

Comparative Analysis of Host Utilization Thresholds in Cloud Datacenters

Kritika Sharma

M.Tech Student,
Department of Computer Engineering,
Punjabi University, Patiala,
Punjab, India

Raman Maini

Professor,
Department of Computer Engineering,
Punjabi University, Patiala,
Punjab, India

ABSTRACT

Cloud Computing provides various services to its users just over an internet connection. For this the Cloud providers have to set up large datacenters which consists of number of computing nodes. These datacenters consumes a huge amount of energy. This not only leads to higher expenses but it also causes harmful effects on our environments. Thus, the attention is shifting towards energy efficiency in Clouds. This is done by optimizing VM allocation and migrating them to some other hosts. For VM migration various utilization thresholds are required. In this paper, we have done a comparative study of some of these utilization thresholds. The simulation is done on CloudSim toolkit. It has been concluded that for the same configurations of datacenter and for the same workload IQR utilization thresholds gives the best results for energy efficiency as well as lowest SLA violations.

Keywords

Cloud computing, Datacenters, Virtual Machines, VM migrations

1. INTRODUCTION

With the advancement in times, there had been many advancements taking place in the technology as well. One such advancement is the introduction of clouds in the IT sector. Clouds in general represent the collection of the computing nodes that are used for processing or storing the client's data. It also implies that the client need not bother where their data is saved. It may be at any location. The whole process of using clouds for performing various computing tasks is termed as Cloud Computing.

Cloud Computing is an emerging technology. It helps the users to use various services provided by the Cloud providers using the internet. Users or the clients need not install any of the softwares in their systems to use any of the services that the Cloud providers provide.

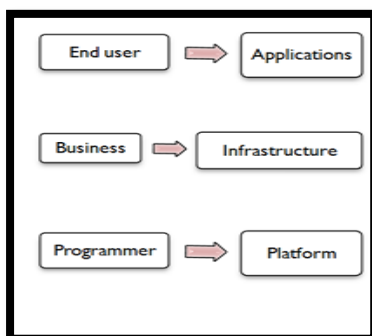


Fig 1. Different users and their requirements

The concept of Cloud Computing can be made clearer in the following scenario. Considering three kinds of users: a simple user with basic requirements like email, a developer and a business person. All these users have their own needs. A simple user may need many applications like an application to send emails, an application for video conferencing and many more. Similarly for a developer there is the need to download and install many softwares and also update them from time to time. On the other hand, if a business person wants to start a new website, then he may need to buy proper hardware and have to build a large infrastructure to keep these hardware. Thus, for all the three users there is a lot of expenditure to efforts, time and money as well.

But now in the Cloud Computing scenario all the different kinds of users are considered as clients and all their needs and requirements are fulfilled by the Cloud providers. The applications like email, video conferencing etc. are all maintained and provided by the Cloud providers. They develop these applications and provide them to their clients over the internet. Some of these applications are free while some of them are charged on the utility basis. Even for developers the Cloud providers provide the services. The providers maintain proper platform for developing the applications and provide them to the developers. The maintenance of these platforms is the sole responsibility of the provider. For business person the providers provide infrastructure facilities. They lend them physical server instances or just the virtual server instances on the pay-as-you-go basis. Thus, the Cloud technology helps to save a lot of time, effort and money of the clients.

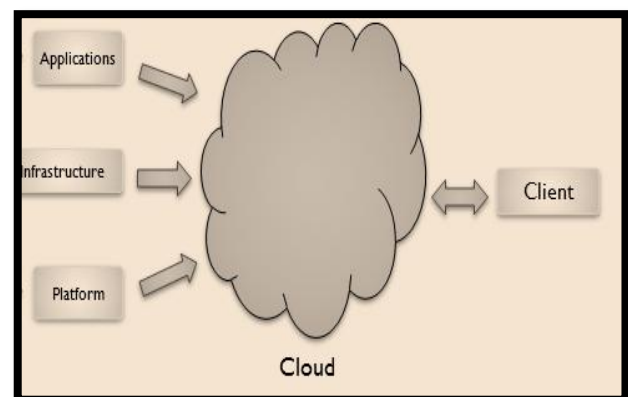


Fig 2. Cloud Concept

NIST[16] states the Cloud service models, Cloud characteristics and Cloud deployment models. Software as a service, platform as a service and infrastructure or hardware as a service are the three service models of Clouds. Here,

Software as a service provides software applications as service, *platform as a service* provides the development environment in the form of service, and *infrastructure as a service* provides hardware as the service to the customers.

