Load Balancing In Oracle Cloud

Zarina Mathew
Department of Computer Applications
Amal Jyothi College of Engineering
Kanjirappally, India
zarinamathew111@gmail.com

Jetty Benjamin
Department of Computer Applications
Amal Jyothi College of Engineering
Kanjirappally, India
jettybenjamin@amaljyothi.ac.in

***Abstract*- This study proposes the Oracle load balancing solution for cloud computing. Load balancing is a strategy for distributing work load and resources in cloud computing. The users can get access to the resources they required and services are available on demand as a Pay-For-What You use model. The Oracle Cloud Infrastructure load balancer provides excellent availability, scalability, and resource usage. The paper discuss the load balancing concept in cloud computing and presents the step by step process of implementing load balancer in OCI.**

Keywords—Oracle Cloud Infrastructure, Load balancer

I.INTRODUCTION

Enterprise workloads today increases, consequently requirements in challenges getting complex, to overcome the complexity and ensure high availability of resources for enterprise applications we can use the oracle cloud infrastructure.OCI load balancer provides:

- Automatic traffic distribution from one entry point to multiple servers.
- Makes better use of resources.
- Facilitate scaling.
- Ensures high availability.

Load balancing can be used to boost utilization and throughput, reduce latency, shorten response times, and prevent system overload. It makes certain that no single node is overburdened. It finds out overloaded and under loaded nodes and then balances the load among them. In cloud computing to optimize different performance parameters, researchers proposed various load balancing approaches.

Virtual Cloud Network(VCN) A private network that you create in Oracle data centers, complete with firewall restrictions and various sorts of communication gateways. In the authorized IP address ranges, a VCN covers a single, continuous IPv4 CIDR block of your choice.

Before you launch a load balancer, you'll need at least one virtual cloud network.

In a virtual cloud network (VCN), subnet is a subdivision, which likes 10.0.0.0/24 or 10.0.1.0/24. A subnet can span multiple regions or be contained inside a single availability domain. A subnet is a set of IP addresses that do not overlap with those of other subnets in the VCN. You define the routing and security rules that apply to each subnet.

Backend server weighting allows you to fine-tune each of these policy kinds when processing load or capacity differs among backend servers. Weighting has an impact on the proportion of requests directed to each server. You apply weights depending on your own criteria, such as the traffic- handling capacity of each server. The weights must be between 1 and 100.

TCP load balancers, cookie-based session persistent HTTP requests (sticky requests), and non-sticky HTTP requests all have various load balancer policy determinations. When directing an initial incoming request to a backend server, a TCP load balancer evaluates policy and weight criteria. This connection's following packets all go to the same endpoint. Requests are forwarded to the backend server specified by the cookie's session information via an HTTP load balancer configured to support cookie-based session persistence.

A. Round Robin Algorithm

Round Robin is the default load balancing policy. Weighted Round Robin algorithm is used in oracle cloud load balancer to balancing the load. It is a simple way to distribute the request to a group of servers. Requests are cyclically directed to accessible servers in this technique. It allows the site administrator to assign weights to every server based on criteria such as traffic handling capacity. Incoming traffic is distributed sequentially to each server in a backend set list by this policy. The load balancer repeats the list in the same order once each server receives a connection. A simple load balancing algorithm is Round Robin. It works best when all of the backend servers have equal capacities and the processing burden for each request is consistent.

