International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395 -0056
Volume: 02 Issue: 09 | Dec-2015 www.irjet.net p-ISSN: 2395-0072

Article on Grid Computing Architecture and Benefits

Ms. K. Devika Rani Dhivya[1], Mrs. C. Sunitha[2]

[1]Assistant Professor, Dept. of BCA & MSc.SS,
Sri Krishna Arts and Science College, CBE, Tamil Nadu, India.
[2] Head, Dept. of Dept. of BCA & MSc.SS,
Sri Krishna Arts and Science College, CBE, Tamil Nadu, India

Abstract: One of the most popular forms of Networking is the Grid computing, which mean unlike things to different persons. The opinion of grid computing becomes worldwide and an individual user can gain access to any type of computing resources. The concept of grid computing has arisen as an innovative method to high concert distributed computing infrastructure. The most important pros of Grid are handling and organizing computer networks and essentially their deeper resource sharing and storing in a new way. It has evolved into an important discipline with in data analysis in the computer industries. Nowadays grid computing are used with big data analysis for reducing the complexity of Storage overhead. This article consists of basic architecture, methods and benefits of grid computing and analysis of grid with Big Data.

Key words: Networking, distributed computing, Grid computing

1. INTRODUCTION

1.1 Grid computing

Grid computing is a processor architecture that associates computer resources from various areas to reach an objective. In grid computing, an individual computer can connect with networks of computers that can perform the task together, thus working as a super processor. It is a form of interconnected computer systems where the machines utilize the same resources collectively. The concept of grid computing invented in the early 1990s as a symbol for making powerful computers as easy as to work with power grid.

A computational grid is a collection of mixed type of computers and the resources spread across numerous administrative fields with the resolute of providing users easy access to these resources. Speaking technically, Grid computing enables the virtualization of distributed computing and data resources such as processing, network bandwidth and storage capacity to create a single system image, granting users and applications seamless access to vast information technology (IT) capabilities. [4].

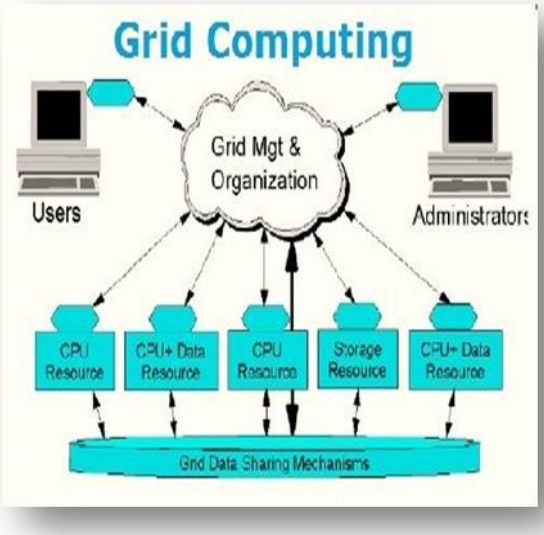

Fig -1: **Grid Computing**

It is a form of distributed computing that contains organizing and sharing computational power, data storage and interconnections across dynamic organizations. It is a model which is used to provide solutions for data sharing and analysis for engineering sciences, industry and commerce. It can be considered as data sharing systems with non-interactive workloads which involves a large number of files (Collection of data).

Due to increasing in the number of applications, the utilization of Grid Infrastructure has radically better to meet the need of data sharing, computational, storage and other needs. All the resource needs of today's demanding applications cannot simply meet in a single location or a site, therefore by using distributed resources can carry many benefits to the users of applications. It can be an effectual organization of heterogeneous, geographically distributed and dynamically available resource by deploying in *Grid systems*. [1] [2]

In the current world data analysis plays a major role in major industries, thus handling all the data sharing and storing can give up to Grid Computing for better performance. Big data is business transformation. So every organization is trying to analyze their big data. Big data poses implementation problems in extreme conditions. There are some methods to use grid computing along with hadoop. Grid computing provide large storage capability and computation power. Some open source toolkit use to implement solution such as Hadoop, Globus Toolkit.[6]

Hadoop analyze big data but faces some problem of storage and computation power. Grid helps hadoop to overcome these problems. So combination of Hadoop and grid for implement big data can be used.

1.2 Grid Computing Architecture

Grids focused on integrating existing resources with their hardware, operating systems, local resource management, and security infrastructure. In order to support the creation of the so called "Virtual Organizations"—a logical entity within which distributed resources can be discovered and shared as if they were from the same organization, Grids define and provide a set of standard protocols, middleware, toolkits, and services built on top of these protocols. Interoperability and security are the primary concerns for the Grid infrastructure as resources may come from different administrative domains, which have both global and local resource usage policies, different hardware and software configurations and platforms, and vary in availability and capacity. Grids provide protocols and services at five different layers as identified in the Grid protocol architecture (see Fig 2). At the Fabric layer, Grids provide access to different resource types such as compute, storage and network resource, code repository, etc. Grids usually rely on existing fabric components, for instance, local resource managers.

