National Center for Computational Sciences Snapshot The Week of October 15, 2007

<u>Dissecting Fire: Model Fully Resolves Diesel Ignition Process for the First Time</u>
Research may fuel low-temperature combustion technology to save energy, cut emissions

To flip an old saying, where there's fire, there's smoke. No one knows that better than Jacqueline Chen, a mechanical engineer at Sandia National Laboratories (SNL) who employs some of the world's fastest supercomputers to model combustion.

With mechanical engineer Chun Sang Yoo at SNL in California and computational scientist Ramanan Sankaran at the National Center for Computational Sciences (NCCS) in Tennessee, Chen used the Cray XT4 Jaguar supercomputer at the NCCS to generate 35 terabytes (trillion bytes) of data about flames similar to those occurring during ignition and stabilization of diesel-engine jets.

Engineers are using Chen's data library to develop predictive models to optimize designs for diesel engines and industrial boilers with reduced emissions and increased efficiency. Diesel fuel powers most semi trucks; delivery vehicles; buses; trains; boats; as well as farm, construction, and military vehicles in the United States, so development of advanced diesel technology is a leading near-term option by which the country could reduce its fuel consumption and greenhouse gas emissions.

"If low-temperature compression ignition systems employing lean, dilute fuel mixtures make their way into next-generation autos, fuel efficiency could increase by as much as 25 to 50 percent," Chen says. That also means meeting future low-emission vehicle standards with almost undetectable emissions of nitrogen oxide, a major contributor to smog, she adds.

Chen, Yoo, and Sankaran created the first three-dimensional simulation that fully resolves flame features such as chemical composition, temperature profile, and flow characteristics. The model shows ignition feature detail on all size scales—the biggest, the smallest, and everything in between—of a turbulent fuel jet in a hot co-flowing airstream.

The researchers modeled in unprecedented detail what happens in the so-called "lifted flames" relevant to industrial boilers and diesel engines. Unlike spark-plug ignition systems in automobiles powered by gasoline in which the fuel and oxidizer (air) are premixed, diesel-injection systems have the diesel fuel entering the engine full of hot air via jet nozzles. Turbulence mixes the fuel and air. Pistons subject the air/fuel mixture to

pressure, and the mixture heats further, spurring a chemical pathway that sharply increases the concentration of a highly reactive chemical, hydroperoxyl radical, Chen says. The hydroperoxyl radical produces heat that spurs the production of other radicals. At about 1,520 degrees Fahrenheit, the fuel/air mixture auto-ignites, or bursts into flame, as a result of the rapid, heat-producing oxidation of its own constituents, regardless of heat from external sources. This process creates a lifted flame. The temperature peaks around 3,140 degrees Fahrenheit.

“Auto-ignition is helpful because it stabilizes the flame,” Chen says. “A high-temperature flame base supports its existence.”

Before this work by Chen and her colleagues, scientists had modeled only large eddies, or turbulent curlicues, in a burning fuel. They had not simulated the full range of scales down to the smallest eddies, which dictate the viscosity of the system and dissipate heat. Now researchers can resolve the nitty-gritty of the small eddies responsible for flame extinction and reignition.

The Department of Energy (DOE) Office of Basic Energy Sciences and DOE Scientific Discovery through Advanced Computing (SciDAC) program supported the work at SNL. The DOE Office of Science supported the work at the NCCS, located at DOE’s Oak Ridge National Laboratory (ORNL), through its Innovative and Novel Computational Impact on Theory and Experiment (INCITE) program. INCITE grants huge allocations of supercomputing time to a handful of scientists addressing grand challenges in physics, chemistry, biology, and beyond. The NCCS is unique in providing Ph.D.-level computational scientists, such as Sankaran, to optimize tools, help with data analysis, improve algorithms, and aid in other ways so researchers get the most out of high-performance computers.

ORNL Researchers Contribute to 2007 Nobel Peace Prize

Computer at ORNL a key player in U.N. report

When you think of winners of the Nobel Peace Prize, you’re more likely to think of Mahatma Ghandi than the IBM pSeries Cheetah high performance computing system at the NCCS, a DOE Office of Science user facility housed at ORNL.

Gandhi never won the peace prize, though he was nominated in 1937, 1938, 1939, 1947, and a few days before his assassination in 1948. And true, a supercomputer can never be a Nobel laureate. But Cheetah provided more than one-third of the simulation data for the joint DOE/National Science Foundation (NSF) data contribution to the United Nation’s Intergovernmental Panel on Climate Change (IPCC) Fourth Assessment Report (AR4), a feat that earns it a place in world history.

This year the IPCC—a group of more than 2,000 scientists and policy experts—will share the 2007 Nobel Peace Prize with former Vice President Al Gore “for their efforts to build up and disseminate greater knowledge about man-made climate change, and to lay the foundations for the measures that are needed to counteract such change,” according to the Nobel announcement.

