

Optimizing earthquake data flow allows scientific research on 'The Big One'

May 29 2024, by Leila Okahata



Credit: Wilson Malone from Pexels

No one can precisely predict when an earthquake is to happen. Since the 1994 6.7-magnitude Northridge earthquake in Los Angeles County that caused 72 deaths, 9,000 injuries and \$25 billion in damages, Southern

California has been anxiously waiting for "The Big One": a devastating quake predicted to be at least a 7.8 magnitude and 44 times stronger. Seismologists can only say that it may happen within the next 30 years.

Although scientists cannot forecast when and where earthquakes will strike, preparation is key to improving society's resilience to [large earthquakes](#). In particular, the USC-based Statewide California Earthquake Center (SCEC) developed [CyberShake](#), a computational platform that simulates hundreds of thousands of earthquakes to calculate regional seismic hazard models.

Revealing geographical areas in Southern California most at risk for intense shaking, its results have influenced Los Angeles building codes and the design of the earthquake models at the U.S. Geological Survey, the nation's largest earth and geological science mapping agency.

CyberShake studies—and much of modern science, however—are highly data and computing-intensive. With multi-step calculations that feed into numerous interconnected computational tasks executing on local and national supercomputers to simulate 600,000 different earthquakes, CyberShake's scientific workflow is complex. USC Viterbi's Information Sciences Institute (ISI) houses the tools to generate and manage such massive data.

Ewa Deelman, a research professor in [computer science](#) and research director at ISI, has continuously designed and updated, since 2000, an automated workflow management system called Pegasus.

Optimized workflows

[Pegasus](#)—named after Planning for Execution and Grids (PEG) and Deelman's love for horses—turns research experiments into optimized workflows. It can be used by scientists in various fields from seismology

to physics to bioinformatics because of its abstract design.

Deelman likens it to a cooking recipe: "You can use the same recipe in different kitchens. Different users can run the recipe (the workflow) but with their own cookware (computational resources). When you design things in a broad enough way, they become widely applicable."

In 2016, scientists from the Laser Interferometer Gravitational-Wave Observatory (LIGO) utilized Pegasus to capture [gravitational waves](#) in the universe, confirming Albert Einstein's General Theory of Relativity and earning the 2017 Nobel Prize for physics. During the 16-year collaboration between ISI computer scientists and LIGO members, the software managed thousands of workflows with millions of tasks.

The Collaborative and Adaptive Sensing of the Atmosphere (CASA), an engineering research center dedicated to improving hazardous weather prediction and response, has also ported its pipelines into Pegasus. As severe weather can slow and compromise local resources and computing capacity, the program sends CASA's data into cloud infrastructures to ensure continuous workflow.

Inspired by animal behaviors

CyberShake has relied on Pegasus for the past 15 years, including its most recent study with its largest set of earthquake simulations yet. Pegasus managed 2.5 petabytes of data and ran 28,120 workflow jobs over 108 days to produce seismic hazard maps in 772,000 node-hours.

"Without Pegasus, there's no way we'd be able to do this kind of science," said Scott Callaghan, a computer scientist at SCEC and lead developer on CyberShake. SCEC will be expanding CyberShake to Northern California, now using the fastest supercomputer in the world, Frontier. Pegasus will continue to remain at their side.

"Every time we do one of these studies, we always encounter unexpected challenges. But I'm confident that, with any workflow issues, the Pegasus team will be able to help us work through them so that we can continue getting cutting-edge science done," Callaghan said.

Deelman is now conducting research and conceptualizing SWARM, another workflow management system inspired by the savvy coordination of group behaviors among social animals, like ants. She also plans to enhance Pegasus' decision-making with artificial intelligence, reimagining how workflow systems will operate in the future.

Provided by University of Southern California

Citation: Optimizing earthquake data flow allows scientific research on 'The Big One' (2024, May 29) retrieved 21 June 2024 from <https://phys.org/news/2024-05-optimizing-earthquake-scientific-big.html>

This document is subject to copyright. Apart from any fair dealing for the purpose of private study or research, no part may be reproduced without the written permission. The content is provided for information purposes only.