UnPMD: An Effective Power Saving Underloaded Physical Machine Detection Method in the Cloud Computing Environment

Deafallah Alsadie

Department of Information Systems
Umm Al-Qura University
Makkah, Saudi Arabia

Summary
Workload consolidation is a significant task in effectively using Cloud Computing resources over a minimum number of the physical machine. It enables reduction in cloud data center's power consumption while meeting service level agreements between cloud users and service providers. It requires accurate determination of over utilized and underutilized physical machines that allows transfer of suitable virtual machines to other machines and ensure the least count of active physical machines. To that end, this work proposes a new method for detecting underutilized physical machines (UnPMD) based on the overall workload of the cloud data center. Besides, it also determines the coutn of active physical machines that can be vacated. It also avoids overloading the active physical machines while shifting the virtual machines from under utilized physical machines. The proposed (UnPMD) method is experimentally validated through simulations using cloudsim software. The obtained results demonstrate the superiority of the proposed (UnPMD) method over existing methods in terms of power consumption, Virtual Machine migration, service level agreement violation, and the number of physical machine shutdowns. The proposed (UnPMD) method can result into a 62% improvement in power consumption, a 109 % decrease in the number of service level agreement violations, and a 54% increase in the number of physical machine shutdowns over the existing methods under different scenarios.
Keywords:
Cloud computing; Energy efficiency; VM consolidation; Underloaded server; Utilization.

1. Introduction

In recent years, Cloud Computing has become the most popular model for provisioning a shared pool of configurable resources such as processing power, storage, and bandwidth [1]. Cloud Computing has enabled access to shared configurable resources in different models for providing different types of services to cloud users. The famous Cloud Computing models include infrastructure as a service, platform as a service, software as a service that enables to use of hardware, software development platform, and a variety of software as a service. These models provide on demand services to cloud users that can be quickly provisioned and revoked with minimum management efforts. The cloud users pay for cloud services on a pay as you use basis model.

Cloud Computing services have been widely accepted in academia and industry due to the availability of many pervasive applications at economical prices compared to conventional computing models [2, 3]. Cloud Computing services have been deployed in different deployment models [4]. The most common Cloud Computing models include private clouds, public clouds, hybrid clouds and community clouds. Cloud users can avail any Cloud Service deployment based on their requirements without any restriction on their location through the internet [5].

The Cloud Computing environment contains multiple cloud data centres distributed over a geographical area. Cloud data centres are equipped with multiple physical machines [6]. The physical machines comprise the installation of virtual machines to fulfil cloud users' computing requirements [7]. The tremendous growth in the use of cloud computing services and the number of cloud users have resulted in the development of large scale cloud data centres with a large number of physical machines to handle a massive amount of computing requirements of cloud users.

A tremendous increase in physical machines in different cloud data centres has increased the use amount of power consumption and caused the generation of a high amount of carbon dioxide. It has been predicted that cloud data centres consume about 3% of total power over the globe in 2020 [8]. This consumption of power resulted in 200 Mio. tons of carbon dioxide. The high carbon dioxide emission contributes to increasing global warming.

The used amount of power consumption and consequently emission of carbon dioxide has resulted in energy wastage, system instability and reduced return on investment [9]. It has been observed that most physical machines located at cloud data centres run under 10% to 50% of their capacity. Therefore, most of the active physical machines can we optimized for their utilisation that can help reduce the cloud data center's power consumption [10]. The process of optimising computing resource utilisation of cloud data center's physical machines with the aim of reducing power consumption as workload consolidation. The techniques used to consolidate physical machines and cloud data centre's virtual machines focus on services and applications with minimum computing resources. These techniques help optimize computing resource utilisation

Manuscript received April 5, 2022
Manuscript revised April 20, 2022
https://doi.org/10.22937/IJCSNS.2022.22.4.84

while meeting service level agreements between cloud users and service providers without degrading computing performance [11]. A promising consolidation method involves reallocating virtual machines regularly for minimizing the active physical machines which can help in reducing the cloud data center's power consumption. These consolidation techniques use online virtual machine migration from one physical machine to other while maintaining appropriate computing performance and providing services to cloud users with minimum down time.

A typical consolidation method involves four primary tasks [12]. The first most important task is the determination of over utilized physical machines violating service level agreement between cloud user and service provider. Secondly, the consolidation method detects underutilized physical machines that require shifting all hosted virtual machines to other physical machines. There is a minimum number of active cloud data center's physical machines. The third step includes the selection of a proper virtual machine for shifting from one over utilized physical machine to another. The final task involves selection of appropriate physical machine for placing the selected virtual machine without overloading it while maintaining service level agreement between cloud user and Cloud Service Provider. Detection of under utilized physical machines during the consolidation contributes to minimize the active physical machines and optimizes cloud data centers' power consumption [13].

Most existing methods for consolidating cloud data centre's virtual machines focus on power consumption management by setting up threshold levels of power consumption statically or dynamically. These methods decide threshold limits based upon current usage of physical machines in detecting underutilized physical machines. The existing methods suffer from certain limitations of ignoring overall cloud data center's power consumption while determining under utilized and over utilized physical machine in cloud data center.

This work proposes a power saving method for detecting underutilized physical machines (UnPMD) based on the total workload of the cloud data centre. It computes the number of physical machines that must be vacated to optimise the power consumption of the cloud data centre based upon lower utilisation and upper utilisation threshold values of the physical machine. It also identifies new physical machines for placing the migrating virtual machine. It maintains two separate lists for physical machines as over utilised and underutilized physical machines.

The rest of the paper is organised as follows. Section 2 to describe the state of art in detecting underutilized physical machines of cloud data centre. Section 3 represents the proposed system model and power saving underutilized physical machine detection method. It also explains the algorithm for the proposed (UnPMD) method. Section 4 highlights the setup for conducting experiments in this work, followed by definition of performance metrics used to measure the proposed (UnPMD) method's performance along with the existing methods. It provides experimental results and discussion. Finally, the paper is concluded in section 5.

