

National Energy Research Scientific Computing Center

2017 ANNUAL REPORT



U.S. DEPARTMENT OF
ENERGY

Office of Science



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

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

Image Credits: Marilyn Chung, Berkeley Lab

National Energy Research Scientific Computing Center

2017 ANNUAL REPORT

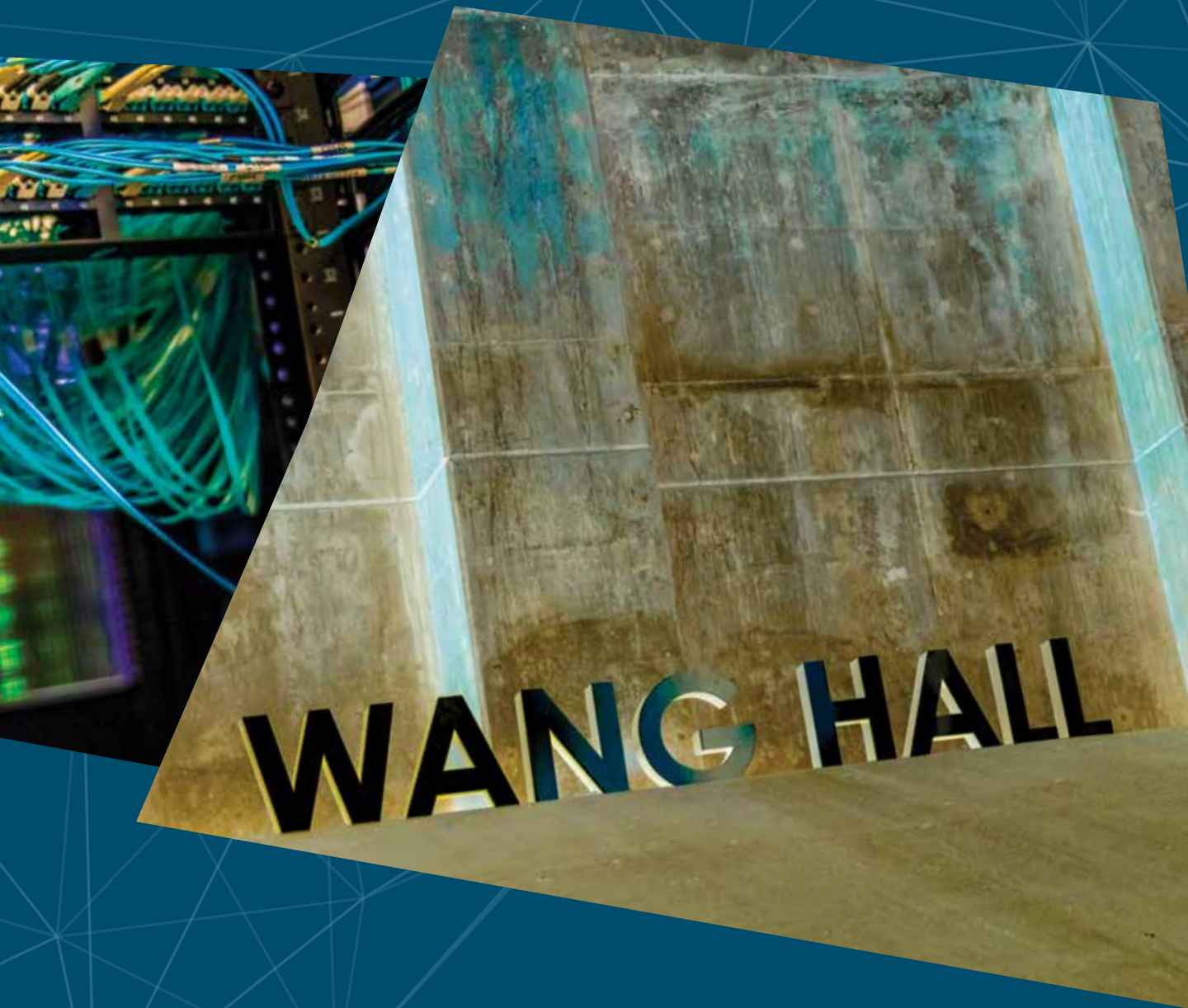
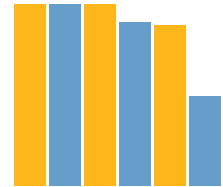


Table of Contents

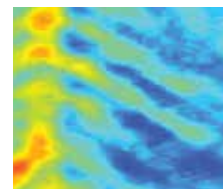
DIRECTOR’S NOTE	6
2017 USAGE DEMOGRAPHICS	10
SCIENCE HIGHLIGHTS	14
Assessing Regional Earthquake Risk in the Age of Exascale (ECP)	15
HPC for Better Computer Memory Technologies (BES)	16
TOAST: Preparing for Next-Generation CMB Experiments (HEP)	17
The Origin of Heavy Elements (HEP)	18
A Record Quantum Circuit Simulation (ASCR)	19
How Catalysts Convert CO ₂ into a Useful Fuel (BES)	20
‘Hindcasting’ Extreme Weather Events (BER)	21
More Dust, Better Air Quality (BER)	22
Shedding Light on Mysterious Plasma Flows (FES)	23
First Complex Calculations of Sigma Particle (NP)	24
Predicting Defect-free Solar Cell Nanomaterials (BES)	25
INNOVATIONS	26
Supporting Gordon Bell Runs	27
ExaHDF5 Team Partners with the Exascale Computing Project	28
NERSC Develops Storage Roadmap for 2020 – and Beyond	28
New Website Facilitates Performance Portability	29
Boosting Data Center Energy Efficiencies	30
Streamlining NERSC Systems’ Data-Collection Capabilities	31
MODS: Business Intelligence for Data Services	31
CI/CD Simplifies Supercomputer Support Environment	32
NERSC Users Take Spin for a Spin	33
Automating Software Management Using Spack	34
TOKIO Framework Simplifies I/O Performance Analysis	34
Hierarchical Roofline Model Development Expands	34
Minimizing the Impact of Cache Thrashing	35
Toward Better Queue Logistics and Interactivity	36
Simplifying Job Script Creation for Users	36
Taking Power Monitoring to the Next Level	37



Page 11

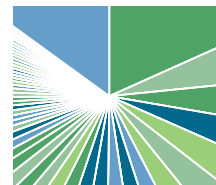


Page 15



Page 23

ENGAGEMENT AND OUTREACH	38
An Exascale Computing Project Update	40
NESAP Continues to Thrive	41
Projecting the NERSC Workload Performance on GPUs in 2020	42
NERSC Data Advisory Committee Looks to the Future	42
NERSC/Intel/Universities Launch ‘Big Data Center’	43
Scaling Up to Petaflop Data Analytics	44
NERSC Takes Superfacility Concept to New Heights	44
NERSC/JGI Collaboration: The FICUS Project	46
OpenMP Standards Update Enhances Runtimes	47
NERSC Hosts 2017 SLURM Meeting	47
Other NERSC Trainings, Workshops and Hackathons	48
CENTER NEWS	50
Shyh Wang Hall Achieves LEED Certification	51
California Senator Tours NERSC	52
NERSC Honors Four Users with 2017 HPC Achievement Awards	52
High-Impact Science at Scale	53
NERSC Shares HPCwire Award for HPC4Mfg Project	54
Berkeley Lab-led Collaborations Earn HPC Innovation Awards	54
NERSC Hosts Jefferson Lab QCD Researcher	55
All-female Student Cluster Team Competes at ISC for Second Time	55
NERSC Prepares for Potential Emergencies	56
Multi-Factor Authentication Added for NERSC Users	57
Personnel Transitions	57
New Hires	58
New NESAP Postdocs	58
PUBLICATIONS	59
DIRECTOR’S RESERVE ALLOCATIONS	60
APPENDIX A: NERSC USERS GROUP EXECUTIVE COMMITTEE	62
APPENDIX B: OFFICE OF ADVANCED SCIENTIFIC COMPUTING RESEARCH	63
APPENDIX C: ACRONYMS AND ABBREVIATIONS	64



Page 42



Page 46



Page 52



Director's Note

The National Energy Research Scientific Computing Center (NERSC) is the mission high performance computing facility for the Department of Energy's Office of Science (DOE SC). NERSC's goal is to accelerate scientific discovery at the DOE SC through high performance modeling, simulation, and data analysis.

In 2017, NERSC supported 7,405 users from universities, national laboratories and industry, providing high-end computing, storage and expertise to the broad science community. NERSC delivered over 7 billion hours to scientists, exceeded all of its operational metrics and maintained exceptionally high user satisfaction ratings.

The NERSC workload represents the wide variety of research performed by users, including simulations that run at the largest scales on the center's world-class supercomputers. In 2017, 40 percent of users reported that their projects were either analyzing experimental and observational data, combining simulations with data analysis or creating tools and algorithms to improve data analysis. SC experiments and facilities that combine HPC simulation and modeling with data analysis from instruments such as light sources, particle accelerators, telescopes, sensors and microscopes are influencing the way NERSC configures systems and supports users.

PREPARING NERSC USERS FOR EXASCALE

The demand from NERSC's SC user community drives NERSC to be forward thinking in key areas, particularly related to preparing our user community for exascale-era

technologies and meeting their surging data needs. Here are a few key projects:

- The NERSC Exascale Science Applications Program (NESAP) reached a culmination point in 2017 with the successful transition of much of the NERSC workload to NERSC's Cori supercomputer. Applications that were part of the NESAP program saw speed-ups of 3x on average. NERSC added data analysis focused projects to the NESAP program in 2017, seeing significant scaling gains of Python-based workflows.
- NERSC, Intel and Cray established a Big Data Center designed to address SC's leading data-HPC intensive science problems at scale on NERSC's Cori system. The collaborators are targeting HPC capabilities to enable workflows that require analysis of over 100TB datasets on 100,000 cores or more. As part of this effort, NERSC is making new inroads in deep learning as well.
- NERSC developed a storage strategy to address the overwhelming data storage challenges the DOE SC user community is expected to face over the next decade and beyond. NERSC developed and released a report, "NERSC



Storage 2020,” presenting a detailed roadmap and storage vision. As input to the report, NERSC used feedback from the DOE Exascale Requirement Reviews.

NERSC convened a Data Advisory Committee — comprising representatives from industry, academia and other national laboratories — to evaluate its data strategy. In a series of meetings held in the Fall of 2017, NERSC staff gave presentations to the committee on key topics related to this strategy and the superfacility initiative, including implementation strategies for systems, software and user engagement.

Transitioning the broad SC workload to more advanced architectures on the path to exascale continues to be a primary focus area for NERSC. NERSC’s strategy is to partner with approximately 20 application teams through NESAP and then transfer lessons learned to the broader SC user community. During the year the NESAP team held four application performance “dungeon sessions” at Intel campuses; three KNL performance hack-a-thons at NERSC; and eight general application performance and performance tool trainings at NERSC. NERSC has been sharing the lessons learned from NESAP to the broader NERSC community through extensive

training, publications and online documentation and case studies. In 2017 NERSC hired a number of new NESAP post-doctoral fellows as others graduated from the program and found career positions in computational science.

PARTNERSHIPS ADVANCE ECP PROGRAM

NERSC continues to partner with ASCR HPC and networking facilities to deliver on ASCR and DOE’s Exascale Computing Project (ECP). NERSC engaged with the ECP projects in a variety of ways throughout the year:

- NERSC collaborated with ECP and the other ASCR facilities to present training on topics such as Python, OpenMP and Kokkos.
- NERSC is collaborating with the ECP application assessment area to bring Roofline to other platforms.
- A significant number of ECP application teams are collaborating with NERSC staff to ready their codes for exascale through the NESAP program.
- NERSC provided 125 million additional NERSC Hours to the ALCC program specifically to support ECP application development and software technology projects.

- NERSC staff were active participants in the PathForward projects with Intel, AMD, Nvidia, IBM and Cray/ARM/Cavium, acting as technical representatives and subject matter experts.
- Multiple NERSC staff took on leadership positions within ECP.

In addition, NERSC — in collaboration with OLCF, ALCF, ESnet, ASCR and the other five Office of Science program offices — completed the Exascale Requirements Reviews with the publication of the Crosscut Report. These reviews brought together the DOE SC user community for discussions and requirements gathering targeting the 2020 and 2025 time frames, and their findings and recommendations are contained in the report. In partnership with ALCF and OLCF, NERSC also developed and deployed a website, “Portability Across DOE Office of Science HPC Facilities” (<http://performanceportability.org>). The website is a documentation hub and guide for application teams targeting the advanced computational resources at the DOE Office of Science facilities.

In 2017, Shyh Wang Hall — the Berkeley Lab building in which NERSC is located — was awarded LEED gold certification for its environmental and energy efficient design, and NERSC continues to work with Berkeley Lab’s Sustainability Office to further improve the energy efficiency of our data center. Seven projects were identified that will reduce energy and water use, and in 2017 two were completed with a net savings of 700,000 KWhr. Additional power metering has been added to the building to improve accuracy

of power usage effectiveness (PUE) measurement, and near real-time PUE calculation is now performed using NERSC’s data-collection system.

TOWARD THE NEXT GENERATION OF NERSC SYSTEMS

NERSC has been actively planning NERSC-9, which will be deployed in 2020, and has started pathfinding for NERSC-10, which will be deployed in 2024/2025. A growing issue for supercomputing centers is the amount of power needed to operate leading-edge systems. NERSC started its transition to energy-efficient architectures with Cori, and our goal is to provide our users a consistent path to exascale technologies in the next decade. The mission need for NERSC-9 was approved in August of 2015:

- Provide a significant increase in computational capabilities over Edison.
- Provide a system that will demonstrate exascale-era technologies and enable the user community to continue transitioning to advanced energy-efficient architectures.
- Provide support for new science initiatives and extreme data analysis from experimental facilities.
- NERSC-9 will be our first supercomputer with an all-flash filesystem to accelerate I/O. It will deploy a flexible, high-speed network to support large-scale simulation and data analysis, as well fast connectivity to experimental and observational facilities. We are very excited about the system; we expect to announce it in late 2018 and issue a new call for NESAP proposals shortly thereafter.

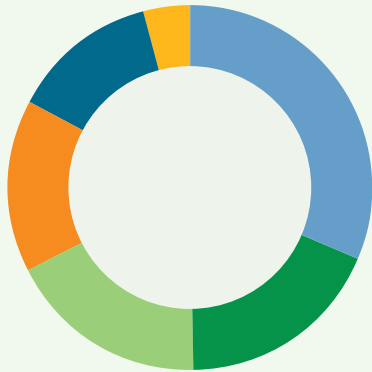
Sudip Dosanjh
NERSC Division Director

2017 Usage Demographics



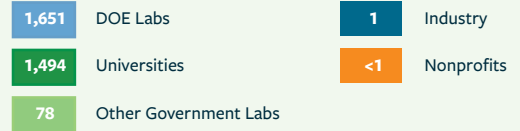
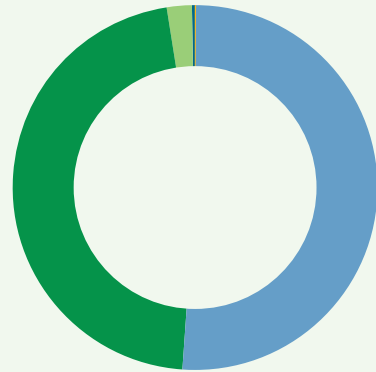
2017 NERSC USAGE BY DOE OFFICE OF SCIENCE PROGRAM

(NERSC Hours Charged in Millions)



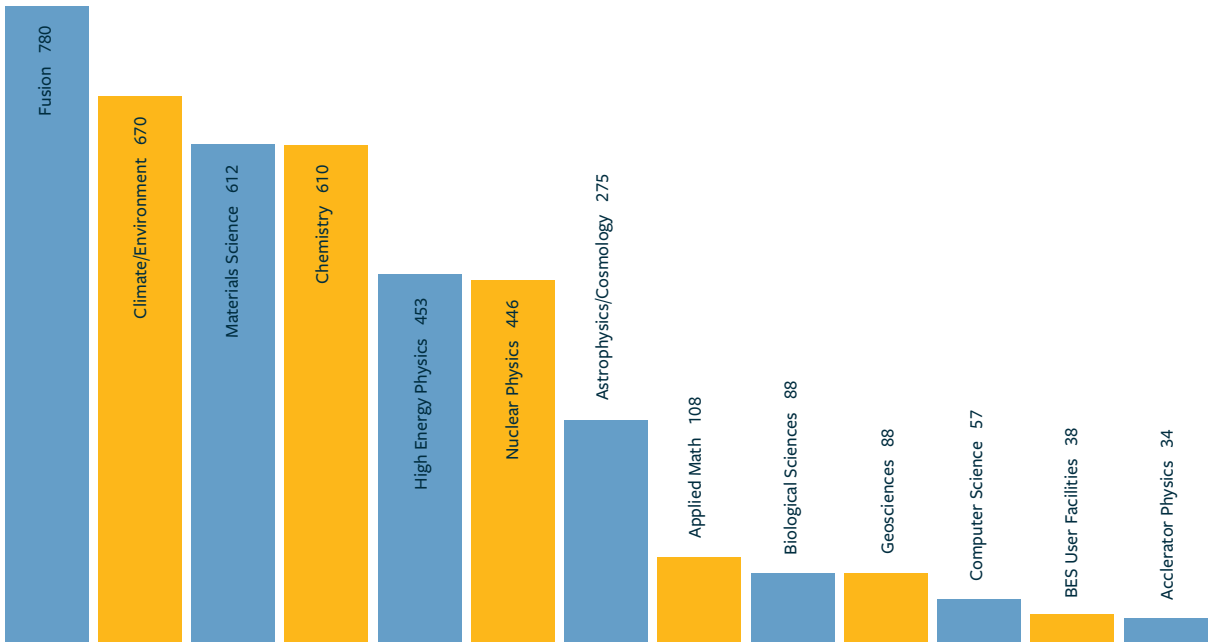
2017 NERSC USAGE BY INSTITUTION TYPE

(NERSC Hours Charged in Millions)



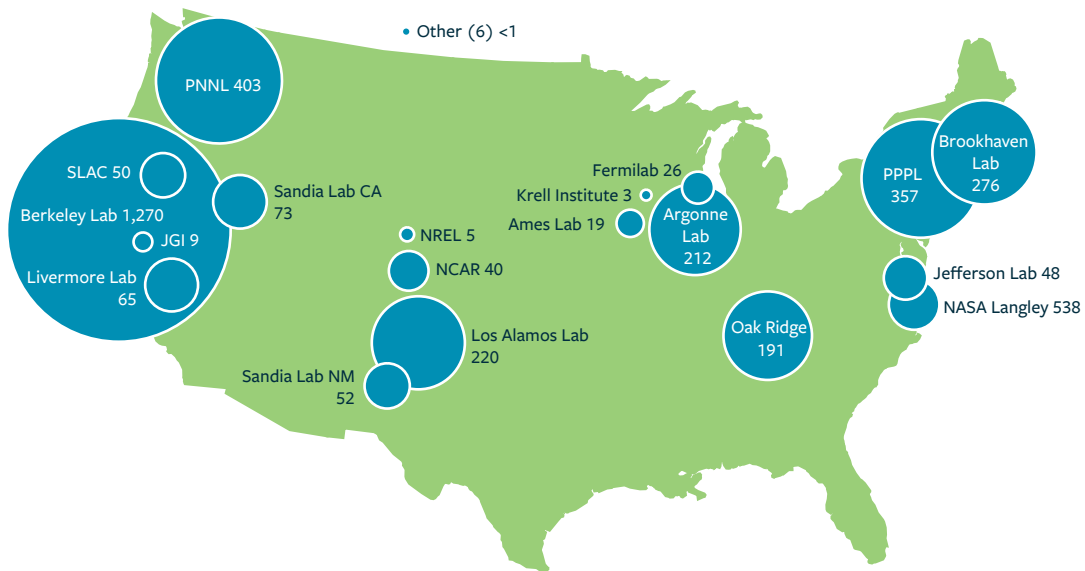
2017 NERSC USAGE BY DISCIPLINE

(NERSC Hours Charged in Millions)



2017 DOE & OTHER LAB USAGE AT NERSC

(NERSC Hours Charged in Millions)

