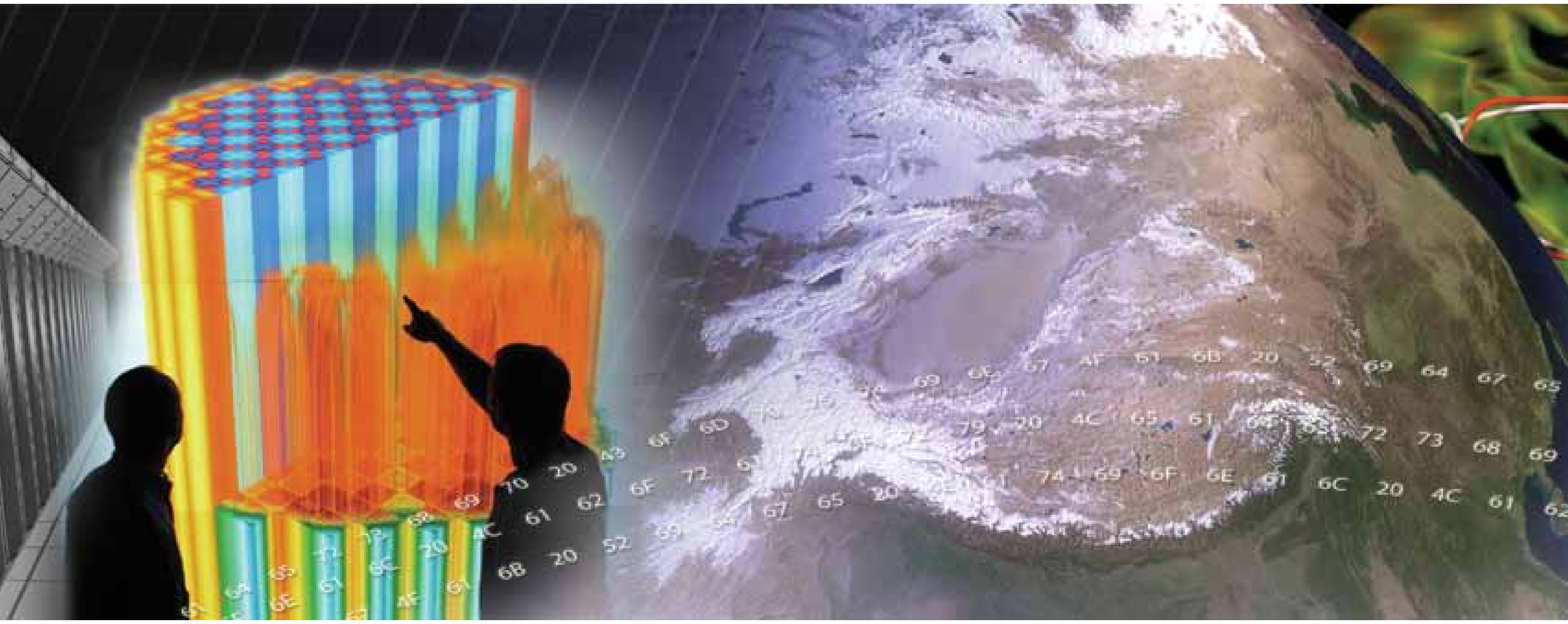
20 Years of Excellence in Computational Science

OLCF

Oak Ridge Leadership Computing Facility

ANNUAL REPORT 2011–2012



Oak Ridge National Laboratory

Titan

OLCF Science

20th Year of Operation

Inside the OLCF

Appendices

Creative Director: Jayson Hines
Writers: Gregory Scott Jones, Dawn Levy, Leo Williams, Eric Gedenk, Sandra Allen McLean
Graphic Design and Layout: Jason B. Smith
Graphic Support: Andy Sproles
Photography: Jason Richards
Additional images: iStockphoto

James J. Hack, Director
National Center for Computational Sciences
Arthur S. Bland, Project Director
Oak Ridge Leadership Computing Facility

Oak Ridge Leadership Computing Facility
P.O. Box 2008,
Oak Ridge, TN 37831-6008

TEL: 865-241-6536
FAX: 865-241-2850
EMAIL: help@nccs.gov
URL: www.olcf.ornl.gov

The research and activities described in this report were performed using resources of the Oak Ridge Leadership Computing Facility at Oak Ridge National Laboratory, which is supported by the Office of Science of the U.S. Department of Energy under Contract No. DE-AC0500OR22725.

We Honor Jaguar Even as We Anticipate Titan

We are excited about the impending arrival of Titan, the serious jump in performance it will give us, and the new insights it will enable for our users. Before we move ahead too far, though, this is also an appropriate time to look back at the accomplishments of Titan's predecessor, Jaguar.

Not only was 2011 Jaguar's last full year in operation; it was also the system's best year, period.

Jaguar leaves the stage having been America's most powerful supercomputer, but more important, it has also been the country's most productive. Jaguar was rock solid in 2011—very stable and likely the easiest-to-use system Oak Ridge National Laboratory (ORNL) has ever stood up.

We are understandably proud of it. Jaguar was the first supercomputer to bring working scientific applications to the petascale, allowing them to complete more than 1,000 trillion calculations each second. As a result simulations became far more complex and realistic than ever before. These applications took us one step closer to realizing our dream of true digital experiments.

Jaguar has hosted petascale applications for nearly 4 years, beginning in 2008. These applications have advanced our understanding in important areas of physics, chemistry, and engineering. Even now they are unlocking the secrets of high-temperature superconductors and the magnetic properties of promising materials, shining light on the energy contained in clusters of water molecules, and exploring ways to advance computer chip design once transistors can get no smaller.

In 2011 Jaguar outdid even itself, delivering nearly 1.5 billion processor hours to the world's most sophisticated computational scientists. If you attempted the same workload on your dual-core laptop, it would take you more than 85,000 years.

Because it is a unique resource, we have reserved Jaguar for the world's most demanding projects—projects that would not even be attempted on other systems. Six of these projects used more than 50 million processor hours each in 2011 (a 2,800-plus-year job on your laptop). These projects span the range of computational science:

- One is dedicated to understanding and eventually designing the catalytic chemical processes that will revolutionize batteries in particular, energy storage in general, and a wide range of other industrial processes.

Jeff Nichols, associate laboratory director for computing and computational sciences

- Another focuses on the conversion of magnetic energy to particle energy in plasmas such as solar flares and fusion experiments.
- The third is deconstructing the turbulent combustion processes that power internal combustion engines.
- The fourth is improving the efficiency of fusion plasma confinement by simulating the edge of a fusion reactor.
- The fifth is working toward a comprehensive description of both stable and unstable atomic nuclei, including nuclei that are rare and short-lived.
- The sixth is improving our understanding of the strong nuclear force.

Making things better right here, right now

We are also pleased with the role our staff and facilities have played in taking the scientific knowledge developed here and elsewhere and applying it to modern challenges. ORNL has a well-deserved and long-standing reputation for translating knowledge into technology, and we are proud to see the National Center for Computational Sciences (NCCS) leading this tradition into the 21st century.

Take the Consortium for Advanced Simulation of Light Water Reactors (CASL). Headquartered at ORNL, CASL is developing a virtual reactor that will help scientists and engineers improve the safety, efficiency, and longevity of America's nuclear power generators. The Department of Energy chose CASL as its first energy innovation hub. We believe the reputations of both the lab and the NCCS were important factors in that win.

> We believe most of the world's powerful systems will be hybrid systems going forward for the foreseeable future. By helping researchers identify opportunities to exploit GPUs on Titan, we are also helping them exploit acclerators on other architectures.

CASL is applying the immense power of computer simulation to three primary efforts. The first will reduce capital and operating expenses by extending the lifetimes and power outputs of both existing and next-generation reactors. The second will enable more complete burnup of reactor fuel, thereby reducing the amount that becomes dangerous and long-lived waste. And the third will ensure greater safety by improving our ability to predict the performance of reactor components.

There are other examples. For instance, the engineering services firm BMI Corporation is using Jaguar to improve the fuel efficiency of long-haul trucks. The trucking industry is absolutely essential to the American economy, but traditional tractor trailers get lamentably low fuel economy. Using the insights of computational fluid dynamics and the power of Jaguar, BMI and its partners were able to design a relatively simple set of attachments that go under the trailer to improve air flow and, thus, fuel economy.

2012: The year of Titan

As we look forward, though, Jaguar's time is coming to a close. By the end of 2012 the system will have been converted to Titan, a supercomputer that will use the power of graphics processing units (GPUs) to give us many times more computing muscle within the same space and power requirements now used for Jaguar.

As of this writing we have completed the first phase of Jaguar's transformation into Titan, increasing the number of AMD Opteron processors by a third and installing nearly 1,000 NVIDIA GPUs. By the end of 2012 many more GPUs will be installed, bringing the system to 20 petaflops or more.

We refer to the use of central processing units (CPUs) and GPUs together as hybrid computing. While the CPUs in Titan will have 16 processing cores, allowing them to do 16 things at a time, the GPUs will be able to do hundreds. In effect they will act as accelerators, giving researchers access to an added layer of parallelism and the opportunity to shorten their time to solution.

It won't be easy. Our users will have to put some serious time and effort into their codes to make the most of the GPUs. Nevertheless, we are working to make their efforts less painful and more worthwhile. In 2009 we identified a diverse group of leading applications to act as test cases, and we're using them to identify best practices for optimizing codes for the new system.

Each of these applications has a group of experts working with it. Besides members of the application team itself, this group includes representatives from Cray and NVIDIA and at least one representative from our Scientific Computing Group. We are confident that by identifying the most effective approach for these applications, these teams will also have identified an effective approach for scientific applications in general.

To ensure that their efforts are worth the time and energy that go into them, we are committed to making the best practices we identify for Titan applicable to other systems as well. This might not be as difficult as it seems. We believe most of the world's most powerful systems will be hybrid systems going forward for the foreseeable future. By helping researchers identify opportunities to exploit GPUs on Titan, we are also helping them exploit accelerators on other architectures.

A code that is optimized for Titan will be well positioned for the future.

Titan

With its revolutionary hybrid architecture, Titan promises to accelerate even the most advanced applications and pave the way for the world on its quest to reach the exascale.

Q&A: James Hack, Director, NCCS

Oak Ridge National Laboratory (ORNL) will be taking a new approach with its next supercomputer, Titan. Like Jaguar, Titan will rely on hundreds of thousands of central processing units (CPUs) working in parallel. But unlike Jaguar, Titan will also contain graphics processing units (GPUs), each of which is able to handle hundreds of tasks simultaneously. The GPUs will function as accelerators, giving Titan an additional layer of parallelism. Jaguar is already one of the United States' most powerful supercomputers, but GPUs will give Titan up to another order of magnitude boost in speed.

National Center for Computational Sciences (NCCS) Director James Hack discusses the reasoning behind this approach to scientific computing, as well as the challenges associated with it and ORNL's response to those challenges.

When it goes online Titan will likely be the most powerful supercomputer ever to use accelerators such as GPUs. Why go this route?

We're at a point where existing technology has reached its scaling limit with regard to power. Jaguar is a 200-cabinet computer. It takes up a very large footprint. It also takes 7 megawatts of power to operate. We could build a bigger version of Jaguar, but it would only be an incremental increase in performance. So, by adopting a heterogeneous architecture which employs disruptive technologies like accelerators, we have a chance in the same footprint, with the same power envelope, to actually enhance the potential performance of the machine by up to a factor of 10. It's a step toward exascale, which is the next thousand-fold increase in computing power. This is really the opportunity for us to transition away from programming paradigms that we've used for the last decade and start exploring a new path forward.

This is a big change for projects that have been running on Jaguar. How can they prepare for Titan, and how will their efforts pay off beyond Titan?

We recognize that the programming paradigm will need to change, which means that the way in which the scientist exploits the hardware will change. This may require new software tools. Our project is currently working with vendors to develop those tools. This may also require new ways of implementing your problem on the architecture.

Most of our users need their codes to be portable across different architectures. So we're working with five of our largest projects running on Jaguar to develop best practices for migrating from the existing architecture to the new architecture.


James Hack, NCCS Director

Our strategy is to provide the framework to exploit the additional performance of Titan without compromising their ability to run on other architectures in other environments. It's really an exercise in generalizing the way in which the user implements a problem on the hardware.

Titan will have the same number of nodes as Jaguar, the same number of degrees of freedom at the interconnect level. The additional performance will come from exploiting parallelism within the node. That's where the real challenge is. We plan to take it on by minimizing the changes users will need to incorporate in their codes through directives, through working with compiler manufacturers, and help them look at fundamental algorithms to see if we can't uncover lightweight parallelism that can be exploited within the node.

How much do we really need all this computing power?

Computing power limits what the computational scientist can do, even at the petascale. We're hoping that for many applications—energy, materials, climate and others—the simulation fidelity will become much more realistic and will advance our fundamental understanding of important scientific problems. The high-performance computational environments that we have been working with for the last decade or so have reached an inflection point. We need to make a jump to a different architecture, a different form of doing computing, and this is really a first step in that direction.

Computational science plays a very important role in many things that we see in our daily life. There's the design of aircraft, for instance, or the fundamental elements of industrial design. The virtual environment provides a much quicker way for us

to improve our understanding of older problems and break ground in our understanding of new phenomena with ramifications for how we live our lives. Machines like Titan and activities like the Titan project will be the vehicles that allow us to explore these fundamental things in computational science. They will provide a framework for better product design, new and innovative technologies, and new materials. And they will enable new insights into how very complex, nonlinear systems work; that, again, has implications for a lot of the technologies that we take for granted.

Besides serving as director of the NCCS you are also director of the Oak Ridge Climate Change Science Institute and actively involved in computational climate research. How will Titan affect your own research activities?

The focus of our work is to push the boundaries of climate simulation to operate at resolutions that are typically reserved for numerical weather prediction. Eventually we also want to increase process complexity.

By increasing resolution we also increase the fidelity of the simulations and capture phenomenology that we miss at lower resolutions. As we conduct these types of simulations, we want to quantify what that higher resolution brings in the way of better projections—of things like extreme events and of trends in climate both natural and induced by humans.

Mature applications like climate will have a very heavy lift in order to exploit Titan. It's a mature science, which means that you're migrating a large body of scientific code from one architecture to another. Some applications are dominated by a kernel, or a very small focus segment of computation that can be optimized for the new architecture. Climate codes are what you would call multiphysics, multiscale codes which in general are very hard to migrate because they don't have any hot spots. They have a very broad range of algorithms embodied in them, and all these things need to be migrated.

Our NCCS application team is working on one climate code, in particular the atmospheric part of it, in order to migrate that simulation component to the new architecture and in the process hopefully establish best practices. Our goal is to apply these migration techniques to other parts of this application, and eventually move the whole application over in a way that takes full advantage of the new architecture.

GPU computing has never before been attempted at the scale of Titan. Does this fact bring particular challenges?

On some level we're feeling our way through the fog. We're trying to figure out where to go next in a way that satisfies the constraints we have on power and space and cooling, and also the scalability of many of our applications. They're not going to continue to scale indefinitely via message-passing paradigms, so this gives us an opportunity to explore and exploit parallelism within the node—really look at issues about data movement and how we can keep data more local, exploit data locality for performance.

We're discovering that the work we're doing to exploit a heterogeneous architecture is actually benefiting applications on the previous generation of architectures. So when application scientists—the computational scientists—rethink their problems to make the most of future architectures, they're also improving their performance on current architectures.—*Q&A with Leo Williams*

Science On Day 1

Titan users will have the skills and resources they need to hit the ground running

ORNL has stood up many world-leading scientific supercomputers over the years.

And yet, installing the system is only half the job. The other half, just as important, is ensuring that researchers are ready for the new hardware.

Accelerators such as those that will be used in Titan present a profound opportunity, but they also bring a substantial challenge. What then, should researchers do to take full advantage of the GPU accelerators that will now be available?

To answer that question, Oak Ridge Leadership Computing Facility (OLCF) staff scientists have been working for more than a year with vendors and some of the facility's most advanced application teams to make the best use of Titan. Their efforts ensure that the system will be producing valuable results right out of the gate.

The challenge of GPUs

In one sense the effort to get users ready for Titan is similar to earlier efforts. Being at the forefront of scientific supercomputing, the OLCF regularly upgrades its capabilities to ensure that researchers have access to the most powerful resources possible.

OLCF Project Director Buddy Bland shows US Energy Secretary Steven Chu models of the OLCF's new NVIDIA GPUs. The GPUs excel at quickly calculating repetitive tasks, freeing up traditional CPUs to do more complex calculations. By adding GPUs to the current Jaguar supercomputer, the OLCF expects to see a ten-fold increase in computation speed.

But Titan is different. Jaguar has gone through six upgrades over the past 7 years, but the challenge there was primarily one of scale, since Jaguar got its power by combining more and more CPUs. Titan, on the other hand, will combine two very different types of processors—CPUs and GPUs—and ask users to coordinate their applications accordingly.

You can't give a GPU the same kind of task you give a CPU and expect positive results. GPUs are lightning fast but incapable of performing the most sophisticated calculations. CPUs are smarter but slower. Allocate your application wisely and you can revolutionize your field of study. Allocate it unwisely and you can have a mess.

A GPU is awesome at churning through simple calculations. Tell it to add, multiply, or do any number of straightforward operations on a lineup of numbers, and it will do you proud. But don't ask it to decide for itself what it should do with these numbers.

CPUs, on the other hand, can make decisions and handle conditional statements (e.g., "If X is true, do Y; otherwise, do Z.") This is not a GPU's strong suit.

"GPUs are fast as lightning," noted ORNL's Bronson Messer, "but you can't ask them to make a decision. All they can do is take a big chunk of data and do the same thing on it. You can have a conditional, but it really slows you down. The easiest way to kill your performance on a GPU is to give it a conditional."

This diversity of strengths between CPUs and GPUs, then, asks researchers to rethink their problems in a way they might not have in recent years.

Center for Accelerated Application Readiness

In response, the OLCF created the Center for Accelerated Application Readiness, or CAAR, a collaboration among application developers, Titan manufacturer Cray, GPU manufacturer NVIDIA, and the OLCF's scientific computing experts.

> OLCF staff scientists have been working for more than a year with vendors and some of the facility's most advanced application teams to make the best use of Titan. Their efforts ensure that the system will be producing valuable results right out of the gate.

CAAR has been working for nearly 2 years to establish best practices that will give users guidance on accelerating their codes. The center is divided into five teams working with five of the OLCF's most advanced and representative applications. Messer, who serves as the center's leader, said the codes were chosen for three primary reasons. First, they use a wide range of computational methods, meaning the lessons of CAAR will apply to a wide range of scientific applications. Second, they are highly advanced, having in recent years used nearly a third of the processing hours allocated on Jaguar through the Innovative and Novel Computational Impact on Theory and Experiment (INCITE) program. And third, they are important to DOE's mission of ensuring American energy independence.

Messer said the process of optimizing an application for Titan begins with an audit. How is the application dividing its work, and where is it spending its time? The answer to these questions will tell the team where it needs to concentrate.

Next, the team looks for effective ways to send work to the GPUs and, indeed, all elements of the system. Titan will have 18,688 nodes. Each node will have one 16-core CPU and one GPU. Each CPU core will have 2 gigabytes of memory, with 6 gigabytes attached to the GPU.

"Applications make use of supercomputers by finding effective ways to spread out the work over the available resources," Messer explained. "On Jaguar, for instance, the first level was to spread it across the nodes. Each node had two six-core AMD processors, so the work at hand would be spread across those.

"With Titan, each node will not only have a 16-core AMD processor, but also an NVIDIA GPU capable of tackling hundreds of tasks simultaneously. Therefore, an application using Titan to the utmost must also find a way to keep the GPU busy, remembering all the while that the GPU is fast, but less flexible than the CPU."

The CAAR applications

The five applications chosen for CAAR are S3D, a leading combustion code; LAMMPS, a molecular dymanics code; WL-LSMS, an application that simulates magnetic systems; Denovo, a code that simulates the inside of a nuclear reactor; and CAM-SE, a climate simulation.

Combustion with S3D. S3D performs direct numerical simulation of fuels burning under highly turbulent conditions.

Three-quarters of the fossil fuel used in the United States goes to cars and trucks, which produce one-quarter of the country's greenhouse gases. As a result, the advances in fuel efficiency and alternative-fuel engines promoted by S3D and its developers benefit both national security and the environment.

This application has made great strides on Jaguar, simulating cells only 5 millionths of a meter wide and modeling syngas, a mix of carbon monoxide and hydrogen that burns more cleanly than coal.

With the 20 petaflops offered by Titan, however, the team will be able to move on to far more complex fuels. Instead of being limited to simple foundational fuels such as hydrogen, syngas, and hydrocarbon fuels with up to four carbons, S3D will be able to explore larger-molecule hydrocarbon fuels such as isooctane (a surrogate for gasoline), commercially important oxygenated alcohols such as ethanol and butanol, and biofuel surrogates (blend of methyl butanoate, methyl decanoate, and n-heptane.)

"These can be used to design future fuel-efficient, clean internal combustion engines, burning at lower temperature and higher pressure than today's engines," noted team member Jacqueline Chen of Sandia National Laboratories.

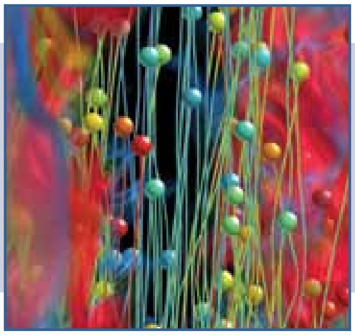
Magnetic systems with WL-LSMS. WL-LSMS analyzes magnetic systems at the nanoscale. It allows researchers to directly and accurately calculate the temperature above which a material loses its magnetism—known as the Curie temperature.

The magnetic properties of materials hold the key to major advances in technology. Magnetism at the atomic scale plays an important role in many materials, including steels and iron-nickel alloys, while lightweight-yet-strong permanent magnets are important components in highly efficient electric motors and generators.

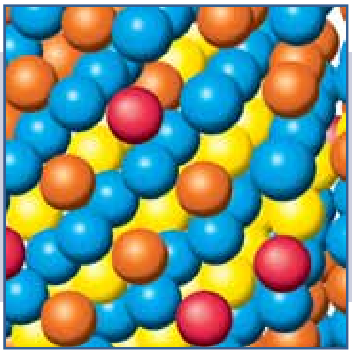
WL-LSMS earned its developers the 2009 Gordon Bell Prize as the world's highest-performance scientific application. On Jaguar, they have been able to calculate the energy of a magnetic state as a function of temperature, a daunting task.

On Titan, however, they will be able to do far more. They will be able to improve calculations of a material's thermodynamics, meaning the magnetic energy states at a given temperature. If they prefer, they can instead calculate the underlying magnetic states known as ground states without the huge margin of error that is currently unavoidable.

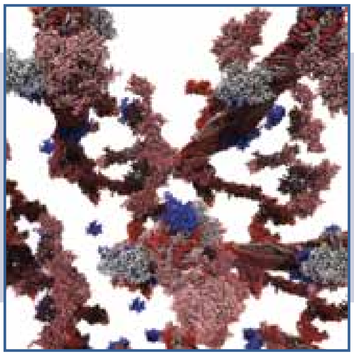
The CAAR applications



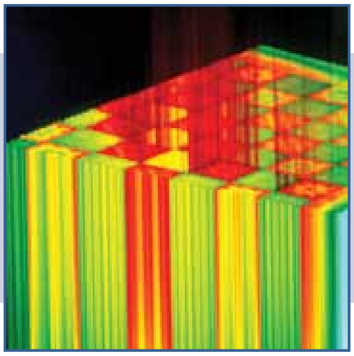
S3D
Combustion simulations to enable the next generation of diesel/biofuels to burn more efficiently.



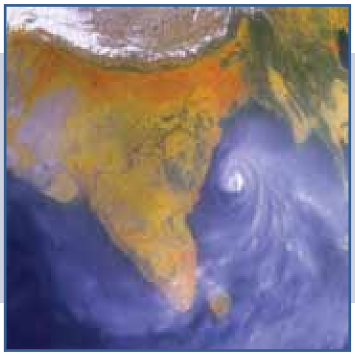
WL-LSMS
Role of material disorder, statistics, and fluctuations in nanoscale materials and systems.



LAMMPS
A simulation model investigating the properties of lignocellulose.



Denovo
High-fidelity radiation transport calculations that can be used in a variety of nuclear energy and technology applications.



CAM-SE
Answers questions about specific climate change adaptation and mitigation scenarios.

"Jaguar was the first system that enabled first-principles-based thermodynamics of magnetic materials," said ORNL's Markus Eisenbach, the application's lead developer. "The order of magnitude increase in computational power available with Titan will allow us to investigate even more realistic models that incorporate the interplay of magnetic fluctuations, magnetic moments, and atomic positions, which will lead to better predictions of material properties such as specific heat and magnetic susceptibility."

Molecular science with LAMMPS. LAMMPS explores molecular science applications through the use of molecular dynamics. It models a wide variety of problem, including membrane fusion, large biomolecular simulations for proteins, and lignocellulose for biofuels.

This research will contribute to the search for more efficient, economical methods for converting woody plants into biofuel. LAMMPS also promises to advance the modeling of proteins, which in turn opens new opportunities for understanding disease and improving medical treatment and targeted drug delivery.

On Jaguar LAMMPS has been able to simulate just 100,000 atoms for charged systems in which the long-range interactions are treated properly. On Titan, in contrast, the application would be able to overcome size limitations and expand to millions of atoms.

Nuclear reactors with Denovo. Denovo is a modeling and simulation tool used in many areas of nuclear technology, including reactor core analysis, radiation shielding, nuclear forensics, and radiation detection.

Nuclear power provides abundant electricity that is not reliant on foreign countries and does not emit greenhouse gases. For it to be effective, though, safety must be guaranteed and the amount of radioactive waste reduced. Denovo is a powerful tool for ensuring the safety and efficiency of nuclear power.

On Jaguar, it takes Denovo 60 hours to simulate a fuel rod through one round of use in a reactor core. On Titan that time will drop to 13 hours.

"Titan's heterogeneous accelerated nodes will make it possible for researchers to perform more accurate reactor calculations and thus accelerate the design of the next generation of safe, efficient reactors," noted Denovo team member Wayne Joubert of ORNL.

Climate change with CAM-SE. The Community Atmosphere Model–Spectral Element simulates long-term global climate to inform public policy and improve scientific understanding of climate changes.

Improved atmospheric modeling will help climate researchers better understand future air quality as well as the effects of particles suspended in the air. The indirect effects of particles are large sources of uncertainty regarding the climate's response to natural and human effects.

On Jaguar this application is able to simulate a grid of 14-kilometer cells at a quarter of a simulated year per day of computation. With Titan, researchers will be able to increase the simulation speed to between one and five years per computing day. The increase in speed is needed to make ultra-high-resolution, full-chemistry simulations feasible over decades and centuries and would allow researchers to quantify uncertainty by running multiple simulations.

"As scientists are asked to answer not only whether the climate is changing but where and how, the workload for global climate models must grow dramatically," noted CAM-SE team member Kate Evans of ORNL. "Titan will help us address the complexity that will be required in such models."— *by Leo Williams*

Tackling Titan

OLCF pulls no punches preparing users for revolutionary architecture

With the advent of Titan, scientific computing is entering a new era. Titan's hybrid architecture is the first step towards the vaunted exascale, a metric universally lauded by computational scientists around the world.

First things first, however. Titan is employing a revolutionary, disruptive technology: its hybrid architecture, which combines CPUs with GPUs, has never been attempted on this scale before.

While standing up a machine the size and scope of Titan with a completely new architecture is an impressive feat in itself, the system is only as good as the people and organizations that utilize it to advance their respective science. Although Titan's hybrid architecture comes with unique challenges, they have been and are being overcome. Moreover, one's challenge is another's opportunity, and the OLCF is beginning to see interest among previously unfamiliar user groups.

Recognizing the priority of quality training in the run-up to Titan's unveiling, the OLCF has embarked on a campaign of user education. The center's numerous research partnerships and workshops over 2011 are a testament to its philosophy of helping users get the most research out of the most stable machine possible.

In preparation for this education overhaul, the OLCF dedicated 2011 to educating users and the wider HPC community so that when Titan arrives, the world will be ready. While the pinnacle of the OLCF's user education schedule was certainly the Titan Summit, which brought a wide range of users and stakeholders together to confront issues from all angles, it was only one piece in a much larger puzzle designed to make Titan the premier scientific simulation research platform.

The OLCF's preparatory path took the organization and its users through a training schedule to prepare Titan's diverse user base for the new programming model. This meant education in three primary areas: identifying greater levels of parallelism, or finding areas in codes where tasks can be subdivided into smaller tasks and run simultaneously; using various tools to parallelize those areas; and debugging. In order for Titan to realize its potential, users will have to be successful in all three pursuits.

