

A Circumscribed Research of Load Balancing Techniques in Cloud Computing

K. Samunnisa, G. Sunil Vijaya Kumar, K. Madhavi

Abstract: Cloud Computing (CC) is a very big field in the computer industry. It has systems that are connected through communication networks such as Internet. The LBA is an important cloud computing area to prevent overloaded workload and provide the same and valuable service. The various algorithms are used to solve the load balancing complexity. In this paper, we explain various standard load balancing techniques in different algorithms and challenges in assigning tasks to Cloud Computing. The load balancing algorithm can be used to better utilize and better understand user needs.

Keywords: Load Balancing in Cloud Computing, Load Balancing Algorithms and techniques, Challenges.

I. INTRODUCTION

In the cloud, the load balancing refers to the allocation of client requests across multiple application servers running in the cloud background. As with other types of load balancing, load balancing in the cloud helps us maximize application performance and reliability [1]. The advantages of distributing tasks from traditional local sources are (usually) reducing costs and simplifying or reducing demand to meet demand [2, 3].

Conventional load balancing solutions are based on data center hardware and employ sophisticated IT staff to install, customize and retain the system. Only large organizations and companies with high IT budgets can achieve better results and greater reliability. Cloud computing is another major disadvantage in hardware solutions: they do not support cloud-based load balancing, since cloud-based infrastructure providers generally do not allow their background to own hardware [4].

Fortunately, software load balancers can provide hardware management and consistency, covering only a small part of the cost. Since they work with basic hardware, they are even available to small businesses. And they are ideally suited to balance the load in the cloud as they can have control in the cloud similar to any other application.

Cloud computing can be a widespread and expandable system in any distributed resource unit in the network. All system resources need to work together to meet consumer demand.

This paper organized as Chapter-2 as Related work, which describes the overview of load balancing, types of load balancing and its importance. Chapter-3 Describes the review of Load balancing algorithms and its classifications. Chapter-4 describes the Challenges in Cloud Load balancing and finally Chapter-5 tells about the conclusion of the paper.

II. RELATED WORK

2.1 Overview of Load Balancing:

With the load balancer you can split the workload and balance it between two or more servers in the cloud. As a result, you can configure your infrastructure to maximize activity, optimize resource allocation, and provide a smallest amount of response time [5].

Using a load balancer is suggested in all cases, whether you require one or more of the following:

By using load balancer in all situations [6], we can avail the following one or more requirements

- Continuity of the guaranteed service
- Manages a lot of traffic
- Prepare sudden peaks in the application

To balance workloads, you need at least two identical cloud servers in both configuration and data mode.

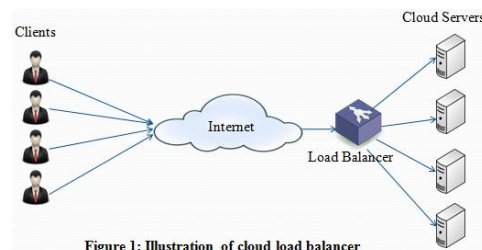


Figure 1: Illustration of cloud load balancer

Using this type of configuration, the load balancer accepts network requests and sorts them between two or more servers in the cloud.

2.2 Types of Load Balancing:

The load balance in the cloud is a process that distributes workloads and calculates resources in a cloud computing environment. With load balancing, organizations can manage application or workload requests by allocating resources to multiple computers, networks, or servers. Load balancing types in Cloud environment with description is explained in table 1.

Table 1: Load balancing types in Cloud environment with description

Types	Description
Software-defined networking (SDN)	This allows you to organize multiple tax balancing. It also helps the network run as a virtual machine and a storage version.

Revised Manuscript Received on April 12, 2019.

K. Samunnisa, Research Scholar, JNTUA, Anantapuramu, A.P, India.
Dr.G. Sunil Vijaya Kumar, Professor of CSE Dept, GPREC, Kurnool, A.P, India
Dr. K. Madhavi, Associate Professor of CSE Dept, JNTUA, A.P. India

User Datagram Protocol (UDP)	UDP load balancing is often used in live and online games when speed is high and there is little need to correct errors. UDP has a low latency because it does not have a lot of time-consuming state control elements.
Transmission control protocol (TCP)	It provides reliable and verified packet error feeds to IP addresses that otherwise could be easily lost or damaged.
Server Load Balancing (SLB)	Provides network services and content delivery using a range of load balancing algorithms. Priorities determine the responses to customer requests on the network. Server load balancing distributes client traffic to servers to provide efficient and consistent applications.
Virtual	<ul style="list-style-type: none"> Manage the physical loading software in the virtual machine It involves limited scalability and automation as well as a lack of central administration.
Load Balancer as a Service (LBaaS)	Load balancing as a service (LBaaS) utilizes load balancing technology to meet the demands of fast traffic and applications from organizations that implement private cloud infrastructure.

