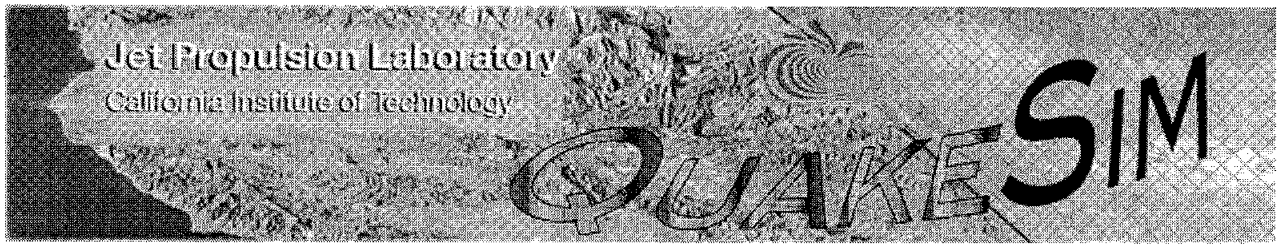
Jet Propulsion Laboratory
California Institute of Technology

+ View the NASA Portal

Search JPL

JPL HOME | EARTH | SOLAR SYSTEM | STARS & GALAXIES | TECHNOLOGY



QUAKESIM HOME | QUAKESIM PORTAL | QUAKETABLES | ISERVO @ JPL | NASA CT

Project information

Goals and Components

Milestones

System Documentation

Download Code

People

Related Links

FORECAST/Scorecard

Animations

Latest news

Read the **Second Annual Report** (posted 3/19/04)

QuakeSim Library is now available *(authorized users only)*

Welcome to the QuakeSim project website

Our objectives

- Develop a solid Earth science framework in order to better understand active tectonic and earthquake processes
- Construct a fully interoperable system of tools for studying these processes.

QuakeSim is sponsored by the NASA Earth Science Enterprise, in partnership with Goddard Space Flight Center, with the full participation of the related science and technology communities.

We are developing three major simulation tools (available fordownload), GeoFEST, PARK, and Virtual California:

- GeoFEST uses stress-displacement finite elements to model stress and flow in a realistic model of the Earth's crust and upper mantle in a complex region such as the Los Angeles Basin. The model includes stress and strain due to the elastic response to an earthquake event in the region of the slipping fault, the time-dependent viscoelastic relaxation, and the net effects from a series of earthquakes. The physical domain may be two or three dimensional and may contain heterogeneous materials and an arbitrary network of faults.

- PARK is a boundary element program that determines the stress on every element of the fault surface due to slip on every other element, using a Green's function approach. For example, it can be used to compute the history of slip, slip velocity, and stress on a vertical strike-slip fault that results from using state-of-the-art rate and state frictional constitutive laws on the fault for a specific geographic setting at Parkfield, California.

- Virtual California is a code that utilizes the Monte Carlo method in order to generate simulated, realistic earthquakes on an arbitrary fault surface mesh. It uses topologically realistic networks of independent fault segments that are mediated by elastic interactions. These segments can be designed to represent fault systems spanning the region of California.

See these links for more details about:

QuakeSim and the Solid Earth Research Virtual Observatory

Andrea Donnellan[(1)], John Rundle[(2)], Geoffrey Fox[(3)], Dennis McLeod[(4)], Lisa Grant[(5)], Terry Tullis[(6)], Marlon Pierce[(3)], Jay Parker[(1)], Greg Lyzenga[(1)]

(1) Science Division, Jet Propulsion Laboratory, California Institute of Technology, Pasadena, CA 91109, USA (e-mail: donnellan@jpl.nasa.gov; phone: +1 818-354-4737). (2) Center for Computational Science and Engineering, University of California, Davis, California, 95616, USA (e-mail: rundle@geology.ucdavis.edu; phone: +1 530-752-6416). (3) Community Grid Computing Laboratory, Indiana University, IN 47404, USA (e-mail: gcf@indiana.edu; +1 852-856-7977). (4) Computer Science Department, University of Southern California, Los Angeles, CA 90089, USA (e-mail: mcleod@usc.edu; 213-740-4504). (5) Environmental Analysis and Design, University of California, Irvine, CA 92697, USA (e-mail: lgrant@uci.edu; 949-824-5491). (6) Brown University, Providence, RI 02912, USA (e-mail: terry_tullis@brown.edu; 401-863-3829).

