BREAKTHROUGHS *2008*

*applied mathematics,
astrophysics, biology, chemistry,
combustion-energy science,
materials science, computer science,
fusion-energy science*

BREAKTHROUGHS

CONTENTS

Introduction

Society in the last decade has benefitted from significant investment in computational science. This investment has elevated scientific computing to the terascale—research conducted at trillions of calculations each second—and the accomplishments detailed in this report are a testament to its efficacy.

As we move forward over the next decade, through the petascale (thousands of trillions of calculations each second) to the exascale (millions of trillions of calculations each second), the opportunities for ground-breaking science will increase even more dramatically. Commensurately, the scale of investment required to realize these opportunities will also increase dramatically. Continued, and certainly growing, investment will require a concerted effort by the computational science community to articulate the major successes achieved through it.

To this end, a panel of computational scientists, applied mathematicians, and computer scientists gathered in February 2008 at the invitation of Dr. Michael Strayer, Associate Director of the Department of Energy Office of Advanced Scientific Computing Research. The panel was charged with identifying recent breakthroughs in computational science and enabling technologies, supported in a broad sense by the Office of Advanced Scientific Computing Research through the Innovative and Novel Computational Impact on Theory and Experiment (IN-CITE) program, the Scientific Discovery through Advanced Computing (SciDAC) program, and/or its base program. This report is the result of that effort; each accomplishment detailed here significantly advanced a scientific frontier or provided a breakthrough in applied mathematics, computer science, or both that enabled such advances. All were achieved and/or published during the past 18 months.

Panel members represented all major areas in the Office of Science's computational-science portfolio: applied mathematics, astrophysics, biology, chemistry, climate science, combustion-energy science, computer science, fusion-energy science, high-energy physics, materials science, and nuclear physics. Among these top researchers were leads from the national computational science centers at Argonne, Lawrence Berkeley, and Oak Ridge national laboratories. The combination of domain expertise and center leadership proved very effective in providing both the needed depth and the ability to make comparative assessments of accomplishments across a broad range of areas.

Each accomplishment in the report was summarized using a standardized template initially provided by the Office of Advanced Scientific Computing Research but then further developed in conjunction with the panel. The template served to ensure all key information was included in order to accurately assess the accomplishments and provide a way to normalize the accomplishments given the diversity of fields from which the panel drew. The template was based on the following key questions: (1) What is the accomplishment? (2) What is the significance of the accomplishment? (3) What impact will the accomplishment have on the Department of Energy mission in science, computing, energy, and/or the environment? (4) What resources and approaches (e.g., computing resources,

codes) were used? (5) What role did university and laboratory collaborations play? (6) What funding sources and computing programs supported this work? (7) What important publications resulted? (8) What future research is planned as a result of the accomplishment?

Given the diversity of fields considered and the number of accomplishments from which to choose, the panel faced a difficult task. Nonetheless, a handful of singular achievements did emerge, telling an incredible story of computational discovery. In this report we describe, among other things,

- the elucidation of the molecular mechanism underlying the progression of Parkinson's disease;

- the discovery of a new and critical phenomenon in the deaths of massive stars, which produce the elements necessary for life, and a new mechanism for the birth of pulsars, nature's lighthouses;

- the first predictions of protein structure with atomic accuracy, a Holy Grail of molecular biology for more than 30 years;

- a breakthrough in understanding how turbulent flames stabilize in combustion devices, with ramifications for the design of gasoline engines, diesel engines, and gas turbines—or most everything we rely on to power transportation;

- a fundamentally new understanding of thermal-energy loss in tokomak fusion reactors, with implications for the design of the $10 billion international ITER device; and

- the fruition of models of high-temperature superconductivity in real materials that position us that much closer to designing superconducting materials and realizing superconductivity in many important practical applications.

Of course, this compelling computational-science story could not be told without the underlying breakthroughs in applied mathematics and computer science that made it possible. The behavior of physical, chemical, and biological systems is often governed by mathematical equations that encode the laws of nature. For example, the dynamic evolution of physical systems in astrophysics, climate science, combustion science, and fusion-energy science are governed by a type of equation known as a partial differen-

tial equation. The complexity of partial differential equations precludes pencil-and-paper solutions, and computational approaches must be used. In turn, computational methods are almost always based on algorithms that solve more fundamental mathematical problems, such as large systems of equations in linear algebra. To advance their respective frontiers, scientists ultimately rely on high-performance software that combines state-of-the-art applied mathematics and computer science to achieve such solutions on today's largest, fastest supercomputers.

In this report we highlight what this panel considered the most significant of such software used by the computational-science community. The Portable Extensible Toolkit for Scientific computation (or PETSc), which is now solving algebraic systems of equations with more than three billion unknowns, has scaled to tens of thousands of processors on leadership-class supercomputers and enabled leadership-class research in many areas of science, such as global climate modeling, as well as significant engineering advances.

Anthony Mezzacappa, Panel Chair
August 5, 2008
Oak Ridge

FOLLOWING ARE THE PROJECTS FEATURED IN THIS REPORT LISTED IN ORDER OF THEIR PANEL RANKING:

1. Scientists Model the Molecular Basis of Parkinson's Disease

2. Astrophysicists Discover Supernova Shock-Wave Instability and a Better Way to Spin Up Pulsars

3. Designing Proteins at Atomic Scale and Creating Enzymes

4. First-Principles Flame Simulation Provides Crucial Information to Guide Design of Fuel-Efficient Clean Engines

5. Breakthrough Fusion Simulation Sheds Light on Plasma Confinement

6. Closing In on an Explanation for High-Temperature Superconductivity

7. Powerful Mathematical Tools Resolve Complex Simulations

8. A Billion-Particle Simulation of the Dark Matter Halo of the Milky Way

9. Exploring the Mysteries of Water

10. Novel Solver Enables Scalable Electromagnetic Simulations
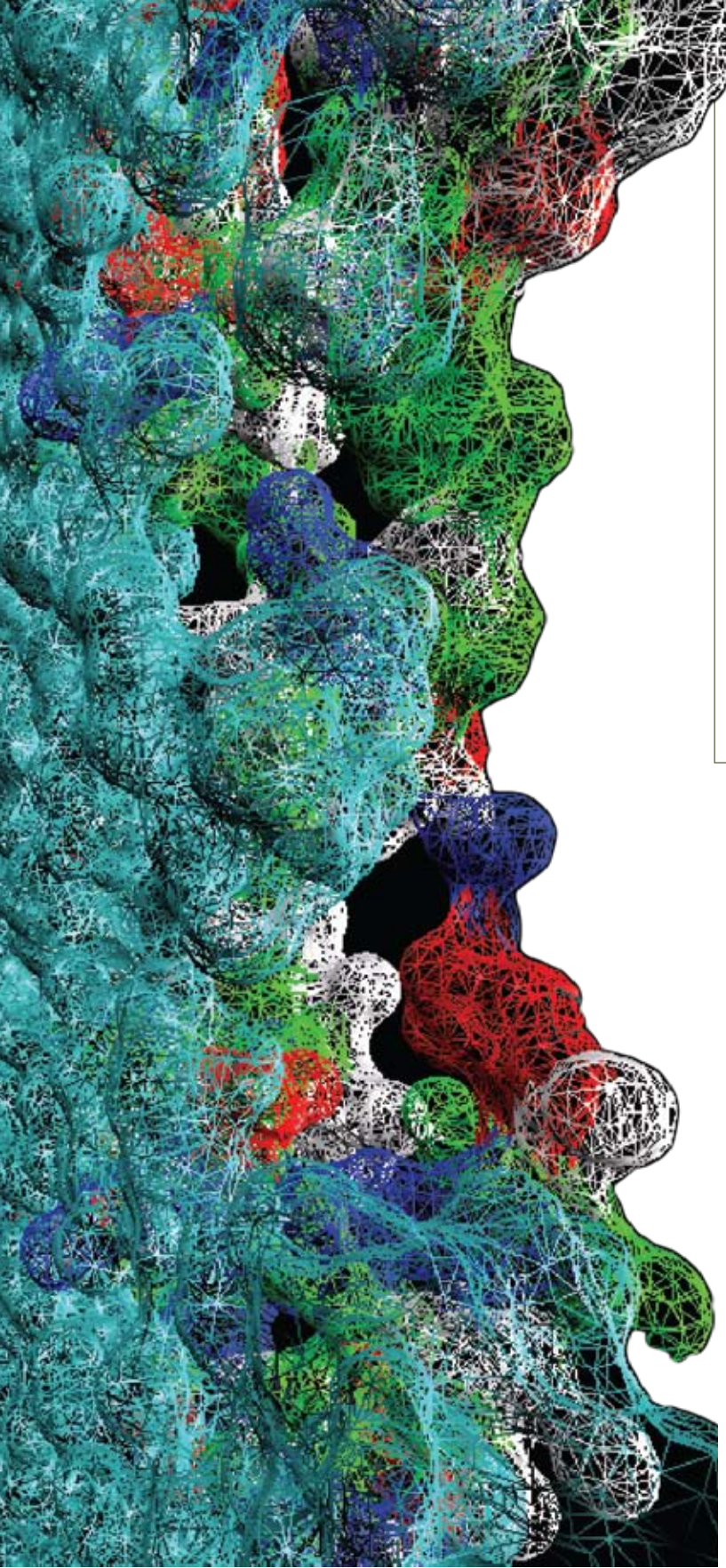
A membrane embedded with alpha-synuclein rings.

Image courtesy I.F. Tsigelny, SDSC at UCSD.

Scientists Model the Molecular Basis of Parkinson's Disease

ABSTRACT: Aided by supercomputers, University of California–San Diego researchers were the first to elucidate the molecular mechanism by which protein units assemble into ringlike structures that spur Parkinson's disease. Their work may guide scientists in disrupting the molecular mechanisms that underlie other neurodegenerative disorders, including Alzheimer's and Huntington's diseases.

Alpha-synuclein is a protein in the brain that may aid communication between nerve cells. Scientists call it an "unstructured" protein because it is not stable enough to remain in one form, or conformation. In fact, genes or environment can set the protein on a path that leads to Parkinson's disease. It was known that alpha-synuclein can organize into fibrils, or long fibers. Until recently, however, no one knew exactly how the protein triggered disease. That changed when computational biologist Igor Tsigelny of the San Diego Supercomputer Center (SDSC) at the University of California–San Diego (UCSD) and neuroscientist Eliezer Masliah of UCSD used supercomputers to model its behavior in the cell and reveal how it causes neuron death.

The team computationally explored alpha-synuclein conformations that were energetically feasible as the protein interacted with the cell membrane. Simulation with biochemical analysis and electron microscopy showed that unstructured units of alpha-synuclein can form ringlike aggregates on membranes.

For a long time physicians believed accumulating fibrils of alpha-synuclein overwhelmed cells to cause Parkinson's disease. Tsigelny and Masliah's work supports a different causal mechanism, proposed by neuroscientist Peter Lansbury in 2002. Their simulation showed that alpha-synuclein can aggregate into a ringlike structure that penetrates the cell membrane to form a pore. Because the concentration of calcium outside cells is much greater than inside, calcium flows through the pore and triggers a cascade of events leading to cell death, Tsigelny says. Electrophysiological studies support the pore theory by showing current flow caused when ions traverse membranes. When the cells that make the neurotransmitter dopamine die, Parkinson's is born. Its symptoms include tremor and rigidity as well as cognitive and mood disorders.

"We work on the drugs that can prevent aggregation to this ringlike structure organization," Tsigelny says. "We clearly find

on the basis of modeling that specific regions need to be isolated from each other, and the drugs that can do that can prevent aggregation."

The findings provide insights into the molecular mechanism for Parkinson's disease progression and may help researchers develop strategies for prevention or cure. Preventing alpha-synuclein from assembling into rings sets a goal for rational design of drugs that treat the disease in an entirely different way than do current treatments, which merely compensate for lost dopamine. Halting pore formation may stop the disease from progressing once it has begun or even prevent the disease from developing if an anti-Parkinson's pill is given prophylactically. Laboratory tests with mice show promise, Tsigelny says, with drugs blocking aggregation of alpha-synuclein. The findings may have broad applicability to other neurodegenerative diseases and provide a test bed for identifying therapeutic interventions through computational modeling. Misfolded or aggregated proteins are linked to Alzheimer's, Huntington's, Creutzfeldt-Jakob, and other neurodegenerative diseases.

