Available Online at www.ijcsmc.com

International Journal of Computer Science and Mobile Computing

A Monthly Journal of Computer Science and Information Technology

ISSN 2320–088X

IJCSMC, Vol. 5, Issue. 2, February 2016, pg.248 – 254

ANALYSIS OF VARIOUS LOAD BALANCING TECHNIQUES IN CLOUD ENVIRONMENT

Suriya Begum[1], Ram Charan[2], Sachin[3], Nikhit[4], Sai Pratap[5]

[1]Prof., Department of Computer Science and Engineering, New Horizon College of Engineering, VTU, India
[2]Student, Department of Computer Science and Engineering, New Horizon College of Engineering, VTU, India
[3]Student, Department of Computer Science and Engineering, New Horizon College of Engineering, VTU, India
[4]Student, Department of Computer Science and Engineering, New Horizon College of Engineering, VTU, India
[5]Student, Department of Computer Science and Engineering, New Horizon College of Engineering, VTU, India
[1] suriyabegumnotes@gmail.com

Abstract— Cloud computing has paved a revolutionary path in direction of distributed environment for accomplishing optimized performance, shortest response time, network resource utilization, and adaptability of service level agreement. Cloud computing has multiple benefits as well as it is also accompanied with certain serious technical loopholes. The proposed paper has focused on one such issue of load balancing. Load balancing is one of the main challenges in cloud computing. It is required to distribute the dynamic local workload evenly across all the nodes to achieve a high user satisfaction and re-source utilization ratio by making sure that every computing re-source is distributed efficiently and fairly.. Hence this paper illustrates various generic issues, and particularly to issues related to load balancing. Various techniques adopted in the past research work have been analyzed in this paper based on advantages and parameters used.

Keywords -- Cloud Computing, Elastic, Load Balancing, Service Level Agreement, Virtual Machine

I. INTRODUCTION

Cloud Computing is a new technology, which provides the various services dynamically to the all type of users in the entire globe. In cloud, data centres plays a very important role in providing the services requested by the users [10][11][12].

Cloud computing can be illustrated as the operation of heavy computing resources that includes hardware and software package that are delivered to clients as a service over an outsized scale network [13].

Cloud Computing main characteristics are convenient and On demand network, Allow access to a shared pool of computing resources, On demand self-service, Global network access, Distributed resource pooling, Scalable, Measured service.

A cloud computing is determined by these characteristics: On demand self-service, Global network access, distributed resource pooling, Scalable, Measured service[16].

This paper is includes Section II describes Cloud Architecture. Section III describes related literature work. Section IV describes existing load balancing techniques in cloud computing.

Section V describes comparison of existing load balancing technique. Section VI conclusion. Section VII future work.

II. CLOUD ARCHITECTURE

The architecture of cloud computing consists of three layers: application, platform and infrastructure[15].

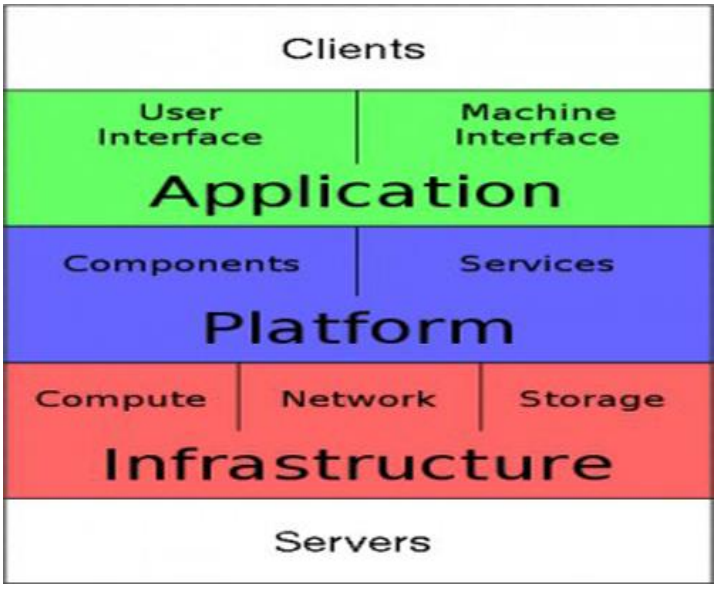

Fig 1:Architecture of Cloud Computing

III. RELATED LITERATURE WORK

The entire cloud environment is organized into centralized, distributed and hierarchical designs. There are various issues & bottle necks in cloud .One of the issue are Load Balancing[11][12].

