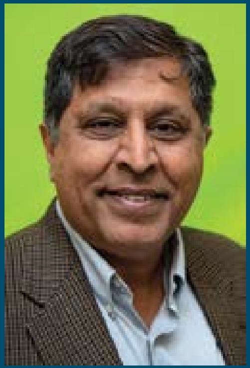
2019 ANNUAL REPORT

National Energy Research Scientific Computing Center

U.S. DEPARTMENT OF ENERGY | Office of Science

Table of Contents

The National Energy Research Scientific Computing Center (NERSC) is the mission high-performance computing facility for the U.S. Department of Energy's Office of Science (DOE SC). NERSC is managed by Lawrence Berkeley National Laboratory (Berkeley Lab) and funded by the DOE SC Advanced Scientific Computing Research Office. NERSC's mission is to accelerate scientific discovery at the DOE SC through HPC and extreme data analysis.



Sudip Dosanjh
NERSC Division Director

Director's Note

NERSC currently serves approximately 8,000 scientists annually who are working on 800 research projects spanning the range of SC scientific disciplines. In addition, NERSC continues to have a large impact on scientific research, with more than 2,400 distinct refereed publications in 2019 citing NERSC.

From both a performance and operational perspective, NERSC hit a number of milestones in 2019. The Edison Cray XC40 system was retired in May to make way for the Perlmutter system in late 2020. Edison was a productive and popular system that, over its lifetime, delivered 12 billion NERSC hours to scientific research. This meant that our Cori supercomputer became the primary system for users for the second half of 2019 and into 2020.

Throughout 2019, NERSC made significant headway in preparing NERSC users and the NERSC facility for Perlmutter, a next-generation heterogeneous system that will incorporate both CPU and GPU processors and is NERSC's first supercomputer to employ a GPU architecture. Giving our users access to the very latest in GPU-accelerated technology is an important step in ensuring they remain productive and are able to utilize the systems to prepare for the exascale era. When Phase 1 arrives in late 2020, researchers computing at NERSC will need to be prepared to use GPUs for their simulation, data processing, and machine learning workloads.

Thus a significant priority for NERSC in 2019 was preparing the facility and user applications for Perlmutter. This included launching the next round of the NERSC Exascale Science Applications Program (NESAP), with applications targeting GPU readiness in three focus areas: Cutting-edge simulations of complex physical phenomena, data analysis science pipelines that process massive data sets from experimental and observational science facilities, and machine learning and deep learning solutions to improve scientific discovery potential on experimental or simulation data.

As part of NESAP, in February 2019 NERSC and Cray began hosting a series of GPU hackathons to help these teams gain knowledge and expertise about GPU programming and apply that knowledge as they port their scientific applications to GPUs. NERSC also hosted a number of trainings and workshops, including the first annual GPUs for Science Day, which was designed to facilitate the transition to GPU systems by giving attendees the motivation, tools, and expertise needed to make this change; and the Monterey Data Conference, an invitation-only meeting that gives researchers from DOE national laboratories, f acilities, universities, and industry the opportunity to share and showcase the latest advances and challenges in scientific data analysis and computing.

Other critical focus areas for NERSC in 2019 included the Superfacility Project. Large-scale analysis of experimental data has revolutionized our understanding of the physical world, from the discovery of elementary particles and the accelerating expansion of the universe to new insights into the microbiome and the ability to better predict extreme weather events. Berkeley Lab's Computing Sciences Superfacility Initiative is a framework for further integrating experimental and observational instruments with computational and data facilities, bringing the power of exascale systems to the analysis of real-time data from light sources, microscopes, telescopes and other devices.

In support of this concept, in 2019 NERSC launched the Superfacility Project, an internal effort to accelerate pipelines from DOE experimental and observational facilities in an automated, high-performing manner. NERSC's superfacility team held a number of outreach events over the past year and has enabled capabilities that improve data movement and sharing for experimental facilities. Targeted science engagements are another key element of the Superfacility Project, which is structured to ensure that our work will be generally useful across multiple science domains as we work to develop a reusable and extensible toolkit.

Similarly, NERSC continues to be active in the Exascale Computing Project, a multi-agency effort spearheaded by the DOE in 2016 to ensure that the U.S. maintains leadership in high performance computing and high-speed networking for science, innovation, and national security. NERSC currently holds key leadership positions within ECP, supports five ECP applications through the NESAP program, and is at the forefront of efforts to develop software integration and continuous integration capability for ECP software teams.

Finally, professional development across the facility is always a priority for NERSC, and we continue to identify ways to provide a work environment where all staff can thrive. A new focus at Berkeley Lab is the creation of a psychologically safe environment for employees. To this end, in July 2019 the lab launched its new IDEA (inclusion, diversity, equity, and accountability) program. NERSC is well represented in this effort, with staff serving as Computing Sciences Area IDEA representatives and leaders of Employee Resource Groups, and many others regularly attending IDEA-themed events.

As the IDEA program continues to roll out, staff are encouraged to attend learning events, and managers at NERSC are beginning to integrate prepared learning materials into normal work activity. NERSC is committed to continued involvement in the IDEA program to foster a work environment where all staff can bring their best contributions forward.

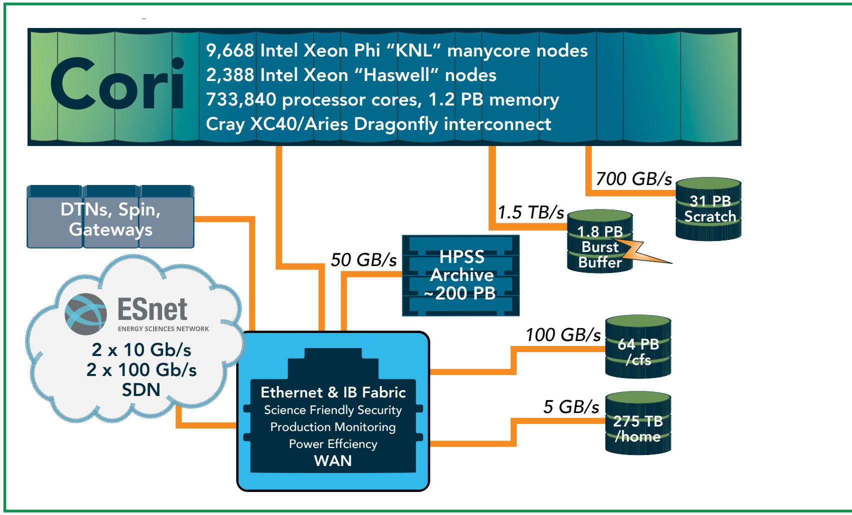


Figure 1. NERSC systems at the end of 2019.

NERSC by the Numbers

2019 NERSC USERS ACROSS US AND WORLD

50 States + Washington D.C.

46 Countries

>7,000 ANNUAL USERS FROM ~1,700 Institutions + National Labs

34% Graduate Students

22% Postdoctoral Fellows

18% Staff Scientists

12% University Faculty

7% Undergraduate Students

7% Professional Staff

62% Universities

29% DOE Labs

5% Other Government Labs

2% Industry

1% Small Businesses

<1% Private Labs

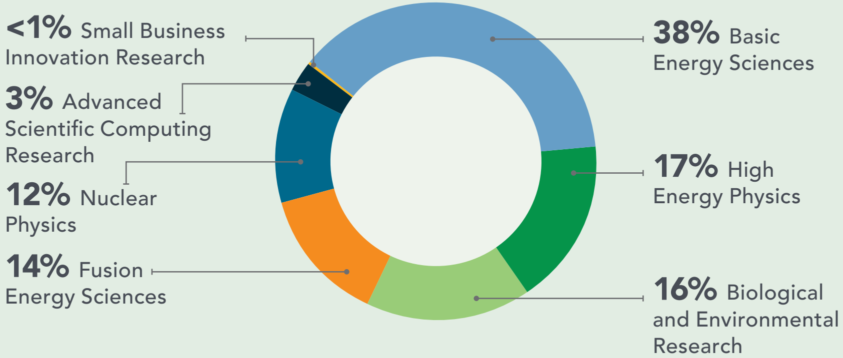
9 BILLION CORE HOURS USED IN 2019

2,435
Publications Involving NERSC Resources

Data Stored 200 Petabytes

900
Projects

2019 DOE Office of Science Program Usage Breakdown



NERSC Data Center PUE

1.08
NERSC Current PUE*

1.67
Industry Standard Data Center PUE**
(as of April 2020)

*Power Usage Effectiveness (1.0 is considered the PUE gold standard)

**Source: https://journal.uptimeinstitute.com/is-pue-actually-going-up/

2019 Top Science Disciplines
(By computational hours used)

Energy Technologies
Mathematics
Earth Systems
Accelerator Science
Engineering
Materials Science
Nuclear Physics
Geoscience
Chemistry
Fusion Energy
Condensed Matter
High Energy Physics
User Facilities
Computer Science
Astrophysics and Cosmology

Innovations

GPUs for Science

NERSC’s next-generation supercomputer, Perlmutter, is a heterogeneous Cray/Hewlett Packard Enterprise Shasta system designed to power data-intensive scientific research and analytics. Perlmutter will be deployed at NERSC in two phases: the first set of cabinets, featuring NVIDIA GPU-accelerated nodes, will arrive in late 2020; the second set, featuring AMD CPU-only nodes, will arrive in mid-2021. A 35-petabyte all-flash Lustre-based file system using HPE’s ClusterStor E1000 hardware will also be deployed in late 2020.

