

# A Case Study of High Performance of Computing Clusters in Cloud Computing

<sup>1</sup>Shaik Hameeda & <sup>2</sup>Dr. R. P. Singh

<sup>1</sup>Research Scholar, Sri Satya Sai University ,Bhopal (India)

<sup>2</sup>Vice Chancellor. SSSUTMS, Bhopal (India)

---

## ARTICLE DETAILS

### Article History

Published Online: 28 January 2018

### Keywords

High performance, application, High Performance computing

---

## ABSTRACT

*A high performance application requires high processing power to compute highly intensive and complex applications for research, engineering, medical and academic projects. In the traditional way, an organization will have to pay very high costs to run an HPC (High Performance computing) application. The organization has to purchase highly expensive hardware for running an HPC application and maintaining it afterwards. The HPC resources on the company premises may not satisfy all the demands of scientific application where resources may not suitable for corresponding requirements. Considering the case of SMEs (small and medium enterprises), an increasing demand is always challenging. Cloud computing is an on-demand, pay-as-you-go- model, that offers scalable computing resources, unlimited storage in an instantly available way.*

---

## INTRODUCTION

High performance computing is more and more important for the economic and technological development. The high performance computing also becomes an indicator to measure the power of a country. Therefore, it is important and meaningful to improve the performance and universality of high performance computing. In addition, cloud computing can provide users with easier computing service through various hardware devices. Applying high performance computing to cloud computing become a meaningful challenge for many scientists and professional talents.

High performance computing (HPC) refers to the computing system, including several processors as part of a single machine or a cluster of several computers as an individual resource. High performance computing owes its feature of high speed computing to its great ability to process information. Therefore the main methodology that is currently applied to high performance computing is parallel computing. In short, high performance computing is legendary for its processing capacity. For instance, it is shown by the latest analysis that machines can perform 1015 floating point operation per second. In a mesh network,[1] the structure of systems can improve the speed of host communication by shortening the physical and the logical distance between the network nodes. Although the network topology and hardware play a necessary role in the high performance computing system, it is the operating system and application software that makes the system so effective and usable. A control node, the interface between system and client computers, manages the distributed computing workload. There are two models for task execution in high performance computing environments: SIMD (Single Instruction Multiple Data) and MIMD (Multiple Instruction Multiple Data). SIMD will execute the same computing instructions and operations across multiple

processes at the same time. MIMD uses multiple processors to asynchronously control multiple instructions, achieving space parallelism. However, no matter which model is employed, the principal of a high performance system is consistent. The operation of a high performance unit (referring to several processors as part of a single machine or a cluster of several computers) is treated as a single computational resource, putting requests to various nodes. The high performance computing solution is an independent unit that is specifically designed and deployed as powerful computing resource.[2]

The birth of high performance computing can be traced back to the start of commercial computing in the 1950's. At that time, the only type of commercially available computing was mainframe computing. One of the main tasks required was billing, a task that almost every type of business needs to perform and is conveniently run as a batch process. Batch processing allows a sequence of several programs or "jobs" to be run without manual intervention. Thus once a job has completed, another job would then immediately start. Since no interaction with the administrator is required, jobs are executed in sequence without the delays created by human interaction. Batch processing saves processing time that is normally wasted with human interaction[3].

## HPC APPLICATIONS IN CLOUD

With increasing demand for high performance, efficiency, productivity, agility and lower cost, since several years, information communication technology are dramatically changing from static silos with manually managing resources and applications, towards dynamic virtual environments with automated and shared services, i.e. from silo-oriented to service oriented architectures. A "traditional" cloud offers features that are attractive to the general public. These services comprise single, loosely coupled instances (an instance of an

OS running in a virtual environment) and storage systems backed by service level agreements (SLAs) that provide the end user guaranteed levels of service. [4]These clouds offer the following features:

Instant availability – Cloud offers almost instant availability of resources.

Large capacity – Users can instantly scale the number of applications within the cloud.

Software choice – Users can design instances to suit their needs from the OS up.

Virtualized – Instances can be easily moved to and from similar clouds.

Service-level performance – Users are guaranteed a certain minimal level of performance.

