Amazon Web Service(AWS

Aws: AWS services is Amazon’s cloud web hosting platform that offers flexible, reliable, scalable, easy-to-use, and cost-effective solutions.

Cloud computing: is a term referring to storing and accessing data over the internet

There are 3 types of cloud services:

1. Private cloud: it is dedicated to a single tenant. It is dedicated in terms of hardware and security.
2. Public cloud: shared with multiple tenants and the cost is lesser. Ex: AWS, GCP (google), Oracle, Microsoft Azure, Alibaba Cloud, etc…
3. Hybrid cloud: hybrid is a combination of both public and private clouds. It is the most successful cloud practice. Example: Openstack & VMware

What is the reason behind the cloud? Answer: Virtualization

Virtualization will transfer hardware into software example: VMware.

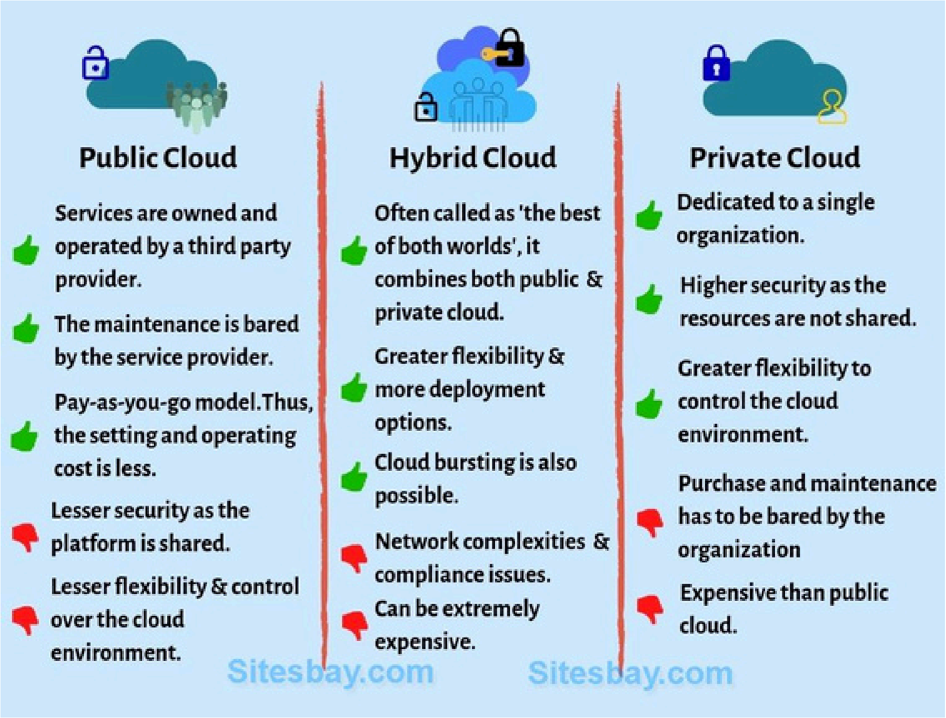
Private cloud: is dedicated to a single organization, highly secured, and Greater flexible.

Public cloud: Third-party provider makes resources and services available to the customer via

Internet

Advantages: cost-effective, reliable, unlimited storage, backup & recovery

Third-party providers: AWS, GCP (google), Oracle, Microsoft Azure, etc…



Public Cloud

👍 Services are owned and operated by a third party provider.

👍 The maintenance is bared by the service provider.

👍 Pay-as-you-go model.Thus, the setting and operating cost is less.

👎 Lesser security as the platform is shared.

👎 Lesser flexibility & control over the cloud environment.

Hybrid Cloud

👍 Often called as 'the best of both worlds', it combines both public & private cloud.

👍 Greater flexibility & more deployment options.

👍 Cloud bursting is also possible.

👎 Network complexities & compliance issues.

👎 Can be extremely expensive.

Private Cloud

👍 Dedicated to a single organization.

👍 Higher security as the resources are not shared.

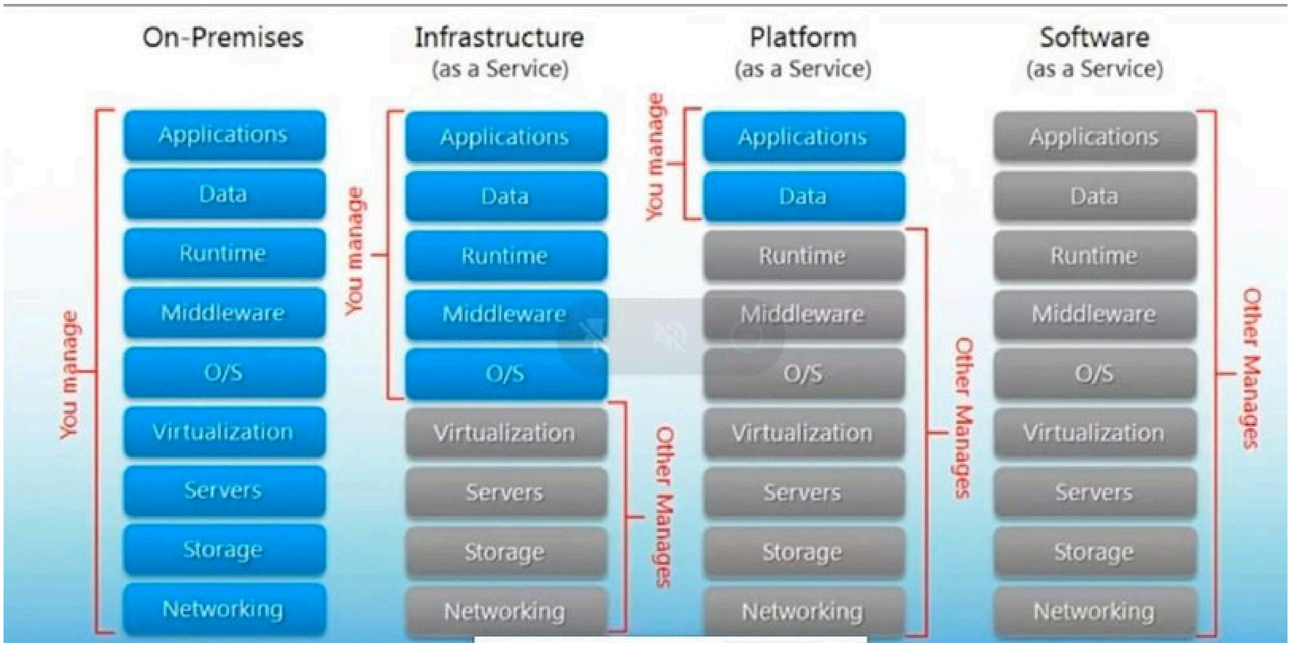
👍 Greater flexibility to control the cloud environment.

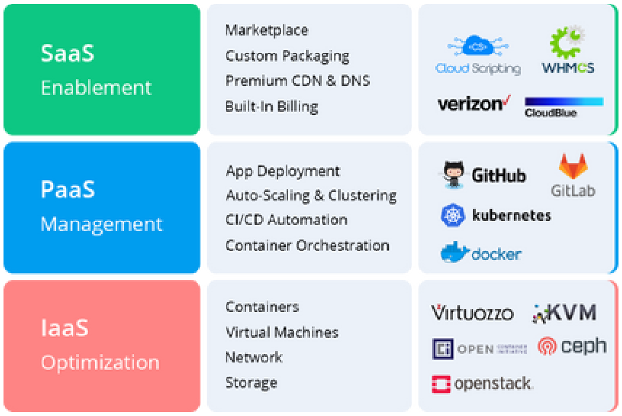
👎 Purchase and maintenance has to be bared by the organization

👎 Expensive than public cloud.

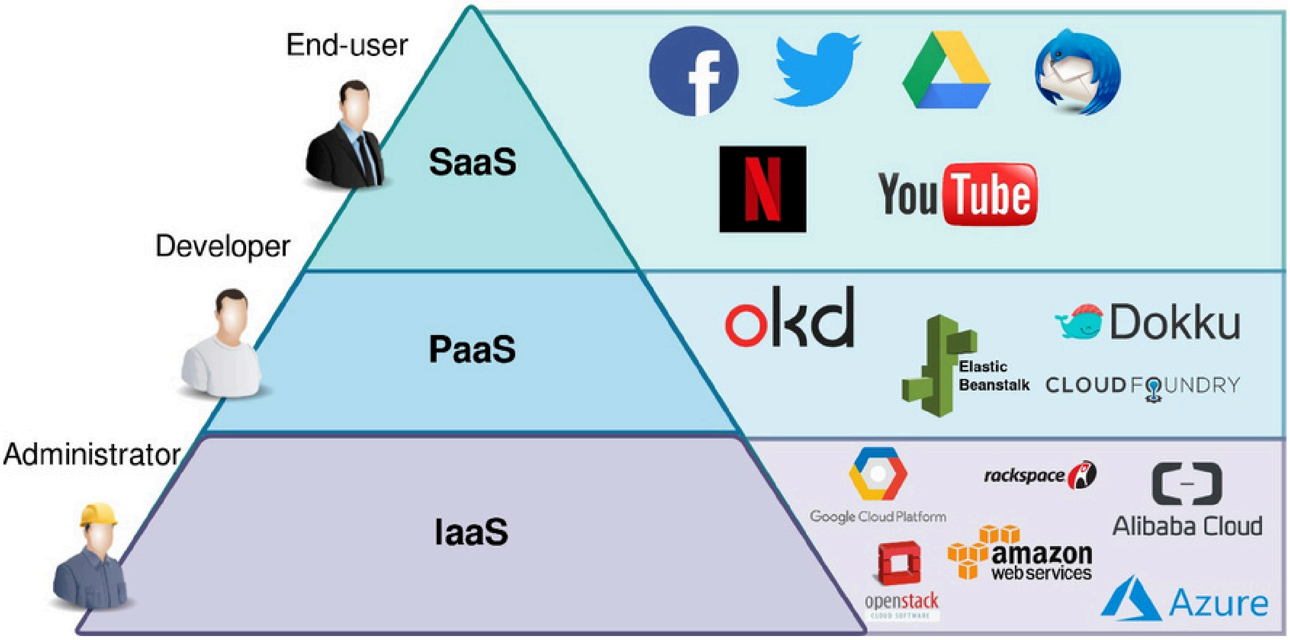
3 types of services

1. IaaS=Infra structure as a service
2. PaaS = Platform as a service
3. SaaS= Software as a service





SaaS Enablement	Marketplace Custom Packaging Premium CDN & DNS Built-In Billing	Cloud Scripting WHMCS verizon CloudBlue
PaaS Management	App Deployment Auto-Scaling & Clustering CI/CD Automation Container Orchestration	GitHub GitLab kubernetes docker
IaaS Optimization	Containers Virtual Machines Network Storage	Virtuozzo KVM OPEN ceph openstack



