International Journal of Scientific Engineering and Research (IJSER)
ISSN (Online): 2347-3878
Index Copernicus Value (2015): 62.86 | Impact Factor (2015): 3.791

A Novel VM Allocation Method for Dedicated Servers in Large Scaled Data Centers

Minu Bala

Assistant Professor in Computer Applications, Higher Education Department, Govt. of Jammu & Kashmir, Jammu, India

Abstract: ***The newly emerged concept of dedicated servers for virtual machines in Cloud Data centers has motivated researchers' community to think more critically on optimal utilization of host resources. VMs hosted on dedicated servers are not allowed to migrate during their lifetime, thus it is very important for service providers to adopt that VM allocation policy which utilizes the resources of servers in an optimal way in order to improve the performance of whole data center. Efficient utilization of resources of data center is quite challenging due to unpredictability of workload. Host machines in a cloud data center deal with heterogeneous VMs which have different RAM and CPU requirements. Compute intensive VMs require more computing power than RAM where as data intensive VMs require less computing power than RAM. Similarly other VMs like Small and Micro too have different RAM and computing capacity requirements. An improper VM allocation policy to place different types of VMs in hosts, sometimes leads to uneven usage of host resources and results in wastage of Hosts' resources like RAM and computing power. In the present work, a novel VM allocation policy has been proposed that allocates VMs to hosts keeping in view the RAM and CPU consumption of host in a proportionate way. It uses the concept of Dis-proportionality Coefficient to measure the unevenness in usage of host resources and allocates VM to that host machine which has least value of Dis-proportionality Coefficient. The results obtained are compared with the simple First Fit VM allocation policy and it has been found that proportionate usage of RAM and CPU capacity of the host machine accommodates more VMs than the Simple First Fit policy and also reduces data center's total energy consumption. Moreover results are also compared with Power Aware VM allocation policy and found better in case of the proposed policy.***

Keywords: Cloud Computing, Load Balancing, Virtual Machine, Energy Consumption, VM allocation, Dis-proportionality Coefficient

1. Introduction

A cloud data center is a farm of networked heterogeneous servers which have a wide variety of computing resources [1]. Virtualization is the key feature of cloud computing model that hides the hardware specifications of host machines in the data center and enables to host multiple virtual machines with potentially different resource requirements. Uneven and low utilization of resources in data center is a crucial cost concern for data center's management team. Uneven utilization of server resources results in wastage of resources and requires more physical machines. This results in increase in expenditures for machine power and capital and operational costs for cooling systems. The servers in the data center have to host heterogeneous VMs (each requiring different computing resources), these VMs face unpredictable workload and may cause a resource usage imbalance within hosts. There can be accumulated resource usage imbalance within hosts due to hosting of VMs with unbalanced computing requirements and/or unbalanced workloads [2]. For instance, a VM running a compute-intensive application with low memory requirements may cause a CPU usage imbalance with respect to memory usage. Then, the mapping between VMs and hosts should take into account Computing Power e.g., number of cores with processing capacity of each core and memory size. The mapping of VMs to hosts based on (static) computing resource requirements can be performed when a VM allocation request is received, the mapping based on the time-varying workload of VMs is a challenging task and should be conducted dynamically. In fact, balancing workloads across heterogeneous hosts by dynamically mapping VMs to a set of hosts is an NP-hard problem [6].

The VM allocation problem can be divided into two parts: Handling new requests of VM provisioning and their placement on Hosts and Optimization of current allocation of VMs at runtime. Virtual machine Migration process is an attractive option to save energy in data centers by consolidating VMs to minimum no. of physical machines and switching off or keeping in sleep mode, the unused servers. VMM process is resource intensive and requires additional resources like CPU cycles, memory and bandwidth and have adverse impact on the delivered Quality of Service. In case of applications which have critical business goals and user SLAs, effects of VM migration cannot be overlooked. Hence VM allocation must find the optimal balance between QoS and energy consumption. There is newly emerging concept of dedicated hosts for VMs which is used and provided by Cloud Service Providers nowadays. VMs hosted on dedicated servers are not allowed to migrate during their lifetime, thus it is very important for service providers to adopt that VM allocation policy which utilizes the resources of servers in an optimal way in order to improve the performance of whole data center. A good VM allocation policy helps to serve as many customer requests as possible with the given set of resources. In the present work, the first part of VM provisioning has been addressed in which VM once placed in a host machine remains in that place during its lifetime. The heterogeneity of VMs and Hosts in a cloud data center makes the process of VM allocation quite challenging. Host machines in a cloud data center deal with heterogeneous VMs which have different RAM and CPU requirements. Compute intensive VMs require more computing power than RAM where as data intensive VMs require less computing power than RAM. Similarly other VMs like Small and Micro too have different RAM and computing capacity requirements. An improper VM allocation policy to place different types of VMs in hosts, sometimes leads to uneven usage of host resources and results in wastage of Hosts' resources like RAM and computing power. The VM allocation can be seen as a bin packing problem with variable bin sizes. In simple first fit case, VM is assigned to the first eligible host. This type of approach leads to uneven usage of host resources and results in wastage of Hosts' resources like RAM and CPU cycles. In the present work, we aim to achieve the following goals:

- To devise a policy of VM Allocation that utilizes the resources of Hosts: CPU and RAM, in a data center in a proportionate way.
- To minimize the number of PMs used in a data center as long as they can still satisfy the needs of all VMs. Idle PMs can be turned off to save energy.

In this paper, we present the design and implementation of VM allocation policy based on the concept of proportionate utilization of resources that achieves the set goals. We make the following contributions:

- We introduce the concept of "Dis-proportionality Coefficient (DP_{coeff})" to measure the uneven utilization of server resources. By lowering the value of DP_{coeff}, we can improve the overall utilization of servers.

We develop a resource allocation system that minimizes the number of servers used and reduces energy consumption of data center.

2. Related Work

In modern large scale virtualization-based data centers, the issue of VM placement has become very critical issue and attracted importance attention recently [8]. The high levels of efficiency can be achieved by robust plans for VM placement in cloud data centers. Among various solutions of bin packing problem, Beloglazov and Buyya, proposed a modification of popular Best Fit Decreasing (BFD) algorithm that was shown to use bins, not more than 11/9.OPT+1 (where OPT is the number of bins provided by the optimal solution) [3]. The modified BFD was named PABFD (power aware best fit decreasing) algorithm which first sorts the VMs according to their CPU utilization in decreasing order and then for each VM it checks all the hosts and find the suitable host where the increase of power consumption is minimum and allocates the VM to that host.

Virtual machine migration is a very important process which helps in acquiring various resource management goals like load balancing, power management, fault tolerance, server consolidation etc. in cloud data centers [4]. This process is resource-intensive and requires additional host machines' resources like CPU cycles, memory and bandwidth and sometimes have an adverse effect on the performance of services which cannot be overlooked in cases where user SLAs and business goals are critical. Various optimization techniques have been proposed to assist data center operators with optimizing resource utilization technologies to reduce VM migration noise. Applications' performance can be improved by incorporating various optimization methods like fine granular reduplication, dynamic write throttling instrumentation, memory contents compression during migration . VM migration process can be made more secure by stopping the compromised entities access to VMM, securing the network connections and isolating the VM boundaries.

A concept of VM multiplexing has been introduced by Meng et. al for efficient resource provisioning in compute cloud [5]. VM multiplexing can be used to save the capacity of hosts by consolidating and provisioning VMs during unaligned peaks and troughs in multiple VMs. Three design modules: a new SLA model, a joint-VM sizing technique and a new VM selection algorithm have been presented to implement the concept and can be plugged into already existing resource provisioning services. It improves the ratio of newly admitted VMs by 16% on an average and up to 75% with stringent SLA requirements.

Structural Constraint-aware virtual machine placement (SCAVP) policy has been proposed by Jaya singhe et. al, 2011 to improve the performance and availability of services hosted on modern data centers [9]. The proposed policy supports three types of constraints: demand, communication and availability. They used hierarchical placement approach with four approximation algorithms that efficiently solves the VM placement problem for large problem sizes.

Nguyen et. al have addressed the problem of automatic virtual resource management in cloud [10]. A two level architecture that separates generic decision making layer from application specific functions has been suggested in present work. Constraint programming and utility functions are used for self-optimization. This approach is better than rule and policy based systems as it avoids the problems encountered in later case.

3. System Model

A large scale data center consists of m heterogeneous servers which provide IaaS to its clients. Each server in the data center consists of processor which has single or multiple processing elements, called cores, whose processing capacity is defined in MIPS (millions of instructions per second), RAM, and network bandwidth. The processing capacity of a server or physical machine with 'a' no. of processing elements in its processor, each having 'b' MIPS of processing power is calculated as a*b MIPS. IaaS provider has no prior knowledge of type and quantity of workload. Various independent users send their requests for creation of n heterogeneous VMs. The VMs are characterized by CPU power in MIPS, RAM and network bandwidth. As VMs are used for various purposes like HPC, storage, web applications etc. and are managed and owned by various users, a mixed kind of workload is faced by cloud data centers. The data center is capable of handling four types of VMs: High-CPU Medium Instance, Extra Large Instance, Small Instance and Micro Instance creation requests. It has been assumed that the processing power of each VM must be less than or equal to processing power of a single CPU core of a host.

