ISSN: 2278-1323
International Journal of Advanced Research in Computer Engineering & Technology (IJARCET)
Volume 1, Issue 9, November 2012

High Performance Computing: A Survey

Mr. Nilesh C. Thakkar, Mr. Nitesh M. Sureja

***Abstract*—This paper surveys techniques used for high performance computing. High performance computing is used to develop machines which provide computing power like super computers. It concentrates on both software as well as hardware development. As the complexity of the computing increases day by day, there is a requirement of having a cost effective computing environment which provide very high computing power. The activities related to research and simulations are the common examples where we require high performance computing. Different types of high performance computing techniques are developed which provides high computing power by using the software/hardware which is easily available and affordable to all. Cluster computing, cloud computing and grid computing are the examples of the techniques developed. In this paper all these methods are discussed and a summary is generated as conclusion.**

***Index Terms*—Cluster Computing, Cloud Computing, Grid Computing, High Performance Computing.**

I. INTRODUCTION

Parallel computing is used by High Performance Computing (HPC) to execute the applications which require advanced computing power. Parallel computing provides efficient, reliable and quick operations. High performance computing normally applies where a system has to work with a teraflop or 10^{12} floating-point operations per second. The term HPC is also used as a synonym for supercomputing as it provides computing power very near to a supercomputer. The only difference is super computers are developed for a special challenged application and used for it for its life span. Academicians, engineers, researchers are the most common users of the HPC. As they can't afford costly supercomputers, new techniques are developed which uses easily available and affordable hardware and software. Some software required to convert them in a HPC platform are also developed. This software is generally known as middleware. Every technique has a special middleware developed for it. Cluster computing [01], Cloud computing [09] and Grid computing [05] are the examples of the some HPC techniques developed. This paper reviews some papers already written by some authors and presents a conclusion in the form of a summary. This paper is organized as follows. Section1 provide a brief introduction. Section 2 discusses Cluster computing followed by Grid computing in section 3. Cloud computing is described in section 4. Finally, section 5 presents conclusion in the form of a summary.

Manuscript received Nov 19, 2012.

***Mr. Nilesh C. Thakkar**, Department of Computer Engineering, G H Patel College of Engineering & Technology, Vallabh Vidyanagar, INDIA, Mobile No: 08866302241*

***Mr. Nitesh M. Sureja**, Department of Computer Engineering, G H Patel College of Engineering & Technology, Vallabh Vidyanagar, INDIA. Mobile No: 09825313936.*

II. CLUSTER COMPUTING

In Cluster computing, a group of computers are connected together in such a way that they can perform as a single computer and provides the computing power of all connected computer for working [1][2][3][4]. The group of computer created is known as a cluster. The main reason for developing this computing is to give a high computing capability to people who can't afford a very costly supercomputer. The other intention was to provide HPC to individual users independently to meet their computing needs. This computing ensures fulltime availability to users. Cluster computing works by distributing the load of the users across an array of computers so that the accommodation of many users can be easily done. Clusters can be developed to balance the load on any network real time. These clusters are known as load balancing clusters. Clusters developed using very cost effective hardware can also be used to achieve high performance cluster computing where the computation is very complex involving very high amount of floating point operations. Cluster computing is very much useful where we perform simulations to verify some theories. A cluster of computers can be a cluster of clusters. In this way we can distribute the load among the clusters of a cluster. Design of a cluster should provide or maintain its performance at the peak. So, it is required to have skilled computer scientists while designing it. Not only the computer hardware, the designers keep operating systems, applications and other specification in consideration while designing a robust cluster. A very close network of same specification of hardware with dedicated connection is used to create a cluster. Security between the nodes of a cluster is also a very important issue in cluster computing. A very high degree of trust between the nodes of a cluster is required to communicate effectively. The architecture of a cluster is shown in figure 01. The back bone of the cluster architecture is its high speed network which is normally an advanced switch. All nodes are connected with each other through this switch. A node can be a simple PC, a Workstation, SMDs or a distributed system. All nodes have their own network interface hardware with communication software to connect themselves with a cluster. A software known as cluster middleware or availability infrastructure is placed as a layer on the personal operating systems of every nodes. This middleware handles all the operations performed on a cluster with the help of operating systems present on all individual nodes. A special parallel computing environment is developed and used to run the applications which run on a parallel computing environment. We can also run simple sequential applications on a cluster.