Although these features serve much of the market, HPC users generally have a different set of requirements:

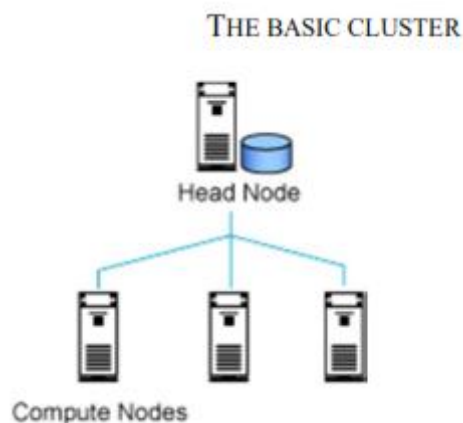
- Close to the “metal” – Many man-years have been invested in optimizing HPC libraries and applications to work closely with the hardware, thus requiring specific OS drivers and hardware support.
- User space communication – HPC user applications often need to bypass the OS kernel and communicate directly with remote user processes.
- Tuned hardware – HPC hardware is often selected on the basis of communication, memory, and processor speed for a given application set.
- Tuned storage – HPC storage is often designed for a specific application set and user base.
- Batch scheduling – All HPC systems use a batch scheduler to share limited resources.

Depending on the user’s application domain, these two feature sets can make a big difference in performance. For example, applications that require a single node (or threaded applications) can work in a cloud. In this case, the user might have a single program that must be run with a wide range of input parameters (often called parametric processing), or they might have dataflow jobs, such as the Galaxy suite used in biomedical research. These types of applications can benefit from most cloud computing resources. Some applications can utilize highly parallel systems but do not require a high-performance interconnect or fast storage. One often cited example is digital rendering, in which many non-interacting jobs can be spawned across a large number of nodes with almost perfect scalability. These applications often work well with standard Ethernet and do not require a specialized interconnect for high performance.[5]

## HPC IN CLUSTERS

A high performance computer appropriate for most small and medium-sized businesses today is built from what are basically many ordinary computers connected together with a

network and centrally coordinated by some special software. Because the computers are usually physically very close together, the common term for a high performance computer today is a cluster[5]. The basic structure of a cluster based computing is shown in the figure[6]



High-Performance Computing (HPC) clusters are characterized by many cores and processors, lots of memory, high-speed networking, and large data stores – all shared across many rack-mounted servers. User programs that run on a cluster are called jobs, and they are typically managed through a queuing system for optimal utilization of all available resources. An HPC cluster is made of many separate servers, called nodes, possibly filling an entire data center with dozens of power-hungry racks. HPC typically involves simulation of numerical models or analysis of data from scientific instrumentation. At the core of HPC is manageable hardware and systems software wrangled by systems programmers, which allow researchers to devote their energies to their code. A successful HPC cluster is a powerful asset for an organization. At the same time, these powerful racks present a multifaceted resource to manage. If not properly managed, software complexity, cluster growth, scalability, and system heterogeneity can introduce project delays and reduce the overall productivity of an organization. A successful HPC cluster requires administrators to provision, manage, and monitor an array of hardware and software components. Clusters are the predominant type of HPC hardware these days; a cluster is a set of MPPs (In clusters, also known as massively parallel processors (MPPs), they don't share the same memory;). A processor in a cluster is commonly referred to as a node and has its own CPU, memory, operating system, and I/O subsystem and is capable of communicating with other nodes.[7]

## CLUSTERING

The term "cluster" can take different meanings in different contexts. There are three types of clusters:

- **Fail-over clusters**
- **Load-balancing clusters**
- **High-performance clusters**

### Fail-over clusters

The simplest fail-over cluster has two nodes: one stays active and the other stays on stand-by but constantly monitors the active one. In case the active node goes down, the stand-by

node takes over, allowing a mission-critical system to continue functioning.

### Load-balancing clusters

Load-balancing clusters are commonly used for busy Web sites where several nodes host the same site, and each new request for a Web page is dynamically routed to a node with a lower load.

### High-performance clusters

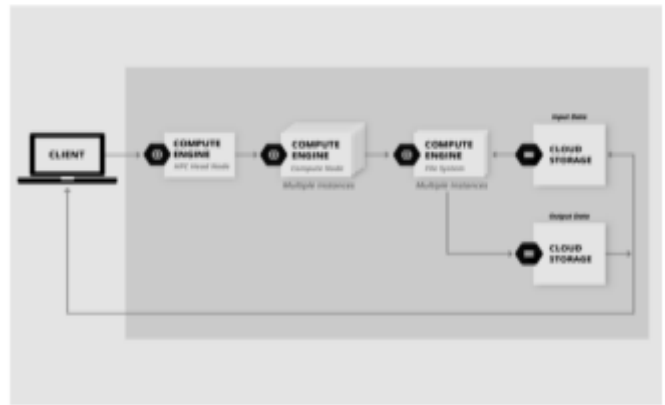
These clusters are used to run parallel programs for time-intensive computations and are of special interest to the scientific community. They commonly run simulations and other CPU-intensive programs that would take an inordinate amount of time to run on regular hardware.

