Grid Computing: An Overview

Prepared By: Nellai Subramaniam
Under the supervision of Dr. Blake Ives

1. Introduction

"Grid Computing is important to Shell because it offers the potential to create a truly unlimited resource, with a uniform interface to a variety of services. This is a significant opportunity for Shell to engage its independent companies in closer cooperation".

J.N. Buur, Principal Research Physicist,
Shell International Exploration and Production B.V.

What is grid computing? Grid computing involves the sharing, selection, and aggregation of a wide array of resources that may include supercomputers, mainframes, storage systems, data sources, distributed applications and management systems. If one considers the internet as a network of communication, grid computing can be considered a network of computation [20].

Grid computing is an extension of scalable and distributed computing concepts. Its objective is to harness the resources of unused CPU cycles of computers to perform computational tasks at a faster rate. The idea of using the unused CPU cycles originated in the early 1970s with the linking of computers by networks. In 1973, scientists at the XEROX Palo Alto Research center developed a worm program that replicated itself in the memory of each of the 100 Ethernet connected computers. The worms used the idle resources to perform shared computations for rendering realistic computer graphics [3].

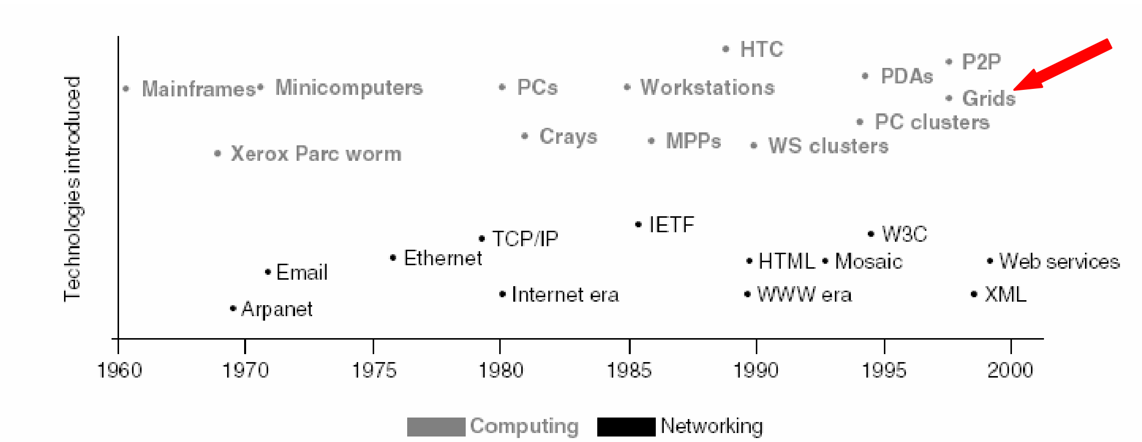

Figure 1

(Source: Buyya et. al [3])

The popularity of the Internet and the technological advances in the last decade has taken the distributed computing concept to a new level. Computers are now hooked together in clusters for high-performance computing. These clusters have resulted in the emergence of peer-to-peer (P2P) networks and computational grids. Figure 1 provides a timeline of the advances in networking and computation leading to the emergence of peer-to-peer networks and computational grids.

Though the terms "peer-to-peer computing" and "grid computing" have been used interchangeably by some, Buyya and Chetty [3] make an interesting distinction between them: P2P computing is concerned with sharing low end systems such as PCs connected to the Internet for amassing computing power (e.g., the SETI@home project[1]), or share contents as done in Napster and Gnutella. On the other hand grid computing is concerned with aggregating the distributed clusters of computers using well defined protocols and standards.

While grid computing was initially used for scientific and academic research, the past two years have seen a significant increase in the use of grid computing by commercial institutions.

Shell uses grid-enabled infrastructure for applications that involve interpretation of seismic data [14]. The grid is expected to cut processing time of seismic data while improving the quality of the output. Further, because of the open standards used, the grid is expected to provide easy integration of existing software.

BankOne is using grid computing to boost performance of analytics for its Chicago based trading floor. Instead of using a super computer to process its massive risk analytics, BankOne uses a lot of small processors and the computing responsibilities have been distributed across them. By doing so, Bank One expects to make significant cost savings on hardware and also utilize the unused processing cycles of the distributed resources [25].

JP Morgan Chase Investment Bank (JPMCIB) is offering its business units online access to raw computing power via an enterprise technology infrastructure, which pools the financial firm's flexible processing resources for compute-intensive applications. The project code-named "Compute Backbone" uses grid computing technology to improve process service levels and to cut costs by charging the business units for their processing, with peak and off-peak pricing [6]. Steven Neiman, head of JPMCIB's high performance computing initiative, described this scheme:

"… we want to make high-performance, distributed computing available in the same way that people use network services. When accessing an application or a device on a network, end-users rarely know or care about how the networking technology works. The bank is attempting to achieve the same thing with its processing resources"

For others in the financial industry, grid computing enables the migration to cheaper computing solutions, such as running Linux on Intel. Charles Schwab converted an existing wealth-management application and grid enabled it using the Globus Toolkit with IBM eServer running RedHat Linux [25]. By running Linux, a free open-source

[1]SETI@home is a scientific experiment that uses Internet-connected computers in the Search for Extraterrestrial Intelligence (SETI). Complete information about SETI@home can be found at http://setiathome.ssl.berkeley.edu/

operating system, Charles Schwab expects to save significant capital expenditure on hardware, such as expensive UNIX servers. The project has helped Charles Schwab reduce the processing time of certain transactions from four minutes to fifteen seconds [13].

