AWS

Cloud Computing

- AWS
- Azure
- GCP (Google Cloud Platform)

Why Cloud ?

For hosting Website

- Buy a stack of servers
- If traffic is high, server requirement is high
- Monitoring and maintaining the server
- Servers are more expensive
- Should more focus on business, not on architecture
- Traffic is varying, so most of the time servers are idle , but you have to pay for the servers
- Data generated before was less, but now a days the data size is high

Cloud

- Collections of data centers
- Used to maintain various functional applications, managing resources by combining all these data centers together through a network and then providing you the control to use these resources and manage them properly
- Cloud let you rent the services that you need and use only those services that you need. So, you ended up paying for the services that you rented and you ended up by saving a lot of money.
- Cloud service providers take care of all the issues like security, infrastructure and al other things

Cloud Computing
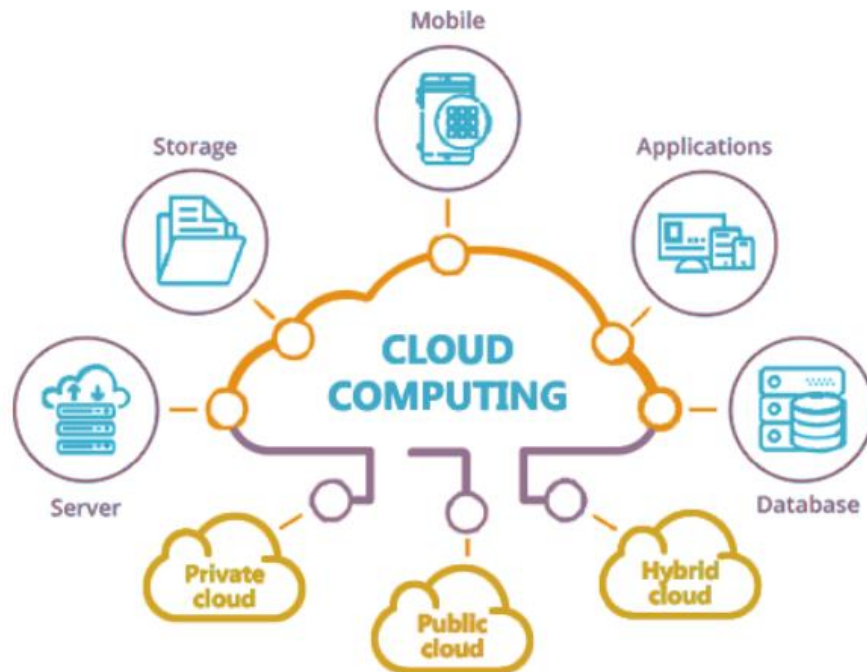
- Storing data/application on remote servers
- Processing data/applications from those servers
- Accessing it from any where

Service Model

Every user has a different requirement, based on these needs we have service models

