Available online at http://www.mecs-press.net/ijem

Enhanced Load Balancing Approach to Optimize the Performance of the Cloud Service using Virtual Machine Migration

Saurabh Jain [a], Dr.Varsha Sharma [b]

[a] *Student at SOIT, Rajeev Gandhi Proudyogiki Vishwavidyalaya,Bhopal,462023,India*
[b] *Assistant Proffesor at SOIT, Rajeev Gandhi Proudyogiki Vishwavidyalaya,Bhopal,462023,India*

Abstract

Cloud computing is a fastest growing technology in the research and industry field.It provides the on demand resources to the customers on the rent basis. These resources are provided through the virtual machines. Resources required by the virtual machines can change dynamically. So load balancing in the cloud is more challenging task as compared to the traditional computing, where the resource requirements are not changed with time. Overall performance of the cloud system can be increased by the efficient load balancing approach. Three steps are involved in the load balancing method i.e., physical machine selection, virtual machine selection and destination physical machine selection. In the past few years a number of load balancing approaches have been proposed to increase the resource utilization and minimize the energy consumption. This paper has proposed a load balancing approach which uses the lower and upper threshold to select the physical machine (PM) for migrating the virtual machine (VM). Then place the selected VM to the PM which consumes minimum power to minimize the energy consumption.

To create the cloud environment, CloudSim simulator is used which provides the interface to deal with the physical and virtual machines. To evaluate the performance, the proposed method is compared with already present load balancing approaches. Simulation result shows that proposed approach minimize the energy consumption, migrations and total simulation time.

Index Terms: Virtualization, migration, energy efficient, virtual machine, physical machine.

1. Introduction

Now days cloud computing is one of the fast growing technologies in the field of computer science and information technology because of online, cheap and pay as use scheme [1, 2]. Virtualization [3, 4]is the key technology in the cloud computing. It increases the resource utilization [5]. Multiple users can share the same

* Corresponding author.
E-mail address:

resources through the virtualization. It is implemented through the hypervisor, which is a small software and responsible for managing all VM related tasks. When user request for the resources, hypervisor create the virtual machine (VM) and assign to the user [6]. VM is a logical machine which behaves like a physical machine (PM). Each VM has its own hardware and software which is created according to the user requirement. Each PM can host multiple VM and these VM may have a different configuration as shown in figure 1.
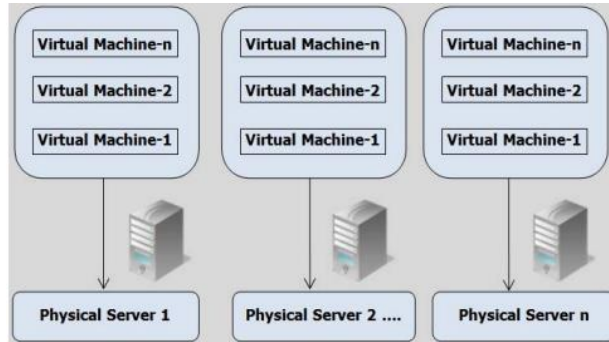

Fig.1. Virtualization Technology

The main focus of the provider is to increase the utilization of the resources. For this purpose provider uses the virtualization. It creates the VM, which must be placed to the appropriate host. So load balancing is an important task in the cloud because it affects the overall system performance [17]. D. Guimaraes et al. [7] states that in ideal state a system consume about 70% of its total energy which it consumed when it is fully utilized. Hence number of active servers plays a vital role in energy consumption. The number of active servers can be minimized through the proper load balancing approaches, which guide to reduce the energy consumption.

Load balancing approach [8] can be static or dynamic. In the static cloud, provider has all resource related information such as resource capacity, processing power, memory etc. which cannot be changed. This static information is used to place the VM. In the dynamic load balancing approach, scheduler uses the run time information and not relay on the historical data. Static approach is more apposite to cloud because user demand changes very frequently. To handle the overloaded and under loaded situation VM migration [9] is used which remap the VM from one PM to another PM. VM migration is the important feature of the virtualization and is used to deal with the overloaded, under loaded and hot spot situations. This paper has proposed a load balancing approach which uses the static value of lower and upper threshold to find the overloading and underloading situation of the PM. Then place the selected VM to the physical machine (PM) which consumes minimum power to minimize the energy consumption.

The rest of this paper is organized as follows. Section 2 describes related work of VM load balancing algorithms. The section 3 describes the structure and the design of proposed model. Section 4 presents experimental result and evaluation. Section 5 gives conclusion and future work.

