Simulation of Large-Scale Network Application Using CloudAnalyst

C. C. Agbo, S. C. Olisa, *C. N. Asiegbu
{chidubem.agbo, samuel.olisa, nnaemeka.asiegbu}@unn.edu.ng
Department of Electronic Engineering, University of Nigeria, Nsukka
*Corresponding author

Abstract— In cloud computing, cloud simulators are employed to characterize and study the different aspects of the cloud computing environment without setting up the actual cloud infrastructure. This is necessary for experimental purposes and research. There are many kinds of cloud simulators, but there is insufficient literature to guide research scholars on how to use these simulators and how to assess them in terms of their suitability for various tasks. In this paper, a cloud simulator, the CloudAnalyst, is used to simulate a large-scale network application to demonstrate its use and to serve as a guide to future users of this cloud simulator.

Keywords—*CloudAnalyst, cloud computing, cloud simulators, performance evaluation, simulation tools.*

I. INTRODUCTION

The cloud computing platform has continued to receive widespread adoption in different application domains because of its cost effectiveness and operational efficiency [1]. Cloud computing, or simply cloud, is a technology that leverages virtualization and grid computing to offer flexible, dynamic, and unlimited computing resources to distributed applications and users. The motivation for the widespread migration to the cloud by both commercial and industrial users can be summed up in the five essential characteristics of cloud computing, namely – on-demand self-service, broad network access, resource pooling, rapid elasticity, and metered service [2].

On-demand self-service means that users of the cloud request and manage their own cloud computing infrastructures and services when and how they want it. Broad network access makes the computing resources available over the internet with broad connectivity options. Resource pooling means that users share from a pool of computing resources, such as data centres. Rapid elasticity and metered service imply that services can be scaled in or out and that the use of services is measurable, and users are billed per the services they consume.

There are different service models available in cloud computing, which include Infrastructure as a Service (IaaS), Platform as a Service (PaaS), Software as a Service (SaaS), and Data as a Service (DaaS) [2]. IaaS involves the outsourcing of the physical hardware resources of the data centre (CPU, Memory, Storage and Bandwidth), giving users the impression of unlimited computing resources. The PaaS model utilizes the services provided by the infrastructure layer to offer platform layer resources, such as operating system support and software development frameworks. In SaaS, a pre-made application, together with any required software, operating system, processing, and storage and network resources is offered as a service. DaaS is a cloud service model where corporate data are housed within a cloud computing environment.

The deployment model of cloud computing is concerned with the way cloud services are designed, deployed, and managed to meet targeted business purposes. There are four deployment models namely private cloud, community cloud, public cloud, and virtual hybrid cloud [3]. Private cloud is a model of cloud computing in which the cloud resources are accessible by only one organization, allowing the organization to have greater control, security, and privacy. Community cloud is established to address the business needs of several organizations that have similar requirements and are willing to share infrastructure to harness some of the benefits of cloud computing. Public cloud describes a model of cloud computing in which computing resources are provisioned and made available on demand to the public through the Internet. Hybrid cloud combines the private and public clouds to harness the benefits of the two models.

To study and experiment with different aspects of the cloud-computing environment, cloud simulators are normally used to simulate the environment without deploying a real cloud platform. The use of cloud simulators is very helpful in the testing and validation of the performance of a cloud application or configuration setting before they are deployed in a real cloud environment. The use of simulators plays an import role in the validation of cloud computing applications and service configurations, as it offers a platform where repeated and controlled testing can be carried out freely. It also makes it possible to identify performance bottlenecks before deployment [4].

There are, however, many cloud simulators available in the market, which makes it difficult for users to decide on the appropriate simulator to use for a given purpose. There is insufficient literature to guide researchers and practitioners on the selection and use of the different cloud simulators. Some authors have tried to classify and evaluate the existing cloud simulators using different criteria, but none of these involves an actual simulation experiment to demonstrate the features of the cloud simulators and how they can be used.

In this paper, these earlier works on the classification and evaluation of the cloud simulators are briefly reviewed, and this is followed by an actual simulation experiment to demonstrate how one of the popular cloud simulators, the CloudAnalyst, can be used to simulate large-scale network applications.