4. Problem Formulation

4.1 VM Allocation

The placement of VM into a PM depends upon the resources requested by VM and resources available with PM. The simple VM Allocation Policy is based on First Fit algorithm in which VM is allocated to the first eligible machine or host in the data center. This policy is quite fast as it searches as little as possible but it has major drawback of disproportionate utilization of Host machine's resources like RAM and processing capacity in terms of MIPS. VMs are heterogeneous and the main elements which differentiate the VM types are the RAM Capacity and the Processing Capacity. The allocation of VMs with higher RAM requirements and lesser processing requirements in the same host machine results in disproportionate utilization of RAM and processing capacity of host and vice versa. This results in wastage of Host machine's resources because it may not be eligible for other VM due to RAM or MIPS requirement failures. In the present work a novel technique for VM Allocation based on the concept of Proportionate Resource Utilization has been proposed.

4.2 Overview of Proposed VM Allocation Policy

Allocation of same type of VM instances in a physical machine makes the disproportionate utilization of host resources. For example, allocation of VM instance of type I consumes more CPU MIPS and less RAM. Allocation of same type of VM instance again to the same node again leaves less no. of CPU MIPS but more RAM. This uneven leftover CPU and RAM combination may not be suitable for some other type of VM instance and remains free. This results in inefficient use of cloud resources. In the proposed policy, a mechanism of allocation/placement has been suggested in which VM is assigned to that physical machine whose allocation makes the ratio of CPU MIPS and RAM less disproportionate than the other eligible physical machines. The CPU MIPS and RAM requirements of candidate VM are subtracted from eligible host's CPU MIPS and RAM and then the ratio of leftover MIPS and Leftover RAM of candidate host is calculated. This process is repeated for all eligible hosts and VM is assigned to the host whose ratio of leftover MIPS and leftoverRAM is more close to 1 i.e. more proportionate. The above mentioned concept can be implemented by calculating the DP_{coeff} of resource usage of each host in the data center and allocating the VM to the host whose DP_{coeff} value is minimum.

4.3 Proposed VM Allocation Policy

The proposed policy works in two phases: (i) Eligibility Check-up (ii) Results Optimization

Eligibility Check-up: The host machine which fulfils the following four pre-laid criteria for accommodating a VM is eligible for VM placement.

- The processing capacity of each PE (Processing Element) of Host machine must be greater than or equal to processing capacity of PE of VM to be allocated.
- Host Machine's available processing capacity must be greater than or equal to the processing capacity of VM to be allocated.
- Host machine must have available RAM greater than or equal to the RAM requested by VM to be allocated.
- Host machine must have available BW (Bandwidth) greater than or equal to the BW requested by VM to be allocated.

Results Optimization: Once all eligible hosts for a VM are found, the DP_{coeff} of resource usage of each eligible host is computed and finally, VM is assigned to the host whose DP_{coeff} value is minimum. In other words, VM is assigned to that host which leaves the resources of host more proportionate than the other hosts in the data center.

Concept behind Optimization: Proportionate Resource Utilization

Each server in a data center has multiple resources. Let m denote the no. of resources of a server under consideration and u_i denote the utilization of ith resource of a server. The Dis-proportionality Coefficient ($DP_{coeff.}$) of resource utilization of server is defined as:

$DP_{coeff.} = \sqrt{\sum_{i=1}^{m} \left(\frac{u_i}{\bar{u}} - 1\right)^2}$, where $\bar{u}$ is the average utilization of all resources u_1, u_2 ... u_m of server. In present study, m = 2 (i.e. CPU & RAM), $\bar{u} = u_1 + u_2$, where u_1 is CPU utilization percentage and u_2 is RAM utilization percentage of a server

Algorithm : Proportionate Resource Utilization Based VM Allocation Algorithm

```
Input:   host_list     List of Hosts in the Data centre
         vm            Virtual Machine to be Allocated/Placed in a Host in
                       the Data centre
Output:  allctd_host   Allocated Host for VM.

1.  Initialize disMax= MAX_value and allctd_host=null
2.  for each host in the host_list do
3.   freeHostMips ← totalHostMips - hostUtilizationMips
4.   freeHostRam ← totalHostRam - hostUtilizationRam
5.   leftOverHostMipsPer ← ((freeHostMips – vmMips)/totalHostMips)*100
6.   leftOverHostRamPer ← ((freeHostRam – vmRam)/totalHostRam)*100
7.   avgResUtiliz← ( leftOverHostMipsPer+ leftOverHostRamPer)/2.
8.   DP ← sqrt (sq((leftOverHostMipsPer/ avgResUtiliz)-1) + sq((leftOverHostRamPer/ avgResUtiliz)-1)).
9.   If (DP < disMax) then
         disMax = DP
         allctd_host=host
10. End if
11. end for
12. return allctd_host
```

