GRID COMPUTING: An Overview to Architecture and Issues

Gurpinder Singh[1]

[1]*Assistant Professor, PUSSGRC (Hoshiarpur, Punjab), Panjab University, Chandigarh, India*
[1]*ggurpinder_singh1989@yahoo.com*

***Abstract*— Grid computing is an idea to use vast and sharable resources on demand, economically and efficiently from around the world. In grid computing companies or institutions can share their resources to solve complex problems. Grid computing is one of technologies that being used for high speed computing for large problems that can't be handled by single infrastructure. This paper presents a state-of-the-art review of architecture and issues of Grid computing.**

***Keywords*— NTP, GRIS, GIDS, X.509, GIIS, MDS, LDAP, QOS.**

I. INTRODUCTION

The ideas of the grid were brought together by Ian Foster, Carl Kesselman, and Steve Tuecke, widely regarded as the "fathers of the grid". They led the effort to create the Globus Toolkit incorporating not just computation management but also storage management, security provisioning, data movement, monitoring, and a toolkit for developing additional services based on the same infrastructure, including agreement negotiation, notification mechanisms, trigger services, and information aggregation. While the Globus Toolkit remains the de facto standard for building grid solutions, a number of other tools have been built for grid services needed to create an enterprise or global grid [7]. As we can see tremendous growth of internet now a days, that changes our society and other research activity in world. The new technologies that are used to cluster resources geographically with low cost and without any hurdle physically or logically are available and further developed, one of which is called Grid Computing. Grid is a parallel and distributed system in which resources spread across multiple administrative domains are able to select, share and integrate based on common rules they accept. Information about the resources hosted by multiple administrative organizations and other networked users is heterogeneous in machine architecture, software & operating system used and rules that gets integrated in Grid [5].

Grid computing appears to be a promising trend for three reasons [2]:

- Its ability to make more cost effective use of a given amount of computer resources.
- As a way to solve problems that can't be approached without an enormous amount of computing power.
- This suggests that the resources of many computers can be cooperatively and perhaps synergistically harnessed and managed as collaboration toward a common objective.

Larger bodies of scientific and engineering applications stands to benefit from Grid Computing, including molecular biology, weather forecasting, financial and mechanical modelling, immunology, circuit simulation, aircraft design, fluid mechanics, biophysics, biochemistry, biology, scientific instrumentation, drug design , high energy physics, data mining, financial analysis, nuclear simulations, material science, chemical engineering, environmental studies, climate modelling, neuroscience/brain activity analysis, structural analysis, mechanical CAD/CAM, and astrophysics.[2]

Fig 1: GRID COMPUTING [5]

II. Small Issues Affecting Grid Computing

A. Security

Security becomes an important factor in the grid computing. If the security of grid is weak then nobody will off course like to use the grid, as its all data will be on high risk it could be hacked and could be use for illegal purposes. So the grid will make sure that if any information is been transferred it should have protection from viruses, access of information will be given to correct identifier, if data changes by mistake then the secure backup will be available in case to recover the data. The record of sent and received data should be kept so in case of any problems the record will be available. [1]

B. Space

If the space factor is considered important in grid computing works for increasing the performance of the network. But in simple as all the space of the machines is not used there is sometimes free space left in the hard disk. Grid computing merges all the free space of those machines so the network becomes more powerful as compared to single machine. This makes the network to execute applications containing huge amount of data as they are difficult to tackle on single machine. [1]

C. Distributed environment

Some applications in Grid computing could consume more time because of distributed environment e.g. A complex problem whose some parts of a task are distributed among various machines and we know that different machines have different processing speed, memory and extra for handling data. If a machine has less processing speed and others depends on it's for their task then they have to wait which means waste of time.[1]

Another issue regarding distributed environment is, some machines are not being used completely. It means that the processor is only used less than 20% and thus it remains idle mostly. The grid computing makes those machines perform more work and it decreases the time of those machines to remain idle, so more work is taken through grid computing. [1]

