A Study on the fundamental properties, features and usage of cloud simulators

Preethi Josephina Mudialba
Electrical and Electronic Engineering
Nanyang Technological University, Singapore
preethi5@e.ntu.edu.sg

Abstract— **Cloud computing is one of the key innovations in this era of technology. Fuelling us into a century of massively distributed cloud applications and services, cloud applications can be accessed with the greatest of ease simply via a web browser. As we advance further into this world of cloud computing, it is highly essential to develop methods to gauge the performance level of the cloud environment. Currently, modelling and simulation technology are the main methods in use to carry out the evaluations. Numerous cloud simulators like CloudSim, techcloud, Greencloud, cloudnet and simizer have specifically been designed for analyzing and examining the performance of the various cloud computing environments. The main objective of this project was to study a few of the existing cloud simulators and to learn about their fundamental properties, features and usage, in order to efficiently use the existing system to experiment and test out applications before their deployment in a real time environment.**

Keywords—cloud computing; cloud simulators; Cloud SIM; Simizer; Green Cloud; Teach Cloud;Cloudnet

I. INTRODUCTION (*CLOUD COMPUTING*)

A blend of the swift pace of technological advancement and the current economic situation has steered us into the era of Cloud Computing. Just as with historic processors, cloud computing is revolutionizing the virtual world as it redefines the way in which computer services are used. Cloud Computing facilitates businesses to share resources efficiently via numerous devices, thereby crafting new opportunities in terms of speed and efficiency [1].The current day's economy is in dire need of change, a fundamental change that can and that will rewrite IT structure costs. Studies reveal that 70% of our IT finances is used in maintaining the traditional IT Systems, while a mere 30% is utilized in enhancing productivity, efficiency and innovation. The need of the hour is to develop superior models which provide innovative services, thus facilitating the increase of IT capabilities and efficiency while simultaneously adding value to the organization [2].

II. CLOUD SIMULATORS

The need for simulators

As we accelerate into a new age of vastly distributed cloud applications and services , which can be accessed via computing devices, with new and improved mobility support ,cloud computing is emerging as a crucial element in the world of computing. Not only is it challenging to get access to realistic test-beds over which to carry out research experiments for testing simulation of distributed soft real-time applications, but also over general-purpose computing platforms and networks due to the lack of proper tools.

Cloud computing is certainly the principal technology involved in the provisioning of dependable, secure, error-tolerant, and scalable features. To ensure the delivery of such services in newly developed cloud systems, prompt, repeatable and easy to deal with procedures are necessary for evaluating the performance of new cloud applications prior to the final deployment of the cloud system

Cloud Simulators the back bone of cloud computing

The use of real test beds for evaluation purposes restricts the range of the experiment to that of the test bed, thereby rendering the reproduction of outcomes an awfully tedious procedure. Alternative methodologies for evaluation and testing purposes will prove to be vital in expanding the horizons of cloud computing.[10]An alternative to the use of real test beds is the use of simulation tools which permit the evaluation of applications, in a repeatable and controllable manner, prior to their development. The use of simulators is especially useful in carrying out cost reduction as the use of the cloud services incurs payments in real currency. Overall, the use of simulators allows cloud customers to experiment and evaluate applications in a governable environment free of cost, and to rectify performance and quality issues before implementing the applications on a real Cloud system. Simulation environments help optimize the use of resources and improve profits, by allowing providers to gauge the various kind of resource leasing scenarios subject to different loads and pricing. Thus with the introduction of Cloud simulators, cloud users no longer need to resort to theoretical evaluations or trial & error methods that result in poor service performance and income generation. [12]

III. STUDY ON CLOUD SIMULATORS

The following section consists of a brief description of a few of the existing cloud computing simulators

CloudSIM

CloudSIM is a novel, all-inclusive, and scalable platform, which serves as a framework for systems as well as behavior modelling of various cloud components, essentially data centers, virtual machines and resource provisioning policies. It implements standard application provisioning techniques which can be easily scaled. Currently, cloudSIM sustains modelling and simulation of single as well as inter-networked clouds.[3]

