Good prospects

by J. William Bell

Seismic modeling and reservoir simulations come to the TeraGrid, improving two workhorses of the oil industry.

Oil prospecting used to rely on hunches as much as anything else. Where to explore? Luck might land you a spot where oil was seeping from the ground. Where to set up the wells, pumps, and other equipment? There were some rules of thumb to fall back on. How to manage production? Perhaps you'd check the site's current performance, compare it to past wells' behavior, and then go by instinct.

Nowadays, however, a golden gut isn't worth what it used to be.

"Twenty or 30 years ago, people started to wonder how to get more from their wells," says Wolfgang Bangerth, a postdoc at the University of Texas at Austin's Center for Subsurface Modeling. "They began to wonder, 'What are the clever ways to do this?'"

Some of these more clever ways are newer techniques in drilling and production. Companies can now drill horizontally and sink wells that branch in multiple directions. They can shut down sections of their wells. Pumps, which force oil toward wells, can be run at different rates and on varying schedules.

Companies also need to choose ideal places for this equipment and to surmise the geological features of the ground beneath it. That's where people like Bangerth and his collaborators come in. Using TeraGrid resources, a multidisciplinary team from UT, The Ohio State University, and Rutgers University is at work on software tools that improve companies' oil reservoir management. The tools allow them to better exploit existing reservoirs, find new reservoirs, and minimize drilling's adverse environmental impact.

"Reservoir models, which help guide production strategies, and seismic imaging, used to determine subsurface properties, are two workhorses of the oil industry," Bangerth says.

