

THE FEASIBILITY STUDY OF RUNNING HPC WORKLOADS ON COMPUTATIONAL CLOUDS

PURVI PATHAK*, KUMAR R

School of Computing Science and Engineering, VIT University, Chennai, Tamil Nadu, India. Email: purviapathak@gmail.com

*Received: 14 June 2017, Revised and Accepted: 16 June 2017***ABSTRACT**

High-performance computing (HPC) applications require high-end computing systems, but not all scientists have access to such powerful systems. Cloud computing provides an opportunity to run these applications on the cloud without the requirement of investing in high-end parallel computing systems. We can analyze the performance of the HPC applications on private as well as public clouds. The performance of the workload on the cloud can be calculated using different benchmarking tools such as NAS parallel benchmarking and Rally. The workloads of HPC applications require use of many parallel computing systems to be run on a physical setup, but this facility is available on cloud computing environment without the need of investing in physical machines. We aim to analyze the ability of the cloud to perform well when running HPC workloads. We shall get the detailed performance of the cloud when running these applications on a private cloud and find the pros and cons of running HPC workloads on cloud environment.

Keywords: Cloud computing, OpenStack, High-performance computing applications, Private cloud.

© 2017 The Authors. Published by Innovare Academic Sciences Pvt Ltd. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>) DOI: <http://dx.doi.org/10.22159/ajpcr.2017.v10s1.20507>

INTRODUCTION

With the increase in popularity of cloud computing many applications in various fields of computing such as social networking and e-commerce are being deployed on the cloud, so there is a possibility that computational clouds can also be utilized for scientific and research applications. High-performance computing (HPC) applications mainly run on supercomputers which usually use distributed computing, dedicated processes in a cluster of computers, or grid computers. The amount of money required to set up such infrastructure is very high. A huge amount of skilled workforce is also required to set up and maintain such infrastructure. Hence, to run applications that require continuous availability, deployment of HPC applications on clouds may be advantageous. Grid technology is used by the scientific community to run and test their HPC applications. When using public grids the researchers have to give a request to use the resources and many times they are not allocated the resources for minor projects. Among other problems, most grids run on a single environment that are a grid makes use homogeneous environment only and is not flexible to accommodate multiple environments depending on the HPC application.

Cloud computing looks like an attractive option to run HPC applications, due to its attractive features which are inclusive of high availability, scalability, pay-per-use model, and the freedom to choose the environment for the installation of HPC applications. Cloud computing makes use of server virtualization. We can launch multiple virtual machines to run the HPC applications. HPC applications may also require scalable environment where the applications requirements can drastically increase or decrease depending on the processes that are being performed, cloud computing has a feature called scalability where the load balancer itself scales the amount of compute power depending on the usage of the resources by the HPC application. Some HPC applications, especially in medical and environmental fields, require high availability where even a little downtime can have catastrophic results. Cloud computing provides an attractive solution to this problem. Most cloud computing platforms are highly available and follow the SLA which is a service level agreement. In the case of cloud computing, if one server goes down in the data center, other unused servers will come up in a matter of seconds ensuring minimum downtime.

In this paper, we shall install an HPC application called open HPC on multiple virtual machines on the OpenStack environment in such a way that one virtual machine will act as the master node and the other virtual machines will act as the slave nodes. The master node will control all the slave nodes, where all the tasks will be given to the slave nodes by the master node and they will perform various tasks in a parallel order and send the output obtained to the master nodes. We shall then perform benchmarking on these nodes to determine whether the use of private cloud environment OpenStack is advantageous to run HPC applications or not. We shall see the feasibility from the results of the benchmarks we have obtained. High-throughput computing (HTC) and many-task computing (MTC) also involve the use of parallel and grid-like systems to perform computation. MTC is a type of computation which tries to reduce the differences between HTC and HPC. MTC differs from HTC in the emphasis of using huge number computing resources over a small amount of time to complete multiple computational tasks. We shall also discuss about HTC and MTC in this paper.

ARCHITECTURE

There will be a graphical user interface through which the user can access the multiple cloud platforms be it OpenStack, Amazon Web Services (AWS), or Azure. After the user will enter the details such as username and password, he will be sent a one-time password (OTP) on his e-mail or mobile device. After entering the OTP, the user will be logged into the system. Keystone is the authentication and authorization interface of OpenStack, keystone will authenticate the user and authorize him to use OpenStack depending on the user's rights.

