International Conference on Systems, Science, Control, Communication, Engineering and Technology 2015 [ICSSCCET 2015]

ISBN	978-81-929866-1-6	**VOL**	01
Website	icssccet.org	**eMail**	icssccet@asdf.res.in
Received	10 - July - 2015	**Accepted**	31- July - 2015
Article ID	ICSSCCET012	**eAID**	ICSSCCET.2015.012

Load Balancing Algorithms in Cloud Environment

M. Aruna[1], Dr. D. Bhanu[2], Dr. S. Karthik[3]

[1]Assistant Professor, Department of Computer Science and Engineering, Sasurie Academy of Engineering,
[2]Professor and Head, Department of Information Technology, Karpagam Institute of Technology,
[3]Professor and Dean, Department of Computer Science and Engineering, SNS College of Technology, Coimbatore, India

Abstract: *Cloud Computing is a pool of resources that can be shared among the users. At present, Cloud Computing is an emerging technology since it provides services at the user level. There are several issues or challenges in the Cloud Computing environment such as availability, security and resource allocation, etc. The paper concentrates on availability of nodes in the cloud. Balancing load under the nodes will increase availability of nodes in Cloud. In order to enhance the performance of the entire cloud environment, efficient Load Balancing techniques are needed. Load Balancing (LB) algorithms distribute the load evenly across all the nodes in cloud. Load Balancing in Cloud Computing will also increase the reliability and user satisfaction.*

Keywords: Cloud Computing, Load Balancing, Virtualization

I. INTRODUCTION

Large business and small business companies are moving to cloud environment because of its scalability. The jobs arriving to the Cloud Environment are executed by the large data centers which have thousands of blade servers. It provides different types of services to the users. Users can get the services with no need to know their infrastructure i.e.,, users do not know where the service is originated and its infrastructure. Users need to pay only for what they used from cloud in the form of services, which is the simplicity of Cloud. The four different types of cloud environment are as follows,

- Public Cloud (Free of Cost, anyone can access)
- Private Cloud (Pay for what you used, only for single organization people)
- Hybrid Cloud (Combined both public & private Clouds)
- Community Cloud(For Communication purpose)

Users can access the cloud resources in the form of services. There are three basic services provided by the Cloud Environment. They are,

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

A. Virtualization

Cloud Computing is based on the concept of Virtualization technology. Virtualization is the software implementation on the bare hardware so that the resources under the hardware can be utilized more effectively. Cloud Computing uses the virtualization technique

Cite this article as: M. Aruna, Dr. D. Bhanu, Dr. S. Karthik. "Load Balancing Algorithms in Cloud Environment." *International Conference on Systems, Science, Control, Communication, Engineering and Technology (2015)*: 50-54. Print.

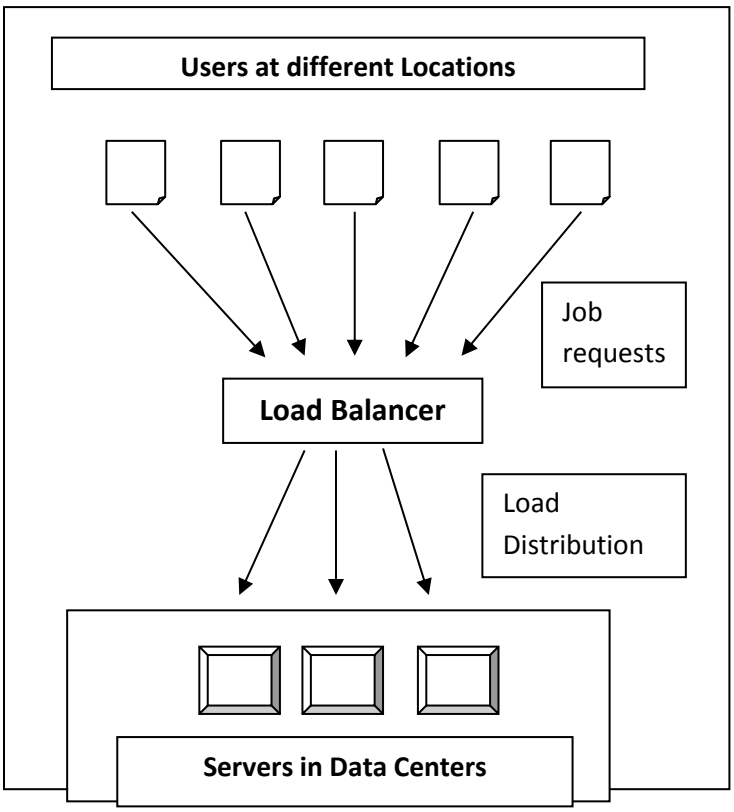
to make use the cloud resources efficiently [3]. Two types of virtualization such as Full Virtualization and Para Virtualization can be used in cloud environment.

