GIRD COMPUTING- A TOOL FOR ENHANCING THE COMPUTING POWER

MANJULA K A
LECTURER
KANNUR UNIVERSITY
KERALA, INDIA
manjulaka@gmail.com

KARTHIKEYAN P
MES COLLEGE OF ENGINEERING
KUTTIPPURAM, KERALA

***Abstract* –** With the enormous increase in the demand for computing capacities, solutions with least investment have to found out. In this direction Grid technology is finding its way out of the academic incubator and entering into commercial environments. Here geographically distributed resources, such as storage devices, data sources, and supercomputers, are interconnected and exploited by users around the world as single, unified resource. This helps in use the idle time of these resources, which is otherwise lost. This article discusses briefly about grid computing and its benefits. Application of this technology is increasing and this article also looks at some of these applications.

Keywords: Distributed Computing, Grid Computing, Benefits of grid Computing, Application of Grid Computing.

1. Introduction

The spread of high-speed broadband networks, the continual increase in computing power, and the growth of the Internet have changed the way in which society manages information and information services. Geographically distributed resources, such as storage devices, data sources, and supercomputers, are interconnected and can be exploited by users around the world as single, unified resource. This has led to the advent of distributed computing, in which different resources both physical and virtual can be shared. While the physical resources include computational power, storage devices and communication capacity, virtual resources include operating systems, software, applications and services, which can be exchanged and are independent from its physical location. One kind of distributed computing is Grid Computing.

2. Grid Computing

Grid computing has recently enjoyed an increase in popularity as a distributed computing architecture. Grids are very large-scale virtualized, distributed computing systems, which allows users to access the computing resources of many different machines distributed around the world. Grid computing got its name because it strives for an ideal scenario in which the CPU cycles and storage of millions of systems across a worldwide network function as a flexible, readily accessible pool that could be harnessed by anyone who needs it, similar to the way power companies and their users share the electricity grid. They cover multiple administrative domains and enable virtual organizations. Such organizations can share their resources collectively to create an even larger grid. Fig.2 shows a sketch of grid computing.

From the Search for Extraterrestrial Intelligent (SETI) to the Search for ways to utilize the unused computing power across the enterprise, grid computing has come of age. It promises to harness the spare clock cycle of all your computers and use this newfound power to speed up the most complex of your computational or data processing demands (Asagba, et al. 2008). It also gives access to all the storage, all the data, of all those PCs and networked systems working on any processor-intensive task (Davey, 2003 & Foster, et al, 1999)

