



Received: 09-04-2020

Accepted: 23-04-2020

Published: 25-05-2020

Editor: Dr. Natarajan Gajendran

Citation: Panhwar MA, Khuhro SA, Mazhar T, Liang DZ, Bilal MEA, Qadir N (2020) Virtual machine optimization to achieve energy efficient optimum resource utilization in cloud data center. Indian Journal of Science and Technology 13(13): 1423-1434. <https://doi.org/10.17485/IJST/v13i13.211>

***Corresponding author.**

Muhammad Aamir Panhwar
School of Electronic Engineering,
Beijing University of Posts and
Telecommunications, China
maamirpanhwar@hotmail.com

Funding: National High Technology
863 Programs of China
(No.2015AA124103) and National
Key R&D Program (No.
2016YFB05502001)

Competing Interests: None

Copyright: © 2020 Panhwar,
Khuhro, Mazhar, Liang, Bilal, Qadir.
This is an open access article
distributed under the terms of the
[Creative Commons Attribution
License](#), which permits unrestricted
use, distribution, and reproduction
in any medium, provided the
original author and source are
credited.

Published By Indian Society for
Education and Environment ([ISSEE](#))

Virtual machine optimization to achieve energy efficient optimum resource utilization in cloud data center

Muhammad Aamir Panhwar^{1*}, Sijjad Ali Khuhro², Tehseen Mazhar³,
Deng Zhong Liang¹, Mughal-E-Azam Bilal³, Nabeel Qadir³

¹ School of Electronic Engineering, Beijing University of Posts and Telecommunications, China

² School of Computer Science and Technology, University of Science and Technology of China, China

³ School of Computer Science, Virtual University of Pakistan, Pakistan

Abstract

Objectives: Cloud offers multiple benefits through its data center-based services. The whole world uses these services that are hosted by physical machines. Millions of virtual machines get an optimized utilization of hardware to utilize these physical machines. The unbalanced distribution of virtual machines to physical machines offers an in-efficient utilization of data center hardware that will also lead to more carbon emission and harm the environment as well. **Methods/material:** A learning function is needed to offer energy efficiency in the cloud data center through VM optimization through its optimal allocation to physical machines. Therefore, an optimal VM placement and migration algorithm is a challenge that is addressed in this paper to reach an efficient energy optimization and resource utilization level. **Finding/Novelty:** The proposed algorithm is led by a learning function that takes into account the available number of physical machines, number of virtual machines, incoming requests and decides to run an optimal number of physical machines to obtain energy efficiency level for the cloud data center by migrating the virtual machines (VMs).

Keywords: Datacenter; Cloud computing; Energy efficiency; VM placement; VM migration; Green Cloud

1 Introduction

Computing has become a new consumer and virtualization model for high-cost computing infrastructures and web-based IT solutions. The cloud provides the appropriate on-demand service, flexibility, extensive access to the network, measurement services and resource pooling⁽¹⁾, with minimal management in the most customizable way. Low-cost computing devices, high-performance network resources, huge storage capacity, semantic web technology, service-oriented architecture SOA, application programming interface (API) are the list of functions in the cloud that can help the rapid development of cloud technology. The data center infrastructure typically

encapsulates all these existing technologies in a web services-based model to provide business agility, enhanced scalability, and on-demand availability. The rapid deployment model, low-commissioning investment, pay-per-use plan, multi-user resource sharing are additional attributes of the data center in the cloud and major industries tend to use these virtualized applications⁽²⁾.

Cloud computing is focused on virtualized data centers, and application providers will be available on a subscription basis. Data centers & cloud computing as a new technology has also raised a major issue about its environmental sustainability. You can save a lot of effort by using large shared virtualized data center cloud computing. However, cloud services can further increase internet traffic and growing information databases, thereby reducing energy conservation. An energy-efficient cloud framework will be able to reduce the energy usage healthily without compromising the quality of service performance, responsiveness, and availability provided by the cloud provider. A unified solution will help achieve energy-efficient cloud computing goals by controlling the cloud's energy consumption. A high-level view of the energy-efficient cloud architecture from an earlier proposal is shown in figure 1. The goal of the architecture is to develop energy-efficiency in the cloud by taking into account both user perspective and provider perspective. In the energy-efficient cloud architecture proposed by⁽³⁾, users submit their cloud service requests through the new middleware energy-efficient proxy, which manages the best energy-efficient cloud provider options to meet the needs of users. The service requires three factors, software, platform, or infrastructure. Cloud providers can register their services in the energy-efficient form in the public directory visited by an energy-efficient broker. Energy-efficient offers include energy-efficient service, pricing and time to minimize energy utilization. Energy-efficient brokers get the energy status of the various cloud services in the energy catalog. The energy usage catalog retains all the data related to cloud service energy efficiency. The data may include the power usage effectiveness (PUE) and cooling efficiency, network cost, and carbon emissions rate of the cloud data center that provides the service and energy-efficient broker calculates the carbon emissions of all cloud providers that provide the requested cloud service. Then it chooses a service that will result in minimal carbon emissions and represents the user to purchase these services. The energy-efficient cloud framework is designed to track the overall energy consumption required by service users. It relies on two major components of the carbon emissions catalog and the energy-efficient cloud, tracking the energy efficiency of each cloud provider and encouraging the cloud provider to be energy-efficient. From the user's point of view, energy-efficient brokers play a vital role in monitoring and selecting the cloud, based on user QoS and ensure that users get the lowest possible carbon emissions. In general, users can use the cloud to access these three types of services (SaaS, PaaS, and IaaS), so the service process should also be energy efficient.

