Dynamic Multi Stage Model Based Load Balancer For Optimal Resource Allocation

[1]Dhatri Prajapati, [2]Gayatri Pandi(Jain)

[1] Student ,Computer engineering, L.J.I.E.T, Gujarat, India
[2] Head of Department, Computer engineering, L.J.I.E.T, Gujarat, India

ABSTRACT

Cloud Computing is the outcome of rapid growth of the internet. Cloud Computing is the forerunner among the technologies emerging today. Two of the main obstacles in the usage of cloud computing are Cloud Security and Performance stability. Load Balancing is the major issue in cloud computing. An efficient Load Balancing technique can improve the performance in terms of efficient resource utilization and higher customer satisfaction. Many Load Balancing algorithms have been proposed for cloud computing. Various Parameters to analyze the performance of load balancing approach are response time, cost, data processing time and throughput. In this paper the topic under discussion is about load balancing and different algorithms that are proposed for distributing the load among the nodes and also the parameters that are taken into account for calculating the best algorithm to balance the load.

Keyword : *Load Balancing; Cloud Computing; Resource Allocation, Cloud simulator analyst*

1. Introduction

1.1 Cloud Computing

- Cloud Computing is a new service delivery model over the internet. Cloud Computing is the model which provides on demand network access to a shared pool of configurable computing resources that can be rapidly provisioned and released with minimal management effort or service provider interaction. Cloud Computing is a product of mixing of grid computing, distributed computing, parallel computing, utility computing, network storage, virtualization and load balancing. Cloud Computing results in cost savings because there is no need of initial installation of much resource. Cloud computing has minimized the payment of premium costs for products with its pay as you go basis which increased the traffic in the IT services making Load Balancing a central point of research. Cloud resources are usually not only shared by multiple users but are also dynamically reallocated per demand. In this paper some of the optimistic algorithms are surveyed which had shown some improvement in load balancing. A cloud computing is a term for anything that involves delivering hosted services over internet.[8]

1.2 Load Balancing

- Load Balancing means distribution of task among different available resources so that no one can over or under utilized. Resources can be data center, physical machine or virtual machine. Load balancing has been taken into consideration so that every virtual machine in the cloud computing system does the same amount

of workload and therefore by increasing the throughput and minimizing the response time, user's satisfactions will be provided. Load balancing is a performance improving method applied in the area of Networking to distribute the work load across multiple resources that are involved in the computation of a networking task. To handle this issue of load balancing in an efficient there is a need of efficient and effective load balancing strategy which is specifically suitable for cloud environment.

- Load Balancing Types

Broadly load balancing strategies for cloud environment can be divided into following categories:

- Static Approach: This approach is suitable for homogeneous and non dynamic environment. In this approach, algorithm is defined during designed time and it remains same throughput. There is no scope of changing scenario.

- Dynamic Approach: This approach considers the current parameters while assigning task to a node. It is more suitable for cloud environment. Such algorithms are hard to implement as they have to constantly monitor the nodes and task progress and take the decision based upon that.[7]

1.3 Challenges in Cloud Computing Load Balancing

Finding a solution for load balancing is never easy process there will be many challenges to be faced while developing a solution. Some of the challenges are discussed.

- Migration time

while serving the client on his demands sometimes there is a need to migrate resources from long distances due to unavailability in near locations. In such cases the time of migration of the resources from far locations will be more which will affect the performance of the system. While developing an algorithm one should note that resource migration time is an important factor that greatly affects the performance of the system.

- Performance of the system

It doesn't mean that if the complexity of an algorithm is high then the performance of the system will be high. If the complexity of algorithm is high then the implementation cost will also be more and even after implementing the system performance will be decreased due to more delays in the functionality of the algorithm.

- Energy Management

A load balancing algorithm should be designed in a way such that the operational cost and the energy consumption of the algorithm must be low. Increase in the energy consumption is one of the main problems that cloud computing is facing today. Even though by using energy efficient hardware architectures which slows down the processor speed and turn off machines that are not under use the energy management is becoming difficult.

