Lawrence Berkeley National Laboratory

Lawrence Berkeley National Laboratory

Title
NERSC Annual Report 2004

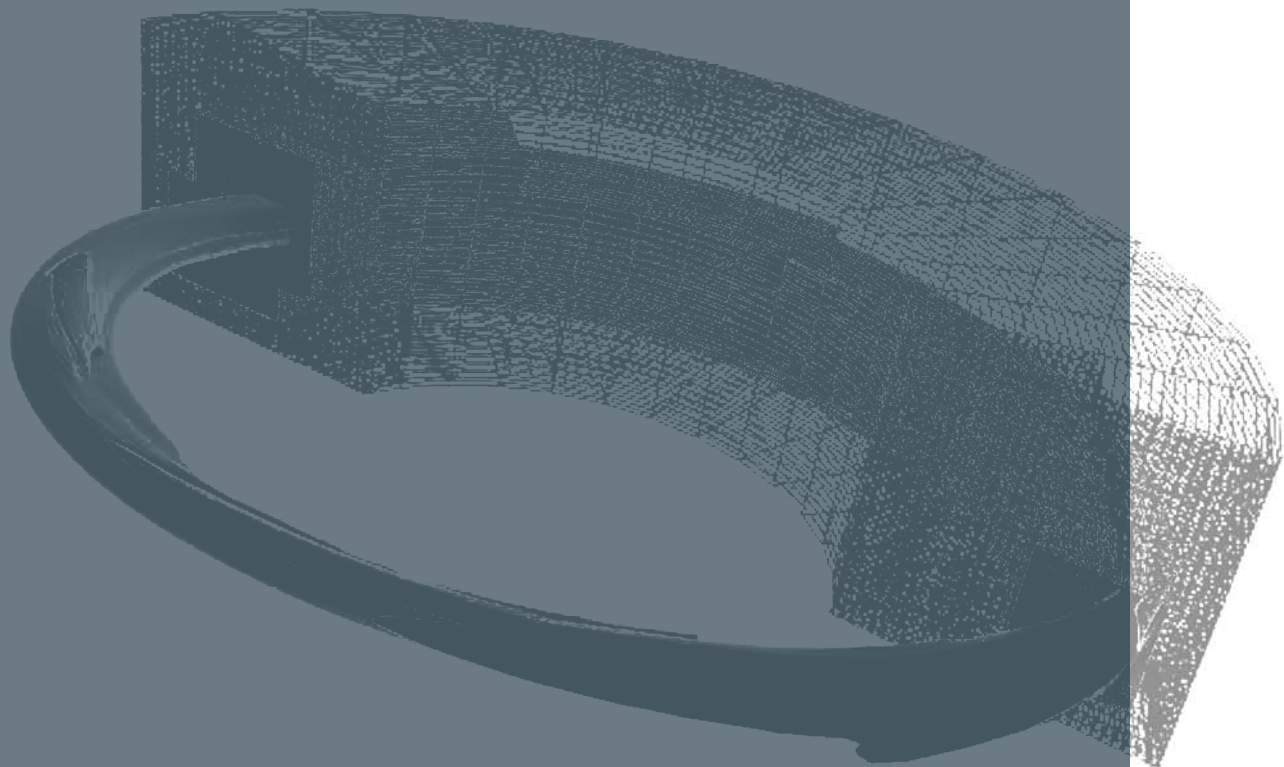**Permalink**
https://escholarship.org/uc/item/0pj4f33c

Authors
Hules, John
Bashor, Jon
Yarris, Lynn
et al.

Publication Date
2005-04-15

National Energy Research Scientific Computing Center
2004 annual report

BERKELEY LAB

NERSC
NATIONAL ENERGY RESEARCH
SCIENTIFIC COMPUTING CENTER

Office of
Science
U.S. DEPARTMENT OF ENERGY

N E R S C   2 0 0 4
Table of Contents

THE YEAR IN PERSPECTIVE

In the future we may remember the spring of 2005 as the time when rapidly rising gas prices brought home to the average consumer the fact that fossil fuels are indeed a limited resource. While a public debate over possible responses and new long-term approaches to this problem has barely begun, we find it fascinating that computational science research carried out using NERSC resources in 2004 is addressing one of the fundamental questions that may hold the key to solving our energy problem. There is growing enthusiasm in the scientific community for a possible solution to this problem through a better understanding of processes in nature that convert solar to chemical energy. Photosynthesis in plants accomplishes this process with 97 percent efficiency, while our best human-engineered solar cells convert only about 9 to 14 percent of the light!

One of the scientific projects at NERSC aims to increase understanding of the complex chemical processes that occur during photosynthesis, taking in carbon dioxide and producing oxygen in the process. This project, led by William Lester of UC Berkeley and described in detail in this annual report, is important on several levels. First, plants and bacteria are the world's foremost means of "carbon sequestration," or removing carbon from the atmosphere in the form of CO_2 — a process which has enormous implications for climate change and global warming. Additionally, photosynthesis is an example of fundamental electron chemistry and is an efficient energy transfer system — processes which are important in many areas of scientific research. But most importantly for the NERSC community, this project was selected under a new competitive program, entitled Innovative and Novel Computational Impact on Theory and Experiment (INCITE), conceived by Dr. Raymond Orbach, Director of the DOE Office of Science. The goal of the program was to select a small number of computationally intensive large-scale research projects that can make high-impact scientific advances through the use of a substantial allocation of computer time and data storage at the NERSC Center.

The first full year of the INCITE program at NERSC in 2004 brought some truly outstanding scientific results using our computational resources. In addition to the better understanding of energy transfer mechanisms in photosynthesis, INCITE researchers investigated stellar explosions and the contribution they make to the abundance of intermediate-mass elements. They also used NERSC to gain greater insight into fluid turbulence and mixing at high Reynolds numbers. From accomplishments such as these, it is clear that we really have entered the new era when computational science is an equal partner with theory and experiment in making scientific progress. All the INCITE results and dozens more are described in our annual report. They demonstrate that NERSC continues to provide one of the most effective and productive high-end capabilities for computational science worldwide.

Supercomputing continues to evolve at a rapid rate. While there are several systems at other sites that claim records of one sort or another, our 10 Tflop/s Seaborg system continues to break new ground when it comes to

reliability and productivity using a unique resource. In 2004 NERSC has moved aggressively to devote a greater share of its processing time to jobs running 512 processors or more. On average, now more than half of all jobs on Seaborg fall into this category, and towards the end of the year these large jobs used around 78% of the system, a dramatic improvement over last year. One of these jobs was a calculation of an entire year's worth of simulated data from the Planck satellite, which ran on 6,000 processors in just two hours.

NERSC users already have the benefit of exploring the scalability of their applications to thousands of processors, and consequently will be ready for the next generation of supercomputer platforms with tens of thousands of processors. Scaling to larger numbers of processors is not only a challenge for algorithms and applications, but it also exposes other limits implicitly built into our thinking and our systems. For example, when trying to tune the performance of codes running on large numbers of processors, most of the existing performance tools exhibited large overhead, and NERSC staff had to develop Integrated Performance Monitoring (IPM). IPM has extremely low overhead and is scalable to thousands of processors. Of course IPM was applied to analyze the INCITE applications, and helped to make the most effective use of their allocations. This is a beautiful example of how the integration of powerful computing systems and excellent staff at NERSC are creating productivity gains for computational scientists that are not easily obtained elsewhere.

On the national scene, a number of new activities that were initiated in 2002 and 2003 were completed in 2004 and will hopefully make a profound impact on computational science and high performance computing in the U.S. The High End Computing Revitalization Task Force (HECRTF) report was released in March 2004, providing the blueprint for further development of high-end computing in the U.S. and defining areas of collaboration for Federal agencies. The second volume of the "Science-Based Case for Large-Scale Simulation" (SCaLeS) report demonstrated the wealth of applications and the potential of scientific progress enabled by next-generation platforms. Finally, the National Research Council study "Getting Up to Speed: The Future of Supercomputing" was released in November 2004, and makes a powerful case for the importance of super-computing both for conducting basic scientific research, as well as for ensuring the economic and physical well-being of the country. These reports give us a glimpse of what simulations are possible at sustained speeds in the range of tens to hundreds of teraflop/s, and lay out the critical research issues that the community needs to address in order to

reach petascale computing performance. This is exciting news, because we are thinking big again in supercomputing. It remains to be seen how well this enthusiasm will translate into future budgets, but it is highly satisfying to see so many NERSC users being deeply engaged in these strategic processes and helping to build our future. Clearly the DOE computational science community, in particular NERSC users, are ready to take the lead in moving up to the next level of computational science.

Several steps towards this bright future will happen at NERSC in 2005. NERSC will acquire and put into production two new computing systems that will take over some of the existing capacity user workload from Seaborg, making even more time available on Seaborg for large-scale capability computing. This will be just an intermediate step towards our next big acquisition, NERSC-5, which will also be initiated in 2005. Thanks to the support from our program management at the DOE Office of Advanced Scientific Computing Research, the NERSC budget has been set at a level that makes all these ambitious plans feasible. Thus we are confidently looking forward to another year of both scientific and computing accomplishments at NERSC. As always, this progress would not be possible without the NERSC staff, who continue to tirelessly dedicate their time, skill, and effort to make NERSC the best scientific computing resource in the world. Our special thanks to all of you.

Horst D. Simon

NERSC Center Division Director

William T. C. Kramer

Division Deputy and NERSC Center General Manager

ADVANCES IN COMPUTATIONAL SCIENCE

As the premiere computational science facility for the DOE Office of Science, NERSC plays a major role in a wide range of discoveries — from predicting the properties of experimental nanostructures, to finding a surprising triggering mechanism for supernova explosions that explains previously puzzling observations. As a partner in discovery, NERSC works with scientists to steadily expand the capabilities of scientific computation. This section highlights some of the results of this partnership in 2004.
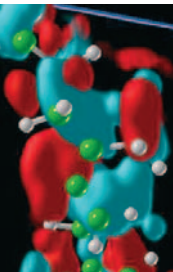
FOLLOW THE ENERGY

SIMULATIONS ARE UNRAVELING THE DETAILS OF ENERGY TRANSFER IN PHOTOSYNTHESIS

Photosynthesis may hold the key to solving one of humanity's most pressing problems: finding a clean, efficient, carbon-neutral, and sustainable source of energy to replace fossil fuels and slow down global warming.

While commercially available photovoltaic cells in solar panels convert only 9 to 14 percent of the sunlight they absorb into electrical energy, the initial photosynthetic steps of energy and electron transfer in green plants and cyanobacteria are 97 percent efficient. Photosynthesis also removes some carbon dioxide, the most important greenhouse gas, from the atmosphere, converting the carbon to carbohydrates.

Nature has a lot to teach us, if we can figure out how photosynthesis works at the level of electrons in molecules.

Researchers are taking a big step toward that understanding with the project "Quantum Monte Carlo Study of Photoprotection via Carotenoids in Photosynthetic Centers," one of the first projects funded by the DOE Office of Science's Innovative and Novel Computational Impact on Theory and Experiment (INCITE) program. Led by theoretical chemist William Lester, Jr. of the University of California, Berkeley and Lawrence Berkeley National Laboratory, the project used about 1.4 million processor hours on NERSC's IBM supercomputer, Seaborg, in 2004. (The "quantum Monte Carlo" in the project title is a statistical model for studying strongly correlated systems such as electrons in molecules. "Photoprotection" is a defense mechanism within the photosynthetic system that protects plants from the oxidation damage that would otherwise result from absorbing more solar energy than they can immediately utilize.)

"The theory behind energy transfer in photosynthesis is more than 50 years old, but some aspects have never been quantitatively tested," says research team member Graham Fleming, who is Berkeley Lab's Deputy Director and a chemistry professor at UC Berkeley. "We need the capabilities of NERSC to provide crucial quantities to test our models for energy transfer in photosynthetic complexes."

Lester says, "Before we had computational capabilities such as those at NERSC, it was not possible to model the energy and electron transfer processes we want to study. NERSC is providing us with the computers and software support that enable us to run codes developed in my laboratory that will give us the information we need and could not otherwise obtain."

Life on Earth depends on the photosynthetic reactions that green plants and cyanobacteria use to convert energy from sunlight into chemical energy. Among other things, these reactions are responsible for the production of all of our planet's oxygen. In high school biology, students learn that nature uses chlorophyll, the family of green pigment molecules, as a light absorber and energy-transfer agent, but the physics and chemistry behind the overall process are extremely complicated. What's more, the elementary photosynthetic steps take place on times as short as a few tens of femtoseconds (a femtosecond being one millionth of a billionth of a second).

"According to the first law of photosynthetic economics, a photon saved is a photon earned," Fleming says. "Nature has designed one of the most exquisitely effective systems for harvesting light, with the reactions happening too fast for any light to be wasted as heat. Current synthetic light-harvesting devices, however, aren't following nature's model.

"The photosynthetic light-harvesting system is so sensitive to changing light conditions, it will even respond to the passing of clouds overhead," Fleming adds. "It is one of nature's supreme examples of nanoscale engineering."

Photosynthesis starts with a light harvesting system, which consists of two protein complexes, Photosystem I and Photosystem II. Each complex features light-absorbing antennae made up of members from two families of pigment molecules, chlorophylls and carotenoids. These pigment antennae are able to capture photons of sunlight over a wide spectral and spatial cross-section.

The chlorophyll and carotenoid molecules gain extra "excitation" energy from the captured photons that is immediately funneled from one neighboring molecule to the next, until it arrives at another molecular complex, which serves as a reaction center for converting energy from solar to chemical (Figure 1). This transfer of excitation energy involves several hundred molecules and hundreds of individual steps along different electronic pathways, yet still transpires within 30 picoseconds for Photosystem I and 200 picoseconds for Photosystem II. By human standards of time, that's instantaneous.

"If we can follow the steps in transferring energy from donor to acceptor molecules, we might be able to design new and much more effective strategies
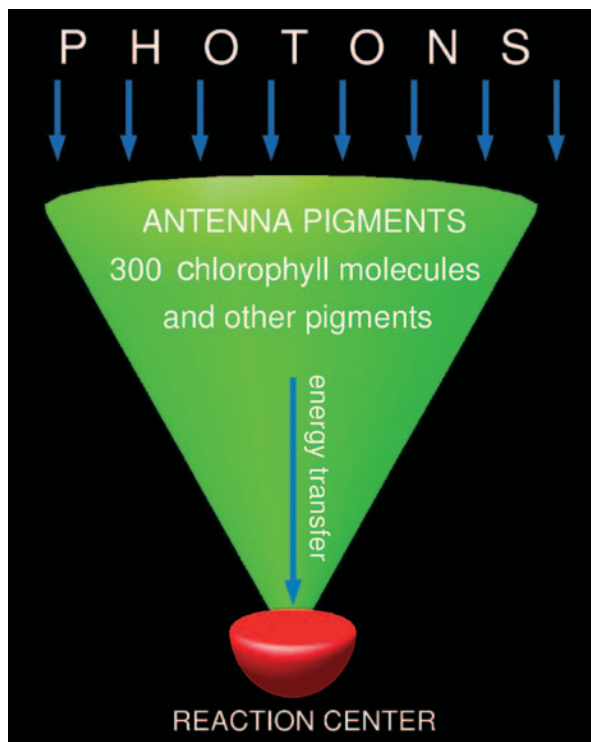
FIGURE 1   A simplified schematic of photon harvesting and energy transfer to the photosynthetic reaction center.

for synthetic light harvesters," Fleming says.

Because the extra energy being transferred from one molecule to the next changes the way each molecule absorbs and emits light, the flow of energy can be followed spectroscopically. However, to do this, Fleming and his experimental research team need to know what spectroscopic signals they should be looking for. This is where the INCITE grant is helping. Lester, Alán Aspuru-Guzik, and other collaborators have developed and are running a quantum Monte Carlo computer code called Zori[1] on NERSC's

[1]Alán Aspuru-Guzik, Romelia Salomón-Ferrer, Brian Austin, Raul Perusquía-Flores, Mary A. Griffin, Ricardo A. Oliva, David Skinner, Dominik Domin, and William A. Lester, Jr., "Zori 1.0: A parallel quantum Monte Carlo electronic structure package," J. Comp. Chem. (in press). Harsha Vaswani and Michael Frenklach contributed to the photoprotection application of Zori. The Zori code and documentation are available at http://www.zori-code.com/.

FIGURE 2   Nuclei of the spheroidene molecule, $C_{41}H_{60}O$. Carbon is shown as green, hydrogen as gray, and oxygen as red.

FIGURE 3   Electron density of the spheroidene molecule.

Seaborg computer to predict the optimal electronic pathways for photosynthetic energy transfer. NERSC consultant David Skinner helped optimize the code to take full advantage of Seaborg's capabilities.
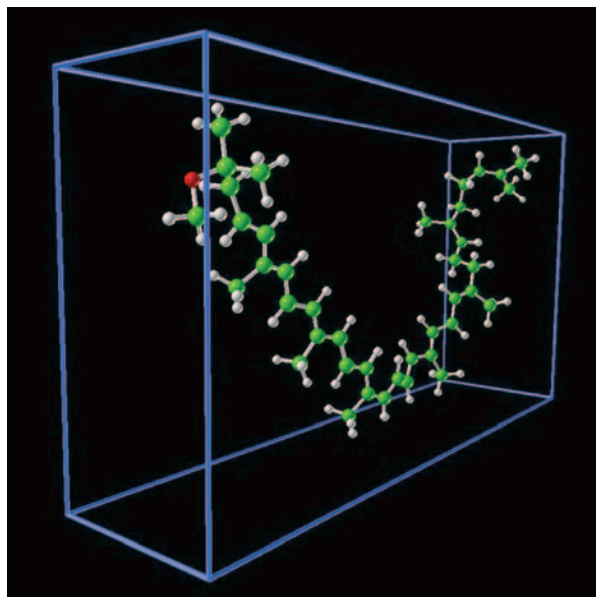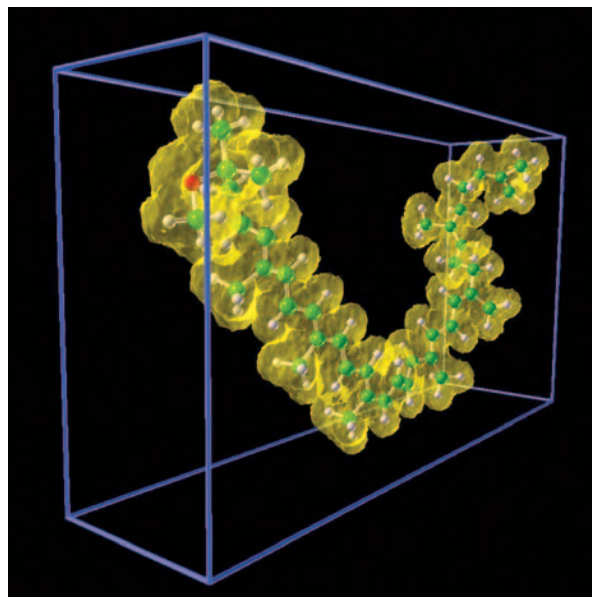
Says Lester, "Most people have long thought of computational chemistry as only being able to tackle simple systems reliably, but we've come a long way with improved implementation of our algorithms in recent years."

Rather than tackling the entire complex process of photosynthesis, Lester's team wanted to work on a manageable piece of the puzzle, so Fleming suggested studying the electronic structures behind photo-protection. Research into photoprotection has focused on the carotenoids in Photosystem II, which appear to be the regulatory agents that dissipate excess energy, thus preventing oxidation damage. (Carotenoids such as beta-carotene and lycopene from the vegetables and fruits we eat perform a sim-

ilar antioxidant function in the human body, protecting us from a variety of illnesses.)

When green plants and photosynthetic bacteria (including the cyanobacteria and purple bacteria families) absorb more energy than they can use, the chlorophyll molecules can undergo a transition to a "triplet" state — a kind of variation in their energy state. This state is very dangerous because it, in turn, can excite molecular oxygen to its excited "singlet" state — an exceedingly reactive chemical that destroys the photosynthetic proteins and, in the end, will kill the plant or bacterium.

Carotenoid molecules are ubiquitous in photosynthetic organisms because they can remove ("quench") chlorophyll triplets, leading to harmless generation of heat instead of dangerous singlet oxygen. In green plants, an additional process called "feedback de-excitation quenching" takes place, in which carotenoids detect changes in the rate of photosynthesis via pH
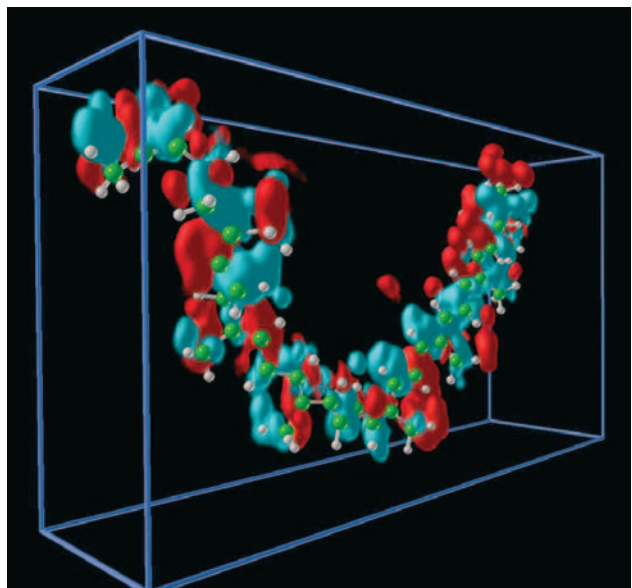
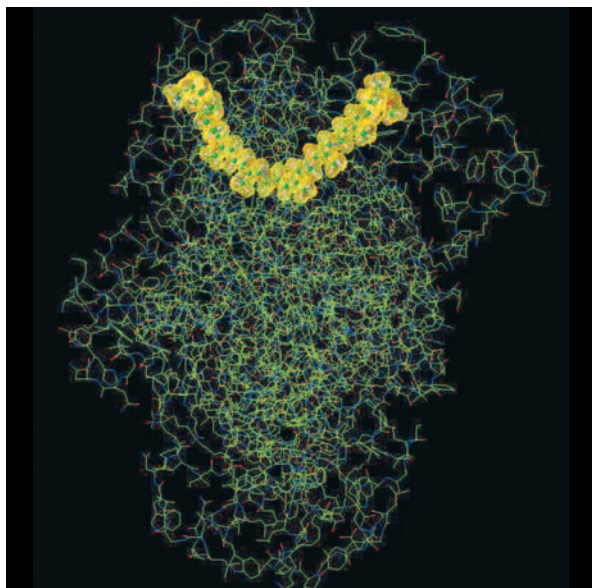FIGURE 4   Spin density of the spheroidene molecule.

FIGURE 5   The spheroidene molecule in the protein environment. The Lester group is developing methods for approximate treatment of the chemical environment of the region surrounding the molecule studied with the accurate quantum Monte Carlo method.

levels (the feedback mechanism) and stop the excitation from reaching the reaction center where oxidizing compounds could be formed. Instead, the excitation is dissipated from one molecular system to another and the energy is safely released as heat.

Earlier studies had already identified spheroidene as a photoprotective carotenoid in bacteria, so Lester's research team set out to calculate the excitation energies and rate of triplet energy transfer between bacteriochlorophyll (a close chemical relative of the chlorophyll in green plants) and spheroidene. They have developed a linear diffusion Monte Carlo method that has high accuracy, scalability, and relative speed. The optimized Zori code now runs 10 times faster than the researchers' original algorithm, and they are now able to study systems 4 times larger than those they worked on before the INCITE grant.[2]

To make the results of their calculations more understandable, the Berkeley Lab/NERSC Visualization Group worked with the researchers to create visual images and movies of the simulation data (see Figures 2–5 and http://www.vis.lbl.gov/Events/SC04/Incite1/). As the research progresses, the visualization techniques developed for this project will be able to illustrate the energy transfer pathways of photosynthesis.

The next step for Lester's research group is to implement an electron pair localization function that describes the pairing of electrons in the spheroidene–bacteriochlorophyll molecular system, further clarifying the energy transfer pathways. Then they will apply their method to other components of photosynthetic processes, such as zeaxanthin, a carotenoid in green plants that was recently proven

[2] A. Aspuru-Guzik, R. Salomón-Ferrer, B. Austin, and W. A. Lester, Jr., "A sparse algorithm for the evaluation of the local energy in quantum Monte Carlo," J. Comp. Chem. (in press).

by Fleming's experimental group to play a photoprotective role in green plants.

The computational scientists are also aiming to develop optimal quantum Monte Carlo methods to calculate the excited state of even larger molecules involving thousands of electrons — a task that will require the capabilities of the next generation of supercomputers. Although the focus of this project is photosynthesis, the work itself is fundamental electron chemistry, so the computational methods and codes developed in this project are likely to find applications in a much wider range of chemical research.

BEHIND THE BIG FLASH
CALCULATIONS REVEAL NEW EXPLANATION OF SUPERNOVA EXPLOSIONS

Type Ia supernovae are stars in more than one sense — they are playing a leading role in revealing the history of the Universe, and they are attracting the devoted attention of a multitude of researchers.

As cosmic distance markers, Type Ia supernovae, combined with other observations, led to the discovery of the mysterious dark energy that is accelerating the expansion of the Universe. These supernovae also play an important role in the chemical evolution of the Universe, fusing carbon and oxygen atoms into heavier elements such as magnesium, silicon, sulfur, iron, and nickel.

Understanding Type Ia supernovae in more detail will provide answers to many cosmological questions. But one of the most basic questions about the supernovae themselves remains unanswered: How do they explode?

Although the origins of Type Ia supernovae have not been demonstrated conclusively, they are generally believed to originate in binary systems in which a dense, compact star (possibly a white dwarf) accretes mass from a companion star (Figure 1). When the compact star's mass approaches the Chandrasekhar limit (about 1.4 times the mass of the Sun), a thermonuclear explosion takes place, consuming the entire star and creating a brilliant flash that traverses the expanse of the Universe. But the triggering mechanism of that explosion, and the details of its evolution, have remained mysteries.

A star can burn in two ways: like a flame, which is called deflagration, or like an explosion, called detonation. Neither deflagration nor detonation by itself can explain everything astronomers see in Type Ia supernovae. So in the past decade, theorists have proposed several "delayed detonation" models, in which a deflagration phase is followed by detonation. These models produce plausible results, but they do not explain why or when the star detonates.

More details on how the explosion might be initiated have emerged from the unprecedented full-star simulations created by one of the 2004 INCITE projects
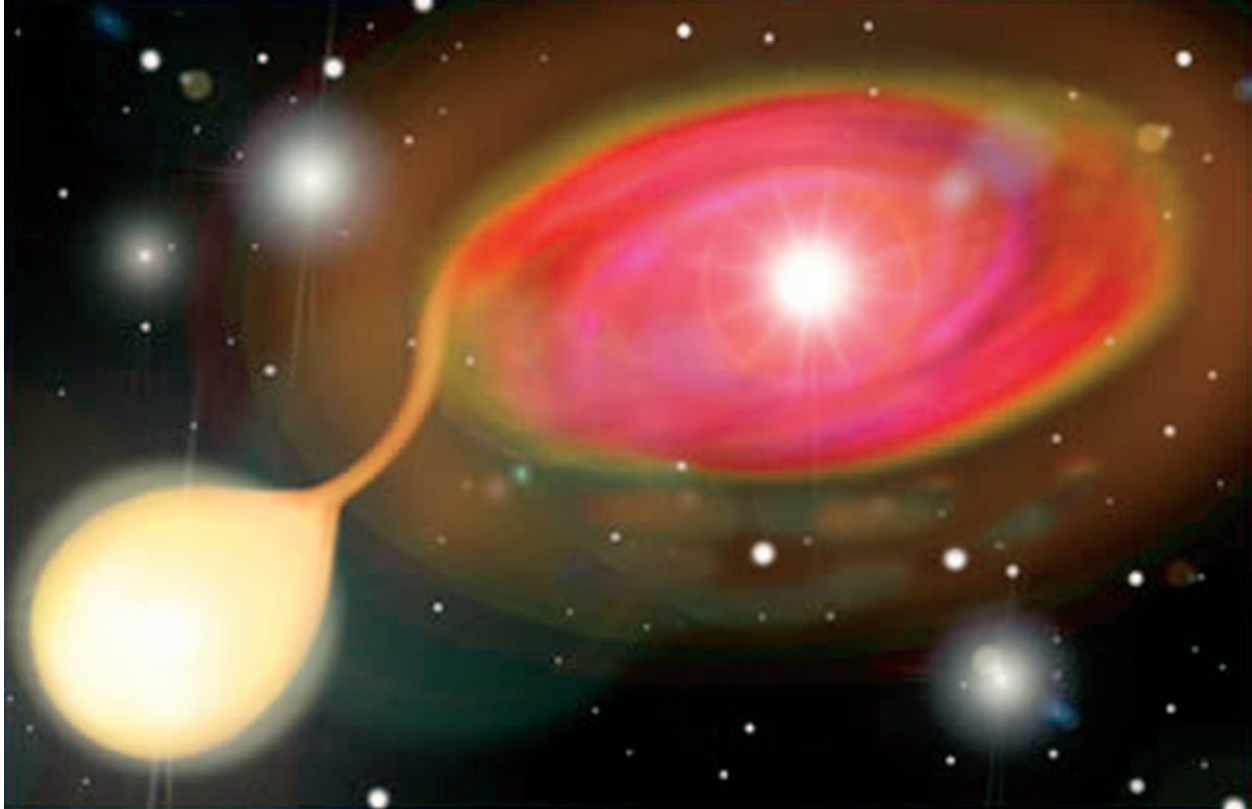
FIGURE 1   Artist's rendition of a white dwarf accreting mass from its companion star.

FIGURE 2   A three-billion-degree bubble of thermonuclear flame mushrooms out of a compact star just seconds before a supernova explosion. Pulled by the star's gravity, the flame will sweep around the star's surface and collide with itself, detonating the explosion.

at NERSC, "Thermonuclear Supernovae: Stellar Explosions in Three Dimensions," led by Tomasz Plewa of the Center for Astrophysical Thermonuclear Flashes at the University of Chicago and the Nicolaus Copernicus Astronomical Center in Warsaw.

The simulations that Plewa and his colleague Timur Linde ran on Seaborg — using 3.2 million processor hours over the course of the year — investigated the birth of the thermonuclear flame, an essential part of an unexpected scenario that the researchers call "gravitationally confined detonation" or GCD.[1] In this scenario, the explosion begins with the ignition of

deflagration slightly off-center in the core of the star, which results in the formation of a buoyancy-driven bubble of hot material. This bubble rushes outward at transonic speeds and breaks through the stellar surface (Figure 2). Confined by the star's gravity, the burning shock wave races around the surface of the

[1]T. Plewa, A. C. Calder, and D. Q. Lamb, "Type Ia supernova explosion: Gravitationally confined detonation," Astrophys. J. **612,** L37 (2004).
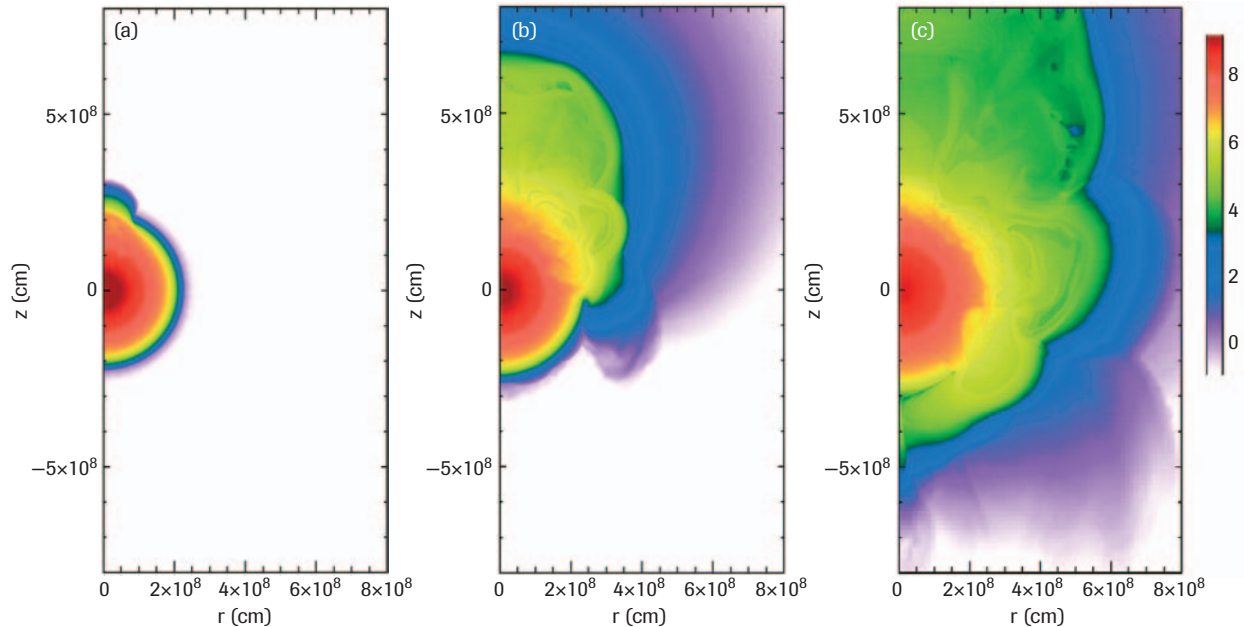
FIGURE 3 This half cross section of the star shows the density evolution of the deflagration phase, from the point of bubble breakout (top of image) at 0.9 seconds after ignition (a), 1.4 seconds (b), and 1.9 seconds (c). Note that most of the material in the outer layers of the star is closely confined as it rushes around the surface and engulfs the star.

star, pushing the fuel-rich outer layers ahead of it like a thermonuclear tsunami (Figure 3).

This flood of nuclear fuel converges at the point opposite the bubble breakout, forming a conical compressed region bounded by the shock (Figure 4). The crashing waves of matter carry enough mass and energy to trigger a detonation just above the stellar surface, resulting in a supernova explosion that incinerates the progenitor star.

Plewa and his team still need to fill in the details of gravitationally confined detonation. For example, although the simulations show the right conditions for detonation in two dimensions (assuming a simplified geometry of the problem), the researchers have not yet been able to simulate the inception of detonation in three dimensions. Such high-resolution

simulations will push current supercomputers to the limits of their capabilities.

In addition, the GCD mechanism needs to be verified by comprehensive spectral and polarimetric studies and observations. The first such study was conducted by Daniel Kasen and Plewa, who used a part of the INCITE allocation to conduct Monte Carlo radiative transport calculations. They demonstrated that certain spectral features of some Type Ia supernovae that cannot be explained by other models can be interpreted as natural consequences of the deflagration phase in the GCD model.[2]

One strength of the GCD model is that it can account for a number of typical characteristics observed in thermonuclear supernovae:

[2] D. Kasen and T. Plewa, "Spectral signatures of gravitationally confined thermonuclear supernova explosions," Astrophys. J. Lett. (in press), astro-ph/0501453 (2005).
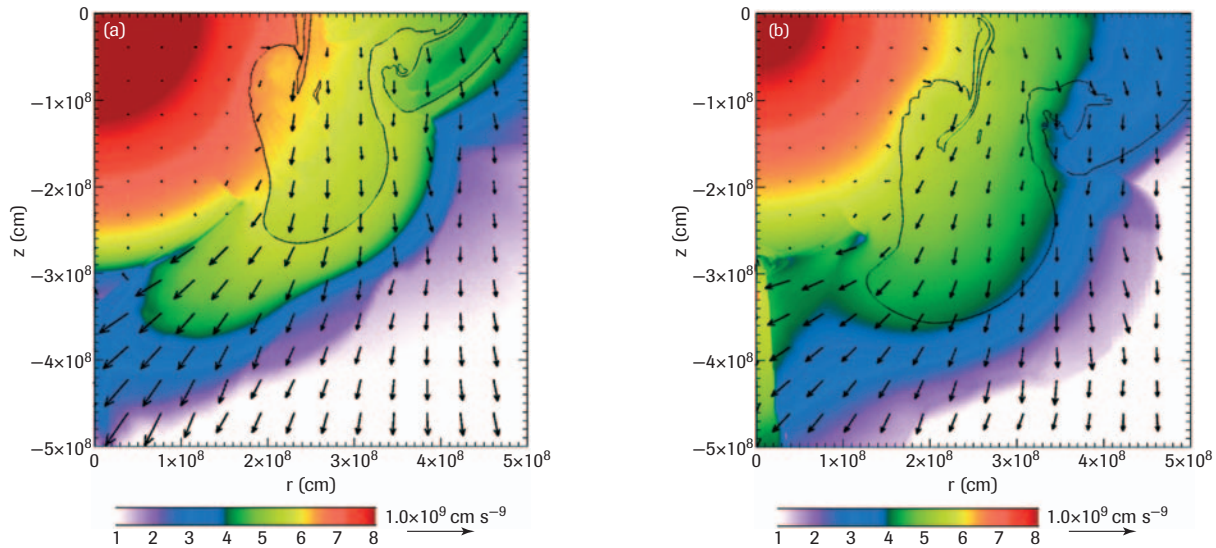
FIGURE 4   This quarter cross section shows the density evolution of the surface flood across the lower portion of the star at 1.85 seconds (a) and 2.005 seconds (b). The conical compressed region is the bright green strip in the lower left of (b); the high-pressure detonation point is the red spot just below -3×10^8.

