Journal of Advances and Scholarly Researches in Allied Education
Vol. XV, Issue No. 5, July-2018, ISSN 2230-7540

Grid Computing Introduction

Abdur Rakib*

VILL- Pat Pukur, P.O-Digha Patpukur, Distric-Nadia State West Bengal, Pin-741154

Abstract* – *Grid computing has been a buzzword in information technology since past few years. Grid computing is an infrastructure involving collaboration of computers, databases & network resources available, to perform manipulation of intensive and large scale data set problems. The hike in the complexities of computational problems in modern era of science and technology forced the engineers and scientists to cross the organizational boundaries to get desired data manipulation. The best logical solution to this issue is distribution of the problem set over multiple computational resources/nodes. Several solutions to grid computing has been developed and are still evolving, since the notion of Grid sprang up in mid 1990s, most of which came from the academic research projects. This paper presents an introduction to Grid computing providing insight into the essential features, architecture, scope and challenges of grid computing.

Keywords* – *Grid Computing, Computation, Introduction, Architecture, Applications, Challenges.

- x -

1. INTRODUCTION TO GRID COMPUTING

Modern scientific problems have become so huge that their computation may take time anywhere between hours and years. The best logical solution to this is distribution of problem's data set over multiple computational resources/nodes. This concept lead to what today we call "Grid Computing". The basic notion of grid came into existence in mid 1990s but still Grid means different to different people. Grid is analogous to the electric power grid which is supposed to provide astonishingly consistent, steady going, transparent access to electricity irrespective of from its source. The consumer just utilizes the electricity plugged through wall sockets. Various definitions for the Grid are provided by various network technology gurus. Ian Foster has indicate a definition of the Grid as "a system that coordinates resources which are not subject to centralized control, using standard, open, general-purpose protocols and interfaces to deliver nontrivial qualities of service . The Globus Project defines Grid as "an infrastructure that enables the integrated, collaborative use of high-end computers, networks, databases, and scientific instruments owned and managed by multiple organizations]**.** Another definition put by Gridbus Project as "Grid is a type of parallel and distributed system that enables the sharing, selection, and aggregation of geographically distributed "autonomous" resources dynamically at runtime depending on their availability, capability, performance, cost, and user's quality-of-service requirements. The capability of modern computer systems and computer networks has increased exponentially as compared to traditional computer systems. This increase in their performance, most of the times lead to wastage of computational resources because most of the time the CPU sits idles. Grid utilizes this idle CPU cycles to perform the computation when requested by the grid users, which otherwise would have been wasted. This enables the users to perform complex computations that in traditional cases would have required large-scale computing resources e.g. image rendering, scientific researches, climatology, etc. There are large groups of people belonging to academia, business houses, and scientific research laboratories working on grid. Taking the basic idea from electric power grid major progress has been under laid, and researchers subsequently have developed various solutions.

Fig. 1 Grid computing

Types of grid:

On the basis of use grid computing can be divided into different types:

- **Computational grids:** These type of grid are meant to provide secure access to computational resources, sufficient enough to perform processing of computational problems which otherwise would have required high computing power machines.

Collaboration grid: With the advances in network hardware resources and internet services, demand for better collaboration has increased. Such desired collaboration is best possible with these kinds of grids.

Utility Grid: In this type of grid not only CPU cycles are shared, also other softwares and special peripherals like sensors are also shared.

Network grid: Even if we have computational machines with enough computational power as a part of grid but with poor network communication one can't utilize those machines optimally. Network grid provides high performance communication using data caching between nodes there by speed-up communication with each cache nodes acting as router.

Data grid: There are two things, data and computation over that data. Data grid provides the support for data storage other data related services like data discovery, handling, publication, etc.

II. GRID ARCHITECTURE:

In grid computing infrastructure resources belong to and come from physically scattered administrative domains to collectively provide various resources (data, computing, and network) to the users. In a grid, computing nodes might not be placed at common physical location but can be independently operated from different locations. Each computer on the grid is a distinct computer [6]. Collection of servers clustered together to work out a common problem forms a grid [7]. The computers joined to form a grid may even have different hardware and operating systems.

Grid consists of a layered architecture model providing protocols and service at five different layers represented by Fig.2.

Fabric layer: Fabric layer sits at the bottom of this layered architecture; it provides shareable resources such as network bandwidth, CPU time, memories, scientific instruments like sensors, telescope, etc. Data received by sensors at this layer can be transmitted directly to other computational nodes or can be stored in the database over grid. Standard grid protocols are responsible for resource control. Accomplishment of sophisticated sharing operation is the measure for quality of this layer. Operating system, queuing systems and processing kernels also form the part of this layer.