The remainder of this paper is, therefore, organized as follows. Chapter II presents the state of the art in the classification and evaluation of cloud simulators. In chapter III, the simulation results obtained from the use of CloudAnalyst are presented. Chapter IV concludes the paper.

II. STATE OF THE ART

As earlier stated, there are many different cloud simulation tools. These cloud simulators are used to test and evaluate the performance of the cloud infrastructures, services and applications to ensure that they meet the quality of service (QoS) requirements specified in the service level agreement (SLA) before they can be deployed on real cloud and offered as services to consumers. As there are many cloud simulators, there is a need to evaluate and document the capabilities and performance levels of these simulators, so that users can easily select the appropriate simulator for a given task. This section therefore presents a literature review of the state of the art in areas of evaluation and comparison of cloud simulators.

In [5], the authors present three cloud simulators – MDCSim, GreenCloud and CloudSim. The description of the simulators and their basic features are given. The authors go further to compare the three simulators using these criteria: language of coding, open source availability, simulation speed and graphical environment. They conclude by stating that none of the presented simulators was best of all, but that each is designed for specific purpose. Georgia S and George L in [6], consider three different cloud simulators. They classify the simulators on the bases of whether they are used in the analysis of energy efficiency or for performance/QoS in cloud computing environment. They also use a similar set of criteria as in [5] to compare the simulators. These include the programming language the simulator is coded in, availability on the web, the type of license (whether it is open source or not) and a short description of the features of the simulator. Wei Zhao et al. also classify and compare eleven cloud simulators using just three criteria: underlying platform; programming language; and whether the simulator is software-based or a combination of software and hardware [7].

A very interesting paper by Utkah and Mayank [8] provides an overview of fourteen different cloud simulators, and compares them using a set of six criteria, namely GUI support, platform, language, support for TCP/IP, S/W or H/W and availability. The authors conclude the work by stating that every cloud simulator has its pros and cons and that the choice of a simulator depends on the user's requirements. In the comparison presented in [9, 10], eleven different criteria are used to evaluate three different cloud simulators. The criteria used are communication networks, graphical support, availability, platform, simulation time, language/script, physical models, energy models, and support for TCP/IP and power saving modes.

Finally, the paper by Abul [11] uses eight criteria to evaluate ten cloud simulators. The eight criteria are provider, license, category, API, OS, services, popularity and comments. These criteria are defined in Table 1.

Table 1: Definition of the Evaluation Criteria

Criteria	**Description**
Provider	Organization(s) involved in the development of the simulator
License	License requirements (e.g. open source, commercial, proprietary)
Category	The category of the simulation tool (simulation software, testbed, etc.)
API	Type of Application Programming Interface provided in the simulator
OS	The Operating Systems which support the installation of the simulator
Services	The type of Cloud Services supported by the simulator (e.g. IaaS, PaaS, etc.)
Popularity	Number of search results on Google Scholar
Comments	Special notable feature or property of the simulator

From the above, it is seen that none of the evaluation and classification criteria is based on actual simulation experiment. In this paper, the existing work in this domain is extended by performing experiments with the CloudAnalyst [12] to further guide future users of the cloud simulator about its capabilities and features in relation to the simulation of large-scale cloud application.

III. SIMULATION RESULTS OBTAINED FROM THE USE OF CLOUDANALYST

Cloud technology is very beneficial for large scale network applications, e.g., Twitter, Facebook. These applications present a non-uniform usage configuration and access to such service varies based on the time of day and geographical location. Let's consider a large network application; number of requests to the application may also be increased due to new functionality that may be only momentary. Scaling capability of cloud allows it to react to the increase in application request and to scale down when the number of requests has been reduced.

Facebook has over 1500 million of active user worldwide. Distribution of Facebook user across the world: Oceania 60 million, Asia 190 million, Africa 35 million, Europe 500 million, south America 150 million and north America 600 million [13]. So, for this case study, the behaviour of Facebook is modeled and CloudAnalyst is used to evaluate response time, data centre processing time and cost related to the use of cloud.

Simulation Configuration

CloudAnalyst divides the globe into six regions; using this capability, six user groups are defined, one for every region and for simplicity, each userbase is considered in a single time zone. It is assumed that most users use the application in the evening time from 1700 to 0000 local time. It is also assumed that only 10% of registered users are online during peak hours and one tenth of that number of users online during off-peak hours. Each user makes an average of 30 new request an hour.

