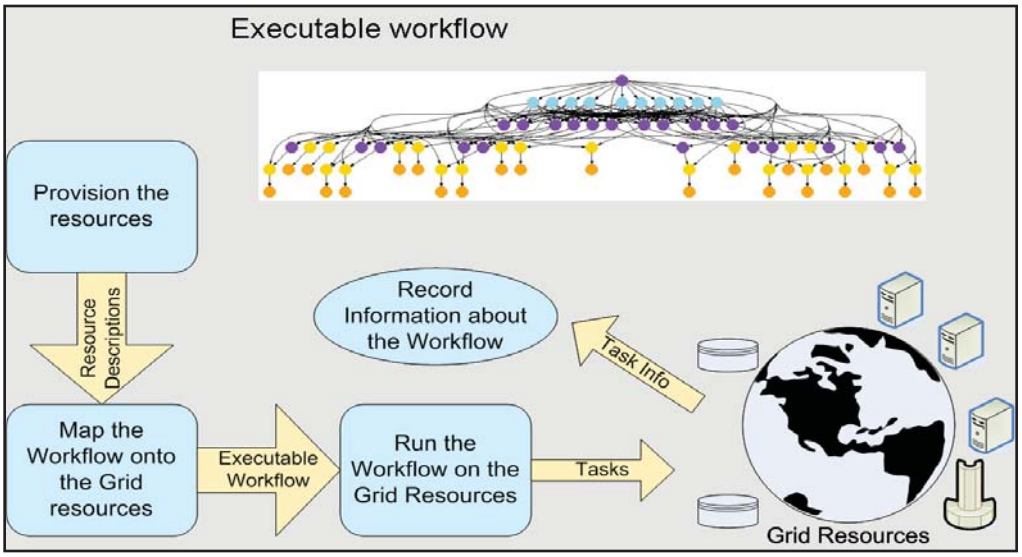
Advancing SCEC Earthquake System Science Research with Grid-based High Performance Computing and Scientific Workflows

The Southern California Earthquake Center (SCEC) is a multi-institution, multi-disciplinary research center bringing together geo-scientists and computer scientists from over 54 institutions to develop a comprehensive and predictive understanding of earthquakes and to disseminate this information in ways that increase awareness, reduce economic losses and save lives. An important component of the SCEC research program is a distributed, grid-based, high performance computing environment - the SCEC grid. The SCEC grid allows scientists to configure and execute large-scale scientific workflows across a heterogenous computing environment.



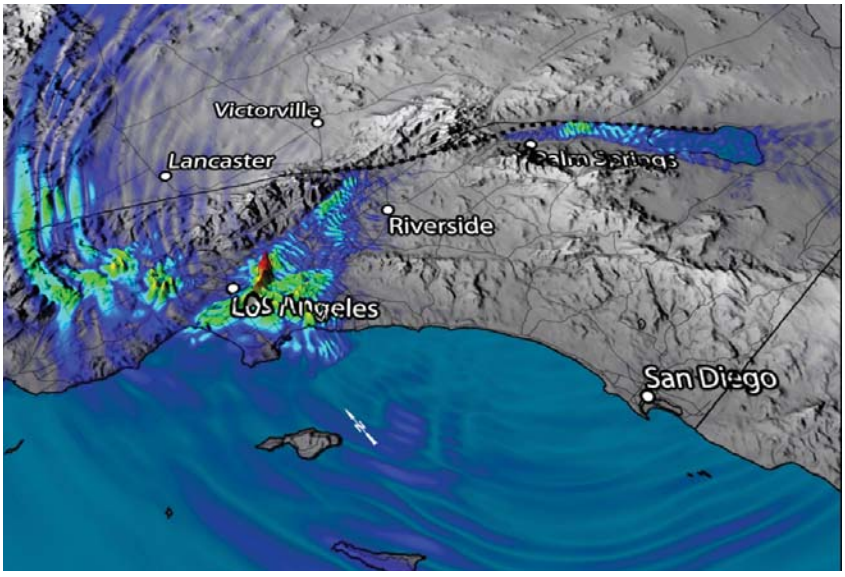
The SCECGrid supports the configuration and executon of distributed, large-scale scientific workflows through the use of the Virtual Data System (VDS) that includes Globus, Condor, Chimera, RLS, MCS, and Pegasus.

The Southern California Earthquake Center (SCEC) is engaged in an active earthquake system science research program that seeks to develop predictive models of earthquake processes. Essential to this research is the development and application of appropriate cyberinfrastructure in order to perform large-scale, physically realistic, numerical simulations. The SCEC collaboration has developed a series of computational platforms (vertically integrated collections of hardware, software, and people) that support large-scale earthquake simulations on high performance computers. The high performance simulation codes, within the SCEC computational platforms, utilize the multi-teraflop computing facilities of the NSF TeraGrid cyberinfrastructure and the computing resources of the USC High Performance Computing and Communications (HPCC) system.

