Ramalingam Sugumar, International Journal of Computer Science and Mobile Applications,
Vol.6 Issue. 3, March- 2018, pg. 1-5 ISSN: 2321-8363
Impact Factor: 5.515

ASLBACC – A Study on Load Balancing Algorithms in Cloud Computing

Ramalingam Sugumar[1]
[1]Professor & Deputy Director, Christhu Raj College, Tiruchirappalli, Tamil Nadu India, E-mail: rssugusakthi1974@gmail.com

Abstract: ***Cloud Computing is change the IT industry, altering the technique software and hardware is used and afford. Hardware infrastructure, software developing platforms and software request are afford as on-demand services to consumers. These services are identified as Infrastructure as a Service (IaaS), Platform as a Service, and Software as a Service. Cloud Computing build the on-demand procedure of computing resources such as bandwidth, storage or computational power and software applications accessible on a pay-as-you-use standard for end users and activity. As the numbers of users are increasing on the cloud, the load balancing has become the challenge for the cloud provider. Load balancing is a main challenge in cloud environment. Hence efficient utilization of resources must be important and for that load balancing plays a vital role to get maximum benefit from the resources. In this paper we are studying various load balancing algorithm and issues related to them in cloud computing.***
Keywords: Hardware, Software, Cloud Computing, Load Balancing.

I. INTRODUCTION

Cloud computing is the best ever developed technology in the IT industry and a new delivery technique for the services on pay per use basis. According to a variety of standard definitions of cloud computing. Cloud Computing has become one of the popular technologies adopted by both industry and academic. It providing a flexible and efficient way to store and retrieve the data. The main problem is to schedule the incoming request in a way so with minimum response time, efficient resource utilization and at the same time resources should not be underutilized. Cloud Computing system are heavily rely on term virtualization that improves the power efficiency of datacenters and enables virtual machines to single physical server.

It delivers all services through the internet dynamically when user demands, such as operating system, network, storage, software, hardware and resources. These services are classified into these types: Software as a Service (SaaS), Platform as a Service (PaaS), Infrastructure as a Service (IaaS), and Network as a service (Naas), Human as a service (Haas) and anything as a Service (Xaas). Cloud computing domain is divided into three categories such as Public, Hybrid and Private cloud. The system of Cloud computing as shown in given figure 1.[1]

Fig1: cloud computing system.

1.1 Infrastructure as a service (IaaS)
It is the most basic service model of cloud known as Infrastructure as a Service (IaaS). It offers hardware as a service to the organization such as Server space, Network equipment, CPU cycles, physical virtual machines and other resources.
1.2 Software as a Service (SaaS)
In SaaS, the various services are provided to the user such as software applications from different cloud providers through the Internet. The customer uses software on a pay-per use basis. It is referred as "on-demand software".
1.3 Platform as a Service (PaaS)
In PaaS models, Cloud service providers deliver a computing platform such as operating system support, programming language execution and web server. Examples are GAE and Microsoft Azure.
Public Cloud: It can be access by anyone from anywhere all over the world. Examples of this cloud are Google's cloud which is open for all after specific SLA between provider and user. It is available on a subscription basis.
Private Cloud: In this type of cloud company's employee can access to company's own data or a partner employment.
Hybrid Cloud: It is a combination of both public as well as private cloud. [1]

II. LOAD BALANCING IN CLOUD COMPUTING

Load balancing is a new technique that provides high resource time and effective resource utilization by assigning the total load among the various cloud nodes, side by side it solves the problem of overutilization and underutilization of virtual machines. Load balancing resolve problem of overloading and focuses on maximum throughput, optimizing resource utilization and minimize response time. Load balancing is the pre requirements for maximizing the cloud performance and utilizing the resources efficiently. In utilization of clouds has been improved by a resource allocation method having preemptable task execution.