A frequent use of cluster computing is to load balance traffic on high-traffic internet sites. An internet page request is forwarded to a "manager" server, that then determines that of many identical or extremely similar internet servers to forward the request to for handling. Having an internet farm (as such a configuration is usually called) permits traffic to be handled additional quickly. Clump has been offered since the Nineteen Eighties once it absolutely was employed in DEC's VMS systems [14].

Load balancing works by assigning the excess work from over performing node to the underperforming node until they become approximately equal. This makes sure that the nodes in the system are neither overloaded nor idle. To reduce the bottle necks, overheads and to improve the overall performance, load balancing is essential. In order to balance the node, we must focus on managing nodes and resources, grids, clusters etc. In Centralized balancing technique, it takes less time to analyse different resources. While in Distributed, if one node fails it doesn't affect other nodes, so no overloading takes place and is scalable, decentralized in nature and support Grid systems. Several Distributed balancing approach examples are Ant Colony, Map Reduce, Cat Swarm Optimization Load Balancing Algorithms[4][5][6]. Max-Min and Min-Min are few examples of Centralized Balancing techniques[3]. Map Reduce Algorithm[9] is an example for Hierarchical balancing.

Several algorithms are proposed by various authors. There aren't many algorithms for general use, so there are many drawbacks. Elastic cloud is a cloud computing system with

elastic features. In order to utilize the full potential of Elastic cloud, Max-Min algorithm is introduced, which improves resource utilization and response time[3].

In order to maximum the utilization of resources being shared is done through load balancing. Performance of a system is determined by the slowest node in the network. Load balancing works by assigning the excess work from over performing node to the underperforming node until they become approximately equal. This makes sure that the nodes in the system are neither overloaded nor idle. Load balancing algorithm has several issues: Assignment, Load Calculation, Job Transfer, System state. Static and Dynamic are the two primary approaches to load balancing algorithm[4]. The information can be collected neither centralized nor distributive approach. To overcome the disadvantages of Dynamic and complexity in Static load balancing algorithm, a new Hybrid algorithm is introduced. It overcomes the disadvantages of both static and dynamic[5].

For implementing a proper load balancing algorithm in cloud computing, we come across several challenges like, task load, CPU load, memory load and bandwidth load to be overcome[5]. In order to provide quick service to the user upon request and to avoid overloading of pipelines, the load should be managed in real time manner. This approach can anticipate the load magnitude and fluctuation among them and thus overcome the problem of unbalanced load by anticipating the task dependency pattern, which results in effective load management and resource monitoring[9].

IV. EXISTING LOAD BALANCING TECHNIQUES IN CLOUD

Following are the various load balancing techniques developed and tested by various researchers that are currently being used in cloud computing:

A. Scheduling Job First (SJF) algorithm :

Singh[1] has experimented with SJF scheduling algorithm to improve the throughput which is better option for private cloud. Since SJF is used, some of the jobs which are having larger burst time cause the other jobs to starve for resources. The jobs are executed when tag becomes zero, the tag is decremented automatically on the arrival of job which reduces the problem of bounded waiting. The author maintains a common queue for all the arriving jobs to ensure proper load balancing. The jobs are allocated to idle Virtual Machine(VM) by cloud manager since it has control to all resources. The VM is required to send a signal upon successful completion of assigned job. The cloud manager assigns process with least burst time to VM if no more processes left with zero tag. The experimental results shows that least loaded VM is chosen for execution for process which will improve the throughput of private cloud. For public cloud they developed an algorithm which provides load balancing and maximizes the profit

B. Round Robin (RR) algorithm :

Akshada[2] has concentrated on assigning the jobs to a cloud partition based on the three cloud partition status, those are idle, normal and overload. When the status of a cloud partition is either normal or idle the partition is done locally, and if it's overloaded then the job is routed to another cloud partition. To achieve this, the authors computed the load parameter, load degree and average of load degree. Then the node status levels are determined and assign the jobs to other nodes based on the status. The nodes can't be engaged to any work until the load status comes back to normal and it will be unavailable since then each load balancer has a status table and is been refreshed in a specific periodic time. And the load balancer use this table to calculate the partition status and the load balancing algorithm for each partition will be different. After the job is assigned, the processing of each job is done with RR algorithm and the Game Theory. The main advantage of this model is that it improves the response time, thus overall performance is improved and even helps in optimum utilization of the cloud resources.

