Server consolidation: A technique to enhance cloud data center power efficiency and overall cost of ownership

Mueen Uddin¹, Mohammed Hamdi², Abdullah Alghamdi², Mesfer Alrizq², Mohammad Sulleman Memon³, Maha Abdelhaq⁴ and Raed Alsaqour⁵

Abstract
Cloud computing is a well-known technology that provides flexible, efficient, and cost-effective information technology solutions for multinationals to offer improved and enhanced quality of business services to end-users. The cloud computing paradigm is instigated from grid and parallel computing models as it uses virtualization, server consolidation, utility computing, and other computing technologies and models for providing better information technology solutions for large-scale computational data centers. The recent intensifying computational demands from multinationals enterprises have motivated the magnification for large complicated cloud data centers to handle business, monetary, Internet, and commercial applications of different enterprises. A cloud data center encompasses thousands of millions of physical server machines arranged in racks along with network, storage, and other equipment that entails an extensive amount of power to process different processes and amenities required by business firms to run their business applications. This data center infrastructure leads to different challenges like enormous power consumption, underutilization of installed equipment especially physical server machines, CO₂ emission causing global warming, and so on. In this article, we highlight the data center issues in the context of Pakistan where the data center industry is facing huge power deficits and shortcomings to fulfill the power demands to provide data and operational services to business enterprises. The research investigates these challenges and provides solutions to reduce the number of installed physical server machines and their related device equipment. In this article, we proposed server consolidation technique to increase the utilization of already existing server machines and their workloads by migrating them to virtual server machines to implement green energy-efficient cloud data centers. To achieve this objective, we also introduced a novel Virtualized Task Scheduling Algorithm to manage and properly distribute the physical server machine workloads onto virtual server machines. The results are generated from a case study performed in Pakistan where the proposed server consolidation technique and virtualized task scheduling algorithm are applied on a tier-level data center. The results obtained from the case study demonstrate that there are annual power savings of 23,600 W and overall cost savings of US$78,362. The results also highlight that the utilization ratio of already existing physical server machines has increased to 30% compared to 10%, whereas the number of server machines has reduced to 50% contributing enormously toward huge power savings.

¹Faculty of Science, Universiti Brunei Darussalam, Bandar Seri Begawan, Brunei Darussalam
²College of Computer Science and Information Systems, Najran University, Najran, Saudi Arabia
³CSE Department, QUEST Nawabshah, Pakistan
⁴Department of Information Technology, College of Computer and Information Sciences, Princess Nourah bint Abdulrahman University, Riyadh, Saudi Arabia
⁵Department of Information Technology, College of Computing and Informatics, Saudi Electronic University, Riyadh, Saudi Arabia

Corresponding author:
Mueen Uddin, Faculty of Science, Universiti Brunei Darussalam, BE 1410 Bandar Seri Begawan, Brunei Darussalam.
Email: mueenmalik9516@gmail.com

Creative Commons CC BY: This article is distributed under the terms of the Creative Commons Attribution 4.0 License (https://creativecommons.org/licenses/by/4.0/) which permits any use, reproduction and distribution of the work without further permission provided the original work is attributed as specified on the SAGE and Open Access pages (https://us.sagepub.com/en-us/nam/open-access-at-sage).
Introduction

Human business activities involving gigantic information technology (IT) infrastructures are causing serious reparations to the environment and climate ecosystems causing severe problems toward the upsurging of global warming, climate control, CO₂ emissions, and environmental degradation. However, these business enterprises have substantial technologies, professional knowledge, creative thinking, and capacity and motivation to understand and achieve environmental sustainability. The environmental impact of IT under the banner of green IT and green cloud computing has emerged as one of the major research areas and is being discoursed by various academic, research, and commercial business entities including the cloud data center (CDC) industry to implement energy-efficient CDCs. CDCs are characterized as one of the key target areas in IT infrastructures where power consumption can easily be reduced and further diminished by applying various cloud-enabled energy-efficient techniques. These cloud infrastructures consume a high density of power results in generating a vast amount of greenhouse gases considered very harmful for environmental sustainability. Power efficiency policy planners and architects have highlighted that power consumption in CDC is increasing alarmingly and will reach a level where it becomes really hard for the cloud owners to sustain their businesses and run the operational capabilities of their businesses. The IT components of a CDC including physical server machines (PSMs) and their networking equipment are considered as major stakeholders, and they further require additional facility components such as cooling, heating, backup, and storage equipment which contribute profoundly toward electricity bills. According to ES-1 report published in 2016, in the United States alone, the CDCs have consumed approximately 75 billion kWh of power in 2015, that is almost 2% of entire power consumption in the whole United States. The report also highlights that the overall CDC power usage has hastily increased to about 4% from 2010 to 2014. Power usage is projected to continue further and will be further increased from 2014 to 2020. The report also forecasts that based on current trend estimations; the CDCs in the United States will consume roughly 74 billion kWh in 2020 as shown in Figure 1. According to IDC, the electronic data growth will expand from 4.5 ZB (zettabytes) in 2014 to almost 45 ZB by 2020. This huge data expansion will require more powerful CDCs to store and update data records.
According to a report published by Pakistan institute of development economics in the year 2018, highlights that the power shortfall in domestic and commercial sectors has become a hindrance for economic and commercial growth. The average shortfall in the commercial and domestic power sector is around 3000 to 4000 MWs. This power shortage leads to load-shedding and power outages across all business firms which costs billions of rupees of losses to Pakistan’s economy every year. The same report also shows that about 55% of the total population, especially in rural areas, is facing an electricity shortage of up to 8–12 h per day. While in the commercial cities, the ratio of electricity shortage ranges from 4 to 8 h per day. This critical and alarming situation in the power sector is considered one of the primary reasons in the country’s slow economic growth and development. The state of the energy sector (and the extension of the economy) is not beyond redemption as shown in Figure 2. This electricity gap is creating an enormous number of problems for the IT industry to run and provide their services especially CDCs, which are considered as the backbone of all major IT multinational companies. Many solutions are being proposed by different researchers to fill this gap especially in the context of green cloud computing, where the solutions are being discourses for having any significant impact on achieving competitive advantage in implementing energy-efficient CDCs. The research proposed in this article attempts to increase the utilization of already installed physical machines in CDCs to support IT industry to overcome some of the power-related issues in Pakistan. The study aims to maximize the performance of CDCs and reduce their power consumption. We proposed a server consolidation technique based on Virtualized Task Scheduling Algorithm to help CDCs in improving their server machine utilization from 10% to 30% and also reducing the number of PSMs from 500 to 96. In our proposed technique, server consolidation was applied on PSMs where different virtual instances were created for several virtual machines (VMs) installed on existing physical machines. This was done using virtualization technology which enables CDCs to create and execute multiple virtual applications onto the same physical server, thereby increasing its utilization ratio. Furthermore, the proposed virtualized task scheduling algorithm was used to encumbered each virtual server machine (VSM) with multiple workloads from PSMs based on their execution capability and total workload available on the PSM. This technique enables us to have less active PSMs being replaced by VSMs which means more power efficiency, fewer power costs, and less CO₂ emission.