Perlmutter is NERSC’s first production system to feature GPUs, which provide a great amount of energy-efficient computing capability for simulation, data processing, and machine learning workloads. This GPU-accelerated and CPU-based system will have at least 1,500 GPU-accelerated nodes with 6,000 GPUS and 3,000 CPU-only nodes, providing 3-4 times the capability of NERSC’s current supercomputer, Cori.

Since announcing this system in October 2018, NERSC has been working with its users to optimize their applications for the next-generation GPU processors that will power Perlmutter.

NESAP Targets GPU Readiness

Nearly half of the workload currently running at NERSC is poised to take advantage of GPU acceleration, and NERSC has played a key role in helping the scientific community leverage GPU capabilities for a broad range of applications. At the core of these efforts is the NERSC Exascale Science Applications Program (NESAP). NESAP partnerships allow projects to collaborate with NERSC and HPC vendors by providing access to early hardware, prototype software tools for performance analysis and optimization, and special training. NESAP efforts targeting Perlmutter started in 2019, and teams have been working with NERSC staff along with NVIDIA and Cray engineers to accelerate as many codes as possible and ensure

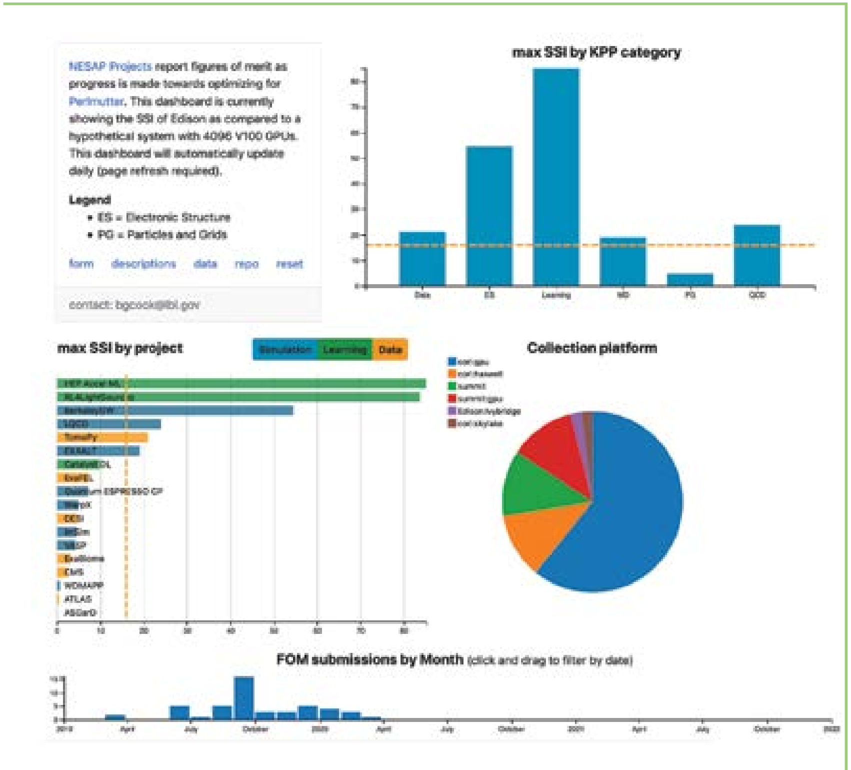


Figure 2. NESAP has developed a custom online dashboard for tracking applications progress toward Perlmutter performance goals.

that the scientific community can hit the ground running when Perlmutter comes online.

NESAP is targeting GPU readiness in three focus areas:

- **NESAP for Simulations:** Cutting-edge simulations of complex physical phenomena
- **NESAP for Data:** Data analysis science pipelines that process massive data sets from experimental and observational science facilities
- **NESAP for Learning:** Machine learning and deep learning solutions to accelerate scientific discovery from experimental or simulation data; or improving HPC applications by replacing parts of the software stack or algorithms with machine learning and deep learning solutions.

NESAP made substantial progress in 2019, with several projects showing improvements in the code performance through combined efforts of NERSC postdocs, NERSC staff liaisons, vendor partners, and project developers. One example is ExaFEL, a NESAP for Data project that aims to accelerate algorithms for X-ray free electron lasers (XFELs) in preparation for the three-order-of-magnitude increase in detector data rates expected by 2025 from the Linac Coherent Light Source (LCLS) at SLAC. Users of an XFEL require an integrated combination of data processing and scientific interpretation, where both aspects demand intensive computational analysis. ExaFEL team members attended a NESAP hackathon to target two of its code bases: MTIP for fluctuation scattering and CCTBX for macromolecular crystals, in particular the nanoBragg benchmark.

MTIP-FXS aimed to reduce the total wall time to complete one reconstruction of the example dataset, which consists of a set of correlations of data from a protein molecule that were collected during 15 minutes of an experiment at LCLS, at roughly 100-120Hz (events/sec). The correlations were based on 6,000 diffraction patterns culled from a raw dataset that was much larger. By October 2019, 90% of the iterative phasing code had been ported to CUDA, a language designed to leverage the computational capabilities of NVIDIA GPUs. Tests on NERSC's Cori GPU test-bed indicate 8-10x speedup when compared to CPU benchmarks.

Also in 2019, NERSC staff developed a custom dashboard that tracks the progress of NESAP projects by maintaining a live view of the current status of each project and a projection of the total scientific throughput, as measured by the Scalable System Improvement metric by project and by category.

Community Outreach: GPU Workshops, Conferences, and Hackathons

As Perlmutter approaches, we see a need for users to build proficiency in using GPUs for scientific computing and a more general need for performance portability; applications need to be able to run, with satisfactory performance, on multiple architectures.

Toward this end, in 2019 NERSC held a number of GPU-centric trainings and workshops. These events included:

- **GPUs for Science Workshop.** Held July 2-3, 2019, this event included more than 20 talks covering a broad overview of GPU topics. Speakers from NVIDIA, Berkeley Lab, other national labs, and several universities were present. More than 120 people attended in person, and more than 250 people attended remotely. The second day of the event included hands-on tutorials and a GPU hackathon for more than 50 in-person attendees.
- **GPU Hackathons.** In February 2019 NERSC and Cray/HPE began hosting a series of GPU hackathons to help NESAP teams gain knowledge and expertise about GPU programming and apply that knowledge as they port their scientific applications to GPUs. These hands-on events are designed to help ensure that NESAP codes and the broader NERSC workload are ready to take advantage of the GPUs when Perlmutter arrives.

Figure 3. In 2019, NERSC hosted a number of GPU trainings and workshops.

Sharing Directive-based GPU Programming Models

Directive-based programming models offer a way to access the power of GPUs and many-core processors in general without relatively fewer code modification compared to other alternatives. Widely accepted such models, like OpenMP, are also portable across various HPC architectures.

NERSC continues to investigate the feasibility of being able to use directive-based programming models on GPU systems with acceptable performance. One research project in 2019 involved porting a widely used block Eigensolver — locally optimal block preconditioned conjugate gradient (LOBPCG) — to NVIDIA GPUs using two directive-based programming models: OpenMP and OpenACC.

The project was novel in two regards. First, it stress-tested the interoperability of OpenMP and OpenACC compiler implementations with CUDA. This was needed because an efficient LOBPCG solver implementation depends on optimized libraries — cuBLAS and cuSPARSE — on NVIDIA GPUs. Second, it provided an opportunity to compare the ease of use and performance of unified memory — a feature that allows programmers to more easily write their codes on GPU-accelerated systems — compared to programmer-managed data movement between CPU and GPU.

The research project was successful, with results presented at the Sixth Workshop on Accelerator Programming Using Directives. The LOPBCG GPU implementations in OpenMP and OpenACC achieved a 2.8x - 4.3x speedup over an optimized CPU implementation when tested with four different input matrices.

Machine Learning and Data-centric Science Support

In conjunction with our emphasis on GPU readiness, NERSC has been expanding its capabilities in using HPC for data analysis and machine and deep learning for science. In 2019, projects primarily focusing on data analysis made up a significant fraction of the projects across all programs in the Office of Science. Heavy users of NERSC data-focused capabilities included experimental projects in High Energy and Nuclear Physics, light sources from the Basic Energy Sciences program, and climate-science observational and simulated data.

Python and Jupyter are especially important for data-centric workflows. In 2019, average monthly Python usage at NERSC was more than 1,000 unique users per month, a slight increase over 2018. Jupyter usage has been trending upwards, as can be seen in Figure 4. At any given time there are, on average, about 200 active Jupyter notebooks running on Cori, excluding notebooks being run by staff. As Jupyter's popularity continues to grow, we have been working on ways to expand and make the service more robust, leveraging compute nodes, GPU nodes, and additional shared/login nodes for Jupyter (currently three are set aside on Cori).

