A Comparitive Study of different Load Balancing Algorithms for Cloud Computing

Mehak Choudhary[1], Dr. Deepti Gupta[2], Dimple Chandra[3], Chitvan Gupta[4]

Assistant Professor, Department of Computer Science, NIET, Greater Noida, India[1,2,3,4]

Abstract: Emerging era in the field of Information Technology is Cloud Computing which provides services to the users over the internet. It is paid services which depends on the usage. The problem associated with this is load balancing which means to use minimum resources to improve the performance of the system.Load Balancing is the process of dynamically allocationg the work load among nodes so that not evena single node would be heavily loaded. In this paper, we have compared the performance of different load balancing algorithms on the basis of their response time.

Keywords: Cloud Computing, Computing Computing Security, Security Threats, VMM.

I. INTRODUCTION

Cloud computing is most widely used term in this era. Large number of organizations currently using cloud computing on large scale. It is based upon the uses and consumption of computing resources. Cloud computing is the computing where computer resources are shared instead of local servers or some personal devices for handling a particular application. In Cloud computing, the term "cloud" stands for   the "internet" so the meaning of term "cloud computing" is "computation depending on internet". Cloud computing is concerned with storage and data accessibility through   internet   instead of using computer's   hard   drive.Cloud Computing   can   be categorized into two ways:
➢ On the basis of services provided
➢ On the basis of locality of cloud computing

Cloud computing is distributed into three types on the basis of services provided:

SaaS: Software as a service is at the top position out of all the three services. Through SaaS, one application can be delivered to many users regardless of their location by the use of cloud computing infrastructure, rather than traditional model that uses one application per desktop.SaaS takes away the organization's need to installation handling, setting and maintaining. Charging on per-user and per-month basis is provided in SaaS and organization has the flexibility to add or remove the users at any time without any extra cost than monthly per-user fee. Example of SaaS are Google+, Gmail email services can be used. Common and popular example of CRM (Consumer Relationship Management) SaaS application is Sales force.

PaaS: Platform as a service is sandwiched between IaaS and SaaS. PaaS is released from the SaaS, where delivered services are cloud computing platform over the web.PaaS have the facility for developing web applications without the need of buying and managing hardware and software. With this technology a third-party provider manages virtualization, operating system, storage, networking and PaaS software itself. Example of PaaS is Microsoft Azure, Force.com, Google.

IaaS: Infrastructure as a service is at the initial level among   the three services.IaaS provides services related to in cloud computing.IaaS provider provides space in virtual datacenter for cloud buyer's rent, rather than leasing space in datacenter, real estate and all utilities to maintain computer server and storage.Cloud Providers allows user to rent virtual server and storage while having network to tie them together. Flexible, standard and virtualized operating environment are the basic goals of IaaS. Examples in IaaS are Amazon, VMware, Microsoft, Rack space and Red Hat.

Cloud computing is divided into following types on the basis locality:

Public Cloud: In Public Cloud, services include resources over an internet connection with pay-per-use fee. Users don't need to purchase hardware to use the services , their use can be scaled on-demand. Infrastructure and pool resources are managed by public cloud providers for the capacity required by user .The infrastructure is owned by organization providing cloud services rather than users. Public clouds are available to large organizations that are owned by third party organizations that offers cloud services. Due to this sharing of resources, public clouds have better performance but a big hindrance is security. Google App Engine, Amazon Elastic Compute Cloud (EC2), Windows Azure Services Platform and Microsoft are some of the public cloud service provider.

Private Cloud: For specific group or organization and for limited access to that group private cloud is used. The purpose is to offer the services within the organization rather than using in public. Data Centre of a company hosts the private cloud and services are provided to the users only within the company. More security and privacy are provided by it than public clouds. It is also called "internal cloud" or "enterprise cloud" . Hybrid Cloud: Composition of two or more cloud models (public, private and community) is called hybrid model. Cost and scale benefits are provided with hybrid cloud with security concerns. For Example- applications with most security concerns are hosted on private clouds while applications which don't require much security concern are mounted on public clouds.

Community Cloud: Community cloud lies between private cloud and public cloud with respect to target set of consumers. Objective of community cloud is to have benefits of public cloud that is shared infrastructure costs and pay-as-you-go billing structure with added private cloud's benefits of security and privacy.

II.   LOAD BALANCING

Load Balancing is a technique which results in maximum throughput and minimum response time to facilitate the networks and resources.Traffic can be divided among servers without any delay.Different Load Balancing algorithms are available that helps to load traffic among servers.

There are two types of load balancing algorithms that are:

- Static Algorithm: Static Algorithm are the load balancing algorithms where servers are allocated with traffic equivalently. In this algorithm traffic is divided equally so it is called round robin algorithm
- Dynamic Algorithm: Dynamic Algorithm are the load balancing algorithms where lightest server is searched over the network to balance the load.

Qualitative metrics of load balancing

- Throughput: It is used for calculation of tasks that completed their execution. Performance of the system is proportioned directly to throughput that is when throughput is high performance is better.
- Migration time: The time required in transferring tasks from one machine to other machine in the system. The time should be less for better system's performance.
- Response time: The minimum time required to run a particular load balancing algorithm to response a task in a system. This metric should be less for better performaygnce.

III.   RELATED WORK

Velagapudi Shreenivas[1] et al. stated the concept of load balancing techniques in cloud computing that distributes the dynamic workload across the multiple nodes evenly so that there will be no overloading in a single node and also there will be improvement in performance and resource utilization. It focuses on minimizing the resource consumption that will further affect the energy consumed and the rate of carbon emitted. In this paper some existing load balancing technologies are discussed on some parameters performance, scalability and overloading. Load balancing algorithm is classified into two categories that are static algorithm and dynamic algorithm. Static algorithm divides the traffic equally among servers. It further has Round-Robin and Weighted Round-Robin. Dynamic algorithm is that through which among whole server lightest server is selected to balance traffic. Searching performance is improved through this technology. Server based load balancing for internet distributes services which further proposed policy of server based load balancing for web server.

