Journal of Science and Technology
ISSN: 2456-5660 Volume 7, Issue 01 (JAN-FEB 2022)

www.jst.org.in *DOI: https://doi.org/10.46243/jst.2022.v7.i01.pp82-91*

Energy-Aware VMs Consolidation Computing Frameworks' of Data Center in Cloud Computing Environment

Dr. Rajesh P. Patel [1], Dr. Harshad B. Bhadka[2]

[1]Department of Computer Engineering, Gujarat Power Engineering and Research Institute,
Gujarat Technological University, Gujarat, India
[2]Faculty of computer science, C.U.Shah University,
Wadhwan, Gujarat, India
[1]rpp.kirc@gmail.com, [2]harshad.bhadka@yahoo.com

To Cite this Article
Dr. Rajesh P. Patel and Dr. Harshad Bhadka, "Energy-Aware VMs Consolidation Computing Frameworks' of Data Center in Cloud Computing Environment", Journal of Science and Technology, Vol. 07,Issue 01, January-February 2022.

Article Info
Received: 02-02-2022 Revised:06-02-2022 Accepted:11-02-2022 Published: 26-02-2022

Abstract: *Cloud computing is a service model that can conveniently access a shared pool of configurable computing resources that can be quickly configured and released on demand. In cloud data centers, the scale and complexity of various computing resources such as servers, network equipment, and cooling systems are constantly evolving, which consumes a lot of power and increases the energy consumption of the data center. Because cloud data center resources are not optimized for maximum utilization, they consume more power. Therefore, it is necessary to integrate virtual machines (VMs) on data center servers to help optimize the use of resources in the cloud, thereby reducing energy consumption. By considering the optimal power consumption of various data center resources, many researchers have proposed various methods and algorithms to reduce the power consumption of servers and network equipment. In this paper, we introduced two energy-saving computing frameworks (1) data center energy-saving server power model, (2) energy-saving VM migration based on Multi-objective to help optimize data center power consumption.*

Keywords**:** *Server consolidation,VM Migration, Quality of Service, virtualized data center, Service Level Agreements, Highest Thermostat Setting*

I. Introduction

Cloud computing is an architecture that provides on-demand computing services through the Internet and accesses a set of shared resources (i.e, networks, storage, servers, services, and applications) at a user fee without physically acquiring them[1]. This type of computing provides many benefits for businesses, shorter start-up time for new services, lower maintenance and operation costs, higher utilization through virtualization, and easier disaster recovery that make cloud computing an attractive option[2]. This technological trend has enabled the realization of a new computing model, in which resources (e.g., CPU and storage) are provided as general utilities that can be leased and released by users through the Internet on-demand fashion [3]. Multi-Tenancy in Cloud Computing occurs when multiple consumers share the same application, running on the same operating system, on the same hardware, with the same data-storage system and both the attacker and the sufferer are sharing the common server [4].

II. Related Work

Nazneen Taj et.al [5] proposed work on energy calculation in a data center with the Losses and the Method calculating the energy losses that occur during scheduling also the losses caused by the components used. They have proposed equations to calculate the total energy consumed. They have discussed how energy supplied at different times or different intervals of time can be saved where some measures have been proposed to minimize the losses. Soumya Ranjan Jena et.al [6] proposed four different power models such as linear model, cubic model, square model, and square root model on an Infrastructure-as-a-Service (IaaS) cloud environment to find out the best one. They found overall CPU utilization, overall power consumption in each case by migration, the relationship between CPU utilization and overall power consumption in each model. It has found the best model out of these four in a predetermined time period based on total power consumption and at the end calculates the accuracy through R-squared and mean square error. They found that the cubic polynomial model is the most efficient one and consumes less power.

Corradi A et al. [7] focus on VM integration from a practical point of view and accurately consider integration features related to power, processor, and network resources. With work-focused networking, host resources, and power consumption, it provides a cloud-based orchestration platform to streamline VM integration. The annotated experimental results indicate that MV interventions should be carefully measured to avoid placing the results. You cannot guarantee the provision of virtual machines with service level guarantees. The virtualization method is to provision virtual machines on the same physical host to use available resources. Cloud providers such as IBM, Google, and Amazon have started the development and implementation of many cloud computing solutions to minimize the use of data centers.

