

# A Comparative Study of Load Balancing Algorithms in Cloud Computing

Reena Panwar  
M.Tech CSE Scholar  
Department of CSE,  
Galgotias College of Engineering and  
Technology, Greater Noida, India

Bhawna Mallick, Ph.D  
Professor  
Department of CSE,  
Galgotias College of Engineering and Technology,  
Greater Noida, India

## ABSTRACT

Nowadays, Cloud computing has become buzzword in the Information Technology and is a next stage in the evolution of Internet. It provides very large amount of computing and storage services to users through the internet. Load Balancing is important for essential operations in cloud virtual environments. As cloud computing has been growing rapidly and many clients all over the world are demanding more services and better results, so load balancing is an essential and important research area. Many algorithms have developed for allocating client's requests to available remote nodes. Efficient load balancing ensures efficient resource utilization of resources to customers on demand basis and enhanced the overall performance of the Cloud. This paper is a brief discussion on the existing load balancing techniques in cloud computing and further compares them based on various parameters like data processing time and response time etc. The paper analyzes the result based on existing Round Robin and Throttled scheduling algorithms.

## General Terms

Cloud Computing, Cloud Environment, Load Balancing, Virtual Machines, Resource allocation.

## Keyword

Cloud Computing, Cloud Environment, Load Balancing, Virtual Machines, Resource allocation.

## 1. INTRODUCTION

Cloud computing is the fastest developed technology in the IT industry and a new delivery method for the services on pay per use basis. According to various standard definitions of cloud computing: "Cloud Computing is the delivery of application software, platform and infrastructure as a service over the Internet accessible from the web browser, as well as software and data are residing on the servers on pay per use basis. Cloud computing provides services by internet.

Cloud Computing has become one of the popular technology adopted by both industry academia providing a flexible and efficient way to store and retrieve files [2]. The main problem is to schedule the incoming request in a way so with minimum response time, efficient resource utilization and at the same time resources should not be underutilized. Cloud Computing system are heavily rely on term virtualization that improves the power efficiency of datacenters and enables virtual machines to single physical server.

Cloud computing is an internet based computing. It delivers all services through the internet dynamically when user demands, such as operating system, network, storage,

software, hardware and resources. These services are classified into these types: Infrastructure as a Service (IaaS), Platform as a Service (PaaS) and Software as a Service (SaaS). Cloud computing domain is divided into three categories such as Public, Hybrid and Private cloud. Cloud computing framework as shown in Fig. 1 depicts that it has four deployment models and three service models which is as follows.

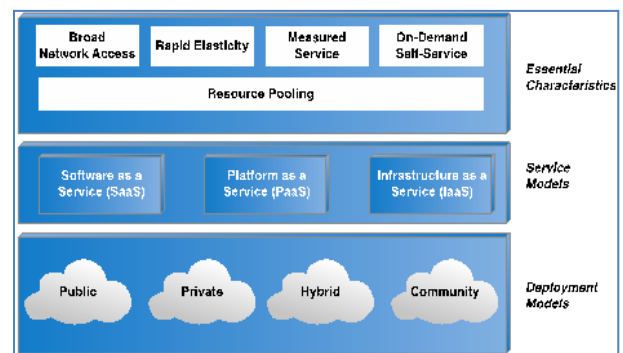


Fig 1: Cloud Computing Framework [2]

### 1.1 Infrastructure as a service (IaaS)

It is the most basic service model of cloud known as Infrastructure as a Service (IaaS). It offers hardware as a service to the organization such as Server space, Network equipment, CPU cycles, physical virtual machines and other resources.

### 1.2 Software as a Service (SaaS)

In SaaS, the various services are provided to the user such as software applications from different cloud providers through the Internet. The customer uses software on a pay-per use basis. It is referred as "on-demand software".

### 1.3 Platform as a Service (PaaS)

In PaaS models, Cloud service providers deliver a computing platform such as operating system support, programming language execution and web server. Examples are GAE and Microsoft Azure[3].