1.1 SaaS level

The vendors require to model and enumerate the energy efficiency of software designing, implement and deployment in a live environment, as SaaS providers primarily provide software installed on their data center or IaaS provider resources. For service users, SaaS providers not only choose energy-saving but also close to the user's data center. You should use energy-saving storage to maintain the minimum number of user confidential data⁽⁴⁾.

1.2 PaaS level

PaaS suppliers give administrative facilities to run their software. They enable the development of applications that guarantee energy efficiency. This should be possible by including different energy-efficiency levels, for example, Joule Sort. A product for energy effectiveness targets the measure of energy required to execute an operation. The PaaS architecture itself might need different code-level advancements that can be necessary for the effective execution of the application in the basic compiler. Application programs execution and advancement in the cloud permits the arrangement of client applications on hybrid clouds. For this situation, keeping in mind the end goal to accomplish extreme energy efficiency, the PaaS plate form is designed with an application and itself decides the processing needs of the application in the cloud⁽⁵⁾.

1.3 IaaS level

Nowadays IaaS level provides autonomous infrastructure services not only, but then also supports other services supported by the cloud, this supplier plays a vital role in the success of the entire energy effective building. The virtualization and association, you can further reduce energy consumption by turning off unused servers. Install a variety of energy meters and sensors to calculate the current energy efficiency of each IaaS provider and its site. This information is regularly published by the cloud provider in the carbon footprint. A selection of energy-efficient scheduling and resource allocation policies will ensure minimal energy consumption. Also, cloud providers have designed a variety of energy-efficient offers and pricing options to provide incentives for users during off-peak hours or at maximum energy savings. A cluster node architecture in the data center is given in Figure 1.

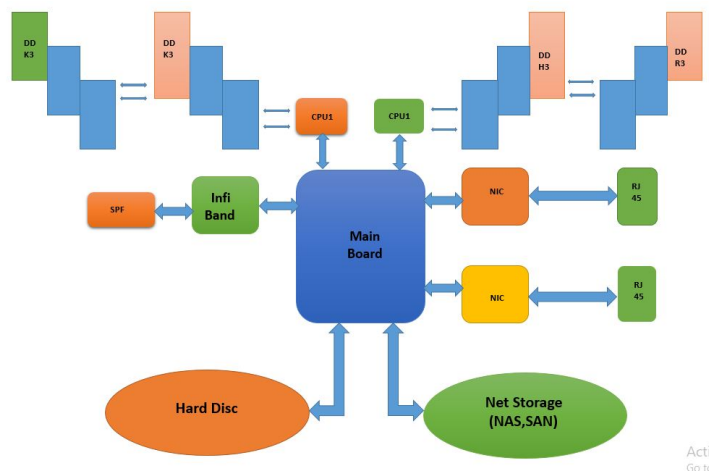


Fig 1. A cluster node architecture in the data center⁽⁶⁾

1.4 Problem statement

Reducing energy consumption now a day has become a major issue because of the economic, Environmental and marketing aspects of energy in all areas⁽⁷⁾. This concern has a great impact on the information and communications technology sector, electronic designers, especially in the network. The network infrastructure and data centers involve high-performance and high-availability machines. Therefore, they contain equipment that requires air conditioning to maintain its normal operation leading to high energy consumption⁽⁸⁾. Reducing energy consumption in the data center is an open challenge and is driving the future of energy-efficient data centers. Researchers like⁽⁶⁾ have pointed to the urgent need to combine the data center's integrated energy efficiency framework to combine energy-efficient IT architectures with specific activities and procedures, with minimal impact on the environment and heat emissions. The work carried out in this study is looking for an energy efficiency framework and should also consider the different types of applications and processes that require long-term resources as an important factor in energy consumption as a potential area for energy efficiency.

2 Review of Literature

Green IT is another name of energy-efficient computing, which enhances the efficiency of energy and lessens the usage of harmful material. Energy-efficient computing concentrates to minimize utilization of resources. To achieve this goal, energy-efficient computing applies to all phases belonging to computer networks like the development of energy-efficient CPUs. Based on the research studies⁽⁹⁾, energy-efficient computing refers to the study of design, manufacture, usage, and wastage of servers, PCs and their insignificant effects on the atmosphere. Cloud computing services are capable of executing the workloads, also consists of processing functionalities. There exist 3 computing services like 'high availability computing node, low availability computing node, and elastic infrastructure'. EI service managed the issues of provision and de-provision of resources.⁽¹⁰⁾ LACC, offers low-cost solutions for services, which are not dangerous. Due to the growing number of web-based and internet applications, the quick development of data-centers has been shown. Server computers range has been increased up to 30 million which are installed in data-centers. But the electricity consumption of new servers is becoming high day by day. If we increase the number of data-centers, then the cost of operating and the consumption of energy will be much higher. For several organizations, electrical energy availability is a serious issue⁽¹¹⁾. For the improvement of data-centers efficiency, few methodologies exist which have been using new equipment, management software, etc.