Advantages: time effectiveness, flexibility and applicability

Novel features:

- Allows simulation and modelling of both small and large scale environments. Serves as an independent, self-contained framework for simulation of clouds services
- Permits simulation of network connections among the system elements. Supports simulation of federated Cloud environment
- Virtualization engine, which facilitates the creation and management of virtualized services on a data centre node.
- Easy to switch between space-shared and time-shared allocation of processing cores to virtualized services.[3]

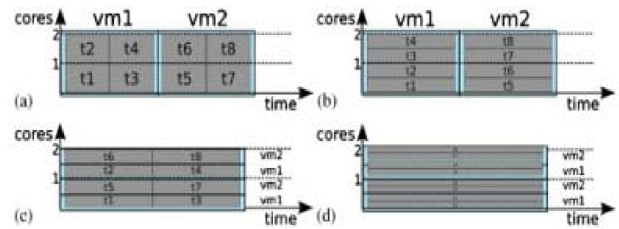
CloudSIM Architecture-The simulation layer of the cloudSIM architecture facilitates modelling and simulation of virtualized Cloud-based data center environments and deals with the primary functions of cloud computing such as allocation of hosts to VMs, managing application execution, and monitoring dynamic system state .

The User Code, which forms the topmost layer of the cloudSIM architecture, contains information regarding the basic entities for hosts, VMs, number of users and their application types, and broker scheduling policies. Using these entities a developer can (i) Generate a mix of workload request distributions, application configurations (ii) Model Cloud availability scenarios and perform robust tests based on the custom configurations (iii) Implement custom application provisioning techniques for clouds and their federation.[4] A Datacenter component in the simulator, in-charge of dealing with service requests models the core hardware framework services associated with the cloud. Service requests are basically application elements that are sandboxed within VMs. The VMs are allocated a part of the Datacenter's host components processing power, where VM processing, refers to the operations related to VM life cycle.[3]

Modelling the VM Allocation -CloudSIM facilitates simulation of diverse provisioning policies under variable levels of performance isolation by sustaining VM provisioning at 2 levels: a) Host level – stipulates the percentage of the overall processing power of each core that will be allocated per VM b) VM level- the VM allocates a fixed percentage of the processing power to the different application services hosted within its execution engine. CloudSIM implements the time-shared and space-shared resource allocation policies at the host as well as VM level.

Modelling the Cloud Market-The Cloud market follows a 2 layer design. All features related to the IaaS models such as the cost per unit of memory, storage, and used bandwidth form the first layer of the market. In this layer cost for usage of memory and storage is incurred when customers create and instantiate VMs ,and the costs for network usage is incurred only when a data transfer event actually happens. The second layer of the cloud market is based on the cost scheme of the SaaS model. Here, the cost incurred is in portion to the task units that are served by the application services. This behavior can be changed by CloudSim users.[3]

Fig.1. The effect of different provisioning policies on task unit execution



Modelling the Network Behaviour -As latency messages can impact the quality of the overall service, modelling of network topologies to link simulated Cloud computing entities is absolutely necessary. An m×n size latency matrix is maintained at all times for all CloudSim entities that are active in the simulation context. An element eij in the matrix represents the delay incurred when a message is transferred from entity i to entity j over the network. It should be noted that CloudSim is an event-based simulator, communication thus occurs with the exchange of events. CloudSIM's event management engine makes use of inter-entity network latency information for inducing delays in transmitting messages to entities. This simple technique of simulating network latencies provides a realistic approach to model practical networking architecture for a simulation environment.

Modelling a Federation of Clouds-To federate or inter-network multiple clouds a coordinator entity is necessary. The coordinator will be in charge of handling the two fundamental aspects of federation & internetworking namely communication and monitoring. The data center deals with the communication, while the cloud coordinator monitors the internal state of a data centre entity[3]

Modelling Dynamic Workloads-As scalability is one of the key features of cloud computing, it is essential for the simulation environment to be able to sustain modelling of dynamic workload patterns. Thus to cater to dynamic behaviors within CloudSIM, an entity, called the Utilization Model was implemented within CloudSIM.

