Work Flow Model, Major Challenges and Issues of Dynamic Load Balancing Scheme in Public Cloud Service

Mangal Nath Tiwari[1], Dr. Nagesh Salimath[2], Dr. Atul Kumar Dwivedi[3], Dr. Sandeep Sharma[4]

[1]Research Scholar, Department of Computer Science and Engineering, Madhyanchal Professional University, Bhopal, India

[2, 3, 4]Associate Professor, Department of Computer Science and Engineering, Madhyanchal Professional University, Bhopal, India

Abstract: *Today Cloud Computing has created a new record in user services, through which the user can access his data at any time and at any place, as well as access information to technology services on the basis of payment and use. Nowadays, due to the flexibility of cloud in every sector, all organizations are transferring their business to cloud, as well as all service providers are building new data centres to serve their users. Along with this, it is also necessary to give that the cost of the works and proper use of the resources can also be done because if the resources are empty and they are not being used, then it is mandatory to make proper arrangements for that too. In the presented research paper, many techniques and information have been given to improve the utilization on the basis of load balancing, scheduling tasks, resource management, quality of service, and workload management. Load balancing in cloud computing allows all data centres to be protected from overloading and under loading virtual machines, which is a challenging task in cloud computing.*

Keywords: Load balancing, cloud computing, resource allocation, task scheduling, virtual machine, workload management, optimization

1. Introduction

Cloud Computing has emerged as a new technology in the last several years. Due to which distributed system systems have evolved towards large scale computing networks. Today, cloud computing organizations like IBM, Amazon, Yahoo or Google are providing cloud services to consumers around the world. In this new technology, the end user is not required to install any kind of apps, instead end - user apps and services are offered on demand. Today, there are various difficulties in achieving the real state of the cloud environment. Among those difficulties, load balancing is a major concern.

In load balancing, it is defined that how to balance the load among different types of servers so that there is no problem of overload on any one server, nor there is any problem of under load on any one server, as well as the user can continue to receive services properly. Under load balancing, the required load is managed on various computer resources, computer clusters, such as servers, network links, disks, CPUs, etc. Load balancing provides methods to maximize system resource utilization, output, and device performance. Also, it gives us an advantage of keeping the data or files in a scalable and easy way and provides them convenience to reach the customer. This paper discusses the major challenges and issues of cloud computing and discusses the load balancing problems.

2. Related Works

Load balancing is a technique which is responsible for distributing the load over one or additional servers, networks or other IT resources to a number of users [6]. Load balancing is the main concerns and need more research. In this section, we have reviewed some articles that have significantly surveyed various challenges and issues of load balancing techniques or schemes under cloud computing environment.

Authors Pawan Kumar and Rakesh Kumar [5] have classified load - balancing techniques into several categories. They have discussed advantages, disadvantages, concepts, and challenges with these techniques. Authors explained that there is a need to improve the techniques for increasing performance of the system in the future. They have stated different challenges to be addressed to improve the load balancing techniques for better of the system in future.

K. Balajiet. al. [11] presented a study for massive research has been done on several load balancing approaches. Author differentiated various load balancing issues for developing algorithms that works effectively for hybrid, and agent based load balancing techniques.

Shalini Joshi et. al [9] presented an overview of cloud computing, cloud computing architecture, virtualization, load balancing and some challenges related to balancing load in cloud computing. They also described major challenges occurs when researchers develop any load balancing algorithm for cloud environment, they also explained how these issues affected the performance of load balancing task.
Authors Ghomi et al. (2017), has discussed various issues and challenges in existing load - balancing techniques along with future directions.

Milani and Jafari (2016) has reviewed different load balancing techniques and also explained the classification of these techniques into dynamic and hybrid approach. They also described the behaviour of these techniques based on different parameters, along with their advantages, disadvantages, and challenges. Authors have explained several issues with these algorithms to develop more

efficient algorithms to make load - balancing techniques more effective.

Almubaddel et. al. [13]Have done a careful review of the terms of service or contracts, various issues &challenges of the load balancing scheme for cloud computing environment. The authors Singh et. al. [22] had stated and described various issues related to task scheduling in the cloud and presented a comparative analysis for both dependent and independent tasks based on the meta - heuristic approach. They also proposed some better approach to develop efficient load balancing algorithms.