**Public Cloud:** It can be access by anyone from anywhere all over the world. Examples of this cloud are Google's cloud which is open for all after specific SLA between provider and user. It is available on a subscription basis.

**Private Cloud:** In this type of cloud company's employee can access to company's own data or a partner employment.

**Hybrid Cloud:** It is a combination of both public as well as private cloud.

## **2. LOAD BALANCING IN CLOUD COMPUTING**

Load balancing is a new technique that provides high resource time and effective resource utilization by assigning the total load among the various cloud nodes[15], side by side it solves the problem of overutilization and underutilization of virtual machines. Load balancing resolve problem of overloading and focuses on maximum throughput, optimizing resource utilization and minimize response time. Load balancing is the pre requirements for maximizing the cloud performance and utilizing the resources efficiently.

In utilization of clouds has been improved by a resource allocation method having preemptable task execution. Adaptive resource allocation algorithm is presented for cloud system with preemptable tasks but this approach does not resolve the problem of response time and effective cost utilization.

In [14], M. Moradi, M.A. Dezfuli, M.H.Safavi, has proposed a New Time Optimizing Probabilistic Load Balancing Algorithm. This algorithm is proposed to establish load balancing and decrease the response time and main purpose is to choose the resources based on least completion time and better past status.

Load balancing has two major tasks, one is the resource allocation or resource provisioning and other is scheduling in distributed environment. Efficient provisioning of resources and scheduling of resources as well as tasks will ensure:

- Resources are available easily.
- Resources are efficiently utilized under condition of low/high load.
- Reduction in cost of using resources.
- Load balancing increases throughput to maximum level and minimum response time.

The load balancing is an efficient and critical concept in cloud computing and it helps in utilizing the resources optimally, therefore minimizing the consumption of resources. Thus load needs to be distributed over the nodes in cloud-based architecture, so that each resource does the equal amount of work at any point of time that is performed by a load balancer. The load balancer determines the various request allocation to different servers. The load balancer uses various scheduling algorithm which server should take the request.

## **3. CLASSIFICATION OF LOAD BALANCING ALGORITHM**

In cloud computing , different load balancing algorithm have been proposed such as Honeybee-based load balancing technique, Random Sampling, Active Clustering, Throttled Load Balancer Active Monitoring Load Balancer, WCAP, JIQ, CLBVM. The main purpose is to achieve high throughput and minimum response time.

Generally load balancing algorithm is of two types:

- A. Static load balancing algorithm
- B. Dynamic load balancing algorithm

### **3.1 Static Load Balancing Algorithm**

The load does not depend on the current state of the system but it requires knowledge about the application and resources of the system. Static load balancing is a load balancing algorithms that distributes the workload based strictly on a fixed set of rules such as input workload.

There are four different types of Static load balancing techniques: Round Robin algorithm, Central Manager Algorithm, Threshold algorithm and randomized algorithm.

### **3.2 Dynamic Load Balancing Algorithm**

Dynamic algorithm are more flexible than the static algorithm and they doesn't rely on prior knowledge but depends on current state of the system. In a distributed system, dynamic load balancing has two different ways: distributed and non-distributed. In the distributed one, this algorithm is executed by all nodes present in the system and the task of load balancing is shared among these servers. The interaction among nodes to achieve load balancing can take two forms: cooperative and non-cooperative. In the first one, the nodes works side-by-side to achieve a common objective which means is to improve the overall response time, etc. In the second form, each node works independently toward a goal local to it.

Here we will discuss two algorithms:

#### **3.2.1 Round Robin Algorithm**

It is one of the simplest scheduling algorithms that utilize the principle of time slices. Here time is divided into multiple slices and each node is given a particular time interval. Each node is given a quantum and in this given quantum node has to perform its operations. If the user request completes within time quantum then user should not wait otherwise user have to wait for its next slot. It means that this algorithm selects the load randomly, while in some case some server is heavily loaded or someone is lightly loaded.