Experiments and evaluation

CloudSIM : scalability and overhead evaluation

The results of an experiments carried out on a machine that had 2 Intel Xeon Quad-core 2.27 GHz and 16 GB of RAM memory is depicted in the figure below

Fig.2. CloudSim evaluation: (a) overhead and (b) memory consumption.[3]

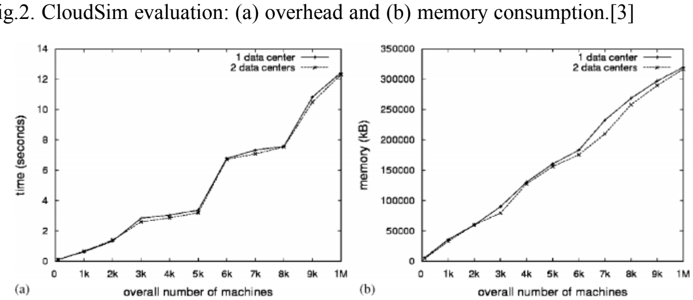
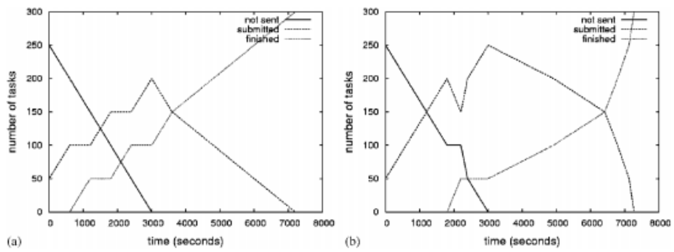
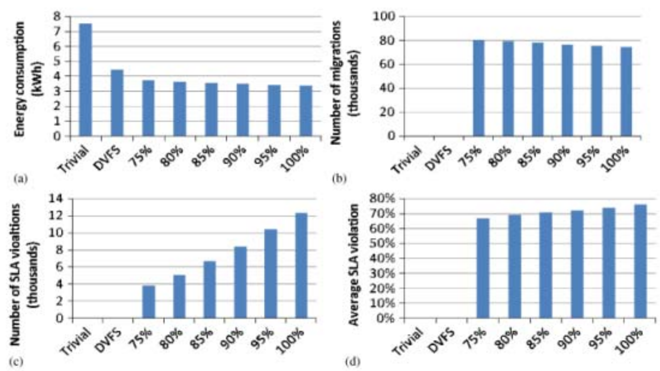


Fig.3. Simulation of scheduling policies: (a) space-shared and (b) time-shared.[3]



Energy-conscious management of data center

Fig.4. Experimental results: (a) total energy consumption by the system; (b) number of VM migrations; (c) number of SLA violations; and (d) average SLA violation.[3]



SIMIZER

Developed to facilitate quick, easy and efficient simulation of distributed computing frameworks, Simizer is an event-based simulator that is used to carry out evaluation and testing of load balancing and data synchronization strategies, prior to their deployment on a cloud based distributed computing system. The result generated from the simulations is comprehensive and accurate and it helps in evaluating the various aspects of the simulated protocols such as scalability and speed. Simizer introduces to the world of cloud computing an intuitive approach to describe workload and infrastructure characteristics, thereby allowing users to concentrate only on the protocol which needs to be evaluated.[5]

Events layer -The event management layers provide two main features:

Event management : In Simizer, a simulated system is made up of numerous entities know as Event Producers (Event Producer class) and each event producer generates a set of specific events. Generally an event is defined by a data, a target object and a virtual time-stamp. An event can be signalled by calling a specific method on the target object and by passing the stored data and time-stamp as arguments. Every time an event is produced, it is placed in a channel (class Channel), and the channel is then linked to a dispatcher. The dispatcher then conducts timely polls on the channel for any incoming events. The simulation is terminated only when no more events are found in the Channel, or when a pre-specified duration of time has lapsed.

Random number generation: As users usually depend on certain typical random distributions. Simizer offers a set of classes for producing discrete values from different random discrete distributions.

Architecture Layer-This layer generally features various entities, that are simulated. For the case of simizer these entities are usually the Virtual Machines executing a set of communicating applications.

