ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Website: www.ijircce.com

Vol. 5, Issue 4, April 2017

A Comparative Study of Load Balancing Algorithms In Cloud Computing.

Aayushi Sharma, Anshiya Tabassum, G.L. Vasavi, Shreya Hegde, Madhu B.R

B.Tech Student, Computer Science & Engineering, Jain University, Bangalore, India

B.Tech Student, Computer Science and Engineering, Jain University, Bangalore, India

B.Tech Student, Computer Science and Engineering, Jain University, Bangalore, India

B.Tech Student, Computer Science and Engineering, Jain University, Bangalore, India

Asst Professor, Computer Science & Engineering, Jain University, Bangalore, India

ABSTRACT: Cloud computing is an emerging technology aimed at allowing users to easily obtain a wide range of web-based services that previously required tremendous hardware and software skills . Load Balancing is an important aspect of cloud computing for efficient operations in distributed environments. Since the Cloud operators are expanding their services rapidly and clients are demanding more services and better results, load balancing for the Cloud resources has become a very integral and important aspect. Load balancing involves dividing the load equally so the throughput is high with less response time. Various algorithms have been proposed to provide efficient load balancing . These approaches aim to enhance the overall performance of the Cloud and provide the user more satisfying and efficient services. In this paper we will study the different types of load balancing techniques and make a comparative analysis amongst all the existing techniques. We discuss and compare these algorithms to provide an overview of the latest approaches in the field.

KEYWORDS: Cloud Computing, Load Balancing, Task Scheduling.

I. INTRODUCTION

Cloud Computing

Cloud computing is an emerging technology aimed at allowing users to easily obtain a wide range of web-based services that previously required tremendous hardware and software skills. Cloud is a pool of heterogeneous resources and a mesh of huge infrastructure[1]. Infrastructure refers to both the applications delivered to end users as services over the Internet and the hardware and system software in datacenter that is responsible for providing those services.
In a nutshell, Cloud Computing refers to applications and services that run on a distributed network.

Fig 1 : Illustration of Cloud Computing

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Website: www.ijircce.com

Vol. 5, Issue 4, April 2017

Load Balancing

It is a process of reassigning the total load to the individual nodes of the collective system to make resource utilization effective and to improve the response time of the job, simultaneously removing a condition in which some of the nodes are over loaded while some others are under loaded [2]. A load balancing algorithm which is dynamic in nature does not consider the previous state or behaviour of the system, that is, it depends on the present behaviour of the system.

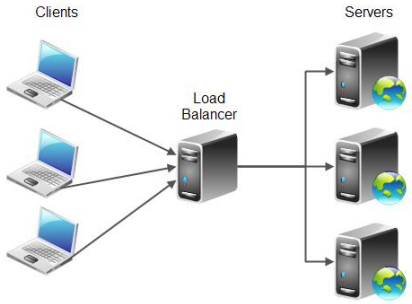

Fig 2: An illustration of Load Balancing

Load Balancing Parameters

For an effective and efficient load balancing , some parameters must be considered to evaluate the load balancing techniques to get better resource optimization . Measurement parameters allow us to see whether the given technique is good enough to balance the load of the traffic on the server or not. Various load balancing measurement parameters are discussed below to evaluate the existing load balancing techniques : [3]

Parameters	Description
1.Throughput	It is calculated as the amount of work completed against time consumed
2. Response time	It is the elapsed time between an enquiry on a sysytem and the response to that enquiry.
3. Fault Tolerance	It is the ability of the load balancing algorithm that allows to keep working properly in some failure condition of the system.
4. Scalability	It is the ability of the computer application or a product to continue to function well when it is changed in size or volume to meet user needs.
5. Performance	It is the overall check of the algorithms working. It comprises the completion of the given task against present known standards like accuracy, cost and speed.
6. Resource Utilization	It is used to keep a check on the utilization of various resources.

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Website: www.ijircce.com

Vol. 5, Issue 4, April 2017

TYPES OF LOAD BALANCING ALGORITHMS [4]

Load balancing algorithms can be broadly categorized into two types:
1) Cooperative: - The common goal of this type of algorithm is to optimize the response time.
2) Non – cooperative: - All running tasks are independent of each other and thus improving the total response time for the local task.

