JOURNAL OF CRITICAL REVIEWS

ISSN- 2394-5125 VOL 09, ISSUE 04, 2022

EFFICIENT LOAD BALANCING TECHNIQUES FOR DYNAMIC WORKLOAD MANAGEMENT IN CLOUD COMPUTING ENVIRONMENTS

Krunal Chandrashekhar Pawar[1] , Rishi Kushwaha[2]

[1]Research Scholar, Department of Computer Science & Engineering ,Sri Satya Sai University of Technology and Medical Sciences, Sehore Bhopal-Indore Road, Madhya Pradesh, India

[2]Research Guide, Department of Computer Science & Engineering ,Sri Satya Sai University of Technology and Medical Sciences, Sehore Bhopal-Indore Road, Madhya Pradesh, India

Abstract

The term "cloud computing" (CC) is used to talk about a part of the computing business that is growing quickly. Cloud computing (CC) makes services like IaaS, SaaS, and PaaS available. These services can be changed to fit the needs of both people and businesses. With the virtualization of the internet, cloud computing has made it easier for developers to make apps and for users to get them. In CC, we have cloud service providers who can handle big computing structures that are completely based on what their users need. Consumers are able to use the services of these providers. But CC is having trouble balancing the load, which can hurt service and speed if the system is overloaded. The final goal of the providers is to make a CC-based load-balancing algorithm that works well and is efficient. VMs need to be made in CC because of several dynamic methods, such as virtualization and scalability. The number of web services and data traffic is growing at an exponential rate. This makes load balancing a big problem in CC and makes job scheduling more important in CC. Because CC is getting better, there is a bigger need for infrastructure and resources to grow. The load-balancing strategy makes sure that the resources are used in the best way possible when cloud customers get services. Load balancing puts subscribers in order of importance based on a good plan. This study wants to add to the research by coming up with a way to get load balancing in CC. The proposed way will make cloud-based apps work better in general.

Keywords—Some of the things that will be talked about are cloud computing (CC), load balancing, methods for machine learning, and organizing tasks.

ISSN- 2394-5125 VOL 09, ISSUE 04, 2022

Introduction

Cloud computing is a new field of study in the IT field as a whole. Cloud computing has made load balance a big problem. Spreading the dynamic job across multiple nodes is a key way to keep any one node from getting too busy. Load-balancing software can be used to make sure that high-traffic websites, apps, and platforms keep serving their users without stopping. Load balancing methods help get the most out of the system's resources, which makes the system as a whole work better. The main goal of load balancing is to use resources as little as possible. This cuts down on energy use and carbon pollution, which are both needed for cloud computing. In order for load balancing in the cloud to be carbon-neutral and save energy, this opens the door for new measures, energy use, and carbon emissions.

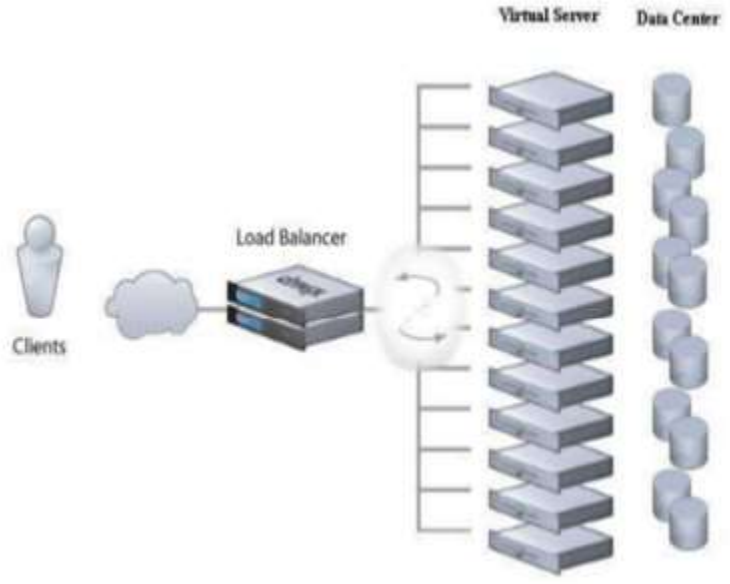

Fig 1.Load Balancing Mechanism

A. The main goals of load balancing

1) Make a method that works even if mistakes are made.

2) Make sure the system stays in balance.

3)Boost output and efficiency to have the most influence.

4) Spending less time getting things done and waiting around

5) To improve the experience of users in general. Here, we talk about two different ways that the new load-balancing method can be used. This was done so that the force would be spread out more evenly.

