Peer-to-Peer Computing and Grid Computing: Towards a Better Understanding.

N.A. Azeez[1]; A.P. Abidoye[1]; A.O. Adesina[1]; K.K. Agbele[1]; Iyamu Tiko[2], and I.M. Venter[1]

[1] Department of Computer Science, University of the Western Cape, Private Bag X17, Bellville, 7535, South Africa.
[2] Tshwane University of Technology, Department of Informatics, Faculty of Information and Communication, Soshanguve South, South Africa.

E-mail: nurayhn@yahoo.ca

ABSTRACT

Currently, both Peer-to-Peer Computing (P2P) and Grid Computing have remained the most vibrant and useful forms of distributed computing all over the world. Their applications are such that they cut across both academia and industry. It has come to the notice of researchers that there are great misunderstanding and misinterpretation on what these forms of distributed computing actually portend and stand for. In this paper therefore, we take a critical look at comparative study of both computing technologies with aim of making readers understand in a clear cut what each really stands for. To have a good comparison, we start by giving a well referenced definition of Grid Computing as well as Peer-to-Peer Computing. Also, we used technical issues and general features in our comparison vis-à-vis the architecture, security issue, data movement, application deployment, and operating system requirement. We also considered the strength of both distributed computing system and finally we considered what could be the future of both technologies.

(Keywords: grid computing, peer-to-peer computing, P2P, resources, architecture, middleware, comparison)

INTRODUCTION

Grid and P2P are both performing the duty of resources distribution across distributed communities (Foster and Adriana, 2003). Despite this similar feature, there are many differences in applications, technologies, security, as well as resources involved between these two virtual organizations.

As a result of continuous innovative development through various applications in the distributed system, some challenges do exist. Some of these challenges include attempts to distinguish and clarify the distinction between the available forms of distributed computing. Grid computing and Peer-to-Peer computing are at the forefront of this innovation. Fulfilling the demands of a virtual world remains the most visible current practical use of Grid Computing (Babak et al., 2006).

Results have shown that the rate of adoption of Grid Computing (Domenico & Paolo, 2004) in both industry and academia is at arithmetic progression. It is therefore observed that:

$$\sum_{i=1}^{i=n} {}^{CG}_{U} = a_1X^i + a_2X^{i+1} + a_3X^{i+2} + a_4X^{i+3} + \ldots\ldots\ldots + a_nx^{i+n}$$

The conclusion to the above was drawn from a survey carried out on developing and developed nations where the usage of Grid Computing gains prominence especially at industrial and academic places. In the course of carrying the research the following factors were considered:

${}^{CG}_{U}$ = the Grid computer usage;
a = the type of application;
x = the rate of demand.

At the end of this research, it is observed that as the application for Grid Computing varies and expands as the demand also increases. This expresses the level at which the adoption is increasing. Also, it underscores the need for making comparison between Grid and other forms of High Performance Computing (HPC); P2P.

P2P on the other hand, though is widely adopted mainly for file sharing purpose, the rate of usage in both academia and industry is below the rate of

using Grid computing at those mentioned places (academia and industry).

Due to the ambiguity as well as lack of proper understanding between Grid Computing and Peer-to-Peer Computing (Babak et al., 2006), the need for comparing these two approaches is inevitable. The comparison shall expose the technical related issues concerning both distributed computing systems. This work will also bring into limelight the comparison of general characteristics. What is more, this work shall explain the strength, weaknesses and motivation which will assist for better understanding the approach to be adopted for implementing any real world problems.

To have a sound understanding of the concept, we adopt a practical approach, that is, we compare both Grid Computing and Peer-to-Peer Computing in the actual way they are in real systems not in theoretical perspective.

The only work that is patterned along this line is the work of (Babak et al., 2006), where a comparison was made between Global Computing and Grid Computing.

The paper gives definitions of both technologies; differentiates both P2P and Grid technologies based on general and technical factors; provides security challenges for both; and provides a brief conclusion for the paper.

PEER-PEER COMPUTING AND GRID COMPUTING: DEFINITION

Grid Computing

To have a reasonable and logical comparative evaluation of Grid Computing and Peer-Peer Computing, we need to first of all define the two concepts.

