Efficient Energy Management in Cloud Data center using VM Consolidation

Swarupa Pattanaik

Department of Computer Science and Engineering
National Institute of Technology Rourkela

Efficient Energy Management in Cloud Data center using VM Consolidation

Dissertation submitted in partial fulfillment

of the requirements of the degree of

Masters of Technology

in

Computer Science and Engineering

by

Swarupa Pattanaik

(Roll Number: 214CS1534)

based on research carried out

under the supervision of

Prof. Pabitra Mohan Khilar

May, 2016

Department of Computer Science and Engineering
National Institute of Technology Rourkela

Department of Computer Science and Engineering
National Institute of Technology Rourkela

Prof. Pabitra Mohan Khilar
Professor

May 25, 2016

Supervisor's Certificate

This is to certify that the work presented in the dissertation entitled *Efficient Energy Management in Cloud Data center using VM Consolidation* submitted by *Swarupa Pattanaik*, Roll Number 214CS1534, is a record of original research carried out by her under my supervision and guidance in partial fulfillment of the requirements of the degree of *Masters of Technology* in *Department of Computer Science and Engineering*. Neither this dissertation nor any part of it has been submitted earlier for any degree or diploma to any institute or university in India or abroad.

Pabitra Mohan Khilar

Dedicated to my family...

Declaration of Originality

I, *Swarupa Pattanaik*, Roll Number *214CS1534* hereby declare that this dissertation entitled *Efficient Energy Management in Cloud Data center using VM Consolidation* presents my original work carried out as a Postgraduate student of NIT Rourkela and, to the best of my knowledge, contains no material previously published or written by another person, nor any material presented by me for the award of any degree or diploma of NIT Rourkela or any other institution. Any contribution made to this research by others, with whom I have worked at NIT Rourkela or elsewhere, is explicitly acknowledged in the dissertation. Works of other authors cited in this dissertation have been duly acknowledged under the sections "Reference" or "Bibliography". I have also submitted my original research records to the scrutiny committee for evaluation of my dissertation.

I am fully aware that in case of any non-compliance detected in future, the Senate of NIT Rourkela may withdraw the degree awarded to me on the basis of the present dissertation.

May 25, 2016
NIT Rourkela

Swarupa Pattanaik

Acknowledgment

It is a great pleasure to express a my deep sense of thanks to GOD for giving me the strength to complete my thesis.

I would like to thank and express the deep sense of gratitude to my philosopher, guide Prof. Pabitra Mohan Khilar. His dedication and keen interest above all overwhelming attitude to help his student had been solely and mainly responsible for completing my work.

I would like to thank our H.O.D Prof. Santanu Kumar Rath for his keen interest at every stage of my research work. His prompt inspiration and timely suggestion have enabled me to complete my thesis.

I thank profusely all the staff, phd scholars of NIT Rourkela for their kind help and cooperates through out my research.

I am extremely thankful to all my classmates for providing me necessary suggestion during my research pursuit.

I would like to gratefully acknowledge various people who have journeyed with me in recent years. I owe an enormous dept of gratitude to my hubby Mr. Lalitendu Bal and all my family members for their constant support and encouragement.

May 25, 2016 *Swarupa Pattanaik*
NIT Rourkela Roll Number: 214CS1534

Abstract

Cloud computing is a model which can fast provisioned and released the computing resources by using minimum number of management effort. This can be done by the user without doing any communication with the cloud service providers. Cloud provide the computing resources , on-demand network access which is pooled together and it can be provisioned dynamically according to the user needs. Due to the large application, more number of computing nodes are required. A large amount of electrical energy is consumed due to the establishment of the data center. There is a problem of carbon dioxide emissions and increasing cost of operation due to the formation of large data center. A consolidation of virtual machines technique is proposed in our thesis to reduce the energy consumption and to maximize the utilization of the computing resources in the data center. Several virtual machines are taken together into a single physical machine in the consolidation technique and it helps to decrease the consumption of energy by putting idle server into inactive mode. A number of active hosts is minimized by continuously reallocating VMs using live migration. In each migration, Service Level Agreement(SLA) violations may occur, hence it is required to reduce the number of migrations.

In order to satisfy quality of services in cloud computing environment, our proposed techniques mainly performs the following functions:(i)reducing the consumption of energy, (ii) minimize the number of migrations and (iii) minimize the percentage of SLA violations. Initially we detect whether any host is overloaded or not. The Overloaded host is detected by considering CPU utilization as a threshold Value. If an overloaded host is detected then some virtual machines are migrated from it by using VM selection policy. After selection of the VMs, the next step is to place the new VMs. For VM placement, the greedy algorithms such as Best Fit Decreasing(BFD) and Modified First Fit Decreasing(MFFD) are used in this thesis. The proposed techniques are compared with the existing EEDVM and PALVM techniques. Using proposed AUTREC technique there is 8% improved in energy consumption, 3% in number of migrations, 10% in SLA violation and 12% in host shutdown as compared to EEDVM technique. Using proposed DUTREC technique there is 9% improved in energy consumption, 6% in number of migrations, 20% in SLA violation and 13% in host shutdown as compared to PALVM technique.

***Keywords*: *VM Consolidation*; *Virtualization*; *VM Migration*; *VM placement*.**

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Introduction

Cloud computing is a style of computing where a collection of IT enable capabilities are delivered as a service using some technologies to the customers. This cloud definition is defined by Gartner[1].

The focus of cloud computing is to design the computing system, develop the applications and take the advantage of existing services for buildong software. Dynamic provisioning is the technique applied to the services, computation capability, storage management and networking[2] in cloud technology. The resources are provided by the cloud vendors as pay per use basis and it is available on the internet. The provided service can be any type application such as email, video conferencing etc.

The computing resources such as servers, networks, storage, applications and services are pooled together in cloud computing. The model of cloud provide the convenient, ubiquitous, on-demand network access to these resources which can be fast provisioned and released with minimum effort of the management or the interaction of the service provider, according to the NIST[3] definition.

Section 1.1 introduces the basics of cloud computing including the characteristic of cloud, cloud models, architecture and cloud data center. section 1.2 specifies motivation and challenges of the consumption of energy, migrations number and viloations SLA. The objective of the research work is specified in section 1.3, section 1.4 introduces the problem statement, section 1.5 specifies thesis contribution, The organization of the thesis is specified in section 1.6 and finally chapter 1 is concluded in section 1.7.

1.1.1 Importance of Reducing Energy Consumption in Data center

A data center consists of large number of computing resources. The cost of the energy and the infrastructure would contribute 75 percentile and 25 percentile would contribute by the Information Technology of the overall cost of operating a data center, given by American Society of Heating, Refrigerating and Air-Conditioning Engineers(ASHRAE)[3].

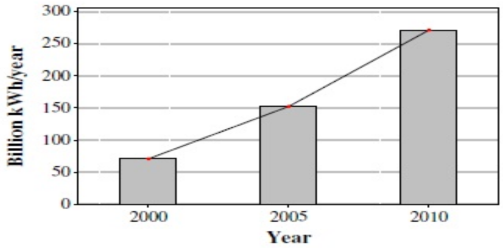

Figure 1.1: Energy Consumption 2000-2010 in world wide[3]

An enormous amount of energy is consumed in a cloud data center because lots of resources are used but they are not properly utilized. Even most of the servers operated at around ten to fifty percent of their full capacity. This leads to more expenses because when servers operate at a lower level of utilization they consumed more power than when they are utilized at hundred percent. Also if the servers are completely idle, they consume around seventy percent power of their highest level of power. Thus for energy efficient computing the focus must be on utilizing the servers at their full capacity and they should also not be kept completely idle. Along with the direct energy consumption of the resources, there is an additional expense of energy through the cooling systems that are required in each data center to keep all the systems cool. For every single watt of power consumed, it requires around 0.5 to 1 watt of power by the cooling systems . Additionally, the more the number of computing resources used, the higher is the carbon dioxide emissions. This carbon dioxide leads to the greenhouse effect.

1.1.2 Cloud characteristics

Cloud characteristics are presented as follows[3]:

- **On-demand self-service:** The cloud services can be provisioned by the client automatically without any communication between the cloud service providers.
- **Dynamic Provisioning:** The dynamic scaling of cloud needs to be done while high level of security and reliability are maintained. Based on current demand requirement it allows the user to provision the services.
- **Broad network access:** The customers can access the cloud services by using their laptop, tablet, mobile, workstation etc. The cloud services are available over the network.
- **Resource pooling:** In cloud computing, the resources are given to multi user. These resources are polled together. According to user demands the virtual resources are assigned or reassigned dynamically.

- **Rapid elasticity:** Resources can be available unlimited and accurate quantity at any time. The resources capabilities can automatically scale up and scale down according to users demand.

- **Measured service:** The resources in a cloud environment can be controlled, monitored and providing the transparency to the users.

1.1.3 Cloud Model

The cloud model is categories into two parts[3].

1. **Cloud Service Model**

   Cloud provide three types of service model. Figure 1.4 shows the different service model of cloud computing. The models are given below:

   - **Infrastructure as a Service (IaaS):**

     It provides the service such as processing, storage, memory, network bandwidth to the users. The users can run or deploy any application or operating systems on the underlying cloud infrastructure[3][4].

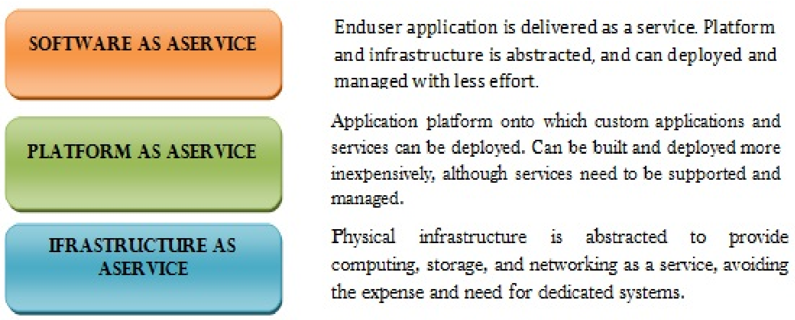

Figure 1.2: Cloud Service Model

   - **Platform as a Service (PaaS):**

     On the cloud environment, this layer provides the platform for building and executing the application. For example, Google App engine provides the environment for the user to build and deploy their application. This layer is built above the infrastructure layer of a cloud environment.

   - **Software as a Service (SaaS):**

     In this layer, application software such as Google Docs, game, e-mail etc. is provided as a service. Users are unable to purchase a high-cost application for their personal use. So the cloud computing provides an application in a few cost to the users. Users do not need to maintain the underlying cloud infrastructure.

2. **Cloud Deployment Model**

   Four types of deployment models are provided by the cloud[3].

   - **Private cloud:**

     In a private cloud, the services can be available within a particular organization. The users who belong to that organization is able to access the service. The organisation itself or by the third party the private cloud is managed.

   - **Public cloud:**

     The services are available to the public, in the public cloud. It can be managed by academic, any business, or government organization.

   - **Community cloud:**

     The cloud is a combination of organizations from a specific community with a common concern. It can be facilitated either internally or externally.

   - **Hybrid Cloud:**

     Two or more model are combined and make hybrid cloud. Many advantages are provided by this cloud due to multiple models.

1.1.4 Cloud Architecture

Cloud architecture is depicted in figure 1.3.

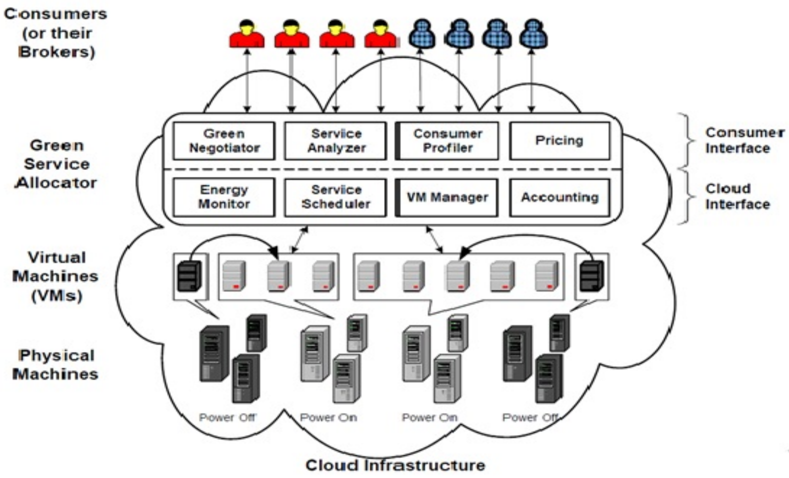

Figure 1.3: Cloud Architecture[2]

The cloud architecture consists of four important entities.

1. **Consumers/Brokers:**

   The consumers or their brokers submit the request to the cloud from anywhere in the world.

2. **Green Service Allocator:**

   GSA act as an interface between the cloud infrastructure and the customers. The GSA need the following component for supporting of efficient management of the resources.

