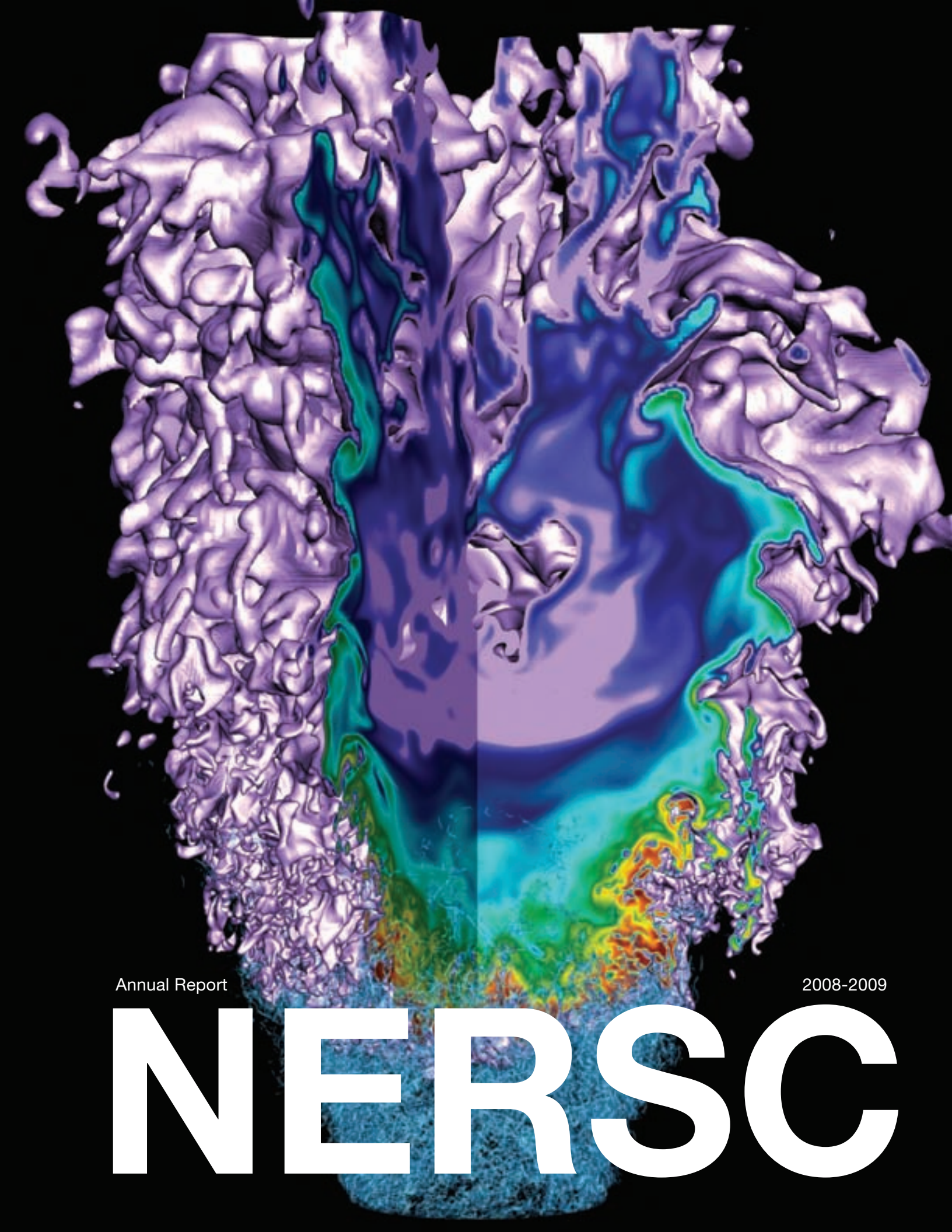


Annual Report 2008-2009

NERSC

Front cover image: Simulated structure of a lean hydrogen flame on a laboratory-scale low-swirl burner. See page 16.

NERSC 2008-2009 Annual Report

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

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

LBNL-3476E, May 2010

Lawrence Berkeley National Laboratory

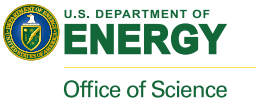


Table of Contents

The Year in Perspective

These last two years have been an exciting time for NERSC. As Director I have endeavored to keep the science goals foremost in NERSC's priorities, while evaluating new technologies that improve NERSC systems and infrastructure, and introducing service innovations to satisfy the increasingly complex needs of our scientific users.

As the primary computing center for the Department of Energy's Office of Science (DOE SC), our mission is serve the computational science needs of DOE SC's six constituent offices. Their science mission currently includes a renewed emphasis on new energy sources, energy production and storage, and understanding the impacts of energy use on both regional and global climate systems, along with mitigation of those impacts through techniques such as carbon sequestration. Climate change researchers have used NERSC systems to study and invent solutions to all aspects of these problems, offering promising evidence that some of the worst impacts of climate change can be avoided with significant reductions in carbon emissions. NERSC users also improved our collective understanding of several promising energy sources, from the nanostructures used within solar panels to the behavior of plasmas in fusion reactors at the scales that will be present in ITER. And in the full tradition of computational science, researchers combined theory, experimentation, and computation to gain insights into the combustion process, which accounts for 85 percent of the world's energy use.

At the same time, NERSC has continued to help researchers in many other fields make basic scientific discoveries such as the first super-bright supernova, screening of color forces between quarks by a background of hadrons, and simulation of a next-generation petawatt laser wakefield accelerator. In a kind of symbiotic relationship, NERSC has provided the computing infrastructure to researchers who are designing the next generation of computing devices, from nanoscale switching elements to materials used for more energy-efficient storage systems.

While NERSC provides help with energy solutions globally, we have also made several contributions locally to make our facility a model of energy efficiency. Starting with our relatively clean supply of electricity—over half from climate-neutral sources—and the cool year-round climate, we have added energy efficiency requirements to our systems procurements. We have also worked with other Berkeley Lab divisions and local industry to develop machine room optimizations, including a system of sensors to provide fine-grained feedback.

But the challenges for energy use by computing centers will continue to grow, and we need solutions that give order-of-magnitude improvements to meet the growing needs for computation within science. Berkeley Lab's Green Flash project started with a clean slate approach, looking at novel hardware technologies for processors, memory and interconnects, as well as new architectural models. The project has demonstrated that by tailoring systems to the needs of scientific computing, order-of-magnitude savings relative to current-day systems are possible. Researchers demonstrated the first Green Flash prototype at the SC09 conference, using hardware emulation on FPGAs and a full software stack, showing a climate modeling application running on the system. One of many articles on the project appeared in the journal IEEE Computer, which was given to all SC09 attendees.

Current hardware trends have resulted in a more immediate challenge for NERSC and its users in the form of multicore processors. As Moore's Law continues on its historical climb, the consequent performance improvements are seen now as an increase in core count rather than clock speed. We saw this trend in practice when, in early 2008, NERSC exercised an option to upgrade its Franklin XT4 system from its original dual-core processors to newer quad-core versions.

Since NERSC has always been focused on application performance rather than peak performance, the decision to upgrade was not an obvious one. In collaboration with Cray, we took a set of application benchmarks and developed estimates of the performance increases. Although the peak performance would increase by nearly four times with the proposed upgrade due to changes in the handling of data-parallel instructions along with the core doubling, the estimated application performance increase was a factor of two. This confirmed the growing gap between application performance and theoretical peak speed and emphasizes the need to measure and evaluate real application performance. The upgrade also took advantage of denser memory technology, which doubled both capacity and bandwidth, allowing users to retain the same parallelization strategy with their codes. In he final analysis, the quad-core upgrade did offer a cost-effective way to double the cycles available to users.

The upgrade process itself was a creation of NERSC and Cray, developed to reduce the risk of the upgrade and minimize the time away from science. Rather than turning off the entire system during the upgrade, we partitioned it to create both a system for science and a system under reconfiguration. These systems changed dynamically as hardware was upgraded, using only short downtimes to repartition the system. This technique has become Cray's standard strategy for upgrades, and has subsequently been used at other computing centers within DOE.

A quantitative approach based on application performance was also used in procuring our next major computing system, a Cray system that we have named Hopper, after pioneering computer scientist Grace Hopper. The Cray technology, in our analysis, offered the best application performance per dollar over the lifetime of the system. The first phase of Hopper was delivered in late 2009, and the second phase will be installed later in 2010. It will utilize new hardware and software technologies, including AMD's new 12-core processors and externalized services to make the system highly available to users. While delivering a peak performance of more than one petaflop and using more than 150,000 cores, the final Hopper system will, more importantly, make a significant impact on the computational capability available for science. To address some of the user challenges posed by the growing number of cores, we have joined with Cray in a Center of Excellence for programming models, leveraging the expertise in both institutions to provide users with sound advice and training on how to best program Hopper and future systems that will likely have many more cores per chip.

Processor technology is not the only recent revolution in computing. Social networking sites have redefined the ways in which people interact and form communities in areas of common interest, breaking

down the usual geographic barriers. The Deep Sky project at NERSC has taken this idea and applied it to science, creating a worldwide community for the exploration of astronomical phenomena. With 8 million images available through the Deep Sky web portal, which sits on top of sophisticated management software, scientists discovered 140 supernovas in 60 nights, with 15 scientists working 24 hours a day worldwide to identify interesting features. The enormous volume of data means that manual image processing is not practical on its own, and advanced machine-learning algorithms automatically identify potential features of interest.

Deep Sky is a prototype example of a scientific web portal for large scientific datasets, and NERSC is providing such services to other science domains in the form of Science Gateways. Our users' needs for processing, analyzing, and sharing massive datasets with their collaborators motivated NERSC's I/O, filesystem, and HPSS upgrades, all of which are described in this Annual Report. Data-intensive science, a field that includes advanced algorithms, data management tools, and storage and communication systems, along with visualization, is one of the centerpieces of NERSC's long-term planning process.

Computing is not just an enabler of social phenomena; it can also be used as a tool for exploring such phenomena. In late 2008 DOE and the National Endowment for the Humanities awarded time on NERSC supercomputers to three projects in the humanities, including cultural analysis based on digitized media. Humanities researchers collaborated with NERSC Analytics Team members who contributed key pattern recognition codes to the cultural analytics image-processing pipeline — yet another way in which NERSC has expanded the usefulness of high performance computing.

In another example of planning how to best meet our users' future needs, we are investigating cloud computing, which has become a prevalent tool for providing computation and storage resources in the commercial world and has frequently been discussed as a possible resource for scientists. This is of particular interest for mid-range problems that are currently run on privately owned clusters managed by individual PIs. NERSC is participating in the American Recovery and Reinvestment Act (ARRA) project called Magellan to begin answering the basic question of what it means to build and manage a cloud computing system for science. NERSC is deploying a testbed to explore whether the cloud computing model can be effectively adapted to mid-range scientific computing needs. Potential benefits, in addition to greater computing capacity, include lower energy use and lower operational costs. NERSC staff will use performance-monitoring software to analyze what kinds of science applications are being run on the Magellan system, how well they perform in a cloud environment, and what hardware and software features are needed for science clouds.

The technology trends and science drivers provide both challenges and opportunities for NERSC, but our reputation as an outstanding center for high performance computing really comes from the staff at NERSC, who are unfailingly dedicated to their jobs and creative when faced with new problems. As part of an effort to develop future experts in computational science, both consultants and users, NERSC is receiving more than $3 million in ARRA stimulus funds over the next two years to hire several post-doctoral researchers into a "Petascale Postdoc" program. These young researchers will help design and modify modeling codes in key research areas such as energy technologies, fusion, and climate, to run on emerging multi-core systems. The postdocs will work with key NERSC users to create new procedures for making their current algorithms suitable for very large numbers of cores, to develop new algorithms to replace ones that do not scale, and to introduce new language constructs that are suitable for multi-core and/or GPU platforms. Under this program, we expect that the same care and consideration that goes into selecting our next-generation systems will also help develop the next generation of computational scientists.

With these exciting times ahead of us, I am again grateful to our DOE Office of Science sponsors for their continued endorsements, to our users for the science they produce using NERSC resources, and to the NERSC staff who make NERSC an excellent high performance computing center.

Katherine Yelick
NERSC Division Director

Climate and Energy Research: Expanding Our Choices

"So we have a choice to make. We can remain one of the world's leading importers of foreign oil, or we can make the investments that would allow us to become the world's leading exporter of renewable energy. We can let climate change continue to go unchecked, or we can help stop it. We can let the jobs of tomorrow be created abroad, or we can create those jobs right here in America and lay the foundation for lasting prosperity."

—President Barack Obama, March 19, 2009

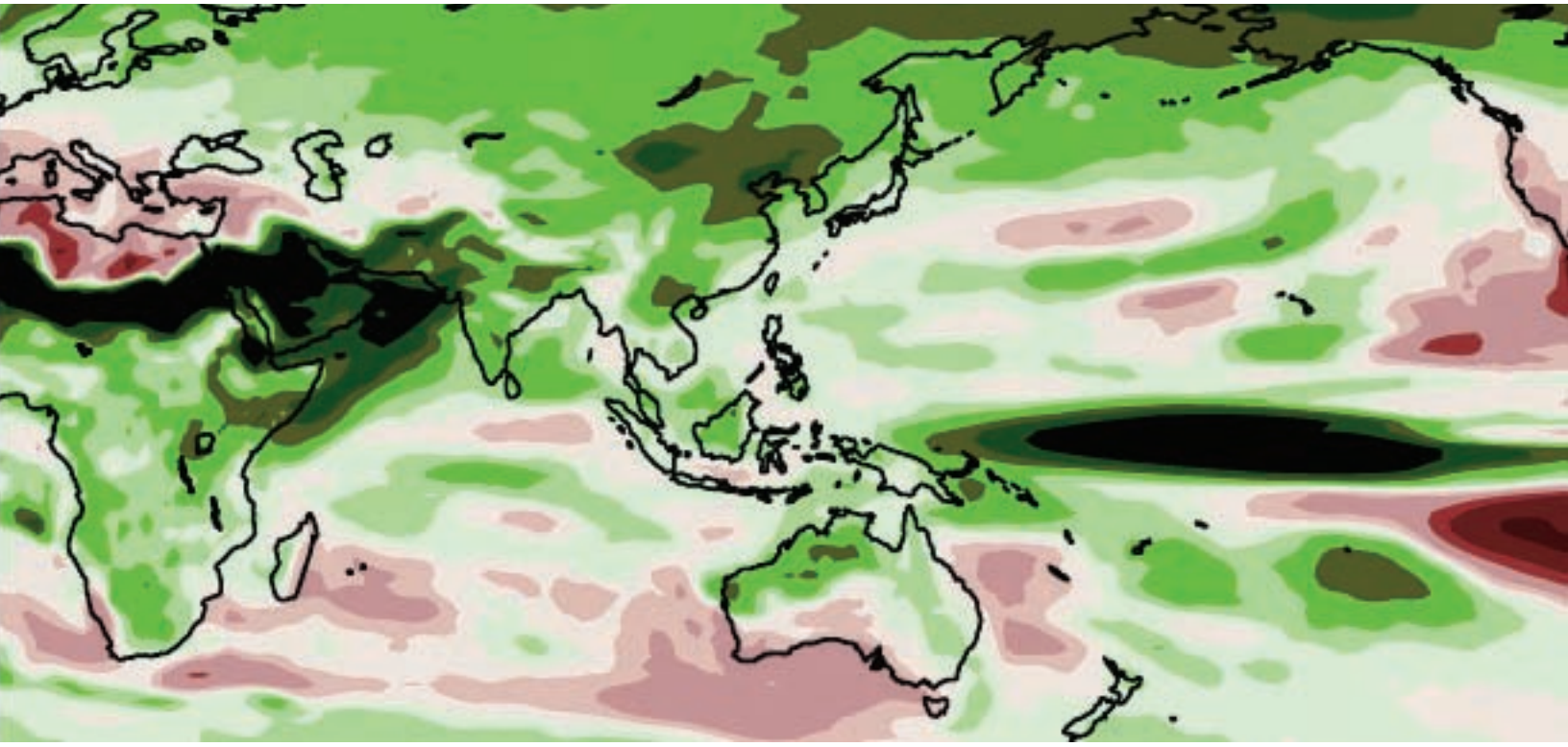
The Obama Administration has pledged to invest $15 billion a year to develop and deploy the next generation of renewable energy technologies here in America. The Administration has already allocated $1.6 billion in Recovery Act funds to the U.S. Department of Energy (DOE) Office of Science. That allocation includes expanded funding for integrated climate research, which blends climate modeling with modeling of human factors such as economics and choices about energy production, consumption, and use.

The National Energy Research Scientific Computing Center (NERSC), as its name suggests, has its roots in the quest for alternative energy sources. NERSC was founded (under a different name) in 1974 as the first unclassified supercomputer center, with the mission of serving the DOE Magnetic Fusion Energy program—a mission that continues today for the Office of Fusion Energy Sciences.

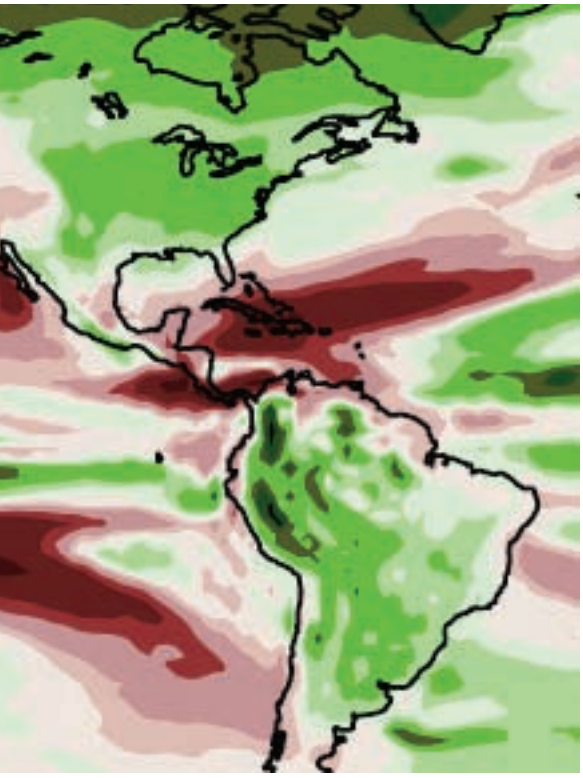
In 1983 NERSC's role was expanded beyond the fusion program, and it began providing general computing services to all of the programs funded by the DOE Office of Energy Research (now the Office of Science). In recent years NERSC has been a major resource for the U.S. and the Intergovernmental Panel on Climate Change (IPCC) climate modeling efforts. NERSC has also supported the basic science research which lays the foundation for energy efficiency and alternative energy technologies, including materials science for catalysts, combustion science, and nanoscience for solar photovoltaics, among many others.

The following pages provide examples of recent scientific accomplishments in climate and energy research, made possible in part by NERSC's computing and scientific resources.

It's Not Too Late



Cuts in greenhouse gas emissions would save arctic ice, reduce sea level rise



Project: Climate Change Simulations with CCSM: Moderate and High Resolution Studies

PI: Warren Washington, National Center for Atmospheric Research

Senior investigators: Gerald Meehl and Lawrence Buja, NCAR

Funding: BER, NCAR

Computing resources: NERSC, OLCF, ALCF, NSF

The threat of global warming can still be greatly diminished if nations cut emissions of heat-trapping greenhouse gases by 70 percent this century, according to a study led by scientists at the National Center for Atmospheric Research (NCAR).[1] While global temperatures would rise, the most dangerous potential aspects of climate change, including massive losses of Arctic sea ice and permafrost and significant sea level rise, could be partially avoided.

"This research indicates that we can no longer avoid significant warming during this century," says NCAR scientist Warren Washington, the lead author. "But if the world were to implement this level of emission cuts, we could stabilize the threat of climate change and avoid an even greater catastrophe."

Average global temperatures have warmed by close to 1 degree Celsius (almost 1.8 degrees Fahrenheit) since the pre-industrial era. Much of the warming is due to human-produced emissions of greenhouse gases, predominantly carbon dioxide. This heat-trapping gas has increased from a pre-industrial level of about 284 parts per million (ppm) in the atmosphere to more than 380 ppm today.

With research showing that additional warming of about 1 degree C (1.8 degrees F) may be the threshold for dangerous climate change, the European Union has called for dramatic cuts in emissions of carbon dioxide and other greenhouse gases. The U.S. Congress is also debating the issue.

To examine the impact of such cuts on the world's climate, Washington and his colleagues ran a series of global supercomputer studies with the NCAR-based Community Climate System Model (CCSM). They assumed that carbon dioxide levels could be held to 450 ppm at the end of this century. That figure comes from the U.S. Climate Change Science Program, which has cited 450 ppm as an attainable target if the world quickly adopts conservation practices and new green technologies to cut emissions dramatically. In contrast, emissions are now on track to reach about 750 ppm by 2100 if unchecked.

The team's results showed that if carbon dioxide were held to 450 ppm, global temperatures would increase by 0.6 degrees C (about 1 degree F) above current readings by the end of the century. In contrast, the study

[1] Warren Washington, Reto Knutti, Gerald Meehl, Haiyan Teng, Claudia Tebaldi, David Lawrence, Lawrence Buja, and Gary Strand, "How Much Climate Change Can Be Avoided by Mitigation?" Geophysical Research Letters **36**, L08703 (2009).

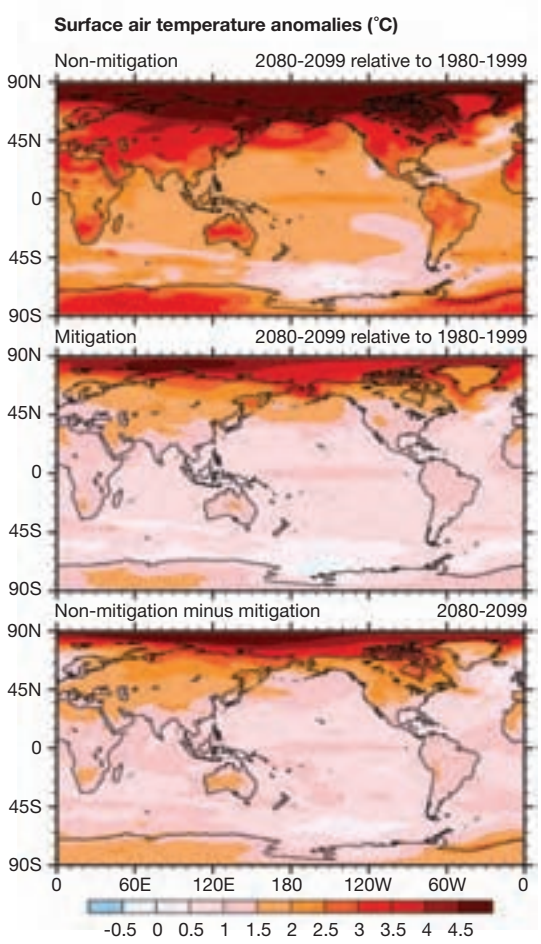


Figure 1. Computer simulations show the extent that average air temperatures at Earth's surface could warm by 2080-2099 compared to 1980-1999, if (top) greenhouse gases emissions continue to climb at current rates, or if (middle) society cuts emissions by 70 percent. In the latter case, temperatures rise by less than 2°C (3.6°F) across nearly all of Earth's populated areas (the bottom panel shows warming averted). However, unchecked emissions could lead to warming of 3°C (5.4°F) or more across parts of Europe, Asia, North America and Australia.

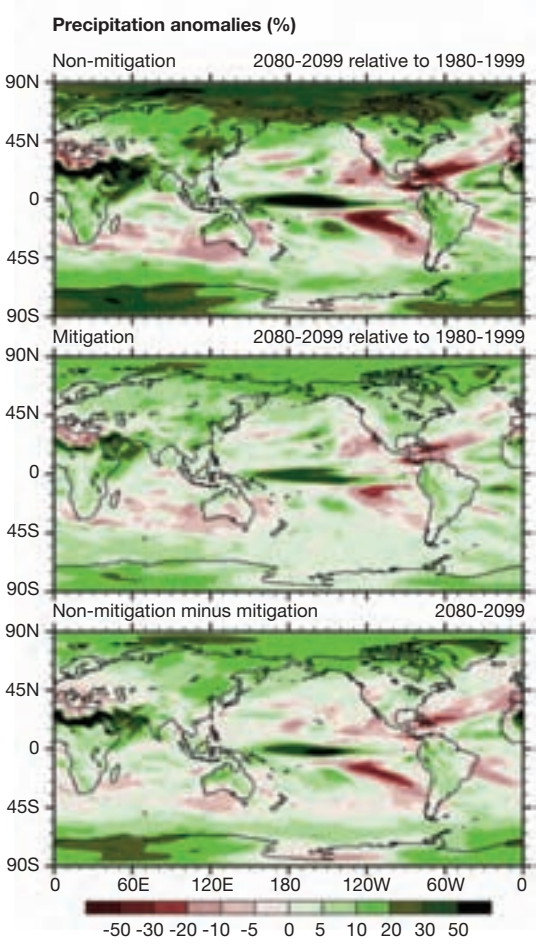


Figure 2. The non-mitigation case (top) shows increased precipitation in the northeast United States, Canada, and other regions, but a significant decrease in precipitation in the southwestern U.S. These changes are cut in half if greenhouse gas emissions are reduced by 70 percent (middle panel). The bottom panel shows the difference between the two scenarios, that is, the changes in precipitation that could be avoided by reducing emissions.

showed that temperatures would rise by almost four times that amount, to 2.2 degrees C (4 degrees F) globally above current observations, if emissions were allowed to continue on their present course (Figure 1).

Holding carbon dioxide levels to 450 ppm would have other impacts, according to the climate modeling study:

- Sea level rise due to thermal expansion as water temperatures warmed would be 14 centimeters (about 5.5 inches) instead of 22 centimeters (8.7 inches). Significant additional sea level rise would be expected in either scenario from melting ice sheets and glaciers.
- Arctic ice in the summertime would shrink by about a quarter in volume and stabilize by 2100, as opposed to shrinking at least three-quarters and continuing to melt. Some research has suggested the summertime ice will disappear altogether this century if emissions continue on their current trajectory.
- Arctic warming would be reduced by almost half, helping preserve fisheries and populations of sea birds and Arctic mammals in such regions as the northern Bering Sea.

- Significant regional changes in precipitation, including decreased precipitation in the U.S. Southwest and an increase in the U.S. Northeast and Canada, would be cut in half if emissions were kept to 450 ppm (Figure 2).
- The climate system would stabilize by about 2100, instead of continuing to warm.

The research team used supercomputer simulations to compare a business-as-usual scenario to one with dramatic cuts in carbon dioxide emissions beginning in about a decade. The authors stressed that they were not studying how such cuts could be achieved nor advocating a particular policy.

"Our goal is to provide policy makers with appropriate research so they can make informed decisions," Washington says. "This study provides some hope that we can avoid the worst impacts of climate change — if society can cut emissions substantially over the next several decades and continue major cuts through the century."

Higher-resolution models

CCSM is one of the world's most advanced coupled climate models for simulating the earth's climate system. Composed of four separate component models simultaneously simulating the earth's atmosphere, ocean, land surface, and sea ice, plus one central coupler component, CCSM allows researchers to conduct fundamental research into the earth's past, present, and future climate states.

CCSM version 3 was used to carry out the DOE/NSF climate change simulations for the Intergovernmental Panel on Climate Change Fourth Assessment Report (IPCC AR4), which was released in 2007. This breakthrough study presents a clear picture of a planet undergoing a rapid climate transition with significant societal and environmental impacts.

"The strength and clarity of the IPCC AR4 report can be attributed, in part, to DOE and NSF supercomputing centers making it possible to deploy climate models of unprecedented realism and detail," Washington says.

"Our next challenge is applying the emerging class of earth system models (CCSM4) that include detailed physical, chemical, and biological processes, interactions and feedbacks in the atmosphere, oceans, and surface, to carry out policy-relevant adaptation/mitigation scenarios. We will also use a higher resolution version of CCSM4 to perform decadal prediction experiments that can provide more regional information about climate change."

For long-term studies with a resolution of 2 degrees for the atmosphere and 1 degree for the ocean, CCSM currently runs on more than 2000 processors of Franklin, NERSC's Cray XT4 system. At that resolution, more than 20 years of climate conditions can be simulated in one day of computing. Recent scaling improvements in the atmosphere component of the code have resulted in a doubling of throughput, allowing more processors to be used for faster simulations.

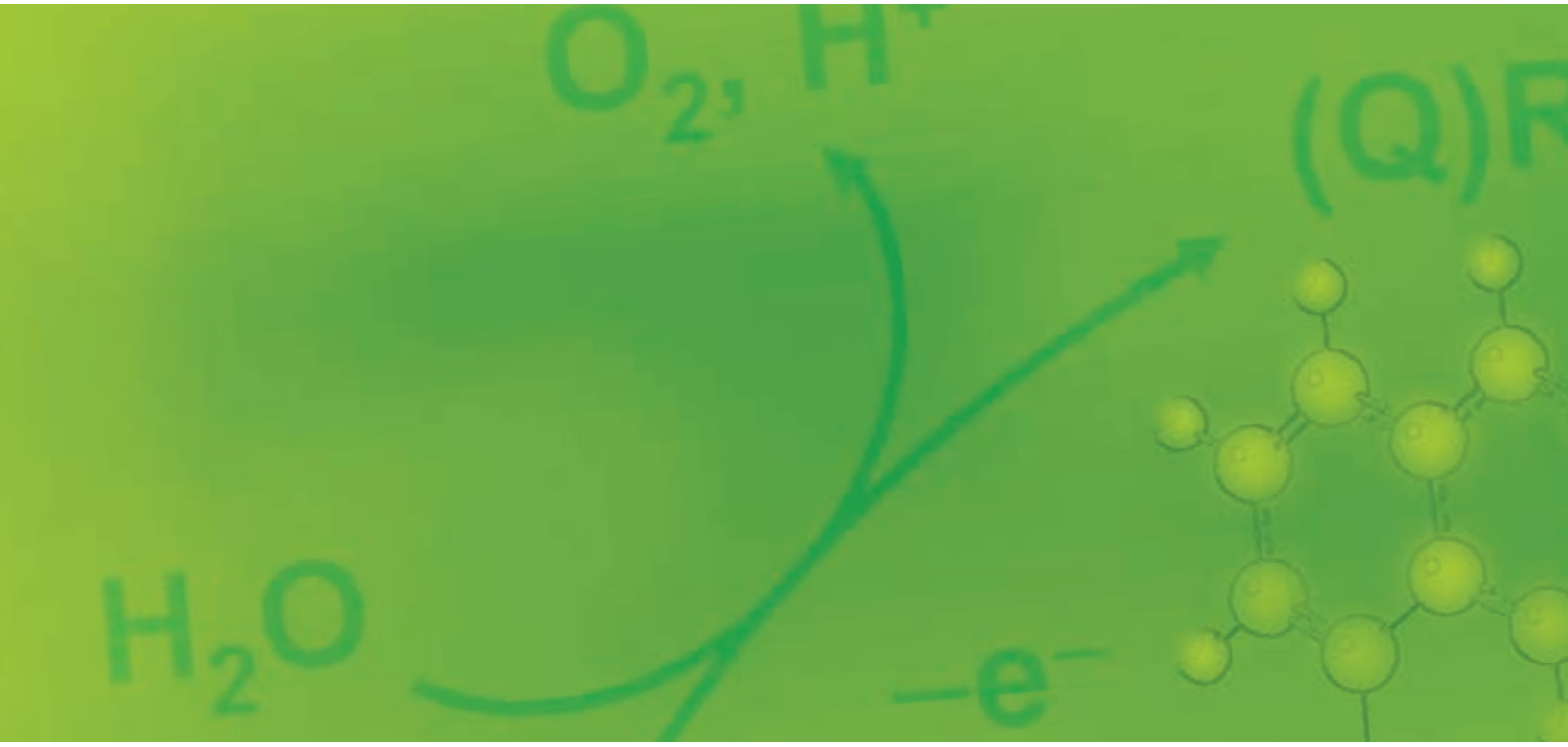
Higher-resolution runs (such as those used for decadal prediction studies) with a resolution of 0.5 degree for atmosphere and land, and 1 degree for ocean and sea ice, will run on up to 6000 processors, simulating about 2 years per day of computing, but offering much more detailed information about the regional impacts of global climate change.

Future work at NERSC by Washington's research group will include further climate change detection/attribution studies for the IPCC AR5 report. A series of simulations will be made to link the observed climate data to specific *forcings*, or factors that influence climate. Ten different forcing combinations will be investigated and compared with observed data for the years 1850–2005. The forcing combinations are:

1. Natural forcings only
2. Anthropogenic forcings only
3. Solar forcing only
4. Volcanic forcing only
5. Greenhouse gas forcing only
6. Ozone forcing only
7. Dust forcing only
8. Aerosol direct forcing only
9. Aerosol indirect forcing only
10. Land-cover changes only.

Other research topics for the coming year will include the impact of the 11-year solar cycle, the impact of collapse of the meridional overturning circulation, climate feedbacks due to permafrost-thaw related methane emissions, and decadal predictability experiments.

Artificial Photosynthesis



Researchers are working to design catalysts that can convert water and carbon dioxide into fuels using solar energy

Project: Computational Studies at BNL of the Chemistry of Energy Production and Use

PI: James Muckerman, Brookhaven National Laboratory

Senior investigators: Ming-Kang Tsai, Altaf Karim, Patrick Achord, Jonathan Skone, and Yolanda Small, BNL

Funding: BES, Hydrogen Fuel Initiative, Solar Energy Utilization Initiative

Computing resources: NERSC, New York Center for Computational Sciences

When scientists use the word *conversion*, they're usually talking about a physical transformation from one material or state to another, such as the conversion of carbon dioxide and water into carbohydrates and oxygen by photosynthesis in plants. But for James Muckerman, a senior chemist at Brookhaven National Laboratory, the word *conversion* could have a more personal meaning.

"About six years ago, I had an epiphany of sorts, in the form of a sudden realization that perhaps the most important problem in the world was staring me in the face," Muckerman says. "This problem was something that I might even have some expertise to help solve, but I was ignoring it. So I decided to stop ignoring it and change the focus of much of my research to work on it."

The problem that loomed in front of Muckerman was actually the twin problems of energy and survival: how to provide the energy needed by a growing world population in a way that does not endanger our survival by changing the world's climate.

Before his epiphany, Muckerman had been an expert in gas-phase molecular dynamics, which explores the energetics, dynamics, and kinetics of chemical reactions resulting from molecular collisions in the gas phase.

"When you change fields like I did, it's pretty scary unless you have a safety net," he says. "My safety net was the three experimental collaborators that I had in this work, who were already experts in the field of renewable energy and catalysis. The problems they were working on needed some theoretical insight, and I was happy to have the opportunity to team up with them and have some really good problems to work on." Those collaborators were Etsuko Fujita at Brookhaven, Dan DuBois at Pacific Northwest National Laboratory, and Koji Tanaka at the Institute for Molecular Science in Japan.

The long-term goal of these chemists' research is to design a bio-inspired system that can produce fuels like methanol from carbon dioxide (CO_2) or hydrogen from water, using solar energy to power the process. Of course, creating fuel from water, CO_2, and solar energy is exactly what photosynthesis does, and fossil fuels are the ultimate product. But Muckerman and his colleagues would like to shorten the process and make fuels more directly by designing catalysts that oxidize water (removing electrons and protons) and reduce CO_2 (adding electrons and protons). Capturing CO_2 at the smokestacks and recycling it by converting it to fuel would go a long way toward solving the greenhouse gas problem.

Carbon dioxide reduction

The first step in CO_2 reduction, converting it to carbon monoxide (CO), is relatively easy. But CO is not very reactive, so adding the hydrogen to make methanol (CH_3OH) is challenging. Rather than trying to add hydrogen atoms, these scientists are researching ionic hydrogenation, which adds hydrogen as a proton and a hydride. (A hydride is the negative ion of hydrogen; it has an extra electron added to the basic structure of one proton and one electron). In natural photosynthesis, the coenzyme NADH can serve as a hydride donor, so the researchers are looking for a catalyst that functions in a similar way.

Molecules containing transition metals are prime candidates for NADH analogues. Transition metals (so called because of their position in the periodic table) absorb light efficiently over a wide spectral range, have long-lived excited states, and have multiple oxidation states, making them excellent catalysts. Ruthenium is one transition metal whose catalytic properties have been studied by inorganic chemists for 25 years.

Tanaka has discovered that a very complicated molecule nicknamed ruthenium-pbn (Ru-pbn) can function as an electrocatalyst for reducing acetone to isopropanol. Fujita, Muckerman and coworkers, including Tanaka, demonstrated that Ru-pbn could be a renewable hydride donor powered by light. When two molecules of Ru-pbn each absorb a photon, they react with each other in a reaction called *disproportionation* — a simultaneous oxidation and reduction — resulting, after several steps, in the conversion of one of the molecules into the protonated hydride Ru-pbnHH, while the other molecule reverts to its original state. A combined experimental and computational study of the mechanistic pathways of Ru-pbnHH formation by Muckerman and his colleagues was featured on the cover of the May 19, 2008 issue of Inorganic Chemistry (Figures 3 and 4).[2]

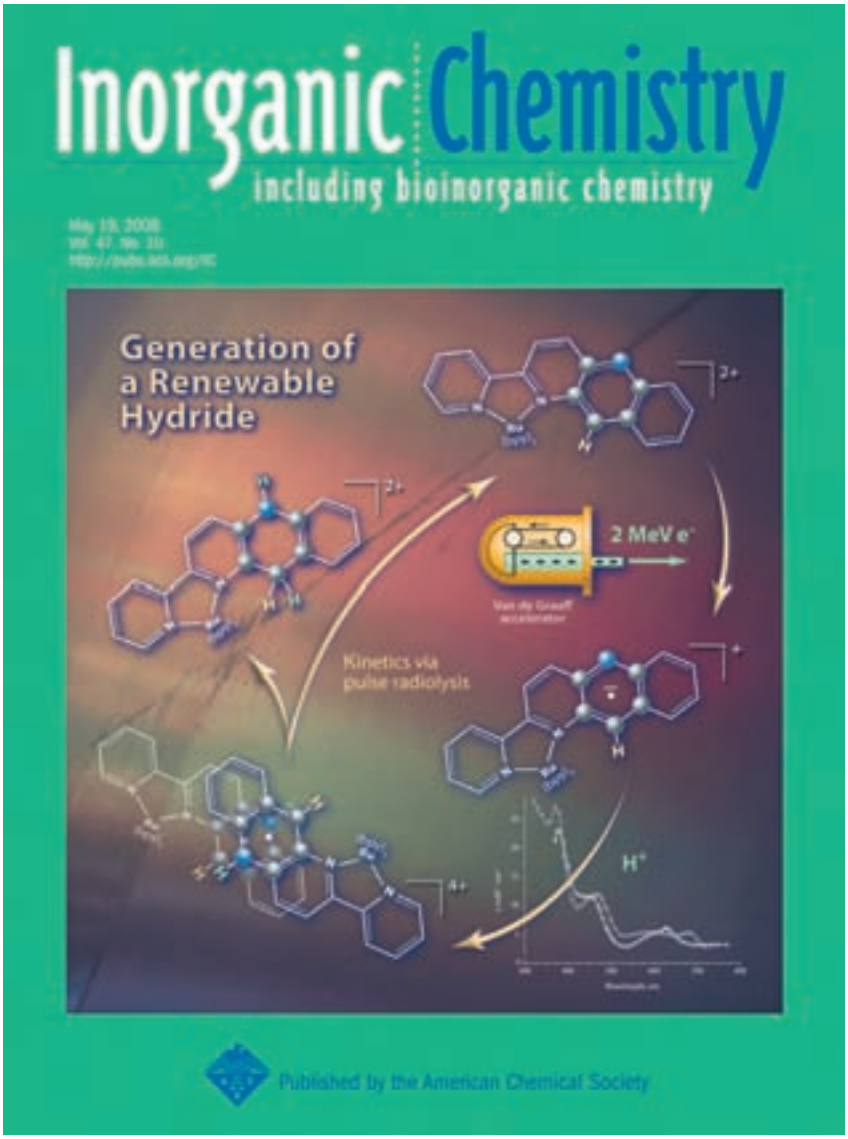


Figure 3. The mechanistic pathways of the formation of Ru-pbnHH from Ru-pbn were studied in an aqueous medium using pulse radiolysis and through theoretical calculations of electronic energies. (Image ©2008 American Chemical Society)

One theoretical contribution that the experimentalists did not expect was the discovery that each Ru-pbnHH molecule, because of its long-lived excited state, can absorb an extra (third) photon in the original reaction time frame, making it an even stronger hydride donor. These results open new opportunities for the photochemical generation of strong hydride donors for the catalytic hydrogenation of CO_2

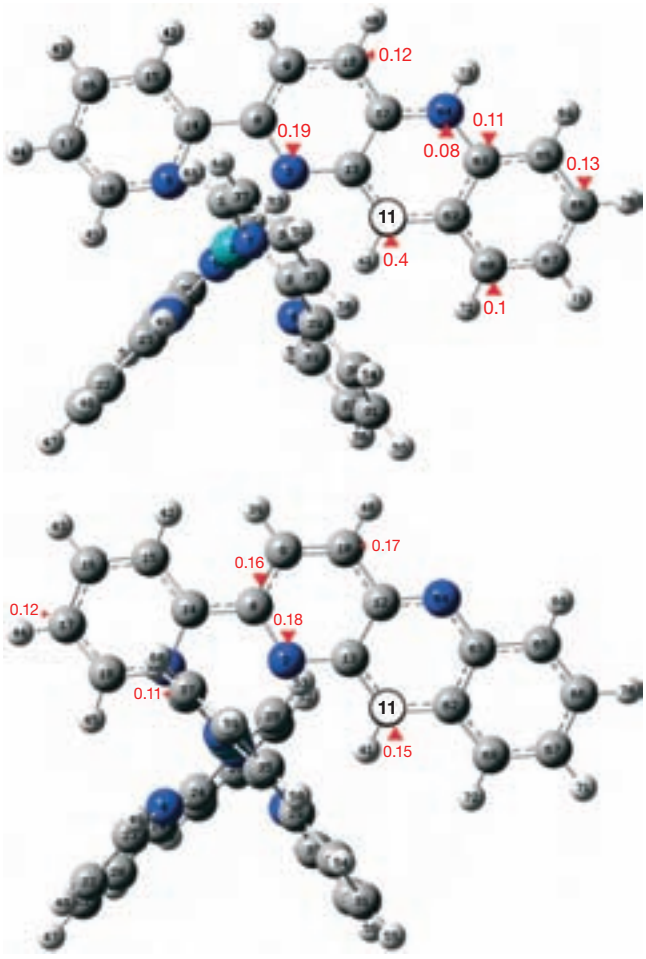


Figure 4. Calculated spin density of an electron in the singly occupied molecular orbital (SOMO) of two intermediate species between Ru-pbn and Ru-pbnHH. The analysis of spin densities reveals substantial differences in unpaired electron location in the two species, elucidating the electronic transitions that take place in the disproportionation reaction.

and related species.

But finding a hydride donor is only part of the solution, because CO does not readily accept a hydride. Binding the CO to a metal (making it a *carbonyl ligand*) makes it much more receptive of hydrides. Hydride donating abilities (or *hydricities*) are very difficult to measure, so Muckerman's group has calculated a table of thermodynamic hydricities of various species involving CO, CO_2, and donor and acceptor molecules for hydrides. That allows them to screen possible systems that might convert CO to methanol by transferring a hydride from a photogenerated hydride donor to a carbonyl ligand of a transition metal complex.

Water oxidation

Another aspect of photosynthesis that the researchers are trying to mimic is water oxidation — using solar energy to split water into molecular oxygen and hydrogen. While plants combine the hydrogen with CO_2 to create carbohydrates, scientists would like to use the hydrogen in fuel cells or combustion devices to produce energy.

Today hydrogen is produced commercially by extracting it from petroleum or natural gas. Using solar energy to oxidize water could be a promising alternative, but water molecules are very stable, so splitting them would require a lot of solar energy as well as a catalyst to activate the reaction. The biggest obstacle to date has been the lack of a robust molecular catalyst that oxidizes water without undergoing oxidative degradation itself.

