International Journal of Applied Engineering Research ISSN 0973-4562 Volume 13, Number 2 (2018) pp. 1139-1143

Experimental Model for Load Balancing in Cloud Computing Using Throttled Algorithm

Gema Ramadhan[1], Tito Waluyo Purboyo[2], Roswan Latuconsina[3]

Research Scholar[1], Lecturer[2,3]

[1,2,3] Computer Engineering, Faculty of Electrical Engineering, Telkom University, Bandung, Indonesia

ORCIDs: [1]0000-0003-2100-904X, [2]0000-0001-9817-3185, [3]0000-0002-8753-6187

ABSTRACT

Cloud Computing is the new generation networking paradigm that appears to provide various services over the internet. Load balancing is a key aspect of cloud computing and avoids situations where some nodes become overloaded while others are idle or busy. Their algorithms are advisable to provide a good mechanism and more efficient algorithm to assign multiple client requests to existing cloud nodes. Results from experiments conducted to determine the level of effectiveness of load balancing algorithm that exists in cloud computing. in this research, know the output based on the overall response time, request servicing, data center loading and cost details in virtual machines from the simulator used.

Keywords: Cloud Computing, Load Balancing, Virtual Machine, Throttled, Cloud Analyst

INTRODUCTION

Cloud Computing system provides a relatively easiest way to save and retrieve data and files. The most important thing if you want to create a huge dataset and the large file size available to prepare the number of users who are scattered throughout the universe [3]. The number of requests of users accessing the server, therefore the impact on the server has a large enough load. So we need a solution to the problem of excessive load on the server so that it can be distributed evenly across multiple servers [1].

Distributed technology in cloud computing systems has so many benefits in terms of giving users needs and applications. the most important thing that clouds computing is to share in terms of software resources, software resources, sharing of information resources through the Internet network, reducing cost-effectiveness, to provide quality of service (QoS) metrics. Including response time, cost, throughput, performance and resource utilization [3].

Therefore, there are several approaches to the method of load balancer in terms of cloud computing. In this experiment, we will discuss load balancing algorithms that are often used in cloud computing. after that, we will do the experiment by using one load balancer technique on a virtual machine in cloud computing.

MATERIAL & METHOD

Simulator

The algorithm is tested in a cloud computing environment. It has two options to test, the first option is to use real tests like Amazon EC2, and the second is to use a simulation tool to simulate the cloud environment. In this experiment, prefers to use a simulator, because using real tests limits experiments on the scale of the infrastructure, and makes reproduction of results a very difficult undertaking [6]. It is also very difficult and time consuming to measure performance in a real cloud environment. In addition, accessing infrastructure actually brings payments in real currency [5].

The simulation framework has the following features [9]:

1. Large-scale support for Cloud computing modeling and installation schemes, including data centers, virtual machines, scheduling, and policy allocation.
2. Support for virtualization services, virtualization, independent, and co-hosted at the data center node.
3. Flexibility to replace the scheme between joint-space allocation and time-sharing.

Cloud Analyst is a Cloudsim-based graphical simulation tool for modeling and analysis the behavior of cloud computing environments, which support visual modeling and large-scale application simulations deployed in Cloud Infrastructure [9].

The main features of Cloud Analyst are as follows:

1. Easy to use Graphical User Interface (GUI).
2. Ability to define simulations with high levels of configurability and flexibility.

Complex system simulations such as internet applications rely heavily on many parameters.

1. The experiment can be repeated.
2. Graphical output.
3. Ease to add extensiom.

Cloud Analysts provide convenience in user location setting schemes, number of users and number of requests per user per hour. And also can do data center location settings, the number

of virtual machines, the number of processors, the amount of storage, network bandwidth and other parameters required.

Method

They are many methods are often used to equalize the queue load on available cloud computing. some of which are presented in this paper, which are:

A. *Round Robin*

The round-robin algorithm is commonly used algorithms in cloud computing for load balancer and is the simplest and easiest to implement. as for several variants of the available round robin algorithm focusing on the measurement of different parameters, and here is the terminology that is familiar with this algorithm [5]:

- Burst Time: BT is the required completion time of the request.
- Time Quantum: TQ is time calculated on request that is allowed to access the VM.

B. *Throttled*

This algorithm uses the system available in the VM status list (BUSY / AVAILABLE.) at the time of the request from the user comes on the load balancer. its checks to find the available VM, after listing the available VM and then allocates to the VM that is available. so these algorithms do not have much of a high degree of complexity in their implementation and their performance is significant [2].