Load Balancing Technique
Its objective is to provide continuous service in case of failure of any service's component by provisioning and de - provisioning the application instances along with proper utilization of resources. Load balancing (LB) is an approach to divide the coming traffic between available servers to handle the requests [10].

The load can be of any type like network load, memory load, CPU load and delay load etc. It is very important to share work load across multiple nodes of system for better performance and increasing resources utilization. In cloud computing environment, various types of algorithms are present that enable load traffic among accessible servers [3]. To balance the entire load, various load balancing algorithms has been designed. We can categorise load balancing algorithms in two ways: A. Static Load Balancing B. Dynamic Load Balancing.
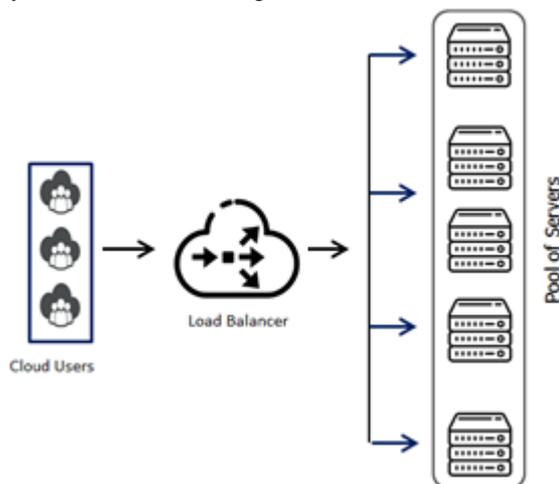

Figure 1: Load balancing

[A] Static Load Balancing (SLB) –static load balancing technique distribute the incoming user traffic equally to all available servers using the specific algorithms, these algorithm has prior knowledge of each resources and servers state. This load balancing technique is suitable for such places where there is little fluctuation of incoming user traffic. One of the biggest problems in this type of method is that the load allocated to a server cannot be transferred to another server when required.

Examples of static load balancing algorithms are: 1. Round robin, 2. Weighted round robin, 3. Source IP hash, 4. URL hash, 5. Randomized algorithm.
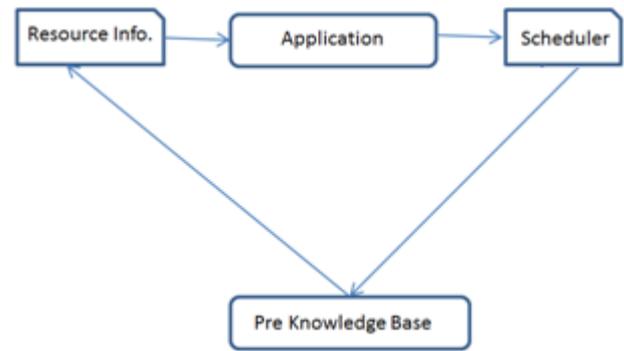

Figure 2: Static load balancing

[B] Dynamic Load Balancing (DLB) – Dynamic load balancing technique divide the incoming users traffic is dynamically between servers. It does not require prior in - depth knowledge of system resources. This technique is designed for systems with high fluctuations of incoming load. If required, allocated load of any server can be reallocated between servers to reduce resource underutilization [14].

There are several dynamic load balancing algorithms exists these are: 1. Least connection method, 2. Weighted least connections method, 3. Least response time method.
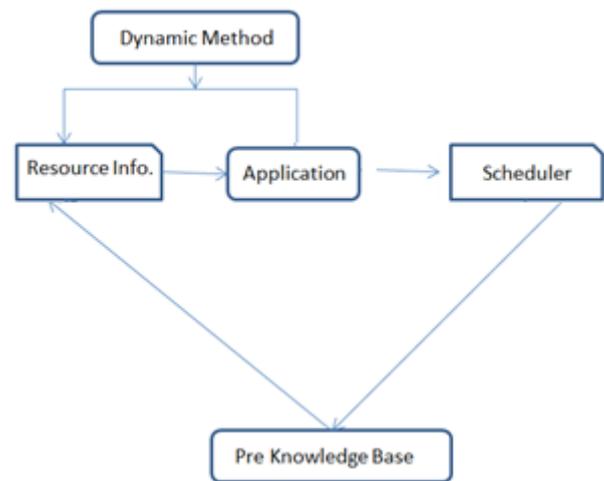

Figure 3: Dynamic load balancing

Major Goals of Load Balancing: The main purpose of load balancing scheme is to speed up applications with different resources and reduce response time. Some major goals of load balancing technique for cloud computing environment are given bellow [12].