Depending on who initiated the process, load balancing algorithms can be of three categories as given:
• Sender Initiated: If the load balancing algorithm is initialized by the sender.
• Receiver Initiated: If the load balancing algorithm is initiated by the receiver.
• Symmetric: It is the combination of both sender initiated and receiver initiated.
Depending on the current state of the system, load balancing algorithms can be divided into 2 categories as given in

1) Static: It does not depend on the current state of the system. Prior knowledge of the system is needed
2) Dynamic: Decisions on load balancing are based on current state of the system. No prior knowledge is needed. It is better than static approach.

III. A COMPARATIVE STUDY OF LOAD BALANCING ALGORITHMS

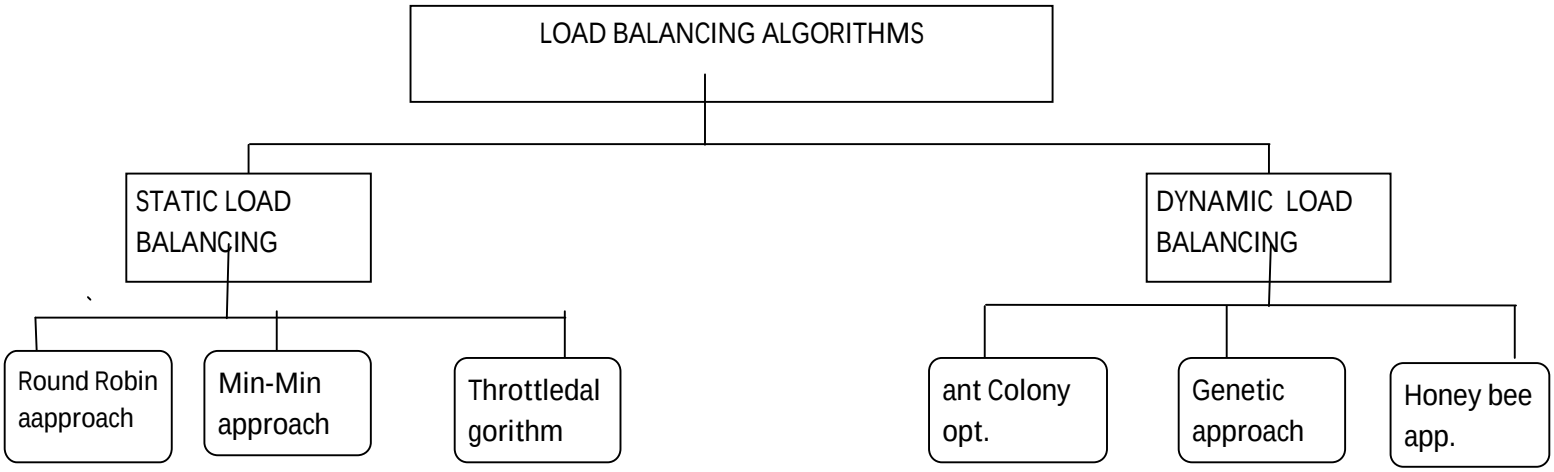

Fig 3: Different load balancing algorithms

A. Static Load Balancing Algorithms:[4]

Static Load balancing algorithms assign the tasks to the nodes based only on the ability of the node to process new requests. The process is based solely on prior knowledge of the nodes' properties and capabilities.
Some of these properties are :
-node's processing power
-memory
- storage capacity
Static load balancing algorithms are non-preemptive i.e once the load is allocated to one node it cannot be transferred to another node.
Following are some of the static load balancing algorithms :

1. Round Robin Algorithm
2. Min-Min load balancing Algorithm
3. Throttled load balancing Alogorithm

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Website: www.ijircce.com

Vol. 5, Issue 4, April 2017

A.1 Round Robin Algorithm [5]

Round-robin load-balancing algorithm (RLBA), is one of the most widely used for distributing loads among the web servers due to its simplicity.

Round Robin (RR) load balancing is one of the most common and still widely used approach. It is a very simple technique which distributes client's requests across the servers. The RR load balancer forwards a client request to each server one by one. Once it reaches the end of the available server list, the RLBA loops back and starts assigning the client requests again from the first server.

The main advantage of RLBA is that it is very easy to implement.

However, in scenarios where request load significantly vary, it cannot distribute the load efficiently. To solve the above problem, weighted round-robin algorithm was proposed . In this approach, load will be assigned depending on the weights of all the nodes. When each node has equal weights, they will receive same traffic. But yet it is difficult to predict the execution time hence, this algorithm is not very suitable for cloud computing environment which requires extensive load balancing.

