Cloud Computing: Performance Analysis of Load Balancing Algorithms in Cloud Heterogeneous Environment

Veerawali Behal, Anil Kumar
Department Of Computer Science & Engineering
Guru Nanak Dev University
behalveerawali@yahoo.in, anil.gndu@gmail.com

***Abstract*—Internet is a great marvel of science. Internet has changed the face of world. It is most essential and revolutionary creation in this world. Cloud computing is associated with internet computing. Cloud computing has undoubtedly benefited both service provider and clients in great extent .There is rapid increase in cloud's customers constantly. Although, the cloud data centers comprised of tremendous power but due to expeditious requests of users there is sudden need of balancing load .However, load balancing emerged as the conspicuous issue in the cloud heterogeneous environment. This study highlights the performance analysis of load balancing policies which are taken in a combination with service broker policy . In Fact, this study addresses that there can be reduction in response time and data center request processing time by using efficient load balancing policies. These all evaluations and results are carried out using cloud analyst simulation tool.**

Keywords- Cloud Computing, Throttled, Round Robin, Response Time, Service Broker, Load Balancing.

1. INTRODUCTION

Cloud Computing is new style of computing which is getting progress constantly. Cloud Computing includes computational and storage services as pay you go model. Cloud Computing is very fascinating to business holders as it eradicates the overhead of provisioning and permits the companies to start from small and expand their resources only when there is great rise in demand of services [1][4]. Cloud Computing is probably the only technology that completely complements the internet. Cloud Computing includes two variants of users: Cloud computing customer and Cloud service providers. Cloud computing customers are those which request for particular service and requires their request to be satisfied as early as possible with minimize expenses strategy whereas cloud service providers offers customers storage or software services available for access via the Internet and requires to satisfy number of customer with number of resources and to have maximum return and follows maximize resource utilization and maximize return on investment strategy. Cloud computing is extremely selected by corporate world. However, there are number of issues such as load balancing, cloud security, management of energy, privacy which hinders its growth.

1.1 Load Balancing & Algorithms

Load balancing algorithm deals with balancing the load on server nodes.The main responsibility of the load balancing algorithms is how to select next server node and to transfer new request to that particular node. In Fact, load balancing is a mechanism of lifting the entire load of a whole system to the individual nodes which are idle or having fewer loads so as to have effective usage of resources and to enhance the response time of the task and concurrently discarding the situation in which a few of the nodes are massively loaded whereas rest of nodes are slightly loaded [6][13]. Load Balancing policies for clouds are used by various datacenters to balance load of requests between available virtual machines. There are some existing algorithms which are:

A. Round Robin Algorithm (RR):

Round Robin is one of the traditional widely used algorithms. In round robin policy, the time slices are allotted to each task in uniform proportion and in circular fashion .Each task is allotted to available virtual machine in circular order .This policy is not considered as priority intended scheduling policy. In it, situation occurs where some nodes are massively loaded and some are slightly loaded. This leads to situation where system load gets imbalance [11].

B. Throttled Algorithm:

Throttled algorithm initiates by assigning favorable virtual machine when customer sends request to load balancer .The role of load balancer is to look after an index table of all virtual machine together with their states depicting busy and available mode. At start, all virtual machines are set to available mode. The data center controller consults balancer for next virtual machine allocation, when it receives a new request. The balancer start checking table thoroughly until a relevant match of virtual machine is found. If favorable virtual machine is found then the balancer returns id of that particular virtual machine to data center controller. At that instant, data center controller sends request to virtual machine identified by that particular id. After that, datacenter controller sends notification to the balancer of new allocation so that it can update the table. If there exists a case, when virtual machine is not found, then the balancer returns -1 value and datacenter queues the request. As soon as virtual machine finishes with the processing of the

978-1-4799-4236-7/14/$31.00 ©2014 IEEE

assigned request, later the datacenter controller receives a response cloudlet and it sends the notification to balancer to virtual machine de-allocation[3][10].

The rest of the study is organized as follows. In Section 2, problem statement is given. Section 3, addresses some recent review of similar work in the cloud computing field. Section 4, discusses the proposed work. Section 5, describes the simulation setup along with evaluation of performance and results. Section 6, concludes the paper.

