



EMERGING APPROACHES FOR LOAD BALANCING IN CLOUD

Roopali Gupta
Department of Computer Science, AKTU
CET-IILM-AHL,
Greater Noida, UP, India

Abstract-Cloud computing is one of the emerging technologies in the last decade. The cloud computing is growing and spreading very fast due to its easy and simple service oriented model provided via internet. As the number of users are growing very fast to access the cloud therefore it is very difficult and challenging task for the service providers to provide maximum resource output to their users. The main objective of the service provider to provide maximum output can be achieved only by implementing load balancing efficient algorithms. Load Balancing in cloud computing will help clouds to increase their capabilities, capacity which results in powerful and reliability cloud. The only objective of the service provider is to provide maximum resource output and this can be achieved by implementing load algorithms which helps in gaining this objectives. This paper proposed efficient algorithms. The first algorithm intends to keep each node busy regardless of the current overload. The second algorithm assigns a task to the node that has the expected minimum completion time. The other algorithm calculates priority during the execution of the system.

Keywords- Cloud computing, load balancing, load balancer.

I. INTRODUCTION

The performances of computational system depend on several concepts, one of which is load balancing. The load balancing mechanism is totally dependent on the amount of work allotted to the system for a specific time period. This is the time where system has to manage and work according to the priority basics. The interaction with factors and some load balancing algorithm which can be applicable for such factors are studied in the current paper. There are several load balancing algorithms for the improvement and optimization of cloud performances. The nature of the algorithm can be dynamic or static, although some algorithms are simple but under some conditions they work more effectively. Cloud computing is a service oriented architecture, which is provided via internet. The only objective of the service provider is to provide maximum resources output and this can be achieved by implementing load balancing algorithm which helps in gaining this objective. The paper proposed an algorithm which is experimented on a virtual machine

environment. The paper is a comparative study of such several algorithms.

1.1 Cloud Computing

Cloud Computing is a general term used to describe a new class of network based computing that takes place over the Internet, basically a step on from Utility Computing. In other words, this is a collection/group of integrated and networked hardware, software and Internet infrastructure (called a platform). Cloud computing provides a variety of computing resources, from servers and storage to enterprise applications such as email, security, backup/DR, voice, all delivered over the Internet. The Cloud delivers a hosting environment that is immediate, flexible, scalable, secure, and available – while saving corporations money, time and resources.

The capability provided to the consumer is to use the provider's applications running on a cloud Infrastructure. The applications are accessible from various client devices through a thin client Interface such as a web browser. The consumer does not manage the underlying cloud.

II. LOAD BALANCING

Cloud computing is one of the fastest implementing technology in the decade. Many companies are trying to implement and introduce clouds, due to its simple and flexible architecture. These result in the increasing number of users reaching cloud. Although clouds are bifurcated in public private and hybrid models but still problem of reliability may arise in these clouds. Cloud computing has been adopted by organization which includes, social networking websites, online application design by Google app managers and by Google doc which are some of the important implementation and a step ahead in cloud computing. Some clouds are also designed for online software testing. This all suggests that cloud computing will change the way we interact with the resources via Internet. Cloud models used virtualization technology; this technology helps in slicing a single data centre or high power server to act as multiple machines. It depends on the hardware configuration of the data centre or server in how many virtual machine they can be divided. To



implement virtualization additional software is also required. This software is the system software an operating system, can be from windows for example windows server 2008 or Hyper-V or for an open source environment like Linux Load balancing is the pre requirements for increasing the cloud performance and for completely utilizing the resources.

Load balancing is a process of reassigning the total load to the individual nodes of the collective system to make resource utilization effective and to improve the response time of the job, simultaneously removing a condition in which some of the nodes are over loaded while some others are under loaded. Load balancing is centralized or decentralized. Server load balancing addresses several requirements that are becoming increasingly important in networks:

- Increased scalability
- High performance
- High availability and disaster recovery

First the requests or job coming from the user side are stored in a job pool or the central middleware then these jobs are partitioned and making the replication of these partitioned jobs into their local middleware. Thus adding or removing of any node does not affect the whole system. And the replication strategy of the partitioned jobs ensures the fault tolerant by the internal interaction among the node i.e. if any of the nodes fail the total system does not affect. The job queue in each middleware are updated the job status at the time when a task is assigned and whenever it is completed the execution.