It all starts with finding the areas of code that can be improved, or exposing additional parallelism, said OLCF user support specialist Bobby Whitten. To accomplish this, the OLCF introduced users to various performance analysis tools such as CrayPAT and Vampir, particularly at the Vampir workshop held at the OLCF in May. These tools look at individual codes or programs and find the most time-consuming sections. Once these areas are spotlighted, users can begin the task of revamping their applications with Titan in mind.

There are several tools at users' disposal to help streamline their codes and get the most from their time on Titan. The standard method for programming the GPUs found in Titan is to use the CUDA programming model from NVIDIA. Unfortunately, this model has the disadvantage of only working with NVIDIA processors. That means that if the architecture were to somehow change in the future, users would be required to rewrite their code, an obviously disruptive liability on the OLCF user community. Therefore, OLCF staff decided to pursue an alternative approach in collaboration with its vendor partners in the form of directive-based compilers.

Compilers are special high-level language instructions. In essence, compilers are translators that convert one language to another. One way to allow compilers to create binary code that Titan's GPUs can understand is with the addition of directives, or special commands inserted into an application's source code that tells the compiler to convert a certain segment into a GPU-readable instruction. By using compilers that recognize directives, users are able to change their codes and are insulated from changes to future hardware. And because directive-based compilers automatically convert certain areas of an application to a GPU-friendly format, users are spared the details of the GPU hardware.

The OLCF offered PGI and HMPP workshops in July and September, respectively, to brief users on both companies' latest compiler technologies, all of which could get them up and running much faster than previously thought. However, different codes can require different approaches, and compilers might not be the best route for every application. That's when CUDA comes in handy.

Using CUDA, programmers are able to precisely target various features of the GPUs to ensure maximum performance. The OLCF's "An Introduction to CUDA" was the first workshop of 2011 and really got the ball rolling, said Whitten. When the OLCF first embarked on the road to Titan, CUDA was the only tool available to researchers to revamp their codes and thus make them capable of harnessing Titan's unprecedented computing power. However, learning the complex subtleties of a tool like CUDA can be overwhelming, and worse, time-consuming. While learning CUDA in-depth may seem like a small price to pay for the potential breakthroughs possible on Titan, OLCF staff immediately realized that it wasn't a practical solution for the diverse user base.

"The OLCF, along with CAPS and Cray, developed directive-based compiler tools and used performance analysis and debugging tools to help ease the transition to hybrid architectures," said Whitten.

However, CUDA remains an important asset in allowing users to harness Titan's raw power to the best of their ability. Almost as important as the final step on the road to tomorrow's most important scientific breakthroughs—debugging. Because Titan's architecture is so new, it is nearly unimaginable that an application that previously ran on Jaguar or another CPU-based system could port to the Titan hybrid platform without plenty of bugs being introduced.

To make the transition for all applications as smooth as possible, the OLCF has worked with Allinea, a leader in parallel software development tools, to enhance DDT, a scalable parallel debugger with support for GPU debugging. No matter the application, debugging will certainly play an important role when the OLCF's entire user base begins making the transition to Titan.

And while debugging may imply that users are primed to begin work on Titan, the OLCF's work never really ends. After the most cumbersome sections of code are identified and repaired and the entire application is enhanced to take advantage of Titan's GPUs and finally debugged, OLCF staff will go back and perform more performance analyses on the separate applications to ensure that they are running at top speed. If not, they keep working. If so, let the breakthroughs begin. But codes will continue to be optimized even as Titan begins to run the most accurate scientific simulations the world has ever seen.

While 2011 was a banner year for user education, there is still plenty to be learned on the road ahead. With the OLCF's help, however, Titan users have an enormous head start. *—by Gregory Scott Jones*

OLCF Titan Training Schedule

January

An introduction to CUDA, a course designed to teach staff about programming in NVIDIA's CUDA GPU framework. OLCF staff began examining issues, developing documentation and tutorials, and providing examples for users.

OLCF staff also began developing the OLCF3 introductory webinar, which culminated in a July 2011 webinar introducing Titan to users.

February

The exascale workshop brought staff and users together to discuss the future of hybrid-architecture supercomputing with respect to exascale computing.

March

OLCF spring training introduced Titan and its architectural details to users.

May

The Vampir workshop saw staff and users trained on the Vampir performance analysis tool. Vampir is an important tool for finding areas of application codes that can benefit from increased levels of parallelism.

July

Staff and users were invited to join a PGI (Portland Group, Inc.) workshop to introduce the latest compiler technologies from PGI. Compilers help port codes to Titan's GPU architecture.

An internal training session helped OLCF staff to develop and examine examples that were to be used to train users on state-of-the-art GPU programming.

August

The Titan Summit brought users, vendors, and staff together to discuss challenges, opportunities, and strategies regarding applications and their stability on Titan's revolutionary architecture.

September

HMPP (compiler) training brought staff and users up to date on HMPP technology.

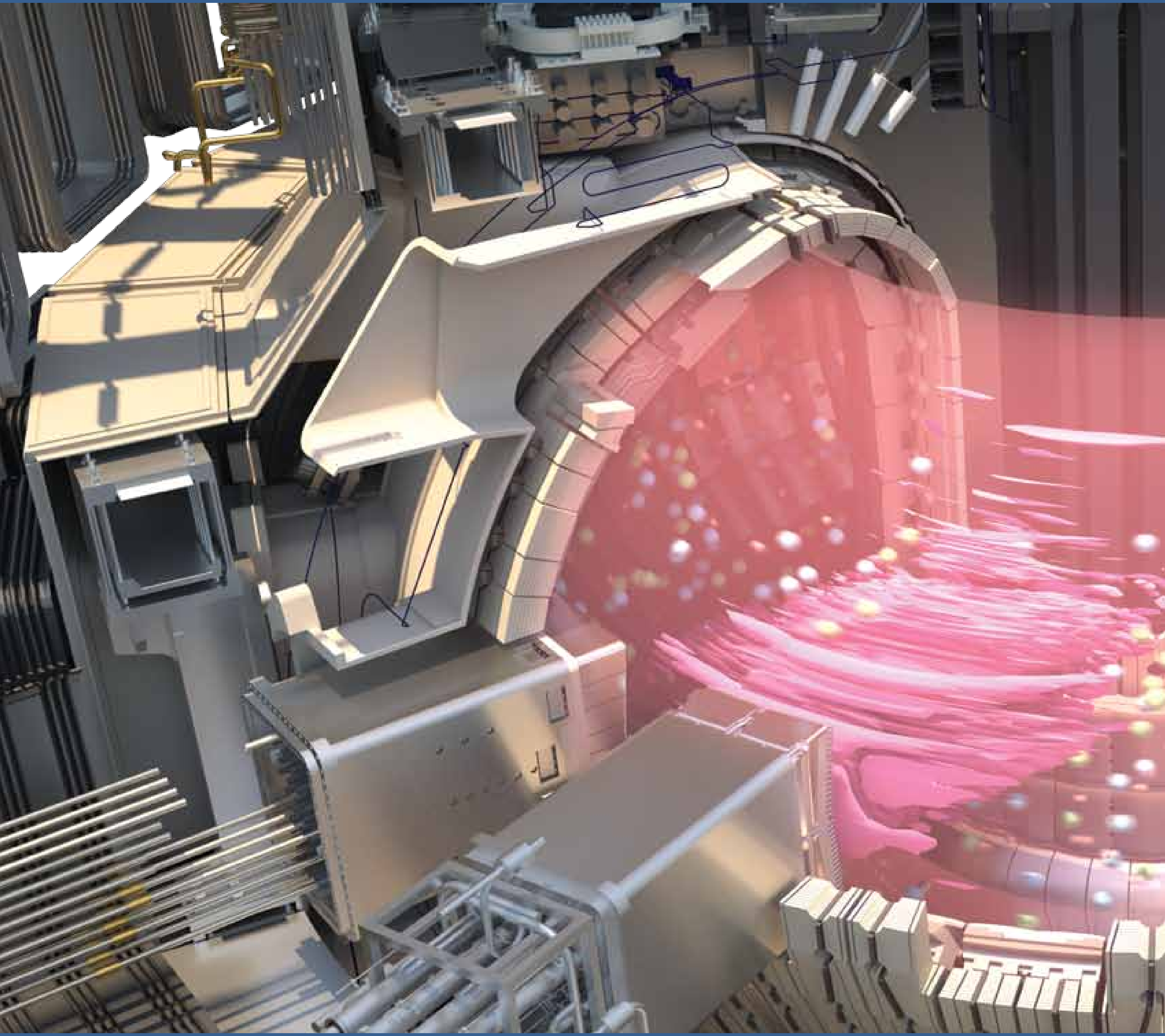
October

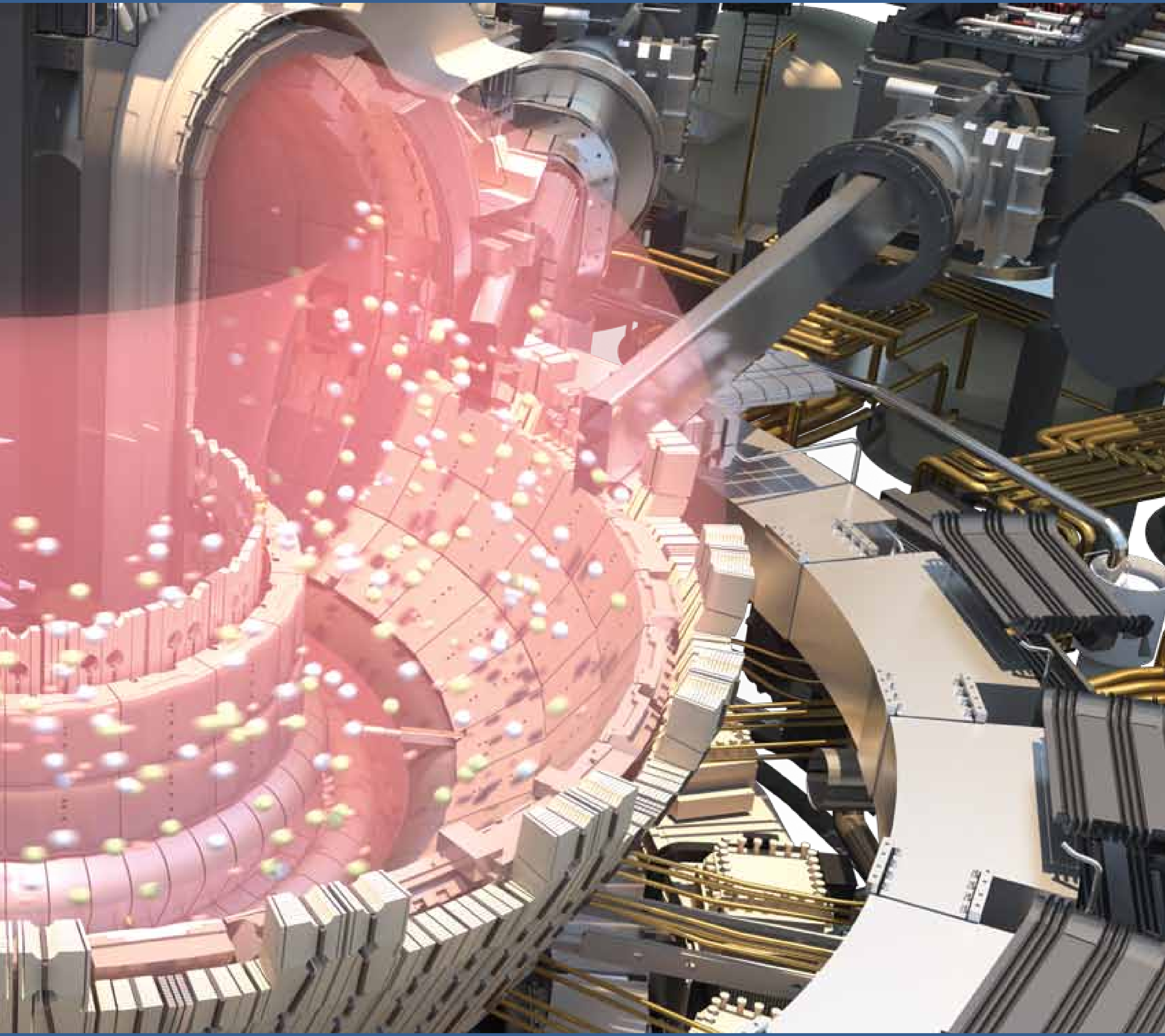
Fall Training saw users tackle more advanced GPU, tools, and technique training.

November

An SC11 birds-of-a-feather discussed Titan and the OLCF's strategy, tools, and techniques for successful utilization of hybrid architecture systems.

OLCF Science





From materials to astrophysics to climate to fusion (shown here), the OLCF has been host to numerous breakthroughs across the scientific spectrum. The year 2011 both continued that legacy and provided a glimpse of the future as teams began to transition to Titan, the next great leap in computational science.

Q&A: Jack Wells, Director of Science, OLCF

On July 1, 2011, Jack C. Wells became the director of science for the NCCS at ORNL.

Wells has led ORNL groups in computational nanotechnology, computational materials science, and nanomaterials theory. He has also served as a scientific advisor to Tennessee Senator Lamar Alexander during an off-site assignment. Most recently, he directed ORNL's Office of Institutional Planning, where his responsibilities included developing a strategic plan for the lab, overseeing discretionary research and development investments, and managing Advanced Research Projects Agency–Energy (ARPA-E) programs. He has a long history with ORNL supercomputing and was a pioneering user of the facility. As principal investigator of a project to understand the chemistry of rechargeable lithium/air batteries, which potentially can store ten times more energy than lithium/ion batteries of the same weight, he and his collaborators received 24 million processor-hours on Argonne and Oak Ridge supercomputers through the INCITE program.

In this interview, Wells describes his vision for executing a scientific strategy for the NCCS that ensures cost-effective, state-of-the-art computing to facilitate DOE's scientific missions. To begin this decade's transition to exaflop computing (capable of carrying out a million trillion floating point operations per second), plans are in the works for a staged upgrade of Jaguar. A high-performance computing (HPC) system employing traditional CPU microprocessors, Jaguar will be transformed into Titan, a hybrid system employing both CPUs and GPUs,energy-efficient number crunchers that accelerate specific types of calculations in scientific application codes. As the OLCF gears up to deliver the system, expected to have a peak performance of at least 20 petaflops, by early 2013, Wells' challenges are many.

Congratulations on your new job. What is your vision for the NCCS?

My vision is a user facility that recognizes that its products are scientific discovery and technical innovation and that we will achieve this vision working with strong partners. These partnerships will be developed through our calls for proposals and our outreach efforts to the scientific, technology, and industrial communities. To broaden the scope of leadership computing, we need to engage through our networks, through our relationships, and encourage users from new communities that can take good advantage of these resources to move us forward in scientific discovery, industrial competitiveness, and sustainability. Partnerships and alliances are a big part of what I think is important about my job. In terms of challenges, clearly we're at a cusp in technology moving to hybrid architectures. This implies a lot of hard work by a lot of people. But it also is a game changer, meaning that all communities are going to have to do a lot of work. No community is already prepared for these new technologies penetrating scientific computing. It's a time when users and developers of traditional supercomputing applications will have to race to catch up, too. It's a good time for users of new applications to get in.

Jack Wells, OLCF Director of Science

What are some examples of new users you hope to entice?

Energy technologies are clearly a topic that is interesting. We want to have impact on society, and energy technology is a compelling mission of the Department of Energy (DOE) in which DOE has a leadership role; modeling and simulation can make a big impact on the rate of innovation going forward. A wonderful example today is the investment that DOE has made through the Office of Nuclear Energy in the modeling and simulation hub, the Consortium for Advanced Simulation for Light Water Reactors (CASL). This is a huge and in some sense groundbreaking investment in modeling and simulation in energy technology. The codes that are being developed can impact nuclear energy, nuclear engineering, and many other areas in energy technology. Accelerated innovation resulting from computing is also starting to happen in areas funded by DOE's Office of Energy Efficiency and Renewable Energy in battery simulation and solar cell simulation. One can imagine a vision in which we have simulation tools for all these energy systems to accelerate innovation and design, help our industry be more competitive, and solve our societal problems.

As the head of an INCITE project to solve the issues that prevent electric cars from running 500 miles on a single battery charge, can you share your experience as a user of the facility you now lead?

This is a project with Argonne National Laboratory, our sister laboratory, and some major industrial partners with strong research and development capabilities, specifically IBM and Bosch, where we focused on what some have called the Mount Everest in battery chemistries, lithium/air. Just from simple chemistry arguments, it has the greatest potential for energy density storage. But by no means do these batteries work today as rechargeable cells in any sense that you would recognize. The fundamental mechanisms of reaction are not understood. The way to build rechargeable energy storage systems that would work is not understood. This INCITE project is trying to make significant progress in the basic chemistry of these reactions to help us understand, is there the potential that one could engineer these systems? Right now, several of the basic issues are being made clear. We shouldn't look for this battery chemistry to impact technology in the near term, and by near term I mean within the next 5 years. But with breakthroughs one could have disruptive change in the marketplace within 10 years. But it's certainly not guaranteed, and other chemistries out there beyond today's lithium/ion batteries are also exciting, like zinc/air or lithium/sulfur batteries. These are topics to which HPC has not been applied historically. There have been modeling and simulation of battery systems, but typically at the level of phenomenological models. Significant progress has been made; but over the coming years, because of the significance of the technology, many, many more people are going to be engaged in this in universities, labs, and companies, and HPC will play a big role. The codes we were running on the leadership computers today are state-of-the-art, first-principles chemistry and materials codes because we're asking basic science questions about these systems. New multiscale battery codes need to be built that can deal with systems-level issues, not just chemical reactions on anodes and cathodes, to be able to describe this as a complex engineering system. No one can do that today.

What was your role in ORNL's Computing and Computational Sciences Directorate before it housed and ran a national user facility?

I came here as a Vanderbilt graduate student working on Office of Science–funded projects in nuclear and atomic physics. My Ph.D. was sponsored by a grand challenge project funded under a program that started with the High Performance Computing and Communications Act of 1992—that's called the Gore Act because Senator Albert Gore, Jr. was the main sponsor in the US Senate. It's through that, as the old story goes, he "invented" the Internet. It was that program, which partnered HPC science teams from around the country with ORNL computer scientists and hardware vendor Intel, that founded the Center for Computational Sciences (CCS) originally in 1992. After a postdoc at Harvard I came back to ORNL in '97 as a Wigner Fellow in the CCS. Buddy Bland, project director of the OLCF-2, which built the petascale Jaguar system, and the OLCF-3, which will build the even more powerful Titan, was my first group leader. I worked in the Scientific Computing Group on parallel code performance optimization and doing my science in theoretical atomic and molecular physics. I did use the CCS computers that we had in my Ph.D. thesis—the Intel iPSC/860 and Intel XP/S 5 Paragon. Then when I came back in '97 we had the XP/S 35 Paragon and XP/S 150, too. We transitioned to the IBM Eagle by about 1999. The point is that we had a CCS even before we had a Leadership Computing Facility. Beginning in 1999, I worked on basic materials and engineering physics programs in DOE's Office of Science–Basic Energy Sciences. And then when the Center for Nanophase Materials Sciences, or CNMS, was constructed at Oak Ridge, I along with my group was matrixed to form the Nanomaterials Theory Institute at the CNMS. During that time, Oak Ridge competed for and won the DOE Leadership Computing Facility in 2004. The significant thing is that CCS has been here for almost 20 years. In 2012 we celebrate our 20-year anniversary.

Did directing institutional planning for ORNL provide lessons that might guide you in your new role?

What I learned from working for our laboratory director's office from August of 2009 through June of 2011—that's the job I was just doing before I came to the NCCS—is that both planning and science are about the future. We need to not be constrained in our thinking by the status quo, but to try to establish a clear and compelling vision for the future for our science programs, for our institution, and ultimately, in collaboration with others, for our nation. We need to not always think about what is, but what could be, and why it would be an attractive future. ARPA-E, a DOE program to spur energy innovations, is an interesting case of a good idea articulated by policymakers that was fairly rapidly put in place. It was authorized by Congress and then implemented by DOE, initially through Recovery Act funding, to bring a new approach to funding high-risk, high reward energy technology research within DOE. It's been reviewed very well by industry and its sponsors in Congress. The ability to take risks and reach for the big payoffs is something that we should think about and try to implement when we can.—*Q&A with Dawn Levy*

Simulations Explore Next-Generation Fuels

Supercomputers help optimize engines, turbines, and other technologies for clean energy

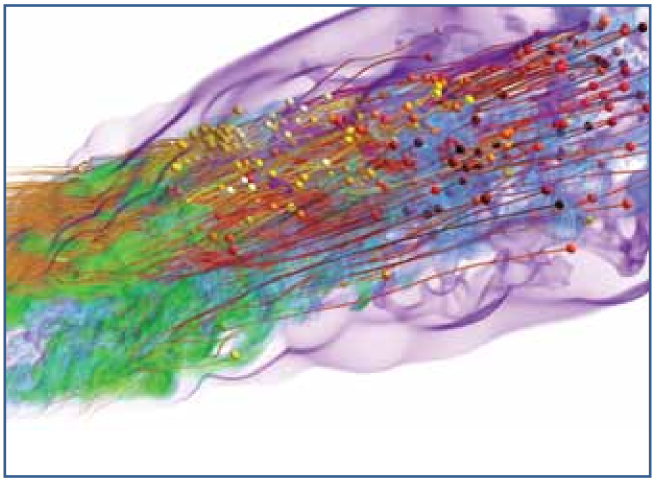
Air and fuel mix violently during turbulent combustion. The ferocious mixing needed to ignite fuel and sustain its burning is governed by the same fluid dynamics equations that depict smoke swirling lazily from a chimney. Large swirls spin off smaller swirls and so on. The multiple scales of swirls pose a challenge to the supercomputers that solve those equations to simulate turbulent combustion. Researchers rely on these simulations to develop clean-energy technologies for power and propulsion.

A team led by mechanical engineers Joseph Oefelein and Jacqueline Chen of Sandia National Laboratories (Sandia) simulates turbulent combustion at different scales. A burning flame can manifest chemical properties on small scales from billionths of a meter up to thousandths of a meter, whereas the motion of an engine valve can exert effects at large scales from hundredths of a meter down to millionths of a meter. This multiscale complexity is common across all combustion applications—internal combustion engines, rockets, turbines for airplanes and power plants, and industrial boilers and furnaces.

Chen and Oefelein were allocated a total of 113 million hours on OLCF's Jaguar supercomputer in 2008, 2009, and 2010 to simulate autoignition and injection processes with alternative fuels. For 2011 they received 60 million processor hours for high-fidelity simulations of combustion in advanced engines. Their team uses simulations to develop predictive models validated against benchmark experiments. These models are then used in engineering-grade simulations, which run on desktops and clusters to optimize designs of combustion devices using diverse fuels. Because industrial researchers must conduct thousands of calculations around a single parameter to optimize a part design, calculations need to be inexpensive.

"Supercomputers are used for expensive benchmark calculations that are important to the research community," Oefelein said. "We [researchers at national labs] use the Oak Ridge Leadership Computing Facility to do calculations that industry and academia don't have the time or resources to do."

The goal is a shorter, cheaper design cycle for US industry. The work addresses DOE mission objectives to maintain a vibrant science and engineering effort as a cornerstone of American economic prosperity and lead the research, development, demonstration, and deployment of technologies to improve energy security and efficiency. The research was funded by DOE through the Office of Science's Advanced Scientific Computing Research (ASCR) and Basic Energy Sciences programs, the Office of Energy Efficiency and Renewable Energy's Vehicle Technologies program, and the Recovery Act–funded Combustion Energy Frontier Research Center.



High-fidelity large eddy simulation (LES) of direct-injection processes in internal-combustion engines provides an essential component for development of high-efficiency, low-emissions vehicles. Here LES reveals how fuel from a state-of-the-art injector mixes with air inside an engine cylinder. Image credit: Joseph Oefelein and Daniel Strong, Sandia National Laboratories.

Making combustion more efficient would have major consequences, given our reliance on natural gas and oil. Americans use two-thirds of their petroleum for transportation and one-third for heating buildings and generating electricity. "If low-temperature compression ignition concepts employing dilute fuel mixtures at high pressure are widely adopted in next-generation autos, fuel efficiency could increase by as much as 25 to 50 percent," Chen said.

Complementary codes

Chen uses a direct numerical simulation (DNS) code called S3D to simulate the finest microscales of turbulent combustion on a three-dimensional virtual grid. The code models combustion unhindered by the shape of a device. These canonical cases emulate physics important in both fast ignition events and slower eddy turbulence and provide insight into how flames stabilize, extinguish, and reignite.

Oefelein, on the other hand, uses a large eddy simulation (LES) code called RAPTOR to model processes in laboratory-scale burners and engines. LES captures large-scale mixing and combustion processes dominated by geometric features, such as the centimeter scale on which an engine valve opening might perturb air and fuel as they are sucked into a cylinder, compressed, mixed, burned to generate power, and pushed out as exhaust.

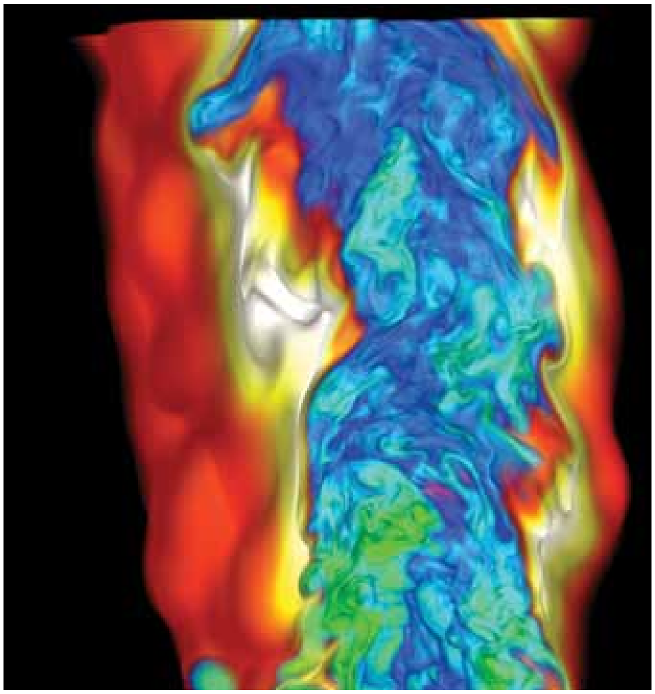
Whereas DNS depicts the fine-grained detail, LES starts at large scales and works its way down. The combination of LES and DNS running on petascale computers can provide a nearly complete picture of combustion processes in engines.

“Overlap between DNS and LES means we’ll be able to come up with truly predictive simulation techniques,” Oefelein said. Today experimentalists burn fuel, collect data about the flame, and provide the experimental conditions to computational scientists, who plug the equations into a supercomputer that simulates a flame with the same characteristics as the observed flame. Researchers compare simulation with observation to improve their models. The goal is to gain confidence that simulations will predict what happens in the experiments. New products could then be designed with inexpensive simulations and fewer prototypes.

In a high-fidelity DNS with cells just 5 millionths of a meter wide, Chen’s team modeled combustion in a canonical high-pressure, low-temperature domain to investigate processes relevant to homogeneous charge compression ignition. Another simulation, using 7 billion grid points, explored syngas, a mix of carbon monoxide and hydrogen that burns more cleanly than coal and other carbonaceous fuels. It used 120,000 (approximately half) of Jaguar’s processors and generated three-quarters of a petabyte of data.

Since their 2008 simulation of a hydrogen flame—the first to fully resolve detailed chemical interactions such as species composition and temperature in a turbulent flow environment—they have simulated fuels of increasing complexity. Because real-world fuels are mixtures, simulations use surrogates to represent an important fuel component. Current simulations focus on dimethyl ether (an oxygenated fuel) and n-heptane (a diesel surrogate). Simulations are planned for ethanol and isooctane (surrogates for biofuel and gasoline, respectively).

Oefelein uses RAPTOR to investigate turbulent reacting flows in engines. In a GM-funded experimental engine designed at the University of Michigan and used by the research community as a benchmark, he explores phenomena such as cycle-to-cycle variations, which can cause megaknock—an extreme form of the “bouncing marbles” sound sometimes heard in older engines. The fuel injection, air delivery, and exhaust systems are interconnected, and pressure oscillations have huge effects on engine performance. “Megaknock can do significant damage to an engine in one shot,” said Oefelein. He noted that combustion instabilities also plague power plants, where fear forces designers of gas turbines to make conservative choices that lessen performance. “Combustion instabilities can destroy a multimillion dollar system in milliseconds,” Oefelein said.



Direct numerical simulation reveals turbulent combustion of a lifted ethylene/air jet flame in a heated coflow of air. Computation by Chun Sang Yoo and Jackie Chen of Sandia National Laboratories, rendering by Hongfeng Yu of SNL. Reference: C. S. Yoo, E. S. Richardson, R. Sankaran, and J. H. Chen, “A DNS Study of the Stabilization Mechanism of a Turbulent Lifted Ethylene Jet Flame in Highly-Heated Coflow,” *Proc. Combust. Inst.* **33** (2011)1619–1627.