III. APPLICATION

A. Supply chain management

Monitoring of goods and other services, application is proposed. We can use sensor Grid (use wireless and wired device and communicate with each for sensing different environments) with each other in supply chain management to optimizing service for customized orders. It keeps track of expensive customized orders. As the goods are unloaded, each item can be sensed by sensors and easily located within the warehouse. [14]

B. Disaster management

The natural disasters like earthquake, flooding, volcanoes, Tsunami, wildfire fighting, etc. cause a serious destruction. As a preservation measure one can set up a sensor Grid that consists of wireless sensors and wired sensors. The properly installed and maintained sensor grid can provide an advance warning of future disaster. The data-driven forest fire simulation presented in predicts the behaviour and spread of wildfires along with detailed information of intensity, propagation speed, and direction. The work considers both dynamic and static environmental conditions. [3]

C. Health care services

Mobile ad hoc grids can be deployed to provide healthcare services during emergency. If a man is serious because of a road accident, temporary ad hoc grid can be created using sensors (embedded at patient body), mobile/PDA, or mobile devices that collaboratively collect, distribute the data to central health care server where it is processed, and obtain an instantaneous suggestion on first aid or pre hospital treatments [14]. Hence it supports continuous monitoring of a patient. The sensor grids deployed on the battlefield as a means of detecting enemy troop movements by the vibrations as they walk or drive through the unseen grid. Sensor networks are currently in use monitoring environmental factors such as temperature or humidity change, motion and light intensity. [3]

D. Bioinformatics

The bioinformatics research area is now faced with a mountain of ever-increasing and distributed information. For example, finding a single gene of the Oryza sativa (rice) genome one must spend weeks, if not months, wandering through approximately 40 million base pairs. These data are scattered in many data repositories. Thus, not only do we need an efficient tool to visualize and analyse DNA data, but the integration and exchange of information on a particular gene or coding regions from different international collaborative databases needs to be done in a careful, but robust manner as well. The emerging grid computing technologies enable bioinformatics scientists to conduct their researches in a virtual laboratory, in which they share public databases, computational tools as well as their analysis workflows. The implementation of existing bioinformatics applications on Grids represent a cost-effective alternative for addressing highly resource-demanding and data intensive bioinformatics tasks .[4]

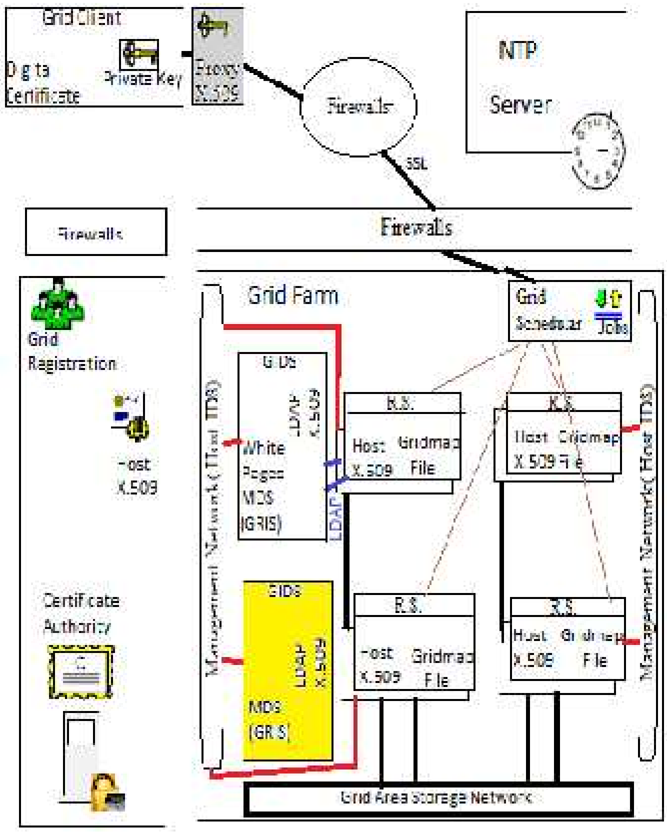
IV. Architecture of Grid

The infrastructure represents the physical hardware and software components used to interconnect different grid computers. These components help support the flow of information between grid systems and provide the basic set of services for connectivity, security, performance availability, and management. While many of these infrastructure components supply basic functionality to the grid, many are optional. It will be up to you to decide on the requirements and how well these components match up to the needs of your design. Architecture of Grid Computing is as with explanation of how we work in Grid environment [8]:

- Single security provider
- Basic set of Globus Toolkit components
    1. Grid clients
    2. Certificate Authority

3. MDS (GRIS,GIIS)
4. Digital certificates for authentication(X.509)

- Grid Scheduler (job scheduler)
- Common infrastructure devices
    1. Management network(GIDS, Resource Management)
    2. Firewalls
- Storage Area Network
- Network Time Protocol Server



R.S. = Resource Management
GIDS= Grid Information and Directory Services
LDAP= Lightweight Directory Access Protocol
MDS= Monitoring and Discovery Service
GRIS= Grid Resource Information Service

Fig 2: Architecture of Grid Computing [8]

A. Security

The Globus Toolkit GSI C component provides APIs and tools for authentication, authorization and certificate management. The authentication API is built using Public Key Infrastructure (PKI) technologies, e.g. X.509 Certificates and TLS. In addition to authentication it features a delegation mechanism based upon X.509 Proxy Certificates. Authorization support takes the form of a couple of APIs. The first provides a generic authorization API that allows callouts to perform access control based on the client's credentials (i.e. the X.509 certificate chain). The second provides a simple access control list that maps authorized remote entities to local (system) user names. The use of firewalls can provide logical and secure segmentation between grid systems [12]. GSI-OpenSSH (open Secure Shell) is a modified version of OpenSSH that adds support for X.509 proxy certificate authentication and delegation, providing a single sign-on remote login and file transfer service. GSI-OpenSSH can be used to login to remote systems and transfer files between systems without entering a password, relying instead on a valid proxy credential for authentication [13]. By using firewalls within your grid design, you can help limit the network communication between grid systems and only use techniques that you specify that the firewall will support. Firewalls work by controlling access to network services that grid computers will be running. Since the network offers a gateway to grid systems, we want to make sure that we control exactly the services and from where to whom on your network. [8]

Some areas you may want to protect within your design are [8]:

- Certificate Authority/Registrant Authority e.g. In the X.509 system, a certification authority issues a certificate binding a public key to a particular distinguished name, or to an alternative name such as an e-mail address or a DNS-entry[10] and Secure Sockets Layer (SSL) which negotiate a stateful connection by using a handshaking procedure. When SSL handshake is complete and the session begins. The client and the server use the session keys to encrypt and decrypt the data they send to each other and to validate its integrity.[9]
- Globus Toolkit components such as MDS, GRIS, and GIIS(explain in next section)
- All grid servers

B. Grid Scheduler

An important function of Grid System is Grid scheduling and load balancing. Here we explain some Schedulers with examples that have the ability to be integrated with some of the products. For example, the Globus Toolkit can be integrated with many schedulers including PBS, Condor[8]. Schedulers are discussed as following:

1) *Condor*

It is workload management software that is developed by a research team located at the University of Wisconsin in Madison. There are three basic components in a Condor: A central manager, execution hosts, and submission hosts and a checkpoint server are an optional fourth component. The central manager serves two main functions in a condor cluster. The first function is collecting the status of all of the nodes in a Condor cluster. The second function is to match up resource requests for job submissions with a Condor node that can fulfill these requirements. Any node in a Condor cluster can be configured to be an execution machine and a submission machine, including the central manager. Execution machines are those nodes that can run Condor jobs and submission machines are those machines where Condor jobs can be submitted. An optional checkpoint server can be added to a cluster to store all of the checkpoint files for the jobs in the cluster. [8]

2) *PBS*

Portable Batch System (PBS) is a scheduler that was developed for NASA Ames Research Center by Veridian. There is an open source version available, OpenPBS, and a commercial version available, PBSPro. There are three main components that make up a PBS: Job server, Job executor, Job scheduler.