- The stellar expansion following the deflagration redistributes mass in a way that ensures production of intermediate-mass and iron group elements.

- The ejected matter has a strongly layered structure, with lighter elements on the outside and heavier elements on the inside.

- The ejected matter is slightly asymmetrical, resulting from bubble breakout and detonation occurring on opposite sides of the star.

- This asymmetry, combined with the amount of stellar expansion determined by details of the evolution (principally the energetics of deflagration, timing of detonation, and original structure of the progenitor star), can be expected to create a family of Type Ia supernova explosions that are almost but not exactly alike — just as astronomers have observed.
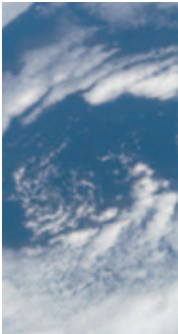
CREATION OF THE WHORLED

COMPUTATIONS PROBE THE SUBTLE PRINCIPLES GOVERNING TURBULENCE

In the world of physics, the difficulty of understanding turbulence in fluids is legendary. A humorous legend has Albert Einstein (or Werner Heisenberg, or Horace Lamb) saying on his deathbed, "I'm going to ask God two questions: Why relativity (or quantum electrodynamics), and Why turbulence? I'm rather optimistic about getting an answer to the first question."

A famous quotation is also variously attributed to Einstein, Heisenberg, Richard Feynman, or Arnold Sommerfeld: "Turbulence is the last great unsolved problem of classical physics." Whether any of those luminaries actually uttered those words is beside the point — everyone agrees with the statement.

Fluids (including air) are everywhere, and their motions are usually turbulent. You can create an exception by turning on a water tap just a little bit: the water flow will be smooth and constant, or in the jargon of physics, "laminar." Turn on the water full blast, and the flow becomes disorderly and turbulent. The same change happens to smoke rising from a cigarette into still air: the smoke immediately above the cigarette is laminar, but a little higher up it becomes rippled, chaotic, and diffusive — turbulent.

Fluid motion becomes turbulent when the speed of the fluid exceeds a specific threshold, below which frictional ("viscous") forces prevent the chaotic behavior. The mathematical expression of that threshold is the Reynolds number, sometimes described in simplified terms as the ratio of inertial to viscous forces (although viscous effects remain significant even at very high Reynolds numbers). The Reynolds number is variable because it is proportional to both the size of the object and the flow velocity. The Reynolds number for air flowing over a cruising aircraft might be in the range of 100 million, while for blood flowing through an artery, it might be around 1,000. But for any given problem, the higher the Reynolds number, the more turbulent the flow — and the more mathematically complicated.

The motion of fluids is described mathematically by the Navier-Stokes equations, which express basic principles of conservation of mass and momentum. Solving the Navier-Stokes equations for high Reynolds number turbulent flows requires massive calculations; so to make the problem manageable, most full-scale engineering calculations use approximations or models of some variables.

But a method called "direct numerical simulation" (DNS) is available that solves the exact conservation equations for mass, momentum, and chemical species concentration and mixing without any approximations. DNS is a powerful research tool that can supply data at a level of detail well beyond what is possible in experiments. Researchers use DNS results to improve their understanding of turbulent phenomena and to develop and test more accurate statistical models of turbulence. Those models in turn can be used to analyze experimental results; to simulate data that cannot be readily obtained by experiment, such as correlations between variables; or to solve engineering problems.

"The complexities of turbulence limit our ability to predict natural phenomena, such as weather, and to design improved engineering devices, ranging from

FIGURE 1   NASA photo of a cloud vortex with smaller vortices over the Madeira Islands.

engines to airplane wings to artificial heart valves," says P. K. Yeung, Professor of Aerospace Engineering at the Georgia Institute of Technology, and principal investigator of the INCITE project "Fluid Turbulence and Mixing at High Reynolds Number." This project used close to 2 million processor hours at NERSC in 2004 in the quest to find the underlying order in apparently disorderly motion and to help develop statistical models of turbulent mixing and dispersion.

"DNS is very CPU intensive, so it provides a grand challenge for high performance computing," Yeung says. "Our INCITE award has allowed us to perform the largest simulation of fluid flow turbulence ever done in the U.S., at a level of detail and within a time frame not possible otherwise. We have used as many as 8 billion grid points to probe deeply into a problem arising in multiple fields of science and engineering. Our simulation has achieved a Reynolds

number comparable to or higher than that observed in many laboratory experiments."

An important feature of turbulent flows is that they are composed of fluctuating eddies or vortices of many sizes, which are constantly forming and breaking down (Figure 1). When the smallest eddies succumb to viscosity, their energy dissipates into heat. The first person to apply mathematics to weather forecasting, British meteorologist Lewis Richardson, described this process in verse:

> *Big whorls have little whorls,*
> *That feed on their velocity;*
> *And little whorls have lesser whorls,*
> *And so on to viscosity.*

The flow of energy between the large, intermediate, and small scales was first expressed mathematically by Kolmogorov in 1941. His theory predicts a uni-

versal constant and a −5/3 exponent for the energy spectrum at intermediate scales for all flows if the Reynolds number is high enough. While the Kolmogorov constant is still considered a good approximation, over the years researchers have found that some features of the theory are not confirmed by experiment.

"High Reynolds number simulations can help answer unresolved questions like this," Yeung says. "Although our focus in this project is on small scales, the high grid resolution allows a wide enough range of scales to test the Kolmogorov constant and other theories and models used to describe flow behavior in applications."

Yeung's DNS simulations examine the most basic, simple form of turbulence, which is isotropic, that is, uniform in all directions. With high resolution and long runs, these simulations produce more detailed and realistic data than can be obtained from shorter, low-resolution simulations. DNS data from multiple computer runs can be analyzed to derive important statistical characteristics of turbulent flows, as well as information that can be used to model complex real-world turbulence. The INCITE project has made significant contributions to a database that now encompasses Taylor-scale Reynolds numbers from 38 to 700 and grid resolutions from 64^3 to 2048^3.

As if turbulence were not difficult enough, the mixing or diffusion of substances or heat in turbulent flows adds further mathematical complications. Diffusion is important for anyone studying the spread of pollutants released into water or the atmosphere, and it is especially important in the study of combustion, where chemical reactions depend on the proximity of the reactants. For example, when a homogeneous mist of fuel is injected into a combustion chamber, the turbulence of the gases already present breaks up the uniformity of

the fuel mist at first, making it uneven and disorderly. But a high level of turbulence eventually tends to mix the fuel with the gases more or less evenly. Better control of diffusion in a combustion chamber results in less unburned fuel and fewer undesirable byproducts.

The mathematical representation of the mixed substance in a turbulent flow is called a "passive scalar." Obukhov and Corrsin independently extended Kolmogorov's theory to the study of passive scalar fluctuations. Of course, turbulent mixing of several different substances requires multiple scalars, resulting in even more complex calculations, so the reliability of the Obukhov-Corrsin constant is critical. One of the earliest results of this INCITE project was the best demonstration to date of Obukhov-Corrsin scaling in the spectrum of fluctuations in turbulent mixing, supporting the validity of this theory.

One of the ways in which scalar dissipation differs from energy dissipation in turbulence is that scalar dissipation is less uniform in both space and time. A phenomenon called intermittency — intense, localized fluctuations of any quantity in a turbulent flow — can result in localized extinction and reignition of a combustion process, so calculating intermittency accurately is important in combustion modeling and engineering applications, as well as many others.

To help Yeung and his collaborators understand the patterns of intermittency in their simulations, the Berkeley Lab/NERSC Visualization Group created images showing energy and scalar dissipation for various "slices" of the data, with data magnitude mapped to both color and height for easy visual recognition of differences. Figure 2 shows that at high Reynolds number, scalar dissipation has much higher "peaks" than energy dissipation. Figure 3 compares scalar dissipation at low and high Reynolds numbers, and shows more intense and localized peaks at the higher Reynolds number.
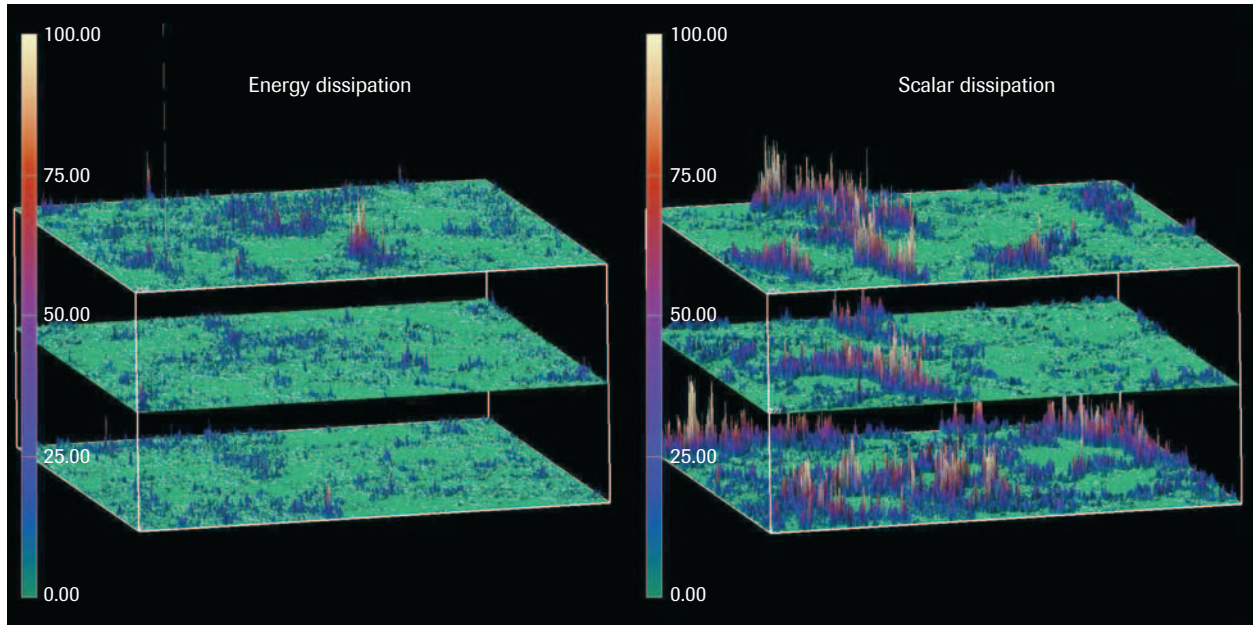
FIGURE 2   Energy and scalar dissipation at Taylor-scale Reynolds number (R_λ) ~700 (2048^3 grid points). Scalar dissipation shows higher peaks.
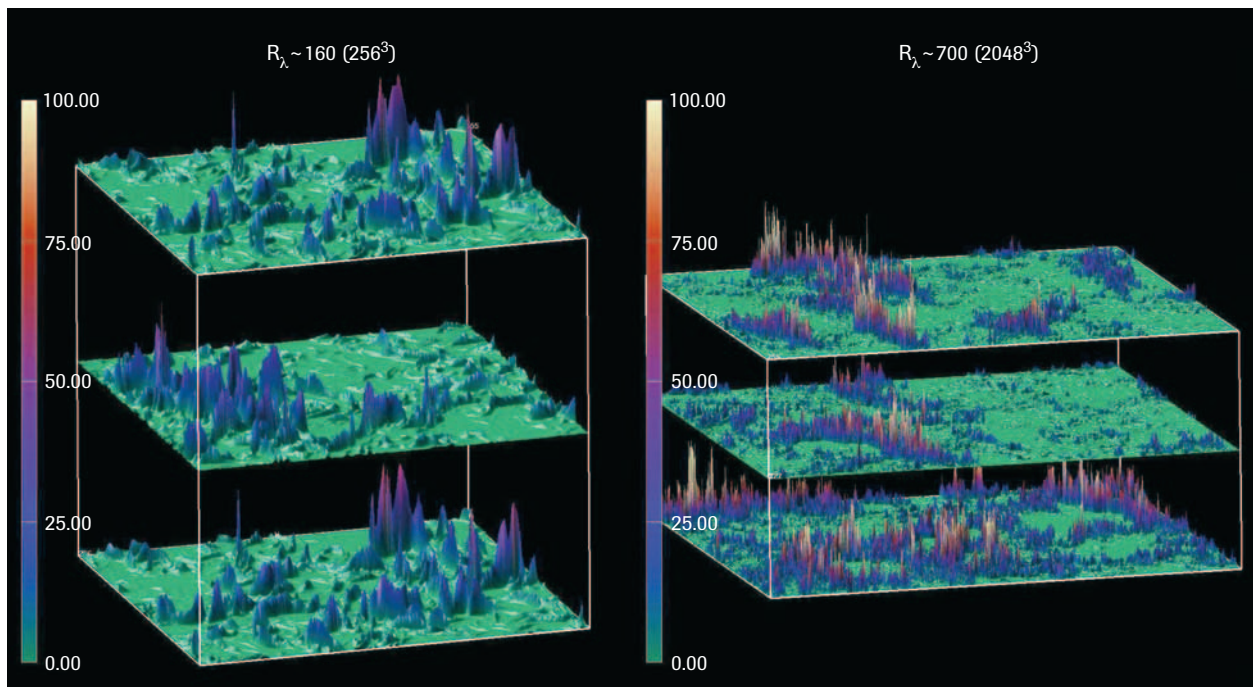
FIGURE 3   Scalar dissipation for low (left) and high (right) Reynolds numbers. The peaks are more intense and localized at the higher Reynolds number.

Figure 4 presents a three-dimensional depiction of scalar dissipation data.

"Our successes with the INCITE program have drawn attention in both the science and supercomputing communities, and created a very broad range of opportunities for the future," Yeung says. "More than twenty leading scientists in the field have indicated a strong desire to access our database to answer questions that they have long sought to resolve. We are extending the simulations to turbulent reacting flows where phenomena such as extinction and reignition are strongly dependent on small-scale mixing and intermittency."

Besides Yeung and his Ph.D. student Diego Donzis of the Georgia Institute of Technology, this project also involves close collaborations with K. R. Sreenivasan, who is both Distinguished University Professor at the University of Maryland and Director of the International Centre for Theoretical Physics in Trieste, Italy, and, more recently, Rodney Fox, who holds both an endowed chair professorship at Iowa State University and an Associate Scientist position at the Ames Laboratory. The first publication of results from this research is P. K. Yeung, D. A. Donzis, and K. R. Sreenivasan, "Turbulence and

FIGURE 4   Using an imaging method called volume rendering, data values from the 3D scalar dissipation data slice are mapped onto color and opacity, revealing more features of the data. A movie of this image rotating can be viewed at http://www-vis.lbl.gov/Events/SC04/Incite3/scdiss2_vol_2.mpg.

scalar transport in numerical simulations at 2048^3 resolution," Bull. Am. Phys. Soc. **49,** 22 (2004).

PARTICLE PHYSICS, PRECISELY

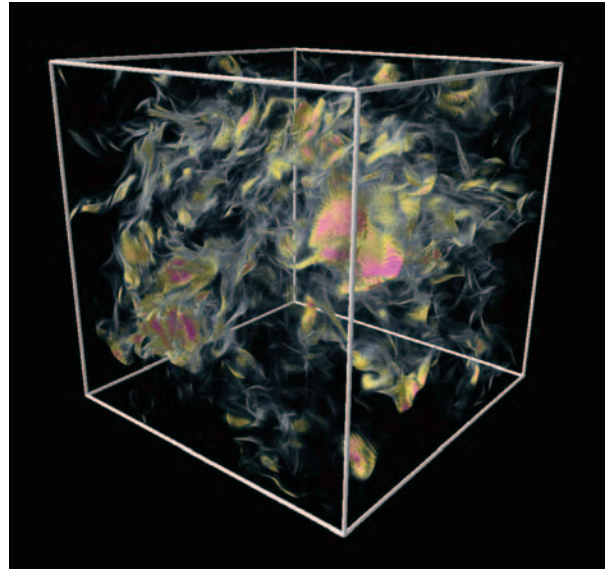## IMPROVED ALGORITHMS AND FASTER COMPUTERS ENABLE FIRST HIGH-PRECISION CALCULATIONS OF HADRON PROPERTIES

Physicists who work in the field of lattice quantum chromodynamics — lattice QCD for short — are nothing if not patient. For 30 years they have been trying to use the basic QCD equations to calculate the properties of hadrons — particles (including protons and neutrons) composed of quarks, antiquarks, and gluons. Limited by the speed of available computers, they have had to simplify their simulations in order to get results in a reasonable amount of time, and those results typically had an error rate of around 15% when compared with experimental data.

Now, with faster computers, improved algorithms that employ fewer simplifications of physical processes, and better-performing codes, four QCD collaborations involving 26 researchers have reported calculations of nine different hadron masses, covering the entire range of the hadron spectrum, with an error rate of 3% or less. This report, published in the journal Physical Review Letters under the title "High-precision lattice QCD confronts experiment,"[1] marks the first time that lattice QCD calculations have achieved such precise results for such diverse physical quantities using the same QCD parameters.

QCD is the theory of the "strong force," by which gluons bind quarks together to form hadrons. QCD theorists use computer simulations to determine the most probable arrangements of gluons and quarks inside a particle and then use these configurations to determine the particle's properties. Calculating all the possible positions of quarks and gluons would be impossible, so theorists simplify the problem by imagining space and time not as a continuum, but

as a lattice — a four-dimensional grid of discrete points at which quarks and gluons can reside. This approach, called lattice QCD, transforms an impossible problem into one that is extremely difficult but, in principle, solvable.

Lattice QCD has already explained why quarks do not exist as separate particles, explored the symmetries of the strong force, and predicted the temperature at which protons and neutrons melt. Soon theorists expect they will be able to make predictions that can be tested experimentally in particle accelerators, probing the limits of the Standard Model of particle interactions.

"Lattice QCD simulations are needed to obtain a quantitative understanding of the physical phenomena controlled by the strong interactions, to determine a number of the basic parameters of the Standard Model, and to make precise tests of the Standard Model's range of validity," explains Doug Toussaint, Professor of Physics at the University of Arizona and a member of the MILC Collaboration,

[1] C.T. H. Davies et al. (HPQCD, UKQCD, MILC, and Fermilab Lattice Collaborations), "High-precision lattice QCD confronts experiment," Phys. Rev. Lett. **92,** 022001 (2004).

one of the four teams that co-authored the "QCD confronts experiment" paper.

"Despite the many successes of the Standard Model," Toussaint says, "high energy physicists believe that to understand physics at the shortest distances a more general theory will be required, which unifies all four of the fundamental forces of nature." Those four forces are the strong force; the weak force, which causes a type of radioactive decay; electromagnetism, which ties electrons to atomic nuclei; and gravity, which theorists have found difficult to integrate with the other three forces.

"The Standard Model is expected to be a limiting case of this more general theory, just as classical mechanics is a limiting case of the more general quantum mechanics," Toussaint says. "A central objective of the experimental program in high energy physics, and of lattice QCD simulations, is to determine the range of validity of the Standard Model, and to search for new physics beyond it. Thus, QCD simulations play an important role in efforts to obtain a deeper understanding of the fundamental laws of physics."

One way of determining the validity of the Standard Model is to tabulate the rates at which heavy quarks decay into lighter ones in an array of numbers known as the Cabibbo-Kobayashi-Maskawa (CKM) matrix. The CKM matrix elements — some of which are still unknown — define the intersection between the strong and weak forces in events that involve both. If the Standard Model is complete, then certain combinations of these rates will add up to 100%. If they add up to something else, then new particles or forces must await discovery.

Unfortunately, it is impossible to directly measure the decay of an individual heavy quark into a lighter one, because every quark is bound by a swarm of gluons to other quarks or to an antiquark. Experimenters have to measure the decays of composite particles, such as B mesons (Figure 1), and

then try to calculate the CKM numbers by subtracting the effects of the extra quarks and gluons. That subtraction is easier said than done, because within each hadron is a dynamic "sea" of gluons exchanging gluons among themselves, as well as quark-antiquark pairs popping in and out of existence. The effect of all this frenetic strong-force activity is called "quark vacuum polarization," and it has been one of the most difficult quantities to include in lattice QCD calculations.

Quark vacuum polarization is the most computationally expensive ingredient in a QCD simulation because moving quarks from real space into lattice space increases the number of quarks by a factor of 16. In the past, most QCD simulations have either omitted quark vacuum polarization altogether ("quenched QCD") or inflated the lightest quark masses by a factor of 10 to 20, reducing both the computational effort and the accuracy of the results.

The "Symanzik-improved staggered-quark discretization" of the lattice, recently developed by the MILC Collaboration and others, includes the vacuum polarization effects of all three light quark flavors (up, down, and strange) and allows QCD simulations with much smaller and more realistic quark masses than previously possible. This algorithm improves on an older method called "staggering" — spreading each quark over four neighboring lattice points, and then eliminating many components to reduce the number of redundant quarks coming from the lattice. Redundant quarks are further reduced by rescaling their effect on the system's evolution. In addition, unphysical interactions produced by staggering have been minimized.

Including quark vacuum polarization has dramatically improved the accuracy of lattice QCD calculations, as Figure 2 shows. These results validate the accuracy of the improved staggered-quark discretization. "In this study, we chose to calculate quantities whose values are well known experimentally," says
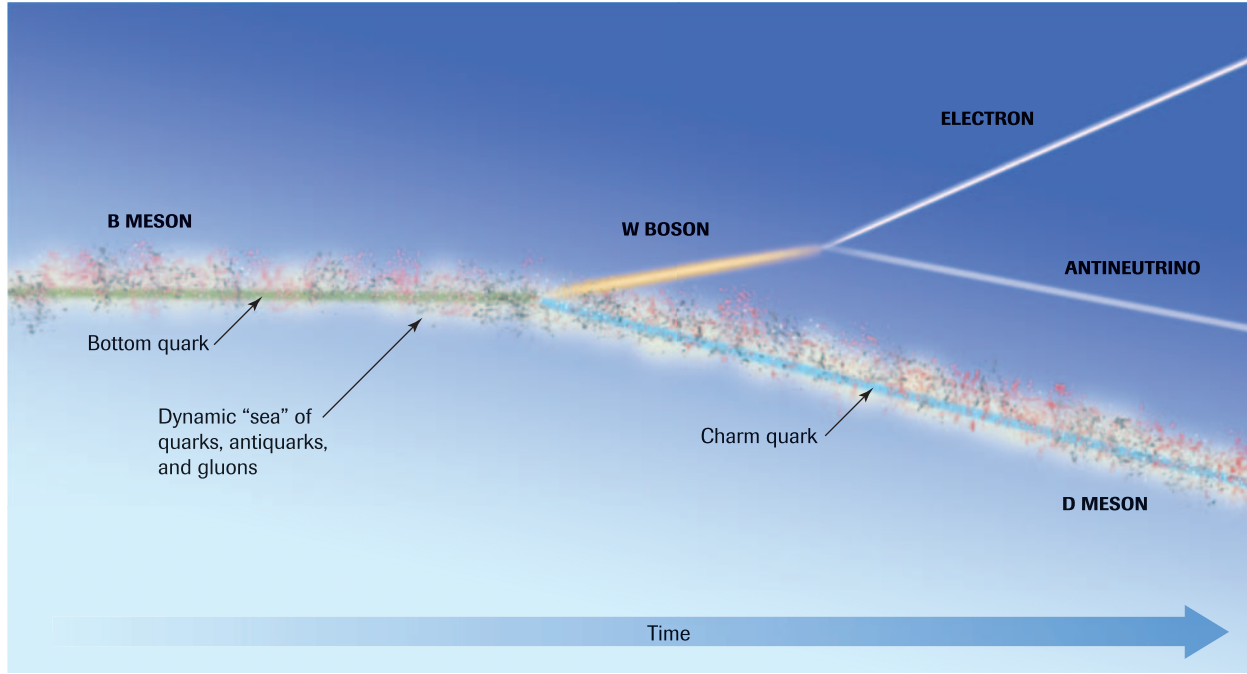
FIGURE 1   The decay rate of a heavy bottom quark in a B meson, a process controlled by the weak force, is obscured by the binding effect of the strong force in a dynamic "sea" of quark-antiquark and gluon activity. Accurate lattice QCD calculations that include this activity are needed in order to correct the uncertainties in experimentally measured decay rates.

MILC collaborator Bob Sugar of the University of California, Santa Barbara. "This allowed us to verify that we have control over all sources of systematic errors." Now that lattice QCD has accurately reproduced a range of well known quantities, researchers will have more confidence in the calculation of other quantities that experiments cannot determine.

Describing the MILC Collaboration's work at NERSC, Sugar says, "The bulk of our allocation goes into the generation of lattices — snapshots of the systems we are simulating — using the improved discretization. Seaborg is one of only two machines available to us that can efficiently handle the large, low-temperature lattices. These lattices are saved, and then used to calculate a wide variety of physical quantities. In order to maximize the physics output from the large investment in the generation of these lattices, we are

FIGURE 2   Lattice QCD calculations divided by experimental measurements for nine different hadron masses, without quark vacuum polarization (left) and with quark vacuum polarization (right). A value of 1 represents perfect agreement.

making them available to other theorists for their research." The MILC collaborators produced the gluon configurations, as well as the raw simulation data for pions and kaons, that were used in the "QCD confronts experiment" paper.
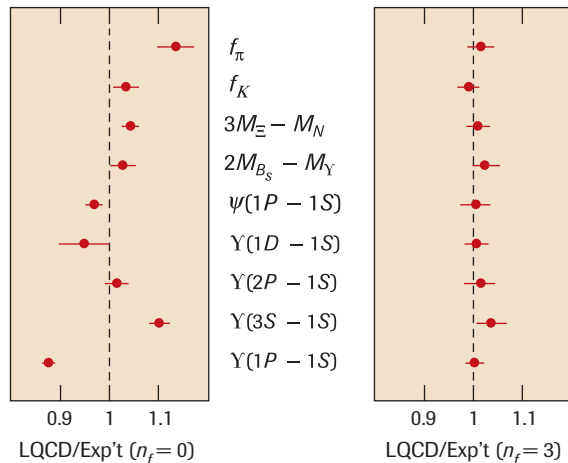
The impact of improved lattice calculations on constraints to the Standard Model is summarized in Table 1, taken from a report by the Lattice QCD Executive Committee.[2] The first column indicates the quantity measured experimentally, the second column the effected CKM matrix element, and the third column the hadronic matrix element which must be calculated on the lattice. The fourth column shows the current non-lattice errors, which are primarily experimental, and the fifth the current lattice errors. The column labeled Lattice Errors Phase I shows the expected lattice errors once the analysis has been completed on existing lattices, and the column labeled Lattice Errors Phase II shows the expected lattice errors once analysis has been completed on the lattices currently being generated or planned for the near future. Note that in all cases for which experiments have been performed, the current lattice error is significantly larger than the experimental one. However, once the Phase II analysis has been completed, the lattice errors will be less than or comparable to the experimental ones. Figure 3 graphically depicts the

FIGURE 3   Constraints on the Standard Model parameters ρ and η. For the Standard Model to be correct, they must be restricted to the region of overlap of all the solidly colored bands. The top figure shows the constraints as they exist today. The bottom figure shows the constraints as they would exist with no improvement in the experimental errors, but with all lattice gauge theory uncertainties reduced to 3%. [Image by R. Patterson from "The SCaLeS Report."[3]]

impact of improved lattice calculations on constraints to the Standard Model.

QCD theory and computation are now poised to fulfill their role as an equal partners with experiment. Sugar comments, "The importance of this can be seen from the fact that a significant fraction of the $750 million per year that the United States spends on experimental high energy physics is devoted to the study of the weak decays of strongly interacting particles. To fully capitalize on this investment, the

Measurement	CKM Matrix Element	Hadronic Matrix Element	Non-Lattice Errors	Current Lattice Errors	Lattice Errors Phase I	Lattice Errors Phase II
ε_K ($\bar{K}K$ mixing)	$\mathrm{Im}V_{td}^2$	$\hat{B}_K$	10%	20%	12%	5%
ΔM_d ($\bar{B}B$ mixing)	$|V_{td}|^2$	$f_{B_d}^2 B_{B_d}$	6%	30%	16%-26%	8%-10%
$\Delta M_d/\Delta M_s$	$|V_{td}/V_{ts}|^2$	ξ^2	—	12%	8%	6%
$B\to(^{\rho}_{\pi})l\nu$	$|V_{ub}|^2$	$\langle^{\rho}_{\pi}|(V{-}A)_\mu|B\rangle$	7%	15%	10%-13%	5.5%-6.5%

TABLE 1   Impact of improved lattices on the determination of CKM matrix elements.

[2] Lattice QCD Executive Committee, "Overview of Lattice Gauge Theory Infrastructure Project" (2004), http://www.physics.ucsb.edu/~sugar/hepap.pdf.
[3] Phillip Colella, Thom H. Dunning, Jr., William D. Gropp, and David E. Keyes, eds., "A Science-Based Case for Large-Scale Simulation" (Washington, D.C.: DOE Office of Science, July 30, 2003).
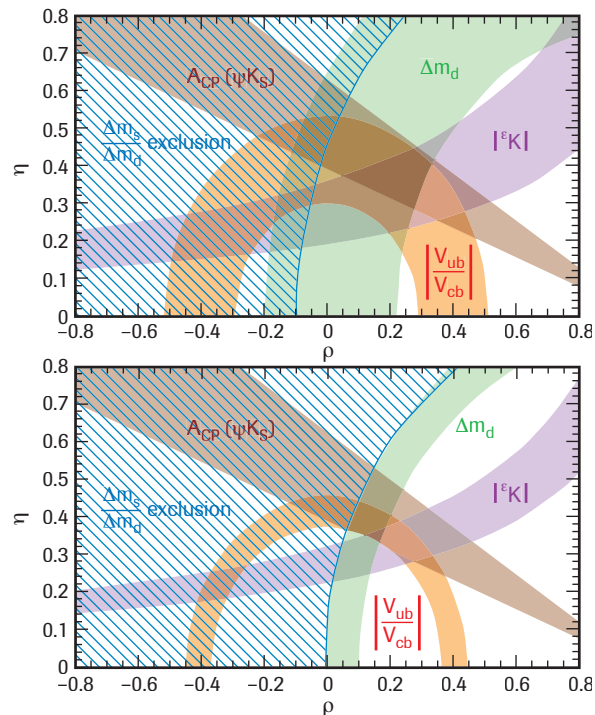
lattice calculations must keep pace with the experimental measurements."

In addition to Toussaint and Sugar, the MILC Collaboration includes Christopher Aubin and Claude Bernard of Washington University, Tommy Burch of the University of Regensburg, Carleton DeTar and James Osborn of the University of Utah, Steven Gottlieb of Indiana University, Eric Gregory of the University of Arizona, Urs Heller of Florida State University, and James Hetrick of the University of the Pacific.
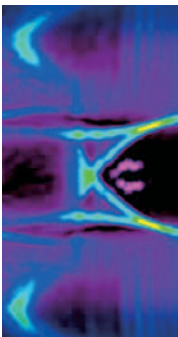
BEST BEAM

LASER WAKEFIELD TECHNOLOGY BLAZES PATH TO COMPACT PARTICLE ACCELERATORS

Researchers at the U.S. Department of Energy's Lawrence Berkeley National Laboratory have taken a giant step toward realizing the promise of laser wakefield acceleration, by guiding and controlling extremely intense laser beams over greater distances than ever before to produce high-quality, energetic electron beams.

The experimental results were then analyzed by running the VORPAL plasma simulation code on supercomputers at NERSC. The results were published in the September 30, 2004 issue of *Nature* magazine,[1] which chose an image from the NERSC simulations for the magazine's cover (Figure 1).

For a quarter of a century physicists have been trying to push charged particles to high energies with devices called laser wakefield accelerators. In theory, particles accelerated by the electric fields of laser-driven waves of plasma could reach, in just a few score meters, the high energies attained by miles-long machines using conventional radiofrequency acceleration. Stanford's linear accelerator, for example, is two miles long and can accelerate electrons to 50 GeV (50 billion electron volts). Laser wakefield technology offers the possibility of a compact, high-energy accelerator for probing the subatomic world, for studying new materials and new technologies, and for medical applications.

In plasmas, researchers have generated electric fields a thousand to ten thousand times greater than in conventional accelerators — but these large fields exist only over the short distance that a laser pulse remains intense; for tightly focused beams, that distance is typically only a few hundred micrometers (millionths of a meter). The resulting beams are of relatively poor quality, with particle energies so widespread that fewer than one percent have enough punch for scientific applications.

[1] Cameron Geddes, Csaba Toth, Jeroen van Tilborg, Eric Esarey, Carl Schroeder, David Bruhwiler, Chet Nieter, John Cary, and Wim Leemans, "High-quality electron beams from a laser wakefield accelerator using plasma-channel guiding," Nature **431,** 538 (2004).

FIGURE 1   Compact particle accelerators were featured as the cover story of the September 30, 2004 issue of *Nature.*

David Bruhwiler, John R. Cary, and Chet Nieter of the Tech-X Corporation of Boulder, Colorado, using the VORPAL plasma simulation code to model their results on supercomputers at NERSC. This modeling allowed the scientists to see the details of the evolution of the experiment.

"With VORPAL, one can see the laser pulse breakup and the injection of particles into the laser-plasma accelerator," said Cary, CEO of Tech-X Corp. and a professor of physics at the University of Colorado. "This allows one to understand how the injection and acceleration occur in detail so that the experiment's designers can figure out how to optimize the process."

VORPAL, jointly developed by Tech-X Corporation and the University of Colorado, is an object-oriented C++ framework for 1D, 2D, and 3D parallel electromagnetic simulation of charged fluids and particles on a structured grid. VORPAL is a relativistic, arbitrary dimensional, hybrid plasma and beam simulation code that contains utilities for data analysis and scripts for data visualization.

"Laser wakefield acceleration works on the principle that by sending a laser pulse through a gas to create

The Berkeley Lab researchers achieve high-quality beams by first shaping a channel through hydrogen gas with powerful, precisely timed laser pulses, then accelerating bunches of electrons through the plasma inside the channel (Figure 2). Because of the controlled accelerator length and the characteristics of the channel, there are several billion electrons in each bunch within a few percent of the same high energy, more than 80 MeV (80 million electron volts).

The work was done by the L'OASIS group (L'OASIS stands for Laser Optics and Accelerator Systems Integrated Studies), led by Wim Leemans of the Center for Beam Physics in Berkeley Lab's Accelerator and Fusion Research Division. To analyze their successful experiment, the group collaborated with
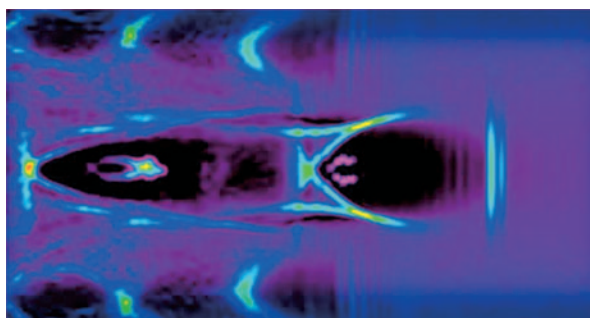

FIGURE 2   A plasma channel, denser toward the edges, guides the laser and allows it to form high-quality electron beams. As the laser pulse travels from left to right it excites a wake in the plasma, trapping and accelerating bunches of electrons to high energies. (Visualization by Cameron Geddes)

a plasma — separating negatively charged electrons from positively charged ions in the gas — some of the free electrons will be carried along in the wake of the plasma wave created by the laser," Leemans explains. "Imagine that the plasma is the ocean, and the laser pulse is a ship moving through it. The electrons are surfers riding the wave created by the ship's wake."

