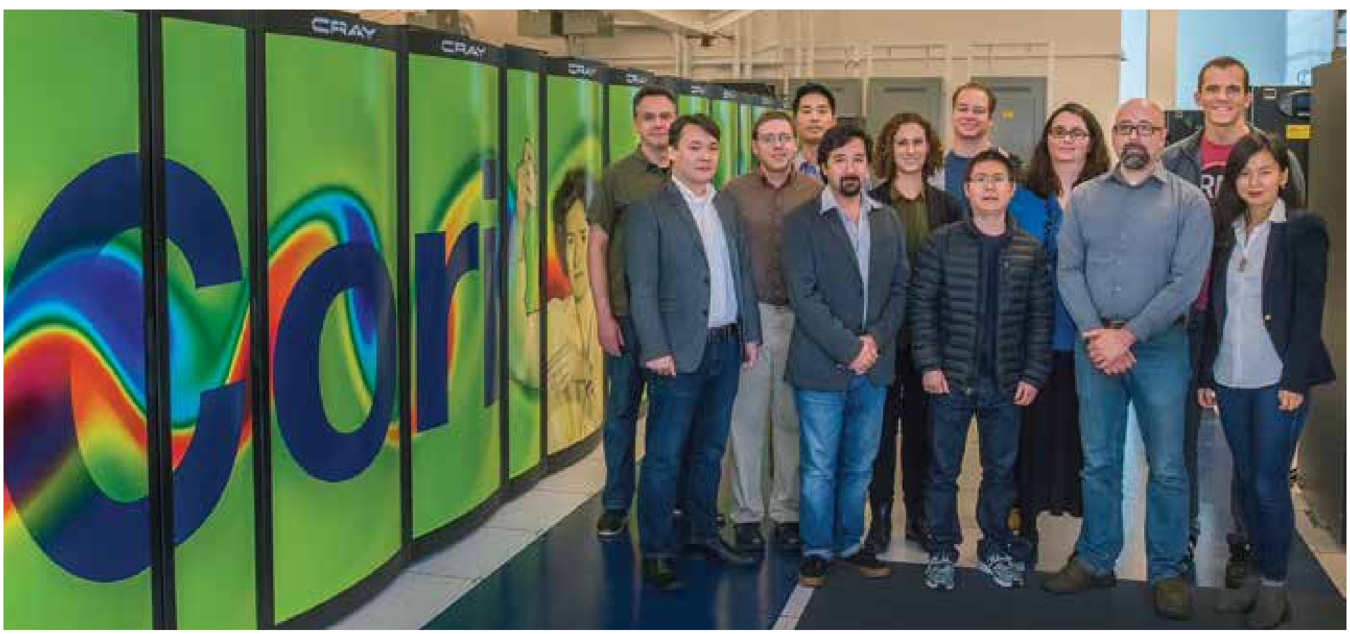
In addition to the new Data Services Department, NERSC now features an HPC Department, which comprises the Advanced Technologies, Application Performance, Computational Systems and User Engagement groups; and the Systems Department, which comprises the Network and Security Group and Operations Technology Group.

Personnel Transitions

A number of long-term NERSC veterans retired in 2015, including:

- **Iwona Sakrejda:** A member of the Technology Integration group at NERSC when she retired, Sakrejda was with Berkeley Lab for 24 years, 15 of those at NERSC. She earned her Ph.D. in particle physics from the Institute of Nuclear Physics in Krakow, Poland.
- **Karen Schafer:** Schafer, a network engineer in the Networking, Security and Servers group, joined NERSC in 2008 and retired in June 2015.
- **David Turner:** An HPC consultant in what was formerly known as NERSC's User Services Group, Turner was with NERSC for 17 years.
- **Francesca Verdier:** Verdier joined NERSC in 1996 as part of the center's massive hiring effort after relocating from Lawrence Livermore National Laboratory to Berkeley Lab. After 19 years at NERSC, she retired as Allocations Manager in June 2015.
- **Harvey Wasserman:** An HPC consultant at NERSC since 2006, Wasserman was involved in workload characterization, benchmarking and system evaluation at NERSC and at Los Alamos National Laboratory for more than 30 years.

