International Journal of Information and Computation Technology.
ISSN 0974-2239 Volume 3, Number 5 (2013), pp. 445-450

http://www. irphouse.com /ijict.htm

A Review on Grid Computing

Keshab Nath

Department of Information Technology, Assam University, Silchar-788011, India.

Abstract

Internet technology is too much popular now a day and due to the availability of high performance computer and network technology at low cost change the view of using computer and internet.
Grid computing is conceptually not like electric grids. In electric grid we can just link to the outlets of an infrastructure and we don't need to know from where and how we are getting electricity. Grid computing led to the possibility of using distributed computers as a single large virtual network that allow sharing and computer power and data storage capacity over the internet.
This paper gives an idea about grid definition, its security challenges and issues. It covers about grid characteristics, types of grids, grid middleware. It gives an overview of grid security followed by challenges in grid.

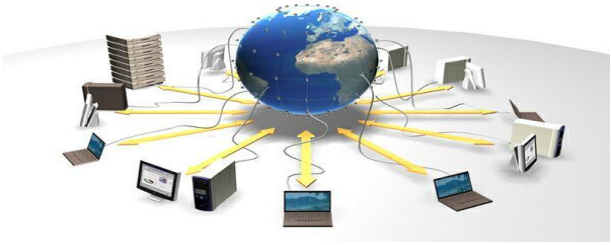
1. Introduction

Grid computing is a computing technology where the CPU cycles and other resources of an idle host in a network, that are not currently using by the system can be efficiently used by connecting those idle host in to a single network to create a virtual system for resource sharing dynamically at runtime

The basic requirements of a grid system is that it can provide the high-level quality of service and access via "The grid" to remote resources of millions of pc's from anywhere to anybody. Grid computing is differing from the Web. In web we can access document by simply giving the IP address or URL of the resource holder but in case of grid we have access to computational resources.

1.1 Grid Computing

Grid computing is a very large virtual network system which allows users to access the computing resources such as database, processor of many different machines distributed around the world. For an example suppose there are 10 pc's and out of them 5 pc's are idle and 5 pc's are not. So the main idea is that we can use the whole CPU cycle of the 5 idle pc's and if the other 5 pc's which are processing some task may or may not use their 100% CPU cycle so we can also use their wasted CPU cycle to perform a large task in combine.



2. Grid Architecture Model [11][13]

Grid architecture is often described in terms of "layers" and each layer performs specific function. Grid architecture can be defined in terms of hourglass model.

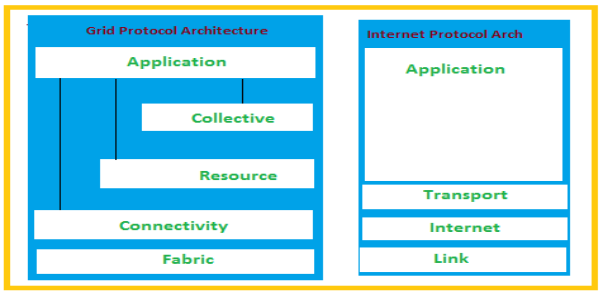


Figure 1: Grid Architecture Model.

The grid architecture deals with the function of each component and their interaction with each other .It establish a protocol between grid user and grid resources for sharing resources, database.

Figure 1 represents the grid architecture consisting several layer and the capabilities of each layer. Each layer shares the behavior of the underlying component layers.

- ***Fabric layer***—The fabric layer provides connection interface on all possible kinds of resources available like computational resources, storage, sensor etc.

- ***Connectivity layer***—The connectivity layer defines basic communication and authentication protocol. Communication protocols provide exchange of data between fabric layer resources. Authentication protocol provide cryptographically secure mechanism for authentication between user and grid resources
- ***Resource layer***—This layer provide grid resources like computers, storage system electronic data catalogues which are connected to the network. The resource layer calls the fabric layer functions to access and control local resources.
- ***Collective layer***—The main purpose of collective layer is to schedule multiple resources, monitoring and file replication services. It's the key layer of the whole architecture
- ***Application layer***—Application layer enables the use of resources in a grid environment through various collaboration and resource access protocols. This is the layer that grid user "see" and interacts with.