**Problem background**

A CDC mainly consists of server machines used for performing data processing, data storage devices responsible for the storage of data and backups, network devices, and equipment responsible for enabling communication between different components of data center. Together with this, all equipment that processes, stores, and transmits digital information are called IT equipment. In a CDC, the majority of the power is consumed by this equipment, especially the server machines which do the majority of the processing for business operations. These server machines consume a lot of power, hence contributing heavily toward carbon emissions and global warming. According to a report, an average 300-W server’s electricity cost is around US$338 and CO₂ emission as much as 1300 kg during a year. The cost of cooling the equipment is not included in these statistics. The exponential growth in Google, Facebook, Amazon, YouTube, and other multinational server farms over the last several years has seen a considerable increase in their number of servers. Google approximately has over 2.5 million servers in its CDCs, and this number is continuously increasing as the demands for huge data storage and processing increase.

The majority of large business enterprises recognize power efficiency as a major challenge in their IT infrastructures. The efficiency of the physical resources is determined by the amount of energy consumption, which depends on the system deployed for the management and efficiency of running applications. Each server is equipped with cooling devices to keep the CPU temperature fix. While the power consumption of cooling equipment greatly changes according to the data center’s temperature as well as the power consumption of CPU. According to Amazon estimates, 53% of the total budget of a CDC is allocated for the expenditures associated with the cost and operations of the servers, whereas the power costs accumulate to about 43% of...
the entire budget. This embraces both the direct power consumption (~18%) and cooling infrastructure (~24%).

The rapid increase in carbon footprints from information communication technology (ICT) infrastructures has motivated the researchers to emphasize and give priority to energy consumption and CO₂ emission in tier-level CDCs. Recent statistics show that in the year 2015, CDCs consume around 1.8% of electricity worldwide, which contributes to almost 3% of CO₂ emissions globally. It is further anticipated that this amount will be increased to 18% (about 6% of the global CO₂ emission) in 2020. Another report on carbon emissions published in 2016 revealed that CDCs globally produce nearly 1.5 Mt of CO₂ in 2016, and this will be tripled by 2025. To solve these problems, energy-aware cloud infrastructure systems should be adopted and implemented to reduce total energy consumption and CO₂ emissions in CDCs. To minimize global warming effects in Pakistan, it was recommended to use cloud-enabled solutions along with renewable energy sources for power generations across IT infrastructures. Furthermore, improving power infrastructure and stepping up electricity preservation techniques can be applied to prevent the negative effects on environment and electricity consumption in CDCs.

Related work

In this section, we review and highlight the related work on CDCs for implementing energy-efficient data centers. During the last decade, cloud computing solutions have been widely applied for sharing resources and services with better efficiency, speed, availability, reliability, and security. The problem comes when the demand for these cloud services from business enterprises increases. This leads toward expanding the currently existing data centers without looking into the aspect of huge power consumption by the server machines and other facilities which ultimately contribute to generating carbon emissions hazardous for environmental sustainability. A recent report suggested that to minimize the negative effects of high-power consumption, cloud-enabled infrastructures and service solutions will lead to almost 36% reductions in global power consumption by 2020. Yang et al. proposed a green power management solution for resource allocations in cloud VMs.