2. Related work

There is plenty of literature on consolidating virtual machines dynamically in cloud data centres. However, limited research efforts have been proposed to consume power in cloud data centres efficiently. For example, Nathuji et al. [14] proposed an effective method for consolidating virtual machines to minimize physical machine power consumption in the cloud data center. They analyzed different power management methods regarding virtual computing resources. They suggested a power management approach or by separating this problem into two levels global and local level. At the local level, their proposed algorithm maintains power management of the guest virtual machines. The research has also highlighted the advantages of Power Management consumption methods.

Belgicus et al. [15] proposed a method for dynamically allocating computing resources for fulfilling cloud user requirements and quick response. They used multi objective anti lion algorithm for optimising makes pan and virtual machine cost.

Zhu et al. [16] analyzed the Virtual Machine consolidation problem based upon CPU usage. They use static threshold values of 85% CPU utilisation. They assume that CPU utilisation above 85% will be considered as over utilized physical machine. There proposed method suffers from limitation of unsuitability for the dynamic workload data centres. The static threshold values of CPU utilisation are unable to manage change in workload in cloud data center.

Moghaddam et al. [17] focused on intelligently migrating cloud data centre's virtual machines. They proposed their method based upon in modified cellular learning automata approach that reduces power consumption and minimizes Virtual Machine migrations. In their method, the authors used an optimized placement approach and propose to delay migration time depending upon predicted future demands of computing resources.

Similarly, Gua et al. [18] also tackled the same problem by allocating cloud data center workload taking into consideration thermal storage and renewable power resources.

The authors of [19, 20] presented the power consumption breakdown of cloud data centre by considering power consumption at CPU level. However, the other network equipment also consumes a major amount of power during their operations. It requires balancing of workload among available network devices to minimise power consumption of cloud data centre.

Liu et al. [21] proposed a distributed flow schedule based method for optimising power consumption of network devices installed in cloud data center. However, this method LAX in considering communication between source and destinations and their computation.

Shang et al. [22] suggested a distributed green routing based method. There proposed method considers temperature and communication and computation happening at cloud data center for making workload balancing decisions in the cloud data center.

Esfandiarpoor et al. [23] suggested a power aware method for consolidating virtual machine in a large scale cloud data center consisting of multiple physical and virtual machines. They considered structural features of cloud data center including network topology and racks etc. They are also considered cooling systems and network structure of cloud data center while consolidating virtual machines and ensuring fulfill of service level agreement between cloud users and service provider.

Beloglazov et al. [24] suggested many threshold values in cloud computing environment for consolidation process based upon statistical features such as median absolute deviation, random strategy, and maximum correlation method. They proposed these heuristics that are adaptable to variations in workload and accordingly their threshold value get changed with time. The authors developed a heuristic-based approach for optimising cloud data center's power consumption. Their approach mainly targets historical values of CPU utilisation for finding upper threshold values to detect over utilized physical machines in cloud data center. Their solution attempt to collect the information and determine appropriate physical machine for consolidating virtual machines of cloud data center.

Von et al. [25] suggested a known adaptive method using different interpolation methods such as polynomial, linear, nearest neighbor and Shepard interpolation methods for estimating cloud data center's power consumption. They proposed to use comparison power consumption for determining accuracy of the proposed approach.

Ruan and Chen [26] suggested a new for allocating virtual machines in cloud data center based upon ratio of performance and power consumption of multiple physical machines. The authors focused on ensuring physical machine working at optimal level of performance and achieving maximum throughput by consuming minimum power. Using the computed ratio, they analyzed physical machine features for making threshold-based decisions. Their approach has resulted in fewer active physical machines and, hence, reduced cloud data centers' power consumption. Experimentally validated a reduction of 60 % power consumption with the reduced Virtual Machine migrations compared to the existing energy efficient virtual machine allocation method.

Patel and Patel [27] suggested an approach for detecting underutilized physical machines and placing virtual machines based upon utilizing physical machines. There proposed approach has been defined based upon over all cloud data center's power consumption for deciding utilisation threshold values. The threshold values are for the used for detecting overloaded cloud data center's physical machines. The authors have validated their approach experimentally for efficient detection of over utilize physical machines, reducing the number of active physical machines and reducing power consumption.

Fard et al. (2017) [28] suggested a new technique for consolidating virtual machines that obtain the quality of service temperature balance. They focus on consolidating virtual machines with high performance physical machines than low performance physical machines. This policy enables less production of heat and less consumption of power. However, this approach resulted in sub optimal utilisation of cloud computing resources.

Farahnakian et al. (2015) [29] applied metaheuristic algorithms, ant Colony Optimization algorithm, for placing virtual machine on appropriate physical machine in cloud computing environment. Their approach is validated for finding near optimal solution that improves cloud data center's power consumption based upon objective functions of ant Colony Optimization algorithm.

Song et al. (2014) [30] designed an approach for allocating computing resources per cloud users' workload needs to optimize the number of active cloud data center's physical machines. In this approach, the authors assumed that all physical machines are of the same configuration in processing power, storage, and network bandwidth. They formulated the Virtual Machine allocation problem as Bin packing problem. They experimentally validated its performance regarding number of virtual machine migrations and workload balancing compared with existing methods in the field.

Zhou et al. (2015) [31] developed a method for deploying virtual machines in a cloud computing environment based on threshold values for conserving power off the data center. They classified cloud data center's physical machines into four types: physical machines with the nominal workload, physical machines with appropriate workload, and machines with middle-level workload and heavy workload. Their approach performs consolidation of physical machines with an appropriate workload that can reduce the power consumption of data center. Virtual machines running on physical machines with nominal workload and heavy workload are shifted to another physical machine with appropriate workload. Experimental results obtained in their experiments demonstrate a considerable reduction in cloud data center's power consumption.

Li et al. (2018) [32] solved the multi-objective problem of virtual machine consolidation by optimising service quality, e power consumption and computing resource uses. They performed an online consolidation of virtual machines by shifting them two other machines while meeting physical machines' overloading constraints using the ant Colony Optimization algorithm. The experimental results have shown that there is a significant quantity of reduction in power consumption using their approach in comparison to the existing approaches.