Despite the multiple retirements and normal turnover and churn from competition with Bay Area technology companies, NERSC expanded its workforce in 2015 with 18 additional staff members, including postdocs hired specifically for the NERSC Exascale Science Application Program. Overall, NERSC currently has more than 60 FTEs, as well as administrative, communications, research and support staff.



NERSC new hires in 2015 totaled 18. Left to right: Cori Snavely, Glenn Lockwood, Taylor Barnes, Jackson Gor, Rollin Thomas, Debbie Bard, Evan Racah, Yunjie Liu, Rebecca Hartman-Baker, Daniel Udwary, Brian Friesen and Indira Kassymkhanova. Not pictured: Wahid Bhimiji, Ankit Bhagatwala, Doug Doerfler, Jialin Liu, Andrey Ovsyannikov and Tavia Stone

Allocation of NERSC Director's Reserve of Computer Time

The NERSC director receives ten percent of the total allocation of computer time, which amounted to 300 million hours in 2015. The Director's Reserve is used to support innovative and strategic projects as well as projects in NESAP (the NERSC Exascale Science Applications Program), a collaborative effort in which NERSC is partnering with code teams and library and tools developers to prepare for the Cori manycore architecture. A number of Director's Discretionary Reserve projects were focused on exploring performance and scaling on emerging applications. An additional part of the Director's Reserve in 2015 was allocated, in conjunction with OS program managers, to high-impact Office of Science mission computing projects, augmenting their base allocations.

CATEGORY	NUMBER OF PROJECTS	ALLOCATED FROM DIRECTOR'S RESERVE (MILLION MPP HOURS)
Director's Reserve Projects	45	160
Mission Computing Augmentation (in collaboration with SC program managers)	221	140
TOTAL		**300**

2015 Director's Reserve Allocations

PROJECT	PI	ORGANIZATION	ALLOCATION
APPLIED MATH			
Anchored Solar-Driven Vortex for Power Generation	Pearlstein, Arne	U. Illinois U-C	5,800,000
NERSC Application Readiness for Future Architectures	Antypas, Katie	Berkeley Lab	1,000,000
High-resolution, Integrated Terrestrial and Lower Atmosphere Simulation of the Contiguous United States	Maxwell, Reed	Colorado School Mines	1,000,000
COMPUTER SCIENCE			
XPRESS Program Environment Testing at Scale	Brightwell, Ron	Sandia Labs NM	12,000,000
High Performance Data Analytics	Prabhat, Mr	NERSC	8,000,000
Performance Analysis, Modeling and Scaling of HPC Applications and Tools	Bhatele, Abhinav	Livermore Lab	3,000,000
Data Intensive Science R&D	Prabhat, Mr	Berkeley Lab	3,000,000
Exploring Quantum Optimizers via Classical Supercomputing	Hen, Itay	U. Southern California	2,000,000
Berkeley Institute for Data Sciences through HPC	Perlmutter, Saul	UC Berkeley	2,000,000
DEGAS: Dynamic Exascale Global Address Space	Yelick, Katherine	Berkeley Lab	1,100,000
Parallel FEC Verification and Scaling Prototype	Bell, Greg	Berkeley Lab	1,000,000
Domain-Specific Language Technology	Quinlan, Daniel	Livermore Lab	1,000,000

