A Review of Cloud Computing Simulation Platforms and Related Environments

James Byrne1, Sergej Svorobej1, Konstantinos M. Giannoutakis2, Dimitrios Tzovaras2, P. J. Byrne1, Per-Olov Östberg3, Anna Gourinovitch1 and Theo Lynn1
1Irish Centre for Cloud Computing and Commerce, Dublin City University, Glasnevin, Dublin 9, Ireland
2Information Technologies Institute, Centre for Research and Technology Hellas, 6th km Xarilaou-Thermi, 57001 Thessaloniki, Greece
3Dept. Computing Science and HPC2N, Umeå University, SE-901 87 Umeå, Sweden

Keywords: Cloud Computing, Cloud Simulation Tools, Data Centre, Fog Computing.

Abstract: Recent years have seen an increasing trend towards the development of Discrete Event Simulation (DES) platforms to support cloud computing related decision making and research. The complexity of cloud environments is increasing with scale and heterogeneity posing a challenge for the efficient management of cloud applications and data centre resources. The increasing ubiquity of social media, mobile and cloud computing combined with the Internet of Things and emerging paradigms such as Edge and Fog Computing is exacerbating this complexity. Given the scale, complexity and commercial sensitivity of hyperscale computing environments, the opportunity for experimentation is limited and requires substantial investment of resources both in terms of time and effort. DES provides a low risk technique for providing decision support for complex hyperscale computing scenarios. In recent years, there has been a significant increase in the development and extension of tools to support DES for cloud computing resulting in a wide range of tools which vary in terms of their utility and features. Through a review and analysis of available literature, this paper provides an overview and multi-level feature analysis of 33 DES tools for cloud computing environments. This review updates and extends existing reviews to include not only autonomous simulation platforms, but also on plugins and extensions for specific cloud computing use cases. This review identifies the emergence of CloudSim as a de facto base platform for simulation research and shows a lack of tool support for distributed execution (parallel execution on distributed memory systems).

1 INTRODUCTION

The definition of cloud computing is widely accepted to be “...a model for enabling ubiquitous, 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” (Mell and Grance, 2009). While the reference architecture for cloud computing has evolved over time the essential characteristics, service models and deployment models have largely remained the same (Liu et al., 2011). The broad cross-domain applicability of cloud computing has led to the emergence of a variety of resource profiles and technological options, with a substantial degree of heterogeneity in data centre resources and service offerings (Östberg et al., 2014). Recently, this trend has also been magnified by increasing demands for dependability and real-time low latency communication, which has driven integration of telecommunications and cloud infrastructure (edge computing), as well as development and integration of applications that make increased use of the capabilities of end-user devices and appliances (fog computing). A general inability to control and process the network environment and predict and control network conditions in hyperscale computing environments has necessitated the development of discrete event simulation (DES) platforms capable of supporting complex decision making within these environments (Jiang et al., 2012), (Tian et al., 2015). IDC (2016) predict rapid and substantial increases in enterprise cloud and the Internet of Things (IOT) adoption with

Byrne, J., Svorobej, S., Giannoutakis, K., Tzovaras, D., Byrne, P., Òstberg, P.-O., Gourinovitch, A. and Lynn, T.
DOI: 10.5220/0006373006790691
In Proceedings of the 7th International Conference on Cloud Computing and Services Science (CLOSER 2017), pages 651-663
ISSN: 978-889-758-243-1
Copyright © 2017 by SCITEPRESS – Science and Technology Publications, Ltd. All rights reserved

651
at least 40% of IoT-created data being stored, processed, analysed, and acted upon close to or at the edge of the network by 2019. These trends are increasing both the range of use cases and features that DES tools are required to support. An overview of the state of the art for such DES tools is presented in this paper.

Earlier efforts for DES in this domain focused on grid computing, whereby simulation tooling support was provided for uniformly aggregating and sharing distributed heterogeneous resources for solving large-scale applications, such as in the fields of science, engineering and commerce (Sulistio et al., 2008). Various grid computing simulators have been developed (Sulistio et al., 2008) and are presented in literature, such as OptorSim (Bell et al., 2002), MONARC (Legrand and Newman, 2000), SimGrid (Legrand et al., 2003), GridSim (Buyya and Murshed, 2002) and MicroGrid (Song et al., 2000). However, these alone do not provide an environment that can be directly used by the cloud computing community (W. Zhao et al., 2012); grid computing simulators assume non-interactive fixed duration whereas cloud simulators typically aim to analyse the behaviour of data centre resources that host virtual machines in multi-tenancy scenarios over non-deterministic timeframes, with highly variable user load taken into consideration. The work presented in this paper focuses on DES tools that support Infrastructure as a Service (IAAS) cloud computing use cases and the related Edge and Fog Computing paradigms.

There are a number of potential advantages to the use and development of such DES tools to support cloud computing. Experimentation in a simulated environment is typically far less expensive economically than using a real testbed. Furthermore, such experimentation is repeatable and potentially scalable in terms of addressing the simulation of larger-scale systems. In addition, experiments can be performed in a timelier fashion, and risks with respect to stochastic compute jobs can be taken into account. However it is noted by (Sakellari and Loukas, 2013) that while simulation offers a number of advantages especially in terms of such scalability and experiment repeatability, it is still based on assumptions and simplifications that might not fully represent an actual cloud. For this reason, it still might be preferable in some circumstances to use real cloud testbeds in place of simulation or to validate results developed in simulated environments. Sakellari and Loukas (2013) provide an overview of such testbeds and software frameworks for setting up such cloud testbeds.