Once the user is logged into the system, he can access the interface through horizon. Now, the user will give parameters to be entered into the HPC application or even the parameters can be automatically generated by the application and the application is run on Nova. Nova is the compute node of OpenStack which manages all the features of the virtual machine such as memory allocation, virtual instance creation, and many more basic tasks such as taking snapshots and creating clones.

The HPC application is then to be run on the virtual machines. Once the application is run, the data are stored in data storage. This data is then fed to the HPC application benchmarking tools such as NAS

parallel benchmarking (NPB) or Rally. NPB is developed by the NASA supercomputer division, which is used to perform benchmarking on parallel applications and Rally is a benchmarking tool created by OpenStack which is used to perform benchmarking on OpenStack and its components. The parameters we are going to use for benchmarking are compute, that is, the usage of RAM and test if the resources available are used properly and optimally, the network performance for transmission of data and the cost analysis of the application but may not be at par with the performance by the grids or the clusters used in dedicated systems for running HPC applications (Fig. 1).

RELATED WORK

There has already been research to evaluate the performance of HPC applications on cloud. These researches have mostly been done on commercial clouds such as AWS.

Sadooghi *et al.* have evaluated the different services available on AWS such as compute, network, memory, and I/O [1]. They first measured the raw performance of the instances running on AWS and after that they measured the performance of the scientific applications running on the cloud. They even measured the performance of the Fermi cloud. They measured the storage services such as EBS and S3 provided by AWS. They came to a conclusion that the actual performance is lower than the expected performance. In this paper, they found that the performance of Fermi cloud is almost to that of AWS. However, the network throughput was better than that of AWS as it uses infiniband network cards. In this paper, they concluded that Amazon EC2 has powerful enough instances to run HPC applications.

Juve *et al.* have examined the performance and the costs of the clouds. They have analyzed the costs incurred to run the HPC workflows on the cloud [2]. The results they found indicated that if proper hardware resources similar to supercomputers are provided on the cloud, very similar results can be achieved. They compared the performance of AWS to NCSA's Abe cluster. In AWS, there was a small virtualization overhead other than that almost all the performance metrics for AWS and Abe Cluster were almost equivalent. They came to a conclusion that clouds are capable of running scientific workloads if the cloud service providers offer high-speed networks and parallel file systems.

Wang and Ng have measured the impact of virtualization on network performance based on various factors such as processor sharing, packet delay, transmission control protocol/user datagram protocol (TCP/UDP) throughput, and packet loss in Amazon EC2 virtual machines [3]. This research tells us that there are widespread processor sharing, abnormal delay variation, and unstable TCP/UDP throughput in EC2 instances. This can degrade the performance and bring up new challenges for different applications.

Walker has used macro- and micro-benchmarks to study the performance of a cluster composed of EC2 high-CPU compute nodes

and compared this against the performance of a cluster composed of equivalent processors available to the open scientific research community. The delivery of HPC performance with commercial cloud computing services such as Amazon EC2 has not yet matured [4]. The tests have shown that there is a performance gap between performing HPC computations on a traditional scientific cluster and on an EC2 provisioned scientific cluster. For cloud computing to be a viable alternative for the computational science community, the cloud service providers will be required to upgrade their service offerings, especially in the area of high-performance network provisioning, to cater to this unique class of users.

Juve *et al.* have tried to understand the best storage system and the cost of the same on Amazon EC2, for running the scientific workflow [5]. Storage systems have a significant impact on the runtime of the workflow, making it necessary to choose the best storage mechanism for running HPC applications so that the runtime of the workflow is optimal. They found out that the performance of smaller large number of files suffered from long latencies. Gluster performed well for the large number of files, whereas EC2 suffered from long latencies. We saw that performance gains could be achieved without increasing cost, or with only a small increase in cost, by provisioning a submit host inside the cloud to manage the workflow.

Vecchiola *et al.* have evaluated the usage of cloud services for HPC applications. They focus on the pay-per-use model of cloud services where we can pay for the resources according to our usage [6]. They also talk about the service level agreement and the quality of service which ensures high availability, which is necessary when running high-performance applications. Aneka is an enterprise cloud computing solution. Aneka harnesses the power of compute resources by relying on both private and public clouds and it delivers the desired quality of service to the users. In the paper, they have shown two case studies, one is the classification of gene expression data and the second one is functional magnetic resonance imaging (MRI) workflows for MRI scanning.

Jackson *et al.* compare the conventional platforms used to run HPC applications and the Amazon cloud services [7]. This study indicates that EC2 is 6 times slower as compared to a Linux cluster, whereas it is 20 times slower as compared to a modern HPC system. The variability in Amazon EC2 can be quite significant. Variability in Amazon EC2 is high due to factors such as the shared nature of the virtualized environment, the network, and the differences in the underlying non-virtualized hardware.