The most acceptable and well referenced definition of Grid Computing was given by (Buyya Rajkumar, 2002) as follow:

The *"Grid is a type of parallel and distributed system that enables the sharing, selection, and aggregation of resources distributed across multiple administrative domains based on their (resources) availability, capability, performance, cost, and users' quality-of-service requirements."* (Buyya Rajkumar, 2002).

Grid Computing is a type of distributed computing which permits and ensures the sharing of aggregated resources across disperse locations (Buyya, 2008). Resources are connected through a required middleware such as Globus, Legion, or gLite with the internet to provide various services for management of resources and security issue (Foster and Kesselman, 1997).

The Grid Computing according to [Foster et al.] can be illustrated with Figure 1. It shows how resources are being shared and how distributed resources are being integrated with general-purpose protocols.

Grid computing is the most recent and most populous computing environment today. It can be regarded as an emergent technology that is being widely accepted for solving various complex problems. Grid computing is being taken and regarded as an extension of distributed computing systems where the number of systems arranged and connected together are now greater than usual. One of the main advantages of Grid computing is that it has the potential and ability of sharing resources as well as direct and monitoring problem solving technique in a dynamic and multi-organizations (Carl Kesselman, 2003).

The concept of Grid system is analogous to the "water Grid system". The facilities of water Grid system make it possible for anyone in his home to open the tap and fetch water without knowing exactly where such water is being processed. The technology behind computing grid is very similar to this (Jacob and Fukui, 2005). Grid computing therefore provides an endless and ubiquitous access to expensive but high quality computing resources to the users wherever they find themselves.

Peer-Peer Computing

P2P is a category of distributed system which optimizes the benefits of resources such as human, storage and cycles that are available on the internet. Implementing P2P requires the creation of overlay networks (Foster and Adriana, 2003).

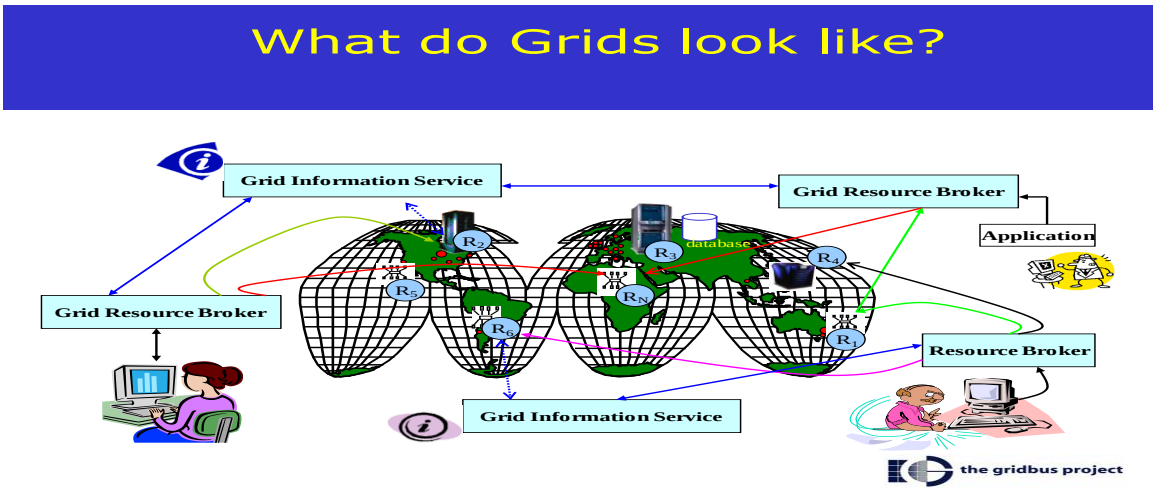

Figure 1: Grid Computing Envirionment.

Peer-to-peer can also be referred to as a communication structure where by each party involved has equal abilities and capabilities and either party can establish a communication process. In most cases, peer-to-peer communications is established by providing client-server capabilities with each of the nodes connected together.

GENERAL AND TECHNICAL FACTORS

At this juncture, we compare both computing technologies with their general qualities; we also take into consideration their technical issues as well.

