Optimum Utilization of Private cloud with predefined Reservation Policy

UDAYKUMAR I. LOKHANDE[#1], DEVEN N. SHAH[*2]

[#]*Terna Engineering College, Information Technology, Mumbai University , Nerul, Navi Mumbai, INDIA*

[1]udayom@gmail.com

[*]*Terna Engineering College, Information Technology, Mumbai University , Nerul, Navi Mumbai, INDIA*

[2]devenshah@ternaengg.ac.in

***Abstract*— Cloud computing is considered as one of the most promising and fast growing technology due to it's flexibility, scalability and cost-reducing Facility. There are two types of cloud one is private and other is Public. The public cloud is considered to have unlimited resource but the private cloud may face some problem since the elastic behaviour property fails due to limited resources. In this paper we study the Infrastructure As A Service cloud environment through our laboratory practical running on private cloud based service, check for legitimately utilized machine and if unnecessary resource is being utilized with no operation being performed then by forcefuly terminating their instance to provide Machine instance to other legitimate users who are waiting for resource by using predefined reservation policy.**

***Keywords*— Parallel computing, machine instance, cloud computing, Virtual Machine, Enforcing policy.**

I. INTRODUCTION

The National Institute of Standards and Technology (NIST)[3] defines defines cloud computing as "A model for user convenience, on demand network access contribute the computing resources (Eg. Network, Storage, Application, Server and services) that can be rapidly implemented with minimal management effort service provider interference." The services in cloud computing are given on X as a Services (XaaS) which is splitted into three services viz. "Platform", "Software" and "Infrastructure".

This is the reason behind the accronym of Services and are called as SaaS ie. Software As A Services, PaaS ie. Platform As A Service, IaaS ie. Infrastructure As A Service. Cloud Service is classified as shown in Fig. 1 called as SPI model. SPI referes Software, Platform and Infrastructure (as a servies) respectively defined.

A. Cloud Service Models

- **Cloud Software As A Service (SaaS):** In the SaaS model, cloud providers install and executes application software in the cloud and their users access the software from cloud clients. The user dont have to care about how the cloud infrastructure has been maintained and how the application is running. Examples of SaaS include: Google Apps, Microsoft Office 365, GT Nexus, Marketo and TradeCard.
- **Platform As A Service (PaaS):** In the PaaS model, the platform is provided to the user which includes OS, database, programming language execution and web server. Application developers can develop and run their application on cloud without cost and complexity of buying and managing the required hardware and software layers.

Examples of PaaS include: AWS Elastic Beanstalk, Cloud Foundry, Heroku, Google App Engine, Windows Azure Compute and OrangeScape.

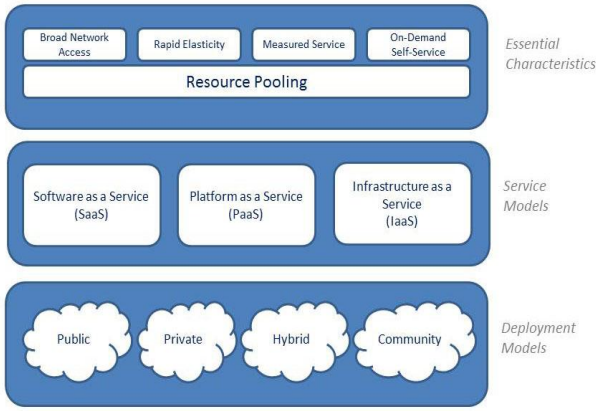

Fig. 1 Cloud Computing Models and Deployment

- **Cloud Infrastructure As A Service (IaaS):** The highly used and widely in use model of cloud is IaaS, which provides physical or virtual machines including other resources to the user by using a concept of Hypervisor. Resources such as images in VM as given in [2] image-library, firewall, IP Address, VLANs and software bundles.

Examples of IaaS providers include Amazon EC2, Windows Azure Virtual Machines, Google Compute Engine, HP Cloud, Oracle Infrastructure as a Service.

B. Cloud Deployment Models

- **Public Cloud:** Public cloud applications, storage, and other resources are made available to the general public by a service provider. These services are free or offered on a pay-per-use model. Generally, public

cloud service providers like Microsoft, Amazon AWS and Google own and operate the infrastructure and offer access only via Internet as given in [4] (direct connectivity is not offered).

- **Community Cloud:** Community cloud shares infrastructure between several organizations from a specific community with common concerns (security, compliance etc.), whether managed internally or externally. The costs are spread over fewer users than a public cloud (more than a private cloud), so only some of the cost savings potential of cloud computing are realized as given in [3].
- **Hybrid Cloud:** Hybrid cloud is a composition of two or more clouds (private, community or public) that remain unique entities but are bound together, offering the benefits of multiple deployments as discussed in [3].
- **Private Cloud:** Private cloud is cloud infrastructure operated only for a single organization, weather they be hosted internally or externally like explained [3]. Private cloud requires a optimum level and degree of engagement to virtualize the infrastructure environment which will also require the organization to re-evaluate decisions about existing resources.

The private cloud is implemented within as Infrastructure but with limited resources. If we consider the IaaS model case in private cloud deployment, the elasticity property does not hold good enough once these cloud services are deployed over the institute of commercial complex.

II. RESERVATION POLICY IMPLEMENTATION

We studied the practical laboratory in our infrastructure where private cloud is implemented. The lab consists of 20 machines and maximum instance that cloud can run is 18. The Fig. 2 shows the average utilization of machine instances when they are started while doing practicals. The part of resource utilization is shown where the machine instance shown in red color launches the instance on cloud server but does nothing while performing practical and due to this the students that are actually in need to perform practical or the legitimate user has to wait till instance of other machine is closed.

