**What is cloud computing?**

Cloud Computing is on-demand delivery of IT resources over the internet with pay-as-you-go model.

# Delivering of computing services across the internet

1. Computing (Ec2, ELB,Autoscaling)
2. storage (database(sql,S3,EBS,EFS)
3. network(VPC, routeTB, IGW,NATGW)
4. security(security grp, Nacl)

Types of top cloud providers:-

1. AWS
2. AZURE
3. GCP

<https://aws.amazon.com/about-aws/global-infrastructure/> Types of cloud

1. Private (no sharing of hardware and software,resources exclusively dedicated, accessible to single customer , single -tenant, on-premises)

1. Public(multi-tenant, customer do not need to maintain the infrastructure)
2. Hybrid(on-premises, and public cloud-AWS,AZURE)

Categories of Cloud

1. Iaas (AWS)
2. Saas (dropbox,google drive,whatsapp)
3. Paas (Used by developer to create a software)

What is AWS?

Amazon Elastic Compute Cloud (Amazon EC2) provides **scalable computing capacity** in the Amazon Web Services (AWS) Cloud. Using Amazon EC2 eliminates your need to invest in hardware up front, so you can develop and deploy applications faster. You can use Amazon EC2 to launch as many or as few virtual servers as you need, configure **security** and **networking**, and manage **storage**.Amazon EC2 enables you to scale up or down to handle changes in requirements or spikes in popularity, reducing your **need to forecast traffic**.

* **Amazon Web Services (AWS)** is a collection of **remote computing services (web services)** that together make up a **cloud computing platform**, offered over the Internet by Amazon.com.

Search for global infrastructure map:-

Regions= 30+

AZs=100+

Min az = 2 Mandatory Max az = 6

245 countries and territory served

What does Aws offer?

1. Availability >> aws
2. Scalability >> Autoscaling
3. pay-as you-go model >>No Upfront cost How to create Aws account:-

<https://aws.amazon.com/console/>

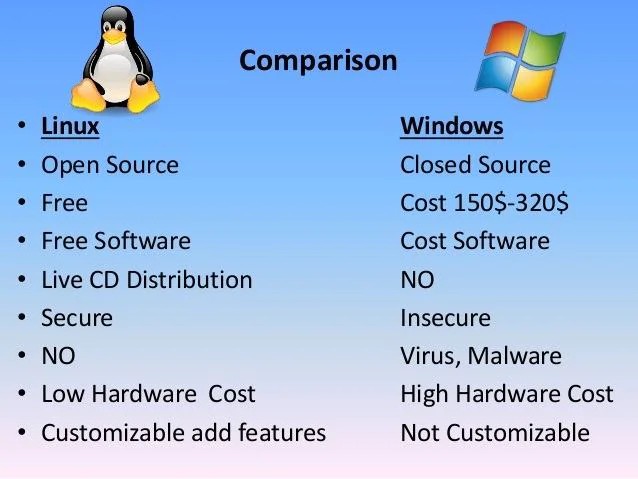
Visa debit card >> we can use it multiple times

Monolithic(Complete website is developed and deployed) vs Microservices(Development and deployment as per modules and functionality like >> login page, account page)

Industry infrastructure:-

1. Dev env
2. Uat env
3. Prod/live env

**Dev**elopment + **Op**erations = **DevOps**



OS :-

1. Windows
2. Linux- amazon linux(free-tier), ubuntu, centos, redhat, fedora
3. Mac- darwin

What is AMI?

An AMI is a **template** that contains the **software configuration** (operating system) required to launch your instance

Types of instances and types

General Purpose

provide a balance of compute, memory, and networking resources,and can be used for a wide range of workloads.

* [Compute optimized](https://docs.aws.amazon.com/AWSEC2/latest/UserGuide/compute-optimized-instances.html)

Compute optimized instances are ideal for compute-bound applications that benefit from high-performance processors.