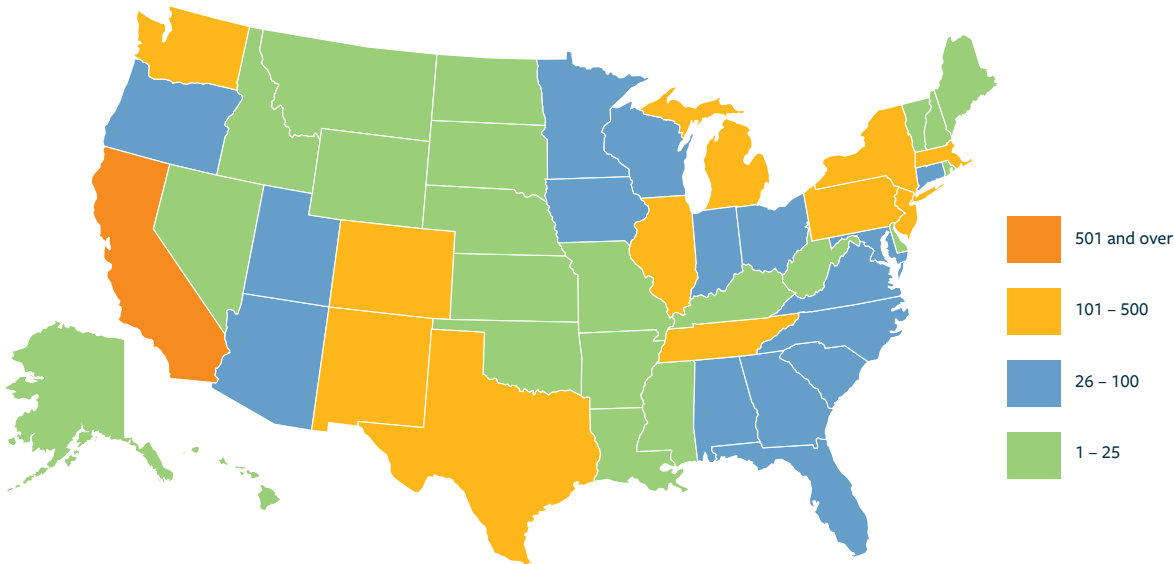


2017 ACADEMIC USAGE AT NERSC

(NERSC Hours Charged in Millions)

U. Wisconsin Madison	222	U. Chicago	33	George Wash Univ	15
MIT	194	U. Oklahoma	33	Penn State	15
U. Colorado Boulder	140	Stanford Univ	27	U. Texas Austin	15
Columbia Univ	121	U. Rochester	27	U. Central Florida	14
UC Berkeley	110	UC Santa Barbara	24	Auburn Univ	13
U. Arizona	91	Iowa State	22	Cornell Univ	13
UC Irvine	77	Michigan State	22	U. Florida	11
Northwestern Univ	65	U. Michigan	22	U. South Carolina	11
UCLA	64	U. Penn	22	U. Houston	9
UC San Diego	55	U. Illinois U-C	19	Indiana State	8
Princeton Univ	53	Yale Univ	19	Scripps Institute	8
U. Kentucky	51	Dartmouth College	17	William & Mary	8
Cal Tech	47	U. Minnesota	17	Harvard Univ	7
U. Washington	40	Carnegie Mellon	16	U. Arkansas	7
Temple Univ	35	U. Maryland	16	U. Kansas	7
Colorado School of Mines	34	Vanderbilt Univ	16	Georgia State	6

2017 NERSC USERS BY STATE



Science Highlights

With more than 7,000 users, 900 projects and 2,400 peer-reviewed published papers reported in 2017, NERSC users' scientific accomplishments could not be fully covered in this report. Instead, here we present some sample summaries chosen to cover the spectrum of scientific research and data-focused projects supported by NERSC.



Assessing Regional Earthquake Risk in the Age of Exascale

Exascale Computing Project

SCIENTIFIC ACHIEVEMENT

Researchers from Berkeley Lab, Lawrence Livermore National Laboratory (LLNL) and the University of California at Davis are building the first end-to-end simulation code designed to precisely capture the geology and physics of regional earthquakes and how the shaking impacts buildings. This work is part of the DOE's Exascale Computing Project.

SIGNIFICANCE AND IMPACT

One of the most important variables that affect earthquake damage to buildings is seismic wave frequency, the rate at which an earthquake wave repeats each second. Buildings and structures respond differently to certain frequencies. Large structures like skyscrapers, bridges and highway overpasses are sensitive to low-frequency shaking, whereas smaller structures like homes are more likely to be damaged by high frequency shaking, which ranges from 2 to 10 hertz and above. Due to computing limitations, current geophysics simulations at the regional level typically resolve ground motions at 1-2 hertz.

The Berkeley Lab/LLNL/UC Davis team is working to achieve motion estimates on the order of 5-10 hertz to accurately capture the dynamic response for a wide range of infrastructure. With emerging exascale supercomputers, they expect to be able to accurately simulate the ground motions of regional earthquakes quickly and in unprecedented detail, as well as predict how these movements will impact energy infrastructure — from the electric grid to local power plants — and scientific research facilities.

RESEARCH DETAILS

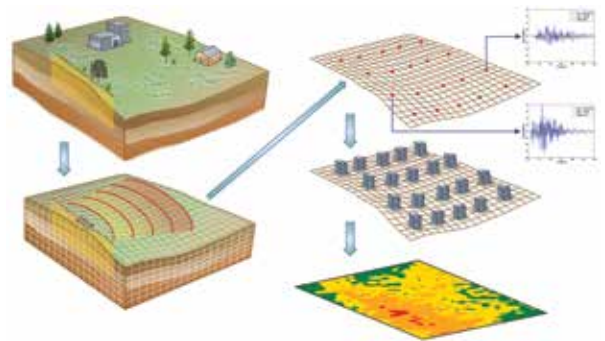
The researchers updated the existing SW4 code — developed at LLNL to simulate seismic wave propagation — to take advantage of the latest supercomputers, like NERSC's Cori system. With the updates to SW4, in 2017 the team

successfully simulated a 6.5 magnitude earthquake on California's Hayward fault at 3 hertz on Cori in about 12 hours on 2,048 Knights Landing nodes. This first-of-a-kind simulation also captured the impact of this ground movement on buildings within a 100-square kilometer (km) radius of the rupture, as well as 30km underground. With future exascale systems, the researchers hope to run the same model at 5–10 hertz resolution in approximately five hours or less.

Principal Investigator: David McCallen, Berkeley Lab

Journal Citation: Johansen et al, *Computing in Science & Engineering*, Vol. 19, Issue 5, September/October 2017

Full Story: <http://bit.ly/NERSCearthquakeexascale>



Transforming hazard into risk: Researchers at Berkeley Lab, LLNL and UC Davis are utilizing ground motion estimates from a regional-scale geophysics model to drive infrastructure assessments.

HPC for Better Computer Memory Technologies

Basic Energy Sciences

SCIENTIFIC ACHIEVEMENT

In a 2017 *Nature Materials* article, a team of researchers from Yale University, the National University of Singapore and the Indian Association for the Cultivation of Science described a new, robust “memristor” device that can last for 1 trillion cycles, far surpassing the endurance of commercial flash memories for computing. They also reported theoretical insights into why it has the size, stability, reproducibility and endurance to supplant flash memory technologies in future generations of digital devices.

SIGNIFICANCE AND IMPACT

A memristor is an electrical resistor component with memory that can regulate the electrical current in a circuit while remembering the level of charge that goes through it. However, previously developed memristors have been too slow to change states and unable to hold the memory of that state for long enough to be useful. This new type of memristor can switch states in 30 nanoseconds or less, comparable to traditional resistors, and is capable of holding that state for over 11 days without any power.

The research team performed calculations at NERSC to gain an understanding of why their memristor device, based on a spin-coated active layer of a transition metal complex, performs so well. The insight may accelerate the deployment of organic resistive memory devices, and the findings have wider applicability for other semiconductor materials, particularly those used in neuromorphic and logic circuits.

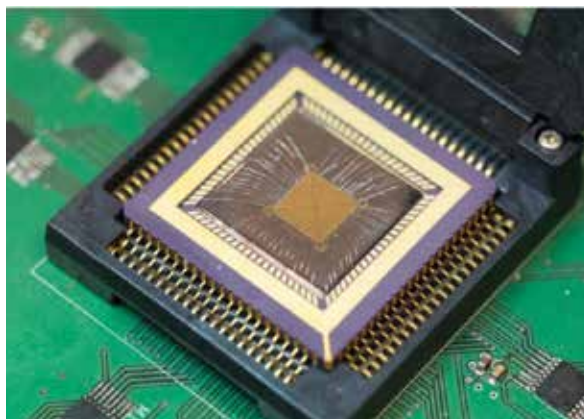
RESEARCH DETAILS

The researchers used a computational methodology that combines large-scale density functional theory molecular dynamics and simulations of electronic relaxation to accurately describe the electronic relaxation processes in functionalized titanium-dioxide surfaces.

Principal Investigator: Victor Batista, Yale University

Journal Citation: S. Goswami, et al, *Nature Materials*, October 2017, doi: 10.1038/NMAT5009

Full Story: <http://bit.ly/NERSCmemristor>



A memristor chip developed at the University of Michigan. Memristors can both perform logic and store data.

TOAST: Preparing for Next-Generation CMB Experiments

High-Energy Physics

SCIENTIFIC ACHIEVEMENT

Berkeley Lab cosmologists achieved a critical milestone in 2017 while preparing for upcoming Cosmic Microwave Background (CMB) experiments: scaling their data simulation and reduction framework TOAST (Time Ordered Astrophysics Scalable Tools) to run on all 658,784 Intel Knights Landing Xeon Phi processor cores on NERSC's Cori supercomputer. The simulated data included the sky signal plus realistic instrumental noise and atmospheric fluctuations.

SIGNIFICANCE AND IMPACT

Next-generation experiments like CMB Stage-4 will probe the Big Bang with unprecedented sensitivity, gathering orders of magnitude more data than previous experiments. The full Cori system — and its successors — will be needed to design the experiments, build and validate their analysis pipelines, reduce the experimental data and deliver the reduced data to the scientific community.

RESEARCH DETAILS

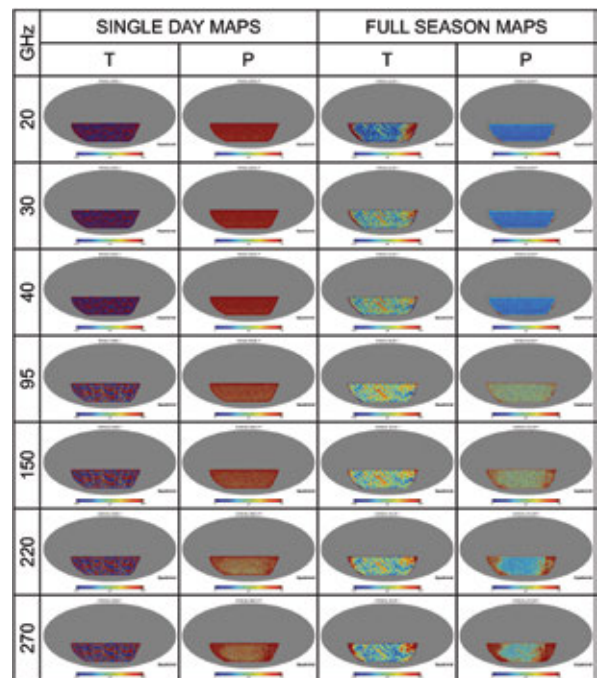
The optimized TOAST code was deployed in a Docker container and the application was run with the NERSC-developed Shifter container-for-HPC software, enabling a start-up time of < 90 sec. Researchers from Berkeley Lab's Computational Cosmology Center (C3) team achieved a balance between computing performance and accessibility by creating a hybrid application. Parts of the framework are written in C and C++ to ensure that it can run efficiently on supercomputers, but it also includes a layer written in Python so that researchers can easily manipulate the data and prototype new analysis algorithms.

The C3 team worked closely with staff from NERSC, Intel and Cray to get the TOAST code to run on all of Cori's processors. This collaboration was part of the NERSC Exascale Science Applications Program, which helps science

code teams adapt their software to take advantage of Cori's manycore architecture. To ensure all of TOAST could effectively scale up to all the cores, NERSC staff helped the C3 team launch their software on Cori with Shifter.

Principal Investigator: Julian Borrill, Berkeley Lab

Full Story: <http://bit.ly/NERSCtoastCMB>



Cumulative daily maps of the sky temperature and polarization at each frequency showing how the atmosphere and noise integrate down over time. The simulated data included the sky signal plus realistic instrumental noise and atmospheric fluctuations.

The Origin of Heavy Elements

Nuclear Physics

SCIENTIFIC ACHIEVEMENT

In a study published in *Nature*, an international team of scientists found that optical and infrared signals from a neutron star merger were found to be consistent with computer model predictions of a kilonova, an event that is 1,000 times brighter than a classical nova. The observations followed the detection of gravitational waves from the neutron star merger GW170817. Such events have been theorized to seed the universe with heavy elements like gold and platinum and radioactive elements like uranium.

SIGNIFICANCE AND IMPACT

Physicists have long discussed how elements heavier than iron could be produced during a neutron star merger. Computer simulations had suggested that, during such a merger, a small fraction of neutron star matter would be flung into surrounding space. Models predicted that this cloud of exotic debris would assemble into heavy elements and give off a radioactive glow over 10 million times brighter than the sun (a kilonova). The observations in the *Nature* paper are the first to provide strong evidence to support this theory.

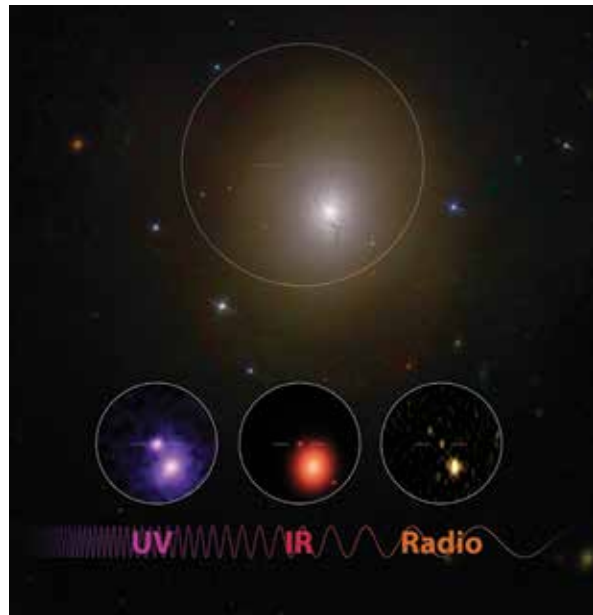
RESEARCH DETAILS

The researchers used 4,000 compute cores and 4 million compute hours on NERSC supercomputers to run the general relativistic Magnetohydrodynamical code HARMPI and the radiation transport code SEDONA. They were able to produce highly resolved simulations of the long-term aftermath and observable light from neutron star mergers.

Principal Investigator: Daniel Kasen, Berkeley Lab

Journal Citation: D. Kasen, et al, *Nature* 551, 80–84 (Nov. 2, 2017), doi:10.1038/nature24453

Full Story: <http://bit.ly/NERSCneutronstarmerger>



An illustration showing a simulated merger of a pair of neutron stars.

A Record Quantum Circuit Simulation

Advanced Scientific Computing Research

SCIENTIFIC ACHIEVEMENT

Researchers from the Swiss Federal Institute of Technology (ETH Zurich) used NERSC's 30-petaflop supercomputer, Cori, to successfully simulate a 45-qubit (quantum bit) quantum circuit, the largest simulation of a quantum computer achieved to date.

This achievement moved the research community another step closer to simulating “quantum supremacy” — the point at which quantum computers become more powerful than ordinary computers.

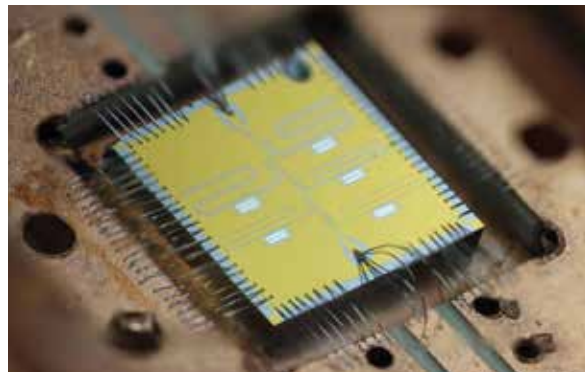
SIGNIFICANCE AND IMPACT

The current consensus is that a quantum computer capable of handling 49 qubits will offer the computing power of the most powerful supercomputers in the world. This simulation was an important step toward achieving “quantum supremacy” — the point at which quantum computers finally become more powerful than ordinary computers.

RESEARCH DETAILS

In addition to the 45-qubit simulation, which used 0.5 petabytes of memory on Cori and achieved a performance of 0.428 petaflops, the researchers also simulated 30-, 36- and

42-qubit quantum circuits. For the 45-bit simulation, they used 8,192 of 9,688 Intel Xeon Phi processors on Cori and 0.5 petabytes of memory. The researchers' optimizations improved the performance — the number of floating-point operations per time — by between 10x and 20x for Cori (depending on the circuit to simulate and the size per node). The time-to-solution decreased by over 12x when compared to the times of a similar simulation reported in a paper on quantum supremacy, which made the 45-qubit simulation possible.



Prototype multi-qubit chip developed at the Quantum Nanoelectronics Laboratory at UC Berkeley.

NERSC PI: Thomas Häner, ETH Zurich

Journal Citation: T. Häner, D. S. Steiger, “0.5 Petabyte Simulation of a 45-Qubit Quantum Circuit,” *SC '17 Proceedings of the International Conference for High Performance Computing, Networking, Storage and Analysis*, Article No. 33, arXiv:1704.01127

Full Story: <http://bit.ly/NERSCquantumqbit>

How Catalysts Convert CO₂ into a Useful Fuel

Basic Energy Sciences

SCIENTIFIC ACHIEVEMENT

Using supercomputers at NERSC in combination with a series of experiments, scientists at Brookhaven National Laboratory identified how a zinc/copper catalyst transforms carbon dioxide (CO₂) and hydrogen into methanol, a useful fuel that could reduce both pollution and our dependence on petroleum products.

SIGNIFICANCE AND IMPACT

Reacting carbon dioxide with hydrogen can produce methanol. But the reaction will not take place on its own; a catalyst is needed to initiate the process. Catalysts bring the reacting chemicals together in a way that makes it easier for them to break and rearrange their chemical bonds. In industrial applications, catalysts made from copper and zinc oxide on alumina supports are often used, but the process itself is not well understood and has been the subject of much debate. Understanding details of these molecular interactions could point to strategies to improve the catalysts for more energy-efficient reactions.

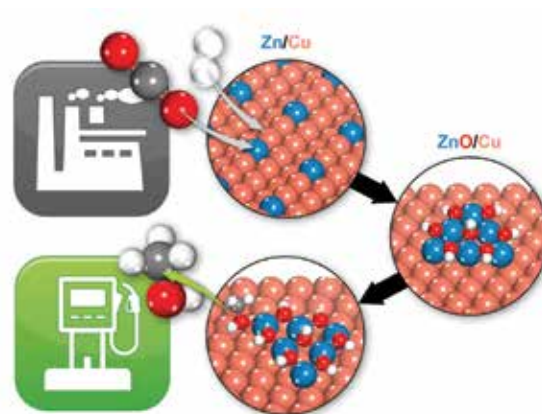
RESEARCH DETAILS

The researchers used computational resources at NERSC to model how two types of model catalysts — one made of zinc nanoparticles supported on a copper surface, and another with zinc oxide nanoparticles on copper — would engage in the CO₂-to-methanol transformations. These theoretical studies used calculations that took into account the basic principles of breaking and making chemical bonds, including the energy required, the electronic states of the atoms and the reaction conditions, allowing scientists to derive the reaction rates and determine which catalyst will give the best rate of conversion.

Principal Investigator: Ping Liu, Brookhaven National Laboratory

Journal Citation: S. Kattel, et al, *Science*, 2017; 355 (6331): 1296

Full Story: <http://bit.ly/NERSCconvertingCO2>



Brookhaven scientists used NERSC supercomputers to identify how a zinc/copper catalyst transforms carbon dioxide (two red balls and one grey ball) and hydrogen (two white balls) to methanol (one grey, one red, and four white balls), a potential fuel.

‘Hindcasting’ Extreme Weather Events

Biological and Environmental Research

SCIENTIFIC ACHIEVEMENT

Using a publicly available climate model, Berkeley Lab researchers “hindcasted” the conditions that led to the September 9–16, 2013 floods around Boulder, Colorado and found that climate change attributed to human activity made the storm more severe than would otherwise have occurred.

SIGNIFICANCE AND IMPACT

Changes in the risk of extreme weather events pose some of the greatest risks to society and the environment. Hindcasting is a way of testing a mathematical model; researchers enter known or closely estimated inputs for past events into the model to see how well the output matches the known results. By running simulations using various climate scenarios, researchers were able to determine that the storm was more severe in today’s climate than it would have been in one without climate change.