Tanaka and his colleague Tohru Wada may have found such a catalyst in a ruthenium complex with quinone ligands. (Quinones are a class of organic compounds having two carbonyl groups in a six-member carbon ring.) What makes this catalyst novel is that in most metal-based catalysts, the electron receptor sites are located on the metal atoms; but in this ruthenium complex, the receptor sites are on the quinone ligands, giving this complex much more stability than previous ruthenium catalysts by avoiding high oxidation states of the metal centers that can

[2] Dmitry E. Polyansky, Diane Cabelli, James T. Muckerman, Takashi Fukushima, Koji Tanaka, and Etsuko Fujita, "Mechanism of hydride donor generation using a Ru(II) complex containing an NAD+ model ligand: pulse and steady-state radiolysis studies," Inorganic Chemistry **47**, 3958 (2008).

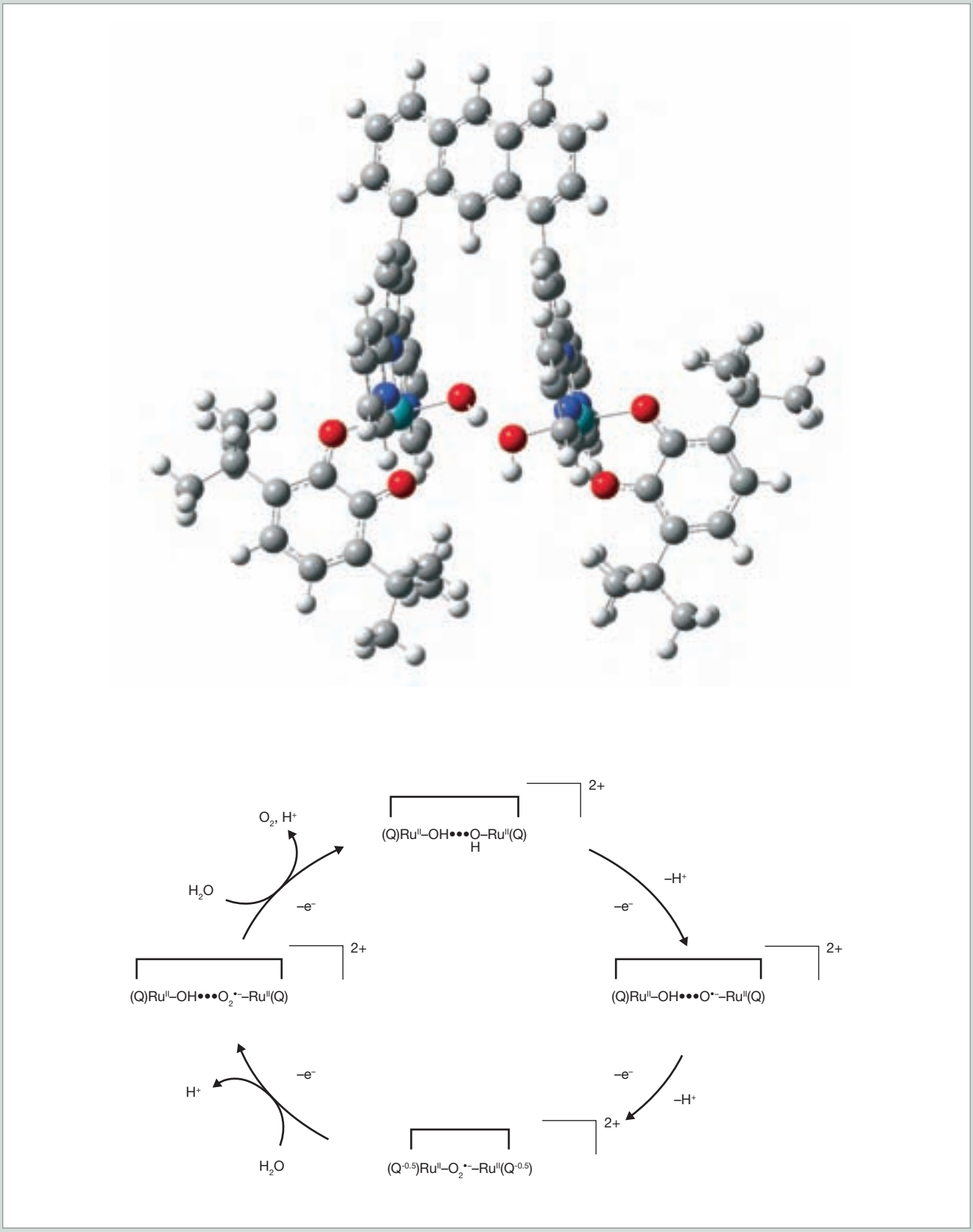


Figure 5. The calculated structure of the Tanaka catalyst (top) and its proposed catalytic cycle for water oxidation in aqueous solution at pH 4 (bottom).

be indiscriminately reactive. These ligands are capable of removing four electrons and four protons from two water molecules in a series of four proton-coupled electron-transfer steps, making this an exceptionally promising water oxidation catalyst.

The theorists and experimentalists are working together to understand the novel mechanisms of this catalyst, and so far have produced three papers on the topic.[3,4,5] It appears that the catalyst holds water molecules in place to make oxygen bonds while the protons and electrons are transferred in a multi-step reaction (Figure 5).

Muckerman's team has developed a theoretical scheme using thermodynamic cycles to calculate the reduction-oxidation (redox) potential of catalysts in aqueous solution at various pH using *Pourbaix diagrams*, phase-diagram-like plots of the dominant species in solution as a function of applied potential and pH. "To the best of our knowledge, this is the first attempt to directly compare theoretical and experimental Pourbaix diagrams," Muckerman says. The researchers have used this theoretical scheme to study monomeric modifications of Tanaka's catalyst (Figure 6).

Because ruthenium is rare and expensive, researchers do not expect to use it in a practical photosynthetic device. "The aim of our work is to understand how these catalysts work and to elucidate the detailed mechanistic steps so we can transfer that chemistry to cheap and plentiful metals like iron, nickel, and cobalt, which are right above ruthenium in the periodic table," Muckerman explains. "Our goal is to design an effective water oxidation catalyst based on an earth-abundant metal."

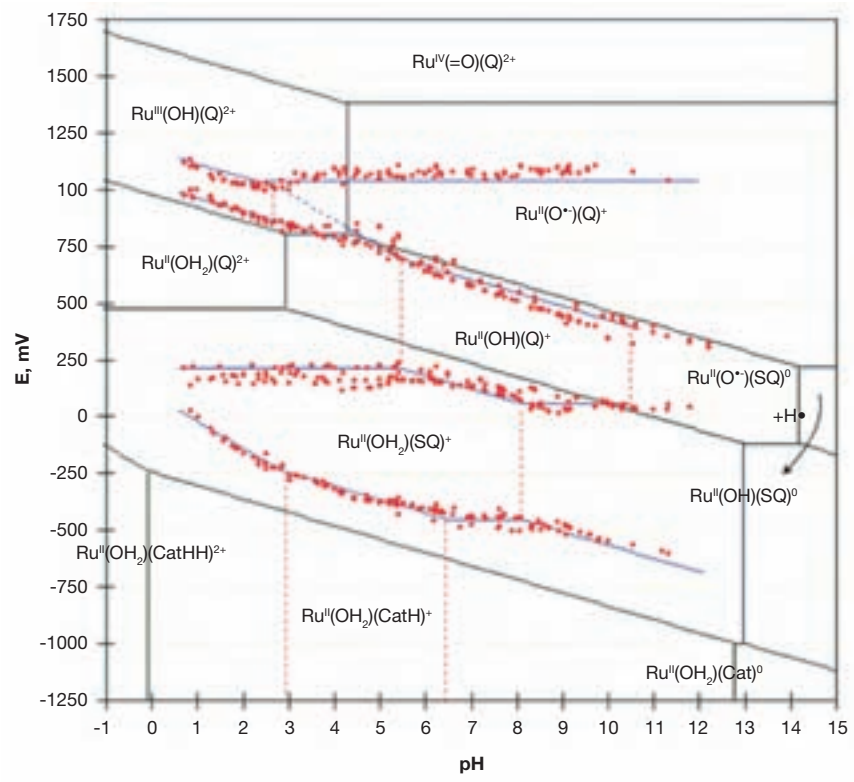
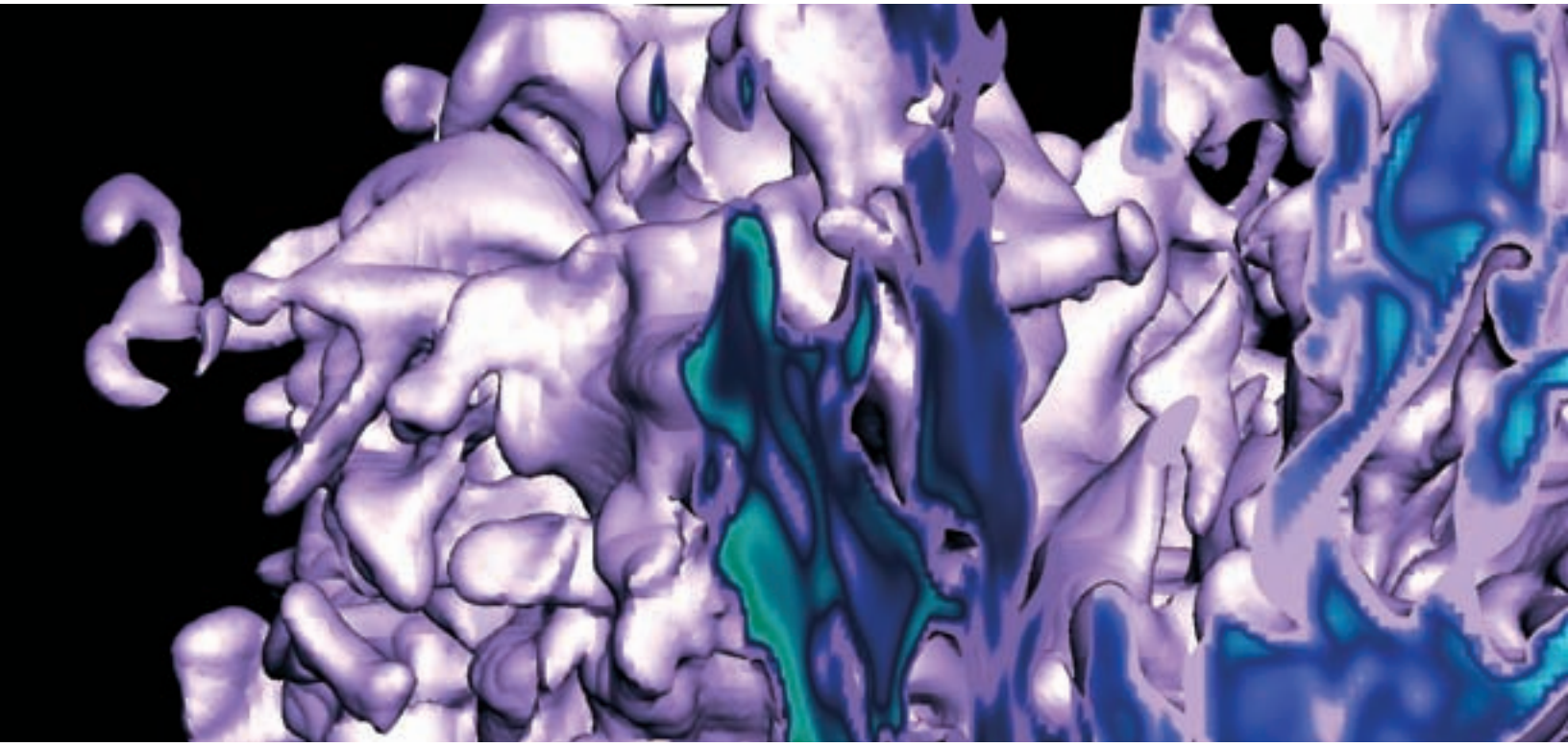


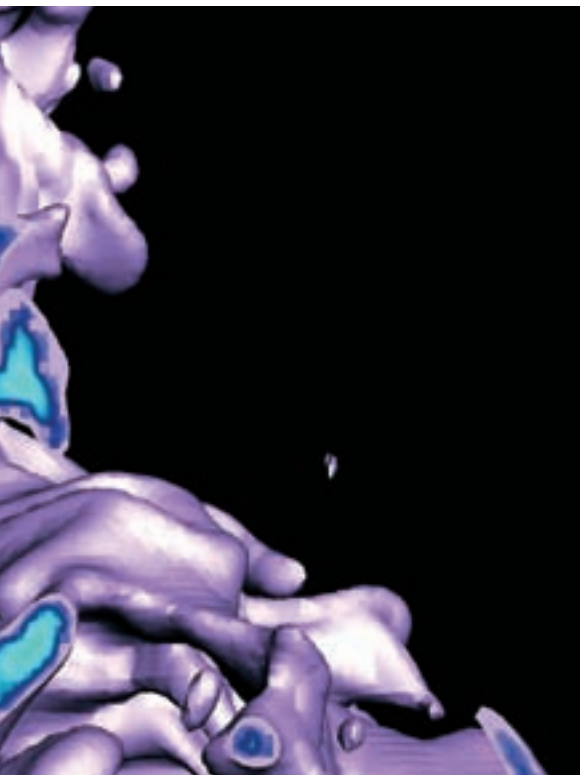
Figure 6. An experimental and theoretical Pourbaix diagram of $Ru(OH_2)(Q)(tpy)^{2+}$. The red dashed and solid blue lines correspond to the experimental pKa and redox potentials. The black lines are the theoretical predictions.

[3] James T. Muckerman, Dmitry E. Polyansky, Tohru Wada, Koji Tanaka, and Etsuko Fujita, "Water oxidation by a ruthenium complex with noninnocent quinone ligands: possible formation of an O–O bond at a low oxidation state of the metal," Inorganic Chemistry **47**, 1787 (2008).
[4] Ming-Kang Tsai, Jonathan Rochford, Dmitry E. Polyansky, Tohru Wada, Koji Tanaka, Etsuko Fujita, and James T. Muckerman, "Characterization of redox states of $Ru(OH_2)(Q)(tpy)^{2+}$ (Q = 3,5-di-*tert*-butyl-1,2-benzoquinone, tpy = 2,2′:6′,2″-terpyridine) and related species through experimental and theoretical studies," Inorganic Chemistry **48**, 4372 (2009).
[5] Jonathan Rochford, Ming-Kang Tsai, David J. Szalda, Julie L. Boyer, James T. Muckerman, and Etsuko Fujita, "Oxidation state characterization of ruthenium 2–iminoquinone complexes through experimental and theoretical studies," submitted to Inorganic Chemistry.

Low-Swirl Combustion



Experiments and simulations join forces to engineer a new type of clean, efficient burner



Project: Interaction of Turbulence and Chemistry in Lean Premixed Laboratory Flames

PI: John Bell, Lawrence Berkeley National Laboratory

Senior investigators: Marcus Day, Michael Lijewski and Vincent Beckner, LBNL

Funding: ASCR, INCITE, SciDAC

Computing resources: NERSC

One promising strategy for reducing U.S. dependence on petroleum is to develop new combustion technologies for burning hydrogen or hydrogen-rich syngas fuels obtained from the gasification of coal and biomass. At Lawrence Berkeley National Laboratory (LBNL), progress in both experiments and simulations is making one new technology, called *low-swirl combustion*, a frontrunner for this application.

Low-swirl combustion belongs to a larger category called *lean premixed combustion*, which is many researchers' technology of choice for increasing combustion efficiency while reducing nitrogen oxide (NOx) emissions. This technology uses a lower fuel-to-oxygen ratio than conventional burners and engines, and the fuel and oxygen are mixed before they enter the combustion chamber. (The mixing of fuel and air molecules is necessary for the chemical reactions of combustion to take place, but this is more commonly achieved by diffusion in the combustion chamber rather than premixing.) Together, these features result in a lower flame temperature, near zero emissions, and maximum fuel efficiency.

Two key characteristics of successful lean premixed burners will be fuel flexibility and flame stability. Researchers and equipment manufacturers want systems that can run on a variety of fuels, including natural gas, hydrogen, syngas (a combination of hydrogen and carbon monoxide), and mixtures of these fuels. However, flames from fuels with high hydrogen content are much more reactive than natural gas flames — they burn in discontinuous segments rather than in the highly wrinkled but continuous reacting layer that we see in most flames. The high reactivity and flame discontinuity can result in problems such as local extinction, where portions of the flame disappear; flashback, where the flame enters the fuel source; or global blowout, the complete extinction of the flame. In diffusion combustion systems, the flame location and stability are controlled by the mixing of fuel and oxidizer; but premixed flames require device-scale stabilization to ensure safety and reliability.

A leading contender among lean premixed burner designs is the low-swirl burner (LSB) invented by Robert Cheng at Lawrence Berkeley National Laboratory (LBNL) and honored with a 2007 R&D100 Award (Figures 7 and 8). The basic LSB principle is fundamentally different than the conventional high-swirl combustion method and defies many established notions of turbulent flame properties and burner engineering concepts. The swirler in the

LSB makes the lean premixed fuel flow expand and slow down as it exits the nozzle, which stabilizes the flame.

Cheng originally developed the low-swirl burner as a versatile laboratory tool for studying turbulence-flame interaction processes, but he soon recognized its practical potential. The LSB has already been scaled for devices ranging in size from home furnaces to industrial boilers and heaters. And it is more economical to manufacture and operate than many conventional burners, while emitting 10 to 100 times less smog-forming NOx.

In the United States alone, Cheng says, LSBs could eliminate about 740,000 tons per year of nitrogen oxides emitted to the atmosphere by boilers, furnaces, and turbines. That's equivalent to the emissions of nearly 100 coal-fired power plants.

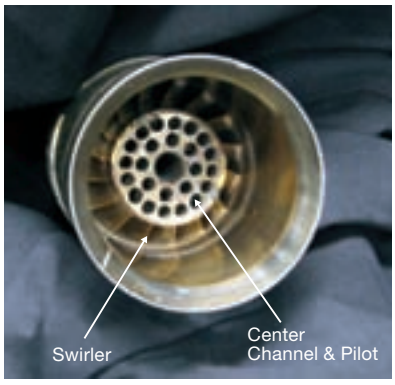


Figure 7. The heart of the low-swirl burner is a vane-swirler that has two flow passages. The fuel/air mixture flows through the openings of the center channel and the gaps between the surrounding swirl vanes. This design creates the low-swirl flow which supports a stable lifted or floating flame.

For much larger turbines capable of generating 250 MW of electricity, the Department of Energy is supporting research to evaluate the low-swirl combustion technology as a candidate for the hydrogen turbines in DOE Office of Fossil Energy's FutureGen Clean Coal Project. FutureGen is a public-private partnership to design, build, and operate the world's first zero-emissions fossil fuel power plant, using the Integrated Gasification Combined Cycle (IGCC) approach to produce hydrogen, which is separated from a concentrated CO_2 stream. The CO_2 is then sequestered in the earth, preventing emissions to the atmosphere that contribute to climate change.

Engineering design of a fuel-flexible gas turbine for utility-scale power plants presupposes a

Figure 8. The low-swirl burner remains cool to the touch because the lifted flame does not heat up its body. The lifted flame originally was thought to be highly undesirable because in other burners it leads to unstable flame behaviors. But the unique divergent flow field of the LSB allows the lifted flame to self-adjust and remain robust even with a very lean fuel mixture.

fundamental understanding of ultra-lean premixed combustion of high-hydrogen fuels that simply does not yet exist. Burning hydrogen in a gas turbine presents significant technical and engineering challenges because of the high reactivity of hydrogen, its fast flame speed, and the propensity of the hydrogen/air mixture to auto-ignite and explode.

Direct numerical simulations

Continued development of low-swirl burners and other advanced combustion technologies for electricity generation depends on improving our understanding of basic flame structure, stabilization mechanisms, emissions, and response to changes in fuel.

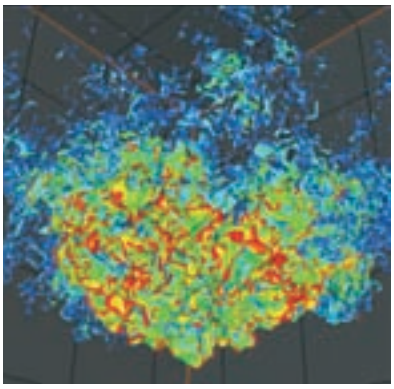


Figure 9. This snapshot of a hydrogen-air flame simulation is taken from below the flame, looking up. It shows an isotherm in the flame zone, colored by the local rate of fuel consumption, with red areas burning intensely and blue areas extinguished.

Numerical simulation has the potential to address some of these issues, but simulation of advanced burners has proven to be difficult because of the large range of spatial and temporal scales in these systems. The bulk of the analysis to date has been based on laboratory experiments.

Despite these difficulties, LBNL's Center for Computational Sciences and Engineering (CCSE), supported by two successive INCITE grants totaling 5.6 million processor hours on Franklin, has succeeded in producing a series of direct numerical simulations of lean, premixed methane and hydrogen flames on laboratory-scale low-swirl burners (Figures 9 to 11).[6,7] CCSE leader John Bell and his team, which includes Marc Day, Mike Lijewski, and Vince Beckner, created the simulations using a low Mach number formulation and adaptive mesh refinement, which dynamically focuses computational effort where it is needed during a simulation.

"There are significant gaps in our understanding of combustion," Bell says. "Theory provides a foundation for basic flame physics but can't address the complexity of realistic flames. Laboratory measurements are difficult to make and limited in the detail they provide. Computation, with its ability to deal with complexity and unlimited access to data, has the potential for closing the gap between theory and experiment and enabling dramatic progress in combustion science."

CCSE's low-swirl flame simulations successfully captured the cellular structure of hydrogen flames and provided a quantitative characterization of enhanced local burning structures, including the fuel consumption rate, which experiments cannot directly measure. The simulations directly modeled detailed chemical reactions and transport, which produced more accurate results than the researchers could have obtained using standard turbulence/chemistry interaction models, which were designed to describe a more continuous flame structure.

Future work will involve incorporating emission chemistry and further characterizing the burner configurations under a variety of fuels relevant to ultra-low emission burner scenarios. Related studies currently under way are examining turbulence/chemistry interactions across a much broader range of fluctuation intensities, and in high-pressure environments.

This work lays the foundation for close collaboration between computational and experimental combustion scientists. Cheng provided key experimental parameters and data that Bell and colleagues used to design their simulations. The simulations,

[6] J. B. Bell, R. K. Cheng, M. S. Day, V. E. Beckner and M. J. Lijewski, "Interaction of turbulence and chemistry in a low-swirl burner," *SciDAC 2008*, J. Phys.: Conf. Ser. **125**, 012027 (2008).
[7] M. S. Day, J. B. Bell, R. K. Cheng, S. Tachibana, V. E. Beckner and M. J. Lijewski, "Cellular burning in lean premixed turbulent hydrogen-air flames: Coupling experimental and computational analysis at the laboratory scale," *SciDAC 2009*, J. Phys.: Conf. Ser. **180**, 012031 (2009).

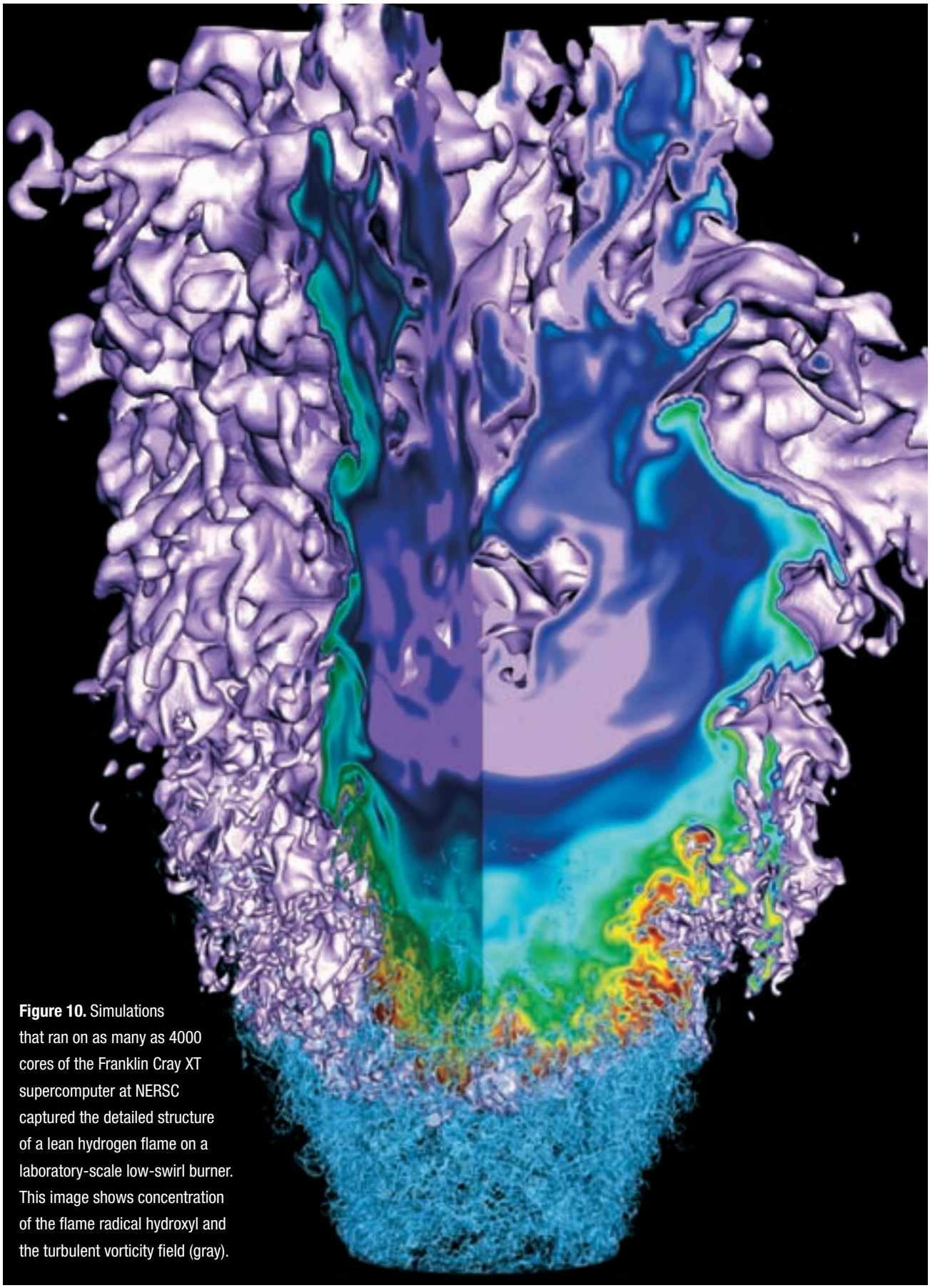


Figure 10. Simulations that ran on as many as 4000 cores of the Franklin Cray XT supercomputer at NERSC captured the detailed structure of a lean hydrogen flame on a laboratory-scale low-swirl burner. This image shows concentration of the flame radical hydroxyl and the turbulent vorticity field (gray).

in turn, provided more detailed data on the combustion process that Cheng is now using to include in the analytical models of low-swirl combustion.

"Even with the best available laser diagnostic techniques, it would be extremely challenging to obtain high fidelity information on the chaotic and highly turbulent flame processes," Cheng says. "John's research gives us a rare window into these 3D time-dependent complex processes to help us gain important insights on the effects of hydrogen on turbulent flame speed and pollution formation." Combustion models that predict these effects will guide engineers in the design of the gas turbines for IGCC power plants.

Recent laboratory studies at simulated gas turbine conditions have shown that the low-swirl burner can operate with natural gas, syngas, and hydrogen while meeting DOE's aggressive emission goal of <2 ppm NOx (corrected to 15% O_2). Cheng is very pleased that the low-swirl combustion method can overcome the challenges associated with very energetic hydrogen flames and looks forward to a long-term collaboration with CCSE to building a scientific foundation for the engineering development. His effort is one of many among universities, national laboratories, and equipment manufacturers to address the challenges of the Integrated Gasification Combined Cycle approach. The impact of IGCC will be significant. Cheng calculates that it would reduce greenhouse gas emissions by an average of 1.8 million metric tons of CO_2 and by 4,000 metric tons of NOx per 250 MW power plant per year.

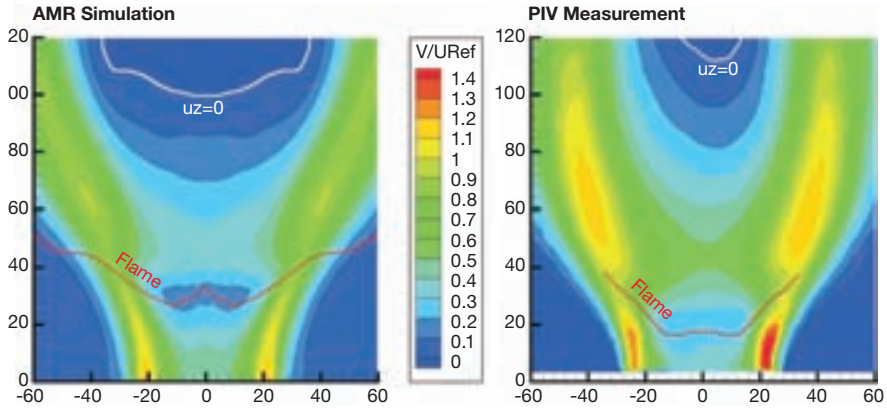
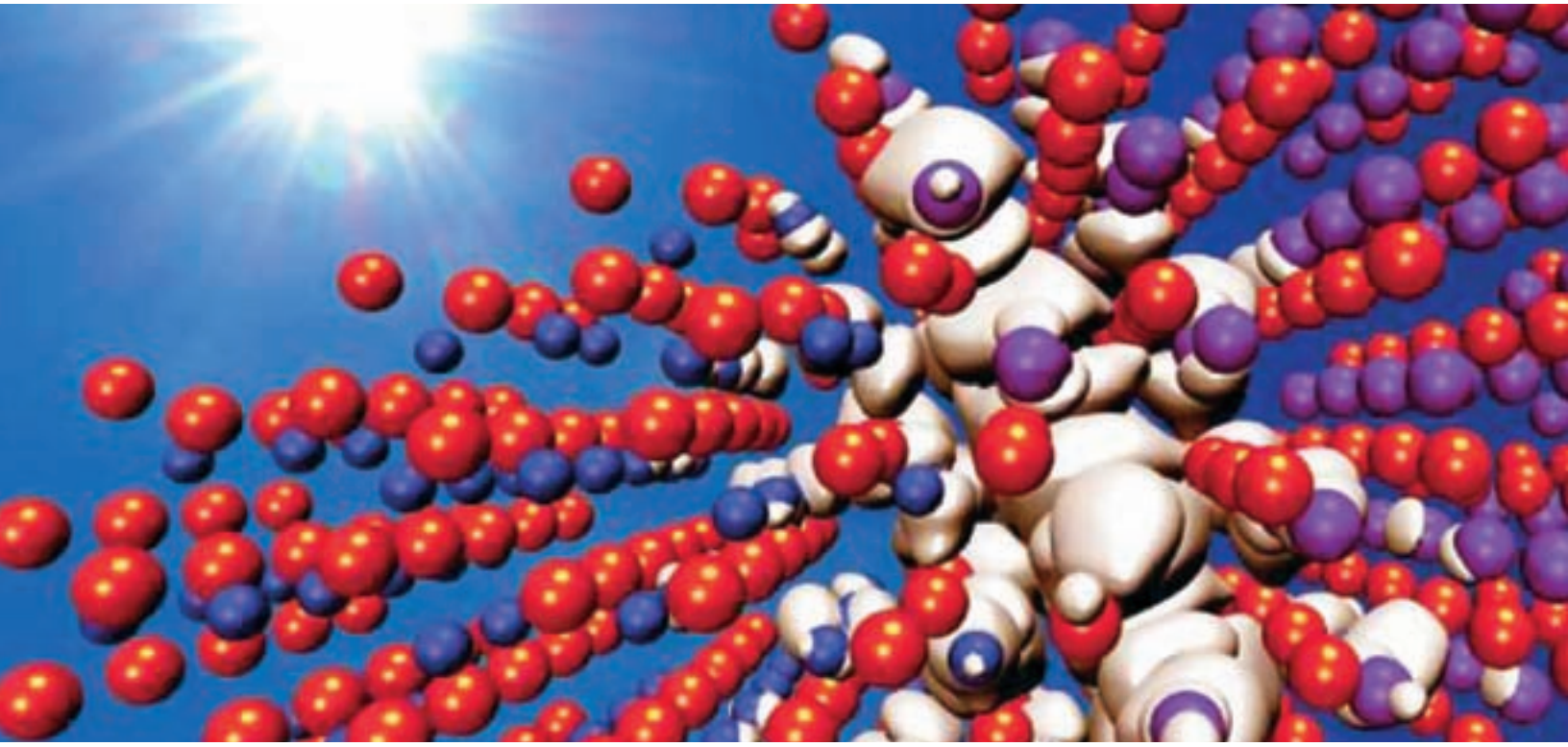
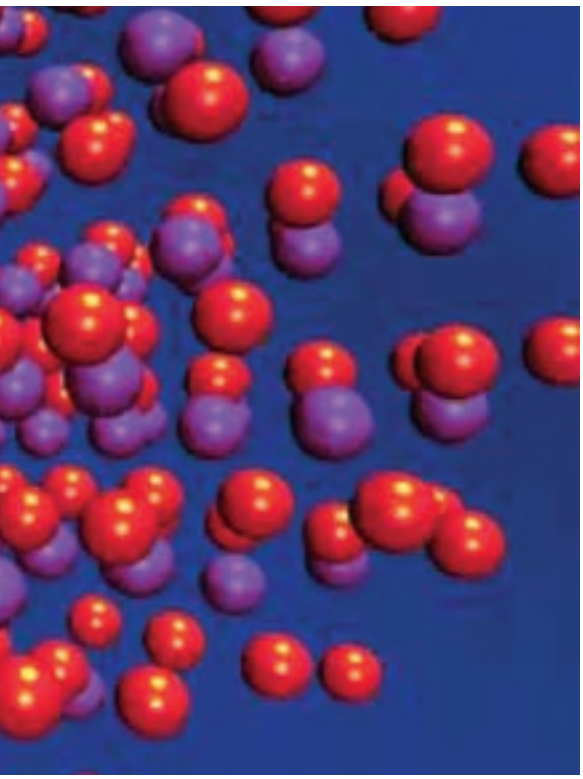


Figure 11. Comparison of the mean vertical velocity (scaled by the mean fueling rate) for the simulated ideal low-swirl nozzle, and a representative experimental profile. A number of key features are recovered in the simulations, including the weak recirculation approximately 10 cm downstream of the nozzle. The computed flame position sits somewhat downstream of the measured data, consistent with the simulation's higher mean velocity in the core.

Divide and Conquer



LS3DF code enables fast and accurate simulations of nanostructures for solar cells



Project: Large-Scale Nanostructure Electronic Structure Calculations

PI: Lin-Wang Wang, Lawrence Berkeley National Laboratory

Senior investigators: Nenad Vukmirovic, Sefa Dag, Maia Garcia Vergniory, and Shuzhi Wang, LBNL

Funding: BES, SciDAC, INCITE

Computing resources: NERSC, OLCF, ALCF

Nanostructures — tiny materials 100,000 times finer than a human hair — may hold the key to energy independence. A fundamental understanding of nanostructure behaviors and properties could provide solutions for curbing our dependence on petroleum, coal, and other fossil fuels by providing new and more efficient technologies for generating, storing, and transmitting electricity. For example, nanostructure systems are cheaper to produce than the crystal thin films used in current solar cell designs for harvesting solar energy, and offer the same material purity. Nanocatalysts may be the key to more cost-efficient hydrogen fuel cells. Even lighting and batteries may be revolutionized by nanotechnology.

Because nanomaterials are so tiny and have such a large surface-to-volume ratio, they behave very differently from the same materials in bulk forms, and numerical simulations are often necessary to understand their behavior. Scientists who create these numerical simulations are interested mainly in the location and energy level of electrons in the system — information that determines a material's properties. But unlike bulk systems, nanostructures cannot be represented by just a few atoms. They are coordinated systems, and any attempt to understand the materials' properties must simulate the system as a whole.

Density functional theory (DFT) allows physicists to simulate the electronic properties of materials, but DFT calculations are time-consuming; and any system with more than 1,000 atoms quickly overwhelms computing resources, because the computational cost of the conventional DFT method scales as the third power of the size of the system. Thus, when the size of a nanostructure increases 10 times, computing power must increase 1,000 times. The nanosystems of interest to energy researchers often contain tens of thousands of atoms. So one of the keys to unleashing the power of nanotechnology is to find a way of retaining DFT's accuracy while performing calculations with tens of thousands of atoms.

With petascale computers now providing more than 100,000 processor cores, the simulation of complete nanosystems should be possible if we can find a way to make the computational cost scale linearly to the size of systems. Researchers at Lawrence Berkeley National Laboratory have demonstrated a way to accomplish this using a divide-and-conquer algorithm implemented in the new Linear Scaling Three-Dimensional Fragment (LS3DF)

method.[8,9,10] In November 2008, this research was honored with the Association for Computing Machinery (ACM) Gordon Bell Prize for Algorithm Innovation. Developers of LS3DF include Lin-Wang Wang, Byounghak Lee, Hongzhang Shan, Zhengji Zhao, Juan Meza, Erich Strohmaier, and David Bailey.

Modeling solar cell efficiency

In a solar cell, there are a few key steps that determine overall efficiency in the conversion of sunlight to electricity: light absorption, exciton generation, exciton dissociation into separated electron and hole, carrier transport, and charge transfer across nanocontacts. A few aspects of nano solar cells often limit their overall efficiency: weak absorption of light, electron–hole recombination, nanocontact barriers, or large overpotentials. Unfortunately, many of these processes are not well understood. In some steps, such as light absorption and exciton dissociation, the quantitative aspects are not well known. In other steps, such as carrier transport, charge transfer, and nonradiative recombination, even the qualitative picture is missing. This lack of theoretical understanding hinders our ability to design more efficient solar cells.[11]

Trial and error can't address the dynamics of these complex nanostructure systems, says Lin-Wang Wang. "To reach that understanding, simulation does play and will continue to play a very important role."

Designing a nanostructure with confidence about its performance is important, because experimentally it might take years to realize one particular solar cell design. Theoretical simulation can play an important role in this design process, and can also help in testing new materials. For example, materials that have separate electron states within the energy band gap, such as zinc tellurite oxide (ZnTeO), have been proposed as second-generation solar cells. Such systems could theoretically increase solar cell efficiencies from 30% to 63%.

The band gap (the energy difference between the highest electron-occupied state and the lowest unoccupied state) for pure zinc tellurite (ZnTe) is about 2.25 eV. That means only the portion of sunlight with a photon energy higher than 2.25 eV can be absorbed and used to generate electricity. However, if there are some electron states in the middle of the ZnTe band gap, these states can be used as a stepping stone for the lower-energy sunlight to be absorbed as well, increasing the efficiency of a solar cell.

The challenge is to find the appropriate material that has the desired electronic property. For the ZnTe system, it has been proposed

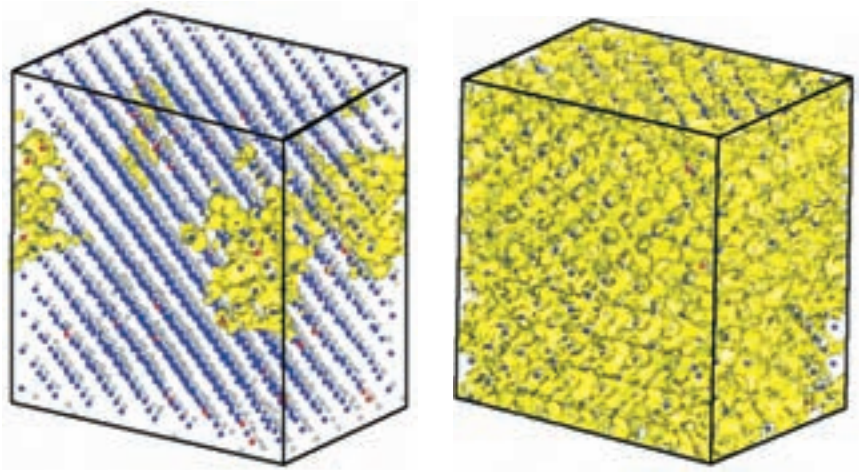


Figure 12. The electron wave functions for an oxygen-induced state (left) and ZnTe conduction band edge state (right) in a ZnTe0 alloy with 3% 0. The grey, blue, and red dots correspond to Zn, Te, and 0 atoms respectively.

[8] L. W. Wang, B. Lee, H. Shan, Z. Zhao, J. Meza, E. Strohmaier, and D. Bailey, "Linear scaling 3D fragment method for large scale electronic structure calculations," *Proceedings of SC08, Nov. 2008*; Lawrence Berkeley National Laboratory technical report LBNL-603E.
[9] L. W. Wang, Z. Zhao, J. Meza, "A linear scaling three dimensional fragment method for large scale electronic structure calculations", Phys. Rev. B **77**, 165113 (2008).
[10] Z. Zhao, J. Meza, and L. W. Wang, "A divide-and-conquer linear scaling three dimensional fragment method for large scale electronic structure calculations," J. Phys: Conds. Matt. **20**, 294203 (2008).
[11] L. W. Wang, "Computational challenges for nanostructure solar cells," Energy & Environmental Science **2**, 944 (2009).

that the addition of oxygen (O) could induce some electron states in the middle of the gap, but it was not known whether such states would be completely separated (by a gap) from the bottom of conduction band. If they were connected, then the excited electrons in the conduction band would relax into the oxygen-induced states by emitting phonons, which would make the system unusable for solar cell applications.

In experiments, only 3% to 6% of O atoms have been incorporated into the ZnTe system. In order to accurately model the random distribution of the O atoms in the system, large supercells need to be simulated. This makes calculations using a direct DFT method impractical; for example, a 13,824-atom ZnTeO supercell would require four to six weeks or more to run on 20,000 processor cores. But using LS3DF, the electron wave function of a 13,824-atom ZnTeO supercell was calculated in just a few hours on 17,280 cores of NERSC's Cray XT4, Franklin.

Simulations of the ZnTeO supercell at NERSC showed that the oxygen-induced electron state is localized around the O clusters, but the electron state in the conduction band is delocalized through the whole system (Figure 12). The clustering of the oxygen-induced states will have an impact on the mobility of the electrons in those states. In other calculations done at NERSC and at the Oak Ridge Leadership Computing Facility (OLCF), Wang and his colleagues found that there is a 0.2 eV band gap between the ZnTe conduction band and the oxygen-induced band, implying that this alloy can indeed be used for solar cell applications. The oxygen-induced states form a very broad band (0.7 eV) inside the band gap of ZnTe; nevertheless, the numerical simulation indicates that the width of the intermediate state will not significantly reduce the overall power efficiency. The theoretical power efficiency of the ZnTeO is estimated to be around 60%.

The approach discussed here can be used not just to analyze materials, but also to simulate nanostructures specifically designed for solar cells. Figures 13 and 14 show one such design using a zinc oxide/zinc sulfide (ZnO/ZnS) core/shell nanowire. Oxides are abundant and chemically stable, but they have large band gaps. One way to achieve a smaller band gap and increase the photovoltaic efficiency is to have a small type-II core/shell structure like ZnO/ZnS. In this case, photon absorption can pump an electron from the valence band of ZnS to the conduction band of ZnO. This effectively reduces the band gap of the whole system from the original 3.4 eV to 2.1 eV, thus significantly increasing the theoretical solar cell efficiency. Calculations show that if the system is small, this absorption can be strong enough for solar cell applications. Another advantage of core/shell nanowire band alignment is that the electron can travel in the ZnO shell while the hole travels in the ZnS core without the danger of optical recombination, which is a major carrier loss mechanism in many oxides.

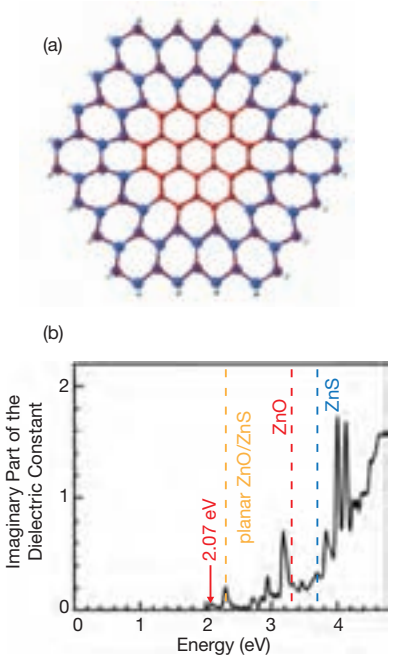


Figure 13. (a) Relaxed geometry of a ZnO/ZnS core/shell nanowire. (b) Calculated optical-absorption using the band-corrected pseudopotentials. The vertical red (ZnO) and blue (ZnS) dashed lines correspond to the bulk band gaps, and the vertical orange line corresponds to the band gap of the planar ZnO/ZnS heterostructure.

