Energy Efficient Task Scheduling Algorithms In Cloud Data Center

A thesis submitted in

in partial fulfillment of the requirements

for the degree of

Master of Technology

by

Devendra Singh Thakur

(Roll: 212cs3128)

Department of Computer Science and Engineering

National Institute of Technology Rourkela

Rourkela – 769 008, India

Energy Efficient Task Scheduling Algorithms In Cloud Data Center

A thesis submitted in
in partial fulfillment of the requirements
for the degree of

Master of Technology

in

Computer Science and Engineering

by

Devendra Singh Thakur

(Roll: 212cs3128)

under the supervision of

Prof. Durga Prasad Mohapatra

Department of Computer Science and Engineering

National Institute of Technology Rourkela

Rourkela – 769 008, India

Certificate

This is to certify that the work in the thesis entitled *Energy Efficient Task Scheduling Algorithms in Cloud Data Centers* by *Devendra Singh Thakur*, bearing roll number 212cs3128, is a record of an original research work carried out by him under my supervision and guidance in partial fulfillment of the requirements for the award of the degree of *Master of Technology* in *Computer Science and Engineering*. Neither this thesis nor any part of it has been submitted for any degree or academic award elsewhere.

Prof. Durga Prasad Mohapatra

Acknowledgment

First of all, I am thankful to God for his blessings and showing me the right direction. With His mercy, it has been made possible for me to reach so far.

Foremost, I would like to express my sincere gratitude to my advisor *Prof. Durga Prasad Mohapatra* for the continuous support of my M.Tech study and research, for his patience, motivation, enthusiasm, and immense knowledge. I am thankful for her continual support, encouragement, and invaluable suggestion. His guidance helped me in all the time of research and writing of this thesis. I could not have imagined having a better advisor and mentor for my M.Tech study.

Besides my advisor, I extend my thanks to our HOD, *Prof. S. K. Rath* and *Prof. B. D. Sahoo* for their valuable advices and encouragement. I express my gratitude to all the staff members of Computer Science and Engineering Department for providing me all the facilities required for the completion of my thesis work.

I would like to say thanks to all my friends especially Dilip Kumar, Alok Pandey for their support.

Last but not the least I am highly grateful to all my family members for their inspiration and ever encouraging moral support, which enables me to purse my studies.

Devendra Singh Thakur
Roll: 212CS3128
Department of Computer Science

Author's Declaration

I hereby declare that all work contained in this report is my own work unless otherwise acknowledged. Also, all of my work has not been submitted for any academic degree. All sources of quoted information has been acknowledged by means of appropriate reference.

Devendra Singh Thakur
Roll: 212CS3128
Department of Computer Science

Abstract

Cloud computing is a technology that provides a platform for the sharing of resources such as software, infrastructure , application and other information. It brings a revolution in Information Technology industry by offering on-demand of resources. Clouds are basically virtualized datacenters and applications offered as services. Data center (Server infrastructure) hosts hundreds or thousands of servers which comprised of software and hardware to respond the client request. A large amount of energy requires to perform the operation. A data center with 500*100 servers consumes around 9Megawatt to perform operation. Energy consumption is a key concern in data center. Energy consumption by Google is 2,675,898 MWh in 2011.

Cloud Computing is facing lot of challenges like Security of Data , Consumption of energy, Server Consolidation, etc. The research work focuses on the study of task scheduling management in a cloud environment. The main goal is to improve the resource utilization and redeem the consumption of energy in data centers. Energy-efficient scheduling of workloads helps to redeem the consumption of energy in data centers, thus helps in better usage of resource. This further reduce operational costs and provides benefits to the clients and also to cloud service provider.

In this report of thesis, the task scheduling in data centers have been compared. Cloudsim a toolkit for modeling and simulation of cloud computing environment has been used to implement and demonstrate the experimental results. The results aimed at analyzing the energy consumed in data centers and shows that by having reduce the consumption of energy the cloud productivity can be improved.

Keywords: Data Center, Cloud computing, Virtual Machines, Physical Machines, Workloads, Energy , Utilization of Resources.