Novartis A.G., the Switzerland based drug and pharmaceutical giant uses grid computing technology to speed up its drug research in a cost-effective way. The grid links about 2700 PCs delivering 5 teraflops of computing power [21]. Manuel Peitsch, head of Novatis AG's department of Informatics and Knowledge Management explains,

"If you look at the desktop PCs in a typical corporation, probably 90 percent of computing cycles are unused. Just by capturing unused cycles on the PCs we have already got, we have created a 5 teraflops supercomputer. We have avoided the expenses of buying an HPC system, building another computer center, and taking on the people to support it. We invested roughly $400,000 in grid software licensing and figure we have saved at least $2M based on the 2700 seats we have currently. We expect to realize more savings of this nature in the future as our grid expands."

Other organizations adopting grid services include the Mayo clinic, which is developing a system for linking its medical database with vast external public and private data sources in order to develop more-effective patient treatments [4]. Pratt & Whitney, a Connecticut based flight technology company, has been using grid technology to model jet engines and gas turbines [22].

The rest of this paper discusses the need for grid computing, grid concepts and components, grid architecture and standards, a case study of adoption of grid computing, and future trends in grid computing.

2. Why should one care about grid computing?

In most organizations there is an enormous amount of computational processing power which is unused. There is also usually a huge amount of unused storage capacity residing on these machines. Grid computing provides a framework for exploiting these underutilized resources.

Another important grid computing contribution is to enable and simplify collaboration among different organizations. This collaboration does not concern file sharing alone. It concerns direct access to computers, software and other resources. Heterogeneous systems distributed globally can work together to create an image of a virtual computing system offering a variety of virtual resources (Figure 2).

The users of the grid can be divided and grouped together based on their functional areas to form virtual organizations (VOs). These virtual organizations may have their own sets of policies and rules. According to Foster, Kesselman[2] and Tuecke in their paper, "The Anatomy of the Grid" [7],

[2]Computer scientists hold the contributions of Foster and Kesselman to grid computing at the same high esteem as the contribution of Tim Berners-Lee to the World Wide Web [17].

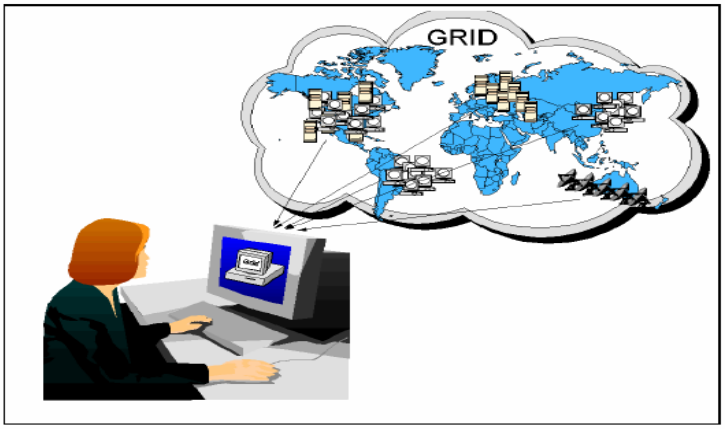

Figure 2 : Simple view presented to the user after virtualization of the distributed resources

(Source: Berstis[2])

"VOs vary tremendously in their purpose, scope, size, duration, structure, community, and sociology. Nevertheless, careful study of underlying technology requirements leads us to identify a broad set of common concerns and requirements. In particular we see a need for highly flexible sharing relationships, ranging from client-server to peer-to-peer; for sophisticated and precise levels of control over how shared resources are used, including fine-grained and multi-stakeholder access control, delegation, and application of local and global policies, for sharing of varied resources, ranging from programs, files, and data to computers, sensors, and networks; and for diverse usage modes, ranging from single user to multi-user and from performance sensitive to cost-sensitive and hence embracing issues of quality of service, scheduling, co-allocation, and accounting."

They further state that current distributed computing technologies do not address the concerns discussed above. According to them:

- *CORBA and Enterprise Java are concerned with resource sharing within a single organization but not across organizations;*
- *Open Group's Distributed Computing Environment (DCE) supports resource sharing across sites but are too cumbersome and inflexible; and*
- *Storage Service Providers and Application Service Providers allow organizations to outsource storage and computing requirements, but only in constrained ways.*

Hence, comprehensive technologies with flexible frameworks and standards are needed to address the concerns discussed above. Years of research on grid computing have now resulted in a set of protocols and rules that address the concerns discussed above. These technologies include security algorithms that support management of resources across multiple locations, protocols for querying information across multiple locations and protocols for configuring resources and machines that form part of the grid.

The last few years has seen a renewed interest in grid projects around the world. Vendors such as IBM and Oracle are positioning grid computing as an important cog in their computing strategies. Grid computing is driving a new evolution in industries such as the bio-medical field, financial modeling, oil exploration, motion picture animation and many others.