Zhang F *et al.* [8] presented a multi-objective scheduling (MOS) scheme that was particularly customized for cloud computing which depends on ordinal optimization which was at first urbanized by the automation society for the design optimization of very compound dynamic systems. They showed the sub-optimality through mathematical analysis. The work was extended for cloud platforms that affect virtual clusters of servers from multiple data centers. The main advantage of MOS method is reduced scheduling overhead time and provides optimal performance. The Wide-ranging tests were performed on the clusters of 16 to 128 virtual machines. LIGO multitasking workload which is of real scientific workload was obtained for earth gravitational wave study. The results of experiments showed that the algorithm quickly and successfully produced a little set of semi-optimal development results. In distributed computing environment the scheduling algorithm of multiple tasks was a famous NP-hard problem.

A new scheduling approach named Pre Ant Policy was introduced by Hancong Duan et.al [9] based on fractal mathematics their method consists of a forecast model and is based on an enhanced ant colony algorithm (ABC) a scheduler. To activate the implementation of the scheduler by virtue of load trend prediction was determined by prediction model and under the premise of guaranteeing the quality-of-service, for resource scheduling, the scheduler is responsible while maintaining energy consumption. The performance results demonstrate that their approach presents resource utilization and excellent energy efficiency. In sequence to elevate the exchange between energy consumption and application performance Rossi et.al [10] presented automated coordination of different energy savings techniques. They implemented Energy-Efficient Cloud Orchestrator-e-eco and by using enlarging applications on a dynamic cloud in a real environment the infrastructure test was carried out to evaluate e-eco. Their evaluation result demonstrates that e-eco was able to reduce energy consumption.

In cloud computing for energy saving, a three-dimensional virtual resource scheduling method (TVRSM) was introduced by Zhu et.al [11]. For the cloud data center, they build the resource and dynamic power model of the PM in their work. There are three stages of the virtual resource scheduling process as follows; virtual resource optimization, virtual resource scheduling, and virtual resource allocation. For the diverse aim of each stage, they design three different algorithms respectively. The TVRSM can successfully diminish the energy consumption of the cloud data center when compared with various traditional methods.

III. Proposed Frameworks

3.1 Energy-Aware Manipulated Framework

The framework proposed for the energy effective cloud consists of a power-efficient data center where servers are a significant component of cloud data centers in the cloud.

(a) Power Calculator

The proposed power calculator component is used to show the power consumption of different components in the data center. The power calculator provides detailed information on the various physical and dynamic elements included in the proposed framework. From these elements, the energy calculator estimates the energy consumption due to data storage and uploading in the cloud, as well as in the physical elements, and similarly, the energy consumption through dynamic elements during the cloud request processing.

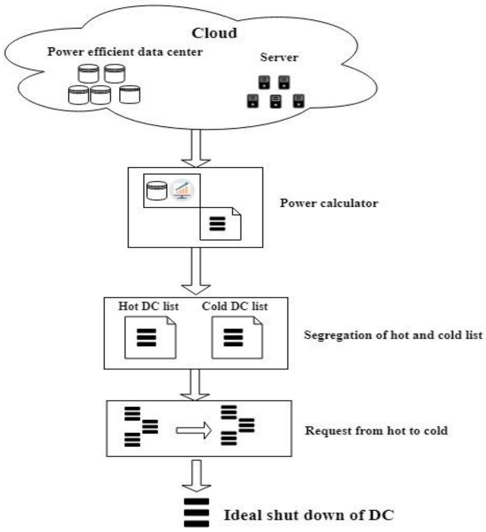

Fig 1: Proposed energy-aware manipulated framework

(b) Hot and Cold DC List Generation

The outcome from the power calculator classifies the existing data centers in the cloud into hot data centers and cold data centers. The hot data centers are those that are heavily loaded and consume more energy in the cloud. Similarly, the cold ones are slightly or zero-loaded data centers. This segregation results in the grouping of hot and cold data centers effectively.