Dumitru et al. presented a technique to increase energy management in data centers but did not give implementation details of the proposed solution. Pedram proposed a solution for energy-efficient data centers based on performance measurement metrics, but the solution was not practical application in large tier-level CDCs. The proposed technique introduced the resources and their power provisioning. The cloud-enabled solution proposed by Brandic was focused on service-level agreement (SLA) management, and it was extremely important for both cloud service providers and consumers to prevent SLA’s violations. Brandic also proposed applications of cloud computing with a proactive solution which led to less power consumption to increase energy efficiency and optimize resource utilization. Maurer et al. proposed an adaptive resource configuration technique for cloud infrastructure management. An energy-efficient mechanism proposed by Valliyammai et al. used the fuzzy c-means algorithm to reduce energy consumption by suspending the idle VMs while minimizing the SLA violations. Uddin and Rahman proposed different energy-saving techniques for CDCs to achieve maximum resource utilization and reduce the effects of global warming.

Warkozek et al. proposed a solution based on CPU power consumption for VMs and services turning on the servers. In this solution, a linear model of electricity consumption was modeled based on running software on the servers. The main aim of this solution is to optimize the power consumption of CDCs due to load on servers. An optimization technique was presented by Cioara aiming to identify non-optimal servers and then bring them to optimal states. This is a two-stage immune-inspired technique including initialization and self-optimization stages. A green aware strategy was designed by Wang et al. to eliminate power consumed by both IT functional devices and cooling systems. Green aware strategies enabled the cloud provider to make data center energy efficient while keeping the required applications at an acceptable level. Le and Wright have proved that scheduling the workload over the data center in the time-period can reduce both electricity cost and CO₂ emission.

However, the problem with most of these techniques is that they do not consider virtualized data centers and also did not focus on how to reduce CO₂ emissions in CDCs. The nucleus research center also emphasized on implementing cloud-based solutions to reduce greenhouse gas emission to about 28% from 2010 to 2020. This article proposes a technique, which applies server virtualization on servers inside the data center and distributes them into measurable units with their associated equipment into different resource pools depending on their workloads.

Proposed power efficiency solution for CDCs

This article proposes and implements a technique to improve the physical server utilization by reducing power consumption. In the proposed server consolidation technique, we consolidated several existing
physical servers to virtual servers called VSMs based on the virtualized task scheduling algorithm. This technique virtually reduces the number of PSMs to almost half or even less than that based on the workloads to be executed. This technique not only reduces power consumption but also CO₂ emissions. The details of the proposed technique are given as follows.

Server consolidation

Server consolidation, a trend sweeping the data center industry, is a technology data center owner is using to reap tremendous financial and capital expenditure savings. This technology instantiates and creates newly configured VMs on the existing PSMs for better utilization of processing capabilities of existing PSM with lower power consumption. Server virtualization maximizes the hardware utilization by aggregating more applications and services onto less hardware, thus enabling applications and services to coexist safely on the same server hardware with multiple operating systems operating simultaneously in the cloud environment. Some of the key benefits of server virtualization are as follows: (1) it enables to increase the resource utilization for both IT infrastructures and non-IT infrastructures to increase the overall proficiency of CDCs by consolidating the workloads of PSMs onto VSMs, (2) it supports to improve the efficacy in conducting business processes more efficiently and with better performance in terms of cost, time, and space, thereby improving the overall business performance, and (3) it diminishes the overall cost of ownership of data center and contributes toward lowering the CO₂ emissions thus helps in achieving sustainable and environment-friendly CDCs.

The increasing electricity and power costs in the CDC industry in Pakistan, such as banks, financial organizations, hospitals, and other multinationals are demanding to adopt and incorporate server virtualization solutions in their existing data center infrastructures to reduce power consumption and its associated costs. The proposed server consolidation technique instantiates and logically creates a virtual instance of an existing PSM. This virtualized instance comprises of three main components called as: VMs, host server machine, and hypervisor also called virtual machine manager (VMM). The VM directly intermingles with VMM and will be characterized as one of the main virtual infrastructure constituents. The host server machine associates itself with the VMM or the original PSM to handle and assign required resources to the newly created VM. The hypervisor or VMM is a cloud management layer between physical machines and their associated virtual infrastructures to handle the coordination between the two entities as shown in Figure 3.

In the proposed server consolidation technique, VSMs are allocated and reallocated between different PSMs using a technique called VM migration. This aids to increase the utilization of existing physical resources which eventually decreases the number of PSMs. The consolidation of VSMs increases the utilization of already installed physical machines, hence reducing power consumption. This reduction in power consumption mainly depending on the workload of several VMs operating on a physical machine as shown in Figure 4. The proposed technique distributes the workloads among VSMs using a virtualized task scheduling algorithm. For each workload (task) that previously requires a dedicated physical machine will now be reallocated to a newly created VSM. This enables us to
assign several tasks from different PSMs onto one VSM, thus enabling us to switch the tasks assigned based on timestamp values and priorities set to process the assigned task. This is a complex process where jobs and tasks scheduled for one PSM are now reallocated and reassigned to different VMs. To achieve this, we proposed a Virtualized Task Scheduling Algorithm, which helps to consolidate multiple workloads from each physical server onto multiple VSMs. The scheduling process was implemented with a converter tool, which automates the conversion process of PSMs onto VSMs and helps to simplify the whole migration and cloning process.