RESEARCH DETAILS

The storm was so strong and intense that standard climate models that do not resolve fine-scale details were unable to characterize the severe precipitation or large-scale meteorological pattern associated with the storm. So the research team used the publicly available Weather Research and Forecasting regional model to study the problem in greater detail, breaking the area into 12-kilometer squares. They used NERSC’s Edison supercomputer to run 101 hindcasts.

This project is contributing to the international C20C+ Detection and Attribution Project. NERSC is hosting the data portal for that project on its Earth System Grid Federation node.

Principal Investigator: Travis O’Brien, Berkeley Lab

Journal Citation: P. Pall, et al, *Weather and Climate Extremes*, July 2017, doi: 10.1061/j.wace.2017.03.004

Full Story: <http://bit.ly/NERSchindcasting>



This picture was taken above Greeley, Colo. on September 16, 2013, following a week’s worth of heavy rainfall that caused historic flooding. Computer models run at NERSC helped researchers better understand what caused the severe weather.

More Dust, Better Air Quality

Biological and Environmental Research

SCIENTIFIC ACHIEVEMENT

Aerosol pollution is one of the most important environmental issues in China. Simulations run at NERSC by researchers from Pacific Northwest National Lab, Scripps Institution of Oceanography and UC San Diego revealed that man-made pollution in eastern China worsens when less dust blows in from the Gobi Desert.

SIGNIFICANCE AND IMPACT

In recent years, eastern China has suffered from heavy haze events with high aerosol concentrations, which has adverse impacts on hundreds of millions of people across the country. The researchers found that dust plays an important role in determining air temperatures and thereby promoting winds to blow away man-made pollution. Less dust means the air stagnates, with man-made pollution becoming more concentrated and sticking around longer.

RESEARCH DETAILS

Using computer models run at NERSC, together with historical data, the researchers found that reduced natural dust transported from the Gobi Desert translates to increased air pollution in highly populated eastern China. The scientists found that reduced dust causes a 13 percent increase in man-made pollution over eastern China during the winter. The results match observational data from dozen of sites in eastern China. The team also found that two to three days after winds had brought dust into the region from western China, the air was cleaner than before the dust arrived.



Gobi Desert dust envelops eastern China.

Less dust means the air stagnates, with man-made pollution becoming more concentrated and sticking around longer.

Principal Investigator: Steven Ghan, Pacific Northwest National Laboratory.

Journal Citation: Y. Yang, et al, *Nature Communications* 8, 15333, May 2017

Full Story: <http://bit.ly/NERSCdustairquality>

Shedding Light on Mysterious Plasma Flows

Fusion Energy Sciences

SCIENTIFIC ACHIEVEMENT

Researchers at Princeton Plasma Physics Laboratory and General Atomics used NERSC supercomputers to simulate a mysterious self-organized flow of the superhot plasma that fuels fusion reactions. Their findings show that pumping more heat into the core of the plasma can drive instabilities that create plasma rotation inside the doughnut-shaped tokamak reactor that houses the hot charged gas.

SIGNIFICANCE AND IMPACT

The findings could lead to improved control of fusion reactions in ITER, the international experiment under construction in France to demonstrate the feasibility of fusion power, and other fusion devices. With careful experiments and detailed simulations of fundamental physics, the researchers are beginning to understand how the plasma creates its own sheared rotation. This is a key step to optimizing the plasma flow to make fusion plasmas more stable and operate with high efficiency.

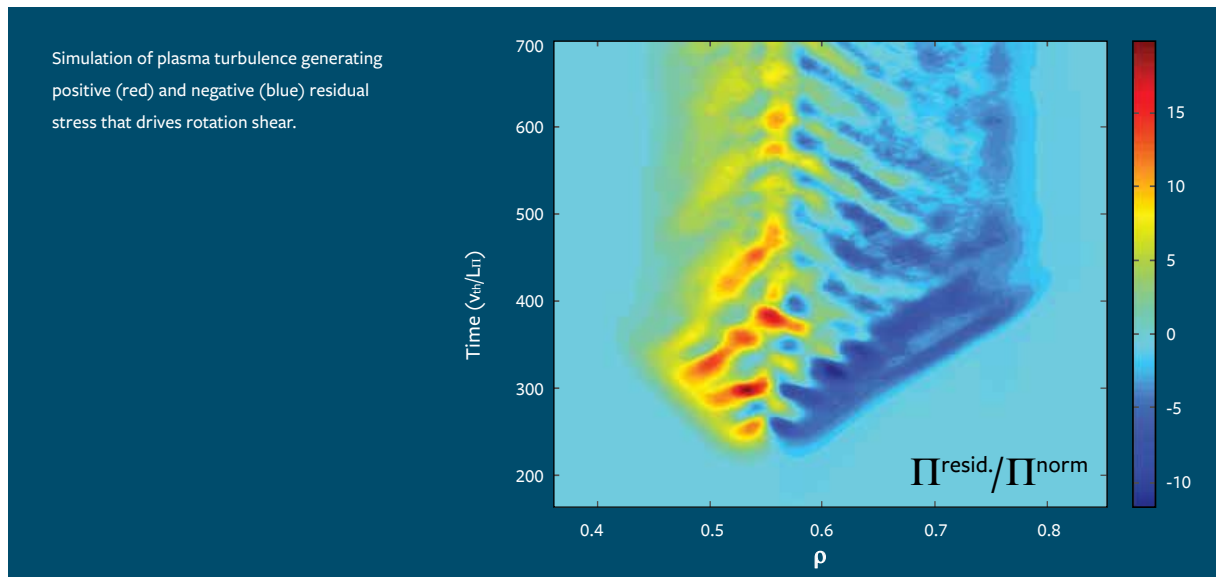
RESEARCH DETAILS

The researchers used GTS, a first principles kinetic code, to simulate the physics of turbulent plasma transport by modeling the behavior of plasma particles as they cycled around magnetic fields. The simulation predicted the rotation profile by modeling the intrinsic torque of the turbulence and the diffusion of its momentum. The predicted rotation agreed well, in shape and magnitude, with the rotation observed in DIII-D experiments.

Principal Investigator: Wei-Li Lee, Princeton Plasma Physics Laboratory

Journal Citation: A. Grierson et al, *Physical Review Letters*, 118, 015002, January 2017

Full Story: <http://bit.ly/NERSCplasmaflows>



First Complex Calculations of Sigma Particle

Nuclear Physics

SCIENTIFIC ACHIEVEMENT

Supercomputing resources at NERSC, Oak Ridge National Laboratory and the University of Illinois at Urbana-Champaign helped an international team of researchers representing the Hadron Spectrum Collaboration achieve the first complex calculations of a subatomic particle called the Sigma. The team did this by generating quantum chromodynamic (QCD) gauge configurations: snapshots of the environment of subatomic particles (the vacuum of space described by QCD).

SIGNIFICANCE AND IMPACT

After decades of catching only brief glimpses of the Sigma's existence from experimental data that showed its effects on other subatomic particles, this calculation gives scientists a new way to study the Sigma and gain new insights about the "strong force" that exists inside all matter.

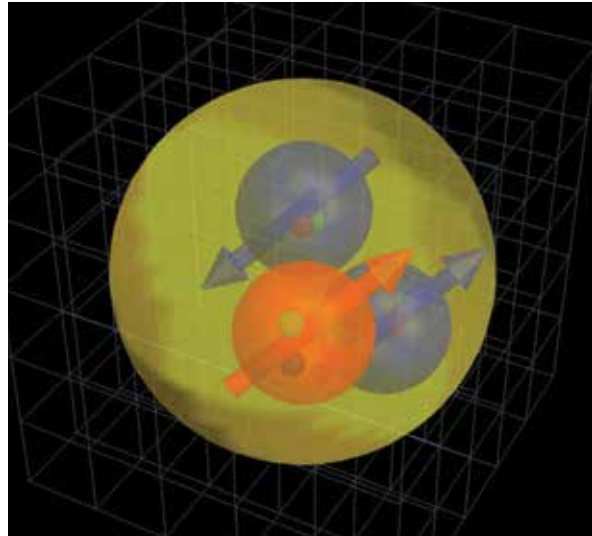
RESEARCH DETAILS

This study is part of a larger effort to investigate QCD, the fundamental theory of strong interactions, and gain deeper insights into the fundamental laws of physics. Experiments at nuclear and high energy physics labs around the world measure the properties of matter with the aim of determining its underlying structure. For this study, the researchers used the Chroma code package, which supports data-parallel programming constructs for lattice field theory and, in particular, lattice QCD.

Principal Investigator: Robert Edwards, Jefferson Lab

Journal Citation: R. Briceño, et al, *Physical Review Letters*, 118, 022002, January 2017

Full Story: <http://bit.ly/NERSCheartofmatter>



Quantum chromodynamics is the fundamental theory of strong interactions (the mechanism responsible for the strong nuclear force) and one of the four known fundamental interactions in physics.

Predicting Defect-Free Solar Cell Nanomaterials

Basic Energy Sciences

SCIENTIFIC ACHIEVEMENT

Heterogeneous nanostructured materials are used in various optoelectronic devices, including solar cells. But the interfaces contain structural defects that can affect device performance. Using atomistic calculations run at NERSC, researchers from Argonne National Laboratory and the University of Chicago found the root cause of the defects in two materials and provided design rules to avoid them.

SIGNIFICANCE AND IMPACT

In the ongoing effort to improve solar cell energy conversion efficiencies, researchers have begun digging deeper to identify material defects that can undermine the conversion process. For this study, the researchers focused on heterostructured nanoparticles, developing a computational strategy to investigate at the atomic level the effect of the structure of the interfaces on the materials' optoelectronic properties, identifying certain atomic "trap states."

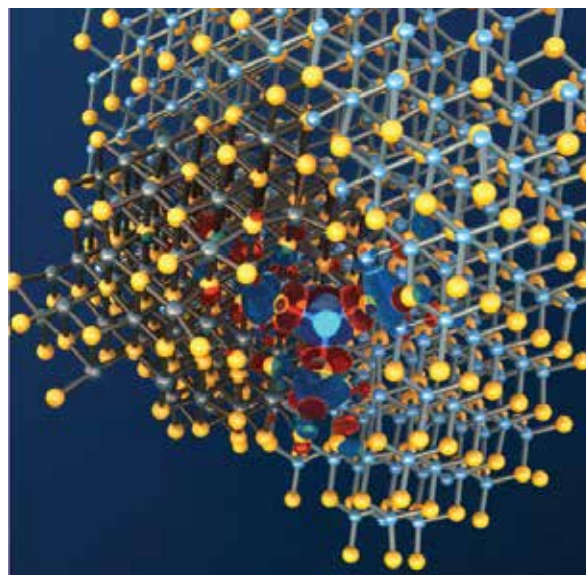
RESEARCH DETAILS

The research team ran a series of atomistic calculations to predict a new material that does not have these trap states and should perform better in solar cells. By using classical molecular dynamics and first principles methods that do not rely on any fitted parameters, their framework allowed them to build computational models of embedded quantum dots. This study — which included studies of atomic and electronic structures — used four million supercomputing hours at NERSC. Most of the atomic structure calculations were run on Cori, NERSC's 30-petaflop system, with the rest being run on Edison.

Principal Investigator: Martin Vörös, Argonne National Laboratory

Journal Citation: F. Giberti, et al, *Nano Letters*, 2017, 17 (4), pp 2547–2553, doi: 10.1021/acs.nanolett.7b00283

Full Story: <http://bit.ly/NERSCatomic-leveldefects>

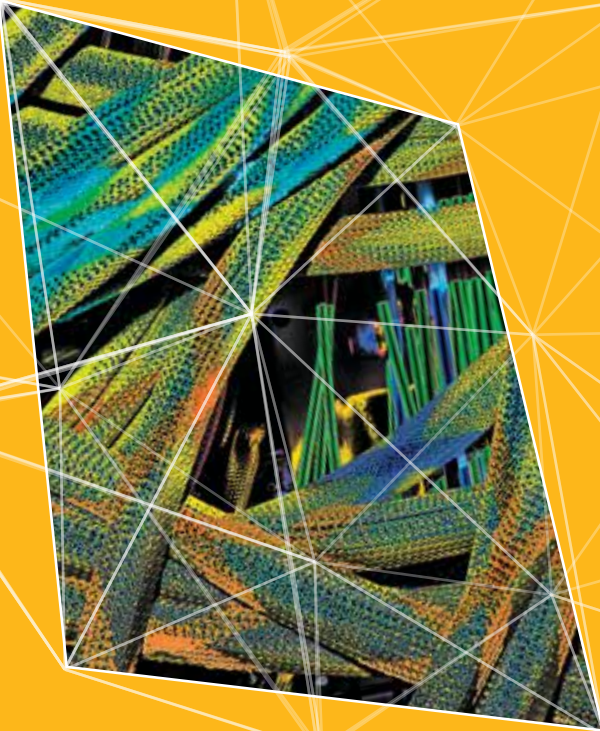


A cross section of the interface between a lead chalcogenide nanoparticle and its embedding cadmium chalcogenide matrix.

Innovations

NERSC supported 7,405 users from universities, national laboratories and industry in 2017, providing high-end computing, storage and expertise to the broad science community. The NERSC workload represents the wide variety of research performed by users, including simulations that run at the largest scales available on the center's world-class supercomputers.

According to the 2017 NERSC User Survey,* 35 percent of users reported that their projects involved either analyzing experimental and observational data, combining simulations with data analysis or creating tools and algorithms to improve data analysis. Office of Science experiments and facilities that combine HPC simulation and modeling with data analysis from instruments such as microscopes, telescopes and particle accelerators are influencing the way NERSC configures systems and supports users, a trend that is reflected in the many innovative technologies and services introduced or expanded in 2017.



*Actual user quotes from the 2017 NERSC User Survey can be found on pages 27, 36, 43 and 45 of this Annual Report.



SUPPORTING GORDON BELL RUNS

In November 2016, NERSC's Cori supercomputer debuted at #5 on the Top500 List, making it an attractive candidate for code teams interested in the Gordon Bell competition. The Gordon Bell Prize is awarded each year to recognize outstanding achievement in HPC. The purpose of the award is to track the progress over time of parallel computing, with particular emphasis on rewarding innovation in applying HPC to applications in science, engineering and large-scale data analytics. Prizes may be awarded for peak performance or special achievements in scalability and time-to-solution on important science and engineering problems.

In 2017, NERSC initiated a call for proposals for the large-scale science program, and eight projects were selected for Gordon Bell runs on the Cori supercomputer: PICSAR/WARP, HipMer, HPX/OctoTiger, 45-qubit Quantum Simulator, BerkeleyGW, Celeste, Galactos and Deep Learning @ 15PF. Projects were selected on the basis of the Gordon Bell criteria, but also to maximize scientific impact. For example, in the traditional HPC space, BerkeleyGW was used in the largest known conventional GW simulations (the current state of the art) of excited state properties of defects in silicon. HPX/OctoTiger is a code written in HPX, a potential

future addition to the C++ standard, and was used here to investigate binary star mergers. HipMer is a high-performance metagenomics assembly pipeline targeting both throughput of production-sized datasets and the largest available metagenomic datasets.

“Excellent, consistent HPC resources and data storage.”

Selected projects were assigned NERSC staff liaisons that included performance engineers, consultants and system administrators. Projects shared reservations and used the Slack messaging system to coordinate across multiple groups debugging codes at scale and to keep all NERSC staff involved in the reservations in close communication. This enabled one team to step in and use the full system if another was unable to run at that time. This level of engagement and coordination allowed multiple groups to do their simulations while minimizing impact to other users and system utilization.

EXAHDF5 TEAM PARTNERS WITH THE EXASCALE COMPUTING PROJECT

In 2017, the ExaHDF5 project — a collaboration of developers and researchers at NERSC, ALCF and the non-profit The HDF Group — was funded to work with ECP application developers to improve their knowledge and effective use of HDF5 and to enhance the HDF5 library for use in exascale environments.

HDF5 (Hierarchical Data Format version 5) is a set of file formats, libraries and tools for storing and managing large scientific datasets. The ExaHDF5 project researches improvements to, implements capabilities in, makes production releases of and supports users of the HDF5 I/O middleware package in HPC environments.

As part of the first year of ECP-ExaHDF5 funding, the team is on track to deliver the following feature milestones:

- **Virtual object layer (VOL)**, which allows interception of HDF5 public API at runtime to access data in alternate ways

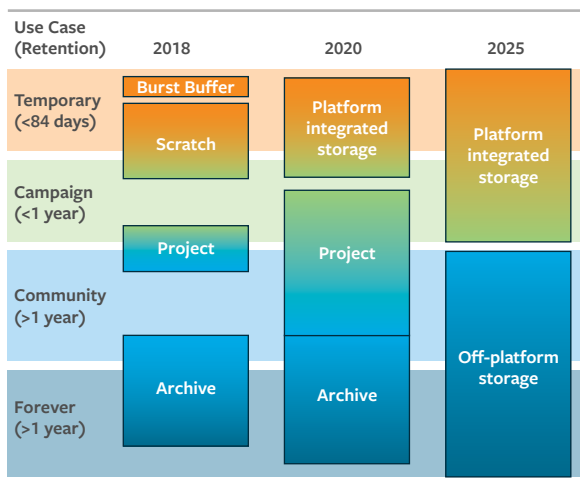
- **Data Elevator**, a write-caching VOL plugin, to intercept HDF5 writes and redirect them to use burst buffer storage transparently
- **Topology-aware I/O in HDF5**, to take advantage of the underlying storage network topology for enhancing the I/O performance of applications
- **Full single writer/multiple readers** application access, which supports advanced workflows by allowing simultaneous, lock-free access to HDF5 data files from one writing application and many reading applications

The ExaHDF5 team has also been funded to expand and enhance HDF5 to store experimental and observational data throughout the DOE complex. As part of this effort, the collaborators will research enhancements to HDF5 to record provenance information, store sparse and stream-oriented data, perform multi-writer concurrent I/O and update the schema of HDF5 files over time.

NERSC DEVELOPS STORAGE ROADMAP FOR 2020 — AND BEYOND

Storage systems play a critical role in supporting NERSC’s mission by enabling the retention and dissemination of science data used and produced at the center. Over the past 10 years, the total volume of data stored at NERSC has increased from 3.5 PiB to 146 PiB, growing at an annual rate of 30%,

driven by a 1000x increase in system performance and 100x increase in system memory. In addition, there has been dramatic growth in experimental and observational data, and experimental facilities are increasingly turning to NERSC to meet their data analysis and storage requirements.



Evolution of the NERSC storage hierarchy between 2018 and 2025.

At the same time, the technologies underpinning traditional storage in HPC are rapidly changing. Solid-state drives are being integrated into HPC systems as a new tier of high-performance storage, shifting the role of magnetic disk media away from performance, and tape technology revenues are slowly declining. To stay ahead of these changes and closely track the changing landscape, NERSC continuously engages with hardware vendors, system integrators and peers in enterprise markets through a variety of channels.

Active participation in PathForward deep-dives, regular attendance and presentations at storage industry conferences such as IEEE MSST and Flash Memory Summit, and informal dialogue with practitioners of advanced storage technologies in commercial sectors have provided NERSC with insight into the greater economic drivers shaping the field. Cloud and hyperscale data center providers are major

forces shaping the future of the mass storage ecosystem, rapidly advancing the state of the art in object-based storage systems over POSIX-based parallel file systems. In addition, non-volatile storage-class memory is emerging as a high-performance, low-latency medium for storage. The combination of these factors broadens the design space of future storage systems in HPC, creating new opportunities for innovation but also introducing new uncertainties.

To set a long-term storage strategy, NERSC developed and released a report, “NERSC Storage 2020,” presenting a detailed roadmap and storage vision to address the overwhelming data storage challenges the science community is expected to face over the next decade and beyond. As input to the report, NERSC used feedback from the Exascale Requirement Reviews, a set of workshops held from 2015–2017 in collaboration with the Oak Ridge and Argonne Leadership Computing Facilities and ESnet. In these workshops, scientists from each Office of Science program office were asked to describe their scientific grand challenges and computational and data requirements. The workshops identified massive data rate increases from various detectors and sensors and the need for analysis, management, archiving and curation capabilities beyond what is common today. The reports that came out of these reviews also emphasized the

growing complexity of scientific workflows from experimental facilities and the need to accommodate them on high performance computing systems.

Using the requirements reviews and a detailed workload analysis, NERSC identified four data storage categories required by the user community: temporary, campaign, forever and community. In parallel, NERSC conducted storage technology deep dives and held discussions and presentations with staff at other HPC facilities to determine how these four categories can map to physical storage systems. The roadmap sets a target of implementing three tiers by 2020 and two tiers by 2025, ultimately combining different types of storage media to simplify data management for users. The performance and scalability requirements of future systems will drive the industry toward object stores by 2025, and HPC centers such as NERSC will rely on middleware to provide familiar interfaces like POSIX and HDF5 for users who aren’t ready to change the way they perform I/O.