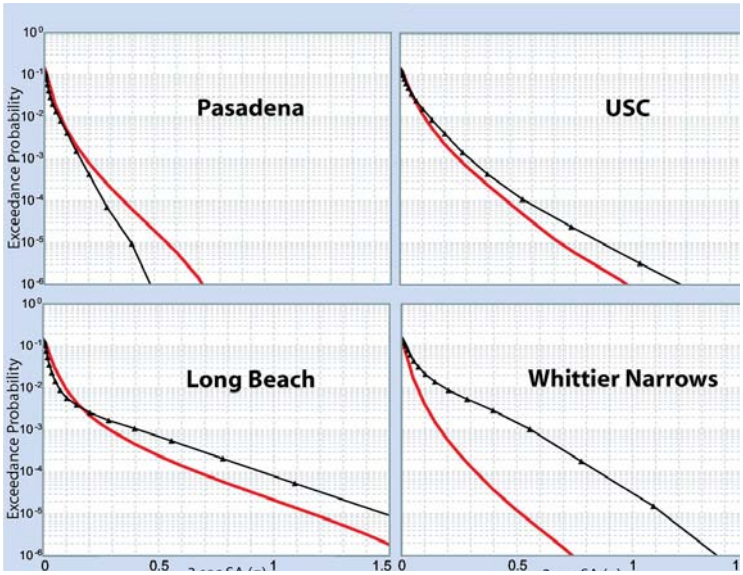
There are currently four SCEC computational platforms and each platform provides a widely useful, reusable simulation capability that supports SCEC science goals The SCEC computational platforms include (1) dynamic rupture and wave propagations capabilities (TeraShake Platform), (2) waveform-based probabilistic seismic hazard curve calculation capabilities (CyberShake Platform), (3) verification and validation capabilities (SCEC Earthworks Platform), and (4) delivery of probabilistic seismic hazard analysis (PSHA) results to end users (OpenSHA Platform).

The SCEC computational platforms have been used to produce important SCEC research results including a Puente Hills Earthquake Loss Estimate study (Field et al (2005), Earthquake Spectra) and the TeraShake simulations of large earthquakes on the Southern San Andreas fault (Olsen et al (2006), Geophysical Research Letters).

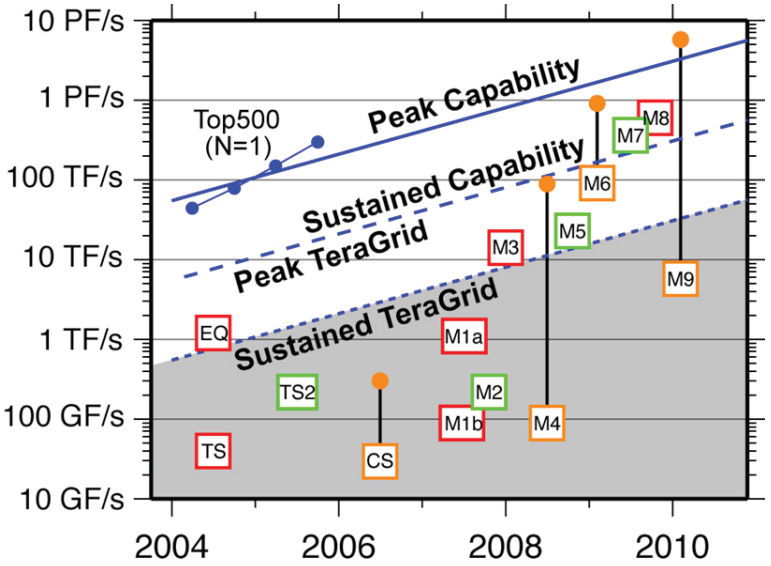
All of the SCEC computational platforms require high performance computing and large scale data management tools. The SCEC computational platforms share common cyberinfrastructure that includes grid-based scientific workflow capabilities based on the Virtual Data System (VDS)



This velocity magnitude map of TeraShake 1.3 data shows how strong ground motions are channeled into the Los Angeles region for a large Southern San Andreas Eartquake. (Image: SDSC)



The output of a CyberShake calculation is a Probabilistic Seismic Hazard curve – the probability of exceeding an Intensity Measure Level (IML) at a site. This image compares several CyberShake hazard curves to attenuation relationship-based hazard curves, for spectral acceleration at a period of 3.0 seconds.