Abstract

We are developing simulation and analysis tools in order to develop a solid Earth science framework for understanding and studying active tectonic and earthquake processes. The goal of QuakeSim and its extension, the Solid Earth Research Virtual Observatory (SERVO), is to study the physics of earthquakes using state-of-the-art modeling, data manipulation, and pattern recognition technologies. We are developing clearly defined accessible data formats and code protocols as inputs to simulations, which are adapted to high-performance computers. The solid Earth system is extremely complex and nonlinear resulting in computationally intensive problems with millions of unknowns. With these tools it will be possible to construct the more complex models and simulations necessary to develop hazard assessment systems critical for reducing future losses from major earthquakes. We are using Web (Grid) service technology to demonstrate the assimilation of multiple distributed data sources (a typical data grid problem) into a major parallel high-performance computing earthquake forecasting code. Such a linkage of Geoinformatics with Geocomplexity demonstrates the value of the Solid Earth Research Virtual Observatory (SERVO) Grid concept, and advances Grid technology by building the first real-time large-scale data assimilation grid.

Introduction

QuakeSim is a new Problem Solving Environment for the seismological, crustal deformation, and tectonics communities for developing an understanding of active tectonic and earthquake processes. One of the most critical aspects of our system is supporting interoperability given the heterogeneous nature of data sources as well as the variety of application programs, tools, and simulation packages that must operate with data from our system. Interoperability is being implemented by using distributed object technology combined with development of object API's that conform to emerging standards. The full objective is to produce a system to fully model earthquake-related data. Components of this system include:

- A database system for handling both real and simulated data
- Fully three-dimensional finite element code (FEM) with an adaptive mesh generator capable of running on workstations and supercomputers for carrying out earthquake simulations
- Inversion algorithms and assimilation codes for constraining the models and simulations with data
- A collaborative portal (object broker) for allowing seamless communication between codes, reference models, and data
- Visualization codes for interpretation of data and models
- Pattern recognizers capable of running on workstations and supercomputers for analyzing data and simulations

Project details and documentation are available at the QuakeSim main web page at http://quakesim.jpl.nasa.gov.

Integrating Data and Models

The last five years have shown unprecedented growth in the amount and quality of space geodetic data collected to characterize geodynamical crustal deformation in earthquake prone areas such as California and Japan. The Southern California Integrated Geodetic Network (SCIGN), the growing EarthScope Plate Boundary Observatory (PBO) network, and data from Interferometric Synthetic Aperature Radar (InSAR) satellites are examples. Hey and Trefethen (http://www.grid2002.org) [1] stressed the generality and importance of Grid applications exhibiting this "data deluge."

Multiscale integration for Earth science requires the linkage of data grids and high performance computing. Data grids must manage data sets that are either too large to be stored in a single location or else are geographically distributed by their nature (such as data generated by distributed sensors). The computational requirements of data grids are often loosely coupled and thus are embarrassingly parallel. Large-scale simulations require closely coupled systems. QuakeSim and SERVO support both styles of computing. The modeler is allowed to specify the linkage of descriptions across scales as well as the criterion to be used to decide at which level to represent the system. The goal is to support a multitude of distributed data source, ranging over federated database, sensor, satellite data and simulation data, all of which may be stored at various locations with various technologies in various formats. QuakeSim conforms to the emerging Open Grid Services Architecture.

Computational Architecture and Infrastructure

Our architecture is built on modern Grid and Web Service technology whose broad academic and commercial support should lead to sustainable solutions that can track the inevitable technology change. The architecture of QuakeSim and SERVO consists of distributed, federated data systems, data filtering and coarse graining applications, and high performance applications that require coupling. All pieces (the data, the computing resources, and so on) are specified with URIs and described by XML metadata.

Web Services

We use Web services to describe the interfaces and communication protocols needed to build our. Web services, generally defined, are the constituent parts of an XML-based distributed service system. Standard XML schemas are used to define implementation independent representations of the service's invocation interface (WSDL): messages (SOAP) exchange between two applications. Interfaces to services may be discovered through XML-based repositories. Numerous other services may supplement these basic capabilities, including message level security and dynamic invocation frameworks that simplify client deployment. Implementations of clients and services can in principle be implemented in any programming language (such as Java, C++, or Python), with interoperability obtained through XML's neutrality.