Advantages of IAS:

1. Shared infrastructure
2. Pay as per you use model
3. Focus on core business
4. On demand scalability

Disadvantage of IAS:

1. Security
2. Maintenance & upgrade

PAS

Ex: Google app engine, salesforce, windows azure etc

Advantages of PAS:

1. Simplified Development
2. Lower risk
3. Scalability

Disadvantages of PAS:

1. Vendor locking /flexibility
2. Integrating with rest of the applications

SAS

Example: Google, Microsoft office 365

Advantages of SAS:

1. Reduced time to benefit
2. Lower costs
3. Scalability and integration
4. Trouble-free Upgradation
5. Easy to use and perform proof-of-concepts

Disadvantages

1. Insufficient Data Security
2. Difficulty with Regulations Compliance
3. Cumbersome Data Mobility
4. Low Performance

Ip address:

Class A : 1.X.X.X TO 126.X.X.X

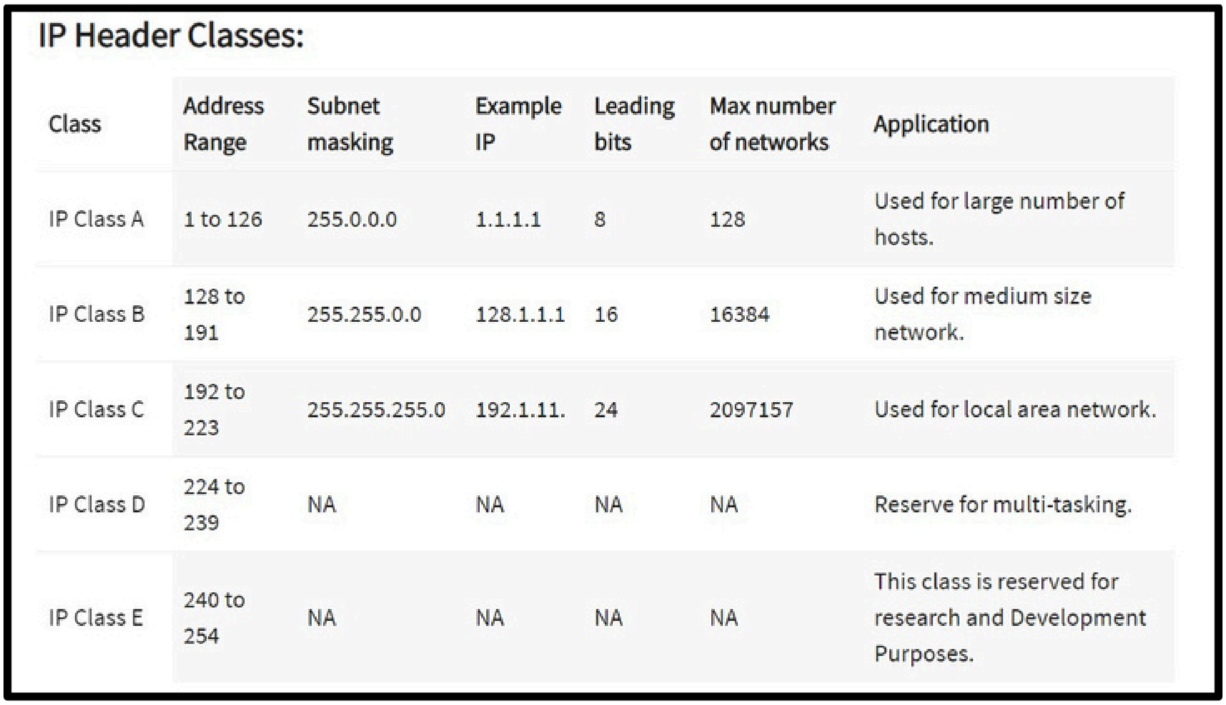
Class B : 128.X.X.X TO 191.X.X.X

Class C: 192.X.X.X TO 223.X.X.X

Class D: 224.X.X.X TO 239.X.X.X

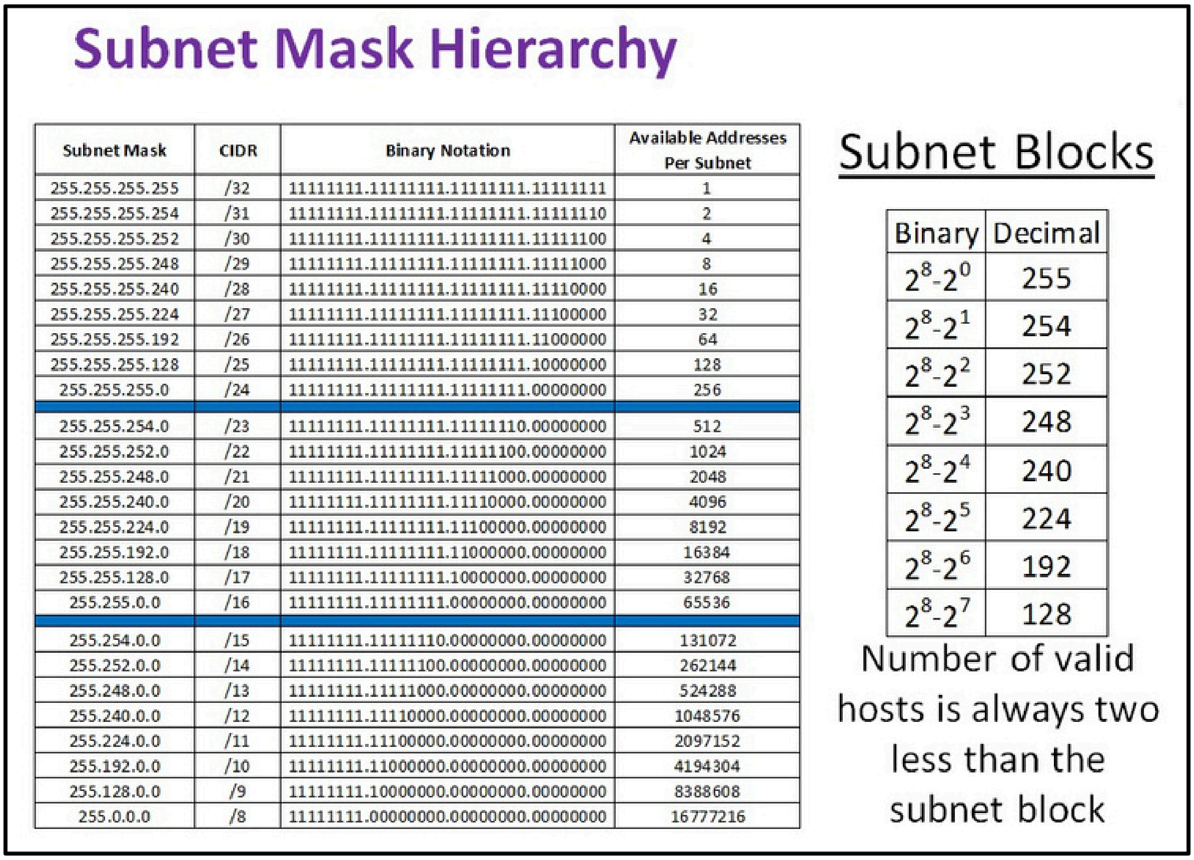
Class E: 240.X.X.X TO 254.X.X.X

127.0.0.0 is called loopback ip (it is reserved)



IP Header Classes:

Class	Address Range	Subnet masking	Example IP	Leading bits	Max number of networks	Application
IP Class A	1 to 126	255.0.0.0	1.1.1.1	8	128	Used for large number of hosts.
IP Class B	128 to 191	255.255.0.0	128.1.1.1	16	16384	Used for medium size network.
IP Class C	192 to 223	255.255.255.0	192.1.11.	24	2097157	Used for local area network.
IP Class D	224 to 239	NA	NA	NA	NA	Reserve for multi-tasking.
IP Class E	240 to 254	NA	NA	NA	NA	This class is reserved for research and Development Purposes.



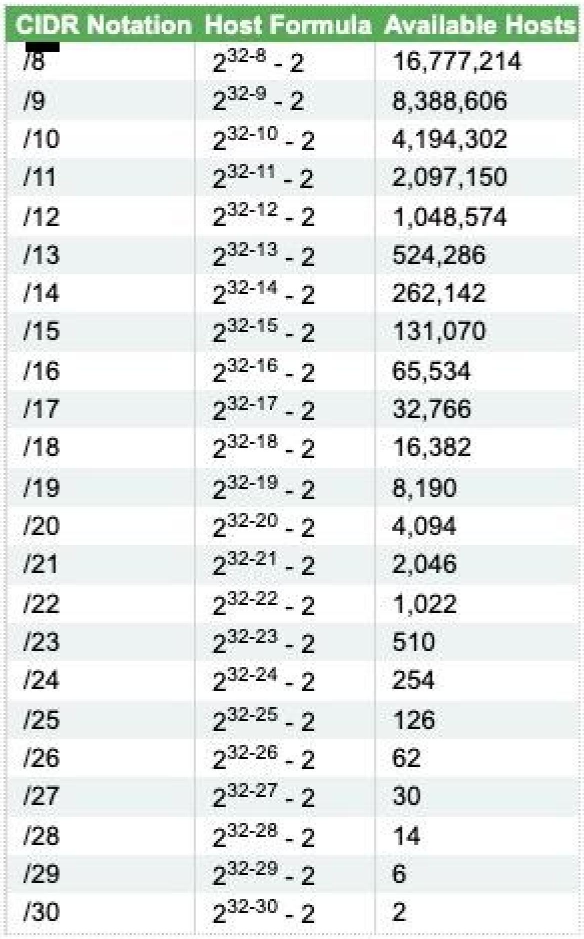
Subnet Mask Hierarchy

Subnet Mask	CIDR	Binary Notation	Available Addresses Per Subnet
255.255.255.255	/32	11111111.11111111.11111111.11111111	1
255.255.255.254	/31	11111111.11111111.11111111.11111110	2
255.255.255.252	/30	11111111.11111111.11111111.11111100	4
255.255.255.248	/29	11111111.11111111.11111111.11111000	8
255.255.255.240	/28	11111111.11111111.11111111.11110000	16
255.255.255.224	/27	11111111.11111111.11111111.11100000	32
255.255.255.192	/26	11111111.11111111.11111111.11000000	64
255.255.255.128	/25	11111111.11111111.11111111.10000000	128
255.255.255.0	/24	11111111.11111111.11111111.00000000	256
255.255.254.0	/23	11111111.11111111.11111110.00000000	512
255.255.252.0	/22	11111111.11111111.11111100.00000000	1024
255.255.248.0	/21	11111111.11111111.11111000.00000000	2048
255.255.240.0	/20	11111111.11111111.11110000.00000000	4096
255.255.224.0	/19	11111111.11111111.11100000.00000000	8192
255.255.192.0	/18	11111111.11111111.11000000.00000000	16384
255.255.128.0	/17	11111111.11111111.10000000.00000000	32768
255.255.0.0	/16	11111111.11111111.00000000.00000000	65536
255.254.0.0	/15	11111111.11111110.00000000.00000000	131072
255.252.0.0	/14	11111111.11111100.00000000.00000000	262144
255.248.0.0	/13	11111111.11111000.00000000.00000000	524288
255.240.0.0	/12	11111111.11110000.00000000.00000000	1048576
255.224.0.0	/11	11111111.11100000.00000000.00000000	2097152
255.192.0.0	/10	11111111.11000000.00000000.00000000	4194304
255.128.0.0	/9	11111111.10000000.00000000.00000000	8388608
255.0.0.0	/8	11111111.00000000.00000000.00000000	16777216

