

Performance of Several Load Balancing Techniques and Algorithms in Cloud Environment

Sujan Narahari, Jaswanth Boyapaty, Tushar Kumar Pandey, Deeban Chakravarthy

Abstract: Cloud computing is an upcoming technology which is used mainly for parallel computing and large scale distributed computing. This concept provides various techniques like shared resources, software packages, information and other resources as per requirements of clients at specific time. The concept resource sharing is a technique that distributes dynamic local workload across all the nodes. Cloud computing is rapidly growing technology and more users are attracted towards utility computing, better service needs to be provided. For better management of the workload at all the nodes, load balancing techniques are used, so the load balancing plays a crucial role in cloud computing. Here in this paper we are discussed various load balancing techniques which will solve the major issues in cloud.

Index Terms: Cloud Computing, Load Balancing, Load Balancing Algorithms, Resource Allocation

I. INTRODUCTION

Cloud computing is a method for delivering information technology (IT) services in which resources are retrieved from the Internet through web-based tools and applications [1]. Cloud Computing is an online platform which allows us to access the application as utilities with the help of internet. It is basically divided into four models like Public, Private, Hybrid and Community Clouds. Each cloud type has its own advantages and disadvantages. Based on Services provided by the cloud they are further divided into 3 types: Infrastructure-as-a-Service (IaaS), Platform-as-a-Service (PaaS), Software-as-a-Service (SaaS).

In an infrastructure-as-a-service (IaaS) cloud, the cloud provider divides the physical machines dynamically into various segments using various technologies, so that the Virtual machine (VM) accepts various requests coming in from the customers.

The basic idea in load balancer is to make sure that the work is disturbed equally across all nodes in order to decrease the response time and to make better use of available resources. This way all the nodes are equally worked and no node is overworked nor less worked. A load balancing algorithm should always be dynamic in nature

Revised Manuscript Received on May 05, 2019.

Sujan Narahari, Computer Science & Engineering, SRM University, Chennai

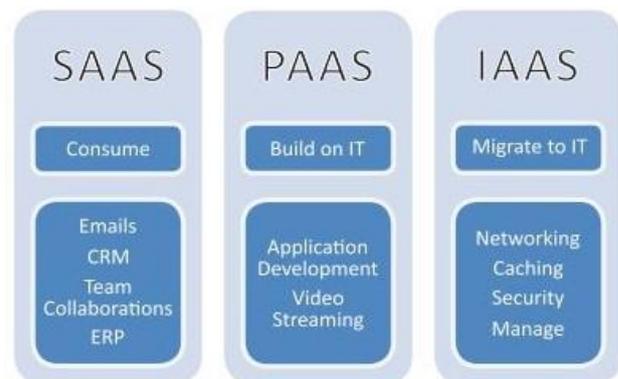
Jaswanth Boyapaty, Computer Science & Engineering, SRM University, Chennai

Tushar Kumar Pandey, Computer Science & Engineering, SRM University, Chennai

Deeban chakravarthy, Computer Science & Engineering, SRM University, Chennai

because these algorithms only consider the present state without checking the previous state or how the system behaves.

There are certain things to consider while developing that algorithm like stability of the system, interactions between the nodes, estimation of the load and nature of work to be transferred. Load balancing is the concept of enhancing the performance of the current system by moving the workload among the available processors. The load of a current machine means total processing time it require to process all the tasks that are assigned to the machine. This load balancing concept increase the work performance of the cloud service provider. Balancing the load of virtual machines uniformly means that anyone of the available machine is not idle or partially loaded while others are heavily loaded.



II. CLOUD PROVIDER

Cloud providers are responsible for building the cloud. They can offer different types of clouds such as Public or Private or Hybrid cloud. Based on requirement and necessities public clouds can be used by individual or organisation [6]. IBM's Blue Cloud, Sun Cloud, Google AppEngine are some of the public clouds.

Private clouds are the ones which owned by enterprises for the internal uses like store and manage of their personal data. These types of cloud offer high level of security for protecting the data. OpenStack, VM ware.

