Analytical Study of Various High Performance Computing Paradigms

Rashmi Gupta
M. Tech Scholar
Banasthali Vidyapeeth
Jaipur, India

Omesh Kumar
M. Tech Scholar
A. K. G Engineering college
Ghaziabad, India

Aditi Sharma
M. Tech Scholar
Banasthali Vidyapeeth
Jaipur, India

ABSTRACT

Now-a-days various computing paradigms are present in IT industry. Cloud, Grid, Cluster and General Purpose-Graphical Processing Unit (GP-GPU) computing are High Performance Computing (HPC) technologies and are growing very quickly. These are undoubtedly today's most enticing technology areas due to the various benefits offered by them such as virtualization, high performance and less managerial overhead at client side etc. However despite of the various advantages offered by these high performance computing paradigms, there are significant drawbacks of each computing technology such as highly dependability over network services and security concerns. In this paper we discuss about Cloud, Grid, Cluster and GP-GPU computing. In addition and equally importantly, this article focuses on the pros and cons of each of these computing paradigms as well as their application areas. This will help in better understanding of these computing paradigms.

General Terms

Parallel Computing, High performance computing.

Keywords

Cloud computing, Grid computing, Cluster computing, GP-GPU computing, Compute Unified Device Architecture.

1. INTRODUCTION

Initially high performance computing was very expensive to be affordable by everyone. Its advantages of having high speed were gained by only those organizations and institutions who could afford to have expensive supercomputers at that time. But the need for having it in small scale and at a lower cost forms the basis for the introduction of the technologies like cloud, grid, cluster and GP-GPU. These computing paradigms offer several advantages that add significant value to the company's bottom line. These computing paradigms have become buzzwords in the IT industry for a long time. But, many IT people are asking the question of what these computing paradigms actually are and where these can be used? To answer this question, we discuss each of these technologies and here is a short description of these high performance computing paradigms-

Cloud Computing is a computing paradigm which is based on internet. It delivers on-demand IT services to users on pay-per-use basis i.e. at a much lower cost. You have to pay for what and how much you use. It offers reduced investment, expected performance (that is why it is high performance computing), high availability, scalability, accessibility and mobility (accessible from anywhere) and many more services. It enables users to enjoy less expensive and more reliable services that are easily accessible and user friendly. Cloud computing involves heterogeneous computers.

Grid computing provides users with the resources they need, at the time they need them. It aims to aggregate the computing power so that the problems which are not solvable in small computers can be computed using this technology. To solve a single large task, it combines computers of heterogeneous nature and the task is divided among them. Grid computing is much like an electrical grid where we do not need to know about the details. We just use the services (current) and do not need to know where the resources are (power plant) and how we are getting the service (current).

A cluster is a collection of parallel or distributed computers which are interconnected using high-speed networks, such as Gigabit Ethernet [1] to solve problems that are either not solvable by a single computer or takes too much time to be solved. These interconnected computers work together to execute such tasks since it is not feasible to execute them on a single computer. Clusters offer high availability and load balancing services. When many homogeneous computers are connected and the load is shared between them, they make a cluster.

GP-GPU is based on the many-core trajectory which aims to increase the execution speed of parallel applications [3]. GPU is used in conjunction with a CPU. The CPU handles the sequential part of the application while GPU solves scientific extensive computation part. The many core GPU is made up of a large number of much smaller cores, each of which is heavily multithreaded. There is a significant increase in the computing speed of a CPU which is equipped with GPU. When the comparison between the speed of GPUs and CPUs is made, the ratio comes out to be 10:1. This ratio shows why GPU computing has been categorized as high performance computing.

This paper is organized as follows:

Section 2 discusses the cloud computing, its pros and cons and application areas. Grid computing, its pros and cons and application areas are listed in section 3. Similarly section 4 and 5 discusses the cluster and GP-GPU computing and their advantages and disadvantages respectively.

2. CLOUD COMPUTING

Cloud computing provides pool of resources or services to the user on pay per use basis. The user needs not to concern about the implementation, update requirements and other managerial issues of resources. The only thing a client needs is the personal computer and an internet connection to access the network. Cloud computing evolves from Grid computing

[2]. Before understanding what cloud computing actually is, we must first have a look at how it is different from traditional computing.