SUPERCOMPUTERS ARE INDISPENSABLE

This research supports the primary mission of the Department of Energy's (DOE's) Office of Advanced Scientific Computing Research (OASCR) program to discover, develop, and deploy computational and networking tools that enable scientists to analyze, model, simulate, and predict complex phenomena. By using the most powerful openly available computers and programs in the world to gain breakthrough insights into an important disease mechanism, the research strengthens U.S. scientific discovery and has the potential to boost economic competitiveness and accelerate innovations likely to improve the quality of life.

Allocations of supercomputing resources enabled researchers to perform complex calculations using Blue Gene/L computers at the Argonne Leadership Computing Facility (ALCF) at Argonne National Laboratory (ANL) and at SDSC. The sophisticated simulations of molecular dynamics required 962,757 processor hours, using 512 processors per run. "The experiments simply could not have been done without supercomputers," Tsigelny says. "Unstructured proteins have zillions of possible conformations, and nobody knows which

conformations will actually bind to the membrane and create these pores."

The team used and developed highly parallel programs as part of a multiprogram, multidata system for studying large multimolecular biological systems. These programs included parallel versions of the NAMD code for molecular dynamics calculations, the DOT code for molecular docking calculations, and the MAPAS code for calculation of membrane-association scores for proteins and protein aggregates. The programs enabled the researchers to predict conformational changes of proteins, protein–protein interactions including aggregation, and interaction of proteins as individuals or as part of a membrane complex. The software developed for the alpha-synuclein studies might provide the basis for a package for study of other unstructured proteins, such as beta-amyloid, which is responsible for Alzheimer's disease, and prions, which cause many diseases including bovine spongiform encephalopathy ("mad cow disease") in cattle and Creutzfeldt-Jakob disease in humans.

It is also important to note that the developed program system can be used for computational construction of specialized channels in microorganisms. Tsigelny says this approach could be useful for designing microorganisms specialized for absorption of various environmental wastes.

As the work progresses, the researchers will focus on a more comprehensive investigation of alpha-synuclein penetration into the membrane, including a thorough study of pore creation. The scope has increased in both the number and scale of simulations, which now model a system containing approximately 800,000 atoms. In addition, systems including beta-amyloid peptide will be studied. The simulations will focus on membrane interactions with higher-level alpha-synuclein aggregates. Given agreement between computational predictions and experimental observations, Tsigelny and Masliah say they expect to make steady progress modeling the disease and designing drugs. "We can create a drug-processing machine on which we can run any of the amyloids and try drugs to see whether they will fit or not," Tsigelny says. "We can probe specific places of interactions, where we can put drugs and check them. We don't claim to have invented these porelike structures, but we first managed to use them for rational and computer-aided drug design."

Contributors/Authors

Igor Tsigelny, SDSC at UCSD; Eliezer Masliah, UCSD

Principal Collaborators

Tsigelny collaborated with several UCSD colleagues; Leslie Crews, Pazit Bar-On, and Makoto Hashimoto did experiments with alpha-synuclein molecules. Yuriy Sharikov of SDSC did extensive programming and simulations. Mark Miller of SDSC organized computational work. Steven Keller performed biochemical experiments. Oleksandr Platoshyn and Jason Yuan did electrophysiological experiments. Eliezer Masliah organized the entire set of experiments and contributed to understanding of the molecular mechanism of the disease.

Publications

Y. Sharikov et al. 2008. "MAPAS: A tool for predicting membrane-contacting protein surfaces," *Nature Methods* **5**(2): 119.

I.F. Tsigelny et al. 2007. "Dynamics of α-synuclein aggregation and inhibition of pore-like oligomer development by β-synuclein," *Federation of European Biochemical Societies Journal* **274**(7): 1862–1877.

I.F. Tsigelny et al. 2008. "Mechanism of alpha-synuclein oligomerization and membrane interaction: Theoretical approach to unstructured protein studies," *Nanomedicine* (in press).

Astrophysicists Discover Supernova Shock-Wave Instability and a Better Way to Spin Up Pulsars

ABSTRACT: Researchers using Oak Ridge National Laboratory's Jaguar and Phoenix supercomputers have provided two key pieces in the puzzle of the core-collapse supernova, demonstrating that the supernova shock wave is unstable and showing how the shock-wave instability may spin up the leftover pulsar. Their efforts help us understand how some of the universe's most dramatic catastrophes are responsible for producing and spreading the building blocks of life.

John Blondin of North Carolina State University has used supercomputer simulation to make two revolutionary advances in our understanding of core-collapse supernovas—the exploding stars that provide most of the elements in the universe. Working in a team led by Oak Ridge National Laboratory's (ORNL's) Anthony Mezza-cappa, Blondin discovered that the shock wave created by the star's collapsing iron core becomes inherently unstable as it is stalled by infalling material from the star. In a second major discovery, Blondin showed that the evolution of this stalled shock wave may well be responsible for the spin of the spinning neutron star—or pulsar—that remains after most of the star is blown into space.

Core-collapse supernovas take place about twice per century in our galaxy. Stars 10 or more times the mass of the sun die in catastrophic stellar explosions and, in so doing, seed the galaxy with the building blocks of planets and life. The star's death is, in fact, an inevitable outcome of its evolution. Over millions of years, the star burns through its nuclear fuel, fusing atoms to create increasingly heavy elements. These elements form into layers, from hydrogen at the surface through helium, carbon, oxygen, and silicon, to iron at the core. The process, however, hits a roadblock at iron, which does not create energy when the atoms are fused. Eventually the iron core becomes so heavy that it collapses under its own weight and becomes a mass of neutrons.
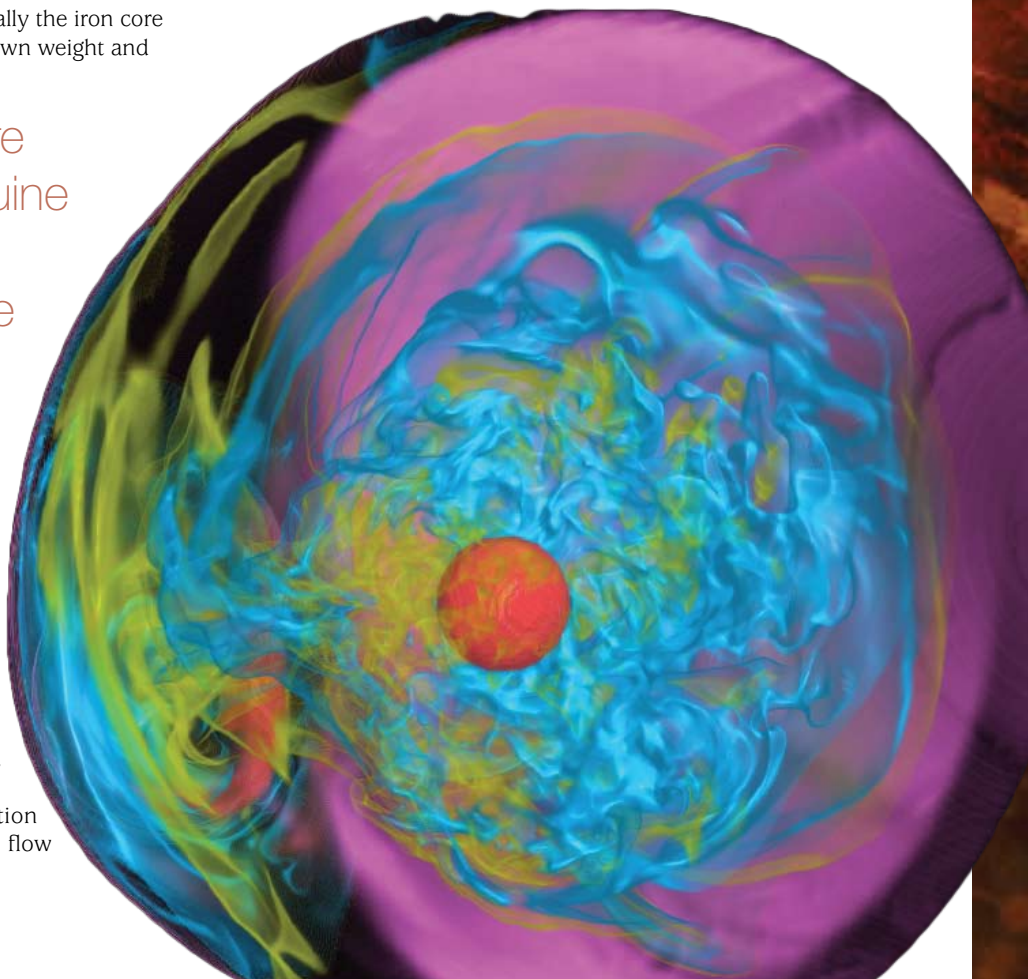
Blondin's discoveries are indispensable to a genuine understanding of the evolution of the universe and the origins of life.

Once the core has collapsed to the highest densities possible, it rebounds and sends out a shock wave, similar to the shock wave of a dropped bowling ball hitting a floor. This shock wave eventually blows most of the star into space, but it first stalls as infalling material from the star rains into the collapsed core. Blondin's first major discovery was that the stalled shock wave becomes unstable. This standing accretion shock instability (SASI) causes the shock wave to eventually become distorted, even cigar-shaped, instead of spherical and create two rotating flows—one in one direction directly below the shock wave and an inner flow that travels in the opposite direction.

Three-dimensional rendering of entropy in a core-collapse supernova showing the development of the supernova shock-wave instability.

Image courtesy K.L. Ma, University of California–Davis.

Blondin's second major discovery was that the inner flow may be responsible for the spin of the leftover pulsar as the inner flow settles onto the core. When a massive star explodes in a core-collapse supernova, the remaining core becomes a neutron star, made up primarily of neutrons; and if that neutron star spins, it becomes a pulsar. Pulsars appear to blink because radiation shoots out of their magnetic poles, which, as with Earth, can be tilted a little from the axis of the spin. As a result, the pulsar behaves like a stellar lighthouse, pointing at an observer once with each rotation.

Blondin's pulsar discovery came at an opportune time because astronomers did not have a workable explanation for how the pulsar gets its spin. To that point astronomers had assumed the spin of the pulsar was a relic of the spin of the original star. If that was true, however, the small pulsar would spin much faster than the original star, in much the same way an ice skater spins faster by pulling in his or her arms. In fact, this explanation accounts for only the fastest observed pulsars. The simulations conducted by Blondin, on the other hand, predict spin periods that are within the observed range, between 15 and 300 milliseconds.

SOPHISTICATED TOOLS ANSWER FUNDAMENTAL QUESTIONS

The simulations performed by Blondin and associated theoretical research efforts have fundamentally contributed to our understanding of these events and therefore of the origins of life in the universe, which is a central goal of the Nuclear Physics program within DOE's Office of Science.

While core-collapse supernovas may seem remote, life would not exist without them. These exploding stars are the principal source of elements in the universe, providing most of those between oxygen (atomic number 8) and iron (atomic number 26) and about half of the elements heavier than iron. Therefore, Blondin's discoveries are indispensable to a genuine understanding of the evolution of the universe and the origins of life.

The discoveries were made entirely through computer simulation; they would have been impossible without the power provided by modern supercomputers able to perform many trillions of calculations each second. The SASI is a fundamentally new piece of supernova theory that has become the framing concept for most current supernova

simulations. It is not only key to an understanding of how the collapsing iron core generates enough energy to blow most of the star into space, but also provides the first explanation for pulsar spin that matches the observations made by astronomers. This validation gives researchers a measure of confidence they are correctly simulating the complex physical mechanisms involved in a core-collapse supernova.

In addition, the counterrotating currents that may give pulsars their spin had escaped researchers until Blondin was able to perform precision simulations in three dimensions with both sufficient resolution and a well-constructed physical model of a supernova. These spins did not exist in two-dimensional simulations, which impose axial symmetry on the exploding star that is not found in nature.

