HPC as a Service on Cloud: A Market Oriented Approach

Manoj H. Radadiya[1], Ashok K. Yadav[2], Meet A. Tilva[3], Dr. Vandana Rohokale[4]

[1, 2, 3] PG Student, Network Security, GTU PG School, Ahmedabad, Gujarat, India
[4] Dean R&D, SKNSITS, Lonavala, Maharashtra, India

ABSTRACT

Cloud is a latest technology used in most of enterprise. Every enterprise is moving towards the cloud platform. On the other hand High Performance Computing (HPC) is a necessity for many industries for their data analytics or huge computations tasks. HPC needed large set of resources so, all enterprises cannot afford supercomputers. HPC as a service (HPCaaS) can deal with these issues. Moving towards HPC systems on the cloud environment is cost effective for small scale industry. Costing is major area in HPC as a service on cloud. This paper addresses some issues and challenges. This paper provides solution for one of these problems as a costing model for HPC as a Services on the cloud environment so that it will enable the small organization to have their own HPC systems at the low costs.

Keyword: - *Cloud Computing, HPCaaS (HPC as a Service), Costing Model for HPCaaS;*

1. INTRODUCTION

1.1 What is Cloud Computing?

Cloud Computing is a technology that maintain data and applications using the internet and central remote servers. Cloud computing allows customer to use applications without any installation and access their files from anywhere with internet access. This type of technology allows for much more efficient computing by centralizing data storage, processing and bandwidth.

An example of cloud computing are Gmail, Hotmail etc. All users need is only an internet connection and they can sending emails. The main server and email management software is on the cloud and is managed by the cloud service provider Google, Hotmail etc. The customers gets to use the software individual and enjoy the services. The analogy is, 'If you need milk, would you buy a cow?' All users required is to get the services of using the software or hardware of the computer like sending emails etc. Just to get services why should a customer buy a (cow) software /hardware? [1]

Cloud computing provides many attractive benefits for enterprise and users. Main benefits of cloud computing include [2]:

Self-Service Provisioning: Users can take resources for any type of workload on demand.
Elasticity: Enterprise can scale their computing requirements increase and decrease.
Pay per Use: Computing resources are measured at a granular level, users have to pay only for the resources which are used by them.

Cloud computing have three deployment model: Private cloud, Public cloud and Hybrid cloud. In the Private cloud model, cloud is located at internal organization for internal users. This model offers very high finance with easy user access, while doing management, control and security. In the public cloud model, a third-party provider provides cloud service over the Internet. Public cloud model are available on-demand, within the minute or the hour. Users only pay for the resources which are used by them. Well known public cloud providers are Microsoft Azure, Google Compute Engine and Amazon Web Services (AWS).The Hybrid cloud model is a combination of public cloud

model and private cloud model. Enterprises can run on their private cloud services and if they need an extra resources then they can use a public cloud services. The goal of hybrid cloud services is to create an automated, scalable and unified environment which has advantage of a public cloud model and private cloud model [2].

Although cloud computing has changed over time to time. Cloud Computing provides three service model: Infrastructure as a Service, Platform as a Service and Software as a Service. The IaaS providers for example - AWS gives a virtual server instance and storage, as well as APIs (application program interfaces) and allow users to migrate workload to a virtual machine (VM). IaaS providers allow users to small, medium, large, extra-large compute optimized instances. In the PaaS model, service providers host or install development tools or software on their infrastructures. Users can access these tools using APIs, Web portals or software gateway via internet. PaaS model is used for software development and many PaaS service providers will host the software after it's developed. Well known PaaS service providers are Amazon Elastic Beanstalk, Google App Engine etc. The SaaS model is a distribution model and gives software applications via the Internet, in other words these are called as Web services. Microsoft Office is a best example of SaaS service which is offering productive software and email services. Users can use SaaS applications or services from anywhere using a computer or mobile device which has Internet access [2].