5. Performance Evaluation

The experiment has been designed to evaluate the proposed PRU based VM allocation policy. The simulation experiment has been conducted on a single computer using Cloudsim - 3.0.3 on Eclipse SDK [11]. The hardware configuration of the computer is shown as follows: Intel(R) Core(TM) i5-3770 CPU @ 3.40 GHz, the OS is Windows 8, RAM is 8 GB and its system architecture is 64bit. The cloud scenario that was created for experimentation consists of one data center with 4 heterogeneous hosts. The two systems which have been modeled in the present study are: HP ProLiant ML110 G4 and HP ProLiant ML110 G5. The frequencies of the servers CPUs were mapped onto MIPS ratings: HP ProLiant ML110 G4 consists of 1860MIPS Processing Speed, 2 Processing Elements, 4 GB RAM, 1 Gbits/s Bandwidth, 1 GB Storage and HP ProLiant ML110 G5 consists of 2660MIPS Processing Speed, 2 Processing Elements, 4 GB RAM, 1 Gbits/s Bandwidth, 1 GB Storage. Four types of Virtual Machine instances, considered for simulation are: i) High-CPU Medium Instance (2.5 EC2 Compute Units i.e. 2500MIPS, 0.85 GB RAM), ii) Extra Large Instance (2 EC2 Compute Units i.e. 2000MIPS, 3.75GB RAM), iii) Small Instance (1 EC2 Compute Units i.e. 1000MIPS, 1.7 GB RAM) and iv) Micro Instance (0.5 EC2 Compute Units i.e. 500MIPS, 0.633 GB RAM). The workload samples that have been considered for the performance evaluation of proposed model are the workload traces taken from CoMon project, a monitoring infrastructure for PlanetLab[11]. The workload traces consist of CPU utilization of thousands of virtual machines from hosts located at different places in different geographical areas of the world which have features like large data volume, various data types, low value density and fast processing speed

5.1 Results and Analysis

With proposed PRU based VM allocation policy, the average CPU Utilization of Host 0 is 20.49%, Host 1 is 40.47%, Host 2 is 20.34% and Host 3 is 41.33% where as in case of other scenario, the average CPU Utilization of Host 0 is 27.28%, Host 1 is 45.99%, Host 2 is 13.56%, Host 3 is 17.50% and Host 4 is 18.41%.

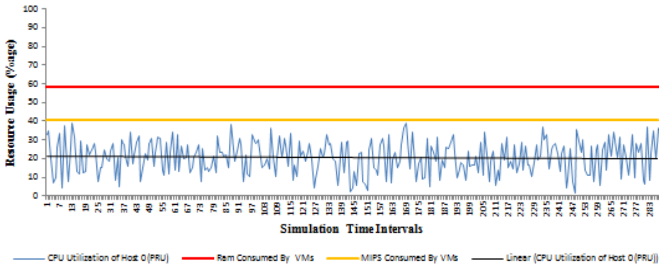

Figure 3(a): CPU Utilization, RAM & MIPs consumed by VMs allocated to Host0 on PRU Based VM Allocation Policy

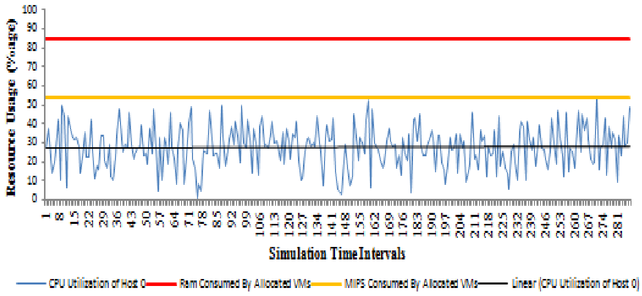

Figure 3(b): CPU Utilization, RAM & MIPs consumed by VMs allocated to Host0 on Simple First Fit Based VM Allocation Policy.

