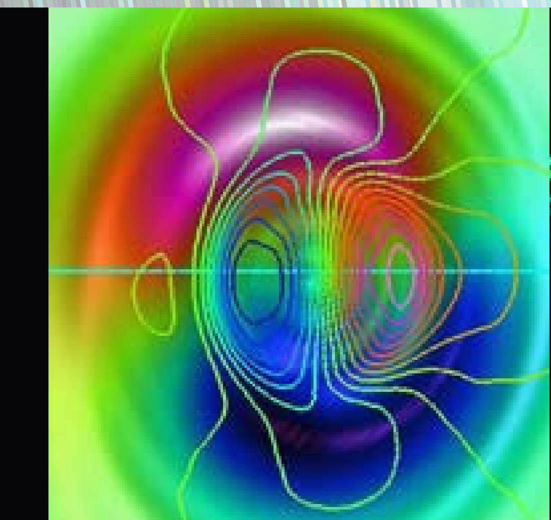
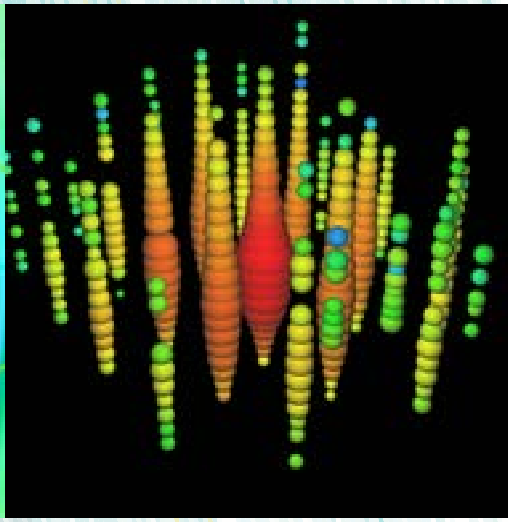
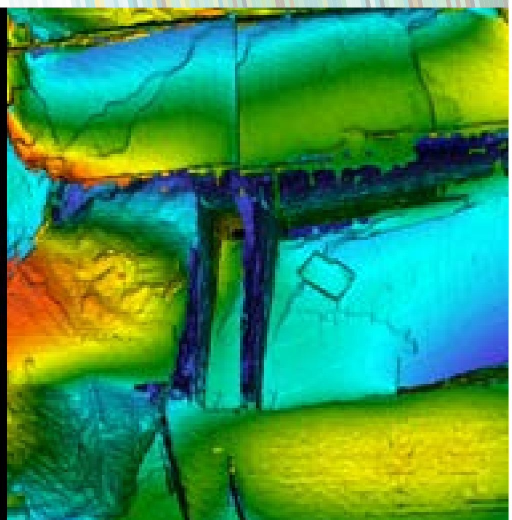
2013 Annual Report

National Energy Research Scientific Computing Center







2013 Annual Report

National Energy Research Scientific Computing Center

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.

Cover Image Credits:

FRONT COVER, LEFT TO RIGHT: Linda Sugiyama, Massachusetts Institute of Technology; IceCube Collaboration; David Trebotich, Lawrence Berkeley National Laboratory. BACK COVER, LEFT TO RIGHT: Menghao Wu, Virginia Commonwealth University; Planck Collaboration; James Sethian, Robert Saye, Lawrence Berkeley National Laboratory, UC Berkeley

Table of Contents

The Year In Perspective

My first full year as Director of NERSC has been an exciting and gratifying one. I have endeavored to keep NERSC focused on enabling scientific discovery, while evaluating and overseeing the implementation of new technologies that are improving NERSC systems and infrastructure, and new service innovations designed to meet the increasingly complex needs of our scientific users.

As the high-end scientific computing facility for the Department of Energy's Office of Science (DOE SC), we serve the computational science needs of DOE SC's six program offices. With more than 5,000 users from universities, national laboratories and industry, we support the largest and most diverse research community of any computing facility within the DOE complex. NERSC provides large-scale, state-of-the-art computing for DOE's unclassified research programs in alternative energy sources, climate change, energy efficiency, environmental science and other fundamental science areas.

NERSC's primary mission is to accelerate scientific discovery at the DOE SC through high performance computing and data analysis. In 2013 we provided 1.25 billion computer hours to our users, and taking advantage of uncharged time on NERSC's newest supercomputing system they were able to use 2.1 billion hours. 2013 proved to be a productive year for scientific discovery at the center; our users published 1,977 refereed papers and 18 journal cover stories based on computations performed at NERSC. In addition, Martin Karplus of Harvard University, who has been computing at NERSC since 1998, was awarded a Nobel Prize in Chemistry for his contributions to the field of computational chemistry.

Several other NERSC projects were also honored in 2013. In December, Physics World's "Breakthrough of the Year" went to the IceCube South Pole Neutrino Observatory for making the first observations of high-energy cosmic neutrinos, an achievement enabled in part by NERSC resources. In fact, NERSC was involved in three of Physics World's top 10 breakthroughs of 2013—the other two being the European Space Agency's Planck space telescope, which revealed new information about the age and composition of the universe, and the South Pole Telescope, which made the first detection of a subtle twist in light from the Cosmic Microwave Background. The Planck Collaboration also received the NERSC Achievement Award for High-Impact Science.

In addition, WIRED magazine listed two NERSC-related projects as its top science discoveries of 2013: the IceCube results and the final findings of the NASA Kepler space telescope that one in five Sun-like stars in our galaxy has an Earth-sized planet orbiting it at a habitable distance. And the Materials Project, one of our most popular science gateways, was featured as a "world changing idea" in a November 2013 Scientific American cover story, "How Supercomputers Will Yield a Golden Age of Materials Science."

We also continued our close collaboration with the Joint Genome Institute, which has given NERSC important insights into the needs of DOE experimental facilities. Our JGI database now has over 3 billion genes, and JGI scientists can process 100 million genes in a few days. We are also working closely with the Advanced Light Source (ALS) at Berkeley Lab and deployed a prototype data pipeline that automatically transfers data from ALS to NERSC for analysis and sharing. We successfully completed a data pilot with the SLAC National Accelerator Laboratory in which we processed a 150 terabyte dataset at NERSC from the Linac Coherent Light Source. We are looking forward to expanding these and other ongoing collaborations in 2014 and beyond.

Another highlight of 2013 was the installation of our newest supercomputing platform, a Cray XC30 system named Edison in honor of American inventor Thomas Alva Edison. Scientists from around the globe eagerly queued up to take advantage of the new supercomputer's capabilities. Edison was the first Cray supercomputer with Intel processors, a new Aries interconnect and a dragonfly topology. The system was designed to optimize data motion—the primary bottleneck for many of our applications—as opposed to peak speed. It has very high memory bandwidth, interconnect speed and bisection bandwidth. In addition, each node has twice the memory of many leading systems. This combination of fast data motion and large memory per node make it well suited for both our traditional HPC workload and newly emerging data intensive applications.

Other technical highlights in 2013 included the upgrade of our border network to 100 gigabits per second and the increase in performance to up to 80 gigabytes per second in the globally accessible scratch file system. Innovations in 2013 included achieving high energy efficiency for running Edison, introducing new tools to support data-intensive science and preparing our users for the transition to exascale.

As always, user support is a high priority at NERSC. In 2013 our efforts to enhance NERSC's user support services continued to prove effective: The "Overall Satisfaction with NERSC" score for 2013 was the second-highest ever recorded in the 15 years our user survey has been in its current form. User support highlights included working with Cray to improve library performance on Edison, providing the results of I/O monitoring to help users diagnose and fix I/O performance issues, assisting users in transitioning to energy-efficient manycore architectures, providing a new parallel scientific database (SciDB) to users and helping users create new science portals to enhance collaboration and team science.

The demands from larger and more detailed simulations, massive numbers of simulations and the explosion in the size and number of experimental datasets mean there is no end in sight to the need for NERSC resources. Overcoming the challenges of energy efficiency, concurrency, memory balance and resilience will require the creativity and expertise of the NERSC staff and greater collaborations with vendor partners, the user community and other supercomputing centers. Staying ahead of technology, anticipating problems and developing solutions that are effective for the broad science workload are part of NERSC's culture.

As we look ahead to 2014—the 40th anniversary of NERSC—we are excited about our forthcoming move into the new Computational Research and Theory (CRT) facility on the main Berkeley Lab site and preparing to deploy our next supercomputer, NERSC-8, which is named Cori in honor of Gerty Cori, the first Amercian woman to win a Nobel Prize in science. CRT is a highly energy-efficient, state-of-the-art computing facility that will provide over 40 megawatts of power and 30,000 square feet of space for computing and storage. In 2013 we released the request for proposals for NERSC-8 in collaboration with the Alliance for Computing at the Extreme-Scale (ACES), a partnership between Los Alamos and Sandia National Laboratories. ACES is planning to deploy their next supercomputer, Trinity, in the same time frame as Cori. NERSC has been working with ACES for several years—we deployed Hopper at the same time ACES installed a very similar system, Cielo—and we look forward to continuing this very successful collaboration.

With all the advanced technologies we deploy, it is still people who make the difference. As always, I am grateful to our DOE SC sponsors for their continued support, our users for the science they produce using NERSC resources and the NERSC staff who make NERSC such a successful high performance computing center.

Sudip Dosanjh
NERSC Division Director

Research News

The National Energy Research Scientific Computing Center (NERSC) is the high-end scientific computing facility for the Department of Energy's Office of Science (DOE SC). With more than 5,000 users from universities, national laboratories and industry, NERSC supports the largest and most diverse research community of any computing facility within the DOE complex.

NERSC provides large-scale, state-of-the-art computing for DOE's unclassified research programs in alternative energy sources, climate change, energy efficiency, environmental science and other fundamental science areas within the DOE mission.

NERSC's primary mission is to accelerate scientific discovery at the DOE SC through high performance computing and data analysis. In 2013, our users reported 1,977 refereed papers and 18 journal cover stories based on computations performed at NERSC.

This section presents a selection of research highlights showing the breadth of science supported by the center.

Edison Opens New Doors for Early Users

From improving carbon sequestration models to creating virtual universes, NERSC's newest supercomputer is a hit

"The memory bandwidth makes a huge difference in performance for our code."

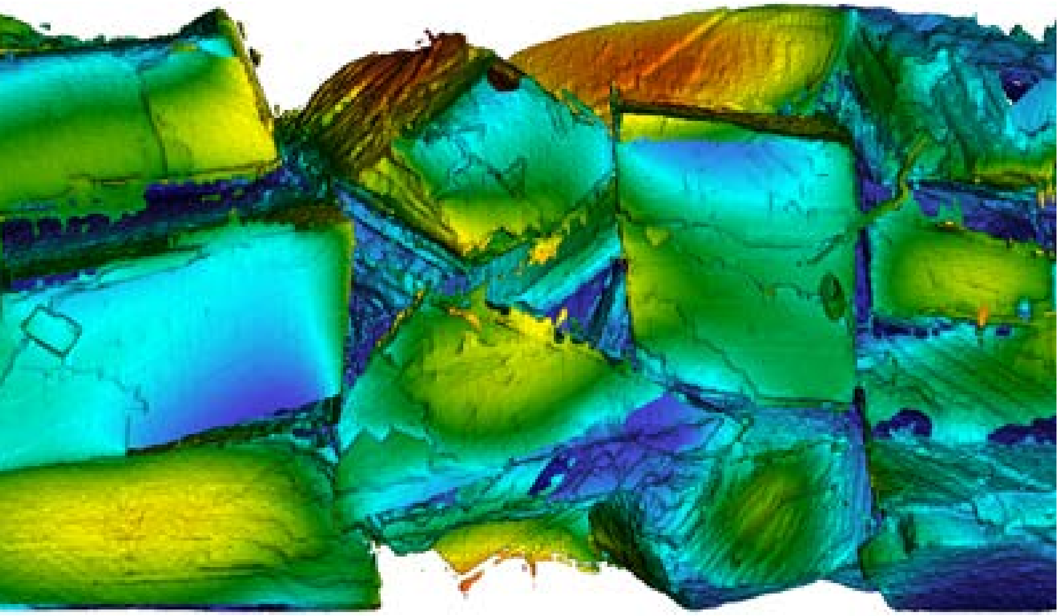
BEFORE ANY SUPERCOMPUTER IS ACCEPTED AT NERSC, the system is rigorously tested as scientists are invited to put the system through its paces during an "early science" phase. While the main aim of this period is to test the new system, many scientists are able to use the time to significantly advance their work. Here's a look at the work NERSC's new Edison Cray XC30 helped some early users accomplish in 2013.

The Fate of Sequestered CO_2

David Trebotich is modeling the effects of sequestering carbon dioxide (CO_2) deep underground. The aim is to better understand the physical and chemical interactions between CO_2, rocks and the minute, saline-filled pores through which the gas migrates. This information will help scientists understand how much we can rely on geologic sequestration as a means of reducing greenhouse gas emissions, which cause climate change.

Unlike today's large-scale models, which are unable to resolve microscale features, Trebotich models the physical and chemical reactions happening at resolutions of hundreds of nanometers to tens of microns. His simulations cover only a tiny area—a tube just a millimeter wide and not even a centimeter long—but in exquisitely fine detail.

"We're definitely dealing with big data and extreme computing," said Trebotich, a computational scientist in Berkeley Lab's Computational Research Division. "The code Chombo-Crunch generates datasets of one terabyte for just a single, 100-microsecond time-step, and we need to do 16 seconds of that to match a 'flow-through' experiment." Carried out by the Energy Frontier Research Center for Nanoscale Control of Geologic CO_2, the experiment captured effluent concentrations due to injecting a solution of dissolved carbon dioxide through a tube of crushed calcite.



◀ The fine detail of this simulation—which shows computed pH on calcite grains at 1 micron resolution—is necessary to better understand what happens when the greenhouse gas carbon dioxide is injected underground rather than being released into the atmosphere to exacerbate climate change. DAVID TREBOTICH, LAWRENCE BERKELEY NATIONAL LABORATORY

Edison's high memory-per-node architecture means that more of each calculation (and the resulting temporary data) can be stored close to the processors working on it. Trebotich was invited to run his codes to help test the machine and his simulations ran 2.5 times faster than on Hopper, a Cray XE6 and the previous flagship system, reducing the time it takes him to get a solution from months to just weeks.

"The memory bandwidth makes a huge difference in performance for our code," Trebotich said.

The aim is to eventually merge such finely detailed modeling results with large-scale simulations for more accurate models. Trebotich is also working on extending his simulations to shale gas extraction using hydraulic fracturing. The code framework could also be used for other energy applications, such as electrochemical transport in batteries.

Better Combustion for New Fuels

Jackie Chen and her research team at Sandia National Laboratories are using Edison to investigate how to improve combustion using new engine designs and fuels such as biodiesel, hydrogen-rich "syngas" from coal gasification and alcohols like ethanol. Her group models the behavior of burning fuels by simulating some of the underlying chemical and mixing conditions found in these combustion engines in simple laboratory configurations, using a direct numerical simulation code developed at Sandia.

During Edison's early science period, Chen and colleagues modeled hydrogen and oxygen mixing and burning in a transverse jet configuration commonly employed by gas turbine combustors in aircraft engines and industrial power plants. Their simulations were performed in tandem with experimentalists to understand the complex flow field affected by reaction and how it manifests itself in instabilities. This information is critical to understanding both the safety and performance of the device.

The team was able to run their direct numerical simulation code, S3D, on 100,000 processors at once on Edison and saw a four- to five-fold performance improvement over Hopper, Chen said.

"In the end, what really matters isn't the technical specifications of a system," said Chen. "What really matters is the science you can do with it, and on that count, Edison is off to a promising start."

Our Universe: The Big Picture

Zarija Lukic, a cosmologist with Berkeley Lab's Computational Cosmology Center (C3), used Edison to model mega-parsecs of space in an attempt to understand the large-scale structure of the universe.

Since we can't step outside our own universe to see its structure, cosmologists like Lukic examine the light from distant quasars and other bright light sources for clues. When light from these quasars passes through clouds of hydrogen, the gas leaves a distinctive pattern in the light's spectrum. By studying these patterns (which look like clusters of squiggles called a Lyman alpha forest), scientists can better understand what lies between us and the light source, revealing the process of structure formation in the universe.

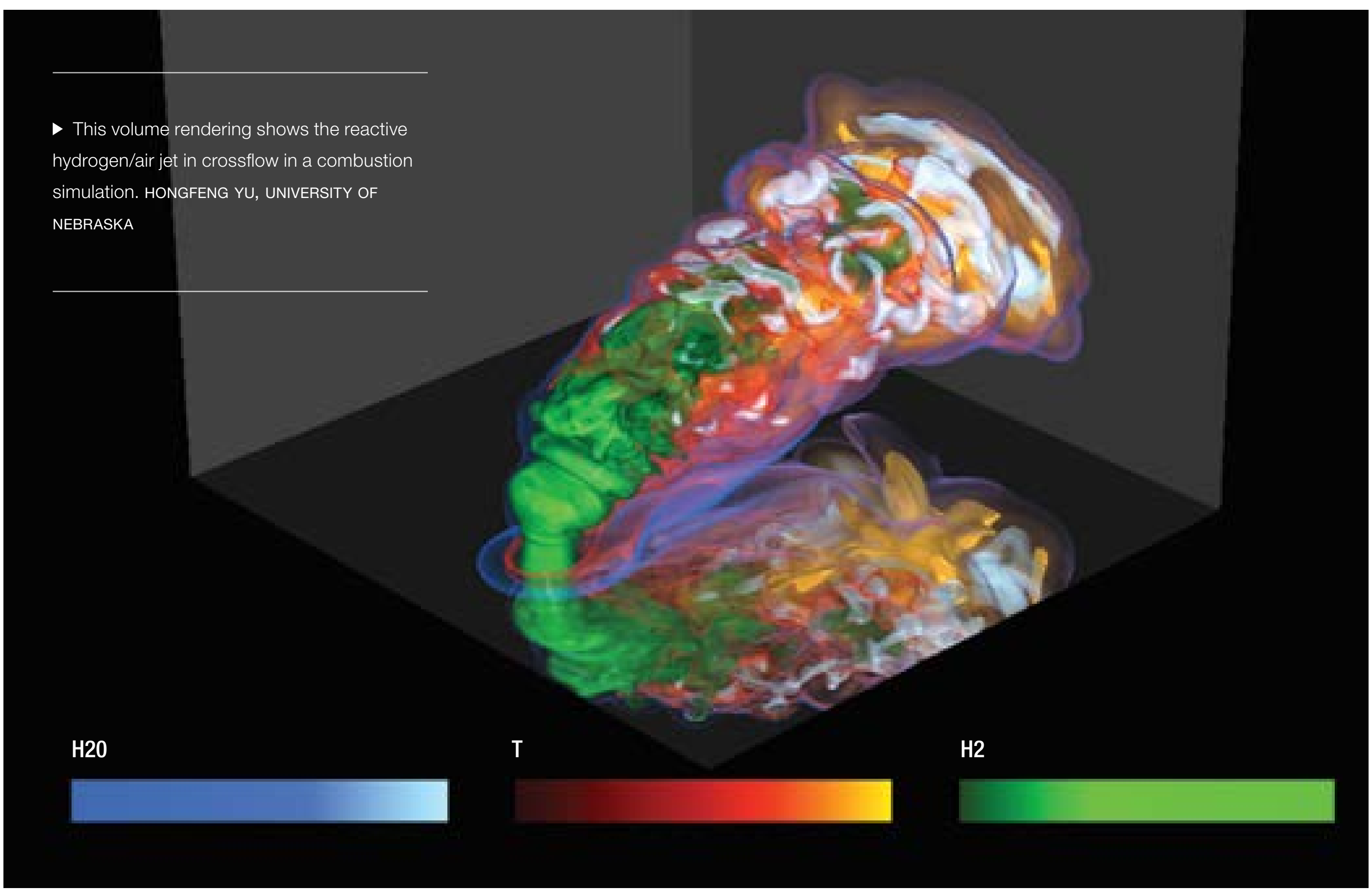
Researchers use a variety of possible cosmological models, calculating for each the interplay between dark energy, dark matter and the baryons that flow into gravity wells to become stars and galaxies. Cosmologists can then compare these virtual universes with real observations. Using the Nyx code developed by Berkeley Lab's Center for Computational Sciences and Engineering, Lukic and colleagues are able to create virtual "what-if" universes to help cosmologists fill in the blanks in their observations.

"The ultimate goal is to find a single physical model that fits not just Lyman alpha forest observations but all the data we have about the nature of matter and the universe, from cosmic microwave background measurements to results from experiments like the Large Hadron Collider," Lukic said.

Working with 2 million early hours on Edison, Lukic and collaborators performed one of the largest Lyman alpha forest simulations to date: the equivalent of a cube measuring 370 million light years on each side.

"With Nyx on Edison we're able—for the first time—to get to volumes of space large enough and with resolution fine enough to create models that aren't thrown off by numerical biases," he said. Lukic expects his work on Edison will become "the gold standard for Lyman alpha forest simulations."

Large-scale simulations such as these will be key to interpreting the data from many upcoming observational missions, including



▶ This volume rendering shows the reactive hydrogen/air jet in crossflow in a combustion simulation. HONGFENG YU, UNIVERSITY OF NEBRASKA the Dark Energy Spectroscopic Instrument, he added.

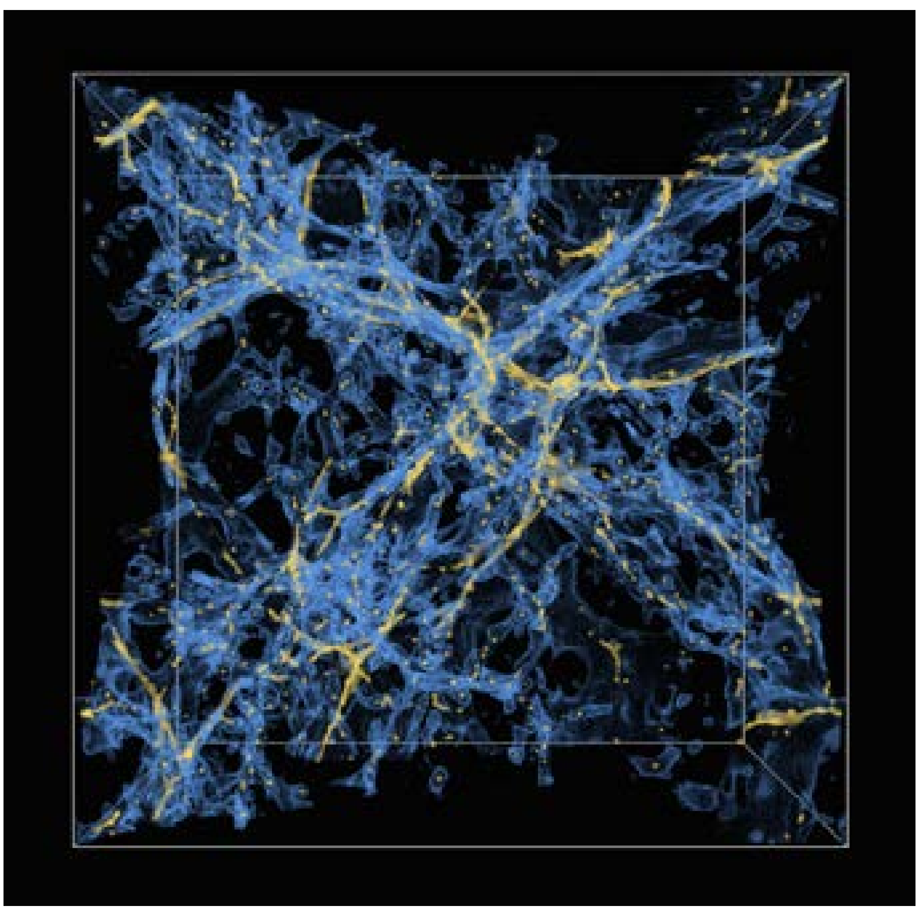
Growing Graphene and Carbon Nanotubes

Vasilii Artyukhov, a post-doctoral researcher in materials science and nanoengineering at Rice University, used 5 million processor hours during Edison's testing phase to simulate how graphene and carbon nanotubes are "grown" on metal substrates using chemical vapor deposition.

Graphene (one-atom thick sheets of graphite) and carbon nanotubes (essentially graphene grown in tubes) are considered "wonder materials" because they are light but strong and electrically conductive, among other properties. These nanomaterials offer the prospects of vehicles that use less energy; thin, flexible solar cells; and more efficient batteries—that is, if they can be produced on an industrial scale.

Today's methods for growing graphene often produce a mish-mash of nanotubes with different properties; for instance, semiconducting types are useful for electronics and metallic types for high-current conduits. For structural purposes the type is less important than the length. That's why the group is also investigating how to grow longer tubes on a large scale.

For Artyukhov's purposes, Edison's fast connections between processor nodes has allowed him to run his code on twice as many processors at speeds twice as fast as before. The Rice University researcher is also using molecular dynamics codes, "which aren't as computationally demanding, but because Edison is so large, you can run many of them concurrently," he said.



◀ Using the Nyx code on Edison, scientists were able to run the largest simulation of its kind (370 million light years on each side) showing neutral hydrogen in the large-scale structure of the universe. The blue webbing represents gas responsible for Lyman-alpha forest signals. The yellow are regions of higher density, where galaxy formation takes place. CASEY STARK, LAWRENCE BERKELEY NATIONAL LABORATORY

IceCube IDs ‘Bert’ and ‘Ernie’

NERSC computers essential for finding neutrino event in timely manner at South Pole

“The large amount of computing resources at NERSC ... was essential to finding these neutrino events in a timely manner.”

DATA BEING COLLECTED AT “ICECUBE,” the neutrino observatory buried in the South Pole ice, nd analyzed at NERSC is helping scientists better understand cosmic particle accelerators.

In 2013, the IceCube Collaboration announced that it had observed 28 extremely high-energy events that constitute the first solid evidence for astrophysical neutrinos from outside our solar ystem, according to Spencer Klein, a senior scientist with Berkeley Lab and a long-time member of the IceCube Collaboration. These 28 events include two of the highest energy neutrinos ever reported, dubbed Bert and Ernie.

NERSC’s supercomputers were used to sift out neutrino signals from cosmic “noise” in the IceCube observations, which were published in the journal Science in November 2013.

“The large amount of computing resources at NERSC allowed me to set up my simulations and run them right away, which was essential to finding these neutrino events in a timely manner,” said Lisa Gerhardt, who at the time of her discovery was with Berkeley Lab’s Nuclear Science Division and is now with NERSC.

These results provide experimental confirmation that something is accelerating particles to energies above 50 trillion electron volts (TeV) and, in the cases of Bert and Ernie, exceeding one quadrillion electron volts (PeV). While not telling scientists what cosmic accelerators are or where they’re located, the IceCube results do provide them with a compass that can help guide them to the answers.

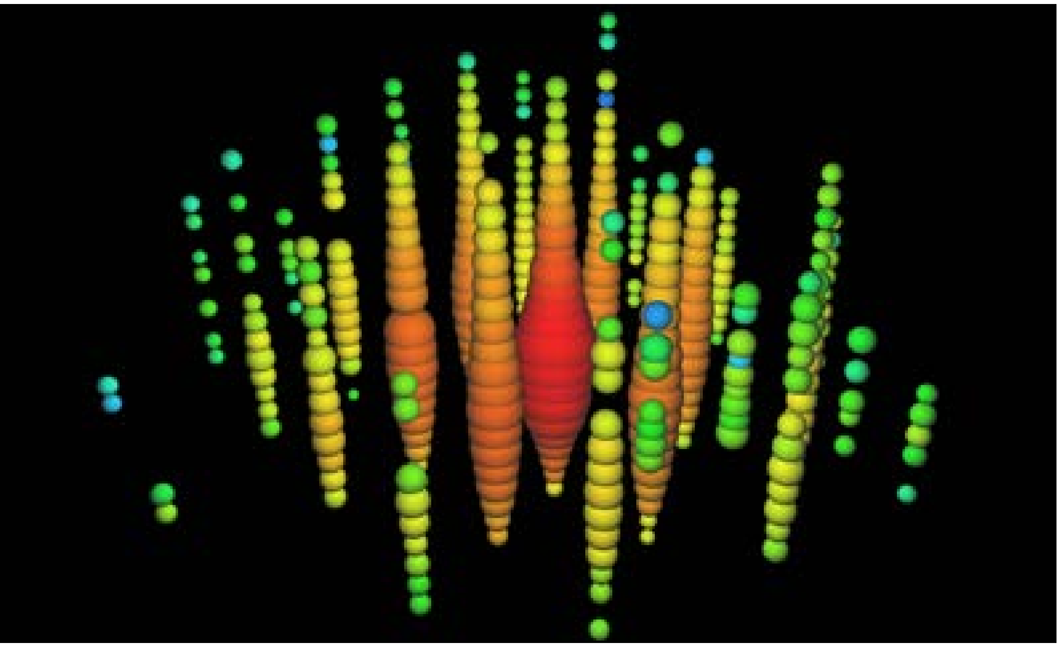
Cosmic accelerators have announced their presence through the rare ultra-high energy version of cosmic rays, which, despite the name, are electrically charged particles. It is known that ultra-high energy cosmic rays originate from beyond our solar system, but the electrical charge they carry bends their flight paths

PROJECT TITLE
Simulation and Analysis for IceCube

NERSC PI
Chang Hyon Has

LEAD INSTITUTION
Lawrence Berkeley National Laboratory

NERSC RESOURCES USED
Carver, Dirac, PDSF



◀ This event display shows "Ernie," one of two neutrino events discovered at IceCube whose energies exceeded one petaelectronvolt (PeV). The colors show when the light arrived, with reds being the earliest, succeeded by yellows, greens and blues. The size of the circle indicates the number of photons observed. ICECUBE COLLABORATION

as they pass through interstellar magnetic fields, making it impossible to determine where in the universe they originated. However, as cosmic ray protons and nuclei are accelerated, they interact with gas and light, resulting in the emission of neutrinos with energies proportional to the energies of the cosmic rays that produced them. Electrically neutral and nearly massless, neutrinos are like phantoms that travel through space in a straight line from their point of origin, passing through virtually everything in their path without being affected.

The IceCube observatory consists of 5,160 basketball-sized detectors called Digital Optical Modules (DOMs), which were conceived and largely designed at Berkeley Lab. The DOMS are suspended along 86 strings embedded in a cubic kilometer of clear ice starting one and a half kilometers beneath the Antarctic surface. Out of the trillions of neutrinos that pass through the ice each day, a couple of hundred will collide with oxygen nuclei, yielding the blue light of Cherenkov radiation that IceCube's DOMs detect.

The 28 high-energy neutrinos reported in Science by the IceCube Collaboration were found in data collected from May 2010 to May 2012. In analyzing more recent data, researchers discovered another event that was almost double the energy of Bert and Ernie, dubbed "Big Bird."

As to the identity of the mysterious cosmic accelerators, Klein thinks these early results from IceCube favor active galactic nuclei, the enormous particle jets ejected by a black hole after it swallows a star.

"The 28 events being reported are diffuse and do not point back to a source," Klein said, "but the big picture tends to suggest active galactic nuclei as the leading contender with the second leading contender being something we haven't even thought of yet."

DOE PROGRAM OFFICE
NP—Nuclear Theory

FULL STORY
http://www.nersc.gov/news-publications/news/science-news/2013/searching-for-cosmic-accelerators-via-icecube/

PUBLICATION
Ice Cube Collaboration, "Evidence for High-Energy Extraterrestrial Neutrinos at the IceCube Detector," Science, November 22, 2013, 342(6161), doi: 10.1126/science.1242856

Spotting Hotspot Volcanoes

NERSC computers helped scientists detect previously unknown channels of slow-moving seismic waves in Earth's upper mantle

"Without these new techniques, the study's analysis would have taken about 19 years to compute."

USING SUPERCOMPUTERS AT NERSC, scientists have detected previously unknown channels of slow-moving seismic waves in Earth's upper mantle. This discovery helps to explain how "hotspot volcanoes"—the kind that give birth to island chains like Hawaii and Tahiti—come to exist.

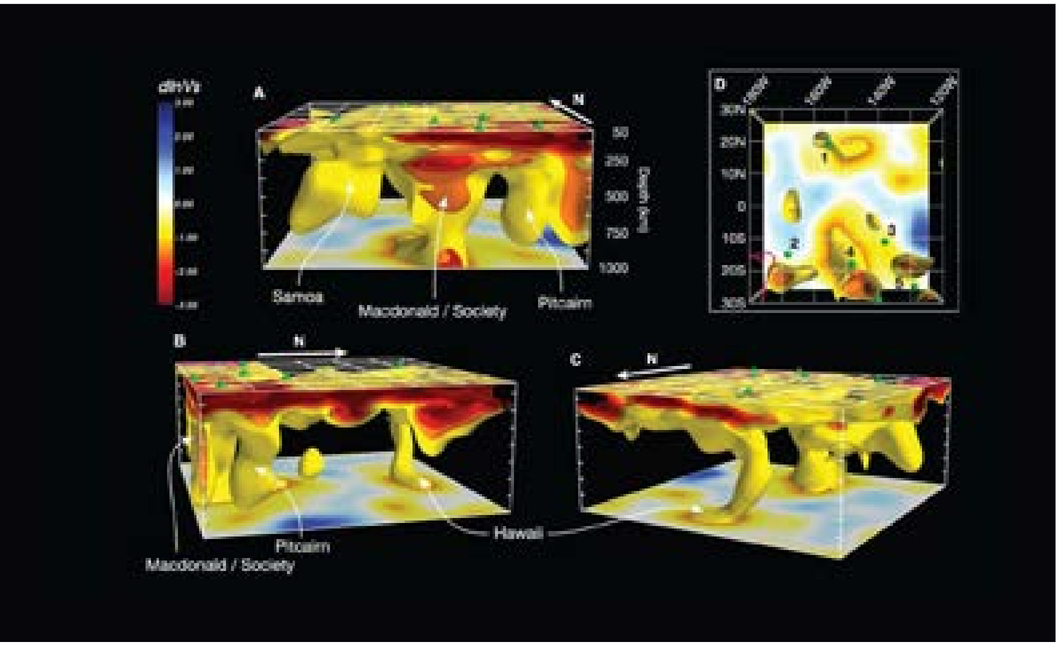
The researchers found these channels by analyzing seismic wave data from some 200 earthquakes using highly accurate computer simulations of how these waves travel through the Earth's mantle—the layer between the planet's crust and core. These analyses allowed them to make inferences about the structure of convection in the mantle, which is responsible for carrying heat from the planet's interior to the surface to create hotspot volcanoes.

