Grid Computing

M.K.Vedhasree (Author)

BCA LLB(HONS)

School Of Excellence In Law,

Chennai, India

**ABSTRACT**

The grid concept is motivated by a real and specific problem – the coordinated resource sharing and problem solving of dynamic, multi-institutional virtual organizations. Now, a combination of technology trends and research progress results in an increased focus on grid technology industry and commercial domain. technologies have enabled the cooperative use of a wide variety of geographically distributed resources as a single more powerful computer. Grid computing is one of the newest methods of pooling resources so as to solve large-scale problems. This chapter explains the concepts underlying Grid computing.

**Keywords**— Grid Computing, Data, Computing, Resources

1. **INTRODUCTION**

 The term “the grid” has emerged in the mid-1990s to denote a proposed distributed computing infrastructure which focuses on large-scale resource sharing, innovative applications, and high-performance orientation [Foster et al., 2001]., The phrase "grid computing" was used as an analogy for rendering computational power as accessible as an electricity network.Ian Foster and Steve Tuecke of the University of Chicago and Carl Kesselman of the University of Southern California's Computer Sciences Institute gathered together the grid's concepts (which included those from parallel development, object-oriented development, and online services). The three are popularly considered as the "fathers of the grid".

1. **GRID COMPUTING**

 Grid computing is a collection of computer resources from multiple locations to reach a common goal. It is a computer network in which each computer’s resources are shared with every other computer.in grid computing the computer work on a task together, thus functioning as a supercomputer. We can infer grid computing as a distributed system with non-interactive work load that involves a large number of files. It is a type of management and commercial infrastructure, designed as a support primarily for science research. General-purpose grid network application packages are frequently used to create grids. The size of the grid might be extremely enormous. It is a sort of parallel processing that uses entire devices (with onboard CPUs, storage, power supply, network connectivity, and so on) linked to a network connection (private or public) via a traditional network connection, like Ethernet, for specific applications. A typical grid computing network consists of three machine types:

* **Control node/server:** A control node is a server or a group of servers that administers the entire network and maintains the record for resources in a network pool.
* **Provider/grid node:**A provider or grid node is a computer that contributes its resources to the network resource pool.
* **User:** A user refers to the computer that uses the resources on the network to complete the task.
1. **KEY COMPONENTS OF GRID COMPUTING**

**1. User interface**

 Today, users are well-versed with web portals. They provide a single interface that allows users to view a wide variety of information. Similarly, a grid portal offers an interface that enables users to launch applications with resources provided by the grid.

 The interface has a portal style to help users query and execute various functions on the grid effectively. A grid user views a single, large virtual computer offering computing resources, similar to an internet user who views a unified instance of content on the web.

**2. Security**

 Security is one of the major concerns for grid computing environments. Security mechanisms can include authentication, authorization, data encryption, and others. Grid security infrastructure (GSI) is an important ingredient here. It outlines specifications that establish secret and tamper-proof communication between software entities operating in a grid network.

 It includes OpenSSL implementation and provides[a single sign-on mechanism](https://www.spiceworks.com/it-security/vulnerability-management/articles/what-is-single-sign-on/) for users to perform actions within the grid. It offers robust security by providing authentication and authorization mechanisms for system protection.

**3. Scheduler**

 On identifying the resources, the next step is to schedule the tasks to run on them. A scheduler may not be needed if standalone tasks are to be executed that do not showcase interdependencies. However, if you want to run specific tasks concurrently that require inter-process communication, the job scheduler would suffice to coordinate the execution of different subtasks.

 Moreover, schedulers of different levels operate in a grid environment. For example, a cluster may represent an independent resource with its own scheduler to manage the nodes it contains. Hence, a high-level scheduler may sometimes be required to accomplish the task done on the cluster, while the cluster employs its own separate scheduler to handle work on its individual nodes.

**4. Data management**

 Data management is crucial for grid environments. A secure and reliable mechanism to move or make any data or application module accessible to various nodes within the grid is necessary. Consider the Globus toolkit — an open-source toolkit for grid computing.

 It offers a data management component called grid access to secondary storage (GASS). It includes GridFTP built on the standard FTP protocol and utilizes GSI for[user authentication](https://www.spiceworks.com/it-security/identity-access-management/articles/what-is-two-factor-authentication) and authorization. After authentication, the user can move files using the GridFTP facility without going through the login process at every node.

