ISSN (e): 2250 – 3005 || Volume, 08 || Issue, 03|| March – 2018 ||
International Journal of Computational Engineering Research (IJCER)

"Network of Computers to Analyze Data Using Grid Computing"

Subhashree sukla[1],Satyabrata dash[2], Rakhi jha[3]
[1,3]Assistant Professor, Department of Computer Science Engineering, Gandhi Institute For Technology (GIFT), Bhubaneswar
[2]Assistant Professor, Department of Computer Science Engineering, Gandhi Engineering College, Bhubaneswar

Abstract- *A scientist studying proteins logs into a computer and uses an entire network of computers to analyze data. A businessman accesses his company's network through a PDA in order to forecast the future of a particular stock. An Army official accesses and coordinates computer resources on three different military networks to formulate a battle strategy. All ofthese scenarios have one thing in common: They rely on a concept called grid computing. At its most basic level, grid computing is a computer network in which each computer's resources are shared with every other computer in the system. Processing power, memory and data storage are all community resources that authorized users can tap into and leverage for specific tasks. A grid computing system can be as simple as a collection of similar computers running on the same operating system or as complex as inter-networked systems comprised of every computer platform you can thinkof.*
Keywords- *Grid, Cloud, Distributed computing*

I. INTRODUCTION

Grid computing requires the use of software that can divide and farm out pieces of a program to as many as several thousand computers. Grid computing can be thought of as distributed and large-scale cluster computing LEARN MORE Data center outsourcing, collocation and cloud Public cloud computing services and as a form of network-distributed parallel processing.[1,3] It can be confined to the network of computer workstations within a corporation or it can be a public collaboration (in which case it is also sometimes known as a form of peer-to-peer computing).A grid computing has attracted great attention. In grid computing, by using multiple computers and executing processes in parallel, we can get powerful computational power. Grid computing systems often employ regular computers such as personal computers or workstations. However, wecan exploit computational power more effectively by employing compact computers embedded to many home electronics such as refrigerators or air- conditioners.[2] In this paper, we propose a grid computing system using compact computers. Generally, embedded compact computers almost always execute their assigned processes. Therefore, our proposed system controls their computational powers so that processes required by the grid computing system do not interrupt assigned processes for compact computers. Grid computing (or the use of a computational grid) is applying the resources of many computers in a network to a single problem at the same time - usually to a scientific or technical problem that requires a great number of computer processing cycles or access to large amounts of data. A well-known example of grid computing in the public domain is the ongoing SETI (Search for Extraterrestrial Intelligence) @Home project in which thousands of people are sharing the unused processor cycles of their PCs in the vast search for signs of "rational" signals from outer space. AccordingtoJohn Patrick, IBM's vice-president for Internet strategies, "the next big thing will be grid computing."The grid computing concept isn't a new one. It's a special kind of distributed computing.[1] In distributed computing, different computers within the same network share one or more resources. In the ideal grid computing system, every resource is shared, turning a computer network into a powerful supercomputer. With the right user interface, accessing a grid computing system would look no different than accessing a local machine's resources. Every authorized computer would have access to enormous processing power and storagecapacity

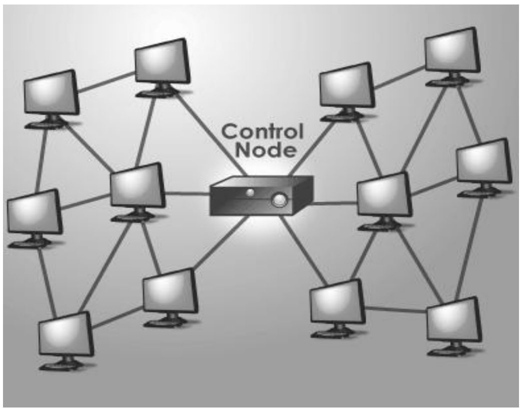

Figure 1.0 Grid Computing Basic Overview [2]

Grid computing systems work on the principle of pooled resources. Let's say you and a couple of friends decide to go on a camping trip. You own a large tent, so you've volunteered to share it with the others. One of your friends offers to bring food and another says he'll drive the whole group up in his SUV. Once on the trip, the three of you share your knowledge and skills to make the trip fun and comfortable. If you had made the trip on your own, you would need more time to assemble the resources you'd need and you probably would have had to work a lot harder on the tripitself.