3. Grid Concepts and Components (adapted from Berstis [2])

A grid consists of many resources. These resources are sometimes addressed by different names such as "nodes", "resources", "members", "donors", "clients", "hosts", "engines". "Computational Grid", "Scavenging Grid" and "Data Grid" are other terms used in the grid computing world. In a computational grid, resources, usually high performance servers, are set aside for computing power. In a scavenging grid, resources, usually desktop machines, are scavenged for unused CPU cycles and other resources. In a data grid, the main focus is to provide access to data. Users are unaware of the location of data and are only concerned with access to data [15]. This section discusses the meaning and application of concepts such as "Computation", "Storage", "Communications", "Jobs", "Scheduling", etc., to the realm of grid computing. The various software components that are used in grid computing are also discussed.

Grid Concepts

- Computation: The computing cycles provided by the processors on the grid.

- Storage: The data storage present in the grid. The storage may refer to the primary storage such as memory attached to the processor or it may be secondary storage such as hard disk drives and other permanent storage media such as networked file systems. Data may also be stored on storage devices that span several machines in a grid. A unifying file system may be used to provide a single uniform name space for this storage. By having a unifying file system, the user can reference the data in the grid without regard to its exact location. Grids may also have an independent scheduler that can select the appropriate storage devices based on usage patterns and storage needs.

- Communications: This refers to communication within the grid and communication external to the grid. Internal communication within the grid is essential, when one considers the fact that many jobs in the grid have to access data that reside on multiple machines. The criticality of the bandwidth available for this internal communication depends on the amount of data that has to be transferred between different machines on the grid. External communication refers to access to the Internet. If an organization's grid has to communicate with other grids, then external communication becomes critical. Communication between grids that are spatially and geographically distributed usually takes place through the Internet. Redundant communication paths are sometimes needed in a grid to provide a fault tolerant system to guard against network failure or excessive data transfer.

- Software and licenses: Grid computing provides the opportunity of installing specific software on only a few machines. If a job requires the software to be used, the job can be sent to the machine in which the software is installed. By doing this, an organization can save significant expenses on licensing fees.

- Special equipment, capacities, architectures, and policies: A grid may contain platforms with different architectures, operating systems, devices, capacities, and equipment. All these attributes must be considered before assigning jobs to the grid. Some software may run only on Windows machines while others may run only on Linux machines. Further, some machines may be used only for number crunching financial research because of the hardware specifications of those machines. Some machines may be used only for search purposes. By connecting each of these machines to the Internet using separate high speed connections, instead of sharing a single connection, the bandwidth is increased and would facilitate faster search results. Hence the grid administrator should ensure that policies and rules are available to schedule jobs based on the capacities and purpose of the resources.

- Jobs and Applications: The term "application" is used to refer to the highest level of a piece of work on the grid. An application may be broken down into jobs and sub-jobs. Terms such as transaction, work unit, and submission are commonly used as equivalents for job. The jobs in an application may be programmed to execute in parallel on different machines in the grid. Some jobs may have specific dependencies and hence may execute in a sequential order with the output produced by one job being used as the input for a certain job. The grid should have provisions to collect and appropriately assemble the results of the jobs to produce the required output for the application (Figure 3).

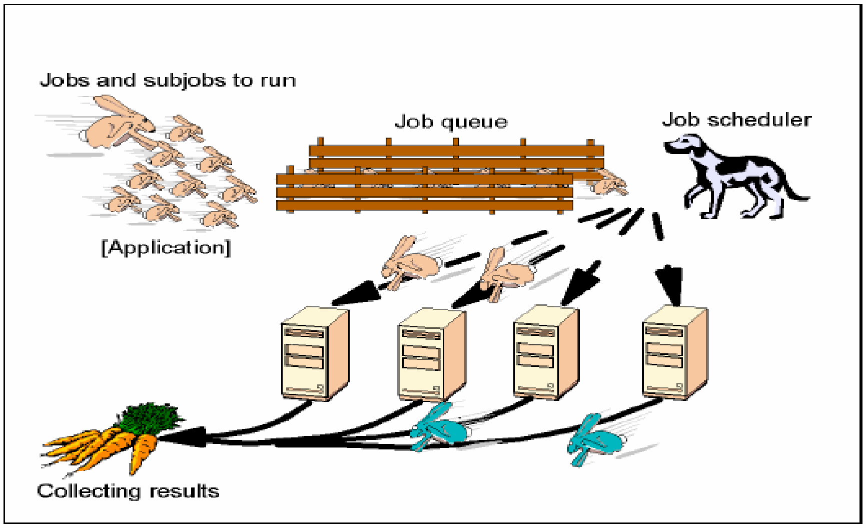

Figure 3 : Jobs and Applications

(Source: Berstis[2])

- Scheduling, Scavenging and Reservation: A grid may contain a "job scheduler" that schedules a job based on the availability and appropriateness of the resources. The term "resource broker" may also be used as an equivalent to "job scheduler". A "scavenging grid" system is one where a machine reports its idle time to the grid management node. The management node would assign to this idle machine the next job that meets the constraints of the machine's resources. But the downside is, if this machine becomes busy with a local non-grid job, then the grid job gets delayed or suspended. This can be overcome by having "dedicated" machines on the grid that are not preempted by local non-grid work. In some cases, grid resources may be "reserved" in advance for a designated set of jobs. "Reservation" can aid in meeting deadlines and also guarantee quality of service. Some of the 'reserved" resources may be "scavenged" for their idle cycles to run low-priority jobs during a reservation period. For maximum efficiency, a combination of "scheduling", "scavenging" and "reservation" may be used. The type of software used for this purpose is discussed later in this section.

