BULGARIAN ACADEMY OF SCIENCES

CYBERNETICS AND INFORMATION TECHNOLOGIES • Volume **21**, No 3
Sofia • 2021 Print ISSN: 1311-9702; Online ISSN: 1314-4081
DOI: 10.2478/cait-2021-0035

VM Consolidation Plan for Improving the Energy Efficiency of Cloud

Satveer, Mahendra Singh Aswal

Department of Computer Science, Gurukula Kangri Vishwavidyalaya, Haridwar, 249404, India
E-mails: rs.satveer@gkv.ac.in msa@gkv.ac.in

Abstract**:** *Achieving energy-efficiency with minimal Service Level Agreement* (*SLA*) *violation constraint is a major challenge in cloud datacenters owing to financial and environmental concerns. The static consolidation of Virtual Machines* (*VMs*) *is not much significant in recent time and has become outdated because of the unpredicted workload of cloud users. In this paper, a dynamic consolidation plan is proposed to optimize the energy consumption of the cloud datacenter. The proposed plan encompasses algorithms for VM selection and VM placement. The VM selection algorithm estimates power consumption of each VM to select the required VMs for migration from the overloaded Physical Machine* (*PM*)*. The proposed VM allocation algorithm estimates the net increase in Imbalance Utilization Value* (*IUV*) *and power consumption of a PM, in advance before allocating the VM. The analysis of simulation results suggests that the proposed dynamic consolidation plan outperforms other state of arts.*

Keywords**:** *VM consolidation, VM selection, VM allocation, resource allocation, power consumption, energy efficiency.*

1. Introduction

Cloud computing runs over the parallel, distributed and grid systems and massive interconnected and virtualized servers or PMs. These PMs are dynamically provisioned and allocated, considered as one or more unified computing resource(s) based upon Quality of Service (QoS) formulized in SLA document. The major services offered through cloud data centers are applications, development platforms and computing power running inside the datacenters. However, these datacenters are consuming high electrical energy in allocating the resources to users' service request. [1]. This immense energy consumption is not only increasing the operating cost of data centers but also polluting the living environment by emitting the carbon dioxide.

In cloud datacenter, hypervisors enable sharing of same resources among user's applications by creating Virtual Machines (VM) on servers using virtualization. The resource allocation can be defined as the process of determination of how much, what, when and where to allocate the available resources to the user's application.

Energy-efficient resource allocation is the significant and motivating issue in the cloud [2, 3].

To provide high performance to user applications, service providers often overprovision the resources of the data centers, which results in low resource utilization and huge power consumption. Thus, the resource utilizations of virtual machines is much lower than the actual capacity of the PM (server). Moreover, servers consume maximum power when they are idle [5]. Thus, minimization of power consumption and SLA violation are two serious concerns, which need addressing during the resource allocation in cloud.

VM consolidation is an important approach to reduce energy consumption by migrating VMs from an under-utilized or overloaded PM to other PM. Dynamic VM Consolidation (DVMC) tries to allocate the maximum VMs over least possible PMs dynamically and puts the unused PMs into the energy saving mode [8]. VM migration is the key enabling technology that makes possible the existence of a dynamic VM consolidation process by transferring the virtual machines from one PM to another PM without rebooting the operating system of VMs [4]. However, aggressive VM consolidation may violate the SLA.

DVMC plan can be implemented by solving the four sub-problems: (1) Finding the overloaded PM; (2) Finding an under-loaded PM; (3) Selecting the VMs from the overloaded and under-loaded PMs; (4) Reallocating the selected VMs to the suitable PMs [5]. In this paper, we propose a new DVMC plan for improving the energy efficiency of a datacenter. Our research contributions are for the last two sub-problems of DVMC plan. The salient contributions of the paper are as follows:

- Proposed an energy aware VM selection algorithm (High Power – HP) that selects the most power intensive VM from overloaded PM for migration.
- Proposed a VM allocation algorithm named as Modified Energy-efficient VM Placement (MEVMP) to map the selected VMs to the suitable PM.
- Simulation results confirm the significant reduction in energy consumption and performance degradation as compared to the other state of arts.

Rest of sections of the paper are arranged as follows. Section 2 reviews the related work on energy aware resource management. Section 3 elaborates the evaluation models and performance metrics. Section 4 discusses the proposed VM consolidation plan and algorithms. In Section 5, simulation process and performance study of the proposed plan have been presented. The paper is summarized in Section 6.

2. Related work

Many researchers have proposed energy-efficient resource management algorithms to achieve the best energy and QoS trade-off. The works of various researchers are being reviewed in this section.

Beloglazov and Buyya [5] propose an architecture for green resource management by developing algorithms such as Minimum Utilization (MU), and Power Aware Best Fit Decreasing (PABFD) for VM selection and allocation respectively. MU selects a virtual machine, which has the least CPU utilization.

PABFD allocates a virtual machine on such a server, which incurs least increase in power. However, VM mapping may need to migrate multiple VMs simultaneously, which may enhance the performance degradation due to migrations. In order to get the optimized solution, researchers have used the particle swarm optimization approach to propose a new VM placement algorithm in [9] by considering the multiple resources of the system. Proposed algorithm minimizes the energy consumption of the datacenter. However, it generates large number of iterations to find the optimal solution, which leads to a high computational complexity. Moreover, the researchers have also not analyzed the performance of the algorithm during the resource allocation. Hence, the proposed solution does not ensure the QoS.

