

Anticipating “The Big One”

TACC’s Ranger supercomputer helps the Southern California Earthquake Center simulate most realistic earthquake to date

At 10 a.m. on Nov. 13, 5.3 million people in Southern California participated in the largest earthquake preparedness activity in U.S. history.

The Great Southern California ShakeOut — a collaboration between the United States Geological Society, the National Science Foundation, and the Southern California Earthquake Center (SCEC), among others — provided detailed information to the public about what would happen in the event of a magnitude 7.8 earthquake, (approximately 5,000 times larger than the magnitude 5.4 earthquake that shook southern California on July 29), and asked Californians to “drop, cover, and hold on,” practicing the behavior that could save their lives during an actual event.

Because, according to Thomas Jordan, director of SCEC, “the big one” is coming.

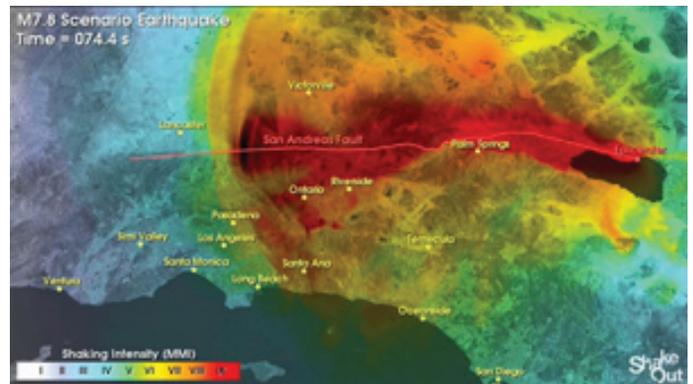
“We know that strain is accumulating on the San Andreas Fault, the stresses are increasing, and we expect that sucker to go at any time,” Jordan said.

If history is any indication, in the next thirty years there will be an earthquake in the region on a scale that hasn’t been experienced in 150 years — one that is expected to cause as many as 2,000 deaths, 50,000 injuries, \$200 billion in damage, and long-lasting disruption. The ShakeOut, both the precise predictions and the readiness exercise, was designed to prepare Southern California for this eventuality.

From an emergency preparation perspective, the activity was an extraordinary success. But perhaps equally impressive was the work that went into creating the most detailed simulation of an earthquake ever accomplished, in order to predict how a massive

upheaval would impact the specific topography of the 500-square-mile Los Angeles basin.

The effort required hundreds of the nation’s top seismologists, thousands of years of collective research, and the combined computational capability of some of the world’s most powerful supercomputers, including Ranger at the Texas Advanced Computing Center (TACC), Kraken at the National Institute of Computational Sciences (NICS) at University of Tennessee, and Datastar and Intimidata at San Diego Supercomputer Center at University of California, San Diego. All three supercomputer sites are part of TeraGrid, the nation’s largest open scientific discovery infrastructure.



This movie shows a view of southern California with the seismic waves radiating outward from the fault as the rupture propagates towards the northwest along the San Andreas fault. Simulations developed by the Southern California Earthquake Center ShakeOut Simulation workgroup. Animations courtesy of the U.S Geological Survey and the Southern California Earthquake Center.

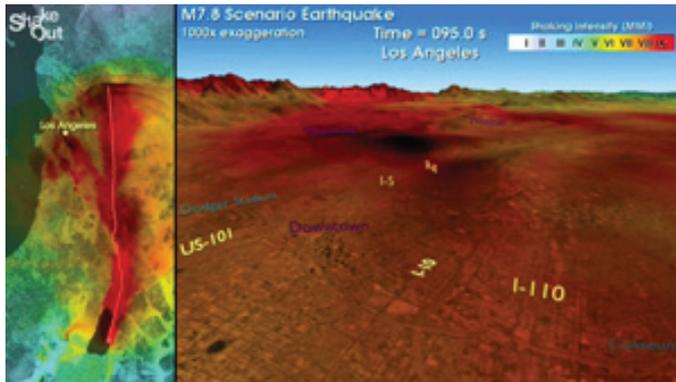
In computing terms, the simulation was unprecedented. But according to Jordan, it was not nearly comprehensive enough, making the exercise a dry run for even larger predictions in the future, with the abil-

Anticipating “The Big One”

ity to determine exactly how each meter of earth will shake when the San Andreas Fault gives.

When the Earth Moves

Traditionally, earthquake predictions are based on empirical data from past tremors, which scientists analyze to determine where the epicenter was and how different areas, and different topographies, reacted to the quake.



The detailed, perspective views show the ground shaking from a viewpoint two miles (three kilometers) above the earth looking towards each location. The left panel shows a map view of the area with the fault rupture highlighted in red, the epicenter (location where the rupture starts) identified by the red ball, and the location shown in the right panel labeled in yellow. In the right panel the deformation of the ground associated with the propagation of the seismic waves is exaggerated by a factor of 100. Animations courtesy of the U.S Geological Survey and the Southern California Earthquake Center.

However, over the last decade, SCEC has begun to transform seismic hazard analysis — the art of predicting and mapping the effect of an earthquake — from an empirical backward-looking discipline to a predictive science where researchers actually simulate the earthquake process in advance.

“We model how waves are generated by the fault, and how they travel outward from the fault and interact with complex three-dimensional geological structures, such as the sedimentary basins,” Jordan said.