#### **3.2.2 Throttled Load Balancing Algorithm**

This algorithm is totally based on the allocation of request to virtual machine. Here client will first request the load balancer to check the right virtual machine which access that load easily and performs the operations request by client or user. In this algorithm the load balancer maintains an index table of virtual machines as well as their states (Available or Busy)[11]. Therefore the client first requests the load balancer to find a suitable Virtual Machine to perform the required operations. These dynamic algorithms are being experimentally performed using the cloud analyst tool which gives the output with respect to virtual machine.

## **4. INTRODUCTION TO CLOUD ANALYST**

The simulation of problem, result and performance analysis is carried out by using this simulator cloud analyst tool.

### **4.1 Cloud Analyst**

Cloud Analyst tool is a GUI based tool based on Cloud Sim architecture. It is developed at the University of Melbourne whose goal is to support evaluation of social networks tools according to geographic distribution of users and data centers. CloudSim is a framework developed by the GRIDS laboratory of University of Melbourne which enables modelling, simulation and experimenting on designing Cloud computing infrastructures[16]. CloudSim enables modeling and simulation of Cloud Infrastructure.

Cloud Analyst separates the simulation experimentation exercise from a programming, such as modeler has to focus on the simulation complexities. The Cloud Analyst also provides modeler to operate over a series of simulation experiments with a little bit change in respective parameters in easy way.

## 4.2 Cloud Analyst Design

The Cloud Analyst is built on top of Cloud Sim tool kit, by extending Cloud Sim functionality with the introduction of concepts that model Internet and Internet Application behaviours.

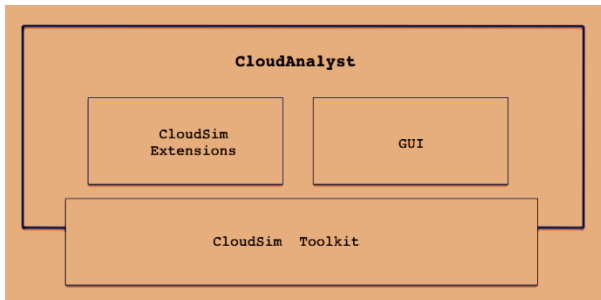


Fig 2: Architecture of Cloud Components

### 4.2.1 The basic components of cloud analyst are [4]

- **GUI Packages:** It is mainly responsible for the graphical user Interface.
- **User Base -** A User Base models a group of users that is considered as a single unit in the simulation and generates traffic for the simulation.
- **Internet-** It models the Internet traffic routing around the globe by introducing transmission and data transfer delays.
- **DataCenterController-**It generally controls Data center activities.
- **VmLoadBalancer-** The Data center controller generally uses a VmLoadBalancer to determine which VM should be assigned the next Cloudlet for processing and models load balancing policies
- **Simulation-**This is used for creating and executing the request.
- **CloudAppServiceBroker –** This component model service brokers that handle traffic routing between user bases and data centers.

Now we will analyse the various load balancing policies by setting the configurations of the various components of the cloud analyst tool. We have set the parameters for the application deployment, user base configuration, Data center configuration and load balancing policy as shown in Figs 3-6 respectively. As shown in the Fig 2, we have been defined the locations of user base in six different regions of the world. We have taken four data centers to handle the requests of these users. One data center is located in region 2, second in region 1, third in region 4 and fourth in region 3. There are total 20 VMs in each of DC1 to DC4. As shown in Fig 6, you can select a load balancing policy, RR algorithm is selected. The duration of simulation is 24hrs. Cloud analyst enables the modeller to execute the simulation repeatedly with change of the parameters time to time. The graphical output of the

simulation results of these figures can be seen more easily and efficiently.