Fig 4: Round Robin Scheduling

A.2 Min-Min load balancing algorithm[6]

The Min-Min algorithm is simple and most basic out of all present algorithms . It starts with a set of all unmapped tasks. Then the resource which has the minimum completion time for all tasks is found. Next, the task with the minimum size is selected and assigned to the corresponding resource (hence the name Min-Min is given). Last, the task is removed from set and the same procedure is repeated by Min-Min until all tasks are assigned (i.e., when the initial set is empty).

This algorithm considers all jobs at a time. So it produces a better makespan and the performance of Min-Min scheduling algorithm is considered to minimize the completion time of all works. Time complexity of Min-Min algorithm when we have R resources and T tasks is : $O(T^2R)$.

In some cases, Min-Min algorithm fails to utilize the resources efficiently which lead to a load imbalance. However ,the biggest weakness of Min-Min algorithm is it does not considers the work load of each resource. Therefore, some resources maybe always get busy but some nodes maybe still.

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Website: www.ijircce.com

Vol. 5, Issue 4, April 2017

A.3 Throttled load balancing algorithm [7]

In this algorithm the load balancer maintains an index table of virtual machines with their states (2 states can be there: Available or Busy). The client/server first makes a request to data centre to find a suitable virtual machine (VM) to perform the recommended job. The data centre queries the load balancer for allocation of the VM. The load balancer scans the index table from top until the first available VM is found or the index table is scanned

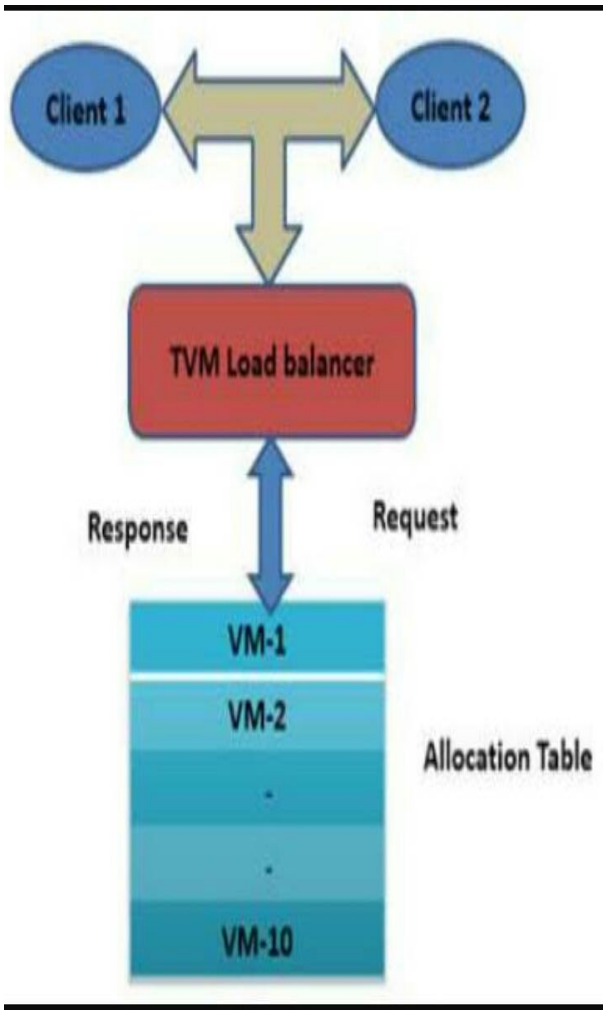

Fig 5: Throttled load balancing algorithm

fully. If the VM is found, the data centre is loaded . The data centre communicates the request to the VM identified by the load balancer. Further, the data centre acknowledges the load balancer of the new allocation and the data centre revises the index table accordingly.

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Website: www.ijircce.com

Vol. 5, Issue 4, April 2017

B. Dynamic load balancing algorithms:

Dynamic load balancing algorithms take into account the different attributes of the node's computing capabilities and network bandwidth [8]. Most of these algorithms work on a combination of knowledge based on prior gathered information about the nodes in the Cloud and run-time properties collected as the selected nodes process the task's components. These algorithms assign the tasks and may dynamically reassign them to the nodes based on the attributes gathered and calculated. Such algorithms require constant monitoring of the nodes and task progress and are usually harder to implement. However, they are more accurate than static approaches and thus, could result in more efficient load balancing.
Following are some of the dynamic algorithms used in load balancing:
1. Ant Colony Optimization
2.Genetic Algorithm
3.Honey Bee Foraging