To reach these answers and others, Blondin and his teammates have developed a sophisticated suite of two- and three-dimensional simulation tools that incorporate leading astrophysical theories, high-fidelity numerical solution techniques, parallel algorithms, and efficient and robust modern software implementations. The tools must model an extremely complex array of physics, including shock physics, neutrino transport with complex neutrino–matter interactions, magnetohydrodynamics, nuclear fusion, and general relativistic gravity.

Some of the simulations were conducted on ORNL's Cray X1E Phoenix platform and consumed well over 1 million processor hours. A typical simulation was executed on about a third of the total 1,024 multistreaming processors. The simulations were partitioned into more than 2 billion cells, with a mesh of 1,296 × 1,296 × 1,296.

The real prize in core-collapse supernova research, however, is a complete explanation of how the collapse of a star's core leads to the explosion that ejects most of its layers. So far that explanation has proved elusive; core-collapse supernovas are complex, three-dimensional events that depend on a vast array of physics, and as yet no core-collapse supernova model is sufficiently realistic. Researchers have not been able to determine the explosion mechanism with confidence, and they are limited in their ability to predict the production of new elements and other observable quantities. In the future Mezzacappa's team plans to perform the first three-dimensional core-collapse supernova simulations that include the multitude of physics believed to be important in these stellar explosions, especially neutrino transport across the range of possible neutrino energies. Neutrinos, nearly massless particles that are copiously produced and behave like radiation in collapsed stellar cores, power the supernova shock wave through neutrino heating and play a central role in the explosion. To provide a definitive explanation of core-collapse supernovas, the team will have to successfully simulate these events across the range of exploding stars seen in nature for a variety of initial conditions. The results of this exploration will likely necessitate revisions to the core-collapse-supernova paradigm. The successes delineated here have already led to important revisions of supernova theory.

Contributors/Authors

John Blondin, North Carolina State University; Anthony Mezzacappa, ORNL

Principal Collaborators

In the areas of data management, networking, and visualization, John Blondin received extensive assistance from Scott Atchley, Micah Beck, and Terry Moore of the University of Tennessee–Knoxville; Steven Carter, Nageswara Rao, and Ross Toedte of ORNL; and Kwan-Liu Ma of the University of California–Davis.

Publications

J. Blondin and A. Mezzacappa 2007. "Pulsar spins from an instability in the accretion shock of supernovae," *Nature* **445**: 58–60, January 4.

Designing Proteins at Atomic Scale and Creating Enzymes

ABSTRACT: In achievement of what has been a goal of structural biology for decades, a team at the University of Washington has successfully predicted protein structures at the atomic level, generating computer models of unprecedented accuracy. They then designed new enzymes for two nonbiological reactions, a promising development for making biofuels, for bioremediation, and for industrial chemistry.
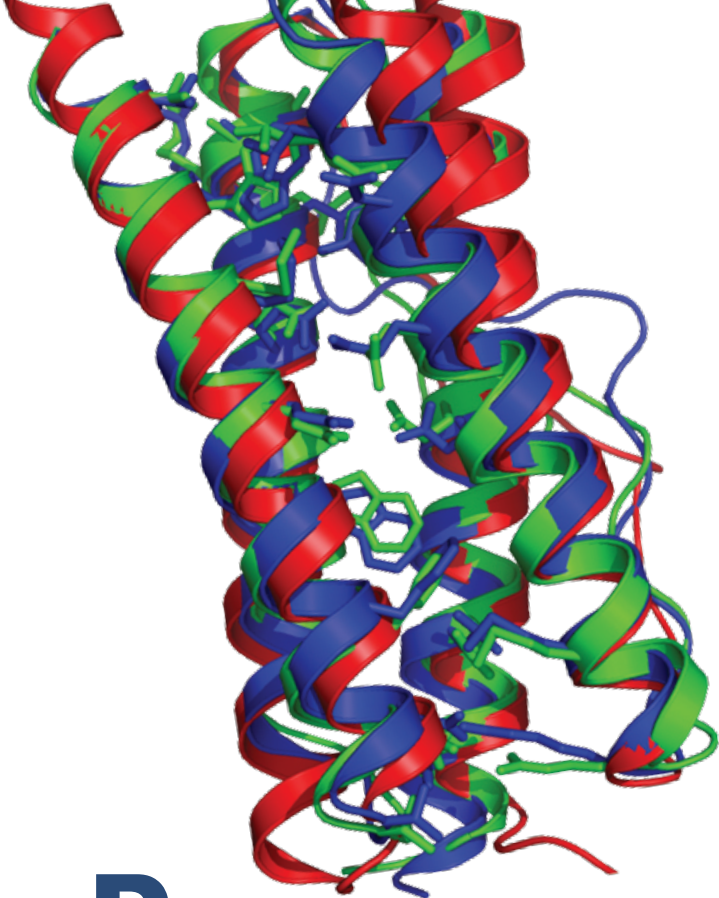
Proteins are tiny, organic machines made of interlocking amino acids that assemble within the cell to carry out the chemical and biological processes that sustain life. Made primarily from carbon, hydrogen, oxygen, and nitrogen, these lean, mean nanomachines are composed of just 20 amino-acid molecules—elemental chemicals that form the basic blueprint of all living things. Yet so powerful are these proteins that they can speed up chemical reactions by as much as a quadrillion to nearly a sextillion (a thousand billion to a thousand trillion) times.

They begin life as a cluster of amino acids, but proteins cannot work until they "fold," changing from two-dimensional threads of amino acids to three-dimensional structures of bewildering complexity. They must fold correctly to avoid serious repercussions—researchers now think Alzheimer's disease may be due to faulty protein folding. Proteins must fold to their favored configuration in nature, called their "ground state," their lowest-energy, three-dimensional atomic structure. One of the central problems of biology is to unravel the rules that govern the folding of a protein's backbone and the packing of its side chains.

Working with amino-acid-sequence information alone, a group led by principal investigator David Baker of the University of Washington–Seattle developed Rosetta, a software package of computational algorithms to design proteins at the atomic scale and predict how they would fold. The software can identify low free-energy sequences for protein backbones, predict complex structure from amino acids, generate a fragments database, design proteins that target DNA sequences, and assemble fragments of RNA.

Using high-performance computing (HPC), they reconfigured and refined their models, selecting the best candidates. To predict the ground state, researchers solved the Schrödinger equation for crucial components of the protein. (This equation gives the quantum mechanical prediction for the ground state of any system.) When tested in the Critical Assessment of Techniques of Protein Structure Prediction, the industry standard, the models were found to be of unprecedented accuracy.

Images are predictions produced by comparative modeling, with crystal structure (blue), the best template in the Protein Data Bank (red), and the best of five submitted Rosetta models (green) all superposed.

Image courtesy B. Qian, S. Raman, and R. Das, University of Washington–Seattle.

When tested in the industry standard, the models were found to be of unprecedented accuracy.

The group then created new enzymes from scratch for chemical reactions not found in nature. Enzymes are proteins that draw molecules into reactions, splitting off atoms from some, fusing others.
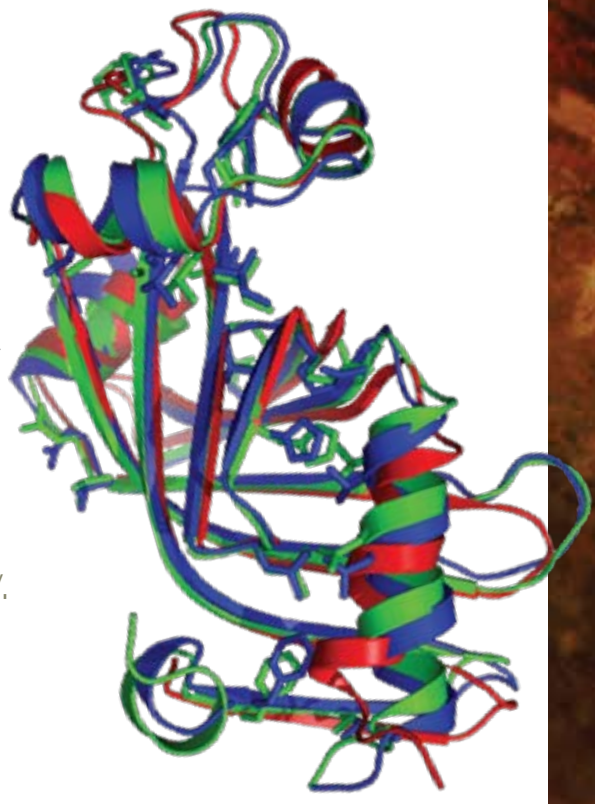
The active site of an enzyme is where a protein docks and catalysis occurs. First, using Rosetta software and HPC, the team generated 180,000 possible enzyme models, eventually whittling these down to 72 whose active sites promised good results. Of these, 32 catalyzed the breaking of a carbon–carbon bond. Some accelerated the reaction by four orders of magnitude.

The group then designed enzymes to remove a proton attached to a carbon atom. After generating, refining, and selecting models with promising active sites, they used directed evolution (selecting for ideal properties using protein engineering) to increase efficiency at the enzyme's active site. Analysis confirmed that the catalysis occurred precisely at the computationally designed sites. The models were validated experimentally; Rosetta generated the amino-acid sequences, and these were grown in the lab. X-ray crystallography confirmed that the cyberproteins and the experimental ones were identical.

For decades molecular biologists have pursued atomic-level protein prediction and enzyme design. Millions have been spent on X-ray crystallography and nuclear magnetic resonance, the current methods. A computational method would both speed the process and reduce costs. Further, many reactions are needed today for which no catalysts exist. Designing new chemical fuels, accelerating complex steps in creating drugs, and destroying toxic compounds are examples. This research raises the possibility that catalysts can one day be designed for any necessary reaction.

The work also addresses societal goals. In DNA sequencing the gap has been growing between science's ability to generate protein-sequence data and its ability to characterize three-dimensional structure. This research begins to bridge that gap and will hasten the interpretation of the genome-sequence information in the Human Genome Project, an international project to determine the sequence of chemical base pairs that make up human DNA and identify the approximately 25,000 genes of our species.

DEVELOPING NEW FORMS OF ENERGY

The research into proteins and their dynamics is also a step in the creation of biofuels, cleaning up waste, and sequestering carbon. Medical applications include the synthesis of new drugs.

Protein-prediction research advances DOE's Biological and Environmental Research mission to transform scientific understanding of energy and matter to advance economic and energy security. The research into new catalysts advances DOE's Basic Energy Sciences program, which is looking for new energy technologies and to mitigate the environmental impact of conventional energy.

In the future Baker's group hopes to transform structural biology from an experimental discipline to a primarily computational science.

The group used an INCITE program allocation of 12 million processor hours to perform computationally intensive calculations on the 5.4 teraflops Blue Gene/L at the ALCF at ANL in Chicago and the IBM Blue Gene Watson in Yorktown Heights, New York.

In the future Baker's group hopes to transform structural biology from an experimental discipline to a primarily computational science. It is refining methods for computing the structures of proteins and protein complexes and developing methods for computing structures from amino-acid sequences. The group is now seeking catalysts to generate new fuel molecules, fix carbon dioxide, and destroy toxic compounds. Because the computational method does not need any naturally occurring enzyme as a starting point, the team can create catalysts for reactions quite different from any that now exist. Its members are also improving Rosetta's algorithms to increase enzyme performance.

Contributors/Authors

David Baker, Eric Althoff, Philip Bradley, Rhiju Das, Lin Jiang, Bin Qian, Srivatsan Raman, Daniela Röthlisberger, Andrew Wollcott, Alexandre Zanghellini, and Howard Hughes, University of Washington–Seattle; Kendall Houk, University of California–Los Angeles; Olga Khersonsky and Dan Tawfik, Weizmann Institute of Science; and Barry Stoddard, Fred Hutchinson Cancer Research Center

Publications

L. Jiang et al. 2008. "De novo computational design of retro-aldol enzymes," *Science* **319**(5868): 1387–1391.

B. Qian et al. 2007. "High-resolution structure prediction and the crystallographic phase problem," *Nature* **450**: 259–264.