C. Max-Min Task Scheduling (MMTS) algorithm :

Y.mao[3] has come with the Max-Min algorithm for the Elastic Cloud (ECMM). The authors main idea is to schedule the task to the virtual machines with the help of two tables, one is the execution task status table and another one is virtual machine status table inside a load balancer. The task status table contains parameters like task execution time, completion time and the latest update time, meanwhile in virtual machine status table contains existing task in Virtual Machine(VM), total execution time of task, status of VM and the latest update

time. How basically the task is allocated with the help of this algorithm is that primarily the task with the maximum execution time (Max) is selected, thereafter computing the execution time in each VM, then select the VM which has the lowest completion time (Min) and assign that particular job to that VM, meanwhile we need to update the virtual machine status table with the relevant details. The main advantages of this algorithm are that it has better response time.

D. **Cat Swarm Optimization algorithm :**

Priyadarshini[4] has experimented with Cat Swarm Optimization technique which is one of the distributed balancing algorithm ,and it is one of the newly emerged algorithms which is cost effective and has very less response time .The authors has proposed this algorithm for load balancing in order to get the trivial solution in very short span of time. In this algorithm they have taken cat as an example, because the cat has its own position composed of any dimension, computed the velocity, fitness value and the flag for all the cats .The author has assumed that the cat can be neither in seeking mode nor in tracing mode. The cat will spend more time in tracing mode .To combine the two modes they define a mixture ratio. While in seeking mode they move their position carefully or even stay where they are. For applying this sort of behaviour into this algorithm they took seeking mode to represent. The behaviour of the cat to take its prey is called the tracing mode .The authors considered N number of cats then the cats are randomly placed in any dimension with max value velocity for each. Randomly some cats are put up into seeking mode and some into tracing mode. Each cat is evaluated using the fitness function and the best among that is taken. Again the cats are randomly distributed to seeking and tracing mode. If the termination condition satisfy then the algorithm stops. The main advantage of this algorithm is the low response time.

E. Particle Swarm Optimization:

Ramzani[5] has designed an algorithm which migrates the processes from the overloaded virtual machines to achieve load balancing. Computing and data intensive task are being implemented in this method. In case of data intensive task, data transfer time is reduced by taking account bandwidth as a variable, which minimizes the data movement. For computing intensive, the scheduling of data to achieve high performance can be done by considering the number of CPUs in the VMs. During VM migration process there is a possibility of losing the most recent activities. Large quantity of memory is required.

F. **Ant colony algorithm :**

Dam[6] proposed an Ant colony algorithm which creates random number of ant each of which having same pheromone value because all ants initially wander randomly, and once the food is found by them, they return to their colony laying down pheromone trails which is chemical substance that attracts other ants. If other ants find such a path, they follow the previously laid path, returning and reinforcing it if they eventually find food. Gradually, the pheromone trail starts to evaporate, thus diminishing its attractive strength. A short path gets marched more frequently, and thus the pheromone density becomes higher on shorter paths than longer ones. If evaporation of pheromone doesn't exist, the paths chosen by the first ants be excessively attractive to the following ones. The underutilized VM's are searched by creating artificial ants and making to disperse to roam across the network. The ants start its journey across the network and moves from one node to another node to search the underutilized node with optimal distance. The experimental results shows that strategy ignores the fault tolerance but assure the QoS requirement for customer**.**

G. **Spring tensor model:**

Aslanzadeh[7] has assessed the fluctuations and capabilities of the load changes in cloud environment by Spring Tensor Model(STM). The authors have experimented with set of workflow tasks, where jobs have dependencies to one another. The model consists of Gaussian Network Model (GNM) which estimates the capabilities of the load and Anisotropic Network Model (ANM) focuses on direction of fluctuation. The associatively among the nodes can be explained with Kirchhoff Matrix. The load fluctuation in multi-dimensional environment can be estimated using capabilities and direction of fluctuation of load using node interactions. The model was experimented with hundreds of nodes to capture the response of load in the cloud. The main advantage of this algorithm is the low response time.

