ISSN 2348 – 9928
IJAICT Volume 4, Issue 6, June 2017 Doi:01.0401/ijaict.2014.08.01 Published on 05 (6) 2017

AN ASSESSMENT ON VARIOUS LOAD BALANCING TECHNIQUES IN CLOUD COMPUTING

Dr. V. Venkatesa Kumar
Assistant professor,

Ms. R. Revathi
PG Scholar,

Mr. M. Newlin Rajkumar
Assistant professor,

Computer Science and Engineering,
Anna University Regional Centre Coimbatore,
Coimbatore, Tamilnadu, India

Abstract— In the recent trend, the need of computing and storage are increasing tremendously. Thus leads to a big challenge in maintaining the load in cloud server, which can be resolved by using various load balancing techniques like security, fault-tolerance etc. To improve the load balancing, many researchers have proposed various techniques. In this paper, here the work proposes various load balancing schemes including Fast adaptive load balancing, Throttled Load Balancing, Active Monitoring Load Balancing, Min-Min, Max-Min, Dynamic and Adaptive Load Balancing, Opportunistic Load Balancing, Honeybee Foraging Load balancing, Ant colony optimization and listed out the comparisons of the above schemes.

Keywords— Min-Min, Max-Min, Honeybee, Load balaning

I. INTRODUCTION

Cloud computing is a term which involves delivering of the hosted services to the client over the internet in a secured way. The cloud computing can be categorized into four categories which mainly focuses on three categories namely Software as a service (SaaS), Platform as a service (PaaS), Infrastructure as a service (IaaS) [5].

Infrastructure as a service which provides the customer with application program interfaces (APIs) that allow the users to access, start and stop the virtual servers and configure those servers and storage. It also provides virtual server instances for the efficient access of the resources. Example for Iaas is Amazon web services. It is also called as pay-for-what-you-use model. [5]

Platform-as-a-service in the cloud will have the set of software development tools that can be hosted on the service provider's infrastructure [5]. Software Developers can create the applications on the service provider's platform over the Internet. Examples for Paas providers are Apprenda, RedHat OpenShift.

In the ***software-as-a-service*** cloud model, the vendor will supply the software products, hardware infrastructure and interactions with the user through a front-end portal. Benefits of the SaaS model include global accessibility, compatibility, easier administration, patch management and automatic software update. Fig1 indicates the cloud architecture.

The cloud can be of the following three types.

- Public Cloud
- Private Cloud
- Hybrid cloud

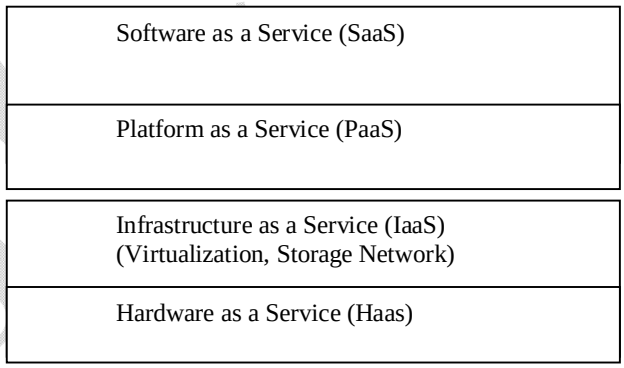

Fig 1. Cloud Architecture

II. LOAD BALANCING IN CLOUD ENVIRONMENT

Load balancing is also defined as new technique that are used to provide the maximum throughput and with minimum response time. Load balancing is divides the traffic between network and interfaces on a network socket (OSI layer).It is the main reason for the computer server clustering [6]. It is used to divide the work of a computer between two or more computers that may get work done in the same amount of time.
All the users get server faster. It can be implemented by using hardware, software or a combination of both. Usage of internet, website is more reason for load balancing, it distributes workloads across the computing resources like a computer, cluster, network link etc. It helps to avoid overload of any single resources. Load balancing is the great impact for performance.

The load can be also termed as CPU load, memory capacity, and delay or network load [8]. Process of distributing the load

ISSN 2348 – 9928
Doi:01.0401/ijaict.2014.08.01 Published on 05 (6) 2017

among various nodes ensures that all processor in system or everyone in network has the equal amount of work at any instant of time. The example for load balancing is shown in Fig 2.

