A

Seminar report

on

Grid Computing

Submitted in partial fulfillment of the requirement for the award of
Degree of Computer Science

SUBMITTED TO: **SUBMITTED BY:**
www.studymafia.org   **www.studymafia.org**

<u>Preface</u>

I have made this report file on the topic **Grid Computing**, I have tried my best to elucidate all the relevant detail to the topic to be included in the report. While in the beginning I have tried to give a general view about this topic.

My efforts and wholehearted co-corporation of each and everyone has ended on a successful note. I express my sincere gratitude to …………..who assisting me throughout the preparation of this topic. I thank him for providing me the reinforcement, confidence and most importantly the track for the topic whenever I needed it.

Grid Computing

Introduction

Since its introduction, the concept of grid computing has acquired great popularity, even greater than the Web itself had at its beginning. The concept has not only found its place within numerous science projects (in medicine e.g.), but is also being used for various commercial applications.

Furthermore, grid computing is particularly suitable for resource-demanding projects and enables scientists to work in a completely new way. Despite all its advantages, there are still features which have not yet been developed, and there are also certain disadvantages of grid computing which are discussed within this article.

What is Grid Computing?

Grid computing is a type of data management and computer infrastructure, designed as a support primarily for scientific research, but, as said in the introduction, also used in various commercial concepts, business research, entertainment and finally by governments of different countries.

On its simplest level, the grid computing concept integrates four components: information, computation, networking and communication. When these components are connected into the grid, the result is a virtual platform which allows an advanced data and computation management.

If this concept is implemented into the areas mentioned above (science, etc.), it provides a platform where resources can be dynamically linked together, while these resources are then used to support the execution of applications that require significant amounts of computer resources.

Who can use grid computing

Scientists use grid computing for their research. But what about you? And who else might be interested? Who wants to invest time and money in "resource sharing"?

- **Governments and International Organizations?** Problems like disaster response, urban planning and economic modeling are traditionally assigned to national governments or coordinated by International Organizations like the United Nations or the World Bank. Imagine if these groups could apply the collective power of the nation's fastest computers and share their data archives more simply and effectively...

- **The military?** It's a pretty safe bet that the military in many countries is already developing grid technology. The United States have traditionally used their most powerful computers for military applications. But this Virtual Organization is unlikely to let other users access its grid!

- **Teachers and educators?** Education involves students, teachers, mentors, parents and administrators and so is a very natural application of grid technologies. E-libraries and e-learning centers are already benefiting from grid-based tools for accessing distributed data and creating virtual classrooms with distributed students, resources and tutors.

- **Businesses?** Global enterprises and large corporations have sites, data, people and resources distributed all over the world. Grids will allow such organizations to carry out large-scale modeling or computing by simultaneously using the resources at their many sites.

Grid Computing Applications

One of the most tantalizing applications of radio astronomy is the observation of radio signals as part of Searches for Extra Terrestrial Intelligence (SETI). The first SETI project in the late 1950's used a radio astronomy dish. Some years later the SETI Institute, a privately-funded organization, was formed to pursue more extensive searches using various radio telescopes in the U.S. The SETI Institute is now building its own array, again with private funds but in close collaboration with radio astronomy engineers and researchers at many observatories and universities.

The vast amount of computing capacity required for SETI radio signal processing has led to a unique grid computing concept that has now been expanded to many applications. SETI@home is a scientific experiment that uses Internet-connected computers to download and analyze radio telescope data for the SETI program. A free computer software program harnesses the power of millions of personal computers, and runs in the background using idle computer capacity. More than 5.2 million participants have logged over 2 million years of aggregate computing time.

Grid computing is now being used for other applications that include biology, medicine, earth sciences, physics, astronomy, chemistry, and mathematics. The Berkeley Open Infrastructure for Network Computing (BOINC) is free, open-source software for volunteer computing and desktop grid computing. Running the BOINC platform allows users to divide work among multiple grid computing projects, choosing to give only a percentage of CPU time to each.

