A SCALABLE GRID INFRASTRUCTURE IN THE EUCALYPTUS CLOUD ENVIRONMENT

Mohan Krishna Varma N, Eunmi Choi*

Department of Information System,
School of Business IT, Kookmin University,
Jeongneung-Dong, Seongbuk-Gu, Seoul, 136-702, Korea
mohankvarma@gmail.com, emchoi@kookmin.ac.kr

Abstract. A powerful and scalable infrastructure, which determines computational and data-intensive tasks and delivers on-demand scalable resources, is very much essential in the future. Integration of current technologies on grid and cloud infrastructures is able to fulfill these kinds of demand. We are investigating the potential and significant association developments of grid and cloud systems. Extending an existing grid system with resources provided by a cloud infrastructure delivers the resource infrastructure environment for cloud computing demands. To achieve this goal, we work on the integration of *Globus Toolkit 4 (GT4)* grid system and *Eucalyptus (Elastic Utility Computing Architecture Linking Your Programs to Useful Systems)* cloud infrastructure. Extending Globus Toolkit 4 grid system with a Virtual Runtime Environment provided by *XEN Virtual Machine Monitor (VMM)* using Eucalyptus framework is the major focus of our research. A hybrid platform that associates grid computing and cloud computing technologies would provide benefits to service providers and users. By using cloud computing technologies, grid computing systems can moderate boundaries for scalable infrastructure.

Keywords: Virtual Runtime Environment, Virtual Machine Monitor, Virtualization, Grid Computing, Cloud Computing, Globus Toolkit, Eucalyptus, XEN

1. Introduction

Grid computing has its traditional way of submitting, managing and executing jobs. According to Ian Foster [6], remote execution environment remains an issue because it offers diverse environments such as operating systems, different middleware architecture and variations in file system results in little usage of available resources by the users. Multiple users share and run their jobs or processes in the end resource, which shares the same operating system of the resource, sometimes this leads to performance issues. Hence there is a need for customizable, dedicated, controllable environment for the user, while Virtualization decouples the running environment from the hardware. Providing such environments can reduce the burden of the middleware by shifting some components such as monitoring to the resource end rather than functioning at the middleware. Cloud computing [3] has the new promising idea of providing the virtual environment for the users. Users can have a dedicated, on-demand, controllable environment with the help of Virtual Machines (VMs)

*Corresponding Author. This research was supported by Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education, Science and Technology (2012-0002774).

[2]. User can have a complete control over the VMs by deploying it or running it or suspending it when needed.

The Globus Toolkit 4 (GT4) is used to implement the Grid computing environment with the help of Linux machines. Globus Toolkit (GT) [9] is an open source toolkit organized as a collection of loosely coupled components. These components consist of services, programming libraries and development tools designed for building Grid-based applications GT components fall into five broad domain areas: Security, Data Management, Execution Management, Information Services, and Common Runtime.

Cloud vendors provide different customized interfaces to the users. The lack of research tools motivates the companies to develop a common and flexible cloud computing framework to formulate experiments and to address open questions in cloud computing. Eucalyptus [1] is a research oriented open source cloud computing system, which utilizes resources that are available in the user environment. It implements Infrastructure as a Service (IaaS) model to provide a solid foundation of cloud computing systems that can be developed and extended. Eucalyptus enables users to create and control Virtual Machine (VM) instances deployed across different physical resources within cloud environment.

XEN is a widely-used high performance open source hypervisor [4], which is installed above the hardware layer and has a direct access to the most of the hardware devices, while running in the most privileged processor level. XEN domains run at the top of hypervisor.

In this paper, we build the three layered architecture to deliver the on-demand scalable resources to process the computational and data-intensive tasks. To achieve scalable infrastructure architecture, a set of technologies of Eucalyptus, XEN and GT4 are integrated together. The paper is organized as follows: Section 2 deals with the System Architecture, Section 3 explains about the Results, and Section 4 concludes the paper.

2. System Architecture of Infrastructure

The design architecture of the infrastructure system shows the three-layers to integrate the cloud computing and the grid computing as in Figurel. Eucalyptus is used to implement the concept of cloud computing, and in the same way GT4 is used to implement the Grid Computing environment. This architecture is able to process the computational and data-intensive tasks and delivers on-demand scalable resources.

2.1 Infrastructure Integration

Eucalyptus and GT4 are integrated to developing the new middleware layer in the architecture. Resource configuration and resource provisioning components of the Eucalyptus communicates with the resource monitor component of the GT4 to monitor the resources for a particular time interval in order to check the behavior the re-

sources. Resource monitor communicates with the resource manager to scale the resources according to the user needs.