Unfortunately, simply punching a laser pulse through a plume of gas makes for a very short trip. "The acceleration distance is limited to what's called the Rayleigh length — the length over which the laser remains focused — which is determined by the wave-length of light and the size of the spot to which the beam is focused," says Leemans. "For optimum acceleration, you have to keep the wake going for many Rayleigh lengths, until just before the electrons start to get ahead of the wave and lose energy" — a distance otherwise known as the dephasing length.

To achieve this extended acceleration length, Leemans, working with his graduate student Cameron Geddes and other colleagues, implemented what he refers to as "the plasma analogue of an optical fiber" (Figure 3). The basic principle of this technique was developed a decade ago, using less powerful lasers, by Howard Milchberg from the University of Maryland, then improved upon by Leemans and his former student Pavel Volfbeyn in 1998. But it took nearly another six years before the L'OASIS lasers and instrumenta-tion were sufficiently advanced to try guiding extremely intense laser pulses with the same method.

By tailoring the plasma channel conditions and laser parameters, the researchers are first able to achieve clean guiding of laser beams of unprecedented high intensity, while suppressing electron capture. This paves the way for using laser-powered plasma chan-nels as ultra-high-gradient accelerating structures.

Next, by using even higher peak powers, plasma waves are excited that are capable of picking up background
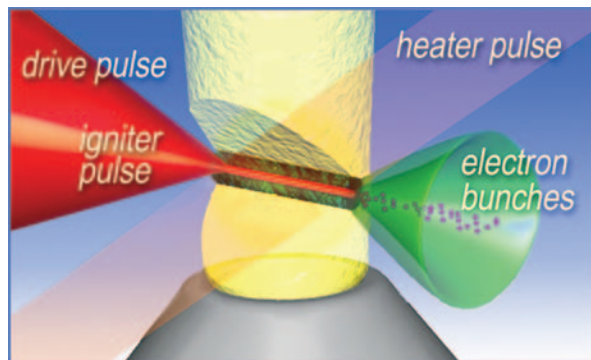
FIGURE 3   An igniter laser pulse forms a "wire" of plasma in a plume of hydrogen gas; a heater pulse expands the wire to a plasma channel; the drive pulse accelerates bunches of electrons inside the channel to nearly uniform high energy. (The green cone represents associated terahertz radiation.)

plasma electrons, rapidly accelerating them in the wake's electric field, then finally subsiding just as the surfing electrons reach the dephasing length, when they are on the verge of outrunning the wake.

How is it that the electrons in the bunch are accel-erated to nearly the same energy? Following the experiments, Geddes and the other Berkeley Lab researchers, with the help of their Tech-X collabora-tors, modeled the evolution of the high-quality elec-tron bunches on supercomputers at NERSC. They found that the formation of large bunches of elec-trons with nearly the same high energy is due to sev-eral factors. Most critical is matching the long accel-eration path to the dephasing length. Other factors include plasma density and laser power, and how the laser pulse evolved during its propagation in the plasma channel.

"The plasma-channel technique is an essential step toward developing compact laser wakefield accelera-tors with multiple stages, which can produce focused, ultrafast, high-energy bunches of electrons to com-pete with state-of-the-art machines using conven-tional radio-frequency acceleration," says Leemans.

THEY'RE COOL, BUT ARE THEY USEFUL?

COMPUTATION SHEDS LIGHT ON NEW NANOSTRUCTURES

Start with a quantum dot — a conglomeration of a few hundreds or thousands of atoms in which it is possible to control a single electron. Attach to the dot four rods of another material which are electronically coupled to the central quantum dot, then have the rods start branching. You're creating nanocrystals in various shapes that may have useful electronic or optical properties. But how do you find out what those properties are?

Ask the computational scientists on your team.

That's what Paul Alivisatos and his colleagues did. Alivisatos, director of Berkeley Lab's Materials Sciences Division and professor of chemistry at the University of California, Berkeley, and his research team have demonstrated a general approach for fabricating inorganically coupled colloidal quantum dots and rods (Figure 1). These remarkable new branched structures, whose dimensions are measured in billionths of a meter, were described in the July 8, 2004 issue of *Nature.*[1]

Lin-Wang Wang and postdoctoral fellow Jingbo Li of Berkeley Lab's Computational Research Division have worked closely with the Alivisatos group, using advanced computational techniques to identify the properties of new nanostructures coming out of the laboratory. Calculating the electronic properties of the structures depends on which compounds are involved, how they are joined, the number and arrangement of their atoms, and their proximity to other structures.

Analysis of the current batch of nanocrystals suggests that some of the new composite forms may have future applications in quantum computers,
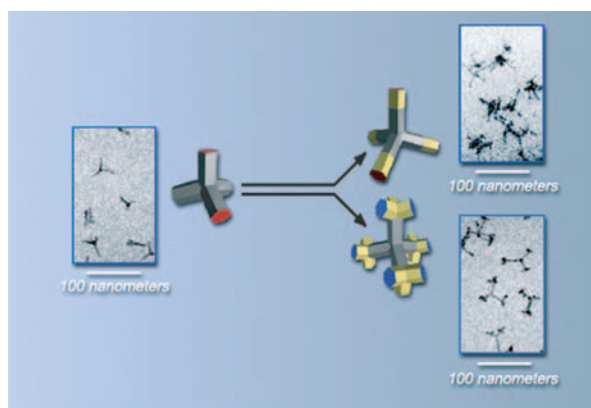
FIGURE 1   Examples of nanostructures made possible by the Alivisatos group's new method include tetrapods of cadmium selenide (left) extended with segments of cadmium telluride (upper right), which can be made to branch in turn (lower right).

which operate by controlling the coherence (relatedness) of small collections of electron wavefunctions. And the unusual electronic band structures of other forms have possible implications for photovoltaic energy conversion in solar cells.

Wang defines a nanostructure as "an assembly of building blocks on the scale where their properties become different from their bulk counterparts." A

---

[1] Delia J. Milliron, Steven M. Hughes, Yi Cui, Liberato Manna, Jingbo Li, Lin-Wang Wang, and Paul Alivisatos, "Colloidal nanocrystal heterostructures with linear and branched topology," Nature **430,** 190 (2004).

nanostructure may be composed of anywhere from a few hundred to a million atoms. While existing methods of calculation are well developed for very small collections of atoms, as in a molecule, or very large collections (virtually infinite) in bulk materials, the region in between is where calculations run into problems.

"One would like to use *ab initio* methods, which start by simply inputting a few atomic numbers and deriving the properties and behavior of the system directly," Wang says. "The Schrödinger equation theoretically makes that possible, but in practice calculating anything larger than a single helium atom necessarily involves simplifications and approximations."

Physicists use a modified *ab initio* technique called the local density approximation (LDA) to calculate the electronic structures of small systems (density refers to the electron charge density in specific regions of the sample).

"Even with LDA, you could spend months doing a straightforward calculation of the electron wavefunctions and their charge densities for a nanosystem with thousands of atoms," Wang says. The memory needed for calculation increases as the square of the number of atoms in the system — while the needed processing power increases as the cube! A million-atom system is far out of reach of LDA.

So Wang and his colleagues developed a hybrid method, which draws on LDA to determine the charge density in one small region of a crystal, then by "charge patching" extends the charge density to the entire system, in this case a quantum dot or nanorod.

To accurately model a nanostructure, Wang begins by "passivating" its surface with fractionally charged pseudo-hydrogen atoms — mathematical entities that assign the surface atoms the same reactivity as that of a real nanoparticle in suspension. The positions of the interior atoms are calculated with the valence force field method, which models the strength, elasticity, and direction of bonds among atoms.

Then LDA is used to determine the charge "motifs" around a number of representative atoms, including the surface pseudo-hydrogens. Through charge patching, the calculation is extended to include the entire nanostructure. In a final step, a "folded spectrum" method that Wang developed 10 years ago is used to determine the material's electronic states near the band gap, including the highest-energy state of the valence band (which in an ideal semiconductor is filled with electrons) and the lowest energy state of the conduction band (which is empty of electrons) (Figure 2).

The various ways that compounds can be assembled into different structures yield very different optical and electronic properties. "The charge-patching method allows us to model thousand-atom structures with *ab initio* accuracy in about one hour, with the help of the Seaborg supercomputer at NERSC," says

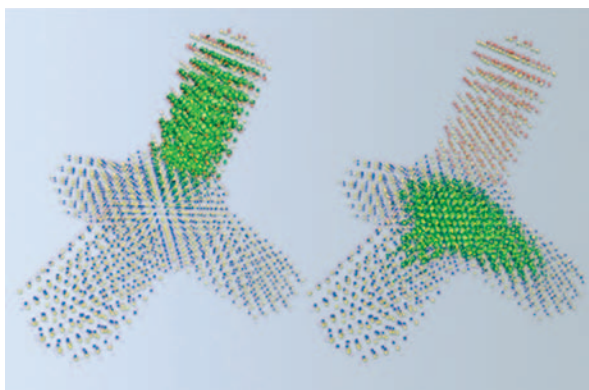FIGURE 2   NERSC calculations using local density approximation, the charge patching method, and the folded spectrum method yield atom-by-atom electronic maps of a tetrapod with one leg of cadmium selenide and three of cadmium telluride. On the left, green marks the conduction band's lowest energy state, which is physically separated in the structure from the valence band's highest energy state, shown in green on the right.

Wang. "This gives us unprecedented power to predict the electronic and optical properties of a given nanostructure."

The marriage of nanostructure fabrication with the ability to precisely calculate electronic properties opens possibilities for collaboration in more ways than one. "In this case the researchers had already synthesized the structures, and we were able to tell them the electronic properties and how they change with alterations," Wang says. "In the future, by modeling a proposed system in advance, we could help decide what's possible and how to control their results."

The ability to independently tune the properties of each component and the nature of their interactions will enable researchers to create electronic devices tailored to a variety of uses. The structures can be chemically manufactured in large quantities, for potential applications ranging from ultrafast transistors to artificial photosynthesis.

Wang's research on nanoscale electronic structure calculations has attracted international attention, including an Overseas Outstanding Young Researcher Award from the National Natural Science Foundation of China, China's equivalent of the U.S. National Science Foundation. While continuing to collaborate with the Alivisatos group, Wang is also making his expertise available to users of Berkeley Lab's new Molecular Foundry (foundry.lbl.gov). One of five DOE Nanoscale Science Research Centers, the Molecular Foundry is focused on the dissemination of techniques and methods for fabricating nanoscale patterned materials. As a computational science liaison to the Foundry's affiliated theory laboratory, Wang will help Foundry users apply LDA, empirical pseudopotential, and charge patching methods, and will continue working to develop new computational methods to meet the needs of the burgeoning fields of nanoscience and nanotechnology.

PREDICTING PLASMA PERFORMANCE

FUSION SIMULATIONS ADVANCE SPHEROMAK RESEARCH

In pursuing the potential of fusion energy, experimentalists try to understand precisely what is going on inside the fusion plasma so they can tune the system to improve the conditions that would lead to a sustained reaction within the heated plasma. If the temperature can be kept high enough and the energy contained, net power can be produced.

A major problem is that determining exactly what is happening inside a fusion plasma is very difficult experimentally. A conventional probe inserted into the hot plasma is likely to sputter and contaminate the plasma, leading to a loss of heat. Experimentalists must use nonperturbative diagnostics, e.g., laser scattering, and measurements with probes and magnetic loops around the edge of the plasma to deduce the plasma conditions and the magnetic field structures inside the plasma.

An important aid to the experiments is work undertaken with computational scientists to create detailed simulations of fusion plasmas. That's what scientists working on the Sustained Spheromak Physics

Experiment (SSPX, Figure 1) at Lawrence Livermore National Laboratory have done in collaboration with researchers at the University of Wisconsin-Madison.

The resulting simulations produced using the NIMROD code on NERSC's Seaborg supercomputer are very close to the observed results from actual runs on SSPX, giving the researchers confidence in the accuracy of the simulations and increased understanding of the physics in the spheromak (Figure 2).

FIGURE 1   The SSPX at Lawrence Livermore National Laboratory.
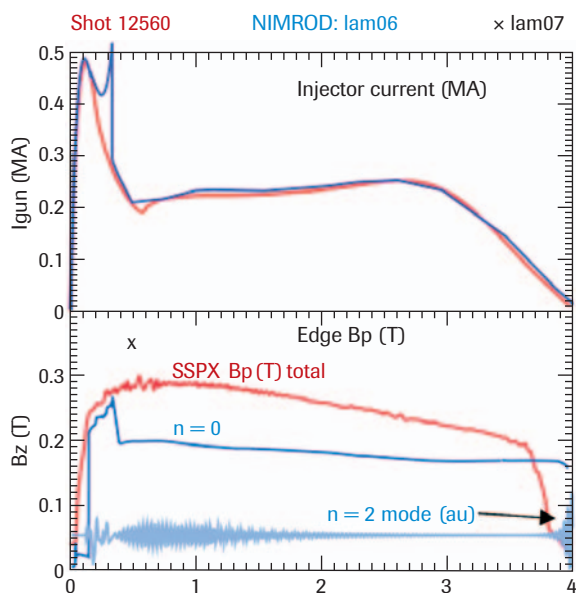
FIGURE 2   These graphs illustrate the improved match of NIMROD simulations (in blue) with results from an actual experiment (in red).

"These simulations are very important for supporting our experiments," said LLNL's Bruce Cohen, who has been computing at NERSC and its predecessors since the earliest days of the center. "The experimental team has been upgrading the capacitor bank used to drive the reactor, and our simulations confirm how careful tuning of the current pulses can improve plasma performance, such as achieving higher temperatures."

Cohen presented some of the group's latest results in an invited talk at the 2004 annual meeting of the American Physical Society's Division of Plasma Physics, held Nov. 15-19, 2004. In a close collaboration led by Carl Sovinec of the University of Wisconsin-Madison, the Wisconsin-LLNL group recently had a paper accepted for publication by Physical Review Letters.[1]

While the principal magnetic fusion approach focuses on the tokamak, fusion energy scientists are also revisiting the spheromak and other alternative concepts for attaining magnetic fusion. The SSPX spheromak at LLNL is a series of experiments designed to determine the spheromak's potential to efficiently create and confine hot fusion plasmas. The systems differ in that the tokamak's magnetic fields are generated by large, external magnetic coils surrounding the doughnut-shaped plasma (which make the tokamak more complex and expensive), while spheromaks confine hot plasma in a simple and compact magnetic field system that uses only a small set of external magnet coils. The necessary strong magnetic fields are generated inside the spheromak plasma by plasma currents and what's known as a magnetic dynamo.

A spheromak can be formed and sustained by injecting magnetic helicity and energy from a magnetized coaxial plasma gun (powered by a capacitor bank)

into a conducting shell or flux conserver. Although the physical spheromak design is simple, its dynamo activity produces plasma behavior that is extremely complex and more difficult to predict and control than that found in tokamaks.

"How high a temperature you can achieve and how well you can hold that heat in the plasma are two of the key points on the short list of things you want to understand and optimize in a fusion device," Cohen said. "Minimizing the energy leakage rate relaxes the requirements on how much energy you need to put into the fusion plasma to drive it and makes it easier to achieve sustained thermonuclear reactions."

With recent changes to their code, the Wisconsin-LLNL collaborators have created simulations with temperature histories — measured in milliseconds — that are closer to the temperature histories
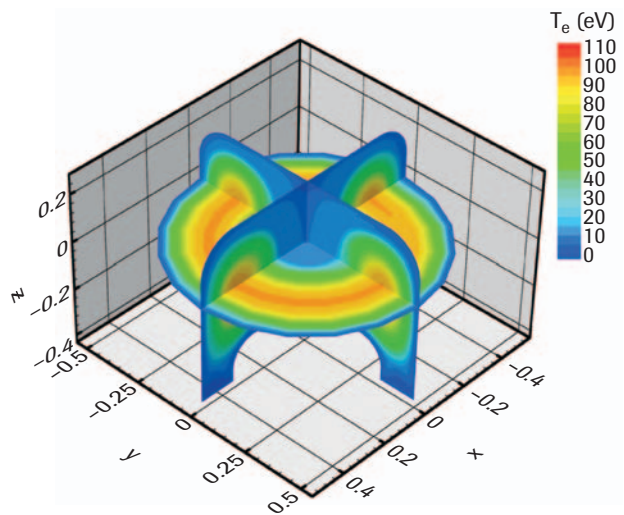

FIGURE 3 This partial-drive simulation of the evolution of conditions inside the spheromak shows the peak temperature of 98.8 eV occurring 1 millisecond into the simulation.

[1]C. R. Sovinec, B. I. Cohen, G. A. Cone, E. B. Hooper, and H. S. McLean, "Numerical investigation of transients in the SSPX spheromak," Phys. Rev. Lett. **94,** 035003 (2005).

observed in experiments (Figure 3). This follows the group's prior success in simulating the magnetics of the experiment. "We have the magnetic history captured pretty well," Cohen said. "Energy confinement is closely coupled to the detailed magnetics."

Plasma and energy confinement in the spheromak is quite sensitive to the quality and symmetry of the magnetic field lines. As long as the plasma and its confining magnetic field remain relatively axisymmetric with a large volume of closed, nested magnetic flux surfaces, a hot fusion plasma can be confined away from the walls of the device, and a high temperature can be maintained. However, due to magnetohydrodynamic (MHD) instability, small fluctuations and islands develop in the magnetic fields. This disrupts the axisymmetry, undercutting the confinement of the plasma and allowing heat loss. Finding ways to eliminate large-scale magnetic fluctuations and the associated islands would allow the experimenters to improve the energy and plasma confinement, and thereby increase the operating temperature.

"The simulations support experimental findings that controlling magnetic fluctuations is central to improving the quality of the magnetic flux surfaces and the energy confinement," Cohen noted in his talk at the APS meeting.

Future work will also address controlling magnetic fluctuations with a conducting insert along the geometric axis and additional physics improvements to the simulation model.

"We really need supercomputer resources to do these kinds of simulations, which involve massive calculations," Cohen said. Although the simulations cover only four milliseconds in physical time, there are of order 10^5 time steps involved in the simulations. As a result, the group ran each of their simulations in 50 to 80 shifts of 10 to 12 hours each on Seaborg, consuming more than 30,000 processor hours in each complete simulation; and multiple simulations were needed.
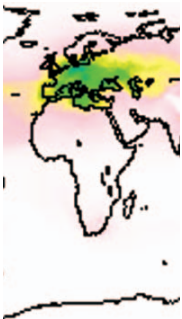
AEROSOL EFFECTS
SIMULATIONS SHOW THAT TINY PARTICLES CAN INFLUENCE CLIMATE

While greenhouse gases have been the focus of climate change research for a number of years, DOE's Tropospheric Aerosol Program (TAP) is now examining how aerosols (small particles of approximately micron size) affect the climate on both global and regional scales.

Scientists in the Atmospheric Science Division at Lawrence Livermore National Laboratory (LLNL) used NERSC's Seaborg supercomputer and LLNL's IMPACT atmospheric chemistry model to perform simulations showing the historic effects of sulfate aerosols at a finer spatial resolution than ever done before. Simulations were carried out for five different decades, from the 1950s through the 1990s (Figure 1).

The results clearly show the effects of the changing global pattern of sulfur emissions. Whereas in 1950 the United States emitted 41 percent of the world's sulfate aerosols, this figure had dropped to 15 percent by 1990, due to conservation and anti-pollution policies. By contrast, Asian emissions jumped six-fold during the same time, from 7 percent in 1950 to 44 percent in 1990.

Under a special allocation of computing time provided by the Office of Science INCITE program, Dan Bergmann, working with a team of LLNL scientists including Cathy Chuang, Philip Cameron-Smith, and Bala Govindasamy, was able to carry out a large number of calculations on 128 and 256 processors.

The objective was to assess the effects of anthropogenic (man-made) sulfate aerosols. The IMPACT model calculates the rate at which SO_2 (a gas emitted by industrial activity) is oxidized and forms particles known as sulfate aerosols. These particles have a short lifespan in the atmosphere, often washing out in about a week. This means that their effects on climate tend to be more regional, occurring near the area where the SO_2 is emitted.

To accurately study these regional effects, Bergmann needed to run the simulations at a finer horizontal resolution, as the coarser resolution (typically 300 km by 300 km) of other climate models is insufficient for studying changes on a regional scale. Livermore's use of the Community Atmosphere Model (CAM3), a high-resolution climate model developed at the National Center for Atmospheric Research (NCAR) with collaboration from DOE, allows a 100 km by 100 km grid to be applied. NERSC's terascale computing capability provided the needed computational horsepower to run the application at the finer level.

Depending on their composition, aerosols can either scatter or absorb sunlight, thereby cooling the Earth and acting as a counter to the warming effects of greenhouse gases. Greenhouse gases such as carbon dioxide are much more long-lived, so they stay in the atmosphere and have more uniform distribution. But since global warming has not increased as much as some computer models predict based on greenhouse gases alone, scientists have become more interested in aerosols and their possible role in countering some effects of greenhouse gases.

The atmospheric concentrations of both aerosols and greenhouse gases have increased over the past cen-
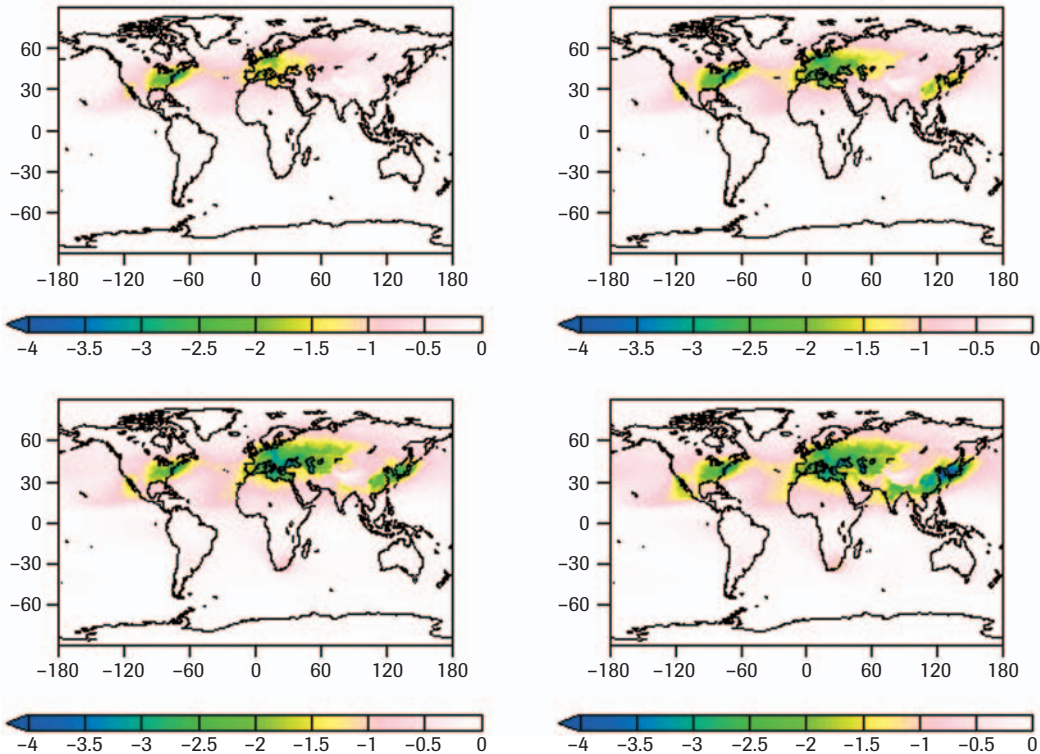
FIGURE 1   Predicted effects on shortwave radiation (watts/m^2) due to anthropogenic sulfate aerosols in the 1950s (top left), 1970s (top right), 1980s (bottom left), and 1990s (bottom right). Note the increased regional cooling effects in central Asia, India, and particularly eastern Asia.

tury and, because of their effects on shortwave and longwave radiation (which are partially offsetting), have presumably upset to some degree the thermal equilibrium of the climate system. While satellite measurements and field studies provide vital information to document the global and regional climate impacts by aerosols and greenhouse gases, accurate climate model simulations are an equally important tool in quantifying the radiative forcing as well as identifying and analyzing the climate response.

By running numerous calculations at finer resolution, the team was able to see the cooling effects by region, and to see more differences in the effects of varying concentrations.

The simulations also gave the team an opportunity to test new meteorological data generated using LLNL's version of NCAR's CAM3. This model calculates the various factors influencing climate, such as winds, temperatures, and precipitation. This data, which bordered on the terabyte scale, was also generated at NERSC.

The LLNL researchers then used the climate data to drive their chemistry model, known as IMPACT. By varying the SO_2 emissions levels while running IMPACT, the team was able to simulate different scenarios for each decade, analyze aerosol distributions from each simulation, and then examine their effects on radiative forcing at regional scales.

WHAT THE PLANCK SATELLITE WILL SEE
SIMULATING A MAP OF THE COSMIC MICROWAVE BACKGROUND

In the year 2007 the European Space Agency, with substantial NASA participation, will launch the Planck satellite (Figure 1) on a mission to map the cosmic microwave background (CMB), the remnant radiation believed to be an echo of the Big Bang that started the Universe.
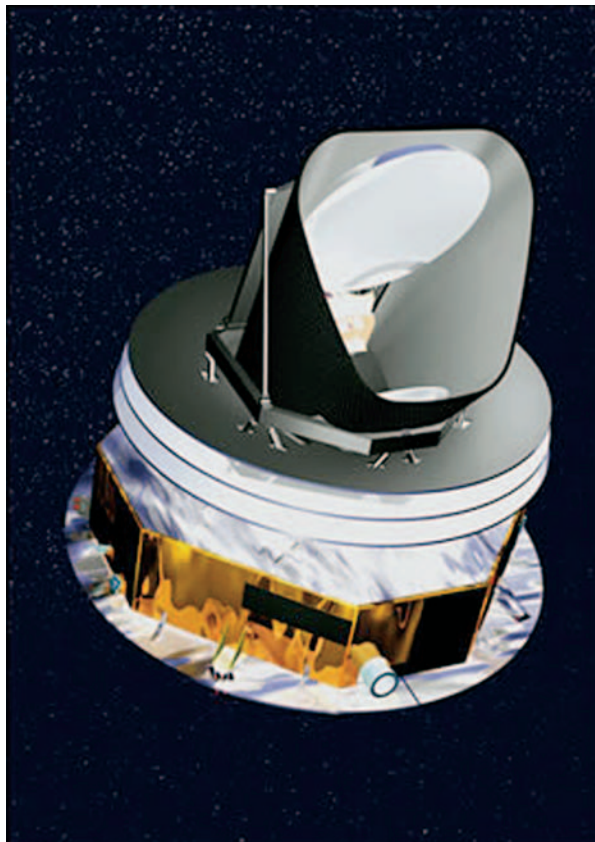

FIGURE 1   The Planck satellite is a collaboration between the European Space Agency and NASA, scheduled for launch in 2007. Planck's mission is to map temperature and polarization fluctuations in the cosmic microwave background with unprecedented resolution and sensitivity. (Image: Alcatel, ESA)

Planck is designed to map CMB temperature and polarization fluctuations with unprecedented resolution and sensitivity, but the enormous volume of data this will generate poses a major computational challenge. Can such a mountain of data be efficiently processed?

Affirmative.

Berkeley Lab researchers led by astrophysicist Julian Borrill of the Computational Research Division (CRD) have achieved what they say is a "significant milestone" by processing an entire year's worth of simulated Planck data at the single most CMB-sensitive frequency, producing a high-resolution CMB map in just two hours. Working on the Seaborg supercomputer at NERSC, the researchers ran their data on 6,000 of Seaborg's 6,080 processors at once, mapping 75 billion observations to 150 million pixels.

"We've shown that we can make maps even with the volume of data we will get from Planck," says Borrill. "There's no computational show-stopper to prevent us from working with the necessary resolution and accuracy. We'll be able to make CMB maps from the Planck data that astronomers will be able to use for science."

Collaborating with Borrill on the project were Radek Stompor of CRD and Christopher Cantalupo of the Space Sciences Laboratory at UC Berkeley, where Borrill and Stompor also hold appointments (Figure 2).
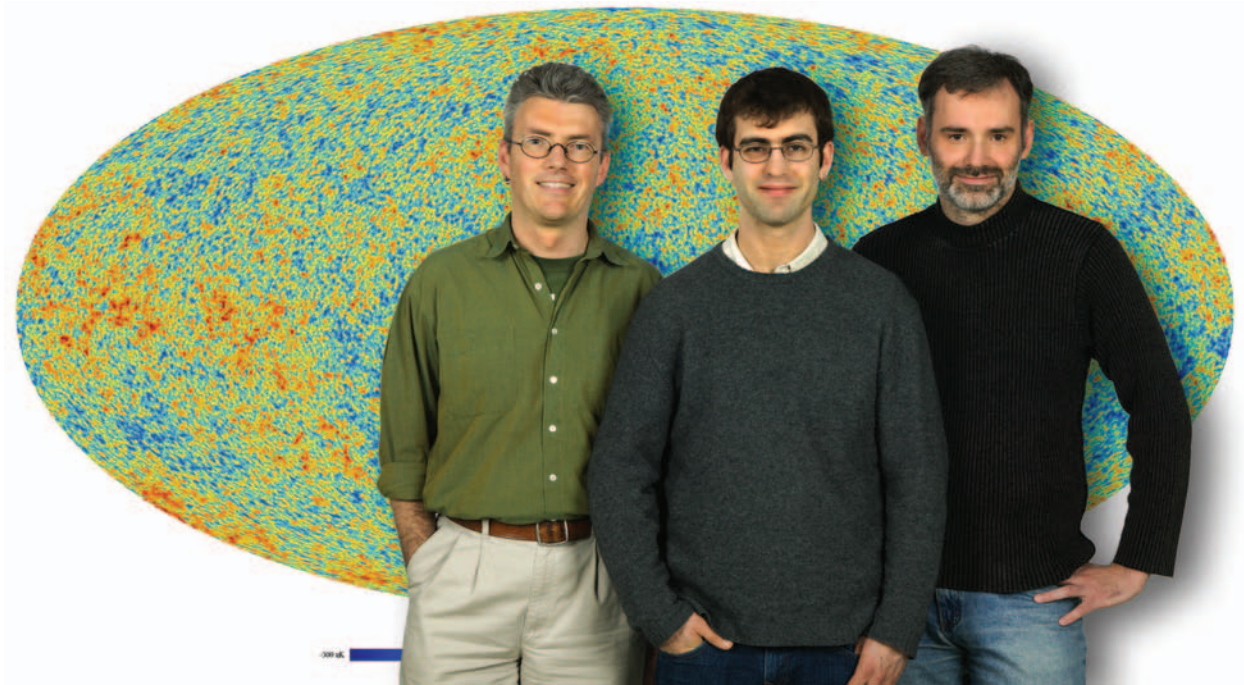
FIGURE 2   Julian Borrill, Radek Stompor, and Christopher Cantalupo used the Seaborg supercomputer at NERSC and a year's worth of simulated Planck satellite data to produce the high-resolution map of the CMB behind them. The colors show tiny temperature fluctuations in the CMB, echoing the Big Bang.

The three colleagues give much credit to the staff at NERSC, who helped shepherd them through all the potential bottlenecks of running 6,000 processors simultaneously, the first time data has ever been processed on this many of Seaborg's processors at once.

"The support that the NERSC staff gave us was incredible," says Borrill. "We've shown that we can make high-level, optimal CMB maps at the 217 megahertz frequency" — 217 million cycles per second — "which is the most CMB-sensitive of Planck's eight detectors and the frequency that will generate the most data. Such maps will be critical for maximizing the scientific return of the Planck mission."

According to the standard model of cosmology, the Universe was created 14.7 billion years ago in a titanic explosion dubbed the Big Bang. In the imme-

diate aftermath of this explosion, the Universe underwent inflation, an enormous expansion of space at a rate many times faster than the speed of light. Rapid expansion led to a cooling of the Universe.

Approximately 400,000 years after the Big Bang, temperatures had cooled enough for protons to capture electrons and form hydrogen atoms. At this point photons, no longer scattered by the dense fog of unattached electrons, were free to move about the Universe. Today, photons from the Big Bang have cooled to the very low energy of 2.7 Kelvin, but they permeate the entire Universe in the form of microwaves of about one millimeter in wavelength.

When the CMB was first accidentally discovered in 1965 by radio astronomers Arno Penzias and Robert Wilson, it was observed to be isotropic, meaning uniform in all directions — which ran contrary to theory.

Subsequent, increasingly sensitive observations revealed that the CMB is actually anisotropic after all: indeed, tiny temperature fluctuations in the earliest radiation were first detected in 1992 in an experiment led by Berkeley Lab physicist George Smoot aboard NASA's Cosmic Background Explorer satellite, COBE. These temperature fluctuations gave rise to the galaxies and clusters of galaxies that populate our Universe today.

The resolution of the COBE images was not sufficient to see any but the largest temperature fluctuations (Figure 3). Subsequent ground and balloon based-experiments, such as MAXIMA and BOOMERANG (whose data were also analyzed at NERSC), have measured smaller-scale temperature fluctuations and some polarization fluctuations as well. But these experiments measure only portions of the sky and are not sensitive enough to see all the spectral data that the CMB should harbor — data that could help answer a number of important cosmological questions about such things as the Hubble constant and the cosmological constant, important for understanding the expansion of the Universe, and the abundance of baryons, the class of massive fundamental particles that includes protons and neutrons.

The Planck satellite, named for Max Planck, the father of quantum theory, is designed to provide enough data to yield full-sky maps in eight frequency bands in the microwave regime between 30 and 857 gigahertz (billions of cycles per second). Such maps would boast an angular resolution of 5 to 10 arc minutes and a sensitivity of a millionth of a degree. (An arc minute is a 60th of a degree; the width of the sun or the full moon viewed from Earth is about half a degree. COBE's best angular resolution was about 7 degrees.)
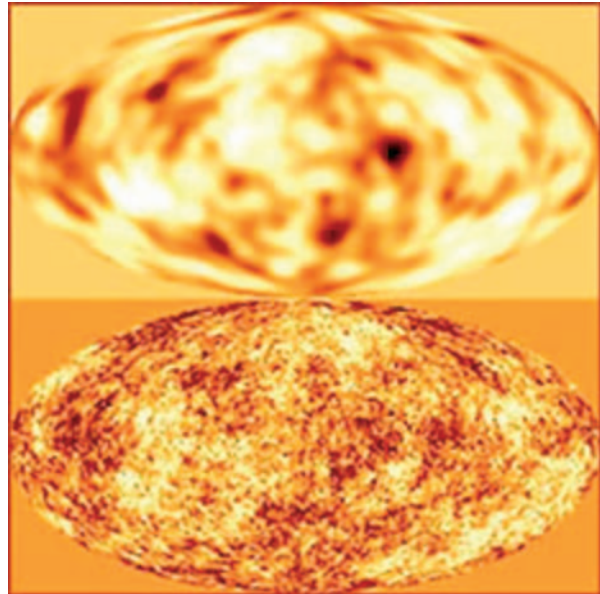
FIGURE 3   These CMB maps show COBE's angular resolution of 7 degrees (top) compared to a simulation of the 5 to 10 arc-minute resolution expected from the Planck satellite. (Images: NASA, ESA)

To produce these maps from the Planck data, Borrill and his colleagues have developed software called MADmap, which stands for Microwave Anisotropy Dataset map-maker. MADmap capitalizes on massively parallel computational capabilities to create millions of pixels from billions of observations.

"We made this map from simulated data that used all 12 of Planck's detectors at the 217-megahertz frequency," Borrill says. "If we can map at this frequency, we should be able to map for all the CMB frequencies that Planck will observe."

SEEING IS
UNDERSTANDING

VISUALIZATIONS HELP SCIENTISTS MAKE SENSE OF COMPLEX DATA

It is often difficult to understand or communicate information from complex or large datasets generated by computational simulations, experiments, or observational instruments. Scientific visualization is the computer science discipline that provides the techniques, algorithms, and software tools needed to extract and display important information from numerical data.