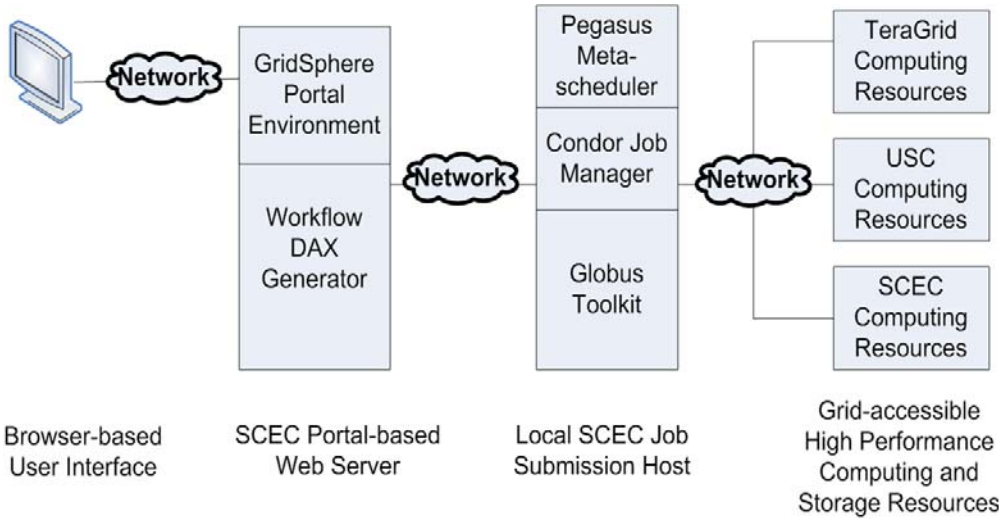
SCEC has outlined a series of simulations that will advance the SCEC earthquake system science research goals. This graph shows the estimated computing performance required to complete each of these simulations within 100 hours. Vertical bars indicate simulations that can use capacity computing.

and the Storage Resource Broker (SRB) data management software.

Complex, multistep computations performed by the SCEC platforms are modeled as scientific workflows and the SCEC grid-based workflow tools securely distribute the computations across a heterogeneous collection of computing resources. This system allows SCEC to share computing and storage resources across the multi-institution SCEC collaboration.

The SCEC earthquake system science program requires a wide variety of high performance computing capabilities including capability computing, capacity computing, and data intensive computing as well as scientific workflow capabilities. The SCEC cyberinfrastructure must also support the highly diverse SCEC research community in which researchers have a wide range of high performance computing expertise.

The SCEC Earthworks science gateway was developed in order to provide non-traditional supercomputer users with access to the high performance computing and scientific workflow capabilities of the SCEC grid. The SCEC Earthworks system allows users to configure, execute, and monitor earthquake simulations through a browser-based interface. Multi-step earthquake simulations are converted into workflows by the Earthworks system. The SCEC Earthworks system provides users with access to the high performance computing, scientific workflow, and data management capabilities of the SCEC grid including access to computing resources at SCEC, USC/HPCC, and the TeraGrid.



The SCEC Earthworks Science Gateway allows SCEC researchers to configure and run an earthquake simulation using grid-based scientific workflow technologies on the TeraGrid national supercomputer facilities through a web-based portal interface.

Southern California Earthquake Center
(213) 740-5843 www.scec.org

Core Organizations
University of Southern California
California Institute of Technology
Columbia University
San Diego State University
University of California, Los Angeles
University of California, Riverside
University of California, San Diego
University of California, Santa Barbara
University of California, Santa Cruz
University of Nevada, Reno
Harvard Univ.
Massachusetts Institute of Technology
Stanford University
U.S. Geological Survey, Golden
U.S. Geological Survey, Menlo Park
U.S. Geological Survey, Pasadena

Participating Organizations
Academia Sinica (Taiwan)
Arizona State University
Boston University
Brown University
Cal-State, Fullerton
Cal-State, Northridge
Cal-State, San Bernardino
California Geological Survey
Carnegie Mellon University
Case Western Reserve University
Central Washington University
CICESE (Mexico)
ETH Zurich (Switzerland)
Inst. of Geol. and Nuclear Sciences (New Zealand)
Jet Propulsion Laboratory
Lawrence Livermore National Laboratory
National Central University (Taiwan)
National Chung Cheng University (Taiwan)
National Taiwan University
Oregon State University
Pennsylvania State University
Rensselaer Polytechnic Institute
Rice University
SUNY Stony Brook
Texas A&M University
University of California, Berkeley
University of California, Davis
University of California, Irvine
University of Colorado
University of Kentucky
University of Massachusetts
University of New Mexico
University of Oregon
University of Western Ontario
URS Corporation
Utah State University
Whittier College
Woods Hole Oceanographic Institution