B. Types Of Load Balancers

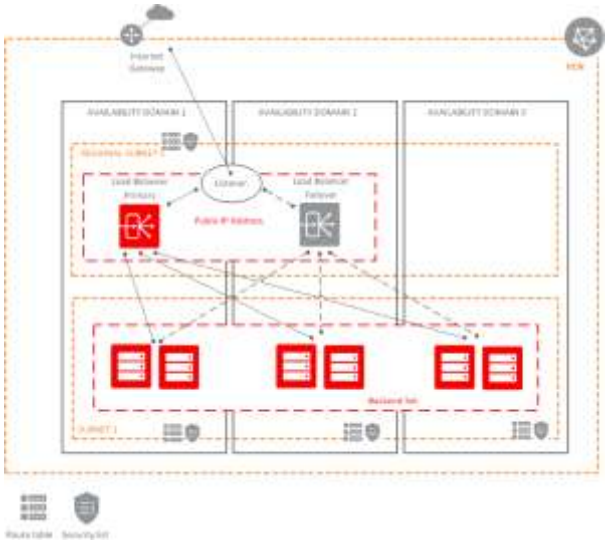
1. Public Load Balancer

   Public load balancer is used when the traffic is originating from public internet. It accepts traffic from the internet. Public load balancers use two subnets for primary and secondary load balancer. We cannot determine which subnet act as primary load balancer and which is secondary.

DOI: 10.5281/zenodo.6227939
ISBN: 978-93-5607-317-3 @2022 MCA, Amal Jyothi College of Engineering Kanjirappally, Kottayam

2. Private Load Balancer

The virtual cloud network is where the traffic originates.It is used when the traffic is not coming from the internet.it requires only one subnet to host both primary and standby load balancers. Highly available within the availability domain.



The figure shows the system configuration of public load balancing.

II. LITERATURE SURVEY

As Noman Islam ,et al havesaid in their paper file load balancing is the process of spread the burden across network of multiple devices for minimizing the overall waiting time of device and reduce the response time. Load Balancing is aimed at minimize response time for resources and work that increases system efficiency at a lower cost.

Sambit Kumar Mishra, Bibhudatta Sahoo and Priti Paramita Parida, "Load balancing in cloud computing: A big picture". The cloud computing is defined as an Internet-based computing model for share the resources, software, and other datas to multiple devices when the user demand it. The load balancing in cloud computing has a sound role on the performance of the system. There are mainly two types of load balancing algorithms which are static and dynamic. Static-based load balancing methods are best suited for environments with a uniform system. Dynamic-based balancing algorithms are mostly adaptable for both heterogeneous and homogeneous environments.

Foram F Kherani, Prof.Jignesh Vania, "Load Balancing in cloud computing", in virtual machines balancing load means that anyone of the available machine is partially loaded while others are heavily loaded. Load balancing is one of the major factors to enhancing the working performance of the cloud service provider.

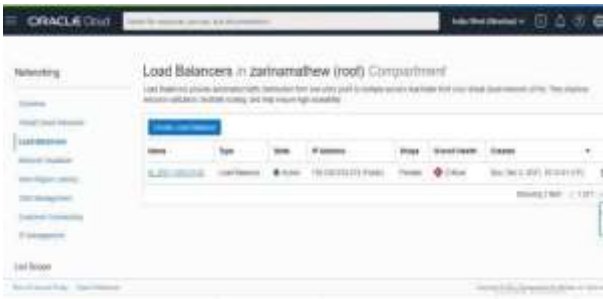
III. METHODOLOGY

Steps for creating load balancer in OCI:

- Initially create a virtual cloud network and subnets for that VCN.
- Create two compute instances(Webserver1 and Webserver2)
- In OCI console go to networking and click on load balancer
- Select public load balancer
- Select appropriate VCN in which we want to create the load balancer
- Select the subnets
- Create a backend set
- Add backend(Servers)to your backend set
- Create a listener for the load balancer
- Installing http server on the instances

Screenshots

1. Oracle Cloud Infrastructure (OCI) console.



2. Make a load balancer.

DOI: 10.5281/zenodo.6227939
ISBN: 978-93-5607-317-3 @2022 MCA, Amal Jyothi College of Engineering Kanjirappally, Kottayam

3. Add backend set and configure listener, select the load balancer shape (load balancer capacity or bandwidth).

4. Load balancer details

5. Add backend servers to the backend set.

6. Create listener

Load balancer distributes the incoming application traffic on the basis of two policies:

- Health check policy
- Load balancing policy

The health check policy is a check to ensure that backend servers are available. Based on time interval specified, the load balancer applies the health check policy to continuously monitor the
backend servers. If a server fails the health check, the load balancer takes the server temporarily out of the rotation. If the server passes the health check later, the load balancer puts it back into rotation. The health status of your load balancer is an indicator of its overall health.

