ISSN:0975-9646

Self Improvement of Cloud Operation Using Dynarep Algorithm and Skewness Algorithm

Prof. M. Senthamarai
Computer Science and Engineering,
Nandha Engineering College, Erode, Tamilnadu, India

Prof.J. Sakunthala
Information Technology
Nandha Engineering College, Erode, Tamilnadu, India

Abstract — **A system that uses virtualization technology to allocate data centre resources dynamically based on application demands and support green computing by optimizing the number of servers are already in use. For quantifying the unevenness in the utilization of multiple resources on a server, known as skewness, the server makes the average utilization of resources. Fault tolerance is also a major concern to guarantee availability and reliability of critical services as well as application execution. To formalize the resource allocation problem as that of dynamically maximizing the cloud utility under CPU and memory constraints, an algorithm called Dynarep is proposed to automatically determine optimal fault tolerance strategy for these significant components.**

Keywords: **Dynarep, Quantifying, Skewness, Optimal fault tolerance**

I.INTRODUCTION

Cloud computing is the access to computers and their functionalities via the Internet or a local area network. Users of a cloud request this access from a set of web services that manage a pool of computing resources (i.e., machines, network, storage, operating systems, application development environments, application programs). When granted, a fraction of the resources in the pool is dedicated to the requesting user until he or she releases them.

It is called "cloud computing" because the user cannot actually see or specify the physical location and organization of the equipment hosting the resources they are ultimately allowed to use. That is, the resources are drawn from a "cloud" of resources when they are granted to a user and returned to the cloud when they are released.

Virtualization is the ability to run "virtual machines" on top of a "hypervisor." A Virtual Machine (VM) is a software implementation of a machine (i.e., a computer) that executes programs like a physical machine.

Each VM includes its own kernel, operating system, supporting libraries and applications. A hypervisor provides a uniform abstraction of the underlying physical machine. Multiple VMs can execute simultaneously on a single hypervisor. The decoupling of the VM from the underlying physical hardware allows the same VM to be started on different physical machines. Thus virtualization is seen as an enabler for cloud computing, allowing the cloud computing provide the necessary flexibility to move and allocate the computing resources requested by the user wherever the physical resources are available.

Given the broad definition of the term "cloud," the current taxonomy differentiates clouds both in terms of cloud service offerings and cloud types. When categorizing cloud service offerings, clouds are referred in terms of "service style" depending on the portion of the software stack delivered as a service.

The most common service styles are referred to by the acronyms IaaS, PaaS, and SaaS. Cloud "types" (including public, private, and hybrid) refers to the nature of access and control with respect to use and provisioning of virtual and physical resources.

Infrastructure as a Service (IaaS)

IaaS clouds provide access to collections of virtualized computer hardware resources, including machines, network, and storage. With IaaS, users assemble their own virtual cluster on which they are responsible for installing, maintaining, and executing their own software stack.

Platform as a Service (PaaS)

PaaS style clouds provide access to a programming or runtime environment with scalable compute and data structures embedded in it. With PaaS, users develop and execute their own applications within an environment offered by the service provider.

Software as a Service (SaaS)

SaaS style clouds deliver access to collections of software application programs. SaaS providers offer users access to specific application programs controlled and executed on the provider's infrastructure. SaaS is often referred to as "Software on Demand."

1.1 Cloud Roles

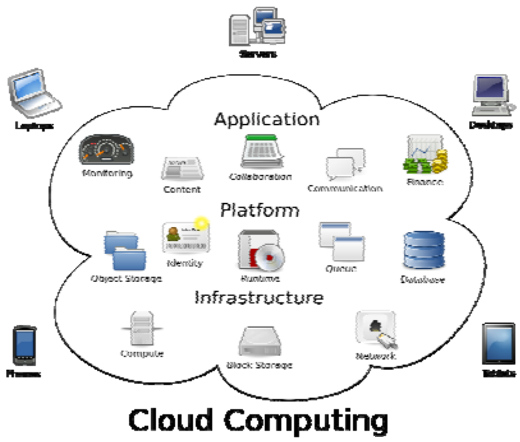

Figure 1.1: Cloud Roles

The Cloud Roles are shown in Figure 1.1 which helps to find out how IaaS and cloud computing applies to Managers, Architects, Administrators, Developers, and Users.