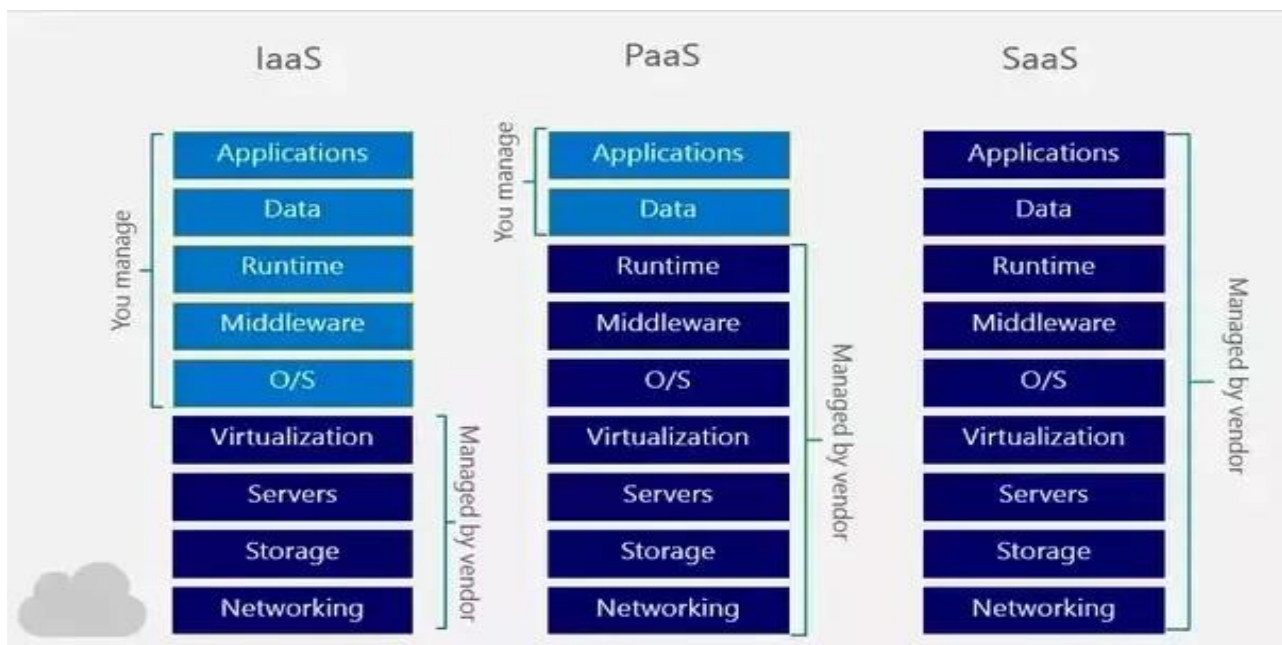**SaaS** (Software as a Service)

- Gmail
  - Only thing to send and receive mails and whatever functionality, uses as a service
  - You don't have to maintain and worry about anything
  - Everything is taken care by Google
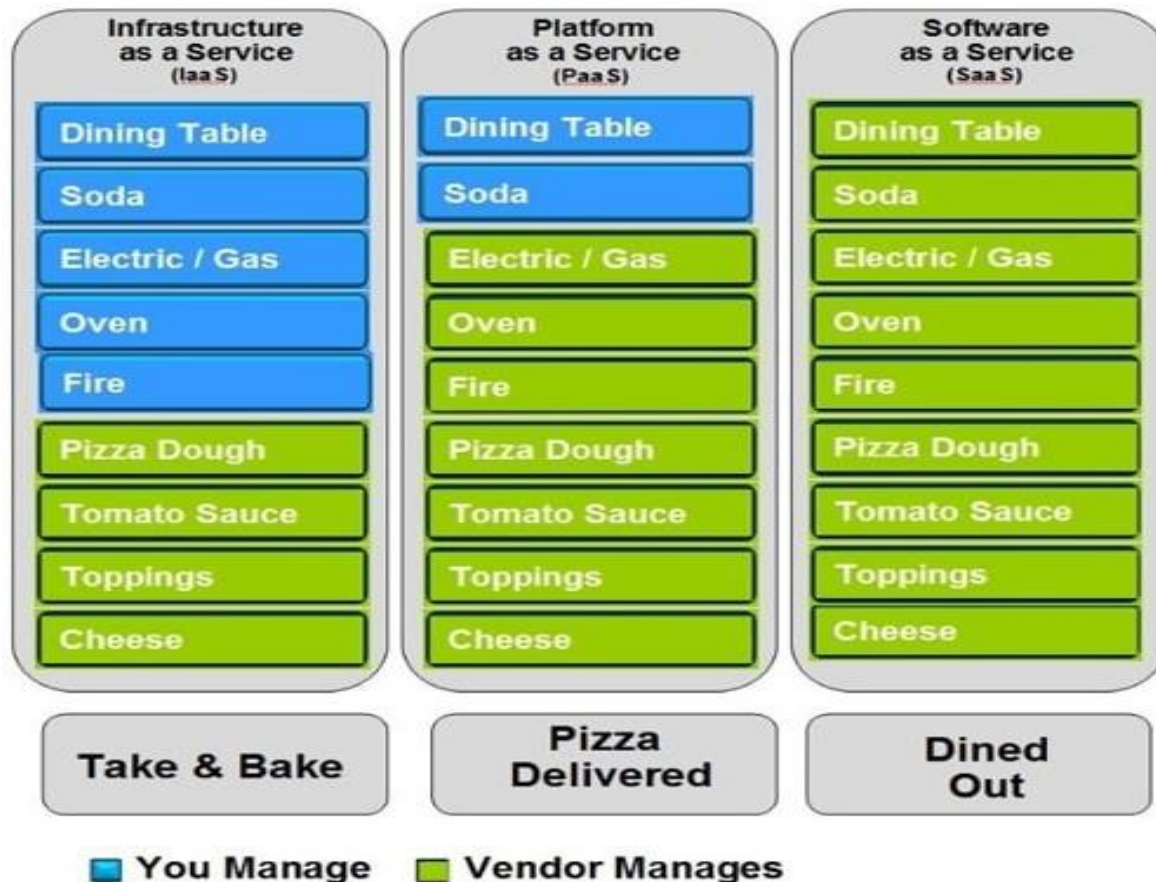
PaaS (Platform as a Service)

