Volume 6, Issue 6, June 2016 ISSN: 2277 128X
International Journal of Advanced Research in Computer Science and Software Engineering
Research Paper
Available online at: www.ijarcsse.com

Live Virtual Machine Migration to Support Load Balancing in Cloud Computing Environment

Prakash Chandra
M.Tech Student,
SET, Jain University,
Bengaluru, India

Madhu BR
Assistant Professor,
SET, Jain University,
Bengaluru, India

Chidananda Murthy P
Assistant Professor,
SET, Jain University,
Bengaluru, India

Abstract*— The key idea of cloud computing to provide the pool of computing resources to its user on demand over the internet is made possible because of the virtualization technology. Virtualization is the process to convert the physical entity into the software component. The resulting logical entity is called as virtual machine (VM). The main advantage of the virtualization technology is that multiple VMs can run over the single physical host. The virtual machine sometimes needs to be migrated from one physical host to another host if the physical host is down for the maintenance or to support load balancing. The process to move the virtual machine from one physical host to another while the client is still connected with the negligible down time and total migration time is known as Live Virtual Machine Migration. This paper presents the details about the live virtual machine migration technique to support load balancing in cloud computing environment.

Keyword: Cloud Computing, Virtualization, Live VM migration, Load Balancing, Virtual machine;

I. INTRODUCTION

Cloud computing is large scale parallel processing and distributed computing platform which uses virtualization technology to provide the pool of computing resources to its user on demand over the internet. The cloud service provider provides mainly Platform as a Service (Paas), Infrastructure as a service (IaaS) and Software as a Service (SaaS) to its users based on the Service Level Agreement (SLA) signed between both the parties. The main advantages of cloud computing is reduction in initial investment and maintenance cost. Most of the companies have started using the cloud services and solutions as part of their IT infrastructure so that they can focus on their key business goals and leave the burden of managing the IT infrastructure to the cloud service providers. Most of the cloud users are the startups and the web service providers company which operates 24 X 7, which results in unpredictable amount on load on the server at any point of time. The computing resources need to be efficiently utilized in order to service the entire request on time. In this case the load balancing techniques comes for the rescue, which dynamically distributes the load across the computing resources to provide the uninterrupted services to its clients. This paper focuses on the live virtual machine migration technique to solve the problem of load balancing in cloud.

II. BACKGROUND

A. Virtualization

The fundamental technology that enables the cloud to provide the dedicated resources to its end user is Virtualization. Virtualization is the software that uses the virtual machine monitor (VMM) , generally called hypervisor to provide abstraction for the physical infrastructure from the computing environment. It is very important technology that helps in running multiple Operating systems (OS) and application over the single physical host. The main advantage of the virtualization technology is that it provides resource sharing, fault tolerance, application isolation and cost efficiency in the data centers.

Virtual machines (VM) are the software component which is the result of the virtualization process. Multiple VMs can run over the single physical host, which are managed by the hypervisor. Cloud service provider provides the computing resources in the form virtual machines.

B. Live VM migration

The process to move the virtual machine from one physical host to another while the client is still connected without the perceptible downtime and total migration time is called live VM migration. In the process of migration the memory, storage, processor state and network connectivity associated with the virtual machines are moved from host machine to destination. The live migration of virtual machine in cloud helps in server consolidation, online maintenance, load balancing and mitigating hotspot and cold-spot in cloud server.

C. VM Memory Migration

The post-copy and pre-copy are the two memory migration technique used in live VM migration which are discussed below:

1. Pre-copy memory migration: The pre-copy memory migration consists of following two phases;
   Warm-up phase: In this phase, the hypervisor copies all the memory pages from the source to destination while the VM is still running at source, during this process if any of the page become dirty (i.e. not copied properly) they are re-copied again until the rate of re-copying is not less than page dirtying rate.
   Stop-and-copy phase: After all the dirty pages are copied from the source to the destination, the virtual machine is stopped at the source and resumed at the destination, the time taken between this operation is called downtime which should be kept as minimal as possible. The various techniques have been proposed to minimize the downtime such as probability density function of memory change.
2. Post-copy memory migration: The post-copy memory migration is initiated by transferring the VM's execution state (e.g. CPU state, register, etc.) from the source to destination and then starting the VM at destination. Then it uses the pre-paging technique to transfer all the memory pages associated with the VM from the source host to the target. If the VM demand for the page that is not available at the destination then it raises the page fault, which is redirected to the source which sends the demanded page over the network. The large number of page fault will result in performance degradation of the application running on VM, so efficient pre-paging techniques should be applied to avoid the same.

