On-Demand Resource Provisioning In Sky Environment

Neha Marria
M.tech Student (Computer Science)
Lovely Professional University
Phagwara, India
neha.marria@gmail.com

Pushpendra Kumar Pateriya
Assistant Professor (Computer Science)
Lovely Professional University
Phagwara, India
pushpendra.mnnit@gmail.com

***Abstract*—In cloud computing resource provisioning is a critical issue and in the environment of heterogeneous clouds (sky) its criticality increases exponentially. In this paper a rule based resource manager is proposed for sky environment which increases utilization of resources in private clouds and reduce the cost. It priorities requests and provide resources accordingly and try to serve requests by using private cloud. To forward a request to public cloud it takes Δt time by doing this it reduce the cost up to a great extent. Performance of resource manager is also evaluated by using CloudSim on the basis of resource utilization and cost in sky environment.**

***Keywords*—*CSM, SSM*,Provisioning,VIM.**

I. INTRODUCTION

Cloud Computing is an ascension trend towards IT Infrastructure. Reducing costs, accelerating processes and simplifying management are all critical to the success of an effective IT infrastructure. Companies are increasingly turning to more flexible IT environments to help them realise these goals. Cloud computing enables tasks to be assigned by combination of software and services over a network. This network of servers is the cloud. Cloud computing can help businesses transform their existing server infrastructures into dynamic environments, expanding and reducing server capacity depending on their requirements. A more formal definition of cloud computing as per Gartner is: "A style of computing where massively scalable IT-enabled capabilities are delivered as a service to external customers using internet technologies" [1].

Today's rapid growth of Cloud Computing touches sky. Burns and Gartner analyst David Smith predicts *Fortune* of cloud computing takeover targets as: Zuora, Nimbula, Heroku. Zuora provides online services to manage and automate consumer subscriptions and payments. Nimbula is a cloud infrastructure and services system that was founded by Amazon's EC2 public cloud service. Heroku is a PaaS provider that emphasizes ease of use, automation, and reliability for web app developers.

This flexibility is the key advantage to cloud computing, and what distinguishes it from other forms of grid or utility computing is providing services that enable the IT functionality to be exposed as services in a multitenant manner. Cloud 'Services' refer to those types of services that are exposed by cloud trafficker and that can be used by cloud consumer on a 'pay-per-use' basis. The types of these services [2] are: Software as a service (SaaS), Platform as a Service (PaaS) and Infrastructure as a service (IaaS).

Software as a Service also refers to Application as a Service. Where the application runs on the cloud provider i.e. own cloud infrastructure and that accessible by cloud consumer through web browser. SaaS is the service based on the concept of renting software from a service provider rather than buying it yourself. Basically it eliminates the need to install, and most recently used because of its high flexibility, great services, enhanced scalability and less maintenance. Yahoo mail, Google docs applications, Salesforce.com CRM apps, Infosys SaaS offerings-Social Platform, E-commerce, Flicker, Facebook are all instances of SaaS.

Platform as a Service, user has access to abstract middleware infrastructure. PaaS are computing platforms where the end users write their own code and the PaaS provider uploads that code and presents it on the web. These are development platforms for which the development tool itself is hosted in the cloud and accessed through a browser. With PaaS, developers can build web applications without installing any tools on their computer and then deploy those applications without any specialized systems administration skills. AppEngine ,Force.com, Bungee Connect, LongJump, WaveMaker are all instances of PaaS.

Infrastructure as a Service is the concept of providing Hardware as a Service. IaaS provides the underlying hardware and operating system resources to do anything you and offers CPU, memory, storage, networking and security as a package for a specific time of service. IaaS is the virtual machine in the sky. This service helps reduce maintenance and usability costs, considering the need for infrastructure management & upgrade. Example includes virtualization, grid computing, and paravirtualization. There are a number of successful IaaS providers: Amazon, Joyent, GoGrid and FlexiScale. While Amazon web services is the best known of the providers.

Cloud computing still has its own characteristics which make cloud computing is cloud computing not anything else. The key characteristics of cloud computing [3] are Service Oriented, Loose Coupling, Strong Fault Tolerant, Business Model, Ease Use, TCP/IP Based, High Security, and Virtualization. These characteristics come under various

considerations like Conceptional, Technical, User Experience and Economic Sector.