Grid computing offers services to low-populated cities (Foster and Adriana, 2003) and integrate large amount of resources to sustain high qualities of service within the cities. INDIANA Grid is a common example.

Grid computing are basically dealing with the deployment of sophisticated applications which is usually achieved by aggregating various sites to handle and solve complex scientific applications (Buyya, Grid Computing). SAGrid (Grid, 2010) is a common example where five universities and one institute are collaborating for resource sharing.

However, P2P deals with large amounts of people but offers very few and limited services. It is less concerned about the quality of service being offered, unlike Grid computing (Foster and Adriana, 2003). P2P computing, however, deals with issues and services that are common and popular such as file sharing which are considered as not complex and sophisticated.

Proficiency of Users/Target Audience: Since Grid was initially developed to handle complex scientific problem (Foster and Adriana, 2003) users of Grid computing are therefore required to have certain level of expertise in computing, in some cases they are usually expert in computing, this is because of the sophisticated nature of the technology while the users of Peer-to-Peer can be end users since no professional knowledge and skill are required. It involves activities such as file sharing.

Middlewares: These are open software that provide access for a grid user to effortlessly share, distribute, and aggregate the computing resources on a grid system across multiple administrative domains (Buyya, Grid Computing Info Center (GRID Infoware), 2008). This is one of the requirements that differentiates grid from P2P and cluster computing. Without middlewares, sharing and distribution of resources cannot take place on the grid system. Some of the examples of grid middlewares are PUNCH, GLOBUS, OGSA-DAI, Alchemi, Legion, and gLite (The Gridbus Project, 2005). Grid middleware is very important if Grid is to be established.

This is, however, not required in P2P computing. They are basically used to establish a connection between nodes that are on a grid. P2P on the

other hand uses protocols designed to make available a particular integrated functionality. It has focus for aggregating and assembling limited resources

Application: The application deployed on Grid computing has to do with complex scientific ones while P2P only handles solutions to specialized resource sharing services, File Sharing, Private File Sharing, and Peercasting. The *SETI@home Project* is one of the popular applications of grid computing which is run by SETI (Search for Extraterrestrial Intelligence). This application searches for radio waves to establish the presence of extra-terrestrial intelligence. Other popular applications are: LHC Computing Grid and NFCR Centre for Computational Drug Discovery which was developed at computational grid of Oxford University. Some of the smaller applications of grid computing include weather forecasting and earthquake simulation. Grid systems aggregate resources that are more sophisticated and well connected than a P2P system (Foster and Adriana, 2003).

Security: A reliable and dependable security infrastructure is one of the most important aspects of grid systems (Ellahi, 2008). Whereas, P2P does not require high levels of security as a major concern. This is because P2P has basically being used for file sharing among its users. Grid can be at anywhere. P2P operates within a restricted domain.

Security is very important for any distributed system. In Grid computing, the security issue is a bit simpler. This is because, the participants are fewer and users are group of professionals, hence security is easy to enable and adopt. However, in P2P, since participants are many, therefore it is difficult to put in place adequate security measure. Therefore, there is absence of trust among P2P users (Foster and Adriana, 2003).

Users' Motivation: Achieving a virtual organization's objective implies that user's motivation and determination has been met (Babak et al., 2006) in Grid computing. Hence, in Grid environment, there is no need for tasking a user for optimizing the benefit of Grid or compensate him for using it. In contrast, in P2P computing, the self interest of a user should be placed first.

Architecture: An abstract overlay network is usually implemented in any Peer-to-Peer system which is always established on the Application Layer. In order to make the P2P independent, the use of overlay for indexing cannot be left out. Within a P2P computing, the contents are usually shared through the underlying Internet Protocol (IP) network.

P2P can be categorized into two: *structured* peer-to-peer networks and *Unstructured peer-to-peer* networks. The *Structured peer-to-peer* networks are usually established following adherence to some criteria and algorithms. Efficient routing of search to a desired file is very important issue in P2P. This is achieved in a P2P via a globally consistent protocol. Distributed Hash Table (DHT) is the popular of all type of structured P2P network. A distributed hash table (DHT) provides a mean for any participating node to conveniently and quickly retrieve the value associated with a given key.

