International Journal of Trend in Research and Development, Volume 6(3), ISSN: 2394-9333
www.ijtrd.com

Cloud Computing Based on Load Balancing

[1]Giripai U and [2]Mohanaprasath S,
[1,2]Student, Dept. of Software Systems, Sri Krishna Arts and Science College, Coimbatore, Tamil Nadu, India

Abstract— The cloud computing technology is a big step in virtualization. Cloud load balancing is one of the process of distributing workloads and computing resources in a cloud computing environment. Load balancing allows enterprises to manage application or by allocating resources among multiple computers, networks or servers. Cloud load balancing is involved in hosting the distribution of demands that reside over the Internet and the workload traffic.

Keywords- *Cloud Computing, Load Balancing Algorithms Virtualization, Task allocation, Static and Dynamic Load Balancing and Hierarchical Load Balancing.*

I. INTRODUCTION

Cloud Computing gives an easy method to retain the data that involves virtualization, distributed computing and web services. Due to the easy method of access to the Internet, each individual, organization uses Cloud computing services. NIST(National Institute of Standards and Technology) defines Cloud computing is a computing model used everywhere and provides convenient, on-demand access to a shared pool of computing resources such as networks, servers, storage, applications, etc. These resources can be assigned and released with minimal management effort or service provider interaction. It provides us with 3 services such as Software as Service (SAAS), Platform as a Service (PAAS) and Infrastructure as a Service (IAAS). Different virtual and physical resources are provided to the users on demand. In CC access to the resource is based on Virtualization. Virtualization is an abstraction of real machines. Virtual Machine has ability to run applications like any real function. Virtualization provides us with facilities similar to real machines. We can create larger number of less powered servers through virtualization, which reduces the overall cost in space, infrastructure and power. Cloud resources can be scaled using virtualization technique. Cloud resources are dynamically allocated to users because of demand. As the number of user for the cloud increases, the available resources decrease *dynamically.*

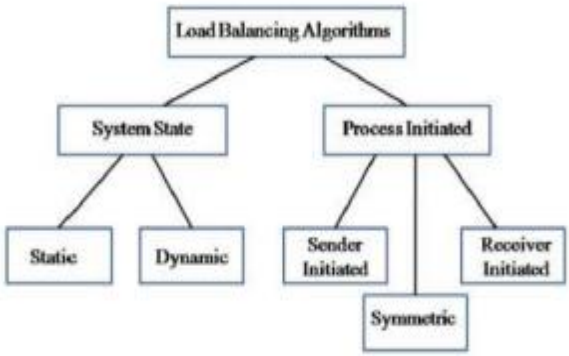
Because of its efficiencies, cloud computing has become a new technology trend, a solution in which all computing resources (hardware, software, networking, storage, etc) are provided promptly to the user as per the request. But it is due to the explosion of data exchange that this technology is posing challenges for developers, experts and researchers around the world, especially the load balancing on cloud data center. Because load balancing is all about improving the quality of service, optimizing computing resources, or otherwise improving the efficiency of the cloud system of the Cloud Service Provider. A few years ago, the amount of data transmitted on a global network if stored on DVDs, the number of disks lined up would be double the distance to the moon. Loading balancing helps us to save computing resources and increase user services.

II. LOAD BALANCING

Load balancing is a process of distributing the workload uniformly and dynamically across all the available nodes in the cloud. By shifting the workloads among different nodes we can improve the overall system performance. By utilizing the resources properly carbon emission can be minimized. Throughput, performance, response time, scalability, resource utilization and fault tolerance are some of the measurement parameters that can be used to evaluate the load balancing techniques. By minimizing the response time the load balancing will have to maximize the throughput. It also helps in clean and green environment. The energy consumption and carbon emission is reduced due to loading balance.