H. **Efficient RID diffusion algorithm:**

Rafiqul[8] has come up with an efficient RID algorithm which reduces the overheads on the load . RID algorithm comes under the diffusion algorithms. Diffusion algorithm basically has two categories, the first one is the sender initiated diffusion (SID) algorithm and the next one is the receiver initiated diffusion (RID) algorithm. The basic working of the RID approach is that the under loaded node will initiate for transferring the

load from the neighbouring over loaded nodes. And in this approach the under loaded nodes which needs to take part for the load transfer do consider all the neighbouring nodes without even considering whether the nodes are ready to transfer the load or not. In this new algorithm , the authors has considered only those nodes which are interested in transferring or receiving the load rather than including all the reluctant nodes .In this model, the under loaded nodes will send its status only to the immediate nodes .The most highlighted part in this algorithm is that the under loaded nodes will not communicate with the immediate nodes until its overloaded .Firstly the global average of that under loaded node is calculated and then decided whether it is under loaded, overloaded or moderate loaded ,and then this status is sent to the nearest neighbouring nodes and it will receive the acknowledgment from the overloaded nodes .The advantage of this algorithm is that in distributed computing environment it is very important to collect the status of other nodes in the network and hence determine whether it is overloaded or not and eventually reduce the overheads of the nodes .

I. **Effective Joint Routing and Scheduling algorithm :**

Suriya at.el.[9] has proposed a mathematical model JRS, exclusively considering virtual machines for performing load balancing. The system jointly addresses the routing as well as task scheduling and also focuses on the issues pertaining to resource allocation. The simulation results shows better throughput by highlighting minimized waiting time for jobs with faster completion of task.

V. COMPARISON OF EXISTING LOAD BALANCING TECHNIQUES

The table below shows the comparative analysis of different load balancing. Techniques that are used in various algorithms. The analysis is made based on advantages and parameters.

TABLE 1: COMPARISONS OF DIFFERENT LOAD BALANCING ALGORITHMS

Algorithm	**Description**	**Advantages**	**Parameters**	
SJF Scheduling SJF[1]	SJF with starvation is used for private cloud and.another algorithm for public cloud	Increases throughput for private cloud and maximizes the profit for public cloud	Throughput and cost	
Round Robin algorithm[2]	Assigning the jobs to a cloud partition based on the three cloud partition status, those are idle, normal and overload.	Improves response time, Overall performance and optimum utilization of the cloud resources.	Response time and resource utilization	
Max-Min algorithm[3]	To schedule the task to the virtual machines with the help of two tables, one is the execution task status table and another one is virtual machine status table inside a load balancer.	Better response time.	Task execution time, completion time, the latest update time, existing task in VM, total execution time of task, status of VM and the latest update time.	
Cat Swarm Optimization[4]	N number of cats has been taken.The cats are randomly placed in any dimension with max value velocity for each. Randomly some cats are put up into seeking mode and some into tracing mode.	Low response time and decreases the cost	Response time and cost	
Particle Swarm Optimization:[5]	It migrates the processes from the overloaded virtual machines to achieve load balancing. Computing and data intensive task are being implemented in this method.	Minimizes the data movement.	Bandwidth	

Ant colony [6]	It provides Quality Of Service as per the requirement of the Customers. Here all the jobs are predicted with same priority and has no fault tolerance.	Improved QOS	QOS	
Spring Tensor[7]	It uses two Models, the Gaussian Network Model (GNM) and Anisotropic Network Model (ANM)	Low response time	Response time	
RID[8]	Diffusion algorithm basically has two categories, the first one is the sender initiated diffusion (SID) algorithm and the next one is the receiver initiated diffusion (RID) algorithm.	Reduces the communication overheads	Communication overheads	

VI .CONCLUSIONS

Cloud Computing is a new paradigm where computer services can be provisioned dynamically. Data centers plays a vital role in cloud , consists of virtual machines for fulfilling the requests of the users. There are various issues in Cloud Computing. Load balancing is one of the major issue in cloud environment. It is a technique in which distribution of the dynamic local workload is done equally across the nodes in a cloud in order to avoid the situation where few nodes are overloaded while few are idle. In this paper various Load balancing algorithms are compared based on advantages and parameters like response time, throughput, cost and QoS.

