Grid Computing: A Paradigm for Virtual Collaboration and Resource Sharing in a Global Context

Daramola, J.O. and Ayo, C.K.*

Abstract

The advent of globalization has facilitated increased interdependence, integration between people and organizations in disparate locations. Grid computing which is a relatively new and evolving computing paradigm that facilitates coordinated resource sharing and collaborative problem solving is presented in this paper as a tool for improved academic and research collaboration in a global context.

1. Introduction

Globalization is the umbrella term that describes the increasing interdependence, integration and interaction between people and organizations in disparate locations. The term was first used by Theodore Levitt in 1981 to describe the new world order that facilitates increasing ties and links that spans the realms of economy, trade, technology, education, culture and social interaction between nation states, corporate organizations and individual world entities [15].

Generally, globalization is seen as a catalyst for commerce, one which promotes trade, social and economic prosperity among nations, facilitating the transfer of wealth from the First World countries to the developing Third World countries.

However, one of the potential benefits of globalization not yet fully explored is the positive impact it can have on the transfer and spread of technology. As at today, the Internet remains the most dominant technology for facilitating cross-boarder, cross-culture, and cross-regional globalization. Presently, the Internet is being used to foster close interaction, integration and information exchange in such areas as e- business, e-commerce, e-government, e-learning, and distance learning to mention just a few. The core features of the Internet include: electronic mails services, instant messaging, the World Wide Web (for hosting web sites), remote login functions (telnet), and data exchange and content transfer through the File Transfer Protocol (FTP). Other specialized features available on the Internet include web meeting, teleconferencing and so on.

Expectedly, the Internet is not without its shortcomings, which also puts a limit on its capability as a medium for resource sharing and collaboration. Some of the limitations of the internet include: 1) access of individual computers on the internet due to limited bandwidth; 2) involves only remotely located computers and not other specialized devices; and 3) inability to initiate concurrent processes to boost the throughput of individual computing machines. A credible solution to these limitations is offered by the grid computing paradigm which is an extension of distributed computing that facilitates co-ordinated resource sharing and collaboration. It offers new possibilities for interaction and virtual collaboration by allowing resources in disparate locations to be pulled together to solve a common problem.

2. Existing Virtual Collaborative methods

The essence of virtual collaboration is to create an environment for electronic interaction that will facilitate the active participation of a virtual community in pursuit of a common objective. Some of the existing methods range from simple approaches such as web meetings and instant messaging services to more advanced methods such as blackboard systems and groupware technology [2]. The groupware has the most potential for collaborative problem solving because it offers features like document sharing, computer conferencing, e-mails and messaging, group or meeting scheduling, project management, idea organization and shared virtual space team building [2]. One common characteristic of these existing collaborative methods is that the nature of sharing is limited to data resources, unlike grid computing which facilitates coordinated resource sharing of all available resources (hardware, software, data etc.) in a virtual collaboration.

3. Grid Computing

3.1 What is Grid Computing?

A Computational grid is a hardware and software infrastructure that provides dependable, consistent, pervasive and inexpensive access to high-end computational capabilities [3]. Grid computing is concerned with the coordinated resource sharing and problem solving in dynamic, multi-institutional virtual organization [1].

Also, according to [4], grid computing was defined as flexible, secure, coordinated resource sharing among dynamic collections of individuals, institutions, and resources.

Putting together other sundry definitions, it can be defined as a computing paradigm for collaborative problem solving within a virtual organization facilitated through coordinated resource sharing and the rendering of mutual non-trivial services.

3.2 Characteristics of a Grid

The characteristics of a grid system are listed as follows [1], [4]:

- The grid must have the ability to negotiate resource-sharing arrangements among set of participating providers and consumers and to use the resulting resource pool for some purpose.
- The resource sharing must allow direct access to computer, software, data and other resources as is required by a range of collaborative problem solving and

Department of Computer and Information Sciences Covenant University, Nigeria.

resource-brokering strategies emerging in industry, science and engineering.

- The resource sharing is highly controlled with resource providers and consumers defining clearly and carefully just what is shared, who is allowed to share, and the conditions under which sharing occurs.
- All grid transactions occur within a virtual organization i.e. a set of individuals and/or institutions defined by such sharing rules.