Castro *et al.* have presented an effort between a private and public institution to offer scientific applications as a service taking advantage of an existing infrastructure to create a private IAAS using OpenStack environment and offering scientific applications through a user-friendly interface [8]. The problems faced by the small research groups include high costs in capital investments, management and the maintenance of the infrastructure, operation and the execution of complex processes, configuration of servers and distributed systems, and the requirement of large computing capabilities and the restrictions as running HPC problems can take a huge amount of time. Hence, the research community has started to opt for cloud-based infrastructure to run their HPC applications. They have presented a collaboration strategy that was developed by two different institutions to offer scientific applications as a service to small and medium research groups so that they can run their HPC workloads on a private IAAS solution in an easily and in a transparent way. They showed how the configuration and operation of a private cloud on existing infrastructures, using open source middleware such as OpenStack can increase their usage and also how its integration with solutions such as e-clouds can facilitate the execution of HPC workloads on private infrastructure.

Dhuldhule *et al.* have tried to give a platform as a service perspective to run HPC applications on the cloud. Their architecture comprises

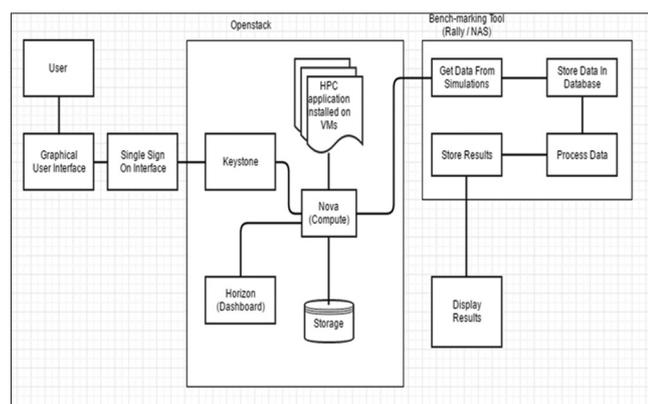


Fig. 1: Architecture

on-demand provisioning for the hardware and cloud computing environment [9]. The architecture suggested by them builds the HPC platform by creating a dedicated node or a group of nodes booted with the HPC environment without the requirement of the virtualization layer. They evaluated performance of the application across different resources and also studied the effect of virtualization on each of the nodes. They came to a conclusion that HPC applications do not give fair performance on the virtualized environment. They have presented a architecture to overcome the performance degradation that is caused by the virtualization layer. Their goal of providing a dedicated and on-demand platform for HPC workload was achieved with the architecture.

Zhao *et al.* have discussed in this paper a scripting language called Swift which can be used for parallel computing by defining operations on datasets based on logical organization [10]. Swift addresses issues occurring in large-scale loosely coupled parallel computation. Swift imposes order on a messy complex world of distributed applications which are prone to failure. It provides bifurcation between physical storage formats and logical data structures. It is a scripting language which is used for specification and composing complex workflows. Swift is a scalable runtime system which can manage and dispatch multiple tasks onto various types of parallel and distributed computation systems.

Most of the research work has been done on public clouds; in this research, we are going to try to find out the feasibility of running these HPC applications on a private cloud set up. If one research organization owns enough infrastructure, it can set up a private cloud in its organization and the private cloud environment can also be used by multiple research teams in collaboration with each other, so private cloud set-up may be advantageous for multiple research teams which have similar requirements.

APPLICATIONS

The applications that we are going to use include HPC applications such as open HPC, OpenStack cloud environment, and benchmarking tools such as NPB and Rally. These benchmarking tools are necessary to test the feasibility of running HPC workloads. The applications we are going to use are as follows.

Parameter space and test bed

In this research, we are going to use OpenStack to test HPC applications. We are going to see the performance of the application on the cloud. HPC applications are the applications that usually require high amount of resources than conventional desktop applications. HPC applications need to be run on dedicated parallel systems, which otherwise run on regular desktop systems will require days to complete. The data that are generated by HPC applications are large, so the standard file system is not enough for storing data. Many scientists also keep data that have previously been generated by the application for comparison with the newly generated data. We have preferred to use a private cloud set up because most research organizations have confidential data. When using public clouds we have to upload all the data to the cloud, after which it is processed and the results are given. However, if there is confidential data required to be processed, there is a risk of all the data falling into wrong hands. When we use a private cloud in which all the infrastructure belongs to the research organization itself, so the risks of the data being accessed and used by an unknown entity are very less.