6) Static load balancing: For this method, you need to know ahead of time what the system can do. So, the current state of the system has nothing to do with how the work should be split.The environment doesn't change the running process to balance the workload, and the performance of the processors is known before execution starts. This method works best in situations where the system is stable.

7) Dynamic Load Balancing: This method decides how to divide the load based on the current state of the system, instead of using information about system resources that was already known. This choice about how to split up the work can be made even if no one knows what system resources are available. This is the best choice for a machine with many different parts. During the dynamic load balancing process, changes are made to the load that is being moved. This method is much better than the static approach and makes a huge difference in how well it works.

Related work

Load balancing is an important part of cloud computing because it makes sure that resources are used as efficiently as possible and that services are delivered on time. This piece looks at a few of the many studies that have been done on on-demand load balancing in cloud computing systems and talks about them.

Manjula Shanbhog et al. (2020) In a study paper from 2020, the load-balancing methods that can be used in cloud computing systems are described in depth. The writers look at the problems with load balancing and give several ways to solve them.

Yelchuri Venkata et al. (2021) In cloud computing settings, you should offer a load-balancing method that uses fuzzy logic to maximize efficiency without increasing reaction times.

Shahbaz Afzal et al. (2019) There are three main types of cloud computing load balancing: static, dynamic, and mixed. They compare and contrast the different ways that load balancing can be done in the cloud and suggest a taxonomy based on levels of order.

Febina K S et al. (2023) Look at the different ways that load balancing can be done in cloud computing settings. Pay special attention to the problems and challenges that come up when putting these plans into action. The writers look at different ways to study and say what they think the direction of future studies in this field should be.

Narayan A. Joshi et al. (2022) describes a load-balancing approach for cloud computing that uses both static and dynamic algorithms to get the most out of the resources available. The author looks at the proposed method and shows how well it works to improve the performance of cloud computing environments.

Overall, the works we've talked about here teach us useful things about load balancing in cloud computing environments and show how important it is to use resources and offer services efficiently. They also show that more study and development is needed in this area to find the best load-balancing solutions for the new problems.

Methodology

A modernized genetic method has been used to count all of the virtual machines that are available. When a new job comes in, the system checks immediately to see if there are any free virtual machines that can be used. If a virtual machine that can do the job is available, the task is given to it. If there isn't a free virtual machine, the work will be done on the machine whose job is expected to finish first. So, all VMs are used well, and neither are they underused nor are they used too much. So, the results of this method are better than the results of any of the other options that are now available, both in terms of cost and energy economy.

The Enhanced Genetic Algorithm is made up of a number of steps.

1. **What to Do First:**

   Most of the time, the first set of possible solutions in the search space are chosen at random. But it is easy to add in domain-specific knowledge or other data.

2. **Second, an answer:**

   The fitness values of the potential solutions are checked as soon as a population starts or a new population is born, whichever comes first.

3. **Choose Your Own #3:**

   Because of this, the survival-of-the-fittest process is used on the candidate solutions, with more copies of the selection going to the solutions that were shown to have better fitness values. The goal of selection theory is to change a bad choice into a good one. To do this, many different methods have been tried, such as random universal selection. Also explained in depth below are the ranking selection, the roulette-wheel selection, and a few less common selection methods, such as the tournament selection.

4. **Step 4: Mixing back together**

   Recombination is the process of making new solutions (offspring) by combining parts of two or more current solutions. There are different ways to do this, and a well-thought-out recombination process is important for getting good overall results. When a couple's genes are mixed, the child will have traits from both parents, but will be a unique mix of the two.

During the installation part, the following steps could be taken: In the first step, a small number of virtual machines are added to the cloud network. The process is now under way. phase 2: If there is a problem with the network, the process moves on to the next step after figuring out which virtual machine is best for running the cloudlet. The third step is to run the new genetic algorithm, which will move the work to a different virtual machine if something goes wrong.

Load balancing

One of the main goals of LB is to make sure that no single cloud node is either too busy or not busy enough. Load balancing (LB) is the process of spreading a workload across many network connections on multiple nodes or groups of nodes to make the best use of resources and reduce

ISSN- 2394-5125 VOL 09, ISSUE 04, 2022

latency. Less time is wasted waiting for the tool, and resources aren't used twice for no reason. By spreading requests across a number of servers, this method makes it so you don't have to wait for data to be handled. Engineers use a method called load balancing (LB) to get the most out of a system. Figure 2 shows the LB at CC for your ease of use.

