Reprint Information

Permission to Copy

By Roxane Richter,
Technology/
Financial Editor

Welcome To The Mega-Grid

How the new eon of virtual shared Grid computing communities will alter and accelerate the next Internet decade

Imagine an era when your PC can e-mail Mom, pay bills and check your portfolio status – all while running a portion of the protein sequencing software from the Human Genome project, or searching for extraterrestrial intelligence by analyzing telescope data from Berkeley.

If it all sounds a little too sci-fi for you, companies around the globe today are signing up their computing brawn to help muscle-up corporate and research brain. The cooperative use of geographically distributed resources into a single powerful virtual computer has been dubbed metacomputing, seamless scalable computing, global computing, and more recently, Grid computing. The Grid may give birth to a global file-swapping network supporting data-intensive financial, oil/gas exploration and research and development companies all over the world.

Here's how it works: Think of Grid computing like a screensaver; it only activates when the PC is not in use for some specified (pre-set) amount of time. And so it goes with peer-to-peer (P2P) distributed computing and Grid computing. And while P2P pools cluster(s) or server(s) into a linked heterogeneous network, Grid computing enables "virtual communities" to share Web-linked resources as they pursue common goals and coordinate the maximum use of geographically distributed resources. (Interestingly, the term 'Grid' was chosen to suggest the idea of a 'power grid': a grid where scientists and users could "plug" into the computing infrastructure, much like plugging into an electrical power grid.)

Users handle the Grid as a single computing resource while management software (available from vendors such as Sun Microsystem, IBM and others) accepts jobs submitted by users and schedules them for execution on appropriate systems in the Grid, based upon pre-configured policies. In essence, users can submit thousands of jobs at a time, without ever being concerned where (on whose system) they'll run.

But no two grids are alike; one size doesn't fit all. There are three key classes of grids, explains Sun Microsystems, which scale from single systems to supercomputer-class compute farms that utilize thousands of processors:

Cluster Grids – are the simplest, consisting of one or more systems working together to provide a single point of access to users in a single project or department.

Campus Grids – enable multiple projects or departments within an organization to share computing resources. Organizations can use campus grids to handle a wide variety of tasks, from cyclical business processes to rendering, data mining, and more.

Global Grids – are a collection of campus grids that cross organization boundaries to create very large virtual systems. Users have access to compute power that far exceeds the resources available within their own organization.

And while Grid computing faces a mammoth technical task, seemingly the thorniest issues will involve social and political dimensions — in other words, how to facilitate sharing between strangers who have no established safety protocol, no prior relationship and, all too often, no trust.

150% Internet Traffic Growth

In February, Global Grid Forum 4 economist Robert Cohen said Internet traffic could grow eight times more than previously forecast over the next decade due to the commercial adoption of Grid computing and related P2P applications.

Cohen, of New York-based Cohen Communications Group, explained that Internet traffic would begin to double in 2003 and speed up to yearly rates of growth 150% or more from 2005-08, with Grid computing the key driver of bandwidth use. With rapid growth of Grid computing, P2P and server-to-server (S2S) traffic could account for nearly 90% of Internet traffic by 2008, he noted. (Current forecasts, such as McKinsey-JP Morgan, assume that annual growth will slow to about 60% by 2005, with traffic growth driven largely by video streaming and video file transfers, as reported by gridcomputing.com.)

"This is where Grid computing comes in, because it's the open protocols of the Grid running on every single system that enable 'intelligent' management capabilities to cut across different systems. Using these Grid protocols, we can realize this vision of creating, over time, an increasingly self-managing infrastructure," said Irving Wladawsky-Berger, IBM server group vice president of technology and strategy.

Future Formula

As with any newfangled, new-to-the-job technology, starting out is never easy. And Grid computing has proven no exception. The lack of standards for connecting the computing resources and applications that are outsourced has proven time-consuming, expensive and frustrating to both the vendors and the users. And until the computing community adopts global-based standards, each distributed engagement will take on a complex, solitary, costly and protracted labor-intensive effort.