The Berkeley Lab/NERSC Visualization Group maintains a suite of commercial visualization software for NERSC users, provides assistance in using it, and conducts research and development to expand the capabilities of visualization tools. These pages provide a sampling of the results of their collaborations with NERSC researchers during the past year. Other examples of their work can be found on pages 8, 9, 17, 18, 56 (Figure 9) and 76, and at http://www-vis.lbl.gov/Vignettes/, where there are also links to several animations.
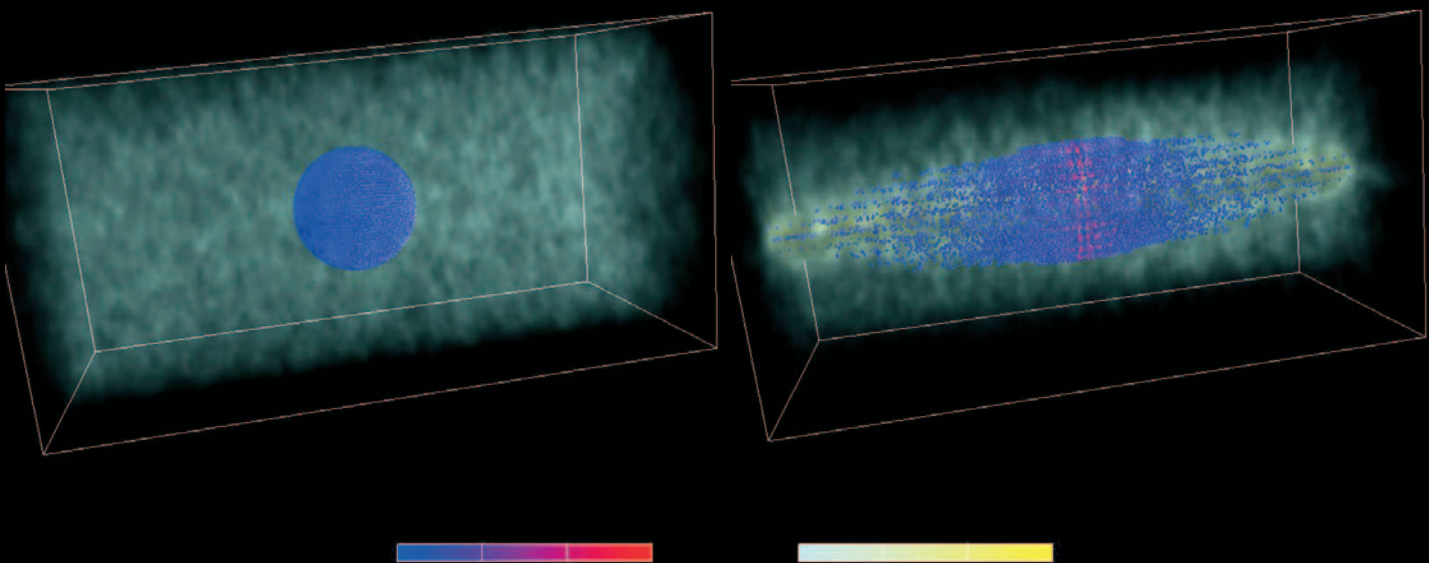

FIGURE 1   Electron cloud simulation with electrons rendered as volume density and protons rendered as particles.
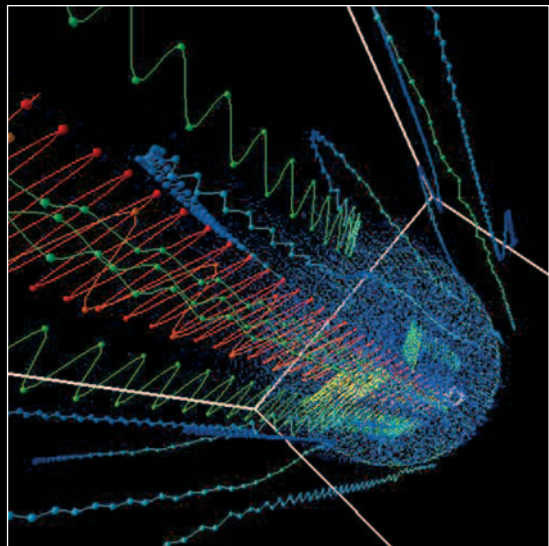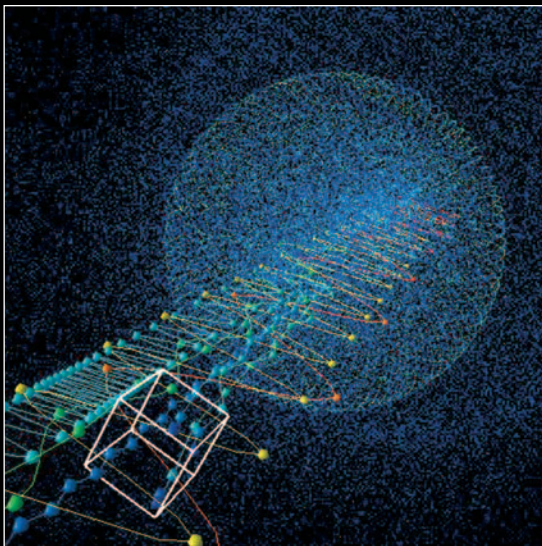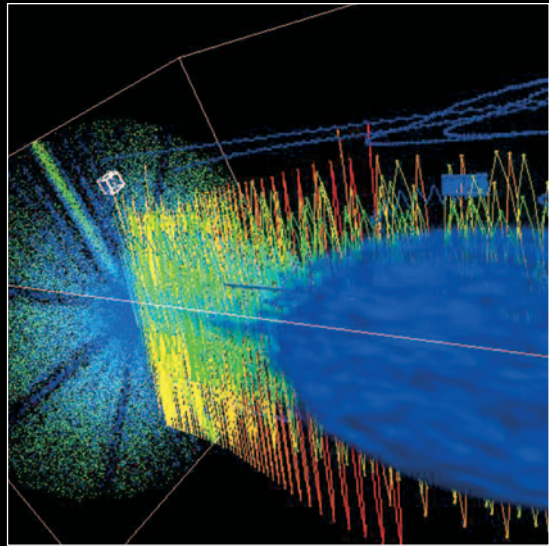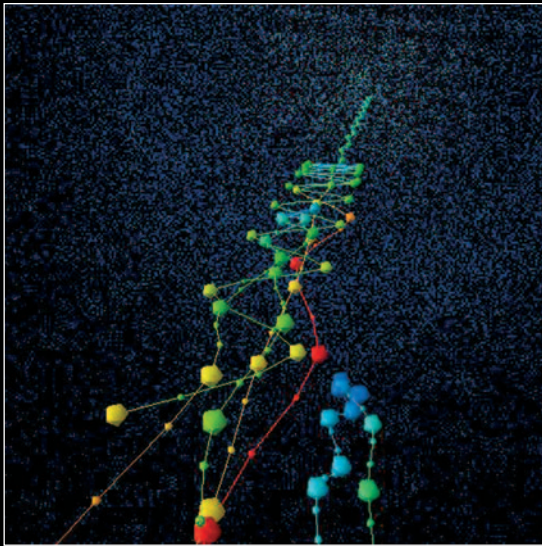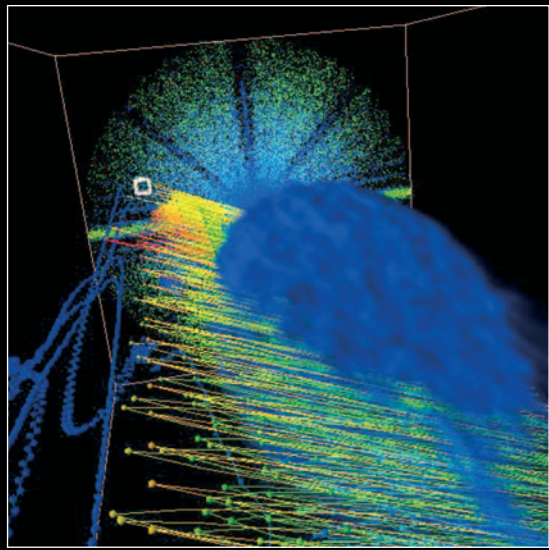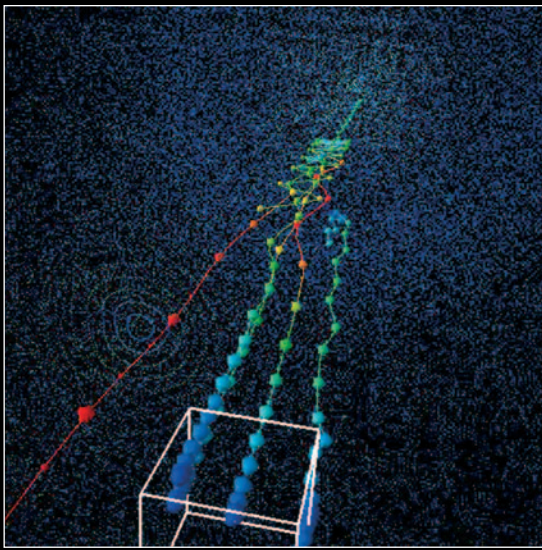
FIGURE 2   Trajectories of electrons are rendered here as splines colored by the magnitude of the velocity, while protons are rendered as particles. .

ELECTRON CLOUD VISUALIZATION

The SciDAC Advanced Computing for 21st Century Accelerator Science and Technology project, led by Rob Ryne of Berkeley Lab, is studying electron cloud instabilities that can disrupt the main proton accelerator beam. Collaborators Andreas Adelmann of the Paul Scherrer Institut in Switzerland and Miguel Furman of Berkeley Lab have developed a simulation tool that calculates the positions and velocities of protons and electrons, but the large and complex datasets are difficult to manipulate and understand.

The Visualization Group is working with this team to develop and apply several different approaches for exploration and analysis of accelerator modeling results. Using AVS/Express software has allowed the researchers to move beyond simple scatter plots and actually track the movement of groups of particles in three dimensions (Figures1 and 2). Two other tools, PartView and PPaint, were developed by the Visualization Group specifically for this type of research. PartView (Figure 3) provides a simple, interactive interface for routine data visualization and analysis. PPaint (Figure 4) enables researchers to "paint" selected particles in order to track their position, orientation, and motion through the beamline, both forwards and backwards in time.
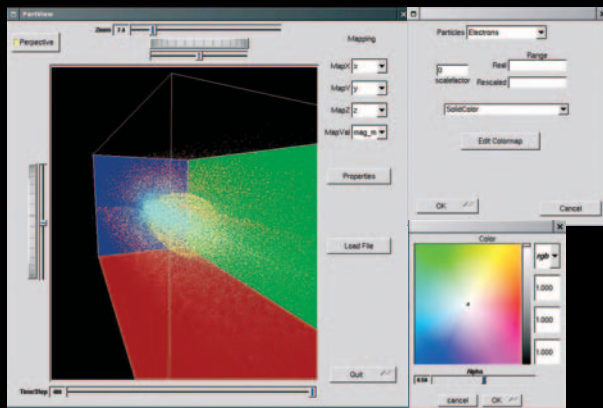
FIGURE 3   PartView provides an easy-to-use interface for routine analysis of accelerator simulation data.
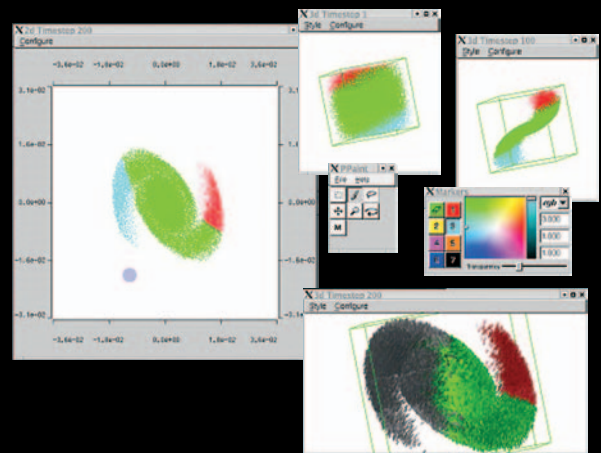
FIGURE 4   PPaint allows users to "paint" selected groups of particles in one of the 2D projected views, then view the selected particle groups in other projections. This approach facilitates understanding of time-varying particle behavior.

GETTING AMR DATA INTO EXISTING VISUALIZATION TOOLS

The Visualization Group is assisting partic-
ipants in the SciDAC Applied Partial
Differential Equations Center (APDEC) by
providing a variety of data analysis options
for understanding adaptive mesh refine-
ment (AMR) datasets. These options include
AMR Volume Renderer, ChomboVis, and
file format converters that enable scien-
tists to visualize their AMR data with their
choice of off-the-shelf applications (Figure
5). The converters are necessary because
most visualization algorithms assume a
uniform grid, while AMR uses a hierarchy
of grids of different cell sizes. The conver-
sion software converts the grid hierarchy
into a hexahedral unstructured mesh.
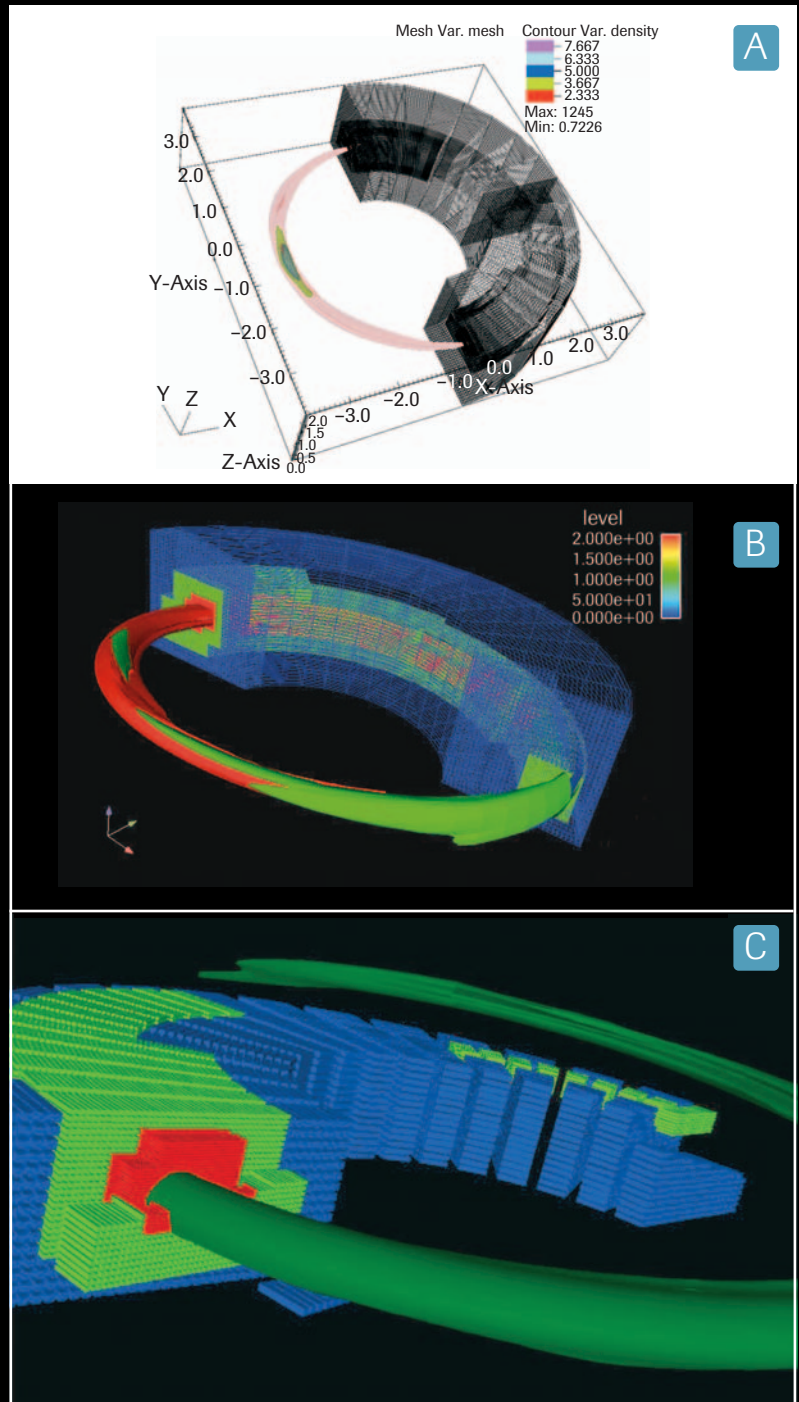
FIGURE 5   The same AMR dataset visualized using (a) VisIt,
(b) Ensight, and (c) AVS/Express. The dataset, produced by Ravi
Samtaney of Princeton Plasma Physics Laboratory using the Chombo
AMR framework, simulates a pellet of fuel being injected into a toka-
mak fusion reactor. The isosurfaces show the density of the pellet
material after it has been injected into the plasma from the interior
wall of the reactor.
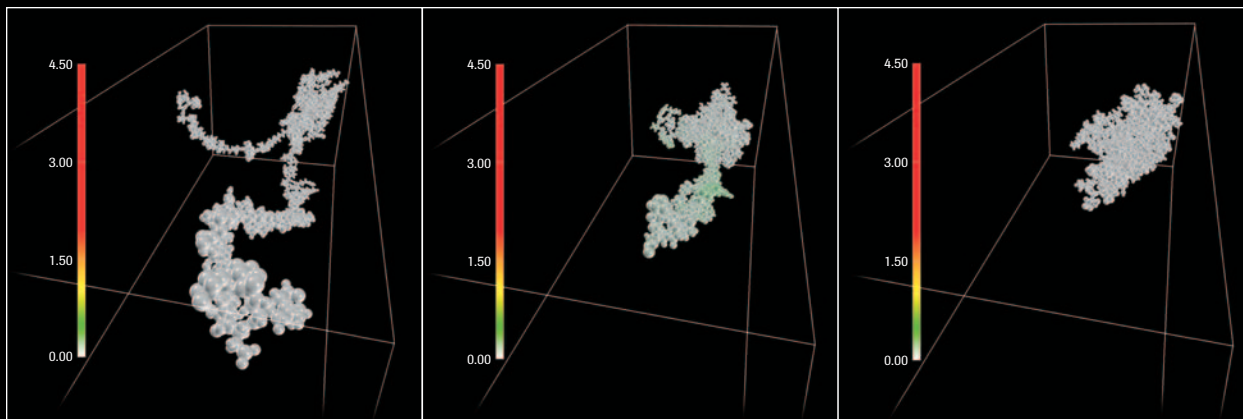
FIGURE 6   Calculating the minimum interaction energies for regions of a protein molecule and for the molecule as a whole results in a sequence of data that can be translated into images of protein folding.
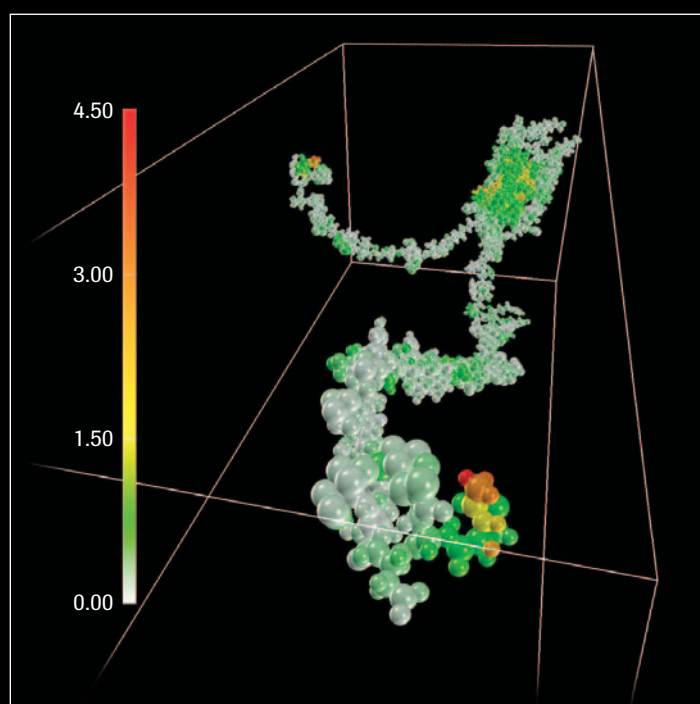
FIGURE 7   The color scale shows the speed at which atoms are repositioning themselves into lower-energy configurations.

SEEING HOW A PROTEIN FOLDS

Protein folding is the process by which a protein molecule — a simple unbranched chain of amino acids — assumes the complex shape, called its "native state," that determines its biological function. The native state is stable because its shape minimizes the interaction energy of the atoms in the molecule. Computational modeling of energy minimization is one way of identifying possible folding pathways for a particular protein.

The Visualization Group collaborated with Ricardo Oliva, Juan Meza, and Silvia Crivelli of Berkeley Lab to visualize the energy minimization and folding of protein T209 (Figure 6), calculated with OPT++ from an initial configuration created with ProteinShop. In this visualization, the color of the atoms shows the relative speed of their motion at any given point in the process (Figure 7). Several movies of these simulations can be found at http://www-vis.lbl.gov/Vignettes/.

A BETTER VIEW OF PLASMA FLOW

The Quasi-Poloidal Stellarator (QPS), an experimental fusion reactor that is expected to be built at Oak Ridge National Laboratory by 2008, is being designed with the help of a computer optimization model that uses an extensive suite of physics design tools. With its compact, highly shaped configuration (sometimes described as a twisted donut), the QPS is expected to eliminate the violent plasma disruptions common in conventional research tokamaks at high plasma pressures because it will have only a fraction of the plasma current. If successful, the QPS may usher in a new class of smaller and more economical fusion reactor designs.

Cristina Siegerist of the Visualization Group worked with Oak Ridge physicist Don Spong to resolve some difficulties in visualizing vector fields with AVS/Express. Using an AVS network that evolved out of the one Siegerist created, Spong produced this visualization of simulated plasma flow streamlines in the QPS (Figure 8).

0.0e+00	3.3e+03	6.7e+03	1.0e+04
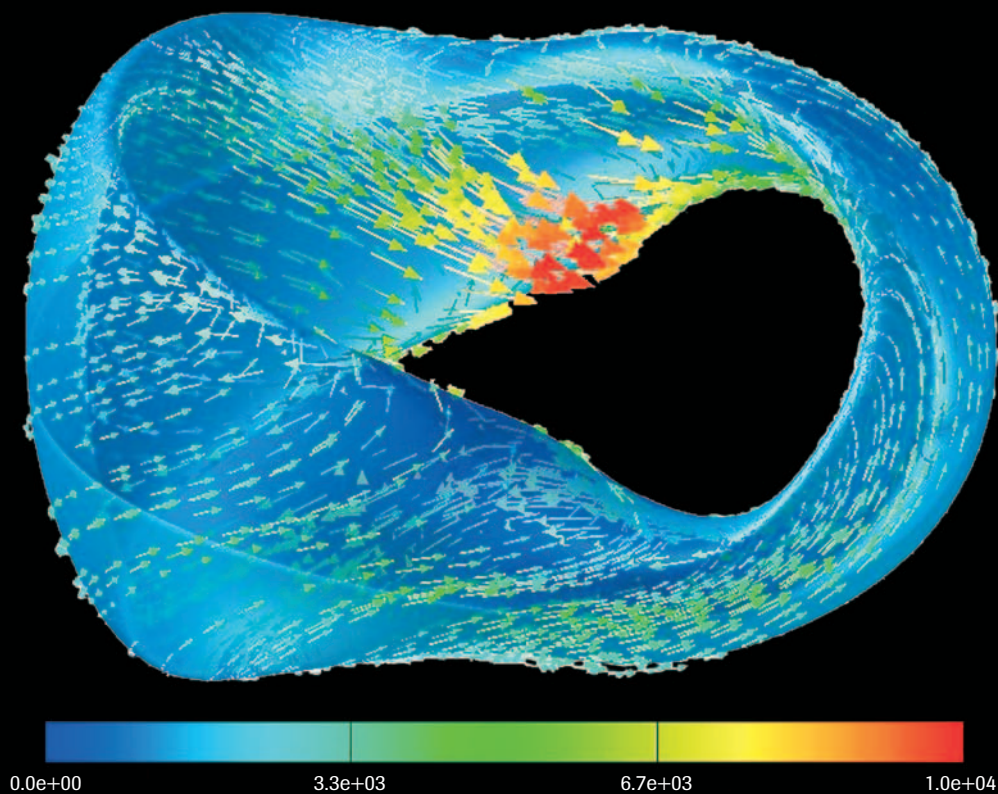
FIGURE 8   Simulated plasma flow streamlines in the QPS.

RESEARCH NEWS

This section of the Annual Report provides a representative sampling of recent discoveries made with the help of NERSC's computing, storage, networking, and intellectual resources. News items are organized by science category, with funding organizations indicated by acronyms in blue type. Organizational acronyms are spelled out in Appendix H.
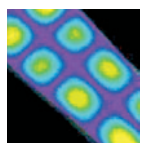
Accelerator Physics

Accurate Simulation of Beam–Beam Interactions

J. Qiang, M. A. Furman, and R. D. Ryne, "A parallel particle-in-cell model for beam–beam interaction in high energy ring colliders," J. Comp. Phys. **198,** 278 (2004). HEP, SciDAC

Accurate simulations of the electromagnetic interaction between two beams are needed to help optimize the luminosity in high energy accelerators. While beam–beam interactions have been studied for many years, the extreme computational cost has caused most previous studies to use simplified models. Qiang et al. have developed the first fully self-consistent model of colliding beams that can simultaneously handle all of the physical processes accurately. They tested the performance of the model on several architectures and applied it to studying the flip-flop instability in an electron–positron collider.

X-Band Linear Collider R&D

Z. Li, N. T. Folwell, L. Ge, A. Guetz, V. Ivanov, M. Kowalski, L.-Q. Lee, C.-K. Ng, G. Schussman, R. Uplenchwar, M. Wolf, and K. Ko, "X-band linear collider R&D in accelerating structures through advanced computing," *Proceedings of the 9th European Particle Accelerator Conference* (2004). HEP, SciDAC

A major computational effort is addressing key design issues in the high gradient accelerating structures for the proposed X-band linear collider, GLC/NLC (Global Linear Collider/Next Linear Collider). A new suite of parallel tools based on unstructured grids has been applied successfully to the design of complex cavities with high accuracy, and to realize end-to-end simulation of accelerator systems. The new simulation tools have played an important role in the R&D of X-band accelerating structures, in cell design, wakefield analysis, and dark current studies (Figure 1).
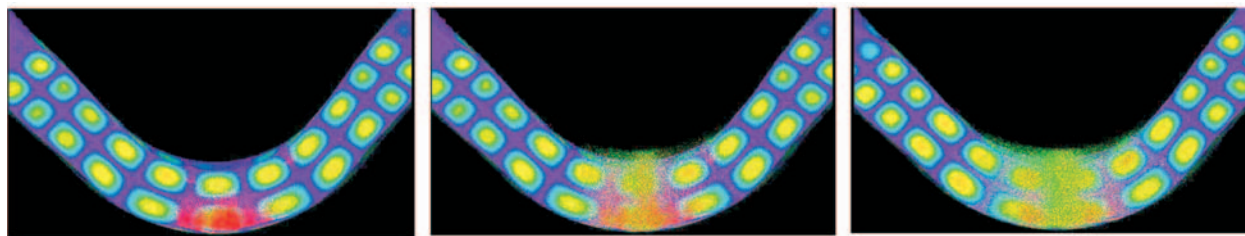
FIGURE 1 Evolution to steady state of surface emissions in NLC square waveguide bend at high power. The wave travels from left to right. Primary particles are red and secondaries are green.

A Multibunch Plasma Afterburner

R. Maeda, T. Katsouleas, P. Muggli, C. Joshi, W. B. Mori, and W. Quillinan, "Possibility of a multibunch plasma afterburner for linear colliders," Phys. Rev. STAB **7**, 111301 (2004). HEP, SciDAC, NSF

Recently a plasma wakefield afterburner concept was proposed for doubling the energy of a linear collider, in which single electron and positron bunches collide at the interaction point of the collider. Maeda et al. considered the possibility of extending the afterburner concept using multiple bunches. Simulation results indicated that an energy-of-collision/energy-of-linac ratio of 2.8 could be obtained with 4% energy spread and 0.29 relative luminosity by utilizing five drive bunches per accelerated bunch.

Proton Shock Acceleration

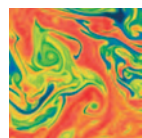L. O. Silva, M. Marti, J. R. Davies, R. A. Fonseca, C. Ren, F. S. Tsung, and W. B. Mori, "Proton shock acceleration in laser–plasma interactions," Phys. Rev. Lett. **92,** 015002 (2004). HEP, NSF, FCT

Recent experimental results show that ultra-intense laser–solid interactions can produce proton beams, with potential applications in proton imaging and proton therapy, but there is still debate about where the protons originate and about the acceleration mechanisms. Using 1D and 2D simulations, Silva et al. have identified two acceleration mechanisms: (1) proton acceleration due to the ambipolar fields arising in the free expansion of the strongly heated electrons at the front and rear of the target, and (2) proton acceleration in a collisionless, electrostatic shock formed at the front of the target.

Applied Mathematics

Crystallization *in Silico*

C. Yang, E. G. Ng, and P. A. Penczek, "Unified 3-D structure and projection orientation refinement using quasi-Newton algorithm," J. Struct. Biol. **149,** 53 (2005). BER, ASCR-MICS, NIH, LBNL

Single-particle electron cryomicroscopy (cryo-EM) offers a promising alternative to x-ray crystallography for discovering the structures of proteins and other macromolecules that cannot be crystallized without distorting their structures. Cryo-EM creates 2D images of thousands of randomly oriented "single particles" of the macromolecule. But constructing an accurate 3D model from noisy images of 2D single particles has been difficult as well as computationally infeasible for large molecules. Yang et al. found a new way of formulating the problem mathematically — an algorithm that simultaneously refines the 3D model while tightening the parameters for the orientation of the individual 2D projections used to reconstruct the model. The method is faster, more efficient, and more accurate than any of its predecessors.

Simulating Supernova Flames

J. B. Bell, M. S. Day, C. A. Rendleman, S. E. Woosley, and M. Zingale, "Direct numerical simulations of Type Ia supernovae flames II: The Rayleigh-Taylor instability," Astrophys. J. **608,** 883 (2004). ASCR-MICS, HEP, SciDAC, NASA, NSF

Accelerating a thermonuclear flame to a large fraction of the speed of sound (possibly supersonic) is one of the main difficulties in modeling Type Ia supernova explosions, which likely begin as a nuclear runaway near the center of a carbon-oxygen white dwarf. The outward propagating flame is unstable, which accelerates it toward the speed of sound. Bell et al. investigated the unstable flame at the transition from the flamelet regime to the distributed-burning regime through detailed, fully resolved simulations. At the low end of the density range, the instability dominated the burning (Figure 2), whereas at the high end the burning suppressed the instability. In all cases, significant acceleration of the flame was observed.
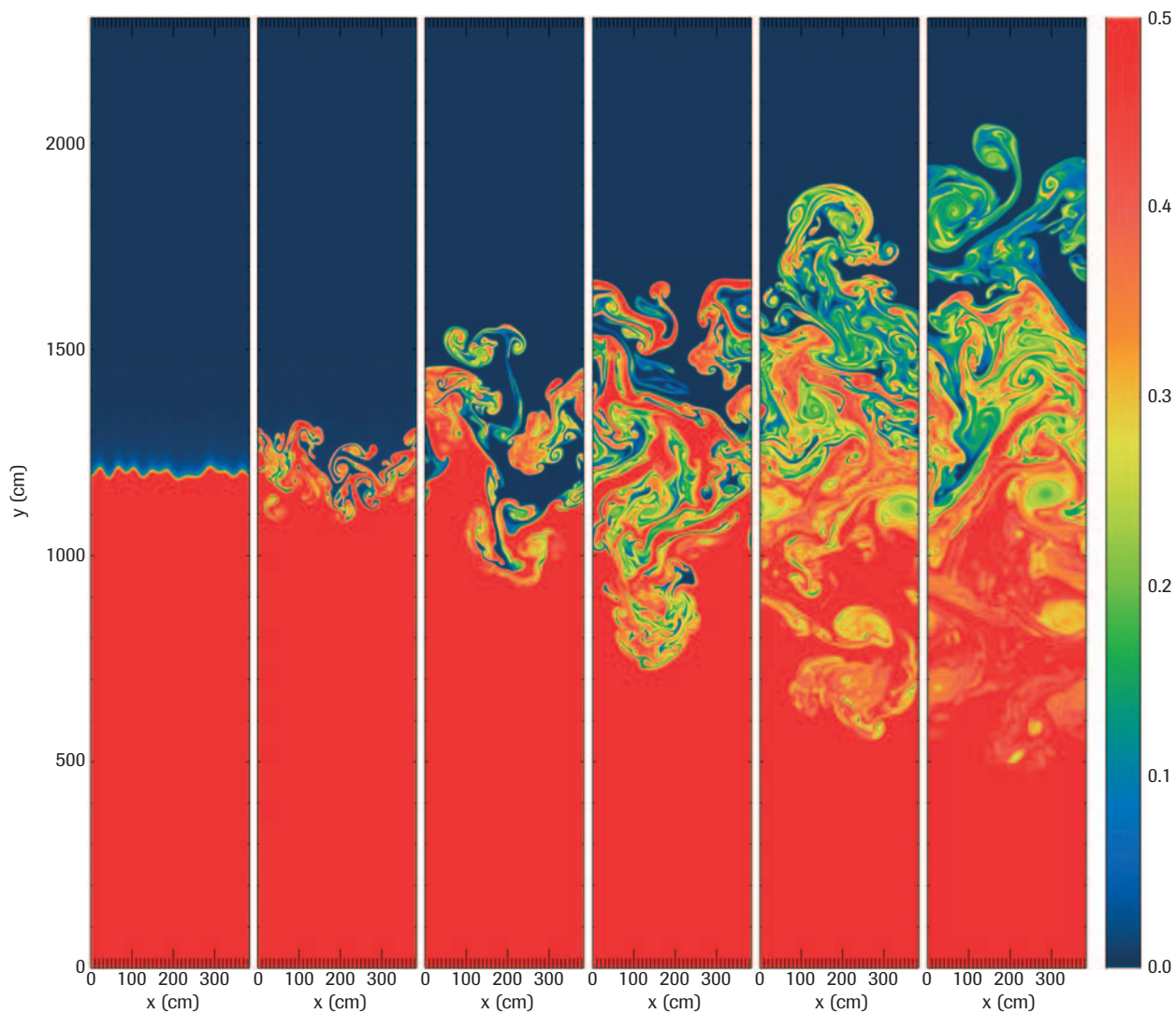
FIGURE 2   Carbon mass fraction for a 384 cm wide, 6.67×10^6 g cm^{-3} C/O flame shown every 1.6×10^{-3} s until 8.12×10^{-2} s. The fuel appears red (carbon mass fraction = 0.5), and gravity points toward increasing y. At this low density, the instability dominates over the burning, and a large mixed region develops.

