Advances in Computer Science and Information Technology (ACSIT)
Print ISSN: 2393-9907; Online ISSN: 2393-9915; Volume 2, Number 9; April-June, 2015 pp. 77-78

http://www.krishisanskriti.org/acsit.html

Cost Optimization in Cloud Computing Environment

Niharika[1] and Jitender Kumar[2]

[1](M.tech Scholor) C.S.E. Department D.C.R.U.S.T Murthal
[2]C.S.E. Department D.C.U.R.S.T Murthal

Abstract—*Cloud computing is trend of today and demand of coming future but as usual every technology comes with its pros and cons, with the ease of its use cloud is also bounded with different challenges like Security, Scaling, bandwidth management, services delivery and proper utilization of resources. Virtualization is one of the attribute that helps to reduce the wastage of hardware and cost of operation performed .As the number of user grows proportionally size of cloud also grows and there are more chances of SLA violations which is an emerging issues for the service providers as SLA violations are directly related to penalty .At the user side choosing the right plan for VM purchasing is the thing need to work on. In the current work we have focused on purchasing optimum usage plan in respect of Software as a Service (SaaS) providers.*

1. INTRODUCTION

Cloud computing is layered model where different services like hardware, software and delivery of application and their management are provided as per user's query. In today's world cloud is an important part of our life we can see the influence of cloud in business applications, social networking sites and web map). Cloud computing gives us the liberty of using online applications without any basic structural knowledge of hardware used, software is updated or not. Each request on cloud is responded automated according to query.

According to [1] "**NIST** definition of Cloud Computing"

"Cloud computing is a model for enabling ubiquitous, convenient, on-demand network access to a shared pool of configurable computing resources (eg. networks, servers, storage, applications, and services) that can be rapidly provisioned and released with minimal management effort or service provider interaction."

The NIST definition [1] is one of the most referred definition which gives a brief about **(1)** Deployment models of cloud according to user's privacy (Private Cloud), organization's privacy (community cloud), public welfare (public cloud) and combination of first three approaches (hybrid cloud).**(2)** Characteristics of cloud (automated services, equal distribution of resources, large bandwidth, proper monitoring of resources, elasticity in work environment) .**(3)** Services provided at software level (Google docs), hardware level (Google App Engine), infrastructure level (Amazon Web Services).Infrastructure of cloud is [2] Ubiquitous, it's like resources are present everywhere in cloud but provisioning of resources should be cyclic with minimum SLO violations so this scenario is termed as NP hard problem. In cloud it is very hard to predictable dynamic load, data availability, available operating and computational systems. To enhance the system performance resources are pooled and provided on demand. Our contribution in this paper is to present a plan that minimizes the cost application management and delivery.

2. VIRTUAL RESOURCE MANAGEMENT AND DEPLOYMENT

There are three basic steps of virtual resources on deployment [3] on cloud (1) start a new virtual machine according to specified hardware and software requirement (2) mapping of virtual resources to the physical machine (3) subsequently deploy the asked application only. Different techniques are used for Cloud Provisioning, [4] in dynamic environment automated approach is used where virtual resources are ranked according to their size, performance and system goals, with all the data collected a learning model is prepared which helps in resource management. In [5] author proper cloud Provisioning will affect the budget and performance of the system. If resources are not scheduled properly SLO violations will be high which lead to system failure. Application service provider can control only application provisioning therefore in this paper we concentrate only on better utilization of virtual resources and virtual application.

3. RELATED WORK

This paper is mostly related to cost minimization at the service provider side by improvising existing techniques. Other related work is also discussed but our main goal would be cost minimization. Li [6] gives a theory that adaptive management of virtual resources in cloud computing can be more effective if feedback control theory is used. His purposed work dynamically adjusts the requirement of virtualized recourses. But he doesn't specify the type of virtual instances used. Sadhasivam [7] worked on proper utilization of resources using meta scheduler and backfill strategy .h purposed a centric virtual machine scheduler for better resource utilization. Fabricio A.B da silva [8] state that Bag of Task application at client-server platform can be improved by full use of file affinity concept. But for scaling purpose author

didn't show clarity in resource type whether reserved or on demand.