Contents

List of Figures

Chapter 1

Introduction

This chapter comprise of cloud computing, the evolution of cloud computing , other technologies related like grid. It also discuss with characteristics of cloud, cloud computing services, followed by the research motivation and thesis organization.

1.1   Evolution of Cloud Computing

The evolution of the cloud goes phase by phase that include the Grid computing , distributed computing. Cloud computing is used first in 1950s, the time during which large-scale mainframes were available in the business industry. The hardware used by the mainframe were installed in a big room and all users are accesing the mainframe through terminals.

Later in the year 1970, the IBM launches OS having a number of virtual machines at a single machine. The Virtual machine Os has taken the application of 1950s, that is of sharing the access to a mainframe to a higher level by considering a number of virtual machines providing different accessible machines at a single physical machines.

Idea of cloud computing was first showed by J.C.R Licklider and John McCarthy in 1969. The vision behind this is that everyone goes interconnected and thus able to access data through anywhere.

The data are stored in a data-center(a centralized infrastructure) which is a vast data storage space. The processing of the request or data performed through servers thus availability and security of the data will be addressed. The service provider and the clients has an agreement for the usage known as SLA(Service level Agreement). Then in 1999 , salesforce.com put this idea to an application. Then in 2002, a Cloud based services of web launched by amazon.

It provide on demand services to the subscribed users.There are many proposed definitions of the Cloud computing due to its growing popularity defining its characteristics. Some of the definitions given by many well-known scientists and organizations are:

- Rajkumar Buyya defines the Cloud computing in terms of its utility to end user as A Cloud is a type of parallel and distributed system consisting of a collection of interconnected and virtualized computers that are dynamically provisioned and presented as one or more unified computing resources based on service-level agreements established through negotiation between the service provider and consumers [1].

- National Institute of Standards and Technology (NIST) defines Cloud computing as follows: Cloud computing is a model for enabling convenient, on-demand network access to a shared pool of configurable computing resources (e.g., networks, servers, storage, applications, and services) that can be rapidly provisioned and released with minimal management effort or service provider interaction. This Cloud model promotes availability and is composed of five essential characteristics, three service models, and four deployment models [2].

Cloud computing also defined as A style of computing where IT-enabled cababilities are delivered as a service to end users using internet.

1.1.1   Cloud Computing characteristics

The characteristics of cloud computing are:-

- **Reduction of Cost** There are a number of reasons to attribute Cloud technology with lower costs. The billing model is pay as per usage; the infrastructure is not purchased thus lowering maintenance. Initial expense and recurring expenses are much lower than traditional computing.

- **Elasticity** Services offered by cloud are rapidly provisioned, and in few cases by itself, rapidly released to quickly scale in. To consumer, the capabilities available for provisioning appears to be endless and could be purchased whenever required.

- **Security and Availability** The cloud should authorize the data access by the end users. However a fear of security always there with the end users . The requests have to be fulfilled in every case , the data and infrastructure should always available.

- **Flexibility** Cloud computing mainly stress on deployment of applications in market as quickly possible , by using the most appropriate building blocks necessary for deployment.

- **Geographical independence** A user can access the data shared on cloud from any location across the globe

1.1.2   Services of Cloud Computing

The meaning of cloud computing services is to use reusable, and fine grained components on a network provided by CSP(Cloud service provider). Cloud computing generally offers three types of services:

- **Software as a service** In Software as a service an application is provided as a service to the customers who can access it through the network. The application is hosted by cloud data centres. Since the application is hosted not on customer site so the customer doesnt have to bother about the mainteanance and support of application . But the customer cant be able to make changes

in application while the service provider can make change in it. The thing is that the customer can only use the software while all changes will be done by provider. The biggest benefits of software as a service is costing less money than to buy the software application. Ex salesforce.com: for buying softwares on demand.

- **Platform as a service** Platform as a service model provides all the resources required to build applications and services through the internet. You dont need to install or download the software. The Paas Services include application design ,development, testing, deployment, and hosting. The hurdles in Paas is that the developers are not having interoperability and portability among the providers. The cost of changing the application to different provider is very high. Example of Paas is Azure services and Amazon web services.

