N news

　　　　　Kirk L. Kroeker

Grid Computing's Future

Outreach programs and usability improvements are drawing many researchers to grid computing from disciplines that have not traditionally used such resources.

IN RECENT YEARS, several powerful research grids consisting of thousands of computing nodes, dozens of data centers, and massive amounts of bandwidth have emerged, but few of these grids have received much attention in the mainstream media. Unlike seti@home, folding@home, and other highly focused grid projects that have captured the popular imagination by allowing home users to donate compute cycles, the big research grids are not accessible to the public and their fame does not extend far beyond the researchers who use them. Outreach teams and usability engineers at the largest of these new grids, such as Naregi, Egee, and TeraGrid, are trying to change that reality by helping to facilitate the adoption of grid technologies in fields that have not traditionally used grid-based supercomputing resources.

TeraGrid, said to be the world's largest distributed network infrastructure for open scientific research, is one such network that has quietly been making waves in research communities outside computer science, and is helping to solve complex problems in biology, physics, medicine, and numerous other fields. TeraGrid consists of 11 data-center sites located around the U.S.
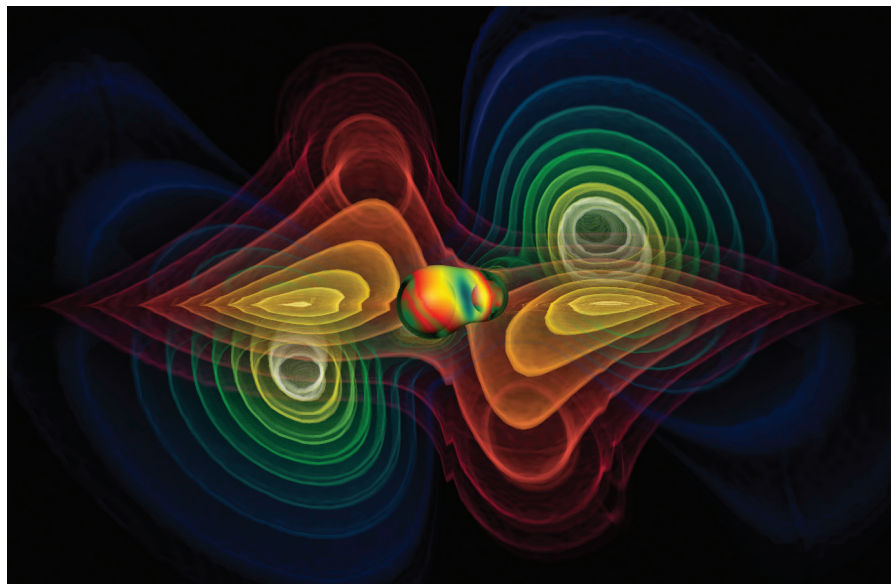
A grid-based computer simulation of the gravitational waves produced as two black holes merge with each other to form a larger black hole.

and is tied to a 10-gigabit backbone that connects primary network facilities in Los Angeles, Denver, and Chicago. At maximum capacity, TeraGrid can produce more than a petaflop of computing power and store more than 30 petabytes of data. The project, started by the National Science Foundation in August 2001, has grown in the past five years from fewer than 1,000 active users in 2005 to nearly 5,000 active us-

ers working on almost 1,600 projects at the end of 2009.

Matthew Heinzel, director of TeraGrid's grid infrastructure group, says that TeraGrid's outreach teams have done an excellent job drawing researchers from fields outside computer science. To obtain CPU time on the grid, scientists simply submit a request that describes the work to be done; extensive CPU-time requests are subject

Computers in Scientific Research

In 2009, Greg Wilson published details of a research project designed not only to determine how scientists learn what they know about using computers and how widely they share their software, but also to gauge the ratios of scientists who use desktop systems versus local clusters and grid-computing resources. Conventional wisdom suggests that as clusters and grids become more sophisticated and increasingly user friendly, more scientists will use them instead of less powerful desktop PCs. The findings Wilson and his colleagues produced run counter to that thinking.

While hard numbers indicate that more scientists each year are using grid-based resources to conduct their research, Wilson's project indicates that amount of work is relatively small compared to the vast number of projects still being run on the desktop. Some 81% of those participating in the research said they primarily use desktop machines, with only 13% saying they use intermediate-sized machines such as local clusters, and 6% saying they use supercomputers for their research.

Wilson, until recently a computer science professor at the University of Toronto and now at work full time on a course designed to address what he perceives as a computer-skills shortfall among scientists, says these figures might very well change in the years ahead. "If we're lucky, though, most scientists won't notice," he says. "Behind the scenes, an increasing number of data services will migrate upward, but the vast majority of scientists will reach them through desktop interfaces, because that's all they'll have the skills for."