2.1 Software optimization by algorithmic optimization

The proficiency of algorithms greatly affects cloud resources which are necessary to execute computer programs. For instance, to change the search algorithms from linear search to hash, index and binary search might minimize cloud resources used for a given activity. It is described in the literature that google search release 7 gram CO₂, a query search emits 0.2 gram of CO₂ ⁽¹²⁾.

2.2 Virtualization

For the reduction of power utilization of data-centers, the virtualization strategy is suitable. The theory behind this concept is that one physical server can be used to host tremendous servers. Also, it shortens data-centers.

2.3 Virtual machine management in cloud computing for energy efficiency

In the past, efficient energy processing and cloud computing have been identified as the following two ideas. What can be help full for the association to locate the most 'energy-saving' project for cloud applications? This question opens up the foundation of energy-efficient cloud computing and green computing. The point-to-point conversation will provide some ideas on the link between the two ideas and give a preface to the idea that will be clarified in the following sections.

2.4 Consumption analysis of cloud

Cloud Stack is an application architecture that connects many systems at the IAAS provider level to plan and manage cloud resources, thereby reducing the usage of resources of the cloud. The combination of virtual machines, virtual machine movements, standby heat management, and awareness of temperature distributions is the case of such procedures that cause low power consumption. Virtualization is a key innovation for these specific programs because it has some elements, for example, real-time migration of cloud resources and a combination of servers. The combination contributes to the trade off between asset use and energy use. Similarly, VM movement⁽¹³⁾ allows the dynamic management of resources while reducing support costs. On the other hand, changes in virtualization leads to a significant decline in virtual machine overhead, thus increasing the cloud's energy productivity.

Till now, in the case of distributed servers, there are some risks and research involving cloud resource management⁽¹⁴⁾. Recommends a method of producing energy to calculate a management problem that attempts to reduce the use of energy in the cloud⁽¹⁵⁾. Proposed another solution to the cloud's resources and servers using over-provisioning issues. On the other hand, it is difficult to maintain the information of each virtual machine in the cloud data center due to different levels of abstraction. In this way, the virtual machine combination uses a different load estimation procedure.

3 Monitoring an energy efficient grid through vm management

Energy-efficient grid introduces measurements like and data center infrastructure Efficiency (DciE) and PUE, for the betterment of data-centers and to achieve measurement ratio⁽¹⁶⁾.

$$PUE = Total\ facility\ power / IT\ equipment\ power$$

$$DciE = \frac{1}{PUE} = IT\ Equipment\ power / Total\ facility\ power * 100\%$$

For the comparison of data-centers' efficiency, *PUE* and *DciE* are the major measurement instruments. From the equations, TFP is referred to be the power which is measurable for data-center, on the other side IT Equipment is said to be the total power utilized within the storage and processing of data.

3.1 Modeling energy utilization for VMs

VMs power usage is imperative to better sort out and plan their usage to reach the goal of an efficient data-center concerning energy utilization. Through VM, power utilization of CPU can be computed. Frameworks depending on data, for example, asset usage is referred to likewise as resource usage counters, which is proposed in⁽¹⁷⁾. They also introduce Joule meter which is a software of virtual machine power approximation and virtual machine power measurement technique. This software can precisely derive the power utilization without including any extra equipment by an application.

3.2 Modeling power utilization by virtual machine migration

Virtual machine relocation comprises shifting of a running VM between the server without any interference. The system enables VM combination to accomplish better energy effectiveness. It lessens extra power utilization and its cost in the form of energy is insignificant. The energy cost of relocation is never practically considered while moving VMs from one server to another. The key focus for productive VM combination faces issues like how to gauge the energy utilization of each VM migration and how to take relocation choices⁽¹⁸⁾.

⁽¹⁹⁾ Presented a numerical model with less overhead to assess the energy costs of virtual machine movements. The model is based on direct relocation between servers without an intermediate node, and the energy cost after the relocation is associated with the network transmission rate. The connection between the VMs in the formula whereas describes the total VM, b is capacity and A , B and C are the constants describing different factors of the system like bandwidth, server utilization, and size.

$$Emig = A + B * s + C * b$$

3.3 Effect of virtualization on cloud energy efficiency

Virtualization is specified as a key factor in cloud computing from a cost and energy proficiency point of view. As indicated by the meaning of virtualization, the technology fundamentally decreases the number of working PCs by making their execution as a simulation to reduce energy usage. Virtualization can be associated with conventional and cloud servers. In conventional server farms, contingent upon your strategies and necessities, you may utilize virtualization, however, cloud computing virtualization is imperative regarding energy production, so it is suggested that you utilize virtualization. Every part of IT can be virtualized, including servers, desktops, applications, administration tests, input/output (I/O), LAN, switches, storage, WAN advancement controller (WOC), application conveyance control (ADC) and a firewall. Here the three fundamental types of virtualization are: servers, desktops, and apparatuses. As a result of the connection between them, the focus should be on server virtualization and algorithm for resource scheduling since it will be the most critical. The explanation behind virtualization is to spare resources and powerfully relocate and design VMs between physical servers. Virtualization is a standout amongst the best methods for energy effectiveness⁽¹⁹⁾.

3.4 Virtual machine allocation

Virtual Machine is a programming execution that can scale equipment assets with the goal that numerous working frameworks can keep running on the two PCs in the meantime. Each working framework keeps running in its virtual machine. Assets, for example, hard disks, memory processors, and so on are assigned to each virtual machine in a consistent case.