Unstructured peer-to-peer networks do not require any algorithm for network establishment.

Grid computing architecture can be divided into four categories. They are:

i. the application layer;
ii. the middleware layer;
iii. computing infrastructure layer and
iv. distributed communication/fabric

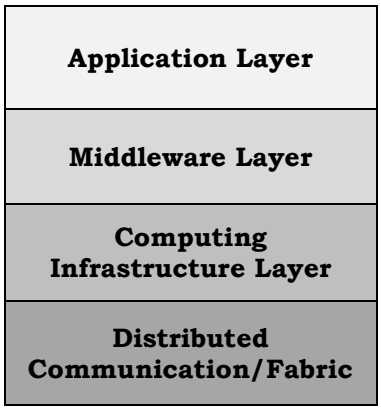

Figure 2: Grid Computing Architecture.

Resource Management: The principal difference between Grid Computing (JATIT, 2005) and P2P lies solely in manner resources are being managed. Resource management in peer to peer is carried out by a centralized resource manager and all connected nodes jointly work as a single entity as shown in the Figure 3. However, in Grid Computing , each node (Jacob and Fukui, 2005) on the grid has its resource manager.

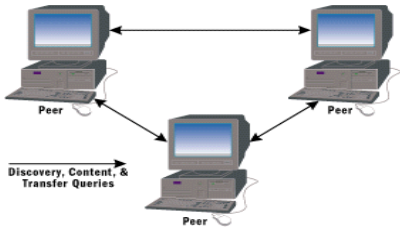

Figure 3: A Pure P2P [3].

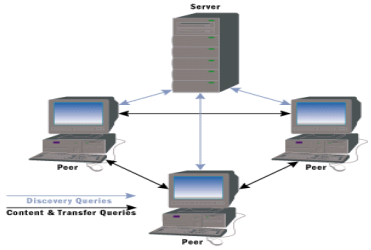

Figure 4: P2P with a Simple Discovery Server [3].

Data Movement: The potential of moving and transporting data from various sites and location remains a crucial matter in the foundation of distributed system. The requirements (Rajkumar, 2009) for data movement between these two technologies are quite different from one another. Some of the factors that determine the movement of data are the protocols and service for data transfer for application deployment, the amount of data to be transferred and the principal actor requesting the transfer of data.

From reliability (Yanxiang, Fei, and Wensheng, 2008) performance as well as security point of view both the FTP and SCP which are considered as simple file transfer mechanisms are not sufficient for file transfer. However, these mechanisms could be used for P2P computing. GridFTP which is a product of Globus middleware remains the only open source quality data mover for Grid Computing.

GridFTP therefore is considered as a high performance, dependable data mover standard protocol optimized for high bandwidth across various administrative domains. This is a recommended data movement protocol for Grid Computing.