D. Röthlisberger et al. 2008. "Kemp elimination catalysts by computational enzyme design," *Nature* **453**: 190–195.

First-Principles Flame Simulation Provides Crucial Information to Guide Design of Fuel-Efficient Clean Engines

ABSTRACT: To create combustion models for designing tomorrow's vehicle engines and power-generation devices, researchers computationally study how turbulence affects chemistry in flames. Insight into how flames stabilize, extinguish, and reignite may spawn new predictive models that guide the design of engines that burn less fuel and generate fewer pollutants and greenhouse gases.

Volume rendering of a lifted autoigniting hydrogen/air jet flame with hydroperoxy radical (ignition marker, red and yellow) and hydroxyl radical (flame marker, blue).

Image courtesy K. L. Ma, University of California–Davis.

A team led by mechanical engineer Jacqueline Chen of Sandia National Laboratories (SNL) used some of the world's fastest supercomputers to simulate key underlying processes in combustion. With mechanical engineer Chun Sang Yoo at SNL and computational scientist Ramanan Sankaran at ORNL, Chen created the first three-dimensional simulation to fully resolve flame and ignition features including chemical composition, temperature profile, and flow characteristics. Their simulations reveal details of these features on all size scales—the biggest, the smallest, and everything in between—of a turbulent hydrogen fuel jet in a hot coflowing airstream as it ignites. The data became a library that engineers are using to develop predictive models to optimize designs for diesel engines and industrial boilers with reduced emissions and increased efficiency.

The team's direct numerical simulations allowed analysis of a system with a moderate Reynolds number, a fluid-dynamics parameter that indicates the range of scales in a system. A system with a lot of turbulence, for example, may have tiny eddy currents with small-scale effects and huge eddies exerting large-scale effects. To simulate phenomena at all length scales required a three-dimensional grid with more than a billion points spaced 15 microns apart. That enabled the world's first fully resolved picture of the physics of so-called lifted flames in direct-injection diesel-engine jet flames. Prior to this work, scientists had simulated only the large eddies in a burning fuel. They had not simulated the full range of turbulence scales down to the smallest eddies, which dissipate heat and interact with reactions in the flame and thus are responsible for flame extinction and delays in self-ignition of a hot mixture.

When a jet of cold fuel and hot air ignites, increasing the speed of the fuel or air streams can lift a flame off a burner. The flame stabilizes, or continues to burn without blowing out, at a region downstream from the burner. It is able to do so because a balance is struck between turbulence, which mixes fuel with air to enable burning, and key ignition reactions, which precede the location of the flame. The lifted flame exists over a range of jet

velocities until further increases in the jet speed result in the flame blowing out.

Lifted flames are important to the functioning of industrial burners for power generation, where they reduce thermal stresses to the nozzle by minimizing contact between flame and nozzle. They are also integral to the workings of direct-injection gasoline engines, compression ignition diesel engines, and gas turbines, in which streams of cold fuel and hot oxidizer are partially premixed prior to combustion. The position downstream of a fuel injector where a diesel fuel jet establishes a flame influences the degree of premixing between fuel and air prior to combustion, which in turn affects combustion and soot-formation processes. Fundamental knowledge of the mechanism by which a lifted flame stabi-

at SNL, in collaboration with ORNL's Sankaran, to optimize S3D on multiple processing cores to model compressible, reacting flows with detailed chemistry. Simulations of the simplest fuel—hydrogen molecules—required 2.5 million computing hours and generated 35 terabytes (trillion bytes) of data about flames similar to those occurring during ignition and stabilization of diesel-engine jets. More recent simulations of a more complex hydrocarbon fuel, ethylene, required 4.5 million hours running on 30,000 processors and generated more than 50 terabytes of data, which is more than five times as much data as contained in the printed contents of the U.S. Library of Congress.

Future simulations will target fuels of increasing complexity and diversity.

Computing at the petascale (a quadrillion calculations per second, available in 2009) and even exascale (a thousand quadrillion calculations per second, possible within a decade) will be required for increasingly complex simulations. Ignition of hydrogen generates eight chemical species and 20 chemical reactions that must be plugged into the model, but more complex fuels such as n-heptane, iso-octane, diesel fuel, or kerosene may generate hundreds of species and thousands of reactions.

If low-temperature compression ignition concepts employing lean, dilute fuel mixtures are widely adopted in next-generation autos, fuel efficiency could increase by as much as 25 to 50 percent.

lizes may aid design and optimization of fuel-efficient, clean-burning combustion devices. For example, development of advanced diesel technology is a leading near-term option for reducing fuel consumption and greenhouse-gas emissions.

HIGH-IMPACT CODE

Transportation accounts for 60 percent of petroleum use by the United States—an amount equivalent to all of the oil imported into our country. Today virtually all transportation energy comes from petroleum. Significant improvements in efficiency are possible through strategic technical investments in both advanced fuels and new low-temperature engine concepts for clean combustion of lean fuels. "If low-temperature compression ignition concepts employing lean, dilute fuel mixtures are widely adopted in next-generation autos, fuel efficiency could increase by as much as 25 to 50 percent," Chen says. That would help meet future low-emission vehicle standards with almost undetectable emissions of nitrogen oxide, a major contributor to smog, she adds.

The simulations were performed on the Cray XT4 Jaguar supercomputer at ORNL's National Center for Computational Sciences (NCCS). The researchers ran the S3D software code developed

The researchers will model fuels with a wide range of ignition characteristics—dimethyl ether, diesel surrogates such as n-heptane, and renewable biofuels such as ethanol—and explore high-pressure, low-temperature, fuel-lean environments representative of advanced compression ignition engines. "Direct numerical simulation is our numerical probe to measure, understand, or see things in great detail at the finest scales where chemical reactions occur," Chen says. "That's particularly important for combustion because reactions occurring at the finest molecular scales impact global properties like burning rates and emissions."

Contributors/Authors

Chun Sang Yoo and Jacqueline Chen, SNL; Ramanan Sankaran, ORNL
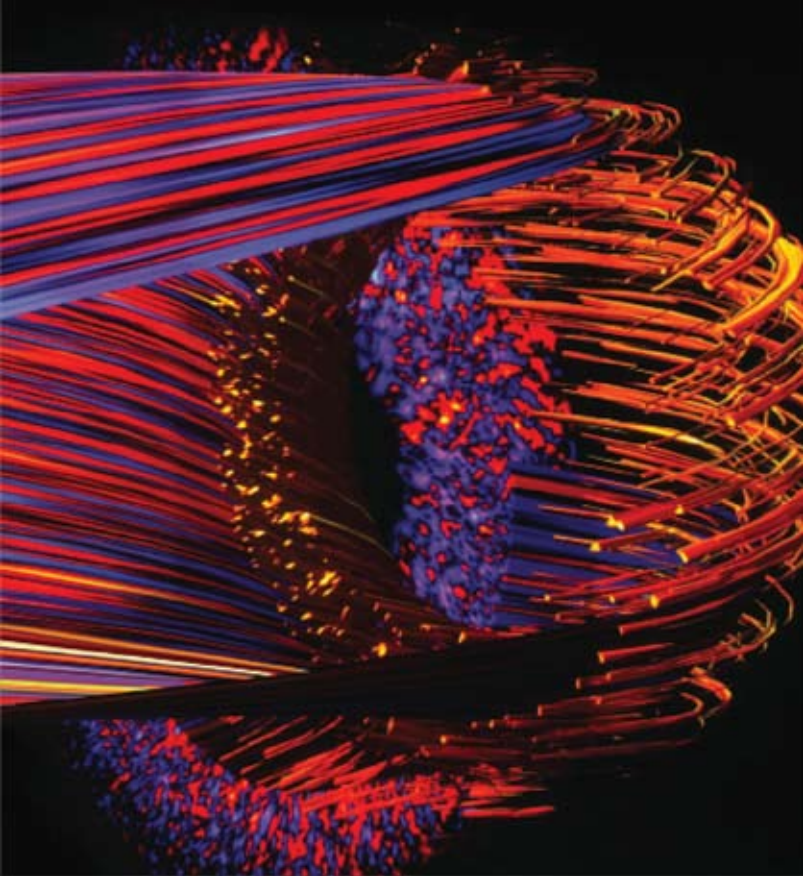
Principal Collaborators

Valerio Pascucci of the SciDAC Visualization and Analytics Center for Enabling Technologies is working with the combustion scientists to develop topology-based feature segmentation and software to track the temporal evolution of intermittent ignition structures that stabilize the lifted flame. Kwan-Liu Ma of the SciDAC Institute for Ultrascale Visualization performed volume rendering of key scalars, which varied over time in the simulation.

Publications

E.S. Richardson, C.S. Yoo, and J.H. Chen 2008. "Analysis of second-order conditional moment closure applied to an autoignitive lifted hydrogen jet flame," *Proceedings of the Combustion Institute* **32** (in press).

Breakthrough Fusion Simulation Sheds Light on Plasma Confinement

ABSTRACT: A team of researchers from the Princeton Plasma Physics Laboratory, University of California–Irvine, University of California–San Diego, Columbia University, and Oak Ridge National Laboratory recently achieved a simulation milestone in modeling instabilities and plasma confinement in a fusion reactor. Fusion power could one day provide the world with a cleaner, more abundant, renewable energy source, greatly reducing harmful emissions and avoiding conventional storage and proliferation issues associated with nuclear fission. This research represents some of the largest fusion simulations in the world to date.

The structure of electrostatic potential in a three-dimensional simulation of plasma microturbulence in a fusion reactor.

Image courtesy S. Klasky, ORNL.

There is no question that new, viable solutions are necessary to meet the world's energy demands and curtail many of the harmful side effects of today's energy-production practices. One of the more promising solutions being explored is fusion, the process by which hydrogen isotopes combine to make helium in a plasma, a very hot (approximately 100 million degrees kelvin) ionized gas. The power source of the sun and other stars, fusion is remarkably clean and plentiful because the primary fuel sources are hydrogen isotopes (deuterium and tritium). While there are storage issues with fusion by-products, the half-lives involved are orders of magnitude smaller than those associated with the uranium and plutonium used in fission power production. Therefore, the radioactive materials will decay much faster, posing less risk over time.

Though plasma makes up 99 percent of the visible universe, understanding its properties has proved a difficult task. To help realize the potential of commercially viable fusion power, an intense research and development program is under way. Simulation, along with theory and experiment, figures to play a major role in the achievement of this goal.

To generate a fusion reaction, the plasma must be magnetically confined, which goes against its natural thermodynamic inclination to escape. These breakthrough simulations demonstrated that the plasma particles will interact with the electrostatic and electromagnetic waves in the confining magnetic field to significantly reduce the efficiency of confinement. For example, it is possible for the hot particles to interact with the electromagnetic waves, cooling the particles as the waves grow in amplitude and creating a less efficient magnetic trap.

Using the state-of-the-art particle-in-cell (PIC) codes GTC and GTS, researchers tracked the movements of billions of particles in magnetically confined fusion plasmas, incorporating the realistic three-dimensional geometry of the containment device. The results were a significantly improved understanding of confinement dynamics and a simulation that is, in terms of validation, closer to what is being observed in fusion experiments. In previous simulations failure to

…the importance of this simulation, which addresses the confinement properties of the hot plasma, is difficult to overstate.

account for the dynamic properties and realistic geometry of the plasma produced results in significant disagreement with those observed in experiments. With this new approach, simulations of plasma fluctuations—believed to be responsible for the degradation of confinement—have reached a new level of realism.

Because the size and cost of a fusion reactor will be largely determined by the balance between the thermal loss rates of the confined plasma and the fusion self-heating rates, the importance of this simulation, which addresses the confinement properties of the hot plasma, is difficult to overstate. The research adds significantly to the knowledge base needed to support major upcoming fusion projects such as ITER, a multinational collaboration to test the feasibility of fusion power production.

SIMULATING THE FUTURE

These simulations are directly related to DOE's mission in a number of ways. Fusion power could revolutionize the way the world meets its energy demands, and advanced simulation is a key component of its development, for both its ability to harvest the physics insights from experiments and its potential to move beyond them in predicting physics performance in future devices. To simulate such complex processes, however, access to the world's best computing systems is essential.