Our approach to Web services divides them into two major categories: core and application. Core services include general tasks such as file transfer and job submission. Application services consist of metadata and core services needed to create instances of scientific application codes. Application services may be bound to particular host computers and core services needed to accomplish a particular task.

XML-Based Metadata Services

In general, SERVO is a distributed object environment. All constituent parts (data, computing resources, services, applications, etc.) are named with universal resource identifiers (URIs) and described with XML metadata. The challenges faced in assembling such a system include a) resolution of URIs into real locations and service points; b) simple creation and posting of XML metadata nuggets in various schema formats; c) browsing and searching XML metadata units.

XML descriptions (schemas) can be developed to describe everything: computing service interfaces, sensor data, application input decks, user profiles, and so on. Because all metadata are described by some appropriate schema, which in turn derive from the XML schema specification, it is possible to build tools that dynamically create custom interfaces for creating and manipulating individual XML metadata pieces.

After metadata instances are created, they must be stored persistently in distributed, federated databases. On top of the federated storage and retrieval systems, we are building organizational systems for the data. This requires the development of URI systems for hierarchically organizing metadata pieces, together with software for resolving these URIs and creating internal representations of the retrieved data. It is also possible to define multiple URIs for a single resource, with URI links pointing to the "real" URI name. This allows metadata instance to be grouped into numerous hierarchical naming schemes.

Federated Database Systems and Associated Tools

Our goal is to provide interfaces through which users transparently access a heterogeneous collection of independently operated and geographically dispersed databases, as if they formed a large virtual database [2,3]. There are five main challenges associated with developing a meta-query facility for earthquake science databases: (1) Define a basic collection of concepts and inter-relationships to describe and classify information units exported by participating information providers (a *"geophysics meta-ontology"*), in order to provide for a linkage mechanism among the collection of databases. (2) Develop a "meta-query mediator" engine to allow users to formulate complex meta-queries. (3) Develop methods to translate meta-queries into simpler derived queries addressed to the component databases. (4) Develop methods to collect and integrate the results of derived queries, to present the user with a coherent reply that addresses the initial meta-query. (5) Develop generic software engineering methodologies to allow for easy and dynamic extension, modification, and enhancement of the system.

Interoperability Portal

QuakeSim demonstrates a web-services problem-solving environment that links together diverse earthquake science applications on distributed computers. For example, one can use QuakeSim to build a model with faults and layers from the fault database, automatically generate a finite element mesh, solve for crustal deformation and produce a full color animation of the result integrated with remote sensing data. This portal environment is rapidly expanding to include many more applications and tools.

Our approach is to build a three-tiered architecture system. The tiers are: 1) A portal user interface layer that manages client components; 2) A service tier that provides general services (job submission, file transfer, database access, etc.) that can be deployed to multiple host computers; and 3) Backend resource, including databases and earthquake modeling software.

The user interacts with the system through the Web Browser interface. The QuakeSim portal effort has been one of the pioneering efforts in building Computing Portals out of reusable portlet components.

Earthquake Fault Database

The "database system" for this project manages a variety of types of earthquake science data and information. There are pre-existing collections, with heterogeneous access interfaces; there are also some structured collections managed by general-purpose database management systems. The QuakeSim database is an objective database that includes primary geologic and paleoseismic fault parameters (fault location/geometry, slip rate at measured location, measurements of coseismic displacement, dates and locations of previous ruptures) as well as separate interpreted/subjective fault parameters such as characteristic segments, average recurrence interval, magnitude of characteristic ruptures, etc. The database is updated as more data are acquired and interpreted through research and the numerical simulations. Our fault database is searchable with annotated earthquake fault records from publications.

Application Programs

QuakeSim is developing three high-end computing simulation tools: GeoFEST, PARK and Virtual California. We have demonstrated parallel scaling and efficiency for the GeoFEST finite element code, the PARK boundary element code for unstable slip, and the Green's functions Virtual California. GeoFEST was enhanced by integration with the Pyramid adaptive meshing library, and PARK with a fast multipole library. Other codes carry out forward or inverse elastic dislocation calculations, pattern recognition, or visualization.

Data Assimilation and Mining Infrastructure

Solid earth science models must define an evolving, high-dimensional nonlinear dynamical system, and the problems are large enough that they must be executed on high performance computers. Data from multiple sources must be ingested into the models to provide constraints, and methods must be developed to increase throughput and efficiency in order for the constrained models to be run on high-end computers.