C. *ESCE (Equally Spread Current Execution Load)*

This algorithm works continuously based on the queue and passes it to a different virtual machine randomly and distributes it by checking the size of the load coming, then distributes it to a virtual machine that has light work. [4]. In this experimental paper, the method to be used is from one of the above algorithms (ESCE) which will be used in the CloudAnalyst simulator.

RESULTS AND ANALYSIS

Scenario

In this scenario, for our software using eclipse OXYGEN.1 with Java programming language, and for hardware used in the form of Intel® Core ™ i3 Processor with 6GB RAM we assume the use of the large-scale internet that is on the application, for example Youtube, Facebook, Instagram etc. in this case we assume the use of more than 500 million users worldwide who access the application. And only executed in one Data Center.

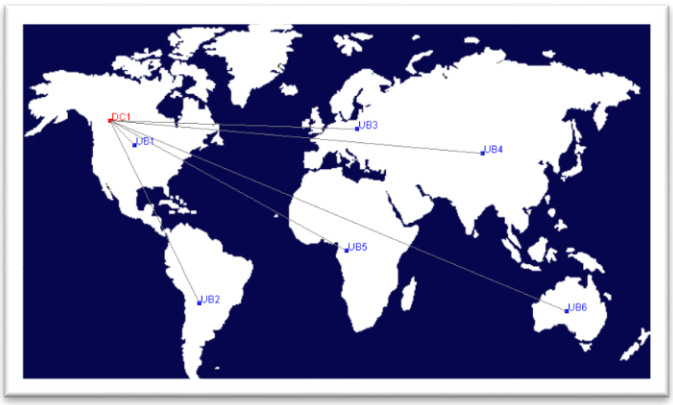

Figure 1: Scenario Cloud Computing

Figure 1 shown the network scenarios to be performed based on one data center used in this experiment and requests from many User Bases. the algorithm to be used in this experiment is *Throttled* on load balancer.

Table 1: Assumptions of the scale of application usage

Region	**CloudAnalyst Region Id**	**Users**
Africa	4	20
Oceania	5	30
Asia	3	80
North America	0	70
South America	1	50
Europe	2	50

Table 1. shows the assumption of user usage based on the geographical location of the user. And in the above is what initials should be input into cloud computing simulator ie CloudAnalyst.

Simulation Configuration

In this scenario, we assume that at the time the clock worked for 2 hours. users use the app at night. And it is also assumed that every 5 minutes the user loads a new request.

Table 2: Parameter on the User Base

User Base	Region	Time Zone	Peak Hours (GMT)	Peak Hours (Local Time)	Online Users Simultaneously During Peak Hour	Online Users Simultaneously During Off-Peak Hours
UB1	0	GMT – 6.00	14:00-16:00	8.00-10.00 pm	500,000	50,000
UB2	1	GMT – 4.00	16:00-18:00	8.00-10.00 pm	200,000	20,000
UB3	2	GMT + 1.00	21:00-23:00	8.00-10.00 pm	400,000	40,000
UB4	3	GMT + 6.00	02:00-04:00	8.00-10.00 pm	250,000	25,000
UB5	4	GMT + 2.00	22:00-24:00	8.00-10.00 pm	150,000	15,000
UB6	5	GMT + 10.00	10:00-12:00	8.00-10.00 pm	180,000	18,000

Table 3: Assumption of Cost

Cost per 1 Gb of Data Transfer (from/to Internet)	$0.11
Cost VM per hour (1024Mb and 100MIPS)	$0.11
Memory Cost	$0.05
Storage Cost	$0.1

Table 2 displays the data on each user base and at what time is the active user accessing the cloud server at the same time.

Table 3. shows In this experiment, we assume an approach with the current cloud service provider service costs such as Amazon EC2. It can affect the performance that will be generated by the virtual machine and the level of satisfaction of the consumer. Because in the case of this simulation we adjust to the actual price specification.

Table 4. shows the assumption at the level of virtual machine specification and physical machine that will be used in this experiment and will be input into cloud computing simulator engine.