B.1 Ant Colony Optimization
This algorithm is inspired by one complex behavior of ants : the ability to find shortest paths, this has become the basic strategy of ant colony optimization (ACO) which is the most successful and widely recognized algorithmic technique based on ant behavior.[3]
The ants leave a pheromone trail upon moving from one node to another. By following the pheromone trails, the ant subsequently came to the food sources. The intensity of the pheromone can vary on various factors like the quality of food sources, distance of the food, etc. The ants use these pheromone trails to select the next node. The ants can even modify their paths upon encountering any obstacles in their path. This phenomenon of the ants is used in ant colony optimization (ACO) algorithms for where the ants follow each other through a network of pheromone paths[9]. The ants upon traversal from one node to another update the pheromone trail of that path, so a path becomes more feasible if more ants traverse upon it. Paths that have the highest pheromone intensity have the shortest distance between the point and the best food source. The movements of these ants independently update a solution set.
The movement of ants in this system is of two types: [9]

1. Forward movements -In this type of movement, the ants move/traverse for searching the food sources or for extracting the food.
2. Backward movements-In this type of movement, the ants traverse back to their original place for storing the food after picking up food from the food sources.

The main task of ants in the ACO algorithm is to redistribute work among the nodes. The ants traverse the cloud network, selecting nodes for their next step according to their needs.

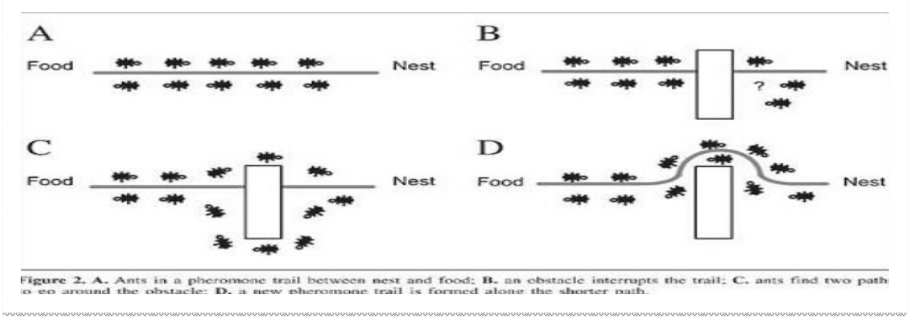

Figure 2. A. Ants in a pheromone trail between nest and food; B. an obstacle interrupts the trail; C. ants find two paths to go around the obstacle; D. a new pheromone trail is formed along the shorter path.

Fig 6: Ant Colony Optimization

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Website: www.ijircce.com

Vol. 5, Issue 4, April 2017

B.2 Genetic Algorithm [10]

Genetic algorithms (GAs) are computer programs that mimic the processes of biological evolution in order to solve problems and to model evolutionary systems. These are adaptive, heuristic search algorithms based on the evolutionary ideas of natural selection and genetics. They are a part of evolutionary computing, rapidly growing area of Artificial Intelligence(AI) and are inspired by Darwin's theory of evolution-"Survival of the fittest". Genetic Algorithm (GA) is much popular for solving NP-Complete problems. Some common key terms used in GA are :

Population- it is a set of possible solutions for proposed problem.

Chromosome-the individuals in the population.

Gene-a variable in a chromosome.

Fitness Function- a type of an objective function used to figure out how close the solution is achieving the set aim.

--Following are the operators used in genetic algorithm:

Selection- solutions with best fittest are selected.

Crossover- for generation of Child, more than one parent is selected.

Mutation- altering the gene value in chromosome.

A simple GA is composed of three operations: selection, genetic operation, and replacement.

The advantage of this technique is that it can handle a vast search space and is applicable to complex objective function .[11]

1. Initial population generation: GA works on fixed bit string representation of individual solution. So, all the possible solutions in the solution space are encoded into binary strings. From this an initial population of ten (10) many chromosomes are selected randomly.
2. Crossover: The objective of this step is to select most of the times the best fitted pair of individuals for crossover. The fitness value of each individual chromosome is calculated using the fitness function as given in 3. This pool of chromosomes undergoes a random single point crossover, where depending upon the crossover point, the portion lying on one side of crossover site is exchanged with the other side. Thus it generates a new pair of individuals.
3. Mutation: Now a very small value (0.05) is picked up as mutation probability. Depending upon the mutation value the bits of the chromosomes, are toggled from 1 to 0 or 0 to 1. The output of this is a new mating pool ready for crossover.