III. RELATED WORK

The live virtual machine migration is very important area for research to improve the efficient functioning of cloud environment. Many research papers and scholarly articles have been written and they have proposed the unique and innovative method to improve the live VM migration techniques. The paper [1] describes about the live migration of the multiple virtual machine with the resource reservation technology. The model to support live migration of VM consists of four main components: migration decision maker, migration controller, resource monitor, resource reservation controller.

a) Migration decision maker is the key module to support the live vm migration. It implements the various strategies to make the effective migration decision. It can use the machine learning technique to derive the effective migration strategies based on the historical data.
b) Migration controller controls the live migration process of the virtual machines and is responsible for migrating the vm from the source machine to the target machine based on the migration strategies implemented by the migration decision maker.
c) Resource reservation controller implements the various resource reservation strategies to dynamically allocate the resources on the target machine to avoid the live migration failure due to insufficient amount of resources in the target machine.
d) Resource monitor is responsible for continuous monitoring of the resource usage by both the physical as well as virtual machine. This module plays very important role to support making the migration decisions. This paper also describes the resource reservation method which can be implemented for both source and the target machine.

The resources of the source machine include CPU and memory, while the resources of the target machine include the whole resources of the virtual machine.

a) Resource reservation in the source machine: As the live migration of the virtual machine consumes the resources of the source machine, the process can be improved by controlling the system resources of the source machine reserved for the virtual machine.
b) Resource reservation in the target machine: The insufficient amount of the resources in the target machine will lead to failure of the live VM migration process. To avoid that the CPU work load and the memory requirement of the migrating VMs are calculated and the target machine uses this information to create the running virtual machine consuming the same amount of CPU and memory. When the VM get migrated to the target machine these running VMs are shut down to vacant the system resources, which is then occupied by the migrated VMs. The experiment conducted has shown that the efficiency of live multiple virtual machine migration is dependent on factors like memory size, CPU resource and the workloads of these virtual machines.

It has also suggested some of the optimization technique to improve the efficiency of the live migrations:

a) Optimization in the source machine, one method is to allocate more CPU and memory resource to the VM being migrated and another method suggest altering the migrating sequence by letting the VM with small memory requirement to be migrated first.
b) Parallel migration of virtual machine can achieve better efficiency than the sequential migration if there are sufficient amount of resources in the source machines.
c) Workload aware migration suggests migration efficiency can be achieved by migrating the virtual machines according to workload characteristics of VM running on target machines.

