GRID Computing: Related Issues and Business Applications

Sanoara Yasmin and Musharuf Hossain Mollah
School of Business, University of Information Technology and Science
Chittagong Campus, Bangladesh
Department of Marketing
Southeast University, Dhaka, Bangladesh

Abstract
Grid computing is a virtual computing system developed through the coordination and sharing of computing power, application, data, and storage facilities of different computers of geographically dispersed organizations. This paper is an effort to highlight the conceptual aspects, key characteristics, benefits and business applications of Grid computing. It focuses on different applications of Grid computing in Bangladesh perspective. Relevant peripheral issues of Grid computing like storage, security, integration and transparency of data, etc. are also discussed in the paper. Finally, the paper focuses on different strategies to implement Grid computing along with the future scope of Grid and a brief account of its limitations.

INTRODUCTION
Grid computing has been positioned as a promising next generation computing platforms due to the accelerated development of sharing the computation resources from peer computers as well computers across geographically dispersed organizations. It is a geographically distributed environment with autonomous domains that share resources amongst themselves (Azzedin and Maheswaran 2002). Foster et al. (1998) defined grid computing as a hardware and software infrastructure that provides dependable, consistent, pervasive and inexpensive access to high-end computational capabilities. It coordinates and shares computing, application, data, storage, or network resources across dynamic and geographically dispersed organizations. It is a “virtualized” computing system comprised of distributed system components (e.g., processors and storage) interconnected by a high-speed network. Grid computing enables creation of virtual enterprise for sharing and aggregation of millions of resources of geographically dispersed organizations (Buyya et al. 2004). It also suggests a computing paradigm similar to an electric power grid, which is a shared pool of resources that can be accessed on an as-needed basis. Thus grid computing has become a mainstream technology for large scale resources sharing and distributing system integration (Foster et al. 2002).

EVOLUTION OF GRID COMPUTING
Traditionally, computing typically was performed within highly integrated host-centric enterprise computing centers. The rise of the internet and the emergence of e-business have led to a growing awareness that an enterprise’s IT infrastructure encompasses external networks, resources, and services (Foster et al. 2002). In the early 1970s when computers were first linked by networks, the idea of harnessing unused CPU cycles was born. In 1973, the Xerox installed the first Ethernet network and the first full-fledged distributed computing effort was underway. Scientists John F. Shoch and Jon A. Hupp created a worm, as they called it, and envisioned it moving from machine to machine using idle resources for beneficial purposes. Distributed computing scaled to a global level with the maturation of the Internet in the 1990s. One of the most famous projects of Grid computing in those years was the ‘dnet’. It used thousands of independently owned computers across the Internet to crack encryption codes.

The origin of the grid concept was established in the early 90’s when great efforts were made to deploy several gigabyte testbeds such as CASA that linked super computer link across the USA (Lyster et al., 1992). The successes of these testbeds inspired the I-WAY experiment in North America in 1995 where a large scale testbed where multiple supercomputers and advanced visualization devices at 17 different places were connected through ATM networks in order to allow high performance applications to be studied in a controlled environment (Bote-Lorenzo et al. 2004). Grid computing entered the mainstream of research with the establishment of Global Grid Forum (www.gridforum.org) by early grid developers and practitioners. The publication of the book titled “The Grid: Blueprint for a Future Computing Infrastructure” by Foster et al. (1998) laid the groundwork of the field.

TYPES OF GRID
From an application perspective, there are two types of grids: ‘compute grids’ and ‘data grids’ and from a topology view there are intra-grids (enterprise), extra grids/utility), clusters and inter-grids(partner).

As the name suggests the ‘compute grid’ is essentially a collection of distributed computing resources, within or across locations, that are combined together to act as a unified processing resource or a virtual supercomputer. The main aspect of a compute grid is that it eliminates the binding of specific machines to certain computing processes by allowing the aggregated pool to service sequential and other attributes that need the power.

Similarly a ‘data grid’ provides wide area, secure access to current data. These grid enables users and applications to manage ad efficiently use database information from distributed locations. Importance of data
grid is that it eliminates the need to move, replicate or centralize data and thereby saves costs. The above grids rely on software for secure access and usage policies and utilize the existing resources.