4. PROPOSED METHODOLOGY

What we think about pricing and using application hasn't change with time and technology.[9] Price optimization for web applications are very challenging as multiple features are associate with them. If we talk about price at service provider side he will first think about his profit which is not possible if we dependant only on static instances or on demand instances. We need a hybrid approach for maximum benefits. Data extracted from [3] we can see that in static VM allocation policy utilization rate and rejection rate depends on number of VM's .first experiment with 50 VM's utilization rate is 77 and rejection rate is 55 whereas with 150 VM's utilization rate is 60 and rejection rate is 0.5. [10] In dynamic environment we know the lower bound of VM required and is very hard to predict the upper bound. From [3] we can analysis we required minimum 55 and maximum 153 VM's are provided than utilization rate is 80 and rejection rate is 2. From [9] we can conclude that if we use reserved resources over on demand instances we can save 60 percent in a year but rejection rate would be so high that may lead to service failure. So we need to mix both the approaches for optimum plan .If we see the algorithm proposed in [3] it works on on-demand instances with initial value of VM as zero which would increase the price and complexity of application. From the past record we find that minimum number of VM's required is 55. If we reserved minimum VM's required with some resource cap it will lead to a long term benefits.

5. CONCLUSION AND FUTURE WORK

Cost optimization is a complex aspect of cloud computing as it is connected from service provider as well user side. It also includes SLA whose violations leads to penalty directly or indirectly in terms of money. So we further need to explore more data for reserved instances at IaaS platform. Further we can divide users as they are reserved or on demand customers so that service providers can also gain profit.

6. ACKNOWLEDGMENTS

This work was supported in part by a grant from the National Science Foundation.

REFERENCES

[1] P.Mell and T.Grance,"The NSIT Definition of CloudComputing(Draft),"http://csrc.nist.gov/publications/draft/800-145/Draft-SP-800-145_cloud-definition.pdf.

[2] R.N Calheiros,R.Ranjan and R.Buyya," Virtual Machine Provisioning Based on Analytical Performance and QoS in Cloud Computing Environments, " International Conference on Parallel Processing ,pp295-304,2014

[3] A. Quiroz, H. Kim, M. Parashar, N. Gnanasambandam, and N. Sharma, "Towards autonomic workload provisioning for enterprise grids and clouds," in Proceedings of the 10th IEEE/ACM International Conference on Grid Computing (GRID'09), 2009

[4] Piyush Shivam , David Irwin, Laura Grit, Ayden Yumerefendi,"Automated and on demand Provisioning of Virtual Machines for Database Applications," *SIGMOD'07* Beijing, China .June 11–14, ACM 978-1-59593-686-82007

[5] A kansal, N. khothri, A.A, Bhattacharya, "Virtual Machine Power Metering and Provisioning,"Indiana USA,June 10-11,2010

[6] Qiang Li,Qinfen Hao,Limin Xiao and Z.li," Adaptive Management of Virtualized Resources in cloud computing Using Feedback Control,"1st International Conference on Information Science and Engineering,pp99-102

[7] S.Sadhasivam,R.jayarani,N.Nagaveni and R.Vasanth Ram,"Design and Implementation of an efficient Two-level scheduler for cloud Computing Environment,"ICRTCC,2009, pp884-886

[8] F.A.B da Silva,H.Senger,"Improving scalability of Bag of Tasks applications running on master-slave platforms, " Parallel Computing system and applications Journal, 2009, pp.57-71

[9] Amazon Web Services-How AWS Pricing works, July 2014 ,available at http://aws.amazon.com/whitepapers/

[10] Z.Shen, S.sethuraman, Xi gu,"CloudScale:Elastic Resource scaling for Multi-Tenant Cloud Systems," Cascais,Portugal,2011