Public cloud, Private cloud and virtual cloud are types of Cloud Computing. Public Clouds [4] are external service clouds based on pay-per-use model. It's useful for small scale organizations for computation that reduces to manage the infrastructure and resources are dynamically provisioned via web applications that are provided by third party service providers. Whereas Private Clouds [5] are exclusively built by organization itself and not shared by other organizations, here internal infrastructure, virtualization seemed by organization to a limited number of peoples behind a firewall. It provides high security and flexibility for computing. Scalability of private cloud is beyond certain limit due to upfront cost of building infrastructure. Virtual Cloud Contrive by Public Cloud or Private Cloud for own computation. Amazon EC2 offers Services through Virtual Cloud by dedicating virtual server to hosting plan.

II. RELATED WORK

The cloud of clouds is the innovative level of cloud computing where organization interweaved multiple clouds into single précis cloud adequate to abide the failure of cloud computing. Combining the ability to trust remote sites with a trusted networking environment; we can now lay a virtual site over distributed resources. Because such dynamically provisioned distributed domains are built over several clouds, we call this kind of computing sky computing [6] where multiple clouds become one. By Functional point of view, the cloud of clouds seems as single unified cloud. The heterogeneity is obscure from you and gives more effective computing process.

In the operative Cloud Computing Model, Split up into Front end and Back End. Front End at the side of cloud user and back end is cloud itself where all the servers, virtualization, resources allocation done. As shown in figure 1, Cloud Consumer selects service from service listing.

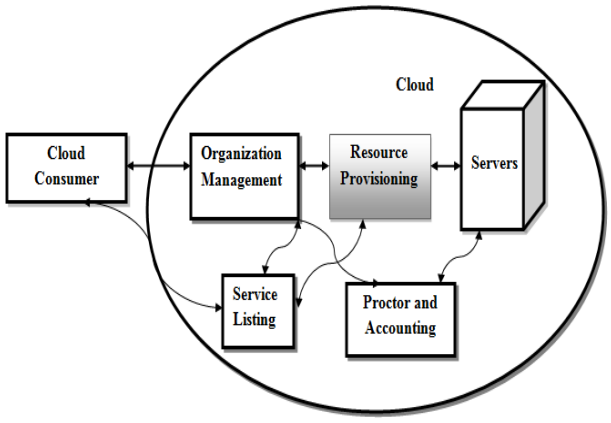

Figure1. Cloud Computing Model

Consumer Request is then forward to organization management tool to find appropriate resources. Further request is transformed provisioning tool to grant resources from the cloud. Service Listing offered the services that are provided by cloud and organization management manages various resources. Resources are allocated by Resource provisioning tool to cloud consumers. Proctor and accounting keep track to various resources and billing of various resources that are used by cloud consumer. Servers are physical resources resides in cloud and handled by organization management tool.

III. RESOURCE PROVISIONING

In resource provisioning, resources are allocated with either a fixed percentage of each partitioning or a relative threshold (dynamic allocation) that can grow and shrink depending upon the application's needs [7]. Resource provisioning gives administrators the ability to manage hardware resources. Resources can be allocated with a fixed percentage of each (partitioning) or a relative threshold (dynamic allocation) that can grow and shrink, within configured limits, as required by cloud consumer. Once allocated, resources are dedicated to a particular consumer and cannot be taken away or used by other consumer or services, unless the resource provisioning configuration is changed to reflect the new resource allocations. Dynamic allocation can makes efficient use of resources.

As Already discussed cloud computing ascension towards IT infrastructure. By time to time Size and complexity of cloud is growing rapidly. So challenge is to manage and utilize cloud computing infrastructure in cost-effective manner. Infrastructure of cloud surrounded towards how effectively cloud consumer service is fulfilled by cloud. Management, utilization and computation all is done at cloud back end side.