- Provided with the platform where you can go ahead and build your own application
- Google App Engine
  - Create your own application and put it on Google app engine so that others can use it
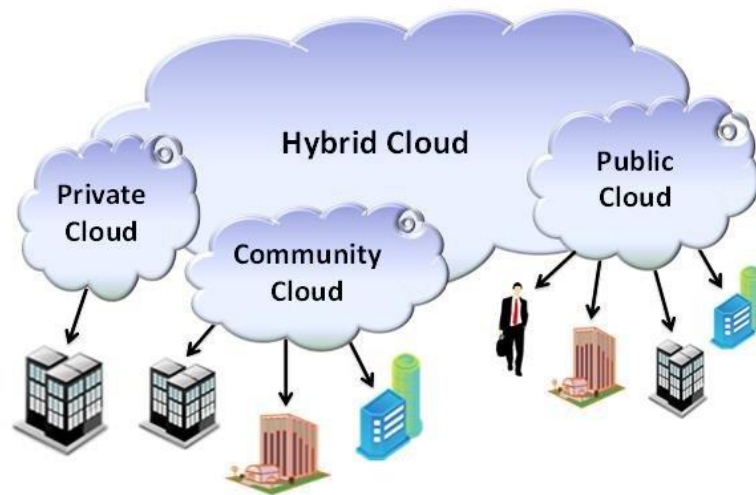
IaaS (Infrastructure as a Service)

- Provided with an infrastructure where you can choose your own OS, language, kind of technology etc

Pizza as a Service

Infrastructure as a Service (IaaS)	Platform as a Service (PaaS)	Software as a Service (SaaS)
Dining Table	Dining Table	Dining Table
Soda	Soda	Soda
Electric / Gas	Electric / Gas	Electric / Gas
Oven	Oven	Oven
Fire	Fire	Fire
Pizza Dough	Pizza Dough	Pizza Dough
Tomato Sauce	Tomato Sauce	Tomato Sauce
Toppings	Toppings	Toppings
Cheese	Cheese	Cheese
Take & Bake	**Pizza Delivered**	**Dined Out**

☐ You Manage   ☐ Vendor Manages

Deployment Models

Public Cloud

- A service provider make resources and make it available to the general public
- Easy and inexpansive way of dealing with the situations

Private Cloud

- Offer services to a limited number of people behind a firewall

Hybrid Cloud

- Combination of both public and private cloud
- You can build your own applications privately and use them efficiently
- When you sense peak in traffic then you can make it to public