Hosting application cost should follow amazon EC2 price plan. The assumed plan is cost per 1 GB of data transfer (from / to internet):$0.10; cost per VM per hour (8192 MB, 600 MIPS): $0.10. VM size 100 GB with 8 GB of memory and 10 Gb of bandwidth. Simulated host has x86 architecture, Xen VMM and Linux operating system. Physical machine has 32 GB memory, 8 processors 100 Gb bandwidth and 4000 GB HD and a time-shared policy is used to schedule resources across VMs. Requests are grouped by factor of 100, each user request required 250 instructions to be executed and users are grouped by factor of 100.

Simulated Scenarios for CloudAnalyst

Different scenarios are considered in this case study.

Scenario-1

Evaluation of response and data centre processing time based on data centres at various geographical places and comparison of various load balancing policies using service proximity-based routing service broker policy.

In this scenario, various cases are selected based on load balancing policy. In first case, one data centre with 50 virtual machines are allocated to the application. In the second case two data centres are considered, each having 50 VMs and one data centre in North America and the second one in Europe. In the third case, three data centres are used,with first data centre in North America with 50 VMs, second data centre in Europe with 50 VMs and third data centre in Asia with 50 VMs. The service proximity-based routing service broker policy is selected and the application is run for 24 hours for the following three load balancing algorithms (1) Round Robin (2) Equal spread current execution load (3) Throttled.

From Figure 1, it is observed that average response time is very high when one data centre is used; response time is better, when two data centres are used and much better when three data centres are used. Figure 2 shows data centre processing time; from the result, the data centre processing time is constant. Therefore, the main cause of the delay in the response time is the transmission delay. Transmission delay can be more apparently observed for different userbase response time as shown in Figure 3.

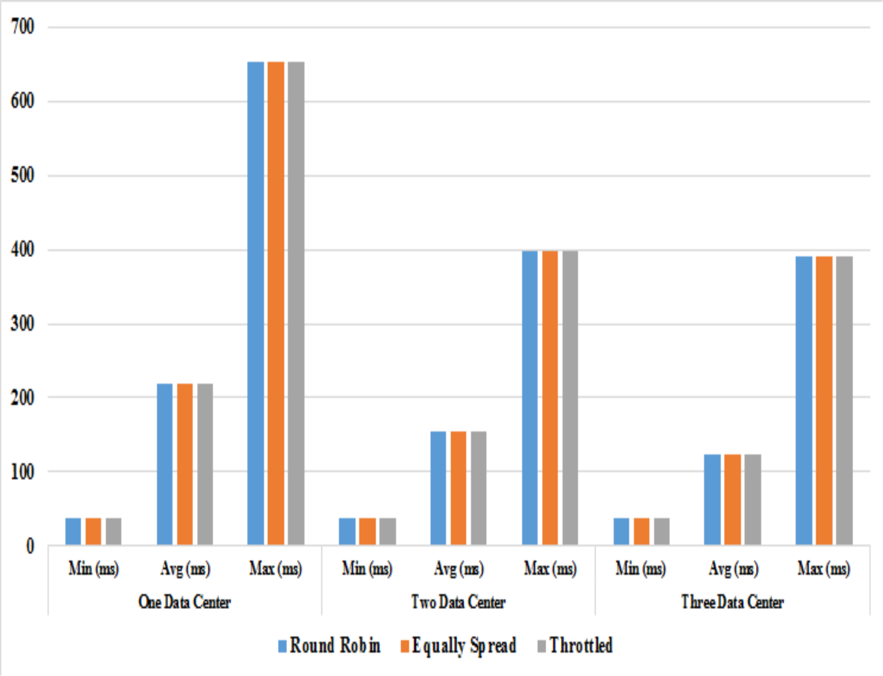

Figure 1: Overall Response Time

Figure 3 shows the effects of size of data center and data centre location on repsonce time for various userbases across the globe. Cost for two data centers is much higher than the cost for three data centers and the main reason for this high cost is the larger amount of data transfer to or from internet. Load balancing algorthims have very small impact on the response time and data centre procesing time, which is neglible. Load balancing algorthim has no impact on total cost of data center. Figure 4 shows cost of data center using different load balancing algorithms.