- Security

Security is one of the problems that cloud computing has in its top most priority. The cloud is always vulnerable in one or the other way to security attacks like DDOS attacks etc. While balancing the load there are many operations that take place like VM migration etc at that time there is a high probability of security attacks. So an efficient load balancing algorithm must be strong enough to reduce the security attacks but should not be vulnerable. [9]

2. Related work

Anurag Jain and Dr. Rajneesh Kumar[1] proposed the Two level load balancer approach by combining join idle queue and join shortest queue approach. Authors have used cloud analyst simulator to test proposed two level load balancer approach.

Two level approach taken the characteristics of join idle queue and join shortest queue. Join idle queue is the first level, while join shortest queue approach is used at second level.[1]

Reena Panwar and Prof. Dr. Bhavna Mallick[2] proposed a dynamic load management algorithm has for distribution of the entire incoming request among the virtual machines effectively.

Dynamic Load Management algorithm which will distributes the load uniformly at the servers by considering the current status of all the available virtual machines intelligently and later response time of this algorithm is compared with the existing VM-Assign Algorithm. This algorithm solves all the issues of ineffective usage of the present VMs. [2]

Mrs. Nagamani H. Shahapure and Dr. Jayarekha P[3] proposed Optimal Cost Scheduling algorithm to understand the existing processor scheduling techniques and develop an optimized load balancing algorithm which gives maximum benefit to the cloud service provider. It reduce the cost and processing power.

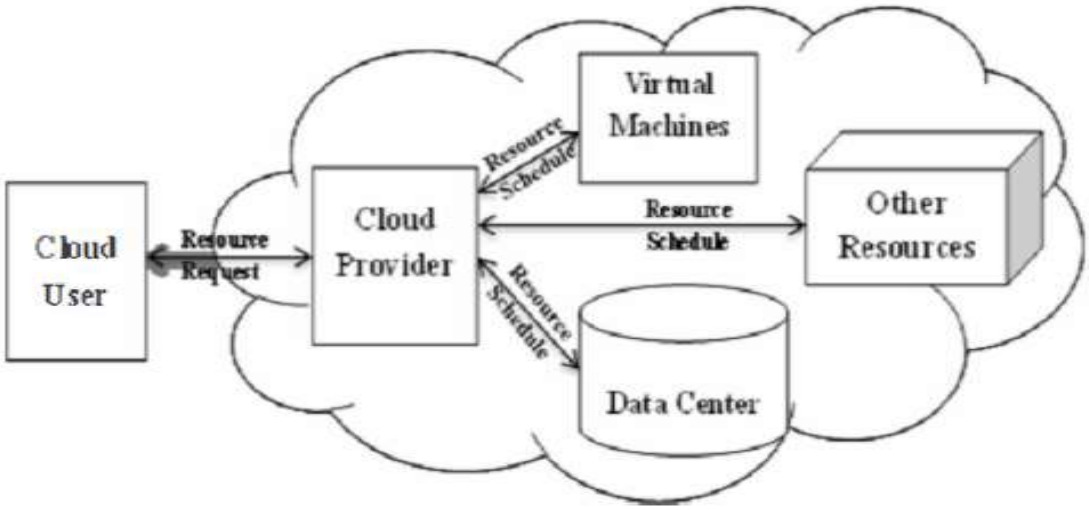

Fig -1 :Cloud scheduling Architecture[3]

Architecture gives overall view of the scheduling process .The user interacts with the provider for receiving the services/resources. When the cloud provider receives the request it forwards it to the Virtual machine for processing. The scheduler applies a proper load balancing and scheduling techniques to process the requests. The processed resources are sent to the cloud user through the cloud provider.[3]

Mohammad Reza Mesbahi, Mahnaz Hashemi and Amir Masoud Rahmani[4] proposed an analytical comparison for the combinations of VM load balancing algorithms and different broker policies. This approachevaluate these approaches by simulating on Cloud Analyst simulator and the final results are presented based on different parameters.