Table 4: Physical parameter of Data Center

Parameter of Data Center	The Value Used in The Simulation
VM Image Size	100,000
VM Bandwith	1,000
VM Memory	1024
Architecture of Data Center	X86
OS of Data Center	Linux
VMM of Data Center	Xen
Number of Machines of Data Center	20
Memory per Machine of Data Center	4096
Storage per Machine of Data Center	100,000
Available per BW of Data Center	10,000
Number of Processors per Machine of Data Center	4
Processors Speed of Data Center	100
VM Policy of Data Center	1,000
Grouping Factors on Request	100
Executable Instruction Length	250

Table 5: Assumption value of matrix latency (milliseconds)

Region/Region	0	1	2	3	4	5
0	35	110	160	260	260	110
1	110	35	260	510	360	210
2	160	260	35	160	160	210
3	260	510	160	35	510	510
4	260	360	160	510	35	510
5	110	210	210	510	510	35

Table 5. shows 1 of 2 categories in the matrix parameters of the existing Internet network is the latency of the internet. and will influence the simulation because it is adjusted to the actual state.

Table 6: Assumption value of matrix latency (Mbps)

Region/Region	0	1	2	3	4	5
0	3000	1500	1500	1000	1000	1500
1	1500	700	1000	1000	1500	1000
2	1500	1500	2500	1000	1500	1000
3	1000	1000	1500	2000	1000	1500
4	1500	1000	1000	1500	600	1000
5	1000	1500	1000	1000	1500	2500

Table 6. shows 1 of 2 categories of matrix parameters on the existing Internet network is the bandwidth parameters. and will influence the simulation because it is adjusted to the actual state.

The Results of Experiments

Table 7: Overall response time summary

	Avg (ms)	Min (ms)	Max (ms)
Overall response time	534750.97	564.98	1062618.97
Data Center processing time	534515.92	500.24	1062540.24

Table 7. Shows a summary of the average response time and data processing time of the simulated output used in this experiment. Can be seen from the simulation level efficiency of the algorithm that we use in load balancer to equalize incoming requests from users to the Data Center.

Table 8: Response time based on region

Userbase	Avg (ms)	Min (ms)	Max (ms)
UB1	532309.45	564.98	1062618.97
UB2	533961.01	34633.81	1028050.50
UB3	540151.29	45705.30	1037462.50
UB4	536031.90	49007.53	1043035.70
UB5	526644.95	69799.72	1011642.72
UB6	534195.94	37186.35	1025327.45

Table 8. Shown a summary of the average response time based on each region of simulation output used in this experiment. we can also see the results of the algorithmic efficiency level simulation we use in load balancer to equalize incoming requests from users to the Data Center.

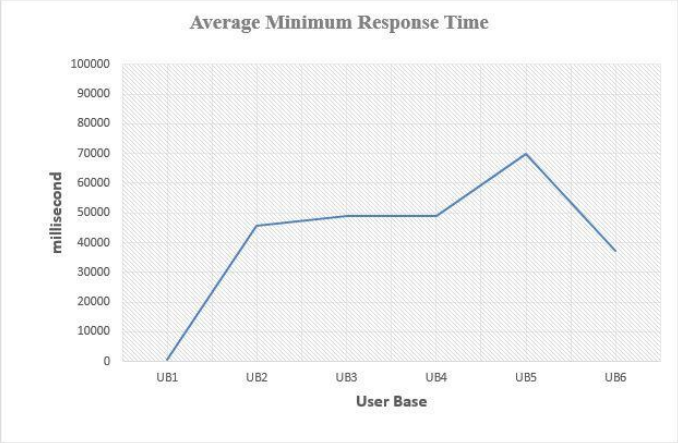

Figure 2: Average response time minimum chart by region

Figure 2. shows the average result of minimum response time by region, it can be said that the minimum time taken on the load balancing algorithm used before reaching the load stage equally in the Data Center. From the graph above can be seen also in each User Base distance based on geographical location affect the response time level.

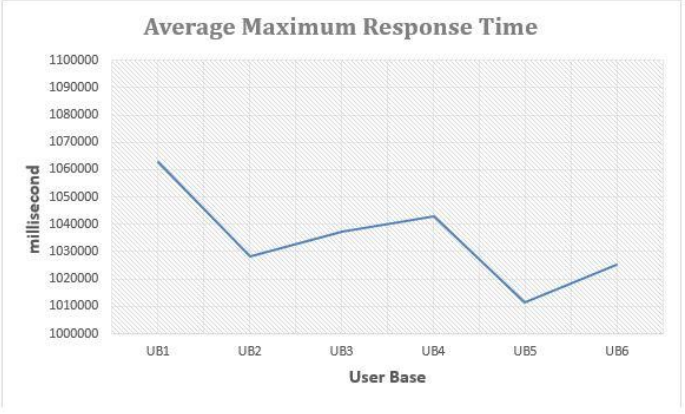

Figure 3: Average response time maximum chart by region

Figure 3. shows the average result of the maximum response time by region, it can be said that the maximum time taken on the load balancing algorithm is used before it reaches the load stage equally in the data center.