In [10], the concept of linear programming has been adapted to model VM placement problem as bin packing problem for improving the energy-efficiency. The proposed algorithm allocates the highest priority to those VMs, which have the more stability of the workload. However, the proposed VM placement plan has good proficiency to reduce the energy consumption only without taking into account the degradation of QoS. Bruno et al. [11] have worked on minimized active hardware using artificial intelligence based on Pseudo-Boolean (PB) constraints. The proposed work reduces power consumption by defining the Boolean variable for the active and inactive hardware without considering the SLA factor. Fard, Ahmadi and Adabi [12], propose new algorithms for VM selection and VM placement by optimizing the thermal state of the servers. The proposed work is able to mitigate the temperature and power consumption. However, proposed idea is not practical due to the PM's heterogeneity.

Researchers have developed a dynamic VM consolidation model by taking into consideration both the compute and cooling systems in [13]. The resource allocation has been carried out based on analyzing the different parameters of the system such as server processor, low power states, transition latency and integrated thermal controls. The model developed comprehensively optimizes the energy and temperature of the datacenter and reduces the power consumed by computing and cooling infrastructure.

In [14], authors have developed three prediction-based algorithms, Maximum Requested Resource, Minimum Downtime Migration and Multi-criteria Technique for order of preference by similarity to ideal solution. All the proposed algorithms select the virtual machines based on the predicted value of resource utilization instead of VM's current CPU utilization. The algorithms proposed reduce the energy consumption but consider only CPU to optimize the energy while avoiding other system's parameters [10]. Okada et al. [15] propose Global Power Aware Best Fit Decreasing algorithm for VM placement. The proposed work measures the power consumption of whole datacenter before allocating virtual machine and places virtual machines on such a PM, which requires least power. However, the proposed algorithm does not consider QoS.

Researchers in [16] use multi-objective modified differential evolution algorithm for resource allocation. The proposed algorithm is able to manage the energy and resource utilization efficiently with accounting different resources of the system. A VM selection method based on server utilization and minimum correlation coefficient has been proposed in [17]. Researchers in [18] propose a VM selection algorithm for selecting the virtual machine from overloaded server. The developed algorithm computes the CPU utilization of each virtual machines on the overloaded server and selects that VM which has the maximum CPU utilization. The proposed algorithm reduces the power consumption of the datacenter while providing a satisfactory QoS.

Wood et al. [19] propose a VM selection algorithm by measuring the V-value for selecting the virtual machine from the overloaded server. This V-value is calculated by calculating the average utilization of CPU, memory and bandwidth of the server. The proposed work improves the energy-efficiency. Tian et al. [20], introduce a load imbalance aware VM allocation algorithm and have named it dynamic and integrated resource scheduling. The proposed algorithm is able to manage the load of the system efficiently. However, there is no energy optimization in the developed solution.

An Energy-Efficient VM Placement (EVMP) algorithm is proposed by Lin, Liu and Guo [21]. The proposed algorithm calculates the **imbalance utilization value** and the power consumption of each server and selects a server, which has minimum IUV and power consumption. However, allocation of workload according to the current status of server, is not so appropriate and further leads to more power dissipation and performance degradation. Researchers in [19-21] indicate that the energy consumption of a server increases with increasing Imbalance Utilization Value (IUV) of the physical machine.

M selection algorithms proposed in [5, 12, 14, 17, 18] select a virtual machine based on the current resource utilization of the server and VM placement algorithms presented in [5, 10, 11, 12, 15, 21] estimate the instant power consumption of each server for deciding the most efficient server to allocate the virtual machines. However, these algorithms have different constraints while saving the energy of cloud data centers when allocating the resources.

In our proposed VM consolidation framework, strategies for selecting a VM from an overloaded server and then allocating the selected VM to an appropriate server have been presented. The framework is elaborated in section IV. The proposed dynamic VM consolidation framework reduces the energy consumption and SLA violation to a significant level.

3. Power modelling and performance metrics

3.1. Power modeling of the system

In this section, power consumption models used by our proposed DVMC plan are represented. The CPU is the main of power consumer in a PM. As the CPU utilization increases, power consumption of the server increases linearly [24, 25]. In order to measure the power consumption of a server, we have used the linear power model given in [5]:

(1) $$P(u) = kP_{\max} + (1-k)P_{\max}\,u.$$

Here: k is the fraction of the power consumed by the idle server; $P_{\max}$ is the maximum unit of power consumed at full CPU utilization of the PM; u represents the CPU utilization.

Let $\mathrm{VM}_{\mathrm{Util}_\mathrm{x}}$ is CPU utilization of VM_x, $\mathrm{VM}_{\mathrm{Count}}$ is the number of VM_s, $\mathrm{VM}_{\mathrm{Util}_\mathrm{y}}$ represents the CPU usage of VM_y on a particular server, and $\mathrm{PM}_{\mathrm{\_Power}}$ is the power consumption of the PM. The power consumption of a particular VM on a physical machine can be measured as given below referring to [26]:

(2) $$\mathrm{VM}_{\mathrm{\_Power}_i} = \mathrm{PM}_{\mathrm{\_Power}} * \frac{\mathrm{VM}_{\mathrm{Util}_\mathrm{x}}}{\sum_{j=1}^{\mathrm{VM}_{\mathrm{Count}}} \mathrm{VM}_{\mathrm{Util}_\mathrm{y}}}.$$

Equation (2) is used by our proposed Algorithm 1 (HP) in order to evaluate the power consumption of a VM on a particular PM, whereas Equation (1) is utilized by Algorithm 2 (MEVMP) for measuring the power consumption of a PM or server.

3.2. SLA violation metrics

Cloud computing is multi-tenant environment, where multiple users request for the resources and services at the same time. User's applications compete with each other for acquiring the required resources with users wishing to execute their applications with high Quality of Service (QoS). The QoS requirements are normally defined in Service Level Agreement (SLA) document. The three SLA violation metrics are defined here under three equations.