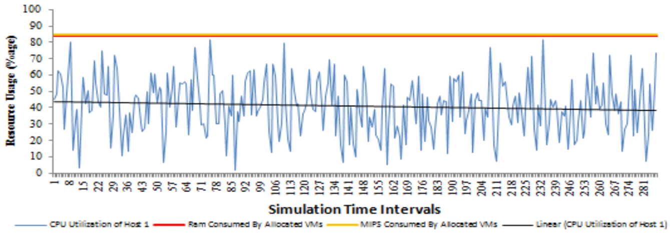

Figure 4 (a):CPU Utilization, RAM & MIPs consumed by VMs allocated to Host1 on PRU Based VM Allocation Policy

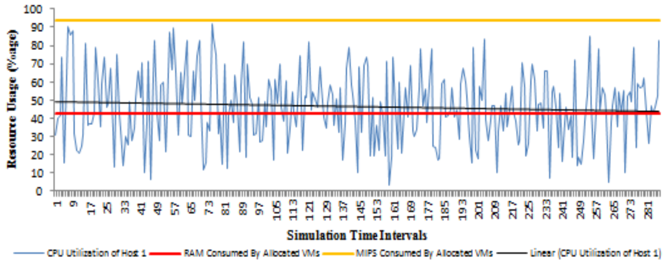

Figure 4 (b): CPU Utilization, RAM & MIPs consumed by VMs allocated to Host1 on Simple First Fit Based VM Allocation Policy

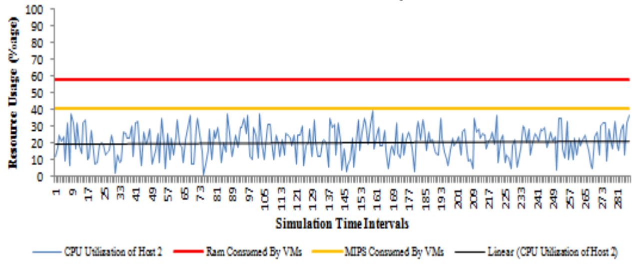

Figure 5 (a): CPU Utilization, RAM & MIPs consumed by VMs allocated to Host2 on PRU Based VM Allocation Policy

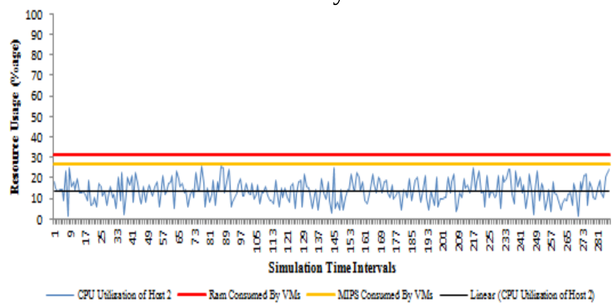

Figure 5 (b): CPU Utilization, RAM & MIPs consumed by VMs allocated to Host2 on Simple First Fit Based VM Allocation Policy.

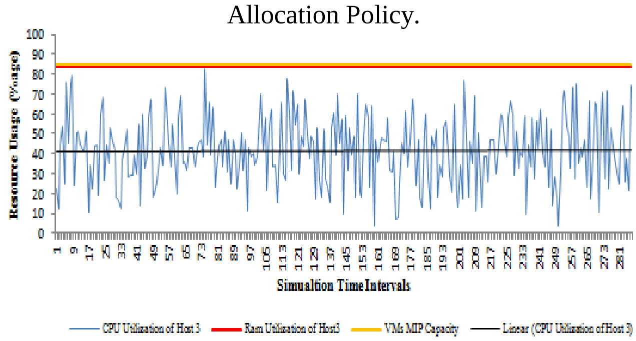

Figure 6 (a): CPU Utilization, RAM & MIPs consumed by VMs allocated to Host3 on PRU Based VM Allocation Policy.

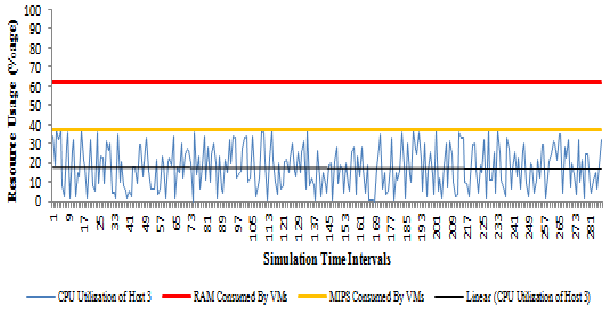

Figure 6 (b): CPU Utilization, RAM & MIPs consumed by VMs allocated to Host3 on Simple First Fit Based VM Allocation Policy

