ISSN (Online) : 2278-1021
ISSN (Print) : 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering
Vol. 3, Issue 4, April 2014

Deployment of Virtual HPC Clusters on Demand from Volunteer Computing Resources

L. Rahul[1], K. Sharavana[2]

MTECH, CSE, MVJCE, Bangalore, Karnataka, India[1]

Associate Professor, CSE, MVJCE, Bangalore, Karnataka, India[2]

ABSTRACT: Volunteer Computing allows individuals to provide their computing resources voluntarily to a community working towards a common goal. Therefore, universities and research institutions can use Volunteer Computing to pool their underlying computing infrastructure into a High Performance Computing (HPC) Cloud, which can later be used to deploy virtual HPC clusters on demand. These virtual HPC clusters allow researchers and scientists to run computationally large parallel applications in the HPC Cloud, which can elastically grow and shrink as per the requirements, at minimal costs. In this paper, we discuss a mechanism that can generate virtual HPC clusters on demand in the Cloud from the pooled Volunteer Computing resources.

Keywords: Cloud Computing; High Performance Computing; Cluster as a Service; HPC as a Service; Volunteer Computing

I. INTRODUCTION

Cloud Computing allows underlying computing resources to be virtualized and then provide that virtualized resource to clients. Cloud Computing can also extend and enhance the virtualized resources at an agile and elastic pace as required on demand. However, it is often costly to invest on the infrastructure demanded by the Cloud Computing paradigm to support virtual High Performance Computing (HPC) clusters in the Cloud. Therefore, in order to reduce the infrastructure costs, the required infrastructure can be provided voluntarily by universities and research institutions that have excess computing resources that can otherwise be used by the demanding large parallel HPC applications. Thus, the computing resources obtained from the universities and research institutions can be pooled into Volunteer Computing resources. The Volunteer Computing resources can be used to generate a Cloud infrastructure that is optimized for HPC applications. Scientists can request an HPC environment to test a hypothesis on demand, whenever and wherever, as required at nominal cost and swift pace. Simulations can be run on the allocated environment, whose resources can later be expanded or shrunk as needed by the applications or models being tested. Furthermore, universities and research institutions can reduce their investment and operation costs, as well as maintenance costs for an HPC environment that can be generated on demand from Volunteer Computing infrastructure with flexible pricing schemes and Service Level Agreements (SLA) [6]. Hence, the virtual HPC instances generated in the Cloud using Volunteer Computing (HPC – VC Cloud [5]) aim to provide an HPC environment to the customers with scalable and elastic computing power as needed, along with necessary adaptable SLAs to service customers towards their consumption of resources. The paper is structured as follows: section II presents the solution provided by HPC – VC Cloud, section III describes a method involved in generation of virtual HPC Clusters in the HPC – VC Cloud, section IV describes the related work similar to HPC – VC Cloud, and finally section V summarizes the paper into a conclusion.

II. HPC – VC CLOUD SOLUTION

Since the advent of supercomputers, they had been used for complex calculations ranging from weather forecasting to molecular dynamics simulation. Earlier supercomputers relied on specialized architecture requiring specialized hardware and operating system to perform scientific simulations, such as nuclear test simulations and brute force code breaking [1]. However, as cost became a limiting factor to increase computational power, opportunistic approaches – such as grid computing – paved the way to increase required computational power without incurring extra costs. HPC – VC Cloud combines both the traditional supercomputing and grid computing, along with additional benefits from Cloud Computing, to provide HPC as a Service to its customers. Another major player in HPC – VC Cloud is the notion of volunteer computing. Volunteer computing allows universities to register their entire computer lab infrastructure as a service to the HPC – VC Cloud; thereby, allowing the HPC – VC Cloud resources to be allocated and expanded dynamically. The HPC – VC Cloud combines several computers across the universities into a HPC Cluster and offers it as a service to its customers [3, 9]. Scientists can rent resources based on a pricing model and SLAs and can later test a weather forecasting model or perhaps even a particle physics simulation on – demand as needed. Businesses can rent server and data farms, which can be scaled up and down as required, to achieve high availability, resiliency, and fault tolerance to keep their services running at minimal operational and maintenance costs.

Figure 1 shows the architecture of HPC – VC Cloud implementation used to provide HPC as a Service to its customers. A set of physical machine resources exist in the HPC – VC Cloud. With the help of virtualization [2] and some clustering tools, the physical machine resources are combined into a HPC cluster based on customer's demand, which can later be scaled up or down as needed. The final HPC Cluster is then offered as a service to the

ISSN (Online) : 2278-1021
ISSN (Print) : 2319-5940
International Journal of Advanced Research in Computer and Communication Engineering
Vol. 3, Issue 4, April 2014

customer using VPN Tunneling to allow the customer to connect to the corresponding HPC Cluster created. Once the customer finishes the work on the allocated HPC Cluster, the cluster is freed back into physical machine resources to be used by other customers.