“It’s those simulations that require the use of super-computing.”

Two factors determine the computational workload of these problems: 1) the complexity of the physics (including both the earthquake itself and the correlated factors that determine how waves travel through the ground), and 2) the resolution of the simulations. To create accurate predictions, scientists need both the most inclusive model of earthquake physics and the highest-resolution simulation possible.

The ShakeOut project accomplished both objectives. Run several times on several different high-performance computing systems, each ShakeOut simulation used increasing physical reality and resolution. The most advanced of these simulations, “ShakeOut D,” performed on the Ranger supercomputer at TACC, simulated the rupture and its impact with a 100-meter resolution over a three-dimensional area approximately 600 kilometers long by 300 kilometers wide by 80 kilometers deep. It stands as a milestone for seismology and for computational science generally.

Yifeng Cui, senior computational scientist at the San Diego Supercomputer Center, was responsible for the computational aspects of these runs. “This simulation is very data-intensive, one of the largest scale ever, not only for earthquake simulations but for all science,” he said.

By means of comparison: the previous iteration of this analysis, performed three years ago, used 1.8 billion grid points and took almost five days to run. ShakeOut D simulated 14.4 billion grid points, utilized between 16,000 and 60,000 cores simultaneously, and was accomplished in well under 12 hours — all of which indicate great leaps in computational science.

Even more important than the scale of the ShakeOut simulation, from a scientific perspective, was the fact that, for the first time, the seismic analysis used a dy-



Anticipating “The Big One”

dynamic, as opposed to a kinematic rupture, which has far greater fidelity and leads to dramatically different predictions.

In a kinematic rupture, the break is pre-determined and only the rippling impact is simulated on a supercomputer. In a dynamic rupture, by contrast, every aspect is driven by first-principle physics. Consequently, “the dynamic rupture simulation is a highly non-linear problem and therefore is a harder problem to solve on a computer,” Jordan said. “And the simulation that was done on Ranger was the largest dynamic rupture simulation we’ve ever done.”

The ShakeOut D scenario showed that an earthquake on the southern end of the San Andreas Fault shoots energy into the Los Angeles basin and then funnels energy into the Los Angeles region, which is why the city is so greatly impacted by this type of earthquake.

“These are exactly the type of effects that are not well represented in the empirical seismic hazard analysis, which shows far less shaking than we see from our simulations,” Jordan affirmed.

The ability to produce simulations with this resolution and complexity was no simple matter, requiring more than just a powerful supercomputer. “When you port the code to run on a big machine there are always new technical challenges,” Cui said. “As we approach the petascale level, a lot of changes have to be made. One size doesn’t fit all, so for example, the algorithms have to be adapted to the architecture for extreme levels of scalability.”

The simulation was not only a milestone in terms of processing power, said Yifeng Cui. It was also a significant advance in terms of memory usage, input and output, data transfer and storage, each of which required unique solutions.

The Real Aim

Jordan and his SCEC colleagues are not ones to rest on their laurels, however.

“The ShakeOut has been a very successful exercise because it was used by five million people and helped prepare L.A. better. But what I want to paint for you is the story of where this is going, because this is small beer compared to what we’re hoping to do,” he said.

For his next-generation seismic hazard mapping, Jordan has a much more ambitious goal. “For every site within a region, we want to generate on the order of half a million rupture models, which will require the same work we’re doing on Ranger, but doing it 500,000 times per site,” Jordan explained. “That gets us to a physics-based, probabilistic seismic hazard analysis that will transform the way we build buildings, the way we actually use seismic hazard analysis, and provide us with much better information.” The SCEC collaboration has already completed probabilistic assessments for 25 sites, however accomplishing his goals before the big one comes will require the compute power of the emerging petascale systems just now coming online, and better methods for using them.

“What is a challenging calculation today — in fact one we can’t yet do — will become commonplace in the future,” Jordan said.

Despite the computational cost of performing such simulations, the social impact of SCEC’s research endeavor outweighs any difficulties. “We’re trying to change the way people think about and live with earthquakes. Instead of something coming out of the blue, it’s something that’s part of your environment,” Jordan said. “In the future, the ShakeOut project is going to save lots of lives and dollars.”



Anticipating “The Big One”

And at the center of the seismic hazard analysis effort are Ranger and the other massive NSF-supported HPC systems that contributed to the simulation and analysis. “Supercomputers were the starting point for the whole thing,” Jordan said. “All of it depends on the continued improvement and creation of massive supercomputers like Ranger.”

###

The Ranger supercomputer is funded through the National Science Foundation (NSF) Office of Cyberinfrastructure “Path to Petascale” program. The system is a collaboration among the Texas Advanced Computing Center, The University of Texas at Austin’s Institute for Computational Engineering and Science, Sun Microsystems, Advanced Micro Devices, Arizona State University, and Cornell University.

Ranger is a key resource of the NSF TeraGrid (www.teragrid.org), a nationwide network of people, resources and services, also sponsored by the NSF Office of Cyberinfrastructure, which enables discovery in U.S. science and engineering. The TeraGrid provides scientists and researchers expertise in and access to large-scale computing power, networking, data-analysis, and visualization systems.

Aaron Dubrow
Texas Advanced Computing Center
Science and Technology Writer
January 21, 2009