General-purpose components such as GARA (general architecture for advanced reservation) [5], and specialized resource management services. Connectivity layer defines core communication and authentication protocols for easy and secure network transactions. The GSI (Grid Security Infrastructure) protocol underlies every Grid transaction. The Resource layer defines protocols for the publication, discovery, negotiation, monitoring, accounting and payment of sharing operations on individual resources.

The GRAM (Grid Resource Access and Management) protocol is used for allocation of computational resources and for monitoring and control of computation on those resources, and GridFTP [5] for data access and high-speed data transfer. The Collective layer captures interactions across collections of resources, directory services such as MDS (Monitoring and Discovery Service allows for the monitoring and discovery of VO resources, Condor-G and Nimrod-G are examples of coallocating, scheduling and brokering services, and MPICH for Grid enabled programming systems, and CAS (community authorization service)for global resource policies. [5]

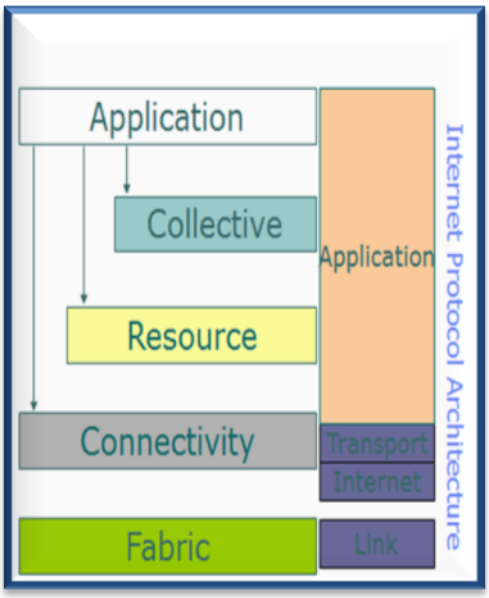

Fig - 2: Architecture of Grid Computing

1.3 Standards and Methods For Grid Computing Environments

- ***Open Grid Service Architecture***

OGSA defines requirements for these core capabilities and thus provides general reference architecture for grid computing environments.

- ***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.

- ***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.

- ***GridFTP***

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. [5]

1.4 Methods of Grid Computing

- ***Distributed Supercomputing***

It combines multiple high-capacity resources on a computational grid into a single, virtual distributed supercomputer. And Tackle problems that cannot be solved on a single system.

- ***High-Throughput Computing***

Using the grid to schedule large numbers of loosely coupled or independent tasks, with the goal of putting unused processor cycles to work. Thus the high-throughput computing was achieved.

- ***On-Demand Computing***

It has a capability to meet short-term requirements for resources that are not locally accessible and its models real-time computing demands.

- ***Data-Intensive Computing***

The focus is on synthesizing new information from data that is maintained in geographically distributed repositories, digital libraries, and databases. Data intensive computing is particularly useful for distributed data analysis.

- ***Logistical Networking***

It concern with global scheduling and optimization of data movement. It is contrasts with traditional networking, which does not explicitly model storage resources in the network and by called "logistical" because of the analogy it bears with the systems of warehouses, depots, and distribution channels.

1.5 Benefits of Grid Computing [5]

- Exploiting underutilized resources
- Parallel CPU capacity
- Virtual organizations for collaboration and virtual resources
- Access to additional resources
- Resource balancing
- Reliability
- Management

- ***Exploiting underutilized Resources***

One of the basic uses of grid computing is to run an existing application on a different machine. The machine on which the application is normally run might be unusually busy due to a peak in activity.

- ***Parallel CPU capacity***

The potential for massive parallel CPU capacity is one of the most common visions and attractive features of a grid. This computing power is driving a new evolution in many fields.

- ***Virtual resources and virtual organizations for collaboration***

Another capability enabled by grid computing is to provide an environment for collaboration among a wider audience.

- ***Access to additional resources***

As already stated, in addition to CPU and storage resources, a grid can provide access to other resources as well. The additional resources can be provided in additional numbers and/or capacity.

- ***Resource balancing***

For applications that are grid-enabled, the grid can offer a resource balancing effect by scheduling grid jobs on machines with low utilization.[5]

- ***Reliability***

High-end conventional computing systems use expensive hardware to increase reliability.

- ***Management***

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.

2. IMPLEMENTATION OF BIG DATA WITH GRID COMPUTING ENVIRONMENT

As earlier definite the main benefits offered by Grid computing are the storage capabilities and the processing power. To efficiently process and manage big data in grid environment, alterations are required on the recent grid infrastructure. General resources used in big data management and processing should be in collaborated with the grid environment.