Petascale simulations will light the way for clean-energy devices and fuel blends that lessen combustion instabilities. The data they generate will inform and accelerate next-generation technologies that increase energy security, create green jobs, and strengthen the economy.

Chen and Oefelein’s collaborators include Gaurav Bansal, Hemanth Kolla, Bing Hu, Guilhem Lacaze, Rainer Dahms, Hongfeng Yu, and Ajith Mascarenhas of Sandia; Ramanan Sankaran of ORNL; Evatt Hawkes of the University of New South Wales; Ray Grout of the National Renewable Energy Laboratory; and Andrea Gruber of SINTEF (the Foundation for Scientific and Industrial Research at the Norwegian Institute of Technology).—*by Dawn Levy*

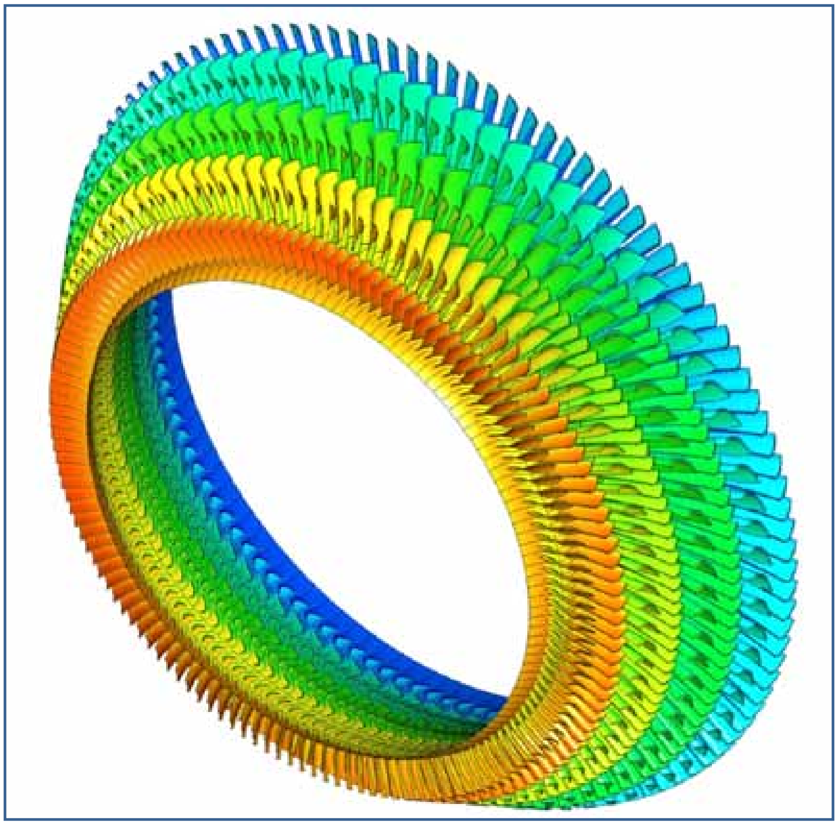
Jaguar Accelerates Design of GE Turbomachinery

Researchers simulate the design of next-generation turbines

Few technologies are more vital to modern day life than turbomachines, the bladed devices used to convert fluid power to and from mechanical power to drive propellers, wind turbines, gas turbines, steam turbines, fans, and compressors. These engineering marvels are literally responsible for keeping the lights on, given that most of the world's electricity is generated by various kinds of turbines. And jets couldn't stay aloft without them to drive a plane's large-diameter fans.

General Electric (GE), the iconic technology, service, and finance company and American industrial leader, has been building turbomachines for nearly a century and is currently a major producer for the electric power generation and aircraft engine industries. Recently, however, GE took its turbomachinery research and development to the fast lane with the help of one of the fastest computers in the world.

Because of the amount of research invested in today's turbomachinery and the sophisticated nature of its engineering, companies have repeatedly grabbed the low-hanging fruit to achieve increased efficiencies, and competition is fierce.



A visualization of a four-stage low-pressure turbine modeled on the Jaguar supercomputer. With their flagship code known as Tacoma, GE ran its largest ever computational fluid dynamics simulation on Jaguar. Image courtesy GE Global Research.

"It's a very competitive business," said principal engineer Graham Holmes of GE Global Research, noting that the company is up against Rolls Royce and Pratt & Whitney in the aviation market and Siemens and Alstom in the power generation arena. "If you could achieve a 1 percent increase in efficiency for a turbomachine, the market would be yours." That 1 percent fuel-burning advantage would, over time, add up to enormous energy and cost savings for GE's customers and provide GE with a business-critical advantage, according to Holmes.

Enter Jaguar, DOE's flagship Cray XT5 HPC system, located at ORNL. Through ORNL's HPC Industrial Partnerships Program, GE recently harnessed Jaguar's power to study the unsteady fluid flows in turbomachines in greater detail than ever before accomplished. Understanding these flows is essential to achieving greater efficiency and for GE to gain an edge in an intensely competitive global marketplace.

It's all in the blades

The basic physics of turbomachinery operations have been well understood for years—jet engines and gas turbines go back to the mid-twentieth century. Essentially, turbomachines feature alternating rows of stationary and moving blades either expanding or compressing gas. The design process has evolved from experimentation and highly simplified analytical models to increasingly sophisticated simulations carried out on increasingly powerful computers. Engineers typically shape blades, run a combination of simulations and experiments, tweak the design, and repeat, which is an expensive path to production by any measure. GE runs these simulations on its in-house Linux clusters. But even with these systems, turbomachinery designers have had to assume that the velocity of air around and across the blades remains steady in the reference frame of the blades, or as seen from the point-of-view of the actual blades.

Turbomachinery designers have always understood that this air flow is unsteady and that it has to be unsteady for a turbomachine to work. But the assumption that the flow, as seen by each blade row, can be approximated as steady has proved to be remarkably powerful. The designs of all the most efficient turbomachines, such as the turbines that drive the large-diameter fans in modern jet engines, have been created using this paradigm.

Any future efficiency improvements will likely depend on deciphering the unsteady nature of the fluid flow. For example, in jet engines, low-pressure turbines drive increasingly larger diameter fans, rotating at lower revolutions per minute to increase efficiency. But running low-pressure turbines at lower speeds presents a severe technical challenge—the lower the relative velocity between the blade rows, the harder it is to extract energy to drive the fan, requiring more rows of blades. Unfortunately, as more rows of blades are added, the turbine becomes heavier and therefore less efficient. Understanding the unsteady flows should allow designers to make needed adjustments without adding to the weight of the turbine, resulting in overall greater fuel efficiency.

Unsteady flow analysis is also essential to understanding other phenomena such as blade flutter, or the blade vibration induced by the fluid flow. Blade flutter can be catastrophic in an aircraft engine if it results in damage to one or more blades and turbine failure.

Ready to ramp up

GE was eager to compare its long-standing steady flow assumptions with actual unsteady flow calculations, but this posed a major computational challenge. Unsteady simulations are orders of magnitude more complex than simulations of steady flows and beyond the capability of GE's Linux clusters. So GE applied for and received access to Jaguar, DOE's most powerful supercomputer at the time and one of the most powerful in the world.

Holmes was joined by Branden Moore of GE Global Research's Advanced Computing Lab to make the most of this opportunity. They further expanded the collaboration with senior engineer Stuart Connell on the visualization front to better decode the mountain of data being produced in action.

"We definitely have the benefit of cross-discipline work at GE," said Moore, adding that this collaborative nature ultimately helps GE's HPC applications run faster. And running faster is a key metric in HPC. The faster a code runs, the more simulations that can be run in a given timeframe. And more simulations increase the opportunity for new insights and new scientific discovery. GE used its flagship code Tacoma, which features specialized computational fluid dynamics (CFD) for turbomachinery. When they paired Tacoma with Jaguar, GE researchers ran their largest-ever CFD calculation and were able to investigate for the first time the unsteady flows in turbomachinery. Simulations were then ramped up from three to four dimensions, and researchers were able to look at the time-resolved unsteady flows in the moving blades.

Ultimately, said Holmes, GE's goal is to make better machines. The suite of simulations, which will benefit GE Aviation and GE Energy, was "an attempt to explore a better way of analyzing flow in those machines," he said. "Designers want to know how to shape the blades in a turbine to improve performance. With Jaguar our team did just that, investigating various blade interactions and the unsteady flow around them to see how the overall machine performed."

Tomorrow's turbine

With access to Jaguar, the team was able to examine a turbine test rig and compare steady and unsteady flows. In the two analyses, the efficiency remained the same, which is "an extremely valuable piece of information," said Holmes. Furthermore, the team found plenty of interesting unsteady phenomena occurring throughout the device. For instance, the interactions between the blade and hub created unsteady secondary flows, which behaved differently from those witnessed in the steady analysis.

The team does not yet know why the secondary flows appear unique, but they definitely want to, said Holmes. "We need to dig in and use this information to improve our visualization skills," he said, adding that the team needs to develop a deeper comprehension of the differences in flows and decipher whether the environment or various test factors are to blame. After all, in a global industry chasing a 1 percent increase in efficiency, every little bit counts.

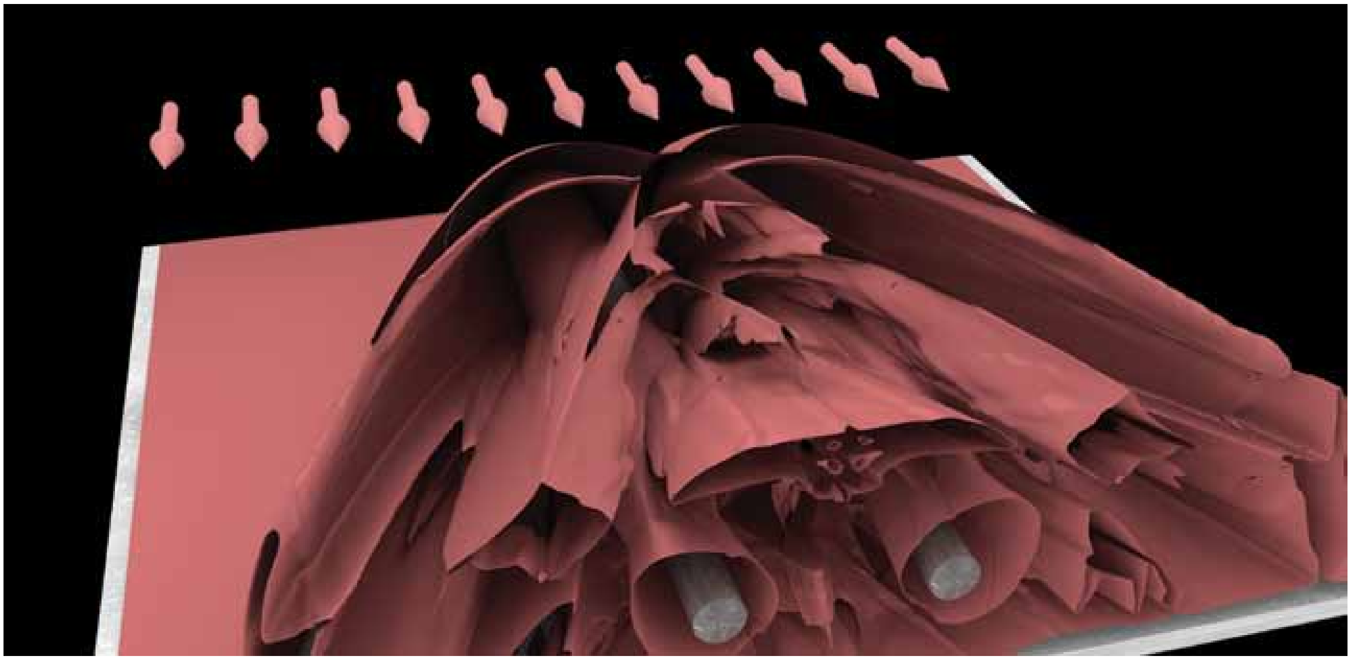
Overall, said Holmes and Moore, GE's simulations on Jaguar advanced the company's research and development in the turbomachinery arena and are providing it with a distinct competitive advantage as the company pores over the simulation results.

The team believes that GE and its competitors will move further into unsteady flow analysis to achieve the final point in efficiency, a move that will require substantial HPC resources. In fact, largely as a result of these calculations, GE recently purchased its own Cray system, a move that significantly ramps up its in-house HPC capability.

According to Holmes, "we simply could not design a competitive turbine or jet engine without these CFD tools. We would have exited the business a long time ago on both aviation and power generation."—*by Gregory Scott Jones*

Ramgen Simulates Shock Waves, Makes Shock Waves Across Energy Spectrum

Simulation on Jaguar shows potential for cheaper, more efficient carbon capture



Visualization of mach 3.75 isosurfaces in a two-body simulation showing complex fluid structures.
Image courtesy Michael Matheson, ORNL, and Ramgen Power Systems

One of the most pressing scientific challenges facing the United States and the world is reducing greenhouse gas emissions. Compounding that challenge is the fact that power plants burning fossil fuels account for more than 40 percent of the world's energy-related CO_2 emissions and will continue to dominate the supply of electricity until the middle of the century. There is an urgent need for cost-effective methods to capture and store their carbon emissions.

DOE is currently sponsoring large-scale demonstration projects to prove the viability of carbon capture and sequestration (CCS). The principal barrier to widespread application of CCS is its cost. Once the CO_2 is captured, compressing it to the required 100 atmospheres represents approximately 33 percent of the total cost of CCS.

Ramgen Power Systems, a small, Seattle-based energy research and development firm, is developing a novel gas compressor system based on shock-wave technology used in supersonic flight applications. This technology holds important promise for the turbomachinery industry of engines and compressors. Ramgen is a world leader in applying this shock-wave-based compression technique to gases, including CO_2—a more challenging application than air because of CO_2's larger molecular weight. DOE's National Energy Technology Laboratory is providing cost-shared support for this project.

The traditional process to design and optimize new turbomachinery, or machines that transfer energy between a rotor and a fluid, involves the testing of multiple physical prototypes, which is expensive and takes more time than DOE's demonstration schedule permits. Ramgen modified this conventional development process by using more extensive computer simulation validated by test results. DOE leadership determined that applying the most capable HPC systems and modern CFD analysis would further accelerate optimizing the turbomachinery's performance.

Ramgen turned to Jaguar, one of DOE's flagship supercomputers located at OLCF. With a peak speed of 3.3 petaflops, Jaguar is currently one of the fastest computers in the world. Ramgen was granted time on Jaguar through DOE's ASCR Leadership Computing Challenge allocation program (http://science.energy.gov/ascr/facilities/alcc/). As a result, the company has been making dramatic progress in designing its advanced turbomachinery for compressing CO_2 and for generating electricity with a breakthrough engine capable of using dilute methane for fuel.

Ramgen has made enormous leaps in its ability to apply the formidable capability of Jaguar to solve complex engineering problems. The accelerated research and development effort is a major factor in the significant reduction in the projected time required to optimize the technology performance for commercial applications; in fact, simulations at the OLCF are guiding testing of prototypes in 2012. These results are an excellent example of what can be accomplished by combining HPC with industry research and development.

Power of Jaguar advances technology

With the expertise and computing power of the OLCF, Ramgen has been able to speed up development of its breakthrough technology through a much more rapid assessment of different design modifications and operating conditions. Only a machine with Jaguar's computing power is capable of accurately simulating the complex fluid dynamics involved in shock-wave compression technology. Scaling up applications to more effectively use Jaguar's power is a core capability of the OLCF's computational scientists.

The OLCF aided Numeca, Ramgen's software vendor, by providing consultation and direction in improving the performance of Ramgen's code. Diagnostic tools like Vampir (designed and managed by Technische Universität Dresden), which measures the parallel performance of certain codes running at large scales, allowed for the identification of code bottlenecks; and the adoption of ORNL's Adaptable I/O System (known as ADIOS) for high-performance file reading and writing reduced problems with initialization and restarts and improved Ramgen's use of Jaguar.

These steps resulted in a 100-fold speedup in time-to-solution and a 2-fold decrease in memory usage per core, which allows the use of higher-resolution computational meshes. Additionally, OLCF's recent upgrade of Jaguar to the Cray XK6 architecture doubled available memory per core, further improving memory use and enabling greater levels of performance.

Using Jaguar, Ramgen has employed sophisticated optimization algorithms, known as "design-of-experiments" techniques, to achieve aerodynamic performance improvement for shock-wave compression. The firm ran hundreds of design combinations with up to 50 design parameters simultaneously ("ensembles") to find the optimal design solution for the technology. Such a process requires a system like Jaguar with hundreds of thousands of processors, along with sophisticated mathematical algorithms to analyze and predict this optimal solution. Other firms have tried similar experiments but have not been successful. "The use of Jaguar has cut the projected time from concept to a commercial product by at least 2 years and the cost by over $2 million," said Ramgen's chief executive officer and director, Doug Jewett.

Ramgen's design ensembles effectively used 80 percent of Jaguar's available computing resources. This intelligently driven optimization process reduced what used to be months of work to a mere 8 hours. More recently Ramgen succeeded in running ensembles of 240,000 cores that again effectively used 80 percent of the available XK6 system. Ramgen's Allan Grosvenor noted, "Two years ago it was impossible for us to run an ensemble of simulations like this. The sophisticated design analyses that we are running now are having a significant impact on our turbomachine development work."

Jewett added, "Jaguar provides, in a remarkably short time, data that we use to predict optimal designs. It's enabling us to advance the design of our equipment in a timeframe that simply would not have been possible without Jaguar and the assistance of Mike Matheson [of the OLCF's Scientific Computing Group] and the OLCF."

Accelerating this timeline is critical to the nation's effort to curb the amount of CO_2 emitted into the atmosphere. Thanks to this effective industry–lab partnership and Jaguar's scalable architecture, Ramgen will initiate testing of a 13,000-horsepower CO_2 compressor this year. This compressor is projected to reduce the capital costs of CO_2 compression by 50 percent and produce a minimum of 25 percent savings in operating costs. Applying these cost savings to a new 400-megawatt clean coal plant would result in capital cost savings of approximately $22 million and an annual operating cost savings of approximately $5 million.

Capturing more than carbon?

Applying the power of Jaguar to shock-wave-based compressors at the scale now being achieved represents a paradigm shift in how new turbomachinery is developed. "I believe that by applying the optimization enabled by Jaguar, Ramgen will accomplish levels of aerodynamic refinement with shock-wave-based technology in 5 years that took 50 years with gas turbines," said Ramgen's chief technology officer, Shawn Lawlor.

There is enormous synergy in the work done on shock-wave-based compression for CO_2 and the Integrated Supersonic Component Engine (ISC Engine), another product Ramgen is developing. The ISC Engine will generate electricity—efficiently and cost effectively—using dilute methane gases released during coal mining operations and from landfills. Currently approximately 90 percent of this methane is simply vented to the atmosphere. Methane, per volume, traps 21 times more heat in the atmosphere than CO_2 but has a shorter atmospheric lifetime. Engineers from Jim Walter Resources, a mining company, and Ramgen staff have developed an approach that can use up to 75 percent of the methane being vented to the atmosphere worldwide, as fuel to generate electricity.

This compression process has additional exciting possible applications with other gases and for other products. "What all these technologies have in common," said Jewett, "is that they'll boost the gas pressure significantly at lower cost than alternatives." With the help of Ramgen's engineers and the increasing power of supercomputers such as Jaguar, these technologies can enter the marketplace sooner, proving that together industry and HPC can change the world in profound ways. *—by Gregory Scott Jones*

Climate Scientists Compute in Concert

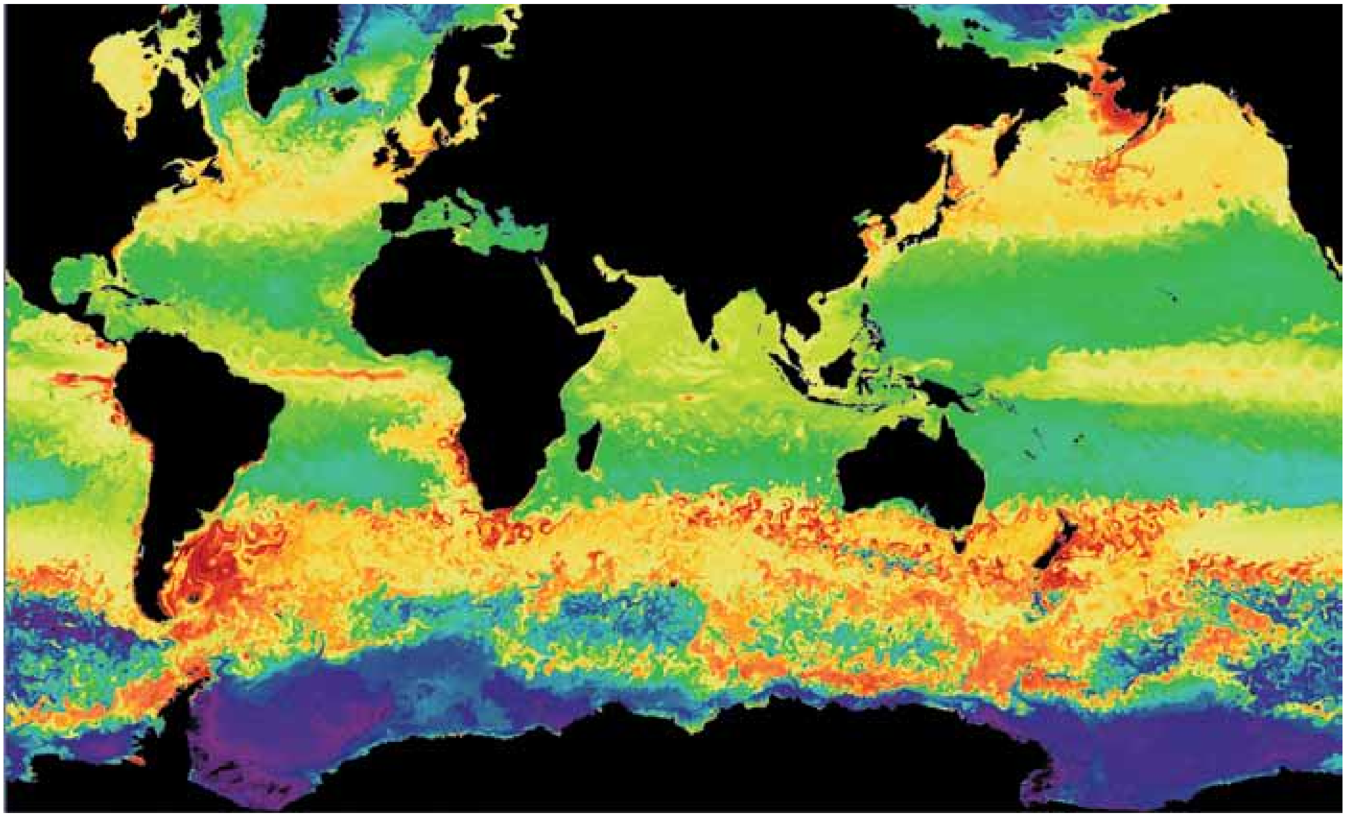
Collaborators work to reduce uncertainty in ice, Earth, atmosphere, and ocean simulations

Researchers at ORNL are sharing computational resources and expertise to improve the detail and performance of a scientific application code that is the product of one of the world's largest collaborations of climate researchers. The Community Earth System Model (CESM) is a mega-model that couples components of atmosphere, land, ocean, and ice to reflect their complex interactions. By continuing to improve science representations and numerical methods in simulations, and exploiting modern computer architectures, researchers expect to further improve the CESM's accuracy in predicting climate changes. Achieving that goal requires teamwork and coordination rarely seen outside a symphony orchestra.

"Climate is a complex system. We're not solving one problem, but a collection of problems coupled together," said ORNL computational Earth scientist Kate Evans. Of all the components contributing to climate, ice sheets such as those covering Greenland and Antarctica are particularly difficult to model—so much so that the Intergovernmental Panel on Climate Change (IPCC) could not make any strong claim about the future of large ice sheets in its 2007 Assessment Report, the most recent to date.

Evans and her team began the Scalable, Efficient, and Accurate Community Ice Sheet Model (SEACISM) project in 2010 in an effort to fully incorporate a three-dimensional, thermomechanical ice sheet model called Glimmer-CISM into the greater CESM. The research is funded by DOE's ASCR office. Once fully integrated, Glimmer-CISM will be able to send information back and forth among other CESM codes, making it the first fully coupled ice sheet model in the CESM.

Through the ASCR Leadership Computing Challenge, the team of computational climate experts at multiple national labs and universities received allocations of processor hours on OLCF's Jaguar, which at the time was capable of 2.3 petaflops, or 2.3 thousand trillion calculations per second.



Ocean simulations that include ecosystems are important for determining how much carbon the oceans can sequester in current and future climates. Chlorophyll, shown in green, is a useful measure of biological activity that can be compared with satellite ocean color measurements. Image credit: Matthew Maltrud, Los Alamos National Laboratory

ORNL climate scientist Kate Evans (left) and computer scientist Patrick Worley discuss ways to improve a climate simulation projected on the EVEREST visualization wall. By scaling up climate codes, researchers are able to simulate individual climate components in higher resolution and get more precise results. Photo credit: Jason Richards, ORNL

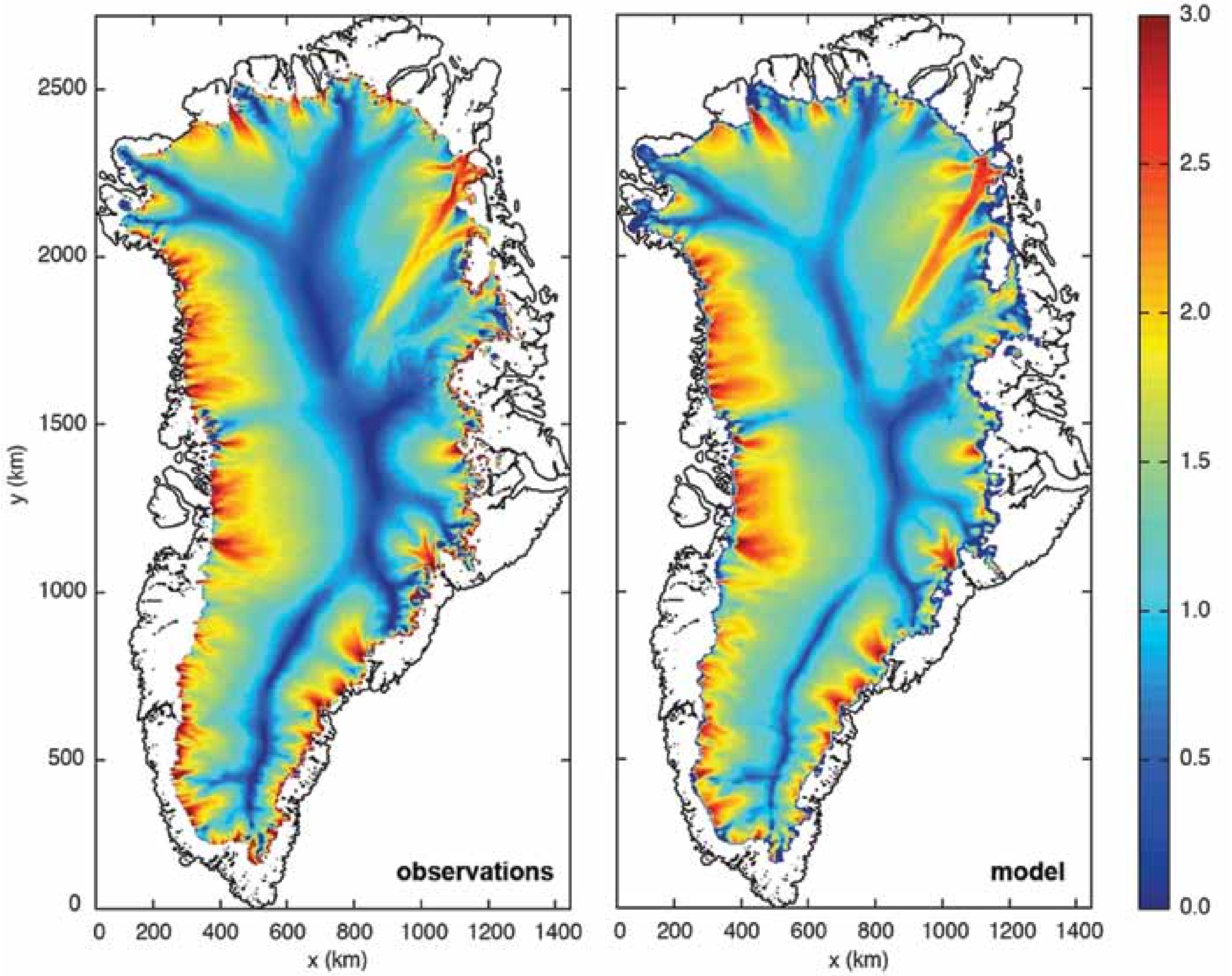
Evans said the team is on track to have the code running massively parallel by October of 2012. Currently, simulations of a small test problem have employed 1,600 of Jaguar's 299,008 processors. Evans said the team expects that number to expand substantially in the near future when they begin simulating larger problems with greater realism.