The characteristics that are essential for the Cloud Computing environment are: service on user demand, ubiquitous network access, independence of location, scalable and infinite resources. The establishment of Clouds is categorized into four models of deployment. The Clouds environment can be public, private, community and hybrid.

Cloud Computing offers a number of benefits to the IT sector, since it helps to relieve the companies from unnecessary jobs like setting the basic hardware infrastructure or maintaining the softwares. But establishment of the Cloud environment involves setting up of large number of computing nodes. These nodes consume the enormous amount of energy. Thus, the overall power consumption of datacenters is gets too high. So far the focus of Cloud Computing has been on increasing the performance of the Cloud environment. But there is a need to shift the attention towards reducing the power consumption of the datacenters.

But the high number of computing nodes cannot be said solely responsible for the huge power consumption of datacenters, but it is also because the computing nodes are not properly utilized. When a node is not properly utilised, it is found to consume even more power. So the focus should be shifted towards optimizing the nodes to consume lowest amount of power possible and also maintaining the SLAs.

Reduction in power consumption in datacenters can be done by using anyone of the following[12]:

- Creating virtual machines.
- Switch off the idle hosts or use lower power consumption modes.
- Using live migration of VMs.

For energy efficiency of the Cloud datacenters there can be basically four main areas that need to be optimized[4]. These are as follows:

- Detection of whether the host is overloaded or not.
- Detection of whether the host is underloaded or not.
- Selecting the VM to be migrated from the underloaded or the overloaded host.
- Selecting the new host for the migrating VM.

In this paper we will be focussing on the first part that is detection of whether the host is overutilized or not. In section II we discuss the related work. Section III gives the details about the utilization thresholds that we are studying. Section IV gives an overview of the CloudSim simulator. Section V explains the simulation and simulation environment created. In section VI and VII we discuss the results and conclusions.

2. RELATED WORK

In [1], the comparative study of various energy efficiency approaches in Cloud Computing like DVFS is done. The study has shown that these approaches help to reduce power consumption and also help to maximize the performance of the Cloud environment.

[2], this paper focusses on the virtual machine migration policies. It also discusses the fault tolerance in Clouds and

services that are offered by the Clouds. The overview of various VM migration policies is presented.

In [4], the authors have explained the Cloud model. They have conducted the competitive analysis and found the competitive ratios of algos for the dynamic VM migrations as well as consolidations. They have also proposed various heuristics for dynamic consolidation of VMs based on the historical data.

[13], here the authors have proposed an energy model to verify the performance of the simulation. A new algorithm Best Fit Decreasing Power is also introduced. This proposed algorithm has been found to give the best results for the light workloads.

In [20], the authors present the review of the various algos that are used for balancing the load among the different computing nodes. The paper gives the corresponding advantages and disadvantages of these algorithms and also studies the various performance metrics related with them.

3. OVERUTILIZED HOST DETECTION

The optimization of datacenters to reduce power consumption involves four basic steps. Firstly the over-utilized and under-utilized hosts are detected. Then, the VMs are selected to be migrated from the over-utilized as well as under-utilized hosts. And finally, the host to which the VMs have to be migrated is selected. Here, the main focus is on finding whether a host is over loaded or not. For this [4] has defined some thresholds which are used to decide whether a host is overloaded. These thresholds are discussed below.

The threshold parameters given below are used to determine the host utilization threshold to decide whether the host is over loaded or not. If the host utilization value is higher than the utilization threshold then that host is considered overloaded and some of its VMs are migrated to some other host.

- MAD: MAD is the short form of Median absolute deviation. MAD is a robust measure. It is better than standard deviation because it does not give errors due to the outliers. MAD is a statistical measure and it is used to calculate dispersion. Let us consider a univariate data set, $Y_1, Y_2, Y_3, \dots, Y_n$, the MAD of such data is defined as the median of all the deviations from the median of the data, and these deviations are the absolute deviations.[4]

$$MAD = \text{median}_k(|Y_k - \text{median}_l(Y_l)|) \quad (1)$$

The value of Utilization Threshold is calculated as"

$$\text{Utilization Threshold} = 1 - s \cdot MAD \quad (2)$$

Here, s is the safety parameter. It is some positive real number. The lower value of s implies that the energy consumption is lower but the value of SLA violations is higher. The SLA violations may be caused due to dynamic consolidation of the virtual machines [4].

