Performance Evaluation of Cloud Computing Simulation Tools

Dhahi Alshammari, Jeremy Singer, Timothy Storer
School of Computing Science
University of Glasgow, UK
e-mail: d.alshammari.1@research.gla.ac.uk, jeremy.singer@glasgow.ac.uk, timothy.storer@glasgow.ac.uk

***Abstract*—Cloud computing experiments can be conducted on simulators before they are tested on real data centre infrastructure. However, the validation of cloud system simulations is not a well-investigated topic in the literature. It can be difficult to select the suitable tools of simulation to predict infrastructure performance due to the variety of existing cloud computing simulators and their varying levels of accuracy. This paper applies an approach to evaluate any cloud simulator through a comparative validation method against micro data centre infrastructure. We have extended earlier work in evaluating cloud computing simulators by applying the method we used to evaluate CloudSim on other tools. We present the level of accuracy for the GreenCloud simulator and Mininet platforms by applying the comparative validation method and sensitivity analysis for the results on actual and simulated performance. According to our findings, GreenCloud currently does not predict the energy consumption for a micro datacenter accurately. On the other hand, Mininet shows reasonable accuracy in modelling the network performance.**

***Keywords*-CloudSim; Mininet; GreenCloud; simulation; performance evaluation**

I. Introduction

Cloud computing infrastructure is composed of distributed systems with complex platform stacks [1], [2]. This complexity results from the existence of numerous abstraction layers of virtualisation technology. Given the high overheads in terms of time and cost, it is not always possible for researchers to deploy their cloud computing experiments on real cloud infrastructures. Therefore, researchers often use *simulators* to model the pertinent aspects of a cloud computing infrastructure, thus establishing a platform on which their experiments can be conducted. Software simulation tools are especially valuable in providing the opportunity to repeat experiments and to control variables with respect to the architectural model of a realistic data centre environment. It is also noteworthy that developers of cloud computing systems may choose from a wide variety of simulators to use in their empirical research [3], [4]. In view of this, it is important to recognize that simulators may have specific features and variants based on target use cases for simulation. Furthermore, there are many limitations and trade-offs when investigating data centre systems issues by means of simulation. This is one explanation for the vast number of cloud simulators available.

Simulators have been used to test the validation of algorithms and the workload efficiency in cloud computing. Use of cloud computing has rapidly increased in many applications for different kinds of deployment. Since it is difficult to measure and test the application performance on a real cloud environment, simulators have been developed to evaluate any algorithm or policy easily [5].

This research is an extension of our previous work on evaluating the level of accuracy for modelling tools such as simulators and emulators that used in cloud computing experiments [6]. In this work, we focus on a simulator called GreenCloud [7] and an emulator called Mininet [8]. We have applied a cross validation method to compare the predicted performance from these tools against actual performance of a Raspberry Pi micro data centre [9].

The rest of the paper provides details about verification of computer system simulations in section II and details about modelling tools in III. Our method for evaluating the selected tools is described in section IV and its results are shown in V. Section VI delivers an overview of the related work and we present our outcomes from the evaluating method in section VII.

II. Verification of Computer System Simulations

As indicated in [10], [11], the verification of computer system simulations is not a well-investigated topic in the literature. To be more precise, specifying particular techniques or algorithms to be applied for any computing simulator to determine the level of accuracy is not adequately explained. The work in [10] and outlines a range of techniques that can be used to verify computer simulation models, and these studies refer to empirical verification techniques which have been applied in this research.

Works in [12], [11] gathered approaches for verifying models of scientific, economic, statistical concepts, which could also be applied in computing models. These works delivered three positions on verification: (a) Rationalism, which is based on the assumption of the system behavior that the simulator can predict; (b) Empiricism, which is based on a comparative study against real work observations; and (c) Positive Economics, which seeks to verify the dependent variable that the model treats rather than the assumption established in the model. The fourth approach that has been delivered is referred to as multi stage verification" (MVS). This approach integrates the three previous approaches, the chief point of value being that it is essential for a computing model to be verified in relation to each approach.