- The job server is responsible for all batch processes in PBS, including creation, modification, running, and monitoring of batch jobs. A PBS cluster contains only one server machine.
- The job executer is the machine in the cluster where a batch job is running.
- The job scheduler is responsible for scheduling jobs on the cluster; it can query the status of execution nodes and also query the server to determine what jobs need to be run.[8]

C. Grid Information and Directory Services (GIDS)

The information services are an effective way for resources within the grid to cope with the dynamic nature of the grid. Within any grid, both CPU and data resources will fluctuate, depending on their availability to process and share data. As resources become free from jobs within the grid, they can update their status within the grid information services. This provide clients information to make intelligent decisions on which grid resources are free to use [8].Following are the components of GIDS:

1) *Lightweight Directory Access Protocol (LDAP)*

LDAP[11] is used to access directory for information. To understand the role of LDAP in the Globus toolkit, we must first define the necessary terminology (Directory, Directory service, and LDAP) [11].

- **Directory:** Directories are used to store and retrieve information. Thus, directories are similar to databases. Special characteristics of directories include following:
  1. Directories are designed for reading more than for writing.
  2. Directories offer a static view of the data.
  3. Updates in directories are simple without transactions.
- **Directory Service:** A Directory Service provides a directory that can be accessed via a network protocol. Often, directory services include mechanisms for replication and data distribution. An example of a directory service is Domain Name System (DNS), which resolves the names of computers to appropriate addresses. Programs using this service are countless: for example mail.
- **LDAP:** The abbreviation LDAP stands for *lightweight directory access protocol*. LDAP defines a standard directory protocol that includes the following features:
  1. A network protocol for accessing information in the directory.
  2. An information model defining the form and character of the information.
  3. A namespace defining how information is referenced and organized.
  4. An emerging distributed operation model defining how data may be distributed and referenced (LDAP version 3).

The information model and namespace are based on entries. An entry is used to store attributes. An attribute has an associated type and can have one or more values. Each entry in the namespace has a distinguished name that allows easy identification. The access of the data in an LDAP-based directory is accomplished by using the following:

1. Base DN: This indicates where in the hierarchy to begin the search.
2. Filter: This specifies attribute types, assertion values, and matching criteria.
3. Scope: This indicates how many levels of the directory tree are searched, relative to the base DN. [11]

2) *White Pages*

The information service associated with a Grid environment must have the functionality of a White pages directory. In this we look up the IP number, amount of memory, and so forth, associated with a particular machine. [11]

3) *Monitoring and Discovery Service (MDS)*

MDS [8] provides access to static and dynamic information of resources. Basically, it contains the following components [8]:

- **Resource information:** Resource information contains the objects managed by MDS, which represent components resources are Infrastructure components e.g. name of the job manager or name of the running job and Computer resources e.g. network interface, IP address, or memory size.

- **Grid Resource Information Service (GRIS):** GRIS is the repository of local resource information derived from information providers. GRIS is able to register its information with a GIIS, but GRIS itself does not receive registration requests. The local information maintained by GRIS is updated when requested, and cached for a period of time known as the time-to-live (TTL). If no request for the information is received by GRIS, the information will time out and be deleted. If a later request for the information is received, GRIS will call the relevant information provider(s) to retrieve the latest information.
- **Grid Index Information Service (GIIS):** GIIS is the repository that contains indexes of resource information registered by the GRIS and other GIISs. It can be seen as a grid wide information server. GIIS has a hierarchical mechanism, like DNS, and each GIIS has its own name. This means client users can specify the name of a GIIS node to search for information.
- **Information providers:** The information providers translate the properties and status of local resources to the format defined in the schema and configuration files. In order to add your own resource to be used by MDS, you must create specific information providers to translate the properties and status to GRIS.
- **MDS client:** The MDS client is based on the LDAP client command, **ldapsearch**. A search for a resource information that you want in your grid environment is initially performed by the MDS client.[8]

