



## GRID COMPUTING YIELDS EARTHQUAKE FORECAST

By Pam Frost Gorder

On the second day of an earthquake research meeting in Tokushima, Japan, John Rundle sits in his hotel room, making a phone call through his laptop computer. Although it's only 6:00 a.m. local time, he's long been awake, still anchored in California time. In a few hours, he'll share some of the latest earthquake research from his laboratory at the University of California, Davis, with the conference attendees. His forecast, set to appear in the 1 December 2006 issue of *Physical Review Letters*, is an eerie one: northern California is likely to experience an earthquake of at least magnitude 6 on the Richter scale in the next 18 months.

The statement is astonishing, not just because it predicts a major earthquake, but because it does so with such precision. The physical forces that govern seismic events are so complex that scientists have struggled to assemble forecasts on the scale of decades. Now a new statistical technique has enabled Rundle's team to narrow the forecast window to less than three years—and in the present case, to only a year and a half. The calculations require a day of supercomputing time, but he and his partners plan to automate research-quality forecasts and make them available online. "Every day, we will recompute forecasts for these large earthquakes for northern and southern California," he says. "You can only do this with grid computing."

The idea of linking supercomputers into a computational grid to confront big problems isn't a new one, but Rundle and others in Earth science are doing something different. They're linking grids together—effectively, using grids of grids—thanks to a software movement that takes its name from the musical phenomenon known as the *mashup*.

### Data on Demand

Artists have made musical mashups since at least the 1950s, combining portions of two or more songs into one. But the genre achieved new popularity in the 1990s, when music-editing software became widely available, and people could publish homemade mashups on the Internet. As music critic Sasha Frere-Jones explains, "The most celebrated mashups are melodically tuned, positing a harmonic relationship between, say, Madonna's voice and the Sex Pistols' guitars" ("1

+ 1 + 1 = 1," *The New Yorker*, 10 Jan. 2005).

The philosophy of software mashups isn't so different: two complementary Web applications merge to create a custom tool. The number is growing every day, as companies like Google and Yahoo make the APIs for their programs publicly available. (Google offers its developer tools at <http://code.google.com/> and Yahoo at <http://developer.yahoo.com/>.) As of mid-November 2006, the comprehensive ProgrammableWeb site ([www.programmableweb.com](http://www.programmableweb.com)) lists more than 1,200 mashups, nearly two-thirds of which are map-related. The most commonly mashed-up API is Google Maps. Users have combined it with everything from apartment listings to crime statistics—any data set with a geographic component will do.

Marlon Pierce, assistant director of the Community Grids Lab (CGL) at Indiana University, was already developing grid applications when Google Maps debuted in the summer of 2005. That December, at the meeting of the American Geophysical Union in San Francisco, he reported CGL's first Google mashup: a seismic data map of the American Southwest.

"A lot of people are interested in Google Maps and Google Earth," Pierce says. "Google did a really good job making them both easy to use, but they also made it easy for people to develop new applications with them. So that was a lesson for the rest of us: a particular application can be complicated, but the face it presents to the outside world should be simple."

The mashup helps get the most out of computing grids. An application can reside on one grid, and a data set on another. The person who wants to use both could be anywhere—as long as all the applications and data sets follow compatible standards, and toolkits exist to bring them together, the combinations are endless.

Pierce is project director for the CrisisGrid ([www.crisisgrid.org](http://www.crisisgrid.org)), a project at CGL that aims to build and integrate such services on the Web, especially for Earth science. "This is our basic quest, to integrate those applications—what I'd call 'Execution Grid' services—with 'Data Grid' services. We want to provide a common framework," he says. CrisisGrid data could be anything from GPS station output to

## WEB TRENDS

For a brief look at current events, including program announcements and news items related to science and engineering, check out the following Web sites:

- Networking Technology and Systems (NeTS; [www.nsf.gov/publications/pub\\_summ.jsp?ods\\_key=nsf07507](http://www.nsf.gov/publications/pub_summ.jsp?ods_key=nsf07507)). The US National Science Foundation (NSF) is soliciting proposals from the networking community for its Networking Technology and Systems (NeTS) program. Proposal deadline is 22 January 2007.
- Science of Design (SoD; [www.nsf.gov/publications/pub\\_summ.jsp?ods\\_key=nsf07505](http://www.nsf.gov/publications/pub_summ.jsp?ods_key=nsf07505)). Proposals for the NSF's SoD program are due 5 February 2007.
- Software Development for Cyberinfrastructure (SDCI; [www.nsf.gov/publications/pub\\_summ.jsp?ods\\_key=nsf07503](http://www.nsf.gov/publications/pub_summ.jsp?ods_key=nsf07503)). Proposals for the SDCI program are due 22 January 2007.
- "NSF Awards Texas Advanced Computing Center" ([www.nsf.gov/publications/pub\\_summ.jsp?ods\\_key=pr06137](http://www.nsf.gov/publications/pub_summ.jsp?ods_key=pr06137)). The NSF's award of US\$59 million will support a high-performance computing system to aid the nation's scientists and engineers.
- "NSF Awards \$76 million for 2006 Science and Technology Centers" ([www.nsf.gov/publications/pub\\_summ.jsp?ods\\_key=pr06146](http://www.nsf.gov/publications/pub_summ.jsp?ods_key=pr06146)). The NSF's award of \$US76 million supports 17 science and technology centers in areas such as next-generation polymers, climate modeling, oceanography, and coastal environments.

roadmaps to census data—all of which can be managed by geographical information system (GIS) services. "These problems may take a long time to calculate, and some of them have to be solved on supercomputers, but the data you get out of them is not at all abstract. You can display them with standard GIS-type services."

He and his colleagues use a distributed computing methodology called publish/subscribe. Users subscribe to a data set, which is pushed to them and published in a way similar to a subscription newsfeed. Middleware filters convert the data to desired formats, and other filters insert the data into Web tools such as Google Maps. In this way, the data set and the Web application that created it are out of the hands of the originators and into users' hands. Pierce says that's the idea of the mashup: "You just provide a service; you don't really care what anybody does with it."

Andrea Donnellan leads the Interferometric Synthetic Aperture Radar (InSAR) science team at NASA's Jet Propulsion Laboratory (JPL), which gathers satellite imagery of seismic activity. She predicts that interoperable Web services will be even more common in the future than they are to-

# computing

in SCIENCE & ENGINEERING

## EDITOR IN CHIEF

**Norman Chonacky**  
[cise-editor@aip.org](mailto:cise-editor@aip.org)

## ASSOCIATE EDITORS IN CHIEF

**Jim X. Chen, George Mason Univ.**  
[jchen@cs.gmu.edu](mailto:jchen@cs.gmu.edu)

**Denis Donnelly, Siena College**  
[donnely@siena.edu](mailto:donnely@siena.edu)

**Douglass E. Post, Carnegie Mellon Univ.**  
[post@ieee.org](mailto:post@ieee.org)

**John Rundle, Univ. of California, Davis**  
[rundle@physics.ucdavis.edu](mailto:rundle@physics.ucdavis.edu)

## EDITORIAL BOARD MEMBERS

Klaus-Jürgen Bathe, MIT, [kjb@mit.edu](mailto:kjb@mit.edu)

Antony Beris, Univ. of Delaware, [beris@che.udel.edu](mailto:beris@che.udel.edu)

Michael W. Berry, Univ. of Tennessee, [berry@cs.utk.edu](mailto:berry@cs.utk.edu)

Bruce Boghosian, Tufts Univ., [bruce.boghosian@tufts.edu](mailto:bruce.boghosian@tufts.edu)

Hans-Joachim Bungartz, Institut für Informatik,  
[bungartz@in.tum.de](mailto:bungartz@in.tum.de)

George Cybenko, Dartmouth College, [gvc@dartmouth.edu](mailto:gvc@dartmouth.edu)

Jack Dongarra, Univ. of Tennessee, [dongarra@cs.utk.edu](mailto:dongarra@cs.utk.edu)

Rudolf Eigenmann, Purdue Univ., [eigenman@ecn.purdue.edu](mailto:eigenman@ecn.purdue.edu)

David Eisenbud, Mathematical Sciences Research Inst.,  
[de@msri.org](mailto:de@msri.org)

William J. Feiereisen, Los Alamos Nat'l Lab, [wjf@lanl.gov](mailto:wjf@lanl.gov)