III. Selected Tools to Be Evaluated

A. Mininet

As a network emulator, Mininet [8] runs a collection of end-hosts, switches, routers, and links on a single Linux kernel. Capitalising on the cryptographic network protocol Secure Shell (SSH), hosts on Mininet can be connected as if they were real hosts. Furthermore, Mininet hosts are described in [13] to be able to run real workloads for web servers, a case in point being the HTTP benchmarking tool Apache Bench (ab) [14] is used in our research. It is also noteworthy that Mininet permits the installation of several profiling tools, including iperf and perf, which can be run to provide measurements of the application performance. In view of the way in which Mininet affords the possibility of creating a realistic virtual network that resembles a hardware network owing to several features (or, alternatively, a hardware network that resembles a Mininet network in a variety of ways), we have developed a Python script that models ab. This Python script can be run on Mininet and, following this, the results can be comparatively examined against the actual performance using the Raspberry Pi data centre, with the latter running the actual benchmark. Mininet has been run on a machine with intel Core i5-4690 processor with 16GiB of memory and 500 GB of disk.

B. Greencloud

As a packet-level simulator, the fundamental use case for Greencloud [7] relates to the domain of energy-aware cloud computing data centres. For the most part, the data centres the Greencloud simulator targets are associated with cloud communications as their primary focal point. At the outset, it is important to recognise that the Greencloud simulator represents an extension of the widely-used NS2 network simulator, and it is released by way of the General Public License Agreement. One of the defining strong points of the Greencloud simulator stems from the potential it affords in modelling the level of energy consumption associated with information technology equipment within a data centre. The equipment that Greencloud is effective at modelling in this respect includes computing servers, communication links, and network switches. Furthermore, GreenCloud can be used to develop new techniques for resource allocation and workload scheduling, and it has proven its value in the optimisation of communication protocols and network infrastructures. Regarding the latter two applications, these are closely linked to the efficient power management of data centre infrastructure. In the present research, the authors have fed the features of the Raspberry Pi data centre (which runs MPI) into Greencloud, where the features range from MIPS, algorithms, and execution time. This decision was made because it enables the comparative examination of the simulated results against the real measurements of the servers energy consumption. It should be noted that we focus to those server parts of the simulated results. The rationale for doing so stems from the fact that within our infrastructure, we only measure the energy consumption of Raspberry Pi nodes.

IV. Method to Validate Simulation Tools Based on Micro Data Centre

In addition to drawing on the cross-validation method we applied in a previous study in [6]. the present research implements a similar overall approach to validating the above mentioned modelling tools. One notable point of difference between [6] and the present research is that we here employed a greater number of tools. For example, this study used power consumption monitoring tools for the Raspberry Pi cluster, thereby enabling the comparative examination with Greencloud. In addition, the nodes in our Raspberry Pi Cloud were instrumented with power monitoring tools for the power consumption measurements. In terms of the method applied here, while running Message Passing Interface (MPI) job in conjunction with the round robin resource allocation algorithm, we collected power consumption measurements from the Raspberry Pi nodes for both running time status (namely, ideal and full load). The workloads and the infrastructure configurations and specifications were fed into Greencloud simulators to evaluate the simulated performance.

According to our measurements, fairly relevant results were obtained with respect to the Raspberry Pi measurements reported in the literature, including [15], [16]. However, we took as our focal point the servers power consumption, which meant that we did not measure the total power consumption for the other components in our data centre. Nevertheless, this variable was taken into consideration for the comparative examination with the Greencloud simulator. One of the ideas for a future project on Mininet, as discussed in [17], involved validating the network emulator against real infrastructure. In view of this consideration, by developing a Python script and using it to serve as a ab benchmark [14] (to be run on both Mininet and the Raspberry Pi Cloud), and we take into consideration the requirement of the validation test for Mininet.