This paper gives an overview of current work in cloud computing simulation tool development. It categorizes and reviews DES tools for cloud computing, identifies application DES tools for cloud computing environments, and provides a multi-level feature comparison of identified simulation tools plugins and extensions. This multi-level comparison concerns a general high level comparison as well as comparing high level technical characteristics for classifying the tools.

The remainder of the paper is structured as follows: Section 2 provides and overview related research to position the contribution of this work. Section 3 introduces the tools identified in the review. Section 4 presents a multi-level feature analysis of the tools. The paper concludes with a discussion of key findings and areas for future research.

2 RELATED WORK

There are a number of existing papers that provide overviews of DES tools to support cloud computing. Zhao et al. (2012) present a summary of tools to model and simulate cloud computing systems, including both software and hardware simulators. They give a feature description for 11 tools, and provide a comparison based on the underlying platform, programming language, and whether they are software or hardware-based. Sinha and Shekhar (2015) present a high level overview of 15 cloud simulation tools, and provide a tabular comparison of these based on graphical user interface support, platform used, language used, support of TCP/IP, whether they are software or hardware-based, and their availability (software license type). As part of their work, Sakellari and Loukas (2013) provide an overview of cloud simulation software. They present an overview of eight tools, and provide a tabular comparison based on whether they support energy efficiency modelling, performance/quality of service (QoS), programming language, availability (on the web), and license type. Malhotra and Jain (2013) provide an overview of five cloud simulation tools, and compare them based on underlying platform, programming language, networking support, the type of simulator (event versus packet based), and license type. Similarly, Mohana, Saroja, and Venkatachalam (2014) provide an overview of six cloud simulation tools and compares them by underlying platform, simulator type, language, networking, and availability. Ahmed and Sabyasachi (2014) give an overview of 12 cloud simulators and
compare these based on underlying platform, availability, programming language, whether or not they provide cost modelling, if they have a GUI, if they have communication models or energy models, the simulation time and whether they model federation policies.

The work presented in this paper builds on these previous related works by extending both the breadth and depth of analysis. 33 platforms, plugins and extensions are introduced and analysed including many which have not been analysed and compared previously e.g. CactoSim (Östberg et al., 2014), DISSECT-CF (Kecskemeti et al., 2014), iFogSim (Gupta et al., 2016) and CloudEXP (Jararweh et al., 2014). For each tool, a multi-level feature analysis is provided, for high-level comparison of the frameworks.

3 PLATFORMS FOR CLOUD COMPUTING SIMULATION

Table 1 lists current identified tools that support DES for cloud computing, in alphabetical order.

<table>
<thead>
<tr>
<th>Table 1: Identified cloud computing-related DES tools.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bazaar Extension*</td>
</tr>
<tr>
<td>CACTOSim</td>
</tr>
<tr>
<td>CDOSim*</td>
</tr>
<tr>
<td>CEPSim*</td>
</tr>
<tr>
<td>Cloud2Sim*</td>
</tr>
<tr>
<td>CloudAnalyst*</td>
</tr>
<tr>
<td>CloudEXP*</td>
</tr>
<tr>
<td>CloudNetSim++</td>
</tr>
<tr>
<td>CloudReports*</td>
</tr>
<tr>
<td>CloudSched</td>
</tr>
<tr>
<td>CloudSim</td>
</tr>
<tr>
<td>CloudSimDisk*</td>
</tr>
<tr>
<td>CloudSimSDN*</td>
</tr>
<tr>
<td>CMCloudSimulator*</td>
</tr>
<tr>
<td>DartCSim*</td>
</tr>
<tr>
<td>DCSim1</td>
</tr>
<tr>
<td>DCSim2</td>
</tr>
</tbody>
</table>

*Derivatives or extensions of CloudSim
1 This refers to DCsSim by Tigue, (2012)
2 This refers to DCsSim by Chen et al., (2012)

Bazaar-Extension (Pittl et al., 2016) is a CloudSim extension for simulating resource allocations by negotiation processes. The negotiation process is realised between provider and consumer using the offer-counteroffer negotiation protocol for resource allocation, while the authors simulate different negotiation strategies. The architecture of Bazaar-Extension (which is built on CloudSim) consists of the Datacenter broker and the Negotiation Manager that handles the auctioning process for forming service level agreements.

CactoSim was developed as part of CACTOS, a European Union Framework 7 project (CACTOS Consortium, 2016). CACTOS aimed to deliver a set of integrated tools for analysing application behaviour and infrastructure performance data, mathematical models and their realization to determine the best fitting resource within a provider context, and a prediction and simulation environment for diverse application workloads (Östberg et al., 2014). To this end, CactoSim is a DES framework built on top of Palladio (Becker et al., 2009), and Simulizar (Becker et al., 2013) which was developed as part of CloudScale (Brataas et al., 2013). It is used to evaluate the effectiveness of optimization strategies for the cloud, as well as for iterative resource planning and operations decision support.