Cloud can fulfill all of the above requirements of the HPC applications as we can use multiple VMs with large computing speed and also storage mechanisms are provided by OpenStack such as Cinder and Swift. Cinder is the block storage provided by OpenStack which stores all the VMs and the data generated by the VMs. Using cinder, we can create and manage the service that provisions and manages the storage in OpenStack [11]. Swift is the object storage mechanism provided by OpenStack. Swift is best suited for backing up the data and creating archives of unstructured data.

Benchmarking tools and applications

To test the feasibility of running HPC applications on the cloud, we are required to perform benchmarking on the HPC application running on OpenStack cloud infrastructure.

The HPC application we are going to use to test the feasibility of the cloud is open HPC [12]. Open HPC has pre-built modules that are required to deploy and manage an HPC Linux cluster which includes provisioning tools, resource management, I/O clients, development tools, and a number of scientific libraries. The main features of open HPC include system stability, performance scalability, community collaboration, traceability, validated and reproducible recipes, component interoperability, knowledge base for configuration recipes, and focusing on the user experience and convenience. We shall run massively parallel programs on open HPC to test the advantages of running these applications on the cloud.

The NPB is a small set of programs designed to help in evaluation of the performance of parallel supercomputers [13,14]. The benchmarks used in this benchmarking tool have been derived from computational fluid dynamics applications. They were developed at the NASA Ames Research Center to assess high-end parallel supercomputers.

Rally is an OpenStack project to create a framework for validating, performance testing, and benchmarking OpenStack [15]. Rally performs the task of automatically installing and configuring tempest, it also automates running tempest tests. Tempest is OpenStack's official test suite. It is a powerful tool for running a set of tests against an OpenStack cluster. Tempest runs against every patch in every project of OpenStack, which allows us to avoid merging the changes that break things.

IMPLEMENTATION PLAN

To test the applications, we first need to create an environment that will run the HPC applications. Running HPC applications require huge amount of compute power, powerful network for communication among nodes, and large amount of data storage. Hence, we need to set up such an OpenStack environment which will support the HPC application.

Once we are done with the installation of OpenStack, we will install multiple nodes of virtual machines of which one will be the master node and the others will be slave nodes. These virtual machines need to be installed based on the requirements of the HPC application. Once the virtual machines have been installed we have to check for the connection between them, once all the virtual machines are connected, we can start with the installation of the HPC application (Fig. 2).

Once the HPC application open, HPC has been installed we have to run the application with different types of loads. We have to give all types of parameters possible and run the HPC application. Once we are running the application we have to run the benchmarking tool such that it will measure the performance metrics of the application running on the cloud. We can collect all the data that have been gathered by the benchmarking tool and perform a detailed analysis on it. We have to compare the performance of the raw capacity of the cloud, the capacities

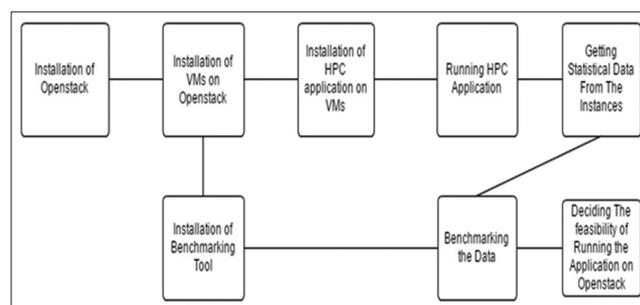


Fig. 2: Implementation plan

and performance when an HPC application is running on the cloud, and the performance of an HPC application on the dedicated systems, on which it conventionally runs. We shall test massively parallel programs and compare their performance with the sequential programs and also the costs incurred in terms of memory, network, and all other factors affecting the application.

We shall install multiple HPC applications and test their performance on simple nodes as well as the cloud environment. We shall also benchmark the performance of these applications on the cloud.

FUTURE WORK

The proposed work for this research is the testing of the application on OpenStack, benchmarking the performance of the application to calculate the practicality of running the HPC applications on cloud. Running the HPC application on private cloud and checking if it affects the performance of the application. In the beginning, we shall try to run the application on a single node, and then check if running the application on multiple nodes affects the performance of the application. After benchmarking the performance of the HPC application on the private cloud, we shall compare its performance with that of a public cloud.

In this research, we shall run parallel programs with open Message Passing Interface and compare their run time with serially running applications. These results will be run in benchmarking tools to check for the feasibility of running these applications on the cloud. If the performance and the results we get from these applications are similar or performing better than the conventional systems, we can use cloud for running HPC applications.