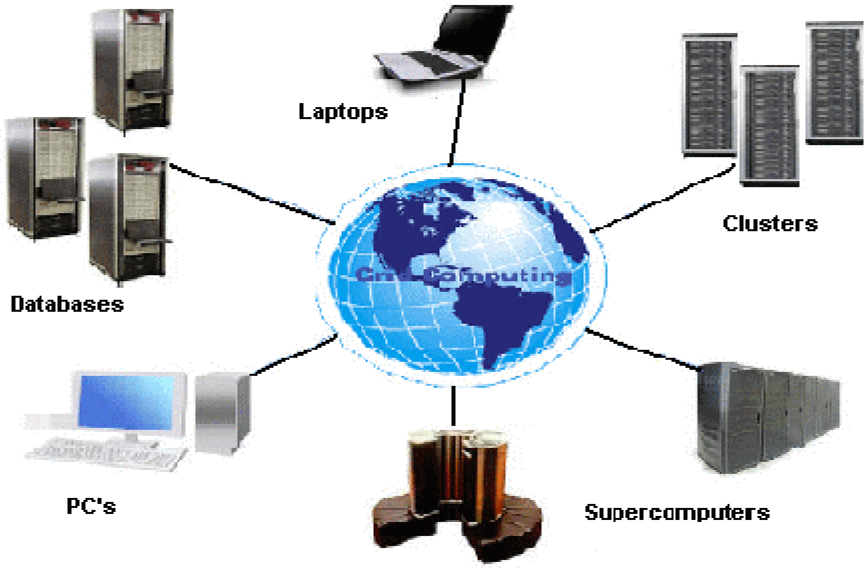

Fig. 2: Grid Computing

2.1 Evolution of Grid Computing

The term Grid Computing originated in the early 1990s as a metaphor for making computer power as easy to access as an electric power grid. CPU Scavenging and Volunteer became popular in 1997 by distributed.net and later in 1991 by SETI@home to harness the power networked PCs worldwide, in order to solve CPU-intensive research problems (Berman, et al, 2003). The concept of grid computing was the brainchild of Ian Foster, Carl Kesselman and Steve Tuecke, and it was made known in one of their seminal presentations, "The Grid: Blueprint for a new computing infrastructure." Their efforts created the Globus Toolkit, which contains computation management, storage management, security management, monitoring and other related services. They are widely acclaimed as the “father of the grid” (Asagba, et al. 2008).

A Grid system is a virtual organization comprising several independent autonomous domains (Akinyemi, et al, 2007). The Grid community often refers to the notion of a “virtual organization” (VO). A virtual organization exists as a corporate, not-for-profit, educational or otherwise productive entity that does not have a central geographical location and exists solely through telecommunication tools (Meliksetian, et al, 2004). Grid enables people to be members of many VOs which gives one access to various computational, instrument-based data and other types of resources. VO makes its resource much more useful and accessible for their users (Joseph, et al, 2004).

2.2 Benefits of Grid Computing

Grid Computing has proved benefitial in many ways, some of these benefits are (Asagba, et al. 2008):

1. *Exploitation of Under-Utilized Resources:* Exploitation of under-utilized resources by running an existing application on different machines, exploiting idle times on other machines and aggregating unused disk drive capacity.
2. *Reduces Computational Time:* Computational time is reduced for complex numerical and data analysis problems.
3. *Provide Information Access:* Information accessibility to maximize the exploitation of existing data assets by providing unified data access during the querying process of non-standard data formats (Farago-Walker, 2001).
4. *Reduces cost by optimising existing IT infrastructure:* The grid facilitate reduction of costs by optimising the use of existing IT infrastructure investments and by enabling data sharing and distributed workflow across partners, and therefore enabling faster design processes. (Foster, et al, 2005B).
5. *Providing access to parallel CPU capacity:* Grid computing offers potential access for large-scale parallel computation to enhance performance in computationally intensive applications.

6. *Offers improved reliability:* Grid technology offers alternate approach to achieving improved reliability. Parallelization can boost reliability by having multiple copies of important jobs run concurrently on separate machines on the grid. Their results can be checked for any kind of inconsistency, such as failures, data corruption and tempering.
7. *Provision of resource balancing:* The grid offers good resource balancing measures that can handle occasional peak loads, job prioritization, and job scheduling.
8. *Effective management of resources:* With grid technology, management of organization can easily visualize resource capacity and utilization to effectively control expenditures for computing resources over a larger organization.
9. *Interoperability of virtual organizations:* The grid offers collaboration facilities and interoperability of different virtual organization by allowing the sharing and interoperation of the heterogeneous resources available.
10. *Access to additional resources:* The grid offers access to other specialized devices such as the cameras and embedded systems.
11. *Harnessing heterogeneous systems together:* Grid computing can be used to harness heterogeneous systems together into a mega computer by applying greater computational power to a task.
12. *Grid virtualization:* Grid computing offers grid virtualization, thereby making a single, local computer to undertake powerful applications.

2.3 Applications of Grid Computing

Grid Computing is being extensively used in various areas like science, business, governments etc. The areas and scope of application is increasing at a very fast pace. In this section a brief discussion on the application of Grid Computing in various domains is done. This is intended to give only an idea of the application of Grid Computing and represents only the tip of the iceberg. The fig. 1 shows some of the possible application areas.