Figure 14. The exciton wave function (the white isosurfaces) at the interface of a ZnO/ZnS nanowire. (Image ©2009, RSC Publishing)

The LS3DF method

A divide-and-conquer scheme is a natural approach for mapping the physical locality of a large problem to the architectural locality of a massively parallel computer. LS3DF is based on the observation that the total energy of a given system can be broken down into two parts: the electrostatic energy and the quantum mechanical energy (i.e., kinetic and exchange-correlation energy). While the electrostatic energy is long-range and must be solved by a global Poisson equation, the computationally expensive quantum mechanical energy is short-range and can be solved locally. The idea is to divide the whole system into small fragments, calculate the quantum mechanical energies of these fragments, and then combine the separate fragment energies to obtain the energy of the whole system (Figure 15).

A critical issue in a divide-and-conquer scheme such as this is how to combine (patch) the fragments. The core of LS3DF is a novel patching scheme that cancels out the artificial boundary effects caused by the division of the system into smaller fragments. As a result of this cancellation, the results are essentially the same as a direct calculation on the large system. In LS3DF, once the fragment sizes are chosen to obtain a given numerical accuracy, the computational cost is proportional to the number of fragments—hence the name, Linearly Scaling Three-Dimensional Fragment method. In contrast, the cost of conventional methods scales as $O(N^3)$, where N is the size of the system in atoms.

Using a small group of computing cores to solve the quantum mechanical part of each fragment independently (Figure 16), LS3DF also scales in performance almost perfectly with the number of cores. Only a small overhead is needed to patch the fragment charge densities into a global charge density, and to solve the Poisson equation for the whole system. As a result, this method can be employed on computer systems with more than 100,000 cores.

On 36,864 cores of the NERSC dual-core Cray XT4 (Franklin), LS3DF achieved 135 Tflop/s or 40% of the peak floating-point performance of the machine. In a separate test, LS3DF ran on 147,456 cores on the quad-core Cray XT5 (Jaguar) at the Oak Ridge Leadership Computing Facility, reaching 442 Tflop/s at 33% of the theoretical peak. In a run on the IBM BlueGene/P system (Intrepid) at the Argonne Leadership Computing Facility, the code achieved 224 Tflop/s on 163,840 cores, or 40% of peak. This makes LS3DF the first variationally accurate linearly scaling *ab initio* electronic structure code that has been efficiently parallelized to such a large number of processors.

"The efficiency of LS3DF on these large computer systems is impressive, but the real story is the power of algorithms," said LS3DF team member Juan Meza, who heads Berkeley Lab's High Performance Computing Research Department. "Using a linear scaling algorithm, we can now study systems that would otherwise take over 1,000 times longer on even he biggest machines today. Instead of hours, we would be talking about months of computer time for a single study."

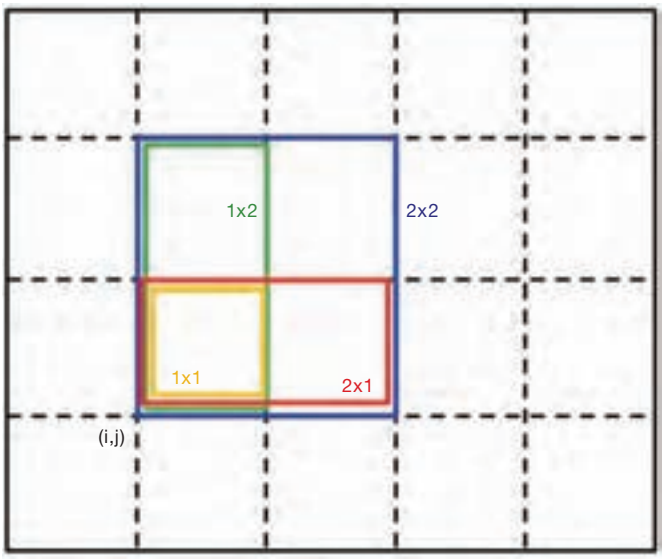


Figure 15. A two-dimensional schematic illustrating how LS3DF divides the space into grids and fragments of different sizes.

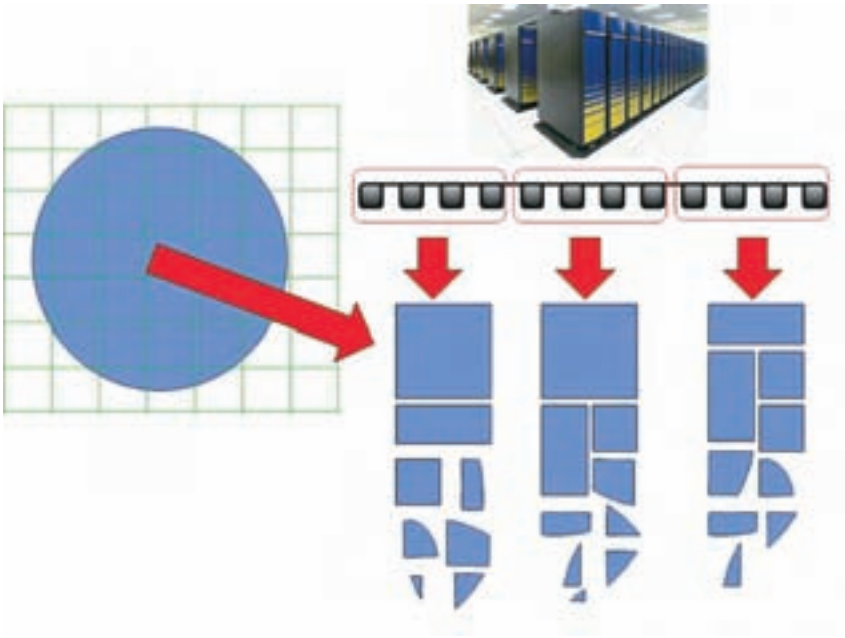


Figure 16. A schematic of the LS3DF parallelization method and distribution of fragments among processor groups.

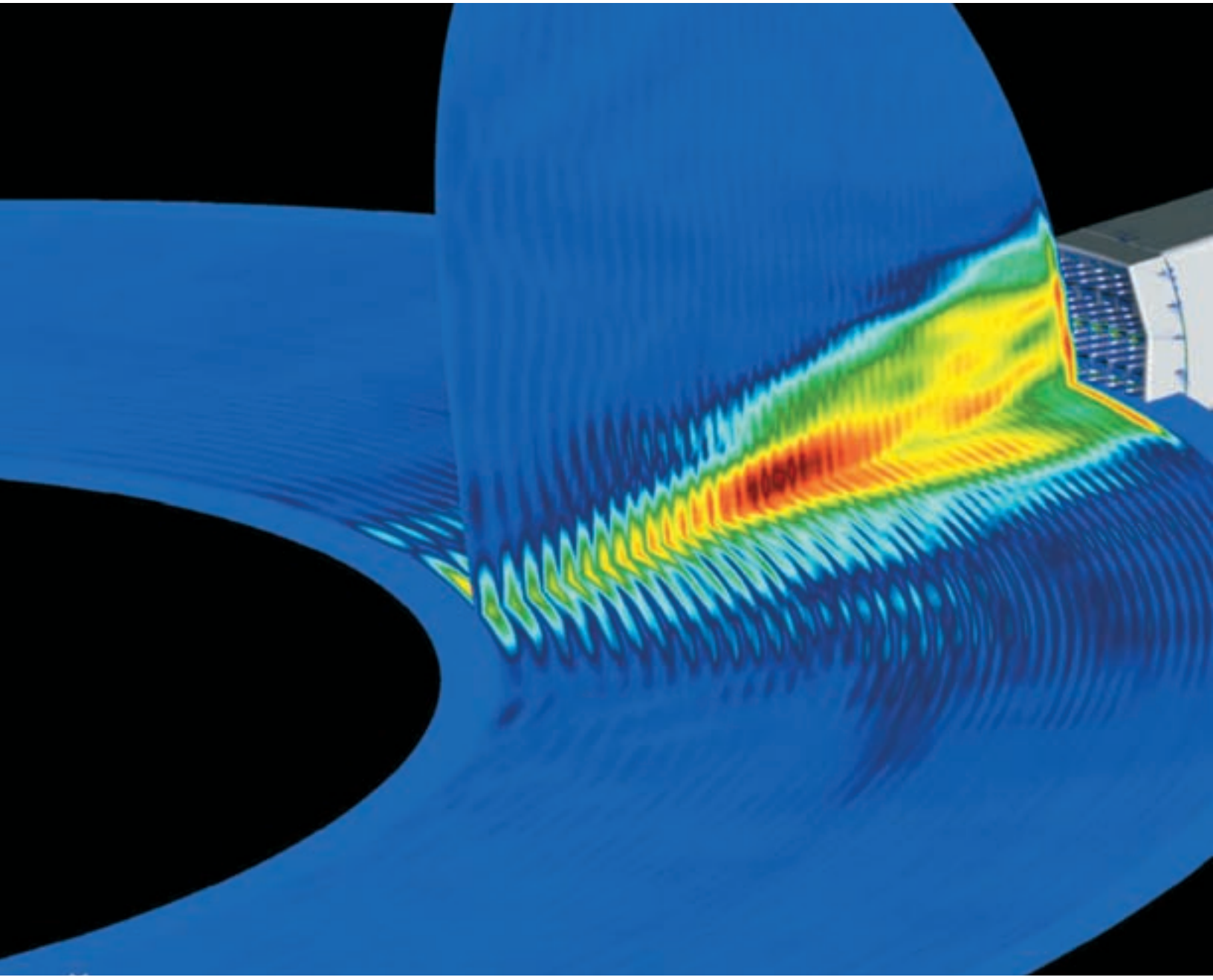
Getting codes to run with such high efficiencies on massively parallel machines is not a trivial task. The LS3DF code is based on a standard plane wave DFT code, PEtot, which is used to calculate the fragment wave functions. The idea of LS3DF was proposed by Lin-Wang Wang and is based on the nearsightedness principle of Prodan and Kohn, which states that many physical properties, such as density, depend only on the nearby environment. Zhengji Zhao implemented the first version of LS3DF on the IBM SP3 (Seaborg) at NERSC, where the initial tests were carried out. Subsequently, Byounghak Lee improved the PEtot part of the code by changing one of the algorithms to use the more efficient BLAS3 libraries instead of BLAS2. As a result, the efficiency of the PEtot code improved from 15% to over 50%.

Hongzhang Shan, David Bailey, and Erich Strohmaier of the SciDAC Performance Engineering Research Institute (PERI) worked hand in hand with Wang and his colleagues to analyze the performance of LS3DF on NERSC's Cray XT4 (Franklin) and to identify potential performance improvements. Shan wrote two data communication routines to compute the global charge density and distribute the global potential to each fragment, allowing the code to run without having to do any significant input/output (I/O).

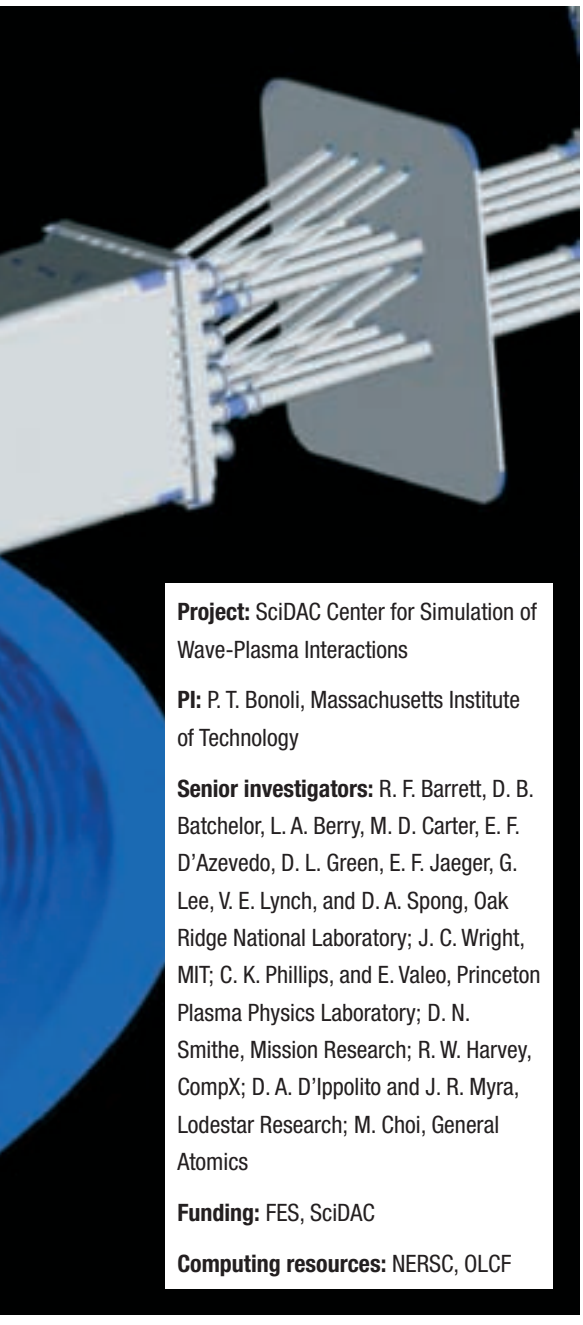
Later, when the code was tested on the Intrepid IBM BlueGene/P machine at ALCF, memory became an issue. This was overcome by carefully reviewing the memory needs and removing unnecessary arrays. With the help of ALCF staff, it was also found that a careful mapping of the processors on the BlueGene/P machine can significantly improve the efficiency of LS3DF.

Once the LS3DF code had been optimized, it was a matter of days before it was running at each of the DOE supercomputing facilities. Development and testing of LS3DF on several different architectures has contributed to its robustness and efficiency.

Harnessing Plasma Waves



Simulations of wave-plasma interactions pave the way toward practical fusion energy



Project: SciDAC Center for Simulation of Wave-Plasma Interactions

PI: P. T. Bonoli, Massachusetts Institute of Technology

Senior investigators: R. F. Barrett, D. B. Batchelor, L. A. Berry, M. D. Carter, E. F. D'Azevedo, D. L. Green, E. F. Jaeger, G. Lee, V. E. Lynch, and D. A. Spong, Oak Ridge National Laboratory; J. C. Wright, MIT; C. K. Phillips, and E. Valeo, Princeton Plasma Physics Laboratory; D. N. Smithe, Mission Research; R. W. Harvey, CompX; D. A. D'Ippolito and J. R. Myra, Lodestar Research; M. Choi, General Atomics

Funding: FES, SciDAC

Computing resources: NERSC, OLCF

The next step toward fusion as a practical energy source is currently under way in the design and construction of ITER, a magnetic confinement device in southern France, supported by an international collaboration, which will be capable of producing and controlling the high performance plasma required for self-sustaining fusion reactions, known as *burning plasma*.

High-power radio frequency (RF) waves in the ion-cyclotron range of frequencies (ICRF) will heat the deuterium and tritium fuel in the ITER plasma in the same way that microwaves transfer heat to food in a microwave oven: the energy carried by the RF waves will be transferred to particles in the plasma, increasing their velocity and temperature. Rf waves can also be used in several different ways to control plasma behavior. But realizing this potential requires theoretical understanding and accurate modeling of these processes and their coupling to other plasma processes such as plasma transport, magnetohydrodynamic (MHD) stability, microstability, and turbulence.

The SciDAC Center for Simulation of Wave-Plasma Interactions (CSWPI) — a team of plasma scientists, computer scientists, and applied mathematicians led by Paul Bonoli of the Massachusetts Institute of Technology — is laying the foundations for this understanding by developing and applying advanced wave solvers, optimizing these solvers on the most powerful computers, and demonstrating the feasibility of coupling these solvers to Fokker-Planck solvers for particle distribution and to Monte Carlo orbit codes.

CSWPI is building these codes into a comprehensive wave-plasma simulation capability that can accurately determine the spectrum of waves launched into the plasma from antenna structures, self-consistently treat the nonlinear interaction of the waves with the plasma kinetics, and accurately calculate power, current, and flow deposition for plasma control. They are validating these models by application to experiments, and applying them to support the burning plasma effort through RF system design and analysis of operating scenarios. The results of this research will eventually be linked with other SciDAC efforts in plasma transport, MHD, and turbulence, to create an integrated plasma simulation capability.

Two recent cover articles in the journal Physics of Plasmas have highlighted major accomplishments by the CSWPI collaboration (Figure 17).

Breakthrough simulations

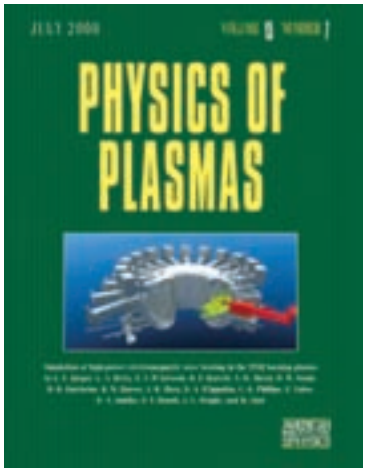
The first article, "Simulation of high-power electromagnetic wave heating in the ITER burning plasma" by E. F. Jaeger et al., reports on the first self-consistent three-dimensional simulations of this important phenomenon in ITER.[12] In order to heat the fuel effectively, the RF wave fields must couple efficiently to the dense core plasma. Calculations in this paper support the argument that this will be the case.

Extending the 3D full-wave calculations to the burning plasma regime of ITER, with its combination of physical size and high plasma density, required an order of magnitude higher resolution than previous calculations. To meet this challenge, the AORSA global wave solver was tuned for use on multicore processing architectures, with the dense linear system of complex coefficient equations solved using SCALAPACK. The result of this approach is that execution on dual- and quad-core Cray XT4 systems — Franklin at NERSC and Jaguar at Oak Ridge National Laboratory — has demonstrated excellent scaling up to tens of thousands of processors, achieving up to 154 trillion calculations per second (154 teraflops).

The visualization in Figure 18 shows the 3D wave fields from AORSA in the framework of the ITER vacuum vessel and antenna structure (~2 m high). A close-up of the antenna region in Figure 19 shows the highly focused wave fronts in both poloidal and equatorial planes. A correlation between the high RF field regions at the plasma edge and the antenna can clearly be seen. An ongoing task is to minimize these "hot spots" and to provide adequate protection and cooling for the antenna.

In Figures 18 and 19, the CSWPI team used the VisIt 3D visualization and analysis technology, developed in part by DOE's SciDAC program, to enhance their understanding of the complex physical phenomena. Three-dimensional visualizations such as these are becoming increasingly important for understanding and using the results of large simulations like AORSA. The capability to efficiently compute the 3D wave fields is adding to plasma researchers' understanding of antenna-edge interactions in ITER and will be important for coupling to MHD simulations.

Figure 20 compares the 3D wave field and plasma heating contours in ITER with those in two U.S. tokamak experiments currently in operation: the National Spherical Tokamak Experiment (NSTX) and Alcator C-Mod. Much of the difference between ITER and these smaller tokamaks can be attributed to the higher magnetic field strength and higher plasma pressure in ITER. The waves in ITER propagate radially inward with strong central focusing and little toroidal spreading, resulting in more efficient heating of the plasma core than in current tokamaks.



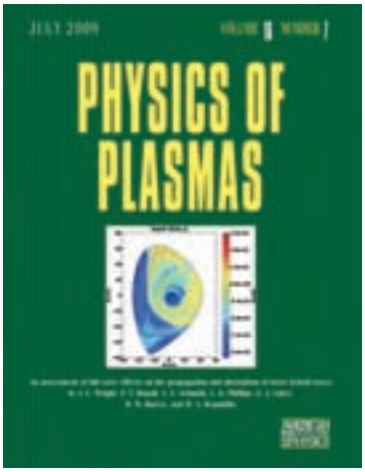


Figure 17. Two recent cover articles in the journal Physics of Plasmas have highlighted major accomplishments by the SciDAC Center for Simulation of Wave-Plasma Interactions: (left) "Simulation of high-power electromagnetic wave heating in the ITER burning plasma" and (right) "An assessment of full wave effects on the propagation and absorption of lower hybrid waves." (Images ©2008, 2009, American Institute of Physics)

[12] E. F. Jaeger, L. A. Berry, E. F. D'Azevedo, R. F. Barrett, S. D. Ahern, D. W. Swain, D. B. Batchelor, R. W. Harvey, J. R. Myra, D. A. D'Ippolito, C. K. Phillips, E. Valeo, D. N. Smithe, P. T. Bonoli, J. C. Wright, and M. Choi, "Simulation of high-power electromagnetic wave heating in the ITER burning plasma," Physics of Plasmas **15**, 072513 (2008).

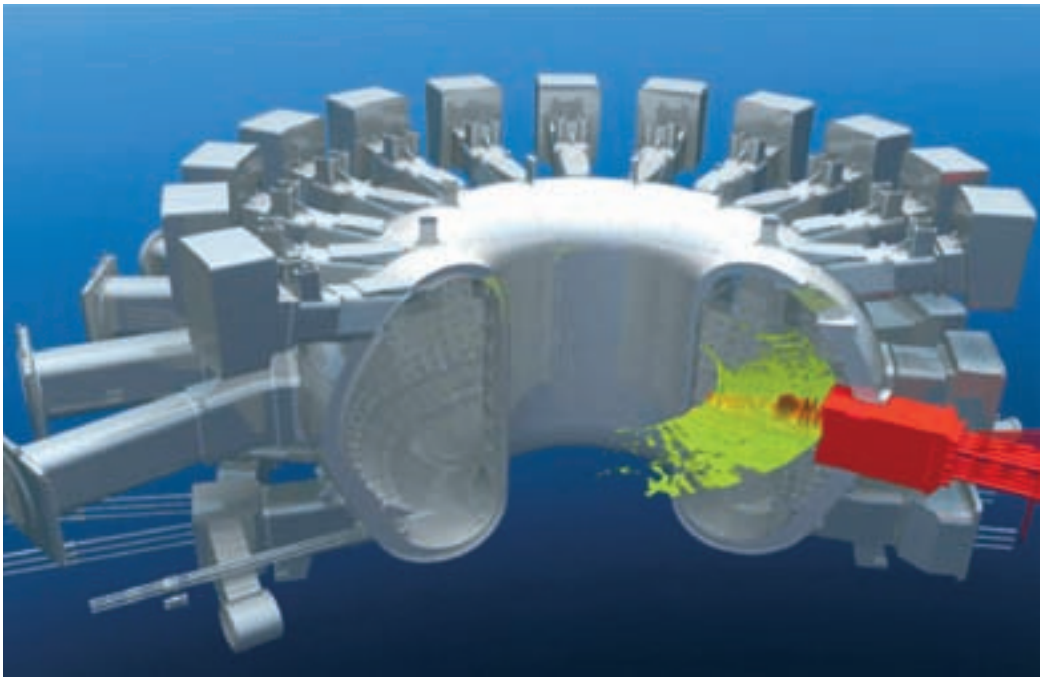


Figure 18. Visualization of the 3D wave fields in the equatorial plane of ITER showing the vacuum vessel and RF antenna structures.

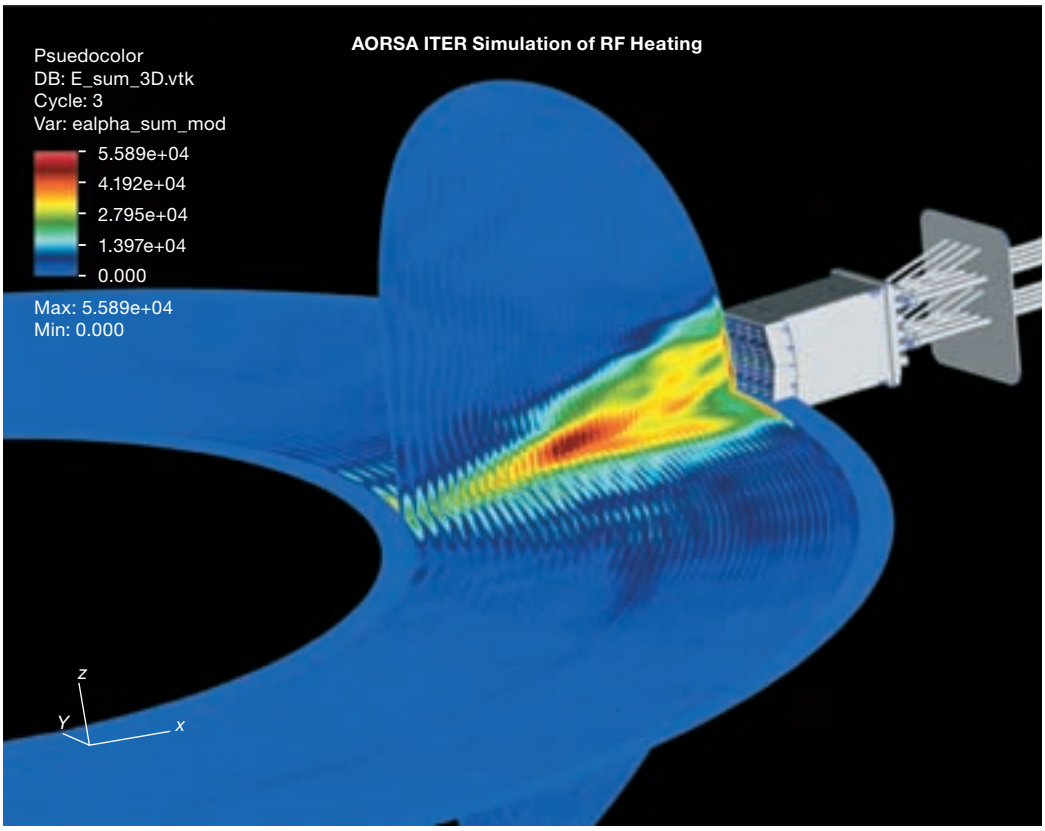


Figure 19. Close-up of the 3D wave fields in front of the ITER RF antenna (~2 m high).

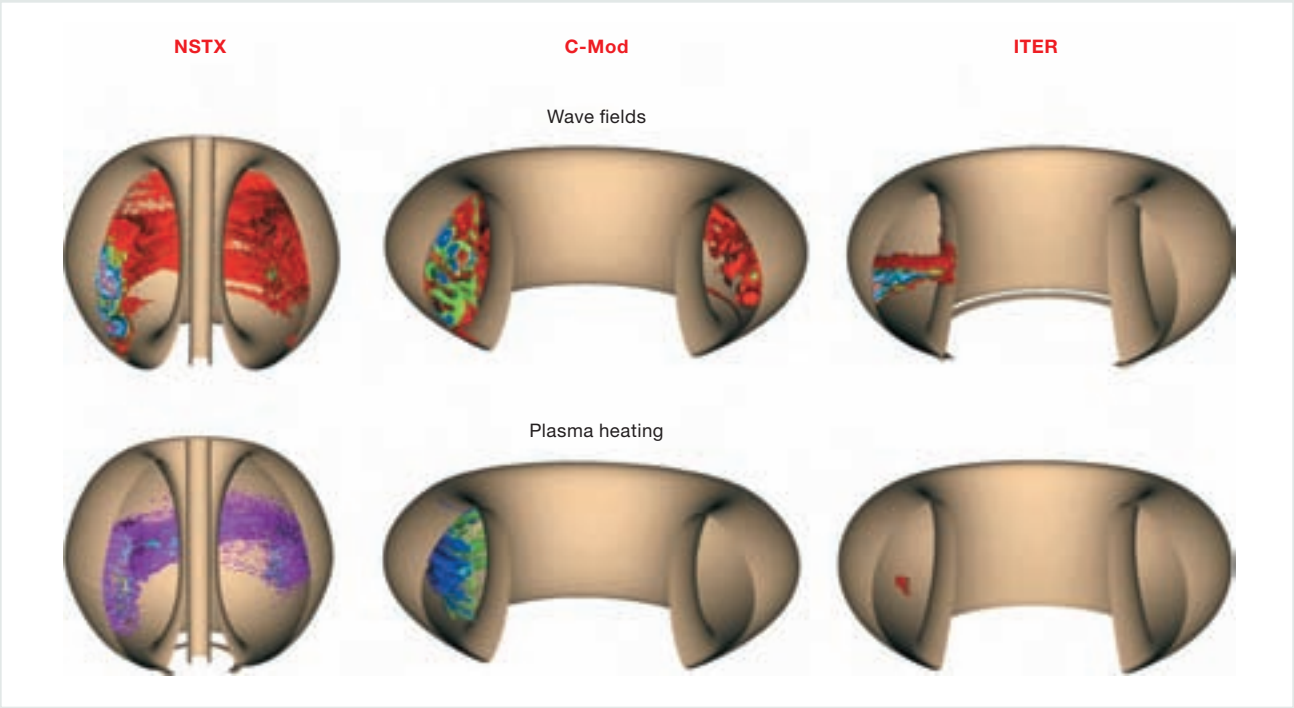


Figure 20. Three-dimensional wave fields (top) and plasma heating (bottom) for NSTX, Alcator C-Mod, and ITER. Figures are not to scale.

Understanding experimental results

The second paper, "An assessment of full wave effects on the propagation and absorption of lower hybrid waves" by J. C. Wright et al., presents the first ever self-consistent simulations of lower hybrid current drive (LHCD) utilizing both a full wave and Fokker-Planck calculation in toroidal geometry (Figure 21).[13]

LHCD has been shown to be effective for controlling the current profile in a tokamak in order to access regimes of improved energy confinement time and high fractions of non-inductive current driven by the pressure gradient, know as the "bootstrap" current. The current profiles needed to access these regimes are peaked far off-axis, and LH waves are especially effective at driving steady state currents at these locations. Thus it is important to develop and validate a predictive capability in this area.

For this study, the researchers used a massively parallel version of the TORIC full wave electromagnetic field solver coupled to the electron Fokker-Planck solver CQL3D to understand LH wave experiments in the Alcator C-Mod device. The TORIC code has achieved almost linear scaling from 32 to over 9000 Cray XT3/XT4 processors.

Scattering at cutoffs in the edge of the plasma has been identified as an important mechanism in LH wave diffraction and was shown to affect the power deposition location. Scattering manifests itself in low single-pass absorption regimes that are common in present day tokamaks with modest electron temperature where LHRF power is used with a low parallel index of refraction to maximize the driven current. Full wave simulations are necessary to accurately predict the location of the current driven in these devices. In higher-temperature devices such as ITER, there is high single-pass absorption, and scattering effects may play a minimal role; future studies will explore this possibility.

During the coming year, Bonoli and the CSWPI team plan to use computing resources at NERSC on two problems.

[13] J. C. Wright, P. T. Bonoli, A. E. Schmidt, C. K. Phillips, E. J. Valeo, R. W. Harvey, and M. A. Brambilla, "An assessment of full wave effects on the propagation and absorption of lower hybrid waves," Physics of Plasmas **16**, 072502 (2009).

They first plan to study the interaction of RF waves in the ion cyclotron range of frequencies (ICRF) with supra-thermal ion populations or "ion tails" that either have been created by the ICRF waves themselves, or are present due to neutral beam injection heating, or are created by plasma fusion processes. Recently they extended a combined full-wave and Fokker-Planck calculation to three spatial dimensions by including the full spectrum of toroidal harmonics for specific ICRF antenna geometries. They have also coupled the full-wave solver AORSA to two Monte Carlo orbit codes, sMC+rf and ORBIT-RF. About half of their NERSC allocation will be used to run this combined 3D model to simulate ICRF heating in the Alcator C-Mod device; to simulate high harmonic fast wave (HHFW) heating in the presence of fast neutral beam ions in the NSTX and DIII-D tokamaks; and to calculate ICRF heating in the ITER burning plasma.

The researchers also plan to use about half of their NERSC allocation to carry out 3D full-wave simulations of LH current drive using the combined full-wave and Fokker-Planck model that was described above. Three-dimensional solutions will be constructed by summing the RF diffusion coefficients from multiple toroidal modes that comprise the LH antenna spectrum. In LH current drive experiments, hard x-rays are produced by the relativistic motion of fast electrons that are generated by the LHRF power. In order to validate this simulation capability, they plan to compare simulated hard x-ray spectra with spectra measured during LH current drive experiments in the Alcator C-Mod tokamak.

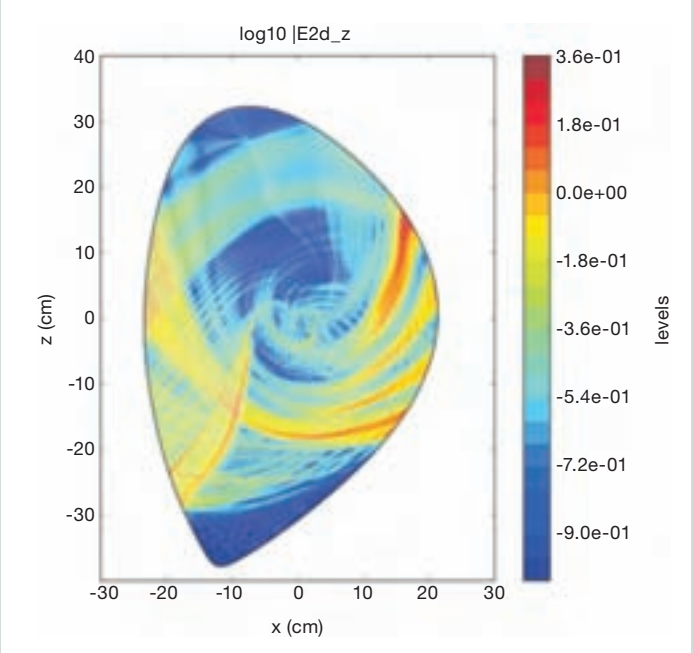
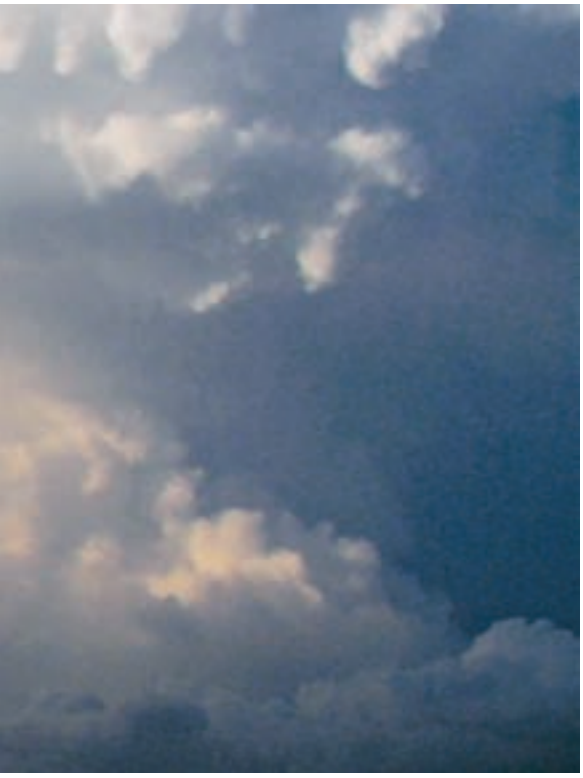


Figure 21. Full wave calculation of LH waves in Alcator C-Mod at electron density of 7×10^{19} m^{-3} using a self-consistent non-Maxwellian dielectric response achieved after four iterations. The electric field component parallel to the equilibrium magnetic field is shown.

Bringing Clouds into Focus

By simulating individual cloud systems, a new global climate model promises to reduce the uncertainty of climate forecasting



Project: Global Cloud Modeling

PI: David Randall, Colorado State University

Senior investigators: Akio Arakawa, University of California, Los Angeles; Celal Konor, Ross Heikes, Joon-Hee Jung, and Hiroaki Miura, Colorado State University

Funding: BER, SciDAC

Computing resources: NERSC, OLCF

Clouds exert two competing effects on the Earth's temperature: they cool the planet by reflecting solar radiation back to space, but they also warm the planet by trapping heat near the surface. These two effects coexist in a delicate balance.

In our current climate, clouds have an overall cooling effect on the Earth. But as global warming progresses, the cooling effect of clouds might be enhanced or weakened — global climate models are evenly divided on this issue. In fact, inter-model differences in cloud feedbacks are considered the principal reason why various models disagree on how much the average global temperature will increase in response to greenhouse gas emissions, when it will happen, and how it will affect specific regions.

Clouds also affect climate in other ways, such as transporting heat and moisture from lower to higher altitudes, producing precipitation, and many other interrelated mechanisms. Current global climate models are unable to directly simulate individual cloud systems from physical principles, because the size and speed of supercomputers place a limit on the number of grid cells that can practically be included in the model. As a result, global models do not have fine enough horizontal resolution to represent large clouds.

Instead, global climate models must rely on parameterizations, which are statistical representations of phenomena, such as cloud cover or precipitation rates, that cannot be directly modeled. Different models use different parameterizations, which is an important reason why their results differ. Cloud parameterizations are the greatest source of uncertainty in today's climate models.

David Randall, a Professor of Atmospheric Science at Colorado State University, is working to clear up that uncertainty by developing and testing a new kind of global climate model, called a global cloud resolving model (GCRM) — a model that's designed to take advantage of the extreme-scale computers expected in the near future.

Randall is the principal investigator of the "Global Cloud Modeling" project that computes at NERSC, and was one of two coordinating lead authors of Chapter 8, "Climate Models and Their Evaluation," in the Intergovernmental Panel on Climate Change's (IPCC's) Fourth Assessment Report, which was honored with the 2007 Nobel Peace Prize. He also directs the Center for Multiscale Modeling of Atmospheric Processes, sponsored by the National Science Foundation.

From a single thunderstorm to the whole earth

"The GCRM is a computer model that simulates the motions of the atmosphere on scales from a single thunderstorm all the way up to the size of the entire earth," Randall explains. "It has about a billion little grid cells to represent the three-dimensional structure of the air. Each grid cell has a wind, a temperature, a humidity, and some other things that are needed. So the number of numbers involved is in the tens of billions, just as a snapshot of what's going on at a given second."

Large thunderstorms play an important role in global atmospheric circulation (Figure 22). They pack a lot of energy in the form of updrafts that move, in extreme cases, 30 to 40 meters a second — "scary fast," Randall says. They "lift air from near Earth's surface to way up near the stratosphere in just a few minutes." In this way, thunderstorms carry moisture, momentum, carbon dioxide, and other chemical species through great depths of the atmosphere very quickly.

Cumulus clouds, Randall says, make the upper troposphere wet by transporting water from its source, the oceans. "A lot of it will rain out along the way, but some of it is still left, and it gets spread out up there and makes cirrus clouds, comprised largely of ice, which are very important for climate. We're especially interested to see how storms that create cirrus affect the climate." Cirrus clouds block Earth's infrared radiation from flowing out to space, and that tends to warm the climate. "If we have more cirrus in the future, that will enhance warming. If we have less, it will reduce the warming."

The GCRM also will give scientists new insights into tropical cyclones, which, Randall says, "are much bigger than thunderstorms, and in fact they contain many thunderstorms simultaneously. They affect the climate in part by cooling the sea surface as they move over the ocean."

The GCRM, supported by DOE's Scientific Discovery through Advanced Computing (SciDAC) program, is built on a geodesic grid that consists of about 100 million mostly hexagonal columns, each with 128 levels, representing layers of atmosphere that reach 50 kilometers above the Earth. For

Figure 22. Supercell storms have such powerful updrafts that the top of the cloud (or anvil) can break through the troposphere and reach into the lower levels of the stratosphere. During the late afternoon and early evening of April 3, 2004, this supercell thunderstorm dropped 2-inch-diameter hail over Chaparral, N.M., causing widespread damage. (Image: Greg Lundeen, NOAA)

each of these grid cells, the model predicts the wind, temperature, and humidity at points just 4 kilometers apart (with a goal of 2 kilometers on the next generation of supercomputers). That's an unprecedented resolution—most global atmospheric models provide detail at a 100-kilometer scale.

"No one has done this before in quite this manner, and it's our hope that our project will point the way to future generations of models," says Randall.

The geodesic grid used in the GCRM, also developed with SciDAC support, is itself quite innovative. If you want to tile a plane with regular polygons, you have only three choices: triangles, squares, or hexagons (Figure 23). Most climate models use some form of square (or rectangular) grid; but the geometry of the grid complicates the calculations, because each square has two different kinds of neighbors—four wall neighbors and four corner neighbors—which require different treatment in the equations. In addition, a square grid poses complications in modeling the Earth's polar regions, where grid cells lose symmetry because of longitudinal convergence. There are solutions to these problems, but they are computationally expensive.

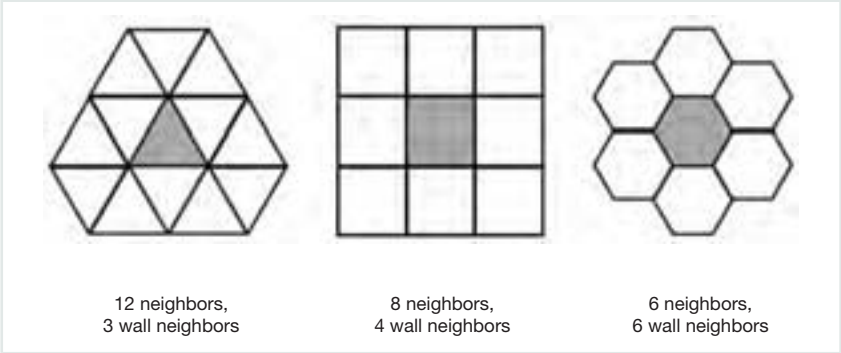


Figure 23. Small grids made up of equilateral triangles, squares, and hexagons. The hexagonal grid has the highest symmetry because all neighboring cells of a given hexagonal cell are located across cell walls. In contrast, with either triangles or squares, some neighbors are across walls, while others are across corners.

The GCRM, in contrast, uses a geodesic, hexagonal grid (Figure 24). In a hexagonal grid, all neighbors of a given cell lie across cell walls; there are no corner neighbors. A geodesic grid on a sphere has twelve pentagonal cells in addition to the many hexagonal cells; but each cell still has only wall neighbors, and all cells are roughly the same size. This type of grid also eliminates the pole problem.

As a result, equations constructed on hexagonal grids treat all neighboring cells in the same way, reducing the complexity and increasing the speed, productivity, and accuracy of the code. The number of cells (both grid columns and levels) can easily be changed for a particular computer run, depending on what the researchers want to simulate. Models based on geodesic grids are now being used by several major weather and climate modeling groups around the world.

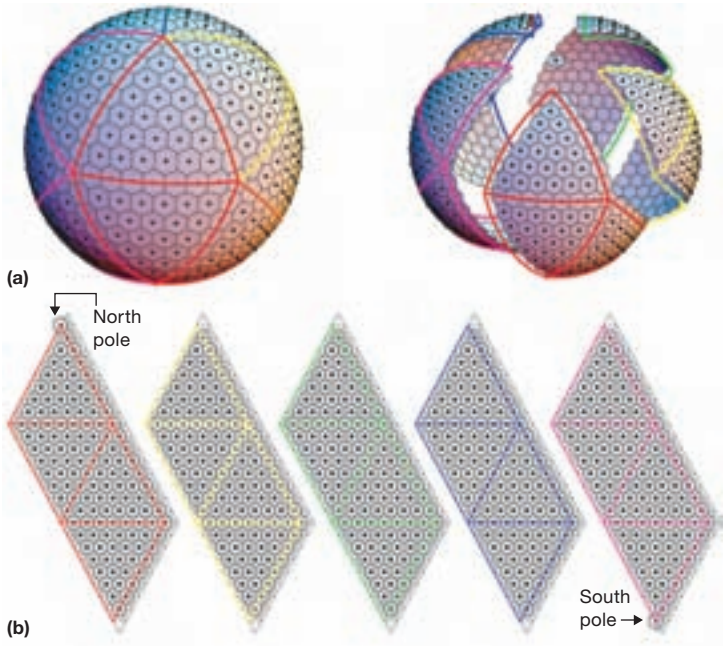


Figure 24. A spherical geodesic grid (a) can be cut into logically rectangular panels (b), which offers a convenient way to organize the data in a computer's memory. For visual clarity, this depiction shows a very low resolution grid.

