Analysis of Cloud Sim Toolkit for Implementing Energy Efficient Green Cloud Data Centers

Sajitha A V¹, Dr. A C Subhajini²

¹Research Scholar, Department of Computer Applications, Noorul Islam Centre for Higher Education, Tamil Nadu, India
²Assistant Professor, Department of Computer Applications, Noorul Islam Centre for Higher Education, Tamil Nadu, India

Abstract: Cloud computing is an emerging technology which emphasis on delivering software, platform, infrastructure to the user, based on demand by cloud services providers through general internet-based applications. Consumer can pay for using these resources on the basis of pay as you go model. Cloud Computing is facing many challenges from its seedling stage to yet. Energy management, Load balancing, Server consolidation, Resource utilization, VM Migration cost and time modeling, Security and Network throughput achievements are various known issues. Out of this energy consumption awarding the first position because of the number of energy hungry data centres are growing rapidly due to the increasing popularity of the Cloud Computing among all levels of organizations. But running the applications in the real cloud for testing are much more expensive and researchers often cannot attain the real cloud environment. To conquer these challenges, a test bed of cloud simulator is necessary. Simulation is the act of imitating or pretending like real environment. This paper introduces the characteristics and working of a cloud simulator framework named CloudSim which provides simulation, supremacy to manage cloud services as well as modeling of cloud infrastructure.

Keywords: Cloud Computing, Data Centre, Energy Efficiency, Cloud Sim, Simulation, Green Cloud.

I. INTRODUCTION

Cloud computing is a versatile technology which can bear the wide spectrum of resources as well as applications. It is gaining its esteem in the midst of different types of organizations that have the aspirations to eliminate the strain of managing their own IT infrastructure via outsourcing the computational resources. It is a technological framework which supplies computing resources as a utility like electricity on a metered basis that authorizes the operations of the workloads similar to grid computing [1]. It provides high reliability, broad network access, fault tolerance, on demand services etc as compared to traditional computing. Cloud provides Infrastructure as a Service (IaaS), Platform as a Service (PaaS) and Software as a Service (SaaS) as service models. Amazon EC2 is an eminent IaaS provider, Google App is a well known PaaS provider and Salesforce.com is a prominent SaaS provider. Cloud Computing distributes its resources, such as, software, storage, databases and network in reliable, economical and secure manner to a wide number of users via conventional internet protocols and these resources are accumulated in geographically dispersed energy hungry data centers.

The conventional cloud applications are social networking, content delivery, web hosting, real time processing of data etc. These kinds of applications have dissimilar requirements for their set up and organization. The system volume, power performance and unreliable loads are the various issues of these applications’ deployment under cloud federation.

The amount of energy consumption of data centers is doubling in every four years. In 2015, U.S. data centers consumed around 70 billion KWh, that in lieu of approximately 2% of aggregate electricity consumption of U.S.A. Based on recent trend, U.S. data centers are projected to expend around 73 billion KWh in coming year 2020 [2]. Each datacenter is emitting 170 million metric tons of Carbon Dioxide per year.

The anticipated global carbon emissions by data centers in 2020 will be 670 million metric tons per annum. Presently, data centers devour up to 3% of whole worldwide power generation while delivering 200 million metric tons of CO₂. Studies shows that the average consumption of data centers can be nearly 20% and energy utilized by the idle resource is can be as regard as 80% of its own peak power. Low Return on the Investment (ROI), system instability, as well as much more CO₂ emissions can be created because of this high amount of energy usage of data centers. Thus an immediate and improved means is necessary to overcome this crisis to achieve Green Cloud data centers. Green Cloud computing is an innovative model which foreseen to accomplish not only efficient processing with the optimum utilization of computing resources, but also decrease the energy consumption.
A. Research Issues Of Cloud Computing

Cloud architecture is a very complex and heterogeneous model where resources are scattered globally. Cloud Computing research tends to the issues of meeting the necessities of next generation private, public and hybrid cloud computing deployments, likewise the issues of enabling applications and advancement platforms to exploit the benefits of cloud computing. Some of the challenging research issues in cloud computing are given below [20].