Cloud computing is different from traditional hosting with its essential characteristics such as [5]:

1. Cloud computing provides **on demand** services to the clients without any human interaction at service provider's site.
2. Cloud computing provides large pool of resources to the client to use as **utility computing**.
3. Cloud computing provides **elasticity** in its services which lets a user to have as much or as little of a service as they want at any instant of time.
4. The service offered by cloud computing are **fully managed by the provider.** The user doesn't need to concern about that. The only thing user has to do is to use the services pay as per the usage, nothing more. These all properties make the system more efficient.

Cloud computing can be best understood using two examples:

1. A simple example of cloud computing is Yahoo email, Gmail or Hotmail etc. To use the services of cloud computing what we need is just a personal computer and an Internet connection to access your mails. We need not worry about anything like how to manage the mails because the server and email management software exists on the cloud and is totally managed by the cloud service provider like Yahoo, Google etc. We can just logon to the email providing service and get our mails. This can be compared with the analogy "If you need milk, will you buy a cow?" Of course not. One would like to use the milk (service) and get its benefits provided by the cow (hardware/software) on the pay-per-use basis instead of buying the cow (a hardware/software). This also decreases the maintenance cost.
2. Next example is suppose you are an executive at a large organization and you are responsible for checking out whether proper hardware/software is in place to meet the organization's objectives. Now let a new employee has joined the organization and he wants the hardware/software. Purchasing and installing them is easier for a single person but what if 100 employees have joined the organization and the hardware/software is now needed for those 100 employees. This would be rather difficult in terms of time and cost. So cloud computing comes with a solution. Instead of installing a suite of hardware/software for each employee, you would have to load only one application. That particular application would allow employees to log into a web based service that will hosts all the applications an employee would be needed for her/his job.

Cloud computing gets its name from the drawings typically used to describe the Internet. It is made up of two words- **cloud and computing**. Cloud can be defined as a pool of resources which includes storage, servers, database, network; software etc and computing can be defined as the activity of using and improving the computer hardware and software. Thus the terms cloud and computing makes up the definition of cloud computing-

"Cloud Computing is the delivery of applications, infrastructure and platform as a service over the Internet accessible from the web browser and desktop with the end user not having any knowledge of the service providing

system, as well as of where the software and data are residing on the servers on a pay-per-use basis."

Based on the type of services provide by cloud, the cloud computing services are broadly divided into three categories-

1. Infrastructure as a Service (IaaS) - It offers the required hardware like server space, network equipment and storage.
2. Platform as a Service (PaaS) - It supplies resources required to build applications and services like design, testing, deployment and hosting.
3. Software as a Service (SaaS) - The applications such as videoconferencing, CRM and accounting are hosted as service.

2.1 Pros of cloud computing

1. On demand- whenever the user demands the service it is ready to provide the services.
2. Reduced cost- cloud computing offers cost much less than you would spend on desktop software and its upgrades and modifications.
3. Increased mobility- With cloud computing you can access your information anywhere there is a network connection (the Internet).
4. Storage capacity- it is almost unlimited so you can store your data with no worries to increase your system's storage capabilities.
5. Elasticity- you can access as much or as little a service as per your need or you can scale your needs up and down depending on your service demand.
6. Fully managed by service provider- In the cloud, software upgrades and enhancements are no-cost, no-hassle, automatic.
7. Availability-Your data is still available even if you lose your laptop.
8. Easy Implementation- As organizations do not have to purchase the required software, hardware, install or upgrade them so the implementation part becomes easy.
9. Support services- The cloud vendors provide a range of services to their clients.

2.2 Cons of cloud computing

1. Accessibility concerns- Although all the resources in a cloud computing system are handled by the service provider, the user has to concern about accessing the services. Due to any technological problem such as loss of internet connectivity, we will be unable to get the services of the cloud. In the worst case scenario the user can lose the access to data he has stored on the cloud.
2. Security concerns- As cloud computing shifts the computing and storage etc from local machine to network, the security is much more critical issue. Sometimes it is difficult to trust on the service provider in case of critical and confidential data which has to be stored at the site of cloud service provider. Compromise with password may lead the leakage of confidential data.
3. Limited flexibility- With the use of cloud computing services, upgrading and new software features are ultimately in the hands of the service provider you have chosen. Although the service provider has various things to offer to the client, it still limits the

flexibility of the user, as compared to the services which all are local in absence of cloud computing.