Because of the diversity of NERSC user workloads across scientific domains, this analysis and the reference storage architecture should be relevant to HPC storage planning outside of NERSC and the DOE, the report concludes.

NEW WEBSITE FACILITATES PERFORMANCE PORTABILITY

As the HPC community prepares for exascale and the semiconductor industry approaches the end of Moore’s Law in terms of transistor size, we have entered a period of time of increased diversity in computer architectures for HPC, with relatively new designs joining mature processor and memory technologies. These technologies include GPUs, manycore processors, ARM, FPGAs, ASICs, and new memory technologies in the form of high-bandwidth memory often incorporated on the processor die and non-volatile memory and solid-state disk technology for accelerated I/O.

With these trends in mind, NERSC, ANL and ORNL developed and implemented a website to facilitate

performance portability: “Portability Across DOE Office of Science HPC Facilities” (<http://performanceportability.org>). The website provides an evolving, ever-growing documentation hub and guide for application teams targeting the advanced computational resources at the DOE Office of Science facilities.

In addition, NERSC was part of the organizing committee for the DOE Centers of Excellence Performance Portability meeting held in Denver, Colorado in August 2017, where NERSC staff presented their work via talks, panels and poster sessions. NERSC will organize the next Performance Portability meeting, to be held in 2019.

BOOSTING DATA CENTER ENERGY EFFICIENCIES

NERSC has been working with Berkeley Lab's Chief Sustainability Officer and the DOE Center of Expertise for Energy Efficiency in Data Centers in Berkeley Lab's Energy Technology Area (ETA) to improve the energy efficiency of the data center. Seven specific projects have been identified that will reduce energy and water use, and to date two have been completed with a net savings of 660,000 kWh. The first of these, related to cooling efficiency, involved optimizing cooling tower fan and water pump speeds to reduce total energy consumption for cooling systems. (It takes less energy to run fans faster to make cooler water than to run pumps faster to cool with warmer water.) The second, related to power consumption, changed the mode of operation of the

building uninterruptible power supply to reduce conversion loss from DC power circuitry.

There is a potential of an additional 800,000 kWh savings from other work in progress; for example, NERSC has added more power metering to the building to improve accuracy of power usage effectiveness (PUE) measurement, and near real-time PUE calculation is now being performed using NERSC's data collection system. Other projects under way include optimizing closed loop pump control, replacing bypass valves and installing firmware to enable ESS mode for UPSs.

Summary of savings from improved energy efficiency of the NERSC data center.

	Measure Title	Energy Savings (kWh/year)			PUE	Water Savings
		Estimated	Verified	Total	Reduction	Gal/Year
1	Optimize Cooling Tower Fan and Pump Controls	-	310,000	310,000	0.006	-300,000
2	Optimize Closed Loop Pump Control	240,000	-	240,000	0.005	110,000
3	Optimize AHU SAT and Flow Control	300,000	-	300,000	0.006	-
4	Reset Cooling Water Supply Temperature	600,000	-	600,000	-0.001	220,000
5	Install Firmware to Enable ESS Mode for UPSs	-	350,000	350,000	0.007	120,000
6	Replace Bypass Valves	250,000	-	250,000	0.005	0
7	Cold Aisle Partial Containment	100,000	-	100,000	0.002	0
Total Savings		1,490,000	660,000	2,160,000	0.030	150,000
Baseline Non-IT Load				3,420,000		
Savings %				64%		

STREAMLINING NERSC SYSTEMS' DATA-COLLECTION CAPABILITIES

The amount of data flowing into the NERSC systems and environmental data-collection scheme — initially rolled out in 2016 — continues to grow substantially. In order to scale and ensure continuous availability, in 2017 NERSC separated the system's virtual machine (VM) cluster into two components: system-related data are now collected into one VM cluster, while environmental-related data are collected into another. This separation prevents system issues from impacting data-collection activities. With this scaling enhancement, NERSC was able to accommodate growth in usage of the data-collection system while avoiding a commercial software license — a potential cost of approximately \$200,000/year.

Within the new VM structure, NERSC was able to integrate the monitoring and notification functions. This new infrastructure has thus been re-named the Operations Monitoring and Notification Infrastructure (OMNI), which comprises the data-collection activities that monitor

everything in the NERSC facility and the notification system that sends alerts based on thresholds. It is an infrastructure that has become a more efficient way to ensure the health of the systems being monitored.

The Building Management System (BMS) monitors the data center environment, including air- and water-cooling systems, temperature and humidity. Any critical alerts have previously been sent to the on-call site facility technicians or specific facility staff. Using OMNI, the Operations Technology Group created an API interface between the BMS and the notification system that allows for correlation between BMS and OMNI environmental data to detect exceeded thresholds and send critical alerts to the 24x7 monitoring team at NERSC. This process has been extremely helpful in instances where the extreme exterior climate impacted the cooling system of the Crays, allowing staff to protect valuable hardware from overheating or freezing with minimal impact to users.

MODS: BUSINESS INTELLIGENCE FOR DATA SERVICES

NERSC's Data and Analytics Services (DAS) team has partnered with NERSC Operations staff to bring its favorite tool — data — to bear on the questions that inform its strategic decisions. In the course of developing a system to monitor usage of the many data-related tools and services DAS manages, it has also illuminated a path to monitoring opportunities throughout the center.

The group maintains a long list of software and services for users. As DAS is a relatively new group, one early challenge has been to determine the best portfolio of tools and services to support. Without data, the team needed to rely on intuition and experience to make its initial choices. But intuition and experience are much less useful than real data in a fast-moving field, so the Monitoring of Data Services (MODS) project was born to tap into usage patterns and generate visualizations to provide insight.

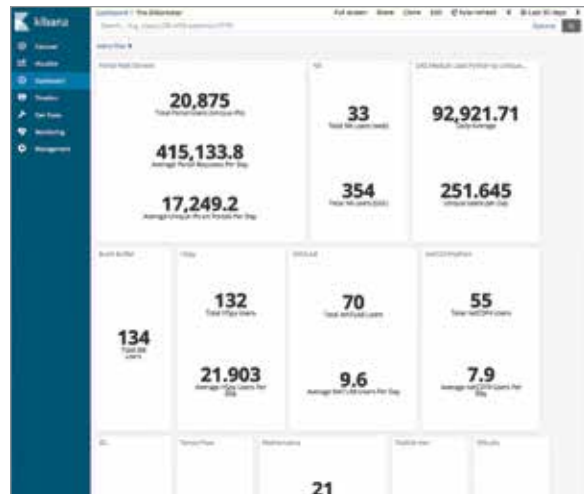
The MODS project leverages the centralized data collection system built by the Operations Group and detailed in NERSC's 2016 Operational Assessment. That system employs RabbitMQ to safely queue up a plethora of concurrent

messages to an Elasticsearch database, which is available centerwide. MODS builds on the ends of that system, shipping data in from service-specific scripts and pulling it out to custom-built dashboards built in the Elastic visualization tool Kibana. The DAS group now monitors 20 separate services via DASboard and DASometer. The former shows time series graphs of usage, enabling identification of trends and usage patterns, and the latter collects numeric metrics that show how many unique individuals have used each service in a given time span. The DAS group also holds a collection of working monitoring scripts and documentation that others can use to set up their own monitoring systems. Already, a similar system has gone into place for tracking usage of queues on Cori.

The DASboard and DASometer are already yielding insights into NERSC's data strategy. The DAS group has determined how to minimize disruption to NX users by shifting maintenance times to when it is least used and is seeing an uptick in use of Python libraries for HDF5. They've learned that the main science gateway servers are answering requests for almost 20,000 unique IP addresses per day, on average. In

addition, nearly all conflicts over use of MATLAB licenses have been eliminated simply by reminding users to close long-running sessions. The group continues to add monitoring

for more services and tools, aiming for complete coverage of the DAS portfolio.



Screenshots of the MODS DASboard (left), which monitors usage of data tools and services over time, and DASometer (right), which reports total counts of unique users of data tools and services.

CI/CD SIMPLIFIES SUPERCOMPUTER SUPPORT ENVIRONMENT

NERSC supports two large-scale HPC systems (Cori and Edison), one cluster system for the JGI and High Energy Physics workloads and a Test and Development System (TDS) for each. These six systems (the three computing platforms and the three TDSs) are of varying complexity, architecture and technology-vintage. Historically they used disparate configurations, and there was minimal reuse across systems. In 2017 NERSC implemented a process of continuous integration and continuous deployment (CI/CD) that simplifies the support environment and helps build toward a future vision of a single unified system configuration.

In conjunction with the CI/CD, the TDSs pre-stage updates (software, firmware, bug fixes) prior to release to the main production systems. The systems utilize a common set of package repositories, Ansible plays and Cray configuration files. The configurations and plays are managed in a private Git revision control system where changes utilize modern software development tools and models. Once merged,

changes can quickly be tested on the TDS systems. All changes are first vetted on the TDS system prior to deployment on the production system. This methodology both improves efficiency and reduces the impact to system availability for NERSC users.

Using these industry-standard approaches and tools, NERSC has attained a level of consistency across the Cray XC systems. Of note, both Cori and Edison are currently on the same operating system release environment for the first time. This common release environment improves user experiences and software portability; shortens time to solution for software patches and release enhancements; and reduces system downtime leading to increased system availability to the users. The CI/CD process reduces the workload on scarce system administration staff. The results of each pass through CI/CD releases additional staff cycles to focus on the next improvement or task, in a compounding fashion.

NERSC USERS TAKE SPIN FOR A SPIN

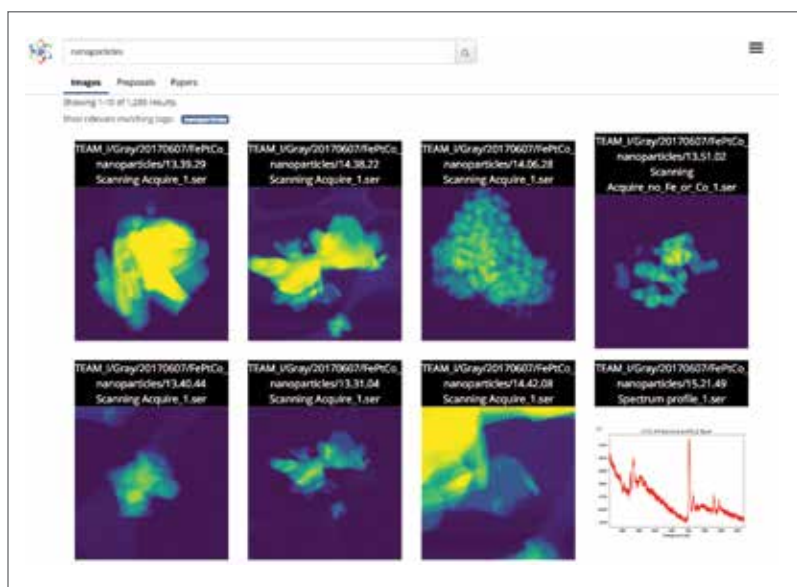
In 2016, NERSC reported early progress on a working prototype of Spin, a new system based on Docker containers that is designed to help NERSC users and staff quickly and easily create science gateways, workflow managers, web applications, databases and other “edge services.” These types of services are increasingly in demand in conjunction with computational projects and automated workflows, and Spin (Scalable Platform Infrastructure at NERSC) offers a scalable, cost-effective, low-overhead solution to meet this emerging need.

In 2017, NERSC evolved Spin from a prototype into a fully functional pilot production system that now underlies a broad variety of edge services at NERSC. The system is running on new hardware and networking, is connected to center storage systems and uses a commercially supported container management framework called Rancher at its core. A truly flexible system, Spin currently supports the user-facing tools JupyterHub and RStudio, a specialized cache service for the CernVM File System in use on Cori, numerous science gateways and network services created by users, several databases and even the license managers for our compiler and debugger packages. The system architects and engineers building Spin partnered closely with the

sponsors of these initial services — both NERSC staff and selected users — for their implementation. In the process, we developed additional knowledge and a documented set of conventions and best practices that will provide fundamental guidance for further deployments.

The most exciting aspect of Spin is how it empowers users to build the customized services they need to conduct their research. One of Spin’s early adopters is Gonzalo Rodrigo Alvarez, a postdoctoral researcher in Berkeley Lab’s Computational Research Division (CRD) who is using Spin to power his group’s first deployment of a software tool called Science Search. Science Search uses machine-learning techniques to extract metadata to build a search index across multiple scientific data sets. According to Alvarez, doing this requires a platform for web-based services that is tightly integrated to NERSC compute and storage resources. By deploying Science Search on Spin, they were able to seamlessly access NERSC resources while also enjoying the simplicity that comes from a Docker-based environment.

Another CRD project that was built on Spin is ESS-DIVE (Environmental Systems Science Data Infrastructure for a Virtual Ecosystem). This data archive is leveraging NERSC



Spin was instrumental in facilitating Berkeley Lab’s Science Search, a web-based search engine for scientific data that requires access to large datasets and analytical data that can only be stored on large supercomputing storage systems. In this screenshot of the Science Search interface, the user did an image search of nanoparticles.

storage via Spin to offer a scalable, reliable repository for earth science data. Data packages in the repository are also immediately accessible to NERSC compute systems, enabling the deep analysis and processing of data sets and metadata that is envisioned by ESS-DIVE project personnel.

The next phase for Spin will be to make it fully self-service, opening new possibilities for users to independently build and deploy the services they need on demand. User-facing tools, security controls and a training program are under active development and were expected to debut in May 2018.

AUTOMATING SOFTWARE MANAGEMENT USING SPACK

NERSC manages about 500 software packages, each with different versions, built with different compilers and for different architectures. Maintaining such a large software stack can lead to loss of consultation time due to the difficult nature of compiling software for a bleeding-edge architecture. In the past, NERSC consultants would manually download and install dependencies of scientific software. As an example, the package qt, a comprehensive cross-platform C++ application framework, has around 40 dependencies. Managing those dependencies manually leads to loss of important time that could be used to help users with more pertinent problems in research.

In order to mitigate the time lost to maintaining builds, NERSC has adopted Spack (Super Computing Package Manager), to install software on Edison and Cori. Typically packages need to be built for multiple compilers and library version combinations, which results in a combinatorial problem. Spack overcomes this issue by providing a

framework to automate building software across many configurations. Currently Spack manages around 910 different packages for each of the different programming environments (Cray, Intel and GNU). Consultants are able to login and, with a simple command, install complex scientific software in a “push of a button” manner.

NERSC’s adoption of Spack has also initiated a collaboration with LLNL. Spack was initially developed at LLNL and much of the early development was focused on supporting the BlueGene- and Cluster-based systems in use there. NERSC and LLNL worked together to port Spack over to Cray machines. NERSC has also worked with staff at other sites, including ORNL, ANL and LANL to help them utilize Spack, especially on their Cray machines. By using a common platform like Spack, the sites are able to easily share recipes for building different common software. This provides a more consistent experience for users while reducing staff burden. Spack is also part of the Exascale Computing Project and is being used by development teams to deploy their software.

TOKIO FRAMEWORK SIMPLIFIES I/O PERFORMANCE ANALYSIS

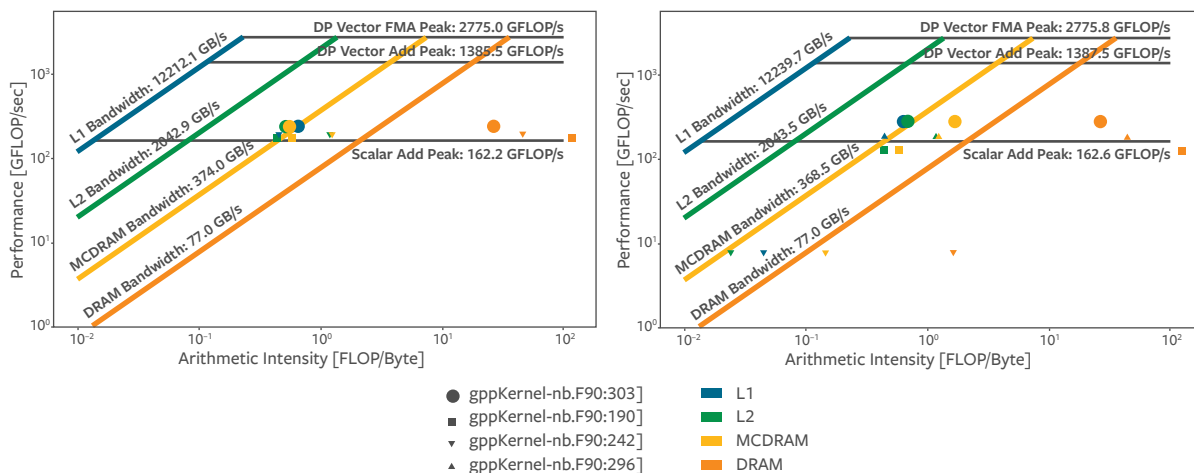
In 2017 NERSC published the Total Knowledge of I/O (TOKIO) framework, which simplifies the process of performing I/O performance analysis across the entire I/O subsystem. TOKIO connects to existing I/O monitoring infrastructures deployed at HPC facilities and provides the tools to combine these data sources to enable integrated, holistic performance analysis. NERSC staff demonstrated the utility of the TOKIO framework at ALCF and Berkeley Lab

for a paper that was presented at the 2nd Joint International Workshop on Parallel Data Storage & Data Intensive Scalable Computing Systems in November 2017. They showed statistically significant sources of I/O performance variation on Edison and Mira, ranging from bandwidth and metadata contention to file system health problems. NERSC is now working with early users to develop self-service interfaces for these multicomponent analyses.

HIERARCHICAL ROOFLINE MODEL DEVELOPMENT EXPANDS

The Roofline Model is a widely used performance model for characterizing and optimizing HPC codes. It was integrated into Intel Advisor as a feature through the collaborative work between NERSC and Intel in 2016, and in 2017 the Roofline

team further extended this model from measuring one level of cache to measuring a full hierarchy of caches and memory components.



Integrated Roofline Model for GPP kernel on KNL before (left) and after (right) cache blocking.

Intel then implemented Roofline with a cache simulator to automate the data-collection process and provide the multi-level roofline charts. This hierarchical Roofline Model has significantly increased the granularity of information that the model can provide and yields more insights into where the code’s performance bottlenecks are and where to target to optimize them.

Compared to the earlier work, the hierarchical Roofline Model not only helps identify a code as compute bound or memory bandwidth bound, but also pinpoints which level of cache/ memory is the limiter of the performance, thus guiding code developers to focus on those areas of their code. With more

detailed data movement captured, the hierarchical Roofline Model also shows the effect of a particular optimization, such as cache blocking.

An example code is the GPP (General Plasmon Pole approximation) kernel. The increase of MCDRAM arithmetic intensity (AI) for loop at line 303 clearly shows that the blocking has happened on the L2 cache level and not the other levels. In addition, the distance between the L2 AI and MCDRAM AI shows that there’s roughly 3x increase in the AI hence 3x data reuse. With the original Roofline Model, this would not have been identifiable.

MINIMIZING THE IMPACT OF CACHE THRASHING

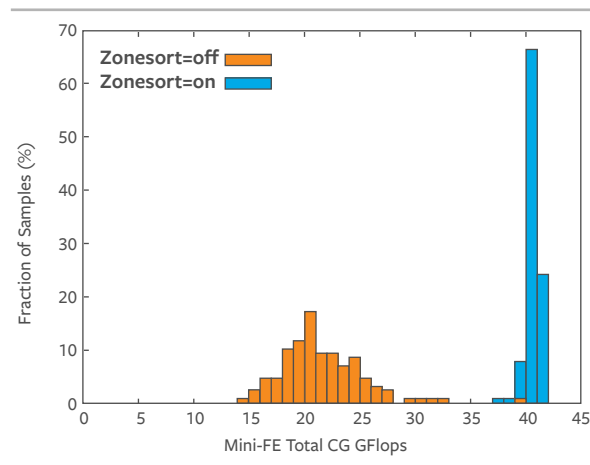
Cori’s KNL nodes have a two-level memory subsystem: the double data rate (DDR) memory layer provides relatively high capacity and modest bandwidth, while the multi-channel DRAM (MCDRAM) offers more limited capacity but high relative bandwidth. There are numerous ways to configure the MCDRAM, but NERSC has found that using it as a cache for the DDR is the easiest approach for the majority of application teams. Performance of the cache mode approaches — and in

many cases exceeds — what can be achieved by using MCDRAM as a separate, application-managed memory space. However, because the Intel Knights Landing’s MCDRAM is a direct mapped cache, it may incur higher rates of cache thrashing than more sophisticated associative caches.

