Supercomputing On Demand: SDSC Supports Event-Driven Science

Pioneering System Offers New Resources for Rapid Response

July 10, 2007

Paul Tooby

Somewhere in Southern California a large earthquake strikes without warning, and the news media and the public clamor for information about the temblor -- Where was the epicenter? How large was the quake? What areas did it impact?

Frame from a movie of a â€œvirtual earthquakeâ€# simulation of the type that will be run on SDSCâ€™s new OnDemand system to support event-driven science. The movie shows the up-and-down velocity of the Earthâ€™s surface as waves radiate out from a magnitude 4.3 earthquake centered near Beverly Hills, California. Strong blue waves indicate the surface is moving rapidly downward, while red/orange waves indicate rapid upward motion. Courtesy of Joroen Tromp, ShakeMovie, Caltech.

A picture is worth a thousand words - or numbers - and the San Diego Supercomputer Center (SDSC) at UC San Diego is helping to provide the answers. Caltech computational seismologist Jeroen Tromp can now give the public movies that tell the story in a language that's easy to understand, revealing waves of ground motion spreading out from the earthquake -- and he can deliver these movies in just 30 minutes with the help of a supercomputer at SDSC. But he can't do it by submitting a job to a traditional computing batch queue and waiting hours or days for the results.

Tromp is an example of the new users in today's uncertain world who require immediate access to supercomputing resources. To meet this need, SDSC has introduced OnDemand, a new supercomputing resource that will support event-driven science.

"This is the first time that an allocated National Science Foundation (NSF) TeraGrid supercomputing resource will support on-demand users for urgent science applications," said Anke Kamrath, director of User Services at SDSC. "In opening this new computing paradigm we've had to develop novel ways of handling this type of allocation as well as scheduling and job handling procedures."

The system is already in operation and formal allocations of time for the OnDemandsystem will begin in October, with proposals due July 13. In addition to supporting important research now, this system will serve as a model to develop on-demand capabilities on additional TeraGrid systems in the future. TeraGrid is an NSF-funded computing grid linking some of the nation's largest supercomputer centers including SDSC.

Urgent applications that will make use of OnDemand range from making movies of Southern California earthquakes to systems that will help give near real-time warnings based on predictions of the path of tornados or a hurricane, or foretell the most likely direction of a toxic plume released by an industrial accident or terrorist incident.

When an earthquake greater than magnitude 3.5 strikes Southern California, typically once or twice a month, Tromp expects that his simulation code will need to use 144 processors of the OnDemand system for about 28

minutes. Shortly after the earthquake strikes a job will automatically be submitted and immediately allowed to run. The code will launch and any "normal" jobs running at the time will be interrupted to make way for the on-demand job.

"SDSC's new OnDemand system is an important step forward for our event-driven earthquake science," said Tromp. "We're getting good performance that will let us cut the time to deliver earthquake movies from about 45 to 30 minutes or less, and every minute is important."

The movies that result from the computations are made available as part of the ShakeMovie project in Caltech's Near Real-Time Simulation of Southern California Seismic Events Portal. But behind the scenes of these dramatic earthquake movies, a great deal of coordinated activity is rapidly taking place in a complex, automated workflow.

The system springs to life every time an earthquake occurs in Southern California. When an event takes place, thousands of seismograms, or ground motion measurements, are recorded at hundreds of stations across the region, and the earthquake's epicenter, or location, as well as its depth and intensity are determined.

The waiting ShakeMovie system at Caltech collects these seismic recordings automatically over the Internet. Then, for events greater than magnitude 3.5, to fill in the gaps between the actual ground motion recorded at specific locations in the region, the scientists use the recorded data to guide a computer model that creates a "virtual earthquake," giving an overall view of the ground motion throughout the region.

The animations rely on the SPECFEM3D_BASIN software, which simulates seismic wave propagation in sedimentary basins. The software computes the motion of the earth in 3-D based on the actual earthquake recordings and what is known about the subsurface structure of the region, which greatly affects the wave motion -- bending, speeding or slowing, and reflecting energy in complex ways.

After the full 3-D wave simulation is run on the OnDemand system at SDSC and a system at Caltech for redundancy, data that captures the surface motion (displacement, velocity, and acceleration) are collected and mapped onto the topography of Southern California, and rendered into movies. The movies are then automatically published via the portal, and an email is sent to subscribers, including the news media and the public.

In between the urgent jobs that use SDSC's OnDemand resource, other users will run on the system in a normal way. The system has StarP installed, a parallel version that provides a high-performance backend to packages such as Matlab.

Media Contact: Paul Tooby, (858) 822-3654 Comments: Anke Kamrath, (858) 534-5140