Geoffrey Fox, Indiana State Univ., [gcf@grids.ucs.indiana.edu](mailto:gcf@grids.ucs.indiana.edu)

Sharon Glotzer, Univ. of Michigan, [sglotzer@umich.edu](mailto:sglotzer@umich.edu)

Steven Gottlieb, Indiana Univ., [sg@indiana.edu](mailto:sg@indiana.edu)

Anthony C. Hearn, RAND, [hearn@rand.org](mailto:hearn@rand.org)

Charles J. Holland, Darpa, [charles.holland@darpa.mil](mailto:charles.holland@darpa.mil)

M.Y. Hussaini, Florida State Univ., [myh@cse.fsu.edu](mailto:myh@cse.fsu.edu)

Rachel Kuske, Univ. of British Columbia, [rachel@math.ubc.ca](mailto:rachel@math.ubc.ca)

David P. Landau, Univ. of Georgia, [dlandau@hal.physast.uga.edu](mailto:dlandau@hal.physast.uga.edu)

R. Bowen Loftin, Texas A&M University, Galveston,  
[loftin@tamug.edu](mailto:loftin@tamug.edu)

B. Vincent McKoy, California Inst. of Technology,  
[mckoy@its.caltech.edu](mailto:mckoy@its.caltech.edu)

Jill P. Mesirov, Whitehead/MIT Ctr. for Genome Research,  
[mesirov@genome.wi.mit.edu](mailto:mesirov@genome.wi.mit.edu)

Constantine Polychronopoulos, Univ. of Illinois,  
[cdp@csr.d.uiuc.edu](mailto:cdp@csr.d.uiuc.edu)

William H. Press, Los Alamos Nat'l Lab., [wpress@lanl.gov](mailto:wpress@lanl.gov)

John Rice, Purdue Univ., [jrr@cs.purdue.edu](mailto:jrr@cs.purdue.edu)

Ahmed Sameh, Purdue Univ., [sameh@cs.purdue.edu](mailto:sameh@cs.purdue.edu)

Henrik Schmidt, MIT, [henrik@keel.mit.edu](mailto:henrik@keel.mit.edu)

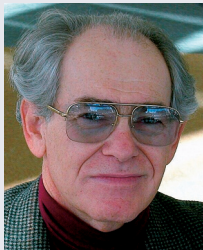
Greg Wilson, Univ. of Toronto, [gwilson@third-bit.com](mailto:gwilson@third-bit.com)

## CONTRIBUTING EDITORS

Francis Sullivan, The Last Word, [fran@super.org](mailto:fran@super.org)

Paul F. Dubois, Café Dubois, [paul@pfdubois.com](mailto:paul@pfdubois.com)

## OBSERVATOIRE LANDAU



By Rubin Landau, news editor

**C**iSE, the name of this publication, stands for *Computing in Science & Engineering*. Although you can interpret this literally to mean that *CiSE* describes computing activities in science and engineering—which it does—I like to interpret it more

broadly to mean that *CiSE* describes and promotes the integration of computing into science and engineering. Consequently, I see part of *CiSE*'s value as presenting the multidisciplinary viewpoint known as “computational science,” which presumes that it’s important to understand some of the computer science and applied mathematics that are part of doing science and engineering on a computer.

In my first column, I reported on several conferences this past summer that provided evidence of a wide interest in having the next generation of students learn more about computational science and computational physics. I tried to make the reporting objective and accurate, with my opinions kept to a minimum (not something I am noted for). In contrast, this column presents a variety of stories relating to computational science and what I have observed as people’s views of it. In some cases, these episodes might indicate that we still have a way to go before computational science is broadly accepted as a bona fide way to do science. In others, they indicate the environment in which we do computational science. In case some of my colleagues read this column, I won’t cite original sources. Likewise, although the essence of these episodes reflects actual events, I can’t claim that the actual words are accurate, given the vagaries of memory (do we sometimes remember what we wish we had said instead of what was?).