1) Energy Management
2) Virtual Machine Migration
3) Service Level Agreement (SLA)
4) Resource Provisioning
5) Server Consolidation
6) Network bandwidth
7) Latency
8) Cloud Data Management & Security
9) Interoperability
10) Multi-tenancy

B. Use Of Simulation

The modeling and testing in a real cloud environment is:

1) Costly due to the enormous number of resources is required for a long period.
2) Repetition of testing in real environment is not possible.
3) By depending on the service provider’s particular infrastructure may not be supported scalable application

Simulation is the act of imitating or pretending like real environment[3]. The primary objectives of simulation are:

4) Elasticity of defining configurations.
5) Easy to use and customization.
6) Economical by using the miniature of real cloud for creation and upgrading phases

For Modeling and Simulation of Cloud Computing Infrastructures and Services, a simulator is available in market is named as CloudSim created by University of Melbourne, Australia [6]. CloudSim is a toolkit not only for modeling and simulating cloud environments but also for assessing resource provisioning algorithms. The code of this open source simulator is designed in JAVA language and can be used in Windows/Linux/UNIX operating system. The pioneer version of CloudSim 1.0 is came into existence in 2009. In May 24, 2016 the latest version of CloudSim is released with advanced features such as: added support for Container virtualization and lots of bug fixes from CloudSim 3.0.3 which was released in the year March 19, 2015.

II. RELATED WORK

In the last decade of 20th century, Grids[4] have monopolized the infrastructure of distributed system for conveying high performance services for computing and data demanding scientific applications. There are numerous Grid simulators such as, GridSim, SimGrid, OptorSim, GridNet, GridLAB-D, GangSim etc. have been proposed for developing and testing such grid based environment. But unfortunately they are not suitable to obviously isolate the multi-layer service abstractions (SaaS, PaaS, and IaaS) of cloud computing which is the leader of this century. None of them supports the designing of the virtualized cloud computing environment. Thus a new simulator was necessary to support and model the virtualized infrastructure in multi-layered architecture of virtualized cloud environment. CloudSim offers an excellent modeling and simulation capabilities to study energy-aware data centre management in multilayered architecture of virtualized cloud environment.

1) Wang, L et al. [5], have described the basics of cloud computing and the importance of various services as well as deployment models in virtualized environment.
2) Calheiros, R. N. et al. [6], have presented a Simulation technique - CloudSim. They have described the various attractive features of CloudSim that it supports for modeling and testing for large-scale cloud computing infrastructure including data centers on a single physical computing node.
3) Chowdhury, M. R. et al. [7], have analyzed various VM placement strategies in CloudSim simulation environment. They have argued with result that their algorithms outperforms the indigenous algorithms of CloudSim. This proved that CloudSim has a flexibility to execute foreign algorithms also.
4) Ouarnoughi, H et al. [8], have discussed the processing of I/O workload generate by virtual machines within a data center, and evaluates the overall performance and energy consumption as an extension of CloudSim. They modified the time computation model and CPU utilization model for calculating energy consumption of the CloudSim relating to I/O operations held in data centre. This is a good sign of CloudSim for its adaptability.

5) Yang, W et al. [9], have analyzed the Network Function Virtualization (NVF) cloud framework named CloudSimNVF in an energy efficient model in multi-dimensional extension based on CloudSim. They achieved an accurate result as same as real cloud environment.

6) Khoshkholghi, M.A et al. [10], have proposed energy-efficient and SLA-aware dynamic VM consolidation Mechanism (PCM) which was the combination of four algorithms: Over loading host detection, Underloading host detection, VM Selection and VM Placement algorithms. In this approach they consider multiple parameters for predicting the host utilization based on main memery, CPU and network Bandwidth. They proved the effectiveness of their algorithm by means of CloudSim toolkit simulation with other four benchmarking algorithms by improving energy efficiency and evading SLA violation.

7) Son, J.et al. [11], have presented Dynamic Overbooking algorithm which jointly controls virtualization potentials as sell as Software Defined Networking (SDN) for VM along with traffic consolidation. Through energy efficient network overbooking capacity, SDN can unite network traffic along with Quality of Service (QoS). In this approach they projected resource allocation ratio based on the historical data monitoring from the online analysis of the host and network usage with no pre-knowledge about the workloads. They can be achieved energy efficiency and a enormous energy cost savings by minimizing the abusement of over provisioned resources, then mutually reduced the SLA violation by distributing adequate resources for the real workload. They have implemented their algorithms in CloudSimSDN which is CloudSim based simulation tool which supports a variety of SDN behaviors for example, dynamic network design, programmable controller etc.