4. An online service is more prone to threats than your local machine. Thus cloud computing is more vulnerable.

5. Dependency- It means the loss of control which further results in quality problems with cloud service providers.

2.3 Application areas of Cloud computing

Cloud computing is suitable both for commercial and non-commercial clients. It is believed that cloud computing has been one of the main sources for success in several major companies such as Google and Amazon [8]. Last year Google launched a service named as Google Wave that allowed groups of people to work on the same document, idea in real time or whenever convenient to each participant [4]. The use of cloud computing lowers the Total Cost of Ownership (TCO), thus benefits private as well as public sectors including healthcare, education and the activities of government agencies.

In both academia and industry, cloud computing has been recently attracting significant momentum and attention as one of those opportunities that could prove to be of immense benefits and empowering in some situations, due to its flexibility and pay-per-use cost structure, for organizations. In the educational and learning arena, this will be called "Education and Learning as a Service" (ELaaS) [9].

3. GRID COMPUTING

The grid computing is simply a way to execute compute jobs across a distributed set of resources instead of one central resource. This way the load is shared among several resources using the load sharing software. Instead of executing the job on one heavy burdened server, the load is distributed among several small computers. The user need not take care of which node is going to execute the job; this will be decided by the "grid software". The distributed nature of grid computing is transparent to the user.

"Grid computing, in short is a simple way to carry out computing jobs in a distributed manner with the help of compute resources which execute the jobs instead of using a huge, single central server. It is much like an electrical grid."

Grid computing was evolved because there was a need for more processing power, greater expertise, and access to data to solve large, complex scientific problems. Grid computing offers sharing of resources among individuals, organizations and institutions over different locations. Furthermore collaborative nature of Grids leads to the concept of virtual organizations consisting of a dynamic set of resources which belongs to a set of organizations to solve a specific problem. Grid computing divides a large program into much smaller fragments and assigns each fragment to an individual node [6]. Each node would now process the fragment it has been assigned and would return the result. Even if one node fails, the result would not get affected because the task it should do will be reassigned to other node. To divide and apportion pieces of a program among several computers grid computing uses a middleware. Middleware includes the methods required to manage the resources so that they can be used in an integrated manner. Thus grid computing is a method of distributed computing since it enables users to share the resources which are geographically distributed. A grid is basically the one that uses the processing capabilities of different computing units which can be at different locations, for processing a single task. Resource sharing is the fundamental concept of grid computing. A variety of resources may be shared, including computers, storage devices, sensors, scientific instruments, network, software or data. In Grid, every computing node is autonomous i.e. it has its own resource manager and behaves like an independent entity.

3.1 Pros of Grid computing

1. Efficient utilization of idle resources- jobs can be given to the computers which are idle at the time.

2. Grid computing does not have single point of failure- Even if one node fails other nodes take care of the processing part of the failure node.

3. Scalability- grid model scales very well. If we need more computing resources, we just plug them in by installing grid client on additional desktops.

4. Policies can be managed by the grid software.

5. 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.

6. It can solve large and complex problems in much shorter time.

7. Makes better use of existing resources.

8. Easier to collaborate with other organizations.

9. The need for huge servers is eliminated and cost efficient small computers are used.

3.2 Cons of Grid Computing

1. Non-interactive job submission.

2. A fast interconnection between computing resources (gigabit Ethernet at a minimum) is needed.

3. Grid software and standards (specially middleware) are still evolving.

4. Computers actually performing computations might not be trustworthy.

5. Due to the lack of central control over the hardware, there is no way to guarantee that nodes will not drop out of the network at random times.

6. Resources shared issues of trust, risk and costs for the people or organizations involved in sharing.

7. Since resources are heterogeneous in nature so managing changes among the heterogeneous computers may be a challenging task.