AWS

- Amazon Web Services is a secure cloud service provider offering compute power, DB storage, content delivery and other functionality to help business scale and grow.
- AWS is a complete suit which is highly secure and provides us a number of services based on our requirements.
- AWS's flexibility, scalability and pricing are best compared to others
- AWS's compute capacity is 6X better than others

AWS Architecture

- Different locations across the world having various data centers are called regions
- These data centers are known as availability zones
- Choosing availability zones:
    a. Pricing
    b. Performance
    c. Customer Base
    d. Latency

Domains of AWS

1. Compute   2. Migration   3. Security and Compliance   4. Storag

5. Networking   6. Messaging   7.Database   8. Monitoring

Compute Service

It's a virtual server for running applications on Amazon EC2. It can also be understood like a tiny part of a large computer, a tiny part which has it's own hard drive, network connection, os etc. But it's actually all virtual.

Amazon EC2

Amazon Elastic Compute Cloud, EC2 is a web service from Amazon that provides resizable compute services in the cloud.

- Scalable → depending on traffic
- Flexible
- Cost efficient

Type of Instances

- General purpose → email
- Compute instance → streaming data
- Memory instance → more memory capable
- Storage instance → huge data storage
- GPU instance → heavy graphics requirement eg – 3D modelling

EC2 Pricing

- On-demand
- Dedicated
- On-spot
- Reserved

Classification of Instances

- Burstable → CPU usages 20 to 200
- EBS optimized → processing data at higher speed
- Cluster networking → different-2 clusters that serves
- Dedicated

AWS AMI (Amazon Machine Image)

Once you create an instance, you can also create an image of that instance as well i.e. a template of that instance

Elastic Beanstalk

- Automated version of EC2
- With this you don't get access to os but you still have the control over configuration
- Used to deploy an application
- It's a PaaS service used for deploying and scaling web applications written in any language on familiar servers such as tomcat, apache etc

    If you need a computer today

    Option: 1

    I.   Will go to computer shop and buy different components and assemble them
    II. Deploying an application without using elastic bean

    Option : 2

    I.   Will go to the shop and buy a computer a/c to your requirement
    II. Deploying an application using elastic bean

    Paas Example :

    - Open Shift → Red Hat
    - Google App Engine
    - Python Anywhere
    - Elastic Bean Stalk
    - Azure App Service

Features of EBS

- Fastest and simplest way of deploying your application on AWS
- Enables you to focus on writing the codes rather than spending time on servers
- Automatically scales up and down your application based on the traffic
- Give freedom to select your resources, like EC2 instance type which are optimal for your application

Components of EBS

- Application Versions
  - Used for storing multiple versions of the application
- Environments
  - Where you can actually run your application
  - EBS provides at a single point of time of time you can run multiple environments with different versions for your single application
- Environment Tier
  - Web Server Tier → used for HTTP request
  - Worker Tier → used for background task
- Environment Health
  - Grey, Green, Yellow, Red

Architecture of EBS

Web Server Tier

- Handles HTTP request from clients
- Environment is a place where you can run your application and this particular environment can run only one version of your application
- Elastic load balancer distributes all it's request among different EC2 instances of auto scaling group
- If your application is facing a lot of traffic and there is a scarcity of new instances then auto scaling group automatically installs few EC2 instance

- Whenever you try to launch an EBS environment, Beanstalk will assign your application with a suitable EC2 instance
- Every EC2 instance has a host manager which does various tasks
  - Provides detail reports of the performance of the application
  - Provides instant level event
  - Provides application log
- Security group is a firewall to your instance. Not anybody can access your instance. You can also define more security groups a/c to your requirements

