NERSC NEWS
NATIONAL ENERGY RESEARCH SCIENTIFIC COMPUTING CENTER
Berkeley Lab
Office of Science
U.S. DEPARTMENT OF ENERGY
JANUARY 2008

Visualizing This

NERSC'S NEW CRAY AND ANALYTICS EXPERTS PRODUCE ANIMATIONS OF CLIMATE MODELS WITH FINE DETAILS

A team of climate researchers who obtained early access to NERSC's new Cray XT4 said the powerful system produced simulations that offered details of oceanic and atmospheric phenomena, results that were difficult to obtain from other supercomputers before.

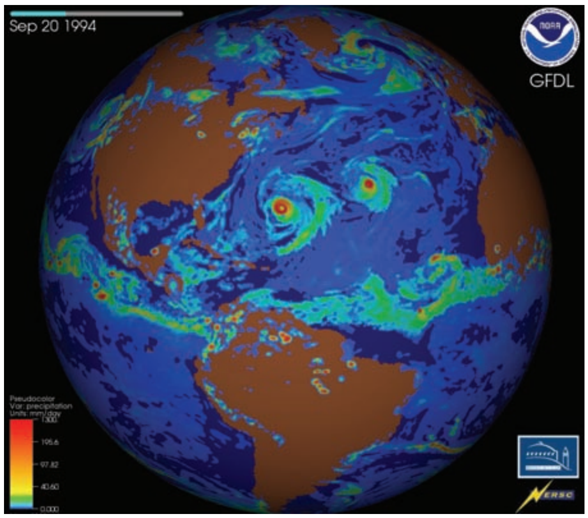
At the DOE's behest, scientists from the National Oceanic and Atmospheric Administration's Geophysical Fluid Dynamics Laboratory (GFDL) proposed a set of experiments using climate models with resolutions many times higher than those in the standard models, such as those used by the IPCC. The high-resolution models offer not only a closer look at physical elements of the climate, such as tropical storms, but they also enable researchers to conduct a more in-depth analysis of climate change as higher-resolution phenomena in the ocean and atmosphere are resolved.

For years, scientists worldwide have relied on simulations with resolutions in the 100-kilometer range for studying forces that shape the oceans and the atmosphere. But the resolution isn't high enough to model details such as ocean vortices and clouds, phenomena that are critical for understanding regional climate variations. Developing a climate model is a computationally intensive task, and getting enough time on powerful supercomputers has always been a challenge.

GFDL scientists, located in Princeton, New Jersey, had developed models capable of modeling the global atmosphere at resolutions down to 5 km, and the ocean at resolutions between 10 km and 20 km. They also have designed experiments which generated 1 to 4 terabytes of data for every year of simulation. NERSC provided GFDL with the computation resources for this challenge by setting aside over 800,000 CPU hours on the new Cray XT4 named Franklin.

Franklin has nearly 20,000 processor cores and a top processing speed of

continued on page 3



The image shows hurricanes formed in the Atlantic Ocean on September 20, 1994.

NERSC 2016

NEW NERSC DIRECTOR KATHY YELICK OUTLINES HER VISION FOR EXPANDING SERVICES AND PROMOTING JOINT RESEARCH

Providing exceptional tools and services to researchers and working with technology companies to develop energy-efficient computers and software were some of the goals outlined by Kathy Yelick during her first staff meeting as the new NERSC Director in January.

Yelick presented "NERSC 2016 Mission," a blueprint for expanding and promoting NERSC's services to meet the changing needs of the scientific community. She noted that NERSC already has earned high marks from users for providing timely support and enabling them to advance their research.

"One of my goals for NERSC is to have it thought of as the most efficient science facility within DOE and the country. Efficiency covers all aspects of NERSC, from the human services side to the energy used in the machines and facility. We have

continued on page 3

Amped Up for Science

MORE PROJECTS, CPU HOURS WILL BE DEVOTED TO INCITE PROJECTS AT NERSC

The U.S. Department of Energy has allocated about 10.4 million CPU hours on supercomputers at NERSC as part of a program to accelerate scientific discoveries in multiple disciplines, including climate, physics, combustion and material science.

The one-year allocations will go to 11 projects by researchers in universities, national labs and industry. Last year, the DOE allotted nearly 9 million CPU hours at NERSC to seven projects.

