A Review on Different Energy Efficient VM Placement Approaches with their Anomalies in Cloud computing Environment

Shivani Goel
er.shivani25@gmail.com

Tripti Arjariya
tripti.beri1@gmail.com

Abstract- **Cloud is a business oriented model, which provides the on demand computing resources to the client. It is a pay as you go model, so client need to pay only for the used resources like electricity bill. Cloud providers are interested to gain more benefit that can be possible only by the effective utilization of the computing resources. For this purpose virtualization technology is used by the cloud provider. Virtualization allows the sharing of the physical resources. It's enabled the service provider to create multiple virtual machines in a single physical machine. When the user request for the resources, virtual machine manager select the appropriate host according to the constraint i.e. energy consumption, cost, downtime etc. Mapping between the physical machine (PM) and virtual machine (VM) is known as a VM placement. VM placement in cloud is a challenging task due to un predictive and dynamic nature of the VM behavior. Numbers of VM placement approach have been introduced in the past decade. Main aims of these approaches are either to increase the resource utilization or reduce the power consumption by minimizing the number of active server. In this paper we discuses some existing VM placement approach with their anomalies.**

Keywords- VM migration, downtime, Energy efficient, VM placement, Server consolidation.

I. INTRODUCTION

Cloud is a fast growing technology in the field of IT industries [1]. It's become to famous because of their attractive service such as easy to use, flexibility, cheapest etc. According to the NIST definition [2] "Cloud computing is a model for enabling convenient, on-demand network access to a shared pool of configurable computing resources (e.g., networks, servers, storage, applications, and services) that can be rapidly provisioned and released with minimal management effort or service provider interaction" . It supports for the three types of services and can be deploy in four different ways [3, 4].

As show in figure 1 cloud can be deploy in four different way i.e. private, public, hybrid and community. Private cloud is a cloud which is use by single organization with in the network. User can not access the cloud services from outside the network. Whereas public cloud access anywhere in the world.

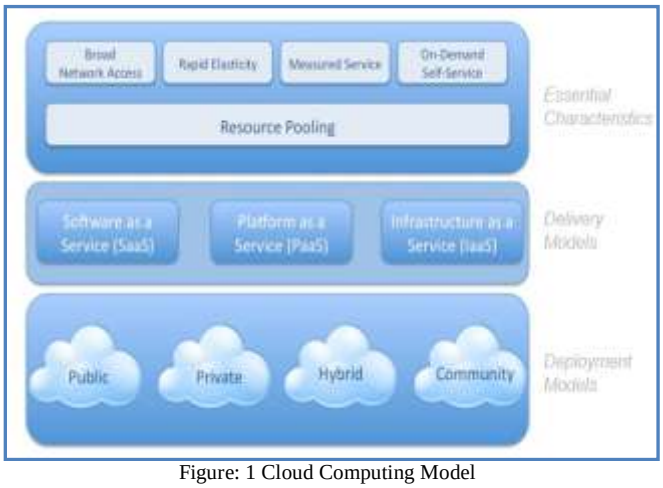

Figure: 1 Cloud Computing Model

Community cloud is a cloud which is share by the multiple organizations and hybrid cloud is a combination of one or more cloud. Cloud supports three types of services Software as a Service (SaaS), Platform as a Service (PaaS) and Infrastructure as a Service (IaaS). In SaaS, only software or applications are provide to the client as a service. Client use these software without ant installation such as Gmail, Facebook. But user does not control the operating system, hardware or network infrastructure on which it's running. In PaaS all computing resources such as hardware, software, network etc. are provide as a service. It is mainly use by the developer not by the normal user. In this type of cloud user does not have any control on the operating system, hardware or network infrastructure on which they are running. Whereas in IaaS all computing resources such as hardware, software, network etc. are provide as a service and user have full control on the operating system, hardware or network infrastructure on which they are running. These services are made available to the users through virtualization techniques.

Virtualization [5, 6] is the backbone of the cloud computing. It is technologies which divide the physical resources and allow creating multiple VM in a single PM. Virtualization is implemented through the hypervisor, which is installed in each physical server and responsible for monitoring the resource utilization.