PROJECT	PI	ORGANIZATION	ALLOCATION
Gamma-Ray Data Cloud	Quiter, Brian	Berkeley Lab	1,000,000
Traleika Glacier X-Stack	Borkar, Shekhar	Intel Inc.	500,000
Point Cloud Mapping of Energy Assets and Risks	Chraim, Fabien	Berkeley Lab	500,000
Application Benchmarking and SSP Development	Oppe, Thomas	HPCMPO	500,000
Performance Profiling and Assessment of MFIX Using Trilinos Linear Solvers Versus Using Native Linear Solvers	Gel, Aytekin	ALPEMI Consulting, LLC	100,000
BIOSCIENCES			
Integrated Tools for Next-Generation Bio-Imaging	Skinner, David	Berkeley Lab	13,500,000
Multi-scale Multi-Compartment Computational Models of the Human Immune Response to Infection with M. tuberculosis	Kirschner, Denise	U. Michigan	2,500,000
Collaborative Research in Computational Neuroscience - CRCNS	Bouchard, Kristofer	Berkeley Lab	1,500,000
Enhanced Sampling Simulations of G-protein Coupled Receptors	Miao, Yinglong	H. Hughes Med Inst	1,000,000
CLIMATE & ENVIRONMENT			
FP-RadRes: High Performance Flux Prediction Towards Radiological Resilience	Steefel, Carl	Berkeley Lab	10,000,000
Next Generation Global Prediction System (NGGPS) Benchmarking	Michalakes, John	NOAA	4,000,000
Greening the Grid USAID/India	Purkayastha, Avi	Berkeley Lab	2,500,000
Database of Energy Efficiency Performance	Hong, Tianzhen	Berkeley Lab	2,000,000
Integrating Climate Change into Air Quality Modeling	Kaduwela, Ajith	UC Davis	1,000,000
A Computational Study of the Atmospheric Boundary Layer with Mesoscale-Microscale Coupling	Balakrishnan, Ramesh	Argonne Lab	1,000,000
A Multi-Decadal Reforecast Data Set to Improve Weather Forecasts for Renewable Energy Applications	Hamill, Thomas	NOAA	100,000
Investigating Bioenergy with Carbon Capture and Sequestration using the California TIMES Model Projected through 2050	Jenn, Alan	UC Davis	100,000
Cyclotron Road - Wave Carpet Wave Energy Converter	Lehmann, Marcus	Berkeley Lab	100,000
Monte Carlo Simulation of Greenhouse Gas Emissions from Biofuels-Induced Land Use Change	Plevin, Richard	UC Berkeley	100,000
GEOSCIENCES			
Viscoelastic Relaxation and Delayed Earthquake Triggering Along the North Anatolian Fault, the Japan Subduction Zone and the Himalayan Range Front	DeVries, Phoebe	Harvard Univ	500,000
Emergent Fracture Propagation in Porous Solids	Edmiston, John	Berkeley Lab	250,000
Feresa: HPC Scaling of Realistic Geothermal Reservoir Models	Bernstein, David	Berkeley Lab	150,000

PROJECT	PI	ORGANIZATION	ALLOCATION
MATERIALS SCIENCE			
Exploring Dynamics at Interfaces Relevant to Mg-ion Electrochemistry from First Principles	Prendergast, David	Berkeley Lab	14,500,000
Quantum Simulations of Nanoscale Energy Conversion	Grossman, Jeffrey	MIT	12,000,000
Next Generation Semiconductors for Low Power Electronic Devices	Lee, Byounghak	Berkeley Lab	12,000,000
Computing Resources for the Center for the Computational Design of Functional Layered Materials	Perdew, John	Temple University	10,000,000
Hot Carrier Collection in Thin Film Silicon with Tailored Nanocrystalline/ Amorphous Structure	Lusk, Mark	Colorado School Mines	8,000,000
Testing the Fundamental Limits on the Number of Inorganic Compounds in Nature	Ceder, Gerbrand	MIT	7,500,000
Accelerated Discovery and Design of Complex Materials	Kent, Paul	Oak Ridge	2,000,000
Data Mining Discovery of Thermoelectrics and Inverse Band Structure Design Principles	Jain, Anubhav	Berkeley Lab	1,000,000
Topological Phases in Correlated Systems	Vishwanath, Ashvin	Berkeley Lab	700,000
ACCELERATOR DESIGN			
Atomic-Scale Modeling of Fracture Along Grain Boundaries with He Bubbles	Xu, Steven	Kinectrics Inc.	1,000,000
ASTROPHYSICS			
Radiation Hydrodynamic Simulations of Massive Star Envelope	Jiang, Yanfei	Smithsonian Astro Observatory	6,169,000