The awards are part of a program called Innovative and Novel Computational Impact on Theory and Experiment (INCITE), launched in 2003. INCITE, supported by the DOE Office of Science, selects projects that not only require large-scale and intensive use of supercomputers but also promise to deliver a significant advance in science and engineering. Overall, the INCITE program is awarding more than 265 million CPU hours to 55 projects for 2008, up from 95 million CPU hours for 45 projects in 2007.

"The Department of Energy's Office of Science has two of the top 10 most powerful supercomputers, and using them

continued on page 4

Speed Demon

RESEARCHERS SET PERFORMANCE RECORDS FOR MODELING WEATHER ON A NEXT-GENERATION SUPERCOMPUTER

A team of researchers has set a speed performance record for a U.S. weather model by running on Franklin, the new Cray XT4 system at NERSC.

For the highly detailed weather simulations, the researchers used the sophisticated Weather Research and Forecast (WRF) model, widely used for continuous weather forecasting by government, military and commercial forecasters as well as for weather and climate research in hundreds of universities and institutions worldwide.

Scientists from the National Center for Atmospheric Research (NCAR), the San Diego Supercomputer Center (SDSC) at UC San Diego, Lawrence Livermore National Laboratory (LLNL) and IBM Watson Research Center made up the research team that carried out the weather simulations, setting national records not only in speed performance but also in size and fidelity of computer weather simulations.

For setting the speed record using Franklin, the scientists ran the WRF code on 12,090 processors of this 100 peak teraflops system. They achieved the important milestone of 8.8 teraflops — the fastest performance of a weather or climate-related application on a U.S. supercomputer. One teraflops is one trillion, or a thousand billion, calculations per second. It would take a person operating a hand-held calculator more than 30,000 years to complete one trillion calculations.

The team's efforts open the way for simulations of greatly enhanced resolution and size, which will serve as a key benchmark for improving both operational forecasts and basic understanding of weather and climate prediction.

The scientific value of the research goes hand-in-hand with the computational achievements. The "non-hydrostatic" WRF weather code is designed for greater realism by including more of the physics of weather and capturing much finer detail than simpler models traditionally used for global scale weather prediction. Running this realistic model using an unprecedented number of computer processors and simulation size enabled researchers to capture key features of the atmosphere never before represented in simulations covering such a large part of the Earth's atmosphere. This is an important step towards understanding weather predictability at high resolution.

"The scientific challenge we're addressing is the question in numerical weather prediction of how to take advantage of coming petascale computing power," said weather scientist Josh Hacker of NCAR. "There are surprisingly complex questions about how to harness the higher resolution offered by petascale systems to best improve the final quality of weather predictions." Petascale computing refers to next generation supercomputers able to compute at a petaflop (10^{15} calculations per second), equivalent to around 200,000 typical laptops.

The team also set a record for parallelism, or harnessing many computer processors to work together to solve a large scientific problem, running on 15,360 processors of the 103 peak teraflops IBM Blue Gene/L supercomputer at Brookhaven National Laboratory, jointly operated by Brookhaven and Stony Brook University.

"We ran this important weather model at unprecedented computational scale," added Hacker. "By collaborating with SDSC computer scientists to introduce efficiencies into the code, we were able to scale the model to run in parallel on more than 15,000 processors, which hasn't been done with this size problem before, achieving a sustained 3.4 teraflops."

Added John Michalakes, lead architect of the WRF code, "To solve a problem of this size, we also had to work through issues of parallel input and output of the enormous amount of data required to produce a scientifically meaningful result."

continued on page 6

TOP TREND

A report co-authored by Kathy Yelick and John Shalf at NERSC was among the top issues and trends singled out by the editor of HPCwire for its last issue in 2007.

Editor Michael Feldman wrote: "'The Landscape of Parallel Computing Research: The View from Berkeley' became a wake-up call to the computing community about the perils and pitfalls of our manycore destiny. Was anyone listening? Maybe. In the past year, both Intel and Microsoft spent a gazillion dollars for parallel computing R&D and education. Universities like Purdue, LSU, the University of Manchester, MIT and many others are expanding their HPC curriculums for the next crop of students. By the time these kids start to graduate in 2010, the manycore chips will be spilling out of the fabs."

The report was published by the Electrical Engineering and Computer Sciences at UC Berkeley. Yelick, a professor at the university, also is the NERSC Director. Shalf is the head of the NERSC's Science-Driven System Architecture Team. David Patterson, a UC Berkeley professor and a researcher in the Computational Research Division (CRD) at Berkeley Lab, along with former CRD research Parry Husbands also co-authored the report.