The proposed server consolidation technique categorizes the underutilized PSMs into three resource pools of physical machines called PSM-1, PSM-2, and PSM-3 depending on different parameters such as the number of workloads, utilization ratio, performance, usage, and applications executed. These PSM classifications are defined as follows:

PSM-1: these servers are the simplest type of servers and are extensively available in the majority of CDCs. These are mostly installed in shared computing environments, where provisioning for new servers is always available. The workloads processed by this category of servers require less processing speed and power, and they are considered as the idlest servers in the CDC.

PSM-2: these type-2 server machines are more controllable and powerful servers compared to type-1. The typical tasks processed by these servers include testing and evaluating applications and processes. These servers are usually fewer in number compared to type-1 servers.

PSM-3: these are the most powerful machines in any CDC. They are assigned the most complex, critical, and problematic tasks compared to types 1 and 2. These servers consume the maximum power and also emit the maximum CO₂ emissions. These servers are usually few in quantity and have a substantial effect on the total business processes.

**Virtualized task scheduling algorithm**

Business organizations are under tremendous pressure to report and account for their environmental footprints. Currently, business enterprises are forced to improve their sustainability efforts and recognize the importance of CO₂ emissions in building an environment-friendly ecosystem. To achieve this, the proposed server consolidation technique was employed in a CDC in Pakistan. During this process, initially, the CDC was restricted from adding any new hardware or other equipment especially the accumulation of new physical servers. In the next phase, a networked and integrated solution was instigated, where the VMs were created on already existing PSMs and workloads are migrated to more efficient VSMs by applying the proposed Virtualized Task Scheduling Algorithm. The aim was to reduce environmental footprints and decrease power consumption.

In the proposed algorithm, a data center component called load balancer creates and maintains a database containing an index of VSMs and their state information. The state can be either available or busy. In the first step, the VSM generates a request to the data center scheduler to know which VM is currently available in an accessible state to execute a specific task. The

![Figure 4. Server machine virtualization (SMV).](image-url)
Pseudocode: virtualized task scheduling algorithm.

Line 1. Phase 1: At the beginning, the index value for all SVM machines will be set to 0 as the current state of all machines is accessible.

Line 2. Phase 2: The client sends a new request to the DataCenterController.

Line 3. Phase 3: The LoadBalancer will be queried for the subsequent provision by the DataCenterController.

Line 4. Phase 4: The VSM index value will be parsed by DataCenterController to find the subsequent accessible VSM.

Line 5. Phase 5: The found VSM Id is sent to DataCenterController by the LoadBalancer.

Line 6. Phase 2 will be continued.

Line 7. By applying the Round Robin algorithm, if idle VSM is not found, the index value of VSM machine will be reinitialized to “0” and this cycle will be patterned again to find out accessible state VSM machine.

Line 8. Phase 5: When the VSM finishes processing the request, and the DataCenterController receives the cloudlet response, it notices the load balancer of the VSM de-allocation.

Line 9. Phase 6: The Load Balancer updates the status of VSM in the allocation table to available.

Line 10. Phase 7: Continue from Phase 2.

Line 11. The purpose of the algorithm is to find the expected Response Time of each VSM and is calculated as Response Time = Fint − Arrt + TrDelay(1).

Line 12. Where Arrt is the time when user request was received, and Fint is the time when user request was complete, and if there is any delay it can be found by applying below-given procedures:

Line 13. TrDelay = TrLatency + TrTransfer(2).

Line 14. Where TrDelay: time required to transmit data (TD) of a single application from source VSM to destination VSM.

Line 15. TrTransfer = TD/Bdtperuser (3).

Line 16. Bdtperuser = Bdttotal/Nuser (4).

Line 17. Where Bdttotal: total accessible bandwidth and Nuser: the number of requests from the user in the system.

The use of virtualization technology in an IT portfolio improves the cost per transaction, lessens data center space, power, and cooling requirements, and deploys resources easily in a fast-changing environment. It over-dramatically changes the operating capability of CDCs by creating virtual instances on physical servers and the resources available in the CDC. In the case study, servers are categorized according to the workloads they perform, applications they execute, and the power they consume. Their power consumption was measured before and after the implementation of the proposed server consolidation technique. Annual power consumption in different segments of the CDC was measured in watts. The power required to maintain the cooling and other facility infrastructures was also measured using power measuring metrics. Other power consumptions in the form of additional leakages and other services, such as generators and uninterruptible power supply (UPS), are also measured. Finally, the overall annual cost applied to total power consumed in the CDC was calculated to measure the entire cost of ownership.

In the case study, server consolidation was practically initiated and employed on all types of PSMs to create logical instances of VSMs and to shift their workload onto VSMs. The efficiency and power consumption of the existing PSMs was measured after applying the proposed technique to obtain the overall power feasting costs. The outcomes of the case study were analyzed with more than 500 PSMs installed in the CDC. Although the data center was large and
contains hundreds of other servers, the top-level management agreed to implement the proposed server consolidation technique on 30% of the designated PSMs and their associated devices. Finally, all values before and after the installation of the server consolidation are gathered, compared, and analyzed to obtain total power consumption costs. Benchmarks are set to measure the overall efficiency using the standardized metrics from the Green Grid.