"We now have a clearer picture that can help us understand the 'plumbing' of Earth's mantle responsible for hotspot volcano islands like Tahiti and Samoa," said Scott French, a University of California at Berkeley graduate student and lead author on the study, which was published in the journal Science.

Unlike volcanoes that emerge from collision zones between tectonic plates, hotspot volcanoes form in the middle of plates. Geologists hypothesize that mantle plumes—hot jets of material that rise from deep inside the mantle, perhaps near the core-mantle boundary—supply the heat that feed these mid-plate volcanic eruptions. But that model does not easily explain every hotspot volcano chain or why large swaths of the sea floor are significantly hotter than expected.

To learn more about the mantle's structure and convection, French looked at seismic data from hundreds of earthquakes around the world. These energetic waves travel for long distances below Earth's surface and are tracked by a global network of seismographs. Because the waves change as they travel through different materials, scientists can gain insights about the structure and composition of the substances beneath the

PROJECT TITLE	NERSC PI	LEAD INSTITUTION	NERSC RESOURCES USED
Global Full-Waveform Seismic Tomography	Barbara Romanowicz	UC Berkeley	Hopper



◀ This 3D view of the top 1,000 kilometers of Earth's mantle beneath the central Pacific shows the relationship between seismically slow "plumes" and channels imaged in the UC Berkeley study. Green cones on the ocean floor mark islands associated with "hotspot" volcanoes, such as Hawaii and Tahiti. BERKELEY SEISMOLOGICAL LABORATORY, UC BERKELEY

planet's surface by analyzing the altered waves. But because they cannot directly observe the structures lurking below Earth's surface, seismic tomography can be harder to interpret.

"To simulate seismic data for interpretation, people have traditionally used broad layers of approximation because it would have been too computationally heavy to run an exact model using numerical simulations," said French.

As a graduate student at UC Berkeley in 2010, Vedran Lekic used supercomputers at NERSC to develop a method to accurately model seismic wave data while keeping computing time manageable. French recently refined this approach, further improving the accuracy of the tomography, while also scaling the technique to solve larger problems. Together, these efforts resulted in higher-resolution images of patterns of convection in the Earth's mantle.

One simulation of one earthquake takes about 144 computer processor hours on NERSC's Hopper system, French said, "But we needed to run this simulation for 200 individual earthquakes to get an accurate seismic model. Our model improves with each run."

This tomographic model eventually allowed the team to find finger-like channels of low-speed seismic waves traveling in the mantle about 120 to 220 miles beneath the sea floor. These channels stretch about 700 miles wide and 1,400 miles apart. Seismic waves typically travel about 2.5 to 3 miles per second at this depth, but waves in the channels traveled about 4 percent slower. Because higher temperatures slow down seismic waves, scientists inferred that material in the channels must be hotter than surrounding material.

Without these new techniques, the study's analysis would have taken about 19 years to compute, Lekic noted.

DOE PROGRAM OFFICE

BES—Geosciences

FULL STORY

http://www.nersc.gov/news-publications/news/science-news/2013/new-model-of-earth-s-interior-reveals-clues-to-hotspot-volcanoes/

PUBLICATION

Scott French, Vedran Lekic and Barbara Romanowicz, "Waveform Tomography Reveals Channeled Flow at the Base of the Oceanic Asthenosphere," Science, 342(6155), 227-230 (2013), doi: 10.1126/science.1241514

Unlocking the Secrets of Solar Wind Turbulence

NERSC resources, visualizations help scientists forecast destructive space weather

"These simulations opened up a whole new area of physics for us."

AS INHABITANTS OF EARTH, our lives are dominated by weather. Not just in the form of rain and snow from atmospheric clouds, but also charged particles and magnetic fields generated by the Sun—a phenomenon called solar wind.

When strong magnetic storms occur on the Sun, tons of highly energetic particles are released into the solar wind. If these particles were free to hit the Earth, the radiation would cause life-threatening damage to our DNA, debilitate power grids, disrupt communications networks and damage electronic devices. Fortunately for us, the Earth's magnetic dipole field and magnetosphere act as an invisible shield barring these particles from entering the atmosphere.

However, this shield is not perfect; during particularly intense solar storms the magnetosphere can "crack," allowing charged particles to seep in and wreak havoc on the Earth's technological infrastructure—an event known as space weather. Scientists currently do not have the ability to accurately predict the severity of a space weather event or where it will have the most impact. But a team of researchers led by University of California, San Diego's (UCSD's) Homa Karimabadi is hoping to change that.

"There is an urgent need to develop accurate forecasting models, but one of the challenges in developing accurate forecasts is that the solar wind is turbulent, and the details of turbulence are not well understood," said Karimabadi, who heads the space plasma simulation group at UCSD. "A severe space-weather event can have dire financial and national security consequences."

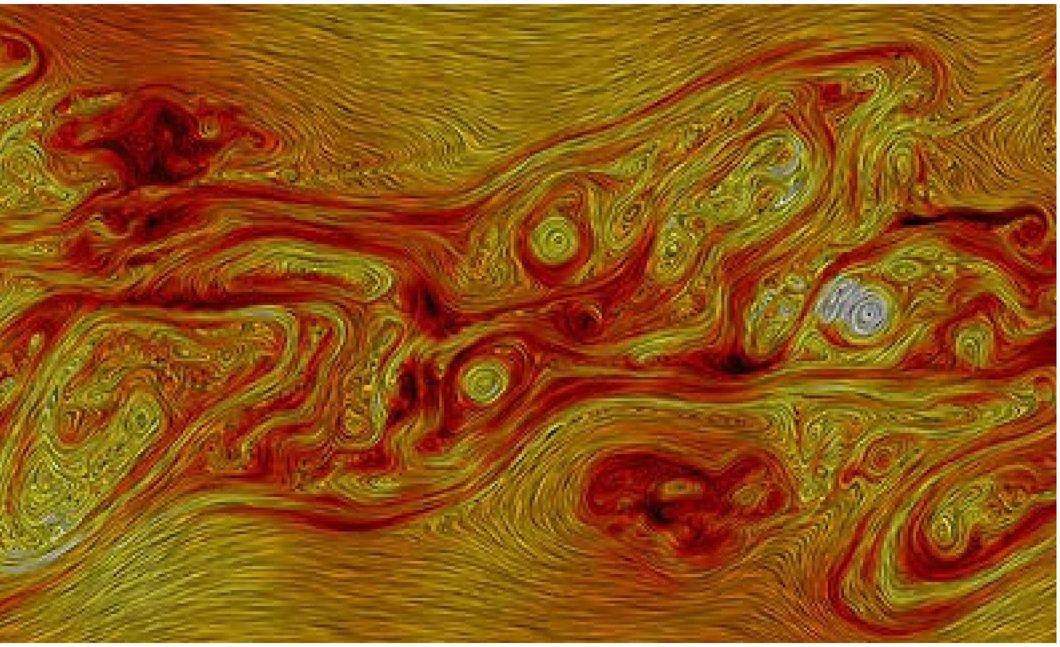
Because turbulence in the solar wind occurs on widely different scales of physics, it is especially difficult to study. But using supercomputers at the National Institute of Computational Sciences, Karimabadi and his colleagues for the first time managed to simulate all the scales of solar wind turbulence at once.

PROJECT TITLE
Petascale Kinetic Simulations in Laboratory and Space Plasmas

NERSC PI
Homa Karimabadi

LEAD INSTITUTION
University of California, San Diego

NERSC RESOURCES USED
Hopper, Edison



◀ This visualization of the magnetic field shows the hierarchical development of coherent structures in formation of current sheets and magnetic islands in fully developed turbulence. BURLEN LORING, LAWRENCE BERKELEY NATIONAL LABORATORY

▼ With help from Berkeley Lab visualization experts, astrophysicists found that as the turbulence evolves, it is actually creating waves that launch into the ambient plasma. BURLEN LORING

To make sense of this massive dataset, they tapped Berkeley Lab Visualization Specialist Burlen Loring, who developed custom analysis tools using supercomputers at NERSC. Loring's work allows researchers to study turbulence in unprecedented detail, and the results may hold clues about some of the processes that lead to destructive space-weather events.

The Sun is constantly blasting solar wind into space. As this wind blows across the solar system, it engulfs all planets in its path. Explosive storms on the Sun's surface—solar flares—are another threat. Occasionally, the Sun emits enormous outbursts—coronal mass ejections—that can send up to 10 billion tons of plasma surging into our solar system. These blasts can produce an extreme space-weather event or even trigger geomagnetic storms.

Our only protection is the Earth's magnetic field, which creates a "bubble" (the magnetosphere) that deflects the solar wind. Although most charged particles don't easily cross the Earth's magnetic field lines, some do penetrate this shield by a process called magnetic reconnection. This occurs when a magnetic field embedded in solar wind points in the opposite direction of Earth's magnetic field. The opposing field lines connect in an explosive process that allows charged particles to reach Earth.

Today, scientists use a number of tools to measure solar activities, such as sending satellites into space to take measurements in the solar wind. "One would expect plasma to cool off as it leaves the Sun and travels across the solar system, but satellite observations show that this is not the case—something heats it up along the way," Karimabadi explained.

Because satellites also reveal significant amounts of turbulence in solar wind, many scientists suspect that a "cascade of turbulence" generates the heat. But until Karimabadi's simulations and Loring's visualizations, it was not possible to identify the dissipation physics down to the smallest electron scales.

"These simulations opened up a whole new area of physics for us," Karimabadi said. "For the first time, we got a glimpse of multi-scale physics in the solar wind and could watch the cascade of turbulence. Now everyone in the field is recognizing the benefit of this kind of simulation."

DOE PROGRAM OFFICE

FES—Fusion Base Program

FULL STORY

http://www.nersc.gov/news-publications/news/science-news/2013/supercomputers-capture-turbulence-in-the-solar-wind/

PUBLICATION

Homa Karimabadi, Vadim S. Roytershteyn, Minping Wan, William H. Matthaeus, William Daughton, Pin Wu, Michael Shay, Burlen Loring, Joseph E. Borovsky, Ersilia Leonardis, Sandra C. Chapman, Takuma Nakamura, "Coherent Structures, Intermittent Turbulence and Dissipation in High-Temperature Plasmas," Physics of Plasmas, January 16 (2013), doi: 10.1063/1.4773205.

With Nanoparticles, Slower May Be Better

Molecular dynamics simulations at NERSC provide unprecedented understanding of nanoparticle structure and symmetry

"These are very large simulations —in our research, one set of simulations used 10 million hours alone and some ran for almost a year."

THE LAST DECADE HAS SEEN A FLURRY of research and development involving nanoparticles—chemicals or objects in the 1-100 nanometer range. These tiny structures are of great scientific interest because they are effectively a bridge between bulk materials and atomic and molecular structures.

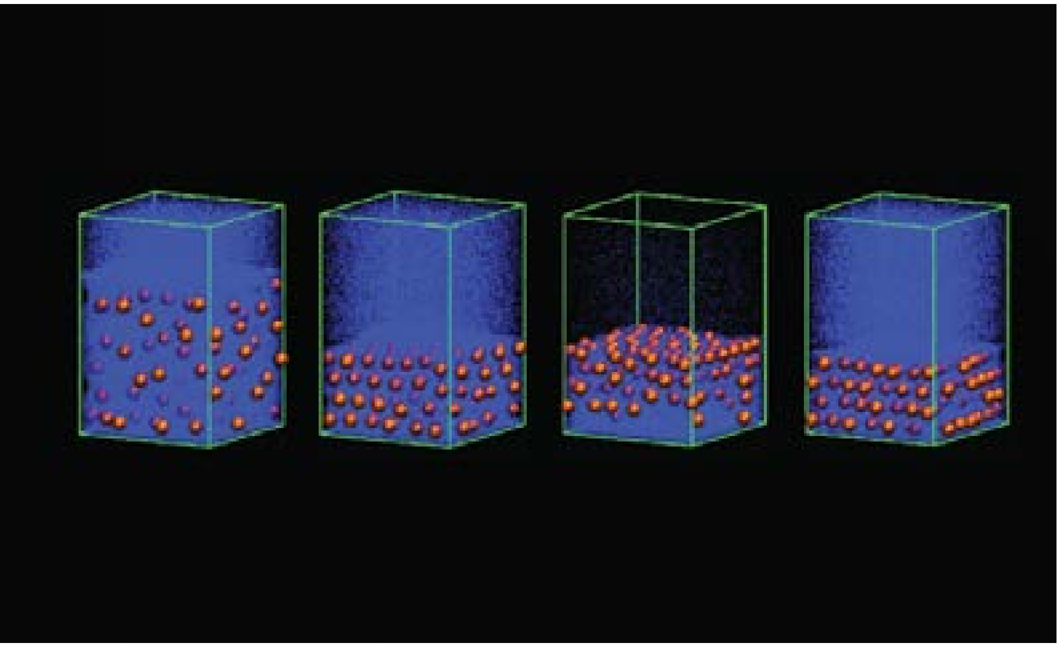
And thanks to their promise of tunability, nanoparticle-based composites are also of great commercial interest for applications ranging from medicine and manufacturing to energy and electronics. The enhanced electro-optical responses of metallic particles, for example, could result in the development of new photonic and plasmonic sensing devices, while magnetic nanoparticles have demonstrated potential for fast memory and storage devices.

Although the concept of using nanoparticles for such applications is well accepted, the ability to control and manipulate their behavior remains a challenge. Experimental studies have shown that nanoparticles tend to randomly aggregate into clusters or migrate to interfacial regions, resulting in the loss of tunability. So researchers are working on methods that will allow them to control how different kinds of nanoparticles assemble and disperse.

"We are mostly interested in how to assemble them into organized arrays," said Gary Grest, a computational physicist at Sandia National Laboratories who is using NERSC supercomputers to run simulations that help explain nanoparticle behavior. "How do you control the interactions, the lattices and structures they build? How do you keep them dispersed? If you want to make, say, a photonics crystal, you want to get them arranged into a nice 2D or 3D array. You want a nice ordered grain crystal structure with no defects."

Over the last two years he and his colleagues at Sandia have performed a series of simulations at NERSC using LAMMPS (Large-scale Atomic/Molecular Massively Parallel Simulator), a molecular dynamics code that is widely used on large-scale parallel machines such as NERSC's Hopper and Edison systems. In 2014, the third year of their

PROJECT TITLE	NERSC PI	LEAD INSTITUTION	NERSC RESOURCES USED
Controlling Nanoparticle Assembly to Engineer New Materials	Gary Grest	Sandia National Laboratories	Edison, Hopper



◀ These visualizations show nanoparticles suspended in a solution before evaporation of the solvent (far left) and how they self-assemble at various stages of solvent evaporation. SHENGFENG CHEN, GARY GREST, SANDIA NATIONAL LABORATORIES

three-year ASCR Leadership Computing Challenge grant, they will continue this research to further explore ways to control the symmetry of nanoparticle assemblies.

As part of this effort, Grest and Shengfeng Cheng, formerly a post-doctoral researcher at Sandia and now an assistant professor of physics at Virginia Tech, used millions of computing hours at NERSC to conduct the first large-scale molecular dynamics simulations of the evaporation-induced assembly of nanoparticles. Their findings were published in The Journal of Chemical Physics (JCP).

Evaporation-induced assembly of nanoparticles —also known as solvent annealing—is commonly used to organize nanoparticles into structures such as nanoclusters, rings, wires, stripes, films and superlattices, Grest explained. In this process, the nanoparticles are suspended in a solvent, the suspended solution is spread on a surface and the solvent is then evaporated.

But few studies have looked at how the assembled structures are affected by the evaporation process, primarily because of the huge computing resources required to run numerical simulations of solvent effects on nanoparticles, Grest said.

"These are very large simulations—in our research, one set of simulations used 10 million hours alone and some ran for almost a year," he said. "The issue is that you've got a nanoparticle that is only 5-10 nanometers in size, but even if you only simulate a few, you still need a large amount of solvent around them—hundreds of thousands to millions of atoms that you are following, and these are big particles that don't move that fast. So NERSC's facilities have been essential for these studies, allowing us to follow lots of particles for long periods of time."

In the JCP study, the LAMMPS simulations revealed that the strength of the polymer/nanoparticle interaction controls the position of the nanoparticles at the interface after the solvent is removed. In particular, the researchers found that the slower the evaporation rate, the better the quality of the nanoparticle crystal.

"We found that by controlling the evaporation rate we could control how the nanoparticles organize and control the quality of the crystal structure in the arrays," Grest said. "If you evaporate the solvent slowly enough, you can organize the nanoparticles nicely."

Defining the role of specific interactions such as this in nanoparticles will help researchers design nanocomposites that can be integrated into large-scale devices without compromising the mechanical and optical features offered at the nanoscale level, he added.

"These computational methods provide an unprecedented understanding of the structure and symmetry of individual nanoparticles," Grest said.

DOE PROGRAM OFFICE
BES Materials Sciences

FULL STORY
https://www.nersc.gov/news-publications/news/science-news/2013/with-nanoparticles-slower-may-be-better/

PUBLICATION
Shengfeng Cheng, Gary S. Grest, "Molecular Dynamics Simulations of Evaporation-Induced Nanoparticle Assembly," The Journal of Chemical Physics, February 8, 2013, 138(064701), doi: 10.1064/1.4789807.

Calculations Confirm 150 Years of Global Land Warming

Land surface air temperatures estimated using NERSC resources match existing temperature measurements from 1901 to 2010

"If, for some reason, you didn't believe global warming was happening, this confirms that global warming really has been occurring since the early 20th century."

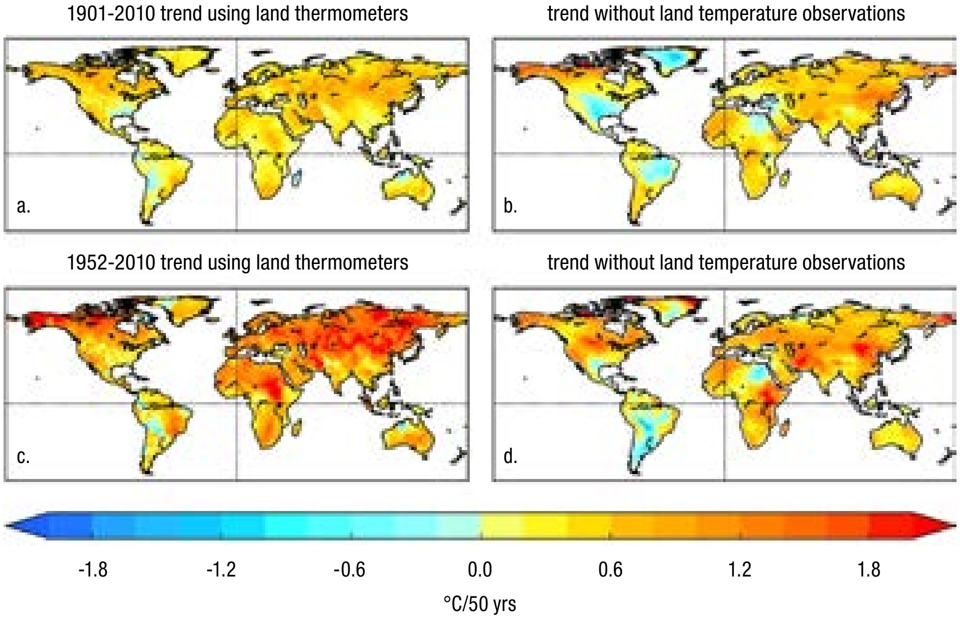
IN MAKING THE CASE FOR GLOBAL CLIMATE WARMING, researchers rely on long-term measurements of air temperature taken around the world, all at a height of two meters or six-and-a-half feet. However, some climate researchers argue that these same measurements are an inconsistent record and unreliable indicators of actual temperature variability and long-term trends.

But a team of scientists led by Gil Compo of the University of Colorado and the National Oceanic and Atmospheric Administration's Earth System Research Laboratory used the Hopper supercomputer at NERSC to demonstrate that land surface air temperatures estimated using a number of other historical observations, including barometric pressure, largely match the existing temperature measurements spanning the years 1901 to 2010.

This work was part of the ongoing 20th Century Reanalysis project, which uses barometric pressure and other climate-related data from the past 150 years to reconstruct the world's climate from its weather using computer models. By plugging this data into climate prediction models, Compo and his team have successfully "predicted" how the earth's weather and climate have varied since the late 1800s. The goal is to demonstrate the accuracy of these models for predicting future warming patterns.

When looking at the land surface air temperatures, critics of the idea that human activity is a factor in global warming note that the land surrounding the measuring stations has often changed significantly since the early 20th century. Agriculture has replaced forests and prairies, farmland has been developed and small towns have grown into cities. Compo and his team also point out that stations have moved and that the

PROJECT TITLE	NERSC PI	LEAD INSTITUTION	NERSC RESOURCES USED
20th Century Reanalysis Project	Gil Compo	University of Colorado, Boulder and National Atmospheric and Oceanic Administration	Hopper



◀ These maps show the changes in air temperatures over land as measured using thermometers (left side) and as calculated by the 20th Century Reanalysis project (left side). The results show interesting differences in some regions, such as the midwestern United States, Argentina and eastern Brazil. 20TH CENTURY REANALYSIS PROJECT

instruments and techniques for measuring temperature have changed. Combined, these factors can undermine confidence in estimates of past, present and future climate change.

"The perception is that there is something wrong with the instruments and that's why there appears to be warming," he said.

To perform their check of the accuracy of the land surface air temperatures, Compo and his team realized they could get most of the answers from their data on the monthly sea surface temperatures and monthly sea ice distribution—two key factors affecting climate over land areas. This yielded seasonal, annual and decadal patterns. By adding in barometric pressure data, they were able to reconstruct even the day-to-day and month-to-month variations.

"This is really the essence of science," Compo said. "There is knowledge 'A' from one source and you think, 'Can I get to that same knowledge from a completely different source?' Since we had already produced the dataset, we looked at just how close our temperatures estimated using barometers were to the temperatures using thermometers."

Although other researchers have used the team's dataset representing climate conditions from 1979 to the present to look at temperature, Compo said this research is unique as they used the reanalysis data stretching back to the early 20th century. The results, he said, are clear.

"If, for some reason, you didn't believe global warming was happening, this confirms that global warming really has been occurring since the early 20th century," Compo said.

DOE PROGRAM OFFICE
Biological and Environmental Research (BER)

FULL STORY
https://www.nersc.gov/news-publications/news/science-news/2013/nersc-calculations-provide-independent-confirmation-of-global-land-warming-since-1901/

PUBLICATION
Gilbert P. Compo, Prashant D. Sardeshmukh, Jeffrey S. Whitaker, Philip Brohan, Philip D. Jones and Chesley McColl, "Independent Confirmation of Global Land Warming without the use of Station Temperatures," Geophysical Research Letters 40(12), 3170–3174 (2013), doi: 10.1002/grl.50425

Physicists ID New Molecules with Unique Features

Hollow magnetic cage molecules may have applications in technology, healthcare

"The potential benefit of this finding is that it provides a route to the synthesis of molecular magnets with colossal magnetic moments."

NERSC SUPERCOMPUTING RESOURCES HELPED Virginia Commonwealth University (VCU) researchers determine it may be possible to create large, hollow magnetic cage molecules that could be used in medicine as a drug delivery system to noninvasively treat tumors and in other emerging technologies.

About 25 years ago, scientists discovered the C60 fullerene—better known as the Buckminster Fullerene or buckyball—a molecule composed of 60 carbon atoms that formed a hollow cage. Due to its unique structure the molecule offers serious technological potential because it could hold other atoms or small molecules inside.

That potential has since spurred worldwide interest among scientists who have been searching for similar molecules. Although some hollow cage structures have been found, none are magnetic. Magnetic properties of the structure are of particular interest because the cage carrying an embedded atom or molecule can be guided by an external magnetic field.

In a study that appeared as the cover story in the July 28, 2013, The Journal of Chemical Physics, two VCU scientists showed that magnetic hollow cages larger than the original C60 fullerene that carry giant magnetic moments are possible. A magnetic moment refers to the measure of the magnetic strength of a cluster.

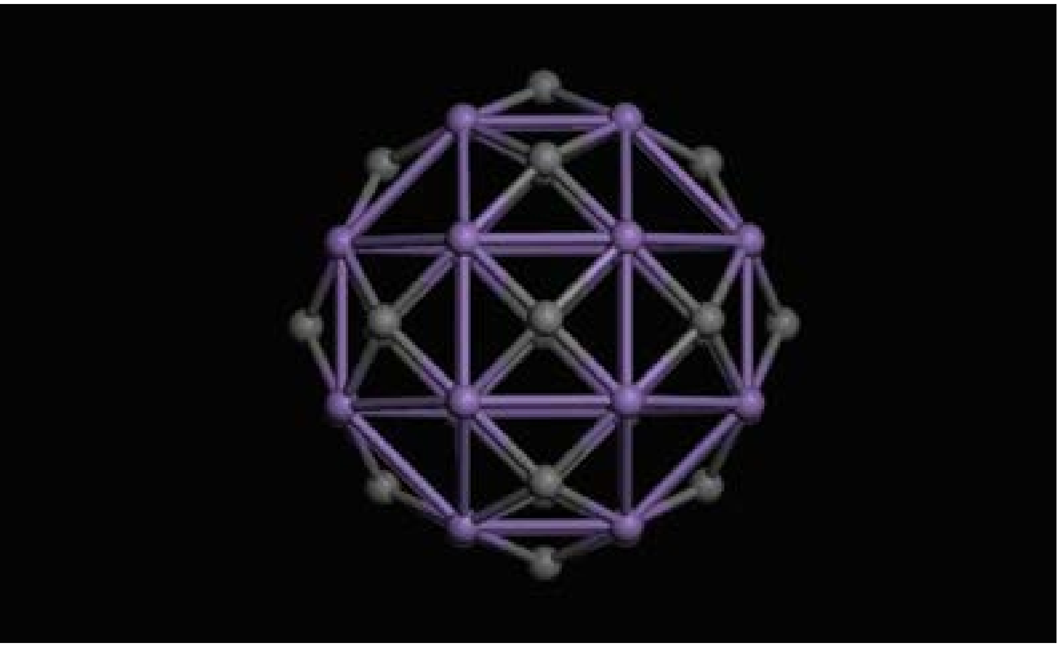
"The potential benefit of this finding is that it provides a route to the synthesis of molecular magnets with colossal magnetic moments," said co-lead investigator Puru Jena, distinguished professor of physics in

PROJECT TITLE
Cluster & Nanostructure for Energy & Bio Applications

NERSC PI
Puru Jena

LEAD INSTITUTION
Virginia Commonwealth University

NERSC RESOURCES USED
Carver



◀ Illustration of a Mn24C18 cluster carrying a magnetic moment of 70 Bohr magnetons, which was featured on the July 28, 2013 cover of The Journal of Chemical Physics. MENGHAO WU, VIRGINIA COMMONWEALTH UNIVERSITY

the VCU College of Humanities and Sciences. Jena collaborated with Menghao Wu, co-author of the paper and a postdoctoral scholar in the VCU Department of Physics.

Using density functional theory and NERSC's Carver IBM iDataPlex system, the researchers theorized that metal-based hollow cages can be formed and, depending on their composition, some of them can even carry very large magnetic moments. "The space inside some of these hollow cage clusters is large enough to accommodate additional atoms or molecules, raising the possibility that they may have technological applications," they wrote.

"These molecules can be used for targeted non-invasive drug delivery. When assembled, the molecules can also form new high strength magnets for device application," Jena said.

However, the team is still early in its discovery process, he emphasized.

"There is a long way to go. Experiments first have to be carried out to prove the predictions of our theory," said Jena. "Ways must be found to synthesize large quantities of these molecules and study their magnetic properties once they are assembled. Finally, these molecules need to be functionalized by embedding desired atoms/molecules for practical applications."

In 2013, NERSC's computational support enabled the VCU research team to publish 13 papers in major international journals, he noted.

DOE PROGRAM OFFICE
BES Materials Science

FULL STORY
https://www.nersc.gov/news-publications/news/science-news/2013/hollow-magnetic-cage-molecules-could-aid-drug-delivery/

PUBLICATION
Menghao Wu, Puru Jena, "Magnetic hollow cages with colossal moments," The Journal of Chemical Physics, July 22, 2013, 139(044301), doi: 10.1063/1.4813022

Why Foams Foam and Bubbles Pop

Mathematicians use supercomputers to illuminate the all-too-brief lives of foamy bubbles

"While different problems are going to require different physics, chemistry and models, this sort of approach has applications to a wide range of problems."

RESEARCHERS FROM BERKELEY LAB and the University of California, Berkeley used the Hopper supercomputer at NERSC to create a mesmerizing computer-generated visualization of the slow and sedate disappearance of wobbly foams, one burst bubble at a time.

James Sethian, who leads the Mathematics Group at Berkeley Lab and is a professor of mathematics at UC Berkeley, and Robert Saye, who earned his Ph.D. in applied mathematics at UC Berkeley, first described mathematically the successive stages in the complex evolution and disappearance of foamy bubbles. They then used these equations to model the foams and their processes.

Their findings, which were published in the journal Science, could help in modeling industrial processes in which liquids mix or in the formation of solid foams such as those used to cushion bicycle helmets.

"This work has application in the mixing of foams, in industrial processes for making metal and plastic foams and in modeling growing cell clusters," said Sethian. "These techniques, which rely on solving a set of linked partial differential equations, can be used to track the motion of a large number of interfaces connected together, where the physics and chemistry determine the surface dynamics."

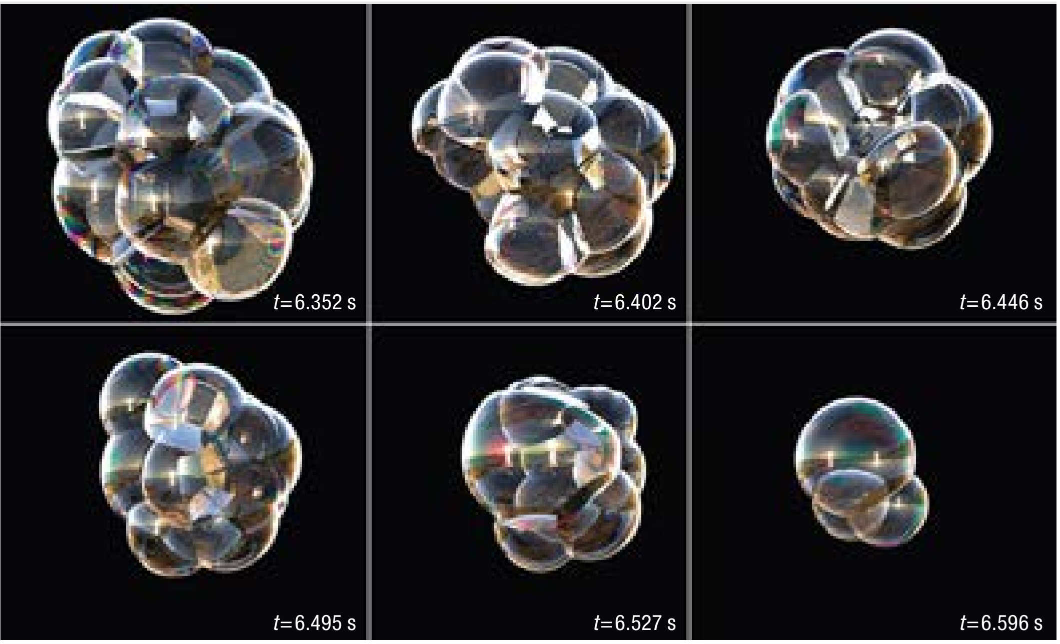
The problem with describing foams mathematically has been that the evolution of a bubble cluster a few inches across depends on what's happening in the extremely thin walls of each bubble, which are thinner than a human hair.

PROJECT TITLE
Numerical Studies in Complex Multiple Interface Phenomena

NERSC PI
James Sethia
Robert Saye

LEAD INSTITUTION
Lawrence Berkeley National Laboratory/UC Berkeley

NERSC RESOURCES USED
Hopper



◀ Mathematicians used a set of equations to solve the physics of a sunset reflected in a cluster of 27 bubbles. They found that thin-film interference in the bubble membranes can create rainbow hues like an oil slick on wet pavement. JAMES SETHIAN, ROBERT SAYE, LAWRENCE BERKELEY NATIONAL LABORATORY, UC BERKELEY

"Modeling the vastly different scales in a foam is a challenge, since it is computationally impractical to consider only the smallest space and time scales," Saye said. "Instead, we developed a scale-separated approach that identifies the important physics taking place in each of the distinct scales, which are then coupled together in a consistent manner."

Saye and Sethian discovered a way to treat different aspects of the foam with different sets of equations that worked for clusters of hundreds of bubbles. One set of equations described the gravitational draining of liquid from the bubble walls, which thin out until they rupture. Another set of equations dealt with the flow of liquid inside the junctions between the membranes. A third handled the wobbly rearrangement of bubbles after one pops. Using a fourth set, the mathematicians solved the physics of a sunset reflected in the bubbles, taking into account thin film interference within the bubble membranes, which can create rainbow hues like an oil slick on wet pavement.

They then used Hopper to solve the full set of equations of motion.

"Solving the full set of equations on a desktop computer would be time-consuming. Instead, we used massively parallel computers at NERSC and computed our results in a matter of days," said Saye.

"Foams were a good test that all the equations coupled together," Sethian said. "While different problems are going to require different physics, chemistry and models, this sort of approach has applications to a wide range of problems."

DOE PROGRAM OFFICE
ASCR

FULL STORY
https://www.nersc.gov/news-publications/news/science-news/2013/math-of- popping-bubbles-in-a-foam/ *(includes video)*

PUBLICATION
Robert Saye and James Sethian, "Multiscale Modeling of Membrane Rearrangement, Drainage, and Rupture in Evolving Foams," Science 340(6133), 720-724, doi: 10.1126/science.1230623

Reading the Cosmic Writing on the Wall

NERSC supercomputers key to Planck's revision of universal recipe

"We've reduced the time it takes to run these calculations from an impossible 1,000 years down to a few weeks."