3. The proposed system model

Most research attempted to determine underutilized cloud data center's physical machines have minimum utilisation. They focused on shifting the running virtual machines from underutilized physical machines iteratively. Generally, researchers proposed to create a physical machine's list in reverse order of utilization. It is followed by shifting all virtual machines from the least utilized physical machine so that the target physical machine can handle the workload without overloading. This process is iterated until all virtual machines are placed on other physical machines. These proposals help to achieve improvement in the utilisation of physical machines. However, most researchers ignored considering the total utilisation of all physical machines are available in the cloud data center while determining underutilized physical machines and finding new physical machines for placement of the selected virtual machines. Configuration of the total utilisation of all cloud data center's physical machines can enable efficient and accurate determination of underutilized physical machines and new physical machines for placing migrating virtual machines.

In order to address the challenges of state of art approaches, this work designs a practical and accurate approach for determining underutilized cloud data center's physical machines based on utilisation thresholds. It considers the utilisation of all physical machines available in the cloud data center. It also helps find the physical machines that must be vacated still supporting the cloud data center's current workload.

So, the proposed approach involves processing two types of physical machines: machines with utilisation below the lower threshold values and machines with utilisation equal to or above the lower threshold value. Therefore, the proposed approach attempts to migrate virtual machines running on physical machines below threshold utilisation values to two physical machines running equal to a lower utilisation threshold value in such a way that there is negligible impact on the power consumption of the cloud data center. The detailed algorithm of the proposed power saving under the loaded physical machine detection method is presented in the following subsection.

TABLE 1: POWER CONSUMPTION OF HOSTS AT DIFFERENT LOADS

		Host			
		HP ProLiant G9	HP ProLiant G7	HP ProLiant G5	HP ProLiant G4
Power consumption (in Watts) at different load on hosts	0%	83.7	55.6	93.7	86
	10%	101	95.4	97	89.4
	20%	118	107	101	92.6
	30%	133	115	105	96
	40%	145	124	110	99.5
	50%	162	133	116	102
	60%	188	142	121	106
	70%	218	155	125	108
	80%	248	173	129	121
	90%	276	192	133	114
	100%	83.7	216	135	117

3.1 The Proposed Power Saving Underloaded Physical Machine Detection (UnPMD) Method

Algorithm 1 outlines the algorithm for the proposed power saving under the loaded physical machine detection method. This method involves maintaining two lists for physical machines that can be used for finding under loaded and finding new machines for placement of virtual machines. For this purpose, it detects overloaded physical machines and switched off physical machines and add them into these two lists. Further, it computes the number of physical machines that must be vacated. It is computed based upon the cloud data center's current workload and lower utilisation threshold values. The computed sum of physical machines to be vacated is for the utilized in finding under loaded physical machines with utilisation less than a lower utilisation threshold value. It is followed by selecting a minimum to utilize the physical machine to place the migrating virtual

TABLE 2: CHARACTERISTICS OF THE VM TYPES

	VM Type			
	CPU (MIPS)	**#Of Cores**	**RAM**	**Bandwidth(Mbps)**
Type 1	500	1000	2000	2500
Type 2	613	1740	1740	870
Type 3	100	100	100	100
Type 4	2500	2500	2500	2500

machines from detected the under loaded physical machine. While determining the physical machine to place the migrating virtual machine, we considered the available computing resources of the physical machine along with its utilisation so that the machine does not get over utilized after placing the virtual machine.

The proposed approach also considers the increase in power utilisation due to the placement of a new virtual machine. Finally, it ensures that all virtual machines of underutilized physical machines are successfully placed on an appropriate physical machine. Minimum impact on power consumption and leads to the least active physical machines.

4. Performmance Evaluation

This section provides experimental setup, metrics to measure the performance of the proposed method and the existing methods and simulation results followed by their analysis in the following subsections.

4.1 Experimintal Setup

The proposed power saving underutilized physical machine detection (UnPMD) method is implemented and validated using cloudSim simulation software. It mainly focuses on reducing the power consumption of data center by identifying under loaded physical machines based upon workload of data center and shifting virtual machines from under loaded physical machines to other. The vacated virtual machines are turned off to power saving mode that leads to overall power consumption of data center. Also, fine number of physical machines to be vacated and find appropriate physical machines for shifting the migrating virtual machine without overloading them while meeting service level agreement.

This work assumes the cloud computing environment as infrastructure as a service model that enables scalable resources to cloud customers [33]. We analyse the performance of the proposed (UnPMD) method using virtualised cloud data center using cloudSim software [34] for repeatable and large scale experiments considering configuration of the system and cost parameters for each set of experiments. This set of experiments assumes 800 physical machines with different Cloud Computing capacities simulated in cloudsim software. We considered two types of physical machines, each of 400 units, with different configuration, HP Pro Liant M L 110 G4 powered by Intel Xeon3040 clocked at 1860 MHz and HP Pro Liant ML 110 G5 powered by Intel Xeon3075 clocked at 2660 MHzs. Each physical machine considered in this work consists of 4GB ram. We measured the processing power / frequency of the physical machine in terms of millions instruction per second (MIPS). Table 1 shows the physical machine characteristics regarding power consumption used in this work. We assume available network bandwidth of 1 GB/s.

Different types of virtual machines are depicted in Table 2. We assume standard Amazon ec2 [35] virtual machine instances installed on physical machine. Therefore 4 categories of virtual machines have been availed for handling workload of the cloud users. Experiments conducted in this work are based upon workload traces obtained from real word systems. It helps to to achieve acceptable experimental results using simulation based experiments. We used experimental workload traces conducted by planet lab in their project called CoMon [36]. The simulation have been performed on 1516 data instances collected from different types of virtual machines present in cloud data center 5 minute interval. The workload based upon its characteristics are allocated to different virtual machines.

4.2 Performance metrics

Several metrics have been proposed to measure the newly proposed approaches and comprehensively compare them with the existing approaches in the cloud computing environment, considering different factors. The most important metrics used for measuring performance includes power consumption, service level agreement violations, number of virtual machine migrations, number of active physical machines in cloud data center, number of physical machines shut down for conserving power etc. In this work, we use the 4 metrics, power consumption, number of service level agreement violations, number of virtual machine migrations and number of physical machine shutdown as explained below.