V. Results

Our datasets for the actual and predicted performance of all tools followed the normal distribution. Considering this, we determined that using the technique of calculating the root-mean-square error (RMSE) was a sufficient way to demonstrate the degree to which the tools used in this

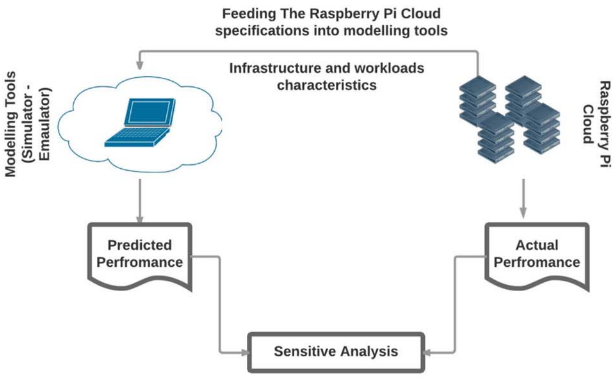

Figure 1. Cross Validation Method

research were accurate [18]. To be more specific, we applied RMSE to compare the predicted performance of the selected tools against the actual performance of the Raspberry Pi cloud. The formula given in V.1 comparatively examines the difference between the values for the actual execution time in $x_{1,t}$ and the predicted performance in $x_{2,t}$. With respect to the ratio of energy consumption for servers in the data centre (i.e., those built on top of Raspberry Pi devices), the RMSE results for Greencloud showed a high relative error when considered in relation to the actual measurements from the Raspberry Pi cloud. For more information, see Figure 2. With respect to the performance of Mininet, we conducted an experiment to model the network performance. The rationale for doing so stemmed from the fact that Mininet focuses on communication rather than computation in experimental situations. Based on the collected data, we confirmed that Mininet was not able to run the ab benchmark with 10,000 concurrent requests in the benchmark. Therefore, we focused on the latency between nodes in running the benchmark, and we sought to calculate the time via iperf on the actual Mininet experiments. The reader is given a representation of the differences between the two experiments (running the ab benchmark with 1,000 concurrent connections) in Figure 3. By applying the sensitive analysis, we present the (RMSE) for the simulated results againest the actual measurements. It shows that GreenCloud gives 49.5 % RMSE Relative Error and mininet gives 27.05 % RMSE Relative Error. Given in Table I are the outcomes from the investigation of the selected tools. Information pertaining to their advantages, drawbacks, and their levels of accuracy when modelling the performance of the Raspberry Pi Cloud is presented.

$$RMSE = \sqrt{\frac{\sum_{t=1}^{n}(x_{1,t} - x_{2,t})^2}{n}} \tag{V.1}$$

VI. Related Work

The DISSECT-CF simulator was one of the notable deliveries to arise as a result of the work reported in [20]. In the study, the simulator was validated against the performance

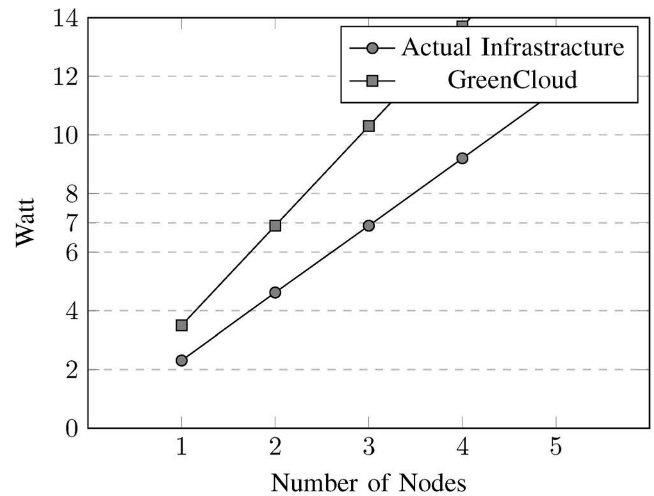

Figure 2. Actual and simulated Power Consumption (Raspberry Pi Cluster vs GreenCloud)