   - **Green Negotiator:** To finalize the SLAs between the consumer and the CSP, the green negotiator negotiates with specified penalties and prices. The SLA is depending on the schemes for energy saving and QoS need of consumer.
   - **Service Analyser:** The work of service analyser is to analyses the service requirements and interprets it after accepting of the submitted request.
   - **Consumer Profiler:** It collects information of users. Give special prioritized and privileges to important consumers/customers over other consumers.
   - **Pricing:** The charging of service requests depends upon the management of supply and demand of the computing resources.
   - **Energy Monitor:** To make energy-efficient resource allocation decisions, the energy monitor observes and provide the information of the consumption of energy caused by the physical machines and virtual machines.
   - **Service Scheduler:** Determine resource entitlements for the allocated VMs and make the assignment request to VMs.
   - **VM Manager:** The virtual machine manager keeps the information of the virtual machines availability and the resource utilization. It is the duty of the VM manager to provision of the new VMs and the VMs which are migrated across physical machines.
   - **Accounting:** The virtual machines check the utilization of the resources and calculates the costs of the resources used. To maximize the allocation of the resource decisions, the historical data can be used.

3. **VMs:** A number of virtual machines can be stared and stopped dynamically and it depends on the requests that are coming to a virtual machine. The partitions of the resources are configured on the same physical machine and it depends on the need of service requests.

4. **Physical Machines:** The hardware infrastructure is provided by the physical machines for building up the virtualized resources to meet the service demands.

1.1.5 Cloud Data center

Under cloud computing, many data centers are operated by hosting a variety of applications. The requirement to deal with this application in the data center creates challenges for on-demand resources provisioning and allocation of time-varying workloads[5]. A data center consumes energy an average of 25,000 households. As the costs of the energy increases, there is a need of optimizing the resource usage in a data center. So that the energy efficiency can be maintained along with the service level performance. Data centers are costly to maintain as well as unfriendly to the environment[6]. Figure 1.1 depicts the world wide energy consumption in the year 2000 to 2010. To minimize the energy consumption different approaches has been proposed such as DVFS, Virtualization, VM Migration server consolidation etc. The approaches are described in the following:

- **DVFS**

  It is a hardware technology[7]. The voltage and frequency of the processor can be automatically adjust in the running time. No need to start the power supply again and again if the system has DVFS configuration. The consumption of power by the CPU can be calculated by:

  $$P = V^2 * F * C \tag{1.1}$$

  Where V is the voltage, P is the power, F is the operating frequency and C is the capacitive load of the system. By reducing the supply voltage the power can be conserved. Automatically the voltage is dropped by reducing the frequency of CPU.

- **Virtualization**

  The classification of the virtualization technique is depicted in figure 1.4. Virtualization allows sharing of resources among multiple virtual machines in a single physical machine. The consumption of energy can be minimized by switching the idle physical machines to inactive state. The virtual machines number can be consolidated to minimum number of physical machines by using the live migration process[8]. The performance degradation occurs when the usage of resources maximized due to the increasing demand of applications.

  The classification of virtualization technology is categorized into two section.

  1. **Full virtualization:**

     In full virtualization the physical machines are heavily loaded due to the reliable compatibility. One merit of full virtualization is that it supports in any hardware environment[9].

  2. **Para virtualization:**

     The operating system may be shared with the resource hardware but the kernel of the system has to be changed. The hardware support in para virtualization is not sufficient. But the para virtualization is better than the full virtualization if we consider the performance parameter.

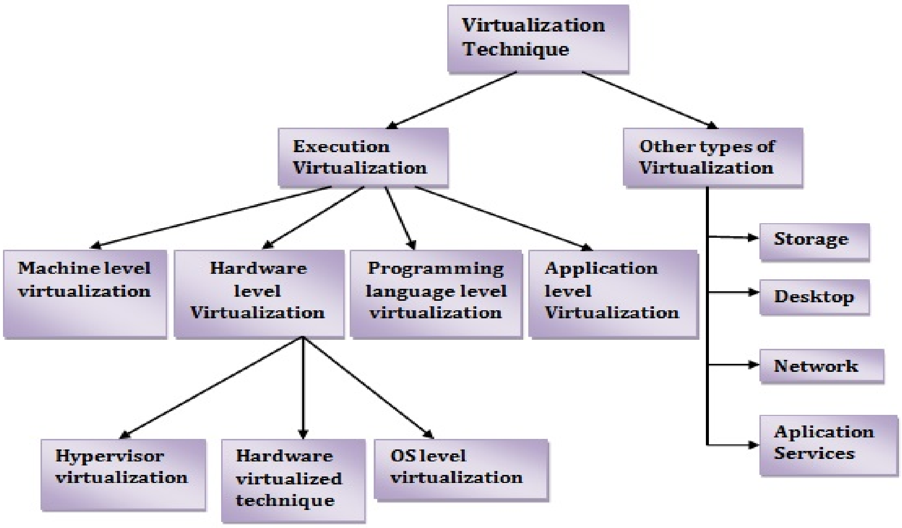

Figure 1.4: Classification of Virtualization technique

In the architecture of virtualization the virtual machine monitor(VMM) is the core implementation. It is the responsibility of the VMM to manage the hardware resources and provide virtualization. Migration is one of the important concept of virtualization and its meaning is to transferring of VMs from one host to another.

- **VM Migration**

  The virtualization system has a unique capability, that is known as virtual machine migration[9] [10]. It allow applications to be moved from one host to another transparently and continue its process along with no loss of progress after migration. The migration process can broadly categorize into off-line migration and on-line migration. An action can not be taken by the user in case of off-line migration because the current user's state must be shutdown before migration. At the user unknown situation the task can be migrated. There is no requirement of shut down or suspend the virtual machines.

  – **Categories of VM Migration**

    The virtual machine migration technique is classified into 3 types[11].

    * **Energy Efficient Migration Techniques:** For saving of energy this migration technique is required by optimizing the utilization of the resources. A huge amount of power can be consumed by the data center although the utilization level is low.
    * **Load Balancing Migration Techniques:** The scalability of the physical servers can be maximized by distributing the load has across the physical host. This can be done by using load balancing migration techniques in the cloud system.

* **Fault Tolerant Migration Techniques:** The virtual machine continue its execution process even if any parts of the system fails. This can only be possible by using fault tolerance techniques.

1.2 Motivation and Challenges

- ### Motivation

  The major motivation that leads to study the energy management in cloud data center are listed below:

  - **Energy saving**

    There is a negative impact on the environment due to the huge consumption of energy and emissions of carbon dioxide. This is estimated that to be 2 percentile of the global emission. Due to the under utilized resources and their inefficient use a large amount of energy is waste. There is a requirement of 0.5 to 1 watt of consumption of power due to the cooling and storage system.

  - **VM Migration**

    The performance degradation arises in a data center as the demand of applications gets maximized and this is due to the downtime arises in the migration process. The more number of migrations will cause more downtime. So minimize the number of migrations in the VM migration technique is a challenging task.

  - **SLA Violation**

    The allocation of the CPU is not performed even if it is demanded. This situation is known as violations of SLAs. The low SLA value give the better quality of services. Hence to minimize the SLA percentage of the SLA violation.

- ### Challenges

  1. The execution of the different workloads is the main reason of low utilizations of the data center. The expected quality of service is not provided by the data center due to the change in non-significant workload. The performance degradation is also occurs.

2. Due to the background tasks, distributed data bases or file systems, the physical machines in a non virtualized data center are totally idle.
3. The application performance will degrade and also SLA violation occurs if the physical machine does not contain idle or enough nodes to satisfy the VM needs.
4. To accomplish a trade-off between the application performance and energy efficiency.
5. To optimize the cost and the number of migration.

1.3 Objective

To reduce the consumption of energy in the data center the following objectives are undertaken:

1. To design and evaluate an energy efficient technique to reduce the consumption of energy in cloud data center.
2. The quality of service can be maintained for the cloud computing and it can be measured in the form of SLA(service level agreement). The main goal is to reduce the consumption of energy.
3. To validate our proposed technique with cloudsim simulator version 3.0.3.
4. To evaluate and compare the proposed technique with some existing approaches which is defined in chapter 2.

1.4 Problem Statement

To reduce the data center energy consumption, virtual machines of overutilized host are migrated to another host and the hosts which are underutilized goes to an inactive state. The VMs of the underloaded host are also migrated to another host. The downtime of migration process maximized due to the more number of VM migrations. Each migration causing SLA violation. To achieve low energy consumption, minimum number of VM migrations and SLA, host utilizations are selected and perform optimum number of migrations.

1.5 Performance Parameters

The following performance parameters are considered to evaluate the performance of the proposed techniques and comparing with the existing techniques such as EEDVM and PALVM.

- **Energy Consumption**

  The total consumption of energy of a data center can be calculated[12] by,

$$total\_energy\_consumption = \sum_{i=1}^{m} energydc(i) \tag{1.2}$$

  Where energy(i) is the energy of i^th data center. The energy of the data center can be calculated by the following equation:

$$energydci = ((dcenmax_i - dcenidle) * maxutildc(i)) + dcenidle_i; \forall i \in 1, ..., m \tag{1.3}$$

  Where $dcenmax_i$ is the maximum energy consumed by i^th data center. utildc(i) is the utilization of a data center. $dcenidle_i$ is the idle data center. The utilization of a data center can be calculated as the fraction of all VMs requirements of all the placed VMs in the data center to the total resources of the data center.

$$utildc(i) = \frac{\sum_{k=1}^{n} CVM(k) * V_{ik}}{CDC(i)}; \forall i \in 1, ..., m \tag{1.4}$$

  where $V_{ik} = 1$ if VM is placed in data center. CVM is the capacity of the virtual machine. CDC is the capacity of the data center.

- **VM Migration**

  To enhance the resource utilization and energy efficiency virtualization technique is used[6]. This technique allows a user to run a number of applications as the virtual machine on the same physical machine. The utilization of a virtual CPU can be calculated as

$$V_{cpui} = \frac{VM_{MIPSi}}{HOST_{MIPSi}} \tag{1.5}$$

$$CPU_{utilization} = \frac{\sum V_{cpui}}{n} \tag{1.6}$$

  Where V_{cpui} is the utilization of a virtual machine. $CPU_{utilization}$ is the utilization of the host machine. VM_{MIPSi} is the MIPS required by a VM. $HOST_{MIPSi}$ is the MIPS required by a host machine. n is the total number of host machines.

- **SLA Violation Calculation**

  The SLA can be calculated as:

$$SLA = \frac{\sum Requested_{mips} - \sum Allocated_{mips}}{\sum Requested_{mips}} \tag{1.7}$$

$Requested_{mips}$ is the sum of requested MIPS. $Allocated_{mips}$ is the amount of allocated MIPS from the requested MIPS.

1.6 Thesis Contribution

The key contributions of the thesis are as follows:

- A classification and survey of different energy efficient techniques.
- Analysis of consumption of energy, migrations number and percentage of SLA violations in our proposed techniques AUTREC and DUTREC.
- The proposed techniques give the better result over existing EEDVM and PALVM techniques.
- Our proposed technique is feasible for the real cloud data center.

1.7 Thesis Organization

The organization of the thesis is given into 5 chapters. They are:

- **Chapter 1** gives the introduction of cloud computing with its characteristics, models, architecture, and about cloud data center. Furthermore motivation and challenges, the objective of our research work, problem statement are included here.
- **Chapter 2** specifies the literature review where we have discussed some existing works on different approaches of energy consumption technique. The classification of scheduling technique is also included into this part.
- **Chapter 3** presents a model for VM migration technique and by using an existing adaptive threshold IQR method to find out a host is overload or not . Once an overloaded host is detected, then the migration can be done by using the MMT policy. In this chapter, an algorithmic approach is proposed that calculate the consumption of energy of a given data center including migrations number and SLA violations.
- The work in **chapter 4** is same as chapter 3. Only the proposed approaches are different. This chapter presents a dynamic utilization threshold method for selection of virtual machines . The migration can be done by the MMT policy. An algorithmic approach is proposed that calculate the consumption of energy of a given data center including the migrations number and SLA violations.

- **Chapter 5** give the suggestion foe the future work on the efficient utilization of energy in a data center with minimum migrations number and violation of SLA. It also conclude the work done in chapter 3 and 4, highlighting the contributions.

1.8 Summary

In this chapter, we introduce about the basics of cloud computing such as the characteristics, model, architecture, cloud data center. We also discuss the motivation and challenges and of cloud computing. The objective of the research work is given in this chapter.

Chapter 2

Literature Review

2.1 Introduction

Cloud computing is a high-performance computing infrastructure for new generation computing and service distribution[16]. To reduce the consumption of power of the computational cloud environment, it is required to schedule VMs efficiently[13]. VMs are the software implementation of the particular computer system and provide sharing of physical hardware resources to the users. Cloud data center frequently uses the virtual machine migration and allocation techniques to provide efficient computing. Basically, cloud data centers are composed of a number of virtual machines and these virtual machines are free to select different CPUs.