1) Providing system stability and scalability.
2) Increase efficiency of the services on demand.
3) Improve the performance and efficiency with fault tolerant system.
4) Enhance user satisfaction with virtualization approach.
5) Having a backup plan for systems that fail.
6) Increase the resource utilization ratio.
7) Minimizing the job execution time and waiting time in queue
8) Maximize the overall system performance.

Load Balancing Model:

The Cloud technology works on the principle of providing accessibility to a group of resources i. e. physical servers, Virtual servers, memory and networks to its users as and when required. To manage the user requests effectively, a quality load balancing algorithm needs to assign the jobs to respective VM's depending on Quality of Service (QoS) requirements. A model and work flow of the load balancing are laid out below in fig.4.

In this model, many requests are received from different user bases for the services of cloud service providers. When the cloud detects a user request from a user base, the cloud broker first examines and analyzes the resources, cost, and performance demanded by that request. After resource analysis, the service broker transfers that request to the selected data center, Where the Data Center controller (DCC) accepts that request for further process. Data center controller sends all requests to available load balancers on available servers. After that, each load balancer distributes all the requests received to virtual machines with the help of a certain load balancing algorithm based on VM availability. In the user request distribution process, it is determined that which request will be processed or executed by which Virtual Machine (VM). Through the load balancer, all incoming user requests are transferred to the virtual machine under a certain policy for execution.
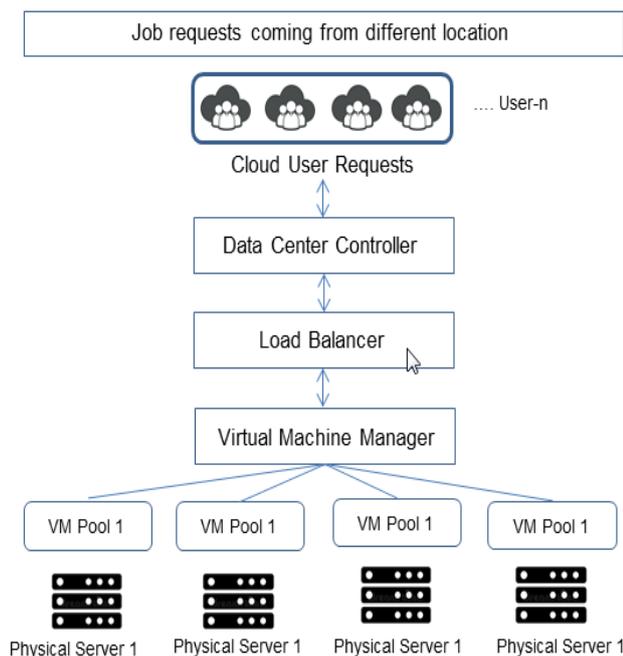

Figure 4: Load balancing model

During this process, the load balancer maintains a state table, in which the state of all available and allocated virtual machines is maintained and load balancer updates the status table from time to time based on availability and allocation. During request distribution or VM allocation, if any VM is not available, then the load balancer stores all the received requests separately in a queue, and stores it until the requested resource is not available. Once a VM finishes the execution of a request, then the load balancer allocate that available VM to another requests. During the whole process,

it is the responsibility of the load balancer, that no VM is neither over - loaded nor under loaded.

3. Major Challenges & Issues of Load Balancing Scheme

Cloud Computing technology is emerging as the most targeted research area. In that particular field, load balancing is appearing as one of the major challenging high - end concern from the researchers. Following are some load balancing challenges and issues that are mentioned below:

a) *Heterogeneity of Computing Nodes:* Since cloud computing environment deal with the heterogeneity of virtual nodes for computing resources, many researchers developed load balancing algorithms for heterogeneous nodes in cloud computing to achieve better performance. In present scenario it is needed that new algorithms are to be developed that full fill re requirement of heterogamous systems for carrying out the user requested operation very efficiently.

b) *Single Point of Failure:* In cloud computing environment, load balancing process is controlled by a single central system or virtual machine. Dynamic load balancing task is completely monitored by this central virtual machine. In the case when central master virtual machine gets crashed, the whole process gets affected. Hence there is a need of such load balancing algorithms so that if central server fails it does not affect all commuting resources.