3.3 The Grid Architecture

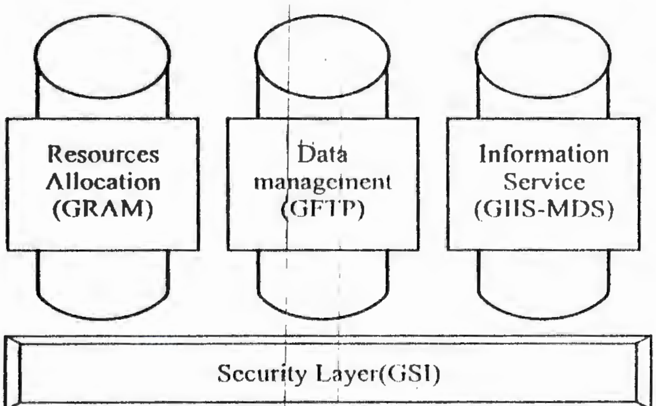
The basic components of grid architecture are:

i. **Grid Security Infrastructure (GSI)**: This provides authentication, single sign-on, delegation through proxy certificates and other security interfaces.

ii. **Grid Resource Allocation Manager (GRAM)**: This handles resource allocation and job management, and file I/O staging provided by the Global Access to Secondary Storage (GASS) interface.

iii. **Monitoring and Discovery Services (MDS)**: This provides information services, it is composed of a Grid Resource Information Service (GRIS) at each resource and of one or more information aggregation services referred to as **Grid Index Information Services** (GIIS)

iv. **Grid File Transfer Protocol (GridFTP)**: This, together with the **Replica Location Service (RLS)** is responsible for File-transfer and File-replica management.



3.4 Key Attributes of a Grid System

The key attributes of a grid system are listed as follows:

- Coordination of resources that are not subject to centralized control: A grid integrates and coordinates resources and users that live within different control domains. For example, the users desktop versus central computing; and different administrative units of the same company or different companies. It also addresses the issues of security, policy, payments, membership and so forth.
- **The use of standard, open, general purpose protocols and interfaces:** A grid is built from multi-purpose protocols and interfaces that address such fundamental issues as authentication, resource discovery and resource Access such as the Open Grid Services Architecture (OGSA) [1].
- Delivery of non-trivial quality of service: A grid allows its constituents resources to be used in coordinated fashion to deliver various qualities of service, relating to response time, through put, availability and security and/or co-allocation of multiple resource types to meet complex user demands, so that the utility of the combined system is significantly greater than that of the sum of its parts.

3.4 Capabilities of the Grid

The sundry capabilities of a grid are summarized as follows:

1) **Exploiting Underutilized Resources**: these could be in the form of 1) running an existing application on different machine; 2) exploiting idle times on other machines; 3) aggregating unused disk drive capacity into much larger virtual storage, thereby boosting performance; 4) creating better balance of resource allocation; and 5) enhancing management view of usage patterns of resources in an organization.

2) **Parallel CPU Capacity**: Grid computing offers potential for massively parallel computation to boost performance in computationally intensive applications e.g. bioinformatics, data mining, financial modeling, oil exploration, motion picture, animation and so on. However, this functionality demands that such algorithm be sufficiently scalable (capable of being be broken into independent running parts).

3) **Virtual Resources and Virtual Organizations**: The grid facilitates collaboration and interoperability of different virtual organizations with the heterogeneous resources available.

4) **Access to Additional Resources**: The grid provides access to other specialized devices [11].

5) **Resource Balancing**: The grid provides healthy resource balancing through activities like handling of occasional peak loads, job prioritization and job scheduling.

6) **Reliability:** Grid technology offers alternate approach to achieving improved reliability through software other than hardware. 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 tampering. Autonomic computing can be utilized such that when there are problems in the grid it can be healed automatically even before the operator or manager is aware of them.

