A Survey on Application of Cloudsim Toolkit in Cloud Computing

R.Kanniga Devi 1, S.Sujan 2

Assistant Professor, Department of Computer Science and Engineering, Kalasalingam University, Tamil Nadu, India 1

P.G Student, Department of Computer Engineering, Kalasalingam University, Tamil Nadu, India 2

ABSTRACT: Cloud computing is a paradigm of large scale distributed computing which is a repackaging of various existing concepts/technologies such as utility computing, Grid computing, Autonomic computing, Virtualization and Internet technologies. Cloud computing use the Internet technologies for delivery of resources 'as a service' to users on demand. As it is a developing technology, various issues such as security, energy management, resource provisioning, reliability need to be addressed. The objective of the work is to study a simulation toolkit known as CloudSim and some of the research works done to address some of the above mentioned issues using this toolkit. Most of the studies in Cloud rely on simulation based experiments, since using real cloud infrastructures such as Amazon EC2, Google are cost and time consuming tasks and also large number of cloud resources is required to achieve realistic results which is impractical for many researchers.

KEYWORDS: Cloud Computing, CloudSim toolkit.

I. INTRODUCTION

Cloud computing is a paradigm of large scale distributed computing which is a repackaging of various existing concepts/technologies such as utility computing, Grid computing, Autonomic computing, Virtualization and Internet technologies. Users can access Cloud infrastructure from anywhere at any time in an on demand manner and pay for what they use. Cloud uses virtualization technology in distributed data centers to allocate resources to the customers over the Internet in the form of configurable virtual machines and provides an illusion of infinite resources. The main principle behind this model is offering computing, storage, networking and software "as a service" and it can be termed as "Computing as a utility".

This paradigm is used to deploy applications with high scalability, processing and storage requirements. It helps organizations in avoiding capital costs of hardware and software. Web browser operates as an interface between clients and cloud and the cloud provides the entire infrastructure. The end user has no idea of where the data is stored, which processor is executing the task and how the entire process is done. Users just sign up to an application.

The rest of the sections is organized in the following way: Following the introduction to cloud computing, related work is given in section II. Research issues in cloud is summarized in section III. Need of simulation is discussed in section IV. CloudSim tool kit and its application in various research areas is described in section V. Finally, Conclusion and scope for future work is given in section VI.

II. RELATED WORK

Several Grid simulators such as GridSim, SimGrid, OptoSim and GangSim have been proposed for modeling of Gridbased environments . The toolkits are not capable of isolating the multilayer service abstractions (SaaS, PaaS, and IaaS) required by Cloud computing environments. There are no support given for modeling virtualized environment. So we find that CloudSim toolkit is the right tool for modeling large scale virtualized cloud environment which has components for creating data centers, hosts, virtual machines, brokers and service requests.
III. GENERAL RESEARCH ISSUES IN CLOUD COMPUTING

Cloud infrastructure is a very complex and heterogeneous system where resources are distributed globally. It has some general issues as discussed in the following:

a) Security challenges
Users suspect on security aspects of cloud environment since their data may sit next to their competitors data when data is outsourced to an unknown environment. Most important security and privacy attributes are confidentiality, integrity, availability, accountability, and privacy-preservability. Some care has to taken to preserve these attributes.

b) Cost
The public cloud offers pay per use model, some time the usage cost may exceed purchase cost. Need of algorithms for optimized resource allocation and usage is mandatory.

c) Reliability
Cloud is a high complex system in nature which may lead to large number of failures in distributed data centers. Fault tolerance serve as an effective means to address reliability concerns. Fault tolerance means that system should continue to operate under fault presence [26].New Fault tolerance approaches for cloud is needed since traditional fault tolerance approaches are less effective because of the abstraction layers of cloud computing. Fault tolerance that uses the Virtualization Technology (VT) can increase the reliability of applications but VM migration and consolidations are difficult to achieve.

d) Network Bandwidth
Massive amount of data movement to and from a data center increases the need of more bandwidth and need proper allocation algorithms to reduce latency issues.

e) Resource Provisioning
Complex distributed environment of cloud has to incorporate efficient resource provisioning algorithms to meet Service Level Agreement (SLA) requirements. Virtualized data centers are providing better management flexibility, lower cost, scalability, better resources utilization, and energy efficiency, but virtualization is not an easy task to do.

f) Latency
Latency can affect the performance and reliability of application content delivery to end user. One way to reduce latency is to choose a cloud service provider who is nearer to the user using edge catching techniques and designing optimal routing technologies.

g) Interoperability
Cloud providers should be interoperable so that they can hire resources from other providers when their resources are not sufficient to service the consumers. But distinct VM technologies will prevent them from being interoperable hence open standards are required to solve the problem.

h) Energy
Large scale data centers are consuming enormous amounts of electrical power dissipating much amount of heat and emit carbon dioxide which is unsafe for environment. Hence number of practices need to be applied to achieve energy efficiency, such as improvement of applications’ algorithms, energy efficient hardware and energy efficient resource management strategies on a virtualized data center.
IV. WHY SIMULATION

Simulation tools are widely used to study cloud environments and it is necessary to know the pros and cons of these simulators.