Need of Load balancing:

- To maximize user satisfaction
- To increase resource utilization.
- To minimize execution time and waiting time.
- To maximize performance
- To increase fault tolerance
- To maintain stability of system



Load balancing uses metrics for improving its performance. These metrics are explained below:

Throughput: It is measure of execution of a process and amount of time required for it.
Response Time: This is the time taken by the task to reach the system.
Resource Utilization: It provides services to a huge number of users assigned to it.
Fault Tolerant: The ability of the infrastructure to continue to provide service to underlying applications even after the failure of one or more component layer.
Migration Time: Migration Time is the time taken to move from one node to another.
Performance: It measures the time of execution of a result.
Scalability: It is used to decrease or expand the number of machines or process influenced by it.
Waiting Time: It is the waiting time spent by the task for Arriving.

III. STATIC ALGORITHMS

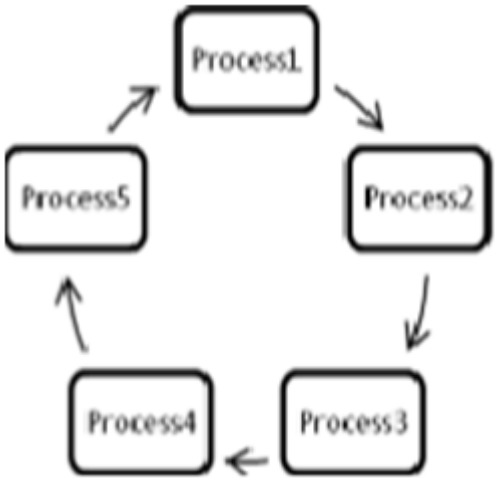
Static algorithms are best in homogeneous environments. However, they are not flexible and cannot consider the dynamic changes to the attributes. While assigning some works to the nodes, static load balancing algorithms will not check the state and functions of the node in previous tasks. Some Static

Algorithms are:

- Round Robin Load Balancing Algorithm
- Load Balancing Min-Min Algorithm
- Load Balancing Min-Max Algorithm

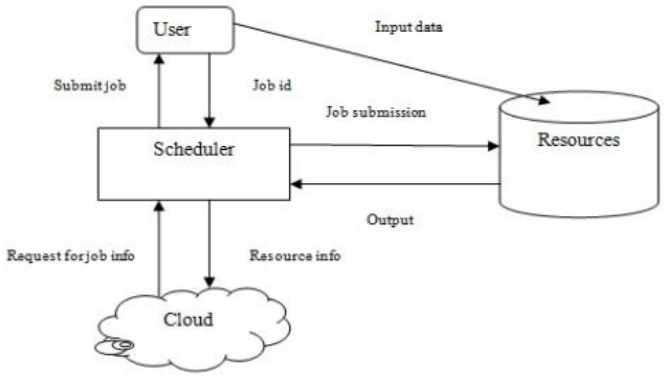
A. Round Robin Load Balancing Algorithm

In this Round Robin algorithm, fixed quantum of time is given to the job. Round Robin is used for network communication. It allocates jobs to all nodes in a circular motion. Processors are assigned in a circular fashion and hence there is no starvation. This algorithm provides quick response in the case of equal workload distribution among processes. However, some nodes may be over loaded while others remain idle and unused.



B. Load Balancing Min-Min Algorithm

A list of task is maintained in order and minimum time completion time is calculated for all the available nodes. A task with minimum completion time is assigned to the requested machine. Hence the name of the algorithm is Load balancing Min-Min. Running time of the machine and the list should be updated. It provides good results when small tasks are assigned more.



C. MIN-MAX Load Balancing Algorithm

A list of task will be updated and minimum completion of time is calculated for all the available nodes. A task with intense as possible completion time is assigned to the machine. Hence the name of the algorithm is Load Balancing Min-Max.

IV. DYNAMIC ALGORITHMS

Dynamic algorithms provide better results in heterogeneous-environment. These algorithms are more stretchable. Dynamic algorithms can take in charge of the dynamic changes to the attributes. However, these algorithms are more composite in nature. Main advantage of this algorithm is that, the selection of task is based on present state and this will help to improve the performance of the system. Dynamic algorithms can be categorized in the following two forms:

1. Distributed System:

Here all the nodes can interact with each other and load balancing algorithm is executed by using all the nodes in the system. The work of load balancing is distributed among all the nodes. Interaction among nodes can be both cooperative or non-cooperative. If any of the node fails in the system, it will not stop the function. i) In cooperative distributed system, all node works together to solve a particular problem, if a problem is raised at any situation. ii) In non-cooperative distributed system, each node works on its own.