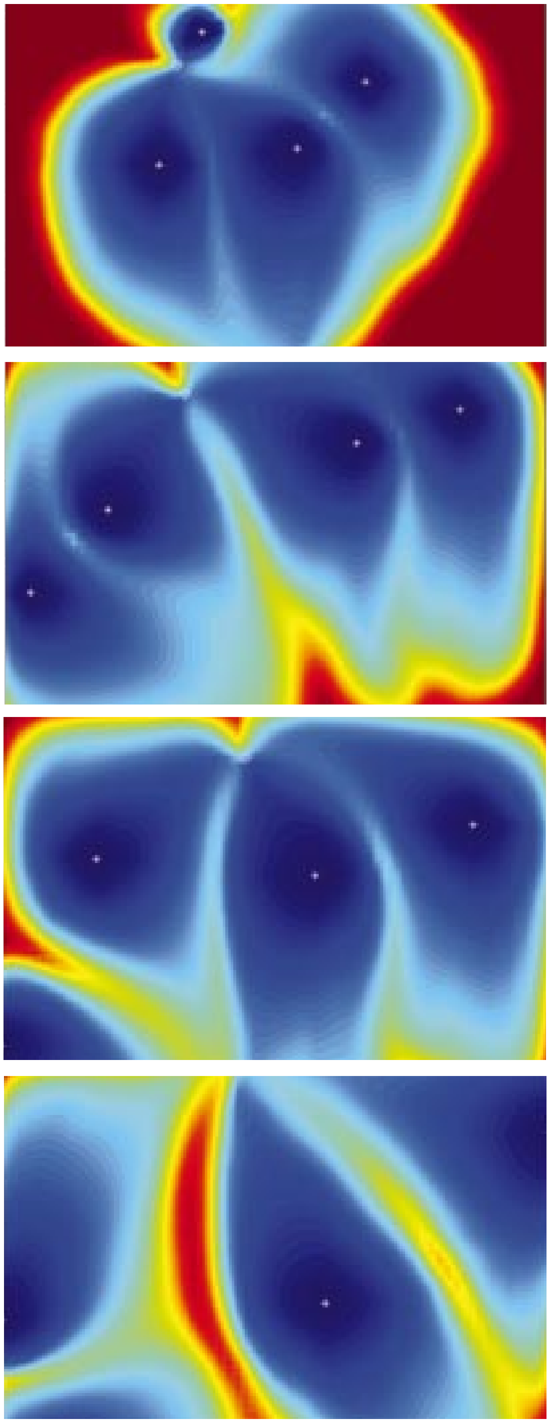
A reservoir of knowledge

Reservoir models are quite complicated. They account for different fluids like water, oil, and gas, different rock properties, the features of pumps and production wells, and sometimes complicated chemical reactions. To make modeling these complex systems computationally tractable, mathematicians subdivide the reservoir into a mesh of blocks. They then associate wells, pumps, and other equipment with individual blocks and solve an approximate model of the systems' fluid dynamics. The team uses IPARS, a multi-model, multi-phase reservoir simulator developed at the Center for Subsurface Modeling under the direction of Mary Wheeler. The output of IPARS is translated into production rates and ever-important revenue levels. Equipment is then moved around the mesh in order to compare different configurations and to find the best one.

The possibilities aren't endless, but their sheer number sometimes makes researchers pine for the days of hunch-based prospecting.

"You can have billions of possible configurations that need to be examined, so you can't just do an exhaustive search of the parameter space [the collection of all possible configurations within a given grid]," according to Tahsin Kurc, an assistant professor at Ohio State and part of the Multiscale Computing Lab that is led by Joel Saltz. A single IPARS run usually takes hours. If it's really difficult, hours can bleed into days.

"Complexity usually translates into precision," Kurc says. "We want to move intelligently through the search space."



Visualizations of oil reservoir simulations at various stages in an optimization. Pumps, represented by asterisks, push oil toward wells, represented by small white circles, that draw oil from the ground. Blue areas indicate areas of high water concentration. Brown areas indicate areas of high oil concentration. By the end of the simulations, there is not much oil left in the reservoir, and what is left has been swept toward the wells.

Intelligent movement relies on a dynamic, data-driven optimization system. Large volumes of data obtained from earlier simulations and dynamically updated by new simulations or experimental measurements are stored, queried, and analyzed to find promising initial configurations. These configurations are then refined with on-the-fly monitoring and steering of the simulation and optimization processes.

A set of simulations provides a rough sampling of the search space. Middleware tools from Saltz's team, called STORM and DataCutter, manage the very large amounts of data produced by these simulations. These tools are also used to identify good starting points for more comprehensive searches. Dynamic steering and collaboration tools—AutoMate and DISCOVER from Associate Professor Manish Parashar's lab at Rutgers—allow on-the-fly searches within these subsections. Sophisticated optimization algorithms coupled with IPARS models guide these searches by comparing configurations in the subsections.

'Do the math on that'

High oil output isn't always the objective. Sometimes companies want to modulate output over time to allow for changes in markets and prices. Sometimes they want to find a design that comes with the lowest risk of failures or mishaps. Regardless of what result they are looking for, they want to identify that result in the shortest possible amount of time. Distributed computing is key.

"We're supporting this now in a grid environment, modeling multiple configurations and multiple points concurrently," Kurc says.

The team recently completed a set of about 25,000 runs—each taking about two hours on a single processor—in less than a week. "Do the math on that," Bangerth says. "You're talking 200 to 400 runs going at any one time...That's not something we're used to having."

These calculations were completed using the TeraGrid cluster at NCSA. More runs are ongoing at NCSA and other TeraGrid sites across the country, including the San Diego Supercomputer Center and the Texas Advanced Computing Center. Machines at UT's Institute for Computational Engineering and Sciences, host institution for the Center for Subsurface Modeling, are also in use. The TeraGrid is the world's largest, most comprehensive infrastructure for open scientific research. It includes 20 teraflops of computing power distributed at nine sites, facilities capable of managing and storing nearly one petabyte of data, high-resolution visualization environments, and toolkits for grid computing.

NCSA's Bruce Loftis and Byoung-Do Kim helped port the reservoir simulation and optimization codes to the TeraGrid machines and built a toolkit that simplifies execution across multiple systems. The toolkit makes it easier for researchers to "babysit," as Kurc calls it, these large, distributed runs. It shows where calculations are taking place, which are complete, which have failed, and which are ongoing. That's no small feat when orchestrating work across hundreds of processors.

"The TeraGrid is well suited to this project—massive numbers of independent jobs," Loftis explains. "There are lots of these sorts of problems out there."