Chemical Shifts in Amino Acids

Y. Yoon, B. G. Pfrommer, S. G. Louie, and A. Canning, "NMR chemical shifts in amino acids: Effects of environments in the condensed phase," Solid State Comm. **131,** 15 (2004). ASCR-MICS, SciDAC, NSF

Nuclear magnetic resonance (NMR) chemical shifts can be useful as fingerprints for detailing the structure and chemical composition of biomolecules, especially proteins and amino acids. Yoon et al. used the parallel electronic structure code PARATEC to calculate NMR chemical shifts in crystalline phases of the amino acids glycine, alanine, and alanyl-alanine. They explored the effects of environment on the chemical shifts in selected glycine geometries. In the crystalline and dilute molecular limits, the calculated NMR chemical shifts were attributed to intermolecular hydrogen-bonds and to dipole electric field effects, respectively.
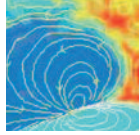
Astrophysics

Modeling Relativistic Pulsar Winds

J. Arons, D. C. Backer, A. Spitkovsky, and V. M. Kaspi, "Probing relativistic winds: The case of PSRJ07370-3039 A & B," in *Proceedings of the 2004 Aspen Winter Conf. on Astrophysics: Binary Radio Pulsars,* F. A. Rasio and I. H. Stairs, eds.; astro-ph/0404159 (2004). NP, NASA, NSF

In late 2003, radio astronomers discovered the first binary star system containing two neutron stars, both of which are observable as radio pulsars. The pulsars eclipse each other, even though the neutron stars are much too small to be the obscuring obstacles. The luminosity of pulsar A suggests that its relativistic wind is confining the magnetosphere of pulsar B to be 30% of the size it would have if B existed in isolation. Within several weeks of the discovery, Arons et al. created a simulation of A's eclipses based on synchrotron absorption in a shocked pulsar wind colliding with the magnetosphere of pulsar B (Figure 3). So far their model is the best developed model of the system, and their predictions are continuing to be confirmed by observations.
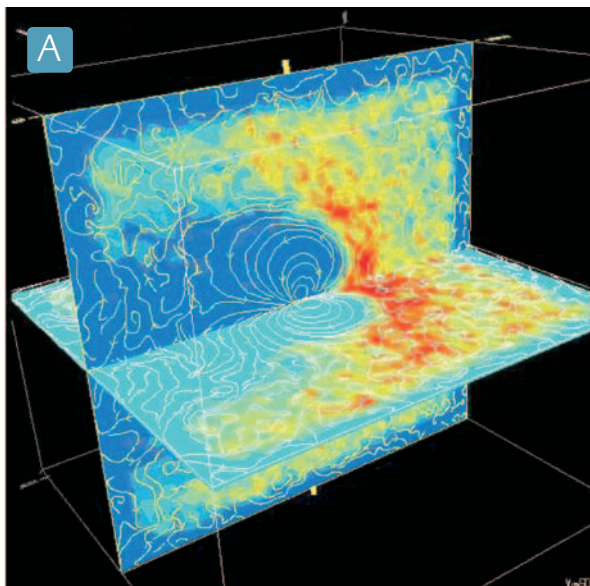
Disappearing Neutrinos Reappear

The KamLAND Collaboration, T. Araki et al., "Measurement of neutrino oscillation with KamLAND: Evidence of spectral distortion," Phys. Rev. Lett. **94,** 081801 (2005). NP, JMECSST, FEPCJ, KMSC

In previous experiments at the KamLAND underground neutrino detector, it was reported that the electron neutrinos were disappearing, i.e., oscillating into the neutrino flavors that are not detectable. With more data and more precise measurements, the disappearing neutrinos have been shown to be oscillating back into the detectable electron neutrinos. This is the most direct evidence yet of neutrino oscillation. These results constitute further proof that neutrinos have mass and that the Standard Model describing fundamental particles will need to be amended.
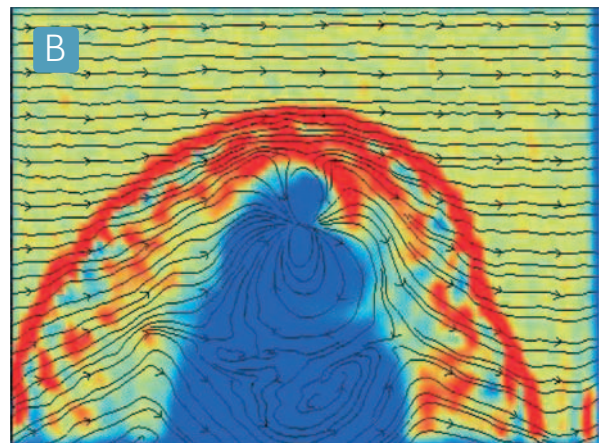
FIGURE 3   (a) Relativistic 3D particle-in-cell simulation of pulsar B's rotating magnetosphere immersed in an unmagnetized wind. (b) Relativistic 3D PIC simulation of B's rotating magnetosphere immersed in a magnetized wind — a snapshot of the equatorial plane.

High-Resolution Cosmology

C. L. Kuo, P. A. R. Ade, J. J. Bock, C. Cantalupo, M. D. Daub, J. Goldstein, W. L. Holzapfel, A. E. Lange, M. Lueker, M. Newcomb, J. B. Peterson, J. Ruhl, M. C. Runyan, and E. Torbet, "High-resolution observations of the cosmic microwave background power spectrum with ACBAR," Astrophys. J. **600,** 32 (2004). HEP, NASA, NSF, CARA

The Arcminute Cosmology Bolometer Array Receiver (ACBAR) is an instrument designed to produce detailed images of the cosmic microwave background (CMB) in three millimeter-wavelength bands. Kuo et al. employed new analysis techniques designed specifically for high-sensitivity ground-based CMB observations and reported the first measurements of CMB anisotropy from ACBAR — the highest signal-to-noise ratio observations to date. Overall, the resulting power spectrum appears to be consistent with the damped acoustic oscillations expected in standard cosmological models.

The Physics of Star Formation

M. R. Krumholz, C. F. McKee, and R. I. Klein, "Bondi accretion in the presence of vorticity," Astrophys. J. (in press); astro-ph/0409454 (2004). NP, NASA, NSF

The formation of high-mass stars remains one of the most significant unsolved problems in astrophysics. The classical Bondi-Hoyle formula for the accretion of gas onto a point particle is incomplete because it does not take vorticity into account, and even a small amount of vorticity can substantially affect accretion. Using a combination of simulations and analytic treatment, Krumholz et al. have provided an approximate formula for the accretion rate of gas onto a point particle as a function of the vorticity of the surrounding gas. Their results have potential implications for models of star formation in which protostars gain mass through a process of competitive accretion.

Neutrinos in Core-Collapse Supernovae

A. Juodagalvis, K. Langanke, G. Martinez-Pinedo, W. R. Hix, D. J. Dean, and J. M. Sampaio, "Neutral-current neutrino–nucleus cross sections for $A \sim 50$–65 nuclei," Nucl. Phys. A **747,** 87 (2005). NP, SciDAC, NSF, DRC, MCyT, ERDF, PFST

Neutrino–nucleus reactions play an essential role in the explosions of core-collapse supernovae, but inelastic neutrino–nucleus scattering has not yet been incorporated into supernova simulations, which may be one reason why these simulations often fail to yield explosions. Juodagalvis et al. studied neutral-current neutrino–nucleus reaction cross sections for Mn, Fe, Co and Ni isotopes, presenting the cross sections as functions of initial and final neutrino energies and for a range of supernova-relevant temperatures. These cross sections will allow improved estimates of inelastic neutrino reactions on nuclei to be used in supernova simulations.

Modeling Galaxy–Mass Correlations

A. Tasitsiomi, A. V. Kravtsov, R. H. Wechsler, and J. R. Primack, "Modeling galaxy–mass correlations in dissipationless simulations," Astrophys. J. **614,** 533 (2004). HEP, NSF, NASA, KICP

Understanding the processes that shape the clustering of dark matter and galaxies is one of the main goals of observational cosmology. Tasitsiomi et al. used high-resolution, dissipationless simulations of the concordance flat LCDM model to predict galaxy–mass correlations and compare them with the recent Sloan Digital Sky Survey weak-lensing measurements. They found that assigning a luminosity to a dark-matter halo of a certain maximum circular velocity by matching the simulated subhalo velocity function to the observed luminosity function leads to good agreement with the observed galaxy–mass correlation and its dependence on luminosity, if an observationally motivated amount of scatter between luminosity and circular velocity is introduced.
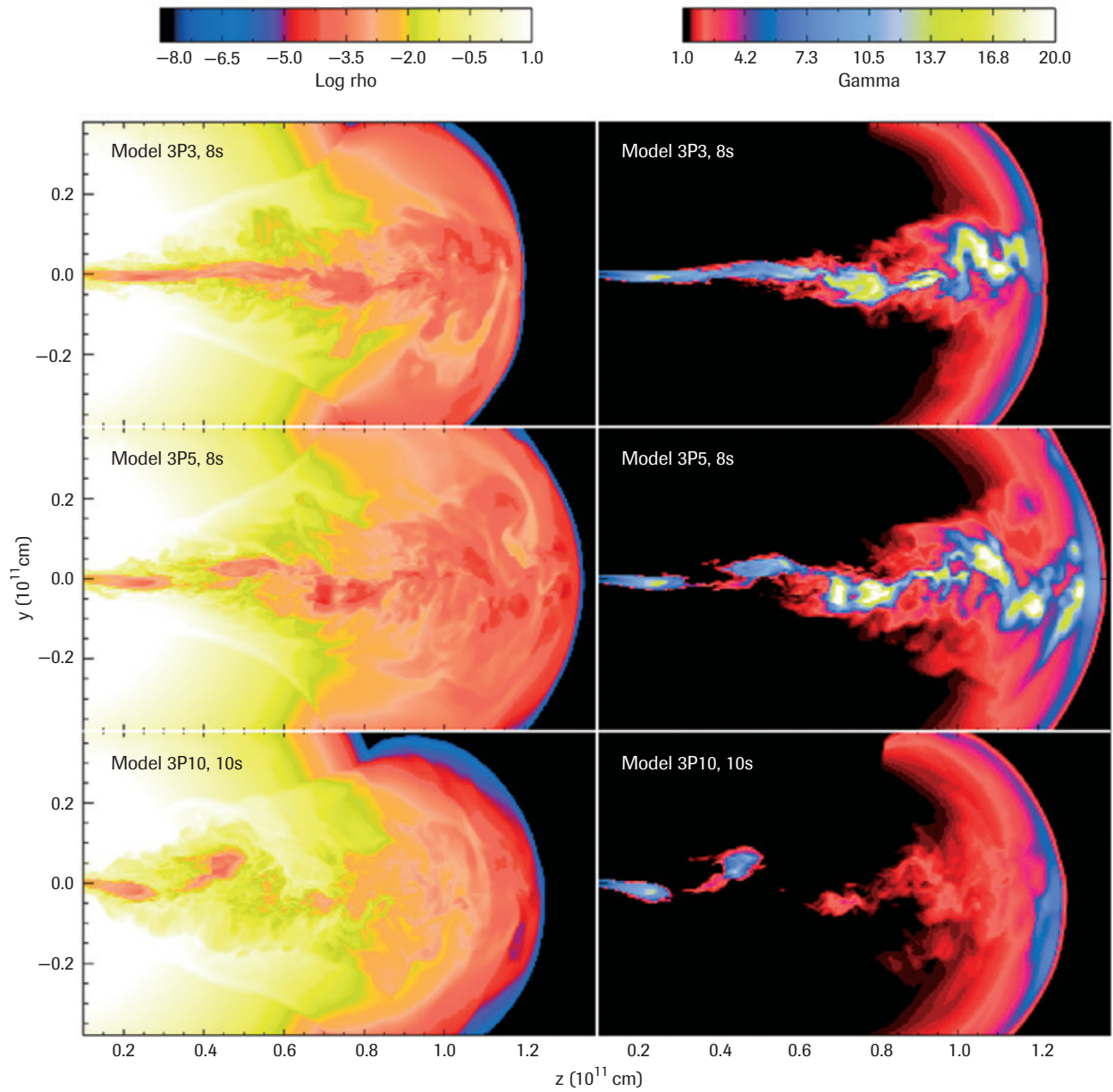
FIGURE 4   Jet angle sensitivity study. Slices of the precessing jet models are shown just after breakout. For model 3P3, the jet emerges relatively intact and might produce a GRB. The jets in models 3P5 and especially 3P10 dissipate their energy before escaping and are unlikely to produce GRBs, although they may still produce hard X-ray flashes.

High–Energy Stellar Jets

W. Zhang, S. E. Woosley, and A. Heger, "The propagation and eruption of relativistic jets from the stellar progenitors of gamma-ray bursts," Astrophys. J. **608,** 365 (2004). HEP, SciDAC, NASA

Zhang et al. have conducted 2D and 3D calculations of relativistic jet propagation and breakout in massive Wolf-Rayet stars, which are thought to be responsible for gamma-ray bursts (GRBs). As it erupts, the highly relativistic jet core is surrounded by a cocoon of less energetic but still moderately relativistic ejecta that expands and becomes visible at larger polar angles. These less energetic ejecta may be the origin of X-ray flashes and other high-energy transients. Calculations of jet stability showed that if the jet changes angle by more than 3° in several seconds, it will dissipate, producing a broad beam with inadequate Lorentz factor to make a common GRB, but possibly enough to make an X-ray flash (Figure 4).

Holes in Type Ia Supernovae

D. Kasen, P. Nugent, R. C. Thomas, and L. Wang, "Could there be a hole in Type Ia supernovae?" Astrophys. J. **610,** 876 (2004). NP, NASA

Type Ia supernovae arise from a white dwarf accreting material from a companion star. Soon after the white dwarf explodes, the ejected supernova material engulfs the companion star, which carves out a conical hole in the supernova ejecta. Kasen et al. used multi-dimensional Monte Carlo radiative transfer calculations to explore the observable consequences of an ejecta-hole asymmetry. They found that when one looks almost directly down the hole, the supernova is relatively brighter and has a peculiar spectrum characterized by more highly ionized species, weaker absorption features, and lower absorption velocities. Ejecta-hole asymmetry may explain the current spectropolarimetric observations of Type Ia supernovae.

Improving the Efficiency of Simulations

F. D. Swesty, D. C. Smolarski, and P. E. Saylor, "A comparison of algorithms for the efficient solution of the linear systems arising from multigroup flux-limited diffusion problems," Astrophys. J. Supp. **153**, 369 (2004). NP, SciDAC, NASA, ASAP

Multi-group flux-limited diffusion (MGFLD) is a popular method for modeling the multidimensional flow of radiation in a wide variety of astrophysical phenomena, but solving the sparse linear systems that arise in this method is computationally expensive. Swesty et al. compared the effectiveness of certain iterative sparse linear system methods for the solution of implicitly differenced MGFLD equations. They found that certain combinations of algorithms and preconditioners consistently outperform others for a series of test problems (Figure 5), and that the method of preparing the linear system for solution by scaling the system has a dramatic effect on the convergence behavior of the iterative methods.
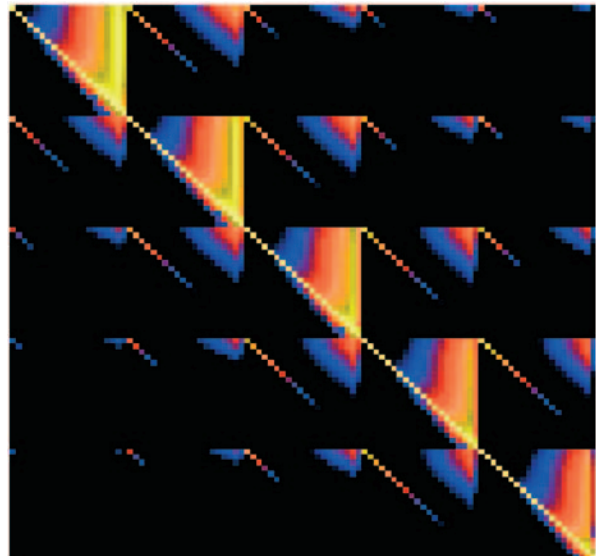
FIGURE 5   False color image of the absolute value of the upper left (100 × 100) corner of the inverse of the coefficient matrix for a test problem. Brighter colors indicate larger elements.

NERSC Users Win Awards for Scientific Accomplishments

During the past year several scientific organizations have honored NERSC users with prizes and awards in recognition of significant contributions to their areas of research.

Chandler

Landman

Souza

Woosley

Fung

Sposito

Wirth

Torn

Wang

The American Physical Society awarded four prizes to NERSC users:

David Chandler of the University of California, Berkeley and Lawrence Berkeley National Laboratory won the Irving Langmuir Prize in Chemical Physics "for the creation of widely used analytical methods and simulation techniques in statistical mechanics, with applications to theories of liquids, chemical kinetics, quantum processes, and reaction paths in complex systems."

Uzi Landman of the Georgia Institute of Technology was awarded the Aneesur Rahman Prize for Computational Physics "for pioneering computations that have generated unique insights into the physics of materials at the nanometer length scale, thereby fostering new theoretical and experimental research."

Ivo Souza of the University of California, Berkeley, received the George E. Valley Prize "for fundamental advances in the theory of polarization, localization and electric fields in crystalline insulators."

Stanford Woosley of the University of California,

Santa Cruz, received the Hans A. Bethe Prize "for his significant and wide ranging contributions in the areas of stellar evolution, element synthesis, the theory of core collapse and type Ia supernovae, and the interpretation of gamma-ray bursts — most notably, the collapsar model of gamma-ray bursts."

The American Geophysical Union honored two NERSC users, both of whom have joint appointments at the University of California, Berkeley and Lawrence Berkeley National Laboratory: Inez Yau-Sheung Fung was awarded the Roger Revelle Medal for laying the groundwork for the emerging area of biogeoscience and for "shaping our current view of the global carbon cycle"; and Garrison "Gary" Sposito received the Robert E. Horton Medal for his extensive contributions to subsurface hydrology, especially "for bridging the areas of aqueous geochemistry and physical hydrology."

Two NERSC users received the Presidential Early Career Award for Scientists and Engineers: Brian Wirth of UC Berkeley for his work in computational modeling and experiments to determine the mechanisms responsible for defects in metals, and Margaret Torn of Berkeley Lab for her innovative research on climate change and the terrestrial carbon cycle. And the National Natural Science Foundation of China honored Lin-Wang Wang of Berkeley Lab with an Overseas Outstanding Young Researcher Award for his role in developing computational methodologies to study nanosystems.

Chemical Sciences

Simulating an Ignition Process

R. Sankaran, H. G. Im, E. R. Hawkes, and J. H. Chen, "The effects of nonuniform temperature distribution on the ignition of a lean homogeneous hydrogen-air mixture," Proc. of the Combustion Institute **30** (in press). BES, SciDAC

To characterize the ignition process in homogeneous charge compression ignition (HCCI) engines, Sankaran et al. performed high fidelity simulations to study the effects of different temperature distributions on the autoignition of a turbulent homogenous mixture. The results showed that the temperature distribution and mixing rate have a major influence on the location of the first ignition sites and the subsequent combustion and heat release (Figure 6). It was found that the presence of a hot core gas leads to an increase in burn duration, while a cold core gas may lead to an undesirable slow combustion of the end gas. A predictive criterion to determine the two ignition regimes, the spontaneous propagation and the deflagration, was defined based on the ratio of their propagation speeds, suggesting a potential modeling strategy for large-scale simulations of HCCI engines.

Controlling the Primary Event of Vision

S. C. Flores and V. S. Batista, "Model study of coherent-control of the femtosecond primary event of vision," J. Phys. Chem. B **108,** 6745 (2004). BES, RC, ACS, FWH, Yale

The cis/trans photoisomerization reaction in rhodopsin is the first step in vision, and as an ultrafast photo-transduction mechanism, it has raised significant interest in the field of bioelectronics. Flores and Batista have computationally demonstrated the feasibility of coherently controlling this femtosecond reaction as modeled by an empirical Hamiltonian. Their approach involves selective photoexcitation of multiple vibrationally coherent wave packets by using two chirped femtosecond pulses. Control over product yields is achieved by externally changing the relative phases of the photoexcitation pulses and consequently affecting the interference phenomena between individual wave packet components.
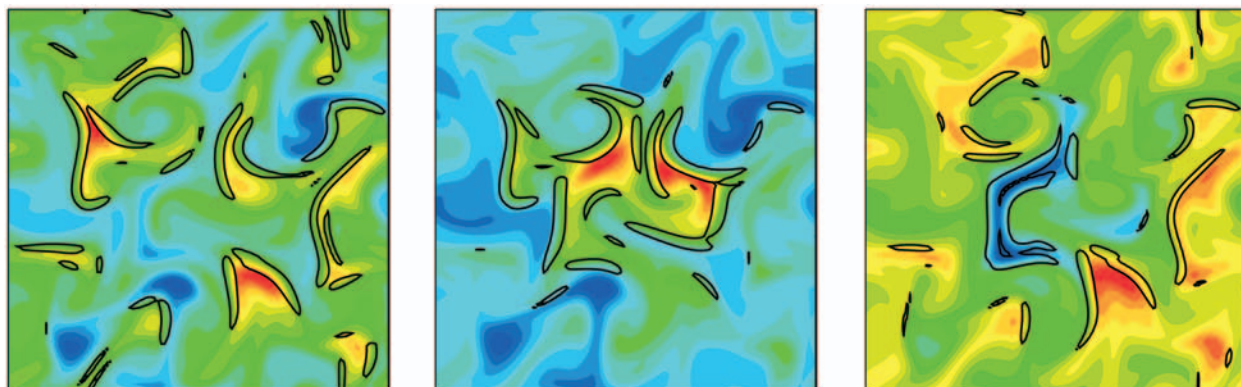
FIGURE 6   Application of the ignition regime criterion on the solution fields at 1% heat release. The color background indicates the temperature field, and line contours of $\beta = 1$ based on $C_\beta = 0.5$ denote the regions where a flame front will be formed.

Symmetric and Buckled Dimers on Silicon

Y. Jung, Y. Shao, M. S. Gordon, D. J. Doren, and M. Head-Gordon, "Are both symmetric and buckled dimers on Si(100) minima? Density functional and multireference perturbation theory calculations," J. Chem. Phys. **119,** 10917 (2003). BES, SciDAC

The silicon (100) surface undergoes reconstruction when cleaved — surface atoms dimerizing to form energetically more favorable bonds. However, the energy levels of these bonds, and whether they are symmetric or buckled (asymmetric), remain open questions. Head-Gordon et al. used density functional theory and multireference perturbation theory calculations to determine whether symmetric or buckled dimers are on the minimum (the energy "valley") of the Si(100) surface. Results varied as to which type of structure was higher in energy; however, overall results indicate that both symmetric and buckled dimers are located on Si(100) minima.

Enhanced Reaction Rates in Fuel Cells

Y. Xu, A. Ruban, and M. Mavrikakis, "Adsorption and dissociation of O_2 on Pt–Co and Pt–Fe alloys," J. Am. Chem. Soc. **126,** 4717 (2004). BES, NSF, 3M

Recently, several platinum–base metal (Pt–M) alloys have been found to possess greater activity than pure Pt for catalyzing the oxygen reduction reaction (ORR) in acidic electrolytes, a primary electrochemical reaction in low-temperature fuel cells. Mavrikakis et al. studied the adsorption of O and O_2 and the dissociation of O_2 on the Pt–M alloys and the Pt "skins" covering these alloys. They found that thin Pt skins, though less reactive than pure Pt for O_2 dissociation, are nonetheless more active toward the ORR because they are less poisoned by O and because they facilitate the activation of O and O-containing intermediates in bond-making elementary reaction steps in the ORR mechanism. The Pt skins also bind CO less strongly than pure Pt does, indicating that they may be more resistant to CO poisoning, a problem relevant to direct methanol fuel cells.

Theorists Challenge Experimental Model

L. D. Speakman, B. N. Papas, H. L. Woodcock, and H. F. Schaefer, "The microwave and infrared spectroscopy of benzaldehyde: Conflict between theory and experimental deductions," J. Chem. Phys. **120,** 4247 (2004). BES, SciDAC, UGA

Recently, it has been proposed that *ab initio* calculations cannot accurately treat molecules comprised of a benzene ring with a π-conjugated substituent, for example, benzaldehyde. Theoretical predictions of the benzaldehyde barrier to internal rotation are typically a factor of 2 too high in comparison to the experimental values of 4.67 (infrared) and 4.90 (microwave) kcal mol^{-1}. However, both experiments use Pitzer's 1946 model to compute the reduced moment of inertia and employ the experimentally observed torsional frequency to deduce benzaldehyde's rotational barrier. When Pitzer's model is applied to a system with a nonconjugated functional group, such as phenol, the model and theoretical values are in close agreement. Therefore, Speakman et al. concluded that (1) the model may not account for conjugation between the substituent and the π-system of benzene; (2) the experimental values of the benzaldehyde rotational barrier are misleading; and (3) the true rotational barrier should be closer to the theoretically extrapolated limit of 7.7 kcal mol^{-1}, based on coupled cluster theory.

Acoustic Speed Reduction Method

Y. Wang and A. Trouvé, "Artificial acoustic stiffness reduction in fully compressible, direct numerical simulation of combustion," Combustion Theory and Modelling **8,** 633 (2004). BES, SciDAC

Direct numerical simulation (DNS) studies of laminar or turbulent flames play a central role in our understanding of basic combustion phenomena, and different mathematical formulations are used for the DNS description of flow and combustion processes. Wang and Trouvé have developed a pseudo-compressibility method, called the Acoustic Speed Reduction (ASR) method, to allow for more efficient computations of slow flow problems using an explicit compressible

flow solver. They tested the performance of this method in a series of problems ranging from isothermal sound propagation to laminar premixed flame problems. In all tested cases, the ASR method proved successful at improving the computational efficiency while maintaining solution accuracy.
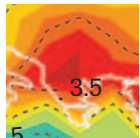
Climate Research

Mapping Future Heat Waves

G. A. Meehl and C. Tebaldi, "More intense, more frequent and longer lasting heat waves in the 21st century," Science **305,** 994 (2004). BER, NSF

Using a global coupled climate model, Meehl and Tebaldi determined that there is a distinct geographic pattern to future changes in heat waves. Model results for Europe and North America indicate that future heat waves in these areas will become more intense, more frequent, and longer lasting in the second half of the 21st century. Observations and the model show that present-day heat waves over Europe and North America coincide with a specific atmospheric circulation pattern that is intensified by ongoing increases in greenhouse gases, indicating that it will produce more severe heat waves in those regions in the future (Figure 7).

Aerosols Influence Radiative Fluxes

J. E. Penner, Y. Chen, and X. Dong, "Observational evidence of a change in radiative forcing due to the indirect aerosol effect," Nature **427,** 231 (2004). BER, NASA

Anthropogenic aerosols enhance cloud reflectivity by increasing the number concentration of cloud droplets, resulting in cooling from the "indirect aerosol effect," but it has been difficult to determine the impact of these indirect effects on radiative forcing. Penner et al. examined the effect of aerosols on cloud optical properties at two North American sites, determined the cloud optical depth required to fit the observed shortwave downward surface radiation,
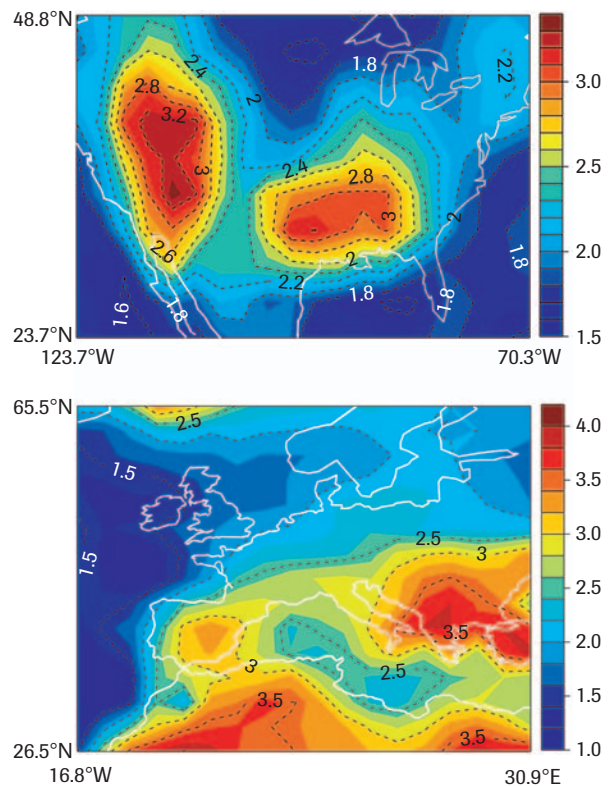
FIGURE 7   Future simulated changes in worst three-day heat waves are not uniformly distributed geographically but instead show a distinct pattern, as seen for North America and Europe from 2080 to 2099.

then used a warm-cloud adiabatic parcel model to simulate cloud optical depth. Good agreement between the simulation and observed surface radiation provided evidence that the indirect aerosol effect has a significant influence on radiative fluxes.

Faster Modeling of Long-Term Climate Evolution

Z. Liu, W. Lewis, and A. Ganopolski, "An acceleration scheme for the simulation of long term climate evolution," Climate Dynamics **22,** 771 (2004). BER

Understanding climate evolution on millennial (and longer) time scales will improve our understanding of past and future global climate changes. Liu et al. used a new acceleration scheme in a coupled ocean-atmosphere model to look at what forces long-term climate evolutions such the one that followed the Last Glacial Maximum 21,000 years ago. This new coordinated acceleration scheme reduces oceanic computation time as much as atmospheric time. For millennial climate evolution, the model produces reasonably good simulations, with an acceleration factor of about 5. For climate evolution of even longer time scales, the acceleration factor can be increased.

More Realistic Precipitation Models

J. Iorio, P. Duffy, B. Govindasamy, and S. Thompson, "Effects of increased resolution on the simulation of daily precipitation statistics in the U.S.," Climate Dynamics **23,** 243 (2004). BER

Changes in precipitation, both daily and seasonal, will account for some of the most important societal impacts of anthropogenic climate change. Increased spatial resolution models are needed to predict these possible effects. Iorio et al. analyzed global climate simulations (performing at a range of spatial resolutions) to assess the effects of horizontal spatial resolution on the ability to simulate precipitation in the continental U.S. By increasing the spatial resolution in the model (CCM3), they obtained more realistic representations of observed present-day precipitation, including both seasonal-mean precipitation and daily precipitation.

Parameterizing Subgrid Processes

J. N. Cole, H. W. Barker, D. A. Randall, M. F. Khairoutdinov, and E. Clothiaux, "Global consequences of interactions between clouds and radiation at scales unresolved by global climate models," Geophys. Res. Lett. **32,** L06703 (2005). BER

General-circulation models (GCMs), which test hypotheses regarding the Earth's climate, see the world in terms of 100–500 km grid spacings, which leaves many "subgrid" processes unresolved and in need of parameterization. Cole et al. tested the assumption that accurate domain averages are sufficient for satisfactory simulation of climatic change. They used a 2D cloud system resolving model (CSRM) with a horizontal 4 km grid spacing to investigate whether local interactions between clouds and radiation are important for successful GCM simulations. Results show that the unresolved interactions are at least as important as getting accurate domain averages.

Computer Science

Self-Tuning Software

Z. Chen, J. Dongarra, P. Luszczek, and K. Roche, "The LAPACK for Clusters project: An example of self adapting numerical software," *Proc. 37th Hawaii International Conf. on System Sciences* (2004). ASCR-MICS, NSF

As computing systems become more powerful and complex, tuning applications for high performance becomes a major challenge. The LAPACK for Clusters project has developed a self-adapting framework for the popular linear algebra software library that automates performance tuning, adapting to the user's problem and the computational environment in an attempt to extract near-optimum performance. Test results show that self-adaptation can come very close to matching the performance of a custom-tuned application, and with minimal overhead.

Mapping the Network Topology

D. Turner, B. Tong, and M. Sosonkina, "Mapping algorithms to the network topology in a portable manner," *Proc. PARA'04 Workshop on the State-of-the-Art in Scientific Computing* (2004). ASCR-MICS

Most scientific applications are unable to take advantage of the changing topology of the system network, resulting in loss of performance and scalability. The NodeMap project is being developed to automatically determine the network topology at runtime and provide an optimal mapping of a 2D or 3D algorithm. In a test of the concept underlying NodeMap, Turner et al. showed that the performance of a classical molecular dynamics code could be greatly enhanced by simply remapping the node arrangements to avoid saturation costs as much as possible. Parallel efficiency improved from 50% to 70% for 10,000,000 atoms on 1024 processors.

Compiling Irregular Applications

J. Su and K. Yelick, "Array prefetching for irregular array accesses in Titanium," IPDPS Workshop on Java for Parallel and Distributed Computing, Santa Fe, New Mexico, April 2004. ASCR-MICS, NSF, NDSEG

Compiling irregular applications, such as sparse matrix and unstructured mesh codes, is a challenging problem. In a compiler for Titanium, a dialect of Java designed for high performance computing, Su and Yelick developed an inspector-executor style optimization framework, which iterates through a loop once, collecting information about random memory access to local or remote memory, and then aggregates and schedules the required communication. A novel aspect is the use of a profile-based performance model to automatically select the optimal communication method. This advance allows application programmers to write Titanium code in a straightforward way, and get performance comparable to a popular hand-tuned library.

Fusion Energy

Electromagnetic Gyrokinetic Simulations

S. E. Parker, Y. Chen, W. Wan, B. I. Cohen, and W. M. Nevins, "Electromagnetic gyrokinetic simulations," Phys. Plasmas **11,** 2594 (2004). FES, SciDAC

A new electromagnetic kinetic electron δf particle simulation model has been demonstrated to work well at large values of plasma β times the ion-to-electron mass ratio. The model shows accurate shear Alfvén wave damping and microtearing physics. Zonal flows with kinetic electrons were found to be turbulent, with the spectrum peaking at zero and having a width in the frequency range of the driving turbulence. In contrast, with adiabatic electron cases the zonal flows were near stationary, even though the linear behavior of the zonal flow was not significantly affected by kinetic electrons. Zonal fields were found to be very weak, consistent with theoretical predictions for β below the kinetic ballooning limit.

Fast Wave Mode Conversion

J. C. Wright, P. T. Bonoli, M. Brambilla, F. Meo, E. D'Azevedo, D. B. Batchelor, E. F. Jaeger, L. A. Berry, C. K. Phillips, and A. Pletzer, "Full wave simulations of fast wave mode conversion and lower hybrid wave propagation in tokamaks," Phys. Plasmas **11,** 2473 (2004). FES, ASCR-MICS, SciDAC

Two full wave codes — a parallel version of the TORIC-2D finite Larmor radius code and an all-orders spectral code, AORSA2D — have been developed

which for the first time are capable of achieving the resolution and speed necessary to address mode conversion phenomena in full 2D toroidal geometry. These codes have been used in conjunction with theory and experimental data from the Alcator C-Mod to gain new understanding into the nature of fast wave mode conversion in tokamaks (Figure 8). Advanced scenarios in burning plasma devices such as the International Thermonuclear Experimental Reactor (ITER) can now be modeled with the new resolution capabilities.
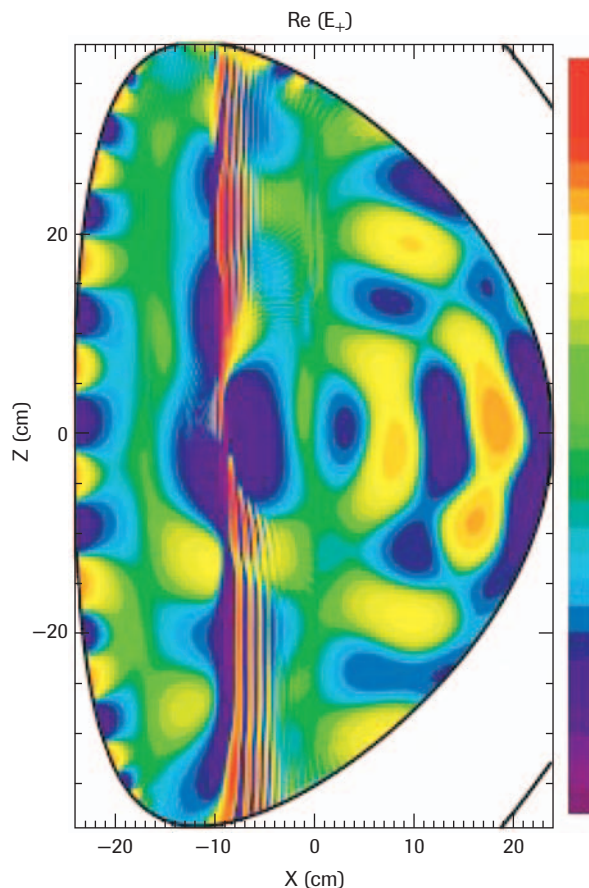
FIGURE 8 TORIC simulation of a D(^{3}He) mode conversion case in the Alcator C-Mod. Note the well resolved presence of the midplane ion Bernstein wave propagating to the left, the ion cyclotron wave above and below the midplane going to the right, and the large fast wave propagating from the right to the layer, with a small amount transmitting through.