1. SLA violation: It is calculated as the product of two submetrics [5] as defined in

(3) $$\mathrm{SLAV} = \mathrm{SLATAH} * \mathrm{PDM}.$$

2. SLA Violation Time per Active Host (SLATAH): This is the percentage of time, in which a PM experiences 100% utilization of the CPU as defined in

(4) $$\mathrm{SLATAH} = \frac{1}{N}\sum_{i=1}^{N}\frac{T_{si}}{T_{\mathrm{a}i}}.$$

3. Performance Degradation due to Migrations (PDM): It indicates the overall performance degradation due to VM migrations and is shown in

(5) $$\mathrm{PDM} = \frac{1}{M}\sum_{j=1}^{M}\frac{C_{\mathrm{d}j}}{C_{\mathrm{r}j}}.$$

Here: N, M are the numbers of PMs and VMs, respectively; T_{si} is the sum of all time frames, in which PM has experienced 100% CPU utilization; $T_{\mathrm{a}i}$ is the total active time of the PM; $C_{\mathrm{d}j}$ represents the performance degradation due to migrations by VM_j; $C_{\mathrm{r}j}$ is the total requested amount of CPU usage by VM_j during its life time; $C_{\mathrm{r}j}$ has been set to 10% of the CPU utilization during all migrations of the VM_j.

3.3. Performance metric

B e l o g l a z o v and B u y y a [5] derive new performance metric Energy (E) and SLA Violation (ESV) by combining energy consumption and SLA violation to assess the overall effectiveness of the proposed work. The ESV is represented in the equation

(6) $$\mathrm{ESV} = E * \mathrm{SLAV}.$$

4. Proposed framework for energy aware dynamic VM consolidation

In order to shape the cloud datacenter more energy-aware and cost-effective, a framework for the Dynamic VM Consolidation (DVMC) is presented in this section and is illustrated in Fig. 1.

The presented plan consists of four major phases namely monitoring, workload analysis, decision and actuation. In the monitoring phase, the information of the system regarding resource utilization, power consumption of virtual machines, workload, increase in Imbalance Utilization Value (IUV) and power consumption of physical machines, are monitored and collected. Workload analysis determines the status of a physical machine, whether it is overloaded or under-loaded and optimizes the resource utilization of each physical machine. If resource utilization of a physical machine exceeds the upper threshold, then it is declared as overloaded. For finding an under-loaded physical machine, the underutilized physical machine from the list of active PMs by comparing the CPU utilization of each physical machine is determined. The decision phase consists of two engines, VM selection and placement and both of these exploit the collected knowledge from the monitoring phase to perform the DVMC operation. VM migration engine optimizes which virtual machine should be migrated and the placement engine determines the optimal physical machine for the allocation of selected virtual machine. Actuation phase carries out the operations for the DVMC plan and switches the PMs on-off according to resource usage. The nomenclature and input parameters are explained in Table 1.

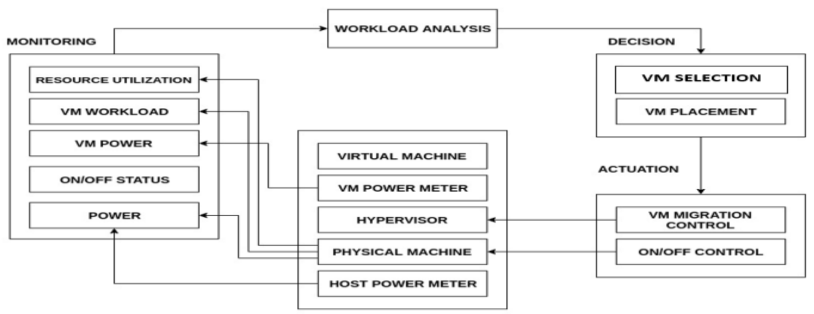

Fig. 1. Proposed framework for dynamic VM consolidation

Table 1. Nomenclature and used parameters

Parameters	Meaning
$v_{\text{cpu}}^{\text{r}}, v_{\text{mem}}^{\text{r}}$ and $v_{\text{net}}^{\text{r}}$	Current requested MIPS, Memory and Network Bandwidth by VM, respectively
$\text{VM}_{\text{\_Power}_i}$, $\text{PM}_{\text{\_Power}}$	Power consumption of VM and PM
$h_{\text{cpu}}^{\text{c}}, h_{\text{mem}}^{\text{c}}$ and $h_{\text{net}}^{\text{c}}$	Computing, Memory and Network Bandwidth capacity of PM, respectively
$h_{\text{cpu}}^{\text{u}}, h_{\text{mem}}^{\text{u}}$ and $h_{\text{net}}^{\text{u}}$	CPU, Memory and Network Bandwidth utilization of PM, respectively
$h_{\text{p\_cpu}}^{\text{u}}, h_{\text{p\_mem}}^{u}$ and $h_{\text{p\_net}}^{\text{u}}$	CPU, Memory and Network Bandwidth potential utilization of PM, respectively
$\text{Avg\_CPU}_i^{\text{u}}$, $\text{Avg\_MEM}_i^{\text{u}}$ and $\text{Avg\_NET}_i^{\text{u}}$	Average CPU, Memory and Network Bandwidth utilization of PM, respectively

4.1. Consolidation plan

The DVMC manages the resources of datacenter in an energy-efficient way through reallocating the workload. Our proposed plan performs reallocation by segregating the running servers or PMs as overloaded and under-loaded.

In case of overloaded server, the proposed plan accomplishes the following steps: (i) selection of required virtual machines for migration; (ii) allocates the selected VMs to the appropriate server. While in the case of under-loaded server, two steps are performed: (i) select all the running VMs; (ii) migrate all selected VMs to selected server or PMs.

