



## Load Balancing Algorithms in Cloud Computing: A Review

D. Saranya, L. Sankara Maheswari

Assistant Professor,

Department of Computer Science,

Sri G.V.G Visalakshi College for Women,

Udumalpet, Tamil Nadu, India

---

**Abstract:** *In the current scenario important focus is on the development load balancing algorithms for cloud computing. In this paper, different kinds of load balancing algorithms are analyzed for cloud computing, to make the network load equally divided in order to provide faster connectivity for all the devices that need the service. Among implemented load balancing algorithms, ant colony optimization algorithm has achieved a better performance, when compared to others. Load balancing algorithms play a challenging task in cloud computing.*

**Keywords:** *Cloud, Load balancing, Servers, Nodes, Resources*

---

### I. INTRODUCTION

Cloud computing is an evolving area that allows users to organize applications with enhanced scalability, availability and fault tolerance. Cloud computing provides internet based platform that is used for computer technology. It describes a diversity of computing concepts [8]. Cloud computing accumulates all the computing resources and manages them automatically. Nowadays world depends on cloud computing to store the public as well as personal information. Cloud computing provides relevant hardware, software and service according to the requirement that users put forward. A cloud computing structure is categorized by its on-need self service, access over internet, pooling of resources, elasticity of service availability and measurement of services utilized by individual users. Cloud computing provides a collective group of resources, including data storage space, networks, computer processing power and specialized corporate and user application. There are four deployment models in cloud computing. They are

- ✓ Public
- ✓ Private
- ✓ Community and
- ✓ Hybrid.

#### *Cloud Services*

Cloud computing provides a number of clouds for providing services. Services can be larger or smaller, and use of a service is measured and customers are billed accordingly [9]. Service Models are the orientation models on which cloud computing is based. These can be categorized into three basic service models as listed below:

- ✓ Infrastructure as a Service (IaaS)
- ✓ Platform as a Service (PaaS) and
- ✓ Software as a Service (SaaS).

### II. LOAD BALANCING

Load balancing is a method for reassigning the entire workload to the separate nodes for a group of systems, in order to make the time more efficient and utilize the resources in efficient manner [1]. Load balancing is the basic technique for scaling out an application server infrastructure. As request demand increases, new servers can be added to the resource pool, where the load balancer will directly begin sending traffic to the new server.

There are two different kinds of algorithm in load balancing. They are static and dynamic algorithm.

#### **A. Static Algorithms**

Static algorithms partition the traffic uniformly among the servers. In this approach, the traffic on the servers will be contemptibly easy and subsequently it will make the circumstances imperfectly [2].

#### **B. Dynamic Algorithms**

A dynamic algorithm selects the appropriate weights on server, by searching through the whole network and prefers the lightest server to balance the traffic. But selecting suitable server needs valid communication within networks that lead to extra traffic being added on system [2].

### C. Benefits of Load Balancing [3]

#### Redundancy

It describes the process of running two or more, the same servers thus providing a guaranteed event that one server becomes occupied.

#### Scalability

Even though modest resources requirements are offered, scalability must always be considered for finding the correct host solution.

#### Resource Optimization

Through load balancing, one can optimize how traffic is circulated to the server cluster so, that it guarantees the best performance.

#### Security

In security, only one IP is exposed to the web with load balancing, which significantly reduces the amount of break points in case of attack.

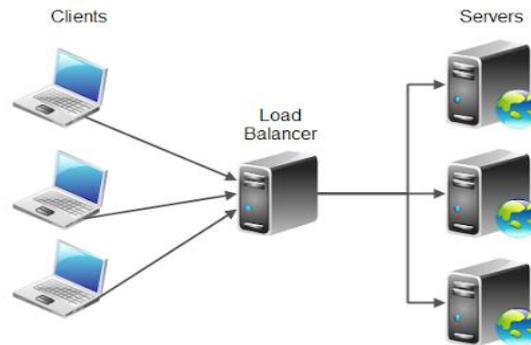


Figure: 2-Example of Load Balancing

### D. Load balancing in cloud computing

Load balancing in cloud is a method, which distributes the overloaded, active local workloads uniformly, across all the nodes. It achieves good utilization of resources and better user satisfaction, efficiently improving the overall performance of the system. When the tasks are arrived from different locations the load balancer receives them and distributes to data centre for appropriate load distribution. Vendors in cloud are usually based on automatic load balancing service, which allows entities to add the number of CPUs for their resources to scale with the increased demands.