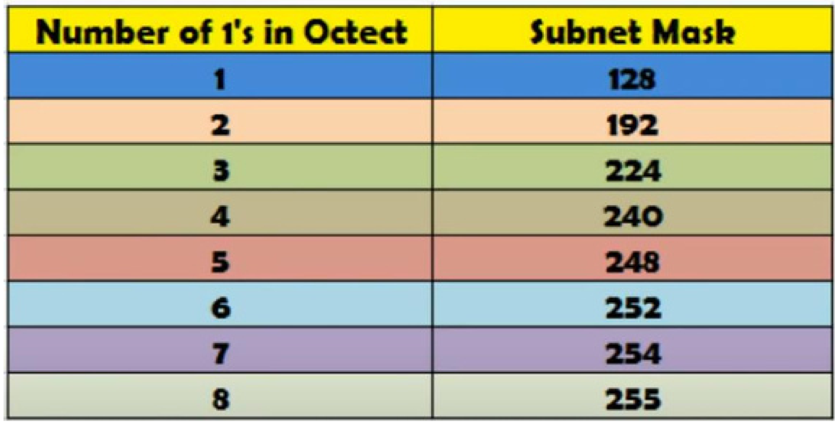
Subnet Blocks

Binary	Decimal
2^8-2^0	255
2^8-2^1	254
2^8-2^2	252
2^8-2^3	248
2^8-2^4	240
2^8-2^5	224
2^8-2^6	192
2^8-2^7	128

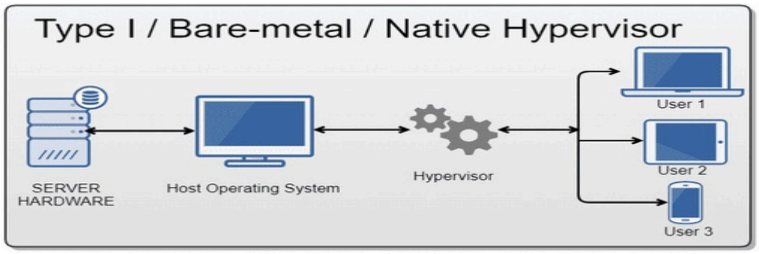
Number of valid hosts is always two less than the subnet block



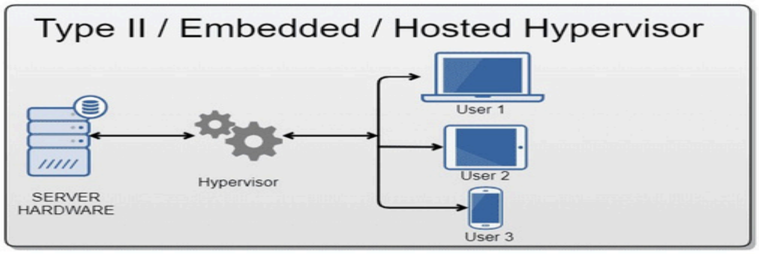
CIDR Notation	Host Formula	Available Hosts
/8	$2^{32-8} - 2$	16,777,214
/9	$2^{32-9} - 2$	8,388,606
/10	$2^{32-10} - 2$	4,194,302
/11	$2^{32-11} - 2$	2,097,150
/12	$2^{32-12} - 2$	1,048,574
/13	$2^{32-13} - 2$	524,286
/14	$2^{32-14} - 2$	262,142
/15	$2^{32-15} - 2$	131,070
/16	$2^{32-16} - 2$	65,534
/17	$2^{32-17} - 2$	32,766
/18	$2^{32-18} - 2$	16,382
/19	$2^{32-19} - 2$	8,190
/20	$2^{32-20} - 2$	4,094
/21	$2^{32-21} - 2$	2,046
/22	$2^{32-22} - 2$	1,022
/23	$2^{32-23} - 2$	510
/24	$2^{32-24} - 2$	254
/25	$2^{32-25} - 2$	126
/26	$2^{32-26} - 2$	62
/27	$2^{32-27} - 2$	30
/28	$2^{32-28} - 2$	14
/29	$2^{32-29} - 2$	6
/30	$2^{32-30} - 2$	2



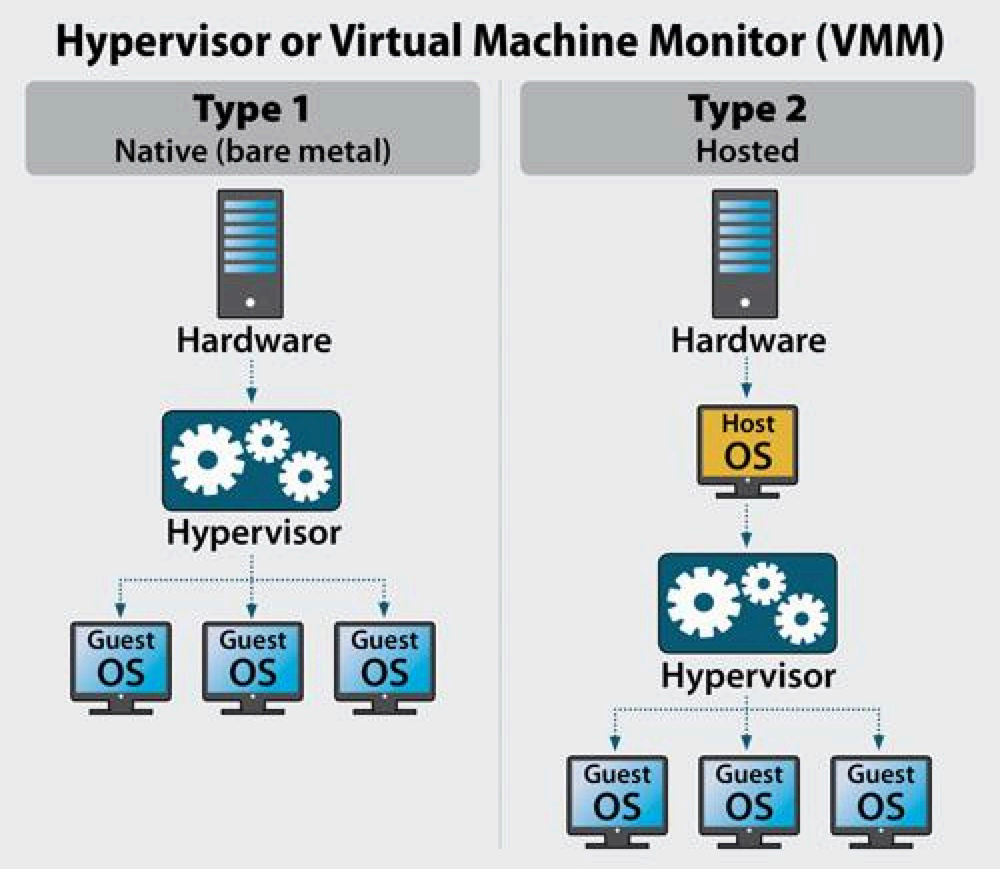
Number of 1's in Octect	Subnet Mask
1	128
2	192
3	224
4	240
5	248
6	252
7	254
8	255



Type I / Bare-metal / Native Hypervisor



Type II / Embedded / Hosted Hypervisor



1) **What is VPC (virtual private cloud)?**

It is a virtual network dedicated to your AWS account, it logically isolates from other virtual networks in the AWS cloud, and where you can launch your AWS instance.
VPC consists of subnets, network gateway & Routing table.

What is a subnet?

The subnet is a logical subdivision of the IP network. The practice of dividing a network into 2 or more networks is called subnetting.

2) **What is a Route table?**

A set of rules called routes that are used to determine where network traffic is directed.

3) **What is an Internet gateway?**

A gateway that you attach to your VPC to enable communication between resources in your VPC and the internet.

4) **What is a NAT gateway?**

NAT Gateway is used to connect to the Internet from instances within a private subnet in the VPC.
Nat gateway will be created from the public subnet and attached to the private subnet.

5) **What is the VPC endpoint?**

It enables you to privately connect your VPC to supported AWS services and VPC endpoint services powered by private links without requiring an internet gateway, NAT device, VPN connection, or AWS direct connection.

6) What is VPC peering connection?

A VPC peering connection is a network connection between 2 VPC that enables you to route traffic between them using a private IPV4 or IPV6 address.
Instances in either VPC can communicate with each other as if they are within the same network.
You can create a VPC peering connection between your own VPC or with a VPC in another AWS account.

7) Limitations of VPC peering.

I. You cannot create a VPC peering connection between VPCs that have matching or overlapping IPV4 CIDR blocks. Amazon always assigns a unique IPV6 CIDR block.
II. You have a quota on the number of active and pending VPC peering connections that you can have per VPC.
III. VPC peering does not support transitive peering relationships. In VPC peering connection, one VPC does not have access to any other VPC with which the peer VPC may have peered.
IV. Cannot have more than 1 VPC peering connection between the same 2 VPCs at the same time.

8) What is NACL (Network ACL)? Network Access Control Lists

NACLs are firewalls at the subnet level.
You can use a network ACL to control the traffic to and from the subnet in which the NAT gateway is located.

9) What is a security group?

A security group acts as a virtual firewall for your instance to control inbound and outbound traffic.
When you launch an instance in a VPC, you can assign up to 5 security groups to the instance.
Security group act at the instance level and not at the subnet level.

10) Difference between public subnet and private subnet

Public Subnet – Users can access resources from the internet. Internet traffic is routed via an internet gateway. Applications are stored in a public subnet.

