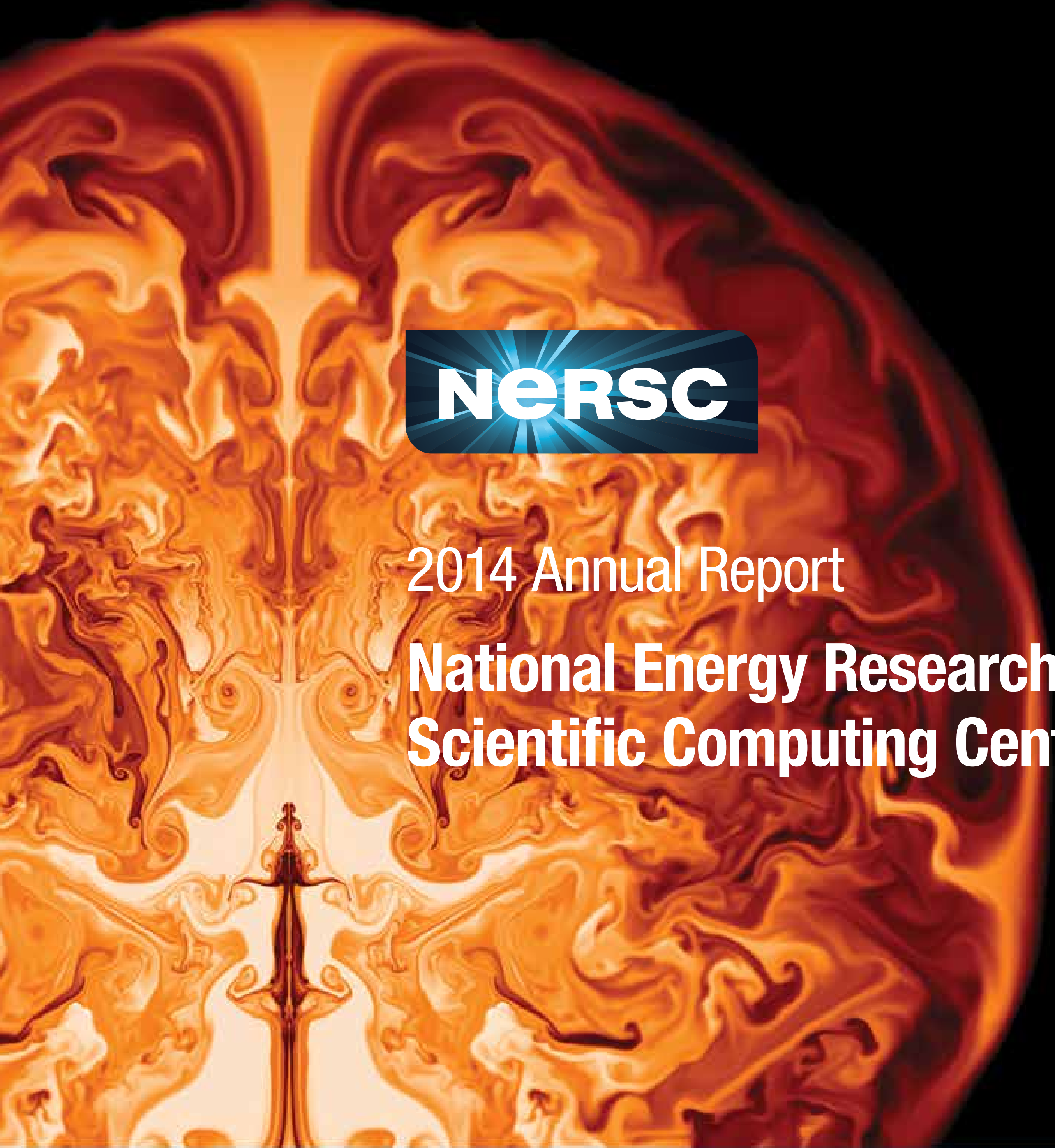
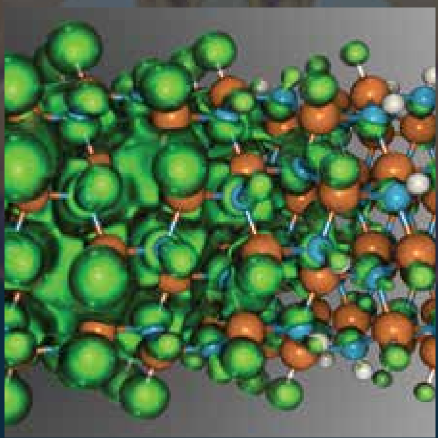
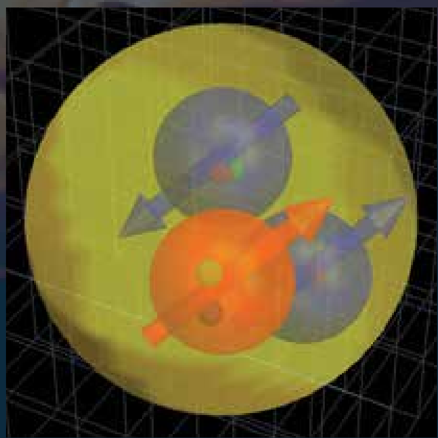


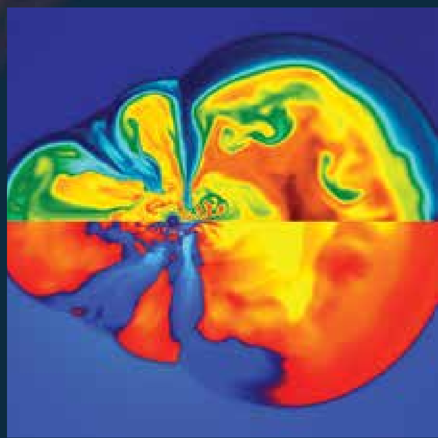
NERSC

2014 Annual Report

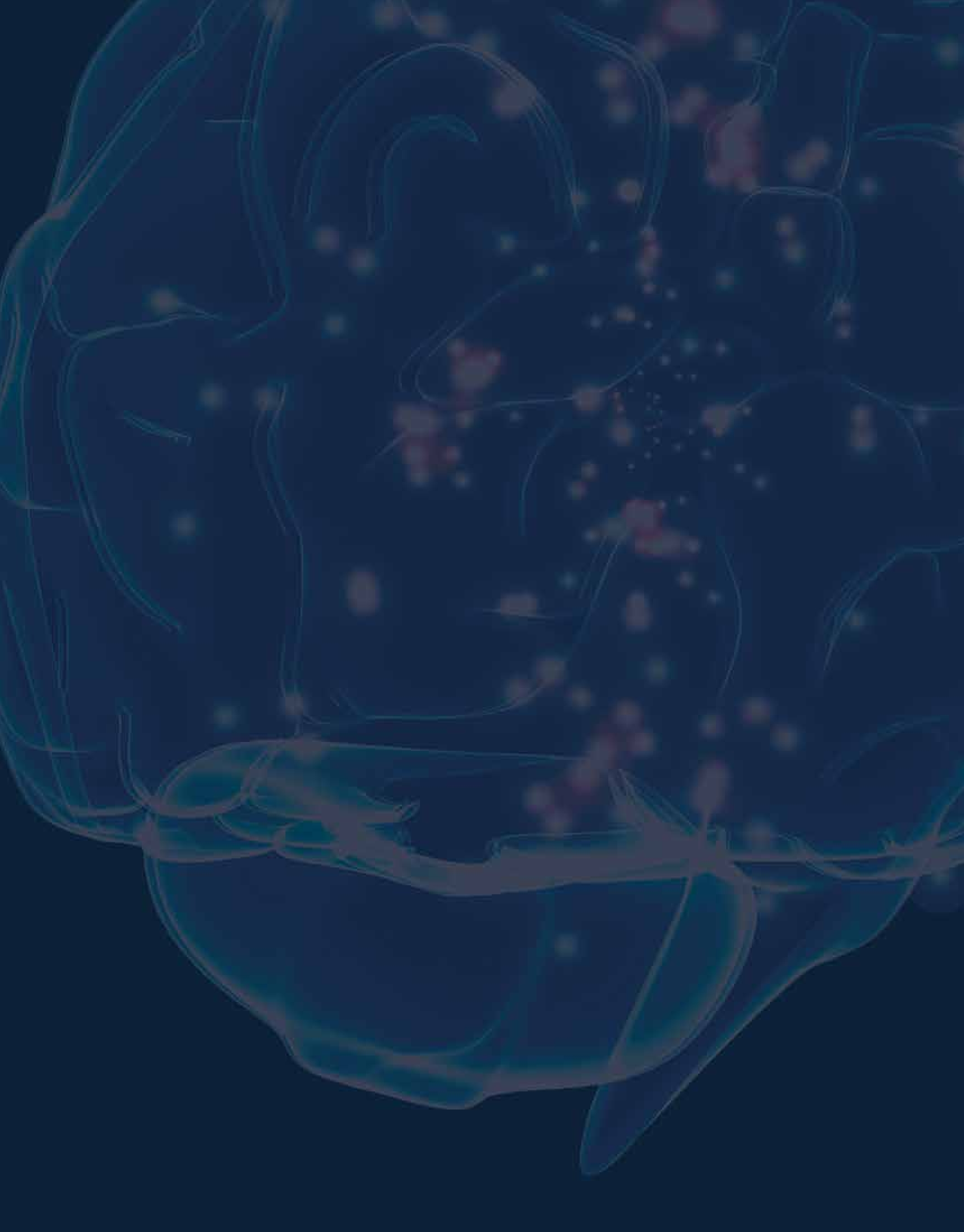
National Energy Research Scientific Computing Center







NERSC 2014 Annual Report



2014 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, MAIN IMAGE: Ken Chen, University of California, Santa Cruz (story, p. 34)

FRONT COVER, LEFT TO RIGHT: Burlen Loring, Lawrence Berkeley National Laboratory (story, p. 42); William Detmold, Massachusetts Institute of Technology (story, p. 36); Raphael Hix, Oak Ridge National Laboratory (story, p. 62)

BACK COVER, LEFT TO RIGHT: Matt Landreman, University of Maryland (story, p. 44); Matthew Emmett, Weiqun Zhang, Lawrence Berkeley National Laboratory (story, p. 50); European Space Agency (story, p. 46)

Table of Contents

The Year in Perspective

2014 was another exciting and productive year at the National Energy Research Scientific Computing Center (NERSC), the high-end scientific computing facility for the Department of Energy's Office of Science. We celebrated our 40th anniversary, saw our user community grow to nearly 6,000, brought our newest supercomputer into full production, signed a contract for our next-generation system, prepared for our move to the new Computational Research and Theory (CRT) facility on the main Berkeley Lab campus and laid the groundwork for our users to begin preparing for next-generation computing architectures.

NERSC's primary mission is to accelerate scientific discovery at the DOE Office of Science through high performance computing and data analysis. NERSC supports the largest and most diverse research community of any computing facility within the DOE complex, providing large-scale, state-of-the-art computing for DOE'S unclassified research programs in alternative energy sources, climate change, environmental science, materials research, astrophysics and other science areas related to the DOE mission.

Throughout 2014, NERSC ramped up its efforts to help users prepare for extreme scale computing and streamline their transition to the next generation of supercomputing. In January we brought into full production our newest computer: Edison, a Cray XC30 comprising 30 cabinets using Intel Ivy Bridge processors. Two months later we announced the contract for our next-generation supercomputer, a Cray XC40 scheduled to go into full production in 2016 that will be the first supercomputer in the new CRT facility. The XC40 system—named "Cori" in honor of bio-chemist and Nobel Laureate Gerty Cori, the first American woman to receive a Nobel Prize in science—is expected to deliver over 10 times the sustained computing capability of NERSC's Hopper system, a Cray XE6 that will be retired when Cori comes online.

The Cori procurement was done in collaboration with Los Alamos National Laboratory (LANL) and Sandia National Laboratories, whose Trinity system is very similar to Cori. We will continue to work with LANL and Sandia as we deploy both systems, and going forward the partnership has been named the Alliance for Application Performance at Extreme-scale, or APEX.

A key feature of Cori is the manycore Intel Xeon Phi processor Knights Landing, which introduces several technological advances, including higher intra-node parallelism; high-bandwidth, on-package memory; and longer hardware vector lengths. These enhanced features are expected to yield significant performance improvements for applications running on Cori.

In order to take advantage of the new features, however, application developers will need to make code modifications because many of today's applications are not optimized to take advantage of the manycore architecture and on-package memory. Thus, to help users transition to the new architecture, in 2014 NERSC established the NERSC Exascale Scientific Applications Program (NESAP). Through NESAP, several code projects are collaborating with NERSC, Cray and Intel with access to early hardware, special training and "deep dive" sessions with Intel and Cray staff. Eight of the chosen projects also will be working with a postdoctoral researcher to investigate computational science issues associated with manycore systems. The NESAP projects span a range of scientific fields—including astrophysics, genomics, materials science, climate and weather modeling, plasma fusion physics and accelerator science—and represent a significant portion of NERSC's current and projected computational workload.

NERSC also formed the Data and Analytics Services (DAS) Group, designed to help NERSC's users address data and analytics challenges arising from the increasing size and complexity of data from simulations and experiments. The DAS Group will play a key role in developing and implementing NERSC's strategic initiative of increasing the productivity, usability and impact of DOE users' data-intensive science.

Throughout 2014 NERSC continued preparations for moving to the new CRT facility on the main Berkeley Lab site. Funded jointly by DOE and the University of California, CRT has many innovative features designed to streamline NERSC workflows and improve overall efficiencies. It has 12.5 megawatts of power (upgradable to over 40 megawatts) and has been designed to be an extraordinarily energy-efficient building, including both ambient "free" cooling and the ability to reclaim heat from the computers to heat the building. The new machine room measures 20,000 square feet (upgradable to 30,000) and features a novel seismic isolation floor for the computers. In addition, moving from our current location in downtown Oakland, Calif., back up to the main Berkeley Lab campus means NERSC will once again be co-located with ESnet and the Computational Research Division. The move, which is expected to take place in mid-2015, will also make it is easier for us to collaborate with other divisions of the Laboratory and with the University of California, Berkeley.

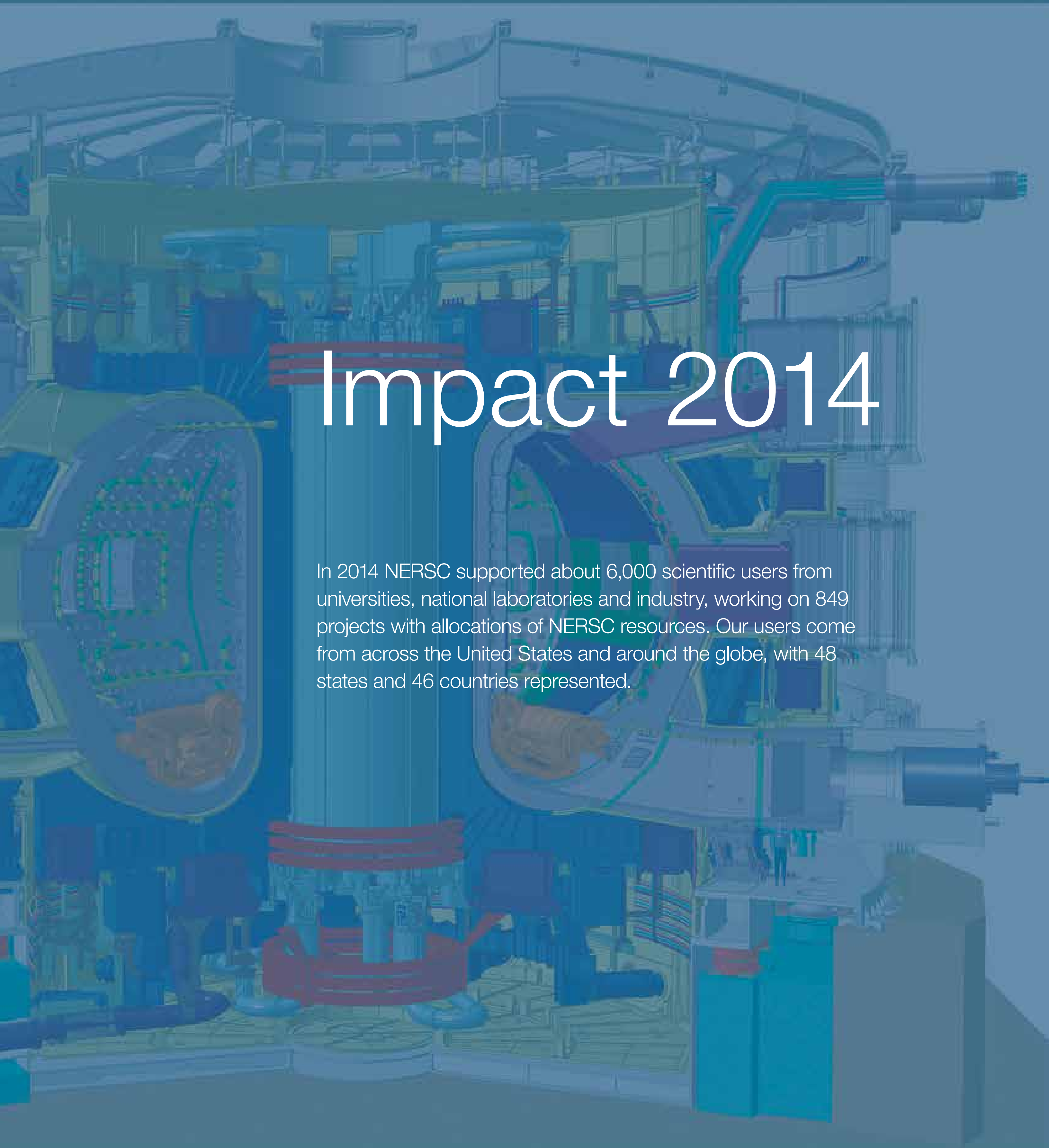
NERSC's 40th anniversary celebration provided the opportunity to highlight NERSC's contributions to high performance computing and scientific advances since its founding in 1974. Perhaps the most exciting aspect was a series of invited talks at Berkeley Lab featuring NERSC users who have won Nobel prizes: George Smoot (2006 Nobel Prize in Physics), Warren Washington (2007 Nobel Prize as a member of the Intergovernmental Panel on Climate Change), Saul Perlmutter (2011 Nobel Prize in Physics) and John Kuriyan (a member of Martin Karplus' group, 2013 Nobel Prize in Chemistry).

And 2014 was another banner year for scientific research at NERSC. Our users reported 1,808 refereed papers and 22 journal cover stories based on work performed at NERSC, once again reflecting the breadth and excellence of scientific research supported by the center.

Finally, NERSC continues to maintain an outstanding reputation for providing comprehensive scientific client services, as evidenced by our latest user survey. The Overall Satisfaction with NERSC score for 2014 was equal to the highest ever recorded in the 15 years the survey has been in its current form. The introduction of Edison, the upgrade of our border network to 100 gigabits per second and the increase in performance to up to 80 gigabytes per second for certain aplications in the globally accessible scratch file system all contributed to user productivity.

With all the advanced technologies we deploy, it is still people who make the difference. As always, I am grateful to our DOE Office of Science sponsors for their continued support, our users for the science they produce using NERSC resources and our amazing staff, who year after year make NERSC one of the most successful high performance computing centers within the DOE complex.

Sudip Dosanjh
NERSC Division Director

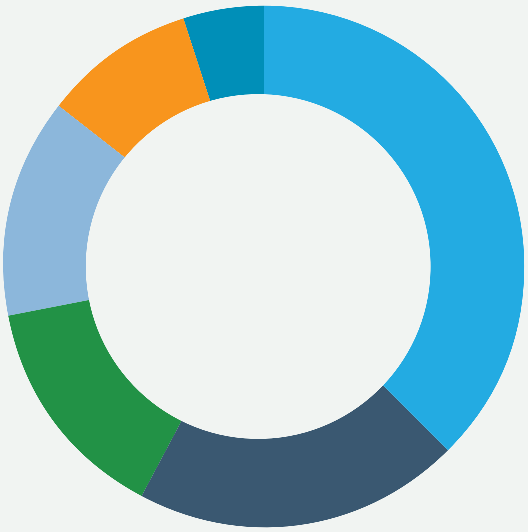


Impact 2014

In 2014 NERSC supported about 6,000 scientific users from universities, national laboratories and industry, working on 849 projects with allocations of NERSC resources. Our users come from across the United States and around the globe, with 48 states and 46 countries represented.

2014 NERSC Usage by DOE Program Office

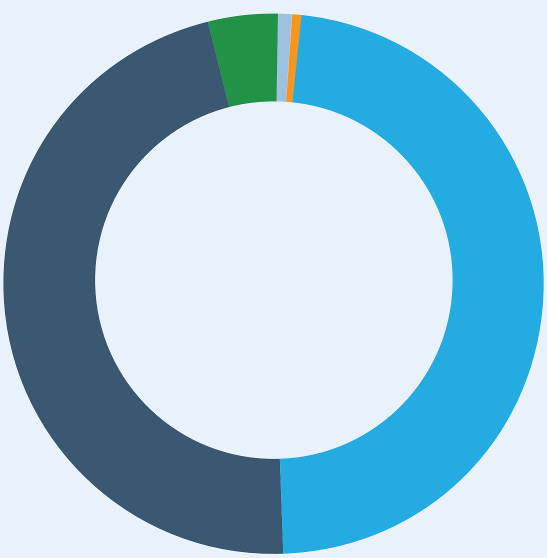
(MPP Hours in Millions)



BES	1,250
FES	677
HEP	474
BER	452
NP	320
ASCR	159

2014 NERSC Usage by Institution Type

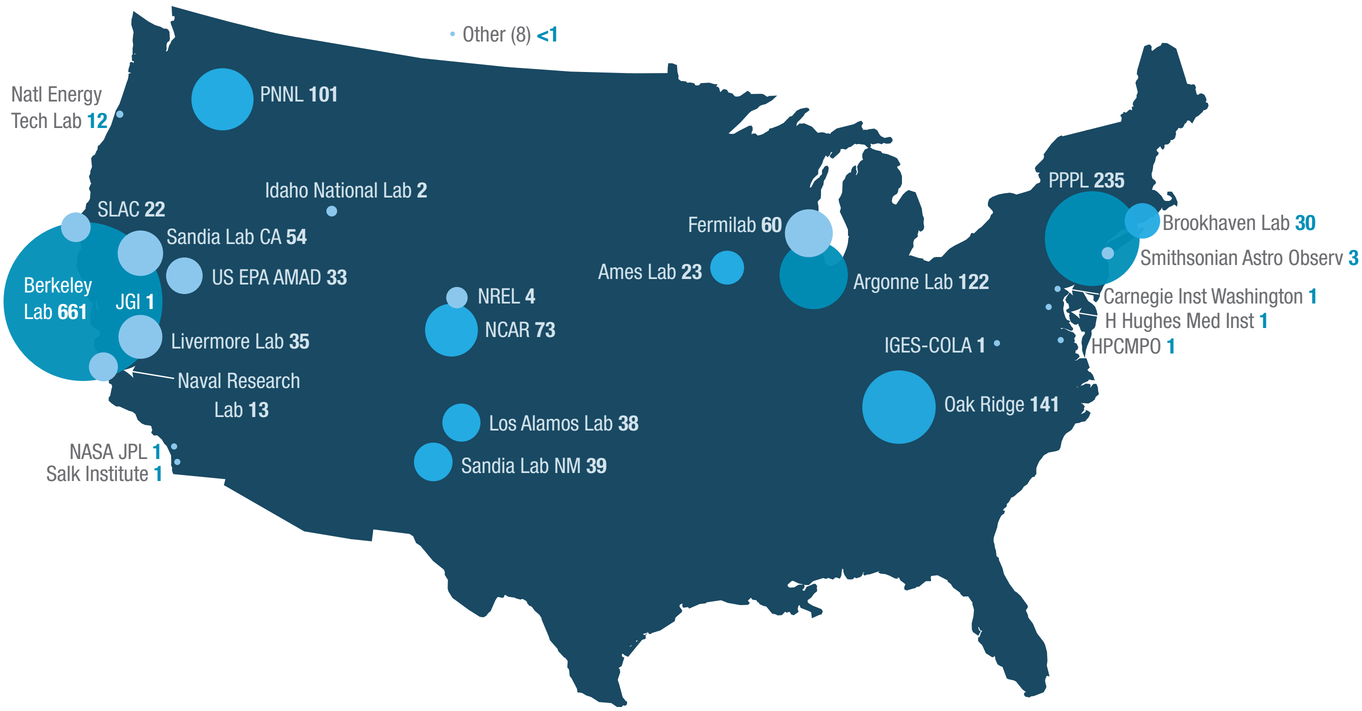
(MPP Hours in Millions)



Universities	1,607
DOE Labs	1,580
Other Government Labs	125
Industry	32
Nonprofits	3

2014 DOE & Other Lab Usage at NERSC

(MPP Hours in Millions)

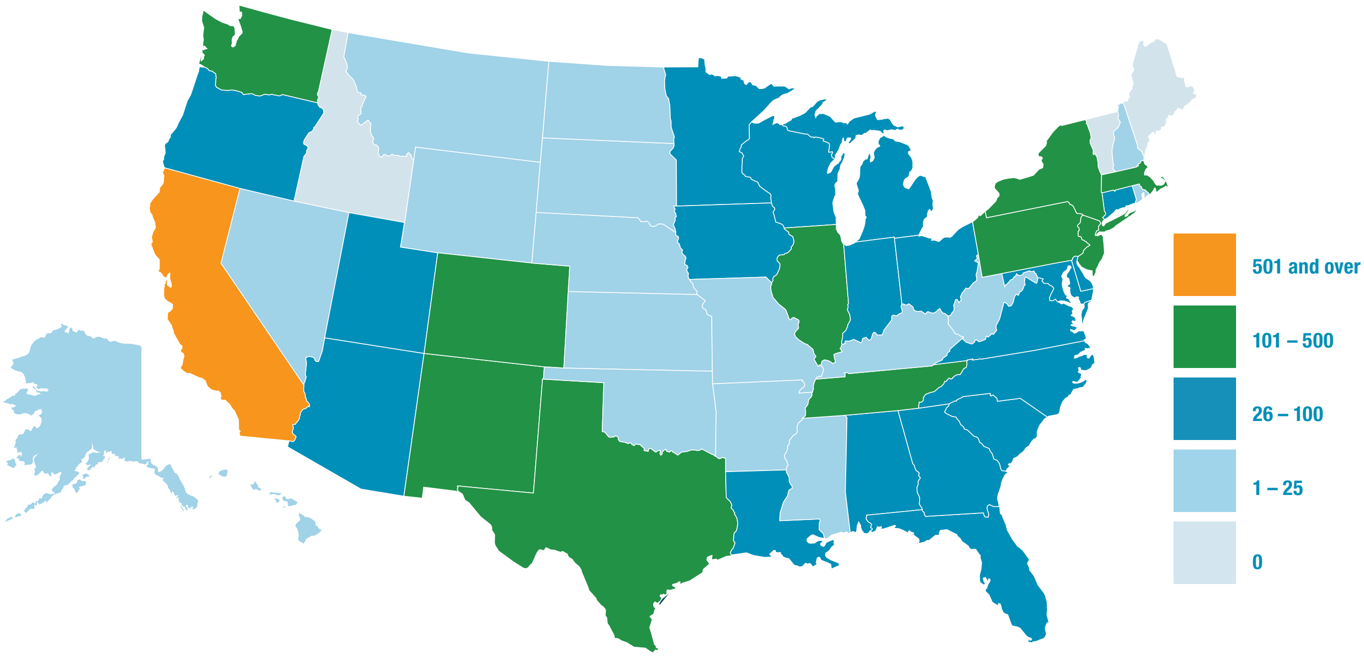


2014 Academic Usage at NERSC

(MPP Hours in Millions)

MPP Hours	Institution
158	MIT
129	UC San Diego
111	U. Washington
110	U. Arizona
63	UC Irvine
60	UC Berkeley
58	U. Colorado Boulder
53	Princeton Univ
49	Temple Univ
43	U. Kentucky
40	Cal Tech
38	Northwestern Univ U. Oklahoma
35	U. Illinois U-C
32	UCLA
29	U. Wisc. Madison
28	Cornell Univ U. Texas Austin
27	UC Davis
26	U. Michigan
24	U. Rochester Auburn Univ
23	Iowa State
21	U. Central Florida
16	Clemson Univ Tulane Univ
15	U. Penn
14	Rice Univ U. Maryland

2014 NERSC Users by State



California	1,885	New Jersey	110	Iowa	54	Arizona	20	Mississippi	6
Illinois	320	Maryland	90	Ohio	37	South Dakota	16	Rhode Island	6
Tennessee	220	Connecticut	71	District of Columbia	36	Kentucky	14	Hawaii	4
New York	207	Michigan	67	Minnesota	35	Kansas	13	Wyoming	4
Washington	187	North Carolina	62	Utah	34	Missouri	12	Nevada	3
Massachusetts	155	Florida	61	South Carolina	27	Oklahoma	11	Alaska	2
Texas	141	Indiana	61	Oregon	26	New Hampshire	10	Nebraska	2
Colorado	139	Georgia	59	Louisiana	25	West Virginia	10	Montana	1
Pennsylvania	132	Virginia	57	Alabama	23	North Dakota	9	Idaho	0
New Mexico	117	Wisconsin	57	Delaware	22	Arkansas	7	Maine	0

13 Colorado School Mines

12 UNC Chapel Hill
Penn State

11 Columbia Univ
U. Missouri KC

10 UC Santa Cruz
Vanderbilt Univ
U. Florida

9 U. Tennessee
William & Mary

8 UMass Amherst
Johns Hopkins Univ
U. South Carolina

7 U. Minnesota
SUNY Stony Brook
Stanford Univ
Indiana St U

6 Harvard Univ
U. Delaware
North Carolina State
Georgia State

5 UC Santa Barbara

4 Mississippi State
U. New Mexico
U. New Hampshire
Georgetown Univ
Louisiana State

3 Marquette Univ
U. Illinois Chicago
Florida State
U. Tulsa
N Dakota State Univ
U. Chicago
U. Utah
Kansas State
Rensselaer

2 Northeastern
Purdue Univ
Colorado State
LA Tech Univ
Yale Univ
Oregon State
New Mexico State
U. Notre Dame
U. Georgia
U. Houston
Dartmouth College
Illinois IT
Georgia Tech
U. Iowa
U.Toledo
Michigan State
U South Dakota

1 N. Illinois University
West Virginia Univ
Hebrew Univ Jerusalem
U. Sussex UK
New Jersey IT
Virginia Comm Univ
U. Alaska

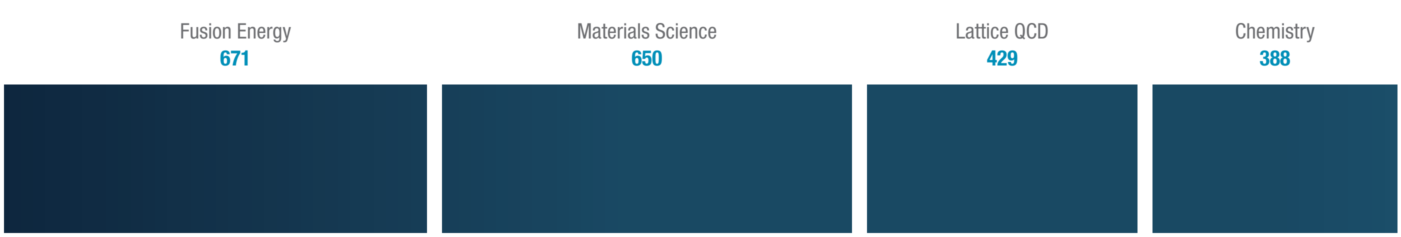
8 Other (42)

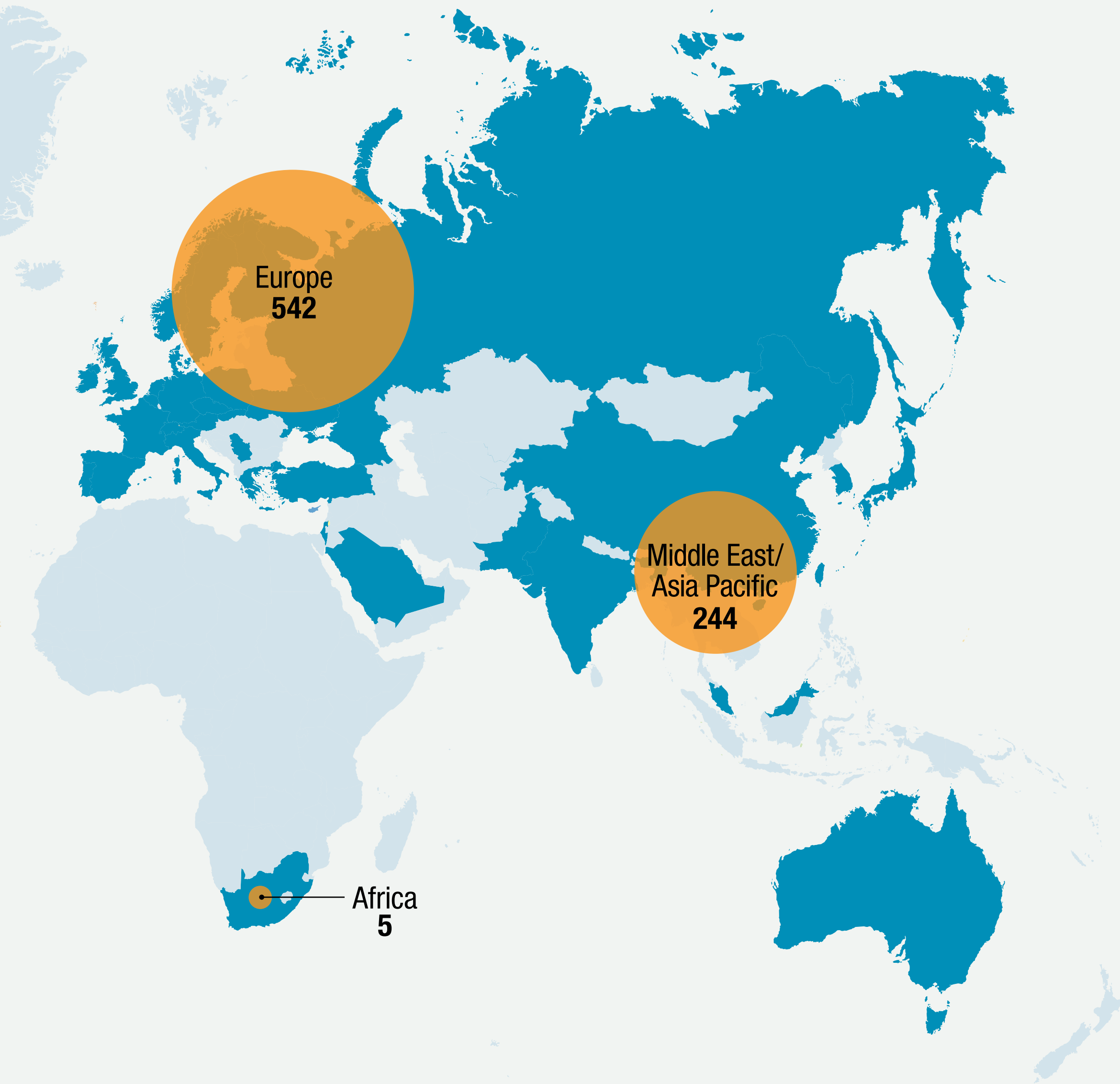
2014 NERSC Users by Country

North America:	**5,156**
United States of America	5,134
Canada	24
Puerto Rico	4
South America:	**22**
Brazil	12
Chile	4
Argentina	2
Colombia	2
Ecuador	1
Venezuela	1

2014 NERSC Usage By Discipline

(MPP Hours in Millions)

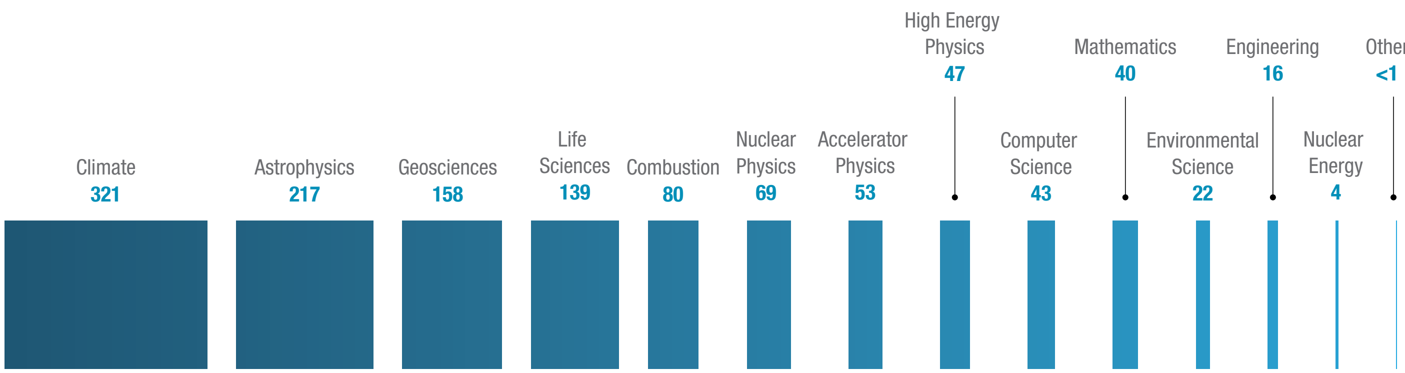




Europe:	542
United Kingdom	106
Germany	82
France	68
Italy	63
Switzerland	40
Spain	33
Denmark	32
Poland	22
Czech Republic	17
Russian Federation	13
Norway	11
Netherlands	9
Sweden	9
Finland	6
Greece	6
Cyprus	5
Portugal	4
Ireland	3
Turkey	3
Belgium	2
Scotland	2
Slovenia	2
Ukraine	2
Austria	1
Serbia	1

Africa:	5
South Africa	5

Middle East/ Asia Pacific:	244
People's Republic of China	107
Republic of Korea (South)	29
Japan	27
India	25
Israel	17
Australia	13
Taiwan, Province of China	12
Singapore	7
Saudi Arabia	5
Malaysia	1
Pakistan	1



Innovations

Throughout 2014 NERSC introduced a number of technology advances to help our users better capture, analyze, manage and store the increasing stream of scientific data coming out of experiments, simulations and instruments. In addition, with exascale supercomputing looming on the horizon, much of the roadmap is dotted with questions about hardware design and how to make these systems energy efficient enough so that centers can afford to run them. Thus in 2014 NERSC took a leadership role in helping users begin to make the transition to next-generation computing architectures.