4.3 Power Consumption

Power consumption is considered an essential metric for measuring the performance of cloud computing policy as it is being consumed by all computational units of cloud data center such as processing units, storage disks, cooling systems, power supplies etc. [37, 38]. It has been observed that power consumption of a physical machine has a linear Association with CPU usage. More the CPU is used more power is consumed. Many researchers believe that idle physical machines consume 70% of the power used while it is fully utilized. Therefore, best work computes power consumption as a function of CPU usage as per equation 1.

$$PC(usage) = r * PC_{max} + (1 - r) * PC_{max} * usage \quad (1)$$

Where, PC_{max} is the maximum consumed power by a physical machine when fully utilized. Usage indicates the current utilization of physical machines. R gives the ratio of physical machine‘s power consumption when it is in an idle state.
However, CPU uses of a physical machine varies with the time due to change in work load of cloud requests. So, the aggregate amount of power consumed is computed as per equation over a period.

$$PC(total) = \int PC\big(usage(t)\big)dt \quad (2)$$

4.4 SLA Violation

In cloud computing environment, cloud user and cloud service providers agree on certain quality features of the cloud services such as quality of service in an agreement called service level agreement [4]. Thereafter, service level agreement metrics have been defined and computed that approximate the performance of service provided by the cloud service provider. The power saving Cloud Computing policies consider service level agreement violations. Service level agreement violations are computed as per equation 3

$$SLAV = \frac{Service_{requested} - Service_{allocated}}{Service_{requested}} \tag{3}$$

Where, $Service_{requested}$ Represents the MIPS of the virtual machine requested by the cloud user and $Service_{requested}$ gives the MIPS allocated to the cloud user.

Number of VM Migrations

The virtual machine consolidation process involves shifting virtual machines from overloaded physical machines to others. It is desired that shifting of virtual machine should be performed with minimum down time. However, transferring virtual machines has a direct impact on performance of the application running on different virtual machines [39]. It has been observed that transferring of virtual machine also has an interference with virtual machines at

ALGORITHM 1: THE PROPOSED POWER SAVING UNDERUTILIZED PHYSICAL MACHINE DETECTION (UNPMD) METHOD

```
          Input: PML = List of Active physical machine
                 VML = list of virtual machines on each physical machine
                 ThsUp = upper threshold of physical machine utilization
                 ThsLow = lower threshold of physical machine utilization
                 PML_UnU = physical machine list for detecting under utilized machines
                 PML_NP = physical machine list for placing the selected virtual machines
                 PML_Sorted = Sorted list of active physical machines in order of utilization
                 Total_WL = Total workload of cloud data center
                 PUk = utilization of kth physical machine
          Output: VM_allocationL
Step 1.   Start
Step 2.   Initialize PML
Step 3.   Initial VML
Step 4.   Set Total_WL = 0
Step 5.   For loop pm in PM_list do
Step 6.     | Total_WL = Total_WL + PUpm
Step 7.   End for
Step 8.   PM_total ← sizeof (PML)
Step 9.   PMmax ← PM_total – (Total_WL / ThsUp) // maximum physical machines to be vacated
Step 10.  ThsLow ← PML_sorted [PMmax]
Step 11.  For loop pm in [1 to PMmax] do
Step 12.    | PML_NP.add (pm)
Step 13.  End for
Step 14.  PM_minU ← 1
Step 15.  Repeat loop
Step 16.    | If (PM_total = = sizeof (PML_UnU) then
Step 17.    |   | Break
Step 18.    | End If
Step 19.    | If (PM_minU ∉ PML_UnU) then
Step 20.    |   | PML_UnU.add (PM_minU)
Step 21.    | End if
Step 22.    | If (PM_minU ∉ PML_NP) then
Step 23.    |   | PML_NP.add (PM_minU)
Step 24.    | End if
Step 25.    | For loop vms in PM_minU do
Step 26.    |   | vmu ← VM_util (vms)
Step 27.    |   | minPD ← MAX
Step 28.    |   | PM_A ← NULL
Step 29.    |   | For loop pm in PML do
Step 30.    |   |   | If pm ∉ PML_NP then
Step 31.    |   |   |   | Continue
Step 32.    |   |   | End If
Step 33.    |   |   | If pm is equipped with resources then
Step 34.    |   |   |   | If PUpm > ThsUP then
Step 35.    |   |   |   |   | Continue
Step 36.    |   |   |   | End if
Step 37.    |   |   |   | If (Powerpm (after_placement) - Powerpm (before_placement)) < minPD then
Step 38.    |   |   |   |   | minPD ← Powerpm (after_placement) - Powerpm (before_placement)
Step 39.    |   |   |   |   | PM_A ← pm
Step 40.    |   |   |   | End If
Step 41.    |   |   | End if
Step 42.    |   | End for loop
Step 43.    |   | If PM_A ≠ NULL then
Step 44.    |   |   | If PML_A ∉ PM_UnU then
Step 45.    |   |   |   | PM_UnU.add (PML_A)
Step 46.    |   |   | End if
Step 47.    |   |   | Migration_L.add (vms, PM_A)
Step 48.    |   | Else
Step 49.    |   |   | Break
Step 50.    |   | End If
Step 51.    | End for loop
Step 52.    | PM_minU = PM_minU + 1
Step 53.  End repeat loop
```