Data Assimilation

Data assimilation is the process by which observational data are incorporated into models to set these parameters, and to "steer" or "tune" them in real time as new data become available. The result of the data assimilation process is a model that is maximally consistent with the observed data and is useful in ensemble forecasting. Data assimilation methods must be used in conjunction with the dynamical models as a means of developing an ensemble forecast capability. Data to be assimilated into the models include earthquake occurrence time, location, and moment release, as well as surface deformation data obtained from InSAR, GPS, and other methods. The data will ultimately be assimilated into viscoelastic finite element and fault interaction models and into a Potts model in which data must be assimilated to determine the universal parameters that defined the model. Once these parameters are determined, the Potts model represents the general form of a predictive equation of system evolution.

Datamining: Pattern Analysis as a General Course Graining

In many solid earth systems, one begins by partitioning (tesselating) the region of interest with a regular lattice of boxes or tiles. Physical simulations demonstrate that the activity in the various boxes is *correlated* over a length scale ξ, which represents the appropriate length scale for the fluctuations in activity. The *analysis* (as opposed to the simple *identification*) of space-time patterns of earthquakes and other earth science phenomena is a relatively new field. Several techniques of pattern analysis have been recently developed and are being applied: *Eigenpattern Method ("Karhunen-Loeve" or "Principal Components")* uses catalogs of seismic activity to define matrix operators on a coarse-grained lattice of sites within a spatial region. *Hidden Markov Models (HMM)* provide a framework in which given the observations and a few selected model parameters it is possible to objectively calculate a model for the system that generated the data, as well as to interpret the observed measurements in terms of that model. *Wavelet-Based Methods* emphasize the use of wavelets to 1) search for scale dependent patterns in coarse-grained space-time data, and 2) analyze the scale dependence of the governing dynamical equations, with a view towards developing a reduced set of equations that are more computationally tractable yet retain the basic important physics. One goal of QuakeSim is to automate *Ensemble Forecasting* using data assimilation and pattern recognition.

On Demand Science

The QuakeSim/SERVO portal is intended to enhance science activities and is available to the science community. The system allows for users to link applications, which allows for ingestion of data, more complete modeling, and visualization of data and results. It has been used for inverting GPS data to determine the rupture plane of the 2003 San Simeon Earthquake. In another application of the portal, GPS data show an anomalous band of compression in the northern part of the Los Angeles basin. Scientists at UC Davis used QuakeSim tools to investigate both elastic and viscoelastic models of the basin to test what set of model parameters best reproduced the observed GPS data. They used SIMPLEX for elastic models and GeoFEST for viscoelastic models. The models that yielded the best results featured a vertically strong crust, a single fault, and a low-rigidity, anelastically deforming basin abutting the rigid San Gabriel Mountains block (see Figure 4 below). These models produced a horizontal velocity gradient that matched the observed gradient across the LA basin both in magnitude and shape. While the match is not perfect, the models suggest that the sedimentary basin may be controlling the deformation in northern metropolitan Los Angeles, rather than faults south of the Sierra Madre fault as some have claimed.

Acknowledgments

We thank Peggy Li for integrating the visualization tools with QuakeSim and Servo, Charles Norton for adapting GeoFEST to the PYRAMID library, and our project review board led by Brad Hager from MIT. We thank the graduate students, Anne Chen, Sang-Soo Song, and Shan Gao from USC for their work on the database systems. Robert Shcherbakov from UC Davis assessed the data assimilation methods. Teresa Baker and Maggi Glasscoe have carried out science models using this system and developed documentation for various components. Michele Judd has proved invaluable in the management of this project. This research was carried out at the Jet Propulsion Laboratory, California Institute of Technology, and at the institutions of the co-authors under contract with NASA through their Computational Technologies and Advanced Information Systems Technologies programs.

References

[1]*Grid Computing: Making the Global Infrastructure a Reality* edited by Fran Berman, Geoffrey Fox and Tony Hey, John Wiley & Sons, Chichester, England, ISBN 0-470-85319-0, March 2003.

[2]*Federated Database-Systems For Managing Distributed, Heterogeneous, And Autonomous Databases,* Sheth AP, Larson JA, Computing Surveys, 22 (3): 183-236, Sep 1990.

[3]*Federating Neuroscience Databases*, Computing the Brain: A Guide to Neuroinformatics (editors Arbib, M., and Grethe, J.), Academic Press, 2001.