A Survey of Various Load Balancing Algorithms In Cloud Computing

Narendra Rao Tadapaneni
Sr. Software Engineer
T.Rowe Price, Inc. MD, USA

Abstract- *Cloud Computing is becoming the most advanced and popular technology providing the world improved methods of storage, effectiveness of the shared resources like storage, computation power, dynamic allocation of resources based on the demand. This emerging technology has lots of challenges to share the resources, increase the availability and maintain the load between the resources. Cloud load balancing is one of the main challenge which will distribute the dynamic workloads and computing resources in cloud environment between the resources evenly. This paper emphasize on a various existing load balancing algorithms, advantages.*

Keywords- cloud computing, load balancing algorithms, cloud service challenges

I. INTRODUCTION

Cloud Computing grew as favorites in recent times. Due to service's part, this provides flexible in retrieving data and easy way for keeping files for making large files and sets of data accessible for numerous consumers to overall world. Managing such kinds of huge sets of data call for many approaches for enhancing and simplifying operations as well as provide perfect efficiency levels for consumers. Load balancing is technique which distributes workload among several nodes within presented workspace so that this ensures no nodes within system is idle or overloaded for every moment [1]. Efficient algorithm of load balancing might clarify each single node within system might have less or more identical quantity of work.

Accountability of algorithm of load balancing is this is managing assignments that are ahead of cloud area of unused services. Hence, overall accessible time for reactions could be enhanced. Additionally, this gives proficient use of resources. Balancing workload continue as one of worries within cloud computing as quantity of the demands could not be figured out which are released in cloud environment [2]. Load balancing's fundamental consideration within platform of cloud is in appointing as well as distributing load dynamically through nodes with certain end goal for satisfying consumer necessities as well as for giving optimal use of resource by arranging overall obtainable load into diverse nodes. Cloud computing and Artificial intelligence are emerging technology there are many security concerns also with these technologies [3].

II. LOAD BALANCING

It is fact to disperse load through many resources within each system. In such manner, load must be distributed over resources in construction modeling based on cloud, as for all resources do around task's identical quantity at each point of time. The elementary need is delivering few approaches for stabilizing demands for giving choice of application quicker [4]. Method of load balancing which makes every processor busy along with for completing works around within same time.

A. Goals

Load balancing's goals are as follows:

- System's stability remains in track.
- Have capability in altering this as per extend or modification within setup of system.
- Promote system of fault tolerant with respect to stamina, performance under system's partial failure.
- Achieve huge improvement within performance.
- Fully utilization of shared resources.
- Increase system's adoptability for adjusting to modifications.

B. Demand

It is technique where this delegates for accomplishing task equally in each attainable node which is shown within the system. Fulfillment of high user is aspect definitely around this. Due to client's large number and their needs are growing, clouds would need in supplying products to visitors. Desirable process of load balancing helps in utilizing available resources, verifying not any node is under load or over load [5]. It facilitates scalability, minimizes time period that is consumed for actually giving responds and prevents

difficulties. Numerous algorithms of load balancing are developed during the order for planning load among maximum machines.

III. TYPES

Load balancing could be classified as system's current state, like dynamic and static load balancing.

A. Static

Load balancing is knowledge of past related to resources and software of the system. Choice for moving workload should not exactly depend on system's present status. It relates with load balancing which distributes workload derived which is related to policies' constant set, associated to qualities of workload. It is not defensive. Hence, each system has minimum one allotted assignment for itself [6]. The static algorithms don't consider dynamic modifications during runtime. Few algorithms of static load balancing are Min-Max, Round Robin and Min-Min algorithm.

B. Dynamic

It doesn't consider system's prior state and no previous understanding is needed. It depends on machine's present status. Usual method is permitted by it for relocating from the machines that are heavily loaded dynamically for obtaining quick execution. In such conditions, there is rise of case communication and turns in more if there is enhancement in variety of the processors [7]. Few dynamic load balancing are the honey-bee foraging, joint-idle queue, biased random sampling and active clustering.

C. Needs

Most people could equalize system's work through relocating workload dynamically to system nearby to the faraway nodes or the systems which are used less. Performing this enhances client's fulfillment, reducing time of reaction, maximizing utilization of resource, decreasing set of quantity of refusals of tasks and boosting system's efficiency stability. In addition, green computing within cloud could be achieved by using the load balancing [8]. Load balancing could cut down easily capacity of utilization of power by keeping distance from machines' overheating due to excess workload. Carbon discharge and energy absorption are two sections for same stage. These are both specifically relative with each other. Minimizing utilization of energy by use of load balancing would cut down automatically discharge of carbon and for such reason produce green computing. Cloud computing technology is emerging [9] and there are many challenges to benefit the from the technology [10].

