

Survey on Load Balancing using Parameters in Cloud Computing

Mr. Prathmesh Nayak¹ Mr. Jignesh Vania²

¹M.E Student ²P.G Student

^{1,2}Department of Information and Technology

^{1,2}LJIET, Ahmedabad, Gujarat, India

Abstract— Cloud computing is a kind of distributed computing where different massively scalable IT related resources or capabilities are provided to a number of external users as a service using internet. Cloud computing is the delivery of computing services over the Internet. Cloud services allow duality and business to use software and hardware that are managed by third parties at remote locations. Here in this paper we have discussed many different load balancing techniques used to solve the issue in cloud computing and also discuss with quality of services in load balancing in cloud environment.

Key words: Cloud Computing, Load Balancing, Quality of Services, Literature Survey

I. INTRODUCTION

The definition of cloud computing is the cloud itself. The cloud is a large group of inter connected computers. These computers are personal computer or network server. Cloud computing is the delivery of computing services over the Internet. The main feature of the Cloud computing is it makes all the resources available at one place in the form of a cluster and the resources are allocated to the users according to their requests. This cluster based approach helps in achieving the maximum CPU utilization and reduces the efforts of users to access the cloud resources. Cloud is a collection of computer and server that publically accessible via internet. Cloud can be public or private. Cloud computing is not network computing.

Cloud computing is widely used in almost every organization. It provides the usage of its virtual resources and scalability. There are various pros and cons of cloud computing.

A. Pros and Cons of Cloud Computing:

1) Cloud computing: Advantages:

- Lower-cost Computer for users
- Improved performance
- Lower IT Infrastructure Costs
- Fewer Maintenance Issues
- Lower Software costs
- Increased Computing Power

2) Cloud computing: Disadvantages:

- Required Unchanging internet connection
- Features Might Be Limited
- Stored Data Might Not Be Secure

B. Cloud Computing Architecture^[1]

In general the architecture of a cloud computing environment can be divided into 4 layers: the hardware layer, the infrastructure layer, the platform layer and application layer as shown in Fig. 1. The cloud computing provide to the different three Services: software as a service (SaaS), platform as a service (PaaS), and infrastructure as a service (IaaS).

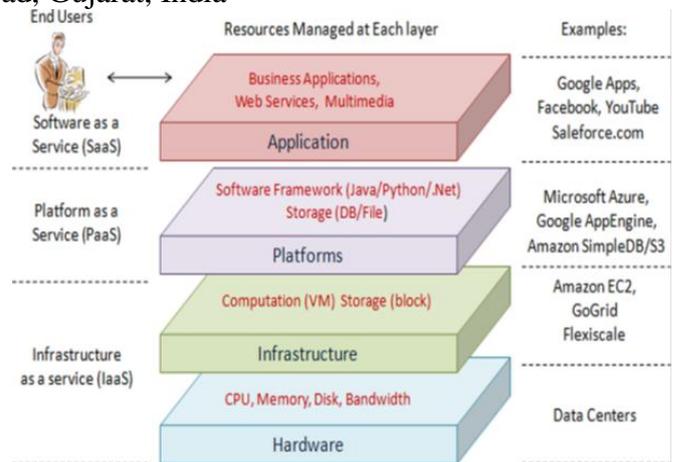


Fig. 1: Cloud Architecture [1]

II. LOAD BALANCING

The main issue related to cloud computing is load balancing. When the number of job increases than Load is occurs. Load balancing is the pre requirements for increasing the cloud performance and for completely utilizing the resources. Load balancing aims to optimize resource use, maximize throughput, minimize response time, and shun overload of any single resource.

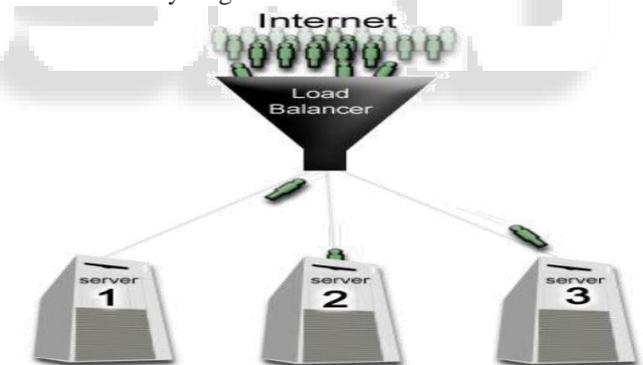


Fig. 2: Load Balancing[2]

A. Goals of Load Balancing:

- To improve the performance substantially
- To Build fault tolerance system
- To maintain system stability
- To accommodate future modification.