- IQR: IQR is the short form for Inter Quartile Range. It is also a measure of statistical dispersion. It also has names like midspread and middle fifty. IQR is equal to the difference between quartiles. These quartiles are first and third.

$$IQR = Q_3 - Q_1 \quad (3)$$

If the data is symmetrically distributed then half value of IQR is same as the MAD value for that data.

The value of Utilization Threshold is calculated as”

$$Utilization\ Threshold = 1 - s \cdot IQR \quad (4)$$

Here, s is the safety parameter.

- THR: It is the static threshold policy. In this policy the static threshold value is used to generate the utilization thresholds. Then this threshold value is used to estimate whether the host is over utilized or not.

The value of Utilization Threshold is calculated as:

$$Utilization\ Threshold = 1 - s \cdot THR \quad (5)$$

Here, s is the safety parameter.

4. CLOUDSIM SIMULATOR

To test which of the above stated utilization threshold gives best results for the specified workload, we use the CloudSim Simulator. We use the simulator because all the new policies or algorithms need to be tested before we are actually implementing them in the Cloud environment. But creating a real world Cloud environment to test the newly suggested policies and algorithms gets too costly in terms of money. And it wastes a lot of time as well. Thus we simulate the Cloud environment and run our policies on the simulations. In this way the efficiency of the policies can be determined at a very low cost of time and money.

CloudSim is a simulation toolkit which has been written in Java. CloudSim involves four layers:

The lowermost SimJava layer is responsible for all the core activities. The GridSim layer helps to model the grid infrastructure. The CloudSim layer is responsible for the actual implementation of the CloudSim. And the topmost UserCode layer contains information like no. of hosts, their configurations, no. of users, etc.

The CloudSim simulator consists of various inbuilt classes that are used to simulate a Cloud environment. Some of the inbuilt classes of CloudSim are:

- Datacenter
- DatacenterBroker
- Vm
- PowerModel
- CloudletScheduler
- VmScheduler,

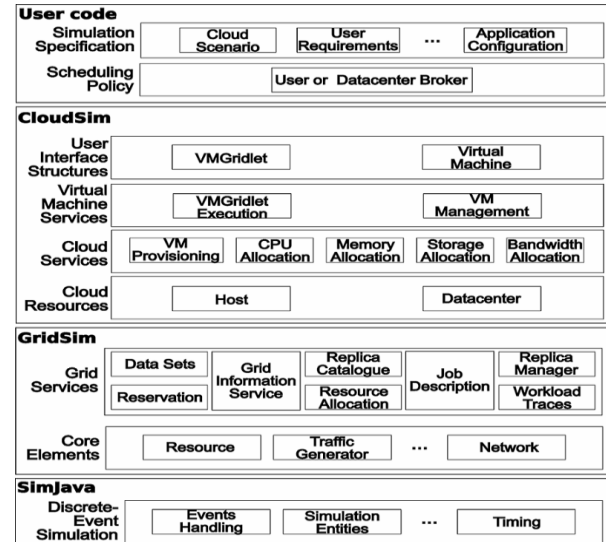


Fig. 3 Layered CloudSim architecture[18]

and many more. All these inbuilt classes in CloudSim package have their own predefined functionalities. Like, the Datacenter class is used to build a datacentre for a list of virtual machines on a list of hosts. CloudSim also provides a method `getPower()` which is used to generate the power consumption for a given time period.

Quantitative parameters comparing utilization thresholds:

- *Energy consumption:* Energy consumption of the datacenter is calculated using the linear interpolation function.

$$Energy\ Consumption = previousPower + (presentPower - previousPower) / 2) * time \quad (6)$$

Here,

previousPower : power consumed due to previous CPU utilization

presentPower : power consumed due to present CPU utilization

time : is the time difference between the current time and last process time

- *SLA Violations:* SLA violations are the violations in the Service Level Agreement. If the requirements of the customers are not met, then we get SLA violations.

$$SLA\ violations = (SLA\ per\ active\ host + SLA\ degradation\ due\ to\ migration) * 100 \quad (7)$$

Here,

SLA per active host: is the sum of all the violations for all the hosts when allocated mips is less than requested mips.