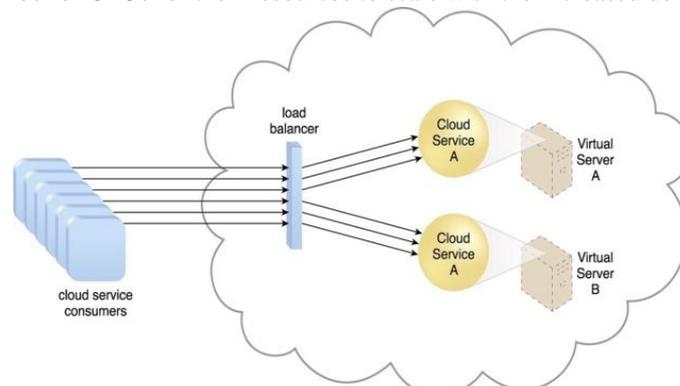


Figure: 3- Example of Load Balancing in Cloud Computing

## III. REVIEW

### A. Round Robin Algorithm [4]

Round Robin is one of the static load balancing algorithms, where preceding states are not taken into account. It is simple and uses the Round Robin Method for job allocation. It selects the first node at random and then allocates the job to all the other nodes evenly in Round Robin Method. The main advantage of Round Robin is that it does not need any interprocess communication. There is no prior information about the processors' running time, so that some tasks may get heavily loaded.

To overcome this, weighted Round Robin algorithm is being proposed. Here each node assigned has a specific weight. Based on the nodes weight, they would get the requests. If all nodes are equal, then the node is indicated to traffic.

### B. Opportunistic Load Balancing Algorithm [4]

It is also one of the static load balancing algorithms, which do not consider the present workload of the VM. It usually keeps each and every node busy. This deals with the unexecuted tasks quickly and in random order to the current node, where each one of task is assigned to the node randomly. This algorithm provides a load balancing schedule but does not produce a good result. The tasks are processed in a slow manner, where the current execution time of the node is not calculated.

### **C. Min-Min Load Balancing Algorithm [4]**

Min-Min is a static load balancing algorithm, where the parameters associated to the job are recognized in advance. In Min\_Min algorithm, the execution and completion time of the unassigned waiting in queue are identified by the cloud manager. The jobs with minimum execution in time are being assigned first to the processors, so that the task is completed in time. But the tasks with maximum execution need to wait for a specific period of time. As such, all the all the tasks in the processor must be updated and the tasks in the queue must be removed. The task with minimum time execution performs better than the maximum time execution. The main disadvantage of this algorithm is that it leads to starvation. The terminology related to static load balancing for Min-Min is [5]

- Excepted Time of Compute (ETC) - The running time excepted for tasks in all the nodes are stored in ETC,
- Minimum Execution Time Algorithm (MET) – It finds the best job-processor-pair, were current load is not considered, and
- Minimum Completion Time Algorithm (MCT) –It allocates the tasks based on the minimum completion time.

### **D. Max-Min Load Balancing Algorithm [4][5]**

Max-Min is same as the Min-Min algorithm, where the maximum time jobs are selected. Only once the minimum time jobs are completed. When the minimum jobs are completed, the tasks that are in the queue are assigned to the processor. Execution time of all tasks is being updated to the processor. Since it is a static algorithm, the time of every task is calculated in advance and performed in a correct manner. An improved version of max-min algorithm was projected in [6].

### **E. Ant Colony Optimization Based Load Balancing Algorithm**

This algorithm is designed to seek out the optimal path among the food and colony of ant, based on its actions. The main aim of this approach is to distribute the work load among the nodes in an efficient manner. The regional load balancing node is preferred as head node in Cloud Computing Service Provider [4]. As the request is being sent, the ant starts its first movement from the head node [5]. The ants collect the information from the cloud node and assign the tasks to the particular node. Once the task is assigned to the head node, the ant moves in a forward direction with the overloaded node to the next node checking whether the node is overloaded or not. During the movement, if it finds any loaded node again it moves in a forward direction, else it finds the overloaded node, it moves in backward direction and replaces were the node found before [6]. Once the job gets successful it is updated, then the result is reported based on the individual result of the ant. After receiving the individual result they are combined together to build the complete report. The solution set is updated automatically, when the ant updates the result for every movement. To prevent backward movement, the ant commits suicide when it reaches the target node.

### **F. Honeybee Foraging Load Balancing Algorithm**

It is one of the dynamic load balancing algorithms, where they are designed based on the behaviour of honey bees. Honey bees have been classified into two types. They are: finders and reapers. The finder honeybee helps in finding the honey source. Once honey source is found, they do the waggle dance to indicate the quality and quantity of available honey. After that, the reapers gather the honey from the sources. Then, again they go for the waggle dance to specify the honey that is left. In load balancing, the servers are combined together as virtual servers, where each and every virtual server has a process queue. Once the request is received from the queue, it calculates the profit quality as the bee does in waggle dance [4]. The server stays only when the profit is high, or else proceeds to forage by indicating that whether the state is loaded, overloaded, under loaded or balanced. Based on this, the current virtual machines are combined. It needs to maintain a separate queue for each and every node. Depending on the priority, the task is taken into concern, by removing the task that is waiting in the overloaded machine. The tasks removed are loaded into lightly loaded machine. Those tasks are known as scout bee for the next step [5]. The behaviour of honey bee in load balancing technique has stimulated to reduce the response time of virtual machine, which also reduces the waiting time. The main disadvantage of this algorithm is, it does not show any improvement in throughput.