- Intragrid and Intergrid: The simplest grid may consist of a few machines having the same hardware architecture and operating system and connected on a local network. Choosing application software for these machines is simple since they have the same type of hardware and operating system. In most cases, a grid would include machines from other departments and also the machines may be heterogeneous (dissimilar hardware architecture and operating systems). An "Intragrid" usually refers to machines from different departments within a single organization forming the grid (Figure 4). "Intergrid" on the other hand refers to a grid that crosses organization boundaries (Figure 5).

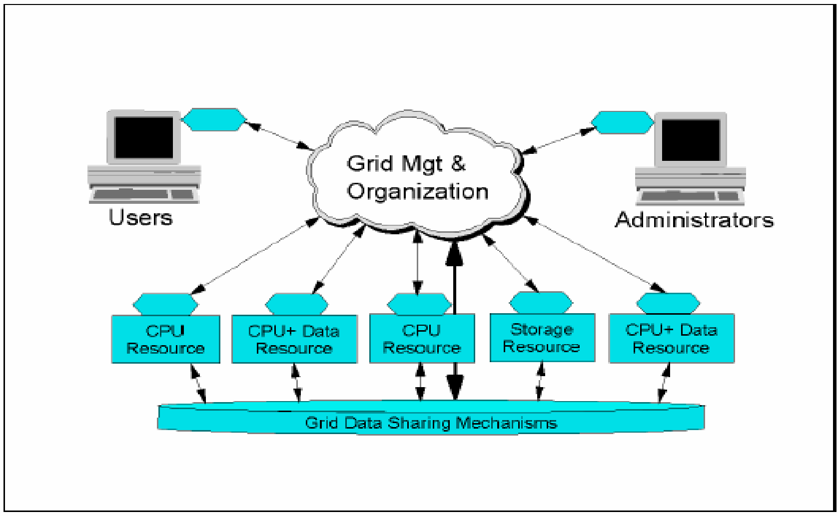

Figure 4: Intragrid

(Source: Berstis[2])

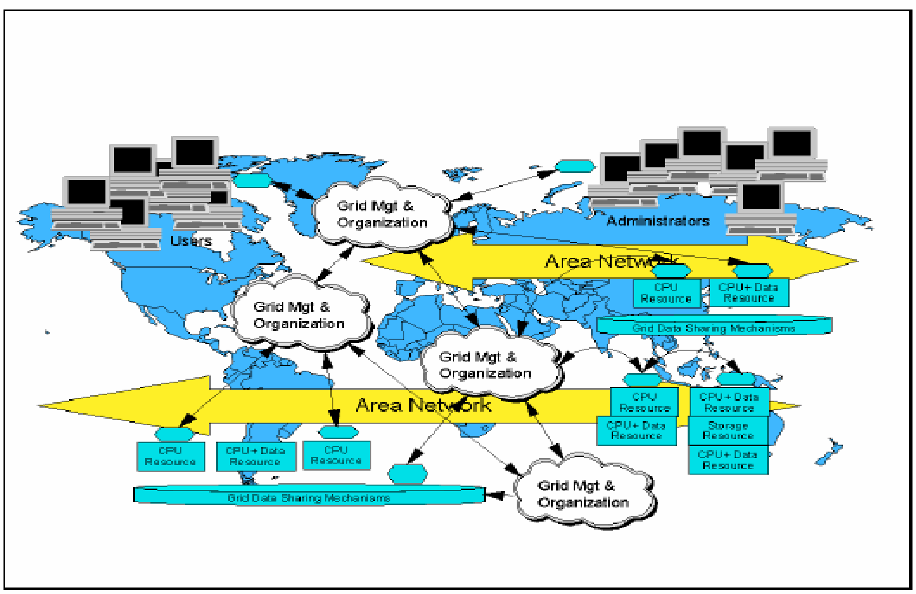

Figure 5: Intergrid

(Source: Berstis[2])

Grid Software Components

- Management Components: Every grid system must have some kind of management software. This software is used to keep track of resources available to the grid and also to keep track of the different users on the grid. Other jobs of this software include measurement of: capacities of the resources, utilization rate of the resources, traffic congestion and bottlenecks, overall usage patterns of resources. Such software also generates statistics, provides for recovery from grid failures, finds alternate ways to get the jobs done.

- Donor and Submission Software: A machine (donor) that contributes resources to the grid must be enrolled as a member of the grid and must have software installed that allows the grid to better manage its resources. The software installed on the donor machine can help to monitor its resources and also to send the resource information to the grid management software. For example, in a "scavenging" grid this information is used to inform the grid management software of the availability of idle time in a machine. Also, this software is used to accept jobs from the grid and execute it on the donor machine. Any member machine in a grid can submit jobs to the grid. But, in some grid systems, specialized software called "submission software" may be installed on individual machines to allow the machines to submit jobs to the grid.

- Job Scheduling Software: Job scheduling software is used to assign a specific job to a machine on the grid. This assignment is usually based on some scheduling algorithm. For example, a priority algorithm with priority queues may be used to schedule jobs. As grid resources become available, the jobs from higher priority queues are scheduled to be executed first. Other information such as grid traffic and network outages is taken into account before scheduling jobs.