## Opinions and War Stories

One of my first, and probably formative, experiences came from my thesis advisor. I recognized early in my career that it was unlikely for me to find success in attacking hard, unsolved physics problems analytically when those smarter than I had been unable to do so. However, I also realized that I could solve some of these numerically, something that hadn’t been done well before due to insufficient computing power. I tried to solve my problems with Fortran (the other choices were Cobol and Pascal), only to discover my advisor’s disapproval that I was willing to stake my professional reputation on computer code I had never seen. I was instructed to look, at the very least, at the assembly language code to know what was being computed, but if I were really a serious scientist, I should look at the basic machine. (Although good advice, it doesn’t mean I followed it.)

More recently, some graduate students came to me to inquire about thesis research in computational physics and appeared to be doing so with second thoughts. When I finally asked why, the students told me they had also spoken to other theoretical physicists who told them that they would always be viewed as second-class theorists if they had to resort to computers to solve problems. (I guess the establishment always tries to retain the established.)

In sort of the converse to the previous episode, I heard a talk by a renowned author of theoretical physics texts. In this case, the lecturer had a student explore charge distributions near the edges of a conductor. Because they had the “analytic” solution of Laplace’s equation in the form of a Fourier series (known to exhibit Gibb’s overshoot), to them it was just a question of adding up enough terms in the series to obtain a convergent answer on the computer. It was painful to hear how this problem, which could have been solved rather simply and quickly using a finite difference algorithm and a relaxation technique, required large amounts of the student’s time and required the student to get multiple hours

day, and she expects more scientists will apply pattern-recognition methods to grid data. Donnellan looks forward to a future InSAR mission that will cover the entire globe and produce roughly 200 Gbytes of data per day. “Computational methods will be necessary for ingesting and understanding the data,” she says.

### Real Hazards, Virtual California

Unlike hurricanes, which form seasonally and can be spotted before they endanger life, earthquakes can happen any time, anywhere without warning. On 15 November 2006—less than a week after Donnellan and Rundle returned to the US from the meeting in Tokushima—an 8.1-magnitude earthquake struck the northernmost part of Japan. The next day, a 5.2-magnitude quake struck Taiwan. As of this writ-

ing, no casualties have been reported from either quake, although the first one sent tsunami waves across the Pacific Ocean that damaged docks in northern California.

Experts don’t question whether the “Big One”—an earthquake of magnitude 8 or more—will strike California via the San Andreas Fault. It’s only a question of when, where, and how big. The 1906 San Francisco earthquake was blamed on the fault, and probably would have registered near magnitude 8 had the scale existed then. Earth scientists are working hard to develop forecasts to help California residents and government agencies prepare for future quakes. As a result, the entire fault system and much of the southwestern US are dotted with permanent GPS stations that enable the monitoring of every tiny tremor. All that seismic data is archived, and much of it is publicly available.

of supercomputer time. Moreover, serious questions about the solution still persisted after everything was completed.

Just this past summer, after recollecting some of these “war stories” in a talk, an audience member came up to me to relate his experiences. He told me that after retiring from a career of doing state-of-the-art multiphysics simulations at a military laboratory, he accepted a faculty position in a physics department. Upon arriving in his new department, he informed the chair that he was experienced in physics simulations and would be happy to develop some new courses along those lines. He was promptly informed that “here at Cupcake U, we do real physics, not simulated physics.” (He has now outlasted the chairman and teaches simulations.)

### Idealism Meets the Real World

A most interesting finding by the American Institute of Physics is that physics graduates rate scientific problem solving with computers and programming as the two most valuable skills they take to the workplace. My surveys indicate that only a very small fraction (roughly 2 percent) of many physics curricula are devoted to computation, and that students learn the physics better when they must apply it in a problem-solving context on the computer. Ironically, I have been told that my computational physics developments are harmful as they “detract from the time students have to spend on physics.” (I suspect I might find this more humorous if it hadn’t appeared in a review of my grant proposal.)

I suspect that most researchers view their work as a search for the truth, even if its scope might be rather narrow. Likewise, in the interest of safety and responsibility, I suspect that most engineers place a high value on obtaining reliable numbers. Accordingly, I was taken aback when, during a class on parallelization taught at a leading supercomputer center, the instructor indicated how important it is to remove all synchronization points in a program because they slow it down. When I interrupted the lecture to point out that synchroniza-

tion might be necessary to ensure that the computed results were valid, the instructor seemed perplexed. After some thought, he replied that he “saw where I was coming from, but that it was speedup that gets you the next grant.”

