National Center for Computational Sciences Snapshot
The Week of July 23, 2007

Researchers Use Jaguar to Tackle Groundwater

Project focuses on groundwater contamination and carbon storage

Earth scientists are taking advantage of the National Center for Computational Sciences (NCCS) Jaguar supercomputer to divine just what happens under our feet.

A project led by Peter Lichtner of Los Alamos National Laboratory is working to improve the science of groundwater modeling. As it succeeds, the project promises to improve efforts to clean up contaminated groundwater and even aid in the battle against global warming.

The project is applying a groundwater application known as PFLOTRAN to two specific cases: the journey of uranium in groundwater at the Department of Energy's Hanford Site and the possible storage of carbon dioxide in the groundwater near power plants. The task requires leadership computing because the scales range from millimeters to kilometers, and simulations must be able to track the groundwater for hours at a time.

At Hanford the team is examining groundwater at the 300 Area, a 1.5-square-mile site that housed an industrial complex and disposal areas next to the Columbia River about 1 mile north of Richland, Washington. The complex was used to fabricate fuel for nuclear reactors, and the groundwater is contaminated with uranium.

Previous studies failed to predict what would happen to the levels of groundwater contamination at the site.

"They did a study 15 years ago and concluded that within 10 years the uranium would be flushed out into the Columbia River," Lichtner noted. "Now it's 15 years later, and the uranium concentration still exceeds the maximum permissible levels. The plume is still as it was 15 years ago."

Lichtner's team is working to improve modeling of the site by taking into account factors that might have been left out of the earlier studies. For instance, one hypothesis suggests that uranium is locked in the interior of mineral grains before it leaches into the groundwater. Another issue is fluctuation of the river stage, which raises and lowers the water table at the site and changes the direction of flow to and from the Columbia River.

"It's a problem of getting a finer scale resolution, and it requires taking many time steps," Lichtner explained. "The fluctuations are on a scale of hours, so it's a perfect problem for massively parallel computing."

The other problem being tackled by the group involves storing carbon dioxide in groundwater near power plants. Carbon dioxide is the principal culprit in global warming, and the ability to sequester the chemical in groundwater might help provide some relief from climate change.

The team is looking at ways to model instabilities in the carbon dioxide as it develops into fingers in the groundwater. These fingers range from tens of centimeters to tens of meters. Lichtner noted that the success of carbon dioxide storage depends at least in part on the salinity of the groundwater, which reduces the amount of carbon dioxide it can hold—an issue being tackled in the simulations.

Lichtner said he has had a very positive experience working with the NCCS and with Richard Mills of the Scientific Computing Group, who is also a member of the project team.

"We've been very impressed running on Jaguar," he said, "and we're lucky to have Richard Mills as one of our members."

Jaguar Highlights Limits of Benchmark Test

High-Performance Linpack benchmark to be modified in response to ORNL run

The NCCS's Cray XT4 Jaguar supercomputer ran the largest problem ever seen for the test used to evaluate the world's fastest systems and in doing so revealed weaknesses in the test itself.

The Top500 List of the most powerful supercomputers ranks Jaguar as the world's second most powerful system. In May Jaguar demonstrated that it belongs at the top of the computing world by performing the largest ever High-Performance Linpack (HPL) benchmarking problem.

HPL was created for an earlier generation of supercomputers, and at first the May 2007 test appeared to generate an incorrect result, one that might have pointed to a hardware problem. In fact, Jaguar went through 2 days of hardware tests to verify that there were no hardware problems.

The explanation was eventually uncovered by mathematicians at manufacturer Cray, Inc., and computer scientists at the University of Tennessee's Innovative Computing Laboratory. They concluded that the test failed because the HPL pseudo-random number generator uses 32-bit mathematics, limiting the generator to 4 billion possible numbers, which caused the numbers to cycle in a way that caused the test to fail. Such cycling is increasingly common at very large matrix sizes, like those used on Jaguar.

When the HPL test was run again with a slightly larger problem, the system was able to solve a matrix containing nearly 5 trillion elements, nearly half again as big as the next largest problem. Because of its massive size, the test took 20 hours to complete, and Jaguar was able to achieve 101.7 trillion calculations a second (101.7 teraflops), or 86 percent of its peak performance of 119 teraflops.

"It is the largest problem that I know of that's ever been done," said Jack Dongarra of the University of Tennessee–Knoxville and Oak Ridge National Laboratory, cocreator of the list. Dongarra noted that adjustments will be made to the benchmark in response to Jaguar's results.

Users Share Advances—from Fusing Atoms to Exploding Stars

SciDAC 2007 brings together leading computational scientists

NCCS users and staff presented their achievements in nuclear fusion, astrophysics, and other fields to colleagues at the 2007 meeting of the Scientific Discovery through Advanced Computing (SciDAC) program held June 24–28 in Boston.

As many as 300 researchers attended SciDAC 2007, which gave NCCS users and other computational scientists an opportunity to share their results with their peers. Fusion scientists John Candy of General Atomics and Paul Bonoli of the Massachusetts Institute of Technology gave invited talks, as did astrophysicists Stan Woosley of the University of California–Santa Cruz and Joan Centrella of the National Aeronautics and Space Administration's Goddard Space Flight Center. All four are users of the NCCS's Cray XT4 Jaguar supercomputer. Candy discussed simulations of microturbulence in fusion plasmas, while Bonoli discussed heating of the plasmas with radio waves. Woosley reviewed simulations of Type Ia supernovas—which he describes as "nature's biggest bombs"—and Centrella talked about simulations of black holes in binary systems.

NCCS staff were also active at the meeting. Sean Ahern, head of the visualization task in the NCCS Scientific Computing Group (SCG) gave an invited talk on visual data analysis at the petascale and a tutorial on scientific data understanding using the VisIt and SCIRun visualization and analysis tools. SCG member Bronson Messer presented a poster on supernova simulation at the petascale using the CHIMERA application. Richard Mills, also of the SCG, presented a poster on the simulation of groundwater and other subsurface flow using the PFLOTRAN application.

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. "What we don't want them to worry about is how to move data from here to there."

The system will arrive in stages, with the first already in place for limited testing. That first stage, which should be complete by the end of the month, will offer about 80 terabytes of disk space and a bandwidth of about 10 gigabytes a second. A second phase, planned for late 2007 or early 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.

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.