Simulating Resource Management across the Cloud-to-Thing Continuum

Created by: Malika Bendechache [1] , Sergej Svorobej [2] , Patricia Takako Endo [3] , Theo Lynn [4]

1, School of Computing, Dublin City University; malika.bendechache@dcu.ie
2, School of Computer Science and Statistics, Trinity College Dublin; sergej.svorobej@tcd.ie
3, Universidade de Pernambuco; patricia.endo@upe.br
4, Dublin City University; theo.lynn@dcu.ie

Version received: 15 July 2020

In recent years, there has been significant advancement in resource management mechanisms for cloud computing infrastructure performance in terms of cost, quality of service (QoS) and energy consumption. The emergence of the Internet of Things has led to the development of infrastructure that extends beyond centralised data centers from the cloud to the edge, the so-called cloud-to-thing continuum (C2T). This infrastructure is characterised by extreme heterogeneity, geographic distribution, and complexity, where the key performance indicators (KPIs) for the traditional model of cloud computing may no longer apply in the same way. Existing resource management mechanisms may not be suitable for such complex environments and therefore require thorough testing, validation and evaluation before even being considered for live system implementation. Similarly, previously discounted resource management proposals may be more relevant and worthy of revisiting. Simulation is a widely used technique in the development and evaluation of resource management mechanisms for cloud computing but is a relatively nascent research area for new C2T computing paradigms such as fog and edge computing. We present a methodical literature analysis of C2T resource management research using simulation software tools to assist researchers in identifying suitable methods, algorithms, and simulation approaches for future research. We analyse 35 research articles from a total collection of 317 journal articles published from January 2009 to March 2019. We present our descriptive and synthetic analysis from a variety of perspectives including resource management, C2T layer, and simulation.

1. Introduction

Today, we are in the midst of a new evolution of the Internet, the Internet of Things (IoT), “a global network and service infrastructure of variable density and connectivity with self-configuring capabilities based on standard and interoperable protocols and formats. IoT consists of heterogeneous things that have identities, physical and virtual attributes, and are seamlessly and securely integrated into the Internet” [1]. Forecasts suggest that the number of connected things will continue to explode over the next five years reaching 42 billion and generating 79.4 zettabytes of data [2]. This explosion of data is changing the characteristics of Internet traffic fundamentally.

The traditional model of cloud computing was designed and built out using substantially different economic and technical paradigms than IoT, with centralised storage and processing in the cloud and associated economies of scale from centralising these functions in multi-tenant data centers. In contrast, IoT is characterised by extremeness-extremely peaky traffic, extreme heterogeneity of device form, function and latency requirements, as well as extreme geographic distribution of infrastructure from the cloud to the device, the so-called cloud-to-thing (C2T) continuum. To support computing across the C2T continuum, new paradigms of computing have emerged, e.g., fog and edge computing. These new paradigms facilitate the deployment of distributed, latency-aware applications and services by enabling local computing capability on a network-accessible device (thing) or somewhere between the cloud and the edge, the fog [3]. The complexity of such distributed infrastructure mandates automated orchestration of applications and end-to-end management of networking, infrastructure and workload placement. This is required to efficiently and dynamically deploy workloads that meet specific targets (many of which may have geographic, latency, or other user profile idiosyncrasies), while at the same time (1) minimising the cost and energy consumption of finite physical hardware resources, and (2) meeting service level agreement (SLA) commitments [4].

Designing resource management mechanisms that identify and select resources (resource provisioning), and then map and execute workloads based on those resources and the required service levels (resource scheduling) across the C2T continuum, at hyper-scale, increases complexity at orders of magnitude significantly greater than traditional cloud computing. As such, there is a substantial ongoing effort to review the efficacy of extant resource management mechanisms in the IoT context and propose new mechanisms that can accommodate the specific constraints and requirements of edge and fog computing [5]. The evaluation of such mechanisms is not without its challenges. Due to the cost, complexity and risk associated with both experimenting on and replicating modern hyper-scale infrastructure, simulation software tools are used extensively to investigate different aspects of resource management and impact, under different scenarios [6]. Even so, such tools only recently evolved to simulate geographically distributed cloud infrastructure, warehouse-scale computing, and heterogeneous resources [7].

This article presents a methodical survey of scholarly literature on the use of simulation software tools for evaluating resource management methods across the C2T continuum. There have been a number of literature surveys on (i) resource management in the cloud, and more recently fog and edge computing [5][8][9][10][11][12], (ii) open-source cloud simulators [13][14], and (iii) simulating fog and edge computing applications[15]. These surveys often conflate mathematical modelling techniques with simulation software tools or do not identify or discuss the use of the simulation software tools and techniques in sufficient detail. No dedicated literature survey could be identified on the use of simulation software for evaluating resource management methods across the C2T continuum. Not only has the research base grown in both resource management and simulation software, but the emergence of joint cloud-fog-edge computing use cases requires a methodical literature survey to evaluate and integrate the existing research available in this field. Our survey meets this call. Our contribution is as follows: (1) extending and complementing existing surveys with a dedicated survey on the use of simulation software tools for evaluating resource management methods across the C2T continuum; (2) identifying simulation software tools for researchers to employ or extend in order to implement and evaluate their resource management proposals in this domain; and (3) categorising extant research to enable researchers efficiently identify gaps in the literature and thereby inform relevant research agendas.