In this approach, three round robin, throttled and Equally Spread Current Execution VM load balancing algorithm and three different datacenter broker policies in cloud computing environments. By this combination we generate

nine different possible load balancing approaches which simulates each one about five iterations with different workloads.[4]

R Krishnam Raju Indukuri, Dr. Suresh Varma Penmasta, Dr. M V Rama Sundari and Dr. G. Jose Moses[5] proposed Multi Stage scheduling in cloud computing to schedule Virtual Machines (VM) for the requested jobs received from customers. This model extended for Deadline Aware Multi Stage Scheduling with respect to response time and waiting time. This gives the better performance evolution metrics when compared to other scheduling algorithm.

	FCFS Scheduling	SJF Scheduling	Multi Stage Scheduling	Deadline Aware Multi Stage Scheduling
AWT	1305	885	668	591
ATT	3609	3189	2972	2896
ADVW	538	188	142	29
ADVR	35	35	19	16
v_w	19	13	7	3
v_r	10	8	5	5

Fig -2: Comparison of performance evaluation matrics[5]

In this table Average waiting time, Average turnaround time, Average deadline violation with respect to waiting time and turnaround time of the Deadline aware Multistage Scheduling algorithm is reduced than other scheduling algorithms.[5]

3. PROPOSED WORK

Cloud user sends the request for resources. The request is goes to the Load Balancer on the basis of Shortest Job First (SJF) algorithm.SJF works on different parameters like Request time, Service time. The Load Balancer works on availability of queue and request time and allocate request to the particular VM. The VM Status Analyzer checking the status of the VM on the basis of the CPU Utilization of each Machine. After checking Status Calculate the Average CPU utilization of each machine. Then Generate the Queue which having CPU utilization less than the average CPU utilization of VMs. **:** If no. of nodes in the Queue is less then the 50% nodes of the Average queue then Scale UP otherwise Scale Down.