CDOSim (Fittkau et al., 2012) is a simulation framework based on CloudSim and focuses on evaluating competing cloud deployment options. It simulates response times, SLA violations and costs of various deployment options. Its purpose is to assist cloud users to find the best ratio between high performance and low costs.

The CEPSim (Higashino et al., 2016) simulator is also an extension to CloudSim that focuses on supporting cloud-based Complex Event Processing (CEP) and Stream Processing (SP) systems that related to big data technologies. CEPSim transforms user queries into directed acyclic graph representations. The modelled queries can be simulated on different deployment models including private, public, hybrid and multi-clouds.

Cloud2Sim (Kathiravelu and Veiga, 2014) is a concurrent and distributed cloud and MapReduce simulator that is built on top of CloudSim, using the distributed shared memory from Hazelcast and the in-memory key-value data grid of Infinispan. The motivation for the development of this simulator was the long execution time and limited simulation size on uniprocessor systems. It provides the functionality to execute CloudSim in parallel and thus scale up simulations.

CloudAnalyst (Wickremasinghe et al., 2010) is a Cloud simulation tool developed on the Java platform for the simulation of large-scale cloud applications with the purpose to study and analyse the behaviour of such applications under various deployment scenarios. It extends the functionality of the CloudSim toolkit through the introduction of concepts that model the Internet and Internet
application behaviour. It allows description of application workloads including information on the geographic location of users generating traffic, the location of data centres, the number of users and data centres, and the number of resources in each data centre. Provided with this information, metrics such as the response and processing time of requests are generated. The main features of CloudAnalyst are: the easy to use Graphical User Interface, the ability to define a simulation with a high degree of configurability and flexibility, the repeatability of experiments, its graphical output, the use of consolidated technology, and ease of extension.

The CloudExp framework is Java-based and again is built on top of CloudSim (Jararweh et al., 2014). CloudExp can be used to evaluate cloud components such as processing elements, data centers, storage, networking, SLA constraints, web-based applications, Service Oriented Architecture (SOA), virtualization, management and automation, and Business Process Management (BPM) components. In addition, CloudExp introduces the Rain workload generator which emulates real workloads in cloud environments.

CloudNetSim++ (Malik et al., 2014) is designed to allow researchers to incorporate their custom protocols and applications for analysis under realistic data centre architectures with various network traffic patterns. It provides a framework that allows users to define SLA policies, scheduling algorithms and models for different components of data centres. The energy utilization is computed in three components: servers, communication links and data centre infrastructures (such as routers and switches). It is built on top of OMNeT++ and provides a rich GUI to simplify analysis, debugging and addition of hardware components into the simulation.

CloudReports (Teixeira Sá et al., 2014) is an extensible simulation tool for energy-aware cloud computing environments to enable researchers to model multiple complex scenarios through a GUI. It provides four layers on top of the CloudSim simulation engine: Reports manager, Simulation manager, Extensions and Core entities. The main advantage of CloudReports is its modular architecture that allows the extension of its API for experimenting with new scheduling and provisioning algorithms.

CloudSched (Tian et al., 2015) is a simulation tool for the evaluation and modelling of cloud environments and applications with a focus on comparing different resource scheduling algorithms in IaaS with regards to both computing servers and user workloads. CloudSched was introduced as a means to provide better cloud performance compared to CloudSim and CloudAnalyst. Unlike traditional scheduling algorithms that consider only one factor (such as CPU), which can cause hotspots or bottlenecks in many cases, CloudSched treats multi-dimensional resources such as CPU, memory and network bandwidth integrated for both physical machines and virtual machines for different scheduling objectives. The main CloudSched features are: its focus on infrastructure as a service (IaaS) layer, the provision of a uniform view of all resources, the lightweight design and scalability, its high extensibility and ease of use.

CloudSim (Calheiros, 2011) is an open source and extensible Java simulation platform for enabling continuous modelling, simulation, and experimentation of cloud computing and application services. CloudSim is the de facto platform of choice for open source simulation tool development; 18 of the tools analysed were derivatives or extensions of CloudSim. The CloudSim architecture follows a layered approach. At the fundamental layer, management of applications, hosts of VMs, and dynamic system states are provided. By extending the core VM provisioning functionality, the efficiency of different strategies at this layer can be studied. At the top layer, the User Code represents the basic entities for hosts, and through extending entities at this layer, one can enable the ability to generate requests using a variety of approaches and configurations, model cloud scenarios, implement custom applications and so on. In the CloudSim implementation, there are no actual entities available for simulating network entities, such as routers or switches. Instead, network latency between two components is simulated based on the information stored in a latency matrix. The event management engine of CloudSim utilizes the inter-entity network latency information for inducing delays in transmitting message to entities. This delay is expressed in simulation time units such as milliseconds. The CloudSim framework provides basic models and entities to validate and evaluate energy-conscious provisioning of techniques/algorithms.

CloudSimDisk (Louis et al., 2015) is a CloudSim extension focusing on modelling and simulating energy aware storage hardware components in cloud infrastructures. The implementation of CloudSimDisk is based on analytical models that were tested against hard disk drive manufacturer specifications. It includes HDD power models, disk array management algorithms and energy-aware data
center storage. Experimentation with CloudSimDisk shows good results in terms of validation, while the scalability of the extension allows future implementations of more complex systems.