It is a very complex and challenging issue to save energy because the development of data and the computation applications are growing so rapidly. Large number of servers and disk are required to process the data and its applications. To achieve the resource utilization and the efficient processing is the main task of Green cloud computing. It also reducing the consumption of energy[2]. Our work is based on consolidation of virtual machines and retaining a strict SLA to consumers while considering the minimization of energyconsumption[19]. and management by performing static and dynamic consolidation of VMs and servers. Different types of algorithms are used to minimize the consumption of energy in a cloud data center. Such as Dynamic scheduling, Live migration, Single Threshold, Double Threshold, Sleep, Genetic algorithm. Figure 2.1 depicts the existing techniques that are used in our proposed technique.

- **Live Migration:**

  The process of transferring an application or running VM between the physical machine is known as live migration. In this process the the transferring is done without disturbing the client or the application. The VM consists of memory, storage and network which are transferred from the guest machine to the destination machine. The network file system is the place where the virtual disks are stored. It is a shared storage and accessible throughout the network[14]. Because of three reasons the VM

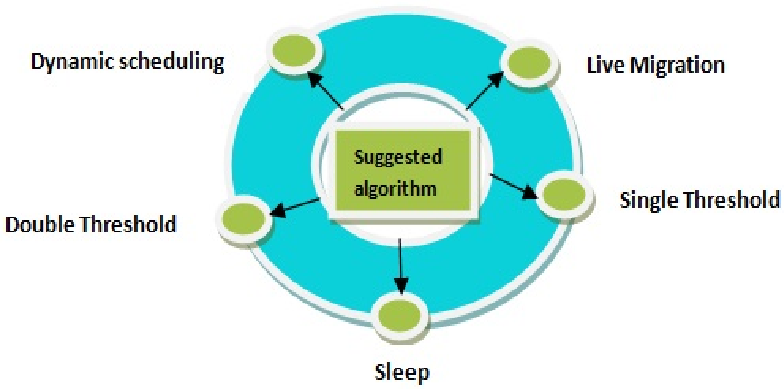

Figure 2.1: Suggested algorithm [20]

live migration can be done, they are (i) requirement of resources, (ii) consumption of power and (iii) VM affinity. If we consider the migration for load balancing then it is categories into two parts. Overloaded VM and Underloaded VM. In case of underloaded host, all the virtual machine are migrated to another host and the underloaded host is switched off to conserve power. In case of overloaded host, some VM are migrated to another host. If the live migration process is carried out repeatedly then it leads to performance degradation[11].

- **Server Consolidation**

  Server consolidation is a key technology of virtualization. This technology is used to minimize the number of physical servers by making a group of virtual machines, in order to reduce server collapse in a data center[15][7][16]. The energy consumption can also be reduced by putting the idle physical server into sleep mode using the server consolidation technique. Server consolidation can be either static or dynamic.

  - **Static consolidation:** In this process, there is a chance of resource wastage. Because the VM manager allocates physical resources to the virtual machines based on peak workload demand.
  - **Dynamic Consolidation:** The utilization of data center resources is better in dynamic consolidation as compared to static consolidation because according to the current workload demand the VM capacities are changed by the VM manager.

Section 2.1 introduces the different reason of consumption of energy problem in the data center. The classification of different energy management techniques define in section 2.2. The classification of VM scheduling techniques are given in section 2.3. Section 2.4 gives the summary of this chapter.

2.2 Classification of Energy Management Technique

Figure 2.2 depicts the energy management techniques in the cloud data center[9].

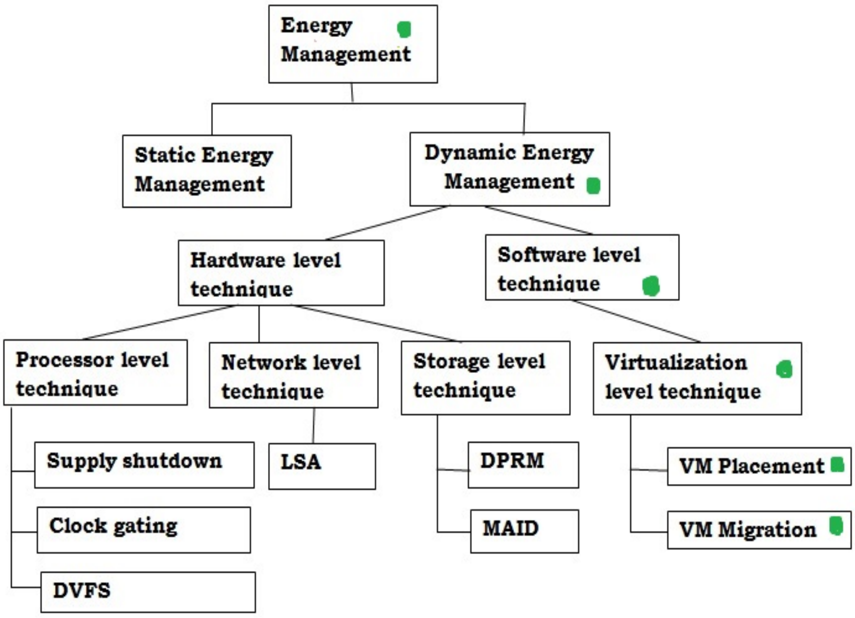

Figure 2.2: Energy management technique[9]

- **Static Energy Management Technique**

  It is an expensive technique for saving energy. For energy conservation at the data center, this technique uses network component having low power. The cloud services that are running in the data center is volatile.

- **Dynamic Energy Management Technique**

  Based on the current resource requirements of the requested service, the dynamic energy management technique dynamically reconfigures the system. This technique utilizes the hardware component and the software methodologies at the data center to reduce the consumption of energy.

  1. **Hardware Level Energy Management Technique**

     To reduce the consumption of energy at the data center, also on each hardware component the energy management technique has to be applied.

     – **Processor Level Energy Management**

     The consumption of energy in the processor is consists of static and dynamic power. The power drawn at the idle time of the processor then it is known as static power. Depending upon the utilization of the various resources the dynamic power is calculated.

(a) **Network level clock gating:** The idle component is also consume some amount of power. So to reduce the supply voltage and the frequency also.

(b) **DVFS:** According to the processing request the voltage and frequency of a processor can automatically be adjusted.

(c) **Supply Shutdown:** To minimize the power dissipation, the inactive components are switched off.

– **Network Level Energy Management**

It is important to address the consumption of energy done by the network devices because the data center and the network infrastructure size is exploding day by day.

(a) **LSA:** For dynamically changing the link speed of the network one technique is used that is known as LSA(Link State Adaption). To reduce the consumption of energy this method is used.

– **Storage Level**

(a) **DPRM:** It stands for Dynamic Rotation Per Minute. The speed of the disk rotation can be changed dynamically per minute.

(b) **MAID:** The Massive Array of Idle Disk is used to place the copies of the required files on a subset of files called the cache disk and the non-cache disks are spin down for saving of energy.

2. **Software Level Energy Management Technique**

– **Virtualization Level**

To achieve better utilization of resources virtualization technology is used in the data center for conservation of energy. Multiple virtual machines can share the resources in a single physical machine bu using virtualization. The data center's energy consumptions depends on resources that are efficiently used. For better utilization of resources VM consolidation technique is used in the data center.

(a) **VM Placement**

The energy performance trade-of can be managed by using the VM placement technique. In this technique one mapping is done between the virtual machine and the physical machine. The placement technique consists of two steps.

(i) According to the capacity requirement resources can be provisioned for the VMs for the corresponding application. (ii)Placement of VMs to PMs is done actually.

(b) **VM Migration**

The migration of virtual machine can be done in the data center by considering a number of concept such as server consolidation, hotspot mitigation, fault tolerance, load balancing etc. By using the VM migration technique the management of resources can also be done. The migration process can be carried out either offline or online[9][10].

2.3 Classification of VM Scheduling Technique

Figure 2.3 shows the different approaches of VM scheduling.

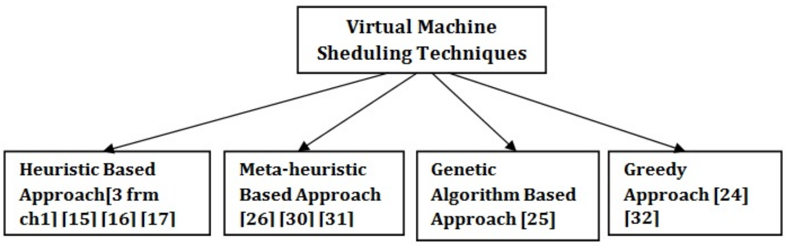

Figure 2.3: Different approaches of VM scheduling

- **Heuristic based approach**

  To reduce the consumption of energy, operating costs, migrations number as well as the SLA violation different heuristic-based approach is used in the cloud research area. The author in[17], proposed a heuristic based resource allocation VM selection and VM allocation method. The goal of the proposed technique is to minimize the energy consumption and the operating costs along with satisfying the client level SLA. In their work, they considered the VM resource allocation problem into MDBP (multidimensional bin packing) problem. In this paper, they provide an architectural framework for optimal resource allocation and management. They also develop efficient algorithms for consolidation of VM and mapping hosts to minimize the consumption of energy and operational costs.

  Chaima Ghribi et al.[18] present two algorithms for efficient energy VM scheduling in the cloud data center. They are the exact allocation and exact migration. The exact allocation is the extension of Bin-packing approach[19][20]. The major goal of this algorithm is to minimize the energy consumption in the data center. They compared their algorithm with the BFD (best fit decreasing), to solve the bin packing problem and shown the efficiency of consolidation of VM and minimize migration. They manage the VM consolidation in server side and also reduce the energy cost of the migration.

The authors in[8], have proposed an Adaptive Migration threshold based approach with Minimum Migration Time (MMT) technique which reduces the number of migrations & improves Load balancing in data centers & reduces SLA violation. They have simulated their proposed approach in cloudsim. The proposed mechanism can reduce the consumption of energy, reduces the number of migrations and also it meet with the service level agreement.

The problem of resource management in the virtualized cloud data center have been addressed byin[21]. The goal of the proposed technique is to solve the problem of resource allocation of the VM which are selected for migration. By using one resource allocation method TOPSIS and TPSA method the above problem can be solved. To minimize the energy consumption, the number of VM migrations and SLA violations, both the techniques are used. A methodology is also proposed to find the Available capacity(AC), Migration delay(MDL) and TOPSIS-Available Capacity-Number of VMs-Migration Delay(TACND) method.

In the paper[21], the main goal of the author is to solve the problem of resource allocation. They have proposed different method to find the solutions such that TOPSIS, SLA-Aware allocation, Available capacity, Migration delay, TOPSIS-Available capacity, number of VMs migration delay techniques. These solution are used to reduce the consumption of energy, migrations number and violations of SLA.

The authors in[22], proposed an energy efficient task consolidation algorithm named as max-max-util to reducing the consumption of energy in the cloud computing system. The authors of this paper make a comparison of their proposed approach with the existing energy aware task consolidation heuristics by taking a different number of tasks. The cloud scheduling model which is captured by the authors.

A technique is proposed by Anton Beloglazov and Rajkumar Buyya[23] for virtual machine consolidation. Their concept is based on the adaptive utilization threshold which is a heuristic-based method. The proposed algorithm reducing the consumption of energy and service level agreements (SLA). They evaluate their technique across different types of workloads using a thousand of PlanetLab servers. They have compared their result with the existing technique by using cloudsim simulator.

In the paper[24], the authors proposed an approach based on energy efficient heuristic using Minimum Correlation Coefficient (MCC) method for the placement of virtual machine. The proposed technique reduces the consumption of energy and SLA violations by using fuzzy analytic hierarchy technique. The VM placement problem is also solved by using this technique. for the validation of the proposed technique, the author use the cloudsim simulator.

- **Meta heuristic based approach**

  To lower the costs of energy and minimize the utilization of resource in the virtualized data center, VM consolidation technique has been used very widely. Recently the metaheuristic technique ACO have been successfully used to address the one-dimensional bin packing and VM consolidation problem[25]. Md Hasanul Ferdaus et al. [26] proposed the ACO based consolidation of server scheme that proposed a balanced utilization of resource on servers across different computing resources their goal is to reduce the consumption of power and wastage of resources[27].

  The work load consolidation problem is an instance of MBDB problem and design of this algorithm is based on the ant colony optimization technique[19]. One comparison is made between the proposed technique and greedy algorithm. For saving of energy the proposed algorithm is used,

  To minimize the consumption of energy in the data center and also maintaining the quality of service requirement live migration technique is used. Using the VM consolidation technique some virtual machines of the overloaded host can be migrated to another host. The underloaded host can either be switched off or put it in a low power mode. A distributed architecture system has been proposed by author in[28]. They perform VM consolidation technique dynamically to reduce the consumption of energy in the cloud data center. They proposed a meta-heuristic based ACS(Ant Colony System) which is a online optimization technique. Their proposed technique minimize the energy consumption also maintaining the required performance level in the data center.