Vorticity: Where the action is

Climate models are systems of partial differential equations that simulate how the atmosphere and oceans move and change over time, based on the laws of physics, fluid motion, and chemistry. Since the equations are all interrelated, the dynamical core of the model has to solve these equations simultaneously for every grid cell at each time step — which is why climate models require massive computing power.

Because the GCRM has such high resolution, Randall's research team knew they needed to use equations that reproduce accurate motions at a wide range of scales to get the most realistic results; so team members Akio Arakawa of the University of California, Los Angeles (UCLA), and Celal Konor of Colorado

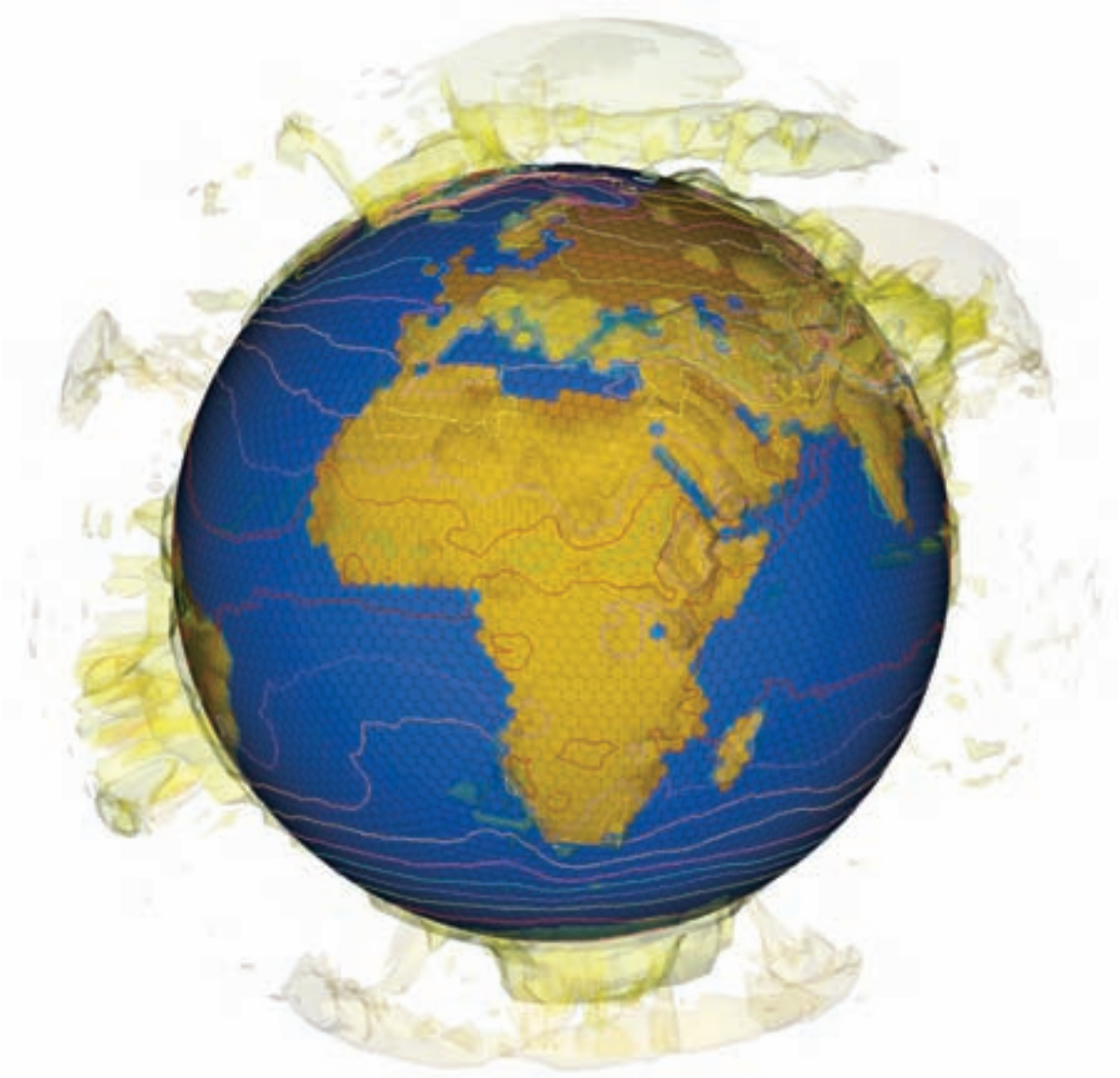


Figure 25. The large data sets generated by the GCRM require new analysis and visualization capabilities. This 3D plot of vorticity isosurfaces was developed using VisIt, a 3D visualization tool with a parallel distributed architecture, which is being extended to support the geodesic grid used by the GCRM. (Image: Prabhat, NERSC Analytics Team)

[14] Akio Arakawa and Celal S. Konor, "Unification of the anelastic and quasi-hydrostatic systems of equations," Monthly Weather Review **137**, 710 (2009).

State University (CSU) developed the Unified System of governing equations (so called because it unifies the quasi-hydrostatic compressible system with the nonhydrostatic anelastic system).[14] The Unified System can cover a wide range of horizontal scales, from turbulence to planetary waves. It also filters out vertically propagating sound waves of all scales, without excluding relevant waves such as inertia-gravity waves, Lamb waves, and Rossby waves.

"The atmosphere can make lots of different kinds of waves," Randall says, "but in choosing equations we knew we wanted to avoid those that include sound waves, because sound waves are completely irrelevant to weather and climate. Because sound moves too fast, if you include sound waves in your model, you have to take very small time steps. If you eliminate sound waves completely, then you can take much longer time steps. There have been other ways to get rid of them in the past, but they've been considerably less accurate. The new method that we've developed does involve approximations, because you're leaving something out, but it has much smaller errors that are, we believe, quite acceptable."

Another key feature of the Unified System is that it solves the three-dimensional vector vorticity equation rather than the vector momentum equation. Vorticity, or spinning motion, "is really at the core of much of the important fluid dynamics in the atmosphere," Randall says. "Vortices move around and maintain their identities and live a life, like little animals. Sometimes two vortices will merge and make a bigger one. Almost everything that is interesting and important in the motion of the atmosphere predominantly involves the spinning part."

Most climate models use the momentum equation because it is easier to solve than the vorticity equation, and vorticity can be derived from momentum. But Akio Arakawa of UCLA and Joon-Hee Jung of CSU found a more efficient way of solving the vorticity equation that represents the important spinning motions much more directly and explicitly than the momentum equation does. "You really have to get that spinning part right, because that's where most of the action is," Randall explains. "Working with the vorticity equation directly means focusing in on the part of the physics that is most important to what we care about." Figure 25 presents a snapshot of atmospheric vorticity simulated by the GCRM.

The component algorithms in the GCRM were selected for their good scaling properties, so the model scales linearly with the number of grid cells used. "Depending on the details of the configuration, we can do a few simulated days per wall clock day on 40,000 processors of Franklin," Randall says. "Which means that doing a whole year is a very big calculation — it might be like a hundred days continuously around the clock on 40,000 processors or more — a big chunk of a very expensive machine. So what we're doing is just barely doable now."

"But in ten more years," he adds, "we expect computers to be a hundred times faster, whether it's Green Flash [see page 91] or some other system. Then we'll be getting, say, a simulated year for a wall-clock day. That's a big improvement. You can start thinking about doing simulated decades or even longer. You're almost getting into the climate regime of about a century. So that's exciting.

"This project could not have happened without a lot of support from the federal government, especially the Department of Energy. We have to use the very fastest, most powerful machines in the world, and DOE, of course, is where you go for that. They're 'Supercomputing Central.' We've been computing at NERSC for more than a decade, and it's been an excellent experience. We have a lot of respect for and gratitude to everyone at NERSC for all the excellent support they have given us over the years."

Further computational challenges

The development of a geodesic dynamical core with a unique system of equations was the major, but not the only computational challenge. Other challenges include parallel input/output (I/O), including storage, management, and distribution of the voluminous output, and visualization of the results. The SciDAC Scientific Application Partnership titled "Community Access to Global Cloud Resolving Model and Data," led by Karen Schuchardt of Pacific Northwest National Laboratory, has been working to address those issues (see sidebar).

As for Randall's group, they are

now adding parameterizations of various physical processes, such as cloud microphysics, to the dynamical core of the GCRM, and they are also working on a method to include topography in the model, which will add vertically propagating waves produced by air flow over mountains. While continuing to run various tests on Franklin at NERSC, including numerical accuracy, stability, and parallel scaling performance, they are also running larger tests on up to 80,000 cores of Jaguar, a Cray XT system at the Oak Ridge Leadership Computing Facility (OLCF).

Early tests of the model will span just a few simulated days and will focus on short-range global weather prediction, starting from high-resolution analyses produced by working weather prediction centers. Tropical cyclones and other extreme weather events will be particular areas of focus. By 2011, the researchers plan to use the GCRM to perform two or more annual-cycle simulations, at least one of which will be coupled to the geodesic ocean general circulation model that they developed under SciDAC Phase 1.

Within the next ten years or so, models similar to the GCRM will be used for operational weather prediction, and eventually GCRMs will be used for multi-century climate simulations. The Green Flash project (page 91) may make this possible sooner rather than later. The long-term target resolution for a Green Flash system is a horizontal grid spacing of about 1 km, which will require approximately 671 million grid columns, each with about 100 layers.

Managing and Analyzing Petabytes of Data

The high spatial and temporal resolution of GCRM simulations will result in volumes of data output, expected to be on the order of 1 terabyte per hourly snapshot or 8.6 petabytes per year of continuous simulation time. It would be impractical to run the model repeatedly just to save the output required for a particular analysis. Therefore, it is necessary to store model results throughout the computation and provide flexible tools to extract subsets of the data required for a wide range of analyses. Developing those tools is the goal of the SciDAC "Community Access to Global Cloud Resolving Model and Data" Scientific Application Partnership (SAP), led by Karen Schuchardt of Pacific Northwest National Laboratory.

"We cannot easily generate this data every time someone needs it, so we view each dataset as an extremely valuable resource and want to make it available to as many collaborators as possible," Schuchardt says.

Project: Development and Test of an I/O API for the Global Cloud Resolving Model

PI: Karen Schuchardt, Pacific Northwest National Laboratory

Senior investigators: Bruce Palmer, Annette Koontz, and Todd Elsethagen, PNNL

Funding: ASCR, SciDAC

Computing resources: NERSC, EMSL

The main tasks of this partnership include developing a web portal that enables users to browse, search, and make specific data subset requests; developing tools to efficiently access, analyze, and visualize subsets of data; and developing a high performance input/output (I/O) application program interface (API) and data format definition. As illustrated in Figure 26, paradigm changing models such as a GCRM require coupled compute, storage, and analysis resources. The software services that provide data access to the broad community are a vital link in the flow of information.

These goals are a perfect match for NERSC's Science Gateways project [see page 79], which is developing custom web interfaces for computing, data distribution, collaboration, and analytics. NERSC's Outreach, Software, and Programming Group has collaborated with Schuchardt's team on portal development. In addition, the NERSC Analytics Group and the SciDAC Visualization and Analytics Center for Enabling Technologies (VACET) are assisting with troubleshooting and improving I/O and with evaluating and developing visualization tools for GCRM data.

When a high-resolution model like

the GCRM is running on 40,000 processors of a fast computer like Franklin, outputting the high volume of data and writing it to disk can become a bottleneck, slowing down the entire computation. The I/O API being developed for GCRM allows the data to be efficiently output in parallel streams to local storage on Franklin, in a data format (netCDF) that is common in the climate modeling community.

So far the researchers have achieved an effective aggregate I/O bandwidth of 5 gigabytes per second for writing GCRM output on Franklin. The increased bandwidth was achieved by consolidating I/O on an optimal number of processors, aggregating writes into large chunks of data, and making additional improvements in the filesystem and parallel I/O libraries. From Franklin, the data is copied to the NERSC Global Filesystem (NGF), where it can be shared with or transferred to scientists around the world.

The large datasets generated by the GCRM also require new analysis and visualization capabilities, including parallel processing and rendering. The NERSC Analytics/VACET team has developed a GCRM plug-in for the VisIt visualization tool that supports the geodesic grid used by the GCRM (Figure 25). Rather than transferring huge datasets between sites, scientists can choose to keep their simulation output at NERSC and use VisIt's client/server architecture to do remote visualization.

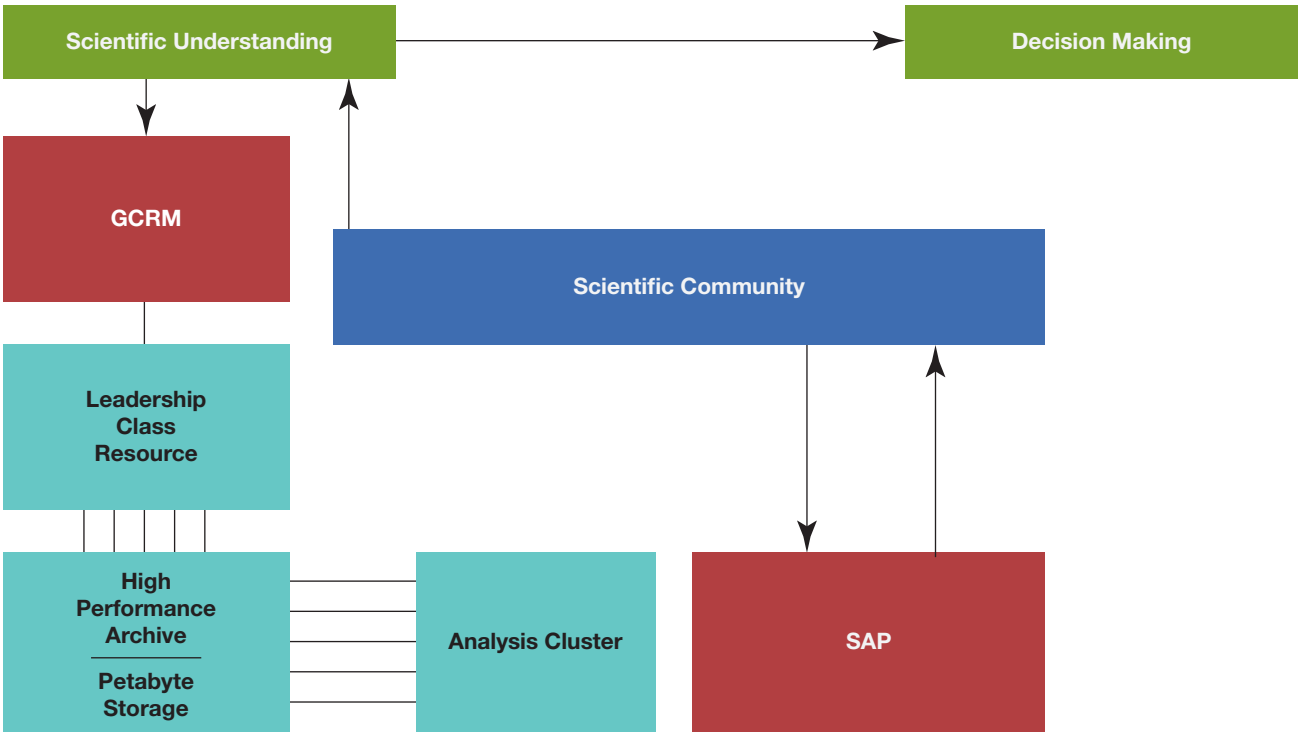
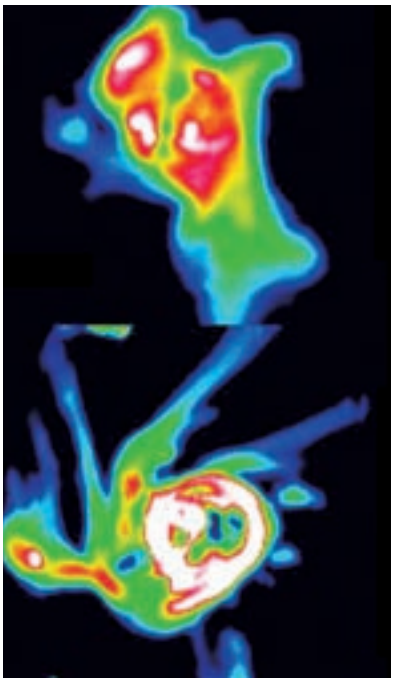


Figure 26. The Global Cloud Resolving Model requires high-performance, highly coupled compute, storage, and analysis resources, which are accessed by the climate science community through tools and services provided by the Scientific Application Partnership.

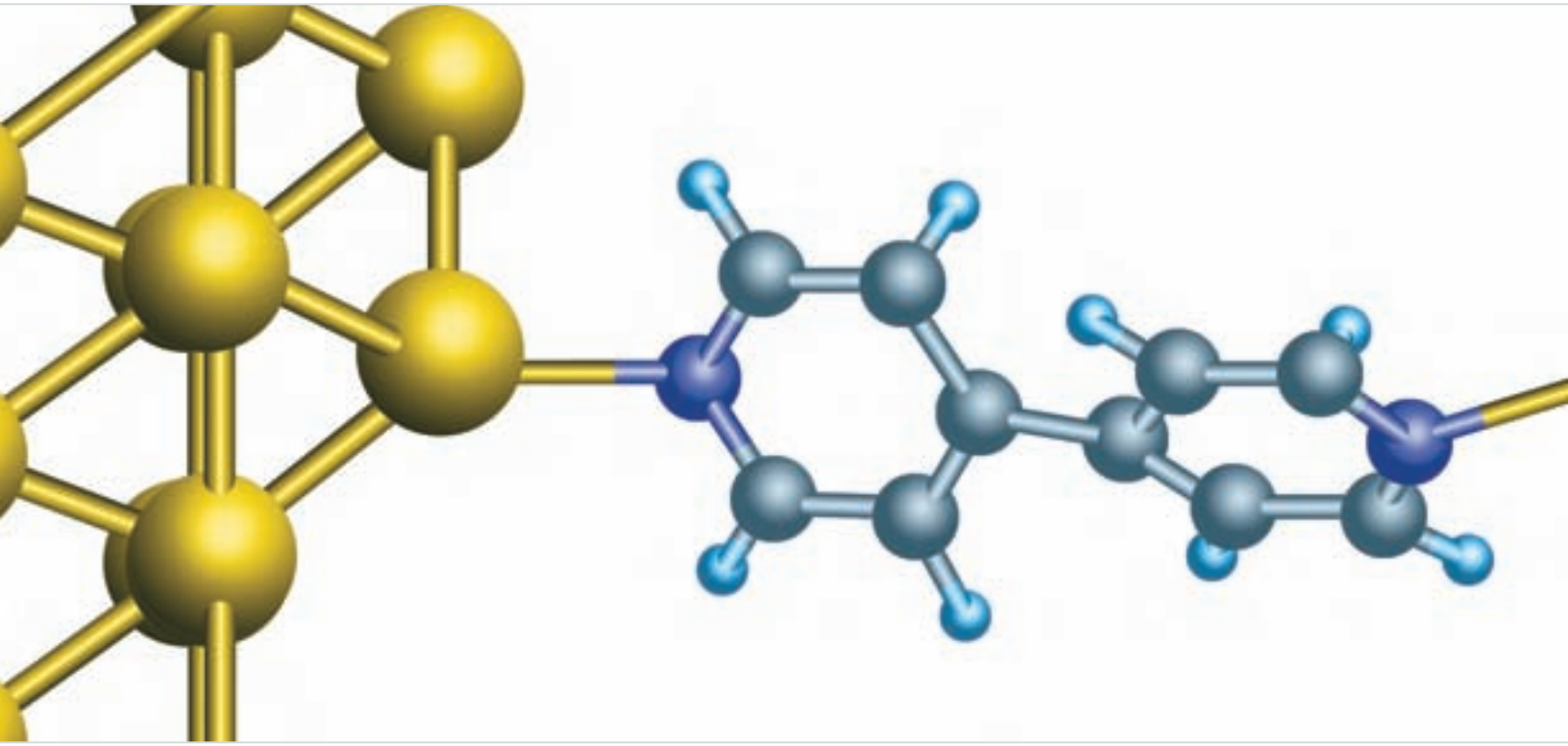
Research News: From Nanoelectronics to Galaxy Formation



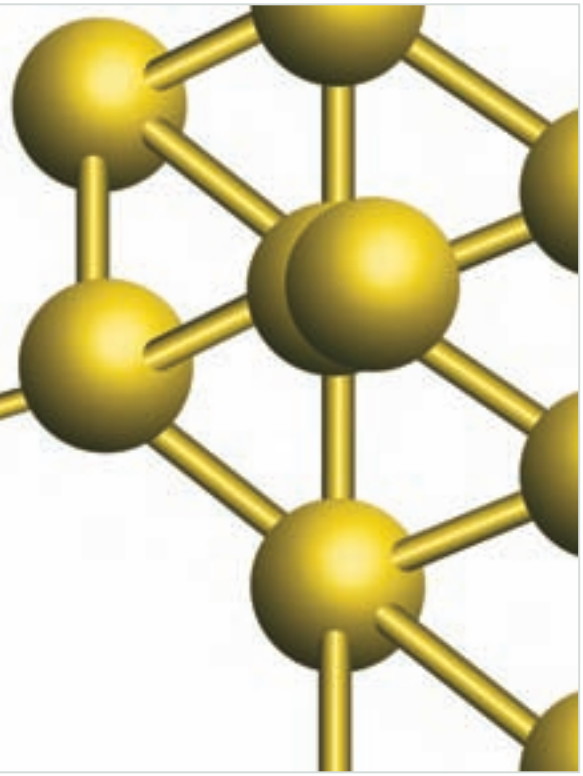
With 3,731 users in 2009 from universities, national laboratories, and industry, NERSC supports the largest and most diverse research community of any computing facility within the DOE complex. Home to some of the world's most efficiently managed supercomputers, NERSC is a leader in providing systems, services, and expertise to advance computational science.

In their annual allocation renewals, users reported 1,464 refereed publications (accepted or submitted) enabled at least in part by NERSC resources in 2008, and 1,646 publications in 2009. These scientific accomplishments are too numerous to cover in this report, but several representative examples are discussed below, ranging from nanoscience and nuclear physics to accelerator physics, astrophysics, and even the humanities. These discoveries are improving our understanding of the universe, from the smallest to the largest scales, and even broadening our understanding of our own culture and history.

Push-Pull



DFT calculations help researchers discover a potential on-off switch for nanoelectronics



Researchers at the Lawrence Berkeley National Laboratory's (Berkeley Lab) Molecular Foundry and Columbia University found that electrical resistance through a molecular junction — a nanometer-scale circuit element consisting of a single molecule contacted with gold wires — can be turned on and off by simply pushing and pulling the junction. Experts believe that this newly demonstrated molecular-scale control could be leveraged for future nanoscale electronic devices.

The switching phenomenon was initially discovered in experiments conducted by a team of researchers led by Latha Venkataraman of Columbia University. But the underlying physics would only later be understood when a group of theorists led by Jeffrey Neaton, Facility Director of the Theory of Nanostructured Materials Facility at the Molecular Foundry, teamed up with the Columbia researchers to develop a theory to describe the conductance of individual molecules trapped between gold electrodes. This work was done with the help of density functional theory (DFT) calculations on NERSC's Cray XT4 system, Franklin.

"If we wish to ultimately design circuit elements at the molecular scale, we need to understand how the intrinsic properties of a molecule or junction are actually connected to its measured resistance," said Neaton. "Knowing where each and every atom is in a single-molecule junction is simply beyond what's possible with experiments at this stage. For these sub-nanometer scale junctions — just a handful of atoms — theory can be valuable in helping interpret and understand resistance measurements."

Project: Theory of Nanostructured Materials

PI: Jeffrey Neaton, Lawrence Berkeley National Laboratory

Senior investigator: David Prendergast, LBNL

Funding: BES, NSF, DARPA, NYSTAR

Computing resources: NERSC, LBNL

Experiments meet theory

In traditional electronic devices, charge-carrying electrons diffuse through circuits in a well understood fashion, gaining or losing energy through transactions with impurities or other particles they encounter. Electrons at the nanoscale, however, can travel by a mechanism called quantum tunneling in which, due to the small length scales involved, it becomes possible for a particle to disappear through an energy barrier and suddenly appear on the other side, without expending energy. Tracking such tunneling of electrons through individual molecules in nanoscale devices has proven difficult.

"You can't use a microscope to see that a molecule is trapped, you can only sense it indirectly through electrical conductivity data, for example. For more than a decade, researchers have been 'wiring up' individual molecules and measuring their electrical conductance," said Neaton.

He notes that routine formation of high quality electrical contacts, or "alligator clips," between nanostructures and electrical leads is extremely challenging. This makes experiments difficult to interpret; as a result, the reported conductance values-in experiment and theory often varied by an order of magnitude or more. The time was ripe for a quantitative comparison between theory and an experiment with well defined contacts, a collaboration made possible by the Molecular Foundry User Program.

One tool experimental researchers use to probe changes in nanostructure currents is called a scanning tunneling microscope (STM), which has a conductive gold tip. Previous work had shown a gold STM tip could repeatedly be plunged into a gold surface containing a solution of molecules and retracted, until the contact area between the tip and the gold surface reduces to a single strand, like a necklace. When this strand finally breaks, nearby molecules can hop into the gap between strands and contact the gold electrodes, resulting in a sudden change in conductance. Using this technique, Venkataraman and her colleagues discovered that the conductance of molecules terminated by amines (related to ammonia) in contact with gold electrodes could be reliably measured.

Using a new theoretical approach, Neaton and his collaborators also began to study the conductance of a junction between gold electrodes and bipyridine — an aromatic molecule with two benzene-like rings, each containing one nitrogen atom. They hypothesized about the conductance of junctions arranged vertically between gold molecules and sandwiched at angles, working closely with the laboratory researchers to compare their computer-generated predictions with experiments.

The team used computational methods based on parameter-free first-principles calculations to test these hypotheses. First-principles methods are atomic-scale computational approaches with the ability to predict measurable properties of materials with good accuracy from scratch, i.e., through solution of the quantum mechanics of a system of interacting electrons in a field of nuclei. Using a new approach involving a self-energy correction, the researchers evaluated the electrical resistance of bipyridine-gold molecular junctions in different conformations (Figure 1).

A self-energy correction

First-principles transport calculations were based on density functional theory (DFT) within the generalized gradient approximation. The SCARLET code was used to calculate the electron transmission for many-junction geometries. The alignment of the frontier molecular energy levels in the junction relative to Fermi energy can show significant errors in DFT, resulting in calculated conductance values that are too large. A self-energy correction was calculated and added to the molecular orbital energies in the junction to account for many-electron effects. As the DFT orbital energy is close to the electrode Fermi energy in this system, the self-energy correction was quantitatively quite important to the predicted conductance value.

"Computing properties of bipyridine-gold molecular junctions related to its electrical conductance for different conformations using first-principles approaches is computationally demanding," said Neaton. "Each iteration requires solution of several hundred linear systems of equations, one at each

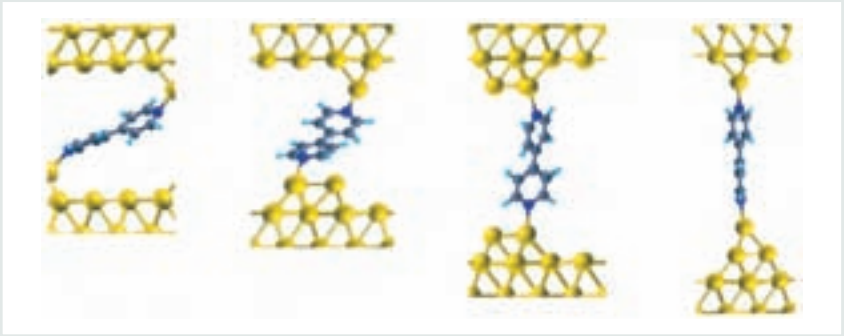


Figure 1. Examples of calculated junction geometries relaxed at different tip–sample distances.

energy point on a dense grid, each linear system typically consisting of 1,000 to 10,000 unknowns. Thus on average, transport calculations tend to be more than 100 times slower than conventional DFT calculations. The high throughput and magnitude of the NERSC resources facilitated a highly interactive back-and-forth with experimentalists, allowing us to rapidly modify our calculations and compare with experimental results as they became available."

A mechanical switch

After the researchers had plugged away for more than a year, the story that emerged was surprisingly detailed: if bipyridine bonded at an angle, more current could flow through the junction compared with the bipyridine bonded in a vertical configuration. This suggests the conductance of bipyridine was linked to the molecule's orientation in the junction. In experiments, the scientists noticed that when the final strand of gold atoms breaks and snaps back, the vertical gap is not big enough for bipyridine to insert itself in line, so instead it bonds at an angle. As the gap increases, the molecule jumps to a vertical configuration, causing the conductance to plummet abruptly. Eventually, the molecule straightens even more, and the contact breaks.

"Once we determined this, we wondered, Could you reverse this behavior?" said Su Ying Quek, a postdoctoral researcher at the Molecular Foundry who was involved in the theory development and was first author of the paper reporting the results.[1]

In experiments, Quek and Neaton were able to demonstrate why pushing the junction to an angle and pulling it straight could repeatedly alter the conductance, creating a mechanical switch with well defined on and off states (Figure 2).

"One of the fascinating things about this experiment is the degree to which it is possible to control the 'alligator clips,'" said Neaton. "For this particular molecule, bipyridine, experiments can reproducibly alter the atomic-scale interface between the molecule and its gold contacts back and forth to switch the conductance of the junction."

Quek and Neaton hope to refine and apply their theoretical framework to promising systems for solar energy conversion and multi-electron heterogeneous catalysis, where controlling charge dynamics at nanoscale interfaces is central.

"Understanding how electrons move through single-molecule junctions is the first step," said Neaton. "Organic-inorganic interfaces are everywhere in nanoscience, and developing a better picture of charge transport in hybrid materials systems will certainly lead to the discovery of new and improved electronic devices."

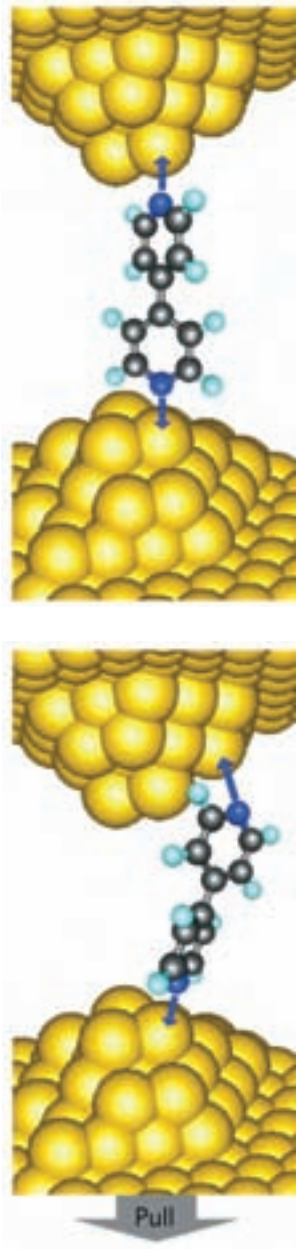
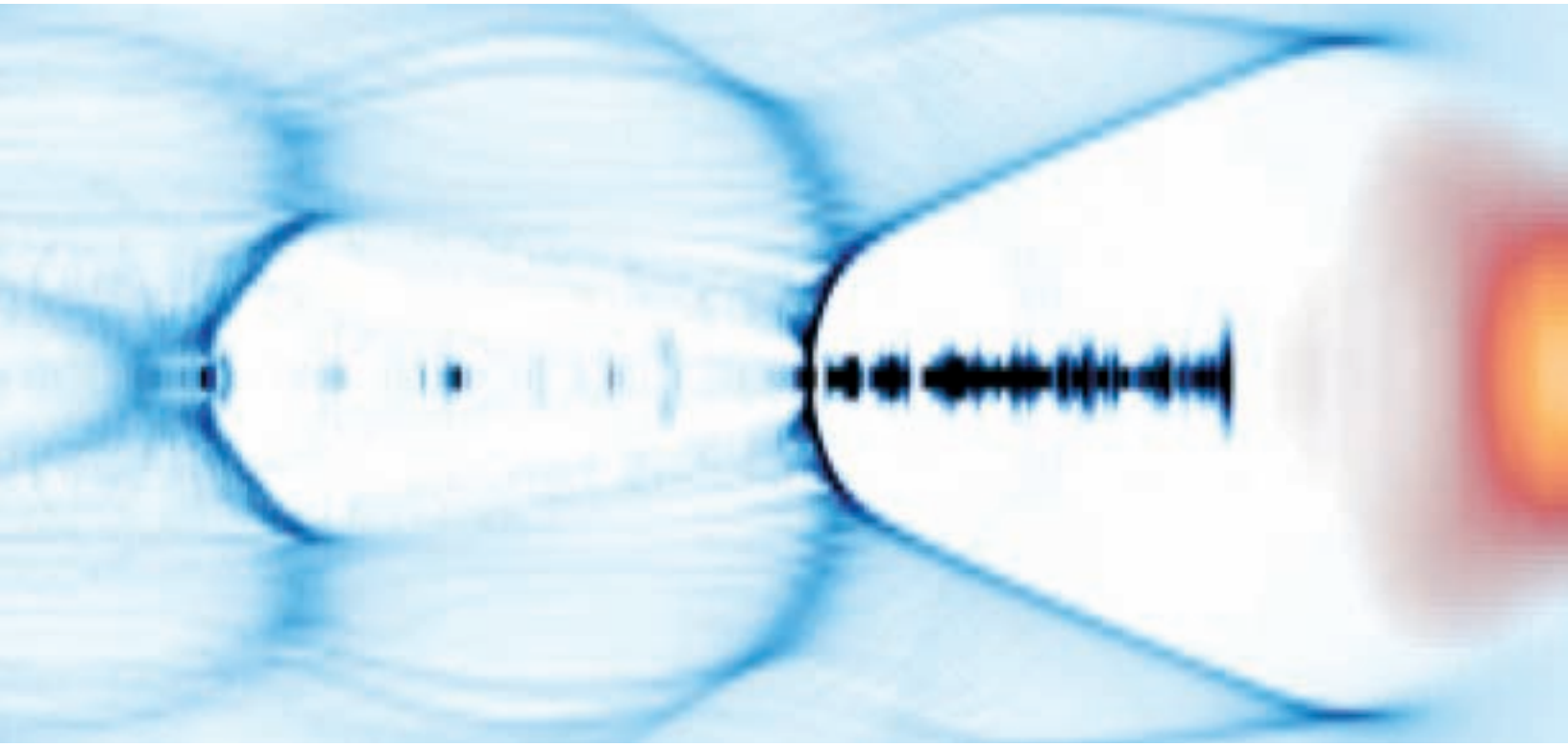


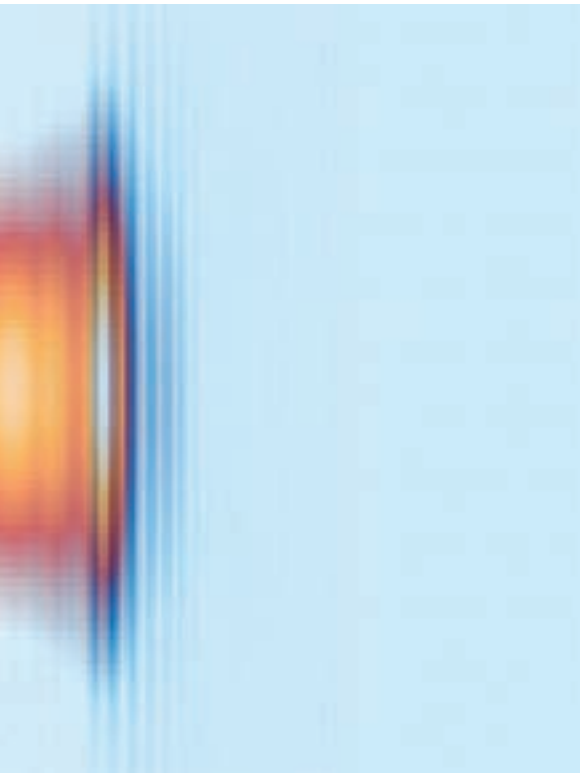
Figure 2. These schematics illustrate the vertical and angled molecular junction configurations for mechanically induced switching. A study has revealed that electrical resistance through such a junction can be turned on and off simply by pushing (top) so that the configuration is vertical or by pulling the junction (bottom) so that the configuation is angled.

[1] Su Ying Quek, Maria Kamenetska, Michael L. Steigerwald, Hyoung Joon Choi, Steven G. Louie, Mark S. Hybertsen, J.B. Neaton, and L. Venkataraman, "Mechanically controlled binary conductance switching of a single-molecule junction," Nature Nanotechnology **4**, 230 (2009).

Reframing Accelerator Simulations



Taking advantage of special relativity, a new computational method results in up to 300-fold speedup of plasma-based accelerator simulations



Project: Petascale Particle-in-Cell Simulations of Plasma Based Accelerators

PI: Warren B. Mori, University of California, Los Angeles

Senior investigators: Frank Tsung and Cheng Kun Huang, University of California, Los Angeles; Thomas Katsouleas, Duke University

Funding: HEP, INCITE, NSF, FCG, FCT, LLNL

Computing resources: NERSC, OLCF, UCLA, IST

For the past 80 years, the tool of choice in experimental high energy physics has been particle accelerators; but as they grew in size and sophistication, accelerators have also grown in cost. The construction cost alone for the Large Hadron Collider (LHC) at CERN, which came online in 2008, was nearly 10 billion dollars. If the same technology is used, the world's next "atom smasher" will cost several times as much. The long-term future of experimental high energy physics research using accelerators depends on the successful development of novel ultra-high-gradient acceleration methods.

A newer technique called *plasma-based acceleration* — in which electrons or positrons gain energy by surfing on a wave generated by a laser pulse or particle beam in an ionized gas, or plasma — could potentially slash the size and expense of high-energy accelerators. Such laser-wakefield or plasma wakefield accelerators could be used for research in particle physics, materials science, and structural biology, as well as for practical applications such as cancer treatment. So far, in small laboratory experiments, plasma-based acceleration has exhibited velocity gradients and focusing forces more than 1000 times greater than conventional technology. Computational simulations have played a major role in understanding and advancing this new technology.

In the past several years, computational simulation tools for plasma-based acceleration have been verified against each other, against experimental results, and against theory. A project called "Petascale Particle-in-Cell Simulations of Plasma Based Accelerators," led by physicist Warren Mori of the University of California, Los Angeles (UCLA), is using these tools not just to understand current experiments, but especially to simulate experimental regimes that will not be accessible for years to come. With DOE INCITE grants of 7.6 million processor hours on NERSC computers from 2007 to 2009, Mori's team, in collaboration with researchers from Instituto Superior Tecnico (IST) in Lisbon, Portugal, has developed the capability to test key plasma accelerator concepts computationally before tens to hundreds of millions of dollars are spent on experimental facilities.

The particle-in-cell (PIC) method, in which individual particles are tracked in continuous phase space, plays an important role in laser wakefield simulations because the algorithm includes all the essential physics. But when large space and time scales are being studied, PIC calculations demand huge computational resources. Mori and his collaborators have developed and use

the three-dimensional PIC code, OSIRIS, which is over 80% efficient running on more than 4000 nodes of NERSC's Cray XT system, Franklin. But with the longer length of next-generation petawatt laser experiments, conventional PIC modeling is simply impossible on today's supercomputers.

One solution is to simplify the algorithm using approximations, which produces acceptable results in some but not all scenarios. But another solution that does not require approximations is to minimize the time and space scales to be resolved in the simulation by using an *optimal Lorentz frame* — a method that takes advantage of the properties of special relativity and the fact that a short-pulse laser is not reflected.

A moving frame

The *Lorentz transformation*, named after the Dutch physicist Hendrik Lorentz, is a set of equations that describes how, according to the theory of special relativity, two observers' varying measurements of space and time can be converted into each other's frames of reference. It reflects the fact that observers moving at very different velocities report different distances, passage of time, and in some cases even different orderings of events (with no causal links), while the physics equations do not change between frames.

Recently physicist Jean-Luc Vay of Lawrence Berkeley National Laboratory suggested that the Lorentz transformation could be used to drastically reduce computer run times for the simulation of relativistic interactions like those found in laser wakefield accelerators.[2,3] "By describing the space and time coordinates from the point of view of a moving observer, rather than a stationary observer, relativistic effects set in and there is, for each object, depending on its velocity in the laboratory, a contraction or dilation of space and time," Vay says. Within an infinite number of possible moving frames will be an optimal one that minimizes the range of space and time scales in which the objects interact, reducing the computations required to describe the system.

Mori's research team quickly saw the advantage of this approach and together with the IST team applied it to laser wakefield simulations using the tools they had developed. Mori notes, "We had tried this idea with limited success more than 10 years ago but were not able to overcome numerical issues. After reading Jean-Luc's paper, we realized that we needed to try it again."

In a benchmark test, they compared computational results in different frames with data from a current state-of-the-art experiment. The simulations with the optimal Lorentz frame ran 20 times faster, with no loss of relevant information.[4] Simulations of a next-generation petawatt laser wakefield accelerator ran up to 300 times faster with the optimal frame.[5]

Figure 3 gives a schematic depiction of how the frame transformation works. On the left (a) is the laboratory frame—the viewpoint of the observing scientist —from which the plasma appears stationary as the laser pulse approaches. In this configuration, the laser wavelength is the smallest spatial structure to be resolved, so it defines the longitudinal grid resolution. However, this fine grid over-resolves the plasma structures, increasing the computational cost without adding any significant information.

Now imagine shifting to the viewpoint of an observer traveling in the direction of the laser, toward the plasma, at nearly the speed of light. This is the relativistic Lorentz frame (Figure 3b) — also called a *boosted frame*, meaning a directional Lorentz transformation with no rotation. In such a frame, the laser pulse wavelength increases, while the plasma contracts and moves toward the laser pulse. This frame reduces the number of grid cells, because the plasma length is shorter, and reduces the number of time steps, since the

[2] J.-L. Vay, "Noninvariance of space- and time-scale ranges under a Lorentz transformation and the implications for the study of relativistic interactions," Phys. Rev. Lett. **98**, 130405 (2007).
[3] J.-L. Vay et al., "Simulating relativistic beam and plasma systems using an optimal boosted frame," J. Phys. Conf. Ser. **180**, 012006 (2009).
[4] S. F. Martins, R. A. Fonseca, L. O. Silva, W. Lu, and W. B. Mori, "Numerical simulations of laser wakefield accelerators in optimal Lorentz frames," Computer Physics Communications, in press.
[5] S. F. Martins et al., "Exploring LWFA regimes for near term lasers using particle-in-cell simulations in Lorentz boosted frames," Nature Physics, in press.

crossing time between the laser and plasma is also smaller. Running the PIC calculations in this frame greatly reduces the number of iterations necessary to produce meaningful scientific results.

A breakthrough in simulation scale

Multiple-frame simulations of three parameters sets for a next-generation petawatt laser wakefield accelerator5 included the first fully kinetic simulations for a meter scale device — in fact, the standard laboratory frame only allowed a plasma length of less than 1 centimeter. The goal of these simulations, performed using the OSIRIS code, was to explore options for obtaining the desired electron beam energy gain and output beam charge.

The first option is to compress all of this laser energy into a very short pulse laser (duration < 30 fs), and then choose a plasma density (10^{19} cm^{-3}) so that the laser duration is approximately a plasma period. This configuration was simulated in the laboratory frame with a 87 × 82 × 82 μm^3 computational window resolved with 3000 × 256 × 256 cells. A total of 8.4×10^8 particles were pushed for 10^5 iterations, corresponding to 0.25 cm of pre-formed plasma. Results show that injection is very strong (>14 nanocoulombs, nC) and that a quasi monoenergetic beam is obtained (~3 GeV), so the output is typically a low energy output beam with maximum charge.

By lowering the laser intensity and plasma density in the second option, the laser spot size and length can be properly matched, and a more controlled blowout of the electrons is achieved.[6] In this regime, self-injection is not so strong, and higher quality beams with smaller charge are obtained. This regime was modeled in a boosted frame with relativistic factor (gamma) of 10, and with a moving window 3.1 × 0.5 × 0.5 mm^3 divided into 7000 × 250 × 250 cells. A total of 3.5×10^9 particles were pushed for 3×10^4 iterations, corresponding to 2.17 cm of pre-formed plasma in the boosted frame (21.7 cm in the laboratory). These conditions lead to the acceleration of 1–2 nC beams above 11 GeV, in agreement with an earlier theoretical prediction.