NERSC closely monitors the state of the art in deep learning applications through user events, engagement, and surveys, in addition to having heavy involvement in projects using state-of-the-art techniques. To support these applications, we provide optimized installations of the most popular deep-learning frame-works: currently PyTorch and Tensorflow. These frame works are optimized for both single-node performance on CPU (in collabora-tion with Intel) and GPU (in collaboration with NVIDIA) and provide distributed capabilities. Various projects have exploited these up to the full scale of Cori, and performance of these applications is closely tracked with benchmarks.

Distributed deep learning at scale remains an open research area and a challenge for users, so in 2019 NERSC led the organization of the inaugural Deep Learning for Science School, a week-long course in Neural Networks and Deep Learning, especially as applied to scientific problems on HPC systems. In addition, NERSC and the Computing Sciences Area at Berkeley Lab joined forces with the Association for High Speed Computing to organize and support the first Monterey Data Conference in Monterey, Calif. The theme for 2019 was Deep Learning for Science. The program featured talks about the impact of deep learning on a broad set of applications, including precision agriculture, personalized cancer treatment, materials by design, detecting extreme climate events, controlling fusion reactors, managing networks, and tracking particles in advanced physics experiments. Panel discussions from industry and DOE Labs explored hardware, software, and methods challenges.

During 2019 NERSC also hosted tutorials on deep learning at scale at SC, ISC, ECP, CUG, and the annual meetings of the American Geophysical Union and the American Meteorological Society. In collaboration with Intel, NERSC has also compiled related guidelines, which are available at https://docs.nersc.gov/machinelearning/distributed-training/.

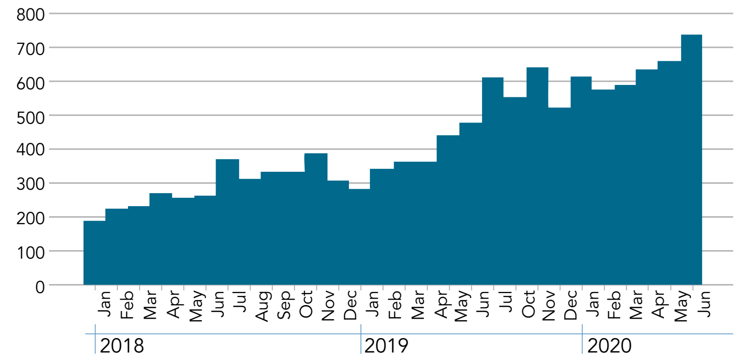


Figure 4. Unique Jupyter users on NERSC systems, January 2018 through June 2020.

NERSC ADDS ITS EXPERTISE TO EXASCALE COMPUTING PROJECT

Established in 2016, the Exascale Computing Project (ECP) is a multi-agency effort within the DOE designed to ensure that the U.S. maintains leadership in high performance computing and high-speed networking for science, innovation, and national security.

As the project matures, NERSC staff continue to be actively involved on multiple levels; NERSC staff currently hold several key positions in the ECP organization:

- **Katie Antypas** is director of the Hardware and Integration focus area
- **Jack Deslippe** is the L3 Application Development lead for Materials Science and Chemistry
- **Debbie Bard** is an L4 lead for Application Integration at the Facilities, representing pre-exascale facilities.

In addition, five ECP applications are integrated into NERSC's NESAP program: AMReX, ExaFEL, EXAALT, WDMApp, and ExaBiome. NERSC is also deeply involved in the ExaI/O project — along with colleagues from LBNL, The HDF Group, and ANL — and is working with the UnifyFS team at LLNL and ORNL. The I/O work also supports a number of ECP-related projects in adaptive mesh refinement, natural hazards, particle accelerators, subsurface wellbores, climate, and ExaFEL. NERSC is also at the forefront of efforts to develop software integration and continuous integration capability for ECP software teams, as well as continuous integration and Spack software integration development. NERSC is the only center with ECP continuous job execution "runners" able to access its production computing system, Cori.

In 2019 NERSC also focused on developing Technical Execution Plans for work to be carried out for the duration of the ECP program, and profiling and optimizing ECP codes. As part of the project's hardware and integration, application performance engineers are now in place at NERSC, ALCF, and OLCF. NERSC also participates in ECP through facilities coordination efforts, including the ECP annual meeting, which attracts hundreds of representatives from DOE national labs, universities, and industry; and the creation and management of the DOE SC Facilities Engagement Plan with the ECP.

Optimizing Operations

Iris: A New System for Account and Identity Management

Iris, a new identity management system and user interface at NERSC that enables users to manage their accounts, projects, and personal information, went live in December 2019, replacing its predecessor, the NERSC Information Management system that was originally deployed in 1999. A ground-up redesign, Iris provides users with job-level allocation tracking; improved reporting and visualization; and a fast, accessible, and responsive user interface that works on computers, tablets, and smart phones.

Initial development of Iris began in mid-2018 with requirements gathering and technology selection. In March 2019, NERSC conducted a technical design review with peers from ALCF and OLCF, as well as the LIGO facility, which was able to offer special insight to NERSC after having recently implemented a system with architectural similarities.

The final stages of development and testing continued through the summer and fall of 2019 and followed an agile software engineering methodology to allow for active feedback and testing during the coding process. A pre-release of Iris was demonstrated during a NERSC User Group webinar; this demonstration was repeated shortly afterward for DOE allocation managers, giving opportunities for further feedback. Iris went live on December 1, 2019, giving ample time for issues to be identified and resolved before the allocation year transition in mid-January 2020.

All Iris components are based on modern, mature, open-source technologies with large user bases. For instance, the web front-end leverages a popular web framework, reducing lines of code written by

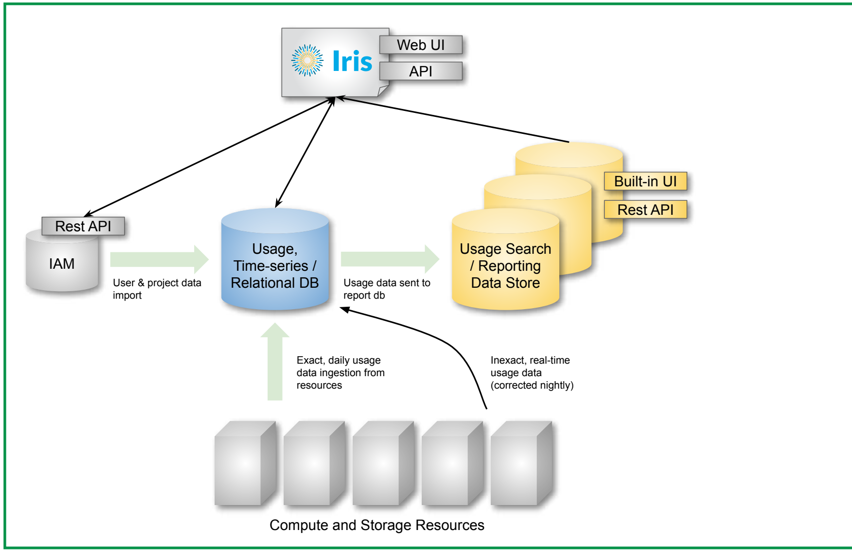


Figure 5. This Iris diagram illustrates how data flows from IAM, compute, and storage resources and is used to populate time-series and report databases.

a factor of 10 compared to NIM; the underlying PostgreSQL open-source database with a time-series extension, instead of NIM's Oracle database, improves performance dramatically for users while saving thousands of dollars in annual licensing costs over NIM's commercial database implementation. The project has benefited throughout its early lifecycle from documentation and easy availability of online support for these and other architectural underpinnings.

The Iris system is composed of the following parts:

- Web front-end built with the React framework, loosely coupled with the application backend via API
- API backend, which uses GraphQLProcessing workflows for user state management, accounting rollups, etc.
- Datastores optimized for different types of queries
- Identity and Access Management (IAM) components from the InCommon Trusted Access Platform (TAP), a widely adopted toolset in research and academia for managing identities, group authorization and collaboration data.

The design of Iris, and in particular the TAP suite of underlying system components, also positions NERSC to support federated identity and federated authentication systems in the future, which will help to simplify authentication and interoperability for distributed workflows and experimental and observational science projects. With an improved data model, more granular usage tracking, and modern tools for visualization, Iris is beginning to have a transformational effect by giving users and staff clearer insight into the use of NERSC resources.

ReFrame: Improving Reliability with Automated Testing

Regression testing is critical to ensuring a stable and reliable user experience on NERSC resources. For many years NERSC has used an informal internal testing infrastructure, and we realized we needed to take the next step and standardize the testing to enable analysis of the results as a single, collective data set.

The Swiss National Supercomputing Centre (CSCS) has developed a Python-based framework called ReFrame to address precisely these challenges. Following ReFrame's publication, NERSC recognized that adopting ReFrame could reduce testing complexity, consolidate disjointed testing efforts, and automate logging test results. We began developing NERSC-specific ReFrame tests in early 2019. In less than one year, ReFrame has replaced much of the older testing infrastructure, reduced the time to evaluate system health following a maintenance, and detected several bugs in Cray system software.