‘Intra grids’ or ‘enterprise grids’ work within organizations that want to share information for better use of resources and tangible business benefits. Examples of companies that have deployed Enterprise Grids include JP Morgan Chase, General Motors, Monsanto, Toshiba, etc. ‘Extra’ or ‘utility grids’ basically are computing resources accessible on the internet and delivered by ASPs (Application Service Providers). This is an expansive grid spanning different business models, applications, operating platforms, etc. Example: IBM has successfully implemented an extra grid with its subsidiaries all across the world using the Internet. Many companies around the world are in the implementation process as yet.

CHARACTERISTICS OF GRID COMPUTING
Extant literature (e.g. Foster, 2002; Foster et al., 2001; Grimshaw, 2002) pointed out different characteristics of Grid Computing. Bote-Lorenzo et al. (2004) compiled the key characteristics and outlined the following characteristics of Grid –

a) A grid must be seen as a single virtual computer.

b) Grid computing must be able to manage large scale resources from just a few to millions.

c) Grid resources may be located at geographically distant places.

d) A grid should be able to host heterogeneous resources such as both hardware and software resources varying from data, files, programs, to sensors, scientific instruments, display devices, super-computers, networks, etc.

e) A grid should involve resources of many organizations that allow other user organizations to access the resources.

f) Each organizations involved with a particular grid may have different administration in order to establish and administer different customized security and administrative policies to access and to use the resources of the grid.

g) Resources in a grid should be coordinated in order to provide aggregated computing capabilities.

h) A grid should assure the delivery of services under established Quality of Service requirements.

i) A grid must be built with standard services, protocols, and inter-faces and should hide the heterogeneity of the resources while allowing its scalability.

j) A grid should be accessible to available resources by adapting to a dynamic environment. The grid must tailor its behavior to extract the maximum performance from the available resources.

USES AND BENEFITS OF GRID COMPUTING
Many companies want to take advantage of the cost and efficiency benefits that come from a grid infrastructure. The benefits of Grid computing can be more extensive. They include increased speed of computing, improved productivity and collaboration, more flexible and resilient infrastructure, etc. The benefits of Grid computing are discussed below -

Utilization of Computing Power
The main benefit from the full utilization of computing power of the PCs. The desktop PCs of an organization are often underutilized. “PCs and Windows servers are about 5 percent utilized; Unix servers are 15 percent utilized.” said Dan Powers, vice president of grid strategy at IBM. Even IBM mainframes are only utilized about 65 percent of the time.1 By harnessing these plentiful underused computing assets and leveraging them for revenue-driving projects, the Grid MP platform provides immediate value for companies. Moreover, PCs are not in use on evenings, weekends, and daytime hours can provide significant computational resources.

Faster Project Results
Increased processing speed is an enormous benefit of Grid computing. It squeezes more computing speed from an organization's machines. The increased processing power of a grid enables applications run faster and delivers results more quickly. It has a direct impact on an organization's ability to win in the marketplace by shortening product development cycles and accelerating research and development processes.

Reduced Operational Costs
Grid computing enables to reduce the operational cost. On a price-to-performance basis, the Grid platform gets more work done with less administration and budget than dedicated hardware solutions. Depending on the size of the network, the price-for-performance ratio for computing power can literally improve by an order of magnitude2

Optimized Capital Investment
Reduced hardware capital investment is another significant and attractive cost saving means yielded by grid computing. As applications increase in breadth, volume and complexity, they increasingly require more compute horsepower. This necessitates that more and more funds be spent on hardware. However, with resource utilization across most enterprises, funding the purchase of new hardware is not required. Thus the company may have a lower TCO (total cost of ownership) for its IT projects through effective sharing of investments in storage with other enterprises.

Increased Productivity
Under a Grid infrastructure, productivity jumps due to increased computational activity. Besides, conservation of resources from the underutilized CPUs also contributes to the productivity the whole system. It is aggregating computing power and converting it into value to the

1 http://www.networkmagazine.com/shared/article/showArticle.jhtml?articleId=17602029

2 (www.grid.org)
Sharing Data and Resources

Grid computing allows widely dispersed departments and business to create virtual organizations to share data and resources. As a result, it is of immense help for the successful implementation of international projects. In addition to above mentioned benefits, grid computing facilitates an infrastructure that bonds and unifies globally remote and diverse resources to provide computational support (Bote-Lorenzo et al. 2004). Hence different types of computational support offered by grids can be used in different purposes in business. Following discussions outline different computation supports offered by Grid and their respective uses -

Distributed Supercomputing Support: Grid computing provides distributed supercomputing support that combines computational resources in order to reduce the completion time of a job (Krauter et al. 2002). It also helps to tackle problems that cannot be solved by a single computer system (Foster 1998).