The workload analyzer securitizes the usage of servers and gets a list of overloaded servers. Next, the VM selection algorithm determines which virtual machine should be selected for migration and VM placement algorithm investigates the suitable destination for placing selected VM from the list of active PMs. This process of VM selection and allocation continues until an overloaded server becomes non-overloaded. The other way round, all running virtual machines are selected for migration using VM selection algorithm from under-loaded PMs. These selected virtual machines, are taken away from under-loaded servers, by the VM placement algorithm to an appropriate server in the list of alive PMs. Thereafter, these empty servers are set to sleep or power-saving mode in order to save power.

4.2. VM Selection Algorithm

Once, it is ensured that a particular server is overloaded with the help of any server overloaded detection algorithm, the next step selects a particular VM(s) from the overloaded server using a VM selection algorithm.

In order to reduce the power consumption of a PM instantly and to change it from overloaded state to normal-loaded state, we propose a new VM selection algorithm namely High Power (HP) selection for selecting the virtual machine(s). The proposed algorithm first estimates the power consumption of a VM using Equation (2) and then selects a VM, which consumes highest amount of power on the overloaded server. If the server is still overloaded and not deemed to be normal loaded after VM migration, then the VM selection algorithm selects again the second-highest power consuming VM and this process continues until the PM is deemed to be not overloaded.

The proposed VM selection algorithm decreases power consumption, workload of the server and performs a number of migrations. Minimizing the number of VM migrations will in turn reduce the energy consumption while preserving the SLA. The pseudo code of the proposed work is presented in Algorithm 1.

Algorithm 1. HP (VM Selection Algorithm)

Input: Overloaded_PM_list, vm_List

Output: vm_To_migrate

```
  vm_To_migrate = null
   vm_List ← PM. getvmlist ()
    min_ Power←double. MIN_Value
double VM_Util_x , VM_Util_y,PM_Power
   For each vm in vm_List do
```

vm_Utily ← util_SumofAllVMs ()
vm_Utilx ← util_ofSingleVM ()
$\mathrm{PM}_{\_\mathrm{Power}}$ ← estimatePower ()
$\mathrm{VM}_{\_\mathrm{Power_i}}$ ← estimateVmPower ($\mathrm{VM\_Util_x}$, $\mathrm{VM\_Util_y}$, $\mathrm{PM}_{\_\mathrm{Power}}$)
Power_i = $\mathrm{VM}_{\_\mathrm{Power_i}}$
If Power_i>min_ Power then
min_Power = Power_i
vm_To_migrate = vm
EndIf
EndFor
Return vm_To_migrate

4.3. VM Placement Algorithm

VM placement improves the resource utilization through reallocating the workload from one PM to other PM to save energy. Additionally, it can be viewed as the bin-packing problem with variable bin sizes and prices. Here, server is represented as the bins and VMs are items that need to be packed into minimum servers.

We propose a new VM placement algorithm as Modified Energy-Efficient Virtual Machine Placement (MEVMP). The proposed strategy searches optimal server for allocating the VM via performing two steps: (i) First it checks which server can satisfy the CPU, memory, and network bandwidth requirements of migrating VM; (ii) Estimates in advance the increase in imbalance resource utilization and power consumption of each server assuming virtual machine allocation to it and then (iii) Selects a server, which incurs the least increase in imbalance resource utilization and power consumption taken together.

The power consumption of a server is estimated using the Equation (1). The Integrated Resource (IR) utilization and Imbalance Utilization Value (IUV) of a server considering multi-dimensional resource constraints can be computed by referring to [28, 29] as given in the next equations:

(7) $$\mathrm{IR} = \frac{\mathrm{Avg\_CPU_i^u} + \mathrm{Avg\_MEM_i^u} + \mathrm{Avg\_NET_i^u}}{3},$$

(8) $$\mathrm{IUV}_i = \frac{\left(\mathrm{Avg\_CPU_i^u} - \mathrm{IR_i^u}\right)^2 + \left(\mathrm{Avg\_MEM_i^u} - \mathrm{IR_i^u}\right)^2 + \left(\mathrm{Avg\_MEM_i^u} - \mathrm{IR_i^u}\right)^2}{3}.$$

The placement engine of the proposed plan finds a suitable physical machine for allocating a VM, by using a variation of selection criteria function as proposed in [21] and is given as

(9) $$Z = w_{\mathrm{i}}.\,\mathrm{IUV^{In}} + w_{\mathrm{p}}.\,P_{\mathrm{u}}^{\mathrm{In}},$$

where, w_{i} and w_{p} are the weights associated with resource imbalanced rate and power consumption values such that, $w_{\mathrm{i}} + w_{\mathrm{p}} = 1$. Where $\mathrm{IUV^{In}}$ and $P_{\mathrm{u}}^{\mathrm{In}}$ represent the net increase in the value of IUV and power consumption of a physical machine respectively after allocation of virtual machine.

The proposed algorithm restricts a virtual machine to migrate on such a physical machine, which is already producing a high value of IUV and power consumption. Thus, it causes reduction in energy consumption, virtual machine migrations and SLA violation. The pseudo code of the proposed work is given in Algorithm 2.

Algorithm 2. MEVMP (VM Placement Algorithm)
Input: PM_List, VMs
Output: Allocation_of_VMs
Migration_Vm_List
For each vm in Migration_Vm_List **Do**
    min_Increase_Power_IUV = **Max**
    For each PM in PM_List do
        If PM has enough resources for **vm Then**
            Increase_Power_IUV = estimateIncreasein_Power_ IUV
            If Increase_Power_IUV < min_Increase_Power_IUV
                min_Increase_Power_IUV = Increase_Power_IUV
                allocated_PM = PM
            EndIF
        EndIF
    EndFor
EndFor
Return allocation

5. Simulation testbed and result analysis

Several simulation-based experiments have been carried out to analyze, optimize and validate the efficiency of research work using the CloudSim [30]. CloudSim is widely used simulator to model and simulate resource allocation and provisioning policies to applications particularly to optimize the power consumption of the datacenters. Additionally, it provides extensible APIs, which can be extended to required level. We have considered that each physical machine is equipped with multi-cores CPU having *n* cores and a single core having *m* MIPS. In the simulation environment, 100 heterogeneous PMs and 200 VMs of four types have been taken [31]. The configuration of simulated physical machines is shown in Table 2. The specifications of simulated VMs are shown in Table 3 [32].