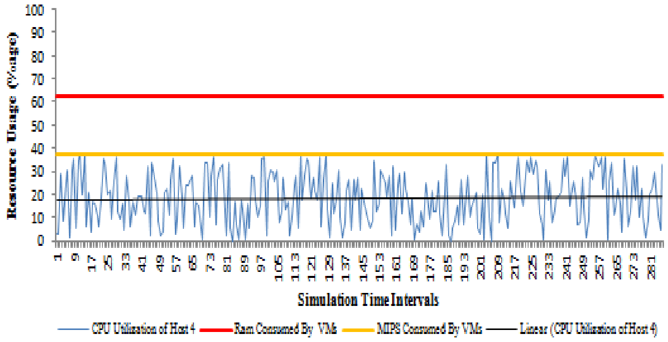

Figure 7 (a): CPU Utilization, RAM & MIPs consumed by VMs allocated to Host4 on PRU Based VM Allocation Policy

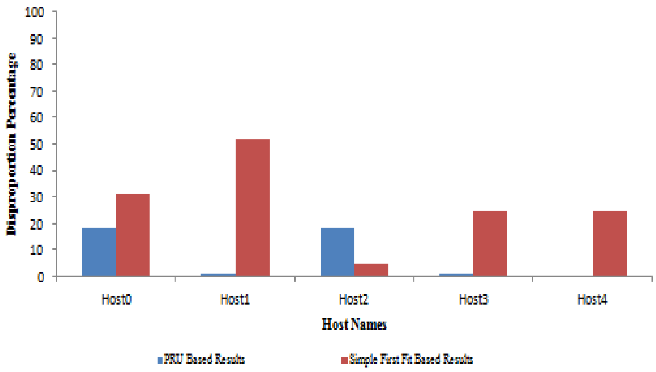

Figure 7 (b): Percentage of Dis-proportionate Resource Utilization (CPU & RAM) in various Hosts in data center based on PRU and Simple First Fit Based Policy.

Figures 3(a), 4(a), 5(a) & 6(a) depict the CPU utilization, Ram consumed by VMs, and the processing capacity (in MIPS) allocated to VMs placed on Host0, Host1, Host2 & Host3, respectively, with proposed PRU based VM allocation policy and figures 3(b), 4(b), 5(b) & 6(b) depict the same with simple First Fit VM allocation policy. VM4 and VM6 are assigned to Host0, VM0 and VM2 are assigned to Host1, VM5 and VM7 are assigned to Host2 & VM1 and VM3 are assigned to Host3 in case of the PRU based VM allocation policy where as VM4 and VM5 are assigned to Host0, VM0 and VM1 are assigned to Host1, VM6 and VM7 are assigned to Host2 & VM2 is assigned to Host3 in case of simple First Fit VM allocation policy. In case of simple First Fit VM allocation policy, there was no eligible host left for VM2 to be allocated, hence a new host i.e. Host4 was added during runtime and VM2 was allocated to it. Figure 7(a) depicts the CPU utilization, Ram consumption, and the processing capacity consumption of newly added Host4.

Moreover, lesser is the distance between the lines depicting MIPS consumed and RAM Consumed, better is the usage of resources of Host machine and greater distance between the lines depicts the amount of disproportion in usage of Host machine's resources. It is clear from the graph that placement of VMs in PRU case is better than the second case because the distance between lines depicting RAM and CPU MIPS consumption is less in PRU case and more in simple First Fit case. The average CPU utilization of Host0, Host1, Host2 and Host 3 is 20.62%, 40.92%, 20.47% and 41.49%, respectively, in case of PRU based policy and 27.45%, 46.28%,13.63% and 17.62%, respectively, in case of simple First Fit VM allocation policy. The average CPU utilization of additional machine i.e. Host4 in case of simple First Fit case is 18.52%.

In figure 4(a) the distance between the lines is almost negligible in case of PRU based policy where as there is significant distance between the lines in figure 4(b). Similar results have been found in case of 6(a) and 6(b). From figure 7(b), it is clear that percentage of disproportionate usage of Host Machine's resources is more in simple Fist Fit policy and is significantly less in case of PRU based policy.

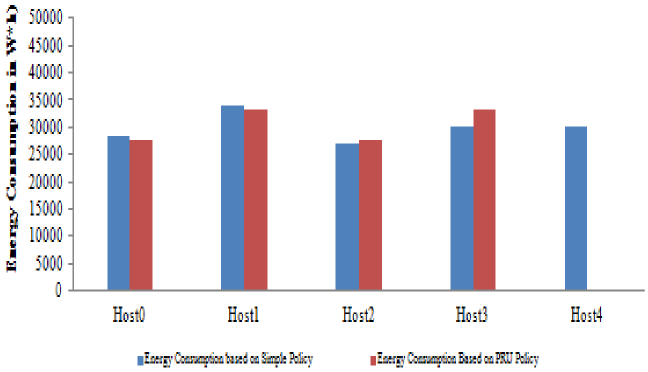

Figure 8: Energy Consumption of data center with PRU based and Simple First Fit based VM Allocation Policies