The Experimentation in a real cloud environment is [25]

1. Expensive, since huge amount of resources is needed for a long period.
2. Repetition of experiments is not possible.
3. Dependence on the provider’s specific infrastructure may not support scalable application.

Due to these reasons, simulation has been chosen as the right environment in which users experiment risk-free. There are few simulation tools available for cloud environment, of which some are commercial and some are open source. In this paper an open source cloud simulation toolkit, CloudSim is chosen and studied various works done to address some of the research issues mentioned in section III.

V. CLOUDSIM TOOLKIT

CloudSim [2] is a Cloud computing modelling and simulation tool that was developed in the University of Melbourne, Australia. It is an experimental tool to conduct new research approaches. It supports the modelling and simulation of large scale Cloud computing environments, including data centers, computing nodes, Virtual machines, resource provisioning, and virtual machine provisioning, and power management.

CloudSim is written in Java and initially was built on top of GridSim, a tool for Grid environment simulation. We need at least basic understanding of how to program in Java some basic Object Oriented Programming (OOP) concepts to work with CloudSim. Current version available is 3.0.3.

a. CloudSim in Cloud research

This section presents a broad survey of research works carried out using CloudSim toolkit.

<table>
<thead>
<tr>
<th>S.No</th>
<th>Author</th>
<th>Research issue considered</th>
<th>Research Work</th>
<th>Technique applied</th>
</tr>
</thead>
<tbody>
<tr>
<td>5.</td>
<td>Manasa H.B et al., [2013]</td>
<td>Energy aware resource Provisioning</td>
<td>Energy aware allocation of data center resources</td>
<td>Modified Best Fit Decreasing algorithm and Migration Algorithm</td>
</tr>
<tr>
<td>Provisioning</td>
<td>(DVFS)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>--------------</td>
<td>--------</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7. Weiwei Chen et al., Resource Provisioning</td>
<td>Workflow modelling</td>
<td>Extended CloudSim to model workflow management systems (WMS)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>8. Bhathiya Wickremasinghe et al., Resource Provisioning</td>
<td>Social network analysis (FaceBook)</td>
<td>Extended CloudSim for GUI support</td>
<td></td>
<td></td>
</tr>
<tr>
<td>9. Gaetano F. Anastasi et al., Resource Provisioning</td>
<td>Smartcloud federation</td>
<td>Federation modelling</td>
<td></td>
<td></td>
</tr>
<tr>
<td>12. Foued Jrad et al., Resource Provisioning</td>
<td>SLA based Intercloud service broker</td>
<td>New Broker design</td>
<td></td>
<td></td>
</tr>
<tr>
<td>13. Vincent C. Emeakaroha et al., Resource Provisioning</td>
<td>SLA based resource allocation</td>
<td>Scheduling heuristic</td>
<td></td>
<td></td>
</tr>
<tr>
<td>20. Ts'epoMofolo et al., [2013] Resource Provisioning</td>
<td>Heuristic based resource allocation</td>
<td>Modified Best Fit Decreasing Algorithm</td>
<td></td>
<td></td>
</tr>
<tr>
<td>23. Palvinder Singh Mann et al., [2011] Security</td>
<td>Handling DDoS attack</td>
<td>Mitigation Algorithm based on analytical method</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Table 1: CloudSim in Cloud research*
Table 1 and Figure 1 lead to a conclusion that more concentration is given to resource provisioning issue and each of these works is summarized as below:

R. N. Calheiros et al., [1] propose CloudSim: an extensible simulation toolkit that enables modelling and simulation of Cloud computing systems and environments. A hybrid Cloud model has been evaluated using CloudSim and the performance gains as regards to response time and cost saving of using such as Cloud model is evaluated under different workload scenarios.