- **Infrastructure as a service** It simply offers the hardware so the customer can keep anything onto it. Iaas allows the customer to take resources like Server space , cpu cycles , memory space, network equipment on rent. Based on requirement the infrastructure can be enhance up or down. VMware and EC2 cloud offered by Amazon as a Iaas.

1.1.3   Cloud Computing Deployment Model

The deployment model are shown in the diagram and explained as follows:

- **Public Cloud**   A public cloud network enables users to distribute and access data from anywhere at any given point in time. This means that public cloud computing systems are incredibly accessible and can be shared with third parties. Based on the standard cloud computing model, in a public cloud the service provider makes its applications, storage or other resources, available to the general public. Examples of the public cloud include Google AppEngine.

  The main benefits of a public cloud service are: easy and inexpensive to set up, scalability, and a pay per what you use model (no wasted resources).

- **Private Cloud** Availability and distribution mediums in a private cloud network are limited only for authorized users from behind a firewall. This form of cloud computing is specifically designed for companies that do not want to distribute their internal work information to third parties. Nonetheless, these outside users can still access or distribute data provided they are authorized by the main client to access. Private cloud computing networks are much safer to use than public ones since they require all users to be authorized.

- **Hybrid Cloud** Hybrid cloud is developed with both public and private cloud characteristics. While public and private cloud systems are more prevalent, hybrid types have been growing in demand. Hybrid cloud systems occur when an organization provides some cloud services in-house and has others provided externally.

  The advantage to this approach is that companies are able to host external data off-site with an external provider , while maintaining control over internal customer data.

1.2 Issues for Research

To reach the extent of CLOUD computing, major aspects have not yet been developed and realised and in some cases not even researched. Many issues are still needs to be know.

- **Server Consolidation** Energy optimization is an important issue in cloud computing environments. To reduce this , a idea is to redeem the idle power going waste by underutilized servers. The fact is that a server even runs a very small workload, then also it consume over 50% of the peak power [6]. Thus the focus of conserving energy is to turn on as few servers as possible by consolidating the workload. This condition is referred to as server consolidation. It is a nice approach to better utilize the resources and to

redeem the consumption of power.

- **Security of Data** Security of data is a vital and important research issue in cloud computing. It hamper the expansion of the cloud. Usually, the services of cloud computing are delivered by a third party known as service provide, who owns the infrastructure. Even for a virtual private Cloud, the service provider can only specify the security setting remotely, without knowing whether it is fully implemented. It is hard to make the trust at each layer of the Cloud. Firstly, the hardware layer must be trusted using hardware trusted platform module. Secondly, the virtualization platform must be trusted using secure virtual machine monitors. VM migration should only be allowed if both source and destination servers are trusted.

- **Energy Consumption** To reduce the energy consumption in data centres is another issue in cloud computing. It has been found that that a server even runs a very small workload, then also it consume over 50% of the peak power. Energy consumption by Google 2,675,898 MWh in 2011 and it is increasing regularly. Thus there is a need to control the consumption of energy otherwise the cost of cloud computing will increase tremendously. The aim is to reduce the energy consumed in data centers by following the service level agreement. This issue is now started gaining importance. This issues can be addressed by a lot of approach. For example by selectively shutting down the unutilized server the power consumption can be greatly reduced. The utilization of resources can be enhanced by addressing the problem.

- **VM Migration** In a cloud environment havina more number of datacenters, virtual machines have to be migrated between physical machines located in any(same or different) datacenter in order to achieve better provisioning of resources. Migration of Vms which means to transfer a virtual machine from one to another physical machines helps in greatly reducing the energy consumption. The benefits for Vm migration is it avoid the hotspots.

1.2.1   Motivation

Cloud computing offers software, infrastructure and platform as a service in a pay as you go model to end users. There are various research issues in a cloud computing environment such as Vm Migration, Server Consolidation, Security of Data, Energy Consumption, etc. as explained in section before.

One of the core issue is the Management of Energy. Typically size of data centers are consists of hundreds or thousands of servers and resources, and the size of data centres are getting a huge expansion due to the rapid increase in use of Cloud computing technology. This rapid growth has made a increase in energy consumed in clouds.

