A Preliminary Systematic Review of Computer Science Literature on Cloud Computing Research using Open Source Simulation Platforms

Theo Lynn[1], Anna Gourinovitch[1], James Byrne[1], P. J. Byrne[1], Sergej Svorobej[1],
Konstaninos Giannoutakis[2], David Kenny[1] and John Morrison[3]
[1]Irish Centre for Cloud Computing and Commerce, DCU, Dublin, Ireland
[2]The Centre for Research and Technology, Hellas, Thessaloniki, Greece
[3]Irish Centre for Cloud Computing and Commerce, UCC, Cork, Ireland

Keywords: Cloud Computing, Open Source, Simulation, Simulator, Research Survey, Literature Review.

Abstract: Research and experimentation on live hyperscale clouds is limited by their scale, complexity, value and and issues of commercial sensitivity. As a result, there has been an increase in the development, adaptation and extension of cloud simulation platforms for cloud computing to enable enterprises, application developers and researchers to undertake both testing and experimentation. While there have been numerous surveys of cloud simulation platforms and their features, few surveys examine how these cloud simulation platforms are being used for research purposes. This paper provides a preliminary systematic review of literature on this topic covering 256 papers from 2009 to 2016. The paper aims to provide insights into the current status of cloud computing research using open source cloud simulation platforms. Our two-level analysis scheme includes a descriptive and synthetic analysis against a highly cited taxonomy of cloud computing. The analysis uncovers some imbalances in research and the need for a more granular and refined taxonomy against which to classify cloud computing research using simulators. The paper can be used to guide literature reviews in the area and identifies potential research opportunities for cloud computing and simulation researchers, complementing extant surveys on cloud simulation platforms.

1 INTRODUCTION

Cloud computing is increasingly a mainstream technology for consumers and enterprises alike. While the market is larger and growing, the public cloud is dominated by a small number of extremely large cloud service providers, most notably Amazon Web Services, Microsoft and Google (Gartner, 2016). The scale, complexity, value and commercially sensitive nature of the technology these hyperscale cloud providers and the datacenters that these providers operate means that enterprises and researchers cannot easily undertake experimental research on these platforms. Even if access was provided, application developers would be stimied by their inability to contol and process the network environment and predict and control network conditions (Tian et al. 2015).

Thus in tandem with the rise of and interest in cloud computing, there has been a similar increase in cloud simulators and analysis tools. Whereas there has been numerous survey papers on simulators and their features, there are few papers that explore what researchers are using these simulators for. This paper focuses on open source cloud simulation platforms, toolkits and extensions to those platforms. We make a preliminary attempt to understand the type and focus of research on and using open source cloud simulation platforms using both descriptive and synthetic analysis. In our synthetic analysis, we assess the efficacy of using Rimal, Choi and Lumb's (2009) taxonomy of cloud computing to classify research undertaken using cloud simulation platforms. Finally, we seek to identify trends and potential gaps in research in this field, and contribute to better quality research.

2 METHODS

Simulation of cloud computing remains an emerging topic. Its evolution is impacted by both

Lynn, T., Gourinovitch, A., Byrne, J., Byrne, P., Svorobej, S., Giannoutakis, K., Kenny, D. and Morrison, J.
A Preliminary Systematic Review of Computer Science Literature on Cloud Computing Research using Open Source Simulation Platforms.
DOI: 10.5220/0006351805650573
In *Proceedings of the 7th International Conference on Cloud Computing and Services Science (CLOSER 2017)*, pages 537-545
ISBN: 978-989-758-243-1

developments in cloud computing technologies and simulation technologies and techniques. As such, extant surveys are prone to be outdated regularly. Furthermore, extant literature surveys tend to focus exclusively on the features and performance of simulators, and typically those to support discrete event simulation, and not necessarily research gaps in the cloud computing literature. With this in mind, this preliminary systematic literature review to describe the current state of computer science research on the use of simulation for cloud computing is appropriate at this time. And while atheoretical, such literature reviews will assist in the transfer and assimilation of related knowledge and understanding on the topic (Rowe, 2014).

While Rowe (2014) suggests authors aim for complete coverage, such coverage is neither possible nor appropriate for a conference paper. Thus, we limit this preliminary analysis to the computer science discipline only and only publication outlets featured in the IEEE Xplore digital library. Furthermore, we focus our survey only on open source simulation platforms or toolkits for cloud computing identified systematically from the literature. Using a variety of terms, we identified 281 articles in relation to the topic. After further scrutiny, the final list was reduced to 256 (Lynn et al. 2017). Papers were omitted on the grounds that their main focus was not specifically cloud computing, were errata, notices, keynotes or other documents, or were not open source. The full reference list for these papers and associated data are available as an online dataset at http://cloudlightning.eu/dissemination/publications/simulation-platforms/ and is referenced herein at Lynn et al. (2017). In line with Rowe (2014), we present a descriptive analysis in Section 3 and a synthetic analysis in Section 4.