The most common solution to handle Big Data is Hadoop, an open source project based on Google's MapReduce and Google File System. Hadoop was founded by the Apache Software Foundation. Hadoop is a distributed batch processing infrastructure which consists of the Hadoop kernel, Hadoop Distributed File System (HDFS), MapReduce and several related projects [6][3].

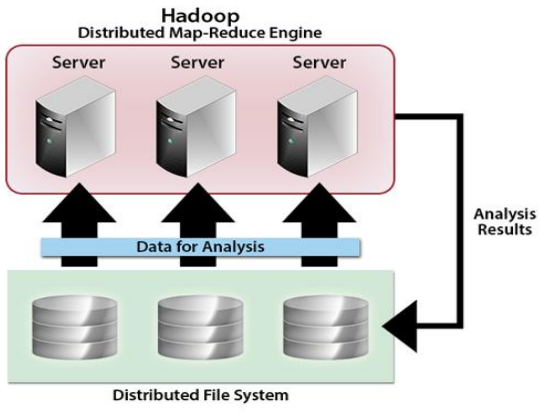

Fig - 3: Hadoop Distributing File system

The foundation of Hadoop lies in HDFS (Hadoop Distributed File System), The need for it comes from the fact that Big Data is stored on many machines.

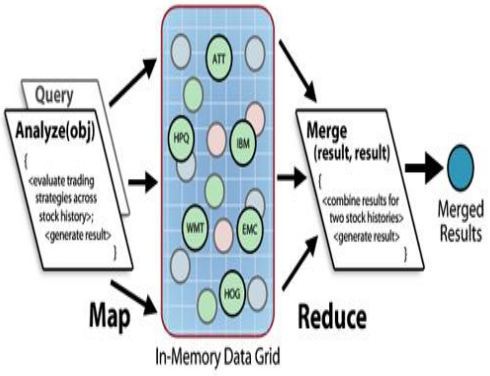

Fig - 4: Big data in memory grid

Here Grid Computing plays a major role in data distributions from one type of system to other which consist of different configuration. And it is useful in Providing Storage place. But recently the infrastructure of grid was not supporting for the Big data.

3. 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.

REFERENCE

[1] Ch v t e v Laxmi1, Dr. K.Somasundaram, "Application Level Scheduling (Apples) In Grid With Quality Of Service (Qos) ", International

Journal of Grid Computing & Applications (IJGCA) Vol.5, No.2, June 2014 .

[2] Ian F., and Carl K., "The Grid: Blueprint for a New Computing Infrastructure," Elsevier Inc., Singapore, Second Edition, 2004.

[3] Lawal Muhammad Aminu1,"Implementing Big Data Management on Grid Computing Environment ", International Journal Of Engineering And Computer Science ISSN:2319-7242 Volume 3 Issue 9, September 2014 Page No. 8455-8459

[4] Rahul Kumar, I. A. Khan and V. D. Gupta 2,"Literature review on grid computing", Vol. 6(7), pp. 144-148, July, 2013 DOI: 10.5897/AJMCSR11.154 ISSN 2006-9731 © 2013 Academic Journals.

[5] Shruti N. Pardeshi1, Chitra Patil, Snehal Dhumale „ "Grid Computing Architecture and Benefits", International Journal of Scientific and Research Publications, Volume 3, Issue 8, August 2013, ISSN 2250-3153 , www.ijsrp.org

[6] Yuvraj S. Sase , Pratik A.Yadav, "Big Data Implementation Using Hadoop and Grid Computing ",International Journal of Innovative Research in Science, Engineering and Technology, An ISO 3297: 2007 Certified Organization, Volume 3, Special Issue 4, April 2014.

BIOGRAPHIES

Ms. K Devika Rani Dhivya, Msc., M.Phil., M.B.A., is working as an Assistant Professor in the Dept. of BCA & MSc SS at Sri Krishna College of Arts and Science, Coimbatore. She is pursuing her Ph.D. at Bharathiar University, Coimbatore and doing research work on Software Engineering. She has published papers in International journals and presented at various Seminars, National and International Conferences and wishes to contribute to the Computing arena..

Mrs. C Sunitha, MCA, M.Phil., is working as the Head, Dept. of BCA & MSc SS at Sri Krishna College of Arts and Science, Coimbatore. With rich teaching and administrative experience, she is also the Controller of Examinations of the College. She is guiding many M.Phil. Research scholars and has published papers in International journals. She has presented papers and organized various Seminars and Forums. She is pursuing her Ph.D. at Bharathiar University, Coimbatore and doing research work on Speech Recognition. She is a Life Member of Computer Society of India (CSI), Coimbatore Chapter and The Indian Science Congress Association (ISCA).