2. Related Work

In this chapter we discuss the work that has been carried out in the past in the field of Virtual Machine Migration along with differences and similarities between the previous work and ours. We also discuss various performance metrics which is important for VM migration, below we are presenting a review of some of the noteworthy contributions

Y. Fang et al. [10], proposed a two stage load balancing algorithm for the cloud environment. One stage scheduling is responsible for the mapping between the task and VM, The second state scheduler map the VM to the PM resources. This approach considered only task response time. Main problem with this approach is that

task migration is done which is more complex than VM migration.

K. K.Pathak et al. [11], present an extended honey bee load balancing algorithm for the cloud. This approach used the bee behavior for placing the VM. In this approach, VM are assigned to each host and calculate the response time and execution time. Based on these calculations they ranked the hosts. When any request arrived for the placement, scheduler place the VM to the most suitable host i.e. which has higher rank. This approach is suitable for the homogeneous environment, where all VMs and PMs have the same configuration which is not suitable for the cloud.

N. Krishna et al. [12], proposed a priority based VM scheduling algorithm for the cloud. Priorities are assigned to the hosts based on the available host resources. VM is assigned to the host which has higher priority. If the highest priority host has a load factor above 80%, then it finds the node which has the load factor less than 80%. Then the VM is scheduled to that node and the search for the next maximum resource node with the load factor less than 80% takes place. This approach used the fixed load factor that is 80% to define the priority of the node.

Mohammad H. AL et al. [13] work on integrated algorithm that takes different server's load and effectively migrates the virtual machines. They used important factor for reducing energy cost as well as cooling cost at data centers by migrating virtual machines such that energy consumption by virtual machine is minimized and also switching off underutilized servers. When the load on the PM is goes below the lower threshold then scheduler shutdown the PM to save the power consumption. But when the resource demand is increased then the PM is activated again which will take time. This situation can be avoided by switching the PM into the power saving mode rather than shut down.

Gaston Keller et al. [14] introduced a problem in consolidation environments which is how to deal when number of virtual machines demand exceeds the resource capacity of the host. If we know better that which virtual machine to migrate and to which hosts to migrate them, it can ease the situations. In his paper authors proposed a First Fit-based relocation policy, which reckon hosts and virtual machines in dissimilar order. Authors presented the simulation results that exhibit the policies depending on the scenario and the metrics ascertained.

R. Addawiyah et al [15], proposed a load balancing approach based on the threshold. They used two thresholds named lower and upper threshold for scheduling and migrating the VM. If the current resource's CPU usage is greater than 90% (which means that it is overloaded), then the virtual machine will be migrated to another resource with the CPU usage that is less than 50%. If the current resource's CPU usage is less than 10% (which means that it is underutilized), then the virtual machine will be migrated to another resource with CPU usage that is less than 70%. After selecting the overloaded or under loaded host, this approach selects largest VM for the migrations. Main limitation of this approach is that it will increase total migration time as it selects the higher utilize VM for the migration.

3. Proposed Work

VM placement is a challenging task in the cloud computing environment, because resource requirement of the VM can be changed dynamically. CPU, RAM and bandwidth are the three critical resources in the cloud [18]. Main objective of proposed approach is to reduce the energy consumption, which can be reduced by minimizing the number of active serves and by increasing the resource utilization. Resent study [7] reveals that on an average idle server consumes 70% of their total energy, which it consumed at fully loaded. So energy consumption can be reduced by minimizing the number of active servers. In proposed approach, to minimize the number of active servers we firstly choose the largest server available into the data center to place the VM. Main reason to choose the largest server is that if we select the small server than it will increase the number of active servers. Three steps are involved in the load balancing method i.e., source host (PM) selection, virtual machine selection and destination host selection. VM and PM load metrics are used to take the migration decision. So, first we find the load on the PM and VM by using following equation.

$$VM_{CPU}^{util} = \frac{RequestedCPU}{ToatlCPUavailable}$$

$$PM\ (j)_{CPU}^{util} = \frac{\sum_{i=1}^{n} VM(i)_{CPU}^{util}}{TotalCPUofthePM}$$

where n is the number of VM running on the PM

3.1. Source host Selection

To decide when to initiate the migration, we are using lower and upper thresholds. When the utilization of the PM is less than the lower threshold then PM is declared as an under loaded PM and when the utilization of the PM is more than the upper threshold then PM is declared as over loaded PM. In both cases, PM is selected as a source PM and initiates the migrations. Lower and upper threshold of 10 and 90 respectively are used in most of the exiting related works. So, we also set the value of lower and upper threshold are 20 and 80 to decide whether the PM is overloaded or under loaded.