2. PROBLEM STATEMENT

In order to perform load balancing, several algorithms have been proposed to combat this hottest issue. In the case of heterogeneous datacenters which includes virtual machines having different variants of processors, the round robin virtual machine load balancing policy doesn't works very well and causes drastic rise in response time and datacenter request servicing time .By Contrast, the combination of throttled virtual machine load balancing policy and optimize response time service broker policy works very well and causes drastic reduction in response time and datacenter request servicing time. The system performance is examined on the basis of comparative study of different parameters such as datacenter processing time, response time, datacenter hourly loading, and total virtual machine cost, total data transfer cost.

3. REVIEW OF RELATED WORK

To analyze the available facts on the field of cloud computing and load balancing, related work will be conducted using efficient approach. Several studies have been reported that have been focused on load balancing in cloud environment. In this section, numbers of papers are studied which relates to cloud computing and its central concern of load balancing.

Martin Randles, et al. addresses the performance of three dynamic load balancing algorithms, discusses the drawbacks and examined centralized scheduling policy is not feasible in the cloud environment. They also reviewed about distributed solutions suggested for load balancing[9]. Authors have addressed the state-of-the-art of load balancing in cloud environment. They surveyed cloud computing in detail and states classification and examples of its implementation in traditional distributed system [5].Authors exchange views about concept of virtualization, discusses the various interfaces in detail. They also discuss the compatibility concerns of processors which supports ISA with some operating systems [7]. Authors addresses the reasons of cloud adoption and discusses how cloud services are assisting various enterprises[15].Authors describes important notions of cloud computing, services provided by cloud providers, operations provided by the cloud and likewise, reviewed some well-known current load balancing algorithms which can be employed on clouds [3]. Authors have reviewed the architectural design of cloud computing, its benefits and some complications such as security, privacy, authenticity etc. and some of its major applications [2].Authors have discussed the detailed functioning of GUI based tool called as Cloud Analyst which can be used for studying the behaviour of huge scaled internet applications [10]. Author discussed the various scheduling policies and computed comparative study among scheduling algorithms in cloud computing and addresses their need in cloud environment [14].

4. PROPOSED WORK

This study analyzed the comparative study of two virtual machine load balancing algorithms and incorporation of each with service broker policy to determine overall response time, overall data processing time and total cost. In this study, the Round Robin virtual machine load balancing policy is used along with optimize response time service broker policy and it is compared with the throttled load Balancing policy which is also engaged along with optimize response time service broker policy and simulation is performed by adjusting parameters to inspect overall response time, data center request servicing time, response time according to region, data center hourly average processing times , user base hourly response times and total cost which has significant effect on performance.

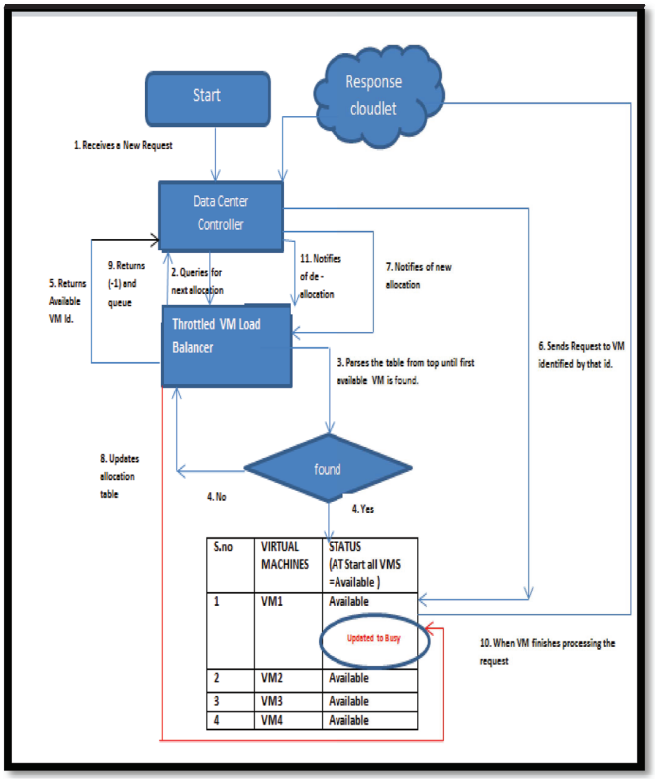

Figure 1: Theoretical Framework

5. SIMULATION SETUP AND PERFORMANCE ANALYSIS

This study make use of well-known tool called as cloud analyst for evaluation of the two renowned algorithms namely round robin, throttled. Simulation organization is done by using four different data centers recognized as DC1, DC2, DC3, DC4 having 100, 75 and 50, 25 numbers of virtual machines respectively. The six user base recognized as USB1, USB2, USB3, USB4, USB5, and USB6 are also included. Each user base is located to different region.