The Path to Exascale

> **"By electing to deploy Intel Ivy Bridge processors with 1866Mhz memory, we were able to provide more than a 2X increase in performance per node compared with Hopper."**

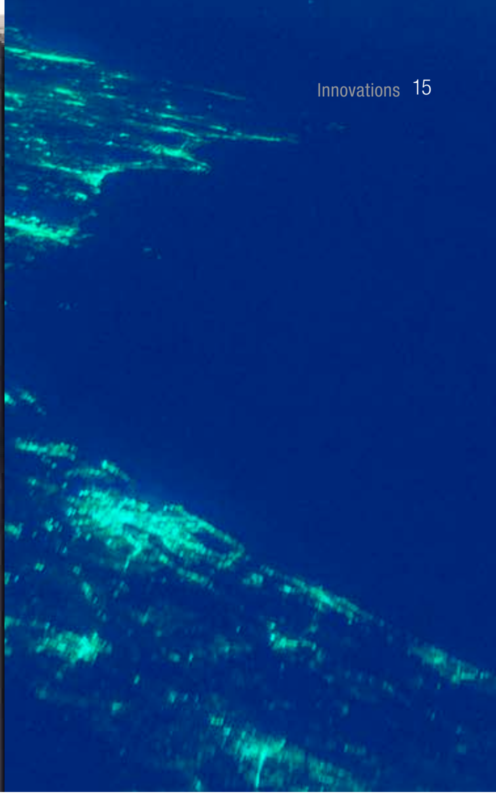
Edison Goes Into Full Production

In January NERSC brought into full production its newest computer: Edison, a Cray XC30. The Edison system has a peak performance of 2.57 petaflops/s, 133,824 compute cores, 357 terabytes of memory, and 7.4 petabytes of online disk storage with a peak I/O bandwidth of 164 GB/s. To enable this performance, Edison features a large number of high-performance compute nodes and a high-bandwidth, low-latency inter-node communications network that offers very high levels of sustained performance on real user codes. It incorporates Intel Ivy Bridge processors and the next-generation Aries interconnect, which uses a dragonfly topology instead of a torus.

A key decision in the procurement of Edison was whether to stick with a conventional Intel-architecture system for ease of adoption by users, or to include accelerators to increase peak performance. Given our broad and diverse application base, we felt it would have been extremely time consuming to port hundreds of applications to an accelerator architecture, with little payoff for most users. By electing to deploy Intel Ivy Bridge processors with 1866Mhz memory, we were able to more than double the increase in performance per node compared with Hopper. Some applications report as much as a 4X increase in performance per node due to the much higher memory bandwidth.

Edison also utilizes a novel water-cooling mechanism that can operate with much warmer water temperatures than earlier supercomputers. The Cray XC30's hybrid water/air cooling system, transverse airflow and high operating temperatures allow NERSC to use the moist, cool air that flows through the Golden Gate from the Pacific Ocean to help reduce cooling costs and increase energy efficiency. Using cooling towers only—without any mechanical refrigeration—Edison can be cooled for one-third the cost of earlier systems.

Edison was installed at the University of California Oakland Scientific Center (OSF) in Oakland, Calif. However, in 2015, NERSC and Edison will move to the new Computational Research and Theory (CRT) facility on the Berkeley Lab main campus in the hills of Berkeley. In order to be



installed in CRT, it was critical that Edison operate with the higher temperature water and air provided by the "free cooling" mechanisms that are key elements of the CRT design. The new cooling design of the XC30 meets this requirement. NERSC took advantage of this capability to modify OSF to mimic the CRT cooling mechanism and provide an early test of the CRT design. Not only was this test successful, we saved enough electricity to earn a $435,000 energy efficiency rebate from our power utility provider.

System Type	Cray XC30
CPU Type	Intel Ivy Bridge
CPU Speed (GHz)	2.4
Compute Nodes	5,576
Service Nodes	118
Cores per socket/node	12/24
Total Compute Cores	133,824
Flops/Core (Gflops/sec)	19.2
Peak Performance (Tflops/sec)	2,569.4
Memory per node	64 GiB
Aggregate Memory	349 TiB
Avg. Memory/Core	2.67 GiB
Memory Speed	1866 MHz
Memory Bandwidth per socket (STREAM)	105 GiB/s
Node Interconnect	Aries
Scratch Disk	7.56 PB (local) + 3.9 PB (global)
Disk Bandwidth	1043 GiB/s
Avg. Power (KW)	1,600

▶ NERSC's next-generation supercomputer, a Cray XC40, will be named after Gerty Cori, the first American woman to be honored with a Nobel Prize in science. She shared the 1947 Nobel Prize with her husband Carl (pictured) and Argentine physiologist Bernardo Houssay.

Cori Acceptance: Transitioning to Manycore

Throughout 2014, NERSC ramped up its efforts to help users prepare for extreme scale computing and streamline their transition to next-generation computing architectures.

In March 2014 NERSC announced the contract for our next-generation supercomputer, a Cray XC40 scheduled for delivery in mid-2016. The system—named "Cori" in honor of bio-chemist and Nobel Laureate Gerty Cori, the first American woman to receive a Nobel Prize in science—is expected to deliver over 10 times the sustained computing capability of NERSC's Hopper system, a Cray XE6.

A key feature of Cori is the manycore Intel Xeon Phi processor Knights Landing, which introduces several technological advances, including higher intra-node parallelism; high-bandwidth, on-package memory; and longer hardware vector lengths. These enhanced features are expected to yield significant performance improvements for applications running on Cori. For example, the extended hardware vector lengths and greater intra-node parallelism should help computations perform with greater efficiency and tackle ever-larger problems more quickly. And new high-bandwidth memories between main memory and cache offer the promise of much higher performance for memory bandwidth-limited codes, a situation becoming increasingly common for scientific computing applications as more cores are added to each processor.

In order to take advantage of the new features, however, application developers will need to make code modifications because many of today's applications are not optimized for the manycore architecture. To help users transition to the new architecture, in 2014 NERSC established the NERSC Exascale Scientific Applications Program (NESAP).

NESAP: Leading the Way in Code Optimization

Through NESAP, several code projects are collaborating with NERSC, Cray and Intel with access to early hardware, special training and "deep dive" sessions with Intel and Cray staff. Eight of the chosen projects also will be working with a postdoctoral researcher to investigate computational science issues associated with manycore systems.

The NESAP projects span a range of scientific fields—including astrophysics, genomics, materials science, climate and weather modeling, plasma fusion physics and accelerator science—and represent a significant portion of NERSC's current and projected computational workload. A variety of numerical methods and computational approaches are also represented, including particle-mesh algorithms, adaptive mesh refinement techniques, molecular dynamics, eigenvalue problems for complex molecular systems and *ab initio* computational chemistry methods. In addition, several codes are directly involved in DOE Scientific Discovery through Advanced Computing (SciDAC) activities.

Here is a list of the NESAP projects:

Advanced Scientific Computing Research (ASCR):

- **Optimization of the BoxLib Adaptive Mesh Refinement Framework for Scientific Application Codes,** Ann Almgren (Lawrence Berkeley National Laboratory)
- **High-Resolution CFD and Transport in Complex Geometries Using Chombo-Crunch,** David Trebotich (Lawrence Berkeley National Laboratory)

Biological and Environmental Research (BER)

- **CESM Global Climate Modeling,** John Dennis (National Center for Atmospheric Research)
- **High-Resolution Global Coupled Climate Simulation Using The Accelerated Climate Model for Energy (ACME),** Hans Johansen (Lawrence Berkeley National Laboratory)
- **Multi-Scale Ocean Simulation for Studying Global to Regional Climate Change,** Todd Ringler (Los Alamos National Laboratory)
- **Gromacs Molecular Dynamics (MD) Simulation for Bioenergy and Environmental Biosciences,** Jeremy C. Smith (Oak Ridge National Laboratory)
- **Meraculous, a Production de novo Genome Assembler for Energy-Related Genomics Problems,** Katherine Yelick (Lawrence Berkeley National Laboratory)

Basic Energy Science (BES)

- **Large-Scale Molecular Simulations with NWChem,** Eric Jon Bylaska (Pacific Northwest National Laboratory)
- **Parsec: A Scalable Computational Tool for Discovery and Design of Excited State Phenomena in Energy Materials,** James Chelikowsky (University of Texas, Austin)
- **BerkeleyGW: Massively Parallel Quasiparticle and Optical Properties Computation for Materials and Nanostructures**, Jack Deslippe (NERSC)
- **Materials Science using Quantum Espresso,** Paul Kent (Oak Ridge National Laboratory)
- **Large-Scale 3-D Geophysical Inverse Modeling of the Earth,** Greg Newman (Lawrence Berkeley National Laboratory)

Fusion Energy Sciences (FES)

- **Understanding Fusion Edge Physics Using the Global Gyrokinetic XGC1 Code,** Choong-Seock Chang (Princeton Plasma Physics Laboratory)
- **Addressing Non-Ideal Fusion Plasma Magnetohydrodynamics Using M3D-C1,** Stephen Jardin (Princeton Plasma Physics Laboratory)

High Energy Physics (HEP)

- **HACC (Hardware/Hybrid Accelerated Cosmology Code) for Extreme Scale Cosmology,** Salman Habib (Argonne National Laboratory)
- **The MILC Code Suite for Quantum Chromodynamics (QCD) Simulation and Analysis,** Doug Toussaint (University of Arizona)
- **Advanced Modeling of Particle Accelerators,** Jean-Luc Vay, Lawrence Berkeley National Laboratory)

Nuclear Physics (NP)

- **Domain Wall Fermions and Highly Improved Staggered Quarks for Lattice QCD,** Norman Christ (Columbia University) and Frithjof Karsch (Brookhaven National Laboratory)
- **Chroma Lattice QCD Code Suite,** Balint Joo (Jefferson National Accelerator Facility)
- **Weakly Bound and Resonant States in Light Isotope Chains Using MFDn—Many Fermion Dynamics Nuclear Physics,** James Vary and Pieter Maris (Iowa State University)

NERSC is also working with other national labs to help coordinate strategies for application readiness and platform portability across the Office of Science. Similar to NESAP, each of the leadership computing facilities will launch its own application readiness and platform portability efforts aimed at preparing application codes for next generation systems. In September 2014, NERSC hosted a meeting that included staff from the Oak Ridge and Argonne Leadership Computing Facilities and Sandia, Los Alamos and Lawrence Livermore national laboratories. Each facility presented an overview of its expected supercomputing architecture and programming model, followed by an open discussion on application and performance portability across architectures. The attendees committed to sharing best practices for application readiness, platform portability and user training.

Intel Center of Excellence

In 2014 NERSC and Cray established a joint Center of Excellence to help users port and optimize target applications that will run on Cori, and Intel named Berkeley Lab an Intel Parallel Computing Center (IPCC). The Berkeley Lab IPCC will be led by Nick Wright of NERSC and Bert de Jong and Hans Johansen of the Computational Research Division.

The Berkeley Lab IPCC is aimed at adapting existing scientific applications to run on future supercomputers built with manycore processors. The IPCC will focus on increasing the parallelism of two applications: NWChem, a leading application for computational chemistry, and CAM5, part of the Community Earth System Model that is widely used for studying global climate. The goal is to deliver enhanced versions of NWChem and CAM5 that at least double their overall performance on manycore machines of today. The research and development will be focused on implementing greater amounts of parallelism in the codes, starting with simple modifications such as adding or modifying existing components and going as far as exploring new algorithmic approaches that can better exploit manycore architectures.

Modernizing these codes to run on manycore architectures will enable the scientific community to pursue new frontiers in the fields of chemistry, materials and climate research. Because both NWChem and CAM5 are open source applications, any improvements made to them will be shared with the broader user community, maximizing the benefits of the project.

◂ Berkeley Lab's Intel Parallel Computing Center will be led by Bert de Jong (left), Hans Johansen (not pictured) and Nick Wright (right).

▶ The new Computational Research and Theory facility, currently under construction on the main Berkeley Lab campus, will open in 2015.

Preparing for the Move to CRT

Throughout 2014 NERSC continued its preparations for relocating to the new Computational Research and Theory (CRT) facility, located on the main Berkeley Lab campus, in 2015. Moving from our current location in downtown Oakland back to the main Berkeley Lab campus means NERSC will once again be co-located with ESnet and the Computational Research Division. The move will also enhance collaborations with other divisions of the Laboratory and with the University of California campus.

CRT, which was funded jointly by DOE and the University of California, has many innovative features designed to streamline NERSC workflows and improve efficiencies. 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. CRT has 12.5 megawatts of power (upgradable to over 40 megawatts) and has been designed to be an extraordinarily energy efficient building, including both ambient "free" cooling and the ability to reclaim heat from the computers to heat the building. The new machine room measures 20,000 square feet (upgradable to 30,000).

The Bay Area is a well-known seismically active area. The CRT was thus designed to include a seismically isolated data-center floor to protect the computer systems and the people working in the vicinity. In this design, developed with Dynamic Isolation Systems (DIS), a conventional raised floor is mounted on a steel substructure that rolls on casters. While based on an earlier system developed by DIS for containerized data centers, the CRT isolated floor is a first-of-a-kind design. Computer cabinets sit on the raised floor. In the event of an earthquake, a combination of the inertia of the floor (including the computer systems riding on top) and springs dampen accelerations and vibrations from the earthquake, thus mitigating damage. The floor structure is designed to handle a maximum credible event occurring on the nearby Hayward Fault.

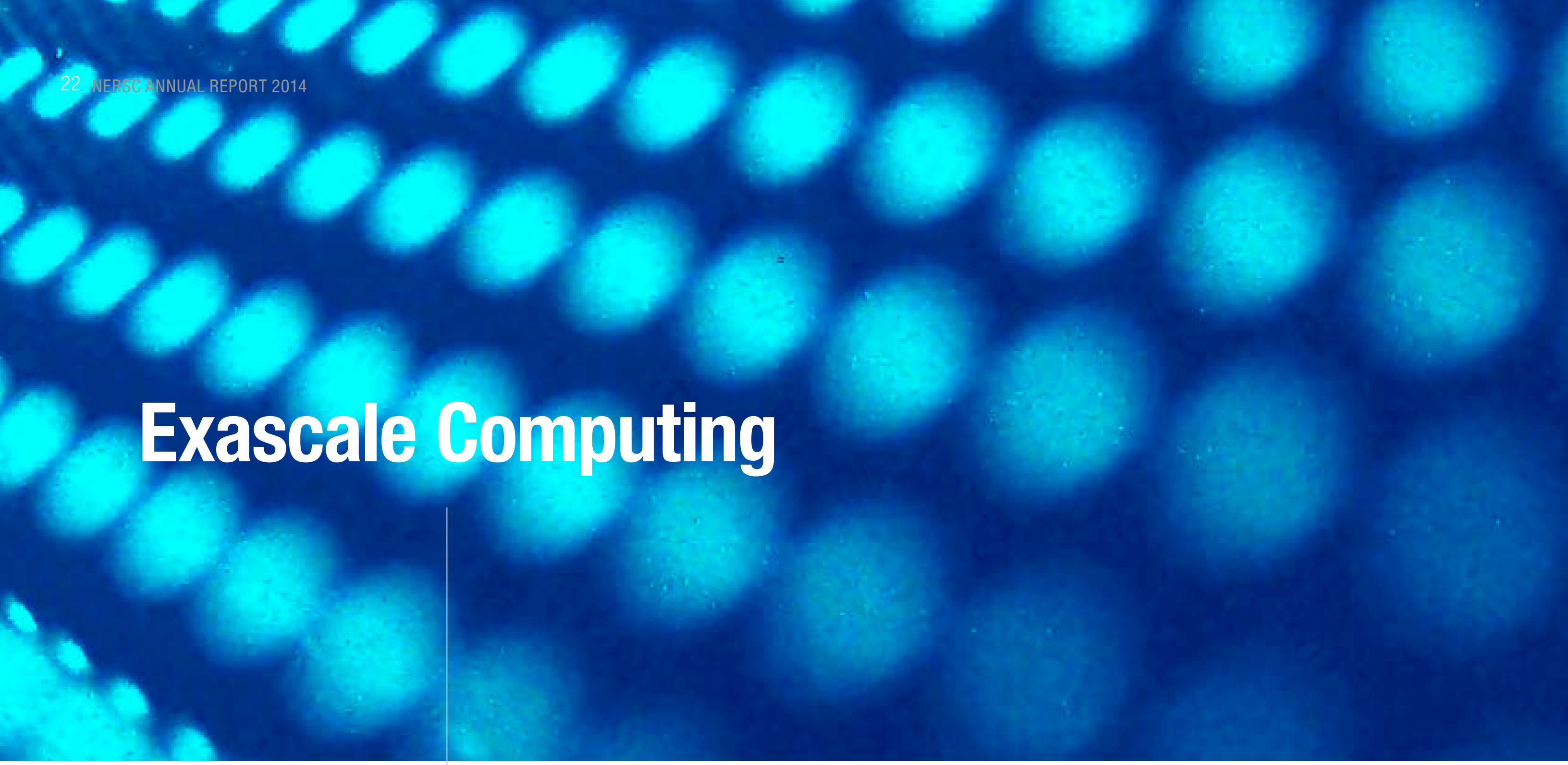
To determine if the floor would function correctly, NERSC and Berkeley Lab leveraged our connections with UC Berkeley to arrange for at-scale testing by DIS at the UCB Pacific Earthquake Engineering Research Center, the world's largest 3D shake table. There were two primary goals for this testing:

1. Study the performance of the isolated floor for seven earthquake motions under three different loading configurations: Fully loaded, half-loaded and unbalanced-loading. Performance parameters studied were overall stability of the system under earthquake loading, ease of movement of the isolation system, amount of movement, stresses in various framing members and reduction of earthquake accelerations by the isolation system.

2. Performance evaluation of components above the DIS Floor Isolation System. These include the tile-stringer-pedestal system that directly connects to the DIS Floor Isolation System and the computer racks that are placed on top of the tiles. Performance parameters studies were overall integrity of the tile-stringer-pedestal system under gravity and seismic loads, stability of the computer racks in anchored and unanchored conditions, accelerations on top of the computer racks and overall system performance.

A test structure was run through three days of testing equivalent to approximately 3,000 years of earthquakes. Overall, the test showed that the isolated floor system worked well. The racks sustained only minor damage despite undergoing dozens of tests, including multiple simulated Richter 7+ events. A few minor design deficiencies were identified. Installation of the isolated floor began in fall of 2014 after the design changes were implemented. CRT completion is anticipated for late spring 2015, and the first computer systems will be installed on the floor soon thereafter.

◀ The CRT facility features a novel seismic isolation floor in the machine room to protect the computers in the advent of an earthquake. The raised floor is mounted on a steel substructure that rolls on casters.



Exascale Computing

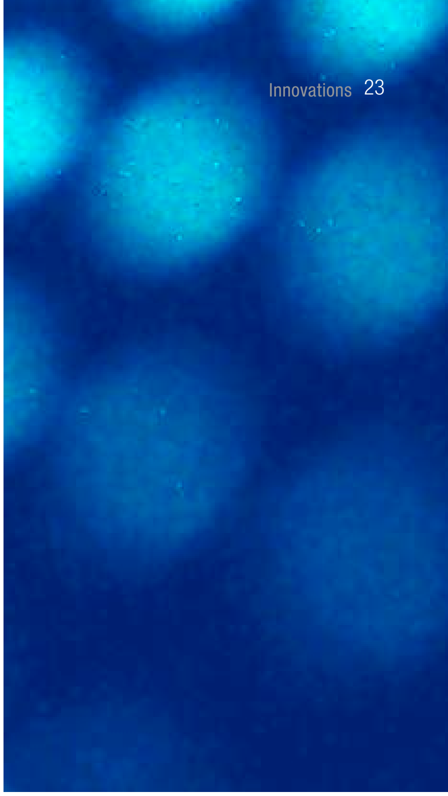
“Of special interest to NERSC users are enhancements that allow them to benefit from the burst buffer without having to modify their code.”

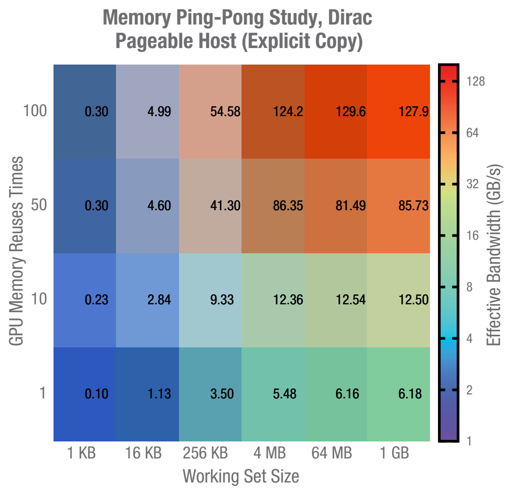
Alternative Use Cases for Burst Buffer

Included in the request for proposals for the NERSC-8 (Cori) system was a requirement for a “burst buffer,” a layer of NVRAM flash that sits between memory and disk and serves to accelerate I/O application performance. As part of the burst buffer activities at NERSC, we entered into a non-recurring engineering (NRE) project with Cray, the winning vendor, and our partners at Sandia and Los Alamos national laboratories.

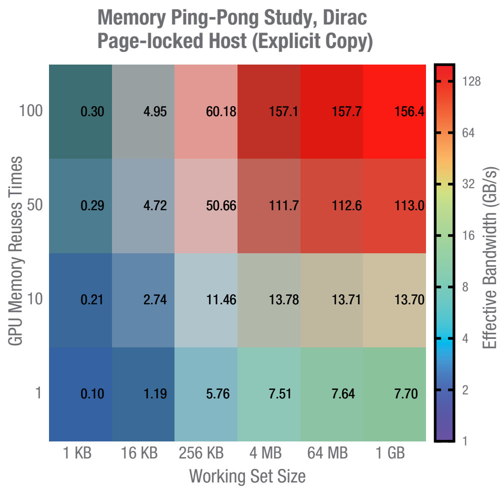
To define the burst buffer features that would most benefit our users beyond checkpoint restart, NERSC undertook an extensive requirements gathering exercise and found that NERSC users are interested in enhanced I/O, in terms of both performance and usability. Of special interest to them are enhancements that allow them to benefit from the burst buffer without having to modify their code. Thus we requested that the burst buffer support a caching mode, where data is automatically migrated to and from the burst buffer on the user’s behalf. This feature request was placed into the NRE contract and will be delivered in the summer of 2016.

NERSC also consulted with members of the Computational Research Division at Berkeley Lab to help determine which enhancements of burst buffer technology we could enable to allow the exploration of advanced concepts likely to impact future machines. Our discussions determined that enabling “compute near memory” features of the burst buffer would allow for the exploration of advanced features likely to be important for exascale-class machines. The basic idea here is to facilitate the use of the processor on the burst buffer node itself to filter, index, annotate or otherwise process data while it is in flight to the burst buffer flash file system. This feature request was placed into the NRE contract and will be delivered in the summer of 2016. This will enable users and researchers to experiment and prototype concepts that will enhance user productivity and scientific output.

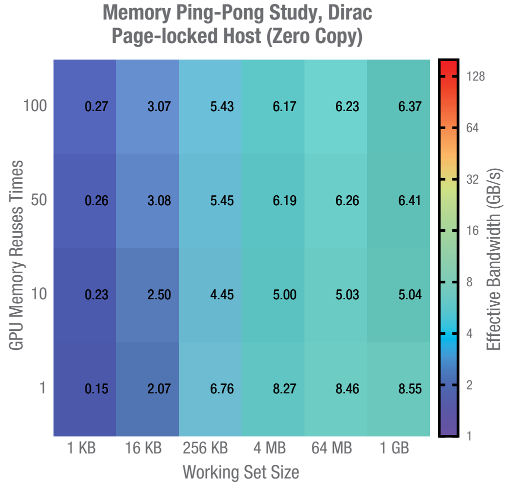




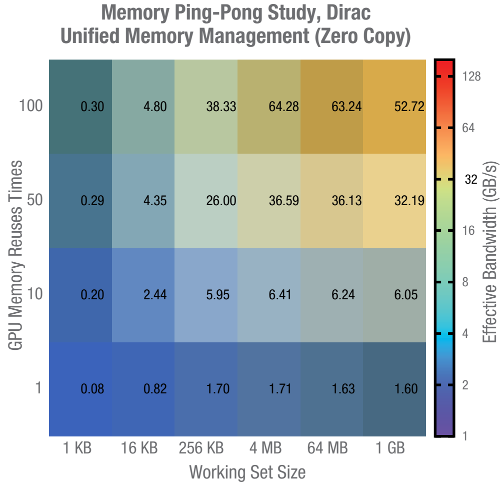
(a) Pageable host with explicit copy between CPU and GPU



(b) Page-locked host with explicit copy between CPU and GPU



(c) Page-locked host with zero copy



(d) Unified (managed) memory

◀ The Empirical Roofline Toolkit, showing effective bandwidth as a function of the Babbage GPU temporal locality (reuse) and working set size for four different GPU memory management mechanisms.

Roofline Tool

Another collaboration between NERSC and Berkeley Lab's Computational Research Division focused on developing the Empirical Roofline Toolkit (ERT) for practical architecture and program analysis. The growing complexity of HPC architectures makes it difficult for users to achieve sustained application performance across different architectures. In addition, quantifying the theoretical performance and the resulting gap between theoretical and observed performance is becoming increasingly difficult.

Performance models and tools that facilitate this process are crucial. A prime example is the Roofline Model, which combines arithmetic intensity, memory performance and floating-point performance together in a two-dimensional graph using bound and bottleneck analysis. Unfortunately, this process requires expert users to manually create and leverage this modeling approach. Thus the goal of ERT is to create an automatic and user-friendly approach to leverage Roofline modeling technology.

We initially constructed the Roofline model using an automated characterization engine. Subsequently, we extended the Roofline formalism to address the emerging challenges associated with accelerated architectures and constructed three benchmarks designed to drive empirical Roofline-based analysis. The first two represent the conventional memory hierarchy bandwidth and floating-point computation aspects of the Roofline. The third is a novel and visually intuitive approach to analyzing the importance of locality on accelerated architectures like GPUs. It quantifies the performance relationship between explicitly and implicitly managed spatial and temporal locality on a GPU.

To understand the impact of these benchmarks, we evaluated these experiments on four platforms: Edison (Intel Xeon CPU), Mira (IBM BlueGene/Q), Babbage (coprocessor only, Intel MIC Knights Corner) and Titan (GPU only, NVIDIA Tesla K20x). Additionally, we used the resulting empirical Rooflines to conduct preliminary analysis on three HPC benchmarks: HPGMG-FV, GTC and miniDFT.

Our future goals include publicly releasing the ERT coupled with an integrated visualizer and enabling database support to track the performance characteristics of evolving architectures and software stack implementations. Overall, we believe ERT has the potential to effectively guide important coding tradeoff design decisions to maximize programmer and performance efficiency.

Exascale Archiving

NERSC is contributing to the development of specific features designed to take its High Performance Storage System (HPSS) into the exascale era of computing. In 2014, NERSC developed the first of several versions of HPSS software (to be delivered in 2015) that parallelize operations, toward the eventual goal of having physically separate core storage servers to scale both metadata and data capabilities by orders of magnitude.

The HPSS software collaboration, a joint venture between IBM and five DOE national laboratories (including NERSC/LBNL), recently renewed its plans to provide new major releases approximately every two years. The major releases typically contain features at system-scale that change the way DOE users archive their HPC data. Examples of such features include:

- Partitioned metadata servers, enabling the mapping of different parts of the HPSS global namespace to different servers
- End-to-end checksums for user data, enabling users to store or create checksums with HPSS files and validate the checksum upon data retrieval
- Undelete operations, to enable users to undo delete operations on HPSS files.

Usability is a major factor of exascale success that NERSC needs to continue improving in HPSS through our software development collaboration. NERSC was first to develop gridFTP and Globus Online endpoints with its HPSS systems, thus enabling scientists to use these important software tools in managing their data across multiple DOE sites. However, much work remains to productionize this critical data storage infrastructure. NERSC funds HPSS development to ensure it scales to meet the demand of our largest computational systems. Through our collaborative development efforts, NERSC continues to support exponential growth of about 50 percent more archived data per year by NERSC users.

Data-Intensive Computing

Science Data Pilot Demos

In 2014, NERSC—along with several other Office of Science labs—participated in a series of demonstrations to explore how enhanced data capabilities can enable science. Demos were carried out in multiple science areas and made use of resources and services from multiple facilities. Many of the demos coupled experimental user facilities such as Berkeley Lab's Advanced Light Source and other light sources with advanced networking, computing and storage. NERSC contributed to these demos by offering resources and staff expertise and also prototyped new capabilities. Two examples of these prototype capabilities were a data replication service and portable execution through user-defined images.

"NERSC contributed to these demos by offering resources and staff expertise and also prototyped new capabilities."

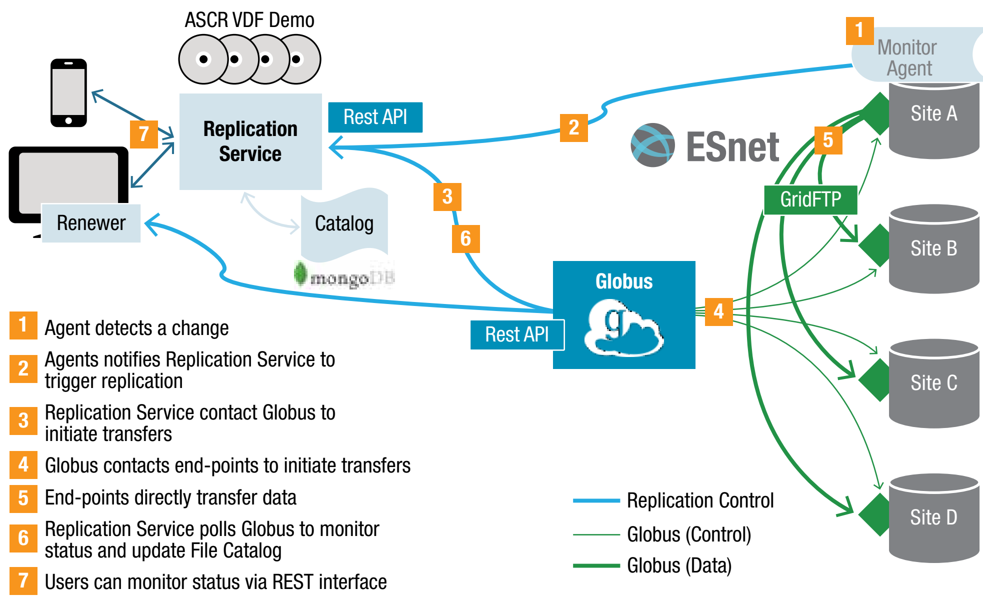
Data Replication Service

Data replication is seen as a core capability of a distributed data facility. Data replication provides a foundation for a robust data access service and for performing distributed data analysis. NERSC led an effort to develop a prototype data replication service for the data demonstrations.

For this demonstration, data end points were established at Argonne, Brookhaven, Lawrence Berkeley, Oak Ridge and Pacific Northwest national laboratories, and the service demonstrated data sets were replicated automatically from one site to the other four. The service also provided a metadata service that could be used to create a data catalogue. While this prototype service was only intended to be a proof of principle, it followed many of the guiding principles that will be important in creating a virtual data facility and illustrates the potential value of such a facility. For example, the service was provided via a RESTful interface and leveraged existing services in a layered architecture that will enable the service to be easily integrated into other applications and frameworks.

Portable Execution with Docker

Quickly and reliably instantiating a functional user environment for carrying out simulation and analysis is a major challenge in data-intensive computing. This affects productivity because it can



▸ NERSC's new Data Replication Service offers robust data access and distributed data analysis.

often be difficult for users to build and deploy all of the applications needed for a particular workflow at a new facility. These applications may have a large set of dependencies that are challenging to satisfy on a new target system. It also has ramifications for the scientific process because scientists need to know that the applications, tools and libraries are functioning correctly for reproducibility.

One way to address this challenge is to allow scientists and scientific communities to create reference images that they can test and certify, and then enable users to easily instantiate these images on various computer systems. As part of the 2014 data demonstrations, NERSC explored the use of Docker to realize this goal.

Docker is an emerging platform that allows users to easily develop, build and ship images that can be quickly instantiated on many systems. Docker builds heavily on Linux containers, which are similar to virtual machines since they allow process compartmentalization and custom environments; however, containers are lighter weight since they don't use full system-level virtualization. Instead, processes running on a shared system execute within the same kernel but have unique views of certain kernel-level data structures to provide the illusion of a dedicated system. Containers can more efficiently access network and file system resources, which provides higher performance for many data-intensive workloads compared with virtual-machine-based approaches.

Docker has several characteristics that make it difficult to run out of the box in a secure manner in a shared user facility like NERSC. Thus NERSC made modifications to Docker to allow users to instantiate containers in a secure manner while still providing flexibility. The modifications force processes to run as the user (versus a privileged user like root) and provide easy access to the user's home directory and global scratch space. This means that Docker processes have access to the same data as processes running on non-Dockerized systems. This allows a user to run a large-scale application on a system like Edison, then use Docker to perform analysis or other operations that may require more customized software.

Recognizing the value of Docker, NERSC has engaged with vendors such as Cray to explore how these approaches can move beyond the testbed phase and be applied to large-scale systems like Edison and the future Cori system.

Robinhood

Managing storage resources like scratch space on large systems is critical to ensuring that the systems can effectively meet user demands. NERSC purges local scratch file systems to ensure end user usability and availability as well as file system reliability. This is accomplished through quotas and by removing files that are older than a specified period. Prior large-scale systems at

NERSC utilized Lustre file systems comprising servers directly connected to Lustre storage nodes. As such, NERSC had full access to the Lustre storage nodes, allowing us to develop programs using the tools and libraries within the Lustre environment/infrastructure.