2. Non-distributed System:

Non-distributed can be the control of the system or semi-distributed. In centralized system, central node is responsible for load balancing of the entire system. The other nodes will interact with this central node. If central node fails, it will stop the function. In case of failure, it is not easy for a recovery. In semi-distributed system, nodes are grouped together to form a cluster. A central node of each cluster performs load balancing of the entire system. If central node of cluster fails, it will stop the function of that cluster only. Multiple central nodes can manage the load balancing.

Some dynamic algorithms are:

- Honeybee Foraging Behavior Load Balancing Algorithm
- Throttled Load Balancing Algorithm.
- ESCE (Equally Spread Current Execution) Load Balancing Algorithm.
- Ant Colony Load Balancing Algorithm
- Biased Random Sampling Load Balancing Algorithm
- Modified Throttled Load Balancing Algorithm.

A. Honeybee Foraging Behavior Load Balancing Algorithm

This algorithm was derived from the behavior of real honey bees in search of their food sources. After finding the food sources, the honey bees come back to the bee hive to make an information about the food source. They perform this in group movement. This is also known as "waggle Dance". They do this waggle dance to inform the other bees for the exact location of the food source. This waggle dance shows the quality, amount of the food and the distance of the food source from the bee hive.

1) Bee foraging behavior:

The artificial bee colony algorithm (ABC) is an optimization algorithm based on the intelligent hunt behavior of honey bee in large group and was proposed by Karaboga in 2005. The algorithm is completely inspired by natural hunt behavior of honey bees.

1. Initialization process:

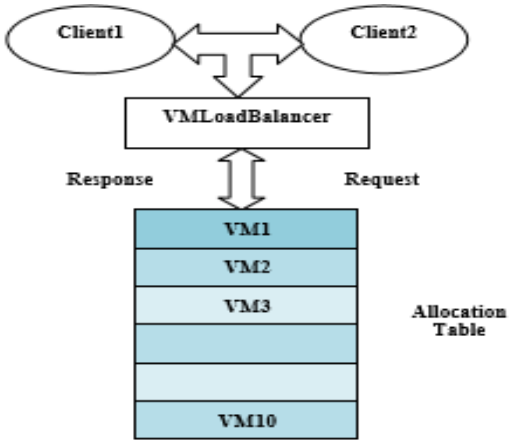
During initialization process, Artificial Bee Colony (ABC) algorithm commence by correlating all the bees without any authority produced for food sources. Certain food sources are randomly selected by bees and the nectar amount of honey bees is determined. These bees come onto the hive and share the information with bees waiting in waggle dance area.

Each food source is generated as:

$$X_{ab} = X_{mn\,b} + \text{rand}\,[0,1]\,(X_{max\,j} - X_{min\,j})$$

2. Throttled Load Balancing Algorithm

Throttled load balancer is suitable for virtual machines. Load balancing algorithm maintains the list of entire virtual machines in the system. It scans the indexing table, when it receives a request. If virtual machine is available, the job will be assigned to the requested machine. After each allocation and de-allocation of resources, the update process will be done.



Throttled Load Balancing

C. ESCE (Equally Spread Current Execution) Load Balancing Algorithm

Load balancing algorithm maintains the list of entire virtual machines and jobs. It scans the entire list of virtual machine, when it receives a request. The request is allocated to that particular VM, which can satisfy the clients request and this process can be done only after a particular VM is found. This algorithm distributes the equal load among all Virtual Machine's.

D. ANT COLONY Load Balancing Algorithm

Real ant selects a shortest path in search for its food. This algorithm depicts the behavior of real ants. When request is initiated, ant starts to move. Ant continuously checks whether the node is overloaded or not overloaded. If ant finds any overloaded node, it returns back. And if ant finds any under loaded node, it tends to continue its process. Is used to collect the information from different nodes, according to the behavior of the ant.