c) *Virtual Machine Migration:* Any physical server have more than one virtual machines installed or configured on it. These single virtual machines (VM) act as a separate individual working system. If in a case where any physical servers gets overloaded during the load balancing then some virtual machines installed in it, needed to be migrated or transferred from one machine to another dynamically. Dynamic load balancing algorithm development for such case is a very challenging task to be addressed. Any dynamic load balancing technique should have the ability to handle such virtual machine (VM) migration issue.

d) *Algorithm Complexity:* Any algorithm which has more complexity, it has provide lesser performance and efficiency for load balancing approached in cloud computing. So any load balancing algorithms should be always less complex (simple) and easy to implement to get better performance of overall system.
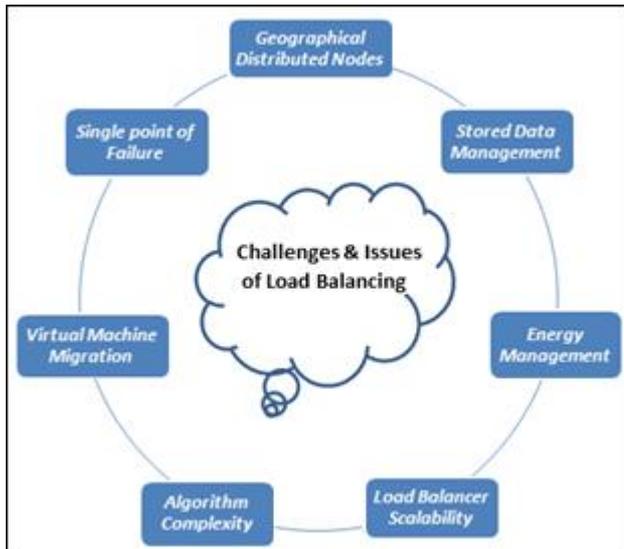
Figure 5: Major Challenges and issues of cloud load balancing

a) ***Load Balancer Scalability:*** Scalability is a major concerning issue in workload management task done by any load balancer. Any user wants to access required resources efficiently from any location and in any time demand to facilitate the system work effectively. So today, it is demand of such dynamic algorithms used by load balancers that provide more scalability.

b) ***Energy Management:*** Any physical server with many virtual machines (VMs) installed on it, consumes massive source of energy. So researchers are expected to develop such dynamic algorithms that deal withserver energy consumption issues during load distribution process. Energy saving is a key point of all dynamic algorithm used to deal with load balancing in cloud computing.

4. Conclusion and Future Work

This research paper is focused on cloud computing problems and its major challenges. Cloud computing is state - of - the - art computer technology which delivers customer support at all times. Load Balancing is one of the biggest problems with Cloud Computing, as overloading a device will lead to terrible results that could create technology obsolete. So there is always a need for an effective Load Balancing algorithm for efficient use of resources. In this study we have reviewed several massive researches has been done related to load balancing scheme in cloud computing. This paper explains about some major challenges and issues facing by load balancing management for cloud servers. Since cloud computing technology is a new field of research and load balancing is a major concern under this technology, lots of research works are being supposed to carry out for developing out good quality algorithms and techniques. Definitely this work will be very helpful for the researches to solve the various discussed challenges and issues of load balancing in cloud environment.

References

[1] Buyya R., R. Ranjan and RN. Calheiros, "InterCloud: Utilityoriented federation of cloud computing environments for scaling of application services, " in proc.10th International Conference on Algorithms and Architectures for Parallel Processing (ICA3PP), Busan, South Korea, 2010

[2] Foster, I., Y. Zhao, I. Raicu and S. Lu, "Cloud Computing and Grid Computing 360 - degree compared, " in proc. Grid omputing Environments Workshop, pp: 99 - 106, 2008.

[3] Muhammad AsimShahid, Noman Islam, Muhammad Alam, M. S. Mazliham, Shahrulniza Musa. "A comprehensive study of load balancing approaches in the cloud computing environment and a novel fault tolerance approach", IEEE Access, 2020

[4] Almubaddel, M., &Elmogy, A. M., "Cloud Computing Antecedents, Challenges, and Directions. Proceedings of the International Conference on Internet of Things and Cloud Computing – ICC", 2016

[5] Pawan Kumar and Rakeshkuamr, "Issues and challenges of load balancing techniques in cloud computing: A survey". ACMComput. surv.51, 6, Article 120, 2019.