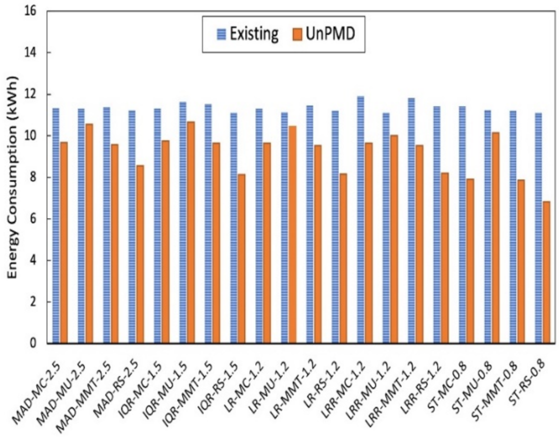

Figure 1. Comparative analysis of power consumption

TABLE 3. IMPROVEMENT IN POWER CONSUMPTION

Policy with Safety Parameter	Energy Consumption (kWh)		%age Improvement
	Existing	UnPMD	
MAD-MC-2.5	11.33	9.67	-17.2%
MAD-MU-2.5	11.31	10.56	-7.1%
MAD-MMT-2.5	11.38	9.58	-18.8%
MAD-RS-2.5	11.22	8.56	-31.1%
IQR-MC-1.5	11.32	9.76	-16.0%
IQR-MU-1.5	11.64	10.66	-9.2%
IQR-MMT-1.5	11.53	9.65	-19.5%
IQR-RS-1.5	11.11	8.13	-36.7%
LR-MC-1.2	11.31	9.65	-17.2%
LR-MU-1.2	11.14	10.45	-6.6%
LR-MMT-1.2	11.45	9.53	-20.1%
LR-RS-1.2	11.21	8.16	-37.4%
LRR-MC-1.2	11.91	9.65	-23.4%
LRR-MU-1.2	11.11	10.01	-11.0%
LRR-MMT-1.2	11.81	9.53	-23.9%
LRR-RS-1.2	11.42	8.2	-39.3%
ST-MC-0.8	11.42	7.91	-44.4%
ST-MU-0.8	11.23	10.14	-10.7%
ST-MMT-0.8	11.21	7.86	-42.6%
ST-RS-0.8	11.11	6.83	-62.7%

source and destination machine [40]. Therefore, number of virtual machine migrations has a direct Association with performance of cloud computing policies. It is desired that and effective Cloud Computing policy should result minimum number of virtual machine migrations.

In this work, we computed number of virtual machine migrations as a performance metric. We compute the cost of virtual machine migration as proposed in [41]. The virtual machine migration causes performance degradation that can be computed using equation 4 and 5.

$$P_{deg} = 0.1 * \int_0^{timej} CPU_j(t)dt \tag{4}$$

$$Timej = \frac{Memj}{Bwidthj} \tag{5}$$

Where CPUj gives the cpu utilization of jth virtual machine over a period of timej. Mj is the memory consumed by jth virtual machine and Bwidthj denotes bandwidth utilized by jth virtual machine

5. Simulation Results and Analysis

We conducted a comprehensive set of experiments for evaluating the performance of the proposed underutilized physical machine detection method and compared results with the existing techniques, maximum correlation, minimum migration time, maximum utilisation and random strategy for cloud computing under different scenarios of Interquartile Range (IQR), Median Absolute Deviation (MAD), Local Robust Regression (LRR), Static Threshold (ST) and Local Regression (LR) [42].

The performances were measured in terms of power consumption, number of virtual machine migrations, number of service level agreement violations, and number of physical machines shutdowns using the proposed (UnPMD) method and the existing methods. Following subsection present search result obtained in this experiment and provide their discussion.

5.1 Power Consumption Based Analysis

Figure 1 presents a comprehensive comparison of power consumption of the existing and the proposed (UnPMD) method for detecting underutilized physical machines. It can be noted that the proposed (UnPMD) method resulted in reducing the power consumption under different scenarios over the existing methods.

The percentage improvement in power consumption of the proposed (UnPMD) method over the existing methods under different scenarios is presented in table 3.

Table 3 clearly indicates a significant improvement in power consumption using the proposed (UnPMD) method.

Under the MAD scenario, the percentage improvement in power consumption using the proposed (UnPMD) method is 31%, 18% over the random strategy method and the minimum migration time method. Under the IQR scenario, 36% improvement in

power consumption has been reported over the random strategy method. In contrast, 37% improvement have been achieved under the LR scenario. Whereas 39 and 62% improvement in power consumption have been received over random strategy in cloud computing under LRR and ST scenarios. A similar improvement in the performance in terms of cloud data center power consumption has been received over different methods under different scenarios as depicted in table 3.

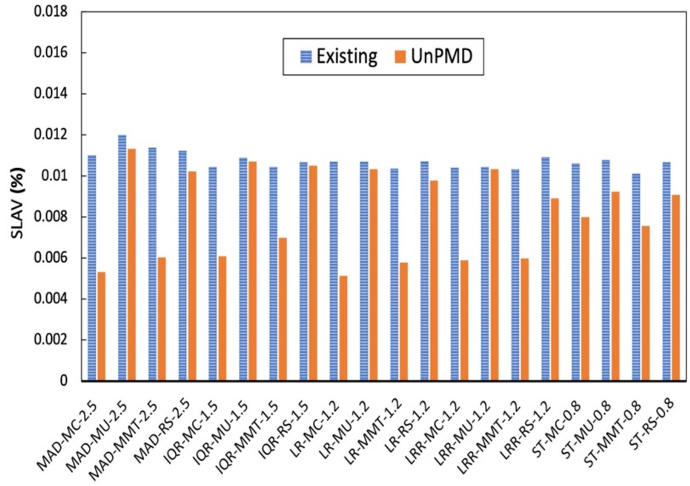

Figure 2. Comparative analysis of Service level agreement violations

TABLE 4. IMPROVEMENT IN SERVICE LEVEL AGREEMENT VIOLATIONS

Policy with Safety Parameter	SLAV (%)		%age Improvement
	Existing	UnPMD	
MAD-MC-2.5	0.011	0.00531	-107.2%
MAD-MU-2.5	0.012	0.01131	-6.1%
MAD-MMT-2.5	0.01138	0.00602	-89.0%
MAD-RS-2.5	0.01122	0.01022	-9.8%
IQR-MC-1.5	0.01044	0.00609	-71.4%
IQR-MU-1.5	0.01088	0.0107	-1.7%
IQR-MMT-1.5	0.01043	0.00698	-49.4%
IQR-RS-1.5	0.01067	0.0105	-1.6%
LR-MC-1.2	0.0107	0.00512	-109.0%
LR-MU-1.2	0.01069	0.01033	-3.5%
LR-MMT-1.2	0.01035	0.00577	-79.4%
LR-RS-1.2	0.01071	0.00977	-9.6%
LRR-MC-1.2	0.0104	0.00588	-76.9%
LRR-MU-1.2	0.01044	0.01033	-1.1%
LRR-MMT-1.2	0.01032	0.00598	-72.6%
LRR-RS-1.2	0.01091	0.0089	-22.6%
ST-MC-0.8	0.01061	0.00799	-32.8%
ST-MU-0.8	0.01077	0.00923	-16.7%
ST-MMT-0.8	0.01011	0.00756	-33.7%
ST-RS-0.8	0.01067	0.009077	-17.5%