The real PlanetLab workload provided by the CoMon project lab has been used in the experiment to check the effectiveness of the proposed work [33]. The workload consists of periodic samples of CPU utilization of VMs collected at interval of each five minutes and workload allocation to virtual machines takes place randomly at initial level.

The performances of proposed algorithms HP and MEVMP have been evaluated by alternately executing these algorithms along with some server overload detection algorithms like Median Absolute Deviation (MAD) and Inter Quartile Range (IQR). Accordingly, other VM selection algorithm MU and VM placement algorithm PABFD have been also executed [5]. The experimental results pertaining to various consolidation plans for the given workload have been generated and compared. The experimental results are shown in Table 4.

Table 2. The characteristics of PMs

Server type	CPU	Cores	MIPS	RAM (MB)	BW
Fujitsu M1	1230	4	2700	8192	1
Fujitsu M3	1230	4	3500	8192	1

Table 3. Types of VMs and their specifications

VM type	No of cores	MIPS	RAM (MB)
#1	1	2500	850
#2	1	2000	1740
#3	1	1000	1740
#4	1	500	613

Table 4. Experimental results

DVMC Plan	IQR_HP_MEVMP	MAD_HP_MEVMP	IQR_HP_PABFD	MAD_HP_PABFD	MAD_MU_PABFD	MAD_MU_MEVMP	MAD_MU_EVMP	IQR_MU_PABFD	IQR_MU_MEVMP	IQR_MU_EVMP
Energy (kW.h)	4.6	4.28	10.5	4.62	21.89	18.85	5.2	23.97	18.47	5.14
SLAV ($\times 10^{-3}$)	0.23	0.16	3.45	0.32	5.65	4.77	15	6.06	4.15	13.54
VM Migrations	657	516	2312	655	5717	4932	13036	5974	4608	12116
PDM	0.01	0.01	0.06	0.02	0.09	0.08	0.25	0.09	0.07	0.24
SLATAH (%)	2.28	2.16	6.51	2.88	7.15	6.23	6.16	7.35	6.64	5.96
ESV	0.0105	0.0068	0.3622	0.0147	1.236	0.899	0.7935	1.4525	0.655	0.6959

5.1. Energy savings

The proposed DVMC plan HP_MEVMP consumes less electrical energy (4.6 kW.h) as compared to IQR_MU_PABFD (23.97 kW.h) and MAD_MU_EVMP (5.2 kW.h) respectively as shown in Fig. 2. Thus, the proposed plan maximizes the energy-efficiency of the cloud datacenter.

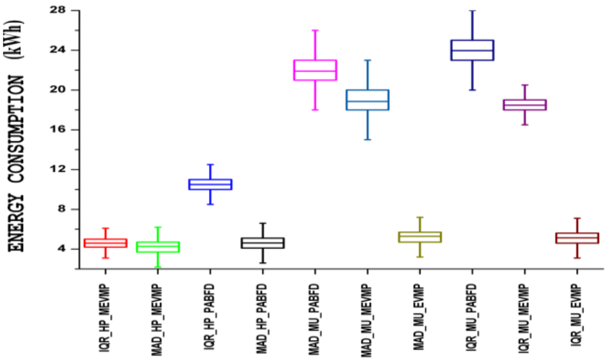

Fig. 2. The energy consumption

5.2. SLA violation

According to Fig. 3, the HP_ MEVMP produces minimum value of SLA violation (0.23×0.001) in comparison of IQR_MU_PABFD (6.06×0.001) and

MAD_MU_PABFD (5.65×0.001). This means that both our algorithms are capable to assure the high QoS as compared to MU and PABFD.

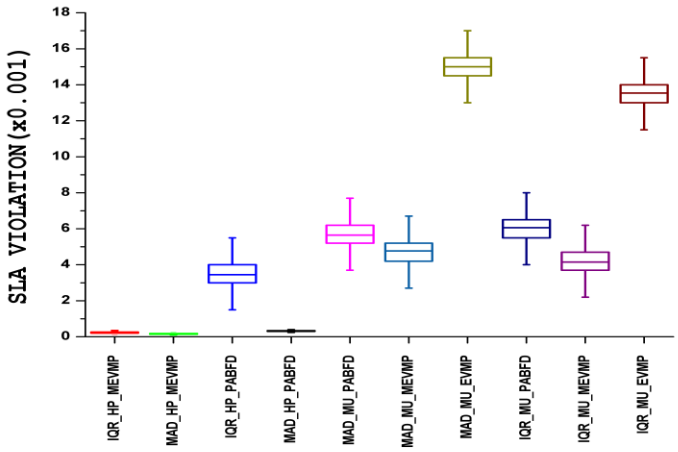

Fig. 3. The SLA violation

5.3. VM migrations

The high number of VM migrations always leads to high SLA violation of the deployed service. As evident from Fig. 4, the proposed plan HP_ MEVMP along with MAD (516) and IQR (657) produces the minimum number of VM migrations. Hence, the proposed algorithms are more adequate to decrease the energy required for migrations and the SLA violation thereof.