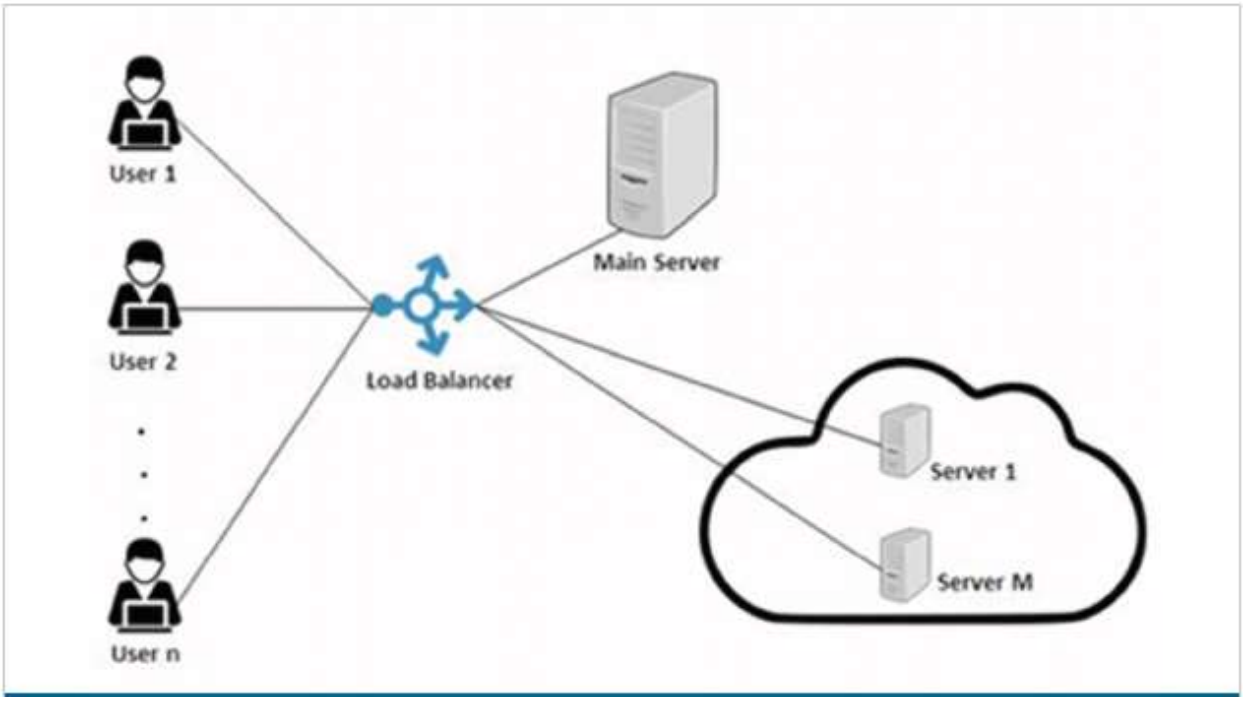

FIGURE 2 .The ways that load balancing is done in the cloud

LB gives a methodical way to divide duties fairly based on the available resources. The goal is to keep services running smoothly, which means making good use of resources in the event of a disaster. In order to reach this goal, the device instance will be set up and then removed. LB also wants to make devices more efficient at a lower cost by cutting down on the time it takes to do chores and making better use of resources.

AN INTENDED SCHEDULE-FORMULA

We came up with a new way to handle load balancing to make cloud-based apps work better as a whole. Here is a full explanation of the pseudo-code method that the virtual machine uses to divide up work.

Proposed Algorithm:

1. Sorting of tasks based on the deadline ascendingly
Tasks if have the same deadline, then Pick the task with the earliest arrival time
Else
Priorities based on (deadline)
For each virtual machine
2. *Utilization computation*
3. *Sort virtual machines according to their Utilization*
Repeat
If a virtual machine is available and a task is allocated to a heavy virtual machine.
then
Migrate task to the less utilized virtual machine
Else
Start scheduling
Till all tasks allocated to a virtual machine
End

Our proposed way helps keep load balancing based on priorities.

Its completion time, which is recorded in milliseconds, is the time between when it finishes and when it is put into the queue.

If you know how long it will take you to finish the job, you can decide whether it is hard, medium, or easy.

The time limit for finishing a job is also given in milliseconds, and it shows how much time has passed since the job was supposed to be done.

Conclusion

People often use the word "cloud computing" to talk about a system that lets several users share computing resources over the internet. Even so, cloud computing has a lot of problems that haven't been fixed yet.