NERSC's adoption of ReFrame has enhanced system stability and monitoring efforts in several ways. First, by merging most of the existing system tests into the same Python framework, it has reduced test fragmentation and redundancy and become the single "point of truth" for system regression tests. It has also significantly improved the "system checkout" process whereby, following a maintenance, reboot, or upgrade, Cori's different user-facing components are evaluated to ensure they work the same way (or better than) they did before. Additionally, NERSC has implemented workflow tests from other user facilities and DOE experiments, including JGI and DESI.

These workflows are sensitive to many different Cori components working correctly simultaneously, and so are invaluable additions to the NERSC test repository. Finally, NERSC has upstreamed several contributions to the CSCS ReFrame repository on GitHub, which benefit any center operating a Cray HPC system.

An additional essential element of system regression testing is automation. This applies particularly to performance tests, which evaluate components of Cori that can change dynamically without a disruptive event like a maintenance or reboot. NERSC has automated ReFrame performance tests by integrating them into NERSC's internal GitLab continuous integration system, which is fully integrated into the Slurm workload manager thanks to enhancements from Exascale Computing Project.

SENSE: Software-defined Network for End-to-end Networked Science at Exascale

The SENSE (Software-defined network for End-to-end Networked Science at Exascale) project led by ESnet is a key part of NERSC's larger superfacility plan. NERSC is leveraging SENSE as a gateway service to help large scientific facilities streamline and automate their data movement workflow to NERSC.

The SENSE project provides the provisioning of multi-domain, multi-resource, end-to-end Layer 2 and 3 services, including end-site resources such as data transfer nodes (DTN) — servers in a data center that are optimized for bandwidth and tuned to transfer data efficiently. The SENSE Layer 2 services include deterministic end-to-end resource guarantees and automated DTN private address configuration. The SENSE Layer 3 service provides mechanisms for directing traffic onto a specific Layer 3 virtual private network for policy or quality-of-service reasons. The system also allowsworkflow middleware to redirect traffic at granularities of a single flow, a specific end-system, or an entire end-site onto a SENSE provisioned service.

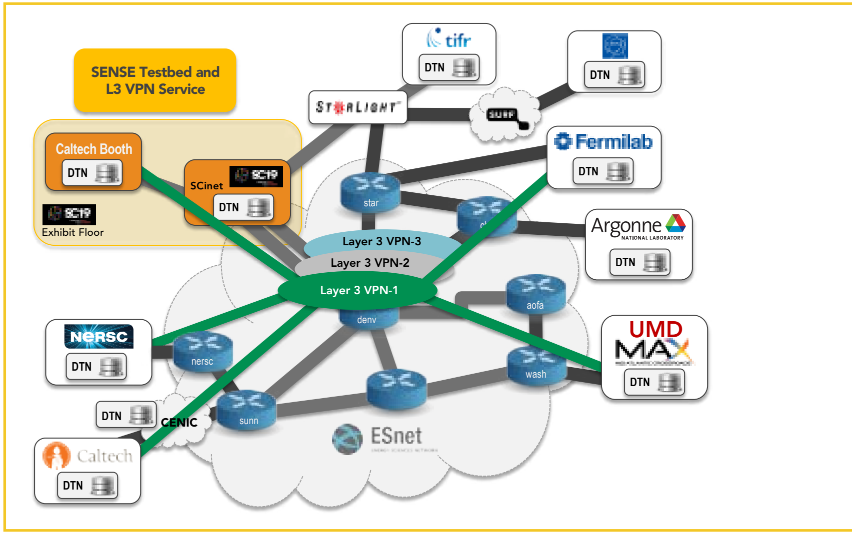


Figure 6. Schematic of the Software-defined Network for End-to-end Networked Science at Exascale (SENSE) project.

NERSC has been integrating SENSE to perform data movement directly to the DTNs for quick and efficient caching of data to the supercomputers. In NERSC's next-generation supercomputer, Perlmutter, the DTNs will be tied directly to the internal interconnect of the supercomputer, thereby allowing streamed data to be as logically close as possible to the compute partition.

Projects such as LCLS-II will be utilizing SENSE to transfer their experimental data into NERSC. SENSE will help expedite the transfer of data by signaling both the endpoint (NERSC in this case) and the intermediate network (ESnet) about the data transfer. This allows for ESnet to provision a link with all the necessary characteristics (bandwidth, jitter, latency) for the data transfer while also allowing NERSC to ensure that the data lands on the correct endpoint (e.g., least loaded, closest to the compute, etc.). The ability to give hints to the network and storage for large transfers allows the system to better utilize all the available resources.

Figure 7. A new detector is being built for NCEM's Transmission Electron Aberration-corrected Microscope 0.5 (TEAM 0.5) that allows researchers to access single-atom resolution for some samples. The new detector, unveiled in February 2019, will generate 4 terabytes of data per minute.

The NCEM-to-NERSC Data Pipeline

The National Center for Electron Microscopy (NCEM) at LBNL has started work on a new technique in electron microscopy called 4D-STEM, a term that means "2D raster of 2D diffraction patterns using scanning transmission electron microscopy." While this technique shows promise, it is hindered by the immense amounts of data that the process generates and needs to analyze and store in near real time.

The NCEM data pipeline is designed to have the 4D-STEM processing done on FPGAs attached to the imaging boards co-located with the electron microscope. Work to implement data processing on FPGAs is still in progress, so NCEM is using compute resources at NERSC to analyze the data in near real time. NCEM is on the Berkeley Lab campus, and this proximity provides a unique opportunity to test high flow rates in preparation for what will become common with the next generation of ESnet. To facilitate this, NERSC provisioned a temporary dedicated network path that connects directly into the same network that supports NERSC's Cori supercomputer. This allows NCEM to stream data from the

4D-STEM instrument directly into NERSC, bypassing normal routes that would take the data through Sunnyvale (a city more than 50 miles south of Berkeley) before it arrived back at NERSC.

Performance Modeling: Roofline Goes to the Next Level

The Roofline Model, developed by Berkeley Lab researchers, helps supercomputer users assess the performance of their applications by combining data locality, bandwidth and parallelization paradigms into a single figure that exposes performance bottlenecks and potential optimization opportunities. NERSC has been working with NVIDIA to create a methodology for Roofline data collection on NVIDIA GPUs, and a set of performance metrics and hardware counters have been identified from the profiling tools, nvprof and Nsight Compute, to construct a hierarchical Roofline. This helps users gain a holistic view of their application and identify the most immediate and profitable optimizations. The methodology has been validated with applications from various domains, including material science, mesh and particles-based codes, image classification and segmentation, and natural language processing.

The methodology is still being improved, and integration of Roofline into NVIDIA's Nsight Compute tool is also in discussion. The goal is that an automated Roofline capability will be added to Nsight Compute so that data collection and performance modeling can be done with ease and production quality for all kernels in a large-scale HPC application.

GPCNet Benchmark Targets Network Congestion

Congestion on the internal high-speed network is one of the biggest problems facing HPC systems today, affecting system throughput, performance, user experience, and consistency in time to solution. Congestion manifests as run-to-run variability due to contention for shared resources or routes between compute endpoints.

In collaboration with Cray and the ALCF, NERSC developed a new open-source benchmark suite called the Global Performance and Congestion Network Tests (GPCNeT) to advance the state of the practice in this area. The benchmark was designed to capture representative HPC traffic as well as adversarial traffic conditions. The output of the benchmark quantifies the performance degradation observed in communication in the presence of congestion. The team ran this benchmark across a wide set of DOE systems, demonstrating the congestion impact as well as the benefits of congestion management in future systems with the Slingshot network. (The congestion impact indicates the factor of slowdown for a given test — latency, bandwidth, and allreduce.)

For more details, see our publication and Github repository: "GPCNeT: designing a benchmark suite for inducing and measuring contention in HPC networks," in Proceedings of the International Conference for High Performance Computing, Networking, Storage and Analysis (SC19): https://github.com/netbench/GPCNET.

Advanced Light Source

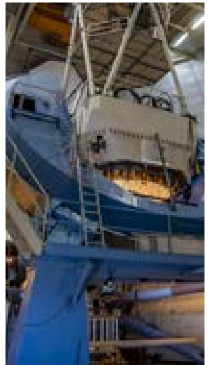
LZ Dark Matter Experiment

Facilities Integrating Collaborations for User Science

Molecular Foundry

Dark Energy Science Collaboration

Linac Coherent Light Source



The Dark Energy Spectroscopic Instrument

NERSC'S SUPERFACILITY PROJECT GAINS MOMENTUM

Large-scale analysis of experimental data has revolutionized our understanding of the physical world, from the discovery of elementary particles and the accelerating expansion of the universe to new insights into the microbiome and the ability to better predict extreme weather events.

Berkeley Lab's Computing Sciences Superfacility Initiative is a framework for further integrating experimental and observational instruments with computational and data facilities, bringing the power of exascale systems to the analysis of real-time data from light sources, microscopes, telescopes and other devices. Data from these experiments will stream to large computing facilities where it will be analyzed, archived, curated, combined with simulation data and served to the science user community via powerful computing, storage, and networking systems.