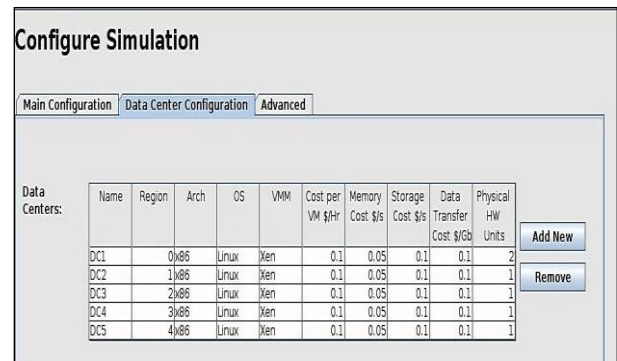


Fig 3: Data Center Configuration

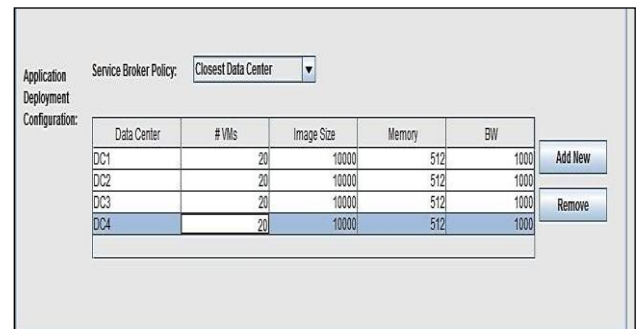


Fig 4: Application Deployment Configuration

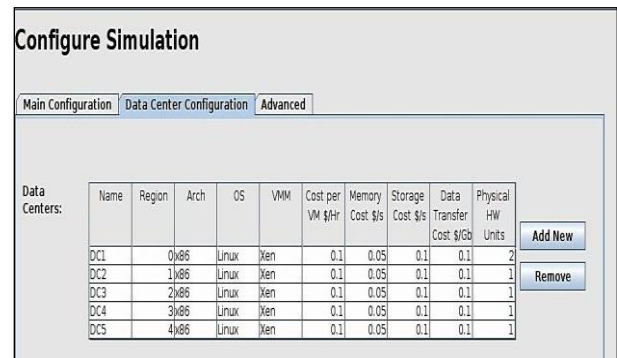


Fig 5: User Base Configuration

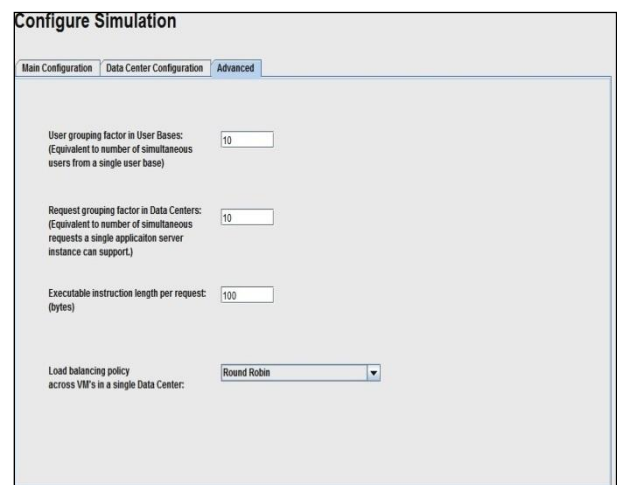


Fig 6: Load Balancing Policy

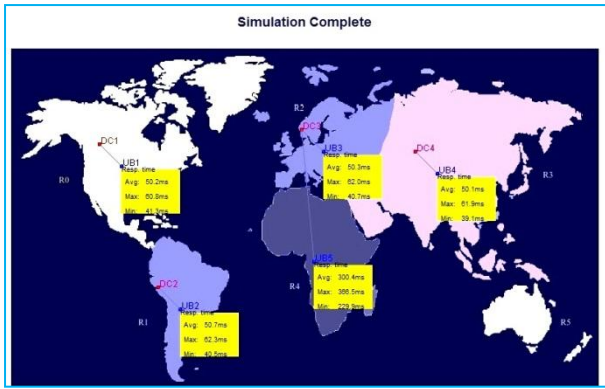


Fig 7: Output Screen of Cloud Analyst

Here after simulation the result computed by cloud analyst is as shown in the following Fig6. We have used the above defined configuration for each load balancing policy one by one and depending on that the result calculated for the metrics like request processing time, response time and cost in fulfilling the request has been shown in Tables I, II, III.