3 DESCRIPTIVE ANALYSIS

In this section, we present a descriptive analysis of literature on cloud computing simulation research in the IEEE Xplore Digital Library between 2009 and 2016 through three key lenses: (1) year of publication, (2) publication outlets and (3) simulator platform or toolkit. These lenses provide insights into the trajectory of computer science research on this topic as well as insights into platform selection.

The first publication on cloud computing simulation research using open source platforms appears in 2009 with Buyya, Ranjan and Calheiros' introduction of the CloudSim toolkit. From 2009 onwards, publications on the topic increase consistently, largely driven by papers relating to CloudSim or the introduction of new cloud simulation platforms. By 2015, cloud simulation papers using open source platforms is a regular topic in computer science publications having grown from 2 in 2009 to 78 in 2014 (See Figure 1). This reflects the interest in cloud computing generally and the growth of cloud computing adoption (Markets and Markets, 2015).

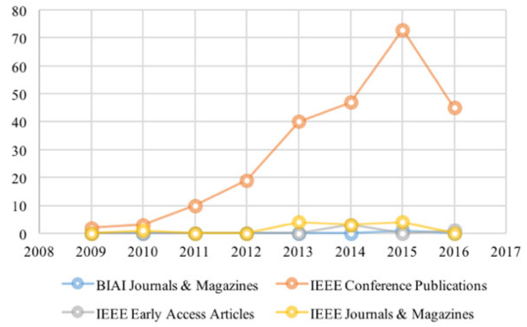

Figure 1: Cloud simulation research using open source platforms by publication outlet and year.

Figure 1 presents a descriptive analysis from a publication outlet perspective. This analysis shows that the overwhelming majority of papers on the topic in the IEEE Xplore Digital Library are conference papers (93%). Only 18 (7%) of papers are published in BIAI and IEEE journals or magazines and these are relatively recent; the earliest being 2013. This is not unsurprising given the nascent stage of cloud computing, open source simulation platforms and toolkits, and simulation research on cloud computing generally.

As cloud computing is a relatively new field, high ranking outlets dealing specifically with the topic are scarce and those that do exist may not be affiliated with IEEE or may require longer turnaround times for acceptance. Given the high number of conference papers since 2013, one would expect a greater number of journal articles in the coming years.

Finally, we look at publications by simulation platform, extension or toolkit. For the purpose of this analysis, we remove five papers that survey the field, provide an overview, or compare one or more simulation platforms. Of the remaining 256, 85% (218) relate to CloudSim, extensions or derivative simulators. Given the seminal nature of Buyya et al.'s work on CloudSim, this is unsurprising. While the sample features 25 simulation platforms, extensions and toolkits, 12 are derivatives or

Table 1: Total publications by open source platform, toolkit or extension by year.

Simulator, Toolkit or Extension (Base Platform)	2009	2010	2011	2012	2013	2014	2015	2016	Total
CACTOS (Palladio, SimuLizar, CloudScale)						2			2
CDOSim (CloudSim)				1					1
CEPSim (CloudSim)							1		1
Cloud2Sim (CloudSim, AEF)						1			1
CloudAnalyst (CloudSim)		1		1	4	3	6	3	18
CloudNetSim++						1			1
CloudReports (CloudSim)							1		1
CloudSched							1		1
CloudSim	1	1	7	9	29	43	61	37	188
CloudSimDisk (CloudSim)							1		1
CloudSimSDN (CloudSim)							1		1
CMCloudSimulator (CloudSim)								1	1
DartCSim (CloudSim)				1	1				2
DCSim(1)				1	2				4
DCSim(2)				1					1
GDCSim (Blue Tool)			1						1
GreenCloud		1		1	3	2	1	4	12
GroudSim/DISSECT-CF								1	1
iCanCloud (SimCan)			1	1					2
MDCSim (CSIM)	1								1
MR-CloudSim (CloudSim)				1					1
NetworkCloudSim (CloudSim)			1						1
SimGrid				1	2		4		7
SimIC (SimJava)					3	1	1		5
SPECI (SimKit)		1							1
WorkflowSim (CloudSim)				1					1
Total	2	4	10	19	44	53	78	46	256

extensions of CloudSim (See Table 3). Other platforms and toolkits are introduced; however, few have gained the traction of CloudSim; Green Cloud lags significantly in publications with 12.

