



Reserve Based Approach for Effective Resource Provisioning in Cloud Computing

Harjinder Singh*

M. Tech. Student, Department of Computer Science
Lovely Professional University,
Punjab, India

Gaurav Raj

Ph. D Scholar, Department of Computer Science
Punjab Technical University,
Punjab, India

Abstract— *In cloud computing environment effective resource provisioning is major issue. In the reserve based approach both security and cost factors are consider by using Hybrid cloud environment. which is combination of both private cloud and public cloud. Our approach is to execute more number of requests at private level only so as to manage the cost and security. private cloud is further divided into two sub groups. Group 1 is reserved for high security requests and group 2 is reserved for low security requests. Every high security should execute at private level so as to maintain confidentiality of data. if both groups are filled and request comes for high security then it sends to the waiting queue and execute later and if both groups are occupied and if any high priority request occupied a resource in group 2 then it shifts to the waiting queue and assign that resource to low priority request otherwise migrate that low priority request to public cloud. In this way Effective resource provisioning should be done by using more resources of private cloud as compared to public so as to maintain cost with security.*

Keywords—*Hybrid cloud, Private cloud, Public cloud, Resource reservation*

I. INTRODUCTION

As Cloud computing is a very new concept of computing. It has been thought of as the future of computing. In cloud computing, software, hardware and network play the main role. The collective efforts of these entities make cloud computing possible. We can visualize the cloud as the cluster of computers which are based on contributed systems that provide services in real time over a network.

Cloud computing is provided everything as a service rather than provided the whole hardware or software to a particular external user. such as Infrastructure as a Service(IaaS), Platform as a Service(PaaS), Storage as a Service(STaaS), Software as a Service(SaaS), Security as a Service(SEaaS), Data as a Service(DaaS), DataBase as a Service(DBaaS), Backend as a Service(BaaS). All these are the basic services provided by the cloud computing. In the business user uses software as a service, with which users are having access to application software and databases. The cloud providers manage the infrastructure and platforms on which the applications run. Software as a Service is referred to as “on-demand software” and is usually priced on a pay-per-use basis. Software as a Service Providers generally price applications using a subscription fee.

Cloud computing providers offer their services according to three fundamental models: Infrastructure as a service (IAAS), platform as a service (PAAS), and software as a service (SAAS) where IAAS is the most basic and each higher model abstracts from the details of the lower models Infrastructure as a Service (IaaS). In this most basic cloud service model, providers offer computers, as physical or more often as virtual machines, and other resources. The virtual machines are run as guests by a hypervisor, Management of pools of hypervisors by the cloud operational support system leads to the ability to scale to support a large number of virtual machines. Other resources in IAAS clouds include images in a virtual machine image library, raw (block) and file-based storage, firewalls, load balancers, IP addresses, virtual local area networks (VLANs), and software bundles.

To deploy their applications, cloud users then install operating system images on the machines as well as their application software. In this model, it is the cloud user who is responsible for patching and maintaining the operating systems and application software. Cloud providers typically bill IAAS services on a utility computing basis, that is, cost will reflect the amount of resources allocated and consumed Platform as a Service(PaaS) In the PaaS model, cloud providers deliver a computing platform typically including operating system, programming language execution environment, database, and web server. Application developers can develop and run their software solutions on a cloud platform without the cost and complexity of buying and managing the underlying hardware and software layers. With some PAAS offers, the underlying computer and storage resources scale automatically to match application demand such that cloud user does not have to allocate resources manually.

Software as a service (SaaS). In this model, cloud providers install and operate application software in the cloud and cloud users access the software from cloud clients. The cloud users do not manage the cloud infrastructure and platform on which the application is running. This eliminates the need to install and run the application on the cloud user's own computers simplifying maintenance and support. What makes a cloud application different from other applications is its scalability. This can be achieved by cloning tasks onto multiple virtual machines at run-time to meet the changing work

demand. Load balancers distribute the work over the set of virtual machines. This process is transparent to the cloud user who sees only a single access point. To accommodate a large number of cloud users, cloud applications can be multitenant, that is, any machine serves more than one cloud user organization. It is common to refer to special types of cloud based application software with a similar naming convention: desktop as a service, business process as a service, test environment as a service, communication as a service Cloud computing is the combination of various technologies. Not a single technology is enough to provide smooth function of cloud computing. The technology as grid computing , virtualization, distributed system, system engineering and service oriented architecture play an important role in cloud computing.

Cloud computing is divided into three types that are private cloud, public cloud and Hybrid cloud, Private cloud is cloud infrastructure operated solely for a single organization and having their own infrastructure that is hardware , software and having its own datacenter. In case of public cloud an organization does not have its own infrastructure to save such higher cost of constructing own cloud environment it purchase cloud services from the third parties on pay per user basis. Hybrid cloud is the combination of both private and public cloud computing so it is cost effective as well as having high security.

II. RELATED WORK

Few researchers have given their views in cloud computing technology. following is described the related research work done in cloud computing.