Thus in order to reduce energy consumption there is a need to effectively utilize the resources and executes the requests. There are various existing methods to manage the workloads but are not able to effectively utilize the resources. The goal of thesis is to effectively manage the workloads so that the utilization of the resources could be maximized and could redeem the consumption of energy. The less consumption of energy provides a Green Computing environment.

1.2.2   Power Consumption Sources

According to data provided by Intel Labs [8], the mainly consumption of power by a server has been used by the CPU, then it is memory in utilizing the power and after that the loss due to the inefficiency in power supply. The data shows that the CPU does not dominates the power consumption by a server.

Current desktop and server CPUs can consume less than 30% of their peak power in low-activity modes, leading to dynamic power ranges of more than 70% of the peak power [9].the power consumed by the CPU further reduced with the evolution of multi core technology. The reason is that the fraction of power consumed by the CPU relatively to the whole system is the adoption of multi-core architectures. Multi-core processors are substantially more efficient than conventional single-core
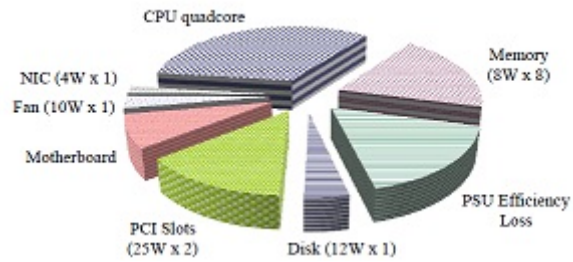
processors.

Figure 1.1: Power Consumption by CPU

1.2.3   Thesis Organisation

The rest of the thesis is organized as follows:

Chapter 2 This chapter describes in detail the literature survey done to study the concept of virtualization, existing techniques of workload consolidation and the green Cloud architecture.

Chapter 3 This chapter describes the problem Analysis of the thesis work. It gives the gap analysis and problem statement.

Chapter 4 This chapter describes in detail the solution of the problem with the help of the Proposed Technique and DFD diagrams.

Chapter 5 This chapter focus on the implementation details and experimental results CloudSim description , Netbeans and snapshots of the simulation.

Chapter 6 This chapter describes the conclusion, contribution to the work done and future research work possible.

Chapter 2

Literature review

This chapter discusses the state of the art and research issues of Virtualization, Resource Allocation policies, Workload Consolidation Techniques and Architecture of Cloud.

2.1 Virtualization in Cloud

Virtualization is the abstraction of physical network, server, and storage resources and it has greatly increased the ability to utilize and scale compute power. It is a technology that allows running two or more operating systems side-by-side on just one PC or embedded controller [7]. Virtualization greatly helps in effective utilization of resources and build an effective system. Many applications are having a limited number of concurrent tasks, thus having a number of unused(idle) cores. This problem can be solved by using virtualization, allocating a group of cores to an OS(Operating system) that can run it concurrently. It enables the service providers to offer virtual machines for work rather than the physical server machines. It forms the basis of Cloud computing on-demand, pay-as-you-go model. The physical server is called the host. The virtual servers are called guests. The virtual servers behave like physical machines. Each system uses a different approach to allocate physical server resources to virtual server needs.

Virtualization also helps in reducing power consumption by reducing the number of physical machines since it provides a number of virtual machines per physical machines and in this way helps in effective utilization of resources. Migration of Vms which means to transfer a virtual machine from one to another physical machines helps in greatly reducing the energy consumption. There are 2 ways to perform migration

(1) Regular migration moves the Vms by pausing the server currently in use, copying the contents then resumes on the moved one machines.

(2) Live migration moves the Vms without pausing the server currently in use, and copying the contents then resumes on the moved one machines. The source server keep on running without intercepting the moved Vms perform its functions.

There are four effects or attributes of IT virtualization;-

- **Density rise** Virtualization results in Higher power density , in some of the racks of server. The areas having high density will pose cooling challenges and if that left unsolved, could hamper the reliability of the overall data-center.

- **IT load Reduction can affect -** Power   Usage   Effectiveness   After virtualization, the data centers power usage effectiveness (PUE) is likely to worsen.