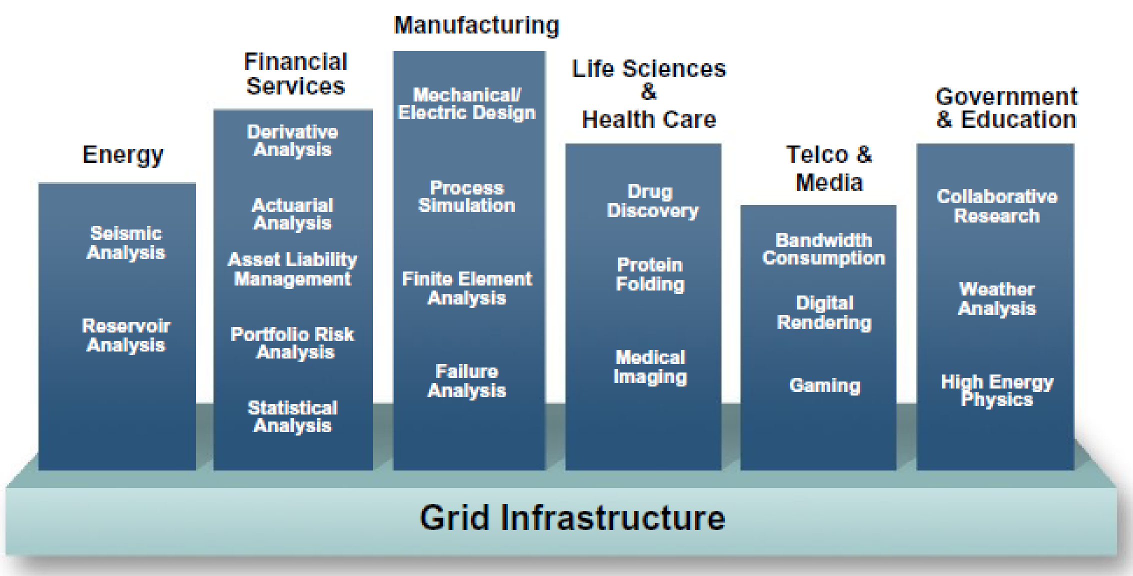

Fig 1: Grid Computing Applications (Source: Watkins, 2006).

2.3.1 Energy

Grid Computing have been extensively used in analysis of seismic data coming from under earth's surface. Seismic data from earth movements are being used for understanding, modeling and forecasting of earthquakes. The discovery and evaluation of oil and gas reservoirs have long required powerful computers and processing of very large amounts of data. These requirements have intensified as the industry has increasingly explored for hydrocarbons under ocean waters. Offshore oil and gas exploration requires immense levels of information processing, seismic imaging is the key technology in discovering these oil and gas fields. This technology has become steadily more sophisticated requiring increased data-processing capacity. Grid Computing has been of immense help in processing these huge volumes of data at a very high speed.

2.3.2 Financial Services

With the emergence of a competitive market force, customer satisfaction, and reduction of risk are the most competitive areas financial communities continually strive to achieve. The requirements related to sophistication, accuracy, and faster executions are among the more salient objectives across financial communities. These objectives are now achieved by real-time access to the current and historical market data, complex financial modeling based on the respective data, and faster response times to user queries. Grid Computing provides this sector with advanced systems delivering all the competitive solutions (Jia, 2006). These solutions are dependent upon providing increased access to massive amounts of data, real-time modeling, and faster execution by using the grid job scheduling and data access features. Various grid solutions were offered including, among others, the implementation of the Open Grid Service Environment (Hochreiter , 2005) that provides an abstract service stack used to model large-scale computational financial problems as abstract workflows. Grid computing has been used for developing data mining applications for supplying effective online financial services for both retail and corporate customers. This is useful for intelligent risk management and decision making for financial enterprises and partners (Fan, 2000).

2.3.3 Manufacturing

The enormous competitive pressure in the business and industry sectors today afford most engineering and design far less turnaround time. They need mechanisms to capture data, speed up the analysis on the data, and provide faster responses to market needs. These engineering activities and design solutions are inherently complex and include, the analysis of real-time data to find a specific pattern within a problem, the parametric studies to verify different aspects of the systems, the modeling experiments to create new designs and the simulation activities to verify the existing models for accuracy. Here Grid Computing systems provide a wide range of capabilities that address these kinds of analysis and modeling activities. These advanced types of solutions also provide complex job schedulers and resource managers to deal with computing power requirements.

In the case of engineering work, virtual collaboration on the design, production and maintenance of products that are described in complex structured product model databases is a major challenge (Turk et al., 2004, Dolenc et al., 2004 & IntelGrid). Existing web service paradigms can weakly support such collaboration (Maad et al., 2005). The potential grid solution lies in an appropriate implementation of the VO concept which provides a robust framework for engineering collaboration due to its flexibility, adaptability, and security.