The GTC research accomplishments were carried out under an INCITE program award on the Cray XT4 Jaguar system at the NCCS at ORNL. The project used nearly 40 billion particles on 6,400 cores for 70 hours per simulation—the largest known plasma simulations to date worldwide. The GTS accomplishments were achieved on the National Energy Research Scientific Computing Center (NERSC) Seaborg system, with associated data analysis and visualization carried out with the support of the NCCS at ORNL.

Future challenges in this area will include the thermal confinement of electrons, which is arguably more important for ITER because fusion products first heat the electrons. The associated simulation of electron turbulence is also more demanding because of the shorter time and space scales involved. In addition, the inclusion of electromagnetic forces, which have not yet been incorporated into the global PIC simulations in satisfactory detail, will greatly increase the demand on future computational resources. In the coming petascale and exascale eras,

however, leadership machines such as that at NERSC will be capable of handling the calculations necessary for more computationally intense simulations that will allow researchers to better understand these processes.

Both GTC and GTS will be used to simulate ITER-relevant electron thermal dynamics, with GTS focusing on the validation of results on currently operating toroidal experiments that have ITER's cross-sectional shape, such as General Atomic's DIII-D in San Diego, the Massachusetts Institute of Technology's Alcator C-Mod, and Princeton Plasma Physics Laboratory's (PPPL's) National Spherical Torus. Incorporating the precise geometry of the tokamak containment device, along with comprehensively accounting for the electromagnetic forces and other phenomena in the system, will be essential for more realistic future simulations and subsequent progress in fusion research. The continuing improvement in the capabilities of the GTC and GTS codes will rely in the near future on petascale systems such as the upcoming Cray at ORNL and eventually on exascale systems.

Contributors/Authors

William Tang, Weixing Wang, Wei-li Lee, Stephane Ethier, and Taik-Soo Hahm, PPPL; Zhihong Lin and Igor Holod, University of California–Irvine; Patrick Diamond, UCSD; Mark Adams, Columbia University; and Scott Klasky, ORNL

Publications

Z. Lin et al. 2007. "Wave-particle decorrelation and transport of anisotropic turbulence in collisionless plasmas," *Physical Review Letters* **99**: 265003.

W.X. Wang et al. 2007. "Nonlocal properties of gyrokinetic turbulence and the role of ExB flow shear," *Physics of Plasmas* **14**: 072306.

Closing In on an Explanation for High-Temperature Superconductivity

ABSTRACT: Materials scientists are on the threshold of a firm theory of high-temperature superconductivity as a result of a series of computational breakthroughs between 2005 and 2008, beginning with the computational solution of the leading theoretical model for high-temperature superconductivity and culminating with recent simulation results that can now be validated by experiment.

How some materials superconduct (conduct electricity with no resistance) at temperatures well above absolute zero is one of the key mysteries of modern physics. A series of computational discoveries represents a leap forward in insight—it makes possible the first experiments on actual materials to test one of the main theories of high-temperature superconductivity (HTSC) behavior.

These insights have tantalizing potential; they may eventually enable scientists to design materials that can be used to build electricity superhighways that revolutionize the way we transport and use electrical power. Experimental validation of the results would supply a definitive theory to undergird the design of new materials that superconduct at significantly higher temperatures.
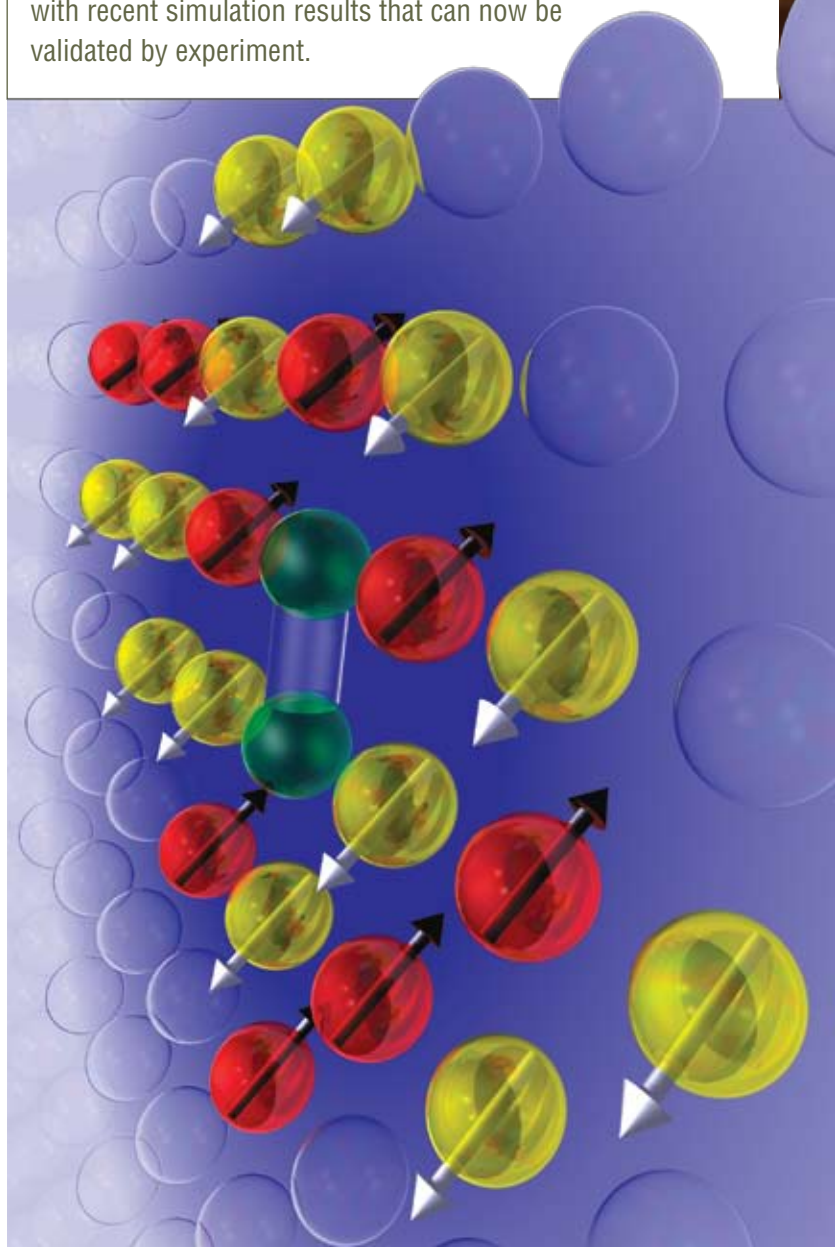
Scientists have long been familiar with materials that superconduct at temperatures near absolute zero (0 K or $-420°F$). Then in 1986, researchers discovered that some materials, notably copper oxide alloys called cuprates, superconduct at more manageable temperatures as high as -150 K. However, scientists remained baffled as to how electrons in those materials bond to form Cooper pairs, the state in which they can stream through the atomic framework (lattice) of the material without resistance. In conventional superconductors, waves of vibrations in the atomic lattice create an attraction between electrons. Cooper pairs (named for the physicist who described them) are electrons that form a weak pair bond due to that attraction, enabling them to avoid scattering off other electrons in the lattice and thus flow through the material without resistance.

The two-dimensional (2D) Hubbard model was the most promising model for describing HTSC in materials, but its complex equations could not be solved for a large enough set of atoms to validate it. In 2005 a team of scientists using the terascale Cray X1E supercomputer at the NCCS solved the 2D Hubbard model and presented evidence that it does predict HTSC behavior. This groundbreaking discovery resolved a decades-long debate in the physics community. Lattice vibrations are absent from the Hubbard model, so it was clear they did not drive the Cooper pairing, but the simulations did not clearly identify a pairing mechanism. In 2006 a second set of simulations focusing on that question presented strong evidence that spin fluctuations (a magnetic effect associated with the rotation of electrons) are responsible for Cooper pairing.

Simulations of embedded atom clusters revealed that spin fluctuations cause electrons to form a superconducting state in the Hubbard model of cuprate superconductors.

Image courtesy T. Maier, T. Schulthess, and L. Meredith, ORNL.

The real test of a theory comes through experimentation on real materials. Follow-on simulations in 2007 and 2008 moved researchers toward that end. The team completed a detailed numerical analysis of the dynamics of the pairing interaction in the 2D Hubbard Model and the related t-J model. The new simulations solidified the case that spin fluctuations are the dominant pairing mechanism in HTSC. (They indicated a smaller role for an alternate theory.) The analysis showed that Cooper pairing is closely predicted by a property of materials called spin susceptibility, which can be measured through neutron-scattering experiments. The team then developed a framework by which the spin susceptibility can be used to calculate the transition temperature (the temperature at which a material becomes superconducting) in a cuprate. Thus for the first time since the discovery of HTSC, a path is open for experimental validation of a theoretical prediction of transition temperature in a cuprate material.

Preparations are being made to measure the spin susceptibility in a cuprate at the Spallation Neutron Source at ORNL. If the experiment confirms theoretical predictions, the spin-fluctuation theory of HTSC will be validated. These discoveries open the way to a much broader understanding of HTSC and represent a significant step toward developing a canonical set of equations elucidating the behavior of materials at all scales. This advance is significant for exploring not only superconductivity, but also other important quantum-physics issues in complex materials.

HIGHER TRANSITION TEMPERATURES = REVOLUTIONARY ENERGY SAVINGS

A major breakthrough in HTSC research will be an enormous boon to an energy-challenged world. Materials that superconduct at higher temperatures would result in incalculable energy savings. This work moves scientists a significant step closer to being able to design superconductors practical for widespread use in real-world applications such as power cables, electric vehicles and trains, and electric machinery. The ultimate goal is materials that superconduct at ambient temperatures and thus require no cooling.

The implications for the power grid alone are enormous. Resistance losses would be cut in transmission lines, generators, transformers, and switches. Power-distribution systems could be significantly smaller, less expensive to maintain and operate, and less vulnerable to breakdown. The cost of electricity and the difficulty of bringing it to remote areas would be reduced worldwide.

Success in developing HTSCs will add dramatically to DOE's energy-efficiency missions and reduce environmental damage from power generation. In addition, the basic scientific knowledge gained will contribute to other broad areas of materials research important to DOE.

The simulations were conducted on the Cray XT3/XT4 and Cray X1E supercomputers at the NCCS. The research used improved numerical techniques and a new code called DCA developed by Mark Jarrell at the University of Cincinnati.

The research team is turning its focus to employing extensions of the Hubbard model together with a newly developed generic implementation of the algorithm in the DCA++ code to describe the behavior of specific materials and shed light on why different materials become superconducting at different temperatures. A petaflops computer scheduled for installation at the NCCS in 2009 will be used. The effects of inhomogeneities on the transition temperatures of alloys are also being analyzed. This work will help researchers develop a calculation framework for screening alloys to search for materials that transition to a superconducting state at higher temperatures.

The team recently used the DCA code to perform the first computations of the transition temperatures of Hubbard models with the model parameters determined entirely computationally. Beyond the input structure of the material, these calculations used no experimental input. This is a necessary step toward eventually predicting or designing new materials computationally. The results, accepted for publication in *Physical Review B*, found the transition temperatures greatly depended on the input parameters. The team is now focusing on improved methods for determining the model parameters. As simulations reveal what parameters push a material to a higher transition temperature, researchers can use that information to design materials with those parameters.

Contributors/Authors

Douglas Scalapino, University of California–Santa Barbara; Thomas Maier, Paul Kent, and Thomas Schulthess, ORNL; Mark Jarrell and Alexandru Macridin, University of Cincinnati; Didier Poilblanc, Laboratoire de Physique Théorique, CNRS and Université de Toulouse

Publications

T. A. Maier, D. Poilblanc, and D. J. Scalapino 2008. "Dynamics of the pairing interaction in the Hubbard and t-J models of high-temperature superconductors," *Physical Review Letters* **100**, 237001.

T. A. Maier et al. 2007. "Systematic analysis of a spin-susceptibility representation of the pairing interaction in the 2D Hubbard model," *Physical Review B* **76**, 144516.