Private Subnet - Users cannot access resources from the internet. Internet traffic is routed via the NAT gateway. Data is stored in a private subnet (database, API calls, passwords)

AWS has 25 regions

80 availability zones(az)

230+ edge locations (network boosters/content delivery networks)

12 edge caches => regional cache

AMI => Amazon Machine Image : Master image for the creation of virtual servers called EC2.

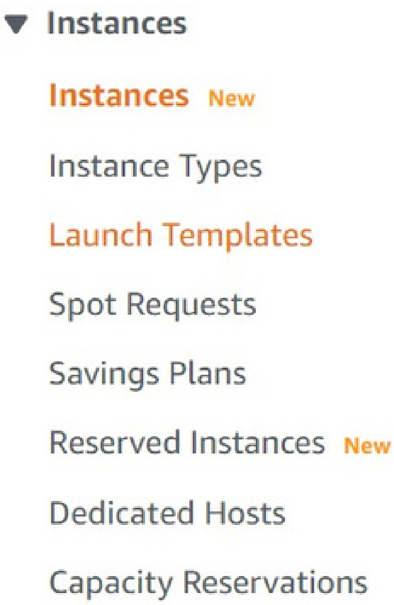
AWS Instance classes :

1) General Purpose => It is used for all kind of basic testing and lower environments. Ex: T2, T3, M6
2) Compute Optimized=> It is used for all kind of gaming servers and most of the machine learning servers. Ex:C6 C5.
3) Memory Optimized => It is used in terms of faster perf and ex: R6 series.
4) Accelerated Computing => high graphic processors. Basically used in handling gaming s/w. Ex: P3, P2 series.
5) Storage Optimized=> used in terms of handling heavy workloads and also for archiving huge data. Ex: L series

Security groups : firewall at a instances level. Are statefull. we can define inbound and outbound rules.

NACL: firewall at a subnet level. Stateless. Define only inbound rules.

Configuration of EC2 Models:



Types of EC2 payment models:

1. On demand or capacity reservation:
   a. These instances work as pay as you go model.
   b. Long time commitment is not required for this instance.
   c. These are bit costlier
2. Spot instances:
   a. These are called as bidding instances.
   b. We can bid these instances as per the requirement.
3. Dedicated hosts:
   a. Basically, dedicated instances are provided with the hardware configurations.
4. Reserved instances:
   a. These are utilized for the longer time frame.
   b. These would be having discounts in terms of pricing because of longer utilization time frame.
   c. Will get to know platform, tenancy, instance type and payment options for reserved instances.

AWS Storage:

1. Elastic Block Storage (EBS)
2. Elastic file storage (EFS)
3. Simple Storage Service (S3)

1. EFS (Elastic File System):

It is like a shared disk, EFS is more used in case of sharing the disk space. EFS storage is used in cluster management to have availability.
Storage space sharing is possible in EFS and not possible in EBS
Eg – Cassandra cluster, Kubernetes, Machine learning, and AWS lambda.

Advantages of EFS:

cost-effective, Speed, Disk share.

2. **EBS (Elastic Block Storage):**

   EBS provides simple, scalable, high-available block storage. EBS can only attach to one EC2 system. It is block storage.
   Storage space sharing is not possible in EBS

Important 3 models of EBS:

1. Provision IOPS (64000)
2. General purpose (16000)
3. Magnetic (5000)

Note: EBS can be attached to only one instance, EBS and Volume should be there in the same region to avail of the EBS

Benefits of EBS:

1. SSD storage technology (solid-state drive)
2. Highly available, fast, and scalable

3. **S3 (Simple Storage Service) :**

   It is basically object storage. S3 cannot hold data, but it can store & hold data that are in the form of objects. S3 is not region specific. We can host a static website on S3.
   Types of S3 bucket -

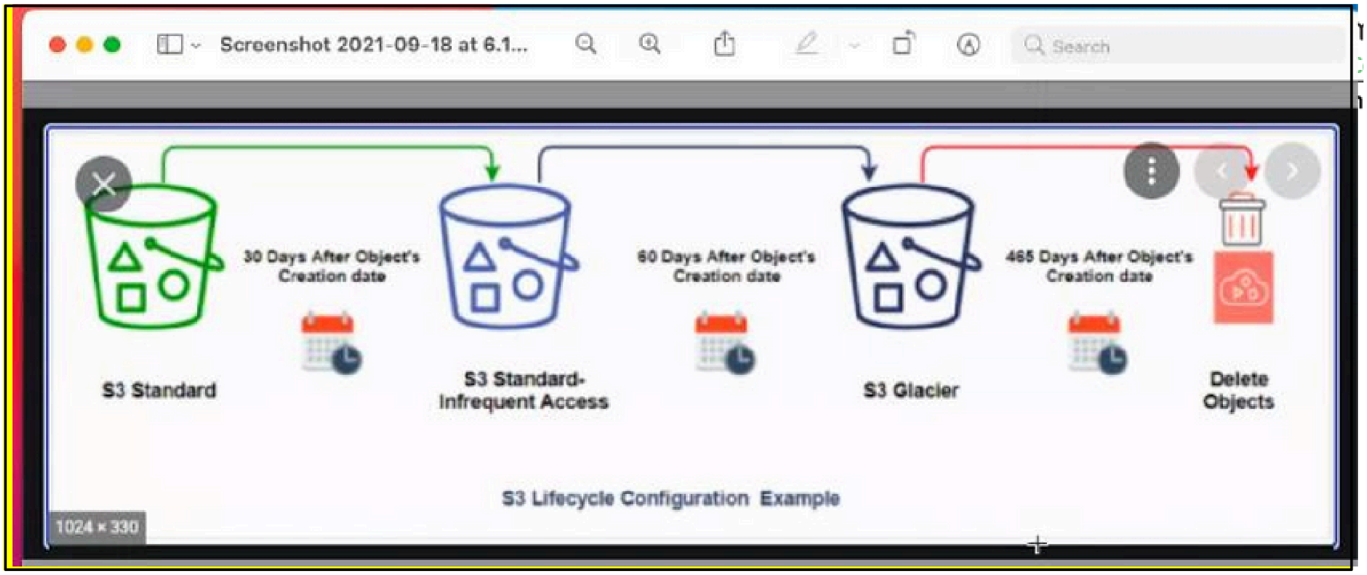
   a) S3 Standard IA
   b) S3 Standard
   c) S3 intelligent tiering
   d) S3 Glacier
   e) S3 Deep archive

4. Key Features of S3:

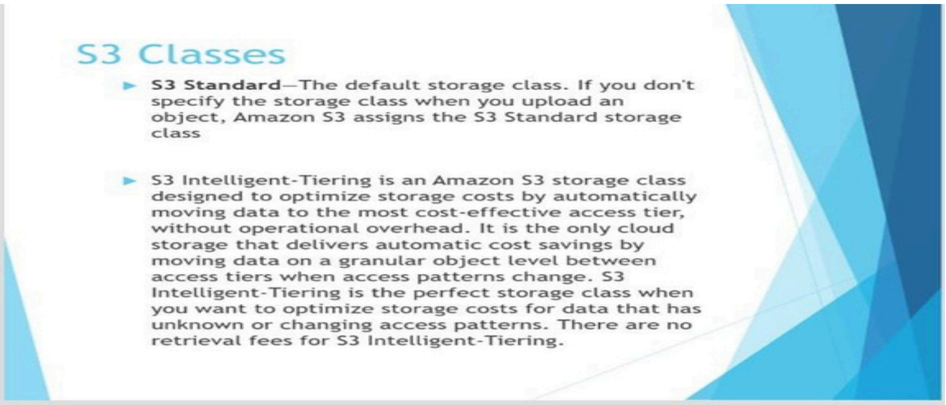
   a. Versioning – AWS S3 is a means of keeping multiple variants of an object in the same bucket, you can use the S3 versioning feature to preserve, retrieve and restore every version of every object stored in your bucket. Versioning-enabled buckets can help you recover objects from accidental deletion or overwrite.
   b. Life cycle management.
   c. Encryption
   d. Multifactor authentication for object deletion.

5. S3 bucket Security:

   There are two types of bucket security:
   a. Bucket Policies – Json based scripts which are embedded in IAM policies of AWS which can be utilized for S3 bucket security.
   b. Access control List.

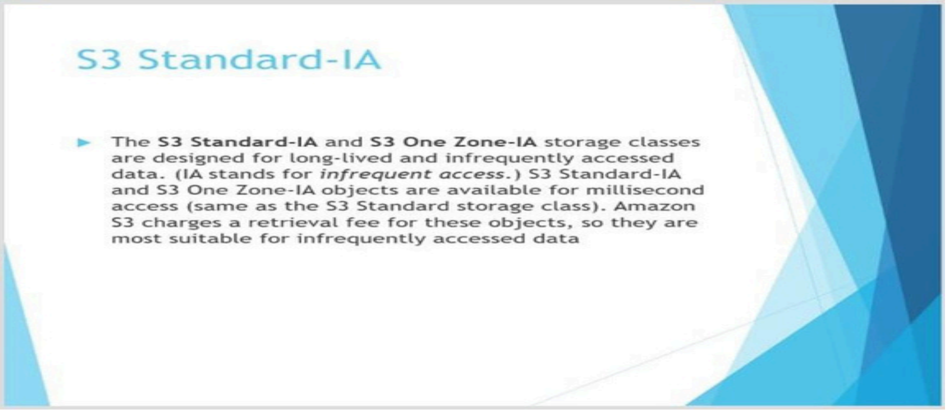

Screenshot 2021-09-18 at 6.1...
Search
S3 Standard
30 Days After Object's Creation date
S3 Standard-Infrequent Access
60 Days After Object's Creation date
S3 Glacier
465 Days After Object's Creation date
Delete Objects
1024 × 330

S3 Lifecycle Configuration Example



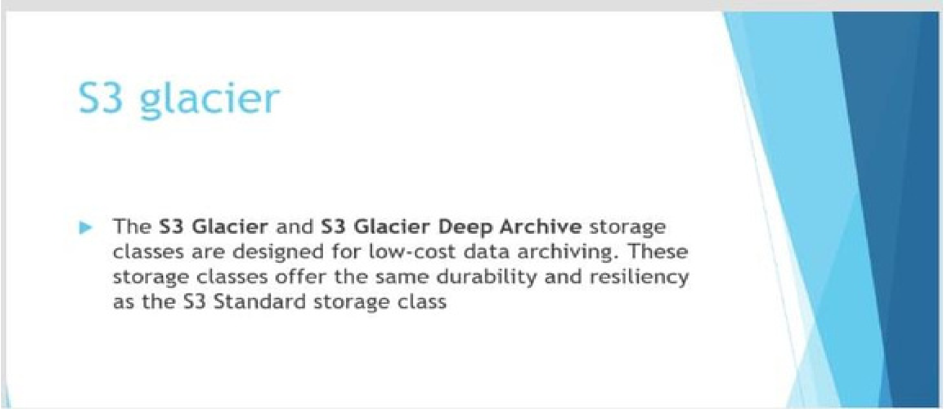
S3 Classes

- **S3 Standard**—The default storage class. If you don't specify the storage class when you upload an object, Amazon S3 assigns the S3 Standard storage class
- S3 Intelligent-Tiering is an Amazon S3 storage class designed to optimize storage costs by automatically moving data to the most cost-effective access tier, without operational overhead. It is the only cloud storage that delivers automatic cost savings by moving data on a granular object level between access tiers when access patterns change. S3 Intelligent-Tiering is the perfect storage class when you want to optimize storage costs for data that has unknown or changing access patterns. There are no retrieval fees for S3 Intelligent-Tiering.