Our latest file systems use Cray's Sonexion Lustre storage nodes, which limit access to the underlying Lustre environment because they are designed to be appliance-like. Since APIs and libraries previously used for purging were not available on Sonexion, it was necessary for NERSC to develop new purge mechanisms. To help address this, Cray provided Robinhood, which collects file system metadata into a database that can be queried for either data or metadata operations without adversely affecting file system performance.

However, using Robinhood alone wasn't sufficient since NERSC prefers to perform additional verification to ensure files are not purged unnecessarily. For instance, if a user downloaded and untarred a set of files, the dates on the files may be older than the purge age, but if they were purged they would need to be downloaded again. To deal with these issues, Cray and NERSC created tools that leverage the Robinhood database but handle the purge as a separate step to ensure NERSC's purge policies are maintained.

Portal for Data Analysis and Cosmological Simulations

Accessing and analyzing data from cosmological simulations is a major challenge due to the large size (hundreds of terabytes) of the datasets and the diversity of the associated large-scale analysis tasks. Analysis of the simulated models requires direct access to the datasets, considerable compute infrastructure and storage capacity for the results.

To help users address these needs, NERSC collaborated with Argonne National Laboratory to develop the Portal for Data Analysis services for Cosmological Simulations (PDACS), a web-based workflow service and scientific analysis platform. PDACS provides access to shared repositories for datasets, data transfer and analysis tools, cosmological workflows and the infrastructure required to perform a variety of serial and large-scale parallel analyses. It is a repurposed implementation of the Galaxy workflow engine—an open, web-based platform for data-intensive biomedical research—that supports a rich collection of cosmology-specific data types and tools.

Working closely with engineers at NERSC, the PDACS team successfully leveraged the Galaxy framework and adapted it for cosmological workflows. Galaxy supports the assembly of complex, multi-step workflows and automates their execution over an abstracted infrastructure. PDACS provides the ability to use and share existing cosmology analysis tools and link them in a workflow with coherent metadata propagation and provenance tracking. Particular features of PDACS include:

- **New cosmology analysis tools:** PDACS provides a set of frequently used cosmology tools: Friends-of-Friends and Spherical Overdensity halo finders, halo profile measurements, concentration measurements, binning routines to calculate mass functions from the halo finder outputs, a concentration-mass relation emulator, a mass power spectrum emulator and measurement tools for two-point statistics (the fluctuation power spectrum and correlation function). Computationally intensive tools, such as the halo finder, are automatically submitted as a batch job to be run on supercomputing resources, while less-intensive tools, such as the emulators, are executed on the same node that manages the Galaxy instance. One of the goals of PDACS is to provide a platform that encourages the community to contribute new tools and promote code reuse.
- **New data types for metadata propagation:** Galaxy can recognize data that are introduced into it. This is accomplished by a combination of what Galaxy calls the data type of the file and the format of the file. PDACS provides more sophisticated data types and a complex hierarchy of formats to support cosmological workflows and ensure that the user of a given data file can identify the application that created the file, all the inputs that were used to generate the file and the set of configuration parameters used by the application that created it.

- **Data access—upload and share:** Currently, PDACS exposes the set of 37 cosmological models known as the Coyote Universe. Users can upload and share their own datasets with the PDACS community.
- **Job submission: O**n NERSC, PDACS leverages the NEWT API for user authentication and dispatching jobs. Authenticating users with NERSC credentials enables jobs and workflows that are dispatched through PDACS to be associated with a NERSC user, allowing resource accountability and the enforcement of compute and data restrictions.
- **Graphing/plotting:** Visualization is generally the final step of a workflow. PDACS supports the ability to plot multiple columns from the same, or multiple, datasets. Multiple visualization tools have been integrated into the PDACS platform, including a JavaScript-compiled Gnuplot, providing powerful graphical support to users, and a Shiny-based application as an interactive visualization tool. Shiny is a web application framework for R. Our Shiny server runs alongside the PDACS instance on the host node, permitting immediate access to data.

SPOT: XMAS, VisIt on the Web

SPOT is a data-analysis framework created by staff at NERSC and Berkeley Lab's Computational Research Division. It is a portal for light-source data analysis powered by NERSC's compute and data services. As of early 2014, the framework included more than 173,000 datasets (over 1 PB) from 153 users. In the past year, we have made multiple enhancements to this framework.

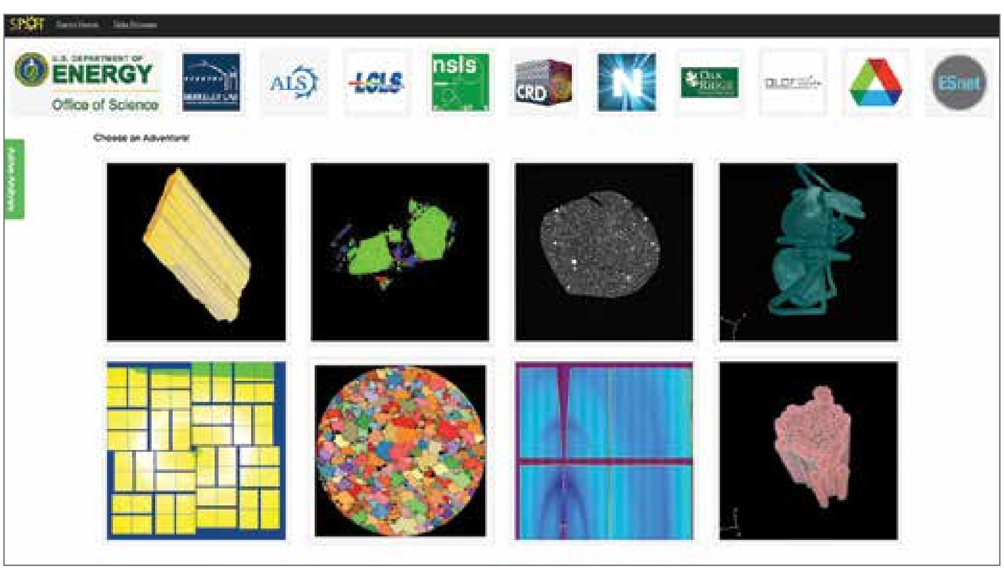
XMAS

SPOT Suite continues to expand its capabilities into additional research techniques, such as X-ray microdiffraction. With this technique, soft X-rays from the Advanced Light Source (ALS) at Berkeley Lab are used to generate a series of diffraction images, one at each point in a fine array rastered over a sample of interest. Once analyzed as a single image, the collection of pixel-by-pixel data gives rich, spatially resolved information about the sample's structure. Processing the data from individual pixels into a single image is most efficiently done via parallel computation. The X-ray Microdiffraction Analysis Software (XMAS), developed at the ALS, enables processing of microdiffraction images in parallel on high-performance systems.

In 2014 NERSC worked with ALS staff at Beamline 12.3.2 to run XMAS analysis jobs via the web. The work focused on two pieces of the puzzle: writing middleware code to serve as a robust interface between the XMAS Fortran code, the NERSC HPC systems and the web front end; and building web-based controls for customizing XMAS jobs and viewing results. Having already achieved a proof-of-concept—triggering customized XMAS jobs via the web—we hope to enable as many of the base code's features as possible via the SPOT interface.

Multi-Facility Use

SPOT has also been expanded, again through collaboration with researchers in the Computational Research Division, to incorporate experimental data from other light-source facilities around the country, including the Linac Coherent Light Source at SLAC, the Advanced Photon Source at



▼ SPOT Suite enhances cross-facility collaboration.

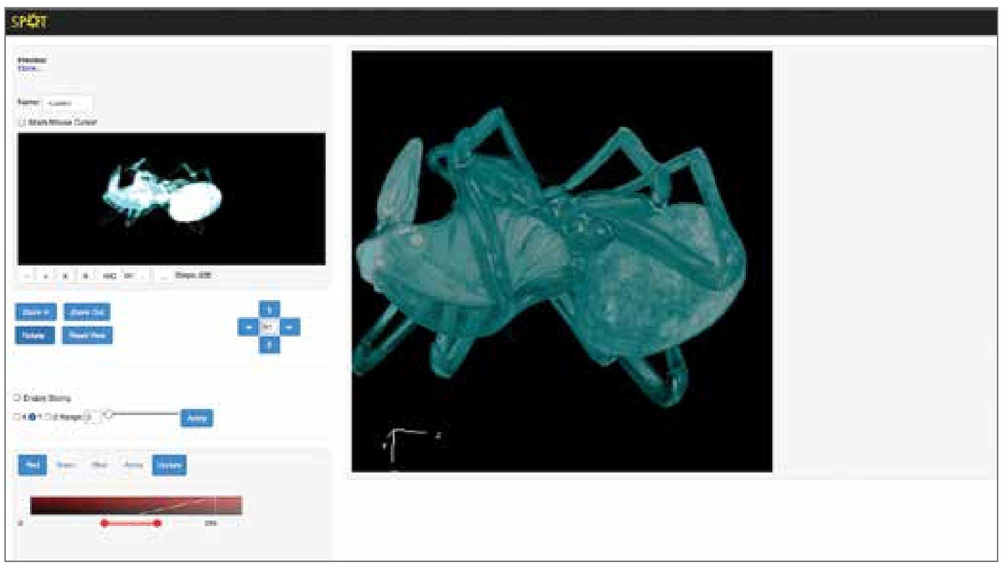
Argonne National Laboratory and the National Synchrotron Light Source at Brookhaven National Laboratory. This cross-facility functionality was demonstrated at the SC14 conference.

Improvements to SPOT targeting multi-facility use include on-demand computing functionality for the XMAS, HIPGISAX and MBIR analysis codes and improved visualization over the web powered by VisIt, a scalable visualization toolset available at NERSC. This work, also done in collaboration with researchers in the Computational Research Division, combines the powerful backend capabilities of VisIt running on NERSC's Edison system with modern web technologies such as WebGL to allow users to seamlessly (and without the need for training) visualize datasets of sizes on the order of 50 GB.

VisIt on the Web

Visualization toolkits such as VTK Paraview and VisIt support hundreds of scientific data formats and provide a large suite of analysis and visualization algorithms. These algorithms are optimized for scalability and performance and designed to work on everything from local desktops to large supercomputing systems. The comprehensive support for a suite of formats, analysis algorithms and visualizations often leads to a complex and unwieldy product that has a high learning curve—one of the most common complaints from users. Additionally, these tools often have complex dependency requirements and require considerable amount of care to be successfully deployed at supercomputing facilities.

To enable effective utilization of these visualization tools, we enhanced the tools, in particular VisIt, to support web-based environments. Almost all modern web browsers now support the ability to render intricate visual scenes through WebGL and connecting to remote services through WebSockets. The combination of the two technologies allows a web browser to provide a level of capability that rivals traditional applications, yet with the ease of web deployment. This approach provides scientists with a web tool optimized for their domain science, access to their data and the scalable infrastructure of a supercomputing facility natively within any modern web browser.



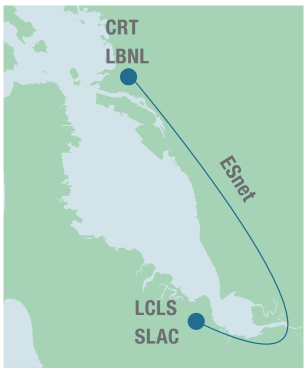
▲ VisIt, a scalable visualization toolset available at NERSC.

The SPOT portal at NERSC now has a custom VisIt interface that supports volume rendering, slicing and updating color maps. Communication and syncing web views among collaborating clients is also supported. This new capability on SPOT enables beamline scientists to visualize, analyze and interact with their data using a modern web browser while using NERSC resources in the backend. Custom visualization interfaces expose functionality and algorithms of interest to the target audience. Built-in collaboration services allow geographically separate scientists to interact with and analyze data simultaneously.

A prototype for this interface was demonstrated at the SC14 conference and at a live cross-country demo at DOE headquarters with scientists working in real time between Washington, D.C., and Berkeley Lab. We are working toward automating the visualization service entirely at NERSC.

Toward a Photon Science Speedway

▼ The photon science speedway concept links DOE's most powerful X-ray laser with DOE's most scientifically productive supercomputers and the world's fastest science network.



A series of experiments conducted by Berkeley Lab and SLAC National Accelerator Laboratory (SLAC) researchers in 2014 demonstrated how light sources and supercomputing facilities can be linked via a "photon science speedway" to better address emerging challenges in massive data analysis.

The researchers led a protein crystallography experiment at SLAC's Linac Coherent Light Source (LCLS) to look at the different photoexcited states of photosystem II, an assembly of large protein molecules that play a crucial role in photosynthesis. Subsequent analysis of the data on supercomputers at NERSC helped explain how nature splits a water molecule during photosynthesis, a finding that could advance the development of artificial photosynthesis for clean, green and renewable energy.

The experiment was also notable as an early demonstration of a "photon science speedway" connecting the LCLS and NERSC. Historically, light source users had to travel to facilities like the LCLS and Berkeley Lab's Advanced Light Source to run experiments. They would then download the raw data to an external hard drive and take this information home for processing and analysis on their personal computers. In this case, however, the researchers saw their data arrive at NERSC in real time using the DOE's Energy Sciences Network (ESnet).

A total of 114 TB of data was collected over a five-day period; during each day's 12-hour experiment run, the data was transferred over ESnet at 7.5 Gbps, which enabled the data processing team at NERSC to determine each night how good the data was from the previous day.

In a follow-on experiment that ran at LCLS in July 2014, the researchers added another twist: SPOT Suite, a set of software tools developed at Berkeley Lab that give scientists access to automated data management, data analysis and simulation tools when working with instruments such as the ALS and LCLS.

All of this bodes well for the creation of a photon science speedway that connects beamlines at facilities like LCLS and ALS to supercomputing centers like NERSC—a scenario that will become increasingly critical as the datasets generated by experiments at these facilities grow ever larger.

BrainFormat

Management and storage of neuroscience data is challenging due to the complex and multi-modal nature of the data. A fundamental challenge toward analysis, sharing and management of such complex data is the lack of common data standards. To address this issue, researchers at Berkeley Lab developed BrainFormat, a library that specifies a general data format standardization framework based on HDF5.

With this framework, the developers have defined and implemented a novel file format for management and storage of neuroscience data, while initially focusing on electrocorticography. Important advantages and features of the format and library include:

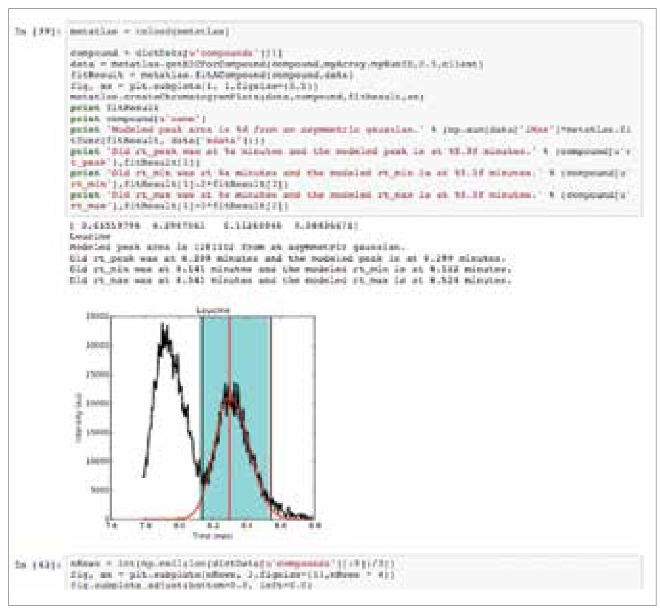
- **Easy to use:** User-friendly design and object-oriented python file API
- **Formal specification:** All components of the format have a formal specification that is part of the library as well as the files (JSON)
- **Verifiable:** The library supports verification of format compliance of complete files and components of files
- **Modular:** Managed objects allow semantic components of the format to be specified as self-contained units
- **Extensible:** New components can be easily added to the format, and different components can be designed independently
- **Reusable:** Existing components can be nested as well as extended through inheritance, and the library provides a number of base building blocks

- **Data annotation:** Reusable modules for annotating data subsets are available that support searching, filtering and merging of annotations and organization of annotations into collections
- **Supports self-contained as well as modular file storage:** All data can be stored in a single HDF5 file, or individual managed object containers can be stored in separate files that can be accessed via external links directly from the main HDF5 file
- **Application-independent design concepts and application-oriented modules:** The library provides a number of core modules that can be used to define arbitrary application file formats based on the concept of managed objects. Based on the concept of managed objects, the library then defines the application-oriented BRAIN file format
- **Portable, scalable and self-describing:** Built on top of HDF5 using best practices
- **Detailed developer and user documentation**
- **Open Source:** Available at https://bitbucket.org/oruebel/brainformat

As part of outreach to the community, the developers of BrainFormat have engaged with the Neurodata without Borders (NWB) initiative, which is supported by the Kavli Foundation, GE, Janelia Farm, the Allen Institute for Brain Science and the International Neuroinformatics Coordinating Facility. NWB hosted a hackathon in November 2014 to consolidate ideas for designing and implementing a standard neuroscience file format, and BrainFormat is among the candidates selected for further investigation. It is now a strong contender to contribute to and develop a community-wide data format and storage standard for the neuroscience research community.

Metabolite Atlas

This past year a group of biologists at Berkeley Lab that operates a number of mass spectrometry machines to explore the various byproducts of microbial metabolism approached NERSC for help in dealing with data analysis challenges. These modern, high-precision instruments are capable of generating up to 1 TB of raw data per day. But the process and tools weren't able to keep pace with and adequately manage this output. This included relying on USB drives and manually labeled hard disks to move data from the instruments to analysis systems. Significant amounts of time and effort were spent on tracking the progress and waiting for analysis to complete.



▲ Screenshot of an iPython Notebook retrieving data and visualizing it.

One of the scientists from this group approached NERSC's Data Analytic Services Group seeking a solution to the problem. NERSC staff engaged with the scientist to analyze the data pipelines, identify bottlenecks and suggest solutions to various aspects of the problem. With the help of NERSC staff, the group built a solution that utilizes the NERSC science gateway service, MongoDB service, SciDB service, global storage, HPSS and Edison. Currently the amount of data that has been stored and processed is approaching 10 TB, and the entire group of approximately a dozen scientists is starting to use the system for their day-to-day analytics. As a result of the new methods and approaches developed jointly with NERSC, a 10-100 times speedup has been observed for most of the operations.

The most innovative part of the project is in how users interact with the system. Traditionally a custom, web-based GUI is developed for each project. However, GUI development and maintenance require a large amount of effort. In this project, RESTful API was developed that exposed a limited number of operations to the end users, including authentication, select and process data. The end user interfaces to this web service via an iPython script. Many, if not most, scientists are familiar with using iPython Notebook for analysis, and they can easily write simple Python scripts to interact with this new service to extract and visualize data sets from these advanced instruments. As a result, this method has been quickly adopted by the scientists. In addition, using iPython Notebook adds the flexibility for sharing work and tracking provenance of analysis and results.

Science and Research

NERSC's primary mission is to accelerate scientific discovery at the DOE Office of Science through high performance computing, data analysis and comprehensive user support. In 2014, our users reported 1,808 refereed papers and 22 journal cover stories based on work performed at NERSC.

This section presents a selection of research highlights from 2014, illustrating the breadth of science supported by the center for all six DOE program offices.

Simulations Reveal Unusual Death for Ancient Stars

Findings Made Possible with NERSC Resources and Berkeley Lab Code

"To model the death mechanisms of these stars, Chen and his colleagues used CASTRO—a multidimensional compressible astrophysics code developed at Berkeley Lab."

Certain primordial stars—those between 55,000 and 56,000 times the mass of our Sun, or *solar masses*—may have died unusually. In death, these objects—among the universe's first-generation of stars—would have exploded as supernovae and burned completely, leaving no remnant black hole behind.

Astrophysicists at the University of California, Santa Cruz (UCSC) and the University of Minnesota came to this conclusion after running a number of supercomputer simulations at NERSC and the Minnesota Supercomputing Institute. They relied extensively on CASTRO, a compressible astrophysics code developed at Berkeley Lab. Their findings were published in *Astrophysical Journal*.

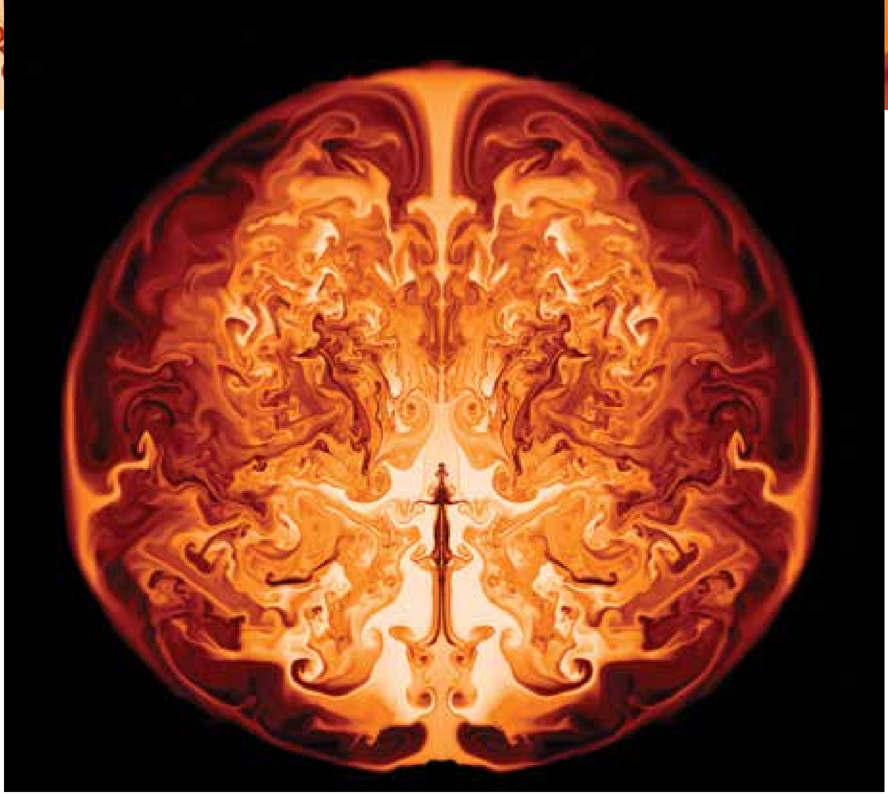
First-generation stars are interesting because they produced the first heavy elements, or chemical elements other than hydrogen and helium. In death, they sent their chemical creations into outer space, paving the way for subsequent generations of stars, solar systems and galaxies. With a greater understanding of how these first stars died, scientists hope to glean some insights about how the universe, as we know it today, came to be.

"We found that there is a narrow window where supermassive stars could explode completely instead of becoming a supermassive black hole, and no one has ever found this mechanism before," said lead author Ke-Jung Chen, a postdoctoral researcher at UCSC. "Without NERSC resources, it would have taken us a lot longer to reach this result."

The Simulations: What's Going On?

To model the life of a primordial supermassive star, Chen and his colleagues used a one-dimensional stellar evolution code called KEPLER. This code takes into account key processes like nuclear

PROJECT TITLE	NERSC PI	LEAD INSTITUTION	NERSC RESOURCES USED	DOE PROGRAM OFFICE
Type Ia Supernovae and X-ray Bursts	Stan Woosley	University of California, Santa Cruz	Edison, Hopper	HEP—Cosmic Frontier; ASCR



◀ This image is a slice through the interior of a supermassive star of 55,500 solar masses along the axis of symmetry. It shows the inner helium core in which nuclear burning is converting helium to oxygen, powering various fluid instabilities (swirling lines). This snapshot from a CASTRO simulation shows one moment a day after the onset of the explosion, when the radius of the outer circle would be slightly larger than that of the orbit of the Earth around the Sun. *Image: Ken Chen, University of California, Santa Cruz*

burning, stellar convection and, relevant for massive stars, photo-disintegration of elements, electron-positron pair production and special relativistic effects. The team also included general relativistic effects, which are important for stars above 1,000 solar masses.

They found that primordial stars between 55,000 to 56,000 solar masses live about 1.69 million years before becoming unstable due to general relativistic effects and then start to collapse. As the star collapses, it begins to rapidly synthesize heavy elements like oxygen, neon, magnesium and silicon starting with helium in its core. This process releases more energy than the binding energy of the star, halting the collapse and causing a massive explosion: a supernova.

To model the death mechanisms of these stars, Chen and his colleagues used CASTRO—a multidimensional compressible astrophysics code developed by Berkeley Lab scientists Ann Almgren and John Bell. These simulations show that once collapse is reversed, Rayleigh-Taylor instabilities mix heavy elements produced in the star's final moments throughout the star itself. This mixing should create a distinct observational signature that could be detected by upcoming near-infrared experiments such as the European Space Agency's Euclid and NASA's Wide-Field Infrared Survey Telescope, according to the researchers.

Depending on the intensity of the supernovae, some supermassive stars could, when they explode, enrich their entire host galaxy and even some nearby galaxies with elements ranging from carbon to silicon. In some cases, a supernova may even trigger a burst of star formation in its host galaxy, which would make it visually distinct from other young galaxies.

"My work involves studying the supernovae of very massive stars with new physical processes beyond hydrodynamics, so I've collaborated with Ann Almgren to adapt CASTRO for many different projects over the years," says Chen. "Before I run my simulations, I typically think about the physics I need to solve a particular problem. I then work with Ann to develop some code and incorporate it into CASTRO. It is a very efficient system."

PUBLICATION

K. Chen, A. Heger, S. Woosley, A. Almgren, D. J. Whalen, J. L. Johnson, "The General Relativistic Instability Supernova of a Supermassive Population III Star, *The Astrophysical Journal*, 790(2), 2014, doi: 10.1088/0004-637X/790/2/162

FULL STORY

http://www.nersc.gov/news-publications/news/science-news/2014/simulations-reveal-unusual-death-for-ancient-stars/

Pinpointing the Magnetic Moments of Nuclear Matter

Lattice QCD Calculations Reveal Inner Workings of Lightest Nuclei

"Savage and his colleagues believe a mathematical technique known as lattice QCD is the key to unlocking the remaining mysteries of nuclear forces."

A team of nuclear physicists has made a key discovery in its quest to shed light on the structure and behavior of subatomic particles. Using supercomputing resources at NERSC, the Nuclear Physics with Lattice QCD (NPLQCD) collaboration demonstrated for the first time the ability of quantum chromodynamics (QCD)—a fundamental theory in particle physics—to calculate the magnetic structure of some of the lightest nuclei.

Their findings, published December 16, 2014, in *Physical Review Letters*, are part of an ongoing effort to address fundamental questions in nuclear physics and high-energy physics and further our understanding of the universe, according to the NPLQCD. The collaboration was established in 2004 to study the properties, structures and interactions of fundamental particles, including protons, neutrons, tritons, quarks and gluons—the subatomic particles that make up all matter.

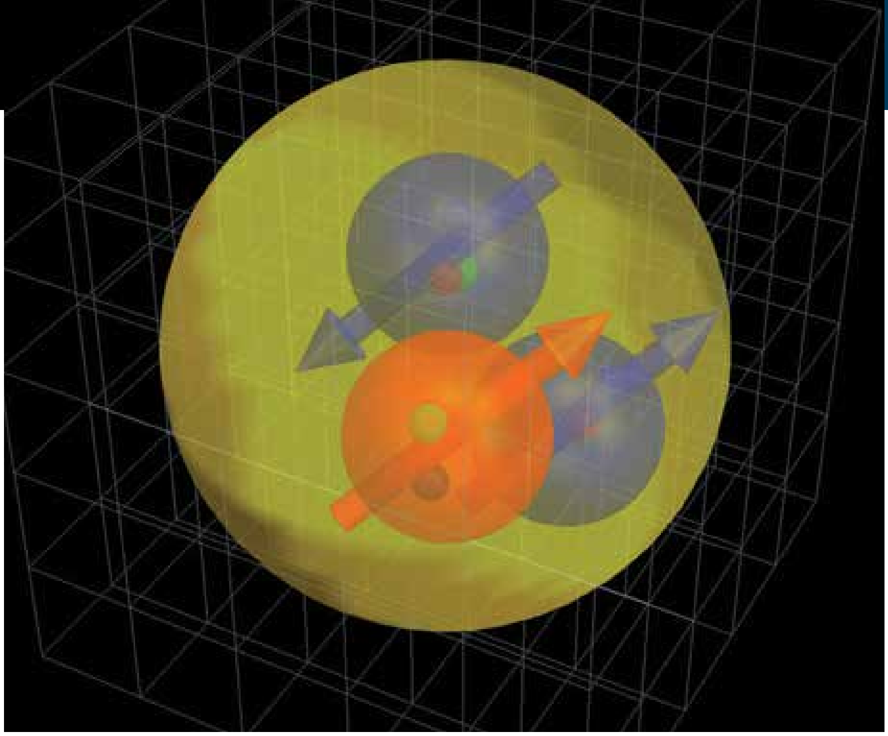
"We are interested in how neutrons and protons interact with each other, how neutrons and strange particles interact with each other and more generally what the nuclear forces are," explained Martin Savage, senior fellow in the Institute for Nuclear Theory and professor at the University of Washington and a founding member of the NPLQCD. "If you want to make predictions for complex nuclear systems, you have to know what the nuclear forces are."

While some of these interactions are well measured experimentally—such as the strength with which a neutron and a proton interact with each other at low energies—others are not. Savage and his colleagues believe a mathematical technique known as lattice QCD is the key to unlocking the remaining mysteries of nuclear forces.

QCD emerged in the 1970s as a theoretical component of the Standard Model of particle physics—our current mathematical description of reality. The Standard Model, sometimes described as "the theory of almost everything," is used to explain electromagnetic, weak and strong nuclear interactions and to classify all known subatomic particles. The only force not described is gravity.

Lattice QCD is an approach to solving the QCD theory of quarks and gluons. Researchers use it to break up space and time into a grid (lattice) of points and then develop and solve equations connecting the degrees of freedom of those points, explained William Detmold, assistant professor of Physics at the Massachusetts Institute of Technology and a co-author on the *Physical Review Letters* paper.

PROJECT TITLE	NERSC PI	LEAD INSTITUTION	NERSC RESOURCES USED	DOE PROGRAM OFFICE
Hadron-Hadron Interactions with Lattice QCD	Martin Savage	University of Washington	Edison	NP—Lattice Gauge Theory



◂ Artist's impression of a triton, the atomic nucleus of a tritium atom. The image shows blue neutrons and a red proton with quarks inside; the arrows indicate the alignments of the spins. *Image: William Detmold, Massachusetts Institute of Technology*

"Lattice QCD is the implementation of QCD on the computer. It is the tool we use to solve QCD," Detmold said.

Over the last two decades, increasing compute power and refinements in lattice QCD algorithms and codes have enabled physicists to predict the mass, binding energy and other properties of particles before they're measured experimentally and calculate quantities already measured.

"There are a number of really critical areas in nuclear physics and particle physics that require lattice QCD to succeed," Savage said. "If you want to make predictions, for instance, of what happens on the inside of a collapsing star, what dictates whether a supernova collapses through to a black hole or ends up as a neutron star is determined by how soft or how hard nuclear matter is. Lattice QCD is a technique that will greatly improve our understanding of the nuclear interactions relevant to these matter densities."

In recent years the NPLQCD and other QCD research groups have made steady progress in their efforts to tease apart the behavior of subatomic particles. For the *Physical Review Letters* study, the NPLQCD team investigated the magnetic moments of neutrons, protons and the lightest nuclei: deuterons, tritons and helium-3 (3He). They used the Chroma lattice QCD code to run a series of Monte Carlo calculations on NERSC's newest supercomputer, Edison, and demonstrated for the first time QCD's ability to calculate the magnetic structure of nuclei.

"This is the first study where we ask the question about what the structure of the nucleus is, and the first study probing what nuclei are at the level of quarks," Detmold explained. "If you can understand the magnetic moments and the more complicated magnetic structure of nuclei, you can come up with a topographical map of what, for example, the current distribution inside a nucleus actually is. The magnetic moment is a first step toward that."

The NPLQCD team also discovered something quite unexpected, according to Savage: that the theoretical magnetic moments of these nuclei at unphysically large quark masses are remarkably close to those found in physical experiments.

"These are the very first calculations of nuclear magnetic moments and a first glimpse of nuclear structure to emerge from lattice QCD calculations," Savage said.

It was only possible to achieve these findings with the advent of the latest supercomputers, Detmold added.

"NERSC resources are critical to our calculations," Savage said. "The NERSC machines are able to run the largest jobs we presently need to undertake and are very effective computational platforms."

PUBLICATION

S.R. Beane, E. Chang, S. Cohen, W. Detmold, H.W. Lin, K. Orginos, A. Parreño, M.J. Savage, B.C. Tiburzi (NPLQCD Collaboration), "Magnetic Moments of Light Nuclei from Lattice Quantum Chromodynamics," *Physical Review Letters,* 113(25), December 16, 2014, doi: http://dx.doi.org/10.1103/PhysRevLett.113.252001

FULL STORY

http://www.nersc.gov/news-publications/nersc-news/science-news/2015/pinpointing-the-magnetic-moments-of-nuclear-matter/

Berkeley Lab Particle Accelerator Sets World Record

Simulations at NERSC Help Validate Experimental Laser-Plasma Design

> **"The computational power of Edison allowed us to perform extensive parameter scans, facilitating our understanding of the physics."**

Using one of the most powerful lasers in the world, Berkeley Lab researchers have accelerated subatomic particles to the highest energies ever recorded from a compact accelerator.

The team used a specialized petawatt laser and charged-particle plasma to get the particles up to speed. The setup is known as a laser-plasma accelerator, an emerging class of particle accelerators that physicists believe can shrink traditional, miles-long accelerators to machines that can fit on a table.

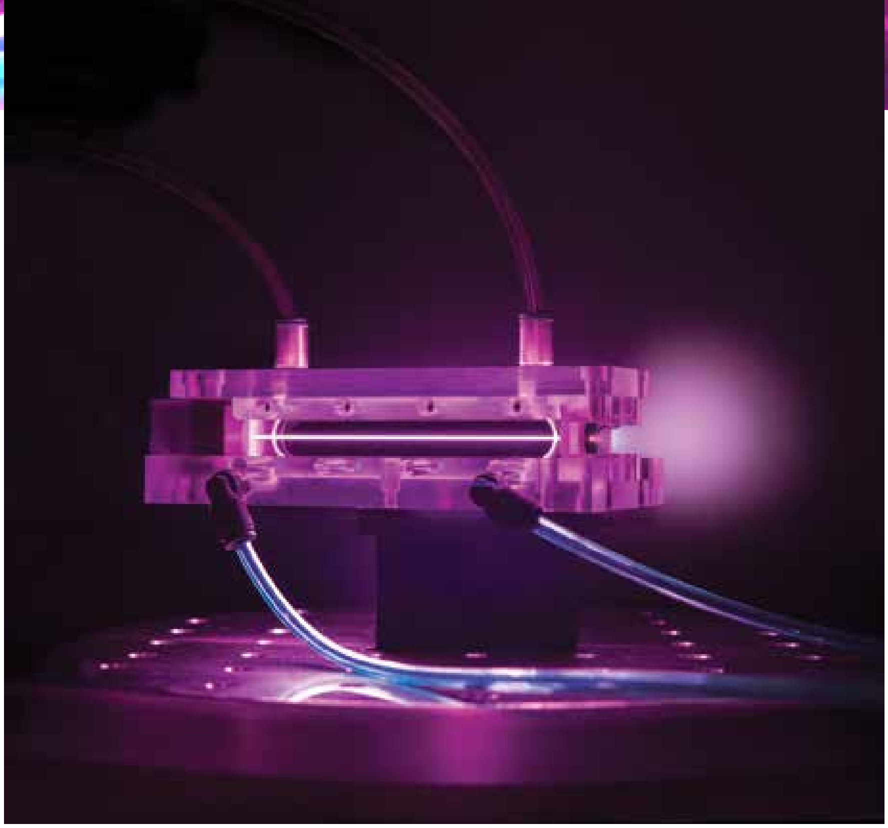
The researchers sped up the particles—electrons in this case—inside a nine-centimeter long tube of plasma. The speed corresponded to an energy of 4.25 giga-electron volts. The acceleration over such a short distance corresponds to an energy gradient 1,000 times greater than traditional particle accelerators and marks a world record energy for laser-plasma accelerators.

"This result requires exquisite control over the laser and the plasma," said Wim Leemans, director of the Accelerator Technology and Applied Physics Division at Berkeley Lab and lead author on the paper, published December 12, 2014, in *Physical Review Letters*.