T. A. Maier, M. Jarrell, and D. J. Scalapino 2007. "Spin susceptibility representation of the pairing interaction for the two-dimensional Hubbard model," *Physical Review B* **75**, 134519.

Powerful Mathematical Tools Resolve Complex Simulations

ABSTRACT: PETSc (Portable, Extensible Toolkit for Scientific computation) is a new software package created to simulate and solve complex climate, combustion, fusion-science, and geoscience research problems. Solving massive algebraic systems for more than 3 billion unknowns, PETSc is now used on parallel computers throughout the Department of Energy's science and engineering research complexes.

When researchers want to simulate a system on the computer, they make a model using mathematical language to describe its behavior. Many of these models are composed of what mathematicians call partial differential equations (PDEs). These equations are used to formulate and help solve problems in areas such as climate, combustion, fusion science, and the geosciences.

PETSc is a computing software package, a large and powerful library of mathematical routines for parallel (where many instructions are carried out simultaneously) computing currently being developed by Barry Smith and his collaborators at ANL and the Illinois Institute of Technology. It provides sets of tools for the parallel numerical solution of PDEs that require solving large-scale systems of equations. It has taken modeling of complex multivariable processes to a new level, solving algebraic systems for more than 3 billion unknowns and scaling to more than 27,000 processors on DOE leadership-class computers. The software allows previously unreachable resolution in modeling of important basic energy science problems including climate, subsurface flow, reactor modeling, and fusion science.

Computer modeling of a problem with many variables requires determining relationships among them and establishing how changes in one quantity can affect others. An example is climate modeling, in which wind velocity, atmospheric pressure, and temperature are all important measurements that vary relative to one another—not just when sampled from time to time, but even in one place when researchers draw a trajectory in time and sample multiple points along its length. Similarly, in subsurface fluid-flow dynamics, in which potentially hazardous fluids are sequestered in underground mineral beds, researchers must model predictions of changes in complex minerals in the subsurface, the biological and
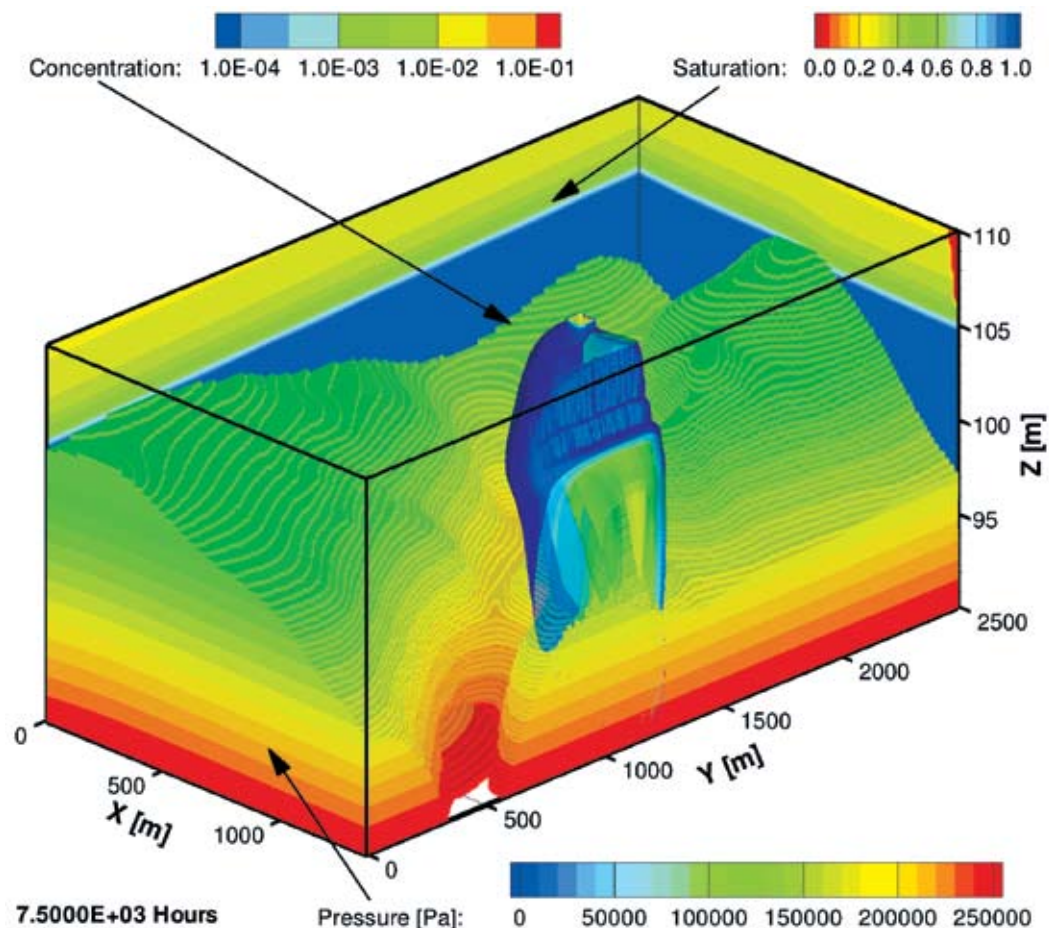
Computation using the PETSc-powered PFLOTRAN subsurface flow simulator and its parallel performance on the DOE leadership-class Cray XT3 system.

Image courtesy G. Hammond, PNNL, and P.C. Lichtner, LANL.

The PETSc researchers are educating the next generation of DOE application scientists.

chemical interactions within multiple fluid phases held within these minerals, reaction scales from the pore scale to the macroscale, and time scales of years to tens of thousands of years. Modeling such systems results in enormous numbers of algebraic equations for all the variables such systems present.

PETSc provides algorithms and software that enable scientists and engineers to solve these types of systems of equations. Exploiting the IBM Blue Gene/P at the ALCF at ANL and the Cray XT4 at the NCCS at ORNL, both DOE leadership computing facilities, the new software allows far more accurate simulations than ever before. The resulting simulations are leading to engineering advances and new scientific understanding in almost all areas of science. A measure of the success is that earlier versions of PETSc were used in three of the Gordon Bell prize-winning applications in 1999, 2003, and 2004 at Supercomputing, the world's premier HPC conference.

The PETSc solvers are now used to model complex phenomena in virtually all areas of DOE-sponsored science and engineering research, including climate science, both fission and fusion, nanosimulations, subsurface flow, oil-reservoir modeling and optimization, combustion, fracture mechanics, and micromagnetics. As the research scientists and engineers continue to increase the quality of their simulations, PETSc is ready to solve their increasingly difficult problems.

ATTACKING PROBLEMS ONCE THOUGHT UNSOLVABLE

The parallel computing infrastructure of PETSc enables DOE scientists and engineers to focus on their primary scientific interests while having access to state-of-the-art solvers, thereby reducing implementation costs and achieving faster and better results. Moreover, the numerical solvers in PETSc are scalable to terascale (and soon to petascale) simulations, thus enabling scientists to use the DOE leadership-class facilities to attack complex problems previously considered intractable. DOE scientific application codes that use PETSc include UNIC

(nuclear-reactor simulation), PFLOTRAN (reservoir modeling), STOMP (subsurface transport), M3D (plasma simulation), and GTC (gyrokinetics). Several of these codes are running on the DOE leadership computing centers under the INCITE program, which is supporting research to enable high-impact advances in science and engineering.

This project was developed using the Oak Ridge leadership computing facilities, as well as on the Linux clusters at ANL. The PETSc developers continue to push the limits of petascale computing on two fronts: increasing the resolution they can provide to their simulation partners (which requires solving larger and larger systems and doing so faster) and solving progressively more difficult problems. Both will require algorithmic and software advances. The team works with application partners to prepare simulations that will make full use of the DOE leadership-class systems with hundreds of thousands of processors.

Contributors/Authors

Barry Smith, Matthew Knepley, Satish Balay, and Lois McInnes, ANL; Hong Zhang, Illinois Institute of Technology

Principal Collaborators

The collaborators are currently working with geoscientists Peter Lichtner at Los Alamos National Laboratory and Glenn Hammond at Pacific Northwest National Laboratory, nuclear engineers Michael Smith and Won Sik Yang at ANL, and fusion physicist Stephen Jardin at PPPL. They also have active collaborations in the geosciences with Brad Aagaard at the California Institute of Technology and Marc Spiegelman at Columbia University. In addition, they continue to pursue software design research with Robert Kirby and Ridgeway Scott at the University of Chicago and Victor Eijkhout at the University of Texas–Austin. On the industrial side, the principals are working with Tech-X Corporation to develop advanced coupled fusion simulations.

The PETSc researchers are also educating the next generation of DOE application scientists. PETSc was used in work reported in two "best student papers" at the annual Supercomputing conference. In addition, more than a dozen DOE Computational Science Graduate Fellows, a fellowship program of DOE's Office of Science and the National Nuclear Security Administration that funds the nation's best science graduate students, have incorporated PETSc into their work.

Publications

A. Cubero and N. Fueyo 2007. "A compact momentum interpolation method for unsteady flows and relaxation," *Numerical Heat Transfer, Part B–Fundamentals* **56**(6): 507–529.

R.F. Katz, M. Spiegelman, and B. Holtzman 2006. "The dynamics of melt and shear localization in partially molten aggregates," *Nature* **442**: 676–679.

V. Kolehmainen et al. 2006. "Bayesian inversion for three-dimensional dental X-ray imaging with limited data," *IEEE Transactions on Medical Imaging* **25**(2): 218–228.
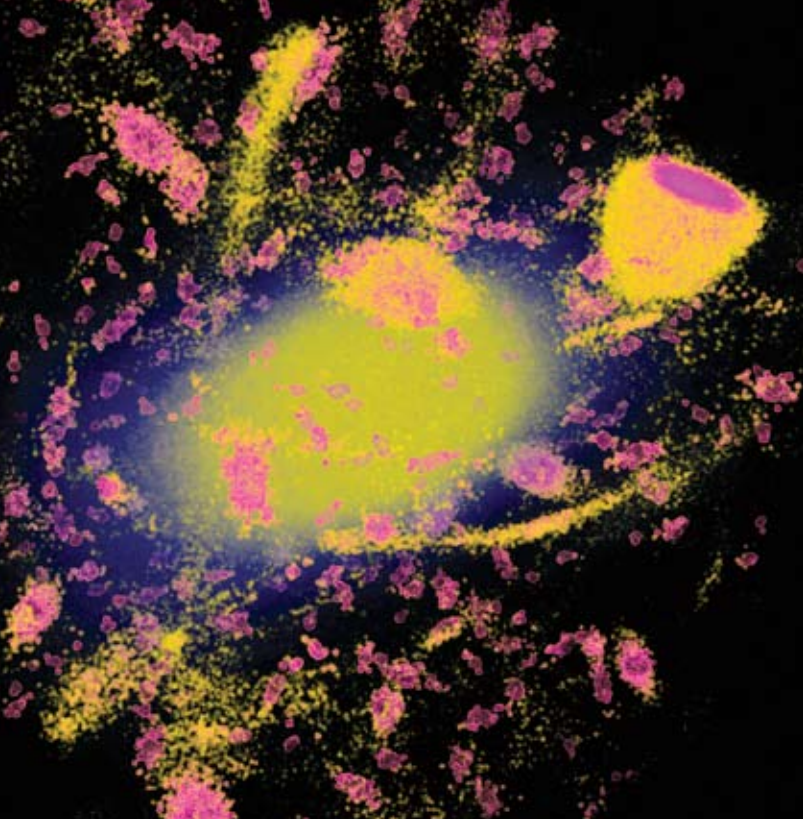
Researchers used ORNL's Jaguar supercomputer to simulate the evolution of the Milky Way's dark matter halo.
Image courtesy P. Madau, UCSC, and S. Ahern, ORNL.

A Billion-Particle Simulation of the Dark Matter Halo of the Milky Way

ABSTRACT: Researchers using Oak Ridge National Laboratory's Jaguar supercomputer performed the largest simulation ever of the dark matter cloud holding our galaxy together. Their accomplishment at the galactic scale may prove groundbreaking in determining the nature of the heretofore undetectable particles that make up most of the universe's mass.