2. Systematic Literature Review Results

Table 1 summarises the selected research works found. As can be seen, there is a significant imbalance in the evaluation of resource management across the C2T continuum. At a high level, every paper included resource scheduling. More specifically, 28 of the 35 papers (80%) focus on resource allocation; it was the primary focus of all papers regardless of C2T layer. This is not wholly unsurprising given that resource allocation is the central task in resource scheduling. This suggests a significant opportunity for research in evaluating resource mapping, resource monitoring and load balancing using simulation, in all layers of the C2T continuum. More research works are also needed With regard to resource provisioning, addressed both detection and selection tasks. Furthermore, there were only ten fog and edge computing articles in total and the remainder of the articles focus on the cloud layer.

Publication	Year	C2T Context			RM Context						Simulation Methodology			Simulator Tool	Variable or KPI under Study
		Cloud	Fog	Edge	Provisioning: Detection	Provisioning: Selection	Scheduling: Mapping	Scheduling: Allocation	Scheduling: Monitoring	Scheduling: Load Balancing	DES	DTS	Hybrid		
Samimi et al. [42]	2016	✓						✓			✓			CloudSim	Defining auction strategies
da Silva and da Fonseca [46]	2016	✓						✓			✓			CloudSim *	VM placement
Cai et al. [48]	2017	✓			✓	✓	✓				✓			ElasticSim *	Rental time and rental cost
Castro et al. [47]	2016	✓						✓			✓			CloudSim	VM placement
Ranjbari and Torkestani [52]	2018	✓						✓			✓			CloudSim	Energy consumption and SLA violations
Gawali and Shinde [53]	2018	✓						✓			✓			CloudSim	Bandwidth and workload
Malawski et al. [41]	2015	✓			✓	✓	✓				✓			CloudSim	Cost (budget), Run time
Heilig et al. [49]	2017	✓			✓	✓		✓			✓			CloudSim *	Cost and latency
Moghaddam et al. [63]	2019	✓						✓	✓		✓			CloudSim *	Autoscaling (vertical and horizontal)
Al-Mansoori et al. [55]	2018	✓						✓			✓			FogNetSim++	VM placement
Higashino et al. [27]	2016	✓						✓			✓			CEPSim *	Latency, response time and accuracy
Kecskemeti [35]	2015	✓							✓		✓			DISSECT-CF	Energy consumption
Bux and Leser [43]	2015	✓						✓			✓			DynamicCloudSim *	Run time
Mishra et al. [56]	2018	✓						✓			✓			CloudSim	Energy consumption
Zakarya and Gillam [64]	2019	✓			✓	✓		✓			✓			CloudSim *	Energy consumption
Kumar and Sharma [57]	2018	✓						✓			✓			CloudSim	Energy consumption, run time and cost
Lin et al. [50]	2017	✓			✓	✓	✓	✓			✓			MultiRE-CloudSim *	Efficiency, power consumption
Arianyan et al. [51]	2017	✓						✓			✓			CloudSim	Energy, SLA and run time
Arianyan et al. [44]	2015	✓						✓			✓			CloudSim	Energy consumption, SLA and VM migration
Fernández-Cerero et al. [58]	2018	✓					✓	✓					✓	SCORE	Energy consumption and task makespan
Magalhães et al. [45]	2015	✓						✓			✓			CloudSim *	accuracy
Priya et al. [65]	2019	✓			✓	✓				✓	✓			CloudSim	Load balance
Madni et al. [68]	2019	✓					✓				✓			CloudSim	Makespan, cost and enhance resource utilisation
Filelis-Papadopoulos et al. [37]	2018	✓						✓				✓		CloudLightning	Energy consumption, task throughput and computational efficiency
Filelis-Papadopoulos et al. [22]	2018	✓						✓				✓		CloudLightning	Energy consumption and Scalability
Sood [54]	2018		✓					✓			✓			CloudSim	Detect deadlocks
Mahmoud et al. [60]	2018		✓					✓			✓			iFogSim *	Energy consumption of the IoT devices
Qayyum et al. [61]	2018		✓					✓			✓			FogNetSim++	Efficient utilization of fog nodes
Talaat et al. [67]	2019		✓							✓	✓			iFogSim *	Average Turnaround Time, Failure Rate, and reliability
Naranjo et al. [62]	2018		✓		✓	✓		✓			✓			iFogSim *	Energy, QoS, and networking delay reduction
Guerrero et al. [69]	2019		✓		✓	✓					✓			iFogSim *	Service placement, service latency and network usage
Gupta et al. [28]	2017		✓					✓			✓			iFogSim *	Cost and latency
AbdElhalim et al. [70]	2019		✓					✓			✓			iFogSim *	Energy consumption and network delay
Pu et al. [59]	2018			✓				✓			✓			ONE	Energy and latency
Filelis-Papadopoulos et al. [66]	2019			✓	✓	✓		✓		✓		✓		CloudLightning	Energy, memory requirement, and scalability

Table 1: Selected articles: Comparative study (* denotes an extension or variant of CloudSim).