Connectivity layer: This layer specifies the protocols for secure and easy access. Protocols related to communication and authentication required for transactions are placed in this layer. These communication protocols permit the exchange of data between resource layer and fabric layer. Authentication protocols are meant to provide secure cryptographic mechanisms for identification of users and resources. E.g. GSI –Grid Security Infrastructure (built around existing TLS protocols).

Resource layer: This layer specifies the protocols for operating with shared resources. Resource layer build on the connectivity layer's communication and authentication protocols to define Application Program Interfaces (API) and software development kit (SDK) for secure negotiation, accounting, initiation, control, monitoring and payment of sharing resources. E.g. GRIP (Grid resource Information Protocol; based on LDAP), GRAM (Andre Rigland, 2010). (Grid Resource Access and Management) for allocation and monitoring of resources.

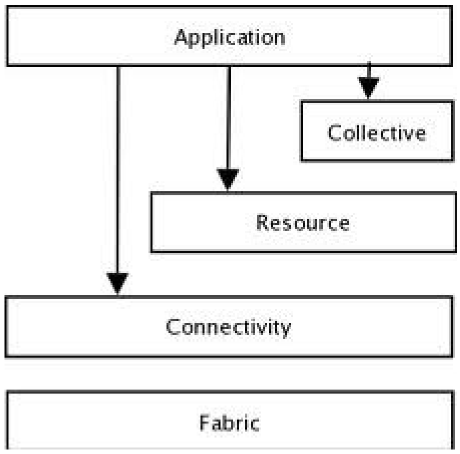

Fig. 2 Grid architecture

Collective layer: This layer consists of general purpose utilities. Any collaborative operations in the shared resources are placed in this layer and it coordinates sharing of resources like directory services, co-allocation, scheduling, brokering services, monitoring and diagnostic services, data replication services.

Application layer: At the top of the grid layered architecture sits the application layer. This layer consists of application which the user will implement. Moreover, this layer provides interface to the users and administrators to interact with the grid.

III. APPLICATION OF GRID COMPUTING

Grid makes an optimized use of resources, utilizing the CPU cycles which otherwise would have been wasted. With this the users can get extra computation resource and thus can process their large-scale computational problems thus solving a complex problem to the computational level of a supercomputer. With this fundamental work scenario of Grid computing, it has a long litany of usages going from small academic project requirement to scientific laboratories used for extra-terrestrial activity.

There are four important procedures which must be carried out in a distributed computing system before it can be called The Grid. These are the Authorization, Authentication, Resource Access and Resource Discovery. These four vital procedures lead to the idea of Virtual Organizations of collaborators who share resources over a Grid. Above mentioned four procedures are the series of steps too from task submission to the grid and getting task executed over grid.

Major benefits which can be utilized by application of grid are the following (Ferreira et. al., 2005).:

- ► Improved efficiency/reduced costs.
- ► Optimized utilization of underutilized resources.
- ► Exploiting underutilized resources
- ► Virtual resources and Virtual Organization (VO)
- ► Increase capacity and productivity
- ► Parallel processing capacity
- ► Resource balancing
- ► Heterogeneous system support
- ► Reduced time of result.

Grid computing provides a way for computation of high data intensive problems like financial modelling, protein folding, climate/weather modeling, image rendering, earthquake simulation, etc.

The various domains in which grid computing is utilized are the following:

Engineering Design and Automation: Computational aerodynamics, artificial intelligence and automation, finite-element analyses, remote sensing applications, pattern recognition, computer vision, image processing, etc.

Medical, Military and Basic Research: Polymer chemistry, medical imaging, nuclear weapon design, problem of quantum mechanics, etc.

Predictive Modelling and Simulation: Flood warning, socio-economic and government use, numerical weather forecasting, astrophysics (Modelling of Black holes and Astronomical formations), semiconductor simulation, Oceanography, human genome sequencing, etc.

Energy Resource Exploration: Plasma Fusion power, seismic exploration, nuclear reactor safety, reservoir modelling, etc.

Visualization (Jacob, et. al., 2005): computer-generated graphics, films and animations, data visualization, etc.

Grid is becoming a culture among researchers and people in academia changing the way they collaborate to make scientific discoveries (Jiang, et. al., 2007).