3. Grid Middleware:

Grid middleware is nothing but software that organized and integrates the resources in a grid, in short we can say that it is system software between the application and the operating system. It provides services from grid resources to grid user.

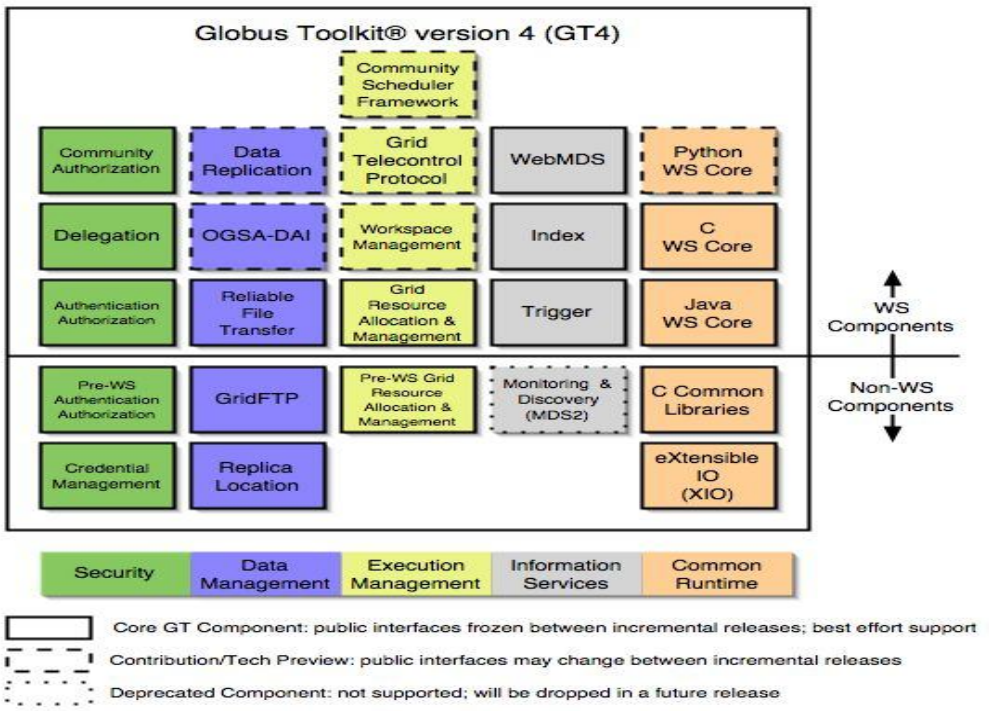
Middleware is basically consisting of "resource and connectivity layer" and higher layer of "collective services". Resource layer and connectivity layer specially deals with sharing resources between computers and grid resources.

Resource layer define protocols for the secure negotiation, monitoring, sharing etc. Basically there are two primary classes of resource layer protocol-information and management protocols.

Information protocol gives the information about the state and structure of a resource and management protocols used to negotiate the access to grid resources.

4. Globus Toolkit 4: [12][14]

The Globus Toolkit of the Globus-Alliance is a Middleware for Grid Systems. Globus Toolkit comes with set of tools for constructing a grid, services such as resource access, data management and security infrastructure. Using globus toolkit we can perform large dataset job on different machine that are geographically far apart and may be owned by different organization at the same time. It facilitates the scientist to deal with large dataset and complex remote collaboration.



Picture: Componentent of the Globus Toolkit 4

Globus includes following components-

GRAM (*Globus Resource Allocation Manager*):GRAM provides remote control function like remotely submission of job and controlling them. It provides a GUI for submission and execution of jobs remotely.

GSI (Grid Security Infrastructure): authenticates users and determines their access rights.

MDS (*Monitoring and Discovery Service*): collects information about resources such as status, processing capacity, bandwidth capacity, type of storage, and so on. **GRIS** (*Grid Resource Information Service*): queries resources for their current configuration, capabilities, and status.