II. CLOUD VIRTUAL MACHNIE ARTCHICTURE

Architects

The roles of cloud architects are:

- Understanding all functional aspects of cloud IaaS, PaaS, SaaS
- Being an expert on cloud resource orchestration
- Platform design for configuring, provisioning, elasticity, and managing
- Identity and access control and implementation

Virtualization technology allows servers and storage devices to be shared and utilization be increased. Applications can be easily migrated from one physical server to another.

Multi tenancy enables sharing of resources and costs across a large pool of users thus allowing for centralization of infrastructure in locations with lower costs (such as real estate, electricity, etc.)

Reliability is improved if multiple redundant sites are used, which makes well-designed cloud computing suitable for business continuity and disaster recovery.

Security could improve due to centralization of data, increased security-focused resources, etc., but concerns can persist about loss of control over certain sensitive data, and the lack of security for stored kernels. Security is often as good as or better than other traditional systems, in part because providers are able to devote resources to solving security issues that many customers cannot afford. In addition, user access to security audit logs may be difficult or impossible. Private cloud installations are in part motivated by users' desire to retain control over the infrastructure and avoid losing control of information security.

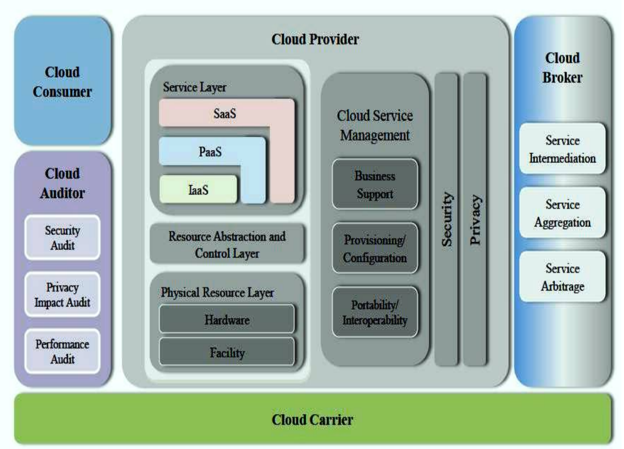

Figure 1.2 Cloud Virtual Machine

The problem of resource management is considered for a large-scale cloud environment. Such an environment includes the physical infrastructure and associated control functionality that enables the provisioning and management of cloud services. This work contributes towards engineering a middleware layer that performs resource allocation in a cloud environment, with the following design goals:

Performance objective: The objective is to achieve max-min fairness among sites for computational resources under memory constraints. Under this objective, each site receives CPU resources proportional to its CPU demand.

Adaptability: The resource allocation process must dynamically and efficiently adapt to changes in the demand from sites.

Scalability: The resource allocation process must be scalable both in the number of machines in the cloud and the number of sites that the cloud hosts. This means that the resources consumed (by the process) per machine in order to achieve a given performance objective must increase sub linearly with both the number of machines and the number of sites.

III. Problem Definition

The problem in large cloud environment is dynamic resource allocation. Such an environment includes the physical infrastructure and associated control functionality that enables the provisioning and management of cloud services. The cloud service provider owns and administers the physical infrastructure, on which cloud services are provided. It offers hosting services to site owners through a middleware that executes on its infrastructure. Site owners provide services to their respective users via sites that are hosted by the cloud service provider.

The cloud provider is having the problem on two aspects, one is overloading and another one is green computing. In overloading problem, there is more number of VMs in a single PM; it may cross the PM's limit. So, the PMs are overloaded, and it will decrease the performance of VMs. Another one is green computing, the cloud manager has to create more number of PMs for small amount of VMs. For minimize the power consumption of PMs, but still satisfying the needs of all VMs, the idle PMs can be turned off to power savings.

In the existing system, another problem is only one cloud provider is used for resource allocation process. This provider has one cluster management controller to find out the load balancing of VMs and PMs. This load balancing is a centralized load balancing mechanism.