- Communications Software: A grid system may contain communication software to help jobs communicate with each other. An application that consists of many sub-jobs may require some of these sub-jobs to communicate information with each other. The sub-jobs should be able to locate other sub-jobs and transfer information to them. For example, communication software that follows the open standard Message Passing Interface (MPI) can be used for this kind of communication.

4. Architecture and Standards

In order to aggregate distributed and heterogeneous high end machines, standards are needed. The standardization of network communication between heterogeneous systems led to the explosion of the Internet [2]. Similarly the emerging standardization for sharing resources will lead to the explosion of grid computing.

Architecture

The components that are required to form a grid can be divided into four layers based on their role in the grid system [9]. They are:

- Fabric*:* At the lowest level this consists of resources users share and access including computers, storage systems, catalogs, networks, sensors, processors.
- Connectivity and Resource Layers: Contains the core communication and authentication protocols and also protocols that enable secure initiation, monitoring and control of resource-sharing operations.
- Collective Layer: Contains protocols, services, and APIs that implement interactions across collections of resources.
- User Applications: Includes user applications which call on the components in other layers to complete their tasks.

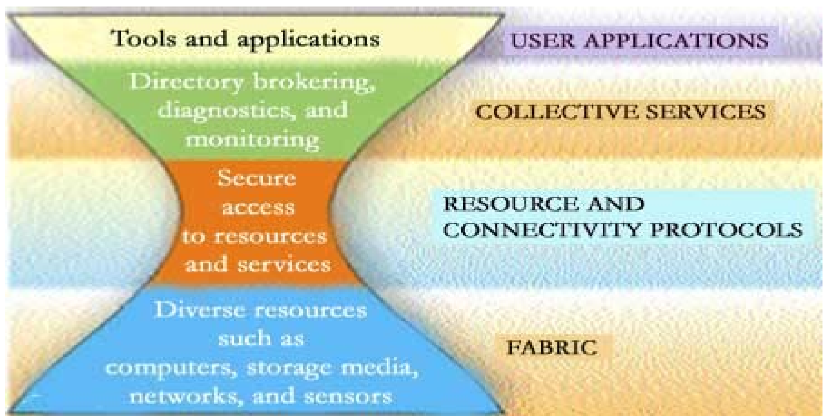

Figure 6: Architecture

(Source: Foster [9])

Standards

- Globus Toolkit: This is an open source toolkit released by the Globus Alliance[3], whose members include researchers at Argonne National Laboratory, the University of Southern California's Information Sciences Institute, the University of Chicago, the University of Edinburgh, and the Swedish Center for Parallel Computers. Some of the prominent corporate sponsors of the alliance include IBM, Microsoft and Cisco. The toolkit includes software for security, information infrastructure, resource management, data management, communication, fault detection and portability.

- Open Grid Services Architecture [OGSA]: This is an evolving standard with significant industry support [15]. OGSA defines what grid services are, what they should be capable of, what types of technologies they should be based on, but doesn't give a technical and detailed specification for the implementation [24].

- Open Grid Services Infrastructure [OGSI]: OGSI gives a formal and technical specification of the standards defined by OGSA [24].

The grid services defined by OGSA and OGSI are extensions of web services[4]. The Globus Toolkit Version 3(GT3) is implemented based on the specification of OGSI standards and hence OGSA standards. Figure 7 presents the interrelationship between OGSA, OGSI, Web Services and the Globus Toolkit implementation. The core services supported by the Globus Toolkit include [1, 10]:

- Grid Security Infrastructure (GSI) – Authentication and related security services.
- Grid FTP - Grid based file transfer protocol built on standard FTP protocol.
- Globus Resource Allocation Manager (GRAM) – Resource Allocation and Process Management.
- Metacomputing Directory Service (MDS) – Provides information about the available resources within the grid and their status.
- Global Access to Secondary Storage(GASS) – Remote access to data via sequential and parallel interfaces

With evolving standards, IT vendors are competing with each other to provide grid services based on these standards. Appendix A contains a list of leading grid solution providers. The relative market positioning of major OEMs-Original equipment manufactures is listed in Table1.

---

[3]More information about the Globus Alliance and the Globus Toolkit can be found at http://www.globus.org/

[4]The paper "The Physiology of the Grid" (Foster et. al [8]) which is included in the reference provides a description of how grid computing can be implemented in a web services environment.

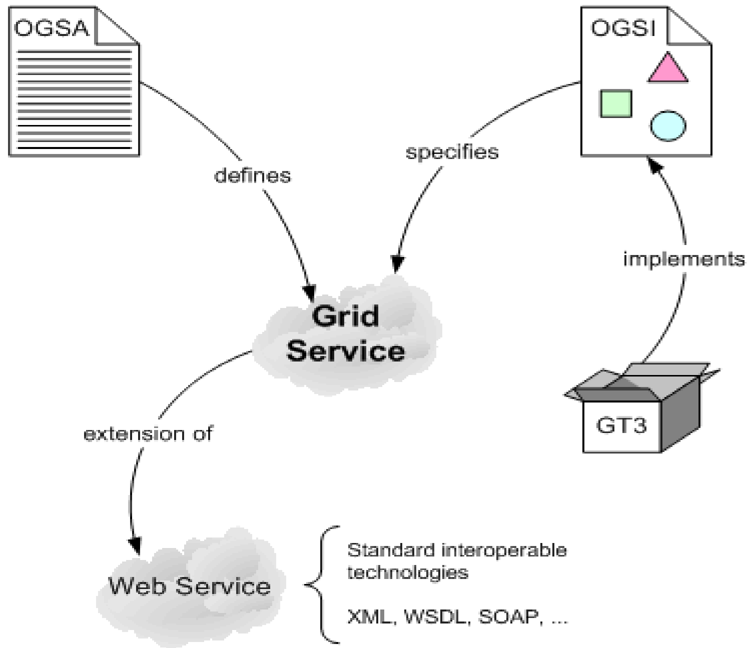

Figure 7 : OGSA, OGSI, Web Services, Globus Toolkit

(Source: Sotomayor [24])

