A Comprehensive approach on Grid Computing

Pallavi Gore
Sayali Phanse
Sourabh Shrivastava
Mr. Akhilesh K Sharma*

Abstract

Today we are in the Internet world and everyone prefers to enjoy fast access to the Internet. But due to multiple downloading, there is a chance that the system hangs up or slows down the performance that leads to the restarting of the entire process from the beginning. This is one of the serious problems that need the attention of the researchers.So we have taken this problem for our research and in this paper we are providing a layout for implementing our proposed Grid Model that can access the Internet very fast. By using our Grid we can easily download any number of files very fast depending on the number of systems employed in the Grid. We have used the concept of Grid Computing for this purpose.

1. Introduction: Until recently, application developers could often assume a target environment that was homogeneous, reliable, secure, and centrally managed. Increasingly, however, computing is concerned with collaboration, data sharing, and other new modes of interaction that involve distributed resources. Grid computing is the act of sharing tasks over multiple computers. Tasks can range from data storage to complex calculations and can be spread over large geographical distances. In some cases, computers within a grid are used normally and only act as part of the grid when they are not in use. These grids scavenge unused cycles on any computer that they can access, to complete given projects. These computers join together to create a virtual supercomputer. Modern supercomputers are built on the principles of grid computing, incorporating many smaller computers into a larger whole.

2. The Need for Grid technologies: The rise of the Internet and the emergence of e-business have, however, led to a growing awareness that an enterprise's IT infrastructure also encompasses external networks, resources, and services.
Problems due to multiple downloading:

The burden of multiple downloading and in particular with downloading huge files i.e., there can be a total abrupt system failure while a heavy task is assigned to the system. The system may hang up and may be rebooted while some percentage of downloading might have been completed. This rebooting of the system leads to download of the file once again from the beginning, which is one of the major problems everyone is facing today. Let us consider N numbers of files of different sizes (in order of several MBs) are being downloaded on a single system (a PC). This will take approximately some minutes or even some hours to download it by using an Internet connection of normal speed with a single CPU. This is one of the tedious tasks for the user to download multiple files at the same time. Our Grid plays a major role here.

*Assistant Professor, IES-IPS Academy, Rajendra Nagar Indore – 452012 , India

Initially, this new source of complexity was treated as a network-centric phenomenon and attempts were made to construct “intelligent networks” that intersect with traditional enterprise IT data centers only at “edge servers”: for example, an enterprise’s Web point of presence, or the virtual private network server that connects an enterprise network to service provider resources.

3. Why grid and why now?

- A biochemist exploits 10, 000 computers to screen 100,000 compounds in an hour.
- 1,000 physicists worldwide pool resources for & analyses of peta bytes of data
- Civil engineers collaborate to design, execute, & analyze shake table Experiments.
- An emergency response team couples real time data, weather model, population data.
- A multidisciplinary analysis in aerospace couples code and data in four companies.
- A home user invokes architectural design functions at an application service provider.
- Scientists working for a multinational soap company design a new product
- A community group pools members’ PCs to analyze alternative designs for a local road.

4. The Evolution of Grid Computing : The term grid computing originated in the early 1990s as a metaphor for making computer power as easy to access as an electric power grid in Ian Foster's and Carl Kesselman's seminal work, "The Grid: Blueprint for a new computing infrastructure" (1999). The ideas of the grid (including those from distributed computing, object-oriented programming, and Web services) were brought together by Ian Foster, Carl Kesselman, and Steve Tuecke, widely regarded as the "fathers of the grid".[6] They led the effort to create the Globus Toolkit incorporating not

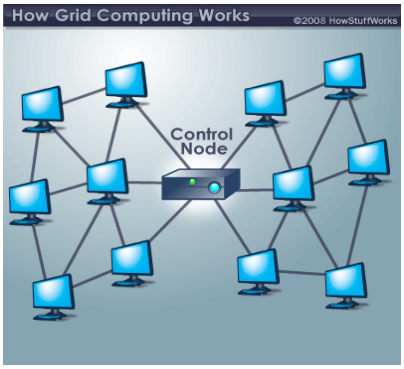
just computation management but also storage management, security provisioning, data movement, monitoring, and a toolkit for developing additional services based on the same infrastructure, including agreement negotiation, notification mechanisms, trigger services, and information aggregation If I have been able to see farther, it was only because I stood on the shoulders of giants. - Sir Isaac Newton, 1675

The software field seems to be operating on a different set of rules, perhaps because of the rush to patent...To his credit, Ian Foster is reluctant to be called the father of grid computing. Perhaps it's because he comes from an academic background with a tradition of giving credit where credit is due.