CloudSimSDN (Son et al., 2015) is a CloudSim extension for Software Defined Networking (SDN)-enabled cloud environments. It provides a lightweight and scalable simulation environment to evaluate the network allocation capacity policies. It simulates cloud data centres, physical machines, switches, network links and virtual topologies for measuring performance metrics, and energy consumption. It also provides a GUI for simplifying the simulation configuration.

CMCloudSimulator (Alves et al., 2016) focuses on simulating applications with various deployment configurations. It incurs the cost it would require when implemented in a cloud provider according to the cost model of any service provider. It is built as a CloudSim extension and supports various cost models that can be designed using XML. With CMCloudSimulator, one can estimate the total cost of the resulting simulation and compare the results with different cloud providers, by obtaining the best price from them dynamically.

DartCSim (Li et al., 2012) is a GUI layer on top of CloudSim providing a more user-friendly interface. This allows the user to configure all the simulation data easily including the configuration of network cloudlets, network topology, and the algorithms for managing the cloud data center.

DCSim (Tighe, 2012), (Keller et al., 2013) is an extensible framework for simulating a multi-tenant, virtualized data centre with special purpose of dynamically managing hardware resources. DCSim provides an application model that can simulate the interactions and dependencies between many VMs working together to provide a single service, such as in the case of a multi-tiered web application. DCSim simulates a virtualized data centre operating an Infrastructure as a Service (IaaS) cloud. Virtual machine management operations, such as VM live migration and replication, are supported within DCSim. The resource needs of each VM in DCSim are driven dynamically by an application class component, which varies the level of resources required by the VM to simulate a real workload. DCSim reports a number of metrics in order to help determine the behaviour of the data centre during the simulation such as SLA violations, data centre utilization, active hosts, host-hours, active host utilization, number of migrations, and power consumption. A visualization tool is included with DCSim which automatically generates a set of graphs based on the simulation log files.

There is an additional simulation platform also called DCSim (Data Centre Simulator) as referred to by Chen et al., (2012). They use this to model a small-scale operating system, HDD and SSD towards achieving a multi-layer heterogeneous system simulation.

DISSECT-CF (DIScrete event baSed Energy Consumption simulator for clouds and Federations) is a simulation framework capable of simulating the internal components and processes of cloud infrastructures allowing the evaluation of energy consumption, network behaviour and the effects of cross virtual machine CPU sharing (Kecskemeti et al., 2014). In their paper, (Kecskemeti et al., 2014) introduce techniques for unifying DISSECT-CF with GroudSim, thereby providing GroudSim with the ability to model the internals of infrastructure clouds (such as energy models and more complex networking), as DISSECT-CF is more focused on the internal organization and behaviour of IaaS systems. This improves the modelling of resource usage, network usage, power consumption and data centre configurations.

EMUSIM (Callieiros et al., 2013) is a tool built on top of CloudSim that automatically extracts information from application behaviour via emulation, and uses this information to generate a corresponding simulation model. This process is performed order to better predict the service’s behaviour on cloud platforms; increased accuracy in an application behaviour model leads to higher accuracy in simulated system resource utilization estimation on cloud platforms.

GDCSim (Gupta et al., 2014) is a simulation tool for studying the energy efficiency of data centres under various data center geometries, workload characteristics, platform power management schemes, and scheduling algorithms. The main focus of GDCSim simulator is the energy efficiency analysis and its functional behaviour can be characterised by: automated processing, online analysis capability, iterative design analysis, thermal analysis capability, workload and power management and consideration of cyber-physical interdependency.

GreenCloud (Kliazovich et al., 2012) is an open-source cloud computing simulator, specifically designed for data centre simulation by implementing detailed modelling of communication aspects of the data centre. It is classified as a packet-level simulator, and, along with the workload distribution, the simulator is designed to capture details of the energy consumed by data centre components.
(servers, switches, routers, and connection links between them) as well as packet-level communication patterns in realistic setups. *GreenCloud* also allows analysis of the load distribution through the network, as well as communication with high accuracy (TCP packet level). It implements a simplistic application model without any communicating tasks or limited network model within the data centre. *GreenCloud* simulator is an extension of ns-2 simulator, which is used in computer networking. Using ns-2 as the foundation, *GreenCloud* implements a full TCP/IP protocol reference model, which allows seamless integration of a wide variety of communication protocols including IP, TCP, and UDP with the simulation.

*GroudSim* (Ostermann et al., 2011) is an event-based Java-based simulation toolkit, mainly focused on scientific applications running on combined Grid and cloud infrastructures. *GroudSim* supports modelling of cloud compute and network resources, job submissions, file transfers, as well as integration of failure, background load, and cost models.

*iCanCloud* (Núñez et al., 2012) is aimed at simulating cloud resources as provided by the Amazon Elastic Compute Cloud (EC2); although its creators claim it can be extended to simulate other environments using the provided API. Its primary aim is to predict the trade-offs between cost and performance of a given application executed in a specific hardware. *iCanCloud* is based on various platforms: OMNeT++, MPI, and C++. The *iCanCloud* architecture follows a layered approach with four layers: VMs repository, Application repository, Cloud hypervisor, and Cloud system. It provides configurations for storage systems, which include models for local storage systems, remote storage systems and parallel storage systems.