The report also acknowledged input from other researchers, including NERSC General Manager Bill Kramer and CRD researchers Jim Demmel and Lenny Oliker.

Kathy Yelick *continued from page 1*

to support users who are running very computationally efficient algorithms and building software that makes efficient use of the hardware resources. We also need to work with vendors to ensure that future architectures support these algorithms and provide facilities that maximize efficiency," Yelick said.

Yelick brings with her decades of experiences in computer science research. She has received a number of research and teaching awards and is the author or co-author of two books and more than 75 refereed technical papers. Yelick earned her Ph.D. in computer science from MIT. She has been a professor at UC Berkeley since 1991, with a joint research appointment at Berkeley Lab since 1996. She headed the Future Technologies Group within the Computational Research Division at Berkeley Lab before becoming the NERSC Director.

In her presentation to the NERSC staff, Yelick emphasized four goals for NERSC:

- Accelerate science by providing high performance computing, information, data and communications resources for all open, applied and basic research and engineering sponsored by DOE.
- Provide unique expertise and assistance to the open science community and enable it to make effective use of HPC resources.
- Accelerate the introduction of ultra efficient low power technology.
- Introduce novel data analysis system and tools that are scalable and flexible.

Yelick also laid out several strategies for achieving the goals. One is to work more closely with researchers in the Computational Research Division (CRD) at Berkeley Lab to explore power-saving chip and computer architectures and algorithms. The construction of the Computational Research and Theory building, which is in the early planning stages, would bring CRD researchers and NERSC engineers together and encourage collaboration and innovation. NERSC is currently located off Berkeley Lab's main campus because of a lack of space.

The building project also provides an opportunity to deploy new energy-saving technologies for running the supercomputer center.

NERSC is enabling new science in all disciplines, with over 1,400 refereed publications annually

Working closely with computer hardware and software companies in improving parallel computing is another key strategy. For the first time, the computer industry is selling parallel systems to the mass consumer market, not just research centers. The industry now has more incentives to improve parallel computing by improving programmability and performance. This trend opens a door for NERSC and CRD researchers to work closely with technology companies to design systems that can benefit both the research community and the consumer market.

"There is an opportunity to influence the broader marketplace and to influence the kinds of machines and software we have access to," said Yelick, adding that the large NERSC user community (nearly 3,000 scientists) makes NERSC an important resource for high-performance computing research.

Providing educational and other training opportunities to NERSC staff also is key to continue NERSC's success. Yelick said she also will encourage the staff to take leadership roles within the high-performance computing community and pursue funding for computing research.

Climate Visualization *continued from page 1*

more than 100 teraflops, making it one of the largest in the world. It arrived at NERSC early last year and passed rigorous testing, which was announced last November.

In addition to carrying out successful runs on Franklin, GFDL also received strong support from NERSC's Analytics Team in using animations to illustrate the results. Prabhat from the Analytics Team created a series of visual renderings of data that included sea surface temperatures and clouds and precipitations in different parts of the world.

"We are able to increase our models' resolutions because of our access to the NERSC machine. One of the results is we can see category 4 or 5 hurricanes in a 20-km model, and they are what we would expect to see in the real world," said V. Balaji, head of the Modeling Services Group at GFDL. Senior software developer Christopher Kerr at GFDL and other members of Balaji's team were responsible for enabling the software infrastructure to perform these scientific experiments. Richard Gerber, a NERSC consultant, resolved system-related issues so that the experiments could be performed on Franklin.

"The results of the visualization collaboration have been outstanding," Balaji added.

The NERSC Analytics Team used the VisIt visualization and analysis package to create images and movies for climate scientists at GFDL. VisIt, which recently won an R&D 100 award, was developed by the DOE Advanced Simulation and Computing Initiative (ASCI) to visualize and analyze the results of large-scale simulations. The team accelerated time to discovery by developing software that eliminates costly data format conversion barriers, namely extra computation, extra data storage and more manual processing steps. As a result, it was possible to

continued on page 5

INCITE Awards *continued from page 1*

through the INCITE program is having a transformational effect on America's scientific and economic competitiveness," DOE Under Secretary for Science Raymond L. Orbach said. "Once considered the domain of only small groups of researchers, supercomputers today are tools for discovery, driving scientific advancement across a wide range of disciplines. We're proud to provide these resources to help researchers advance scientific knowledge and understanding and thereby to provide insight into major scientific and industrial issues."