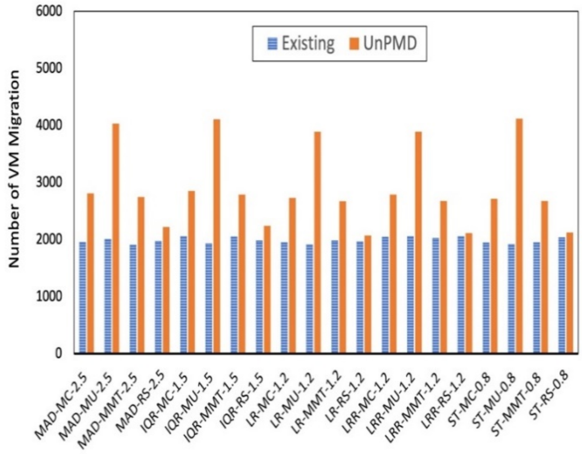

Figure 3. Comparative analysis of Virtual Machine migrations

5.2 The Service Level Agreement Violation Based Result Analysis

Figure 2 compares the number of service level agreement violations due to the proposed (UnPMD) method and the existing methods of random strategy, maximum correlation, maximum utilisation and minimum migration time in combination with other techniques of MAD, IQR, ST, LR and LRR. It can be observed from figure to that there is a significant reduction in number of service level agreement violation using

TABLE 5. IMPROVEMENT IN VIRTUAL MACHINE MIGRATIONS

Policy with Safety Parameter	Number of VM Migration		%age Improvement
	Existing	UnPMD	
MAD-MC-2.5	1962	2811	30.2%
MAD-MU-2.5	2011	4034	50.1%
MAD-MMT-2.5	1912	2745	30.3%
MAD-RS-2.5	1975	2220	11.0%
IQR-MC-1.5	2064	2856	27.7%
IQR-MU-1.5	1932	4111	53.0%
IQR-MMT-1.5	2056	2788	26.3%
IQR-RS-1.5	1988	2244	11.4%
LR-MC-1.2	1955	2734	28.5%
LR-MU-1.2	1919	3889	50.7%
LR-MMT-1.2	1989	2675	25.6%
LR-RS-1.2	1971	2071	4.8%
LRR-MC-1.2	2051	2788	26.4%
LRR-MU-1.2	2061	3891	47.0%
LRR-MMT-1.2	2029	2678	24.2%
LRR-RS-1.2	2061	2113	2.5%
ST-MC-0.8	1951	2718	28.2%
ST-MU-0.8	1921	4121	53.4%
ST-MMT-0.8	1956	2677	26.9%
ST-RS-0.8	2039	2123	4.0%

the proposed underutilized physical machine detection method in comparison to the existing method. Table 4 shows the percentage improvement of the proposed (UnPMD) method over existing methods. It can be concluded from table 4 that improvement in number of service level agreement violations of 109 percent can be achieved over maximum correlation method

under scenario of LR, followed by 89% improvement over minimum migration time method under the MAD scenario. A considerable performance has also been achieved over all the

testing methods used in this experiment under different scenarios as depicted in table 4. testing methods used in this experiment under different scenarios as depicted in table 4.

5.3 Virtual Machine Migrations Based Result Analysis

Figure 3 compares the number of virtual machine migrations using the proposed method and existing methods. It can be observed from figure 3 that the proposed method has increased the number of virtual migrations because of vacating more physical machines while meeting service level agreements. It is evident from figures 2, 3 and Table 5 that in spite of the rise in the Virtual Machine migrations, the total number of service level agreement violations has decreased. So, increase in the number of virtual machine migrations does not affect the number of service level agreement violations. Finally, it reduced power consumption by optimising the number of active physical machines supporting the current workload of the cloud data centre.

5.4 Number of Physical Machine Shutdowns based Result Analysis

A comprehensive comparison of the number of physical machine shutdowns using the proposed (UnPMD) method and the existing methods is presented in figure 4. It can be observed from figure for that a significant large number of physical machines can be shut down by shifting their virtual machines to other machines. More number of physical machine shutdowns resulted in reducing the cloud data center's power consumption considerably.

Table 6 presents the percentage improvement in shutting down the physical Machines by using the proposed (UnPMD) method over the existing methods. It can be observed from table 6 that approximately 53% more physical machines can be shut down over maximum utilisation method under LR scenario and 54% over maximum utilisation method under static threshold scenario. This considerable amount of increase in physical machine shutdowns have a significant effect on cloud data center's power consumption and hence return on investment.

Experimental results demonstrate the superiority of the proposal underutilized physical machine detection method over the existing methods of random strategy, minimum migration time, maximum correlation, and maximum utilisation method under the scenario of MAD, IQR, LR, LRR and ST.

6. conclusion

In cloud computing, the Virtual Machine consolidation process has a significant impact on data center power consumption. An effective Virtual Machine consolidation method involves detecting under utilized and over utilized physical machines, migrating correct virtual machines from over utilized ones to others, shifting all virtual machines of underutilized physical machines to others, and turning off the vacated physical machines to a power saving mode.