One of the biggest problems with cloud computers is load balancing. In this study, both static algorithms and dynamic algorithms are looked at. In fact, the parts of clouds can be very different from one to the next. Static methods make modeling and watching the environment easier, but they don't take into account the fact that clouds are always changing. Dynamic

methods to load balancing have shown to be very adaptable to the different ways that clouds are set up, even though they are hard to model.

In this post, we'll talk about a wide range of load-balancing methods and give a review of their pros, cons, and real-world uses.

Reference

1. Manjula Shanbhog,Kalpna. LOad balancing research paper,July 2020,DOI:10.35940/ijrte.B1176.0782S619.
2. Yelchuri Venkata Sai Harsha, Nagaraj G Cholli. Load Balancing in Cloud Computing. International Journal of Advanced Research in Computer and Communication Engineering Vol. 10, Issue 5, May 2021 DOI 10.17148/IJARCCE.2021.105146
3. Shahbaz Afzal and G. Kavitha. Load balancing in cloud computing – A hierarchical taxonomical classification. Afzal and Kavitha Journal of Cloud Computing: Advances, Systems and Applications (2019) 8:22 https://doi.org/10.1186/s13677-019-0146-7.
4. Febina K S, Tanuja Hemchand, Shamimunnisabi , Roopa P, Aishwarya P. A Survey on Issues and Challenges of Load Balancing Techniques in Cloud Computing Environment DOI Link: https://doi.org/10.22214/ijraset.2023.48878.
5. Narayan A. Joshi. Technique for Balanced Load Balancing in Cloud Computing Environment. Digital Object Identifier (DOI) : 10.14569/IJACSA.2022.0130316 Article Published in International Journal of Advanced Computer Science and Applications(IJACSA), Volume 13 Issue 3, 2022.
6. Abhishek Kumar Tiwari,Sreeja Nair. A Review of Load Balancing Technique of Cloud Computing Using Swarm Intelligence. / (IJCSIT) International Journal of Computer Science and Information Technologies, Vol. 7 (2) , 2016, 741-744
7. Mahalingam, Nandhalakshmi Nithya, “Efficient Load Balancing in Cloud Computing Using Weighted Throttled Algorithm”, International Journal of Innovative Research in Computer and Communication Engineering, 2015,vol.3, 5409 – 5415. [12] Keke Gai, Meikang Qiu,
8. Muhammad Asim Shahid, Noman Islam, Muhammad Mansoor Alam, Mazliham Mohd Su’ud, Shahrulniza Musa,”A Comprehensive Study of Load Balancing Approaches in the

Cloud Computing Environment and a Novel Fault Tolerance Approach " . DOI: 10.1109/ACCESS.2020.3009184

9. Hui Zhao, "Cost-Aware Multimedia Data Allocation for Heterogeneous Memory Using Genetic Algorithm in Cloud Computing", IEEE Transactions on Cloud Computing, 2015. [13]
10. Mr. Mayur S. Pilavare, Mr. Amish Desai, "A Novel Approach Towards Improving Performance of Load Balancing Using Genetic Algorithm in Cloud Computing", IEEE Sponsored 2nd International Conference on Innovations in Information Embedded and Communication Systems, ICIIECS'15, 2015.
11. Tang L, Li Z, Ren P, Pan J, Lu Z, Su J, Meng Z (2017) Online and offline based load balance algorithm in cloud computing. Knowl-Based Syst 138: 91–104.
12. Vanitha M, Marikkannu P (2017) Effective resource utilization in cloud environment through a dynamic well-organized load balancing algorithm for virtual machines. Comp Elec Eng 57:199–208
13. Singh AN, Prakash S (2018) WAMLB: weighted active monitoring load balancing in cloud computing. In: Big data analytics. Springer, Singapore, pp 677–685.
14. R. S. Sajjan, "Load Balancing and its Algorithms in Cloud Computing : A Survey International Journal of Computer Sciences and Engineering Open Access Load Balancing and its Algorithms in Cloud Computing : A Survey," no. January 2017, 2019.
15. Devaraj, A., Saviour, F., et al.: Hybridization of firefly and improved multi-objective particle swarm optimization algorithm for energy efficient load balancing in cloud computing environments. J. Parallel Distrib. Comput. 142, 36–45 (2020).
16. Shahdi-Pashaki, S., Teymourian, E., Tavakkoli-Moghaddam, R.: New approach based on group technology for the consolidation problem in cloud computing-mathematical model and genetic algorithm. Comput. Appl. Math. 37(1), 693–718 (2016). https://doi.org/10.1007/s40314-016-0362-4.
17. Ghumman, N.S., Kaur, R.: Dynamic combination of improved max-min and ant colony algorithm for load balancing in cloud system. In: 2015 6th International Conference on Computing, Communication and Networking Technologies (ICCCNT). IEEE (2015).
18. Hashem, W., Nashaat, H., Rizk, R.: Honey bee based load balancing in cloud computing. KSII Trans. Internet Inf. Syst. 11(12), 5694–5711 (2017).