At the refinery

Oil reservoirs are generally inaccessible, being thousands of feet under the ground or the ocean. Unsurprisingly, little is typically known about their exact geological features. Scientists cope with this problem in two ways. Either they solve their equations on hundreds of geological models that are equally compatible with the existing knowledge of a site, or they try to come up with additional information.

Some information on geological conditions can come from real-world tests. Sound waves are blasted through the earth and bounce back to receivers on the ground or the ocean's surface. This echo can be translated into information about things like the type, density, and permeability of the rock and the amount of oil contained within. These soundings are taken over and over from different positions for a single area. They're expensive propositions, especially if they are to capture the level of precision researchers really want for their optimization.

The team's seismic models are used to develop likely geological conditions, based on simulated soundings. These conditions fine tune the reservoir models, making them as realistic as possible from the get-go.

The results of a single sounding passing into the ground and bouncing back can consume 20 gigabytes of space. Currently, more than eight terabytes of seismic simulation data sit on clusters at NCSA, ready to be integrated into the reservoir models. With the distributed storage and computing power of the TeraGrid and the middleware tools STORM and DataCutter, the team is looking to create more than 10 times that amount in the short term.

"And a good, big [seismic survey of an area] would be into the petabytes," Kurc says.

The team hopes to end up with a system that allows those prospecting for oil to build a database of possible conditions that have already had reservoir optimizations run. Companies will assess the geological features of the site they are interested in, query the database for the description that most closely resembles the site, and receive an already-completed optimization in return.

It might take some of the romance out of the process when compared to the make-or-break old days. But what's a little romance in the face of fewer failures, fewer environmental problems, and more dollars?

This research is supported by the National Science Foundation's Information Technology Research program, the Department of Energy, and the Department of Defense.

Access Online: http://access.ncsa.uiuc.edu/CoverStories/oil/

For further information: http://www.ices.utexas.edu/csm/

http://multiscalecomputing.org/

http://automate.rutgers.edu/

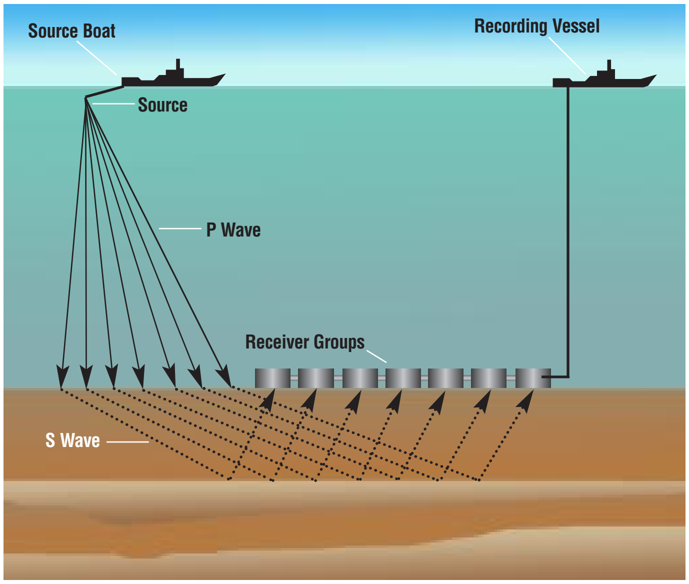


Illustration of soundings being taken on the ocean. Sound waves are projected from a ship, pass through the water, and are altered by the rock in the ocean floor. The echo from this process is picked up by receivers on the ocean floor or seismic monitoring platforms. It can be translated into information about things like the type, density, and permeability of the rock and the amount of oil contained within. These soundings are taken over and over from different positions for a single area.

Team members

Wolfgang Bangerth
Viraj Bhat
Umit Catalyurek
Shannon Hastings
Hector Klie
Vijay S. Kumar
Tahsin Kurc
Steve Langella
Vincent Matossian
Scott Oster
Manish Parashar
Benjamin Rutt
Joel Saltz
Roustam Seifoullaev
Mrinal Sen
Krishnan Sivaramakrishnan
Paul Stoffa
Mary Wheeler
Michael Zhang