The CESM began as the Community Climate Model in 1983 at the National Center for Atmospheric Research (NCAR) as a means to model the atmosphere computationally. In 1994, NCAR scientists pitched to the National Science Foundation (NSF) the idea of expanding their model to include realistic simulations of other components of the climate system. The result was the Climate System Model—adding land, ocean, and sea ice component models—which was renamed the Community Climate System Model (CCSM) to recognize the many contributors to the project. Development of the CCSM also benefitted from DOE and National Aeronautics and Space Administration (NASA) expertise and resources.

The CCSM developed into the CESM as its complexity increased. Today the model is a computational collection of the Earth's oceans, atmosphere, and land, as well as ice covering land and sea. Its components calculate in chorus. "The model is about getting a higher level of detail, improving our accuracy, and decreasing the uncertainty in our estimates of future changes," said Los Alamos National Laboratory climate scientist Phil Jones, who leads that laboratory's Climate, Ocean, and Sea Ice Modeling group. The group develops the first-principles ocean, sea ice, and ice sheet components of CESM. Its members are interested in sea-level rise, high-latitude climate, and changes in ocean thermohaline circulation—aqueous transport of heat and minerals around the globe.

The team is changing its ocean code, the Parallel Ocean Program, to become the Model for Prediction Across Scales–Ocean (MPAS–Ocean) code. Unlike the Parallel Ocean Program, MPAS-Ocean is a variable-resolution, unstructured grid model. It will allow researchers to sharpen simulation resolution on a regional scale when they want to look at climate impacts in particular localities.

The CESM has continually grown in intricacy, enabling researchers to calculate more detail over larger spatial scales and longer time scales. Further, researchers are able to introduce more complex physics variables and simulate in greater detail Earth's biogeochemical components—chemical and ecosystem impacts on the climate. But these advances do not come without a price.



The SEACISM project will improve the fidelity of ice-sheet models. The simulated (right) versus observed (left) flow velocities for Greenland ice sheets match well. Simulation data from Steve Price, Los Alamos National Laboratory; observational data from Jonathan Bamber and colleagues in Journal of Glaciology, ***46***, 2000

"At any given time when we're integrating the model forward, it's very [computationally] expensive—very time consuming and using large amounts of memory," Evans said. "It all has to work in concert to generate huge amounts of data that we then need to analyze afterward."

Conducting climate codes

Researchers and code developers for the CESM are scattered around the United States. One of their biggest challenges is tying together separate climate code components created in different places on different computer architectures. What's more, climate researchers usually focus on a specific aspect of the climate, such as ocean or atmosphere. An atmosphere scientist, for example, needs the ability to raise resolution in the atmosphere but may want to lower the resolution used in the ocean to minimize the computational cost of the simulation. ORNL computer scientist Patrick Worley helps researchers optimize their codes. That makes him one in a small group in the CESM community who conduct the climate code orchestra.

"There are many scientific issues with getting simulations right, and the computer scientists are involved with helping the scientists test and optimize them," said Worley. He serves as a co-chair in the CESM software engineering working group, which is dedicated to solving the unique challenges climate research imposes on computing resources.

A number of computational issues distinguish climate science from other scientific disciplines that make heavy use of simulations, Worley said. First, if researchers want to change their problem size, they must rework the simulations' new physical processes in correspondingly increased or decreased resolution. Second, many researchers focus only on particular areas of the Earth during their simulations, meaning they need only a particular part of the CESM framework running in high resolution. Finally, climate simulations run at varying time scales, sometimes spanning several thousand years. The required time to solution and available computing resources can force researchers to choose between high spatial resolution in their models or an extended observation period.

> By continuing to improve science representations and numerical methods in simulations, and exploiting modern computer architectures, researchers expect to further improve the CESM's accuracy in predicting climate changes.

According to Evans, computer scientists like Worley help climate scientists deal with that complexity in the coupled model. "Pat does bridge across many components. He can look at an ice code, which solves very different equations and is structured in a very different way than an atmospheric model, and be able to help us run both systems not only individually at their maximum ability but coupled together," she said. "There are few in the climate community that have an understanding of how all of it works."

A hybrid horizon

With software improvements and increasingly more powerful supercomputers, resolution and realism in climate simulations have reached new heights. But there is still work to be done. "What we can't do yet in the current model is get down to regional spatial scales so we can tell people what specific impacts are going to happen locally," said Jones. "The current IPCC simulations are at coarse enough resolution that we can only give people general trends."

Climate research that concluded in 2010 makes up the final pieces of information that will go into the next IPCC Assessment Report, due for release in 2013. Meanwhile, climate researchers are preparing for the future. "The climate research community doesn't have a single climate center where they run all of their climate simulations," Worley said. "They run wherever there is supercomputing time available. So it's important for the codes to run on as many different platforms as possible." Currently, US climate codes are running on National Oceanic and Atmospheric Administration, NASA, DOE, and NSF supercomputing resources.

One of the biggest challenges facing climate research, according to Worley, is writing the computational score for hybrid architectures. Next-generation supercomputers, such as the OLCF's Titan, a 20 petaflop machine, will use both CPUs and GPUs to share the computational workload. This novel approach will require closer attention to all levels of parallelism and will alter the approach to computing climate. "There are a number of people that want to make sure we are not surprised by the new machines," Worley said. If all goes well, CESM researchers may hear calls for an encore.—*by Eric Gedenk*

Supernovas Explode in 3D Detail at ORNL

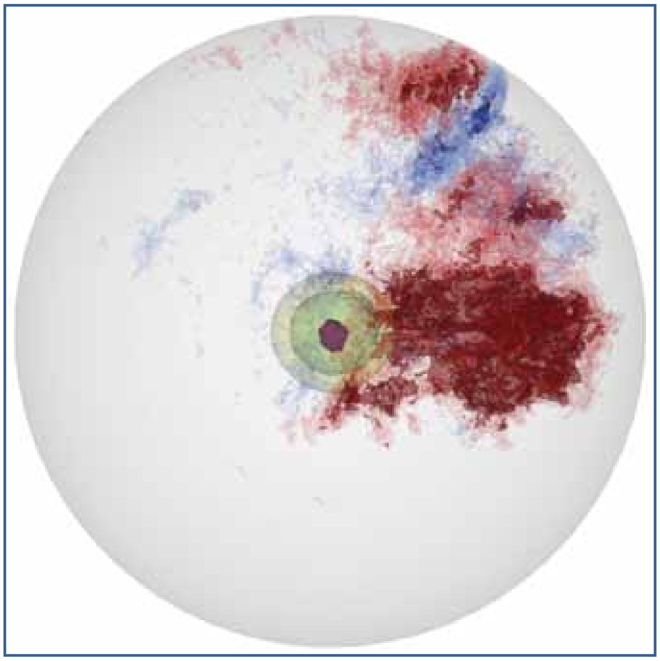
Simulations of standard candles will help astrophysicists understand dark energy

Understanding a Type Ia supernova—an exploding white dwarf star—requires supercomputers. A team of astrophysicists and computational scientists is using the power of Jaguar, a Cray XT5 high-performance computer at ORNL, to virtually blow up these white dwarfs. In the process the researchers are revealing the secrets of the biggest thermonuclear explosions in the universe and finding the answers needed to measure the size of the universe.

"The physics of supernova explosions is something astrophysicists have been trying to figure out for about 50 years now," said Stan Woosley, principal investigator of an INCITE project and professor of astrophysics at the University of California–Santa Cruz. "It is an interesting physics problem in turbulent combustion, but it is also important because—as the 2011 Nobel Prize attested—Type Ia supernovae can be used to show that the expansion of the universe is accelerating." This is because the supernovas function as "standard candles," brilliant lights of known properties usable as measuring tools because their distance can be inferred by how bright they appear.

Using an empirical correlation between the peak luminosity of Type Ia supernovas and the rate at which their brightness declines, three Nobel Prize winners—Saul Perlmutter, Adam Reiss, and Brian Schmidt—determined that distant galaxies were farther away than expected for a universe expanding at a constant rate. Their work showed that the expansion of the universe is, in fact, accelerating and refuted a long-held belief that the universe was expanding at a constant rate or even slowing down. It is as if a tennis ball, thrown straight up, did not fall back to Earth but continued to move away faster and faster. The driver behind this acceleration, called dark energy, is a force not yet fully understood that counters the pull of gravity on matter.

Woosley said his team seeks to understand the nature of the supernovas, find other ways of predicting a supernova's peak luminosity, and explain the empirical relation between brightness and decline rate (the ratio of maximum light declining over time). Even with the world's most powerful telescopes, the data obtained through observations is limited and sometimes corrupted by the effects of interstellar dust or the particular characteristics of the telescopes used. Given a valid physical description, the supercomputer can model the supernova without these observational biases. The simulations not only give the supernova's light curve (the total intensity of light over time) and spectrum (the intensity of light at certain frequencies) at all angles over time, but also allow researchers to look for other measurable quantities that might correlate with peak luminosity.



In the inner 1,000 kilometers of a 1,700-kilometer-diameter white dwarf star, color changes show the progression rate of nuclear energy generation. The simulation, by the MAESTRO code on the Jaguar supercomputer, indicates the ignition will happen off-center (red). Image courtesy M. Zingale, Stony Brook University, and A. Nonaka, Lawrence Berkeley National Laboratory

Presently cosmologists use Type Ia supernovas to measure distances to about 10 percent accuracy. The immediate goal is to reduce the error below 5 percent. To do that, they need a better yardstick, and numerical experiments on the supercomputer are a way to build one.

A second life

A white dwarf star begins, as do all stars, when gravity compresses a cloud of helium, hydrogen, and dust. As the density increases, so does the force of gravity. The temperature and density are highest at the center of the cloud. The pressure of the compaction increases the internal temperature. As the pressure and temperature climb, the compaction and heat set off the star's core fusion reaction. The star begins to glow. After shining 10 billion years or so (for a star of our sun's mass), its fuel is spent and the star—cooling, fading, and dwindling to the size of Earth—becomes a white dwarf.

With its own energy depleted, the star will continue cooling into a cold, black dwarf unless it can accrete mass from a nearby star.

The white dwarf can accumulate mass in two ways. It can steal mass slowly, from a younger, vibrant companion star still undergoing nuclear fusion, until the dwarf reaches its maximum mass while remaining stable—a relationship known as the Chandrasekhar limit. Or, in what is known as the double-degenerate model, two white dwarves can merge their masses into one.

The increasing mass adds to the star's density and internal pressure, and its core temperature climbs. After centuries the star reaches critical mass. Its carbon and oxygen ignite in a fusion reaction deep inside, and a thermonuclear runaway begins. For roughly a century the dwarf interior bubbles and churns as convection carries away the energy generated by carbon fusion. Hot gases rise, cool, drift back to the reaction zone, and are reheated. Eventually convection cannot carry away the heat fast enough. The burning becomes isolated and rapidly escalates. Temperature rises in a small region from hundreds of millions of degrees to ten billion degrees. A deflagration—a moving flame front of burning gases—is born.

At this point the star has reached full runaway mode and will soon explode. In the second before the final runaway, ashes rapidly rise like a giant mushroom cloud. The floating deflagration approaches the speed of sound as it rips through the surface of the star. Simulations suggest that this breakout accelerates the subsonic burning into a supersonic detonation. The burning happens in a strong shock wave that rebounds to the unburned areas, completely incinerating the star. It then explodes violently, leaving behind rapidly expanding radioactive gases. An isotope of nickel, which decays first to cobalt and then to iron, keeps the debris hot and as bright as a medium-sized galaxy before fading away over months. The star's life is over.

Replicating supernovas in the lab

Previous simulations in two dimensions by Woosley and Daniel Kasen from the University of California–Berkeley and Lawrence Berkeley National Laboratory (LBNL) have explained the observed asymmetrical explosions of the Type Ia.

"We found the most popular kind of supernova model—the Chandrasekhar mass model—doesn't ignite in the dead center of the star," said Woosley. "The explosion, once ignited off-center, stayed off-center and the supernova blew up asymmetrically, one side before the other." This knowledge explains the differences in the appearance of supernovas when viewed from various directions and must be considered when calibrating the standard candle.

Jaguar is now running three-dimensional (3D) simulations with varying ignition-point locations, focused on the two-stage, deflagration-to-detonation sequence.

Because of the size of the computational runs, the studies are decoupled and done in three successive stages—ignition, explosion, and supernova. The team is using three codes developed for this project: MAESTRO, CASTRO, and SEDONA. John Bell and Ann Almgren at LBNL wrote MAESTRO to model the ignition and CASTRO to model the explosion. SEDONA, written by Kasen, is a hydrodynamics code that calculates the spectra and light curves of the supernovas.

The multiyear project was allocated 50 million computing hours on Jaguar in 2011 and 47 million in 2012 through the INCITE program, which is jointly managed by DOE's Argonne and ORNL Leadership Computing Facilities.

Additionally, Woosley has begun investigating another variety of Type Ia supernova—a white dwarf that accretes a thick layer of helium from a companion star. This helium-rich celestial object may detonate in a straightforward blast after an ignition in the star's outer layer compresses its inner core.

After the simulations at ORNL blow up several 3D models of white dwarf stars and their light curves and spectra have been determined, the findings will be put into an archive available to researchers. Observers will especially analyze the data to find other, more accurate indicators of luminosity to compare with a rapidly growing library of observational data. The data will also be used to help plan future large surveys of supernovas by missions such as the Large Synoptic Survey Telescope and the Palomar Transient Factory. Said Woosley, "It is our eventual ambition to have hundreds of model supernovae sampled at all wavelengths, angles, and times available for comparison by machine intelligence with thousands of observations in order to get the very best distance measurements possible."—*by Sandra Allen McLean*

Materials Modeling Shows Big Future for Boron Nitride Nanoribbons

Simulation on supercomputer predicts new desirable properties

ORNL's Alejandro Lopez-Bezanilla was smiling as Konstantin Novoselov, the 2010 Nobel Prize co-winner in physics, lectured at this year's American Physics Society meeting.

Novoselov's topic was graphene, the current darling of the nanoelectronic world. The reason for Lopez-Bezanilla's smile? He was already months deep into a computational study of the material being proposed as a graphene substrate: the compound boron nitride.

Graphene, which is carbon in the form of freestanding one-atom-thick sheets, is a natural for next-generation computer chips, communications equipment, and solar energy devices. Electrons flow through the material at an astonishing 1 million meters per second.

To live up to its potential, however, graphene needs support. On its own, its edges wrinkle, tear, or roll up. The silicon dioxide substrate used for today's microchips is not a good partner for graphene; it creates vibrations that slow the electrons, and its surface is too bumpy.

An ideal substrate would not physically interfere with the graphene and would have a low electrical resistance. "The substrate has to be a dielectric material, a material which is insulating and that can be polarized [positive and negative charges grouped in two distinct regions] by an applied electrical field," says Lopez-Bezanilla, a research associate in ORNL's Computing and Computational Sciences Directorate. He is funded by the Petascale Initiatives program of DOE ASCR.

Any substrate will affect the electrons in the adjacent graphene layer, but boron nitride interferes less than silicon dioxide. Also, boron nitride resists chemical change and is unaffected by high temperatures, leading many researchers to believe that it could be that consummate base.

"Boron nitride is a covalent material with atoms tightly bonded to each other, but it also presents a strong ionic behavior," making it a great insulator and poor conductor, explains Lopez-Bezanilla. With the help of computer simulations run on Jaguar, ORNL's petascale supercomputer, Lopez-Bezanilla took a closer look—a nanoscale look—at boron nitride's properties, and discovered boron-nitride monolayers are an ideal dielectric substrate material for graphene nanodevices.

Boron nitride, like graphene, can be formed as one-atom thick sheets or as nanotubes, then cut into nanoribbons with their atoms arranged in a hexagonal lattice (imagine chicken wire). After a boron nitride nanotube is cut, the resulting edge shape will affect the behavior of the nanoribbon. A slice straight through the lattice halves every other hexagon, making armchair-shaped edges (aBNNR); zigzagged (zBNNR) edges are made by cuts along the hexagonal borders. The zigzag-cut ribbon has one edge lined with nitrogen atoms and the other with boron atoms. A boron nitride nanoribbon (BNNR) is a quasi-one-dimensional structure, one atom thick, with a width-to-length ratio of 1 to 1,000-plus.

The size of a nanoribbon can be compared with a strand of spider's silk, which is one thousand times thicker than one BNNR. If the BNNR were scaled up to the thickness of the spider silk filament, a correspondingly scaled silk strand would be as wide as a coin's edge.

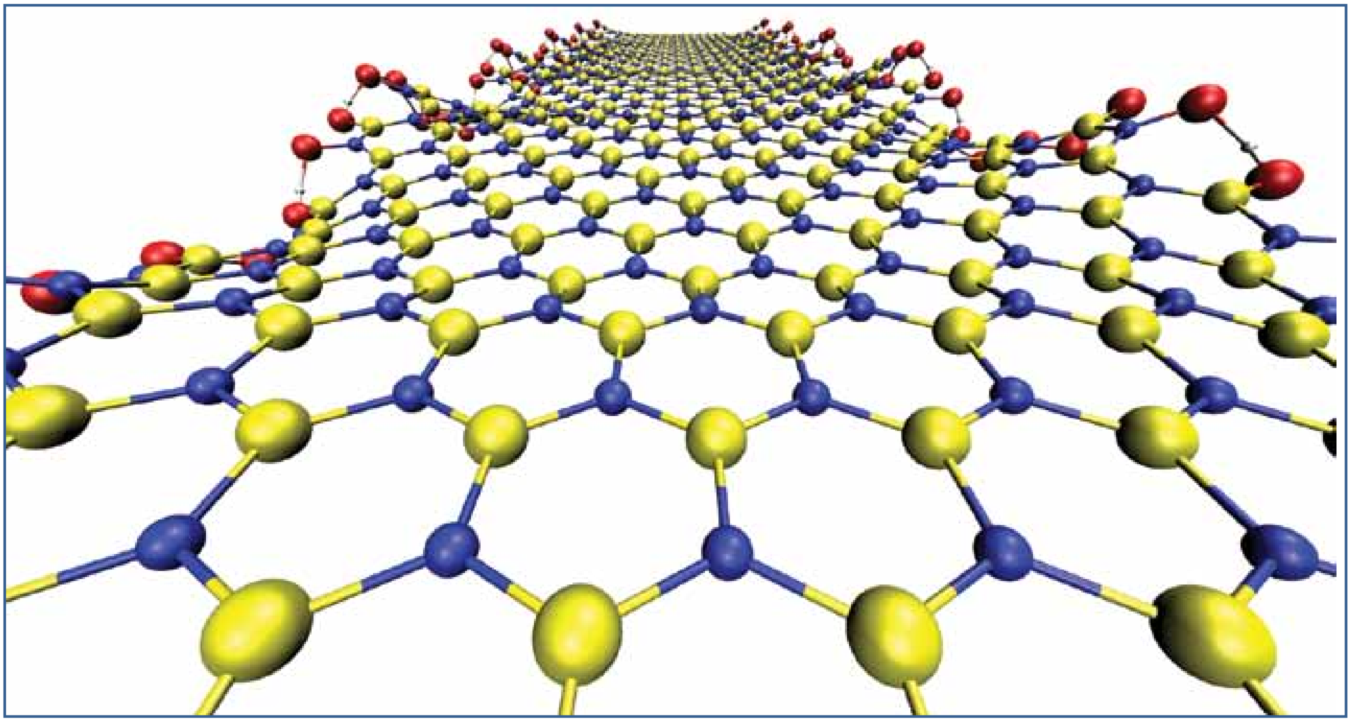
Changing boron nitride's properties

The judicious application of elements to a material—a process known as functionalization—can change the material properties at the nanoscale. Lopez-Bezanilla pushed boron nitride's boundaries by modeling the compound under different conditions to see what new properties it might have at the nanoscale, expanding its usefulness. Working with colleagues at ORNL, he simulated zBNNRs modified by oxygen or sulfur atoms attached to each boron and nitrogen atom along the cut edges. Under these conditions, the insulating boron nitride becomes metallic.

The team used two systems for the simulation. Jaguar ran the Vienna Ab-initio Simulation Package (VASP) and the ORNL Institutional Cluster ran the SIESTA code. Both are density functional theory (DFT)–based codes.

"VASP needs the outstanding computational capabilities of Jaguar," said Lopez-Bezanilla. "To verify the metallic features of the oxygenated zBNNRs with a higher degree of accuracy, we performed spin-polarized calculations with the help of VASP on large unit cells, which needed thousands of processors due to the extensive calculation of large systems at the level of a DFT code based on plane waves."

Typical job sizes were 2,048 cores. "The project consisted of doing a large number of calculations at different levels of detail," said team member Bobby G. Sumpter, group leader of Computational Chemical and Materials Science and the Nanomaterials Theory Institute at ORNL. "Jaguar was needed to perform certain calculations that used a different level of detail for the electronic structure—in particular to check system



The computational models run on the Jaguar supercomputer showed that the edge shape, the choice of added elements, and the location of the elements all play a role in creating different behaviors in boron nitride nanoribbons. A boron nitride nanoribbon with armchair edges adopts an undulating shape when the edges are terminated with hydroxyl groups (oxygen shown in red, hydrogen in white, boron in yellow, and nitrogen, blue). Image courtesy Alejandro Lopez-Bezanilla

size effects and for more rigorous treatment of electron exchange and spin, but also to examine the use of more complete basis sets. Basically Jaguar provides the required computational engine that enables us to do a complete study, without which we would not have been able to fully develop an understanding of the boron nitride nanoribbons within months."

The computational models showed that the edge shape; the choice of added elements—oxygen, sulfur, or hydrogen; and the location of the elements all play a role in creating different behaviors in the nanoribbons.

The repetitive zigzag edge shape sets the added oxygen atoms at the same position—always parallel to the edge and equidistant from the adjacent oxygen atoms, so that the electron shells overlap, allowing electrons to move between atoms. In other words, it gives the nanoribbon metallic qualities. If BNNRs with armchair edges are modified, the oxygen atoms are arranged differently, preventing the emergence of metallic characteristics.

Both oxygen and sulfur render the zBNNRS metallic, whereas the control group of hydrogen-edged zBNNRs was semiconducting and nonmagnetic.

The placement of the elements also affects the metallic qualities and conductivity. "To show the importance of the edges [in converting] the zBNNRs into a metallic compound, we modeled the doping of the ribbon when the oxygen atoms form an extended wire embedded in the center of the ribbon," says Lopez-Bezanilla. "The result is that it conserves its insulating features."

The team outlined its findings in a July 7, 2011, paper in the American Chemical Society journal *Nano Letters*.

The simulation results have not yet been tested in a laboratory.

"So I encourage experimentalists: take zBNNR, put oxygen on the edges, and test what I said. We encourage chemists to do that, and if it is eventually synthesized, we will have some nanoelectronic applications," says Lopez-Bezanilla.

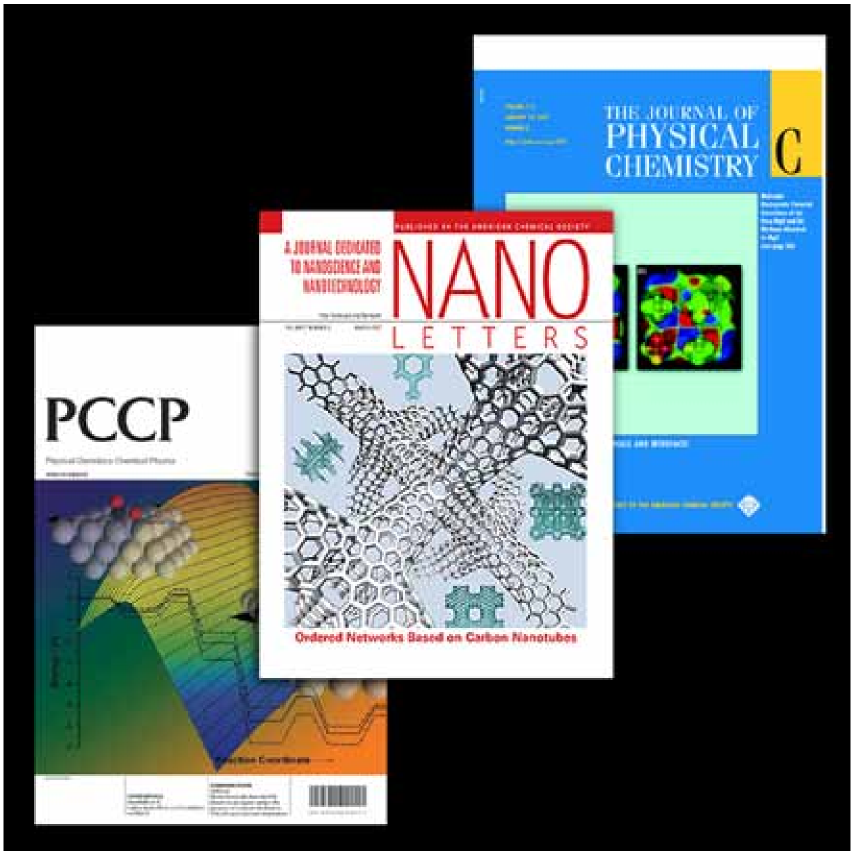
Simulating the transformation of a known material into one with novel features opens the door to possible new uses—optical, magnetic, electronic. There is a new world of opportunities for the features in zBNNRs. Reproducing Lopez-Bezanilla's results in the lab will put a smile on many faces.—*by Sandra Allen McLean*

Chemistry and Materials Simulations Speed Clean Energy Production and Storage

Researchers solve electronic structures to explore biomass, supercapacitors, and more

Enterprises from energy production to environmental cleanup depend on chemistry. Catalysts, which make chemical reactions more likely, contribute to nearly 20 percent of all industrial products. In the United States, three industries relying on catalysts—petroleum, chemicals, and pharmaceuticals—account for $1 trillion of the gross national product.

Catalysts are just one area of investigation for a multi-institutional team whose 70 publications in 3 years detail prodigious scientific output from the world's fastest chemistry simulations. Its work integrates experiment, theory, and simulation to explore a continuum from chemistry to materials science. The researchers aim to understand, control, and design technologies needed for clean energy.



Chemistry and materials science simulations made possible through Robert Harrison's INCITE project have generated 70 scientific publications since 2008, many of which graced the covers of prestigious journals and dealt with topics from production of hydrogen for clean energy to development of graphene nanoribbons for power delivery. Collage courtesy Robert Harrison and Jason Smith

"Our long-term goal is enabling the design of new generations of clean and sustainable technologies to produce, transmit, and store energy," said team leader Robert Harrison, a computational chemist at ORNL and the University of Tennessee who directs the Joint Institute for Computational Sciences, a partnership between the two organizations. "Key elements are a fundamental understanding of chemical and electronic processes at the atomic scale and ultimately effectively transferring this knowledge in partnership with experiment into the hands of people and industries interested in the next steps of R&D."

The research closely aligns with DOE's missions and its 20-year priority to support nanoscale science for new materials and processes. Two DOE Office of Science programs, Basic Energy Sciences and ASCR, fund the work.

Through the INCITE program, the researchers have been awarded more than 100 million processor hours since 2008. At the OLCF, they calculate the electronic structures of large molecules and surfaces using scientific application codes called NWChem and MADNESS. The findings inform the development of processes, such as biomass conversion and fuel combustion, and products, such as batteries, fuel cells, and capacitors.

The electronic structure allows scientists to determine the positions and binding energies of atoms within molecules and responses to perturbations. Petascale computers speed complex calculations of molecular dynamics and quantum mechanics as substances undergo chemical transformation.

"Some of the largest calculations are only feasible on the leadership computers, not just because of speedy processors, but because of other architectural features—the amount of memory, the amount and speed of the disks, the speed and other characteristics of the interprocessor communication," Harrison said.

A scientific paper by team member Edoardo Apra of ORNL was a 2009 Gordon Bell finalist after he and co-authors reported scaling NWChem to use most of Jaguar, ORNL's Cray XT5 supercomputer, to calculate the electronic structure of water clusters, which is important in chemistry at interfaces and nucleation in the atmosphere. Because NWChem is a flagship application used throughout the community, getting it to run at petascale will have high impact. That was a special challenge, though, because the code employs distributed, shared memory instead of the message passing used by most codes. As a result of the team's efforts, NWChem has now joined four other scientific applications sustaining more than a petaflop on Jaguar.