Today's most of the business enterprises use Public cloud for the sake of high flexibility, elasticity as well as pay-per-use model. Most of the small scale enterprises use public cloud in their business environment to acquire high productivity by cut down building cost, managing computing of infrastructure. For large scale enterprises adoption of public cloud leads to possible issue. For critical application, data control under cloud service provider, effectuate issue to Privacy and Protection concerns. To overwhelm these exceptions of public cloud, the concept of private cloud originates based on dedicated virtual machines and is limited to organization beyond the firewall. High Security and Flexibility by according to need of computing configuration are key benefit of private cloud. Drawback of Private cloud is scalability i.e. Private Cloud can be scaled beyond certain limit and not much elastic. Another Key issue of Private Cloud is cost of building the infrastructure is too costly rather than adopting from public cloud.

So Resource Provisioning in Cloud Computing is a big deal for small scale enterprises as well as large scale enterprises. For efficient utilization of resources and flexibility to cloud consumer, according to need of user take benefits of public cloud or private cloud or by combining both 'Hybrid Cloud' by mitigating their limitation. Because these Public and private cloud provides complementary benefits. Most of the enterprises uses hybrid model to acquire full utilization of public and private cloud. Because critical applications are serviced by private cloud due to security and privacy concerns and it never cross organization firewall. On another side non

critical applications handled by public cloud due to elastic, interoperability.

By using hybrid model uses of public cloud gives greater extent [8], to overwhelm this problem concept of 'Sky Environment' came into existence where multiple clouds becomes one. To follow up sky environment is too hard. In here complexity increases, and resource on demand provisioned by private, virtual or public cloud. The selection of suited cloud depends upon many circumstances such as data and resource applications.

If there is a large amount of data and processing resource need is small, computation such a request on public cloud won't be efficient than computation it on private cloud. If the data is confidential and critical applications, resources provisioning from private cloud should always be provided so that it never crosses the organization's firewall. Likewise if data and application not critical processing, it may be relocated to public cloud in order to fulfil other critical task request and for better utilization of its own cloud. For efficient and quick utilization relocating on public cloud switch the application to virtual cloud, rather than borrowing from public cloud.

IV. RESOURCE MANAGER

Hadoop, Eucalyptus, 10gen, Zenoss, Puppet are open source cloud computing framework having their own system architecture. Here in these models, all user requests for resources handle through cloud service manager (CSM) manage the resources, proctor and accounting. Further CSM interacts with Virtual Infrastructure Manager (VIM) to handles user's resource request. User request can be for new resource or to scaling up and existing resource request. Further virtual machine allocations done by interacting to service broker. Data centre and physical host machines allocates by broker. All of these CSM, VIM and broker contribute to make cloud as given in figure 2. These components are the key elements of cloud at Infrastructure as a service layer as shown in Figure. In this layer all hardware components are resides also known as Hardware as a service layer. All of cloud services are provides by cloud service manager.

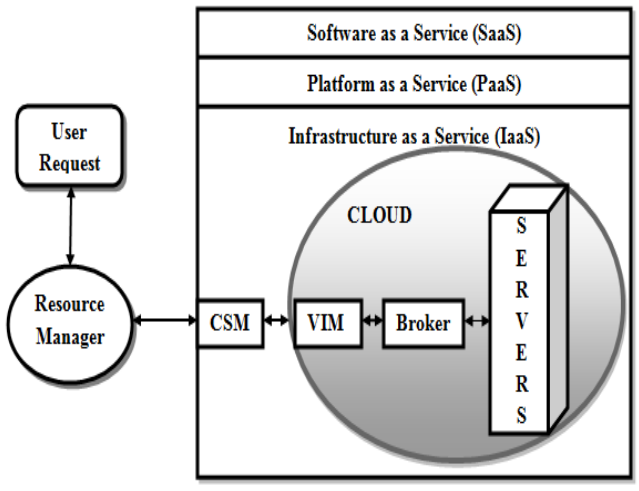

Figure2. Resource Manager

V. MANAGEMENT OF RESOURCES

In hybrid model, Resource manager is used for resource provisioning. User request satisfy by private or public cloud [8]. Concerns towards security aspects, critical application and critical data processing as well as to take benefits of cloud computing, enterprises have to turn their data centre in a private cloud. To provide high scalability, the resources from virtual cloud or public cloud or both are borrowed forming a Sky Environment. The main concern is to efficiently manage the allocation of resource as user request. The effectiveness is achieved by the resource utilization and money spent for using public cloud. The resources of a private cloud must be used maximum and public cloud's usage cost should be minimized. In the next portion a Rule Based Resource Manager is represented for resource provisioning in the Sky Environment.