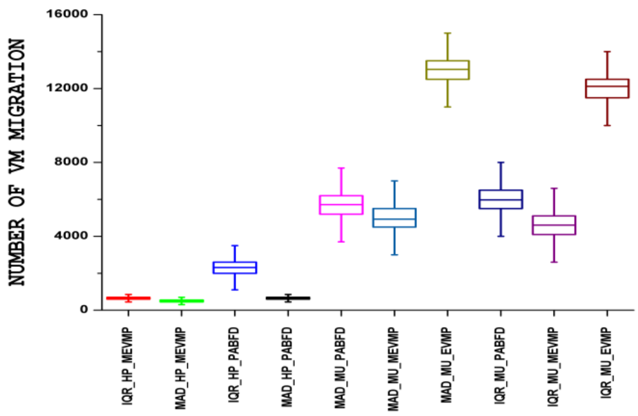

Fig. 4. Number of VM migration

5.4. PDM

The Performance Due to Migration (PDM) rises as the VM migrations grow. The proposed algorithms, HP and MEVMP along with MAD (0.01), produce fairly low values of PDM in contrast to IQR_MU_PABFD (0.08), and MAD_MU_PABFD (0.08), as shown in Fig. 5.

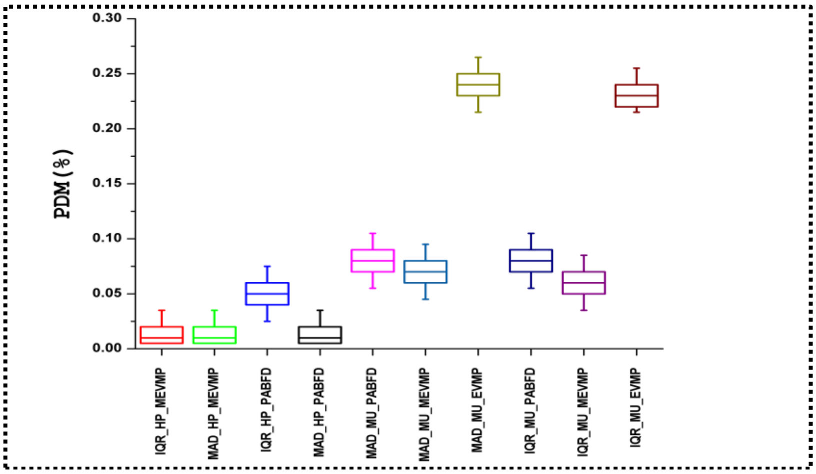

Fig. 5. The PDM

5.5. SLATAH

The experimental results of HP_ MEVMP when integrated with MAD (2.16) and IQR (2.08) produce comparatively minimum values of SLATAH as shown in Fig. 6. This means that proposed consolidation plan HP_MEVMP reduces the chances of 100% CPU utilization of active server and executes the application workloads with low SLA Violation.

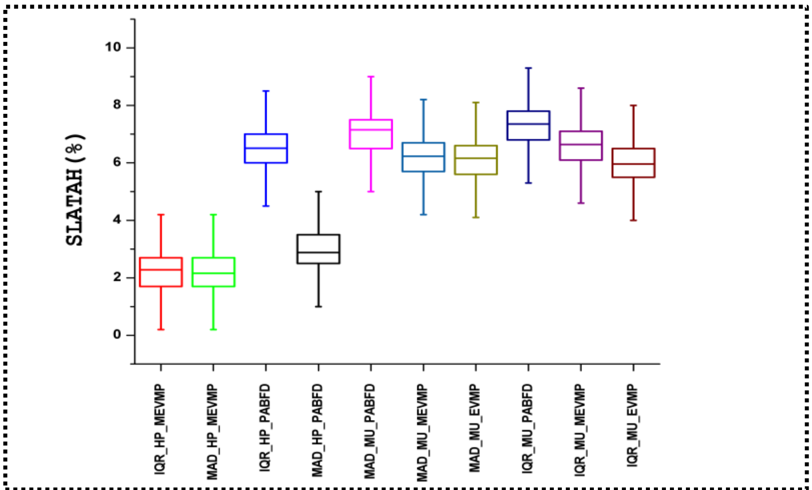

Fig. 6. The SLATAH

5.6. ESV

Energy and SLA Violation (ESV) is the composite metric of energy consumption and SLA violation. The minimum value of ESV is better for the overall performance of the datacenter. The experimental results of simulation in Fig. 7 indicate that HP_MEVMP along with MAD (0.001) produces minimum value of ESV as compared to the IQR_MU_PABFD (0.145) and MAD_MU_PABFD (0.123).

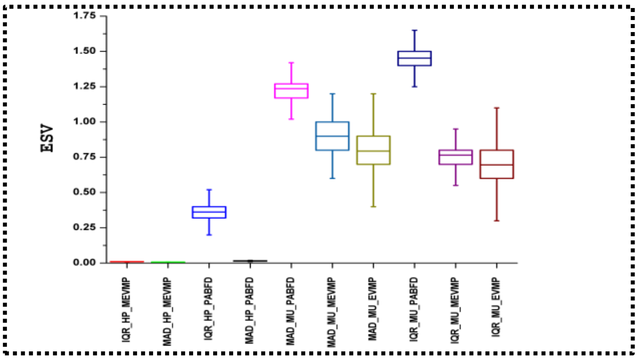

Fig. 7. The ESV

6. Conclusion

Recently, the issue of energy-efficiency has assumed significance while allocating the resources in cloud datacenters. We have proposed an energy-aware dynamic consolidation plan, comprising of two new algorithms for VM selection (HP) and VM Placement (MEVMP), resulting in the reduction of energy consumption and SLA violation of a datacenter. In particular, the VM selection algorithm being proposed selects virtual machine(s) from an overloaded physical machine consuming maximum power to reduce the instant power consumption of the physical machine. While on the other side, the proposed placement algorithm allocates the selected virtual machine on a PM, which generates the least increase in Imbalance Utilization Value (IUV) and power consumption after allocating the selected VM. The proposed placement algorithm estimates the increase in IUV and power consumption by considering the potential use of CPU utilization, memory and network bandwidth of a physical machine.