S3 Standard-IA

- The **S3 Standard-IA** and **S3 One Zone-IA** storage classes are designed for long-lived and infrequently accessed data. (IA stands for *infrequent access.*) S3 Standard-IA and S3 One Zone-IA objects are available for millisecond access (same as the S3 Standard storage class). Amazon S3 charges a retrieval fee for these objects, so they are most suitable for infrequently accessed data



S3 glacier

- The **S3 Glacier** and **S3 Glacier Deep Archive** storage classes are designed for low-cost data archiving. These storage classes offer the same durability and resiliency as the S3 Standard storage class

What is partitions? How many AWS partitions are there? **

A Partition is a group of AWS Region and Service objects.

You can use a partition to determine what services are available in a region, or what regions a service is available in.

AWS accounts are scoped to a single partition. You can get a partition by name. Valid partition names include:

1."aws" - Public AWS partition

2. "aws-cn" - AWS China

3. "aws-us-gov" - AWS GovCloud

4. "AWS-ISO,

5. "AWS-ISO-b"

The last two are only for Secret and Top-Secret US Government data.

Define Auto-scaling.

Auto-scaling is an activity that lets you dispatch advanced instances on demand.

Moreover, auto-scaling helps you to increase or decrease resource capacity according to the application.

Can you illustrate the relationship between an instance and AMI?

With the help of just a single AMI, you can launch multiple instances and to even different types.

At the same time, an instance type is characterized by the host

What is a default storage class in S3?

The standard frequency accessed is the default storage class in S3.

What is the standard size of an S3 bucket?

The maximum size of an S3 bucket is five terabytes.

Is Amazon S3 an international/Global service?

Yes. Amazon S3 is an international service.

Its main objective is to provide an object storage facility through the web interface,

and it utilizes the Amazon scalable storage infrastructure to function in its global network.

Can you name some AWS services that are not region-specific?

o IAM

o Route 53

● S3

o Web application firewall

o CloudFront

Can you define EIP?

EIP stands for Elastic IP address.

It is a static Ipv4 address that is provided by AWS to administer dynamic cloud computing services.

IAM (Identity Access Management):

IAM allows you to manage users and their level of access to the AWS console.

Key components of IAM

1) Users – Users are end users within an organization.

eg – developers, testers, and infrastructure ppl.

2) User group – User groups are collections of users, each user in the group will inherit the permission of the group.

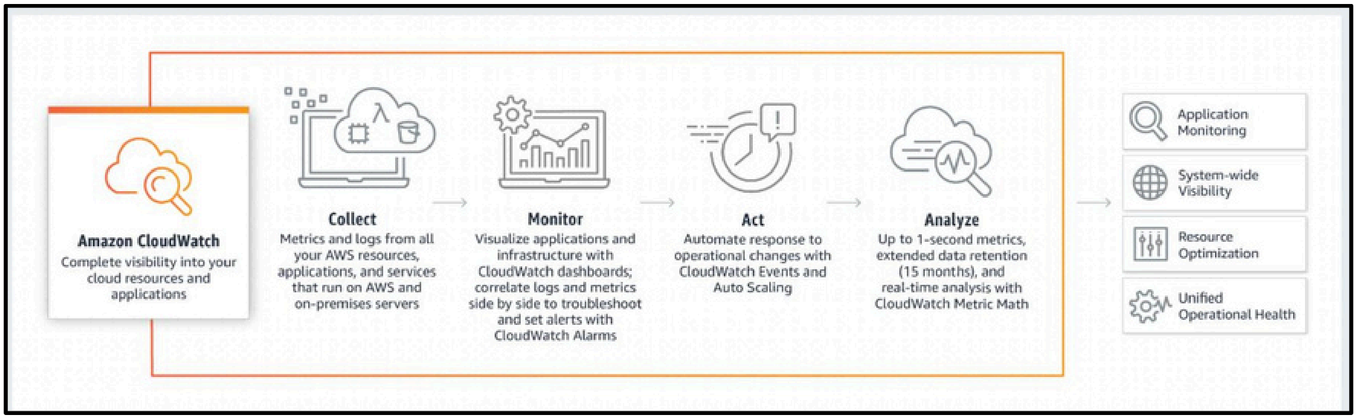
3) Policies – Policies are made up of documents called policy The document, these documents are in the format of JSON, and they give permission as to what a user group or a role can do.

4) IAM role – It is an IAM entity that defines a set of Permission-making AWS service requests. It is not associated with specific users or groups.

Advantages of IAM

- It provides centralized control of your AWS account.
- Shared access to your AWS account.
- Multi-factor authentication.
- Identity federation

Cloud watch:


Amazon CloudWatch
Complete visibility into your cloud resources and applications
Collect
Metrics and logs from all your AWS resources, applications, and services that run on AWS and on-premises servers
Monitor
Visualize applications and infrastructure with CloudWatch dashboards; correlate logs and metrics side by side to troubleshoot and set alerts with CloudWatch Alarms
Act
Automate response to operational changes with CloudWatch Events and Auto Scaling
Analyze
Up to 1-second metrics, extended data retention (15 months), and real-time analysis with CloudWatch Metric Math
Application Monitoring
System-wide Visibility
Resource Optimization
Unified Operational Health

Amazon cloud watch is a monitoring and observability service built for all the application team members. Cloud watch collects monitoring and operational data in the form of logs, matrices, and events. Cloud watch is useful in setting up alarms, visualizing logs, and matrix side by side.

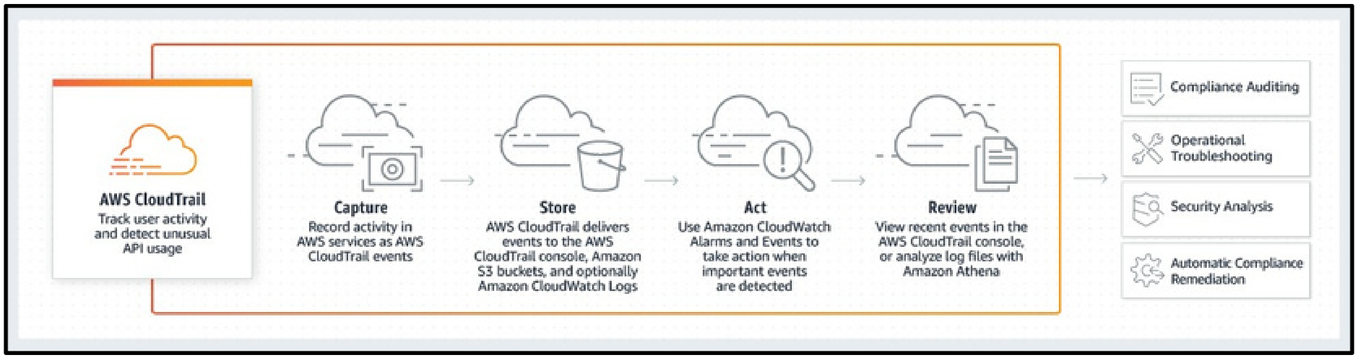
SNS (simple notification service):

Amazon SNS is a fully managed messaging service for both application-to-person and application-to-application communication.

Use cases:

1. Send messages directly to millions of users
2. Reliably deliver messages
3. Automatically scaling workload.

Clod Trail:


AWS CloudTrail
Track user activity and detect unusual API usage
Capture
Record activity in AWS services as AWS CloudTrail events
Store
AWS CloudTrail delivers events to the AWS CloudTrail console, Amazon S3 buckets, and optionally Amazon CloudWatch Logs
Act
Use Amazon CloudWatch Alarms and Events to take action when important events are detected
Review
View recent events in the AWS CloudTrail console, or analyze log files with Amazon Athena
Compliance Auditing
Operational Troubleshooting
Security Analysis
Automatic Compliance Remediation

AWS CloudTrail is a service that enables governance, compliance, operational auditing, and risk auditing of your AWS account.

Cloud trail will store in S3 bucket.

Benefits:

1. Simplified compliance
2. Visibility into user and resource activity
3. Security analysis and troubleshooting
4. Security automation

Amazon RDS -

It is a relational database, RDS is fully managed with fast and creditable performance.
RDS is simple and scalable.
RDS is low-cost and pays for what we use.
Eg – MySql, Postgrace SQL, MariaDB, Oracle, Amazon Aurora

Amazon Aurora -

Is RDS reinvented for cloud, Aurora is 5 times better performance than MySql.
Aurora is available at 1/10 the cost of commercial db.

RDS -

It is easy to administer, RDS is highly scalable.
RDS is available & durable.
RDS provides a feature called a ready replica.
Ready Replica – Amazon RDS synchronously replicates the data to a standby instance in a different availability zone.

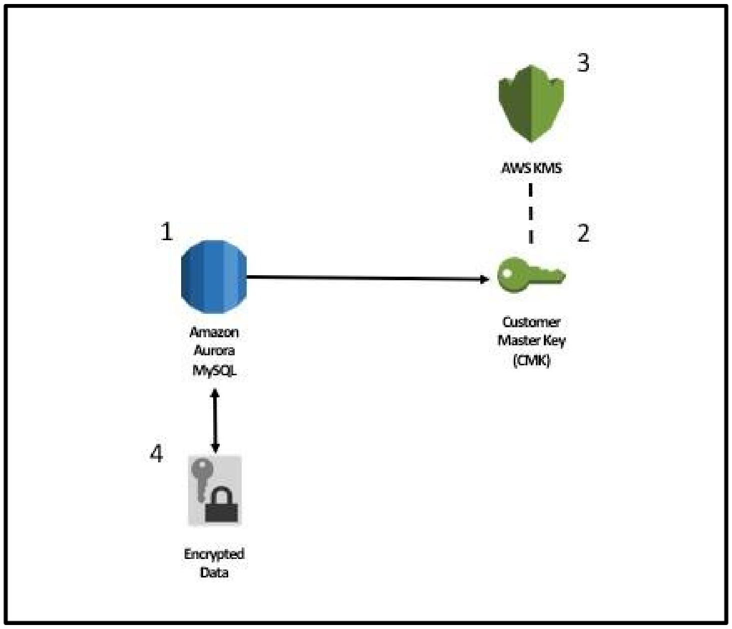
Amazon RDS supports the most demanding applications and is fast.

