Comparison of Various Load Balancing Techniques in a Centralized and Distributed Cloud Computing Environment

Maria Tariq¹, Khadija javed²

¹ Department of Computer Science, Lahore College University for Women Lahore

² Department of Computer Science, GC University Lahore
khadijajaved17@yahoo.com

Abstract:

Cloud computing presently plays a significant job in our lives. The cloud is a platform that provides high accessibility, dynamic asset pools, and virtualization. By improving, it enables a huge reduction in cost. Load Balance in cloud computing is very important. For cloud computing, there are different techniques available that have different parameters. Many methods are used centrally and some are used in cloud computing distributed environments. In this exploration, we have thought about both, brought together and conveyed methods after picking the best load-balancing procedure that depends on various parameters, for example, execution, over-burden, throughput, reaction time, and movement time.

Keywords: Cloud Computing, Load Balancing, Centralized, Distributed, Token Routing, Round Robin, Central Queue

1. INTRODUCTION

Cloud computing is a straightforward, productive and promising innovation in which all administrations are accessible on the server-side and the end-client can undoubtedly get to those assets whenever, anywhere from anyplace by sharing code, resources, and data starting from one terminal to the next [1]. There are plenty of resources available in the data center and different applications are accessible when the end-user needs that tool. Cloud computing offers a business model in the computer world at a very low cost[2]. Cloud computing is a blend of parallel registering, appropriate figuring, ubiquitous processing, and system registration. Information can be moved without much of a stretch from versatile gadgets to enormous server sides in distributed computing [3]. The fundamental point of cloud computing is to improve the administration's condition with rapid registering in a minimal effort figuring element as per the client request. Along these lines, to decrease the information preparing trouble, use distributed computing. The principle focal points of distributed computing

are [4].

- Easy adaptability
- High accessibility
- Cost-sparing
- Reduce information handling
- Autoscaling
- Improved testability
- More Effective advancement life cycle

2. LOAD BALANCING METRICS IN CLOUD COMPUTING

There are different load balancing metrics for load balancing techniques[5-8]:

- Performance is used to calculate the system's output. If the price is fair, it is increased.
- Time of Response, is the time that gadget can take to react to a particular burden. The reaction time taken in calculating static burden adjustment is short, however, the reaction time in powerful burden adjustment calculation is high

because the undertaking movement requires a great deal of time.

- **Stability**, It is a term of delay between the processor and the additions in load adjusting calculations to get the best execution in a particular measure of time.
- **Overload Rejection** In extra over-burden, load adjustment is beyond the realm of imagination so there is a need to gauge dismissal. When this circumstance closes at that point, quit estimating over-burden dismissal.
- **Throughput**, is the measure of information that moves through one spot to somewhere else in a particular given time.
- **Resource Utilization**, In programmed load adjustment, asset usage incorporates. So a dispersed framework has an unforeseen number of procedures that request a lot of handling power. On the off chance that the given calculation handles the used assets, they can move under the stacked processor more effectively. Static burden adjustment calculations have the least asset use yet powerful burden adjustment calculations have the most extreme asset usage ability.
- **Fault-Tolerant**, This measurement gives that the calculation is utilized to endure convoluted shortcomings or not. This calculation empowers the working framework to work appropriately even if disappointment comes.
- **Migration Time**, It is a time in which the resources or jobs can be moved from one node to another node.
- **Load Balance**, Scheduling of tasks in the cloud environment is responsible for optimum task and resources matching. A task scheduling algorithm can keep the balance of workload. Load balancing in the cloud, therefore, becomes another important measure [17].

3. LOAD BALANCE IN CLOUD COMPUTING

Burden adjusting is a technique for systems administration that is utilized to disperse the remaining task at hand over numerous PCs, plate drives, focal handling units, and different assets to get to the base reaction time, boost throughput, most extreme asset use and maintain a strategic distance from over-burden in distributed computing. All the dedicated software and hardware provide the load balancing services, such as system server, multilayer switch and domain name.

The focal issue in cloud computing is load balancing[9].it is a component wherein conquer all the remaining task at hand of the considerable number of hubs in the entire cloud.it gets the greatest asset use and uses fulfillment. The primary point of burden adjusting is to improve the exhibition of equalization of the heap that relies upon various assets like plate drives, focal handling units, and system joins. Every one of these assets is utilized to get the most extreme throughput, least reaction time and maintain a strategic distance from over-burden in cloud computing.