8. Tweaking of some applications will be required to make full use of grid computing [7].

3.3 Application areas of Grid computing

The size of a grid may vary from small a network of computer workstations within a corporation to large collaborations across many companies and networks. Grid computing has several application areas. Resources that can be shared can be as simple database or it can be some instruments. The sharable resources can be tangible or intangible. For example if researchers in different locations want to have access to an instrument say telescope, they can use grid computing for this purpose. In some applications, chemical databases in different locations are searched to find information needed to test molecular blocking, a factor in drug efficacy. Expertise is another resource shared in grid computing. Students can easily interact with the renowned scientists without travelling to their locations. The creation of dynamic and virtual multi-institutional organizations is yet another application area of grid computing. Many grids can be used in physics, chemistry

and genetics. Animation is another application area for grid computing.

4. CLUSTER COMPUTING

The term 'cluster computing' describes the network of computers that run distributed applications and shared resources. Cluster computing is a platform that combines commercial off-the-shelf computers with a high speed network to form a single powerful super computer. The use of clusters as computing platform is not just limited to scientific and engineering applications; there are many business applications that get benefit from the use of clusters. This technology improves the performance of applications by approaching parallelism, simultaneously on different machines and also enables shared use of distributed resources.

Basically, a cluster is a type of parallel or distributed computer system, which consists of a collection of inter-connected stand-alone computers working together as a single integrated computing resource. Clusters provide good platform for solving range of parallel and distributed applications. When several computers are connected to a high speed network they consume a single resource pool and these grouped computers are called clusters. Clusters are mainly used for eliminating single point failure i.e. if one node fails then other nodes can work as alternative. In cluster computing environment several computers are linked together and they share their workload as a single virtual computational platform.

Computers in a cluster network are attached in tightly coupled fashion i.e. they all use the same subnet and follow the same domain and networked with very high bandwidth connection and they all use the same hardware and software.

Computers connected in a cluster perform different roles such as coordinator, worker etc, so that they schedule processes with higher performance. In cluster computing environment users request is distributed among various computers in a network. The advantage of doing this is to increase the performance and processing of the task becomes fast.

Cluster Networks are of two types and are described as follows-
1. Software Clustering-In this type of cluster network, clusters are formed by connecting multiple servers that act like a single system and clustering software are installed in each of the servers in a group.
2. Hardware Clustering-In this type of clustering network, a special hardware device is installed on one of the servers which controls the whole network.

4.1 Pros of cluster computing

1. Reduced Cost-Because all the computers in a cluster use single resource pool which reduces the cost of cluster computer than general purpose computer.
2. Scalability-Computers in a cluster can be easily expanded as per requirement.
3. High Performance-Because of multiple computers connected together, cluster computing leads to speedy task processing.
4. High Availability-If a node in a cluster fails, its work can be transferred to another node with in a cluster ensuring that there is no interruption in the work.
5. Improved Network technology-Computers in a cluster are connected to single virtual network,

information can flow in the network without lag of time.
6. Large capacity-Because in a cluster network several computers are connected, which increase their capacity and they also solve the most challenging problem.
7. Message passing Interface (MPI)-In cluster network this programming interface allows distributed applications to communicate each other and share information.
8. Incremental growth.

4.2 Cons of cluster computing

1. Implementation and Configuration complexity-Implementation of cluster environment is complex, managing the connectivity among hosts and configuring them requires big team and is not an easy task for many organizations.
2. Update and Upgrade factors-Updating and upgrading multiple computers with newer version or new technology are cumbersome.
3. Cost Factor- To implement a clustered virtual host environment, we need to duplicate parts of the infrastructure. Also, most vender implementations require a storage area network or separate disk subsystem.

4.3 Application areas of cluster computing

Cluster computing environment can be used in science and engineering applications that require high computational power. For Example, it is used in execution of scientific applications used in predicting life threatening system such as earthquake and hurricanes. It is not possible to forecast earthquake using a single personal computer. For such type of applications cluster computing environment deliver better performance.

Cluster computers are also used in commercial applications e.g. in a bank where server halts and customer transactions fail and also some of confidential data is lost. A bank could operate after failure of server by migrating alternative resources as means of service.