To mitigate the effects of cache thrashing, Intel developed the “Zonesort” kernel module which reorders Linux’s free page

list to preferentially allocate blocks of memory that do not conflict with each other in the cache. NERSC then initiated and led a Center of Excellence between Intel, Cray, LANL and ALCF to evaluate Zonesort and investigate approaches to avoid cache thrashing on KNL. This effort demonstrated that the average performance of some applications is doubled and that runtime variation is significantly reduced by running Zonesort.

Zonesort is now incorporated into Cori’s workload manager and runs before every job. However, there is also evidence that the effectiveness of Zonesort diminishes with time until the node is rebooted. NERSC’s continuing efforts to research and understand the causes of this gradual diminution will benefit direct mapped cache performance and, possibly, low-associativity caches of future processors.



Impact of Zonesort on the Mini-FE mini-application performance distribution.

TOWARD BETTER QUEUE LOGISTICS AND INTERACTIVITY

One of the most important questions for many NERSC users is “How soon will my job run?” The queuing problem is extremely difficult, and we rely on complex heuristics and policies to maximize throughput. Sometimes the queue appears unfair, when a job submitted long after another job fits into backfill and runs first. To help users better understand how soon their job is likely to run, in 2017 NERSC consultants implemented a script called “sq.”

“The best invention since Nutella.”

The sq script was built on top of the Slurm squeue command. The initial version provided a priority ranking of jobs, plus a backfill priority ranking. But user feedback indicated that users were most interested in how soon their job would run,

not how it compared to other jobs. The second version, developed over the summer, now provides an estimate of when a job will run or, if the job is not yet eligible to be scheduled, the time/date when the job will be eligible to be scheduled. If a full-system maintenance reservation has been put in place, a warning at the bottom of the list indicates that the estimate may be inaccurate due to these circumstances.

In addition, in response to feedback from users needing better turnaround for debugging at scale, NERSC deployed a dedicated interactive queue. Made up of 192 Haswell nodes and 192 KNL nodes, this queue offers a resource for interactive debugging at large scale for up to four hours. Because this is an interactive resource, users can quickly debug new scripts (without having to wait for the next iteration to move through the queue) and can correct I/O problems that only show at larger scales. The interactive queue was used by roughly 600 unique users in 2017, and every Haswell node is frequently fully utilized during the day. NERSC has gotten overwhelming positive feedback on the interactive queue, with one user styling it as “the best invention since Nutella.”

SIMPLIFYING JOB SCRIPT CREATION FOR USERS

One of the top subjects for user tickets is “running jobs,” particularly as the complexity of the machines grows. For

example, Cori users now have to consider which type of node they will use, what mode (for the KNL nodes) and what type

of thread binding they want. There are many examples on the NERSC website to help guide users, but while these are sufficient for many users, others still struggle.

In an effort to make it easier for users to create their own job scripts, in 2017 NERSC created the Jobscrip Generator at

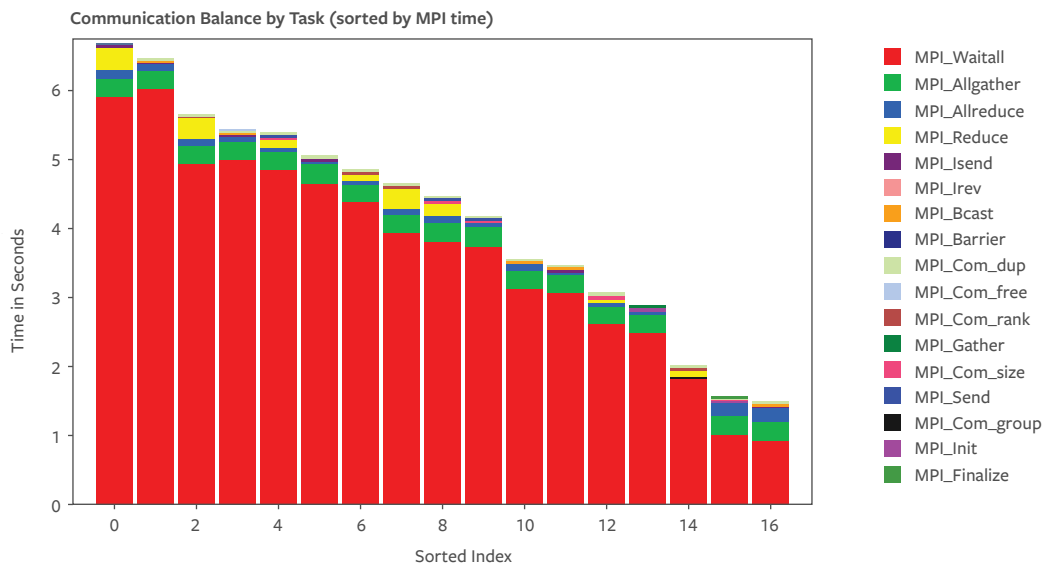
my.nersc.gov. The tool generates a batch script template with the correct KNL mode (if applicable) and process and thread-binding configurations. The tool was advertised in the NERSC weekly email when it first came out and has been widely adopted. In addition, an HPC center in Europe subsequently asked if NERSC would let them access the source code.

TAKING POWER MONITORING TO THE NEXT LEVEL

Manycore architectures are an energy-efficient step toward exascale computing within a constrained power budget. The Intel KNL manycore chip used in Cori is a specific example of this architecture. It is therefore important to understand the performance and energy usage characteristics of KNL. Toward this end, NERSC enhanced the Integrated Performance Monitoring (IPM) tool to evaluate energy usage of key applications.

With this new capability a study was performed that looked at the performance and energy efficiency of KNL in contrast to Cori's Xeon (Haswell) architecture for applications

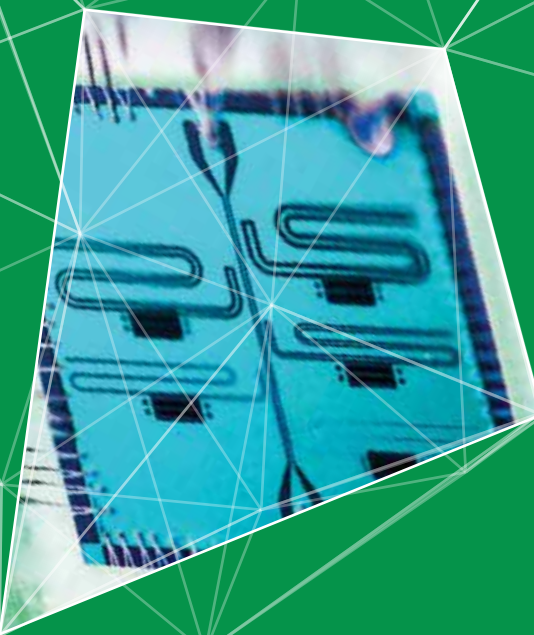
representative of the NERSC workload. The study looked at the optimal MPI/OpenMP configuration of each application and used the results to characterize KNL in contrast to Haswell. It also considered the differences in performance and energy usage when using the KNL's traditional DDR memory when compared to the on-chip MCDRAM "fast" memory. Results showed that on average the KNL is 1.8 times more energy efficient than a Haswell-based node and has 1.3 times greater performance. The new IPM power capability has been pushed upstream to the publicly available version found on GitHub.



NERSC has enhanced the Integrated Performance Monitoring tool to evaluate energy usage of key applications.

Engagement and Outreach

In 2017 NERSC supported 7,405 active users across 994 projects, with active users representing all 50 U.S. states (plus Washington D.C.) and 44 foreign countries. The number of projects and users was about 5% higher in 2017 than 2016, but this hides a more dynamic picture: about 30% of all projects ended after 2016 and were replaced by a slightly larger number of new projects.



NERSC has long been recognized for the depth and breadth of its user support.

A cross-group set of consultants from the User Engagement, Application Performance and Data Science Engagement groups provide the first line of support for NERC'S user community. These consultants, all of whom have Masters degrees and/or PhDs in science or computational science, are experts in HPC as well as various science domains. They are responsible for problem management and consulting, helping with user code optimization and debugging, strategic project support, web documentation and training, third-party applications and library support, running Office of Science-wide requirements reviews and coordinating the NERSC User Group. NERSC's consultants and account support staff are available to users via email and an online web interface 24x7, 365 days a year.

NERSC also regularly communicates with program managers from ASCR through weekly calls and face-to-face meetings at headquarters. The center and its staff also engage and collaborate with vendors and other HPC centers to support existing systems, develop and add new features and discuss technology roadmaps.

AN EXASCALE COMPUTING PROJECT UPDATE

Berkeley Lab is one of six major national laboratories playing a key role in the Department of Energy's Exascale Computing Project (ECP), which aims to deliver breakthrough modeling and simulation solutions that analyze more data in less time, providing insights and answers to the most critical U.S. challenges in scientific discovery, energy assurance, economic competitiveness and national security. In 2017, NERSC engaged with the ECP in numerous ways:



- **Access to NERSC Systems for Application**

- **Development and Software Technology Projects:**

NERSC provided 125 million additional NERSC Hours – above the nominal 10% of NERSC time – to the ALCC program to support ECP application development and software technology projects. One half of the time (62.5 million hours) was provided in NERSC allocation year 2017, and ECP teams used 54 million hours of that allocation. In addition, NERSC made time available to any ECP project on Cori Intel Xeon Phi nodes during “free time” in January through June 2017. Usage by ECP teams in 2017 totaled about 60 million NERSC Hours.

- **PathForward:** NERSC staff have been active in the PathForward projects with Intel, AMD, Nvidia, IBM and Cray/ARM/Cavium, acting as technical representatives and subject matter experts. NERSC and Berkeley Lab CS staff have also been involved in deep dive meetings held at Intel, AMD and Cray.

- **Software technologies:** NERSC identified packages of interest and ranked the relevance of the software development projects in the ECP Software Technologies area.

- **Training:** NERSC collaborated with ECP and the other ASCR facilities to present training classes. For example, “Python in HPC” was held in June with an emphasis on using Python for HPC at NERSC, ALCF and OLCF; an OpenMP tutorial was held in July; and Kokkos training took place in May.

- **Applications:** NERSC is collaborating with the ECP application assessment area to bring Roofline to other platforms. Roofline is a software toolkit developed at Berkeley Lab to better understand supercomputer performance and boost application performance. In addition, a significant number of ECP application teams are collaborating with NERSC staff to ready their codes for exascale through the NESAP program. These include: Lattice QCD, NWChemEx, WDMApp, WarpX, ExaStar, ExaSky, Subsurface, Exabiome, E3SM and ExaFEL. In addition, multiple NERSC staff contribute to the AMREX codesign center.

- **Multiple NERSC staff took on leadership positions within ECP:**

- **Jeff Broughton** — PathForward Technical Representative for Intel
- **Jack Deslippe** — L3 manager for Area 2.2.1 Materials Science and Chemistry Application Development
- **Thorsten Kurth** — PathForward point-of-contact for the Nvidia project
- **Brian Friesen** — Berkeley Lab Subject Matter Expert (SME) for the IBM PathForward project
- **Taylor Groves** — Berkeley Lab SME for Cray

- **NERSC staff attended the ECP RFI meeting in March.**

- **NERSC staff participated in the ECP 2017 All Hands Meeting** in Knoxville and contributed to discussions in a number of sessions, including a number related to facilities.

NESAP CONTINUES TO THRIVE

The NERSC Exascale Science Applications Program (NESAP) reached a culmination point in 2017 with the successful transition of Cori to production. The program hired a number of NESAP post-doctoral fellows, including Zahra Ronaghi (AMREX/Tomopy), Rahul Gayatri (SW4/Performance-Portability), Jonathan Madsen (TOAST) and Laurie Stephey (DESI). In addition, several postdocs graduated from the program, going on to find career employment within the field of computational science:

- **Mathieu Lobet** — La Maison de la Simulation (CEA) (France)
- **Tareq Malas** — Intel (Applications Engineer)
- **Taylor Barnes** — MOLSSI (NSF Funded Materials Software Institute)
- **Andrei Ovsyannikov** — Intel (Applications Engineer)

High-level NESAP activities in 2017 included:

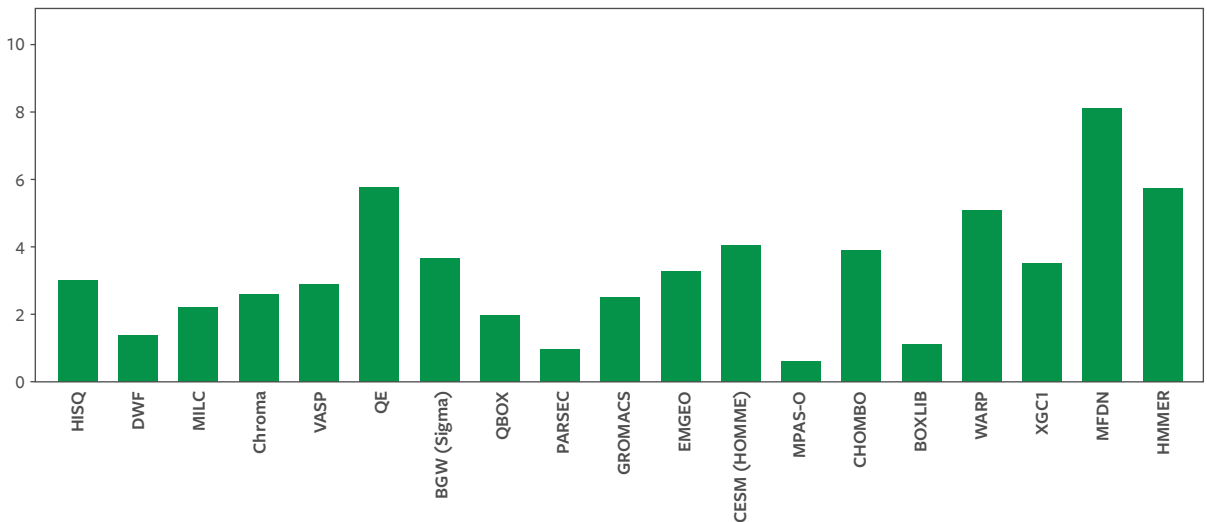
- Four application performance “dungeon sessions” (including one “super dungeon” and two targeting data applications) were held at Intel campuses.

- Three KNL performance hack-a-thons were held at NERSC, including two emphasizing the use of Intel’s Advisor Roofline tool.
- Eight general application performance and performance tool trainings were held at NERSC.

NERSC staff also collected a set of performance data from 18 NESAP applications before Cori went into production that include baseline (pre-NESAP) and optimized code performance on both Edison and Cori. The numbers show an average speedup (across an equivalent number of nodes) of approximately 3x between baseline codes on Edison and NESAP optimized codes on Cori. In addition, optimization made to the codes targeting the Cori KNL system, on average, sped up performance on Edison by an average of approximately 2x — a significant boost in scientific throughput.

NERSC has been distilling the lessons learned from the NESAP program to the broader NERSC community through extensive training, publications and a large amount of online documentation and case studies.

NESAP Optimized Code Performance versus Edison Baseline Performance (Node per Node)

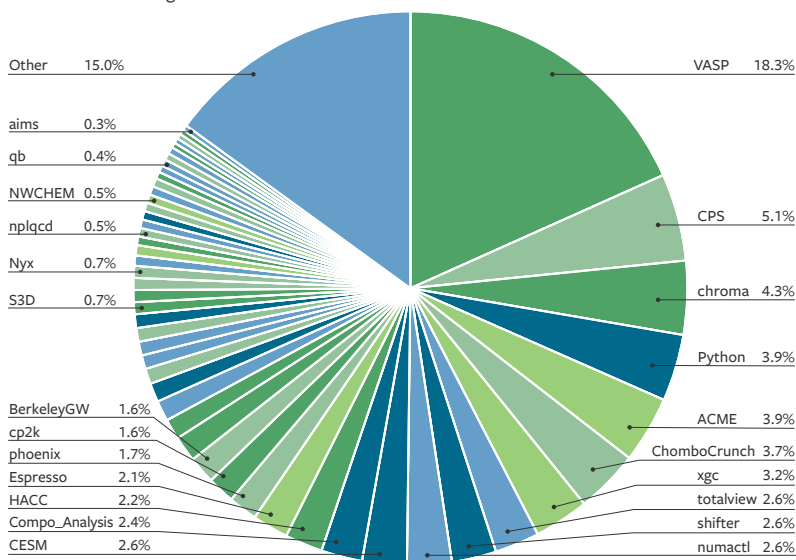


PROJECTING THE NERSC WORKLOAD PERFORMANCE ON GPUS IN 2020

In preparation for NERSC-9, NERSC’s next-generation supercomputer due in 2020, a study was performed to characterize the readiness of the NERSC application base for GPU-based accelerators. The study included a survey of codes to determine workload coverage in addition to literature searches and empirical studies of key codes on current GPU-based platforms. The results showed that 46% of the codes had at least some kernels that were accelerated already; 19% of the codes had a proxy application that was algorithmically similar; 11% were determined not to have a GPU port and it was likely that a GPU port would require a significant effort; and 24% were unclassified as their readiness cannot be assessed at this time.

The study also looked at a representative set of GPU-ready applications, benchmarked on CPU-accelerated platform at NVIDIA, the Swiss National Supercomputer Centre’s Piz Daint and the OLCF’s Summitdev. The results showed that for well-optimized codes, single-node GPU-accelerated performance compared to a Cori node is typically 4 to 6 times greater for a compute-bound code and 2 to 4 times greater for a memory-bound code. However, at scale these speedups decrease depending on the respective communication characteristics of the application. From this study, NERSC concluded that it is feasible to consider GPUs in the NERSC-9 architecture.

GPU Readiness among NERSC Codes — Breakdown of Hours at NERSC



GPU Status	Description	Fraction
Enabled	Most features are ported and performant.	32%
Kernels	Ports of some kernels have been documented.	10%
Proxy	Kernels in related codes have been ported.	19%
Unlikely	A GPU port would require major effort.	14%
Unknown	GPU readiness cannot be assessed at this time.	25%

NERSC DATA ADVISORY COMMITTEE LOOKS TO THE FUTURE

The Exascale Requirements Reviews conducted with all six offices highlighted a growing need for DOE HPC facilities and DOE experimental user facilities to co-evolve capabilities to support the growing workload, datasets and analysis requirements from DOE experimental facilities. In addition, the number of NERSC users dealing with experimental and observational data in some way has grown to 40%, according to NERSC’s 2017 user survey.

These trends prompted the center to convene a Data Advisory Committee to evaluate its data strategy and receive feedback from stakeholders in the community. In a series of meetings held in the Fall of 2017, NERSC staff developed presentations on key topics related to its data strategy and the superfacility initiative, including implementation strategies for systems, software and user engagement.

“NERSC is a great collaborator.”

On October 4, NERSC hosted the Data Advisory Committee for a review of the center’s resulting data strategy vision: to transform science by enabling seamless, large-scale data pipelines and analysis on leading-edge HPC systems and platforms. The committee comprised Katrin Heitmann, Argonne; Gil Compo, University of Colorado; Loren Frank, University of California, San Francisco; Dula Parkinson,

Berkeley Lab; Tom Abel, SLAC; Sudarshan Vashkudai, Oak Ridge National Laboratory; David Schissel, General Atomics; and Peter Nugent, Berkeley Lab.

The committee issued a closeout report on October 12. Among their comments: “During the presentations, NERSC showed extraordinary awareness of the user issues and concerns around data. The goal of a ‘superfacility’ built from common APIs is ambitious and commendable. The identification of the building blocks seems well thought out and organically evolved from engagement with the experimental and observational facilities and collaborations. ... We applaud NERSC for ... taking the initiative to make a plan and to convene this Data Advisory Committee.”

NERSC/INTEL/UNIVERSITIES LAUNCH ‘BIG DATA CENTER’

A collaboration between NERSC, Intel and five Intel Parallel Computing Centers (IPCCs) resulted in the establishment of a Big Data Center (BDC) that is working on code modernization and tackling data-intensive science challenges. The five IPCCs who are part of the BDC program are the University of California-Berkeley, University of California-Davis, New York University (NYU), Oxford University and the University of Liverpool.

The goal of the BDC is to solve DOE’s leading data-intensive science problems at scale on the NERSC Cori supercomputer. The BDC, in collaboration with Intel and the IPCCs, is testing to see if the current HPC systems can support data-intensive workloads that require analysis of more than 100 terabyte datasets on 100,000 cores or greater. The BDC will optimize and scale the production data analytics and management stack on Cori. All the optimizations done at the BDC will be open source and made available to peer HPC centers as well as the broader HPC and data analytics communities.