**5. Workload & resource management**

 The workload & resource component enables the actual launch of a job on a particular resource, checks its status, and retrieves the results when the job is complete. Say a user wants to execute an application on the grid. In that case, the application should be aware of the available resources on the grid to take up the workload.

So, it interacts with the workload manager to determine the resource availability and updates the status accordingly. This helps in efficient workload and resource management for various nodes on the grid.

1. **CLASSIFYING GRID SYSTEMS**

**1. Computational grid**

 It refers to systems that harness machines of an administrative domain in a “cycle-stealing” mode to have higher computational capacity than the capacity of any constituent machine in the system.

**2. Data grid**

It denotes systems that provide a hardware and software infrastructure for synthesizing new information from data repositories that are distributed in a wide area network.

**3. Service grid**

 It refers to systems that provide services that are not provided by any single local machine. This category is further divided as on demand (aggregate resources to provide new services), collaborative (con- 4 Fundamentals of Grid Computing net users and applications via a virtual workspace), and multimedia (infrastructure for real-time multimedia application

**Figure 1: Types of Grid computing**

1. **APPLICATIONS OF GRID COMPUTING**

**1.Financial services**

 Financial institutions use grid computing primarily to solve problems involving risk management. By harnessing the combined computing powers in the grid, they can shorten the duration of forecasting portfolio changes in volatile markets.

**2. Gaming**

The gaming industry uses grid computing to provide additional computational resources for game developers. The grid computing system splits large tasks, such as creating in-game designs, and allocates them to multiple machines. This results in a faster turnaround for the game developers.

**3. Entertainment**

 Some movies have complex special effects that require a powerful computer to create. The special effects designers use grid computing to speed up the production timeline. They have grid-supported software that shares computational resources to render the special-effect graphics.

**4. Engineering**

 Engineers use grid computing to perform simulations, create models, and analyze designs. They run specialized applications concurrently on multiple machines to process massive amounts of data. For example, engineers use grid computing to reduce the duration of a Monte Carlo simulation, a software process that uses past data to make future predictions.

1. **DIFFRENCE BETWEEN CLOUD AND GRID COMPUTING**

| **S.NO** | **Cloud Computing** | **Grid Computing** |
| --- | --- | --- |
| 1. | Cloud computing is a Client-server computing architecture. | While it is a Distributed computing architecture. |
| 2. | Cloud computing is a centralized executive. | While [grid computing](https://www.geeksforgeeks.org/grid-computing/) is a decentralized executive. |
| 3. | In cloud computing, resources are used in centralized pattern. | While in grid computing, resources are used in collaborative pattern. |
| 4. | It is more flexible than grid computing. | While it is less flexible than cloud computing. |
| 5. | In cloud computing, the users pay for the use. | While in grid computing, the users do not pay for use. |
| 6. | Cloud computing is a high accessible service. | While grid computing is a low accessible service. |
| 7. | It is highly scalable as compared to grid computing. | While grid computing is low scalable in comparison to cloud computing. |
| 8. | It can be accessed through standard web protocols. | While it is accessible through grid middleware. |
| 9. | Cloud computing is based on service-oriented. | Grid computing is based on application-oriented. |
| 10. | Cloud computing uses service like [IAAS](https://www.geeksforgeeks.org/difference-between-iaas-paas-and-saas/), PAAS, SAAS. | Grid computing uses service like [distributed computing](https://www.geeksforgeeks.org/what-is-distributed-computing/), [distributed pervasive](https://www.geeksforgeeks.org/introduction-to-pervasive-computing/#:~:text=Pervasive%20Computing%20is%20also%20called,are%20unaware%20of%20their%20presence.), distributed information. |

1. **HOW GRID COMPUTING WORKS**



**Figure 2: How Grid computing works**

**1.Task Submission**

 A user submits a computational task or job to the grid network through the control node. The task can be a complex calculation, data analysis, simulation, or other computationally intensive work.

**2.Task Segmentation**

The control node receives the submitted task and breaks it into smaller subtasks. These subtasks are designed to be independent and can be executed in parallel.