Electron–Impact Ionization of Helium

Pindzola et al. have developed a nonperturbative lattice solution of the time-dependent Schrödinger equation that appears capable of yielding accurate cross sections for Coulomb four-body breakup. They applied this method to obtain single and double ionization cross sections for helium that are in excellent agreement with experimental measurements. The total wave function for the three-electron system was expanded in nine dimensions: three dimensions were represented on a radial lattice, and a coupled channels expansion was used to represent the other six dimensions, resulting in a simulation that captured all of the scattering processes (Figure 9).
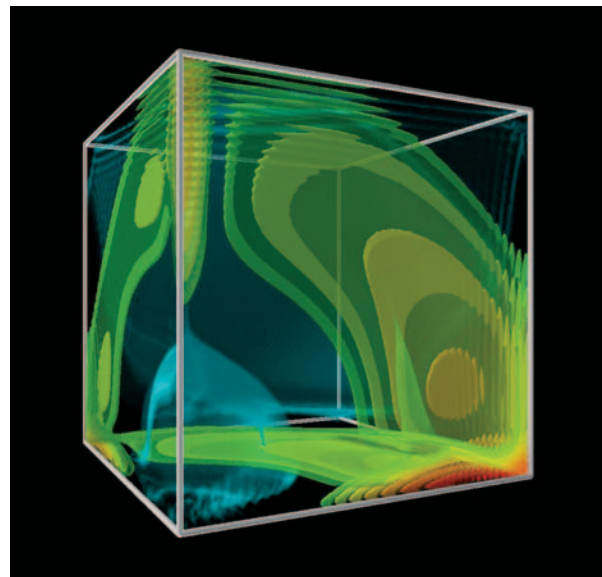
FIGURE 9 Simulation of a wavepacket scattering from a helium atom.

Whistler Wave Turbulence

S. Galtier and A. Bhattacharjee, "Anisotropic weak whistler wave turbulence in electron magnetohydrodynamics," Phys. Plasmas **10,** 3065 (2003). FES, SciDAC, NSF, INSU, ECRTN

Whistler waves have been observed in many spacecraft traversals of the Earth's magnetotail and the solar wind. Galtier and Bhattacharjee have developed a weak turbulence theory for electron magnetohydrodynamics in which the turbulence is mediated by the nonlinear interaction of whistler waves and is dominated by three-wave interactions. They demonstrated that the nonlinear interactions of whistler waves transfer energy and magnetic helicity mainly in the direction perpendicular to the external magnetic field. The anisotropic turbulence thus generated has exact stationary power-law solutions for energy and magnetic helicity.

Turbulence Spreading and Transport Scaling

Z. Lin and T. S. Hahm, "Turbulence spreading and transport scaling in global gyrokinetic particle simulations," Phys. Plasmas **11,** 1099 (2004). FES, SciDAC

An intriguing observation in tokamak experiments and in simulations of ion temperature gradient turbulence is that the fluctuations are microscopic, while the resulting turbulent transport is not gyro-Bohm. Lin and Hahm have identified a possible solution to this puzzle: turbulence spreading from the linearly active (unstable) region to the linearly inactive (stable) region. Large-scale gyrokinetic simulations found that transport driven by microscopic fluctuations is diffusive and local, whereas the fluctuation intensity is determined by nonlocal effects. Turbulence spreading reduces the fluctuation intensity in the unstable region, especially for a smaller device size, and thus introduces a non-local dependence in the fluctuation intensity. The device size dependence of the fluctuation intensity, in turn, is responsible for the observed gradual transition from Bohm to gyro-Bohm transport scaling.

A Triangular Finite Element for MHD

S. C. Jardin, "A triangular finite element with first-derivative continuity applied to fusion MHD applications," J. Comp. Phys. **200,** 133 (2004). FES, SciDAC

A 2D triangular finite element known as the reduced quintic has been used in structural engineering studies since the late 1960s, but has apparently been overlooked by the extended magnetohydrodynamics (MHD) community. Jardin has shown that the reduced quintic is well suited for many problems arising in fusion MHD applications, including a 2D elliptic problem, the solution of the anisotropic heat conduction problem, a time-dependent reduced-MHD problem (Figure 10), and the 2D axisymmetric toroidal equilibrium problem. The element requires only three unknowns per triangle, which is considerably less than other high-order elements of comparable accuracy.

Figure 10. Poloidal flux at times $t = 0$ (a) and $t = 5$ (b) and plasma current at times $t = 0$ (c) and $t = 5$ (d) for the tilt mode problem with $N = 40$. Singular currents can be seen developing in (d).

Trapped Electron Mode Turbulence

D. R. Ernst, P. T. Bonoli, P. J. Catto, W. Dorland, C. L. Fiore, R. S. Granetz, M. Greenwald, A. E. Hubbard, M. Porkolab, M. H. Redi, J. E. Rice, K. Zhurovich, and the Alcator C-Mod Group, "Role of trapped electron mode turbulence in internal transport barrier control in the Alcator C-Mod Tokamak," Phys. Plasmas **11**, 2637 (2004). FES

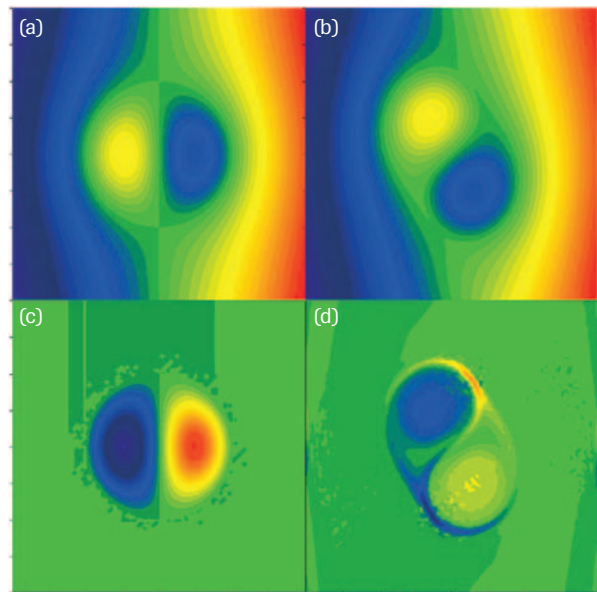To understand basic particle transport processes underlying spontaneous formation and subsequent control of the internal particle and energy transport barriers in the Alcator C-Mod Tokamak, Ernst et al. simulated trapped electron mode (TEM) turbulence using the nonlinear gyrokinetic turbulence code GS2. They found that toroidal ion temperature gradient driven modes are suppressed inside the barrier foot, but continue to dominate in the outer half-radius. As the density gradient steepens further, TEMs are driven unstable. The onset of TEM turbulence produces an outflow that strongly increases with the density gradient, upon exceeding a new nonlinear critical density gradient, which significantly exceeds the linear critical density gradient. The TEM turbulent outflow ultimately balances the inward Ware pinch, leading to steady state.

Using Finite Elements in NIMROD

C. R. Sovinec, A. H. Glasser, T. A. Gianakon, D. C. Barnes, R. A. Nebel, S. E. Kruger, D. D. Schnack, S. J. Plimpton, A. Tarditi, and M. S. Chu (the NIMROD Team), "Nonlinear magnetohydrodynamics simulation using high-order finite elements," J. Comp. Phys. **195**, 355 (2004). FES, SciDAC

The NIMROD Team has applied a conforming representation composed of 2D finite elements and finite Fourier series to 3D nonlinear non-ideal MHD using a semi-implicit time-advance. The self-adjoint semi-implicit operator and variational approach to spatial discretization are synergistic and enable simulation in the extremely stiff conditions found in high temperature plasmas without sacrificing the geometric flexibility needed for modeling laboratory experiments. Benchmark cases and a nonlinear simulation demon-

strated the effectiveness of the algorithm, which is suitable for many applications in magnetic confinement fusion.

Smoothness of Turbulent Transport

J. Candy, R. E. Waltz, and M. N. Rosenbluth, "Smoothness of turbulent transport across a minimum-q surface," Phys. Plasmas **11**, 1879 (2004). FES, SciDAC

Some controversy exists over the role of weak or reversed shear in the formation of internal transport barriers. One theory attributes the formation of internal transport barriers to a gap in global wave structures in the minimum-q region. In simulations made with the GYRO gyrokinetic code, Candy et al. showed that in general no such gap exists. All of their findings — linear and nonlinear, local and global — supported the conclusion that transport generally decreases steadily as shear is decreased from positive to negative values.

Understanding Plasma Pedestals

C. S. Chang, S. Ku, and H. Weitzner, "Numerical study of neoclassical plasma pedestal in a tokamak geometry," Phys. Plasmas **11**, 2649 (2004). FES, KBSI

The increase in plasma energy content associated with the formation of a pedestal in the plasma edge is becoming a crucial performance measure for future tokamak reactors such as ITER. Consequently, the pedestal's formation mechanism and behavior are high research priorities. Chang et al. used the parallel numerical guiding center code XGC to investigate the fundamental neoclassical properties of a steep plasma pedestal and found that its properties are closely tied to neoclassical orbital dynamics. The orbit effect across a steep pedestal generates a self-consistent radial electric field and dramatically reduces the pedestal width to a level comparable with experimental observations.

High Energy Physics

Twisted Mass Lattice QCD

S. R. Sharpe and J. M. S. Wu, "Phase diagram of twisted mass lattice QCD," Phys. Rev. D **70,** 094029 (2004). HEP

There has been considerable interest recently in the twisted mass formulation of lattice QCD, which has several advantages over "untwisted" Wilson fermions, including simplified calculations of weak matrix elements. Sharpe and Wu used the effective chiral Lagrangian to analyze the phase diagram of two-flavor twisted mass lattice QCD as a function of the normal and twisted masses, generalizing previous work for the untwisted theory. Their calculations generated graphs of the condensate and pion masses for both scenarios, and they found a simple mathematical relation between them — a result that may be important to numerical simulations.

Numbers of Flavors in QED$_3$

S. J. Hands, J. B. Kogut, L. Scorzatod, and C. G. Strouthos, "Non-compact QED$_3$ with N_f = 1 and N_f = 4," hep-lat/0404013 (2004). HEP, NSF, PPARC

If three-dimensional quantum electrodynamics (QED$_3$) is a relevant theory for the underdoped and non-superconducting region of the phase diagram of high-T_c superconducting cuprate compounds, then the number of fermion species N_f is an issue of concrete significance. Hands et al. presented numerical results for non-compact QED$_3$ for numbers of flavors N_f = 1 and N_f = 4 and found that chiral symmetry is spontaneously broken in the continuum limit for N_f = 1, implying that the critical number of flavors N_{fc} is >1. While some qualitative differences between N_f = 1 and N_f = 4 emerged, it was difficult to determine whether N_f = 4 lies above or below N_{fc} .

Materials Sciences

Total Negative Refraction in Crystals

Y. Zhang, B. Fluegel, and A. Mascarenhas, "Total negative refraction in real crystals for ballistic electrons and light," Phys. Rev. Lett. **91,** 157404 (2003). BES

Zhang et al. have found theoretically that a ballistic electron beam passing a domain twin in an ordered semiconductor results in a negative refraction with 100% transmission (i.e., zero reflection). This result, which applies in principle to the full spectrum of electromagnetic waves, has led to the first experimental demonstration of total negative refraction of light in real crystalline materials. The ability to steer light without reflection could be extremely valuable for high-power optics.

Excitonic Effects in Nanotubes

C. D. Spataru, S. Ismail-Beigi, L. X. Benedict, and S. G. Louie, "Excitonic effects and optical spectra of single-walled carbon nanotubes," Phys. Rev. Lett. **92,** 077402-1 (2004). BES, NSF

Optical response of individual single-walled carbon nanotubes (SWCNTs) can now be measured, but measurements have deviated greatly from theory. To understand why, Spataru et al. calculated electron–hole interaction effects (excitonic effects) on optical spectra of five SWCNTs. Results show bound excitons can exist in metallic SWCNTs — surprising news, since they do not exist in bulk metallic systems. The binding energy of excitons can be as large as 1 eV in semiconducting SWCNTs (at least one order of magnitude

larger than exciton binding energy in bulk semiconductors of similar band gap) and 100 meV in metal-lic SWCNTs. These large many-electron effects explain discrepancies between previous theories and experiments.

Elucidating Amontons' Law

J. Gao, W. D. Luedtke, D. Gourdon, M. Ruths, J. N. Israelachvili, and U. Landman, "Frictional forces and Amontons' Law: From the molecular to the macroscopic scale," J. Phys. Chem. B **108,** 3410 (2004). BES, AF

Amontons' law states that for any two materials, friction force is directly proportional to applied load, with a constant of proportionality independent of the contact area, surface roughness, and sliding velocity (Figure 11). No one knows why this law works on macroscopic, microscopic, and nanoscopic scales. Using molecular dynamics simulations, Gao et al. found that local energy-dissipating mechanisms are thermodynamic as well as mechanical, and that a proper statistical description can be formulated through the use of the Weibull distribution of local friction forces. They also concluded that although the "real" area of contact is a nonfundamental quantity, it may be a convenient scaling parameter for describing the really fundamental parameters — the number density of atoms, molecules, or bonds involved in an adhesive or frictional interaction.

Self–Healing of CdSe Nanocrystals

A. Puzder, A. J. Williamson, F. Gygi, and G. Galli, "Self-healing of CdSe nanocrystals: First-principles calculations," Phys. Rev. Lett. **92,** 217401 (2004). BES

CdSe semiconductor nanocrystals are a promising building block for new nanoscale materials. Puzder et al. completed an *ab initio* study that shows how the surface reconstructions of CdSe nanoparticles affect their electronic and optical properties. Atomic structures of the clusters are relaxed both in vacuum and in the presence of surfactant ligands, indicating significant geometrical rearrangements of the nanoparticle surface in both environments while the

FIGURE 11   Adhesive interaction between hexadecane molecules and gold atoms. The solid surfaces were modeled as a rough gold surface (a) or a flat gold (111) plane (b). The bottom panel shows ordered domains corresponding to the flat surface junction. The coefficient of friction is independent of the detailed nature of surface roughness.
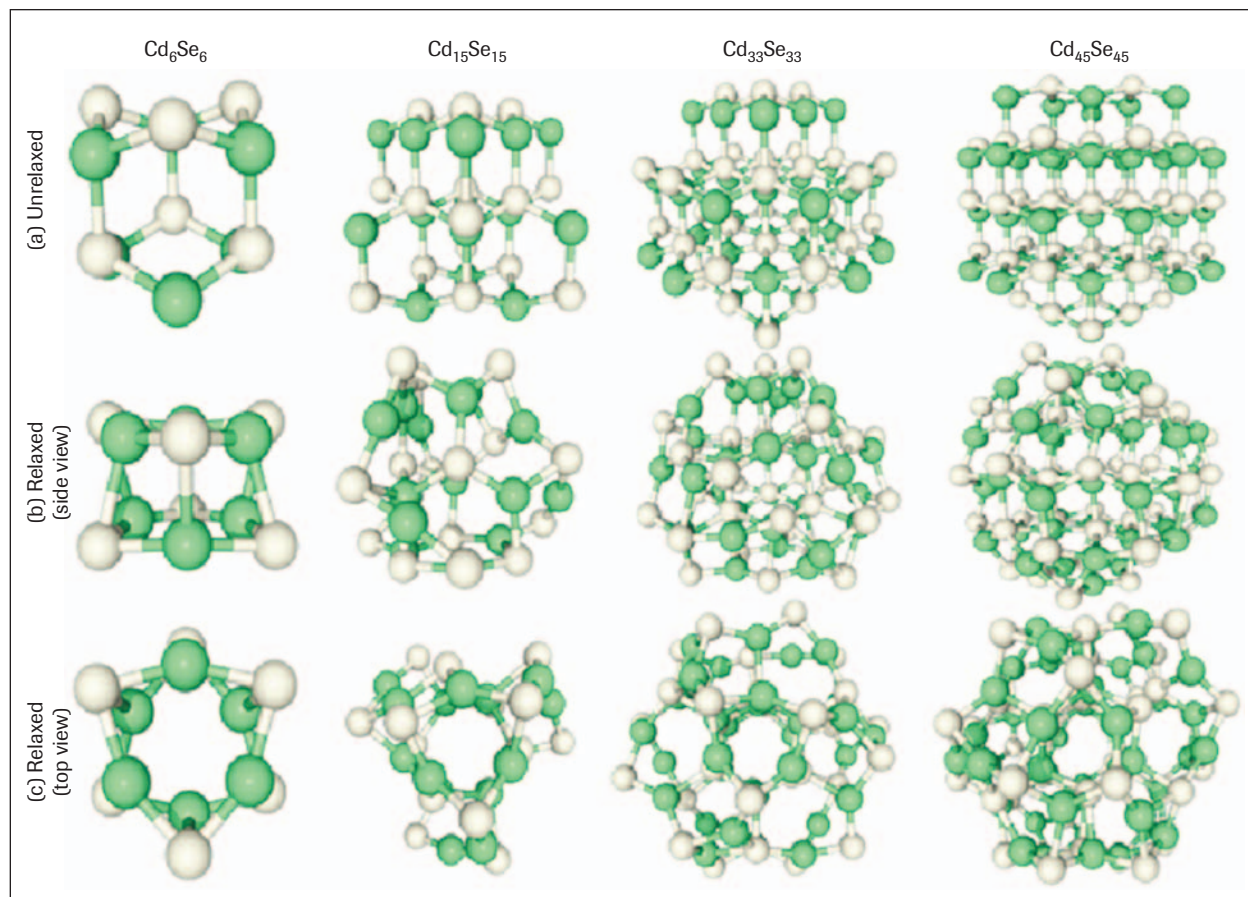
FIGURE 12   Unrelaxed (a) and relaxed (b),(c) wurtzite structures of Cd_xSe_x, x = 6, 15, 33, and 45. The Cd is green and the Se is white in the models. The side view is parallel to the c axis, while the top view is along the c axis.

wurtzite core is maintained (Figure 12). These reconstructions lead to the opening of an optical gap without the aid of passivating ligands, thus "self-healing" the surface electronic structure. Calculations also predict the existence of a midgap state responsible for recently observed subband emission.

Excitonic Entanglement in Quantum Dot Pairs

Bester et al. calculated the correlated pseudopotential wave function of an exciton in a pair of vertically stacked InGaAs/GaAs dots, and found that competing effects of strain, geometry, and band mixing led to many unexpected features missing in contemporary models. The first four excitonic states were all optically active at small interdot separation, due to the broken

symmetry of the single-particle states. The authors quantified the degree of entanglement of the exciton wavefunctions, showed its sensitivity to interdot separation, and suggested ways to spectroscopically identify and maximize the entanglement of exciton states.

Segregation of Platinum in Nanoparticles

G. Wang, M. A. Van Hove, P. N. Ross, and M. I. Baskes, "Monte Carlo simulations of segregation in Pt–Re catalyst nanoparticles," J. Chem. Phys. **121,** 5410 (2004). BES

Bimetallic mixtures of platinum (Pt) and rhenium (Re) on alumina are widely used as catalysts in petroleum refining and show promise for fuel cell applications. Wang et al. used the Monte Carlo simulation method to investigate the surface segregation and the core–shell structures of Pt–Re nanoparticles. They found that regardless of the shape of the nanoparticles, the Pt atoms segregate preferentially to the facet sites, less to edge sites, and least to vertex sites in the outermost atomic layer. Particularly interesting was the discovery of surface reconstruction on some of the facets, reflecting the preference of Pt atoms to form close-packed hexagonal structures at surfaces.
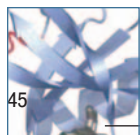
Medical and Life Sciences

45

Unfolding a Misfolded Protein

A. van der Vaart, J. Ma, and M. Karplus, "The unfolding action of GroEL on a protein substrate," Biophys. J. **87,** 562 (2004). BER, NIH, MCIF

Protein folding is complicated by high concentrations of other molecules in the cell, which can cause misfolding or aggregation of denatured proteins. Chaperones, such as the well studied bacterial chaperone GroEL, are known to protect folding proteins, but the protective mechanisms have been unclear. Van der Vaart et al. used a molecular dynamics simulation to show how GroEL can actively unfold denatured rhodanese as the apical domains of GroEL move to the open position (Figure 13). This finding confirms that multidomain interactions of the chaperone system with the protein substrate play an essential role in the protein-folding process.

Understanding Protein–Folding Intermediates

S. Brown and T. Head-Gordon, "Intermediates in the folding of proteins L and G," Protein Sci. **13,** 958 (2004). BER, ASCR-MICS, UCB, NSF

Understanding how and why proteins fold through intermediates (partially folded chains) is important for understanding larger proteins, more complicated topologies, and their possible connection to misfolding processes responsible for disease. Brown and Head-Gordon simulated the folding of Ig-binding proteins L and G, which have distinctly different folding mechanisms but identical folding topologies. Results indicate that protein G folds through an early intermediate while protein L does not. Specifically, protein G folding involves an on-pathway early intermediate; a barrier separates the unfolded state from the early folding intermediate; and this barrier is lower in free energy relative to the unfolded state.

Closed   10% Open   20% Open

30% Open   40% Open   *r'*
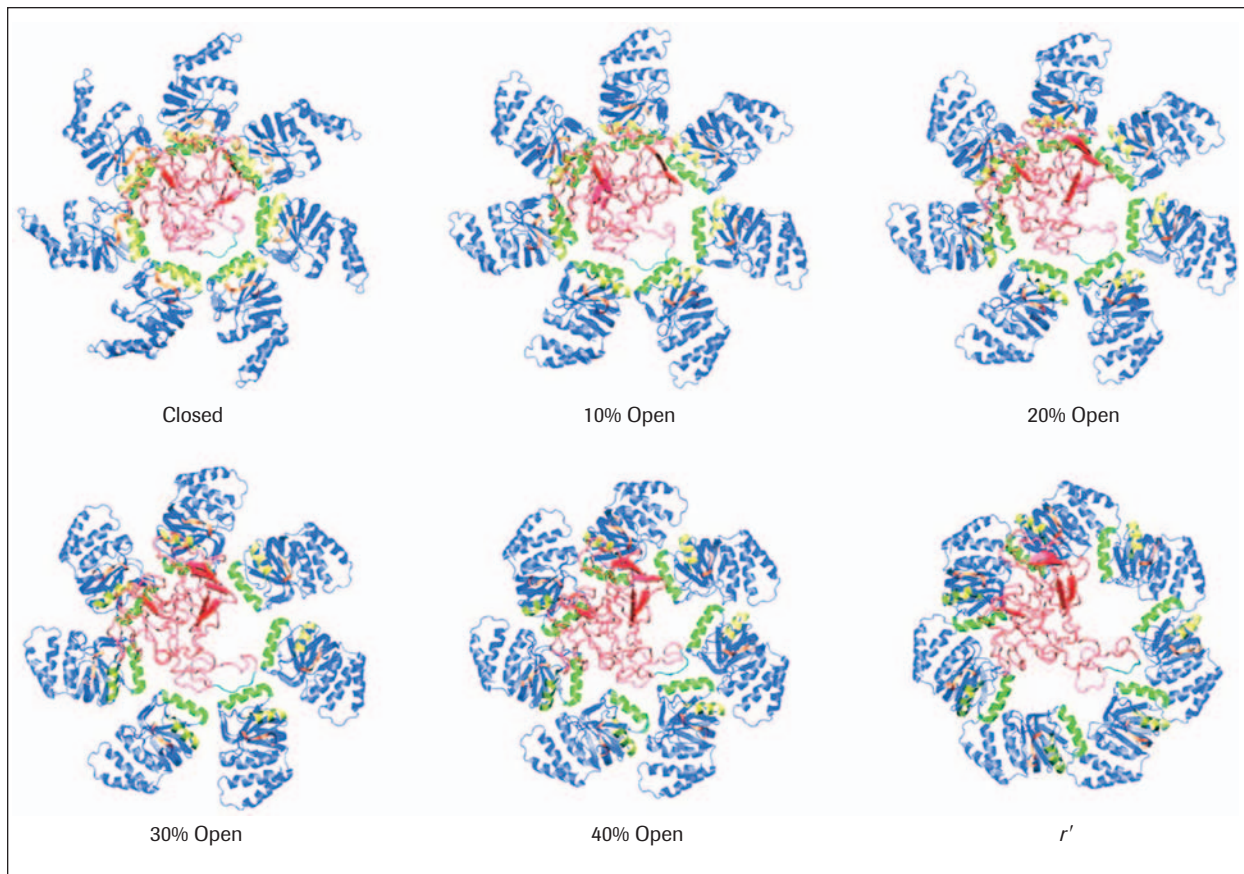
FIGURE 13   Snapshots during the unfolding simulation. The H helices of GroEL are colored yellow, the I helices are green, the loops formed by residues 310–315 are orange, and the rest of GroEL is blue. GroEL is viewed from the top, looking down into the *cis* cavity. Rhodanese is shown in red, except for the loop consisting of residues 45–50, which is in light blue.
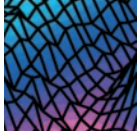
Nuclear Physics

Progress in QGP Search

The STAR Collaboration has reviewed the most important experimental results from the first three years of nucleus–nucleus collision studies at RHIC and assessed their interpretation and comparison to theory. They found that central Au+Au collisions at RHIC produce dense, rapidly thermalizing matter characterized by: (1) initial energy densities above the critical values predicted by lattice QCD for establishment of a quark-gluon plasma (QGP); (2) nearly ideal fluid flow, marked by constituent interactions of very short mean free path, established most probably at a stage preceding hadron formation; and (3) opacity to jets. Many of the observations are consistent with models incorporating QGP formation in the early collision stages (Figure 14), but the measurements themselves do not yet establish unequivocal evidence for a transition to this new form of matter.
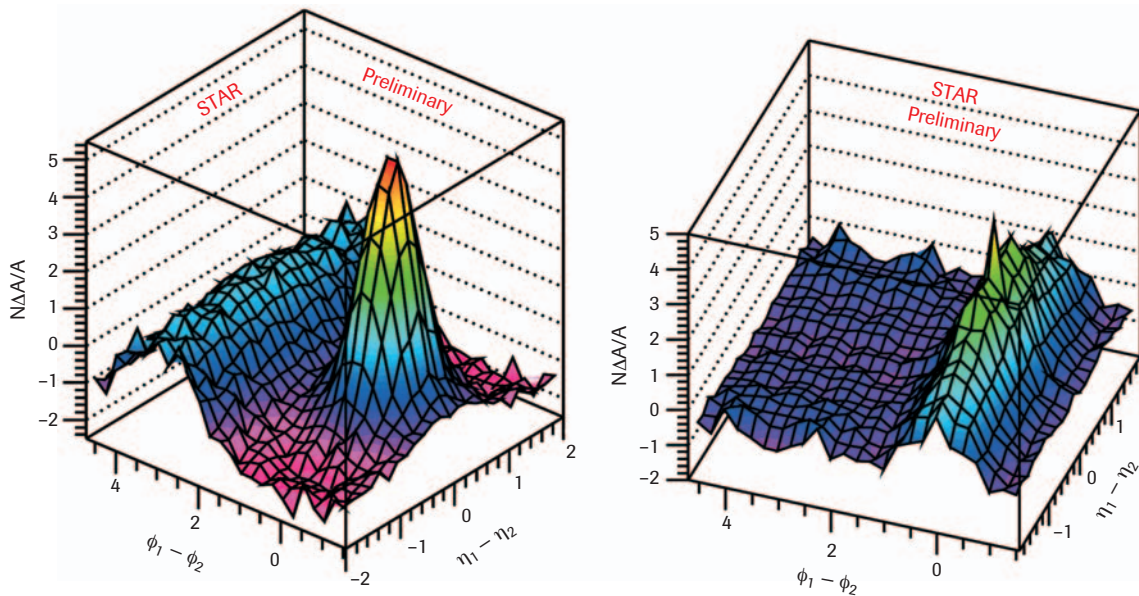
FIGURE 14   STAR angular correlation results suggest that there is appreciable soft hadron emission before the attainment of local thermal equilibrium. (a) Measurements for p+p collisions at RHIC of the joint autocorrelation on the angular difference variables pseudorapidity and azimuthal angle for all but the softest pairs of charged hadrons. This plot illustrates the central role of parton fragmentation in p+p collisions, resulting in a prominent near-side jet peak and a broad away-side jet ridge. (b) The observed soft-hadron-pair correlation for central Au+Au collisions exhibits a substantially modified remnant of the jet correlation on the near side, affecting typically 10–30% of the detected hadrons. This trend suggests that while some parton fragments are not yet fully equilibrated in the soft sector, they are nonetheless rather strongly coupled to the longitudinally expanding bulk medium.

Understanding Superfluid Fermi Gases

S. Y. Chang, V. R. Pandharipande, J. Carlson, and K. E. Schmidt, "Quantum Monte Carlo studies of superfluid Fermi gases," Phys. Rev. A **70,** 043602 (2004). NP, NSF

Chang et al. have produced quantum Monte Carlo calculations of the ground state of dilute Fermi gases with attractive short-range two-body interactions. The strength of the interaction was varied to study different pairing regimes. They successfully calculated the ground-state energy, the pairing gap Δ, and the quasiparticle spectrum for the entire region ranging from free fermions to tightly bound Bose molecules.

Studying Pentaquarks on the Lattice

N. Mathur, F. X. Lee, A. Alexandru, C. Bennhold, Y. Chen, S. J. Dong, T. Draper, I. Horváth, K. F. Liu, S. Tamhankar, and J. B. Zhang, "Study of pentaquarks on the lattice with overlap fermions," Phys. Rev. D **70,** 074508 (2004). NP

The reported discovery two years ago of an exotic five-quark resonance has spawned intense experimental and theoretical interest. Mathur et al. have produced a quenched lattice QCD calculation of spin-1/2 five-quark states with $uudd\bar{s}$ quark content for both positive and negative parities. They did not observe any bound pentaquark state in these channels for either $I = 0$ or $I = 1$. The states they found are consistent with KN scattering states, which were shown to exhibit the expected volume dependence of the spectral weight. The results were based on overlap-fermion propagators on two lattices, $12^3 \times 28$ and $16^3 \times 28$, with the same lattice spacing of 0.2 fm, and pion mass as low as ~180 MeV.

Calculating Excited States in Nuclei

Steven C. Pieper, R. B. Wiringa, and J. Carlson, "Quantum Monte Carlo calculations of excited states in A = 6–8 nuclei," Phys. Rev. C **70,** 054325 (2004). NP

Pieper et al. have demonstrated that it is possible to use the Green's function Monte Carlo (GFMC) method to compute the energies of multiple nuclear states with the same quantum numbers. The success of this method substantially increases the number of nuclear-level energies that can be compared to experimental values in the light p-shell region. For cases in which the physical excited state is reasonably narrow, the GFMC energy converges to a stable result. The results for many second and higher states in A = 6–8 nuclei are close to the experimental values.

Unraveling the Roper Resonance

Y. Chen, S. J. Dong, T. Draper, I. Horváth, F. X. Lee, K. F. Liu, N. Mathur, and J. B. Zhang, "Roper Resonance and S_{11}(1535) from Lattice QCD," hep-ph/0306199 (2004). NP

Unraveling the nature of the Roper resonance, the first excited state of the nucleon, has direct bearing on our understanding of the quark structure and chiral dynamics of baryons, which is one of the primary missions at labs like Jefferson Lab. Chen et al. performed the first lattice QCD calculation to simulate the Roper resonance. They concluded that spontaneously broken chiral symmetry dictates the dynamics of light quarks in the nucleon.

Calculating Transition Form Factors

C. Alexandrou, P. de Forcrand, T. Lippert, H. Neff, J. W. Negele, K. Schilling, W. Schroers, and A. Tsapalis, "N to Δ electromagnetic transition form factors from lattice QCD," Phys. Rev. D **69,** 114506 (2004). NP, LF, AVHF, ESOP, UC

Alexandrou et al. have demonstrated the feasibility of a new method for calculating the N to Δ electromagnetic transition form factors. For the first time in lattice QCD calculations, this new method enabled measurement of the Coulomb and electric quadrupole transition form factors up to a momentum transfer of 1.5 GeV^2 with sufficient accuracy to determine that they are both negative and statistically distinct from zero. Rough extrapolation from the high values of the quark mass at which the exploratory calculations were carried out to the physical quark mass yielded values for these quadrupole transition form factors as well as the dominant magnetic dipole form factor in qualitative agreement with experiment.

THE NERSC CENTER

In support of the DOE Office of Science's mission, the NERSC Center served 2,819 scientists throughout the United States in 2004. Appendix B provides statistical information about NERSC's clients, their primary affiliations, and their scientific disciplines. Hundreds of scientific publications in 2004 were based at least in part on calculations done at NERSC; a list is available at http://www.nersc.gov/news/reports/ERCAPpubs04.php.

Making this scientific productivity possible requires the hard work, expertise, and ingenuity of the NERSC Center and Computational Research Division (CRD)[1] group and team leaders (Figure 1) and their dedicated staff, backed up by skilled business and administrative specialists (Figure 2). Highlights of their accomplishments during the past year are described on the following pages.

JIM CRAW

Advanced Systems,
Computational Systems,
and PDSF Systems

NANCY MEYER

Mass Storage

FRANCESCA VERDIER

User Services

LYNN RIPPE

Procurement
Specialist

ZAIDA MCCUNNEY

Administrator

HOWARD WALTER

Networking, Security,
Servers and
Workstations

STEVE LOWE

Control Room,
Operations and ESnet
Support Team

WES BETHEL

Visualization Services
(CRD)

HEIDI HALLETT

Business Analyst

FIGURE 2   NERSC's business and administrative specialists provide crucial support that helps the Center run smoothly and efficiently.

ESMOND NG

Scientific Computing
Support (CRD)

KATHY YELICK

Future Technologies
(CRD)

FIGURE 1   NERSC and CRD group and team leaders work behind the scenes to make the Center a world leader in high performance computing (HPC) and computational science.

[1]Selected CRD staff work on advanced development projects for the NERSC Center and provide focused high-end support for NERSC clients.

High Performance Systems for Large-Scale Science

Meeting the Challenge of Running Larger Jobs

As home to one of the largest supercomputers available for unclassified research — the IBM SP "Seaborg" with 6,080 computing processors — the NERSC Center has moved aggressively to devote a greater share of its processing time to jobs running on 512 or more processors, in accordance with DOE and Office of Management and Budget goals.

Among these jobs was a calculation of an entire year's worth of simulated data from the Planck satellite, which ran on 6,000 processors in just two hours (see page 34). This was the first time data has ever been processed on this many of Seaborg's processors at once.

Figures 3 and 4 show NERSC's progress over the course of the year in raising the proportion of large jobs. Achieving this goal required major efforts from NERSC staff in both systems and applications.

From the perspective of the Computational Systems Group, the main issue posed by large jobs is scheduling. "Given the nature of a system like Seaborg, this is a difficult task," said group leader Jim Craw.

Just as nature abhors a vacuum, Seaborg is programmed to dislike idle nodes. Once processors are freed up, the system tries to fill them up with the next appropriate job in the queue. Left unchecked, this would result in lots of small jobs running, filling the nodes and rarely freeing up enough processors to run the larger jobs.

Craw's group created a customized scheduling algorithm to run the LoadLeveler workload management system, giving priority to larger jobs, allowing them
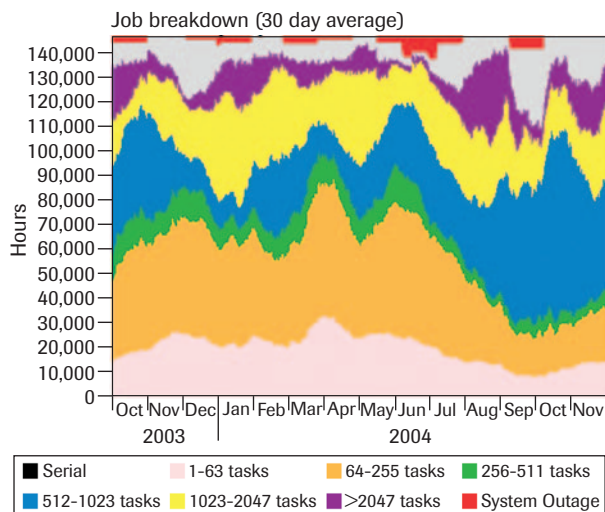
FIGURE 3   In 2004 **NERSC** continued to make improvements to Seaborg's ability to schedule larger job sizes.
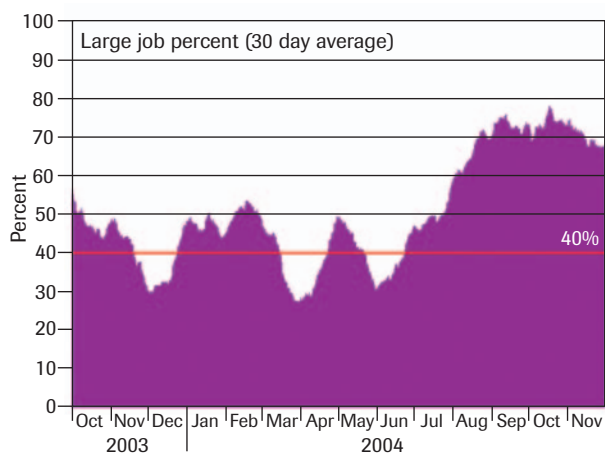
FIGURE 4   Toward the end of 2004, jobs with a processor count of 512 or more peaked around 78% of system utilization — a dramatic improvement from earlier in the year.

to work their way to the head of the queue. The system first calculates how many nodes the large job will need, then determines when the required number of nodes will be available. In the meantime, the system keeps all the nodes utilized by assigning smaller jobs that will be completed before all the needed nodes are available

While it represents a challenge to the staff, the need for such prioritizing is a testimony to the success of NERSC as an HPC center of choice. Because there is consistently more demand for computing time than can be allocated, NERSC needs to maintain a very high utilization rate, meaning that as many nodes are in use for as many hours as possible.