In support of this concept, 2019 NERSC launched the internal Superfacility Project, an effort to accelerate pipelines from DOE experimental and observational facilities in an automated, high-performing manner. NERSC's superfacility team held a number of outreach events over the past year and has enabled capabilities that improve data movement and sharing for experimental facilities. Current superfacility engagements include the ALS, LCLS, and National Center for Electron Microscopy (BES); DESI, LZ, and Legacy Survey of Space and Time Dark Energy Science Collaboration (LSST-DESC, HEP); and the Joint Genome Institute (BER). NESAP efforts addressed code performance issues of ALS users, the DESI and LZ experiments, and LSST-DESC specifically in preparation of Perlmutter.

The Superfacility Project consists of four areas of work:

- Applications requirements and deployment, including targeted science engagements, Jupyter, policy, scalable code development via NESAP, and user outreach.
- Scheduling and middleware, including Spin, federated ID, workflow resiliency, scheduling and an API interface to NERSC.
- Automation, including software defined networking (SDN), self-driving facilities, and wide-area network (WAN) projects.
- Data management, including internal data movement and the data dashboard.

Targeted science engagements are a key element of the Superfacility Project. The project has selected seven strategic partnerships to drive our work, chosen because they use and challenge NERSC in different ways. The project is therefore structured to ensure that our work will be generally useful across multiple science domains, so that rather than designing one-off solutions we are developing a reusable and extensible toolkit. To date, the Superfacility Project has made significant progress in all of these areas, including:

Enhanced Globus Sharing: NERSC has deployed Globus Sharing to allow NERSC users to easily share large volumes of data with the wider scientific community. Using a Globus Sharing endpoint, scientists can control what data they want to share and who they want to share it with. In 2019, 275 TB of data was shared using Globus Sharing at NERSC. Beamline scientists at the ALS used the Globus Sharing endpoint for simultaneously distributing 7 TB of data with users from UCLA and CU Boulder while the experiment was still ongoing. All transfers were user-initiated and involved minimal NERSC staff intervention. NERSC's Globus Sharing is now routinely used for data distribution at ALS beamlines.

Workflow resiliency: In 2019 the Superfacility Project engaged closely with our partner teams to ensure they have appropriate plans in place for operations when NERSC is unavailable. For example, NERSC team members worked closely with LZ to include resiliency as a topic in their 2020 computing readiness review. NERSC also led an application for compute time under ALCC across the three ASCR HPC facilities, with the aim of providing compute time for portable workflows. NERSC has also focused on making its own systems more resilient.

Data management and movement: NERSC deployed command-line tools for large-scale data movement using Globus. Intended for optimized movement between the different file systems at NERSC, this command-line interface can be incorporated into users' workflows. We are also exploring integrating our HPSS file system with our Spectrum Scale file system to provide an easier-to-use HPSS interface. The features of NERSC's data dashboard were enhanced to include a sunburst visualization of a user's data on the Community File System, and a prototype "PI Dashboard" that allows PIs to do simple changes to their data was tested with friendly users.

HDF5 extensions: As part of a jointly funded ASCR/BES grant, the LLAna pilot project is funding research into tools and technologies to support data analysis from LCLS-II. One of the research thrusts in LLAna involves changes to HDF5 to support two new functionalities: a virtual object layer that will allow LCLS-II data to be read into HDF5 software, opening the way for HDF5-designed tools to be used to analyze XTC2 data directly; and support in HDF5 for writing variable-length data.

NERSC'S RESPONSE TO CALIFORNIA'S MANDATED POWER SHUTDOWNS

In October 2019, Berkeley Lab experienced two Public Safety Power Shutdown (PSPS) events due to California wildfires. PSPSs are part of a new policy implemented by California's primary power supplier, Pacific Gas & Electric (PG&E), that calls for large-scale power shutdowns during hot, windy weather to help prevent wildfires.

During the first PSPS event, NERSC used generator power to keep login nodes, the Spin edge-services infrastructure, and some file systems available to users. Although rising temperatures during this event eventually forced NERSC to shut down Cori's login nodes, during the 36 intervening hours almost 400 users logged into Cori, and NERSC received high praise from users for keeping data accessible. During the second event, NERSC performed a full shutdown.

In response to each event, Berkeley Lab activated its Emergency Management Team (EMT) to monitor the impact of the power outage on critical laboratory infrastructure and equipment, ensure the safety of laboratory personnel, and plan for the return of power to the laboratory. Several NERSC staff are part of the Lab EMT, serving as communications experts and subject matter experts for the larger Computing Sciences area.

Because no power was available on Lab property, Berkeley Lab was formally closed. Any staff who could work from home were required to do so if possible. Critical staff (primarily operations and a few leadership personnel) remained on site to monitor the equipment. The Operations Technology Group served as the communications hub for NERSC during the event to ensure everyone was kept up to date on the status of equipment and each PSPS event. Daily stand-ups were held to provide updates and plan for the return to service.

After the power lines were reenergized, Berkeley Lab personnel restored power systematically to one building at a time. The building housing NERSC was the first to have its power restored due to NERSC's national importance.

Based on learnings from the 2019 PSPS events, we are hardening our infrastructure to maximize the number of subsystems that can remain up during the event. The laboratory as a whole is undertaking a utility resilience program to harden lab-wide infrastructure and to work with PG&E to minimize the impact and extent of PSPS events. Many of the infrastructure hardening and mitigations are expected to be in place by the fall 2020 fire season.

Cybersecurity

NERSC's security strategy enables science through a focus on system-level security, in-depth monitoring and intrusion detection, and vulnerability management. NERSC's security group maintains a strong security culture through a good working relationship with the various NERSC technical groups. The NERSC Security Team, consisting of representatives from all NERSC stakeholders, meets regularly to coordinate tactical responses to new threats and communicate out to technical staff.

NERSC separates public (anonymous) services into a separate Science DMZ and keeps those services partitioned off from staff and administrative hosts behind a firewall. Staff and visitor workstations and laptops are connected to the Berkeley Lab network and are covered under the Berkeley Lab security plan.

Multi-factor Authentication: A Success Story

One contributing factor to NERSC's ability to minimize security incidents is last year's deployment of multi-factor authentication (MFA). By requiring essentially all NERSC users to use MFA, NERSC has been able to mitigate one of the most significant risks in our threat model; specifically, that of stolen credentials. In the past, the vast majority of reportable security incidents at NERSC have been the result of stolen credentials. With MFA fully deployed, the NERSC security team has been able to confirm that attackers have

Figure 8. Each November, Berkeley Lab participates in the DOE's Annual CyberForce Competition, a multi-lab event designed to inspire the next generation of energy sector cybersecurity professionals. The competition was established in 2014 to address the cybersecurity capability gap in the U.S. and increase hands-on cyber education for college students, awareness of critical infrastructure, and basic under-standing of cybersecurity in a real-world scenario.

been denied access to NERSC systems despite compromising external user accounts. This is due to the fact that a stolen password is no longer sufficient to gain access to NERSC systems.

NERSC's deployment of MFA began in early 2018 as a security-enhancing feature that users could optionally choose to enable. In mid-2018 NERSC made the decision to require MFA for all users beginning with the allocation year rollover in early January 2019. NERSC successfully transitioned all but a handful of users to using MFA at the new allocation year. Both the design of the MFA implementation and the plan to transition users contributed to NERSC's success in rolling out this new feature.

Vulnerability Management

NERSC has implemented a vulnerability management program in which vulnerabilities are identified and their risks evaluated. This evaluation leads to identifying, tracking, and correcting the vulnerabilities and removing or accepting the risk after evaluation.

NERSC performs a periodic vulnerability scan of the whole network in addition to scans on demand for the different stakeholders. In addition to regular scanning, we encourage staff in different teams to use the scan infrastructure in order to perform their own vulnerability assessments. NERSC has deployed Tenable. sc as its primary scanning tool; it is supplemented by secondary tools like SQLmap to focus on specific vulnerabilities. We have developed our own software to automate security analysis efforts.

The vulnerability remediation process promotes a strong cybersecurity culture at NERSC. It includes presenting the information to a multidisciplinary audience and helping system engineers in the remediation process, and taking appropriate action if necessary.

Another critical component of NERSC's cybersecurity arsenal is the ability to monitor activity on user-facing systems. As part of this process, in 2019 NERSC partnered with a startup company, Prismo Systems, to test and further develop the technique of monitoring computational systems by instrumenting the operating system kernel and the system calls executed by every process running on the system. Prismo's system collects the data coming from these sensors, analyzes it, correlates activity of users as they move from one system to another, and presents the data in a way that is most useful for security analysts. NERSC has deployed these sensors on Cori service nodes and continues to work with Prismo to improve the tool, making it more effective and efficient. Ultimately, we will make this data available for analysis and for applying machine learning algorithms to detect anomalous behavior.


CRAY


Cori

DATA CENTER OPTIMIZATION: NERSC EFFICIENCY METRICS PROJECT

Figure 9. The NERSC cooling towers next to Shyh Wang Hall on the Berkeley Lab campus.