Simulator usage -Simizer can work in 3 different modes:

Trace generation mode: This mode enables users to generate a request sequence based on a set of random distributions specified in a configuration file and the result along with the information regarding the time when each request is sent to the servers is saved in a CSV file. Traces can be updated manually, and aid in testing specific requests sequence that are likely to cause trouble to the system. By means of this file, a simulation can be launched by specifying the number of servers the system requires and by selecting a request distribution strategy.

Trace replay mode: in this mode Simizer sends the requests as stated in the CSV file.

Generative mode: in this mode the control of the behaviour of the entities is decided by a set of configuration files containing the characteristics of the simulated system [6]

Green Cloud

With the growth of data centers and parallel computing paradigms cloud computing is gradually becoming more widespread. The operation of geographically extensively distributed data centers involve a considerable amount of energy, and currently this amounts to approximately 10% of the overall data center operational expenses (OPEX)).The high power consumption also results in a release of a large amount of heat , necessitating the use of a cooling system which incurs a cost of $2 to $5 million per year for classical data centers. Failure to maintain the temperature within the operational range will reduce the hardware reliability and may result in violation of the Service Level Agreement.

At first the power saving solutions focused on improving the power efficiency of data centre hardware Components. These solution usually try to: (a) accumulate the workload in the least set of computing resources possible and (b) maximize the quantity of resource that can be put into sleep mode. Almost 30% of the overall energy is spent on the communication links, switching and aggregation elements

The energy consumed can be decreased by reducing the communication speed and by cutting operational frequency along with the input voltage for the trans receivers and switching elements, but this may result in a bottleneck, consequentially affecting the overall performance as such reducing the need to be based on the demands of user applications Green Cloud, developed as an extension of a packet-level network simulator, serves a simulation environment for advanced energy-aware studies of cloud computing data centres in realistic setups. It particularly concentrates on accurately capturing the communication patterns of the current and future architecture of the data centre.[7]

Data centre architecture

Two-tier data center architecture – Depending on the type of switches used in the access network, the two-tier data center architecture can sustain up to 5500 nodes and this was enough to meet the needs of the initial data centers which had a smaller number of computing servers [8].

The maximum network bandwidth allotted to each computing server is determined by the number of core switches and capacity of the core links.

Three-tier data center architecture as the name suggest the three-tier data center architecture has 3 layers, namely access, aggregation, and core layers.

Three-tier high-speed data center architecture optimizes the number of nodes, core capacity, and aggregation networks that form a bottleneck, thus limiting the maximum number of nodes in a data center or a per-node bandwidth

Simulation of energy-efficient data centre

Energy efficiency-The energy efficiency of a data center is measured in terms of the performance delivered per watt and is quantified by two metric units (a) Power Usage Effectiveness (PUE) and (b) Data Center Infrastructure Efficiency (DCiE) [7]

Structure of the simulator-The Green Cloud provides its users, a comprehensive model of the energy consumed by the elements of the data centre as well as a thorough analysis of the workload distributions.

Green cloud also provides the finest-grain control by particularly focusing on packet-level simulations of communications in the data centre infrastructure. [7]

Servers (S) are in charge of task execution .The switches implement single core nodes that have a fixed processing power limit, associated size of memory resources, and contain task scheduling mechanisms .

Switches and Links constitute the interconnection fabric that deal with the delivery of workload to the computing servers. Transmission rate and the link distance are the two factors that govern the cost and quantity of the energy consumed by trans receivers, thus the quality of signal transmission is determined by a trade-off between these two factors.

Workloads are objects designed for modelling of different cloud user services. Workloads are usually modelled as a sequence of jobs that can be split into tasks which are dependent or independent. In general, due to the nature of grid computing applications, the number of existing jobs usually tends to be greater than the computing resources available. The three types of jobs defined are: Computationally Intensive Workloads (CIWs) model, Data-Intensive Workloads (DIWs), Balanced Workloads (BWs) [7]

Performance evaluation- Presented below is a case study of simulations of an energy-aware data center for two-tier (2T), three-tier (3T), and three-tier high-speed (3Ths) architecture. Table 1 provides a summary of the setup parameters. Unlike the 3T and 3Ths architecture, a 2T data center does not include aggregation switches.