*iFogSim* (Gupta et al., 2016) is a simulator built on top of CloudSim specifically for supporting the modelling of IoT and Fog computing environments, in order to measure the impact of resource management techniques in terms of latency, network congestion, energy consumption and cost.

*MDCSim* (Lim, 2009) is a flexible and scalable simulation platform for in-depth analysis of multi-tier data centres. It implements all the important design specifics of communications, kernel level scheduling artefacts and application level interactions among the tiers of a three-tier data centre.

*MR-CloudSim* is primarily concerned with designing and implementing the *MapReduce* computing model on *CloudSim* (Jung and Kim, 2012), in order to provide an easier way to examine a *MapReduce* model in a data centre.

*NetworkCloudSim* (Garg and Buyya, 2011) is an extension of *CloudSim* that supports a scalable network model of a data centre and generalized applications such as high-performance computing (HPC), e-commerce, social networks, and web applications. *NetworkCloudSim* can simulate a cloud data centre network and applications with communicating tasks with accuracy. It provides models to support realistic, multi-tier applications that comprise several tasks that communicate with each other. In the original *CloudSim* implementation, it was assumed subtly that each VM is connected with all other VMs. The drawback of this is that it fails to model a realistic data centre environment. To tackle this issue, *NetworkCloudSim* provides three types of switches in the corresponding levels: root, aggregate and edge level. Users can design customized types of switches and their ports according to the data centre environment they want to simulate.

*SimGrid* (Casanova et al., 2008) is a simulation toolkit for the study of scheduling algorithms for distributed applications. Originally designed for simulating grid computing, it has been extended to support a variety of cloud computing use cases including multi-purpose network representation (Bobelin et al., 2012); VM abstraction (Hiroya and Lebre, 2013); live migration (Hiroya et al., 2013); virtual machine support (Hiroya et al., 2015), and storage simulation (Lebre et al., 2015).

Sotiriadis et al. (2013) present *SimIC* (Simulating the Inter-Cloud) which is a DES toolkit based on the process oriented simulation package of *SimJava* (Howell and McNab, 1998). It aims to replicate an inter-cloud facility wherein multiple clouds collaborate with each other for distributing service requests with regard to the desired setup of the simulation.

According to (Sriram, 2009) *SPECI* (Simulation Program for Elastic Cloud Infrastructures) is a simulation tool that allows exploration of aspects of scaling as well as performance properties of future data centres. *SPECI* simulates the performance and behaviour of data centres given the size and middleware design policy as an input.

*TeachCloud* is a tool designed to overcome challenges in teaching cloud computing (Jararweh et al., 2013). Based on *CloudSim*, the authors developed a GUI for the toolkit. They also integrated the MapReduce framework, and added a *rain cloud workload generator*, modules relating to SLA and BPM, cloud network models, a monitoring outlet for most of the cloud system components, and an action
model to enable students to reconfigure the system and study impact on the total system performance. UCloud (Sqalli et al., 2012) was also developed for educational purposes. Built on CloudSim, UCloud’s architecture is based on the hybrid cloud model and therefore supports both public and private clouds. It comprises two parts, the Cloud Management System and the Hybrid Cloud. WorkflowSim (Chen and Deelman, 2012) extends CloudSim through the provision of a higher layer of workflow management. This enables researchers to evaluate their workflow optimisation techniques with more accuracy and support.

4 CLOUD COMPUTING DES FEATURE MATRICES

In order to compare the identified DES platforms presented in Section 3, a multi-level approach is employed. Two feature matrices have been produced: Table 2 presents a general high level feature matrix, whereas Table 3 presents a high level technical feature matrix.

Table 2 presents the following key features for comparison of general high level aspects:

- Underlying Software Stack. Any major 3rd party dependencies that are required for software to function.
- License(s). The software license type of the simulation platform and the underlying software stack.
- Initial Publication Year. The year when the first academic publication became available describing features and usage scenarios of simulation platform.
- Lines of Code (LOC). The number of lines of code, determined by using Cloc v1.64. Comments and empty lines are not included in this calculation. Also, the authors made the best judgement to exclude any 3rd party source code that also was distributed in a bundle. For example, for all CloudSim based simulators the actual CloudSim code (usually located in src/org/cloudbus) was removed from calculations.
- Last Code Update. The identified year that the last commit of the source code was carried out.
- User Documentation Availability. The identified availability of separate documentation that explains how to install and use the relative DES platform.
- Source Code Availability. The identified availability of an online repository with the latest source code that can be downloaded and used by anyone.

- Binary availability. The availability of pre-compiled executable code.

Table 3 summarises the high level technical features as follows:

- Language(s). The major identified programming language(s) that were used in the development of the simulation platform.
- Platform Portability. The ability to use the simulation platform under multiple operation systems (e.g. MS Windows, Linux) without significant effort and performance difference.
- Distributed Architecture. The ability of software to be executed on more than one host. This category includes a single simulation run being distributed among multiple hosts as well as scaling up for load balancing if the multiple simulation runs need to be executed at the same time.
- Model Persistence Type. The identified persistence format of the experiment scenarios that the simulation platform requires in order to execute simulation runs.
- Web API Availability. The identified availability of a web-based API for controlling the simulation platform remotely.
- User Documentation Availability. The identified availability of separate documentation that explains how to install and use the relative DES platform.
- Graphical User Interface Availability. The availability of a graphical user interface that enables the graphical modelling of experiments, simulation execution and the presentation of simulation results.
- Headless Execution. The identified ability to run the simulation platform without a user interface, using only command line arguments.