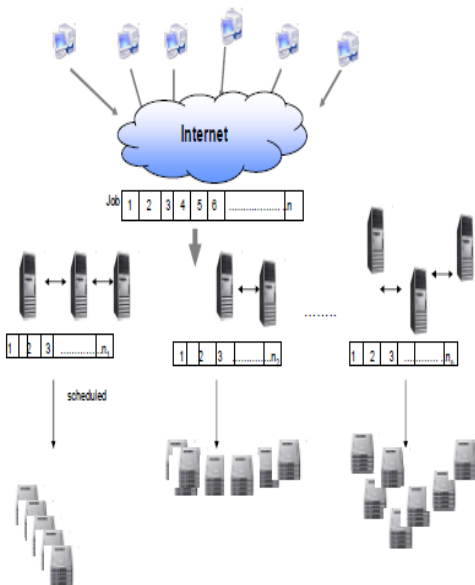


Figure 1: Load balancing in cloud

Load Balancing algorithms are used for implementing. Several load balancing algorithm are introduced like round robin algorithm a mining improvement in the performance. The only differences with this algorithm are in their complicity. The effect of the algorithm depends on the architectural designs of the clouds. Today cloud computing is a set of several data centres which are sliced into virtual servers and located at different geographical location for providing services to clients. The objective of paper is to suggest load balancing for such virtual servers for higher performance rate.

III. LOAD BALANCING ALGORITHMS

The paper describes about three load balancing algorithms which are Equally Spread Current Execution Load, Minimum Completion Time Algorithm, and Priority Scheduling Algorithm.

Equally spread current execution load: This algorithm requires a load balancer which monitors the jobs which are asked for execution. The task of load balancer is to queue up the jobs and hand over them to different virtual machines.

Minimum Completion Time Algorithm: This algorithm assigns a task to the node that has minimum expected completion time. The node which has minimum expected completion time would be given highest priority.

Priority Scheduling Algorithm: In this algorithm, priorities are calculated during the execution of the system.

During the execution, priorities are defined. It allows out-of-order processing according to priority. In real life, systems typically use a mixture of algorithms that together allow for prioritized and out-of-order processing.

IV. PROPOSED ARCHITECTURE

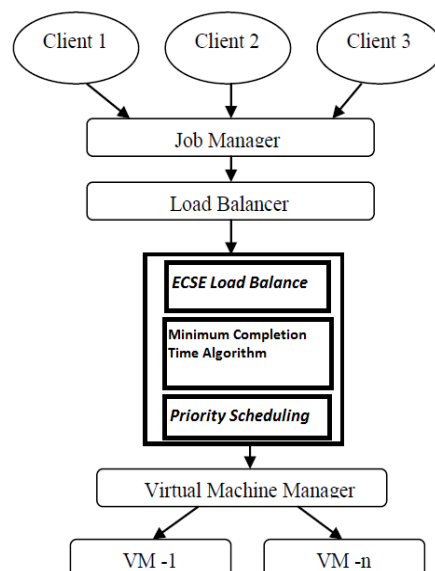


Figure 2: Load balancing Algorithms Execution



V. CONCLUSION AND FUTURE SCOPE

Load Balancing is a very essential task in Cloud Computing environment for achieving maximum utilization of resources. 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 difficult to simulate but are best suited in 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 and on the other hand, hierarchical algorithm divide the load at different levels of hierarchy with upper level nodes requesting for services of lower level nodes in balanced manner. Hence, dynamic load balancing techniques in distributed or hierarchical environment provide better performance. However, performance of the cloud computing environment can be further maximized if dependencies between tasks are modelled using workflows. In next level, we are going to compare all the load balancing algorithms which are existing in Cloud analyst tool and propose a new improved algorithm which will give better results in terms of response time and will reduce cost.

VI. REFERENCES

- [1] Wei-Tek Tsai*, Xin Sun, Janaka Balasooriya "Service-Oriented Cloud Computing Architecture" Computer society 2010.
- [2] Zenon Chaczko, Venkatesh Mahadevan, Shahrzad Aslanzadeh and Christopher Mcdermid," Availability and Load Balancing in Cloud Computing" IPCSIT vol.14 (2011).
- [3] Anthony T. Velte ,Toby J. Velte, Robert Elsenpeter, " Cloud Computing: A Practical Approach ", The McGraw-Hill Companies(2010), [Book]
- [4] Zenon Chaczko, Venkatesh Mahadevan , Shahrzad Aslanzadeh and Christopher Mcdermid," Availability and Load Balancing in Cloud Computing" , IACSIT 2011.
- [5] Ratan Mishra and Anant Jaiswal "Ant colony Optimization: A Solution of Load balancing in Cloud", IJWesT, 2012
- [6] Jasmin James, Dr. Bhupendra Verma "Efficient VM load balancing algorithm for a cloud computing environment" IJCSE, 2012