Publications

Staying ahead of the technological curve, anticipating problems and developing proactive solutions are part of NERSC's culture. Many staff members collaborate on computer science research projects, scientific code development and domain-specific research. They also participate in professional organizations and conferences and contribute to journals and proceedings. The NERSC user community benefits from these activities as they are applied to systems, software and services at NERSC and throughout the HPC community. In 2015, NERSC users reported 2,078 refereed papers based on work performed at the center, reflecting the breadth of scientific activity supported by NERSC. In addition, NERSC staff contributed some 120 papers to scientific journals and conferences in 2015, showcasing our increasing involvement in technology and software development designed to enhance the utilization of HPC across a broad range of science and data-management applications. To see a comprehensive list of publications and presentations by NERSC staff in 2015 go to http://1.usa.gov/1f3uFIv.

Appendix A

NERSC Users Group Executive Committee

OFFICE OF ADVANCED SCIENTIFIC COMPUTING RESEARCH

Milan Curcic, *University of Miami*
Jeff Hammond, *Intel*
Brian Van Straalen, *Lawrence Berkeley National Laboratory*

OFFICE OF BASIC ENERGY SCIENCES

Eric Bylaska, *Pacific Northwest National Laboratory*
Monojoy Goswami, *Oak Ridge National Laboratory*
Paul Kent, *Oak Ridge National Laboratory*

OFFICE OF BIOLOGICAL AND ENVIRONMENTAL RESEARCH

Samson Hagos, *Pacific Northwest National Laboratory*
Gary Strand, *National Center for Atmospheric Research*
David Wong, *U.S. Environmental Protection Agency*

OFFICE OF FUSION ENERGY PHYSICS

Christopher Holland, *University of California, San Diego*
Nathan Howard, *Massachusetts Institute of Technology*
David Green, *Oak Ridge National Laboratory*

OFFICE OF HIGH ENERGY PHYSICS

Ted Kisner, *Lawrence Berkeley National Laboratory*
Zarija Lukic, *Lawrence Berkeley National Laboratory*
Frank Tsung, *University of California, Los Angeles (Chair)*

OFFICE OF NUCLEAR PHYSICS

Balint Joo, *Jefferson Lab*
Nicolas Schunck, *Lawrence Livermore National Laboratory*
Michael Zingale, *Stony Brook University*

MEMBERS AT LARGE

James Amundson, *Fermilab*
Carlo Benedetti, *Lawrence Berkeley National Laboratory*
David Hatch, *University of Texas, Austin*

Appendix B

Office of Advanced Scientific Computing Research

The mission of the Advanced Scientific Computing Research (ASCR) program is to discover, develop and deploy computational and networking capabilities to analyze model, simulate and predict complex phenomena important to the Department of Energy (DOE). A particular challenge of this program is fulfilling the science potential of emerging computing systems and other novel computing architectures, which will require numerous significant modifications to today's tools and techniques to deliver on the promise of exascale science.

To accomplish its mission and address those challenges, the ASCR program is organized into two subprograms: Mathematical, Computational and Computer Sciences Research; and High Performance Computing and Network Facilities

The Mathematical, Computational and Computer Sciences Research subprogram develops mathematical descriptions, models, methods and algorithms to describe and understand complex systems, often involving processes that span a wide range of time and/or length scales. The subprogram also develops the software to make effective use of advanced networks and computers, many of which contain thousands of multi-core processors with complicated interconnections, and to transform enormous data sets from experiments and simulations into scientific insight.

The High Performance Computing and Network Facilities subprogram delivers forefront computational and networking capabilities and contributes to the development of next-generation capabilities through support of prototypes and testbeds.