A third regime is obtained by further increasing the spot and duration of the laser pulse. The matched conditions require a propagation of over 5 meters, but the laser intensity is no longer enough to allow for self-guiding nor self-injection. This implies that a plasma channel is required to guide the laser, and that an electron beam must be externally injected into the wake. This simulation was also performed in a boosted frame with gamma of 10, with a moving window of size 6.5 × 1.4 × 1.4 mm^3 divided into 8000 × 128 × 128 cells. A total of 5.2×10^8 particles were pushed for 2.1×10^5 steps, corresponding to 5.28 m of preformed plasma in the laboratory frame. Results confirmed the possibility of accelerating ~1 nC of charge up to 40 GeV. These simulations provide guidance for future laser wakefield experiments.

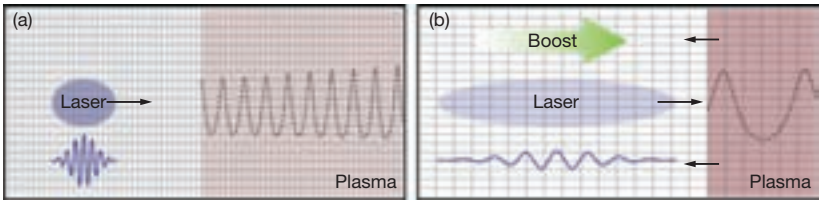


Figure 3. Illustration of a numerical grid for a laser wakefield accelerator simulation in (a) laboratory frame and (b) relativistic Lorentz frame (boosted frame). When going from the plasma rest frame to a frame moving in the direction of the laser, the pulse length increases and the intensity decreases; in contrast, the plasma contracts, becomes more dense, and propagates to the left at the boost speed. Therefore, the same number of points per laser wavelength may be used to define the grid, which now correspond to a smaller number of points in the plasma. The plasma wake (dashed line) also stretches similarly to the laser pulse, since it moves relativistically in the boost direction. *(Image from ref. 4)*

[6] W. Lu et al., "Generating multi-GeV electron bunches using single stage laser wakefield acceleration in a 3D nonlinear regime," Phys. Rev. ST AB **10**, 061301 (2007).

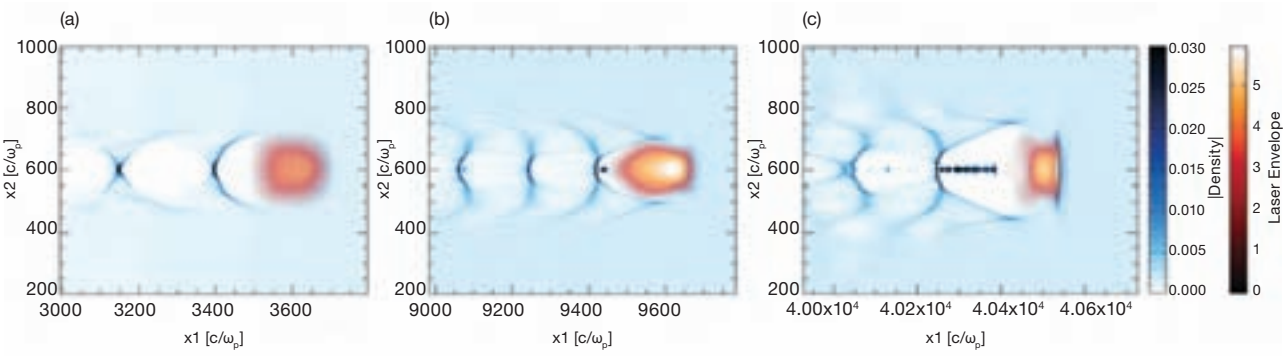


Figure 4. Snapshots of the density profiles (blue color table) from the simulation; the laser pulse (red color table) is superimposed on the density map. (a) At $z = 0.5$ mm, the laser pulse has completely blown out the electrons, forming a nearly spherical wakefield. (b) As the laser pulse evolves, a stable accelerating structure is formed and self-injection occurs ($z \sim 3$ mm). (c) By the end of the plasma ($z = 8.5$ mm), the highest energy electrons have begun to dephase, producing a 700 MeV electron beam. *(Image from ref. 7)*

Modeling an innovative experiment

OSIRIS simulations were also performed to understand the underlying physics of an experiment done as a collaboration between UCLA and Lawrence Livermore National Laboratory (LLNL).[7] This experiment, performed at LLNL's Jupiter Laser Facility, was the first GeV-class laser wakefield acceleration experiment where a self-injection threshold was demonstrated for densities below 5×10^{18} cm^{-3} in a gas jet without a guiding structure. In order to probe wider ranges of laser and plasma parameters, a large set of simulations was performed in a boosted frame, which enabled ultra-fast turnaround times.

The simulations used a $100 \times 150 \times 150$ μm computational window corresponding to $4000 \times 256 \times 256$ grid points with a resolution along the laser propagation direction of 25.4 and 595 nm in the transverse plane, with four particles per cell. Figure 4 shows PIC simulation results where the laser beam completely blows out the electrons, producing a stable accelerating structure. These simulations show that the laser pulse is nearly matched to the plasma conditions in that the laser pulse is completely contained within the ion bubble and continues to be guided over the entire length of the plasma. While a nearly constant plasma density (3×10^{18} cm^{-3}) was maintained, the electron energy gain increased from 75 to 720 MeV when the plasma length was increased from 3 to 8 mm.

Electron injection and trapping

In addition to the technology developments that will enable laser-plasma acceleration of higher charge beams to higher energies, significant effort is currently being directed towards new and improved methods for particle injection into these accelerators. The objective is to improve the reliability and efficiency of the accelerators and enable better charge, divergence, and energy of the accelerated electrons.

It was recently shown that in a partially ionized plasma, the electric field of the wake driven by an electron beam could become large enough to tunnel ionize and inject electrons directly into the wakefield. Mori's research team has extended this idea to laser-driven

[7] A. Pak, K. A. Marsh, S. F. Martins, W. Lu, W. B. Mori, and C. Joshi, "Injection and trapping of tunnel-ionized electrons into laser-produced wakes," Phys. Rev. Lett. **104**, 025003 (2010).
[8] D. H. Froula et al., "Measurements of the critical power for self-injection of electrons in a laser wakefield accelerator," Phys. Rev. Lett. **103**, 215006 (2009).

accelerators, where the field of the laser is used to control the injection process (as shown in Figure 5).[8]

They have explored these new configurations with full scale 3D simulations in OSIRIS. Results were then used to support design of experiments at UCLA, using a Ti-sapphire laser with energy <500 mJ and pulse length of ~45 fs. The laser is focused to a spot size of 6 μm onto the edge of a 2 mm wide and 1.9 mm long gas column created by a conical gas jet. The results obtained from the experiment and the one-to-one simulations are in good agreement, in particular the final energy spectra of the injected beams with 60–100 MeV, the relative amount of trapped charge (tens of pC), and the beam divergence of a few mrad. When compared with standard self-trapping regimes, the tunnel ionization injection requires lower laser intensity for trapping to occur, and can also constitute a useful mechanism for injection into lower density wakefields.

"The common goal of all these studies," Mori says, "is to understand the physics of plasma-based acceleration in order to determine if it could be used to reduce the size of the largest accelerators from tens of kilometers to ten meters in length. We have developed the right tools to do this — a three-dimensional full PIC code (OSIRIS) that includes the Lorentz frame capability, as well as a quasi-static PIC code (QuickPIC). The quasi-static algorithm can also provide a savings of 100–10,000 in computer time without loss of accuracy for certain problems. The tools and the experiences of our group put us in a unique position to make an impact in this field."

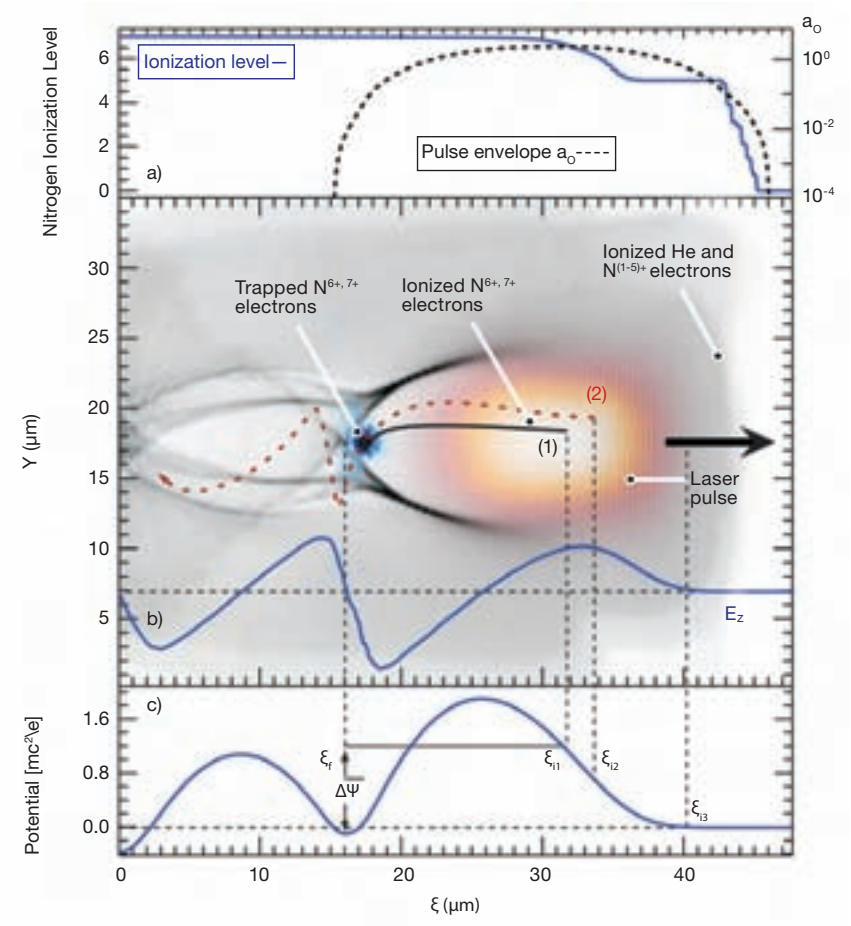
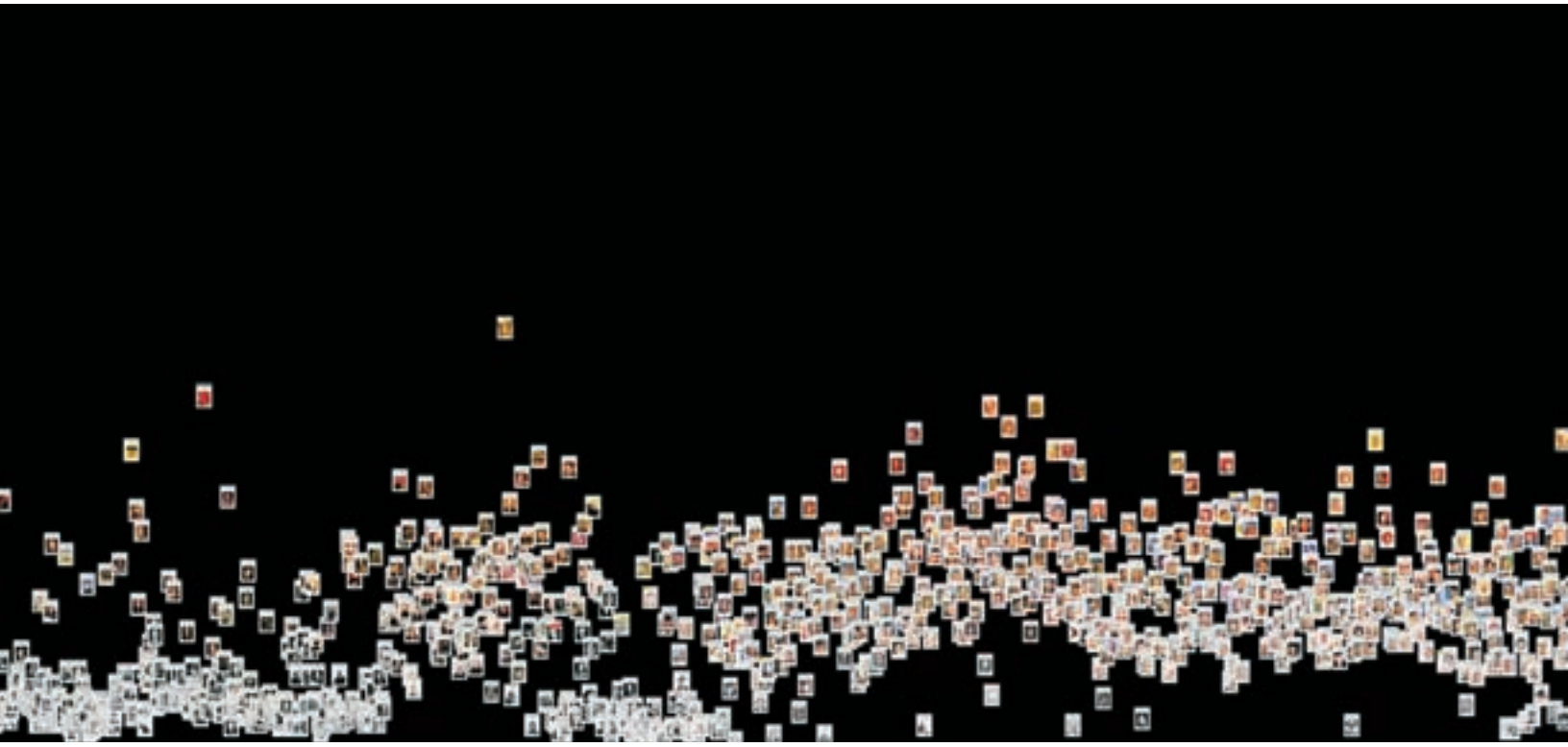


Figure 5. Results from the OSIRIS simulation of the injection of tunnel-ionized electrons. As the laser pulse propagates to the right, it ionizes a 9:1 mix of helium and nitrogen and drives a wake. (a) The envelope of the normalized vector potential a_o of the laser (dashed line) and the ionization state of nitrogen atoms (solid line) on axis. The superimposed trajectories (1) and (2) in frame (b) represent simulation electrons ionized into the wake from the *K* shell of nitrogen. Electron (1) is ionized close to the axis and is trapped by the wakefield, while electron (2), ionized earlier and off axis, slips over the potential well and is not trapped. (c) The normalized wake potential on axis, with particular points relevant to the physics of the trapping mechanism depicted. *(Image from ref. 8)*

Cultural Analytics



NERSC computers help to develop new paradigm for studying culture

Project: Visualizing Patterns in Databases of Cultural Images and Video

PI: Lev Manovich, University of California, San Diego

Senior investigator: Jeremy Douglass, UCSD

Funding: ASCR, National Endowment for the Humanities High Performance Computing Initiative

Computing resources: NERSC

Just as slide projectors revolutionized the study of art by allowing billions of people in classrooms all over the world to scrutinize sculptures and paintings outside of art galleries, researchers from the Software Studies Initiative at the University of California at San Diego (UCSD) believe that a new paradigm called *cultural analytics* will drastically change the study of culture by allowing people to quantify evolving trends across time and countries.

Inspired by scientists who have long used computers to transform simulations and experimental data into multi-dimensional models that can then be dissected and analyzed, cultural analytics applies similar techniques to cultural data. With an allocation on NERSC's supercomputers and help from the facility's Analytics Team, UCSD researchers recently illustrated changing trends in media and design across the 20th and 21st centuries via Time magazine covers and Google logos.

"The explosive growth of cultural content on the web, including social media together with digitization efforts by museums, libraries, and companies, make possible a fundamentally new paradigm for the study of both contemporary and historical cultures," says Lev Manovich, Director of the Software Studies Initiative at UCSD.

"The cultural analytics paradigm provides powerful tools for researchers to map subjective impressions of art into numerical descriptors based on intensities of color, textures and shapes, as well as the organization of images using classification techniques," says Daniela Ushizima of the NERSC Analytics Team, who contributed pattern recognition codes to the cultural analytics image-processing pipeline.

Manovich's research, called "Visualizing Patterns in Databases of Cultural Images and Video," is one of three projects participating in the Humanities High Performance Computing Program, an initiative that gives humanities researchers access to some of the world's most powerful supercomputers, typically reserved for cutting-edge scientific research. The program was established in 2008 as a unique collaboration between DOE and the National Endowment for the Humanities.

"For decades NERSC has provided high-end computing resources and expertise to over 3,000 science users annually. These resources have contributed to a number of science breakthroughs that have improved our understanding of nature. By opening up these computing resources

to humanities, we will also gain a better understanding of human culture and history," says NERSC Division Director Kathy Yelick.

Visualizing changes in Time covers and Google logos

As relatively cheap hardware and software empowers libraries, museums and universities to digitize historical collections of art, music, and literature, and as masses of people continue to create and publish their own movies, music, and artwork on the Internet, Manovich predicts that the biggest challenge facing cultural analytics will be securing enough computing resources to process, manage, and visualize this data at a high-enough resolution for analysis.

Manovich and his collaborators leveraged the expertise of NERSC's Analytics Team to help them to develop a pipeline for processing cultural images and to scale up their existing codes to run on NERSC's high-end computing systems. To test their pipeline, the team mapped out all 4,553 covers of Time magazine from 1923 to 2008 and all the Google logos that have been published on the search engine's homepage around the world from 1998 to 2009.

In the Time magazine visualization (Figure 6), the X axis represents time in years, from the beginning of the 20th century on the left to the early 21st century on the right. The Y axis measures the brightness and saturation hue of each cover, with the most colorful covers appearing toward the top.

"Visualizing the Time covers in this format reveals gradual historical changes in the design and content of the magazine," says Manovich. "For example, we see how color comes in over time, with black and white and color covers co-existing for a long period. Saturation and contrast of covers gradually increases throughout the 20th century — but surprisingly, this trend appears to stop toward the very end of the century, with designers using less color in the last decade."

He notes that there are also various changes in magazine content revealed by the visualization. "We see when women and people of color start to be featured, how the subjects diversify to include sports, culture, and topics besides politics, and so on. Since our high-resolution visualizations show the actual covers rather than using points or other graphical primitives typical of standard quantitative graphics, a single visualization reveals many trends at once. It is also accessible to a wider audience than statistical graphs," adds Manovich.

Using NERSC computers, Jeremy Douglass, a postdoctoral researcher with Software Studies, also applied cultural analytics techniques to hundreds of Google logos that have appeared on the search engine's homepage all over the world from 1998 to the present (Figure 7). Artists periodically reinterpret logos on the Google homepage to mark a cultural milestone, a holiday, or a special occasion. In the visualization, the X axis measures how much the various logos deviate from the original design. Images toward the left show very little modification, and those toward the right have been significantly modified.

Figure 6. 4553 covers of Time magazine, 1923–2008. The X axis is time; the Y axis is a composite dimension of brightness, hue, and saturation measures that were automatically extracted from the images. The graph shows the complicated transition from black and white to color printing and indicates some basic design trends in various eras of the publication. *Image: Jeremy Douglass and Sunsern Cheamanunkul*

Meanwhile, the images toward the bottom of the Y axis illustrate artistic changes that affect the bottom of the word "Google," and those toward the top show changes in the upper portion of the word.

"Google logos are relatively small, and there have so far only been less than 600 of them, so the act of rendering full-resolution maps is quite doable with a desktop workstation. However, we were interested in using data exploration to tackle ideas about visual composition and statistical concepts like centroid, skewness, kurtosis [distribution of data around the mean], and so forth," says Douglass.

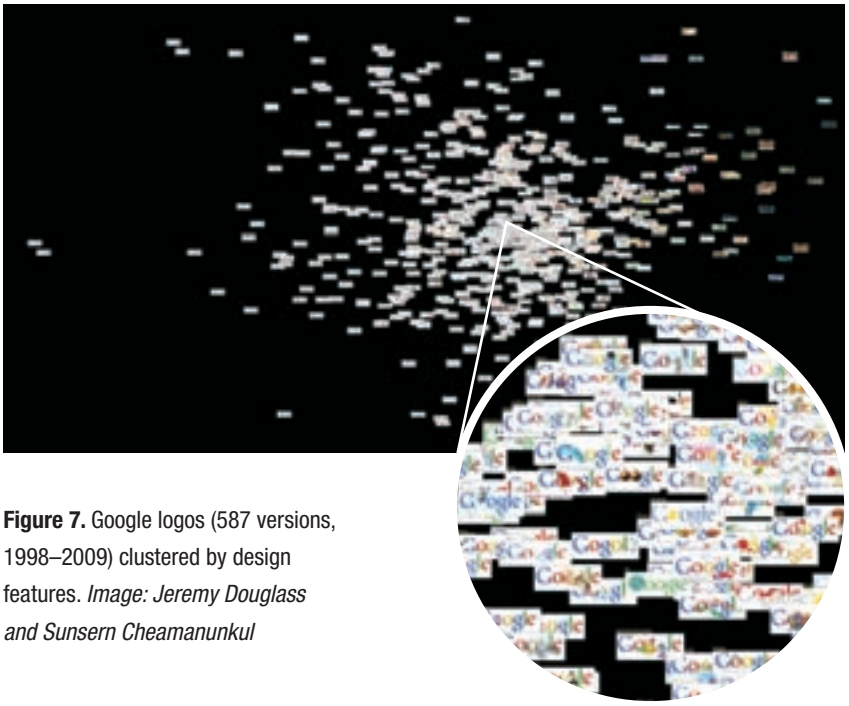


Figure 7. Google logos (587 versions, 1998–2009) clustered by design features. *Image: Jeremy Douglass and Sunsern Cheamanunkul*

A game changer

To explore how these and other concepts might contribute to mapping the "space of aesthetic variation," Douglass notes that the team needed to experiment. "That's where the ability to iterate with NERSC's Analytics Team becomes so important. When you want to repeatedly re-render 4,553 high-resolution images and be able to see how they evolve over time according to various features, 'How long will this take?' become a very big deal. And there's no such thing as too much power," he adds.

The cultural analytics team ultimately hopes to create tools that will allow digital media schools and universities to compare hundreds of thousands of videos and images in real time to facilitate live discussions with students. They also hope to create visualizations that will measure in many millions of pixels and show patterns across tens of millions of images. Although they are currently only beginning to develop software for cultural analytics, Manovich says the NERSC runs showed him that supercomputers are a game changer and will be vital to achieving this goal.

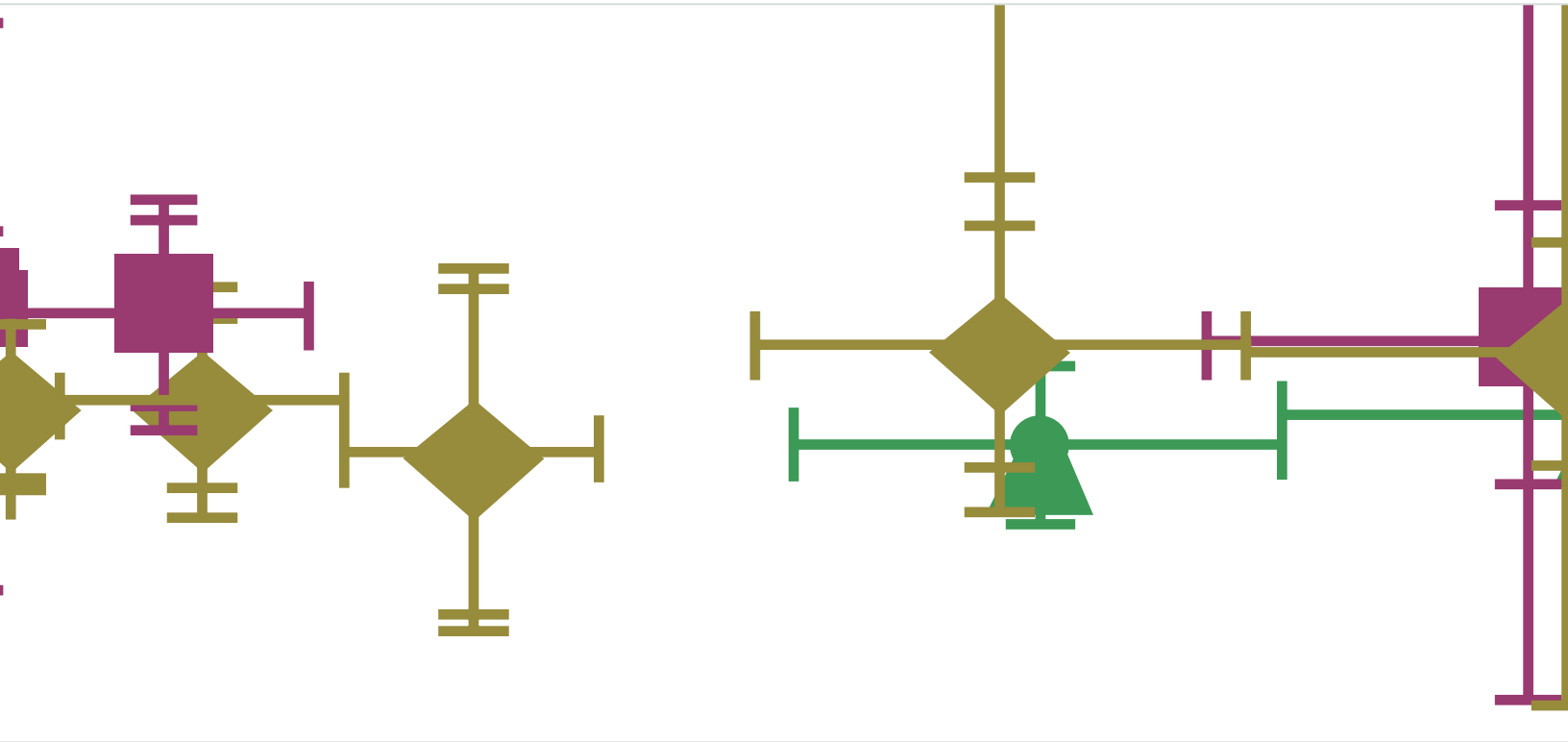
"Datasets that would take us months to process on our local desktop machines can be completed in only a few hours on the NERSC systems, and this significantly speeds up our workflow," says Manovich. "The NERSC Analytics Team has been incredibly helpful to our work. In addition to helping us develop the technical tools to process our data, they also share in our excitement, often sending us information that might be useful to the project."

Currently Manovich and his collaborators are using NERSC computers to analyze and visualize patterns across 10 million comic book images from around the globe.

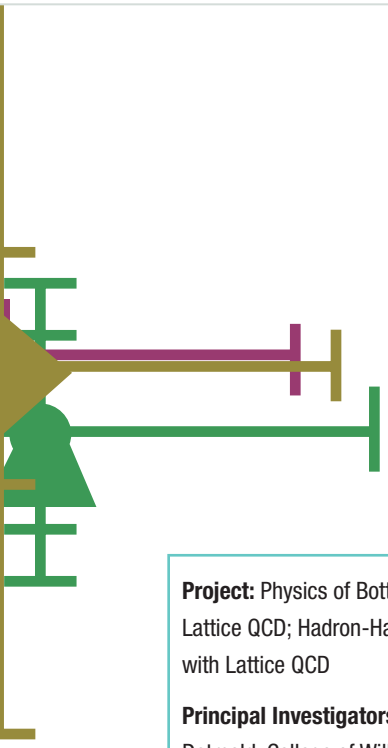
"Contemporary culture is constantly evolving with the advancement of technology. Billions of pictures, video and audio files are uploaded to the Internet every day by ordinary people all over the world. Cultural analytics is an emerging paradigm to make visible patterns contained in this ocean of media," says Manovich.

"Working with Lev Manovich gave me a lot to think about in terms of how high performance computing is underutilized in the humanities and how it could potentially accelerate knowledge discovery," says Janet Jacobsen of the NERSC Analytics Team. "All in all, this has been a really fun project to be involved with. These researchers are doing something that no one else in their field is doing."

Brace for Impact



The search for why matter dominates our universe



Project: Physics of Bottom Baryons in Lattice QCD; Hadron-Hadron Interactions with Lattice QCD

Principal Investigators: William Detmold, College of William and Mary, Thomas Jefferson National Accelerator Facility; and Martin J. Savage, University of Washington

Senior Investigators: Chi-Jen David Lin, National Chiao Tung University; Matthew Wingate, University of Cambridge; Kostas Orginos, College of William and Mary, Thomas Jefferson National Accelerator Facility

Funding: NP, DOE OJI, NSF, JMT, TNSC

Computing Resources: NERSC, BSC-CNS, FNAL, HPCS, JLab, LLNL, NCHC, NCSA, UW

While the fireworks at CERN's Large Hadron Collider (LHC) transfix the world, theorists are quietly doing some computational heavy lifting to help understand what these particle smash-ups might reveal about the fundamental mystery of existence: Why is there anything at all?

The Standard Model of particle physics can't explain why there exists more matter than antimatter in the universe. At the LHC and other colliders, scientists sift the debris of high-energy particle collisions searching for clues to physics that lie beyond our current understanding.

However, in order for scientists to "claim they've seen something beyond the Standard Model—a very important claim—they would need to know with quite high precision what the Standard Model predicts," said William Detmold, an assistant professor of physics at the College of William and Mary and senior staff scientist at the Department of Energy's Thomas Jefferson National Accelerator Facility. "That's what we try to calculate," said Detmold, whose group computes at NERSC.

Detmold works in the field of quantum chromodynamics (QCD), the mathematical theory describing the strong force that binds quarks together into protons, neutrons and other less-familiar subatomic particles. QCD also governs how particles interact with each other. Through a computational method called lattice QCD (LQCD), Detmold's team painstakingly calculates the characteristics of subatomic particles in various combinations.

Using NERSC systems, Detmold and colleagues achieved the first-ever QCD calculations for both a three-body force between hadrons and QCD calculations for a three-body baryon system. They reported their findings in the journal *Physical Review Letters*.[9] In a subsequent paper,[10] Detmold and Savage also reported another first: a QCD calculation that could help scientists better understand the quark soup that was our universe milliseconds after its birth.

A better understanding of these interactions will help physicists build

[9] Silas R. Beane, William Detmold, Thomas C. Luu, Kostas Orginos, Martin J. Savage, Aaron Torok, "Multipion Systems in lattice QCD and the three-pion interaction," Physical Review Letters **100**, 082004 (2008).
[10] William Detmold and Martin J. Savage, "Color screening by pions," Physical Review Letters **102**, 032004 (2009).

The ABCs of Lattice QCD

Lattice QCD is a computational method for working around the mathematical messiness of the quantum world.

Rather than try to calculate how a roiling soup of quarks, antiquarks, and gluons interact over all space, the method freezes those elements on a limited-volume, four-dimensional space-time grid or "lattice" in a controlled way. Quarks are stationed at each of the crosspoints or "nodes" on the grid, and forces are only calculated along the connections between nodes. QCD interactions within this space-time lattice are calculated again and again using importance sampling until statistically valid averages emerge.

The size of the cube itself is minute (about two or three times the size of the proton in each direction in Detmold's experiments); yet the number of calculations required to advance a simulation by a fraction of a second can be staggering. Detmold used a lattice with 24 nodes along each spatial dimension and 64 time nodes. That's a total of more than 880,000 nodes and 30 million connections (each node has four connections and eight colors).

Lattice QCD was proposed in the 1970s, but until supercomputers powerful enough to do the necessary calculations came on the scene in the late 1990s, it has not been a precise calculation tool. Increasing compute power, combined with recent refinements in lattice QCD algorithms and codes, has enabled physicists to predict the mass, binding energy, and other properties of particles before they're measured experimentally, and to calculate quantities already measured, confirming the underlying theory of QCD.

Lattice QCD also allows physicists to research scenarios impossible to create experimentally, such as conditions at the heart of a neutron star, said Detmold. "Our understanding of the evolution of stars through supernovae and then neutron stars depends on how neutrons and lambda baryons, for example, interact," said Detmold. "Experimentally, this is very hard to study, but in our models, we just change a line of code," he said.

Particle Roll Call

Six kinds of quarks and eight types of gluons can interact to create a virtual particle zoo, but most matter is made of just two: the neutrons and protons that comprise an atomic nucleus. Protons and neutrons are in turn comprised of quarks (held together by the exchange of gluons).

However, physicists also use a panoply of categories and particle names that sound confusingly similar: hadrons, baryons, mesons, pions, and on and on. Below are defined some of the terms used in this article.

Quark: The smallest, discrete division of matter that we know of, quarks come in six varieties (whimsically named up, down, top, bottom, strange, and charm) and three "colors," which really aren't colors at all, but rather the properties that influence how quarks group.

Hadron: Any subatomic particle made up of quarks.

Baryon: A hadron made up of three quarks.

Neutron and proton: The subatomic particles that comprise the nucleus of an atom. The neutron has no electrical charge, while the proton has a positive charge. Neutrons and protons are also baryons (have three quarks) and nucleons (because they comprise the nucleus of an atom).

Pion: The lightest naturally occurring hadron, the pion contains only a single quark and antiquark pair.

Bottom baryon: These particles contain bottom quarks and are a major area of investigation at the LHC and other particle colliders.

better models of atomic nuclei. The findings also help scientists better understand what they ought to see when certain particles collide. Anything outside those values could indicate new phenomena.

Know thyself

While Detmold's research may help form a jumping-off point for the discovery of exotic physics, he emphasizes that the work is confined to the Standard Model we know today: "The overall goal of our project is to provide a QCD-based understanding of the basics of nuclear physics, how protons and neutrons interact with each other and with other particles," Detmold said.

That's more easily said than done. Calculating the properties of subatomic particles is a fiendishly difficult business that requires billions of calculations consuming millions of processor hours. Particles can't simply be torn apart and studied quark by quark, their parts summed to make a whole. Instead, quarks exist in a seething, quantum soup of other quarks, antiquarks, and gluons, all of which must be taken into account (see sidebar "The ABCs of Lattice QCD"). Also, the quantum nature of quarks requires simulations to be run over and over (using random starting points) to derive average values.

"Ideally we'd simulate the whole nucleus of a carbon atom inside our computer and try to directly calculate from QCD its binding energy," said Detmold. Even with

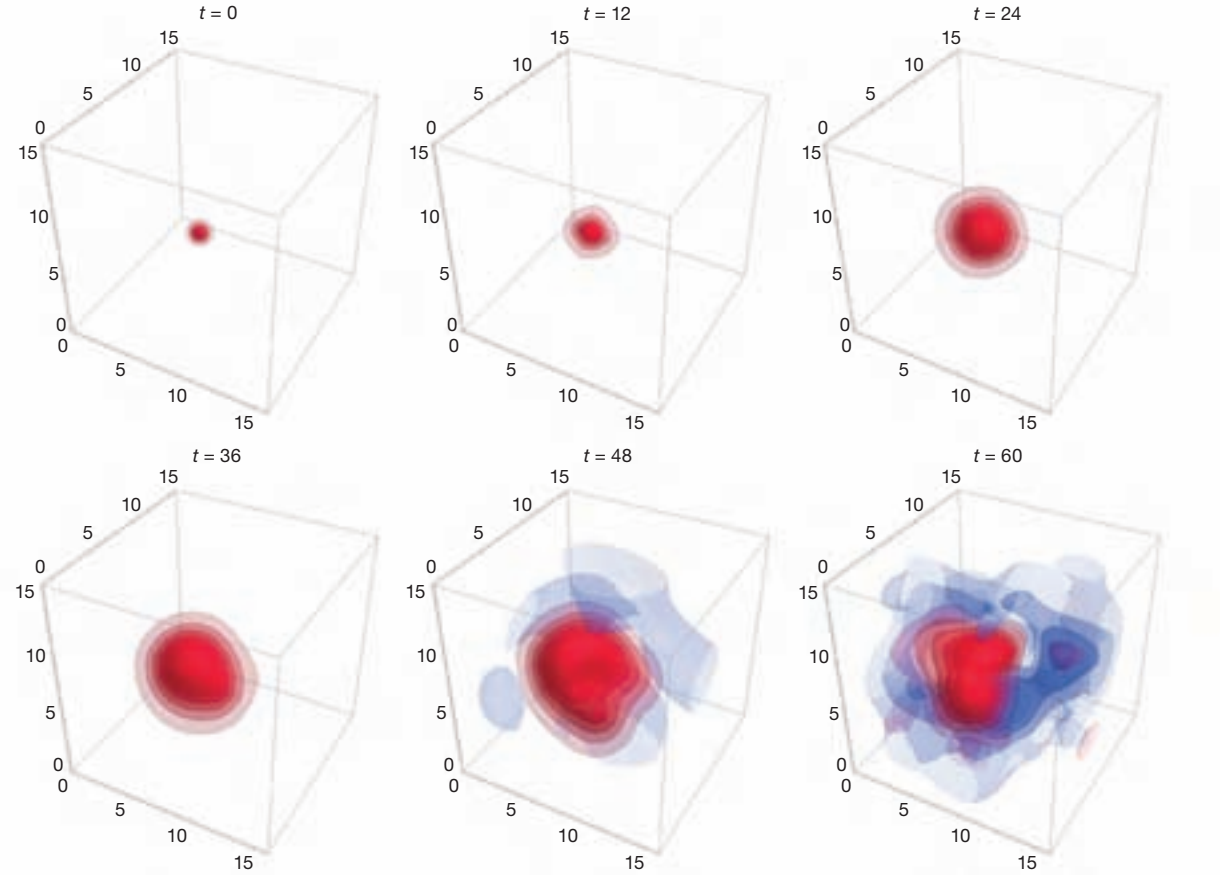


Figure 8. This selection of time-slices from a lattice QCD calculation shows the evolution of the proton correlation function over 7.02 septillionths of one second (7.02 × 10-24 seconds). To enhance detail, the colors — red being more positive and blue more negative — have been normalized on each time-slice. Information from these types of calculations enhances our understanding of the nature of matter while helping scientists search for more exotic physics. *(Image: H. W. Linn, University of Washington)*

today's computing power and algorithms, that's just not possible. "We're talking exascale-sized computations here," he said.

Instead, Detmold and colleagues substituted simpler proxies in their three-body interactions. One calculation used pions, the simplest composites formed from quarks in nature. Physicists can use these calculations to predict the properties of similar but more complex structures, atomic nuclei.

Even at their simplest, however, LQCD calculations require massively parallel supercomputers: "Machines like Franklin are very important because they have a large amount of processing power, and the fact that they are highly parallel lets us do our calculations much faster and allows us to do calculations not possible on smaller computers," Detmold said. At NERSC, Detmold's calculations ran on 4,000 cores at once. In 2009 alone, Detmold's team consumed about 30 million processor hours on systems around the world, of which NERSC provided over a third (12 million).

The search goes on

Scientists at the Relativistic Heavy Ion Collider (RHIC) at Brookhaven National Laboratory recently reported that they had produced a quark-gluon plasma, the same cosmic soup that existed milliseconds after the Big Bang. The plasma itself lasted a tiny

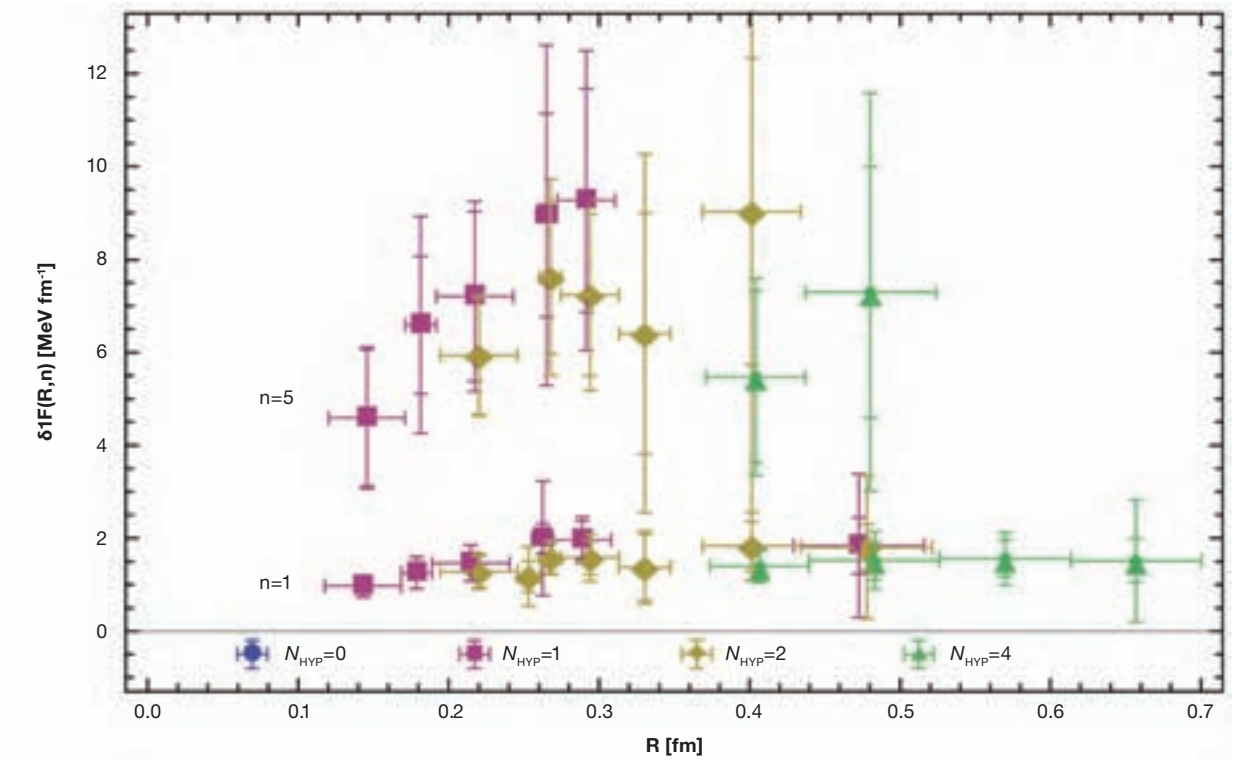


Figure 9. This chart shows the change in the contribution to the radial quark-antiquark force at two pion densities. The attractive force is slightly reduced by the presence of a pion gas, similar to what has been observed in quark-gluon plasmas.

[11] William Detmold, C. J. David Lin, Matthew Wingate, "Bottom hadron mass splittings in the static limit from 2+1 flavor lattice QCD," Nuclear Physics B **818**,17 (2009)

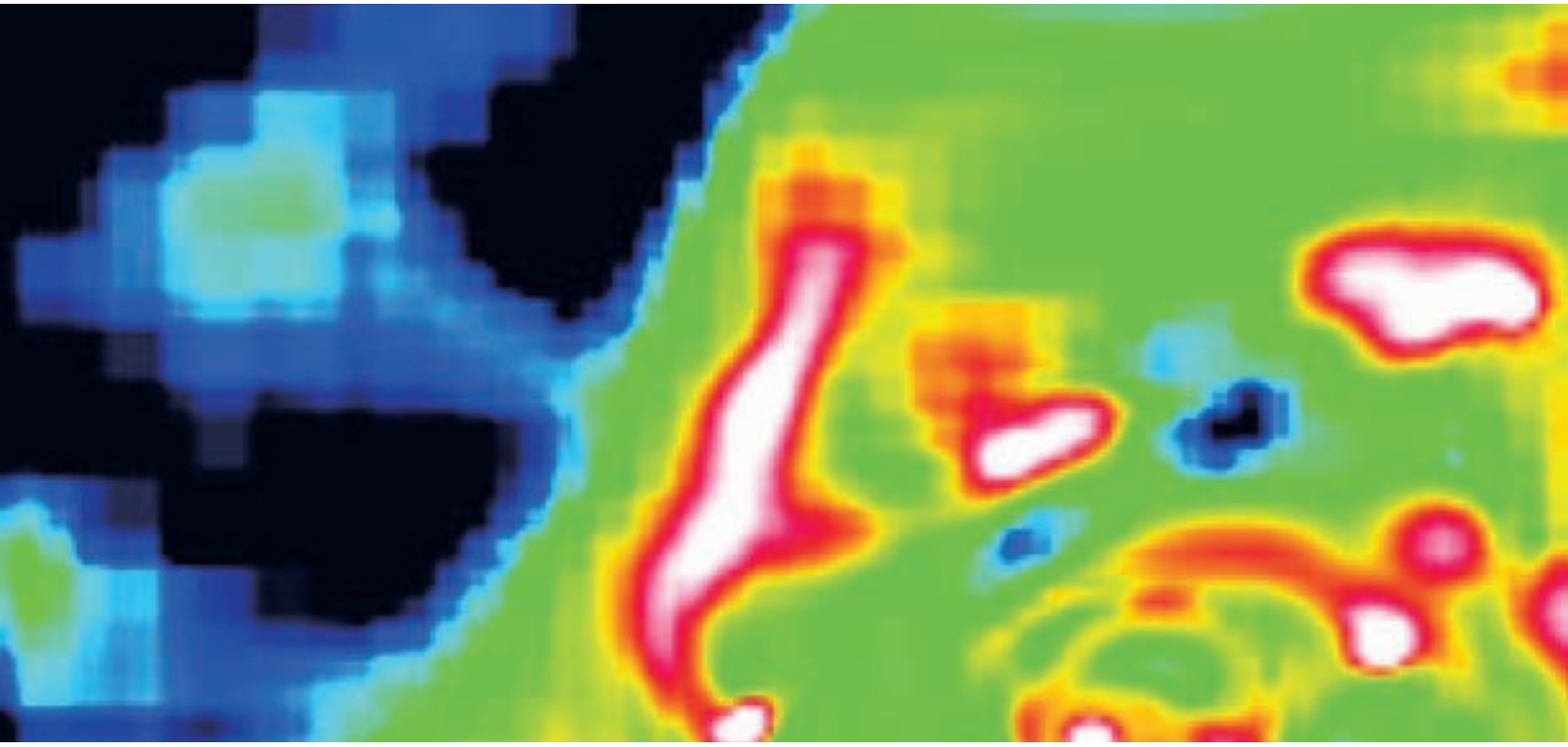
fraction of a second. A major part of the evidence for this exotic state of matter was that it emitted too few J/psi particles, the quark-antiquark pairs that shower from heavy particle collisions. This is known as J/psi screening.