IV. Proposed System

The goal of the skewness algorithm is to mix workloads with different resource requirements together, so that the overall utilization of server capacity is improved. By using this skewness algorithm the unevenness (variance) of the resource utilization of a server can be found. To minimize this unevenness and to improve the overall resource utilization of the server, multi- dimensional resource constraints are used. Time complexity of this algorithm is O(n+m), and also load prediction algorithm is used to capture the usage of the resource for the particular application, and find out the load of VMs.

To promote the fault tolerant, DYNAREP (DYNAmic REPlication) algorithm is used. This algorithm incorporates the reliability analysis into the active replication schema, and exploits a dynamic number of replicas for different tasks. The main benefits of implementing fault tolerance in cloud computing include failure recovery, lower cost, improved performance metrics etc.

To overcome the single cloud provider problem, multi cloud providers can be used. Multi-cloud providers can receive more resources from various cloud providers and that resources can be used in VMs for various applications. By implementing Dynamic Distributed load balancing mechanism, load balancing can be dynamically calculated. Therefore it results in the increase of resource utilization.

V. Result

Our proposed system is doing the dynamic replication by the DYNARep algorithm and it is applied for the virtual machines under the cloud for minimizing the system resource utilization and the continues skewness in the virtual machine was monitored and controlled.

VI.Conclusion

In the cloud environment, it is difficult to change the resource demand dynamically. The skewness metric is used to combine virtual machines with different resource characteristics appropriately, so that the capacities of servers are well utilized. Autonomic fault tolerance management is proposed, to facilitate fault tolerance framework for building reliable systems in voluntary resource cloud infrastructure. The basic mechanism to achieve the fault tolerance is replication or redundancy.

Sometimes the virtual machine in the cloud may fail to ping while sending client process for resource allocation. When a virtual machine in the cloud fails to ping, another virtual machine will be provided for allocating client's process in cloud. To improve the fault tolerant, DYNAREP (DYNAmic REPlication) algorithm will be implemented. For dynamic resource allocation, multi cloud providers can be used. Those providers will allocate the resources in the cloud virtual machines. Hence, all the virtual machines in the cloud can be utilized efficiently.

REFERENCES

1. Adam. C and Stadler. R, “Service middleware for self-managing largescale systems,” IEEE Trans. Network and Service Management, vol. 4, no. 3, pp. 50–64,Dec 2007.
2. Armbrust. M Armando Fox, Rean Griffith, Anthony D. Joseph, Randy Katz, Andy Konwinski, Gunho Lee, David Patterson, Ariel Rabkin, Ion Stoica, and Matei Zaharia “Above the clouds: A berkeley view of cloud computing,” University of California, Berkeley, Tech. Rep., Feb 2009.
3. Clark. C., Fraser. K, Hand. S, Hansen. J. G, Jul. E, Limpach. C, Pratt. I and Warfield. A, “Live migration of virtual machines,” in Proc. of the Symposium on Networked Systems Design and Implementation (NSDI’05), May 2005.
4. Jelasity. M., Montresor. A and Babaoglu. O, “Gossip-based aggregation in large dynamic networks,” ACM Trans. Computer Syst., vol. 23, no. 3, pp. 219–252, 2005.
5. Jelasity. M., Alberto Montresor, Ozalp Babaoglu “T-Man: gossip-based fast overlay topology construction,” Computer Networks, vol. 53, no. 13, pp. 2321–2339, 2009.
6. McNett M., Gupta. D, Vahdat. A and Voelker. G. M, “Usher: An extensible framework for managing clusters of virtual machines,” in Proc. of the Large Installation System Administration Conference (LISA’07), Nov. 2007.
7. Pacifici. G., Segmuller. W, Spreitzer. M and Tantawi. A, “Dynamic estimation of CPU demand of web traffic,” in 2006 International Conference on Performance Evaluation Methodolgies and Tools.
8. Rerngvit. Y., Wuhib. F and Stadler. R, “Gossip-based resource allocation for green computing in large clouds,” in 2011 InternationalConference on Network and Service Management.
9. Wood. T., Shenoy. P, Venkataramani. A and Yousif. M, “Black-box and gray-box strategies for virtual machine migration,” in Proc. Of the Symposium on Networked Systems Design and Implementation (NSDI’07), Apr. 2007