This GA process is repeated till either the fittest chromosome (optimal solution) is found or the termination condition (maximum number of iteration) is exceeded.

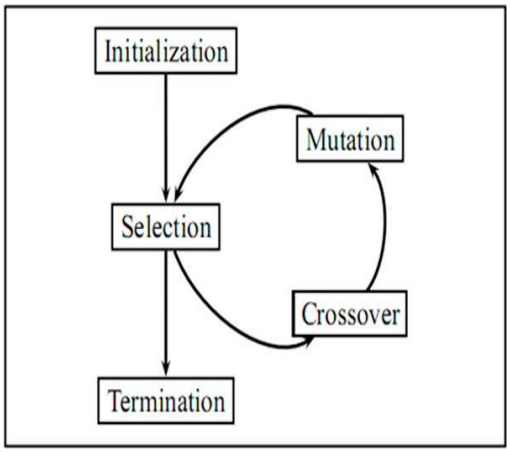

Fig 7. Steps involved in Genetic Algorithm

A colony of honey bee can extend itself over long distances to find the food sources such as flower patches and then these bees harvests nectar or pollen from these flowers. A small fraction of the colony finds the environment looking

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Website: www.ijircce.com

Vol. 5, Issue 4, April 2017

for new flower patches. When food source is encountered the scout bees go in the field surrounding the hive and check for quality beneficial. When they return to the hive, the scouts collect the food harvested[12]. After finding the source it will come back to its hive and perform vibration or waggle dance. By seeing this dance other bees can easily identify the amount and distance of food. Then another type of bee called forager bee will moves to that source by following the scout bees. This process will repeat until the food in that new source gets reduced. This behavior is known as honey bee foraging behavior.[13]
Following are the 3 phases in Honey Bee Foraging-[14]

1. **Employed bee phase :** Employed bees stay on the food source and provide the neighborhood of the source in its memory. After sharing the information in the dance area, employed bees go to food source visited by its previous cycle and choose new food source by using the information in the neighborhood. Then onlooker prefers a food source depending on nectar information provided by employed bees.
2. **Onlooker bee phase:** Onlooker bees get the information about food sources from employed bees in hive and select one of the sources. Onlooker bee is waiting for a dance to choose a food source. Waggle/tremble/Vibration dances are performed by the bees to give an idea about quality and quantity of food and its distance from bee hive.
3. **Scout bee phase** : Scout bee carries out random search. When the nectar source is abandoned by the bees, a new food source is randomly determined by a scout bee.

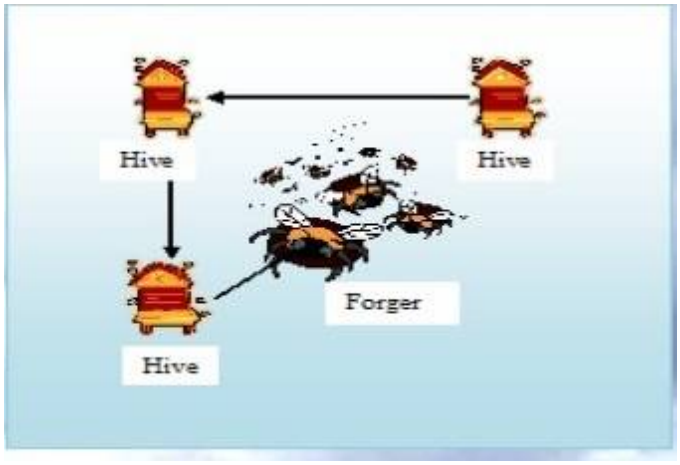

Fig 8. A pictorial representation of Honey Bee Foraging Behavior

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Website: www.ijircce.com

Vol. 5, Issue 4, April 2017

IV. COMPARISON

Algorithms Metrics	Round Robin Approach	Min-Min Approach	Throttled Algorithm	Ant Colony Optimization	Genetic Approach	Honey bee foraging Approach
Throughput	LOW	LOW	MEDIUM	HIGH	HIGH	HIGH
Fault Tolerance	LOW	LOW	LOW	HIGH	HIGH	HIGH
Response Time	SLOW	FAST	FAST	FAST	FAST	FAST
Scalability	LOW	LOW	LOW	MEDIUM	LOW	LOW
Performance	MEDIUM	HIGH	MEDIUM	HIGH	HIGH	LOW
Resource Utilization	IMPROPER	PROPER	PROPER	PROPER	PROPER	PROPER

V. CONCLUSION

In this paper, we surveyed multiple algorithms for load balancing for Cloud Computing. We discussed the metrics that are used to select a suitable and efficient load balancing algorithm for any problem.. Then, we compared the existing algorithms based on the metrics we discussed. We also discussed the advantages and disadvantages of these algorithms. From the aforementioned assessment, we come to a conclusion that static load balancing algorithms are stable than the dynamic ones.