For decades, Lawrence Berkeley National Laboratory has been at the forefront of efforts to design, build, and optimize energy-efficient hyperscale data centers. The Lab's Energy Technologies Area (ETA) has supported NERSC and other Berkeley Lab facilities on energy issues for some 30 years.

In 2007, for example, Berkeley Lab was lead author on a U.S. Environmental Protection Agency (EPA) study of the national impact of data center energy use and on the resulting DOE report. In addition, ETA has been involved in EPA's Energy Star program for more than 25 years, developing specifications for servers, data storage, and network equipment. Much of this work has been funded through the Federal Energy Management Program, including establishing a Center of Excellence at Berkeley Lab that serves as a repository of knowledge and consulting for federal agencies on data centers in energy and water management.

More recently, this expertise has been applied to the design, construction, and operation of Berkeley Lab's Shyh Wang Hall and the computing, storage, and networking systems hosted in the building's data center. These efforts also led to the NERSC Efficiency Optimization Project, a collaboration between NERSC, ETA, and Sustainable Berkeley Lab that is applying emerging operational data analytics (ODA) methods to optimize cooling systems and save electricity.

One of the key performance metrics for any data center is power utilization effectiveness (PUE), which measures the ratio of total data center electricity draw to the electricity used directly for computing only. Shyh Wang Hall was designed to have a target PUE of less than 1.1. This means the support and cooling of the computing equipment was just 10% or less of the direct computing energy use, putting NERSC on par with the most efficient data-processing facilities in the world. At present, NERSC's PUE is 1.08.

But the Berkeley Lab team felt there was room for improvement. So NERSC, ETA, and Sustainable Berkeley Lab set out to find additional ways to further optimize performance. In 2016, working with the mechanical engineering firm kW Engineering, a number of potential savings opportunities were identified, and the team began analyzing and applying the metrics to various systems and processes in the NERSC data center, according to John Elliott, chief sustainability officer at Berkeley Lab. To date the overall energy savings realized is about 37% of the non-IT electricity consumption at NERSC. In addition, the project has saved 1.8 million kWh and more than a half million gallons of water per year.

These efforts included implementing all of the metrics into a system called SkySpark, which collects and analyzes building data and sets up calculated points derived from that data. SkySpark works in conjunction with NERSC's OMNI data collection system, a platform of applications that combines HPC and IT systems data with cooling and facility systems performance data from the Building Management System, plus supplemental rack-level IT sensors and computer syslog data. Together, these systems enable NERSC's Energy Efficiency team to use sophisticated ODA tools to continuously monitor operations and implement energy efficiency reliability measures. The initial emphasis has been on the facility's cooling systems. For example, most facilities calculate PUE as an annual average. But NERSC can now consistently calculate this metric every 15 minutes.

Another key achievement has been the creation of active communications between NERSC's Cori supercomputer's internal blower fans and the cooling towers and pumps of the traditional cooling plant. The Cray XC Dynamic Fan Speed Control feature automatically varies the cabinet blower fan speeds based on processor temperatures. This new ODA link method provides continuous feedback, which enables the Cray system and building controls to shift the load ratio between the blower fans and the cooling water loop depending on outdoor environmental conditions. And they are in constant communication every 15 minutes so that they are always working together.

This kind of internal communications between systems and facilities is the wave of the future, providing the opportunity to merge the compute systems and the facility plant silos into a unified holistic system and thus advance the state of computing science.

Science Highlights

NERSC runs highly available, stable, complex, world-class HPC resources with extremely usable development and software environments, allowing its thousands of users to be very scientifically productive. With more than 2,400 publications in 2019, we can only share a sample of NERSC users' science here. The following were chosen to represent the breadth of scientific research and data-focused projects supported by NERSC.

How California Wildfires Can Impact Water Availability

THE SCIENCE. Berkeley Lab researchers used NERSC supercomputer resources to show that conditions left behind by California wildfires led to greater winter snowpack, greater summer water runoff, and increased groundwater storage.

THE IMPACT. In recent years, wildfires in the western United States have occurred with increasing frequency and scale. Even though California could be entering a period of prolonged droughts with potential for more wildfires, little is known about how wildfires will impact water resources. This study provided critical new insights for planners and those who manage California's water resources.

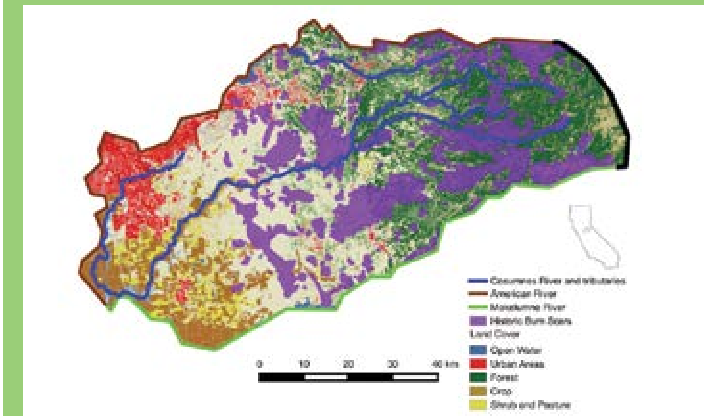
SUMMARY. The researchers modeled California's Cosumnes River watershed, which extends from the Sierra Nevada mountains down to the Central Valley, as a prototype of many other California watersheds. Using about 3 million hours on NERSC's Cori supercomputer to simulate watershed dynamics over a period of one year, the researchers were able to identify regions most sensitive to wildfire conditions, as well as the hydrologic processes that are most affected.

NERSC PI: Erica R. Siirila-Woodburn, Lawrence Berkeley National Laboratory

PROJECT FUNDING: Berkeley Lab Early Career LDRD

ALLOCATION AWARD: DOE Office of Science Biological and Environmental Research

PUBLICATION: Maina, F.Z. and Siirila-Woodburn, E.R., "Watersheds dynamics following wildfires: Nonlinear feedbacks and implications on hydrologic responses," *Hydrological Processes*, 34, 33 (2020). [DOI: 10.1002/hyp.13568]

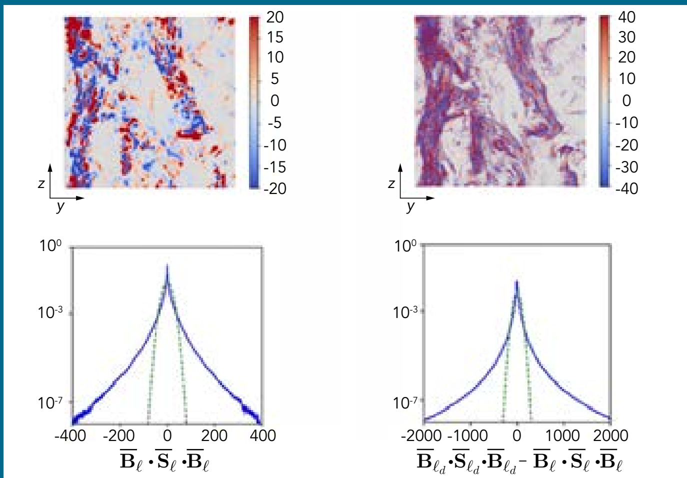


Berkeley Lab researchers built a numerical model of the Cosumnes River watershed, extending from the Sierra Nevada mountains to the Central Valley, to study post-wildfire changes to the hydrologic cycle.

New Conservation Laws in Turbulent Magnetized Flows

THE SCIENCE. University of Rochester researchers used a novel coarse-graining framework for disentangling multiscale interactions and discovered two separate conservation laws over a large range of length scales in turbulent magnetized plasma flows.

THE IMPACT. Flows coupled to magnetic fields, known as magnetohydrodynamic (MHD) flows, are central to our understanding of an astonishing variety of physical systems, including galaxies and galaxy clusters, gaseous nebulae and the interstellar medium, stellar evolution, solar winds and space weather, nuclear fusion, and metallurgy. This study identified magnetic and kinetic energies as separate dynamical invariants over a range of scales in turbulent MHD flows; such invariants are highly prized physical quantities vital to understanding the evolution of dynamical systems. These findings have broad implications for understanding MHD flows, including understanding magnetic reconnection in solar flares, amplification of magnetic field strength in dynamo action, and the physics of astrophysical accretion disks.



Using a novel multiscale analysis framework and massively parallel simulations, UofR researchers were able to disentangle the energy pathways in turbulent magnetized flows and find that there are two separated conservation laws instead of just one over an entire range of scales.

SUMMARY. This study relied on a suite of massively parallel simulations run at NERSC using the DiNuSUR code developed at the University of Rochester to perform careful convergence studies and explore the robustness of the results in systems under different physical conditions. Simulations used up to 8.5 billion grid points on 16,384 compute cores over a month's worth of computing, consuming about 6.3 million CPU hours and generating 6TB of data. NERSC staff also provided I/O troubleshooting for this particular project.