8) Koutsandria, G et al. [12], have suggested a task scheduling algorithm (MinDelay) that provide accommodation for service request and Energy Conscious Algorithm (ECTC) that decreases energy consumption in a cloud system in a live VM migration procedure. They have conducted their experiments with the help of CloudSim toolkit which provide an extensive support to accomplish their goals.

9) Abdullah, M.et al. [13], have investigated a novel heuristic approach for the dynamic VM consolidation of virtual machines (VMs) in terms of energy efficient cloud datacenters. The Fast Best Ft Decreasing(FBFD) algorithm for intellectual VMs assigning into hosts and Dynamic Utilization Rate (DUR) algorithm for utilization space and VM migration are successfully evaluated in CloudSim simulation environment

10) Nguyen, T. H. et al. [14], have discussed a virtual machine consolidation algorithm with multiple usage prediction (VMCUP-M) in context of various resource type and its predicted utilization to improve the energy efficiency of cloud data centers. They have reached their goals with real and synthesis workload

Analysis with the support of CloudSim simulation toolkit. They stated that CloudSim is an excellent toolkit for handling multiple types of resources.

Different research work carried out using CloudSim

<table>
<thead>
<tr>
<th>Author</th>
<th>Scheme</th>
<th>Simulator</th>
<th>Achievement</th>
</tr>
</thead>
</table>
| Farahnkian, F et al.[15] | Ant Colony System (ACS)- online optimization metaheuristic algorithm | CloudSim   | i. Reduces the energy consumption of data centers by consolidating VMs into a reduced number of active hosts  
ii. Preserved the QoS requirements |
| Han, G. et al.[16] | Remaining Utilization-Aware (RUA)algorithm, Power-Aware (PA) algorithm | CloudSim   | i. Reduction of SLA violations  
ii. Improvement of the system performance  
iii. Saving energy by switching off under utilized hosts. |
| Singh,S.et al. [17] | Energy-aware Autonomic Resource management Technique(EARTH) | CloudSim   | Perform better in terms of:  
i. Resource utilization  
ii. Energy consumption |
| Farahnkian, F. et al.[18] | Utilization Prediction-aware VM Consolidation (UP-VMC) | CloudSim   | i. Reduces the energy consumption  
ii. Diminishes number of VM migrations  
iii. Minimizes the number of SLA violations. |
| Abdelsamea, A et al.[19] | Multiple Regression Host Overload Detection (MRHOD) | CloudSim   | i. Reduces energy consumption  
ii. Ensuring SLA |
III. CLOUDSIM TOOLKIT

CloudSim legitimize researchers to inspect and scrutinize every layer (Infrastructure-as-a-Service, Platform-as-a-Service, and Software-as-a-Service) of the cloud computing environment in a cost effective, secure and scalable way.

A. Characteristics Of Cloudsim

The kernel of CloudSim toolkit is mentioned below [21]. It provides:

1) Assistance for modeling and simulation of wide ranging Cloud computing data centers.
2) Assistance for modeling and simulation of application repositories.
3) Assistance for modeling and simulation of energy-conscious computational assets.
4) Assistance for modeling and simulation of virtualized hosts, with tailor-made policies for provisioning host machines to virtual machines.
5) Assistance for modeling and simulation of federated clouds.
6) Assistance for modeling and simulation of data center network topologies as well as message-flowing applications.
7) Assistance for active incorporation of simulation parameters, stop and continue of simulation.
8) Assistance for user-defined strategies for provision of physical machine to virtual machines and strategies for allocation of physical machine assets to virtual machines.

B. Cloudsim Architecture

The CloudSim layer offers support for modeling and simulation of cloud environments including memory, storage, bandwidth and VMs. It also supplies hosts to VMs, application execution supervision and dynamic system status monitoring. The layered CloudSim architecture is shown in Fig1. A cloud service provider can implement customized methods at this layer to study the competence of different policies in VM provisioning [6].