VI. PROPOSED RULE BASED RESOURCE MANAGER

In sky environment, to effectively utilizing the private cloud resources and considering the security requirements of critical application and data, we propose a Rule Based Resource Manager. With proposed resource manager a private cloud can be scaled up to allocate resources on-demand even if it is overloaded. Decision is made on the basis of rules presented in the following section.

As shown in figure 3, user's request is enter to the Resource Manager. In our approach we categorized user's request into two types based on resource requirements, critical data processing and data security. These two types of requests are assigned priority accordingly, if the user's need is to perform critical data processing or the security demand is high, the request is classified as a high priority and if the request is to run non critical tasks such as routine maintenance, it is classified as low priority. The Resource Manager identifies the suitable cloud to be used to fulfil a request, such as a high priority request must always be acquire resources from the private cloud itself, because it may have confidential information that enterprise would never cross its firewall. Further low priority requests can be fulfilled from either private cloud, virtual cloud or public cloud. But if the private cloud resources are available, it must always be used first as these resources are possess by the enterprise and should be maximally utilized. Determination of provisioning resource acquired through Resource Manager (SSM). Further decision is transferred to Sky Service Manager that switches the request to appropriate cloud. In Sky environment, all clouds are interacts with each other through own cloud service manager. Private cloud generates virtual cloud, if needed.

Farther in our method we consider VM switching, in which the whole VM image is switched from one cloud to another based on priority of requests and the resource availability in private cloud. VM switching is referred by most of the commercial cloud service providers, e.g. Amazon EC2, Hadoop.

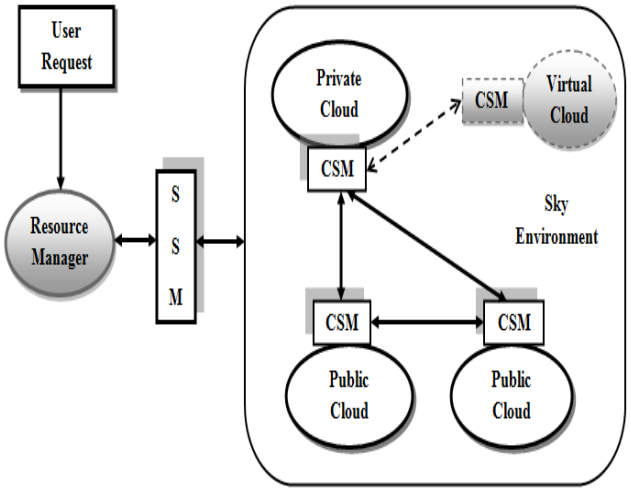

Figure3. Resource Provisioning in Sky Environment

Sometimes it may occur that a high priority request is needed to be fulfilled by the private cloud but its resources are already occupied to fulfil earlier requests of both low and high priority. In this situation, we find those already allocated low priority requests reallocate on virtual cloud. Virtual cloud utilizes free spaces that are available in high priority request, by computing as job pool. As shown in figure 4 Ri requests are in under processing i.e. utilizing or computing the resource and Free space as depicted Fi is merges to fulfil critical application that not done by virtual cloud. Virtual cloud is engendering by private cloud if private cloud is overloaded. Or if number of requests increases and not computed by private or virtual cloud then reallocate low priority requests to public cloud with minimum cost.

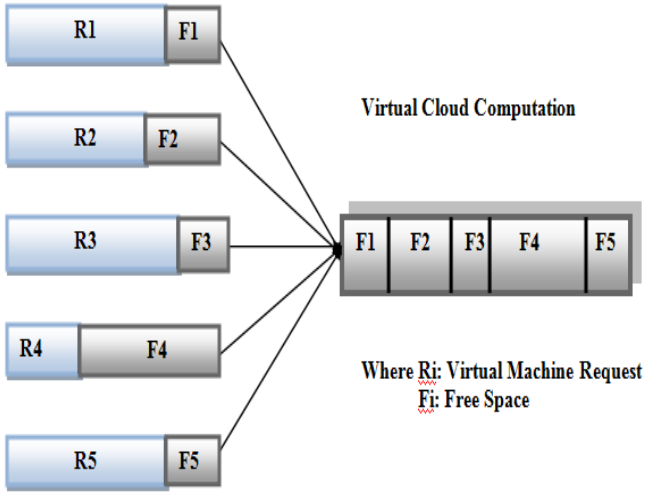

Figure4. Virtual Cloud Computing

The algorithm for Resource Manager is as below:

VM = Virtual Machine

New_VM_Request$_{c+s}$ = New Virtual Machine Request for Compute and Storage Resources

Available$_{c+s}$ = Available Compute and Storage Resources in Private Cloud

High_Priority_Request = Request which Performs Critical Data Processing and Needs Good Response Time

Low_Priority_Request = Request which does not Perform Critical Data Processing

Δ**th** = High_Priority_Request time

Δ**tl** = Low_Priority_Request time

Where $\Delta th > \Delta tl$

x = Allocated Low Priority Virtual Machines

x$_{c+s}$ = Compute and Storage Resources of x

Relocation_Cost_on_Public_Cloud = Public Cloud Usage Cost of a VM, when It is Reallocated from Private Cloud to Public Cloud.

Waiting_Queue = VM Requests are Put in Waiting Queue for Allocation on Private Cloud When Its Resources Become Available.

Algorithm

RESOURCE_MANAGER (New_VM_Request$_{c+s}$**)**

{

RULE 1:

{

If (New_VM_Request$_{c+s}$ $\leq$ Available$_{c+s}$)

Then:

Allocate New_VM_Request$_{c+s}$ on Private_Cloud

}

RULE 2:

If(New_VM_Request$_{c+s}$>Available$_{c+s}$)AND High_Priority_Request

Then:

{

Choose x such that (x$_{c+s}$ $\geq$ New_VM_Request$_{c+s}$) To Relocation_on_Virtual_Cloud

If (x is Non-Empty Set)

Then:

{

Re-Allocate x on Virtual_Cloud

Allocate New_VM_Request$_{c+s}$ on Private_Cloud

}

Else:

Put New_VM_Request$_{c+s}$ in Waiting_Queue

}

RULE 3:

If($New\_VM\_Request_{c+s}$ > $Available_{c+s}$)AND Low_Priority_Request

Then:

{

Put $New\_VM\_Request_{c+s}$ in Waiting_Queue

Until $\Delta th = \emptyset$

}

Else

{

Choose x such that ($x_{c+s} \geq New\_VM\_Request_{c+s}$)

To Minimum $\sum x$ Relocation_cost_on_Public_Cloud

If (x is Non-Empty Set)

Then:

{

Re-Allocate x on Public_Cloud

Allocate $New\_VM\_Request_{c+s}$ on Private_Cloud

}

}

}

VII. EXPERIMENTAL EVALUATION

CloudSim v2.1.1 is used to implement Resource Manager. It is a simulation framework which supports seamless modelling and experimentation of cloud computing infrastructure on a single computer [9]. The simulation is performed on a computer running Ubuntu 10.1 with Linux Kernel 2.6.31 and Pay-per-use model is used. In order to get few requests when private cloud is overloaded, we made 10 resource allocation requests at different time intervals, which are of both types – low priority and high priority. For less than 3 requests, the cloud will not be overloaded, hence performance cannot be evaluated. The priority of these requests are uniformly randomly distributed which is generated by using rand() function of FreeMat [10]. FreeMat is an open source alternative to Matlab. These requests are made for VM varying in computing and storage power.

We simulated different combinations of priorities among these requests for Rule Based and Non-Rule Based approaches. Computing resource is more costly than storage resource, in our simulation we consider only computing resource for resource utilization. In both Rule based and non-rule based approaches the private cloud is getting overloaded, the resource utilization is shown in graph. Private cloud utilization is found to be 65% in Non-Rule Based approach and 74% for Rule Based approach.