NERSC PI: Hussein Aluie, University of Rochester

PROJECT FUNDING AND ALLOCATION AWARD: DOE Office of Science Fusion Energy Sciences

PUBLICATION: Bian, X. and Aluie, H., "Decoupled Cascades of Kinetic and Magnetic Energy in Magnetohydrodynamic Turbulence," *Physical Review Letters*, 122, 135101 (2019).
[DOI: 10.1103/PhysRevLett.122.135101]

Simulating Vulnerability to Antarctic Regional Ice Shelf Collapse

THE SCIENCE. A team led by Berkeley Lab performed the first fully resolved systematic study of millennial-scale ice-sheet response to regional ice-shelf collapse based on all major Antarctic drainage basins and the resulting potential for large contributions to sea level rise.

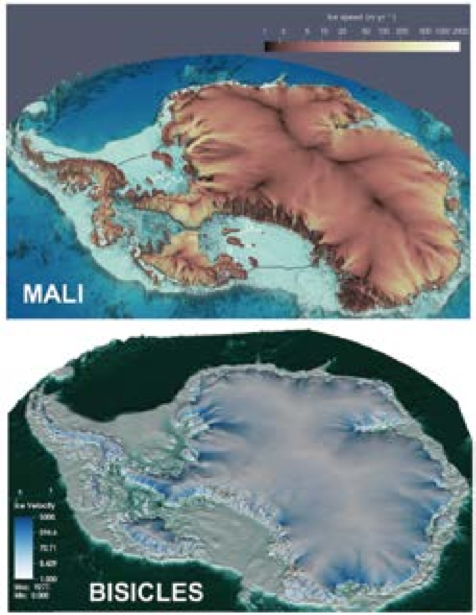
THE IMPACT. The researchers found that sustained ice-shelf loss in any of the Amundsen Sea, Ronne, or Ross sectors can lead to wholesale West Antarctic ungrounding and collapse. However, even with unrealistically extreme forcing, loss is relatively modest for the initial century, increasing markedly afterward in West Antarctic collapse scenarios. The results indicate that Antarctic drainage basins are dynamically independent for 1-2 centuries, after which dynamic interactions between basins become increasingly important and regional modeling results will be increasingly inaccurate.

SUMMARY. The study divided Antarctica into 14 sectors corresponding to large-scale drainage basins, simulating thinning to each sector's floating ice shelves and running the high-resolution BISICLES ice flow model 1,000 years into the future for each case. This work, entailing 35,000 years of Antarctic simulations, was enabled by the combination of DOE SciDAC-funded BISICLES ice sheet model (part of the ProSPect partnership) and NERSC's computational resources. BISICLES uses adaptive mesh refinement (AMR) to deploy sufficient (~1km or finer) model resolution where needed to resolve grounding-line dynamics.

NERSC PI: Esmond Ng, Lawrence Berkeley National Laboratory

PROJECT FUNDING AND ALLOCATION AWARD: DOE Office of Science Advanced Scientific Computing Research

PUBLICATION: Martin, D.F., Cornford, S.L., and Payne, A.J., "Millennial-scale Vulnerability of the Antarctic Ice Sheet to Regional Ice Shelf Collapse," *Geophysical Research Letters*, 46, 1467 (2019). [DOI: 10.1029/2018GL081229]

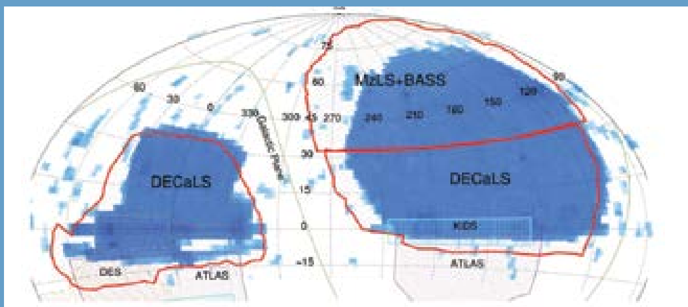


Simulated Antarctic ice sheet 200 years after all floating ice shelves are instantaneously removed.

Pointing DESI in the Right Direction

THE SCIENCE. In 2020, the Dark Energy Spectroscopic Instrument (DESI) will begin creating the largest 3D map of the universe's distribution of "ordinary" matter through a massive campaign of spectroscopic observations of distant galaxies. To prepare, the DESI team has been using NERSC to build catalogues of the most interesting observational targets, modeling the shapes and colors of more than 1.6 billion individual galaxies detected in 4.3 million images collected by three large-scale sky surveys.

THE IMPACT. The statistical distribution of galaxy separations has a preferred length scale that is governed by cosmological physics, so measuring that length scale (a "standard ruler") constrains the fundamental nature of dark energy, a mysterious form of energy that drives the accelerating expansion of the universe and is unlike "ordinary" matter of people, places and things. The nature of dark energy is one of the deepest mysteries of modern physics. Before it could begin operations, this phase of the DESI project needed to know to exquisite precision where to position each of its 5,000 fibers per pointing to collect each galaxy's spectrum, and avoid spending precious observing time on the wrong kinds of objects. The most recent catalogue of 2D sky positions and object descriptions was created at NERSC in 2019; the final catalogue will be prepared in 2020.



This map shows the sky areas covered (blue) by three surveys conducted in preparation for DESI.

SUMMARY. The DESI Legacy Imaging Surveys have performed all their catalogue generation at NERSC over the course of eight data releases (DRs) over the past several years. DR8 was the first to perform a combined analysis of all three surveys. In addition to using all the available data, DR8 involved significant code improvements over DRs 6 and 7 and better consistency between reductions and calibrations. More than 10 million NERSC hours on both Cori Haswell and KNL partitions were used, and the resulting rendered images, models, and catalogues are available at https://legacysurvey.org (see also its Google Maps-like viewer). The data are hosted at NERSC in the Cosmology Data Repository, which contains about 800 TB of cosmological data. NERSC staff provided the project with Spin training and support, and worked with Cray to resolve file system issues that the DESI data processing pipeline encountered.

NERSC PI: David Schlegel, Lawrence Berkeley National Laboratory

PROJECT FUNDING AND ALLOCATION AWARD: DOE Office of Science High Energy Physics

PUBLICATION: Dey, A., et al., "Overview of the DESI Legacy Imaging Surveys," *The Astronomical Journal*, 157, 168 (2019). [DOI: 10.3847/1538-3881/ab089d]

Connecting Supernova Composition Models with Progenitors

THE SCIENCE. Researchers from the University of California, Santa Cruz used simulations run at NERSC to interpret a new, growing data set of a newly discovered type of supernova explosion (Type "Iax") and constrain its chemical composition. This study was the first of its kind to analyze all of the available data for a particular type of supernova.

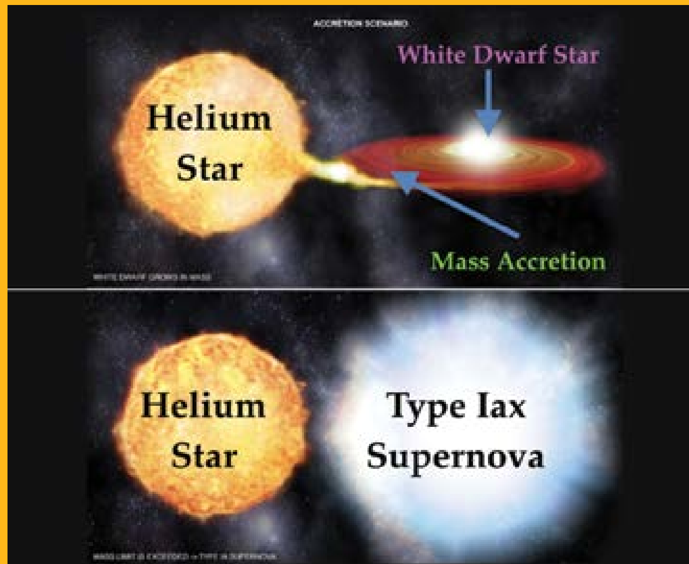
THE IMPACT. These supernovae and their more energetic cousins (Type "Ia") are thermonuclear supernovae, thought to arise from the explosion of at least one white dwarf in a binary star system and responsible for the synthesis of heavy elements such as iron, nickel, and cobalt. Understanding these supernovae is critical to under standing the formation of the elements and their dispersal into the interstellar medium and the evolution of binary star systems and compact objects, and they can provide insight into the as yet mysterious thermonuclear explosion mechanism.

SUMMARY. The team systematically searched for signatures of helium in Type Iax supernova spectra to constrain the proposed white dwarf + helium star progenitor channel for these objects. Over the course of a six-month investigative campaign, the team used the SYNAPPS MPI+OpenMP-parallel fast supernova spectrum synthesis code on Edison to fit 110 Type Iax spectra from 44 such objects. NERSC staff helped build the SYNAPPS code and provided consulting expertise to help the users organize their modeling campaign.

NERSC PI: Daniel Kasen, Lawrence Berkeley National Laboratory

PROJECT FUNDING AND ALLOCATION AWARD: DOE Office of Science Nuclear Physics

PUBLICATION: Jacobson-Galan, W., et al., "Detection of circumstellar helium in Type Iax progenitor systems," *Monthly Notices of the Royal Astronomical Society*, 487, 2538 (2019). [DOI: 10.1093/mnras/stz1305]



Type Iax supernova explosions are thought to arise from a binary system of a white dwarf and helium star. To confirm this progenitor channel, NERSC simulations modeled the elemental composition of all known objects within this supernova class. The ions detected in these peculiar supernovae helped to explain the explosion scenario shown above.

