Competitive Review
August 2002   Issue 5

Grid Computing: Contender or Pretender?

Part 2: What does it all mean?

By Charles King

In the first part of this report, ***Will computational grids be bigger than the Internet?*** *we discussed the historical issues, technical advances, and industry efforts that paved the way for what has become known as grid computing. But is grid's development as linear as it might appear? As has been proven true in past occasions, a technology's logical development can veer severely off course or even off kilter as it reaches the often illogical mainstream. Though computational grids share the same university and research laboratory origins as the Internet and many industry watchers have claimed that grids represent the next technological and economic phase of Internet development, is there any earthly reason to expect that the commercial uptake of grid computing will be as dramatic or enthusiastic as what the Internet has enjoyed? At this juncture, we believe it wise to remember just how opaque the Internet's origins have become since it was exposed to and embraced by the commercial marketplace. That is a natural enough development, since markets tend to focus more on effect than cause, and most consumers are as curious about the technologies and industry standards underlying their forays on the Web as they are about the R&D efforts behind their automobile transmissions. We believe that a similar evolution is likely to occur as grid computing solutions enter the mainstream ...*

The Sageza Group, Inc.
836 W El Camino Real
Mountain View, CA 94040-2512
650·390·0700   fax 650·649·2302
London +44 (0) 20·7900·2819
Munich +49 (0) 89·4201·7144
Amsterdam +31 (0) 35·588·1546

In the first part of this report, **Will computational grids be bigger than the Internet?** we discussed the historical issues, technical advances, and industry efforts that paved the way for what has become known as grid computing. But is grid's development as linear as it might appear? As has been proven true in past occasions, a technology's logical development can veer severely off course or even off kilter as it reaches the often illogical mainstream. Though computational grids share the same university and research laboratory origins as the Internet and many industry watchers have claimed that grids represent the next technological and economic phase of Internet development, is there any earthly reason to expect that the commercial uptake of grid computing will be as dramatic or enthusiastic as what the Internet has enjoyed?

At this juncture, we believe it wise to remember just how opaque the Internet's origins have become since it was exposed to and embraced by the commercial marketplace. That is a natural enough development, since markets tend to focus more on effect than cause, and most consumers are as curious about the technologies and industry standards underlying their forays on the Web as they are about the R&D efforts behind their automobile transmissions. We believe that a similar evolution is likely to occur as grid computing solutions enter the mainstream. While in their purest forms computational grids offer once inconceivable solutions and opportunities for sharing and leveraging computing resources, this notion means more to data center staffers than it does to the business owners and managers who sign the purchase orders for those grids to be built. Given that, how will computational grid solutions (and their vendors) find success in the marketplace? First and foremost, they must be designed and developed to accomplish and enhance specific business tasks. To that end, current and future grid deployments have begun to follow a number of models with commercially recognizable benefits.

Grid Models

Despite similarities of technical demands, grids are deployed in a wide variety of form factors. Though current computational grids are largely designed and built to order, most follow one of the following models:

- Local or "cluster" grids: Denote the smallest grid infrastructures that tend to be deployed within single departments or to support specific industrial applications such as volume visualization tools for product design. Many in the grid community question whether there is any significant difference between cluster grids and conventional clustering solutions. Proponents suggest local grids can incorporate heterogeneous elements that are seldom seen in traditional cluster environments.
- Enterprise, campus, or intra-grids: Are trans-institutional, connecting multiple departments or workgroups within individual enterprises, and enabling more efficient use of computing infrastructures by allowing interdepartmental sharing of resources. For example, an intra-grid might enable multiple sales and marketing workgroups to share and leverage sophisticated data mining resources. Enterprise grids are entirely enabled behind corporate firewalls, utilizing existing campus networks and secure connections to remote offices and worksites.
- Partner grids: Refer to grid infrastructures that are shared between enterprises and their favored partners or suppliers for mutually supported projects. For example, a partner grid might be used to enable product design processes and engineering change orders between a vendor and its

manufacturing partners. Because they allow access to information and/or applications that would normally be kept secure behind company firewalls, partner grids are often built on dedicated, secure private networks.

- Data grids: Can be considered a subset of partner grids, but focus on sharing specific, often high volume data files along with infrastructure resources. For example, the National Digital Mammography Archive will centrally store, administer, and distribute mammogram files and patient records from the medical school hospitals at the Universities of Pennsylvania, North Carolina, Chicago, and Toronto. Because of patient privacy concerns and the sheer volume and size of the files involved (an average hospital collects 7+ terabytes of CT, MRI, and other electronic images annually) a data grid provides more muscular and flexible solutions for the Archive than would a simple private network.