The IPCC AR4 effort is an unprecedented coordinated climate change assessment study involving 16 climate modeling centers from 11 countries around the world running the same set of climate change scenarios with 23 different models (see Figure 1). The IPCC AR4 produced four reports in 2007 that featured the involvement of ORNL scientists such as Corporate Fellows Tom Wilbanks and David Greene and the Environmental Sciences Division's Paul Hanson, Virginia Dale and Gregg Marland. The 31 terabytes (31 trillion bytes) of model data collected and distributed by the Program for Climate Model Diagnosis and Intercomparison at Lawrence Livermore National Laboratory using DOE's Earth System Grid have been accessed by more than 1,200 scientists, resulting in the more than 200 refereed papers that went into the IPCC AR4.

The ORNL Cheetah IPCC simulations were led by computational scientist John B. Drake of ORNL and atmospheric scientist Lawrence Buja of the National Center for Atmospheric Research (NCAR, an NSF center). In 2004, the NCCS granted the scientists a huge allocation of supercomputing time to run simulations on Cheetah, a system that performed calculations at a rate of 4 teraflops, or 4 trillion floating point operations per second. The runs occupied half of Cheetah's processors for the better part of a year and required the efforts of five ORNL staff members to help develop the model and two dozen NCCS staff members to move enormous amounts of simulation data. "We don't always know what the answer will be when we start a large computational study, or that anyone will pay much attention to the results, but in the case of the climate change modeling for the IPCC, it has had a big impact," Drake says. "Having the computer resources at ORNL made it possible to carry out a more comprehensive and detailed study than ever before. This improved the level of certainty for some of the conclusions in the IPCC reports and enabled breakthroughs to new climate results such as the prediction of regional heat waves."

These runs were part of a collaborative DOE/NSF project headed by NCAR's Warren Washington. Washington's project involved running the IPCC simulations at three centers—NCAR, ORNL's NCCS, and the National Energy Research Scientific Computing Center (NERSC, a DOE Office of Science user facility at Lawrence Berkeley National Laboratory). International collaborators extended these simulations using Japan's Earth Simulator Center, the most powerful supercomputer at the time. The IPCC runs were made with the Community Climate System Model (CCSM), a fully coupled global climate model that provides state-of-the-art simulations of the Earth's past, present, and future climate states. The CCSM is an open, communitywide, collaborative project involving NCAR, DOE (Los Alamos National Laboratory, ORNL, Lawrence Berkeley National Laboratory, Pacific Northwest National Laboratory, and Argonne National Laboratory), the National Aeronautics and Space Administration, and many university and industrial partners. With more than 10,000 model years simulated, at a higher resolution than had ever before been attempted at this scale of production, Washington's group made the largest data contribution to the IPCC project of any other center in the world.

"Access to DOE leadership-class, high-performance computing assets at ORNL and NERSC significantly improved model simulations," says NCAR's Buja. "These computers made it possible to run more realistic physical processes at higher resolutions, with more ensemble members, and longer historical validation simulations. We simply couldn't have done this without the strong DOE/NSF interagency partnership."

The DOE studies climate because of the close link between energy production, energy use, and carbon dioxide concentrations in the atmosphere. Climate simulations require the fastest computers because the projections must cover hundreds to thousands of years. The skill of the model projections is checked by comparing with the historical climate data of the past 100 years. Today's atmospheric carbon dioxide levels are 380 parts per million—and rising. "Even if levels were to stabilize at their current numbers, warming is projected to continue over the next century before achieving a steady state in following centuries," Drake says.

The IPCC AR4 is the culmination of a six-year international effort. It concludes the 1.5 degrees Fahrenheit of planetary warming observed during the 20^{th} century has a more than 90 percent chance of being the result of human activities. The authors predict large-scale changes in food and water availability, dramatic changes in ecosystems, and increased flooding and extreme weather. They advise quick action to avoid some devastating effects and say existing technologies can balance climate risks with economic competitiveness.

Blue Gene/P Comes to ORNL

Chemistry and materials applications show promise

ORNL and IBM are teaming to bring the next generation of the IBM Blue Gene supercomputer, the Blue Gene/P, to ORNL.

ORNL's new system was accepted in late September. It features 8,192 compute cores and will be capable of more than 27 trillion calculations a second, or 27 teraflops.

The Blue Gene/P system at ORNL comprises two cabinets, each holding 1,024 nodes. Each node contains four 850-MHz PowerPC 450 processors with 2 gigabytes of memory per node. Nodes are interconnected with three high-performances networks that will support a variety of interprocessor communication techniques—a three-dimensional torus for point-to-point message passing, a global collective network supporting one-to-many communication, and a global interrupt network enabling low-latency barriers and interrupts. The Blue Gene/P's low-latency communications networks will benefit researchers for which communication between processors is especially important. The system's quad-core configuration will also benefit users as multi-core architectures become more pervasive.