The load balancing is an efficient and critical concept in cloud computing and it helps in utilizing the resources optimally, therefore minimizing the consumption of resources. Thus load needs to be distributed over the nodes in cloud-based architecture, so that each resource does the equal amount of work at any point of time that is performed by a load balancer. The load balancer determines the various request allocation to different servers. The load balancer uses various algorithm which server should take the request.[2]

III. CLASSIFICATION OF LOAD BALANCING ALGORTIHM

In cloud computing, different load balancing algorithm have been proposed. The main purpose is to achieve high throughput and minimum response time.
Generally load balancing algorithm is of two types
A. Static load balancing algorithm
B. Dynamic load balancing algorithm
3.1 Static Load Balancing Algorithm
The load does not depend on the current state of the system but it requires knowledge about the application and resources of the system. Static load balancing is a load balancing algorithms that distributes the workload based strictly on a fixed set of rules such as input workload.
There are four different types of Static load balancing techniques: Round Robin algorithm, Central Manager Algorithm, Threshold algorithm and randomized algorithm.

3.2 Dynamic Load Balancing Algorithm

Dynamic algorithm are more flexible than the static algorithm and they doesn't rely on prior knowledge but depends on current state of the system. In a distributed system, dynamic load balancing has two different ways: distributed and non-distributed. In the distributed one, this algorithm is executed by all nodes present in the system and the task of load balancing is shared among these servers. The interaction among nodes to achieve load balancing can take two forms: cooperative and non-cooperative. In the first one, the nodes works side-by-side to achieve a common objective which means is to improve the overall response time, etc. In the second form, each node works independently toward a goal local to it.[2]

Existing load balancing algorithms

To design an effective load balancing policy and to determine how to increase the cloud resource usage are the two main goals of a cloud service provider. The VM scheduling algorithms for load balancing helps in allotment of VMs efficiently on need. Basically, a VM load balancing algorithm decides which VM is to allocate when request is made by cloud consumer. Numerous VM load balancing algorithms that have been proposed are discussed here

Round Robin VM Load Balancing Algorithm

It is a very simple load balancing algorithm that places the newly coming cloudlets on the available virtual machines in a circular manner. The major advantages of this algorithms is its simplicity and easy implementation. The main drawbacks are that it requires the prior knowledge of user tasks and system resources & it do not make use of current state of the system.

Throttled VM Load Balancing Algorithm

It is a dynamic approach. In this, user submits its request to the Data Center Controller (DCC). Data Center Controller asks the VM Load Balancer to determine the appropriate virtual machine that can handle that much workload easily. Throttled VM Load Balancer keeps a virtual machine list and their status (available/busy). If a suitable VM is found on memory space, cores or availability basis, then throttled VM Load Balancer accept the cloudlet request and allot the cloudlet request over that virtual machine. Otherwise, client have to wait in the waiting queue untill a suitable VM becomes available. Among all, it is best approach for load balancing, since it maintains the present state of all VMs in data center. But the major drawback is that it works properly only if all VMs in a data center have same hardware configuration.[3]

ESCE VM load Balancing Algorithm

ESCE stands for Equally Spread Current Execution. It is also called Active VM Load Balancing algorithm. This algorithm is based on spread spectrum technique. As the name indicate, it equally distribute the workload on each VM in data center. A job queue keeps all the cloudlet requests that needs the VM for their execution. ESCE VM Load Balancer (VMLB) also maintains a list of virtual machines. The VM Load Balancer continuously check the job queue and VM list. If a VM is found free, then cloudlet request will be allotted over that VM. At the same time, VMLB inspect the overloaded VMs. If any virtual machine is found overloaded, then VMLB move some load to an idle or an under loaded virtual machine, so as to reduce some load of overloaded VM. The main drawback is high computational overhead.[3]

IV. RELATED WORK

An enhanced priority based HTV load balancing algorithm to perform the effective and reliable resource allocation of the tasks on the servers in cloud computing environment. This algorithm considers the three parameters load on the server, current performance of server and time limit of the tasks as its prime parameters. In this, algorithm is calculating the load and performance factor of each virtual machine and then allocating the incoming tasks to various virtual machines according to their time limit and stand-by time. According to the algorithm it conclude that by considering the time limit of the tasks, it will increase the throughput and performance.[4]