3.2. Virtual Machine Selection

After deciding the host is over loaded or under loaded, some VM need to be migrated to the other PM. Selecting VM for the migration has a significant effect on the performance of the load balancing approach. Since a number of VMs are running in each PM, so deciding which VM to migrate is a challenging task. In the overload situation we choose a VM whose size is equal to the variation between the PM utilization and upper threshold. Following algorithm is used to select the VM from the over loaded or under loaded host.

1.   **for** each h in hostList**do**
2.   hUtil←h.getUtil()
3.   **if** hUtil<lower_threshold**then**
4.   select all VM for the migration
5.   **if** hUtil>upper_threshold**then**
6.   vmList← {All running VM}
7.   **while** hUtil>upper_threshold**do**
8.   temp ←hUtill- upper_threshold
9.   **for** each vm in vmList**do**
10.   vmUtil←vm.getUtill
11.   select the VM whose size is greater than or qual to temp
12.   Add vm to the migration list
13.   PM (MIPS) ← PM(MIPS) - VM(MIPS)
14.   PM (BW) ← PM(BW) - VM(BW)
15.   PM (RAM) ← PM(MIPS) - VM(RAM)

3.3. Destination Host Selection

Suitable PM selection for migrating the VM has a significant role in reducing energy consumption. If VM is migrated to the unsuitable PM then it will increase the energy consumption and number of migrations. Hence,

performance of the load balancing approach mostly depends on the suitable PM selection. Main concept which is used to find the suitable PM for the VM is that, power consumed by the PM before and after placing the VM should be minimum. Following algorithm is used for the destination host selection.

1.   Input: hostList, vmList Output: allocation of VMs
2.   **for each**vm in vmList**do**
3.   **if** t = 0 **then**
4.   arrange all host into the decreasing order of their capacity
5.   **for each** host in hostList**do**
6.   hUtil ← h.getUtil()
7.   **if**hUtill< 100 **then**
8.   Assign vm to the host
9.   PM (MIPS) ← PM(MIPS)+ VM(MIPS)
10.   PM (BW) ← PM(BW)+ VM(BW)
11.   PM (RAM) ← PM(MIPS)+ VM(RAM)
12.   hUtil1 ← h.getUtil()+vm.getUtill()
13.   **else**
14.   Check vmList after t time unit
15.   if vmList≠ Null
16.   **for** all vm in vmList**do**
17.   allocatedHost← Null
18.   **for** all PM in the PMlist**do**
19.   **if** PM has the enough resources **then**
20.   hUtil ← h.getUtil()
21.   hUtil1 ← h.getUtil()+vm.getUtill()
22.   diffList←hUtil1- hUtil
23.   arrange all PM in the increasing order of their diff value
24.   select first PM from the diffList
25.   allocatedHost← host
26.   **if** allocatedHost← Null **then**
27.   Active new host and assign VM.