On-Demand Computing Support: Grid computing offers on-demand computing support that helps to retrieve resources that cannot be cost-effectively or conveniently located locally (Foster 1998). For example users who intend to perform accurate stock market analysis and price prediction in their home desktop can be benefitted from on-demand computing by employing grid connectivity (Bote-Lorenzo et al. 2004).

Data Intensive Computing Support: It allows applications to use grids to synthesize new information from distributed data repositories, digital libraries, and databases (Foster 1998). The creation of a new database using data mining from a number of online databases would be an example of data-intensive computing applications (Bote-Lorenzo et al. 2004).

Multimedia Computing Support: It allows applications to use grids to deliver contents assuring end-to-end quality service (Krauter et al. 2002). For example, a video conferencing application is a typical example of application requiring multimedia computing support (Bote-Lorenzo et al. 2004).

APPLICATIONS OF GRID COMPUTING – BANGLADESH PERSPECTIVE

Though Grid computing started in laboratories facilitating researchers to collaborate among themselves and utilize the unused computational power to perform complex mathematical operations, the concept has real business applications right now. Grid technologies can be applied in the following industries -

Universities and Academic Institutions

Universities can utilize the benefits of grid for the greater wellbeing of the all universities of Bangladesh. Universities possess plenty of computers used both for academic and administrative purpose, which can be brought under grid network to utilize their unused resources. Specially, the after hours idle administrative computer resources can be used for academic purposes by throughout the country. In addition, universities can also share their library resources, which will be beneficial to all the parties involved in the Grid network. University Grant Commission (UGC) may wish to conduct a through study and prepare a central plan in this regard. Similar university grid is proposed for Canadian universities to utilize their idle resources for the greater wellbeing (Schick, 2005).

Financial Services

It is a common situation in financial services industry that subsidiaries, divisions and line of business maintain their own set of system and data. Standardizing them by a common technology and harnessing its underutilized resources through Grid Computing can result in huge improvement in operational efficiency. Different financial service companies such as JP Morgan, Charles Schwab, etc. already implemented grid computing and reaped benefit out of it (Glasgow 2003).

Pharmaceutical Industry

Most Pharmaceutical companies derive competitive advantage from faster discovery of new drugs. However such discoveries almost always require molecular modeling and computational chemistry. These tasks needs huge amount of computing power. Pharmaceutical companies can harness the power from unused resources in their organization by adopting grid technology. This can greatly reduce the cost of drug discovery and provide significant competitive advantage. For example, Novartis already implemented grid computing and using it for about a decade by linking together more than 2700 personal computers (PCs) available in its research and development and already came up with several interesting molecules as a result of incorporating the great amount of computing power tied into the grid (Glasgow 2003).

Geosciences

Upstream operations are absolutely vital for petroleum companies and hence analysis of large amount of geophysical datasets for reservoir modeling is critical. It requires huge computing power and companies can significantly reduce seismic imaging turnaround time and maximize reservoir simulation productivity by implementing Grid Computing Technology.

Industrial Engineering

Many aerospace companies and automobile companies use computational fluid dynamics and finite element analysis to improve the design of an automobile part or build a more sustainable aircraft. Employing grid computing can result in running simulation and modeling faster and with more precision. This improvement will result in better quality products with shorter development time, significantly boosting the bottom line.

PERIPHERAL ISSUES OF GRID COMPUTING

Peripheral issues of Grid computing deals with the factors related with the implementation of Grid computing in an organization. To build a Grid, the development and deployment of a number of services is required. However,
the key issues where managers need to focus in this regard are - Service-Oriented Architecture, Security issues and Storage\(^3\).

**Service-Oriented Architecture**

A service-oriented architecture (SOA) is essentially a collection of services. These services communicate with each other. The communication can involve either simple data passing or it could involve two or more services coordinating some activity. It uses web services to connect the service provider and service consumer.\(^4\) It is designed in such a way that it can seamlessly connect with existing platforms and management solutions.

For the successful implementation of Grid, the organizations should have service-oriented architecture as it enables the grid architecture to work flexibly on real world heterogeneous operating platforms and manage the entire infrastructure of a dynamic business. Besides, scalability is also important in an on-demand environment. The architecture should be scalable enough to support the needs of the business as it grows, and maximize administrative productivity in a dynamic environment. A SOA can meet such need of the Grid architecture.