Grid Type	**Environment**	**Who Uses Them?**	**Vendors**
Local Cluster	Single department or specific application	Any type of organization or enterprise	Sun
Enterprise Campus Intra-Grids	Multiple departments or workgroups	Any type of organization or enterprise	IBM, HP, Intel, Platform Computing, Sun
Partner	Multiple organizations	Enterprise and manufacturing partner; enterprise and supplier	Platform Computing, Sun
Data	Multiple organizations	Database providers or repositories	IBM
Research/HPC	Multiple organizations	Research organizations	IBM, HP, Intel, Microsoft, Sun, Fujitsu
Commercial Technical	Multiple organizations	Commercial and technical development organizations	IBM, Intel
Utility Service	Multiple organizations with different types of architectures	All types of organizations and enterprises	IBM, HP, Microsoft

- Research/HPC grids: Might also be called early-adopter grids, since they include many experimental and ambitious early grid deployments. Research grids are usually designed to enable the sharing of HPC or supercomputing resources between scientists and staff at cooperating national and international research institutions. Because of their focus on HPC and other supercomputing processes, research grids take advantage of dedicated, high security, high throughput networking infrastructures.
- Commercial/technical grids: Share many of the same characteristics as

research/HPC grids, but they tend to be deployed to support applications in ROI-intensive commercial/industrial environments such as life sciences and pharmaceutical product development. For example, a technical grid could be useful for providing access to high-end virtual modeling, data mining, or bioinformatics applications and hardware to scientists or work groups at widely geographically distributed commercial research facilities. Like research/HPC grids, commercial/technical grids often take advantage of high security, high throughput networking infrastructures.

- Utility/ service grids: Represent a future dreamed of by many proponents of Web services and what Sageza refers to as Service Computing: centrally controlled and provisioned heterogeneous computing infrastructures that can provide computing power, services, and applications on demand via a model similar to the way electricity is delivered to homes and businesses. If this concept is to succeed, it will depend on computational grids as the foundation infrastructure for delivering computing services in much the same way an electrical utility utilizes switching stations and power lines. Implementing utility grids will remain a dream until robust, secure, industry standard protocols make the public Internet an environment that can support and secure such services.

How Grid Vendors Approach the Market

Given the near-unanimous support computational grids boast across the high tech community, one might assume that vendors are following somewhat similar paths in their development of grid solutions. But such an assumption would be wrong, since vendors' grid development work tends to complement or simply parallel their other product research and development efforts. So what are vendors doing to make grid a reality? In November 2001, SGI announced that working with CANARIE (Canada's advanced Internet development organization) and McGill University, the company had successfully completed experiments that allowed workers to interactively run graphics applications on a SGI Onyx visualization system from 100 to 1,900 miles away. In April 2002, Fujitsu announced plans to introduce the Grid Solution for the Sciences, which could be used to deploy Fujitsu UNIX and Linux servers and supercomputers in computational grids to support scientific research. Following are short summaries of the grid activities of some enterprise IT vendors.

HP

Though HP and Compaq are now one company, the two followed separate paths in their grid development efforts. Compaq's extensive background in clustering technologies makes it an obvious grid player, and the company worked closely with the GGF and was a primary contributor to the Globus Toolkit. Additionally, Compaq worked closely with Platform Computing in pushing forward grid industry standards and developing specific commercial grid solutions. HP has taken a somewhat different approach, pursuing grid research and focusing its efforts on university and lab-related grid projects. HP is a platinum sponsor of the GGF, plans to port V2 of the Globus Toolkit to HP-UX, and is working with Globus on the OGSA-enabled Toolkit V3. Additionally, HP is promoting its Utility Data Center (UDC) product as an extension of and enterprise management solution for grid computing environments and related Web services deployments. It is not yet clear how these efforts will be integrated and their value proposition articulated in the newly combined company.

Intel

At one level, Intel is pursuing a conventional course in grid development. The company is a sponsor of the GGF, and is supporting or providing solutions for a number of grid and HPC computing projects around the world including the

Singapore Biomedical Grid and CERN's OpenLab project. Intel is also deeply involved in a number of P2P-model grid efforts including the Intel Philanthropic Peer-to-Peer Program, which uses technologies similar to the SETI Project to assist various medical research efforts. Intel's partnerships with IT vendor grid enthusiasts such as IBM and HP places the company in what could be a pivotal role in assisting and influencing the future direction of grid. Additionally, Intel's development of a wide range of products with grid implications, from embedded systems chips to HPC processors, along with its unique position as an open source solution enabler suggests the company will likely maintain a high profile in grid design and deployment.

