International Journal of Science and Research (IJSR)
ISSN (Online): 2319-7064
Index Copernicus Value (2015): 78.96 | Impact Factor (2015): 6.391

Survey of Load Balancing on Cloud Computing

Prabha Patel[1], Umesh Lilhore[2]

[1]M. Tech Student, Computer Science and Engineering Department, NIIST Bhopal Madhya Pradesh, India

[2]A.P. & M. Tech Coordinator, Computer Science and Engineering Department, NIIST Bhopal Madhya Pradesh, India

Abstract: ***Cloud computing is a most modern development in huge level data processing. It helps in providing shared resources. It gives sustain to the distributed parallel processing. Alternatively providing a well-organized response time is a most important demanding concern in the distributed systems. Cloud computing make available data on the pay per use basis and removes they require of having one's own device. Less significant response time is required for distributed computing and well-organized load balancing is one of the majority essential problems that can develop response time. In this paper, we first understand about load balancing, its importance and desired characteristics in cloud. Here discuss the proposed load balancing algorithm, which give good response time.***

Keywords: Cloud computing, Load Balancing, Data Center, Virtual Machine (VM), Round Robin

1. Introduction

Cloud computing is the ability of using various computing resources through the internet including applications and storage services. The distributed collection of the resources is hosted by the cloud provider. According to the National Institute of Standard and Technology [1], Cloud Computing is defined as a model for providing convenient and on demand access to the distributed pool of resources including networks, storage, services etc.These services require minimal effort. From the business point, cloud computing is a new business model that has not been possible without the cloud. The main attribute of the cloud computing is the elasticity, means its tendency to grow and reduce the computation as per the requirements. Cloud technology is getting popular and large more number of users are using the cloud services. Balancing load is most important task of cloud service provider with smallest amount reaction time and enhanced approval of service level agreement of consumers. Many algorithms like Round Robin, Active Monitoring and Throttled are utilized to distribute the consignment to virtual machine running in data centre with smallest amount answer time and enhanced throughput. But limitations such as over/under use of virtual machine and communication interruption are not addressed noticeably. Usually load balancing algorithms can be confidential into static and dynamic algorithms.

Static load balancing algorithms do not depend on the state and behavior of system while allocating the consumer demand. These algorithms are mostly suitable for homogeneous and stable atmosphere but in dynamic background they generate a reduced response time. Dynamic Algorithms depends on the state and behavior of the system, while allocating the user request they are more flexible and take different types of system attributes. The other attribute is it is scalable, means its ability to balance the increased demands of the CPU storage, bandwidth etc. Cloud Computing provides the secure access to the applications. But high level security is a challenge for the cloud developers.

Figure 1: Cloud Computing

The term "cloud" was originated in 1990"s, when for the data communication Virtual Private Networks (VPN) was used to provide services [2]. Prior to the VPN, companies provided the dedicated point to point service which resulted in the waste of the bandwidth and thus resulted in the increase of the overall cost. It is a recent trend in the IT sector.

2. Load Balancing

A web server has the dispatcher to balance the incoming request to the servers. The main aim of dispatcher in load balancing is to transfer the request to the server that is available at that time. The front end is responsible to balance the requests by making decisions regarding the transfer to load is transferred efficiently to the server which can process the request at that moment. Web server's load information is used in making the decisions by the front end. The front end may send a series of the requirements to the number of the web servers. An example of the load balancing has exposed in the Figure 2. In this, the front end distributes the load to the server with least load at that moment. The server and the front end exchanges the in sequence concerning the load with each other to make the effective decisions. With the help of the correct decision about the load balancing, the service quality is improved and the

system becomes more robust. The load balancer decides how to forward the requests and the decision is made correspond to the CPU load percentage on the particular virtual machine.

Load balancing is the procedure of improving the presentation of a parallel and distributed system all the way through a reorganization of load between the processor. The most important objective of load balancing is to match the workload among the nodes by reduce execution time, minimizing communication interruptions, take full advantage of resource consumption and maximizing throughput.