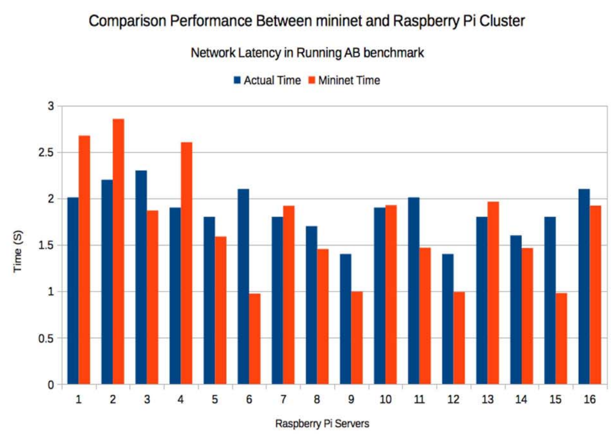

Figure 3. Network Latency Between a Client and Servers (Raspberry Pi Cluster vs Mininet)

of a low power computer cluster, one built on top of Raspberry Pi 2. In view of the experiments, DISSECT-CF was verified on one case of a number of nodes running the workloads, which involved 12 Raspberry Pi 2 devices. One of the limitations of the research highlights a potential area for improvement: in particular, it would have been more advantageous for the work to provide the performance of the simulators on different cluster sizes. This would have illuminated the relationship between the simulator and the actual performance of the low power computing servers.
In the study conducted by [21], an evaluation experiment was conducted for Mininet regarding its emulated performance. Specifically, the evaluation experiment took place with a focus on the comparison between expected theoretical values and Mininets emulated performance. According to the findings reported by the researchers, the reliability of

TABLE I. FEATURES AND ACCURACY RESULTS OF GREENCLOUD AND MININET

-	Area	Advantages	Drawbacks	Actual Experiment	Modelled Experiment	Accuracy
GreenCloud [7]	Energy Consumption	Modelling all communications processes on packet level	Does not consider different architecture for CPU and workloads characteristics.	Running MPI job Data Traffic (DT) from NAS Benchmark [19]. Measuring Real Power consumption by a monitoring tool	Simulating Raspberry Pi MIPS and File specifications on the simulator. Selecting Round-Robin algorithm for resource allocation	49.5 % RMSE Relative Error
Mininet [8]	Network	Running real code including standard Unix and Linux	Limited to the physical resource. Dose not support CPU computations	Running Apache Bench (AB) [14] on Raspberry Pi Cloud. Using iperf tool to profile network performance.	Developing python script involving AB benchmark and iperf commands line to be run on Mininet.	27.05 % RMSE Relative Error

Mininet was attested to with respect to two scenarios: firstly, round-trip delay time; and secondly, throughput. One of the fundamental differences between [21] and the present study is that the latter compares the emulated performance of Mininet against micro servers, namely, clusters of Raspberry Pi devices.

Even a brief consultation of the literature demonstrates the broad variety of ways in which cloud computing simulators have been reviewed and analysed in the past. To give a case in point, studies including [22], [23] and [24] have delivered systematic reviews addressing the existing modelling tools, with these ranging from CloudSim [25] and Greencloud [7] to GroudSim [26] and a host of others. Nevertheless, it is important not to overlook the fact that none of the studies in the extant and related literature have evaluated the selected tools with respect to the issue of their accuracy. In view of this, the present research represents a valuable contribution to the literature, primarily in that it gives an empirical examination and analysis of notable tools, based on the degree to which they are accurate.