5. Case Study

Hewitt Associates (adapted from Information Week [19])

Hewitt Associates is a global outsourcing and consulting firm delivering a complete range of human capital management that includes HR and Benefits Outsourcing, HR Strategy and Technology, Health Care, Organizational Change, Retirement and Financial Management, and Talent and Reward Strategies. Their clients include more than half of the Fortune 500 companies and several Fortune Global 500 companies. They handle more than 53 million HR-related customer interactions a year, and are characterized by the following:

Revenue (last 4 quarters) $ 1.98 billion
Business Human Resource Outsourcing (65% of revenue)
Consulting (35 % of revenue)
Income (last 4 quarters) $79 million
Employees 14, 600
IT Employees 2,700

Challenge: Hewitt processed HR benefits for more than 16 million people employed in companies that included Sony Electronics and Johnson and Johnson. Business process

Bauer College of Business Administration
University of Houston, Houston, Texas, 77204-6283
Phone: 713-743-4691; Fax: 713-743-4693; e-mail: isrc@uh.edu

outsourcing expertise was one of Hewitt's main competitive advantages. More than 2000 of Hewitt's IT staff members work on outsourcing.

Hewitt had a problem that was centered on figuring the pension benefits of a client company's employees based on the client company's specific plan. Traditionally, Hewitt used to use a mainframe to do this calculation. But about 1% of its transactions involved pension calculations of customers whose pension rules had changed because of mergers and acquisitions. These calculations consumed about 25% of the mainframe processing power. Hewitt was wasting its mainframe capacity on this 1% of interactions.

In the past HR managers at the client companies were content with occasional reports of pension calculations. But with the development in internet technologies, HR managers either wanted to do it themselves or give their employees self-serve access. According to Jim McGhee, technology development leader at Hewitt,

"A lot of clients are starting to say, 'Give us Web-services access to calculations'. I don't want to say, 'You can have it for $4 per calculation' ".

<u>Grid Computing Solution</u>: Hewitt decided to implement a grid computing solution to solve its pension calculation problem. The grid included 10 IBM server blades[5] with Red Hat Linux operating system, a software program from DataSynapse[6] to distribute the task across the grid, and Java based middleware to connect the grid to the IBM mainframe. The results of the grid implementation demonstrated that the pension calculations were accurate. The grid based pension calculations were 0.2 seconds slower for the easy ones and 10 to 15 seconds faster for the tough ones. The successful functioning of the dedicated grid meant a 90% savings over running the pension calculations on the mainframe. The standardized Java middleware and the SOAP messaging protocol opened up other opportunities for harnessing the grid. According to Tony Bishop, chief business architect of DataSynapse,

"*They now have a way to migrate the right type of request to the right type of computing source. Since they have built it through a Web service, open program, they have allowed this mainframe to talk to the distributed world".*

Hewitt realizes that the greatest payoff won't come from cutting costs of a particular type of pension calculations. The next step for Hewitt would be to build its business competency around grid technology and then to migrate some of the other functions and tasks from its mainframe based systems to the grid. This is best summarized by Hilgenberg, Hewitt's application strategist,

"If we build a competency around grid, we will start to see opportunities we don't now have."

[5]A server blade is an independent server, containing one or more processors with associated memory, disk storage and network controllers and running its own operating system and applications. A blade server is an optimized chassis with many server blades. More information can be found at ftp://ftp.pc.ibm.com/pub/pccbbs/bp_server/why_blade_servers.pdf

[6]DataSynapse is a leading provider of grid computing software, http://www.datasynapse.com/

7. Future Trends

According to Gartner, by 2006, 5 percent of commercial organizations that routinely use supercomputing cycles will turn to cheaper grid computing for non-security sensitive resources, a number they expect to see increase to 25% by 2011[11]. An interesting recent research endeavor at Virginia Tech University seems to validate the above prediction. Researchers at the Virginia Tech University have assembled a supercomputer by linking together 1100 Apple Macintosh computers at a cost of about $5 million, a fraction of the cost of a traditional supercomputer ($100 to $200 million price tag). This super computer was able to compute at 7.41 trillion operations a second, a speed that is inferior only to three other super computers in the world 18].

According to a ground breaking study by The Insight Research Corporation [26], financial applications and professional business services will account for the highest grid expenditures by 2008. Total worldwide grid spending is expected to grow from $250 million in 2003 to approximately $4.9 billion in 2008, as shown in Figure 8.