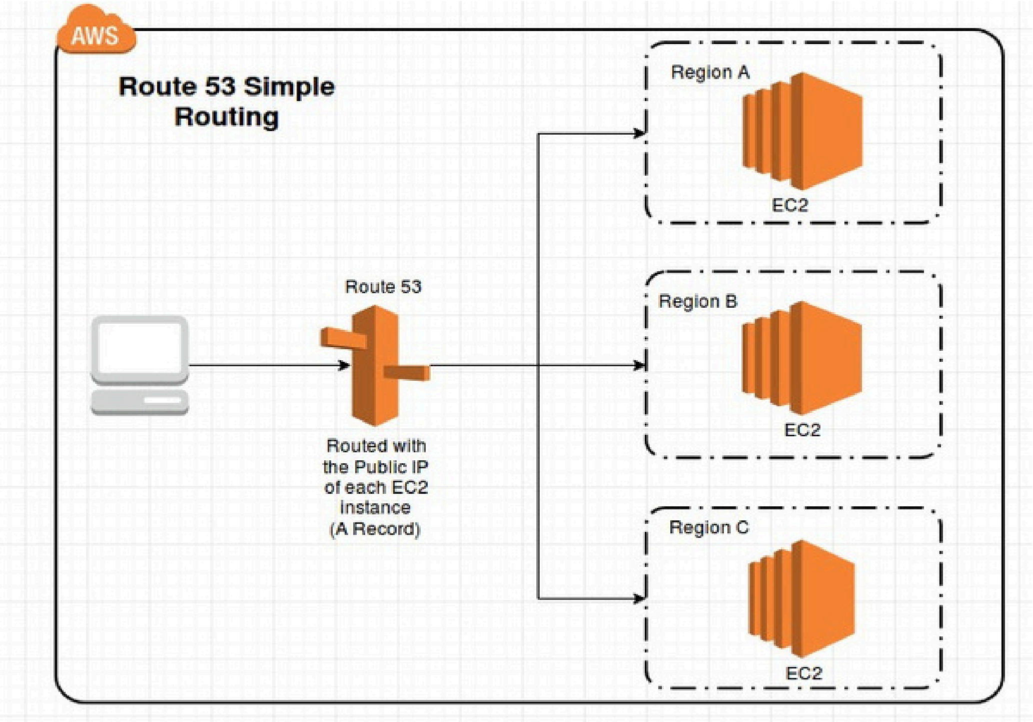
It is made easy to control n/w access to your DB.
Amazon RDS lets you run your database in your instance in VPC.
It isolates the DB and makes it secure.

AWS KMS (Key Management Service): AWS KMS makes it easy for us to create and manage cryptographic keys and control their use across a wide range of use and their applications.

Benefits:

1. Fully managed: you control access to your encrypted data by defining permission to use keys while aws
2. Centralised Key management:
3. Manage encryption for AWS services.
4. Low cost



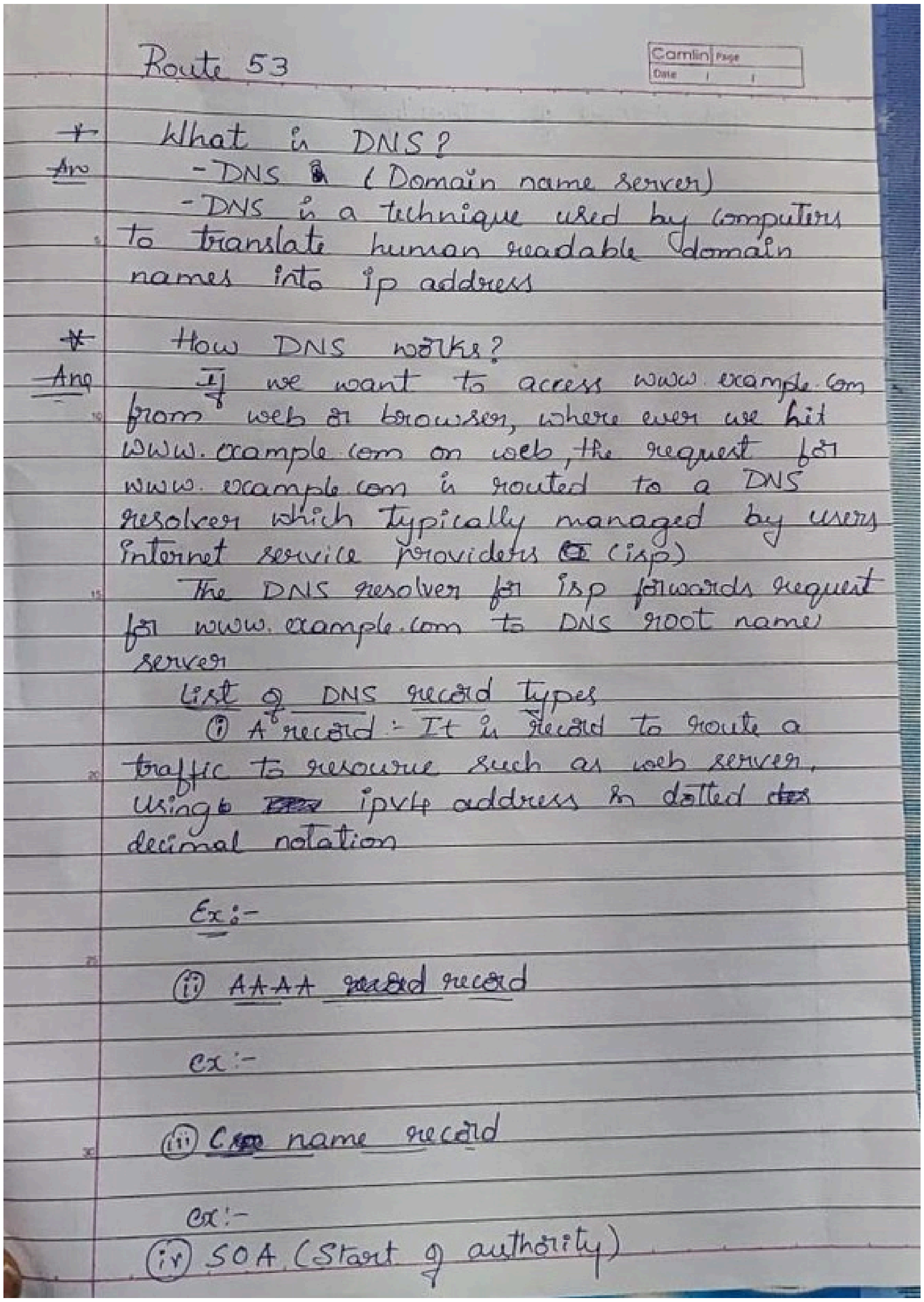


Route 53:

What is DNS (domain name server): it is a technique used by computers to translate human-readable domain names into IP addresses.

How does it work? Let's take an example, we want to access www.example.com from web or a browser whenever we hit www.example.com on the web/browser/address bar the request for www.example.com is routed to a DNS resolver which is typically managed by the user's internet service provider.

The DNS resolver for



Route 53

Camlin Page
Date / /

* What is DNS?

Ans - DNS (Domain name server)

- DNS is a technique used by computers to translate human readable domain names into ip address

* How DNS works?

Ans If we want to access www.example.com from web or browser, where ever we hit www.example.com on web, the request for www.example.com is routed to a DNS resolver which typically managed by users Internet service providers (isp)

The DNS resolver for isp forwards request for www.example.com to DNS root name server

List of DNS record types

(i) A record :- It is record to route a traffic to resource such as web server, using ipv4 address in dotted ~~dec~~ decimal notation

Ex:-

(ii) AAAA ~~record~~ record

ex:-

(iii) C name record

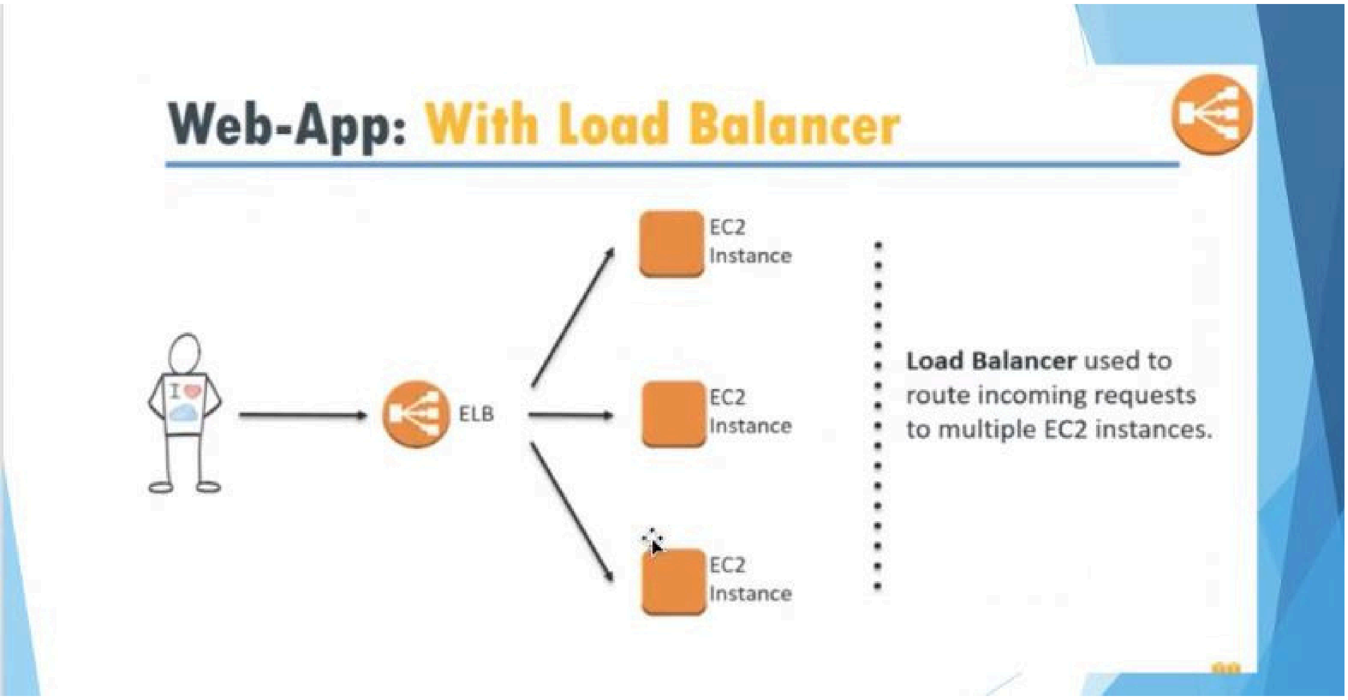
ex:-

(iv) SOA (Start of authority)