1.2 What is HPC?

High Performance Computing (HPC) has been a major resource for many enterprise from many years to resolve a larger range of problems which needs huge range of computing resources. HPC systems are mostly made for a specific task that can itself present issues around their restricted application by a small set of those within an enterprise. This has allow differences in adjusting the total cost of ownership of these systems.

Most of us heard many time about supercomputing or super computer that deal with human's largest problems of science and technology. These machine uses virtualization technology inside it and high infrastructure like 10000 of processor on single machine because of this features super computers are very expensive. This type of computing is related to HPC. An HPC system is more complex than our desktop systems or computers. The basics of HPC aren't much more difficult to learn because there are many enterprise either big or small can provide the information or help to anyone [3].

Public and Private clouds have changed our thinking about computing and also HPC. Public clouds provides huge numbers of compute nodes which can be easily available for users with some amount of money for a period of time. For HPC jobs which have require frequently, not always then this funda can be used for huge money savings. Private clouds provides self-service, flexibility and cost savings to on premise computing. HPC sites are giving productivity and cost efficiency using private cloud in HPC [4].

1.3 What is HPCaaS?

HPC system is remote and batch. The desktop system is local and interactive. The cloud system is remote and interactive. So if service provider merge HPC and Cloud then they will get the system which is interactive, remote and batch. The main issue is it takes the hours or days to run an HPC job or transfer data sets. But people have a solutions for example, an engineer can solve his problems so that he complete overnight. He can start a batch before leaving office and he have answers in the morning. In this scenario one set of HPC tools deal with the batch work. Another set of desktop tools deal with interactive analysis [5].

2. STATE-OF-THE ART

As our Research area is based on HPC as a Services on the cloud computing platform, So Authors refer many other research papers from this specific domain. While referring a research papers authors are getting some basic of clouds and High Performance Computing and also get some useful research areas in this specific domain.In this chapter authors discuss the literature review about our interested domain.Table-1 shows the main contain about the all the papers which authors refereed during survey on HPCaaS on the Cloud.

Development of Generic Virtualization Service (GVirtuS) [6] in this paper author developed a framework for GPU (Graphics Processing Units) virtualization based on split driver for cloud virtualization Technology. Author focused on GPU Virtualization that allows for the development of split-drivers in a straight forward fashion by requiring to only implement stub routines in the frontend and service routines in the backend. A SDN Empowered Task Scheduling System [7], a SET system which targets dynamic configuration and monitoring of cloud networking

using SDN that improve the performance of HPC application and task scheduling of HPCaaS. SETSA (SDN-Empowered Task Scheduling Algorithm) is used as task scheduling algorithm that monitor network bandwidth, improve the bandwidth efficiency and reduce task turnaround time. Author mention new SETSAW (SETSA Window) mechanism as an improvement of SETS algorithm. Using SETSA get turnaround time is reduced till 75% by better utilizing the network bandwidth.

UnaCloud Platform Architecture for HPC [8], a Platform as a Service (PaaS) model constructed over a desktop-based Cloud infrastructure for developing high performance computing applications taking advantage of unused resources opportunistically. UnaCloud PaaS represents a next step on opportunistic use of unused infrastructures. They use the High- Performance Linpack Benchmark for implementation. Dynamic Resource Management in a HPC [9], deal with utilization of cluster for hardware resources in data center. In this paper author also mention AR based periodic decomposition algorithm for prediction accuracy of jobs arrival rate, Queuing model for QoS. Due to they used a hybrid environment, they get least 30% resources utilization.

Energy-Efficient Thermal-Aware Autonomic Management of Virtualized HPC Cloud Infrastructure [10], provide active cross layer thermal management solution and Proactive application centric strategy for VM allocation. Author Guarantees QoS. They gives solution significantly contributes to energy efficient 12% reduce & optimize application performance in VM allocation. Moldable Job Scheduling for HPC as a Service [11], paper is related job submission and scheduling issues. Author provides better system performance, and decreases user burden. As a result author get 75% performance improvement than the traditional rigid processor allocation method and 3% improvement than previous moldable method.