The proposed plan has been simulated using CloudSim simulator. The experimental results of simulation suggest that the framework being proposed outperforms other static power management and dynamic VM consolidation plans. Thus, our solution being proposed maximizes the energy-efficiency of the datacenter and reduces the SLA violation by optimizing the multiple resources of the system. Further, it can be inferred that the proposed solution also results in reducing the carbon footprints in the environment.

References

1. S h y n u, P. G., K. J. S i n g h. A Comprehensive Survey and Analysis on Access Control Schemes in Cloud Environment. – Cybernetics and Information Technologies, Vol. **16**, 2016, No 1, pp. 19-38.
2. B e l o g l a z o v, A., R. B u y y a. Managing Overloaded Hosts for Dynamic Consolidation of Virtual Machines in Cloud Data Centers under Quality-of-Service Constraints. – IEEE Transactions on Parallel and Distributed Systems, 2013. DOI: 10.1109/TPDS.2012.240.
3. P u h a n, S., D. P a n d a, B. K. M i s h r a. Energy Efficiency for Cloud Computing Applications: A Survey on the Recent Trends and Future Scopes. IEEE Xplore, 2020.
4. W a n g, H., H. T i a n f i e l d. Energy-Aware Dynamic Virtual Machine Consolidation for Cloud Datacenters. – IEEE, Vol. **6**, 2018, pp. 15259-15273. DOI: 10.1109/ACCESS.2018.2813541.

5. B e l o g l a z o v, A., R. B u y y a. Optimal Online Deterministic Algorithms and Adaptive Heuristics for Energy and Performance Efficient Dynamic Consolidation of Virtual Machines in Cloud Datacenters. – Concurr. Comput. Pract., 2013, pp. 1397-1420.
https://doi.org/10.1002/cpe.1867
6. M a s d a r, M., M. Z a n g a k a n i. Green Cloud Computing Using Proactive Virtual Machine Placement: Challenges and Issues, Springer Nature B. V. 2019. – J. Grid Computing.
https://doi.org/10.1007/s10723-019-09489-9
7. L e e, H. M., Y. J e o n g, H. J. J a n g. Performance Analysis-Based Resource Allocation for Green Cloud. – J. Supercomput., Vol. **69**, 2014, pp. 1013-1026.
https://doi.org/10.1007/s11227-013-1020-x
8. E s f a n d i a r p o o r, S., A. P a h l a v a n, M. G o u d a r z i. Structure-Aware Online Virtual Machine Consolidation for Datacenter Energy Improvement in Cloud Computing. – Comput. Electr. Eng., Vol. **42**, 2015.
https://doi.org/10.1016/j.compeleceng.2014.09.005
9. M a d h u m a l a, R. B., H. T i w a r i, C. D e v a r a j V e r m a. Virtual Machine Placement Using Energy – Efficient Particle Swarm Optimization in Cloud Datacenter. – Cybernetics and Information Technologies, Vol. **21**, 2021, No 1, pp. 62-72.
10. F e r r e t o, T. C., M. A. S. N e t t o, R. N. C a l h e r i o u s, C. A. F. D e R s o e. Server Consolidation with Migration Control for Virtualized Data Centers. – Future Generation Computer Systems, October 2011.
https://doi.org/10.1016/j.future.2011.04.016
11. B r u n o, B. C., C. R i b a s, R. M. S u g u i m o t o, R. A. N. R. M o n t a ñ o, F. S i l v a, L. D. B o n a, M. A. C a s t i l h o. On Modelling Virtual Machine Consolidation to Pseudo-Boolean Constraints. – J. Pavon, Ed. 2012. pp. 361-370.
https://doi.org/10.1007/978-3-642
12. F a r d, S. Y. Z., M. R. A h m a d i, S. A d a b i. A Dynamic VM Consolidation Technique for QoS and Energy Consumption in Cloud Environment. – J. Supercomput., Vol. **73**, 2017, pp. 4347-4368.
https://doi.org/10.1007/s11227-017-2016-8
13. K u m a r, M. R. V., S. R a g h u n a t h a n. Heterogeneity and Thermal Aware Adaptive Heuristics for Energy Efficient Consolidation of Virtual Machines in Infrastructure Clouds. – Journal of Computer and System Sciences, March 2016.
https://doi.org/10.1016/j.jcss.2015.07.005
14. A r i a n y a n, E., H. T a h e r i, S. S h a r i f i a n. Novel Energy and SLA Efficient Resource Management Heuristics for Consolidation of Virtual Machines in Cloud Data Centers. – Comput. Electr. Eng., Vol. **47**, 2015, pp. 222-240.
https://doi.org/10.1016/j.compeleceng.2015.05.006
15. O k a d a, T. K., A. D e L a F u e n t e V i g l i o t t i, D. M. B a t i s t a, A. G o l d m a n vel L e j b m a n. Consolidation of VMs to Improve Energy Efficiency in Cloud Computing Environments. – In: Proc. of 2015 XXXIII Brazilian Symposium on Computer Networks and Distributed Systems, Vitoria, 2015, pp. 150-158. DOI: 10.1109/SBRC.2015.27.
16. K o l l u, A., V. S u c h a r i t a. Energy-Aware Multi-Objective Differential Evolution in Cloud Computing. – In: S. Dash, S. Das, B. Panigrahi, Eds. Proc. of International Conference on Intelligent Computing and Applications. Advances in Intelligent Systems and Computing. Vol. **632**. Singapore, Springer, 2017.
https://doi.org/10.1007/978-981-10-5520-1_40
17. H o r r i, A., M. S. M o z a f a r i, G. D a s t g h a i b y f a r d. Novel Resource Allocation Algorithms to Performance and Energy Efficiency in Cloud Computing. – J. Supercomput., Vol. **69**, 2014, pp. 1445-1461.
https://doi.org/10.1007/s11227-014-1224-8
18. M a n d a l, R., M. K. M o n d a l, S. B a n e r j e e, U. B i s w a s. An Approach toward Design and Development of an Energy-Aware VM Selection Policy with Improved SLA Violation in the Domain of Green Cloud Computing. Springer Science+Business Media, LLC, Part of Springer Nature 2020, The Journal of Supercomputing.
https://doi.org/10.1007/s11227-020-03165-6

