

A Literature Survey on "Load Balancing by Partitioning the Public Cloud"

Pallavi T. Dhotre, Asst. Professor IT Department, PGMCOE, Wagholi, Pune.

pallavidhtr03@gmail.com

Abstract— Load balancing is an important factor in today's technological era. Where almost people are making use of internet for getting the new information, cloud computing has become very important factor for storing and accessing the information they need. But while accessing the public cloud network congestion has become the main issue here. Now a day's where workload is increasing day by day the network traffic has become main issue in this area and therefore degraded performance of the network, reduced efficiency, server overloading etc problems came into existence. Many times server overloading causes to crash the system. To avoid the network traffic related problems take the advantage of the concept of "Load Balancing by Partitioning the Public Cloud Model". The primary aim to review the concept in this paper is to know the performance of load balancing and to understand concept of partitioning depend on the geographical area. To partition the public cloud in Load Balancing Model the main role is performed by main controllers, balancers and the servers.

Keywords—load balancing, public cloud, controller, balancer, server.

I. INTRODUCTION

Cloud computing is nothing but storing the data globally and access it from anywhere at any time. We can store large amount of data using clouds. Quality of service in the cloud computing is mainly depend on the Quality of traffic, low throughput, errors, dropped packets, latency, jitter etc. To improve the performance and avoid the traffic problem load balancing is the main concept which can be used.

While accessing data from the cloud many users can request to the same server for the same data this will cause the network congestion and will result into degradation of performance. To maintain the performance and to avoid the congestion Load balancing concept is very useful. In load balancing requests are distributed in such a manner that every server will get an equal workload from the user so that each server will provide the equivalent result to the user within the less amount of time.

Load balancing can be done with the help of partitioning the public cloud. Where partitioning of the cloud is done on the basis of geographical locations. The better load balancing scheme results into the improved performance of cloud computing and also results into more satisfying users.

This paper tries to review the various Load balancing schemes in cloud computing. The main purpose of using the load balancing is to remove the degradation of performance, to handle the workload by load balancing and to speed up the user-server interaction. Main controllers, balancers and servers perform the main role in load balancing scheme.

II. LITERATURE SURVEY

1. Ms. Shilpa D.More, Prof. Arti Mohanpurkar, "LOAD BALANCING STRATEGY BASED ON CLOUD PARTITIONING CONCEPT"[1]

Main aim of load balancing model in this paper is in order to improve performance and maintain stability of processing so many jobs in public cloud. Load balancing is reduce processing and response time which is having impact on cost. This objective is achieve by constructing good balancing model for public cloud based on cloud partitioning with switch mechanism to choose different strategy to improve the efficiency in public cloud environment.[1]

2. Akshada Bhujbal, Prajakta Jakate, Manasi Wagh, Madhura Pise, Prof.M.V.Marathe, "Load Balancing Model in Cloud Computing," [2]

The overall goal of this project is the balancing of load on clouds. This will improve the performance of cloud services. And prevent the overloading of server, which would otherwise degrade the performance. The response time will also improve. Thus overall performance of cloud services will remains unchanged. In case of the system fails even also partially, it having backup plan. It will maintain the stability of the system. There are provisions to accommodate future modifications in the system. This algorithm will ensure the optimum utilization of cloud resources. This algorithm will cut the economic cost for an organization because less resources will be required than static algorithms to handle the user requests.[2]

3. Surbhi Kapoor, Dr. Chetna Dabas," Cluster Based Load Balancing in Cloud Computing,"[3]

In this paper, it is studied that aspect of load balancing which deals with assigning tasks to appropriate VMs in order to maintain the load on all the VMs. Various exiting algorithms have been reviewed for this work. The problems with existing throttled and modified throttled algorithms is that these approaches do not consider resource specific demands of the tasks, are not suitable for heterogeneous VMs environment and have additional overhead of scanning the entire list of VMs. These problems have been addressed in the proposed approach by clustering the VMS into groups such that VMs in same group have similar capacities. Experimental results have shown that our proposed approach gives better results than throttled and modified throttled algorithms when compared on the basis of waiting time, execution time, turnaround time and throughput.[3]

Following table gives the results of various load balancing algorithms used in cloud computing:

Sr.	Name of the Paper	Authors	Result
No			
1	LOAD	Ms. Shilpa D.More,	improve
	BALANCING	Prof. Arti	performance
	STRATEGY	Mohanpurkar	and maintain
	BASED ON		stability of
	CLOUD		processing so
	PARTITIONING		many jobs in
	CONCEPT		public cloud
2	Load Balancing	Akshada Bhujbal,	prevent the
	Model in Cloud	Prajakta Jakate,	overloading
	Computing	Manasi Wagh,	of server,
		Madhura Pise,	ensure the
		Prof.M.V.Marathe	optimum
			utilization of
			cloud
			resources
3	Cluster Based Load	Surbhi Kapoor, Dr.	assigning
	Balancing in Cloud	Chetna Dabas	tasks to
	Computing		appropriate
			VMs in order
			to maintain
			the load on all
			the VMs

Table 1 - Results of various Load balancing algorithms in cloud computing

III. SYSTEM ARCHITECTURE



Fig. 1 Alpha α, betaβ, gamma γ status of the cloud nodes

Before understanding the concept of load balancing using main controller balancer and the clouds first we will have a look on the concepts alpha α , beta β and gamma γ .[2]

Alpha (α): maximum nodes are idle in the cloud. There is no load on that specific cloud.

Beta (β) : Nodes are having normal load no overloading and not idle.

Gamma (γ): Nodes are overloaded with the user requests.

Main components of the architecture of load balancing in cloud computing

1. Main controller:

Main controller directly deals with the incoming jobs or the requests sent by multiple users. Cloud partitioning is done on the basis of various parameters. Here we are focusing on the geographical locations. On the basis of geographical location public cloud partitioning is done and the load is distributed over various balancers which are present in each partition and are connected to the main controller.[1]

2. Balancers:

There are various balancers which are connected to the main controller. One balancer is present in each partition. Depending upon the alpha, beta, gamma status of the balancer the job is given to that particular balancer by the main controller. If the nodes connected to the balancer are in alpha state that is they are idle then the job is allocated to that balancer. If the nodes connected to the balancer are in beta state that is they are having the normal workload then the job is assigned to that balancer depending upon the capacity of that balancer. And if the nodes connected to the balancer are in the gamma state that is they are already overloaded then the job requests are given to the another balancer via main controller.[1]

3. *Nodes* : Various nodes are connected to the balancers for dealing with the incoming requests[1].





IV. CONCLUSIONS

This paper reviews the load balancing techniques in which public cloud is partitioned by geographical area and then load balancing is done. Load balancing is nothing but distributing the load evenly on all nodes so that no node will stay idle or no node will be overloaded. There are various load balancing techniques but the partitioning the public cloud is the easiest method and efficient as compare to other methods. In today's world big data concept comes in existence so for storing this large amount of data we require cloud and while storing the data on cloud it is essential to check the load on the particular node. And if any node is overloaded and again new request is forwarded to that node then network congestion may occur so in order to avoid the overhead and the network congestion in the cloud computing Load balancing techniques are important.



REFERENCES

1] Ms. Shilpa D.More, Prof. Arti Mohanpurkar, "LOAD BALANCING STRATEGY BASED ON CLOUD PARTITIONING CONCEPT, "*Multidisciplinary Journal of Research in Engineering and Technology, Volume 2, Issue 2,2015*

2] Akshada Bhujbal, Prajakta Jakate, Manasi Wagh, Madhura Pise, Prof.M.V.Marathe, "Load Balancing Model in Cloud Computing ,"International Journal of Emerging Engineering Research and Technology Volume 3, Issue 2, February 2015

3] Surbhi Kapoor, Dr. Chetna Dabas," Cluster Based Load Balancing in Cloud Computing," IEEE, 20-22 Aug. 2015.

4] Mohamed Belkhouraf, Ali Kartit and Hassan Ouahmane et. Al., "A secured load balancing architecture for cloud computing based on multiple clusters," IEEE, 2015

5] Jessica Oueis, Emilio Calvanese Strinati, and Sergio Barbarossa" The Fog Balancing:Load Distribution for Small Cell Cloud Computing," IEEE,2015.

6] Gaochao Xu, Junjie Pang, and Xiaodong Fu "A Load Balancing Model Based on Cloud Partitioning for the Public Cloud ," IEEE TRANSACTIONS ON CLOUD COMPUTING YEAR 2013

7] Radha Krishna Palivela, U. Chandra Sekhar Reddy "LOAD Balancer Strategy Based On Cloud Computing ," International Journal of Research in Computer and Communication Technology, Vol 3, Issue 10, October – 2014

8] PUVVALA SUPRIYA, K.VINAY KUMAR "Cloud Partitioning Based Load Balancing Model for Cloud Service Optimization," *IJCSMC, Vol. 3, Issue. 12, December 2014*

for Research in Engineering