II. LOAD BALANCING

Load Balancing [1] is a technique to balance the load across cloud environment. It is the process of transferring load from heavily loaded nodes to low loaded nodes. As a result, no node should be heavily loaded which will increase the availability of nodes. If all the jobs are arrived to the single node, then its queue size is increased and it becomes overloaded. There is a need to balance the load across several nodes, so that every node is in running state but not in overloaded state. The goals of load balancing are as follows[2]:

- To increase the availability
- To increase the user satisfaction
- To improve the reource utilization ratio
- To minimize the waiting time of job in queue as well as to reduce job execution time
- To improve the overall performance of Cloud environment

Figure 1 shows the major works of Load Balancing Technique. The Load Balancer may be any software or hardware which receives jobs from different users in different locations. The received loads are distributed evenly across all the servers in Data Center. Table 1 shows the comparison of LB algorithms.



A. Basic Types of Load Balancing Algorithms

Depending on the initiator of the algorithm, Load Balancing algorithms can be categorized into three types [3]:

Sender Initiated - Sender identifies that the nodes are overwhelmed so that the sender initiates the execution of LB algorithm.

Receiver Initiated - The requirement of Load balancing situation can be identified by the receiver/server in cloud and that server initiates the execution of LB algorithm.

Symmetric -It is the combination of both the sender initiated and receiver initiated types. It takes advantages both types.

Based on the current state of the system, load balancing algorithms can be divided into two types:

Static Schemes - The current status of the node is not taken into account [6]. All the nodes and their properties are predefined. Since it does not use current system status information, it is less complex and it is easy to implement.

Dynamic Schemes - This type of algorithm is based on the current system information [6]. The algorithm works according to the changes in the state of nodes. Dynamic schemes are expensive one and are very complex to implement but it balances the load in effective manner.

Cite this article as: M. Aruna, Dr. D. Bhanu, Dr. S. Karthik. "Load Balancing Algorithms in Cloud Environment." *International Conference on Systems, Science, Control, Communication, Engineering and Technology (2015)*: 50-54. Print.

Status Table - Status table [1] is a data structure to maintain the current status of all the nodes in the cloud environment. This information can be used by some of the dynamic scheme algorithms to allocate jobs to the nodes that are not heavily loaded.