7) **Management:** The existence of grid technology will

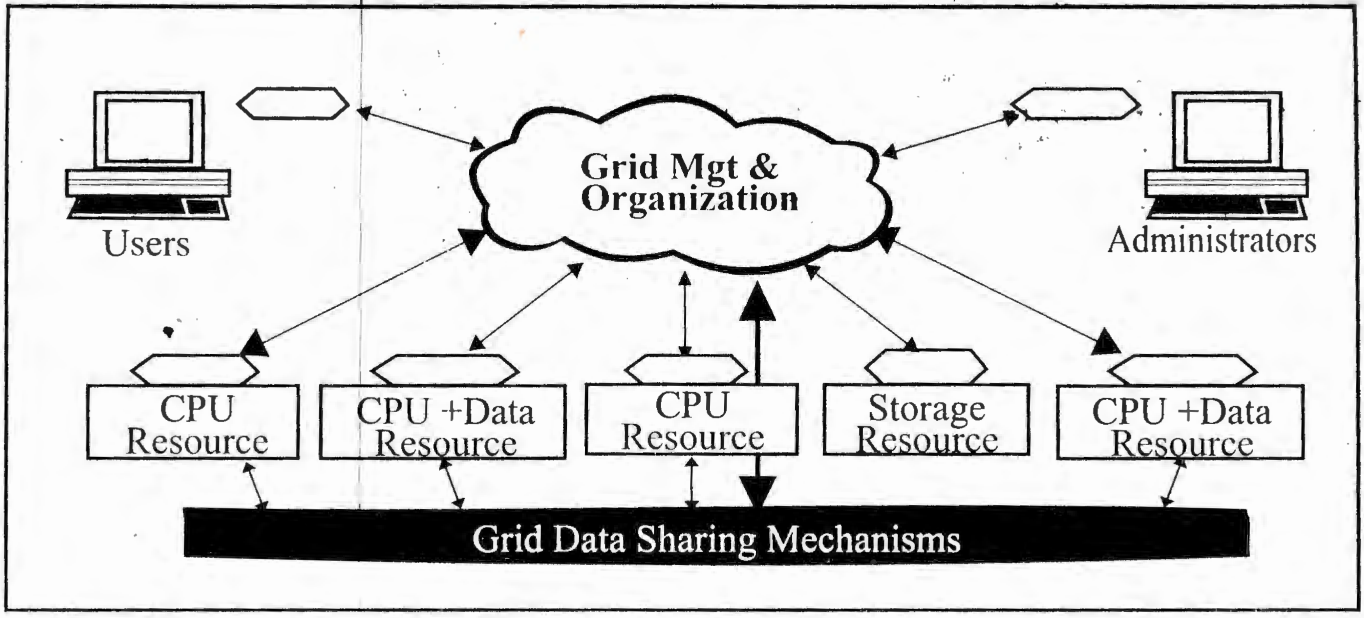


Figure 2: Overview of a Simple Grid [1]

3.6 Grid Topologies

The three main grid topologies are *Infragrid (Cluster), Intragrid and Intergrid*.
The key characteristics of a *Cluster (Infragrid)* are: 1) few machines; 2) homogenous systems; 3) exists within a department of an organization; 4) may not require policies or security concerns and 5) does not normally include scheduling concerns.

An *Intragrid* consists of: 1) heterogeneous systems; 2) have more resources available for sharing; 3) includes some scheduling components; 4) file sharing may still be accomplished through networked file systems; and 5) nodes in the grid may include those for multiple departments but within the same organization.

An *Intergrid* on the other hand has the following characteristics: 1) consists of heterogeneous systems; 2) consists of different organizations forming a virtual family; 3) has more concerns for policy of coordinated sharing, security and scheduling; 4) have security as a major concern; 5) has dedicated grid machines that may be added to increase the quality of service for grid computing, rather than depending entirely on scavenged resources; 6) dedicated communications may have to be provided; and 7) VPN tunneling or other technologies may be used over the internet to connect the different parts of the organization. The intergrid offers the prospect of trading or brokering resources over a much wider audience which makes it an ideal tool for virtual collaboration in global context.

4. The Covenant University Intragrid Model

Grid computing offers enormous positive potentials for virtual collaborative research between Nigerian Universities and other world class institutions, because it offer the opportunity to share remotely located resources that may be too expensive for this part of the world, particularly the developing nations. Covenant University as an academic institution thus becomes an ideal enterprise where the practice of the grid computing paradigm is poised to thrive.

As a first step to achieving this, we hereby present, a model design of a proposed enterprise intragrid for the University consisting of her major departments and units. The implementation of the Intragrid will put the University in vintage position to virtually collaborate with other international organizations through an intergrid arrangement. Some of the anticipated benefits include [10],[13]:

- increased capacity for collaborative problem solving;
- improvement is overall computational throughput;
- the opportunity of leveraging available resources in all units and departments of the University being underutilized;
- coordinated sharing of resources for mutual benefits;
- optimal usage of available hardware resources; and
- effective management and prioritization of resources

The *Covenant Grid* (see Figure 3) as being proposed will be an intragrid of major units and departments of the University