As both the policies are non power aware, the energy consumption in both the cases is almost same. The average energy consumption of Host0, Host1, Host2 & Host3 is 27760.94 W/h, 33188.94 W/h, 27749.07 W/h and 33272.84 W/h, respectively in case of PRU based policy and 28434.6266 W/h, 33953.02 W/h, 27069.65 W/h & 29958.799 W/h, respectively in case of simple Fist Fit VM Allocation policy. The addition of new host machine in runtime i.e. Host4 in the data center has increased the total energy consumption. It is clear from fig. 8 that addition of new host increases the total energy consumption of data center in case of simple First Fit VM Allocation policy. The total Energy Consumption of data center in PRU based policy is 9.72 KW/h and in case of simple First Fit case is 11.92 KW/h.

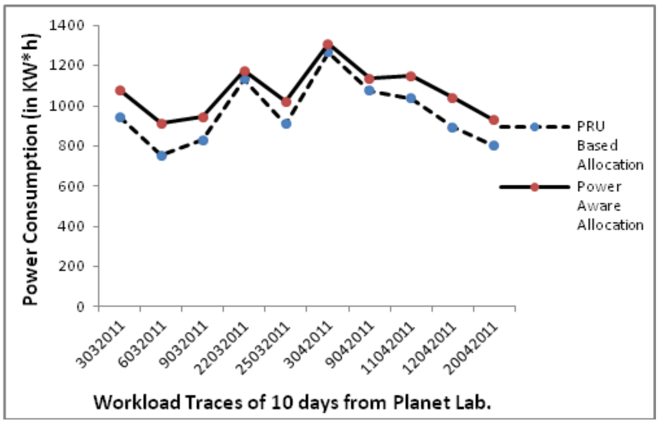

Figure 9: Power Consumption of data center with PRU based and Power Aware VM Allocation Policies

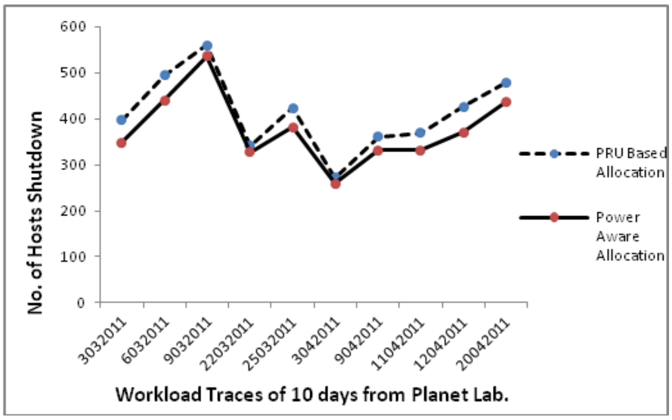

Figure 10: Migration with PRU based and Power Aware VM Allocation Policies

Figures 9&10 depict the experimental results obtained from implementing proposed PRU based VM Allocation policy and traditional Power Aware VM allocation policy available in CloudSim toolkit using 10 days' workload traces from PlanetLab. There is significant decrease in Power consumption of the data center in case of proposed policy than Power Aware policy. For each respective day, there is 12.54%, 17.23%, 12.21%, 3.51%, 10.97%, 3.10%, 5.21%, 9.53%, 14.13% and 13.76% decrease in power consumption in case of the PRU based policy as compared to the power aware policy. The graph in fig.10 shows that there is more no. of server shutdown in case of PRU based policy as compared to the other one and hence VMs in case of PRU based policy are more optimally packed in servers which results in minimum no. of active servers and reduced power consumption of the data center as compared to the Power Aware policy. For each respective day, there is 12.10%, 10.91%, 4.11%, 4.39%, 9.69%, 4.41%, 8.03%,10.03% 12.91% and 8.96% increase in server shutdown in case of PRU based policy than the power aware policy.