C. Class Diagram Of Cloudsim

CloudSim is defined as “a library for the simulation of cloud scenarios. It provides essential classes for describing data centres, computational resources, virtual machines, applications, users, and policies for the management of various parts of the system such as scheduling and provisioning.”[22]. The design of the class diagram is shown below in Fig2. Some of the important classes of the framework are described in short.
Regions: It models geographical regions in which cloud service providers assign resources to their clients. In cloud analysis, there are six regions that correspond to six continents in the world.

Data centres: It models the infrastructure services offered by a variety of cloud service providers. It summarizes a set of computing hosts or servers that are either heterogeneous or homogeneous in nature, according to their hardware configurations.

Hosts: It models physical resources (compute or storage).

Cloudlet: It states the set of user requests. It consists of the application ID, name of the user base that is the persons to which the responses have to be routed back, and the size of the request execution commands, as well as input and output files. It models the cloud-based application services.

Service broker: The service broker settles on which data centre should be chosen to provide the services to the requests from the user base.

Virtual Machine: It represents the characteristics of Virtual machine such as RAM, CPU, Network bandwidth, virtualization monitor, hosting Physical Machine etc are encapsulated here.

VMM allocation policy: It generates provisioning policies on how to assign VMs to hosts.

VM scheduler: It models the time or space shared, scheduling a policy to assign processor cores to VMs.

Network Topology: It implements network layer in CloudSim. It reads a BRITE file and generates a to pological network from it which is used to simulate latency in network traffic of CloudSim.

IV. INDIGENOUS ENERGY-AWARE ALGORITHMS AND POLICIES OF CloudSim

CloudSim has its own default VM Placement algorithms and VM allocation as well as VM Selection policies. These methods help to achieve energy efficient data centres in cloud environment through proper resource utilization. The following sections giving a brief description about them.

A. VM Placement Algorithms

The placement of VMs among PMs is treated as a bin packing problem[24]. Some popular solutions of this problem are:

1) The First Fit (FF): First Fit begins with the most lively bin and tries to pack every item in it before going into next bin. If appropriate bin is not to be found for the item, then the subsequent bin is elected to locate as the new bin.

2) First Fit Decreasing (FFD): In FFD the items are arranged in descending order and after that items are processed as in the method of using First Fit algorithm. 

3) Best Fit Decreasing (BFD): Like FFD, BFD also arranges items in descending order and afterwards for packing items it prefers a bin with minimum vacant space to be left there after the item is being packed.

4) Worst Fit Decreasing (WFD): Worst Fit Decreasing works accurately equal to BFD apart from in one thing, rather than selecting bin with least empty space it opts bin with greatest empty space to be left there after the allotment of item on that bin.

5) Second Worst Fit Decreasing (SWFD): Commensurate WFD, it just select bin with second least empty space. It is also called as Almost Worst Fit Decreasing (AWFD).
B. VM Allocation Policies

At the point when a host is treated as being over utilized, this implies the VM is as of presently utilizing more MIPS than areas of presently accessible on that host or that it is above the threshold value set by the individual allocation policies. Some of VM allocation policies are:

1) Median Absolute Deviation (MAD) – MAD is a powerful measurement which measures the changeability of a set of facts. This is like a standard deviation however is measured to be a more tough measurement as it is less delicate to exceptions in a data index. This is because of the way that standard deviation utilizes a squared mean which can significantly change the variation if the information point is significantly adequate. For MAD, a vast deviation or even a couple of extensive deviations have no impact on the median.

2) Interquartile Range (IQR) - IQR is a measurement for unevenness, in light of isolating a data set into quartiles. Quartiles partition a position-oriented data set into four equivalent parts. The values that split each element are known as the primary, secondary, and third quartiles; then they are denoted by Q1, Q2 and Q3, correspondingly. The IQR is a symmetric distribution because of its connection to the MAD.

\[ 0.5 \times IQR = \text{MedianAbsoluteDeviation} \]

3) Local Regression (LR) - LR technique is depends on the LOESS local regression algorithm. This algorithm takes a restricted set of the data and endeavors to discover the deviation in the data from point to point. This is done by utilizing locally weighted polynomial regression methods. While utilizing the weighted least squares technique, the data points that are nearer to the approximate to get weighted greater than the data points that are auxiliary from the approximation.

4) Robust Local Regression (RLR) - RLR is as stated to be robust when its algorithm is not sensitive to an off base presumptions. By means of the similar LOESS regression algorithm as the local regression data set, but changing the system to which the values are weighted enables this regression algorithm to be viewed robust. This implies that the algorithm could be deaden to an outlying model which might be conciliation the strength of a standard regression algorithm.