- **Dynamic Workloads**

- **Low redundancy**

Each data center will have a higher or lower PUE curve depending upon the efficiency of its individual devices and the efficiency of its system configuration but the curve always has this same general shape

While virtualization may reduce overall power consumption in the room, virtualized servers tend to be installed and grouped in ways that create localized high-density areas that can lead to hot spots. This cooling challenge may come as a surprise given the dramatic decrease in power consumption that is possible due

to high, realistically achievable physical server consolidation ratios of 10:1, 20:1 or even much higher.

2.2 Survey

The techniques explains the scheduling of workloads considering a number of parameters such as length, number of CPUs required and buffer size of input and output

- M. Steinder et al. [10], proposed the method of how to manage heterogeneous workloads in a data center consists of virtualization. They convey a method of placement of workload dynamically, on the same hardware, in order to increase utilization of resources. The performance function in the method used to check out the performance of different workloads and analyse them.

- Gaweda et al. [11], proposed the method that manage the data centers running the workloads according to the situations. In this method the workloads are located at the server then assigned the resources based on the better utilization of them. The author has used the real system and also the simulated environment to verify the result.

- Yatendra et al. [12], proposed a dynamic compare and balanced algorithm that works on dynamic threshold values[12]. The method is an algorithm uses load balancing and consolidation of server techniques. In the technique, the resource consumption is noted and whenever needed process are migrated so that the load get balanced and thus minimizing the power consumption.

- Liu et al. [13], proposed the technique of assigning the tasks to the most efficient server. The total energy that the datacenter consumes is defined as the sum of energy consumed by task proceeding of all tasks at the data center. The problem solved using a greedy approach. Here the servers are sorted on the

based of their energy efficiency and then the most efficient server is assigned the task.

- Zhibo et al. [14], proposed an algorithm with smart human intelligence to shuffle the task and relocate and then to reduce energy consumption manage the speeds within the constrained. This method is used for heterogeneous workloads.

- Beloglazov et al. [15], focusses on provisioning of resources in a dynamic manner and provides algorithm for efficiently handling of workloads between the datacenter. They proposed (i) architectural principles to efficiently manage the clouds (ii) policies to effectively and efficientyly utilize resource and also scheduling algorithms for that which consider Qos expectations, and data center characteristics of using power. and (iii) a novel software technology for energy-efficient management of Clouds.They have used the cloudsim simulation environment to obtain the performance. Their results obtained after simulation explains that it is a best algorithm to save energy and the migration of virtual machines according to this approach provides the higher energy saving for CPU.

- Ahuja et al. [16] have explores the dynamic placement of applications in a system comprised of Vms , while optimizing the consumption of power and within the SLA constrained. They have proposed the pMapper application placement framework, which comprises of 3 managers and an arbitrator, and they coordinates their respective actions and then make the allocation decisions.

2.3   Cloud Architecture

Green Cloud an architecture of Data Center, whose aims is to reduce data center power consumption, and also to guarantee the users performance lifting the live

migration of virtual machine. There is a challenge for the architecture of making the decision for scheduling on dynamically migrating Virtual machines between the physical machines by reducing the energy uses.

Figure shows the architecture for Green Cloud computing supporting energy-efficient allocation. Basically there are four main entities :

- **DEnd-users** Consumers or end-users they submit their requests from any place in the world to Cloud. A consumer could be a single user or a company having deployed a Web application, that have workload changes based on users online accessing it.

- **Resource Allocator** It acts as an interface between the infrastructure of cloud and end users. It gathers the interaction of the various components to support energy-efficient utilization of resources.

- **Virtual Machines** On a single physical machine many VMs could be managed for working or stop, in order to meet accepted requests, thus it provides maximum flexibility to configure various partitions of resources on the same physical machine to different specific requirements of service requests. Multiple Virtual Machines can simultaneously process the user requests depending on the different Os environments . Also , by the migration of virtual machines across physical machines, taskloads can be gathered and resources which are unused have to made at a state of consuming less power , or turned off for the sake of saving energy.

- **PMs** The physical servers are providing hardware infrastructure which forms the baseline for generating virtualized resources to serve the customers demands.