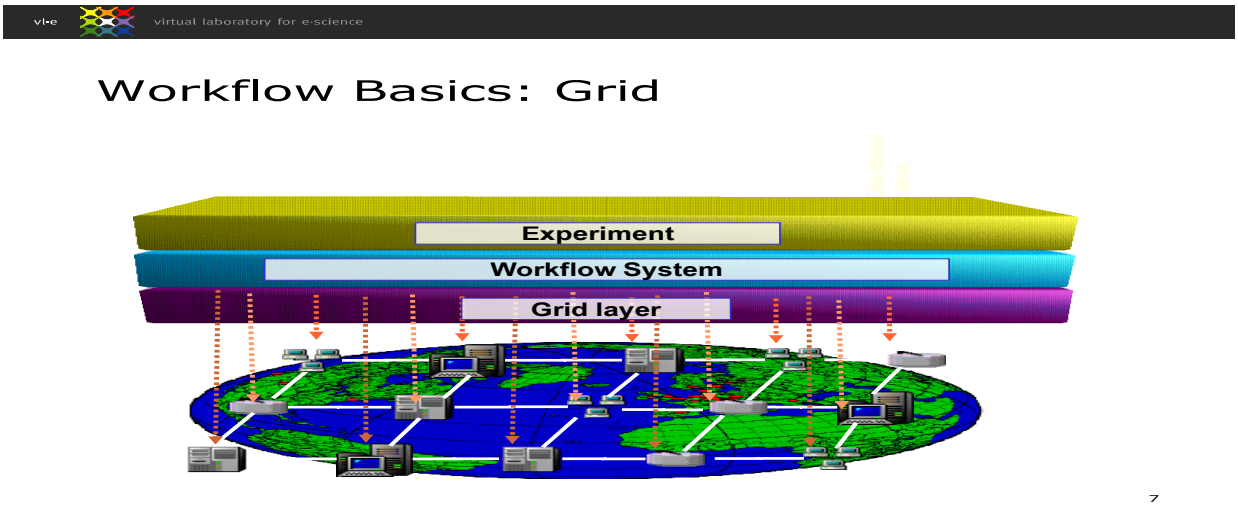

Figure 5: Data Movement in a Grid Based Environment.
Source: (Frank, 2008)

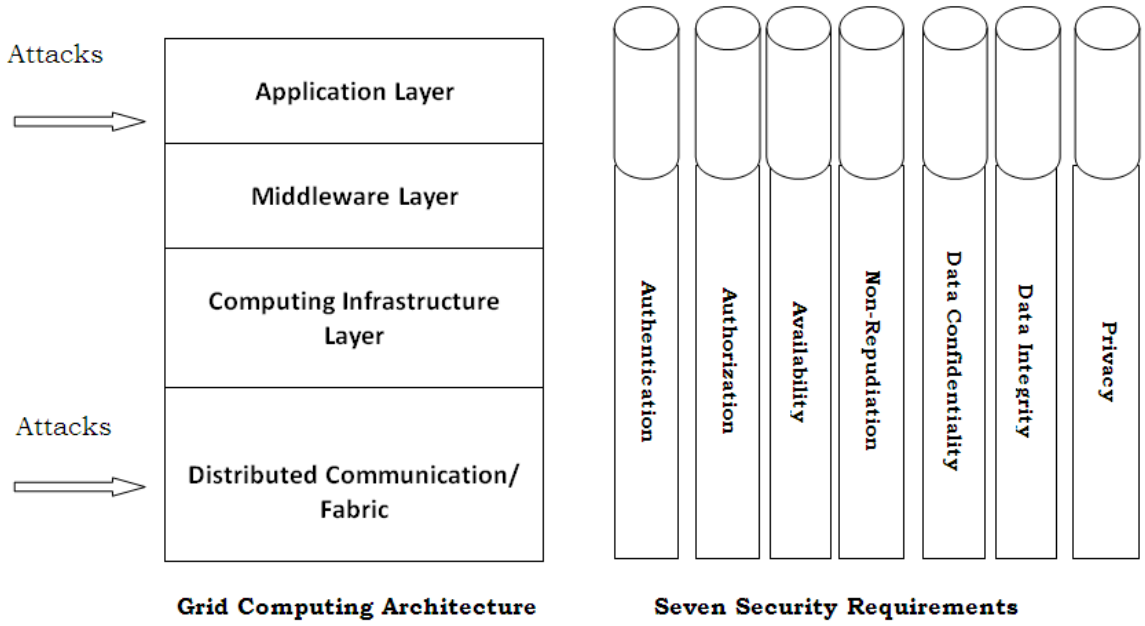

Figure 6: Showing Grid Computing Architecture and their 7 Security Requirement.

Challenge of Grid Computing and P2P Computing

P2P is a type of distributed system whereby each node has duties, responsibilities and capabilities (Rajkumar, 2009). This is complete another architecture from common client-server where some nodes are basically meant for assisting and serving other nodes on the same network. P2P model solely rely on the network carrying capacity (bandwidth) of participants and the power.

The security measures which are monitored and governed by a set of security policies are therefore important to be put in place in any distributed environment (Ali, Hussein, and François, 2009). The security architecture has specified seven important security needs to protect Grid information as well as P2P data transfer across the peers against any attack either minor or major. These requirements are: authentication, authorization, availability, Non-Repudiation, Data Confidentiality, data integrity and privacy. This is depicted in the Figure 6 shown above. Ensuring the implementation of these security requirements has been identified a major task in any distributed computing environment especially Grid computing (Ali, Hussein, and François, 2009).