THANKS TO A SUPERSENSITIVE SPACE TELESCOPE and some sophisticated supercomputing, scientists from the international Planck collaboration have made the closest reading yet of the most ancient story in our universe: the cosmic microwave background (CMB).

In March 2013, the team released preliminary results based on the Planck observatory's first 15 months of data. Using supercomputers at NERSC, Planck scientists have created the most detailed and accurate maps yet of the relic radiation from the Big Bang. They reveal that the universe is about 100 million years older than we thought, with more matter and less dark energy and a slower expansion rate.

"These maps are proving to be a goldmine containing stunning confirmations and new puzzles," said Martin White, a Planck scientist and physicist with University of California, Berkeley and Berkeley Lab. "This data will form the cornerstone of our cosmological model for decades to come and spur new directions in research."

Written in light shortly after the Big Bang, the CMB is a faint glow that permeates the cosmos. Studying it can help us understand how our universe was born and its nature, composition and eventual fate. However, CMB surveys are complex and subtle undertakings. Even with the most sophisticated detectors, scientists still need supercomputing to sift the CMB's faint signal out of a noisy universe and decode its meaning. Julian Borrill, a Planck collaborator and cosmologist in Berkeley Lab's Computational Research Division and co-founder of the Computational Cosmology Center (C3) at the lab, has been developing supercomputing tools for CMB experiments for over a decade.

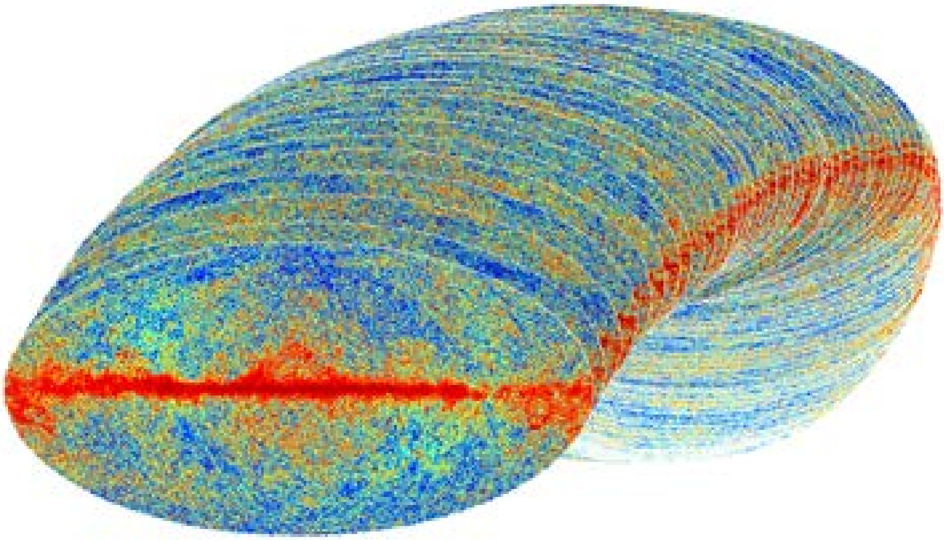
Parked in an artificial orbit about 800,000 miles away from Earth, Planck's 72 detectors complete a full scan of

PROJECT TITLE
Cosmic Microwave Background Data Analysis for the Planck Satellite Mission

NERSC PI
Julian Borrill

LEAD INSTITUTION
Lawrence Berkeley National Laboratory

NERSC RESOURCES USED
Hopper, Carver



◀ Using supercomputers at NERSC and the Planck observatory's first 15 months of data, scientists with the international Planck collaboration have created the most detailed and accurate maps yet of relic radiation from the big bang. PLANCK COLLABORATION

the sky once every six months or so. Observing at nine different frequencies, Planck gathers about 10,000 samples every second, or a trillion samples in all for the 15 months of data included in this first release. In fact, Planck generates so much data that, unlike earlier CMB experiments, it's impossible to analyze exactly, even with NERSC's resources.

Instead, CMB scientists employ clever workarounds. Using approximate methods they are able to handle the Planck data volume, but then they need to understand the uncertainties and biases their approximations have left in the results.

One particularly challenging bias comes from the instrument itself. The position and orientation of the observatory in its orbit, the shapes and sizes of detectors and even the overlap in Planck's scanning pattern affect the data.

To account for such biases and uncertainties, researchers generate a thousand simulated copies of the Planck data and apply the same analysis to these. Measuring how the approximations affect this simulated data allows the Planck team to account for their impact on the real data.

With each generation of NERSC supercomputers, the Planck team has adapted its software to run on more and more processors, pushing the limits of successive systems while reducing the time it takes to run a greater number of complex calculations.

"By scaling up to tens of thousands of processors, we've reduced the time it takes to run these calculations from an impossible 1,000 years down to a few weeks," said Ted Kisner, a C3 member at Berkeley Lab and Planck scientist.

DOE PROGRAM OFFICE

HEP—Cosmic Frontier

FULL STORY

https://www.nersc.gov/news-publications/news/science-news/2013/planck- results/

The Life and Times of Massive Stars

An automated supernova hunt is shedding new light on the death sequence of massive stars

> **"This discovery shows that there is still a lot that we can learn from the archival data at NERSC."**

DIGGING THROUGH THE PALOMAR TRANSIENT FACTORY (PTF) data archive housed at NERSC, astronomers have found the first causal evidence that massive stars shed huge amounts of material in a "penultimate outburst" before final detonation as supernovae.

A focused search for Type IIn SN precursor bursts, conducted by Eran Ofek of Israel's Weizmann Institute and the PTF team, led to this finding. PTF is an international collaboration that brings together researchers at universities, observatories and Berkeley Lab to hunt for supernovae and other astronomical objects that appear for awhile, then disappear.

Massive stars—somewhere between eight and 100 times the mass of our Sun—spend much of their lives fusing hydrogen into increasingly heavier elements, like helium, oxygen, carbon and so on. In the end, there is almost nothing left but an iron core. Eventually, that core collapses, releasing a tremendous amount of energy as neutrinos, magnetic fields and shock waves and destroying the star in the process. From Earth, this explosive event is observed as a supernova. If astronomers detect hydrogen, the event is classified as a Type II supernova. And if the hydrogen-emission line is narrow, the event is classified as a Type IIn (for "narrow").

In the case of Type IIn events, scientists suspected that the narrow emission line that occurs as light from the event passes through a thin sphere of hydrogen that was already surrounding the star before it went supernova. Some believed that the dying star might have shed this shell of material before it self-destructed, but until recently there was no evidence to link such an outburst to an actual supernova.

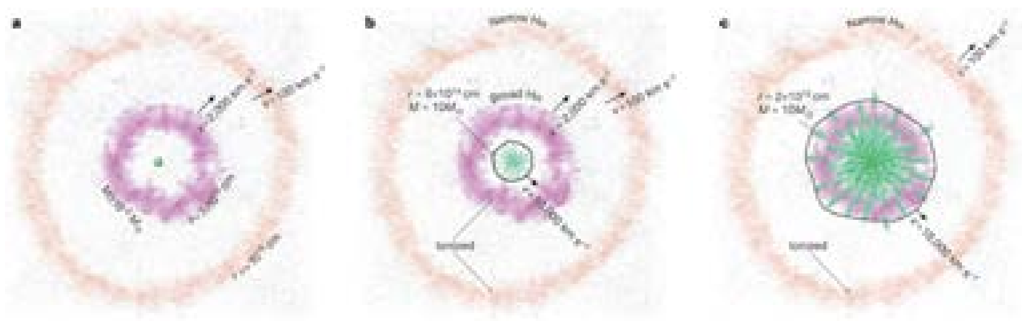
That's where PTF comes in. For almost four years the PTF team has relied on a robotic telescope mounted

PROJECT TITLE
Palomar Transient Factory

NERSC PI
Peter Nugent

LEAD INSTITUTION
Lawrence Berkeley National Laboratory

NERSC RESOURCES USED
Archives, Carver, Data transfer nodes, Science Gateway nodes



◀ A qualitative sketch of a proposed model for SN2010mc. Panel A shows the supernova on the day it detonated. An inner shell (purple) represents the material ejected by the precursor star about one month earlier from the penultimate outburst. An outer shell (orange) is made up of material ejected by the precursor star prior to the penultimate burst. Panel B shows SN2010mc at day five. The supernova shock front (grey line) is moving at 10,000 kilometers per second, ionizing the inner and outer shells along the way, producing the broad and narrow hydrogen emission lines that astronomers on Earth detect. Panel C shows the object at day 20, when the supernova shock engulfs the inner shell. At this point, astronomers only detect a narrow hydrogen emission line. E. O. OFEK, WEIZMANN INSTITUTE OF SCIENCE

on the Palomar Observatory's Samuel Oschin Telescope in Southern California to scan the sky nightly. As soon as observations were taken, the data traveled to NERSC via ESnet, where computers running Real-Time Transient Detection Pipeline software screened the data and identified events for astronomers to follow up on. NERSC also archived this data and allowed collaborators to access it through DeepSky, a web-based science gateway.

On August 25, 2010 the PTF pipeline detected a Type IIn supernova half a billion light years away in the constellation Hercules. Shortly afterward, Ofek led a search of previous PTF scans of the same stellar neighborhood and found the supernova's likely precursor, a massive variable star that had shed a huge amount of mass only 40 days before the supernova was detected. They labeled the event SN 2010mc.

Ofek and the PTF team developed a scenario and tested it against competing theoretical ideas, using evidence from several sky surveys that were triggered to observe SN 2010mc once it was detected by the NERSC pipeline. They concluded that the "penultimate outburst" had blown off a hundredth of a solar mass in a shell expanding 2,000 kilometers per second, already 7 billion kilometers away from the supernova when it exploded. Earlier ejecta were detected 10 billion kilometers away, having slowed to a hundred kilometers per second.

After the supernova explosion, high-velocity ejecta passing through shells of earlier debris left a record of varying brightness and spectral features. The observations pointed to the most likely theoretical model of what happened: turbulence-excited solar waves drove successive episodes of mass loss, finally culminating in the collapse and explosion of the core. Because the stellar outburst occurred very shortly before the supernova, the astronomers suspected that the events were causally linked.

Once the team found SN 2010mc's precursor, they sifted through stellar neighborhoods in the PTF archival data where other Type IIn supernovae had previously been detected.

"Although the PTF project is no longer collecting data every night, we are still relying on NERSC resources to sift through our archival data," said Peter Nugent, a Berkeley Lab senior staff scientist and member of the PTF collaboration. "This discovery shows us that there is still a lot that we can learn from the archival data at NERSC, and gives us insights into how we may design future experiments to further investigate these events."

DOE PROGRAM OFFICE

High Energy Physics

FULL STORY

https://www.nersc.gov/news-publications/news/science-news/2013/a-massive- stellar-burst-before-the-supernova/

PUBLICATION

Eran Ofek, Mark Sullivan, Brad Cenko, Mansi M. Kasliwal, Avishay Gal- Yam, Shrinivas Kulkarni, Iair Arcavi, Lars Bildsten, J.S. Bloom, Assaf Horesh, Andy Howell, Alexi Filippenko, Russ Laher, D. Murray, Ehud Nakar, Peter E. Nugent, Jeffrey M. Silverman, Nir J. Shaviv, Jason Surace, Ofer Yaron, "An outburst from a massive star 40 days before a supernova explosion," Nature 494, 65–67 (2013), doi:10.1038/ nature11877

Thinnest Solar Cells Ever?

2D materials could yield thinner, more lightweight solar panels with much higher power densities

"The only way we could get these algorithms to run was at NERSC."

EFFORTS TO IMPROVE SOLAR CELLS have historically focused on improving energy conversion efficiencies and lowering manufacturing costs. But computer simulations conducted at NERSC have shown how using a different type of material could yield thinner, more lightweight solar panels that provide power densities—watts per kilogram of material—orders of magnitude higher than current technologies.

Solar cells are most often made from multi-crystalline silicon semiconductors, which have energy conversion efficiencies in the 15 to 20 percent range. But silicon-based photovoltaics, typically hundreds of microns thick, do not absorb light very efficiently and must be mounted between layers of glass, which can be heavy and expensive. In fact, about half the cost of today's panels is in support structures, installation, wiring and control systems.

Now, using NERSC's Hopper and Carver systems to run density functional theory calculations, scientists have demonstrated that two-dimensional materials—specifically, transition metal dichalcogenides (TMDs) and grapheme—could produce 1-nanometer thick solar cells that yield 1 to 2 percent energy conversion efficiency.

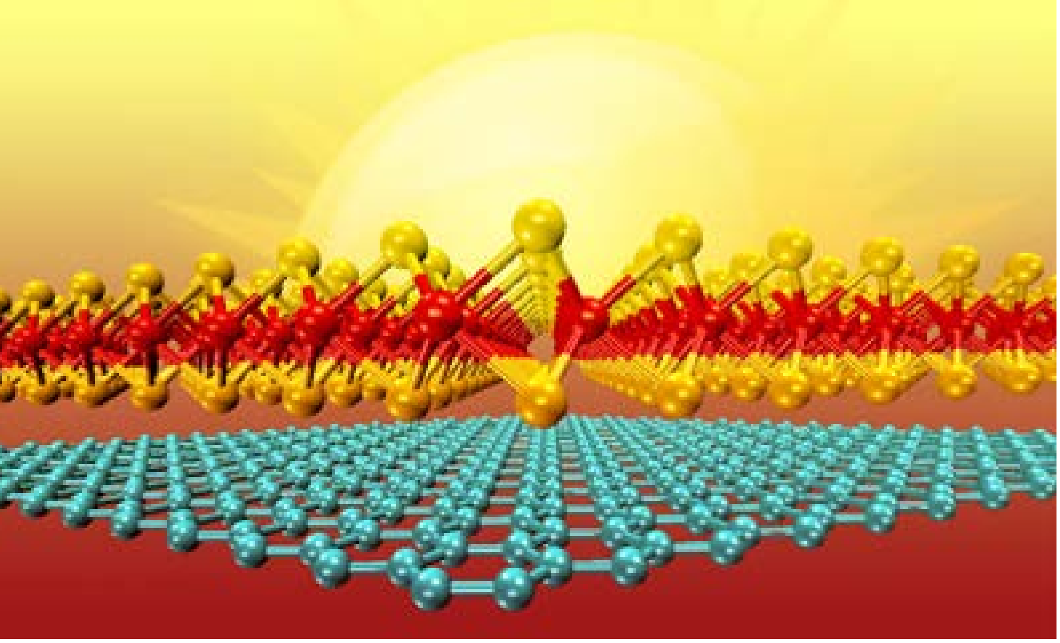
While that conversion rate may not seem impressive at first glance, it is achieved using material that is 20 to 50 times thinner than the thinnest solar cell made today, according to Jeffrey Grossman, an engineering professor at the Massachusetts Institute of Technology (MIT) and co-author on this research. Because the cells would be so thin, several could be stacked to improve the energy efficiency while still being much smaller and lighter than existing solar cells.

PROJECT TITLE
Quantum Simulations of Nanoscale Energy Conversion

NERSC PI
Jeffrey Grossman

LEAD INSTITUTION
MIT

NERSC RESOURCES USED
Hopper, Carver



◀ Using NERSC supercomputers and density functional theory calculations, researchers demonstrated that an effective solar cell could be made from a stack of two 1-molecule-thick materials: graphene (a one-atom-thick sheet of carbon atoms, shown in blue) and molybdenum disulfide (with molybdenum atoms shown in red and sulfur in yellow). JEFFREY GROSSMAN, MARCO BERNARDI, MIT

"Our calculations show that, theoretically at least, one can stack 2D materials to make something a little thicker and get up to 20 percent efficiency—the same as current cutting-edge photovoltaics," he said. In addition, the 2D materials—which in this case were molybdenum disulfide and molybdenum diselenide—have very good thermal and UV robustness, he noted.

They also have keen light-absorption capabilities. Solar cells using these monolayers as active layers could show significantly higher power density than existing ultra-thin solar cells because TMDs can absorb sunlight effectively despite being extremely thin.

"The question we wanted to answer was how well do these materials absorb light," Grossman said. "And that is something that surprised us: how with just two layers of material it is possible to absorb 10 percent of the solar spectrum and have a cell that is 2 percent efficient as a result."

NERSC supercomputers were critical in enabling these findings, which open up new avenues for nanoscale solar energy conversion, Grossman added.

"In order to model accurately enough the optical properties of these two materials, we couldn't use traditional methods. There were too many electrons and we needed too much accuracy," he said. "So basically the only place we could get these algorithms to run was at NERSC."

DOE PROGRAM OFFICE

NP—Nuclear Theory

FULL STORY

https://www.nersc.gov/news-publications/news/science-news/2013/2d-monolayers-could-yield-thinnest-solar-cells-ever/

PUBLICATION

Marco Bernardi, Maurizia Palummo and Jeffrey C. Grossman, "Extraordinary sunlight absorption and one nanometer thick photovoltaics using two-dimensional monolayer materials," Nano Letters 13(8), 3664–3670 (2013), doi: 0.1021/nl401544y

Taming Plasma Fusion Snakes

Supercomputer simulations help explain a plasma instability mystery

"We had no really good simulations of snakes prior to this."

CONTROLLED NUCLEAR FUSION HAS OFFERED THE PROMISE of a safe, clean, sustainable energy resource for decades. Now, with concerns over global climate change growing, the ability to produce a reliable carbon-free energy source has taken on new urgency.

Fifty years of fusion research has brought many advances, but technical hurdles remain. Since NERSC was founded in 1974, researchers have used the center's supercomputing resources to address these issues and move fusion energy closer to reality.

In a fusion reaction, energy is released when two hydrogen isotopes (tritium and deuterium) are fused together to form a heavier nucleus, helium. To achieve high enough reaction rates to make fusion a useful energy source, the hydrogen gas must be heated to extremely high temperatures—more than 100 million degrees Celsius—which transforms it into a hot gas known as plasma.

Unfortunately, the high energy required to bring the plasma to these temperatures also drives the formation of plasma instabilities that need to be controlled in order to hold the plasma together and allow fusion to take place.

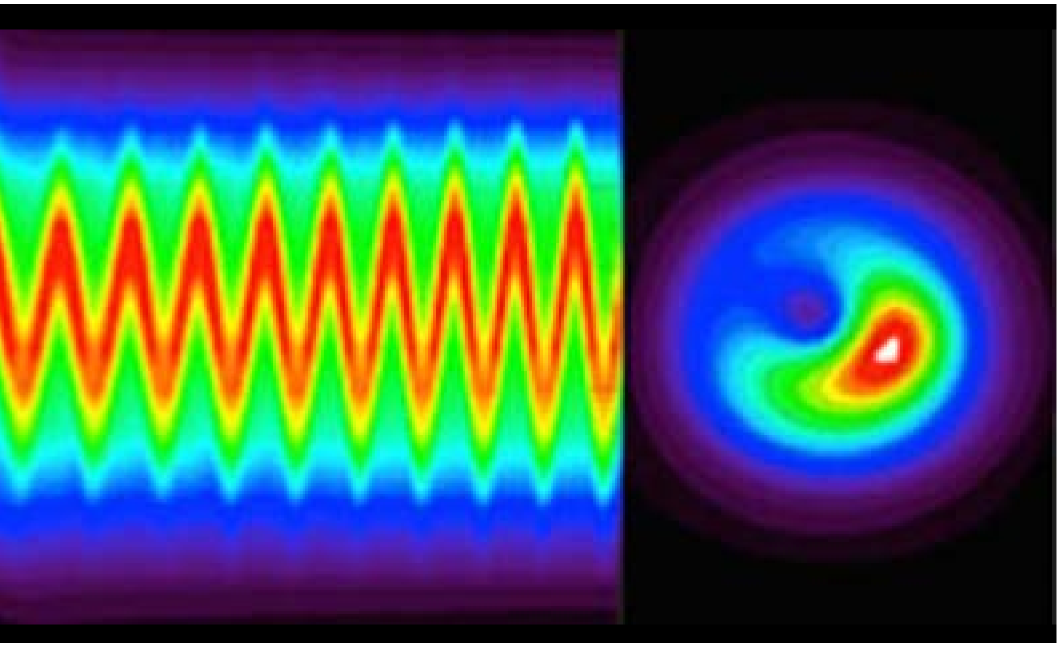
One commonly observed instability is the plasma density snake, named for its corkscrew-shaped appearance. Impurity snakes have been a regular feature in every major tokamak fusion experiment of the last 25 years, according to Linda Sugiyama, a principal research scientist in the Laboratory for Nuclear Science at the Massachusetts Institute of Technology (MIT) and a long-time NERSC user. But only recently have Sugiyama and colleagues been able to explain the formation and stability of these snakes, which could lead to insights into how to control them and help maximize the safety and performance of next-generation reactors

PROJECT TITLE
3D Extended MHD Simulation of Fusion Plasmas

NERSC PI
Stephen Jardin (PPPL)

LEAD INSTITUTION
Princeton Plasma Physics Laboratory

NERSC RESOURCES USED
Edison, Hopper



◀ A measured time sequence of a plasma snake (side and edge views).

▼ 3D nonlinear plasma simulations conducted at NERSC are giving researchers new insights into the formation and stability of plasma snakes.
LINDA SUGIYAMA, MIT

such as ITER, a tokamak facility under construction in France.

A tokamak is a type of fusion reactor considered by many to be the best candidate for producing controlled thermonuclear fusion power. It uses a toroidal (doughnut-shaped) vessel and extremely strong magnetic fields to confine the plasma as it is heated up. Unfortunately, snakes can trap impurities released from the plasma-facing walls of the tokamaks, and these impurities radiate large amounts of energy that cool the plasma and undermine its ability to create a fusion reaction.

"When you try to heat up magnetically confined plasma, the temperature in the center builds up," Sugiyama explained. "While this is good for fusion reaction, if it reaches a certain point it can trigger various 3D magnetohydrodynamic instabilities (such as snakes), which are the fastest growing and most dangerous."

To date, the most commonly studied snakes have been those induced by the injection of deuterium pellets into the plasma. But novel spectroscopic imaging diagnostics have enabled researchers to identify a new type of snake with unprecedented resolution: the impurity snake, which is produced by an accumulation of impurity ions rather than deuterium ions from injected fueling pellets.

Now simulations run on NERSC's Hopper system have confirmed the existence of impurity snakes and yielded new insights into the formation and features of these plasma instabilities.

"Because the mechanism for forming these snakes was so different from the assumptions that had been made previously, we did simulations of the snakes to show that it was possible for them to form and be self-sustaining in a way that was seen in the experiment," Sugiyama explained.

"What made these simulations different was that we did both temperature and density evolution; you really need to know the plasma density separate from the temperature to understand the snakes. We had no really good simulation of snakes prior to this, and this is where the NERSC computers come in."

The codes she and her colleagues ran at NERSC are capable of rapidly running and re-running hundreds or thousands of virtual experiments on plasma behavior in tokamaks in the time it would take to construct just one physical experiment, at a fraction of the cost and with far more detailed information about the results, she emphasized.

"Using these codes, we have already learned much about how to avoid or control various instabilities in order to maximize the performance of a future reactor, and we expect to continue to improve," Sugiyama said.

DOE PROGRAM OFFICE
FES—Fusion SciDAC

FULL STORY
https://www.nersc.gov/news-publications/news/science-news/2014/taming-plasma-fusion-snakes/

PUBLICATION
Linda E. Sugiyama, "On the formation of m = 1, n = 1 density snakes," Physics of Plasmas, March 4, 2013, 20(032504), doi: 10.1063/1.4793450

Getting More from Natural Gas Reserves

New structural information could yield more efficient extraction of gas and oil from shale

"This is a crucial step toward predicting and designing materials with enhanced gas adsorption properties."

SUPERCOMPUTERS AT NERSC HELPED SCIENTISTS at Oak Ridge National Laboratory (ORNL) study gas and oil deposits in shale and collect structural information that could lead to more efficient extraction of these resources from shale.

It could also enable environmentally benign and efficient energy production from coal and viable CO_2 sequestration technologies, according to Yuri Melnichenko, an instrument scientist at ORNL's High Flux Isotope Reactor.

In a paper published in the Journal of Materials Chemistry A, Melnichenko and colleagues from ORNL's Materials Science and Technology Division described a small-angle neutron scattering technique that, combined with electron microscopy and theory, can be used to examine the function of pore sizes.

Using their technique in conjunction with the High Flux Isotope Reactor's General Purpose SANS instrument, the scientists demonstrated that there is significantly higher local structural order in nanoporous carbons than previously believed. This is important because it enables them to develop modeling methods based on local structure of carbon atoms. The research team also probed distribution of adsorbed gas molecules at unprecedented smaller length scales, allowing them to devise models of the pores.

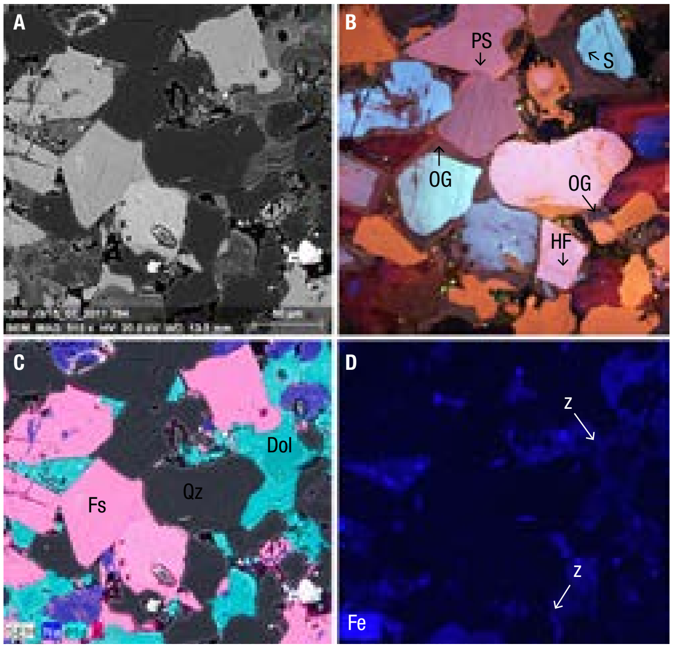
"We have recently developed efficient approaches to predict the effect of pore size on adsorption," said James Morris, co-author on the paper and a member of ORNL's Materials Science and Technology Division. "However, these predictions need verification, and the recent small-angle neutron experiments are ideal for

PROJECT TITLE
Theoretical Studies of Complex Materials

NERSC PI
Valentino Cooper

LEAD INSTITUTION
Oak Ridge National Laboratory

NERSC RESOURCES USED
Edison, Hopper



◀ A scanning electron microscope image illustrating mineralogy and texture of an unconventional gas reservoir. OAK RIDGE NATIONAL LABORATORY

this. The experiments also beg for further calculations, so there is much to be done."

While traditional methods provide general information about adsorption averaged over an entire sample, they do not provide insight into how pores of different sizes contribute to the total adsorption capacity of a material. Using supercomputing resources at NERSC, the ORNL researchers ran large-scale simulations of carbon cooling to produce the nanoporous carbon structures. This approach, in conjunction with previous work, allows scientists to analyze two-dimensional images to understand how local structures can affect the accessibility of shale pores to natural gas.

"Combined with atomic-level calculations, we demonstrated that local defects in the porous structure observed by microscopy provide stronger gas binding and facilitate its condensation into liquid in pores of optimal sub-nanometer size," Melnichenko said. "Our method provides a reliable tool for probing properties of sub- and super-critical fluids in natural and engineered porous materials with different structural properties. This is a crucial step toward predicting and designing materials with enhanced gas adsorption properties."

Together, the application of neutron scattering, electron microscopy and theory can lead to new design concepts for building novel nanoporous materials with properties tailored for the environment and energy storage-related technologies, the researchers noted. These include capture and sequestration of man-made greenhouse gases, hydrogen storage, membrane gas separation, environmental remediation and catalysis.

DOE PROGRAM OFFICE

BES Materials Science

FULL STORY

https://www.nersc.gov/news-publications/news/science-news/2013/nersc-supercomputers-help-reveal-secrets-of-natural-gas-reserves/

PUBLICATION

J. R. Morris, C. I. Contescu, M. F. Chisholm, V. R. Cooper, J. Guo, L. He, Y. Ihm, E. Mamontov, Y. B. Melnichenko, R. J. Olsen, S. Pennycook, M. B. Stone, H. Zhang and N. C. Gallego , "Modern approaches to studying gas adsorption in nanoporous carbons." Journal of Materials Chemistry A, April 21, 2013, 1, doi: 10.1039/C3TA10701A

Policing Air Quality in California

NERSC computations help policy makers in the Golden State address critical environmental issues

"If California's efforts in reducing black carbon can be replicated globally, we can slow down global warming."

TWO STUDIES PUBLISHED IN 2013 as part of a NERSC project titled "Investigation of the Magnitudes and Probabilities of Abrupt Climate TransitionS (IMPACTS)" are helping policy makers address critical environmental issues in California. The studies were funded by the California Air Resources Board (CARB).

Reductions in emissions of black carbon since the late 1980s have resulted in a measurable reduction of concentrations of global warming pollutants in the atmosphere, according to researchers from UC San Diego, UC Berkeley, Berkeley Lab and Pacific Northwest National Laboratory (PNNL).

The study's findings support a growing body of scientific evidence that suggests it is possible to immediately slow the pace of climate change regionally by reducing emissions of short-lived climate pollutants. The major sources of black carbon, or soot, in California are diesel-burning mobile sources, residential wood burning in fireplaces and heaters, agricultural burning and wildfires.

Their three-year study, led by Veerabhadran Ramanathan of the Scripps Institution of Oceanography at UC San Diego, estimated that reductions in black carbon as a result of clean air regulations were equivalent to reducing carbon dioxide emissions in California by 21 million metric tons annually, or taking more than 4 million cars off California roads every year. It is the first comprehensive regional assessment of the climate impact of black carbon on California.

According to co-author Tom Kirchstetter of Berkeley Lab, black carbon levels have decreased by about 90 percent over a 45-year period, beginning with the establishment of CARB in 1967, mostly as a result of state regulations for diesel engine emissions. Researchers found the state's efforts to reduce diesel emissions have lessened the impact of global warming on California.

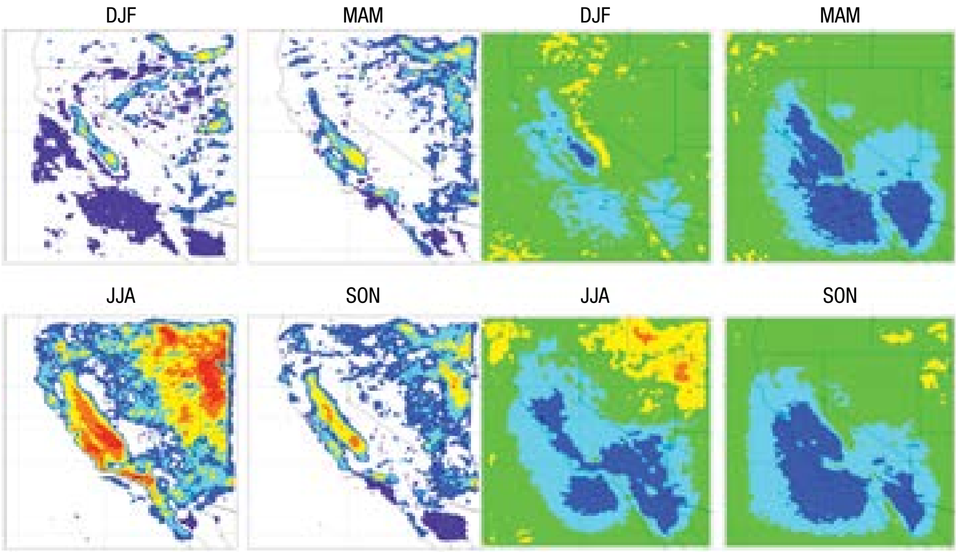
PROJECT TITLE
Investigation of the Magnitudes and Probabilities of Abrupt Climate TransitionS (IMPACTS)

NERSC PI
William Collins

NERSC RESOURCES USED
Hopper

LEAD INSTITUTION
Lawrence Berkeley National Laboratory

DOE PROGRAM OFFICE
BER—Climate and Environmental Sciences



◀ Changes in atmospheric temperature at 2 meters (left) and 2 km (right) comparing the 2000s to the 1960s, in Celsius. CALIFORNIA AIR RESOURCES BOARD

As CARB's current efforts to clean up trucks and buses move forward, resulting in the continued cleanup and turnover of older heavy-duty diesel vehicles, California should continue to see declines in particulate matter emissions. Advanced engine emissions control systems and filters are expected to dramatically reduce emissions from all new diesel engines. Current diesel truck engines, for example, are over 90 percent cleaner than models from years when they were unregulated.

"If California's efforts in reducing black carbon can be replicated globally, we can slow down global warming in the coming decades by about 15 percent," Ramanathan said.

Aerosols Over California

In a related study, scientists from PNNL, Colorado State University and CARB used NERSC's Hopper supercomputer to characterize the roles of various particles as atmospheric change agents on a regional scale. These simulations have, for the first time ever, characterized the relative, direct influence of different aerosol species on seasonal atmospheric warming and cooling over California. The scientists found that aerosols have a net cooling effect on California's atmosphere, but individual species contribute differently. While sulfates contributed the most to cooling, black carbon particles were responsible for up to 95 percent of countervailing warming.

"With a better understanding of how each particle affects the atmosphere, we can assess the success of regional emissions controls of anthropogenic, or human-caused particles," said Chun Zhao, a PNNL climate scientist.

Zhao and his colleagues applied the Weather Research and Forecasting model with coupled chemistry (WRF-Chem) to the atmosphere above California to simulate the amount of warming or cooling caused by each particle species by season, as well as its distribution in the atmosphere.

They found that combined aerosols of all species had an overall cooling effect at the top of the atmosphere through all seasons in California, and that sulfates were the largest contributors to cooling in summer and winter. They also found that dust and carbon-containing aerosols were the largest contributors to warming among aerosol species. The researchers corroborated their model simulations with various datasets of meteorological and aerosol field measurements taken during 2005 and 2008.

