Simulation Modelling Practice and Theory 49 (2014) 180–192

Contents lists available at ScienceDirect

Simulation Modelling Practice and Theory

journal homepage: www.elsevier.com/locate/simpat

CloudExp: A comprehensive cloud computing experimental framework

Yaser Jararweh [a,*], Moath Jarrah [b], Mazen kharbutli [b], Zakarea Alshara [a], Mohammed Noraden Alsaleh [c], Mahmoud Al-Ayyoub [a]

[a] Department of Computer Science, Jordan University of Science and Technology, 22110, Jordan
[b] Department of Computer Engineering, Jordan University of Science and Technology, 22110, Jordan
[c] University of North Carolina Charlotte, Charlotte, NC, USA

ARTICLE INFO

Article history:
Received 5 February 2014
Received in revised form 4 August 2014
Accepted 10 September 2014
Available online 15 October 2014

Keywords:
Cloud computing
Cloud computing modeling and simulation
CloudSim
Network topologies
MapReduce
SLA management
Rain workload generator

ABSTRACT

Cloud computing is an emerging and fast-growing computing paradigm that has gained great interest from both industry and academia. Consequently, many researchers are actively involved in cloud computing research projects. One major challenge facing cloud computing researchers is the lack of a comprehensive cloud computing experimental tool to use in their studies. This paper introduces *CloudExp*, a modeling and simulation environment for cloud computing. *CloudExp* can be used to evaluate a wide spectrum of cloud components such as processing elements, data centers, storage, networking, Service Level Agreement (SLA) constraints, web-based applications, Service Oriented Architecture (SOA), virtualization, management and automation, and Business Process Management (BPM). Moreover, *CloudExp* introduces the Rain workload generator which emulates real workloads in cloud environments. Also, MapReduce processing model is integrated in *CloudExp* in order to handle the processing of big data problems.

1. Introduction

Cloud computing is an emerging computing paradigm that is continuously evolving and spreading. Many experts believe it will become the dominant IT service delivery model by the end of the decade [1]. Cloud computing is built on a wide range of different computing technologies such as high-performance computing, grid and utility computing, distributed systems, virtualization, storage, networking, security, management, automation, Service-Oriented Architecture (SOA), etc. Moreover, other concepts like Business Process Management (BPM), Service-Level Agreement (SLA), Quality of Service (QoS). This complexity presents a major challenge for researchers to conduct comprehensive cloud computing-related experiments for two main reasons: First, conducting experiments on real systems is expensive and the criticality and frangibility of the system poses many limitations and risks. Such experiments may adversely affect the system's availability, reliability, and security. Second, while simulation tools would mitigate the problems that arise from using a real system, there are no comprehensive cloud computing experimental tools that cover the wide spectrum of cloud computing components. Unfortunately, current

* Corresponding author.
E-mail addresses: yijararweh@just.edu.jo (Y. Jararweh), mjarrah@just.edu.jo (M. Jarrah), kharbutli@just.edu.jo (M. kharbutli), maalshbool@just.edu.jo (Z. Alshara), malsaleh@uncc.edu (M.N. Alsaleh).

http://dx.doi.org/10.1016/j.simpat.2014.09.003
1569-190X/

cloud computing research tools only cover a subset of the different components listed above, which make them limited and not applicable when considering new cloud computing technologies.

One of the earliest and most popular cloud computing simulators is CloudSim which an environment developed at the University of Melbourne [14]. CloudSim is a very promising tool for conducting cloud computing experiments. However, CloudSim has several limitations and shortcomings: First, it is built on top of a grid computing environment which puts limitations on the infrastructures that can be simulated. This is a major limitation since current cloud systems are built using a wide range of different hardware configurations. Second, it only includes a basic and simplified network model with limited workload traffic generator. Third, it does not provide efficient modeling for embarrassingly parallel data problems. Fourth, it lacks several important cloud computing components such as BPM and SLA. Fifth, it lacks a Graphical User Interface (GUI) which can be useful and easier for researchers to conduct their experiments.

As a result of the lack of a comprehensive cloud computing experimental tool and due to the limitations in CloudSim, we set out to develop *CloudExp* as a comprehensive and effective cloud computing experimental framework. CloudExp capabilities cover most of cloud computing related technologies such as big data management and mobile cloud computing, etc. CloudExp uses CloudSim as the base design platform and introduces many new enhancements and extensions on top of it. These enhancements and extensions include:

- Integrating of the Rain cloud workload generator from Berkeley into CloudSim simulator [16].
- Integrating of a MapReduce framework into CloudSim to handle embarrassingly parallel data processing paradigms and Big Data problems.
- Adding new modules related to SLA and BPM.
- Adding new cloud computing network models (VL2, BCube, Portland, and DCell) that represent actual topologies in real cloud environments.
- Introducing of the Mobile Cloud Computing (MCC) simulation framework.
- Introducing a monitoring outlet for most of the cloud system components.
- Adding an action module to allow researchers to reconfigure a cloud system and study the overall system behavior and performance.
- Adding a GUI for the simulator to make it more user-friendly.

This paper presents CloudExp and discusses its different components in detail. The rest of the paper is organized as follows: Section 2 provides a comprehensive survey in the field of cloud computing simulation and modeling. CloudSim, which is the base for CloudExp, is introduced in sub-Section 2.1. CloudExp is fully detailed in Section 3. A use case scenario for CloudExp is presented in Section 4 where different simulation results are presented. Finally, Section 5 concludes the paper.

2. Related work

Simulations tools are essential for carrying out research experiments in cloud computing. Several cloud computing simulators have been developed, each of which supports certain aspects/components of cloud computing. This section describes some of the most popular simulators in the fields of cloud computing and distributed systems.

2.1. CloudSim: The underlying framework for CloudExp

CloudSim is a cloud computing modeling and simulation tool that was developed at the University of Melbourne [14]. It aims to provide cloud computing researchers with an experimental tool to conduct cloud computing-related research. It supports the modeling and simulation of various cloud computing components, including power management, performance, data centers, computing nodes, resource provisioning, and virtual machines provisioning [12].

CloudSim is a layered design framework written in Java and was initially built on top of SimJava and GridSim [4]. Both SimJava and GridSim are discrete event simulators that are widely used to simulate parallel and distributed computing systems. However, SimJava and GridSim have several scalability limitations that led the developers of CloudSim to implement a new discrete event management framework which is the CloudSim core simulation engine [14]. The following is a brief description of the different cloud computing features and components modeled by CloudSim.

The *Data centers* represent the core infrastructure in a cloud system and include the hardware and software stack. In defining the data centers, the modeler needs to identify the number of hosts in each data center. The *Hosts* are modeled using the class *Host* which models a physical resource such as a computer or a server. A data center can have multiple hosts. *Cloud tasks* are represented by *Cloudlets* which are the cloud-based application services. In this class, computational metrics are used to model the complexity of applications. *Brokering* is performed through the use of the *DatacenterBroker* class. *Virtual machines (VM)* are modeled in the class *VirtualMachine*. *Coordination* is performed using the *CloudCoordinator* class which models the communication between different cloud coordinators and cloud brokers. In addition, it monitors the internal state of each data center that is connected to it. *Bandwidth provisioning services* are modeled using the *BWProvisioner* class. *Memory provisioning* and the policies for allocating memory to VMs are modeled using the *MemoryProvisioner* class. *Virtual machine provisioning* is performed using the *VMProvisioner* class which allocates VMs to hosts.*VM allocation policies* models