That trend might already be materializing. Wilson says that in the year since the details of his research were published, he has witnessed an increased interest in commodity clouds, such as Amazon's Elastic Compute Cloud. "The course Titus Brown just taught at Michigan State [http://bit.ly/aE5Qpg] would have been a lot harder on everyone without pay-and-play," says Wilson, who predicts that more researchers will begin using such systems, which can be run independently from supercomputing-style applications.

As for the efforts to make grids easier to use and to draw researchers from fields that don't typically train scientists in supercomputer programming, Wilson says these efforts will fail outright unless scientists first invest in basic computing skills. "Asking someone who doesn't know how to modularize or test a program to use clouds and GPUs is as sensible as taking someone who's just learning how to drive a family car and putting them behind the wheel of a transport truck on the highway at rush hour," he says.

The best solution, according to Wilson, is to teach scientists some basic computational skills so they can tackle the challenges posed by advanced resources without expending heroic effort, which is the primary goal of the project Wilson is working on now, called Software Carpentry. As for future research to explore ways of assessing the impact of computers on scientists' productivity, Wilson wasn't able to find funding. "This is one of the reasons I left academia and returned to industry," he says. "As in health care, it seems that most people are more interested in pushing new things to market than in finding out what works and what doesn't." —*Kirk L. Kroeker*

to a quarterly peer-review process. Surprisingly, molecular biosciences, physics, astronomical sciences, chemistry, and materials research top the list of disciplines whose researchers are using the grid. "In some cases, we are victims of our own success," says Heinzel. "We've extended our user base well beyond our goals." But the downside to this success, he says, is that TeraGrid is being asked for far more CPU cycles than it can provide.

Still, Heinzel says that having more demand than capacity is a positive force that generates an impetus for the network to grow regularly. But the challenges facing Heinzel and TeraGrid's infrastructure group, which is run through the University of Chicago in partnership with the Argonne National Laboratory, extend beyond mere compute cycles and bandwidth. As more users from outside computer science are drawn to TeraGrid to run complex computational problems, one ongoing and key challenge is usability for those who do not have the necessary computer science skills. (The sidebar "Computers in Scientific Research," above, describes one project designed to address this skills shortfall.)

TeraGrid's resources, coordinated by UNIX-based Globus grid software, are integrated through a service-oriented architecture, with all systems running a common user interface. While the interface includes job-submission and data-management tools, the network requires problems to be translated into specific types of grid-ready code. To accommodate those who do not have the skills needed to translate their problems into compatible code, TeraGrid has allocated funding to embed dedicated supercomputer specialists in research teams for up to one year. Heinzel says that these specialists, who write and refine project-specific code, are among TeraGrid's most valuable resources, and represent a large part of the infrastructure group's budget.

"You can't just put a big piece of hardware on the floor and say, 'OK guys, here you go,'" says Heinzel. "You need somebody with the skills to help people not only run the code but also improve the code."

Modeling Black Holes

One researcher who has conducted extensive work on TeraGrid is Erik Schnetter, an assistant research professor in the department of physics and astronomy at Louisiana State University. His research has modeled black holes, neutron stars, and other highly complex astrophysical objects. The most interesting of these objects, he says, are gamma-ray bursts, which are bright bursts of high-energy photons that are visible from Earth and are said to be generated by the most energetic events in the universe. "It turns out that these bursts emanate from billions of light years away, essentially at the other end of the universe," says Schnetter. "The fact that they are still so brightly visible here means that they must come from truly tremendous explosions."

The mechanism that creates these explosions, the source of the energy, is not completely understood. After decades of research, the astrophysics community found that one model, called the collapsar model, might help explain the gamma-ray bursts. "What we do in one of our projects is model stars that form a supernova, then form a black hole at their center, and then we study how the remaining material behaves," says Schnetter. "These are very complex systems, and modeling them is a large task."

The computer code used to calculate these models not only is complex, but also requires significant computational power. Schnetter's group has performed simulations on local workstations and clusters, but he says that any kind of production work that has a

high level of accuracy requires systems that are too large for a single university. For example, the nodes processing Schnetter's modeling code require communication with other nodes several times per second and a data-exchange rate of about one gigabyte per second.

"To finish a simulation in a reasonable time, we need to use hundreds or thousands of cores," he says. "That means we need to split the problem into many pieces, and we need to ensure that each of these pieces remains as independent from the others as possible." By using this modular technique, Schnetter's team can replace or exchange grid code if it becomes necessary to apply new physics or use a different hardware architecture.