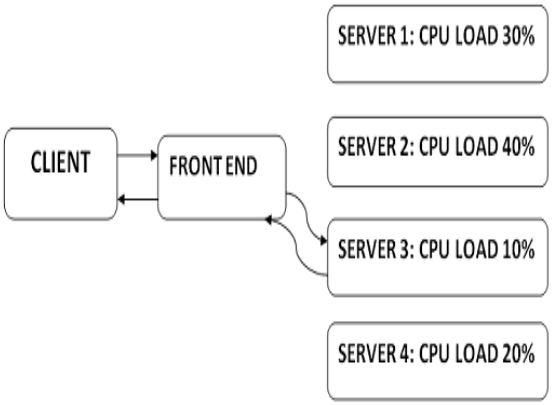

Figure 2: Load Balancing

A fundamental illustration of load balancing in our daily life can be associated to websites. Without load balancing, customers could knowledge setbacks, timeouts and achievable long scheme reactions. Load balancing solutions more often than not are appropriate unnecessary servers which help an improved distribution of the communication load so that the website ease of use is decisively resolved [5]. In a distributed computer system background, where two or more independent computers are related via a communication network, resource sharing is a large amount attractive characteristic. Accepting sharing data and I/O devices, joins of a distributed system could additional get better scheme presentation by sharing their computational power. Load balancing is a method that allows works to move from one computer to another within the share out system. These produces earlier job services e.g., minimize job response time and develops resource utilization. Various learning has shown that load balancing among nodes of a distributed scheme highly advances system presentation and enhances resource utilization.

To ability to distribute the coming load among the virtual machines in cloud computing environment with the intention that the general cloud presentation can be enhanced. For this purpose, researchers have suggested numerous load balancing algorithms such Round Robin, Active Monitoring Load Balancer [4], Throttled load balancer [4] etc. and these algorithms are broadly used for executing the users requests within the minimum response time. Load balancing algorithms are categorized into two kinds that are static and dynamic algorithms. Static algorithms are used in an environment where dynamic changes are not made at run time e.g. Round Robin (RR) [4]. Dynamic load balancing algorithms are used where the changes are made at run time e.g. weighted least connection (WLC).

3. Existing Technique for Load Balancing Algorithms

a) **Round Robin Load Balancer** This load balancing rule [6] attempts to sustain identical workloads on all the available VMs. These algorithm efforts on random selection of the virtual machines. The datacenter organizer allocates the demands to a list of VMs on a revolving basis. The initial demand is assigned to a VM picked indiscriminately from the group and then the Data Center controller allocates the demands in a rounded order. Once the VM is allocated the demand the VM is moved to the end of the list.

b) **Active Monitoring Load Balancer** Active VM Load Balancer sustains data about each VMs and the number of calls at present assigned to which VM. When a request to distribute a novel VM appears it recognize the smallest amount weighted down VM. If there is more than one, the primary recognized is chosen. Active VM Load Balancer precedes the VM id to the Data Center Controller the data Center Controller throws the demand to the VM recognized by that id. Data Center Controller informs the Active VM Load Balancer of the novel allocation [8].

c) **Throttled Load Balancer** Throttled algorithm is entirely supported on virtual machine. In this user initial demanding the load balancer to ensure the accurate virtual machine which right to use that load easily and perform the operations which is give by the client or customer. In this algorithm the client initial demands the load balancer to discover an appropriate Virtual Machine to complete the need operation [8].

Techniques	Nature	Response Time	Performance
Round Robin	static	no	yes
Throttled	dynamic	no	yes
Active Monitoring	dynamic	yes	yes
Least Weighted Connection	dynamic	yes	no

4. Related Work

Ankit kumar,Mala kalra et al.[9] In this research work, author have presented Modified Active Monitoring Load Balancer algorithm for efficiently distributing the jobs among the available VMs. Load balancing gives a method of accomplishing the appropriate use of resources and enhanced customer approval. Consequently, make use of a suitable load balancing algorithm is necessary for choosing the virtual machines or servers. This paper center of attentions on the load balancing method which allocates the external jobs among VMs optimally in cloud data centers. The proposed algorithm is implemented using Cloud Analyst simulator and the presentation of the recommended algorithm is compared with the existing three algorithms which are preexists on the basis of response time. The experiment accepted out in the paper give you an idea about that the recommended algorithm achieves better than the existing algorithms.

Asha ML,Neethu Myshri R et al.[10] Here author has comparing cloud environment with different methods by service Broker policy by giving various services to client demands; here they are directed on neighboring Data center.

If similar requests has come from unusual Data centers from unusual client Bases then examination will be make available by recognizing which is adjacent to Data center. And as well they are using various progress methods like how many instructions has to be executed per second be appropriating exacting load balancing methods like uniformly increase existing implementation load, Throttled Load Balancer and here they are focused on Round Robin Policy means services of different users will be provided based on Fixed time slices called Round Robin policies. They are also judgment different Internet characteristic like Transmission Delay, Bandwidth available for different areas. Overall response time by area and average processing time for Data center loading.

In accessible Virtual Machine Load balancing rule drawback of cloud is that they don't keep the circumstances of the earlier allocation of virtual machine to a request from the customer and the VM Load balancing algorithm requires implementation each time a novel demand for VM allocation received from user.