- **Genetic Algorithm based Approach**

  One host machine can enable many number of virtual machines in server consolidation mechanism and used to maximize the efficiency of the data center. To optimize the work of data center many technique has been introduced. For cost reduction virtualization technology is used to achieve energy saving. Work in [29], proposed a two-level control system that manages the mapping of workloads to VM and from the virtual machine to resources. A fuzzy multi objective optimization problem is proposed by the author for VM placement and also minimize the thermal dissipation, consumption of power and total wastage of resources.

  The authors in[30], developed a GA based approach for VM placement and load balancing which uses the previous history of PM to effectively utilize the resource. The proposed algorithm controls the continuous stream of VM requests and dynamic workload of VMs and this approach is trying to minimize the energy consumption by maintaining the load among the physical machines. They also give a mathematical model for the above said problem with the system architecture. But the problem with this approach is that they did not focus on the performance trade-off as well as the quality of service required.

- **Greedy based Approach**

  Data center consumed a huge amount of energy. Recently many works have been focus to reduce the consumption of energy by doing efficient jod scheduling, consolidation of virtual machines and optimized allocation of resources[4]. A greedy scheduling of task has been introduced in[31]. This is the first proposed technique that efficiently minimize the energy consumption on the cloud data center. The result produced by the author 70 times lower then the Random task scheduling scheme. By reducing the number of active server the energy is minimized in this proposed technique.

  The data center size is increasing day by day, as the number of resources requested by the users is increased. The consumption of energy in a data center is directly proportional to the size of a data center. The required amount of energy can be reduced by distributing the VM requests over the data center along with the costs of distribution of VM is to be considered. These two parameters have been taken in[12], for solving the problem of VM distribution across the data center. The authors have classified the servers and workloads as IO bound and CPU bound. they have used the problem model of bin packing problem and proposed an optimized VMD problem. They present the solution of VMD by using a greedy algorithm and give an analysis by comparing the simple greedy algorithm (BF, NF, WF, RA).

- **Other Approaches**

  Table 2.1 gives the summary of the review works. The research work of the paper[32], is based on load balancing algorithms and the simulation is carried out by using cloudsim simulator[33]. In this paper, they provide a comparative study on different load balancing technique such as token ring, round robin, randomized, central queuing and least connection. To observe what is practically going on they execute three algorithms from the above using cloudsim. The research work of the proposed technique considers data center, virtual machines(VM), host and cloudlet components from cloudSim. The service request are hndle by the data center component.The application elements of VM are connected with these requests, so data center's host components should allocate VM process sharing. Three existing algorithms are executed by the authors such as Round robin, Central queuing and Randomized. The result analysis is based on the response time with respect to VM, HOST, and MIPS.

  An architectural framework is provided by the author in[34] for efficient energy management in cloud computing. They also present open research challenges, resource allocation and provision algorithm for efficient enrgy management. The performance analysis of different VM placement techniques are also discussed here. The techniques that are discussed such as MMT technique, HPG technique and RC technique in term of consumption of power, number of migrations and SLA violarions. The author conducted a performance evaluation by using cloudsim simulator for the validation of the proposed work [33].

Resource utilization and energy saving is now the main research area of cloud computing. Many works have been devoted, to maximize the utilization of resources. In the paper[35], the author proposed a method for the adjustment of resources automatically by scaling the voltage. This can be achieved by inspecting the CPU utilization. To achieve the energy conservation, the voltage of idle nodes can be reduced and the loads of over utilized nodes can be migrated to lightly loaded nodes.

The authors in[16], present the different types of research issues , techniques, and software for dynamic consolidation of the virtual machines in the cloud data center.

Table 2.1: Summary of Review Works

REF.NO	APPROACH	SCHEME	PERFORMANCE METRICS
[17]	Heuristic Based approach	VM selection and Allocation with MDBP	Energy Consumption, Cost, SLA
[20] [18]	do	Exact allocation and migration with Bin packing	Energy consumption, Energy cost
[8]	do	Adaptive threshold with MMT policy	Number of Migrations and load balancing
[21]	do	TOPSIS, TPSA, EO	Energy consumption, SLA, Number of migrations
[22]	do	$max - max$ utilization with economic based optimization model	Energy consumption
[23]	do	VM consolidation	Energy consumption, SLA
[24]	do	MCC policy for VM placement	Energy consumption, SLA
[26] [27]	Meta Heuristic Approach	ACO based Consolidation	Power consumption, Resources
[19]	do	Work consolidation with MDBP problem	Energy consumption
[28]	do	Dynamic VM consolidation, ACS	Energy consumption, QoS
[30]	Genetic Algorithm based	VM placement and Load balancing	Distribution of Load, Energy consumption
[29]	do	Two level control system VM placement with Fuzzy multi objective	Resource usage, Power consumption
[31]	Greedy based approach	Greedy task scheduling model	Energy consumption
[12]	do	VM Distribution	Energy consumption, Cost
[34]	Other	Energy efficient allocation policies and Scheduling algorithm	Energy consumption, SLA
[35]	do	DVFS with VM migration)	Energy consumption
[16]	do	DVFS with Live migration	Computing resources, Energy consumption

The proposed technique uses the concept of live migration, single threshold, double threshold and dynamic scheduling in order to reduce the energy consumption, the number of migrations and the percentage of SLA violations in this thesis.

2.4 Summary

In this chapter, we discussed different approaches to VM migration techniques. The classification of energy management technique is also discussed. We discussed the VM scheduling techniques for virtual machine placement and also some other approaches for energy management, load balancing, server consolidation and live migration.

Chapter 3

AUTREC

Adaptive Utilization Threshold for Reducing Energy Consumption in Cloud Data center

3.1 Introduction

Cloud computing is an important area of research which is based on Internet and Web Technology. The recent and most challenging topic of cloud computing is the conservation of energy in cloud data center. It is a complicated task to reduce the consumption of energy in a data center. However, in this chapter the following techniques are proposed to solve the consumption of energy problem in cloud computing. They are virtualization and VM consolidation[36]. The consumption of energy is achieved by switching the idle nodes into sleep mode so that the idle consumption of power is eliminated. According to the current resource needs the virtual machines are consolidated to the minimum number of physical nodes by using the live migration technique[37]. To reduce the consumption of energy and efficient usage of the resources VM migration technique is used. However, each migration causes SLA violations. Therefore, it is also a challenging task to minimize migrations number.

In this chapter, we proposed an adaptive threshold based approach which reduces the consumption of energy of a data center and also minimizes the number of migrations due to which there is a less percentage of SLA occurs. Our proposed method consists of the steps which are defined in the VM consolidation technique.

3.1.1 Adaptive Utilization threshold

Based on the historical data that are collected during the lifetime of the VMs, the adaptive utilization threshold technique provide auto adjustment of the threshold value. The main key concept is to adjust the upper threshold value depending on the CPU utilization. Figure 3.1 depicts the classification of the proposed technique.

In our proposed work, we used an adaptive threshold based IQR method which finds

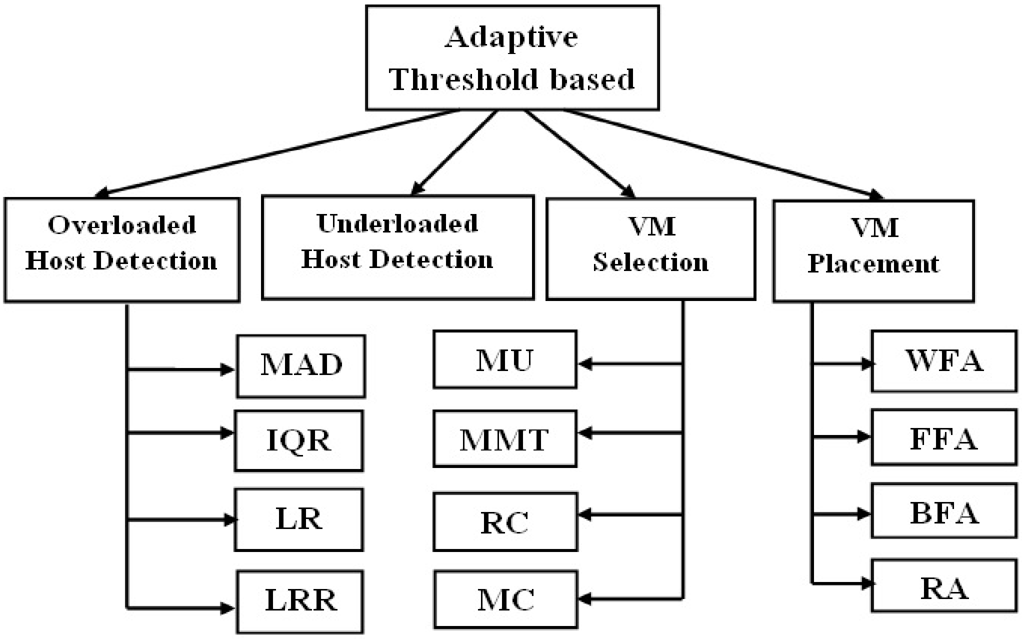

Figure 3.1: Classification of proposed technique

whether a host is overloaded or not. If an overloaded host is detected, then some VM should be migrated from it by using MMT policy.

Section 3.1 introduces the energy consumption problem with proposed AUTREC technique. The classification of Adaptive utilization threshold is also included here. The system model is given in section 3.2. Section 3.3 specifies the proposed technique. The simulation environment and results analysis are defined in section 3.4 and 3.5 respectively. The summary of the chapter is given in section 3.6.

3.2 System Model

The main component of system model are local and global manager. The physical machine contains the local manager which is associated with the VMM(virtual machine monitor). The work of local manager is to deciding when and which virtual machines have to be transferred, monitoring the CPU utilization and the virtual machine resizing according to the resource requirement. The local manager gives the details such as the updation in utilization of the resources, virtual machines migration and the state of the node inside the master node. The nodes which are not active are kept switched off. The system model is depicted in figure 3.2.

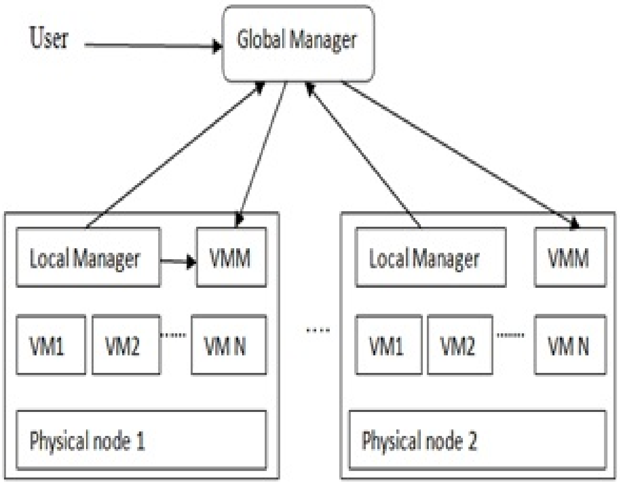

Figure 3.2: System model[12]

Table 3.1: Power consumption by at different load levels in Watt[23]

Server	**0%**	**10%**	**20%**	**30%**	**40%**	**50%**	**60%**	**70%**	**80%**	**90%**	**100%**
HP ProLiant G5	93.7	97	101	105	110	116	121	125	129	133	137
HP ProLiant G4	86	89.4	92.6	96	99.5	102	106	108	112	114	117

3.2.1 Power and Energy Model

Table 3.1 gives the consumption of power by selected server at different load levels in watt. In the data center, the power consumption can be done by computing nodes such as CPU, disk storage, network interfaces. To minimize the consumption of energy our technique is used to improve the utilization of CPU of the physical nodes in a data center. The utilization of power can be calculated by[9];

$$En = \int P_t(U(t)dt) \tag{3.1}$$

Where En is the energy, P is the power and U(t) is the CPU utilization of a physical machine. The consumption of power can be defined by a linear relationship between the power consumption and utilization of CPU.

$$P(u) = k * P_{max} + (1 - k) * P_{max} * u \tag{3.2}$$

Where P_max is the maximum consumed power when processor is fully utilized. K is the fraction of power consumed by the idle server and u is the CPU utilization.

The energy consumption can be calculated by considering both the static power and the dynamic power. Different notations parameters that are used in the system model are depicted in figure 3.2.

$$E = P_{dynamic} + P_{static} \tag{3.3}$$

Where $P_{dynamic}$ is the dynamic power, P_{static} is the static power and E is the energy. The total energy consumption of the CPU is obtained as the sum of these two components. For calculation of dynamic power DVFS technology is used and it depends upon an operational frequency f. DVFS enabled processor execute a task by using a discrete set of voltage and frequency pairs (f_i, v_i), in which $v_1 < v_2 < v_n$ and $f_1 < f_2 < f_n$.

$$P_{static} = v * I_{leak} \tag{3.4}$$

Where v is the voltage, I_{leak} is the leakage current.

$$P_{dynamic} = c * f * v^2 \tag{3.5}$$

where

$$f\alpha\frac{(v - v_t^2)}{v_t} \tag{3.6}$$

Where c is capacitive load, f is frequency, v_t is the threshold voltage.