A team led by astrophysicist Piero Madau of the University of California–Santa Cruz (UCSC) has deepened our understanding of dark matter—the invisible material that provides most of the universe's mass—by performing the largest supercomputer simulation ever of dark matter evolving in a galaxy such as the Milky Way. By dividing the galaxy's envelope of dark matter into more than a billion parcels and simulating their evolution over 13 billion years, the team's Via Lactea II simulation showed that small dark matter structures from early in the galaxy's history survived, even as they were incorporated over billions of years into progressively larger structures. In fact, the entire galaxy is decidedly "clumpy," including the inner reaches and our own neighborhood. This discovery was enabled by ORNL's Cray XT4 Jaguar, which at the time of the simulations in November 2007 was capable of nearly 120 trillion calculations per second. Earlier simulations on less powerful systems had shown the dark matter smoothing out, especially in the galaxy's dense inner reaches, because they did not have the resolution to resolve unevenness.

Scientists have realized since the 1930s that visible matter provides far too little gravitational force to keep stars and galaxies tethered to the orbits we observe.

To explain the orbital velocities of stars traveling around galaxy centers and galaxies traveling around one another, they concluded that galaxies are enveloped in halos of invisible matter. In fact, researchers have concluded that what we see makes up less than a fifth of the matter in the universe. The remaining 83 percent is made up of a substance known as cold dark matter—cold because it moves slowly and dark because it interacts extremely weakly with ordinary matter except through the force of its gravitational pull. There is so much dark matter in the universe that its gravitational force controls the lives of stars and galaxies.

> Researchers have concluded that what we see makes up less than a fifth of the matter in the universe. The remaining 83 percent is made up of a substance known as cold dark matter.

These results give astronomers a valuable tool in their search for dark matter. While researchers do not know what the dark matter particle is, they do have candidates, including the neutralino, a theoretical particle predicted by the latest particle theories. Some of these candi-

dates may annihilate when they come into contact with one another, creating gamma rays that can be detected by instruments such as the recently launched Gamma-Ray Large Area Space Telescope (GLAST). Using data from the Via Lactea II simulations, Madau and his colleagues have produced detailed predictions of the gamma radiation that should be detected from dark matter annihilation.

The predictions of Madau's team may also be seen through a phenomenon known as gravitational lensing, in which gravity bends light. In this case the gravity that bends light from distant quasars comes from dark matter found in galaxies along the way. If the dark matter halos of galaxies are as clumpy as this simulation suggests, the light from a distant quasar should be broken up, like a light shining through frosted glass.

GUIDANCE FOR SEEING THE INVISIBLE

DOE is a sponsor of GLAST, which will give physicists a powerful tool for studying the nature of dark matter as well as black holes, pulsars, and other catastrophes that spew gamma rays. Madau and his teammates have provided quantitative predictions for the gamma radiation GLAST will detect from dark matter annihilation. Because this annihilation depends on particles coming into contact, the gamma-ray signals produced in this process should reflect the clumpy structure of the dark matter itself predicted by Madau's team's high-resolution simulation.

The team is also working with members of the GLAST collaboration to develop optimal analysis methods for detecting these signals. If the project is successful, the indirect detection of dark matter by GLAST will be a major breakthrough in our understanding of the universe and a huge step forward in both particle physics and cosmology. Madau's team has shown that GLAST may discover several tens of dark matter clumps orbiting the Milky Way given reasonable values for physics parameters of the dark matter particle.

DOE is also a sponsor of the Cryogenic Dark Matter Search II experiment, in which cryogenic germanium and silicon instruments are attempting to directly detect the flux of dark matter particles hitting Earth as it moves through our galaxy's halo. The underground experiment, run from deep within a former iron mine in Minnesota, may solve the dark matter problem and revolutionize

particle physics and cosmology. Once again, it is the poorly known distribution of dark matter in the solar neighborhood that should determine the strength of the dark matter signal.

Madau's simulation itself begins about 20 million years after the Big Bang and calculates the gravitational interactions of 1.1 billion particles of dark matter over 13.7 billion years to produce a Milky Way–sized halo. Initial conditions for the simulation were generated with a modified, parallel version of the code GRAFIC2 using data from the Wilkinson Microwave Anisotropy Probe, a satellite mission to survey the sky and measure the temperature of radiant heat left over from the Big Bang. The simulation used more than 1 million processor hours on Jaguar, running on as many as 3,000 cores at once. It was performed with an application code known as PKDGRAV2, which ignored visible matter and focused entirely on the gravitational interaction between more than a billion dark matter particles.

Via Lactea II demonstrates the existence of dark matter substructure in the neighborhood of our solar system and provides the first constraints about its properties. Madau and his teammates hope next to calculate this local dark matter structure in more detail, focusing most of the supercomputer's power on resolving our neighborhood of the galaxy.

The team received a grant of 1.5 million processor hours on Jaguar in 2007 through DOE's INCITE program.

In addition, the analysis and interpretation of this simulation were supported by the National Aeronautics and Space Administration.

There is so much dark matter in the universe that its gravitational force controls the lives of stars and galaxies.

Publications

J. Diemand et al. 2008. "Clumps and streams in the local dark matter distribution," *Nature* (in press).

M. Kuhlen, J. Diemand, and P. Madau 2008. "The dark matter annihilation signal from galactic substructure: Predictions for GLAST," *The Astrophysical Journal* (in press).

Exploring the Mysteries of Water

ABSTRACT: Computer simulations resolved some important unknowns about how water behaves at interfaces with other materials at the microscopic scale. This knowledge is essential to understanding biological functions such as protein folding, which is thought to hold the key to nervous-system disorders such as Alzheimer's disease and water transport inside cells.

Despite its familiarity and ubiquity, water's basic chemistry and its behavior in highly confined environments remain a mystery to scientists. A research team led by Giulia Galli of the University of California–Davis used first-principles simulations (calculations containing no experimental data) of water at interfaces to uncover some secrets of how water behaves at the atomic level, yielding results with important implications for fields ranging from biochemistry to materials science.

The simulations explored water at interfaces with graphene, carbon nanotubes, hydrogenated diamond, and biocompatible materials such as silicon carbide, looking specifically at how water behaves in confined spaces only a few nanometers in size. The results clarified the microscopic structure of the water–substrate interfaces and identified the role played by electrons in determining the arrangement of water molecules near surfaces. In addition, they provided predictions of what should be seen experimentally when measuring how water molecules vibrate in contact with a surface.

Being able to predict the behavior of water in confined spaces is especially important because many natural and constructed systems contain water in nanoscale environments.

The confinement of water with both hydrophilic (water-attracting) and hydrophobic (water-repelling) surfaces was modeled with prototypical materials. The research team was surprised by one of the key findings: that properties of confined water at the nanoscale are affected more profoundly by hydrophobic than by hydrophilic interfaces. At these interfaces a thin layer of water molecules forms that has structural properties different from the properties of water at the bulk level. The simulations indicate that layer is about 3 Å thick at hydrophilic interfaces and about 5 Å at hydrophobic surfaces. (An angstrom is 1.0×10^{-10} meters.)

Simulation of sodium chloride dissolved in water inside a carbon nanotube.

Image courtesy of G. Galli, University of California—Davis, and E. Schwegler, LLNL.
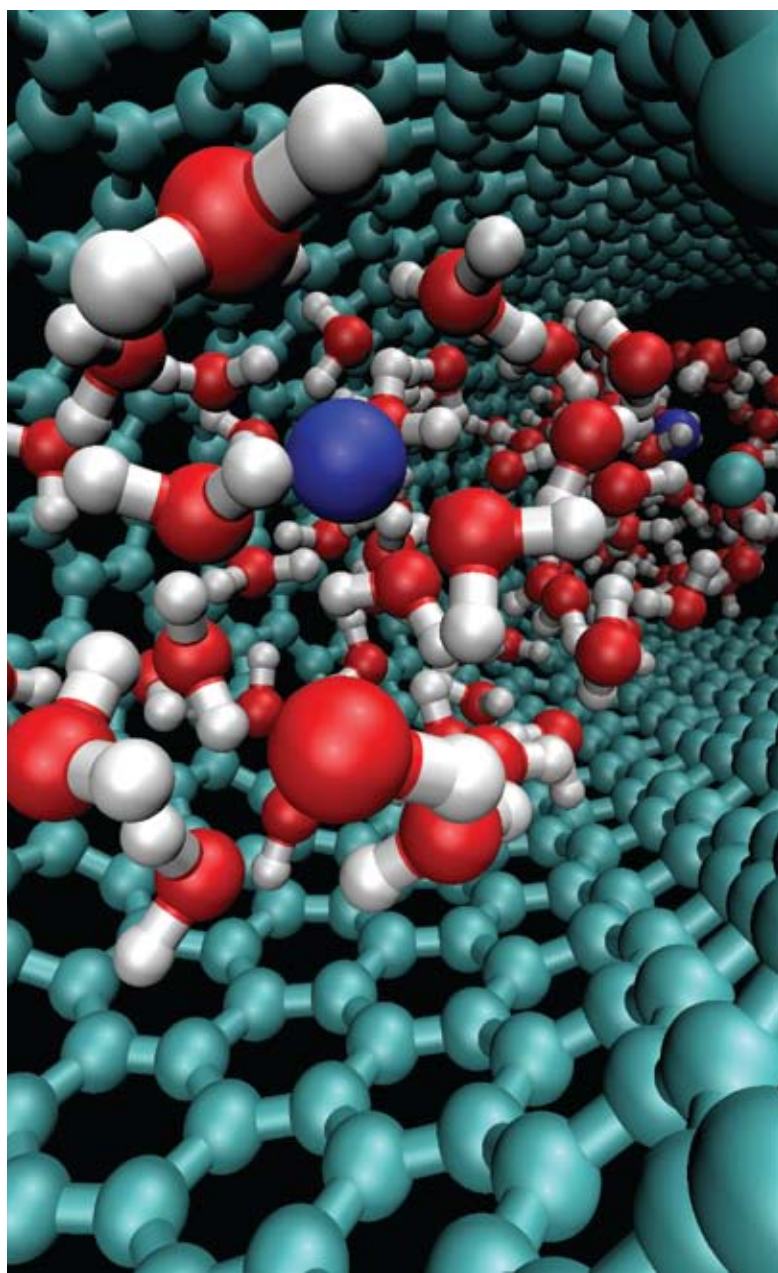
Examining the properties of liquid water near an interface is a very challenging task from both experimental and theoretical standpoints. The challenge is even greater when water is confined in very small spaces, in the range of a few nanometers. Common experimental techniques for probing the properties of liquids have often proved to be exceedingly difficult to interpret, and simple empirical models often do not apply. First-principles methods like those used in this effort (a first principle is a basic assumption that cannot be deduced from any other assumption) are based on the fundamental laws of quantum mechanics and have been used with great success to aid in the interpretation of the experimental measurements that exist for these complex interfacial systems.

Knowing how the structure of water rearranges itself when it comes into contact with other materials is critical to understanding key functions of biological systems and predicting how materials will behave in different types of environments. Being able to predict the behavior of water in confined spaces is especially important because many natural and constructed systems contain water in nanoscale environments similar to the carbon nanotubes examined in the project: proteins, zeolites and clay soils as well as nanofluidic devices for electronics and medical applications.

For example, the ability to predict how water molecules react at hydrophobic interfaces inside proteins could shed light on the dynamics of protein folding, the process by which chains of amino acids form proteins. Research indicates that how tiny hydrophobic areas of amino-acid chains react with water helps determine how they fold. (Correct folding is essential for proteins to function properly, and misfolded proteins are thought to cause a variety of disorders, including Alzheimer's disease.) In cell membranes, tiny pores transport enormous quantities of water in and out; determining the properties of water at the nanoscale will help to unravel how such complex biological systems function. It will also be useful in technological applications such as developing improved membranes for purifying and desalinating water. And interpreting how the surfaces of materials react with water is a key issue in materials design. Every surface in a natural environment, no matter how dry it appears, is covered with a microscopic film of water. If materials scientists can understand how water behaves at those interfaces, they will be better able to predict the mechanisms that cause materials to degrade and how to protect them.