(c) Flow of request from Hot DC to Cold DC

Based on the list of hot and cold DC, the load from the hot DC is transferred into the cold DC. This transfer process helps to reduce the energy consumed by hot DC and also minimizes to a greater extent the number of requests to be processed. This may result in the formation of idle DC.

(d) Idle Shut Down

The final stage of the proposed framework is to shut down the DC that is observed idle after the request transfer through the power saving scheme.

Consumed network power PW_{NET} is stated as:

$$PW_{NET} = PW_{TOR} + PW_{AGR} + PW_{CR} \quad (1)$$

P_{ToR} PW_{TOR}, PW_{AGR} and PW_{CR} represent power consumption at rack switches at top, switches at aggregation-layer, and core router correspondingly. For total consumption of cooling power PW_C when turning on the j^{th} rack, the related cooling rack can be switched on and consumes PWj power. Therefore, PW_C is defined as

$$PW_C = \sum_{j=1}^{Nr} PWj \quad (2)$$

Where N_r determines the total number of racks. Finally, the overall DC power consumption PW_{DC} is given as:

$$PW_{DCi} = PW_s + PW_c + PW_{NET} \quad (3)$$

$$PWs = \sum_{i=1}^{Ns} PWi \quad (4)$$

Where PW_s represents total power consumption of servers and N_S specifies the total number of servers in all racks. The power consumption of each DC has been measured. After finding the individual DC power consumption, DC are distinguished on Minimum and Maximum consumption of power.

A. Maximum Power Consumption of DC

Maximum power consumption DCs are separated if the Power of that DC is greater than the limit.

$$\max PW_{DCi} = PW_{DCi} > L \tag{5}$$

B. Minimum Power Consumption of DC

Minimum DCs power consumption is detached when the DC Power is lower than the limit.

$$\min PW_{DCi} = PW_{DCi} < L \tag{6}$$

The limit (L) employed here for evaluating maximum and minimum power consumption of DC is calculated by Normal power consumption of DC NPW_{DC} multiplied by the power required for each process $PW_{process}$ divided by a total process ($N_{process}$)

$$L = \frac{NPW_{DC} \times PW_{process}}{N_{\text{Process}}} \tag{7}$$

(e) Dynamic Components of DC

The Total Number of Requests (R) is collected from GIS (Global Information System)

$$R = GIS \tag{8}$$

The utilized capability of the DC_i is calculated using eq.

$$UC_{DCi} = \frac{OC}{ET} \tag{9}$$

OC- Original Capacity, ET- Execution Time

For all the jobs in the queue *WT* from *total Request R*. The difference of estimated execution time EET and the completed execution time CET is calculated using eq. (10)

$$WT = EET - CET \tag{10}$$

The job execution time *ET* determines the speed of performing a job *T*. It can be calculated by finding the difference between the current time and entry time of the job into VM using eq. (11)

$$ET = Current\ time - Entry\ time\ of\ job\ into\ VM \tag{11}$$

The job completion time is the total time for both waiting and execution.

$$CT = WT + ET \tag{12}$$

(f) Estimated Completion Time

The estimated completion time of an un-started job is calculated by counting the number of floating-point operations (supplied by the project) and CPU floating-point benchmark.

$$ECT = FA + (1 - F)B \tag{13}$$

Where F is the generated fraction, A is the CPU elapsed time and fraction generated estimate, and B is the floating-point count and benchmarks estimate.