Rajkumar Buyya et al., [2] propose CloudSim: an extensible simulation toolkit that enables modelling and simulation of Cloud computing environments. The CloudSim toolkit supports modelling and creation of one or more virtual machines (VMs) on a simulated node of a Data Center, jobs, and their mapping to suitable VMs. It also allows simulation of multiple Data Centers to enable a study on federation and associated policies for migration of VMs for reliability and automatic scaling of applications.

Rodrigo N. Calheiros et al., [3] propose CloudSim toolkit which supports both system and behaviour modelling of Cloud system components such as data centers, virtual machines (VMs) and resource provisioning policies. It implements generic application provisioning techniques that can be extended with ease and limited efforts. It supports modelling and simulation of Cloud computing environments consisting of both single and internetworked clouds (federation of clouds).

A. Venkat Reddy et al., [4] propose a framework which serves as an excellent platform for making cloud services intrusion tolerant. The feasibility of the framework has been tested by making cloud’s Infrastructure as a Service (IaaS) and Data Storage Service intrusion tolerant. The proposed framework has been validated by integrating Intrusion Tolerance via Threshold Cryptography (ITTC) mechanism in the simulated cloud’s IaaS. For this, the data centre authentication key is distributed among the hosts using Shamir Secret Sharing algorithm.

Manasa H.B et al., [5] propose that the energy aware resource allocation of data center resources to client applications in a way that improves energy efficiency of the data center, while delivering the negotiated Quality of Service (QoS). They conduct a survey of research in energy efficient computing and propose: architectural principles for energy efficient management of Clouds; energy efficient resource allocation policies and scheduling algorithms considering QoS expectations and power usage characteristics of the devices.

LiangTeh Lee et al., [6] propose a dynamic resource management with energy saving mechanism, which presents a method of dynamic voltage scaling for dynamic adjustment of resources by inspecting CPU utilization in the cloud computing environment. The voltage of the idle or light loaded computer can be reduced and heavy loaded works can be migrated to those machines with lighter loading for achieving the purpose of energy saving. The experimental results show that energy consumption can be saved significantly by applying the proposed mechanism.
Weiwei Chen et al., [7] propose WorkflowSim: the scientific workflow studies, which extend the existing CloudSim simulator by providing a higher layer of workflow management. It assists researchers to evaluate their workflow optimization techniques with better accuracy and wider support than existing solutions.

Bhathiya Wickremasinghe et al., [8] propose Cloud Analyst, extended from CloudSim. It was developed to simulate large scale Cloud applications with the purpose of studying the behaviour of applications under various deployment configurations. Cloud Analyst helps developers with insights in how to distribute applications among Cloud infrastructures and value added services such as optimization of applications performance and providers incoming with the use of Service Brokers.

Gaetano F. Anastasi et al., [9] propose SmartFed, a simulator for cloud federations that is able to model the richness typical of an environment with multiple cloud providers. They show the capability of SmartFed by simulating a sample mapping process for assigning applications to providers.

K. Thyagarajan et al., [10] propose a model for an optimal approach in job scheduling in cloud computing for effective gain of the profits to the cloud computing service provider.

Kavita Bhatt et al., [11] propose that there is a concept Particle Swarm Optimization (PSO) based on swarm intelligence and is a good algorithm to find the near optimum. The algorithm is good for searching so as for load balancing and even for workflow scheduling. This paper provides the run of PSO on CloudSim with comparison analysis from different simulators.

Foued Jrad et al., [12] propose that Intercloud computing allows users to easily migrate their application workloads across Clouds regardless of the underlying Cloud provider platform. A very promising use case of Intercloud computing is Cloud services brokerage. In this paper, they propose a generic architecture for a Cloud service broker operating in an Intercloud environment by using the latest Cloud standards. The broker aims to find the most suitable Cloud provider while satisfying the users’ service requirements in terms of functional and non-functional Service Level Agreement parameters.

Vincent C. Emeakaroha et al., [13] propose that service provisioning in Clouds is based on Service Level Agreements (SLAs). On the one hand SLA violation should be prevented to avoid costly penalties and on the other hand providers have to efficiently utilize resources to minimize cost for the service provisioning. Thus, scheduling strategies considering multiple SLA parameters and efficient allocation of resources are necessary. Scheduling and deploying service requests considering multiple SLA parameters such as amount of CPU required, network bandwidth, memory and storage are still open research challenges. In this paper, they present a novel scheduling heuristic considering multiple SLA parameters for deploying applications in Clouds.

Soumya Ray et al., [14] propose that Load balancing is a methodology to distribute workload across multiple computers to achieve optimal resource utilization, maximize throughput, minimum response time, and avoid overload. This paper presents a review of a few load balancing algorithms in cloud computing. The objective of this paper is to identify qualitative components for simulation in cloud environment and then based on these components, execution analysis of load balancing algorithms are also presented.