Berkeley Lab thanks the program managers with direct responsibility for the NERSC program and the research projects described in this report:

ASCR PROGRAM

Steve Binkley, *Associate Director, ASCR*

Michael Martin, *Fellow*

Julie Stambaugh, *Financial Management Specialist*

Lori Jernigan, *Program Support Specialist*

FACILITIES DIVISION

Barbara Helland, *Director*

Vince Dattoria, *Computer Scientist, ESnet Program Manager*

Betsy Riley, *Computer Scientist, ALCF Program Manager*

Robinson Pino, *Computer Scientist, REP Program Manager*

Claire Cramer, *Physical Scientist, REP Program Manager*

Carolyn Lauzon, *Physical Scientist, ALCC Program Manager*

Dave Goodwin, *Physical Scientist, NERSC Program Manager*

Christine Chalk, *Physical Scientist, ORLC Program Manager, CSGF Program Manager*

Sally McPherson, *Program Assistant*

RESEARCH DIVISION

William Herrod, *Director*

Teresa Beachley, *Program Assistant*

Randall Laviolette, *Physical Scientist, SciDAC Application Partnerships*

Thomas Ndousse-Fetter, *Computer Scientist, Network Research*

Ceren Susut, *Physical Scientist, SC Program SAPs*

Rich Carlson, *Computer Scientist, Collaboratories/ Middleware*

Steven Lee, *Physical Scientist, SciDAC Institutes*

Lucy Nowell, *Computer Scientist, Data and Visualization*

Sonia Sachs, *Computer Scientist, Extreme Scale*

Angie Thevenot, *Program Assistant*

Appendix C

Acronyms and Abbreviations

ACM
Association for Computing Machines

ACS
American Chemical Society

ALCC
ASCR Leadership Computing Challenge

ALS
Advanced Light Source, Lawrence Berkeley National Laboratory

ANL
Argonne National Laboratory

API
Application Programming Interface

APS
American Physical Society

ASCII
American Standard Code for Information Interchange

ASCR
Office of Advanced Scientific Computing Research

BELLA
Berkeley Lab Laser Accelerator

BER
Office of Biological and Environmental Research

BES
Office of Basic Energy Sciences

BNL
Brookhaven National Laboratory

C^3
Computational Cosmology Center, Lawrence Berkeley National Laboratory

CAL
Computer Architecture Laboratory

CARB
California Air Resources Board

CCM
Cluster Compatibility Mode

CdTe
Cadmium Telluride

CeO_2
Cerium Dioxide

CERN
European Organization for Nuclear Research

CESM
Community Earth Systems Model

CFD
Computational Fluid Dynamics

CLE
Cray Linux Environment

CMB
Cosmic Microwave Background

CMIP5
Coupled Model Intercomparison Project, 5th Phase

CO_2
Carbon dioxide

CPU
Central Processing Unit

CRD
Computational Research Division, Lawrence Berkeley National Laboratory

CRT
Computational Research and Theory Facility, Lawrence Berkeley National Laboratory

CSE
Computational Science and Engineering

DARPA
Defense Advanced Research Projects Agency

DESI
Dark Energy Spectroscopic Instrument

DFT
Density Functional Theory

DME
Dimethyl Ether

DNS
Direct Numerical Simulation

DOE
U.S. Department of Energy

DOI
Digital Object Identifier

DSL
Dynamic Shared Library

DTN
Data Transfer Node

DVS
Data Virtualization Service

EFRC
DOE Energy Frontier Research Center

EMSL
Environmental Molecular Science Laboratory, Pacific Northwest National Laboratory