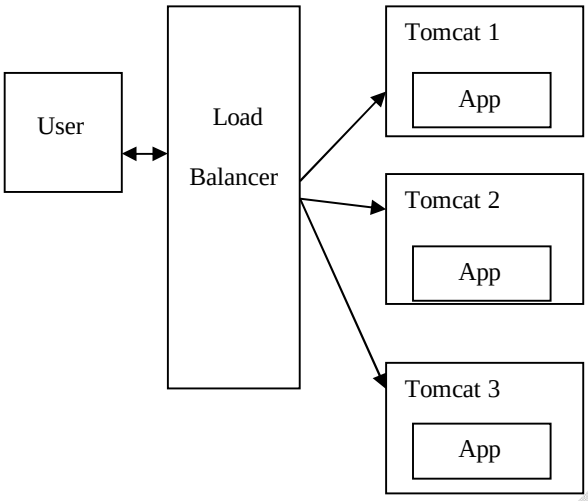

Fig 2. Load Balancing Example

Metrics for load balancing in cloud:

- Scalability
- Performance
- Resource Utilization
- Overhead Associated
- Response Time
- Fault Tolerance
- Throughput
- Migration time

III. RELATED WORK

3.1 A Fast adaptive load balancing method (FALB)

In Fast Adaptive Load Balancing method a binary tree structure is used to partition the region into sub-domains [1]. It is used to adjust the workload between the processors from local areas to global areas. In Fast Adaptive Load Balancing the binary tree mode can be used to partition the region. So that the tree contains parent nodes, leaf nodes and child nodes. The binary tree indexes of the cells on the left side are smaller than the indexes of the cells on the right side and the indexes on the top cells are smaller than the bottom cells. Based on the balancing structure, we can calculate the workload. The Fast Adaptive Load Balancing method has a faster balancing speed, and it takes less elapsed time and less communication cost. The advantages of Fast Adaptive Load Balancing method is having less communication overhead, higher efficiency and faster balancing speed. The Disadvantages of this method includes it cannot able to maintain the neighboring cells.

3.2 Throttled Load Balancing Algorithm (TLB)

In Throttled Load Balancing Algorithm, the load balancer maintains an index table which contains the number of virtual machines and their states. The State indicates that whether the Virtual Machine (VM) is available to use or it is busy. The Client will send a request to the data center to find a suitable VM for performing a particular job [4]. The data center will query the Load balancer for the allocation of a VM for that client. The load balancer will check for the index table for any of the available VM. If any VM is found, then it will load data centre. Then the data centre communicates the request to the VM and it will identify by the id. Then the data centre will acknowledge the load balancer of the new allocation and it will revise the index table for the new updates. If no VM is found for the client request, it will return -1 to the data centre. The data center will put the client request in the queue. If the VM completes the allocated task, then the request is acknowledged to the data centre and the load balancer will de-allocate the VM and update the index table. The example for Throttled load balancer is shown in Fig 3.

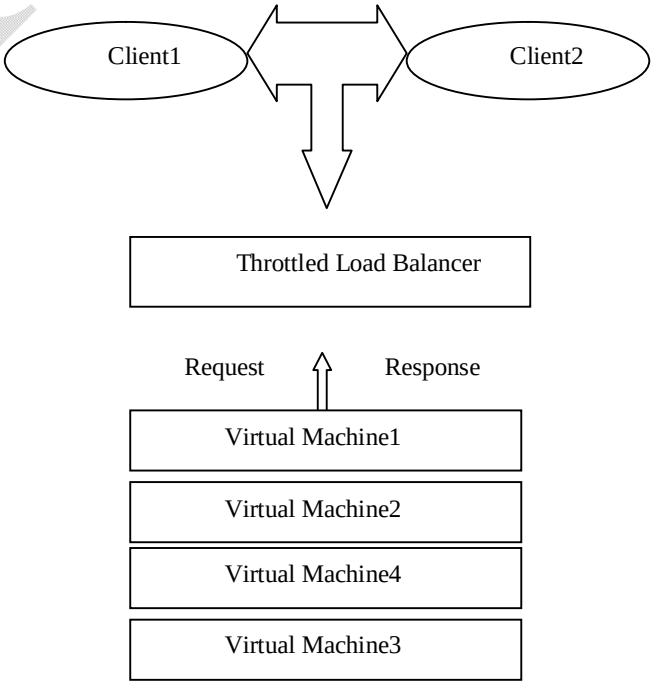

Fig 3. Throttled Load Balancer

ISSN 2348 – 9928
Doi:01.0401/ijaict.2014.08.01 Published on 05 (6) 2017

The Execution Time is calculated in three phases. First, the formation of the virtual machines and they are waiting for the scheduler to schedule the jobs in the queue. In Second phase, once jobs are allocated, the virtual machines in the cloud will start processing the jobs. The Third phase which will perform the cleanup and destruction of the VM's. The Advantages of this method provides the better performance and optimal job allocation.