6. Conclusion

We proposed a resource efficient VM allocation policy to place VMs in the host machines in cloud data center. The policy is based on the concept of proportionate utilization of resources. According to the proposed policy, a VM is assigned to that host machine in the data center which leaves the remaining resources (RAM & CPU) less disproportionate than the other host machines in the data center. The CloudSim simulator based test bed has been used to create one data center with four host machines of heterogeneous type. The simulation started with request of placement of eight VMs of heterogeneous type to host machines in the data center. The experimental results obtained in case of the proposed policy are compared with simple first fit VM allocation policy and reveals that:

- Proportionate utilization of RAM and CPU capacity of Host machine has more chance to accommodate VMs and less chance of failures of allocation of VMs due to insufficiency of either of RAM or of CPU capacity.
- The results obtained in PRU based allocation are better than simple First Fit based allocation.
- There is 31.20%, 51.50%, 4.76%, 24.91% & 24.91% disproportionate usage of RAM and CPU capacity of Host0, Host1, Host2, Host3 & Host4 in case of simple First Fit based VM allocation policy where as it is 17.98%, 0.85%, 17.98% & 0.85% disproportionate usage of Host0, Host1, Host2 & Host3 in case of PRU based allocation, which are better than the first one.
- Furthermore, the addition of new host at runtime to meet the requirements of VM3, in case of simple First Fit VM allocation policy, adds to the total energy consumption of the data center.
- The energy consumption of data center is 18.46% more in case of simple First Fit policy than the proposed PRU based policy.

VMs are more closely and optimally packed in servers in case of PRU based policy which results in reduces power consumption of the data center.

References

[1] Vaquero, Luis M., Luis Rodero-Merino, Juan Caceres, and Maik Lindner:A break in the clouds: towards a cloud definition. *ACM SIGCOMM Computer Communication Review* 39, no. 1 (2008): 50-55.

[2] Chekuri, Chandra, and Sanjeev Khanna: On Multi-Dimensional Packing Problems. In *SODA*, vol. 99, pp. 185-194. 1999.

[3] Beloglazov, Anton, and Rajkumar Buyya: Adaptive threshold-based approach for energy-efficient consolidation of virtual machines in cloud data centers. In *MGC@ Middleware*, p. 4. 2010.

[4] Ahmad, Raja Wasim, Abdullah Gani, Siti Hafizah Ab Hamid, Muhammad Shiraz, Feng Xia, and Sajjad A. Madani: Virtual machine migration in cloud data centers: a review, taxonomy, and open research issues. *The Journal of Supercomputing* 71, no. 7 (2015): 2473-2515.

[5] Meng, Xiaoqiao, Canturk Isci, Jeffrey Kephart, Li Zhang, Eric Bouillet, and Dimitrios Pendarakis: Efficient resource provisioning in compute clouds via VM multiplexing. In *Proceedings of the 7th international conference on Autonomic computing*, pp. 11-20. ACM, 2010.

[6] Hyser, Chris, Bret McKee, Rob Gardner, and Brian J. Watson. Autonomic virtual machine placement in the

data center. *Hewlett Packard Laboratories, Tech. Rep. HPL-2007-189* 189 (2007).

[7] Bobroff, Norman, Andrzej Kochut, and Kirk Beaty: Dynamic placement of virtual machines for managing sla violations. In *Integrated Network Management, 2007. IM'07. 10th IFIP/IEEE International Symposium on*, pp. 119-128. IEEE, 2007.

[8] Jung, Gueyoung, Kaustubh R. Joshi, Matti A. Hiltunen, Richard D. Schlichting, and Calton Pu.: Performance and availability aware regeneration for cloud based multitier applications. In *Dependable Systems and Networks (DSN), 2010 IEEE/IFIP International Conference on*, pp. 497-506. IEEE, 2010.

[9] Jayasinghe, Deepal, Calton Pu, Tamar Eilam, Malgorzata Steinder, Ian Whally, and Ed Snible: Improving performance and availability of services hosted on iaas clouds with structural constraint-aware virtual machine placement. In *Services Computing (SCC), 2011 IEEE International Conference on*, pp. 72-79. IEEE, 2011.

[10] Nguyen Van, Hien, Frederic Dang Tran, and Jean-Marc Menaud: Autonomic virtual resource management for service hosting platforms. In *Proceedings of the 2009 ICSE Workshop on Software Engineering Challenges of Cloud Computing*, pp. 1-8. IEEE Computer Society, 2009.

[11] Calheiros, Rodrigo N., Rajiv Ranjan, Anton Beloglazov, César AF De Rose, and Rajkumar Buyya: CloudSim: a toolkit for modeling and simulation of cloud computing environments and evaluation of resource provisioning algorithms. *Software: Practice and experience* 41, no. 1 (2011): 23-50.

Author Profile

Minu Bala is working as an Assistant Professor in Computer Applications with Higher Education Department, Jammu and Kashmir Government, Jammu, India. She completed her MCA from University of Jammu in 1999 and is Post Graduate in Mathematics also. Presently, she is pursuing her Ph.D. from University of Jammu under Teacher Fellowship Programme, sanctioned by University Grants Commission, New Delhi. Her field of research is cloud computing.