"To see whether our emission control policies are making a difference on air quality and regional climate, and to make informed policy recommendations for the future, we need to understand the seasonal variation and movement of aerosols over California," said Zhao.

FULL STORIES

https://www.nersc.gov/news-publications/news/science-news/2013/emission-regulations-reduced-impact-of-climate-change-in-ca/

https://www.nersc.gov/news-publications/news/science-news/2013/researchers-model-impact-of-aerosols-over-california/

PUBLICATIONS

Veerabhadran Ramanathan, Ranjit Bahadur, Kimberly Prather, Odelle Hadley, P.S. Praveen, Alberto Cazorla, Chun Zhao, Tom Kirchstetter, and Ruby Leung, "Black carbon and the regional climate of California," California Air Resources Board report, Contract 08-323, April 15, 2013

Chun Zhao, L. Ruby Leung, Richard C. Easter, Jenny L Hand and Jeremy Avise, "Characterization of speciated aerosol direct radiative forcing over California," Journal of Geophysical Research, doi:10.1029/2012JD018364 (in press

Turning Greenhouse Gases into Gold

Computer simulations reveal reaction mechanism behind CO_2 conversion into carbon-neutral fuels and chemicals

"These findings will be useful in the design and development of ... fuels with a carbon-neutral footprint."

ENVIRONMENTALISTS THE WORLD OVER HAVE LONG LAMENTED the destructive effects of greenhouse gases, with carbon dioxide (CO_2) often accused of being the primary instigator of global climate change. As a result, numerous efforts are under way to find ways to prevent, capture and sequester—perhaps even bury—CO_2 emissions and reduce their negative effects.

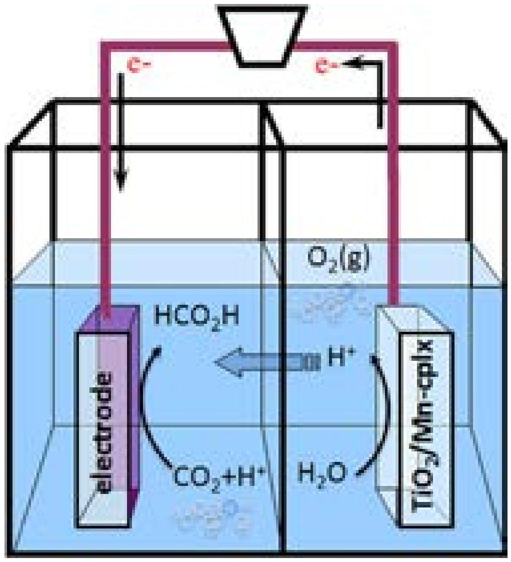
But some researchers say CO_2 is getting a bad rap. They contend that it represents a virtually unlimited energy resource that can be recycled into carbon-neutral fuels and chemicals, reducing both the amount of CO_2 in the Earth's atmosphere and our dependence on fossil fuels.

To achieve this role-reversal requires a process known as conversion reaction—the electro- or photochemical reduction of CO_2. The goal of conversion reaction is to break CO_2's molecular bonds and make it more flexible so it is easier to transform it into something else. This requires exposing the CO_2 to an energy source in combination with a catalyst such as pyridinium.

"CO_2 has two double bonds and is very rigid, so it doesn't want to react," said Victor Batista, a professor of chemistry at Yale University whose research group has been using NERSC computing resources to study novel ways to do this. "So the question is, how do you make it reactive and convert it into something that can be used as a fuel, and do this with as little energy as possible?"

Unfortunately, early conversion reaction efforts required high electrode overpotential (the charge transfer of an electrochemical reaction), leading to very energy-inefficient processes that defeated the purpose of using CO_2 to create an alternative energy source. But in 2008 researchers demonstrated that by submerging a platinum semiconductor into an acidic solution of pyridine and CO_2 and charging it with just 600 millivolts of electricity, the CO_2 could be transformed into formic acid, formaldehyde and methanol.

PROJECT TITLE	NERSC PI	LEAD INSTITUTION	NERSC RESOURCES USED
QM/MM Studies of High-Valent Manganese Complexes	Victor S. Batista (Yale)	Yale	Hopper



◀ Researchers have shown that by submerging a platinum semiconductor into an acidic solution of pyridine and CO_2 and charging it with just 600 millivolts of electricity, the CO_2 could be transformed into formic acid, formaldehyde and methanol.
VICTOR BATISTA, YALE UNIVERSITY

Still, a key question remained: what causes the conversion reaction?

Batista and colleagues from Yale and Brookhaven National Laboratory initially speculated that the pyridine was serving as a catalyst and that the platinum surface assisted in the generation of the pyridine radical. They also thought the reduction of pyridine at low overpotentials might be generating other species involved in the CO_2 reduction.

But through computer simulations involving a free-energy analysis of reaction intermediates, they discovered a simpler explanation: the CO_2 is reduced by platinum-bound hydrogen atoms, not the pyridinium itself. These atoms migrate from the platinum surface to the CO_2 surface via a proton-coupled hydride transfer mechanism activated by the acidic pyridinium ions.

"From our calculations, everything we are observing so far can be explained by the mechanism we have proposed," Batista said.

The calculations included a reaction energy profile, which explores how the CO_2 molecule reacts at the surface—a process that took several weeks of computing on NERSC'S Hopper system, he added.

These findings will be useful in the design and development of new technologies that can generate fuels with a carbon-neutral footprint, turning CO_2 into an alternative feedstock, according to Batista and his co-authors. They are now looking at using different electrocatalysts and metal surfaces through an ongoing CO_2 reduction collaboration with researchers at University of California, San Diego and Emory University.

DOE PROGRAM OFFICE

BES—Chemical Sciences

FULL STORY

https://www.nersc.gov/news-publications/news/science-news/2013/turning-greenhouse-gases-into-gold/

PUBLICATION

Mehmed Z. Ertem, Steven J. Konezny, C. Moyses Araujo and Victor S. Batista, "Functional role of pyridinium aqueous electrochemical reduction of CO2 on Pt(111)," Journal of Physical Chemistry Letters 4(5), 745-748 (2013), doi: 10.1021/jz400183z

Helping a Catalyst Reach its Full Potential

Placing protons in the right spot prevents wasteful reactions in chemical and energy production

> "This work is a prime example of how theory and simulation can help us gain greater insights into chemistry."

CHEMICAL REACTIONS FACILITATED BY CATALYSTS are crucial to many industrial processes. In fertilizer production, chemical companies combine copious amounts of molecular hydrogen with nitrogen to produce ammonia—a process that currently consumes about 1 percent of the world's energy. To produce higher-octane gasoline, petroleum companies will remove hydrogen from hydrocarbon molecules during the refining process. And molecular hydrogen is used to store electrical energy generated by renewable sources like the Sun and wind.

Although many catalysts used in industry work just fine, researchers at the Pacific Northwest National Laboratory (PNNL) want to help them reach their full potential. So they combined supercomputing simulations and laboratory experiments to capture, at an atomic scale, the reactions that occur when a well-known nickel-based electrocatalyst dissolves into a liquid that has been bubbled with hydrogen gas. The catalyst used in this study pulls hydrogen molecules apart.

"The catalyst we studied is the fastest of its type with hydrogen, but it still isn't fast enough to put in a fuel cell and drive down the road," said Wendy Shaw, a biophysical chemist at PNNL. "To get the catalysts to achieve their full potential, we need to understand all of the bottlenecks and how to overcome them."

Using computing resources at NERSC, PNNL and Oak Ridge, she and her colleagues demonstrated that proton delivery or removal determines if the nickel-based catalyst takes a highly productive form or twists into a less useful structure. The scientists found that the most productive catalytic structure has key nitrogen-hydrogen bonds close to the nickel center. In this form, called endo/endo, the reaction occurs in a fraction of a second. If the catalyst is in any other form, the reaction takes days to complete.

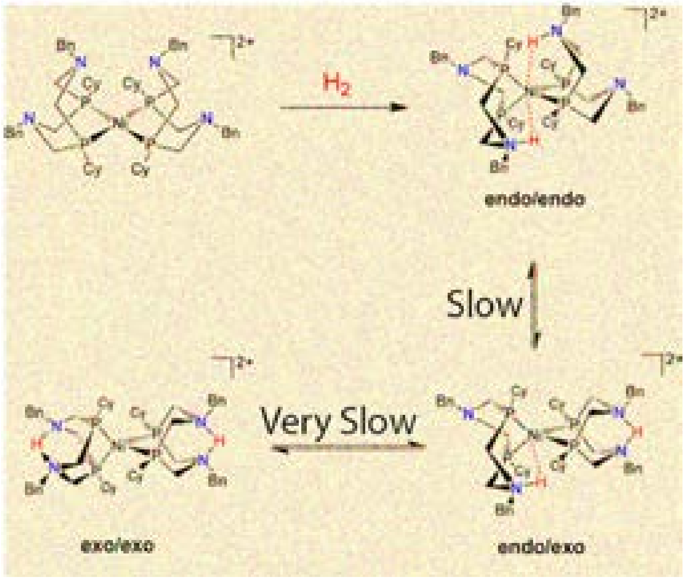
"When we started on the research, there was the belief that breaking or forming hydrogen was the crucial step—it isn't," said Simone Raugei, a theoretician at PNNL. "It is putting the protons in the right spot on

PROJECT TITLE
Molecular Electrocatalysis

NERSC PI
Shentan Chen
Michel Dupuis
Roger J. Rousseau
Simone Raugei
Sharon Hammes-Schiffer

LEAD INSTITUTION
Pacific Northwest National Laboratory

NERSC RESOURCES USED
Hopper, Carver, HPSS



◀ While one configuration (endo/endo) of a popular nickel catalyst can produce thousands of hydrogen molecules a second, the other configurations that place the proton farther from the center are slower and less efficient. PNNL

the catalyst. Once you have them in the right spot, everything goes very quickly."

At an atomic level, the nickel-based catalyst used in this study has four "floppy arms," called ligands, around the nickel center. These ligands twist into three distinctive isomers, which are essentially different arrangements of the same atoms. The team set out to see how protons move on the ligands of all three isomers.

"Once we realized that H_2 breaking and forming steps were not rate determining, we started focusing on proton delivery to and from the catalyst. We ran large-scale *ab initio* molecular dynamics simulations on Department of Energy supercomputers to shed some light on this process," said Raugei. "These calculations are computationally expensive, but they allow us to capture the full complexity of the system and match experimental conditions as much as possible."

The supercomputers helped the theoreticians identify barriers that kept the catalyst in less-productive isomeric forms. By combining these results with nuclear magnetic resonance spectrometry observations of the three different isomers, the team found that initially protons preferred to be positioned away from the nickel. Getting them from this place to a position near the nickel center takes several chemical steps, requiring energy and slowing down the catalyst. They also found that it is difficult to remove protons when they are next to nickel.

"Proton delivery is key to catalysis," said Raugei. "This is a very complicated catalytic environment, and we achieved this level of understanding by following the full chemistry of the system and comparing with the laboratory experiments. This work is a prime example of how theory and simulation can help us gain greater insights into chemistry and guide new chemistry research."

DOE PROGRAM OFFICE

BES—Chemical Sciences

FULL STORY

https://www.nersc.gov/news-publications/news/science-news/2013/supercomputers-help-a-catalyst-reach-its-full-potential/

PUBLICATION

O'Hagan M, MH Ho, JY Yang, AM Appel, M Rakowski DuBois, S Raugei, WJ Shaw, DL DuBois, and RM Bullock, "Proton Delivery and Removal in [Ni(PR2NR´2)2]2+ Hydrogen Production and Oxidation Catalysts," Journal of the American Chemical Society 134, 19409-19424, doi: 10.1021/ja307413x.

How Cells Interact with their Surroundings

New computer model of integrin has implications for cancer, atherosclerosis research

“Our all-atomic simulations would take forever if they were to be run on anything but a supercomputer.”

CELLS ARE SOCIAL BUTTERFLIES. They constantly interact with their surroundings, taking in cues on when to divide and where to anchor themselves, among other critical tasks.

This networking is driven in part by proteins called integrins, which reside in a cell’s outer plasma membrane. Their job is to convert mechanical forces from outside the cell into internal chemical signals that tell the cell what to do. But when they misfire, integrins can cause diseases such as atherosclerosis and several types of cancer.

Despite their importance, scientists don’t exactly know how integrins work. That’s because it’s very difficult to experimentally observe the protein’s molecular machinery in action.

But a computer model of integrin developed by Berkeley Lab researchers using supercomputers at NERSC is making it easier. Like its biological counterpart, the virtual integrin snippet is about 20 nanometers long. It also responds to changes in energy and other stimuli just as integrins do in real life. The result is a new way to explore how the protein connects a cell’s inner and outer environments.

“We can now run computer simulations that reveal how integrins in the plasma membrane translate external mechanical cues to chemical signals within the cell,” said Mohammad Mofrad, a faculty scientist in Berkeley Lab’s Physical Biosciences Division and associate professor of Bioengineering and Mechanical Engineering at UC Berkeley.

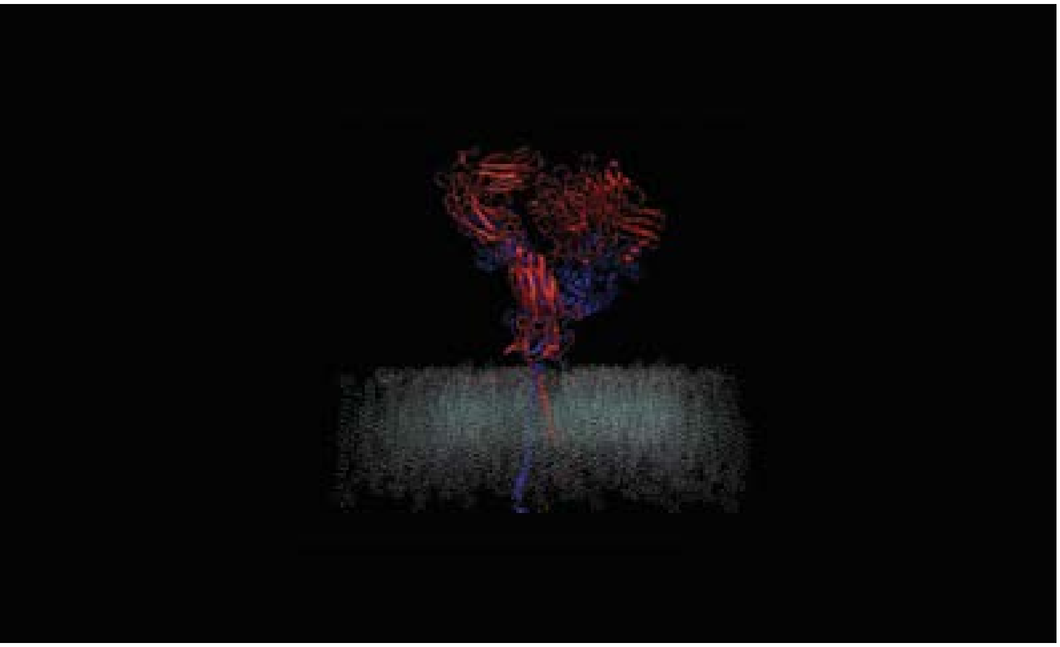
The project involved several simulation runs that took approximately 250,000 CPU hours on NERSC’s Carver system. “Our all-atomic simulations would take forever if they were to be run on anything but a supercomputer,” said Mofrad. “Thanks to NERSC, we managed to extract the essence of this biology problem within several months.”

PROJECT TITLE
Focal Adhesion Mechanotransduction

NERSC PI
H. Shams (Mofrad Lab)

LEAD INSTITUTION
Lawrence Berkeley National Laboratory

NERSC RESOURCES USED
Carver



◀ Two transmembrane integrin β-domains (in red) interact at their tails. The domains are embedded in a lipid layer that mimics the cell membrane.
MOFRAD LAB

The model is already shedding light on what makes integrins tick, such as how they "know" to respond to more force with greater numbers. When activated by an external force, integrins cluster together on a cell's surface and join other proteins to form structures called focal adhesions. These adhesions recruit more integrins when they're subjected to higher forces. As the model indicates, this ability to pull in more integrins on demand may be due to the fact that a subunit of integrin is connected to actin filaments, which form a cell's skeleton.

The model may also help answer a longstanding question: Do integrins interact with each other immediately after they're activated? Or do they not interact with each other at all, even as they cluster together? To find out, the scientists ran simulations that explored whether it's physically possible for integrins to interact when they're embedded in the plasma membrane. They found that interactions are likely to occur only between one compartment of integrin called the β-subunit.

These computationally obtained insights could guide new experiments designed to uncover how integrins do their job. "Our research sets up an avenue for future studies by offering a hypothesis that relates integrin activation and clustering," said Mofrad.

DOE PROGRAM OFFICE
BER—Biological Systems Science

FULL STORY
https://www.nersc.gov/news-publications/news/science-news/2013/simulations-yield-clues-to-how-cells-interact-with-surroundings/

PUBLICATION
Mehrdad Mehrbod, Mohammad Mofrad, "Localized Lipid Packing of Transmembrane Domains Impedes Integrin Clustering," PLoS Computational Biology 9(3): e1002948 (2013), doi:10.1371/journal.pcbi.1002948

An Inside Look at an MOF in Action

New insights into a metal-organic framework's structure could improve greenhouse gas capture

"We will be relying on NERSC's high-performance supercomputing resources ... as we look at even more complex MOFs."

NERSC SUPERCOMPUTING RESOURCES HELPED PROVIDE a unique inside look at the electronic structure of a highly touted metal-organic framework (MOF) as it is adsorbing carbon dioxide gas—findings that should help lead to more efficient methods of capturing and containing greenhouse gases.

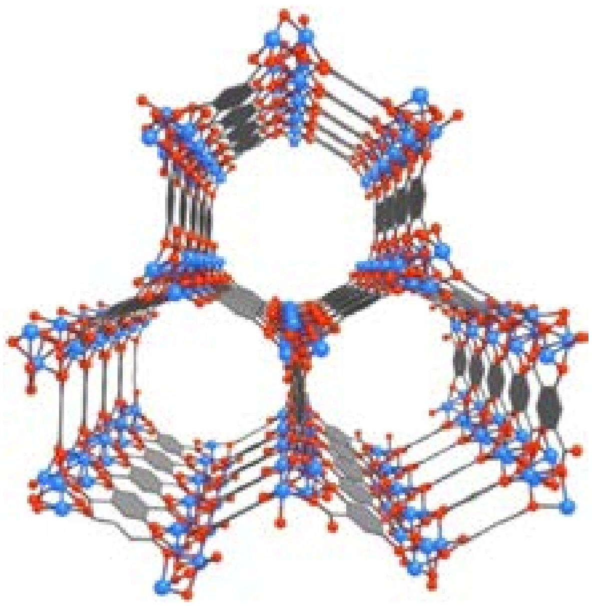
As part of the Energy Frontier Research Center (EFRC) for Gas Separations Relevant to Clean Energy Technologies, Berkeley Lab researchers recorded the first *in situ* electronic structure observations of the adsorption of CO_2 inside Mg-MOF-74, an open metal site MOF that has emerged as one of the most promising strategies for capturing and storing greenhouse gases.

A team led by Jeff Kortright of Berkeley Lab's Materials Sciences Division combined first-principles calculations performed at NERSC and the Molecular Foundry with experiments conducted at Berkeley Lab's Advanced Light Source (ALS), where they used an X-ray spectroscopy technique known as Near Edge X-ray Absorption Fine Structure (NEXAFS) to obtain what are believed to be the first ever measurements of chemical and electronic signatures inside of an MOF during gas adsorption.

MOFs are molecular systems consisting of a metal oxide center surrounded by organic "linker" molecules that form a highly porous three-dimensional crystal framework. This microporous crystal structure enables MOFs to serve as storage vessels with a sponge-like capacity for capturing and containing greenhouse gases. When a solvent molecule applied during the formation of the MOF is subsequently removed, the result is an unsaturated "open" metal site MOF that has a strong affinity for CO_2.

"We've demonstrated that NEXAFS spectroscopy is an effective tool for the study of MOFs and gas adsorption," Kortright said. "Our study shows that open metal site MOFs have significant X-ray spectral signatures that are

PROJECT TITLE	NERSC PI	LEAD INSTITUTION	NERSC RESOURCES USED
Simulating Core-Level Spectroscopy of Energy-Relevant Materials	David Prendergast	Lawrence Berkeley National Laboratory	Hopper



◀ Mg-MOF-74 is an open metal site MOF whose porous crystalline structure could enable it to serve as a storage vessel for capturing and containing the carbon dioxide emitted from coal-burning power plants. NATIONAL ACADEMY OF SCIENCES

highly sensitive to the adsorption of CO_2 and other molecules."

NERSC's Hopper was crucial to the first phase of this work, the researchers noted.

"The spectral calculations are highly sensitive to the positions of the atoms in the MOF, so starting from the relaxed, lowest-energy configuration is critical," said co-author Walter Drisdell. "Moving forward, we will be relying on NERSC's high-performance supercomputing resources, like Hopper and Edison, even more as we look at more complex MOFs and simulate how they behave at ambient temperature."

To examine adsorption in Mg-MOF-74, Drisdell designed a special gas cell that enabled NEXAFS measurements to be made as CO_2 pressure was varied from vacuum up to 100 Torr at ambient temperature. This capability provided the means to make direct comparisons between empty and bound sites of the same Mg-MOF-74 sample. NEXAFS measurements were made at ALS beamline 6.3.1, a bending magnet beamline optimized for X-ray absorption spectroscopy.

Spectral simulations were performed in collaboration with co-author David Prendergast, a nanostructures theorist at the Molecular Foundry.

"The calculations were a great aid in interpreting our spectra," Drisdell said. "Not only could we reproduce the spectral signatures we observed upon adsorption, but we could show that these signatures arise from a specific, distorted electronic state at the open metal sites that displays a unique interaction with different adsorbed molecules."

DOE PROGRAM OFFICE

BES—Materials Sciences

FULL STORY

https://www.nersc.gov/news-publications/news/science-news/2013/an-inside-look-at-a-mof-in-action/

PUBLICATION

Walter S. Drisdell, Roberta Poloni, Thomas M. McDonald, Jeffrey R. Long, Berend Smit, Jeffrey B. Neaton, David Prendergast, Jeffrey B. Kortright, "Probing Adsorption Interactions In Metal-Organic Frameworks Using X-ray Spectroscopy," Journal of the American Chemical Society, November 13, 2013, 135(48), doi: 10.1021/ja408972f

NISE Program Encourages Innovative Research

The NERSC Initiative for Scientific Exploration (NISE) is a mechanism used for allocating the NERSC reserve (10 percent of the total allocation). It is a competitive allocation administered by NERSC staff and management. In 2013, NERSC awarded this time by making the second year of the two-year NISE awards made in 2012, supplemented by projects selected by the NERSC director. In 2013, NERSC awarded 125 million hours under NISE to 27 projects, as listed below.

Project	**Investigator**	**Hours Awarded**	**DOE Office and Program**
NERSC Application Readiness for Future Architectures	Katie Antypas, *Lawrence Berkeley National Laboratory*	250,000	ASCR Applied Mathematical Sciences
Carbon Data Assimilation with a Coupled Ensemble Kalman Filter	Inez Fung, *University of California Berkeley*	750,000	BER Climate and Environmental Sciences
A Multi-decadal Reforecast Data Set to Improve Weather Forecasts for Renewable Energy Applications	Thomas Hamill, *National Oceanic & Atmospheric Administration*	10,000	BER Climate Research
Numerical Simulations of Nanocrystals, Interfaces, and Grain Boundaries in Complex Materials	Juan Idrobo, *Oak Ridge National Laboratory*	4,000,000	BES Materials Science
First Principles Molecular Dynamics Investigation of Carbon Dioxides	Yosuke Kanai, *University of North Carolina at Chapel Hill*	1,250,000	BES Materials Science
Turbulent Reacting Flows for Multi-physics Model Development	Colleen Kaul, *Stanford University*	700,000	BES Chemistry
Electronic Properties of Novel Nitride Nanostructures	Emmanouil Kioupakis, *University of Michigan*	7,000,000	BES Materials Science
Multi-scale Multi-compartment Computational Models of the Human Immune Response to Infection with M. tuberculosis	Denise Kirschner, *University of Michigan*	250,000	BER Biological Systems Science
Multi-scale Modeling of Nanoparticle Self-assembly and Molecular Electronics in Nanocarbons	Petr Kral, *University of Illinois*	1,500,000	BES Materials Science

Project	Investigator	Hours Awarded	DOE Office and Program
Sampling Diffusive Dynamics on Long Timescales, and Simulating the Coupled Dynamics of Electrons and Nuclei	Thomas Miller, *California Institute of Technology*	15,000,000	BES Chemistry
Large Eddy Simulation of Turbulence-Chemistry Interactions in Reacting Multiphase Flows	Joseph Oefelein, *Sandia National Laboratories*	4,400,000	BES Chemistry
Turbulence over Complex Terrain: A Wind Energy Perspective	Edward Patton, *National Center for Atmospheric Research*	3,000,000	BER Climate Research
Joint Center for Energy Storage Research	Kristin Persson, *Lawrence Berkeley National Laboratory*	4,000,000	BES Materials Science
The Material Genome	Kristin Persson, *Lawrence Berkeley National Laboratory*	4,000,000	BES Materials Science
Calibration of 3D Upper Mantle Structure in Eurasia Using Regional and Teleseismic Full Waveform Seismic Data	Barbara Romanowicz, *University of California Berkeley*	3,000,000	BES Geosciences
Quantum Transport Simulation of Nanoscale Electronic Devices for Ultra Low Power Computing	Sayeef Salahuddin, *University of California Berkeley*	200,000	BES Materials Science
Surface and Interface of Photocatalytic Metal Oxide Materials	Annabella Selloni, *Princeton University*	3,275,000	BES Materials Science
Thermodynamics of Secondary Aerosol Formation	George Shields, *Bucknell University*	100,000	BER Climate Research
Next Generation Bioimaging Institute	David Skinner, *Lawrence Berkeley National Laboratory*	200,000	BER Biological Systems Science
Static and Dynamic Solutions for Heavy Nuclei	Ionel Stetcu, *Los Alamos National Laboratory*	5,000,000	NP Nuclear Theory
Attribution of Extreme Weather Risk to Anthropogenic Emissions	Daithi Stone, *Lawrence Berkeley National Laboratory*	9,500,000	BER Climate Research
Pore Scale Reactive Transport Processes Associated with Carbon Sequestration	David Trebotich, *Lawrence Berkeley National Laboratory*	25,000,000	ASCR Applied Mathematical Science
Computational Prediction and Discovery of Magnet Materials	Cai-Zhuang Wang, *Ames Laboratory, Iowa State University*	12,000,000	BES Materials Science
Spin-lattice Coupling in Magnetic Phase Transition	Yi Wang, *Pennsylvania State University*	8,800,000	BES Materials Science
Multiple Exciton Generation and Charge Extraction in Nanoparticle-based Solar Cells	Stefan Wippermann, *University of California Davis*	5,000,000	BES Materials Science
Guest-Host Interactions in Aqueous Systems and Hydrate Lattices: Implications for H2 Storage and CO_2 Sequestration	Sotiris Xantheas, *Pacific Northwest National Laboratory*	8,000,000	BES Chemistry
Development of ITM Oxygen Technology for Integration with Advanced Industrial Systems	Bi-Cheng Zhou *Pennsylvania State University*	265,000	BES Materials Science

NERSC Users' Awards and Recognition

Long-time NERSC User Martin Karplus Wins Nobel Prize in Chemistry

In October 2013, the Nobel Prize in Chemistry was awarded to three scientists for pioneering methods in computational chemistry that have brought a deeper understanding of complex chemical structure and reactions in biochemical systems. These methods can precisely calculate how very complex molecules work and even predict the outcomes of very complex chemical reactions.

One of the laureates—Martin Karplus of Harvard University—has been using supercomputers at NERSC since 1998. The other laureates were Michael Levitt of Stanford University and Arieh Warshel of the University of Southern California.

According to the Royal Swedish Academy, these accomplishments have opened up an important collaboration between theory and experiment that has made many otherwise unsolvable problems solvable.

"Today the computer is just as important a tool for chemists as the test tube. Simulations are so realistic that they predict the outcome of traditional experiments," the Royal Swedish Academy wrote in its announcement of the winners.

Martin Karplus ▶

Long gone are the days when chemists used plastic balls and sticks to create models of molecules. Today, modeling is carried out on computers, and Karplus' work helped lay the foundation for the powerful programs that are used to understand and predict chemical processes. These models are crucial for most of the advances made in chemistry today.

Karplus, Levitt and Warshel revolutionized the field of computational chemistry by making Newton's classical physics work side-by-side with fundamentally different quantum physics. Previously, researchers could only model one or the other. Classical physics models were ideal for modeling large molecules, but they couldn't capture chemical reactions. For that purpose, researchers instead had to use quantum physics. But these calculations required so much computing power that researchers could only simulate small molecules.

By combining the best from both physics worlds, researchers can now run simulations to understand complex processes like how drugs couple to its target proteins in the body. For example, quantum theoretical calculations show how atoms in the target protein interact with the drug. Meanwhile, less computationally demanding classical physics is used to simulate the rest of the large protein.

Karplus began computing at NERSC in 1998, with an award from Department of Energy's Grand Challenges competition. The Grand Challenges applications addressed computation-intensive fundamental problems in science and engineering, whose solution could be advanced by applying high performance computing and communications technologies and resources.

At the time, Karplus and his colleague, Paul Bash who was at Northwestern University, were looking to understand chemical mechanisms in enzyme catalysis, which they couldn't investigate experimentally. So they ran computer simulations at NERSC to gain a complete understanding of the relationship between biomolecular dynamics, structure and function.

One of the enzymes they looked at was a class called beta-lactamases. Researchers knew that these enzymes were responsible for the increasing resistance of bacteria to antibiotics, but the precise chemical resistance mechanisms were unknown. So Karplus and Bash ran simulations on NERSC supercomputers to investigate this mechanism at an atomic level of detail.

In his 15 years as a NERSC investigator, Karplus and his research group have explored everything from how the molecule ATP synthase acts as a motor that fuels cells to how myosin, the molecular engine behind muscles, operates. Today, Karplus' group is tackling the science behind molecular machines, which may someday power man-made systems, for example by converting sunlight into biofuels; working as tiny "molecular motors" capable of performing chemical analyses or other tests for "lab-on-chip" devices; or even "manufacturing" nanodevices.

NERSC Supports Top Science Discoveries of 2013

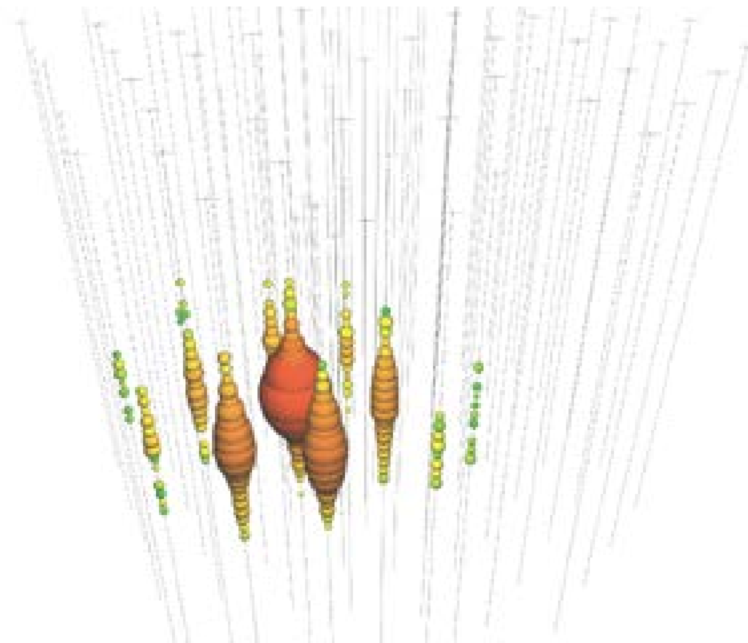
Research supported by NERSC was honored by end-of-year reviews in two leading magazines: Physics World and WIRED. The IceCube South Pole Neutrino Observatory was notably named to both lists, being honored as the most important discovery by Physics World.

Three of Physics World's top 10 breakthroughs of 2013 went to discoveries that used NERSC resources. In addition to the IceCube South Pole Neutrino Observatory's top honor, "Breakthrough of the Year," the magazine named the European Space Agency's Planck space telescope, which revealed new information about the age and composition of the universe; and the South Pole Telescope, which made the first detection of a subtle twist in light from the CMB, known as B-mode polarization.

WIRED magazine's Top Science Discoveries of 2013 named two findings NERSC had a hand in, including IceCube and the final findings of the NASA Kepler space telescope: One in five Sun-like stars in our galaxy has an Earth-sized planet orbiting it at a habitable distance.

Breakthrough of the Year: IceCube South Pole Neutrino Observator

The magazine's "Breakthrough of the Year" honor went to the IceCube South Pole Neutrino Observatory for making the first observations of high-energy cosmic neutrinos—particles that originate from far beyond the solar system. Neutrinos are notoriously difficult to detect, and IceCube achieved this result by building a colossal detector deep under the ice at the South Pole. The observatory can also determine the neutrino's direction, making it an incredibly useful telescope.



▶ In November 2013, the IceCube Collaboration announced the observation of 28 very high-energy particle events—the first solid evidence for astrophysical neutrinos from cosmic accelerators. The neutrinos were found in data collected by the IceCube detector from May 2010 to May 2012 and analyzed for neutrino events that exceed 50 teraelectronvolts (TeV). The colors show when the light arrived, with reds being the earliest, succeeded by yellows, greens, and blues. The size of the circle indicates the number of photons observed. ICECUBE COLLABORATION

In November 2013, the IceCube Collaboration publicly announced the observation of 28 extremely high-energy events that constitute the first solid evidence for high-energy astrophysical neutrinos from outside our solar system, including two of the highest energy neutrinos ever reported, which have been named "Bert" and "Ernie." Published in Science, the results provide experimental confirmation that something is accelerating particles to energies above 50 trillion electron volts (TeV) and exceeding one quadrillion electron volts (PeV). The team used NERSC supercomputing resources, located at Berkeley Lab, to sift out neutrino signals from cosmic "noise" in the IceCube observations.