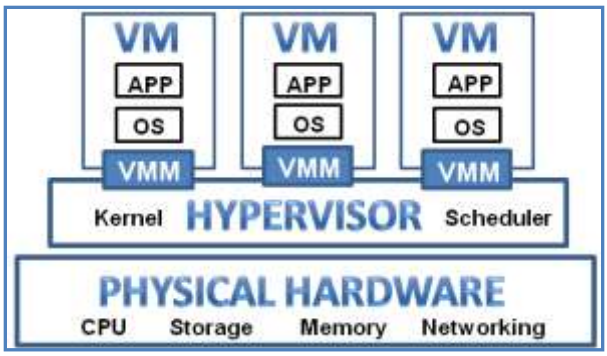

Figure 2: Virtualization

There is a great demand of virtualization in dynamic resource management. It reduces the server sprawl, minimize power consumption, balances load across physical machines. Virtualization technology is so flexible that it transfers the data when a machine hardware is overloaded, thus making easy work for hardware.

In cloud when user request for the resources, virtual machine manager has to decide which host is suitable for hosting this new VM? This decision is taken on the basis of requirement constraint such as energy consumption, service level agreement (SLA) violation, downtime etc. VM placement is a NP hard problem [6]. Proper placement of the VM is a very important task in cloud, because wrong PM selection may increase the number of active server which lead to increase the energy consumption and increase the number of migration. Due to the dynamic nature of the VM, this PM selection is very challenging task. This paper explains some exiting VM placement approach with their anomalies.

II. Related Work

Numbers of VM placement approach have been introduced in the past decade. Main aims of these approaches are either to increase the resource utilization or reduce the power consumption by minimizing the number of active server. First condition which is check by the virtual machine manager is that total resources capacity used by all VM running on that host should not exceed the total capacity of that host. This can be explain as

$$\sum_{j=1}^{m_i} u^i_{cpu} < C^i_{cpu} \qquad \forall j \in m_i$$

$$\sum_{j=1}^{m_i} u^i_{ram} < C^i_{ram} \qquad \forall j \in m_i$$

$$\sum_{j=1}^{m_i} u^i_{bw} < C^i_{bw} \qquad \forall j \in m$$

Where C^i_{cpu}, C^i_{ram} and C^i_{bw} represent total CPU, RAM and bandwidth capacity of the ith host respectively and m_i is the number of VM in i^{th} host.

Cost of power is increasing very sharply [13, 14], so reducing the energy consumption is one of the major issues in clouds. It has been estimated that the cost of powering and cooling accounts for 53% of the total operational expenditure of data centers. In 2006, data centers in the US consumed more than 1.5% of the total energy generated in that year, and the percentage is projected to grow 18% annually [7].

K. Mills et al. [8] proposed a VM Placement Approach based on the First Fit. In this approach when the VM request arrive to the scheduler, it start scanning from the first PM and place the VM to the first PM which have the sufficient resources. If none of the PM satisfies the VM resource requirement, then new PM is activated. This approach may be suffered by the load balancing problem, because some host is overloaded and other are underutilized.

E. G. Coffman et al. [9] proposed a VM Placement Approach which use the First Fit Decreasing (Single Dimension) strategy. In this approach both physical and virtual machine are arrange into the increasing order of their capacity in the term of single dimension. After that first fit approach is used to place the VM. Resources in the cloud are multi dimensional, whereas this approach used the single dimension for placing the VM. So this approach may select the wrong PM.

K. Mills et al. [8] introduce a energy efficient VM Placement Approach which select the VM randomly. In this approach scheduler selects any PM randomly and checks whether the selected PM have the sufficient resources. If yes then assign the VM to the selected PM. If there is no PM with sufficient resources then activate new PM. This approach may increase the number of active server which increased the energy consumption.

Bo. Li et al. [10], proposed an energy efficient VM placement approach. This approach used the migration approach to reduce the number of active PM. Lower and upper threshold are use to define the amount of resources that can be assign to the VM. When the load on the PM is more than the upper threshold or less than the lower threshold migration is trigger. This approach seams good, but use the static threshold which is not suitable for the cloud where the load on the VM is change continuously.