As outstanding amongst other elements and capacities of virtualization, VM De allocation/allocation guarantees application accessibility. On account of support or investigating, the framework should be closed down for a time frame. With VM de-allocation, we can move virtual machines running various working frameworks and applications to another physical server without interfering with application operations. Keeping in mind the end goal to expand the up time of the application, it is a continuous allocation that happens when a virtual machine is running on the server and proceeds on the objective framework. VM allocation must be made under the accompanying conditions. Asset use is not sufficient, so the virtual machine should move to another server, this is server downtime. VM has a great deal of communication with different VMs on the server. Because of the workload, the VM temperature surpasses, so it should move to another server to cool the overheated server.

By the above criteria, it is discovered that VM De-Allocation likewise has the extra preferred standpoint of giving convenience by lessening the cost by killing the underutilized server and fulfilling the predefined execution. VM De-Allocation conveys a few advantages to cloud computing.

4 Hybrid box method for energy resource allocation

Cloud centers for data collection are power consumption organizations in a bulk manner, particularly if resources are all-time active, regardless of the possibility that they are not utilized. The idle server expends around 70% of the energy. The misuse of this idle energy is the primary explanation behind the low efficiency or productivity. A vital approach to bring energy effectiveness into the cloud condition is to present energy-efficient planning and awareness algorithm to improve resources administration. This work is finished by utilizing energy-efficient allotment and de-allocation of resources to lessen this commitment to extreme energy utilization. This will result in a large number of idle servers entering the rest mode. Intel's cloud computing 2015 vision likewise stresses the requirement for this dynamic asset management to enhance the power proficiency of server and data centers by killing and setting idle servers. This work utilizes the recipe of the boxed issue to show an exact energy-efficient assignment

algorithm. The logic behind it is to diminish the number of servers utilized or to amplify the number of inactive servers entering the rest mode. Keeping in mind the end goal to consider the workload and administration time, utilize a linear programming algorithm to persistently enhance the number of servers utilized after the administration begins. This de-allocation method is consolidated with asset allocation to decrease the aggregate power utilization within the server center.

The proposed algorithm works on the basis of the VM scheduler that is energy efficiency in nature. It can improve the present framework, by the supervision of a scheduler, for example, Open Nebula and Open Stack. Power utilization pointers can be given for energy utilization estimation of different instruments (e.g. Joules). It utilizes a committed test system to judge its performance and validate it⁽²⁰⁾. The evaluation tells that mixing of resource assignment algorithm and de-allocation method incredibly diminishes the number of servers serving a given load. Along these lines, limiting force utilization in the server center as well.

4.1 Proposed system framework

The model considers the way that the infrastructure vendor allocates the resources to request instances of user applications. It is equivalent of proportionate to working VMs for this reason. Physical resources are dealt with as servers. Except the application is bundled into a virtual machine facilitated by the framework supplier. Cloud suppliers aim to save energy and diminish power utilization by integrating and combining VMS allocation to limit idle servers from entering rest or sleep mode.

The accompanying figure portrays a framework, that shows how power utilization estimator can be administered in cloud resources (for resource instantiation & administration) works under proposed energy proficiency allotment and de-allocation algorithms. A concise portrayal of every module sets a phase for the demonstration of energy proficiency asset management issues in the cloud.

- IaaS management module in the cloud (for example, 'Open Stack, Open Nebula & Eucalyptus') will control & oversee resources within the cloud as per incoming customer demands, VM planning, and manage storage space.
- This estimation manager is a middle ware among cloud administration & energy-sensitive scheduler. This estimation will utilize some sensors i.e. Joule meter for the accurate estimation of energy in cloud servers, which utilizes a power model to surmise the power utilization of the VM or server in resource utilization.

Energy-aware VM planning for energy-mindful VMs in the server clusters is the concentration of our energy utilization display. The energy-efficient scheduler comprises basically of two modules. Distribution module and de-allocation sub-module. The distribution module is to utilize our proposed VM allotment algorithm to play out the underlying VM administration. The dynamic combination of virtual machines managed by the de-allocation module, and on account of our proposed VM de-allocation algorithm, the number of servers that can be utilized or initiated can be limited. Unused servers are closed down or go into rest mode. All the required data (the two servers and VMs running the algorithm can be recovered through the cloud LaaS Manager that likewise performs virtual machine management and split activities. This model will consider the resources demands of the request by the client requests with the quantity of VMs required and the sort of VM objects required (for instance, little, medium, extensive). Each VMI is described by a working time t_i and a max power scattering p_i . Every server or host hub (let say j) in the server center will observe a threshold limit that will be maximum power utilization constraint, marked P_{jMax} . This limit is described by the administrator of the cloud. We are expecting the isomorphic properties of the servers. Because it is difficult to stretch out this utilization conceptual model to heterogeneous servers, however including complexity and quality does not give an extra benefit. The best approach to execute is a hybrid combination of the de-allocation algorithm with the two allocation algorithms as given in Figure 2.

Energy proficiency in our suggestions is to utilize the Boxed algorithm to optimize the format of client requests for and to go for merging of *Vms* when requester is low in number, dynamic coordination is performed by the de-allocation algorithm, which reassembles the VM as much as possible, however much as could be expected to release whatever number of servers are in sleep or shutdown model.