"We wondered if the protons and neutrons not caught up in this plasma could also cause J/psi screening," Detmold said. Calculating a similar interaction with a lattice containing 12 pions, they found a similar J/psi screening effect, albeit to far lesser degree.

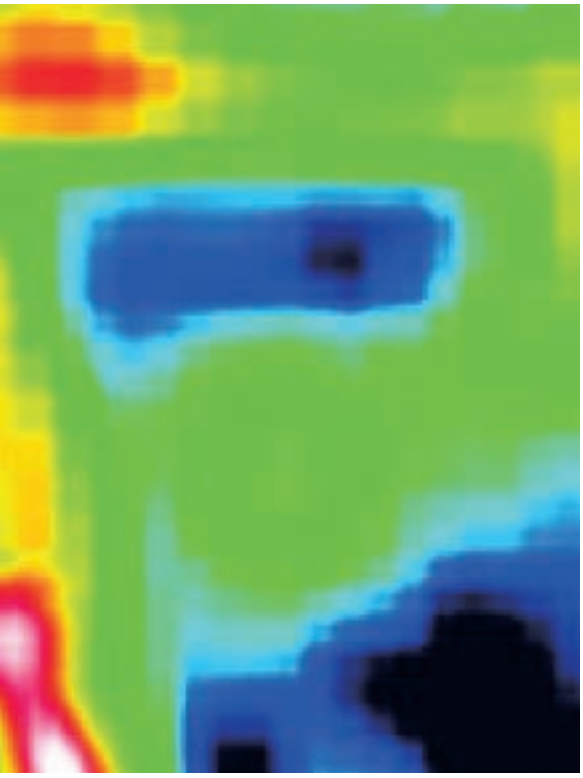
"At the RHIC, what's being collided are nuclei, which are all protons and neutrons. The pion system we're looking at is not the system that's there, but it's the simplest multi-hadronic system we can look at," said Detmold. "It indicates that at least some of the screening [at the RHIC] could be coming from hadrons," he concluded.

At NERSC, Detmold's group also concentrates on particles containing bottom-type quarks, called bottom baryons.[11] One of the four major experiments at the LHC is dedicated to probing for novel physics among these particles. Detmold said: "Our studies will contribute an important ingredient in LHC searches for physics beyond the Standard Model in the bottom-baryon sector."

Baby Brutes



Simulations help solve the mysteries of massive young star-forming galaxies



Project: Galaxy Formation Simulations

PI: Joel Primack, UC Santa Cruz

Senior Investigators: Joel Primack, UC Santa Cruz; Anatoly Klypin, New Mexico State University

Funding: DOE, NSF, NASA, ISF, GIF, DIP, FITS, ECHU, ERC, MFP

Computing Resources: NERSC, NAS, BSC-CNS

Astronomers have in recent years been surprised to find hulking brutes among the baby galaxies of the early Universe. Studded with bright, giant clumps of rapidly forming stars, these galaxies hail from a time when the cosmos was less than 4 billion years old, yet each contains about the mass of a modern Milky Way, which took 10 billion years to form.

Once considered oddities, these galaxies are now thought to be the engines that drove the Universe's most active period of star formation. It remains a mystery, however, how such massive galaxies came to be so quickly and what has happened to them in our modern Universe. Attempts to model the evolution of these so-called star-forming galaxies (SFGs) have failed. When modelers feed gas into a typical galaxy simulation at the high rates required for an SFG, the resulting galaxy can be misshapen or form too many stars too quickly, or both.

"There have been a lot of problems trying to understand galaxies. In fact, there are many aspects of galaxies that are still quite mysterious," said Joel Primack of the University of California, Santa Cruz, who is the principal investigator for the Galaxy Formation Simulation project that computes at NERSC. His group uses computer simulations and visualizations to explore questions about the structure, nature, and fate of the cosmos.

One of the team, Daniel Ceverino, may have cracked the SFG case. Zeroing in on single galaxies with high-resolution simulations and compensating for the effects of supernovae and runaway stars, Ceverino successfully simulated SFG-like galaxies using NERSC's Bassi system.

"These are the best physics, highest resolution simulations anyone has run to date of forming galaxies," said Primack, who helped originate cosmology's standard model (Lambda Cold Dark Matter) under which simulations are run. "The Bassi machine at NERSC has been one of the best machines available for doing that kind of simulation," he added.

Ceverino of The Hebrew University, Jerusalem and collaborator Anatoly Klypin of New Mexico State University outlined their groundbreaking simulations in The Astrophysical Journal.[12] In a second paper[13] published in the same

[12] Daniel Ceverino and Anatoly Klypin, "The Role of Stellar Feedback in the Formation of Galaxies," Astrophysical Journal **695**, 292 (2009).
[13] Avishai Dekel, Re'em Sari, and Daniel Ceverino, "Formation of Massive Galaxies at High Redshift: Cold Streams, Clumpy Disks, and Compact Spheroids," Astrophysical Journal **703**, 785 (2009).

journal, Avishai Dekel, Re'em Sari and Ceverino, all of The Hebrew University, Jerusalem, used the simulations to explore a theory for how these unusual galaxies formed.

The secret lives of galaxies

According to the Cold Dark Matter model, a web of mysterious dark matter underpins the visible Universe. Galaxies form when the gravity of dense dark-matter knots snags streams of relatively cold gases (approximately 10,000 Kelvin). Compressed by gravity, the gases fuel star creation. Massive stars, in turn, slow the formation of additional stars by injecting energy, mass, and metals back into the galaxy-formation process through stellar winds and supernova explosions—a process called *stellar feedback.*

Because of the large scale and subsequent low resolution of many simulations, however, stellar feedback has been poorly accounted for in typical models, especially the feedback of runaway stars. Runaway stars are massive stars ejected from their natal molecular clouds by gravitational effects. (A dramatic, though uncommon example would be when one of a binary pair goes supernova: losing the mass to hold its partner in co-orbit, it slings the other star out into space.) When some of these runaway stars explode as supernovae outside the damping effects of a giant molecular cloud, the released energy can ripple out for hundreds of light years, allowing these relatively small-scale events to have a significant impact on galaxy formation.

Ceverino, a PhD candidate at New Mexico State University at the time, and Klypin, his advisor, adjusted their simulations to account for the feedback effects of both massive star clusters and runaway stars. Testing the changes on a galactic plane with a volume of 4,000 parsecs on a side (1 parsec = 3.26 light years), the researchers successfully reproduced the characteristics they expected of stellar feedback: chimneys of fast, hot outflows of gas, and smaller, hot gas bubbles (Figure 10) resulting in galactic winds.

A more realistic model

Using the adjusted simulation, Ceverino and Klypin then modeled three massive galaxies at "high redshift." Astronomers measure the distance (and thus the age) of galaxies based on how much their once-blue light has stretched, shifting it into the red spectrum. The redder the light, the higher the redshift value, the further away (and the longer ago) is an observed galaxy.

Unlike previous models, the galaxies in the adjusted simulation looked very much like the star-forming galaxies astronomers observe (Figure 11). Instead of smooth, large central bulges, their visualizations formed thickened disks with clumpy arms, much like SFGs (Figure 12).

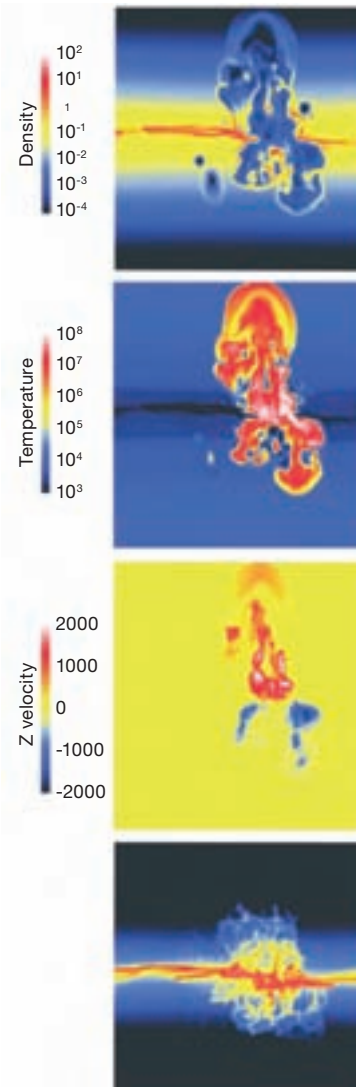


Figure 10. In this model galactic plane (seen in cross section), the energy input of stellar superclusters causes gases to shoot out at speeds up to 1000 kilometers per second at temperatures reaching 100 million Kelvin. These appear as plumes or "chimneys," visible in the top three frames (showing density, temperature, and velocity; the bottom frame shows gas column density). Small bubbles of hot gas in the field (visible in the top two frames) are the result of stellar feedback from runaway stars.

[14] A. Dekel et al., "Cold Streams in Early Massive Hot Haloes As the Main Mode of Galaxy Formation," Nature **457**, 451 (2009).

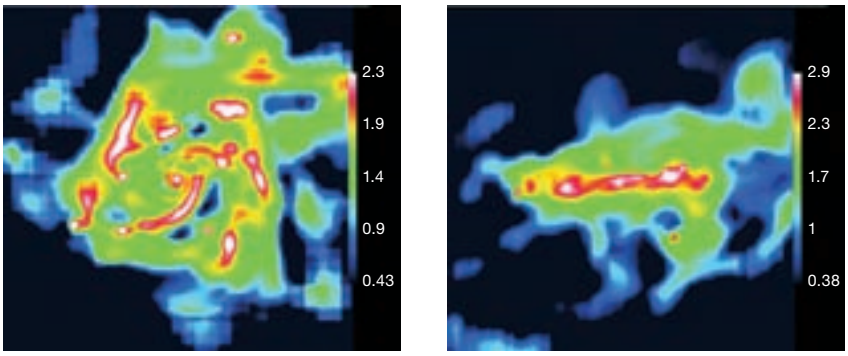


Figure 11. This simulated galaxy is smaller than the typical star-forming galaxy, but it resembles an SFG's appearance and demonstrates the general behavior predicted by theoretical analysis. In the face-on view (left) the extended disk is broken into several giant clumps and sheared perturbations, similar to many observed star-forming galaxies. The edge-on view (right) shows a well-defined disk, resembling observed "chain" systems. The disk is fed by streams, which are largely below the threshold density for the plot and so are not emphasized in these pictures.

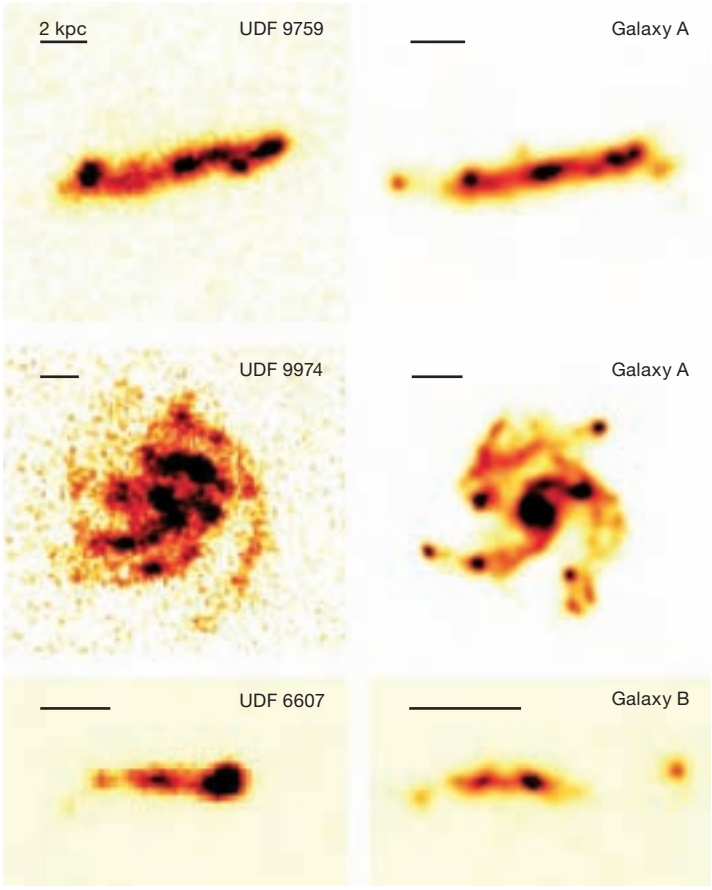


Figure 12. Models of simulated high-redshift galaxies (on the right) are compared to three galaxies observed by the Hubble Ultra Deep Field (shown left). The modeled galaxies for the first time reproduce key morphological properties of high-redshift (ancient) galaxies, including clumps of rapidly forming stars, indicated by black in these false color images.

The simulations also supported the theories of astronomer Avishai Dekel, who had proposed in a January 2009 Nature article an explanation of how SFGs form.[14] If incoming gas streams could remain "cold" (about 10,000 Kelvin) and stream continuously into the galaxy core, Dekel proposed, SFGs could form. Ceverino's models showed streams of cold, often clumpy gas streams feeding early-stage galaxies.

The models also suggest that these wild, violent galaxies may have settled down considerably in the space of a few billion years. In the simulations, star-forming clusters make a few galactic rotations before being drawn into a central bulge, which slows star formation and "quenches" galaxy growth.

Cranking up the resolution was key. Lower-resolution simulations didn't properly account for stellar feedback. At resolutions poorer than 50 parsecs, the models lost resemblance to observed galaxies. Ceverino ran simulations at resolutions as high as 35 parsecs, ten times that of typical simulations.

On both cosmological and computational scales, that's unusually fine. Our own Milky Way galaxy measures 30,000 parsecs across. At the finest resolutions, these galactic models were imaging only portions of a single galaxy. "If you're doing simulations that fine, then you can resolve the regions where stars form, though not the individual stars themselves," Primack said.

The models may also help scientists finally identify some of the oldest and most mysterious cosmic structures ever observed:

Lyman-alpha blobs. Primack and colleagues are exploring whether these glowing globs of gases may actually be evidence of the cosmic streams that fed SFGs.

Ceverino's close-up, high-resolution models will also help increase the accuracy of larger-scale simulations aimed at understanding how dark matter and visible, or baryonic, matter interact. "Nobody has the computer power or the understanding of underlying physics to try to do everything in one go," said Primack. "We learn what we can from the small-scale simulations and extend that up to larger-scale models."

The next challenge: black holes

Meanwhile, Ceverino's simulations have reached 6.2 billion years, almost half the age of the Universe, and continue to run at NERSC. The Bassi system is set for early 2010 decommissioning, but his work will continue on the Franklin (Cray XT4) and Hopper (Cray XT5) systems. Primack's group may also test this and other codes on NERSC's Magellan, a scientific cloud computing testbed.

Primack's next challenge is to incorporate the effects of black holes found at the center of massive galaxies into models. "Black holes are 1,000 times more effective than stars at releasing energy from matter, so there's no doubt these black holes play a huge part" in galaxy formation, he said.

Methods, codes, and computation

Galaxy simulations were performed on the NERSC Bassi system using more than 130,000 processor hours to run the Eulerian Hydrodynamics plus N-body Adaptive Refinement Tree (ART) code.

The code produces a tree of cubic cells. Partial differential equations (such as the Poisson equation of the equations of hydrodynamics) are solved on this adaptive mesh. Dark matter and stellar components are implemented as particles, which use the particle-in-mesh technique for density assignment and for force calculations.

This simulation has 3.4 million dark matter particles and 4 million stellar particles in a high-resolution region of 2.1 comoving megaparsecs in radius. The hydrodynamics are resolved with more than 17 million gas cells. The mass resolution is 700,000 solar masses for dark matter and 10,000 solar masses for stars. The maximum resolution is between 35 and 70 proper parsecs.

Parallelization is implemented differently for different variants of the code. For large-volume N-body-only simulations, the code uses adaptive domain decomposition to split the computational volume. Each domain is handled by one MPI task. After one zero-level time-step, domains exchange information. Boundaries of the domains (parallelepipeds) are adjusted to spread the CPU load evenly. Domains send particles to other domains to create buffer zones around each domain. Then another time-step is made, and so on.

NERSC Users' Awards and Honors

Every year a significant number of NERSC users are honored for their scientific discoveries and achievements. Listed below are some of the most prominent awards given in 2008 and 2009.

Members of the National Academy of Sciences

Emily Carter, *Princeton University*

George Smoot, *Lawrence Berkeley National Laboratory*

Fellows of the American Academy of Arts & Sciences

Emily Carter, *Princeton University*

Steven G. Louie, *University of California, Berkeley, and Lawrence Berkeley National Laboratory*

Warren M. Washington, *National Center for Atmospheric Research*

Fellows of the American Association for the Advancement of Science

Liu Chen, *University of California, Irvine*

Michael Dupuis, *Pacific Northwest National Laboratory*

Graham R. Fleming, *University of California, Berkeley, and Lawrence Berkeley National Laboratory*

L. Ruby Leung, *Pacific Northwest National Laboratory*

Juan C. Meza, *Lawrence Berkeley National Laboratory*

Joyce E. Penner, *University of Michigan*

Sotiris S. Xantheas, *Pacific Northwest National Laboratory*

Fellows of the American Physical Society

Rana Biswas, *Iowa State University and Ames Laboratory*

Jeff Candy, *General Atomics*

Thomas Devereaux, *Stanford University and SLAC National Accelerator Laboratory*

Chong Long Fu, *Oak Ridge National Laboratory*

Beate Heinemann, *University of California, Berkeley, and Lawrence Berkeley National Laboratory*

Spencer Klein, *Lawrence Berkeley National Laboratory*

Chung-Pei M. Ma, *University of California, Berkeley*

Patric Muggli, *University of Southern California*

Scott E. Parker, *University of Colorado*

Gregory K. Schenter, *Pacific Northwest National Laboratory*

Carl R. Sovinec, *University of Wisconsin, Madison*

Sergei Voloshi, *Wayne State University*

Nu Xu, *Lawrence Berkeley National Laboratory*

Fellows of the Society for Industrial and Applied Mathematics

John Bell, *Lawrence Berkeley National Laboratory*

Andrea L. Bertozzi, *University of California, Los Angeles*

Phillip Colella, *Lawrence Berkeley National Laboratory*

James Demmell, *University of California, Berkeley, and Lawrence Berkeley National Laboratory*

Christopher R. Johnson, *University of Utah*

Yousef Saad, *University of Minnesota*

James Sethian, *University of California, Berkeley, and Lawrence Berkeley National Laboratory*

U.S. Department of Energy Ernest Orlando Lawrence Award

William Dorland, *University of Maryland*

Wim Leemans, *Lawrence Berkeley National Laboratory*

IEEE Computer Society Sidney Fernbach Award

Roberto Car, *Princeton University*

American Chemical Society (ACS) Peter Debye Award in Physical Chemistry

Richard J. Saykally, *University of California, Berkeley, and Lawrence Berkeley National Laboratory*

ACS Joel Henry Hildebrand Award in the Theoretical and Experimental Chemistry of Liquids

Graham R. Fleming, *University of California, Berkeley, and Lawrence Berkeley National Laboratory*

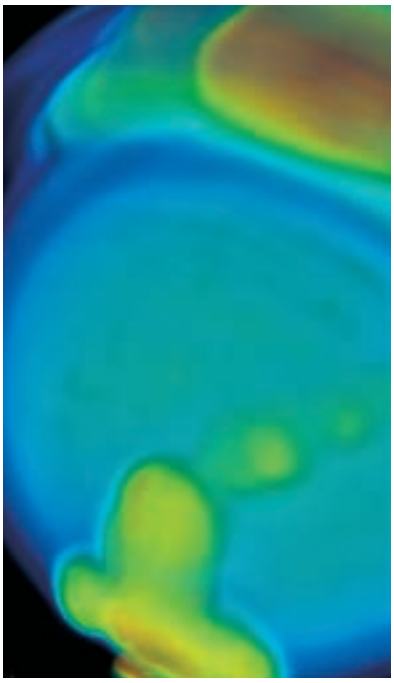
ACS Ahmed Zewail Award in Ultrafast Science and Technology

Graham R. Fleming, *University of California, Berkeley, and Lawrence Berkeley National Laboratory*

ACS "Legends of Environmental Chemistry" Award

Garrison Sposito, *University of California, Berkeley, and Lawrence Berkeley National Laboratory*

The NERSC Center: Innovation in the Service of Science



As the primary computing resource for the DOE Office of Science research community, NERSC has a tradition of providing systems and services that maximize the scientific productivity of its user community. NERSC takes pride in its reputation for the expertise of its staff and the high quality of services delivered to its users. To maintain its effectiveness, NERSC proactively addresses new challenges in partnership with the larger high performance computing (HPC) community. The following pages describe some of the ways NERSC is meeting current challenges and preparing for the future.

Upgrade doubles Franklin supercomputer's scientific capability

In July 2009, NERSC officially accepted a series of upgrades to its Cray XT4 supercomputer (Figure 1), providing the facility's 3,000 users with twice as many processor cores and an expanded file system for scientific research. NERSC's Cray XT4 system is named Franklin in honor of Benjamin Franklin, the United States' pioneering scientist.

"Franklin's upgrade has already provided a tremendous benefit to the DOE computational science community, which now has a system in which the aggregate system performance is double that of the original Franklin system," says Kathy Yelick, NERSC Division Director. "The key to improving application throughput is to maintain balance for that workload, so when we doubled the number of cores, we also doubled the memory capacity and bandwidth, and tripled the I/O bandwidth."

In a strategic effort to maintain scientific productivity, the upgrades were implemented in phases. The quad-core processor and memory upgrade was done by partitioning Franklin and performing an upgrade and test on one part while the rest of the system was available to users (Figure 2).

"The phased upgrade was engineered by Cray and NERSC staff specifically for the Franklin upgrade, and is now a model for upgrading future systems," says Wayne Kugel, Senior Vice President of Operations and Support at Cray.

A later upgrade increased the file system capacity to 460 terabytes, and the speed at which data is moved in and out of the system increased threefold. As a result of these upgrade efforts, the amount of available computing time roughly doubled for scientists studying everything from global climate change to atomic nuclei.

Figure 1. Franklin, NERSC's Cray XT4 system.

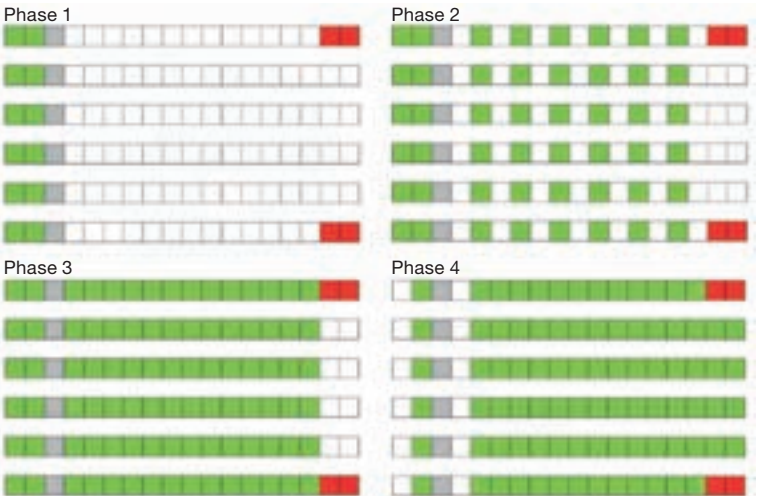


Figure 2. Four intermediate stages of the Franklin quad-core upgrade. The boxes represent Franklin system cabinets: white boxes for cabinets with dual-core nodes, green boxes for quad-core nodes, red boxes for cabinets containing both service nodes (which remained dual-core) and compute nodes, and gray boxes for columns in the NERSC machine room. Nodes were upgraded in place while Franklin remained in service for the users. Note that in Phase 4, NERSC exchanged the dual-core nodes located between the "red" service nodes with previously upgraded quad-core nodes to minimize downtime to users. In the final configuration (not shown), the last remaining dual-core nodes were upgraded with quad-core.

The final Franklin system has a theoretical peak performance of 355 teraflop/s, three and half times that of the original system. The increase in peak performance comes from doubling of cores, doubling the number of floating-point operations per clock cycle, and a modest drop in the clock rate.

"Our focus has always been on application performance, and we estimated the doubling of total system performance prior to the upgrade; the growing gap between application performance and peak emphasizes the need to measure and evaluate real application performance," says Yelick.

"This acceptance is a great achievement for the NERSC facility, and it wouldn't have been possible without a team of diligent staff from NERSC and Cray to quickly perform the hardware upgrades, identify and respond to necessary software changes, and guide our users throughout the process," she adds.

Early scientific results from the Franklin upgrade

Many NERSC users saw immediate benefits from the Franklin quad-core upgrade. Not only were codes running faster, but the researchers were also able to tackle larger, more complex problems.

"We are very pleased with the improvements on Franklin," says Gil Compo, a climate researcher at the University of Colorado at Boulder CIRES Climate Diagnostics Center and NOAA Earth System Research Laboratory. "Since these upgrades were completed, our scientific output has increased by 25 percent — meaning we can make more historical weather maps in significantly less time."

A NERSC user since 2007, Compo is leading the 20th Century Reanalysis Project to reconstruct global weather conditions in six-hour intervals from 1871 to the present. These weather maps will help researchers assess how well computational tools used in climate projections can successfully recreate the conditions of the past.

"NERSC's excellent computing facilities and phenomenally helpful consulting staff have been extremely important to our research," Compo says.

Another project, "The Role of Eddies in the Meridional Overturning Circulation" led by Paola Cessi of the Scripps Institute of Oceanography, performed over 15,000 years worth of deep ocean circulation simulations on the upgraded Franklin system. These simulations show how mesoscale oceanic flows, which are driven by surface winds and differences in solar heating, bring heat from the deep ocean to the surface (Figure 3).

"As a result of the Franklin experiments, we have been able to demonstrate that the Southern Ocean exerts remarkable control over the deep stratification and overturning circulation throughout the ocean," says Christopher Wolfe, a researcher at Scripps and a member of Cessi's team. "The results of our simulations indicate that changes in the Southern Ocean forcing could also have a large impact on the Meridional Overturning Circulation and, therefore, on global climate."

James Vary, a professor of physics at Iowa State University, led a research team that used Franklin to calculate the energy spectrum of the atomic nucleus of oxygen-16, which contains eight protons and eight neutrons. This is the most common isotope of oxygen, and it makes up more than 99 percent of the oxygen that humans breathe. Physicists currently do not fully understand the fundamental structures of

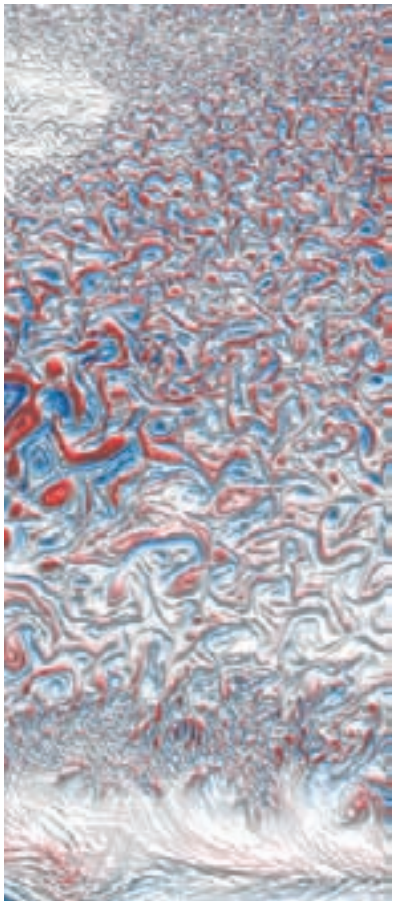


Figure 3. Results of a simulation designed to systematically explore the effect of various parameters, such as diffusivity, the pole-to-pole temperature gradient, and wind stress, on the ocean's Meridional Overturning Circulation. An interesting feature is the abundance of eddies away from the equator, which is at the bottom of the image.

atomic nuclei, especially the basic forces between the protons and neutrons. They use both experiments and calculations to expand their knowledge.

"The quad-core upgrade had a significant impact on our ability to push the frontiers of nuclear science," says Vary. "The expanded machine made 40,000 compute cores available for our research, which allowed us to calculate the energy spectrum of a nucleus that would have been previously impossible. The calculations that we did on Franklin gave us very valuable information about how to refine our theories and methods, and tell us that we have a long way to go before we can begin to understand extremely complex systems like the nucleus of a sodium atom."

Franklin upgrades improve I/O performance

Throughout the month of March 2009, Franklin underwent a major input/output (I/O) upgrade involving both hardware and software. The disk capacity of the scratch file system was increased by 30% to 460 TB, and the I/O bandwidth was nearly tripled to an aggregate write performance of 32 GB/sec, compared to 11 GB/s before the upgrade. Instead of adding the new hardware to the existing Lustre scratch file system, NERSC chose to implement a second Lustre scratch file system and reconfigured both, so that each of Franklin's two scratch file systems now has a peak write bandwidth of 16 GB/sec.

"We doubled the amount of I/O hardware and nearly tripled the bandwidth, which was a pleasant surprise," said Kathy Yelick, NERSC Division Director.

The extra boost came from a set of hardware and software changes that included a re-arrangement of the I/O nodes based on an analysis of an optimum layout for the particular torus network configuration on Franklin.

"The I/O upgrade will not only improve the peak I/O performance of applications, but should also result in more predictable performance and less network congestion even under heavy I/O workloads," said Katie Antypas of the NERSC User Services Group.

Next-generation supercomputer delivered to NERSC

Figure 4. Hopper, NERSC's new Cray XT5 system, during installation.

The first phase of NERSC's next-generation supercomputer was delivered to Berkeley Lab's Oakland Scientific Facility on October 12, 2009 (Figure 4). NERSC awarded the contract for this system to Cray Inc. in August 2009.

The system that was delivered is a Cray XT5, which will be upgraded to a future-generation Cray supercomputer. When completed, the new system will deliver a peak performance of more than one petaflops, equivalent to more than one quadrillion calculations per second.

This machine is named "Hopper" after Rear Admiral Grace Murray Hopper, the American computer scientist and United States Naval officer who invented the first compiler for a programming language, led the development of COBOL, and popularized the word *debugging*.

"As NERSC is the primary supercomputing center for DOE's Office of Science, making Cray's latest technology available to our users will accelerate innovation across a wide range of scientific disciplines, helping scientists tackle problems of vital importance to our nation's future," said Michael Strayer, Associate Director of DOE's Office of Advanced Scientific Computing Research.

According to NERSC Director Kathy Yelick, Cray was awarded the contract based on several factors, including performance and energy efficiency, on a set of application benchmarks that capture the challenging workload of the 3,000 NERSC users.

"Because we serve such a large and scientifically diverse user community, it's critical that our systems deliver the best performance while running real-world applications—especially as users scale their codes to run on tens of thousands of processor cores," Yelick said.

The new Cray system will provide many pioneering features, including the ability for users to customize the operating system for their own codes and to schedule jobs and access their data without logging in to the supercomputer. Yelick adds, "Cray's new cooling system and interconnect network technology mesh well with our research efforts into energy-efficient computing and programming models."

"We are proud that NERSC chose Cray as its ongoing partner to provide its diverse and demanding users with advanced scientific computing capabilities," said Cray President and CEO Peter Ungaro. "Our partnership with NERSC expands beyond our supercomputers to working together to get the most efficient and effective use of the systems as possible — a partnership that benefits all Cray customers around the globe. We are excited at the scientific achievements that NERSC's users have made on the Cray XT4 'Franklin' system, and we are looking forward to the advancements that will be made on both our Cray XT5 system as well as our future systems that are part of this contract."

Consisting of products and services, the multi-year contract is valued at over $50 million. The full system is expected to go into production in late 2010.

Magellan: cloud computing for science

Cloud computing is gaining traction in the commercial world, but can such an approach also meet the computing and data storage demands of the nation's scientific community? A new program funded by the American Recovery and Reinvestment Act through the U.S. Department of Energy will examine cloud computing as a cost-effective and energy-efficient computing paradigm for scientists to accelerate discoveries in a variety of disciplines, including analysis of scientific data sets in biology, climate change, and physics.

Cloud computing refers to a flexible model for on-demand access to a shared pool of configurable computing resources (e.g., networks, servers, storage, applications, services, and software) that can be easily provisioned as needed. While shared resources are not new to high-end scientific computing, smaller computational problems are often run on departmental Linux clusters with software customized for the science application. Cloud computing centralizes the resources to gain efficiency of scale and permit scientists to scale up to solve larger science problems while still allowing the system software to be configured as needed for individual application requirements.

To test cloud computing for scientific capability, NERSC and the Argonne Leadership Computing Facility (ALCF) in Illinois will install similar mid-range computing hardware, but will offer different computing environments. The combined set of systems will create a cloud testbed that scientists can use for their computations while also testing the effectiveness of cloud computing for their particular research problems. Since the project is exploratory, it has been named Magellan in honor of the Portuguese explorer who led the first effort to sail around the globe and for whom the Clouds of Magellan—two dwarf galaxies in the southern sky—were named.

One of the goals of the Magellan project is to explore whether cloud computing can help meet the overwhelming demand for scientific computing. Although computation is an increasingly important tool for scientific discovery, and DOE operates some of the world's most powerful

Figure 5. The Magellan system at NERSC will be IBM's newest iDataPlex dx360 M2 server, with 5,760 Intel "Nehalem" cores and a theoretical peak speed of more than 60 teraflops. The system will be available to users in early 2010.

supercomputers, not all research applications require such massive computing power. The number of scientists who would benefit from mid-range computing far exceeds the amount of available resources.

"As one of the world's leading providers of computing resources to advance science, the Department of Energy has a vested interest in exploring new options for meeting the overwhelming demand for computing time," said Michael Strayer, associate director of DOE's Office of Advanced Scientific Computing Research. "Both NERSC and ALCF have proven track records in deploying innovative new systems and providing essential support services to the scientists who use those systems, so we think the results of this project will be quite valuable as we chart future courses."

DOE is funding the project at $32 million, with the money divided equally between Argonne National Laboratory and Lawrence Berkeley National Laboratory, where NERSC is located.

"Cloud computing has the potential to accelerate discoveries and enhance collaborations in everything from optimizing energy storage to analyzing data from climate research, while conserving energy and lowering operational costs," said Pete Beckman, director of Argonne's Leadership Computing Facility and project lead. "We know that the model works well for business applications, and we are working to make it equally effective for science."

At NERSC, the Magellan system will be used to measure a broad spectrum of the DOE science workload and analyze its suitability for a cloud model by making Magellan available to NERSC's 3,000 science users. NERSC staff will use performance-monitoring software to analyze what kinds of science applications are being run on the system and how well they perform on a cloud.

"Our goal is to get a global picture of Magellan's workload so we can determine how much of DOE's mid-range computing needs could and should run in a cloud environment and what hardware and software features are needed for science clouds," said NERSC Director Kathy Yelick. "NERSC's users will play a key role in this evaluation as they will bring a very broad scientific workload into the equation and help us learn which features are important to the scientific community."

Looking at a spectrum of DOE scientific applications, including protein structure analysis, power grid simulations, image processing for materials structure analysis, and nanophotonics and nanoparticle analysis, the Magellan research team will deploy a large cloud testbed with thousands of Intel Nehalem CPU cores. The project will also explore commercial offerings from Amazon, Microsoft, and Google.

In addition, Magellan will provide data storage resources that will be used to address the challenge of analyzing the massive amounts of data being produced by scientific instruments ranging from powerful telescopes photographing the universe to gene sequencers unraveling the genetic code of life. NERSC will make the Magellan storage available to science communities using a set of servers and software called "Science Gateways" (see page xx) as well as experiment with flash memory technology to provide fast random access storage for some of the more data-intensive problems.

The NERSC and ALCF facilities will be linked by a groundbreaking 100 gigabit-per-second network, developed by DOE's ESnet (another DOE initiative funded by the Recovery Act). Such high bandwidth will facilitate rapid transfer of data between geographically dispersed clouds and enable scientists to use available computing resources regardless of location.

"It is clear that cloud computing will have a leading role in future scientific discovery," added Beckman. "In the end, we will know which scientific application domains demonstrate the best performance and what software and processes are necessary for those applications to take advantage of cloud services."

The Magellan system at NERSC will be IBM's newest iDataPlex dx360 M2 server, which features double the memory and even higher power efficiency than previous versions (Figure 5). The iDataPlex is on the list of most energy efficient computers in the world, the Green500 List. NERSC's Magellan system has 5,760 processor cores and a theoretical peak speed of more than 60 teraflops. Magellan will be subdivided into smaller resource pools based on the requirements of different cloud testbeds and types of cloud research being done. The system will be available to users in early 2010.

Jeff Broughton Brings 30 Years of HPC Experience to NERSC as New Head of Systems Department

"High-end computing is something that gets your juices flowing—you're working with people dealing with some of the most complex and important problems of our time and using supercomputers to try to find the answers."

Jeffrey M. Broughton, who has 30 years of high performance computing (HPC) and management experience and most recently served as senior director of engineering at QLogic Corp., joined NERSC as Systems Department Head on August 3, 2009.

"I'm very pleased to announce that Jeff Broughton has accepted the position of Systems Department Head at NERSC," said Kathy Yelick, NERSC Director. "This is a key position for us, and Jeff will be responsible for management of the computing, storage, networking and security groups. His high-performance computing experience spans architectures, operating systems, compilers, interconnects, chips, and board layouts. He understands how the HPC components fit together and interact."

"I've had a career-long interest in HPC, from my days at Livermore to Amdahl, Sun, PathScale and QLogic, so when I heard about this opportunity, I thought that NERSC would be a great match for my skills and interests," Broughton said. "High-end computing is something that gets your juices flowing—you're working with people dealing with some of the most complex and important problems of our time and using supercomputers to try to find the answers."

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

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

Broughton said that his experience working for different vendors to develop both hardware and software has given him insight into what vendors can do, and he is looking to develop closer partnerships with key vendors. This is especially important as both HPC vendors and architectures consolidate, leaving centers such as NERSC with a narrower set of options.

"It's becoming more important to work with vendors to develop novel and interesting approaches, and I think that the national laboratories can take a lead role in this area," Broughton said.

New mid-range system available soon

NERSC's next medium-sized, general purpose scientific computing system will, like Magellan, be an IBM iDataPlex Linux cluster. This system will replace Bassi (an IBM Power5) and Jacquard (a Linux cluster) in early 2010. The new IBM system, selected in a competitive procurement, provides excellent performance, good energy efficiency per flop, and a familiar environment for mid-range parallel applications.

The system, named after the American scientist and educator George Washington Carver, will consist of 3,200 computational cores, configured as 400 nodes with two quad-core Intel Nehalem 2.67 GHz processors and 24 GB of memory per node. Carver's peak performance will be 34.2 Tflops, which is about 3.5 times more powerful than Bassi and Jacquard combined. Every node will have a full-featured Linux OS, and all file systems will be hosted on the NERSC Global Filesystem.

Increasing the filesystem bandwidth

In April 2009, an additional 110 terabytes of storage was added to the NERSC Global Filesystem (NGF), which was launched to facilitate data sharing between science users and machines. The system currently contains close to 300 terabytes of user accessible storage, allowing users to store larger datasets without having to move data between disk storage and the archival tape storage system.

"We worked very hard to ensure minimal disruptions to our users, and we succeeded," says Shane Canon of the NERSC Technology Integration Group. "The additional space was added without taking the file system offline."

Canon credits the seamless upgrade to the advanced capabilities of NGF's underlying file system, IBM's General Parallel File System (GPFS). NGF is mounted on all of NERSC's computing systems, allowing users who run applications on multiple machines to access information from one place, instead of copying large datasets from one machine to the next. For a handful of large-scale science projects, NGF also provides permanent online storage.

The NGF network architecture was significantly enhanced to increase the bandwidth to NGF from the computational systems. This included the rollout of special file system gateways. The gateways rely on software changes to GPFS that were the result of close work between NERSC and the IBM GPFS developers. As a consequence of these enhancements, the single stream bandwidth on many systems was improved by a factor of 5.

In addition, Franklin was converted to running native GPFS clients, which improved the performance by over 10 times on the login nodes.

NERSC staff worked closely with Cray on the development and testing of the Data Virtualization Service (DVS) software that will allow the NERSC Global Filesystem to be accessed from the Franklin Cray XT4 compute nodes. This collaboration included discussions with Cray developers to explain how to best interface with GPFS along with joint testing efforts at scale on Franklin.

Providing 60 petabytes of storage with cutting-edge technology

NERSC's High Performance Storage System (HPSS) can now hold 60 petabytes of scientific data—equivalent to all the music, videos or photos that could be stored on approximately 523,414 iPod classics filled to capacity. This is important because NERSC users are storing more and more data in HPSS. Production usage of HPSS took a marked jump of 50 percent more data daily since Franklin went into production with an average of 15 TB per day and peak loads of over 30 TB per day.

This 37 PB increase in HPSS storage capacity was made possible by deploying cutting-edge technologies—the SunT10000 generation B tape drive, which holds 1 TB of data on a single cartridge, and three Sun Slimline 8500 tape libraries, which add 15,000 new cartridge slots to the system.

"Since NERSC's Cray XT4 system, Franklin, went into production in March 2008, we've seen a 50 percent increase in new data archived to HPSS, " says Jason Hick, who heads NERSC's Storage Systems Group. "The new tape technology allows us to fill the ever-increasing demand for scientific archive space, with a negligible increase in our power consumption and very little impact on our floor space budget."

According to Hick, one SunT10000 cartridge can hold the equivalent of all the data stored on three to five cartridges built with technology from 10 years ago. In addition to being cost and space efficient, tape storage is also a very green solution to the ever-increasing amount of scientific data generated as technology improves.

"At NERSC, scientists have the option to keep what they immediately need on disk, and store the information that they are not currently analyzing on tape for later use. Since the tape cartridge sits in a library slot offline, there is no power associated with that," says Wayne Hurlbert, also of the Storage Systems Group. "Tape technology allows us to increase our storage capacity without increasing our power bill."

Figure 6. Sun SlimLine 8500 tape library, external and internal views.

Science Gateways — innovating how scientists connect to NERSC

The traditional way of connecting the computational and data resources that NERSC provides the science community is through UNIX accounts, shell-level commands, and application program interfaces (APIs). NERSC is working to build more efficient web-based interfaces through its Science Gateways. NERSC staff are working with science teams to craft web page interfaces that allow groups to submit jobs, move data, and analyze data more easily. The Science Gateway project is taking NERSC in a direction that allows people to access huge data resources in ways that are more familiar to them, through common web interfaces. The core functions addressed so far through this work address management of data, processing of data, and managing jobs submitted to NERSC production queues.

In some cases these interfaces also provide a "home on the web" for a science team where they can merge, analyze, and refine their data sets. When requested, those data resources can be made available to the public. Since NERSC's gateways are well connected to the NERSC Global Filesystem and production computing machines, they provide added value over building such interfaces remotely or using grid APIs.

Innovation in bringing HPC to the web is part of a broader strategy to make facilities funded by the DOE Office of Advanced Scientific Computing Research (ASCR) into scientific instruments that are widely useful to a large community of researchers. The web is an important addition to the traditional services NERSC provides through shell-level access.