* [Memory optimized](https://docs.aws.amazon.com/AWSEC2/latest/UserGuide/memory-optimized-instances.html)(RAM)

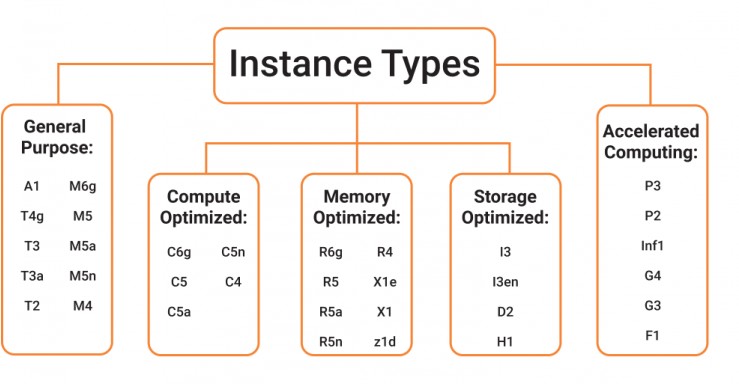
Memory optimized instances are designed to deliver fast performance for workloads that process large data sets in memory.

* [Storage optimized](https://docs.aws.amazon.com/AWSEC2/latest/UserGuide/storage-optimized-instances.html)

Storage optimized instances are designed for workloads that require high, sequential read and write access to very large data sets on local storage. They are optimized to deliver tens of thousands of low-latency, random I/O operations per second (IOPS) to applications

* GPU Graphic instances(Accelerated computing)

GPU-based instances provide access to NVIDIA GPUs with thousands of compute cores



Ways to launch instance :-

1. On-demand
2. spot(testing,Low price)
3. Reserved



Dedicated host : mac os

Which terminal do you use to ssh into the server(ec2 instance)?

1. Mobaxterm(Windows)
2. Git Bash(Windows)
3. Cmd/powershell(Windows)
4. Terminal(mac/Linux)
5. Putty

SG – It acts as a virtual firewall for your instances to control inbound and outbound traffic.

AMI – It is template that contains software configuration required to launch your instance

Key-pair

* 1. Private (hamare system pe download hoga )
  2. Public key (aws ke pass rahega)

Status check

1. (1 /2) >> hardware(remedy> system status check)
2. (2 /2) >> software(remedy>instance status check)

Types of volumes:-

1. SSD(General purpose, provisioned IOPS)
2. HDD(Throughput optimised, cold HDD)

Linux flavors:-

Amazon linux 2,3 > yum

1. Ubuntu > apt , apt-get
2. Suse > yum
3. Redhat > yum
4. Centos > yum

* Open source >> source code is available to all
* Secure, simplified updated,light weight
* multiuser and multitasking is easy

Windows:-

* Is not open source.
* Not secure , not lightweight,
* limited multitasking
  1. Login to AWS
  2. Choose a region
  3. Services -- EC2
  4. Launch instance

# Name and tags

* A tag is a label that you assign to an AWS resource. Each tag consists of a key and an optional value, both of which you define.

# Application and OS Images (Amazon Machine Image) ( Note: Select free tier eligible )

* An AMI contains the software configuration (operating system, application server, and applications) required to launch your instance

# Instance type> t2.micro

* Select an instance type that meets your computing, memory, networking, or storage needs.

# Key pair (login)

* You can use a key pair to securely connect to your instance.
* Download the .pem file in your local system.

# Network settings

* Select the virtual private network in which you want to launch your instance(VPC,subnet,security group)

# Configure storage

* Specify the storage options for the instance.

>> Launch instance

SG :-

Acts as a virtual firewall for your instances to control inbound and outbound traffic.

Two ways to connect :-

1. Ec2 instance connect
2. Ssh client(putty,git bash,mobaxterm)

# EBS Volume

# Ebs Volume is nothing but is a persistent storage disk which is available even when the instance will be turned off.

There are five types of EBS volumes available as below:

1. General Purpose SSD (gp2)

SSD (Solid State Drive) is the volume with which EC2 chooses as the root volume of your instance by default. For small input/output operations, SSD is many times faster than HDD (Hard Disk Drive). It gives a balance between price and performance (measured in IOPS - Input-Output Operations per second).

1. Provisioned IOPS SSD (io1)

This is the most expensive and fastest EBS volume. They are intended for I/O-intensive applications.

1. Throughput Optimized HDD (st1)

These are low-cost magnetic storage volumes whose performance is measured in terms of throughput.

1. Cold HDD (sc1)

These are even less expensive magnetic storage options than Throughput

Optimized. They are intended for large, sequential cold workloads, such as those found on a file server.