5 Energy-saving dynamic resource allocation algorithm

The VMs are placed on the server as they arrive. They keep on running on the server until they achieve their maximum burst, and they exit from the system as the associated job ends. These opportunities for resignation are re-optimized by consolidating virtual machines with a minimum number of fully encapsulated servers in the system. The re-optimization is dependent on the de-allocation algorithm. This algorithm is an integer linear program (ILP) that uses the integration mechanism to allocate resources and then combine the similar kinds of VMs. The ILP introduces some identification of servers based on inequalities of energy usage to apply minimum de-allocation. VM merging is shown through a mathematical model. This model shows the VM merging. The merging of VMs between different servers is dependent on de-allocation from the first server after the

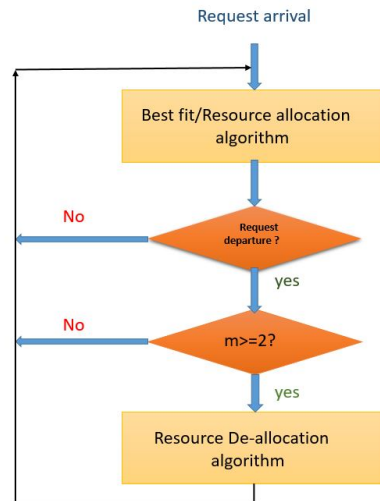


Fig 2. A hybrid combination of the de-allocation algorithm with the two allocation algorithms

calculation of average efficiency achieved using ILP formulas. This algorithm is focused to transfer or move a VM from the source server to another server who can inculcate the VM. The source node is chosen who is intended to become empty so that it can be closed or put to sleep. The target is achieved by migrating the VM to other selected target nodes (the algorithm is designed to fill them so that they are in the maximum number of virtual machines served by cloud resources until cloud capacity reaches). The algorithm will try to reach an ideal situation.

The de-allocation will be performed on a set of active servers that are denoted as m_i . Every server m_i belongs to the server list M . The constraint is on the power utilization limit which should be less than the threshold limit P_j , in such case the de-allocation or migration of virtual machines will be performed to some other server which has the capacity greater than the current power utilization greater than P_j and Max in units of m .

The issues are the number of allocation on servers which are finite however the actualization of a virtual machine is greater in number. This makes this problem NP-hard. That's why the allocation algorithm of resources is based on linear integer programming taking the current utilization energy and threshold limit as input for the de-allocation. This also takes into account the actual size of virtual machines that are associated with demanded resources. The optimality of this function can be expressed as maximizing the number of servers that are in-active along with the allocation of VM to the server to a maximum level with the help of a combined effort by resource allocation and VM- de-allocation. This will achieve the maximum level of energy efficiency at both the server level and the data center level. This can be done with the help of accurate utilization of resources with boxed allocation to ensure optimal configurations.

5.1 Working of resource allocation algorithm

Energy-efficient VM scheduler in charge of the VM management of the server center. This VM scheduler is responsible to perform the energy optimization through VM placement and de-allocation. This scheduler is essentially made out of two sub-modules. A module named as VM allocation module that is responsible for receiving VMs and allocating random servers to them and another de-allocation and migration sub-module that re-distributes the VMs to reach an optimized level of active servers. The part of the VM assignment module is to manage the underlying VM utilizing VM distribution. The consolidation and de-allocation of VMs are taken off during relocation. This module limits the quantity of active servers because of optimized VM placement. The servers that are unused or inactive are sent to the rest mode. There is another manager module that provides initial data for several active and executing VMs. This data is taken as a baseline to execute the algorithm. This algorithm

additionally executes the VM, some necessary arrangements for allocation for resources management & relocation activities. To describe the framework considered, resources are needed by the VMs that describes the client requests. The virtual machines can be divided into different types (e.g., little, medium, vast) as per their required resources list.

Every VMI is described by a required bust time denoted as t_i and a power utilization requirement denoted as p_i and every server or facilitating hub j . This server has a power utilization limit denoted as P_j , Max. This can be set up by cloud managers or administration staff. We expect that all servers are homogeneous; as stretching out the model to heterogeneous servers is inconsequential, however it will expand the difficulty and decreases the general understanding. The method to accomplish energy optimization in our proposal is to utilize a boxed model for the ideal arrangement of client requests and combining VM de-allocation and consolidation using the number of VM migrations from one server to another. The hybridized algorithm takes care of consolidation using a VM shifting algorithm and then regrouping the virtual machine to make max server free. The max server free is then shutdown or put to sleep to decrease energy utilization.