B. Load Balancing Algorithms Are divided into 2 Categories:

1) Static Load Balancing:

Static load balancing algorithm requires knowledge about the application and resource of system. It does not depend on the current state of the system. The System is needed by using prior knowledge. The advantages of static methods are simplicity in implementation and low communication overheads.

2) *Dynamic Load Balancing:*

Dynamic load balancing it depends on current state of the system. No Prior knowledge is needed. A dynamic load balancing algorithm does not consider the previous state of the system. Improves overall system performance.

C. *Load Balancing Algorithms:*

This survey paper describes about three load balancing algorithms which are Round robin algorithm, equally spread current execution load and Throttled Load balancing.

1) *Round Robin Algorithm:*

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 less compared to the other two algorithms and default algorithm used in simulation. This algorithm simply allots the job in round robin fashion which doesn't consider the load on different machines. In this algorithm It will not check whether the server is heavy loaded or not, it will directly assign the request whenever its turn comes so that is the reason, some server are heavy loaded while some are lightly loaded.

2) *Dynamic Round Robin:*

Dynamic Round Robin, strategy, weights assignments is based on continuous monitoring of the servers and is therefore constantly changing. Distributions of connections is done on the basis of server performance analysis such as the current number of instances i.e. connection per node or the response time of a fastest node.

3) *Weighted Round Robin:*

Weighted round robin was defined to improve the critical challenges associated with round robin. In this algorithm each server has been assigned a weight and according to the highest weight they receive the more connection.

4) *Equally Spread Current Execution:*

Improves performance by transferring load from heavily loaded server. In this algorithm process are handle with priorities. Resource allocation of cloud computing is based on which the performance of the system is estimated. In spread spectrum technique load balancer makes effort to preserve equal load to the VMs connected with the data center.

5) *Throttled Algorithm:*

Client first requests the load balancer to find a suitable Virtual Machine which access that loads easily and perform the required operation. These algorithms are completely base on virtual machine. Does not consider the advanced load balancing requirements such as processing times for each individual requests.

D. *Comparison of Quality of Service in Existing Load Balancing Technique*^[6]

Metrics/Techniques	Throughput	Overhead	Fault tolerance	Migration time	Response time	Resource Utilization	Performance
Round Robin	Yes	Yes	No	No	Yes	Yes	Yes
Dyn	Yes	Yes	Yes	Yes	No	Yes	No

amic RR							
Active monitoring	Yes	Yes	No	Yes	Yes	Yes	No
Throttled	No	No	Yes	Yes	Yes	Yes	Yes

Table. 1: Various Qualities Of Services

In This table various Qualities of services have been considered to compare different techniques. The Quality of service on which the existing load balancing techniques are discussed below: [5]

1) *Throughput:*

This Qos or metrics are used to estimate the total number of tasks, whose execution has been completed successfully. High throughput is need full for overall system performance.

2) *Overhead:*

In this metric or Qos any load balancing algorithm are indicates the extra cost involved in implementing the algorithm. It must be as low as possible.

3) *Fault Tolerance:*

Fault tolerance can perform uniform load balancing in case of any failure. A good load balancing algorithm must be highly fault tolerable.

4) *Migration Time:*

Migrating the jobs or resources required the total time from one node to another. It should be minimized.

5) *Response Time:*

This Qos are used to time interval between sending a request and receiving its response.

6) *Resource Utilization:*

It is used to ensure the proper utilization of all those resources. This factor must be optimized to have an efficient load balancing algorithm.

7) *Performance:*

It is used to check, how efficient the system is. This performance can br improved at a reasonable cost e.g. reducing the response time though keeping the acceptable delays.

III. LITERATURE SURVEY

A. *State-of-the-art and research challenges by cloud computing*^[1]

By Qi Zhang Lu Cheng Raouf Boutaba 20 April 2010 springer. In this paper, we have surveyed the state-of-the-art of cloud computing cover concepts, architectural designs, characteristics, key technologies and research directions. cloud computing technology is provide a better understanding of the design challenges of cloud computing.

B. *A Comparative Study of Load Balancing Algorithms in Cloud Computing Environment*^[2]

By Mayanka Katyral, Atul Mishra In this paper, we discussed various load balancing schemes, each having some pros and cons. On one hand static load balancing scheme provide easiest simulation and monitoring of environment but fail to model heterogeneous nature of cloud. On the other hand dynamic load balancing algorithm are used for difficult simulate but are used for best

heterogeneous environment of cloud computing. Also the level at node which implements this static and dynamic algorithm plays a vital role in deciding the effectiveness of algorithm. Unlike centralized algorithm, distributed nature of algorithm provides better fault tolerance but requires higher degree of replication. dynamic load balancing techniques in distributed or hierarchical environment provide better performance.