Chapter 3

Problem Analysis

3.1 Problem Statement

The problem is to have a technique that can effectively allocated task so that the energy wastage will reduce and resources can be effectively utilized. The problem is to determine what kind of applications can be allocated to a single host that will provide the most efficient overall usage of the resources. The approach provided for energy efficient scheduling in data centers do not deals with the problem of gathering different workloads. Those focusses on a single type of workload and also not consider various kinds in the application by considering different workload.

The aim of this thesis work is to propose an efficient task scheduling technique so that the utilization of the resources can be enhance and the consumption of energy in data center will be minimized thus developing a Green-computing environment . A existing techniques is also implemented and compared with the proposed techniques.

Chapter 4

Proposed Technique for Scheduling Workloads

This chapter discusses about how the problem stated in previous chapter can be solved with the help of the proposed Technique.

4.1 Design of Solution

The solution to the problem (Energy Efficient Scheduling of Workload) has been designed by the proposed Technique, and DFD. Following section presents the design of the solution of the proposed Technique and Data Flow Diagrams

Figure 4.1 shows the layered architecture of Cloud computing. PaaS Layer includes the Heterogeneous Workload Consolidation technique to calculate the energy consumption of the data center and also gives the information about the SLA violation, as the allocation policies are implemented on PaaS layer which are followed by the IaaS layer.

Below is the proposed algorithm to calculate the power consumption of the the data center .

- Input: Set of tasks and servers Output: Scheduling of tasks to servers

- for Each Task x of type i do

- for Each S_j do

- Calculate server energy consumption $[E_{i,j}) = P_{i,j} * t_{i,j}$

- if $E(i,j) \leq E(a,b)$ then

- $a = i, b = j$

- end

- end

- end

- while unscheduled tasks remain do

- for Each

$$S_j$$

  do

- Calculate energy consumption

- Assigned task to efficient server

- end

- end

- Schedule Task

- end

4.2 Data Flow Diagram

Data Flow Diagrams (DFD) of Heterogeneous Workload Consolidation Technique implemented by designing a Workload Consolidation CloudPortal(WC CloudPortal), designed for thesis is shown in Figure 4.2, which shows the context level (level 0) DFD of the WC CloudPortal with three entities- new user (want to register), member (registered user) and administrator.
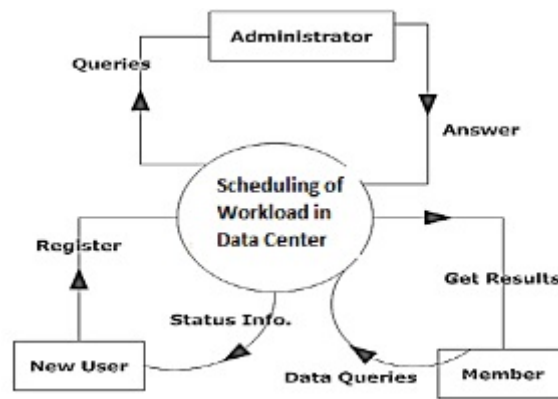

Figure 4.1: Level O DFD

4.3 Proposed Technique with Better Utilization

Below is the proposed algorithm to calculate the power consumption of the the data center by the Workload . The algorithm takes hostlist as an input and gives VMlist of the VMs and power consumed by the data center.

- Input: Set of tasks and servers Output: Scheduling of tasks to servers

- for Each Task x of type i do

- Input the M number of Clouds with L number of Virtual Machines associated with Each cloud.

- Input N number of user process request with some parameters specifications like arrival time, process time, required memory etc, i.e hostlist.

- Arrange the process requests in order of memory requirement

- $Power \leftarrow estimatePower(host, Vm)$

- for Each Task x do

- For each Vm in Vmlist Do

- if host has enough resourse for Vm Then

- $Power \leftarrow estimatePower(host, Vm)$

- else

- Calculate server energy consumption $[E_{i,j}) = P_{i,j} * t_{i,j}$ t $\leftarrow m.getUtil()hostUtil$

- • end

  - end

  - return Vmlist, Energy

  - end

Chapter 5

Result

This chapter focuses on tools for setting Cloud environment, implementation of heterogeneous workload consolidation technique on CloudSim Toolkit and experimental results of this approach.