IV. ALGORITHMS

Following algorithms of load balancing are used currently within cloud computing:

A. Round Robin Algorithm

It is algorithm for static load balancing which uses fashion of round robin to allocate jobs. This scheduling is quite efficient and effective time scheduling policy. The algorithm randomly selects nodes for load balancing. Here, essential role is played by data centers in handling process of the load balancing within cloud computing [11]. When data center's controllers receives request from the user, then this passes request to algorithm of round robin. Within the algorithm, there is division of time in small units which is known as time slice. Hence, the algorithm is specially designed for sharing of time.

Firstly, every processor which could be run, are stored inside circular queue. In defined slot of time, server is allocated by scheduler to every process within queue. When there are new processes, this would be added at queue's end. First process is selected by scheduler from queue randomly. As there is end of time slot of process, process is passed on from server and then attached at queue's tail. If this process is totally completed before time slot, process is voluntary released by it. Server is assigned by scheduler to ready the process within queue. In such way, there is processing of user request in a circular way through using the algorithm [12]. However, due to server's random selection several times, few servers could be overloaded that results in decrement of load balancing's performance. For overcoming this issue, better technique of allocation is introduced and is called weight round robin load balancing algorithm.

B. Min-Min Load Balancing Algorithm

This is fast or simple algorithm which gives improved performance. The algorithm includes task set. There is not assigned tasks initially to any nodes. Hence minimum time for completion is calculated for every available node within system. After calculation, task is chosen which have minimum time for completion and assign to separate node. The time that is currently available for execution is uploaded, then there is removal of task from available set of task. The process is performed until every task would be allocated in equivalent machines [13]. The algorithm works better for situations where there is more number of smaller tasks than

the larger tasks. This algorithm's disadvantage is that this leads into starvation as this assigning smaller tasks firstly, making large tasks wait in waiting stage.

C. Opportunistic Load Balancing Algorithm

The algorithm doesn't analyze virtual machine's current state as this is static load balancing algorithm. This makes effort in keeping all nodes busy. The algorithm manages unexecuted tasks rapidly to nodes that are available within system. Every task could be randomly assigned to node. The algorithm doesn't give load balance good results [14]. Due to such reason, this doesn't calculate current time of execution of node, hence task would process slow with this manner.

D. Max Min Load Balancing Algorithm

The algorithm is same as algorithm of Min-Min Load Balancing. At beginning every task that is available is submitted to system and calculation is done for minimum time for completion for every available task. After the calculation, a task is selected which have maximum time for completion and the task is allocated to corresponding machine [15]. This algorithm's performance is better when compared with Min-Min algorithm as if only single large task is there in task set, then short tasks are run parallel by Max Min algorithm with the large task.

E. Active Monitoring Load Balancing Algorithm

It is algorithm of dynamic load balancing where load is allocated to virtual machine through finding out least loaded virtual machine or idle virtual machine in list. Initially, there is search for null virtual machine if no null virtual machine is there. Further the least loaded virtual machine is chosen. Here index table for every requests and servers which are assigned to servers currently is maintained with help of load balancer. When there is new request, the servers' index table is scanned by the data center which is least loaded or idle. The algorithm uses concept of first come first serve to assign load to server having least index number for greater than two servers [16]. By using the server id, load is allocated to server as well as index table of server is incremented. After completing the task, the data center is forwarded the information as well as server's index table is decremented. When a new request comes, index table is rescanned with load balancer and process allocation takes place.

F. Equally Spread Current Execution Algorithm

It is algorithm of dynamic load balancing where effort is made by load balancer for distributing equal quantity of load between every server which is available in data center. The processes are assigned priority at starting of the algorithm, it then checks capacity and size for transferring load to server that could handle the load in smaller time period [17]. At such point, there is measure of capacity of virtual machine and estimation of load. Load is allocated as per capacity and size of the matching virtual machine.

G. Active Clustering Algorithm

The algorithm defines virtual machine's clustering to balance load within cloud computing. For the algorithm, the clustering is grouping of the objects together that have similar kind of properties [18]. Hence, virtual machines having similar properties are together grouped in the cluster for handing kind of load.

H. Throttled Load Balancing Algorithm

The algorithm is about virtual machine. Throttled Load Balancer (TLB) maintains every process as well as monitors work on servers. Hence, in the algorithm, best virtual machine is found by load balancer for client request which could handle load in an effective way and quite easily. Different virtual machines have different properties and capacity for handling different loads. Hence, as per load, right virtual machine should be selected for load. There is maintenance of maintenance of index table for every server and when the data center I sent request by the client, data center's controller forward request to throttled load balancing. To find idle server that is available, index table is scanned by TLB and send back server id to data center and task is allocated to the servers [19]. Index table after allocation is updated. Whenever controller of data center gets information of task completion there is decrement again in index table. In the algorithm, if no server is there in the idle state, request remains in queue.