5.1 Cloud Analyst [10] [12]

Cloud Analyst is the GUI based tool .It is an open source toolkit which allows simulating and evaluating the performance of various cloud services .It is an extended version of cloudsim. It is built on the top of cloudsim. Response time and data centre request processing time are the parameters for evaluating performance. Response time is considered to be time taken by an internet application and is defined as delay time between sending of a request and receiving of a response. Several other parameters values are fixed which are shown in Table 1.

Table1: Various Parameters and their values

Parameter	Value Passed
VM-image size	10000
VM-memory	512 MB
VM-bandwidth	1000
Service broker policy	Dynamic Service Broker
Data center architecture	x86
Data center-OS	Linux
Data center -VMM	Xeon
Data center –No. of VMs	DC1-100 ,DC2-75 DC3-50,DC4-25
Data center-memory per machine	2 GB
Data center–storage per machine	1 TB
Data center-available bandwidth per machine	1000000
Data center –processor speed	10000
Data center –VM Policy	Time shared
User grouping factor	1000
Request grouping factor	250
Executable instruction length	100

It includes three essential facilities options .These options are utilize for framing the whole simulation process. Simulation setup and study of results are accomplished for duration of 120 hrs. This study examines comparison analysis of two load balancing policies when optimize response time service broker policy is taken .The simulation is being configured by setting various parameters as shown in figures 2, 3, 4, 5 for case 1 and as shown in figures 8, 9, 10, 11 for case 2. This study takes into account two cases which are follows:

CASE-1: VMs are having equal number of Processors:
In this case, Quad core processors are embedded within each virtual machine of every datacenter.

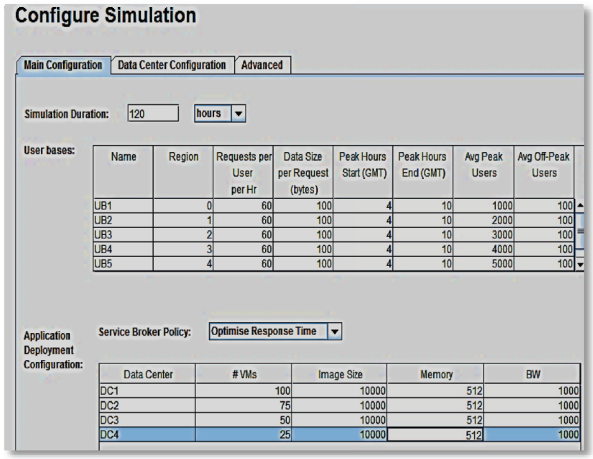

Figure 2: Setting Main Configuration

Figure 2, 8 depicts configuration simulation screen which includes three configuration options on main tab namely simulation time, user bases table application deployment configuration. Simulation time option defines the duration of simulation which can be allotted in mins, hrs. or days. User bases table contains list of all the user bases along with their regions in the simulation and Application deployment configuration defines the table which contains the data regarding number of virtual machines allocated for the application in each data center. Moreover, details of virtual machine are also included.

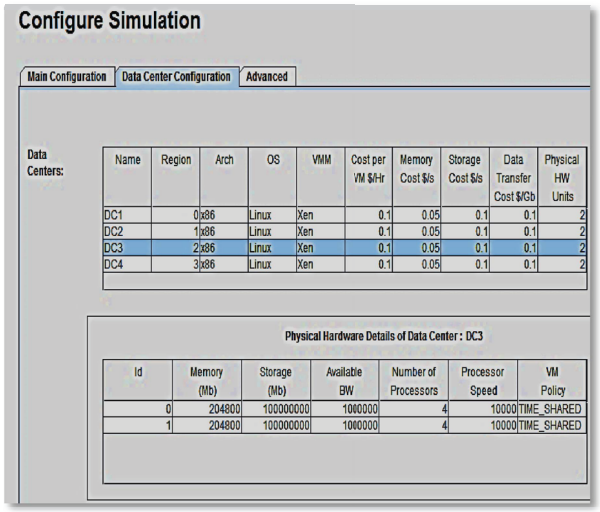

Figure 3: Data Center Configuration

Figures. 3, 9 depicts configuration simulation screen which includes data center tab which permits user to define the whole configuration of the datacenter. Figures. 4,5,10,11 depicts the advance tab which comprised of some essential parameters such as user grouping factor, request grouping factor, executable instruction length along with load balancing policy.