Jasmin James et al., [15] propose a new VM load balancing algorithm i.e. ‘Weighted Active Monitoring Load Balancing Algorithm’ and implemented for an IaaS (Infrastructure as a Service) framework in Simulated cloud computing environment, in order to achieve better performance parameters such as response time and Data processing time.

Linan Zhu et al., [16] propose a new business calculation mode for resources scheduling. The key research is the process of work scheduling and resource allocation problems based on ant colony algorithm. Using CloudSim simulation environment and simulation experiments, the results show that the algorithm has better scheduling performance and load balance than general algorithm.
Ghalem Belalem et al.,[17] propose that in cloud computing, the availability and performance of services are two important aspects to be lifted, because users require a certain level of quality service in terms of timeliness of their duties in a lower cost. To ensure the quality of service, they extend and enrich the simulator CloudSim by auction algorithms inherited from GridSim simulator

M. S. Kavith et al., [18] present a SaaS (Software as a Service) in Pcloud. Pcloud is a distributed system which contains a number of cooperated nodes. In pcloud, the programming software is distributed to a number of nodes so that queries can be executed in parallel manner. In this distributed system, communication overhead between the cooperated nodes is very low. To minimize the effect of node failures, PCloud includes data replication. The Load Balancing can also be achieved by loosely coupled and nearest neighbour property of clouds.

Helmut Hlavacs [19] propose that distributed dynamic VM consolidation can be an effective strategy to improve energy efficiency in cloud environments. The goal is to consolidate virtual machines dynamically in a way that optimizes the energy performance trade off online. In this paper they concentrate on the VM selection task and propose a Fuzzy QLearning (FQL) technique so as to make optimal decisions to select virtual machines for migration. They validate their approach with the CloudSim toolkit using real world PlanetLab workload.

Ts’epoMofolo et al., [20] propose that in cloud computing environments the cost of VM migration requires thorough consideration. Each VM migration may result in SLA violation; hence it is essential to minimize the number of migrations to the extent possible. Failure to do so will result in performance degradation and the cloud provider will have to incur the cost in monetary terms. In this work, they devise an algorithm that will keep the migration time minimum as well as minimizing the number of migrations. This will play a major role in avoiding the performance degradation encountered by a migrating VM.

Tejprakash Singh [21] propose that to access the cloud services users require license, a legal authorization through which a cloud service provider authorized the user to use its resources on pay per basis and resources are provided on demand by the service provider. But due to lack of license management in cloud computing it is a challenge to provide availability of resources to the user. This paper presents an algorithm to provide license to the user. This paper provides authenticated access to the resources through remote license server. This license management architecture can be deployed in any distributed environment.

Reeea S L [22] propose that at present there is no authorization recycling approach in cloud computing. The aim of the work is to develop an authorization recycling approach using CSAR (Cooperative Secondary Authorization Recycling) in Cloud computing Systems. By this, each application server recycles previously received authorizations and shares them with other application servers to mask authorization server failures and network delays. With role based access control, access decisions are based on the roles that individual users have as part of an organization. The process of defining roles should be based on a thorough analysis of how an organization operates and should include input from a wide spectrum of users in an organization. This is performed by using a simulation tool known as CloudSim.

Palvinder Singh Mann et al.,[23] presents a novel algorithm, which is based on an analytical approach to mitigate Distributed Denial of Service (DDoS) attacks on Cloud, and the simulation results based on CloudSim proves it as an effective measure to counterpart the DDoS attacks.

VI. CONCLUSION AND SCOPE FOR FUTURE WORK

In this paper, a broad survey on application of CloudSim toolkit in various research areas of cloud is carried out and also there is a scope for research in reliability in which no works have been done using CloudSim toolkit to the best of our knowledge. Cloud is a high complex system in nature which may lead to large number of failures in distributed data centers. Fault tolerance serve as an effective means to address reliability concerns. Fault tolerance means that system should continue to operate under fault presence. New Fault tolerance approaches for cloud is needed since
traditional fault tolerance approaches are less effective because of the abstraction layers of cloud computing. Fault tolerance that uses the Virtualization Technology (VT) can increase the reliability of applications but VM migration and consolidations are difficult to achieve. Hence we like to design a reliable cloud system with ease of VM migration and consolidation.

In future, we will achieve reliability by means of fault tolerance using CloudSim toolkit by improving the toolkit.

REFERENCES