### Power savings

The CDC industry took a lot of time to realize the importance of aligning their processes and operations conducive to power efficiency and reduced CO₂ emissions. Sustainable development requires a supply of energy resources that are sustainably available at an equitable cost. To achieve power savings, server consolidation was implemented to reduce the underutilization of already installed PSMs. The proposed technique showed a dramatic reduction in IT power systems, almost substantially more than the cost of servers. Table 1 shows the total number of PSMs in the CDC, their classification along with their utilization ratio, and power consumed in watts before applying server consolidation. A total of 500 servers are selected for performing the case study and are classified into three classes depending on the workloads by applying the proposed task scheduling algorithm. The total utilization ratio of all servers was about 10% on average, whereby each server averagely consumed about 100 W of power, as it was hard to get the accurate power consumption of each server, the total power consumed by all servers was approximately 50,000 W.

\[ 500 \times 100 = 50,000 \text{ W} \]

Table 1 illustrates that in the CDC, on average, the PSMs are utilized for only 10%–12%. The proposed server consolidation technique increased the efficacy of these underutilized server machines by merging the tasks onto less PSMs. After applying server consolidation with ratios of 7:1 for PSM-1 servers, 5:1 for PSM-2, and 3:1 for PSM-3 servers, a decrease of almost half of the power consumption by these servers was observed, as shown in Table 2. The average consolidation ratio for all servers after applying consolidation was 5:1. The consolidation ratios are selected based on the tasks and processing capabilities of these servers. Table 2 also shows that after applying server consolidation, the total number of servers is reduced to 96 from 500. After applying server consolidation, each server averagely consumed 275 W of power compared to 100 W. This upsurge in power consumption was instigated by the intensification in the workloads and

| Types of servers | PSM-1 | PSM-2 | PSM-3 | Total power usage |
|------------------|-------|-------|-------|------------------|
| Number of servers | 250   | 175   | 75    | 500              |
| Average utilization ratio | 8%    | 12%   | 10%   | 10%              |
| Average watts per server = 100 × server count | \(250 \times 100 = 25,000 \text{ W}\) | \(175 \times 100 = 17,500 \text{ W}\) | \(75 \times 100 = 7500 \text{ W}\) | \(50,000 \text{ W}\) |

PSM: physical server machine.

---

### Table 1. Pre-consolidation ratio of servers and their power consumption.

| Types of servers | PSM-1 | PSM-2 | PSM-3 | Total power usage |
|------------------|-------|-------|-------|------------------|
| Number of servers | 250   | 175   | 75    | 500              |
| Average server consolidation ratio | 7:1   | 5:1   | 3:1   | 5:1              |
| Number of servers after consolidation | 36    | 35    | 25    | 96               |
| Power consumption before consolidation | 25,000 W | 17,500 W | 7500 W | 50,000 W |
| Power consumption after consolidation | \(36 \times 275 = 9900 \text{ W}\) | \(35 \times 275 = 9625 \text{ W}\) | \(25 \times 275 = 6875 \text{ W}\) | \(26,400 \text{ W}\) |
| Total power savings | \(25,000 - 9900 = 15,100 \text{ W}\) | \(17,500 - 9625 = 7875 \text{ W}\) | \(7500 - 6875 = 625 \text{ W}\) | \(23,600 \text{ W}\) |

PSM: physical server machine.

---

### Table 2. Consolidation ratio of servers and their power consumption after server consolidation.

| Types of servers | PSM-1 | PSM-2 | PSM-3 | Total values |
|------------------|-------|-------|-------|--------------|
| Total number of servers | 250   | 175   | 75    | 500          |
| Average server consolidation ratio | 7:1   | 5:1   | 3:1   | 5:1          |
| Number of servers after consolidation | 36    | 35    | 25    | 96           |
| Power consumption before consolidation | 25,000 W | 17,500 W | 7500 W | 50,000 W    |
| Power consumption after consolidation | \(36 \times 275 = 9900 \text{ W}\) | \(35 \times 275 = 9625 \text{ W}\) | \(25 \times 275 = 6875 \text{ W}\) | \(26,400 \text{ W}\) |
| Total power savings | \(25,000 - 9900 = 15,100 \text{ W}\) | \(17,500 - 9625 = 7875 \text{ W}\) | \(7500 - 6875 = 625 \text{ W}\) | \(23,600 \text{ W}\) |

PSM: physical server machine.
processing capability of these servers. The utilization ratio also increased to almost 30% compared to 10% before consolidation.

Table 2 further describes the average power consumed after applying server virtualization. The total power consumed by all PSMs after consolidation was only 26,400 W compared to 50,000 W before server consolidation, thus a saving of almost 23,600 W of power, which is almost half of the total power consumed before applying server consolidation was achieved. Furthermore, if the utilization ratio is further increased to 50%, about 350 W of power on average will be consumed by every server compared with 275 W. However, a decrease of almost two to three times of power consumption was possible to be achieved in these PSMs. In the CDC, where this case study was performed, the target utilization for PSMs was 30%, but in actual they are performing at only 5%, we can certainly decrease the power consumption to almost 90% by shutting down idle server machines. Thus, the proposed technique was beneficial in implementing power and energy-efficient and environmentally friendly data center as power consumption was almost reduced to half. The above results demonstrate that, when PSMs are consolidated with an average ratio of 5:1, a substantial decrease in the power consumption is obtained, that is, five physical servers (100 W each server on average) consumed a total of 500 W. However, after consolidating the load of these five PSMs, the average power consumption of one server becomes 275 W and its utilization ratio also upsurges to virtually 30%. It practically saves 225 W, which is nearly 50% of the total power consumed by five PSMs.