Traditional particle accelerators—such as the Large Hadron Collider at CERN, which is 17 miles in circumference—speed up particles by modulating electric fields inside a metal cavity. It's a technique that has a limit of about 100 mega-electron volts per meter before the metal breaks down.

Laser-plasma accelerators take a completely different approach. In the case of this experiment, a pulse of laser light was injected into a short, thin, straw-like tube that contained plasma. The laser created a channel through the plasma as well as waves that trapped free electrons and accelerated them to high energies—similar to the way a surfer gains speed when skimming down the face of a wave.

PROJECT TITLE	NERSC PI	LEAD INSTITUTION	NERSC RESOURCES USED	DOE PROGRAM OFFICE
Simulation of Laser-Plasma Particle Accelerators	Cameron Geddes	Lawrence Berkeley National Laboratory	Edison, Hopper	HEP—Accelerator Simulations



◂ A 9 cm-long capillary discharge waveguide was used in BELLA experiments to generate multi-GeV electron beams. The plasma plume has been made more prominent with the use of HDR photography. *Image: Roy Kaltschmidt, Lawrence Berkeley National Laboratory*

The record-breaking energies were achieved with the help of BELLA (Berkeley Lab Laser Accelerator), one of the most powerful lasers in the world, producing a quadrillion watts of power (a petawatt). In addition to packing a high-powered punch, BELLA is renowned for its precision and control.

"We're forcing this laser beam into a 500-micron hole about 14 meters away," Leemans said. "The BELLA laser beam has sufficiently high pointing stability to allow us to use it." Moreover, the laser pulse, which fires once a second, is stable to within a fraction of a percent, he added.

At such high energies, the researchers also needed to see how various parameters would affect the outcome. So they used computer simulations at NERSC to test the setup before ever turning on a laser. The multi-dimensional particle-in-cell (PIC) simulations they ran on NERSC's Edison system were fundamentally important to modeling the propagation of an high-intensity laser in the plasma, characterizing the nonlinear wakefield excitation and studying the details of particle self-trapping, according to Leemans.

"PIC simulations of the laser-plasma interactions are usually very computationally demanding," he said. "The computational power of Edison allowed us to perform extensive parameter scans, facilitating our understanding of the physics."

In order to accelerate electrons to even higher energies—Leemans' near-term goal is 10 giga-electron volts—the researchers will need to more precisely control the density of the plasma channel through which the laser light flows. In essence, the researchers need to create a tunnel for the light pulse that's just the right shape to handle more-energetic electrons. According to Leemans, future work will demonstrate a new technique for plasma-channel shaping.

"Small changes in the setup give you big perturbations," said Eric Esarey, senior science advisor for the Accelerator Technology and Applied Physics Division at Berkeley Lab, who leads the theory effort. "We're homing in on the regions of operation and the best ways to control the accelerator."

PUBLICATION

W. P. Leemans, A. J. Gonsalves, H.-S. Mao, K. Nakamura, C. Benedetti, C. B. Schroeder, Cs. Tóth, J. Daniels, D. E. Mittelberger, S. S. Bulanov, J.-L. Vay, C. G. R. Geddes, E. Esarey, "Multi-GeV Electron Beams from Capillary-Discharge-Guided Subpetawatt Laser Pulses in the Self-Trapping Regime," *Physical Review Letters*, 113(24), December 12, 2014, doi: http://dx.doi.org/10.1103/PhysRevLett.113.245002

FULL STORY

http://www.nersc.gov/news-publications/nersc-news/science-news/2014/berkeley-lab-particle-accelerator-sets-world-record/

Chombo-Crunch Sinks Its Teeth into Fluid Dynamics

Decade of Development Yields Novel Code for Energy, Oil & Gas Applications

"Chombo-Crunch could enhance efforts to develop carbon sequestration as a way to address Earth's growing carbon dioxide challenges."

For more than a decade, mathematicians and computational scientists have been collaborating with earth scientists at Berkeley Lab to break new ground in the modeling of complex flows in energy and oil and gas applications.

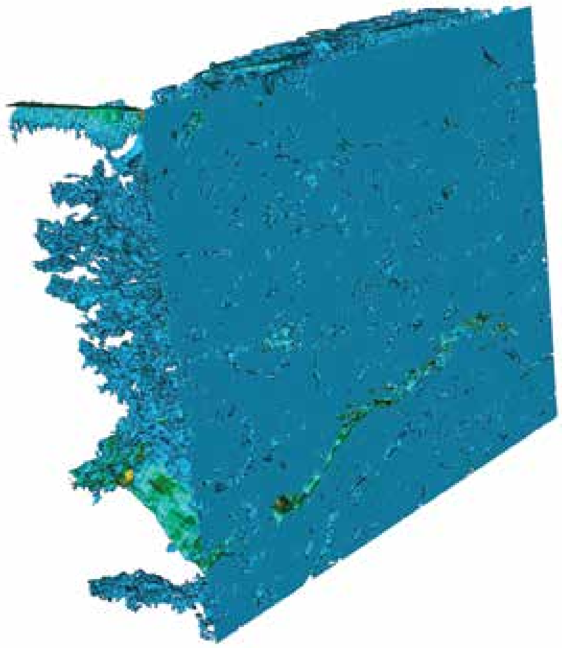
Their work has yielded a high-performance computational fluid dynamics (CFD) and reactive transport code dubbed "Chombo-Crunch" that could enhance efforts to develop carbon sequestration as a way to address the Earth's growing carbon dioxide challenges. It could also lead to new safety measures in the oil and gas industry and aerospace engineering.

Chombo-Crunch—the brainchild of David Trebotich, a computational scientist in the Computational Research Division (CRD), and Carl Steefel, a computational geoscientist in the Earth Sciences Division (ESD)—has its algorithmic roots in LDRD-funded microfluidics research Trebotich was involved in back in 2002 while he and Steefel were at Lawrence Livermore National Laboratory. At the time there was national interest in developing miniaturized microfluidic devices to detect chemical attacks via air samples, so Trebotich began modeling flow in a packed bed design—a small cylinder filled with spheres (essentially an engineered porous medium) that could function as a flow-through device. The challenge was to overcome the limitations of traditional finite difference methods (FDMs) used in simulating complex geometries.

Using CFD tools based on embedded boundary technology, Trebotich found that he could determine the optimal flow conditions for capturing a polymer molecule by fluid mechanical forces alone. This work, coupled with Steefel's vision for using the finite volume approach to perform direct numerical simulation of reactive transport at the pore scale, laid the foundation for Chombo-Crunch.

Chombo-Crunch addresses the limitations of FDMs in simulating complex geometries by combining two sets of code in an embedded boundary approach: Chombo, a suite of multiscale, multiphysics simulation tools for solving applied partial differential equations such as those that describe fluid dynamics and species transport in complex geometries; and CrunchFlow, a code that provides mixed equilibrium-kinetic treatment of geochemical reactions, developed by Steefel and interfaced to Chombo by Trebotich and Sergi Molins of Berkeley Lab's Earth Sciences Division. Together they are able to provide more detailed analysis and modeling of complex flows and subsurface reactive transport than previously possible using lab experiments or larger scale models.

PROJECT TITLE	NERSC PI	LEAD INSTITUTION	NERSC RESOURCES USED	DOE PROGRAM OFFICE
Chombo-Crunch: Advanced Simulation of Subsurface Flow and Reactive Transport Processes Associated with Carbon Sequestration	David Trebotich	Lawrence Berkeley National Laboratory	Hopper, Edison	ASCR—Applied Mathematics, BES—EFRC



◄ Simulation of resolved steady-state flow in fractured Marcellus shale based upon FIB-SEM imagery obtained by Lisa Chan at Tescan USA, processed by Terry Ligocki (Lawrence Berkeley National Laboratory), courtesy Tim Kneafsey (Lawrence Berkeley National Laboratory).

Fine-Tuning the Engine

For the last 12 years, Trebotich has been testing and tweaking the simulation engine that is under the hood of Chombo-Crunch—a Chombo-based flow and species transport solver—including running some impressive carbon sequestration simulations in 2013 at NERSC and the Oak Ridge Leadership Computing Center. More recently he's been looking at new geometry-based applications that can benefit from this CFD simulation prowess, primarily in energy and oil and gas.

In a paper published in *Communications in Applied Mathematics and Computational Science*, Trebotich described how the robustness of the flow solver in Chombo can enhance CFD modeling for a broad range of flow geometries, from creeping flow in realistic pore space to transitional turbulent flows past bluff bodies and turbulent pipe flow. For example, the Berkeley Lab team used the CFD algorithm to study turbulent flow past a sphere, a model problem that Trebotich believes could help aerospace engineers optimize takeoff and landing patterns through more accurate prediction of aircraft wakes.

The high-resolution CFD simulator also helped shed new light on the Deepwater Horizon Macondo oil well explosion and fire in April 2010, which resulted in the deaths of 11 workers and caused a massive oil spill into the Gulf of Mexico. After the failure of the blowout preventer allowed oil and natural gas to gush into the Gulf at the sea floor, the Department of Energy established the Flow Rate Technical Group, comprising scientists from multiple national labs, to estimate the oil flow rate based on the physical properties and behavior of the oil and gas in the reservoir, the wellbore and seafloor attachments.

Another series of simulations showed how this modeling approach could help geoscientists address some of the unknowns of hydrofracturing, or "fracking." Using the algorithm to solve the incompressible Navier-Stokes equations in complex geometries, Trebotich worked with Berkeley Lab geoscientist Tim Kneafsey to achieve the first-ever computer simulation of fully resolved flow in fractured shale from image data of rock samples. The ability to simulate flow in shales at high resolution could help geoscientists better understand the effects of fracking and develop safe and reliable methods for storing carbon dioxide underground.

"We just have a proof of concept here, but we can actually simulate flow in fractured shale and resolve flow features in microfractures," Trebotich said. "Coupled with adaptive mesh refinement, our embedded boundary approach is a powerful tool for doing high-performance, multiscale, multiphysics simulations."

PUBLICATION

D. Trebotich, D.T. Graves, "An Adaptive Finite Volume Method for the Incompressible Navier-Stokes Equations in Complex Geometries," *Communications in Applied in Mathematics and Computational Science*, 10(1), 2015, doi: 10.2140/camcos.2015.10.43

FULL STORY

http://www.nersc.gov/news-publications/nersc-news/science-news/2015/chombo-crunch-sinks-teeth-into-fluid-dynamics/

To Bridge LEDs' Green Gap, Scientists Think Small

Nanostructures Half as Wide as a DNA Strand Show Promise for Efficient LEDs

"Indium niatride at the few-nanometer size range offers a promising approach to engineering efficient, visible light emission at tailored wavelengths."

Nanostructures half the breadth of a DNA strand could improve the efficiency of light emitting diodes (LEDs), especially in the "green gap," a portion of the spectrum where LED efficiency plunges.

Using NERSC supercomputing resources, University of Michigan researchers Dylan Bayerl and Emmanouil Kioupakis found that the semiconductor indium nitride (InN), which typically emits infrared light, will emit green light if reduced to 1 nanometer-wide wires. Moreover, by varying their sizes, these nanostructures could be tailored to emit different colors of light, which could lead to more natural-looking white lighting while avoiding some of the efficiency loss today's LEDs experience at high power.

"Our work suggests that indium nitride at the few-nanometer size range offers a promising approach to engineering efficient, visible light emission at tailored wavelengths," said Kioupakis.

Their results were featured on the cover of the July 2014 issue of *Nano Letters*.

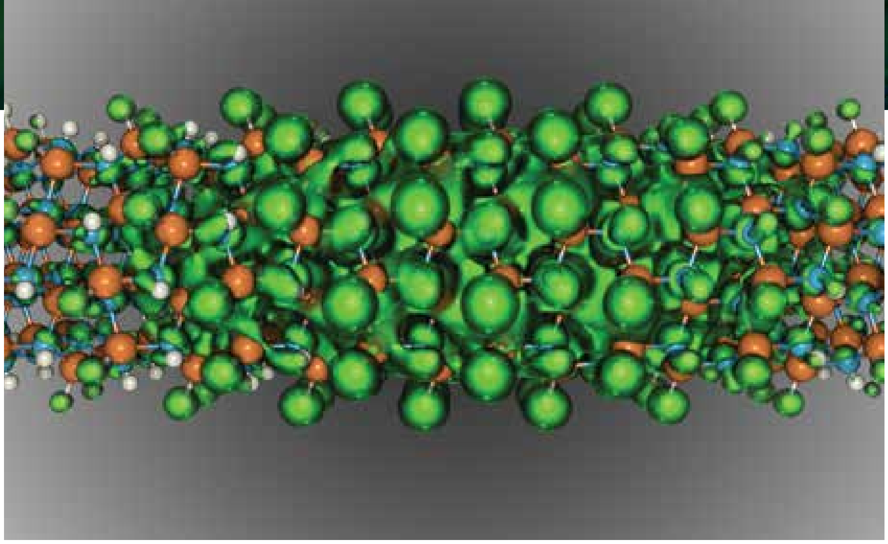
A Tantalizing Prospect

Today's LEDs are created as multilayered microchips. The outer layers are doped with elements that create an abundance of electrons on one layer and too few on the other. The missing electrons are called holes. When the chip is energized, the electrons and holes are pushed together, confined to the intermediate quantum-well layer where they are attracted to combine, shedding their excess energy (ideally) by emitting a photon of light.

At low power, nitride-based LEDs (most commonly used in white lighting) are very efficient, converting most of their energy into light. But turn the power up to levels that could light up a room and efficiency plummets, meaning a smaller fraction of electricity gets converted to light. This effect is especially pronounced in green LEDs, giving rise to the term "green gap."

Nanomaterials offer the tantalizing prospect of LEDs that can be "grown" in arrays of nanowires, dots or crystals. The resulting LEDs could not only be thin, flexible and high-resolution, but very efficient as well.

PROJECT TITLE	NERSC PI	LEAD INSTITUTION	NERSC RESOURCES USED	DOE PROGRAM OFFICE
Electronic and Optical Properties of Novel Photovoltaic and Thermoelectric Materials from First-principles	Emmanouil Kioupakis	University of Michigan	Edison	Basic Energy Sciences



◀ This simulation of a 1 nanometer-wide indium nitride wire shows the distribution of an electron around a positively charged "hole." Strong quantum confinement in these small nanostructures enables efficient light emission at visible wavelengths, researchers found. *Image: Burlen Loring, Lawrence Berkeley National Laboratory*

"If you reduce the dimensions of a material to be about as wide as the atoms that make it up, you get quantum confinement. The electrons are squeezed into a small region of space, increasing the bandgap energy," Kioupakis said. That means the photons emitted when electrons and holes combine are more energetic, producing shorter wavelengths of light.

The energy difference between an LED's electrons and holes, called the bandgap, determines the wavelength of the emitted light. The wider the bandgap, the shorter the wavelength of light. The bandgap for bulk InN is quite narrow, only 0.6 electron volts (eV), so it produces infrared light. In Bayerl and Kioupakis' simulated InN nanostructures, the calculated bandgap increased, leading to the prediction that green light would be produced with an energy of 2.3eV.

"If we can get green light by squeezing the electrons in this wire down to a nanometer, then we can get other colors by tailoring the width of the wire," said Kioupakis. A wider wire should yield yellow, orange or red. A narrower wire, indigo or violet.

That bodes well for creating more natural-looking light from LEDs. By mixing red, green and blue LEDs, engineers can fine tune white light to warmer, more pleasing hues. This "direct" method isn't practical today because green LEDs are not as efficient as their blue and red counterparts. Instead, most white lighting today comes from blue LED light passed through a phosphor, a solution similar to fluorescent lighting and not a lot more efficient. Direct LED lights would not only be more efficient, but the color of light they produce could be dynamically tuned to suit the time of day or the task at hand.

Upping Efficiencies

Using pure InN rather than layers of alloy nitride materials would eliminate one factor that contributes to the inefficiency of green LEDs: nanoscale composition fluctuations in the alloys. These have been shown to significantly affect LED efficiency. Also, using nanowires to make LEDs eliminates the "lattice mismatch" problem of layered devices.

"When the two materials don't have the same spacing between their atoms and you grow one over the other, it strains the structure, which moves the holes and electrons further apart, making them less likely to recombine and emit light," said Kioupakis, who discovered this effect in previous research that also drew on NERSC resources. "In a nanowire made of a single material, you don't have this mismatch so you can get better efficiency."

NERSC's Edison supercomputer has been instrumental in this research, said Bayerl. The system's thousands of compute cores and high memory-per-node allowed Bayerl to perform massively parallel calculations with many terabytes of data stored in RAM, which made the InN nanowire simulation feasible.

PUBLICATION

D. Bayerl, E. Kioupakis, "Visible-Wavelength Polarized-Light Emission with Small-Diameter InN Nanowires," *Nano Letters* 14(7), 3709-3714 (2014), doi: 10.1021/nl404414r

FULL STORY

http://www.nersc.gov/news-publications/news/science-news/2014/to-bridge-leds-green-gap-scientists-think-small/

Hot Plasma Partial to Bootstrap Current

Calculations Show How Plasma Phenomenon Could Reduce Fusion Costs

"If we ever want to make a tokamak fusion plant down the road, for economic reasons the plasma will have to supply a lot of its own current."

Supercomputers at NERSC are helping plasma physicists create a potentially more affordable and sustainable fusion reaction that utilizes some of the reaction's own energy.

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

But there's a lot going on inside the plasma as it heats up, not all of it good. Driven by electric and magnetic forces, charged particles swirl around and collide into one another, and the central temperature and density are constantly evolving. In addition, plasma instabilities disrupt the reactor's ability to produce sustainable energy by increasing the rate of heat loss.

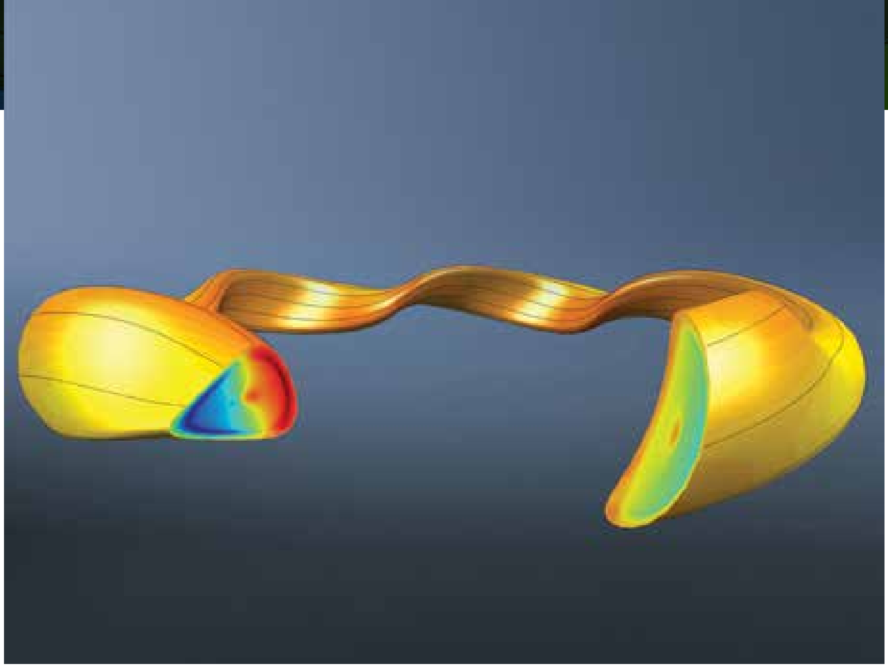
Fortunately, research has shown that more beneficial forces are also at play within the plasma. For example, if the pressure of the plasma varies across the radius of the vessel, a self-generated current will spontaneously arise within the plasma—a phenomenon known as the "bootstrap" current.

NERSC supercomputers are enabling further study of the bootstrap current, which could help reduce or eliminate the need for an external current driver and pave the way to a more cost-effective fusion reactor. Matt Landreman, research associate at the University of Maryland's Institute for Research in Electronics and Applied Physics, collaborated with two research groups to develop and run new codes at NERSC that more accurately calculate this self-generated current.

"It turns out that plasma in a curved magnetic field will generate some average electric current on its own," Landreman said. "If we ever want to make a tokamak fusion plant down the road, for economic reasons the plasma will have to supply a lot of its own current."

The Plasma Physics and Controlled Fusion paper focused on plasma behavior in tokamak reactors using PERFECT, a code Landreman wrote. First introduced in the 1950s, tokamaks today are considered by many to be the best candidate for producing controlled thermonuclear fusion power. A tokamak features a torus (doughnut-shaped) vessel and a combination of external magnets and a current driven in the plasma to create a stable confinement system.

PROJECT TITLE	NERSC PI	LEAD INSTITUTION	NERSC RESOURCES USED	DOE PROGRAM OFFICE
Turbulence, Transport and Magnetic Reconnection in High Temperature Plasma	William Dorland	University of Maryland	Edison, Hopper	FES—Fusion Base Program
Computational Studies in Plasma Physics and Fusion Energy	Abhay Ram	Massachusetts Institute of Technology	Hopper	FES—Fusion Base Program



◀ A calculation of the self-generated plasma current in the W7-X reactor, performed using the SFINCS code on Edison. The colors represent the amount of electric current along the magnetic field; the black lines show magnetic field lines. *Image: Matt Landreman, University of Maryland*

PERFECT was designed to examine the plasma edge, a region of the tokamak where "lots of interesting things happen," Landreman said. For example, in most of the inner part of the tokamak there is a fairly gradual gradient of the density and temperature.

"But at the edge there is a fairly big jump in density and temperature—what people call the edge pedestal. What is different about PERFECT is that we are trying to account for some of this very strong radial variation," he explained.

This study's findings are important because researchers are concerned that the bootstrap current may affect edge stability. PERFECT is also used to calculate plasma flow, which also may affect edge stability.

"My co-authors had previously done some analytic calculations to predict how the plasma flow and heat flux would change in the pedestal region compared to places where radial gradients aren't as strong," Landreman said. "The analytic calculations provide insight into how the plasma flow and heat flux will be affected by these strong radial gradients."

Tokamak vs. Stellarator

In the *Physics of Plasmas* study, the researchers used a second code, SFINCS, to focus on related calculations in a different kind of confinement concept: a stellarator.

First introduced in the 1950s, stellarators have played a central role in the German and Japanese fusion programs and were popular in the U.S. until the 1970s when many fusion scientists began favoring the tokamak design. In recent years several new stellarators have appeared, including the Wendelstein 7-X (W7-X) in Germany.

"In the W7-X design, the amount of plasma current has a strong effect on where the heat is exhausted to the wall," Landreman explained. "So at Max Planck (where W7-X is being built) they are very concerned about exactly how much self-generated current there will be when they turn on their machine. If the plasma makes more current than expected, the heat will come out in a different location, and you don't want to be surprised."

These concerns stemmed from the fact that the previous code was developed when computers were too slow to solve the "real" 4D equation, he added.

"The previous code made an approximation that you could basically ignore all the dynamics in one of the dimensions (particle speed), thereby reducing 4D to 3D," Landreman said. "Now that computers are faster, we can test how good this approximation was. And what we found was that basically the old code was pretty darn accurate and that the predictions made for this bootstrap current are about right."

PUBLICATIONS

M. Landreman, F. I. Parra, P. J. Catto, D. R. Ernst, I. Pusztai, "Radially global δf computation of neoclassical phenomena in a tokamak pedestal," *Plasma Physics and Controlled Fusion* 56 (2014), 045005, doi: 10.1088/0741-3335/56/4/045005

M. Landreman, H.M. Smith, A. Mollen, P. Helander, "Comparison of particle trajectories and collision operators for collisional transport in nonaxisymmetric plasmas," *Physics of Plasmas* 21(4), 042503 (2014), doi: http://dx.doi.org/10.1063/1.4870077

FULL STORY

http://www.nersc.gov/news-publications/news/science-news/2014/hot-plasma-partial-to-bootstrap-current/

Human-Induced Climate Change Cuts Flooding in African Delta

Simulations Run at NERSC Help Shed Light on Evolving Weather Patterns in Okavango River Basin

"This study, a first of its kind carried out on the African continent, produced a unique set of simulations generated by two computer models of the region's climate system."

Researchers at the University of Cape Town, Berkeley Lab and the United Nations Development Programme analyzed how human-induced climate change affects flooding in the Okavango River, an ecologically and geographically unique river basin in southern Africa.

After seasonal rains fall in southern Angola, floodwaters flow slowly down the Okavango River into semi-arid northwestern Botswana, where the river spreads, floods and eventually evaporates within the inland Okavango Delta. The annual floods of 2009, 2010 and 2011 reached extents last seen decades ago. While deaths were minimal, villages and houses were flooded, bridges closed or washed away and water and electricity supplies interrupted.

In a world where greenhouse gas emissions from human activities are causing the climate to change, were these unusually high floods related to human-induced climate change?

A study published in the *Journal of Hydrology*—a first of its kind carried out on the African continent—addressed this question by producing a unique set of simulations generated by two computer models of the region's climate system. In one set of simulations, the models were driven with "real-world" conditions observed over the 2009-2011 period, including concentrations of greenhouse gases. In other simulations, the models were driven with "non-greenhouse-gas" conditions. These simulations were then used to run a model of the Okavango River system to predict how floods would change under different climate conditions.

"By comparing the number of simulations under the 'real world' conditions that produced high floods with the number of simulations that did so in the 'non-greenhouse-gas' world, we were able to estimate the degree to which greenhouse gas emissions have altered the chance of a high flood occurring," said lead author Piotr Wolski, who started the research while at the Okavango Research Institute and is now at the University of Cape Town.

After running a number of simulations, Wolski and his colleagues ultimately found that greenhouse gas emissions have substantially reduced the chance of the floods in the Okavango Delta. To find out

PROJECT TITLE	NERSC PI	LEAD INSTITUTION	NERSC RESOURCES USED	DOE PROGRAM OFFICE
Attribution of Extreme Weather Risk to Anthropogenic Emissions	Daithi Stone	Lawrence Berkeley National Laboratory	Hopper	BER—Biological and Environmental Research

◀ This image is a compilation of three images from Envisat's radar and shows where southwestern Africa's Okavango River empties into the inland Okavango Delta in northern Botswana. The Okavango River originates in Angola, forms part of the Angola Namibia border and then ends in northern Botswana. *Image: European Space Agency*

why, the researchers looked at the different climates to dissect and understand the mechanism responsible for the reduced chance of flooding.

"It was fairly simple," explained Wolski. "The air is warmer in the climate we are experiencing and the river takes a long time to flow down to the delta, so you get more evaporation occurring before the river even reaches the delta and thus fewer high floods. Rainfall was not very different between the two sets of simulations, but in any case it would have to have been quite a big change in rainfall in order to match the difference in evaporation."

The findings appear counterintuitive, given the high floods of 2009-11. "It is indeed an unusual situation," said Wolski. "On the one hand, we have unusually high flooding and the general understanding that anthropogenic climate change leads to an increase in frequency and magnitude of extreme weather events (such as floods). On the other hand, we have our results, which indicate that greenhouse gas emissions actually led to a decrease in flooding. What makes the two consistent with each other is the natural variability occurring in the system."

This variability manifests itself as 20-30 year periods of above-average or below-average flooding, and the 2009-11 years fall within the former. "If not for climate change, the Okavango system would have experienced even larger flooding in 2009-11 than it actually did," Wolski said.

The University of Cape Town team initially generated a set of simulations to assess the role of greenhouse gases in the chance of extreme events. The Berkeley Lab researchers then mirrored their efforts using the CAM5 atmospheric model, which is part of the Community Earth Systems Model, developed by the National Center for Atmospheric Research for the wider climate research community. The Cape Town team built up the simulations of their climate model gradually over a period of years using an in-house computing cluster, then matched them over a shorter time by running their climate model on NERSC's Hopper supercomputer.

PUBLICATION

P. Wolski, D. Stone, M. Tadross, M. Wehner, B. Hewitson, "Attribution of floods in the Okavango basin, Southern Africa," *Journal of Hydrology*, 511, 350-358, April 16, 2014, doi: 10.1016/j.jhydrol.2014.01.055

FULL STORY

http://www.nersc.gov/news-publications/news/science-news/2014/human-induced-climate-change-reduces-chance-of-flooding-in-okavango-delta/

Stellar Behemoth Self-Destructs in Type IIb Supernova

Astronomers Catch Rare Star Just Hours After it Explodes

"For the first time ever, in 2014 scientists directly confirmed that a Wolf-Rayet star—sitting 360 million light years away in the Bootes constellation—died in a violent explosion known as a Type IIb supernova."

Our Sun may seem pretty impressive: 330,000 times as massive as Earth, it accounts for 99.86 percent of the solar system's total mass, generates about 400 trillion trillion watts of power and has a surface temperature of about 10,000 degrees Fahrenheit.

But the real cosmic behemoths are Wolf-Rayet stars, which are more than 20 times as massive as the Sun and at least five times as hot. Because these stars are relatively rare and often obscured, scientists don't know much about how they form, live and die. But this is changing, thanks to an innovative sky survey called the intermediate Palomar Transient Factory (iPTF) that uses resources at NERSC and ESnet to expose fleeting cosmic events such as supernovae.

For the first time ever, in 2014 scientists directly confirmed that a Wolf-Rayet star—sitting 360 million light years away in the Bootes constellation—died in a violent explosion known as a Type IIb supernova. Using the iPTF pipeline, researchers at Israel's Weizmann Institute of Science caught supernova SN 2013cu within hours of its explosion. They then triggered ground- and space-based telescopes to observe the event approximately 5.7 hours and 15 hours after it self-destructed. The findings were published May 22, 2014, in *Nature*.

Their detailed study of SN 2013cu's spectrum (the distribution of colors comprising the light from the supernova) using a technique called "flash spectroscopy" revealed the signature of a wind blown by the aging star just prior to its terminal explosion, and allowed the researchers to determine what elements were abundant on the surface of the dying star as it was about to explode as a supernova. This provided important information about how massive stars evolve just prior to their death, and the origin of crucial elements such as carbon, nitrogen and oxygen.

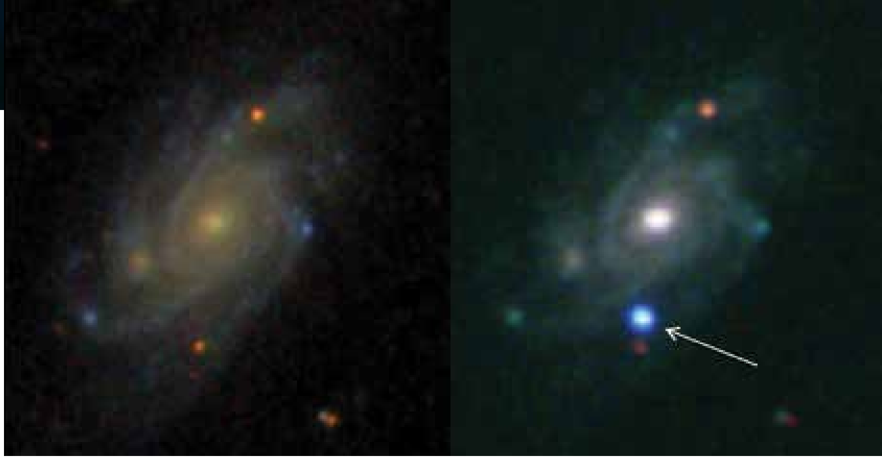
Some supermassive stars become Wolf-Rayets in the final stages of their lives. Scientists find these stars interesting because they enrich galaxies with the heavy chemical elements that eventually become the building blocks for planets and life.

"We are gradually determining which kinds of stars explode and why, and what kinds of elements they produce," said Alex Filippenko, professor of astronomy at UC Berkeley and a study co-author. "These elements are crucial to the existence of life. In a very real sense, we are figuring out our own stellar origins."

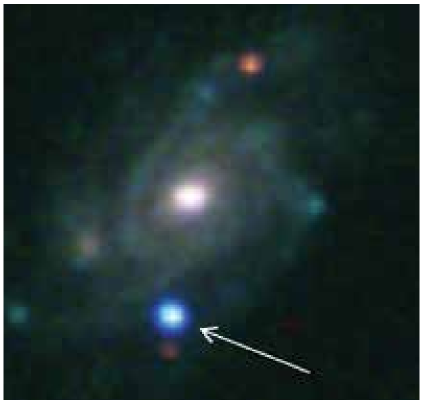
Going Supernova

All stars—no matter what size—spend their lives fusing hydrogen atoms to create helium. The more massive a star, the more gravity it wields, which accelerates fusion in the star's core, generating energy

PROJECT TITLE	NERSC PI	LEAD INSTITUTION	NERSC RESOURCES USED	DOE PROGRAM OFFICE
Palomar Transient Factory	Peter Nugent	Lawrence Berkeley National Laboratory	Carver	HEP—Cosmic Frontier



◂ A star in a distant galaxy explodes as a supernova: While observing a galaxy known as UGC 9379 (left; image from the Sloan Digital Sky Survey; SDSS) located about 360 million light years away from Earth, the team discovered a new source of bright blue light (right, marked with an arrow; image from the 60-inch robotic telescope at Palomar Observatory). This very hot, young supernova marked the explosive death of a massive star in that distant galaxy. *Images: Avishay Gal-Yam, et al, Weizmann Institute of Science*



◂ A detailed study of SN 2013cu's spectrum (the distribution of colors composing the light from the supernova) using a technique called "flash spectroscopy" revealed the signature of a wind blown by the aging star just prior to its terminal explosion, and allowed scientists to determine what elements were abundant on the surface of the dying star as it was about to explode as a supernova, providing important information about how massive stars evolve just prior to their death, and the origin of crucial elements such as carbon, nitrogen and oxygen.

to counteract gravitational collapse. When hydrogen is depleted, a supermassive star continues to fuse even heavier elements until its core turns to iron. At this point, atoms (and even subatomic particles) are packed in so closely that fusion no longer releases energy into the star. It is now solely supported by electron degeneracy pressure—the quantum mechanical law that prohibits two electrons from occupying the same quantum state.

When the core is massive enough, even electron degeneracy won't support the star and it collapses. Protons and electrons in the core merge, releasing a tremendous amount of energy and neutrinos. This, in turn, powers a shockwave that tears through the star ejecting its remains violently into space as it goes supernova. The Wolf-Rayet phase occurs before the supernova. As nuclear fusion slows, the heavy elements forged in the star's core rise to the surface, setting off powerful winds. These winds shed a tremendous amount of material into space and obscure the star from prying telescopes on Earth.

"When a Wolf-Rayet star goes supernova, the explosion typically overtakes the stellar wind and all information about the progenitor star is gone," said co-author Peter Nugent, who headed Berkeley Lab's Computational Cosmology Center and leads the Berkeley contingent of the iPTF collaboration. "We got lucky with SN 2013cu—we caught the supernova before it overtook the wind."

The iPTF team managed to capture SN 2013cu's chemical light signatures with the ground-based Keck telescope in Hawaii and saw the telltale signs of a Wolf-Rayet star. When they performed follow-up observations 15 hours later using NASA's Swift satellite, the supernova was still quite hot and strongly emitting in the ultraviolet range. In the following days, iPTF collaborators rallied telescopes around the globe to watch the supernova crash into material that had been previously ejected from the star. As the days went by, the researchers were able to classify SN 2013cu as a Type IIb supernova because of the weak hydrogen signatures and strong helium features in the spectra that appeared after the supernova cooled.

"With a series of observations, including data I took with the Keck-I telescope 6.5 days after the explosion, we could see that the supernova's expanding debris quickly overtook the flash-ionized wind that had revealed the Wolf-Rayet features," said Filippenko.

PUBLICATION

Avishay Gal-Yam, I. Arcavi, S. Ben-Ami, S.B. Cenko, M.M. Kasliwal, Y. Cao, O. Yaron, D. Tal, J.M. Silverman, A. Horesh, A. De Cia, F. Taddia, J. Sollerman, D. Perley, P.M. Vreeswijk, S.R. Kulkarni, P.E. Nugent, A.V. Filippenko, J.C. Wheeler, "A Wolf-Rayet-like progenitor of SN 2013cu from spectral observations of a stellar wind," *Nature*, May 22, 2014, 509, 471-474, doi: 10.1038/nature13304

FULL STORY

http://www.nersc.gov/news-publications/news/science-news/2014/confirmed-stellar-behemoth-self-destructs-in-type-iib-supernova/

Optimized Algorithms Boost Combustion Research

Methane Flame Simulations Run 6x Faster on NERSC's Hopper Supercomputer

"Modeling and simulation have become integral parts of the combustion design process."