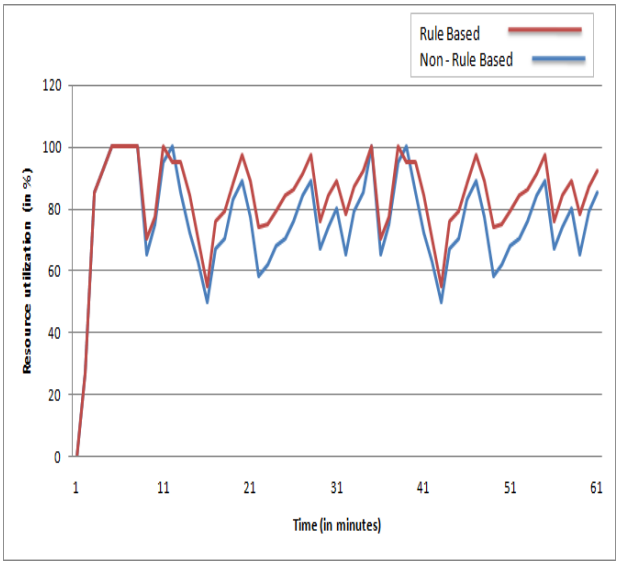

Figure5. Resource Utilization of Private Cloud

So our approach based on Rule Based Resource Manager proves to be cost effective in terms of the most utilization of private cloud resources. Also our approach considers security requirements of confidential data and never permits them to cross the organization's firewall.

VIII. CONCLUSION

Cloud Computing is new technology widely studied in recent years. In this paper, proposed methodology is based on Infrastructure as a Service layer to quickly access of resources on demand. Although study of sky computing provide concurrent access to multiple clouds based on user requirement. Proposed algorithm for resource manager in sky environment gives cost effectiveness by proper utilization of resources from private cloud or virtual cloud. It decreases the borrowing cost of public cloud. Resource Manager provides enhanced resource availability in sky environment. In this domain, framework of sky environment greater extent towards upcoming research work.

Acknowledgment

There are a bunch of people to thank for this paper, including Pushpendra Kumar and Yasir Imteyaz. This paper would not exist but for their faith in me, and I offer them my heartfelt thanks.

References

[1] Gartner Research. Cloud Computing: Defining and describing an emerging Phenomenon.http://www.gartner.com/it/page.jsp?id=1035013, Last Accessed on 20th March,2011.

[2] Minqi Zhou, Rong Zhang, Dadan Zeng and Weining Qian, "Services in the cloud computing era: A Survey", in 4th International conference Universal Communication Symposium (IUCS), 2010

[3] Chunye Gong, Jie Liu, Qiang Zhang, Haitao Chen and Zhenghu Gong, "The characteristics of cloud computing," in Parallel Processing workshops, 2010.

[4] Li, Ang Yang, Xiaowei Kandula, Srikanth Zhang and Ming, "Comparing Public-Cloud Providers," Internet Computing Duke University, March 2011

[5] Sotomayor, B. Montero, R.S. Llorente and I.M. Foster, "Virtual Infrastructure Management in Private and Hybrid Clouds " in Internet Computing, IEEE Univ. of Chicago, Chicago, IL, USA 2009

[6] Katarzyna Keahey, Maurício Tsugawa, Andréa Matsunaga, and José A.B. Fortes, "Sky Computing," in University of Florida, 2009

[7] Hu, Y, Wong, J, Iszlai, G, and Litoiu, "Resource provisioning for cloud computing". In Proceedings of the 2009 Conference of the Center For Advanced Studies on Collaborative Research (Ontario, Canada, November 02 - 05, 2009). CASCON '09. ACM, New York, NY, 101-111. DOI = http://doi.acm.org/10.1145/1723028.1723041

[8] Yasir Imteyaz, Krishna Kant, "On-Demand Resource Provisioning in Hybrid Cloud Environment", Proceeding CNC International Conference on Advances in Communication, Network, and Computing, 2010.

[9] Buyya, R., Ranjan, R. and Calheiros, R.N. 2009. Modeling and Simulation of Scalable Cloud Computing Environments and the CloudSim Toolkit: Challenges and Opportunities. In Proceedings of the 7th High Performance Computing and Simulation Conference (HPCS 2009, ISBN: 978-1-4244-4907-1, IEEE Press, New York, USA), Leipzig, Germany, June 21-24, 2009.

[10] FreeMat. http://freemat.sourceforge.net/, Last Accessed: March 13, 2011.