You use an A record to route traffic to a resource, such as a web server, using an IPv4 address in dotted decimal notation.
192.0.2.1
AAAA record type You use an AAAA record to route traffic to a resource, such as a web server, using an IPv6 address in colon-separated hexadecimal format.
2001:0db8:85a3:0:0:8a2e: 0370:7334

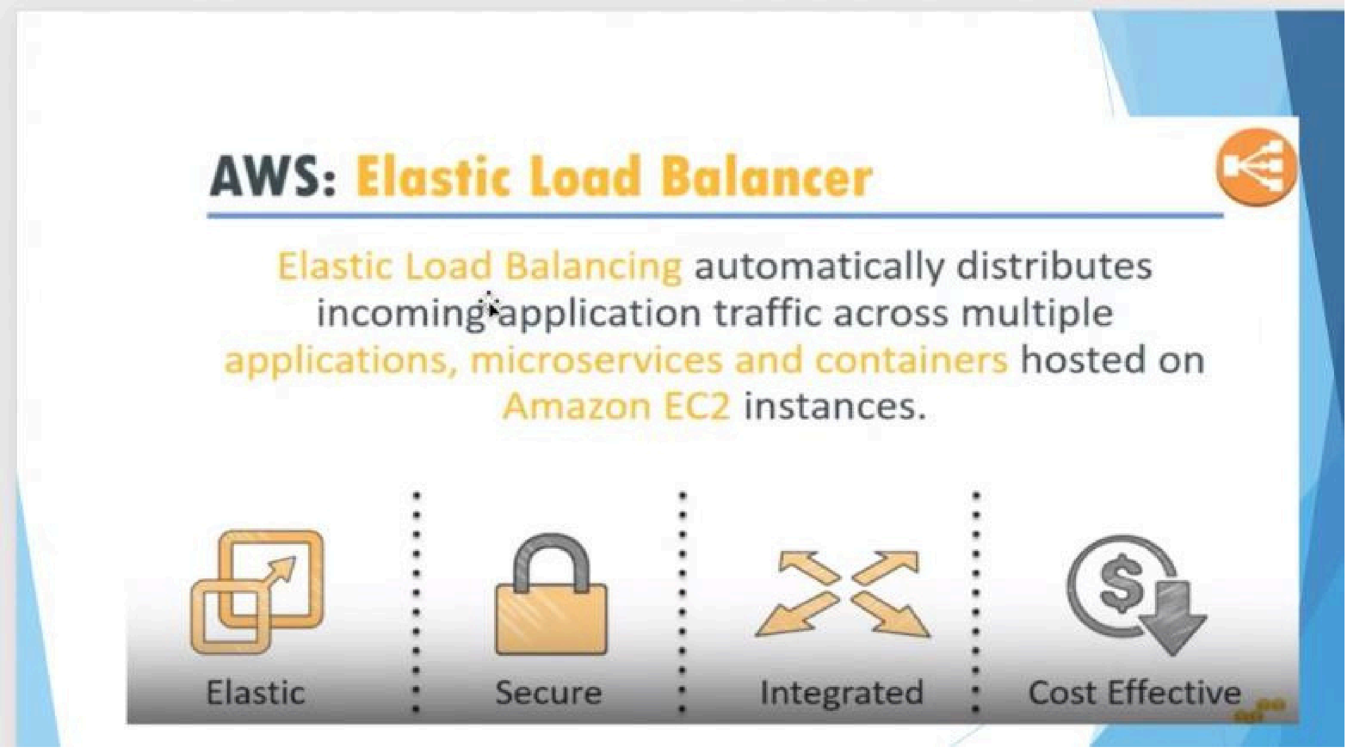
A CNAME record maps DNS queries for the name of the current record, such as acme.example.com, to another domain (example.com or example.net) or subdomain (acme.example.com or zenith.example.org).
hostname.example.com
SOA record type

A start of authority (SOA) record provides information about a domain and the corresponding Amazon Route 53 hosted zone. For information about the fields in an SOA record,
example: ns-2048.awsdns-64.net hostmaster.awsdns.com 1 1 1 1 60



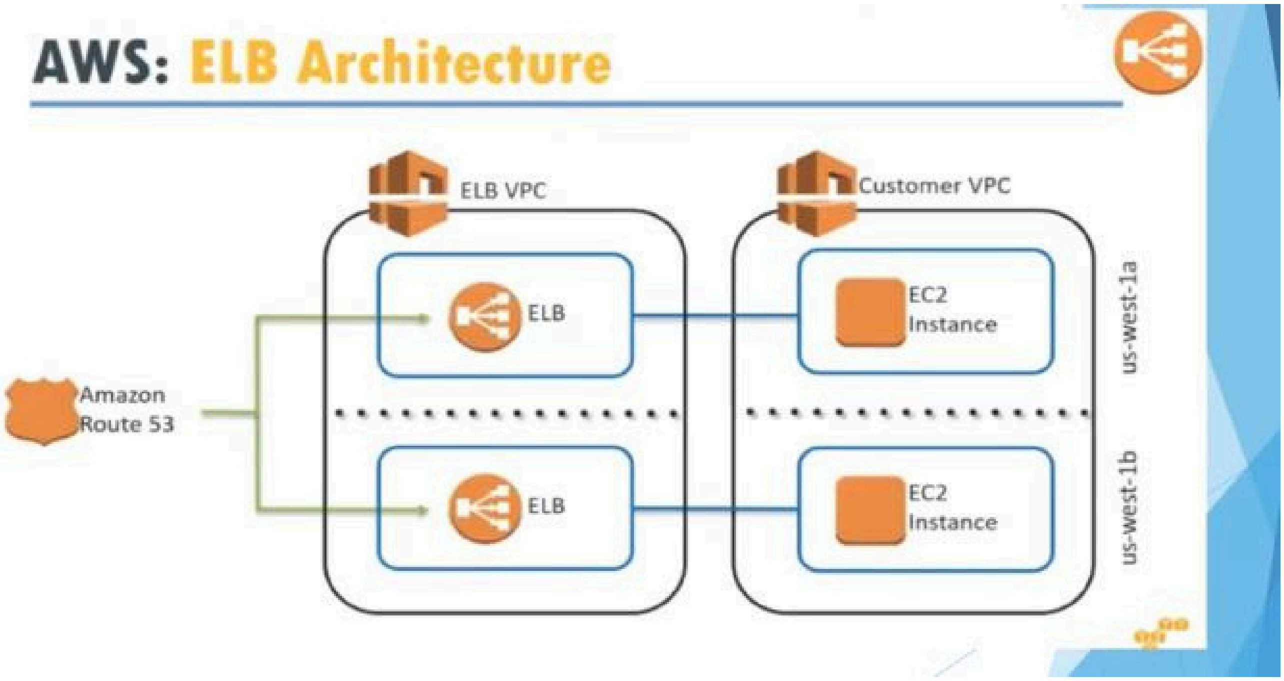
Web-App: With Load Balancer

Load Balancer used to route incoming requests to multiple EC2 instances.



AWS: Elastic Load Balancer

Elastic Load Balancing automatically distributes incoming application traffic across multiple applications, microservices and containers hosted on Amazon EC2 instances.


AWS: ELB Architecture

Elastic Load Balancing automatically distributes incoming application traffic across multiple targets, such as Amazon EC2 instances, containers, and IP addresses.

It can handle the varying load of your application traffic in a single Availability Zone or across multiple Availability Zones.

Elastic Load Balancing offers three types of load balancers that all feature the high availability, automatic scaling, and robust security necessary to make your applications fault tolerant.

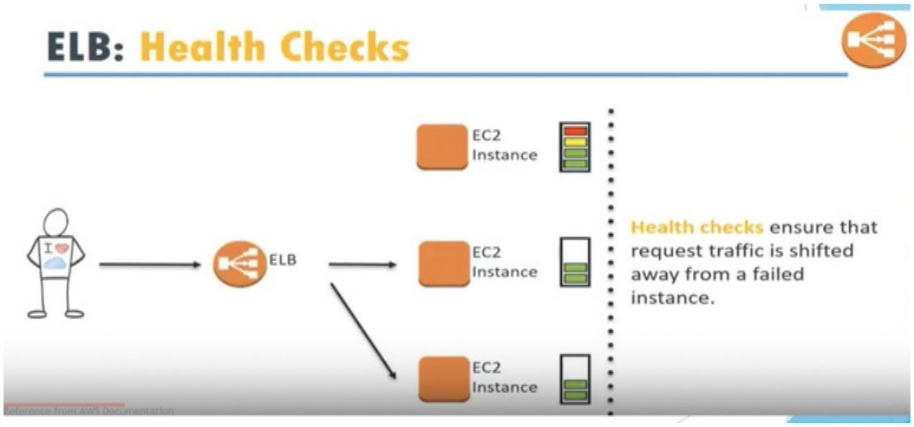
ELB Features:

ELB is used to load balance over EC2 instances within a VPC. Support both public and private IPS.

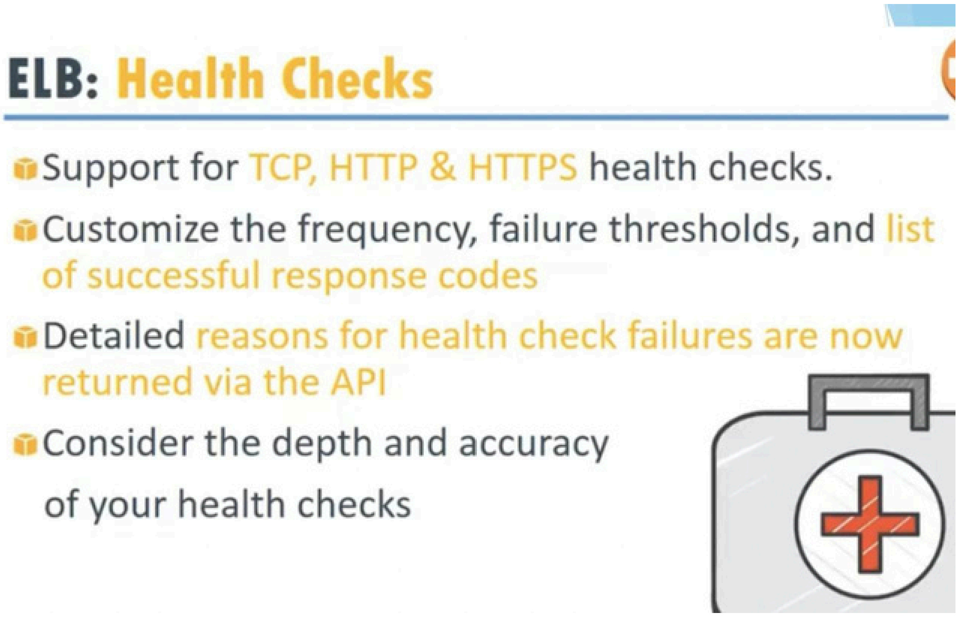
Full control over load balancer and security group.