\[ 500 \text{ W} - 275 \text{ W} = 225 \text{ W} \]

### Cost savings in PSMs

A typical data center has \( n \) different application environments and \( m \) servers. Each application executes several classes of transactions and requires several servers for processing. Servers can be dynamically provisioned among applications to optimize server utilization in the CDC. Server consolidation decreases the number of PSMs to reduce the power consumption and costs required to manage those servers. The power reduction has a great effect on the overall data center heating requirements, cooling load, and UPS backup time. It helps to increase performance proficiency, protracts generator backup times, and diminishes IT configuration of diverse interconnecting types of equipment and devices. This technique also reduces overall cost in terms of annual savings linked with cooling and other costs, such as generators, UPS, and power leakages. These cost-saving techniques have a great effect on the purchasing and maintenance of data center equipment and services.

### Cost saving in PSM-1 servers

PSM-1 servers are the largest pool of servers in the data center under study. Therefore, their cost savings are also huge because their consolidation ratio was higher compared to other types of servers. Table 3 shows that by applying server consolidation, the number of servers was abridged to 36 from 250. Each server machine was consuming 250 W of power, and total power consumed by these servers including cooling, UPS, power leakage, and other additional electrical consumptions was approximately 17,187.5 kWh of total power consumption before applying server consolidation. The total cost for this power consumption was PKRs. 10,070,000. After applying consolidation, the number of servers was reduced to 36 and the total power consumed was 594 kWh and the costs are PKRs. 2,890,800. So, a total cost saving of PKRs. 7,179,200 (US$46,618) was achieved (US$1 = PKRs. 154) as given in Table 3.

### Cost saving in PSM-2 servers

Table 4 shows that by applying server consolidation, the number of servers was reduced to 35 from 175. The power consumed by
each server was 175 W; hence, the total power consumed by these servers including cooling, UPS, power leakage, and other additional electrical consumptions are approximately 6,878 kWh of total power consumption before applying server consolidation. The total cost for this power consumption was PKRs. 8,427,120. After applying consolidation, the number of servers was reduced to 35 and the total power consumed was 289 kWh and costs are PKRs. 2,715,600. So, a total cost saving of PKRs. 6,711,520 (US$43,581) was achieved (US$1 = PKRs. 154).

**Cost saving in PSM-3 servers.** Table 5 demonstrates that by employing the proposed methodology, the PSMs are reduced to 25 compared to 75 and each machine was consuming 75 W of power. The total power consumed by these servers including cooling, UPS, power leakage, and other additional electrical consumptions was approximately 1,541 kWh of total power consumption before applying server consolidation. The total costs for this power consumption were PKRs. 3,600,360. After applying server consolidation, the number of servers was reduced to 35 and the total power consumed was 289 kWh and costs are PKRs. 2,027,064. So, a total cost saving of PKRs. 1,573,296 (US$10,216) was achieved (US$1 = PKRs. 154).

**Total cost savings for all server machines.** Table 6 shows the annual cost savings of PKRs. 12,067,776 (US$78,362) in all types of servers in terms of power consumption per kWh, including cooling costs, power leakage costs, additional costs (generators and UPS), and overall cost savings before and after server consolidation. Therefore, the power utilization ratio was greatly reduced. Lessening PSM diminishes power consumption as well as has a great effect on CDC IT equipment. It also helps in reducing CO₂ emissions to attain environmentally friendly and sustainable CDCs. From the results obtained, it is visible that the algorithm improved the efficacy and performance of the CDC. The algorithm performed much better during the three designated phases in which VSMs are created, tasks are allocated and reallocated to them by the scheduler, and finally, their annihilation process.
algorithm helps the CDC to increase the utilization of already installed PSMs and also save huge power consumption.

**Key findings**

1. By implementing server consolidation on an efficient platform, the data center annually saves 23,600 W of total power and the annual cost of PKRs. 12,067,776 kWh being wasted on idle servers and other cooling and infrastructure facilities.

2. The total number of processors within servers under management was reduced by up to 50% compared with the total value before workload consolidation; data centers were able to reduce their server hardware and software costs.

3. The overall utilization ratio of underutilized PSMs increased to 30% compared to only 10%, and this ratio can further be increased by involving more servers.

4. The overall task completion time of different IT services was increased to almost 50%.

**Conclusion**

In this article, we have discussed the enactment of server consolidation in a CDC in Pakistan. It has been pragmatic in this case study that the proposed technique along with the proposed algorithm leads to cost and power savings of almost 50% or even more. The article proposed a technique to classify CDC servers into three classifications depending on the applications and workloads they execute. A virtualized task scheduling algorithm was proposed to help data center managers to properly distribute the workload among different servers. Server consolidation was applied on all PSMs to migrate their workload onto VSMs by applying the proposed algorithm. The results indicate a power saving of nearly 23,600 W and an increase in the utilization ratio of already installed servers to almost 30% from 10%. The results also demonstrate a cost saving of PKRs. 12,067,776 kWh (US$78,362) being wasted on idle servers and other cooling and infrastructure facilities before implementing server virtualization.