TABLE 1
COMPARISON OF LB ALGORITHMS

S.No.	Algorithm	Description	Advantages
1.	Dynamic round robin algorithm [7]	1. Uses two rules to save the power consumption 2. Works for consolidation of virtual machine	Reduce the power consumption
2.	Hybrid algorithm [1]	1. Combination of dynamic round robin and first-fit algorithm 2. Applied in non-rush hours and rush hours	1. Improved resource utilization 2. Reduced power consumption
3.	ESCE algorithm [8]	Estimate the size of job and look for availability of resources	Improved response time and processing time
4.	Central Load Balancing policy for VM	Balances the load evenly	Improves overall performance
5.	Enhanced Equally Distributed Load Balancing Algorithm [9]	Based on the counter variable, the job is allocated by Central Server	1. Computing Resource is distributed efficiently and fairly 2. Reduces request to response ratio
6.	Decentralized Content Aware Load Balancing Algorithm [9]	1. Uses Unique and Special Property (USP) of nodes 2. Uses content information to narrow down the search	1. Improves the searching performance hence increasing overall performance 2. Reduces idle time of nodes
7.	Join-Idle Queue Algorithm [10]	1. Assigns idle processors to dispatchers for the availability of idle processors 2. Then assigns jobs to processors to reduce average queue length	1. Reduces system load 2. Less communication overhead
8.	Honeybee Foraging Behavior [10]	Achieves global load balancing through local server actions	Improved scalability
9.	Min-Min Algorithm [7]	1. Estimates minimum execution time and minimum Completion time 2. Jobs having minimum completion time is executed first	Smaller tasks are executed quickly
10.	Max-Min Algorithm [11]	1. Same as Min-Min 2. Gives more priority to larger tasks than smaller one	Larger tasks are executed quickly and efficiently
11.	RASA Algorithm [11]	Combination of both Min-Min and Max-Min Algorithms	1. Efficient resource allocation 2. Minimum execution time
12.	Improved Max-Min Algorithm [12]	1. Improved version of Max-Min Algorithm 2. Assigns task with minimum execution time	Scheduling jobs effectively
13.	2-Phase Load Balancing Algorithm [7]	1. Uses OLB to keep each node busy 2. Uses LBMM to achieve minimum execution time of each job	1. Efficient utilization of resources 2. Enhances work efficiency
14.	PALB Algorithm [13]	1. Implemented in Cluster Controller 2. Use Job Scheduler to simulate requests from users for virtual machine instances	Physical Machines that are in idle state are move to power off state to conserve energy

II. PROPOSED WORK

Existing Load balancing algorithms have some drawbacks in improving overall performance of the cloud environment. Still there is a problem of overloading nodes in the Cloud environment. It is very difficult to manage entire cloud environment. Hence the proposed idea is to divide the entire cloud environment into several partitions based on its geographical locations [1]. Now the Load balancing algorithm can be applied only to the partitions, not to the entire cloud. Fig 2 shows the cloud environment after partitioning is done. The load balancing algorithm is applied to each partition in order to avoid overloading of nodes.

Cite this article as: M. Aruna, Dr. D. Bhanu, Dr. S. Karthik. "Load Balancing Algorithms in Cloud Environment." *International Conference on Systems, Science, Control, Communication, Engineering and Technology (2015)*: 50-54. Print.

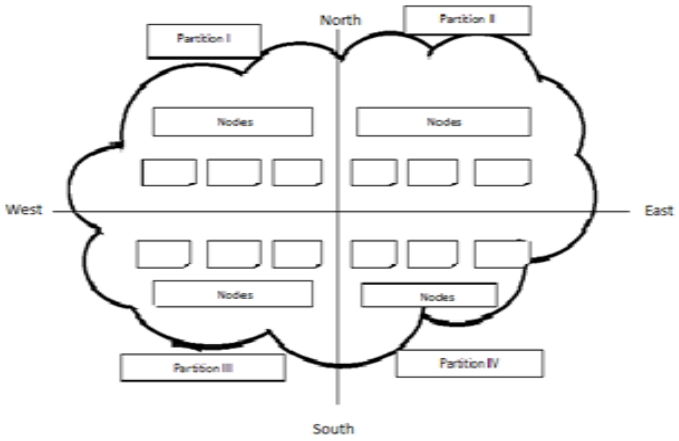


Figure 2. Partitioned Cloud Environment

For balancing load in cloud partition model [1] there are two important components needed:

- Load Balancer
- Main Controller

Load Balancer is associated with each partition whose work is to maintain the state information and is to be updated in periodic intervals. Whenever the controller receiving a job it has to communicate with each partition to collect state information. Then the job is allocated to the partition if it is in idle or normal state. After assigning job to the partition, the balancer has to update the status information of each node in that partition.

Main controller receives all the jobs that arrive from the cloud. Whenever the main controller receives the job it has to decide, which partition to receive the job. Each partition has the state information associated with it. It may be in idle state, normal state or heavily loaded state. The state of the node in particular partition is set by considering several parameters of that node and the parameters may be static or dynamic. Static parameters include number of CPUs, memory size and speed of the processor or CPU. Dynamic parameters include CPU utilization ratio and memory utilization ratio.
An algorithm will be designed for the nodes that are idle or normal and it has to update the status information of each node periodically. Load Balancer in each partition maintains the status table. The status table contains information about the load of all the nodes in that partition. This table is updated by a Load Balancer periodically. For better efficiency, balancer maintains the two status table and each of which are associated set by the "Flag".