Minimum Make span algorithm produces higher throughput by migrating resource to unallocated node. This algorithm compares with other load balancing algorithms like Min-Min, Max-Min and RASA. It produces the minimum make span.[5]

A load balancing approach that amplify the physical resource utilization and curtail the energy consumption. To calculate the performance of the approach it is compare with the existing load balancing approach and judge against the number of migration, energy consumption. CloudSim simulator is use as a simulation tool to create the cloud environment. Experiment results say that this approach gives better result as compare to the existing load balancing method.[6]

Load Balancing is required to properly manage the resources of the service contributor. Load balancing is a technique to distribute the workload among many virtual machines in a Server over the network to achieve optimal resource consumption, decrease in data processing time, decrease in average response time, and avoid overload. The objective of effective and efficient and enhanced composite scheduling algorithm that can be used to maintain the load and provides efficient resource allocation techniques. This Composite approach is applied for load balancing using Equally Spread Current Execution (ESCE) and Throttled algorithms.[7]

The distributed dynamic priority based algorithm is used for balancing the load on instances effectively and to improve the system consistency, minimum response time and increase the throughput. Allocating the resources on virtual machines based on priority achieves the better response time and processing time. Load balancing ensures all instances in a node in the networks do the equal amount of work at any instant of time. Priority based resource provision to improve the utilization of resources and reducing response time of cloud services.[8]

A new burstness-aware load balancing algorithm which can adapt to the variation in the request rate by adopting two load balancing algorithms: RR in burst and Random in non-burst state. Fuzzy logic is used in order to assign the received request to a balanced VM. The algorithm has been evaluated and compared with other algorithms using Cloud Analyst simulator. Results show that the algorithm improves the average response time and average processing time in comparison with other algorithms.[9]

The load balancing model use improved throttle algorithm. This improved throttle algorithm works well even though underlying capacity of each VM is different because the hardware configuration of VMs is different. So improved throttle algorithm is taking decision of VM selection with hash table with more parameters such as expected response time and loading condition. Now Expected response time can be calculated using CPU utilization of VM. Using improved throttled load balancing algorithm with less overhead, results better VM allocation and increased number of user request handling, thus reducing the rejection in the number of requests arrived at datacenter of cloud. So this throttled algorithm is implemented and tested. Both of these algorithms are compared in terms of Response time, Datacenter Request Servicing time and Cost in Cloud Analyst and Results prove the performance of algorithm.[10]

As cloud computing is a vast and emerging area, there are many issues related to it. Load balancing is one of them and it needs to be solved to provide better customer service. A new enhanced load balancing algorithm is implemented in heterogeneous cloud environment using Cloud Analyst tool. By considering the desired parameters in graphs and tables it easily identify that the overall response time and data centre processing time is improved. Simulation result shows reduction up to 20-30%in the time. According to the experiment and analysis proposed enhanced algorithm has the best integrate performance.[11]

A weighted based optimized load balancing approach is for distributing of incoming jobs uniformly among the servers or virtual machines. The performance is analyzed using Cloud simulator and compared with existing Round Robin and EIPR algorithms. Simulation results have demonstrated that the algorithm has distributed the load uniformly among virtual machines.[12]

An improved scheduling algorithm is introduced after analyzing the traditional algorithms which are based on user priority and task length. The approach considers two parameters: (1) Task Length and (2)User Priority. The algorithm is based on credit system. Each task is assigned a credit based on their task length and priority. In the actual scheduling of the task, these credits will be considered. The experimental results show a considerable improvement in the utilization of resources.[13]

A framework for global server load balancing of the Web sites in a cloud with two-level load balancing model. This framework is intended for adapting an open-source load-balancing system and the framework allows the network service provider to deploy a load balancer in different data centers dynamically while the customers need more load balancers for increasing the availability.[14]

V. CONCLUSION

The study focus on Cloud computing load balancing algorithms. This is the main issue. The Load balancing is a main task in cloud computing for efficient utilization of resources. The main objective of load balancing is to maximize resource utilization, increase the performance of the cloud system thereby reduction in the energy consumption and satisfy the user requirement. The performances of various algorithms have been discussed in this paper.