"It really boils down to a balancing act between small jobs and large jobs," Craw said. "We want to fill all the small holes as soon as we can, but we also need to have them available for large jobs. To achieve this, over the years we've had to do a lot of tuning of the LoadLeveler scheduling software."

The LoadLeveler customization was just one part of a major upgrade to Seaborg's system software. The migration from the AIX 5.1 operating system to version 5.2 was a significant undertaking for both NERSC and IBM. NERSC wanted the numerous improvements and fixes available only in AIX 5.2, including current and future enhancements to LoadLeveler, the GPFS file system, the Parallel Environment, and checkpoint/restart capability. However, NERSC was the first site to migrate to AIX 5.2 on a system with Power3 processors, and there were several difficulties that had to be overcome. But in true NERSC character, and with lots of help from IBM, the Systems staff solved all the problems and successfully deployed AIX 5.2.

"September 2004 was a very intensive month," Craw recalled, "with many NERSC and IBM staff working around the clock to resolve problems encountered. In the end, NERSC's staff and users provided excellent bug reports and feedback to IBM so they could improve the migration process and make it easier for other sites in the future."

The Computational Systems Group has also been working to deploy checkpoint/restart functionality on Seaborg. This functionality, which brings all jobs to a stopping point and restarts them from that point when the system is restored, will primarily be used for system administration tasks, such as shortening drain time before system or node maintenance.

Over the past two years, NERSC staff have been working with IBM to resolve a significant number of technical and design issues before checkpoint/restart can be put into production use. Additionally, NERSC staff had to make changes to the job submission mechanism used on Seaborg and developed a custom program to enable all LoadLeveler jobs to be checkpointable before they are submitted to the batch system.

"These changes are transparent to the users, so they will not notice any difference in how the batch system accepts, queues, and runs their jobs," said Jay Srinivasan, the project leader. "However, most or all jobs submitted to LoadLeveler will be checkpointable, which will allow us to use checkpoint/restart to migrate jobs from nodes and shorten queue drain times before system maintenance." Once checkpoint/restart for LoadLeveler proves to be reasonably stable, it is expected to go into production use during the first half of 2005.

Also being deployed early in 2005 is a new Fibre Channel disk array that provides an additional 12 TB of storage for users' home directories on Seaborg. Additionally, user home areas were reduced from 13 independent file systems to only three, recouping about 2 TB of unused free space from the separate areas. Together, these two changes allowed user home storage quotas to be increased by 50%. The disks freed up by the new array will be added to existing GPFS scratch and other file systems.

On the applications side, "Running larger jobs is a matter of removing bottlenecks," said David Skinner of the User Services Group. "There is always some barrier to running at a higher scale."

On Seaborg, choosing the right parallel I/O strategy for a large job can make a big difference in that job's performance. "If you have a code that runs on 16 tasks, there are a lot of ways to do I/O that will perform roughly the same," Skinner said. "But when you scale up to 4,000 tasks, there is a lot of divergence between the different I/O strategies."

There are two frequently encountered bottlenecks to scaling that come from the computational approach itself. NERSC consultants address removing these bottlenecks by rethinking the computational strategy and rewriting portions of the code.

The first bottleneck is synchronization, in which all of the calculations in a code are programmed to meet up at the same time. As the code scales to more tasks, this becomes more difficult. Skinner likens it to trying to arrange a lunch with various numbers of people. The larger the desired group, the harder it is to get everyone together at the same time and place.

"People think in a synchronous way, about closure," Skinner said. "But in a code, you often don't need synchronization. If you remove this constraint, the problems can run unimpeded as long as necessary."

The other bottleneck is in load balancing. By dividing a large scientific problem into smaller segments — and the more uniform the segments, the better — the job can often scale better, Skinner said. "Domain decomposition is important," he added.

Storage Upgrades Reduce Costs

Since relocating to Berkeley Lab almost 10 years ago, one of NERSC's goals has been to consistently introduce new technologies without creating bottlenecks to users' ongoing work. In a program of planned upgrades to tapes with greater density, the Mass Storage Group (Figure 5) has not only achieved this goal, but will save $3.7 million over the course of five years. Over that same time, from 2003 to 2007, the amount of data stored will grow from just over 1 petabyte to 11.3 petabytes.

As computational science becomes increasingly data intensive, the need for storage has ballooned. When the NERSC Center moved into the Oakland Scientific Facility in 2000, the archival data were stored on 8,000 tape cartridges. Four years later, the silos housed 35,000 cartridges. While this increase is dramatic, the real growth indicator is that the amount of data stored on each cartridge has also risen by a factor of 10.

This planned switch to higher-density cartridges is at the heart of the Mass Storage Group's investment strategy. The strategy of saving money by adopting the latest technology as soon as it is available flies in the face of conventional wisdom, which holds that by waiting for the technology to become pervasive, the price will drop. Because the storage group deploys newer technology incrementally rather than all at once, the center benefits from the technological improvements but does not get locked into one format.

"This allows us to grow to the next dense technology or build the current dense technology gradually by 10 to 15 drives a year, rather than replacing hundreds of drives all at once," said Mass Storage Group Lead Nancy Meyer.

"Our rule of thumb is to maintain the archive with 10 percent of the data going to the lower density cartridges and repack 90 percent onto the higher density cartridges. This frees up the lower density

NANCY MEYER

MATTHEW ANDREWS

SHREYAS CHOLIA

DAMIAN HAZEN

WAYNE HURLBERT

NANCY JOHNSTON

FIGURE 5   By taking advantage of new technologies, the Mass Storage Group saves money while meeting the ballooning need for data storage.

cartridges and we can reuse them without purchasing more media. And when data is growing by petabytes per year, this adds up to significant savings."

NERSC's investment strategy is already paying off, saving an estimated $135,000 in 2004. By 2007, the annual saving is estimated at $2.3 million. In all, the program is expected to save $3,739,319 over five years.

At the start of the plan, data were stored on a mix of cartridges, each capable of holding either 20 or 60 gigabytes. To keep up with data growth, the group could either stick with the current cartridges — and buy a lot of them — or adopt the newly available 200 GB cartridges. While this switch would also require the center to buy new tape drives to read the denser tapes, the long-term costs showed that higher density meant lower overall costs.

For example, in 2004, data grew by 832,000 terabytes. Buying enough 60 GB cartridges would have cost $497,952, while buying enough higher-density tapes and tape drives cost $135,000 less than the lower-density option. In 2005, when the group

expects to upgrade to 500 GB cartridges, the estimated savings will be $308,000, even after accounting for new tape drives.

"It's important to note that these savings only reflect the cost of the tapes," said Nancy Meyer, leader of the Mass Storage Group. "It doesn't include the money saved by not adding more silos or moving other systems to create more adjacent floor space."

Other computing centers, according to Meyer, take the more traditional approach. Not only is this more expensive in the long run, but it also results in wholesale rather than incremental upgrades, and these can be more disruptive to users.

NERSC users, Meyer said, don't see much difference when storage equipment is upgraded. Smaller files, amounting to about 10 percent of the total data, are still stored on 20 GB cartridges. Finding a file on one of these tapes takes about 10 seconds. Locating a file on the denser tape takes about a minute, possibly 90 seconds. "While it takes a little longer to locate files, we think it's a fair tradeoff for almost $4 million in savings," Meyer said.

PDSF Boosts Processors, Performance

Researchers using the PDSF cluster managed by NERSC benefited from a system upgrade in 2004, giving them access to more processing power and a higher-speed network connection for accessing archived data. The new hardware, shipped in September, included 48 nodes of dual Xeon processors and 10 nodes of dual Opteron processors. These were added to the existing 550-processor cluster.

The system's connection to NERSC's internal network was upgraded from a 2 Gigabit switch to a 10 Gigabit switch. This will improve the transfer rate of data between PDSF and NERSC's HPSS archival storage.

Finally, some of the PDSF disk drives were also upgraded, adding 12 terabytes of disk capacity for a total of 135 TB.

DaVinci Expands Visualization and Analysis Capabilities

In response to requests from the NERSC community, the Center has expanded its interactive analysis capability by acquiring a new platform called DaVinci. DaVinci will be NERSC's primary resource for general purpose, interactive data-intensive visualization and analysis work. It provides a large amount of high performance scratch storage, an expandable architecture that supports shared- or distributed-memory parallelization, a 64-bit operating system, and a large amount of memory.

At the end of 2004, DaVinci, an SGI Altix system, consisted of eight 1.4 GHz Itanium2 processors, 48 GB of RAM, 3 TB of attached storage, and a combi-nation of bonded Gigabit Ethernet and 10 Gigabit network interface connectors. Significant expansion of the system is expected during 2005, subject to the availability of funding.

DaVinci is well suited for interactive visualization and analysis tasks, which require a different system balance than NERSC's primary computational resources — a balance biased towards larger memory sizes and higher I/O throughput per processor. The Altix's single-system-image architecture makes all system resources, including I/O bandwidth and memory, available to any process in the system. When the analysis task is serial in implementation — as is the case with many commercial, legacy, and "quick and dirty" codes — more memory for the task translates directly into increased processing capability. The combination of a large, fast scratch file system and high internal bandwidth makes DaVinci perfect for data-intensive applications.

DaVinci is expected to go into full production in May 2005, when it will be used primarily for interactive, data-intensive visualization and analysis applications. During off-peak hours, it will be available for use by batch applications to help fulfill allocation requests. The interactive vs. batch use policy will be outlined when full production is announced.

DaVinci will be outfitted with the same visualization applications present on the older Escher visualization server, which will be decommissioned shortly after DaVinci goes into full production mode. The visualization tools include an array of libraries and applications for plotting and charting, imaging, command-line and GUI-based interactive development environments, and point-and-click applications.

Comprehensive Scientific Support

Support for Large-Scale Projects

Many of the important breakthroughs in computational science come from large, multidisciplinary, multi-institutional collaborations working with advanced codes and large datasets, such as the SciDAC and INCITE collaborations. These teams are in the best position to take advantage of terascale computers and petascale storage, and NERSC provides its highest level of support to these researchers. This support includes special service coordination for queues, throughput, increased limits, etc.; and specialized consulting support; which may include algorithmic code restructuring to increase performance, I/O optimization, visualization support — whatever it takes to make the computation scientifically productive.

The principal investigators of the three 2004 INCITE projects were unanimous in their praise of the support provided by NERSC staff. William A. Lester, Jr., PI of the "Quantum Monte Carlo Study of Photoprotection via Carotenoids in Photosynthetic Centers" project, said:

> "The INCITE Award led us to focus on algorithm improvements that significantly facilitated the present calculations. One of these improved quantum Monte Carlo scaling … others included improved trial function construction and a more efficient random walk procedure. We have benefited enormously from the support of NERSC staff and management … and the Visualization Group has provided us with modes of presenting our work beyond our wildest imagination."

Tomasz Plewa, leader of the project "Thermonuclear Supernovae: Stellar Explosions in Three Dimensions," commented:

> "We have found NERSC staff extremely helpful in setting up the computational environment, conducting calculations, and also improving our software. One example is help we have received when implementing an automatic procedure for code checkpointing based on the access to the queue remaining time. We have also been able to resolve problems with large I/O by switching to a 64-bit environment. Making this environment available on short notice was essential for the project."

And P. K. Yeung, PI of the "Fluid Turbulence and Mixing at High Reynolds Number" INCITE project, offered these comments:

> "The consultant services are wonderful. Our designated consultant has worked hard to help us, including responding to emails every day of the week. We have benefited from consultants' comments on code performance, innovative ideas for improvement, and diagnostic assistance when user or system problems occur. The visualization staff have also been very good…."

> "We really appreciate the priority privilege that has been granted to us in job scheduling. This has allowed most of our jobs to start relatively quickly compared to what we experience at other sites…. Our INCITE Award has had tremendous impact on our work, and by opening up many exciting opportunities for collaborations and long-term partnerships, can be expected to bring many long-lasting advances in both turbulence and computational science in general."

Integrated Performance Monitoring Simplifies Code Assessment

As the HPC center of choice for the DOE research community, NERSC consistently receives requests for more computing resources than are available. Because computing time is so valuable, making the most of every allocated processor-hour is a paramount concern. Evaluating the performance of application codes in the diverse NERSC workload is an important and challenging endeavor. As NERSC moves toward running more large-scale jobs, finding ways to improve performance of large-scale codes takes on even greater importance.

For this reason identifying bottlenecks to scalable performance of parallel codes has been an area of intense focus for NERSC staff. To identify and remove these scaling bottlenecks, David Skinner of NERSC's User Services Group has developed Integrated Performance Monitoring, or IPM. IPM is a portable profiling infrastructure that provides a performance summary of the computation and communication in a parallel program. IPM has extremely low overhead, is scalable to thousands of processors, and was designed with a focus on ease of use, requiring no source code modification. These characteristics are the right recipe for measuring application performance in a production environment like

NERSC's, which consists of hundreds of projects and parallelism ranging from 1 to 6,000 processors.

Skinner points to the lightweight overhead and fixed memory footprint of IPM as one of its biggest innovations. Unlike performance monitoring based on traces, which consume more resources the longer the code runs, IPM enforces strict boundaries on the resources devoted to profiling. By using a fixed memory hash table, IPM achieves a compromise between providing a detailed profile and avoiding impact on the profiled code.

IPM was also designed to be portable. It runs on the IBM SP, Linux clusters, Altix, Cray X1, NEC SX6, and the Earth Simulator. Portability is key to enabling cross-platform performance studies. Portability, combined with IPM's availability under an open source software license, will hopefully lead to other centers adopting and adding to the IPM software.

Skinner characterizes IPM as a "profiling layer" rather than a performance tool. "The idea is that IPM can provide a high-level performance summary which feeds both user and center efforts to improve performance," Skinner said. "IPM finds 'hot spots' and bottlenecks in parallel codes. It also identifies the overall characteristics of codes and determines which compute resources are being used by a code.
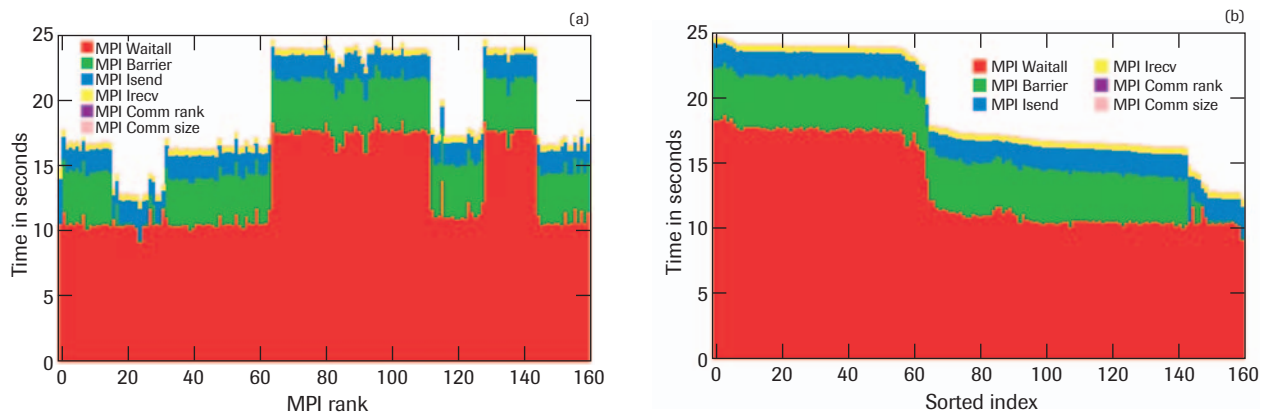

FIGURE 6   IPM can graphically present a wide range of data, including communication balance by task, sorted by (a) MPI rank or (b) MPI time.

It really provides a performance inventory. Armed with that information, users can improve their codes and NERSC can better provide compute resources aligned to meet users' computational needs."

IPM automates a number of monitoring tasks that Skinner and other HPC consultants used to perform manually. By running a code with IPM, NERSC staff can quickly generate a comprehensive performance picture of a code, with the information presented both graphically (Figure 6) and numerically.

The monitors that IPM currently integrates include a wide range of MPI communication statistics; HPM (Hardware Performance Monitor) counters for things like flop rates, application memory usage, and process topology; and system statistics such as switch traffic.

The integration in IPM is multi-faceted, including binding the above information sources together through a common interface, and also integrating the records from all the parallel tasks into a single report. On some platforms IPM can be integrated into the execution environment of a parallel computer. In this way, an IPM profile is available either automatically or with minor effort. The final level of integration is the collection of individual performance profiles into a database that synthesizes the performance reports via a Web interface. This Web interface can be used by all those concerned with parallel code performance: users, HPC consultants, and HPC center managers. As different codes are characterized, the results are posted to protected Web pages. Users can access only the pages for the codes they are running.

One of the first uses for IPM was to help the initial three INCITE projects make the most effective use of their large allocations. Subsequently it has been expanded to other projects. Even a small improve-

ment — say 5 percent — in a code that runs on a thousand processors for millions of processor-hours is a significant gain for the center. "Our primary goal is to help projects get the most out of their allocated time," Skinner said.

But the same information is also interesting to the center itself. Obtaining a center-wide picture of how computational resources are used is important to knowing that the right resources are being presented and in the right way. It also guides choices about what future NERSC computational resources should look like. For example, IPM shows which parts of MPI are widely used by NERSC customers and to what extent. "It's good to know which parts of MPI our customers are using," Skinner said. "As an HPC center this tells us volumes about not only what we can do to make codes work better with existing resources as well as what future CPUs and interconnects should look like."

"We are looking for other programmers to contribute to IPM," Skinner added. "IPM complements existing platform-specific performance tools by providing an easy-to-use profiling layer that can motivate and guide the use of more detailed, in-depth performance analysis."

More information about IPM is available at http://www.nersc.gov/projects/ipm/.

New Capabilities for Remote Visualization of Large Data Sets

In 2002 NERSC hosted a workshop to solicit input from the user community on the subject of future visualization and analysis requirements. The workshop findings report[2] cited several areas of user concern and need. One is the need for more capable visualization and analysis tools to address the chal

---

[2] B. Hamann, E. W. Bethel, H. D. Simon, and J. Meza, *NERSC "Visualization Greenbook": Future Visualization Needs of the DOE Computational Science Community Hosted at NERSC,* Lawrence Berkeley National Laboratory report LBNL-51699 (2002), http://vis.lbl.gov/Publications/2002/VisGreenFindings-LBNL-51699.pdf.

lenges posed by ever-larger and complex data being collected from experiments and generated by simulations. Another is the need for centralized deployment and management of general-purpose visualization resources. Yet another is the need for remote visualization software that can be used by multiple members of distributed teams.

During 2004, the NERSC Center took steps towards addressing these concerns. First, the Center consolidated management of licensed visualization software to a set of redundant license servers, thus streamlining and simplifying license management. Consolidating licenses from different machines into a larger license pool has increased the number of licenses available to users for certain applications, such as IDL. The result is increased availability of licensed visualization software at a reduced cost of managing licenses.

An additional benefit of the central license servers is that remote users can now check out a license for a commercial visualization application and run it on their remote desktop. In some cases, the commercial applications are quite expensive, so the user community benefits by being able to take advantage of resources paid for and maintained by NERSC. More information is available at the "Remote License Services at NERSC" Web page.[3]

To address the need for more capable visualization tools that can be used in a remote and distributed setting, NERSC has purchased and deployed EnSight Gold, an interactive visualization application from CEI. EnSight runs in a pipelined fashion, with visualization processing taking place at NERSC, and rendering taking place on the user's local desktop, which makes it particularly useful when data files are too large to download for local analysis. The "back end" that performs the visualization processing at NERSC can be run in parallel, scaling to

accommodate large data files. EnSight also supports collaborative interaction — multiple persons at different locations may run the EnSight display client simultaneously, all viewing and interacting with the same data set running at NERSC (Figure 7).

The Berkeley Lab Visualization Group is evaluating several additional new visualization applications that may be useful for remotely visualizing large scientific data sets in the NERSC environment.
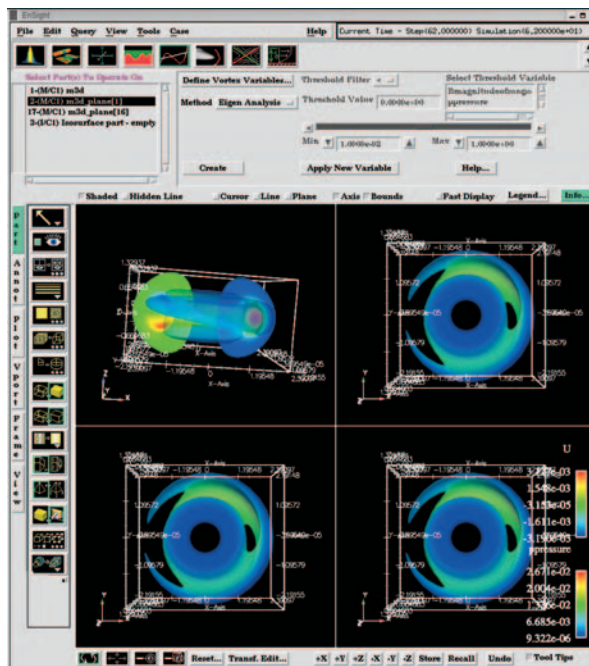

FIGURE 7   Screen shot of the EnSight Gold display client rendering multiple views of GTK fusion simulation results. (Image: Berkeley Lab Visualization Group and Scott Klasky, Princeton Plasma Physics Laboratory)

---

[3] http://www.nersc.gov/nusers/services/licenses/overview.php

User Surveys Prompt Improvements

For the sixth consecutive year, NERSC conducted a survey and invited all users to provide input on how well the Center is meeting their HPC requirements and expectations. Not only do the survey responses indicate how well the Center is performing and where it can improve, but the feedback is also used to implement changes to improve systems and services. In 2004, 209 users responded to the survey.

Areas with the highest user satisfaction include the HPSS mass storage system, both its reliability and its availability, account support services, and HPC consulting. In addition to overall consulting support, particular areas rated highly were the timely initial response to consulting questions and follow-up to initial consulting questions. The largest increases in satisfaction over last year's survey came from training classes attended in person, visualization services, the HPSS and Seaborg Web pages, and software bug resolution.

Areas with the lowest user satisfaction include the Seaborg's batch turnaround time and queue structure, as well as services used by only small numbers of users. These services include the math and visualization servers, Grid services, and training classes presented over the Access Grid. The areas rated significantly lower this year include Seaborg and available computing hardware.

In the section "What does NERSC do well?" 118 respondents gave top marks to access to powerful computing resources, excellent support services and staff, reliable hardware and a well-managed center, an easy-to-use environment for users, general satisfaction, HPSS, and the documentation on the NERSC Web site.

When it came to "What should NERSC do differently?" 94 responses focused mainly on issues involving Seaborg. The top request was to improve Seaborg's turnaround time, followed by changing the job scheduling policies for Seaborg, especially midrange jobs. There were also requests to provide more computing resources and to improve the allocations process.

Such comments often lead to real change. As a result of the 2003 survey, NERSC made the following improvements:

- In striving to balance sometimes competing users' expectations and DOE priorities, such as the Office of Science's emphasis on capability computing (including large jobs and INCITE jobs), in 2004 NERSC implemented several measures which should help improve turnaround for all jobs: Per-user run limits were decreased from six to three, and per-user idle limits (the number of jobs that are eligible for scheduling) from 10 to two. This provides fairer access to Seaborg's processors.

- NERSC is working to add additional computational resources for jobs that do not scale to 512 or more processors.

- Working closely with IBM's compiler support group, NERSC is now able to more quickly resolve reports of bugs and compiler problems. Compiler upgrades have resulted in better run-time performance.

- TotalView upgrades this past year have resulted in more stability and usability, and TotalView can now debug a wider range of parallel codes.

- In 2004 NERSC organized 20 user training lectures in seven separate events. All were presented via the Access Grid and were captured as streaming videos so that users can replay them at any time. These lectures have been added to the tutorials page for one-stop access to training materials.

- The PDSF support team has made it possible to run interactively on the batch nodes. PDSF also

added new nodes with much faster processors and twice as much memory, as well as 20 terabytes of additional disk space.

- NERSC finalized consolidation of license management for all visualization software hosted at the Center. The new system also supports remote use of licensed visualization software.

- Finally, NERSC reorganized its Web site last year, merging several different sites into one site that meets the needs of our users, our DOE sponsors, and the general public.

Complete results of the 2004 user survey can be found at http://www.nersc.gov/news/survey/2004/.

Secure Connectivity

All NERSC Production Systems Now Grid-Accessible

Early in 2004 came good news for NERSC users — all of NERSC's production computing and storage resources are now Grid-enabled and can be accessed by users who have Grid applications using the Globus Toolkit. Grid users now have access to Seaborg, HPSS storage, the PDSF cluster, and the visualization and math servers. Users can also get support for Grid-related issues from NERSC's User Services Group.

"Now that the core functionalities have been addressed, our next push is to make the Grid easier to use and manage, for both end users and the administrators who handle applications that span multiple sites," said Steve Chan, who coordinated the Grid project at NERSC.

One of the major challenges faced by NERSC's Grid team was installing the necessary software on Seaborg, which operates using a complicated software stack. Additionally, the system needed to be configured and tested without interfering with its heavy scientific computing workload.

"With Steve Chan's leadership and the technical expertise of the NERSC staff, we crafted a Grid infrastructure that scaled to all of our systems," said Bill Kramer,

general manager of the NERSC Center. "As the driving force, he figured out what needed to be done, then pulled together the resources to get it done."

To help prepare users, Center staff presented tutorials at both Globus World and the Global Grid Forum, with training specifically designed for NERSC users.

When authenticated users log into NIM (NERSC Information Management), they are now able to enter their certificate information into the account database and have this information propagated out to all the Grid-enabled systems they normally access, superseding the need to have separate passwords for each system.

Because the Grid opens up a wider range of access paths to NERSC systems, the security system was enhanced. Bro,[4] the LBNL-developed intrusion detection system which provides one level of security for NERSC, was adapted to monitor Grid traffic at NERSC for an additional layer of security.

Full Grid access is the culmination of a two-year effort by the staff members pictured in Figure 8.

Keeping Out Cyber Intruders

NERSC, which has shared its cyber security expertise with the HPC community for the past three years,

[4] http://www.bro-ids.org/

MIKHAIL AVREKH

CLAYTON BAGWELL

MARK HEER

NICK BALTHASER

SCOTT CAMPBELL

SHANE CANON

STEVE CHAN

STEVE LAU

KEN OKIKAWA

R. K. OWEN

DAVID PAUL

IWONA SAKREJDA
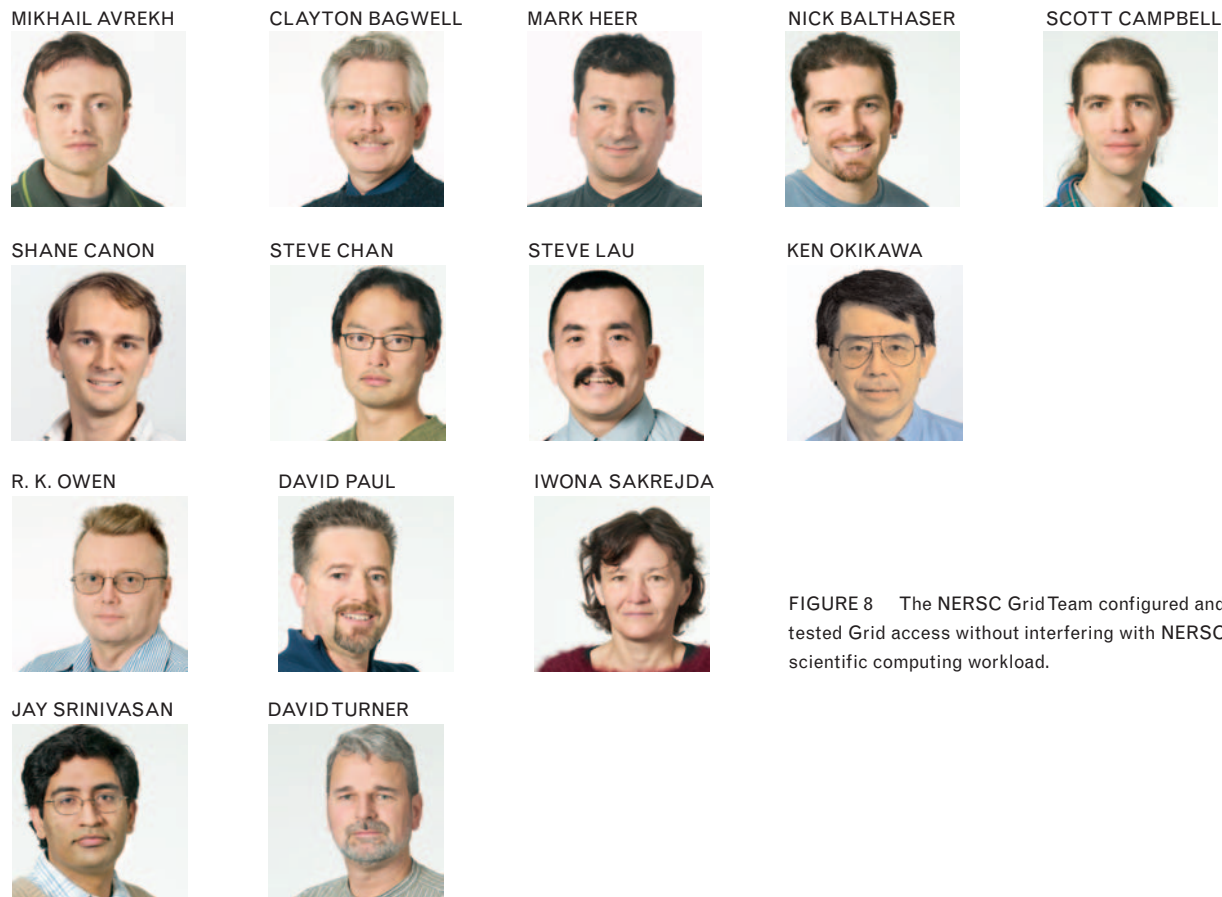
FIGURE 8   The NERSC Grid Team configured and tested Grid access without interfering with NERSC's scientific computing workload.

JAY SRINIVASAN

DAVID TURNER

unintentionally emerged as a true leader in the field in March and April 2004, when a notorious hacker broke into many of the nation's leading supercomputing centers and forced them to take their systems offline. Thanks to its Bro intrusion detection system, NERSC withstood the hacker's attacks and remained in service to users.

The attacks prompted the organization of "Cybersecurity Summit 2004," an invitation-only workshop held in Arlington, Virginia on September 27–28. This meeting brought together cyber security experts from some of the nation's top research institutions to better prepare for future cyber attacks.

Sharing their experience and expertise at the workshop were Howard Walter and Scott Campbell of NERSC's

Networking, Security, Servers and Workstations Group, and Berkeley Lab's Computer Protection Program Manager Dwayne Ramsey, who was an invited speaker.

"The biggest benefit to holding a meeting like this is building the trust relationships that are critical when it comes to developing a coordinated strategy against cyber attacks," said Walter, a member of the summit program committee. "When an attack is detected by multiples sites at the same time, knowing who to talk to — and who to trust — can make all the difference."

While the attacks in spring raised the visibility of cyber security, the issue is a constant threat. The NERSC Center is scanned for vulnerabilities about 5,000 times a month, while another 70 worms and viruses attempt to penetrate the systems each month.

Providing an invisible shield around NERSC is the Bro intrusion detection system developed at Berkeley Lab. Bro constantly monitors network links, searching for traffic that potentially violates a site's access and usage policies. Bro automatically blocks incoming malicious traffic and notifies cyber security staff.

Steve Lau of NERSC's Networking, Security, Servers and Workstations Group has also developed a visualization tool for Bro to graphically show how dangerous the Internet has become. First deployed at the SC2003 conference and then again at SC2004, the "Spinning Cube of Potential Doom" (Figure 9) attempts to display the overall level of malicious traffic in a fashion that can be easily understood by those without a computer security or networking background.

The Cube leverages Bro's capability to log all instances of completed or attempted TCP connections, displaying this information within a three-dimensional cube. Each axis represents a different component of a TCP connection: the X axis represents the local IP address space, the Z axis represents the global IP address space, and the Y axis represents the port number. Port numbers are used in connections to locate services and coordinate communication (e.g., 22 for ssh and 80 for http).

Each TCP connection, whether attempted or successful, is displayed as a single point in the Cube. Successful connections are shown as white dots, while incomplete connections are colored using a rainbow color map, with the color varying by port number. This color mapping was used to assist the viewer in locating the point in 3D space. The vast majority of colored dots can be considered malicious traffic searching for potentially vulnerable systems.

One of the more interesting findings, even surprising to those with backgrounds in computer security, is that visual patterns emerge from the data. Various distinct patterns are easily discernable. Port scans
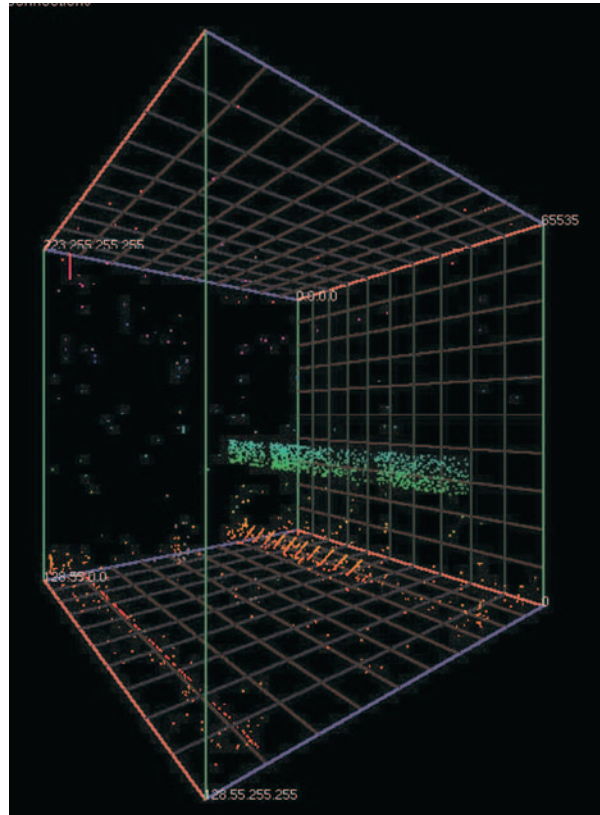
FIGURE 9   The Spinning Cube of Potential Doom reveals surprising patterns in malicious Internet traffic.

appear as lines, with vertical lines representing scans directed at a particular host searching for any listening port, and horizontal lines being scans directed at the entire local address space on a particular port.

Another unexpected pattern emerging from the data was dubbed "barber pole" because it looks like the striping on barber poles. These scans vary their port numbers and IP addresses in an attempt to elude detectors. Although they may be capable of evading detectors, they stand out when visualized by the Cube of Doom.

Another type of scan detected was dubbed a "lawn-mower" scan. These scans are quite noisy, in that

they scan a wide range of contiguous ports while simultaneously marching across the entire local address space. Some of these scans are quite rapid, occurring within a few seconds, while others are more leisurely, ranging on the order of minutes.

The Cube's visual display received much interest from SC conference attendees, many of whom were surprised at the amount of potentially malicious traffic on the Internet. Many attendees were curious as to what portion of the data represented attempts against their systems and expressed surprise that they had not noticed it themselves. Many of them said that the data had inspired them to make sure their systems are up to date.

Making Access More Secure and Efficient

One of the most popular displays at the past few SC conferences has been a monitor displaying unprotected passwords as they were sent over the network. The data, gathered by Berkeley Lab's Bro intrusion detection system, was of great interest because is showed both how little effort some people put into choosing their passwords, and how easily those passwords could be intercepted.