1. Magnetic (standard)

These are older generation magnetic drives that are best suited for workloads with infrequent data access.

server apache :-

1. Web Server >> static content >> eg. httpd, apache2(linux), ,nginx,
2. Application server >> dynamic content >> tomee , jboss

Shebang line >> #!/bin/bash Bootstrap script

#!/bin/bash yum update -y

yum install httpd -y

systemctl start httpd # apache server systemctl enable httpd

systemctl status httpd

echo “my web server” >> /var/www/html/index.html

Load balancer >> >>

1. Application lb >> mbl or web appl > 7th layer of OSI(open system interconnection model how the systems communicate with each other, it is a concept of networking not related to devops) model > http/https
2. Network lb >> low latency efficient network> 4th layer of OSI model>>tcp/udp/tls>> response time kam chaiye
3. Classic lb >> decommissioned
4. Gateway lb >> new addition in aws

TCP(connection oriented)/UPD(connectionless internet protocol) >> transport layer ke protocol he

Private IP range:-

1. 10.0.0.0 — 10.255.255.255;

2. 172.16.0.0 — 172.31.255.255;

3. 192.168.0.0 — 192.168.255.255

ELB :-

1. EC2/web-server/target
2. Target group
3. Load balancer

Load Balancer (LB)

It is automatically distributing incoming traffic across multiple targets in one or more availability zones.

Load balancer is used to distribute traffic received from the internet to multiple machines attached to it.

1. Application Load Balancer

For creating we have created 2instances 1 in zone C and 1 in zone D. Security group we have added SSH,HTTPS,TCP and HTTP rules.

In instance we need to install apache server >> check above given installation steps. We need to create a target group so we created a target group and registered the target in the Target Group.

Then we created Application Load Balancer, selected zones from which your target group is, added a listener, security group, selected target group and created ALB.

1. Network Load Balancer

We can create Network LB, Low latency LB, used 4th layer osi model i.e tcp/udp protocol

Same procedure as ALB but create target group should have tcp protocol.

Then go inside vi sample.sh

#write

#!/bin/bash

nwlb="here copy paste your DNS whole name"

for((i=0;i<100;i++))

do

curl ${nwlb}

done

Save and quit this

Bash sample.sh

Auto-scaling :-

It is cloud computing features that automatically adjust the number of compute resources allocated to an application based on its current load or demand.

Auto Scaling helps you ensure that you have the correct number of EC2 instances available to handle the load for your application.

You create collections of EC2 instances, called Auto Scaling groups.

You can specify the minimum number of instances and the maximum number of instances in the AutoScaling group, and Amazon EC2 Auto Scaling ensures that your group never goes above this size.

**Vertical** >> scale up / scale down instance configuration change >> ram, storage,processor(cpu)

**(ASG)Horizontal** >>scale in / scale out >> number of instance increase/decrease

EBS Snapshot

You can create a snapshot of EBS Volume, right click on volume and create a snapshot.

From that snapshot you can again create volume

We generally use snapshot for taking backup of volumes

Snapshots are incremental and store it on S3

What is auto scaling ?

I am running a web application on an EC2 and I want to increase the load on an EC2 instance whenever the CPU utilization crosses 80% what will I use?

Advantages of auto scaling ?

Elastic :- automatically adapt to the requirements

Reliable:- counteracts the failure of instance or AZs (t2.micro)

What are the **scaling plans** in auto scaling ?

* Predictive policy
* dynamic 1. Target tracking 2. Simple 3. Step
* scheduled

Support center

* + Account and billing
  + Service limit increase
  + Technical support

Type of storage Storage

* 1. Instance type(root)
  2. EBS Volume >> 20 volume storage >> one instance >> zone specific
  3. S3 >> 100 bucket >> global service
  4. EFS >> multiple instance attach

AWS S3(Simple Storage Service):-

S3

It is storage service that makes web scale computing easier.The simple interface is used to store and retrieve data at any time from anywhere over the internet.