Other types of computing similar to HPC are MTC and HTC. MTC evaluates technologies that are simple and advanced in form, to compose highly scalable applications by making use of the sequential codes that are available. MTC comprises applications that are loosely coupled, which are made of many tasks to achieve a larger goal for the application. MTC represents high-performance computations that comprise of many different activities that are coupled using file system operations.

Whereas, HTC is defined as the concept of using distributed computing to enable multiple jobs to be scheduled to resources that are available to complete the job as fast as possible [16,17]. With the help of HTC, we can integrate multiple computing systems to enable a huge number of tasks to be scheduled and completed in a short amount of time. HTC contrasts the workloads for which the key metric is not a floating point operation which is calculated using per second but measured using per month or year metric.

We can use HPC, MTC, and HTC applications on the cloud environment to test the performance of the applications that require huge amount of resources and also consume petabytes of data. These applications require large amount of resources, so we shall test the performance of these applications on cloud environment to check the feasibility of running these applications on the cloud environment.

CONCLUSION

Cloud computing is used in a wide range of fields such as social media, education, e-commerce, and many more areas. Many organizations are

moving toward cloud computing and are moving their business to the cloud. Hence, with the scope of research getting wider, large amounts of servers and computing power are required by scientists to run large HPC applications. Hence, instead of investing in newer infrastructure for parallel computing, creating a private cloud environment and running those applications on the cloud will be advantageous. In all of the above papers, we have read about the performance of AWS, in which network speed was an issue, but in private cloud set up the network speed can be worked on and made faster depending on the requirement of the application because the infrastructure is present with the scientific organization itself and the use of external networking devices is not required. We shall also check the feasibility of running HTC and MTC applications on the cloud so that these applications can utilize cloud resources and test their performance on the cloud.

The results we will get from running massively parallel applications on the cloud and the same applications on the conventional systems will prove the feasibility of running these applications on computational clouds. When using OpenStack, we can also decide the hardware requirements of the application. We can also understand the performance of the cloud environment and whether it is feasible to run HPC, HTC, and MTC applications on the cloud.

REFERENCES

1. Sadooghi I, Martin JH, Li T, Brandstatter K, Maheshwari K, de Lacerda Ruivo TP. Understanding the Performance and Potential of Cloud Computing for Scientific Applications. *IEEE Transactions on Cloud Computing*; 2015.
2. Juve G, Deelman E, Vahi K, Mehta G, Berman BP, Berrima B, Maechling P. Scientific Workflow Applications on Amazon EC2.
3. Wang G, Ng TS. The Impact of Virtualization on Network Performance of Amazon EC2 Data Center. *IEEE*; 2010.
4. Walker E. Benchmarking Amazon EC2 for High-Performance Scientific Computing. In: *USENIX; Login: Magazine*, October; 2008.
5. Juve G, Deelman E, Berman BP, Berriman B, Maechling P. An Evaluation of the Cost and Performance of Scientific Workflows on Amazon EC2. *IEEE*; 2009.
6. Vecchiola C, Pandey S, Buyya R. High-performance Cloud Computing: A View of Scientific Applications.
7. Jackson KR, Muriki K, Canon S, Cholia S, Shalf J, Wasserman HJ, Wright NJ. Performance Analysis of High Performance Computing Applications on the Amazon Web Services Cloud. *IEEE*; 2010.
8. Castro H, Villamizar M, Garc'es O, P'erez J, Caliz R, Arteaga PF. Facilitating the Execution of HPC Workloads in Colombia through the integration of a Private IaaS and a Scientific PaaS/SAAS Marketplace. *IEEE*; 2016.
9. Dhuldhule P, Lakshmi J, Nandy SK. High Performance Computing Cloud - a Platform-as-a-Service Perspective. *IEEE*; 2015.
10. Zhao Y, Hategan M, Clifford B, Foster I, von Laszewski G, Nefedova V, *et al.* Swift: Fast, Reliable, Loosely Coupled Parallel Computation. *IEEE*; 2007.
11. Available from: <http://www.docs.openstack.org/ops-guide/arch-storage.html>.
12. Available from: <https://www.github.com/openhpc/ohpc>.
13. Available from: <https://www.nas.nasa.gov/publications/npb.html>.
14. Available from: <http://www.davidhbailey.com/dhbpapers/npb-encycp.pdf>.
15. Available from: <https://www.rally.readthedocs.io/en/latest/>.
16. Available from: <http://www.rc.arizona.edu/hpc-htc/high-performance-computing-high-throughput-computing>.
17. Available from: <http://www.citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.160.9183&rep=rep1&type=pdf>.