CloudSim's dominance may be interpreted in a variety ways. It may reflect ease of use, platform stability, feature quality, the size of the user community or a publication bias. It certainly indicates CloudSim as the leading open source platform for cloud modelling and simulation if not the de facto standard. Of these 256, 45 papers relate to the design, development and extension of simulation platforms (see Table 3). The complete reference list for these papers and associated data is available at Lynn et al. (2017). Three platforms do not feature in this sample (Groudsim/DISSECT-CF, SPECI and CloudReports). The design papers for these projects may feature in other digital libraries. It should be noted that the overwhelming majority of these platforms support discrete event simulations and not continuous or real-time simulations although Malik et al. (2014) and Aguero et al. (2015) suggest that CloudNetSim++ and CloudSim respectively can be used for near real-time simulations of applications.

The remaining 211 papers in the sample relate to the use of the platforms for research (See Lynn et al. (2017) for complete reference list and associated data). Again, the overwhelming majority are undertaking research using CloudSim or derivatives and extensions to CloudSim (92%). The increasing use of CloudAnalyst can be explained by its utility in providing a GUI for CloudSim.

4 SYNTHETIC ANALYSIS

To review the status and research trends in the existing computer science literature on cloud computing research using open source simulation

Table 2: Research papers undertaking research using an open source simulation platform.

Simulator	2010	2011	2012	2013	2014	2015	2016	Total
CloudAnalyst*			1	4	2	6	3	16
CloudReports*						1		1
CloudSim	1	6	9	27	43	60	32	178
GreenCloud				3	2		4	9
GroudSim/ DISSECT-CF							1	1
SimGrid						2		2
SimIC				2		1		3
SPECI	1							1
Total	2	6	10	36	47	70	40	211

Table 3: Research papers on the design and development of open source simulation platforms by simulator type.

Simulator	Papers	References
CACTOS	2	Ostberg et al. (2014); Wesner et al. (2014)
CDOSim*	1	Fittkau et al. (2012)
CEPSim*	1	Higashino et al. (2015)
Cloud2Sim*	1	Kathiravelu and Veiga (2014)
CloudAnalyst*	2	Wickremashinge et al. (2010); Mahajan and Dahiya (2014)
CloudNetSim++	1	Malik et al. (2014)
CloudSched	1	Tian et al. (2015)
CloudSim*	10	Buyya et al. (2009); Werner et al. (2011); Long et al. (2013); Suciu et al. (2013); Nagamani et al. (2016); Kouba et al. (2016); Pittl et al. (2016a); Pittl et al. (2016b); Aguero et al. (2015); Chavan et al. (2016)
CloudSimDisk*	1	Louis et al. (2015)
CloudSimSDN*	1	Son et al. (2015)
CMCloudSimulator*	1	Alves et al. (2016)
DartCSim*	2	Li et al. (2012); Li et al. (2013)
DCSim*	3	Tighe et al. (2012); Tighe et al. (2013); Keller et al. (2013)
DCSim**	1	Chen et al. (2012)
GDCSim	1	Gupta et al. (2011)
GreenCloud	3	Kliazovich et al. (2010); Kilazovich et al (2012); Sharkh et al. (2015)
iCanCloud	2	Nunez et al. (2011); Castane et al. (2012)
MDCSim	1	Lim et al. (2009)
MR-CloudSim*	1	Jung and Kim (2012)
NetworkCloudSim*	1	Garg and Buyya (2011)
SimGrid	5	Bobelin et al. (2012); Hirofuchi and Lebre (2013a) Hirofuchi et al. (2013b); Lebre et al. (2015); Hirofuchi et a. (2015)
SimIC	2	Sotiriadis et al. (2013); Sotiriadis et al. (2014)
WorkflowSim*	1	Chen and Deelman (2012)

*denotes a simulator based on CloudSim. **Two simulators called DCSim were launched in 2012 independently of each

platforms, we employ the taxonomy of cloud computing systems proposed by Rimal et al. (2009) to classify the 211 papers dealing with the utilisation of simulation platforms rather than the platforms per se. For reference purposes, the classified data is available at http://cloudlightning.eu/dissemination/publications/simulation-platforms/ (Lynn et al. 2017).