ISSN: 2278-1323
International Journal of Advanced Research in Computer Engineering & Technology (IJARCET)
Volume 1, Issue 9, November 2012

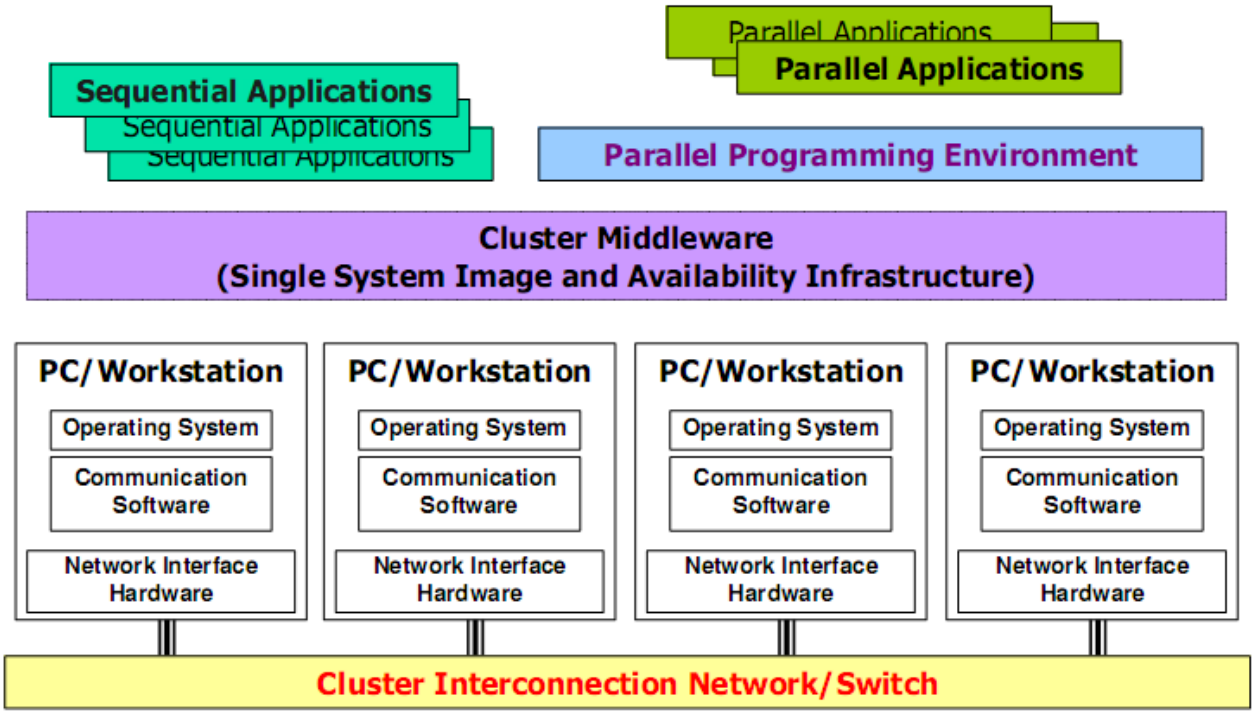

Figure 01 Cluster Architecture (R. Buyya [01])