However dynamic algorithms are always better compared to static ones because of the same above mentioned parameters. In future work, we need to implement all these algorithms and check for specific parameters to choose a good load balancing algorithm.

REFERENCES

[1].Mayanka Katyal and Atul Mishra, "A Comparative Study of Load Balancing Algorithms in Cloud Computing Environment", International Journal of Distributed and Cloud Computing ,Volume 1 ,Issue 2, December 2013.

[2]. Reena Panwar and Bhawna Mallick, "A Comparative Study of Load Balancing Algorithms in Cloud Computing",International Journal of Computer Applications Volume 117 – No. 24, May 2015 .

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Website: www.ijircce.com

Vol. 5, Issue 4, April 2017

[3]. Abhinav Hans and Sheetal Kalra, "Comparative Study of Different Cloud Computing Load Balancing Techniques.", 2014 International Conference on Medical Imaging, m-Health and Emerging Communication Systems .
[4]. Ms. Kunjal Garala and Ms. Namrata Goswami, "A Performance Analysis of Load Balancing Algorithms in Cloud Environment.", 2015 International Conference on Computer Communication and Informatics (ICCCI -2015), Jan. 08 – 10, 2015 .
[5]. Furqan Alam, Vijey Thayananthan and Iyad Katib , "Analysis of Round-robin Load-balancing Algorithm with Adaptive and Predictive Approaches.", 2016 UKACC 11th International Conference on Control Belfast, UK, 31st August - 2nd September, 2016.
[6]. Huankai Chen, Professor Frank Wang, Dr Na Helian and Gbola Akanmu, "User-Priority Guided Min-Min Scheduling Algorithm For Load Balancing in Cloud Computing."
[7]. Mr. Durgesh Patel and Mr. Anand S Rajawat, "Efficient Throttled Load Balancing Algorithm in Cloud Environment", International Journal of Modern Trends in Engineering and Research, 2015.
[8]. Klaithem Al Nuaimi, Nader Mohamed, Mariam Al Nuaimi and Jameela Al-Jaroodi , "A Survey of Load Balancing in Cloud Computing: Challenges and Algorithms.", 2012 IEEE Second Symposium on Network Cloud Computing and Applications.
[9]. Kumar Nishant, Pratik Sharma, Vishal Krishna, Chhavi Gupta, Kunwar Pratap Nitin and Ravi Rastogi, "Load Balancing of Nodes in Cloud Using Ant Colony Optimization", 2012 14th International Conference on Modelling and Simulation.
[10]. Mr.Hussain A Makasarwala and Mr.Prasun Hazari, "Using Genetic Algorithm for Load Balancing in Cloud Computing", ECAI-2016, International Conference 8th Edition, Electronics Computers And Artificial Intelligence.
[11]. Kousik Dasgupta, Brototi Mandal, Paramartha Dutta, Jyotsna Kumar Mondal and Santanu Dam, "A Genetic Algorithm (GA) based Load Balancing Strategy for Cloud Computing." ,International Conference on Computational Intelligence: Modeling Techniques and Applications (CIMTA) 2013.
[12]. Harshit Gupta and Kalicharan Sahu, "Honey Bee Behavior Based Load Balancing of Tasks in Cloud Computing.", International Journal of Science and Research (IJSR) .
[13]. Anju Baby and JenoLovesum, "A Survey on Honey Bee Inspired Load Balancing of tasks in Cloud Computing.", International Journal of Engineering Research & Technology (IJERT) Vol. 2 Issue 12, December – 2013.
[14]. Ms. Anna Baby, Dr. Joshua Samuel Raj, "An efficient load balancing using Bee foraging technique with Random stealing", IOSR Journal of Computer Engineering (IOSR-JCE), Volume 17, Issue 2, Ver. V (Mar – Apr. 2015).