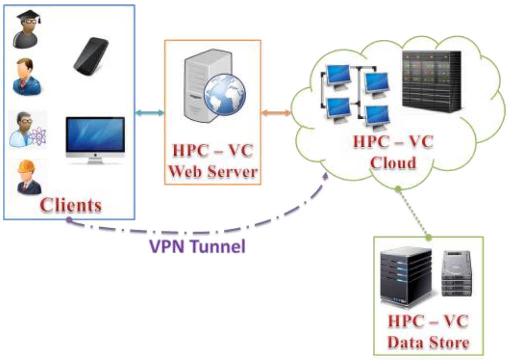

Figure 1 – HPC – VC Solution

The HPC – VC Cloud also includes scheduling and monitoring mechanisms to meet the agreed upon SLAs and generate reports for pricing as needed. Thus, volunteer computing allows university computer lab resources to be dynamically added to the HPC – VC Cloud as needed; thereby, increasing the resources in the HPC – VC Cloud to generate more HPCs in the Cloud.

III. GENERATING VIRTUAL HPC CLUSTERS

In order to generate virtual HPC clusters from the Volunteer Computing resource pool, we need to employ several principles, as shown in Figure 2. The physical resources obtained from the Volunteer Computing resources need to be clustered into an HPC cluster. Then, we can generate virtual HPC clusters from the physical resource pool using a virtualization hypervisor. The generated virtual HPC cluster is then provisioned and made available to the requesting client via a Virtual Private Network (VPN) tunnel. In addition, an HPC – VC Store Access enables the clients to connect to their online HPC – VC Data Store to upload their data to be used for their HPC application and then later download the results of the HPC application's execution. Finally, the Price Monitor watches the resources being consumed by the client and applies pricing charges according to a pre – determined pricing plan. Several tools and technologies exist and can be enhanced to achieve the process of generating virtual HPC clusters from the physical resources, which are explained in subsequent sub sections.

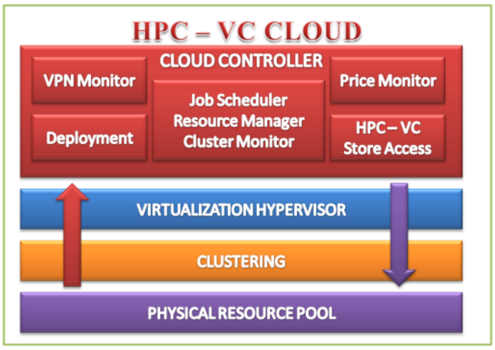

Figure 2 – A Mechanism for Generating Virtual HPC Clusters

A. *Clustering Tools*

A large number of tools, including both open source and commercial, exist for clustering several physical computers into an HPC cluster. Some are turn – key solutions, which can generate an HPC cluster with minimal configuration; while others are light – weight clustering solutions, which perform minimal detection and generation of HPC clusters that are extremely flexible and dynamic. Regardless of the clustering tools used, the clustering process must be flexible enough to allow the clustering of heterogeneous computers; be able to dynamically adjust computing resources attaching and leaving the Volunteer Computing resource pool; be able to accurately perform assigned jobs with fault tolerance and process migration; and finally, be able to accurately monitor the resources and assign jobs accordingly to satisfy the HPC application needs. Some commonly used clustering tools include NPACI Rocks Cluster Distribution, Warewulf, and Open Source Cluster Application Resources (OSCAR) distribution.

B. *Virtualization Hypervisor*

A virtualization hypervisor is responsible for virtualization of physical computing resources into virtual computing resources, which can later be expanded and shrunk as needed. A virtualization hypervisor can either run directly on hardware as an operating system or as an application on top of another operating system. The former are called bare – metal hypervisors, since they run directly on hardware; while, the later are called para-virtualization hypervisors, since they run on top of another operating system. Nevertheless, the virtualization hypervisor must be able to detect the underlying generated physical HPC cluster created from the Volunteer Computing resource pool and then generate virtual HPC clusters from the physical HPC cluster. The virtual HPC clusters should be able to expand and shrink as needed by adding or removing more physical computing resources from the Volunteer Computing resource pool. These generated virtual HPC clusters are then deployed and made accessible to requesting clients to run their HPC applications. Some of the most commonly used virtualization hypervisors include VMware vSphere Hypervisor, Kernel – Based Virtual Machine (KVM), Citrix XenServer, and Microsoft Hyper – V.

C. *Job Schedulers, Resource Managers, and Cluster Monitors*

In order to run HPC applications, they must be scheduled accurately to run in parallel with minimum delay and error. Thus, a parallel job scheduler is necessary to schedule the client jobs on the HPC clusters. Furthermore, in order for job schedulers to allocate jobs accurately, the HPC clusters must be monitored carefully with resource managers and cluster monitors. Moreover, it must also be possible to migrate jobs from one HPC to another in the event of a lack of resources and failure of the HPC. Thus, appropriate tools are required to ensure guarantee of the HPC application's execution on the allocated virtual HPC clusters. Some commonly used job schedulers for HPC clusters include TORQUE Resource Manager, Simple

ISSN (Online) : 2278-1021
ISSN (Print) : 2319-5940
International Journal of Advanced Research in Computer and Communication Engineering
Vol. 3, Issue 4, April 2014

Linux Utility Resource Manager (SLURM), Maui and Moab Cluster Suite, High Throughput Condor (HTCondor), and Univa Grid Engine. The most widely used cluster monitoring tool is the Ganglia Monitoring System.