4. Simulation Result

In this paper, CloudSim [16] is used as a simulation tool to create the cloud environment and evaluate the performance of the proposed approach. For this purpose one data center and 10 physical machines are created. In CloudSim, CPU capacity is represented by MIPS (Million instructions per second). CPU capacity of each PM is either 1000, 2000 or 3000 MIPS, RAM capacity of each PM is 10000MB and BW given to each PM is 100000bit/sec. Number of VMs created in those PMs vary from 20-30. MIPS is given to each VM vary from 250, 500, 750, 1000. RAM assigned to VMs is 128, 512, 1024 and BW allocated to VMs is 2500, 5000, 7500 respectively. Size of the cloudlet (cloudlet is the task that perform by VMs) is 150000 instructions. To evaluate the performance of the proposed approach it is compared with the existing load balancing approach.So first experiment is performed to compare the number of migrations in both approaches (proposed and existing approach).
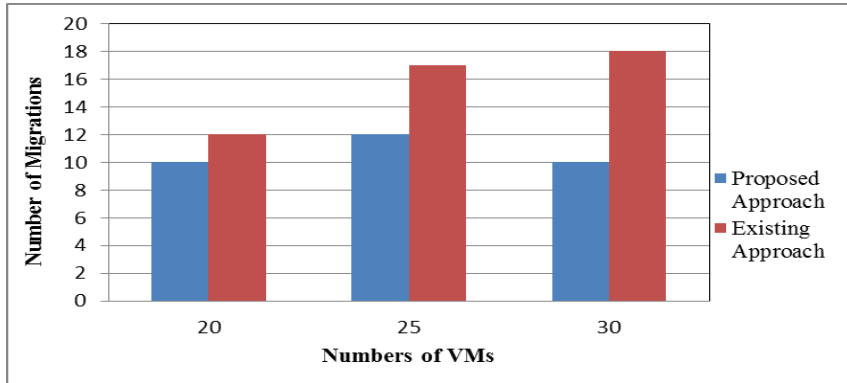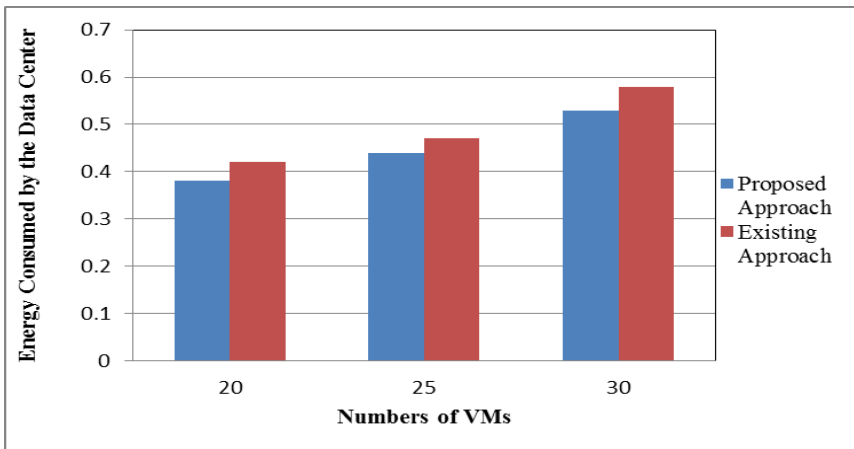
Fig.2. Number of Migrations
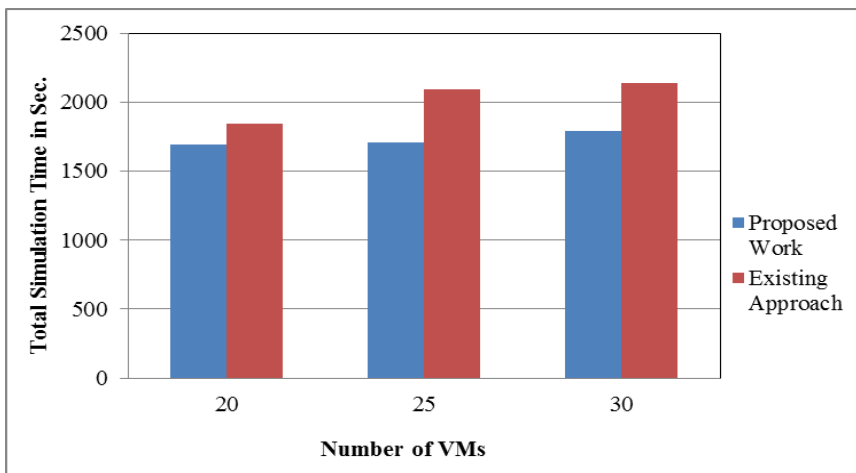

Fig.3. Energy Consumption

Fig.4. Total Simulation Time

Figure 2, 3 and figure 4 shows that proposed approach gives better results compared to the exiting approach. The proposed approach minimizes the number of migrations, simulation time and energy consumption. Hence, performance of the physical machines is increased.

5. Conclusion

Load balancing is a very challenging task in cloud computing because of a dynamic change in the VM resource requirements. It is the essential step for the cloud provider because it affects the overall system performance. An efficient load balancing approach can diminish the number of migration and diminish the energy consumption by minimizing the number of active servers. In this paper, we proposed a load balancing approach, which used the host utilization for migrating the VMs. Proposed approach selects the VM when the PM is overloaded or underloaded and migrates it to the PM which consumes less power. This paper proposed an algorithm for selecting and scheduling the VM. CloudSim simulator is used as a simulator tool to evaluate the performance of the proposed approach. Experiment results shows that proposed approach perform better as compared to the exiting approach and improves the PM performance.