Implementation of these security requirements will go a long way at ensuring privacy, authorization, authentication, as well as data integrity on both platforms.

CONCLUSION

Since P2P and Grid computing are both designed to fulfill the same objective for harnessing and applying various resources across various organizations it will be more useful if both technologies can be married together to have absolute and optimal benefits. P2PGrid can be a very good hybrid for handling solution to a large-scale resource distribution and sharing as well as system integration. Incorporating the two technologies will ensure the elimination of weaknesses inherent as well as removing the service bottleneck and therefore ensure and improve the scalability, interoperability and dynamism of the system and resources.

REFERENCES

1. Ali, A.B., Hussein, Z., and François, S. 2009. "Access Control Mechanism for Mobile *ad hoc* Network of Networks (MANoN)". Software Technology Research Laboratory, De Montfort University: Leicester, UK.

2. Babak, et al. 2006. *Comparison of Global Computing with Grid Computing*. IEEE Computer Society: New York, NY.

3. Buyya, R. 2002. *National Partnership for Advanced Computational Infrastructure in Data Grid Managent System (DGMS)*. Melbourne, Australia.

4. Buyya, R. 2008. "The Gridbus Project". Retrieved July 28, 2010, from Grid Computing Info Center (GRID Infoware): http://www.gridcomputing.com/gridfaq.html

5. Buyya, R. (n.d.). "Grid Computing". Retrieved 11 12, 2010, from http://www.gridcomputing.com/gridfaq.html

6. Carl-Kesselman, I. F. 2003. *The Grid2: Blueprint for a New Computing Infrastructure*. Morgan Kaufmann Publishers: San Francisco, CA.

7. Domenico, T. and Paolo, T. 2004. "A P2P Grid Services-Based Protocol: Design and Evaluation". Euro-Par 2004. University of Calabria: Pisa Italy.

8. Ellahi, E.A. 2008. "Security Framework for P2P Based Grid Systems".

9. Foster, I. and Adriana, I. 2003. *On Death, Taxes, and the Convergence of Peer-to-Peer and Grid Computing*. Springer-Verlag: Berlin, Germany. 118-128.

10. Frank, T. 2008. "NBIC_Introduction Workflow". Retrieved November 19, 2010.

11. Foster, I. and Kesselman, C. 1997. "Globus: A Metacomputing Infrastructure Toolkit". *The International Journal of Supercomputer Applications and High Performance Computing*. 11(2):115-128.

12. Jacob, K. and Fukui, B. 2005. "Introduction to Grid Computing". IBM International Technical Support Organization (Redbooks): New York, NY.

13. JATIT. 2005. "Comparison of Grid Computing vs. Cluster Computing". Retrieved July 27, 2010, from *Journal of Theoretical and Applied Information Technology*. www.jatit.org/research/introduction_grid_computing.htm

14. NWICG. 2008. "Northwest Indiana Computational Grid (NWICG)". (Purdue University-Calumet, Purdue University – West Lafayette, and the University of Notre Dame). Retrieved August 2, 2010, from http://www.nwicgrid.org/

15. Rajkumar, K. 2009. "OSC Special Presentation". Retrieved November 19, 2010

16. The Gridbus Project. 2005. *Grid Computing Info Centre (GRID Infoware)*. (R. Buyya, Editor) Retrieved August, 1st, 2010, from http://www.gridcomputing.com/

17. Yanxiang, H., Fei, L., and Wensheng, H. 2008. "The Design and Implementation of Security Communication Model in Grid Networks". *Proceedings of the 2008 International Conference on Computer Science and Information Technology*. 421-424. IEEE Computer Society: New York, NY.

SUGGESTED CITATION

Azeez, N.A., A.P. Abidoye, A.O. Adesina, K.K. Agbele, Iyamu Tiko, and I.M. Venter. 2011. Peer-to-Peer Computing and Grid Computing: Towards a Better Understanding". *Pacific Journal of Science and Technology*. 12(1):270-276.

Pacific Journal of Science and Technology