

# A Survey: Load Balancing in Cloud Environment

Sukrati jain<sup>1</sup>, Dr. Ashendra k saxena<sup>2</sup>

<sup>1</sup>CCSIT Teerthanker Mahaveer University Moradabad

<sup>2</sup>CCSIT Teerthanker Mahaveer University Moradabad

[sukratijain0918@gmail.com](mailto:sukratijain0918@gmail.com)

[ashendrasaxena@gmail.com](mailto:ashendrasaxena@gmail.com)

**Abstract** - Today is the world of cloud computing. In the field of computer science, cloud computing is the most interesting technology. Although there are many issues such as power management, security and the most challenging issue is Load Balancing. Several algorithms are developed for balancing the load in cloud computing. Load Balancing is basically the procedure for distributing the workload across multiple servers. For research scholars, it has ever been an interesting subject for attaining a high user gratification and resource utilization. In this review paper, we have studied several approaches related to load balancing and how to compute load degree to check the system status to know it is overloaded or not .

**Keywords**— Ant Colony Optimization, Load Degree, Load Balancing, Cloud Environment;

## I. INTRODUCTION

Cloud computing has become the most popular field for research scholars over the last few years. Cloud computing is a central remote servers where a programs or applications run on server machine [1]. It is just like a new image of computing. Cloud Computing consists of Virtual machine, Host, Servers. User need not to worry about maintenance of any software, it is the responsibility of cloud service provider to provide the needed resources to user.

One of the most challenging issue in the cloud environment is load balancing. It is achieved with the help of multiple resources i.e. multiple servers are available and they have ability to fulfil the request [2]. There is lots of data stored on cloud and multiple requests come for resources, due to this increase load on data servers and sometimes unable to provide resources on time. Therefore, to manage the load by scheduling task appropriately

researchers have introduced some load balancing techniques that can improve the performance of cloud computing along with optimal resource utilization.

This paper organized in to 4 sections: Section 1 presents a study of cloud computing and load balancing and classification of load balancing algorithm. Section 2 gives the answer of the question what is the role of load balancing algorithm in cloud environment, the basic goals and parameters for balancing the load and how to balance the load using ant colony optimization. Section 3 provides reviews and study of different existing methods of load balancing algorithm, drawn from literature survey. Finally, conclusion is presented in Section 4.

### A. Load Balancing Algorithm Classification:

Load balancing algorithms can be categorized into two types which are as follows [3]:

- Static algorithm: The present position of the node is not occupied into deliberation. All the nodes and their characteristics are recognized in advance. The working of this type of algorithm is based on this predefined knowledge. It does not use existing system status information and it is easy to implement.
- Dynamic algorithm: It is based on the present status of the system. It works according to the changes in the state of nodes. Dynamic

algorithms are difficult to implement but it balances the load in efficient manner. Dynamic load balancer uses policies for keeping track of modified info.

On the basis of initiation of the process, load balancing algorithms can be categorized into three classes which are as follows:

1. Sender Initiated: In this type of load balancing algorithm, algorithm is initialized by the sender. Here, the sender leads request messages till it finds a receiver that can receive the load.
2. Receiver Initiated: In this type of load balancing algorithm, algorithm is initiated by the receiver. Here, the receiver leads request messages till it finds a sender that can get the load.
3. Symmetric: It is the blend of both the sender.

## II. LOAD BALANCING IN CLOUD COMPUTING

Load balancing is one of the critical aspects in cloud computing environment that can significantly improve resource utilization, performance and save energy by properly assigning/reassigning computing resources to the incoming requests from users [4]. As technology is growing faster, there are huge amount of users on internet so, to manage and fulfil their requirements, load balancing comes into the picture which ensure that workload is spread equally to all of the available servers without any delay to accomplish higher user satisfaction and maximum throughput with minimum response time [5].

### A. Compute Load Degree

We have to check the status of system to know system is overloaded or not because on the basis of status we are transferring the request to balance load. So we have to compute load degree of server. There are three types of server status [6]:

- Idle: When the percentage of idle nodes exceeds alpha ( $\alpha$ ), change to idle status. If it is in idle status, this request should be transferred to another.  
 $Load\_Degree(n)=0;$

- Normal: When the percentage of the normal nodes exceeds beta ( $\beta$ ), change to normal load status.

$$0 < Load\_Degree(n) \leq Load\_Degreehigh$$

- Overload: When the percentage of the overloaded nodes exceeds gamma  $\gamma$ , change to overloaded status. If it is overload, this request should be transferred to another.

$$Load\_Degreehigh \leq Load\_Degree(n).$$

The parameters alpha ( $\alpha$ ), beta ( $\beta$ ) and gamma ( $\gamma$ ) are to be set balancers. When the load status is idle or normal, then this can be done locally. If the load status is not normal or idle, then this request should be get transferred to another.