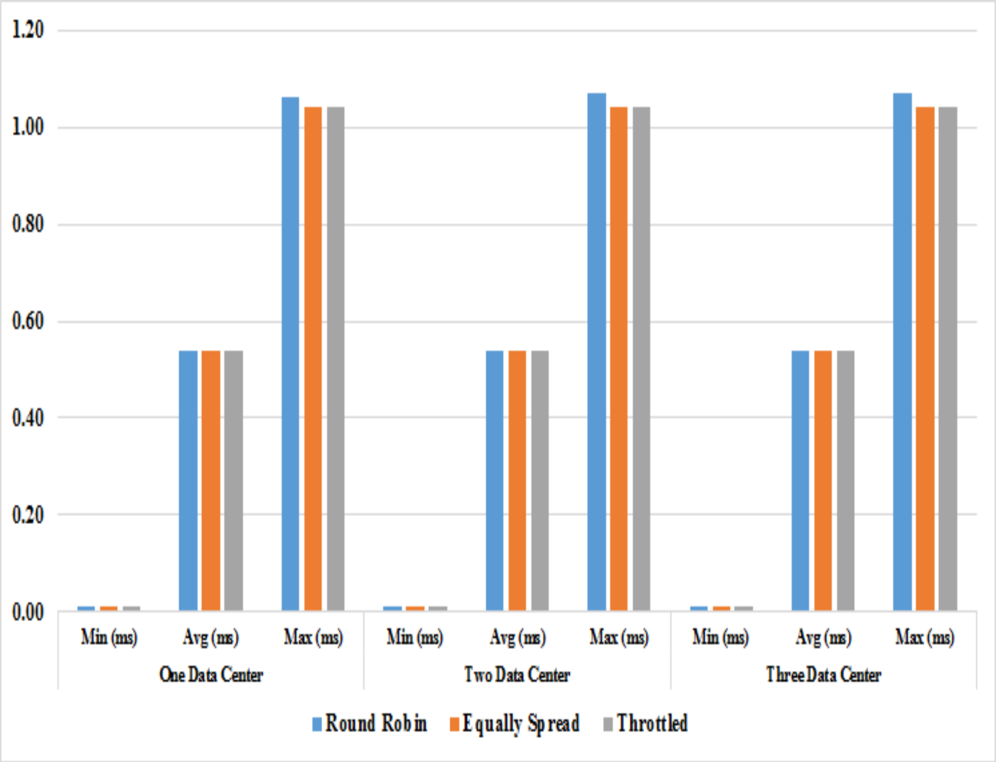

Figure 2: Data Centre Processing Time

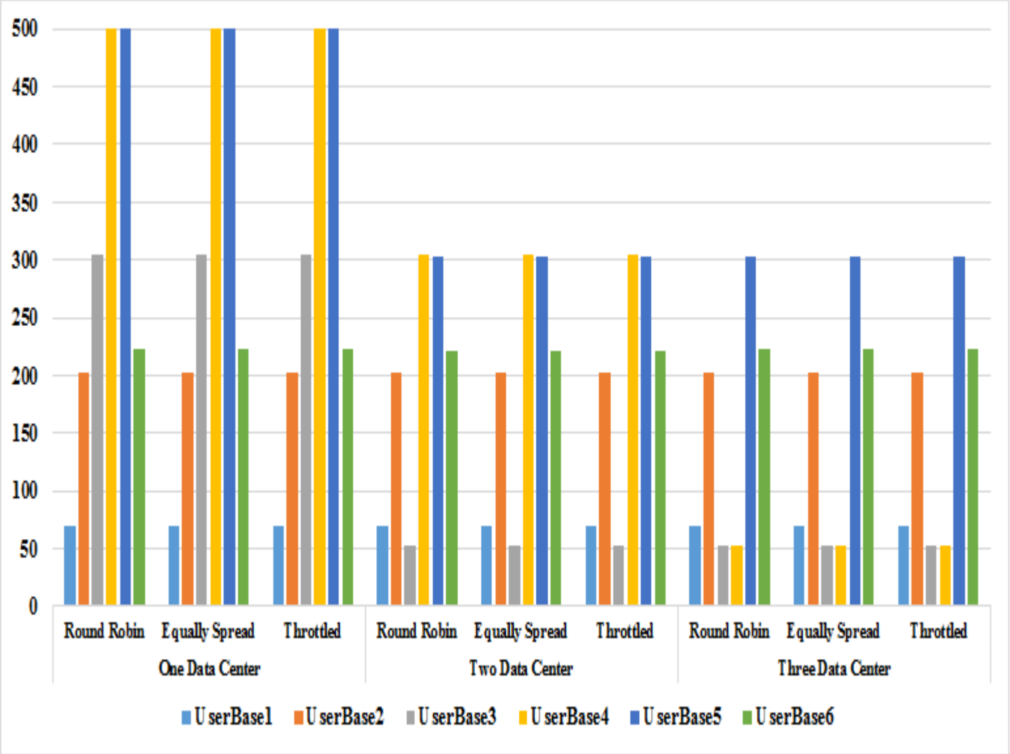

Figure 3: Avg Response Time by Region using different Number of Data Center and service proximity-based routing service broker policy

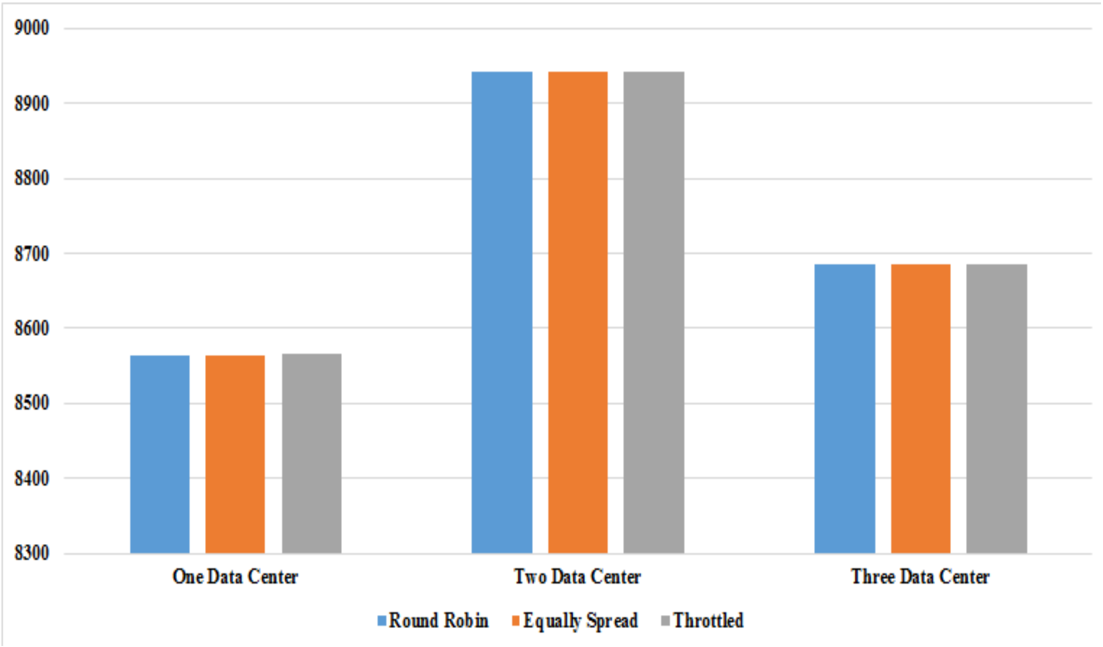

Figure 4: Cost

Scenario-2

Evaluation of response and data centre processing time based on data centres at various geographical places and comparison of various load balancing policies using performance optimized routing service broker policy.

In this scenario, only the service broker policy I changed and by selecting performance optimized routing service broker policy, the application is run for the various load balancing algorithms; repeat same process as for scenario-1. Response time and data centre processing time can be more luminously seen from Figure 5 & 6 respectively.