D. *Resource Management*

The activity of GRAM based on *QoS(Quality of Service)* exists in the process of service application, service execution and service close. At the stage of service application, the resource allocation management system will define user's resource demands according to specific grid service, and convert user's *QoS* demands to particular grid *QoS* parameters, and then it will take these *QoS* parameters as constrained factors to search for available resources that satisfy the requirement in grid. In the process of searching, systems maybe negotiates with the user, then get final result: not being able to supply, being able to supply or reducing *QoS* demands to supply. If the negotiation is successful finally, it means that service provider can provide resource with final *QoS* demands, and then system can carry out the admission control according to specific strategy. After that, it reserves the resource and sent the grid job to resource waiting pool, waiting to be scheduled and executed. If not successful, system should inform the user and terminate the application. At the executive stage of service, the resource allocation management system will monitor the resource which at the reserved state, and renew the relevant *QoS* information of them. Work flow of *GRAM with QOS* as follows [15]:

i. Grid users sign-on the grid system.
ii. Grid users inquire and apply for grid service within the Grid using GRIS.
iii. System confirms user's resource demands according to the specific grid service.
iv. The user puts forward QoS demand.
v. The QoS Mapping & Converting module implements the mapping conversion from user's QoS demand to particular QoS parameters.
vi. System chooses logical resources according to that needed by grid service in *Grid Resource Information Service Centre*. System assures that selected logical resources can satisfy user’s QoS demands. If not finding resource that can satisfy the demands, the Error Process module would inform the user and terminate the applying.
vii. According to the result from step **vi**, systems inquire about the information of physical resources according the mapping relation between logical resource and physical resource.
viii. Through the QoS Negotiation module in Grid Resource Node, system judges whether all QOS can satisfy the user’s demands. When presenting resources cannot fully satisfy the user’s demand, QoS Negotiation module should interact with relevant modules and inquire whether the user can reduce QoS demand, working flow go back to step **vi**.
ix. Then the QoS Admission Control module implements affirmation work requested by grid job.
x. The Trade Server module determines the using price of resource and records the information about the used resource.
xi. The Resource Reservation module sets up resource reservation flag and record QoS demands.
xii. Grid job enters Waiting-job Pool. The Waiting-job Pool module takes charge of dynamic adjusts the priority of grid jobs in Waiting-job Pool.
xiii. The Resource Monitor module monitors the state of reserved resources.
xiv. The Scheduler schedules local jobs and grid jobs in Waiting-job Pool according to particular strategy.
xv. The QoS Control module adjusts QoS parameters according to the result of QoS negotiation and reacts the requirement from the Grid Resource Information Service module and renews its information about resource state.[15]

E. *Network Time Protocol Server*

NTP is a TCP/IP protocol for synchronizing time over a network for server. Basically a client requests the current time from a server, and uses it to set its own clock. Behind this simple description, there is a lot of complexity - there are tiers of NTP servers, with the tier one NTP servers connected to atomic clocks, and tier two and three servers spreading the load of actually handling requests across the Internet. Client software is a lot more complex, it has to factor out communication delays, and adjust the time in a way that does not upset all the other processes that run on the server [16]. NTP client/server model a client sends an NTP message to one or more servers and processes the replies as received. The server interchanges addresses and ports, overwrites certain fields in the message, recalculates the checksum and returns the message immediately. Information included in the NTP message allows the client to determine the server time with respect to local time and adjust the logical clock accordingly. In addition, the message includes information to calculate the the expected timekeeping accuracy and reliability, thus select the best from possibly several servers [17].