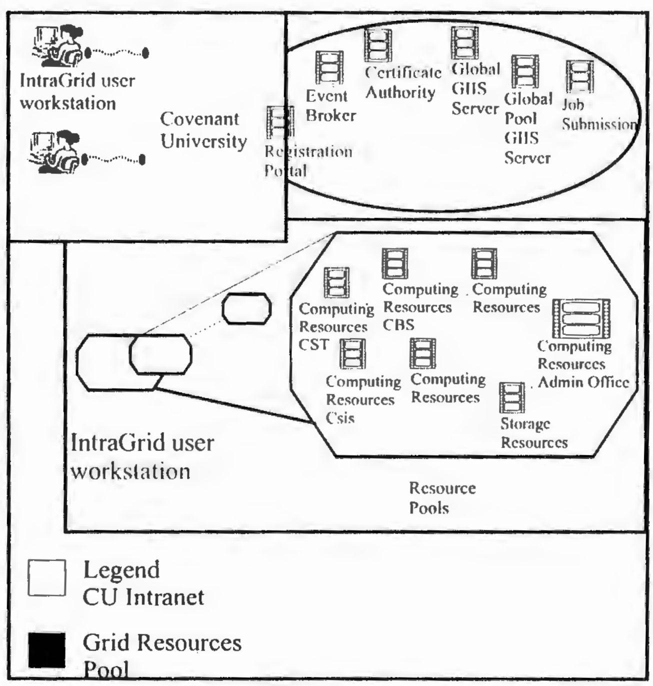


Figure 3: Architecture of the Covenant Grid

built on an expansive University wide intranet, which will offer the platform for several grid capabilities.

5. Conclusion

The potential of the grid computing paradigm as a vehicle for technological globalization is quite enormous. The myriad of existing grid projects all around the world covering so many areas of the Sciences, Engineering, Computer Science, Particle Physics, and Meteorology to mention just a few, is a pointer to this fact [14]. Further more, other extensions of the native grid computing like the mobile grid [5],[7],[8], autonomic grid, semantic grid and cognitive grid [12] are currently being researched. Nigerian and African academic institutions must also endeavor to be a part of this trend, if we must thrive on the positive side of technological globalization.

Generally, grid computing will foster collaboration between institutions in the developing and the developed world, for improved quality of research, improved academic standard, and academic networking through global resource sharing.

References

[1] Berstis, V. (2002): Fundamentals of Grid Computing, IBM Redbooks, www.ibm/redbooks

[2] Doggett, M. (2004): Selected Collaborative Problem-Solving methods for Industry, http://www.nait.org

[3] Foster, I, Kesselman, C. and Tuecke, S. (2001): Anatomy of the GRID: Enabling Scalable Virtual organizations, Proc. Euro-par Parallel Processing, LCNS 2150, Springer-Verlag, 1-4.

[4] Foster, I. (2002): What is the grid? A three point checklist, www.gridtoday.com/02/0722/100136.html

[5] Gaynor, M, Welsh, M, Lacombe, E., Rowan, A, and Wynne, J. (2004): Integrating Wireless Sensor Networks with the Grid, IEEE Internet Computing, Vol. 8. No. 4, 32-39.

[6] Joseph, J., Ernest, M.,Fellenstein, C. (2004): Evolution of grid computing architecture and grid adoption models, IBM System Journal, Vol. 43, No. 4, 624-644

[7] Kyung, H. (2002), Mobile Grid, Real-Time Multimedia Lab., Www.gridforumkorea.org/eng/workshop/20022002_summer/lee_sy.ppt.

[8] McKnight, L.W., Gaynor, M., (2003): "Wireless Grid Issues", Proc. 8th Global Grid Forum (GGF8), www.wirelessgrids.net/docs/draft-ggf-lwmcknight-wgissues-0.pdf

[9] McKnight, L.W., Howison, J., Bradner, S. (2004): Wireless Grids: Distributed Resource Sharing by Mobile, Nomadic, and Fixed Devices, IEEE Internet Computing, Vol. 8, No. 4, 24-31

[10] Meliksetian, D. S., Prost, J.P., Bahl, A. S., Boutbou, I., Currier, D. P., Fibra, S., Girard, J.Y. Kassab, K. M. , Lepesant, J.L., Malone, C., Manesco, P. (2004): Design and implementation of an enterprise grid, IBM System Journal, Vol. 43, No. 4, 645-664

[11] Millard, D., Woukeu, A., Tao, F. B. and Davis, H. (2005): Experiences with Writing Grid Clients for Mobile devices. In Proceedings of 1st International ELeGI Conference on Advanced Technology for Enhanced Learning BCS Electronic Workshops in Computing (eWiC), Vico Equense, (Napoli), Italy.

[12] Parasher, M, and Lee, G.A.(2005): Scanning the special issue on Grid Computing, Proceedings of IEEE Vol. 93, No. 3, 2005.

[13] http: //www.globus.org/toolkit/

[14] http: //www.globusalliance.org

[15] http://www.wikipedia.com, "Globalization"