Chung-Cheng Li[2] et al. proposed, proposed SLA-aware load balancing scheme for cloud datacenter which further uses tldlb(two –level decentralized load balancer) algorithm for balancing the virtual machine's load using nn-wrr(neural-network weighted round robin).Loads are in the form of batches of job. Here nn-wrr will increase the overhead in one virtual machine because while assigning jobs to virtual machine it follows a circular fashion so the virtual machine having node already will be assigned with upcoming job again so their will more jobs than it can handle. So, nn-wrr is required to be replaced with some other algorithm for balancing the load. Also this model considers all the metrics that are CPU, memory, network bandwidth and disk input output without the violation of SLA.

Bhaskar Prasad Rimal[3] et al. discussed about cloud computing, its type, features and their architecture. This paper helped researchers and developer for the ideas of issues of cloud computing for research. Cloud computing offers services that are used on the basis of pay-per-use. The services offered by it are SaaS (Software as a Service), Paas (Platform as a Service) and IaaS (Infrastructure as a Service). Cloud is of four types on the basis of location public, private, hybrid and community. Private and public clouds are considered for analysis and energy used for transmission, switching and storage of data.

Jihua Hu[4]et al. presented the strategy of scheduling on virtual machine for virtual machine's load balance on the basis of genetic algorithm that is on the basis of data collected historically and system variation. The strategy

used in this paper will help to reduce migration costs using genetic algorithm which further uses coding of population, then population initialization then particular fitness function is applied on the basis of which selection strategy is performed then crossover operation is performed followed by mutation strategy.

Mohammad Alhamad[5] et al. discussed the SLA framework which is the negotiation between clients and service provider for the particular services. It includes three criteria for the negotiation between client and service providers. It is used further as a tool for trust management system in cloud computing for the selection of reliable services.

IV.   COMPARATIVE ANALYSIS

This section mainly concerned with the comparison and its analysis based on the simulation performed on CloudSim. All results of simulation are performed on the basis of different load balancing algorithms that are nn-dwrr, round robin, random while providing user defined cloudlets (request for resources) on 5 virtual machines by one user. The performance parameters which are compared are the results of average response time.

TABLE I. Load balancing simulation parameters

Simulation Parameters	Values
Virtual Machine Hyperviser	Xen
No. of Virtual Mahines	5
Duration(time limit)	60 milliseconds
Length of Cloudlets(user requets)	5000
Total no. of cloudlets	50

Three algorithms are being compared that are round-robin, random, nn-dwrr algorithms on the basis of their average response time by simulating on cloudsim.We will analyse the output on the basis of 50 Cloudlets are being entered by user, 1 user and 5 VMs.

A. Graphical Analysis
Graphs are plotted for the performance evaluation of three different algorithms random, round-robin and nn-dwrr on the basis of parameters average response time. Results are evaluated on the basis of 50 cloudlets.
Following table is used for the graph analysis.

TABLE II. Average Response Time

Algorithms	Average Response Time
Random	98.81
Round-Robin	82.5
NN-DWRR	78.86

Response time is the amount of time server takes to return the results of requests to the user
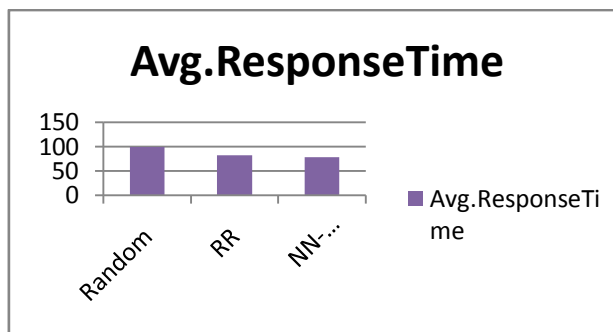

Fig. 1.   Comparison of three algorithms on the basis of avg.response time

Above graph shows the average of response time produced by three different load balancing algorithms. Average response time for the system should be minimum for the better performance. Results of graph shows that algorithm NN-DWRR performs better than Round Robin, on the basis of average response time. Here also we have used 50 cloudlets for the evaluation.

V.   CONCLUSION AND FUTURE SCOPE

This research presented shows the comparative analysis of different load balancing algorithms. Experimental results have shown that NN-DWRR performs really better than other two algorithms that is random and round robin algorithms because it has very small average response time for handling the particular task.
We have discussed the cloudlet (request for resources) assignment to virtual machine without the overloading among one virtual machine and with small response time. In future, we can perform the algorithm which takes very small response time for performing particular task.

REFERENCES

[1] Velagpudi Sreenivas,Prathap.M and Mohammad kemal "Load Balancing Techniques: Major Challenge in Cloud Computing – A Systematic Review", IEEE, 2014.
[2] Chung-Cheng Li and Kuochen Wang "A SLA-aware Load Balancing Scheme for Cloud Datacenters",IEEE,pp:58-63,2014.
[3] Bhaskar Prasad Rimal ,Eunmi Choic and Ian Lumb "A Taxonomy and Survey of Cloud Computing", IEEE.pp:44-51, 2009.
[4] Jinhua Hu,Jianhua Gu,Guafei Sun and Tianhai Zhao "A Scheduling Strategy on Load Balancing of Virtual Machine Resources in Cloud Computing Environment", IEEE, pp: 89-96, 2010.
[5] Mohammad Alhamad,Tharam Dhillon and Elizabeth Chang "SLA-Based trust Model for Cloud Computing", IEEE, pp: 321-324, 2010.