D. Integration with other vendors

Apart from universities and research institutions, several vendors, including Amazon, Google, Microsoft, VMware, Rackspace, GoGrid, etc. provide commercial Infrastructure – as – a – Service (IaaS), which can be used to extend resources in the Volunteer Computing resource pool to guarantee some SLAs. However, most of these vendors do not provide resources voluntarily and have vendor lock-in SLAs, which make it difficult for HPC – VC Cloud to integrate with them. However, these vendors have decided to willingly provide resources without vendor lock-in by voluntarily sponsoring an IaaS solution led by Rackspace and NASA known as OpenStack. Thus, in order to guarantee resources for some SLAs in the Volunteer Computing resource pool, HPC – VC Cloud can integrate with OpenStack without the fear of vendor lock-in and with minimal effort.

IV. RELATED WORK

HPC – VC Cloud follows the "HPC as a Cloud" approach to offer HPC-as-a-Service [7, 8] to the customers from the Volunteer Computing resource pool. HPC – VC Cloud provides an end to end solution offering all the three dimensions of Cloud Computing involving: Infrastructure-as-a-Service (IaaS) to provide customized HPC clusters; Platform-as-a-Service (PaaS) to allow customers to build and run their applications for their HPC clusters; and Software-as-a-Service (SaaS) to provide thin client through which customers access their HPC clusters. Hence, HPC – VC offers services similar to ideas based on HPC-as-a-Service (HPCaaS) [4] and Berkeley Open Infrastructure for Network Computing (BOINC).

V. CONCLUSION

The HPC – VC Cloud solution provides an end – to – end product to generate virtual HPC Clusters and present them as a service to the demanding customers at minimal costs, along with agreed upon SLAs and flexible pricing models. The HPC – VC Cloud generates virtual HPC Clusters from the physical computing resources voluntarily provided by the universities and research institutions via Volunteer Computing. We have also discussed how to generate virtual HPC clusters from the physical resource pool using several tools and technologies. Furthermore, HPC – VC Cloud can also integrate with other existing Cloud vendors by using a solution provided by OpenStack without the fear of vendor lock-in. The virtual HPC resources can also be scaled up and down as needed based on the client requirements.

REFERNCES

[1] A. Ramirez, "Supercomputing: Past, present, and a possible future", 2011 International Conference on Embedded Computer Systems (SAMOS), July 2011, doi: 10.1109/SAMOS.2011.6045437.

[2] A. Reuther, P. Michaleas, A. Prout, J. Kepner, "HPC-VMs: Virtual machines in high performance computing systems," High Performance Extreme Computing (HPEC), 2012 IEEE Conference on, vol., no., pp.1,6, 10-12 Sept. 2012, doi: 10.1109/HPEC.2012.6408668.

[3] F. Doelitzscher, M. Held, C. Reich and A. Sulistio, "ViteraaS: Virtual Cluster as a Service", 2011 IEEE Third International Conference on Cloud Computing Technology and Science (CloudCom), pp. 652-657, December 2011, doi: 10.1109/CloudCom.2011.101.

[4] G. Shainer, Tong Liu, J. Layton and J. Mora, "Scheduling strategies for HPC as a service (HPCaaS)", 2009 IEEE International Conference on Cluster Computing and Workshops CLUSTER '09, pp.1-6, August 2009, doi: 10.1109/CLUSTR.2009.5289158.

[5] L. Rahul and Sharavana K., "A Survey on Deploying HPC in the Cloud Using Volunteer Computing", The International Journal of Computer Engineering and Applications (IJCEA), ISSN: 2321 – 3469 vol.5, no.1, pp.34-38, January 2014.

[6] M. AbdelBaky, M. Parashar, Hyunjoo Kim, K.E. Jordan, V. Sachdeva, J. Sexton, H. Jamjoom, Zon-Yin Shae, G. Pencheva, R. Tavakoli and M. F. Wheeler, "Enabling High-Performance Computing as a Service", Computer, vol.45, no.10, pp.72-80, October 2012, doi: 10.1109/MC.2012.293.

[7] R. Kubert and S. Wesner, "High Performance Computing as a Service with Service Level Agreements", 2012 IEEE Ninth International Conference on Services Computing (SCC), pp.578-585, June 2012, doi: 10.1109/SCC.2012.43.

[8] S. Yokoyama and N. Yoshioka, "Cluster as a Service for Self-Deployable Cloud Applications", 2012 12th IEEE/ACM International Symposium on Cluster, Cloud and Grid Computing (CCGrid), pp.703-704, 13-16 May 2012, doi: 10.1109/CCGrid.2012.64.

[9] Xiaorong Li, H. Palit, Yong Siang Foo, T. Hung, "Building an HPC-as-a-Service Toolkit for User-Interactive HPC Services in the Cloud," Advanced Information Networking and Applications (WAINA), 2011 IEEE Workshops of International Conference on , vol., no., pp.369,374, 22-25 March 2011, doi: 10.1109/WAINA.2011.116.