5. Working of Grid computing: 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 trip itself. 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:

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 disk drive.
Storage: In grid computing terms, storage refers to permanent data storage devices like hard disk drives or databases



In a basic grid computing system, every computer can access the resources of every other computer belonging to the network.

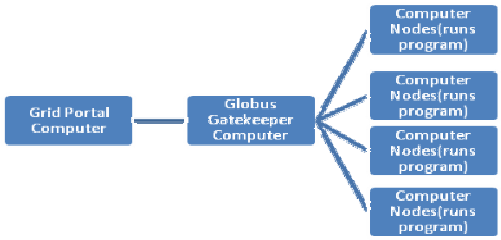
6. The Many Faces of Grid Computing: Grid computing is still a developing field and is related to several other innovative computing systems, some of which are subcategories of grid computing. Shared computing usually refers to a collection of computers that share processing power in order to complete a specific task. Then there's a software-as-a-service (SaaS) system known as utility computing, in which a company offers specific services (such as data storage or increased processor power) for a metered cost. Cloud computing is a system in which applications and storage "live" on the Web rather than on a user's computer.

7. Applications:

1. Fabric layer: interfaces to control layer.
2. connectivity layer: connecting easily and securely.
3. Resource layer: sharing single resource
4. Collective: coordinating multiple resources

8. Grids versus conventional supercomputers: "Distributed" or "grid" computing in general is a special type of parallel computing that relies on complete computers (with onboard CPU, storage, power supply, network interface, etc.) connected to a network (private, public or the Internet) by a conventional network interface, such as Ethernet. This is in contrast to the traditional notion of a supercomputer, which has many processors connected by a local high-speed computer bus.

9. Grid Hardware View:



10. Advantages of grid computing:

- Can solve larger, more complex problems in a shorter time
- Easier to collaborate with other organizations
- Make better use of existing hardware
- Much more efficient use of idle resources. Jobs can be farmed out to idle servers or even idle desktops. Many of these resources sit idle especially during off business hours. Policies can be in place that allow jobs to only go to servers that are lightly loaded or have the appropriate amount of memory/CPU characteristics for the particular application.
- Grid environments are much more modular and don't have single points of failure. If one of the servers/desktops within the grid fail there are plenty of other resources able to pick the load. Jobs can automatically restart if a failure occurs.

- Jobs can be executed in parallel speeding performance. Grid environments are extremely well suited to run jobs that can be split into smaller chunks and run concurrently on many nodes. Using things like MPI will allow message passing to occur among compute resources.

11. Disadvantages of grid computing

- Grid software and standards are still evolving
- Learning curve to get started
- Non-interactive job submission
- Licensing across many servers may make it prohibitive for some apps. Vendors are starting to be more flexible with environment like this.

12. The future: All software is network- centric

- We don't build or buy "computers" anymore, we borrow or lease required resources.
- When I walk into a room, need to solve a problem, need to communicate.
- A "computer" is a dynamically, often collaboratively constructed collection of processors, data sources, sensors, networks, Similar observations apply for software

13. Conclusion: Grid computing was once said to be fading out but due to the technological convergence it is blooming once again and the Intranet Grid we have proposed adds a milestone for the Globalization of Grid Architecture, which, leads to the hasty computing that is going to conquer the world in the nearest future. By implementing our proposed Intranet Grid it is very easy to download multiple files very fast and no need to worry about the security as we are authenticating each and every step taking place in our Grid and in particular user to access the database. Further implementations could be carried out in the nearest future.

14. References:

1. Jia Yu and Rajkumar Buyya1, **Gri**d Computing and **D**istributed **S**ystems (GRIDS)Laboratory, Department of Computer Science and Software Engineering, The University of Melbourne, Australia
2. http://www.gridbus.org
3. http://www.citeseer.ist.psu
4. http://www.google.com
5. http://www.wikipedia.com