2.3.4 Life Science & Health Care

2.3.4.1 Life Science

There have been dramatic advances in the life sciences sector, which have in turn provided rapid changes in the way that drug treatment and drug discovery efforts are now being conducted. These advances have now presented a number of technical challenges to the information technology sector. These challenges include huge amounts of data analysis, data movement, data caching, and data mining. In addition to the complexity of processing data, there needs to be additional requirements surrounding data security, secure data access, secure storage, privacy, and highly flexible integration. The Grid Computing systems can provide a common infrastructure for data access, and at the same time, provide secure data access mechanisms while processing the data. Today, life sciences utilizes the Grid Computing systems to execute sequence comparison algorithms and enable molecular modeling using the above-collected secured data. This now provides the Life Sciences sector the ability to afford world-class information analysis respective to this discussion, while at the same time providing faster response times and far more accurate results. The analytics and system efforts' surrounding genomic, proteomics, and molecular biology efforts provides the basis for many of these Grid Computing advancements in this sector., and especially the Grid Computing disciplines.

2.3.4.2 Health Care

One of the most obvious applications of grid computing is in medicine. Imagine if your doctor had access to a grid that could handle administrative databases, medical image archives and specialized instruments such as MRI machines, CAT scanners and cardio angiography devices... This could enhance diagnosis procedures, speed analysis of complex medical images, and enable life-critical applications such as telerobotic surgery and remote cardiac monitoring. Also grid computing has been useful in a variety of situations like producing interactive medical simulations like heart simulation and in analyzing and managing medical images. Grids have also been successfully used in supporting virtual collaboration in e-Hospitals (a virtual network of hospitals that delivers medical training, e-surgery, and medical analysis services).

2.3.5 Telecom & Media

2.3.5.1 Media

The major challenge facing media applications is the production, broadcasting, delivery and playout of interactive media content (audio, video, image) in real time. Grid solutions have been developed for this, these include, the Grid Visualization Kernel (GVK) [22] which allows the visualization pipeline [23] (data enrichment/reduction, followed by mapping of the data to an abstract form, and finally composing the visual image) to be ported to grid resources, and also handles the communication between the simulation generating the data and the visualization of the simulated data; and G-Vid [24] which is based on GVK which allow the production of real time interactive MPEG4 compliant video content on the grid.

2.3.5.2 Collaborative Games

There are collaborative types of Grid Computing disciplines that are involving emerging technologies to support online games, while utilizing on-demand provisioning of computation-intensive resources, such as computers and storage networks. These resources are selected based on the requirements, often involving aspects such as volume of traffic and number of players, rather than centralized servers and other fixed resources. These on-demand-driven games provide a flexible approach with a reduced up-front cost on hardware and software resources. We can imagine that these games use an increasing number of computing resources with an increase in the number of concurrent players and a decrease in resource usage with a lesser number of players. Grid Computing gaming environments are capable of supporting such virtualized environments for enabling collaborative gaming (Maad et al., 2005).

2.3.6 Government & Education

2.3.6.1 Government

The Grid Computing environments in government focus on providing coordinated access to massive amounts of data held across various agencies in a government. This provides faster access to solve critical problems, such as emergency situations, and other normal activities. These key environments provide more efficient decision making with less turnaround time. Grid Computing enables the creation of virtual organizations, including many participants from various governmental agencies (e.g., local, state and central). The formation of virtual organizations, and the respective elements of security, is most challenging due to the high levels of security in government and the very complex requirements (Maad et al., 2005). Another area of application is e-government, where the major *challenge* is the development of an e-government infrastructure that supports the shift to a service-oriented e-government model (Maad et al., 2005b).