19. Mukundha, C., Venkatesh, N., Akshay, K.: A comprehensive study report on load balancing techniques in cloud computing. Int. J. Eng. Res. Dev. 13(9), 35–42 (2017).
20. S. Srivastava and R. Kumar, "Indirect method to measure software quality using CK-OO suite," 2013 International Conference on Intelligent Systems and Signal Processing (ISSP), 2013, pp. 47-51, doi: 10.1109/ISSP.2013.6526872.
21. Ram Kumar, Gunja Varshney , Tourism Crisis Evaluation Using Fuzzy Artificial Neural network, International Journal of Soft Computing and Engineering (IJSCE) ISSN: 2231-2307, Volume-1, Issue-NCAI2011, June 2011
22. Ram Kumar, Jasvinder Pal Singh, Gaurav Srivastava, "A Survey Paper on Altered Fingerprint Identification & Classification" International Journal of Electronics Communication and Computer Engineering Volume 3, Issue 5, ISSN (Online): 2249–071X, ISSN (Print): 2278– 4209
23. Kumar, R., Singh, J.P., Srivastava, G. (2014). Altered Fingerprint Identification and Classification Using SP Detection and Fuzzy Classification. In: , et al. Proceedings of the Second International Conference on Soft Computing for Problem Solving (SocProS 2012), December 28-30, 2012. Advances in Intelligent Systems and Computing, vol 236. Springer, New Delhi. https://doi.org/10.1007/978-81-322-1602-5_139
24. Gite S.N, Dharmadhikari D.D, Ram Kumar," Educational Decision Making Based On GIS" International Journal of Recent Technology and Engineering (IJRTE) ISSN: 2277-3878, Volume-1, Issue-1, April 2012.
25. Ram Kumar, Sarvesh Kumar, Kolte V. S.," A Model for Intrusion Detection Based on Undefined Distance", International Journal of Soft Computing and Engineering (IJSCE) ISSN: 2231-2307, Volume-1 Issue-5, November 2011
26. Vibhor Mahajan, Ashutosh Dwivedi, Sairaj Kulkarni,Md Abdullah Ali, Ram Kumar Solanki," Face Mask Detection Using Machine Learning", International Research Journal of Modernization in Engineering Technology and Science, Volume:04/Issue:05/May-2022
27. Kumar, Ram and Sonaje, Vaibhav P and Jadhav, Vandana and Kolpyakwar, Anirudha Anil and Ranjan, Mritunjay K and Solunke, Hiralal and Ghonge, Mangesh and Ghonge, Mangesh, Internet Of Things Security For Industrial Applications Using Computational

Intelligence (August 11, 2022). Available at SSRN: https://ssrn.com/abstract=4187998 or http://dx.doi.org/10.2139/ssrn.4187998

28. Kumar, Ram and Aher, Pushpalata and Zope, Sharmila and Patil, Nisha and Taskar, Avinash and Kale, Sunil M and Gadekar, Amit R, Intelligent Chat-Bot Using AI for Medical Care (August 11, 2022). Available at SSRN: https://ssrn.com/abstract=4187948 or http://dx.doi.org/10.2139/ssrn.4187948
29. Kumar, Ram and Patil, Manoj, Improved the Image Enhancement Using Filtering and Wavelet Transformation Methodologies (July 22, 2022). Available at SSRN: https://ssrn.com/abstract=4182372
30. Ram Kumar, Manoj Eknath Patil ," Improved the Image Enhancement Using Filtering and Wavelet Transformation Methodologies", Turkish Journal of Computer and Mathematics Education ,Vol.13 No.3(2022), 987-993.
31. Ram Kumar, Jasvinder Pal Singh, Gaurav Srivastava, "A Survey Paper on Altered Fingerprint Identification & Classification" International Journal of Electronics Communication and Computer Engineering ,Volume 3, Issue 5, ISSN (Online): 2249–071X, ISSN (Print): 2278–4209.
32. Muhammad Asim Shahid,Md parves AlamMazliham Bin Mohd Su'ud. Performance Evaluation of Load-Balancing Algorithms with Different Service Broker Policies for Cloud Computing January 2023Applied Sciences 13(3):1586 DOI:10.3390/app13031586 LicenseCC BY 4.0