A Novel Approach for Load Balancing in Cloud Data Center

Gulshan Soni[1],
[1]*M.E. Student, Department of Computer Science,*
NITTTR, Chandigarh, India
[1]gsoni260@gmail.com,

Mala Kalra[2]
[2] *Assistant Professor, Department of Computer Science,*
NITTTR, Chandigarh, India
[2] malakalra2004@gmail.com

Abstract- **In a large-scale cloud computing environment the cloud data centers and end users are geographically distributed across the globe. The biggest challenge for cloud data centers is how to handle and service the millions of requests that are arriving very frequently from end users efficiently and correctly. In cloud computing, load balancing is required to distribute the dynamic workload evenly across all the nodes. Load balancing helps to achieve a high user satisfaction and resource utilization ratio by ensuring an efficient and fair allocation of every computing resource. Proper load balancing aids in minimizing resource consumption, implementing fail-over, enabling scalability, avoiding bottlenecks and over-provisioning etc. In this paper, we propose "Central Load Balancer" a load balancing algorithm to balance the load among virtual machines in cloud data center. Results show that our algorithm can achieve better load balancing in a large-scale cloud computing environment as compared to previous load balancing algorithms.**

Keywords— ***Load balancing, Cloud Data Center, Live Virtual Machine Migration, Virtualization, CloudAnalyst***

I. INTRODUCTION

The cloud means the applications and services that are offered from data center to all over the world. These applications and services are offered over the internet. The services provide by cloud computing are infrastructure as a service (IaaS), platform as a service (PaaS) and software as a service (SaaS) that are made available as pay-as-you-go model to clients. Cloud Computing Deployment Model refers to the location and management of the infrastructure cloud services. The Deployment Model of cloud computing are Private Cloud, Community cloud, Public cloud and Hybrid cloud. Cloud Computing contain some essential characteristics that are rapid elasticity, on-demand self-service, resource pooling, broad network access, and measured service. Cloud computing is based on the concept of virtualization. Virtualization is a method for creating what are called virtual servers that run on a cluster of a number of real servers. Virtualization allows for a smaller number of high-powered servers to create a larger number of less-powered servers while reducing the overall cost in space, power, and other infrastructure.

Currently cloud computing is becoming popular among users and corporate world but despite of growing uses of cloud technology, many crucial problems still need to be solved for the realization of cloud computing. Load balancing is one of these problems, it plays a very important role in the realization of Cloud Computing. Load balancing means the ability to distribute the load over a number of separate systems therefore the overall performance of processing the incoming requests increased. There are four major resources processor (CPU), memory (RAM), network and storage (Disk). In traditional computing environments, researchers [16, 17, and 18] have proposed various static, dynamic and mixed load balancing policies. Static load balancing algorithm assign load to machines according to their processing capability but do not consider dynamic changes of these attributes at run-time. Commonly used static algorithms are Round Robin (RR) & Weighted Round Robin (WRR). Dynamic load balancing algorithm collects information and run times conditions of machines and according to gathered characteristics assign and dynamically reassign the load among machines. Least connection (LC) and weighted least connection (WLC) are dynamic load balancing algorithms commonly used.

In the cloud computing environment, load balancing is required to achieve short response time and high system throughput. For cloud environment various load balancing approaches have been proposed such as Honeybee-based load balancing technique [3], Active Clustering [3], Random sampling [3], Active Monitoring Load Balancer [4], Throttled Load Balancer [4], WCAP [6], JIQ [7], CLBVM [13] etc.

The rest of this paper is organized as follows: Section II gives a review of the related work which realizes load balancing importance in cloud computing. Section III introduces the proposed approach for load balancing in cloud. Section IV describes the experimental setup for implementation of the proposed algorithm. Section V analyses the performance of the mechanism. Finally conclusion of the work is discussed in Section VI along with the envisaged future work.