CENTER FOR COMPUTATION & TECHNOLOGY
Interdisciplinary | Innovative | Inventive

LOUISIANA STATE UNIVERSITY

Events

Current Events
Lectures ▼
Events Archive ▼

CCT Colloquium Series

Seismic Hazard Modeling using Heterogeneous Scientific Workflows

Philip Maechling, University of Southern California

Information Technology Architect, Southern California Earthquake Center

Johnston Hall 338
September 14, 2007 - 02:00 pm

Abstract:

As a part of the Southern California Earthquake Center (SCEC) program of seismic hazard research, we are using scientific workflow technologies to run large-scale high performance and high throughput scientific applications. As we construct scientific workflows to execute our seismic hazard computations, we face many types of heterogeneity that complicate the construction and execution of our research workflows. The seismic hazard applications programs that are run as a part of our workflows vary from large-scale MPI-codes, written in C, to small, shared-memory codes, written in FORTRAN 77. Run-times for different portions of the workflow may vary from multiple hours on a large cluster to less than a minute on a single processor. The scientific application codes may output a small number of very-large files, or many very-small files. In addition, the distributed computing environments in which these workflows run include a wide range of computers, operating systems, networks, and software environments and the users of workflow technology may have widely different backgrounds and levels of computing expertise. We view these issues of heterogeneity as essential and unavoidable challenges for the application and use of scientific workflow technology. In this talk, I will discuss the computational challenges of SCEC seismic hazard research and will describe our use of heterogeneous scientific workflows using a workflow lifecycle perspective. The scientific workflow lifecycle is initiated when a scientific question is posed that can be addressed through computational methods. The workflow lifecycle then proceeds to include the preparation of the scientific application codes for use in workflows, the construction of workflows, the instantiation of workflows, the execution of the workflows in a distributed high performance computing environment, the discovery of results, and the access to results by scientists.

Speaker's Bio:

Philip Maechling is the Information Technology Architect at the Southern California Earthquake Center (SCEC) and is currently the Project Manager on the NSF-funded project called "A Petascale Cyberfacility for Physics-based Seismic Hazard Analysis." At SCEC since 2002, Mr. Maechling has led the development of an integrated geophysical simulation modeling framework called the SCEC Community Modeling Environment (CME) that automates the process of selecting, configuring, and executing numerical models of earthquake processes. The SCEC CME system integrates high performance Geoscientific application programs into a distributed, grid-based, scientific workflow system that provides scientists with the ability to perform large-scale and highly complex research simulations and to organize and analyze the simulation results. Prior to his role at SCEC, Philip worked as a Member of the Professional Staff in the Caltech seismological laboratory where he coordinated the development of the next generation real-time earthquake monitoring system for southern California called the TriNet system. Philip has authored and co-authored several Geoscientific research publications on topics such as probabilistic seismic hazard analysis and large-scale earthquake wave propagation simulations as well as journal articles and book chapters on computer science research topics including distributed object technology, grid computing, large-scale data management, and scientific workflow technologies.