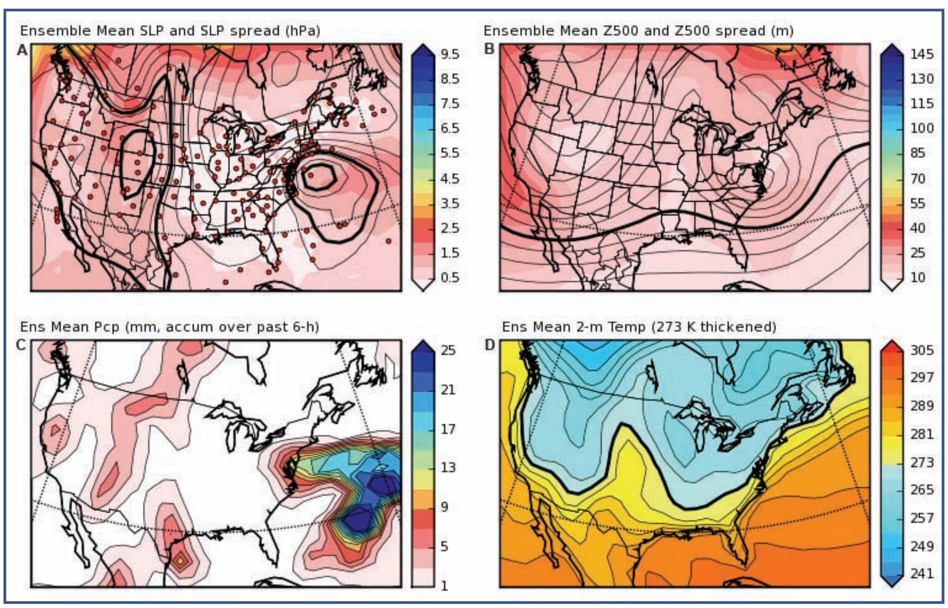
In addition to the projects at NERSC, other INCITE projects were awarded time at DOE's Leadership Computing Facilities at Oak Ridge National Laboratory in Tennessee and Argonne National Laboratory in Illinois, and the Molecular Science Computing Facility at Pacific Northwest National Laboratory in Washington. As the flagship facility for the Office of Science, NERSC provided the only computing resources available during the first two years of the program.

One of the projects at NERSC is led by Gilbert Compo from the University of Colorado who will produce a global tropospheric circulation dataset dating back to 1892 (see figure). The dataset will help validate the climate models being used to make climate projections for the 21st century. The only dataset available for the early 20th century consists of error-ridden, hand-drawn analyses of the mean sea level pressure field over the Northern Hemisphere.

"The allocation has been invaluable. Without it, we could not have generated a dataset of the 6-hourly global weather maps spanning 1918 to 1949 that will be used to understand the Dust Bowl and dramatic Arctic warming of 1920–1940s, among other climate and weather anomalies of the period," said Compo, who also works with the National Oceanographic and Atmospheric Administration (NOAA). Compo, who received an INCITE allocation last year, will be computing at NERSC again this year.

Here are descriptions of other 10 projects awarded computing time at NERSC:

- John Bell from Berkeley Lab will lead a computational study to enable a fundamental understanding and characterization of thermo-diffusively unstable flames in both atmospheric and high-pressure regimes relevant to ultra-lean turbulent premixed burners. The research will aid the development of near-zero-emission combustion devices, a goal of the FutureGen power plant project sponsored by the DOE's Office of Fossil Energy.
- Hong Im from the University of Michigan will lead the work on developing three-dimensional simulations of turbulent nonpremixed flames in the presence of a mean flow strain and fine water droplets. The simulations will help address important issues on energy and environmental research.
- Warren Washington from the National Center for Atmospheric Research (NOAA) will continue the development of the Climate Science Computational End Station (CCES), models for simulating and predicting climate change. The research will examine the human impact on the climate and improve the accuracy of climate models, including the simulation of the global carbon cycle.
- David Randall from Colorado State University will head a project to simulate the global circulation of the atmosphere with roughly a 2-kilometer grid spacing. Understanding the role of clouds in the global atmosphere is key to developing more accurate climate models. This research will not only contribute to that understanding, it also will improve capability for both weather prediction and the simulation of climate change.
- Warren Mori from the University of California at Los Angeles will develop simulations to answer questions about plasma-based particle accelerators that currently cannot be answered through experiments. The project will contribute to the development of better acceleration methods, which are critical for the future of experimental high-energy physics research. New acceler-