Uncovering Uncultivated Microbes in the Human Gut

THE SCIENCE. Researchers from the Joint Genome Institute, the Gladstone Institutes, and the Chan- Zuckerberg Biohub used NERSC to help computationally reconstruct 61,000 microbial genomes from 3,810 publicly available human gut metagenomes, which are datasets of all the genetic material present in a microbiome sample.

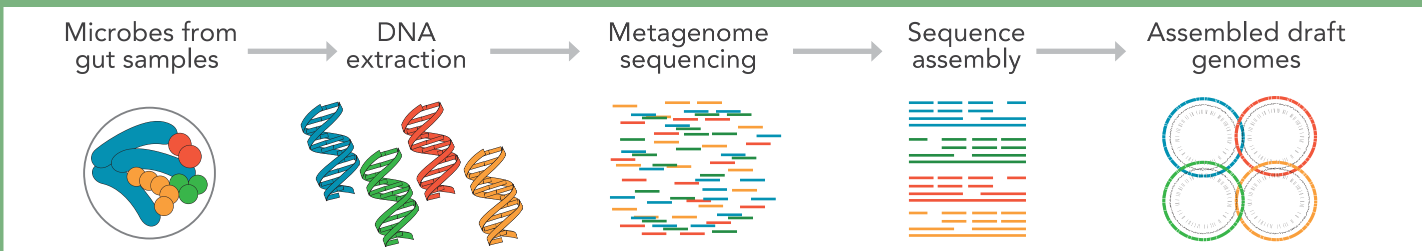
THE IMPACT. Human health is shaped both by environmental factors and the body's interactions with the microbiome, particularly in the gut. Genome sequences are critical for characterizing individual microbes and understanding their functional roles. However, previous studies have estimated that only 50% of species in the human gut microbiome have a sequenced genome, in part because many species have not yet been cultivated for study. This study included metagenome-assembled genomes of 2,058 previously unknown species, bringing the number of known human gut species to 4,558 and increasing the phylogenetic diversity of sequenced gut bacteria by 50%.

SUMMARY. This project extensively used large-scale computing on the Denovo and Cori systems, associated scratch space, and the NERSC Global File System. Nearly 10 million inodes and 40 TB of disk space were needed to manage the file system. Over 1 million compute hours were consumed for the study going back to 2017. The largest jobs used approximately 2,000 cores. NERSC provided extensive consulting support on this project, primarily for resolving issues arising in the installation of scientific software.

LEAD RESEARCHER: Stephen Nayfach, Joint Genome Institute

PROJECT FUNDING AND ALLOCATION AWARD: DOE Office of Science Biological and Environmental Research (BER)

PUBLICATION: Nayfach, S., et al., "New insights from uncultivated genomes of the global human gut microbiome," *Nature*, 568, 505 (2019). [DOI: 10.1038/s41586-019-1058-x]



The general process of reconstructing genomes from metagenomes.

Machine-Learned Impurity-Level Prediction in Semiconductors

THE SCIENCE. Argonne National Laboratory researchers ran high-throughout atomistic simulations on NERSC supercomputers and generated comprehensive computational datasets of impurity properties in two classes of semiconductors: lead-based hybrid perovskites and cadmium-based chalcogenides. These data sets led to machine-learned models that accelerate prediction and design for the entire chemical space of materials and impurities in these semiconductor classes.

THE IMPACT. Impurity energy levels in semiconductors can change their optoelectronic behavior, which has consequences for solar cell applications. The ability to instantly and accurately estimate impurity levels is key, and the current research combines computational data and machine learning to facilitate that. These models can potentially transform the design of novel semiconductors that are defect-tolerant or have tailored impurity properties.

SUMMARY. The researchers performed density functional theory (DFT) calculations for hundreds of impurity atoms in selected semiconductors to determine their formation enthalpies and energy levels. The generated data was transformed into predictive models using nonlinear regression-based machine learning algorithms. The DFT simulations modeled systems containing ~100 atoms over several months, using ~1.5 million CPU hours at NERSC.

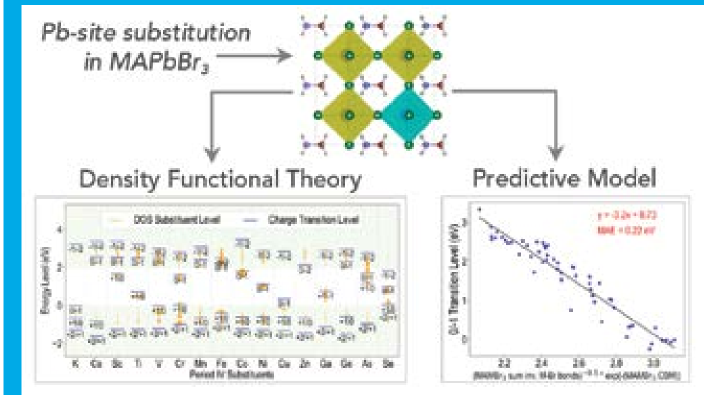
NERSC PI: Maria Chan, Argonne National Laboratory

PROJECT FUNDING: DOE Office Science Basic Energy Sciences and DOE Office of Energy Efficiency and Renewable Energy

ALLOCATION AWARD: DOE Office Science Basic Energy Sciences

PUBLICATIONS: Mannodi-Kanakkithodi, A., et al., "Comprehensive Computational Study of Partial Lead Substitution in Methylammonium Lead Bromide," Chemistry of Materials, 31, 3599 (2019). [DOI: 10.1021/acs.chemmater.8b04017]

Cao, D.H., et al., "Charge Transfer Dynamics of Phase-Segregated Halide Perovskites: CH3NH3PbCl3 and CH3NH3PbI3 or (C4H9NH3)(2)(CH3NH3)(n-1)Pb(n)I(3n+1) Mixtures," *ACS Applied Materials & Interfaces*, 11, 9583 (2019). [DOI: 10.1021/acsami.8b20928]



Impurities at the Pbsite were modeled in a hybrid perovskite material, Methylammonium Lead Bromide, and highthroughput DFT computations were launched on NERSC to calculate their formation enthalpies and energy levels. The computed DFT data has been shown for period IV substituents, and a machine learned model has been shown for one impurity level.

Acronyms and Abbreviations

ALS
Advanced Light Source, Lawrence Berkeley National Laboratory

AMR
Adaptive Mesh Refinement

ANL
Argonne National Laboratory

API
Application Programming Interface

ASCR
Office of Advanced Scientific Computing Research

BER
Office of Biological and Environmental Research

BES
Office of Basic Energy Sciences

BNL
Brookhaven National Laboratory

CERN
European Organization for Nuclear Research

CMB
Cosmic Microwave Background

CPU
Central Processing Unit

CSCS
Swiss National Supercomputing Centre

DESI
Dark Energy Spectroscopic Instrument

DFT
Density Functional Theory

DTN
Data Transfer Node

ECP
Exascale Computing Project

FES
Office of Fusion Energy Sciences

GB
Gigabytes

Gbps
Gigabits Per Second

GPCNet
Global Performance and Congestion Network

GPU
Graphics Processing Unit

HDF5
Hierarchical Data Format 5

HEP
Office of High Energy Physics

HPC4Mfg
High Performance Computing for Manufacturing

JGI
Joint Genome Institute

KNL
Knights Landing Processors

LANL
Los Alamos National Laboratory

LCLS
Linac Coherent Light Source

LLNL
Lawrence Livermore National Laboratory

LZ Dark Matter Experiment
LUX-Zeplin Dark Matter Experiment

MFA
Multi-Factor Authentication

MHD
Magnetohydrodynamic

NCEM
National Center for Electron Microscopy

NESAP
NERSC Exascale Scientific Application Program

NIM
NERSC Information Management

NOAA
National Oceanic and Atmospheric Administration

NP
Office of Nuclear Physics

OLCF
Oak Ridge Leadership Computing Facility

OpenMP
Open Multi-Processing

OpenMSI
Open Mass Spectrometry Imaging

PB
Petabytes

PNNL
Pacific Northwest National Laboratory

PPPL
Princeton Plasma Physics Laboratory

PUE
Power Usage Effectiveness

SENSE
Software-defined Network for End-to-End Networked Science at Exascale

SciDAC
Scientific Discovery Through Advanced Computing

SDN
Software-defined Networking

SLURM
Simple Linux Utility for Resource Management

TAP
Trusted Access Platform

TB
Terabytes

FOR MORE INFORMATION ABOUT NERSC, CONTACT:

NERSC Communications
Lawrence Berkeley National Laboratory
1 Cyclotron Road
Berkeley, CA 94720-8148
Email: CSComms@lbl.gov

NERSC's Web Site
www.nersc.gov

NERSC Annual Report Editor
Kathy Kincade

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

20-CSAO-9030