GIIS (*Grid Index Information Service*): coordinates arbitrary GRIS services.

GridFTP (*Grid File Transfer Protocol*): provides a high-performance, secure, and robust data transfer mechanism.

Replica Catalog: provides the location of replicas of a given dataset on a grid.

The **Replica Management system**: manages the Replica Catalog and GridFTP, allowing applications to create and manage replicas of large datasets.

5. Grid Security:

Grid operation need to be very secure, because if globus isn't secure then no one willing to share their data. So Globus security is mainly based on authentication and authorization.

Authentication of every user is done by providing an grid certificate, that uniquely identified each and every user. Whenever a user tries to access grid resources, he needs to send his grid certificate along with his access request command to the target computer system. So the target computer can easily recognise who has issued the request.

After authenticating the user, we need to authorize that particular user, authorization means to check if that particular user has the permission to access the resources which he asking for access.

6. Grid Computing Challenges and Issues:

Till now there is no clear standard for grid computing, It may or may not possible to use different operating system and different grid middleware at the same machine in the same time .

Still there is lots of debate on concept of grid computing. There is always confusion between distributed computing Vs grid computing.Grid development is too difficult task due to a single person need to know lots of programming language.

From architecture point of view grid should access only by trusted one. There should be an access policy so that user data can be protected from unauthorized access. Need separate access policy for accessing both personal data and system data. Security policy should be there for each and every resource level. Issue related to network and client component infrastructure.

No user should not wait for a very long period for getting services i,e there should not be any job starvation. Local jobs should get higher priority than system jobs. Resources need to distribute globally so that it will be available to any body from anywhere. Managing grid is too much difficult due to the heterogeneous nature of the grid.

References

[1] Browne, J.C. "*Grid computing as applied distributed computation: a graduate seminar on Internet and Grid computing***",** Cluster Computing and the Grid, 2004. CCGrid 2004. IEEE Int'l Symposium, 19-22 April 2004, pp 239 – 245.

[2] Irving, M; Taylor, G.; Hobson and P."*Plug in to Grid Computing***",** *Power and Energy Magazine*, IEEE, Volume 2, Issue 2, Mar-Apr 2004, pp 40 – 44.

[3] Dai, Y.S.;Xie,M.; and Poh, K.L. "*Reliability Analysis of Grid Computing Systems***"** Dependable Computing, 2002. Proceedings. 2002 Pacific Rim International Symposium, 16-18 Dec. 2002, pp: 97 – 104. [4] A Networking Approach to Grid Computing , pp 1-100, by Daniel Minoli, John Wiley & Sons.

[4] Oracle 10g: Infrastructure for Grid Computing, Oracle Corporation (2003)

[5] Oracle and the Grid, Oracle Corporation http://otn.oracle.com/products/oracle9i/grid_computing/OracleGridWP.pdf
[6] The Grid, Oracle Magazine, Volume 17 Issue 5http:/otn.oracle.com/tech/grid/collateralGrid_Feature.pdf
[7] http://ir.canterbury.ac.nz/bitstream/10092/3060/1/12618906_12618906_Pawlikowski.pdf.
[8] Grids Portals: Frameworks, Middleware or Toolkit by Xavier Medianero-Pasco, Belén Bonilla-Morales Miguel Vargas-Lombardo/ IJCSI International Journal of Computer Science Issues, Vol. 7, Issue 3, No 4, May 2010.
[9] http://www.cdac.in/html/events/betatest/archives/gripsi-2007/gripsi-2007-gridcomp-overview.html.
[10] http://www.instant-grid.org/?q=en/node/323.
[11] http://www.gridcafe.org/EN/globus-toolkit.html.
[12] Grid Computing/KavitaAgrawal/Department of Information Technology, Ajay Kumar Garg Engineering College, P.O. Adhyatmic Nagar, Ghaziabad 201009.
[13] http://gdp.globus.org/gt4tutorial/multiplehtml/ch01s04.html