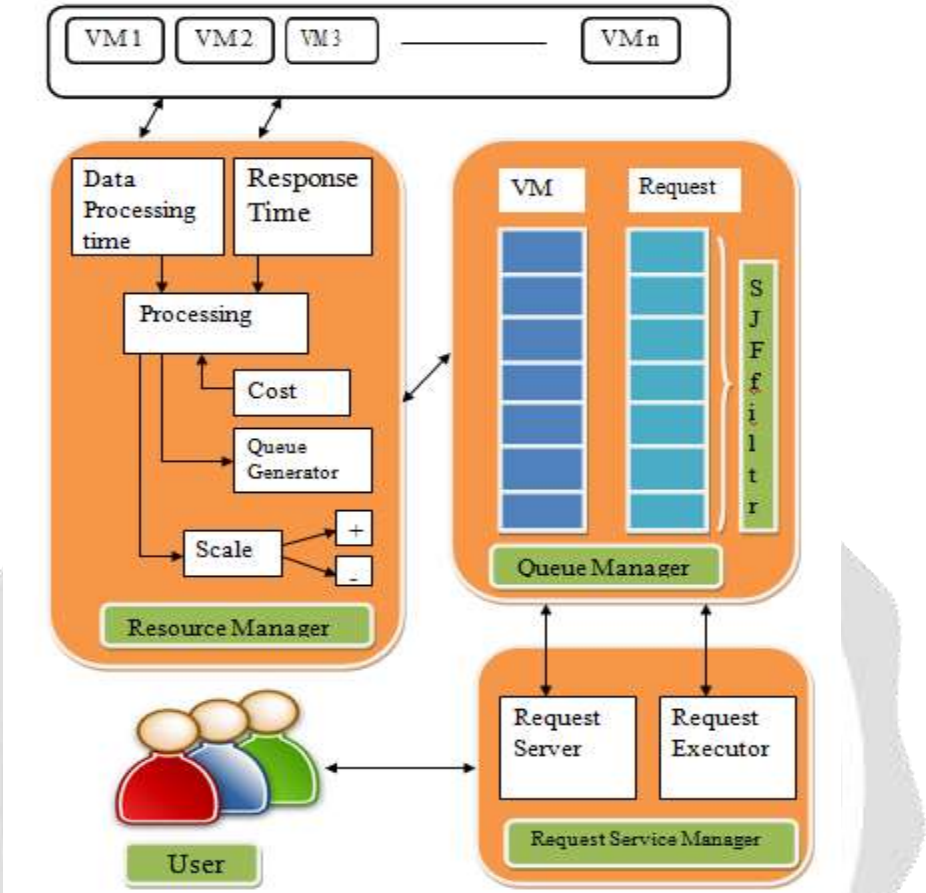

Fig -3 : Propose model

4. RESULT & ANALYSIS

```
Info: VM Response time Info...
i-0351ca2a2a7ac8a2f Response time: 0s
i-003e48eca289a896c Response time: 0s

Info: VM Cost Info...
i-0351ca2a2a7ac8a2f Cost: 0.195$
i-003e48eca289a896c Cost: 0.195$

Info: Calculating average responce time
Info: Average response time: 0.0s

Info: No Up-down scaling requqired.

Info: Printing queue....
i-0351ca2a2a7ac8a2f is in average queue with ip 35.154.156.212
i-003e48eca289a896c is in average queue with ip 35.154.179.96

Info: Wait time... 2min
```

Fig -4 : Result of response time of different machine

The fig. 4 describes the result of this proposed work which determines that the response time and cost of each VM updates time to time. The values of it changes as time interval. In this research we take the time interval between two iteration is 2. We get result for multiple machines.

5. CONCLUSION

In this dissertation focus on the balancing load among the different virtual machine by allocating the resources to the different VMs and using the prior testing mechanism to inform us about overloading and under loading virtual machine and datacenter in advance. Using the different parameters we can distribute the load among different VMs. This gives Better response time.

5. REFERENCES

[1] Anurag Jain and Dr. Rajneesh Kumar, "A Multi stage Load Balancing Technique for Cloud Environment " IEEE ISBN 978-1-5090-2552-7, 2016.
[2] Reena Panwar and Prof. Dr. Bhawna Mallick, "Load Balancing in cloud Computing Using Dynamic Load Management Algorithm " 978-1- 4673-7910-6/15/$31.00 ©2015 IEEE.
[3] Mrs. Nagamani H. Shahapure and Dr. Jayarekha P, "Load Balancing with Optimal Cost Scheduling Algorithm", 978-1-4799- 38261114/$31.00©2014 IEEE.
[4] Mohammad Reza Mesbahi, Mahnaz Hashemi and Amir Masoud Rahman, "Performance Evaluation and Analysis of Load Balancing Algorithms in Cloud Computing Environments" 978-1-5090-2166-6/16/$31.00 ©2016 IEEE 145 978.
[5] Dr. Suresh Varma Penmasta and Dr. G. Jose Moses, "Performance Evaluation of Deadline Aware Multi Stage Scheduling in Cloud Computing" 978-1-4673-8286-1/16 $31.00 © 2016 IEEE DOI 10.1109/IACC.2016.51.
[6] V RaviTeja Kanakala,V.Krishna Reddy and K.Karthik, " Performance Analysis of load balancing techniques in cloud computing environment", 978-1-4799-6085-9/15/$31.00©2015 IEEE.
[7] Amandeep, V. Yadav and F. Mohammad," Different Strategies for load Balancing in cloud Computing Environment: A Critical Study", International Journal of Scientific Research Engineering & Technology(IJSRET), Vol. 3,No. 1, pp. 85-90,April 2014.
[8] https://en.slideshare.org/load balancing /Cloud_computing . (Accessed:05/10/2016 at 2:00 PM)
[9] https://www.google.co.in/search?q=challanged+of+load+balancing+incloud%3Acomuting&aqs=chrome..69i5 7.18306joj4&client(Ac cessed:10/10/2016 at 12:30 PM)
[10] Mayank Mistra, Anwesha Das, Purushottam Kulkarni and Anirudha Sahoo, "Dynamic Resource Management Using Virtual Machine Migrations" , Sep 2012. 0163-6804/12,IEEE Communications Magazine.