Yun tan ,Ying Dong Lu IBM Thomas J. Watson , Cathy H. Xia. Provides a solution for provisioning for large scale Cloud Services. This provisioning solution provides better quality of service guarantees at all load conditions. The instance are composed of various resources such as memory, CPU, storage. Another important component in the service offerings are the service level agreement(SLA),which specify the desired targets on various performance metrics that the service provider should meet. This paper focuses on service availability. The objective of resource provisioning is then to plan sufficient capacity of all resource types so as to guarantee the service availability defined in the SLA. For this they present mathematical models that capture the key features of cloud computing.[1]

Renngvit Yanggratoke, Fetahi Wuhib and Rolf stadler provides Gossip-based resource allocation for green computing in large clouds. They propose a generic gossip protocol for resource allocation, which can be instantiated for specific objectives which aims at minimizing power consumption through server consolidation.under overload ,the protocol gives a fair allocation of CPU resources to clients. Simulation results suggest that key performance metrics do not change with increasing system size making the resource allocation process scalable to well above 100,000 servers. server consolidation aims at concentrating the workload onto a minimal number of servers. It is effective because utilization level in datacenters today are often low, around 15%. [2]

Mohammed Alhamad, Tharam Dillon, Chen Wu, Elizabeth Chang researched on Response time for Cloud Computing providers. In this paper they develop performance metrics to measure and compare the scalability of the resources of virtualization on the cloud data centers. First they discuss the need for a reliable method to compare the performance of cloud Second, develop a different types of metrics and propose a suitable methodology to measure the scalability using these types of metrics Finally compare the well known cloud providers using the proposed approach and conclude the recommendation Intelligent usage of resources in cloud computing may help cloud customers to reduce the large amount of IT investments as well as operational costs. However, for critical business application and more sensitive information, cloud provider must be selected based on high level of performance and trustworthiness[3].

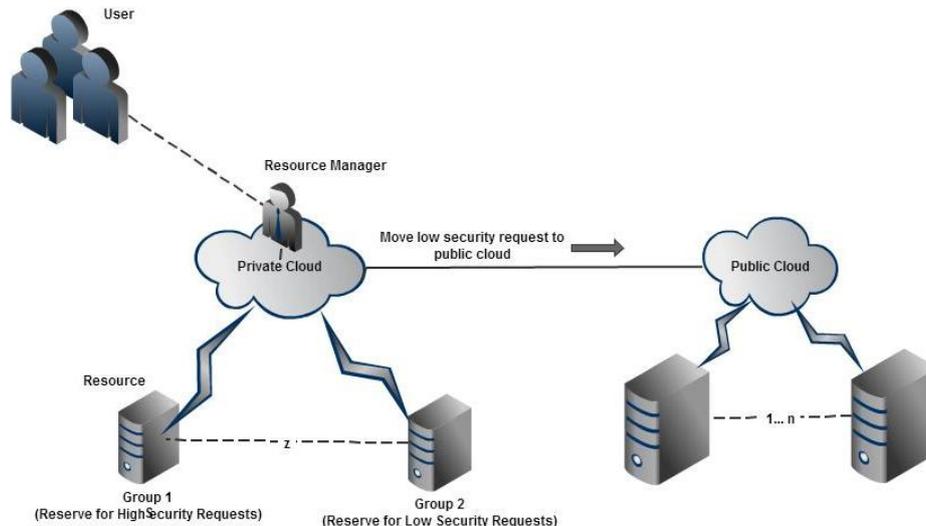
Kyong Hoon Kim, RajKumar Buyya, Jong Kim researched on power aware scheduling of Bag-of-tasks Applications with deadline Constraints on DVS-enabled Clusters. They develop a dynamic voltage scaling(DVS) scheduling Algorithm. It controls appropriate voltage levels for bag of tasks, Minimize the power consumption Also meet the deadline specified by application users. Manage the power consumption using power aware ability of system devices. Lowering the supply voltage of system Due to two main reasons that are **Operational cost and system reliability** This scheduling algorithm supply appropriate voltage to processors to minimize the power consumption. Two types of DVS scheduling algorithms has consider in this paper that are:-Space shared policy and Time shared policy[4]

Rajkamal kaur grewal and pushendra kumar pateria provide a research on rule based resource provisioning in cloud computing. In this paper Rule based resource manager increase the scalability and also the security for resource request. It utilizes both private cloud and public cloud and executes each request which requires very high security at private level only because if that request moves beyond private cloud then security reduces. for low security required requests it only uses public cloud.[5]

Gaurav Raj and Ankit Nischal provides a research on Efficient Resource allocation in Resource Provisioning Policies over Resource Cloud Communication Paradigm. It provides Efficient resource allocation Algorithm which provides complete optimization checking due to which wastage of resources decreases. With the help of performance factor helps in cost optimization. Performance factor depends upon the execution time. So it reduces the time taken by any task to execute. on Efficient Resource allocation(ERA) can tested on different Load conditions and it gives better performance [6]