Though grid computing has opened up a number of opportunities for many organizations to better manage their resources, it is not a panacea for all computational problems nor appropriate for every organization. The configuration of a grid can influence the benefits that accrue to an organization. Most current business applications are not written for parallel processing. Parallelizing these applications can involve a major investment. Standards for ensuring security in grids, monitoring user access and charging for user access are still evolving. Hence it is imperative for organizations to understand what the grid can and cannot do as well as the associated risks.

All the major IT providers are involved in promoting grid computing in one form or another. With more and more companies embracing grid technology, it is slowly moving from research labs to commercially viable solutions. Acknowledging this scenario, Larry Ellison, CEO of Oracle, who has been championing his company's grid enabled database Oracle 10g states that[16, 23],

"It's not going to happen overnight, but there will be an inexorable move [to grid computing]. The economics are compelling; the reliability is compelling."

Bauer College of Business Administration
University of Houston, Houston, Texas, 77204-6283
Phone: 713-743-4691; Fax: 713-743-4693; e-mail: isrc@uh.edu
ISRC

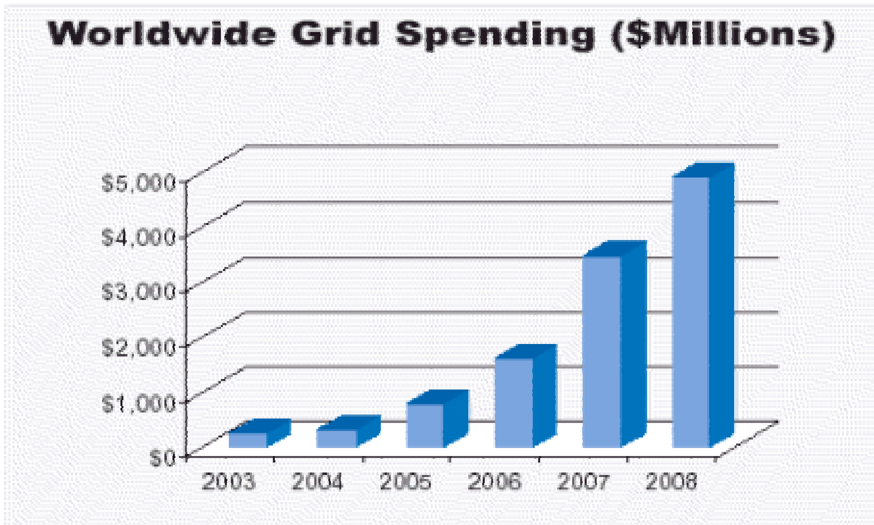

Figure 8

(Source: The Insight Research Corporation Report [26])

References

1. Baker, M., Buyya, R., Laforenza, D., "Grids and Grid Technologies for Wide-area Distributed Computing", 2002, http://www.cs.mu.oz.au/~raj/papers/gridtech.pdf

2. Berstis, Viktors, "Fundamentals of Grid Computing", Redbooks Paper, IBM Corporation, 2002.

3. Buyya, R., Chetty, M., "Weaving Computational Grids: How Analogous Are They with Electrical Grids?", IEEE Technical Reports, 2001.

4. Chui, Willy, "Grid Computing: Fulfilling the Promise of the Internet" , GridComputingPlanet.Com, July 14, 2003

5. Clabby Analytics, Independent Technology Research and Analysis, "Competitive Positioning: IBM in Grid Computing", September 2003

6. Davidson, Clive, "JP Morgan unveils Project Compute Backbone", Financial Technology Intelligence, October, 2002.

Bauer College of Business Administration
University of Houston, Houston, Texas, 77204-6283
Phone: 713-743-4691; Fax: 713-743-4693; e-mail: isrc@uh.edu

7. Foster, I., Kesselman ,C., Tuecke,S., "The Anatomy of the Grid, Enabling Scalable Virtual Organizations", http://www.globus.org/research/papers/anatomy.pdf

8. Foster, I., Kesselman, C., Nick, J., Tuecke, S., "The Physiology of the Grid: An Open Grid Services Architecture for Distributed Systems", Draft Paper, http://www.globus.org/research/papers/ogsa.pdf

9. Foster, I., "The Grid, A new infrastructure for 21st Century Science", Physics Today, February 2002, http://www.aip.org/pt/vol-55/iss-2/p42.html

10. Foster, I., Knesselman, C., "The Grid: BluePrint for a New Computing Infrastructure , Morgan Kaufmann Publishers, 1999

11. Gartner Report, "Emerging Core Computing Technologies", October 2001.

12. Hewitt Associates, http://was4.hewitt.com/hewitt/about/overview/index.htm

13. IBM Case Study – Charles Schwab, http://www-1.ibm.com/grid/pdf/schwab.pdf

14. IBM Case Study - Royal Dutch/Shell, http://www-1.ibm.com/grid/pdf/royaldutchshell.pdf

15. Jacob, Bart, "Grid Computing, What are the key components?" , June 2003, http://www-106.ibm.com/developerworks/grid/library/gr-overview/

16. Krill, Paul, "Ellison champions grid", InfoWorld, September 09, 2003, http://www.infoworld.com/article/03/09/09/HNellison_1.html?s=feature

17. Lohr, Steve, "Teaching Computers to Work in Unison", NYTimes, July 2003.

18. Markoff, John, "Low Cost Supercomputer Put Together From 1100 PC's", NYTimes, October 2003,

19. Murphy, Chris, "The Guts to Say Go: How Hewitt's grid-computing project got the green light", Information Week, September 2003, http://www.informationweek.com/story/showArticle.jhtml?articleID=15200497