It is observed from the simulation results that average response time is very high (220 milliseconds) when one data centre was used, response time is better (153 millisecond) when two data centres were used and much better (122 millisecond) when three data centres were used. Data centre processing time remains constant. So the main cause of delay in the response time is the transmission time. Transmission delay can be more apparently observed from different userbase. Figure 7 shows effect of data centre location and capacity on various userbase across the globe.

It is also observed from simulation results that load balancing algorthim has very small impact on the response time and data procesing time which is neglible. Load balancing algorthim has no impact on total cost of data center.

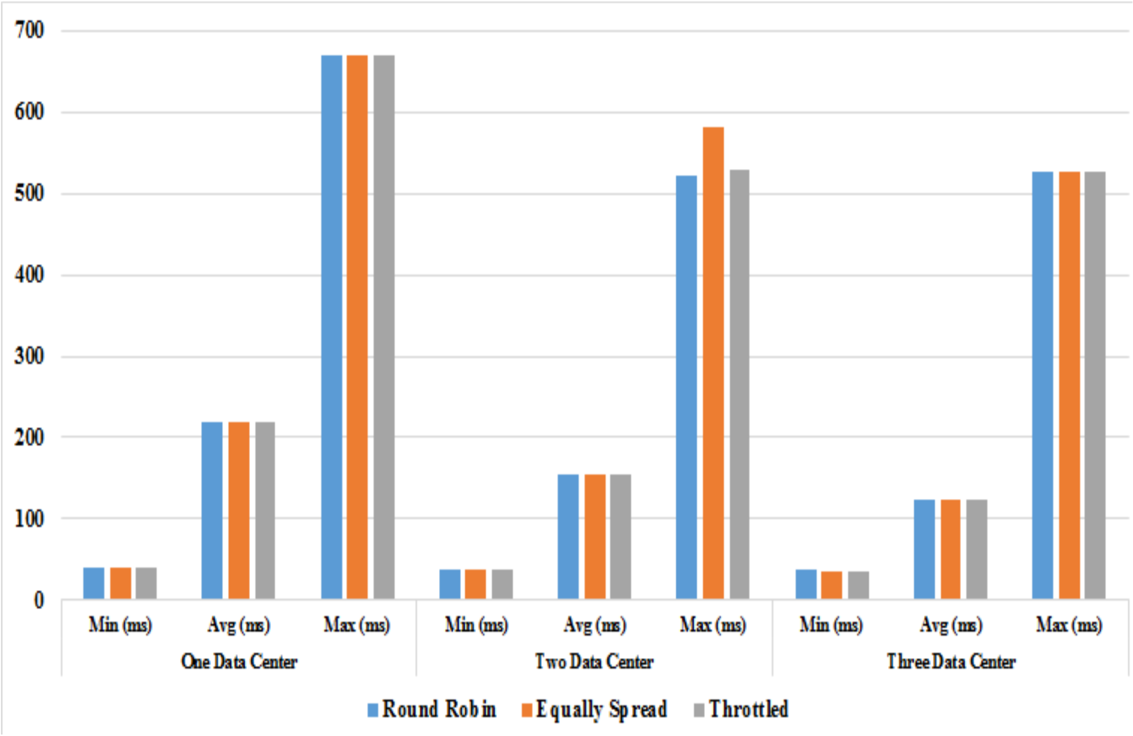

Figure 5: Overall Response Time

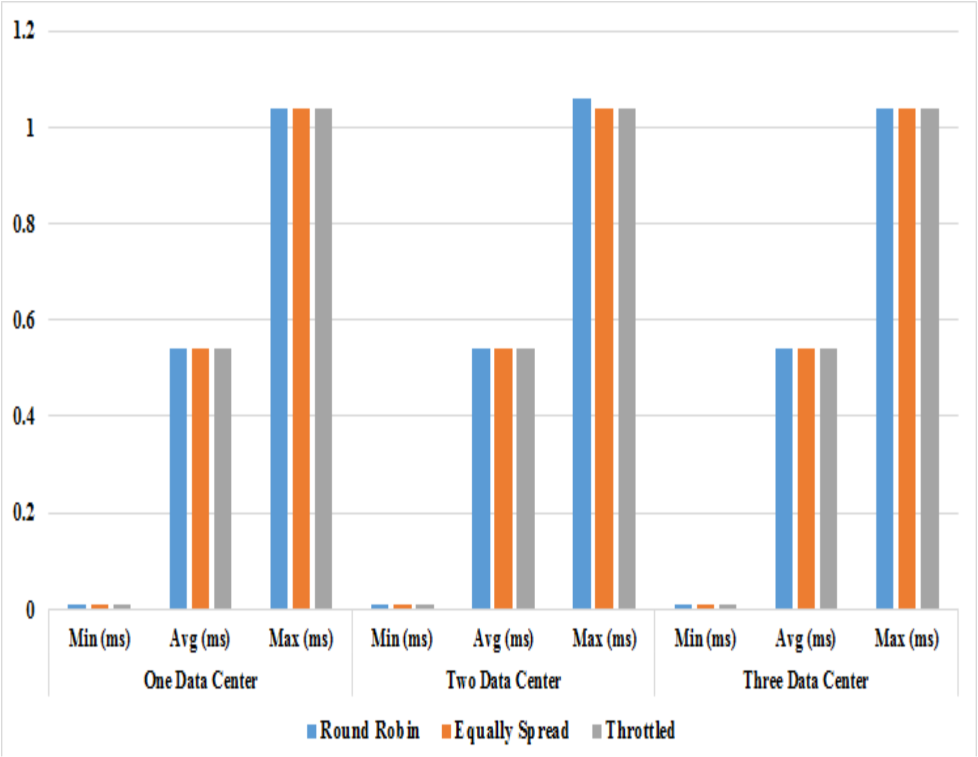

Figure 6: Data Centre Processing Time

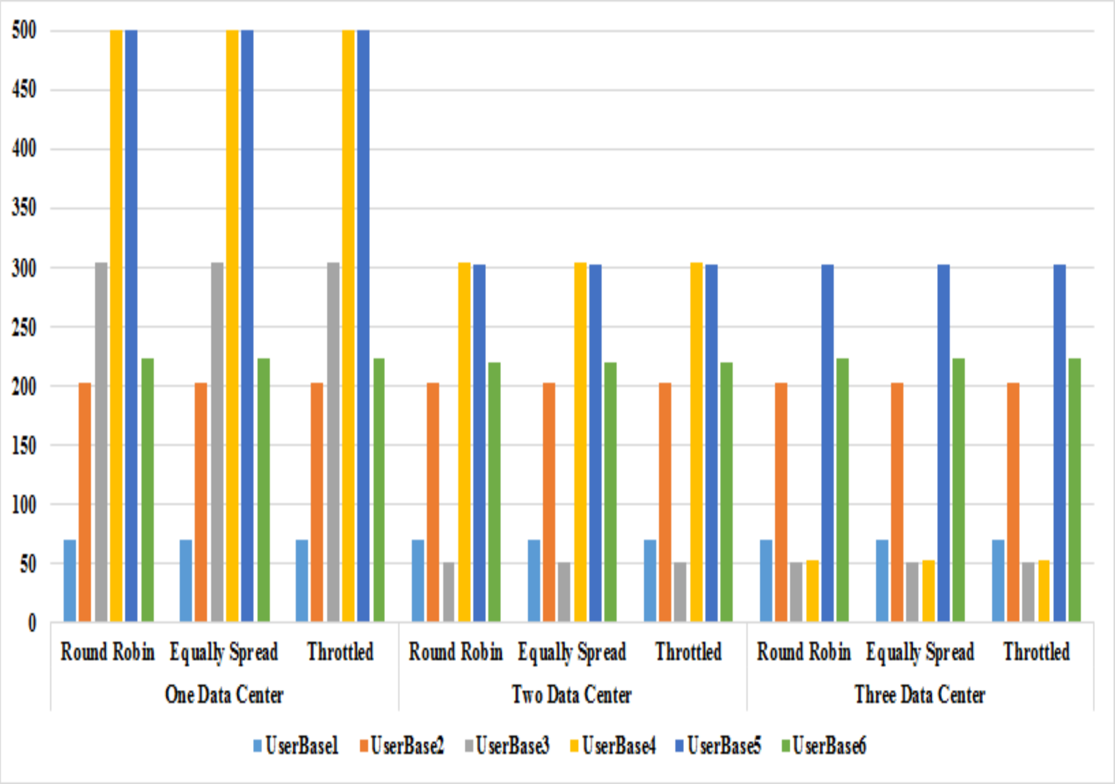

Figure 7: Avg Response Time by Region using different Number of Data Center and performance optimized routing service broker policy