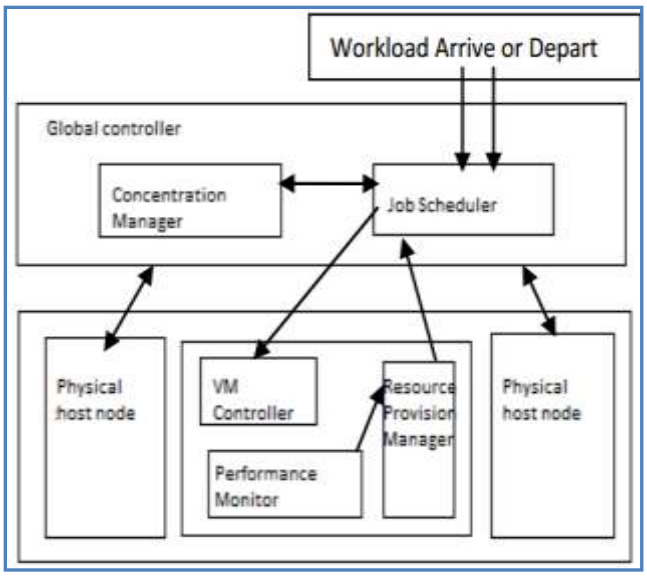

Figure 3: EnaCloud Architecture

Ching-Chi Lin et al. [11], proposed Energy-Aware Virtual Machine Dynamic Provision and Scheduling approach for Cloud Computing. In this approach they use the time interval called retirement threshold. When the host utilization is below the lower threshold, system move to the retirement state? In the retirement state system will not accept new VM and remain in this state till the retirement threshold. After the retirement threshold interval all VM running on that PM are forcefully migrated to the other machine. Limitation of this approach is that they use only CPU as a metric for placing the VM to the PM.

R.Buyya et al. [12], proposed a dynamic threshold based VM placement approach for the cloud. Lower and upper threshold are use to define the amount of resources that can be assign to the VM. CPU utilization of the PM must be in the range of lower and upper threshold value. When CPU utilization of the PM is more than the upper threshold or less than the lower threshold migration is trigger. In this approach they use the concept of server consolidation. So when the load on PM is less than the lower threshold all VM running on that host are migrated to the other host and the host is shut down. T-distribution is use to calculate the lower and upper threshold. This approach seems good because it use the dynamic lower and upper threshold, but main limitation of this approach is that it use only CPU for calculating the lower and upper threshold.

D. Versick et.al [13], proposed an consolidation based approach for reducing the energy consumption. Main idea which is use to reduce the energy consumption is to minimize the number of active server. In this approach initially they calculate the number of host in a cluster based on the resource requirement of the VM. Then K- Mean cluster algorithm is use to assign the VM to the PM.

R. Addawiyah Mat Razali [16], Proposed a load balancing approach for the cloud environment. In this approach all VM are arranged into the increasing order of their CPU utilization. If numbers of VM are waiting for a same resource, even though there are other hosts available to execute the task. This approach migrate some VM to the other host to reduce the waiting time. This approach use 10% and 90% of a CPU utilization as a lower and upper threshold respectively. When the host CPU utilization is below the 10 % of their total CPU capacity then all VM running on that host have to be migrated to the other host and if the host CPU utilization is above the 90 % of their total CPU capacity then some VM running on that host need to be migrated.

Main limitation this approach is that it use the static threshold value as a lower and upper threshold, which is not suitable for the cloud. Second limitation of this approach is that they only use the CPU utilization for calculating the threshold.

III. Conclusion

Energy consumption is one of the major issue in the cloud computing. Cost of power is increase day by day and cloud data center consume too much power which increase the maintenance cost for the cloud provider. It has been estimated that the cost of powering and cooling accounts for 53% of the total operational expenditure of data centers. In 2006, data centers in the US consumed more than 1.5% of the total energy generated in that year, and the percentage is projected to grow 18% annually. Power consumption can be reduced only by minimizing the number of active server that is possible by the VM on the appropriate host. Number of VM placement exist that works under the cloud environment. This paper explains some existing VM placement approach with their anomalies.

References

[1]. R. Buyya et al., “Cloud Computing and Emerging IT Platforms: Vision, Hype, and Reality for Delivering Computing as the 5th Utility", Future Generation Computer Systems, 2011.

[2]. Mell, P. et al.”The NIST Definition of Cloud Computing”, NIST Special Publication, 2011.

[3]. O. D, “Cloud computing service models, part1: Infrastructure as a service,” CEO, Creative RIA, Jan. 2011.