Load balancing policy tells the load balancer how to distribute the incoming traffic to the backend server. And the load balancer accepts which incoming traffic is determined by the listener. In OCI we can provide pre-provisioned load balancing capacity (bandwidth) by selecting the load balancer shape. We cannot change the shape of the load balancer after create it.

IV. RESULT

After creating the load balancer we can see the details of the load balancer. Each load balancer has a certain configuring limits such as one IP address maximum of 16 backend set, 512 backend servers per backend set and maximum of 16 listeners. Verify the load balancer by checking the public IP is open for the port number given. That is, test connection to load balancer.

OCI load balancer distribute the work load and resources in cloud computing environment to multiple servers using backend set consist of backend servers load balancing policy, health check policy, SSL handling.Round Robin algorithm is using for the traffic distribution.

V. CONCLUSI

ON The unique capabilities of load
balancers are:

- Choice of bandwidth: One can provision the load balancer as per the application traffic density.
- High availability: One can distribute the traffic across multiple availability domain.
- Access control: We can configure load balancer to adhere to specified network security policy.
- Support for network (TCP/IP) and application traffic load balancing.

Oracle Cloud Infrastructure Load balancer is an ideal front end for the application.

VI. FUTURE WORK

In this paper we discussed the concept of load balancing in cloud computing environment and how it is implementing in oracle cloud infrastructure (OCI). The creation of load balancer and the effective utilization of resources are done in OCI.There are various cloud platforms that provide load balancing for efficient resource utilization and thus enhances the availability. So, the future work we can do a comparative study on load balancing in various cloud platforms such as

DOI: 10.5281/zenodo.6227939
ISBN: 978-93-5607-317-3 @2022 MCA, Amal Jyothi College of Engineering Kanjirappally, Kottayam

Amazon Web Services (AWS), Microsoft Azure, Google Cloud Platform (GCP) and Oracle Cloud Infrastructure (OCI). In terms of load balancing, we can see how one cloud service platform differs from the others.

VII. REFERENCES

[1] Sambit Kumar Mishra, Bibhudatta Sahoo "Load balancing in cloud computing: A Big Picture". February 2020[Journal of King Saud University-Computer and Information Sciences]. https://doi.org/10.1016/j.jksuci.2018.01.003

[2] Rajgopal K T,Dr.KR.Anil Kumar,Nagesh Shenoy H,"Load Balancing in Cloud Computing:A survey on popular techniques and comparative analysis". 2018-06-08{Global Jounal of Computer Science and Technology].
https://computerresearch.org/index.php/computer/article/view/1743

[3] Dalia Abdulkareem Shafiq, N.Z. Jhanjhi, "Load balancing techniques in cloud environment:A review". 2 March 2021[Journal of King Saud University-Computer and Information Sciences].
https://doi.org/10.1016/j.jksuci.2021.02.007

[4] Shahbaz Afzal,G.Kavitha-"Load balancing in cloud computing-A hierarchical taxonomical classification".23 December 2019.
https://doi.org/10.1186/s13677-019-0146-7

[5] Pandaba Pradhan,Prafulla Ku.Behera,B.N.B.Ray-"Modified Round Robin Algorithm for Resource Allocation in Cloud Computing"-2016[Procedia Computer Science].
https://doi.org/10.1016/j.procs.2016.05.278

DOI: 10.5281/zenodo.6227939
ISBN: 978-93-5607-317-3 @2022 MCA, Amal Jyothi College of Engineering Kanjirappally, Kottayam