REFERENCES

[1] Harshada Raut , Kumud Wasnik, "A SURVEY ON LOAD BALANCING ALGORITHMS IN CLOUD COMPUTING" International Journal of Advanced Technology in Engineering and Science, Volume No 03, Special Issue No. 01, March 2015 .

[2] Reena Panwar ,Bhawna Mallick, "A Comparative Study of Load Balancing Algorithms in Cloud Computing", International Journal of Computer Applications (0975 – 8887) Volume 117 – No. 24, May 2015 .

[3] Mamta Khanchi, Sanjay Tyagi, "An Efficient Algorithm for Load Balancing in Cloud Computing" International Journal of Engineering Sciences & Research Technology, June, 2016.

[4] Divya Garg, Urvashi Saxena," Dynamic Queue Based Enhanced HTV Dynamic Load Balancing Algorithm in Cloud Computing", International Journal of Innovative Research in Science, Engineering and Technology ,Vol. 5, Issue 1, January 2016.

[5] Harish Chandra, Pradeep Semwal, Sandeep Chopra," load balancing in cloud computing using a novel minimum makespan algorithm", International Journal of Advanced Research in Computer Engineering & Technology ,Volume 5, Issue 4, April 2016.

[6] Sachin Soni, Praveen Yadav, "A Load Balancing Approach to Minimize the Resource Wastage in Cloud Computing", International Advanced Research Journal in Science, Engineering and Technology ,Vol. 3, Issue 3, March 2016.

[7] Navtej Singh Ghumman, Rajesh Sachdeva," an efficient approach for load balancing in cloud computing using composite techniques", International Journal of Research in Engineering and Applied Sciences, volume 6, issue 2 February, 2016.

[8] G.Suryadevi , D.Vijayakumar, R.SabariMuthuKumar, Dr. K .G. Srinivasagan, "An Efficient Distributed Dynamic Load Balancing Algorithm for Private Cloud Environment", International Journal of Innovative Research in Science, Engineering and Technology, Volume 3, Special Issue 3, March 2014.

[9] Sally F. Issawi, Alaa Al Halees, Alaa Al Halees, "An Efficient Adaptive Load Balancing Algorithm for Cloud Computing Under Bursty Workloads", Engineering, Technology & Applied Science Research Vol. 5, No. 3, 2015.

[10] Slesha Nayak , Mr.Pragnesh Patel, "Analytical Study for Throttled and Proposed Throttled Algorithm of Load Balancing in Cloud Computing using Cloud Analyst", IJSTE - International Journal of Science Technology & Engineering | Volume 1 | Issue 12 | June 2015.

[11] Lakhvir Kaur, Simarjit Kaur, "Enhanced Load balancing Algorithm in Heterogeneous Cloud Environment", international journal for advance research in engineering and technology, Volume 3, Issue VIII, Aug 2015.

[12] B. Nithya Nandhalakshmi , Mahalingam, "Efficient Load Balancing in Cloud Computing Using Weighted Throttled Algorithm", International Journal of Innovative Research in Computer and Communication Engineering ,Vol. 3, Issue 6, June 2015.

[13] Antony Thomas, Krishnalal G, Jagathy Raj V P, "Credit Based Scheduling Algorithm in Cloud Computing Environment", International Conference on Information and Communication Technologies (ICICT 2014).

[14] Po-Huei Liang1 and Jiann-Min Yang. "Evaluation of two-level Global Load Balancing Framework in cloud environment" International Journal of Computer Science & Information Technology (IJCSIT) Vol 7, No 2, April 2015.

Navtej Singh Ghumman, Rajesh Sachdeva," an efficient approach for load balancing in cloud computing using composite techniques", International Journal of Research in Engineering and Applied Sciences, volume 6, issue 2 February, 2016.