[6] Muhammad AsimShahid, Noman Islam, Muhammad MansoorAlam, MazlihamMohdSu'ud, Shahrulniza Musa. "A Comprehensive Study of Load Balancing Approaches in the Cloud Computing Environment and a Novel Fault Tolerance Approach", IEEE Access, 2020

[7] Joshi, S., &Kumari, U., "Load balancing in cloud computing: Challenges & issues.2016 2nd International Conference on Contemporary Computing and Informatics (IC3I) ", 2016.

[8] Somani, Rajkumar, and J. Ojha. "A Hybrid Approach for VM Load Balancing in Cloud Using CloudSim. " International Journal of Science, Engineering and Technology Research (IJSETR), pp: 1734 - 1739, 2014.

[9] K. Balaji, P. SaiKiran, M. Sunil Kumar, "Load balancing in Cloud Computing: Issues and Challenges", Turkish Journal of Computer and Mathematics Education, Vol.12 No.2 (2021), 3077 - 3084.

[10] Rajguru, Abhijit A., and MrsSulabha S. Apte. "Various Strategies of Load Balancing Techniques and Challenges in Distributed Systems. "

[11] Almubaddel, Majed, and Ahmed M. Elmogy. "Cloud Computing Antecedents, Challenges, and Directions. " Proceedings of the International Conference on Internet of things and Cloud Computing, ACM, 2016.

[12] Raghava, N. S., and Deepti Singh. "Comparative study on load balancing techniques in cloud computing. " Open Journal of Mobile Computing and Cloud Computing 1.1 (2014).

[13] Salimath, N., Mallappa, S., Padhy, N., Sheetlani, J., "Scrambling and Descrambling of Document Image for Data Security in Cloud Computing", vol 160. Springer, Singapore, 2020, 283 - 290.

[14] Salimath, Nagesh and Kavitha, Dr. C. and Sheetlani, Dr. Jitendra, Analysis of Resource Management and Security Management in Cloud Computing Environment (May 17, 2019). Proceedings of the Second International Conference on Emerging Trends in Science & Technologies for Engineering Systems (ICETSE - 2019).

[15] Salimath, Nagesh and Kavitha, Dr. C. and Sheetlani, Dr. Jitendra, "Analysis of Resource Management and Security Management in Cloud Computing Environment", Proceedings of the Second International Conference on Emerging Trends in Science & Technologies for Engineering Systems, ICETSE, 2019.

[16] Nagesh Salimath & C. Kavitha, "Secure data compression and error detecting codes for networks and cloud storage", International Journal of Research and Analytical Reviews (IJRAR), Vol.6, Issue 2, page no.452 - 458, May - 2019.

[17] Nagesh Salimath & C. Kavitha, "Review of error detecting codes for networks and cloud storage", International Journal of Emerging Technologies and Innovative Research, ISSN: 2349 - 5162, Vol.6, Issue 5, page no. pp 370 - 375, May - 2019.

[18] Singh, A., Sahu, S., Gautam, K., &Tiwari, M., "Private Cloud Scheduling with SJF, Bound Waiting, Priority and Load Balancing", International Journal of Advanced Research in Computer Science and Software Engineering, 4 (1), 367 - 372, 2014.

[19] A. K. Singh et al, "Scheduling with load balancing in Cloud computing", International Journal of scientific engineering and research, vol.2, no.1, pp.38 - 43, Jan - 2014.

[20] Mangal Nath Tiwari, Vijay Anand, "Perspective study of Public Cloud: A highly scalable Cloud deployment model", International Journal of Science & Engineering Development Research ISSN: 2455 - 2631, Vol.7, Issue 9, pp.848 - 853, 2022.

[21] Mangal Nath Tiwari, Nagesh Salimath, Vijay AnandSullare, "Analysis of Various Performance Measurement Parameters (PMP) of Load Balancing Algorithms Used in Public Cloud Environment", International Journal of Computer Sciences and Engineering, Vol.11, Issue.2, pp.12 - 17, 2023.

[22] Tiwari M., Gautam K., and Katare K., "Analysis of Public Cloud Load Balancing using Partitioning Method and Game Theory", International Journal of Advanced Research in Computer Science and Software Engineering, 4 (2): pp.807 - 812, 2014.

[23] Mangal Nath Tiwari, Tripathi A., and Sharma S., "Cost And Power Consumption Analysis Of Smartphone Users In India", International Journal for Science and Advance Research In Technology, Vol.8, Issue.10, pp.12 - 17, 2022.