20. Myer, Thomas, "Grid Computing: Conceptual Flyover for Developers", May 2003, http://www-106.ibm.com/developerworks/library/gr-fly.html

Bauer College of Business Administration
University of Houston, Houston, Texas, 77204-6283
Phone: 713-743-4691; Fax: 713-743-4693; e-mail: isrc@uh.edu

21. Novartis, Intel Business Center Case Study, http://www.intel.com/ebusiness/pdf/cs/cs_Novartis.pdf

22. Robb, Drew, "Plugging Into Computing Power Grids" , ComputerWorld, April 22, 2003

23. Songini, Marc, "Oracle Grid Computing Plan Faces Obstacles", ComputerWorld, September 15, 2003, http://www.computerworld.com/databasetopics/data/story/0,10801,84878,00.html

24. Sotomayor, Casa, "The Globus Toolkit – Programmers Tutorial", http://www.casa-sotomayor.net/gt3-tutorial/

25. Schmerken, Ivy, "Girding for Grid", Wall Street and Technology Online, March 12, 2003

26. The Insight Research Corporation Report, "Grid Computing: A Vertical Market Perspective 2003-2008"

Useful Websites

1. http://www.ggf.org/ - This is a grid community initiated forum of more than 5000 researchers whose main purpose is to promote the deployment and implementation of grid technologies and to document best practices, technical specifications and implementation guidelines.
2. http://www.gridcomputingplanet.com/ - An online news portal for grid computing.
3. http://www.gridcomputing.com/ - A great site maintained by Dr. Rajkumar Buyya, Dept. of Computer Science, University of Melbourne. This site contains a variety of grid computing links.
4. http://www-fp.mcs.anl.gov/~foster/ - This is Dr. Ian Foster's home page at the Argonne National Laboratory. Apart from links to Dr. Foster's papers, this site also contains links to other grid resources.

Appendix A

- Avaki
- CapCal
- Centrata
- DataSynapse
- DELL
- Distributed Science
- Elepar
- Entropia.com
- Grid Frastructure
- GridSystems
- Groove Networks
- HP
- IBM
- Intel
- Jivalti
- Mind Electric
- Mithral
- Mojo Nation
- NewsToYou.com
- NICE, Italy
- Noemix, Inc.
- Oracle
- Parabon
- Platform Computing
- Popular Power
- Powerllel
- ProcessTree
- Sun Gridware
- Ubero
- United Devices
- Veritas
- Xcomp

Bauer College of Business Administration
University of Houston, Houston, Texas, 77204-6283
Phone: 713-743-4691; Fax: 713-743-4693; e-mail: isrc@uh.edu
ISRC

Table 1

Dell	Dell's grid strategy is very clear: Dell is seeking to supplant Unix high performance computing solutions with Dell Intel-based Linux high performance computing clusters (HPCC). In short, Dell's grid computing strategy is based on targeting higher cost HP-UX, Solaris, and AIX platforms ¯ providing low-cost, high performance Linux cluster alternatives that run on commodity Intel hardware.
HP	HP's grid strategy focuses on grid resource management, utility resource provisioning, and homogeneity. The company has "grid-enabled" its entire line of server products ¯ and has implemented a centralized grid management system called the "Utility Data Center" (a product that has met with very limited success). HP has also established partner relationships with the AVAKIs, Platforms, and Entropias of the grid world in order to help provide specialized solutions on HP systems. HP is also well represented on standards committees.
IBM	IBM focuses on homogeneity, standards, packaging, compute and data grids, and professional services delivery. IBM has grid-enabled its entire server line, has established partner relationships will all of the major grid solutions providers, has integrated its management and data integration products into its grid product offerings, and has the best equipped grid professional services delivery organization in the industry.
SGI	Silicon Graphics (SGI) focuses on visualization (three-dimensional computer modeling), high-performance computing, and the management of complex data. SGI's stated mission is to provide scalable computing, collaborative visualization, and complex data management solutions for technical and creative users. In a nutshell, SGI's grid strategy focuses on providing collaborative grid solutions that address visualization output. (SGI is known for its work producing highly optimized visualization hardware and software ¯ the company uses grids to share advanced 3D computer model output amongst communities of researchers, explorers, educators, designers, scientists, and commercial

Sun	Sun builds its own grid resource/policy management software (contained in Sun Grid Engine (SGE) and Sun Grid Engine Enterprise Edition (SGEEE)). When coupled with Sun's systems, application development environments, operating systems, storage subsystems, and systems and network management software Sun puts forward a very potent *cluster-level* grid offering. To move beyond grid clusters, Sun frequently partners with AVAKI (a supplier of grid software that includes support for large, heterogeneous environments as well as rich naming conventions and other important enterprise-level functionality). Sun also regularly participates in industry-driven grid standards organizations such as the Global Grid Forum (the GGF). The GGF is the primary standards setting organization for grid architecture. (Note: Sun was the first sponsor of the GGF).

Grid Strategy: Major Original Equipment Manufacturers (Source: Clabby Analytics [5])

Bauer College of Business Administration
University of Houston, Houston, Texas, 77204-6283
Phone: 713-743-4691; Fax: 713-743-4693; e-mail: isrc@uh.edu