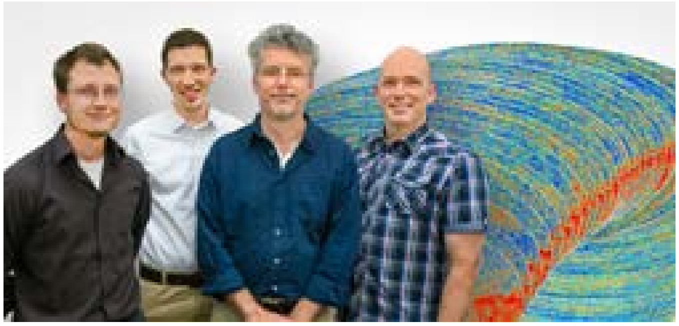
Planck Reveals 'Almost Perfect' Universe

The European Space Agency's Planck space telescope also made this year's Physics World breakthroughs list for producing the most precise measurement ever of the cosmic microwave background (CMB) radiation. Written in light shortly after the big bang, the CMB is a faint glow that permeates the cosmos.

Thanks to Planck, which launched in 2009, we now know that the proportion of the universe made up of dark energy is slightly less than previously thought, but there is more dark matter and ordinary matter than previous studies of the CMB had suggested. Planck also found that the universe is about 100 million years older than previously thought. In addition, the Planck data contain tantalizing hints of anomalies in the temperature of the CMB in different parts of the universe, which could point toward new physics.

CMB surveys are complex and subtle undertakings, and NERSC supercomputers were crucial to sifting the CMB's faint signal out of a noisy universe and decoding its meaning. Using NERSC resources, Planck scientists created the most detailed and accurate maps yet of the relic radiation from the big bang.

"CMB data sets grow with Moore's Law, increasing a thousand-fold over the last 15 years and projected to do the same over the next 15 too," said Julian Borrill, a Planck scientist and co-founder of the Computational Cosmology Center (C3) at Berkeley Lab. His group has been developing the supercomputing tools needed to analyze the big data from CMB experiments for over a decade.



◀ From left, Reijo Keskitalo, Aaron Collier, Julian Borrill, and Ted Kisner of the Computational Cosmology Center with some of the many thousands of simulations for analyzing Planck data. ROY KALTSCHMIDT

"These maps are proving to be a goldmine containing stunning confirmations and new puzzles," says Martin White, a Planck scientist and physicist with University of California Berkeley and at Berkeley Lab. "This data will form the cornerstone of our cosmological model for decades to come and spur new directions in research."

B-mode Polarization Spotted in Cosmic Microwave Background

The IceCube collaboration may have bagged the Physics World 2013 Breakthrough of the Year award, but another discovery from the South Pole also made Physics World's top 10 list. This year, the South Pole Telescope made the first detection of a subtle twist in light from the CMB, known as B-mode polarization. This twist has long been predicted and its detection paves the way for a definitive test of inflation—a key theory in the Big Bang model of the universe.

◀ South Pole Telescope at night. ICECUBE COLLABORATION

The South Pole Telescope team used NERSC resources, through an allocation shared by a dozen suborbital CMB experiments, including data analysis on the Hopper and Edison supercomputers and archiving on the High Performance Storage System.

Earth-like Planets Not Uncommon in Our Galaxy

One out of every five Sun-like stars in our Milky Way galaxy has an Earth-sized planet orbiting it in the Goldilocks zone—not too hot, not too cold—where surface temperatures should be compatible with liquid water, according to a statistical analysis of data from NASA's Kepler spacecraft by Erik Petigura, a graduate student at the University of California, Berkeley.

Petigura and his colleague Andrew Howard, now at the University of Hawaii, Manoa, spent three years developing a transit search pipeline called TERRA, which is optimized for finding small planets. When they used this tool on supercomputers at NERSC to analyze nearly four years of Kepler observations, the scientists determined that our galaxy could contain as many as 40 billion habitable Earth-sized planets.

▲ The artist's concept depicts NASA's Kepler mission's smallest habitable zone planet. Seen in the foreground is Kepler-62f, a super-Earth-size planet in the habitable zone of a star smaller and cooler than the sun, located about 1,200 light-years from Earth in the constellation Lyra. NASA

Collaboration Shines in Materials Project Success

Scientific American featured a Computing Sciences-powered project on its December cover as a top world-changing idea of 2013. The Materials Project, an open, web-hosted service, allows scientists to use quantum mechanical equations to design new materials atom by atom, before ever running an experiment, the magazine noted.

"Materials science is on the verge of a revolution," wrote co-authors Kristin Persson of Berkeley Lab and Gerbrand Ceder of MIT. "We can now use a century of progress in physics and computing to move beyond the Edisonian process."

The Materials Project aims to take the guesswork out of finding the best material for a job—be it a new battery electrode or a lightweight spacecraft body—by making the characteristics of every inorganic compound available to any interested scientist. With 35,000 materials and 5,000 users, the once small, experimental project has grown to become what is likely the largest and arguably the most sophisticated open materials database yet fielded. The ultimate goal is to cut in half the amount of time it typically takes to bring new materials to market, which is currently about 18 years.

NERSC serves as the computing and data engine for the project. It provides the software and hardware infrastructure for the web gateway and databases that serve up the Materials Project data. In addition to supporting calculations on supercomputers, the NERSC Division maintains cluster nodes purchased by and dedicated solely to the Materials Project. ESnet connectivity enables access to the NERSC science gateways, and serves as the platform to access data resources at NERSC.

Persson, of Berkeley Lab's Environmental Energy Technologies Division, brought the Materials Project with her from MIT, where as a post-doctoral researcher she co-founded it with Ceder. At Berkeley Lab, Computing Sciences' world-class talent and resources combined to help grow the already promising project into the world-changing idea it is today.

As the Materials Project transitioned from its experimental roots to a full production service for thousands of users, its SQL database emerged as a key difficulty. Dan Gunter of Berkeley Lab's Computational Research Division (CRD) led the transformation of the project's data model from traditional SQL to MongoDB, with assistance from Shreyas Cholia, acting group leader of NERSC's Data and Analytics Group, and Monte Goode, also of CRD.

The Mongo DB model lets researchers add new data and then generates its structure (or schema) from that. This technology is used as the workhorse of the Materials Project to schedule and track quantum mechanical calculations of materials properties on supercomputers, to store and search the results of these computations, and to perform advanced analytics on the computed materials properties.

Cholia and David Skinner, strategic partnerships lead for NERSC, were instrumental in creating the NERSC science gateway that hosts the Materials Project. Using an interactive, web-based interface, researchers can peruse compounds, access applications to explore and visualize materials and even submit new calculations to NERSC computers.

The Materials Project also provides an HTTP REST application programming interface (API) that makes it possible to programmatically interact with the project's database directly. Using the API, researchers can run their own analyses directly on the data, an extremely powerful tool for collaborating with other projects that wish to consume Materials Project data.

Four NERSC Users Honored with Annual HPC Achievement Awards

NERSC announced the winners of its 2013 High Performance Computing (HPC) Achievement Awards on Feb. 4, 2014, during the annual NERSC User 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 presented two early career awards.

"Choosing the recipients of such achievement awards is no easy task, as all 5,000 of our users are tackling important scientific challenges and we can only honor four of them," said NERSC Director Sudip Dosanjh. "The accomplishments of our 2013 NERSC Achievement Award winners really highlight the wide range of research supported by NERSC, in this case studying problems ranging from the beginning of our universe to developing innovative algorithms for accelerator design, from creating a public database for developing new batteries and solar panels to modeling the transitions that underlie many processes in biology."

"While many of NERSC's users are longtime users who have a consistent record of accomplishment, it's important that we also acknowledge the innovative work of those in the formative years of their work, as evidenced by our Early Career Award winners," said Richard Gerber, leader of NERSC's User Services Group. "These young researchers are the future of science and they bring new ideas, new perspectives and new energy to their chosen fields."

Here are the winners for 2013:

NERSC Award for High-Impact Science—Open Division

Julian Borrill ▶

The Planck Collaboration is the recipient of the 2014 NERSC Achievement Award for High-Impact Science. The Planck satellite, a joint mission between the European Space Agency and NASA, made news in 2013 when its science team released the most detailed map ever made of the Cosmic Microwave Background—the remnant radiation from the Big Bang—and refined some of the fundamental parameters of cosmology and physics. Long-awaited by many constituencies in cosmology, astrophysics, and theoretical physics, these results, comprising 31 scientific papers, capped a 10-year collaboration between

NERSC and the Planck team, with Julian Borrill of Berkeley Lab's Computational Cosmology Center serving as the principal investigator at NERSC.

CMB data analysis is computationally challenging. For over a decade Planck has relied on NERSC to p rovide the necessary computational capabilities, including tens of millions of CPU hours, hundreds of terabytes of spinning disk space, and support for hundreds of Planck data analysts. The culmination of this work to date, and the most computationally challenging part of the entire CMB analysis pipeline, has been the production of the sixth Full Focal Plane simulation (FFP6), comprising 1,000 realizations of the Planck mission reduced to 250,000 maps. FFP6 is essential both to quantify uncertainties in and remove biases from the analysis of the Planck data, and to validate and verify the analysis pipelines themselves. Considerable effort has gone into the code optimization and system-specific tuning needed to make these simulations tractable, providing a 1,000-fold speed-up from the first single-frequency mapmaking on Seaborg in 2006 (also the first application code test of the newly-enabled 6,000-way parallelism) to 70,000-way parallel runs of FFP6 on Hopper in 2012-13.

NERSC Award for High-Impact Science—Early Career

Victor Ovchinnikov, a post-doctoral fellow working with 2013 Nobel Laureate Martin Karplus, is the recipient of the 2014 NERSC early career award for high impact scientific achievement. According to Karplus, during the years 2011-13, with the help of NERSC resources, Ovchinnikov made outstanding contributions to the field of computational modeling of conformational transitions in large biological molecules, such as proteins and DNA. Conformational transitions underlie a wide range of biological processes, such as enzyme catalysis, DNA replication, refolding of misfolded proteins, ATP synthesis, muscle contraction and multidrug resistance. But these transitions are extremely difficult to investigate by experimental methods because of the small spatiotemporal scales involved.

Computer simulation of conformational transitions in biomolecules is a generally unsolved problem in computational science because spontaneous transitions in proteins often require millisecond simulation times. To overcome the time-scale limitation of standard molecular dynamics simulations, various enhanced techniques have been developed. However, most of the methods require a priori information, which is generally unknown except for simple molecular systems.

Ovchinnikov's contribution to the field was to extend a family of transition path algorithms under the umbrella of the "string method" beyond proof-of-principle biological models to transitions in real proteins of biological importance. The Karplus group is currently collaborating with other researchers who are using the string methods to understand a variety of biological and chemical problems. More generally, Ovchinnikov's developments of the string method are expected to become routinely used for other types of transitions in biology and chemistry.

NERSC Award for Innovative Use of HPC—Early Career

Anubhav Jain, a research scientist/chemist in Berkeley Lab's Environmental Energy Technologies Division is recipient of the 2014 NERSC early career award for innovative use of HPC. Jain's work focuses on new materials discovery using high-throughput computations. Jain is the primary author of the FireWorks code for automating calculations at supercomputing centers. A free, open-source code for defining, managing and executing scientific workflows, FireWorks can be used to automate calculations over arbitrary computing resources, including those that have a queuing system.

He is pursuing major projects under the auspices of both the Materials Project and the Joint Center for Energy Storage Research (JCESR). Both of his projects are aimed at developing more efficient batteries. The Materials Project is a multi-institution effort to accelerate materials discovery by computing the properties of all known inorganic properties. This is achieved by automating first-principles calculations on a massive scale at DOE supercomputing centers. The Materials Project currently provides more than 35,000 pre-computed compounds via a NERSC science gateway to over 5,000 users from all over the world. JCESR is a DOE Energy Innovation Hub led by Argonne National Laboratory and supported by DOE's Office of Science. Jain's role in JCESR is similar: to uncover novel electrolytes for next-generation batteries using high-throughput computations to efficiently screen amongst thousands of candidates.

Jain, a recipient of the DOE Computational Science Graduate Fellowship and Berkeley Lab Computing Sciences Luis W. Alvarez Postdoctoral Fellowship, currently serves as an executive member of the NERSC Users Group.

NERSC Award for Innovative Use of HPC—Open Division

Jean-Luc Vay of Berkeley Lab's Accelerator and Fusion Research Division is the recipient of the 2013 NERSC award for innovative use of HPC in the open division.

▶ Ted Kisner (International Planck Collaboration), Jean-Luc Vay and Anubhav Jain accept their NERSC Achievement awards.
ROY KALTSCHMIDT

An internationally recognized pioneer in the development of novel algorithms for simulation of beams and particles, Vay is highly esteemed by fellow theoretical/computational physicists at many institutions worldwide. Vay's work in developing innovative algorithms has greatly improved the use of high performance computing to advance the simulation of charged particles, beams and plasmas.

Perhaps the most far-reaching advance made by Vay during the past few years is the development of the Lorentz-boost algorithm as a practical tool to perform multi-scale simulations. While the idea had been conceived earlier elsewhere, Vay thought of it independently, and then contributed the critical elements to make it work in practice and turn it into an impressive tool. As a result, certain multi-scale/multi-species simulations are speeded up by several orders of magnitude without loss of accuracy. Adopted at many institutions worldwide, this technique has been successfully applied to free electron lasers, laser-plasma acceleration and particle beams interacting with electron clouds. It enables massive parallelism to be efficiently exploited with moderate numbers of time steps to make routine simulations that were impossible until recently.

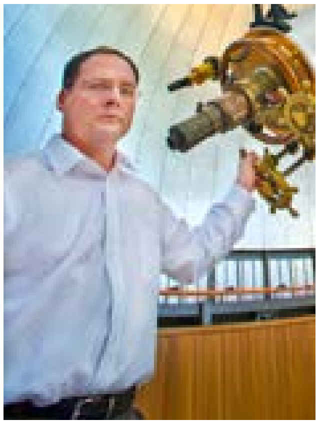
Vay's current project is to improve the scaling of simulations using certain electromagnetic solvers. A well known issue in simulations of electromagnetics, charged particles and plasmas is that spectral electromagnetic solvers are accurate but do not scale well, while finite differences scale well but are much less accurate. Vay exploited the finite propagation of the speed of light to allow decomposition of spectral solvers in a way similar to (and with similar scaling to) finite differences while maintaining the superior accuracy of the spectral method. This effectively adapts a highly accurate method to efficiently use high performance computing. He has recently published this method, working with some of the developers of early spectral methods, and is collaborating with computer scientists at NERSC and Berkeley Lab's Computational Research Division to further improve the decomposition in ways suitable for system architectures.

Peter Nugent, Elizabeth Bautista Honored for Exceptional Achievement

In August 2013, Peter Nugent and Elizabeth Bautista of NERSC were recognized with Lawrence Berkeley National Laboratory Director's Awards of Exceptional Achievement.

Nugent, co-leader of the Computational Cosmology Center in the Computational Research Division and NERSC Analytics Group Lead, was honored for scientific achievement in helping create the Palomar Transient Factory, a scientific pipeline of images from the universe, streamed from the Palomar Observatory in Southern California and then archived at NERSC and made available to scientists around the world as they seek to increase our understanding of supernovae. In presenting the award, Associate Laboratory Director for General Sciences James Symons called Nugent's work "groundbreaking" and setting a new standard for supporting scientific research.

Bautista, NERSC's Operations Technology Group Lead, was recognized for her work in diversity, particularly her efforts to encourage students to pursue careers in computer science and to give the students hands-on experience through expanded internship opportunities. As she presented the award, Associate Laboratory Director for Computing Sciences Kathy Yelick said that Bautista has "made the lab a better and friendlier place to work."



◀ Left, Peter Nugent; Right, Elizabeth Bautista ROY KALTSCHMIDT

NERSC Center News

Jeff Broughton Promoted to Division Deputy for Operations

In August 2013, Jeff Broughton was appointed the new NERSC Division Deputy for Operations.

"Rather than this being a new position, the division deputy title is a fitting recognition of the duties and responsibilities Jeff has taken on since he joined NERSC four years ago," NERSC Division Director Sudip Dosanjh said. "When he was hired as head of the Systems Department in 2009, Jeff brought 30 years of HPC and management experience to the position."

Broughton's career includes nine years at Lawrence Livermore National Laboratory, where he served as both a project leader and a group leader in computing. He also spent ten years at Amdahl Corporation, where he worked in both computer architecture development and marketing. During a two-year stint at Sun Microsystems, he was awarded five system architecture patents and played key role in developing a massively parallel system architecture for Sun.

Broughton was recruited by the startup firm PathScale Inc. in 2001 and helped build an organization of 50 employees to develop cluster computer systems. In 2005, he won the HPCWire "Most Significant New HPC Software Product for 2005" for delivering a commercially viable compiler based on open source technology. In 2006 PathScale was acquired by QLogic, and Broughton continued to lead the hardware and software organization for InfiniBand-related products.

"Jeff's ability to build effective partnerships with both hardware and software vendors will be critical as NERSC moves into the realm of exascale computing," Dosanjh said.

Katie Antypas Leads NERSC Services Department

Katie Antypas, who led NERSC's User Services Group since October 2010, became Services Department Head in September 2013. Antypas took over for Francesca Verdier, who will serve as Allocations Manager until her planned retirement in June 2014.

Antypas is also the project lead for the NERSC-8 system procurement, a project to deploy NERSC's next generation system in 2016. In addition to increasing the computational capability available to NERSC's 5,000 users, the system will also serve as a platform that will begin to transition DOE scientific applications to more energy-efficient, manycore architectures.

"Katie's leadership in ensuring that NERSC users are able to maximize their use of both our current and future systems has positioned her well to help lead NERSC users and staff into the next era of extreme scale computing," said NERSC Division Director Sudip Dosanjh.

Antypas joined NERSC's User Services Group in 2006, serving as a consultant to users. She was the co-implementation team lead on Hopper, NERSC's first petaflop/s system.

Before coming to NERSC, Antypas worked at the ASC Flash Center at the University of Chicago supporting the FLASH code, a parallel adaptive mesh refinement astrophysics application.

Nick Wright Tapped to Head Advanced Technologies Group

Nick Wright was appointed head of NERSC's Advanced Technologies Group (ATG) in February 2013. ATG focuses on understanding the requirements of current and emerging applications to make choices in hardware design and programming models that best serve the science needs of NERSC users. The team engages with vendors and the general research community to advocate technological features that will enhance the effectiveness of systems for NERSC scientists.

Wright works closely with NERSC management to set strategic directions for the facility. To provide input for HPC system procurements, he collaborates with Lawrence Berkeley National Laboratory's Future Technologies Group to assess emerging technologies in architecture, algorithms, parallel programming paradigms and languages.

Wright joined NERSC in 2009 and initially worked on the Integrated Performance Monitoring framework. More recently, he worked on performance measurement and optimization as part of the NERSC-Cray Center of Excellence. He is the architecture and performance lead for the NERSC-7 (Edison) and NERSC-8 procurement projects and has published more than 20 papers related to high performance computing.

Richard Gerber New NERSC User Services Group Lead

Richard Gerber was named leader of the NERSC User Services Group in 2013, succeeding Katie Antypas, who was appointed head of the NERSC Services Department in September.

In his new role, Gerber manages a group that works to increase the scientific productivity of NERSC's 5,000 users. In addition to providing consulting services to users, he oversees and contributes to the NERSC web site, coordinates training efforts, serves as the liaison to the NERSC User Group, manages NERSC's trouble ticket system and designs and implements the annual user survey.

"Richard is the innovator behind many of NERSC's most popular web services that have been mimicked by HPC centers around the world, including the live Running Jobs page, the online queue look and the completed jobs workload statistics pages," Antypas said.

Gerber also served as the NERSC-7 (Edison) project deputy and co-leads the NERSC Program Requirements Reviews, which bring together leading scientists, DOE program managers and NERSC staff to gather future HPC requirements for the DOE Office of Science.

Gerber joined NERSC in 1996 as a consultant in the User Services Group. He also holds the title of NERSC Senior Science Advisor.

NERSC Helps Launch DesignForward Exascale Partnership

In November 2013, the DOE Office of Science and the National Nuclear Security Administration (NNSA) announced that they had awarded $25.4 million in research and development contracts to five leading companies in high performance computing (HPC) to accelerate the development of next-generation supercomputers.

Under DOE's new DesignForward initiative, AMD, Cray, IBM, Intel Federal and NVIDIA are working to advance extreme-scale, on-the-path to exascale computing technology that is vital to national security, scientific research, energy security and the nation's economic competitiveness.

The DesignForward contracts, which cover a two-year performance period, will support the design and evaluation of interconnect architectures for future advanced HPC architectures. The interconnects will tie together hundreds of thousands or millions of processors as building blocks of supercomputers to be used in studying complex problems in unprecedented detail. The DesignForward focus will be on developing interconnects that are energy efficient, have high bandwidth and minimize the time needed to move data among processors.

"A major disruption is facing high performance computing because energy constraints are causing our building blocks—microprocessors and memory—to change dramatically," said NERSC Director Sudip Dosanjh. "We need to collaborate with computer companies to ensure that future supercomputers meet DOE's mission needs in science, energy and national security. Berkeley Lab is pleased to place these contracts on behalf of DOE and its laboratories."

Under the contract, Intel will focus on interconnect architectures and implementation approaches, Cray on open network protocol standards, AMD on interconnect architectures and associated execution models, IBM on energy-efficient interconnect architectures and messaging models and NVIDIA on interconnect architectures for massively threaded processors. The vendors will collaborate with DOE's Exascale Co-design Centers to determine how changes in the system architectures will affect how well the scientific applications perform.

DesignForward is funded by DOE's Office of Science and NNSA and technically managed by Argonne, Lawrence Berkeley, Lawrence Livermore, Los Alamos, Oak Ridge, Pacific Northwest and Sandia national laboratories.

Petascale Post-Doc Project an HPC Success Story

The first post-doctoral research project centered at NERSC is being credited with helping its participants advance their careers while boosting the state of the art in high performance scientific computing software applications.

The Computational Science and Engineering Petascale Initiative—also referred to as the Petascale Post-doc Project—was unveiled in 2009. The goal was to hire several post-doctoral researchers to work with science teams on research in key application areas that support DOE's energy mission and encourage the use of NERSC supercomputers for energy research. The program was funded through the American Recovery and Reinvestment Act (ARRA).

The Computational Science and Engineering Petascale Initiative identified key application areas with specific needs for advanced programming models, algorithms and other support. The applications were chosen to be consistent with the current mission of the DOE, with a particular focus on applications that benefit energy research, those supported by other ARRA funding and Energy Frontier Research Centers (EFRCs). The initiative paired post-doctoral researchers at NERSC with these high-impact projects.

In all, nine post-docs participated in the project, each for about two years: Brian Austin, Kirsten Fagnan, Christos Kavouklis, Jihan Kim, Bobby Liu, Filipe Maia, Praveen Narayanan, Robert Preissl and Rebecca Yuan. The program was centered at the Lab's Oakland Scientific Facility (OSF), where NERSC resides, and overseen by principal investigator Alice Koniges of NERSC.

The post-docs collaborated with DOE-funded scientists on projects ranging from carbon capture and sequestration to fusion energy simulations, advanced accelerator modeling and geophysical imaging. They worked to increase the performance and science content of energy-related codes running on NERSC systems and provide feedback to NERSC users on application acceleration techniques, new programming models and algorithmic enhancement.

The post-docs generally spent a third to half of their time at OSF and the remainder of their time embedded in the research project area. Alternatively, if the research project area scientists were not located in commuting distance to the OSF, the post-docs traveled for periods of time (1–3 weeks) to work directly at the research project facilities.

An integral part of the program was to provide the post-docs a productive environment for working with the application code teams. Despite the variety of application areas, there was considerable overlap in the methods and tools for achieving petascale application performance. Through interaction with the NERSC staff as well as the Berkeley Lab computational science groups, the work produced a significant benefit to NERSC staff and users.

▲ Alice Koniges (third from left) led the Computational Science and Engineering Petascale Initiative, which paired post-doctoral researchers with high-impact projects at NERSC. Post-docs pictured are (from left) are Jihan Kim, Filipe Maia, Robert Preissl, Brian Austin, Wangyi "Bobby" Liu, Kirsten Fagnan and Praveen Narayanan. (Not pictured: Christos Kavouklis and Rebecca Yuan)

Here is a brief overview of each post-doc's research project(s) and achievements:

BRIAN AUSTIN worked with Jonathan Wurtele and Ji Qiang at Berkeley Lab's Center for Beam Physics to model a next-generation light source. He focused on improving the parallel performance of one of their simulation codes, IMPACT; very fast reduced physics models for the Next-Generation Light Source; and helped piece together different parts of MATLAB codes to create a complete image of the system rather than various components. He is now employed full-time at NERSC as a member of the Advanced Technologies Group.

KIRSTEN FAGNAN worked with John Bell in the Center for Computational Sciences and Engineering on a carbon sequestration/groundwater flow project. She focused on improving the Porous Media Adaptive Mesh Refinement (PMAMR) code developed at Berkeley Lab to model the movement of compressible and incompressible fluids. Fagnan is currently a bioinformatics consultant with DOE's Joint Genome Institute.

CHRISTOS KAVOUKLIS collaborated with Phillip Colella, leader of Berkeley Lab's Applied Numerical Algorithms Group, to develop a new version of the method of local corrections (MLC) code (developed by Colella) for the numerical solution of Poisson's equation on 3D structured grids. He continues to work with Colella's group in the Computational Research Division.

JIHAN KIM collaborated with UC Berkeley professors Martin Head-Gordon and Berend Smit on two carbon capture projects. In his work with Head-Gordon, he used graphics processing units (GPUs) to speed up some of the quantum chemistry routines that can be used to accurately compute the energy interactions between the carbon dioxide and the host carbon capture materials. In his work with Smit, he developed methods and software code to speed up the Monte Carlo code to efficiently characterize zeolite materials for carbon capture. Kim currently holds a faculty position in the Department of Chemical and Biomolecular Engineering at the Korean Advanced Institute of Science and Technology in South Korea.

WANGYI "BOBBY" LIU worked to improve the ALE-AMR code to model the Neutralized Drift Compression Experiment II (NDCX-II), an ARRA-funded induction accelerator commissioned in 2012. The NDCX-II will be used to study material in the warm dense matter regime, and in ion beam/hydrodynamic coupling experiments relevant to heavy ion-based inertial fusion energy. Lie focused on surface tension modeling and implemented two different surface tension models. Liu currently works for Google in Shanghai.

FILIPE MAIA worked on two projects during the petascale post-doc program—one assisting Stefano Marchesini of the Advanced Light Source in developing software for data analysis of ptychography experiments, the other converting a linear solver used in geological problems from running on CPU to GPUs. He is now on a tenure track at Upsala University in Sweden.

PRAVEEN NARAYANAN focused on performance characterization and benchmarking for HPC applications. His initial focus was on fusion applications on the Cray XT/XE platform at NERSC, followed by an examination of the user workload on the NERSC IBM iDataPlex platform (Carver) to carry out a machine-wide performance analysis. Narayanan is currently employed by NVIDIA.

ROBERT PREISSL focused on optimizing the Gyrokinetic Tokamak Simulation code, developed to study plasma microturbulence in tokamak reactors. Following the post-doc project, Priessl was employed at IBM and is currently involved with an HPC startup.

REBECCA YUAN worked with David Keyes from King Abdullah University of Science and Technology in Saudi Arabia and Columbia University and Stephen Jardin from the Princeton Plasma Physics Laboratory on the numerical simulation of four-field extended MHD equations in dynamically adaptive curvilinear coordinates. Yuan is currently doing data analysis for Bank of America.

Computational Systems and Facilities

Navigating the Road to NERSC-8

Throughout 2013, NERSC made significant headway in its efforts to unveil a new next-generation high performance computing (HPC) system by 2016.

Internally known as NERSC-8, this project is designed to support the rapidly increasing computational and storage demands of the entire spectrum of the DOE SC's computational research. In order to meet this goal, NERSC previously identified the need for an HPC system that would provide at least 10 times the sustained computing capability of Hopper, its Cray XE6 system.

In addition to increasing the computational capability available to DOE computational scientists, the NERSC-8 system will be instrumental in transitioning DOE scientific applications to more energy-efficient, manycore architectures. Energy-efficient computing is a cornerstone technology of exascale computing and represents the only way of continuing NERSC's historic performance growth in response to science needs.

Toward this end, NERSC achieved several project milestones in 2013:

- Completed a NERSC workload analysis and gathered system requirements from users
- Created an application benchmark suite comprising seven applications representing the NERSC workload that stress different components of HPC systems
- Launched an Application Readiness effort to prepare NERSC users for the new architecture
- Partnered with the Alliance for Computing at Extreme Scale (ACES), a collaboration between Los Alamos National Laboratory and Sandia National Laboratories, to release a joint Request for Proposals (RFP) for NERSC-8 and a second next-generation system, Trinity
- Evaluated vendors' proposals in response to the RFP
- Successfully negotiated and awarded a contract to Cray Inc., the HPC vendor
- Continued construction of the new Computational Research and Theory facility at the Berkeley Lab main campus, which will house the new system

As a result, in April 2014 NERSC finalized a $70 million contract with Cray to develop and deliver a new Cray XC system to NERSC in mid-2016. The system will be called Cori, in honor of biochemist and Nobel Laureate Gerty Cori, and will feature a number of new, cutting-edge technology features.

The Cori system will have over 9,300 Knights Landing compute nodes and provide over 400 gigabytes per second of I/O bandwidth and 28 petabytes of disk space. The Cray contract also includes an option for a "Burst Buffer," a layer of NVRAM that would move data more quickly between processor and disk, allowing users to make the most efficient use of the system while saving energy. The Cray XC system features the Aries high-performance interconnect linking the processors, which also increases efficiency.

The Knights Landing processor used in Cori will have over 60 cores, each with multiple hardware threads with improved single thread performance over the current generation Intel Xeon Phi co-processor. The Knights Landing processor is "self-hosted," meaning that it is not an accelerator or dependent on a host processor. With this model, users will be able to retain the MPI/OpenMP programming model they have been using on NERSC's previous generation Hopper and Edison systems. The Knights Landing processor also features on-package high bandwidth memory that can be used either as a cache or explicitly managed by the user.

As part of the Application Readiness effort, NERSC plans to create teams composed of NERSC principal investigators along with NERSC staff and newly hired postdoctoral researchers. Together they will ensure that applications and software running on Cori are ready to produce important research results for the Office of Science. NERSC also plans to work closely with Cray, Intel, DOE laboratories and other members of the HPC community who are facing the same transition to manycore architectures.

"We are committed to helping our users, who represent the broad scientific workload of the DOE Office of Science community, make the transition to manycore architectures," said Katie Antypas, NERSC's Services Department Head. "We recognize some applications may need significant optimization to achieve high performance on the Knights Landing processor. Our goal is to enable performance that is portable across systems and will be sustained in future supercomputing architectures."

NERSC Users Take Edison For a Test Drive

Beginning in early 2013, NERSC users were instrumental in testing the center's newest flagship supercomputer: Edison, a Cray XC30 named in honor of American inventor Thomas Alva Edison.

Edison can execute nearly 2.6 quadrillion floating-point operations per second at peak theoretical speed. While theoretical speed is impressive, NERSC's longstanding approach is to evaluate proposed systems by how well they meet the needs of its diverse community of researchers.

"We focus on sustained performance on real applications," said NERSC Division Deputy for Operations Jeff Broughton, who led the Edison procurement team.

Edison was also configured to handle multiple kinds of computing equally well: simulation and modeling and data analysis, both of which rely heavily on moving data.

"Edison has been optimized for this," said Sudip Dosanjh, director of NERSC. "It has a really high-speed interconnect, lots of memory bandwidth, lots of memory per node, and very high input/output speeds to the file system and disk system."

Because Edison does not employ accelerators, such as GPUs, scientists have been able to move their codes from NERSC's previous flagship system, Hopper, to Edison with few or no changes, another consideration meant to keep them doing science instead of rewriting code.

Historically, NERSC was an exporter of data as scientists ran large-scale simulations and then moved that data to other sites. But with the growth of experimental data coming from other sites, NERSC is now a net importer, taking in a petabyte of data each month in fields such ranging from biosciences and climate to high-energy physics and cosmology.

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

Computational Research and Theory Facility Nears Completion

Construction of the new Computational Research and Theory (CRT) facility on the Berkeley Lab main campus was ahead of schedule during 2013, aided by unseasonably balmy weather throughout the fall and into early 2014.

The new CRT facility is slated to be completed in early 2015.

On Dec. 6, 2013, the final steel girder of the CRT was hoisted into place in a "topping off" ceremony that marks the beginning of the end of the heavy construction phase. As tradition dictates, the beam was decorated with an evergreen and American flag and signed by both the builders and some of the eventual occupants, including Kathy Yelick, Associate Laboratory Director for Computing Sciences, and Division Directors Sudip Dosanjh (NERSC), David Brown (Computational Research) and Greg Bell (Scientific Networking).

CRT is designed to provide the kind of world-class energy efficiency necessary to support exascale systems. It will be a highly energy-efficient, state-of-the-art computing facility that can provide over 40 MW of power and 30,000 square feet of space for computing and data storage. The additional power and space within CRT will allow NERSC to deploy pre-exascale and exascale systems that are usable by the broad scientific community, and to meet the exponentially growing data needs of DOE.

When construction is complete, CRT will house all three divisions of Berkeley Lab's Computing Sciences—NERSC, ESnet and the Computational Research Division. NERSC has been located at the University of California Oakland Scientific Center in downtown Oakland since 2001.

▲ The new Computational Research and Theory facility, currently under construction at Lawrence Berkeley National Laboratory, is slated to be completed in early 2015.
ROY KALTSCHMIDT

Global Scratch Gets an Upgrade

The most-used file system at NERSC—global scratch—went through a major upgrade in 2013. As a result, some users' data output to global scratch can now reach up to 80 gigabytes per second. Although users will probably not see their 20-terabyte storage quotas increase, the upgrade ensures that global scratch remains flexible and paves the way for the Parallel Distributed Systems Facility (PDSF) to eventually use the file system.