IV. CHALLENGES TO GRID COMPUTING:

Although tremendous benefit can be drawn from grid computing but road to grid is not free of jerks. Inherent nature of grid i.e. heterogeneity of software and hardware, handling wide spread resources, control of different organizations pose serious challenges before the researchers. Many scientific problems which cannot be tested practically are simulated over grid. But this simulation in itself is a challenge for grid because no standard has yet been developed for simulation over grid. Moreover, simulation models developed for traditional hardware systems are not valid for the modern systems [11]**.** No doubt grid computing seems like a promising solution to eliminate the resource islands and to provide resources and services over internet in a transparent way. But to achieve all above, we need to analyses the current barriers and challenges in developing, deploying, promoting and use of grid computing.

The various challenges to the grid are the following:

- Grid reliability:
- Scheduling of tasks
- Load balancing
- Resource monitoring
- Service availability
- Distributed management (Foster, 2002).
- Availability of data

- Uniform user friendly environment.
- Grid application development
- Standard protocols
- Efficient algorithms and problem solving methods
- Programming models and tools
- Management and administration of grid
- Performance analysis and resource monitoring
- Centralized management

The list of grid challenges is not limited to the above. There are various other challenges to grid like hidden costs, no widely accepted definition and scope of grid computing, lack of significant applications for grid, availability of widely accepted standard protocols to control and manage grid and much more. All these leading to the shift of user's attention from grid computing to service computing. Above mentioned issues forces to rethink, "for what indeed the grid computing is".

V. CONCLUSION AND FUTURE SCOPE:

Despite substantial research in grid computing certain issues are still holding back grid computing from becoming widespread in its use. Today Grid computing has been utilized by most of the scientific domains like astronomy, biological science, climatology, and much more. But the prime requisite for using grid is high speed internet, if one does not have a high speed internet one cannot get the best benefits from grid. This has been holding back grid for long. On one end of grid is high computation and optimized utilization of resources and at other the ability to manage distributed and heterogeneous systems. We need security with high availability of data and resources on demand and at the same time ease of access to implement these. On-demand provisioning of resources and more secure protocols for users and service providers to monitor and arrange payment is still a necessary requirement for commercial utilization of grid. No doubt grid computing is still evolving and better protocols and standards will be implemented in near future, culled from academia and business houses to speed up the grid evolution process. It will be of interest how various researchers would find new ways to improve grid and move along parallel paths. There is a lot of scope for Grid Computing as more and more complex algorithms are being developed. It is being speculated that grid computing through dedicated fiber optic links, might change the current scenario of broadband and other services providing bandwidth ranging from gigabytes to terabytes.

VI. ACKNOWLEDGMENT

I would like to express my special thanks and gratitude to my guide Ms. Poonam Dabas who gave me the golden opportunity to write this paper on the topic grid computing, which also helped me in carrying out an intensive research through which I came across so many new concepts and terminologies. I am really thankful to her.

REFERENCES:

Krishna Nadiminti, Srikumar Venugopal et. al.: "The Gridbus Grid Service Broker and Scheduler (v.2.2) User Guide", [Online] Available: http://www.cloudbus.org/broker/2.2/manual v2.2.pdf

Ovaise Khan (2005). "Types of Grid", [Online] Available: http://thegridweblog.blogspot.in/2005/10/types-of-grids.html.

http://cloudcomputingapps.org/what-is-the-difference-betweencluster-computing-and-grid-computing.html

http://www.linkedin.com/answers/technology/enterprisesoftware/TCH_ENT/552522-14941035

Luis Ferreira, Fabiano Lucchese, Tomoari Yasuda, Chin Yau Lee, Carlos Alexandre Queiroz, Elton Minetto (2005). Antonio Mungioli First Edition (April 2005). Grid Computing in Research and education.

http://www.jatit.org/research/introduction_grid_computing.htm

Bart Jacob, Michael Brown, Kentaro Fukui (2005). Nihar Trivedi First Edition (December 2005). Introduction to Grid Computing.

Congfeng Jiang, Cheng Wang, Xiaohu Liu and Yinghui Zhao (2007). "A survey of Job Scheduling in Grid."

I. Foster (July 20, 2002). "What is Grid? A three point checklist, Grid Today " vol. 1; no. 6; pp. 22-25

R. Esposito, P. Mastroserio et. al., (March 2003). "Standard FTP and Grid FTP Protocols for International Data Transfer in Pamela Satellite Space Experiment". [Online]. Available: http://arxiv.org/ftp/hep-ex/papers/0305/0305084.pdf

Andre Rigland (2010-12-17). "Cloud Computing-How it is Different from Grid Computing".

Trial lecturer in conjunction with PhD thesis defence.

Corresponding Author

Abdur Rakib*

VILL- Pat Pukur, P.O-Digha Patpukur, Distric-Nadia State West Bengal, Pin-741154

E-Mail – drrakib1978@rediffmail.com