In February, InternetNews.com reported that Big Blue (IBM) and two other Grid computing organizations were working with the Global Grid Forum, a broad-based organization working to establish Grid standards, and the Globus Project to expand Grid technology beyond the scientific and academic worlds to better support business applications.

Perhaps IBM's Wladawsky-Berger best summarized the future promise and pitfalls of Grid computing: "There is an incredible amount of potential in this technology of Grid computing," he said. "I feel a little bit as if I could have been addressing you in 1994 or 1995. It's as if back then I was saying, 'You know, there is this thing called the Internet that the research world has been making a lot of progress with'…I think the feeling now is similar, that the Grid is opening up a whole set of really important capabilities we have all been after for a long while. And now of course, starts the hard work..which is to take these technologies and translate them into real business value."

And that real business value definitely translates into the data-intensive environment of energy commodity risk management, oil and gas exploration and more.

A Sample Of Current Grid Projects:

Distributed.net (www.distributed.net): founded in 1997, this project has grown to encompass thousands of users; computing power has grown to become equivalent to that of more than 160,000 PII 266MHz computers working 24 hours a day, 7 days a week, 365 days a year.

SETI@home (setiathome.ssl.berkeley.edu): a Berkeley-led scientific experiment that uses Internet-connected computers in the Search for Extraterrestrial Intelligence (SETI). Anyone can participate by running a free program that downloads and helps analyze radio telescope data. Every four hours the amount of work done by SETI@home participants is posted, broken down according to various criteria.

FightAIDS@Home (www.fightaidsathome.org): a computational research project partnership between Entropia, the Olson Laboratory and the Scripps Research Institute, the site asks users "Go to battle against AIDS with your PC!" and use your computer to assist AIDS research. Through the donation of massive computing power, scientists have constructed software and a system to model the evolution of drug resistance and design anti-HIV drugs.

Folding@Home (www.gridcomputing.com): The "folding" of proteins can have serious effects, including many well known diseases like Alzheimer's, Mad Cow and Parkinson's disease. Since Oct. 1, 2000, over 100,000 worldwide users have participated in Folding@Home.

Genome@Home (www.nhgri.nih.gov/HGP): The Human Genome Project is nearing completion, and scientists need to put the huge amounts of data (over 50,000 genes and 3 billion nucleotide base pairs) from the Human Genome Project into biological and medical context, giving it real meaning. By running the Genome protein sequencing software, users lend the project their computer while they're not using it. Donators can see a small window showing the protein sequences that are being designed.

United Devices, Intel and University of Oxford Cancer Research Project (www. members.ud.com/projects/cancer): With the slogan "Don't just make a donation, make a difference," this project asks users to volunteer their PCs to help process molecular research being conducted by the Department of Chemistry at the University of Oxford in England and the National Foundation for Cancer Research.

GridSim (www.csse.monash.edu.au/~rajkumar/gridsim): the primary objective of this project is to investigate effective resource allocation techniques based on computational economy through simulation. The organization explores how significantly the local economy and the global positioning (e.g., the time zone) of a particular resource play in securing jobs under various pricing and demand/supply situations and other simulations.

GIMPS, the Great Internet Mersenne Prime Search (www.mersenne.org): If you own a computer, you could discover one of the most coveted finds in all of mathematics - a new Mersenne prime number. In addition to the joy of making a mathematical discovery, you might win the award and egg-head accolades that goes to the first person or group to discover a ten million-digit prime number.

EUROGRID (www.eurogrid.org): from Manchester to Zurich and Bergen to Warsaw, EUROGRID is a mega-project made possible through a European Commission grant. The EUROGRID project will demonstrate the use of Grids in selected scientific and industrial communities, address the specific requirements of these communities, and highlight the benefits of using Grids.

Other grid projects include EU DataGrid, Grid Physics Network (GriPhyN), Particle Physics Data Grid, UK Grid Support Centre, e-Sceince at Oxford University and more. For more information on Grid computing, see www.gridcomputing.com and www.gridcomputingplanet.com).