HPC Performance and Energy-Efficiency of the OpenStack Cloud Middleware [12], this paper abstract from the specifics of a single architecture and the benchmarks were run using two different hardware configurations, based on Intel and AMD processors. Author used HPCC (which includes the reference HPL benchmark). There is a substantial performance impact Introduced by the cloud middleware layer across the considered hypervisors. Enterprise HPC on the Clouds [13], discuss the major issues that must be addressed in order to make clouds viable for enterprise HPC and review research based on existing or simulated cloud systems that hints as to how the problems can be solved. Author conclude that Integration & customization handle by existing technology but security & availability will need the improvement of existing technology to achieve scalability.

The VCL Cloud Architecture [14], discuss a very successful production level architecture and policy framework for supporting HPC services within a more general cloud computing infrastructure. Author used VCL Cloud Architecture. They mention non-HPC & HPC cloud computing architecture can provide both in a cost effective manner in this paper. Cloud Computing Based HPC: A Framework for Ethiopian Universities [15], handling of large data sets in the cloud. Author discussed private cloud infrastructure for each university using OpenStack by considering HPC and create a collaboration mechanism for each University private clouds throughout the country using federated cloud architecture. High Performance Computing in a Cloud Using OpenStack [16], deals with some challenging features in cloud computing, such as performance. Methods proposed in this paper allow customizing and optimizing cloud infrastructure to increase its performance and to meet certain requirements, drawn from the analysis of case study simulation codes.

3. OPPORTUNITIES AND CHALLENGES

Several researches and studies compared the performance of HPC applications on the cloud with on-premises infrastructure and reached to this conclusion that clouds of today cannot compete with supercomputers. This section describes the primary reasons.

Multi-tenancy [17]: It is one of the characteristics of the cloud. It is also one of the profit making features of the cloud for cloud providers. It enables cloud providers to share resources between multiple users. Degree of multi-tenancy refers to the number of users sharing a same resource on the cloud. By increasing the degree of multi-tenancy, cloud providers are able to overprovision the resources to users. Overprovisioning allows cloud providers to maximize benefits, though with the risk of reducing QoS. Nevertheless, multi-tenancy is in direct contrast with what HPC needs. HPC applications demand direct access to dedicated hardware using some sort of batch scheduling While shared resources on the cloud complicates the performance of HPC applications. Shared resources make the simultaneous applications compete over resources resulting in lower performance of the applications as well as the

Table 1 – State of the Art for HPCaaS on the Cloud

Research Paper	Year	Mechanism	Key Points
Integration of High-Performance Computing into Cloud Computing Services [6]	2010	VCL manager software	Non-HPC & HPC cloud computing architecture can provide both in a cost effective manner
A General-Purpose Virtualization Service for HPC on Cloud Computing: An Application to GPUs [7]	2012	GPU virtualization based on split driver virtualization Technology	1.Virtual parallel high performance file system 2. A MPI channel for low-latency communication among virtual machines running in a virtual HPC cluster environment.
Energy-Efficient Thermal Aware Autonomic Management of Virtualized HPC Cloud Infrastructure [8]	2012	An innovative application-centric energy-aware strategy for VM allocation is proposed	Proactive VM allocation model that uses the concept of heat imbalance and place workload on server.
Enterprise HPC on the Clouds [9]	2012	Issues	Security & Availability will need the improvement of existing technology to achieve scalability for thousands of nodes
Building Platform as a Service for High Performance Computing over an Opportunistic Cloud Computing [10]	2013	Deploy and manage virtual clusters over the available infrastructure	1. Will increase the amount of platforms and software modules on UnaCloud, 2. Planning to implement UnaCloud SaaS Model
Dynamic Resource Management in a HPC and Cloud Hybrid Environment [11]	2013	A dynamic resource allocation algorithm	AR based periodic decomposition algorithm for prediction accuracy of jobs arrival rate
Moldable Job Scheduling for HPC as a Service [12]	2014	Processor Allocation for Moldable Job Scheduling	Better system performance , less user burden, 75% performance, 3% improvement than previous moldable method
HPC Performance and Energy-Efficiency of the OpenStack Cloud Middleware [13]	2014	HPCC	Larger scale experiments over the cloud environment
Cloud Computing Based HPC: A Framework for Ethiopian Universities [14]	2014	Swift, Nova, Apache Hadoop, and MapReduce	Private cloud deployment model that enables HPC, there are issues related to cost, performance and other parameters
High Performance Computing in a Cloud Using OpenStack [15]	2014	Scientific Cloud Based on OpenStack	Designing and Implementing one or more of the proposed solutions from the method set.
ASETS: A SDN Empowered Task Scheduling System for HPCaaS on the Cloud [16]	2015	Software Define Network and HPC	1. Provide proof of concept, 2. Performance evaluation, 3. Adding the Cost Model, 4. Distributed file system architectures, 5. Estimate the overhead cost with the cost of virtualization Compared with a traditional HPC platform.