E. Biased Random Sampling Load Balancing Algorithm

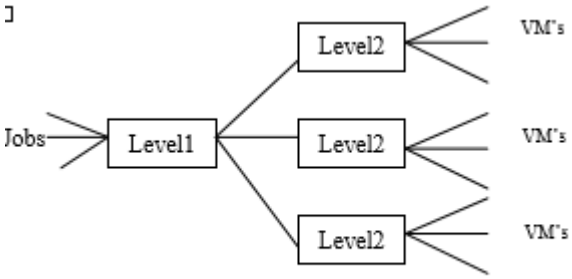
The load is balanced by this algorithm through random sampling of system domain. Construction of the Virtual Graph is done. In a directed graph, each node is represented by a vertex and which is figured in degree represents the free resources of that node. The load balancer allocates the job to the node which has minimum of one in-degree. The in-degree of the node is expanded and decremented once when the job is completed and simultaneously when the job is allocated. This is handled by the process of random sampling.

F. Modified Throttled Load Balancing Algorithm

This algorithm is mainly focused on how jobs are allocated to the available Virtual Machine's. This algorithm maintains an index table of VM's and also the current state of VMs (BUSY/AVAILABLE). This algorithm is initiated by VM at the first index depending upon the state of the VM. According to the request the available VM is assigned. If the VM is not available it returns -1. If the new request arrives, the VM at the previous VM index + 1 is chosen depending on the nature of VM.

V. HIERARCHICAL LOAD BALANCING ALGORITHM

Hierarchical Load Balancing involves heterogeneous levels in load balancing decisions. Every node is maintained or balanced by its parent node. Parent node takes the responsibility for load balancing. Hierarchical load balancing can be used in both homogeneous and heterogeneous environment. Cluster can also play a vital role in hierarchical load balancing. Clustering is the process of organizing same type of objects into groups. Virtual Machine's having similar characteristics are grouped logically. The last level is the Virtual Machine.



Hierarchical Load Balancing

VI. COMPARISON BETWEEN VARIOUS LOAD BALANCING ALGORITHMS

Algorithm	Advantages	Disadvantages
Round Robin Load Balancing Algorithm	• It is Simple algorithm and importance is on fairness. • It works in round fashion. • Fast reply in the case of equal workload distribution • There is no malnourishment.	• Each node is immovable with a time slice. • It is not flexible and scalable. • Some node may have heavy load and some nodes are idle. • Does not save the state of earlier allocation of a VM. • Pre-emption is required.
MIN-MIN Load Balancing Algorithm	• It is simple and reckless algorithm. • It works better for smaller task	• Selects the task having least completion time. • There is malnourishment. Smaller tasks will get executed first, while the larger tasks keep on in the waiting stage. • Poor load balancing • Does not contemplate the existing load on a resource.
MIN-MAX Load Balancing Algorithm	• It is simple algorithm. • It runs short tasks in parallel	• Selects the task having the maximum completion time • There is a starvation. Larger

		tasks will execute first, while the smaller tasks need to wait. • Poor load balancing
Honeybee Foraging Behavior Load Balancing Algorithm	• Self-organizing, nature enthused algorithm. • Performance will be attained by increasing the system size. • Suitable for varied environment	• Rise in resources will not rise the overall throughput
Throttled Load Balancing Algorithm	• List of VMs is preserved along with the status of each VM • Good performance • Better resource utilization	• Scans the whole list of VMs from the beginning • Does not reflect the current load on VM.
ESCE Load Balancing Algorithm	• Upholds equal load at all VMs • Maximize the throughput	• Dominant point of failure • Not fault accepting.
ANT COLONY Load Balancing Algorithm	• Under loaded node is found at beginning of the search • Dispersed	• Network overhead • Delay in moving forward and backward.
Biased Random Sampling Load Balancing Algorithm	• Fully decentralized • Suitable in large network	• Performance is tainted with an increase in diversity
Modified Throttled Load Balancing Algorithm	• Index table is parsed from the index next to already assigned VM. • Faster response throttled algorithm • Well-organized usage of available resources than	• Does not reflect the current load on VM.
Hierarchical Load Balancing	• Faster response • Suitable for homogeneous and heterogeneous environment	• Less fault tolerant

CONCLUSION

Cloud computing allows wide range of users to access distributed, scalable, virtualization, hardware and software resources over the network. Load balancing is one of the most discussed issue of cloud computing. It is a process, which distributes workload evenly across all the nodes in the entire cloud. we can achieve a high user satisfaction and resource utilization, through efficient load balancing algorithms. Hence, this improves the resource utility and overall performance of the system. With proper load balancer, resource consumption can be kept at a low state, which will further reduce energy consumption and carbon emission rate. Through hierarchical structure of system, performance of the system will be maximised.