Cluster Computing is also provided on windows by Windows Compute Cluster Server 2003(CCS) from Microsoft CCS supports cluster of servers which include one or more head node and more than one compute nodes. In CCS user first assign the job to the head node, then job identifies an application that run on cluster. The job scheduler on the head node assign each of the task (that defined by the job) to the node and then starts each of the application instance on the assigned node. Results from application instance are transferred to the clients in the form of files and databases.

Cluster-based systems can be used to execute many Internet applications: Web servers, search engines, email, and security, proxy, and database servers. With the escalating popularity of the Web, computer system availability is becoming critical, especially in e-commerce application. Popular and free email service providers such as Hotmail and search engines– that use Inktomi solutions are hosted by clusters.

5. GP-GPU COMPUTING

Before using the many-core GPUs in laptops/desktops, multi-core CPUs were used to increase the execution speed of sequential programs. While the multi-core CPUs focus on the speed of sequential programs, the many-core GPUs focus more on the execution throughput of parallel applications. It is used to accelerate the computing speed. General Purpose-Graphical Processing Unit (GP-GPU) computing makes use

of a Graphics Processing Unit to do general purpose scientific and engineering calculations. GPU is used in conjunction with CPU. The computing model uses a CPU and GPU together in a heterogeneous co-processing computing environment so that whenever the user requests the CPU to execute the application, its sequential part runs on the CPU **Table -1.**

Comparisons of Cloud, Grid, Cluster and GP-GPU computing paradigms.

	Cloud	Grid	Cluster	GP-GPU
Definition	Refers to both applications delivered as service as well as the hardware and software that provide those services.	Enables sharing of resources such as databases and includes heterogeneous computers.	Homogeneous computers work together to solve a problem	Use of many core graphics processing unit to increase the execution throughput of parallel applications.
Features	Internet based delivery of IT services on pay per use basis.	End users do not have knowledge of where the resources are and how the services they are getting	Good platform for solving range of parallel and distributed applications.	GPU is used in collaboration of CPU where sequential part of application runs on CPU and compute intensive part runs on GPU.
Used for	Healthcare, education, academia, industry, government agencies.	Physics, chemistry, genetics, expertise and databases can be shared.	Science and engineering applications, to execute Internet applications.	Biology, video & audio coding & manipulation, scientific & engineering computing.
Merits	On-demand reduced cost, increased mobility, reduced investment, elasticity, availability.	Solve large and complex problems, efficient utilization of resources, scalability.	Offers reduced cost, high performance, and high availability.	Teraflops speed, able to solve scientific calculations, SIMD structure.
Demerits	No or less Security, less flexibility	Non-interactive job submission, fast interconnection needed.	Implementation complexity.	Small memory, graphics programming languages need to be mastered.

and the computationally-intensive part is accelerated by the GPU. This way high computing speed can be achieved because the work can be done in parallel.

Initially GPU was not introduced to perform such massive computations. It was used to do graphics related operations only so that the load from the CPU can be shared and it can focus on the computation part. But later when the need to accelerate the computing speeds was felt; it was used to perform scientific calculations. An example of GPU is GTX 280 with 240 cores, each of which is heavily multithreaded, in order, single instruction issue processor that shares its control and instruction cache with seven other cores. NVIDIA is the GPU manufacturer and accelerated the computing speed by introducing GPUs that are built on its new massively parallel architecture called CUDA (Compute Unified Device Architecture). CUDA is the architecture which enables GPUs to perform graphics related operations as well as scientific calculations. CUDA C is the language used to program CUDA GPUs.

A GPU is a highly parallel computing device for the execution of highly compute intensive applications in which the same operations are carried out on many elements of data in parallel. The use of GPUs for computing purpose contributed a lot in high performance computing. It has evolved to have teraflops of floating point performance. The ratio of executing the floating point calculation for GPUs and that of CPUs is 10:1[3]. The difference in the computing speed between CPUs and GPUs is due to the differences in the fundamental design philosophies between the two types of processors.

