

September 26, 2003

SCEC/CME Pathway 1 Grid Computing Benchmark Exercise

Hongsuda Tangmunarunkit, Vipin Gupta, Nitin Gupta, Edward Field, Philip Maechling
hongsuda@isi.edu, field@usgs.gov, {niting,vgupta,maechlin}@usc.edu

Abstract:

The SCEC/CME Pathway 1 and Grid development working groups performed a series of Probabilistic Seismic Hazard Assessment (PSHA) Map calculations using the SCEC/CME computational testbed. PSHA Map calculations were performed using a standalone SCEC/CME server, and identical calculations were performed using the SCEC-USC Grid system. Each Hazard Map calculation consisted of 11760 individual hazard curve calculations. Using the standalone SCEC/CME server, the best case run-time was 430 wall-clock minutes (7.2 hours). Using the Condor Pool capabilities of the SCEC-USC Grid system, the best case run-time was 27 wall-clock minutes.

1. Pathway 1 Application Program:

The SCEC/CME Pathway 1 group has developed Java application programs that calculate the data sets required to plot PSHA Hazard Maps. These PSHA Hazard map data sets consist of a series of hazard curves, with one hazard curve data file for each site in a gridded region. For the benchmark exercise, we selected a southern California region of 4.5 degrees by 6.5 degrees with spacing of 0.05 degrees resulting in 11760 sites of interest. To generate the Hazard Map data set for this gridded region, we needed to perform 11760 individual hazard curve calculations.

2. Calculation Characteristics:

The PSHA hazard map calculations are typically run on shared memory machines as a series of hazard curve calculation. The same hazard curve calculation is performed for each site in the map. There is no time, or computational, dependency between the Hazard curve calculations. The hazard curves at each site differ because of the distance from the site being calculated and the predefined earthquakes used in the Hazard forecasts.

The inputs to the PSHA hazard curve calculations are a latitude and longitude. No programmatic I/O is done until the calculation is completed. Each hazard curve calculation outputs an ASCII File containing 20 data points.

These characteristics indicated that this application was a good candidate for the Condor Pool capabilities of the USC Grid system. Since each Hazard curve calculation was independent, the curves calculations could be run in parallel, limited only by the number of nodes available.

Since the PSHA application is written in Java, Java must be available on any Condor nodes used for scheduling. Output files should be returned to node from which the jobs are submitted.

3. Comparative Systems:

The Hazard Map calculation were run on the SCEC/CME server, a Dell Poweredge 2650 running Linux Redhat 8.0. This system is a development system for the SCEC/CME and it was reserved for benchmarking use during the time of the measurements.

The SCEC/CME server is a grid client to the USC Grid system with user accounts on the SCEC/CME server configured to submit jobs to the USC node Almaak.

4. Computational Work Units:

The time required to create a Hazard curve is the sum of the time required to instantiate a Java virtual machine plus the time required to perform the actual calculation. If the hazard curve calculations are short, the JVM instantiation time could require a large fraction of the total computation time.

All our benchmarks runs consisted of 11760 hazard curve calculations. We divided these calculations into groups of calculations that we called Work Units. A Work Unit, in this benchmark, is one, or more, hazard curve calculations. Because a Work Unit is a series of hazard curve calculations, Work Units vary in size. Benchmarks with a low number of Work Units will instantiate the JVM a low number of times. Benchmarks with a high number of Work Units will instantiate a high number of JVMs. A Work Unit must be performed on a single shared memory machine and cannot be divided by the scheduler.

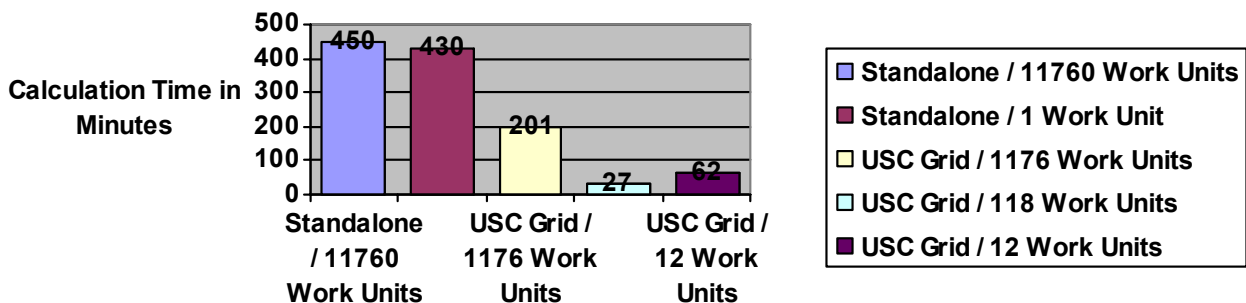
5. Results of Benchmarks Run:

We ran five benchmark map calculations. Two of the calculations were done on the SCEC/CME standalone server. Three of these calculations were done using the USC Grid Condor Pool.

On the SCEC/CME server, we ran one benchmark consisting of 1 Work Unit. We ran a second benchmark consisting of 11760 Work Units.

On the USC Grid, we ran three benchmarks: 1176 Work Units, 118 Work Units, and 12 Work Units. During the time period we were running these benchmarks, the number of nodes available to us in the USC Grid Condor Pool was 100 nodes. Our Work Units were schedule across these nodes by the Grid scheduler.

PSHA Hazard Map Calculation Time



6. Discussion:

Our benchmark measurements confirm our expectation that instantiating a JVM for each hazard curve calculation causes measurable overhead. JVM instantiation in the standalone Map calculation took 30 minutes, more than the total calculation time in our best case Grid benchmark.

The benchmarks also show that the optimal number of Work Units is the number of nodes available for use in the calculation. Since the number of nodes available in the Condor Pool is variable, it appears we should plan on dynamically adjusting the number of our Work Units to approximately match the number of nodes available in the Pool at run-time.

From these initial benchmarks, it appears that the Condor Pool capability of the USC Grid matches the computing requirements of our Hazard Map calculations. Our run-time for Hazard Map calculations was improved by more than an order of magnitude by using this USC Grid capability.