Table 1: Comparison of Different Existing Energy Efficient Load Balancing Approach

Paper Name	System Resources	Virtualization	Goals	Technique
Comparing VM placement algorithms for on-demand clouds [8]	CPU, Memory	Yes	Comparisons of different VM placement approach	Energy Aware Heuristic approaches
Validating heuristics for virtual machine consolidation [9]	CPU	Yes	Minimize the number of active server	VM Consolidation approach
An Energy-saving Application Live Placement Approach for Cloud Computing Environments [10]	Memory Storage	Yes	Minimize energy consumption, application scheduling	Energy Aware heuristic algorithm
Energy-Aware Virtual Machine Dynamic Provision and Scheduling for Cloud Computing [11]	CPU	Yes	Power saving, scheduling, consolidation	VM Consolidation approach
Adaptive Threshold Based Approach for Energy Efficient Consolidation of Virtual Machines in Cloud Data Centers [12]	CPU	Yes	Dynamic consolidation of VM with minimum SLA violation and no. of VM migration	Dynamic consolidation of VM based on adaptive utilization threshold
Reducing Energy Consumption by Load Aggregation with an Optimized Dynamic Live Migration of Virtual machine [13]	Memory, storage	Yes	Minimized energy consumption, minimum running physical machines	Clustering Algorithm

[4]. Sosinsky et al." Cloud Computing Bible", Wiley Publishing Inc 2012.

[5]. S. H. J. G. H. E. J. C. L. I. P. A. W. Christopher Clark, Keir Fraser, "Live migration of virtual machines," in Proceeding NS DI'05 Proceedings of the 2nd conference on Symposium on Networked Systems Design & Implementation, vol. 2, pp. 273– 286, 2010.

[6]. P. Barham, B. Dragovic, K. Fraser, S. Hand, T. Harris, A. Ho, R. Neugebauer, I. Pratt, and A. Warfield, "Xen and the art of virtualization," in Proceedings of the nineteenth ACM symposium on Operating systems principles, ACM, pp. 164 –177, 2003.

[7]. Kim et al., "Experimental Study to Improve Resource Utilization and Performance of Cloud Systems based on Grid Middleware" proceeding in first International Conference on Internet 2013.

[8]. K. Mills, J. J. Filliben, and C. Dabrowski, "Comparing vm-placement algorithms for on-demand clouds." in CloudCom, C. Lambrinoudakis, P. Rizomiliotis, and T. W. Wlodarczyk, Eds. IEEE, 2011, pp. 91–98.

[9]. E. G. Coffman, Jr., M. R. Garey, and D. S. Johnson, "Approximation algorithms for np-hard problems," proceeding in the 7th ACM conference on cloud and approximation algorithms for bin packing: a survey, pp. 46–93, 2010.

[10]. Bo Li, Jianxin Li, Jinpeng Huai, Tianyu Wo, Qin Li, Liang Zhong, "EnaCloud: An Energy-saving Application Live Placement Approach for Cloud Computing Environments", International Conference on Cloud Computing IEEE 2009.

[11]. Ching-Chi Lin, Pangfeng Liu, Jan-Jan Wu, "Energy-Aware Virtual Machine Dynamic Provision and Scheduling for Cloud Computing", 4th International Conference on Cloud Computing, IEEE 2011.

[12]. Anton Belaglozav, R. Buyya, "Adaptive Threshold-Based Approach for Energy-Efficient Consolidation of Virtual Machines in Cloud Data Centers", Proceedings of the 8th International Workshop on Middleware for Grids, Clouds and e-Science, ACM 2010.

[13]. Daniel Versick, Djamshid Tavangarian, "Reducing Energy Consumption by Load Aggregation with an Optimized Dynamic Live Migration of Virtual Machines", International Conference on P2P, Parallel, Grid, Cloud and Internet Computing, IEEE 2010.

[14]. G. Xu et al. "A Load Balancing Model Based on Cloud Partitioning for the Public Cloud" in the proceeding of IEEE conference on TSINGHUA SCIENCE AND TECHNOLOGY, pp. 34-39, 2013..

[15]. A Jain et al., "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.

[16]. R. Addawiyah Mat Razali et al., "Virtual Machine Migration Implementation in Load Balancing for Cloud Computing" 5th IEEE International Conference on Intelligent and Advanced Systems (ICIAS), pp. 1-4, June 2014.

[17]. R.N. Calheiros et al., "CloudSim: a toolkit for modeling and simulation of cloud computing environments and evaluation of resource provisioning algorithms, Software: Practice and Experience", pp. 23–50, 2012.