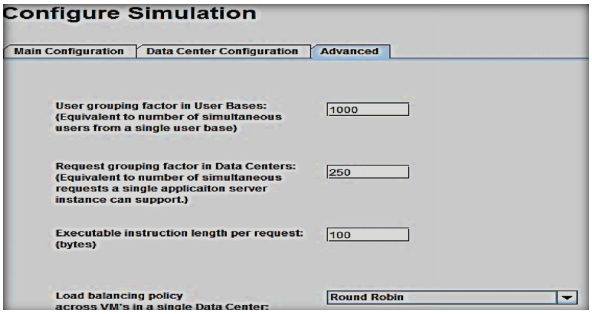

Figure 4: Advance Settings

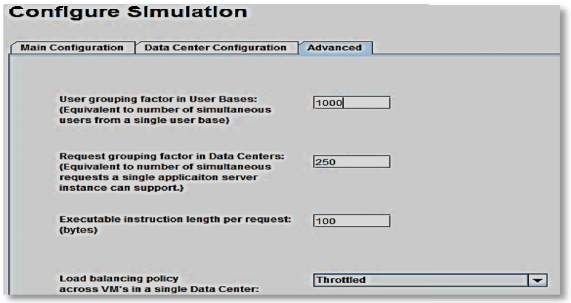

Figure 5: Advance Settings

5.2 Experimental Results

After performing the simulation the results computed by cloud analyst are as shown in the figures 6 and 7 and tables 2, 3 and 4 for case 1 and likewise results are computed for case 2 and are shown in figures 12, 13 and tables 5, 6, 7. The result is determined for the parameters like response time, data center request servicing time and cost in accomplishing the request by using the above specified configuration for each load balancing policy again and again.

Table 2: Comparison of load balancing policies (Average response time)

UBS	Round Robin	Throttled
UB1	63.07	63.07
UB2	60.54	60.64
UB3	57.34	57.34
UB4	53.86	53.96
UB5	308.25	308.26
UB6	216.85	216.85

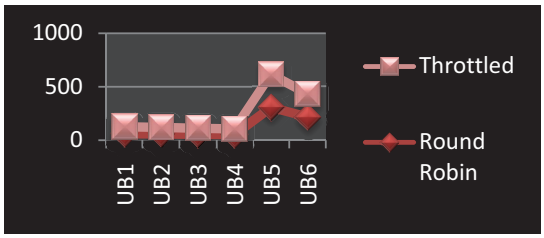

Figure 6: Analysis showing the average peak users vs. average response time

Table 3: Average data center request servicing time among LB policies

Data Center	Round Robin	Throttled
DC1	15.25	15.25
DC2	10.85	10.96
DC3	8.09	8.09
DC4	4.22	4.32

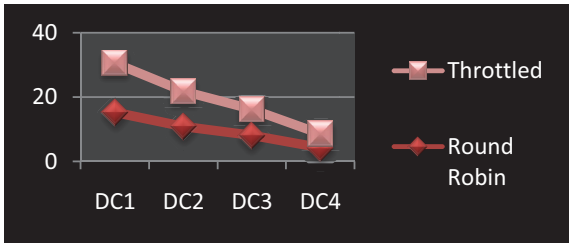

Figure 7: Analysis by taking average data center serving time and data centers

Table 4: Comparison of load balancing policies (Overall response time, Data Center Processing time, Total Cost)

Parameter	Round Robin	Throttled
Overall Response Time	150.06	150.10
Data Center Processing Time	10.10	10.13
Total Cost	$3111.30	$3111.30

CASE-2: VMs having unequal number of processors:
In this case, DC1 is having the combination of Quad core and dual core processors, while DC2 is having only Quad core processors, DC3 have Quad core processors and DC4 is having Dual core, Quad core and Hexa core processors.