Worker Tier

- It handles background task which are time consuming and resource consuming
- If web server tier comes across a task which is time consuming, it hands over that task to worker environment
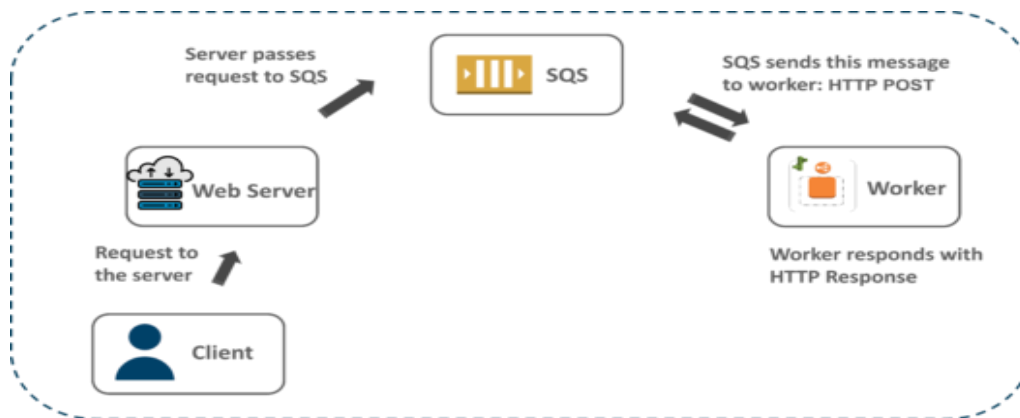
AWS Lambda

- Again an automated version of EC2
- Don't have the access to os and configuration both
- With AWS lambda you can upload your code and it executes that
- Only used to execute background task
- You can't deploy an application
- Serverless compute service

Storage Services

Cloud Storage

A storage which is available as a service connected over a network

> Service Provider → Google Cloud Platform, Digital Ocean, IBM Cloud, terremark, Azure, AWS

Amazon S3

- Simple Storage Service is used to store and retrieve any amount of data, at any time, from anywhere on the web
    - Highly durable → uses checksum
    - Highly flexible
    - Highly available
    - Cost efficient
    - Scalable
    - Secure
- S3 work with the concept of objects and buckets
- Here bucket is a container and object is a file which is stored in a bucket
    - Objects → data with key and version id
- In which region you have to create the bucket that depends on requirement. You need to decide which region is more accessible for your customer and how much cost that region charges you.
- Depending upon region cost may very
- S3 has 3 storage classes

Feature	Standard	Infrequent	Glacier
Durability	99.99%	99.99%	99.99%
Availability	99.99%	99.90%	N/A
Min. Object Size	no limit	128 kb	No limit
Min. Storage Duration	no limit	30 days	90 days
First Byte Latency	milli secs	milli secs	4 hours
Retrieval Fee	no fee	per GB	per GB

Versioning

- Keeping multiple versions of an object in one bucket
- By default versioning is disabled
- Prevent overwriting and accidental deletion
- Get non concurrent version by specific version id

Cross Region Replication

It enables automatic asynchronous copying of objects across buckets in different AWS regions. Buckets configured for cross region replication can be owned by the same AWS account or by different accounts

S3 Transfer Acceleration

- If you transfer data from one region to another region which is much far away, it may take a long time
- S3 transfer acceleration enables fast, easy and secure transfer of all files over long distance b/w your client and S3 bucket

Networking Services

**VPC (Virtual Private Cloud)**

- It's a virtual network
- If you include all your AWS resources that have launched inside one VPC then all these resources become visible to each other or can interact with each other
- When you have a private data center and you are using AWS infrastructure as well and you want your AWS resources to be used as if they were on your own network, in that case you will establish a virtual private network connection your VPC in which you include all your services that you want in your private network

Direct Connect

- Replacement of internet connection
- It's a leased line using which you can directly connect to the AWS infrastructure