Figure 1: Scheduling and load balancing Design[18]

For the most part, load adjusting calculations have two fundamental characterizations: Centralized methodology and Distributed methodology [15-16].

In a **Centralized methodology**, a solitary hub is dependable inside the entire framework to deal with the circulation in distributed computing.

In a distributed methodology, every hub autonomously handles its heap that is gathered by different hubs. By utilizing nearby burden vectors, choices are made locally.in distributed computing, this is progressively reasonable for

dispersed frameworks.

4. LOAD BALANCING ALGORITHMS

There is the accompanying burden adjusting calculations of load balancing algorithms[10]:

4.1. Token Routing

The calculation's principal objective[11] is to limit the expense of the framework by moving the tokens around the framework. Be that as it may, because of the correspondence bottleneck, specialists in a versatile cloud framework might not have enough data to convey the outstanding task at hand. So the circulation of the outstanding burden between the specialists isn't set. The token steering calculation's disadvantage can be expelled with the assistance of a token-based burden adjusting heuristic methodology. This calculation gives the choice to course rapidly and effectively. There is no compelling reason to have a comprehension of the total information on their worldwide state and the working heap of neighbors in this calculation administrator. They are building their insight base to settle on their choice where to pass the token. This information base is gotten from the tokens recently got. Along these lines, no overhead correspondence is produced in this methodology.

4.2. Round Robin

In this calculation [12], the procedures are separated between all processors. Each procedure is allowed to the processor in a cooperative request. The procedure portion request is kept up locally autonomous of the designations from remote processors. Even though the outstanding task at hand disseminations between processors is equivalent to the activity handling time for various procedures isn't the equivalent. So a few hubs might be vigorously stacked anytime and others may stay inactive. This calculation is for the most part utilized in web servers where solicitations are comparable and are similarly conveyed[19].

Figure 2: Round Robin Processing Cloud Computing [20]

4.3. Central Queue

This algorithm[13] deals with the dynamic standard. Each new movement that arrives at the administrator of the line is embedded into the line. At the point when the line director gets a solicitation for a movement, the principal action is expelled from the line and sent to the requester. On the off chance that there is no prepared action in the line, the solicitation will be cushioned until there is another movement.

Figure 3: Work flow of central queue [14]

Yet, on the off chance that there are unanswered demands in the line, the main such solicitation will be pulled back from the line and new activity will be dispensed to it. At the point when a processor load falls beneath the edge, a solicitation for the new movement is sent to the focal burden chief by the nearby burden director. The supervisor will at that point react to the inquiry if the prepared movement is discovered generally until the new action shows up.

4.4. Connection Mechanism:

The heap adjusting calculation [15] can likewise be founded on the least connection system that is a piece of the calculation for

dynamic booking. To evaluate the heap, it should progressively tally the number of associations for every server. The heap balancer records every server's contact number. At the point when another connection is sent to it, the quantity of associations increments and diminishes when the association ends or a break happens.

5. COMPARISON OF DIFFERENT LOAD BALANCING TECHNIQUES:

Table 1 demonstrated the comparison of different burden adjusting procedures in centralized and distributed approaches, in centralized approach token Routing is the best approach for load balancing and in distributed approach central queuing approach is the best approach.

Table 1: Comparison of various load balancing techniques

Environment	Techniques	Nature	Resource Utilization	Advantages	Disadvantages
Distributed	Token Routing	dynamic	Maximum	-less communication delay -minimize the cost	-not use for process migration
	Round Robin	Static	Minimum	-No inter-process -communication Simple and easy to implement. -Minimum system overhead.	-Not useful if tasks have varying execution time. -Not useful if nodes have different capacities.
Centralized	Central Queuing	dynamic	Maximum	-Tasks assigned after execution -useful for heterogeneous nodes.	-No task migration in inter process -single point of failure
	Least Connection	dynamic	Minimum	-useful for nodes of similar capabilities	-not useful when task duration is different.

However, after comparison of both these two best approaches central queuing approach is the best approach for load adjusting in cloud computing.

6. CONCLUSION:

Different techniques are available in cloud computing that has different parameters. Some techniques are used in centralized and some are used in a distributed environment of cloud computing. In this research, after comparing both centralized and distributed techniques central queuing is the best load Balance technique that is based on different parameters such as performance, overload, throughput, response time, and migration time.