3. Discussions and Final Considerations

This paper provides two key contributions to the research community. First, to the best of our knowledge, this is the first methodical review of its kind on the evaluation of resource management mechanisms across the C2T continuum that focuses exclusively on studies using simulation software. Second, it categorises extant research and highlights areas where future research contributions in resource management, (ii) cloud, fog and edge computing and their interaction, and (iii) simulation, are required. Our suggestions for research contributions are based on the idea that every key task in resource provisioning and resource scheduling is important for investigation for the C2T continuum and the future IoT. This review also provides an insight into one part of the simulation, resource management, and cloud, fog, and edge computing research community, and how it is evolving. Finally, it is worth noting that the Internet of Things has the potential to change how society interacts and operates. Therefore, it will be useful for future research efforts to combine insights from other disciplines into these research communities and research agendas.

References

1. Tarkoma, S.; Katasonov, A. Internet of Things Strategic Research Agenda; Finnish Strategic Centre for Science, Technology and Innovation: Las Palmas, Spain, 2011.
2. IDC. Worldwide Global DataSphere IoT Device and Data Forecast 2019–2023. Available online: https://www.idc.com/getdoc.jsp?containerId=US45066919 (accessed on 29 May 2019).
3. Iorga, M.; Goren, N.; Feldman, L.; Barton, R.; Martin, M.; Mahmoudi, C. Fog Computing Conceptual Model; NIST: Gaithersburg, MD, USA, 2018.
4. Loomba, R.; Ellis, K.A.; Forsman, J.; Fowley, F.; Lynn, T.; Svorobej, S.; Willis, P. Optimisation of Edge Networks and Their Distributed Applications; Technical Report; Intel: Leixlip, Ireland, 2020.
5. Hong, C.H.; Varghese, B. Resource management in fog/edge computing: A survey on architectures, infrastructure, and algorithms. ACM Comput. Surv. (CSUR) 2019, 52, 1–37.
6. Buyya, R.; Ranjan, R.; Calheiros, R.N. Modeling and simulation of scalable Cloud computing environments and the CloudSim toolkit: Challenges and opportunities. In Proceedings of the 2009 International Conference on High Performance Computing & Simulation, HPCS'09, Leipzig, Germany, 21–24 June 2009; pp. 1–11.
7. Filelis-Papadopoulos, C.K.; Giannoutakis, K.M.; Gravvanis, G.A.; Kouzinopoulos, C.S.; Makaratzis, A.T.; Tzovaras, D. Simulating Heterogeneous Clouds at Scale. In Heterogeneity, High Performance Computing, Self-Organization and the Cloud; Lynn, T., Morrison, J.P., Kenny, D., Eds.; Palgrave Macmillan: Cham, Switzerland, 2018; pp. 119–150.
8. Singh, S.; Chana, I. A survey on resource scheduling in cloud computing: Issues and challenges. J. Grid Comput. 2016, 14, 217–264.

9. Singh, S.; Chana, I. Cloud resource provisioning: Survey, status and future research directions. Knowl. Inf. Syst. 2016, 49, 1005–1069.
10. Kumar, M.; Sharma, S.; Goel, A.; Singh, S. A comprehensive survey for scheduling techniques in cloud computing. J. Netw. Comput. Appl. 2019, 143, 1–33.
11. Duc, T.L.; Leiva, R.G.; Casari, P.; Östberg, P.O. Machine learning methods for reliable resource provisioning in edge-cloud computing: A survey. ACM Comput. Surv. (CSUR) 2019, 52, 1–39.
12. Ghobaei-Arani, M.; Souri, A.; Rahmanian, A.A. Resource management approaches in fog computing: A comprehensive review. J. Grid Comput. 2019, 18, 1–42.
13. Tian, W.; Xu, M.; Chen, A.; Li, G.; Wang, X.; Chen, Y. Open-source simulators for cloud computing: Comparative study and challenging issues. Simul. Model. Pract. Theory 2015, 58, 239–254.
14. Zhao, W.; Peng, Y.; Xie, F.; Dai, Z. Modeling and simulation of cloud computing: A review. In Proceedings of the 2012 IEEE Asia Pacific cloud computing congress (APCloudCC), Shenzhen, China, 14–17 November 2012; pp. 20–24.
15. Svorobej, S.; Takako Endo, P.; Bendechache, M.; Filelis-Papadopoulos, C.; Giannoutakis, K.M.; Gravvanis, G.A.; Tzovaras, D.; Byrne, J.; Lynn, T. Simulating fog and edge computing scenarios: An overview and research challenges. Future Internet 2019, 11, 55.

Keywords

Simulating Cloud Computing; Simulating Resource Management; Simulating Resource Provisioning; Simulating Resource Scheduling; Cloud-to-Thing Continuum; C2T; Systematic Review; Cloud Computing; Edge Computing; Fog Computing; Simulation Tools