Route 53

- It's a Domain Name System
- DNS converts domain name to IP address
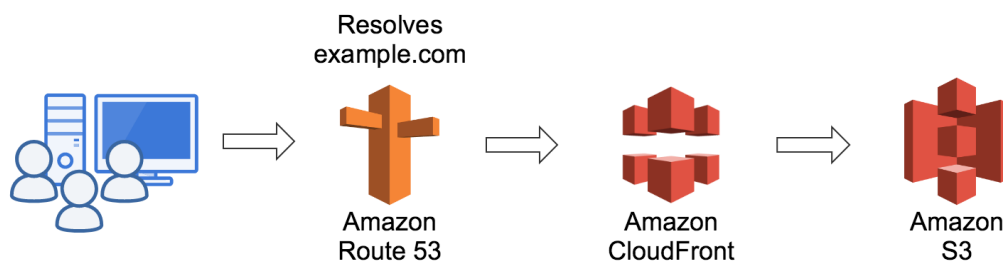
AWS CloudFront

- A web service which speeds up distribution of your data both static and dynamic
- CloudFront focus on 3 points
  - Routing
  - Edge locations
  - More availability
- By caching it delivers data more frequently
- CloudFront uses edge locations for storing cached data, if you want to fetch the data which is not available at the cached location then it request to the original server for the data

Monitoring and Management

CloudWatch

- Performance
- Cost
- Detecting errors
- Monitors client experience
- CloudWatch is a component of AWS that provides real time monitoring of AWS resources and computer applications running on AWS infrastructure
- Collect, monitor and share logs
- Create alarm and send notifications
- Send system events from AWS resources to AWS lambda, SNS etc

CloudWatch Logs

Used to monitor, access and store logs files from AWS resources like Amazon EC2 instance, Route 53 etc

Cloud Formation

- A computer application on AWS can have many resources and managing all these resources can be a difficult task
- By using AWS cloud formation, you can manage all these resources at a single place
- Cloud formation is a service that helps you to setup your AWS resources so that you can spend less time on managing those resources and spending more time on focusing on your application
- For using cloud formation you need a JSON script
- Lamp Stalk is a sample template created by cloud formation

Auto Scaling & Load Balancer

Snapshots

- Not a bootable copy. It only captures data that has been written to your Amazon EBS volume at the time the snapshot command is issued

AMI

- Bootable copy
- It's an executable image of your existing EC2 instance
- Copy of data from one drive to another external drive is called snapshot. But if you boot from that external drive so that the whole OS comes up on Some other m/c then it becomes an AMI.

Auto Scaling

It analyzes the load that is coming in and it deploys the server a/c to that. Extra servers are deployed based on AMI's original server

Load Balancer

- A load balancer is a device which acts as reverse proxy and distributes the network traffic across a number of servers. Load balancer are used to increase capacity and reliability of the application

Classic LB	Application LB
Divide traffic equally	Divide traffic based on request made
Less intelligent	More intelligent
Less used	More used

Cloud Security

- Companies like LinkedIn, Sony, iCloud etc already faced cyber attack. Due to that cloud security is required
- Cloud security is the use of latest technology and security techniques to protect your data, applications and infrastructure
- We should have to upgrade our security time to time

Cloud Trail

- It's a logging service which can be used to log the history of API calls
- It can also identify which user from AWS console requested the particular service
- It also helps to identify the hacker

AWS IAM (Identity & Access Management)

- Used to provide granular permissions
- Secure access to applications running on EC2 environment by giving you a private file
- Free to use
- AWS IAM is a web service that helps you securely control access to AWS resources for your users.
  - Based on users
  - Based on groups
  - Based on roles
  - Based on policies

➔ Multifactor authentication work as same as 2 factor authentication

Database Service

AWS RedShift

- A data warehouse is a subject oriented, integrated, time variant and non volatile collection of organizations data
- Amazon redshift is a fast scalable data warehouse that makes it simple and cost effective for you to analyze all your data across your data warehouse
- Column oriented