IBM

That IBM has taken a leading role as a grid supporter and enabler is hardly surprising, given the company's prominence in the research labs and HPC facilities where much of the pioneering grid-related work is being done and where many grid projects are either operating or under construction. But the company has also been especially proactive in establishing leadership in grid development. IBM is a Platinum sponsor of the GGF and is the only IT hardware vendor listed as a Globus Project Collaborator. Additionally, the company is deeply involved in OGSA development; IBM Fellow Jeffery Nick was one of four authors credited in *The Physiology of the Grid*, an influential research paper published in February 2002 that expounded on the issue of leveraging computational grids and Web services technologies into broader solutions. Overall, grid technologies are central to IBM's visionary utility computing and autonomic computing initiatives. From a more practical standpoint, the company plans to grid-enable its hardware lines, will provide middleware and services for clients' grid projects, and supports ISVs such as Platform Computing and Avaki that develop commercial grid solutions.

Microsoft

If any major vendor's involvement in grid computing qualifies as being in "stealth" mode, it is likely Microsoft's. The company declined invitations to be interviewed for this report, and a search of the Microsoft Web site produced few relevant responses. At the same time, Microsoft is one of two IT vendors (along with IBM) listed as Collaborators in the Globus Project, and has invested $1 million in the Project, reportedly to ensure that the Globus Toolkit works with Windows XP and the .NET software infrastructure. The company is also a Platinum sponsor of the GGF. From a purely practical standpoint Microsoft's interest in grid makes perfect sense, given the company's deepening involvement in Web services delivery via its .NET initiative, its long term interest in and push toward high end computing and datacenter products, and its close working relationship with Intel, HP, IBM and other grid players. At the same time, Microsoft's general animus toward Linux and open source issues lends a healthy touch of irony to its warm support of the Globus Project.

Platform Computing

While most of the recent media ink and hoopla around grid computing has been spent on large scale projects like the TerraGrid and iVDGL, less notice has been paid to Canada-based Platform Computing, which has created successful enterprise/corporate grids for scores of companies including Texas Instruments, General Motors, and Pharmacia. Platform was launched in 1992 as a distributed computing ISV that developed proprietary solutions for performance, resource, and workload management, and the company's first grid solution was introduced in 1996. Platform is a sponsor of the GGF and works closely with the Globus Project. In fact, the company introduced Platform Globus, which Platform describes as the "first commercially supported distribution of the Globus Toolkit," in February 2002. In June 2002, Platform announced enhancements of its grid products that the company claims essentially doubles their performance and scalability. In all, over 150 of Platform's 400 employees are involved in grid-related solutions and projects, and the company's strategic partner roster includes IBM, HP, SGI, Cisco, NEC, Fujitsu, and Sun, among many others.

Sun

Sun's official interest in computational grids accelerated in July 2000, when the company acquired German ISV Gridware and began freely distributing the company's resource management application as the Sun Grid Engine. The product appears complementary to the company's Sun Cluster solution set, and Sun claims that over 5000 grids have been enabled since Grid Engine's introduction. In general, Sun's approach to grid has been highly pragmatic; Sun insists that the Grid Engine freeware offers businesses an opportunity to take immediate advantage of grid solutions that can benefit day-to-day operations, and the majority of installations qualify as cluster or departmental grids with an average size of about forty CPUs. Sun also strongly suggests that customers utilize the company's support services for grid deployments. In June 2002, the company announced the Grid Engine Enterprise Edition 5.3, a new commercial grid solution that is priced between $20,000 and $80,000.

What Is the Short Term Prognosis for Grid?

To begin, we believe that a good deal of heat lies beneath the smoke swirling around grid computing, and with good reason. If there is a supportable truism regarding the current state of business computing, it is that more is not necessarily better. Enterprise computing environments appear to be increasing in both size and complexity on a near-constant basis. In a sense, grid can be regarded as a highly technical solution that offers the same promise of bringing practical order to business computing that collaborative Shaker families incorporated into their daily lives. That same Shaker practicality and its attendant increased productivity also resonate among enterprise technology users. At least part of the allure of grid solutions can be found in their promises of easing resource management and increasing the use (and therefore the business value) of technology investments. That notion of demonstrable cost/efficiency enhancement is likely enough to send most IT staff (along with CTOs and CFOs) into palpitations, and provides at least one of the reasons so many IT vendors have climbed aboard the grid wagon.