III. Grid Computing

Grid computing is a method used to share computer power and data storage through the Internet. This method now a days makes a big contribution to research by making available, analyze and storing the large amount of data all around the world[5][6][7][8]. A powerful grid consists of many smaller grids connected together, forming a global network of computer. This network operates as one big computational resource. Already, there are hundreds of grids present across world. Each one of them is created for a special purpose by a specific group. Researchers and engineers are working to bring closer the several million computers present across the globe, and owned by millions different owners. In a grid, the computers may be desktops, laptops, supercomputers, data vaults, various instruments like mobiles, sensors, telescopes etc.. The grid is what all these computers connect themselves and forms a very huge single powerful computer. The name "Grid" is taken from an analogy with the electrical "power grid". The researchers want to access computer power as simple as accessing electrical power from an electrical grid. You don't have to worry about from where the computer power you are using comes from. You have to simply put your computer in to the Internet; it will get the computer power you need to do the job. The grid is the environment which makes this possible. Grid links together all computing resources, and provides the mechanism needed to access them. The Grid portal makes all remote computing resources accessible from different platforms whether it is a PC, a laptop, a PDA or mobile phone. All of those simply can be accessed through your web browser. We can use grid as a utility, you ask for a computational resource and you get it. You have to pay for what you use from a grid. There are some very important issues related to grid. Global sharing of computational resources is the very essence of grid computing. As the users don't know each other, trust between users and providers is essential. There may be Sharing resources conflicts due to different security policies on individual computer centers, on individual PCs and other computational resources, so getting grid security right is crucial. It is also essential that a balanced and efficient use of the resources is maintained. You should be able to access computer resources from where ever you are. A grid should adopt open standards while the development to achieve the interoperability between different grids. Use of standards encourages industry to invest in developing commercial grids. Figure 02 gives a brief idea about the layered architecture of a community grid model.

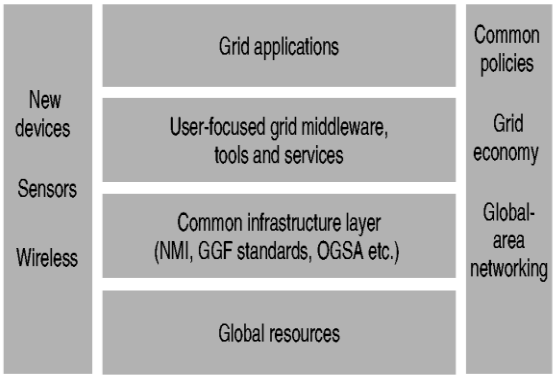

Figure 02 Layered Architecture of a community Grid [05]

A grid computing system requires a computer which is a server to handle all the administrative tasks of the system. This computer some times also known as a control node. Grid also requires some application as well as Web servers to provide specific services to the system. A network using special software (Middleware) is required to run a grid system. Nodes of this network act as user interface as well as the resources also. A grid can be of two types, homogeneous

and heterogeneous grid. Homogeneous grid is made of very similar software as well as hardware, while it is not here in the case of a heterogeneous grid. A grid also consist a collection of computer software called middleware. The task of middleware is to make possible to run a process or application across the grid. Middleware can be called a heart of the grid computing scheme. Without a middleware, it is impossible to achieve the communication across the system. Middleware is available in different formats. The control node is known as a dispatcher. It gives the priority and schedule tasks on the network. It determines what resources each task will be able to access. It also monitors the system and ensures that the system never becomes overloaded. It also maintains each user's computer's performance.