POWERFUL TOOLS FOR POWERFUL RESULTS

Understanding and controlling aqueous environments are important to DOE missions in energy efficiency, environmental protection and mitigation, and biological and chemical sciences. First-principles methods provide powerful tools to complement experiments in those areas. For example, the formation of ceramic materials in aqueous environments is being explored as a more environmentally benign way of developing new materials for energy and environmental applications. This research can help predict how species dissolved in water form clusters, then nanoparticles, and eventually solids, and how the solids formed will later eventually degrade in the environment with exposure to moisture.

The project used quantum calculations—simulations based on quantum mechanics without any experimental data as inputs. The simulations were conducted on the BlueGene/L supercomputers at the ALCF at ANL and IBM's T. J. Watson Research Laboratory and on the Thunder Linux cluster at Lawrence Livermore National Laboratory (LLNL).

QBox, the primary code for the simulations, is a parallel, scalable first-principles molecular dynamics code developed by François Gygi, a member of the research team. QBox is unique in that it was written specifically for massively parallel computers rather than developed over time and tweaked to run on massively parallel systems. It runs efficiently on a very large number of processors, providing excellent scalability. In fact, it won the 2006 Gordon Bell prize for peak performance based on floating point operations per second.

Future simulations for this project will explore prototypical systems and more complex materials to see how they react with water at interfaces. The results will suggest to materials scientists how to design new materials that will react with water in desired ways—for example, better fuel-cell membranes for more efficient fuel cells, thermoelectric materials that can convert waste heat directly to electricity, hydrogen-storage materials to enable hydrogen-burning vehicles, and niobates for applications in solar cells and ferroelectric materials.

Contributors/Authors

Giulia Galli and François Gygi, University of California–Davis; Eric Schwegler, LLNL

Principal Collaborators

Collaborators on this project include Giancarlo Cicero, previously at LLNL and now at the Polytechnical School in Turin, Italy; Jeffrey Grossman at the University of California–Berkeley; and Davide Donadio at the University of California–Davis.

Publications

G. Cicero et al. 2008. "Water confined in carbon nanotubes and within graphene sheets: A first principle study," *Journal of the American Chemical Society* **130**: 1871–1878.

D. Lu, F. Gygi, and G. Galli 2008. "Dielectric properties of ice and liquid water from first principle calculations," *Physical Review Letters* **100**: 147601.

M. Sharma et al. 2008. "Probing properties of water under confinement: Infrared spectra," *Nanoletters* (in press).

Novel Solver Enables Scalable Electromagnetic Simulations

ABSTRACT: A team based at Lawrence Livermore National Laboratory has developed the first provably scalable solver code for Maxwell's equations, a set of partial differential equations that are fundamental to numerous areas of physics and engineering. This new software technology enables researchers to solve larger computational problems with greater accuracy.
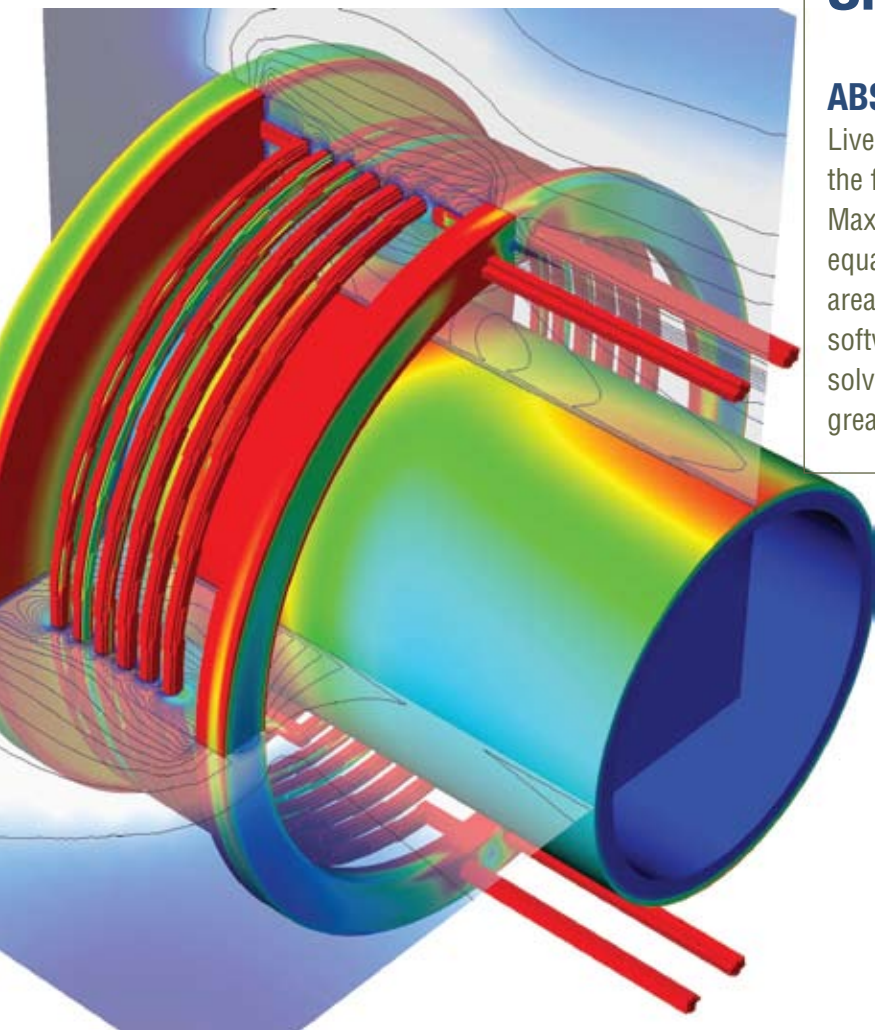
AMS computation of a bifilar helical coil problem used in pulsed-power experiments.

Image courtesy R.N. Rieben, LLNL.

Large-scale electromagnetic simulations are often bottlenecked by slow linear solvers. In simulations conducted at LLNL, a newly developed algorithm, known as the auxiliary-space Maxwell solver (AMS), outperforms earlier solution techniques by as much as 25 times, giving researchers a significant advantage in the future as the questions they seek to answer inevitably become larger and more complex.

The solver's capability is the result of its scalability. Specifically, AMS exhibits "weak" parallel scalability, meaning that the solution time is constant as the problem size and processor workload simultaneously increase. The new algorithm is able to handle complex geometries and problems with large jumps in the material coefficients. In contrast some of the old solvers take more time and produce less accurate results when faced with systems that consist of materials of widely different electromagnetic properties, which are common in engineering.

AMS works by reducing the original problem to a series of equations that can be individually handled using classical techniques. A major advantage of this approach is that its performance is backed by a solid theoretical framework. Thus, AMS is a perfect example of how fundamental mathematical research can lead to important software advances in HPC. In this effort Panayot Vassilevski and Tzanio Kolev of LLNL collaborated with Jinchao Xu of Penn State University and Ralf Hiptmair of ETH, Zurich.

Electromagnetic simulations have a wide range of physical and engineering applications such as in the development of semiconductor chips, stealth aircraft, and electrical generators. As the ability of supercomputers to tackle ever bigger problems grows, researchers need to be able to efficiently take advantage of this newfound computing power.

The AMS solver does just that, solving ever more complex simulations with greater accuracy. It gives researchers an edge by taming solution time, which enables a greater number of simulations,

and allowing mesh refinement, which reveals a more detailed, accurate picture of the system being studied, just as a digital camera with more megapixels captures a more detailed image.

AMS is the first solver for Maxwell's equations that demonstrates theoretically supported weak parallel scalability. It has now been incorporated into several LLNL physics codes previously limited by their Maxwell solvers, most noticeably in terms of resolution. AMS has been tested in a number of applications and demonstrated a significant (4 to 25 times) improvement in the solution time when run on a large number of processors. For example, it takes less than 6 minutes to solve a complicated multimaterial, unstructured electromagnetic diffusion problem with 1.2 billion unknowns on approximately 2,000 processors. Maxwell problems of this size previously were not tractable in LLNL's state-of-the-art physics codes.

LARGE-SCALE PROBLEM "SOLVER"

AMS, with its parallel scalability, is central to the DOE mission because it allows researchers to conduct electromagnetic simulations on a much larger scale than ever before, more accurately representing the physics involved. The new solver was specifically designed for large-scale problems. Because electromagnetic simulations are crucial to a wide variety of systems and phenomena, the application of this solver is relevant to numerous DOE initiatives.

The team used two production Linux clusters at LLNL to create and test the solver: Thunder, with 4,096 Itanium 2 processors, and Zeus, containing 2,304 AMD processors. The software was implemented in and relied heavily on the hypre library developed in the Center for Advanced Scientific Computing at LLNL.

AMS is only one part of the research in this area being conducted at LLNL. The hypre team performs other long-term research on scalable solvers for a variety of applications supported by OASCR and is primarily concerned with algebraic multigrid methods, scalable solvers that rely on as little information as possible from the user. Ultimately the researchers aim to develop new multigrid methods to address new classes of physics and engineering applications, such as contact mechanics and gas dynamics.

This research was funded by OASCR. DOE's INCITE program also provided funding, while the National Nuclear Security Administration Advanced Simulation and Computing program provides funding support for the hypre library, which also receives support from the SciDAC program.

AMS is a perfect example of how fundamental mathematical research can lead to important software advances in high-performance computing.

Contributors/Authors

Tzanio Kolev and Panayot Vassilevski, LLNL; Jinchao Xu, Penn State University; Ralph Hiptmair, ETH, Zurich; Joseph Pasciak, Texas A&M University

Publications

R. Hiptmair and J. Xu 2007. "Nodal auxiliary space preconditioning in H(curl) and H(div) spaces," *SIAM Journal on Numerical Analysis* **45**: 2483–2509.

T.V. Kolev, J.E. Pasciak, and P.S. Vassilevski 2008. "H(curl) auxiliary mesh preconditioning," *Numerical Linear Algebra with Applications* **15**: 455–471.

T.V. Kolev and P.S. Vassilevski 2008. "Parallel auxiliary space AMG for H(curl) problems," *Journal of Computational Mathematics* (in press).

List of Acronyms

USED IN THIS REPORT

ALCF   Argonne Leadership Computing Facility

ANL   Argonne National Laboratory

DOE   Department of Energy

HPC   high-performance computing

INCITE Innovative and Novel Computational Impact on Theory and Experiment

LLNL   Lawrence Livermore National Laboratory

NCCS   National Center for Computational Sciences

NERSC National Energy Research Supercomputing Center

OASCR Office of Advanced Scientific Computing Research

ORNL   Oak Ridge National Laboratory

PETSc   Portable, Extensible Toolkit for Scientific computation

PPPL   Princeton Plasma Physics Laboratory

SciDAC Scientific Discovery through Advanced Computing

UCSD   University of California–San Diego

The Panel

Pete Beckman, acting director, Argonne Leadership Computing Facility, Argonne National Laboratory (computer science)

Jacqueline Chen, Sandia National Laboratories (combustion-energy science)

Giulia Galli, University of California–Davis (chemistry and materials science)

James Hack, director, National Center for Computational Sciences, Oak Ridge National Laboratory (climate science)

David Keyes, Columbia University (applied mathematics)

Douglas Kothe, science director, National Center for Computational Sciences, Oak Ridge National Laboratory (computational fluid dynamics and nuclear engineering)

Paul Messina, interim director of science, Argonne Leadership Computing Facility, Argonne National Laboratory (computer science)

Anthony Mezzacappa, Oak Ridge National Laboratory (astrophysics and panel chair)

Claudio Rebbi, Boston University (high-energy and nuclear physics)

William Tang, Princeton Plasma Physics Laboratory (fusion-energy science)

Nagiza Samatova, Oak Ridge National Laboratory (biology)

Katherine Yelick, director, National Energy Research Supercomputing Center, Lawrence Berkeley National Laboratory (computer science)

Production Team

Agatha Bardoel, science writer

Deborah Counce, science writer

LeJean Hardin, page layout

Priscilla Henson, editor

Jayson Hines, creative director

Scott Jones, science writer

Dawn Levy, science writer

Julia White, production team leader

Leo Williams, science writer