**Security**

Another most important factor to consider in the utility computing model is its security issue. Grid computing calls for servers to be dynamically configured and shared by different partners. As a result, there should be proper deployment and maintenance of security policies. There should be an agreed understanding among the partners to what extent and how much time the resources of the grid will be shared by a specific party. There should be a centralized security department committed to ensuring that these security issues are enforced across the organization and end-users conform to guidelines set forth by the Security department.

**Storage**

As grid computing enables the participants to share data, computational power and other resources, consequently, there should be a storage system of all these data or other resources. The term storage here essentially means a virtual storage, which acts as one of the main enablers of the utility computing model. As the amount of data companies need to store has increased, companies have to buy huge storage devices to keep all those electronic records and information. The two key technologies that are enabling the virtualization of storage for on-demand computing are Storage Area Networks (SAN) and Network Attached Storage (NAS).\(^5\) *Storage Area Network (SAN)* is a high-speed sub network of shared storage devices. A storage device is a machine that contains nothing but a disk or disks for storing data. A SAN’s architecture works in a way that makes all storage devices available to all servers on a LAN or WAN. As more storage devices are added to a SAN, they too will be accessible from any server in the larger network. In this case, the server merely acts as a pathway between the end user and the stored data. Because stored data does not reside directly on any of a network's servers, server power is utilized for business applications, and network capacity is released to the end user.\(^6\) On the other hand, a *Network-Attached Storage (NAS)* device is a server that is dedicated to nothing more than file sharing. NAS does not provide any of the activities that a server in a server-centric system typically provides, such as e-mail, authentication or file management. NAS allows more hard disk storage space to be added to a network that already utilizes servers without shutting them down for maintenance and upgrades. With a NAS device, storage is not an integral part of the server. Instead, in this storage-centric design, the server still handles all of the processing of data but a NAS device delivers the data to the user. A NAS device does not need to be located within the server but can exist anywhere in a LAN and can be made up of multiple networked NAS devices.\(^7\)

**Integration and Transparency of Data**

Another key issue to consider here is the integration and transparency of the data. A strict enforcement of rules for data management needs to be established across the organization. Because of open standards and a distributed environment, any updates made by end-users are available in real time across the organization and business units. But there is no control mechanism to check for errors or duplications in the database. As a result the integration of the data and its transparency is a vital issue here. The end users need to be more responsible in this regard.

**STRATEGIES FOR IMPLEMENTING GRID COMPUTING**

Grid computing is still in infancy and very few companies have embraced these technology full heartedly. However these early adopters have learnt from their mistakes. This continuous learning will someday evolve into best practices once the technology is mature. Therefore, by analyzing how the early adopters like IBM, GM, JP Morgan Chase, and others have implemented Grid Computing and what were their mistakes, some strategies for the successful implementation of Grid Computing are outlined below -

**Evaluate Business Needs**

Though it might sound cliché, this is the classic mistakes corporations have been doing over time while adopting new technology. Grid computing is certainly a “buzzword” in technology arena today; however companies

---


\(^4\) [SOA definition at http://www.service-architecture.com/web-services/articles/service-oriented_architecture_soa_definition.html](http://www.service-architecture.com/web-services/articles/service-oriented_architecture_soa_definition.html)


\(^7\) [http://www.webopedia.com/TERM/N/network-attached_storage.html](http://www.webopedia.com/TERM/N/network-attached_storage.html)
have to contemplate whether they are going to get some business value out of it or not. Our reading of various articles gives us an impression that technology is mature enough to move out of research phase and ready to roll into industry, but suitable only for industries that need enormous computational power and perform complex mathematical operations. Grid computing may be a natural fit for aerospace, financial services and petroleum industries.

Go Slow

Grid computing is still in experimental stage and people are exploring the technology in various directions. It will not be prudent to make huge commitment based on early trends. It has been observed that often technologies change the course and move in different direction, which was not thought of in early stages. Moreover as the concept and standard of Grid Computing is still evolving and such evolving standard means companies have to constantly upgrade the technology, the company might not have enough resources to cope with it. Therefore, the companies should follow a go-slow policy to implement Grid Computing.

Pick a partner not a product

As grid computing is gaining attention from corporate world, many small firms are coming up with new products. However quality of the products and after sales services must be evaluated carefully before choosing a product. We rather suggest that it would be very important to choose a right partner rather than a product. This approach can be handy as the technology is still evolving. Companies must be able to upgrade their existing technologies in line with evolving standards. A vendor who is committed to the technology can be of real value in this case.