(g) Best DC

Best DCs are DC with less computation time. If the completion time of an allocated process is lesser or equal to the estimated completion time then it is chosen as best DC

$$BDC_i = CT <= ECT \tag{14}$$

(h) UnBest DC

Unbest DCs are the DC with high computation time or crashed DCs. If the completion time of an allocated process is more elevated than the estimated completion time then it is chosen as Unbest DC

$$UDC_i = CT > ECT \tag{15}$$

(i) Hot DC

It is evaluated by utilizing a comparative performance server present in the list of both max power consumption and time utilization. If a server presents in both max power and unbest DC then it is concluded as hot DC

$$H_{DCi} = \max PW_{DCi} \cap UDC_i \tag{16}$$

(j) Cold DC

If a server is present in both min power and best DC then it is concluded as cold DC (C_{DC_i})

$$C_{DCi} = \min PW_{DCi} \cap BDC_i \tag{17}$$

After listing the hot and cold Datacenters separately further new requests from users are moved from hot DC to cold DC.

3.2 Proposed Energy-Efficient Multi-Objective Based Energy-Efficient Vm Migration Framework

3.2.1 Dolphin Echolocation Optimization Algorithm

It is recently developed for the discrete search space with the advantages of echolocation to discover their environment. The problem of finding some variables' value in search space is like a search of dolphins in their environment. In optimization strategy choosing the best answer for a problem is similar to a dolphin's attempt to find the best target. Dolphins at the outset, look around the search space to find out where the preys are, subsequently they restrict the trace to locate the precise position. The method simulates dolphin echolocation by decreasing the size of the random search space proportional to the distance to the target.

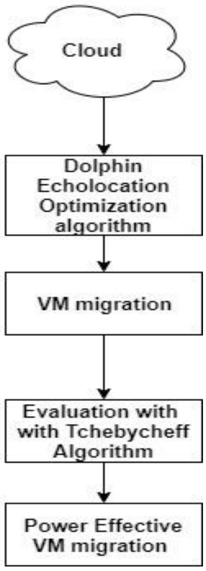

Fig 2: Proposed multi-objective energy-efficient framework

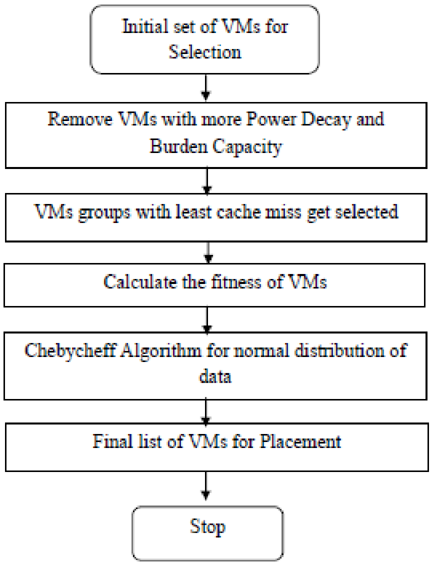

Fig 3: Proposed work Flowchart

Steps in flowchart fig 3 of the proposed work are given below.

First Level: Remove those VM with more Power decay (a threshold would be set) and burden capacity less than the threshold would be selected and asked to move to the next level of selection.

Second Level: VM's would be selected with CPU utilization (in terms of a cache miss) the VM's groups with the least cache miss get chosen.

Third level: DEOA calculates fitness function which can be calculated based on the total number of jobs for each VM, cache size, and memory utilization with their respective parameter. DEOA has separated the best sets of VMs.

Fourth Level: To still find an optimal {ideal sets}, we go for an algorithm called Tchebycheff algorithm, In that data should follow (at least approximately) a normal distribution (we assume that the fitness follows normal distribution), and if it does, find the mean and use it as a threshold. For migration, a VM with a fitness value greater than the threshold will be selected. Algorithm use Burden capacity, Power Decay, least cache miss, Migration time, and temperature parameter to find out the fitness for migration of VMs.

(b) Burden Capacity

The BC of a server or virtual machine is the formation of its CPU, network, and memory capacity like

$$BC = \frac{1}{1-U_c} \cdot \frac{1}{1-U_m} \cdot \frac{1}{1-U_n} \tag{18}$$

Consider
U_C is the utilization of CPU
U_m is the utilization of memory
U_n is network usage.