Acknowledgement

References

[1]   Rajkumar Buyya, Chee Shin Yeo, SrikumarVenugopal, James Broberg and IvonaBrandic, "Cloud Computing and Emerging IT Platforms: Vision, Hype, and Reality for Delivering Computing as the 5th Utility", Future Generation Computer Systems, vol. 25, no. 6, june 2011, pp 599-616.
[2]   Peter Mell and Timothy Grance "The NIST Definition of Cloud Computing". NIST Special Publication, 2011.
[3]   Barrie Sosinsky, "Cloud Computing Bible", 1st ed., USA: Wiley Publishing Inc., 2012.
[4]   VMware Inc. VMware distributed power management concepts and use, 2010.
[5]   Peer1 hosting site puts a survey on "Visualized: ring around the world of data center power usage", from engadget.com, 2011.
[6]   Rajeev Kumar Gupta and R. K. Pateriya, "A Complete Theoretical Review on Virtual Machine Migration in Cloud Environment", International Journal of Cloud Computing and Services Science (IJ-CLOSER), Vol.3, No.3, June 2014, pp. 172-178.
[7]   Daniel Guimaraes do Lago, Edmundo R. M. Medeira and Luiz Fernando Bittencourt, "Power-Aware Virtual Machine Scheduling on Clouds Using Active Cooling Control and DVFS", 10th IEEE/ACM Intl. Symp. on Cluster Computing, 2010.
[8]   Mayanka Katyal and Atul Mishra, "Comparative Study of Load Balancing Algorithms in Cloud Computing Environment" Article can be accessed online at http://www.publishingindia.com, 2013.
[9]   Akshay Jain, AnaghaYadav, Lohit Krishnan and Jibi Abraham, "A Threshold Band Based Model For Automatic Load Balancing in Cloud Environment", in proc. of IEEE International Conference on Cloud Computing in Emerging Markets, pp 1-7, 2013.
[10]   Yiqiu Fang, Fei Wang and JunweiGe, "A Task Scheduling Algorithm Based on Load Balancing in Cloud Computing", 18th IEEE International Workshops on Enabling Technologies: Infrastructures for

Collaborative Enterprises, 2010, pp. 271-277.

[11] Kamlesh kumar Pathak, Prasant Singh Yadav, Rameshwaram Tiwari and Dr. Tarun Kumar Gupta "A Modified Approach for Load Balancing in Cloud Computing Using Extended Honey Bee Algorithm", International Journal of Research Review in Engineering Science and Technology, December 2012, pp. 12-19.

[12] Nitish Krishna G, Subramanian S, Kiran Kumar M, Sreesh P and G. R. Karpagam, "ANAdaptive Algorithm for Dynamic Priority Based Virtual Machine Scheduling in Cloud", International Journal of Computer Science Issue, November 2012, pp. 397-402.

[13] Mohammad H. AL Shayejiand and M.D. Samrajesh, "An Energy-aware Virtual Machine Migration Algorithm", proceeding of the International Conference on Advances in Computing and Communications, August 2012, pp. 242-246.

[14] Gaston Keller, Michael Tighe, HananLutfiyya and Michael Bauer, "An Analysis of First Fit Heuristics for the Virtual Machine Relocation Problem.", proceeding of the 6th International DMTF workshop on Systems and Virtualization Management (SVM)/CNSM,October 2012, pp. 406 - 413.

[15] RabiatulAddawiyah, Mat Razali, RuhaniAbRahman, NorlizaZaini and MustaffaSamad, "Virtual Machine Migration Implementation in Load Balancing for Cloud Computing", 5th International IEEE Conference on Intelligent and Advanced Systems, 2014, pp. 1-4.

[16] R. Calheiros, R Ranjan, César A. F. De Rose, R. Buyya, "CloudSim: A Novel Framework for Modeling and Simulation of Cloud Computing Infrastructures and Services", 2011.

[17] Vijaypal S. Rathor, R. K. Pateriya, Rajeev K. Gupta, "An Efficient Virtual Machine Scheduling Technique in Cloud Computing Environment", IJMECS, vol.7, no.3, pp.39-46, 2015.DOI: 10.5815/ijmecs.2015.03.06.

[18] R. Narayani, W. Aisha Banu, "Framework for Provence based Virtual Machine Placement in Cloud", *IJEME, vol. 5, no. 1. pp.19-26, 2015. DOI: 10.5815/ijeme.2015.01.03.*

Authors' Profiles

Saurabh Jain has completed his Masters in Cyber Forensic from School of Information Technology, RG.P.V. Bhopal. His dissertation work was carried out in the field of virtual machine migration in the cloud environment under the guidance of Asst. Proff. Dr. Varsha Sharma of University Institute of Technology, R.G.P.V.Bhopal.