19. W o o d, T., P. S h e n o y, A. V e n k a t a r a m a n i, M. Y o u s i f. Black-Box and Gray-Box Strategies for Virtual Machine Migration. – In: Proc. of 4th USENIX Symposium on Networked Systems Design Implementation (NSDI'07), 11-13 April 2007, USA, pp. 229-242.
20. T i a n, W., Y. Z h a o, Y. Z h o n g, M. X u, C. J i n g. A Dynamic and Integrated Load-Balancing Scheduling Algorithm for Cloud Datacenters. – In: Proc. of 2011 IEEE International Conference on Cloud Computing and Intelligence Systems, 2011, pp. 311-315. DOI: 10.1109/CCIS.2011.6045081.
21. L i n, X., Z. L i u, W. G u o. Energy-Efficient VM Placement Algorithms for Cloud Data Center. – In: W. Qiang, X. Zheng, C. H. Hsu, Eds. Proc. of Cloud Computing and Big Data. CloudCom-Asia 2015. Vol. **9106**. Cham, Springer, 2015. **https://doi.org/10.1007/978-3-319-28430-9_4**
22. G r e e n b e r g, A., D. H a m i l t o n, A. M a l t z, P. P a t e l. The Cost of a Cloud Research Problems in Data Centers Networks. – In: Proc. of ACM SICOMM, Vol. **39**, 2009, No 1, pp. 68-73.
23. K h o s r a v i, A., S. K G a r g., R. B u y y a. Energy and Carbon-Efficient Placement of Virtual Machines in Distributed Cloud Data Centers. – In: F. Wolf, B. Mohr, D. an Mey, Eds. Proc. of Euro-Par 2013 Parallel Processing. Euro-Par 2013. Lecture Notes in Computer Science, Vol. **8097**. Berlin, Heidelberg, Springer, 2013. **https://doi.org/10.1007/978-3-642-40047-6_33**
24. M a z z u c c o, M., D. D y a c h u k, R. D e t e r s. Maximizing Cloud Providers' Revenues via Energy Aware Allocation Policies. – In: Proc. of 2010 IEEE 3rd International Conference on Cloud Computing, Miami, FL, 2010, pp. 131-138. DOI: 10.1109/CLOUD.2010.68.
25. R i v o i r e, S., P. R a n g a n a t h a n, C. K o z y r a k i s. A Comparison of High-Level Full-System Power Models. – In: Proc. of 2008 Conference on Power Aware Computing and Systems (HotPower'08), USENIX Association, USA.
26. K a v a n a g h, R., D. A r m s t r o n g, K. D j e m a m e, D. S o m m a c a m p a g n a, L. B l a s i. Towards an Energy-Aware Cloud Architecture for Smart Grids. – In: J. Altmann, G. Silaghi, O. Rana, Eds. Proc. of Economics of Grids, Clouds, Systems, and Services (GECON'15). Vol. **9512**. Cham, Springer, 2015. **https://doi.org/10.1007/978-3-319-43177-2_13**
27. V o o r s l u y s, W., J. B r o b e r g, S. V e n u g o p a l, R. B u y y a. Cost of Virtual Machine Live Migration in Clouds: A Performance Evaluation. – In: M. G. Jaatun, G. Zhao, C. Rong, Eds. Proc. of Cloud Computing. CloudCom 2009. Lecture Notes in Computer Science. Vol. **5931**. Berlin, Heidelberg, Springer. **https://doi.org/10.1007/978-3-642-10665-1_23**
28. H o n g y o u, L., W. J i a n g y o n g, P. J i a n, W. J u n f e n g, L. T a n g. Energy-Aware Scheduling Scheme Using Workload-Aware Consolidation Technique in Cloud Data Centres. – China Communications, Vol. **10**, 2013, No 12, pp. 114-124.
29. S i m a r r o, J. L. L., R. M. V o z m e d i a n o, R. S. M o n t e r o, I. M. L i o r e n t e. Scheduling Strategies for Optimal Service Deployment across Multiple Clouds. – Future Generation Computer Systems, Vol. **29**, 2013, No 6, pp. 1431-1441. **https://doi.org/10.1016/j.future.2012.01.007**
30. C a l h e i r o s, R. N., R. R a n j a n, A. B e l o g l a z o v, D e C. A. F. R o s e, R. B u y y a. CloudSim: A Toolkit for Modeling and Simulation of Cloud Computing Environments and Evaluation of Resource Provisioning Algorithms. – Softw. Pract. Exp., Vol. **41**, 2010, No 1, pp. 23-50, **https://doi.org/10.1002/spe.995**
31. **https://www.spec.org/power_ssj/results/**
32. **https://aws.amazon.com/ec2/instance-types/**
33. P a r k, S., V. P a i. CoMon Monitoring System for Planet Lab. – ACM SIGOPS Operating Systems Review, Vol. **40**, January 2006, Issue 1, pp. 65-74. **https://doi.org/10.1145/1113361.1113374**

Received: 18.01.2021; Second Version: 04.07.2021; Accepted: 10.08.2021