To date, the Intel/NERSC BDC Collaboration has enhanced the performance of multiple deep learning frameworks (TensorFlow and PyTorch) by over 10x; deployed the latest optimized builds on Cori; and updated the NERSC

user-facing documentation (<http://bit.ly/NERSCdeeplearning>) with instructions on incorporating deep learning in their workflows. The collaboration has also been tracking the performance of Python builds on Cori.

In December 2017 Cray Inc. also joined the BDC to help advance the adoption of artificial intelligence (AI), deep learning and data-intensive computing. The collaboration is now focusing on three fundamental areas:

- **Advancing the state-of-the-art in scalable, deep learning training algorithms**, which is critical to the ability to train models as quickly as possible in an environment of ever-increasing data sizes and complexity.
- **Developing a framework for automated hyper-parameter tuning**, which provides optimized training of deep learning models and maximizes a model’s predictive accuracy.
- **Exploring the use of deep-learning techniques and applications against a diverse set of important scientific use cases**, such as genomics and climate change, which broadens the range of scientific disciplines where advanced AI can have an impact.

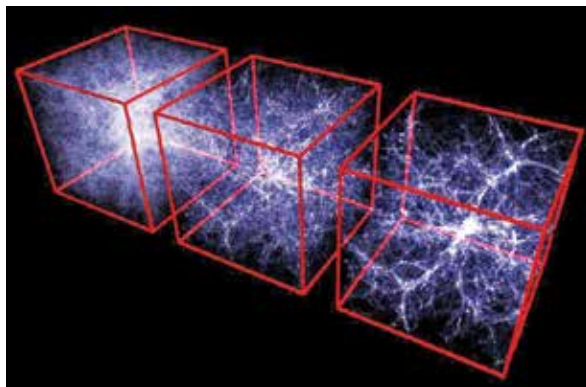
SCALING UP TO PETAFLOP DATA ANALYTICS

In 2017, NERSC made major advances in enabling data applications to run at scale on Cori/KNL. Three exemplar projects in this space were Celeste, Galactos and Deep Learning. All three projects are the first petaflop-class applications in the Data Analytics space at NERSC.

- **The Celeste project** broke new ground along a number of fronts: this is the first application in native Julia to exceed 1PF, the team processed 55TB SDSS data in 15 minutes on Cori/KNL utilizing the Burst Buffer, this is the largest demonstration of stochastic variational inference in the research community and the resulting list of 188M stars and galaxies is the first comprehensive catalog of objects with point and uncertainty estimates. The project comprised staff from NERSC, Berkeley Lab Physics, UC Berkeley, Intel, MIT and Julia Computing.
- **The Galactos project** successfully solved the open problem of computing the 3-pt correlation function for galaxies. The codebase was implemented in C/C++/MPI and demonstrated excellent strong and weak scaling up to 9,600 nodes. The code obtained 9.8PF peak performance at scale on Cori, and processed 2 billion galaxies in under 20 minutes, a new record for the field of cosmology. The project comprised staff from NERSC, Berkeley Lab Physics, Princeton and Intel.
- **The Deep Learning project** was successful in scaling climate science and high-energy pattern recognition applications, implemented in the Caffe framework to an unprecedented 15PF performance level on Cori/KNL. To date, this is the largest demonstration of a deep learning

application on an HPC system. This project comprised staff from NERSC, Intel, Stanford and the Montreal Institute for Learning Algorithms.

These applications are indicative of a new breed of data technologies (Caffe, Julia and Python) that are now being optimized and scaled on HPC systems. In collaboration with industry and academic partners and the Big Data Center, NERSC continues to spearhead innovation in this important area going forward.



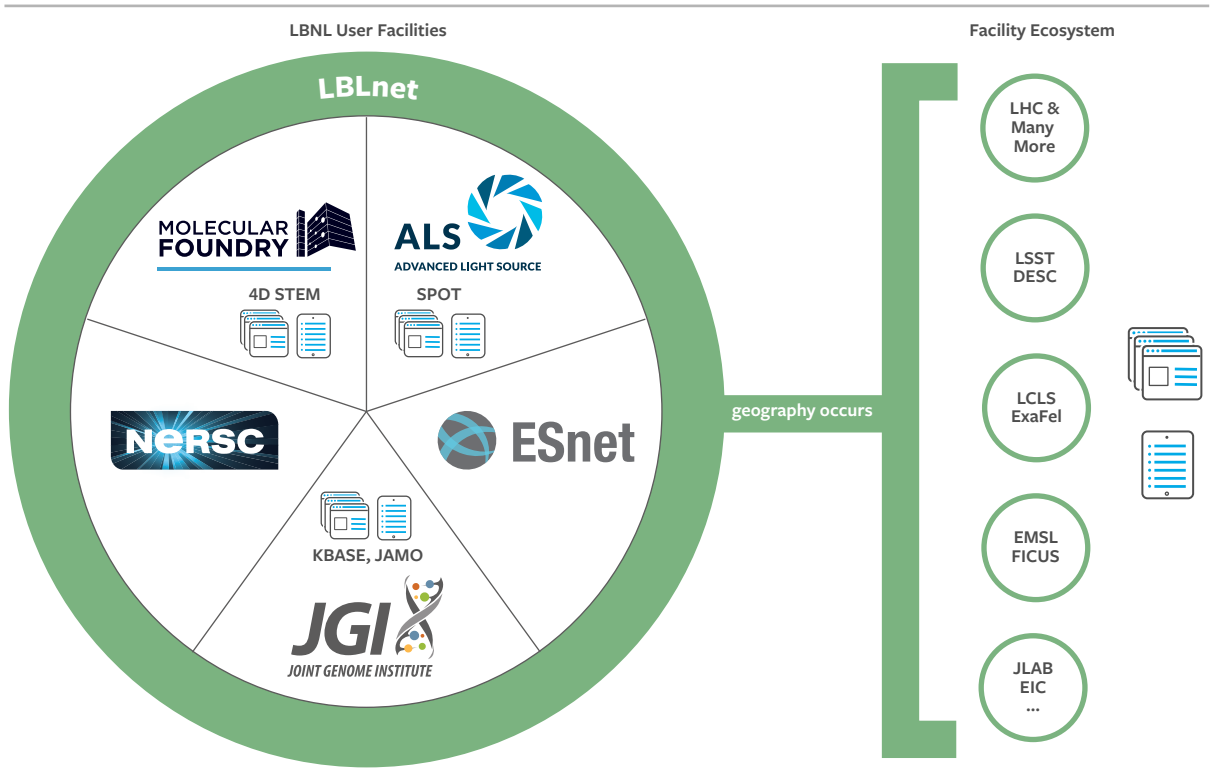
The Galactos team performed an entire 3-point correlation function computation on two billion galaxies in less than 20 minutes on NERSC's Cori system. This image shows that, over time, the attractive force of gravity and the expansive force of dark energy create a web-like structure of matter in the universe.

NERSC TAKES SUPERFACILITY CONCEPT TO NEW HEIGHTS

With scientific datasets growing in size and complexity, experimental facilities are increasingly looking to partner with NERSC for their data analysis requirements. Known as a superfacility or inter-facility, experiments and facilities that combine HPC with instruments such as microscopes, telescopes and particle accelerators are becoming more common as science teams find opportunities to accelerate their workflows. As the mission computing center for the DOE Office of Science, NERSC is a hub for these projects. A 2015 DOE Office of Science report (<http://bit.ly/NERSCuserfacilities2015>) surveying the national user facility

complex found more crossover users (scientists that use more than one facility) at NERSC than any other facility.

As the concept of superfacilities gains traction, NERSC has evolved customized and architected systems and software to better support data analysis workloads, such as through the use of Spin and software defined networking (SDN). It is worth noting the geographic benefits of the facility cluster at Berkeley Lab, which provides both collaborative wins and cost savings. For example, the National Center for Electron Microscopy (NCEM) at Berkeley Lab is able to connect its



The inter-facilities ecosystem — also known as a “superfacility.”

new 4D STEM detector to Cori at 400 Gb/s using an inexpensive local link via LBLnet. Remote instruments connect to NERSC via ESnet; the Linac Coherent Light Source (LCLS) at SLAC, for example, is connected at 200 Gb/s.

“NERSC provides a liaison to interact with high impact science at scale.”

In addition, NERSC’s Cori system has been specially upgraded in terms of WAN connectivity to better connect with the facility ecosystem, with innovations that include adding more gateway nodes and some early work with SDN. New

capabilities, including the ability to stream data directly from an external network connection to compute nodes and the burst buffer, are being provided to inter-facility science teams.

Coordinating experiments that require multiple facilities in real time remains a challenge, however. NERSC’s queue reservation system and real-time queues support many projects’ need to apply HPC capabilities at certain well-scheduled beam times. During experiment shifts usually lasting eight hours, for example, beamline scientists can use Cori to speed their analysis, but this requires co-scheduling and plans to adapt in the case of an unscheduled outage. Some users have adopted resilience planning to throttle or suspend data analysis in case of unplanned interruption. Future resilience plans include adding additional HPC centers as failover options and the use of killable queues to gain access to many nodes quickly.

Here are two examples of NERSC's partnerships with experimental facilities in 2017 that highlight the center's superfacility capabilities:

Using HPC to Innovate Electron Microscopy Data

Acquisition: Electron microscopy (EM) at NCEM involves a rapidly evolving technology track in detectors. Innovation in EM detectors led by Peter Denes at Berkeley Lab has led to a collaboration with NERSC whereby the next-generation data acquisition system (DAQ) plugs directly into Cori. NERSC and NCEM have tested data transfer from the 4D-STEM DAQ through 48 ports of the new switch to four gateway nodes on the Cori system. As this first-of-its-kind detector is being operated, NCEM will use computing power at NERSC to design algorithms for data analysis. These algorithms in some cases will ultimately be ported back to the DAQ to reduce and annotate data near the device. Computational scalability and a robust software environment available through NERSC allow this microscope development team to

get started early with their data analysis from this new detector. It also informs the long-term possibilities for EM data-analysis architectures.

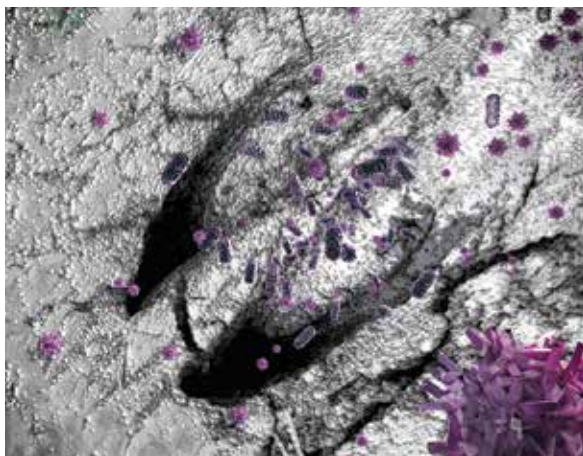
Exascale Data Analytics: In 2016 the ExaFEL project — a free electron laser (FEL) science collaboration between SLAC, Berkeley Lab and LANL — was launched. This project is an evolution of work demoed at SC in 2014 and is now porting three FEL science applications to exascale-era architectures. New detectors with higher data rates, such as the LCLS-II experiment at SLAC, are coming online in the 2020 timeframe, and new tools and capabilities need to be developed in order to analyze XFEL data. Data analysis in nano-crystallography, single particle imaging and diffuse x-ray scattering algorithms are being developed and improved so that they will be when ready when the LCLS-II experiment comes online. The applications will provide useful real-time feedback to users running experiments on the LCLS-II beamlines at SLAC.

NERSC/JGI COLLABORATION: THE FICUS PROJECT

In 2017 six proposals were selected to participate in a new partnership between two DOE user facilities through the “Facilities Integrating Collaborations for User Science” (FICUS) initiative. The expertise and capabilities available at the DOE Joint Genome Institute (JGI) and NERSC will help researchers explore the wealth of genomic and metagenomic data generated worldwide through access to supercomputing resources and computational science experts to accelerate discoveries.

Through this project, users can use NERSC's Cori system to query across all available data to look for patterns across data sets in the DOE JGI's Integrated Microbial Genomes and Microbiomes database. As many of these researchers are new to computing, a member of NERSC's Data Science Engagement Team has been assigned to work with each FICUS project, and DOE JGI staff will assess their needs and help them develop tools and workflows. Ultimately, these tools and scientific findings will be made publicly available via a NERSC science gateway.

The JGI-NERSC FICUS project is the latest partnership since the collaborative science initiative was formed in 2014 by the Office of Biological and Environmental Research.



The JGI-NERSC FICUS project allows users to query across all available data to look for patterns across data sets in the DOE JGI's Integrated Microbial Genomes and Microbiomes database. Shown here: a rumen microbiome.

OPENMP STANDARDS UPDATE ENHANCES RUNTIMES

In 2017, two Berkeley Lab Computing Sciences staff members — Helen He, of NERSC, and Alice Koniges, of the Computational Research Division — were instrumental in getting critical new runtime features added to the OpenMP Technical Report (TR) 6 standard document, the precursor document for the official OpenMP 5.0 release slated for SC18.

OpenMP is an application programming interface (API) that supports multi-platform, shared-memory, multiprocessing programming in C, C++ and Fortran. The most popular node-level parallelization approach used at NERSC, OpenMP comprises a set of compiler directives, library routines and environment variables that influence run-time behavior.

The additions to the 600+ page TR6 document that He and Koniges wrote and shepherded through the OpenMP Architecture Review Board's detailed approval process for the OpenMP 5.0 standard include two OpenMP runtime

environment variables (with their associated integrity check values) and four runtime APIs. These new features enable users to control, collect and verify runtime thread affinity information, which is critical to ensuring optimal performance on any system and is an essential step before starting any code optimization attempts.

Previously, the ability to report and take action on OpenMP thread binding to compute cores had been lacking in the standard. In addition, if it was attainable at all, thread affinity information was either compiler-dependent or required extra libraries or tools to be loaded and used. With the new standard, each OpenMP compiler will provide this information with a uniform user interface. Compiler vendors have started implementing the new compiler/runtime features to fully test the syntax in real compilers. Once finalized and adopted into OpenMP 5.0, the new features will be included in future releases of all compilers that support OpenMP.

NERSC HOSTS 2017 SLURM MEETING

NERSC hosted the 2017 Slurm User Group (SLUG) meeting September 24–26 at Berkeley Lab — the largest SLUG meeting to date, with 85 attendees from the U.S., Europe and Australia.

Slurm (Simple Linux Utility for Research Management) is an open-source workload management and scheduling system

developed and supported primarily by SchedMD, which was founded in 2010 to create and provide services around Slurm.

Initially developed for large Linux clusters at Lawrence Livermore National Laboratory, Slurm is currently used on NERSC's Cori and Edison systems and is increasingly being adopted by supercomputing facilities around the world. “The



NERSC hosted the 2017 Slurm User Group (SLUG) meeting at Berkeley Lab — the largest SLUG meeting to date, with 85 attendees from the U.S., Europe and Australia.

rapid adoption of SchedMD's Slurm resource manager is due in part to the vibrant and expanding Slurm User Group," said David Paul of NERSC's Computational Systems Group, who helped organize the 2017 SLUG meeting, along with NERSC staffers Jackie Scoggins, Kerry Peyovich, Basil Lali and Tony Quan. "Not only is the number of attendees growing, the sites and companies represented is expanding to include prominent labs, universities and research centers from around the world."

This year's attendees came from major universities, including U.C. Berkeley, Stanford, MIT, Harvard, University of Michigan, BYU and NYU; national labs, including Lawrence Livermore, Sandia and Oak Ridge; and industry, including Novartis, Genentech, Cray and Roche. Representatives from several major European computing centers also attended, including CSCS (Switzerland), FZ-Juelich (Germany), AWE (England), CEA France and Bull/Atos (France).

OTHER NERSC TRAININGS, WORKSHOPS AND HACKATHONS

NERSC held 14 trainings in 2017, aimed at different segments of the center's user population.

Because computational chemistry and materials science users comprise a large portion of the center's user population, NERSC hosted trainings on the BerkeleyGW and VASP applications. Combined with the five-day VASP workshop at the end of 2016, the VASP workshops have resulted in the majority of VASP users adopting the new hybrid VASP code that runs efficiently on KNL nodes, and in 2017 VASP accounted for more than 15% of the CPU hours on Cori's KNL nodes.

In conjunction with the NERSC User Group annual meeting, NERSC held its Data Day event in September. The event consisted of talks and tutorials on topics such as the burst buffer, machine learning, Shifter, Python and data management, followed by a data hackathon and data

challenge competition. Competition participants chose one of two problems to solve: analyze the NERSC Slurm batch system output, looking for any insights from the history of the jobs run on Edison and Cori; or analyze astronomy catalogs, differentiating between the images of galaxies and stars. The winning team on the data challenge provided insights into what types of jobs have the best throughput on NERSC resources and the best time and day of the week to submit a job to minimize queue wait time.

NERSC staff also gave trainings to the recipients of the DOE Computational Science Graduate Fellowship and to students during the 2017 Richard Tapia Celebration of Diversity in Computing. The annual Tapia meeting brings together some 1,200 undergraduate and graduate students, faculty, researchers and professionals in computing from diverse backgrounds and ethnicities to learn from thought leaders, present innovative ideas and network with peers.



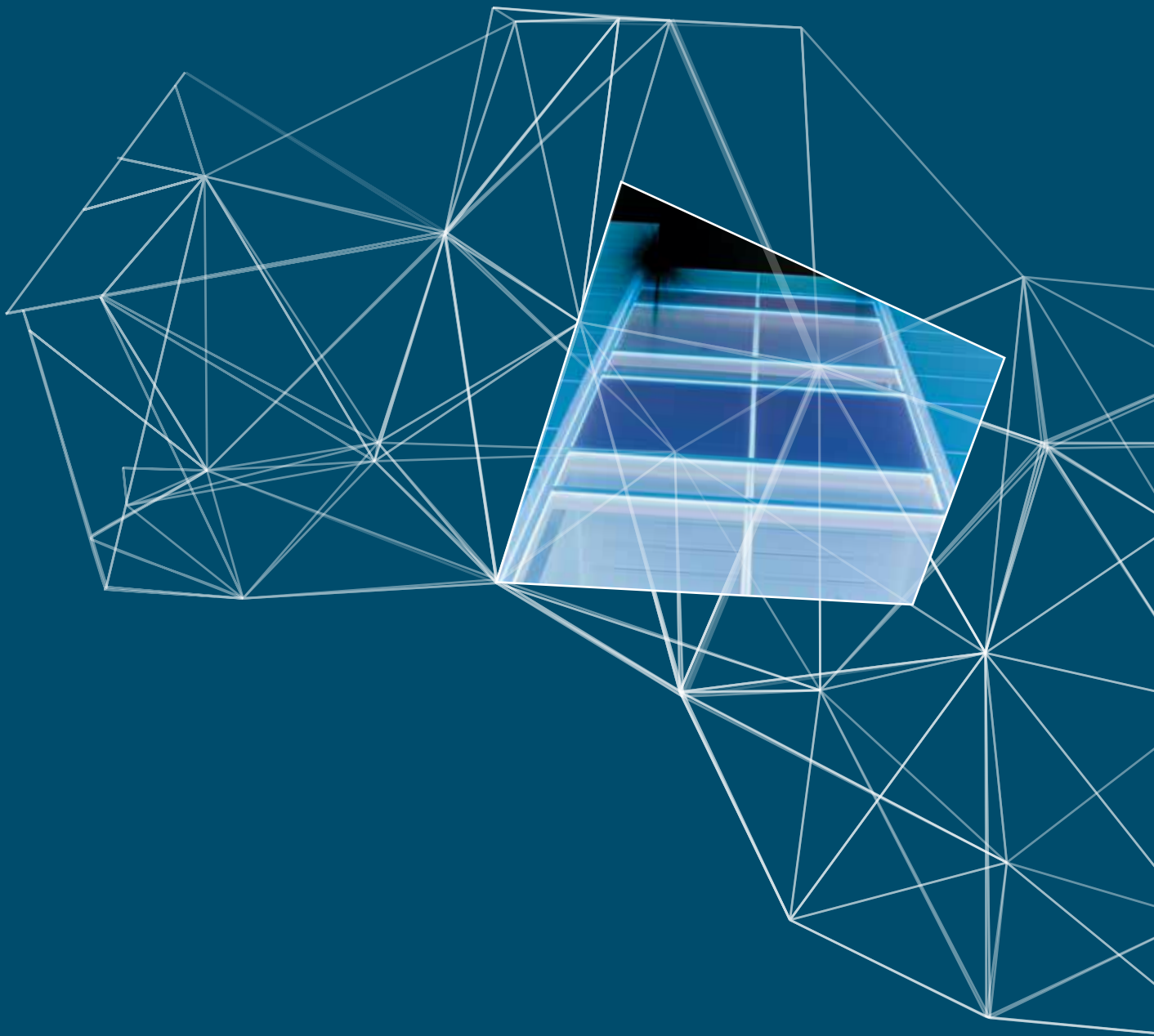
NERSC staff led a hands-on KNL workshop during HPC Day at the 2017 Tapia conference.