network bandwidths.

Virtualization Overhead [17]: Virtualization plays a key role in the cloud helping the cloud to have rapid elasticity, resource pooling and flexibility. However, virtualization and in particular the hypervisor adds unwanted overhead by adding a software layer and preventing applications to have direct access to the hardware resources. This virtualization overhead is not the same for all types of hardware. For example, because of the hardware support, virtualization overhead for processors is significantly less than the overhead of network virtualization. For some hardware types such as GPUs, it is often more efficient to pass through GPUs than to have virtual GPUs.

Network Bandwidth and Latency [17]: Network interconnects and I/O resources in the cloud are shared between several tenants. Therefore, the bandwidth and latency of the network may not be predictable. The bandwidth in most cases is much less than what is expected. Moreover, the latency of the network on the cloud is not stable. Therefore, we will see performance degradation for HPC applications and in particular data intensive ones.

Other Challenges are [18]:

1. Costing or Pricing Model
2. Submission Model **[**Job queuing and reservations towards VM deployment**]**
3. Brining of data in and out from the cloud which is costly and results in data lock-in
4. Security, regulatory compliance and various "-ilities" [Performance, availability, Business continuity, SLA and so on]

4. PROPOSED WORK

In above two sections, authors discussed about the state of the art and issues and challenges in High Performance Computing as a services on the cloud computing environment. As authors find there are some drawbacks in current scenario and research which is done till the date in this domain. Authors will focus on the ASETS: A SDN Empowered Task Scheduling System [16] for more deep research as a future work. As author mention there are many loopholes in this paper. Authors have a three option for the research that are as under:

1. Proof of Concept
2. Add a Costing Model
3. Use a Distributed File System

As authors are early discussed in section III there is a challenges in this is domain and one of them is "Costing or Pricing Model". Based on research work in [16] and other challenges-1 [19], the authors are going to focus on the "Costing Model" for this domain.

Authors decide the title for the future research that is "Costing Model for High Performance Computing as a Services [HPCaaS] on the Cloud." For further work, authors make one costing model for the HPCaaS for the consumer as authors are giving them an HPCaaS on the cloud platform. The costing model will done based of various parameters like [19],

1. Processor, Co-Processor, Network and Hardware,
2. Power Consumption and Cooling Infrastructure,
3. Software Licenses - including Operating System and Software
4. Racks, Cables, and Installation,
5. Maintenance
6. Staffing

5. CONCLUSIONS.

With emergence of new technology there is need to combine two different domains like High Performance Computing and Cloud Computing. With the combination of these two domains service provider and users can get all the benefits of both the domain which are useful for emerging cloud computing environment. Users will be able to get high computing power using the cloud services with lower price. This is very useful mechanism for a small scale

and medium scale enterprises. In this paper the authors have done the survey and come up with some challenges of HPCaaS on the cloud domain. This work will be extended for "Costing Model" development.