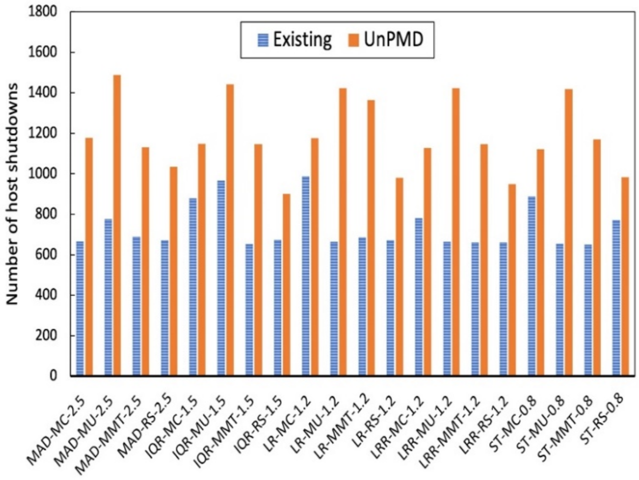

Figure 4. Comparative analysis of the number of physical machine shutdowns

TABLE 6. IMPROVEMENT IN PHYSICAL MACHINE SHUTDOWNS

Policy with Safety Parameter	Number of physical machine shutdowns		%age Improvement
	Existing	UnPMD	
MAD-MC-2.5	666	1177	43.4%
MAD-MU-2.5	776	1489	47.9%
MAD-MMT-2.5	689	1130	39.0%
MAD-RS-2.5	671	1034	35.1%
IQR-MC-1.5	878	1148	23.5%
IQR-MU-1.5	966	1441	33.0%
IQR-MMT-1.5	653	1146	43.0%
IQR-RS-1.5	673	901	25.3%
LR-MC-1.2	988	1176	16.0%
LR-MU-1.2	664	1423	53.3%
LR-MMT-1.2	685	1364	49.8%
LR-RS-1.2	672	979	31.4%
LRR-MC-1.2	781	1128	30.8%
LRR-MU-1.2	664	1423	53.3%
LRR-MMT-1.2	661	1146	42.3%
LRR-RS-1.2	661	949	30.3%
ST-MC-0.8	888	1121	20.8%
ST-MU-0.8	654	1419	53.9%
ST-MMT-0.8	650	1170	44.4%
ST-RS-0.8	771	983	21.6%

Detection of under utilized physical machines is one of tasks in virtual machine consolidation process. The existing methods mainly focus on setting up static threshold values based on current utilisation. However, they ignored considering the overall the cloud data center's workload to determine under utilized physical machines.

To tackle issues with the existing method, this work designs a new method for determine meaning underutilized physical machines called UnPMD based upon data center workload. It also determines the number of physical machines to be vacated and ensures that virtual machine migration does not make the target machine as overloading machine. Therefore, it maintains the active physical machines that are not over utilized. It helps in placing the migrating virtual machines to an appropriate physical machine.

The proposed (UnPMD) method has been experimentally validated for detecting underutilized physical machines and maintaining minimum power consumption compared to the existing machines. Experimental results indicate that the proposed (UnPMD) method can help achieve a considerable change in power consumption, number of service level agreement violations, and Virtual Machine migrations. The proposed (UnPMD) method can improve up to 62% improvement in power consumption over random strategy under static threshold scenarios. Similarly, it has resulted in a 109% decrease in the number of service level agreement violations under the MAD scenario. The proposed (UnPMD) method has increased the number of physical machine shutdowns significantly. The experimental result can observe 54% increase in number of physical machine shutdowns in maximum utilisation method under static threshold scenario.

References

[1] Moghaddam, Seyedhamid Mashhadi, Michael O'Sullivan, Cameron Walker, Sareh Fotuhi Piraghaj, and Charles Peter Unsworth. "Embedding individualized machine learning prediction models for energy efficient VM consolidation within Cloud data centers." Future Generation Computer Systems 106 (2020) : 221-233.

[2] Yadav, Rahul, Weizhe Zhang, Keqin Li, Chuanyi Liu, and Asif Ali Laghari. "Managing overloaded hosts for energy-efficiency in cloud data centers." Cluster Computing 24, no. 3 (2021) : 2001-2015.

[3] Masdari, M., Khezri, H. : Efficient VM migrations using forecasting techniques in cloud computing : a comprehensive review. Clust. Comput. (2020). https : //doi.org/10.1007/s10586-019- 03032-x

[4] Yadav, R., Zhang, W., Kaiwartya, O., Singh, P.R., Elgendy, I.A., Tian, Y.C. : Adaptive energy-aware algorithms for minimizing energy consumption and SLA violation in cloud computing. IEEE Access 6, 55923 – 55936 (2018)

[5] Luo, J.P., Li, X., Chen, M.R. : Hybrid shuffled frog leaping algorithm for energy-efficient dynamic consolidation of virtual machines in cloud data centers. Expert Syst. Appl. 41 (13), 5804 – 5816 (2014)

[6] Saadi, Youssef, and Said El Kafhali. "Energy-efficient strategy for virtual machine consolidation in cloud environment." Soft Computing 24, no. 19 (2020) : 14845-14859.

[7] El Kafhali S, Salah K (2018a) Performance analysis of multi-Core VMs hosting cloud SaaS applications. Comput Stand Interfaces 55 : 126 – 135

[8] Asad Z, Chaudhry MAR (2017) A two-way street : Green big data processing for a greener smart grid. IEEE Syst J 11 (2) : 784 – 795

[9] Rinco´n D, Agustı ´-Torra A, Botero JF, Raspall F, Remondo D, Hesselbach X, Beck MT, de Meer H, Niedermeier F, Giuliani G (2013) A novel collaboration paradigm for reducing energy consumption and carbon dioxide emissions in data centres. Comput J 56 (12) : 1518 – 1536

[10] Barroso LA, Ho¨lzle U (2007) The case for energy-proportional computing. Computer 40 (12) : 33 – 37

[11] El Kafhali S, Salah K (2017) Stochastic modelling and analysis of cloud computing data center. In : Proceedings of 20th conference on innovations in clouds, internet and networks, IEEE,pp 122 – 126

[12] Khoshkholghi MA, Derahman MN, Abdullah A, Subramaniam S, Othman M (2017) Energy-efficient algorithms for dynamic virtual machine consolidation in cloud data centers. IEEE Access 5 : 10709 – 10722

[13] Younge AJ, Von Laszewski G, Wang L, Lopez-Alarcon S, Carithers W (2010) Efficient resource management for cloud computing environments. In : Proceedings of the international conference on green computing, IEEE, pp 357 – 364

[14] Nathuji, R., Schwan, K. : Virtual power : coordinated power management in virtualized enterprise systems. ACM SIGOPS Oper. Syst. Rev. 41 (6), 265 – 278 (2007)

[15] Belgacem, A., Beghdad-Bey, K., Nacer, H., Bouznad, S. : Efficient dynamic resource allocation method for cloud computing environment. Clust. Comput. (2020). https : //doi.org/10.1007/s10586-020-03053-x

[16] Zhu, X., Young, D., Watson, B.J., Wang, Z., Rolia, J., Singhal, S., Mckee, B., Hyser, C., Gmach, D., Gardner, R. : 1000 islands : integrated capacity and workload management for the next generation data center. In : International Conference on Autonomic Computing, pp. 172 – 181 (2008)