Turbulent combustion simulations that provide input to the design of more fuel-efficient combustion systems have gotten their own efficiency boost, thanks to researchers from the Computational Research Division (CRD) at Berkeley Lab.

Matthew Emmett, Weiqun Zhang and John Bell developed new algorithmic features that streamline turbulent flame simulations, which play an important role in designing more efficient combustion systems. They tested the enhanced code on NERSC's Hopper supercomputer and achieved a dramatic decrease in simulation times. Their findings appeared in the June 2014 *Combustion Theory and Modelling*, and some of their simulations were featured on the journal cover.

More than 80 percent of energy consumed in the U.S. occurs via the burning of fossil fuels in transportation systems, heat and stationary power generation systems. This is why research into new fuels and more efficient engine technologies offers enormous potential for savings as well as for pollutant reduction. Shortening the design cycle of new fuels optimally tailored to work with new fuel-efficient engines requires fundamental advances in combustion science.

This is where applied mathematics and supercomputers come in. Modeling and simulation have become integral parts of the combustion design process. A good simulation can inform experimental design to assure the return of high-quality experimental data and enhance the analysis of physical phenomena.

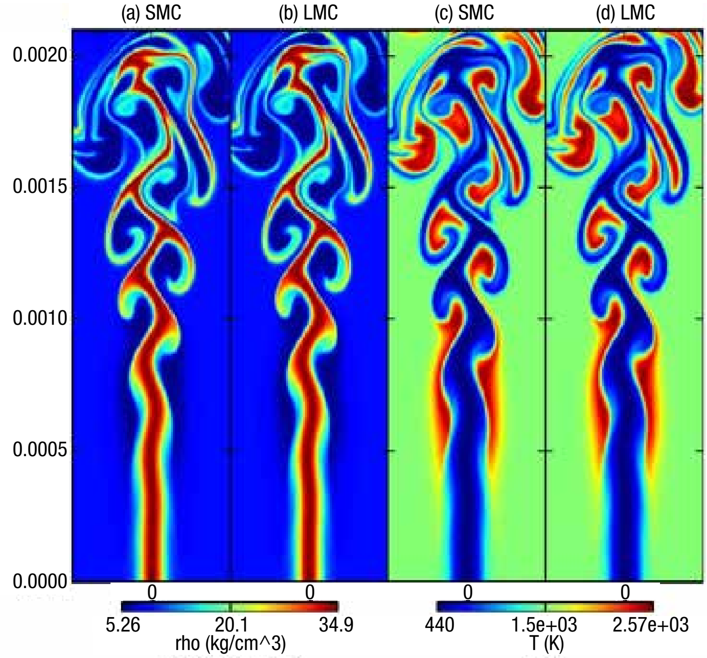
In the supercomputing world, a key component of combustion research is direct numerical simulation (DNS). The combustion community has experimented with various DNS approaches for more than 25 years, but the advent of manycore computer architectures is prompting computational scientists to update DNS codes to ensure that they can run efficiently on next-generation systems.

"The thing that we as applied mathematicians want to get across is that the systems you are running these codes on are changing, and the code you are running now is not going to necessarily be efficient on future machines," said Emmett, a postdoc in the CRD. "And to get it to be efficient isn't just a matter of getting somebody to optimize the code—you need new algorithms."

Time-stepping Strategies

One of the standard computational tools used in DNS studies of combustion is the numerical integration of the reacting, compressible Navier-Stokes equations. Although these types of codes have been successful, they suffer from some inherent weaknesses, according to Emmett. For one thing, the chemical equations function at a different time scale than the flow equations, which affects how

PROJECT TITLE	NERSC PI	LEAD INSTITUTION	NERSC RESOURCES USED	DOE PROGRAM OFFICE
Simulation and Analysis of Reacting Flows	John Bell	Lawrence Berkeley National Laboratory	Hopper, Edison	ASCR



◄ This simulation illustrates the total mass density (left) and temperature (right) of a dimethyl ether jet fuel simulation. It is a snapshot of the solution that corresponds to a physical time of 0.00006 seconds. *Image: Matthew Emmett, Weiqun Zhang, Lawrence Berkeley National Laboratory*

efficiently the code will run. In addition, manycore architectures make it more computationally expensive to access memory, particularly when trying to communicate across the network to access data stored somewhere else.

"What we are seeing is that floating point operations, or flop/s, are becoming relatively cheap, while network operations and memory accesses are becoming more expensive," he said. "So if you rewrite your algorithm to do more floating point operations and do less communication, then you should get a faster code that is more efficient because you're doing the expensive part less frequently."

Thus Emmett and Zhang developed a hybrid OpenMP/MPI parallel DNS code called SMC designed to simulate turbulent combustion on next-generation architectures. In particular, it can more robustly handle cases where the time scales of the advection, diffusion and chemical processes are considerably different, Emmett explained.

Two new algorithmic features of the code are a narrow stencil finite-difference algorithm and a multi-rate time-stepping strategy. Implementing a narrow stencil rather than the traditional wide stencil increases the number of floating point operations but reduces the amount of communication between cores, while the multi-rate integrator enables the chemistry to be advanced on a different time scale than the fluid dynamics.

"The big improvement here is the multi-rate integration and the way we separate the advection-diffusion processes and the chemistry process," Emmett said. "In traditional codes these things are locked together and you have to march them forward at the same time, which has consequences for how efficient your code can be. But we separate them to make the code run more efficiently. And even though the two processes have been separated, overall a very tight coupling is maintained between them. This is one of the improvements over previous DNS codes."

For example, a methane flame simulation run with the multi-rate integrator performed six times faster on Hopper compared to the same code with the multi-rate integrator turned off. And in a series of dimethyl ether (DME) jet simulations designed to demonstrate how the SMC code could be used to study complex new fuels, the multi-rate integrator was able to operate with a larger time-step than single-rate integrators and thus obtained an accurate solution in less time.

"The DME jet is an example of a difficult problem that is essentially infeasible, in the context of DNS codes, without using advanced algorithms such as the multi-rate integration scheme in SMC," Emmett said. "Some of the chemical reactions in a DME flame happen very quickly and restrict the size of the numerical time-step that some codes can take. With a code that doesn't have a multi-rate integrator, trying to run the DME jet simulation would take a very, very long time."

PUBLICATION

M. Emmett, W. Zhang, J. B. Bell, "High-Order Algorithms for Compressible Reacting Flow with Complex Chemistry," *Combustion Theory and Modelling*, 18(3), June 2014, pp. 361-387, doi: 10.1080/13647830.2014.919410

FULL STORY

http://www.nersc.gov/news-publications/nersc-news/science-news/2014/optimized-algorithms-boost-combustion-research/

Supercomputers Fuel Global High-Resolution Climate Models

Climate Science Entering New Golden Age

"I've been calling this a golden age for high-resolution climate modeling because these supercomputers are enabling us to do gee-whiz science in a way we haven't been able to do before."

Not long ago, it would have taken several years to run a high-resolution simulation on a global climate model. But using supercomputing resources at NERSC, in 2014 Berkeley Lab climate scientist Michael Wehner was able to complete a run in just three months.

What he found was that not only were the simulations much closer to actual observations, but the high-resolution models were far better at reproducing intense storms, such as hurricanes and cyclones. The study was published in the *Journal for Advances in Modeling the Earth System*.

"I've been calling this a golden age for high-resolution climate modeling because these supercomputers are enabling us to do gee-whiz science in a way we haven't been able to do before," said Wehner, who was also a lead author for the recent Fifth Assessment Report of the Intergovernmental Panel on Climate Change (IPCC). "These kinds of calculations have gone from basically intractable to heroic to now doable."

Using version 5.1 of the Community Atmospheric Model, developed by the DOE and the National Science Foundation for use by the scientific community, Wehner and his co-authors conducted an analysis for the period 1979 to 2005 at three spatial resolutions: 25 km, 100 km and 200 km. They then compared those results to each other and to observations.

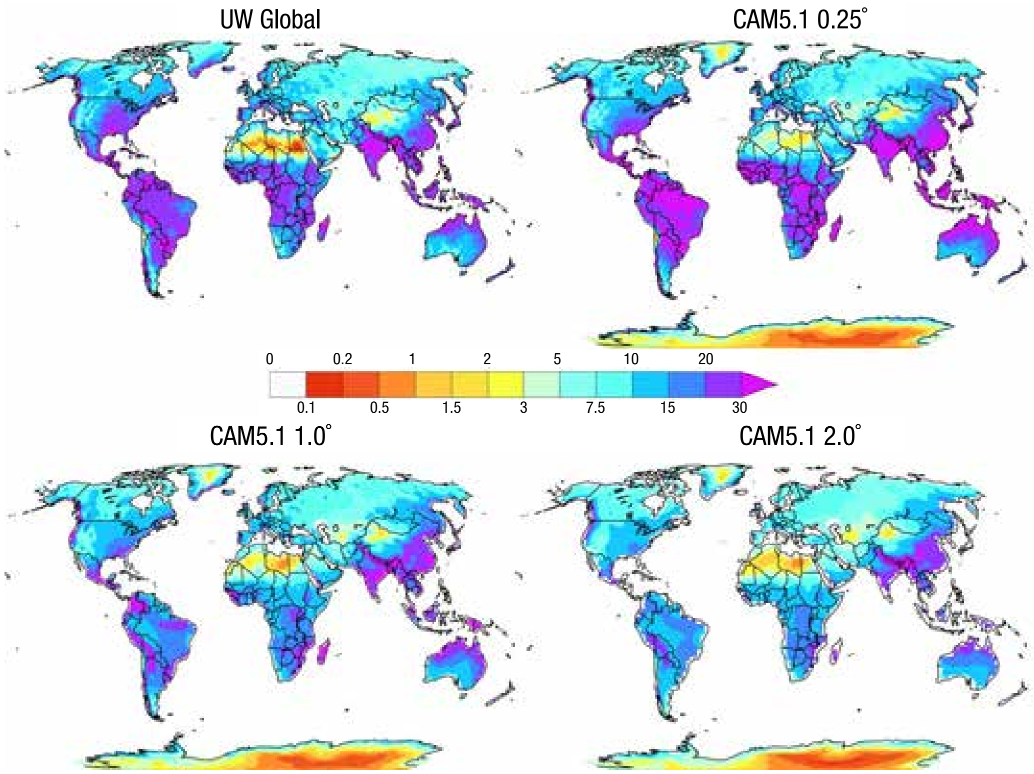
One simulation generated 100 terabytes of data. Wehner ran the simulations on NERSC's Hopper supercomputer.

"I've literally waited my entire career to be able to do these simulations," Wehner said.

The higher resolution was particularly helpful in mountainous areas since the models take an average of the altitude in the grid (25 square km for high resolution, 200 square km for low resolution). With more accurate representation of mountainous terrain, the higher resolution model is better able to simulate snow and rain in those regions.

"High resolution gives us the ability to look at intense weather like hurricanes," said Kevin Reed, a researcher at the National Center for Atmospheric Research and a co-author on the paper. "It also gives us the ability to look at things locally at much higher fidelity. Simulations are much more realistic at any given place, especially if that place has a lot of topography."

PROJECT TITLE	NERSC PI	LEAD INSTITUTION	NERSC RESOURCES USED	DOE PROGRAM OFFICE
CASCADE: CAlibrated and Systematic Characterization, Attribution and Detection of Extremes	Bill Collins	Lawrence Berkeley National Laboratory	Hopper	BER—Climate and Environmental Sciences



▲ Simulated and observed annual maximum five day accumulated precipitation over land points, averaged. Observations are calculated from the period 1979 to 1999. Model results are calculated from the period 1979 to 2005. *Image: National Center for Atmospheric Research*

The high-resolution model produced stronger storms and more of them, which was closer to the actual observations for most seasons. "In the low-resolution models, hurricanes were far too infrequent," Wehner said.

The IPCC chapter on long-term climate change projections concluded that a warming world will cause some areas to be drier and others to see more rainfall, snow and storms. Extremely heavy precipitation was projected to become even more extreme in a warmer world.

"I have no doubt that is true," Wehner said. "However, knowing it will increase is one thing, but having confidence about how much and where as a function of location requires the models do a better job of replicating observations than they have."

Wehner says the high-resolution models will help scientists to better understand how climate change will affect extreme storms. His next project is to run the model for a future-case scenario. Further down the line, Wehner believes scientists will be running climate models with 1 km resolution. To do that, they will have to have a better understanding of how clouds behave.

"A cloud system-resolved model can reduce one of the greatest uncertainties in climate models, by improving the way we treat clouds," Wehner said. "That will be a paradigm shift in climate modeling. We're at a shift now, but that is the next one coming."

PUBLICATION

M. F. Wehner, K. A. Reed, F. Li, Prabhat, J. Bacmeister, C. T. Chen, C. Paciorek, P. J. Gleckler, K. R. Sperber, W. D. Collins, A. Gettelman, C. Jablonowski, "The Effect of Horizontal Resolution on Simulation Quality in the Community Atmospheric Model, CAM5.1," *Journal of Advances in Modeling Earth Systems*, 6(4), 980-997, December 2014, doi: 10.1002/2013MS000276

FULL STORY

http://www.nersc.gov/news-publications/nersc-news/science-news/2014/supercomputers-fuel-global-high-resolution-climate-models/

Decoding the Molecular Mysteries of Photosynthesis

Understanding the Inner Workings of Photosynthesis Is Key to Building New Man-Made Energy Resources

"Our goal was to develop a physical understanding of why the proteins are organized the way they are."

At first glance, photosynthesis seems elegant in its simplicity. Only three components—light, water and carbon dioxide (CO_2)—are needed for plants and other organisms to convert light into chemical energy, all in a matter of minutes.

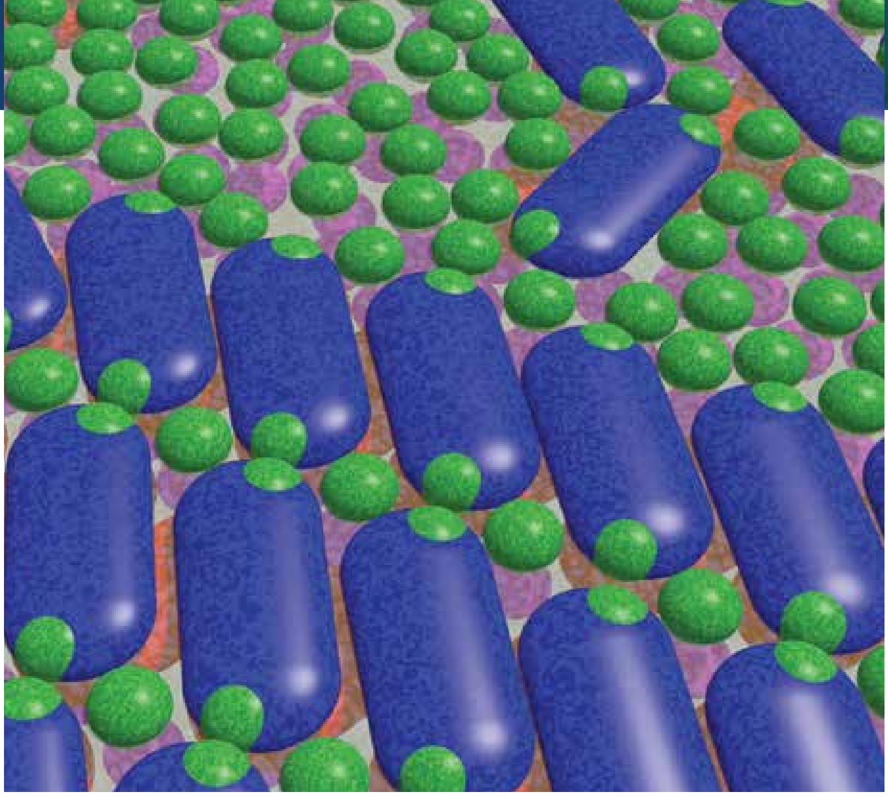
But upon closer inspection, the process is much more complicated. How, for example, does the plant know how much sunlight, water or CO_2 it needs at any given point in time? And what if the system breaks—how does the plant repair it? Scientists are working to answer these questions to enable the development of artificial photosynthesis systems that can generate "green" alternatives to fossil fuels. But first they must pinpoint what happens at the molecular level that allows photosynthesis to occur in the first place.

Toward that end, simulations by researchers from UC Berkeley and Berkeley Lab using supercomputers at NERSC revealed new insights about the inner workings of the photosynthetic process. Their findings were published in *Biophysical Journal*.

"In artificial photosynthesis the systems can't repair themselves if they break and they can't tune their efficiency up and down based on what the system needs," explained lead author Anna Schneider, a graduate of the Biophysics Graduate Group at UC Berkeley. "But these are things that plants can do, and we are trying to figure out the physics that underlie that so someone down the line can build something that takes advantage of this."

All photosynthesis occurs in a plant's chloroplasts—organelles contained within the plant cells. A typical plant cell can contain up to 50 chloroplasts. A chief component of the chloroplasts is the thylakoid membrane, which contains discs of protein assemblies arranged in stacks (grana) that

PROJECT TITLE	NERSC PI	LEAD INSTITUTION	NERSC RESOURCES USED	DOE PROGRAM OFFICE
Self-assembly and Fluctuations in Nanoscale Systems	Philip Geissler	Lawrence Berkeley National Laboratory	Hopper, Scratch	BES—Chemical Sciences



◂ Two protein assemblies in a plant cell's chloroplasts—Photosystem II (blue and red) and light-harvesting complex II (green and purple)—are key to initiating photosynthesis. This visualization illustrates how the proteins organize themselves to make this happen. *Image: Anna Schneider, University of California, Berkeley, and Lester Hodges, Lawrence Berkeley National Laboratory*

resemble towers of tiny green poker chips. In recent years much photosynthesis research has focused on teasing apart one of these protein assemblies: Photosystem II (PSII), a large complex containing dozens of proteins and hundreds of chlorophylls and other pigments that sets an upper limit for overall photosynthetic efficiency and productivity.

Photosynthetic efficiency relies on precise spatial organization of PSII and its associated light-harvesting complex II (LHCII), both of which are concentrated in the grana stacks. But how the proteins know to arrange themselves is unclear, according to Schneider and her co-author, Phillip Geissler, professor of chemistry at UC Berkeley. This makes it difficult to determine exactly how they regulate a plant's light harvesting and energy conversion efficiencies.

"The grana protein organization affects the efficiency of photosynthesis overall," Schneider said. "Our goal was to develop a physical understanding of why the proteins are organized the way they are."

To do this they created a nanoscale computational model of LHCII and PSII and used it to analyze the protein organization. The model is simple enough to access large length scales, yet rich enough to capture structural motifs such as inter-membrane correlations, Schneider explained. In addition, it can predict thermodynamic phase behavior that could affect photosynthetic function. The computer model was designed to enable researchers to make this distinction, and to do so with a very small set of variables.

"That was the point of our simulations: to pin down that it's not just that the PSII and LHCII happen to be next to each other in a way that looks like a crystal—it is a true thermodynamic crystal with true separate phases," Schneider said. This discovery means that scientists can apply everything they already know about the thermodynamics of ice and water to what's going on inside plants, she added.

Schneider collaborated with Lester Hedges, a post-doc working in Berkeley Lab's Molecular Foundry, to create 3D images of the data using a visualization tool she developed and POV-Ray, a free software tool. One of these images was featured on the cover of *Biophysical Journal*.

"It may not look like there is that much data in some of the figures, but it actually took several hundred cores on Hopper running for several weeks to develop that data," Schneider said.

PUBLICATION

Anna R. Schneider, Phillip L. Geissler, "Coexistence of Fluid and Crystalline Phases of Proteins in Photosynthetic Membranes," *Biophysical Journal*, 105(5), 1161–1170, September 3, 2013, doi: http://dx.doi.org/10.1016/j.bpj.2013.06.052

FULL STORY

http://www.nersc.gov/news-publications/news/science-news/2014/decoding-photosynthesis-molecular-mysteries/

Supercomputer Helps Model 3D Map of Adolescent Universe

Researchers Demonstrate Novel Technique for High-Resolution Universe Maps

"Stark and White used large cosmological simulations run at NERSC to construct mock data and to test the robustness of the maps."

Using extremely faint light from galaxies 10.8 billion light years away, scientists have created one of the most complete three-dimensional maps of a slice of the adolescent universe—just 3 billion years after the Big Bang.

The map shows a web of hydrogen gas that varies from low to high density at a time when the universe was made of a fraction of the dark matter we see today. It was created in part using supercomputing resources at NERSC by a team that included researchers from Berkeley Lab's Computational Cosmology Center (C3).

The study, led by Khee-Gan Lee and his team at the Max Planck Institute for Astronomy in conjunction with researchers Berkeley Lab and UC Berkeley, was published October 16, 2014 in *Astrophysical Journal Letters*.

In addition to providing a new map of part of the universe at a young age, the work demonstrates a novel technique for high-resolution universe maps, according to David Schlegel, an astrophysicist at Berkeley Lab and study co-author. Similar to a medical computed tomography (CT) scan, which reconstructs a 3D image of the human body from the X-rays passing through a patient, Lee and his colleagues reconstructed their map from the light of distant background galaxies passing through the cosmic web's hydrogen gas.

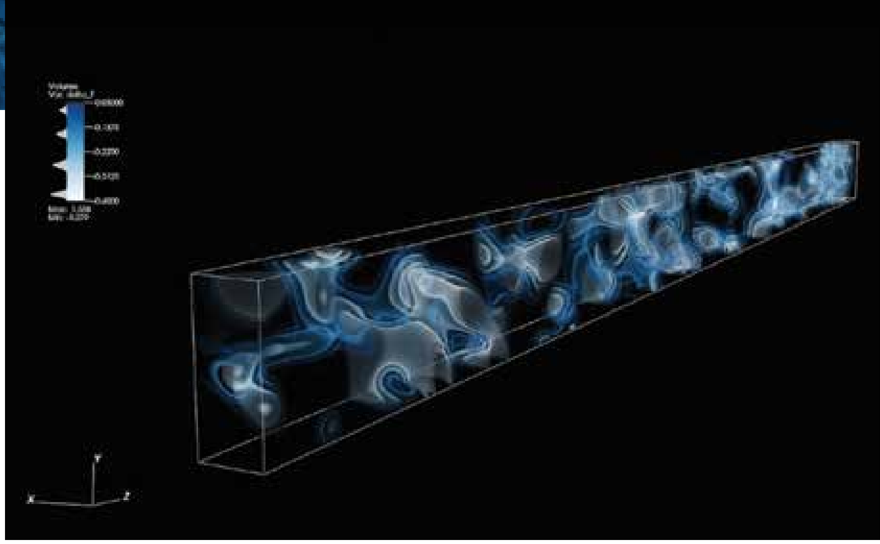
The new technique might inform future mapping projects such as the proposed Dark Energy Spectroscopic Instrument (DESI). Managed by Berkeley Lab, DESI has the goal of producing the most complete map of the universe yet.

"DESI was designed without the possibility of extracting such information from the most distant, faint galaxies," said Schlegel. "Now that we know this is possible, DESI promises to be even more powerful."

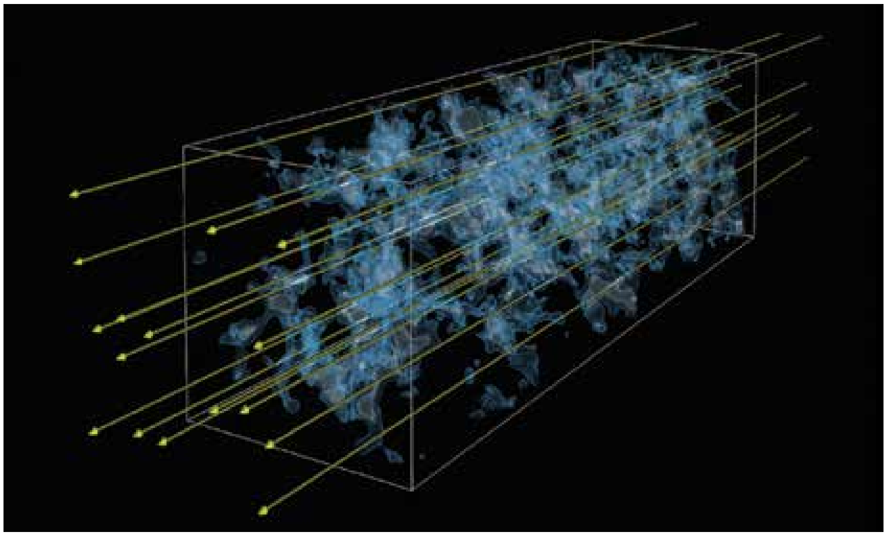
Before this study, no one knew if galaxies farther than 10 billion light years away could provide enough light to be useful, Schlegel said. But in early 2014, the team collected four hours of data on the Keck-1 telescope during a brief break in cloudy skies.

"It turned out to be enough time to prove we could do this," Schlegel said.

PROJECT TITLE	NERSC PI	LEAD INSTITUTION	NERSC RESOURCES USED	DOE PROGRAM OFFICE
Cosmological Simulations for Sky Surveys	Julian Borrill	Lawrence Berkeley National Laboratory	Edison	HEP—Cosmic Frontier



◀ 3D map of the cosmic web at a distance of 10.8 billion years from Earth, generated from imprints of hydrogen gas observed in the spectrum of 24 background galaxies behind the volume. This is the first time large-scale structures in such a distant part of the universe have been directly mapped. *Images: Casey Stark, University of California, Berkeley; Khee-Gan Lee, Max Planck Institute for Astronomy*



◀ Artist's impression illustrating the technique of Lyman-alpha tomography: as light from distant background galaxies (yellow arrows) travel through the Universe towards Earth, they are imprinted by the absorption signatures from hydrogen gas tracing in the foreground cosmic web. By observing several background galaxies in a small patch of the sky, astronomers were able to create a 3D map of the cosmic web using a technique similar to computed tomography scans.

However, the galaxies' light was indeed exceedingly faint. In order to use it for a map, the researchers needed to develop algorithms to subtract light from the sky that would otherwise drown out the galactic signals. Schlegel developed the algorithm to do this, while Casey Stark—a UC Berkeley astrophysics graduate student who works in Berkeley Lab's Computational Research Division and in the C3—and Martin White, a UC Berkeley physicist and theoretical cosmologist, developed a code dubbed "Dachshund" that uses a Wiener filter signal processing technique to reconstruct the 3D field from which the signal is drawn.

Stark and White used large cosmological simulations run at NERSC to construct mock data and to test the robustness of the maps. Running mock reconstructions on NERSC's Edison system was a crucial step in validating the method and code, Stark emphasized. After optimizing the tomography code on Edison, he and White collaborated with Lee to process the data for the *Astrophysical Journal Letters* study.

"The signal processing technique we used is much like what you would do for a CT scan, where you have a sparse sampling of a field and you want to fill in the gaps. In our case, we observe the absorption along lines to distant galaxies and then infer the amount of absorption between the lines," Stark explained. "The technique is simple, but it's an expensive computation. Fortunately, we realized we could simplify the computations by tailoring them to this particular problem and thus use much less memory. With that simplification, the code is very easy to parallelize."

PUBLICATION

K. G. Lee, J. F. Hennawi, C. Stark, J. X. Prochaska, M. White, D. J. Schlegel, A. C. Eiler, A. Arinvo-i-Prats, N. Suzuki, R. A.C. Croft, K. I. Caputi, P. Cassata, O. Ilbert, B. Garilli, A. M. Koekemoer, V. Le Brun, O. Le Fevre, D. Maccagni, P. Nugent, Y.Taniguchi, L. A.M. Tasca, L. Tresse, G.Zamorani, E. Zucca, "Lyα Forest Tomography from Background Galaxies: The first Megaparsec-Resolution Large-Scale Structure Map at z>2," *Astrophysical Journal Letters*, 795, L12 (2014), doi: 10.1088/2041-8205/795/1/L12

FULL STORY

https://www.nersc.gov/news-publications/nersc-news/science-news/2014/supercomputer-helps-model-3d-map-of-adolescent-universe/

Disordered Materials Hold Promise for Better Batteries

Supercomputers Find Surprising Link between Disordered Materials and Lithium Batteries

"The computer runs also helped the researchers determine the concentration of lithium excess required to produce an effective battery."

Because lithium batteries are among the lightest and most energetic rechargeable batteries available, much research is focused on facilitating their use in electronic devices. Using supercomputers at NERSC and other facilities, researchers at the Massachusetts Institute of Technology (MIT) and Brookhaven National Laboratory have found a new avenue for such research: the use of disordered materials, which had generally been considered unsuitable for batteries.

In a rechargeable lithium-based battery, lithium ions—atoms that have given up an electron and thus carry a net charge—are pulled out of the battery's cathode during the charging process and returned to the cathode as power is drained. But these repeated round-trips can cause the electrode material to shrink and expand, degrading performance over time.

In today's lithium batteries, cathodes are usually made of an orderly crystalline material, sometimes in a layered structure. When slight deviations from that perfect order are introduced, the battery's efficiency generally goes down. Thus disordered materials have mostly been ignored in the search for improved battery materials.

But a series of computer models and laboratory experiments have shown that certain kinds of disorder provide a significant boost in cathode performance. These surprising findings were published in January 2014 in *Science*.

New Pathways for Lithium Ions

These days, lithium battery cathodes are striated materials usually made up of alternating layers of lithium and transition-metal oxides. The materials also typically contain an exact balance of lithium and metal atoms. Until recently, scientists believed that this layering was necessary to provide a clear pathway for lithium to pass in and out of the cathode without bumping into the transition metal oxide layer because high lithium ion mobility is essential for an efficient rechargeable battery and disorder—or so it has long been thought—could significantly reduce mobility.

But it turns out that disordered materials with irregular pathways can still act as efficient channels for lithium mobility—there just needs to be an excess of lithium atoms. And because lithium ions don't push layers out of shape in disordered materials as the battery charges, this could mean safer and longer-lasting batteries.

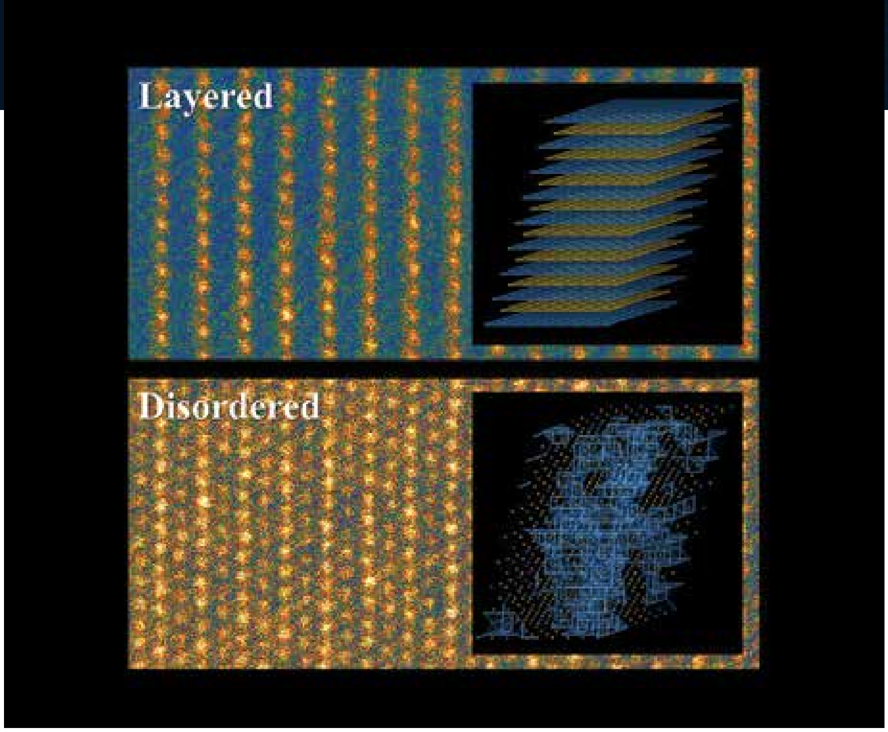
PROJECT TITLE
Testing the Fundamental Limits on the Number of Inorganic Compounds in Nature

NERSC PI
Gerbrand Ceder

LEAD INSTITUTION
Massachusetts Institute of Technology

NERSC RESOURCES USED
Hopper

DOE PROGRAM OFFICE
BES—Materials Science



◀ Conventional layered lithium and transition metal cathode material (top) and the new disordered material studied by researchers at Massachusetts Institute of Technology (bottom) as seen through a scanning transmission electron microscope. Inset images show diagrams of the different structures in these materials. (In the disordered material, the blue lines show the pathways that allow lithium ions to traverse the material.) *Image: Massachusetts Institute of Technology*

"In most other lithium materials, as you pull the lithium in and out, it changes dimension, swelling or contracting, which can cause all sorts of problems, including fatigue that can lead to cracking," said co-author Gerbrand Ceder, professor of materials science and engineering at MIT.

But a new disordered material—lithium molybdenum chromium oxide—has shown a very high dimensional stability. While the dimensional changes in layered materials can be as much as 5 to 10 percent, in the new disordered material it is only about 0.1 percent, according to Ceder.

While lithium molybdenum chromium oxide can hold and release significantly more lithium than existing materials, it produces a lower voltage—meaning its overall performance is about the same as that of existing materials, according to Ceder. Despite this revelation, this new finding is significant because it opens a whole new category of possibilities that had previously been ignored in the area of battery research, he added.

Supercomputers: The Game Changer

To quickly accomplish the high-throughput screening of cathode materials, the researchers relied on the Materials Project, co-founded by Ceder, which aims to take the guesswork out of finding the best material for a job by making the characteristics of every inorganic compound available to any interested scientist.

In this case, once the team had a list of candidates, co-author Jinhyuk Lee of MIT began synthesizing compounds and found that molybdenum chromium oxide worked very well as a cathode. But surprisingly, the material did not have a traditional layered look—it was disordered.

To solve this mystery, the team used NERSC's Hopper system to compute lithium migration barriers in a number of different disordered structures and diffusion channels. The calculations showed that active diffusion channels in disordered materials are different than those in traditional materials. Unlike traditional cathodes, which have an exact balance of transition metals and lithium ions, the disordered materials require an excess of lithium to be effective. The computer runs also helped the researchers determine the concentration of lithium excess required to produce an effective battery.

"Our analysis shows a new direction that we can take in searching for even better materials for batteries," said co-author Alexander Urban, a post-doc at MIT. "It is also a great example of the extraordinary science that can be accomplished when simulations computed in real-time feed into laboratory experiments and vice versa."

PUBLICATION

J. Lee, A. Urban, X. Li, D. Su, G. Hautier, G. Ceder, "Unlocking the Potential of Cation-Disordered Oxides for Rechargeable Lithium Batteries," *Science*, 343, 519-522, January 31, 2014, doi: 10.1126/science.1246432

FULL STORY

http://www.nersc.gov/news-publications/nersc-news/science-news/2014/disordered-materials-hold-promise-for-better-batteries/

Calming Plasma's Stormy Seas

Overcoming Ion Instabilities in Hot Plasma Can Boost Fusion Reactor's Energy Output

"Researchers have been working to pinpoint both what causes this turbulence and how to control or even eliminate it."

For decades, controlled nuclear fusion has held the promise of a safe, clean, sustainable energy source that could help wean the world from fossil fuels. But the challenges of harnessing the power of the sun in an Earth-based nuclear fusion reactor have been many, with much of the progress over the last several years coming in incremental advances.

One of the key technical issues is actually a common occurrence in fusion reactions: plasma turbulence. Turbulence inside a reactor can increase the rate of plasma heat loss, significantly reducing the resulting energy output. So researchers have been working to pinpoint both what causes this turbulence and how to control or even eliminate it.

Simulations run at NERSC are helping researchers better understand a central piece of the puzzle: the relationship between fast ion particles in the plasma and plasma microturbulence.