6. REFERENCES

1. Wikinvest -"Cloud Computing". http://www.wikinvest.com/concept/CloudComputing
2. Margaret Rouse, "Search Cloud Computing", TechTarget. http://searchcloudcomputing.tech-Target.com/definition, 2015
3. Inside HPC, "what is high performance computing?" http://insidehpc.com/hpc-basic- training/what-is-hpc/
4. Adaptive Computing, "Converging Worlds: HPC, Big Data, and Cloud", Sept20013. http://www.adaptiv-Ecomputing.com/blog-cloud/converging-worlds-hpc-bigdata-and-cloud/
5. Dave Hudak, "From On Demand to AweSim: Merging HPC with the Cloud", OHTECH, Oct2013. http://www.ohtech.org/blog/ondemand_awesim_merging_hpc_cloud#.VhDwUvmqqkq
6. Mladen A. Vouk, Eric Sills, and Patrick Dreher, "Integration of High-Performance Computing into Cloud Computing Services" Handbook of Cloud Computing, DOI 10.1007/978-1-4419-6524-0_11 Springer Science Business Media, LLC-2010.
7. Raffaele Montella, Giuseppe Coviello, Giulio Giunta, Giuliano Laccetti, Florin Isaila3, and Javier Garcia Blas, "A General-Purpose Virtualization Service for HPC on Cloud Computing: An Application to GPUs" Springer-Verlag Berlin Heidelberg-2012.
8. Ivan Rodero, Hariharasudhan Viswanathan, Eun Kyung Lee, Marc Gamell, Dario Pompili, Manish Parashar, "Energy- Efficient Thermal-Aware Autonomic Management of Virtualized HPC Cloud Infrastructure" Springer Science Business Media B.V-2012.
9. Ioannis A. Moschakis and Helen D. Karatza, "Enterprise HPC on the Clouds" Computer Communications and Networks, DOI 10.1007/978-1-4471-2236-4_12, Springer- Verlag London Limited-2011.
10. German A. Sotelo, Cesar O. Diaz, Mario Villamizar, Harold Castro, Johnatan E. Pecero, and Pascal Bouvry, "Building Platform as a Service for High Performance Computing over an Opportunistic Cloud Computing" Springer International Publishing Switzerland-2013.
11. Miao Chen, Fang Dong, and Junzhou Luo, "Dynamic Resource Management in a HPC and Cloud Hybrid Environment" Springer International Publishing Switzerland-2013.
12. Kuo-Chan Huang, Tse-Chi Huang, Mu-Jung Tsai, and Hsi- Ya Chang, "Moldable Job Scheduling for HPC as a Service" Future Information Technology, Lecture Notes in Electrical Engineering 276, Springer-Verlag Berlin Heidelberg-2014.
13. Sébastien Varrette, Valentin Plugaru, Mateusz Guzek, Xavier Besseron and Pascal Bouvry, "HPC Performance and Energy-Efficiency of the OpenStack Cloud Middleware" 43rd International Conference on Parallel Processing Workshops-2014.
14. Samuel Fentahuen, Sreenivas Velagapudi, "Cloud Computing Based HPC: A Framework for Ethiopian Universities" IJIRCCE Vol. 2, Issue 5, May-2014.
15. Roman Ledyayev and Harald Richter, "High Performance Computing in a Cloud Using OpenStack" The Fifth International Conference on Cloud Computing, GRIDs, and Virtualization-2014.

16. Saba Jamalian, Hassan Rajaei, "ASETS: A SDN Empowered Task Scheduling System for HPCaaS on the Cloud" IEEE International Conference on Cloud Engineering-2015.

17. Saba Jamalian, Hassan Rajaei, "Data-Intensive HPC Tasks Scheduling with SDN to Enable HPC-as-a-Service", IEEE 8th International Conference on Cloud Computing – 2015.

18. Dejan Milojicic, "High Performance Computing (HPC) in the Cloud", Computing Now, HP Labs, Sept2012. http://www.computer.org/web/computingnow/archive/september2012

19. "Managing the Real Cost of On-Demand Enterprise Cloud Services with Chargeback Models.", White Paper of Cisco Systems-2010.