Store multiple files, folders, videos here

5TB data storage for 1 bucket

Store as a storage or as a backup like terraform.tfstate file

Types of Classes

It is Object Storage

You can create 100 buckets

S3 is less costly than EBS

Scalable, Durable, Flexible, Data Performance, Backup and Restore,

Security provide

S3 is Global Service and Bucket is Region Specific

Name of S3 Bucket should be unique

# Safe place to store your files

- Object based storage >>

# - Objects are stored in bucket

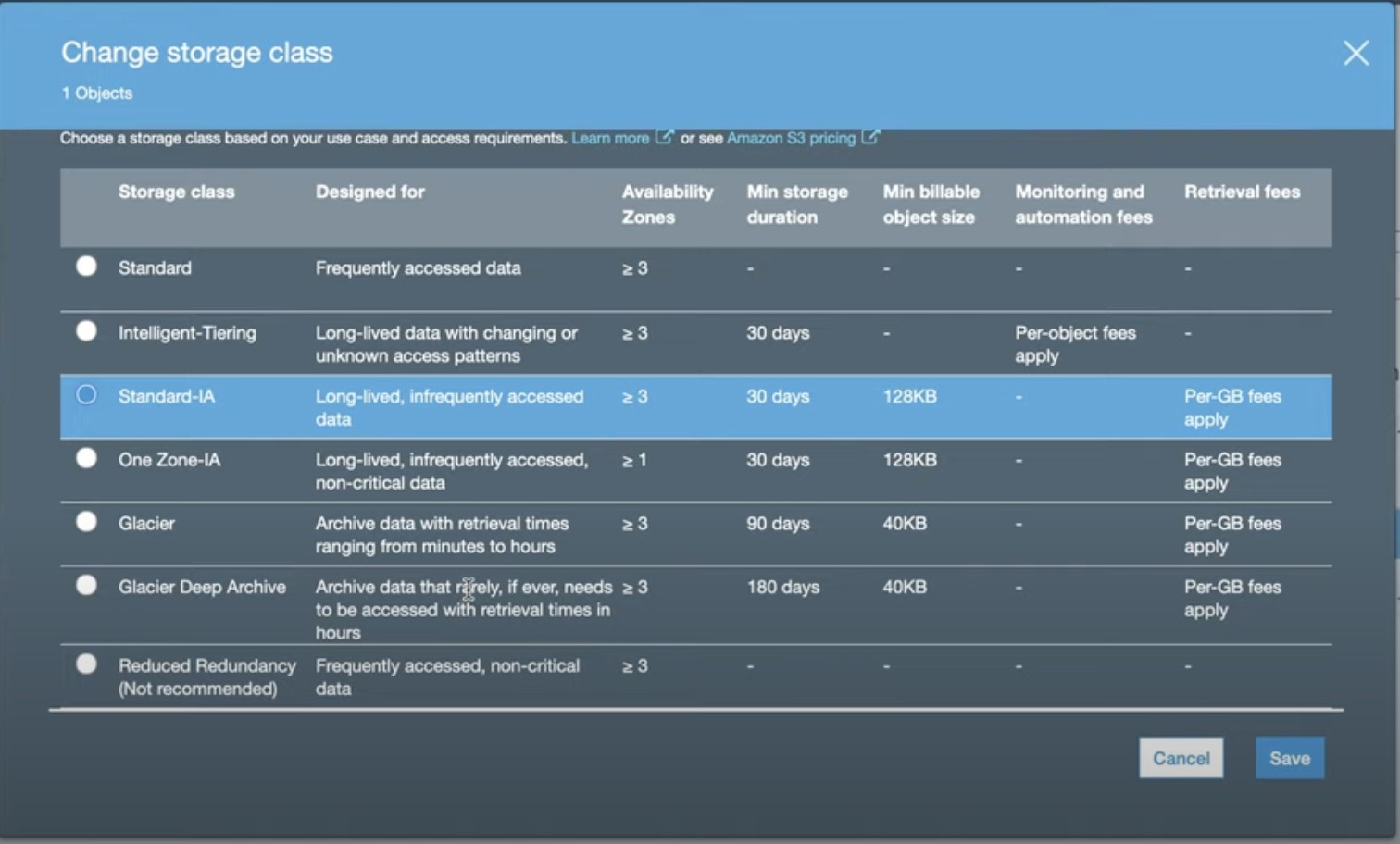
- **File size >> max 5TB**(graphical/console)

# Unlimited storage>> 256Tb

* **S3 bucket names must be unique globally**

# - Versioning

* + 1. Recover the deleted object
    2. Maintain the versions of objects



Ebs >> only one instance attach >> root volume S3

S3

* object storage
* Role based access
* Secret access key , access key

EFS >> elastic file system >> multiple instance attach via security (NFS)

* Limitation > 1. No windows support 2. No system boot volumes 3. Region specific
* Benefit > pay-as you-go ,scalable,secure :- port nfs , dynamic,fully managed by aws.