Tightly integrated into associated VPC and subnet.

How load balancer will do health check


ELB: Health Checks

Health checks ensure that request traffic is shifted away from a failed instance.



ELB: Health Checks

- Support for TCP, HTTP & HTTPS health checks.
- Customize the frequency, failure thresholds, and list of successful response codes
- Detailed reasons for health check failures are now returned via the API
- Consider the depth and accuracy of your health checks

Application layer Load Balancing:

You can load balance HTTP/HTTPS applications and use layer 7-specific features, such as X-Forwarded-For headers.

HTTPS Support:

An Application Load Balancer supports HTTPS termination between the clients and the load balancer. Application Load Balancers also offer management of SSL certificates through AWS Identity and Access Management (IAM) and AWS Certificate Manager for pre-defined security policies.

Server Name Indication (SNI):

Server Name Indication (SNI) is an extension to the TLS protocol by which a client indicates the hostname to connect to at the start of the TLS handshake. The load balancer can present multiple certificates through the same secure listener, which enables it to support multiple secure websites using a single secure listener.

IP addresses as Targets:

You can load balance any application hosted in AWS or on-premises using the IP addresses of the application backends as targets. This allows load balancing to an application backend hosted on any IP address and any interface on an instance. Each application hosted on the same instance can have an associated security group and use the same port. You can also use IP addresses as targets to load balance applications hosted in on-premises locations (over a Direct Connect or VPN connection), peered VPCs, and EC2-Classic (using Classic Link). The ability to load balance across AWS and on-prem resources helps you migrate to the cloud, burst-to-cloud, or failover-to-cloud.

Network load balancer:

Network Load Balancer operates at the connection level (Layer 4), routing connections to targets - Amazon EC2 instances, microservices, and containers – within Amazon Virtual Private Cloud (Amazon VPC) based on IP protocol data. Ideal for load balancing of both TCP and UDP traffic, Network Load Balancer is capable of handling millions of requests per second while maintaining ultra-low latencies. Network Load Balancer is optimized to handle sudden and volatile traffic patterns while using a single static IP address per Availability Zone. It is integrated with other popular AWS services such as Auto Scaling, Amazon EC2 Container Service (ECS), Amazon CloudFormation, and AWS Certificate Manager (ACM).

Network load balances key features:

1. Connection-based Load Balancing

You can load balance both TCP and UDP traffic, routing connections to targets - Amazon EC2 instances, microservices, and containers.

2. High Availability

Network Load Balancer is highly available. It accepts incoming traffic from clients and distributes this traffic across the targets within the same Availability Zone. The load balancer also monitors the health of its registered targets and ensures that it routes traffic only to healthy targets.

3. high Throughput

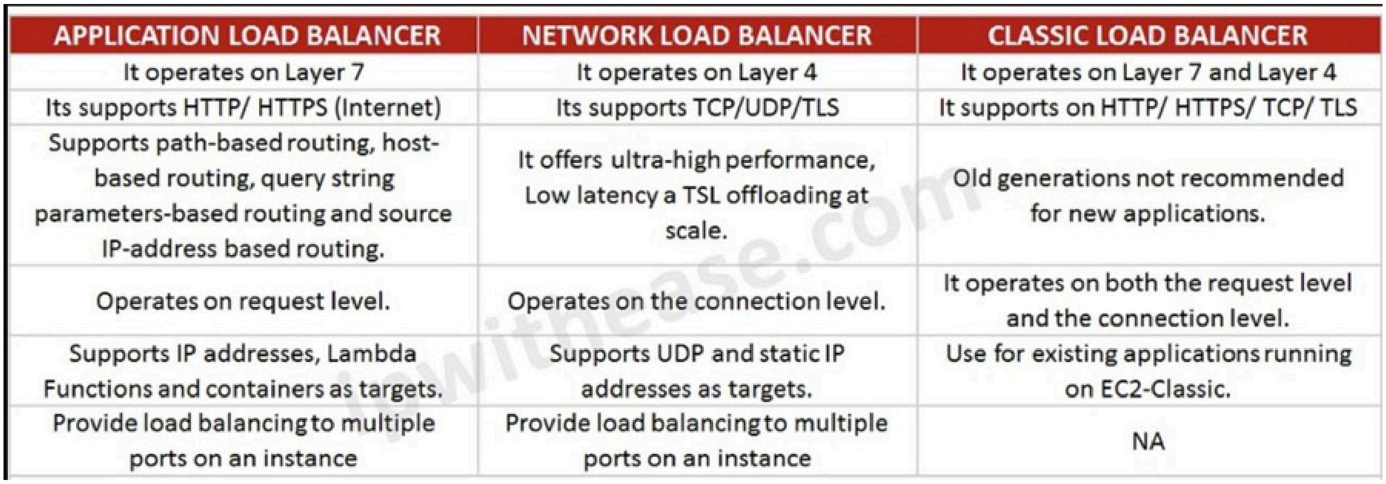
Network Load Balancer is designed to handle traffic as it grows and can load balance millions of requests/sec. It can also handle sudden volatile traffic patterns.

Gateway load balancer:

Gateway Load Balancer makes it easy to deploy, scale, and manage your third-party virtual appliances. It gives you one gateway for distributing traffic across multiple virtual appliances, while scaling them up, or down, based on demand. This eliminates potential points of failure in your network and increases availability.

What is OSI mode?

SDLC life cycle: Agile methodology



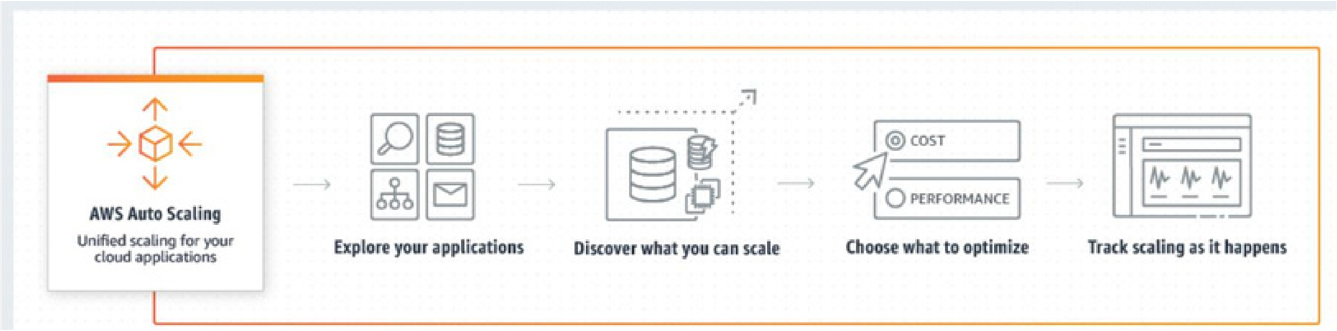
APPLICATION LOAD BALANCER	NETWORK LOAD BALANCER	CLASSIC LOAD BALANCER
It operates on Layer 7	It operates on Layer 4	It operates on Layer 7 and Layer 4
Its supports HTTP/ HTTPS (Internet)	Its supports TCP/UDP/TLS	It supports on HTTP/ HTTPS/ TCP/ TLS
Supports path-based routing, host-based routing, query string parameters-based routing and source IP-address based routing.	It offers ultra-high performance, Low latency a TSL offloading at scale.	Old generations not recommended for new applications.
Operates on request level.	Operates on the connection level.	It operates on both the request level and the connection level.
Supports IP addresses, Lambda Functions and containers as targets.	Supports UDP and static IP addresses as targets.	Use for existing applications running on EC2-Classic.
Provide load balancing to multiple ports on an instance	Provide load balancing to multiple ports on an instance	NA

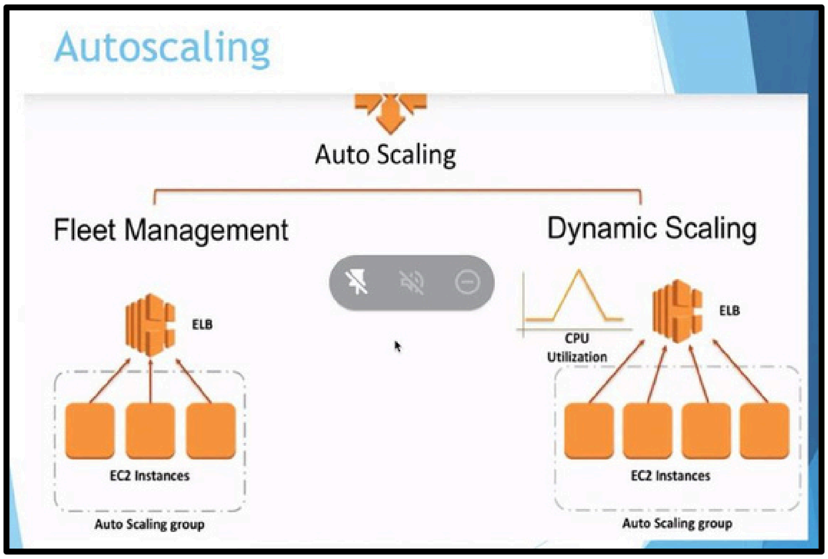
Autoscaling:

AWS Auto Scaling monitors your applications and automatically adjusts capacity to maintain steady, predictable performance at the lowest possible cost.

Autoscaling group:

- Setup scaling quickly
- Automatically maintain performance.
- Make smart scaling decisions.
- Pay only for what you need.
- Aws auto scaling features.


AWS Auto Scaling
Unified scaling for your cloud applications
Explore your applications
Discover what you can scale
COST
PERFORMANCE
Choose what to optimize
Track scaling as it happens


Autoscaling
Auto Scaling
Fleet Management
Dynamic Scaling
ELB
EC2 Instances
Auto Scaling group
CPU Utilization
ELB
EC2 Instances
Auto Scaling group