3.3 Active Monitoring Load Balancer (AMLB):

The Active monitoring Load Balancer maintains the information about each VMs and the number of requests currently allocated to which VM. When a request is made by the client to access the VM, then the AMLB will identify the least loaded VM and to assign the VM to the client. If suppose more than one least loaded VM's are found, then the first identified will be selected. ActiveVmLoadBalancer will return the VM id to the Data Center Controller [3].

The data Center Controller sends the request to the VM identified by that id. Datacenter Controller notifies the ActiveVmLoadBalancer about the new allocation of a VM. The advantages of this method includes, it maximizes the throughput, minimize response time, and avoid overload of any single resource. It provides increased reliability.

3.4 Min-Min Algorithm

The Min-Min algorithm will do the jobs as quickly as possible to dispatch each task to virtual machines which will complete the task in the shortest time period. This Algorithm will execute the shortest jobs first and the long jobs will follow the shortest jobs [11]. Each shortest job will be executed in parallel. The shortest jobs will be scheduled first and to be executed [7]. The Main task of min-min algorithm is to establish the minimum completion time for every unscheduled job and then assigns the job with the minimum completion time.

The advantages of min-min algorithm will increase the resource utilization rate, improved quality of service, dynamic priority model and improved cost of service. Long tasks can be executed at a reasonable time and the user requirements can be fulfilled.

3.5 Max-Min Algorithm

In Max-Min Algorithm, Based on the priority the tasks are allocated for the larger and smaller tasks. This algorithm is used to minimize the waiting time of the shortest jobs through larger tasks can be executed first by slower resources. At the same time, the smaller tasks can be executed concurrently with the fastest resources where the Meta –tasks can have different execution time. The tasks can be completed within the minimum execution time and maximum completion time. The efficiency can be improved in Max-Min algorithm.

3.6 A Dynamic and Adaptive Load Balancing Strategy for Parallel File System (DALB)

In parallel file system the data are transferred between the storage devices and the memory. So the data management plays an important role in parallel file system. Load migration, availability of the system, scalability and network Transmission are the challenges that are faced during load balancing in the parallel file system.[2] In dynamic load balancing algorithm, the load applied in each I/O servers are different because the workload becomes varies continuously. So we need some decision making algorithms to perform better load balancing.

In this decision making system, first we will use a central decision maker. The central node is used as a decision maker so that if the central node fails, then the whole system performance will become low and the reliability of the system will be down. Secondly we will use a group decision maker. In this the total system should be divided into groups.

The advantage of this algorithm will reduce the communication cost. But global optimization is a major problem due to taking decision without considered the whole system load. Finally the Distributed decision maker is used. In this each server can take their own decision. So it provides a better scalability and availability. So the self acting load balancing algorithm (SALB) addresses the load prediction algorithm, effective distributed decision maker, efficient load collection mechanism, migration selection model and dynamic file migration algorithm for a better load balancing.

3.7 Opportunistic Load Balancing Algorithm (OLB)

This algorithm does not consider the current workload of the Virtual Machine. It makes each node to be busy. Every unexecuted task can be done in random order so that each task can be assigned to the node randomly. These processes will be in slow manner because it will not calculate the current execution time of the node.

3.8 Honeybee Foraging Load balancing Algorithm

Load balancing which helps to achieve load balance across heterogeneous virtual machine. The virtual machine is calculated by the current workload and it decides about the load or unloads. The priority of the task is given and it can be taken into consideration after removed from the overloaded VM [10]. The lightly loaded tasks are scheduled so that it will decrease the response time and the waiting time of the task.

Corresponding Author: Ms. R. Revathi, Annauniversity Regional Centre, Coimbatore, Tamilnadu, India.