The proposed correct VM management algorithm is boxed. This algorithm incorporates substantial conditions communicated as limitations or imbalances. The goal is to manage VMs into an arrangement of boxes (servers or hubs facilitating the VMs) as per their energy utilization. Suppose n is the quantity of VM requests, for them the server available is denoted by m that are active. The servers are expected to have similar power utilization limits: P_j , Max, $\{j = 1, 2, \dots, M\}$. Every server j facilitating various VM is portrayed by its current energy usage denoted as P_j current. Keeping in mind the end goal to minimize energy usage within a servers cluster, a key choice is characterized as E_j for every server j . This choice is flagged as I if server J is hosting VMs. This flag is shown as 0 if there is no VM hosted by this server. Likewise, we characterize the bivalent variable x_{ij} to demonstrate that VMI is hosted by server j and set x_{ij} is flagged as I or x_{ij} is in off state. Generally, the target capacity to put every one of the requests in one of the n servers of the initial active list. This streamlining is liable to solve various straight imperatives as the limit reaches the servers and therefore it demands some certainties, for example, a VM must be allocated to one server or a server can just host VMs within the limit of e_j (a measure of outstanding resources):

$$\min Z = \sum_{j=1}^n e_j$$

• Every server has a power limit on its consumption p_j , Max that can't be surpassed when serving or facilitating VMs and this happens as indicated by residual power limit of all servers, described as:

$$\sum_{i=1}^m p_i x_{ij} \leq P_{jmx} \quad e_j - p_j \quad \forall j = 1, \dots, m$$

This equation shows that we have a limitation of server energy consumption that is p_i and our goal is to maximize the e_j which is the efficiency of servers.

$$\sum_{j=1}^m x_{ij} = 1 \text{ for all requests } 1 \text{ to } n$$

If the VM is assigned to exactly one server then the function value is described as 1.

• A cloud supplier needs to satisfy all resource requests inside an endorsed SLA or standard and each asked for VM will be assigned to one single server:

• For all the servers will satisfy the limit P_j , max $> P_j$, current and $P_j! = 0$. There can be a limit on the total number of servers used described as following;

$$\sum_{j=1}^m P_{j, \text{current}} / P_{j, \text{max}}$$

The inequality of power generation that needs to be minimized is shown in the following equation;

$$\sum_{j=1}^m e_j \geq \sum_{j=1}^m P_{j, \text{current}} / P_{j, \text{max}}$$

Therefore we can state that our objective function is to reach;

$$\min Z = \sum_{j=1}^n e_j$$

That is dependent upon the following limitations;

$$\sum_{j=1}^n p_i x_{ij} \leq P_j \max.e_j - P_{j,current}$$

This holds for all $j = 1, \dots, n$. The value of e_j is 1 if the server j is used by some VM otherwise its value will be denoted as 0. Similarly, x_{ij} is denoted as 1 if the virtual machine is placed in some server j . Otherwise, its value is denoted as 0.

The referenced variables are described in the following lines.

- n States the number of requests that are received and stated as VMs.
- m Describes the total servers available in the server center.
- p_i Describes the total power consumed by each VM.
- A dependent variable $x_{i,j}$ describes that VMI is utilizing server resources of a server denoted as j .
- e_j describes that a server is inactive or active
- $p_{j,max}$ describes that maximum power utilized by a server j
- $P_{j,current}$ Describes the currently used energy by the server denoted as j .

$$P_{j,Current} = P_{j,idle} + \sum_K p_k$$

- In this equation, the p_k describes the power used by a VM k that is hosted by some server. And $p_{j,inactive}$ shows that total power used by a server j , which is inactive.

- Resources of the server are already limited concerning CPU, memory and storage capacity

This limitation is shown as;

$$\sum_{i=1}^n c_i x_{ij} \leq c_j e_j$$

Where c represents the number of CPUs. c_i Represents the CPU allotment request by the virtual machine i , and the total number of CPU of the server j are shown as c_j

- Similarly, the limitations of memory and storage are also shown in the following lines.

$$\sum_{i=1}^m m_i x_{ij} \leq m_j e_j$$

- Where the m_i is the requested memory by the virtual machine i and the total number of memory available of a server are represented by m_j by a server j . Memory is associated with the CPU m .

- Storage is represented as S . and the limitation of the storage is shown in the following lines.

$$\sum_{i=1}^M s_i x_{ij} \leq s_j e_j$$

Here s_i represents the requested storage by a VM i and total storage capacity of the server j are represented as s_j .

If the constraints are met in the data center and we have enough storage, memory, and CPU then we need to manage only the energy efficiency requirement of the VMs.

6 Simulation to verify the result

Our proposed algorithm is assessed through a dedicated Java execution using the same model as the linear CPLEX. A committed test system is produced to lead the evaluations and analysis. The goal of the numerical assessment is to measure the level of expected energy saving when joining the correct VM assignment and relocation is performed in the servers with the help of the proposed migration method. The appropriate responses during the mathematical analysis will show high adaptability and low intricacy of the proposal. This algorithm in the server center also shows how to manage the incoming requests to use minimum resources. The developed simulation takes into account the process details as incoming time, burst needed and outgoing time. Along with these, the simulation will also correspond to the ending time of the VMs. In the example scenario, we take 200 servers as an input. We gather the following performance optimization factor-like reduction in the number of utilized servers (which naturally gives the power spared by using the proposed algorithm) & the execution to manage their best relocations through this algorithm. Figure 3 shows that the simulation results of migration algorithm and efficiency have been achieved.

Suppose the power utilization top P_j P_j , and the Max set to 200 watts, we consider it the principal energy to run a server farm 250 watts⁽²⁰⁾. A power estimation model tells us that we need the following energy for high, medium and low power consumption for three SPECcpu2006 workloads. (454.calculix is the workload for high, 482.sphinx is the workload for medium and 435.gromacs is the workload for short power usage) Their related power utilization is near 13, 11, 10 watts individually. The power estimation given by⁽²¹⁾ described the additional details. Another model by⁽²²⁾ described workloads (471.omnetpp for large scale VM, 470.lbm for the average and 445.gobmk for small VMs) that were assessed as power utilization of other SPECcpu2006.