The taxonomic analysis undertaken in this review includes six key elements identified by Rimal et al (2009):

- Architecture – includes private cloud, public cloud, hybrid cloud and federated clouds.
- Virtualisation management – includes any activity related to the abstraction of logical

resources away from their underlying physical resources.

- Services – includes Infrastructure-as-a-Service (IaaS), Software-as-a-Service (SaaS).
- Platform-as-a-Service (PaaS) and other servitised software.
- Fault tolerance - includes simulations of outages
- Security – includes attack simulation and methods for enhanced security or trust.
- Other – includes load balancing, interoperability, and data storage.

The primary focus of the papers analysed was hyperscale data center performance and typically are indiscriminate on whether the data center is an enterprise (private cloud) or hyperscale. It can be reasonably assumed that the primary architectural focus is public cloud. A small number of papers specify explicitly the cloud architecture they are seeking to address e.g. Canedo et al. (2012), Simao and Veiga (2013) and Sujan and Devi (2015) have private clouds as a specific focus and Sqalli et al. (2012) focus on hybrid clouds. Sotiriadis et al. address the issue of inter-cloud simulation in a series of papers using the SimIC simulation toolkit (Sotiriadis et al. 2013a, 2013b, and 2015) and similarly Hamze et al. (2014) and Aazam and Huh (2014) seek to simulate inter-cloud scenarios using CloudSim. A further five papers address related federated cloud simulation scenarios (Patel and Sarje, 2012; Aazam and Huh, 2014; Aral and Ovatman, 2015; Wen et al. 2016; Pacini et al. 2016). The vast majority of papers in our review dealt with some aspect of virtualisation/resource management. For classifying papers, we employ Singh and Chana's high level taxonomy of resource management (Singh and Chana, 2016).

Classifying simulation papers by resource management is difficult due to overlapping between various resource management concerns. Resource scheduling is particularly prevalent; it is not surprising that a significant number of papers seek to address this issue as it is considered as hard as a Nondeterministic Polynomial (NP) optimization problem (Zhan et al. 2015). In contrast, studies using open source simulation platforms to explore monitoring for resource management are relatively recent and few. An additional catch-all category was added to capture papers simulating multiple virtualisation stages and processes in different contexts e.g. mobile (Li and Li, 2013; Artail et al. 2015), IOT (Shaoling et al., 2015) and manufacturing (Dong and Jianling, 2013).

Classification of the cloud simulation literature identified in this paper by service type does not provide any substantial insights. Due to the nature of the simulation platforms and toolkits available and reviewed, the focus is primarily data centers and the IaaS layer. Indeed, only 14 papers specifically identify IaaS. None specify the PaaS layer, and only two specifically identify the SaaS layer other than in the wider layered sense of cloud computing. Two papers, Nuaimi et al. (2013) and Zhihua (2013) address Data-as-a-Service and Network-as-a-Service respectively.

There are few papers on fault tolerance as a discrete topic of study within the papers reviewed. Six papers identify fault tolerance as a focus of study. Four (Wang et al. 2015; Goutam et al. 2014; and Bosilca et al. 2014; Abderrahim and Choukair, 2015) have fault tolerance and fault tolerance mechanisms as a primary focus whereas two papers refer to improved fault tolerance as an outcome of their architecture and algorithms respectively (Pardesi et al. 2014; Yadav and Kushwaha; 2014).

Despite the wider focus on security in cloud computing as a major barrier to adoption and concern to enterprises, the general public and policymakers, the papers reviewed did not feature a significant number of papers on security. The seven papers identified can be classified into five categories: security as a system requirement (Wen et al. 2016), attack simulations (Karthik and Shah; 2014), malicious virtual machines (Bazm et al. 2015), novel methods for secure data management (Hani and Dichter, 2016; Xu et al. 2016; Zardari et al. 2014; Boomija, 2016) and security education (Shi et al. 2016).

43 papers addressed the issue of load balancing. In cloud computing, load balancing occurs in three stages - data center selection, virtual machine scheduling, and task scheduling at a selected data

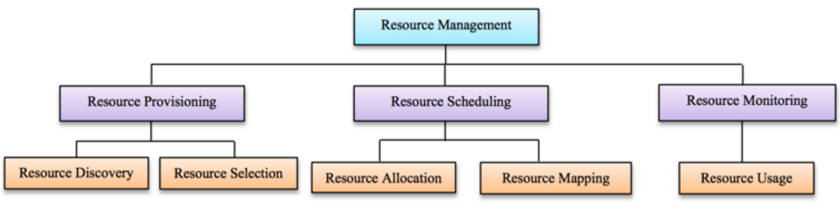

Figure 2: Taxonomy of Resource Management in Cloud Computing (Singh and Chana, 2016).