[17] Moghaddam, M.J., Esmaeilzadeh, A., Ghavipour, M., Zadeh, A.K. : Linear power modeling for cloud data centers : taxonomy, locally corrected linear regression, simulation framework and evaluation. Clust. Comput. 1, 1 – 10 (2019)

[18] Guo, Y., Gong, Y., Fang, Y., Khargonekar, P.P., Geng, X. : Energy and network aware workload management for sustainable data centers with thermal storage. IEEE Trans. Parallel Distrib. Syst. 25 (8), 2030 – 2042 (2014)

[19] Lange, C., Kosiankowski, D., Weidmann, R., Gladisch, A. : Energy consumption of telecommunication networks and related improvement options. IEEE J. Sel. Top. Quantum Electron. 17 (2), 285 – 295 (2011)

[20] Fan, X., Weber, W.D., Barroso, L.A. : Power provisioning for a warehouse-sized computer. In : ACM SIGARCH Computer Architecture News, vol. 35, pp. 13 – 23. ACM, New York (2007)

[21] Liu, R., Gu, H., Yu, X., Nian, X. : Distributed flow scheduling in energy aware data center networks. IEEE Commun. Lett. 17 (4), 801 – 804 (2013)

[22] Shang, Y., Li, D., Xu, M. : Energy-aware routing in data center network. In : Proceedings of the First ACM SIGCOMM Workshop on Green Networking, pp. 1 – 8. ACM, New York (2010)

[23] Esfandiarpoor, S., Pahlavan, A., & Goudarzi, M. (2015). Structure aware online virtual machine consolidation for data center energy improvement in cloud computing. Computers & Electrical Engineering, 42, 74 – 89.

[24] Beloglazov, A., Abawajy, J., & Buyya, R. (2012). Energy aware resource allocation heuristics for efficient management of data centers for cloud computing. Future Generation Computer Systems, 28 (5), 755 – 768.

[25] von Kistowski, J., & Kounev, S. (2016). Univariate interpolation based modeling of power and performance. In Proceedings of the 9th

EAI international conference on performance evaluation methodologies and tools (pp. 212 – 215). ICST (Institute for Computer Sciences, Social-Informatics and Telecommunications Engineering).

[26] Ruan X, Chen H (2015) Performance-to-power ratio aware virtual machine (VM) allocation in energy efficient clouds. In : Proceedings of the international conference on cluster computing, IEEE, pp 264 – 273

[27] Patel N, Patel H (2017) Energy efficient strategy for placement of virtual machines selected from underloaded servers in compute Cloud. J King Saud Univ Comput Inf Sci. https : //doi.org/10. 1016/j.jksuci.2017.11.003

[28] Fard SY, Ahmadi MR, Adabi S (2017) A dynamic VM consolidation technique for QoS and energy consumption in cloud environment. J Supercomput 73 (10) : 4347 – 4368

[29] Farahnakian F, Ashraf A, Pahikkala T, Liljeberg P, Plosila J, Porres I, Tenhunen H (2015) Using ant colony system to consolidate VMs for green cloud computing. IEEE Trans Serv Comput 8 (2) : 187 – 198

[30] Song W, Xiao Z, Chen Q, Luo H (2014) Adaptive resource provisioning for the cloud using online bin packing. IEEE Trans Comput 63 (11) : 2647 – 2660

[31] Zhou Z, Hu ZG, Song T, Yu JY (2015) A novel virtual machine deployment algorithm with energy efficiency in cloud computing. J Central South Univ 22 (3) : 974 – 983

[32] Li Z, Yan C, Yu L, Yu X (2018) Energy aware and multi resource overload probability constraint-based virtual machine dynamic consolidation method. Future Gener Comput Syst 80 : 139 – 156

[33] Patel, Nimisha, and Hiren Patel. "Energy efficient strategy for placement of virtual machines selected from underloaded servers in compute Cloud." Journal of King Saud University-Computer and Information Sciences 32, no. 6 (2020) : 700-708.

[34] Calheiros, R.N., Ranjan, R., Beloglazov, A., Rose, C.A.F.D., Buyya, R., 2014. CloudSim : a toolkit for modeling and simulation of cloud computing environments and evaluation of resource provisioning algorithms. Software : Pract. Exp. 41 (1), 23 – 50.

[35] Verma A, Dasgupta G, Nayak TK, De P, Kothari R "Server workload analysis for power minimization using consolidation". In : Proceedings of the 2009 conference on USENIX annual technical conference. San Diego, USA, 2009

[36] Moges, Fikru Feleke, and Surafel Lemma Abebe. "Energy-aware VM placement algorithms for the OpenStack Neat consolidation framework." Journal of Cloud Computing 8, no. 1 (2019) : 1-14.

[37] Nadimi, Reza, and Reza Ahmadi Khatir. "VM Consolidation by using Selection and Placement of VMs in Cloud Datacenters." Caspian Journal of Mathematical Sciences (CJMS) peer 6, no. 2 (2017) : 120-130.

[38] Minas, L., and Ellison, B. (2009). Energy Efficiency for Information Technology : How to Reduce Power Consumption in Servers and Data Centers. Intel Press.

[39] Fu, Xiong, and Chen Zhou. "Virtual machine selection and placement for dynamic consolidation in Cloud computing environment." Frontiers of Computer Science 9, no. 2 (2015) : 322-330.

[40] Xu F, Liu F, Liu L, Jin H, Li B. Iaware : making live migration of virtual machines interference-aware in the cloud. IEEE Transactions on Computers, 2014, 63 (12) : 3012 – 3025

[41] Song Y, Wang H, Li Y, Feng B, Sun Y. Multi-tiered on demand resource scheduling for VM-based data center. In : Proceedings of the 9th IEEE/ACM International Symposium on Cluster Computing and the Grid. 2009, 148 – 155

[42] Zhang, Q., Zhani, M.F., Zhang, S., Zhu, Q., Boutaba, R., Hellerstein, J.L. : Dynamic energy aware capacity provisioning for cloud computing environments. In : Proceedings of the 9th International Conference on Autonomic Computing, pp. 145 - 154. ACM (2012)