II. GRID COMPUTINGCONCEPT

A grid computing system uses that same concept: share the load across multiple computers to complete tasks more efficiently and quickly. Before going too much further, let's take a quick look at a computer's resources: [3]

- Central processing unit (CPU): A CPU is a microprocessor that performs mathematical operations and directs data to different memory locations. Computers can have more than one CPU.
- Memory: In general, a computer's memory is a kind of temporary electronic storage. Memory keeps relevant data close at hand for the microprocessor. Without memory, the microprocessor would have to search and retrieve data from a more permanent storage device such as a hard diskdrive.
- Storage: In grid computing terms, storage refers to permanent data storage devices like hard disk drives ordatabases.

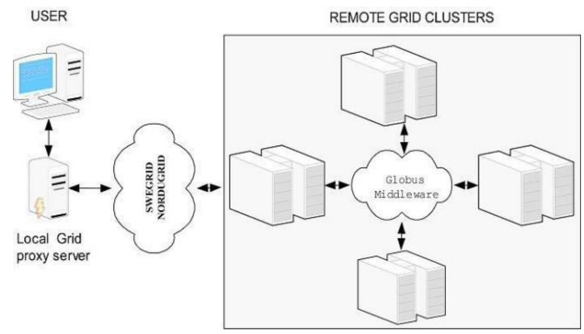

Figure 2.0 Basic Structure of Grid [3]

Normally, a computer can only operate within the limitations of its own resources. There's an upper limit to how fast it can complete an operation or how much information it can store. Most computers are upgradeable, which means it's possible to add more power or capacity to a single computer, but that's still just an incremental increase in performance. Grid computing systems link computer resources together in a way that lets someone use one computer to access and leverage the collected power of all the computers in the system. To the individual user, it's as if the user's computer has transformed into a supercomputer.

III. GRID COMPUTING-HOW IT WORKS?[3]

A scientist studying proteins logs into a computer and uses an entire network of computers to analyze data. A businessman accesses his company's network through a PDA in order to forecast the future of a particular stock. An Army official accesses and coordinates computer resources on three different military networks to formulate a battle strategy. All of these scenarios have one thing in common: They rely on a concept called grid computing. Atits mostbasic level, grid computing is a computer network in which each computer's resources are shared with every other computer in the system. Processing power, memory and data storage are all community resources that authorized users can tap into and leverage for specific tasks. A grid computing system can be as simple as a collection of similar computers running on the same operating system or as complex as inter-networked systems comprised of every computer platform you can think of. The grid computing concept isn't a new one. It's a special kind of distributed computing. In distributed computing, different computers within the same network share one or more resources. In the ideal grid computing system, every resource is shared, turning a computer network into a powerfulsupercomputer.

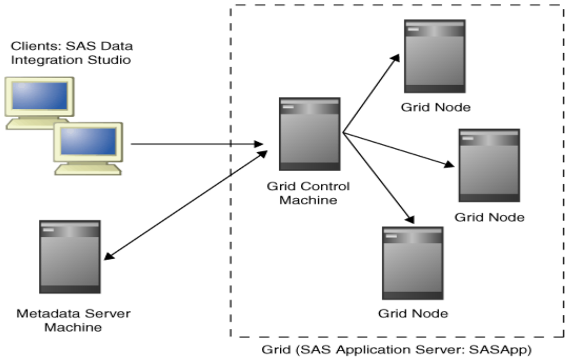

Figure 3.0 How GridWorks[3]

With the right user interface, accessing a grid computing system would look no different than accessing a local machine's resources. Every authorized computer would have access to enormous processing power and storage capacity. Though the concept isn't new, it's also not yet perfected. Computer scientists, programmers and engineers are still working on creating, establishing and implementing standards and protocols. Right now, many existing grid computer systems rely on proprietary software and tools. Once people agree upon a reliable set of standards and protocols, it will be easier and more efficient for organizations to adopt the grid computingmodel.

IV. GRIDCOMPUTING-SHARING RESOURCES [3]

Several companies and organizations are working together to create a standardized set of rules called protocols to make it easier to set up grid computing environments. It's possible to create a grid computing system right now and several already exist. But what's missing is an agreed-upon approach. That means that two different grid computing systems may not be compatible with one another, because each is working with a unique set of protocols andtools.

In general, a grid computing system requires:

- At least one computer, usually a server, which handles all the administrative duties for the system. Many people refer to this kind of computer as a control node. Other application and Web servers (both physical and virtual) provide specific services to the system.
- A network of computers running special grid computing network software. These computers act both as a point of interface for the user and as the resources the system will tap into for different applications. Grid computing systems can either include several computers of the same make running on the same operating system (called a homogeneous system) or a hodgepodge of different computers running on every operating system imaginable (a heterogeneous system). The network can be anything from a hardwired system where every computer connects to the system with physical wires to an open system where computers connect with each other over theInternet.[3]
- A collection of computer software called middleware. The purpose of middleware is to allow different computers to run a process or application across the entire network of machines. Middleware is the workhorse of the grid computing system. Without it, communication across the system would be impossible. Like software in general, there's no single format formiddleware.