The functional PD for a server node encompasses static energy decay E_s, which represents the static power decay P_s of all system components apart from the CPU in fraction to the all set time, and active energy decay E_d of the running applications, which rely mostly on the dynamic power P_d is

$$PD = E_s + E_d \tag{19}$$

The dynamic power P_d of the CPU is generally relative to the CPU frequency f_p given by

$$P_d = k.f_p{}^3 \tag{20}$$

Where k is proportionality constant. Consequently, we deliberate an application with execution time t, and the CPU runs at frequency f_p such that $0 < f_p < f_m$ the execution time is defined by $t/\left(f_p / f_m\right)$. Thus the dynamic energy decay E_d for this application is given by

$$E_d = \int_0^{t/(f_p/f_m)} P_d = k.t.f_m.f_p{}^2 = \alpha.t.S^2 \tag{21}$$

Where α is the proportionality constant, and S is the related CPU speed connected to the f_p $(S = f_p / f_m)$ in equation (21). The energy decay for the system is affected unwaveringly by the BC. Nevertheless, still, the BC of two systems is equivalent; the power decay can contrast under the system ability for every node

(c) Temperature

The temperature for a data center server is relative to the functional power of the system, and the surrounding heat T_{amb} is given by

$$T = P.R + T_{amb} \tag{22}$$

Where R is resistance.

(d) Resource Wastage

The massive amount of wastage of resources in Cloud datacenters results in resource management problems. The challenges related to datacenters with a particular emphasis on how new virtualization technologies can be used to simplify deployment, improve resource efficiency, and reduce the usage of physical servers. The resource wastage can be expressed as,

$$RW = \sum_{i \neq k} (NR_o - NR_s) \tag{23}$$

Where, NRs, recognize the dimensions which have the least normalized remaining ability and 'o' for other dimensions. For instance, for three goals CPU, memory, and network RW can be estimated as,

$$RW = (U_m - U_c) + (U_m - U_n) \tag{24}$$

Where U_m stands for NR_s in this condition.

(e) Rate of Migration

Migration Rate (MR) might contrast suggestively for various workloads because of the differences in VM arrangements and workload qualities. Briefly, the enactment of VM relocation is influenced by numerous components, essentially the amount of the Virtual machine memory, the network communication cost, and energy usage because of relocation, mainly in a significant system. Live immigration has more extra components. Migration cost measurements for VM relocation are incorporated in the accompanying equation:

$$MR = iV_m + jT_m + kE_m + [lT_d] \tag{25}$$

Where $i + j + k + l = 1$ i, j, k, and l are mass of rate metrics, T_m is the period of relocation, V_m is the overall network traffic of the relocation method, E_T is overall energy expended by the relocation method and T_d is the downtime causes in the relocation method. Only for the offline relocation, the total network interchange V_m is the memory amount of the relocated virtual machine V_{mem}. The migration duration T_m is planned by

$$T_m = \frac{V_{mem}}{TR_m} \tag{26}$$

Where TR_m is the communication speed of memory. The total energy E_T will be evaluated by

$$E_T = E_S + E_D + [E_N] \tag{27}$$

Where, E_S, E_D, and E_N declare to the extra energy devoted by the resource, destination, and network switches, individually.

IV. Experimental Setup And Results

It describes the experimental results of the proposed technique executed in a PC with 8 GB physical memory, 64-bit Windows 10 operating system, and Intel(R) Core(TM) i5 processor and for implementation, the IDE tool and Cloud tool is used. Utilizing Netbeans IDE 8.2, CloudSim toolkit, the cloud setup with various configurations is simulated.

The power calculator is used to find the power consumption of DC. The Power Consumption of Data centers is also revealed in fig 4 in which data center 3 and 4 consume more power compared to other data centers. Virtual machines are created and assigned to data center host machines. Based on the power calculation of virtual machines the virtual machines are decided Busy or Idle status, which helps for migration which is shown in fig 5.