IV. CLOUD COMPUTING

Cloud computing is a modern computing method which shares the computing resources and removes the need of having local servers or personal devices to handle applications [9][10][11][12]. It is an internet based computing method which delivers services like servers, storage and software applications to the computers of the organization through the internet. The main goal of cloud computing is to provide computing capabilities like supercomputing to the consumers. Handling finance related activities, providing storage facilities, computer games are some examples of this. Large networks of a very large group of servers running on low cost general PC technology connected with special interfaces to achieve the data processing across them are used by cloud computing. Cloud computing creates a shared IT infrastructure containing a very big group of linked systems. Cloud computing uses virtualization to improve its performance. Cloud computing shifts the significant workload of running applications on the local computers to the network of computers which makes a cloud of the computers. Users connect their computers to the cloud via interface software, and then cloud's network takes care of the rest. A cloud can be dived into two parts namely, frond end and back end. Both the parts connect each other with the internet. Users are generally refers as front end while cloud refers itself as back end. Front end is made of the user's computer/computer network and software required to access the cloud. Various computers, servers and data storage system make the back end which is in fact the cloud of computing services. a cloud computing system could include practically any computer program you can imagine, from data processing to video games. Usually, each application will have its own dedicated server. The system has a server to manage the system centrally. It monitors traffic, administers the system and sees that the clients demands served efficiently. A set of protocols and software called middleware are used to ensure a smooth communication over the network. Clients can access any application or use a resource from the cloud as a service. Client has to pay for the service used from the cloud. Cloud computing comes in three forms: public clouds, private clouds, and hybrids clouds. A public cloud is one in which the services and infrastructure are provided off-site over the Internet. These clouds offer the greatest level of efficiency in shared resources; however, they are also more vulnerable than private clouds. A private cloud is one in which the services and infrastructure are maintained on a private network. These clouds offer the greatest level of security and control, but they require the company to still purchase and maintain all the software and infrastructure, which reduces the cost savings. A hybrid cloud includes a variety of public and private options with multiple providers. By spreading things out over a hybrid cloud, you keep each aspect at your business in the most efficient environment possible. A cloud provides three types of services [13][14][15]. They are: Software as a Service (SaaS), Platform as a Service (Paas) and Infrastructure as a Service (Iaas).

A. Software as a Service (SaaS)

In this Software type of service, a cloud provider provides software as a Service. SaaS is most popular of all type of the services a cloud provides. Storage, data backup solutions, web-based tools, services such as email, and millions of these type services are offered to subscribers for a monthly, quarterly, or annual fee. Users don't have to install or buy anything while they use SaaS service .

B. Platform as a Service (PaaS)

In this type of service a cloud provider provides platform such as Microsoft Access, Microsoft Azure, etc as a service. PaaS is considered as a web-based platform hosted by a secure cloud, to develop applications of all sorts using that platform. Cloud service providers host, manage and give the necessary platform along with resources and other tools for application development as a service.

C. Infrastructure as a Service (IaaS)

Cloud service providers provide a full infrastructure which consists of data centers, servers, networking equipment and other things as IaaS. Iaas removes the need to have above mentioned assets and hundreds of people to maintain these assets at users place. Infrastructure-as-a-service (IaaS) offers you all the resources as and when you need it. Instead of making heavy investments in IT networks and the manpower needed to run it all together, IaaS vendors let you focus on your business while the entire infrastructure, resources, equipment, IT expertise, hardware, software and manpower is managed by vendors themselves. Following figure explains the concept of the cloud computing.

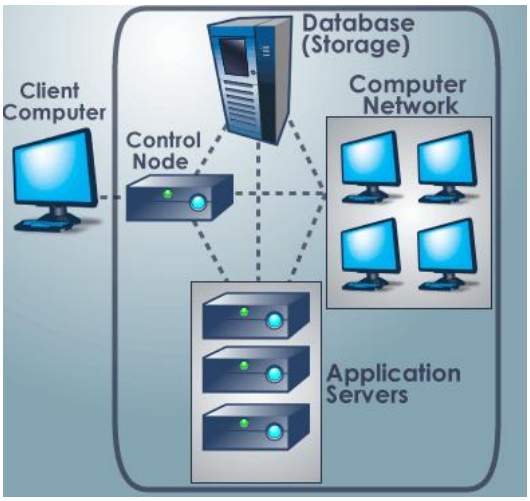

Figure 02 Architecture of a Cloud [11]

V. Summary/Conclusion

In this paper we have surveyed several papers related to the techniques used to achieve the high performance computing on a very cheap and easily available infrastructure. Various papers on the methods like Cluster computing, grid computing and Cloud computing are surveyed. We have found each of these methods very good in their areas. Every method has its advantages and disadvantages. But certainly each of them is going to put a big impact in the areas they are used in very near future. In fact some of them have already started to their effect in the consumer markets. At last, we can conclude with a statement that all of the methods discussed are going to change the scenario of the technology in near future.