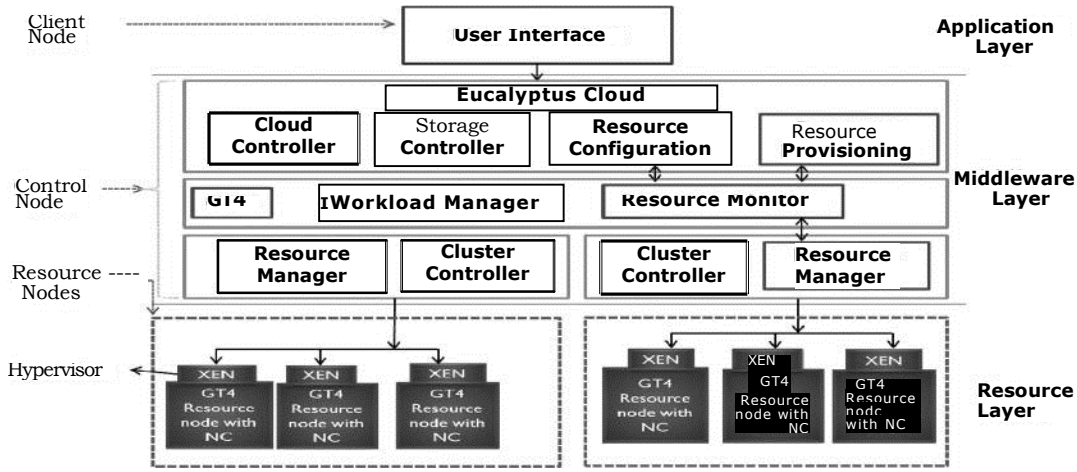

Fig. 1. System Architecture

2.2 **Scalability for VM Resources**

Proper deployment of VMs can handle the resource scalability. To achieve this, each resource node is configured with GT4 and Eucalyptus Node Controller (NC) on the top of XEN. GT4 components help to achieve the capabilities of grid computing and NC with XEN help to achieve the scalable infrastructure. NC controls the execution, inspection and termination of the VM instances on the host where it runs.

2.3 Subsystems to deliver VM tasks

To deliver the VM tasks set of subcomponents work together with middleware. Cloud controller makes high level scheduling decision and implements them by making requests to cluster controllers. Storage controller is a storage service that provides a mechanism for storing and accessing virtual machine images and user data. Workload manager assigns the tasks to appropriate computing elements. Cluster controller gathers the information about the VM execution on specific node controller.

3. Results

Results are experimented using the XEN virtualization environment. To install the Eucalyptus node controller, XEN is one of the prerequisite. XEN environment is tested by installing number of virtual machines with 96MB memory allocation for each guest domain. This experiment is done with 2GB, 3GB, 4GB, 5GB and 6GB Linux systems. The maximum number of VMs running on each system is 15, 24, 34, 43 and 54 respectively as in Figure 2. This experiment supports to estimate the capacity of the XEN virtual machine environment by running a number of VMs within a system.

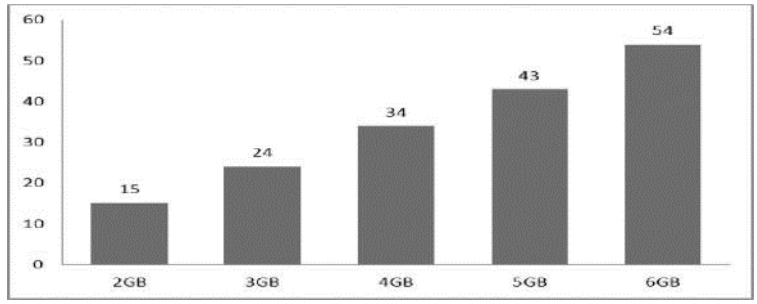

Fig. 2. CPU utilization with maximum VMs

4. Conclusion

Virtualization is a promising technology in Cloud Computing for utilizing the resources effectively. This feature can further be extended in mapping jobs to virtual machines and execute the jobs in a controllable, customizable, secured and virtualized environment. The paper proposes middleware architecture by taking advantages of Cloud Computing and Grid Computing to provide virtualization environment of VM tasks.

Reference

1. Daniel Nurmi, Richard Wolski, Chris Grzegorczyk, Graziano Obertelli, Sunil Soman, Lamia Youseff, Dmitrii Zagorodnov, "The Eucalyptus Open-Source Cloud-Computing System", CCGRID 2009, pages 124-131.
2. M. McNett, D. Gupta, A. Vandat, and G. M. Voelker, "Usher: An Extensible Framework for Managing Clusters of Virtual Machines", In Proceedings of the 21st Large Installation System Administration Conference (LISA), November 2007.
3. Marcos Dias de Assuncao, Alexandre di Costanzo, Rajkumar Buyya, "Evaluating the Cost-Benefit of Using Cloud Computing to Extend the Capacity of Clusters", 2009.
4. P. Barham, B. Dragovic, K. Fraser, S. Hand, T. Harris, A. Ho, R. Neugebauer, I. Pratt, and A. Warfield, "Xen and the art of virtualization", In SOSP'03: Proceedings of the nineteenth ACM symposium on Operating systems principles, pages 164-177, New York, NY, USA, 2003. ACM.
5. Lizhe Wang, Gregor von Laszewski, Marcel Kunze, Jie Tao, Jai Dayal, "Provide Virtual Distributed environments for Grid Computing on demand", Adavances in Software Engineering, Elsevier, 2009.
6. I. Foster and C. Kesselman, "The Grid-Blueprint for a New Computing Infrastructure", Morgan Kaufmann, 1998.
7. Rajkumar Buyya, Suraj Pandey, and Christian Vecchiola, "Cloudbus Toolkit for Market-Oriented Cloud Computing", The University of Melbourne, 2009.
8. Fallenbeck N, Picht HJ, Smith M, freisleben B, "Xen and the arts of Cluster Scheduling", Proceedings of 1st international workshop on virtualization technology and distributed computing, 2006.
9. I. Foster and C. Kesselman, "Globus: A metacomputing infrastructure toolkit", International Journal of Supercomputer Applications, 1997.