So, finding Status of system we require Load degree. Firstly we consider a load Parameter Set:  
 $F = \{f_1, f_2, \dots, f_m\}$

$m =$  represent the total no. of parameters

Compute the load degree as:

$$Load\_Degree(N) = \sum_{i=1}^m \alpha_i f_i$$

Where  $i=1 \dots m$

Load degree are weights that may differ for different kinds of jobs.  $N$  represents the current node.

$$f_i = (1 \leq i \leq m; f_i \in [0, 1])$$

### B. Goals of Load Balancing

The goal of load balancing is to improve the performance by balancing the load among these various resources (network links, CPU, disk drives) to achieve optimal resource utilization, maximum throughput, minimum response time, and avoiding overload. In order to balance the demands of the resources it is important to recognize a few major goals of load balancing algorithms [7]:

- a) Cost effectiveness: The major aim is to achieve an overall improvement in system performance at a reasonable cost.
- b) Scalability and flexibility: The distributed system in which the algorithm is implemented may change in topology or size. Thus the algorithm must be scalable and flexible enough to allow such changes to be handled easily.
- c) Priority: It require prioritization of the resources or jobs need to be done on beforehand through the algorithm itself for better service to the important or high prioritized jobs in spite of equal service provision for all the jobs regardless of their origin.

#### C. Load Balancing Parameters

Some performance metrics for balancing the load are as follows [8]:

- Throughput: It is the amount of work completed in a specified period of time high throughput is sought for higher performance of a system.
- Response time: The time it takes for LB algorithm to start responding to a given instruction is defined as response time for a better performance a short response time is desired.
- Scalability: To manage a growing amount of workload and increasing number of nodes as the cloud expands.
- Fault tolerance: It is the ability of a LB algorithm to continue its operation. If some error occurs, a high fault tolerance is desired for optimal system performance.

#### D. Load Balancing Using Ant Colony Optimization

Each node in ant based control system which was designed to solve the load balancing in cloud environment was configured with:

- 1) Capacity that accommodates a certain.
- 2) Probability of being a destination.
- 3) Pheromone (or probabilistic routing) table.

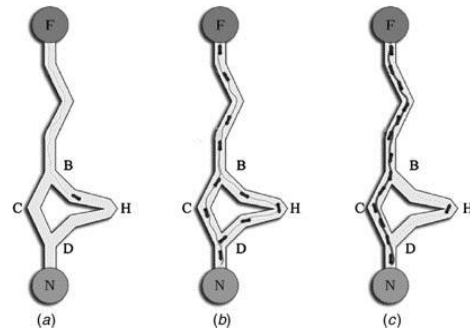


Figure1: Updation through incoming ants

As shown in fig, incoming ants update the entries the pheromone table of the node. The updated routing information can only influences the routing ants and calls that have as their destination [10].

Now we discuss briefly about ant movement, foraging and trailing pheromone, pheromone table which is as follows:-

- Forward and Backward Direction: The routing in a cloud network can be done in both forward and backward direction to discover overloaded and underloaded nodes. The backward ant take the same path as that of its corresponding forward ant but in opposite direction [11].
- Foraging Pheromones: The ants continuously move in the forward direction encountering overloaded node or under loaded node laying down the foraging pheromones [9].
- Trailing Pheromones: If an ant encounters an overloaded node in its movement when it has previously encountered an underloaded node then it will go backward to the underloaded node to check the node is still underloaded or not and if it finds it still underloaded then it

will redistribute the workload to the underloaded node[9].

- Pheromones Tables: The routing tables in the cloud network servers are replaced by the tables of probabilities which are called pheromones tables and the pheromones strengths are exhibited by these probabilities [10].
- Pheromones Updating: The strength of pheromone of each resource was updated after completion of the request [11].

### III. LITERATURE SURVEY

In this research paper, the main focus is on load balancing technique and procedures in cloud environment. There is a complete review on diverse load balancing methods which are present in cloud analyst tool and several policies by different researchers [16].

This paper focuses on to balance the load of entire system while trying to maintain the reliability of the system by creating it as a fault tolerant system. The objective is to study the existing ACO's and to develop an effective fault tolerant system using ant colony optimization. Paper describes an algorithm in order to make the system more reliable and fault tolerant [12].

In this paper, a cloud task scheduling policy based on Ant Colony Optimization and compared with different scheduling algorithms. The aim of these mechanisms is minimizing the make span of a given tasks set. They have been simulated using cloudsims toolkit package [13].

In this paper, Researchers have proposed a new method for load balancing in Cloud Computing. Here we have focused only on accessing different web services. Scheduling of different jobs to different web services in less amount of time. As the number of ants' increases, the time also increases but there is not much difference in time. This shows that the more number of work comes to

job scheduler, the less amount of time to schedule the jobs [14].