Further feats of computational science

Chemistry and materials science are critical for innovation. ORNL theorist Bobby Sumpter with lab colleagues Vincent Meunier and Jingsong Huang ran calculations of hundreds of teraflops on Jaguar to investigate a next-generation capacitor, a device that stores energy through charge separation at an electric double layer formed within porous materials. As transportation systems are electrified, fast charging of vehicles with such devices will be necessary. Iterating between simulation and experiment to reduce intermediate models, within 2 years Sumpter's team and collaborators at Rice University arrived at a practical device for high-power energy delivery. Their supercapacitor stores—and quickly discharges—a thousand times more energy than a conventional capacitor. Several challenges remain to be overcome, however, before a commercial device can be developed and deployed.

> Our long-term goal is enabling the design of new generations of clean and sustainable technologies to produce, transmit, and store energy.

"Because there's very little chemical or material change, supercapacitors can be cycled millions of times without significant degradation. In contrast in a battery you're physically moving considerable matter around and inducing chemical change, so it's much harder to cycle more than a few thousand times," Harrison said. "We're familiar with the batteries in our laptops and cell phones dying after a year or two."

While capacitors don't seem chemically active, in fact a lot of chemistry is going on that's relevant to energy storage. Indeed, Sumpter and collaborators achieved high energy densities in supercapacitors by coupling chemistry with nanomaterials—specifically, graphene nanoribbons structured at the billionth-of-a-meter scale of atoms. They considered the smallest unit of structure needed to achieve an energy-storage function and modeled a supercapacitor as a carbon pore. An ion and the shell of solvent surrounding it can move into the pore. But in materials with pores too small for the entire solvation shell to enter, some solvent gets stripped off. The ions align into a nanowire structure as they pack into the pore. The pore size and the ion's tiny diameter determine the high capacitance.

Whereas Sumpter and collaborators simulate energy storage, other members of Harrison's team address energy production. "When growing corn for conversion into ethanol, a lot of lignin and cellulosic material is left over," Harrison said. "We could get energy out of it, and we could turn it into other useful chemicals—if only we had controllable, efficient processes."

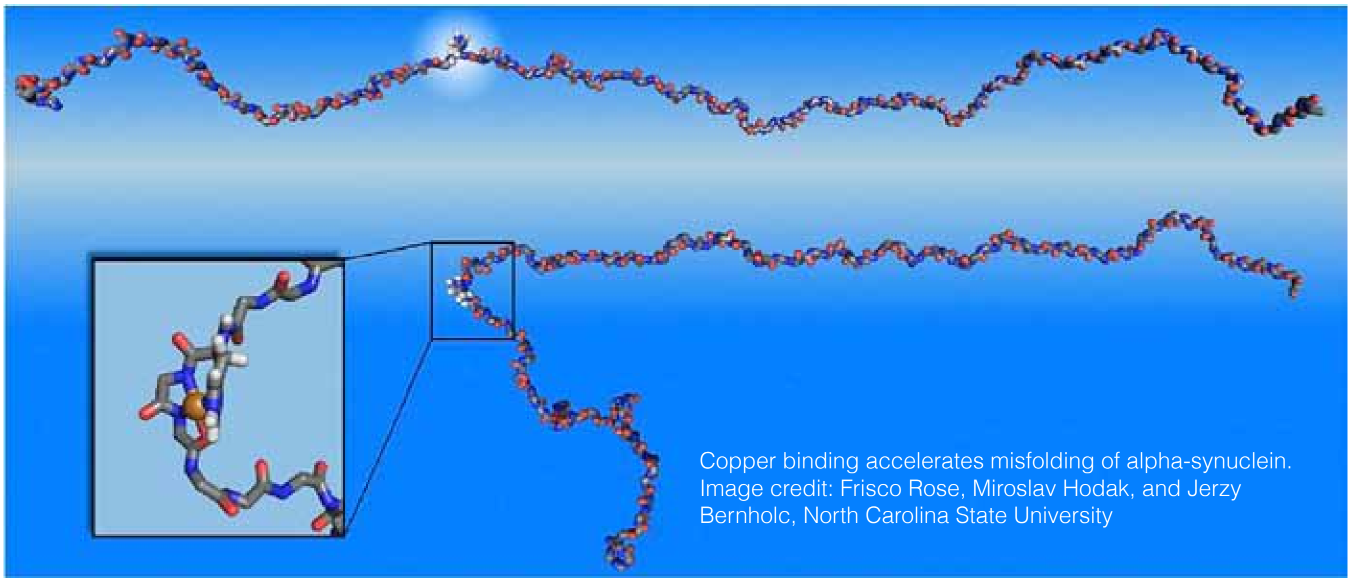
To improve those processes, computational chemist Ariana Beste and experimental chemist A. C. Buchanan, both of ORNL, explore thermochemical degradation of plant materials. They study how molecular structures influence networks of chemical reactions. The rate of a reaction depends on the height of energy barriers along paths between reactants and products and the fraction of molecules with enough energy to hurdle those barriers. One chemical reaction may lead to half a dozen products. Favoring a path that results in a specific product may necessitate understanding a hundred reaction paths.

Petascale simulations can quickly calculate the proportion of molecules with the requisites for a specific reaction—a herculean statistical challenge. Calculating which bonds between atoms in a molecule have the lowest energies, for example, reveals the optimal shape for a molecule to assume. That knowledge can speed design of processes faster than do trial and error or expert insight.

"Chemistry and materials sciences are sciences about the real world, which is complicated and messy, hard to characterize and control," Harrison said. "We're never going to be in a position to say we can replace chemistry with simulation. What we can say, confidently now, is that in increasingly large areas of our discipline, simulation and theory are equal and valued partners with experiment. Petascale computing is critical to accelerating scientific advancements."—*by Dawn Levy*

Researchers Pinpoint How Copper Folds Protein into Precursors of Parkinson's Plaques

Calculations on the Jaguar supercomputer reveal metal binding likely to misfold alpha-synuclein



Copper binding accelerates misfolding of alpha-synuclein. Image credit: Frisco Rose, Miroslav Hodak, and Jerzy Bernholc, North Carolina State University

Aided by the Jaguar supercomputer at ORNL, researchers at North Carolina State University have figured out how copper induces misfolding in the protein associated with Parkinson's disease, leading to creation of the fibrillar plaques that characterize the disease. This finding has implications both for the study of Parkinson's progression and for future treatments.

The protein in question, alpha-synuclein, is the major component of fibrillar plaques found in Parkinson's patients. Researchers had already discovered that certain metals, including copper, could increase the rate of misfolding by binding with the protein, but they were unsure of the mechanism by which this binding took place.

"We knew that the copper was interacting with a certain section of the protein, but we didn't have a model for what was happening on the atomic level," said Frisco Rose, Ph.D. candidate in physics and lead author of a paper in the June 14, 2011, issue of *Nature Scientific Reports*, an online journal of the Nature Publishing Group.

"Think of a huge swing set, with kids all swinging and holding hands—that's the protein," he continued. "Copper is a kid who wants a swing. There are a number of ways that copper could grab a swing, or bind to the protein, and each of those ways would affect all of the other kids on the swing set differently. We wanted to find the specific binding process that leads to misfolding."

Rose and North Carolina State colleagues Miroslav Hodak, research assistant professor of physics, and Jerzy Bernholc, Drexel Professor of Physics and Director of the Center for High Performance Simulation, developed a series of computer simulations designed to ferret out the most likely binding scenario.

According to Hodak, "We simulated the interactions of hundreds of thousands of atoms, which required multiple hundred thousand CPU-hour runs to study the onset of misfolding and the dynamics of the partially misfolded structures."

The number of calculations was so large that Hodak and Bernholc had to devise a new method to make it possible for a computer to process them. Only supercomputers like Jaguar, ORNL's most powerful supercomputer, were up to the task. But the simulations finally revealed the binding configuration most likely to result in misfolding.

The researchers hope that their finding will advance our understanding of Parkinson's, one of the most common—and devastating—neurological diseases. "Understanding the molecular mechanism of Parkinson's disease should help researchers in developing drugs that treat the disease rather than merely alleviate symptoms," Bernholc said.

DOE and NSF funded the research. The INCITE program, jointly managed by the Argonne and Oak Ridge Leadership Computing Facilities, provided the supercomputing time allocation on Jaguar.—*by Tracey Peake, North Carolina State University*

Special Report Highlights Research at America's Leadership Computing Facilities

Supercomputers accelerate science and engineering breakthroughs

INCITE | IN REVIEW

Jaguar Supercomputer Performs Most Advanced Earthquake Simulation

INCITE Accelerates Research Breakthroughs with Leadership Computing

Modeling Turbulent Combustion Speeds Design of Power and Propulsion Devices

A special report highlights the accomplishments of researchers running large, complex, and often unprecedented simulations on DOE Office of Science supercomputers. The research community gains access to these powerful HPC systems through the INCITE program jointly managed by the Argonne Leadership Computing Facility (ALCF) at Argonne National Laboratory and the OLCF at Oak Ridge National Laboratory. The 64-page report, *INCITE in Review*, is available at http://www.doeleadershipcomputing.org/wp-content/uploads/2011/07/INCITE_IR_FINAL_7-19-11.pdf.

"Whether it's gaining a better understanding of the universe or engineering solutions for a growing, energy-hungry populace, the investigations enabled by INCITE are making the world a smarter, more sustainable place," James J. Hack, director of the NCCS, which houses the OLCF; Michael E. Papka, director of the ALCF; and Julia C. White, INCITE program manager, wrote in the report's introduction. "We are proud to tell the story of INCITE's role to date in advancing the frontiers of human knowledge for the benefit of all."

The report highlights 22 select projects, including simulations to make personalized genomics quick and affordable; explore corrosion in building materials; model carbon sequestration underground; and develop next-generation catalysts, semiconductors, and nuclear reactors for fission and fusion. Some projects elucidate biomass conversion for energy production and supercapacitors for energy storage, whereas others contribute to a deeper understanding of phenomena from earthquakes to supernovas.

Since 2003 INCITE has promoted transformational advances in science and technology through large allocations of computer time, supporting resources, and data storage. From modeling hurricanes that put people and property at risk, to simulating combustion instabilities in power plant turbines and vehicle engines, INCITE projects aim to accelerate breakthroughs in fields in which major advancements would not be probable or even possible without supercomputing.

With a modest start by today's standards, in 2004 three INCITE projects received a total of nearly 5 million processor hours from LBNL's National Energy Research Scientific Computing Center. Since 2008 the program has focused on leadership computing facilities, from which researchers can obtain the largest single-award time allocations available on powerful computing systems, including the OLCF's Cray XT5 (Jaguar) with 299,008 processing cores yielding a peak performance of 3.3 thousand trillion calculations each second and the ALCF's IBM Blue Gene/P (Intrepid) with 163,840 processing cores yielding a peak performance of 557 trillion calculations per second.

For the 2011 calendar year, 57 INCITE awardees received a total of 1.7 billion processor hours. The allocations averaged 27 million hours, with one project receiving more than 110 million hours. From INCITE's inception through the end of 2011, researchers from academia, government laboratories, and industry will have been allotted more than 4.5 billion processor hours to speed innovations and discoveries.—*by Dawn Levy*

20th Year of Operation

1992

The DOE Office of Science established the Center for Computational Sciences at ORNL in 1992. Participating in the ribbon-cutting ceremony for the opening of the center are (L to R): Jim Glimm, State University of New York at Stony Brook; Larry Dowdy, Vanderbilt University; Jack Dongarra, University of Tennessee—Knoxville; Justin Rattner, Intel; Dave Nelson, DOE-HQ; Bill Appleton, ORNL; Clyde Hopkins, Martin Marietta Energy Systems; Joe LaGrone, DOE-OR; Al Trivelpiece, ORNL Director; Jim Decker, DOE-HQ; John Marburger, State University of New York at Stony Brook; Billy Stair, State of Tennessee; and Ed Masi, Intel. Image credit: ORNL

Q&A: Arthur Bland, Project Director, OLCF

ORNL's Arthur (Buddy) Bland directs the Titan project to build a hybrid supercomputer with a peak performance of at least 20 petaflops (a thousand trillion calculations per second) at the OLCF. Titan, a resource of the DOE Office of Science, will bring unprecedented computational power to solve pressing problems in science and engineering. Prior to leading Titan's development, Bland headed deployment of Jaguar, a supercomputer powered by CPUs that underwent staged upgrades to reach 3.3 petaflops. Jaguar was #1 on the TOP500 list of world's fastest supercomputers from November 2009 to November 2010. Testament to its importance in research is a 2012 cover of the journal *Science,* which featured an OLCF simulation of a core-collapse supernova explosion. GPUs—energy-efficient accelerators of specific types of calculations in scientific application codes—are currently being added to Jaguar to transform it into Titan, a GPU/CPU hybrid, in fiscal year 2013.

After earning bachelor's and master's degrees in computer science from the University of Southern Mississippi and serving as a captain in the Air Force, Bland joined ORNL in 1984 as the system programmer and administrator for the Cray X-MP system. In 1992 he transferred to ORNL's newly established Center for Computational Sciences, where he managed the deployment and operation of computing resources including the Kendall Square KSR-1; Intel Paragons; IBM Power3 and Power4; Compaq AlphaServer SC; SGI Altix; and Cray XD1, X1, X1E, XT3, XT4, XT5, and XK6. In 2004 the OLCF was established to deliver a supercomputer 1,000 times more powerful than the leading systems of the day, and Bland managed the development and installation of its Cray XT5 Jaguar computer networks and storage systems. The OLCF user facility that Bland manages gives the worldwide scientific community access to capability computing, which solves in the shortest time possible problems of a size or complexity that few or no other computers can approach.

In this interview Bland recalls 20 years of leadership computing. He and other pioneers are paving the way to exaflop computing—or execution of a million trillion calculations per second—to provide a resource up to addressing the considerable research challenges that affect us all, from developing sustainable energy to understanding climate change.

Arthur Bland, OLCF Project Director

What did ORNL's computing landscape look like before 1992, when the lab won the Partnership in Computational Science proposal competition that planted a high-performance computing center amidst the Appalachian foothills?

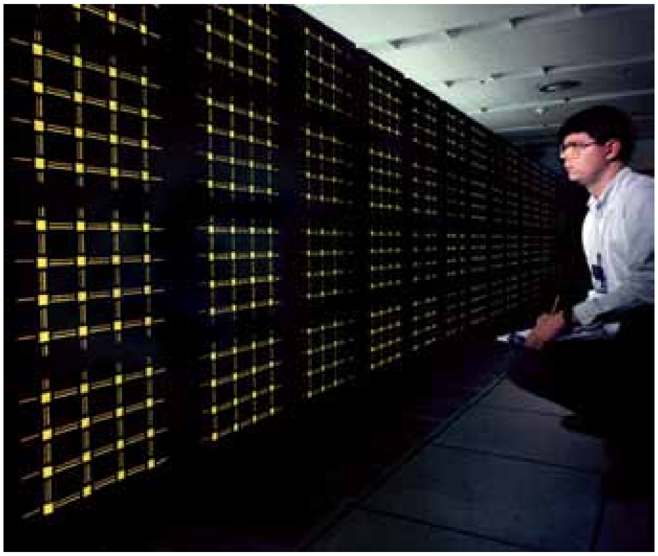
I came to the laboratory in 1984. At that time we had several machines from Digital Equipment Corporation and IBM that were used for scientific calculations. Seymour Cray's company, Cray Research, was building the world's fastest computers—the Cray-1 and then after that the Cray X-MP. I came into the laboratory because I had experience with the Cray-1 from my time in the Air Force. I was hired to help set up a Cray X-MP that the lab had bought. In those days a single contractor ran all three DOE facilities in Oak Ridge, so there was a single computing division for ORNL, K-25 [the current East Tennessee Technology Park], and the Y-12 National Security Complex. ORNL's science programs partnered with a uranium isotope separation program at K-25 to purchase a Cray X-MP system in 1985. It was installed at K-25 because it needed to do classified work for the uranium enrichment program, but it was also doing unclassified work for ORNL. The Cray X-MP and a five-processor DEC PDP-10 system were the major scientific computers at ORNL. By 1992 the Cray system had gotten to the end of its life. At the time there was a new set of computers coming out from IBM called the RS/6000, which used the RISC processor. These were fast minicomputers that did a terrific job; they were taking over. There was no money to buy a big, new replacement supercomputer, so we were using these small computers. "RISC" stands for "reduced instruction set computing." Some people say it means "Really Invented by Seymour Cray." If you look at the architecture of the Cray-1, the reason it was so fast is because it used the same techniques as the RISC machines.

You've deployed generations of supercomputers, from early machines such as the Cray 1 and the KSR-1 computers to ORNL's first production, massively parallel machines, the Intel Paragons and later, Cheetah and Jaguar. Did that machine evolution follow any themes?

The theme for supercomputers has always been speed on applications. The scientists and engineers using these systems need the fastest computers possible to attack time-critical problems. The limits of performance have always been defined by memory latency and bandwidth, processor performance, and application scalability. Almost everything we design into supercomputers is there to address one of these four areas. The other theme is the application domains. Climate modeling, materials science, computational fluid dynamics, and molecular dynamics have dominated high-performance computing, at least in the open-science domains.

What science applications were you running on the early parallel computers?

We had people working on groundwater flow for pollution remediation. We're still working on some of those kinds of problems. We had people working on climate modeling. These climate problems are broken up into discrete elements around the globe, and a particular resolution that they were running all of their climate models on at the time was called T42. [At this resolution, temperature, moisture, and other features were tracked in grid boxes each spanning about 120 by 180 miles at midlatitudes, an area about as large as West Virginia.] At the time the climate scientists were hoping to get to T170 resolution, and they still needed to be able to do five simulated days of climate across the world in one day of real time on the computer. We weren't able to get to T170 resolution on the Paragon, but we were at least able to run a T170 code on the Intel Paragon XP/S 150 simulating a few hours of the world's climate and demonstrate that we could make it work. There had never been a computer that had done that before… That was a very large parallel machine. It had all the kinds of problems that we are still dealing with today: Can you make it stay up? Can you make it resilient? Can you program the thing? Does the operating system work? Can you deal with the operating system jitter? Can you get a file system that's fast enough to get the data in and out of the machine? Does it have enough memory? All of the things that continue to be themes in parallel computing today we were dealing with back then, trying to figure out how to make this work. But it was remarkable. And we did make it work. And people got terrific science out of this. In fact, a scientist at ORNL, collaborating with [researchers at] other labs, won the Gordon Bell Prize for the first code able to sustain 1 gigaflop. This occurred on a Cray machine [Y-MP8] back in 1988. In 1990 Malcolm Stocks, Al Geist, and collaborators again crossed the 1 gigaflop threshold, this time using novel parallel architecture by Intel called the iPSC/860, also winning a Gordon Bell Prize. Stocks and collaborators achieved another milestone in 1998 with the first code, LSMS, able to cross the threshold of 1 teraflop with a Cray T3E. In 2008 successors to Stocks in the ORNL computational materials science enterprise, including Thomas Schulthess, were the first to achieve a petaflop using the superconductivity code DCA++. In 2009 Schulthess and Markus Eisenbach led a team to once again win the Gordon Bell Prize for application performance for the code WL-LSMS, the successor to LSMS, developed at ORNL to compute the magnetic structure of materials from first principles. We were pointing this out at a conference one time, and somebody was asking what's going to be the first code that's going to be an exaflop, and somebody stood up and said, "I can't tell you [what] the name of the code is going to be, but I can tell you it's going to be at Oak Ridge National Laboratory and it's going to be in 2018." Following that line, we're actually trying to hit that projection.



Buddy Bland examines the 19 cabinets of ORNL's Intel Paragon XP/S 150, which in 1995 was the fastest high-performance computer in the world at 150 gigaflops. Today a traditional AMD Opteron CPU—one chip—is as powerful as the Paragon, operating at about that same speed, Bland says. The GPU accelerators that will be installed in the Titan supercomputer will be about ten times as fast. Image credit: ORNL

Six of the top 20 supercomputers on the June 2012 TOP500 list were DOE machines. What's the key to wringing science out of powerful supercomputers?

The other important contribution that has come out of ORNL's leadership computing efforts is not really the series of computers. The most important thing that has come out of this has been the fact that we've figured out how to scale these applications and make them effectively use these massively, massively parallel computers. In the early days of parallel computing, there was much debate about whether going from four to eight processors would even be possible, whether you could get the operating system to run with eight, or whether it would just use so much overhead that you would not effectively use that. We showed all these detailed scaling plots that showed that by the time you got between four and six processors, the performance gains of adding more processors were negligible and there was hardly any point in going beyond that. We have computational scientists here who have spent their careers figuring out how to do this and have come up with algorithms and tools and have figured out how to make this work. And if anybody is an unsung hero, they are because they are the ones who have made these machines useful for scientific computing and given us the ability to solve these enormous problems. Working with our users, we have scaled codes here beyond what anybody else in the world has done. Others are catching up now, and there will be others that have more processors than we do. But much of this early work has been done by people here at Oak Ridge National Laboratory.

Titan's coming. Is it a logical milestone along a tried-and-true path or a disruptive technology that'll become indispensable once scientists master the GPU aspect?

I like to think of Titan as the Rosetta Stone of computing. It has both {GPUs and} the traditional CPUs that we've used for many years, these AMD Opteron x86 processors that are very much in the Intel Paragon, IBM Power, and Cray XT line of computers that we've had at ORNL. People can take their codes from the old Jaguar before the upgrade, put them onto Titan, and they will continue to run on the 2.6 petaflops of CPU computing. Now, you're not going to get a huge benefit from that, but they will run right out of the box. But once you do that and your code is working, you can work on converting your code to be able to use the application accelerators, the GPUs. We can use the application accelerators to figure out how to make these applications run dramatically faster. We're seeing examples of this right now, and it's not just the accelerator. All you're getting from the accelerator is more parallel threads of execution. It's not like any of them are any faster than the others; in fact some of the individual threads in the accelerators are slower than the threads in the Opteron processors. It's just that the accelerator has so many of them for a given amount of power that it is much more efficient to use than traditional CPU threads. The point of Titan is to deliver great science today, but do it in a way that gives the users a transitional step from traditional CPU computing to a more accelerated computing paradigm, which is where we believe things are going to go with exascale computing.

What is the energy difference for the CPU/traditional versus GPU/accelerated path?

If we were to build an exascale computer using today's technology, it would require the power output of two nuclear power plants—the power required by a city of a million people. Clearly that's not sustainable or affordable. At best, if you look at projections for what would happen if we don't do anything, just let technology take its course, we would predict that an exascale computer by the end of this decade would use as much power as a small city, approximately 100 megawatts, about the same amount of power as the city of Oak Ridge uses with 35,000 people. Again, that's not something that we can do. We really need to find new ways of doing much more efficient computing, and these application accelerators are dramatically more efficient than the traditional CPUs that we have today. A traditional AMD Opteron processor today can do about 150 gigaflops. Interesting, because that Intel Paragon occupying 19 cabinets that we had in the early days was 150 gigaflops. Now you can buy a CPU—one chip!—that's as powerful as the fastest computer in the world in 1995. One chip today is about 150 gigaflops, and it uses about 100 watts—the power of one light bulb. The accelerators that we will be deploying will use more power—about 200 watts—but they will be about ten times faster. So the GPUs that we'll be deploying in Titan will be about five times more energy efficient per computation than what we're using today on the CPUs.

What do you hope ORNL's Computing and Computational Sciences Directorate will be celebrating 20 years from now?

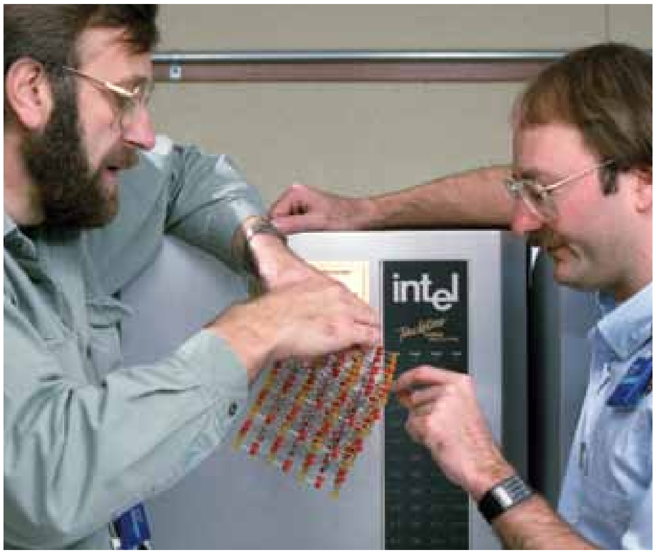
It's got to be all science. Twenty years from now we would anticipate that we would be delivering the first zettaflops computer, which is 10^{21} operations per second. And we'll be using that to solve the most challenging problems of the day. I'm sure we'll still be working on many of the same types of science and engineering problems we explore today, but we'll be doing it with higher resolution and better understanding and more physical fidelity. Today we can develop products like the Boeing 787 completely on a computer, whereas in the past the engineers would build it and go fly it. We need to be able to do predictive science in that way 20 years from now for many, many more things. Twenty years ago we installed a machine that was 150 gigaflops. Today my iPhone is probably pretty close to that. The petaflop computer of today in 20 years will be in your cell phone. Just think about Siri on your cell phone that can answer all those questions. Think about what she can do now with all that power.—*Q&A with Dawn Levy*

America’s Premier Supercomputing Complex Celebrates 20 Years of Discovery

In the 1940s when the US government built a secret research facility in East Tennessee as part of a project to develop the atomic weapons that ended World War II, big computing machines were the exotic and expensive tools of elite labs. The machines used a single processor to carry out one calculation at a time. Over the years computer scientists made that lone processor run faster and faster and in some cases combined a few processors to help solve complex problems. In the ’80s researchers at the DOE Office of Science ORNL decided to explore solving scientific problems with parallel computing at a more massive scale. This many-hands-make-light-work strategy split problems into parts that multiple processors could calculate concurrently, allowing the problems to be scaled up. The approach decreased the time to solution and increased the realism of simulations, and as a result researchers clamored for an infusion of HPC resources.

On December 9, 1991, Congress signed the High Performance Computing and Communication Act (HPCCA) into law. Through government funding of technology, HPCCA launched the world into cyberspace much faster than might have occurred through private-sector efforts alone. It led to the development of the National Information Infrastructure, which we know today as the internet. Building upon the Department of Defense’s ARPANET of the 1960s and the NSF’s NSFnet of the ’80s, it created a seamless web of computers, public and private communications networks, interactive services, interoperable hardware and software, databases, and consumer electronics that put information at the fingertips of the masses. HPCCA also funded the National Research and Education Network, which provided high-speed internet service to research and education communities and enabled innovative internet protocols and architectures.

HPCCA also called for proposals to build new HPC facilities to serve science. Researchers quickly saw the potential for parallel processing to address grand challenges, such as unlocking the secrets of DNA, forecasting severe weather, and inventing superconducting materials. At ORNL, Laboratory Director Al Trivelpiece put Ed Oliver, director of the Office of Laboratory Computing, in charge of assembling a team to propose that DOE’s Office of Science fund a High Performance Computing Research Center at the lab.


In the early 1990s Malcolm Stocks, left, and Al Geist at Oak Ridge National Laboratory examined a physical model of a superconductor’s electronic structure, which was calculated on a microprocessor-based parallel computer from Intel. Image credit: ORNL

In 1992 Oak Ridge’s only HPC footprint was a Cray X-MP computer that had just one CPU and a Kendall Square computer with a novel shared-memory architecture. Since 1985 Bob Ward, leader of the Engineering Physics and Mathematics division, had been purchasing parallel computing prototypes with which his group experimented. ORNL had the first Intel iPSC/1 [Intel Personal SuperComputer/1]. Other early machines included the iPSC/2, iPSC/860, and nCUBE, all of which had 64 processors. Ward’s group learned to program these microprocessor-based parallel computers and developed algorithms and mathematical libraries that could run on them. Al Geist was the main architect behind a software package called Parallel Virtual Machine that enabled message passing between computers so an application could run across more than one compute node. This achievement allowed research teams to create parallel systems using personal computers. With these parallel systems, scientists at ORNL and other locations developed some of the earliest parallel applications to take advantage of the power of these systems.

Based on these experiences, ORNL computer and computational scientists submitted a proposal, “Partnership in Computational Science” (PICS), with colleagues from three national laboratories—Ames, Brookhaven, and Sandia (Albuquerque)—and seven universities—Rice, Stony Brook, Texas A&M, South Carolina, Tennessee (Knoxville), Wyoming, and Vanderbilt.

“To everyone’s amazement, including our own, we won,” recalled Arthur (Buddy) Bland, who came to the lab in 1984 to run the X-MP and now directs the Titan supercomputer project to deploy a 299,008-core hybrid CPU-GPU Cray XK6 system that will deliver 2 billion core hours to the scientific community in 2013.