Evaluated control utilization was observed to be in the range of 25 and 28 watts for these components. We will use these published results for our ease of possible analysis. We will associate the energy consumption with three different types of VMs like small, medium and large and will associate their energy utilization as 10, 20 and 30 watts respectively for small, medium and large size virtual machines. We take the incoming requests as a consistent entry rate from 1 to 3, 1 for small and 3 for large. The VM sizes are taken as uniform. We have drawn a comparative diagram between the two algorithms. The first algorithm that is originally taken in response is the best fit heuristic algorithm and then our proposed boxed algorithms. The simulation will take into account 200 servers request for resources are taken randomly from 1 to 200. The burst time need is taken as uniform from a minimum of 30s and a max of 180s. Our proposed algorithm in the achieved statistics are shown in the Figure 3. The best-fit algorithm uses more resources as compared to our proposed algorithm.

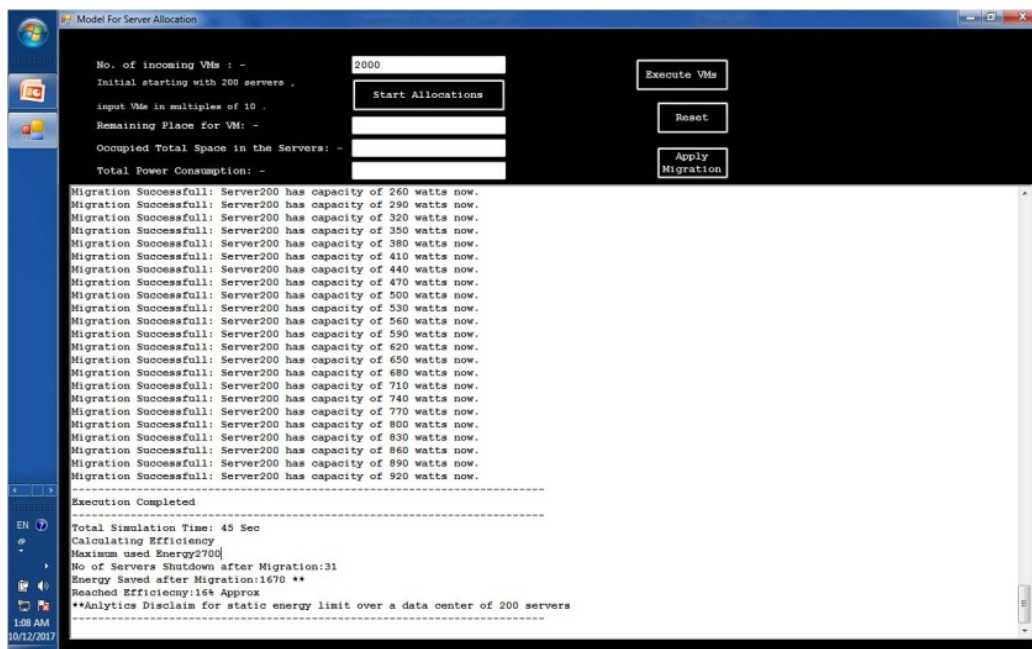


Fig 3. Simulation results of the migration algorithm and efficiency achieved

7 Proposed algorithm compared with best fit heuristics algorithm

The earlier used algorithm for virtual machine allocation is the best-fit heuristic algorithm that is described by⁽¹⁹⁾. The important facts of this algorithm are given as follows:

- It sorts the VMs and keeps the most energy-consuming VMs as a priority. This builds a decreasing stack with the least energy-consuming at the bottom. Then these VMs are allocated to the server. This is somehow the same as box packing of VMs while the boxes represent the server allocation that can accommodate the VMs with one exception that most high consumption VM is taken first.
- The topmost VM in the decreasing stack is the one that will consume the most energy. This way the smallest energy requirement VM is left in the list. When the most energy-consuming VMs are allocated then the least consumption VM is tried to make fit in the remaining slot of allocation. This process repeats for the target server and allows the freed server to go to sleep mode. This algorithm also tries to fill the maximum boxes in the server. This was algorithm was used as a comparison with our

proposed one. The result of this algorithm for the same number of server configurations will help us analyze the performance of our proposal. The results are taken to compare the exact situations of the two algorithms. Our allocation and migration algorithm will have the advantage of migrating the VMs that are not earlier addressed. Our objective was to test a benchmark case for its comparison with best fit heuristic adaptation in the service centers. We will take the best-fit algorithm as classical sub-optimal performance and our proposal as a combined hybrid adaptation to reach optimality in energy consumption.

8 Conclusion and future work

This research focused on the energy efficiency of data center through energy-efficient resource allocations. Our proposal of VMs allocation to the servers and then migration criteria is supported by the simulation results. The simulations describe that as the number of powered-on servers in the data center increases the application of algorithm ensures a higher degree of energy efficiency in the cloud server. A few issues identified with the energy efficiency and resource allotment have not been addressed in this theory. The potential future dimensions incorporate the following:

- Admission control components are vital to choosing which clients' virtual machine is to service. This system will be founded on a transaction procedure to propose an elective scheduling method for incoming VMs.
- Load expectation methods assume an essential part to anticipate the general load in the framework. As future work, it is necessary to upgrade with forecast methods to additionally enhance security.
- Most research on resource planning for cloud conditions concentrate on computational assets. There is a need to investigate the network connections inside the network for energy efficiency.