3.2.2 Application Model

1. The service layers are specified in chapter 1. For our proposed work, we take Iaas as a service layer
2. Different models are also provided by the cloud providers and they can be classified as private cloud model, public cloud model, hybrid cloud model. In the thesis, we chose Public Cloud Model.

3.3 Proposed Technique

Our Proposed technique consists of mainly two section.

- Description part of the proposed technique
- Analysis Part of the proposed technique

The first section describe about the proposed technique and the second part gives analysis of that technique.

Table 3.2: Notation parameters

Symbols	**Meaning**
E_n	Energy
P(u)	Utilization of power
k	constant
P_{max}	Maximum power
u	of power
$P_{dynamic}$	Dynamic power
P_{static}	Static power
P(u)	Utilization of power
c	capacitive load
f	frequency
v	Voltage
V_t	Threshold voltage
H_{utol}	Utilization of host machine
MIPS	Million Instructions per Second
s	Safety parameter

3.3.1 Description of the proposed technique

This part consists of four step.

- The algorithm first searches for an overloaded host which migrate one or more VM to another host.
- Then search for the underloaded host which requires all VM to be migrated from this host and switch the host to sleep mode.
- Selection of virtual machine for migration using VM selection method.
- The new selected VM have to be placed by using VM placement method.

1. **Overloaded host detection**

   According to the CPU usage, we determine a host is overload or not. If an overloaded host is detected, then migrate some VM from it. To define the time to migrate a VM from a host is basically based on the utilization threshold. The Key concept is to set an upper utilization threshold value for host and keeping all the total utilization of the CPU by all VMs between the threshold value[38]. Assume that there are n number of hosts present in a data center, such as $1, 2, ..., n$ and on each host machine, there are p number of virtual machines, such as$1, 2, ..., p$. The cloudlet requests are executed on the host machine. To carried out this process, we take IQR method. In the inter quartile range method, the utilization of the host machines are divided into quartiles according to the upper and lower utilization. Thus the utilization of a host can be calculated by the ratio of the summation of all the cloudlet on a host to total MIPS on that host. It can be represented as,

$$H_{util} = \frac{\sum cloudlet_{req}}{total_{MIPS}} \tag{3.7}$$

H_{util} is the utilization of the host machine, $cloudlet_{req}$ is the requested cloudlet, $total_{MIPS}$ is the total required MIPS. If the H_{util} is greater the upper threshold value, then host is considered as overloaded.

Table 3.3: Approaches for Host overload Detection[38]

Slno	**Algorithm**	**Approach for host overload detection**	**Safety parameter**
1	LRR(Robust Local Regression)	Non-threshold	1.2
2	LR(Local Regression)	Non-threshold	1.2
3	IQR(Inter Quartile Range)	Adaptive utilization threshold	1.5
4	MAD(Median Absolute Deviation)	Adaptive utilization threshold	2.5

- **The Algorithm for IQR Method:**

  The IQR is an adaptive utilization threshold based policy. The method equals to the difference between upper and lower quartiles[39].

$$IQR = Q3 - Q1 \tag{3.8}$$

where Q3 is the upper utilization threshold and Q1 is lower utilization threshold.

Example: Suppose the utilization of host machines are 55%, 67%,78% 85%, 90%, 93% 96%. The median value can be calculated and marked it as Q2. In this example Q2=85%

Now, the single sequence is divided in to parts. lower and upper half. From the lower part calculate Q1 and from upper calculate Q3.

Q1(lower threshold)=67% and Q3(upper threshold)=93%

The upper threshold value for a host is calculated as;

$$T_{upper} = 1 - s.IQR \tag{3.9}$$

where s is safety parameter. T_{upper} is the upper threshold.

2. **Selection of virtual machine for migration**

Algorithm 1 Overloaded Host Detection(IQR Method)

Require: Input : hostlist output: overloaded host

```
host_list.sortDecreasingOrderOfUtilization()
for i=1 to n do
    Calculate the Median;
    Q2 ← Median
end for
for i=1 to Q2 do
    Calculate the Median;
    Q1(lowerthreshold) ← Median
end for
for i=Q2 to n do
    Calculate the Median;
    Q3(upperthreshold) ← Median
end for
```

The allocation of virtual machines is carried out by using a two-step process. (i) Selection of VMs that need to be migrated. (ii) The virtual machines which are chosen in the step one are placed on hosts using VM placement method. To determine when and which VM should be migrated, we use a threshold based VM selection policy known as MMT (Minimum migration time) policy. The main goal is to prevent SLA violations from consolidation. When a host is detected as overload, the next step is to migrate some VMs from that host to some other host. In our proposed technique, we chose the MMT policy of VM selection for migration of virtual machine.

- **MMT Policy**

  When a physical machine is found as overloaded the very next step is to select some virtual machines for migrations. The minimum migration time method is used for migration. The virtual machine which has the minimum time of migration is selected. This method is applied iteratively until the physical machine is not overloaded. The migration time can be determined as follows[34]:

$$v \in V_j \mid \forall a \in V_j, \frac{RAM_u(VM)}{NET_j} \leq \frac{RAM_u(cVM)}{NET_j} \tag{3.10}$$

  Where $RAM_u(VM)$ is the amount of RAM utilized by the VM, $RAM_u(cVM)$ is the amount of RAM currently utilized by the VM and NET_j is the spare network bandwidth available for the host j.

3. **Placement of the selected virtual machine**

In VM placement, We apply modified best fit decreasing algorithm to solve this problem. According to the CPU utilization, all the virtual machines are sorted in decreasing order[34].

Algorithm 2 BFD Algorithm

```
Input : pm_list, vm_list output: migration_list

vm_list. sortDecreasingutilization()
for vm_list=1 to m do
    min_power ← MAX
    allocated_pm ← null
    for pm_list=1 to n do
        if pm has enough resources for vm then
            power ← estimate(pm,vm)
            if power < min_powre then
                allocated_pm ← pm
                min_power ← power
            end if
        end if
    end for
    if allocated_pm! = null then
        allocation . add(vm . allocated pm)
    end if
end for
```

4. **Detection of underloaded host**

   The utilization of the host machine is compared with the utilization of lower threshold. if the utilization of the host is less then it is considered that the host is underloaded. In this case all the virtual machines are migrated from the underloaded host and this host is in switch off mode.

3.3.2 Proposed Algorithm

Algorithm 3 AUTREC Algorithm

```
Input : hostlist, vmlist output: migration list

for host=1 to n do
    Finding overloaded host(using IQR)
    if host is overloaded then
        select vm using MMT algorithm until the host is overloaded.
    end if
    Placing of VM from overloaded host using BFD algorithm.
end for
```

The flow chart of the proposed AUTREC technique is depicted in figure 3.3.

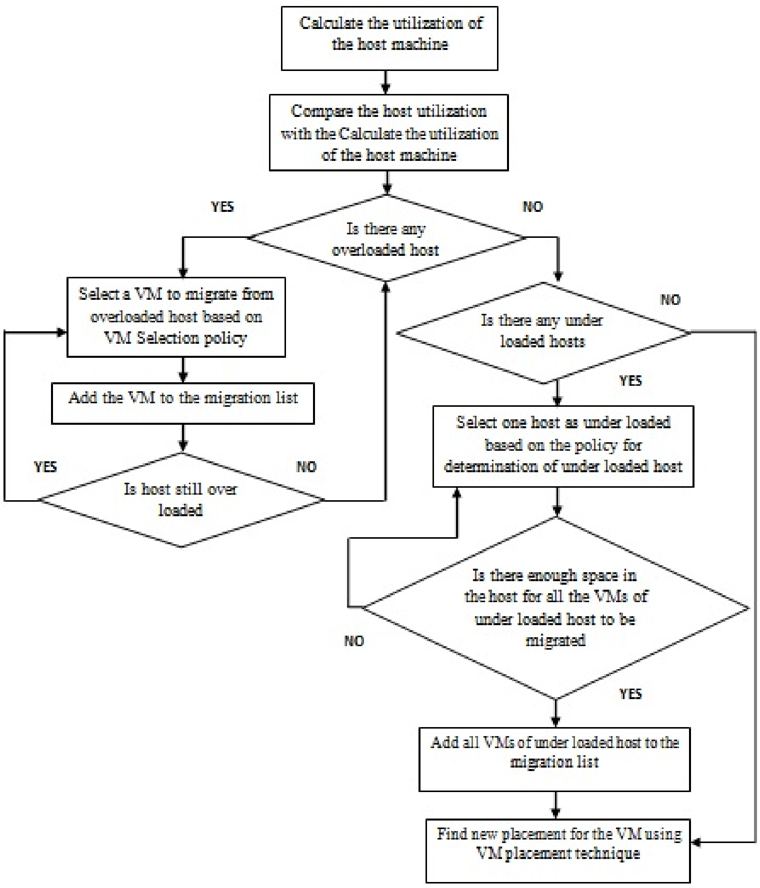

Figure 3.3: Flow chart of the proposed AUTREC technique

3.3.3 Analysis of the proposed technique

In the analysis section, we analyse our proposed AUTREC technique for the operations it performed in the cloud data center.

- Claim 1:The energy management in our proposed AUTREC technique is efficient. The energy consumption in the existing techniques is more as compared to our proposed technique. In the existing NPA technique, no power aware mechanism is used. Similarly, in DVFS and Single threshold also energy consumption is more. In DVFS, no migration method is used. But in our technique, we implement VM consolidation technique which consists of both VM selection and VM placement. Thus AUTREC technique is more energy efficient.

- Claim 2: In the Single threshold method, only one utilization threshold value is used. So it did not consider the host which are underutilized. As this technique uses the VM migration process, the number of migration is more. The performance of the system will degrade if there is more migration number and it will effect on the service level agreement also. There is a direct proportional exist between the migration number and the SLA. If SLA will be more, then the cloud provider give a penalty to its clients. Thus the migration number in AUTREC technique is less as compared to the EEDVM.

- Claim 3: The quality of service is measured in the form of SLA. The SLA is calculated according to the MIPS.If we compare the SLA value of our proposed AUTREC technique with the EEDVM algorithm then it gives better results.

3.4 Simulation Environment

This section consists of the specification of the host machine and Virtual machine.

Table 3.4: Specification for the Host machine

Host_MIPS	1860 2660
Host_RAM	4096 4096
Host_BW	1000000
Host_SIZE	1000000
Host_TYPES	2 HP PoLiant ML110 G4, HP ProLiant ML110 G5

Table 3.5: Specification for the Virtual machine

VM_MIPS	250 500 750 1000
VM_RAM	870 1740 1740 613
VM_BW	10000
VM_SIZE	2500
VM_TYPES	4

3.5 Simulation and Results

In this section, we evaluate our proposed AUTREC technique using the cloud simulator cloudsim 3.0.3 [33]. Cloudsim support both the system and behaviour modelling of cloud components such as data centers, virtual machines and provisioning policies for resources. The toolkit is extended to enable energy aware simulation as many simulators do not provide it.

We have simulated our AUTREC technique by taking a single data center. There are 50 numbers of heterogeneous physical machines are defined in the data center. Two types of host are used in our approach, HP PoLiant ML110 G4, and HP ProLiant ML110 G5. The performance of the host machine depends on the availability of MIPS and it is set to 1860 and 2660, modeled with one CPU core, 4096 MB of RAM for both types of machines and 1 TB of storage. Four types of virtual machines are used for the simulation. The MIPS set for the VMs are 250, 500, 750, 1000. 128 MB of memory, 1GB of storage is defined for the VM. The specification for the Host and virtual machine are defined in the table 3.4 1nd 3.5 respectively.

In our experiment, we considered three existing technique. They are NPA (Non-power aware), DVFS and EEDVM. The first two algorithms are defined in the cloudsim toolkit. We implement our concept of adaptive utilization threshold using IQR, MMT, and BFD algorithm. Initially, we have to calculate the utilization of the host machines for detecting of the overloaded host. The host utilization can be calculated by using equation 3.7. The next step is to sort these value in descending order and apply inter quartile range method on this to find the upper utilization threshold and lower utilization threshold value. The host utilization greater than the upper threshold is considered as overloaded. Some virtual machines are migrated from this overloaded host loaded by using the minimum time migration method. We use the best fit decreasing algorithm for the placement of the new VM. The host utilization below the threshold value is considered as under loaded hosts. The VMs of under loaded host are migrated to some other host and the under loaded hosts are either switched off or goes into an inactive mode.

We compare the result with the existing techniques. We came with some results as shown in table 3.6.