**Acknowledgement**

This research was funded by the Deanship of Scientific Research at Princess Nourah bint Abdulrahman University through the Fast-track Research Funding Program.

**Declaration of conflicting interests**

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

**Funding**

The author(s) received no financial support for the research, authorship, and/or publication of this article.

**References**

1. Khosravi A, Andrew L and Buyya R. Dynamic VM placement method for minimizing energy and carbon cost in geographically distributed cloud data centers. *IEEE Trans Sustain Comput* 2017; 2(2): 183–196.

2. Nasim R, Taheri J, Kassler AJ, et al. Optimizing virtual machine consolidation in virtualized datacenters using resource sensitivity. In: Proceedings of the 2016 IEEE international conference on cloud computing technology and science (CloudCom), Luxembourg, 12–15 December 2016. New York: IEEE.

3. Lee S, Chung BD, Jeon HW, et al. A dynamic control approach for energy-efficient production scheduling on a single machine under time-varying electricity pricing. *J Clean Prod* 2017; 165: 552–563.

4. Shehabi A, Walker B and Masanet E. The energy and greenhouse-gas implications of internet video streaming in the United States. *Environ Res Lett* 2014; 9(5): 054007.

5. Horner N and Azevedo I. Power usage effectiveness in data centers: overloaded and underachieving. *Electr J* 2016; 29: 61–69.

6. Oro E, Depoorter V, Garcia A, et al. Energy efficiency and renewable energy integration in data centers: strategies and modeling review. *Renew Sustain Energy Rev* 2015; 42: 429–445.

7. Zhou Z, Zhi-Gang H, Jun-Yang Y, et al. Energy-efficient virtual machine consolidation algorithm in cloud data centers. *J Central South Univ* 2017; 24(10): 2331–2341.

---

**Table 6. Total cost savings for all types of servers.**

| Total cost savings                  | Before consolidation | After consolidation |
|------------------------------------|----------------------|---------------------|
| Server power consumption cost      | 8,760,000 kWh        | 4,196,040 kWh       |
| Cooling costs                      | 10,941,240 kWh       | 2,088,384 kWh       |
| Additional cost savings (UPS)      | 2,959,420 kWh        | 4,308,460 kWh       |
| Total cost                         | 22,660,660 kWh       | 10,592,884 kWh      |
| Total cost savings                 | PKRs. 22,660,660     | 10,592,884 = 12,067,776 kWh (US$78,362) |

UPS: uninterruptible power supply.