SLA degradation due to migration: is the sum of all the violations when VM cannot be allocated to a host because of VM migration

5. SIMULATIONS

The simulations are done in the CloudSim version 3.0.1. We have created 1000 hosts. All the hosts are considered homogeneous. The host storage is 1 GB. The bandwidth of 1 Gbit/s. For 1000 hosts we have 1052 VMs. Each VM’s bandwidth is 100 Mbit/s and its size is 2.5GB. each VM requires only one processing core. When the simulation is carried out with the same workload for all the three utilization

thresholds, we get the following results.

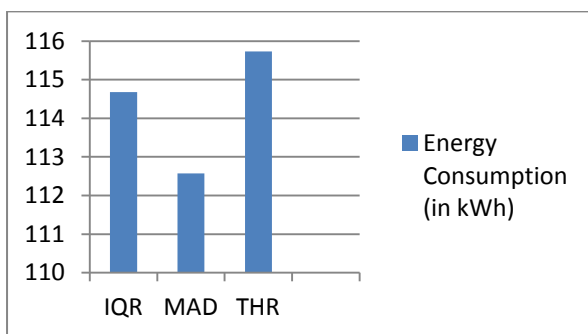
Table. 1

	IQR	MAD	THR
No. of hosts	1000	1000	1000
No. of VMs	1052	1052	1052
Energy consumption (in kWh)	114.68	112.57	115.73
Sla violations (in %)	0.00346	0.00359	0.00352
No. of VM migrations	16332	16433	16181

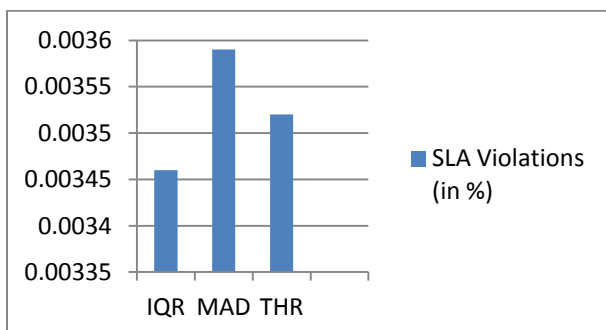
6. RESULTS AND DISCUSSION

It can be seen in Table 1 that for same number of hosts and VMs and with the same configuration of datacenter, the lowest energy consumption value is for MAD which is 112.57 kWh. IQR gives energy consumption of 114.68 kWh while THR gives the highest energy consumption of 115.73 kWh. Graph I shows energy consumption for all the three thresholds.

For SLA violations the lowest value is given by IQR. IQR gives 0.00346 % SLA violations. MAD gives 0.00359 % SLA violations. THR gives the highest number of SLA violations i.e. 0.00352 %. The graph for SLA violations is shown below (Graph II).



Graph I.



Graph II.

7. CONCLUSION

The simulations results show that MAD gives the best results for energy efficiency but it gives high value for SLA violations. On the other hand, THR gives the highest energy consumption as well as very high percentage of SLA

violations. But IQR gives better results for energy efficiency than THR and gives lowest number of SLA violations. Thus, IQR utilization threshold not only helps to reduce power consumption of the datacenter but it also leads to high system performance.