Hybrid is the combination of private and public clouds. It has the features of both public and private clouds and hence it will be more useful in some cases.



III. LOAD BALANCING

Load balancing is the important concept in cloud environment and it is used to effectively maintain speed and consistency of the cloud server. In cloud environment few nodes will be overloaded and other are slightly loaded.

Load balancing is mainly used to remove server clusters. The basic concept is to distribute work load across all the variable nodes in the cloud. It mainly aims to optimize resource use, minimum response time, maximize the rate of production and avoid overload of any single resource. This concept increases reliability and availability through redundancy. It mainly accounts two major task i.e. resource allocation and other one is task scheduling in disturbed environment. CloudSim [3] provides an extensible simulation framework that enables seamless modelling and simulation of performance. By using CloudSim, developers can focus on specific systems design issues that they want to investigate, without getting concerned about details related to cloud-based infrastructures and services.

It is composed of four entities i.e. Datacentres, Virtual Machines, Hosts and Applications as well as System Software. Datacentres are responsible for providing Infrastructure level services to cloud users. Hosts are pre-configured physical servers in cloud environment. Characteristics of load balancing are [4][5]

- To achieve minimal response time.
- To attain high user satisfaction.
- Maximize overall system performance.
- To share the load equally across the nodes.

IV. LOAD BALANCING ALGORITHMS

These algorithms are proposed mainly in attempt to ensure good overall performance of the system by distributing total load across the nodes by transferring the workload from heavily loaded nodes to lightly loaded nodes. There are mainly two types of load balancing. They are

- Static Load Balancing.
- Dynamic Load Balancing.

Static Load Balancing Algorithms

Static Algorithms are designed in such a way that they do not consider the current state of nodes. These algorithms only consider the average behaviour of the system ignoring the current state of the system. These are every simpler as they don't need to maintain and process system state information.

Round Robin Algorithm

Round robin comes under the category static load balancing algorithms in which the previous states are not taken into consideration, and it uses round robin method for allocation of the tasks. The advantage of this method is it does not need internal process communication. In this round robin method it first selects a node randomly and then allocation of work to all other nodes is evenly done. In this some tasks are heavily loaded due to no prior information about the processors, In this round robin algorithm each node is predefined with specific weight, based on the weight of the node they will get the requests.

Opportunistic Load Balancing Algorithm

This algorithm do not consider the present workload of the virtual machine. In this algorithm every node is kept busy by assigning some work. Every task will be processed in slowly, and the current execution time for a node is not calculated.

Min-Min Load Balancing Algorithm

This algorithm come under the static load balancing technique, in this the parameters connected to the job are recognised. In this some complex time calculations are identified by the cloud manager. To complete the task in time the job with minimum execution in time are executed first. The task with maximum execution time will perform poor than the minimum execution time. The disadvantage of this algorithm it lead to long time starvation. It allocates the tasks by considering the completion time, the task with minimum completion time is allocated the job first.

Max-Min Load Balancing Algorithm

Max-Min is a static Algorithm in which the maximum jobs are selected when the minimum time jobs are completed. The tasks are assigned to the processor only when the minimum time jobs are done and then the execution time of all the processors are noted. As this Algorithm is static the time taken to complete the process is fixed and predefined. Presently an advanced Max-Min version is underway.

Dynamic Load Balancing Algorithms

Dynamic Load Balancing considers the current state of the system. These algorithms will search for lightest node in the whole system and prefers this node for balancing the load.

Ant Colony

In Ant Colony, we can solve problems and finding good paths through graphs. It is used to see how different ants seek their food in an optimal way. The main idea has to solve a large class of numerical problems so that several problems which are emerging can be done by drawing on various aspects of the behaviour of ants. In a broader way, ACO performs a model-based search and share some similarities with Estimation of Distribution.

This algorithm is first proposed by Marco Dorigo in 1991. This algorithms can be applied to various problems, varying from quadratic assignment to protein folding and a lot of derived methods have been adapted to dynamic problems in real life. Some of the major applications are traveling sales man, scheduling, Network model problem and vehicle routing to name a few. Some of the major advantages of this algorithm are it gives a positive feedback and distributed computation as it avoids premature convergence.