5.1 Tools for setting Cloud Environment

5.1.1 Cloudsim

CloudSim is an extensible simulation toolkit that enables modeling and simulation of

Cloud computing systems and application provisioning environments. The CloudSim toolkit supports both system and behavior modeling of Cloud system components such as data centers, virtual machines and resource provisioning policies. It implements generic application provisioning techniques that can be extended with ease and limited effort.

CloudSim Architecture

Figure shows the multi-layered design of the CloudSim software framework and its architectural components. The CloudSim simulation layer provides support

for modeling and simulation of virtualized Cloud-based data center environments including dedicated management interfaces for VMs, memory, storage, and bandwidth.
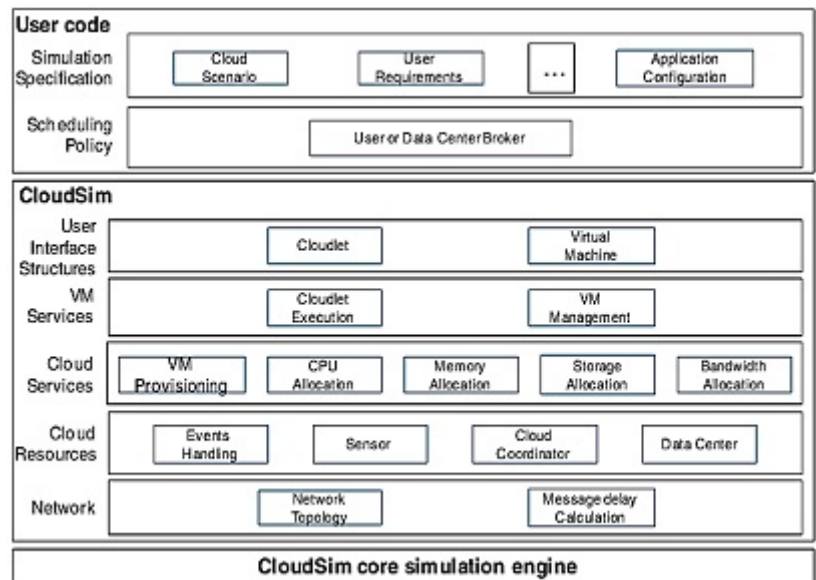

Figure 5.1: Cloudsim Simulation

5.1.2 NetBeans

NetBeans is an integrated development environment for developing primarily with Java, but also with other languages. he NetBeans IDE is written in Java and can run on Windows, OS X, Linux. he NetBeans Platform allows applications to be developed from a set of modular software components called modules. Applications based on the NetBeans Platform including the NetBeans IDE itself can be extended by third party developers.

5.1.3 Implementation of the Proposed Technique

The technique has been implemented the heterogeneous workload in CloudSim Toolkit by using Netbeans We have taken a different number of processors, different

number of tasks and execution times for each task in each processor as input. This is shown in figure

IT shows the generating of virtual machines onto the physical hardware

5.1.4   Results and Conclusions

It can be concluded from the results that the allocation of the virtual machines help to save the energy consumption as the workload will be allocated to the virtual machine having less utilization. So we can set the threshold utilization of the node according to the variability of the workloads as shown in the table.

Chapter 6

Conclusion and Future Scope

This chapter discusses the conclusions of the work presented in this thesis. This chapter ends with a discussion of the future direction which can be taken further.

6.0.5 Conclusion

This thesis gives the introduction of the Cloud computing technique and discusses various workload allocation techniques to efficient manage workloads. In this work a task assignment technique to manage the energy consumption of the data center has been proposed.Technique has been developed in java, deployed on CloudSim toolkit and Experimental results have been gathered.

6.0.6 Thesis Contribution

- In this thesis existing workload assignment techniques have been analyzed and compared according to their features.

- A task assignment technique has been designed and the design also explained through Data Flow Diagrams.

- The designed has been implemented and deployed on CloudSim toolkit by using Eclipse as a run time environment.

- Experimental results have been gathered.

6.0.7   Future Scope

(a)This work shows the energy consumption and SLA violations of the workloads having different characteristics. The other processing element like number of CPUs required by a cloudlet can also be considered to further increase the efficiency of technique.