DevOps on AWS

- Developer + Operator = DevOps
- A set of practices to reduce the time b/w doing a change to the system and reflect that change into the normal production, while ensuring high quality
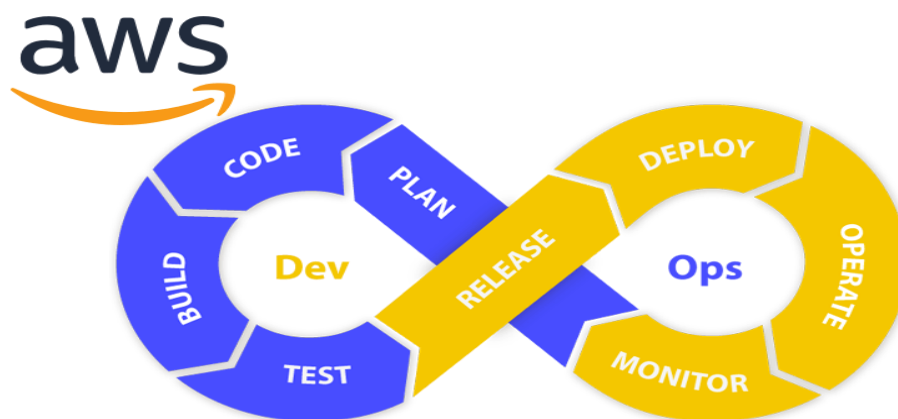

- A platform like AWS which is a popular cloud service provider, AWS ensures that all the requirement of DevOps can be taken care on the platform itself and there are various services that are made available to you
- Continuous integration and delivery does code change on latest code every time
- Source code → build server → test → production

AWS Code Pipeline

- It's a continuous delivery service, you can use to model visualize and automate the steps required to release your software

  Code commit → code build → code deploy
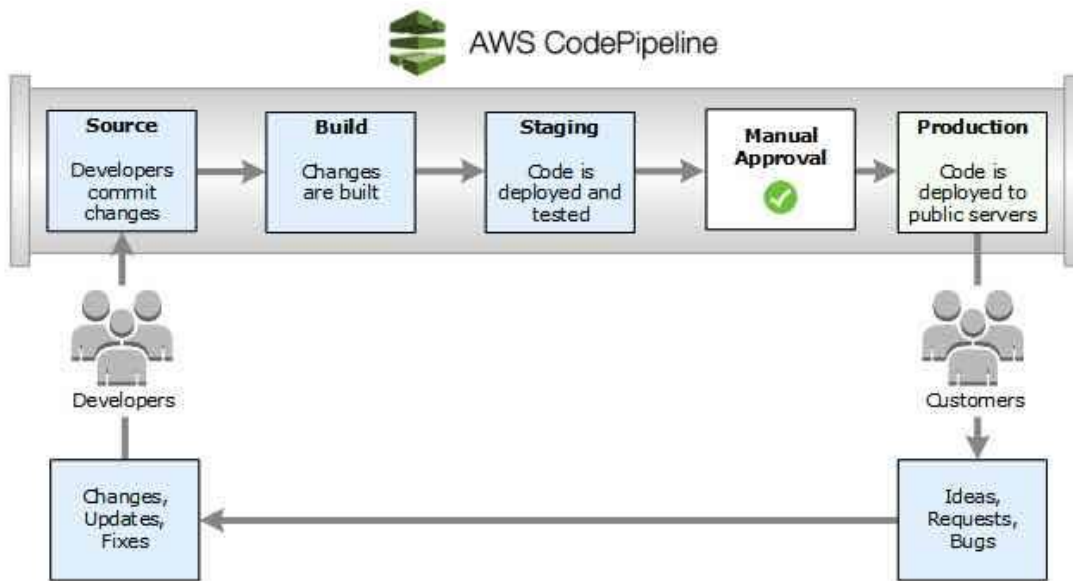
- When an instance is stopped, the instance performs a normal shutdown and then moved into stopped state. All of it's Amazon EBS volume remain attached and you can start the instance again
- When an instance is terminated, the instance performs a normal shutdown, then the attached EBS volume are deleted and the instance itself is also deleted, here you can't restart the instance again
- You are not charged, if only one Elastic IP address is attached with your running instance. But you do charge in following conditions:
  - If you use more than one Elastic IP's with your instance
  - When your Elastic IP's are attached to a stopped instance
  - When your Elastic IP is not attached to any instance