Event	Date(s)	Brief Summary
BerkeleyGW2017	January 4–6	Workshop on ab initio many-electron effects calculations using BerkeleyGW.
VASP KNL Training	January 19	How to get started running the hybrid VASP program on KNL nodes
Cori KNL Programming and Optimization Workshop	February 14–16	Cray presents on Cori KNL programming environment, debugging and optimization.
Roofline Training	February 16–17	Hands-on workshop using the roofline model and new roofline features in Intel Advisor to profile and optimize your code.
New User Training	February 23–24	Training targeted at novice NERSC users to help them navigate the center and its systems.
Intel Compilers, Tools, and Libraries Training	March 9–10	Training on Intel compilers, tools, and libraries, presented by experts from Intel.
Cori KNL Training	June 9	Introducing Cori KNL to all users and an overview of how to run successfully on the KNL architecture
Scaling to Petascale Institute	June 26–30	NERSC was a co-organizer of this event, which reached more than 300 attendees across the US, Canada, Costa Rica, and Brazil.
Debugging and Profiling Party with Allinea Tools	June 28	Hands-on training for using Allinea tools to debug and profile codes
CSGF HPC Workshop 2017	July 26	Training for DOE Computational Science Graduate Fellows on optimizing for KNL and the burst buffer.
Data Day/NUG Meeting User Training	September 19–20	Machine learning, data analytics and data movement training and hackathon held in conjunction with NERSC User Group annual meeting.
Tapia HPC Workshop 2017	September 23	Training for attendees of the Tapia Celebration of Diversity in Computing for using Cori, burst buffer, code optimization, and an introduction to OpenMP
Cori KNL Hands-On Hack-A-Thon	September 27	Hands-on session for users to build, evaluate, and improve their applications and workflows on Cori's KNL nodes.
Roofline Training	November 8–9	Hands-on training for advanced users to optimize their code using the roofline model and new roofline features in Intel Advisor.

In addition, NERSC was a co-organizer of the Scaling to Petascale Institute, along with OLCF, NCSA and TACC. For this event NERSC hosted a training site, contributed training on code optimization and I/O and distributed 400

NERSC training guest accounts to the participants, who were located at sites across the United States, Canada, Costa Rica and Brazil.

Center News



SHYH WANG HALL ACHIEVES LEED CERTIFICATION

Berkeley Lab's Shyh Wang Hall — home to the Computing Sciences organization and NERSC's supercomputing resources — earned a Gold LEED (Leadership in Energy and Environmental Design) certification from the U.S. Green Building Council (USGBC) for its environmental and energy-efficient design.

LEED is a prominent green building certification that rates a building's sustainability aspects. It was implemented as a way to evaluate the environmental performance of a building and to encourage market transformation toward sustainable design. Points are awarded across a variety of categories, including use of renewable energy, water efficiency and innovation in design.

Beyond the basic LEED certification (40-49 points) are the Silver (50-59 points), Gold (60-79) and Platinum (80 points

and above) levels. Wang Hall achieved a total of 69 points, scoring particularly well in Sustainable Sites, Water Efficiency, Indoor Environmental Quality and Innovation in Design.

“With this certification, the USGBC has recognized the unique and sustainable features of this beautiful, high performance computing facility,” said Berkeley Lab's Sheree Swanson, Shyh Wang Hall project director. “Energy efficiency and innovation in design were top priorities throughout the design and construction.” Examples of features that contributed to the Gold certification include innovative cooling that eliminated the need for conventional chillers and maximized the use of outside air, large hydro-modification tanks underneath the facility that mitigate storm water runoff impact and a low-emissivity roof to reduce radiant thermal energy.

Shyh Wang Hall, home to NERSC, earned Gold LEED certification for its environmental and energy-efficient design.



CALIFORNIA SENATOR TOURS NERSC

In September, California State Senator Henry Stern (D-Canoga Park) — who successfully pushed to increase funding for clean energy research from the state’s Greenhouse Gas Reduction Fund — met with Berkeley Lab Director Mike

Witherell and researchers from the Earth and Environmental Sciences, Energy Technologies and Biosciences Areas. He then toured NERSC and attended Cyclotron Road’s annual innovator showcase.

NERSC HONORS FOUR USERS WITH 2017 HPC ACHIEVEMENT AWARDS

NERSC announced the recipients of the 2017 High Performance Computing (HPC) Achievement Awards during the annual NERSC Users Group meeting in September. 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 features two early career awards.

2017 NERSC AWARD FOR HIGH-IMPACT SCIENTIFIC ACHIEVEMENT

A team of researchers from Lawrence Berkeley National Laboratory, University of California Berkeley and Caltech was honored in this category for using NERSC resources to speed up the discovery of commercially viable materials that can be used to produce solar fuels. The group gathered a list of potentially useful compounds and then were able to rapidly screen and test the best materials with NERSC. This process would normally take an immense amount of time to conduct all the tests and experiments by hand. The researchers were able to go through 174 compounds containing vanadium and oxygen, called vanadates, and identify 12 useful materials. These materials will be very useful for developing solar fuels, which are a clean and renewable alternative to fossil fuels. The work was led by Berkeley Lab’s Jeff Neaton, John Gregoire and Qimin Yan. Other members of the team were Jie Yu, Santosh Suram, Lan Zhou, Aniketa Shinde, Paul Newhouse, Wei Chen, Guo Li and Kristin A. Persson.

2017 NERSC AWARD FOR HIGH-IMPACT SCIENTIFIC ACHIEVEMENT — EARLY CAREER

Badri Narayanan, a scientist from Argonne National Laboratory, was honored for developing atomistic models to understand reactive interfaces of energy applications. He ran simulations on NERSC’s Edison system to study how diamond-like carbon tribofilms work on the molecular scale. With these simulations he was able to discover that the metal



The 2017 HPC Achievement Award winners, from left: Jae-Seung Yeom, Lawrence Livermore National Laboratory; Abhinav Bhatele, Lawrence Livermore National Laboratory; Thomas Heller, Friedrich Alexander University Nuremberg; Badri Narayanan, Argonne National Laboratory. Also pictured are NERSC’s Richard Gerber and Sudip Dosanjh, director.

catalysts in the nanocomposite coatings broke down the hydrocarbon chains into smaller ones and then rearranged them to form the tribofilm. His findings helped open up a new field called “tribocatalysis” that could revolutionize the study of lubricants.

2017 NERSC AWARD FOR INNOVATIVE USE OF HPC

A team led by Abhinav Bhatele of Lawrence Livermore National Laboratory was honored for using NERSC resources to determine the scaling for a realistic simulation of an infectious disease on the national level. To predict how much processing power these simulations would take, the team ran experiments on various supercomputers, including NERSC’s Cori system, using the EpiSimdemics simulator. This study will help future scientists predict and contain the spread of infectious diseases on the national scale. The other team members were Jae-Seung Yeom, Nikhil Jain, Chris Kuhlman, Yarden Livnat, Keith Bisset, Laxmikant Kale and Madhav Marathe.

2017 NERSC AWARD FOR INNOVATIVE USE OF HPC — EARLY CAREER

Thomas Heller of the Friedrich Alexander University Nuremberg was honored in this category for his work with simulations of merging stars — notably “for demonstrating that an asynchronous massively parallel tasking runtime system can be used to harness billions of tasks for a scalable hydrodynamics simulation of the merger of two stars.”

He was nominated for using OctoTiger, an adaptive mesh refinement program based on HPX (an asynchronous, massively parallel tasking runtime system) to perform full system runs on NERSC’s Cori system. He proved that these are powerful programs that could be used to assist large-scale tasks such as gathering information from a simulation of merging stars.

HIGH-IMPACT SCIENCE AT SCALE

NERSC’s High-Impact Science at Scale program provides competitively selected projects an allocation of time to use Cori’s unique capabilities at scale to investigate key science problems that they would not otherwise be able to address. Teams were selected following a Call for Proposals, in which responses were reviewed for the ability to scale well on Cori KNL nodes, the potential to deliver a significant science result, a reasonable ratio of feasibility to risk and the appropriateness of approach and resources requested.

In 2017, seven awards were given, each with enough time to make significant progress on the proposed research problem. Highlights include:

- A simulation of the intergalactic medium in the largest volume so far.
- Modeling the break up of a collisional current sheet in plasma conditions similar to a solar flare.
- Emittance degradation observed in a simulation of a laser wakefield accelerator leading to a proposed solution.

2017 HIGH-IMPACT SCIENCE AT SCALE PROJECTS

The following projects were awarded time through the program, which allocated a total of about 450 million NERSC Hours.

PI Name	Organization	Project Title	NERSC Hours Used	Nodes (Cores) Used
Trebotich, David	Berkeley Lab	Chombo-Crunch: Extreme scale simulation of flow and transport in heterogeneous media	143,056,732	9,664 (657,152)
Heitmann, Katrin	Argonne Lab	Knowhere (Cosmology)	92,391,956	6,144 (417,152)
Stanier, Adam	Los Alamos Lab	Probing the physics of magnetic reconnection – from fusion energy to space plasmas	64,170,260	2,048 (139,264)
Louie, Steven	UC Berkeley	Ab initio quasiparticle and optical properties of materials at scale	56,952,502	4,096 (278,528)
Lukic, Zarija	Berkeley Lab	Physical model of the intergalactic medium	51,167,795	8,192 (557,056)
Karsch, Frithjof	Brookhaven Lab	Net strangeness and net electric charge fluctuations in strongly interacting matter	41,614,572	9,178 (624,104)
Vay, Jean-Luc	Berkeley Lab	High-resolution 3D studies of asymmetric effects in the BELLA plasma accelerator experiments	27,886,930	1,274 (86,632)

- Insight into the properties of strongly interacting matter at the time where quarks and gluons — the fundamental degrees of freedom in strong interaction physics — begin forming bound states and give rise to ordinary matter such as protons and neutrons that form the nuclei of all atoms.
- One of the largest GW calculations ever performed: to predict the defect energy states in silicon carbide, which is of interest for the design and development of quantum computers. The team accurately determined energy values, alignments and defect orbitals.

NERSC SHARES HPCWIRE AWARD FOR HPC4MFG PROJECT

A unique collaboration between Berkeley Lab, NERSC, Lawrence Livermore National Laboratory and an industry consortium was honored in 2017 with an HPCwire Editor's Choice Award for using high performance computing to help U.S. paper manufacturers find new ways to reduce production costs and increase energy efficiencies.



Tabor Communications CEO Tom Tabor (far right) presented the HPCwire Editor's Choice award for Best Use of HPC in Manufacturing to, from left: LLNL Deputy Associate Director for Science and Technology in Computation Lori Diachin, LLNL's Associate Director for Computation Bruce Hendrickson, Lawrence Berkeley National Laboratory Department Head for Computational Science Peter Nugent and HPC Department Head for NERSC Richard Gerber.

The project is part of the DOE's HPC for Manufacturing (HPC4Mfg) initiative, a multi-lab effort to use high performance computing to address complex challenges in U.S. manufacturing. Through HPC4Mfg, Berkeley Lab and LLNL are partnering with the Agenda 2020 Technology Alliance, a group of paper manufacturing companies that has a roadmap to reduce their energy use by 20 percent by 2020.

The project combined the researchers' advanced simulation capabilities and high performance computing resources with industry paper press data to help develop integrated models to accurately simulate the water papering process. The idea was to identify ways to increase the paper dryness after pressing and before the drying stage, which would help the paper industry to save energy.

"This was true 'HPC for manufacturing,'" said David Trebotich, a computational scientist in Berkeley Lab's Computational Research Division and co-PI on the project. "We used 50,000-60,000 cores at NERSC to do these simulations. It's one thing to take a research code and tune it for a specific application, but it's another to make it effective for industry purposes. Through this project we have been able to help engineering-scale models be more accurate by informing better parameterizations from micro-scale data."

BERKELEY LAB-LED COLLABORATIONS EARN HPC INNOVATION AWARDS

Two Berkeley Lab-led projects — Celeste and Galactos — were honored with Hyperion Research's 2017 HPC Innovation Excellence Award for "the outstanding application of HPC for business and scientific achievements." According to Hyperion, the awards are designed to showcase return on investment and success stories involving HPC; to help other users better understand the benefits of adopting HPC; and to help justify HPC investments, including for small and medium-size enterprises.

CELESTE: A NEW MODEL FOR CATALOGING THE UNIVERSE: This research collaboration of astrophysicists, statisticians and computer scientists from UC Berkeley, Berkeley Lab, MIT, Julia Computing and NERSC developed Celeste, a statistical analysis model designed to dramatically speed up one of modern astronomy's most time-tested tools: Sky surveys. The goal of the project is to create highly scalable inference methods for extracting a unified catalog of objects in the visible universe from all available astronomy data.

The Hyperion award was presented to the Celeste team during ISC 2017: Jeff Regier, Kiran Pamnany, Keno Fischer, Andreas Noack, Max Lam, Jarrett Revels, Steve Howard, Ryan Giordano, David Paul, David Schlegel, Jon McAuliffe, Alan Edelman, Viral Shah, Rollin Thomas and Prabhat.

GALACTOS PROJECT SOLVES ONE OF COSMOLOGY'S HARDEST CHALLENGES:

Cosmologists and astronomers have wanted to perform the 3-point computation for a long time but could not do so because they did not have access to scalable methods and highly optimized calculations that they could apply to datasets. In 2017, the Galactos project — which teams researchers from Harvard University with the Big Data Center collaboration involving NERSC, Berkeley Lab and Intel — made a major breakthrough in successfully running the 3-point correlation calculation on the Outer Rim, the largest known simulated galaxy dataset that contains information for two billion galaxies.



Two Berkeley Lab-led computing projects received Hyperion Research's 2017 HPC Innovation Excellence Award for "the outstanding application of HPC for business and scientific achievements."

The Hyperion award was presented to the Galactos team during SC17: Brian Friesen, Mostofa Patwary, Brian Austin, Nadathur Satish, Zachary Slepian, Narayanan Sundaram, Debbie Bard, Daniel Eisenstein, Jack Deslippe, Pradeep Dubey and Prabhat.

NERSC HOSTS JEFFERSON LAB QCD RESEARCHER

Balint Joo, a well-known quantum chromodynamics researcher from Jefferson Lab, spent the summer of 2017 at NERSC working on programming models for the NESAP

program. NERSC hosted Joo while he explored the KOKKOS programming environment as part of NESAP's development efforts for future supercomputing architectures.

ALL-FEMALE STUDENT CLUSTER TEAM COMPETES AT ISC FOR SECOND TIME

Five former NERSC interns and one student assistant joined forces to participate in the Student Cluster Competition at ISC17 in Frankfurt, Germany, marking only the second time that NERSC has fielded a student cluster competition team.

The all-female contingent competed against 11 other teams over three days to demonstrate the best performance across a series of benchmarks and applications on a small supercomputer cluster that they designed and built. Five of the six members were also on NERSC's ISC16 team:

- **Kristi Arroyo**, a senior in computer science at the Missouri University of Science and Technology
- **Tiffany Connors**, who graduated in 2017 from Texas State University with a computer science degree and was hired as a post-doc at NERSC

- **Yong Li Dich**, now a junior at Harvard University who plans to pursue computer science
- **Grace M. Rodríguez Gomez**, who graduated in 2017 from the University of Puerto Rico with a computer science degree
- **Elizabeth Wang**, a freshman at the University of Illinois Urbana-Champaign
- **Ruoyun Zheng**, a freshman at Caltech

The team was once again led by Rebecca Hartman-Baker, group lead for NERSC's User Engagement Group. During the competition, the team ran a variety of problem sets given to them by the competition organizers, including WRF, a well-known weather modeling application; SPLOTCh, a visualization program for astronomy; Graph500; HPCG; Tensorflow; and FEniCS.

Hartman-Baker and the team again worked closely with their sponsors, Intel and Cray, to select and procure the parts they used to build their cluster, which featured six nodes plus an

Intel NUC that serves as the head node. The team also used the Open HPC software stack and SPAC for software management and installation.

NERSC PREPARES FOR POTENTIAL EMERGENCIES

NERSC participated in Berkeley Lab's annual earthquake drill in October 2017 and a lab-wide evacuation exercise in December 2017. In the December drill, some lab buildings were evacuated to an off-site assembly area, while some staff left by vehicle and others sheltered in place. Six NERSC staff members who have completed Community Emergency Response Training to support the Lab during an actual disaster event assisted in directing NERSC staff during the evacuation drill.

In 2017, NERSC also hosted a presentation from the Lab's Protective Services Group on the use of Stop the Bleed kits, designed to provide trained emergency responders and the

public with immediate access to products intended to stop traumatic hemorrhaging. One kit is located on each office floor in Wang Hall, along with the Automated External Defibrillators.

NERSC has long been proactive in emergency preparedness, particularly earthquakes, given that the San Francisco Bay Area is well known to be a seismically active area. Thus Shyh Wang Hall was designed to include a custom-built seismically isolated floor to protect the computer systems and personnel. The unique floor is the brainchild of Dynamic Isolation Systems, a pioneer in the development of seismic isolation technology.

Data centers typically feature a raised access floor system in the machine rooms to provide flexibility in wiring, hardware location and air distribution. The design isolates data lines, power cables and piping to create a safe environment for operators and to protect hardware operations. Due to Berkeley Lab's proximity to the nearby Hayward Fault, the Shyh Wang Hall design takes the raised floor concept a bit further, with a seismically isolated floor that protects the supercomputers from damage by decoupling them from the floor slab below. While the computer cabinets sit on a conventional raised floor, the floor is mounted on a steel substructure that rolls on casters and utilizes a sophisticated spring dampening mechanism. This prevents high seismic forces from being transmitted into the computers.



Six NERSC staff members have completed Community Emergency Response Training to support the Lab during an actual disaster event.



NERSC staff evacuated to an off-site assembly location along with other Lab participants during a December 2017 earthquake preparedness drill.

MULTI-FACTOR AUTHENTICATION ADDED FOR NERSC USERS

In 2017, NERSC evaluated and began deploying multi-factor authentication (MFA) for its users. The design goals were to minimize support costs and minimize negative impact on users while improving the overall security of the center. Toward this end, NERSC is taking a phased approach, starting with an opt-in phase where adventurous users can enable MFA for their account. This also provides an opportunity for NERSC to stress-test its implementation. In the future, NERSC will move toward an opt-out approach in which users will need explicit approval to disable MFA for their account. The intent is to eventually get to a mandatory policy, with provisions made to ensure that complex workflows and other automated mechanisms can continue to run smoothly and securely.

Before implementing an MFA solution, NERSC surveyed solutions being used at other DOE centers and NSF sites.

NERSC evaluated several options before deciding on a solution based on LinOTP using TOTP-based soft tokens that users can configure on a standard smartPhone. LinOTP was chosen because it was already in use by Berkeley Lab's IT department and had proven itself to be a robust platform. Rather than standing up a separate instance, NERSC leveraged the Lab's deployment. This required working out SLA agreements and ensuring that NERSC had the ability to monitor the service.

Since usability of this service was a major concern for NERSC staff, NERSC conducted usability reviews with a number of target users as well as internal testing with NERSC staff. This testing helped further refine and improve the behavior of the solution and associated documentation.

PERSONNEL TRANSITIONS



DOE PROGRAM MANAGER DAVE GOODWIN RETIRES

NERSC's long-time DOE Program Manager, Dave Goodwin, retired in 2017.

In addition to being NERSC's program manager, Goodwin was also the ASCR Leadership Computing Challenge (ALCC) program manager and Chair of ASCR's Supercomputing Allocations Committee.

GERBER TAPPED TO HEAD NERSC'S HPC DEPARTMENT

Richard Gerber was named head of NERSC's High-Performance Computing (HPC) Department, formed in early 2016 to help the center's 7,000 users take full advantage of new supercomputing architectures and guide and support them during the ongoing transition to exascale.

Gerber previously served as acting head of the department, which comprises four groups: Advanced Technologies, Application Performance, Computational Systems and User Engagement.

"This is an exciting time because the whole HPC landscape is changing with manycore, which is a big change for our users," said Gerber, who joined NERSC's User Services Group in 1996 as a postdoc, having earned his PhD in physics from the

University of Illinois. "Users are facing a big challenge; they have to be able to exploit the architectural features on Cori, and the HPC Department plays a critical role in helping them do this."

The HPC Department is also responsible for supporting world-class systems in a production computing environment and looking to the future. "We work with complex, first-of-a-kind systems that present unique challenges," Gerber said. "Our staff is constantly providing innovative solutions that make systems more capable and productive for our users. Looking forward, we are evaluating emerging technologies and gathering scientific needs to influence future HPC directions that will best support the science community."