These projects have tremendous humanitarian and economic potential. For example, the malariacontrol.net project is an application that makes use of network computing for stochastic modeling of the clinical epidemiology and natural history of plasmodium falciparum malaria. Simulation models of the transmission dynamics and health effects of malaria are an important tool for malaria control. They can be used to determine optimal strategies for delivering mosquito nets, chemotherapy, or new vaccines which are currently under development and testing.

Grid architecture

A grid's architecture is often described in terms of "layers", where each layer has a specific function. The higher layers are generally user-centric, whereas lower layers are more hardware-centric, focused on computers and networks.

- The lowest layer is **the network**, which connects grid resources.

- Above the network layer lays the **resource layer**: actual grid resources, such as computers, storage systems, electronic data catalogues, sensors and telescopes that are connected to the network.

- The **middleware layer** provides the tools that enable the various elements (servers, storage, networks, etc.) to participate in a grid. The middleware layer is sometimes the "brains" behind a computing grid!

- The highest layer of the structure is the **application** layer, which includes applications in science, engineering, business, finance and more, as well as portals and development toolkits to support the applications. This is the layer that grid users "see" and interact with. The application layer often includes the so-called service ware, which performs general management functions like tracking who is providing grid resources and who is using them.

TYPES OF GRID

Grid computing can be used in a variety of ways to address various kinds of application requirements. Often, grids are categorized by the type of solutions that they best address. The three primary types of grids are

a) COMPUTATIONAL GRID:

A computational grid is focused on setting aside resources specifically for computing power. In this type of grid, most of the machines are high-performance servers.

b) Scavenging Grid:

 A scavenging grid is most commonly used with large numbers of desktop machines. Machines are scavenged for available CPU cycles and other resources. Owners of the desktop machines are usually given control over when their resources are available to participate in the grid.

c) DATA GRID:

A data grid is responsible for housing and providing access to data across multiple organizations. Users are not concerned with where this data is located as long as they have access to the data. For example, you may have two universities doing life science re-search, each with unique data. A data grid would allow them to share their data, manage the data, and manage security issues such as who has access to what data. Another common distributed computing model that is often associated with or confused with Grid computing is peer-to-peer computing. In fact, some consider this is another form of Grid computing.

Advantages

For improved product quality. By employing existing resources, grid computing helps protect IT investments, containing costs while providing more capacity.

- **Increased user productivity**

By providing transparent access to resources, work can be completed more quickly. Users gain additional productivity as they can focus on design and development rather than wasting valuable time hunting for resources and manually scheduling and managing large numbers of jobs.

- **Scalability**

Grids can grow seamlessly over time, allowing many thousands of processors to be integrated into one cluster. Components can be updated independently and additional resources can be added as needed, reducing large one-time expenses.

- **Flexibility**

Grid computing provides computing power where it is needed most, helping to better meet dynamically changing work loads. Grids can contain heterogeneous compute nodes, allowing resources to be added and removed as needs dictate.

Disadvantages

1) for memory hungry applications that can't take advantage of MPI you may be forced to run on a large SMP

2) you may need to have a fast interconnect between compute resources (gigabit ethernet at a minimum). Infobahn for MPI intense applications

3) some applications may need to be tweaked to take full advantage of the new model.

4) Licensing across many servers may make it prohibitive for some apps. Vendors are starting to be more flexible with environment like this.

Areas that already are taking good advantage of grid computing include bioinformatics, cheminformatics, and oil & drilling, and financial applications.

With the advantages listed above you'll start to see much larger adoption of Grids which should benefit everyone involved. I believe the biggest barrier right now is education.

CONCLUSION

Grid computing introduces a new concept to IT infrastructures because it supports distributed computing over a network of heterogeneous resources and is enabled by open standards.

Grid computing works to optimize underutilized resources, decrease capital expenditures, and reduce the total cost of ownership. This solution extends beyond data processing and into in-formation management as well.

Information in this context covers data in databases, files, and storage devices. In this article, we outline potential problems and the means of solving them in a distributed environment.