References

[1] Peter Mell Timothy Grance", The NIST Definition of Cloud Computing", National Institute of Standards and Technology Special Publication 800-145(September 2011) ,csrc.nist.gov/publications/800-145/SP800-145.

[2] Ravinder Kaur," Load Balancing in Cloud Computing", Association of Computer Electronics and Electrical Engineers, 2014, DOI: 02.ITC.2014.5.92.

[3] Asha latha. R," Dynamic load balancing methods for resource optimization in cloud computing environment ", 2015 Annual IEEE India Conference (INDICON), Pages: 1 - 6, DOI: 10.1109/INDICON.2015.7443148.

[4] Garima Rastogi, Dr Rama Sushil, " Analytical Literature Survey on Existing Load Balancing Schemes in Cloud Computing", 2015 International Conference on Green Computing and Internet of Things ICGC loT (Internet Of Things), pages:1506-1510.

[5] K. Karthik, " Performance analysis of load balancing techniques in cloud computing environment", 2015.

[6] Surbhi Kapoor, Dr. ChetnaDabas," Cluster Based Load Balancing in Cloud Computing", 2015 Eighth International Conference on Contemporary Computing (IC3)

[7] AartiVig, Rajendra Singh Kushwah and Shivpratap Singh Kushwah, "An Efficient distributed Approach for Load Balancing in Cloud Computing", 2015 International Conference on Computational Intelligence and Communication Networks

[8] Mohapatra, S., Mohanty, S., and Rekha, K.S., 2013, "Analysis of different variants in round robin algorithms for load balancing in cloud computing," International Journal of Computer Applications, Vol. 69, Issue 22.

[9] Goel, H., Chamoli, N., June 2014, "Job Scheduling Algorithms for Cloud Computing: A Survey," International Journal of Computer Applications, Volume 95, Issue 23.

[10] Ray, S., and Sarkar, A. De., 2012, "Execution analysis of load balancing algorithms in cloud computing environment," International Journal on Cloud Computing: Services & Architecture, Vol. 2, No. 5.

[11] Nishant, K., Sharma, P., Krishna, V., Gupta, C., Singh, K.P., Nitin, N., and Rastogi, R., 2012, "Load balancing of nodes in cloud using ant colony optimization," in Computer Modelling and Simulation (UKSim) 14th International Conference on IEEE, Vol. 4, Issue 4, pp. 3–8.

[12] Subramanian, S., Krishna, G.N., Kumar, M.K., Suresh, P., and Karpagam, G., 2012, "An adaptive algorithm for dynamic priority based virtual machine scheduling in cloud," International Journal of Computer Science Issues (IJCSI), Vol. 9, No. 6.

[13] Kaur, J., 2012, "Comparison of load balancing algorithms in a cloud," International Journal of Engineering Research and Applications, Vol. 2, Issue 3, pages 1169–173.

[14] Moharana, S.S., Ramesh R.D., and Powar, D., May 2013, "Analysis of Load Balancers in Cloud

Computing," International Journal of Computer Science and Engineering, Vol.2, Issue2.

[15] Roy, A., and Dutta, D., April 2013, "Dynamic Load Balancing: Improve Efficiency in Cloud Computing," International Journal of Emerging Research in Management & Technology, Vol. 2. Issue 4.

[16] Khiyaita, A., Zbakh, M., Bakkali, H. El., and Kettani, D.El, 2012, "Load balancing cloud computing: state of art," In National Days of Network Security and Systems (JNS2), IEEE, pages 106–109.

[17] Kansal, N. J., and Chana, I., 2012, "Existing load balancing techniques in cloud computing: A systematic review," Journal of Information Systems & Communication, Vol. 3, Issue 1.