5 DISCUSSION AND CONCLUSIONS

This work provides an overview of 33 cloud simulation tools through an analysis of the available literature. This analysis not only focused on autonomous simulation platforms, but also includes plugins and extensions that many researchers have
Table 2: Identified cloud computing DES platform high-level feature matrix.

<table>
<thead>
<tr>
<th>Simulation Platform</th>
<th>Underlying Stack</th>
<th>License(s)</th>
<th>Initial Publication Year</th>
<th>Lines of Code</th>
<th>Last Update Year</th>
<th>Documentation Available</th>
<th>Source Code Available</th>
<th>Binary Available</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bazaar-Extension</td>
<td>CloudSim, F(X)yz</td>
<td>Apache 2, BSD</td>
<td>2015</td>
<td>N/A</td>
<td>N/A</td>
<td>No</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>CACTOSim</td>
<td>DESMO-J, Palladio, Simulizar, EMF, Eclipse, CDO</td>
<td>GPL, Apache 2, EPL</td>
<td>2014</td>
<td>46914</td>
<td>2016</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>CDOSim</td>
<td>CloudSim, CloudMIG Xpress, Eclipse, EMF</td>
<td>EPL, Apache 2</td>
<td>2012</td>
<td>15619</td>
<td>2012</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>CEPsim</td>
<td>CloudSim</td>
<td>MIT</td>
<td>2015</td>
<td>5564</td>
<td>2015</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Cloud2Sim</td>
<td>CloudSim, Hazelcast, Infinispan</td>
<td>GPL, Apache 2</td>
<td>2014</td>
<td>2994</td>
<td>2015</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>CloudAnalyst</td>
<td>CloudSim</td>
<td>No data, Apache 2</td>
<td>2009</td>
<td>3277</td>
<td>2010</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>CloudEXP</td>
<td>CloudSim</td>
<td>No data, Apache 2</td>
<td>2014</td>
<td>N/A</td>
<td>N/A</td>
<td>No</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>CloudNetSim++</td>
<td>Inet, Omnet++</td>
<td>GNU, Academic, GPL, LGPL</td>
<td>2014</td>
<td>2276</td>
<td>2014</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>CloudReports</td>
<td>CloudSim</td>
<td>GPL 3, Apache2</td>
<td>2011</td>
<td>19274</td>
<td>2015</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>CloudSched</td>
<td>None</td>
<td>No data</td>
<td>2015</td>
<td>16681</td>
<td>2015</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>CloudSim</td>
<td>None</td>
<td>Apache 2</td>
<td>2009</td>
<td>28450</td>
<td>2016</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>CloudSimDisk</td>
<td>CloudSim</td>
<td>LGPL3, Apache 2</td>
<td>2015</td>
<td>1901</td>
<td>2015</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>CloudSimSDN</td>
<td>CloudSim</td>
<td>GPL 2, Apache 2</td>
<td>2015</td>
<td>4006</td>
<td>2015</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>CMCloudSimulator</td>
<td>CloudSim</td>
<td>No data, Apache 2</td>
<td>2016</td>
<td>566</td>
<td>2016</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>DartCSim</td>
<td>CloudSim</td>
<td>No data, Apache 2</td>
<td>2012</td>
<td>N/A</td>
<td>N/A</td>
<td>No</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>DCsim</td>
<td>None</td>
<td>GPL 3</td>
<td>2012</td>
<td>7369</td>
<td>2014</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>DCsim</td>
<td>MicroController-II</td>
<td>No data, Comm.</td>
<td>2012</td>
<td>N/A</td>
<td>N/A</td>
<td>No</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>DISSECT-CF</td>
<td>Trove, Apache commons</td>
<td>GPL 3, LGPL, Apache2</td>
<td>2015</td>
<td>9153</td>
<td>2016</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>EMUSim</td>
<td>CloudSim, Xen</td>
<td>GPL, Apache 2</td>
<td>2012</td>
<td>1369</td>
<td>2012</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>GDCSim</td>
<td>None</td>
<td>GPL 2</td>
<td>2011</td>
<td>3061</td>
<td>2001</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>GreenCloud</td>
<td>NS2</td>
<td>GPL</td>
<td>2010</td>
<td>6543</td>
<td>2016</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>GroudSim</td>
<td>SSJ, DISSECT-CF</td>
<td>GPL, Apache2, GPL 3</td>
<td>2010</td>
<td>8714</td>
<td>2010</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>iCanCloud</td>
<td>Inet, Omnet++</td>
<td>GPL 3, GNU, Academic</td>
<td>2011</td>
<td>38708</td>
<td>2015</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>iFogSim</td>
<td>CloudSim</td>
<td>No data</td>
<td>2016</td>
<td>8397</td>
<td>2016</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>MDCSim</td>
<td>CSIM</td>
<td>Commercial/Education</td>
<td>2009</td>
<td>N/A</td>
<td>N/A</td>
<td>No</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>MR-CloudSim</td>
<td>CloudSim</td>
<td>No data, Apache 2</td>
<td>2012</td>
<td>N/A</td>
<td>N/A</td>
<td>No</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>NetworkCloudSim</td>
<td>CloudSim</td>
<td>No data</td>
<td>2011</td>
<td>N/A</td>
<td>N/A</td>
<td>No</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>SimGrid</td>
<td>None</td>
<td>GPL</td>
<td>2001</td>
<td>94951</td>
<td>2016</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>SimIC</td>
<td>SimJava, jFreeChart</td>
<td>Uni. Of Ed. Acad. Non-Comm., LGPL</td>
<td>2013</td>
<td>N/A</td>
<td>N/A</td>
<td>No</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>SPECI</td>
<td>No data</td>
<td>No data</td>
<td>2009</td>
<td>N/A</td>
<td>N/A</td>
<td>No</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>TeachCloud</td>
<td>CloudSim, Rain</td>
<td>GNU, Apache 2</td>
<td>2013</td>
<td>9891</td>
<td>2014</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>Ucloud</td>
<td>CloudSim</td>
<td>No data</td>
<td>2012</td>
<td>N/A</td>
<td>N/A</td>
<td>No</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>WorkflowSim</td>
<td>CloudSim</td>
<td>LGPL3, Apache 2</td>
<td>2015</td>
<td>5269</td>
<td>2015</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
</tr>
</tbody>
</table>
Table 3: Identified cloud computing DES platform high level technical feature matrix.