Table 4: Classification of papers using open source simulators by Singh and Chana's (2016) Taxonomy of Resource Management.

Category	Topics	Papers	References
Architecture	Public*	NA	NA
Architecture	Private	3	Canedo et al. (2012), Simao and Veiga (2013); Sujan and Devi (2015)
Architecture	Hybrid	1	Sqalli et al. (2012)
Architecture	Other	10	Patel and Sarje (2012); Sotiriadis et al. (2013a) Sotiriadis et al. (2013b); Aazam and Huh (2014); Hamze et al. (2014); Aazam and Huh (2014); Sotiriadis et al. (2015); Aral and Ovatman (2015); Wen et al. (2016); Pacini et al.(2016)
Service	IaaS	13	Kim et al. (2011); Achar and Thilagam (2014); Abar et al. (2014); Rodriguez and Buyya (2014); Hamze et al. (2014); Karthik and Shah (2014) Tian et al. (2015); Luo et al. (2015); Sotiriadis et al. (2015); Pittl et al. (2015); Chowdhury et al. (2015); C. Sequin et al. (2015); Pittl et al. (2016)
Service	PaaS	-	-
Service	SaaS	3	Achar et al. (2012); Huang et al. (2012); Zotkiewicz et al. (2016)
Virtualisation Management	Resource Provisioning	40	Sriram and Cliff (2010); Shi et al. (2011); Bose et al. (2011); Canedo et al. (2012) Patel and Sarje (2012); Huang et al. (2012); Cao and Zhu (2013); Patel et al. (2013); Sotiriadis et al. (2013), Tao and Dong (2013); Kord and Haghighi (2013); Masoumzadeh and Hlavacs (2013); Achar and Thilagam (2014); Sahal and Omara (2014); Udeze et al. (2014); Lo et al. (2014); Abar et al. (2014); Rodriguez and Buyya (2014); Cao et al. (2014); Azzam and Huh (2015a); Aazam and Huh (2015b); Garala and Dobariya (2015); Sotiriadis et al. (2015); Xavier et al. (2015); Fakhfakh et al. (2015a); Fakhfakh et al. (2015b); Thaman and Singh (2015); Monil and Rahman (2015); Chowdhury et al. (2015); Rekik et al. (2015); Alhiyari and El-Mousa (2015); Sharma and Mahrishi (2015); Chen et al. (2015); Li et al. (2015); Xue et al. (2016); Vedova et al. (2016); Pacini et al. (2016); Ranjana et al. (2016); Selim et al. (2016); Shidik et al. (2016); Xavier et al. (2016);
Virtualisation Management	Resource Scheduling	91	Jeyarani et al. (2010); Li et al. (2011); Taheri and Zamanifar (2011); Achar et al. (2012); Simao and Veiga (2013); Pacini et al. (2013); Sotiriadis et al. (2013); Domanal and Reddy (2013); Kilazovich et al. (2013); Dubey et al. (2013); Yu et al. (2013); Yun (2013); Takouna et al. (2013); Tawfeek et al. (2013); Guerout and Alaya (2013); Jung et al. (2013); Vijayalakshmi and Prathibha (2013); Perret et al. (2013); Hu and Yu (2013); Li et al. (2013); Chen et al. (2013a); Chen et al. (2013b); Ru and Keung (2013); Royaee and. Mohammadi (2013); Ming et al. (2014); Limrattanasilp and Gertpho (2014); Li et al. (2014); Garg and Krishna (2014); Haidiri et al. (2014); Gupta et al. (2014); Rodriguez and Buyya (2014); Yadav and Kushwaha (2014); Bagwaiya and Raghuwanshi (2014); Mathew et al. (2014); Saxena and Chouhan (2014); Faria et al. (2014); Sharma and Bharti (2014); Komarasamy and Muthuswamy (2014); Sahal and Omara (2014); Lou et al. (2014); Tsai et al. (2014); Luo and Yi (2014); Malekloo and Nara (2014); Indira and KavithaDevi (2014); Ashwin et al. (2014); Chen et al. (2015); Zhu et al. (2015); Sujan and Devi (2015); Domanal and Reddy (2015); Tian et al. (2015); Wang et al. (2015); Sotiriadis et al. (2015); Whittington et al. (2015); Alahmadi et al. (2015); Khedher and Jarraya (2015); Komarasamy and Muthuswamy (2015); Rajeshwari and Dakshayini (2015); Garala and Dobariya (2015); Bhutani et al. (2015); Saxena and Saxena (2015a); Mennour et al. (2015); Khalili and Babamir (2015); Kumari et al. (2015); Ali and Hamad (2015); Saxena and Saxena (2015b); Adrian and Heryawan (2015); Al-Olimat et al. (2015); Alhiyari and El-Mousa (2015); Bruneo et al. (2015); Khanna and Sarishma (2015); Fareh et al. (2015); Asemi et al. (2015); Tareghian and Bornaee (2015); Thanh et al. (2015); Santra and Mali (2015); Elhady and Tawfeek (2015); Ali and Ozkasap (2016a); Ali and Ozkasap (2016b); Zotkiewicz et al. (2016); Pacini et al. (2016); Simao and Veiga (2016); Vedova et al. (2016); Atiewi et al. (2016); Qiu and Hwang (2016); Gupta et al. (2016); Kimpan and Kruekaew (2016); Ettikyala and Latha (2016); Hguyen et al. (2016); Wang et al. (2016); Nehru et al. (2016); Aslanpour and Dashti (2016)
Virtualisation Management	Resource Monitoring	13	Koushal and Johri (2013); Monil et al. (2014); Quwaider and Jararweh (2014); Abar et al. (2014); Sotiriadis et al. (2014); Lu et al. (2015); Wang et al. (2015); Rekik et al. (2015); Monil and Rahman (2015); Chen et al. (2015); Rajeshwari and Dakshayini (2015); Aazam et al. (2016)
Virtualisation Management	VM Placement	14	Bose et al. (2011); Guerout and Alaya (2013); Kord and Haghighi (2013); Varalakshmi and Maheshwari (2013); Zhang et al. (2014); Malekloo and Kara (2014); Aral and Ovatman (2015); Chowdhury et al. (2015a); Chowdhury et al. (2015b); Chavan and Kaveri (2015); Benali et al. (2016); Ranjana et al. (2016); Alharbi and Yang (2016); Malekzai et al. (2016)
Virtualisation Management	VM Migration	12	Takouna et al. (2012); Takouna et al. (2013); Masoumzadeh and Hlavacs (2013); Razali et al. (2014); Monil et al. (2014); Yakhchi et al. (2015); Ghafari et al. (2015); Chowdhury et al. (2015); Monil and Rahman (2015); Alhiyari and El-Mousa (2015); Selim et al. (2016); Shidik et al. (2016); Maio et al. (2016)