F. Storage Area Network: Storage Mode

Storage systems use the following modes of storage: file system, object-based, or block-based. A *file system* typically organizes files into a hierarchical tree structure. Each level of the hierarchy normally contains zero or more directories, each with zero or more files. A file system may also be flat or use some other organizing principle. An *object-based* storage mode as one that stores discrete chunks of data (e.g., IP datagram or another type of aggregation useful to an application) without a pre-defined hierarchy or meta-structure. A block-based storage mode as one that stores a raw sequence of bytes, with a client being able to read and/or write data at offsets within that sequence. Data is typically accessed in blocks for efficiency. A common example for this storage mode is raw access to a hard disk. Storage systems typically allow a user, content owner, or some other entity to define the access policies for the Network storage system. The Network storage system then checks the authorization of a user before it stores or retrieves content. There are three types of access control authorization: public-unrestricted, public- restricted, and private. "Public-unrestricted" refers to content on a Grid network storage system that is widely available to all clients (i.e., without restrictions). An example is accessing Wikipedia on the Web, or anonymous access to FTP sites. "Public-restricted" refers to content on an network storage system that is available to a restricted (though still potentially large) set of clients, but that does not require any confidential credentials from the client. An example is some content (e.g., a TV show episode) on the Internet that can only be viewable in selected countries or networks (i.e., white/black lists or black-out areas). "Private" refers to content on a network storage system that is only made available to one or more clients presenting the required confidential credentials (e.g., password or key). [18]

V. Conclusion

Grid Computing is the concept of sharing the idle resources leads to vast computing and storage pool, which gives us the power or ability for complex task execution with resources at different locations. The purpose of this paper is to explain the conceptual architecture of Grid Computing and issues for understanding the Grid Computing concept for common users. Architecture explains the mechanism of connecting users to different resources for job execution using services e.g. management of resources, information, services related to jobs provided by Globus toolkit project. Future researches are focused on enhancement of secured communication, resource management and economic models in Grid Computing.

References

1. Muhammad Naqaash et. Al,"Grid Computing Used For Next Generation High Speed Processing Technology", International Journal on Computer Science and Engineering(IJCSE) ,Vol. 02, No.05, 2010, 1926-1933.
2. "A Study on Applications of Grid Computing in Bioinformatics", IJCA Special Issue on "Computer Aided Soft Computing Techniques for Imaging and Biomedical Applications" CASCT, 2010.
3. S. S. Manvi , M. N. Birje, "A Review on Wireless Grid Computing", International Journal of Computer and Electrical Engineering, Vol. 2, No. 3, June, 2010,1793-8163.
4. Manjula.K.A, Dr.G.Raju," A Study on Applications of Grid Computing in Bioinformatics". IJCA Special Issue on "Computer Aided Soft Computing Techniques for Imaging and Biomedical Applications"CASCT, 2010.
5. http://www.adarshpatil.com/newsite/research.htmL
6. http://www.gridcafe.org/grid-in-30-sec.html
7. https://en.wikipedia.org/wiki/Grid_computing
8. Introduction to Grid Computing with Globus, September 2003, ibm.com/redbooks. http://www.redbooks.ibm.com/redbooks/pdfs/sg246895.pdf
9. http://en.wikipedia.org/wiki/Secure_Sockets_Layer
10. http://en.wikipedia.org/wiki/X.509
11. http://www.globus.org/alliance/publications/papers.php #Information Services Components, Usage of LDAP in Globus
12. http://www.globus.org/toolkit/docs/latest stable/gsic/#gsic
13. http://www.globus.org/toolkit/security/gsiopenssh/
14. Gaynor, Mark, Moulton, Steven, Integrating Wireless Sensor Nets with the Grid , IEEE Internet Computing, July-August 2004
15. Xiaozhi Wang, Junzhou Luo," Architecture of Grid Resource Allocation Management Based on QoS", The second international workshop on Grid and cooperative computing, Dec. 2003,Shanghai,china.
16. https://help.ubuntu.com/lts/serverguide/NTP.html
17. http://tools.ietf.org/html/rfc1059
18. http://tools.ietf.org/html/rfc6392