Currently there is no one-size-fits-all approach to Science Gateways, which are crafted in collaboration with science teams to meet their particular needs. A potential future innovation as these web interfaces mature is to provide a standard set of web-based service offerings, similar to Google's application engine, from which researchers using NERSC could build their own gateways from scratch.

Three active Science Gateway projects are described below: Deep Sky, GCRM, and the Gauge Connection.

Deep Sky: web interface for astronomical image data

Deep Sky is an astronomical image database of unprecedented depth, temporal breadth, and sky coverage. Image data are gathered from the Near Earth Asteroid Tracking (NEAT) project from the 3-CCD and Quest 112-CCD cameras on the Samuel Oschin telescope at the Palomar Observatory in San Diego County, California (Figure 7).

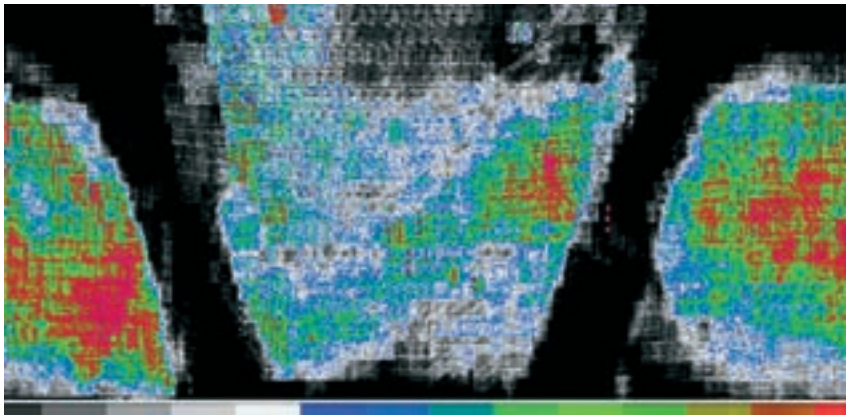


Figure 7. This image is a map of the sky as viewed from Palomar Observatory in California. Each color shows the number of times an area was observed: grey-white is about 30, blue is about 70, green is about 100, and red is about 200 observations. All of these images are stored in the NERSC archive, which contains approximately 11 million cosmic images taken over a period of nine years.

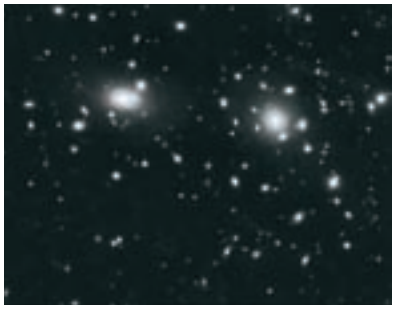


Figure 8. This image of the Coma cluster was made by combining over 500 images collected between 2001 and 2007 and stored in the Deep Sky database.

Containing a total of nine million images, or 60 terabytes of image data, Deep Sky covers nearly the entire northern sky.

"This unique collection of data allows astronomers to track how the sky has changed over the past nine years," says Peter Nugent, an astrophysicist and member of the NERSC Analytics Team, who is the project lead for Deep Sky. "It will serve as an invaluable resource for astronomers who are interested in finding cosmic events like supernovae and gamma ray bursts, or tracking the trajectories of asteroids and comets."

To streamline this effort and make observations instantly accessible and useful to researchers, members of the Deep Sky team developed a system to automatically copy all of the archived raw data and process it. Then they incorporated the processed data into the NERSC Global Filesystem, where the information is stored on disks. In a way similar to how an MP3 device can instantly search a database of music files and immediately play a song, this system allows users to query the Deep Sky database and instantly pull up processed observations for analysis—for example, combining images to improve resolution (Figure 8).

"As an astronomer, instant access to the Deep Sky dataset is extremely valuable. I can search the images for potential events—supernovae or gamma ray bursts—by looking for dots that appear for a while and then disappear," says Nugent. "I can then follow up on any oddities with even more powerful tools like NASA's Hubble Space Telescope to find out what is really going on."

GCRM: interactive subselection of climate data

Karen Schuchardt, a computational scientist at the Pacific Northwest National Laboratory, believes the NERSC Science Gateway tools present an opportunity to help researchers remotely manage large datasets. She is currently heading an effort to create a portal that allows researchers to remotely access data generated by the Global Cloud Resolving Model (GCRM) project. Members of this team seek to develop a model that can simulate Earth's climate at a 2 to 4 kilometer resolution across the entire globe.

"Simulating climate across the entire globe at these resolutions is an extremely complex and resource-intensive process, requiring extensive amounts of computer and human hours," says Schuchardt. "We cannot easily generate this data every time someone needs it, so we view each dataset as an extremely valuable resource and want to make it available to as many collaborators as possible."

Because GCRM datasets are extremely massive, she notes that it would take a long time to transfer an entire dataset across a network. Also, most researchers can only analyze a small portion of the data at a time. So in addition to creating a portal that will deliver this

information to collaborators, Schuchardt is working with NERSC staff to create a gateway tool that will allow the scientists to remotely access the data at the supercomputing centers where it is generated, search the metadata for what they need, and download that portion for analysis.

"Management for these volumes of data is a major challenge, and the Science Gateway tools that we are building will go a long way toward helping our remote researchers get the data they need, which will in turn pave the way for scientific breakthroughs," says Schuchardt.

The Gauge Connection: accessing lattice QCD datasets

"Data-sharing tools create a fertile environment for scientific breakthrough by streamlining the scientific process," says James Hetrick, a professor of physics at the University of the Pacific in Stockton, California, who is currently working on a NERSC Science Gateway project called the Gauge Connection.

According to Hetrick, the Gauge Connection portal is especially useful for researchers interested in quantum chromodynamics (QCD), the theory that describes the complex interactions between quarks —the constituents of protons, neutrons, and certain other subatomic particles. For example, physicists currently do not understand where protons get their mass. Although protons are each made of just three quarks, the quarks' mass actually accounts for less than 2 percent of the proton's mass. Scientists now suspect that the mass may actually come from the "glue," or the strong force, that binds quarks together inside the proton; however, the exact mechanism of this is unclear.

"We study quarks and the strong force because they are among the most basic constituents of matter," says Hetrick. "In the 19th century, nobody knew anything about atoms or the electromagnetic force that holds them together, but once we figured out the basic physics behind this phenomenon, we were able to move into the electric age, and subsequently, with our understanding of quantum physics, the electronic and information ages."

To learn more about the strong force, physicists use supercomputers to create a series of QCD lattices. These are four-dimensional representations of the quantum fluctuations of quarks and the force fields between them in a very tiny space-time region. Analyzing a large set of such lattices allows researchers to understand the physics of quarks.

According to Hetrick, a series of lattices can take years to generate on a supercomputer, even with tens of thousands of processors dedicated to the project. So far his team has generated over 20 TB of data, equivalent to more than 25,600 hours of video, at various computing centers around the county. As part of the Gauge Connection science gateway project, his team will consolidate all of this data at NERSC. When this portal launches, any researcher interested in accessing this information can get it through the Gauge Connection science gateway via any web browser.

Much of the data is currently hosted at the Gauge Connection (http://qcd.nersc.gov/). Of this data collection Hetrick says, "The HPSS staff helped me set up one of the earliest data sharing archives in 1998, now part of a larger national effort toward Science Gateways. This archive has a venerable place in the lattice community and is known throughout the community as the 'NERSC Archive.' In fact, until recently the lingua franca for exchanging lattice QCD data was 'NERSC format,' a protocol developed for the archive at NERSC."

"Once you are done with the lattice production stage, there are many different sorts of analysis projects that one might do, and that's where the science gateways come in. Rather than having to regenerate new lattices with the same kind of quantum fluctuations, which is very costly, other scientists can use existing sets to do the analysis part for their own ideas," says Hetrick. "These gateways greatly expand the scientific process by allowing us to recycle very valuable data."

NERSC Initiative for Scientific Exploration (NISE) encourages innovation

The NERSC Initiative for Scientific Exploration (NISE, pronounced "nice") program was initiated in September 2009, and 17 projects were awarded 9.6 million hours of computer time by the end of the year. NISE allocations are intended for researchers exploring new scientific research areas or programming models, such as:

- A new research area not covered by the existing ERCAP proposal: this could be a tangential research

project or a tightly coupled supplemental research initiative.

- New programming techniques that take advantage of multicore compute nodes by using OpenMP, Threads, UPC or CAF: this could include modifying existing codes, creating new applications, or testing the performance and scalability of multicore programming techniques.
- Code scaling to higher concurrencies: projects which have demonstrated the ability to run codes at higher scales, but do not have a large enough allocation to run at that scale, may apply.

The NISE program is ongoing, and awards will be announced once a month or so for the first half of the allocation year. For 2010, 25 million hours will be available. Researchers can apply online at http://www.nersc.gov/nusers/accounts/NISE.php.

The 2009 NISE awards went to PIs at nine national laboratories and eight universities. Research topics included nanoscience and materials science, biology and medical science, combustion science, climate science, chemistry and chemical physics, and plasma physics. Although many NISE projects are exploratory, some achieved notable scientific results.

For example, the project "Energy Landscape of Nanoclusters," led by De-en Jiang of Oak Ridge National Laboratory, explored the energy landscape of nanoclusters by employing the density functional theory-based basin-hopping technique for global-minimum search. They used plane-wave density functional theory (as implemented in VASP) to do fast geometry optimization, to power the basin-hopping global minimum search. The researchers studied a series of gold sulfide nanoclusters and found very interesting results that provide new insights into gold sulfide nanoclusters at ~1 nm (Figure 9). These results were reported in a paper submitted to the journal Angewandte Chemie.

The NISE award enabled the "Reformation of Methanol on Metal Surfaces" project, led by Hua Guo of the University of New Mexico, to explore the initial step of methanol steam reforming on a PbZn surface using density functional theory (DFT). Previously it was believed that the initial dissociation step for adsorbed methanol has a relatively low barrier. However, the DFT calculations enabled by the NISE award indicated that this is not the case, and that it is quite difficult to dissociate methanol on this catalyst. This conclusion is consistent with a recent experimental observation by John Vohs at the University of Pennsylvania that PbZn does not facilitate significant methanol decomposition. This discovery forces researchers to rethink the current model for methanol reformation. A manuscript based on the calculations is in preparation.

Another NISE project, "Studies of the *Sec translocase* Transmembrane Channel in Protein Synthesis," led by Thomas Miller of the California Institute of Technology, has resulted in one paper in print, one submitted, and one in preparation. A critical step in the biosynthesis of many proteins involves either translocation across a cellular membrane or integration into a cellular membrane. Both processes proceed via the *Sec translocase* — a ubiquitous and highly conserved transmembrane channel. Two key projects in Miller's research group are the direct simulation of coupled electronic and nuclear dynamics in complex systems, and the investigation of channel-facilitated regulation of protein translocation and membrane systems. The NISE award provided computational resources that were critical to the advancement of both projects. It allowed them to investigate the electron injection dynamics in liquid water using the recently developed ring polymer molecular dynamics (RPMD) method, and it allowed them to perform direct simulations of electron transfer dynamics among ions in solution. The NISE award also enabled new insights into the role of peptide substrates

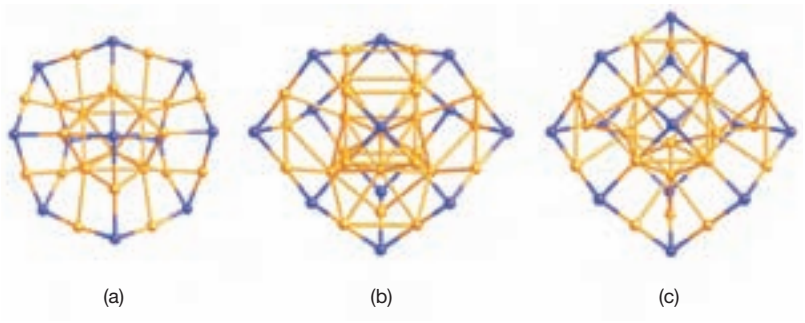


Figure 9. Global minima for (a) $Au_{23}S_{11}^-$, (b) $Au_{25}S_{12}^-$, (c) $Au_{27}S_{13}^-$.

in stabilizing large length scale conformational changes of the Sec protein translocon and in the regulation of protein translocation vs. membrane integration.

Descriptions of all the 2009 NISE projects are available at http://www.nersc.gov/nusers/accounts/NISE2009awards.php.

Speeding up science data transfers between DOE facilities

As scientists conduct cutting-edge research with ever more sophisticated techniques, instruments, and supercomputers, the data sets that they must move, analyze, and manage are increasing in size to unprecedented levels. The ability to move and share data is essential to scientific collaboration, and in support of this activity, network and systems engineers from NERSC, the Department of Energy's Energy Sciences Network (ESnet), and the Oak Ridge Leadership Computing Facility (OLCF) have teamed up to optimize wide-area network (WAN) data transfers.

With the installation and deployment of new dedicated data transfer nodes at NERSC and OLCF linked by ESnet, researchers are now able to move large datasets between each facility's mass storage system at a rate of 200 megabytes per second (MB/sec). At this rate, 74 TB of information in the U.S. Library of Congress' digital collection could be transferred in approximately four days.

According to Jason Hick, NERSC Storage Systems Group lead, WAN transfers between NERSC and OLCF increased by a factor of 20 with the new dedicated nodes, which are tuned and optimized specifically for wide-area transfers. Prior to this installation, wide-area data transfers between the two sites used infra- structure and tools that were tuned and optimized for local-area transfers. This slowed data movement between the two supercomputing centers, creating a bottleneck to scientific progress.

In addition to building the infrastructure, engineers from ESnet, OLCF, and NERSC have been collaborating on strategies for optimizing bandwidth performance between the various data storage systems at the supercomputing sites. Both sites deployed perfSONAR network monitoring applications on their servers during the testing phase to identify the transfer "choke points," where data stalled between the two facilities. The perfSONAR findings allowed staff at both sites to make the necessary adjustments to alleviate congestion. The engineers were also able to identify a variety of user-specific tuning parameters that will enable the best transfer rates possible between the two facilities. These tips are published at http://fasterdata.es.net.

Enhanced SSH daemon allows intrusion detection, improves long-haul transfers

Recent enhancements to NERSC's SSH daemon have improved security and increased the performance of long-haul transfers for users. SSH (Secure Shell) is the security software that protects most inbound connections to the center, and the daemon is the server software that handles those connections.

SSH provides secure communication over insecure channels by encrypting keystrokes and system responses, including passwords. Encryption prevents "password sniffing," one of the most common means by which computer system security is compromised. SSH thus allows users to remotely log onto computers over a network, execute commands, and transfer files from one system to another, with the utmost privacy.

"SSH has become a standard and is required for all users logging into NERSC systems," says Craig Lant, NERSC security analyst. "But because SSH encrypts commands that are typed into NERSC computers by remote users, it makes traditional intrusion detection less effective."

For the past two years, NERSC's security team has been developing and improving their own in-house variant of the SSH deamon that improves intrusion detection by capturing the user keystrokes while still preserving SSH's encryption of data in transit. This modification allows NERSC's intrusion detection system, called Bro, to automatically spot the signs of a security breach and immediately notify the security team. In this way, NERSC is able to detect that hackers have gained access to a user account before any real damage is done.

Lant notes that the keystrokes of a malicious intruder can often be identified automatically because they are somewhat different from those of a legitimate user. The modifications to SSH enable Bro to spot this

suspicious activity immediately, which allows the security team to block the intruder and alert any users whose accounts may have been compromised.

In addition to these modifications, the NERSC security team also installed a patch to the SSH client that improves the performance of long-distance transfers of massive datasets. This particular patch came from the Pittsburgh Supercomputing Center. Lant notes that prior to this patch, SSH was not tuned for managing massive, long-distance data transfers. Major remote experimental devices, such as the Large Hadron Collider in Europe, will be producing petabytes of data that will be transferred to NERSC for analysis and storage, and those transfers will go faster with the SSH upgrade.

User requirements workshops

NERSC has begun a new process for obtaining user input on computational needs at a series of requirements workshops that will be held every three years for each science discipline. The workshops elicit information from program managers and key scientists about the current and future computational requirements for each science area managed by the particular program office. They also give NERSC staff the opportunity to have extended, face-to-face discussions with researchers on how individual science goals drive NERSC center requirements.

In May 2009, NERSC held its first such workshop for DOE's Office of Biological and Environmental Research (BER) with DOE program managers from BER and ASCR and key computational scientists from biology, climate modeling, and environmental modeling. A High Energy Physics workshop was held in November 2009, which discussed accelerator physics, astrophysics, lattice QCD, and detector simulation and data analysis. A Basic Energy Sciences workshop in February 2010 will discuss material sciences, chemistry, geosciences, and combustion. Reports from these workshops are posted online (http://www.nersc.gov/projects/science_requirements/) as they become available.

These workshops are already yielding results. For example, at the BER workshop, scientists said that they wanted to be able to reserve portions of NERSC resources at specific times. As a result, NERSC is now testing a Franklin Dedicated Time Reservation service to allow users fixed periods for debugging codes at very large scale.

NERSC also gathers information via its ERCAP allocations request process, by monitoring the way its computing and storage systems are used, through the NERSC Users Group (NUG), and in day-to-day interaction with scientists who use the facility and their DOE program managers.

Data obtained from requirements studies are combined with monitoring of technology trends to help set enterprise-wide objectives for NERSC's science-driven systems and services. These data have a direct impact on NERSC planning and procurements for computing, storage, networking, and visualization platforms as well as the entire range of support activities that enable a broad range of scientists to effectively use NERSC systems in their research.

Making visualization more scalable

The size and complexity of large-scale simulations and experimental datasets make it challenging to visualize and analyze the data. The typical practice has been to run the simulations on large supercomputers, then transfer the data to a smaller, specialized server (like DaVinci at NERSC) for visualization and analysis. But for massive datasets such as supernova simulations, it makes sense to take advantage of the parallel capabilities of a supercomputer for both simulation and visualization.

NERSC's Analytics Team has demonstrated that VisIt—an interactive, parallel visualization tool that allows researchers to visualize and analyze the results of terascale simulations—can take advantage of the growing number of cores on today's advanced supercomputers, using them to process unprecedentedly large problems. They did this by installing VisIt on Franklin and using it to visualize data from an astrophysics simulation, which was augmented to create a sample scientific dataset of the size expected in the near future. The primary objective of this experiment was to better understand problems and limitations that might be encountered at extreme levels of concurrency and data size.

The team ran VisIt using 8,000

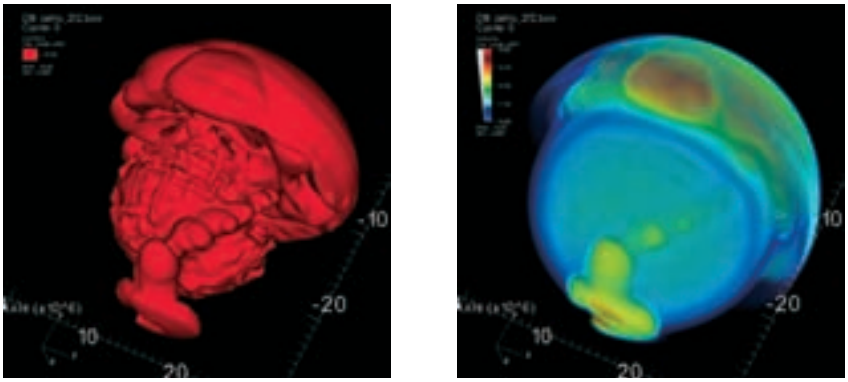


Figure 10. Supernova simulation of two trillion zones on 32,000 cores of Franklin: left, isocontour; right, volume rendering.

to 32,000 processing cores on Franklin to tackle three-dimensional datasets ranging from 500 billion to 2 trillion zones, or grid points. Data was loaded in parallel, with the application performing two common visualization tasks—isosurfacing and volume rendering—and producing images (Figure 10). From these experiments, the team collected performance data that will help them both to identify potential bottlenecks and to optimize VisIt before the next major version is released for general production use at supercomputing centers.

"These results are the largest-ever problem sizes and the largest degree of concurrency ever attempted within the DOE visualization research community. They show that visualization research and development efforts have produced technology that is today capable of ingesting and processing tomorrow's datasets," says E. Wes Bethel, leader of the NERSC Analytics Team and co-leader of the Visualization and Analytics Center for Enabling Technologies (VACET), which is part of DOE's SciDAC program.

VisIt is now available to NERSC users on both Franklin and DaVinci.

Analytics Team develops benchmark for scientific graphics software

The NERSC Analytics Team has developed a performance benchmark for the Open Graphics Library (OpenGL), a graphics industry standard specification defining a cross-language, cross-platform API for writing applications that produce 2D and 3D computer graphics.

The NERSC benchmark, called svPerfGL, focuses on measuring OpenGL rendering performance in the presence of extremely heavy graphics payload with relatively few OpenGL state changes, which is a typical of a workload incurred by scientific visualization applications. In contrast, industry standard benchmarks like SPECviewperf generate workloads typical of CAD and gaming applications, which do not always apply to scientific visualization.

svPerfGL is written in a combination of C and C++ and uses calls consistent with the OpenGL 1.2 specification. Therefore, it is very conservative in its use of the OpenGL API and should run on any modern platform that supports OpenGL. svPerfGL is available under an open source license at http://www-vis.lbl.gov/Research/svPerfGL/.

HDF5 I/O library performance analysis and optimization

NERSC has partnered with the non-profit Hierarchical Data Format (HDF) Group to optimize the performance of the HDF5 library on modern HPC platforms. The HDF5 library is the third most commonly used library package at NERSC and the DOE SciDAC program, according to recent surveys, and the most commonly used I/O library across DOE computing platforms. HDF5 is also a critical part of the NetCDF4 I/O library for the CCSM4 climate modeling code, which provides major input to the Intergovernmental Panel on Climate Change's assessment reports.

Because parallel performance of HDF5 has been trailing on massively parallel HPC platforms, NERSC has worked with the HDF Group to identify and fix performance bottlenecks that affect key codes in the DOE workload, and to incorporate those optimizations into the mainstream HDF5 code release so that the broader scientific and academic community can benefit from the work. Figure 11 demonstrates the performance improvements on two codes.

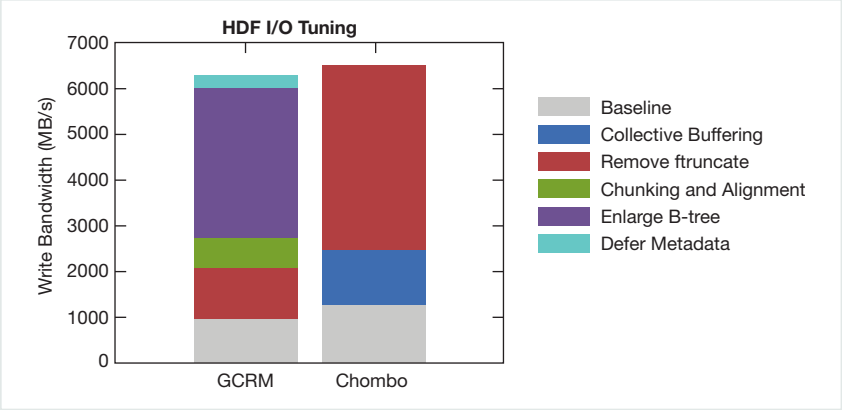


Figure 11. This plot illustrates performance improvements that were derived from the tuning work performed at NERSC in collaboration with the HDF Group. The GCRM is the Global Cloud Resolving Climate Model from Colorado State University, and Chombo is the adaptive mesh refinement framework from LBNL — two very demanding I/O intensive codes. The baseline (gray) is the original performance, and colored bars on top of the baseline show the performance benefits derived from the optimization methods. The optimizations benefit a wide variety of scientific applications and are now part of the standard HDF5 library.

Network-on-chip energy efficiency improvements

As multiprocessors scale to unprecedented numbers of cores in order to sustain performance growth, it is vital that these gains are not nullified by high energy consumption from inter-core communication. Recent advances in 3D integration CMOS technology raise the possibility of hybrid photonic-electronic network-on-chip devices with 10x to 100x improvements in interconnect and DRAM bandwidth with a fixed power budget.

NERSC researchers have collaborated with the Lightwave Research Center (LRC) at Columbia University to study the energy efficiency and performance benefits of CMOS-integrated silicon photonic networking technology for 3D stacked devices. This research is the first comparison of photonic and electronic network designs using both synthetic benchmarks and real applications.

The NERSC scientists worked with the LRC to develop a cycle-accurate simulation of a CMOS-integrated network-on-chip composed of silicon photonic ring-resonator components, and to evaluate its performance using traces of communication from DOE scientific computing applications. The ring resonators act as pure optical switches, capable of reconfiguring optical networks with near zero energy cost.

The NERSC/LRC collaboration was able to demonstrate a 27x improvement in energy efficiency of the ring-resonator network-on-chip device. The work has led to a number of technology transfer discussions with IBM and a joint proposal with LBNL, Columbia, and Aprius Inc., a startup company seeking to develop optical DRAM interface technology.

NERSC helps manage and analyze LHC data

Over 15 million gigabytes of data per year need to be stored, processed, backed up, and distributed to researchers across the world now that the Large Hadron Collider (LHC) is smashing together beams of protons to search for new particles and forces, and beams of heavy nuclei to study new states of matter. Managing this mountain of data requires an international effort, with portions of the results sent to many supercomputing centers across the globe.

Large amounts of data from two of the LHC's detectors, ATLAS and ALICE (Figure 12), are being sent to NERSC, where two systems are being utilized to manage it. The High Performance Storage System (HPSS) archives the raw and processed data, while the Parallel Distributed Systems Facility (PDSF) processes and distributes results to thousands of scientists across the globe.

"You can build a world-class particle accelerator like the LHC, but if you don't have a way to store and analyze the data, it's all for nothing," says Peter Jacobs of the Lawrence Berkeley National Laboratory's Nuclear Science Division, who contributed to the construction of the ALICE detector, one of the four large experiments at the LHC. The LHC particle accelerator is located in Geneva, Switzerland, and is managed by the European Center for Nuclear Research (CERN).

PDSF comes full circle

The LHC is the most powerful particle accelerator in the world.

Capable of smashing together protons at an unprecedented 14 tera-electron volts of energy, it opens up a vast new landscape called the "terascale" for exploration. One exciting possibility is the discovery of the Higgs boson, a fundamental particle predicted to give mass to all matter in the Universe; another is the discovery of a new family of supersymmetric particles, which are predicted to exist by theories that unify all the forces of nature. The LHC is also the world's highest-energy collider of heavy nuclei, generating matter under the extreme conditions that existed a few microseconds after the Big Bang.

Terascale physics may be the new frontier, but the LHC is not the first collider designed to explore it. The U.S. had ambitions to investigate this realm of science in the 1980s, when its scientists began constructing the Superconducting Super Collider (SSC) in Waxahachie, Texas. In fact, the original PDSF was built to analyze SSC data. However, the project was cancelled in the mid-90s, and PDSF was transferred to NERSC, where it underwent multiple upgrades and expansions. Now it is one of the most flexible computing facilitates in the U.S., and has supported the majority of large nuclear and high energy physics projects undertaken by the country's leading scientists.

"PDSF is designed from the beginning to be able to support a wide range of nuclear science and high energy physics, from terascale physics accelerators, to experiments in the wastes of the Antarctic, to space experiments, in different ways. With the LHC, it comes full circle to support ultra-large collider experiments again," says Jay Srinivasan, PDSF system lead.

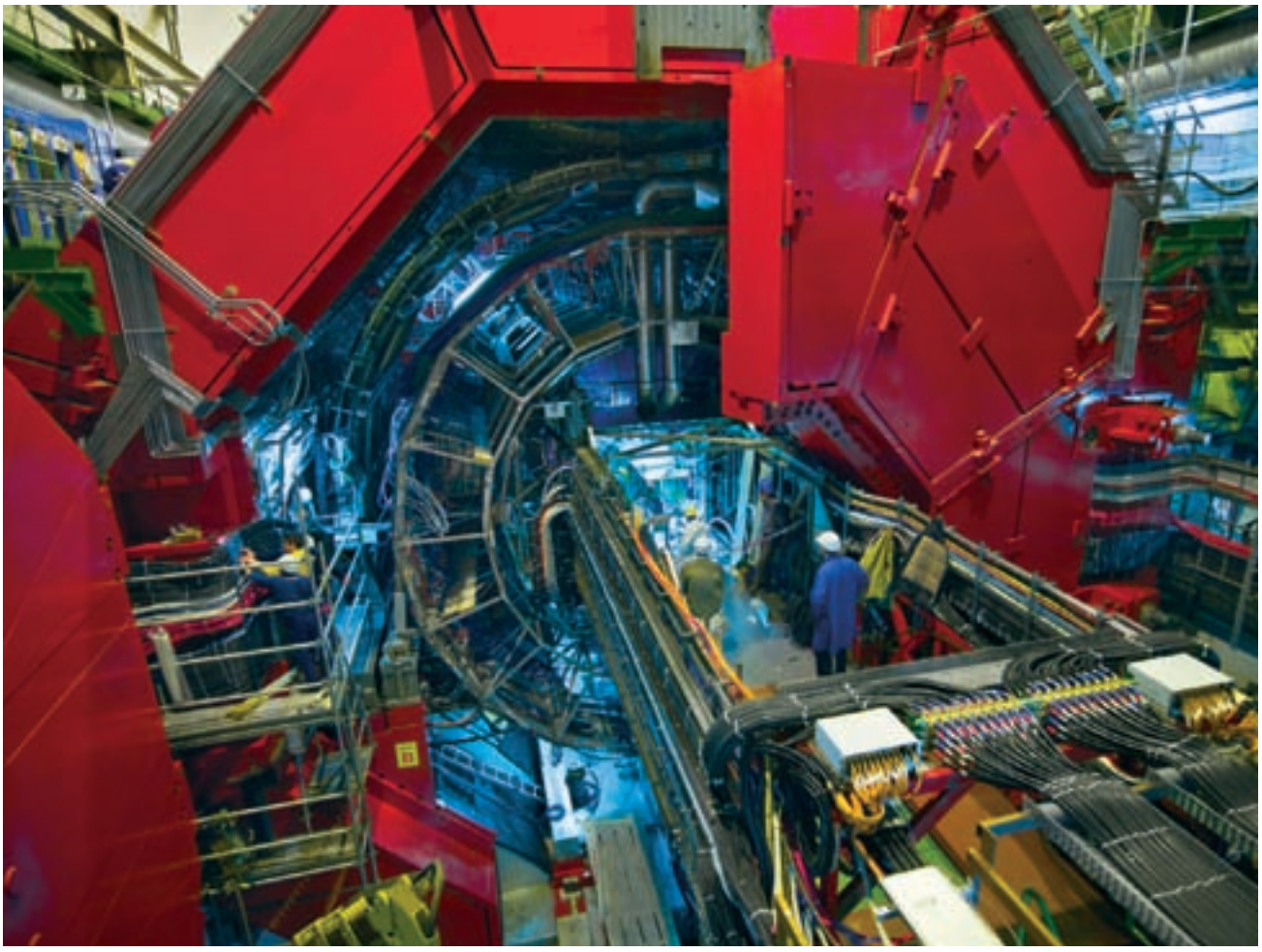


Figure 12. A magnetic door on the LHC's ALICE experiment closes, and an unprecedented opportunity to explore the realm of terascale physics opens.

According to Jacobs, the PDSF architecture is ideal for processing high energy physics data because the different nodes in the cluster do not really need to communicate with each other. Each particle collision is taken as a single event, and only one node is required to process that event. "All we really need is a large set of processors that can access tremendous amounts of data," he says.

"At NERSC we provide scientists the processing capabilities of PDSF, along with user support from consultants who are experts in computational science and performance tuning, visualization assistance, training, customized support, and other services," says Srinivasan. "And because PDSF uses commodity technology, or hardware that is in the marketplace, we offer a very cost efficient service."

PDSF also provides hundreds of terabytes of disk storage for the LHC experiments, as well as archival storage in HPSS. "Both PDSF and HPSS have processed and stored data from experiments similar to the LHC, so we know what to expect," says Srinivasan. "However, we are all really excited to be a part of a major project that will advance our knowledge of the world around us."

ALICE and ATLAS

ALICE is the only one of the four large LHC detectors optimized to study the very complex collisions of heavy nuclei. These collisions will occur inside the LHC for approximately four weeks every year, and will generate a fireball about 100,000 times hotter than the core of our Sun. Such temperatures have not existed in nature since a few millionths of a second after the Big Bang. At these temperatures, scientists expect to see an extended fluid of quarks and gluons, which are typically confined inside of subatomic protons and neutrons. This fluid state is called the quark-gluon plasma.

A group of U.S. institutions is collaborating with the international community to carry out research using ALICE. As part of this collaboration, the institutions are obligated to provide computing resources to the project in proportion to the fraction of authors on ALICE papers that they represent. PDSF is the main site for U.S. computing resources, and scientists suspect that by 2010 there will be about 400 processors at PDSF devoted to this experiment.

ATLAS is one of two general-purpose detectors that will investigate a wide range of physics. When particles collide inside the LHC, ATLAS records the birth of new particles formed by the smashup —their paths, energies, and identities. ATLAS plays an important role in two major searches: the search for the Higgs boson, a fundamental particle predicted to give mass to all matter in the Universe; and the search for a new family of super-symmetric particles, which are predicted to exist by theories that unify all the forces of nature.

In order to efficiently plan and execute experiments, scientists had to fully understand ATLAS detector capabilities beforehand, and they achieved this by running Monte Carlo simulations of the detector on PDSF. Berkeley Lab has been designated as an Analysis Support Center for the U.S. ATLAS project.

Historic Sudbury Neutrino Observatory data lives on at NERSC

Tunneled 6,800 feet underground in Canada's Vale Inco Creighton mine, the Sudbury Neutrino Observatory (SNO) was designed to detect neutrinos produced by fusion reactions in the Sun (Figure 13). Although the observatory officially ended operations in August 2006, a copy of all the data generated for and by the experiment will lives on at NERSC.

"NERSC has been providing great support to SNO for over a decade," says Alan Poon, a member of the SNO collaboration at Berkeley

Figure 13. By filling this 12-meter sphere at SNO with an unusual type of heavy water and surrounding it with light detectors, astrophysicists were able to collect enough data to solve the mystery of the missing solar neutrinos. That data is now archived at NERSC for future re-analyses.

Lab. "We used the PDSF cluster do some of the early analyses and were really appreciative of the support that we received from NERSC staff. When we looked around at different facilities and talked to colleagues that have used the center's High Performance Storage System [HPSS] extensively, we immediately concluded that one copy of our data should be stored at NERSC."

"The Department of Energy invested a lot of resources into SNO, and we believe that preserving these datasets at NERSC will afford the best protection of the agency's investment," he adds.

According to Poon, the SNO experiment made tremendous contributions to humanity's understanding of neutrinos, invisible elementary particles that permeate the cosmos. Before the observatory started searching for solar neutrinos on Earth, all experiments up to that point detected only a fraction of the particles predicted to exist by detailed theories of energy production in the Sun. Results from the SNO experiment eventually revealed that the total number of neutrinos produced in the Sun is just as predicted by solar models, but the neutrinos are oscillating in transit, changing in type or "flavor" from electron neutrinos (the flavor produced in the Sun) to muon or tau neutrinos. In 2001, Science magazine identified SNO's solution to the solar neutrino mystery as one of their ten science breakthroughs of the year.

"SNO data will be unique for decades to come," says Ryan Martin, a postdoctoral researcher at Berkeley Lab. "There will not be another experiment in the foreseeable future that would provide the same measurement with better precision and accuracy. It is important to preserve this data for the scientific community, in case a new theory would require further studies of the data."

Martin worked closely with Damian Hazen of NERSC's Storage Systems Group to transfer 26 terabytes of data from disks at the SNO facility in Sudbury, Canada, across the DOE's Energy Sciences Network (ESnet) to NERSC's HPSS. This transfer includes raw data generated by the experiment, processed data from various scientific analyses, as well as the computer codes and simulations used for detector design and data analysis.

"From testing the transfer speed, tuning the network and identifying packet losses, to the final archiving at HPSS, NERSC's expertise saved us a lot of headache," says Martin. "These are technical issues that laymen like us would take a long time to solve, if at all. We have been really pleased with the help that NERSC staff have provided."

Survey results lead to changes at NERSC

Responses to NERSC's user surveys provide feedback about every aspect of NERSC's operation, help us judge the quality of our services, give DOE information on how well NERSC is doing, and point us to areas we can improve. The 421 users who responded to the 2008/2009 survey represented 13% of authorized users and 77% of those who had used more than 250,000 processor hours.

On a scale of 7, the average score for overall satisfaction with NERSC was 6.21. Smaller users were the most satisfied, and larger users the least satisfied. Areas with the highest user satisfaction (> 6.5) were HPSS reliability and uptime, account and consulting support, grid job monitoring, NERSC Global Filesystem uptime and reliability, and network performance within the NERSC center. Areas with the lowest user satisfaction (4.7 to 4.9) were Bassi queue wait times and Franklin uptime.

Every year we institute changes based on the previous year's survey. In 2008 and 2009 NERSC took a number of actions in response to suggestions from the 2007/2008 user survey.

1. **Franklin disk configuration and I/O performance:**

   In 2009 NERSC and Cray staff worked extensively on benchmarking and profiling collective I/O performance on Franklin, conducting a detailed exploration into the source of the low performance (less than 1 GB/s write bandwidth) reported by several individual researchers.

   A number of issues were explored at various levels of the system/software stack, from the high-level NetCDF calls to MPI-IO optimizations and hints, block and buffer size allocations on individual nodes, Lustre striping parameters, and the underlying I/O hardware.

   These metrics were instrumental in making the case for increased I/O hardware and for making

software and configuration changes. Once implemented, the cumulative effect of the hardware, software and middleware improvements is that a class of applications is now able to achieve I/O bandwidths in the 6 GB/s range.

On the 2009 survey, Franklin's disk configuration and I/O performance received an average score of 5.60, a statistically significant increase over the previous year by 0.46 points.

2. **Franklin uptime:**

In 2009 NERSC and Cray assembled a team of about 20 people to thoroughly analyze system component layouts, cross interactions and settings; to review and analyze past causes of failures; and to propose and test software and hardware changes. Intense stabilization efforts took place between March and May, with improvements implemented throughout April and May.

As a result of these efforts, Franklin's overall availability went from an average of 87.6 percent in the six months prior to April to an average of 94.97 percent in the April through July 2009 period. In the same period, mean time between interrupts improved from an average of 1 day 22 hours 39 minutes to 3 days 20 hours 36 minutes.

The Franklin uptime score in the 2009 survey (which opened in May) did not reflect these improvements. NERSC anticipates an improved score on next year's survey.

3. **PDSF ability to run interactively, disk configuration and I/O performance:**

In 2008 NERSC improved the interactive PDSF nodes to more powerful, larger memory nodes. In early 2009, we reorganized the user file systems on PDSF to allow for failover, reducing the impact of hardware failures on the system. We also upgraded the network connectivity to the filesystem server nodes to allow for greater bandwidth. In addition, NERSC added a queue to allow for short debug jobs.

On the 2009 survey, the PDSF "Ability to run interactively" score increased significantly by 0.60 points and moved into the "mostly satisfied—high" range. The PDSF "Disk configuration and I/O performance" score increased by 0.41 points, but this increase was not statistically significant (at the 90 percent confidence level).

The question "What does NERSC do best? How does NERSC distinguish itself from other computing centers you have used?" elicited 130 responses. Here are a few representative comments:

NERSC's documentation is very good and the consultants are very helpful. A nice thing about NERSC is that they provide a number of machines of different scale with a relatively uniform environment which can be accessed from a global allocation. This gives NERSC a large degree of flexibility compared to other computational facilities.

NERSC has very reliable hardware, excellent administration, and a high throughput. Consultants there have helped me very much with projects and problems and responded with thoughtful messages for me and my problem, as opposed to terse or cryptic pointers to information elsewhere. The HPSS staff helped me set up one of the earliest data sharing archives in 1998, now part of a larger national effort toward Science Gateways.

The quality of the technical staff is outstanding. They are competent, professional, and they can answer questions ranging from the trivial tothe complex.

Getting users started! It can take months on other systems.

Complete survey results are available at http://www.nersc.gov/news/survey/.

Computational Science and Engineering Petascale Initiative

A multi-core revolution is occurring in computer chip technology. No longer able to

sustain the previous growth period when processor speed was continually increasing, chip manufacturers are instead producing multi-core architectures that pack increasing numbers of cores onto the chip. In high performance scientific computing, this revolution is forcing programmers to rethink the basic models of algorithm development, parallel programming, and programming languages.

To ensure that science effectively harnesses this new technology, NERSC is receiving more than $3 million in stimulus funds over the next two years from the American Recovery and Reinvestment Act (ARRA) to develop the Computational Science and Engineering Petascale Initiative. As part of this program, NERSC will hire several post-doctoral researchers to help design and modify modeling codes in key research areas such as energy technologies, fusion, and climate, to run on emerging multi-core systems.

"Emerging multi-core and other heterogeneous architectures provide an opportunity for major increases in computational power over the next several years," says Alice Koniges of NERSC's Advanced Technologies Group, who will be heading the initiative. "Designing effective programming models and new strategies will be the key to achieving next-generation petascale computing."

The introduction of multi-core chips as well as other new hardware designs such as those built on graphics processors, or GPUs, brings a new heterogeneity to the high performance computing community. While during the past 10–15 years, scientists were able to make amazing progress with parallel message passing models that allowed them to harness many thousands of processors, the new architectures have distinct differences in memory capacity and hierarchies that are causing researchers to consider new programming models and languages as a means to effectively exploit the power afforded by millions of cores.

Every year, NERSC provides computing resources to more than 3,000 DOE-supported scientists, who are developing new materials, modeling climate, investigating protein structures and conducting research in a host of other scientific endeavors. As the DOE Office of Science's primary scientific computing facility, the center is also tasked with helping these users adjust their codes to keep up with new trends in high performance computing. Currently, this includes overcoming the multi-core revolution challenges.

Hired on a two-year term assignment, each of the initiative's post-doctoral researchers will work with key NERSC users to create new procedures for making their current algorithms suitable for very large numbers of processors, to develop new algorithms to replace ones that do not scale, and to introduce new language constructs that are suitable for multi-core and/or GPU platforms. The researchers will also develop a framework for enabling a broader scientific community to take advantage of multi-core performance.

The projects selected to participate in this collaboration come from areas directly related to DOE's energy mission with an emphasis on projects that also receive funding from the ARRA stimulus package, those associated with EFRCs (Energy Frontier Research Centers), and projects centered on energy producing technologies such as fusion. Also included in the range of projects that will be considered are those that model national emergencies and projects that will help to broaden the community of researchers capable of using high performance computing facility resources.

"An integral part of this initiative is collaboration," says Koniges. "Because we are hosting this program at NERSC, the post-docs can leverage the expertise of scientists and engineers running a world-class supercomputing facility, as well as members of Berkeley Lab's Computational Research Division, who have extensive experience in creating computational tools and techniques for a wide range of science disciplines."

Green Flash project runs logical prototype successfully

The Green Flash project, which is exploring the feasibility of building a new class of energy-efficient supercomputers for climate modeling, has successfully reached its first milestone by running the dynamical core of the Global Cloud Resolving Model (GCRM, see page 34) on logical prototypes of both single- and dual-core Green Flash processors, with eight-core processors coming soon.

"The logical prototype simulates the entire circuit design of the proposed processor," says

John Shalf, head of NERSC's Advanced Technologies Group and principal investigator of Green Flash.

The prototype was designed in collaboration with Tensilica, Inc., using Tensilica's Xtensa LX-2 extensible processor core as the basic building block, and has been running cycle accurate hardware emulations of the circuit design on a BEE3 FPGA platform, which is used for computer architecture research by the RAMP Consortium (Research Accelerator for Multi-Processors). A next-generation, limited area model version of GCRM has been used as the test code.

David Donofrio of the Advanced Technologies Group, who works on the hardware design of Green Flash, ran the first single-core prototype in a demonstration at the SC08 conference in Austin, Texas in November 2008. This was followed by a multiprocessor demo at the SC09 conference. Green Flash was first proposed publicly in the paper "Towards Ultra-High Resolution Models of Climate and Weather," written by Michael Wehner and Lenny Oliker of Berkeley Lab's Computational Research Division (CRD) and Shalf of NERSC.[1]

One solution for three problems

The Green Flash project addresses three research problems simultaneously—a climate science problem, a computer architecture/hardware problem, and a software problem. This multidisciplinary development process is commonly referred to as hardware/software co-design.