Bhavana Bidarkar et al.[11] The proposed algorithm is better than existing Round Robin algorithm for implementing VM scheduling algorithm in CloudSim toolkit. Assuming the application is positioned in one data centers having virtual machine. Experimental results demonstrate that Round Robin VM Load Balancing method develops the performance by consuming less time for scheduling virtual machine.

Slesha Nayak et al.[12] In this paper author has proposed load balancing representation will utilize advanced throttle algorithm. This advanced throttle algorithm efforts well even though fundamental facility of each VM is different because the hardware configuration of VMs is unusual. So enhanced throttle algorithm is taking assessment of VM selection with hash table with additional limitations for example predictable response time and loading circumstance. At this time predictable response time can be find out using CPU utilization of VM. By means of enhanced throttled load balancing algorithm with less operating cost, enhanced results VM allocation and increased number of user request handling, thus reducing the negative response in the number of requests reach your destination at datacenter of cloud.

Existing effort on load balancing algorithms there are numerous drawbacks like enhance in waiting time, context switching and enhance in response time. For the reason that of the increase in waiting time the charge also increases. Consequently turnaround time also enhances.

Lavanya M. et al.[13] Here author proposes an efficient scheduling algorithm which can sustain the load balancing and can make available enhanced progressed plan through well-organized job scheduling, customized resource allocation methods and decrease the power usage and context switching between the servers. Enhanced Throttled Load Balancing Algorithm provides get better consequences to Throttled Load Balancing Algorithm.

Meenakshi Sharma et al. [14] Proposed a new Efficient Virtual Machine Load Balancing Algorithm. The projected algorithm discovers the projected response time of each resource (VM). When a demand from the data center manager reach your destinations algorithm launch the ID of virtual machine having lowest amount response time to the data center manager for allocation to the novel request. The algorithm informs the allocation table raising the allocation count for that VM. When VM terminates processing of demand data center controller inform algorithm for VM de-allocation. The experimental result evaluates recommended VM load balancing algorithm with the Throttled Load Balancer and Active VM Load Balancer. The well-organized collection of a VM enlarges the general presentation of the cloud environment and also reduces the usual response time and cost evaluate to Throttled Load Balancer and Active VM Load Balancer.

Jasmin James et al. [15] proposed Weighted Active Monitoring Load Balancing (WALB) Algorithm which has an enhancement over the Active VM Load Balancer. This algorithm generates VM"s of unusual dispensation power and assigns weighted calculate according to the computing power of the VM. WALB sustains index table of VM"s, connected weighted count and number of appeal at present allocated to each VM. When a ask for to allocate a VM reach your destinations from the Data Center Controller this algorithm recognize the smallest amount loaded and most influential VM according to the weight assigned and arrivals its VM id to the Data Center Controller. The Data Center Controller throws a demand to the recognized VM and inform the algorithm of allocation. The algorithm increases the count by one for that VM. When VM terminates processing, algorithm decreases the count of that VM by one. The experimental consequence demonstrates that the projected algorithm accomplishes better presentation issues for example response time and processing time but the algorithm does not think about process duration for each entity request.

5. Following Problem Identified Based On Survey

Round Robin-The Round robin uses the time slice mechanism. Each node is allotted with a time slice in which they have to perform their task. This algorithm is fewer throughputs compared to other load balancing algoriths.This algorithm simply allots the job in round robin fashion to virtual machines, this method does not concentrate on load. In this algorithm it will not check how server loaded, it will directly assign the request whenever its turn comes so this is the reason, some server are heavy loaded while some are lightly loaded.

Throttled Method- Throttled algorithm maintains an index table which contains VMid, state of virtual machines and load of virtual machine. When request arrive it allocate to available VMs. In case VMs are not available this method maintains a queue for arrived requests. There is problem such as more priority request is waiting in the queue and least is running. There has been an attempt made to improve the response time and achieve efficient usage of available virtual machines.

Active Monitoring Load Balancing- This method also maintain a parse table, which have VMid, state (available/busy) and load on the VMs. Problem on this method is, when two VMs have same load it allocate request

to near VM , but there is a case may be physical configuration of allocated VM is low.

Modified Active Monitoring Load Balancing- This method maintain a parse table to balance the load on server. When request arrive it check the least loaded VM and allocated VM to that request. In case if two VM have same load it check memory utilization of both and allocate to better VM. Problem of Modified Active load balancing is it does not match the physical configuration (e.g. speed, bandwidth, memory) of two selected VMs.

6. Proposed Methodology

1) The Proposed Methodology implemented here contains a table of Virtual Machines id, State (Busy / Available), and the amount of demands at present assigned to the VMs. Initially all VMs are in available State.
2) The DCC (Data Center Controller) take delivery of a novel request.
3) The DCC then queries the Load Balancer to distribute the upcoming request.