VII. CONCLUSION AND FUTURE WORK

The immediate aim of this research has been to illuminate the degree to which a collection of modelling tools can be viewed as accurate. This aim was achieved by designing a methodology that applied sensitivity analysis for the results, specifically relating to the actual and simulated performance of the modelling tools. In conformance with the theoretical tenets of empiricism, the cross-validation method was applied to derive insights about the tools in question, and the collected results indicate that this was a reliable way in which to verify and validate the computer system simulators. The results of our evaluations indicate that *GreenCloud* needs to be improved to predict the energy consumption for micro server architecture, involving the energy consumption for different components in the servers such as CPU, RAM, and buses. This stems from the fact that these components have a significant impact on the energy consumption of the servers in a way that is dependent on the nature of the architectures and the nature of the workload type. On the other hand, *Mininet* was associated with a higher level of accuracy when modelling the network performance for Apache Bench when comparatively examined against *CloudSim* in the previous work in [6] as there is 60% relative error in CloudSim performance. However, an important consideration that should not be overlooked is the fact that this finding is limited to the capacity of the physical machine that it is installed on. This research also provides general limitations and obstacles for modelling tools by validating the predicted performance of micro servers, thus incorporating more features of their architecture to improve the reality of simulating any type of actual data centre. Moreover, our future work will be to predict the performance of cloud computing infrastructure based on the evaluation measurements of our Raspberry Pi infrastructure. This will lead to investigate the level of accuracy of the prediction models based on Machine Learning techniques by comparing the predicted results with the measurements from the tools that have been evaluated in this research. [14]

REFERENCES

[1] M. Armbrust, A. Fox, R. Griffith, A. D. Joseph, R. Katz, A. Konwinski, G. Lee, D. Patterson, A. Rabkin, I. Stoica, and M. Zaharia, "A view of cloud computing," *Commun. ACM*, vol. 53, no. 4, pp. 50–58, Apr. 2010. [Online]. Available: http://doi.acm.org/10.1145/1721654.1721672

[2] P. Mell and T. Grance, "The nist definition of cloud computing," Computer Security Division, Information Technology Laboratory, National Institute of Standards and Technology Gaithersburg, Tech. Rep., 2011.

[3] A. Barker, B. Varghese, J. S. Ward, and I. Sommerville, "Academic cloud computing research: Five pitfalls and five opportunities," in *Proceedings of the 6th USENIX Conference on Hot Topics in Cloud Computing*, ser. HotCloud'14. Berkeley, CA, USA: USENIX Association, 2014, pp. 2–2. [Online]. Available: http://dl.acm.org/citation.cfm?id=2696535.2696537

[4] C. Li and Z. Deng, "Value of cloud computing by the view of information resources," in *Network Computing and Information Security (NCIS), 2011 International Conference on*, vol. 1, May 2011, pp. 108–112.

[5] A. Ahmed and A. Sabyasachi, "Cloud computing simulators: A detailed survey and future direction," in *Advance Computing Conference (IACC), 2014 IEEE International*, Feb 2014, pp. 866–872.

[6] D. Alshammari, J. Singer, and T. Storer, "Does cloudsim accurately model micro datacenters?" in *Cloud Computing (CLOUD), 2017 IEEE 10th International Conference on*. IEEE, 2017, pp. 705–709.

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

[8] M. Team, "Mininet: An instant virtual network on your laptop (or other pc)," 2012.

[9] F. P. Tso, D. White, S. Jouet, J. Singer, and D. Pezaros, "The glasgow raspberry pi cloud: A scale model for cloud computing infrastructures," in *Distributed Computing Systems Workshops (ICDCSW), 2013 IEEE 33rd International Conference on*, July 2013, pp. 108–112.

[10] R. G. Sargent, "Verification and validation of simulation models," in *Proceedings of the 37th conference on Winter simulation*. winter simulation conference, 2005, pp. 130–143.

[11] T. H. Naylor and J. M. Finger, "Verification of computer simulation models," *Management Science*, vol. 14, no. 2, pp. B–92, 1967.