Because of the upgrade, users will also be able to better access their temporary data files or "scratch data" from any NERSC system, not just the one that generated it. Prior to the upgrade, global scratch typically operated at over 90 percent capacity with data input and output rates around 15 gigabytes per second.

"Utilization of the existing global scratch file system is extremely high," said Jason Hick, who leads NERSC's Storage Systems Group. "The consolidated file system is popular because users can store, analyze and refactor data at high bandwidth from a variety of different systems at the facility without the hassle of transferring it between systems."

By upgrading the consolidated global scratch, users can continue taking advantage of an efficient and scalable storage resource for their scientific storage needs.

"We can slice data differently for various purposes without concern over bandwidth or latency," Hick said. "Efficiency was our key metric with adopting a site-wide storage architecture. By optimizing storage for different requirements, such as large- and small-scale simulations, visualizations or analytics, we could offer our community the most efficient, scalable storage resources possible."

In choosing an embedded storage solution, NERSC has eliminated the need for additional servers, cabling, network switches and adapters, which reduced administrative overhead by hundreds of thousands of dollars.

"NERSC was a pioneer in moving to a site-wide file system architecture and recently for moving toward a consolidated storage architecture," Hick said. "We recognized that centralization could yield substantial storage and network performance improvements while offering us a much simpler, cost-effective approach to deploying HPC resources."

NERSC Now Connecting All Science at 100 Gbps

As of mid-2013, all network traffic flowing in and out of NERSC is now moving at 100 Gigabits per second (Gbps). This includes everything from email to massive scientific datasets. At this speed, 1.8 million people could simultaneously download an eBook in about two minutes.

The impetus for this move comes from an increasing demand for bandwidth from scientific users of the facility. It was made possible through of a redesign of NERSC's border monitoring system (Bro) by engineers in the center's Network and Security Team. Bro can now monitor paths and load-balance the 100 Gbps connection across an array of intrusion detection systems, then automatically terminate and block bad actors.

NERSC is the first computing facility to implement a 100 Gbps security system, and this may serve as a model for other DOE facilities to do the same.

"Before we could move all NERSC traffic to 100 Gbps, we had to make sure that we could securely monitor all the data streaming in at this speed," said Jason Lee, who leads NERSC's Network and Security Team. "Security was our number one requirement, and until recently nobody knew how to do this."

Once NERSC decided to move all of its traffic to 100 Gbps, Lee worked with engineers at the DOE's Energy Sciences Network (ESnet) to set up a 100 Gbps Science DMZ, which gives NERSC network engineers the ability to set up multiple private circuits using software-defined networking (SDN). With these tools, NERSC staff can help remote scientists who may see their data transfers slow down due to firewalls at their local campuses achieve a true 100 Gbps end-to-end connection.

ESnet engineers also helped NERSC set up a system to announce their own address space. This allows the center to separately route traffic to any research and education (R&E) site or separate R&E traffic from the commodity Internet.

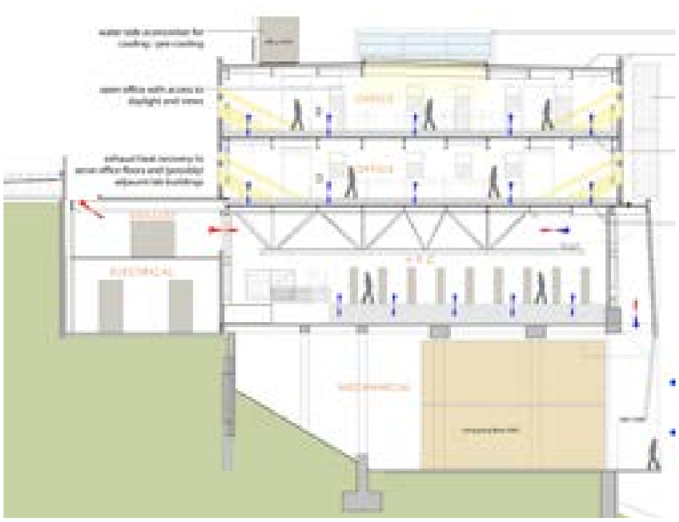
Innovations

Preparing for Exascale: Energy Efficiency

Edison Free Cooling

With Edison—NERSC's newest flagship computer—NERSC has leveraged the Cray XC30's unique water-cooling mechanism and the temperate San Francisco Bay Area climate to eliminate the need for chillers and achieve excellent energy efficiency, with a power usage effectiveness (PUE) ~1.1 for Edison alone.

NERSC has been located at the University of California Oakland Scientific Center (OSF) since 2001. Edison was installed at OSF in 2013 and was targeted to occupy the floor space previously occupied by Franklin, a retired Cray XT4. But the two computer systems have very different cooling requirements, prompting NERSC to investigate a new cooling plant design for Edison. The resulting "free cooling" approach also took into account the unique environmental features of the new Computational Research and Theory (CRT) facility being built at Berkeley Lab, where NERSC and Edison will relocate in 2015 (Figure 1).

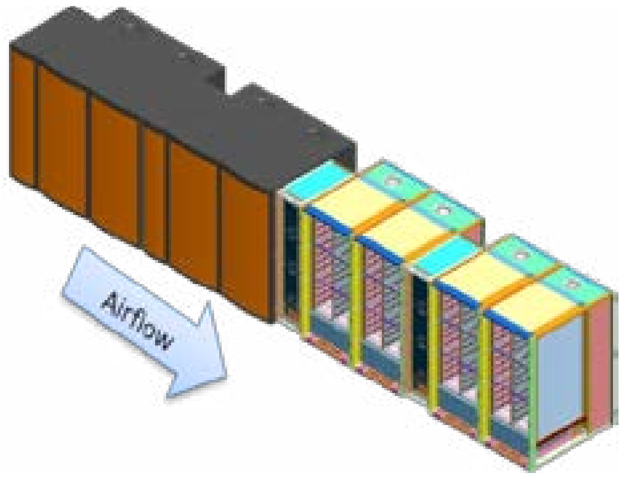


▶ **Figure 1.**
A cross-section of the CRT data center, opening in 2015 on the Berkeley Lab main campus. The bottom floor houses air handlers, pumps, and heat exchangers. The second floor is the data center; the top two floors are offices.

Many data centers utilize free cooling when the conditions are amenable, and fall back on chillers when temperatures and humidity rise. In the San Francisco Bay Area, conditions are favorable all year.

Temperatures remain relatively cool year-round, and on those occasions when they rise above the 70s, humidity stays low. Since the OSF and CRT sites share the same climate, implementing free cooling at OSF is a way to prototype the CRT design and better understand system requirements.

At OSF, Franklin was cooled by a set of chillers and cooling towers in the mechanical yard adjacent to the computer floor. But the Cray XC30 cooling mechanism is designed to operate with much warmer water temperatures than previous supercomputers. It does this by using a hybrid air/water cooling mechanism. Airflow is side-to-side—where the exhaust air from one rack becomes the intake air of the next—rather than a more typical front-to-back airflow (Figure 2). Each cabinet contains a water intercooler (radiator) on the outlet side. This transfers heat from the air blowing through the cabinets to the water loop and cools air delivered to the next cabinet in the row. Fan cabinets are placed at the intake and outlet ends of the row, as well as interspersed between pairs of two compute racks.



▲ **Figure 2.**
Diagram of a Cray XC30. Air flow is through the cabinets from upper left to lower right.

Flowing the air from side to side has unique advantages. First, additional cabinets in a row do not require more air from the computer room than was supplied to the first rack. Lower overall air requirements mean fewer building air handlers and greater overall center efficiency.

The use of free cooling with Edison at OSF has resulted in positive improvements in energy utilization. Prior to the Edison installation, only Hopper and other smaller systems were operating using mechanical chillers. At that time, the center PUE was approximately 1.31, meaning that the energy to cool the systems was 31 percent of the energy used to power the systems. After installation of the free cooling plant and Edison, the PUE dropped to 1.19. This is a blend of the Edison standalone PUE and that for the existing systems.

In addition to the free cooling improvements, air temperatures within the machine room have been raised, computational power efficiency has been improved by installing updated processors and power required for cooling has been reduced. As a result of these improvements, power savings for cooling have been calculated at almost 2 million kWh per year, compared with using a traditional chiller plant. NERSC also received a rebate of $435,000 for the CRT through Pacific Gas & Electric's energy efficiency project rebate program.

Preparing for Exascale: New Computational Models

Babbage and the Early Application Readiness Effort

In the NERSC-8 timeframe (2016), it is expected that all vendor offerings will require modification of user application codes in order to achieve high performance. The amount of code modification and restructuring will depend on the particular architecture chosen and will likely require applications to expose finer levels of parallelism and increased vector lengths. In addition, some architectures may require applications to transition to an offload programming model.

Because NERSC had previously studied application performance on GPUs using the Dirac cluster and because DOE users have experience with the GPU HPC configurations at the Oak Ridge Leadership Computing facility, NERSC focused early application readiness efforts on evaluating application performance on the Intel MIC architecture. A 45-node Intel Xeon Phi (Knights Corner or KNC) test system named "Babbage" was purchased and configured to allow NERSC staff and early users to test, evaluate and port their application codes onto the MIC architecture and gain firsthand knowledge of exposing more loop-level parallelism and vectorization, which are essential for optimal performance on this architecture. Each Babbage node is configured with two KNC cards and two Intel Xeon Sandy Bridge "host" processors. Each KNC card has 8GB of memory and 60 cores, each with four hardware threads. The Intel Parallel Studio XE software suite provides Intel compilers, MPI, and performance tools. Intel provided NERSC staff with training on this new software package in October 2013.

Much effort was put into configuring Babbage with ease of use in mind, including:

- Mounting all production file systems
- Providing seamless access to NERSC's HPSS archival system
- Configuring SSH to allows password-less access to the host nodes and MIC cards smoothly
- Deploying the Intel Parallel Studio XE software suite
- Creating a modules environment to allow users to quickly switch between software versions
- Providing a default environment that does not require users to copy system libraries to each MIC card manually at runtime as pre-steps for running jobs
- Providing Web documentation for running and optimizing on the MIC cards

The Babbage test system was used to study representative applications and kernels in various scientific fields to gain experience with the challenges and strategies needed to optimize code performance on the MIC architecture. Below are the findings of one such case study, involving the BerkeleyGW application.

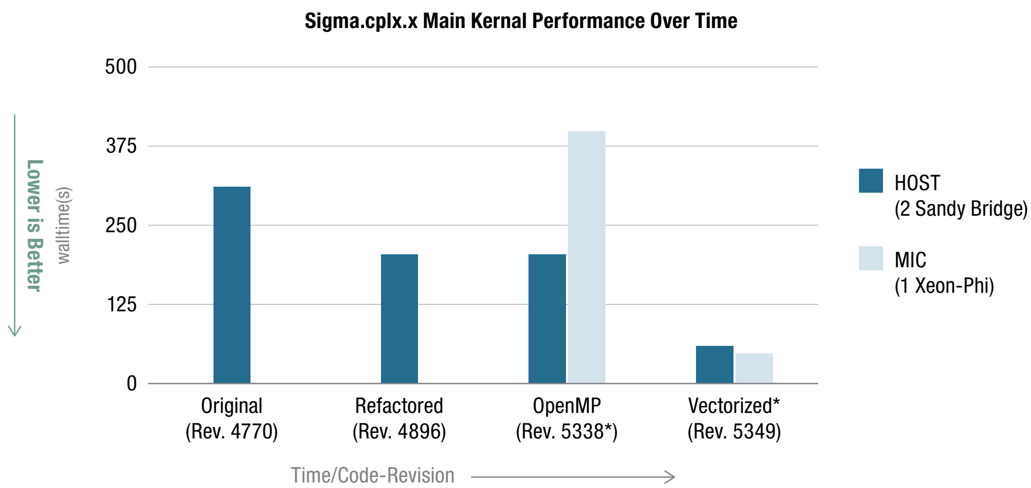
The BerkeleyGW package is a materials science application that calculates electronic and optical properties with quantitative accuracy, a critical need in materials design for more efficient and cost-effective solar light-harvesting and energy conversion. Its use at NERSC has grown such that it is now one of the top 20 codes used by NERSC users based on wall-clock hours.

GW calculations within BerkeleyGW occur across two executables. The first depends heavily on dense linear algebra and FFT math libraries, while the second depends on custom code that expresses large summations in terms of reduction loops. NERSC focused predominantly on optimizing the performance of the main kernel in the custom-coded executable.

The original code (revision 4770 in Figure 3) was an MPI-only code. Using CPU cores in an MPI-only model, we were unable to fit our sample problem into memory on the Xeon Phi. The first step taken during optimization (Rev. 4896) was to refactor the code to have a three-loop form: a large loop to be parallelized with MPI tasks, a large loop to be parallelized with OpenMP threads and a large inner loop that can be vectorized.

Next, the appropriate OMP reduction pragma was added to the loops targeted for OpenMP. While the code scaled well to tens of threads, it still performed poorly on the Xeon Phi because the compiler was not vectorizing the innermost loops due to loop dependency. Simplifying the flow (splitting into multiple loops in some cases) ensured that critical inner loops were vectorized by the compiler. After this optimization, the performance on the Xeon Phi was found to be comparable to the host Sandy Bridge CPUs for the test example. It should be noted that the example wall-clock time decrease between revision 5338 and 5349 in Figure 3 is not entirely attributable to vectorization, but a combination of vectorization, code simplification and restructuring of loops.

This BerkeleyGW kernel example is, in many ways, an ideal case for the Xeon Phi: reduction loops can benefit from the increased memory bandwidth and the code was able to be restructured in a way that left very long inner loops (1,000-10,000 iterations), which is ideal for vectorization. While this is fortuitous, not all codes can be structured in such a way.



▶ **Figure 3.**
BerkeleyGW kernel performance over time.

The Babbage test cases show that in order to achieve performance on the Xeon Phi Knight's Corner architecture, developers had to improve OpenMP scalability and vectorization capabilities by restructuring the application or the data structures. Going forward, the Application Readiness effort will ramp up as a broad, multi-pronged effort that includes user training, access to early development systems, application analysis deep dives and collaborations with vendors and other members of the HPC community who are facing the same transition to manycore architectures. Several NERSC top application codes run on the Leadership Computing Facilities (LCFs) at Oak Ridge and Argonne, and many of them are already well prepared for the NERSC-8 system. Other application teams will need more guidance and assistance to make the transition. NERSC will leverage best practices from the LCFs, ACES and others in the HPC community to prepare its users for the NERSC-8 architecture. NERSC intends to proactively share its experiences with its ACES partners (Sandia, LANL) and with the ALCF and OLCF. NERSC is already working closely with the ALCF and OLCF to develop an SC Applications Readiness plan to ensure that efforts at one site are not unnecessarily duplicated at another.

MiniDFT Benchmark and Innovations in SSP for NERSC-8

The diversity of technology options in the exascale timeframe is too large to be explored using legacy scientific codes; these codes are often large and have many features that obscure their underlying computational patterns. This complexity increases the code's scientific value but also increases the challenges of restructuring the code to expose additional parallelism or harness the capabilities of new programming models or hardware features.

Miniature applications ("mini-apps") that capture the algorithmic and performance characteristics of a real code but are far simpler to understand and modify are gaining traction as tools to rapidly explore software transformations.

NERSC has developed the Mini-DFT mini-app to represent plane-wave density functional theory (PW-DFT) software. Mini-DFT fills a significant gap in the suite of mini-apps and helps an important class of codes prepare for exascale computing. A recent analysis of the NERSC workload, which was conducted as part of the NERSC-8 procurement, showed that PW-DFT accounts for more than 10 percent of the computing resources used at NERSC, and several high profile projects such as the Materials Genome Initiative (http://www.nist.gov/mgi) rely heavily on DFT.

Furthermore, other existing mini-apps do not exercise the dense linear algebra and 3D-FFT algorithms that are central to PW-DFT codes. Mini-DFT was developed by extracting the essential routines from the open-source Quantum Espresso package. The resulting code has 90 percent fewer lines of code than Quantum Espresso, but maintains significant overlap with the original application so that the most beneficial transformations can be adopted with relative ease.

The development of Mini-DFT had immediate and ongoing value. Mini-DFT was one of the application performance benchmarks used for the NERSC-8 procurement, and vendors in the DesignForward project continue to study Mini-DFT in order to understand the performance requirements of future interconnects.

Using Manycore for Hybrid Software-Based Rendering

With the transition to manycore architectures expected in the NERSC-8 time frame, effective use of threading will be key to achieving high performance. This in turn creates a need to address the challenges visualization codes face on manycore architectures.

To answer these challenges requires understanding the specifics of what GPUs provide to key libraries in the HPC rendering stack, as well as all capabilities currently provided in the rapidly evolving open source OpenGL libraries that VTK and other 3D rendering applications depend on for software-based rendering on HPC systems. There is also a need to understand the current and future support, if any, for manycore architectures in these libraries.

For systems without GPUs, Mesa3D is typically used. Its new capabilities provide support for threading and native vectorization of runtime compiled shader code. NERSC computational scientists found a number of issues in VTK preventing the effective use of new software rendering capabilities in Mesa3D. Fortunately, they were able to resolve all of the issues in VTK and submit patches that were adopted for VTK 6.0 and 6.1. They also ported a number of GPU rendering algorithms for use on Edison and investigated the parallel MPI+threads rendering performance on a large PIC plasma simulation dataset provided by a NERSC user as a test case for these efforts.

The early results show that VTK-based applications such as ParaView that incorporate new rendering software known as Gallium provide increased capabilities and performance that approaches GPU-based rendering. In addition, limiting the total number of rendering threads per node to the number of available cores, including hyper-threads if they are available, produces a very reasonable result that scales for large datasets.

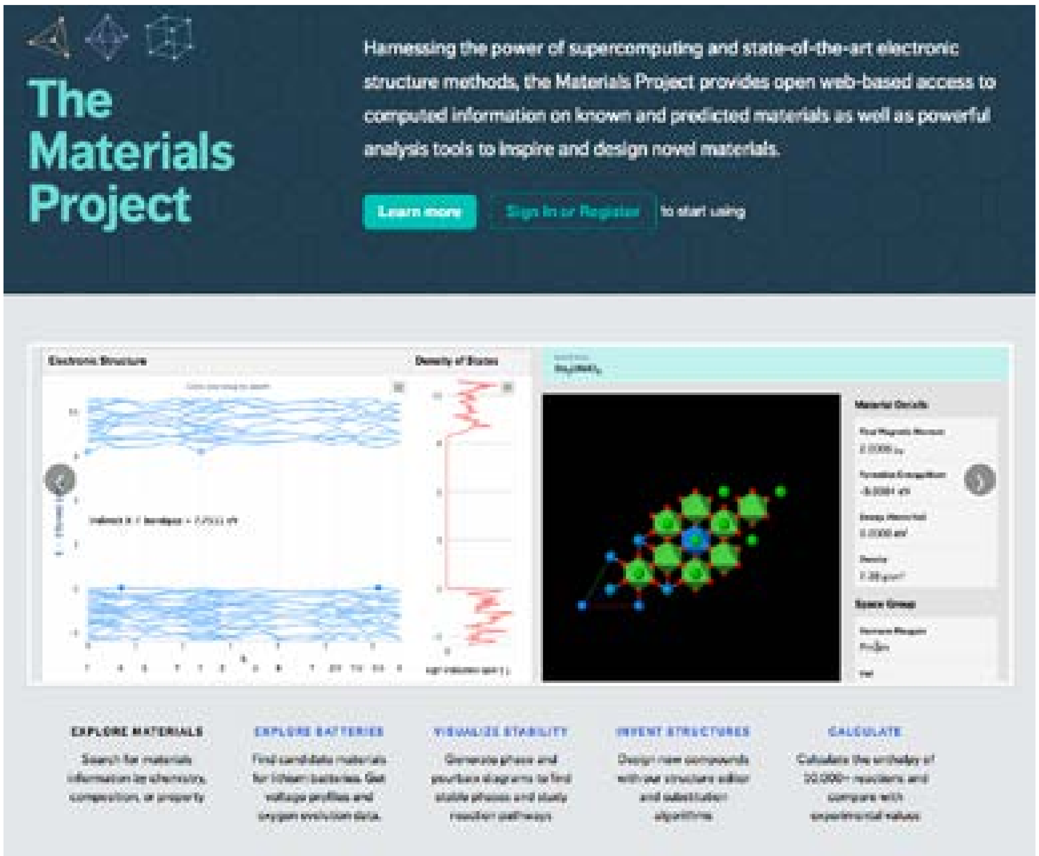
Improving the User Experience

NERSC's Science Gateways

MATERIALS PROJECT: GATEWAY + MONGODB + HTC WORKFLOW

The Materials Project is an open, web-hosted service at NERSC that allows scientists to design new materials atom by atom, before ever running an experiment (Figure 4). It has proven so popular that Scientific American featured this project on its December cover as a top world-changing idea of 2013.

NERSC serves as the computing and data engine for the project. In addition to supporting calculations on supercomputers, the NERSC division maintains cluster nodes purchased by and dedicated solely to the Materials Project. It provides the software and hardware infrastructure for the web gateway and databases that serve up the Materials Project data.



▶ **Figure 4.**
Materials Project gateway screen capture.

The Materials Project requires thousands of calculations, with each running on a handful of processors (this is sometimes called "high-throughput" or "ensemble" computing). The system for submitting calculations in most supercomputing environments is geared toward fewer, larger jobs. The "Fireworks" workflow manager, developed by Anubhav Jain from the Materials Project, adds intelligence and automation to the process of submitting large batches of high-throughput jobs. And because the framework was written to be general purpose, other projects are now considering using it for research that requires an ensemble-computing model.

The Materials Project data is stored in the MongoDB NoSQL Database infrastructure deployed at NERSC. Traditional SQL is focused on upfront design; first you come up with a schema, then you enter the data. In science, however, you rarely know everything you're going to do with the data before you start. The MongoDB model lets researchers add new data and then generates its structure (or schema) from that. This technology is used as the workhorse of the Materials Project to schedule and track quantum mechanical calculations of materials properties on supercomputers, to store and search the results of these computations and to perform advanced analytics on the computed materials properties.

NERSC has now made the MongoDB infrastructure more generally available to all its users who need access to a scalable database that supports flexible schemas. We are already seeing increasing use of MongoDB from JGI, ALS and DESSN users,who have evolving data models.

NEXTGEN BIOIMAGING

Exponential data scaling in bioimaging, combined with the data fusion demands of increasingly multi-modal imaging approaches, require R&D in computational bioimaging solutions that extend the ways in which scientists can leverage image-based data. Bioimaging is leaving an era of images-as-results and entering an era where results come in the form of biological models built on massive gigapixel images and vast image collections. These models require scalable analysis, integration and dissemination of image data delivered from advanced bioimaging instrumentation.

NERSC's NextGen Bioimaging (NGBI) project prototypes big data bioimaging methods in the context of biofilms, also known as microbial communities, and the Protein Atlas fluorescence microscopy project. Biofilms are important to DOE due to their dominance of microbial lifestyle and their significance in microbial mediated processes. Protein expression and its spatial localizations studied by high-throughput automation promise a detailed understanding of the proteome. NGBI is co-designing a modular, web-based computing and data strategy with these two projects and plans to extend this capability to other projects in Big Data bioimaging.

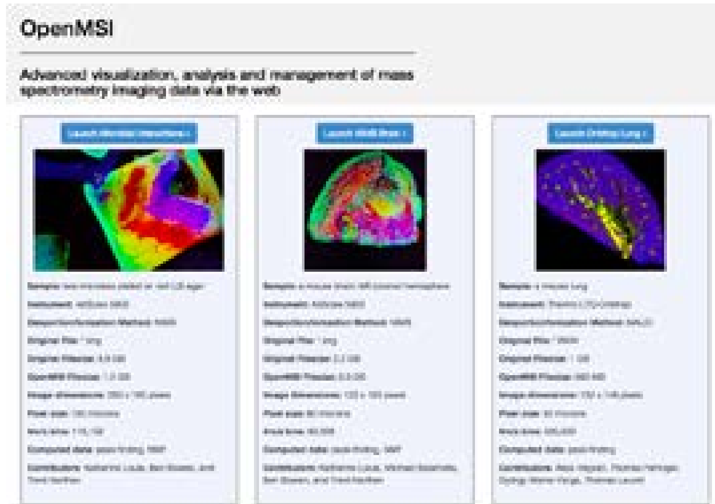
The NGBI team has deployed custom-tailored image-processing algorithms and successfully integrated a suite of universally used processing software among the biology community into an OMERO-based science gateway using NERSC's HPC resources. This shared, scalable web service provides a modular and flexible one-stop-shop solution for producers and consumers of models built on imaging data by refining pixel data into actionable knowledge resources. As a result, NERSC offers a platform that has multi-modal capabilities and is Big Data-ready.

To date, more than 30 LBNL and University of California, San Francisco scientists have adopted the gateway, many to manage electron and fluorescence microscopy workflows.

OPENMSI MASS SPECTROMETRY IMAGING

Mass spectrometry imaging (MSI) plays a key role in the study of the metabolism and energetics of cells and cellular communities. Metabolite imaging and metabolomic analysis are used to map the connectivity, dynamics and spatial topography of complex cellular systems. In contrast to optical imaging, using MSI allows the study of large numbers of metabolic compounds in a single image.

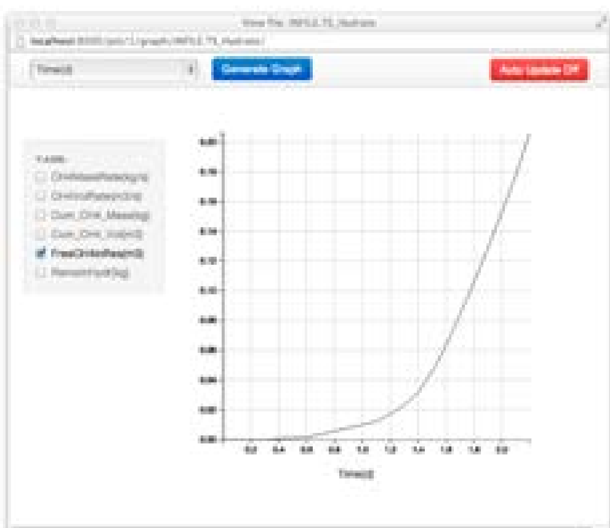
In 2013 the OpenMSI project made its first public release of the OpenMSI science gateway that provides interactive access to MSI data stored at NERSC (Figure 5). The OpenMSI project has gained broad attention and already serves ~20 alpha testers and hosts ~1TB of MSI data. The OpenMSI team has published its results in the journal ACS Analytical Chemistry and has made a range of presentations, e.g., as part of the LBNL Science At the Theatre lecture series and BioImaging seminar series.



▶ **Figure 5.**
The OpenMSI science gateway, hosted by NERSC.

NEW GATEWAY: TOUGH EARTH SCIENCE CODES

Members of the Outreach Software and Programming group have created a new science gateway for users of the TOUGH family of earth science codes (Figure 6).



▶ **Figure 6.**
The TOUGH science gateway.

TOUGH is used for modeling and simulation of fluid flow through fractured and porous media. The software offers capabilities that appeal to nontraditional HPC users, many of whom have never attempted to submit a parallel computing job. The new gateway enables users to create, submit, manage and watch real-time output from TOUGH jobs running on Hopper, all via an interactive graphical interface. Allowing submission via the web enables users who might otherwise be limited to running the code on desktop systems to take advantage of NERSC's much speedier computational resources, attacking more complex problems than they would otherwise be able to address.

Workflow Enhancements

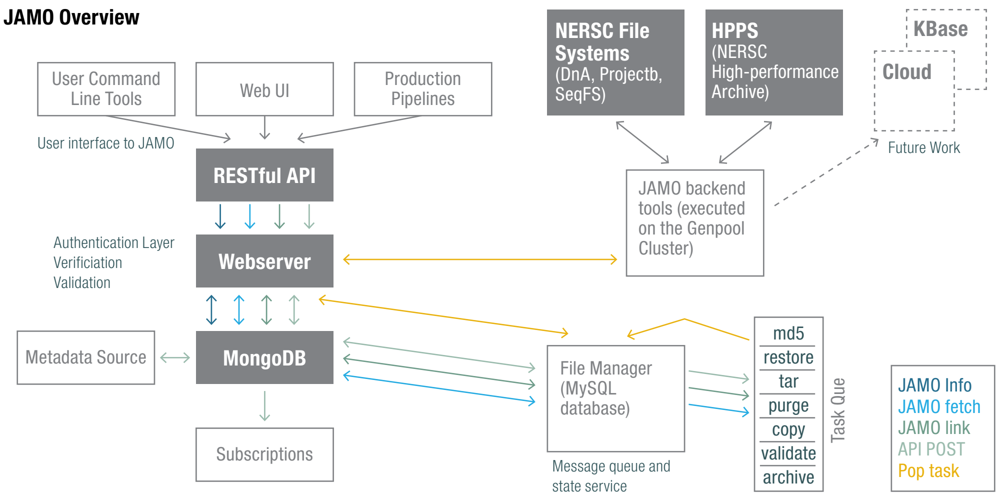
JGI ARCHIVE AND METADATA ORGANIZER

The JGI Archive and Metadata Organizer (JAMO), which came online at NERSC in August 2013, is a centralized data management system designed to help researchers at the Joint Genome Institute (JGI) deal more efficiently and effectively with the vast amounts of data yielded by increasingly complex bioinformatics projects.

JAMO began as a collaboration between the Sequence Data Management, Quality Assurance and Quality Control and Genome Assembly groups at the JGI. These groups recognized the need for a centralized data management system for the data and analysis generated at the JGI. The collaboration also included support staff from NERSC to ensure that the design would make effective use of the center's storage resources.

JAMO is JGI's first implementation of a hierarchical data management system (Figure 7). Users can register important data with JAMO and set an expiration policy. JAMO then migrates the data to the archive and copies it to the DnA (Data n' Archive) file system where the data can be read but not modified. If a user would like to modify the file, they must register a new file with the system, and it will be treated as a new version of the same file. JAMO also tracks how long the data is kept on disk. Once the expiration policy is hit, the file is deleted from spinning disk and only exists in NERSC's High Performance Storage System. If a user requests a file that is only in the archive, JAMO will automatically restore that file to DnA.

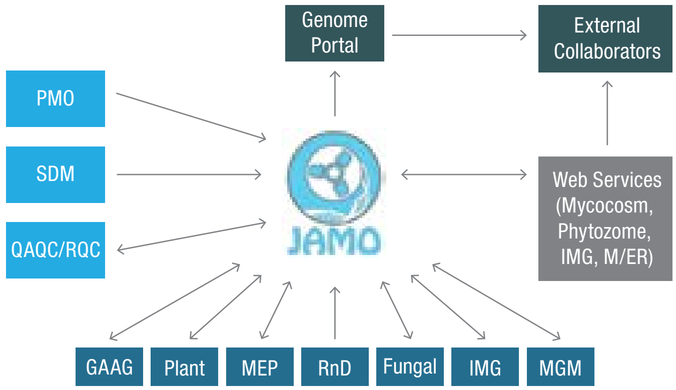
JAMO Overview



▶ Figure 7. JAMO is JGI's first implementation of a hierarchical data management system.

While sequence data is being prepared for JGI's users, it needs to be shared between most groups at JGI because each is responsible for a different component of the analysis. Prior to JAMO, data sharing was accomplished through a complex system of Google documents, emails and databases spread out across the organization. This strategy meant that as the data aged and its location was forgotten it would take longer and longer to find.

With JAMO now in place, JGI is moving toward a system in which each group queries a centralized web service to locate the data needed for calculations or for download by the JGI's users (Figure 8).



▶ **Figure 8.** JGI's new data management scheme, utilizing JAMO.

The feedback so far has been overwhelmingly positive, according to Kirsten Fagnan, bioinformatics and HPC consultant in the User Services Group at NERSC.

"The JGI is excited about being able to find old data in minutes as opposed to hours," she said. "We are also looking forward to using the JAMO system to generate reproducible pipelines and workflows by storing detailed metadata with each project analysts complete. We have made this project feel collaborative, which has provided momentum at JGI for doing software development projects right."

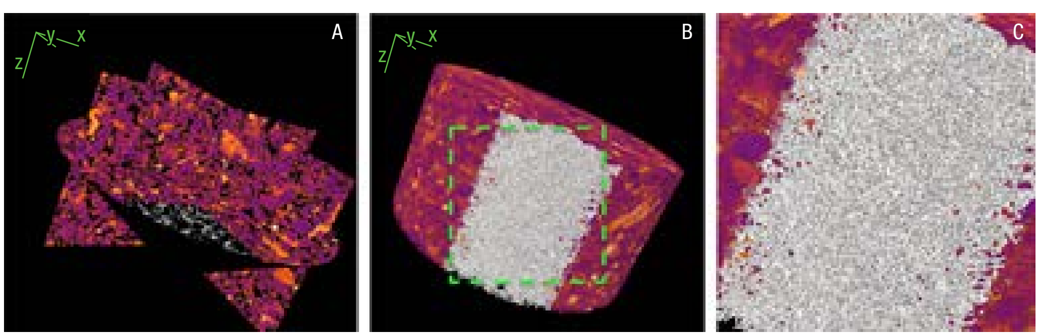
DATA ANALYSIS FOR IMAGING FACILITIES

A key object for imaging facilities is the development of data analysis algorithms, workflows and new software tailored to process image stacks (volumetric raster data) from experimental data. The employment of several computer vision algorithms, including mathematical morphology, anisotropic filters and graph-based algorithms for classification, is necessary to analyze these data sets.

NERSC data analytics staff created solutions by developing ImageJ/Fiji plugins and macros that the science teams being helped incorporated into their workflows. These new image-processing workflows reduced the time required to process tens of gigabytes of data from days to approximately 20 minutes per stack.

Part of this work involved analyzing porous material (Figure 9). Each dataset (~10GB) consisted of microCT data, and NERSC's contribution was to create an analysis prototype supporting beamline facilities, which is available in source-forge and at NERSC.

In support of the "Real-time tomography of biomass deconstruction" project being conducted at the Advanced Light Source (ALS), NERSC provided techniques to analyze breast cancer cell images. Their biomass was imaged at the ALS producing small datasets of large-sized files (~1TB of focused ion beam data). NERSC created a tool that uses hybrid architectures and evaluated the performance on GPU cards to segment microtubules.