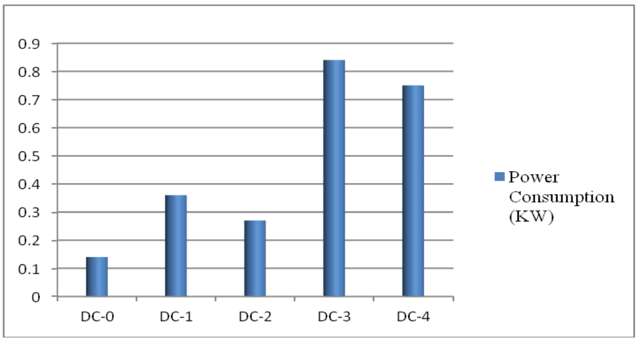

Fig 4: Power consumption of data centers (DCs)

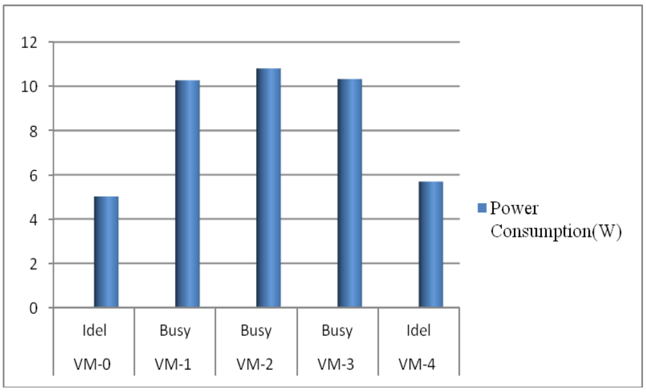

Fig 5: VMs Power consumption and status

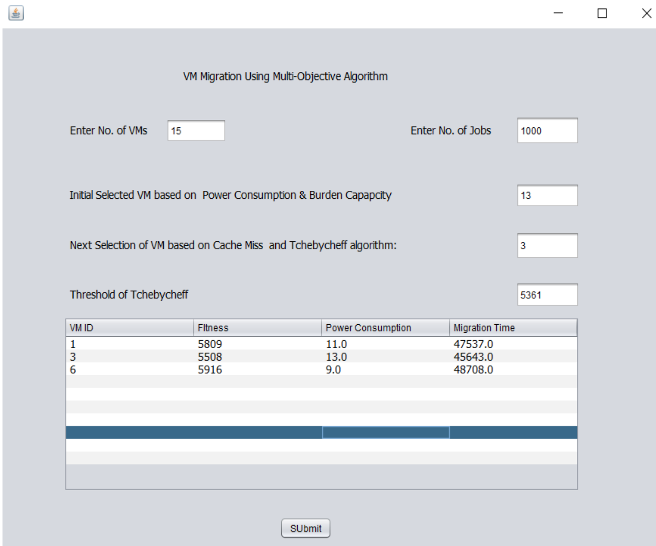

Fig 6: Modified algorithm simulations

The experiment mentioned above, fig.6 was carried out with 15 VMs that have been assigned 1000 jobs. The proposed work initially selects VM for migration which is based on power consumption and burden capacity. The next selection of VM is based on the Minimum cache miss and Tchebychef threshold which chooses the ideal VMs for migration.

In the above fig.7 Tchebycheff threshold is used to identify the final VMs for the migration. The VMs having fitness greater than the threshold will be selected for final migration. In the simulation, we have got a threshold value of 5361. So the Virtual machines (VMs) Id – 01, 03, 06 are finally selected for migration.

