LAB 3

Q) 1) Describe Load Balancing and its significance in Cloud Environment.

Answer –

Load balancing in cloud computing simply means distributing the task when there is a heavy load on one server. This kind of service mainly helps at the time of sale peak time in eCommerce. This can be not only distributed among various servers but between various regions as well. This improves organization performance as well as it helps to provide uninterrupted customer service. Load balancing tries to speed up different constrained parameters like response time, execution time, system stability, etc.

Q) 2) List the Load Balancing Service available in AWS, Azure and GCP.

Answer –

- There are three types of load balancers available in AWS.

Classic Load Balancer (CLB) operates on both the request and connection levels for Layer 4 (TCP/IP) and Layer 7 (HTTP) routing. It is best for EC2 Classic instances.

Application Load Balancer (ALB) works at the request level only. It is designed to support the workloads of modern applications such as containerized applications, HTTP/2 traffic, and web sockets.

Network Load Balancer (NLB) operates at the fourth layer of the (OSI) Open Systems Interconnection model. It is capable to handle millions of requests per second.

- There are three types of load balancers in Azure: **Azure Load Balancer, Internal Load Balancer (ILB), and Traffic Manager**. The various load balancers ensure that the traffic is sent to healthy nodes.

Microsoft's Azure Load Balancer offers a higher level scale with layer 4 load balancing across multiple VMs (virtual machines).

Internal Load Balancer (ILB) has an internal-facing Virtual IP. Meaning, users can apply an internal load balancing for virtual machines (VM) that are connected only to an internal Azure cloud service or a virtual network.

Traffic Manager is an internet-facing solution that balances the traffic loads at various endpoints using a policy engine as well as a set of DNS queries. It can route traffic to any region's service and even to non-Azure endpoints.

- There are three deployment types of load-balancing services in Google: **Global, Network and Internal.**

Global Load Balancing supports HTTP(S) traffic for modern web-based applications. Traffic is distributed to the region that is closest to the calling user, provided the region has available capacity.

Network Load Balancing directs traffic across virtual machine (VM) instances in the same region in a VPC network. Any TCP and UDP traffic can be load balanced on the basis of source, destination port, and protocol so that the traffic from the same connection reaches the same server.

Internal Load Balancing is a regional load balancer that distributes the internal traffic across a set of back-end instances without requiring a public IP address.

Q) 3) Create 2 Identical AWS EC2 / GCP VM Instances (Instance Name: Regno_EC2_VM1, Regno_EC2_VM2) and install a web server of your choice in each of the instances to host the web site of your organization globally.

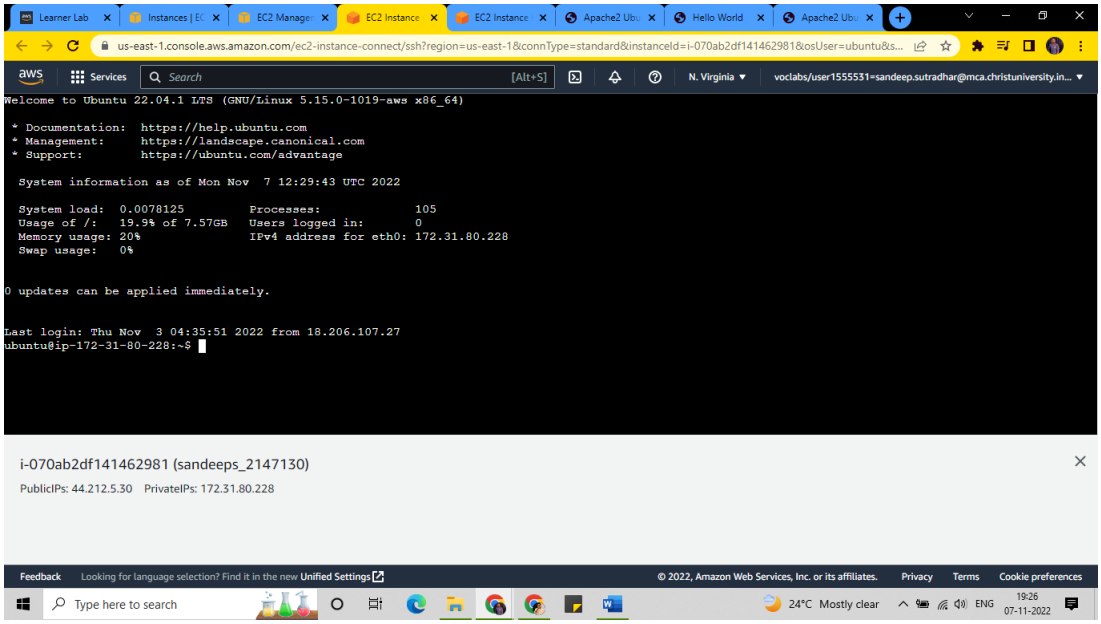
Answer –

Step 1: Click on Launch Instance → Give the name of your instance →Select Ubuntu in Quick Start → Leave everything as by default →I n the key pair login you may create a new key pair (In that also leave everything as same) → In Firewall (security groups) Allow both http and https with ssh as well → Click Launch Instance

Step 2: Wait for some time for launching the instance and make it on Check the security group as well for verifying if port 80 is open.

Step 3: Create instance 2 as the same (Repeat step1 and 2)

Step 4: Select One instance then select Connect option → **EC2 Instance Connect** → select Connect → A new tab will open



Step 5: You may write your HTML CODE here to modify the web page

Step 6: After modifying and saving the Html page Come to instance page and open the public IP to check if it is running properly.

Q) 4) Create a Load Balancer to ensure the fare allocation of tasks among the web servers deployed on the Virtual machine instances.

Answer –

Step 1: In the Load Balancing option (Left side Menu of Instance Page) select Target Groups → Select Create Target Group

Step 2: Give the Target Group Name(like - target-for-app-load-bal)→ Leave everything as default → Next → Select the instances for Load Balancing → Include as pending Below button → Select Create Target Group.

Note: This is only responsible for changing the web server in the specific time which can be changed in the previous step of time.

Step 3: Click on Load Balancer under Load Balancing option → Select Http and HTTPS option→ Give the Load Balancer name → In the mappings select the regions you want to provide (Minimum Two , But we have selected all) → In the Security groups Select the security group which we have selected while creating the instances and keep default security group as well. → Select Target group (Created in step 2) → Click Create Load Balancer

Step 4: Wait for some time to make it Active.

Step 5: Copy the DNS name(app-load-bal-sandeep-1773173604.us-east-1.elb.amazonaws.com) and paste it into the browser to open the DNS. It will open the first instance by default then while you refresh it will open the second instance after the time which by default is 300 miliseconds.