Important Observation from CloudAnalyst Simulation Results

1. Quality of Service can be improved by bringing service closer to the user (i.e. the response time).
2. Quality of Service (i.e. response time in this case) can also be improved by use of load balancing algorithm across data centers.
3. Load balancing also has impact on response time when efficiently used inside a data center.
4. Response time can be improved by using sufficient capacity in the data centers to meet the peak demand.
5. Sufficient data center resources can be allocated to meet peak load throughout; however, this is not economical because there will be time where that capacity is not fully utilized.

IV. CONCLUSION

This paper has demonstrated the use of CloudAnalyst to simulate and study the characteristics of the cloud computing environment for large-scale network application. The results obtained from the simulation will serve as guides to potential users of the cloud simulator.

In future work, other cloud simulators will similarly be evaluated by simulating other relevant cloud computing scenarios.

REFERENCES

[1] D. Hallmans, K. Sandström, T. Nolte and S. Larsson, "Challenges and opportunities when introducing cloud computing into embedded systems," 2015 IEEE 13th International Conference on Industrial Informatics, Cambridge, pp. 454-459, 2015.

[2] A. Shawish, and M. Salama, "Cloud Computing: Paradigms and Technologies," In: Xhafa F., Bessis N. (eds) Inter-cooperative Collective Intelligence: Techniques and Applications. Studies in Computational Intelligence, vol 495. Springer, Berlin, Heidelberg, 2014.

[3] N. Khanghahi and R. Ravanmehr, "Cloud Computing Performance Evaluation: Issues and Challenges," International Journal on Cloud Computing: Services and Architecture, vol. 3, no. 5, October 2013

[4] R. Buyya, R. Ranjan and R. N. Calheiros, "Modeling and simulation of scalable cloud computing environments and the CloudSim toolkit: challenges and opportunities," 7th International Conference on High Performance Computing and Simulation, June 2009.

[5] Mahdi Mollamotalebi, Raheleh Maghami1 and Abdul Samad Ismail, "Grid and Cloud Computing Simulation Tools," International Journal of Networks and Communications, Vols. 3(2): 45-52 DOI: 10.5923/j.ijnc.20130302.02, 2013.

[6] Georgia Sakellari and George Loukas, "survey of mathematical models, simulation approaches and testbeds used for research in cloud computing," Simulation Modelling Practice and Theory, vol. 39, 2013.

[7] Wei Zhao, Yong Peng and Feng Xie, Zhonghua Dai, "Modeling and Simulation of Cloud Computing:A Review," in IEEE Asia Pacific Cloud Computing Congress (APCloudCC), 2012.

[8] Utkal Sinha and Mayank Shekhar, "Comparison of Various Cloud Simulation tools available in Cloud Computing," International Journal of Advanced Research in Computer and Communication Engineering, vol. 4, no. 3, March 2015.

[9] Parveen Kumar and Anjandeep Kaur Rai, "An Overview and Survey of Various Cloud Simulation Tools," Journal of Global Research in Computer Science, vol. Volume 5, no. No. 1, January 2014.

[10] Dzmitry Kliazovich, Pascal Bouvry and Samee Ullah Khan, "GreenCloud: a packet-level simulator of energy-aware cloud computing data centers," J Supercomput DOI 10.1007/s11227-010-0504-1.

[11] C. Yfoulis and A. Gounaris, "Honoring SLAs on cloud computing services: a control perspective," in *Proceedings of European Control Conference*, Budapest, Hungary, 2009.

[12] Arif Ahmed and Abadhan Saumya Sabyasachi, "Cloud Computing Simulators: A Detailed Survey and Future Direction," in *IEEE*, 2014.

[13] Hitesh Marwaha and Dr. Rajeshwar Singh, "A Comprehensive Review of Cloud Computing Simulators," *Journal of Information Sciences and Computing Technologies,* vol. Volume 4, no. Issue 1, June 05, 2015.