-

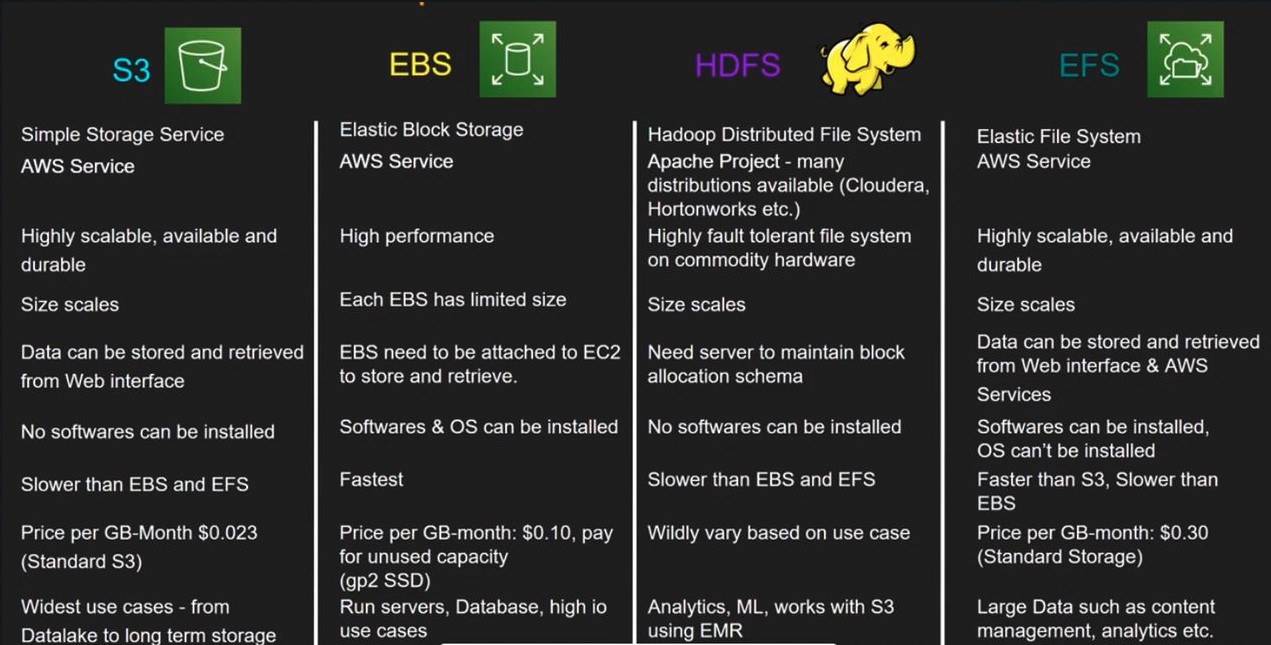
2 instance create >> sg >>nfs port open (2049) storage> EFS >>”sg efs” attach karna he to all AZs

* mkdir efs
* yum install -y amazon-efs-utils Efs >. Attach > mount via dns

RDS

RDS is managed relational database service provided by amazon web service. It simplifies the process of setting up, operating and scaling a relational database in the cloud.

* Also called SQL
* Table based structure
* Scalability
* Tabular data model



[How to delete an AWS s3 bucket?](https://www.onlineinterviewquestions.com/aws-s3-interview-questions/#question8) Aws s3 rb s3://<bucket-name>

Servers:

* 1. Web server(nginx)(icici.com)> frontend web developer >> public access
  2. Application server(tomcat)(application/logic)> backend developer >> private
  3. Database server>>backend developer(dynamo db, mysql)>> private
     + 1. >> ip address /31>>