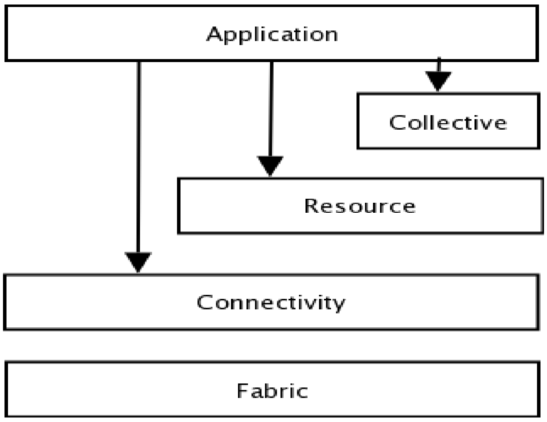

Figure 4.0 Grid Layers[3]

V. ISSUES IN GRID COMPUTING[5]

Grid computing is a highly collaborative distributed computing model; solutions to traditional distributed computing issues, such as security and resource management, do not scale well in grid computing. Furthermore, grid computing introduces other issues, such as information services and data management. We summarize those grid issues in the following and refer readers or related literature for more detailed discussions.

- Security: Most distributed computing systems use identity-based authentication and authorization control. As the typical case, a user is given a username and password foraccessinga computing system; when she is ready to launch her applications, she logs into the system and submits the application jobs. In a grid environment, users or theiragents simultaneously need accesses to multiple resources from different administrative domains that have different security mechanisms. This requirement creates several security issues. The two typical onesare:
- Single sign-on: A user should be able to authenticate once (e.g., when starting a computation) and initiate computations that acquire resources, use resources, release resources, and communicate internally, without further authentication of the user. Interoperability with local security solutions: While the grid security solutions may provide inter-domain access mechanisms, an access to a resource will typically be determined by a local security policy that is enforced by local security mechanisms. It is impractical to modify every local resource to accommodate inter-domain accesses.
- Resource Management: Grid resources are from different administrative domains that have their own local resource managers and a grid does not have full control of these resources. When managing these resources, a grid resource management system should respect the usage policies enforced by local resource managers, and meanwhile, deliver user required quality of services and improve global resource usage. This dilemma, i.e., managing a resource without ownership, is referred to as "site autonomy" and "heterogeneous substrate" issues. Another requirement for grid resource management comes from the fact that some grid applications, such as workflow, require resources to be allocated based on the application execution patterns and coordinated allocations of multiple resources simultaneously are necessary in order to deliver application-level quality of services. Also, resource management should be able to adapt application requirements to resource availability, particularly, when the requirements and resource characteristics change during execution. These issues are referred to as resource co-allocation andonline-control.[5]
- Information Services: Information services play an important role in grids. They indicate the status and availability of grid entities, i.e., compute resources, software libraries, networks, etc., without which there would be little coordination in such a dynamic environment as a grid. A grid information system should provide two types of services, the accounting service and the auditing service. Grid accounting maintains historical informationofresource status and job resource consumption for the purpose of performance prediction, resource allotment, charging and application performance tuning. Grid auditing provides runtime information of resource load status and application resource consumption for the purpose of resource allocation and resource usagecontrol.[4]
- Data Management: Data-intensive, high- performance computing applications require the efficient management and transfer of terabytes or petabytes of information in wide-area, distributed computing environments. Data management is concerned with how to provide secure, efficient and transparent access to distributed, heterogeneous pools of data on wide-area grid resources. In providing such services, grids should harness data, storage, and network resources located in distinct administrative domains, respect local and global policies governing how data can be used, schedule resources efficiently (again subject to local and global constraints), and provide high speed and reliable accesses todata.

- Standardization: Grid computing is a highly integrated system and a grid is built from multi- purpose protocols and interfaces that address those fundamental issues described above. The grid vision requires protocols (and interfaces and policies) that are not only open and general- purpose but alsostandard.

REFERENCES

[1]. "What is the Grid? A Three Point Checklist" Ian Foster,Argonne National Laboratory & University of Chicago
[2]. "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
[3]. http://www.techpdf.in/grid_computing.php
[4]. "A Study on Applications of Grid Computing in Bioinformatics", Manjula K.A, Department of Information Technology, Kannur University.
[5]. "Secure Grid Computing",Jianmin Zhu and Dr. BhavaniThuraisingham, University of Texas at Dallas, IJCSNS International Journal of Computer 216 Science and Network Security, VOL.6 No.8B, August 2006.