TABLE 1. Simulation setup parameters [7]

	Parameter	Data center architectures		
		Two-tier	Three-Tier	Three-tier high-speed
Topologies	Core nodes (C_1)	16	8	2
	Aggregation nodes (C_2)	–	16	4
	Access switches (C_3)	512	512	512
	Servers (S)	1536	1536	1536
	Link (C_1–C_2)	10 GE	10 GE	100 GE
	Link (C_2–C_3)	1 GE	1 GE	10 GE
	Link (C_3–S)	1 GE	1 GE	1 GE
	Link propagation delay	10 ns		
Data Center	Data center average load	30%		
	Task generation time	Exponentially distributed		
	Task size	Exponentially distributed		
	Simulation time	60 minutes		

Fig.5. Server workload distribution with a "green" scheduler [7]

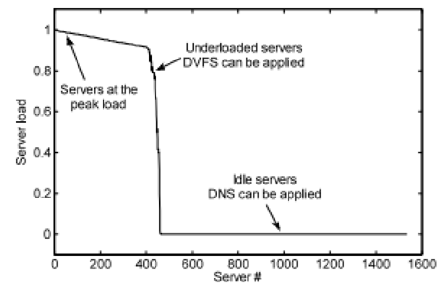


TABLE 2. The distribution of data center power consumption[7]

Parameter	Power consumption (kW h)		
	Two-tier (2T)	Three-Tier (3T)	Three-tier high-speed (3Ths)
Data center	477.8	503.4	508.6
Servers	351	351	351
Switches	126.8	152.4	157.6
Core (C_1)	51.2	25.6	56.8
Aggregation (C_2)	–	51.2	25.2
Access (C_3)	75.6	75.6	75.6

CloudNetSim

CloudNetSim is a project aiming to provide a simulation platform to assist experimentation with resource management and scheduling in cloud computing. At a glance, its main features comprise of: packet-level simulation of end-to-end network communications between clients and servers distributed throughout a cloud infrastructure; simulation of computing resources including but not limited to CPU scheduling both at the hypervisor and at the guest OS levels; support for virtual machine (VM) deployment strategies; modularity and extensibility, with the possibility to introduce additional scheduling policies, VM deployment strategies and application models as needed. [9]

One of the primary goals of the CloudNetSim project is to integrate within a single simulation platform the major factors contributing to end-to-end latency of low-latency cloud applications, namely networking, computing and disk access, including overheads due to virtualization .It was opted to implement the computing part of the simulation on top of OMNEST, due to its relative maturity, modular design and extensibility.

Use -Research and development in the sector of resource management and scheduling for real-time cloud computing applications. The presented simulation models are very important to simulate the impact on performance of sharing physical computing resources within the infrastructure, however, CloudNetSim may also be useful for simulation of soft real-time distributed embedded systems.[9]

TeachCloud

Educational organizations across the globe are progressively introducing cloud computing in their syllabuses. Cloud computing offers numerous services such as distributed systems, virtualization, storage, networking, security, management and automation, etc. This complexity serves as a major barrier to students attempting to gain a thorough understanding cloud computing. Although the students can experiment on a real cloud system, the complexity and frangibility of a cloud system renders such an option usage highly risky, thus creating a need for different options .A solution to this problem was offered by TeachCloud, a full cloud-simulator designed to be a learning tool , an environment where students can conduct experiments in a risk-free manner.[11]Initially, CloudSim was considered as a suitable environment to be used for teaching cloud computing. But, CloudSim portrays numerous limitations:

- Lacks a (GUI), a feature which will drastically improve the learning process of the students
- As clouSIM was developed on top of a grid computing framework, it has numerous restrictions in terms of the infrastructures it can simulate.
- Lacks BPM and SLA components.