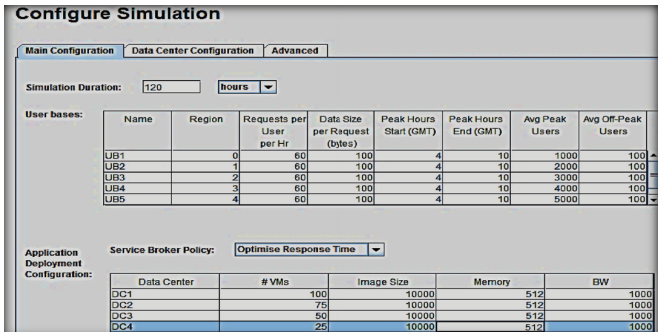

Figure 8: Setting Main Configuration

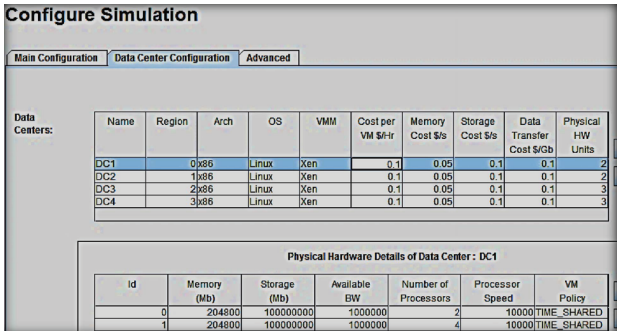

Figure 9: Data Center Configuration

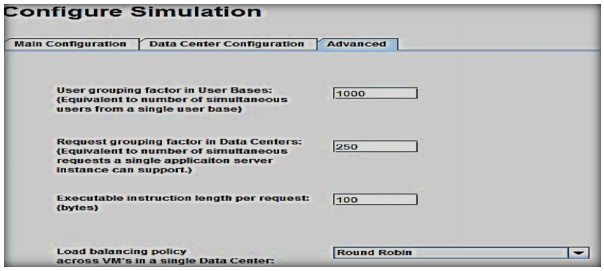

Figure 10: Advance Settings

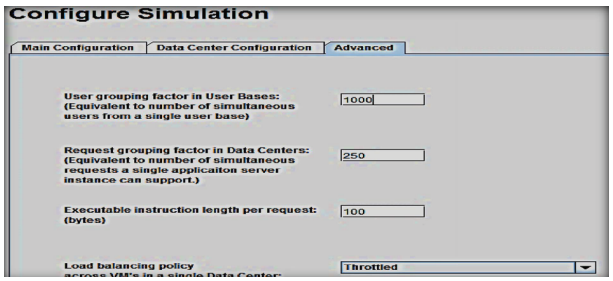

Figure 11: Advance Settings

Table 5: Comparison of load balancing policies (Average response time)

UBS	Round Robin	Throttled
UB1	74.20	63.44
UB2	60.53	60.64
UB3	54.84	54.91
UB4	52.64	52.12
UB5	305.47	305.54
UB6	226.07	217.43

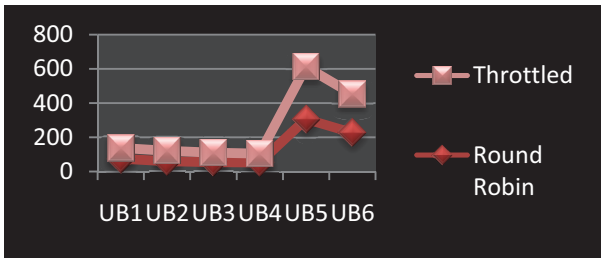

Figure 12: Analysis showing the average peak users vs. average response time

Table 6: Average data center request servicing time among LB policies

Data Center	Round Robin	THROTTLED
DC1	25.30	15.72
DC2	10.85	10.96
DC3	5.45	5.51
DC4	2.99	2.48

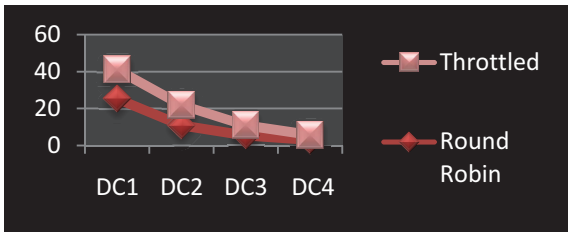

Figure 13: Analysis by taking average data center serving time and data centers

Table 7: Comparison of load balancing policies (Overall response time, Data Center Processing time, Total Cost)