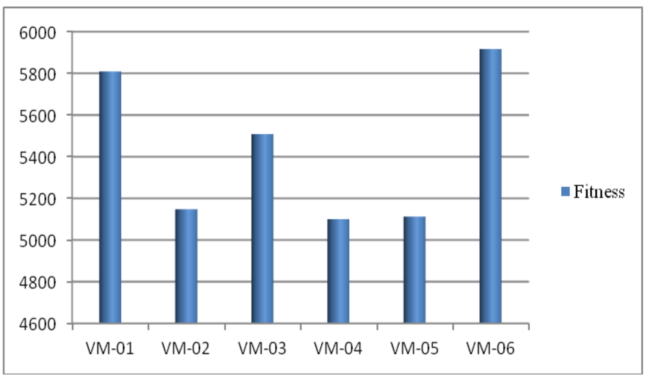

Fig 7: Fitness of VMs

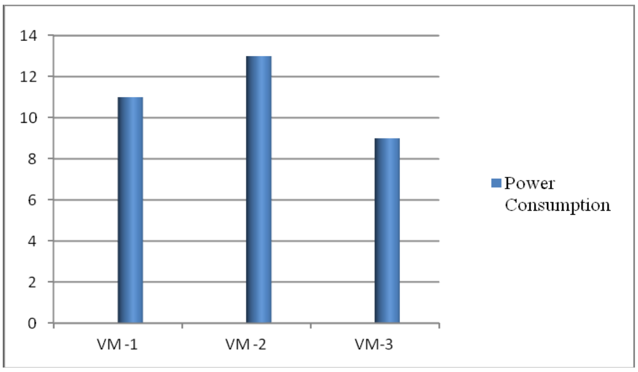

Fig 8: Power consumption of VMs selected for migration

V. Conclusion

The power calculation method recognizes the Best and UnBest data center using their computational time which is also used to differentiate Hot and Cold data centers of the cloud using the power calculation method. User requests placement techniques to find the best location of each request on the servers based on the typical data center configuration. A list of Virtual machines is identified for the migration which consumes more powers is calculated by the power model. This model helps to recognize VMs for migration selection on other servers for load balancing. The proposed framework performs Energy-efficient VM migration in a cloud data center using dolphin echolocation optimization with the Tchebycheff algorithm which optimally selects VMs which can be migrated to a server with maximal workloads, at the same time reducing the usage of resources in a data center.

VI. References

[1] Vivek Raich, Pradeep Sharma, Shivlal Mewada and Makhan Kumbhka "Performance Improvement of Software as a Service and Platform as a Service in Cloud Computing Solution", International Journal of Scientific Research in Computer Science and Engineering,Vol.**1**, Issue-6, pp.13-16, Dec 2013.

[2] Karthik Kumar and Yung-Hsiang Lu, "Cloud Computing For Mobile Users: Can Offloading Computation Save Energy", Published by the IEEE Computer Society, 2010.

[3] Nabil Sultan, "Cloud computing for education: A new dawn", International Journal of Information Management, vol. 30, pp. 109–116, 2010.

[4] Fatma A. Omara, Sherif M. Khattab and Radhya Sahal, "Optimum Resource Allocation of Database in Cloud Computing", Egyptian Informatics Journal, vol. 15, pp.1–12, 2014.

[5]Nazneen Taj, Dr. Anirban Basu, "International Journal of Engineering Research &Technology (IJERT)" ICIOT - 2016 Conference Proceedings, Volume 4, Issue 29, 2016.

[6]Soumya Ranjan Jena, V. Vijayaraja, Aditya Kumar Sahu, "Performance Evaluation of Energy Efficient Power Models for Digital Cloud" Indian Journal of Science and Technology, Vol 9(48), December 2016.

[7] Corradi A, Fanelli M, and Foschini L, "VM consolidation: A real case based on Open Stack Cloud", Future Generation Computer Systems, Vol. 32, pp. 118-127, 2014.

[8] Zhang F, Cao J, Li K, Khan S. U, and Hwang K, "Multi-objective scheduling of many tasks in cloud platforms" Future Generation Computer Systems, Vol. 37, pp. 309-320, 2014.

[9] H. Duan, C. Chen, G. Min and Y. Wu, "Energy-aware scheduling of virtual machines in heterogeneous cloud computing systems", Future Generation Computer Systems, vol. 74, pp. 142-150, 2017.

[10]F. Rossi, M. Xavier, C. De Rose, R. Calheiros and R. Buyya, "E-eco: Performance-aware energy-efficient cloud data center orchestration", Journal of Network and Computer Applications, vol. 78, pp. 83-96, 2017.

[11] W. Zhu, Y. Zhuang and L. Zhang, "A three-dimensional virtual resource scheduling method for energy saving in cloud computing", Future Generation Computer Systems, vol. 69, pp. 66-74, 2017.