<table>
<thead>
<tr>
<th>Simulation Platform</th>
<th>Language(s)</th>
<th>Platform Portability</th>
<th>Distributed Architecture</th>
<th>Persistence Type</th>
<th>Model Availability</th>
<th>Web API Availability</th>
<th>GUI Availability</th>
<th>Headless Execution</th>
<th>Result Output Format</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bazaar-Extension</td>
<td>Java</td>
<td>Yes</td>
<td>No</td>
<td>Java classes</td>
<td>No</td>
<td>Yes</td>
<td>No data</td>
<td>No data</td>
<td>Dashboard plots, F(X)yz 3D renders</td>
</tr>
<tr>
<td>CACTOSim</td>
<td>Java</td>
<td>Yes</td>
<td>No</td>
<td>Ecore</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>EDP2, CSV</td>
</tr>
<tr>
<td>CDOSim</td>
<td>Scala, Java</td>
<td>Yes</td>
<td>No</td>
<td>Java classes</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>Text</td>
</tr>
<tr>
<td>Cloud2Sim</td>
<td>Java</td>
<td>Yes</td>
<td>Yes</td>
<td>Java classes</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>Text</td>
</tr>
<tr>
<td>CloudAnalyzer</td>
<td>Java</td>
<td>Yes</td>
<td>No</td>
<td>XML</td>
<td>No</td>
<td>Yes</td>
<td>No data</td>
<td>No data</td>
<td>No data</td>
</tr>
<tr>
<td>CloudEXP</td>
<td>Java</td>
<td>Yes</td>
<td>No</td>
<td>No data</td>
<td>No</td>
<td>Yes</td>
<td>No data</td>
<td>No data</td>
<td>No data</td>
</tr>
<tr>
<td>CloudNetSim++</td>
<td>C++</td>
<td>Yes</td>
<td>No</td>
<td>NED</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>Text</td>
</tr>
<tr>
<td>CloudReports</td>
<td>Java, JS</td>
<td>Yes</td>
<td>No</td>
<td>SQLite DB</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>Javascript, text</td>
<td></td>
</tr>
<tr>
<td>CloudSched</td>
<td>Java</td>
<td>Yes</td>
<td>No</td>
<td>No data</td>
<td>No</td>
<td>Yes</td>
<td>No data</td>
<td>XLS, Text</td>
<td></td>
</tr>
<tr>
<td>CloudSim</td>
<td>Java</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>No data</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>Text</td>
</tr>
<tr>
<td>CloudSimDisk</td>
<td>Java</td>
<td>Yes</td>
<td>No</td>
<td>No data</td>
<td>No</td>
<td>Yes</td>
<td>XLS, Text</td>
<td>No</td>
<td></td>
</tr>
<tr>
<td>CloudsDyn</td>
<td>Java</td>
<td>Yes</td>
<td>No</td>
<td>CSV</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>CSV, JSON</td>
</tr>
<tr>
<td>CMCCloudSimulator</td>
<td>Java</td>
<td>Yes</td>
<td>No</td>
<td>XML, Java classes</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>Text</td>
<td></td>
</tr>
<tr>
<td>DartCSim</td>
<td>Java, C++</td>
<td>No</td>
<td>No</td>
<td>XML</td>
<td>No</td>
<td>Yes</td>
<td>No data</td>
<td>XML</td>
<td></td>
</tr>
<tr>
<td>DCSim(^1)</td>
<td>Java</td>
<td>Yes</td>
<td>No</td>
<td>Java classes</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>Text</td>
</tr>
<tr>
<td>DCSim(^2)</td>
<td>Java</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>No data</td>
<td>No</td>
<td>No</td>
<td>No data</td>
<td>Text</td>
</tr>
<tr>
<td>DISSECT-CF</td>
<td>Java</td>
<td>Yes</td>
<td>No</td>
<td>Java classes</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>Text</td>
<td></td>
</tr>
<tr>
<td>EMU/Sim</td>
<td>Java</td>
<td>Yes</td>
<td>No</td>
<td>XML</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>Text</td>
<td></td>
</tr>
<tr>
<td>GDCSim</td>
<td>C/C++, Shell</td>
<td>No</td>
<td>No</td>
<td>C code</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>Text</td>
<td></td>
</tr>
<tr>
<td>GreenCloud</td>
<td>C++, TCL, JS, CSS, Shell</td>
<td>No</td>
<td>No</td>
<td>TCL</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Dashboard plots</td>
<td></td>
</tr>
<tr>
<td>GroudSim</td>
<td>Java</td>
<td>Yes</td>
<td>No</td>
<td>XML</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>Text</td>
<td></td>
</tr>
<tr>
<td>iCanCloud</td>
<td>C/C++, Shell</td>
<td>Yes</td>
<td>No</td>
<td>NED</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>Text</td>
<td></td>
</tr>
<tr>
<td>iFogSim</td>
<td>Java</td>
<td>Yes</td>
<td>No</td>
<td>JSON</td>
<td>No</td>
<td>Yes</td>
<td>No data</td>
<td>XLS, PDF</td>
<td></td>
</tr>
<tr>
<td>MDCSim</td>
<td>No data</td>
<td>No</td>
<td>No</td>
<td>No data</td>
<td>No</td>
<td>No</td>
<td>No data</td>
<td>No data</td>
<td></td>
</tr>
<tr>
<td>MR-CloudSim</td>
<td>No data</td>
<td>No</td>
<td>No</td>
<td>No data</td>
<td>No</td>
<td>No</td>
<td>No data</td>
<td>N/A</td>
<td>No data</td>
</tr>
<tr>
<td>NetworkCloudSim</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>SimGrid</td>
<td>C/C++</td>
<td>Yes</td>
<td>No</td>
<td>XML, Java classes</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>Text</td>
<td></td>
</tr>
<tr>
<td>SimIC</td>
<td>Java</td>
<td>Yes</td>
<td>No</td>
<td>text, Java classes</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>Text</td>
<td></td>
</tr>
<tr>
<td>SPECI</td>
<td>Java</td>
<td>Yes</td>
<td>No</td>
<td>No data</td>
<td>No</td>
<td>No data</td>
<td>No data</td>
<td>No data</td>
<td>No data</td>
</tr>
<tr>
<td>TeachCloud</td>
<td>Java</td>
<td>Yes</td>
<td>No</td>
<td>Java classes</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>Text</td>
<td></td>
</tr>
<tr>
<td>Ucloud</td>
<td>Java</td>
<td>Yes</td>
<td>No</td>
<td>No data</td>
<td>No</td>
<td>No</td>
<td>No data</td>
<td>No data</td>
<td></td>
</tr>
<tr>
<td>WorkflowSim</td>
<td>Java</td>
<td>Yes</td>
<td>No</td>
<td>Java Classes</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>Text</td>
<td></td>
</tr>
</tbody>
</table>