5.1 Pros of GP-GPU

1. Parallel computing- Since GP-GPU is able to perform teraflops of operations; it becomes very powerful to enhance the parallel computing power.

2. The computationally intensive part of code that generally has Single Instruction Multiple Data (SIMD) structure may run very efficiently on GPU.

3. The architecture of GPU has streaming multiprocessors (SM) where each SM has a number of streaming processors (SP) and each SP is massively threaded so it can run thousands of threads per application, accelerating the computing speed.

4. The graphics pipeline is well suited for parallelism.

5.2 Cons of GP-GPU

1. The problem with GP-GPU is that to program GPU we require graphics programming languages like OpenGL and Cg. Developers need to learn the graphics programming languages and master them before using this high performance computing paradigm. This limits the accessibility of tremendous performance of GPUs for science.

2. GPU has small memory which is also a limitation as heavy I/O based modules may not run with good efficiency.

3. The memory bandwidth is much greater than the communication bandwidth between GPU and CPU.

4. GPUs are designed as numeric computing engines, and they will not perform well on some tasks on which CPUs are designed to perform well.

5. The GPU computing would benefit if programs written on GPUs can run efficiently on other architectures and vice versa.

5.3 Application areas of GP-GPU

The biology research community focuses more and more on the molecular level. But, since limitations exist to the molecular level observations, the increasing computing speed can contribute in simulating the molecular activities which in turn can address the above limitations. The increasing computing speed contributes in science and medical fields. Some more benefits offered by greater computing speed are much better user interfaces. For example the Apple iPhone interfaces offer much better touch screen interface to users than traditional handsets. For applications such as video and audio coding and manipulation higher computing speed is required and GPUs are providing this much of computing speed.

6. CONCLUSIONS

Cloud, Grid, Cluster and GP-GPU are the most emerging technologies. Many organizations are using these computing paradigms to increase the performance and throughput and to decrease the cost induced. Grid computing is generally used to handle specific problem which are generally large in nature. In this paper, we first discussed about each of these computing technologies. On the positive side each computing paradigm has its own advantages such as cloud offers on-demand services, reduced cost, Grid efficiently utilizes the resources and work on heterogeneous network and there is no dominant programming paradigm in Grid computing, Cluster offers high availability and work efficiently on problems with short task means some of which depend in result of previous task and GP-GPU offers teraflops of speed. But since every coin has two faces, these computing paradigms have disadvantages as well. As the advent of high speed networks accessibility is not a big issue now-a-days but still the security and reliability are the most concerning factors of these network based services. This article while discussing what the high performance computing paradigms are, also discusses the application areas for each of these so that we can know where the high performance computing paradigms can be used. Also, there is a table showing the comparative study of the high performance computing paradigms. This helps in better understanding of the computing technologies, all in one go.

7. REFERENCES

[1] N. Sadashiv, S. M. Dilip Kumar, "Cluster, Grid and Cloud Computing: A Detailed Comparison", 6th International Conference on Computer Science & Education (ICCSE), Singapore, 2011.

[2] S. Zhang, X. Chen, X. Huo, "The comparison between cloud and grid computing", 2010 International Conference on Computer Application and System Modelling (ICCASM), 2010.

[3] D. B. Kirk, W. Hwu, "Programming Massively Parallel Processors ", Morgan Kaufmann, 1-13, 2010.

[4] www.technobuffalo.com/internet/five-examples-of-cloud-computing/

[5] W. Kim, "Cloud Computing: Today and Tomorrow", Jr. of Object Technology, 8(1), 2009.

[6] F. Berman, A. Hey and G. Fox, "Grid Computing: Making the Global Infrastructure a Reality", John Wiley and Sons, 2003.

[7] http://voices.yahoo.com/pros-cons-grid-computing-4008827.html

[8] J. Deng, S. C-H. Huang, Y. S. Han, J. H. Deng, "Fault Tolerant and Reliable Computation in Cloud Computing", 2010.

[9] M. M. Alabbadi, "Cloud Computing for Education and Learning: Education and learning as a service", 14th International Conference on Interactive Collaborative Learning (ICL)-11th International Conference Virtual University (vu'11), Slovakia, 2011.