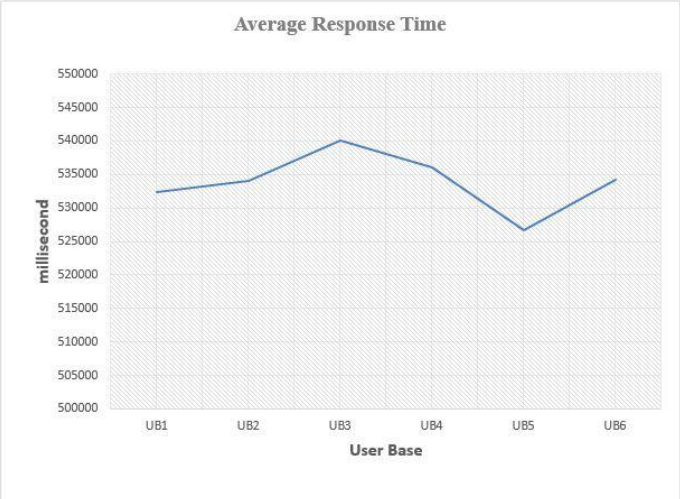

Figure 4: Average response time by region

Figure 4. shows the average result of the overall response time by region, it can be said that the maximum time taken on the load balancing algorithm is used before reaching the same load stage in the data center.

Table 9: Data center request servicing times

Data Center	Avg (ms)	Min (ms)	Max (ms)
DC1	524515.92	500.24	1062540.24

In table 8 can be seen how the data center overcome requests from users based on the average time service. also seen the range of minimum service time up to the maximum service time calculated the average.

Table 10: The Overall cost of the virtual machine ($)

Data Center	VM Cost $	Data Transfer Cost $	Total $
DC1	0.88	60.51	61.39

Table 9. shows the total cost incurred by one Data Center, related to the total amount of costs that is based on virtual machine cost & data transfer cost.

CONCLUSION

In this research, It can be seen that load balancing is a very important aspect on the performance of existing cloud computing. Because the load of request from users is very much and varied. And then from this research, we can conclude that the Throttled algorithm used in load balancer in cloud computing simulations using CloudAnalyst is the average response rate is still within the average range between UB1 and the other. however, in terms of minimum average and maximum response is too far.

REFERENCES

[1]. Maheta, S. D. (2014). Utilizing Round Robin Concept for Load Balancing Algorithm at Virtual Machine Level in Cloud Environment. *International Journal of Computer Application (0975-8887), 94*.

[2]. Klaithem Al Nuaimi, N. M.-J. (2012). A Survey of Load Balancing in Cloud Computing: Challenges and Algorithms.

[3]. Harmandeep Singh Brar, V. T. (n.d.) (2009). A Survey of Load Balancing Algorithms in Cloud Computing. *International Journal of Computer Science Trends and Technology (IJCST), 2*(3).

[4]. Younis, H. J. (2015). *Efficient Load Balancing Algorithm in Cloud Computing.*

[5]. R. Kaur and P. Luthra (2014). *Load Balancing in Cloud Computing.* Recent Trends in Information, Telecommunication and Computing, ITC, Association of Computer Electronics and Electrical Engineers.

[6]. Keith R. Jackson, L. R. (n.d.). Performance Analysis of High-Performance Computing Applications on the Amazon Web Services Cloud. *2nd International Conference on Cloud Computing Technology and Science.* 2010.

[7]. Saxena, S. J. (2016). A Survey of Load Balancing Challenges in Cloud Environment. *5th International Conference on System Modeling & Advancement in Research Trends, 25-27 November.* Moradabad, India.

[8]. Calheiros, R.N., Ranjan, R., De Rose, C.A.F., and Buyya, R., *CloudSim: A Novel Framework for Modeling and Simulation of Cloud Computing Infrastructures and Services*. pp. 1-9, (2009).

[9]. Wickremasinghe, B., Calheiros, R.N., and Buyya, R. *CloudAnalyst: A CloudSim-based Visual Modeller for Analysing Cloud Computing Environments and Applications*. in *24th International Conference on Advanced Information Networking and Applications (AINA)*. IEEE Computer Society, pp. 446-452, (2010).