Biased Random Sampling load balancing Algorithm

Biased Random Sampling Load Balancing Algorithm is a Dynamic Algorithm, it is represented in virtual graph. The basic notations in this algorithm is vertex and in degree



where each server is donated as a vertex and free available resources are donated by in degree. Based on in degree the load balancer allocates the jobs to particular nodes. The job is allocated to a node when there is at least one in degree. As soon as the job is allocates to the node then the in degree is decrement by one.

This technique is mainly used in addition and deletion of the processes. The processes are based on a threshold value. The length of traversal is known as walk length. When the current node is in traversal the neighbouring node is on the waiting list. The basic process of this is Algorithm is the load balancer selects a random node and then it compares that to the current node and if the current walk length is equal to or greater than the threshold value then that job is executed else the walk length is incremented and another node is selected randomly.

V. CONCLUSION

Load balancer is very effective and useful in cloud computing as it is used to effective distribute the load work among all nodes and share the equally across all nodes, so that no node is overworked or under worked. The load balancer is more effective if it's dynamic algorithm as it consider the current state and does the work accordingly whereas static algorithms consider the previous state and aren't much effective when compared to dynamic algorithms.

REFERENCES

1. Basic Definition of Cloud
<https://www.investopedia.com/terms/c/cloud-computing.asp>
2. A. Khiyati, M. Zbakh, H. El Bakkali, D. El Kettani "Load Balancing Cloud Computing: State Of Art", IEEE, 2012.
3. Bhathiya Wickremasinghe, Roderigo N. Calherios "Cloud Analyst: A Cloud-Sim-Based Visual Modeler For Analyzing Cloud Computing Environments And Applications". Proc Of IEEE International Conference On Advance Information Networking And Applications ,2010.
4. L.Giuseppe Aceto, Alessio Botta, Walter de Donato, Antonio Pescapè, "Cloud monitoring: A survey", SciVerse ScienceDirect's Computer Networks 57, PP-2093–2115, ASOC 1894 1–12, © 2013 Elsevier B.V.
5. Chun-Cheng Lin, Hui-Hsin Chin, Der-Jiunn Deng, "Dynamic Multiservice Load Balancing in Cloud-Based Multimedia System", 1932-8184/\$31.00 ©2013 IEEE, DOI 10.1109/JSYST.2013.2256320.
6. L.D. Dhinesh Babua, P. Venkata Krishna, "Honey bee behavior inspired load balancing of tasks in cloud computing environments", SciVerse ScienceDirect's Applied Soft Computing, ASOC 1894 1–12, © 2013 Elsevier B.V.
7. B.Mondal, K. Dasgupta, and P. Dutta, "Load balancing in cloud computing using stochastic hill climbing-a soft computing approach," Procedia Technology, vol. 4, pp. 783–789, 2012.
8. Armbrust, M.; Fox, A.; Griffith, R.; Joseph, A.D.; Katz, R.; Konwinski, A.; Lee, G.; Patterson, D.; Rabkin, A.; Stoica, I.; et al. A view of cloud computing. Commun. ACM 2010, 53, 50–58.
9. Mishra, R.; Jaiswal, A. Ant colony optimization: A solution of load balancing in cloud. Int. J. Web Semant. Technol. 2012, 3, 33–50.
10. Ren, X.; Lin, R.; Zou, H. A dynamic load balancing strategy for cloud computing platform based on exponential smoothing forecast. In Proceedings of the 2011 IEEE International Conference on Cloud Computing and Intelligence Systems (CCIS), Beijing, China, 15–17 September 2011; pp. 220–224.
11. Abhijit A Rajguru, S.S. Apte, "A Comparative Performance Analysis of Load Balancing Algorithms In Distributed Systems Using Qualitative Parameters", International Journal of Recent Technology and Engineering, Vol. 1, Issue 3, August 2012
12. Parin. V. Patel, Hitesh. D. Patel, Pinal. J. Patel, "A Survey on Load Balancing in Cloud Computing" IJERT, Vol. 1, Issue 9, November 2012