2.3.6.2 Education

There is a big difference in education sector between first world and developing world. This difference is based on digital learning resources and computing power. Unfortunately all of these resources are geographically distributed all over the world. There is no doubt that ICT is playing a big role for introducing E-learning around the world. But if we really want to overcome these differences we need a revolutionary approach that support mutual use of geographically distributed computing and learning resources as an aggregated environment that will create new ways of flexibility, interoperability and extensibility. Fortunately Grid is the technologies that can integrate all of these resources of knowledge's and produce super-computing power from those geographically distributed computers to access those knowledge's without sacrificing local autonomy(GF/117). For example, the Indian Institute of Information Technology and Management, Kerala is heading a Grid project known as Kerala Educational Grid that envisages the linkage of colleges and universities to resource centres that will supply education materials on demand and increase cooperation and networking among the affiliated academics (Sherly, 2005).

2.3.6.3 Environment

Environment applications involves parallel execution of hundreds of programs corresponding to large scale air pollution, nuclear waste storage, pollution and climate models. Various grid *solutions* have been offered. For a particular large scale air pollution model, a national VO was formed (Thandavan, 2004) involving researchers and resources from 7 UK institutions. At an international level, a similar VO was formed involving 21 EU institutions within the CrossGrid consortium (CrossGrid).

2.3.6.4 Weather Forecasting

Another field with huge data generation and processing requirements is weather forecasting. Both local weather data stations and satellites collect and transmit large volumes of data for analysis. Networks of detectors have been placed in the ocean in many locations to detect tsunamis and predict their size and course to permit timely evacuations of coastal areas. Grid Computing has been very extensively used for forecasting of weather and other natural catastrophic events.

2.3.6.5 Research Collaboration

Research-oriented organizations and universities undertaking advanced research collaboration require the analysis of tremendous amounts of data. Some examples of such projects are subatomic particle and high energy physics experiments, remote sensing sources for earth simulation and modeling, and analysis of the human genome sequence. These organizations engaged in research collaboration activities generate petabytes of data and require tremendous amounts of storage space and thousands of computing processors. Researchers in these fields must share data, computational processors, and hardware instrumentation such as telescopes and advanced testing equipment. Most of these resources are pertaining to data-intensive processing, and are widely dispersed over a large geographical area. In these situations the Grid Computing discipline provides mechanisms for resource sharing by forming one or more virtual organizations providing specific sharing capabilities. Such virtual organizations are constituted to resolve specific research problems with a wide range of participants from different regions of the world.

In physics, for example the LHC (LHC) is projected to begin pumping out a tsunami of raw data, an annual output of 15 petabytes (15 million gigabytes) will soon dwarf that of any other scientific experiment in history. Scientists at CERN, the European Organization for Nuclear Research, organized the LHC Computing Grid, a network of 100,000 computers in 33 countries, to cope with this unprecedented mountain of data that will be generated by the LHC. Similarly, in the field of chemistry, molecular design and engineering is proving to be a major challenge (Schuller et al, 2005). Grid solutions like OpenMolGRID (OpenMolGRID) which provides the necessary infrastructure for molecular design and engineering is proving to be helpful.

2.3.6.6 Astronomy

The major *challenge* facing the field of astronomy is the analysis of tera-bytes of astronomical image data generated by telescopes. Moreover, astronomical image capturing devices can generate several images, each of hundreds of Mbytes, per single shot (Yamamoto, 2004). This necessitates: data intensive computation; scalable file I/O in the order of GB/s; replica management; and parallel / distributed processing of files. Grid computing is essential and has been successfully used for such data intensive computation.

3. Conclusion

Grid Computing has emerged as a powerful tool for enhancing the computer power. This technology also called 'poor man's supercomputer' uses the idle capacities of various geographically spread resources. These idle capacities can be used to perform calculations and activities which otherwise would take long time to perform. Grid Computing has been briefly explained here. The benefits of Grid Computing and various applications have also been discussed.