Acknowledgment

The research work is financially supported by National High Technology 863 Programs of China (No.2015AA124103) and National Key R&D Program (No. 2016YFB05502001) and carried out in State Key Laboratory Intelligent Communication, Navigation, and Micro-Nano system, Beijing University of Posts and Communications, Beijing.

References

- 1) Laszewski V, Diaz G, Wang J, Fox F, C G. Comparison of multiple cloud frameworks. *IEEE*. 2012;p. 734–741.
- 2) Wang L, Khan SU. Review of performance metrics for green data centers: a taxonomy study. *The journal of supercomputing*. 2013;63(3):639–656.
- 3) Xavier MG, Neves MV, Rossi FD, Ferreto TC, Lange T, Rose CAD. Performance evaluation of container-based virtualization for high performance computing environments. *Parallel, Distributed and Network-Based Processing (PDP)*. 2013;p. 233–240.
- 4) Mishra SK, Puthal D, Sahoo B, Jayaraman PP, Jun S, Zomaya AY, et al. Energy-efficient VM-placement in cloud data center. *Sustainable Computing: Informatics and Systems*. 2018;20:48–55. doi:10.1016/j.suscom.2018.01.002.
- 5) Pavithra B, Ranjana R. Energy efficient resource provisioning with dynamic VM placement using energy aware load balancer in cloud. *IEEE*. 2016;p. 1–6.
- 6) Zakarya M, Gillam L. Energy efficient computing, clusters, grids and clouds: A taxonomy and survey. *Sustainable Computing: Informatics and Systems*. 2017;14:13–33. doi:10.1016/j.suscom.2017.03.002.
- 7) Liu XF, Zhan ZH, Deng JD, Li Y, Gu T, Zhang J. An energy efficient ant colony system for virtual machine placement in cloud computing. *IEEE transactions on evolutionary computation*. 2016;22(1):113–128.
- 8) Kansal NJ, Chana I. Energy-aware Virtual Machine Migration for Cloud Computing - A Firefly Optimization Approach. *Journal of Grid Computing*. 2016;14(2):327–345. doi:10.1007/s10723-016-9364-0.
- 9) Zhan ZH, Liu XF, Gong YJ, Zhang J, Chung HSH, Li Y. Cloud computing resource scheduling and a survey of its evolutionary approaches. *ACM Computing Surveys (CSUR)*. 2015;47(4):63–63.
- 10) Agarwal D, Jain S. Efficient optimal algorithm of task scheduling in cloud computing environment. *arXiv preprint*. 2014;p. 1404.2076.
- 11) Alapati S, Kumar P, Palani GS, Ravindran R, Sadasivam SK. U.S. Patent No. 8,914,515. Washington, DC: U.S. Patent and Trademark Office. 2014.
- 12) Belady C, Rawson A, Pfleuger J, Cader T. Energy efficient grid data center power efficiency metrics: PUE and DCIE. 2008.
- 13) Beloglazov SM, Beloglazov GS. Electrochemical Hydrogen in Metals - Can we Effectively Control this Harm? *Solid State Phenomena*. 2014;225:1–6. doi:10.4028/www.scientific.net/ssp.225.1.
- 14) Beloglazov A, Buyya R, IEEE. Energy efficient allocation of virtual machines in cloud data centers. In Cluster, Cloud and Grid Computing (CCGrid). In: and others, editor. 2010 10th IEEE/ACM International Conference on. IEEE. 2010;p. 577–578.
- 15) Bertran R, Becerra Y, Carrera D, Beltran V, Gonzalez M, Martorell X, et al. Accurate Energy Accounting for Shared Virtualized Environments using PMC-based Power Modeling Techniques. .
- 16) Bohra AEH, Chaudhary V. VMeter: Power modelling for virtualized clouds. In: and others, editor. In Parallel & Distributed Processing, Workshops and Phd Forum (IPDPSW). IEEE. 2010;p. 1–8.
- 17) Chen K, Hu C, Zhang X, Zheng K, Chen Y. Cloud Computing. *IEEE Network*. 2011;p. 4–4.
- 18) David H, Fallin C, Gorbatoev E, Hanebutte UR, Mutlu O. Memory power management via dynamic voltage/frequency scaling. In: and others, editor. In Proceedings of the 8th ACM international conference on Autonomic computing. ACM. 2011;p. 31–40.
- 19) Metzler J. Virtualization: benefits, challenges, and solutions. Riverbed Technology. San Francisco. 2011.
- 20) Nowak A, Leymann F, Schumm D, Wetzstein B. An Architecture and Methodology for a Four-Phased Approach to Energy efficient Business Process Reengineering. In: and others, editor. Proc. of the ICTGLOW. 2011.

- 21) Rabkin I, Stoicaand M, Zaharia, UC Berkeley Reliable Adaptive Distributed Systems Laboratory. Above the Clouds: A Berkeley View of Cloud Computing. In: and others, editor. Tech Report. 2009.
- 22) Pierson JM, Hlavacs H. Introduction to Energy Efficiency in Large-Scale Distributed Systems. In: and others, editor. Large-Scale Distributed Systems and Energy Efficiency: A Holistic View. 2015;p. 1–16.