EPSI
SciDAC Center for Edge Physics Simulations

ERD
Earth Sciences Division, Lawrence Berkeley National Laboratory

ERT
Empirical Roofline Toolkit

ESnet
Energy Sciences Network

eV
Electron Volts

FDM
Finite Difference Method

FES
Office of Fusion Energy Sciences

FLOPS
Floating Point Operations

FTP
File Transfer Protocol

GB
Gigabytes

Gbps
Gigabits Per Second

GPU
Graphics Processing Unit

HEP
Office of High Energy Physics

HPC
High Performance Computing

HPSS
High Performance Storage System

HTML
Hypertext Markup Language

HTTP
Hypertext Transfer Protocol

IEEE
Institute of Electrical and Electronics Engineers

InN
Indium Nitride

IPCC
Intel Parallel Computing Center; Intergovernmental Panel on Climate Change

iPTF
intermediate Palomar Transient Factory

ITER
An international fusion energy experiment in southern France

ITG
Ion Temperature Gradient

IXPUG
Intel Xeon Phi Users Group

JCESR
Joint Center for Energy Research Storage

JET
Joint European Torus

JGI
Joint Genome Institute

LED
Light-emitting Diode

LHCII
Light-harvesting Complex II

LANL
Los Alamos National Laboratory

LLNL
Lawrence Livermore National Laboratory

MANTISSA
Massive Acceleration of New Technologies in Science with Scalable Algorithms