VII FUTURE WORK

In the future work, we will try to analyse the algorithms by taking the parameters for security purpose, elasticity, execution time and scalability and we will compare their results with the existing algorithms.

REFERENCES

[1] A. S. Singh, S. Shu, M. N. Tiwari, and R.K. Katare,"*Scheduling Algorithm with Load balancing in Cloud Computing"*, International Journal of Scientific Engineering and Research, Volume 2 issue 1,January 2014 pp. 38-43

[2] M. V. Marathe, A. Bhujbal, P. Jakate, M. Wagh, and M. Pise, "*Load Balancing Model in Cloud Computing"*, International Journal of Emerging Engineering Research and Technology, volume 3 ,issue 2, February 2015, pp. 1-6

[3] Yingchi Mao, Xi Chen and Xiaofang Li, *"Max-Min Task Scheduling Algorithm for Load Balancing in Cloud Computing"* ,Proceedings of International Conference on Computer Science and Information Technology, Advances in Intelligent System and Computing

[4] A.P. Priyadarshini, R. Sharmistha & K.P. Prasant "Load Balancing Adaption of Some Evolutionary Algorithms In Cloud Computing Environment **",** IEEE **,** Vol 8. No. 2 June, 2015

[5] Jing Deng, Ping Guo, Qi Li, Haizhu Chen ,"*A Load Balancing Strategy with Bandwidth Constraint in Cloud Computing*", he Open Cybernetics & Systemics Journal, 2014, 8, pp. 115-121

[6] Santanu Dam, Gopa Mandal, Kousik Dasgupta, and Paramartha Dutta ,"*An Ant Colony Based Load Balancing Strategy in Cloud Computing", Advanced Computing, Networking and Informatics - Volume 2*, 403 Smart Innovation, Systems and Technologies 28

[7] Fahimeh Ramezani, Jie Lu, Farook Khadeer Hussain "Task-Based System Load Balancing in Cloud Computing Using Particle Swarm Optimization ",Int J Parallel Prog (2014) 42:739-754

[8] Rafiqul Z. Khan , Md F. Ali , "*An Efficient Diffusion Load Balancing Algorithm in Distributed System* ",I.J. Information Technology and Computer Science, 2014, 08, pp. 65-71

[9] Suriya Begum, Dr. Prashanth C.S.R,"Mathematical Modelling of Joint Routing and Scheduling for an Effective Load Balancing in Cloud" in International Journal of Computer Applications (0975 –8887) Volume 104 –No.4, October 2014.

[10]. Jaspreet kaur "Comparison of load balancing algorithms in a Cloud" 2012,International Journal of Engineering Research and Applications (IJERA) ISSN: 2248-9622 www.ijera.com Vol. 2, Issue 3, pp.1169-1173.

[11] Suriya Begum ,Prashanth "Review of Load Balancing in Cloud Computing" ,IJCSI International Journal of Computer Science Issues, Vol. 10, Issue 1, No 2, January 2013 ISSN (Print): 1694-0784 | ISSN (Online): 1694-0814

[12] Suriya Begum,Prashanth "Investigational Study of 7 Effective Schemes of Load Balancing in Cloud Computing ",IJCSI International Journal of Computer Science Issues, Vol. 10, Issue 6, No 1, November 2013 ,ISSN (Print): 1694-0814 | ISSN (Online): 1694-0784

[13] http://www.ibm.com/cloud-computing/us/en/Accessed on 26th Sep, 2012.

[14] David S. Linthicum, Cloud Computing and SOA Convergence in your Enterprise, Pearson, 2010

[15]URL=https://www.google.co.in/search?q=cloud+aRCHITECTURE&biw=1600&bih=787&source=lnms&tbm=isch&sa=X&ved=0ahUKEicg9b69IrLAhWUco4KHQnYAN4Q_AUIBigB#tbm=isch&q=cloud+computing+layers+architecture&imgrc=77eksAteeNU5jM%3A

[16] Ashima Narang , " Various Load Balancing Techniques in Cloud Computing " , International Journal of Computer Science and Mobile Computing, Vol.3 Issue.9, September- 2014, pg. 502-509