Along a similar vein, I recall a colleague who came across a complicated-looking partial differential equation to describe unusual quantum waves. To him it looked like an equation that “just called out to be solved by Matlab” (I think this means that he would otherwise have to get his hands dirty to write a program to do it); to me, it looked like an equation in which you had to be very careful. As might be expected, his solution exhibited some unusual ripples, which he interpreted as a new physics discovery. While I remained quiet, I recalled the adage that “nothing looks more like a new effect than some bad numbers.” (Yes, I know the story about how this attitude kept Lorenz from publishing his chaotic results earlier.)

Let me end by describing a scene that I have seen played out in various ways. A talk begins with an attentive audience, but shortly thereafter, someone in the front row nods off and doesn’t awaken until the question period, at which point they ask some questions that the speaker has difficulty answering. For a computational talk, these questions might include the following: have you compared the solutions from your code to those for a similar analytic problem? Can you describe the algorithm used to solve the equations, and why you chose it? Can you tell me explicitly just what the equations are that you solved or computed? Can you tell me what the error is in those curves you have been showing? Just what is it that is being left out of this solution? Have you checked that the long-range truncations made in your formulation don’t adversely affect your solution? Who actually wrote the programs used to solve this problem?

I suspect that many readers have similar situations that they have experienced, and if you wish to share them, I look forward to hearing from you at [rubin@physics.oregonstate.edu](mailto:rubin@physics.oregonstate.edu).

Such a data set is maintained by the Scripps Orbit and Permanent Array Center (SOPAC) at the Scripps Institution of Oceanography at the University of California, San Diego. Pierce and his team, led by Geoffrey Fox at Indiana University, are working with SOPAC to convert a Web application that previously processed daily data to one that can process real-time data. SOPAC uses data mining software to pick out modes in seismic activity—patterns of GPS station movement.

“Our challenge is to apply this mode of detection to data that comes out once per second instead of once per day. We don’t know what kind of modes to expect, so it will be interesting to see what might be hidden in those GPS signals,” Pierce says.

Rundle’s UC Davis team used data mining—a combina-

tion of algorithmic techniques called Relative Intensity Pattern Informatics (RIPI)—to construct its new forecast. To test the technique, they created a “hindcast” of seismic activity in northern and southern California since 1960. RIPI correctly predicted the time “windows” during which 16 out of 17 major groups of earthquakes occurred. “That’s not a random coincidence. There’s something causal at work,” Rundle says. “We’ve discovered an underlying pattern in the timing of these earthquakes.”

If RIPI is correct, the next major earthquake to strike California will happen in the northern part of the state by mid 2008. Rundle emphasized that the situation for southern California could change at any time, but at least his method gives an indication of whether a region is at high risk. Soon, he and his group believe they might be able to predict the locations



## DEPARTMENT EDITORS

**Book Reviews:** Mario Belloni, Davidson College, mabelloni@davidson.edu

**Computing Prescriptions:** Isabel Beichl, Nat'l Inst. of Standards and Tech., isabel.beichl@nist.gov, and Julian Noble, Univ. of Virginia, jvn@virginia.edu

**Computer Simulations:** Muhammad Sahimi, University of Southern California, moe@iran.usc.edu, and Dietrich Stauffer, Univ. of Köln, stauffer@thp.uni-koeln.de

**Education:** David Winch, Kalamazoo College, winch@TaosNet.com  
**News:** Rubin Landau, Oregon State Univ., rubin@physics.oregonstate.edu

**Scientific Programming:** Konstantin Läufer, Loyola University, Chicago, klauffer@cs.luc.edu, and George K. Thiruvathukal, Loyola University, Chicago, gkt@cs.luc.edu

**Technology:** James D. Myers, jimmyers@ncsa.uiuc.edu

**Visualization Corner:** Claudio T. Silva, University of Utah, csilva@cs.utah.edu, and Joel E. Tohline, Louisiana State University, tohline@rouge.phys.lsu.edu