Table 3.6: Comparision of proposed technique with existing technique

Existing Approach	**Energy Consumption(kWh)**	**Number of Migration**	**SLA Violation**
NPA	150.68	-	-
DVFS	101.33	-	-
EEDVM	11.3	682	.85
AUTREC	10.4	652	0.35

1. **Observation-I Energy Consumption:**

   The observation is based on the comparison of the consumption of energy (kWh)by the data center and the number of virtual machines present in the data center. Energy consumption can be calculated by using equations 1.2, 1.3, 1.4. When the number of VMs are increased, the utilization of the host machine decreases. But if the number VMs increases, then the energy consumption due to the VMs in data center also increases. There is 8% decrease in the consumption of energy as compared to the EEDVM technique depicted in figure 3.4.

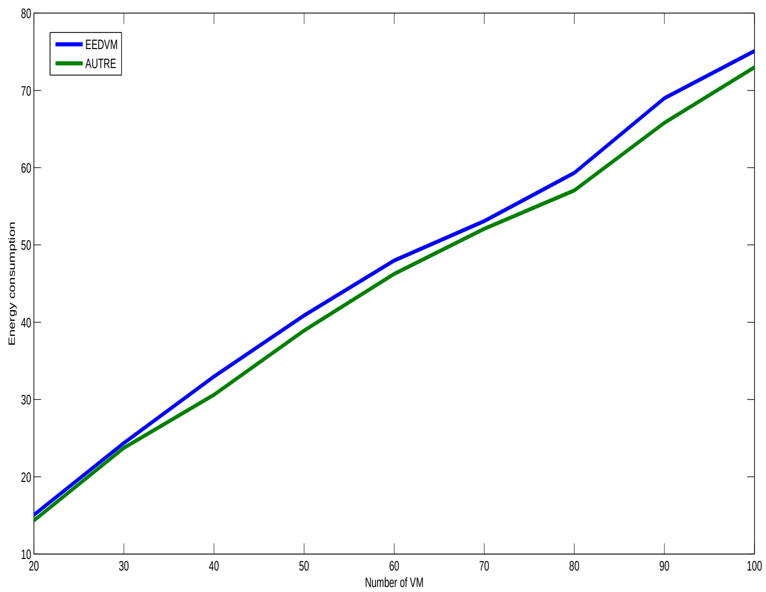

Figure 3.4: Energy consumption

2. **Observation-II Number of Migration:**

The observation is based on the comparison of the number of virtual machines and the number of migrations. It is seen that the migrations number increase if the number of virtual machines increases. There is 3% decrease in the number of migrations as compared to the EEDVM technique depicted in figure 3.4.

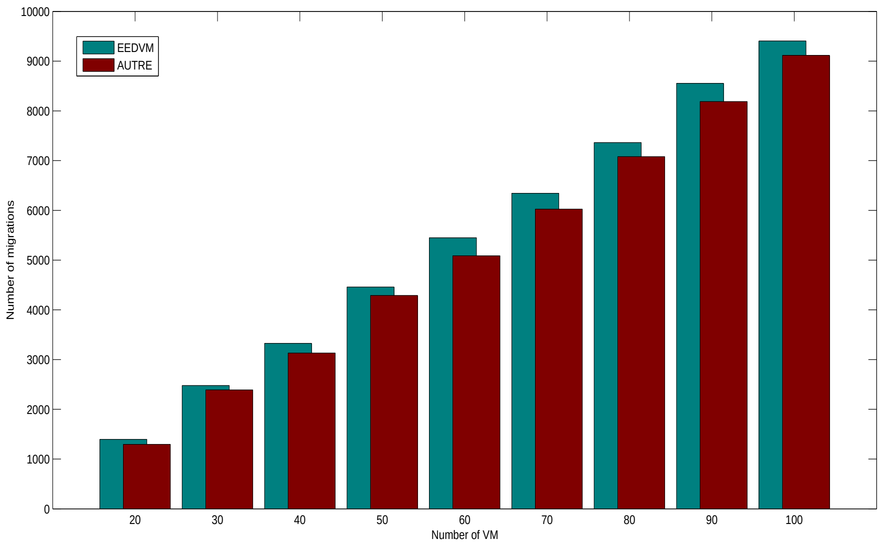

Figure 3.5: Number of migration

3. **Observation-III SLA violation:**

   SLA will occur, when CPU is not allocated even if it is demanded. The QoS is determined by using the SLA value. In equation 1.7 the SLA calculation is defined. This observation is focused on the comparison of the number of migrations and percentage of SLA violation. Our proposed AUTREC technique gives the lower SLA as compared to the EEDVM technique. There is 10% decrease in the percentage of SLA as compared to the EEDVM technique. The result is shown in figure 3.6.

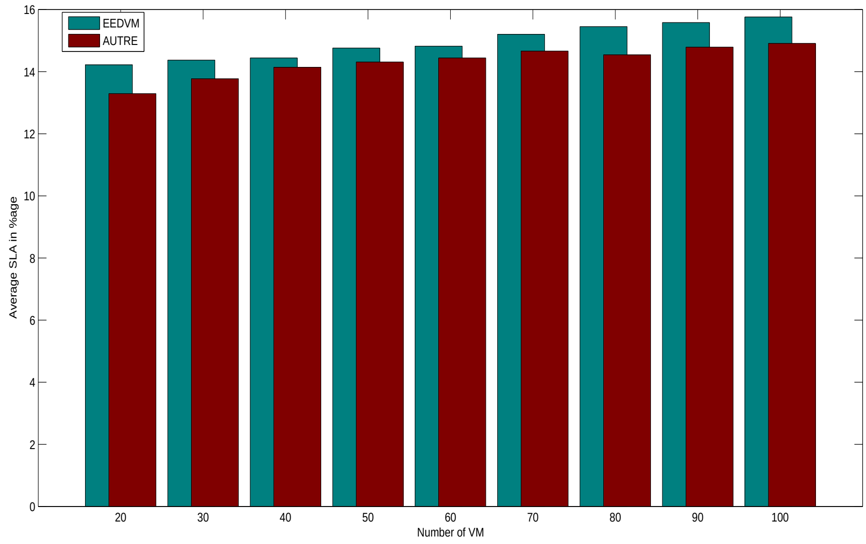

Figure 3.6: Average SLA Violation

4. **Observation-IV Number of host shut down:**

   This observation is based on the comparison of the number of VMs Vs number of host down. After migration how many numbers of physical machines are shut down is calculated here. The physical machines which are underutilized are going in a sleep mode or inactive mode. This can be done by using AUTREC technique. There is 12% decrease in the host shutdown as compared to the EEDVM technique. The result is shown in figure 3.7.

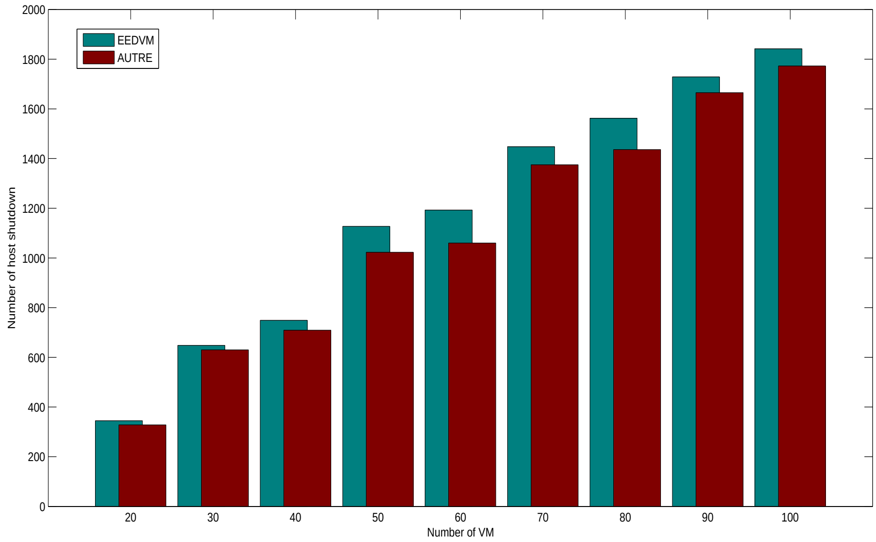

Figure 3.7: Number of Host shut down

3.6 Summary

Management of energy in a data center is a complex issue. As the number of data centers increase day by day, the energy consumption also increases and this leads to increasing of carbon emissions. One more important factor in the cloud data center is SLA (Service Level Agreement). Our technique reduces a significant amount of consumption of energy, migration's number, SLA violation under the certain quality of service requirements. We calculate the utilization of all the host machines and virtual machines. We used an IQR policy for detection of an overloaded host. We used the minimum migration time method for migrating a virtual from an overloaded host. By using the proposed technique placement of virtual machine can be done. The VM placement follows the BFD (Best Fit Decreasing) algorithm. We evaluate our proposed AUTREC technique in cloudsim simulator and compared with other techniques also by taking the random workload defined in the simulator itself. Using the proposed technique, 8% improved in energy consumption, 3% in number of migrations, 10% in SLA violation and 12% in host shutdown.

Chapter 4

DUTREC

Dynamic Utilization Threshold for Reducing Energy Consumption in Cloud Data center

4.1 Introduction

The development of large scale data centers around the world consuming a lot of energy. Increasing in consumption of energy leads to high energy costs that minimize the benefit of cloud providers. Due to the high energy consumption, high carbon emissions are produced. For a data center power has become one of the limiting factor[9]. There is a need of cloud data center for optimizing the management of the resources to minimize the energy efficiency and power, while meeting of SLA guarantees. Thus for energy efficient computing the focus must be on utilizing the servers at their full capacity and they should also not be kept completely idle. Along with the direct energy consumption of the resources, there is an additional expense of energy through the cooling systems that are required in each data center to keep all the systems cool[14].

In this chapter, we proposed a dynamic threshold based technique which is reducing the energy consumption of a data center and also minimizes the number of migration due to which a less percentage of SLA occur. Our proposed technique consists of the steps which are defined in the VM consolidation technique. In our proposed work, we used dynamic threshold based method which find a host is overloaded or not. If an overloaded host is detected, then some VM is migrated from it by using the dynamic utilization policy.

4.1.1 Allocation Policies:

The Classification of the allocation policies is depicted in figure 4.1[23].

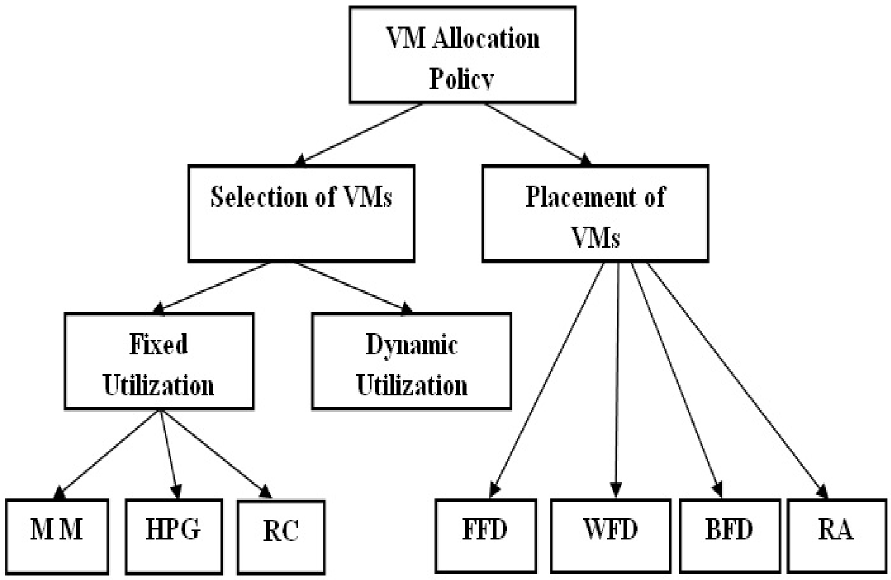

Figure 4.1: Classification of proposed technique

The total process of operation is divided into two parts. Virtual machine selection and placement. The selection process is carried out for migration of the selected VMs. The placement process is used to placed the new VMs in a host machine.

- **Fixed Utilization**

  Fixed utilization means the static utilization. This scheme is based on the concept of setting an upper utilization of threshold value[38]. The single threshold technique is based on this concept. But the rest three techniques such as MM, RC, HPG based on the concept of upper and lower utilization threshold.

  - **MM (Minimum migrations):** The number of migrations in the VM migration process should be less. So that the migration overhead can be decrease.
  - **HPG (Highest potential growth):** In the VM migration process, migrate the VMs which has the low CPU utilization. So that the total potential increase of the utilization can be minimized and also reduction in SLA violation.
  - **RC (Random Choice):** The required amount of VMs can be selected randomly.