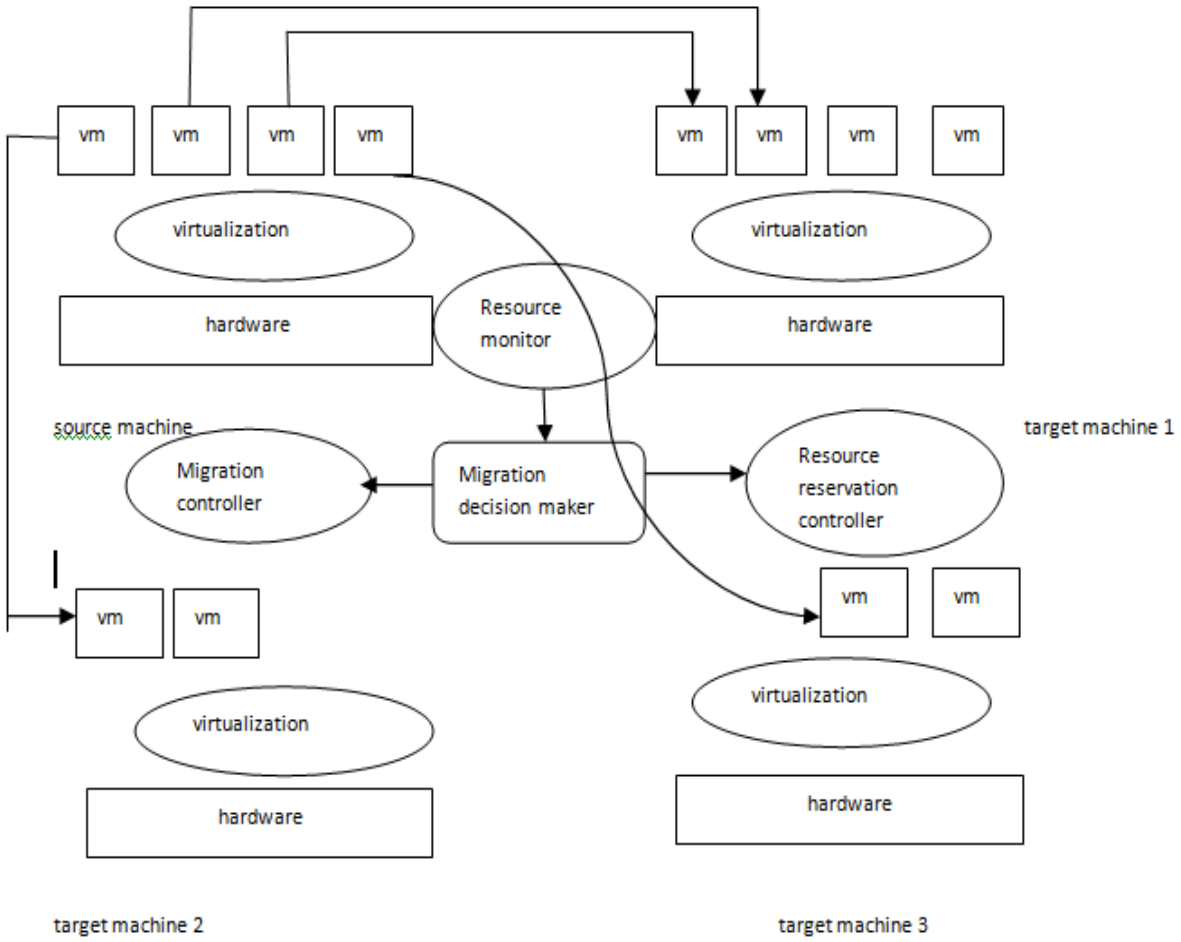

Figure.1 Virtual Machine Migration

The paper [2] has proposed the random graph model to improve the efficiency of the live virtual machine migration. It has proposed the heuristics methods to decide which set of VMs should be migrated to which target server. The VMs having highest correlation of CPU utilization with respect to other VMs is selected as set of migrating VMs. The CPU capacity and the migration opportunity of particular sever is used as the metric to decide the target server. In the graph model, each node is labeled with the number of virtual machine hosted by that server. Each node has certain number of In-degree as well as out-degree , the migration opportunity for that node become higher as the number of In-degrees for that node increases and vice versa. When a node attempts to map the set of VMs to target sever it initiates the random walk, and will decide which is the suitable target by using above mentioned techniques. Once the target sever is selected the VMs are migrated to that server. The most important advantage of this proposed model is that the need for the central VM migration controller is eliminated which can be the performance bottle neck for the migration process as it uses the random walk technique to decide the target server. This paper has also emphasized on the fact that the cloud remain the load balanced although the load characteristics and the migration opportunity of each node vary. In order to maximize the utilization of existing resources and improve the efficiency of live VM migration process this paper [3] has proposed the model which is based on the idea that the available resources are sorted in ascending order according to their execution speed measured in MIPS and the VMs are sorted based on their execution time on these machines in the same order. According to the paper by using this technique the VM having the higher expected execution time will be assigned to the resources having higher execution capability and for the VM having low expected execution time will be assigned to the resources with minimum execution capability. Thus, for the migrating VM their expected execution time is calculated and they are assigned to the suitable resources base on the above mentioned techniques.

This paper has introduced the following three approaches to minimize the migration time of the VM which are as follows:

i) If the host server CPU usage is greater than 90% then target server will be selected with the CPU usage less than 50%.
ii) The resources with the CPU usages of more than 70% will be selected as target if the host CPU is underutilized i.e. usage is less than 10%.
iii) The idle CPU should be switched off(i.e. utilization is 0%) to save power.

By using the proposed model the VM can be migrated to the suitable physical resources thus maximizing its utilization and unnecessarily running resources can be switched off to which lead to less power consumption at the data centers which will help achieve the main objective of GREEN COMPUTING.

IV. CONCLUSION AND FUTURE WORK

Most of the proposed models have been successful in balancing the load of dynamic cloud environment by using live VM migration technique. These models can be further improved for the full automation by using machine learning approach to support the live VM migration. The historical data from different data centers at different point of time can be collected and can be given as a training data to derive a model, which can predict the load at the data centers at different point of time. Accordingly the VM can be migrated reducing the down time and the total migration time.

REFERENCES

[1] Kejiang Ye, Xiaohong Jiang, Dawei Huang, Jianhai Chen, Bei Wang; "Live Migration of Multiple Virtual Machines with Resource Reservation in Cloud Computing Environments"; IEEE 4th International Conference on Cloud Computing, 2011.

[2] Narander Kumar, Shalini Agarwal ; "Self Regulatory Graph Based Model for Managing VM Migration in Cloud Data Centers" ; Advance Computing Conference (IACC), 2014 IEEE International.

[3] Rabiatul Addawiyah Mat Razali, Ruhani Ab Rahman, Norliza Zaini, Mustaffa Samad; "Virtual Machine Migration Implementation in Load Balancing for Cloud Computing" ; 5th International Conference on Intelligent and Advanced Systems (ICIAS), 2014.

[4] Jaspreet Kaur , Manpreet Kaur ,Sahil Vashist ; "Virtual Machine Migration in Cloud Datacenters"; International Journal of Advanced Research in Computer Science and Software Engineering; Volume 4, Issue 8, August 2014

[5] DiegoPerez-Botero; "A Brief Tutorial on Live Virtual Machine Migration From a Security Perspective Princeton University, Princeton, NJ, USA