Table 4: Classification of papers using open source simulators by Singh and Chana's (2016) Taxonomy of Resource Management (cont.).

Virtualisation Management	Other	20	Chen et al. (2011); Sallam and Li (2012); Shiraz et al. (2012); Dong and Jianling (2013); Shiraz et al. (2012)); Guzek et al. (2013); Li and Li (2013); Householder et al. (2014); Moreno et al. (2014); Chavan and Kaveri (2014); Hussein et al. (2014); Nguyen et al. (2015); Li et al. (2015); Artail et al. (2015); Shaoling et al. (2015); Bonacquisto et al. (2015);Yusof et al. (2015); Tsai et al. (2015); Akbari et al. (2016); Rawat et al. (2016); Fang et al. (2016)
Fault Tolerance	Fault Tolerance	6	Wang et al. (2015); Goutam et al. (2014); Bosilca et al. (2014); Pardesi et al. (2014); Yadav and Kushwaha (2014); Abderrahim and Choukair (2015);
Security	Security	8	Zardari et al. (2014); Karthik and Shah (2014); Bazm et al. (2015); Wen et al. (2016) Hani and Dichter (2016) Xu et al (2016); Boomija (2016); Shi et al. (2016).
Other	Load Balancing	43	Jeyarani et al. (2010); Li et al. (2011); Rawat et al. (2012); Chen et al. (2013); Domanal and Reddy (2013); Kliazovich (2013); Deye et al. (2013); Nuaimi et al. (2013); Goutam et al. (2014); Domanal and Reddy (2014); Mesbahi et al. (2014); Roopa et al. (2014); Lou et al. (2014); Bagwaiya et al. (2014); Ashwin et al. (2014); Razali et al. (2014); Dhurandher et al. (2014); Haidri et al. (2014); Soni and Kalra (2014); Aazam and Huh (2014); Tang et al. (2014); Aslanzadeh and Chaczko (2015); Luo et al. (2015); Domanal and Reddy (2015); Rajeshwari and Dakshayini (2015); Garala and Dobariya (2015); Santra and Mali (2015); Dam et al. (2015); Ghumman and Kaur (2015); Kulkarni and Annappa (2015); Panwar and Mallick (2015); Garg et al. (2015); Sowmya et al. (2015); Yakhchi et al. (2015); Qiu and Hwang (2016); Kimpan and Kruekaew (2016); Ye et al. (2016); Atiewi et al. (2016); Mesbahi et al. (2016); Kumar and Kalra (2016); Ettikyala and Latha (2016); Nishad et al. (2016)
Other	Interoperability	13	Sotiriadis et al (2013a); Sotiriadis et al. (2013b); Azzam and Huh (2014a); Hamze et al. (2014); Achar and Thilagam (2014); Mahalle et al. (2015); Pacini et al. (2016); Sqalli et al. (2012); Azzam and Huh (2014b); Sotiriadis et al. (2015); Xavier et al. (2015); Giupta et al. (2016); Benali et al. (2016)
Other	Storage	11	Bose et al. (2011); Nuaimi et al. (2013); Kaveri and Chavan (2013); Roopa et al. (2014); Zhang et al. (2014); Quwaider and Jararweh (2014); Seguin et al. (2015); Guthmuller et al. (2015); Xue et al. (2016); Zhou et al. (2016); Xu et al. (2016)