---

This table shows the total cost savings before and after consolidation, including server power consumption, cooling costs, and additional cost savings due to uninterruptible power supply (UPS). The total cost savings are significant, amounting to 12,067,776 kWh (US$78,362), indicating a substantial reduction in energy consumption and associated costs.
8. Smith A, Horner SJ, Azevedo N, et al. United States Data Center energy usage report. LBNL-1005775, Berkeley, CA: Lawrence Berkeley National Laboratory, 2016, https://wwwosti.gov/servlets/purl/1372902
9. IDC, 2014, https://www.emc.com/leadership/digital-universe/2014view/executive-summary.htm
10. Haroon SA, Ghulam S and Naseem F. Electricity subsidies and welfare analysis: the perspective of Pakistan. Centre for Environmental Economics and Climate Change (CEECC), 2019, https://www.pide.org.pk/pdf/Working%20Paper/WorkingPaper-164.pdf
11. Vasudevan M, Tian Y-C, Tang M, et al. Profile-based dynamic application assignment with a repairing genetic algorithm for greener data centers. J Supercomput 2017; 73(9): 3977–3998.
12. Zhu W, Zhuang Y and Zhang L. A three-dimensional virtual resource scheduling method for energy saving in cloud computing. Future Gener Comput Syst 2017; 69: 66–74.
13. Zhou Z, Abawajy J, Chowdhury M, et al. Minimizing SLA violation and power consumption in Cloud data centers using adaptive energy-aware algorithms. Future Gener Comput Syst 2018; 86: 836–850.
14. Ranjbari M and Akbari Torkestani J. A learning automata-based algorithm for energy & SLA efficient consolidation of virtual machines in cloud data centers. J Parallel Dist Comput 2018; 113: 55–62.
15. EPA. ENERGY STAR Program requirements for computer systems–Draft 4. Washington, DC: Environmental Protection Agency, 2009.
16. Aujla GS and Kumar N. MEnSuS: An efficient scheme for energy management with sustainability of cloud data centers in edge-cloud environment. Future Gener Comput Syst 2018; 86: 1279–1300.
17. Uddin M, Rahman AA and Memon J. Carbon sustainability framework to reduce CO2 emissions in data centres. Int J Green Econ 2011; 5(4): 353–369.
18. Tang X, Liao X, Zheng J, et al. Energy efficient job scheduling with workload prediction on cloud data center. Cluster Comput 2018; 21: 1581–1593.
19. Haque MA, Aydin H and Zhu D. On reliability management of energy-aware real-time systems through task replication. IEEE T Parallel Distrib Syst 2017; 28(3): 813–825.
20. Hameed A, Khoshkbarforoushha A, Ranjan R, et al. A survey and taxonomy on energy efficient resource allocation techniques for cloud computing systems. Computing 2016; 98(7): 751–774.
21. Bawden T. Global warming: Data centres to consume three times as much energy in the next decade, experts warn 2016, http://www.independent.co.uk/environment/global-warming-data-centres-to-consume-three-times-as-much-energy-in-next-decade-experts-warn-a6830086.html
22. Mishra SK, Puthal D, Sahoo B, et al. Energy-efficient VM-placement in Cloud data center. Sust Comput: Inform Syst 2018, https://rranjans.files.wordpress.com/2019/07/26.pdf
23. Al-Dalaimy A, Itani W, Zantout R, et al. Type-aware virtual machine management for energy efficient cloud data centers. Sust Comput: Inform Syst 2018; 19: 185–203.
24. Khan MA, Paplinski A, Khan AM, et al. Dynamic virtual machine consolidation algorithms for energy-efficient cloud resource management: a review. In: Rivera W (ed.) Sustainable cloud and energy services. Cham: Springer, 2017, pp.135–165.
25. Puthal D, Obaidat MS, Nanda P, et al. Secure & sustainable load balancing of edge data centers in fog computing. IEEE Commun Mag 2018; 56(5): 60–65.
26. Raza SA, Shahbaz M and Nguyen DK. Energy conservation policies, growth, and trade performance: evidence of feedback hypothesis in Pakistan. Energy Pol 2015; 80: 1–10.
27. Zhang W, Qi Q and Deng J. Building intelligent transportation cloud data center based on SOA. Int J Ambient Comput Intell 2017; 8(2): 1–11.
28. Pike Research. Cloud computing energy efficiency. Boulder, CO: Pike Research, 2010.
29. Yang C, Wang K, Cheng H, et al. Green power management with dynamic resource allocation for cloud virtual machines. In: Proceedings of the IEEE international conference on high-performance computing and communications, Banff, AB, 2–4 September 2011. New York: IEEE.
30. Dumitru I, Fagarasan I, Iliescu S, et al. Increasing energy efficiency in data centers using energy management. In: Proceedings of the IEEE/ACM international conference on green computing and communications (GreenCom), Chengdu, China, 2011, pp.159–165. New York: IEEE.
31. Pedram M. Energy-efficient data centres. IEEE T Comput-Aid Des Integr Circ Syst 2012; 31(10): 1465–1484.
32. Brandic I. Reducing energy consumption by using clouds. In: Proceedings of the IEEE 15th international symposium on symbolic and numeric algorithms for scientific computing (SYNASC), Timisoara, 26 May 2013, pp.3–4. New York: IEEE.
33. Maurer M, Brandic I and Sakellariou R. Adaptive resource configuration for cloud infrastructure management. Future Gener Comput Syst 2013; 29(2): 472–487.
34. Valliyammai C, Uma S, Dhivy BK, et al. Efficient energy consumption in green cloud. In: Proceedings of the international conference on recent trends in IT (ICR-TIT), Chennai, India, 10–12 April 2014, pp.1–4. New York: IEEE.
35. Uddin M and Rahman AA. Techniques to implement green data center to achieve energy efficiency & reduce global warming effects. Int J Global Warm 2011; 3(4): 372–389.
36. Warkozek G, Drayer E, Debusschere V, et al. A new approach to model energy consumption of servers in data centers. In: Proceedings of the IEEE international conference on industrial technology (ICIT), Athens, Greece, 4 June 2012, pp.211–216. New York: IEEE.
37. Cioara T, Pop CB, Anghel I, et al. Immune-inspired technique for optimizing server’s energy consumption. In: Proceedings of the IEEE international conference on intelligent computer communication and processing (ICCP), Cluj- Napoca, 26–28 August 2010, pp.273–280. New York: IEEE.
38. Wang X, Du Z, Chen Y, et al. A green-aware virtual machine migration strategy for sustainable data centers powered by renewable energy. Sim Model Prac Theor 2015; 58: 3–14.
39. Le T and Wright D. Scheduling workloads in a network of data centres to reduce electricity cost and carbon footprint. *Sust Comput: Inform Syst* 2013; 5: 31–40.
40. Mao L, Qi DY, Lin WW, et al. An energy-efficient resource scheduling algorithm for cloud computing based on resource equivalence optimization. *Int J Grid High Perform Comput* 2016; 8(2): 43–57.
41. Uddin M, Memon J, Alsaqour R, et al. Virtualized load management algorithm to reduce CO₂ emissions in the data centre industry. *Int J Global Warm* 2015; 7(1): 3–20.
42. Osanaiye O, Chen S, Yan Z, et al. From cloud to fog computing: a review and a conceptual live VM migration framework. *IEEE Access* 2017; 5(1): 8284–8300.
43. Nguyen TH, Di Francesco M and Yla-Jaaski A. Virtual machine consolidation with multiple usage prediction for energy-efficient cloud data centers. *IEEE T Serv Comput* 2017; 13: 186–199.
44. Uddin M and Rahman AA. Energy efficiency & low carbon enabler green IT framework for data centers considering green metrics. *Renew Sust Ener Rev* 2012; 16(6): 4078–4094.
45. Dai X, Wang JM and Bensaou B. Energy-efficient virtual machines scheduling in multi-tenant data centers. *IEEE T Cloud Comput* 2016; 4: 210–221.
46. Monil MAH and Rahman RM. VM consolidation approach based on heuristics, fuzzy logic, and migration control. *J Cloud Comput* 2016; 5(1): 591–5918.