To minimize the risk of this happening to both NERSC users and staff, members of the Networking, Security, Servers and Workstations Group have been assessing the use of one-time passwords. Not only was the group interested in the effectiveness of the offerings from the three leading vendors in the field, but staff also wanted to develop an approach that was relatively easy to use.

A major obstacle with implementing one-time passwords is that they are literally just that, said Steve Chan. "The difficulty is that every time you log into a system, you need to use a new password. In the course of a day, you could find yourself having to generate and use one password for your desktop machine, another one to log into Seaborg, yet another one

if you use PDSF, then another one if you want to open a new window on Seaborg. And if you were running batch jobs and wanted to copy files from HPSS or using Grid FTP, you'd need even more passwords," Chan explained.

Such a system would make it difficult to carry out automated processes, such as backups, which are often done late at night.

So, the group looked into how one-time passwords could be integrated with a single sign-on, allowing the user to use one secure password across all NERSC systems. After investigating numerous options, the group developed a plan to integrate the password system with Kerberos, which would generate a single "ticket" and be recognized across most platforms. For password generation, the group is planning to build an initial system using CryptoCard, a one-time password vendor that provides and an open-solution toolkit.

The group, working with ESnet and other labs, also tested a federated system to allow single sign-on across other DOE laboratories. In the spring and summer of 2004, a test program was conducted across ESnet to evaluate the technical feasibility of using a remote laboratory's one-time password system for accessing local systems at Lawrence Berkeley, Argonne, Oak Ridge, and Pacific Northwest national laboratories. The successful tests led ESnet to propose building out an "authentication fabric" linking the individual authentication systems at each lab together, while NERSC sought funding for a local one-time password system that ties the one-time passwords into Kerberos. While still awaiting final word on the funding proposal, Chan said NERSC will begin testing a system in 2005.

"In addition to testing how the system will be used, we also want to make sure that communication across the channels is secure and that nothing sensitive goes across the network without being protected," Chan said.

Ultimately, just as Berkeley Lab and NERSC have done with Bro, the goal is to take a lead role in developing and deploying an effective and affordable system that can be used, added to and extended by other DOE sites — and that does not hinder scientific productivity.

Advanced Development

Architecture Benchmarking and Evaluation

With the re-emergence of viable vector computing systems such as the Earth Simulator and the Cray X1, and with IBM and DOE's BlueGene/L taking the number one spot on the TOP500 list of the world's fastest computers, there is renewed debate about which architecture is best suited for running large-scale scientific applications.

In order to cut through conflicting claims, a team of researchers from Berkeley Lab's Computational Research (CRD) and NERSC Center divisions have been putting various architectures through their paces, running benchmarks as well as scientific applications key to DOE research programs. The team (Figure 10) includes Lenny Oliker, Julian Borrill, Andrew Canning, and John Shalf of CRD; Jonathan Carter and David Skinner of NERSC; and Stephane Ethier of Princeton Plasma Physics Laboratory.

In their first study of 2004,[5] the team put five different systems through their paces, running four different scientific applications key to DOE research programs. As part of the effort, the group became the first international team to conduct a performance evaluation study of the 5,120-processor Earth Simulator. The team also assessed the performance of:

- the 6,080-processor IBM Power3 supercomputer running AIX 5.1 at the NERSC Center

- the 864-processor IBM Power4 supercomputer running AIX 5.2 at Oak Ridge National Laboratory (ORNL)

JULIAN BORRILL   ANDREW CANNING   JONATHAN CARTER   STEPHANE ETHIER (PPPL)

LENNY OLIKER   JOHN SHALF   DAVID SKINNER

FIGURE 10   NERSC's architecture benchmarking and evaluation team analyzes the complex interplay of scientific codes and system designs.

[5] L. Oliker, A. Canning, J. Carter, J. Shalf, and S. Ethier, "Scientific computations on modern parallel vector systems," presented at SC2004, Pittsburgh, Pa., Nov. 6-12, 2004; http://crd.lbl.gov/~oliker/papers/SC04.pdf.

- the 256-processor SGI Altix 3000 system running 64-bit Linux at ORNL

- the 512-processor Cray X1 supercomputer running UNICOS at ORNL.

"This effort relates to the fact that the gap between peak and actual performance for scientific codes keeps growing," said team leader Lenny Oliker. "Because of the increasing cost and complexity of HPC systems, it is critical to determine which classes of applications are best suited for a given architecture."

The four applications and research areas selected by the team for the evaluation were:

- Cactus, an astrophysics code that evolves Einstein's equations from the Theory of Relativity using the Arnowitt-Deser-Misner method

- GTC, a magnetic fusion application that uses the particle-in-cell approach to solve non-linear gyrophase-averaged Vlasov-Poisson equations

- LBMHD, a plasma physics application that uses the Lattice-Boltzmann method to study magneto-hydrodynamics

- PARATEC, a first-principles materials science code that solves the Kohn-Sham equations of density-functional theory to obtain electronic wave functions.

"The four applications successfully ran on the Earth Simulator with high parallel efficiency," Oliker said. "And they ran faster than on any other measured architecture — generally by a large margin." However, Oliker added, only codes that scale well and are suited to the vector architecture may be run on the Earth Simulator. "Vector architectures are extremely powerful for the set of applications that map well to those architectures," Oliker said. "But if even a small part

of the code is not vectorized, overall performance degrades rapidly."

As with most scientific inquiries, the ultimate solution to the problem is neither simple nor straightforward.

"We're at a point where no single architecture is well suited to the full spectrum of scientific applications," Oliker said. "One size does not fit all, so we need a range of systems. It's conceivable that future supercomputers would have heterogeneous architectures within a single system, with different sections of a code running on different components."

One of the codes the group intended to run in this study — MADCAP, the Microwave Anisotropy Dataset Computational Analysis Package — did not scale well enough to be used on the Earth Simulator. MADCAP, developed by Julian Borrill, is a parallel implementation of cosmic microwave background map-making and power spectrum estimation algorithms. Since MADCAP has high I/O requirements, its performance was hampered by the lack of a fast global file system on the Earth Simulator.

Undeterred, the team re-tuned MADCAP and returned to Japan to try again. The resulting paper[6] found that the Cray X1 had the best runtimes but suffered the lowest parallel efficiency. The Earth Simulator and IBM Power3 demonstrated the best scalability, and the code achieved the highest percentage of peak on the Power3. The paper concluded, "Our results highlight the complex interplay between the problem size, architectural paradigm, interconnect, and vendor-supplied numerical libraries, while isolating the I/O filesystem as the key bottleneck across all the platforms."

As for BlueGene/L, currently the world's fastest supercomputer, David Skinner is serving as Berkeley Lab's representative to a new BlueGene/L consortium led by Argonne National Laboratory. (The first Blue Gene

---

[6] J. Carter, J. Borrill, and L. Oliker, "Performance characteristics of a cosmology package on leading HPC architectures," presented at HiPC 2004, The Eleventh International Conference on HPC, Bangalore, India, December 2004; http://crd.lbl.gov/~oliker/papers/HIPC04.pdf.

system is being installed at Lawrence Livermore National Laboratory.) The consortium aims to pull together a group of institutions active in HPC research, collectively building a community focused on the BlueGene family as a next step towards petascale computing. This consortium will work together to develop or port BlueGene applications and system software, conduct detailed performance analysis on applications, develop mutual training and support mechanisms, and contribute to future platform directions.

Science-Driven Architecture Ideas Go into Production

Two years after NERSC led a team from seven national laboratories to propose a new HPC architecture designed specifically to meet the needs of computational scientists, such an architecture is being used for the ASC Purple supercomputer being built by IBM for Lawrence Livermore National Laboratory (LLNL).

The 2003 white paper, entitled "Creating Science-Driven Computer Architecture: A New Path to Scientific Leadership,"[7] was submitted to the High-End Computing Revitalization Task Force (HECRTF), whose goal was to re-establish the United States as the clear leader in high-performance computing. The white paper called for moving away from dependence on hardware that is designed and optimized for commercial applications and to create a new class of computational capability in the United States that is optimal for science — a strategy whose broad outline was adopted in HECRTF's final report.[8]

This comprehensive strategy includes development partnerships with multiple vendors, in which teams of scientific applications specialists and computer scientists will work with computer architects from major

U.S. vendors to create hardware and software environments that will allow scientists to extract the maximum performance and capability from the hardware. The white paper recommended pursuing three different options to determine the best price-performance in a scientific computing environment.

The second option called for using commercial microprocessors in a new architecture known as ViVA or Virtual Vector Architecture. ViVA was based on enhancing IBM's existing Power5 plan in order to deploy a system with high sustained performance (up to one-third of peak) running real scientific codes. Each node would consist of eight single-core CPUs that would provide double the memory bandwidth of the standard Power5 CPUs and would have their own dedicated L1, L2, and L3 caches. The CPUs would have a peak performance of roughly 8–10 Gflop/s.

This architecture caught the attention of ASC (Advanced Simulation and Computing) managers at LLNL, who were working with IBM to design and build ASC Purple. When the final architecture of ASC Purple was agreed upon, it was designed around the eight-way node. As a result, the system is expected to be less expensive than originally expected, as well as more effective for carrying out its intended mission.

"We owe you a debt of gratitude for conceiving the eight-way node with IBM and pushing so strongly for it with IBM," wrote Michael McCoy of LLNL when informing NERSC Center General Manager Bill Kramer of the ASC Purple decision. "This is a clear example of ASC benefiting from Office of Science — and NERSC in particular — working the system for better solutions for science. Also, the cost of the system is much lower than the original thanks in part to the

[7] H. D. Simon, C. W. McCurdy, W. T.C. Kramer, R. Stevens, M. McCoy, M. Seager, T. Zacharia, J. Nichols, R. Bair, S. Studham, W. Camp, R. Leland, J. Morrison, and B. Feiereisen, "Creating Science-Driven Computer Architecture: A New Path to Scientific Leadership," May 2003; http://www.nersc.gov/news/reports/HECRTF-V4-2003.pdf.

[8] *Federal Plan for High-End Computing: Report of the High-End Computing Revitalization Task Force (HECRTF),* (Arlington, VA: National Coordination Office for Information Technology Research and Development, May 10, 2004); http://www.house.gov/science/hearings/full04/may13/hecrtf.pdf.

new node. We all owe NERSC thanks for vision and energy in this singular effort."

The 12,544-processor ASC Purple system is scheduled to be installed at Lawrence Livermore in July 2005.

Streamlining File Storage

In 2001, NERSC began a project to deploy a system to streamline its file storage system using existing and emerging technologies. The Global Unified Parallel File System (GUPFS) project aims to provide a scalable, high-performance, high-bandwidth, shared file system for use by all of NERSC's high-performance production computational systems. GUPFS will provide unified file namespace for these systems and will be integrated with HPSS. Storage servers, accessing the consolidated storage through the GUPFS shared-disk file systems, will provide hierarchical storage management, backup, and archival services. An additional goal is to distribute GUPFS-based file systems to geographically remote facilities as native file systems over the DOE Science Grid.

When fully implemented, GUPFS will eliminate unnecessary data replication, simplify the user environment, provide better distribution of storage resources, and permit the management of storage as a separate entity while minimizing impacts on the NERSC computational systems.

The major enabling components of this envisioned environment are a high-performance shared file system and cost-effective, high performance storage-area networks (SANs) and emerging alternative fabrics. These technologies, while evolving rapidly, are generally not targeted towards the needs of high-performance scientific computing. The GUPFS project is working with vendors to incorporate HPC requirements into their products.

During the first two years of the project, the GUPFS team, led by Greg Butler of the Advanced Systems Group, tested and investigated shared file systems, SAN technologies, and other components of the GUPFS environment. The team has successfully assembled a complex testbed simulating the envisioned NERSC storage environment, with which they have developed the knowledge base needed to select and deploy production-quality solutions.

Now, during the third year, their focus has shifted from component evaluation to deployment planning — evaluating solutions, systems, and deployment scenarios with an eye toward issuing formal requests for information and proposals. The team is also narrowing its focus to candidate file systems.

"Fabric and storage technologies have been advancing and are not major concerns," Butler said. "But file systems, while they are improving, still need work to meet the needs of parallel production environments, especially in terms of stability, parallel I/O performance and functionality, and scalability. File systems also need to support more platforms and operating systems. And vendors should consider open-source software if they want their products to be widely used."

APPENDIX A
NERSC Policy Board

Robert J. Goldston
Princeton Plasma Physics Laboratory

Stephen C. Jardin
(ex officio, NERSC Computational Review Panel)
Princeton Plasma Physics Laboratory

Sidney Karin
University of California, San Diego

William J. Madia
Battelle Memorial Institute

Paul C. Messina
Argonne National Laboratory

Albert Narath
Lockheed-Martin Corporation (retired)

Daniel A. Reed
University of North Carolina

Robert D. Ryne
(ex officio, NERSC Users Group Chair 2004)
Lawrence Berkeley National Laboratory

Tetsuya Sato
Earth Simulator Center/Japan Marine Science and
Technology Center

Stephen L. Squires
Hewlett-Packard

Michael S. Witherell
Fermi National Accelerator Laboratory

APPENDIX B
NERSC Client Statistics

NERSC served 2,819 scientists throughout the United States in 2004. 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.

Industries
(3%)

Universities
(47%)

Other Labs
(5%)

DOE Labs
(45%)
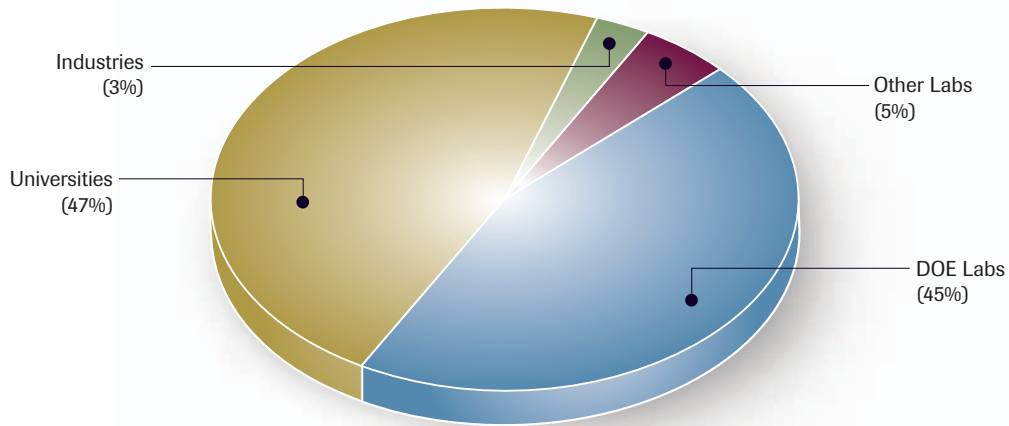
FIGURE 1   NERSC usage by institution type, 2004.

Brookhaven
(654,174)

National Renewable Energy
(892,416)

Stanford Linear Accelerator
(570,067)

Los Alamos
(955,884)

Sandia
(324,510)

Argonne
(1,638,300)

Pacific Northwest
(293,972)

National Center for
Atmospheric Research
(2,403,412)

Ames
(166,182)

Princeton Plasma Physics
(3,113,909)

Eight Others
(223,076)

Oak Ridge
(3,167,098)

Lawrence Berkeley
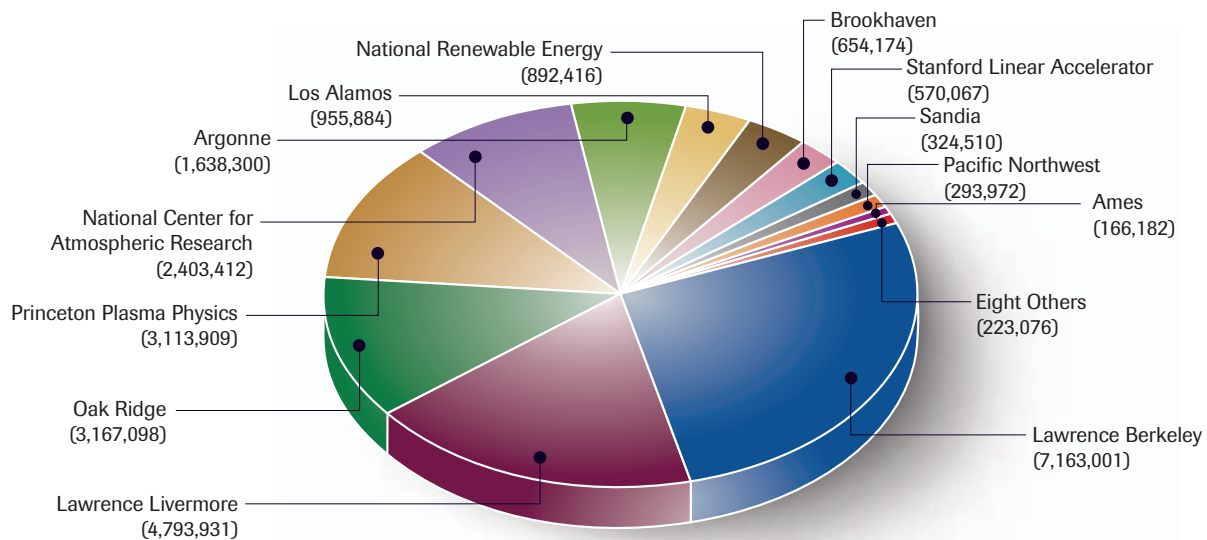(7,163,001)

Lawrence Livermore
(4,793,931)

FIGURE 2   DOE and other Federal laboratory usage at NERSC, 2004 (processor hours).
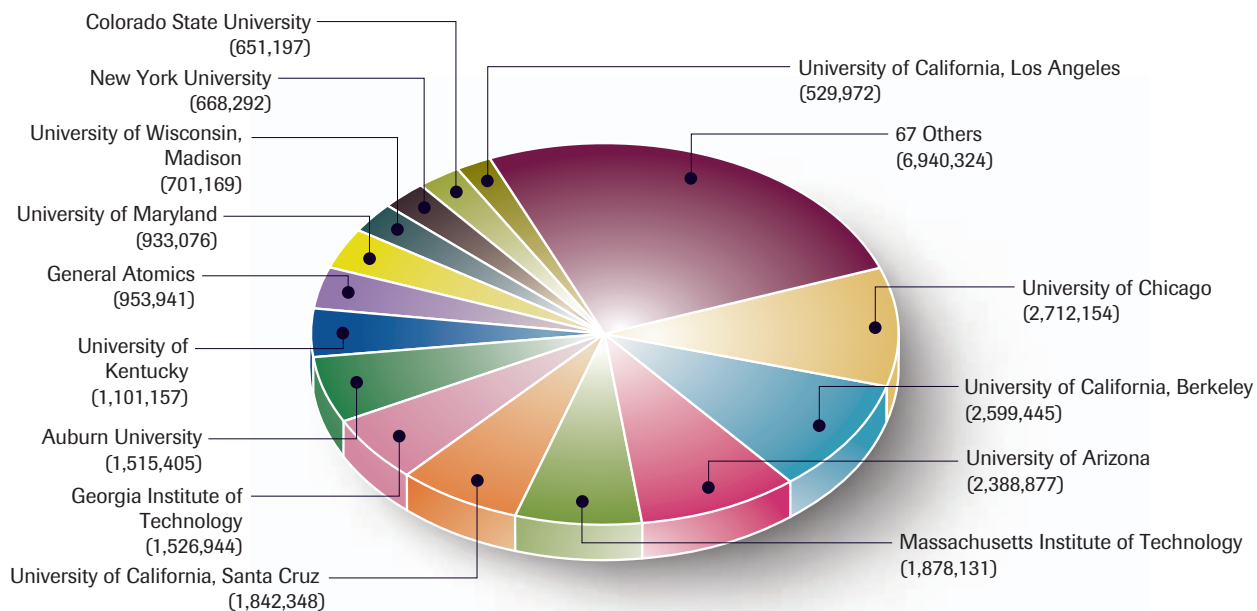
FIGURE 3   Academic and private laboratory usage at NERSC, 2004 (processor hours).

University of California, Los Angeles (529,972)
67 Others (6,940,324)
Colorado State University (651,197)
New York University (668,292)
University of Wisconsin, Madison (701,169)
University of Maryland (933,076)
General Atomics (953,941)
University of Kentucky (1,101,157)
Auburn University (1,515,405)
Georgia Institute of Technology (1,526,944)
University of California, Santa Cruz (1,842,348)
University of Chicago (2,712,154)
University of California, Berkeley (2,599,445)
University of Arizona (2,388,877)
Massachusetts Institute of Technology (1,878,131)

FIGURE 4   NERSC usage by scientific discipline, 2004.

Applied Math and Computer Science (4%)
Accelerator Physics (5%)
Materials Science (8%)
Engineering and Geosciences (4%)
Chemistry (9%)
Nuclear Physics (2%)
Climate Research (9%)
Life Sciences (2%)
Lattice QCD (11%)
Astrophysics (17%)
Fusion Energy (29%)

APPENDIX C
NERSC Computational Review Panel (CORP)

Computational Science

Accelerator Physics
Robert Ryne
Lawrence Berkeley National Laboratory

Applied Mathematics
David Bailey
Lawrence Berkeley National Laboratory

Astrophysics
Douglas Swesty
State University of New York at Stony Brook

Chemistry
David Dixon
University of Alabama

Robert Harrison
Oak Ridge National Laboratory

Climate Research
Ferdinand Baer
University of Maryland

Douglas Rotman
Lawrence Livermore National Laboratory

Computer Science
David Bailey
Lawrence Berkeley National Laboratory

Engineering
Pui-kuen Yeung
Georgia Institute of Technology

Environmental Sciences
Ferdinand Baer
University of Maryland

Douglas Rotman
Lawrence Livermore National Laboratory

Fusion Energy
Stephen Jardin
Princeton Plasma Physics Laboratory

Charlson Kim
University of Wisconsin, Madison

Geosciences
Andrey Kalinichev
University of Illinois at Urbana-Champaign

Pui-kuen Yeung
Georgia Institute of Technology

High Energy Physics
Robert Sugar
University of California, Santa Barbara

Lattice Gauge Theory (QCD)
Robert Sugar
University of California, Santa Barbara

Life Sciences
Brian Hingerty
Oak Ridge National Laboratory

Material Sciences
David Turner
Ames Laboratory

Michael Weinert
University of Wisconsin, Milwaukee

Nuclear Physics
Douglas Swesty
State University of New York at Stony Brook

Applied Mathematics

Jesse Barlow
Pennsylvania State University

Ralph Byers
University of Kansas

Marc Day
Lawrence Berkeley National Laboratory

Eric de Sturler
University of Illinois at Urbana-Champaign

Tony Drummond
Lawrence Berkeley National Laboratory

Joseph Grcar
Lawrence Berkeley National Laboratory

Ilse Ipsen
North Carolina State University

Ricky Kendall
Ames Laboratory

Esmond Ng
Lawrence Berkeley National
Laboratory

Jeffrey Vetter
Oak Ridge National Laboratory

Robert Ward
University of Tennessee

Computer Science

David Bernholdt
Oak Ridge National Laboratory

Brett Bode
Ames Laboratory

Erik DeBenedictis
Sandia National Laboratories

Mary Hall
Information Sciences Institute

James Kohl
Oak Ridge National Laboratory

Shirley Moore
University of Tennessee

Robert Numrich
University of Minnesota

Allan Snavely
San Diego Supercomputer Center

Erich Strohmaier
Lawrence Berkeley National
Laboratory

Rajeev Thakur
Argonne National Laboratory

Mary Zosel
Lawrence Livermore National
Laboratory

NERSC Staff Reviewers

Accelerator Physics
Parry Husbands

Applied Mathematics
Osni Marques

Astrophysics
Julian Borrill
Peter Nugent

Chemistry
Jonathan Carter
David Skinner

Climate Research
Chris Ding
Helen He
Michael Wehner

Computer Science
Parry Husbands

Engineering
Osni Marques

Environmental Sciences
Chris Ding
Helen He
Michael Wehner

Fusion Energy
Thomas DeBoni
Jodi Lamoureux

Geosciences
Osni Marques

High Energy Physics
Iwona Sakrejda

Lattice Gauge Theory (QCD)
Radek Stompor

Life Sciences
Adrian Wong
Jonathan Carter

Material Sciences
Andrew Canning
Lin-Wang Wang

Nuclear Physics
Iwona Sakrejda

APPENDIX D
NERSC Users Group Executive Committee

Members at Large

Yuen-Dat Chan,** Lawrence Berkeley National Laboratory

Stephen C. Jardin,* Princeton Plasma Physics Laboratory

Ricky A. Kendall, Ames Laboratory, Iowa State University

Donald Sinclair,* Argonne National Laboratory

Donald A. Spong, Oak Ridge National Laboratory

Douglas Swesty, State University of New York at Stony Brook

John A. Taylor,* Argonne National Laboratory

Office of Advanced Scientific Computing Research

David Bernholdt, Oak Ridge National Laboratory

David Keyes, Old Dominion University

David Turner, Ames Laboratory

Office of Basic Energy Sciences

Feng Liu, University of Utah

G. Malcolm Stocks, Oak Ridge National Laboratory

Theresa L. Windus, Pacific Northwest National Laboratory

Office of Biological and Environmental Research

Michael Colvin, University of California, Merced, and Lawrence Livermore National Laboratory

Brian Hingerty,** Oak Ridge National Laboratory

Douglas Rotman,* Lawrence Livermore National Laboratory

Vincent Wayland, National Center for Atmospheric Research

Office of Fusion Energy Sciences

Stephane Ethier, Princeton Plasma Physics Laboratory

Alex Friedman, Lawrence Livermore and Lawrence Berkeley National Laboratories

Carl Sovinec, University of Wisconsin, Madison

Office of High Energy Physics

Olga Barranikova,** Purdue University

Kwok Ko,* Stanford Linear Accelerator Center

Warren Mori,** University of California, Los Angeles

Robert D. Ryne,* Lawrence Berkeley National Laboratory

Donald Sinclair,** Argonne National Laboratory

Office of Nuclear Physics

David Dean,** Oak Ridge National Laboratory

Patrick Decowski,** Lawrence Berkeley National Laboratory

Douglas L. Olson, Lawrence Berkeley National Laboratory

* Outgoing members   ** Incoming members

A P P E N D I X   E
Supercomputing Allocations Committee

The Supercomputing Allocations Committee (SAC) is responsible for setting the policies associated with the utilization of computing, data storage, and other auxiliary services available to DOE Office of Science (SC) researchers and otherwise coordinating SC's computational projects. The Committee sets the distribution of NERSC and other available Office of Advanced Scientific Computing Research resources for scientific computing applications every year.

Committee Chair

David Goodwin

Office of Advanced Scientific Computing Research

David Goodwin

Office of Basic Energy Sciences

Richard Hilderbrandt

Office of Biological and Environmental Research

Michael Riches

Office of Fusion Energy Sciences

John Mandrekas

Office of High Energy Physics

Glen Crawford

Office of Nuclear Physics

Sidney Coon

APPENDIX F
Office of Advanced Scientific Computing Research

Mathematical, Information, and Computational Sciences Division

The primary mission of the Advanced Scientific Computing Research (ASCR) program, which is carried out by the Mathematical, Information, and Computational Sciences (MICS) subprogram, 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 MICS research projects described in this report:

C. Edward Oliver
Associate Director, Office of Advanced Scientific Computing Research

Michael Strayer
Acting MICS Director and SciDAC Program Director

David Goodwin
NERSC Program Manager

Daniel Hitchcock
ASCR Senior Technical Advisor

Frederick C. Johnson
Computer Science Research

Gary M. Johnson
Applied Mathematics Research (Acting)

Thomas D. Ndousse-Fetter
Networking

Walter M. Polansky
ASCR Senior Technical Advisor for Project Management

Charles Romine
Special Assignment to the MICS Director

Mary Anne Scott
Collaboratory Research and ESnet

George R. Seweryniak
Program Manager

APPENDIX G
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.

Members

Jill P. Dahlburg
Naval Research Laboratory

David J. Galas
Keck Graduate Institute

James J. Hack
National Center for Atmospheric Research

Thomas A. Manteuffel
Department of Applied Mathematics
University of Colorado at Boulder

APPENDIX H
Organizational Acronyms

DOE . . . . U.S. Department of Energy

ASC . . . . Advanced Simulation and Computing Program, National Nuclear Security Administration

ASCR-MICS . . . Office of Advanced Scientific Computing Research, Mathematical, Information, and Computational Sciences Division

BER . . . . Office of Biological and Environmental Research

BES . . . . Office of Basic Energy Sciences

FES . . . . Office of Fusion Energy Sciences

HEP . . . . Office of High Energy Physics

INCITE . . . . Innovative and Novel Computational Impact on Theory and Experiment Program

NP . . . . Office of Nuclear Physics

SC . . . . Office of Science

SciDAC . . . . Scientific Discovery through Advanced Computing Program

3M . . . . 3M Company

ACS . . . . American Chemical Society

AF . . . . Academy of Finland

ASAP . . . . Academic Strategic Alliances Program of the DOE Advanced Simulation and Computing (ASC) Program

AVHF . . . . Alexander von Humboldt Foundation

BMBF . . . . Bundesministerium für Bildung und Forschung (Germany)

CARA . . . . Center for Astrophysics Research in Antarctica

CSIR . . . . Council of Scientific and Industrial Research (India)

DAE . . . . Department of Atomic Energy (India)

DRC . . . . Danish Research Council

DST . . . . Department of Science and Technology (India)

ECRTN . . . . European Commission Research Training Network

EPSRC . . . . Engineering and Physical Sciences Research Council (UK)

ERDF . . . . European Regional Development Fund

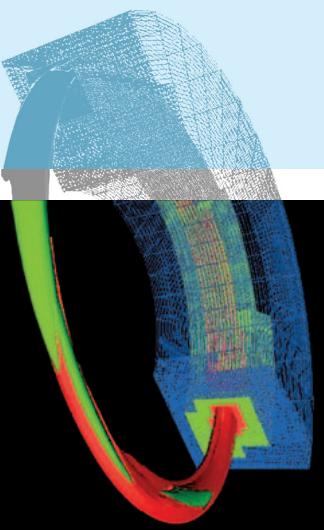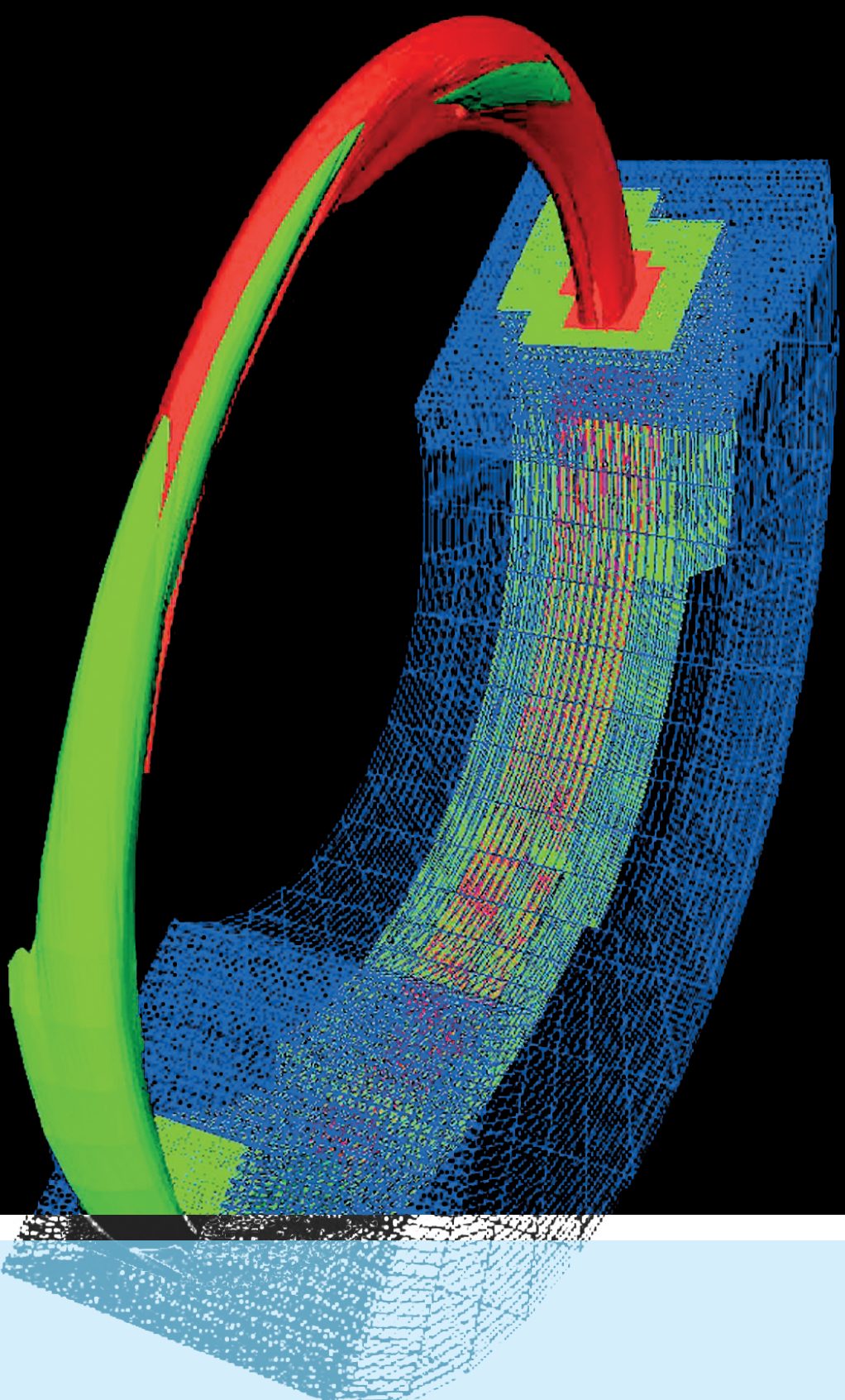ESOP . . . . European Research Training Network: Electromagnetic Scattering off Confined Partons

FAPESP . . . . . . Fundação de Amparo à Pesquisa do
Estado de São Paulo (Brazil)

FCT . . . . Faculdade de Ciências e Tecnologia
da Universidade Nova de Lisboa
(Portugal)

FEPCJ . . . . Federation of Electric Power
Companies Japan

FOM. . . . . Fundamenteel Onderzoek der
Materie (Netherlands)

FWH . . . . F. Warren Hellman Family

GACR . . . . Grant Agency of the Czech
Republic

Hertz . . . . Hertz Foundation

IN2P3 . . . . Institut National de la Physique
Nucléaire et de la Physique des
Particules (France)

INSU . . . . Institut National des Sciences de
l'Univers (France)

JMECSST . . . . . Japanese Ministry of Education,
Culture, Sports, Science and
Technology

KBSI . . . . Korea Basic Science Institute

KICP . . . . Kavli Institute for Cosmological
Physics at the University of
Chicago

KMSC . . . . Kamioka Mining and Smelting
Company

LBNL . . . . Lawrence Berkeley National
Laboratory

LLNL . . . . Lawrence Livermore National
Laboratory

LF . . . . Levendis Foundation

MCIF . . . . Marie Curie Individual Fellowship

MCyT . . . . Ministerio de Ciencia y Tecnología
(Spain)

MEC. . . . . Ministry of Education of China

NASA . . . . National Aeronautics and Space
Administration

NDSEG. . . . . . . National Defense Science and
Engineering Graduate Fellowship
Program

NIH . . . . National Institutes of Health

NNSFC. . . . . . . National Natural Science
Foundation of China

NSF . . . . National Science Foundation

PFST . . . . Portugese Foundation for Science
and Technology

PPARC . . . . Particle Physics and Astronomy
Research Council (UK)

PSCSR . . . . Polish State Committee for
Scientific Research

RC . . . . Research Corporation

RMST. . . . . Russian Ministry of Science and
Technology

SNSF . . . . Swiss National Science Foundation

UC . . . . University of Cyprus

UCB. . . . . University of California, Berkeley

UGA. . . . . University of Georgia

UU. . . . . University of Utrecht (Netherlands)

UWM . . . . University of Wisconsin, Madison

Yale . . . . Yale University

National Energy Research Scientific Computing Center

2004 annual report

DISCLAIMER

Cover image: Visualization based on a simulation of the density of a fuel pellet after it is injected into a tokamak fusion reactor. See page 40 for more information.