Reference

[1] Akinyemi, I.O., Daramola, J.O. and Adebiyi, A. A.(2007), Grid-Enabled e-learning Framework for Nigerian Educational Institutions, Nigeria Computer Society, 21st National Conference Proceedings, Vol. 18, pp. 91-98.
[2] Asagba, Prince Oghenekaro., Ogheneovo, Edward E. (2008). Qualities of Grid Computing that can last for Ages. J. Appl. Sci. Environ. Manage., Vol. 12(4) 47 – 52.
[3] Berman, F., Anthony J. G. H. and Geoffrey C. F. (2003). Grid Computing: Making the global infrastructure a reality, ISBN 0-12-742503-9.
[4] Cross Grid, http://www.crossgrid.org
[5] Davey, W. (2003). Grid computing - spreading the word, PC PRO: Computing in the real world, pp.191-194.
[6] Foster, I. and Carl K. (1999), The Grid: Blueprint for a New Computing Infrastructure, Morgan Kaifmann Publishers. ISBN 1-55860-475-8.
[7] Dolenc M., Turk Z., & Stankovski V. 2004. InteliGrid Project: A Vision of Engineering on the Grid, Proceeding Cracow'04 Grid Workshop, Crakow, Poland, p.19.
[8] Fan, M. 2000. The Internet and the Future of Financial Markets, Communications of the ACM, 43(11), 82-88.
[9] Farago-Walker, S. (2001), Peer-to-Peer Computing: Overview, Significance and Impact, E-learning, and Future Trends, http//www.elearningmag.com.
[10] Foster, I., and Tuecke, S. (2005B), The Globus Project: A Status Report.
[11] Hochreiter R., Wiesinger C., Giczi D. 2005, Large Scale Computational Finance Applications on the Open Grid Service Environment, In proceeding of European Grid Conference, 2005, The Netherlands.
[12] InteliGrid, http://www.inteligrid.com/index2.htm
[13] Jia, H., Muneaki, O. & Ning, Z. 2006. Developing Grid-based E-finance Portals for Intelligent Decision Making. The 20th Annual Conference of the Japanese Society for Artificial Intelligence.
[14] Joseph, J., Ernest, M. and Fellenstein, C. (2004), Evolution of Grid Computing Architecture and Grid Adoption Models, IBM Systems Journal, Vol. 43, No. 4, pp. 624 – 645.
[15] LHC Computing Grid, http://lcg.web.cern.ch/LCG/
[16] Maad, S. and Coghlan, B. and Pierantoni, G. and Kenny, E. and Ryan, J. & Watson, R., 2005a. Adapting the Development Model of the Grid Anatomy to meet the needs of various Application Domains, Cracow Grid Workshop (CGW'05), Cracow, Poland, November.
[17] Maad S., Coghlan B., John R., Eamonn K., Watson R., and Pierantoni G. 2005b. 'The Horizon of the Grid For E-overnment'. Proceeding eGovernment'05 Workshop, Brunel, United Kingdom.
[18] Meliksetian, D. S., Prost, J. P., Bahl, A. S., Boutboul, I., Currier, D. P., Fibra, S., Girard, J. Y., Kassab, K. M., Lepesant, J. L., Malone, C., and Manesco, P. (2004), Design and Implementation of an Enterprise Grid, IBM Systems Journal, Vol. 43, No. 4. pp. 646 – 664.
[19] OpenMolGRID, http://www.openmolgrid.org/
[20] Schuller B., Romberg M. & Kirtchakova L. 2005, Application Driven Grid Developments in the OpenMolGRID project, In proceeding of European Grid Conference, The Netherlands.
[21] Sherly, E. 2005. Data Grid Architecture for a Distributed Data Management system: Kerala Educational Grid Perspective. CSI Communications, July, 2005.
[22] Thandavan A., Sahin C., & Alexandrov V. N. 2004. Experiences with the Globus Toolkit on AIX and deploying the Large Scale Air Pollution Modelling application as a grid service 27, Proceeding Cracow'04 Grid Workshop, Crakow, Poland
[23] Turk, Z. M. Dolenc, J. Nabrzyski, P. Katranuschkov, E. Balaton, R. Balder, & M. Hannus. 2004. Towards Engineering on the Grid, in Dikbas A. and Scherer R. (eds.) Proceedings, ECPPM Conference, Istanbul.
[24] Watkins, R. 2006. IBM Grid Computing & Virtualization, GRID @ Asia Conference**,** Seoul, Korea, December 11-13.
[25] Yamamoto N., Tatebe O., and Sekiguchi S. 2004. "Parallel and Distributed Astronomical Data Analysis on Grid Datafarm", Proceedings of 5th IEEE/ACM International Workshop on Grid Computing (Grid 2004), 461-466.