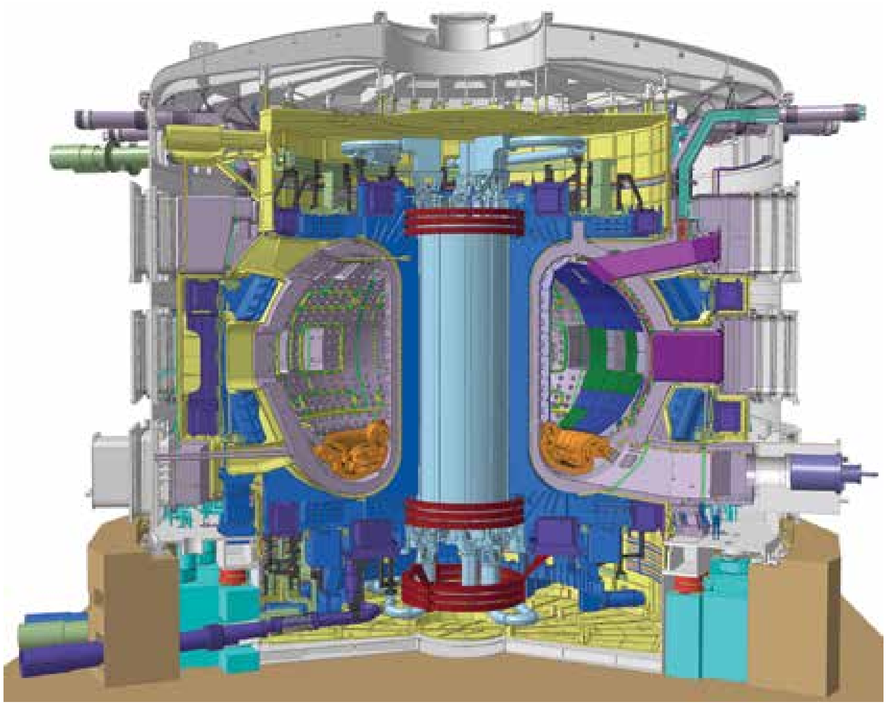
Ion-temperature-gradient Instabilities

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 centigrade—which ionizes the gas and transforms it into plasma.

A tokamak reactor uses a torus (doughnut-shaped) vessel and extremely strong magnetic fields to confine the energy of the heated plasma by a sufficient degree to ensure a net fusion energy gain. But in a tokamak, turbulence caused by microinstabilities in the plasma—particularly ion-temperature-gradient (ITG) instabilities—can impair the energy confinement.

"The ITG mode is a ubiquitous driver of tokamak microturbulence, which in general is driven by the very large gradients of temperature and density in the plasma," explained Jonathan Citrin of the Dutch Institute for Fundamental Energy Research and lead author of a study published in *Physical Review Letters*. "This is essentially one of the limiting factors of the total fusion power one can achieve with tokamaks. The more we can stabilize these ITG modes, the more we can reduce the size and cost of the machine for the same total fusion power."

PROJECT TITLE	NERSC PI	LEAD INSTITUTION	NERSC RESOURCES USED	DOE PROGRAM OFFICE
Experimental Tests of Gyrokinetic Simulations of Microturbulence	David Mikkelsen, PPPL	Princeton Plasma Physics Laboratory	Edison	FES—Fusion Base Program



◀ Interior view of the ITER tokamak reactor under construction in Cadarache, France. In a tokamak, turbulence caused by microinstabilities in the plasma can significantly impact energy confinement. *Image: ITER*

Ion heat transport experiments led by Paola Mantica at the Joint European Torus (JET) tokamak in Oxfordshire, UK, demonstrated that in certain regimes, the ITG turbulence in the tokamak core could be reduced, leading to higher ion temperatures. But the mechanism for this stabilization was not fully understood.

So Citrin and collaborators from the European Union and the U.S. ran a series of computer simulations on NERSC's Edison system to test the findings of the JET experiments. Following extensive nonlinear gyrokinetic turbulence simulations with the GENE code—a software package written by Frank Jenko and his group at the Max Planck Institute for Plasma Physics—Citrin and his colleagues were able to explain the improved energy confinement seen in the JET experiments.

"The hypothesis following these experiments was that the relatively high rotation found in these plasmas was responsible for the reduced turbulence, but we found that this is not actually the case," Citrin explained. "Rather, it was a combination of reduced turbulence due to magnetic fluctuations and the inclusion of suprathermal ion species. When you heat the plasma, you accelerate beams of particles and inject them into the plasma. This also makes the plasma rotate, but correlated with the rotation is a significant source of fast ions that coexist with the thermal ions, and we found that this changed the thermal regime significantly."

Thanks to Edison and the large amount of computing time made available at NERSC, the researchers were able to simulate and explain this observation.

"In the past, most simulations were done where only the electric fields fluctuate. But with Edison we were able to include the magnetic fluctuations as well as additional ion species, which leads to less turbulence," Citrin said. "This is exciting because in future reactors, one does not expect to have strong rotation because of the large size and the amount of torque you can put in. But there will be fast ions due to the fusion reactions themselves. Increased ion temperature due to rotation does not extrapolate well to future devices, but this explanation does extrapolate well."

PUBLICATION

J. Citrin, F. Jenko, P. Mantica, D. Told, C. Bourdelle, J. Garcia, J. W. Haverkort, G. M. D. Hogeweij, T. Johnson, and M. J. Pueschel, "Nonlinear Stabilization of Tokamak Microturbulence by Fast Ions," *Physical Review Letters* 111, 155001, October 7, 2013, doi: http://dx.doi.org/10.1103/PhysRevLett.111.155001

FULL STORY

http://www.nersc.gov/news-publications/nersc-news/science-news/2014/calming-fusion-s-stormy-seas/

The Rise and Fall of Core-Collapse Supernovae

2D and 3D Models Reveal What Fuels an Exploding Star

"Understanding the explosion mechanism of supernovae remains among the great challenges of astrophysics."

Few events in the cosmos are as powerful as the explosion of a massive star in a supernova. For several weeks, the brightness of that single event can rival that of an entire galaxy. As a result, supernovae of various types are often used as "standard candles" in determining the size and expansion rate of a universe powered by dark energy.

Despite decades of research, however, understanding the explosion mechanisms of supernovae remains among the great challenges of astrophysics—in part because of the complexity of the computations involved. For core-collapse supernovae, which mark the death of a massive star, the heart of the matter—the core, if you will—lies in understanding the physics of the neutrinos that are produced in the neutron star that is born during the supernova explosion.

"The numerical study of supernova is a subject that goes back almost 50 years," said Raphael Hix, an astrophysicist at Oak Ridge National Laboratory (ORNL) who studies core-collapse supernovae. "The challenge has always been that the real problem is much more complicated than we can fit in any computer we have. So with each new generation of computers, we get to do the problem with a little more realism than before."

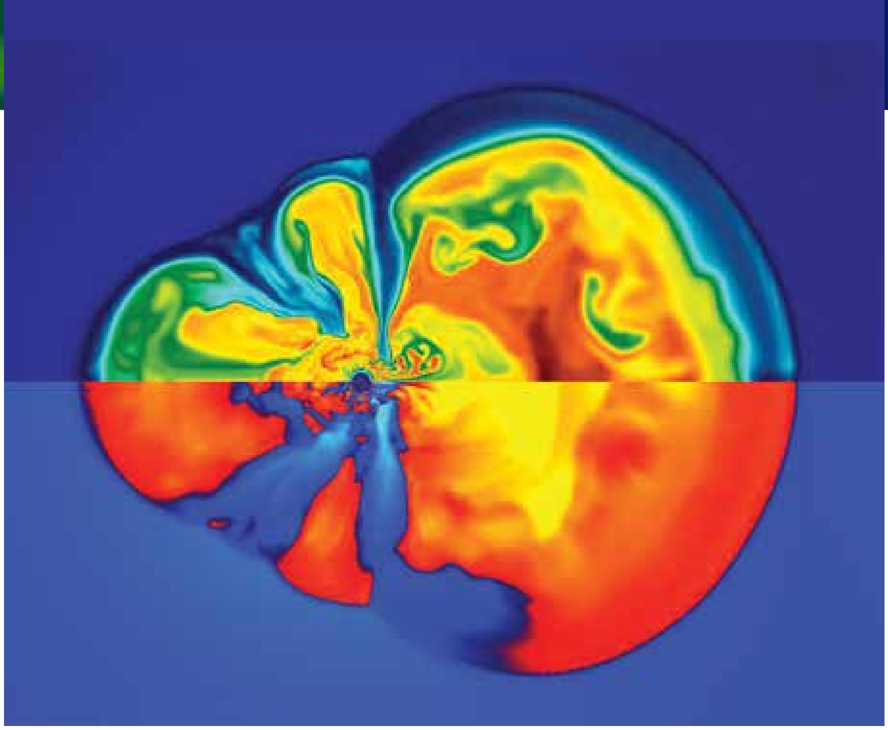
A wide variety of stars can die as core-collapse supernovae—stars anywhere from 10 to perhaps 40 solar masses in size, with different internal structures, Hix explained. Furthermore, stars being formed today in our galaxy are not the same as the stars of the same mass formed 13 billion years ago at the beginning of the universe.

"Our ability to wander through the parameter space of these kinds of stars that die is very limited in 3D, even with good-sized allocations on the fastest machines in the world. It's too expensive," Hix said. "So we've been trying to do one or two 3D simulations a year, and then dozens of 2D simulations to help us figure out what the important 3D simulations are."

Comparisons Critical

In a study published in *AIP Advances*, Hix and colleagues from ORNL, University of Tennessee, Florida Atlantic University and North Carolina State University compared 3D models run at ORNL with 2D models run at NERSC to gain new insights into the explosion mechanism behind core-collapse supernovae.

PROJECT TITLE	NERSC PI	LEAD INSTITUTION	NERSC RESOURCES USED	DOE PROGRAM OFFICE
Developing an Understanding of Core-Collapse Supernova Explosion Systematics Using 2D Chimera Simulations	Raphael Hix	Oak Ridge National Laboratory	Edison, Hopper	NP—Astrophysics



◀ This 2D simulation shows the evolution of the entropy (upper half) and radial velocity (lower half) of a supernova explosion at 300 milliseconds. Red material is moving outward, blue material inward. *Image: Raph Hix, Oak Ridge National Laboratory*

"The principal difference in the supernova research we're doing now, compared to the previous generation (on earlier supercomputers), is spectral neutrino transport," Hix said. "In the past, in multidimensional simulations, you would have treated neutrinos in terms of something like the mean temperature or energy, a single parameter for the whole spectrum of neutrinos as a function of space. But now we actually keep track of more than 20 different bins of neutrinos at different energies."

This advance led Hix and his co-authors to realize that previous core-collapse supernovae models overestimated the speed of the process, producing the explosions too rapidly.

"The process has been called 'delayed explosion mechanism' since the 1980s, but what we are finding is that the explosion is more delayed than we previously thought," he said. "Back then, the delay was 50 milliseconds, now it's 300 milliseconds. So we think—and it seems to be borne out by the 3D model we've run at ORNL and compared to 2D models at NERSC—that this is a general trend, that the better the physics, the slower the explosions are to develop."

This finding has a number of practical issues, he added. For example, the size of the neutron star that gets left behind is dependent upon how soon the explosion goes off because once the explosion pushes away the outer parts of the star, the growth of the neutron star stops. It also has an effect on what layers of the star get ejected versus becoming trapped within the neutron star.

Being able to compare the 2D and 3D models has been critical to enabling Hix and his collaborators to more effectively and efficiently identify trends in observable quantities like the explosion energy.

"Once we have handful of 3D models, we will know a lot more about what the 2D models are telling us, such as which things you can take from a 2D model that are good and which are wrong," he said. "Because the results of the models are close to the observed values like the explosion energy and mass of radioactive nickel expelled, we feel we are making progress, that these models have some semblance of reality."

Hix and his colleagues are looking forward to the next generation of supercomputers, which should allow them to drill down even further into the inner workings of core-collapse supernovae and create more usable 3D models in much less time.

"In the next decade, we will see three-dimensional models that finally include all of the essential physics that our current models suggest is necessary to understand core-collapse supernovae," they concluded in the *AIP Advances* paper.

PUBLICATION

W. R. Hix, E.J. Lentz, E. Endeve, M. Baird, M.A. Chertkow, J.A. Harris, O.E.B. Messer, A. Mezzacappa, S. Bruenn, J. Blondin, "Essential Ingredients in Core-Collapse Supernovae," *AIP Advances* 4, 041013 (2014), doi: http://dx.doi.org/10.1063/1.4870009

FULL STORY

https://www.nersc.gov/news-publications/nersc-news/science-news/2015/clocking-the-rise-and-fall-of-a-core-collapse-supernovae/

Atomic Switcheroo Explains Thin-Film Solar Cell Mystery

Computer Simulations Reveal Detailed Physics for First Time

"The research team's findings, in addition to providing a long-awaited explanation, could be used to guide engineering of higher efficiency CdTe solar cells."

Scientists have known since the 1980s that treating cadmium-telluride (CdTe) solar cell materials with cadmium-chloride improves efficiency, but the underlying physics have remained a mystery—until now. Combining electron microscopy with computer simulations run at NERSC, researchers have finally put this decades-long debate to rest. The finding could also lead to a less-expensive, more easily fabricated thin-film alternative to silicon-based photovoltaics.

Thin-film CdTe solar cells are considered a potential rival to silicon-based photovoltaic systems because of their theoretically low cost per power output and ease of fabrication. Their comparatively low historical efficiency in converting sunlight into energy, however, has limited the technology's widespread use, especially for home systems.

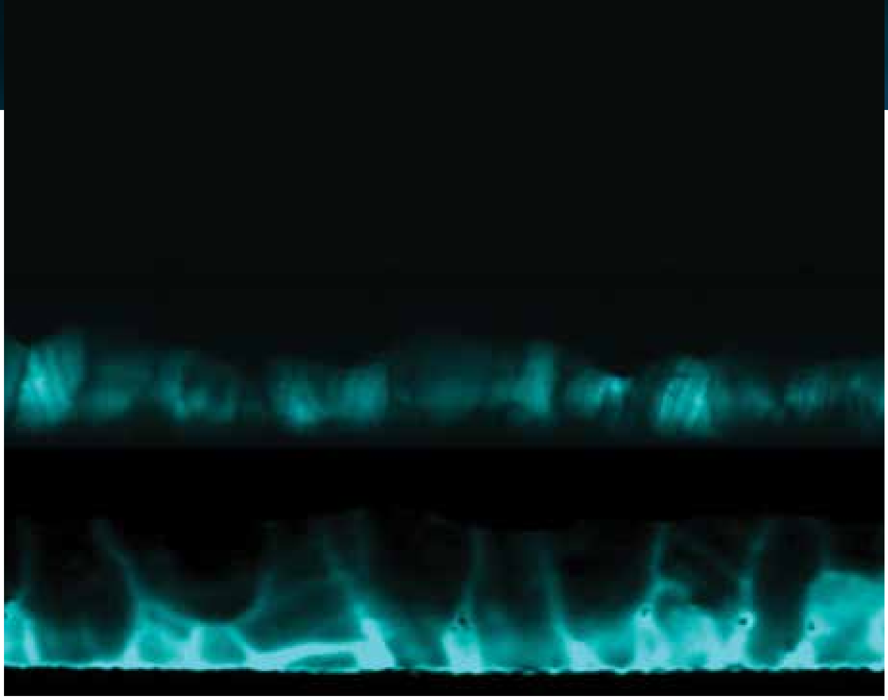
Research in the 1980s showed that treating CdTe thin films with cadmium-chloride significantly raise the cell's efficiency, but scientists have been unable to determine the underlying causes. The answer lies in investigating the material at an atomic level, according to Chen Li, a researcher from Oak Ridge National Laboratory who led this study. The team's findings were published in *Physical Review Letters*.

"We knew that chlorine was responsible for this magical effect, but we needed to find out where it went in the material's structure," said Li.

By comparing the solar cells before and after chlorine treatment, the researchers realized that "grain boundaries" were implicated in the enhanced performance. Grain boundaries are tiny defects that normally act as roadblocks to efficiency because they inhibit carrier collection, which greatly reduces solar cell power.

Using electron microscopy to study the thin films' structure and chemical composition before and after treatment, the researchers found that chlorine atoms replaced tellurium atoms within the grain

PROJECT TITLE	NERSC PI	LEAD INSTITUTION	NERSC RESOURCES USED	DOE PROGRAM OFFICE
Defect Physics of Photovoltaic Materials	Yanfa Yan	University of Toledo	Hopper	BES—Materials Science



◀ Cross-sectional electron beam-induced current maps show the difference in cadmium telluride solar cells before (top) and after (bottom) cadmium chloride treatment. The increased brightness after treatment indicates higher current collection at the grain boundaries. *Image: Chen Li, Oak Ridge National Laboratory*

boundaries. With a clearer picture of the material's atomic structure, the team was then able to construct computer models and calculate what was happening in the material.

"We calculated the electronic properties of the system containing the unit of the grain boundary, first without any chlorine to make sure our models were correct, and then with differing amounts of chlorine," said Yanfa Yan of the University of Toledo, who led the density functional theory calculations performed on NERSC's Hopper system.

His calculations revealed that this atomic substitution—chlorine for tellurium—creates local electric fields at the grain boundaries that boost the material's photovoltaic performance instead of damaging it.

"This grain boundary model has been proposed by others, but there wasn't any theoretical understanding or experimental evidence to support it from the atomic scale," said Yan. "The calculations we performed at NERSC were critical to this result."

The research team's findings, in addition to providing a long-awaited explanation, could be used to guide engineering of higher efficiency CdTe solar cells. According to Li, controlling the grain boundary structure is a new direction that could help raise the cell efficiencies closer to the theoretical maximum of 32 percent light-to-energy conversion. Currently, the record CdTe cell efficiency is only 20.4 percent.

"We think that if all the grain boundaries in a thin film material could be aligned in the same direction, it could improve cell efficiency even further," Li said.

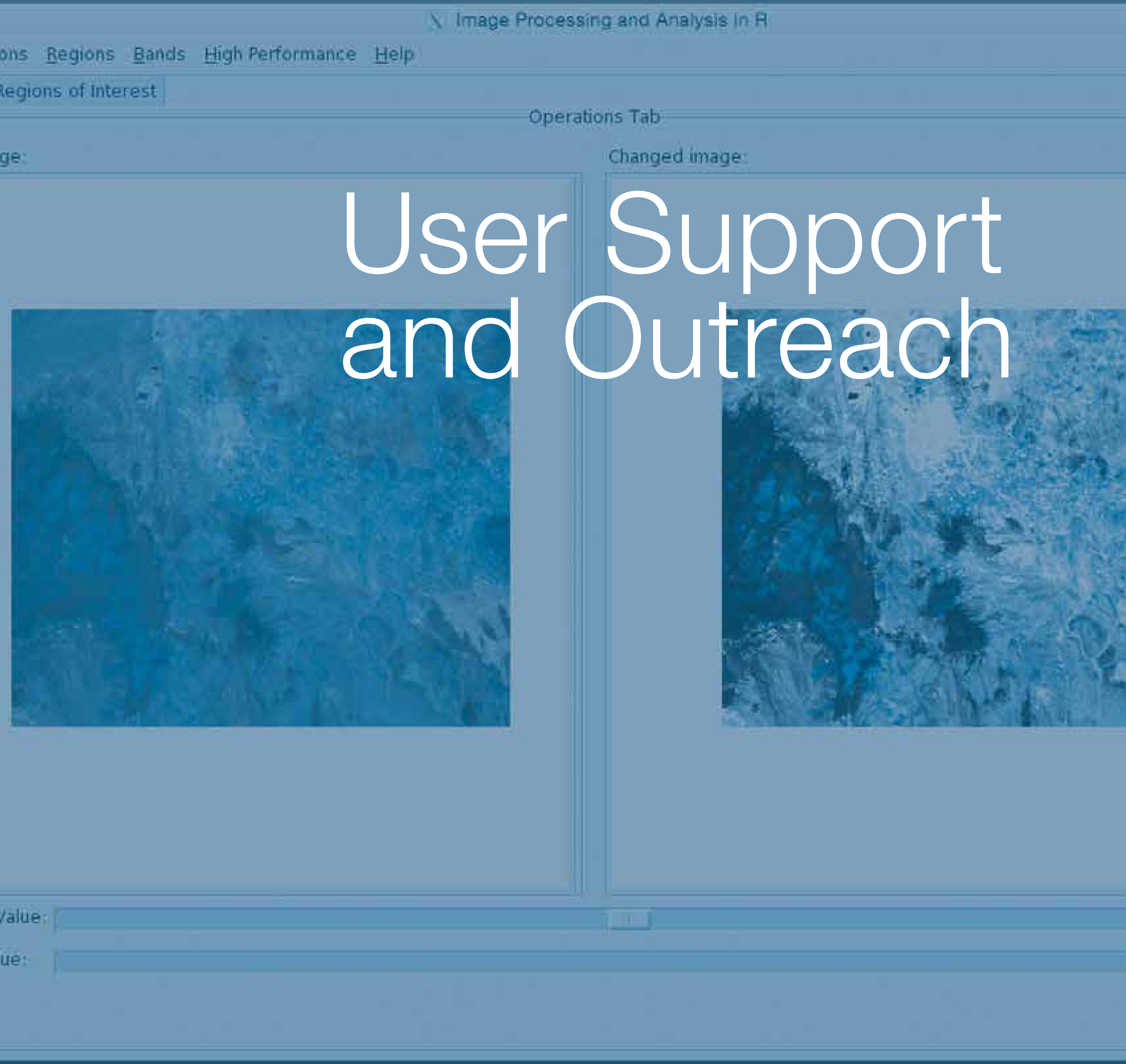
PUBLICATION

C. Li, Y. Wu, J. Poplawsky, T. J. Pennycook, N. Paudel, W. Yin, S. J. Haigh, M. P. Oxley, A. R. Lupini, M. Al-Jassim, S. J. Pennycook, Y. Yan, "Grain-Boundary-Enhanced Carrier Collection in CdTe Solar Cells," *Physical Review Letters* 112, 156103, April 16, 2014, doi: http://dx.doi.org/10.1103/PhysRevLett.112.156103

FULL STORY

http://www.nersc.gov/news-publications/news/science-news/2014/atomic-switcheroo-explains-origins-of-thin-film-solar-cell-mystery/

User Support and Outreach

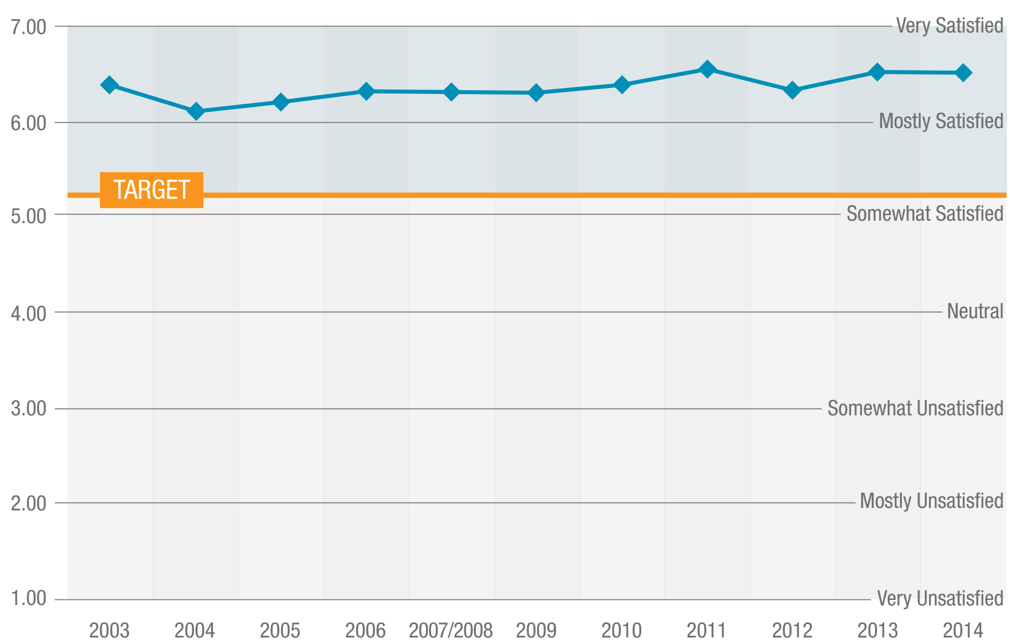


User Support Remains Key Priority in 2014

The first line of support for NERSC's user community is the Account Support team and User Services Group (USG). USG provides technical support, education and advocacy and develops and deploys new computational and data technologies. The group is also responsible for problem management and consulting, helping with user code optimization and debugging, strategic project support, web documentation and training, third-party applications, library support, running Office of Science requirements reviews and coordinating the NERSC Users Group.

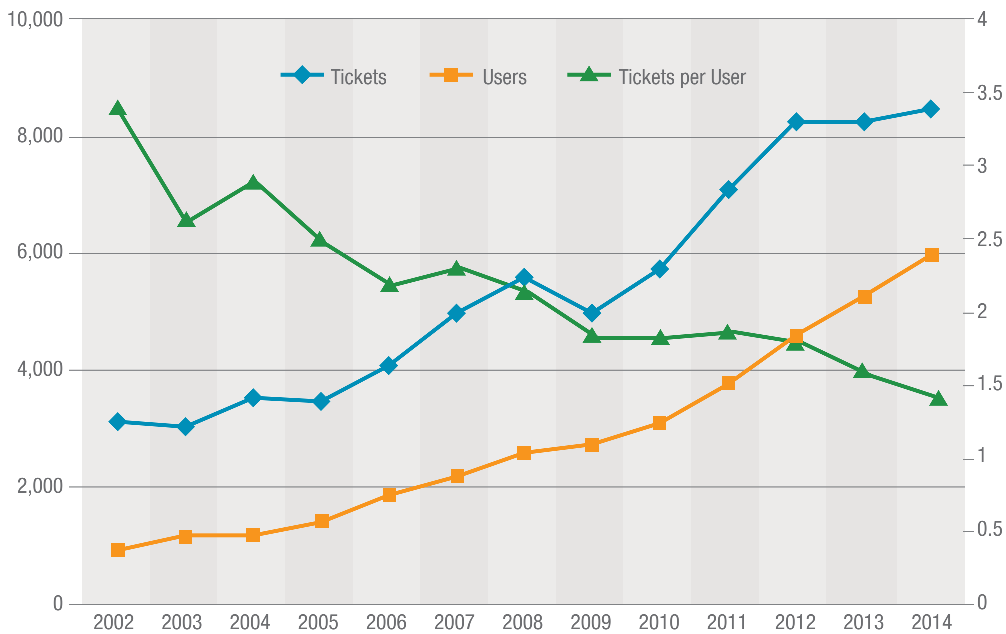
NERSC's consultants and account support staff are available to users via email, an online web interface and on the phone during business hours (8 a.m.–5 p.m. Pacific Time). Basic account support (password resets, resetting login failures) is available online or via the NERSC operations staff 24 x 7, 365 days a year. When users contact NERSC, they immediately are in contact with highly trained HPC specialists who can often solve issues directly or otherwise immediately route the request to the appropriate NERSC staff expert. NERSC's 11 consultants, 10 of whom have Ph.D.s, are all experts in HPC and various science domains. NERSC responds to all inquiries within four business hours and either solves the problem or communicates a work plan to the user within three business days for all reported incidents.

Overall User Satisfaction with NERSC



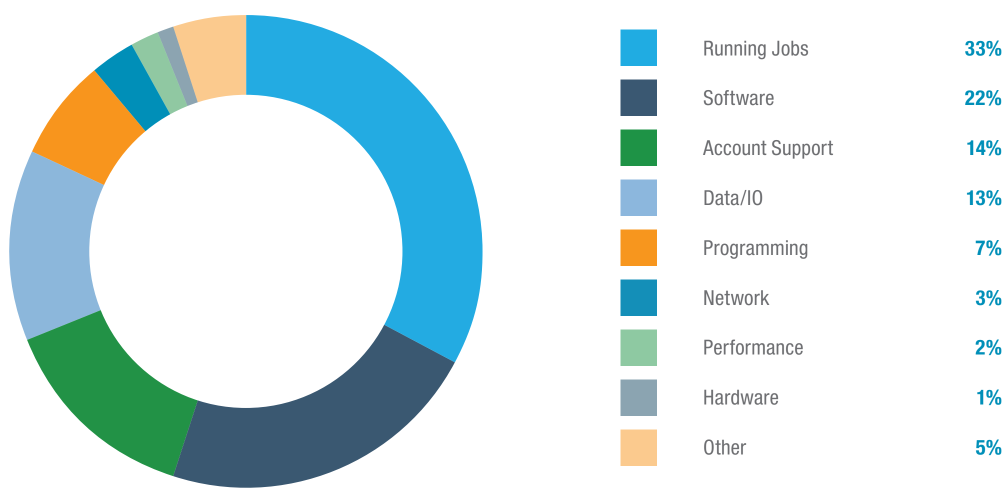
In 2014 NERSC supported about 6,000 scientific users from universities, national laboratories and industry, working on 849 projects with allocations of NERSC resources. The chart on the preceding page shows NERSC users' Overall Satisfaction scores since 2003. The 2014 score of 6.50 was equal to the second highest ever recorded in the 15 years the survey has been in its current form.

Number of NERSC Users and User Tickets Created per Year



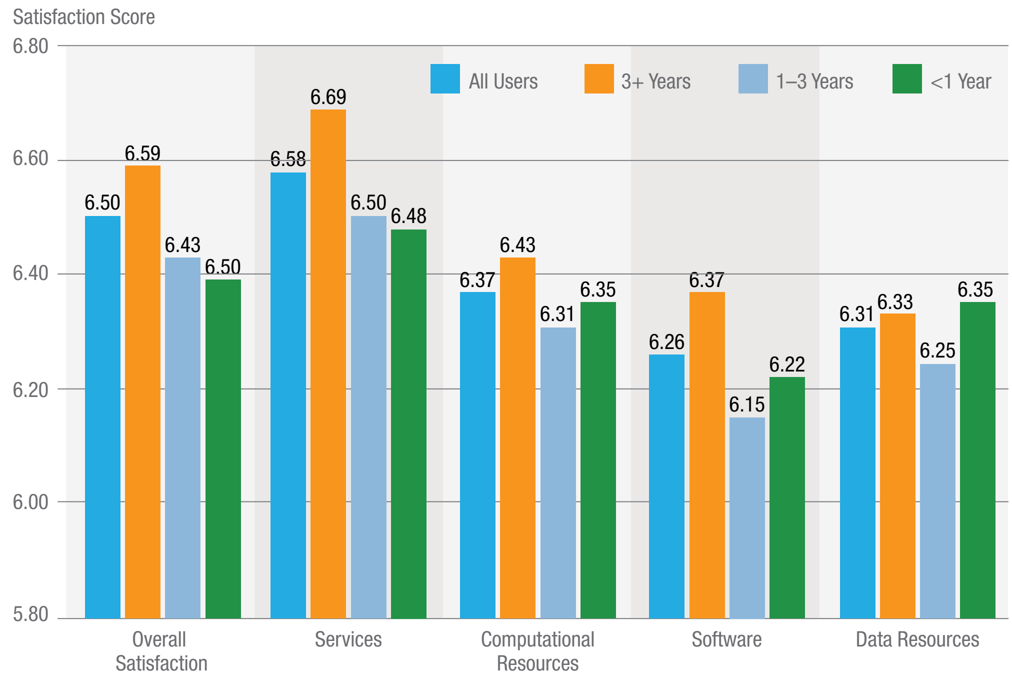
In calendar year 2014, NERSC users generated 8,430 consulting and account support trouble tickets from 2,205 different users. The breakdown of tickets by category was almost identical to that in 2013, as consultants continued to field many questions about running jobs and using software.

User Trouble Tickets by Category



Users who have computed with NERSC the longest expressed the highest satisfaction. NERSC does not collect demographic information that would help determine if this reflects a generational bias.

Satisfaction by User's Experience Using NERSC



User Support Services in 2014

With thousands of users and a limited number of staff, NERSC relies heavily on delivering assistance at scale with a modern web portal (my.nersc.gov), extensive online documentation (www.nersc.gov) and an active training program. This broad training and assistance program frees staff resources to allow NERSC consultants and other staff to work closely with key application developers.

MyNERSC

MyNERSC is a web-based portal that provides NERSC users with real-time information on their jobs, disk usage, allocations and queue wait times. It was originally developed 10 years ago, and in 2014 both the web and mobile versions were revamped to bring a wealth of new interactive content to the web. The new version of MyNERSC is designed to be a one-stop destination for interactive user content on the web. It seamlessly transforms for mobile or desktop use and combines interactive content that used to be spread across MyNERSC, NIM (NERSC's accounting interface), Service Now (NERSC's ticketing system), Mobile NERSC and nersc.gov.

New enhancements include an individualized dashboard for all users, containing repo and disk usage information, job and system status, completed job information and global chat. There is also an advanced queue-time predictor based on historical data, which for the first time can exclude time jobs spend on hold, allowing users to optimize their job scripts to minimize queue latency.

In addition, users can now monitor the backlog of NERSC systems over time and track their (and other) repos' usage over time for multiple years. They can also monitor file system and compute benchmark performance at NERSC using the same data NERSC uses to monitor performance on its HPC systems over time. MyNERSC also enables users to see a rich amount of detail regarding queued, running and completed jobs. This includes tracking the rank in the queue over time and a variety of performance data from Darshan, ALTD, Cray Proc Stats and LMT about completed jobs.

The latest version of MyNERSC also allows users to:

- Plot the rank of their job over time in the queue
- Hold, release and delete their jobs
- Download any of their files and upload new files, including to and from HPSS
- Monitor the NERSC file system performance over time
- Subscribe to the NERSC outages schedule via Google Calendar
- Create and manage their Service Now support tickets
- View a history of all NERSC outages and announcements
- Change their default repo, default shell and other NIM profile information
- See a list of top jobs running at the center and node breakdown by DOE office, science category and class
- Choose an interactive login node to use based on utilization and/or free memory.

User Training Events

NERSC held a number of training events for users in 2014, including three for new users, the second BerkeleyGW workshop and others targeted at helping users prepare for Cori: one on OpenMP and vectorization, another on using the Cray Reveal tool and an instructional Birds-of-a-Feather session at SC15 on Intel Xeon Phi optimization and tuning.

The annual NERSC Users Group (NUG) meeting, held February 3-6 at the center's Oakland headquarters, also provided training opportunities for users. In 2014 the NUG meeting served as the kickoff event for NERSC's 40th anniversary celebration, with a training day for new users, two days of science and technology presentations by NERSC and Berkeley Lab staff, a special "reminiscences" event that drew numerous former employees, presentation of the annual NERSC HPC Achievement Awards and the official dedication of NERSC's Edison supercomputer.

In June 2014, NERSC hosted the Joint Facilities User Forum on Data-Intensive Computing. Attended by more than 140 facility users and staff, this three-day event was organized by leaders from several DOE facilities, including NERSC, Argonne, Oak Ridge, Sandia, Lawrence Livermore and Los Alamos national laboratories. The workshop brought together facility users and HPC center staff to discuss the latest trends and techniques related to data management, analysis and visualization. In addition to training, the event facilitated discussion designed to inform the co-located DOE High Performance Computing Operational Review on data-intensive computing.

The High Performance Computing Operational Review was held to discuss how DOE HPC centers can best provide facilities and services to enable large-scale, data-driven scientific discovery at the DOE national laboratories. Breakout sessions focused around eight topics:

- System configuration
- Visualization/*in situ* analysis
- Data management policies
- Supporting data-producing facilities and instruments
- Infrastructure
- User training
- Workflows
- Data transfer.

The review found that to support the needs of the DOE data-producing facilities—whether they are HPC centers, experimental facilities or other collaborative projects like astronomical sky surveys—will require DOE HPC centers to change the way they have traditionally operated. The breakout groups identified a number of key challenges and issues. The full report is available at https://www.nersc.gov/assets/HPCOR/HPCOR-Data-2014.pdf.

◀ The second annual Berkeley GW workshop was hosted by NERSC in November 2014.

In November, the second annual BerkeleyGW user and developer workshop was held in downtown Oakland, near Berkeley Lab's Oakland Scientific Facility. The BerkeleyGW code—used to compute the electronic and optical properties of a variety of material systems—is being developed as part of a SciDAC collaboration between UC Berkeley and Berkeley Lab's Materials Science Division, Computational Research Division and NERSC.

As at the first workshop, attendee interest exceeded available space, but 50 researchers were able to attend the two-and-a-half-day training workshop, which covered practical aspects of DFT and GW calculations using NERSC HPC resources, including both methodology and performance aspects.

The workshop was well received; an anonymous post-workshop survey showed that all attendees found at least one session at the workshop useful, and the majority felt the hands-on sessions were "very useful."