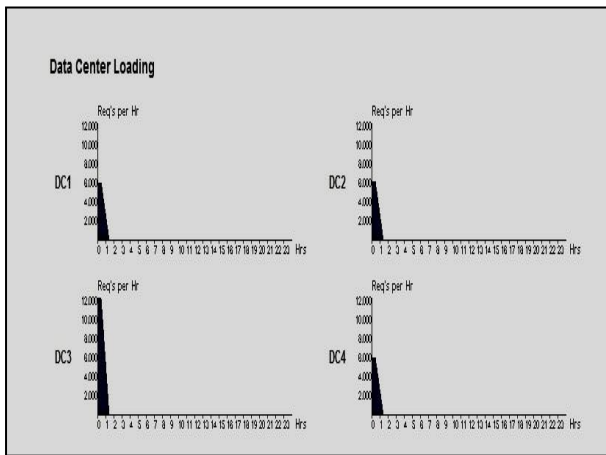


Figure 8: Data Center Hourly Loading

### 4.3 Response Time by Region

Response time all user base and overall response time are calculated by the Cloud Analyst tool for each loading policy has been shown in the Table 1, Fig 8, 9.

Table 1. Response Time By Region

User Base	Round Robin (ms)	Throttled(ms)
UB1	50.194	50.129
UB2	50.722	50.716
UB3	50.361	50.364
UB4	50.097	50.18
UB5	300.346	300.275

### 4.4 Datacenter Request Servicing Times

Data center Request Servicing Time for each data center is calculated by the cloud analyst for each loading policy has been shown in the Table 2.

Table 2. Data Center Request Servicing Times

Data Center	Round Robin (ms)	Throttled (ms)
DC1	.618	.613
DC2	.856	.853
DC3	.782	.782
DC4	.495	.504

### 4.4 Cost

The data transferring cost for each load balancing policy computed by the cloud analyst as can be seen from the Table3.

Table 3. Date Transfer Cost

Data Center	Round Robin (\$)	Throttled (\$)
DC1	0.064	0.064
DC2	0.065	0.065
DC3	0.128	.128
DC4	0.502	0.063

### 4.5 Overall Response Time

The overall responding time for each algorithms calculated by the cloud analyst as can be seen in the Fig 8, 9.

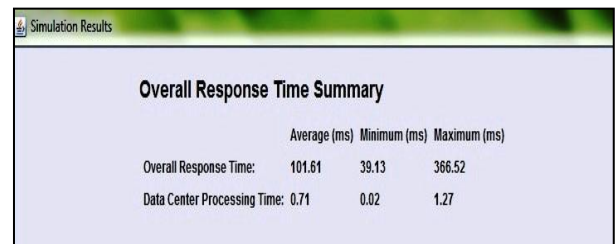


Fig 9: Overall Response Time Summary of RR

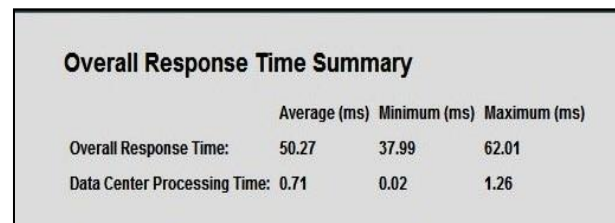


Fig 10: Overall Response Time Summary of Throttled

## 5. CONCLUSION

Load balancing is a main task in cloud computing for efficient utilization of resources. The main goal of load balancing is to achieve higher client satisfaction, maximize resource utilization and increase the performance of the cloud system thereby reduction in the energy consumption and carbon