V. CONCLUSION

Cloud computing is an emerging trend within IT's era having huge requirements for infrastructures, storage and resources. Load balancing is cloud computing's essential aspect for balancing load in system. Numerous users are allowed in accessing distributed, hardware, software, virtualized and scalable resources over internet by cloud computing. Load balancing is major issue for cloud computing. This is mechanism that distributes workload over every node within overall cloud. This would improve resource utility and overall performance of system. This paper

discusses about load balancing and its goals, demands, types and needs. This paper also analyses algorithms of load balancing within cloud computing. These algorithms of load balancing ensure resources' utilization through distributing load between several nodes within system by use of task scheduling.

REFERENCES

[1] Ghomi, E. J., Rahmani, A. M., & Qader, N. N. (2017). Load-balancing algorithms in cloud computing: A survey. *Journal of Network and Computer Applications*, *88*, 50-71.

[2] Mesbahi, M. R., Hashemi, M., & Rahmani, A. M. (2016, April). Performance evaluation and analysis of load balancing algorithms in cloud computing environments. In *2016 Second International Conference on Web Research (ICWR)* (pp. 145-151). IEEE.

[3] Tadapaneni, Narendra Rao, Artificial Intelligence Security and Its Countermeasures (January 30, 2020). *International Journal of Advanced Research in Computer Science & Technology, Vol. 8, Issue 1. Available at 3553065*

[4] Joshi, S., & Kumari, U. (2016, December). Load balancing in cloud computing: challenges & issues. In *2016 2nd International Conference on Contemporary Computing and Informatics (IC3I)* (pp. 120-125). IEEE.

[5] Hota, A., Mohapatra, S., & Mohanty, S. (2019). Survey of different load balancing approach-based algorithms in cloud computing: a comprehensive review. In *Computational Intelligence in Data Mining* (pp. 99-110). Springer, Singapore.

[6] Lakhani, B., & Agrawal, A. (2018). Pros and Cons of Load Balancing Algorithms for Cloud Computing. *International Journal of Computer Science Engineering & Technology*, *8*(4).

[7] Mukati, L., & Upadhyay, A. (2019). A Survey on Static and Dynamic Load Balancing Algorithms in Cloud Computing. *Available at SSRN 3365568*.

[8] Islam, T., & Hasan, M. S. (2017, October). A performance comparison of load balancing algorithms for cloud computing. In *2017 International Conference on the Frontiers and Advances in Data Science (FADS)* (pp. 130-135). IEEE.

[9] Tadapaneni, Narendra Rao, Cloud Computing - an Emerging Technology (March 1, 2020), *International Journal of Innovative Science and Research Technology (pp. 37-40). Available at SSRN 3553550*

[10] Tadapaneni, Narendra Rao, Cloud Computing: Opportunities And Challenges (March 20, 2018), *International Journal of Technical Research and Applications. Available at SSRN 3563342*

[11] Shakir, M. S., & Razzaque, A. (2017, October). Performance comparison of load balancing algorithms using cloud analyst in cloud computing. In *2017 IEEE 8th Annual Ubiquitous Computing, Electronics and Mobile Communication Conference (UEMCON)* (pp. 509-513). IEEE.

[12] Jain, S., & Saxena, A. K. (2016, November). A survey of load balancing challenges in cloud environment. In *2016 International Conference System Modeling & Advancement in Research Trends (SMART)* (pp. 291-293). IEEE.

[13] Xu, M., Tian, W., & Buyya, R. (2017). A survey on load balancing algorithms for virtual machines placement in cloud computing. *Concurrency and Computation: Practice and Experience*, *29*(12), e4123.

[14] Mazher, N., Riaz, F., Tariq, M., Ishaque, F., & Shakir, H. S. (2018). Performance Comparison of Load Balancing Algorithms using Cloud Analyst in Cloud Computing.

[15] Kumar, P., & Kumar, R. (2019). Issues and challenges of load balancing techniques in cloud computing: a survey. *ACM Computing Surveys (CSUR)*, *51*(6), 1-35.

[16] Babu, K. R., & Samuel, P. (2016). Enhanced bee colony algorithm for efficient load balancing and scheduling in cloud. In *Innovations in bio-inspired computing and applications* (pp. 67-78). Springer, Cham.

[17] Milani, A. S., & Navimipour, N. J. (2016). Load balancing mechanisms and techniques in the cloud environments: Systematic literature review and future trends. *Journal of Network and Computer Applications*, *71*, 86-98.

[18] Adhikari, M., & Amgoth, T. (2018). Heuristic-based load-balancing algorithm for IaaS cloud. *Future Generation Computer Systems*, *81*, 156-165.

[19] Khan, R. Z., & Ahmad, M. O. (2017). A survey on load balancing algorithms in cloud computing. *International Journal of Autonomic Computing*, *2*(4), 366-383.