1. Initialize all the VMs by some Constant Values but should be equal at all VM and probably 0 is best.
2. While not stop condition
3. Generate m values for a tour
4. Perform sorting on values by their length such that $l_1<=l_2<=l_m$
5. Value at Parse Table of the VM is update is weighted according to the rank R of the VMs
6. Then best VMs is chosen based on the rank R
7. If W is the weight of the trail level involvement of the best tour length than it should not be exceeded by any other VMs weight.

4) The Load Balancer parses the table from the top and finds least loaded VM whose state is available.
5) If there are additional than one
   a) The Load Balancer checks memory utilization of each VM and finds highest priority virtual machine.
   b) The Load Balancer returns the VMid of highest priority VM to Data Center Controller.
   c) DCC notifies the Load Balancer of new allocation.
   d) Load Balancer updates the information of the table accordingly.
6) Return VMid to Data Center Controller.
7) DCC sends the request to the Chosen VM
8) After finishing the request by the VM, the DCC receives the response cloudlet and notifies the Load Balancer for the de-allocation of the VM.
9) Whenever the request of the VM is processed the allocation table of VM is decremented by one by the Load Balancer.
10) If there are more requests, go to step 3 else continue from 2.

7. Conclusion

In this paper we review various existing load balancing method using on cloud environment but the main issue remained is load balancing and the resource utilization.

References

[1] P. Mell and T. Grance, "The NIST Definition of Cloud Computing", National Institute of Standards and Technology, Sept. 2011.
[2] J. Harauz, L. M. Kaufinan and B. Potter, "Data Security in the World of Cloud Computing", in proc. of IEEE Security & Privacy, co published by IEEE Computer and Reliability Societies, pp. 61-64, July 2009.
[3] Chaczko, Z., Mahadevan, V., Aslanzadeh, S. and Mcdermid, C., "Availability and Load Balancing in Cloud Computing", International Conference on Computer and Software Modeling, IACSIT Press, Singapore, 14, 134-140, 2011.
[4] W. Bhathiya, "CloudAnalyst a CloudSim-based tool for modelling and analysis of large scale cloud computing environments", MEDC Project, Cloud Computing and Distributed Systems Laboratory, University of Melbourne, Australia, pp. 1-44, June 2009.
[5] M. B. Mollah, K. R. Islam and S. S. Islam, "Next Generation of Computing through Cloud Computing Technology", in proc. of 25th IEEE Canadian Conference on Electrical and Computer Engineering (CCECE), pp.1-6, 2012.
[6] Alberto F. De Souza and Rajkumar Buyya, Computer Architecture and High Performance Computing, International Journal of Parallel Programming, Vol. 36, No. 1, Springer, Netherlands, April 2008.
[7] Raj, G. Comparative Analysis of Load Balancing Algorithms in Cloud Computing, 1(3), 120–124, 2012.
[8] Sharma, M., & Sharma, P. Efficient Load Balancing Algorithm in VM Cloud Environment, 8491, 439–441. -43, 2012.
[9] Ankit Kumar, Mala Kalra, "Load Balancing in Cloud Data Center Using Modified Active Monitoring Load Balancer" IEEE, 2016.
[10] Asha M L, Neethu Myshri R, "Performance Evaluation of Round Robin Algorithm in Cloud Environment" International Conference on Information and Communication Technologies (ICICT-2014).
[11] Bhavana Bidarkar, Shakheela Attikeri, Ramya S. Pure, "Round Robin Approach for Better VM Load Balancing in Cloud computing" International Journal of Engineering Innovation & Research Volume 3, Issue 4, ISSN: 2277 – 5668, 2014.
[12] Slesha Nayak Mr. Pragnesh Patel, Analytical Study for Throttled and Proposed Throttled Algorithm of Load Balancing in Cloud Computing using Cloud Analyst" International Journal of Science Technology & Engineering | Volume 1 | Issue 12 | June 2015.
[13] Lavanya M., Sahana V., Swathi Rekha K. and Vaithiyanathan V. "Adaptive Load Balancing Algorithm Using Modified Resource Allocation Strategies On Infrastructure As A Service Cloud Systems" ARPN Journal of Engineering and Applied Sciences, VOL. 10, NO. 10, JUNE 2015.
[14] Meenakshi Sharma, Pankaj Sharma, "Performance Evaluation of Adaptive Virtual Machine Load Balancing Algorithm", International Journal of Advanced Computer Science and Applications, pp 86-88, Volume 3, Issue2, ISSN: 2156-5570, 2012.

[15] Jasmin James, Bhupendra Verma, "Efficient VM Load Balancing Algorithm for a Cloud Computing Environment", International Journal on Computer Science & Engineering, pp 1658-1663, Volume. 4, ISSN: 0975-3397, September 2012.