emission rate. Nowadays Cost and Time are key challenges for every IT engineer to develop products that can enhance the performance of business in the cloud. The performances of two algorithms Round Robin and Throttled have been discussed in this paper. The requested time for both the techniques applied are same which means there is no effect in request time of data centers after changing the algorithms. The cost analysis for each algorithm is calculated with the help of experiments shown in the respective figure. The cost calculated for virtual machine usage per hour is not same for two algorithms, Here Throttled Load balancing algorithm reduce the cost of usage, therefore Throttled Load balancing technique is more efficient in terms of cost for load balancing on cloud data centers.

## 6. REFERENCES

- [1] Ali M Alakeel, "A Guide To Dynamic Load Balancing In Distributed Computer Systems", International Journal of Computer Science and Network Security, Vol. 10 No. 6, pp. 153-160, 2010.
- [2] Anthony T.Velte, Toby J.Velte, Robert Elsenpeter, Cloud Computing A Practical Approach, TATA McGRAW-HILL Edition 2010.
- [3] [http://en.wikipedia.org/wiki/Cloud\\_computing](http://en.wikipedia.org/wiki/Cloud_computing).
- [4] Hamid Shoja, Hossein Nahid "A Comparative Survey On Load Balancing Algorithm In Cloud Computing" IEEE 33044.
- [5] Subasish Mohapatra, K. Smruti Rekha, Subhadarshini Mohanty "Comparison of Four Popular Heuristics for Load Balancing of Virtual Machines in Cloud Computing", International Journal of Computer Applications, Volume 68– No.6, April 2013.
- [6] Radojevic, B & Zagar, M, "Analysis of issues with load balancing algorithms in cloud environment proceedings of 34th International Convention on MIPRO, IEEE
- [7] Nidhi Jain Kansal and Inderveer Chana, "Cloud Load Balancing Techniques: A Step toward Green Computing", International Journal of Computer Science Issues, Vol.9, No.1, 2012 pp.238-246.
- [8] Sandeep Sharma, Sarabjeet Singh, Meenakshi Sharma, "Performance Analysis of Load Balancing Algorithms", World Academy of Science, Engineering and Technology, 2008.
- [9] T. Kokilavani and D. Amalarethnam, "Load balanced min-min algorithm for static meta-task scheduling in grid computing" International Journal Computer Applications, vol. 20, no.2, 2011.
- [10] Z. Bo, G. Ji, and A. Jieqing, "Cloud loading balance algorithm," in Information Science and Engineering (ICISE), 2nd International Conference on, pp. 5001–5004, IEEE, 2010.
- [11] Shridhar G. Domanal and G. Ram Mohana Reddy, "Load Balancing in Cloud Computing Using Modified Throttled Algorithm", IEEE Randles, M. Bendiab, A. T. & Lamb, D. (2008).
- [12] D.A. Menasc'e and P. Ngo, "Understanding Cloud Computing: Experimentation and capacity planning," in Computer Measurement Group Conference, 2009.
- [13] G. Khanna, K. Beaty, G. Kar, and A. Kochut, "Application Performance Management in Virtualized Server Environment," in Network Operations and Management Symposium, (2006). NOMS (2006). 10th IEEE/IFIP, pp 373–381, 2006.
- [14] M. Moradi, M.A. Dezfouli, M.H. Safavi, Department of Computer and IT, Engineering, Amirkabir University of Technology, Tehran, Iran, "A New Time Optimizing Probabilistic Load Balancing Algorithm in Grid Computing" IEEE Vol.1, pp. 232-237, 2010.
- [15] Load Balancing computing [EB/OL], <http://en.wikipedia.org>.
- [16] W. Bhatiya, 2009. CloudAnalyst a CloudSim-based tool for modelling and analysis of large scale cloud computing environments. MEDC Project, Computing and Distributed Systems Laboratory, University of Melbourne, Australia, pp. 1-44, June 2009.