- **Dynamic Utilization**

  The static threshold values are not suitable for all the cases like for dynamic workloads and for the workload which are unpredictable. Because in a dynamic environment many types of application can share a single physical resource. In the dynamic case the system automatically adjusts its behaviour depending upon the workload[38].

Section 4.1 introduces different allocation policies. The proposed DUTREC technique is define in section 4.2. The simulation environment and results analysis are specifies in section 4.3 and 4.4 respectively. Section 4.5 gives the summary of this chapter.

4.2 Proposed Technique

Our Proposed technique consists of mainly two section. VM selection and VM placement. Selection of virtual machines for migration, we used dynamic utilization threshold method. For VMs placement Modified First Fit Decreasing(MFFD) technique is used.

This part consists of four step.

- The algorithm first search for an overloaded host which migrate one or more VM to another host.
- Then search for underloaded host which requires all VM to be migrated from this host and switch the host to sleep mode
- Selection of virtual machine for migration using VM selection method
- The new selected VM have to be placed by using VM placement method.

1. **Overloaded host detection**

   According to the CPU usage, we determine a host is overload or not. If an overloaded host is detected, then migrate some VM from it. The aim is to setting an upper and lower utilization threshold value for host[38].

2. **Selection of virtual machine for migration**

   The selection of virtual machines is carried out by using a two step process. The first one is the selection of VMs that need to be migrated. In the second step, the virtual machines which are chosen in the first step are placed on hosts using VM placement method. A dynamic utilization technique is used for the selection of virtual machines and it determine when and which VM has to be migrate.

   - **MMT Policy**

     When a physical machine is found as overloaded the very next step is to select some virtual machines for migrations. The minimum migration time method is used for migration. The virtual machine which has the minimum time of migration is selected. This method is applied iteratively until the physical machine is not overloaded[34].

Algorithm 4 Selection of virtual machine

Require: Input : hostlist output: overloaded host

```
vm_list. sortDecreasingOrderOfUtilization()
for pmlist=1 to n do
    P_util ← p.util()
    BF_util ← MAX
    while p_util > p.upthrs() do
        for vmlist=1 to m do
            if vm.util > (p_util − upthrs()) then
                t ← vm.util() − p_util + p.upthrs()
                if t < MFF_util then
                    MFF_util ← t
                    MFF_vm ← vm
                else
                    if MFF_util = MAX then
                        MMF_vm ← vm
                        break
                    end if
                end if
            end if
        end for
        p_util ← p_util − MMF_vm.util()
        miglist.add(MMF_vm))
        vmlist.remove(vm)
    end while
    if p_util < lowthrs() then
        miglist.add(p.getp_util())
        vmlist.remove(p.getvmlist())
    end if
    return miglist
end for
```

3. **Placement of the selected virtual machine**

   The virtual machine placement problem is considered as a bin packing problem. We apply modified first fit decreasing algorithm to solve this problem. All the virtual machines are sorted in decreasing order according to the CPU utilization.

Algorithm 5 MFFD

Require: Input : $host\_list, vm\_list$ output: $allocation\_virtual\_machines$

$vm\_list.sort\_DecreasingOrderOf\_Utilization()$
for $vmlist = 1ton$ **do**
  $host\_list.sort\_availablepower\_decreasing()$
  $allocated\_host \leftarrow$ null
  for $host\_list$ =1 to m **do**
    if $host == enough\_resources$ **then**
      $powerafter\_allocation \leftarrow getpower\_afterallocation(host, vm)$
      if $powerafter\_allocation < host.getmaxpower()$ **then**
        $allocated\_host \leftarrow$ host
      end if
    end if
  end for
  if $allocated\_host! =$ null **then**
    $allocation.add(vm.allocated\_host)$
  end if
end for

4. **Detection of underloaded host**

   The underloaded physical machines are detected by comparing the utilization with the lower utilization threshold. If utilization of the host machine is less, of lower threshold then the host is considered as underloaded. In this case all the virtual machines are migrated from the underloaded host and this host is in switch off mode.

4.3 Simulation Environment

This section consists of the specification of the host machine and Virtual machine.

Table 4.1: Specification for the Host machine

Host_MIPS	1960 2700
Host_RAM	4096 4096
Host_BW	1 Gbits per Sec
Host_SIZE	1 TB
Host_TYPES	2

Table 4.2: Specification for the Virtual machine

VM_MIPS	500 575 750 1000
VM_RAM	870 1740 1740 613
VM_BW	1 Mbits per sec
VM_SIZE	2.5 GB
VM_TYPES	4

4.4 Simulation Result and Analysis

In this section, we evaluate our proposed DUTREC model using the cloud simulator cloudsim 3.0.3 [33]. We simulate our DUTREC technique by taking a single data center. There are 50 numbers of heterogeneous physical machines defined in the data center. Two types of host are used in our approach, HP PoLiant ML110 G4 and HP ProLiant ML110 G5. The performance of the host machine is depend upon the availability of MIPS and it is set to 1860 and 2660, modeled with one CPU core, 4096 MB of RAM for both types of machines and 1 TB of storage. Four types of virtual machines are used for the simulation. The MIPS set for the VMs are 250, 500, 750, 1000. 128 MB of memory, 1GB of storage is defined for the VM. The specification for the Host and virtual machine are defined in table 4.1 and 4.2.

In our experiment we consider four existing algorithms. They are NPA (Non power aware), DVFS, PALVM and AUTREC. The first two algorithms are defined in the cloudsim toolkit. We implement our concept of dynamic utilization threshold using dynamic utilization threshold, MMT and BFD algorithm. Initially we calculate the utilization of the host machine for detecting of overloaded host. The host utilization is calculated by using equations 3.7. The next step is to sorting these value in descending order and apply dynamic threshold method to find the upper utilization threshold and lower utilization threshold value. The host utilization greater than the upper threshold is considered as overloaded host. Some virtual machines are migrated from this overloaded host by using the minimum time migration method. We used modified first fit decreasing algorithm for the placement of the new VM. The host utilization below the threshold value can be considered as under loaded hosts. The VMs of under loaded host are migrated to some other host and the under loaded hosts are either switched off or goes into an inactive mode.

We compare the result with PALVM and some other existing technique as well as with the adaptive utilization threshold technique. We came with some results as shown in table 4.3.

Table 4.3: Comparision of proposed approach with existing approach

Existing Approach	**Energy Consumption(kWh)**	**Number of Migrations**	**SLA Violation**
NPA	150.68	-	-
DVFS	613.6	-	
PALVM[14]	10	782	2.68
DUTTREC	9.2	737	2.24

1. **Observation-I Energy Consumption:**

   The following observation is based on the comparison of the consumption of energy (kWh)by the data center and the number of virtual machines present in the data center. Energy consumption can be calculated by using equations 1.2, 1.3, 1.4. There is 9% decrease in the consumption of energy as compared to the PALVM algorithm and it is depicted in figure 4.2.

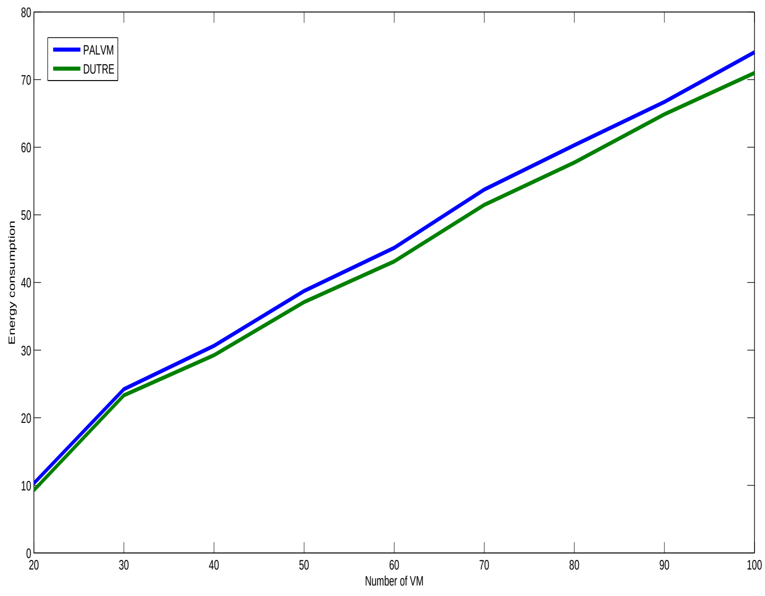

Figure 4.2: Energy consumption

2. **Observation-II Number of Migration:**

   The observation is based on the comparison of the number of virtual machines and number of migrations. The VM migration is calculated by using equations 1.5 and 1.6. It is seen that the migrations number increase, if number of virtual machines increases. There is 6% decrease in the number of migrations as compared to the PALVM algorithm and it is depicted in figure 4.3.

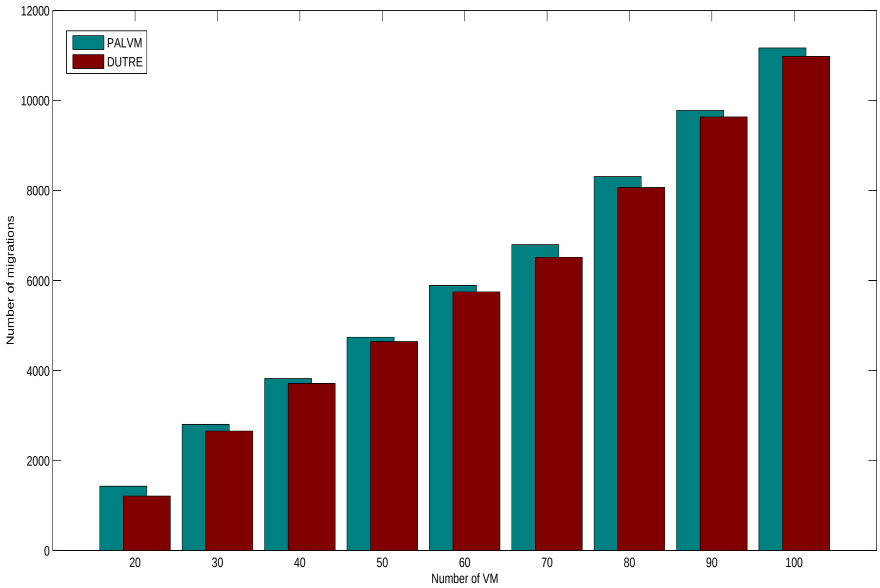

Figure 4.3: Number of migration

3. **Observation-III SLA violation:**

   This observation is focus on the comparison of number of migration and percentage of SLA violation. There is 18% decrease in the percentage of SLA violations as compared to the PALVM algorithm and it is depicted in figure 4.4.

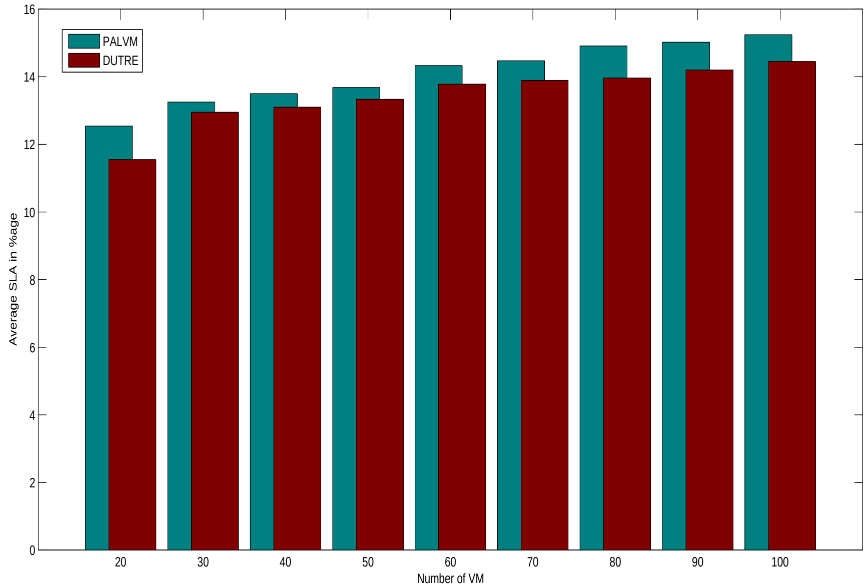

Figure 4.4: Average SLA Violation

4. **Observation-IV Number of host shut down:**

   The following observation is based on the comparison of number of migration Vs number of host down. There is 13% decrease in the number of host shutdown as compared to the PALVM algorithm and it is depicted in figure 4.5.