proposed and target to solve and support different aspects of cloud, edge and fog computing. These features have been presented and compared across these tools with respect to two main categories: general high-level features and high-level technical features of the simulation platforms.

This review identifies the emergence of CloudSim as a de facto base platform for simulation.
development and research. 18 of the platforms analysed were derivatives or extensions of CloudSim. This is not surprising given the early mover advantage CloudSim had, the eminence of the researchers involved, and the quality and timeliness of the release of the simulator platform. There are advantages and disadvantages to such dominance including code reuse, resource efficiencies and development of a wider knowledge base in the use of CloudSim. However, one might also argue that dominance of CloudSim may result in inherited limitations from drawbacks in the CloudSim design.

The multi-level analysis presented identifies some apparent gaps in the features of existing simulation tools. For example, the analysis highlights a gap in the capability of the simulators identified to support distributed execution, i.e. parallel execution on distributed memory systems. Due to the nature of the problem that simulators have to solve, execution and scalability are crucial and are limited by the sequential execution. Similarly, with a number of notable exceptions there are few simulators focussing on emerging cloud use cases e.g. HPC in the cloud, Edge and Fog computing, and IoT. This is unsurprising given the nascent level of these use cases compared to the public cloud IAAS use case.

This review is a significant extension of existing reviews of simulation tools for cloud computing both in terms of breadth and depth however it is not without limitations. Future work is recommended towards a deeper analysis of the tools against alternative real cloud computing scenarios with a focus towards heavy validation of simulated results. Moreover, further analysis can be performed by reviewing simulation models, VM allocation policies, supported cloud services and levels, and in general more cloud oriented specific characteristics. Similarly, whereas this review focuses on simulation tools for cloud computing, an additional survey on the uses to which such tools are employed is warranted and is worthy of investigation.

ACKNOWLEDGEMENTS

This work is partially funded by the European Union’s Horizon 2020 and FP7 Research and Innovation Programmes through CloudLightning (http://www.cloudlightning.eu) under Grant Agreement No. 643946, RECAP (http://www.recap-project.eu) under Grant Agreement No. 732667 and CACTOS (http://www.cactosfp7.eu) under Grant Agreement No. 610711.

REFERENCES


A Review of Cloud Computing Simulation Platforms and Related Environments