The High Performance Computing and Communications Act paved the way to the founding of the Center for Computational Sciences in 1992 and the creation of the National Center for Computational Sciences, shown here, in 2004. Image credit: ORNL

Today ORNL is the home of a world-class scientific supercomputing complex. Researchers from across the globe vie for time on its supercomputers to conduct computational experiments that allow them to be the first to make discoveries or invent products. The lab's expertise and infrastructure enable efficiencies that give a greater return on the public research investment. That confluence of resources took supercomputing at ORNL from a handful of computer and computational scientists to one of the world's premier computational research facilities in just two decades. Currently the OLCF provides the scientific community access to Jaguar, the world's sixth-fastest supercomputer. By the end of 2012, an upgrade will turn Jaguar into the mightier Titan, a hybrid of energy-efficient application-code accelerators called GPUs and traditional CPUs with the ability to execute at least 20 quadrillion calculations per second, or petaflops. The NCCS, which houses the OLCF, also hosts the Gaea supercomputer, operated by ORNL on behalf of the National Oceanic and Atmospheric Administration, and collaborates with the University of Tennessee to house and manage the NSF's Kraken supercomputer.

> "There was more to it than just delivering hardware because what the Partnership in Computational Science proposal was really about was how to do science on these parallel machines." – Arthur Bland

Commemorating the 20-year anniversary of the founding of the CCS at ORNL is an opportunity to pay homage to the past and admire the vision and hard work that enabled accomplishments that now light the way to the future.

"We're celebrating the vision and leadership of the creators of the High Performance Computing and Communications Act, which laid the foundation for the world's most powerful computing complex here at Oak Ridge National Laboratory," Bland said. "We're celebrating the leadership of scientists here at Oak Ridge and partners around the country who wrote the proposal that built this high-performance computing capacity and gave the worldwide scientific community unprecedented tools for discovery. And we're celebrating the science that has been achieved on these machines over the last 20 years."

Creation of the Center for Computational Sciences

In May of 1992 ORNL was awarded a High Performance Computing Research Center named the CCS. Other centers were awarded to university, industry, and government affiliates of the NSF, NASA, and another DOE lab. Ward became CCS's interim head until the hiring of its first employees—Ken Kliewer to direct the center and Bland, transferred from another ORNL group, to assemble and deploy the computer systems.

The proposed centerpiece of the CCS was a very-high-performance parallel computer, an Intel Paragon that represented the state of the art in computing. In those days the Defense Advanced Research Projects Agency had a program called Touchstone that built several prototypes of parallel machines. Intel built one, called Touchstone Sigma, at

ORNL. Later the company commercialized the machine and changed its name to Paragon.

"What we delivered was a series of increasingly more powerful parallel computers from Intel," Bland said. "The first one [the Intel XP/S 5] was a single cabinet and had 66 processors and could do 5 gigaflops. A typical laptop PC is about 50 gigaflops today." The Titan supercomputer that Bland's team will deploy in 2012 will be at least 4 million times more powerful than the XP/S 5. The second machine in the series, the XP/S 35, had 512 processors in nine cabinets and could calculate at 35 gigaflops.

"In 1995 we got the culmination of this series of machines that had started in 1992, the Intel Paragon XP/S 150, which, when it was delivered, was the fastest computer in the world," Bland said. Alas, the XP/S 150, capable of 150 gigaflops, would never head the TOP500 list of world's fastest supercomputers because it was delivered between two lists, and an even faster computer was delivered before the next list was published.

"You'll see this theme of increasingly powerful computers repeated over the years," Bland said. "There was more to it than just delivering hardware because what the Partnership in Computational Science proposal was really about was how to do science on these parallel machines. My part of it was getting the machine stood up. But lots of other people were around whose jobs were to figure out how to program these parallel machines and how to get good performance on these applications."

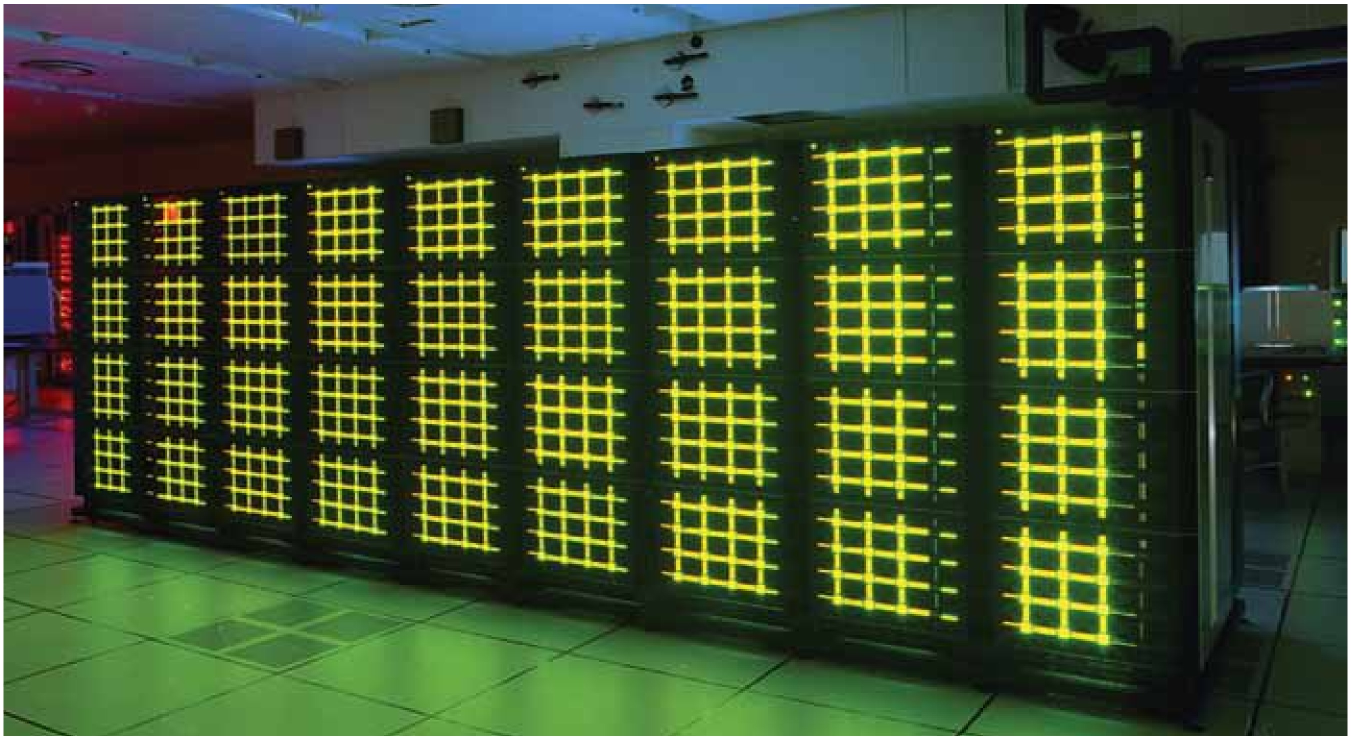
Importantly, the PICS proposal created a national user facility for research called the Computational Center for Industrial Innovation (CCII). The CCII was codirected by Gary Jacobs and Thomas Zacharia. As its name implies, the CCII focused on industrial competitiveness through computational science. Today that focus is preserved through the OLCF's industrial outreach program, named Accelerating Competitiveness through Computational ExceLlence (ACCEL) and headed by Suzy Tichenor.

The CCS enabled users to run scientific applications at almost unimaginable speeds. With each machine upgrade, CCS experts worked closely with science teams to scale applications to run on the increasingly parallel architectures. The partnership was fruitful, and simulations on the Paragons provided insight into how solids melt, pollutants flow through groundwater, materials behave in vehicle collisions, combustion occurs in engines, atoms interact in magnetic materials, and air flows over a plane's wing.

Over the 20 years since its establishment, the CCS has been instrumental in creating a global computational sciences community, deploying some of the world's most powerful supercomputers, and enabling unmatched scientific discoveries and technological innovations. Users have published thousands of scientific journal articles; center staff and partners have published hundreds of technical papers.

In response to the growth in the importance of computing, ORNL created the Computing and Computational Sciences

In 1995 a series of ever-more-powerful parallel computers at ORNL culminated in the Intel Paragon XP/S 150. Capable of 150 gigaflops (150 billion floating point operations per second), it was the fastest supercomputer in the world. Image credit: ORNL



Directorate (CCSD), today led by Jeff Nichols. It has more than 600 staff members in four divisions. The first is the NCCS, which houses the OLCF. Its director is climate scientist James Hack, who also heads the Oak Ridge Climate Change Science Institute. The OLCF is available to computational scientists in industry, academia, and government through three programs that allocate millions of processor hours to users: the INCITE program, which supports high-impact, computationally intensive research; the ASCR Leadership Computing Challenge, which primarily aids research that supports the energy mission of the Office of Science and emphasizes high-risk, high-rewards endeavors; and OLCF's Director's Discretionary program, which helps develop new users of HPC and enhance ORNL's ability to compete for new work. OLCF supercomputers enabled major insights in astrophysics, combustion, nuclear fusion and fission, materials science, biology, and more. Currently the OLCF has 1,003 users worldwide. A key to its success is deep engagement with users to develop and scale their applications. Recognizing such engagement as critical to being able to use extremely powerful supercomputers, other centers subsequently adopted the OLCF's pioneering model to enable success on their HPC systems.

CCSD's second division, led by Barney Maccabe, is Computer Science and Mathematics (CSM). It evolved out of Ward's Engineering Physics and Mathematics division. Whereas NCCS is a scientific user-facilities division with the mission to deploy and operate supercomputers for researchers, CSM is a research division funded by science and engineering programs, largely supported by DOE's Office of Science, in fields including astrophysics, chemistry, materials science, Earth science, engineering, energy sciences, computer science, mathematics, and molecular biology. It is ORNL's premier source of basic and applied research in HPC, applied mathematics, and intelligent systems. Group members assess and develop future technologies, strategies to manage the tsunamis of data that supercomputers generate, and application performance tools.

In 8 hours the lab's Oak Ridge Automatic Computer and Logical Engine, a.k.a. ORACLE, arrived at an answer that would have taken 100 people each working a year to compute.
Image credit: ORNL

CCSD's third division is Computational Science and Engineering, headed by Brian Worley. Its staff members do work for other federal agencies, much of it related to defense and homeland security. The research largely deals with data gathering and analysis—quickly sorting through haystacks of data to find the needles.

The forth division in CCSD is the Information Technology Services Division (ITSD), led by Michael Bartell. ITSD provides the computing and support infrastructure for ORNL, including the high-speed networks for the CCS.

Rise of the machines

ORNL was a leader in computational research in the '50s and early '60s as a result of the math and computing work of Alston Householder and his recruits. For a while in 1954, the Oak Ridge Automatic Computer and Logical Engine (ORACLE) had the fastest speed and largest data-storage capacity of any computer. With vacuum tubes and a storage capacity of 1,024 words of 40 bits each, ORACLE could do 100 person-years of computing in 8 hours.

In 1985 the lab purchased a Cray X-MP system that was shared between classified uranium enrichment programs and open energy research programs. As the Cray approached the end of its life, it was apparent that new microprocessor-based systems were rapidly approaching the performance of large custom HPC systems. Oak Ridge researchers engaged with technologists who were creating clusters of microprocessors to build superfast parallel machines.

The lab subsequently procured KSR1 and Intel Paragon computers. Thomas Zacharia in metals and ceramics led a group that used the Paragons for materials-engineering, car-crash, and welding simulations. Other researchers ran applications to study groundwater, fuel combustion, climate, magnetic interactions, and melting solids.

By early 2000 the Paragon 150, ORNL's biggest machine, was approaching the end of its useful life. Zacharia, who by then had been appointed to head the lab's scientific computing division after Kliewer's retirement, brought in an even more powerful IBM Eagle supercomputer.

CCS's powerful IBM systems were funded in large part by the Research and Evaluations Prototype program until the creation of the NCCS in 2004. Prototypes prodigiously populated the computer room on the second floor of Building 4500 North. First came Eagle, an IBM Power3 machine that was the Office of Science's first teraflop computer and the tenth-fastest computer in the world. Eagle was upgraded with higher-density processors to become Winterhawk 2. When a Power4 called Cheetah was delivered, it was #8 on the TOP500 and went on to provide more than one-third of the simulation data for the joint DOE/NSF data contribution to the United Nation's IPCC

Phoenix, a Cray X1 at left, and Jaguar, a Cray XT3 at right rear, were among the first leadership-class high-performance computing systems deployed in the National Center for Computational Sciences's computer room. Image credit: Larry Hamill

Fourth Assessment Report. "The IBM Power4, Cheetah, was a remarkable machine for us," Bland said. "It was here as part of this Research and Evaluation Prototype program, which means it didn't really have a direct mission to deliver cycles and do science. Yet we worked with the scientists who desperately needed cycles and couldn't get them anywhere else."

Other machines—Falcon, a Compaq AlphaServer, and Phoenix, a Cray X1—filled out the computer room. Cabinets weighing 6,000 pounds each taxed the floor's loading capacity. "We were out of space and had to migrate," Bland said.

Sputnik moment

In 2002 Japan's Earth Simulator eclipsed US supercomputing with its unprecedented speed of 36 teraflops (trillion calculations per second). It had the processing power of America's 14 fastest computers combined and was five times faster than its swiftest, the IBM ASCI White system at DOE's Lawrence Livermore National Laboratory. "The power of the Earth Simulator really took the world by surprise," Bland said.

Just as the shock of Russia's 1957 launch of the Sputnik satellite motivated America to invest to win the space race, the surprise of Japan's Earth Simulator provoked US supercomputing investments beyond those made for stockpile stewardship. It was time to increase investment in next-generation supercomputers for science and engineering. DOE's Office of Science issued a call for proposals to develop a leadership computing capability to support investigations in a broad spectrum of scientific domains.

By this time Cray had been acquired by SGI, which subsequently spun it off. Cray machines such as Phoenix, named in honor of the company's rebirth, employ reduced instruction set computer (RISC) processors. While the Intel x86 microprocessor architecture also in use today employs complex instructions, the RISC processors pare instructions to the essentials to run faster.

In response to DOE's challenge, ORNL partnered with Cray and science teams on an ambitious proposal to increase computing capability a thousandfold in just 4 years. Science applications on the proposed "Jaguar" supercomputer would run at a petaflop. Calculations that once took months would take mere minutes. Jaguar would help achieve DOE's science and energy missions, contribute to the nation's security, and improve standards of living and health.

Field of dreams

At the turn of the millennium, if someone had plunked down, among rusty, ramshackle labs erected at Oak Ridge in the '40s to win a war rather than an aesthetics award, a sign that said "future home of the world's greatest supercomputing complex," the probable reaction would have been laughter. Indeed, when UT-Battelle took over the lab's management and operation from Lockheed Martin in 2000, the lab's infrastructure was in desperate need of modernization if managers expected to continue to attract premier scientists. Before the next winter, Bill Madia, the lab's director at the time, and teammate Jeff Smith had planned the first new office building.

Meanwhile, Zacharia, director of the CCS division, had been trying to figure out how to build the computing program at the laboratory. He had had many discussions with Oliver, who now directed the ASCR office at DOE headquarters, about the HPC needs of DOE's Office of Science. They realized that computer sizes would need to dramatically increase and that the Office of Science would need a place to build these massive machines.

Zacharia persuaded Madia that the nation needed a leadership computing facility and that ORNL, with inexpensive power from the Tennessee Valley Authority, would have a leg up in winning a proposal if it turned part of the new office building

Thomas Zacharia thought inexpensive power from the Tennessee Valley Authority and converting a new office building into a computer room would help ORNL win a proposal for a leadership computing facility. Turns out he was right: in 2004 DOE chose ORNL to deploy America's first leadership computing facility. Image credit: ORNL

into a computer room. "His vision was kind of *Field of Dreams*—if we build it, they will come. That's really the visionary moment that allows this thing to go forward," Bland said. "UT-Battelle management was willing to take a big risk to go after what it thought were important, major programs for the Department of Energy."

The DOE High-End Computing Revitalization Act of 2004 authorized America's leadership computing facilities, and the appropriations process made funds available. The act defined a high-end computing system as "a computing system with performance that substantially exceeds that of systems that are commonly available for advanced scientific and engineering applications." It defined a leadership system as "a high-end computing system that is among the most advanced in the world in terms of performance in solving scientific and engineering problems." The act prioritized capability computing (solving a single large problem in the shortest amount of time) over capacity computing (solving a small number of somewhat large problems or a large number of small problems). Allocations on leadership computing resources would be about 100 times larger than those commonly available.

DOE chose ORNL to deploy the nation's first leadership computing facility. In 2004 the NCCS was established in the new office building. Zacharia was appointed associate laboratory director for ORNL's new CCSD, which he oversaw until his 2009 promotion to deputy director for science and technology.

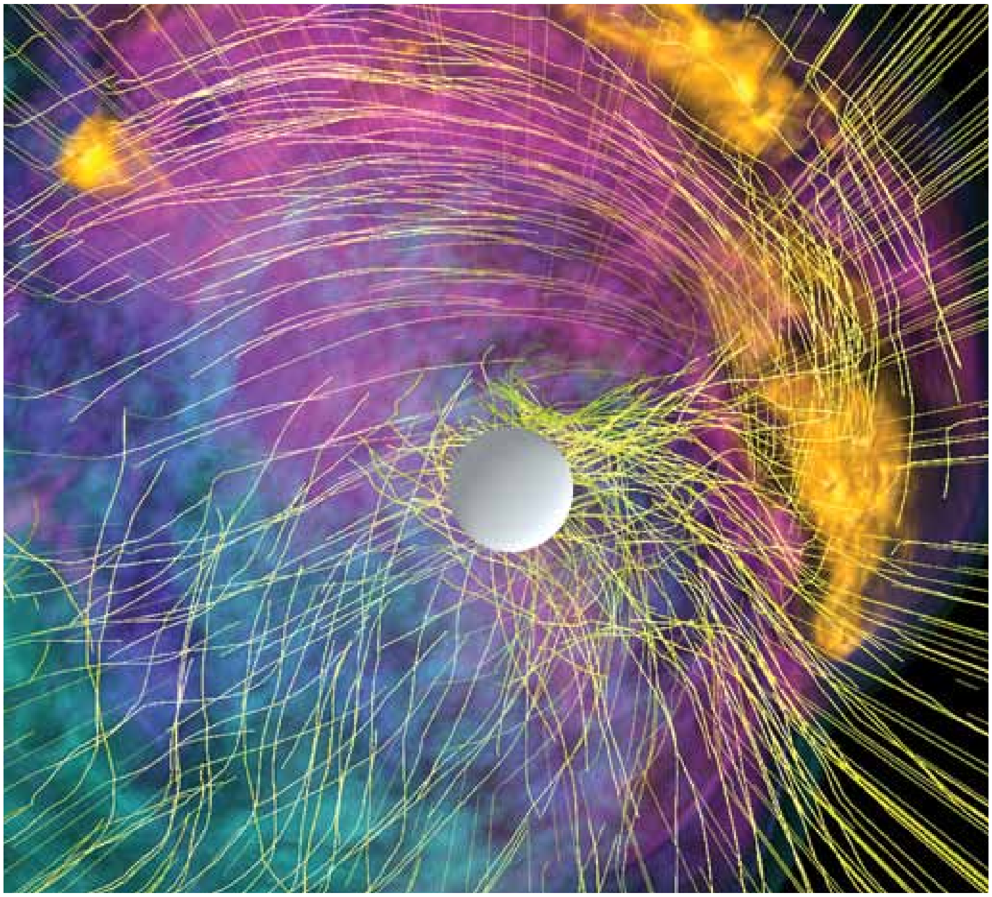
ORNL had won the leadership computing facility competition with Argonne National Laboratory as a partner. But as more money became available for civilian scientific supercomputing, ASCR managers decided to have a two-headed DOE Leadership Computing Facility program to embrace diverse computer architectures and subsequently established the ALCF. Though there is no mandate to embrace any particular vendor or architecture, the OLCF has largely worked with Cray, while the ALCF has mainly collaborated with IBM.

"I remember the day we moved in [to the NCCS building]," Bland said. Its computer room was bigger than a basketball court. "We moved in one cabinet of the Cray X1. We moved in the Cheetah machine. And we moved in some of the IT equipment for the laboratory. [Zacharia] and I walked through that computer room, and he looked at me and said, 'How are we going to ever fill this place up?'"

A star simulator is born

Today the 40,000-square-foot facility is chock-full of computers. It has undergone four upgrades of power and two of cooling to support ever-more-powerful supercomputers, such as the OLCF's Cray XT3/4 Jaguar system, which underwent staged increases to 26, 54, 119, and 263 teraflops. In 2008 a 1.382-petaflop Cray XT5 system was added. The Jaguar XT5 became the first HPC system to run a scientific application at a sustained petaflop (the DCA++ code to simulate high-temperature superconductivity). A 2009 upgrade of the XT5 to 2.332 petaflops made it #1 on that November's TOP500 list, where it remained for a year, after running the Linpack benchmark application at 1.75 petaflops.

A 2012 upgrade took Jaguar to 3.3 peak petaflops, and by the end of 2012, the OLCF will finish adding GPUs during yet another upgrade to reach at least 20 petaflops. Jaguar has been an extremely productive scientific resource for its users. If mathematics is the language of science, computation is its workhorse and Jaguar its thoroughbred.



Oak Ridge simulations have made pioneering contributions in numerous scientific fields. A study published in *Nature* in 2007 caused astrophysicists to revise theories about how stars explode and how pulsars get their spins. In 2012, simulations by this same research group landed on the cover of *Science*. Visualization by Kwan-Liu Ma, UC-Davis

Exploring research frontiers and pushing farther and faster to make breakthroughs requires enormous computing power to solve the equations behind the major enigmas of our time, such as what is happening to our planet's climate and how to accelerate the arrival of sustainable energy technologies. The complexity of these problems gives the scientific community a voracious appetite for computing resources. OLCF supercomputers allow scientists and engineers to approach grand challenges, from understanding the nature of the universe to the molecular basis of disease.

Making the most of supercomputers largely means exploiting parallelism to deliver more science. Scaling efforts in materials science highlight this success in but one scientific domain. The Association for Computing Machinery annually awards Gordon Bell Prizes to recognize outstanding HPC applications and assess progress in parallel computing. Many winning codes have run on ORNL supercomputers. The 1990 winner, an alloy-theory code called MKKR-CPA designed to determine the electronic structure of a high-temperature superconductor, was run by ORNL's Malcolm Stocks and Al Geist with William Shelton of the US Naval Research Laboratory and Beniamino Ginatempo at the University of Messina in Italy. The code ran at 2.5 gigaflops on a 128-node Intel iPSC/860. In 1998 Stocks and Shelton, now at ORNL, ran the winning code, LSMS, with ORNL colleagues Balazs Ujfalussy, Xindong Wang, Xiaoguang Zgang, and Donald Nicholson, as well as Andrew Canning of LBNL, Yang Wang of the Pittsburgh Supercomputing Center, and Balazs Gyorffy of the H.H. Wills Physics Laboratory in the United Kingdom. The code ran at just over a teraflop on a Cray T3E system to calculate the electronic structure of a metallic magnet.

In 2008 the Gordon Bell Prize was awarded to DCA++, an algorithm run at 1.352 petaflops on ORNL's Cray XT Jaguar by Gonzalo Alvarez, Michael Summers, Don Maxwell, Markus Eisenbach, Jeremy Meredith, Thomas Maier, Paul Kent, Eduardo D'Azevedo, and Thomas Schulthess, all of ORNL, and Jeffrey Larkin and John Levesque, both of Cray Inc. The code, used in simulations of disorder effects in high-temperature superconductors, helps scientists understand, predict, and optimize the complex behavior of these materials and accelerate their eventual development in technologies that save energy in power cables, electric vehicles and trains, and electric machinery. And the 2009 winner, the WL-LSMS code to analyze the effect of temperature on magnetic systems, important in many technologically useful materials, achieved 1.84 petaflops on ORNL's Cray XT5 Jaguar system, making use of more than 223,000 processing cores and reaching nearly 80 percent of Jaguar's peak performance of 2.33 petaflops. Led by ORNL's Eisenbach, the team included Chenggang Zhou of J.P. Morgan Chase & Co.; Nicholson of ORNL; Gregory Brown of Florida State University; Larkin of Cray; and Schulthess, now of ETH Zurich and the Swiss National Supercomputing Centre.

If math is the language of science, computation is its workhorse and Jaguar (above) its thoroughbred. From bottom left, simulations improve the safety and efficiency of nuclear power plants, aid our understanding of climate change, and speed development of new classes of drugs. Images courtesy ORNL

Other superfast codes running at the OLCF are taking engineers closer to batteries that can power cars for 500 miles before having to be recharged, commercially viable cellulosic biofuels, advanced internal combustion engines with twice the fuel efficiency of conventional engines, fusion reactors to provide clean energy, and as reported in *Nature* this year, more complete maps of the nuclear landscape so scientists can calculate the number of isotopes allowed by the laws of physics. They also solve the electronic structures of industrially important catalysts and device interfaces to accelerate breakthroughs in chemistry, nanotechnology, and materials science.

Moreover, applications running on OLCF supercomputers have illuminated the molecular mechanism underlying Parkinson's disease and provided the fundamental understanding of quarks and gluons needed to interpret physics experiments. As reported in two more *Nature* articles, they elucidated abrupt climate change at the end of the last Ice Age and showed that clumping of dark matter birthed galaxies. They enabled the most advanced earthquake simulation to date to aid building-design and emergency-response efforts, and they have helped researchers understand, control, and design processes and products for clean energy, such as biomass conversion for energy production and supercapacitors for energy storage.

World-class visualization facilities turn the floods of numerical data generated by OLCF supercomputers into images that can lead to insights. The EVEREST Powerwall, on which data are projected for researchers, is a data-analysis system that triggers the occasional "Aha!" moment. Visualization expert Sean Ahern recalled the first-ever 3D simulation of a core-collapse supernova by Tony Mezzacappa's team. The researchers didn't expect spin given that they started off with a nonrotating system. But when they saw the results played as a movie on EVEREST, they realized they had discovered that the exploding star's shock wave may be unstable, drive stellar magnetic fields, and spin the neutron-rich core remaining after the blast to produce a pulsar. The results were published in *Nature* in 2007, and in June 2012, an OLCF simulation image by this same group, which had by now caused theories about how stars explode and how pulsars get their spins to be revised, landed on the cover of *Science*.

Predictive science

By the end of this decade, supercomputers are expected to reach the exascale—a thousand times faster than today's petascale machines. That acceleration will change what's possible for science, said NCCS Science Director Jack Wells.

It may usher in an era of truly predictive science. Predictive simulations may push progress in scientific fields with long-unsolved issues, such as improving efficiencies of photovoltaic materials or predicting long-range interactions in polymers.

Wells provided an example of how exascale computing will change one field. “In materials today we have state-of-the-art methods that have a really good predictive capability,” said Wells. “But weaknesses in some approximations that are made limit the transferability of these methods to problems for which they haven’t previously been validated. What that means is that people don’t have confidence to apply these methods on materials where they haven’t been tried before. There’s this sense of creeping empiricism where we only really trust our methods where they’ve been validated and tweaked, where we know the experimental truth. A truly predictive method could go out in front of experiment and have results that one can trust.”

David Ceperly of the University of Illinois at Urbana-Champaign is famed for using computers to analyze the electronic structure of models of materials with uniform charge density, called the free-electron gas. He numerically solved this model using very accurate Quantum Monte-Carlo methods and fit equations to extract the so-called local-density approximation for materials, which is often an efficient computational approach to obtaining approximate solutions for materials with uniform charge densities. “We spent the last 40 years basically trying to add incremental accuracy, for example by considering gradients in the charge density, to Ceperly’s approximation,” Wells said. “That’s where we still are.”

Supercomputers can now perform “hero” benchmark calculations of real materials. These benchmarks allow researchers to extract mathematical equations governing materials without the need to resort to model approximations. To achieve this, ORNL’s Fernando Reboredo, Paul Kent, and Jeongnim Kim are collaborating with Ceperly to use supercomputers to validate and improve the last 40 years of materials simulations.

Hero calculations have already benefited GE [see story, pp. 20–21] and Ramgen Power Systems [see story, pp. 22–23], companies that are designing turbomachinery for aircraft engines and gas compressors, respectively. “All of our commercial airplanes effectively are powered by [turbines],” Wells noted. “More than half of our electricity is generated by turbogenerators. [Turbomachines] all share very similar characteristics. It’s a mature technology, so if you improve efficiency by 1 percent, it can [dominate] the marketplace.”

In 2008 the Gordon Bell Prize was awarded to DCA++, an application run at 1.352 petaflops on Jaguar by researchers at ORNL and Cray. Accepting the award from Association for Computing Machinery Executive Director John White, right, is Thomas Schulthess, who today heads Switzerland’s premier supercomputing center.