[12] J. P. Kleijnen, "Verification and validation of simulation models," *European Journal of Operational Research*, vol. 82, no. 1, pp. 145 – 162, 1995. [Online]. Available: http://www.sciencedirect.com/science/article/pii/0377221794000166

[13] R. L. S. De Oliveira, A. A. Shinoda, C. M. Schweitzer, and L. R. Prete, "Using mininet for emulation and prototyping software-defined networks," in *Communications and Computing (COLCOM), 2014 IEEE Colombian Conference on*. IEEE, 2014, pp. 1–6.

[14] Apache. (2017) APACHE HTTP SERVER POJECT ab - apache http server benchmarking tool. [Online]. Available: https://httpd.apache.org/docs/2.4/programs/ab.html

[15] G. Bekaroo and A. Santokhee, "Power consumption of the raspberry pi: A comparative analysis," pp. 361–366, 08 2016.

[16] F. Kaup, P. Gottschling, and D. Hausheer, "Powerpi: Measuring and modeling the power consumption of the raspberry pi," in *Local Computer Networks (LCN), 2014 IEEE 39th Conference on*. IEEE, 2014, pp. 236–243.

[17] Mininet. (2014) Ideas preoject ideas for mininet. [Online]. Available: https://github.com/mininet/mininet/wiki/Ideas

[18] T. Chai and R. R. Draxler, "Root mean square error (rmse) or mean absolute error (mae)?–arguments against avoiding rmse in the literature," *Geoscientific Model Development*, vol. 7, no. 3, pp. 1247–1250, 2014.

[19] D. H. Bailey, E. Barszcz, J. T. Barton, D. S. Browning, R. L. Carter, L. Dagum, R. A. Fatoohi, P. O. Frederickson, T. A. Lasinski, R. S. Schreiber *et al.*, "The nas parallel benchmarks," *International Journal of High Performance Computing Applications*, vol. 5, no. 3, pp. 63–73, 1991.

[20] G. Kecskemeti, W. Hajji, and F. P. Tso, "Modelling low power compute clusters for cloud simulation," in *Parallel, Distributed and Network-based Processing (PDP), 2017 25th Euromicro International Conference on*. IEEE, 2017, pp. 39–45.

[21] J. Ortiz, J. Londoño, and F. Novillo, "Evaluation of performance and scalability of mininet in scenarios with large data centers," in *Ecuador Technical Chapters Meeting (ETCM), IEEE*, vol. 1. IEEE, 2016, pp. 1–6.

[22] G. Sakellari and G. Loukas, "A survey of mathematical models, simulation approaches and testbeds used for research in cloud computing," *Simulation Modelling Practice and Theory*, vol. 39, no. Supplement C, pp. 92 – 103, 2013, s.I.Energy efficiency in grids and clouds. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S1569190X13000658

[23] M. I. Alam, M. Pandey, and S. S. Rautaray, "A comprehensive survey on cloud computing," *International Journal of Information Technology and Computer Science (IJITCS)*, vol. 7, no. 2, p. 68, 2015.

[24] R. Malhotra and P. Jain, "Study and comparison of cloudsim simulators in the cloud computing," *The SIJ Transactions on Computer Science Engineering & its Applications*, 2013.

[25] R. N. Calheiros, R. Ranjan, A. Beloglazov, C. A. F. De Rose, and R. Buyya, "Cloudsim: A toolkit for modeling and simulation of cloud computing environments and evaluation of resource provisioning algorithms," *Softw. Pract. Exper.*, vol. 41, no. 1, pp. 23–50, Jan. 2011. [Online]. Available: http://dx.doi.org/10.1002/spe.995

[26] S. Ostermann, K. Plankensteiner, R. Prodan, and T. Fahringer, "Groudsim: An event-based simulation framework for computational grids and clouds," in *Proceedings of the 2010 Conference on Parallel Processing*, ser. Euro-Par 2010. Berlin, Heidelberg: Springer-Verlag, 2011, pp. 305–313. [Online]. Available: http://dl.acm.org/citation.cfm?id=2031978.2032020