▶ **Figure 9.** Segmentation of X-ray micro-CT to identify grains and channels extracted from soil sample.

PREEMPTIBLE QUEUE

The increasing need for NERSC to support a variety of workloads, including those that have real-time constraints, has prompted the facility to employ preemptible queues. These queues allow users to voluntarily allow jobs with more flexible turnaround to be preempted by higher-priority jobs that require rapid turnaround. Preempted jobs may be subject to a reduced charging factor or other incentives to encourage users to make use of this queue. Higher-priority jobs typically start within a few minutes.

NERSC uses preemption in the Carver cluster to allow jobs submitted to the preemptible queue to run on idle dedicated resources. Jobs submitted to this queue can run on any of the nodes in the cluster. If a preemptible job has started and been allocated some of the dedicated resources, it will be cancelled by the batch system if another job is submitted to the dedicated workflow requiring those resources. For users, jobs submitted to the preemptible batch queue normally start more quickly; only those jobs can take advantage of unused dedicated resources.

The center benefits from increased utilization and throughput. For Carver, NERSC anticipates a 4-5 percent utilization increase from our average 90 percent value once this queue becomes familiar to the user community. The dedicated workflow is only delayed by the time it takes the batch system to cancel the preempted job, typically requiring less than one 10-second scheduling cycle.

RSTUDIO SERVICE

R is an open-source data analytics toolset with a rich collection of data analytics packages that can be used to perform statistical analysis and machine learning on large datasets. R is very popular among genomics and biology communities. However, running R on NERSC systems requires some knowledge of the HPC environment, and a number of R users are not sophisticated HPC users. Consequently, usability is the major roadblock for its adoption at NERSC.

▶ **Figure 10.**
Screenshot of a working RStudio session.

To solve the usability problem, NERSC staff deployed the RStudio service. RStudio is a web-based user interface for R. Any active NERSC user can go to http://r.nersc.gov and login with his/her NIM username and password and start using R in a fully graphical interface (Figure 10).

The RStudio service is presently in a beta stage and initially was announced to a small subset of NERSC users (about 500). As of early 2014, some 150 users had used the service and 10-20 users were using it extensively. In 2014, NERSC plans to upgrade RStudio to a production-grade service and announce it to all NERSC users.

NOMACHINE NX SERVICE

NoMachine NX is a virtual desktop service that provides users with an uninterrupted desktop at NERSC. The most significant benefit of NX is the ability to keep a session alive even when the client disconnects from it. This allows remote users to disconnect/reconnect to sessions without interrupting the applications running in the session (e.g., all SSH connections will be maintained within the session).

NERSC started providing centralized NX service in 2011, but the old system had an outdated window manager and was not user friendly. In 2013, NERSC upgraded the NX service to modern hardware and started supporting the KDE window manager. The usability of the service is greatly improved. As of January 2014, more than 700 unique users have used NX and on any workday over 100 unique users are using the service.

GIVE + TAKE

For many years, NERSC has supported file sharing among users via project directories and appropriately set ownership and permission bits. However, the task of creating project directories and adding all parties to the directory group is fairly involved for use cases when a user simply wants to share a small number of files with a colleague. In the past, users have worked around this by making their personal directories world readable/executable and often writeable, which can lead to unintended consequences such as broad visibility of files or accidental file deletion.

To address the use case of casual sharing of small numbers of files, NERSC has developed and deployed a “Give and Take” system to enable users to quickly and efficiently share files with individual users. Unlike other implementations of Give and Take, NERSC’s implementation does not require any user-privilege escalation, such as setuid scripts, which create additional security risks. The implementation, which relies on filesystem ACLs, is still able to validate the sender’s ID and ensures only the intended recipient can access (take) the shared files.

Give and Take consists of two commands (give and take) that allow users to give individual files or directories. Given files are copied to a staging area that is part of the NERSC Global File System. The recipient is notified via email that a file is available to take. The take command is used to move the files to a desired location.

In some cases, NERSC users wish to be able to receive files from non-NERSC users or data stored outside of NERSC. The NERSC FTP Upload service is designed for external collaborators to be able to send data to NERSC staff and users. It allows collaborators to create a temporary FTP account to upload files that will be delivered to a NERSC user.

MYNERSC AND MOBILE NERSC

NERSC continues to enhance user access to real-time job, usage and status information. In 2013 the center redesigned its mobile application (m.nersc.gov) and added functionality to MyNERSC to improve the user experience and re-use common code between the mobile and desktop portal.

One improvement NERSC made in direct response to a user request was to track and show in these portals the rank of users’ jobs in the queue over time. Users have long asked for the ability to track their job’s progress in the queue and predict start-times. This tool provides an accurate visual way to accomplish this. In many cases, it is more reliable in estimating job start times than the showstart command provided by the scheduler.

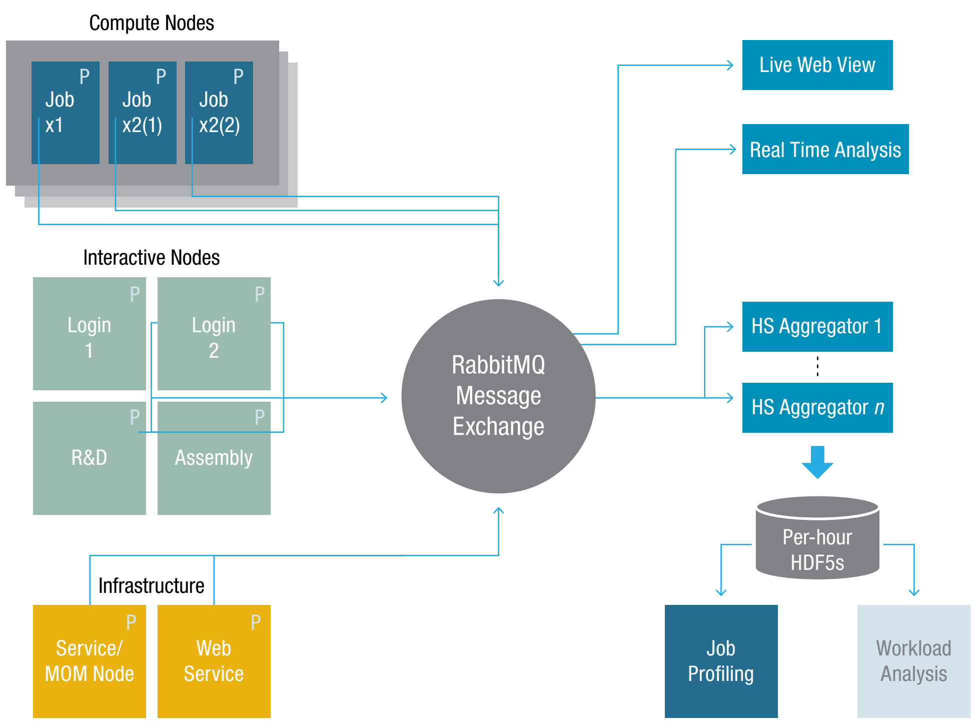
In addition, NERSC's past, current and future outages are now programmatically inserted into Google Calendar for many users to subscribe to. This was a direct response to a user request on the NERSC annual user survey.

PROCMON

With today's increasingly data-intensive workloads, HPC centers often allow multiple different jobs to occupy the same compute nodes, thus increasing utilization of resources and throughput. This sharing of resources can lead to new challenges, including more complex scheduling of jobs, complex aggregate I/O patterns on the compute nodes and potential resource conflicts between jobs running on the same node.

To begin to better characterize this complexity, NERSC developed the procmon system. Procmon closely monitors what is running on each compute node, the resources consumed by each process and batch-job context for each process. Procmon records all these data every 30 seconds, enabling a time-series record of all the sampled processes in the system. Procmon is designed to be extremely scalable, while still collecting large amounts of useful information. Figure 11 shows the overall design of the procmon system and how it operates on different classes of nodes.

Procmon uses a next-generation messaging system, RabbitMQ, to enable communication between the sensors and the data collectors. RabbitMQ also allows messages to be selectively duplicated, enabling targeted, live analysis of data in addition to the aggregate consumption for logging of all process data. Procmon also features a set of programs that multiplex the data from the RabbitMQ messaging system, reduce it and store a minimal copy in per-hour HDF5 files for later analysis. This approach enables more than 4,000 simultaneously connected procmon sensors to have their data reliably logged without incurring any noticeable impact on the network or filesystem.



▶ **Figure 11.**
Procmon closely monitors what is running on each compute node.

User Support and Outreach

NERSC offers a wide range of user support services. The User Services Group (USG) provides the direct interface between NERSC and its user community. USG's mission is to increase the scientific productivity of NERSC users through technical support, education, advocacy and the development and deployment of new computational and data technologies. This group is responsible for problem management and consulting; helping with user code optimization and debugging; strategic project support; web documentation and training; and third-party applications and library support. USG staff serve as advocates for the user community, ensuring that new systems and services are well matched to the needs of scientists using NERSC.

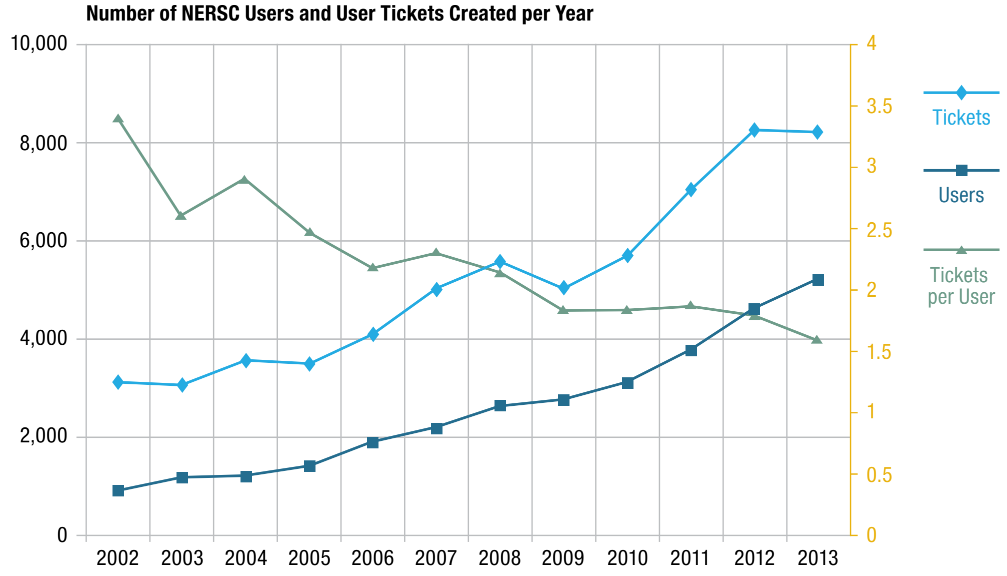
The services and engagements through User Services are manifold: resetting passwords, one-on-one consulting, running Office of Science-wide requirements reviews and coordinating the active NERSC User Group (NUG). NERSC's web site features a prominent "For Users" top-level section that makes it easy for users to find information, and the center's training activities provide additional opportunities for communicating with users.

NERSC's consultants and account support staff are available to users via email, an online web interface and on the phone during business hours (8 a.m.–5 p.m. Pacific Time). Basic account support (password resets, resetting login failures) is available via the NERSC operations staff 24 x 7, 365 days a year.

NERSC's 12 consultants include nine experts in HPC, six with Ph.D. degrees. NERSC's two account support personnel each have 10-plus years of experience. When users contact NERSC, they immediately are connected with highly trained HPC specialists who can usually solve issues directly or route the request to the appropriate systems engineers right away. NERSC's policy is to respond to all inquiries within four business hours and either solve the problem or communicate a work plan to the user within three business days for all reported incidents. The latter is a metric the center reports to DOE.

When users require more extensive assistance—such as programming or optimization help—they are routed to a consultant with the most appropriate skill set who will then engage in a longer-term collaboration with the user. Of course, with thousands of users and only a handful of consultants, extensive one-on-one assistance is not always possible. NERSC's strategy is to deliver assistance at scale by heavily relying on extensive web documentation and an active training program. Likewise, many account support services that were once labor intensive (e.g., password resets, updating contact information) have been converted to self-service operations via the NERSC Information Management (NIM) web interface (nim.nersc.gov). These efficiency improvements have resulted in fewer tickets per user (Figure 12).

Deep engagements with users leverage NERSC's extensive HPC technology expertise. For example, NERSC computational scientists have developed tools to help users create science gateways (web portals). These tools have contributed to NERSC hosting a number of prominent gateways for science teams, including those for the Materials Project, Deep Sky, the Coherent X-Ray Imaging Data Bank and the Daya Bay Neutrino Detector. Work is also under way to set up workflows to connect NERSC with the Advanced Light Source at Berkeley Lab and with users of the LINAC Coherent Light Source at SLAC.



▶ **Figure 12.**
Number of NERSC users and tickets created annually, 2002-2013.

Library Performance Improvements on Edison

A large fraction of the NERSC workload, such as many material science and chemistry codes, depends heavily on math libraries like FFTW, BLAS, LAPACK and ScaLAPACK. For this reason NERSC consultants continually evaluate the performance of target applications under various library versions and distributions.

During early testing on Edison, NERSC's newest supercomputer, NERSC materials science consultants discovered significant deficiencies in the Cray math library software stack. NERSC contacted Cray and pushed for improvements, which Cray implemented by the time Edison entered production in January 2014. In particular, consultants reported and received:

- Performance fixes in Cray LibSci, resulting, for example, in a factor of two speedup in ZGEMM (a matrix multiplication and addition routine) on Edison
- Cray LibSci library support for the Intel programming environment
- Support for the highly performant MKL libraries under the Cray programming environment. The threaded MKL FFT libraries are significantly faster than the Cray-provided FFTW builds in some instances.

Darshan I/O Performance Monitoring Software

In 2013, NERSC and its users began reaping the benefits of having installed the Darshan I/O performance monitoring software. On November 15, 2012, NERSC enabled Darshan by default on Hopper and then in 2013 did the same on Edison. NERSC configured the software to save I/O performance results to a database and provide an interface for both users and staff to query the results.

Through Darshan, NERSC users can view their codes' I/O performance and independently identify poorly performing I/O. In addition, NERSC proactively scans the results to find codes that can immediately benefit from rethinking the way they perform read and write operations.

As one example, NERSC Consultant Yushu Yao found a code running on 35,000 cores on Hopper that spent 33 percent of its time performing I/O. The Darshan data revealed that this code was writing only about 100MB to this file—an extremely small amount to be taking so much time. Yao contacted the user and found that the code was using the MPI I/O library to write to a small file in the user's home directory. The problem was that home directories are not configured to be used for parallel I/O and perform extremely poorly for that task.

When the code was reconfigured to read and write to Hopper's high-performing scratch file system, the percentage of time spent in I/O dropped to 1 percent of the total runtime. The result was a savings of 18,000 hours of compute time per run for this user and more time for other jobs running on Hopper.

BerkeleyGW Coding Workshop

In November 2013, NERSC hosted an event designed to give users assistance and guidance in modifying their codes to run on new energy-efficient manycore architectures. This event targeted the popular BerkeleyGW materials science code, one of the top 10 codes used at NERSC in 2013. At the workshop, NERSC was only able to accommodate 45 of the 120 applicants from around the world. The main BerkeleyGW developers attended from NERSC, Berkeley and MIT.

The workshop, organized by NERSC Consultant Jack Deslippe, consisted of two days of presentations and hands-on sessions (powered by NERSC resources), as well as developer discussion and staff interaction. BerkeleyGW is unique in that NERSC has staff expertise, but this workshop will serve as a prototype for future application workshops from many third party codes. Planning is under way for a similar workshop on Quantum Espresso in 2014.

SciDB Testbed Project

Data coming from simulations or experiments are commonly in the form of multi-dimensional arrays. SciDB is a parallel database that can efficiently manage, analyze and share terabytes of this type of array data. With SciDB, scientists can perform parallel computations on data without having to invest in developing complex parallel applications.

In 2012, NERSC set up a 20-node SciDB testbed cluster and invited science teams to evaluate the technology for their problems. NERSC formed a partnership with each science project and helped jumpstart the effort.

To date, NERSC has initiated more than 10 partnerships covering a broad range of science topics, from astrophysics to high-energy physics to biology. A number of high-profile projects (such as the LUX dark matter search experiment, the DOE Systems Biology Knowledgebase (KBase) and the IMG bioinformatics system) have shown interest in using SciDB as a long-term solution to their data problem and are requesting a production-grade SciDB service at NERSC.

NERSC Supports DARPA Mission Partners on Edison

Edison was the first Cray XC30 supercomputer, which was developed in part with the Defense Advanced Research Projects Agency's (DARPA) High Productivity Computing Systems (HPCS) program. Run by the U.S. Department of Defense, DARPA formed the HPCS program to foster development of the next generation of high productivity computing systems for both national security and industrial user communities. Program goals were to develop more broadly applicable, easier-to-program, failure-resistant high performance computing systems. The Cray XC30 system represents the last phase of the three-phase HPCS program.

As part of the program, DARPA and its mission partners were to have up to 20 million hours of computing time on an XC30 for testing, evaluation and machine characterization. In a collaboration with Cray and DARPA, NERSC hosted mission partners from DARPA, DOE and universities. Forty-four mission partner users were integrated into the NERSC community of users, with full access to NERSC training, consulting and other user benefits.

To ensure the mission partners had adequate access to the machine for their runs, NERSC implemented a fair-share schedule on the pre-production Edison system to assure DARPA got its promised 25 percent share of the system until the 20 million-hour usage level was reached. Mission partners concluded their testing when Edison went into production in January 2014.

Mission partners were active members of the NERSC community while they were using Edison, often attending NERSC User Group teleconferences and training events and submitting 121 consulting trouble tickets. Their problem reports were often among the most challenging for the NERSC consultants, as DARPA was testing novel programming models and stressing the system.

Ten users associated with the mission partner program completed the user survey, rating NERSC very highly: an average of 6.70 (out of 7.0) for overall satisfaction with NERSC and 6.78 for consulting overall.

They also left a number of comments, including:

> *"Customer service was excellent in all respects: timeliness and accuracy of answers, availability, and a helpful attitude. I enjoyed interacting with the customer assistance center."*

> *"Web pages, on account and run information, on system status, on usage tips, etc. are excellent compared to the other centers I work with. User support and consultants are very responsive and conscientious."*

> *"You provided help when I asked for it. Thanks!"*

> *"From my perspective, NERSC did an excellent job of managing Edison."*

Research and Development by NERSC Staff

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, as well as participating in professional organizations and conferences and contributing to journals and proceedings. The NERSC user community benefits from the results of these activities as they are applied to systems, software and services at NERSC and throughout the HPC community.

Publications and presentations by NERSC staff in 2013 are listed below. (Not all co-authors are from NERSC.) Online links to many of these reports and presentations can be found at http://www.nersc.gov/news-publications/publications-reports/nersc-staff-publications-and-presentations/.

IceCube Collaboration: M. G. Aartsen et al, "Search for Time-Independent Neutrino Emission from Astrophysical Sources with 3 yr of IceCube Data," *Astrophysical Journal* 779 132, December 2013.

Massimo Di Pierro, James Hetrick, Shreyas Cholia, James Simone, and Carleton DeTar., "The new 'Gauge Connection' at NERSC," 21st International Lattice Data Grid Workshop, December 2013.

IceCube Collaboration: M. G. Aartsen et al, "Probing the Origin of Cosmic Rays with Extremely High Energy Neutrinos Using the IceCube Observatory," *Physical Review* D88 112008, December 2013.

IceCube Collaboration: M. G. Aartsen et al, "An IceCube Search for Dark Matter Annihilation in Nearby Galaxies and Galaxy Clusters," *Physical Review* D88 122001, December 2013.

IceCube Collaboration: M. G. Aartsen et al, "Evidence for High-Energy Extraterrestrial Neutrinos at the IceCube Detector," *Science* 342 1242856, November 2013.

Richard Gerber (Berkeley Lab), Ken Bloom (U. Nebraska-Lincoln), "Report of the Snowmass 2013 Computing Frontier Working Group on Distributed Computing and Facility Infrastructures," to be included in proceedings of Community Summer Study ("Snowmass") 2013. November 11, 2013

M. J. Cordery, B. Austin, H. J. Wasserman, C. S. Daley, N. J. Wright, S. D. Hammond, D. Doerfler, "Analysis of Cray XC30 Performance using Trinity-NERSC-8 benchmarks and comparison with Cray XE6 and IBM BG/Q," SC `13, November 11, 2013.

Hongzhang Shan, Brian Austin, Wibe De Jong, Leonid Oliker, Nicholas Wright, Edoardo Apra, "Performance Tuning of Fock Matrix and Two-Electron Integral Calculations for NWChem on Leading HPC Platforms," SC 13, November 11, 2013.

Richard Gerber and Harvey Wasserman, eds., "Large Scale Computing and Storage Requirements for High Energy Physics - Target 2017," November 8, 2013.

Richard Gerber, NUG Webinar for November 2013. November 7, 2013.

N. Balthaser, GlobusOnline/HPSS Live Demo, HUF 2013. November 5, 2013.

N. Balthaser, "LBNL/NERSC Site Report: HPSS in Production," HUF 2013. November 5, 2013.

Richard A. Barrett, Shekhar Borkar, Sudip S. Dosanjh, Simon D. Hammond, Michael A. Heroux, X. Sharon Hu, Justin Luitjens, Steven G. Parker, John Shalf, Li Tang, "On the Role of Co-design in High Performance Computing," *Transition of HPC Towards Exascale Computing*, E.H. D'Hollander et. al (Eds.), IOS Press, 2013.

J. Hick, "A Storage Outlook for Energy Sciences: Data Intensive, Throughput and Exascale Computing," FujiFilm Executive IT Summit 2013. October 24, 2013.

Daya Bay Collaboration: F.P. An, A.B. Balantekin, H.R. Band, W. Beriguete, et. al, "Spectral measurement of electron antineutrino oscillation amplitude and frequency at Daya Bay," *Physical Review Letters*, October 24, 2013.

D.C. Eder, A.C. Fisher, A.E. Koniges and N.D. Masters, "Modelling debris and shrapnel generation in inertial confinement fusion experiments," *Nuclear Fusion* 53 113037, 2013.

Richard Gerber, "Edison Overview (Focus on hardware relevant for performance)," October 10, 2013.

Jack Deslippe, "Building Applications on Edison," October 10, 2013.

Trever Nightingale, "Introduction to Google Chromebooks," LBNL Tech Day Lightning Talk, October 10, 2013.

Oliver Rübel, Annette Greiner, Shreyas Cholia, Katherine Louie, E. Wes Bethel, Trent R. Northen, Benjamin P. Bowen, "OpenMSI: A High-Performance Web-Based Platform for Mass Spectrometry Imaging," *Analytical Chemistry*, 2013. 85 (21), pp 10354–10361, October 2, 2013.

IceCube Collaboration: M. G. Aartsen et al, "South Pole Glacial Climate Reconstruction from Multi-Borehole Laser Particulate Stratigraphy," *Journal of Glaciology* 59 1117-1128, October 2013.

Clayton Bagwell, "How to Submit a 2014 ERCAP Request," September 16, 2013.

Joaquin Correa, "Integrated Tools for NGBI—Lessons Learned and Successful Cases," LBNL Integrated Bioimaging Initiative, September 4, 2013.

IceCube Collaboration: M. G. Aartsen et al, "Measurement of the Cosmic Ray Energy Spectrum with IceTop-73," *Physical Review* D88 042004, August 28, 2013.

Rolf Riesen, Sudip Dosanjh, Larry Kaplan, "The ExaChallenge Symposium," IBM Research Paper, August 26, 2013.

Alice Koniges, Jean-Luc Vay, Alex Friedman,Hartmut Kaiser, and Thomas Sterling, "Consideration of Asynchronous Algorithms for Particle-Grid Simulations," 2013.

Joseph Teran, Alice Koniges, "Material Point Methods and Multiphysics for Fracture and Multiphase Problems," 2013.

IceCube Collaboration: R. Abbasi et al, "Measurement of Atmospheric Neutrino Oscillations with IceCube," *Physical Review Letters* 111 081801, August 2013.

Richard Gerber, "Data-Driven Science at NERSC," August 8, 2013.

Richard Gerber, "High-Performance Parallel I/O," August 6, 2013.

Richard Gerber, NERSC/Berkeley Lab and Ken Bloom, University of Nebraska-Lincoln, Snowmass Computing Frontier I2: Distributed Computing and Facility Infrastructures, July 31, 2013.

Anubhav Jain, Shyue Ping Ong, Geoffroy Hautier, Wei Chen, William Davidson Richards, Stephen Dacek, Shreyas Cholia, Dan Gunter, David Skinner, Gerbrand Ceder, Kristin A. Persson, "The Materials Project: A materials genome approach to accelerating materials innovation," *APL Materials* 1, 011002 (2013) http://dx.doi.org/10.1063/1.4812323, July 2013.

IceCube Collaboration: R. Abbasi et al, "First Observation of PeV-energy Neutrinos with IceCube," *Physical Review Letters* 111 021103, July 2013.

Richard Gerber, Dirac Science Highlights 2013. June 25, 2013.

Richard Gerber, Introduction to High Performance Computing, June 10, 2013.

Richard Gerber, Harvey Wasserman, "High Performance Computing and Storage Requirements for Biological and Environmental Research Target 2017," June 6, 2013.

IceCube Collaboration: M. G. Aartsen et al., "Measurement of South Pole Ice Transparency with the IceCube LED Calibration System," Nuclear Instruments and Methods A711 73-89, May 2013.

N. Balthaser, W. Hurlbert, "T10KC Technology in Production," May 9, 2013.

Larry Pezzaglia, "Supporting Multiple Workloads, Batch Systems, and Computing Environments on a Single Linux Cluster," Cray User Group 2013. May 9, 2013.

Zhengji Zhao, Katie Antypas, Nicholas J Wright, "Effects of Hyper-Threading on the NERSC workload on Edison," 2013 Cray User Group Meeting, May 9, 2013.

Brian Austin, Matthew Cordery, Harvey Wasserman, Nicholas J. Wright, "Performance Measurements of the NERSC Cray Cascade System," 2013 Cray User Group Meeting, May 9, 2013.

C. Daley, "Petascale Simulations of Turbulent Nuclear Combustion. ALCF-2 Early Science Program Technical Report," ALCF ESP Technical Report, 2013.

Brian Austin, NERSC, "Characterization of the Cray Aries Network," May 6, 2013.

David Camp, Hari Krishnan, David Pugmire, Christoph Garth, Ian Johnson, E. Wes Bethel, Kenneth I. Joy, Hank Childs, "GPU Acceleration of Particle Advection Workloads in a Parallel, Distributed Memory Setting," *Proceedings of Eurographics Symposium on Parallel Graphics and Visualization* (EGPGV), May 5, 2013.

S. Dosanjh, R. Barrett, D. Doerfler, S. Hammond, K. Hemmert, M. Heroux, P. Lin, K. Pedretti, A. Rodrigues, T. Trucano, J.Juitjens, "Exascale Design Space Exploration and Co-Design," *Future Generation Computer Systems*, special issue on Extreme Scale Parallel Architectures and Systems, in press, available online, May 2, 2013.

Jay Srinivasan, Richard Shane Canon, "Evaluation of A Flash Storage Filesystem on the Cray XE-6," CUG 2013. May 2013.

J. Hick, "Storage at a Distance," Open Fabrics Alliance User Day 2013. April 19, 2013.

Richard A. Gerber, "High Performance Computing and Big Data (for High Energy Physics theory)," April 2, 2013.

Alice Koniges, Wangyi Liu, John Barnard, Alex Friedman, Grant Logan, David Eder, Aaron Fisher, Nathan Masters, and Andrea Bertozzi, "Modeling warm dense matter experiments using the 3D ALE-AMR code and the move toward exascale computing," *Journal of Physics: Conference Series*, in press, 2013.

Richard A. Gerber, Fusion Energy Sciences Requirements Review Overview and Goals, March 19, 2013.

Richard Gerber, NUG March 2013 Webinar, March 7, 2013.

C. Daley, "Preparing for Mira: experience with FLASH multiphysics simulations," Mira Community Conference, 2013.

A. Friedman, R. H. Cohen, D. P. Grote, W. M. Sharp, I. D. Kaganovich, A. E. Koniges, W. Liu, "Heavy Ion Beams and Interactions with Plasmas and Targets (HEDLP and IFE)," NERSC FES Requirements Workshop: Heavy Ion Fusion and Non-Neutral Plasmas, March 2013.

Alice Koniges, Praveen Narayanan, Robert Preissl, Xuefei Yuan, "Proxy Design and Optimization in Fusion and Accelerator Physics," SIAM Conference on Computational Science and Engineering, February 25, 2013.

Richard A. Gerber, "Debugging and Optimization Tools," February 19, 2013.

Richard Gerber, "Trends, Discovery, & Innovation" at NUG User Day 2013. February 19, 2013.

N. Balthaser, "Introduction to NERSC Archival Storage," February 15, 2013.

Katie Antypas, "Best Practices for Reading and Writing Data on HPC Systems," NUG Meeting 2013. February 14, 2013.

Richard A. Gerber, Tina Declerck. Zhengji Zhao, "Edison Update," February 12, 2013.

Richard A. Gerber, Harvey Wasserman, "NERSC Requirements Reviews," February 12, 2013.

Francesca Verdier, "NERSC Accomplishments and Plans," February 12, 2013.

Katie Antypas, "NERSC-8 Project," NUG Meeting, February 12, 2013.

Brent Draney, "NERSC's New Building Update," February 12, 2013.

Nick Wright, "NERSC Initiative: Preparing Applications for Exascale," February 12, 2013.

Yushu Yao, "NERSC Parallel Database Evaluation," February 12, 2013.

Richard Gerber, "Requirements Reviews Update," February 12, 2013.

David Skinner and Shane Canon, "NERSC and High Throughput Computing," February 12, 2013.

David E. Keyes, Lois Curfman McInnes, Carol Woodward, William Gropp, Eric Myra, Michael Pernice, John Bell, Jed Brown, Alain Clo, Jeffrey Connors, Emil Constantinescu, Don Estep, Kate Evans, Charbel Farhat, Ammar Hakim, Glenn Hammond, Glen Hansen, Judith Hill, Tobin Isaac, Xiaomin Jiao, Kirk Jordan, Dinesh Kaushik, Efthimios Kaxiras, Alice Koniges, et al., "Multiphysics Simulations Challenges and Opportunities," *International Journal of High Performance Computing Applications*, February 2013. 27:4--83,

Shyue Ping Ong, William Davidson Richards, Anubhav Jain, Geoffroy Hautier, Michael Kocher, Shreyas Cholia, Dan Gunter, Vincent L. Chevrier, Kristin A. Persson, Gerbrand Ceder, "Python Materials Genomics (pymatgen): A robust, open-source python library for materials analysis," *Computational Materials Science*, Volume 68, February 2013. Pages 314-319, ISSN 0927-0256, 10.1016/j.commatsci.2012.10.028., February 1, 2013.

J. Barnard, R. M. More, P. A. Ni, A. Friedman, E. Henestroza, I. Kaganovich, A. Koniges, J. W. Kwan, W. Liu, A. Ng, B.G. Logan, E. Startsev, M. Terry, A. Yuen, "NDCX-II Experimental Plans and Target Simulations," West Coast High Energy Density Science Cooperative Meeting Berkeley and Palo Alto, California, January 2013.

Richard A. Gerber, "Getting Started at NERSC," January 17, 2013.

Richard Gerber, Kathy Yelick, Lawrence Berkeley National Laboratory, "Data Requirements from NERSC Requirements Reviews," January 9, 2013.

J. Hick, "GPFS at NERSC/LBNL," SPXXL Winter 2013. January 7, 2013.

S. Hachinger, P. A. Mazzali, M. Sullivan, R. S., K. Maguire, A. Gal-Yam, D. A. Howell, P. E., E. Baron, J. Cooke, I. Arcavi, D., B. Dilday, P. A. James, M. M. Kasliwal, S. R., E. O. Ofek, R. R. Laher, J. Parrent, J. Surace, O. Yaron, E. S. Walker, "The UV/optical spectra of the Type Ia supernova SN 2010jn: a bright supernova with outer layers rich in iron-group elements," *Monthly Notices of the RAS*, 2013. 429:2228-2248,

E. O. Ofek, M. Sullivan, S. B. Cenko, M. M. Kasliwal, A., S. R. Kulkarni, I. Arcavi, L. Bildsten, J. S., A. Horesh, D. A. Howell, A. V. Filippenko, R., D. Murray, E. Nakar, P. E. Nugent, J. M., N. J. Shaviv, J. Surace, O. Yaron, An outburst from a massive star 40days before a supernova explosion, Nature, Pages: 65-67, 2013.

Jack Deslippe, Georgy Samsonidze, Manish Jain, Marvin L Cohen, Steven G Louie, "Coulomb-hole summations and energies for GW calculations with limited number of empty orbitals: a modified static remainder approach," Accepted *Physical Review* B (arXiv preprint arXiv:1208.0266), 2013.

Patrick Oesterling, Christian Heine, Gunther H. Weber, Scheuermann, "Visualizing nD Point Clouds as Topological Landscape Profiles to Guide Data Analysis," *IEEE Transactions on Visualization and Computer Graphics*, 19(3):514-526, 2013.

Dmitriy Morozov, Gunther H. Weber, "Distributed Merge Trees," *Proceedings of the 18th ACM SIGPLAN Symposium on Principles and Practice of Parallel Programming (PPoPP 13)*, 93—102. New York, NY, USA, ACM, 2013.

E. Wes Bethel, Prabhat, Suren Byna, Oliver Rübel, K. John Wu, Michael Wehner, "Why High Performance Visual Data Anaytics is Both Relevant and Difficult," Visualization and Data Analysis, IS&T/SPIE Electronic Imaging 2013. San Francisco, CA, USA, 2013.

Babak Behzad, Joseph Huchette, Huong Luu, Suren Byna, Yushu Yao, Prabhat, "A Framework for Auto-tuning HDF5 Applications," HPDC, 2013.

Suren Byna, Andrew Uselton, Prabhat, David Knaak, Helen He, "Trillion Particles, 120,000 cores, and 350 TBs: Lessons Learned from a Hero I/O Run on Hopper," Cray User Group Meeting, 2013.

