HPCELASTIC:AUTOMATIC RESOURCE ELASTICITY MODEL FOR HPC IN THE CLOUD

Kalyani Sudame[1], Shanti Upase[2], Siddhi Yatnalkar[3],Prof.Navnath Kale[4]

[1] *Student, Computer Engineering,* JSPM's PVPIT, *Maharashtra, India*
[2]*Student, Computer Engineering,* JSPM's PVPIT, *Maharashtra, India*
[3]*Student, Computer Engineering,* JSPM's PVPIT, *Maharashtra, India*
[4]*Professor, Computer Engineering* JSPM's PVPIT, *Maharashtra, India*

ABSTRACT

Elasticity is undoubtedly one of the most striking characteristics of cloud computing. Especially in the area of high performance computing (HPC), elasticity can be used to execute irregular and CPU-intensive applications. However, the on- the-fly increase/decrease in resources is more widespread in Web systems, which have their own IaaS-level load balancer. Considering the HPC area, current approaches usually focus on batch jobs or assumptions such as previous knowledge of application phases, source code rewriting or the stop-reconfigure-and-go approach for elasticity. In this context, this article presents HPCElastic:Automatic resource elasticity model for HPC in the cloud. Its differential approach consists of providing elasticity for high performance applications without user intervention or source code modification. The scientific contributions of AutoElastic are twofold: (i) an Aging-based approach to resource allocation and deallocation actions to avoid unnecessary virtual machine (VM) reconfigurations (thrashing) and (ii) asynchronism in creating and terminating VMs in such a way that the application does not need to wait for completing these procedures.

Keywords:- *Cloud elasticity, high-performance computing, asynchronism, resource management, self-organizing Load Balancing*

1. INTRODUCTION

AutoElastic, is an elasticity model that focuses on high performance applications. Elasticity is an essential principle for the cloud model because it not only provides efficient resource sharing among users but also makes it feasible to have a pay-as- you-go computing style. Current state-of-the-art shows that the most common approach for elasticity is the replication of stand-alone virtual machines (VMs), when a particular threshold of a given metric or combination of metrics is reached The load balancer maintained by the cloud provider manages the service calls (which are usually identified by URL or IP address) and redirects demands for the most suitable replica.

2. LITERATURE REVIEW

The manual approach is especially common in private clouds. Users can develop their own application for monitoring services that run on VMs, launching elasticity actions when necessary. However, some middleware, such as Amazon AWS, Windows Azure and Nimbus, configurable subsystems for service provide con monitoring and elasticity management. Users must complete the rules and the limits of a metric to be monitored as well as the

conditions and actions for reconfiguration. AWS provides and actions for recon an API and a graphical tool for these tasks. The monitoring is performed by CloudWatch, and the elasticity itself is managed reactively by Autoscaling

3. WORKING

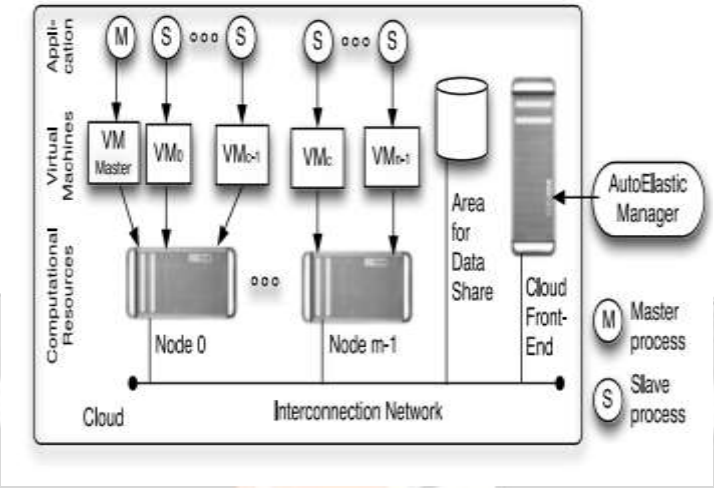

Fig -1: Architecture Diagram

CLOUD COMPUTING is a kind of internet based computing that provides shared processing resources and data to computers and other devices on demand. It is a model for enabling ubiquitous , on demand access to shared pool of configurable computing resources which can be rapidly provisioned and released with minimal management effort In this context, this work presents AUTOELASTIC, a PaaS-level elasticity model for HPC in the cloud.

3. SUMMARY

CLOUD COMPUTING is a kind of internet based computing that provides shared processing resources and data to computers and other devices on demand. It is a model for enabling ubiquitous , on demand access to shared pool of configurable computing resources which can be rapidly provisioned and released with minimal management effort. ELASTICITY is undoubtedly one of the most striking characteristics of cloud computing. Especially in the area of high performance computing (HPC), elasticity can be used to execute irregular and CPU-intensive application..

4. CONCLUSION

- In brief, Auto Elastic self-organizes the number of virtual machines without any user intervention, providing benefits for both the cloud administrator (better energy consumption and resource sharing among the users) and the cloud user (application performance and overhead when migrating to the cloud). It is a model for enabling ubiquitous , on demand access to shared pool of configurable computing resources which can be rapidly provisioned and released with minimal management effort.The Auto Elastics model for applications matches the socket-programming style that is offered by MPI-2, in which new processes can be easily added or removed, connected and disconnected, from the parallel computation.

6. REFERENCES

[1 Rodrigo da Rosa Righi, Vinicius Facco Rodrigues, Cristiano Andre, Automatic Resource Elasticity for High Performance Applications in Cloud, 2016

[2]. B. HimaGiriNandini, 2K. Ravi Kumar.SII: Improving Query Responsive Time to find Spatial Objects with keywords.International Journal of Research in Computer and Communication Technology, Vol 4, October-2015

[3]. C. Usha Rani, N.Munisankar. Spatial Index Keyword Search in multidimensional database.(IJCSIT) International Journal of Computer Science and Information Technologies, Vol. 5 (5), 2014

[4]. X.Cao,L.Chen,G.Cong, and C.S.Jenesen, and Q.Qu, A.Skovsgaard, D.Wu and M.L.Yiu.Spatial Keyword Querying.InER,pages16-29, 2012.

[5]. X.Cao, G.Cong, and C.S.Jensen.Retrieving Top-k prestige based relevant spatial web objects.PVLDB,2010.

[6]. E.Chu, A.Baid, X.Chai,A.Doan and J.Naughton.Combining keyword search and forms for Ad Hoc querying of databases.In proc. of ACM management of data(SIGMOD),2009.

[7]. M. Mao, J. Li, and M. Humphrey, Cloud auto-scaling with dead and budget constraints, in Proc. 11th IEEE/ACM Int. Conf.Grid Comput., 2010, pp. 4148.