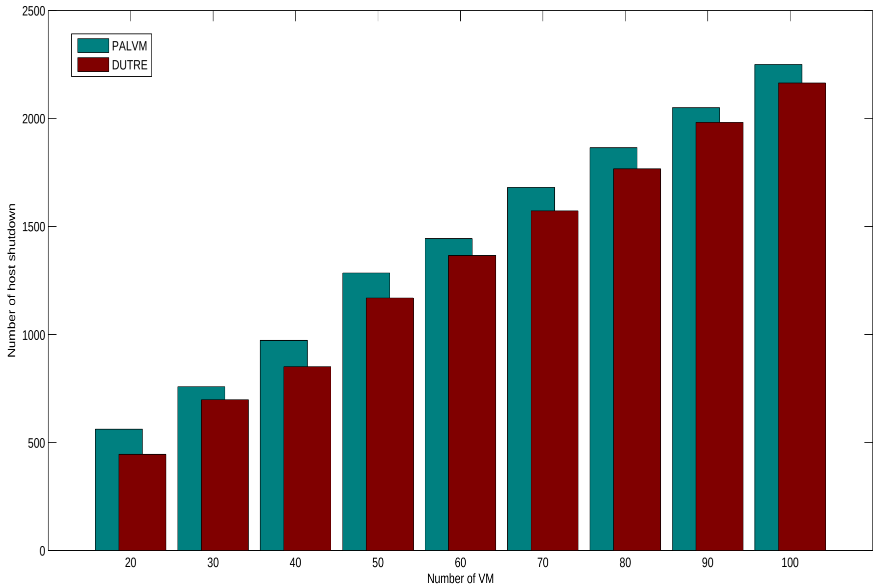

Figure 4.5: Number of Host shut down

4.5 Conclusion

The consumption of energy is a big issue now in the world of cloud computing. Many researchers doing their research to reduce the consumption of energy in the data center. Along with the energy consumption one more important factor is there in the cloud data center that ie SLA. The proposed technique minimize the energy consumption, SLA violation under certain QoS requirements and number of migrations. The utilization of the host and virtual machines are calculated by using the equations defined in the system model. We used a dynamic threshold policy for detection of overloaded host. We used the minimum migration time method for migrating a virtual from an overloaded host. The VM placement follows the MFFD(Modified First Fit Decreasing) algorithm. We evaluate our DUTREC approach in cloudsim simulator and compared with other techniques also by taking the random workload defined in the simulator itself. This technique give better results than the approach AUTREC proposed in chapter 3. Using this technique, 9% improved in energy consumption, 6% in number of migrations, 20% in SLA violation and 13% in host shutdown.

Chapter 5

Conclusion

Management of consumption of energy in a data center is a complex issue. As the number of data centers increase day by day, the energy consumption is also increases and this leads to increasing of carbon emissions. One more important factor in the cloud data center is SLA (Service Level Agreement). To overcome the above issues virtualization concept is introduced by providing consolidation and dynamism. For reducing energy consumption, AUTREC method is proposed. This method is energy efficient. AUTREC reduces a significant amount of consumption of energy, migration's number and SLA violation under certain quality of service requirements.In this method, we first calculate the utilization of all the host machines. An IQR based policy is proposed to find out a host is overloaded or not. The detection of overloaded host is based on the CPU utilization. Some virtual machines have selected from the overloaded host for migration. In AUTREC, we used MMT policy for the migration of virtual machines. For placement of these new VMs, Best Fit Decreasing algorithm is used. The proposed technique is a energy based technique, which has been developed and simulated using the cloud simulator cloudsim 3.0.3. Using this technique, 8% improved in energy consumption, 3% in number of migrations, 10% in SLA violation and 12% in host shutdown.

In this work one more approach is proposed, named as DUTREC. This method is also proposed for reducing the consumption of energy of a cloud data center. DUTREC also follows the steps of VM consolidation. This method is based on the concept of dynamic utilization threshold. For placement of VM, we are using the MFFD algorithm. Using this technique 9% improved in energy consumption, 6% in number of migrations, 20% in SLA violation and 13% in host shutdown.

Scope for Future Research

The algorithm or technique considering the number of parameters responsible for QoS improvement, need to be considered. We will consider for large data center having more number of physical and virtual machines. The consumption of power for cooling, storage is to be included in our future algorithm. For our research work, we consider only one data center. Hence, in future algorithms we will include more than one data center.

References

[1] C. Pettey and B. Tudor, "Gartner says energy-related costs account for approximately 12 percent of overall data center expenditures," 2010.

[2] R. Buyya, A. Beloglazov, and J. Abawajy, "Energy-efficient management of data center resources for cloud computing: a vision, architectural elements, and open challenges," *arXiv preprint arXiv:1006.0308*, 2010.

[3] P. Mell and T. Grance, "The nist definition of cloud computing," 2011.

[4] V. J. Patel and H. A. Bheda, "Reducing energy consumption with dvfs for real-time services in cloud computing."

[5] A. Fox, R. Griffith, A. Joseph, R. Katz, A. Konwinski, G. Lee, D. Patterson, A. Rabkin, and I. Stoica, "Above the clouds: A berkeley view of cloud computing," *Dept. Electrical Eng. and Comput. Sciences, University of California, Berkeley, Rep. UCB/EECS*, vol. 28, no. 13, p. 2009, 2009.

[6] T. Veni and S. Bhanu, "A survey on dynamic energy management at virtualization level in cloud data centers," *Computer Science & Information Technology*, no. 3, pp. 107–117, 2013.

[7] J. Sekhar, G. Jeba, and S. Durga, "A survey on energy efficient server consolidation through vm live migration," *International Journal of Advances in Engineering & Technology*, vol. 5, no. 1, pp. 515–525, 2012.

[8] K. Maurya and R. Sinha, "Energy conscious dynamic provisioning of virtual machines using adaptive migration thresholds in cloud data center," *Int. J. Comput. Sci. Mob. Comput*, vol. 3, no. 2, pp. 74–82, 2013.

[9] M. H. Al Shayeji and M. Samrajesh, "An energy-aware virtual machine migration algorithm," in *Advances in Computing and Communications (ICACC), 2012 International Conference on*. IEEE, 2012, pp. 242–246.

[10] S. K. Mandal and P. M. Khilar, "Efficient virtual machine placement for on-demand access to infrastructure resourcesin cloud computing," *International Journal of Computer Applications*, vol. 68, no. 12, 2013.

[11] P. G. J. Leelipushpam and J. Sharmila, "Live vm migration techniques in cloud environment—a survey," in *Information & Communication Technologies (ICT), 2013 IEEE Conference on*. IEEE, 2013, pp. 408–413.

[12] A. K. Paul, S. K. Addya, B. Sahoo, and A. K. Turuk, "Application of greedy algorithms to virtual machine distribution across data centers," in *India Conference (INDICON), 2014 Annual IEEE*. IEEE, 2014, pp. 1–6.

[13] L. Salimian and F. Safi, “Survey of energy efficient data centers in cloud computing,” in *Proceedings of the 2013 IEEE/ACM 6th International Conference on Utility and Cloud Computing*. IEEE Computer Society, 2013, pp. 369–374.

[14] R. Sinha, N. Purohit, and H. Diwanji, “Power aware live migration for data centers in cloud using dynamic threshold,” 2011.

[15] A. Corradi, M. Fanelli, and L. Foschini, “Vm consolidation: A real case based on openstack cloud,” *Future Generation Computer Systems*, vol. 32, pp. 118–127, 2014.

[16] V. Patel and H. Bheda, “An advanced survey on research issues of energy management in cloud computing,” vol. 4, no. 1. IEEE, 2014, p. 163—166.

[17] M. S. Hasan and E.-N. Huh, “Heuristic based energy-aware resource allocation by dynamic consolidation of virtual machines in cloud data center.” *TIIS*, vol. 7, no. 8, pp. 1825–1842, 2013.

[18] C. Ghribi, M. Hadji, and D. Zeghlache, “Energy efficient vm scheduling for cloud data centers: Exact allocation and migration algorithms,” in *Cluster, Cloud and Grid Computing (CCGrid), 2013 13th IEEE/ACM International Symposium on*. IEEE, 2013, pp. 671–678.

[19] E. Feller, L. Rilling, and C. Morin, “Energy-aware ant colony based workload placement in clouds,” in *Proceedings of the 2011 IEEE/ACM 12th International Conference on Grid Computing*. IEEE Computer Society, 2011, pp. 26–33.

[20] N. Karmarkar and R. M. Karp, “An efficient approximation scheme for the one-dimensional bin-packing problem,” in *Foundations of Computer Science, 1982. SFCS'08. 23rd Annual Symposium on*. IEEE, 1982, pp. 312–320.

[21] E. Arianyan, H. Taheri, and S. Sharifian, “Novel energy and sla efficient resource management heuristics for consolidation of virtual machines in cloud data centers,” *Computers & Electrical Engineering*, vol. 47, pp. 222–240, 2015.

[22] D. Kumar and B. Sahoo, “Energy efficient heuristic resource allocation for cloud computing,” 2014.

[23] A. Beloglazov and R. Buyya, “Adaptive threshold-based approach for energy-efficient consolidation of virtual machines in cloud data centers,” in *Proceedings of the 8th International Workshop on Middleware for Grids, Clouds and e-Science*, vol. 4. ACM, 2010.

[24] N. Kord and H. Haghighi, “An energy-efficient approach for virtual machine placement in cloud based data centers,” in *Information and Knowledge Technology (IKT), 2013 5th Conference on*. IEEE, 2013, pp. 44–49.

[25] M. Dorigo, G. Di Caro, and L. M. Gambardella, “Ant algorithms for discrete optimization,” *Artificial life*, vol. 5, no. 2, pp. 137–172, 1999.

[26] M. H. Ferdaus, M. Murshed, R. N. Calheiros, and R. Buyya, “Virtual machine consolidation in cloud data centers using aco metaheuristic,” in *Euro-Par 2014 Parallel Processing*. Springer, 2014, pp. 306–317.

[27] J. Levine and F. Ducatelle, “Ant colony optimization and local search for bin packing and cutting stock problems,” *Journal of the Operational Research Society*, pp. 705–716, 2004.

[28] F. Farahnakian, A. Ashraf, T. Pahikkala, P. Liljeberg, J. Plosila, I. Porres, and H. Tenhunen, "Using ant colony system to consolidate vms for green cloud computing," *Services Computing, IEEE Transactions on*, vol. 8, no. 2, pp. 187–198, 2015.

[29] J. Xu and J. A. Fortes, "Multi-objective virtual machine placement in virtualized data center environments," in *Green Computing and Communications (GreenCom), 2010 IEEE/ACM Int'l Conference on & Int'l Conference on Cyber, Physical and Social Computing (CPSCom)*. IEEE, 2010, pp. 179–188.

[30] M. Madhusudhan and K. C. Sekaran, "A genetic algorithm approach for virtual machine placement in cloud," 2013.

[31] Z. Dong, N. Liu, and R. Rojas-Cessa, "Greedy scheduling of tasks with time constraints for energy-efficient cloud-computing data centers," *Journal of Cloud Computing*, vol. 4, no. 1, pp. 1–14, 2015.

[32] S. Ray and A. De Sarkar, "Execution analysis of load balancing algorithms in cloud computing environment," *International Journal on Cloud Computing: Services and Architecture (IJCCSA)*, vol. 2, no. 5, pp. 1–13, 2012.

[33] R. N. Calheiros, R. Ranjan, A. Beloglazov, C. A. De Rose, and R. Buyya, "Cloudsim: a toolkit for modeling and simulation of cloud computing environments and evaluation of resource provisioning algorithms," *Software: Practice and Experience*, vol. 41, no. 1, pp. 23–50, 2011.

[34] A. Beloglazov, J. Abawajy, and R. Buyya, "Energy-aware resource allocation heuristics for efficient management of data centers for cloud computing," *Future generation computer systems*, vol. 28, no. 5, pp. 755–768, 2012.

[35] L.-T. Lee, K.-Y. Liu, and H.-Y. Huang, "Dynamic resource management for energy saving in the cloud computing environment," in *Proceedings of the International Conference on Information Science and Industrial Applications*, 2012, pp. 176–181.

[36] W. Chawarut and L. Woraphon, "Energy-aware and real-time service management in cloud computing," in *Electrical Engineering/Electronics, Computer, Telecommunications and Information Technology (ECTI-CON), 2013 10th International Conference on*. IEEE, 2013, pp. 1–5.

[37] S. S. Baghshahi, S. Jabbehdari, and S. Adabi, "Virtual machines migration based on greedy algorithm in cloud computing," *International Journal of Computer Applications*, vol. 96, no. 12, 2014.

[38] A. Abdelsamea, E. E. Hemayed, H. Eldeeb, and H. Elazhary, "Virtual machine consolidation challenges: A review," *International Journal of Innovation and Applied Studies*, vol. 8, no. 4, p. 1504, 2014.

[39] P. Shukla and R. Pateriya, "Iqr based approach for energy efficient dynamic vm consolidation for green cloud data centers," *International Journal of Computer Applications*, vol. 123, no. 9, 2015.

[40] G. Metkar, S. Agrawal, and S. Singh, "A live migration of virtual machine based on the dynamic threshold at cloud data centres," *International Journal of Advanced Research in Computer Science and Software Engineering*, vol. 3, no. 10, pp. 401–405, 2013.