### **G. Biased Random Sampling Load Balancing Algorithm**

Biased Random Sampling is a dynamic load balancing algorithm. Here, random sampling method is being used to achieve the load balancing across all the nodes. In this algorithm, all the servers are treated as nodes [7]. This method is represented in the form of virtual graph, constructed with the connectivity which represents the load on each node. Each node is taken as vertex in a directed graph. When a request is received from the client to the load balancer, the load balancer assigns the job to the node that has a minimum of one in-degree. Once a job is assigned to the node, the server starts executing the job, indicating the reduction in availability of free resources. After the completion of the job, the node gets incremented by one in-degree, indicating the increase in available resources. The addition and deletion of such processes are completed by the process of random sampling technique. [5]. Threshold value is used as a parameter that considers each and every process by representing the maximum walk length. The traversal is from one node to another node until finding a designation is known as a walk. After receiving the request from the load balancer, it compares the current node to the randomly selected node with the threshold value. If the threshold value is equal or greater than the current walk length, the node executes its job, or else it moves to another neighbour node that is randomly selected. The performance decreases as the number of servers increases.

#### H. Active Clustering load balancing Algorithm

Active Clustering is an improved method of random sampling. The concept of clustering is used in this algorithm. The main principle of this algorithm is grouping similar nodes together, and working based on those grouped nodes. Grouping of nodes helps the resources to increase the throughput efficiently. In this algorithm, a method called match-maker is introduced [7]. While an execution starts, the first node selects the neighbour node. The neighbour node is taken as match make node, which connects the neighbour node that is same as initial node. At last the match maker node gets disconnected. And this process is done iteratively to balance the load equally. The system performance is improved highly, by increasing the throughput. There is an efficient utilization of resources when there is an increase in throughput[5].

#### IV. COMPARISON OF LOAD BALANCING ALGORITHMS IN CLOUD COMPUTING

Load Balancing Algorithm	Static/Dynamic	Job Allocation	Advantage	Disadvantage
Round-Robin	Static	Selects the First Node in Random Manner, and allocates other nodes in Round Robin Method.	Treats the entire server equally.	Any process is not known in advance.
Opportunistic Load Balancing Algorithm	Static	Based on the framework of the system.	Keeps every node busy.	The execution time is completed, but the node is still busy.
Min-Min Load Balancing Algorithm	Static	Identifies and completes the job waiting in queue.	Performs better small execution time.	Leads to Starvation.
Max-Min Load Balancing Algorithm	Static	Finding the minimum execution time and deals with the maximum execution time.	Improves efficiency by increasing concurrent execution.	Execution that takes maximum time need to wait for long time.
AntColony Optimization Based Load Balancing Algorithm	Static	Based on actions of ants and seeking an optimal path in collecting their food.	Distributes the workload among nodes in efficient and optimal job scheduling is achieved.	_____
Honeybee Foraging Load Balancing Algorithm	Dynamic	Based on the behaviour of honeybees' and their approach in collecting honey.	The response time and waiting time of the virtual machine are reduced.	There is a decrease in the throughput, when there is an increase in resources.
Biased Random Sampling Load Balancing Algorithm	Dynamic	Based on Random Sampling Method.	Load balancing is achieved through all the nodes in the system.	Corrupts when load increases.
Active Clustering load balancing Algorithm	Dynamic	Grouping nodes together.	Similar nodes are grouped together.	The performance is poor when there is an increase in variety of nodes.

#### V. CONCLUSION

Cloud Computing is a huge concept and load balancing plays a very important role in Clouds. We have discussed and compared various load balancing algorithms, other load balancing algorithms can also be applied. Balancing the network load equally is one of the significant tasks in cloud computing. The ant colony optimization works better and distributes the workload in an efficient manner when compared to other algorithms.

#### REFERENCES

- [1] Rajesh George Rajan, V.Jeyakrishna (Dec 2013), "A Survey on Load Balancing in Cloud Computing Environment", Vol.2, Issue.12, pp.4726-4728.
- [2] ZenonChaczko, VenkateshMadadevan, ShahrzadAslanzadehand, ChristopherMedermind (2011), "Availability and Load Balancing in Cloud Computing", Vol.14, pp.138-140.

- [3] <https://www.edgehosting.com/managed-solutions/managed-hosting-services/server-load-balancing/benefits>.
- [4] Dharmesh Kashyap, Jaydeep Viradiya(Nov 2014),”A Survey of Various Load Balancing Algorithms in Cloud Computing”,Vol.3,Issue.11,pp.115-119.
- [5] Shanti Swaroop Moharana,RajaDeepanD.Ramesh and Digamber Powar(May2013),”Analysis of Load Balancers in Cloud Computing”,Vol.2,Issue.2,pp.101-108.
- [6] Upendra Bhoi,Purvi N.Ramanuj(April 2013) ”Enhanced Max-Min Task Scheduling Algorithm in Cloud Computing”,Vol.2,Issue.4,
- [7] J.Uma,V.Ramasamy,A.Kaleeswaran(Feb 2014),”Load Balancing Algorithms in Cloud Computing Environment-A Methodical Comparison”,Vol.3,Issue.2,pp.272-274.
- [8] [http://en.wikipedia.org/wiki/Cloud\\_computing](http://en.wikipedia.org/wiki/Cloud_computing).