Intel Xeon Phi Users Group Goes International

NERSC, in collaboration with the Texas Advanced Computing Center (TACC), led the expansion of the Intel Xeon Phi Users Group (IXPUG) into a national and international organization during 2014. IXPUG is an independent users group whose mission is to provide a forum for the free exchange of information that enhances the usability and efficiency of scientific and technical applications running on large HPC systems using the Intel Xeon Phi processor. IXPUG will serve as a resource to help NERSC staff and users prepare codes to run on Cori when the Xeon Phi-based system arrives in 2016.

IXPUG was previously hosted by TACC, which has a large system with attached Xeon Phi coprocessors. The new IXPUG held its first event at SC14 with a Birds-of-a-Feather session focused on Xeon Phi code optimization and tuning. The event drew about 150 participants. IXPUG is planning workshops at ISC and SC15, as well as a fall meeting in Berkeley, hosted by NERSC.

Hardware and Software Advances

In addition to supporting and facilitating these and other user training opportunities, in 2014 NERSC staff were instrumental in the development of a number of software and hardware advances designed to improve how users interact with the center, access and utilize NERSC resources, manage their data and jobs and increase productivity.

Batch Scheduler Enhancements

Batch systems are at the core of how NERSC fulfills its mission of providing HPC services to DOE Office of Science programs and increasing our user community's scientific productivity. We continually strive to improve, innovate and streamline operations in this area. We are early adopters of new releases and provide crucial feedback to vendors and the open software developers that serve not only us but the broader HPC community. We are always on the lookout for new job scheduling products, and we deploy, test and evaluate them at scale.

NERSC currently makes use of two batch scheduler/resource managers for the production systems: Torque/Moab and Univa GridEngine (UGE). NERSC is also actively evaluating alternative schedulers such as SLURM. During 2014 we made significant contributions in the following areas:

- **Torque-Moab HA configuration:** We changed the architecture of the batch system used on NERSC's two Cray systems into a high availability configuration.
- **Torque-Moab upgrades:** We upgraded the resource manager and scheduler on the three systems using Torque/Moab (Carver, Hopper and Edison). This resulted in a much more responsive batch system as well as improvements to large job startup and efficiency of scheduling high-throughput workloads.
- **Torque/ALPS integration:** We debugged issues with Torque/ALPS (Cray's Application Level Placement Scheduler) integration, discovering in the process some vendor bugs that were subsequently addressed.

- **Fair share scheduling trial deployment in Torque/Moab on Edison:** While exploring an option to use fair share scheduling we were able to accommodate special runs, meet NERSC's commitment to DARPA and provide computing cycles to early users in proportion to their NERSC allocations. Careful investigation led to a conclusion that the fair share option did not meet NERSC scheduling needs. While most of this work was performed in 2013, our experiences with it were written up and shared with the community in a Cray User Group paper in 2014.
- **SLURM test deployments and testing:** An alternative batch system software, SLURM, is currently under evaluation at NERSC. It has been deployed in test environments similar to the Carver, Hopper and Edison systems and tested under workload conditions that closely emulate production use. We have tested scalability, scheduler efficiency, the SLURM feature set and the effectiveness of Torque wrapper commands on both Cray systems and traditional Linux clusters.
- **Modifications of Globus batch plugins for UGE:** NERSC relies on UGE to manage computations for extreme data processing for the high energy and nuclear physics workloads on PDSF. Institutions all over the world participate in the LHC data analysis, such as the hunt for the Higgs boson. Globus software is used to coordinate that effort, and NERSC was instrumental in writing and debugging Globus plugins for UGE.
- **Integrating management of UGE wrappers/patches for JGI with configuration management (Cfengine3):** Joint Genome Institute workloads require customized modifications to the UGE batch management software. In the past those modifications had to be re-introduced after each UGE release install. NERSC staff modified automated system configuration to manage dependencies with Cfengine 3 and streamlined deployment of new releases, greatly improving staff productivity and preventing inconsistencies.

Automated Deployments with Jenkins

Some of NERSC's partner organizations, such as the Joint Genome Institute, are heavily engaged in software development of their HPC and data analysis codes, having frequent and varied release schedules for much of the software they produce. Production scientific software at NERSC is deployed to a high-performance, purpose-built file system tuned to deliver software to the computational platforms. For security and performance reasons, NERSC does not allow users to install software onto this file system or to manipulate the module files that provide access to the software therein. Therefore, to reduce the NERSC staff burden of installing user-developed software, and to enforce good practices and security of our resources, we implemented a Jenkins Continuous Integration service with custom plugins to automate the building and installation of user software.

The NERSC Jenkins system requires users to include build scripts in their SCM repository. Once this is done, builds and installations are performed automatically whenever the software SCM repository is updated. This allows users to keep the very latest software available at NERSC without having to "ask permission" and await consulting assistance for every update. The automated build system, in turn, manages the file system space, including purging old versions, ensuring that existing installations are kept around for all jobs currently using them and preventing failed builds from reaching the user community.

This system was introduced in late 2013 and updated in 2014, and it has now been used by over 20 software projects, performing thousands of on-demand builds and updates. Users have more control over the software they rely on, while embracing modern software engineering best practices and increasing efficiency of NERSC staff.

Transitioning from CFEngine to Puppet

Configuration management software is critical to the NERSC Server Team's daily tasks. Several NERSC groups have historically used CFEngine for configuration management; however, CFEngine can be unwieldy and does not offer features provided by competitors. Thus we have chosen a new, additional configuration management tool: Puppet, a flexible software product that can be adapted to a variety of use cases. Ample documentation, collaboration and support are provided by the vendor, PuppetLabs, and through the community.

Puppet has become critical in helping us to maintain a consistent infrastructure, which saves time and effort. Puppet is built on top of an existing programming language, ruby, which is easy to understand and program. Puppet also has an easy-to-read syntax, which greatly speeds up development time. In addition, translating CFEngine recipes to Puppet manifests has proven to be very straightforward.

Puppet saves the Server Team time and labor because it is easier to develop and deploy. A service can be defined and rolled out to hosts within hours. If the configuration of a service changes due to upgrades or to changes by another staff member, the Server Team can detect this "configuration drift" and bring the system back to compliance. By comparison, doing the same changes with CFEngine would often take several days, and CFEngine could not consistently detect configuration drift.

NERSC's Server Team currently manages 20 systems using Puppet, and we plan to eventually manage all of our systems using Puppet once NERSC transitions to the new Computational Research and Theory facility. We are also working with other groups at NERSC, such as the Storage, Operations Technology and Data Analytics Services groups, about using Puppet in their environments.

Sponsored Storage for Big Science

In 2013 and 2014 NERSC began seeing a surge of requests for large storage allocations on a file system called /project, which is used for scientific collaborations. The requests range from a few hundred terabytes to a petabyte and often come from high-profile science projects that have large amounts of experimental or simulation data. NERSC is unable to provide these large amounts of non-purged storage to all users and so created a buy-in model called "Sponsored Storage," a method of offering storage allocations through chargeback to recoup costs necessary to provide the allocation.

NERSC instituted the Sponsored Storage program in 2014, and as of early 2015 six projects are already taking advantage of it. The innovation to provide sponsored storage services was partly technical, but primarily procedural. The Lab's financial groups had not been exposed to this type of problem before, and NERSC worked closely with other departments at Berkeley Lab to be able to accept funding from different projects and buy into storage at NERSC for a given duration and support level. Not only has this effort helped the science teams accomplish their work, but it has saved the customized effort required in the past to deploy storage for different users.

Enhanced Monitoring with Automated Tests

With the pending move to CRT and in preparation for the Cori system, NERSC's Operations Technology Group has been recreating its workflows to improve its diagnostic capabilities with enhanced monitoring and the ability to detect problems more quickly to avoid potential system outages. These changes include monitoring areas that historically only alerted the system administrators, collaborating with vendors to improve the process of reporting issues and analyzing data from several sources to correlate multiple events. This effort has contributed to overall system availability and is part of the team's strategic plan of providing a higher level of triage and problem resolution to maximize resource availability to users.

In a recent example of how the new workflow contributed to system availability, a chiller that cools the Hopper system went down, causing Hopper's temperature to rise, an event that historically would have caused the system to power down in a short time to prevent overheating. In this case, the new workflow allowed the staff to restart the chiller, call in the appropriate Cray engineers, stop the queues and prevent the system from powering down. This incident was invisible to users, who were able to submit jobs and, once the queues were restarted, continue to run their jobs.

Data Analytics

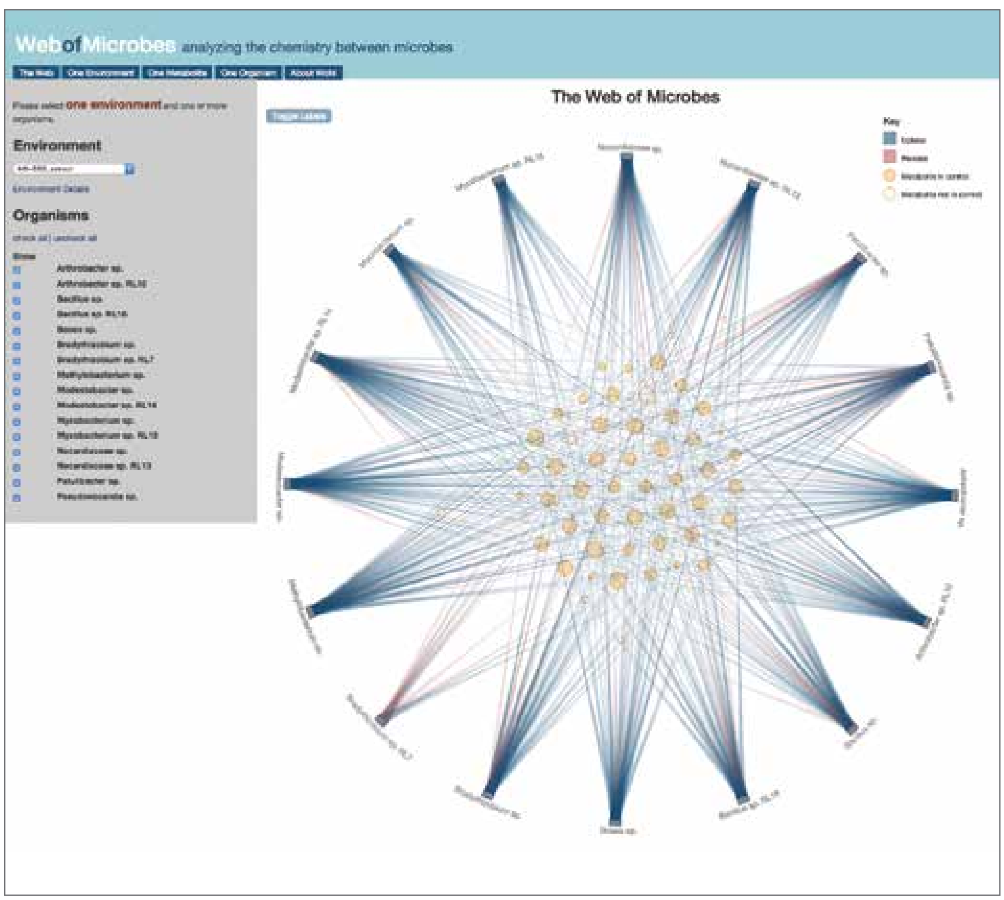
With the increasing emphasis on data-centric science across all disciplines, NERSC is expanding its capabilities to support the increasingly data-intensive computing needs of our users. In 2014 the center formed the Data and Analytics Services Group to help NERSC users address data and analytics challenges arising from the increasing size and complexity of data from simulations and experiments. Group members work directly with NERSC users to help them manage, store, move and analyze data.

Here are three examples of NERSC's data analytics successes in 2014.

Web of Microbes

Microbes can be finicky. In one environment, they may be happy to take up and metabolize a given compound, while in other environments they may prefer to ignore that same compound. The Northen Lab in the Berkeley Lab Life Sciences Division is analyzing the differences in how a variety of microbes interact with a long list of metabolites in different environments. The work has been slowed by the difficulty in analyzing this multidimensional data. Scientists have needed a way to visualize the interactions so that they can begin to answer questions such as which microbes might grow well together in a given environment or which conditions are key to controlling a particular

▼ Web of Microbes screen shot showing microbes (outer ring) growing on M9 minimal media plus BSC extract and metabolites on which they act (central dots).



microbe's growth. Through a collaboration between the Northen group and NERSC's Data and Analytics Services Group, NERSC has helped create a new online data visualization tool for analyzing these microbial interactions: the Web of Microbes (webofmicrobes.org).

The tool offers four different views of the data. The first is a network diagram that gives an overview of what amounts to a microbial food web for each environment. Color-coded connections between microbes and metabolites indicate uptake or release. Line densities immediately indicate key metabolites. Color densities indicate metabolites for which multiple microbes compete or microbes that are responsible for large fluxes of nutrients. One can click a metabolite to highlight all microbes that interact with it, or click a microbe to highlight all the compounds with which it interacts. The user can select a different environment from a drop-down menu and immediately see how the food web changes.

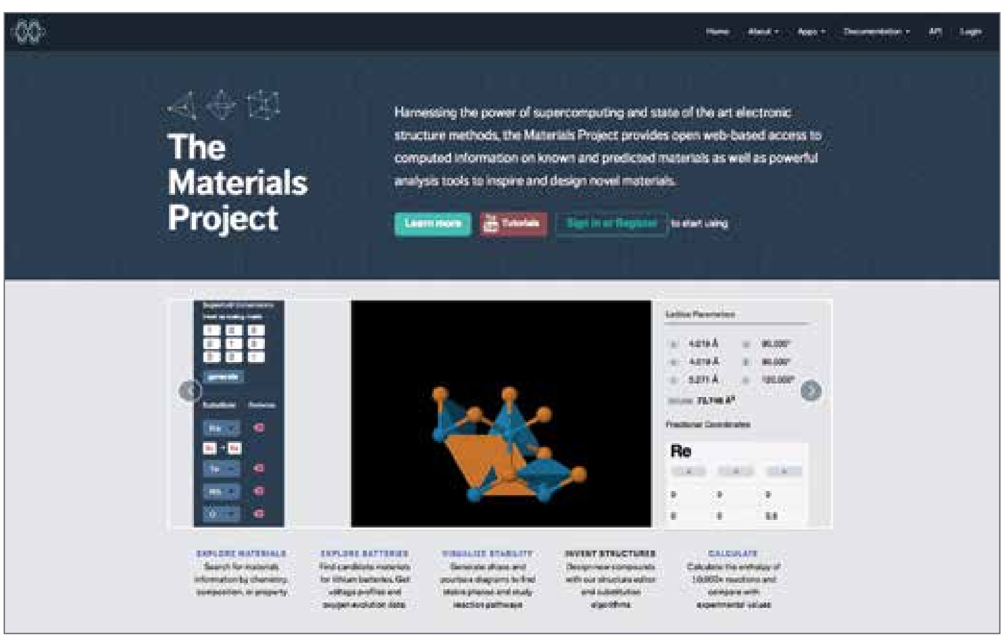
The three remaining views are heatmaps that allow the researcher to hold one dimension constant and analyze changes in the remaining two. In one, the user selects an environment, gets a checklist of metabolites observed in that environment and sees a customizable heatmap with columns representing microbes and rows representing metabolites. The intersection of each microbe with each metabolite is colored to indicate uptake, release or neither, and a level of confidence. One microbe is designated as the reference organism for analyzing which other microbes compete with it or are compatible with it. An overall affinity score is shown for each microbe, and symbols in the heatmap show the metabolites for which each microbe competes—or cooperates—with the reference organism. Clicking a different microbe makes that one the reference organism, immediately updating the competition/compatibility symbols and affinity scores. Another heatmap shows environments as columns and organisms as rows. In this view, the user selects a compound and can then analyze the actions on it by various microbes in different environments. Yet another view allows the user to focus on a single microbe and analyze its uptake and release of individual metabolites in multiple environments.

Operating on data stored in the NERSC Project file system and served by our science gateway hardware, the Web of Microbes has already stimulated new research in the Northen group. It has helped group members quickly identify problems in preliminary data and led them to develop new experiments to approach the analysis of chemical mobility between microbes and environments.

Materials Project

In 2014, NERSC continued its fruitful engagement with the Materials Project, an open, web-hosted database that allows scientists using NERSC supercomputers and quantum mechanical equations to design new materials atom by atom, before ever running an experiment.

NERSC serves as the computing and data engine for the project, providing the software and hardware infrastructure for the web



▼ The new Materials Project web gateway.

gateway and databases that serve up the Materials Project data. In addition to supporting calculations on the Hopper and Edison systems, the NERSC Division maintains cluster nodes purchased by and dedicated solely to the Materials Project.

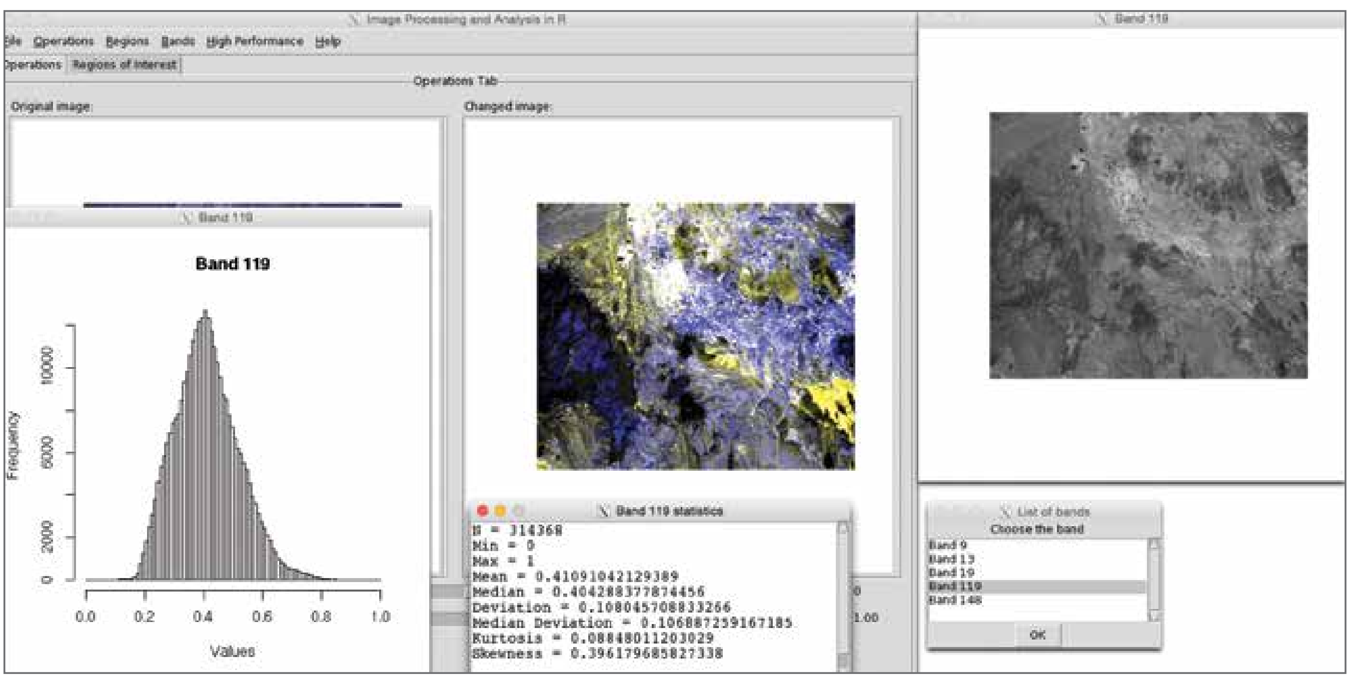
To provide a more powerful resource for users, in 2014 the Materials Project launched a completely revamped version of its web science gateway, with help from engineers in the NERSC Data and Analytics Services Group. The aim of the new web portal is to dynamically explore the data and launch new analyses.

2014 also saw the Joint Center for Energy Research Storage (JCESR) collaboration come on board. The JCESR data was deployed into the Materials Project site and is housed in a special “sandbox” of the MongoDB database.

RIPA: Image Processing and Analysis in R

Scientific imaging facilities produce massive amounts of images and the challenge is to make sense of them automatically. Image analysis tools based on the R statistical framework are essential for this task; however, the user community lacks stable solutions that deal with image processing at scale in R. Delving into R Analytics for Image Analysis (RIPA), NERSC staff—in collaboration with the Computational Research Division’s Scientific Visualization Group—developed and deployed RIPA, a new and unique R package in CRAN repository for image processing and analysis. RIPA allows the construction, test and validation of workflows with user-friendly interaction between R image analysis packages and R HPC packages. In addition to installing the package at NERSC, we have shared the package with the HPC community through CRAN for multiple platforms.

▼ Screen shot of NERSC’s new R Analytics for Image Analysis (RIPA) tool.

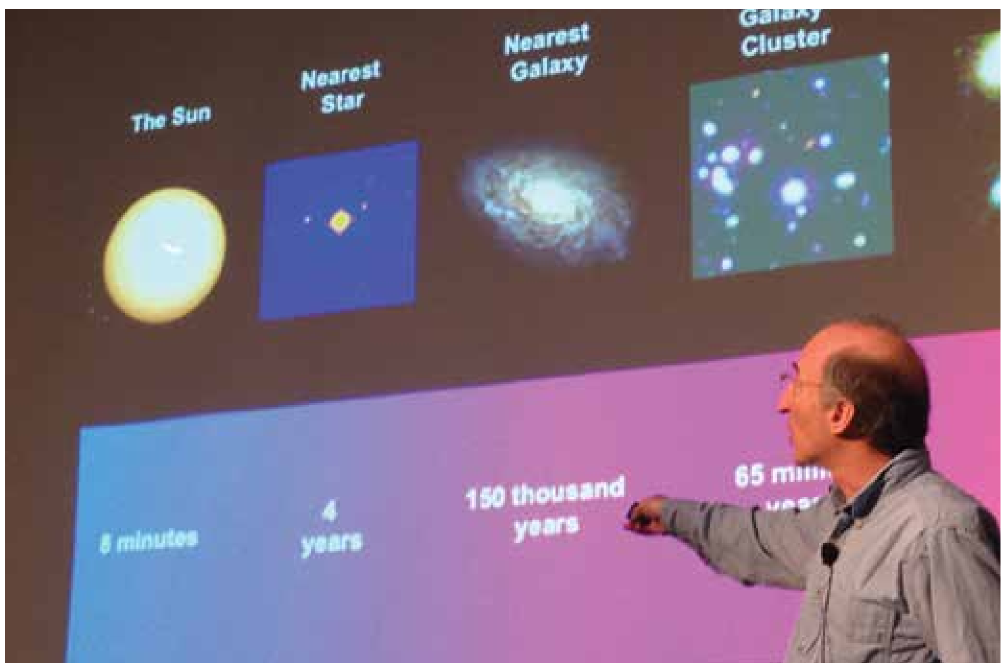


Center News

NERSC Honors 40th Anniversary With Nobel Lecture Series & More

Throughout 2014, NERSC celebrated yet another major milestone: its 40th anniversary.

Established in 1974 as the Controlled Thermonuclear Research Computer Center at Lawrence Livermore National Laboratory, NERSC was initially designed to provide scientific computing support for the fusion energy research community. It was renamed the National Magnetic Fusion Energy Computing Center in 1976, and in 1983 began providing a fraction of its computing cycles to other research areas. To reflect its increasingly broad scientific mission, in 1990 the center was christened the National Energy Research Supercomputer Center. NERSC moved to Berkeley Lab in 1996 and, while keeping the acronym, was renamed the National Energy Research Scientific Computing Center.



▲ Saul Perlmutter shared his thoughts on "Data, Computation and the Fate of the Universe" during the NERSC 40th Anniversary Nobel Lecture Series. *Photo: Margie Wylie, Lawrence Berkeley National Laboratory*

One of the highlights of NERSC's 40th anniversary yearlong commemoration was a series of invited talks featuring NERSC users whose research led to Nobel prizes. The four talks, hosted at Berkeley Lab in May and June, drew standing-room-only crowds and attracted hundreds of remote viewers as well.

The four Nobel talks were:

- **"Molecular Dynamics Simulations and the Mechanisms of Protein Complexes" by Prof. John Kuriyan of UC Berkeley.** Kuriyan is a member of a research team led by Martin Karplus, co-recipient of the 2013 Nobel Prize in Chemistry for his pioneering work in computational chemistry. Karplus has been a NERSC user since 1998.

▶ To commemorate NERSC's 40th anniversary, in February 2014 the California Legislature presented an Assembly Resolution to NERSC Director Sudip Dosanjh, shown here with Mark Chekal-Bain, district director for Assemblymember Nancy Skinner (AD-15). The certificate states: "NERSC is to be commended on the celebration of its fortieth anniversary for the vital role it has played in advancing scientific discovery at the U.S. Department of Energy's Office of Science." *Image: Roy Kaltschmidt, Lawrence Berkeley National Laboratory*

- **"The Brief History and Future Development of Earth System Models: Resolution and Complexity" by Warren Washington, a senior scientist at the National Center for Atmospheric Research in Boulder, Colo.** Washington, a longtime NERSC user who has been a leader in studying climate change for more than 50 years, shared the 2007 Nobel Peace Prize as a member of the Intergovernmental Panel on Climate Change.

- **"Mapping the Universe" by George Smoot, an astrophysicist at Lawrence Berkeley National Laboratory and professor at UC Berkeley.** Smoot was a co-recipient of the 2006 Nobel Prize in Physics for his research into cosmic microwave background (CMB), which produced an image of the infant universe. Today, NERSC is a center for the global CMB research community.

- **"Data, Computation and the Fate of the Universe" by Saul Perlmutter, a Berkeley Lab astrophysicist and professor at UC Berkeley.** Perlmutter was co-recipient of the 2011 Nobel Prize in Physics for his discovery that the universe is continuing to expand and that the rate of expansion is increasing. His initial findings in 1998 were supported by 10,000 simulations of supernovae generated on NERSC computers.

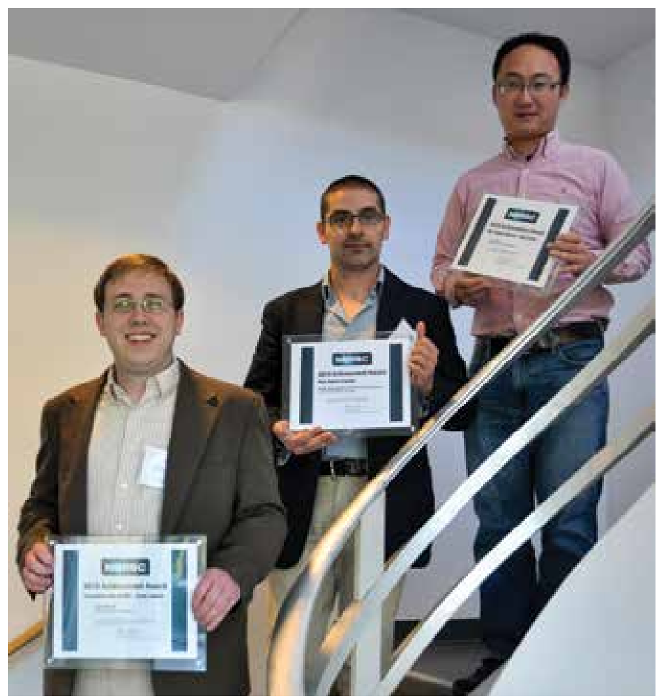
Other anniversary-related activities during 2014 included a dedicated "NERSC@40" web page, where we chronicled the center's history through a comprehensive, interactive timeline and numerous articles looking at how NERSC has worked closely with the various DOE program offices over the years; a "40 Years at the Forefront" brochure detailing NERSC's advances and achievements; an educational calendar featuring historical NERSC images and trivia; and a tribute to former NERSC staff during the annual NERSC Users Group meeting in February.

NERSC Wins HPCWire Editors’ Choice Award

At SC14 in New Orleans in November, Tom Tabor, publisher of HPCWire, presented NERSC with HPCWire’s 2014 Editors’ Choice Award for Best HPC Collaboration Between Government & Industry. The award recognized NERSC’s partnership with Intel and Cray in preparation for Cori, the Cray XC supercomputer slated to be deployed at NERSC in 2016. The HPCwire awards are recognized as one of the most prestigious HPC industry honors, demonstrating excellence and outstanding technological advancements achieved by the HPC community. *Image: HPCwire*

3rd Annual HPC Achievement Awards Honor Innovative Research

NERSC announced the winners of its third annual High Performance Computing Achievement Awards on February 24, 2015, during the annual NERSC User Group meeting at Berkeley Lab.



▲ **Third annual NERSC HPC Achievement Award winners:** (left to right) Taylor Barnes, Caltech; Carlo Benedetti, Berkeley Lab BELLA program; and Ken Chen, UC Santa Cruz. Not pictured: Craig Tull, SPOT Suite project. *Image: Margie Wylie, Lawrence Berkeley National Laboratory*

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.

NERSC 2015 Award for High Impact Scientific Achievement—Open

Berkeley Lab’s BELLA (Berkeley Lab Laser Accelerator) team was honored in this category for its work using NERSC resources to design and configure the world’s most powerful compact particle accelerator.

Traditional particle accelerators, like the Large Hadron Collider at CERN, which is 17 miles in circumference, speed up particles by modulating electric fields inside a metal cavity. It’s a technique that has a limit of about 100 mega-electron volts per meter before the metal breaks down. Laser-plasma accelerators such as BELLA take a completely different approach. A pulse of laser light is injected into a short, thin, straw-like tube that contains plasma. The laser creates a channel through the plasma and waves that trap free electrons and accelerates them to high energies.

In 2014, a team of Berkeley Lab researchers led by BELLA Director Wim Leemans used BELLA to accelerate subatomic particles to the highest energies ever recorded from a compact accelerator. The researchers sped up the particles—electrons in this case—inside a 9-centimeter long tube of plasma. The speed corresponded to an energy of 4.25 giga-electron volts. The acceleration over such a short distance corresponds to an energy gradient 1,000 times greater than traditional particle accelerators and marks a world record energy for laser-plasma accelerators.

NERSC 2015 Award for High Impact Scientific Achievement—Early Career

Ken Chen, a post-doctoral researcher at the University of California, Santa Cruz, was honored in this category for his work using NERSC to study the explosion of very massive stars in multiple dimensions.

One area of Chen's research has focused on pulsational-pair instability supernovae. One-dimensional calculations had shown the existence of a violent pulsational instability in stars with main sequence masses over 90 solar masses that ejects solar masses of material with supernova-like energies in a repeated pulses. The collisions of these shells can make extremely luminous and enduring supernovae. In 1D, however, the collisions pile up most of the ejecta in a dense, thin shell that is unphysical. Using supercomputing resources at NERSC, Chen has shown that the collision actually produces instabilities and large-scale mixing. The big density spikes previously seen in the 1D studies go away. The resulting remnant looks more like the Crab Nebula than a spherical bulls eye.

In 2014, Chen was lead author on a study that revealed that certain primordial stars—those between 55,000 and 56,000 times the mass of the Sun—may have died unusually. In death, these objects would have exploded as supernovae and burned completely, leaving no remnant black hole behind. Previously, it was thought that all stars over about 300 solar masses would collapse to black holes. This study also garnered much attention in the popular press, in large part because of the eye-catching visualizations Chen created using CASTRO—a multidimensional compressible astrophysics code developed at Berkeley Lab—and VisIt. One of his images is featured on the cover of the 2014 NERSC Annual Report.

NERSC 2015 Award for Innovative Use of HPC—Open

Berkeley Lab's SPOT Suite team received this award for their work in beginning to transform the way scientists run their experiments and analyze data collected from DOE light sources.

SPOT Suite is a set of tools for reducing, managing, analyzing and visualizing beamline data. It was developed through a collaboration between researchers from Berkeley Lab's Computational Research Division, NERSC, the ALS and the Molecular Foundry.

One of the first things the SPOT Suite collaboration did was set up data transfer nodes between NERSC and three beamlines at Berkeley Lab's Advanced Light Source (ALS), including the GISAXS/GIWAXS (grazing incidence small-angle X-ray scattering/grazing incidence wide-angle X-ray scattering) beamline. This tool is primarily used to characterize the assembly and shape of nanoscopic objects at surfaces or buried interfaces in thin films, including materials like organic photovoltaics, fuel cell membranes or batteries. In 2014, the ALS became the first and only facility in the world to fully automate GISAXS/GIWAXS measurements.

With SPOT Suite, researchers can run experiments at the ALS and other beamlines from anywhere in the world, provided they have Internet access. "This automated system represents a significant leap forward in terms of labor saving, ease of use and throughput," said Berkeley Lab researcher Alexander Hexemer, who manages the GISAXS/GIWAXS beamline.

NERSC 2015 Award for Innovative Use of HPC—Early Career

Taylor Barnes, California Institute of Technology (Caltech), was honored in this category for using NERSC resources to make outstanding methodological advances that advance our ability to harness large-scale computational resources for important chemical problems.

At Caltech, Barnes has led the development and implementation of powerful new computational methodologies to model complex systems, which will help to advance the understanding of fundamental chemical processes in catalysis, solar energy conversion and battery technologies.

Quantum embedding is a promising multi-scale strategy for vastly reducing the cost of accurate electronic structure theory calculations, allowing high-level quantum methods to be used to describe part of the system, while the remainder is treated with a low-level or continuum-level method. However, prior to Barnes' work in this area, density functional and wavefunction embedding approaches were only applicable to weakly interacting systems, which excluded essentially all condensed-phase and reactive chemical applications. By developing inversion-based and projection-based strategies to enable accurate embedding in the context of strongly interacting systems, Barnes' work has dramatically expanded the applicability of quantum embedding methodologies.

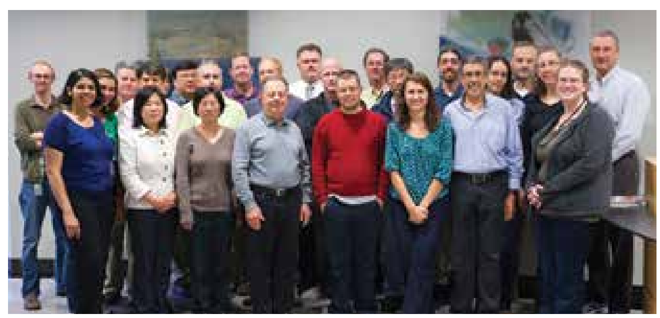
His research has so far led to five publications, with several more articles in preparation. It is also creating new research opportunities for the accurate description of catalytic and electronically non-adiabatic processes in complex systems, which are already being employed in a number of other theoretical chemistry groups across the U.S.

NERSC Hosts Application Readiness and Portability Meeting

More than 30 staff members from NERSC and the Oak Ridge and Argonne leadership computing facilities met at Berkeley Lab's Oakland Scientific Facility in September 2014 to coordinate strategies for application readiness in preparation for the next generation of supercomputers. The meeting also included representatives from Sandia, Los Alamos and Lawrence Livermore national laboratories.

During the meeting, each Office of Science computing facility presented an overview of its expected supercomputing architecture and programming model. An open discussion and brainstorming session on application portability and performance portability across architectures followed. Attendees also discussed how to provide users access to each other's systems so they can test performance portability. The attendees have committed to sharing best practices for application readiness and user training.

Similar to NERSC's Exascale Science Applications Program (NESAP), each of the leadership computing facilities will launch its own application readiness efforts aimed at preparing application codes for next generation systems. NESAP is designed to support NERSC's next-generation supercomputer, Cori, a Cray XC40 system slated to be deployed in 2016. NESAP features partnerships with more than three dozen application code teams and technical support from NERSC, Cray and Intel. Through NESAP, researchers are working with NERSC and vendor staff to optimize application codes for Cori's Knights Landing manycore architecture.



▲ **Attendees of NERSC's application readiness and portability meeting (in alphabetical order):** Katie Antypas, NERSC; Ashley Barker, OLCF; Matt Cordery, NERSC; Jack Deslippe, NERSC; Carter Edwards, SNL; Hal Finkel, ALCF; Fernanda Foertter, OLCF; Scott French, NERSC; Scott Futral, LLNL; Richard Gerber, NERSC; Helen He, NERSC; Oscar Hernandez, OLCF; Judy Hill, OLCF; Wayne Joubert, OLCF; Brian Lally, LANL; Ray Loy, ALCF; Bronson Messer, OLCF; Paul Messina, ALCF; James Osborn, ALCF; Scott Parker, ALCF; Katherine Riley, ALCF; Sreeranjani (Jini Ramprakash, ALCF; Adam Simpson, OLCF; Tjerk Straatsma, OLCF; Harvey Wasserman, NERSC; Timothy Williams, ALCF; Nick Wright, NERSC; Woo-Sun Yang, NERSC; Zhengji Zhao, NERSC.