"Selected chemistry and materials applications especially have shown strong performance on the Blue Gene," said Thomas Zacharia, ORNL's associate laboratory director for computing and computational sciences. "We look forward to seeing researchers produce leading-edge science on this system."

“The Blue Gene/P system is ideally suited for spin-fermion Monte Carlo calculations over several thousand processors, which allow us to answer crucial questions about magnetoelectronics and spintronics,” said Thomas Schulthess, leader of the Computational Materials Science Group in ORNL’s Computer Science and Mathematics Division. “We also see potential for greater understanding of the statistical physics of fracture, which is focused on materials strength and scaling laws of fracture.”

The machine has been dubbed “Eugene,” after Hungarian physicist and mathematician Eugene Wigner, winner of the Nobel Prize in physics and an early advocate of the Manhattan project. Wigner would later become director of research and development at Oak Ridge National Laboratory.

New Storage Device Online at the NCCS

Extra storage keeps NCCS running

When running simulations of the intricacies of the human genome or climate change, supercomputers need to quickly access massive amounts of data. At the NCCS some of the world’s leading supercomputers manage this need for data using the High-Performance Storage System (HPSS).

The HPSS system recently received a helping hand with the addition of the Sun StorageTek SL8500 modular library system, now in use at the NCCS. With 24 T10000 (Titanium) tape drives in the new silo, HPSS is now able to store more overall data, giving one of the nation’s top supercomputing centers even more ammunition with which to tackle today’s Grand Challenge science.

As the need for storage space grows, its capacity can scale to thousands of petabytes ($1x10^{15}$ bytes) with the addition of disks and tapes. Today the system stores 1.6 petabytes of data—the hard-drive contents of a typical laptop 10,000 times over.

At peak performance, the HPSS can store a single stream of data as fast as 90 million bytes per second, the equivalent of storing 750 small novels every second. The average speed of a single data stream is 600 megabits per second. When the storage system is taxed with dealing with numerous requests at once, the speed at which data flows is 3 billion bits (gigabits) per second, says Stanley White, senior storage manager at the NCCS, located at Oak Ridge National Laboratory. With new equipment added this year, that rate will rise to 6 gigabits per second, he says.

HPSS is a collaborative effort of Lawrence Berkeley National Laboratory, Lawrence Livermore National Laboratory, Los Alamos National Laboratory, Sandia National Laboratories, ORNL, and IBM. HPSS is licensed to other users and supported by IBM under an agreement between IBM and the United States Department of Energy.

Spider Will Free Users from Web of Time-Consuming Chores

State-of-the-art file system will let scientists focus on science

NCCS users will find more time to focus on groundbreaking science with the arrival of a new centerwide file system.

Known as Spider, the one-stop file system will eliminate the need for users to move enormous amounts of data between multiple computer and data analysis systems to get their work done.

"What we want is for researchers to focus on the science and the discovery," said Shane Canon, leader of the NCCS Technology Integration Group. "We don't want them to worry about how to move data from system to system."

The system is arriving in stages, with the first already in place for limited testing. That first stage offers about 80 terabytes of disk space and a bandwidth of about 10 gigabytes a second. A second phase, planned for 2008, will increase the disk space to about 1 petabyte and bandwidth to more than 20 gigabytes per second. The third phase, which is expected in summer 2008, will increase storage to as much as 10 petabytes and bandwidth to 200 gigabytes a second. Staff members are currently performing head-to-head comparison tests of the latest storage technology from vendors LSI and DDN.

While the center's Cray XT4 Jaguar system will also have a separate file system, Canon said, the center's upcoming petascale system will use Spider exclusively. By keeping data in a single location, Spider will make it easier for researchers to take advantage of visualization and data analysis tools. If it increases the use of these tools, the system may even increase the value of that data by enhancing scientific discovery.

Canon noted that the system is eagerly anticipated, especially among researchers with data-intensive applications. This group includes astrophysicists, climate scientists, fusion researchers, and combustion researchers.

ORNL Users to 'Try On' Blue Gene

Blue Gene/P Workshop coming to ORNL later this month

ORNL and IBM have teamed to bring the next generation of the IBM Blue Gene/P supercomputer to the NCCS. The Blue Gene/P system configuration, with 8192 compute cores, will provide more than 27 TFLOPS of computing capability. The system is being prepared for production, and is scheduled for release to ORNL and its partners in early November.

The NCCS and IBM are presenting a one-day workshop to provide an introduction to the IBM BG/P hardware, software, and computing environment. The workshop will also describe the policies associated with obtaining access to the system, and demonstrate the features unique to this system that facilitate productive science.

The workshop will be held on October 31, 2007 at the American Museum of Science and Energy, in Oak Ridge Tennessee. Registration is required. There is no registration fee. Registration and an agenda can be accessed at http://nccs.gov/news/workshops/ibmworkshop/index.html.