VI. Acknowledgment

We would like to acknowledge our thanks to all who have helped us directly or indirectly for doing this work.

References

[1] R. Buyya(ed.), "high performance cluster computing: architectures and systems, volume 1." Prentice hall,1999.
[2] H. Jin, R. Buyya and M.A. Baker, Cluster Computing Tools, Applications, and Australian Initiatives for Low Cost Supercomputing, Monitor, The Institution of Engineers Australia (IEAust), Volume 25, No 4, December 2000.
[3] A. Kaur and M. K. Mann, "Cluster Computing", international journal of computer science and its applications.
[4] Apon, A., Buyya, R., Hai Jin and Mache, J., " Cluster computing in the classroom: topics, guidelines, and experiences" In the Proceedings of first IEEE/ACM International Symposium on Cluster Computing and the Grid, Brisbane, Qld, 2001.
[5] F. Burman, G. Fox and T. Hey, "Grid Computing-making the global infrastructure a Reality", John Wiley and son's ltd, 2002.
[6] http://computer.howstuffworks.com/grid-computing.htm
[7] http://searchdatacenter.techtarget.com/definition/grid-computing
[8] http://www.gridcafe.org/grid-in-30-sec.html
[9] J. Hurwitz, R. Bloor, M. Kaufman and F. Halper, "Cloud Computing for Dummies", Wiley Publishing, Inc, 2010.
[10] I. Gandotra, P. Abrol, P. Gupta, R. Uppaland and S. Singh,"Cloud Computing Over Cluster, Grid Computing: a Comparative Analysis" journal of Grid and Distributed Computing, Volume 1, Issue 1, 2011, pp-01-04.
[11] http://computer.howstuffworks.com/cloud-computing/cloud-computing.htm.
[12] M. A. Vouk, "Cloud Computing – Issues, Research and Implementations**",** Journal of Computing and Information Technology - CIT 16, 2008, 4, 235–246
[13] Sriram, I. & Khajeh-hosseini, A., "2010. Research Agenda in Cloud Technologies." In 1st ACM Symposium on Cloud Computing, SOCC 2010.
[14] http://www.contextis.com/research/white-papers/assessing-cloud-node-security/
[15] Qi Zhang, Lu Cheng, Raouf Boutaba: Cloud computing: state-of-the-art and research challenges. J. Internet Services and Applications 1(1): 7-18 (2010)

Mr.Nilesh C. Thakkar is an undergraduate student at G H Patel College of Engineering & Technology, Vallabh Vidyanagr, Gujarat, INDIA. He is pursuing his bachelor degree in Computer Engineering form there. The institute is affiliated with Gujarat Technological University, Ahmedabad, Gujarat, INDIA.

Professor Nitesh M. Sureja is a Diploma, B.E., and M.E., all in Computer Engineering. He has an industrial experience of four years in Raj Infotech as a Sales and Service engineer of Computer Systems and Peripherals. He worked in CCET Wadhwan City (Saurashtra University) as Lecturer. He also worked in SPCE Visnagar (Hemchandracharya North Gujarat University) as Lecturer and incharge of the Department of Computer and IT. His current profile comprises of Associate Professor (Computer Engineering) G.H. Patel College of Engineering and Technology, Vallabh Vidyanagar. He is a life member of ISTE and member of IEANG and IMETE. He has been offering his services as reviewer of various national and international Journals such as IJCSI and UBICC, and reviewer of various international and national conferences. He has guided more than 50 UG Projects and 2 PG. dissertations. He has published 4 International Journal papers and 3 Conference papers. His fields of interest and research are Swarm Intelligence, Nature and Bio inspired Computing and Artifical Intelligence.