References

- [1] Fox, A., Griffith, R., Joseph, A., Katz, R., Konwinski, A., Lee, G., ... & Stoica, I. (2009). Above the clouds: A Berkeley view of cloud computing. Dept. Electrical Eng. and Comput. Sciences, University of California, Berkeley, Rep. UCB/EECS, 28(13), 2009.
- [2] Dikaiakos, M. D., Katsaros, D., Mehra, P., Pallis, G., & Vakali, A. (2009). Cloud computing: Distributed internet computing for IT and scientific research. IEEE Internet Computing, 13(5), 10-13.
- [3] Gong, C., Liu, J., Zhang, Q., Chen, H., & Gong, Z. (2010, September). The characteristics of cloud computing. In 2010 39th International Conference on Parallel Processing Workshops (pp. 275-279). IEEE.
- [4] Sidhu, A. K., & Kinger, S. (2013). Analysis of load balancing techniques in cloud computing. International Journal of computers & technology, 4(2), 737-741.
- [5] Mathew, T., Sekaran, K. C., & Jose, J. (2014, September). Study and analysis of various task scheduling algorithms in the cloud computing environment. In 2014 International Conference on Advances in Computing, Communications, and Informatics (ICACCI) (pp. 658-664). IEEE
- [6] Kushwaha, M., & Gupta, S. (2015). Various schemes of load balancing in distributed systems—A review. International Journal of Scientific Research Engineering & Technology (IJSRET), 4(7), 741-748.
- [7] Soundarabai, P. B., Sahai, R. K., Thriveni, J., Venugopal, K. R., & Patnaik, L. M. (2012). Comparative study on load balancing techniques in distributed systems. International Journal of Information Technology and Knowledge Management, 6(1), 53-60.
- [8] Rimal, B. P., Choi, E., & Lumb, I. (2009, August). A taxonomy and survey of cloud computing systems. In 2009 Fifth International Joint Conference on INC, IMS and IDC (pp. 44-51). Ieee.
- [9] Ray, S., & De Sarkar, A. (2012).

Execution analysis of load balancing algorithms in a cloud computing environment. *International Journal on Cloud Computing: Services and Architecture (IJCCSA)*, 2(5), 1-13.

[10] Chaczko, Z., Mahadevan, V., Aslanzadeh, S., & Mcdermid, C. (2011, September). Availability and load balancing in cloud computing. In *International Conference on Computer and Software Modeling*, Singapore (Vol. 14, pp. 134-140).

[11] Xu, Z., & Huang, R. (2009). Performance study of load balancing algorithms in distributed web server systems. *CS213 Parallel and Distributed Processing Project Report*, 1.

[12] Beltran, M., Guzman, A., & Bosque, J. L. (2011). Dealing with heterogeneity in clusters. In *proceeding of the Fifth International Symposium on Parallel and Distributed Computing*, ISPDC.

[13] Werstein, P., Situ, H., & Huang, Z. (2006, December). Load balancing in a cluster computer. In *2006 Seventh International Conference on Parallel and Distributed Computing, Applications and Technologies (PDCAT'06)* (pp. 569-577). IEEE.

[14] Kanagaraj, G., Shanmugasundaram, N., & Prakash, S. (2012). Adaptive load balancing algorithm using service queue. *Jurnal. ICCSIT*.

[15] Samal, P., & Mishra, P. (2013). Analysis of variants in Round Robin Algorithms for load balancing in Cloud Computing. *International Journal of computer science and Information Technologies*, 4(3), 416-419.

[16] Al Nuaimi, K., Mohamed, N., Al Nuaimi, M., & Al-Jaroodi, J. (2012, December). A survey of load balancing in cloud computing: Challenges and algorithms. In *2012 second symposium on network cloud computing and applications* (pp. 137-142). IEEE.

[17] Kaur, R., & Luthra, P. (2012, December). Load balancing in cloud computing. In *Proceedings of international conference on recent trends in information, telecommunication and computing*, ITC.

[18] Devi, D. C., & Uthariaraj, V. R. (2016). Load balancing in cloud computing environment using improved weighted round robin algorithm for nonpreemptive dependent tasks. *The scientific world journal*, 2016.

[19] Mahajan, K., Makroo, A., & Dahiya, D. (2013). Round robin with server affinity: a VM load balancing algorithm for cloud based infrastructure.

[20] Sameen, A., & Budhwant, D. K. (2015). Comparative analysis of round robin VM load balancing with modified round robin VM load balancing algorithms in cloud computing'. *International Journal of Engineering, Management & Sciences (IJEMS)*, 2(3), 61-67.