Some features of clusters are as follows

- Clusters are built using commodity hardware and cost a fraction of the vector processors. In many cases, the price is lower by more than an order of magnitude.
- Clusters use a message-passing paradigm for communication, and programs have to be explicitly coded to make use of distributed hardware.
- With clusters, you can add more nodes to the cluster based on need.
- Open source software components and Linux lead to lower software costs.
- Clusters have a much lower maintenance cost (they take up less space, take less power, and need less cooling).[8]

### HIGH PERFORMANCE COMPUTING CLUSTER IN A CLOUD ENVIRONMENT

High Performance Computing clusters can be created on the Google Cloud Platform by utilizing Google Compute Engine VMs and Google Cloud Storage[3]. By running HPC workloads in Google's Cloud, customers can augment on-premise HPC clusters or run all their jobs in the cloud[9]



The compute portion of the HPC cluster consists of a Head Node running scheduling and management software on a Google Compute Engine VM. The compute/worker nodes are also running on Google Compute Engine VMs. Instances can be selected to match the workload. Choices include Standard, High Memory or High CPU instances in 1, 2, 4, 8 or 16 core sizes. Instances can also be added or deleted depending on the resources needed. The user has a choice of various commercial packages or open source software components to create the cluster. Compute Engine VMs can also be used to create a file system for the cluster. Two popular options are NFS and Gluster. This is an optional component as the Compute Nodes can also access Google Cloud Storage directly.[10]

### CONCLUSION

High performance computing and cloud computing. The feature, high speed computing capacity, can be used to process big data and complicated information. Cloud computing is based upon pay-per-use model, offering visualized resources through Internet service. The idea of integration of the high performance computing and cloud computing comes up for better computing service. high-performance scientific computing on public clouds. The adoption of Cloud computing as a technology and a paradigm for the new era of computing has definitely become popular and appealing within the enterprise and service providers. It has also widely spread among end users, which more and more host their personal data to the cloud. For what concerns scientific computing, this trend is still at an early stage.

### REFERENCES

- [1] R-Cloud. <http://www.r-hpc.com/>
- [2] Sabalcore HPC Cloud Service. [http://www.sabalcore.com/why\\_on\\_demand.html](http://www.sabalcore.com/why_on_demand.html)
- [3] UniCloud. <http://www.univa.com/products/unicloud.php>
- [4] Ahronovitz, M., et al. (2010). Cloud computing use cases. A white paper produced by the Cloud Computing Use Case Discussion Group. Retrieved from <http://groups.google.com/group/cloud-computing-use-cases>.
- [5] Altintas, I., Berkley, C., Jaeger, E., Jones, M., Ludascher, B., & Mock, S. (2004). Kepler: An extensible system for design and execution of scientific workflows. Proceedings of the 16th International Conference on Scientific and Statistical Database Management (SSDBM), Santorini Island, Greece. Retrieved from <http://keplerproject.org>
- [6] Costaa,P.,CruzB,A.:Migration to Windows Azureanalysis and comparison.In:CENTERIS 2012-Conference on ENTERprise Information Systems/HCIIST 2012 – International Conference on Health and Social Care Information Systems and Technologies.Procedia Technol. 5, 93–102 (2012)
- [7] Mondéjar, R., Pedro, C., Carles, P., Lluís, P.: CloudSNAP: a transparent infrastructure for decentralized

- web deployment using distributed interception. *Future Gener. Comput. Syst.* 29, 370–380 (2013)
- [8] Marozzo, F., Lordan, F., Rafanell, R., Lezzi, D., Talia, D., Badia, R.M.: Enabling cloud interoperability with COMPSs. In: Kaklamanis, C., Papatheodorou, T., Spirakis, P.G. (eds.) *Euro-Par 2012*. LNCS, vol. 7484, pp. 16–27. Springer, Heidelberg (2012)
- [9] Brecher, C., Gorgels, C., Kauffmann, P., Röthlingshöfer, T., Flodin, A., Henser, J.: *ZaKo3Dsimulation possibilities for PM gears*. In: *World Congress on Powder Metal*, Florenz (2010)
- [10] *A Performance and Cost Analysis of the Amazon Elastic Compute Cloud (EC2) Cluster Compute Instance* by Fenn, M., Holmes, J., Nucclarone, J., and Research Computing and Cyberinfrastructure Group, Penn State University