IV. FUTURE WORK

The future work is to develop two algorithms, one for the partitions of Cloud environment that are in idle state and another one for the partitions that are in normal states. Switching mechanism is needed for applying these two alternative algorithms. If the partition is in idle state one simple algorithm is to be used and later the same partition can became normal state and alternative algorithm is to be used. The algorithm designed for normal state partitions should be more efficient so that it avoids the partition becoming overloaded.

V. CONCLUSION

Though there are several issues in cloud environment, it has been widely adopted by many organizations and industries. Researchers are doing many works to resolve those challenges and issues. For Load Balancing issue, the solution is to develop suitable algorithms that balance the load across the partitioned cloud environment. Both the algorithms should work accordingly as the partition status changes. It reduces the server overhead, increase throughput, increase performance, reduce server power consumption and also distribute the load across nodes.

REFERNECES

[1] Gaochao Xu, Junjie Pang and Xiaodong Fu, “A Load Balancing Model Based on Cloud Partitioning for the Public Load”, IEEE Transactions on Cloud Computing, 2013.
[2] Zenon Chaczko, Venkatesh Mahadevan, Shahrzad Aslanzadeh and Christopher Mcdermid, “Availability and Load Balancing in Cloud Computing”, International Conference on Computer and Software modeling IPCSI, 2011.
[3] Ratan Mishra and Anant Jaiswal, “Ant Colony Optimization: A solution of Load Balancing in Cloud”, International Journal of Web & Semantic Technology (IJWesT), April 2012.
[4] Liang Liu, Hao Wang, Xue Liu, Xing Jin, WenBo He, QingBo Wang, Ying Chen, "Green Cloud: A New Architecture for Green Data Center", ACM Journal, 2009.
[5] V.Srimathi, D.Hemalatha, R.Balachander, "Green Cloud Environmental Infrastructure", International Journal of Engineering and Computer Science, December 2012.

Cite this article as: M. Aruna, Dr. D. Bhanu, Dr. S. Karthik. “Load Balancing Algorithms in Cloud Environment.” *International Conference on Systems, Science, Control, Communication, Engineering and Technology (2015)*: 50-54. Print.

[6] Venubabu Kunamneni, "Dynamic Load Balancing for the cloud", International Journal of Computer Science and Electrical Engineering, 2012.
[7] Karanpreet Kaur, Ashima Narang, Kuldeep Kaur, "Load Balancing Techniques of Cloud Computing", International Journal of Mathematics and Computer Research, April 2013.
[8] Dr.Hemant S.Mahalle, Prof. Parag R. Kaveri, Dr. Vinay Chavan, "Load Balancing on Cloud Data Centers", International Journal of Advanced Research in Computer Science and Software Engineering, January 2013.
[9] Shreyas Mulay, Sanjay Jain, "Enhanced Equally Distributed Load Balancing Algorithm for Cloud Computing" International Journal of Research in Engineering and Technology, June 2013.
[10] Nidhi Jain Kansal, Inderveer Chana, "Cloud Load Balancing Techniques: A Step Towards Green Computing", International Journal of Computer Science, January 2012.
[11] S.Mohana Priya, B.Subramani, "A New Approach for Load Balancing in Cloud Computing", International Journal of Engineering and Computer Science, May 2013.
[12] O.M.Elzeki, M.Z.Reshad, M.A.Elsoud, "Improved Max-Min Algorithm in Cloud Computing", International Journal of Computer Applications, July 2012.
[13] Jeffrey M. Galloway, Karl L.Smith, Susan S. Vrbsky, "Power Aware Load Balancing for Cloud Computing", World Congress on Engineering and Computer Science, 2011.

Cite this article as: M. Aruna, Dr. D. Bhanu, Dr. S. Karthik. "Load Balancing Algorithms in Cloud Environment." *International Conference on Systems, Science, Control, Communication, Engineering and Technology (2015)*: 50-54. Print.