**The vast majority of papers address public cloud but do not necessarily specify it*

center (Atiewi et al. 2016). As such it is dealt with independently of resource management and resource scheduling although more efficient load balancing is often an objective of papers in both the resource management and scheduling literature that we reviewed. Accordingly, there may be some misclassification or duplication in classification in this context. It should be noted that while clustering and load balancing are often used interchangeably by practitioners, we have excluded papers on clustering in our count as load balancing can occur without clustering. There were relatively few papers (13) addressing the issue of interoperability specifically and these tended to focus on inter-cloud, federation and brokerage scenarios. Even then, interoperability may be considered a secondary goal or rather a necessity given the context of those studies. Similarly, there are relatively few papers addressing storage (11). These are wider than scalable storage and include studies on addressing the impact of data replication (Bose et al. (2011); Nuaimi et al. (2013); Zhang et al. (2014); Xue et al. (2016)), distributed file systems (Seguin et al. 2015), security (Xu et al. 2016) and wider approaches and techniques for performance optimization (Kaveri and Chavan (2013); Roopa et al. (2014); Quwaider and Jararweh (2014); Guthmuller et al. (2015); and Zhou et al. (2016)).

5 CONCLUSIONS

This paper completed a preliminary systematic literature review of articles on and using open source simulation platforms featured in the IEEE Xplore digital library. We employed two complementary analyses – a descriptive analysis and a synthetic analysis. The synthetic analysis employed a highly cited taxonomy of cloud computing to organise the literature (Rimal et al. 2009). The objectives of this paper were multifold. Firstly, we sought to organise research on and using open source cloud simulation platforms. Secondly, we sought to assess the efficacy of using Rimal et al.'s taxonomy of cloud computing to classify research. Thirdly, we sought to identify trends and potential gaps in research in this field, and to contribute to better quality research.

There are numerous surveys of cloud simulation platform literature. These papers typically focus on the features of the platforms rather than how these platforms have been applied. This paper makes an original contribution by examining the application of the simulation platforms in a cloud computing context. Notwithstanding this contribution, the paper is limited to the IEEE Xplore Library and open source simulators. Further surveys are needed including studies using commercial and proprietary simulation platforms and with a wider set of publication outlets along the lines of Paulsson et al. (2016).

The descriptive analysis identifies clear trends and areas for further research. Cloud simulation platform research has become an established domain now and is consistently featured in IEEE conferences. The momentum developed since 2009 should result in a higher number of journal publications in the coming years. Notwithstanding this, there is a clear need for more comprehensive publications in journals. This may be a factor of journal editorial inflection, the quality of publications or the volume of papers submitted. Clearly, CloudSim is the dominant cloud simulation platform for research and this paper provides strong supporting evidence for the selection of CloudSim for future research initiatives. Such dominance can be perceived as both a positive and negative factor. For example, there is a dearth of research on continuous and (near) real-time simulations, possibly due to limitations by existing platforms including CloudSim.