C. VM Selection Policies

Formerly a PM has been decided as over utilized by the VM allocation policy, the choice strategies will figure out which VM’s will be migrated from the aforementioned PM to bring the PM back to a more competent rank.

1) Minimum Migration Time (MMT) – This elementary selection strategy analyze the amount of RAM consumed by each and every VM inside a PM. The bigger the VM, the higher data there will be to shift, so the littlest VM in the RAM of the PM will be the excellent to migrate.

2) Random Selection (RS) – The RS technique accesses the number of VMs within a PM and forms a random digit generated starting from 0 to that range. It applies the index coupled with each VM in a PM to establish which VM has that allocated digit.

3) Minimum Utilization (MU) – The MU policy brings out the list of VMs within an over utilized PM and locates the VM with the lowest MIPS requisites for that time span and forwards it in to migration process.

4) Maximum Correlation (MC) – The utmost utilization brings out a list of migratable VMs from a PM that has been decided to be over utilized. This list is applied to construct a migration grid in accordance with the volume of the list and the least amount consumption of the list. The correlation coefficient is established by taking the transpose of the migration grid so that it is becoming a linear model and afterward it utilizing linear regression. The LR technique is utilized as the similar linear regression technique for the linear regression allocation policy. While assessing the correlation coefficients of the VMs, the VM among the uppermost coefficient is the VM to be migrated.

V. INSTALLATION Of CloudSim IN Java NetBeans IDE

Cloudsim is a tool for cloud computing simulation written in Java language. So the basic knowledge of JAVA programming and cloud computing is essential for using CloudSim. Cloudsim does not need to be installed, because it is a library. Thus it is only to unzip the package and then add .jar file as a library into new project. It can be worked with any programming IDEs that support Java such as NetBeans or Eclipse. To use CloudSim with NetBeans [23,25]:

1) Download NetBeans IDE from https://netbeans.org/downloads/ and install it.

2) Download CloudSim installable files from https://code.google.com/p/cloudsim/downloads/list and unzip.

3) Open NetBeans IDE

4) Create a new project (File -> New project).

5) Choose “Java” under Categories and “Java Application” under Projects.
6) Name the new project and choose a file path. Uncheck “Create Main Class” box. Then click “Finish” button.
7) Now add CloudSim library into the project.
8) Right click on the project and click on “Properties”. A project properties dialog will appear. Select “Libraries” and click on “Add Library” button at the right. In the “Add Library” dialog, press “Add jars/Folder” button.
9) Goto the folder that CloudSim unpacked. There shows its library at “jars” folder. Then select file “CloudSim.jar” and add it into the project’s library.

VI. EVALUATION AND EXPERIMENTAL SETUP
The simulation is used Windows 10 Excelance 32-bit, Intel® Atom™ processor, 1.336GHz with 2GB memory. For the simulation special toolkit needed known as CloudSim 3.0 toolkit. Java NetBeans IDE8.2 used as programming language. The simulation comprises the implementation of Cloudlets by VMs and the allocation of these VMs on PMs.

A. Initialize the CloudSim package
   - Int num_user = 1; //number of cloud users
   - Calendar calendar = Calender.getInstance();
   - Boolean trace_flag = false; //mean trace events

B. Initialize CloudSim library:
   - CloudSim.init(num_user, calendar, trace_flag)

C. The second steps the creation of Data Centres. Data centres are the resource providers in CloudSim. To create Datacenter, we need the Datacenter Characteristics object that stores the properties of a data centre such as architecture, OS, list of machines, allocation policy that specifies the time or space shared, the time zone and its cost.
   - Datacenter datacenter_0 = createDatacenter(“Datacenter_0”,characteristics);

D. The third step is to create a broker:
   - DatacenterBroker broker = createBroker();

E. The fourth step is to create one virtual machine
   - Vmlist = new ArrayList<Vm>();
   - VM Properties:
     - Int vmid = 0;
     - Int mips= 1000;
     - Long size = 10000; //image size (MB)
     - Int ram = 512; //vm memory (MB)
     - Long bw = 1000;
     - Int pesNumber = 1; //number of cpu
     - String vmm = “Xen”; // VMM name
     - Vm vm = new Vm(vmid, brokerId, mips, pesNumber, ram, bw, size, vmm, new CloudletSchedulerTimeShared());