ISSN 2348 – 9928
Doi:01.0401/ijaict.2014.08.01 Published on 05 (6) 2017

3.9 Ant colony optimization algorithm

This algorithm is mainly proposed for load balancing of nodes. It can be used in both cloud and grid networks. This approach aims efficient distribution of workload among the nodes.

The ant will start to move towards the source of the food from the head node when the request is initialized. Ant records their data for future decision making and it keeps records for every node and it makes visit to the record [9]. Every ant is build with their own individual result set and further built for giving the complete solution. It makes to update continuously with a single result set rather than own result set is updating. The cloud is a collection of many nodes which supports various types of application is used by the client on the basis of pay per use.

This ant colony is that work together into a foraging behavior. This ant works in searching of new sources food with the use of existing food sources to shift the food back to the nest. This mainly aims that efficiently distribution of the load among the nodes. It does not encounter the dead end of the movement to the node for building an optimum solution set. Table1 represents the comparisons of the above mentioned Load balancing schemes.

Techniques/ Parameter	FALB	TLB	AMLB	Min-Min	Max-Min	DALB	OLB	Honey bee	Ant colony
Throughput	YES	YES	YES	YES	YES	YES	NO	YES	YES
Scalability	NO	YES	YES	NO	NO	YES	NO	YES	YES
Response Time	YES	YES	NO	YES	NO	YES	NO	NO	NO
Fault-tolerance	NO	NO	YES	NO	NO	NO	NO	NO	NO
Performance	NO	YES	NO	YES	YES	NO	YES	YES	YES
Resource utilization	YES	NO	YES	YES	YES	YES	YES	NO	YES
Overhead	YES	NO	NO	YES	YES	NO	NO	NO	NO

Table1. Comparison of Various Load Balancing Schemes.

IV. CONCLUSION

In this paper, different load balancing schemes are analyzed. Fast adaptive load balancing, Throttled Load Balancing, Active Monitoring Load Balancing, Min-Min, Max-Min, Dynamic and Adaptive Load Balancing, Opportunistic Load Balancing, Honeybee Foraging Load balancing, Ant colony optimization and listed out the comparisons of the above schemes. Among these analyzes, this survey paper will help to determine the differences in various load balancing techniques and to make future improvements among those techniques.

References

[1] Dongliang Zhang, Changjun Jiang,Shu Li, "A fast adaptive load balancing method for parallel particlebased simulations", Simulation Modelling Practice and Theory 17 (2009) 1032–1042.

[2] Bin Dong, Xiuqiao Li, Qimeng Wu, Limin Xiao, Li Ruan, "A dynamic and adaptive load balancing strategy for parallel file system with large-scale I/O servers", J. Parallel Distribution Computing. 72 (2012) 1254–1268.

[3] Jasmin James, Dr. Bhupendra Verma "Efficient VM load balancing algorithm for a cloud computing environment" 72 (2016) 1258–1368.

[4] Tejinder Sharma, Vijay Kumar Banga,"Efficient and Enhanced Algorithm in Cloud Computing" 2 (2007) 1554–1262.

[5] http://searchcloudcomputing.techtarget.com/definition/cloud-computing

[6] http://searchnetworking.tech.target.com/definition/loadbalancing.

[7] D.Doreen Hephzibah Miriam and K.S.Easwarakumar, "A Double Min Min Algorithm for Task Metascheduler on Hypercubic P2P Grid Systems", IJCSI International Journal of Computer Science Issues, Vol. 7, Issue 4, No 5, July 2010

[8] Dharmesh Kashyap, Jaydeep Viradiya , "A Survey Of Various Load Balancing Algorithms In Cloud Computing" International journal of scientific & technology research volume 3, issue 11, november 2014.

[9] Ratan Mishra and Anant Jaiswal, "Ant Colony Optimization: A solution of Load Balancing in Cloud", International Journal of Web & Semantic Technology (IJWesT), April 2012.

[10] Dhinesh B. L.D , P. V. Krishna, "Honey bee behavior inspired load balancing of tasks in cloud computing environments", in proc. Applied Soft Computing, volume 13, Issue 5, May 2013, Pages 2292-2303.

[11] T. Kokilavani, Dr. D. I. George Amalarethinam "Load Balanced Min-Min Algorithm for Static Meta Task Scheduling in Grid computing" International Journal of Computer Applications Vol-20 No.2, 2011.