2.3 Significance of Load Balancing

Load balancing is the distribution of network traffic by server group. The load balancer is the server that performs this operation. Load balancing serves as a hardware and software performance solution. Significance of Load balancing is explained in table 2.

Table 2: Significance of Load balancing

Significance of Load balancing	Description
Better Performance	Load balancing methods are smaller and easier to implement than their counterparts. Organizations can work their customers' applications more quickly and deliver better performance at relatively low cost.
Maintain Website Traffic	Cloud Balancing provides scalability to manage your site traffic. With efficient load balancing, you can easily manage user traffic at a high level with the presence of servers and network devices.
Handle Sudden Traffic Burst	Load balancers have this ability to handle any sudden traffic received at a specified time. Load balancers distribute the entire site load on different servers to get optimal results with minimal response time.
Flexibility	If the workload is distributed across servers or network devices, even if one node fails, the load can be replaced by another node. It shows scalability, flexibility, and the ability to handle traffic.

III. COMPARE OF LOAD BALANCING ALGORITHMS& RESULTS

numerous sorts of load balancing algorithms are projected, which may be analyzed as follows: [7] you in all likelihood remember the fact that the burden balancer passes purchaser get entry to requests to the servers and cleverly determines which server is most suitable for responding to each request. but how does the weight balancer make this selection? This assessment describes the techniques used. severa load balancing techniques and algorithms may be used at the equal time as loading the patron get proper of access to request to severa servers [8]. the choice of techniques frequently is based totally upon at the beauty of provider, and moreover the u . s .a .of servers and networks on the time of purchaser request. In element, the strategies described underneath may be used collectively to determine the most remarkable server for dealing with new requests. The contemporary degree of load balancing software program software softwaresoftware program software software application is regularly determined through the technique used [9]. If the weight is low, one of the simplest distribution techniques is enough. however, throughout excessive load time, the maximum advanced strategies are used to make certain a steady allocation of precise sooner or later of the community and carrier pressure. desk three compares with numerous load balancing algorithms.

Table 3: Comparison of various Load Balancing Algorithms

LBA	Objective	Merits	Demerits	Nature	Future Work
BRS	To gets LB over all nodes using RS of the system domain	To effective, reliable and scalable load balancing strategy	No capacity is handle to large load	DLB	To apply on heterogeneous system
VD	To managing hierarchal and multi-dimensional resource	To use global state information	Removes overloaded on server switch and storage	DLB	Incorporate proactive migrations based on Statistical model and historical trends
CAB	Zero-downtime reallocation and used adaptive live migration for each equilibrium solution	Very fast and decreases migration time	Let CAB be have a sufficient memory with every physical host and the migration is still slow for Open VZ VM	DLB	Xen virtual technology for live migration and to support affinity and anti affinity
CARTON	To solve optimizing problem and considerable cost for allocation of job server	For implementation is to easy and simple, computation and COM overhead is very low	COM between small UDP packets and it has no extra capacity for loss of COM packet	DLB	To use COM between TCP packets and uses control with infinite bandwidth
CLBVM	To make load balancing decisions use global state information	To improve overall performance of balances the load	It is not consider fault tolerance and load of network is constatnt	DLB	Decentralized load balanced
LBVS	To achieve replica LB module	To provides large scale net data storage and storage as a service	Use writing balancing algorithm	DLB	To apply on virtual storage for large scale
ACCLB	To balance the load for open cloud computing using ant colony and complex network	It provides long distance facility to allow shared computations and storage optimal job scheduling	It takes more time	DLB	To use the probability function for good PR to the ant colony algorithm

3.1 Load Balancer strategies:

a) Weighted round Robin

it is based totally on spherical Robin's easy load balancing technique. in the evaluated version, every corporation server receives a static virtual weight. pinnacle rated servers get preserve of greater requests sent.

b) Agent based completely Adaptive Load Balancing

eachorganisation server has an agent that reviews the current-day load to the weight balancer. This real-time statistics is used to decide which server is maximum suitable for dealing with an software software program. it's miles



achieved in combination with particular techniques, along with spherical weights and the least weighed compound.

c) *Chained Failover (everyday Weighted)*

This approach has configured a pre-described collection of servers in the chain. through chain all requests are forwarded to the number one server. in case you cannot take transport of numerous requests, all requests may be despatched to the following chain server after which to the 1/three server. and so forth

d) *software application utility defined Networking (SDN) Adaptive*

The Adaptive SDN connects records of the pinnacle layers of the network to get information about the bottom line of the network. This lets in you to get facts about the reputation of the servers, their walking fame, the u . s . a .of the community infrastructure, and the degree of network congestion so that everybody can participate in the selection-making tool.

3.2 *magnificence of Load balancing algorithms:*

relying at the contemporary-day-day country of the device, load balancing algorithms may be divided into 2 lessons as tool load and tool topology [10].