F. Submit the VM list to the broker
   - Broker.submitVmList(vmlist);

G. Create a cloudlet with length, file size, output size, and utilization model:
   - cloudLetList = new ArrayList<Cloudlet>();
   - Cloudlet Properties:
     - Int id = 0;
     - PesNumber = 1;
     - Long length = 400000;
     - Long fileSize = 300;
     - Long outputSize = 300;
UtilizationModel utilizationModel = new UtilizationModelFull();
Cloudlet cloudlet = new Cloudlet(id, length, pesNumber, fileSize, outputSize, utilizationModel, utilizationModel1);
Cloudlet.setUserId(brokerId);
Cloudlet.setVmId(vmid);

H. Add The Cloudlets To The List:
Broker.submitCloudletList(cloudLetList);
cloudLetList.add(cloudlet);

I. Submit Cloudlet List To The Broker:
Start Simulation :
CloudSim.startSimulation();
CloudSim.stopSimulation();

J. Final step : Print result when simulation is over
List<Cloudlet> newList = broker.getCloudletReceivedList();
printCloudletList(newList); // Print the debt of each user to each datacenter Datacenter0.printDebts();
r
Starting CloudSimExample2...
Initialising...
Starting CloudSim version 3.0
GlobalBroker is starting...
Datacenter_0 is starting...
Broker_0 is starting...
Entities started.
0.0: Broker_0: Cloud Resource List received with 1 resource(s)
0.0: Broker_0: Trying to Create VM #0 in Datacenter_0
0.1: Broker_0: VM #0 has been created in Datacenter #1, Host #0
0.1: Broker_0: Sending cloudlet 0 to VM #0
Adding: GlobalBroker_
GlobalBroker_ is starting...
200.0: GlobalBroker_: Cloud Resource List received with 1 resource(s)
200.0: GlobalBroker_: Trying to Create VM #100 in Datacenter_0
200.1: GlobalBroker_: VM #100 has been created in Datacenter #3, Host #1
200.1: GlobalBroker_: Sending cloudlet 100 to VM #100
320.1: Broker_0: Cloudlet 0 received
320.1: Broker_0: All Cloudlets executed. Finishing...
320.1: Broker_0: Destroying VM #0
320.1 Broker_0 is shutting down...
520.1: GlobalBroker_: Cloudlet 101 received
520.1: GlobalBroker_: All Cloudlets executed. Finishing...
520.1: GlobalBroker_: Destroying VM #100
GlobalBroker_ is shutting down...
Simulation: No more future events
CloudInformationService: Notify all CloudSim entities for shutting down.
Datacenter_0 is shutting down...
Datacenter_1 is shutting down...
Broker_0 is shutting down...
GlobalBroker_ is shutting down...
Simulation completed.
Simulation completed.

============= OUTPUT =============

Cloudlet ID  STATUS  Data center ID  VM ID  Time  Start Time  Finish Time
0  SUCCESS  3  0  320  0.1  320.1

*****Power Datacenter: Datacenter_0

User id  Debt
2  53.6

CloudSimExample2 finished!
BUILD SUCCESSFUL (total time: 2 seconds)

VII. CONCLUSION

In this work, we conduct a detailed survey of CloudSim tool kit related to energy aware VM selection, allocation and placement strategies. We present a comprehensive information about default energy conscious VM placement algorithms in CloudSim. Through this survey we reached a conclusion that CloudSim is a better simulation kit for energy oriented data centre evaluation in dynamic environment because it is extensible open source software. In short, it is supporting not only its default algorithms but also other foreign algorithms.

In future we will conduct an exhaustive study about the traffic-aware data centre infrastructure with the support of CloudSim toolkit by utilizing its adaptability and extensibility nature.

ACKNOWLEDGMENT

We would like to acknowledge our thanks to Dr. Rajkumar Buyya, CLOUDS Laboratory, University of Melbourne, Australia for providing an excellent research article in his web portal (www.buyya.wordpress.com) as for Cloud Computing. Also wish to convey our gratitude to all people who help us in different ways to fulfill our research work.

REFERENCES