**3.Subtask Assignment**

The control node assigns each subtask to different provider nodes within the grid network. The assignment is based on resource availability, capability, and workload balancing.

**4.Parallel Execution**

 Each provider node receives its assigned subtask and executes it in parallel with others. The provider nodes work simultaneously on their respective subtasks using their local computing resources.

**5.Communication and Coordination**

During the execution phase, the provider nodes communicate with each other and the control node. This communication enables them to share information about the progress of their subtasks, exchange data, and synchronize their activities.

**6.Aggregation of Results**

 Once the provider nodes complete their subtasks, they return the results to the control node. The control node collects and aggregates these results to obtain the final output of the main task.

1. **BENEFITS OF GRID COMPUTING**
* One of the basic uses of grid computing is to run an existing application on a different machine.
* The potential for massive parallel CPU capacity is one of the most common visions and attractive features of a grid.
* Another capability enabled by grid computing is to provide an environment for collaboration among a wider audience.
* In addition to CPU and storage resources, a grid can provide access to other resources as well.
* A grid federates a large number of resources contributed by individual machines into a large single-system image.
* High-end conventional computing systems use expensive hardware to increase reliability. The goal to virtualize the resources on the grid and more uniformly handle heterogeneous systems will create new opportunities to better manage a larger, more distributed IT infrastructure.
1. **STANDARDS OF GRID COMPUTING ENVIRONMENT**

**1.Open Grid Service Architecture (OGSA)** defines requirements for these core capabilities and thus provides general reference architecture for grid computing environments.

**2.Open Grid Service Infrastructure (OGSI)** As grid computing has evolved it has become clear that a service-oriented architecture could provide many benefits in the implementation of a grid infrastructure. The Global Grid Forum extended the concepts defined in OGSA to define specific interfaces to various services.

**3. OGSA- Data Access and Integration** The OGSA-DAI (data access and integration) is concerned with building middleware to support with access and integration of data from distinct data sources through the grid.

**4. GridFTP** is a secure and reliable data transfer protocol providing high performance and optimized for wide area networks that have high bandwidth It uses basic Grid security on both control and data channels.

1. **ADVANTAGES OF GRID COMPUTING**

1. It is not centralized, as there are no servers required, except the control node which is just used for controlling and not for processing.

2. Multiple heterogeneous machines i.e. machines with different Operating Systems can use a single grid computing network.

3. Tasks can be performed parallelly across various physical locations and the users don’t have to pay for them (with money).

1. **DISADVANTAGES OF GRID COMPUTING**
2. The software of the grid is still in the involution stage.
3. A super-fast interconnect between computer resources is the need of the hour.
4. Licensing across many servers may make it prohibitive for some applications.
5. Many groups are reluctant with sharing resources.
6. Trouble in the control node can come to halt in the whole network.
7. **CONCLUSION**

 Grid computing is quickly developing to perform new science and develop new applications. It is a promising trend for some reasons they are, its capability to produce more cost-effective use of a given quantity of computer resources, then as a way to solve problems that cannot be dissolve computing power, and it suggests that the resources of many computers which are not in use can be utilize for other computational task. In this we have described about grid computing and its architecture methods and standards. Grid computing offers storage capabilities and the processing power for data processing. For the grid to support big data management and processing certain requirements based on big data concept have to be considered. Although grid computing provides technology to overcome the hardware limitation in term of storage space, processing power and memory capacity, Implementation of big data processing and management using grid computing requires additional techniques for managing the huge data effectively. In future we will discuss about the issues in security of grid computing and stab to deliver specific key for the problem of security and memory storage.

**REFERENCES**

Mishra, M. K., Patel, Y. S., Ghosh, M., & Mund, G. B. (2017). A review and classification of grid computing systems. International Journal of Computational Intelligence Research, 13(3), 369-402.

Refrence of an article from

<https://www.geeksforgeeks.org/>

Frederic Magoules,”Fundamentals of Grid Computing”.

“A Review on Wireless Grid Computing” S. S. Manvi, Member, IACSIT and M. N. Birje, International Journal of Computer and Electrical Engineering, Vol. 2, No. 3, June, 2010

Article referance https://aws.amazon.com/what-is/grid-computing/