Check Compatibility

The biggest challenge for grid computing is that many existing applications are not compatible with grid technology. In other words it means that if application is not written for multi processing, it requires significant effort to get those applications on the grid. It is very important that while doing cost benefit analysis; companies factor in the cost of modifying the applications to make them compatible with grid technology.

Manage Change

Grid computing is a paradigm shift in the way people used resources and thus it is certain to bring enormous changes in people’s work life. This enormous change has to be managed in order to realize the full potential of the technology. Though resources belong to the company, departmental view of personal ownership is hard to shake. Thus it requires significant effort from management to adjust the mindset of the company. This is critical for the enterprise wide success of the grid computing.

FUTURE SCOPE OF GRID

According to an estimate by Grid Technology Partners, a firm that specializes in grid computing, total worldwide market for grid computing is expected to touch $8 billion by end of 2008. Gartner research report says grid computing is one of the few emerging technologies that will transform enterprise computing over the next 10 years. Thus the question is not whether we will be using grid in future; the question is how we will be using grid in future.

On Demand Computing

Information Technology (IT) as an integral part of a business cannot be an exception. On demand computing is a model in which computing resources are made available to the users as needed. Because an enterprise's demand on computing resources can vary drastically from one time to another, maintaining sufficient resources to meet peak requirements can be costly. Thus if company needs exceptionally high computing power for short period of time, it can get it from vendors like IBM and pay based on the computing power used. The same concept can be applied for applications and network storage. Thus combination of web services, grid computing and open architecture will help IT gain advantage from economies of scale. This also means that over a period of time companies may substantially reduce their fixed cost in IT and reduce the total cost of ownership. This also means that IT staff becomes more of a broker between company and vendors thus significantly changing their role.

Global Grid

Grid computing holds excellent promise as the next generation of the shift from proprietary to open based computing architecture. It is the next level of eliminating the complexities of dependencies. This is how we envision a more promising application of on-demand technology. The computing power on PCs is mostly unutilized and hence concept of applying grid is a possibility. Grid computing can be of great value to students, individual researchers, entrepreneurs and professionals. As a student we come across situations where we need access to very expensive software but the state is not willing to pay the huge licensing fees, especially in case of public universities. What if you are an entrepreneur and you need access to a data warehouse that contains critical information about your target customers. Imagine a researcher who needs access to a large amount of processing power for a highly specialized research project. In the current computing environment, it is not possible for these individuals to reach their goals without a huge corporate or government sponsorship. Now imagine if the university can go to a website and purchase the pricey software for 12 students every Wednesday from 5:00 to 6:30 pm. What if the entrepreneur could access the data warehouse for fifteen minutes and download the information he/she needs. What is the researcher could tap into the grid for a couple of hour’s everyday to perform his calculations at 600 GHz. Though this may sound day dreaming, the apt combination of web services, grid computing and open source architecture can make it happen. A global grid connecting computing resources all over the world accessible through internet is the ultimate destination proponents of grid computing should strive for!
LIMITATIONS AND CONCLUSION

Grid computing is a technology with huge potentials. It could bring a paradigm shift in the way people share computing resources. However, grid computing is still in its infancy and hence standards are not well developed. As grid resources are not entirely dedicated to the environment, computational and networking capabilities can vary significantly over time (Cao et al. 2003). Hence, predicting application performance may become difficult and real-time resource information update within a large-scale global grid may become impossible. In addition, as the process of grid work flow encompass multiple administrative domains or organizations, lack of central ownership and control over the grid may result in incomplete information processing as well as many other uncertainties (Cao et al. 2003).

Despite these limiting issues, it is sure that adopting grid computing will help improving operational efficiency of existing infrastructure and achieve more collaboration among various organizations. Specially, computational grid enables creation of virtual computing environment for sharing and aggregation of distributed resources for solving large-scale problems in science, engineering, and commerce (Buyya et al. 2002). Hence, application of grid computing is limited to businesses where high computational power is required. If web services and grid computing could be merged seamlessly, it can be utilized the productivity of masses and world can reap enormous advantages from it. Consequently, relevant legislations should be enacted in order to ensure proper direction and management of the resources of the virtual organization developed by the GRID itself. Hence further research is warranted in this regard.

REFERENCES


Web sources:

1. Liberated from Light Reading by Rick Thompson
3. (www.grid.org)
5. SOA definition at http://www.service-architecture.com/web-services/articles/service-oriented_architecture_soa_definition.html