NERSC: Who We Are

As one of the world's premier supercomputing centers, NERSC supports perhaps the largest and most diverse research community of any high-performance computing facility, providing large-scale, state-of-the-art computing for DOE'S unclassified research programs. More than 6,000 scientists worldwide use NERSC to conduct basic and applied research in energy production and conservation, climate change, environmental science, materials research, chemistry, fusion energy, astrophysics and other areas related to the mission of the DOE Office of Science.

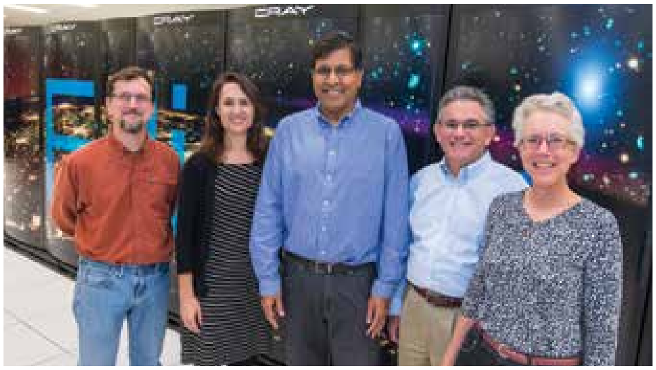
But NERSC is much more than just a collection of computers, servers, routers and software tools. One of its most valuable attributes is its staff, a talented group of computer scientists, mathematicians, engineers and support personnel. More than 50 percent of NERSC staff hold advanced degrees in a scientific or technical field. And collaboration—aka "team science," a concept pioneered by Berkeley Lab founder Ernest O. Lawrence in 1931—is a cornerstone of NERSC's philosophy, both internally and through its engagements with the broader science community.

Due to its proximity to Silicon Valley, NERSC often finds itself competing with the booming tech industry when recruiting new employees. But that's what makes those who choose to work at NERSC instead even more special: they're fundamentally committed to facilitating discovery on some of the nation's most pressing scientific challenges.

"The exciting thing about NERSC, and what we are all really passionate about, is enabling science that benefits society," said Sudip Dosanjh, director of NERSC. "I think a lot of us are really here for public service. There are lots of very high-paying computer industry jobs in the Bay Area, but really what drives us at NERSC is facilitating a broad range of scientific breakthroughs that benefit us all."

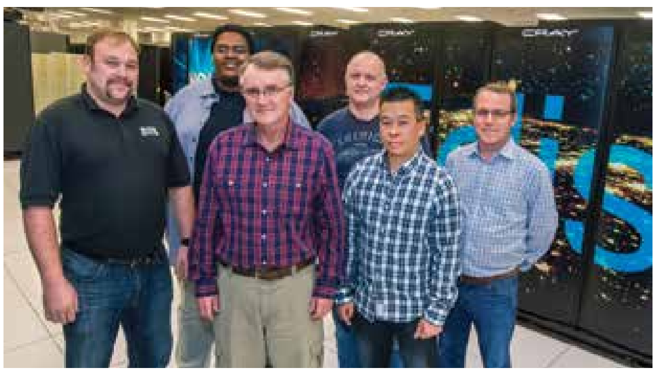
NERSC is divided into nine groups designed to serve the varied needs of its users and maintain the highest quality supercomputing infrastructure possible. There is also a dedicated Cray support team at NERSC.

- **Advanced Technologies:** The Advanced Technologies Group focuses on understanding HPC architecture trends as they relate to the requirements of current and emerging NERSC user applications. The goal is to make hardware design and configuration choices to best serve the needs of our users. The group's efforts include workload characterization, benchmarking, performance analysis and modeling.



Center Leadership: (l-r) John Shalf, Chief Technology Officer; Katie Antypas, Services Department Head; Sudip Dosanjh, NERSC Division Director; Jeff Broughton, Deputy for Operations and Systems Department Head; Francesca Verdier, Allocations Manager

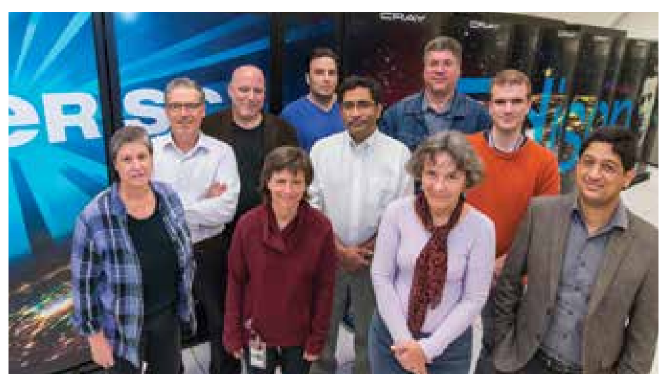
Center Administration/Facilities/Budget: (l-r) Kenley Ngai, Kerri Peyovich, Zaida McCunney, Victor Siqueiros, David Tooker, Norma Early, Jeffrey Grounds, Seleste Rodriguez



Cray Group: (l-r) Kris Howard, Robert Johnson, Randy Palmer, Terence Brewer, Shawn Le, Mark Green. Not pictured: Stephen Luzmoor

Data & Analytics Services: (l-r) Shreyas Cholia, Joaquin Correa, Oliver Ruebel, Annette Greiner, Prabhat (Group Lead), Dani Ushizima, Michael Urashka, Ulugbek Baymuradov, Yushu Yao. Not pictured: Burlen Loring, Peter Nugent, R.K. Owen, Jeff Porter

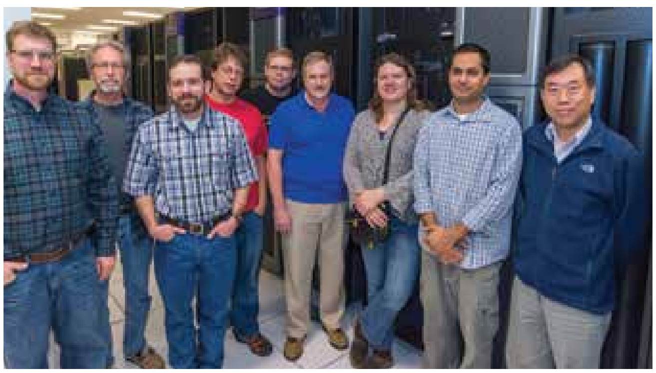
Advanced Technologies Group: (l-r) Nicholas Wright (Group Lead), Brian Austin, Chris Daley, Matthew Cordery



Computational Systems Group: (l-r) Tina Butler, David Paul, James Botts, Tina Declerck, Scott Burrow, Jay Srinivasan (Group Lead), Cary Whitney, Iwona Sakrejda, Douglas Jacobsen, Bhupender Thakur

Networking, Security and Servers: (l-r) Matt Dunford, Aaron Garrett, Craig Lant, Brent Draney (Group Lead), Damian Hazen, Jason Lee. Not pictured: Scott Campbell, Stefan Lasiewski, Karen Schafer, Brian Yumae

Operations Technology Group: (l-r) Tom Davis, Robert Neylan, Wilma Snider, Isaac Ovadia, Cary Whitney, Del Black, Elizabeth Bautista (Group Lead), Andrew Weaver, Alex Ubungen, Basil Lalli, Michelle Phung, Ray Spence, Yulok Lam. Not pictured: Tony Quan



Storage Systems Group: (l-r) Jason Hick (Group Lead), Wayne Hurlbert, Nicholas Balthaser, Patrick Hajek, William Baird, Greg Butler, Allison Andrews, Ravi Cheema, Rei Chi Lee. Not pictured: Fred Loebl

User Services Group: (l-r) Zhengji Zhao, Alice Koniges, David Turner, Helen He, Rebecca Hartman-Baker, Mark Heer, Woo-Sun Yang, Clayton Bagwell, Scott French, Richard Gerber (Group Lead), Harvey Wasserman, Jack Deslippe

Strategic Partnerships & Technology Integration: David Skinner (l), Shane Canon (r)

- **Computational Systems:** The Computational Systems Group deploys and maintains extreme-scale computational systems to support HPC and data-intensive science at NERSC. The group is responsible for providing the best possible technology on reliable, stable computing platforms for all users at NERSC in support of the DOE Office of Science mission. The group also helps support complex user workflows that use a diverse set of resources.
- **Data & Analytics Services:** In 2014 NERSC formed the Data and Analytics Services Group to help NERSC's users address data and analytics challenges arising from the increasing size and complexity of data from simulations and experiments. Group members work directly with NERSC users to help them manage, store, move and analyze data.
- **Networking, Security and Servers:** The Networking, Security and Servers team is responsible for designing and implementing a network that can provide high-performance transport of all scientific data in a seamless manner. The group also safeguards physical and intellectual assets while providing an open and secure computational environment and supporting the core center infrastructure, including databases, web servers and scientific portals.

- **Operations Technology:** The Operations Technology Group ensures the accessibility, reliability, security and connectivity of NERSC and ESnet by architecting and providing a central location for problem reporting, diagnosis, data collection, escalation and resolution to maximize the scientific productivity of users and ensure operational excellence of the facility.
- **Storage Systems:** The Storage Systems Group's mission is to develop and maintain facility-wide data systems and services to support the entire range of science facilitated by NERSC. The group is organized into three teams:
  - The Archival Storage Team focuses on development and deployment of a High Performance Storage System with stewardship of over 40 years of science data.
  - The Facility-wide File System Team deploys and manages the NERSC Global File Systems to provide high bandwidth storage to all computational systems.
  - The Enterprise Database Team supports large science databases on enterprise solutions.
- **Strategic Partnerships:** NERSC's Strategic Partnerships mission is to extend NERSC capabilities toward private sector and facility-to-facility collaborations. Partnerships seek overlapping interests in technological innovation and discovery science between NERSC and research enterprises that extend beyond the single PI model. We focus on collaborative efforts that develop novel operational models for NERSC in serving new science communities and industry market sectors.
- **Technology Integration:** The Technology Integration Group develops and demonstrates crosscutting, innovative solutions to enable NERSC to support more science and make users more productive.
- **User Services:** The User Services Group'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 science teams with code optimization and debugging; strategic project support; web documentation and training; and third-party applications and library support.

Personnel Changes in 2014

Prabhat

Prabhat was named Group Lead of the new Data and Analytics Services (DAS) Group at NERSC.

The DAS group helps NERSC's users address data and analytics challenges arising from the increasing size and complexity of data from simulations and experiments. As the DAS Group Lead, Prabhat is playing a key role in developing and implementing NERSC's data strategic initiative: to increase the productivity, usability and impact of data intensive science by providing comprehensive data systems and services to store, analyze, manage and share data. He was previously a member of the Scientific Visualization group and the NERSC Analytics team.

Prabhat has published numerous papers on extreme scientific data analysis and is the PI on the MANTISSA project funded by the Advanced Scientific Computing Research (ASCR) division of the DOE Office of Science. He is also leading the DOE/ASCR funded ExaHDF5 project, which is tasked with scaling HDF5 on current petascale and future exascale platforms, and the development of pattern detection and spatial statistics tools for the Calibrated and Systematic Characterization, Attribution and Detection of Extremes (CASCADE) Scientific Focus Area.

Prabhat received a master's degree in computer science from Brown University (2001) and a bachelor's degree in computer science and engineering from IIT-Delhi (1999). He is currently pursuing a Ph.D. in the Earth and Planetary Sciences Department at UC Berkeley.

David Skinner

Kjiersten Fagnan

In early 2014, NERSC created a new position—Strategic Partnerships Lead—to identify new science communities that can benefit from NERSC resources. David Skinner, former head of NERSC's Outreach Software and Programming Group (OSP), was selected to fill the role.

In his new post, Skinner is serving as a liaison between NERSC and projects that span many Berkeley Lab divisions. He has also spearheaded a new industrial partnerships initiative and is looking for opportunities to collaborate with ESnet and Berkeley Lab's Computational Research Division.

Skinner holds a Ph.D. in theoretical chemistry from the University of California, Berkeley. His research focused on quantum and semi-classical approaches to chemical reaction dynamics and kinetics. He began working at NERSC/Berkeley Lab in 1999 as an HPC engineer and spent the last eight years leading the OSP group.

Kjiersten Fagnan was named the NERSC/JGI Engagement Lead for Joint Genome Institute (JGI) computational and data analysis efforts carried out at NERSC.

Fagnan, who holds a Ph.D. in Applied Mathematics from the University of Washington, started at NERSC as a petascale postdoc in 2010 and in 2012 joined the team to help biologists at JGI exploit advanced computing systems hosted at NERSC.

In her new role, Fagnan is coordinating NERSC activities carried out in support of JGI's mission and leads NERSC's JGI user support team, which provides expert technical computational science support and advice to JGI staff and users. In addition, Fagnan is the NERSC co-chair with Shane Canon, Group Lead for NERSC's Technology Integration Group, on the JGI/NERSC Coordinating Committee and is leading a number of high performance computing and analysis initiatives at the JGI.

The DOE established the JGI in 1997 to support DNA sequencing, informatics and technology development pioneered at DOE genome centers. NERSC and the JGI are collaborating to develop solutions to the genomic community's growing need for advanced computation, data storage and analysis. JGI researchers use NERSC systems to assemble and perform comparative analysis for genomes from plant, fungal, microbial and metagenome communities.

New Hires in 2014

Doug Doerfler is an HPC Performance Engineer in NERSC's Advanced Technology Group, where he is leading the benchmarking effort for the NERSC 9 procurement. He will also conduct performance analysis and benchmarking of current and future high-performance computing architectures at NERSC. Doerfler came to NERSC after more than 29 years at Sandia National Laboratories, where he most recently served as Chief Architect and Deputy Program Manager for the NNSA's Cielo and Trinity supercomputers, both joint efforts with Los Alamos National Lab.

Scott French is an HPC Consultant in NERSC's User Services Group, where he assists users with code debugging and performance optimization, provides training in new technologies and troubleshoots user-environmental issues, among other things. He is also participating in NERSC's Application Readiness Program. He became a NERSC user in 2009 while working on his Ph.D. at UC Berkeley. Through a CITRIS-sponsored Designated Emphasis in Computational Science and Engineering Program at UC Berkeley, French completed computational science coursework as part of his Ph.D. and the computational aspects of his research were featured prominently in his dissertation. The latter gave rise to collaborations with Berkeley Lab computational researchers.

Rebecca Hartman-Baker joined NERSC's User Services Group as an HPC Consultant. In addition to helping scientists use the facility's machines more efficiently, Hartman-Baker is also helping NERSC develop a robust user training program and procure its next supercomputer. Before joining NERSC, Hartman-Baker performed very similar roles at the Pawsey Supercomputing Center in Perth, Western Australia, where she helped build a new petascale computing facility.

Kerri Peyovich is the new administrative assistant at NERSC, assisting staff with general office support, including travel arrangements, procurement, conferences and requisitions. Before coming to NERSC, Peyovich was an events manager at a private club in Berkeley for 10 years.

Jonathon Rood is a computer science postdoctoral fellow in NERSC's Advanced Technologies Group, where he is researching sequence alignment algorithms used in bioinformatics in an attempt to increase their performance using many-core computer architectures. Before coming to NERSC, Rood was an associate research mathematician at Tech-X Corporation, where he worked to increase the performance of NASA software using GPUs. The software is being used to analyze data from an instrument on the SAGE III mission, which provides accurate, long-term measurements of ozone, aerosols, water vapors and other key parameters of Earth's atmosphere.

Cory Snavely joined NERSC as a member of the Server Team, where he is managing systems and services in the periphery of NERSC's high-performance computing environment, including databases, web sites, storage, authentication and directory services and other core infrastructure. Throughout his career, Snavely has been responsible for architecting, building and maintaining similar environments, and for adapting them as needs have evolved. Before coming to NERSC, he managed Library IT Core Services at the University of Michigan and built large digital library systems.

Wilma Snider is an HPC technician in the Operations Technology Group, the team that ensures accessibility, reliability and connectivity for NERSC and ESnet. Before coming to the Lab, she spent five years at the SF Symphony doing PC support and then completing her AA in Information Systems from Peralta Unified College District. Snider was an intern at Berkeley Lab prior to transitioning to her current position.

Ray Spence started his second tour of duty at NERSC in 2014 as a computer systems engineer in the Operations Technology Group. From 2006-2008, Spence worked in NERSC's Networking Servers and Security group. In his new role, he is introducing new monitoring and alerting services within the group and expanding these capabilities within NERSC. Throughout his career Spence has supported several types of network services in various environments, including Fortune500 and dot-com companies, Lawrence Livermore National Laboratory and UC Berkeley.

Bhupender Thakur is a computer systems engineer in NERSC's Computational Systems Group, where he focuses on keeping resources provided to the Joint Genome Institute up and running and in good health. A native of India, Thakur came to the United States to pursue a Ph.D. in physics at the University of Delaware. In 2010, Thakur took a job at Louisiana State University's HPC center, where he helped researchers use the university's HPC resources, and later served in a systems support role similar to the position he currently holds at NERSC.

Allocations of NERSC Director's Reserve of Computer Time

The NERSC director receives 10 percent of the total allocation of computer time, which amounted to 300 million hours in 2014. In 2014 NERSC decided to phase out the NERSC Initiative for Scientific Exploration (NISE) program and allocated only one-third of its Director's Reserve time to this program. There were two reasons behind this decision: first, the NISE program tended to augment the investigators' DOE allocated projects rather than bring in new scientific explorations as intended; and second, the NERSC director wished to reserve allocation time for projects of strategic importance to DOE, Berkeley Lab and NERSC. NESAP, the new NERSC Exascale Science Applications Program, was one of the strategic efforts that received Director's Reserve allocations in 2014. Since the NERSC Director's awards were not able to use their 200 million hours, NERSC used about 131 million hours from this reserve to augment high-impact projects.

Director's Reserve Allocations

INVESTIGATOR	NERSC REPO	HOURS ALLOCATED	DOE OFFICE	PROJECT TITLE
Katie Antypas, Lawrence Berkeley National Laboratory	m1759	130,000	ASCR Applied Mathematical Sciences	NERSC Application Readiness for Future Architectures
Greg Bell, Lawrence Berkeley National Laboratory	ciena	25,000	ASCR Computer Sciences	Parallel FEC Verification and Scaling Prototype
Julian Borrill, Lawrence Berkeley National Laboratory	mcplanck	23,000,000	HEP Cosmic Frontier	Massive Monte Carlo Simulations for the Analysis of Planck Satellite Data
Kristofer Bouchard, Lawrence Berkeley National Laboratory	m2043	75,000	BER Biosciences	CRCNS
Valerie Daggett, University of Washington	m559	7,750,000	BER Biosciences	Molecular Dynameomics
Thomas Hamill, National Oceanic & Atmospheric Administration	refcst	150,000	BER Climate Research	A Multi-Decadal Reforecast Data Set to Improve Weather Forecasts for Renewable Energy Applications
Tianzhen Hong, Lawrence Berkeley National Laboratory	m2019	4,000,000	BER Climate Research	Database of Energy Efficiency Performance
Yanfei Jiang, Smithsonian Astrophysical Observatory	m1972	3,800,000	NP Astrophysics	Radiation Hydrodynamic Simulations of Massive Star Envelope
Balint Joo, Jefferson Lab	m2176	100,000	NP Lattice Gauge Theory	Modernizing the Chroma LQCD Code for Calculations of Hadron Spectroscopy in Nuclear Physics on Cori
Yinglong Miao, Howard Hughes Medical Institute	m1925	1,025,000	BER Biosciences	Enhanced Conformational Sampling of G-protein Coupled Receptors
John Michalakes, National Oceanic & Atmospheric Administration	m2190	1,000,000	BER Climate Research	Next Generation Global Prediction System (NGGPS) Benchmarking
Thomas Oppe, DoD High Performance Computing Modernization Program	m2035	650,000	BES Chemistry	Application Benchmarking and SSP Development

INVESTIGATOR	NERSC REPO	HOURS ALLOCATED	DOE OFFICE	PROJECT TITLE
Saul Perlmutter, Lawrence Berkeley National Laboratory	bids	25,000	ASCR Computer Sciences	Berkeley Institute for Data Sciences through HPC
Kristin Persson, Lawrence Berkeley National Laboratory	jcesr	21,750,000	BES Materials Science	Joint Center for Energy Storage Research
Richard Plevin, University of California Berkeley	m1496	175,000	BER Climate Research	Monte Carlo Simulation of Greenhouse Gas Emissions from Biofuels-induced Land Use Change
George Shields, Bucknell University	m1226	175,000	BER Climate Research	Thermodynamics of Secondary Aerosol Formation: The Role of Binary and Ternary Nucleation
David Skinner, Lawrence Berkeley National Laboratory	ngbi	4,800,000	BER Biosciences	Next Generation Bioimaging Institute
Katherine Yelick, Lawrence Berkeley National Laboratory	mp309	1,075,000	ASCR Computer Sciences	Class Account for UCB CS267/CS194 "Applications of Parallel Computing"

2014 NISE Program Awards

INVESTIGATOR	NERSC REPO	HOURS ALLOCATED	DOE OFFICE	PROJECT TITLE
Gerbrand Ceder, Massachusetts Institute of Technology	rope	5,250,000	BES Materials Science	Testing the Fundamental Limits on the Number of Inorganic Compounds in Nature
Christopher Cramer, University of Minnesota Twin Cities	m1939	2,300,000	ChemSci	Comparative Theoretical Study of the Optical Response and Ultrafast Excited-state Dynamics of Dye-sensitized ZnO and TiO2 Solar Cell Model Systems
Paul Kent, Oak Ridge National Laboratory	m1994	9,300,000	BES Materials Science	Accelerated Discovery and Design of Complex Materials
Jitendra Kuma, Oak Ridge National Laboratory	m1952	1,750,000	BER Climate Research	Multiscale Modeling of Dynamic Arctic Landscapes in a Changing Climate
Gregory Newman, Lawrence Berkeley National Laboratory	m1941	55,250,000	BES Geosciences	Imaging of Seismic Wavefields in Complex Heterogeneous Media at Unprecedented Details and Scales
Travis O'Brien, Lawrence Berkeley National Laboratory	m1949	3,500,000	BER Climate Research	Calibrated and Systematic Characterization, Attribution and Detection of Extremes
Michael Pindzola, Auburn University	m1943	6,250,000	NP Astrophysics	Quantal Four Body Breakup Problem
David Prendergast, Lawrence Berkeley National Laboratory	m1947	10,000,000	BES Materials Science	Exploring Dynamics at Interfaces Relevant to Mg-ion Electrochemistry from First-Principles
Martin Voros, University of Chicago	m1948	5,600,000	BES Materials Science	Large Scale Calculations on Nanostructured Heterogeneous Interfaces

Transfers to DOE Projects

INVESTIGATOR	NERSC REPO	HOURS ALLOCATED	DOE OFFICE	PROJECT TITLE
Julian Borill, Lawrence Berkeley National Laboratory	planck	15,000,000	HEP Cosmic Frontier	Cosmic Microwave Background Data Analysis for the Planck Satellite Mission
Choong-Seock Chang, Princeton Plasma Physics Laboratory	m499	13,250,000	Fusion Energy Sciences	Center for Edge Physics Simulation: SciDAC-3Center
James Chelikowsky, University of Texas at Austin	m175	2,500,000	BES Materials Science	Scalable Computational Tools for Discovery and Design: Excited State Phenomena in Energy Materials
Teresa Head-Gordon, Lawrence Berkeley National Laboratory	m1876	1,400,000	BES Chemistry	Advanced Potential Energy Surfaces for Condensed Phase Simulation
Steve Jardin, Princeton Plasma Physics Laboratory	mp288	6,100,000	Fusion Energy Sciences	3D Extended MHD Simulation of Fusion Plasmas
Paul Kent, Oak Ridge National Laboratory	m526	3,500,000	BES Materials Science	Computational Resources for the Nanomaterials Theory Institute at the Center for Nanophase Materials Sciences
Michael Klein, Temple University	m1592	17,000,000	BES Chemistry	Advanced Modeling of Ions in Solutions, on Surfaces and in Biological Environments
Duy Le, University of Central Florida	m1996	4,750,000	BES Chemistry	Computational Design of MoS2-based Functional Materials
Zhihon Lin, University of California Irvine	m808	7,300,000	Fusion Energy Sciences	SciDAC GSEP: Gyrokinetic Simulation of Energetic Particle Turbulence and Transport
Jeffrey Neaton, Lawrence Berkeley National Laboratory	m387	16,500,000	BES Materials Science	Theory of Nanostructured Materials
Kristin Persson, Lawrence Berkeley National Laboratory	matgen	14,000,000	BES Materials Science	The Materials Genome
Berend Smit, University of California, Berkeley	m979	5,795,000	BES Chemistry	Computational Characterization of Porous Materials
Brian Wirth, Oak Ridge National Laboratory	m1709	24,000,000	Fusion Energy Sciences	PSI SciDAC: Bridging from the Surface to the Micron Frontier

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. They also participate in professional organizations and conferences and contribute to journals and proceedings. The NERSC user community benefits from these activities as they are applied to systems, software and services at NERSC and throughout the HPC community.

For example, in 2014 NERSC staff participated in several DOE-sponsored Science Data Pilot Projects to demonstrate new approaches for collecting, moving, sharing and analyzing massive scientific data sets.

One of the pilot projects demonstrated the ability to use a central scientific computing facility (NERSC) to serve data from multiple experimental facilities. Data from experiments at the Advanced Light Source (ALS), the Advanced Photon Source at Argonne National Laboratory (ANL), the Linac Coherent Light Source at SLAC and the National Synchrotron Light Source at Brookhaven National Laboratory were moved to NERSC via ESnet. Once the data arrived at NERSC, it was automatically or semi-automatically analyzed and visualized within the SPOT Suite software toolkit developed at Berkeley Lab and hosted by NERSC. While the projects at the ALS and SLAC processed data in real time from ongoing experiments, the other two used real data in simulated real time as a demonstration of automating their real-time workflows using SPOT Suite.

A second pilot project, which involved experiments with organic photovoltaic materials at the ALS, demonstrated the seamless integration of multiple, complementary DOE Office of Science user facilities—NERSC, the ALS, the Oak Ridge Leadership Computing Facility (OLCF) and ESnet—into a virtual "superfacility" offering fundamentally greater capability. Enabled by ESnet's connectivity between ALS, NERSC and OLCF and using specialized software (including ANL's Globus Online), the project showcased how researchers in organic photovoltaics could measure scattering patterns for their samples at the ALS, see real-time feedback on all the samples through SPOT Suite and see near-real-time analysis of their samples running at the largest scale on the Titan supercomputer at OLCF.

To see a comprehensive list of publications and presentations by NERSC staff in 2014 go to http://1.usa.gov/1f3uFIv.

Appendix A:

NERSC Users Group Executive Committee

Office of Advanced Scientific Computing Research

Mark Adams, *Lawrence Berkeley National Laboratory*
Anubhav Jain, *Lawrence Berkeley National Laboratory (Vice-Chair)*
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

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

Office of High Energy Physics

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

Office of Nuclear Physics

Balint Joo, *Jefferson Lab*
Tomasz Plewa, *Florida State University*
Nicolas Schunck, *Lawrence Livermore National Laboratory*

Members at Large

Cameron Geddes, *Lawrence Berkeley National Laboratory*
Konstantin Kemenov, *CRAFT Tech*
James Amundson, *Fermilab*

Appendix B:

Office of Advanced Scientific Computing Research

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

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

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

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

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

ASCR Program

Steve Binkley, *Associate Director, ASCR*

Michael Martin, *Fellow*

Julie Stambaugh, *Financial Management Specialist*

Lori Jernigan, *Program Support Specialist*

Facilities Division

Barbara Helland, *Director*

Vince Dattoria, *Computer Scientist, ESnet Program Manager*

Betsy Riley, *Computer Scientist, ALCF Program Manager*

Carolyn Lauzon, *Physical Scientist, ALCC Program Manager*

Dave Goodwin, *Physical Scientist, NERSC Program Manager*

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

Sally McPherson, *Program Assistant*

Research Division

William Herrod, *Director*

Teresa Beachley, *Program Assistant*

Randall Laviolette, *Physical Scientist, SciDAC Application Partnerships*

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

Karen Pao, *Mathematician, Base/Multiscale Mathematics*

Ceren Susut, *Physical Scientist, SC Program SAPs*

Robinson Pino, *Computer Scientist, HPC Sci and Apps Program Manager*

Rich Carlson, *Computer Scientist, Collaboratories/ Middleware*

Steven Lee, *Physical Scientist, SciDAC Institutes*

Lucy Nowell, *Computer Scientist, Data and Visualization*

Sonia Sachs, *Computer Scientist, Extreme Scale*

Angie Thevenot, *Program Assistant*

Appendix C:

Acronyms and Abbreviations

ACM
Association for Computing Machinery

ACS
American Chemical Society

ALCC
ASCR Leadership Computing Challenge

ALCF
Argonne Leadership Computing Facility

ALS
Advanced Light Source, Lawrence Berkeley National Laboratory

ANL
Argonne National Laboratory

API
Application Programming Interface

APS
American Physical Society

ASCII
American Standard Code for Information Interchange

ASCR
Office of Advanced Scientific Computing Research

BELLA
Berkeley Lab Laser Accelerator

BER
Office of Biological and Environmental Research

BES
Office of Basic Energy Sciences

BNL
Brookhaven National Laboratory

C3
Computational Cosmology Center, Lawrence Berkeley National Laboratory

CAL
Computer Architecture Laboratory

CARB
California Air Resources Board

CCM
Cluster Compatibility Mode

CdTe
Cadmium Telluride

CERN
European Organization for Nuclear Research

CESM
Community Earth Systems Model

CFD
Computational Fluid Dynamics

CLE
Cray Linux Environment

CMB
Cosmic Microwave Background

CO_2
Carbon dioxide

CPU
Central Processing Unit

CRD
Computational Research Division, Lawrence Berkeley National Laboratory

CRT
Computational Research and Theory Facility, Lawrence Berkeley National Laboratory

CSE
Computational Science and Engineering

DARPA
Defense Advanced Research Projects Agency

DESI
Dark Energy Spectroscopic Instrument

DFT
Density Functional Theory

DME
Dimethyl Ether

DNS
Direct Numerical Simulation

DOE
U.S. Department of Energy

DOI
Digital Object Identifier

DSL
Dynamic Shared Library

DTN
Data Transfer Node

DVS
Data Virtualization Service

EFRC
DOE Energy Frontier Research Center

EMSL
Environmental Molecular Science Laboratory, Pacific Northwest National Laboratory

EPSI
SciDAC Center for Edge Physics Simulations

ERT
Empirical Roofline Toolkit

ESD
Earth Sciences Division, Lawrence Berkeley National Laboratory

ESnet
Energy Sciences Network

eV
Electron Volts

FDM
Finite Difference Method

FES
Office of Fusion Energy Sciences

FLOPS
Floating Point Operations

FTP
File Transfer Protocol

GB
Gigabytes

Gbps
Gigabits Per Second

GPU
Graphics Processing Unit

HEP
Office of High Energy Physics

HPC
High Performance Computing

HPSS
High Performance Storage System

HTML
Hypertext Markup Language

HTTP
Hypertext Transfer Protoco

IEEE
Institute of Electrical and Electronics Engineers

InN
Indium Nitride

IPCC
Intel Parallel Computing Center; Intergovernmental Panel on Climate Change

iPTF
intermediate Palomar Transient Factory

ITER
An international fusion energy experiment in southern France

ITG
Ion Temperature Gradient

IXPUG
Intel Xeon Phi Users Group

JCESR
Joint Center for Energy Research Storage

JET
Joint European Torus

JGI
Joint Genome Institute

LED
Light-emitting Diode

LHCII
Light-harvesting Complex II

LANL
Los Alamos National Laboratory

LLNL
Lawrence Livermore National Laboratory

MIT
Massachusetts Institute of Technology

MOF
Metal Oxide Framework

MPP
Massively Parallel Processing

MSI
Mass Spectrometry Imaging

NCAR
National Center for Atmospheric Research

NESAP
NERSC Exascale Scientific Application Program

NEXAFS
Near Edge X-ray Absorption Fine Structure

NGF
NERSC Global Filesystem

NIH
National Institutes of Health

NIM
NERSC Information Management

NOAA
National Oceanic and Atmospheric Administration

NP
Office of Nuclear Physics

NPLQCD
Nuclear Physics with Lattice QCD

NSF
National Science Foundation

NUG
NERSC Users Group

NVRAM
Non-volatile Random Access Memory

NWB
Neurodata Without Borders

OLCF
Oak Ridge Leadership Computing Facility

OpenMSI
Open Mass Spectrometry Imaging

OSF
Oakland Scientific Facility

PDACS
Portal for Data Analysis services for Cosmological Simulations

PDSF
Parallel Distributed Systems Facility, NERSC

PI
Principal Investigator

PIC
Particle-In-Cell Simulations

PB
Petabytes

PSII
Photosystem II

PNNL
Pacific Northwest National Laboratory

PPPL
Princeton Plasma Physics Laboratory

QCD
Quantum Chromodynamics

RIPA
R Analytics for Image Analysis

SC
DOE Office of Science

SciDAC
Scientific Discovery Through Advanced Computing

SDN
Software-defined Networking

SIAM
Society for Industrial and Applied Mathematics

TACC
Texas Advanced Computing Center

TB
Terabytes

URL
Universal Resource Locator

XMAS
X-ray Microdiffraction Analysis Software

For more information about NERSC, contact:

Jon Bashor

NERSC Communications
1 Cyclotron Road
Berkeley Lab, MS 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, Linda Vu and Margie Wylie
Berkeley Lab Computing Sciences Communications Group

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

Julie Chao, Kate Greene
Berkeley Lab Communications Department

Riana Geldenhuys
University of Cape Town

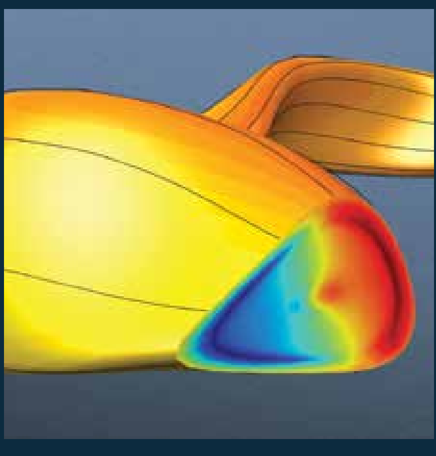
Design

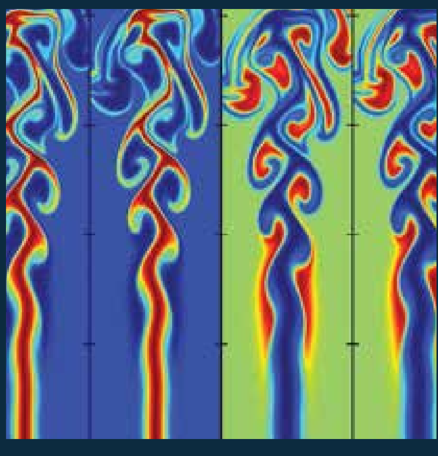
Design, layout, illustrations, photography and printing coordination provided by the Creative Services Office of Berkeley Lab's Public Affairs Department.

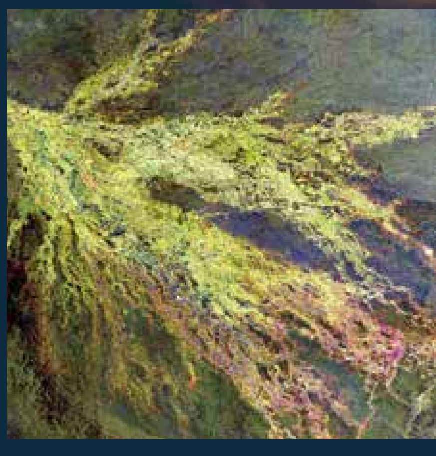
PAC 15-NERSC-1696

BERKELEY LAB

U.S. DEPARTMENT OF ENERGY
Office of Science







NERSC 2014 Annual Report