**Your Homework Assignment:** Dianne P. O'Leary, Univ. of Maryland, oleary@cs.umd.edu

## STAFF

**Senior Editor:** Jenny Stout, jstout@computer.org

**Group Managing Editor:** Steve Woods

**Staff Editors:** Kathy Clark-Fisher, Rebecca L. Deuel, and Brandi Ortega

**Contributing Editor:** Cheryl Baltes and Joan Taylor

**Production Editor:** Monette Velasco

**Magazine Assistant:** Hazel Kosky, cise@computer.org

**Technical Illustrator:** Alex Torres

**Publisher:** Angela Burgess, aburgess@computer.org

**Associate Publisher:** Dick Price

**Advertising Coordinator:** Marian Anderson

**Marketing Manager:** Georgann Carter

**Business Development Manager:** Sandra Brown

## AIP STAFF

**Circulation Director:** Jeff Bebee, jbebee@aip.org

**Editorial Liaison:** Charles Day, cday@aip.org

IEEE ANTENNAS AND  
PROPAGATION SOCIETY LIAISON

Don Wilton, Univ. of Houston, wilton@uh.edu

## IEEE SIGNAL PROCESSING SOCIETY LIAISON

Elias S. Manolakos, Northeastern Univ., elias@neu.edu

## CS PUBLICATIONS BOARD

Jon Rokne (chair), Michael R. Blaha, Mark Christensen, Frank Ferrante, Roger U. Fujii, Phillip Laplante, Bill N. Schilit, Linda Shafer, Steven L. Tanimoto, Wenping Wang

## CS MAGAZINE OPERATIONS COMMITTEE

Bill N. Schilit (chair), Jean Bacon, Pradip Bose, Arnold (Jay) Bragg, Doris L. Carver, Kwang-Ting (Tim) Cheng, Norman Chonacky, George Cybenko, John C. Dill, Robert E. Filman, David A. Grier, Warren Harrison, James Hendler, Sethuraman (Panch) Panchanathan, Roy Want

## EDITORIAL OFFICE

COMPUTING in SCIENCE & ENGINEERING

10662 Los Vaqueros Circle, PO Box 3014, Los Alamitos, CA 90720

phone +1 714 821 8380; fax +1 714 821 4010; www.computer.org/cise/



IEEE Antennas &  
Propagation Society

IEEE

Signal Processing Society



of individual major earthquakes to within 50 kilometers or less.

Real seismic data fueled this new forecast, but the team is also using simulated data in a project called Virtual California. It applies everything that scientists know about the physics of fault systems to simulate millions of years of earthquakes. Rundle hopes to mine the simulations for clues as to how the real San Andreas operates.

Virtual California is one of three simulators under development for the QuakeSim project at NASA (<http://quake.sim.jpl.nasa.gov/>). Donnellan heads this project, which aims for interoperability between all its freely downloadable tools.

## Multidisciplinary, Multigrad


Pierce sees mashups as a democratizing force in the building of grid applications. He also suspects that letting people develop their own tools will be a boon to research. "I think we should encourage scientific mashups. When you look at the grid projects that have been the most successful, like high-energy physics or astronomy, the work has been led by the physical scientists," he says. "They have been making their own software."

Not everyone is a programming expert, and that's why mashups work so well. Scientists don't need to know specific details about how the tools work, as long as they know the parts that matter for the problem they're trying to solve.

Still, Pierce won't call the use of scientific mashups a paradigm shift. "I think the possibility [of a paradigm shift] is there, but I think it will take some very successful applications to convince people that this is what they should do."

JPL's Donnellan agrees that ease of use is the key. "The more off-the-shelf and plug-and-play grid computing is, the more earthquake scientists will make use of the available resources," she says.

To Rundle, the progress being made by his group at UC Davis (<http://hirsute.cse.ucdavis.edu/~rundle/>), as well as by the CrisisGrid and QuakeSim projects, proves that there's a major place for computational science within earthquake forecasting. He calls for more collaboration between the Earth and computational sciences and statistical physics, which is especially suited for dealing with multidimensional data systems such as earthquake phenomena.

"Without the ideas of statistical physics, we would never have been able to do this," Rundle says. "Much of the theory we used came from areas well outside of Earth science, and that has to continue if we're going to make progress." 

**Pam Frost Gorder** is a freelance science writer based in Columbus, Ohio.