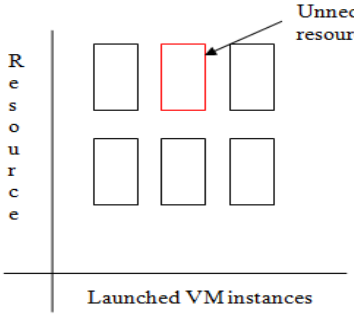

Fig. 2 Resource Utilization in Lab – Partly

III. IMPLEMENTATION

To solve this discussed issue we study this resource utilization content and forcefully terminate those virtual machine instances which are unnecessarily occupying the resources. The solution can be implemented by using simple script which takes the input parameters of utilization of each machine in lab as a value compares it with standard predefined threshold level, and if demand for machine instance comes in between when all Vms (Virtual Machines) are launched, the user is assigned with flag for high priority after any of instance closed. If the demand for machine instance increases we can make the reservation pool in which all machine would be registered and they will be given preferences when next time machine instances are launched.

The input parameter of utilization of each machine will be compared with threshold value and as shown (in fig 2 - the red marked VM instance), if the machine is occupying instance and doing nothing, that instance will be terminated forcefully. Then the reservation list will be checked and the highest priority machine as shown in [8] will have chance to get instance for his work.

The Algorithm is shown in Figure 3, where the discussed scenario is completely shown and represented as a script working at the service provider side. This will keep checking if the instance are actually utilizing the resources of the launched instances or just occupying the resource by holding the instance. And the Highest processing power required block is shown tilted since it requires the Administrator privilege of the root privilege for forcefully terminating the instance.

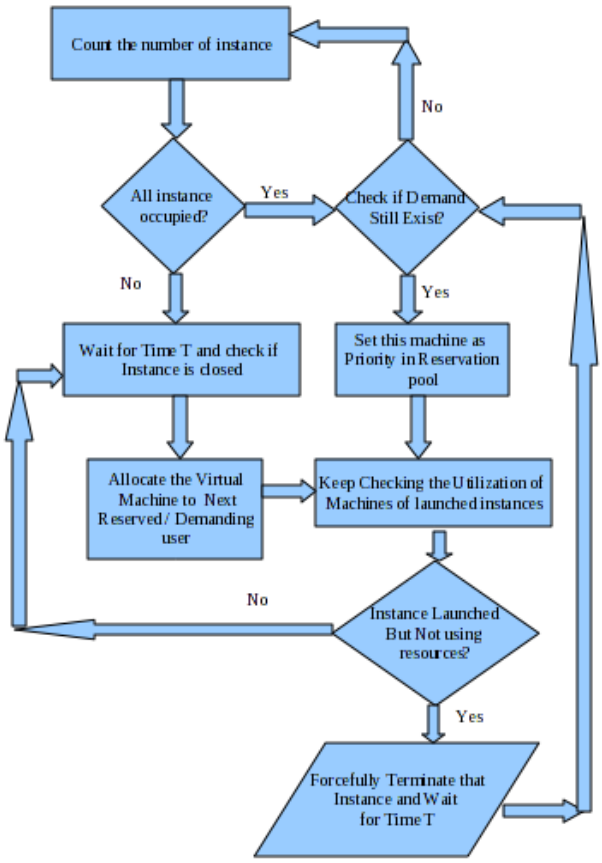

Fig. 3 Implementation Algorithm

This script will keep on executing in the background with root privilege and loop will be maintained including the reservation policy to accurately utilize the resource and give service to legitimate students or users.

IV. CONCLUSIONS

Hence we have studied that the machine instance of cloud server can be forcefully terminated whenever required and the reservation policy can be implemented if necessary so as

to give service to the legitimate user. In this way we can save the cost of modification in infrastructure to accommodate more number of user within institute. This reservation policy also helps in enforcing some user to keep doing there work regularly and in limited time otherwise timer system can be setup that will discard any changes or program that are running on that machine if they take more than required time. This could be the great idea of taking practicals in lab from students within specified time and if student fails to complete his program or his work he have to start it all over again and learn to do it fast by demanding for next machine instance and that will be provided based on reservation policy whether to give him instance or not.

REFERENCES

[1] The Basic Structure of Cloud Service - Models and Deployment, http://cadencema.com

[2] Amies, Alex; Sluiman, Harm; Tong, Qiang Guo; Liu, Guo Ning. "*Infrastructure as a Service Cloud Concepts*". http://www.ibmpressbooks.com/bookstore/=product.asp?isbn=9780133066845.

[3] *"The NIST Definition of Cloud Computing".* National Institute of Standards and Technology. www.csrc.nist.gov/publications/nistpubs/SP800.pdf

[4] *"Defining "Cloud Services" and "Cloud Computing"*. IDC. 2008-09-23 http://blogs.idc.com/ie/?p=190.

[5] Stevens, A(June 29, 2011). "*When hybrid clouds are a mixed blessing*" The Register. www.theregister.co.uk/2011/06/29/hybrid_cloud/

[6] Haff, Gordon (2009-01-27). *"Just don't call them private clouds"* www.news.cnet.com/8301-13556_3-10150841-61.html

[7] InformationWeek.2010-06-30.*"There's No Such Thing As A Private Cloud"*, www.informationweek.com/cloud-computing/theres-no-such-thing-as-a-private-cloud/22920792

[8] Chopde, International Journal of Computer Applications (0975 - 8887) Volume 34- No.9, November 2011

[9] Sambhaji Sarode, Deepali Giri, Khushb, "*The Effective and Efficient Security Services for Cloud Computing*",

[10] Catteddu D. 2010 Cloud Computing. [Online] www.enisa.europa.eu/act/rmfiles/deliverables