Given the fact that far more IT vendors appear to be on that wagon than off, we think it appropriate to consider just how likely it is that the vendors we have discussed will realize the myriad promises of computational grids. Of the larger vendors, we see IBM and Intel as probably the best positioned of the group to benefit from grid developments. IBM's deep involvement in many of the earliest grid deployments has given the company a technical and practical leg up on some other players, offering the company just the sort of advantage it has enjoyed so often in the past. IBM's methodical approach to grid development meshes with the company's larger strategy of providing essential business computing solutions, and also allows IBM to define itself as a key developer of future utility grid infrastructure solutions, which it regards as enablers of entirely new classes of behavior such as autonomic computing. Intel might be best described as a player that regularly wins by covering a lot of different bets. Given the company's wide and deep alliances with a host of vendors, Intel is likely to reap rewards any time an ally does something right (and conversely, to lose when an ally blows it). The question to posit is how large might such a risk be? We would answer: not very large considering the breadth and depth of Intel's myriad other investments and involvements. If grid fails, Intel will move on. If grid succeeds, we believe Intel will be a primary beneficiary.

In a sense, Microsoft's position is quite similar to Intel's; the company will benefit whenever its hardware vendor partners do. Additionally, grid technologies are especially well-suited to support Microsoft's data center products and greater .NET strategies. In all, grid could serve to catalyze and extend future Microsoft efforts, and we expect the company (like every other IT vendor) to maintain its involvement in and influence over grid initiatives, and to incorporate whichever solutions best support its own corporate and technical goals. Platform Computing is another company that we believe is likely to profit substantially whenever its IT allies' grid efforts win. At first glance, the company's proprietary solutions may appear somewhat at odds with its support of the open source Globus Toolkit that vendors are flocking to, but Platform's experience and history of building enterprise grids makes the company a natural partner for grid-happy vendors of every stripe. Additionally, even if the most optimistic three- to five-year calendars for open source grid solutions are met, Platform has a great deal of time to pick a soft spot and land on its feet.

HP is positioned well moving ahead, especially given the quality of the company's enterprise solution set, including UDC and the extent to which it should benefit from Compaq's grid-related resources such as the company's clustering technologies and its Zero Latency Enterprise (ZLE) solutions. However, HP is currently undergoing a massive reorganization that could initially put its grid efforts several months (or more) behind the competition. Additionally, we are uncertain at this juncture how HP's plans to sell its entire middleware division might affect the company's grid plans. Neither are we especially sanguine about Sun's grid efforts. Despite the company's overt optimism and sizable grid deployment claims, we think it unlikely that thousands of freeware downloads will evolve into meaningful numbers of commercial enterprise software purchases, at least in the short term. Additionally, given the small average size of Sun's Grid Engine deployments, we wonder how many of those grid "customers" would have found equal or greater satisfaction with simple cluster solutions.

What Does It All Mean?

So despite whatever bumps and dips will occur (and should be expected) along the way, most if not all current players are well positioned to both enjoy and take advantage of the wholly natural evolution of grid into the next, biggest, brightest techno-wonder on the face of our small blue planet. Right? Not so fast, because from where we sit those bumps look fairly big and those dips run pretty deep. The first problem that should be considered might be called the dynamic of unreasonable expectations. For the past six months and more, enthusiasts have raved about the grid computing model as a likely and worthy successor to the Internet, despite the fact that the rollout of widely accessible grids is at least three to five (or to our minds, more likely five to seven) years away. In other words, grid is currently at about the same evolutionary state that the Internet was in the late 1980s. The problem we see in overzealous hype is that it tends to undermine whatever practical (and often highly specialized, though effective) applications come along in the natural course of things. The result is that if and when grid ever does live up to its promise, its natural audience may already have given it up and drifted away. In this case, grid would hardly be the first high tech wunderkind to die a-borning.

Does that mean that grid computing is on a collision course with the trash heap of history? Not at all, but it does suggest that some grid vendors might do well to exercise a bit of caution in their promotional efforts. For the Shakers, bringing order to life provided the impetus and inspiration for remarkable spiritual and temporal achievements. The sect's beautiful furniture and handiworks are still highly prized by collectors and museums alike. A very few Shaker communities remain today as living museums where a bare handful of Believers still live, and where vacationers can learn about those dead and dying families whose curious, utilitarian habits inspired their myriad accomplishments but failed to prevent their ultimate extinction. While we believe grid computing offers enterprises serious, tangible technical and practical benefits, its current influence is and will continue to be felt primarily in the data center, at least in its present form. The risk, of course, is that if grid never leaves the realm of the data center, if it never develops or inspires the sort of "killer" apps that rocketed the Internet into common usage, grid will likely be thought of and then remembered as yet another cool technology that never lived up to its potential, that was tripped up by its own hype on the way to the mainstream. If that comes to pass, grid computing, like the Shakers, could become a museum piece best remembered as an elegant, ultimately unsuccessful utopian experiment.