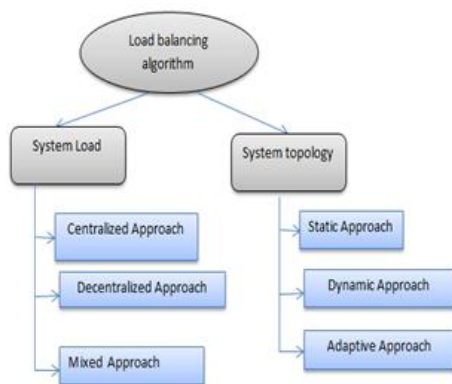


Figure 2: Classification of Load balancing algorithms

A) *beauty of device Load*

Centralized method: One node is accountable for dealing with aid allocation all through the entire network.

allocated method: every node independently generates its load vector with the resource of the usage of gathering tax information from distinct nodes. selections are made domestically the use of close by load vectors. This approach is appropriate for extended not unusual systems, consisting of cloud computing.

mixed technique: use a centralized and allocated method.

B) *kind of device Topology*

B) kind of device Topology

Static technique: Static algorithms are typically suitable for a homogeneous and strong environment and can supply higher consequences in those environments. however, they may be normally non-flexible and can not meet the dynamics of dynamic variables.

Dynamic approach: Dynamic algorithms are extra flexible and may go through in mind unique characteristic sorts within the tool every earlier than and at some point of execution time.

Adaptive technique: these methods are enough to residence the distribution of the burden to the device to dynamically exchange parameters and algorithms. they're able to provide higher typical overall performance at the same time because the country of the system is continuously changing and is appropriate for scalable distribution structures together with cloud computing.

those algorithms can adapt to adjustments and provide better effects in a dynamic and heterogeneous surroundings. as the distribution of features becomes greater complicated and dynamic, a number of these algorithms may moreover grow to be ineffective and generate greater overheads than vital, which generally worsens the wonderful of services.

1. worrying situations in the direction of LOAD BALANCING IN CLOUD COMPUTING

some statistics can be advanced to decorate the distribution of cloud computing duties.

Scalability: research the way the machine achieves a load balancing set of rules with a limited amount of nodes or processor.

Throughput: described as the most important extensive sort of obligations done over a given duration. to enhance device overall performance, the parameter wishes excessive normal overall performance.

Fault Tolerant: described because of the reality the set of guidelines's ability to artwork as it need to be and calmly, despite the fact that the pattern node fails.

Migration time: The time it takes to migrate the approach from one system node to some distinctive is referred to as migration time. This time there can be continuously a small quantity to decorate device overall performance in the cloud.

reaction time: Specifies the shortest time needed to respond to a specific load balancing set of policies in a distributed tool. This time it desires to be decreased to improve performance.

useful aid use: Use tool assets as masses as viable. effective load balancing algorithms make the notable use of available sources.

ordinary overall performance: that is the entire machine performance within the cloud. If all tool configurations are superior, standard everyday normal performance may be superior.

IV. CONCLUSION

the precept hassle in Cloud storage is load balancing of information.Help you make good use of resources and thus improve system performance. Different load balancing algorithms with balancing methods were analyzed in this paper, and also noticed some complicated problems towards load balancing in Cloud Computing. It focuses on reducing overhead costs, improving system performance, maximizing capacity utilization, service intervals and resource utilization, etc. In the future, different parameters and various types of sentimental computing technique can be included in the cloud and wider use of resources.



REFERENCES

1. Cisco (2009) Cisco visual networking index: Forecast and methodology, 2009-2014. White paper.
2. Weiss A (2007) Computing in the clouds. *Networker* 11: 16-25.
3. Hayes B (2008) Cloud computing. *Commun ACM* 51: 9-11.
4. Mell P, Grance T (2009) Draft NIST Working Definition of Cloud Computing. *Nat Inst Standards Technol*
5. Rimal, Prasad B, Choi E, Lumb V (2009) A taxonomy and survey of cloud computing systems. *Proceedings of 5th International Joint Conference on INC, IMS and IDC, IEEE* .
6. Sinha PK (1997) *Distributed operating Systems Concepts and Design*. IEEE Computer Society Press.
7. <https://kemptechnologies.com/in/load-balancer/load-balancing-algorithms-techniques/>
8. Wang SC, Chen CW, Yan KQ, Wang SS (2013) The Anatomy Study of Load Balancing in Cloud Computing Environment. *The Eighth International Conference on Internet and Web Applications and Services* 230-235
9. Gulati A, Chopra RK (2013) Dynamic Round Robin for Load Balancing in a Cloud Computing, *International Journal of Computer Science and Mobile Computing* 2: 274-278.
10. Khan Z, Singh R, Alam J, Saxena S (2011) Classification of Load Balancing Conditions for parallel and distributed systems. *International Journal of Computer Science Issue* 8: 411- 419.