In the light of these limitations, Teach Cloud was designed to serve as an educational toolkit, used by students to gain a comprehensive understating of cloud computing in a risk free manner. TeachCloud, extends the basic framework of CloudSIM and offers several new features on top of it such as: GUI, Rain cloud workload generator, new modules related to SLA and BPM, new cloud network models VL2, BCube, Portland, and DCell ,which represent the topologies that exist in real cloud environments, an action module which allows users to reconfigure the cloud system and study the impact of the changes on the overall system performance.

Workload generation and the Rain workload generator

The Rain workload generator, is an open source generator framework which presents accurate and comprehensive CC workload characteristics. TeachCloud achieves such advanced workload capabilities by integrating CloudSim and the Rain generator.

The main workload characteristics offered by Rain are:

1) Variations in workload intensity and the operations mix performed by the system components

2) Variations in data access patterns and frequency, otherwise known as data hot spots. Prediction of the data hot spots allows for easy access to them, leading to a decrease in access time and overheads.

Network topologies for TeachCloud

TeachCloud integrates the DCell, VL2, Portland and BCube network topologies . [11]

Conclusion

As Cloud applications requirements are diverse by nature evaluating the performance for each application is very challenging. The use of real test beds limit the scale of the test and make the reproduction of results highly difficult. An alternative is the use of simulators which offer repeatable and controllable environment free of cost. Cloud simulators allow users to evaluate different kinds of resource leasing situations under varying load and pricing distributions.

Acknowledgment

I would like to express my gratitude to everyone who helped me through this study.I thank the Almighty, NTU & ISEP, my family & friends for their constant encouragement .

References

[1] "Introduction to Cloud Computing." [Online]. Available at: http://www.priv.gc.ca/resource/fs-fi/02_05_d_51_cc_e.pdf.

[2] R. Calheiros, R. Ranjan, C. De Rose, R. Buyya, Cloudsim: a novel framework for modelling and simulation of cloud computing infrastructures and services

[3] R.N. Calheiros, R. Ranjan, A. Beloglazov, C.A.F. De Rose, R. Buyya, Cloudsim: a toolkit for modelling and simulation of cloud computing environments and evaluation of resource provisioning algorithms, Software: Practice and Experience 41 (1) (2011) 23–50.

[4] Howell F, Mcnab R. SimJava: A discrete event simulation library for java. Proceedings of the First International Conference on Web-based Modeling and Simulation, San Diego, U.S.A., 1998.

[5] Simizer: evaluating consistency trade offs through simulation
Sylvain Lefebvre,Sathiya Prabhu Kumar, Raja Chiky
LISITE - ISEP, 28 rue Notre Dame des Champs, Paris, France.
firstname.lastname@isep.fr

[6] Simizer: Cloud application simulator
Sylvain Lefebvre,Sathiya Prabhu Kumar, Raja Chiky
LISITE - ISEP, 28 rue Notre Dame des Champs, Paris, France.
firstname.lastname@isep.fr

[7] D. Kliazovich, P. Bouvry, Y. Audzevich, S. Khan, Greencloud: a packet-level simulator of energy-aware cloud computing data centers, in: Proceedings of the Global Telecommunications Conference (GLOBECOM 2010), IEEE, 2010, pp. 1–5

[8] Cisco Systems (2008) Cisco Data Center Infrastructure 2.5 Design Guide, Inc., May.

[9] Tommaso Cucinotta, Aram Santogidis ,Bell Laboratories, Alcatel-Lucent, Dublin ,CloudNetSim - Simulation of Real-Time Cloud Computing Applications

[10] G. Sakellari, G. Loukas, A survey of mathematical models, simulation approaches and testbeds used for research in cloud computing, Simulat. Modell. Pract. Theory (2013),

[11] Y. Jararweh, Z. Alshara, M. Jarrah, M. Kharbutli, M. Alsaleh, Teachcloud: a cloud computing educational toolkit, in: Proceedings of the 1st International IBM Cloud Academy Conference (ICA CON 2012), IBM, Research Triangle Park, NC, USA, 2012

[12] Sakellari, G., & Loukas, G. (n.d.). A survey of mathematical models, simulation approaches and testbeds used for research in cloud computing.http://dx.doi.org/10.1016/j.simpat.2013.04.002