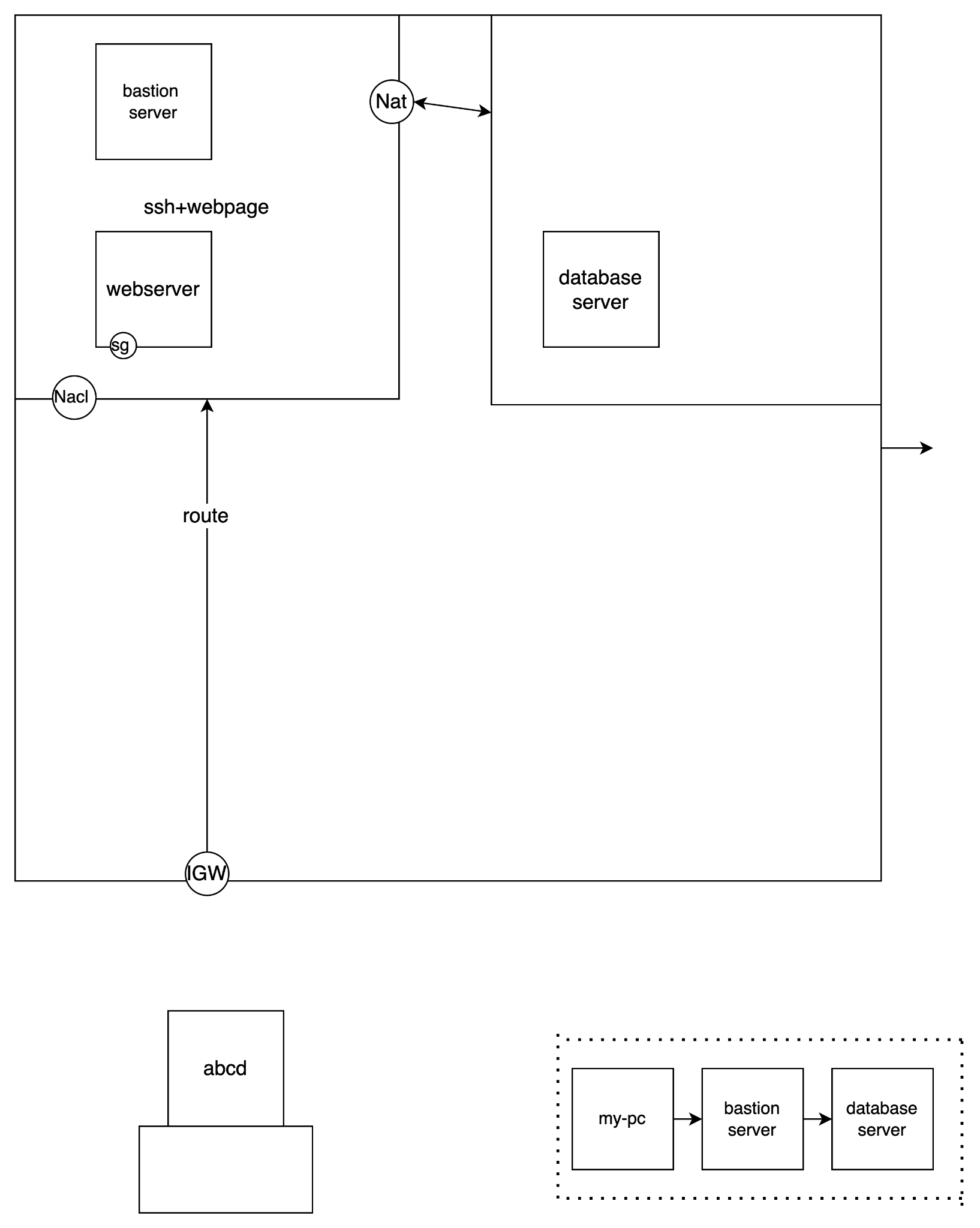
32-31 =1 . >> 2^16 = 65536

Vpc - 100 machines are allotted

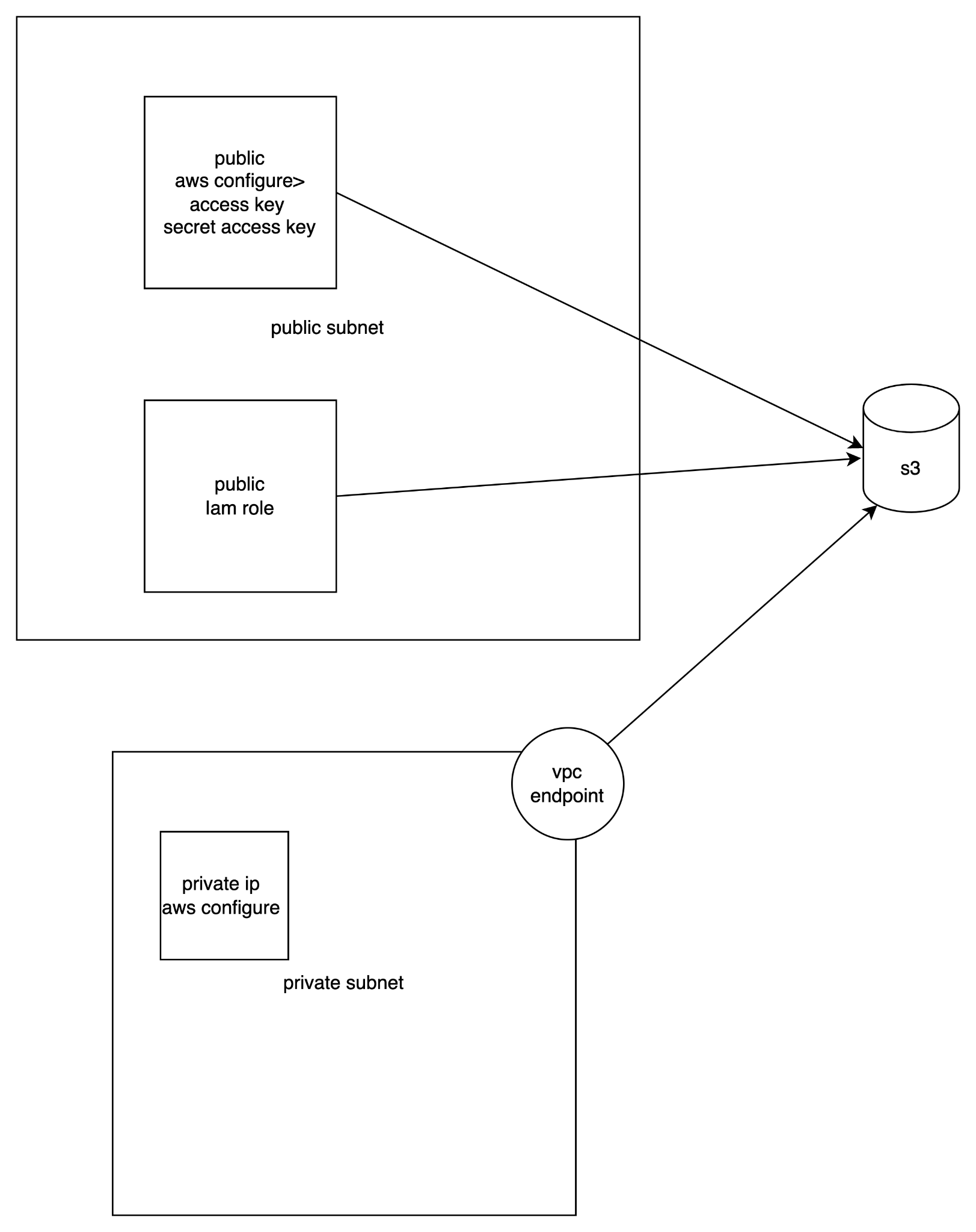
Subnet1- a machines Subnet2- b machines >> A +b >100 nahi hona chaiye

Apt install mysql-community-server-8.0.33/=8.0.3

Apt install mysql-client -y







CLOUDWATCH

Aws CloudWatch enables the monitoring of AWS environments like RDS Instances, Ec2 Instances, CPU Utilization that is nothing but CloudWatch.

Types of cloudwatch

Basic Cloudwatch

Detailed CloudWatch

Difference between basic and detailed cloudwatch

Low component

Free

Less frequent

5 min interval

ROUTE 53

It is scalable and high available Domain Name System (DNS) web service provided by AWS.

Where we can register our domain name(Godaddy, freenon, hostinger)

It convert human friendly ip address to the domain name and vice versa

Phonebook of the internet or Backbone

Advantages

Domain Registration

DNS Management

Traffic Management

Health check

Policies

Simple - Default

Weighted – split the traffic according to weight assign

Failover – active-passive failover

Latency – low network latency for end user

Geolocation – region/country specific

NS - These name servers are responsible for providing DNS resolution for the domain and answering queries about the domain's DNS records

SOA - The SOA record contains essential information about the domain's zone, including details such as the primary name server for the zone, the email address of the domain administrator, the serial number (used for zone updates), and various timing parameters (such as refresh, retry, and expiration intervals).