In this they proposed an approach. It exploits ants conduct to collect data of cloud hub to assign task to the definite hub. In this approach, when the solicitation is ongoing, the ant and the pheromone initiates the onward growth in the pathway from the "head" hub. The ant moves in onward manner from a loaded hub searching for next hub to ensure whether it is a loaded hub or not [15].

### IV. CONCLUSION

In this research paper, we discussed about various types of load balancing techniques which helps to user for selecting a suitable algorithm for balancing the load and discuss the working of ant colony optimization. We also talking about what are the goals, performance metrics for load balancing and reviews various research papers that are related to load balancing. The goal of this analysis is to understand the concept of load balancing in cloud environment.

### REFERENCES

- [1] Harshada Raut<sup>1</sup>, Kumud Wasnik<sup>2</sup>, "Load Balancing in Cloud Computing using Ant Colony Optimization", International Journal of Innovative Research in Computer and Communication Engineering (An ISO 3297: 2007 Certified Organization) Vol. 3, Issue 12, December 2015.
- [2] Er. Pooja, Er. VivekThapar "Survey of VM Load Balancing Algorithm in Cloud Environment", International Journal of Computer Science Trends and Technology (IJCST) – Volume 4 Issue 2, Mar - Apr 2016.
- [3] Sukrati Jain, Ashendra k. Saxena A Survey of load balancing challenges in cloud environment, IEEE, 12april2017,page(s)291-293.
- [4] KalaiSelvi B. Mary L. (2014, August). A Survey of Load Balancing Algorithms using

- VM. , International Journal of Advancements in Research & Technology: IJOART 2014, pp 68-76, ISSN: 2278-7763.
- [5] Kaur R. And Luthra P. (2012) .Load Balancing in Cloud Computing. Association of Computer Electronics andElectrical Engineers: ACEE 2014, pp. 375-381, ISSN: 1899-0142, DOI:02.ITC.2014.5.92.
- [6] Akshada Bhujbal, Prajakta Jakate, Manasi Wagh, Madhura Pise, Prof.M.V.Marathe, Load Balancing Model in Cloud Computing, Volume 3, Issue 2, February 2015, PP 1-6 ISSN 2349-4395 (Print) & ISSN 2349-4409 (Online).
- [7] Soumya Ray and Ajanta De Sarkar ,” Execution Analysis Of Load Balancing Algorithms In Cloud Computing Environment”, International Journal on Cloud Computing: Services and Architecture (IJCCSA), Vol.2, No.5, 2012.
- [8] Amritpal Singh , A Review of Existing Load Balancing Techniques in Cloud Computing (IJARCET) Volume 4 Issue 7, July 2015.
- [9] Kumar Nishant, Pratik Sharma, Vishal Krishna, Chhavi Gupta and Kunwar Pratap Singh, Nitin and Ravi Rastogi. 2012 Load Balancing of Nodes in Cloud Using Ant Colony Optimization. 14th International Conference on Modelling and Simulation.
- [10] Ratan Mishra and Anant Jaiswal, April 2012, Ant colony Optimization: A Solution of Load balancing in Cloud, International Journal of Web & Semantic Technology (IJWesT) Vol.3, No.2.
- [11] Tanya Prashar, Nancy, Dinesh Kumar, Fault Tolerant ACO using Checkpoint in Grid Computing, International Journal of Computer Applications (0975 – 8887) Volume 98– No.10, July 2014.
- [12] Divya Rastogi,Farhat ullah khan, ”Effective Scheduling Algorithm for Load Balancing using ant colony optimization in cloud computing”,International journal of advances in engineering & Technology,july 2014,Vol. 7, Issue 3,pp. 911-916.
- [13] Medhat Tawfeek, Ashraf El-Sisi, Arabi Keshk and Fawzy Torkey, ” Cloud Task Scheduling Based on Ant Colony Optimization”, International Arab Journal of Information technology, vol.12,No.2, March 2015.
- [14] Ranjan Kumar, G. Sahoo, K.Mukherjee, ” Performance Analysis of Cloud Computing using Ant Colony Optimization Approach”, International Journal of Innovative Research in Science, Engineering and Technology vol.2, Issue 6, June 2013.
- [15] Author Nishant, K. P. Sharma, V. Krishna, C. Gupta, KP. Singh, N. Nitinand R. Rastogi, "Load Balancing of Nodes in Cloud Using Ant Colony Optimization.", In proc. 14th International Conference on Computer Modelling and Simulation,IEEE,pp: 3-8, March 2012.
- [16] Shipra Goyal, Manoj K Verma, ”Load Balancing Techniques in Cloud Computing Environment-A Review”, International journal of Advanced Research in computer Science and Software Engineering Research Paper,Volume 6,Issue4, April 2016.