C. A Survey on Load Balancing in Public Cloud^[3]

International Journal of advanced Research in Computer Science & Software Engineering, March 2013, Vol 4, issue 3. In this paper, we have surveyed various load balancing techniques for cloud computing. The main purpose of load balancing is to satisfy the customer requirement by distributing load dynamically among the nodes and to make maximum resource utilization by reassigning the total load to individual node.. So in our paper we have discussed all the existing techniques for Load balancing. And we have also discussed the virtualization and cloud computing.

D. Improving Efficiency of Round Robin Scheduling Using Ascending Quantum And Minumim-Maxumum Burst Time^[4]

By Ali Jbaer Dawood. Round Robin (RR) is a kind of process algorithms, where the time quantum is fixed along the processes execution. RR Performs in timesharing system by given each process static Time Quantum (TQ). In this paper, The TQ studied to improve the efficiency of RR and performs the degrades with respect to Context Switching (CS), Average Wait Time (AWT) and Average Turned Around Time (ATAT) that an overhead on the system. Thus, the new approach was proposed to calculate the TQ, known as Ascending Quantum and Minumim-Maxumum Round Robin (AQMMRR). The processes were ascending with shortest remaining burst time and calculate the TQ from multiply the summation of minimum and maximum BT by (80) percentage. The experimental result shows that AQMMRR performs better than RR and comparing with other two related works.

E. Comparative Study on Load Balancing Techniques in Cloud Computing^[5]

N. S. Raghavaand Deepti Singh. In this paper survey of various load balancing algorithms in the Cloud Computing environment. Discussed about issues related to load balancing in cloud computing and comparison of existing load balancing techniques also discussed the current literature and the load balancing algorithm in cloud. Discuss the Quality of service or metrics like performance, throughput, fault tolerance, Response time etc.

F. Improved Load Balancing model based on Partitioning in Cloud Computing^[8]

In this paper, static load balancing authors have used round robin, equally spread current execution, throttled algorithm on like response time, data center processing time. They have verified total cost of different data centers. Below is the Comparison of load balancing policies (overall response time, Data Center Processing time, Total Cost).

	Round robin	ESCG	Throttled
Overall response time	300.12	300.12	300.12

Datacenter processing time	0.35	0.35	0.35
VM cost(\$)	0.50	0.50	0.50
Total cost	16	16	16

Table 2: Comparison of load balancing policies

In this table there are no changes in response time and no other factors are remained unchanged. So Uses of static algorithm are not preferable. So in dynamic environment are using RR, ESCG, and Throttled are shown in table below:

	Round robin	ESCG	Throttled
Overall response time	223.34	209.75	209.74
Datacenter processing time	13.84	10.11	10.10
VM cost(\$)	5.3	4.45	4.03
Total cost	18.3	18.3	18.3

Table 3: Comparison of load balancing policies

IV. CONCLUSION

In this paper we have survey various load balancing algorithms in cloud computing. In cloud computing load balancing is the main issue. We have discussed the already proposed algorithms by various researchers in literature. A comparison has been done on basic different criteria like scalability, resource utilization, fault torerance, response time, etc.

REFERENCES

- [1] Qi Zhang · Lu Cheng · Raouf Boutaba “state-of-the-art and research challenges by cloud computing “20 April 2010 springer.
- [2] Mayanka Katyal, Atul Mishra A Comparative Study of Load Balancing Algorithms in Cloud Computing Environment.
- [3] A Survey on Load Balancing in Public Cloud International Journal of advanced Research in Computer Science & Software Engineering, March 2013, Vol 4, issue 3.
- [4] Ali Jbaer Dawood. ” Improving Efficiency of Round Robin Scheduling Using Ascending Quantum And Minumim-Maxumum Burst Time”.
- [5] N. S. Raghava and Deepti Singh “Comparative Study on Load Balancing Techniques in Cloud Computing” open journal of mobile computing and cloud computing vol 1,aug-2014.
- [6] Jaspreet Kaur-,”comparison of load balancing algorithms in cloud”, International journal of Engineering Research and Application”, Vol.2,Iss 3,May-June 2012.
- [7] Nusrat pasha and Dr. amit agarwal “Round robin approach for VM load balancing algorithm in cloud computing environment” International Journal of advanced Research in Computer Science & Software Engineering, vol 4, issue 5, May 2014.
- [8] G.Rajiv Ratnakar, CH.Madhu Babu “improved load balancing model based on partitioning in cloud computing” International Journal of Computer Science and Mobile Computing July 2014.