The employment of a taxonomy of cloud computing to classify papers was of benefit. Again, it highlights areas for increased focus and clarity. From a communications perspective, researchers presenting cloud research should possibly provide greater clarity on the applicability of their research for target architectures and services. Cloud simulation platforms provide a valuable service to resource management researchers. The relatively high volume of research reflects both the complexity of the area and the interest of researchers. However, from a market-focussed perspective, one might argue that security, QoS and reliability may be of more interest. This is where Rimal et al.'s taxonomy, while useful as a high-level frame of analysis, is lacking. It does not provide the sufficient granularity and detail needed to provide a more robust classification of literature in this area. Even by augmenting the analysis with Singh and Chana (2015), evidently a new more complete taxonomy is need for cloud computing. Future research should not only develop a more comprehensive taxonomy for classification but accommodate emerging themes. Motivations such as energy efficiency, profitability cost effectiveness feature in the literature as well as new and emerging use cases e.g. the impact of heterogeneous resources, autonomic and self-adaptive management techniques, mobile clouds, IOT and FOG computing, MapReduce and Hadoop, and HPC in the cloud. Content mining and autonomic classification may help identify new insights and relationships in a way that the systematic approach employed in this paper does not clearly implies an imbalance in focus with a heavy emphasis on resource provisioning, scheduling and load balancing. One could argue that the literature reviewed is more academically-focussed than market-focussed. This might explain the relatively few papers on security including the highly topical areas of data protection and security, interoperability and fault tolerance. Similarly, the lack of papers on PaaS and SaaS, while understandable, presents an opportunity for future research on and using open source simulation platforms. Similarly, while the papers feature studies on new and emerging issues and applications such as those mentioned in the previous paragraph, these are relatively few and are areas worthy of greater focus. Finally, the majority of studies focus on discrete event simulations and not continuous or (near) real-time cloud simulations. While these are both conceptually and technically challenging, they should not be disregarded.

Open source cloud simulation platforms will continue to evolve over time. Updated surveys are needed to keep researchers informed on both the evolving features and performance of these platforms. However, such surveys are only one part of the story. There is also a need to present surveys and literature on the use of these platforms in research. This paper provides an initial contribution.

ACKNOWLEDGEMENTS

This work is partially funded by the European Union's Horizon 2020 Research and Innovation Programme through the CloudLightning project (http://www.cloudlightning.eu) under Grant Agreement Number 643946 and the RECAP project (http://www.recap-project.eu) under Grant Agreement Number 732667.

REFERENCES

Due to the large number of papers reviewed in this study and conference page limitations, the full reference list can be found at http://cloudlightning.eu/dissemination/publications/simulation-platforms/ (Lynn et al. 2017).

Buyya, R., Ranjan, R., & Calheiros, R. N. (2009, June).

Modeling and simulation of scalable Cloud computing environments and the CloudSim toolkit: Challenges and opportunities. In High Performance Computing & Simulation, 2009. HPCS'09. International Conference on (pp. 1-11). IEEE.

Gartner, Inc (2016). Magic Quadrant for Cloud.

Lynn, T., Gourinovitch, A., Byrne, J., Byrne, P. J., Svorobej, S., Giannoutakis, K., Kenny, D. and Morrison, J. (2017, January). A Preliminary Systematic Review of Computer Science Literature on Cloud Computing Research Using Open Source Simulation Platforms [Data and Full References]. http://cloudlightning.eu/dissemination/publications/simulation-platforms/

Markets and Markets. (2015). Cloud Services Brokerage Market by Types - Global Forecast to 2020. June 2016. Retrieved from http://www.researchandmarkets.com/research/xwbvvz/cloud_services.

Paulsson, V., Emeakaroha, V., Morrison, J., & Lynn, T. (2016). Cloud Service Brokerage: A systematic literature review using a software development lifecycle. In AMCIS 2016.

Rimal, B. P., Choi, E., & Lumb, I. (2009). A taxonomy and survey of cloud computing systems. INC, IMS and IDC, 44-51.

Rowe, F. (2014). What literature review is not: diversity, boundaries and recommendations. European Journal of Information Systems, 23(3), 241-255.

Singh, S., & Chana, I. (2016). Cloud resource provisioning: survey, status and future research directions. Knowledge and Information Systems, 49(3), 1005-1069.

Tian, W., Zhao, Y., Xu, M., Zhong, Y., & Sun, X. (2015). A toolkit for modeling and simulation of real-time virtual machine allocation in a cloud data center. IEEE Transactions on Automation Science and Engineering, 12(1), 153-161.

Zhan, Z. H., Liu, X. F., Gong, Y. J., Zhang, J., Chung, H. S. H., & Li, Y. (2015). Cloud computing resource scheduling and a survey of its evolutionary approaches. ACM Computing Surveys (CSUR), 47(4), 6.