Study on cloud computing resource scheduling strategy based on the Ant colony optimization Algorithm. This paper is a research on cloud computing technology including its complete structure and mechanism behind cloud computing technology. it has find out that resource scheduling is basic and major function in cloud computing. For detailed analysis of cloud resource scheduling they use Ant colony Algorithm for calculating node distribution and load balancing[7]

III. PROPOSED RESERVATION BASED APPROACH FOR RESOURCE PROVISIONING



In Cloud environment as shown in diagram above two types of requests are came to resource manager first is high Security request and second is low Security request. High Security request required more security as compared to low Security requests. Hybrid cloud is combination of both private cloud and public cloud. our aim is to execute every high Security request at private level so as to maintain the security for such requests because they have high authority data. On the other hand we have to maintain cost as well. So we are using reserve based approach which divides the resources into two groups that is group 1 and group 2. Group 1 is reserved for high priority request and group 2 is reserved for low Security requests. If space in group 2 is available and request comes is of high Security then we can also allocate the resource to high Security request as well. but if after a few second or millisecond a request come for low Security request than we will shift that high priority request to which we have allocated a resource in group 2 just before to the waiting queue and when resource will be available in group 1 or 2 then we will allocate the resource to that high priority request .in another scenario if both group 1 and group 2 is occupied and request comes for low Security request then we will shift that request to public cloud . In this manner we will take advantage of both private cloud and public cloud .brief description of reserve based approach has shown in flow chart and technical working of this approach is as follows

RESOURCE_MANAGER(New_VM_Request)

{

RULE 1:

If(New_VM_Request_Security = High_Security_Request AND
New_VM_Request <= Available_Resource_Group 1)

{
Allocate New_VM_Request on Group 1_Private_Cloud
}

RULE 2:

If(New_VM_Request_Security = High_Security_Request AND
New_VM_Request <= Available_Resource_Group 2)

{
Allocate New_VM_Request on Group 2_Private_Cloud
}

RULE 3:

If(New_VM_Request_Security = High_Security_Request AND
New_VM_Request NOT Available_Resource_Group1 AND
New_VM_Request NOT Available_Resource_Group 2)

{
Insert Request into Waiting Queue
}

RULE 4:

If(New_VM_Request_Security = Low_Security_Request AND
New_VM_Request <= Available_Resource_Group 2)

{
Allocate New_VM_Request on Group 2_Private_Cloud
}

RULE 5:

If(New_VM_Request_Security = Low_Security_Request AND
New_VM_Request NOT Available_Resource_Group 2)

{

```
If(High_Security_Request_Allocate_COUNT > 0)
{
INSERT ALLOCATED_High_Security_Request to waiting Queue AND
Allocate New_VM_Request on Group 2_Private_Cloud
}
```

RULE 6:

```
If(New_VM_Request_Priority = Low_Security_Request AND
New_VM_Request NOT Available_Resource_Group 2)
{
Migrate New_VM_Request to Public_Cloud
};
```

IV. CONCLUSIONS

In Today's environment every investors wants to get more effective services with less cost .Cloud Computing makes this possible by providing everything as a service such as in business environment they mostly access software as a service and Infrastructure as a service . It plays major role in Computer science field. In Hybrid Cloud environment it takes the benefit of both private cloud and public cloud so as to manage the cost as well as security. By using reserve based approach we reserved the resources of private cloud by dividing them into two groups that is Group 1 is reserved for high priority requests which requires high security and having data with high authentication. This approach execute every high priority request at private level only. On the other hand for effective resource provisioning we have to maintain the cost as well as security. For the Remaining Resources that is Group 2 are reserved for low priority requests. If request comes for High priority and resource is available in group 2 then it is assigned to High priority request. And if suddenly low priority request comes then the previous high priority request shifted to waiting queue. and if both groups were occupied then low priority request migrated to public cloud. In this way effective resource provision has done by using reserved based approach which manages both high priority and low priority requests as well.

REFERENCES

- [1] Provisioning for Large Scale Cloud Computing Services Provided by Yue Tan , Yingdong Lu, Cathy H.Xia
- [2] Gossip based Resource Allocation for Green Computing in Large Clouds provided by Renngvit Yanggratoke, Fetahi wuhib and Rolf Stadler
- [3] Response Time for Cloud Computing Providers by Mohammed Alhamed, Tharam Dillon, Chen Wu, Elizabeth Chang
- [4] Power Aware Scheduling of Bag-of-Tasks Application with Deadline Constraints on DVS-enabled Clusters provided by Kyong Hoon Kim, RajKumar Buyya, Jong Kim and Jong Kim
- [5] A Rule Based approach for effective Resource provisioning in Hybrid cloud environment provided by Rajkamal kaur Grewal and Pushpendra Kumar Pateria in 2012
- [6] Efficient Resource allocation in Resource Provisioning Policies over Resource Cloud Communication Paradigm by Gaurav Raj and Ankit Nischal in june 2012.
- [7] Study on cloud computing resource scheduling strategy based on the Ant colony optimization Algorithm by LInan Zhu, Qingshui li and Lingna he.