(b)In future, dynamic resource allocation technique can also be used to manage the heterogeneous workloads and can be validated by working on Cloud environment.

References

1. R. Buyya, C. Yeo, and S. Venugopal, -Market-oriented cloud computing: Vision, hype, and reality for delivering it services as computing utilities, in Proceedings of the 10th IEEE International Conference on High Performance Computing and Communications (HPCC-OB, IEEE CS Press. Los Alamitos, CA. USA), 2008.

2. P. Mell, and T. Grance, - The NIST Definition of Cloud computing, National Institute of Standards and Technology, 2009.

3. Rodrigo N. Calheiros, Rajiv Ranjan, Anton Beloglazov, Cesar A. F. De Rose and Rajkumar Buyya, - CloudSim: a toolkit for modeling and simulation of cloud computing environments and evaluation of resource provisioning algorithms Software: Practice and Experience, Wiley Press, New York, USA, 2010.

4. S. Garg and R. Buyya, - Green Cloud Computing and Environmental Sustainability, Harnessing Green IT: Principles and Practices, S. Murugesan and G. Gangadharan (eds), Wiley Press, UK, in press, accepted on April 2, 2011.

5. , Growth in data center electricity use 2005 to 2010, Analytics Press, Tech. Rep., 2011.

6. G. Chen, W. He, J. Liu, S. Nath, L. Rigas, L. Xiao, and F. Zhao. Energy-aware server provisioning and load dispatching for connection-intensive internet services. In Proceedings of the 5th USENIX Symposium on Networked Systems Design and Implementation, NSDI08, pages 337350, Berkeley, CA, USA, 2008. USENIX Association.

7. Aasys, –Virtualization Basics, Vol. 6, Issue 9, September 2008.

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

9. L. A. Barroso and U. Holzle, The case for energy-proportional computing, Computer, vol. 40, no. 12, pp. 3337, 2007.

10. M. Steinder , D. Carrera, I. Whalley, J. Torres and E. Ayguade, - Managing SLAs of heterogeneous workloads using dynamic application placement, HPDC '08 Proceedings of the 17th international symposium on High performance distributed computing, New York, USA, 2008.

11. . Gaweda , M. Steinder, I. Whalley, D. Carrera, and D. M. Chess, -Server virtualization in autonomic management of heterogeneous workloads, Integrated Network Management, page no. 139-148, 2007.

12. Yatendra Sahu , R.K. Pateriya , Rajeev Kumar Gupta , - Cloud Server Optimization with Load Balancing and Green Computing Techniques Using Dynamic Compare and Balance Algorithm, Proceedings of 5th International Conference on Computational Intelligence and Communication Networks, page no. 527- 531,2013.

13. Ning Liu, Ziqian Dong, Roberto Rojas-Cessa, - Task and Server Assignment for Reduction of Energy Consumption in Datacenters, IEEE 11th International Symposium on Network Computing and Applications ,2012.

14. Zhibo Wang, Yan-Qing Zhang - Energy-Efficient Task Scheduling Algorithms with Human Intelligence Based Task Shuffling and Task Relocation, IEEE/ACM International Conference on Green Computing and Communications, page no 38-43 , 2011.

15. A. Beloglazov ,R. Buyya, and J. Abawajy, Energy-efficient management of data center resources for cloud computing: a vision, architectural elements,

and open challenges, In: International conference on parallel and distributed processing techniques and applications (PDPTA), Las Vegas, USA, 2010.

16. P. Ahuja ,A. Verma, and A. Neogi, pMapper: power and migration cost aware application placement in virtualized systems, in Proceedings of the 9th ACM/IFIP/USENIX International Conference on Middleware. Springer-Verlag New York, page no. 243264, 2008.

17. T. Dillon, C. Wu, and E. Chang, Cloud Computing: Issues and Challenges, in 24th IEEE International Conference on Advanced Information Networking and Applications, page no. 27-33, 2010.

18. J. Hamilton, Cooperative expendable micro-slice servers (CEMS): low cost, low power servers for Internet-scale services, In the Proceeding of CIDR, 2009.

19. J. Koomey, Estimating total power consumption by servers in the US and the world, Final report, vol. 15, February 2007.