Reanalysis of conditions at 7 p.m. on January 28, 1922. (A) Sea level pressure (SLP) measured in hectopascals (hPa): contours show the ensemble mean SLP, with 1000 and 1010 hPa contours thickened; colors show the range of uncertainty; red dots indicate observation locations. (B) Height of 500 hPa pressure in meters: contours show the ensemble mean height, with the 5600 m contour thickened; colors show the range of uncertainty. (C) Ensemble mean precipitation accumulated over 6 hours, in millimeters. (D) Ensemble mean temperature (Kelvin) at 2 meters, with the 273 K (0° F) contour thickened.

continued on page 5

Climate Visualization *continued from page 3*

load the simulation output generated on Franklin directly into VisIt for visual data exploration and analysis.

Developing the visualizations was a collaborative effort. The Analytics Team had extensive interactions with GFDL staff and scientists in order to create visualizations that focus on the most interesting and significant phenomena (e.g., the formation of tropical storms and ocean eddies). The visualizations present the simulation data in an accessible format using conventions familiar to the climate modeling community. Because of the work done by the Analytics Team, VisIt now provides GFDL scientists with the capability to do large-scale visualization of climate data, which was impossible with conventional visualization tools. The Analytics Team has installed VisIt at GFDL to allow scientists to use the new visualization capability on a day-to-day basis.

The climate modeling project, called the Coupled High-Resolution Modeling of the Earth System (CHiMES), began as a collaboration between NOAA/GFDL and DOE. The research uses comprehensive Earth system models (ESMs) and historical data to examine how climate has changed over time and what external forces will likely influence the climate in the future. The models are based on the Flexible Modeling System (FMS) developed by GFDL. FMS is a powerful computational infrastructure for constructing coupled climate models on high-end scalable computer architecture.

The CHiMES project seeks to understand how the overall climate responds to high-resolution phenomena such as ocean eddies, as well as how fine-scale events such as tropical storms respond to climate change. To answer these questions, the project has been divided into two parts. One is to study the climate's predictability over decades or longer using high-resolution coupled models. The second part is to study the correlations between tropical storms and climate change, a hot topic in the research world. Work by GFDL researchers on this subject appeared in over 35 publications in scientific peer-reviewed journals last year.

For the hurricane research, CHiMES uses an atmospheric model based on the cubed sphere grid developed by lead scientist S.J. Lin at GFDL. This projection of a grid over the surface of the Earth, is a more scalable basis than latitudes and longitudes for solving the equations of computational fluid dynamics on a sphere. The researchers have done test-runs using the cubed sphere and found that the highly scalable methodology would enable them to carry out simulations with a 5-km resolution.

"We can go a lot further on this model," Balaji said. "If the coupled model simulations done at NERSC represent today's leading edge, this model is already showing what will be possible when the next generation of hardware becomes available."

INCITE Awards *continued from page 4*

ation techniques using lasers and plasmas could lead to ultra-compact accelerators for applications in science, industry, and medicine.

- Chuang Ren from the University of Rochester will carry out large-scale particle-in-cell (PIC) simulations of the ignition phase in fast ignition (FI), one of the most promising new methods for improving the viability of inertial confinement fusion as a practical energy source. The project will help make fusion energy an environmentally friendly and safe option.
- Lawrence Pratt at Fisk University plans to unravel the mysteries of several lithium compounds that are among the best reagents for forming carbon-carbon bonds in organic synthesis, which can lead to the development of powerful medicines. The project will use ab initio and density functional theory methods to investigate the structure and reactions of the organolithium compounds.
- Ji Qiang from Berkeley Lab will optimize the design and improvement of beam delivery systems for the next-generation X-ray free electron lasers (FELs), which have excellent applications in physics, material science, chemical science and bioscience. Optimizing the beam delivery systems to produce and preserve high intensity and good quality electron beams will not only lower the cost of design and operation of FELs, but also improve the performance of the X-ray light output.
- Leeor Kronik from the Weizmann Institute of Science in Isarel will use the many body perturbation theory in understanding the strucures of novel electronic materials. The results from the research will help clarify pressing issues in figuring out the electronic structure of organic/inorganic interfaces with applications in areas such as semiconductors.
- The project by Fluent Inc., in partnership with General Motors, will use its software to perform computational fluid dynamics and thermal calculations for designing automobiles. The research will tackle five areas, including the full-vehicle open sunroof wind buffeting calculations and the simulations of semi-trucks passing stationary vehicles with raised hoods.