The rich behavior of fluid dynamics is a big computational problem if researchers try to model turbulence at all the levels and resolve all possible phenomena. In GE’s case, company researchers have always made an assumption in engineering simulations that they know is not true but that, for reasons not fully understood, works pretty well. They have assumed airflow is steady over turbine blades. The OLCF resources made it possible to consider unsteady flow in GE’s biggest CFD simulation ever, which revealed unsteadiness builds as gas traverses fan blades and heats up. If materials are not designed to be temperature-resistant, they can fail. Based on the results of that study, GE purchased a Cray XE6 to continue studies in-house.

Said Wells, “For the OLCF that’s something of an accomplishment because we’ve encouraged the adoption of high-performance computing in American industry. We don’t really know what the outcome is going to be for their turbomachinery, but it has introduced a new tool set. Ten years from now I’d expect we’d have significantly more efficient turbomachinery.”—*by Dawn Levy*

Inside the OLCF

OLCF Systems Overview

Introduction

The OLCF is the premier computational research center in the country. From one of America's fastest supercomputers to one of the largest file storage systems in the world, the OLCF gives researchers all the tools they need to produce world-class scientific findings.

ORNL not only provides resources for leadership science but also delivers them as efficiently as possible. The Computational Sciences Building was one of the first Leadership in Energy and Environmental Design (LEED) certified computing centers in the country. The computer room has a vapor barrier that seals it off from the rest of the building, helping to keep a relatively cool and dry climate for efficiency. The air pressure is higher inside the room than outside, allowing air to flow out with none entering from the outside.

Further, the OLCF uses high-efficiency chillers for the massive amounts of refrigerated water needed to keep the various computing resources cool. The ECOphlex cooling system removes heat generated from Jaguar and helps reduce the amount of chilled water needed to keep the machine running effectively. The OLCF uses a quarter of a watt of power for cooling each watt used for computation. Many centers close in upon a one-to-one ratio in this category, making ORNL's computation some of the most efficient in the country.

As the OLCF continues to integrate the Titan supercomputer, a Cray XK6 capable of 20 thousand trillion calculations, efficiency remains a top priority. Titan is expected to not only exceed the world's current fastest supercomputer, the K supercomputer in Japan, in computational speed, but will also be up to three times as efficient.

Jaguar supercomputer.

Jaguar

Jaguar has been one of the most productive supercomputers for science. As the system is slowly transitioned into Titan, it will only become more effective at delivering science on larger scales. During 2011, Jaguar ran at a peak speed of 2.33 petaflops, or 2.33 quadrillion calculations per second. In 2012, an upgrade increased peak speed to 3.3 petaflops.

Jaguar spreads out over 5,000 square feet, with 200 cabinets housing the 18,688 compute nodes. Each node has 32 gigabytes of memory, giving the entire system 600 terabytes of memory. Information can travel in and out of Jaguar at 240 gigabytes per second, with 512 service and input/output (I/O) nodes directing large volumes of data to their respective locations.

All of these processes give Jaguar a peak power density of 1,750 watts per square foot, necessitating a liquid cooling system to dissipate heat. ECOphlex cooling technology, designed by Cray, uses R-134, a high-temperature refrigerant used in automobile air conditioners. This technology conditions air as it is drawn into the machine while removing heat as it enters and exits the machine. This system not only is reliable but also saves 2.5 million kilowatt-hours annually.

Data Management

As supercomputers become increasingly powerful, the need to track and maintain large amounts of data has become imperative. In addition, researchers require quick and efficient means to recall their data remotely. The OLCF uses the Inifiband high-performance network to quickly ship data from one location to the other. The industry standard in HPC, InifiBand connects Jaguar with the High Performance Storage System (HPSS), Spider file system, and visualization resources, in addition to other center platforms. With the arrival of Titan, even more demand will fall on staff associated with data management at the OLCF to continue to expand the capabilities for housing and directing data.

High Performance Storage System

The OLCF maintains one of the largest archival storage systems in the world. The HPSS first writes data onto disks using high-speed data movers; then the data is gradually transferred to tapes. In 2009, the OLCF changed the HPSS server and switched the platform to Linux, improving performance and redundancy. Staff members are constantly adding more disk and tape resources to handle the ever-growing mountain of data produced on the center's resources—HPSS grew from 1,000 terabytes in 2006 to nearly 21,662 in 2011. In addition to HPSS, the OLCF houses three Storage Tek SL8500 storage silos for additional archival data space.

Spider

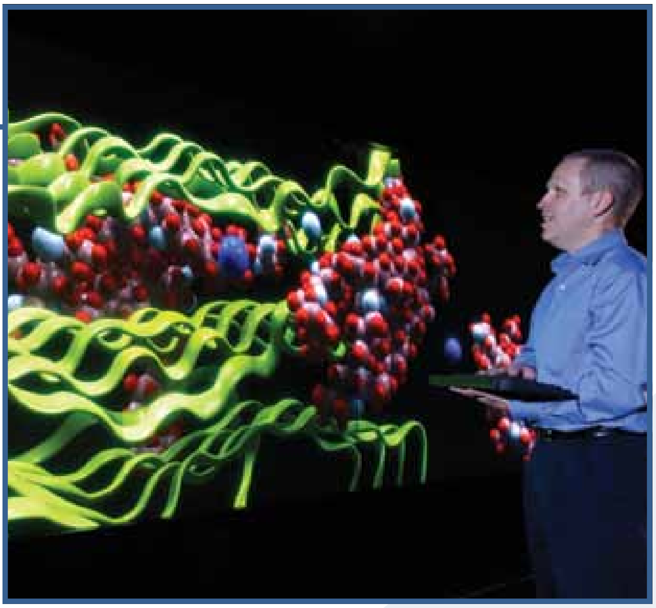
The Lustre-based file system, Spider, lies at the center of the OLCF's technological integration. With more than 26,000 nodes mounting it, Spider is one of the largest file systems in the world. Spider organizes data from the multiple computing platforms into a unified file system. The project began in 2005, and now Spider is the main operational file system connecting Jaguar, the LENS visualization cluster, the Smoky development cluster, and the center's dedicated GridFTP servers—at a blazing-fast bandwidth of 240 gigabytes per second.

Data Analysis

Simulations at the OLCF range from the interactions of subatomic particles to climate simulations spanning the world over millions of years. Jaguar is capable of compiling data, but to get meaningful insight, researchers need to tangibly see their computation in action. All researchers receiving allocations at the OLCF are entitled to use the visualization resources. The data analysis team at the OLCF is trained in a variety of software, including VisIt, EnSight, POV-Ray, AVS/Express, ParaView, and IDL. The team's goal is to help researchers who stand to benefit from visual insights for their research.

EVEREST

The OLCF provides the Exploratory Visualization Environment for REsearch in Science and Technology (EVEREST) and its associated visualization cluster, everest. A large-scale venue for data exploration and analysis, EVEREST measures 30 feet wide and 10 feet tall. Its main feature is a 27-projector Powerwall displaying 11,520 by 3,072 pixels or a total of 35 million pixels, offering tremendous detail. Each of the projectors, arranged in a 9×3 array, provides 3,500 lumens. The everest cluster is a 15 node Linux cluster dedicated to driving the EVEREST Powerwall. Each node contains two dual-core AMD Opteron processors and two NVIDIA 8800 GTX GPUs. Coupled with a dedicated Lustre file system, the EVEREST facility provides a compelling experience for scientists exploring data at extremely fine scales.

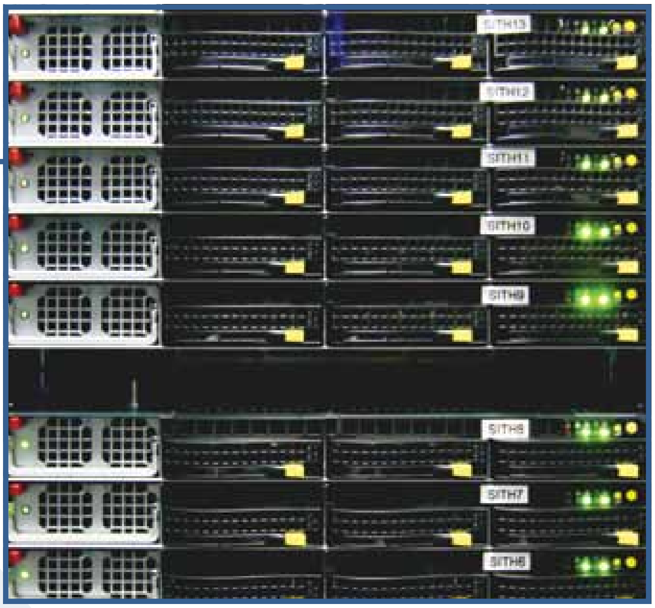


Lens

Lens is a 77-node Linux cluster dedicated to data analysis and high-end visualization. Each of the 77 nodes contains four 2.3 GHz AMD Opteron processors. Of the nodes, 32 are configued with 64 GB of main memory, an NVIDIA 8800 GTX GPU with 768 MB of memory and an NVIDIA Tesla general purpose GPU with 4 GB of memory. The other 45 nodes do not contain GPUs but are configured with 128 GB of memory. The primary purpose of Lens is to enable data analysis and visualization of simulation data generated on Jaguar so as to provide a conduit for large-scale scientific discovery. Members of allocated Jaguar projects will automatically be given accounts on Lens.

Developmental Systems

The OLCF tries to make sure users spend their allocations answering scientific questions, not figuring out how to use computing resources. Two developmental clusters, Sith and Smoky, allow researchers to scale up their algorithms and identify problems before beginning computation on Jaguar. These systems offer many of the same capabilities as Jaguar, albeit on a smaller scale. These devices can also work on auxiliary calculations for larger projects.



Sith

Sith is an Opteron-based Infiniband cluster running Linux. The system is provided as an End-to-End resource for center users. It is used for workflow automation for jobs running from Jaguar and for advanced data analysis. The system contains 40 compute nodes. Each compute node contains four 2.3 GHz 8 core AMD Opteron processors and 64 GB of memory. The system is configured with a 86 TB Lustre file system for scratch space.

Smoky

Smoky is a development resource provided to users needing a comparable system to the larger resources for application development. Smoky's current configuration is an 80-node Linux cluster consisting of four quad-core 2.0 GHz AMD Opteron processors per node, 32 GB of memory (2 GB per core), a gigabit Ethernet network with Infiniband interconnect, and access to Spider, the center-wide Lustre-based file system. Its primary purpose is application development, specifically application scaling development for petascale applications. Its programming environment mirrors the environment available on the Jaguar system. Only a limited number of center users are allowed access to this resource.

OLCF Groups

The OLCF is led by a management team experienced at guiding world-class systems. The management team has five principals: Director Jim Hack guides the overall vision of the center, while Project Director Buddy Bland supervises installation and upgrades of the OLCF supercomputers. Deputy Project Director Kathlyn Boudwin provides project management for a variety of computing and computing system installation and upgrade projects. Director of Science Jack Wells oversees the research teams that use the computing systems. Jim Rogers, as Director of Operations, manages day-to-day operations along with future systems and infrastructure planning. And Julia White, program manager for INCITE, manages one of the leading programs through which researchers gain access to OLCF resources.

The Application Performance Tools Group researches, tracks, purchases, and develops software tools that help researchers improve the performance of their applications on current and emerging OLCF computing systems. The group also manages contacts with vendors for the purchase of new modeling tools, languages, middleware, and performance-characterization tools. The group primarily focuses on issues that arise for research applications when they are run on very large-scale systems such as Jaguar. For more information, contact interim group leader David Bernholdt at bernholdtde@ornl.gov.

The User Assistance and Outreach Group has four responsibilities. The first is user support by account management liaisons and user assistance analysts. The second is effective communication to users, stakeholders, and vendors through emails, posters, and brochures. The third is training for current and future users. And the fourth is outreach, by which information about the facility and its work are shared with external media and the public. Outreach activities also attract new users to the center through tours, summer internships, and education opportunities; in 2011, the OLCF conducted 13 workshops with approximately 700 on-site participants and also began webcasting select training events. These outreach activities attract the next generation of users as well as the next generation of personnel into the OLCF. For more information, contact group leader Ashley D. Barker at ashley@ornl.gov.

The members of Scientific Computing group (SciComp), generally Ph.D.-level scientists, work with INCITE-project researchers to maximize their science output, often serving as embedded members for code development and scientific collaboration. SciComp also provides INCITE projects with visualization specialists, whose work supports the projects' data analysis. SciComp is concerned with the next generation of machines as well, and members work to analyze the performance of codes on current platforms to help inform the design of future systems. Finally, SciComp leads new-machine acceptance efforts, making sure that procured machines meet the users' and the center's requirements. For more information, contact acting group leader Bronson Messer at bronson@ornl.gov.

The High Performance Computing Operations Group (HPCO) keeps the leadership-class supercomputers running. To ensure optimal operations, group members monitor all systems around the clock, anticipating problems before they arise. HPCO is responsible for administration, configuration management, and cyber security; during installations and upgrades, group members rigorously test systems and continually monitor them with diagnostic tools. The group also ensures that all systems conform to ORNL cyber security policy. For more information, contact group leader Kevin Thach at kthach@ornl.gov.

The Technology Integration Group (TechInt) is responsible for updating and integrating the networks, file systems, and archival storage infrastructure into the OLCF computing systems. The group evaluates emerging technologies and provides system programming to seamlessly integrate new technologies and tools into the infrastructure as they are adopted. TechInt co-developed the OLCF's HPSS and is constantly working to increase the speed of data transfer for the OLCF's network. For more information, contact acting group leader Al Geist at gst@ornl.gov.

Education, Outreach, and Training

The OLCF is committed to education year-round. The facility hosted a plethora of classes in 2011—both on site and via webinars—to enhance the experience of its users and strengthen their ability to make the most of leadership-class computational resources.

Workshop on Exascale Data Management, Analysis, and Visualization—February 22–23

DOE and the OLCF co-hosted this workshop in Houston for computational scientists and data experts. Application scientists worked with experts in scientific data analysis and visualization to determine the support necessary to develop scientific discovery at the exascale. Participants produced a report suggesting future directions for data analysis, scientific visualization, and data management to maximize future scientific discovery.

http://www.olcf.ornl.gov/event/exascale-2011/

OLCF/NICS Spring Training—March 7–11

Users of two of ORNL's supercomputers, Jaguar and Kraken, honed their skills at a weeklong "Spring Training" at ORNL. The training, sponsored by the DOE-sponsored OLCF and NSF-sponsored National Institute for Computational Sciences (NICS), offered opportunities for expert users and novices alike. OLCF and NICS experts shared their expertise in using compilers on the systems, optimizing a code's I/O performance, choosing the right libraries, and debugging code. Professionals from Cray and other vendors were also on hand to instruct attendees and answer questions during lab times. The training was followed by the annual OLCF Users' Meeting.

http://www.olcf.ornl.gov/event/olcf-nics-springtraining/

OLCF/NICS Fall Training—October 18–19

In this training, NICS and OLCF staff provided introductory materials to inexperienced users, covering basic utilization of the Oak Ridge leadership-class resources. This event was simultaneously webcast for those unable to attend in person. Topics included connecting, compiling, submitting jobs, running jobs, file systems, software, modules, visualization, and data transfer.

http://www.olcf.ornl.gov/event/olcf-nics-fall-training-2011/

Two Summer Courses for ORNL Interns and Staff

HPC Fundamentals Course—June 16–August 3. An eight-week summer course on HPC fundamentals introduced students and interns with no prior experience to the skills needed to access HPCs, develop applications for them, and execute developed applications.

Crash Course in Supercomputing—June 16. The Crash Course in Supercomputing gave participants—interns and staff with no prior exposure—an overview of HPC concepts and techniques in a one-day intensive class. The course was offered in two sections. In Section 1, Students learned to program, compile, and run code on the Unix operating system. Section 2 covered parallel programming concepts. Students learned to write parallel programs, compile, and run a parallel code on one of the NCCS supercomputers.

http://www.olcf.ornl.gov/event/crash-course-in-supercomputing-summer-2011/

STEM Experience

The OLCF welcomed secondary school students to the ORNL campus for workshops and research experiences. The OLCF has worked to promote science, technology, engineering, and math (STEM) education by providing research experiences for middle- and high-school students.

In July, the Appalachian Regional Commission (ARC) and ORNL continued their partnership to engage the next generation of scientists by putting on two summer programs—a high-school math, science, and technology institute and a middle-school science camp.

High School Institute—July 9–22. For the ninth consecutive year, this summer institute provided high-school students and teachers the unique opportunity to work on inquiry-based, applied projects in science, math, and computer technology with ORNL staffers. The 33 students and 9 teachers from 13 Appalachian states were divided into teams for a challenging two weeks with ORNL/OLCF staffers. For example, teams working with the NCCS's Robert Whitten and ORNL's Bobby Sumpter built a small supercomputer to show how supercomputers are organized and programmed.

Middle School Science Camp—July 17–22. Oak Ridge Associated Universities joined ARC and ORNL to host the middle-school Science Camp for 25 students from 11 Appalachian states. At the end of the week, students presented their research projects, which focused on biofuels, solar power, and wind energy.

The goals of the ARC program are to (1) encourage more high-school students to continue their studies beyond high school; (2) encourage more students to pursue careers in the projected shortage areas of science, technology, engineering, mathematics, and computer sciences; and (3) raise the quality of math, science, and technology instruction in high schools throughout the region.

Titan Takes the Stage

The transformation of Jaguar into Titan challenged OLCF staff to step up the education opportunities offered to users. Accordingly, the OLCF created workshops and webinars to introduce the user community to the new opportunities that the hybrid system offers.

Titan Overview—July 26. The OLCF hosted the Titan Overview webinar to give attendees a glimpse of the OLCF roadmap, a Titan overview, and a description of future training opportunities to be provided by the OLCF. The event brought together OLCF staff, current and prospective users, and vendors to discuss Titan and its implications for application development. The open forum encouraged discussions of techniques, strategies, concerns, and expectations.

Titan Summit —August 15–17. Users, vendors, and OLCF staff shared plans and information regarding the OLCF's next major system at the Titan Summit workshop. The event focused on Titan's ability to enable groundbreaking research in climate, energy creation and storage, biology, chemistry, astrophysics, and materials science. On the summit's first day, ORNL staff presented case studies and methods for exposing parallelism in code. Vendors spoke on the second day, focusing on compilers, debuggers, and optimizers—software to help users with codes. On the final day, users talked about their applications and research topics. The overriding theme of the Titan Summit was the usability of Titan. Users learned that their current codes could run on Titan with little or no modification and that by using the provided tools, they could reap increased performances with little to no hand coding within the CUDA programming framework.

http://www.olcf.ornl.gov/event/titan-summit/

The Research Alliance in Math and Science

Students from eight colleges and universities were selected for summer internship projects ranging from network security, to many-body physics of strongly correlated electronic systems, to computational biology-related data mining, to parallel algorithms, to fluid structure interaction to name a few, through the Research Alliance in Math and Science (RAMS) program. Students were mentored by Computing and Computational Sciences Directorate research staff. The RAMS program is uniquely successful at identifying and mentoring underrepresented undergraduate and graduate students majoring in science, technology, engineering, and mathematics. Students become engaged in hands-on research projects that build technical skills and self-determination to seek advanced degrees. The RAMS program is sponsored by ASCR.

http://computing.ornl.gov/internships/rams/

SC11

In November, OLCF staffers headed to Seattle for the 23rd International Conference for High Performance Computing, Networking, Storage and Analysis, familiarly known as SC11. This annual conference gathers the supercomputing community, bringing together researchers, scientists, computing center staff members, IT and data center managers, application developers, vendors, and journalists. OLCF staff participated in poster sessions, distributed printed materials, exhibited videos, and led panel discussions, workshops, and tutorials. *—by Sandra Allen McLean*

Awards for OLCF Roll In

Dave Pugmire of OLCF won awards for his visualizations “Magnetic Field Outflows from Active Galactic Nuclei” (left) and “Magnetic Fields in Core-Collapse Supernovae” (center). Mike Matheson won for his visualization of “Shock Wave/Turbulent Boundary Layer interactions” (right).

OLCF staff are committed to supporting and achieving excellence every day. And throughout 2011, various judges and awards committees agreed. In numerous areas, the accomplishments of the OLCF have been recognized and heralded as outstanding.

Visualization Awards. At the Electronic Visualization Night of the annual Scientific Discovery through Advanced Computing (SciDAC) conference, three animations created by Scientific Computing Group visualization specialists Dave Pugmire and Mike Matheson took the top awards. The OLCF duo won a total of three OASCRs (Office of Advanced Scientific Computing Research awards) and two juried “Best Presentations.”

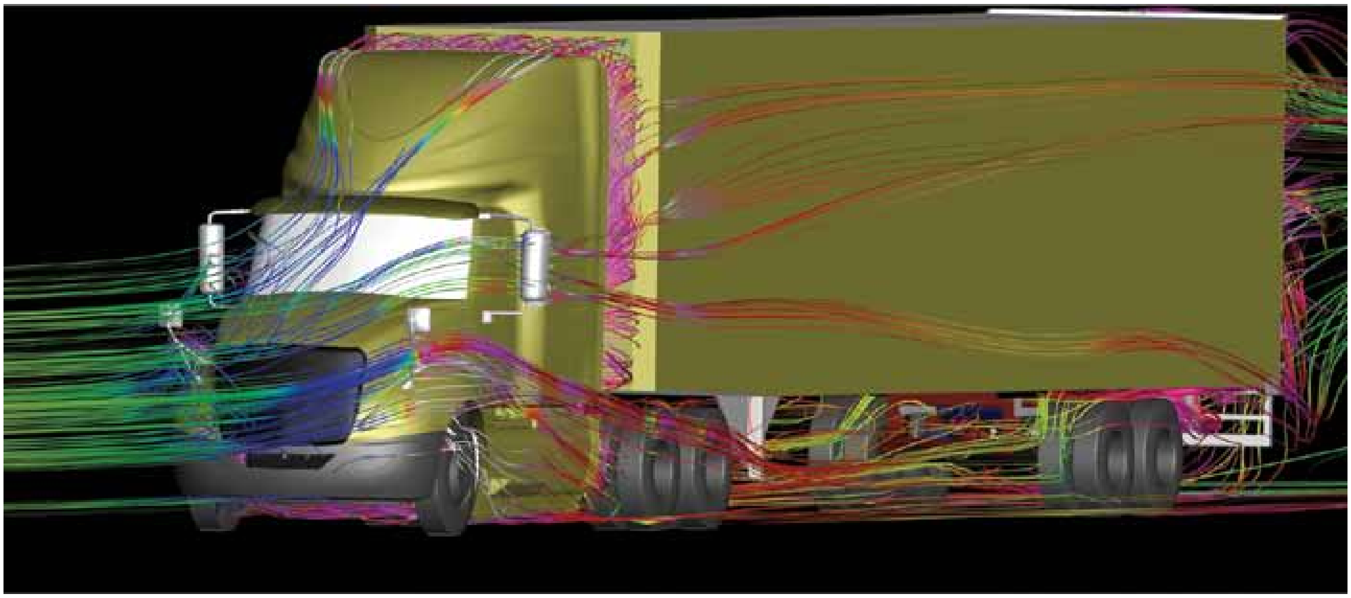
Pugmire received People’s Choice OASCR statuettes for his two movies on vector field analysis. His visualization of magnetic field generation, titled “Magnetic Field Outflows from Active Galactic Nuclei,” was also judged “Best Visual Presentation” in the juried competition. Pugmire’s second OASCR winner, “Magnetic Fields in Core-Collapse Supernovae,” depicted the magnetic field inside the shock surface of a supernova. Matheson’s entry, “Shock Wave/Turbulent Boundary Layer Interactions,” which showed complex fluid flows, received both the People’s Choice OASCR and the juried “Best Presentation of Information” awards.

The Viz Night competition is an annual event at the SciDAC conferences, with conference attendees voting on the submissions and OASCR statuettes going to the top 10 vote-getters. A new competition category was added this year, with three judges considering all submissions for two “best” place prizes.

Gordon Bell Honorable Mention for OMEN. At the SC11 supercomputing conference, the Gordon Bell Prize competition awarded three honorable mentions for the first time. One of the winners, a nanoscale simulation of electronic devices, ran on ORNL’s Jaguar supercomputer and reached 1.44 quadrillion calculations a second, or 1.44 petaflops, using 221,400 of 224,000 processing cores. The team included researchers from Purdue University, the University of Alabama–Huntsville, and Switzerland’s ETH Zurich. Team member Gerhard Klimeck of Purdue University described their application, OMEN, as the first electrical engineering software, maybe even the first engineering code of any type, that runs at the petascale, allowing researchers to explore electronic design spaces with repeated simulations. The Gordon Bell Prize is presented each year to developers of the world’s most advanced scientific supercomputing application.

Aprà Voted One to Watch. ORNL computational chemist Edoardo Aprà was named one of 10 “People to Watch” for their impact and influence on the future of HPC. The group was selected by the publishers of HPC online magazine *HPCwire*. Aprà also received the *HPCwire* Reader’s Choice Award in supercomputing achievement for his work with the computational chemistry application NWChem, developed at Pacific Northwest National Laboratory. Under Aprà’s guidance, NWChem joined the rarefied community of petascale applications, reaching 1.39 petaflops.

Mentors Recognized by DOE. OLCF staff members James Rogers and Bobby Whitten received Outstanding Mentor Awards from DOE in May. Coordinated by DOE’s Office of Science Workforce Development for Teachers and Scientists, the award recognizes mentors for their personal dedication to preparing students for careers in science and science education through well-developed research projects. Rogers, who is the director of operations for the OLCF, mentored a chemical



Trailers equipped with SmartTruck UnderTray components can achieve between 7 and 12 percent improvements in fuel mileage. The Jaguar supercomputer assisted BMI Corporation in its development of the product, earning an HPC Innovation Excellence Award for the work.

engineering junior from Tennessee Technological University for two years. Whitten—a member of the OLCF's User Support Group—acts as a mentor in two programs aimed at educators and students. The DOE-sponsored ACTS (Academies Creating Teacher Scientists) program helps high school teachers grow as leaders of STEM education. To better integrate the practice of science into their curricula, the teachers work with mentors in one-on-one training at national laboratories. The ARC sponsors an annual summer science camp for high school students and a math-science-technology institute for students and teachers. Since 2008, Whitten has mentored 22 students in both the ACTS and the ARC programs.

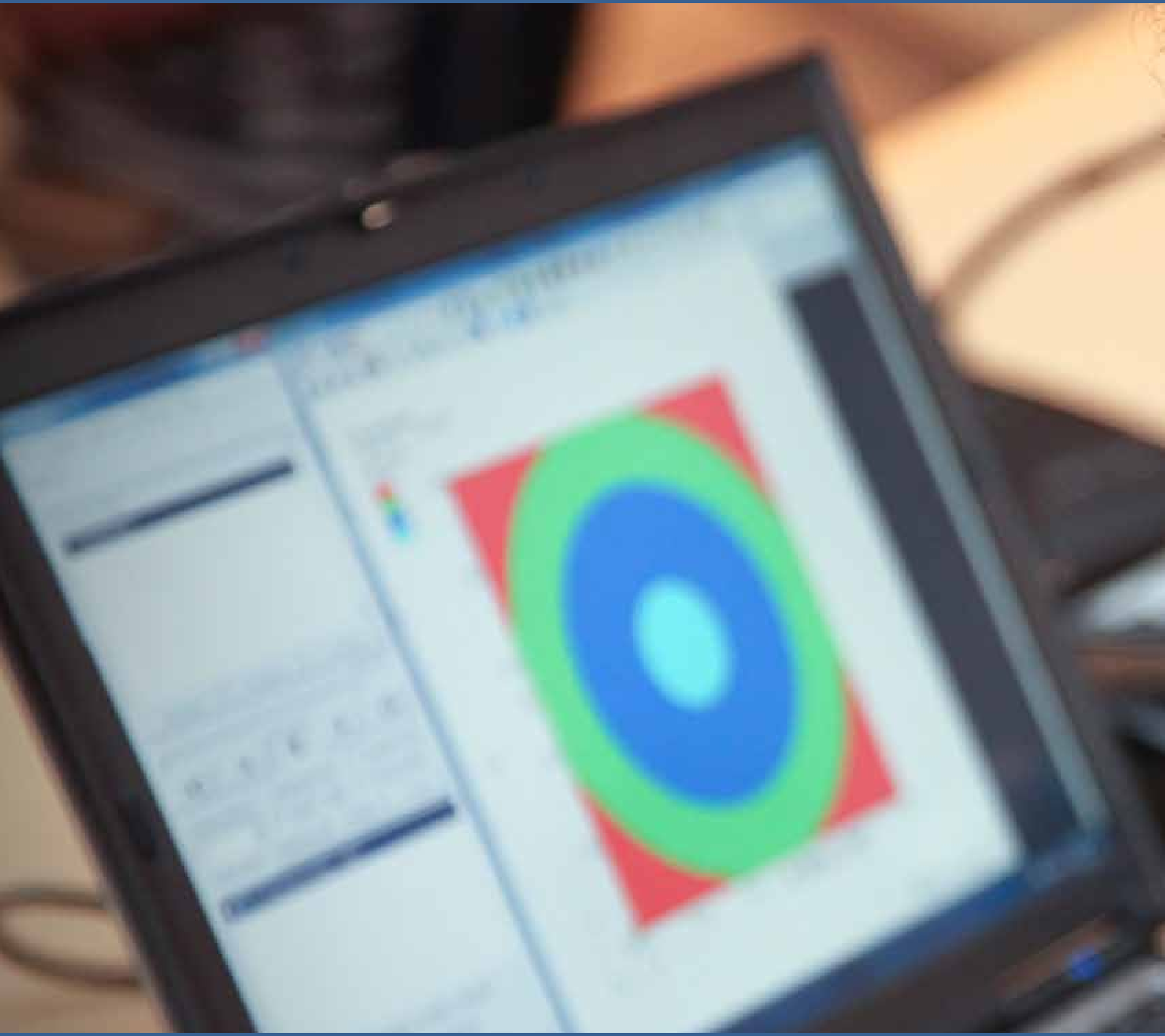
> In numerous areas, the accomplishments of the OLCF have been recognized and heralded as outstanding.