Mark Howison, E. Wes Bethel, "GPU-accelerated Denoising of 3D Magnetic Resonance Images, Journal of Real Time Image Processing," 2013.

Wei-Chen Chen, George Ostrouchov, Dave Pugmire, Prabhat, Michael Wehner, "Exploring multivariate relationships in Large Spatial Data with Parallel Model-Based Clustering and Scalable Graphics," *Technometrics*, 2013.

Daithi Stone, Chris Paciorek, Prabhat, Pardeep Pall, Michael Wehner, "Inferring the anthropogenic contribution to local temperature extremes," *PNAS*, 110 (7), 2013.

Dean N. Williams, Timo Bremer, Charles Doutriaux, John Patchett, Galen Shipman, Blake Haugen, Ross Miller, Brian Smith, Chad Steed, E. Wes Bethel, Hank Childs, Harinarayan Krishnan, Prabhat, Michael Wehner, Claudio T. Silva, Emanuele Santos, David Koop, Tommy Ellqvist, Huy T. Vo, Jorge Poco, Berk Geveci, Aashish Chaudhary, Andrew Bauer, Alexander Pletzer, Dave Kindig, Gerald L. Potter, Thomas P. Maxwell, "The Ultra-scale Visualization Climate Data Analysis Tools: Data Analysis and Visualization for Geoscience Data," *IEEE Special Issue: Cutting-Edge Research in Visualization*, 2013.

Prabhat, William D. Collins, Michael Wehner, "Extreme-Scale Climate Analytics," Google Regional PhD Summit, 2013.

Prabhat, William D. Collins, Michael Wehner, Suren Byna, Chris Paciorek, "Big Data Challenges in Climate Science," Berkeley Atmospheric Science Symposium, 2013.

Prabhat, "Pattern Detection for Large Climate Datasets," *Climate 2013*. 2013.

Prabhat, "Data Formats, Data Models and Parallel I/O," CRD Division Review, 2013.

P. A. Mazzali S. Hachinger, "The UV/optical spectra of the Type Ia supernova SN 2010jn: a supernova with outer layers rich in iron-group elements," *Monthly Notices of the Royal Astronomical Society*, Pages: 490 2013.

Xuefei Yuan, Xiaoye S Li, Ichitaro Yamazaki, Stephen C Jardin, Alice E Koniges, David E Keyes, "Application of PDSLin to the magnetic reconnection problem," *Computational Science & Discovery*, 6:014002, 2013.

D. Eder, D. Bailey, F. Chambers, I. Darnell, P. Di Nicola, S. Dixit, A. Fisher, G.Gururangan, D. Kalantar, A. Koniges, W. Liu, M. Marinak, N. Masters, V. Mlaker, R. Prasad, S. Sepke, P. Whitman, "Observations and modeling of debris and shrapnel impacts on optics and diagnostics at the National Ignition Facility," *J. Physics*, in press, 2013.

A. Horesh, C. Stockdale, D. B. Fox, D. A. Frail, J., S. R. Kulkarni, E. O. Ofek, A., M. M. Kasliwal, I. Arcavi, R. Quimby, S. B., P. E. Nugent, J. S. Bloom, N. M. Law, D., E. Gorbikov, D. Polishook, O. Yaron, S., K. W. Weiler, F. Bauer, S. D. Van Dyk, S., N. Panagia, D. Pooley, N. Kassim, "An early and comprehensive millimetre and centimetre wave and X-ray study of SN 2011dh: a non-equipartition blast wave expanding into a massive stellar wind," *Monthly Notices of the RAS*, Pages: 1258-1267, 2013.

J. M. Silverman, J. Vinko, M. M. Kasliwal, O. D., Y. Cao, J. Johansson, D. A. Perley, D., J. C. Wheeler, R. Amanullah, I. Arcavi, J. S., A. Gal-Yam, A. Goobar, S. R. Kulkarni, R., W. H. Lee, G. H. Marion, P. E. Nugent, I. Shivvers, "SN 2000cx and SN 2013bh: extremely rare, nearly twin Type Ia supernovae," *Monthly Notices of the RAS*, Pages: 1225-1237, 2013.

U. Feindt, M. Kerschhaggl, M. Kowalski, G. Aldering, P., C. Aragon, S. Bailey, C. Baltay, S., C. Buton, A. Canto, F. Cellier-Holzem, M., N. Chotard, Y. Copin, H. K. Fakhouri, E., J. Guy, A. Kim, P. Nugent, J., K. Paech, R. Pain, E. Pecontal, R., S. Perlmutter, D. Rabinowitz, M., K. Runge, C. Saunders, R. Scalzo, G., C. Tao, R. C. Thomas, B. A. Weaver, C. Wu, "Measuring cosmic bulk flows with Type Ia supernovae from the Nearby Supernova Factory," *Astronomy and Astrophysics*, Pages: A90, 2013.

M. Rigault, Y. Copin, G. Aldering, P. Antilogus, C. Aragon, S. Bailey, C. Baltay, S. Bongard, C. Buton, A. Canto, F. Cellier-Holzem, M. Childress, N. Chotard, H. K. Fakhouri, U. Feindt, M. Fleury, E. Gangler,

P. Greskovic, J. Guy, A. G. Kim, M. Kowalski, S. Lombardo, J. Nordin, P. Nugent, R. Pain, E. Pécontal, R. Pereira, S. Perlmutter, D. Rabinowitz, K. Runge, C. Saunders, R. Scalzo, G. Smadja, C. Tao, R. C. Thomas, B. A. Weaver, "Evidence of environmental dependencies of Type Ia supernovae from the Nearby Supernova Factory indicated by local Hα," *Astronomy and Astrophysics*, Pages: A66, 2013.

K. Maguire, M. Sullivan, F. Patat, A. Gal-Yam, I. M. Hook, S. Dhawan, D. A. Howell, P. Mazzali, P. E. Nugent, Y.-C. Pan, P. Podsiadlowski, J. D. Simon, A. Sternberg, S. Valenti, C. Baltay, D. Bersier, N. Blagorodnova, T.-W. Chen, N. Ellman, U. Feindt, F. Förster, M. Fraser, S. González-Gaitán, M. L. Graham, C. Gutiérrez, S. Hachinger, E. Hadjiyska, C. Inserra, C. Knapic, R. R. Laher, G. Leloudas, S. Margheim, R. McKinnon, M. Molinaro, N. Morrell, E. O. Ofek, D. Rabinowitz, A. Rest, D. Sand, R. Smareglia, S. J. Smartt, F. Taddia, E. S. Walker, N. A. Walton, and D. R. Young, "A statistical analysis of circumstellar material in Type Ia supernovae," *Monthly Notices of the RAS*, Pages: 222-240, 2013.

W. Zheng, J. M. Silverman, A. V. Filippenko, D. Kasen, P. E. Nugent, M. Graham, X. Wang, S. Valenti, F. Ciabattari, P. L. Kelly, O. D. Fox, I. Shivvers, K. I. Clubb, S. B. Cenko, D. Balam, D. A. Howell, E. Hsiao, W. Li, G. H. Marion, D. Sand, J. Vinko, J. C. Wheeler, J. Zhang, "The Very Young Type Ia Supernova 2013dy: Discovery, and Strong Carbon Absorption in Early-time Spectra," *Astrophysical Journal Letters*, Pages: L15, 2013.

L. P. Singer, S. B. Cenko, M. M. Kasliwal, D. A., E. O. Ofek, D. A. Brown, P. E. Nugent, S. R., A. Corsi, D. A. Frail, E. Bellm, J., I. Arcavi, T. Barlow, J. S. Bloom, Y., N. Gehrels, A. Horesh, F. J. Masci, J., A. Rau, J. A. Surace, O. Yaron, "Discovery and Redshift of an Optical Afterglow in 71 square degrees: iPTF13bxl and GRB 130702A," *Astrophysical Journal Letters*, Pages: L34, 2013.

Y. Cao, M. M. Kasliwal, I. Arcavi, A. Horesh, P., S. Valenti, S. B. Cenko, S. R. Kulkarni, A., E. Gorbikov, E. O. Ofek, D. Sand, O., M. Graham, J. M. Silverman, J. C. Wheeler, G. H., E. S. Walker, P. Mazzali, D. A. Howell, K. L., A. K. H. Kong, J. S. Bloom, P. E. Nugent, J., F. Masci, J. Carpenter, N. Degenaar, C. R. Gelino, Discovery, "Progenitor and Early Evolution of a Stripped Envelope Supernova iPTF13bvn," *Astrophysical Journal Letters*, Pages: L7, 2013.

J. M. Silverman, P. E. Nugent, A. Gal-Yam, M., D. A. Howell, A. V. Filippenko, Y.-C. Pan, S. B. Cenko, I. M. Hook, "Late-time Spectral Observations of the Strongly Interacting Type Ia Supernova PTF11kx," *Astrophysical Journal*, Pages: 125, 2013.

S. Taubenberger, M. Kromer, S. Hachinger, P. A., S. Benetti, P. E. Nugent, R. A. Scalzo, R., V. Stanishev, J. Spyromilio, F. Bufano, S. A. Sim, B. Leibundgut, W. Hillebrandt, "Super-Chandrasekhar Type Ia Supernovae at nebular epochs," *Monthly Notices of the RAS*, Pages: 3117-3130, 2013.

J. M. Silverman, P. E. Nugent, A. Gal-Yam, M., D. A. Howell, A. V. Filippenko, I., S. Ben-Ami, J. S. Bloom, S. B. Cenko, Y., R. Chornock, K. I. Clubb, A. L. Coil, R. J., M. L. Graham, C. V. Griffith, A., M. M. Kasliwal, S. R. Kulkarni, D. C., W. Li, T. Matheson, A. A. Miller, M., E. O. Ofek, Y.-C. Pan, D. A. Perley, D., R. M. Quimby, T. N. Steele, A. Sternberg, D. Xu, O. Yaron, "Type Ia Supernovae Strongly Interacting with Their Circumstellar Medium," *Astrophysical Journal Supplement*, Pages: 3, 2013.

C. Baltay, D. Rabinowitz, E. Hadjiyska, E. S. Walker, P., P. Coppi, N. Ellman, U. Feindt, R. McKinnon, B. Horowitz, A. Effron, "The La Silla-QUEST Low Redshift Supernova Survey," *Publications of the ASP*, Pages: 683-694, 2013.

E. Terziev, N. M. Law, I. Arcavi, C. Baranec, J. S. Bloom, K. Bui, M. P. Burse, P. Chorida, H. K. Das, R. G. Dekany, A. L. Kraus, S. R. Kulkarni, P. Nugent, Eran O. Ofek, S. Punnadi, A. N. Ramaprakash, R. Riddle, S. P. Tendulkar, "Millions of Multiples: Detecting and Characterizing Close-separation Binary Systems in Synoptic Sky Surveys" *Astrophysical Journal Supplement*, Pages: 18, 2013.

M. J. Childress, G. Aldering, P. Antilogus, C. Aragon, S. Bailey, C. Baltay, S. Bongard, C. Buton, A. Canto, F. Cellier-Holzem, N. Chotard, Y. Copin, H. K. Fakhouri, E. Gangler, J. Guy, E. Y. Hsiao, M. Kerschhaggl, A. G. Kim, M. Kowalski, S. Loken, P. Nugent, K. Paech, R. Pain, E. Pecontal, R. Pereira, S. Perlmutter, D. Rabinowitz, M. Rigault, K. Runge, R. Scalzo,G. Smadja, C. Tao, R. C. Thomas, B. A. Weaver, C. Wu, "Host Galaxy Properties and Hubble Residuals of Type Ia Supernovae from the Nearby Supernova Factory," *Astrophysical Journal*, Pages: 108, 2013.

M. J. Childress, G. Aldering, P. Antilogus, C. Aragon, S. Bailey, C. Baltay, S. Bongard, C. Buton, A. Canto, F. Cellier-Holzem, N. Chotard, Y. Copin, H. K. Fakhouri, E. Gangler, J. Guy, E. Y. Hsiao, M. Kerschhaggl, A. G. Kim, M. Kowalski, S. Loken, P. Nugent, K. Paech, R. Pain, E. Pecontal, R. Pereira, S. Perlmutter, D. Rabinowitz, M. Rigault, K. Runge, R. Scalzo,G. Smadja, C. Tao, R. C. Thomas, B. A. Weaver, C. Wu, "Host Galaxies of Type Ia Supernovae from the Nearby Supernova Factory," *Astrophysical Journal*, Pages: 107, 2013.

S. B. Cenko, S. R. Kulkarni, A. Horesh, A. Corsi, D. B. Fox, J. Carpenter, D. A. Frail, P. E. Nugent, D. A. Perley, D. Gruber, A. Gal-Yam, P. J. Groot, G. Hallinan, E. O. Ofek, A. Rau, C. L. MacLeod, A. A. Miller, J. S. Bloom, A. V. Filippenko, M. M. Kasliwal, N. M. Law, A. N. Morgan, D. Polishook, D. Poznanski, R. M. Quimby, B. Sesar, K. J. Shen, J. M. Silverman, & A. Sternberg, "Discovery of a Cosmological, Relativistic Outburst via its Rapidly Fading Optical Emission," *Astrophysical Journal*, Pages: 130, 2013.

R. Pereira, R. C. Thomas, G. Aldering, P. Antilogus, C. Baltay, S. Benitez-Herrera, S. Bongard, C. Buton, A. Canto, F. Cellier-Holzem, J. Chen, M. Childress, N. Chotard, Y. Copin, H. K. Fakhouri, M. Fink, D. Fouchez, E. Gangler, J. Guy, W. Hillebrandt, E. Y. Hsiao, M. Kerschhaggl, M. Kowalski, M. Kromer, J. Nordin, P. Nugent, K. Paech, R. Pain, E. Pécontal, S. Perlmutter, D. Rabinowitz, M. Rigault, K. Runge, C. Saunders, G. Smadja, C. Tao, S. Taubenberger, A. Tilquin, C. Wu, "Spectrophotometric time series of SN 2011fe from the Nearby Supernova Factory," *Astronomy and Astrophysics*, Pages: A27, 2013.

Barone-Nugent, R. L.; Lidman, C.; Wyithe, J. S. B.; Mould, J.; Howell, D. A.; Hook, I. M.; Sullivan, M.; Nugent, P. E.; Arcavi, I.; Cenko, S. B.; Cooke, J.; Gal-Yam, A.; Hsiao, E. Y.; Kasliwal, M. M.; Maguire, K.; Ofek, E.; Poznanski, D.; Xu, D., "Erratum: Near-infrared observations of type Ia supernovae: The best known standard candle for cosmology," *Monthly Notices of the RAS*, Pages: L90, 2013.

David Levitan, Thomas Kupfer, Paul J. Groot, Shrinivas R. Kulkarni, Thomas A. Prince, Gregory V. Simonian, Iair Arcavi,Joshua S. Bloom, Russ Laher, Peter E. Nugent, Eran O. Ofek, Branimir Sesar, Jason Surace, "Five new outbursting AM CVn systems discovered by the Palomar Transient Factory," *Monthly Notices of the RAS*, Pages: 996-1007, 2013.

S. Tang, Y. Cao, L. Bildsten, P. Nugent, E., S. R. Kulkarni, R. Laher, D. Levitan, F., E. O. Ofek, T. A. Prince, B. Sesar, J. Surace, "R Coronae Borealis Stars in M31 from the Palomar Transient Factory," *Astrophysical Journal Letters*, Pages: L23, 2013.

A. G. Kim, R. C. Thomas, G. Aldering, P. Antilogus, C., S. Bailey, C. Baltay, S. Bongard, C., A. Canto, F. Cellier-Holzem, M. Childress, N., Y. Copin, H. K. Fakhouri, E. Gangler, J., M. Kerschhaggl, M. Kowalski, J. Nordin, P., K. Paech, R. Pain, E. Pecontal, R., S. Perlmutter, D. Rabinowitz, M., K. Runge, C. Saunders, R. Scalzo, G. Smadja, C. Tao, B. A. Weaver, C. Wu, "Standardizing Type Ia Supernova Absolute Magnitudes Using Gaussian Process Data Regression," *Astrophysical Journal*, Pages: 84, 2013.

E. Y. Hsiao, G. H. Marion, M. M. Phillips, C. R. Burns, C. Winge, N. Morrell, C. Contreras, W. L. Freedman, M. Kromer, E. E. E. Gall, C. L. Gerardy, P. Höflich, M. Im, Y. Jeon, R. P. Kirshner, P. E. Nugent, S. E. Persson, G. Pignata, M. Roth, V. Stanishev, M. Stritzinger, and N. B. Suntzeff, "The Earliest Near-infrared Time-series Spectroscopy of a Type Ia Supernova," *Astrophysical Journal*, Pages: 72, 2013.

P. W. Nugent, J. A. Shaw, S. Piazzolla, "Infrared Cloud Imager Development for Atmospheric Optical Communication Characterization, and Measurements at the JPL Table Mountain Facility," *Interplanetary Network Progress Report*, Pages: C1, 2013.

E. O. Ofek, D. Fox, S. B. Cenko, M. Sullivan, O., D. A. Frail, A. Horesh, A. Corsi, R. M., N. Gehrels, S. R. Kulkarni, A., P. E. Nugent, O. Yaron, A. V. Filippenko, M. M., L. Bildsten, J. S. Bloom, D., I. Arcavi, R. R. Laher, D. Levitan, B. Sesar, J. Surace, "X-Ray Emission from Supernovae in Dense Circumstellar Matter Environments: A Search for Collisionless Shocks," *Astrophysical Journal*, Pages: 42, 2013.

C. Buton, Y. Copin, G. Aldering, P. Antilogus, C., S. Bailey, C. Baltay, S. Bongard, A., F. Cellier-Holzem, M. Childress, N., H. K. Fakhouri, E. Gangler, J. Guy, E. Y., M. Kerschhaggl, M. Kowalski, S., P. Nugent, K. Paech, R. Pain, E., R. Pereira, S. Perlmutter, D., M. Rigault, K. Runge, R. Scalzo, G., C. Tao, R. C. Thomas, B. A. Weaver, C. Wu, "Nearby SuperNova Factory, Atmospheric extinction properties above Mauna Kea from the Nearby SuperNova Factory spectro-photometric data set," *Astronomy and Astrophysics*, Pages: A8, 2013.

You-Wei Cheah, Richard Canon, Plale, Lavanya Ramakrishnan, "Milieu: Lightweight and Configurable Big Data Provenance for Science," BigData Congress, 46-53, 2013.

Andrew Uselton, Nicholas J. Wright, "A file system utilization metric for I/O characterization," 2013 Cray User Group Conference, Napa, CA, 2013.

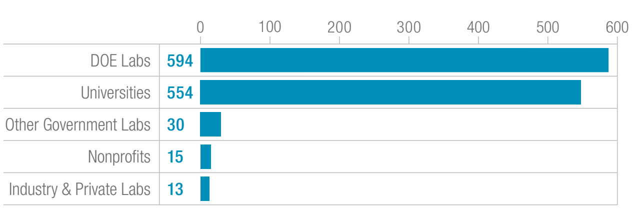
Appendix A

NERSC User Statistics

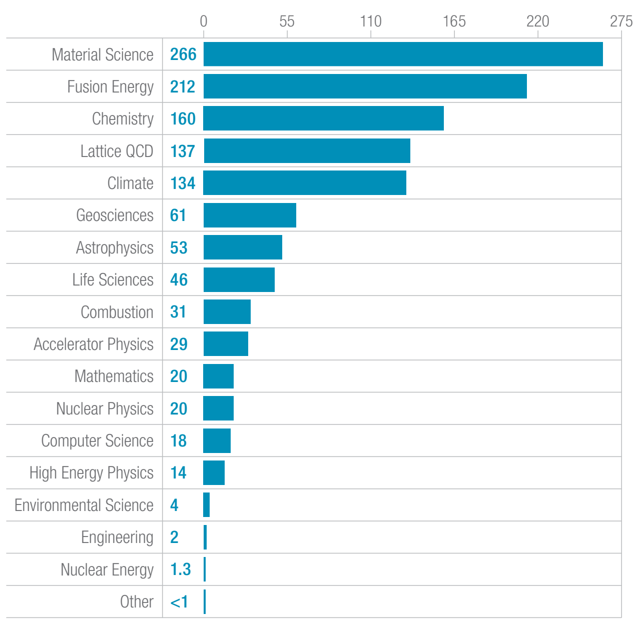
In support of the DOE Office of Science's mission, NERSC served 5,192 scientists in 42 countries in 2013. These researchers work in DOE laboratories, universities, industry and other federal agencies. Figure 1 shows the proportion of NERSC usage by each type of institution, while Figures 2 and 3 show laboratory, university and other organizations that used large allocations of computer time. Computational science conducted at NERSC covers the entire range of scientific disciplines, but is focused on research that supports the DOE's mission and scientific goals, as shown in Figure 4.

On their Allocation Year 2014 proposal forms, principal investigators reported 1,977 refereed publications (published or in press) for the preceding 12 months, based at least in part on using NERSC resources. Lists of publications resulting from use of NERSC resources are available at https://www.nersc.gov/news-publications/publications-reports/nersc-user-publications/.

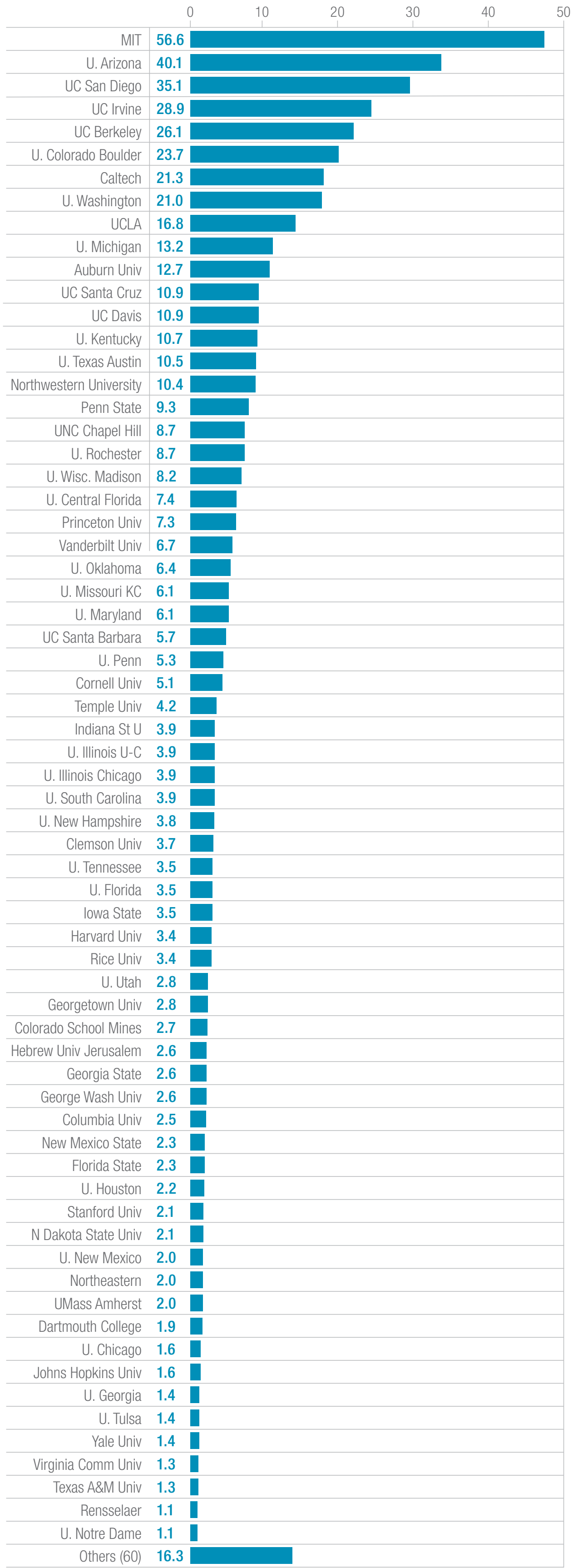
The MPP hours reported here are Cray XE6 equivalent hours.



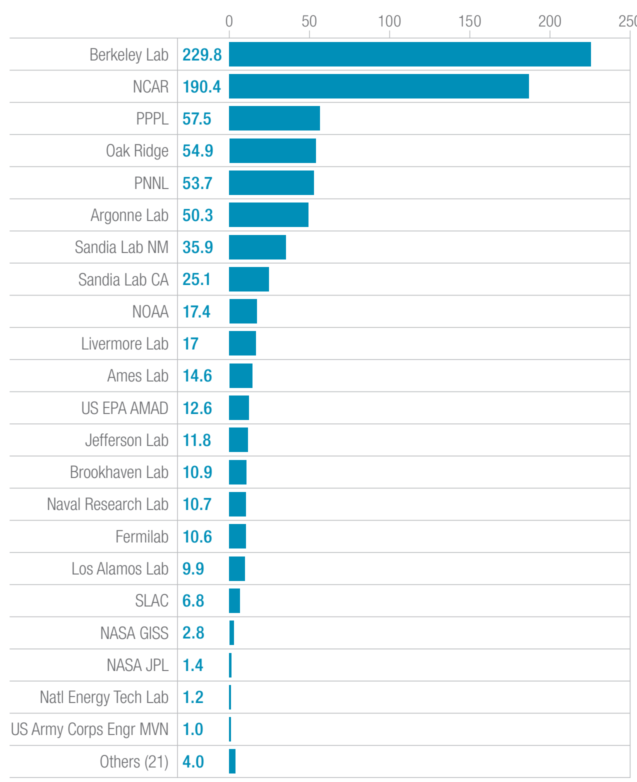
▲ **Figure 1.** NERSC usage by institution type, 2013 (MPP hours in millions).



▲ **Figure 2.** NERSC usage by scientific discipline, 2013 (MPP hours in millions).



▲ **Figure 3.** Academic usage at NERSC, 2013 (MPP hours in millions).



▲ **Figure 4.** DOE and other laboratory usage at NERSC, 2013 (MPP hours in millions).

Appendix B

NERSC Users Group Executive Committee

Office of Advanced Scientific Computing Research

Mark Adams, *Columbia University*
Anubhav Jain, *Lawrence Berkeley National Laboratory*
Vadim Roytershteyn, *SciberQuest, Inc.*

Office of Basic Energy Sciences

Eric Bylaska, *Pacific Northwest National Laboratory*
Gary Grest, *Sandia National Laboratories*
Paul Kent, *Oak Ridge National Laboratory*

Office of Biological and Environmental Research

Thomas Bettge, *National Center for Atmospheric Research*
Rudesh Toofanny, *University of Washington*
Guishan Zheng, *Harvard University*

Office of Fusion Energy Physics

Julian Borrill, *Lawrence Berkeley National Laboratory*
Steven Gottlieb, *Indiana University*
Frank Tsung (Vice Chair), *University of California, Los Angeles*

Office of Nuclear Physics

Daniel Kasen, *University of California, Berkeley*
Tomasz Plewa, *Florida State University*
Martin Savage, *University of Washington*

Members at Large

Cameron Geddes, *Lawrence Berkeley National Laboratory*
Konstantin Kemenov, *CRAFT Tech*
Gregory Newman, *Lawrence Berkeley National Laboratory*

Appendix C

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:

Steve Binkley, *Associate Director, ASCR*
Julie Stambaugh, *Financial Management Specialist*
Lori Jernigan, *Program Support Specialist*
Melea Baker, *Administrative Specialist*

Facilities Division

Barbara Helland, *Director*
Vince Dattoria, *Computer Scientist*
Betsy Riley, *Computer Scientist*
Dave Goodwin, *Physical Scientist*
Sally McPherson, *Program Assistant*
Ravi Kapoor, *Computer Scientist*

Research Division

William Harrod, *Director*
Teresa Beachley, *Program Assistant*
Christine Chalk, *Physical Scientist*
Carolyn Lauzon, *AAAS S&T Fellow*
Randall Laviolette, *Physical Scientist*
Thomas Ndousse-Fetter, *Computer Scientist*
Karen Pao, *Mathematician*
Ceren Susut, *Physical Scientist*
Rich Carlson, *Computer Scientist*
Sandy Landsberg, *Mathematician*
Steven Lee, *Physical Scientist*
Lucy Nowell, *Computer Scientist*
Sonia Sachs, *Computer Scientist*
Robert Lindsay, *Computer Scientist*
Angie Thevenot, *Program Assistant*

Appendix D

Acronyms and Abbreviations

ACM
Association for Computing Machinery

ACS
American Chemical Society

ALCC
ASCR Leadership Computing Challenge

ALS
Advanced Light Source, Lawrence Berkeley National Laboratory

ALTD
Automated Library Tracking Database

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

BER
Office of Biological and Environmental Research

BES
Office of Basic Energy Sciences

CAL
Computer Architecture Laboratory

CARB
California Air Resources Board

CCM
Cluster Compatibility Mode

CLE
Cray Linux Environment

CMB
Cosmic microwave background

CRT
Computational Research and Theory Facility (under construction, Lawrence Berkeley National Laboratory)

CO_2
Carbon dioxide

CPU
Central processing unit

CRD
Computational Research Division, Lawrence Berkeley National Laboratory

CSE
Computational Science and Engineering

CSGF
Computational Science Graduate Fellowship

CTO
Chief Technology Officer

DARPA
Defense Advanced Research Projects Agency

DFT
Density functional theory

DNA
Deoxyribonucleic acid

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 at Pacific Northwest National Laboratory

EPSI
SciDAC Center for Edge Physics Simulation

ESnet
Energy Sciences Network

eV
Electron volts

FES
Office of Fusion Energy Sciences

FIRST
Fluid Interface Reactions, Structures and Transport: a DOE Energy Frontier Research Center

FTP
File Transfer Protocol

GB
Gigabytes

Gbps
Gigabits per second

GPFS
Global Parallel File System (IBM)

GPU
Graphics Processing Unit

GUI
Graphical user interface

HECC
NASA's High-End Computing Capability Project

HEP
Office of High Energy Physics

HMP
Human Microbiome Project

HPC
High performance computing

Appendix D / Acronyms and Abbreviations *continued*

HPSS
High Performance Storage System

HTML
Hypertext Markup Language

HTTP
Hypertext Transfer Protocol

IMPACTS
Investigation of the Magnitudes and Probabilities of Abrupt Climate Transitions

IEEE
Institute of Electrical and Electronics Engineers

I/O
Input/output

ITER
Latin for "the way," an international fusion energy experiment in southern France

JGI
Joint Genome Institute

LBNL
Lawrence Berkeley National Laboratory

LED
Light-emitting diode

LLNL
Lawrence Livermore National laboratory

LMT
Lustre Monitoring Tool

MIT
Massachusetts Institute of Technology

MOF
Metal oxide framework

MPI
Message Passing Interface

MPP
Massively parallel processing

MSI
Mass spectrometry imaging

NCAR
National Center for Atmospheric Research

NCGC
Center for Nanoscale Control of CO_2: a DOE Energy Frontier Research Center

NEMS
Nanoelectromechanical systems

NERSC
National Energy Research Scientific Computing Center

NEWT
NERSC Web Toolkit

NEXAFS
Near edge X-ray absorption fine structure

NGF
NERSC Global Filesystem

NICS
National Institute for Computational Sciences at the University of Tennessee/Oak Ridge National Laboratory

NIH
National Institutes of Health

NIM
NERSC Information Management

NISE
NERSC Initiative for Scientific Exploration

NOAA
National Oceanic and Atmospheric Administration

NP
Office of Nuclear Physics

NSF
National Science Foundation

NUG
NERSC Users Group

OLC
Onion-like carbon

OLCF
Oak Ridge Leadership Computing Facility

OpenMSI
Open Mass Spectrometry Imaging

PDSF
Parallel Distributed Systems Facility (NERSC)

PeV
1 quadrillion electron volts

PI
Principal investigator

PNNL
Pacific Northwest National Laboratory

PPPL
Princeton Plasma Physics Laboratory

PTF
Palomar Transient Factory

QCD
Quantum chromodynamics

RNA
Ribonucleic acid

SC
DOE Office of Science

SciDAC
Scientific Discovery through Advanced Computing (DOE)

SDN
Software-defined networking

SDSS
Sloan Digital Sky Survey

SIAM
Society for Industrial and Applied Mathematics

SQL
Structured Query Language

STM
Scanning tunneling microscope

TACC
Texas Advanced Computing Center

TB
Terabytes

TCP/IP
Transmission Control Protocol/Internet Protocol

TeV
50 trillion electron volts

TMD
Transition metal dichalcogenides

URL
Universal Resource Locator

VASP
Vienna Ab initio Simulation Package

VPC
Virtual private cluster

WRF Chem
Weather Research and Forecasting model with coupled chemistry

Xe
Xenon

For more information about NERSC, contact:

Jon Bashor

NERSC Communications
1 Cyclotron Road
Berkeley Lab, MS 50B2215
Berkeley, CA 94720-8148

email jbashor@lbl.gov
phone (510) 486-5849
fax (510) 486-4300

NERSC's Web Site

http://www.nersc.gov/

NERSC Annual Report Editor

Kathy Kincade

Contributing Writers and Editors

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

Katie Antypas, Shane Canon, Shreyas Cholia, Brent Draney, Richard Gerber, Jason Hick, David Skinner, Jay Srinivasan, Francesca Verdier, Harvey Wasserman, Nick Wright
NERSC

Dan Krotz, Jon Weiner and Lynn Yarris
Berkeley Lab Communications Department

Sathya Achia Abraham
Public Affairs, Virginia Commonwealth University

Ron Walli
Communications Office, Oak Ridge National Laboratory

Design

Design, layout, illustrations, photography and printing coordination provided by the Creative Services Office of Berkeley Lab's Public Affairs Department.

CSO 14-NERSC-0295

DISCLAIMER

This document was prepared as an account of work sponsored by the United States Government. While this document is believed to contain correct information, neither the United States Government nor any agency thereof, nor The Regents of the University of California, nor any of their employees, makes any warranty, express or implied, or assumes any legal responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by its trade name, trademark, manufacturer, or otherwise, does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof, or The Regents of the University of California. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof, or The Regents of the University of California.

Ernest Orlando Lawrence Berkeley National Laboratory is an equal opportunity employer.

BERKELEY LAB

U.S. DEPARTMENT OF ENERGY

Office of Science