Eleven Berkeley Lab researchers also will take part in four INCITE projects that will be carried out in other DOE supercomputer centers. Lin-Wang Wang will lead a project to explore how and which nano-scale materials should be used for designing better electronic devices, including solar cells. Juan Meza and Zhenji Zhao, both from Berkeley Lab, will take part in this research. Other Berkeley Lab researchers participating in the other three INCITE projects are Bill McCurdy, Tom Rescigno, Ann Almgren, John Bell, Marc Day, David H. Bailey, Lenny Oliker and Kathy Yelick.

Weather Code

continued from page 2

"The input data to initialize the run was more than 200 gigabytes, and the code generates 40 gigabytes each time it writes output data," Michalakes said.

With this power the researchers were able to create "virtual weather" on a detailed 5-kilometer horizontal grid covering one hemisphere of the globe, with 100 vertical levels, for a total of some two billion cells — 32 times larger and requiring 80 times more computational power than previous simulation models using the WRF code.

"The calculation, which is limited by memory bandwidth and interprocessor communication, is representative of many other scientific computations," said Allan Snavely, director of the Performance Modeling and Characterization (PMaC) lab at SDSC, whose group helped tune the model to run at these unprecedented scales. "This means that what we learn in these large simulations will not only improve weather forecasts, but help a number of other applications as they enter the petascale realm."

The work was presented last November at SC07, the international conference for high performance computing, networking, storage, and analysis, where it was a finalist in the prestigious Gordon Bell Prize competition in high performance computing.

"Modeling weather systems is an enormously challenging endeavor, and forecast accuracy depends on the ability to represent many components of the environment and their complex interactions," said Fran Berman, director of SDSC. "The WRF team used sophisticated optimizations to create a new breakthrough in resolution which will lead the way to better predictions, and lays the groundwork for runs on next generation 'petascale' supercomputers. We congratulate them on these exciting results."

In preparing for the groundbreaking runs on the Stony Brook-Brookhaven and NERSC systems, the extensive problem-solving required to achieve these results was made possible by running the WRF code on the Blue Gene system at DOE's Livermore lab, the fastest supercomputer on the Top500 list, and the large Blue Gene system at the IBM Watson Research Center.

Tuning and testing were also carried out at the National Center for Computational Sciences at Oak Ridge National laboratory and on SDSC's Blue Gene system, a resource in the National Science Foundation-supported TeraGrid, an open scientific discovery infrastructure combining leadership class resources at nine partner sites. In these ongoing collaborations the team anticipates further record-setting results.

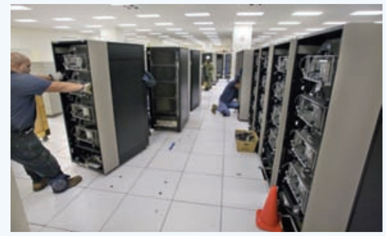
Team members include John Michalakes, Josh Hacker, and Rich Loft of NCAR; Michael McCracken, Allan Snavely, and Nick Wright of SDSC; Tom Spelce and Brent Gorda of Lawrence Livermore; and Robert Walkup of IBM.

Story courtesy of San Diego Supercomputer Center

Seaborg Shutdown

NERSC staff shut down Seaborg, the IBM SP RS/6000, in early January. Seaborg has provided more than 250 million CPU hours to about 3,000 NERSC users since August 2001. Included in these numbers are 26.5 million CPU hours for 22 projects from the Innovative and Novel Computational Impact on Theory and Experiment (INCITE) program, which was created by the DOE Office of Science to support large-scale, high-impact projects.

Work crews tore down Seaborg by removing all the computer racks. Cable recovery and demolition of the seismic bracing for the frames underneath the floor also were underway. NERSC's Bill Isles coordinated the removal work with help from Wallace Haynes in Berkeley Lab's Facilities Division. The work crews came from the Lab's rigger, moving and sheet metal teams. Contractors Doubleday and Trane also participated.



WHAT IS NERSC NEWS?

NERSC News publishes every other month and highlights the cutting-edge research performed using the National Energy Research Scientific Computing Center, the flagship supercomputer facility for DOE's Office of Science. NERSC News editor Ucilia Wang can be reached at 510 495-2402 or Uwang@lbl.gov. Find previous NERSC News articles at http://www.nersc.gov/news/nerscnews.