The climate science problem stems from the resolution of current climate models, which is too coarse to directly calculate the behavior of cumulus convective cloud systems (see "Bringing Clouds into Focus," page 34). Direct numerical simulation of individual cloud systems would require horizontal grid resolutions approaching 1 km. To develop a 1 km cloud model, scientists would need a supercomputer that is 1,000 times more powerful than what is available today.

But building a supercomputer that powerful with conventional microprocessors (the kind used to build personal computers) would cost about $1 billion and would require 200 megawatts of electricity to operate—enough energy to power a small city of 100,000 residents. That constitutes the computer architecture problem. In fact, the energy consumption of conventional computers is now recognized as a major problem not just for climate science, but for all of computing, from cell phones to the largest scale systems.

Shalf, Wehner, and Oliker see a possible solution to these challenges—achieving high performance with a limited power budget and with economic viability—in the low-power embedded microprocessors found in cell phones, iPods, and other electronic devices. Unlike the general-purpose processors found in personal computers and most supercomputers, where versatility comes at a high cost in power consumption and heat generation, embedded processors are designed to perform only what is required for specific applications, so their power needs are much lower. The embedded processor market also offers a robust set of design tools and a well-established economic model for developing application-specific integrated circuits (ASICs) that achieve power efficiency by tailoring the design to the requirements of the application. Chuck McParland of CRD has been examining issues of manufacturability and cost projections for the Green Flash design to demonstrate the cost-effectiveness of this approach.

Meeting the performance target for the climate model using this technology approach will require on the order of 20 million processors. Conventional approaches to programming are unable to scale to such massive concurrency. The software problem addressed by the Green Flash project involves developing new programming models that are designed with million-way concurrency in mind, and exploiting auto-tuning technology to automate the optimization of the software design to operate efficiently on such a massively parallel system.

To meet this challenge, Tony Drummond of CRD and Norm Miller of the Earth Sciences Division are working on analyzing the code

[1] M. Wehner, L. Oliker, and J. Shalf, "Towards Ultra-High Resolution Models of Climate and Weather," International Journal of High Performance Computing Applications **22**, 149 (2008).

requirements; and Shoaib Kamil, a graduate student in computer science at the University of California, Berkeley (UCB) who is working at NERSC, has been developing an auto-tuning framework for the climate code. This framework automatically extracts sections of the Fortran source code of the climate model and optimizes them for Green Flash and a variety of other architectures, including multicore processors and graphics processors.

An innovative aspect of the Green Flash research is the hardware/software co-design process, in which early versions of both the processor design and the application code are developed and tested simultaneously. The RAMP emulation platform allows scientists to run the climate code on different hardware configurations and evaluate those designs while they are still on the drawing board. Members of the RAMP consortium on the UC Berkeley campus, including John Wawrzyneck and Krste Asanovic (both of whom have joint appointments at NERSC), Greg Gibling, and Dan Burke, have been working closely with David Donofrio of NERSC and the Green Flash hardware team throughout the development process. A RAMP test at UCB has successfully emulated more than 1,000 cores.

At the same time, auto-tuning tools for code generation test different software implementations on each hardware configuration to increase performance, scalability, and power efficiency. Marghoob Mohiyuddin, another UCB graduate student at NERSC, has been working on automating the hardware/software co-design process. With a dual-core processor configuration on the RAMP emulator, Mohiyuddin can test more than 200 configurations in one day, which is 125 times faster than conventional approaches to design space exploration. The result will be a combination of hardware and software optimized to solve the cloud modeling problem.

The researchers estimate that the proposed Green Flash supercomputer, using about 20 million embedded microprocessors, would deliver the 1 km cloud model results and cost perhaps $75 million to construct (a more precise figure is one of the project goals). This computer would consume less than 4 megawatts of power and achieve a peak performance of 200 petaflops.

Maximizing efficiency

Following the design philosophy that the best way to reduce power consumption and increase efficiency is to reduce waste, the Green Flash team chose an architecture with a very simple in-order core and no branch prediction. Because the climate model's demands for memory and communication are high, both aspects drive the core design, which includes a local store to maximize use of the available dynamic RAM (DRAM) bandwidth.

As Figure 14 shows, the design uses a torus network fabric with

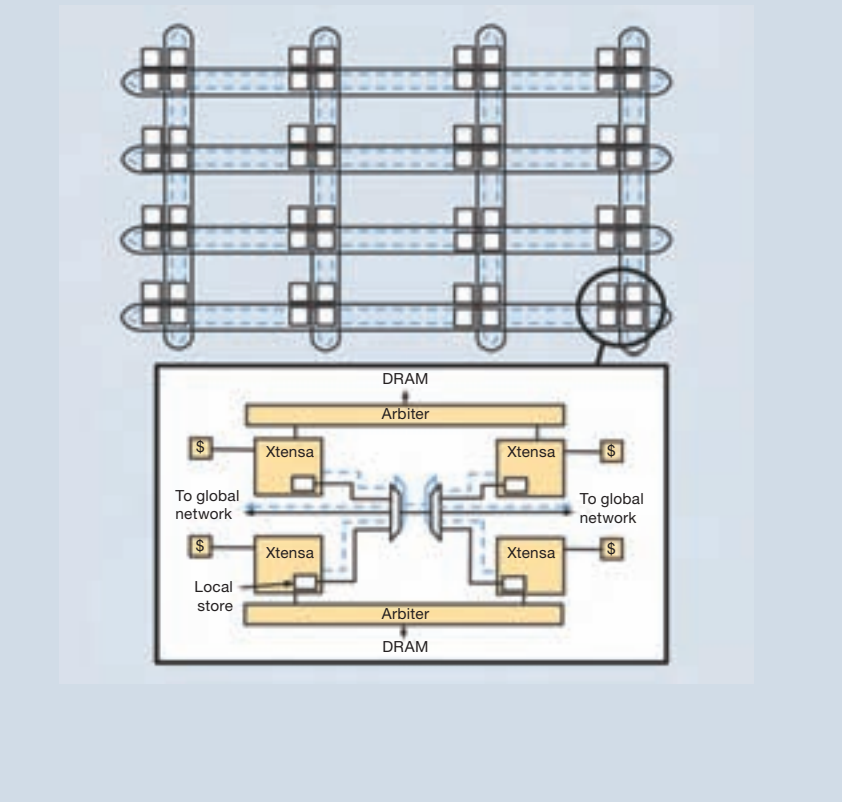


Figure 14. The on-chip network fabric for the Green Flash system-on-chip. A concentrated torus network fabric yields the highest performance and most power-efficient design for scientific codes.

two on-chip networks. Most of the communication among the climate model's subdomains is nearest neighbor, and experiments showed that a concentrated torus topology provides superior performance and energy efficiency for codes in which a nearest-neighbor communication pattern dominates. The researchers are currently targeting a core with a clock speed of 500 MHz, a 32-KB conventional error correction code (ECC)-protected cache per core, and a 128-KB local store. The availability of a conventional cache will allow code to be incrementally ported to use the local store. Each socket of 128 cores will have a 50-GB-per-second interface to DRAM.

Achieving the target execution rate on 20 million processors requires computing on a local mesh size that is 8 × 8 × 10 cells. If the code were to run on conventional cache-based hardware, it would spend 90 percent of its time in communication due to the overhead penalty of exchanging extremely small messages between cores. But Green Flash has added specialized hardware to each core to enable extremely low-overhead messaging between cores, bringing the communication overhead below 20 percent of the total execution time. This ultra-low-overhead streaming interface bypasses the cache to minimize latency and connects to one of the on-chip torus networks. The narrow network is for address exchange; the wider torus network is for bulk data exchange using asynchronous direct memory access (DMA) data transfers. The address space for each processor's local store is mapped into the global address space, and the data exchange is done as a DMA from local store to local store.

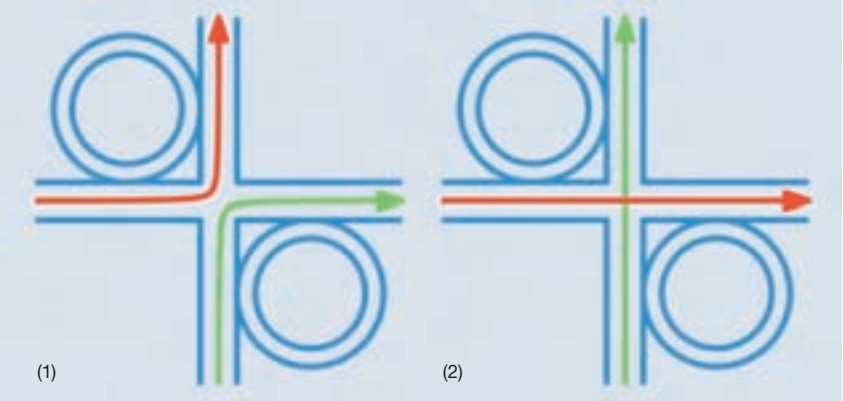


Figure 15. Photonic switching elements. (1) Light is coupled onto a perpendicular path; (2) messages propagate straight through. The lack of distance and complex structures are strong advantages over a purely electrical interconnect.

From a logical programming view, all processors are directly connected to each other, but physically they are connected using a concentrated torus network to the chip's 2D planar geometry. To further simplify programming, a traditional cache hierarchy is also in place to allow the slow porting of codes to the more efficient interprocessor network.

To minimize power, the researchers are investigating the use of hybrid electronic-photonic interconnects for the inter-core network, which could prove to be an efficient way of transferring long messages. Designers place photonic detectors and emitters along with specialized low-power photonic switching elements on a special interconnect layer and interface them with processing elements using conventional electronic routers. Figure 15 shows how the switching elements work. Large-scale communications occur over photonic links, which have several strong advantages over electronic networks. Energy consumption for photonics is less dependent on signaling rate and distance compared to electronics, and the photonic switches are much simpler, as they do not require buffers or repeaters. Preliminary research with messaging patterns from scientific applications shows that such hybrid networks have the potential to bring major gains in efficiency, due to their lower power consumption combined with fast propagation speed. Early research studies done in collaboration with the Lightwave Research Laboratory at Columbia University, for example, show that a hybrid electronic-photonic interconnect composed of ring resonators can deliver 27 times better energy efficiency than electrical interconnects alone.

To optimize the code and reduce the computational burden, the Green Flash team created an autotuning framework that automatically searches a range of optimizations to improve the

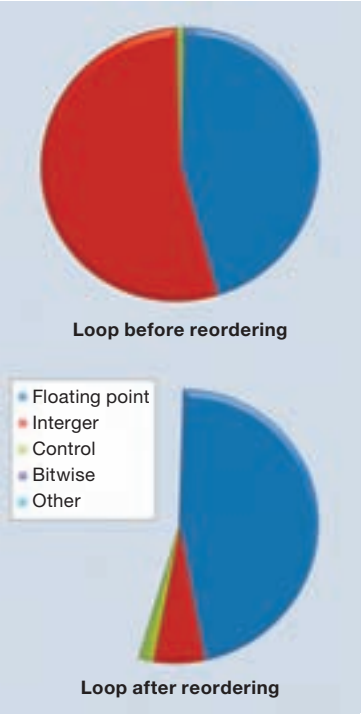


Figure 16. Effect of optimization on a single loop in the climate model. In addition to greatly reducing the instruction count, optimization reduced the cache footprint of this loop by more than 100 times. With software tuning, Green Flash can reduce a per-core computational requirement of 3.5 gigaflops to a more feasible 0.5 gigaflops.

application kernels' computational efficiency. The autotuner first systematically applies compiler optimizations and then uses domain-specific knowledge of the algorithm to take more aggressive steps, such as loop reordering, to produce optimal but functionally equivalent code. In this way, it maintains performance across a diverse set of architectures.

Figure 16 shows the autotuning results for the climate model. The researchers ran the autotuning framework using the Tensilica architectural simulator, reducing the cache footprint and overall instruction count and increasing the kernel's computational density. They first generated the original requirement of 3.5 gigaflops per core using a machine that ran with approximately 5 percent efficiency. Autotuners, combined with hardware optimizations, will play a key role in dramatically increasing the efficiency of Green Flash. Through these combined optimizations, Green Flash is expected to realize a two-orders-of-magnitude increase in efficiency.

The hardware-software codesign method tailors the hardware to autotuned software to get better energy efficiency. The autotuning technology can automate the exploration for the optimal combination of tuned software and hardware in a coordinated design cycle. As Figure 17 shows, this cotuning approach incorporates extensive software tuning into the hardware design process. The autotuned software tailors the application to the hardware design point under consideration by empirically searching software implementations to find the best mapping of software to microarchitecture.

As a demonstration of this cotuning methodology, the Green Flash team used the Smart Memories multiprocessor (based on Tensilica cores) as the target architecture and three widely used kernels from scientific computing: dense matrix-matrix multiplication, stencil codes, and sparse matrix vector multiplication. As part of exploring the hardware design space, they varied four hardware parameters: number of cores, whether caches are managed by hardware or software, cache size per core, and total memory bandwidth available. They estimated the area and power of each hardware configuration that had the corresponding best software configuration, which they obtained through autotuning. As Figure 18 shows, power and area efficiencies improved dramatically for the three kernels.

The hardware-software codesign process enables scientific application developers to directly participate in the design process for future supercomputers in an unprecedented way. With this fast, accurate emulation environment, designers can run and benchmark the actual climate model as it is being developed and use cotuning to quickly search a large design space.

Scaling up

In considering any system of this scale, a myriad of system software issues come to the forefront, such as scalable operating systems, fault resilience infrastructure, and the development of entirely new programming models to make billion-way parallelism more tractable.

Although the fault resilience problem is certainly not trivial, neither is it unusual. Across silicon design processes with the same design rules, hard failure rates are proportional to the number of system sockets and typically stem from mechanical failures. Soft error rates are proportional to the chip surface area, not how many cores are on a chip. And bit error rates tend to increase with clock rate. The Green Flash architecture is

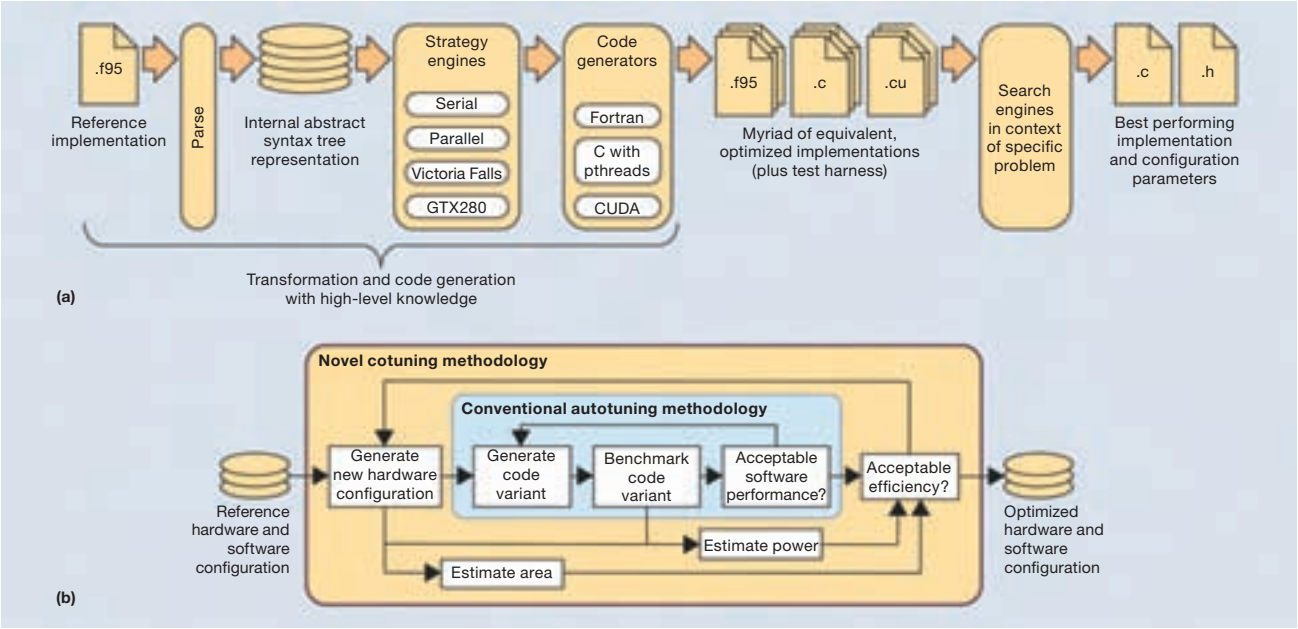


Figure 17. Cotuning in the Green Flash design. (a) Conventional autotuning uses source code generators and search heuristics to empirically choose an efficient software implementation given a high-level representation of a kernel. (b) Hardware-software cotuning extends conventional hardware design space exploration by using autotuning to tailor software to each hardware design point.

unremarkable in all these respects and should not pose challenges beyond those that a conventional approach faces.

To deal with hard errors, designers often add redundant cores per chip to cover defects, a strategy that is entirely feasible for the Green Flash design. Moreover, Green Flash's low power dissipation per chip (7 to 15 W) will reduce the mechanical and thermal stresses that often result in a hard error. To address soft errors, the design includes all the basics for reliability and error recovery in the memory subsystem, including full ECC protection for all hierarchical levels. Green Flash's low target clock frequency provides a lower signal-to-noise ratio for on-chip data transfers. Finally, to enable faster rollback if an error does occur, the design makes it possible to incorporate a nonvolatile RAM controller onto each SMP so that each node can perform a local rollback as needed. This strategy enables much faster rollback than user-space checkpointing.

The researchers are also exploring novel programming models together with hardware support to express fine-grained parallelism. The goal of this development thrust is to create a new software model that can provide a stable platform for software development for the next decade and beyond for all scales of scientific computing.

They have developed direct hardware support for both the message passing interface (MPI) and partitioned global address space (PGAS) programming models to enable scaling of these familiar single program, multiple data (SPMD) programming styles to much larger-scale systems. The modest hardware support enables relatively well-known programming paradigms to utilize massive on-chip concurrency and to use hierarchical parallelism to enable use of larger messages for interchip communication. GCRM's icosahedral formulation of the climate problem can expose a massive degree of parallelism through domain decomposition, which can use a 20-million processor computing system. The autotuning framework is rapidly evolving into a generalized code generator, which allows the programmer to express the solver kernels at a much higher level of abstraction—enabling a productive programming environment that supports portability, performance, and correctness without exposing scientists to the details of the computer architecture. This approach is expected to support

a broad range of codes that have such inherent explicit parallelism.

However, not all applications will be able to express parallelism through simple divide-and-conquer problem partitioning. The Green Flash team is beginning to explore new asymmetric and asynchronous approaches to achieving strong-scaling performance improvements from explicit parallelism. Techniques that resemble class static dataflow methods are garnering renewed interest because of their ability to flexibly schedule work and to accommodate state migration to correct load imbalances and failures.

In the case of the GCRM climate code, dataflow techniques can be used to concurrently schedule the physics computations with the dynamic core of the climate code, thereby doubling concurrency without moving to a finer domain decomposition. This approach also benefits from the unique interprocessor communication interfaces developed for Green Flash. Successful demonstration of the new parallelization procedure for a range of leading extreme-scale applications can then be utilized by other similar codes, accelerating development efforts for the entire field.

Designs that follow the Green Flash approach have the potential to open a market demand for massively concurrent components that can also be the building blocks for mid- and extreme-scale computing systems. "We believe that our decision to draw from the embedded computing industry will produce technology that reduces economic and manufacturing barriers to constructing computing systems useful to science," Shalf says. "It will also ensure that selected technologies have broad market impact for everything from the smallest handheld to the largest supercomputer. The investment will thus be the center of a sustainable software-hardware universe supported by applications across the IT industry."

The Green Flash prototype research is funded by Berkeley Lab's Laboratory Directed Research and Development program.

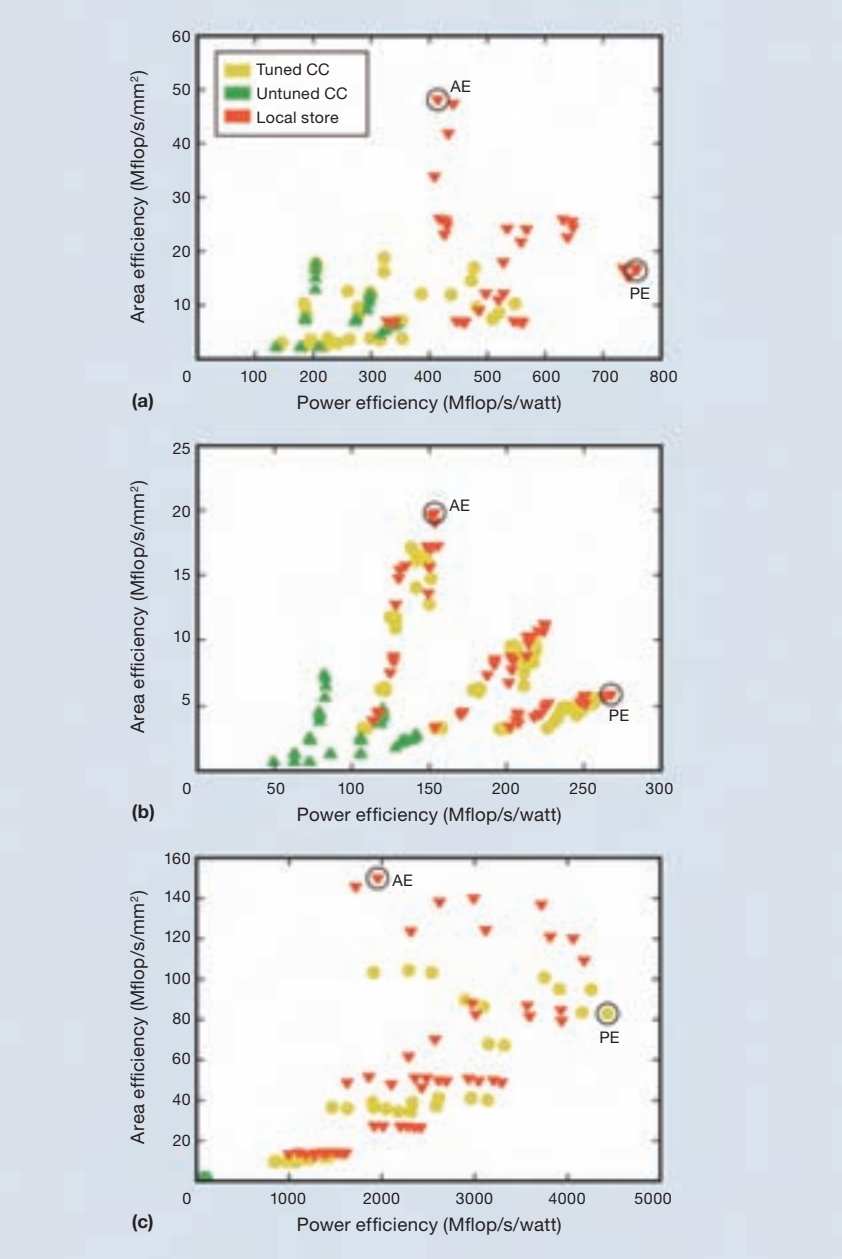


Figure 18. The advantages of cotuning for three kernel types common in scientific applications. AE and PE points denote configurations with highest area and power efficiencies. Improvements varied from 2x to 50x.

Appendix A

NERSC Policy Board

The NERSC Policy Board meets at least annually and provides scientific and executive-level advice to the Berkeley Lab Director regarding the overall NERSC program and, specifically, on such issues as resource utilization to maximize the present and future scientific impact of NERSC, and long-range planning for the program, including the research and development necessary for future capabilities. Policy Board members are widely respected leaders in science, computing technology, or the management of scientific research and/or facilities.

Daniel A. Reed, Policy Board Chair
Corporate Vice President, Extreme Computing Group
Microsoft Corporation

Achim Bachem
Chairman of the Board of Directors
Forschungszentrum Jülich GmH

Jack Dongarra
University Distinguished Professor
Director, Innovative Computing Laboratory
University of Tennessee

Stephane Ethier
Ex officio, NERSC Users Group Chair
Computational Plasma Physics Group
Princeton Plasma Physics Laboratory

Sidney Karin
Professor of Computer Science and Engineering
University of California, San Diego

Pier Oddone
Director
Fermi National Accelerator Laboratory

Edward Seidel
Acting Assistant Director for Mathematics and Physical Sciences
National Science Foundation

Appendix B

NERSC Client Statistics

In support of the DOE Office of Science's mission, NERSC served 3,271 scientists throughout the United States in 2008 and 3,731 scientists in 2009. These researchers work in DOE laboratories, universities, industry, and other Federal agencies. Figure 1 shows the proportion of NERSC usage by each type of institution, while Figures 2 and 3 show laboratory, university, and other organizations that used large allocations of computer time. Computational science conducted at NERSC covers the entire range of scientific disciplines, but is focused on research that supports the DOE's mission and scientific goals, as shown in Figure 4. User demographics and usage for previous years can be found at http://www.nersc.gov/about/users.php.

In their annual allocation renewals, users reported 1,464 refereed publications (accepted or submitted) enabled at least in part by NERSC resources in 2008, and 1,646 publications in 2009. Lists of publications resulting from use of NERSC resources are available at http://www.nersc.gov/projects/reports/.

The MPP hours reported here are Cray XT4 equivalent hours.

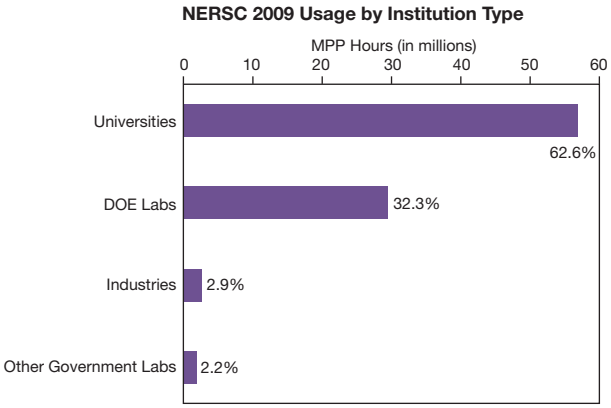


Figure 1. NERSC MPP usage by institution type, 2009.

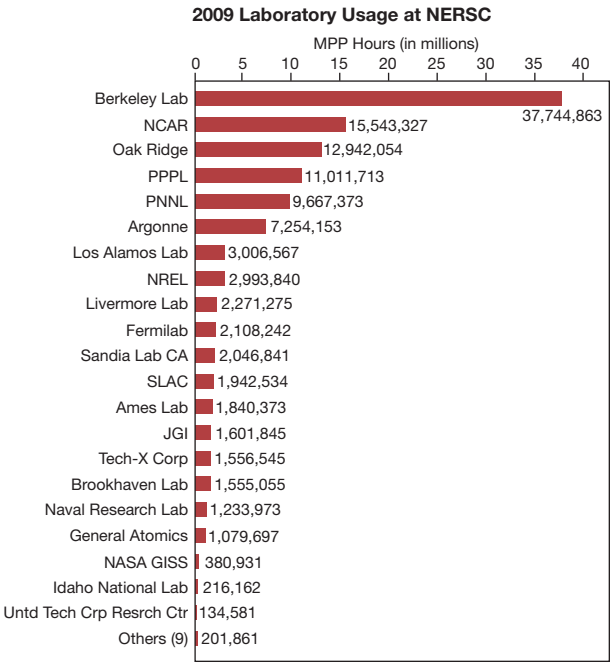


Figure 2. DOE and other laboratory usage at NERSC, 2009 (MPP hours).

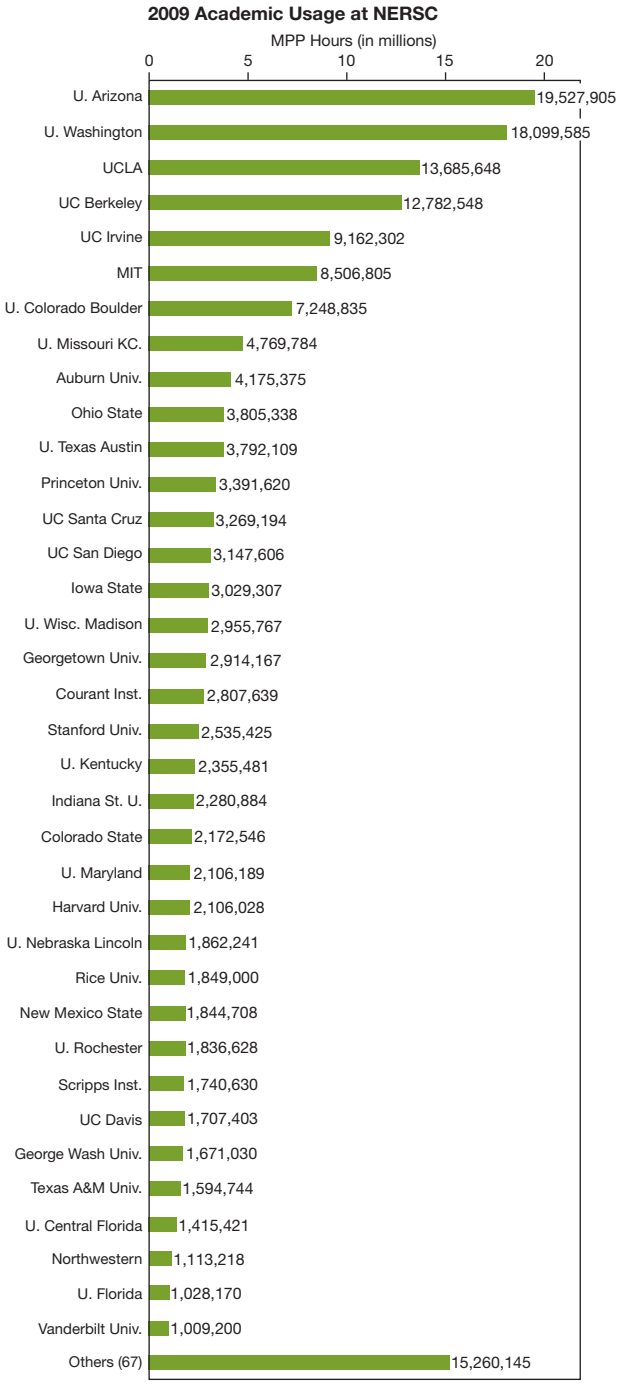


Figure 3. Academic and usage at NERSC, 2009 (MPP hours).

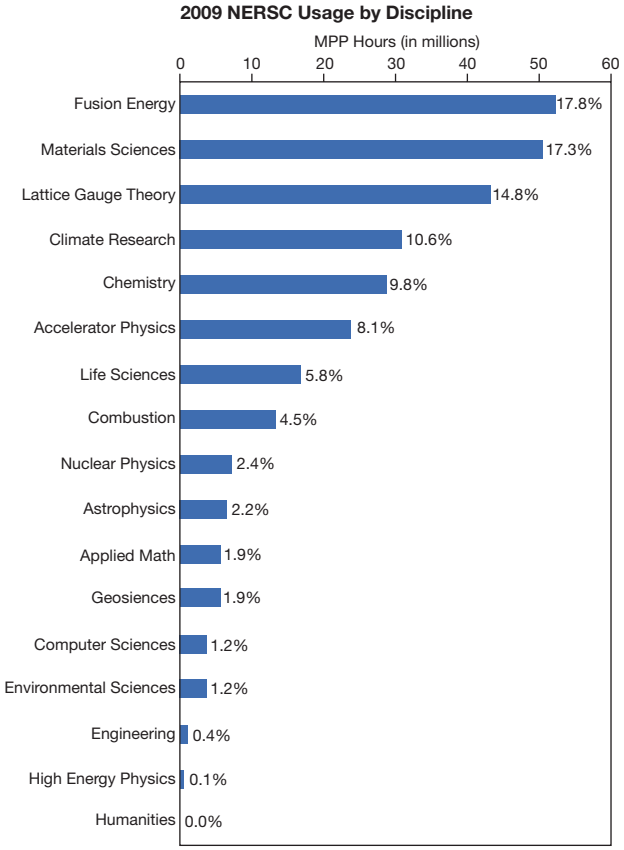


Figure 4. NERSC usage by scientific discipline, 2009

Appendix C

NERSC Users Group Executive Committee

Office of Advanced Scientific Computing Research

Mark Adams, Columbia University
Sergei Chumakov, Stanford University
Mike Lijewski, Lawrence Berkeley National Laboratory

Office of Basic Energy Sciences

Eric Bylaska, Pacific Northwest National Laboratory
Paul Kent, Oak Ridge National Laboratory
Brian Moritz, SLAC National Accelerator Laboratory

Office of Biological and Environmental Research

David Beck, University of Washington
Brian Hingerty, Oak Ridge National Laboratory
Adrianne Middleton, National Center for Atmospheric Research

Office of Fusion Energy Sciences

Stephane Ethier (Chair), Princeton Plasma Physics Laboratory
Linda Sugiyama, Massachusetts Institute of Technology
Jean-Luc Vay, Lawrence Berkeley National Laboratory

Office of High Energy Physics

Julian Borrill, Lawrence Berkeley National Laboratory
Cameron Geddes, Lawrence Berkeley National Laboratory
Frank Tsung (Vice Chair), University of California, Los Angeles

Office of Nuclear Physics

David Bruhwiler, Tech-X Corporation
Peter Messmer, Tech-X Corporation
Tomasz Plewa, Florida State University

Members at Large

Angus Macnab, DyNuSim, LLC
Jason Nordhaus, Princeton University
Ned Patton, National Center for Atmospheric Research
Maxim Umansky, Lawrence Livermore National Laboratory

Appendix D

Office of Advanced Scientific Computing Research

The primary mission of the Advanced Scientific Computing Research (ASCR) program is to discover, develop, and deploy the computational and networking tools that enable researchers in the scientific disciplines to analyze, model, simulate, and predict complex phenomena important to the Department of Energy. To accomplish this mission, the program fosters and supports fundamental research in advanced scientific computing—applied mathematics, computer science, and networking—and operates supercomputer, networking, and related facilities. In fulfilling this primary mission, the ASCR program supports the Office of Science Strategic Plan's goal of providing extraordinary tools for extraordinary science as well as building the foundation for the research in support of the other goals of the strategic plan. In the course of accomplishing this mission, the research programs of ASCR have played a critical role in the evolution of high performance computing and networks. Berkeley Lab thanks the program managers with direct responsibility for the NERSC program and the research projects described in this report:

Michael R. Strayer
Associate Director, ASCR

Barbara Helland
Senior Advisor

Walter M. Polansky
Cyber Security Advisor

Robert Lindsay
Special Projects

Julie Scott
Financial Management Specialist

Lori Jerigan
Program Support Specialist

Melea Baker
Administrative Specialist

Facilities Division

Michael R. Strayer
Acting Division Director

Daniel Hitchcock
Computer Scientist

Sally McPherson
Program Assistant

Barbara Helland
Computer Scientist

Robert Lindsay
Computer Scientist

Yukiko Sekine
Computer Scientist

Vincent Dattoria
Computer Scientist

Betsy Riley
Computer Scientist

Computational Science Research and Partnerships (SciDAC) Division

Walter M. Polansky
Acting Division Director

Teresa Beachley
Program Assistant

Daniel Hitchcock
Computer Scientist

Christine Chalk
Physical Scientist

Barbara Helland
Computer Scientist

Thomas Ndousse-Fetter
Program Manager

Steven Lee
Mathematician, Detailee

Susan Turnbull
Detailee

Bill Spotz
Mathematician, IPA

Sandy Landsberg
Mathematician

Osni Marques
Computer Scientist, IPA

Karen Pao
Mathematician

Lucy Nowell
Computer Scientist

Rich Carlson
Computer Scientist

Sonia Sachs
Computer Scientist

Appendix E

Advanced Scientific Computing Advisory Committee

The Advanced Scientific Computing Advisory Committee (ASCAC) provides valuable, independent advice to the Department of Energy on a variety of complex scientific and technical issues related to its Advanced Scientific Computing Research program. ASCAC's recommendations include advice on long-range plans, priorities, and strategies to address more effectively the scientific aspects of advanced scientific computing including the relationship of advanced scientific computing to other scientific disciplines, and maintaining appropriate balance among elements of the program. The Committee formally reports to the Director, Office of Science. The Committee primarily includes representatives of universities, national laboratories, and industries involved in advanced computing research. Particular attention is paid to obtaining a diverse membership with a balance among scientific disciplines, institutions, and geographic regions.

Roscoe C. Giles, Chair
Boston University

F. Ronald Bailey
Computer Sciences Corporation

Marsha Berger
Courant Institute of Mathematical Sciences

Jackie Chen
Sandia National Laboratories

Jack J. Dongarra
University of Tennessee

Susan L. Graham
University of California, Berkeley

James J. Hack
Oak Ridge National Laboratory

Anthony Hey
Microsoft Research

Thomas A. Manteuffel
University of Colorado at Boulder

John Negele
Massachusetts Institute of Technology

Linda R. Petzold
University of California, Santa Barbara

Vivek Sarkar
Rice University

Larry Smarr
University of California, San Diego

William M. Tang
Princeton Plasma Physics Laboratory

Robert G. Voigt
SAIC

Victoria White
Fermi National Accelerator Laboratory

Appendix F

Acronyms and Abbreviations

ACM	Association for Computing Machinery
ALCF	Argonne Leadership Computing Facility
API	Application program interface
AR4	Intergovernmental Panel on Climate Change Fourth Assessment Report
AR5	Intergovernmental Panel on Climate Change Fifth Assessment Report
ARRA	American Recovery and Reinvestment Act
ART	Adaptive Refinement Tree
ASCR	Office of Advanced Scientific Computing Research (DOE)
ASIC	Application-specific integrated circuit
BER	Office of Biological and Environmental Research (DOE)
BES	Office of Basic Energy Sciences (DOE)
BNL	Brookhaven National Laboratory
BSC-CNS	Barcelona Supercomputing Center - Centro Nacional de Supercomputación
CAF	Co-Array Fortran
CCD	Charge-coupled device
CCSE	Center for Computational Sciences and Engineering (LBNL)
CCSM	Community Climate System Model
CERN	European Organization for Nuclear Research
CIRES	Cooperative Institute for Research in Environmental Sciences
CMOS	Complementary metal-oxide semiconductor
CPU	Central processing unit
CRD	Computational Research Division, Lawrence Berkeley National Laboratory
CSU	Colorado State University
CSWPI	Center for Simulation of Wave-Plasma Interactions (SciDAC)
DARPA	Defense Advanced Research Projects Agency
DFT	Density functional theory
DIP	Deutsch-Israelische Projektkooperation
DMA	Direct memory access
DOE	U.S. Department of Energy
DRAM	Dynamic random access memory
DVS	Data Virtualization Service
ECC	Error correction code
ECHU	Einstein Center at Hebrew University, Jerusalem
EFRC	Energy Frontier Research Center
EMSL	Environmental Molecular Sciences Laboratory (PNNL)

ERC	European Research Council
ERCAP	Energy Research Computing Allocations Process
ESnet	Energy Sciences Network
FCG	Fundação Calouste Gulbenkian (Portugal)
FCT	Fundação para a Ciência e a Tecnologia (Portugal)
FES	Office of Fusion Energy Sciences (DOE)
FITS	France–Israel Teamwork in Sciences
FNAL	Fermi National Accelerator Facility
FPGA	Field-programmable gate array
GB	Gigabyte
GCRM	Global cloud resolving model
GIF	German-Israeli Foundation for Scientific Research and Development
GPFS	General Parallel File System
GPU	Graphics processing unit
HDF	Hierarchical Data Format
HEP	Office of High Energy Physics (DOE)
HHFW	High harmonic fast wave
HPC	High performance computing
HPCS	High Performance Computing Service (University of Cambridge)
HPSS	High Performance Storage System
ICRF	Ion-cyclotron range of frequencies
IEEE	Institute of Electrical and Electronics Engineers
IGCC	Integrated gasification combined cycle
INCITE	Innovative and Novel Computational Impact on Theory and Experiment (DOE)
I/O	Input/output
IPA	Intergovernmental Personnel Act
IPCC	Intergovernmental Panel on Climate Change
IPM	Integrated Performance Monitoring
ISF	Israel Science Foundation
IST	Instituto Superior Técnico (Portugal)
ITER	A multinational tokamak experiment in France (Latin for “the way”)
JLab	Thomas Jefferson National Accelerator Facility
JMT	Jeffress Memorial Trust
KB	Kilobyte
LBNL	Lawrence Berkeley National Laboratory
LH	Lower hybrid

LHC	Large Hadron Collider (CERN)
LHCD	Lower hybrid current drive
LHRF	Lower hybrid radio frequency
LLNL	Lawrence Livermore National Laboratory
LQCD	Lattice quantum chromodynamics
LRC	Lightwave Research Center (Columbia University)
LS3DF	Linear Scaling Three-Dimensional Fragment
LSB	Low-swirl burner
MFP	Minerva Fellowship Program
MHD	Magnetohydrodynamic
MHz	Megahertz
MPI	Message Passing Interface
NAD	Nicotinamide adenine dinucleotide
NADH	The reduced form of NAD in electron transport reactions
NAS	NASA Advanced Supercomputing Division of NASA Ames Research Center
NASA	National Aeronautics and Space Administration
NCAR	National Center for Atmospheric Research
NCHC	National Center for High-Performance Computing (Taiwan)
NCSA	National Center for Supercomputing Applications
NERSC	National Energy Research Scientific Computing Center
NGF	NERSC Global Filesystem
NIM	NERSC Information Management system
NISE	NERSC Initiative for Scientific Exploration
NOAA	National Oceanographic and Atmospheric Administration
NP	Office of Nuclear Physics (DOE)
NSF	National Science Foundation
NSTX	National Spherical Tokamak Experiment
NUG	NERSC Users Group
NYSTAR	New York State Office of Science, Technology and Academic Research
OJI	Outstanding Junior Investigator Program (DOE)
OLCF	Oak Ridge Leadership Computing Facility
OpenGL	Open Graphics Library
ORNL	Oak Ridge National Laboratory
PB	Petabyte
PDSF	Parallel Distributed Systems Facility (NERSC)
PERI	Performance Engineering Research Institute (SciDAC)

Petaflops	Quadrillions of floating point operations per second
Pflops	Petaflops
PGAS	Partitioned global address space
PI	Principal investigator
PIC	Particle-in-cell
PNNL	Pacific Northwest National Laboratory
QCD	Quantum chromodynamics
RAM	Random access memory
RAMP	Research Accelerator for Multiple Processors
RHIC	Relativistic Heavy Ion Collider (BNL)
RF	Radio frequency
RPMD	Ring polymer molecular dynamics
SAP	Scientific Application Partnership (SciDAC)
SC	Office of Science (DOE)
SciDAC	Scientific Discovery through Advanced Computing (DOE)
SFG	Star-forming galaxy
SIAM	Society for Industrial and Applied Mathematics
SMP	Symmetric multiprocessor
SNO	Sudbury Neutrino Observatory
SPMD	Single program, multiple data
SSC	Superconducting Super Collider
SSH	Secure Shell
STM	Scanning tunneling microscope
TB	Terabyte
Teraflops	Trillions of floating point operations per second
TNSC	Taiwanese National Science Council
UC	University of California
UCB	University of California, Berkeley
UCLA	University of California, Los Angeles
UCSD	University of California, San Diego
UPC	Unified Parallel C
UW	University of Washington
VACET	Visualization and Analytics Center for Enabling Technologies (SciDAC)
WAN	Wide-area network

For more information about NERSC, contact:

Jon Bashor
NERSC Communications
Berkeley Lab, MS 50B4230
1 Cyclotron Road
Berkeley, CA 94720-8148

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

NERSC's web site:
http://www.nersc.gov/

NERSC Annual Report Editor: John Hules.

Contributing writers: John Hules, Jon Bashor, Linda Vu, and Margie Wylie, Berkeley Lab Computing Sciences Communications Group; Aditi Risbud, Molecular Foundry; Allen Chen, Environmental Energy Technologies Division.

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

CSO 18473

Back cover image: A snapshot of atmospheric vorticity simulated by the Global Cloud Resolving Model. See page 34.

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

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

Lawrence Berkeley National Laboratory
1 Cyclotron Road
Berkeley, CA 94720
CSO 18473

NERSC

BERKELEY LAB
Lawrence Berkeley
National Laboratory

U.S. DEPARTMENT OF ENERGY
Office of Science