<http://18.224.213.62/> this is my public ip of vm

**Load balancer**

1. **Application load balancer** – it can handle all the http and https requests.

Works on port 443 and 80.

it works on OSI, layer 7th. 7th layer is application.

OSI Layer:

7Application

6 Presentation

5 Session

Transport

Network

Data link

Physical.

1. **Network LB**

Osi layer 3, it can handle millions of requests/ sec.

It can handle multiple users at a time, ex: amazon, flipkart, naukri.com etc

But it has high cost

1. **Classic LB**

OSI layer 4, 7, it can handle http, https and tcp req.

It is rarely used, or not used at all, coz it is only recommended for old EC2 instances.

**LAB**

Putty

Cd /var/www/html

Vi index.html

Type hi and save it.

Service httpd restart

Similarly create one more instance and install httpd and update index.html file .

Day 4:

Every instance have some limit on users, so when users increases, instances needs to be increased.

So there is a need to increase the instance at any point of time , for that case we need system which can auto increase and delete the number of instances automatically

To handle this sudden increase in the users, and handle the environment we need autoscale concept.

**Types of AS 2 :**

**Vertical :** inc/dec size of instances , configuration or cpu, memo , like t2micro to t2 **large**

**Horizontal :** it will pipeline the same kind if environmentautomatically.

T2 large t2 large t2 large.

**Metrix**: based on matrix scaling is done, ex cpu usage, => utilization > 80%, then it creates inst automatically.

It creates and configure the sites and instances at the same time.

Viveversa it will remove the extra inst if < 80% automatically, for **how long setting** can also be done in mints

ELB (Elastic load balancer)

Autoscaling group

**EFS : Elasti file system, stores flat files of instance data , it is also connected through RDS(relational db), and it stores all the data of the instances, and all the authentication is done here.**

**So even whn the instance is deleted the data is always there EFS. It is similar to the shared file system on windows.**

**LAB : Autoscaling**

**Autoscaling is an region specific service similar to ec2, AS is a part of ec2.**

**In order to Configure instances in automatic way**

Click on launch configuration right side

GOAL

ELB (Elastic load balancer)

Autoscaling group

**In putty**

Yum update

Yum install httpd

Service https start

Systemctl enable httpd : this command will will keep the services running even after reboot.

**AIM**: to install the services during instance creation.

Copy the sourcecode of any website, and paste it in the file name index.html

Upload this file in a bucket.

Type this command in order to save the file to any location.

#Aws s3 cp <bucket\_name>

Where cp is copy.

Eg: aws s2 cp s3://<buk\_name> --recursive

and it(index file, or all the files present int the bucket) needs to be pasted to the instances var/www/html location.

**surabhi2893** is my bucket name.

in order EC2 fetch the data from s3, EC2 needs write permission on s3, and this thing can be done using Policy, and this policy needs to be applied over a **role.**

**Note:**  1 aws service can call different aws services on behalf of you.

**Note**: IAM role cannot be applied on user level.

Script needs to be placed in the auto scaling(right botton on ec2 page) >> next >> attach role >> adv ko expnd and the script can be placed in the userdata

Script:

**#!/bin/bash/**

**yum install httpd –y this –y is for the yes/no options**

**yum update -y**

**aws s3 cp s3://surabhi2893 /var/www/html/ --recursive {recursive is used to pick all the files fom the bucket}**

**service httpd start**

**chkconfig httpd on**

all the data gets coiped in the round robin method in all the regions I =n a load balancer.

**We can scale on the basis of CPU and network utilization.**

**Warm up time : say 300 secounds.: time after which the installation will complete and the data will start loading on the newly created instances.**

**Route53: domain service provided by AWS**

Classless inter domain routing (CIDR)

**5 IP which aws reserves 10.0.0.0 network**

1: vpc router

2 dns server

3 future use

255 n/w broadcast address

How to create a 2 tier instance

10.0.01.0/24 tier 1 environment (front end)

10.0.2.0/24 tier 2 (backend)

More tier increases the security of the environment.

By default all the subnets(custom) are private.

Que: how to convert private subnet into public subnet

Ans: Internet gateway.

Create a new IG and attach it to your newly created vpcnet 1 .

Then create route table, and attach it to the vpc.

Select the route new created and attach the subnet1 to it, to make it public.

We can make only subnet 1 public because it is gui and front end.

Qye: how to download pcks from internet without providing internet connectivity to instance?

And: NAT gateway and NAT instance.

Create nAT gateway in Public subnet, and attach it to subnet 2

Then got to route table and attach subnet 2 to it and attach it to the NAT gateway.

Internet gateway makes subnets public and we can only make subnet 1 as public as it has gui etc

But subnet 2 contains data base info and it shdnt be public, that is why we need to use NAT gateway.

NATGateway is a service

NAT Instance is VM

NAT Gateway is preferred over NAT Instance

As nat instance is a vm and it is cost affecting, and we need HA, updates, patching , security etc needs to be managed by us.

Aue: How to access db instances ?

2 types of IP :

Static or elastic

dynamic

ssh -i mykey.pem [ec2-user@10.0.2.84](mailto:ec2-user@10.0.2.84)

subnet 1 – make it public

SN 2make it private

Create vm 1 – in SN 1

VM2 in SN 2

Login to vm1 as ec2 user

Vi key.pem (paste penkey content)

Sudo ssh –i key.pem ec2-user@<private ip of VM2>