8. REFERENCES

- [1] Aditi Mahajan, Anita Ganpati, A Study of Energy Efficiency Techniques in Cloud Computing. IJCSMC, Vol. 3, Issue. 8, August 2014, pg.707 – 712
- [2] Amritpal Singh, Supriya Kinger, Virtual Machine Migration Policies in Cloud, International Journal of Science and Research (IJSR), India Online ISSN: 2319-7064 Volume 2 Issue 5, May 2013
- [3] Anton Beloglazov, Jemal Abawajy, Rajkumar Buyya, Energy-aware resource allocation heuristics for efficient management of data centers for Cloud computing, Future Generation Computer Systems 28 (5) (2012) 755–768.
- [4] Anton Beloglazov_ and Rajkumar Buyya, Optimal Online Deterministic Algorithms and Adaptive Heuristics for Energy and Performance Efficient Dynamic Consolidation of Virtual Machines in Cloud Data Centers , CONCURRENCY AND COMPUTATION: PRACTICE AND EXPERIENCE Concurrency Computat.: Pract. Exper. 2012; 24:1397–1420
- [5] Bhupendra Panchal, Prof. R. K. Kapoor, Dynamic VM Allocation Algorithm using Clustering in Cloud Computing, International Journal of Advanced Research in Computer Science and Software Engineering Volume 3, Issue 9, September 2013 ISSN: 2277 128X
- [6] Buyya R, Yeo CS, Venugopal S, Broberg J, Brandic I. “Cloud computing and emerging IT platforms: Vision, hype, and reality for delivering computing as the 5th utility.” Future Generation Computer Systems 2009; 25(6):599–616
- [7] cloudsim 3.0 API [Online]. Available: <http://www.cloudbus.org/cloudsim/doc/api/overview-summary.html>
- [8] Cyril Onwubiko, “Security Issues to Cloud Computing,” in Cloud Computing: Principles, Systems and Applications, Computer Communications and Networks, DOI 10.1007/978-1-84996-241-4_16, © Springer-Verlag London Limited 2010
- [9] Hetal Joshiyara, Energy Efficient Provisioning of Virtual Machines in Cloud System using DVFS, INDIAN JOURNAL OF RESEARCH Volume : 3 | Issue : 4 | May 2013
- [10] Introduction to Cloud Computing Office of Privacy Commissioner of Canada, 1989.
- [11] Kritika Sharma, Raman Maini, Cloud Computing and the Cloud Simulation, International Journal of Engineering Research & Technology (IJERT) ISSN: 2278-0181 Vol. 3 Issue 5, May – 2014.
- [12] Kritika Sharma, Raman Maini, Comparative Analysis of Energy Efficiency in a datacenter using DVFS with a Non-Power Aware Datacenter (IJERT) ISSN: 2278-0181 Vol. 4 Issue 02, February-2015
- [13] Liang Luo, Wenjun Wua, W.T. Tsai, Dichen Di, Fei Zhang, Simulation of power consumption of cloud data centers, Simulation Modelling Practice and Theory 39

(2013) 152–171

- [14] Nguyen Quang Hung, Nam Thoai, Nguyen Thanh Son, PERFORMANCE CONSTRAINT AND POWER-AWARE ALLOCATION FOR USER REQUESTS IN VIRTUAL COMPUTING LAB,
- [15] Nikzad Babaii Rizvandi, Javid Taheri, Albert Y. Zomaya, Young Choon Lee, Linear combinations of DVFS-enabled processor frequencies to modify the energy-aware scheduling algorithms, in: Proc. of the 2010 10th IEEE/ACM International Conference on Cluster, Cloud and Grid Computing, 2010, pp. 388–397.
- [16] Peter Mell, Timothy Grance, “The NIST Definition of Cloud Computing (Draft),” Computer Security Division, Information Technology Laboratory, National Institute of Standards and Technology, Gaithersburg, January 2011.
- [17] Rajkumar Buyya, Anton Beloglazov, and Jemal Abawajy, Energy-Efficient Management of Data Center Resources for Cloud Computing: A Vision, Architectural Elements, and Open Challenges, 2010, CSREA Press
- [18] Rodrigo N. Calheiros, Rajiv Ranjan, Anton Beloglazov, Cesar A. F. De Rose and Rajkumar Buyya “CloudSim: a toolkit for modeling and simulation of cloud computing environments and evaluation of resource provisioning algorithms”, CLOUDS Laboratory, Department of Computer Science and Software Engineering, The University of Melbourne, Australia , DOI: 10.1002/spe.995.
- [19] Saiqin Long, Yuelong Zhao, Wei Chen School, A three-phase energy-saving strategy for cloud storage systems, The Journal of Systems and Software 87 (2014) 38– 47
- [20] Sukhvir Kaur, Supriya Kinger, Review on Load Balancing Techniques in Cloud Computing Environment. IJSR, ISSN : 2319-7064
- [21] Tom Guérout, Thierry Monteil, Georges Da Costa, Rodrigo Neves Calheiros, Rajkumar Buyya, Mihai Alexandru, Energy-aware simulation with DVFS, Simulation Modelling Practice and Theory 39 (2013) 76–9