ORNL Helps Geology Unit Win Award. The Software and Information Industry Association presented a "CODiE" award to the JASON Project as the nation's best science or health curriculum. ORNL staff members contributed knowledge and support on this project's geology unit. The curriculum unit, *Operation: Tectonic Fury*, helps middle school students solve geological mysteries by researching and analyzing the Earth's past, present, and future through multimedia activities. The JASON Project, a nonprofit organization sponsored by the National Geographic Society, physically and virtually connects students and teachers with researchers to provide enriching scientific experiences.

MADNESS Is an R&D Winner. A team led by Robert Harrison, a computational chemist at ORNL and the University of Tennessee–Joint Institute for Computational Sciences, received an R&D 100 Award from *R&D Magazine* for the development of the "Multi-resolution Adaptive Numerical Environment for Scientific Simulations," or MADNESS. MADNESS is a free, open-source, general-purpose, user-friendly software package for developing scientific simulations. It runs on a range of systems, from laptops to massively parallel supercomputers. MADNESS utilizes the latest parallel computing and solution methodologies to solve many-dimensional integral and differential equations accurately and precisely for real-world problems. The application was developed by an ORNL group directed by George Fann, of ORNL's CSM Division, and Harrison. The team also included Rebecca Hartman-Baker of the OLCF's Scientific Computing Group.

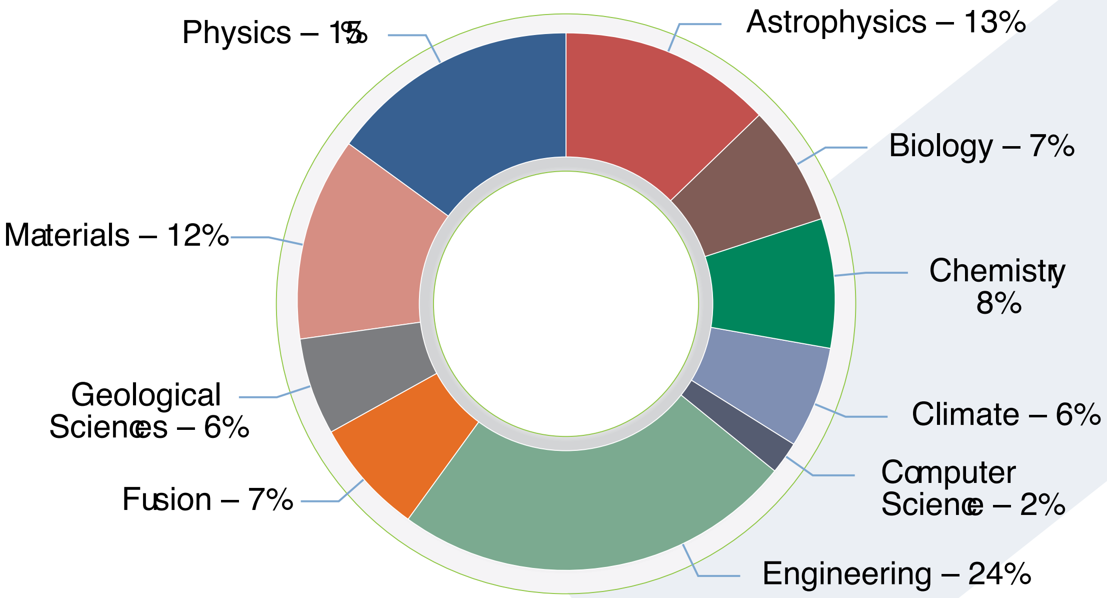
SmartTruck Simulations Earn Innovation Award. One of the nine 2011 HPC Innovation Excellence Awards was presented to Jaguar user Mike Henderson, CEO of BMI Corporation and Smart Truck. Henderson and colleagues at BMI ran simulations on Jaguar to design new add-on parts for long-haul trucks that dramatically decrease drag and increase fuel efficiency, resulting in an estimated fuel savings of $5,000 per truck per year. A key part of the process was creating the most complex model of a trailer to date and studying the airflow around it using NASA's Full Unstructured Navier Stokes, or FUN3D, application. Trailers equipped with BMI Corp. SmartTruck UnderTray components can achieve between 7 and 12 percent improvements in fuel mileage. International Data Corporation sponsors the awards, given to organizations achieving an important, quantifiable achievement with the help of HPC.—*by Sandra Allen McLean*

Appendices

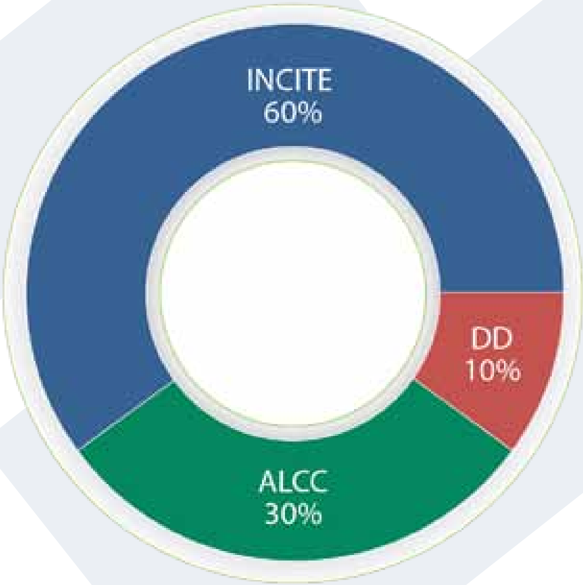


Center Stats

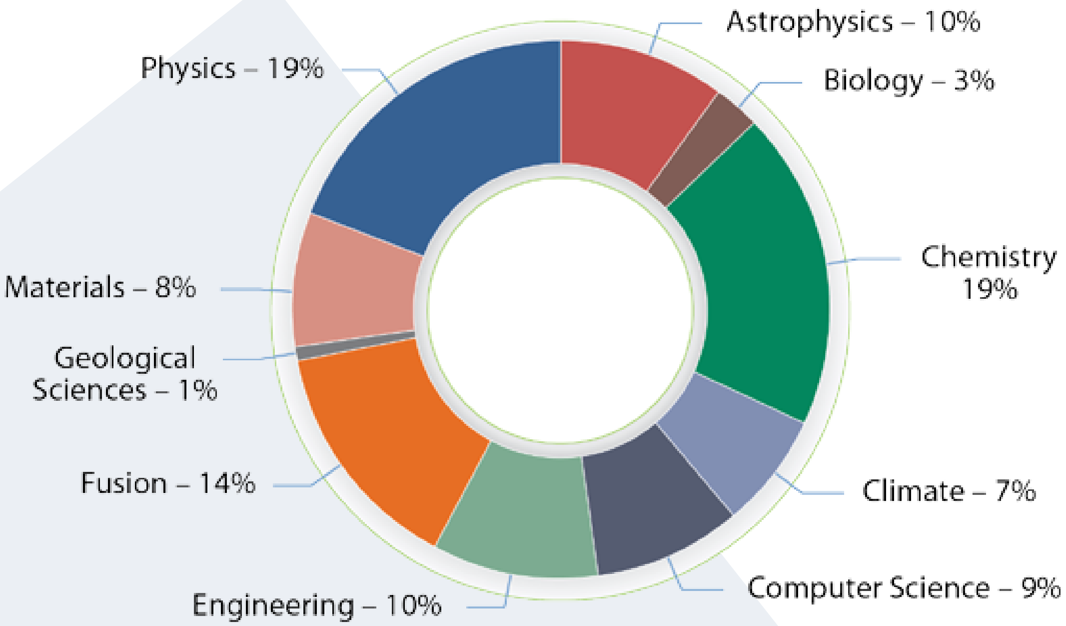
Jaguar Allocation by Domain 2012



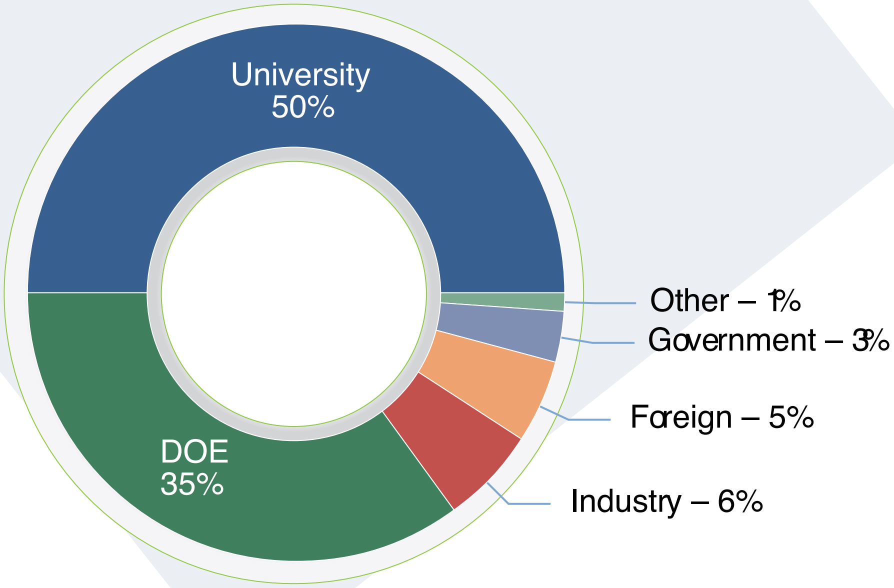
Allocation Hours on Jaguar 2011



Jaguar Usage by Domain 2011



User Allocation on Jaguar 2011



INCITE 2011

doeleadershipcomputing.org

Three-dimensional Simulations for Core Collapse Supernovae:
Anthony Mezzacappa, Oak Ridge National Laboratory
60,000,000 hours

“Turbulent Heating of Astrophysical Plasmas”
Gregory Howes, University of Iowa
10,000,000 hours

“Petascale Simulations of Type 1a Supernovae from Ignition to Observables”
Stan Woosley, University of California—Santa Cruz
50,000,000 hours

“How High Redshift Galaxies Reionized the Universe”
Michael Norman, University of California—San Diego
35,000,000 hours

“Sculpting Biological Membranes by Proteins”
Klaus Schulten, University of Illinois at Urbana—Champaign
5,000,000 hours

“Cellulosic Ethanol: Simulation of Multicomponent Biomass System”
Jeremy Smith, Oak Ridge National Laboratory
30,000,000 hours

“Control of Complex Transformations with Advanced Molecular Simulation Methods”
Christopher Mundy, Pacific Northwest National Laboratory
20,000,000 hours

“Petascale Modeling of Chemical Catalysts and Interfaces”
Robert Harrison, Oak Ridge National Laboratory
75,000,000 hours

“Coarse Grained Molecular Dynamics Studies of Vesicle Formation and Fusion”
Michael Klein, Temple University
30,000,000 hours

“CHiMES: Coupled High-Resolution Modeling of the Earth System”
Venkatramani Balaji, NOAA/GFDL, Princeton University
20,000,000 hours

“Climate-Science Computational Development Team: The Climate End Station II”
Warren Washington, National Center for Atmospheric Research
70,000,000 hours

“Performance Evaluation and Analysis Consortium End Station”
Patrick Worley, Oak Ridge National Laboratory
20,000,000 hours

“Simulation of Turbulent Lean Hydrogen Flames in High Pressure”
John Bell, Lawrence Berkeley National Laboratory
40,000,000 hours

“High-Fidelity Simulations for Advanced Engine Combustion Research”
Joseph Oefelein, Sandia National Laboratories
60,000,000 hours

“Petascale Simulation of Nano-Electronic Devices”
Gerhard Klimeck, Purdue University
15,000,000

“Quantum Monte Carlo Brings New Realism to Surface-Science Modeling”
Dario Alfe, University College London
17,000,000 hours

“Magnetic Structure and Thermodynamics of Low Dimensional Magnetic Structures”
Markus Eisenbach, Oak Ridge National Laboratory
50,000,000 hours

“Explosive Hazard Predictions with the Uintah Framework”
Martin Berzins, University of Utah
15,000,000 hours

“Singularities and Multi-Scaling in MHD”
Annick Pouquet, National Center for Atmospheric Research
15,000,000 hours

Understanding the Ultimate Battery Chemistry: Rechargeable Lithium/Air
Jack Wells, Oak Ridge National Laboratory
10,000,000 hours

“Gyrokinetic Simulation of Energetic Particle Turbulence in ITER Burning Plasmas”
Zhihong Lin, University of California—Irvine
35,000,000 hours

“High-Fidelity Tokamak Edge Simulation for Efficient Confinement of Fusion Plasma”
C.S. Chang, New York University
50,000,000 hours

"Investigation of Multi-Scale Transport Physics of Fusion Experiments Using Global Gyrokinetic Turbulence Simulations"
Weixing Wang, Princeton Plasma Physics Laboratory
20,000,000 hours

"Validation of Plasma Microturbulence Simulations for Finite-Beta Fusion Experiments"
William Nevins, Lawrence Livermore National Laboratory
20,000,000 hours

"Ultrascale Simulation of Basin-Scale CO_2 Sequestration in Deep Geologic Formations and Radionuclide Migration Using PFLOTRAN"
Peter Lichtner, Los Alamos National Laboratory
15,000,000 processor hours

"Electronic Structure Calculations for Nanostructures"
Lin-Wang Wang, Lawrence Berkeley National Laboratory
10,000,000 hours

"Quantum Monte Carlo Simulation of Models of Condensed Matter"
Richard Needs, University of Cambridge
15,000,000 hours

"Uncertainty Quantification for Three-Dimensional Reactor Assembly Simulations"
Thomas Evans, Oak Ridge National Laboratory
18,000,000 hours

"Petascae Particle-In-Cell Simulations of Plasma Based Accelerators"
Warren Mori, University of California—Los Angeles
12,000,000 hours

"Unraveling the Physics of Magnetic Reconnection with 3D Kinetic Simulations"
William Daughton, Los Alamos National Laboratory
30,000,000 hours

"Nuclear Structure and Nuclear Reactions"
James Vary, Iowa State University
43,000,000 hours

Petascale Computing for Terascale Particle Accelerator: International Linear Collider Design and Modeling
Paul Mackenzie, Fermilab
30,000,000 hours

ALCC 2011

science.energy.gov/ascr/facilities/alcc

High Resolution Design-Cycle CFD Analysis, Supporting CO_2 Compression Technology Development
Allan Grosvenor, Ramgen Power Systems
Jaguar: 700,000 hours

Understanding the Factors that Affect the Efficiency of Bio-catalytic Processes
Pratul Agarwal, Oak Ridge National Laboratory
Jaguar: 4,000,000 hours

High Fidelity Simulations of Gas Turbine Combustors for Low Emissions
Venkat Tangiral, GE Global Research
Jaguar: 10,000,000 hours

DOE SC ASCR Joule Software Metric
Jack Wells, Oak Ridge National Laboratory
Jaguar: 100,000,000 hours

Petascale Kinetic Plasma Simulation of the Interaction among Laser Speckles in Laser-driven Inertial Fusion Energy Settings
Brian Albright, Los Alamos National Laboratory
Jaguar: 15,000,000 hours

Uncertainty Quantification in Large-Scale Ice Sheet Modeling and Simulation
Omar Ghattas, The University of Texas at Austin
Jaguar: 10,700,000 hours

Petascale Quantum Monte Carlo Calculations of Strongly Correlated and Energy Storage Materials
Paul Kent, Oak Ridge National Laboratory
Jaguar: 15,000,000 hours

Advanced Simulation of Light Water Reactors
John Turner, Oak Ridge National Laboratory
Jaguar: 30,000,000 hours

Reliable Predication of Performance of High Lift Systems of Commercial Air
John Bussoletti, Boeing
Jaguar: 13,000,000 hours

Non-icing Surfaces for Cold Climate Wind Turbines
Masaka Yamada, GE Global Research
Jaguar: 40,000,000 hours

Large-eddy Simulation for Turbomachinery—Advancing the State-of-the art
Gorazd Medic, United Technologies Research Center
Jaguar: 7,000,000 hours

Controlling Nanoparticle Interactions to Engineer New Materials
Gary Grest, Sandia National Laboratories
Jaguar: 30,000,000 hours

High-Impact Publications

Astrophysics

Howes GG, TenBarge JM, Dorland W, Quataert E, Schekochihin AA, Numata R, Tatsuno T, "Gyrokinetic Simulations of Solar Wind Turbulence from Ion to Electron Scales," *Physical Review Letters* **107** (3), 035004 (2011).

Pogorelov NV, Heerikhuisen J, Zank GP, Borovikov SN, Frisch PC, McComas DJ, "Interstellar Boundary Explorer Measurements and Magnetic Field in the Vicinity of the Heliopause," *Astrophysical Journal* **742** (2), 104 (2011).

Woosley SE, Kasen D, "Sub-Chandrasekhar Mass Models for Supernovae," *The Astrophysical Journal* **734** (38), 1—38 (2011).

Biology

Hsin J, LaPointe LM, Kazy A, Chipot C, Senes A, Schulten K, "Oligomerization State of Photosynthetic Core Complexes is Correlated with the Dimerization Affinity of a Transmembrane Helix," *Journal of the American Chemical Society* **133** (35) 14071—14081 (2011)

Petridis L, Pingali SVm Urban V, Heller WT, O'Neil HM, Foston M, Ragauskas A, Smith JC, "Self-similar Multiscale Structure of Lignin Revealed by Neutron Scattering and Molecular Dynamics Simulation," *Physical Review E* **83** (6), 061911 (2011).

Ramanathan A, Agarwal PK, "Evolutionarily Conserved Linkage between Enzyme Fold, Flexibility, and Catalysis," *PLOS Biology* **9** (11), e1001193 (2011).

Chemistry

Baer MD, Mundy CJ, "Toward an Understanding of the Specific Ion Effect Using Density Functional Theory," *Journal of Physical Chemistry Letters* **2** (9), 1088—1093 (2011).

Goswami M, Kumar R, Sumpter BG, Mays J, "Breakdown of Inverse Morphologies in Charged Diblock Copolymers," *Journal of Physical Chemistry B* **115** (13), 3330—3338, (2011).

Climate

Compo GP, Whitaker JS, Sardeshmukh PD, Matsui N, et al., "The Twentieth Century Reanalysis Project," *Quarterly Journal of the Royal Meteorological Society* **137** (654), 1—28 (2011).

Long MS, Keene WC, Kieber DJ, Erickson DJ, Maring H, "A Sea-state Based Source Function for Size- and Composition-Resolved Marine Aerosol Production," *Atmospheric Chemistry and Physics* **11** (3), 1203—1216 (2011).

Meehl GA, Arblaster JM, Fasullo JT, Hu A, Trenberth KE, "Model-Based Evidence of Deep-Ocean Heat Uptake During Surface-Temperature Hiatus Periods," *Nature Climate Change* **1**, 360—364, (2011).

Fusion

Guttenfelder W, Candy J, Kaye SM, Nevins WM, Wang E, Bell RE, Hammett GW, LeBlanc BP, Mikkelsen DR, Yuh H, "Electromagnetic Transport from Microtearing Mode Turbulence," *Physical Review Letters* **106** (15), 155004 (2011).

Nevins WM, Wang E, Candy J, "Magnetic Stochasticity in Gyrokinetic Simulations of Plasma Microturbulence," *Physical Review Letters* **106** (6), 065003 (2011).

Wang WX, Hahm TS, Ethier S, Zakharov LE, and Diamond PH, "Trapped Electron Mode Turbulence Driven Intrinsic Rotation in Tokamak Plasmas," *Physical Review Letters* **106**, 085001 (2011).

Materials

Cao D, Pang P, He J, Luo T, Park JH, Krstic P, Nuckolls C, Tang JY, Lindsay S, "Electronic Sensitivity of Carbon Nanotubes to Internal Water Wetting," *ACS Nano* **5** (4), 3113—3119 (2011).

Dag S, Wang SZ, Wang LW, "Large Surface Dipole Moments in ZnO Nanorods," *Nano Letters* **11** (6), 2348—2352 (2011).

Lopez-Bezanilla A, Huang JS, Terrones H, Sumpter BG, "Boron Nitride Nanoribbons Become Metallic," *Nano Letters* **11** (8), 3267—3273 (2011).

Santra B, Klimes J, Alfe D, et al., "Hydrogen Bonds and van der Waals Forces in Ice at Ambient and High Pressures," *Physical Review Letters* **107** (18) 185701 (2011).

Shao MH, Peles A, Shoemaker K, "Electrocatalysis on Platinum Nanoparticles: Particle Size Effect on Oxygen Reduction Reaction Activity," *Nano Letters* **11** (9), 3714-3719 (2011).

Steiger S, Povolotskyi M, Park HH, Kubis T, Klimeck G, "NEMO5: A Parallel Multiscale Nanoelectronics Modeling Tool," *IEEE Transactions on Nanotechnology* **10** (6) 1464—1474 (2011).

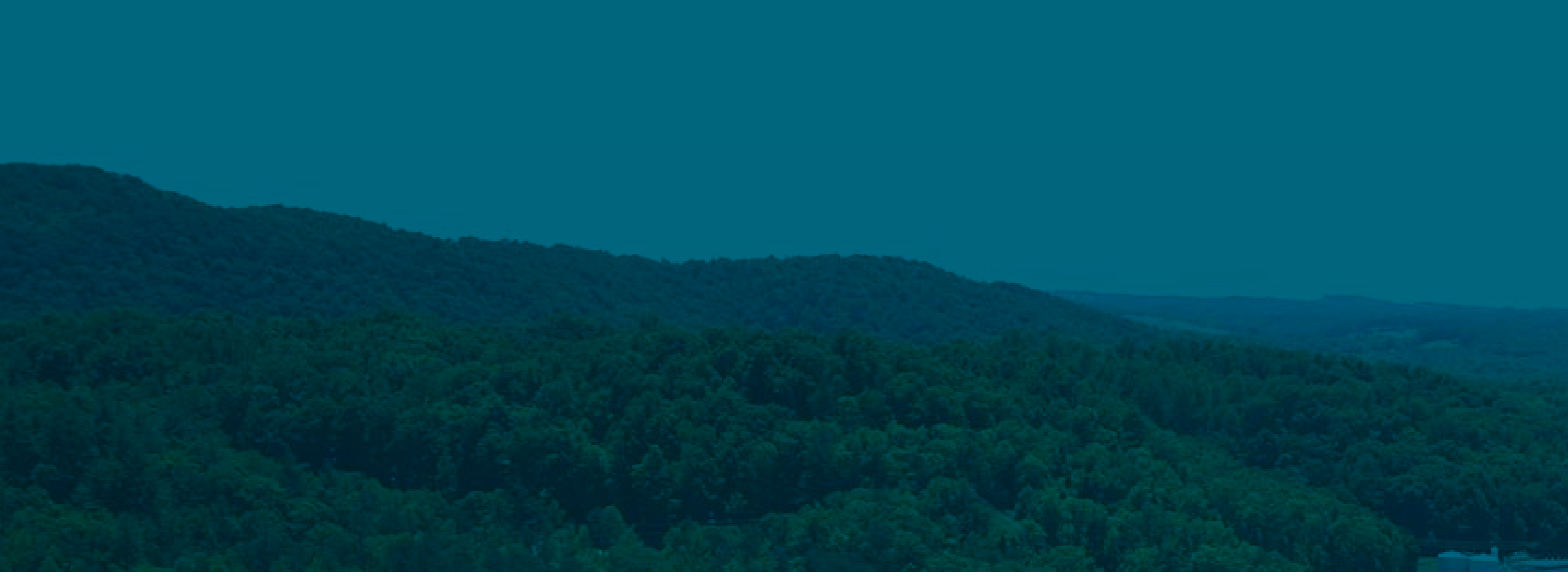
Physics

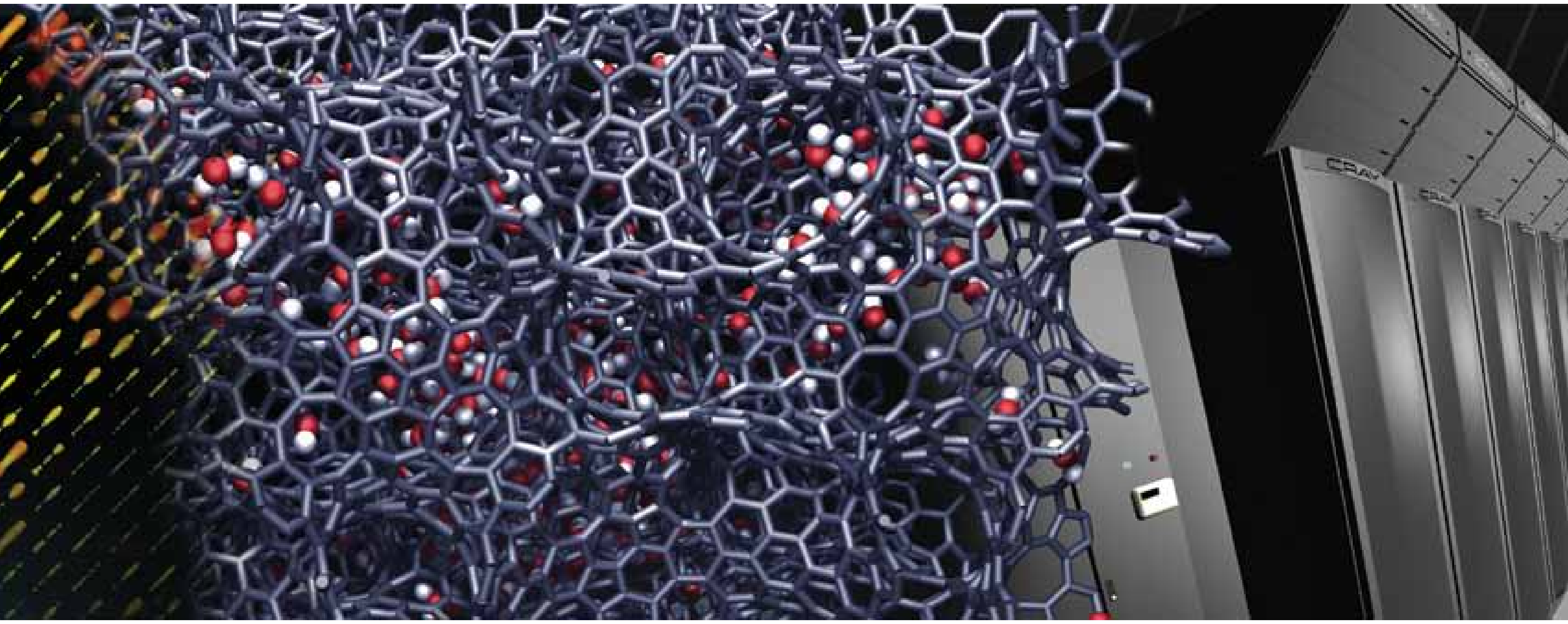
Bulgac A, Luo YL, Magierski P, Roche KJ, Yu YL, "Real-Time Dynamics of Quantized Vortices in a Unitary Fermi Superfluid," *Science* **332** (6035), 1288—1291 (2011).

Jensen O, Hagen G, Hjorth-Jensen M, et al. "Quenching of Spectroscopic Factores for Proton Removal in Oxygen Isotopes," *Physical Review Letters* **107** (3), 032501 (2011).

Maris P, Vary JP, Navratil P, Ormand WE, Nam H, Dean DJ, "Origin of the Anomalous Long Lifetime of C-14," *Physical Review Letters* **106** (20), 202502 (2011).

Vieira J, Martins SF, Pathak VB, Fonesca RA, Mori WB, Silva LO, "Magnetic Control of Particle Injection in Plasma Based Accelerators," *Physical Review Letters* **106** (22), 225001 (2011).




CRAY