Another project run on TeraGrid is an effort to understand the environmental impact of aviation. Conducted out of Stanford University by doctoral candidate Alexander Naiman and overseen by Sanjiva Lele, a professor in the department of aeronautics and astronautics, the project models condensation trails, the ice clouds formed by aircraft emissions. Naiman, whose research group specializes in computational fluid dynamics and turbulence simulations, says the difficulty of contrail modeling becomes increasingly acute as the complexity of the model increases. "The more complex the flow, the higher resolution required to simulate it, and the more resources are needed," he says.

While Stanford has local supercomputing resources, they are in high demand. "TeraGrid provides relatively large and modern supercomputing resources to projects like ours that have no other supercomputing support," he says. The simulation code that Naiman and his team run on TeraGrid was written at the Center for Turbulence Research at Stanford. Naiman says it was easy to get that code running on TeraGrid. The research group parallelized the program, a type of large eddy simulation, using standard message-passing interface strategies that Naiman says have been highly scalable on TeraGrid.

The contrail modeling project is ongoing, but so far the Stanford team has simulated the first 20 minutes of contrail development for several scenarios, producing terabytes of three-dimensional flow fields and other data, such as time histories for cloud ice mass. Naiman says that although the TeraGrid data is still undergoing analysis, it is likely to help improve understanding of the development of contrails and their environmental impact. "TeraGrid performs as advertised, providing us with CPU hours that we would not have had access to otherwise," he says. "We also take advantage of the large archival storage available on TeraGrid to ensure that important data is backed up."

Improving Grid Software

As for the future of research on grid networks, TeraGrid's Heinzel says he remains optimistic, but points out that improvements must be made in grid software not only to enhance ease of use for researchers such as Schnetter and Naiman, but also to take complete advantage of new generations of hardware. "You have to be almost a systems admin to set your parameters on data movement correctly so you can take full advantage of these systems," says Heinzel. "So the software really needs to mature."

Echoing these concerns, LSU's Schnetter points out that his research groups consist of people with widely varying degrees of supercomputer experience. "Teaching everybody how to use the different systems, and staying on top of what works best on what system, and which parameters need to be tuned in what way to achieve the best performance, is like herding cats," he says. "There are almost no GUIs for supercomputers, and most of the ones that exist are really bad, so that using them requires some arcane knowledge."

Schnetter says he hopes that grid-based supercomputing will have a much larger influence on the curriculum than it does today, especially with so few universities teaching scientific programming at the level required to effectively use grid resources. "The good students in my group learned programming by themselves, on the side, because they were interested," he says. Still, Schnetter suggests that such self-taught programming might not be sustainable in a world in which computers are becoming increasingly complex. "I hope that this changes in the next decade," he says.

Stanford's Naiman offers a similar observation. He says that while his work on the grid has been positive, the usability of the technology could be improved. For his part, TeraGrid's Heinzel says he remains optimistic about the accessibility of grids, and predicts that major usability improvements are on the way. He likens the evolutionary pace of these grid developments to how the Web quickly emerged from the Internet and now requires little more than a browser and a basic knowledge of hyperlinks. "If you know exactly how to use grid tools, they work effectively," says Heinzel. "Now we need to make them more user-friendly so we can get a wider audience."

In the future envisioned by Heinzel, grids will be manipulated easily by computer scientists while still providing friendly interfaces for researchers coming from other fields. Rather than predicting that the arrival of such technologies will take decades or more, Heinzel says that much progress will be made in the next few years alone. "We're going to see some big improvements in the usability of the grid and grid software in the next two to four years," he says. "Future systems will be very user-friendly with a high degree of abstracting the inner workings of what's going on from the end users." ▣

Further Reading

Ferreira, L., Lucchese, F., Yasuda, T., Lee, C.Y., Queiroz, C.A., Minetto, E., and Mungioli, A.S.R.
Grid Computing in Research And Education. IBM Redbooks, Armonk, NY, 2005.

Magoulès, F.
Fundamentals of Grid Computing: Theory, Algorithms, and Technologies. Chapman & Hall, Boca Raton, FL, 2009.

Neeman, H., Severini, H., Wu, D., and Kantardjieff, K.
Teaching high performance computing via videoconferencing, *ACM Inroads 1*, 1, **March 2010.**

Scavo, T., and Welch, V.
A grid authorization model for science gateways, *International Workshop on Grid Computing Environments 2007*, **Reno, NV, Nov. 11, 2007.**

Wong, J. (Ed.)
Grid Computing Research Progress. Nova Science Publishers, Hauppauge, NY, 2008.

Kirk L. Kroeker works in communications and has written extensively about the impact of emerging technologies.