Parameter	Round Robin	Throttled
Overall Response Time	152.11	148.99
Data Center Processing Time	12.25	9.01
Total Cost	$3111.30	$3111.30

After performing simulation ,it is analyzed that although, round robin policy works better in homogeneous environment i.e. when each data center is equipped with identical quad core processors as in above case 1 . However, in practical situation each cloud data center is equipped with number of heterogeneous processors. As a result ,round robin does not fit to the situation whereas proposed strategy of throttled algorithm when taken in combination with optimize response time , as in case 2 works more better and shows drastic reduction in overall response time and data center request servicing time. The comparative analysis of overall response time and data center request servicing time of both cases are shown in figure 4 and 7 respectively.

6. CONCLUSION

While arriving at a conclusion, this research work has proposed a novel strategy to analyses the behavior of two different scheduling algorithms in cloud environment. Both algorithms are taken into consideration and their scheduling criteria like average response time, data center request service time and total cost of different data centers are verified. According to the experiment results and performance analysis, the proposed strategy of throttled load balancing when taken in a combination with optimize response time service broker policy has the best integrate performance in heterogeneous cloud environment. In the near future, this research will be extending for evaluating the performance in real time environment.

7. REFERENCES

[1] Qi Zhang · Lu Cheng · Raouf Boutaba, "Cloud computing: state-of-the-art and research challenges", The Brazilian Computer Society, Springer, 2010.

[2] Y.Jadeja and K. Modi," Cloud Computing - Concepts, Architecture and Challenges", International Conference on Computing, Electronics and Electrical Technologies [ICCEET] 2012.

[3] Ram Prasad Padhy and PGoutam Prasad Rao ,"Load balancing in cloud computing system", Department of Computer Science and Engineering National Institute of Technology, Rourkela Rourkela-769 008, Orissa, India May, 2011.

[4] M. D. Dikaiakos et al. "Cloud Computing: Distributed Internet Computing for IT and Scientific Research", IEEE Journal of Internet Computing, Vol. 13, No. 5, September/October 2009, pages 10-13.

[5] Khiyaita et al." Load Balancing Cloud Computing: State Of Art", IEEE, 2010.

[6] R.Shimonski. Windows 2000 & Windows server 2003 clustering and load balancing. Emeryville. McGraw-Hill Professional publishing, CA, USA (2003), p 2, 2003.

[7] J E Smith and Ravi Nayar, "Introduction to virtual Machines", Elsevier Science, Nov 14, 2004.

[8] Jaspreet Kaur, "Comparison Load Balancing Algorithms In A Cloud", International Journal Of Engineering Research And Applications, 2012.

[9] Martin Randles et al."A Comparative Study into Distributed Load Balancing Algorithms for Cloud Computing", IEEE 24th International Conference on Advanced Information Networking and Applications Workshops, 2010.

[10] Bhathiya, Wickremasinghe,"Cloud Analyst: A Cloud Sim-based Visual Modeller for Analysing Cloud Computing Environments and Applications", IEEE, 2010.

[11] P. Goyal et al., "An optimized round robin scheduling algorithm for CPU scheduling", International journal of computer and electrical engineering (IJCEE), vol. 2, No. 7, pp. 2383- 2385, December, 2010.

[12] Tanvee Ahmed, Yogendra Singh, "Analytic Study of Load Balancing Techniques Using Tool Cloud Analyst". International Journal of Engineering Research and Applications. pp. 1027-1030, 2012.

[13] Rich Lee, Bingchiang Jeng, "Load-Balancing Tactics in Cloud", International Conference on Cyber-Enabled Distributed Computing and Knowledge Discovery, 2011.

[14] Isam Azawi Mohialdeen,"Comparative Study of Scheduling Algorithms in Cloud Computing Environment", Journal of Computer Science, 2013.

[15] Bhushan Lal Sahu and Rajesh Tiwari, "A Comprehensive Study on Cloud Computing", International Journal of Advanced Research in Computer Science and Software Engineering, Volume 2, Issue 9, September 2012.