MeV
Mega-electron Volts

MIT
Massachusetts Institute of Technology

MOF
Metal Oxide Framework

MPP
Massively Parallel Processing

MSI
Mass Spectrometry Imaging

NCAR
National Center for Atmospheric Research

NESAP
NERSC Exascale Scientific Application Program

NEXAFS
Near Edge X-ray Absorption Fine Structure

NGF
NERSC Global Filesystem

NIH
National Institutes of Health

NIM
NERSC Information Management

NOAA
National Oceanic and Atmospheric Administration

NP
Office of Nuclear Physics

NPLQCD
Nuclear Physics with Lattice QCD

NSF
National Science Foundation

NUG
NERSC Users Group

NVRAM
Non-volatile Random Access Memory

NWB
Neurodata Without Borders

OLCF
Oak Ridge Leadership Computing Facility

OpenMSI
Open Mass Spectrometry Imaging

OSF
Oakland Scientific Facility

PDACS
Portal for Data Analysis services for Cosmological Simulations

PDSF
Parallel Distributed Systems Facility, NERSC

PI
Principal Investigator

PIC
Particle-In-Cell Simulations

PB
Petabytes

PSII
Photosystem II

PNNL
Pacific Northwest National Laboratory

PPPL
Princeton Plasma Physics Laboratory

QCD
Quantum Chromodynamics

RIPA
R Analytics for Image Analysis

SC
DOE Office of Science

SciDAC
Scientific Discovery Through Advanced Computing

SDN
Software-defined Networking

SDSS
Sloan Digital Sky Survey

Si
Silicon

SiO_2
Silicon Oxide

SIAM
Society for Industrial and Applied Mathematics

TACC
Texas Advanced Computing Center

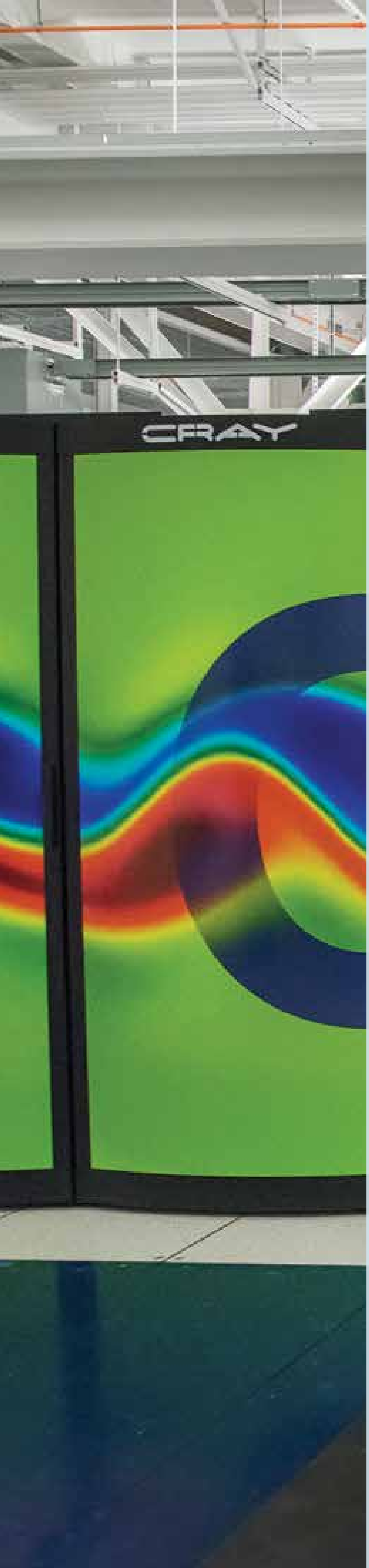
TB
Terabytes

TECA
Toolkit for Extreme Climate Analysis

UO_2
Uranium Dioxide

URL
Universal Resource Locator

XMAS
X-ray Microdiffraction Analysis Software



FOR MORE INFORMATION
ABOUT NERSC, CONTACT:

Jon Bashor
NERSC Communications
1 Cyclotron Road
Berkeley Lab, MS 59R3022B
Berkeley, CA 94720-8148
Email jbashor@lbl.gov
Phone (510) 486-5849
Fax (510) 486-4300

NERSC's Web Site
nersc.gov

NERSC Annual Report Editor
Kathy Kincade

Contributing Writers and Editors
Jon Bashor, Kathy Kincade, Linda Vu and Margie Wylie
Berkeley Lab Computing Sciences Communications Group

Katie Antypas, Elizabeth Bautista, Shane Canon, Jack Deslippe,
Richard Gerber, Rebecca Hartman-Baker, Glenn Lockwood,
Jackie Scoggins
NERSC

Dan Krotz
Berkeley Lab Public Affairs Communications and
Media Relations Group

Robert Sanders
University of California, Berkeley

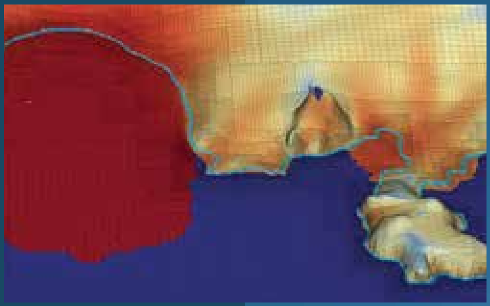
Raphael Rosen
Princeton Plasma Physics Laboratory

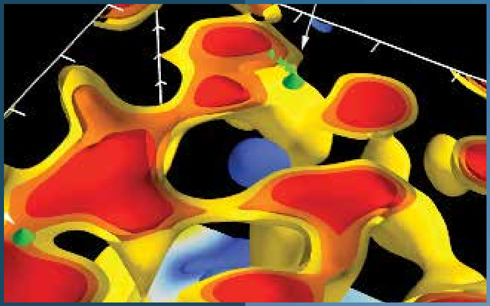
Siv Schwink
University of Illinois at Urbana-Champaign

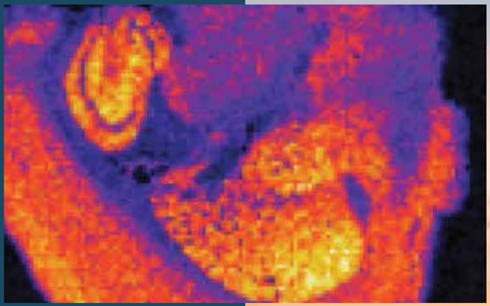
Design
Design, layout, illustraton, photography and printing coordination
Berkeley Lab Public Affairs Creative Services Group

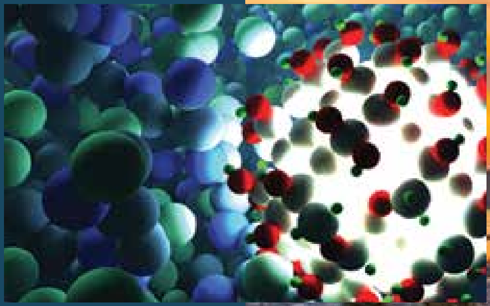
Image: Roy Kaltschmidt, Lawrence Berkeley National Laboratory

16-NERSC-2695









BERKELEY LAB

U.S. DEPARTMENT OF ENERGY

Office of Science