NEW HIRES

ERSHAAD BASHEER, HPC Systems Engineer, Computational Systems Group

MARK DAY, Senior Computer Systems Engineer, Infrastructure Services

ADITI GAUR, Data Systems Engineer

TAYLOR GROVES, HPC Performance Engineer in the Advanced Technology Group, joined NERSC in 2017 to work on a methodology for analyzing networking requirements for the user facility's workload. Before coming to Berkeley, Groves was a Graduate Research Assistant at Sandia National Laboratories' Center for Computing Research in New Mexico where he worked on simulation and modeling of HPC networks.

KRISTY KALLBACK-ROSE joined the NERSC Storage Systems Group in early 2017. In recent years, Kristy has

worked with geographically distributed filesystems and archival storage, including GPFS and HPSS, primarily at Indiana University. Prior to that she worked in a variety of roles including grid computing in support of the ATLAS project, databases, development and instruction. She has undergraduate degrees in Japanese and physics and a masters in physics.

GEORG RATH, Computer Systems Engineer, Computational Systems Group

CHARLENE YANG is an Application Performance Consultant at NERSC. Before joining NERSC in September 2017, she was a Supercomputing Application Specialist at the Pawsey Supercomputing Center in Perth, Australia. Yang has a Ph.D. in signal processing and wireless communication from the University of Western Australia.

NEW NESAP POSTDOCS

ZAHRA RONAGHI joined NERSC as the first NESAP for Data postdoc in January 2017, working with Doga Gursoy at Argonne National Lab on performance optimization of TomoPy, a tomographic reconstruction code that is used by DOE light sources. She is also working with the Application Performance Group at NERSC exploring performance portability of BoxLib/AMReX. She graduated from Clemson University with a Ph.D. in biomedical engineering.

RAHULKUMAR GAYATRI earned his Ph.D. in the area of parallel programming models from Barcelona Supercomputing Center in March 2015. His thesis work was on synchronization of multiple threads on a multi-core processor. Later he worked in the HPC group at Wipro Infotech where he provided parallel programming solutions to clients. At NERSC he is working on the SW4 and performance portability projects.

JONATHAN MADSEN earned his Ph.D. from Texas A&M University in Nuclear Engineering. As part of the NESAP program, Madsen is working on applications developed by the Cosmic Microwave Background group at Berkeley Lab led by Julian Borrill. He has been a member of the Geant4 collaboration since 2011 and is currently a member of the Run, Event, and Detector Responses Working Group.

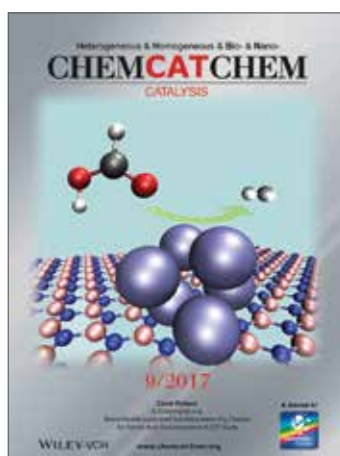
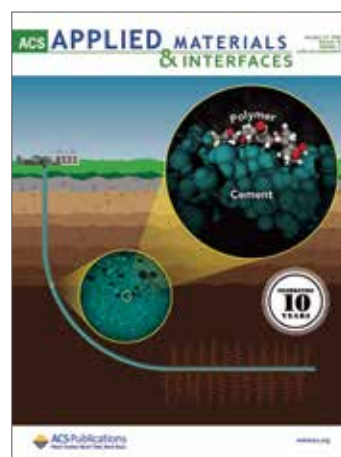
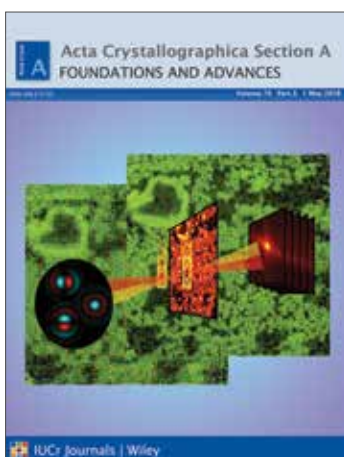
LAURIE STEPHEY, a NESAP for Data postdoc, is working with the DESI (Dark Energy Spectroscopic Instrument) project to help make their large-scale data processing run efficiently at NERSC. She is employing strategies like MPI using mpi4py and Just-In-Time compilation using numba to make the DESI simulation tools run faster. In 2017, she earned her Ph.D. from the University of Wisconsin-Madison studying edge physics in the HSX and W7-X stellarators.

Publications

In 2017, NERSC's 7,000 users reported more than 2,200 peer-reviewed published papers that involved NERSC resources. In addition, NERSC staff contributed more than 150 papers to scientific journals and conferences, showcasing our continued involvement in HPC hardware and software development and our increasing expertise in data-intensive computing, data storage and analysis,

scientific workflows, code optimization, machine learning and quantum computing.

To see a comprehensive list of publications and presentations by NERSC staff in 2017, go to <http://bit.ly/NERSCstaffpublications>



Allocation of NERSC Director's Reserve of Computer Time

In 2017, NERSC used its Director's Discretionary Reserve (DDR) pool of 600 million NERSC hours to enable strategic projects, support readiness efforts of Exascale Computing Project projects and to award time to projects that were selected through the competitive NERSC High-Impact Science at Scale program.

Through the High-Impact Science at Scale program seven projects were given awards that enabled them to address compelling scientific challenges using Cori, the world's largest supercomputer based on the Intel Xeon Phi KNL processors. Using up to 657,000 of Cori's processor cores, these seven projects ran ground-breaking simulations in cosmology, plasma physics, subsurface flow, compact accelerator design,

nuclear physics, and materials science. The Director's Reserve also supported strategic projects that don't cleanly fit into the DOE Office of Science allocation scheme, including research into advanced networking for science, research in computational neuroscience, and impacts of heat on irrigation water demands.

Top 10 Director's Reserve projects for 2017 (in terms of hours used and excluding high-impact science at scale projects)

PI Name	Organization	Project Title	Science Category	NERSC Hours Used
Compo, Gilbert	U. Colorado Boulder	The 20th Century Reanalysis version 3 (1978-2017)	Earth Systems Science	72,351,183
Newman, Gregory A.	Berkeley Lab	Large scale Geophysical Imaging utilizing NERSC's Cori Supercomputer at Scale	Geoscience	35,040,014
Monga, Inder	Berkeley Lab	Parallel Forward Error Correction Verification and Scaling Prototype for Advanced Networking	Computer Science	19,456,877
Antypas, Katie	Berkeley Lab	NERSC Application Readiness for Future Architectures	Computer Science	18,249,367
Bates, Susan	NCAR	High-resolution ocean-atmosphere simulations with the CESM1.3	Earth Systems Science	11,398,130
Borrill, Julian D.	Berkeley Lab	Planck Simulations	Cosmology	10,791,430
Bouchard, Kristofer E.	Berkeley Lab	Collaborative Research in Computational Neuroscience - CRCNS	Biosciences	5,203,507
Sharifzadeh, Sahar	Boston Univ	Large-Scale Many-Body Perturbation Theory Simulations of Optoelectronic Materials	Computer Science	2,951,487
Nogales, Eva	Berkeley Lab	Structuromics: solving structures to bridge the gap between sequence and function	Biosciences	2,861,060
Jones, Andrew D.	Berkeley Lab	Impacts of future heat on irrigation water demand	Earth Systems Science	2,375,469

High-impact science at scale projects for 2017. Each project had a NERSC staff member assigned to help the teams prepare their code and workflows for extreme scale computing. Staff arranged reservations of resources and debugged problems as they arose during the runs.

PI Name	Organization	Project Title	NERSC Hours Used	Nodes (Cores) Used
Trebotich, David	Berkeley Lab	Chombo-Crunch: Extreme scale simulation of flow and transport in heterogeneous media	143,056,732	9,664 (657,152)
Heitmann, Katrin	Argonne Lab	Knowhere (Cosmology)	92,391,956	6,144 (417,152)
Stanier, Adam	Los Alamos Lab	Probing the physics of magnetic reconnection – from fusion energy to space plasmas	64,170,260	2,048 (139,264)
Louie, Steven	UC Berkeley	Ab initio quasiparticle and optical properties of materials at scale	56,952,502	4,096 (278,528)
Lukic, Zarija	Berkeley Lab	Physical model of the intergalactic medium	51,167,795	8,192 (557,056)
Karsch, Frithjof	Brookhaven Lab	Net strangeness and net electric charge fluctuations in strongly interacting matter	41,614,572	9,178 (624,104)
Vay, Jean-Luc	Berkeley Lab	High-resolution 3D studies of asymmetric effects in the BELLA plasma accelerator experiments	27,886,930	1,274 (86,632)

APPENDIX A:

NERSC Users Group Executive Committee

OFFICE OF ADVANCED SCIENTIFIC COMPUTING RESEARCH

Milan Curcic, *University of Miami*

Jeff Hammond, *Intel*

Brian Van Straalen, *Lawrence Berkeley National Laboratory*

OFFICE OF BASIC ENERGY SCIENCES

Eric Bylaska, *Pacific Northwest National Laboratory*

Monojoy Goswami, *Oak Ridge National Laboratory*

Paul Kent, *Oak Ridge National Laboratory*

OFFICE OF BIOLOGICAL AND ENVIRONMENTAL RESEARCH

Samson Hagos, *Pacific Northwest National Laboratory*

Gary Strand, *NCAR*

David Wong, *U.S. Environmental Protection Agency*

OFFICE OF FUSION ENERGY PHYSICS

Christopher Holland, *University of California, San Diego*

Nathan Howard, *MIT*

Orso Meneghini, *General Atomics*

OFFICE OF HIGH ENERGY PHYSICS

James Amundson, *Fermilab*

Wirming An, *UCLA*

Zarija Lukic, *Lawrence Berkeley National Laboratory*

OFFICE OF NUCLEAR PHYSICS

Balint Joo, *Jefferson Lab*

Michael Zingale, *Stony Brook University*

David Tedeschi, *Lawrence Berkeley National Laboratory*

MEMBERS AT LARGE

Carlo Benedetti, *Lawrence Berkeley National Laboratory*

David Hatch, *University of Texas, Austin*

Pieter Maris, *Iowa State University*

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. 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

Barbara Helland, *Associate Director*
 Julie Stambaugh, *Financial Management Specialist*
 Lori Jernigan, *Program Support Specialist*
 Tameka Morgan, *Administrative Specialist*

FACILITIES DIVISION

Ben Brown, *Director (Acting), Physical Scientist, ESnet Program Manager*
 Betsy Riley, *Computer Scientist, ALCC Program Manager*
 Carolyn Lauzon, *Physical Scientist, ALCC Program Manager*
 Sonia Sachs, *Computer Scientist, ALCF*
 Robinson Pino, *Computer Scientist, REP Program Manager*
 Christine Chalk, *Physical Scientist, ORLC Program Manager, CSGF Program Manager*
 Sally McPherson, *Program Assistant*

RESEARCH DIVISION

Robinson Pino, *Director (Acting)*
 Teresa Beachley, *Program Assistant*
 Laura Biven, *Mathematician, Data & Visualization*
 Randall Laviolette, *Physical Scientist, SciDAC Application Partnerships*
 Thomas Ndousse-Fetter, *Computer Scientist, Network Research*
 Ceren Susut, *Physical Scientist, SC Program SAPs*
 Rich Carlson, *Computer Scientist, Collaboratories/Middleware*
 Steven Lee, *Physical Scientist, Base, Math: Algorithms, Models, Data*
 Lucy Nowell, *Computer Scientist, Computer Science*
 Angie Thevenot, *Program Support Specialist*

APPENDIX C:

Acronyms and Abbreviations

ACM Association for Computing Machinery	CERN European Organization for Nuclear Research	DOI Digital Object Identifier	FLOPS Floating Point Operations
ACS American Chemical Society	CESM Community Earth Systems Model	DSL Dynamic Shared Library	FTP File Transfer Protocol
ALCC ASCR Leadership Computing Challenge	CFD Computational Fluid Dynamics	DTN Data Transfer Node	GB Gigabytes
ALS Advanced Light Source, Lawrence Berkeley National Laboratory	CI/CD Continuous Integration and Continuous Deployment	DVS Data Virtualization Service	Gbps Gigabits Per Second
ANL Argonne National Laboratory	CMB Cosmic Microwave Background	ECP Exascale Computing Project	GPU Graphics Processing Unit
API Application Programming Interface	CO₂ Carbon dioxide	EFRC DOE Energy Frontier Research Center	GUI Graphical User Interface
APS American Physical Society	CPU Central Processing Unit	EMSL Environmental Molecular Science Laboratory, Pacific Northwest National Laboratory	HDF5 Hierarchical Data Format 5
ASCII American Standard Code for Information Interchange	CRD Computational Research Division, Lawrence Berkeley National Laboratory	EPSC SciDAC Center for Edge Physics Simulations	HEP Office of High Energy Physics
ASCR Office of Advanced Scientific Computing Research	CSE Computational Science and Engineering	ERD Earth Sciences Division, Lawrence Berkeley National Laboratory	HPC High Performance Computing
BDC Big Data Center	DARPA Defense Advanced Research Projects Agency	ESnet Energy Sciences Network	HPC4Mfg High Performance Computing for Manufacturing
BER Office of Biological and Environmental Research	DAQ Data Acquisition System	eV Electron Volts	HPSS High Performance Storage System
BES Office of Basic Energy Sciences	DDR Double Data Rate	FDM Finite Difference Method	HTML Hypertext Markup Language
BMS Building Management System	DESI Dark Energy Spectroscopic Instrument	FEC Forward Error Correction	HTTP Hypertext Transfer Protocol
BNL Brookhaven National Laboratory	DFT Density Functional Theory	FES Office of Fusion Energy Sciences	I/O Input/Output
CCM Cluster Compatibility Mode	DNS Direct Numerical Simulation	FICUS Facilities Integrating Collaborations for User Science	IEEE Institute of Electrical and Electronics Engineers
	DOE U.S. Department of Energy		InN Indium Nitride
			IPCC Intel Parallel Computing Center; Intergovernmental Panel on Climate Change

IPM

Integrated Performance Monitoring

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

KNL

Knights Landing Processors

LED

Light-emitting Diode

LANL

Los Alamos National Laboratory

LCLS

Linac Coherent Light Source

LLNL

Lawrence Livermore National Laboratory

MIT

Massachusetts Institute of Technology

MODS

Monitoring of Data Services

MOF

Metal Oxide Framework

MPI

Message Passing Interface

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

OLCF

Oak Ridge Leadership Computing Facility

OMNI

Operations Monitoring and Notification Infrastructure

OpenMP

Open Multi-Processing

OpenMSI

Open Mass Spectrometry Imaging

PB

Petabytes

PDACS

Portal for Data Analysis services for Cosmological Simulations

PDSF

Parallel Distributed Systems Facility, NERSC

PF

Petaflop

PI

Principal Investigator

PiB

Pebibyte

PIC

Particle-In-Cell Simulations

POSIX

Portable Operating System Interface

PSII

Photosystem II

PNNL

Pacific Northwest National Laboratory

PPPL

Princeton Plasma Physics Laboratory

PUE

Power Usage Effectiveness

QCD

Quantum Chromodynamics

QUBITS

Quantum Bits

SC

DOE Office of Science

SciDAC

Scientific Discovery Through Advanced Computing

SDN

Software-defined Networking

SIAM

Society for Industrial and Applied Mathematics

SLURM

Simple Linux Utility for Resource Management

SPACK

Super Computing Package Manager

TACC

Texas Advanced Computing Center

TB

Terabytes

TOAST

Time Ordered Astrophysics Scalable Tools

TOKIO

Total Knowledge of I/O Framework

URL

Universal Resource Locator

VASP

Vienna Ab initio Simulation Package

VM

Virtual Machine

WAN

Wide Area NetworkCredits

FOR MORE INFORMATION ABOUT NERSC, CONTACT:

Kathy Kincade

NERSC Communications
1 Cyclotron road
Berkeley Lab, MS593024B
Berkeley, CA 94720-8148
Email: kkincade@lbl.gov
Phone: 510-495-2124
Fax: 510-486-4300

NERSC's Web Site

www.nersc.gov

NERSC Annual Report Editor

Kathy Kincade

Contributing Writers and Editors

Katie Antypas, Brian Austin, Elizabeth Bautista, Deborah Bard, Shane Canon, Jack Deslippe, Doug Doerfler, Richard Gerber, Rebecca Hartman-Baker, Damian Hazen, Doug Jacobsen, Thorsten Kurth, Glenn Lockwood, Betsy MacGowan, Prabhat, David Skinner, Cory Snively, Linda Vu, Margie Wylie

Design

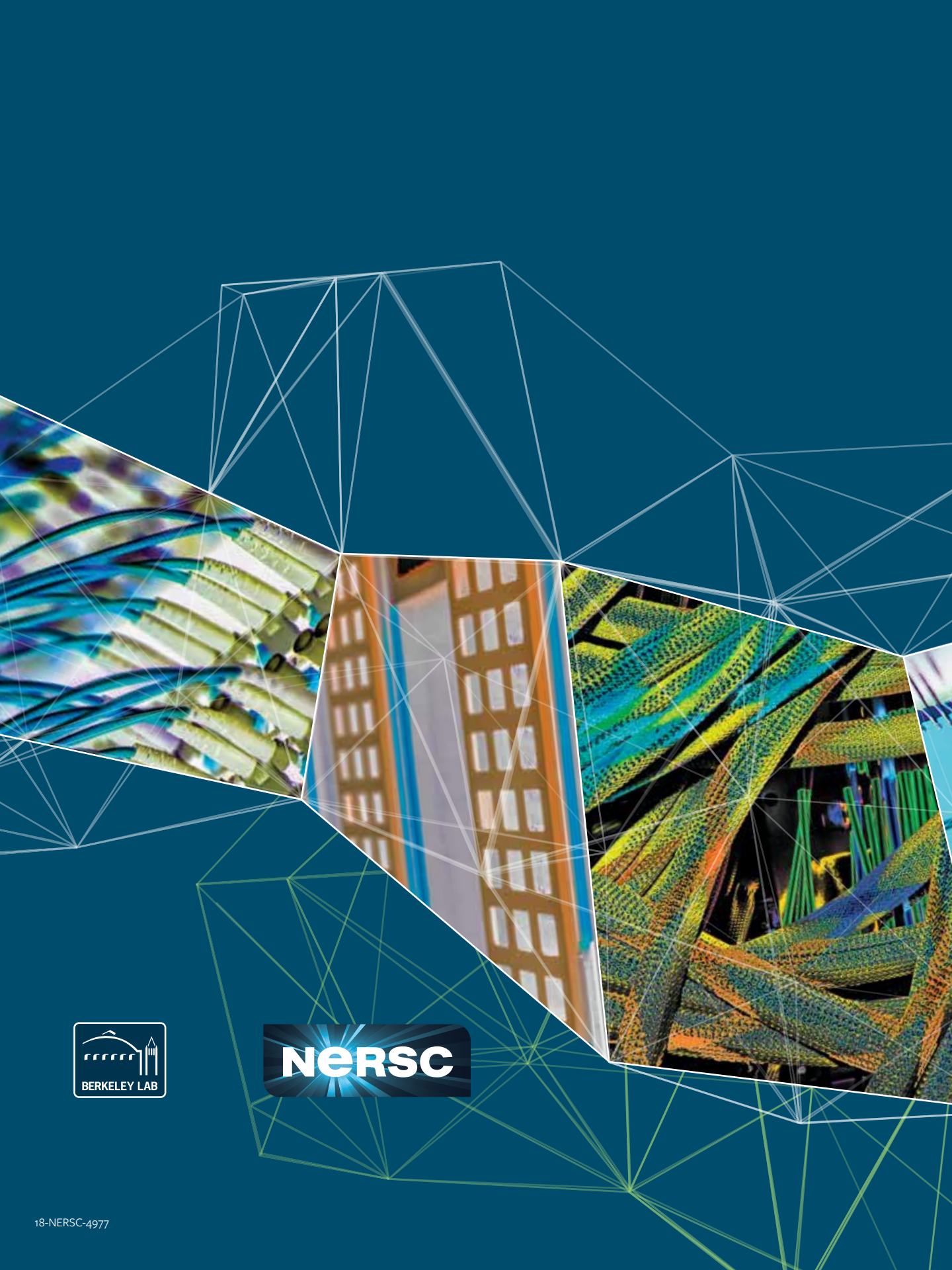
Design, layout, illustration, photography and printing coordination:
Berkeley Lab Creative Services



Image credit: Kelly Owen, Berkeley Lab

DISCLAIMER

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



NERSC