# I Về lý thuyết

## Giao thức cổng đường biên là gì? What is Border Gateway Protocol

<https://aws.amazon.com/what-is/border-gateway-protocol/#:~:text=Border%20Gateway%20Protocol%20(BGP)%20is,%2C%20devices%2C%20and%20communication%20technologies>.

Giao thức cổng đường biên (BGP) là một tập hợp các quy tắc xác định những tuyến mạng tốt nhất để truyền dữ liệu trên Internet. Internet bao gồm hàng nghìn mạng riêng, công cộng, của công ty và của chính phủ được liên kết với nhau thông qua các giao thức, thiết bị và công nghệ giao tiếp được chuẩn hóa. Khi bạn duyệt Internet, dữ liệu sẽ di chuyển qua nhiều mạng trước khi đến đích. Trách nhiệm của BGP là xem xét tất cả các đường dẫn sẵn sàng truyền dữ liệu và chọn tuyến tốt nhất. Ví dụ: Khi một người dùng ở Hoa Kỳ tải một ứng dụng có máy chủ gốc ở châu Âu, BGP sẽ giúp hoạt động giao tiếp đó diễn ra nhanh chóng và hiệu quả.

## What is jiter

<https://www.gearupbooster.com/vi/blog/what-is-jitter-in-networking.html>

Các gói dữ liệu thường được ngăn cách đều đặn trong một khoảng thời gian định trước. Khi có sự dao động hoặc "thay đổi độ trễ của gói" khi truyền một trong các gói dữ liệu này qua kết nối mạng của bạn, điều này dẫn đến hiện tượng giật hình mà bạn có thể nhận thấy trong cuộc gọi VoIP hoặc cuộc họp video dưới dạng âm thanh hoặc video chậm hoặc các biến dạng khó chịu khác.

Để đưa nó vào một định nghĩa, Jitter là độ lệch của tính chu kỳ của tín hiệu hoặc sự kiện định kỳ so với mục tiêu hoặc tần số thực của nó. Jitter cũng được sử dụng trong viễn thông để mô tả sự dao động về độ trễ của các gói cung cấp dữ liệu âm thanh hoặc video thông qua kết nối truyền thông.

Jitter thường bị nhầm lẫn với độ trễ, vì vậy hãy cùng nêu ra sự khác biệt. Jitter là sự thay đổi về lượng thời gian cần thiết để một gói dữ liệu di chuyển qua mạng. Các gói dữ liệu bị dừng trên đường đến người nhận do tắc nghẽn mạng. Chúng sẽ xuất hiện trong các khoảng thời gian ngẫu nhiên, giống như âm thanh bị rè trong cuộc gọi hoặc hiển thị dưới dạng video pixel trong cuộc gọi điện video. Mặt khác, độ trễ là thời gian cần thiết để một gói dữ liệu di chuyển qua mạng. Với độ trễ, toàn bộ quá trình truyền hoàn chỉnh, các bit và các mảnh, sẽ đến muộn hơn dự kiến.

## 3 what is IpSec ( Vpc support protocol security ‘ipSec’)

Link: <https://aws.amazon.com/vi/what-is/ipsec/#:~:text=IPSec%20l%C3%A0%20h%E1%BB%87%20th%E1%BB%91ng%20c%C3%A1c,c%C6%B0%E1%BB%9Dng%20b%E1%BA%A3o%20m%E1%BA%ADt%20giao%20th%E1%BB%A9c>.

-Amazon supports Internet Protocol Security (IPSec) VPN connections. Data transferred between your VPC and datacenter routes over an encrypted VPN connection to help maintain the confidentiality and integrity of data in transit. An internet gateway is not required to establish an AWS Site-to-Site VPN connection.

IPSec là hệ thống các quy tắc hoặc giao thức truyền thông dùng để thiết lập kết nối an toàn qua một mạng. Giao thức Internet (IP) là tiêu chuẩn phổ biến giúp xác định cách dữ liệu truyền qua Internet. IPSec bổ sung khả năng mã hóa và xác thực để tăng cường bảo mật giao thức. Ví dụ: IPSec xáo trộn dữ liệu tại nguồn và khôi phục dữ liệu bị xáo trộn tại đích của giao thức này. IPSec cũng xác thực nguồn dữ liệu.

## 4 what is dual home instance (VPC)

A Host or Virtual machine with more than one network interfaces is called dual-homed Instance. A dual-homed instance is a system fitted with two network interfaces (NICs) that sits between an untrusted network (like the Internet) and trusted network (such as a corporate network) to provide secure access.

## 5. What is proxy protocol ( Loadbanlacer)

<https://www.google.com/search?q=wwhat+is+proxy+protocol&oq=wwhat+is+proxy+protocol&gs_lcrp=EgZjaHJvbWUyBggAEEUYOTILCAEQABgNGBMYgAQyCwgCEAAYDRgTGIAEMgoIAxAAGBMYFhgeMgoIBBAAGBMYFhgeMgoIBRAAGIAEGKIEMgoIBhAAGIAEGKIEMgcIBxAAGO8FMgcICBAAGO8F0gEIODkwNWowajeoAgCwAgA&sourceid=chrome&ie=UTF-8>

Proxy protocol is an internet protocol used to carry connection information from the source requesting the connection to the destination for which the connection was requested. Elastic Load Balancing uses proxy protocol version 1, which uses a human-readable header format.

## 6. Limit params of AWS

<https://docs.bitnami.com/aws/faq/get-started/understand-limits/#:~:text=static%20IP%20addresses.-,EC2%20Instances,up%20on%20your%20AWS%20account>.

Link detail:

<https://docs.aws.amazon.com/organizations/latest/userguide/orgs_reference_limits.html#:~:text=Limit%20increases%20can%20be%20granted,the%20default%20of%2010%20accounts>.

### Max have 20 EC2 in all account with in a region. Request more need create new request

### By default, AWS sets a limit of 5 static IP addresses per region. This includes IP addresses unassigned and currently assigned to a server.To increase IP addresses limit, [request a higher limit](http://aws.amazon.com/contact-us/eip_limit_request/) by providing information about the new limit and regions where it should be applied.

### The AWS default limit for all snapshots is 10000 snapshots per region.To increase the number of snapshots allowed, [contact AWS Support](https://aws.amazon.com/contact-us/) and request a higher limit.

* Max 10 account for org
* Maximum number of security group in a vpc : 5
* Max vpc in a region: 5
* What is the default concurrency limit for AWS Lambda functions: 1000
* maximum number of messages that can be stored in a standard queue in Amazon SQS: unlimit

## 7. What is Subject Alternative Name san?

The Subject Alternative Name (SAN) is an extension to the X. 509 specification that allows users to specify additional host names for a single SSL certificate. The use of the SAN extension is standard practice for SSL certificates, and it's on its way to replacing the use of the common name.

## 8. What is Server Name Indication (SNI) ?

Server Name Indication (SNI) is an extension to the TLS protocol that is supported by browsers and clients released after 2010. If you configure CloudFront to serve HTTPS requests using SNI, CloudFront associates your alternate domain name with an IP address for each edge location.

* Server Name Indication, often abbreviated SNI, is an extension to TLS that allows multiple hostnames to be served over HTTPS from the same IP address.

## 9. Database Clustering là gì?

Link: <https://topdev.vn/blog/database-clustering-la-gi-khi-nao-nen-su-dung/>

## 10 what is s3 transfer acceleration

<https://aws.amazon.com/vi/s3/transfer-acceleration/#:~:text=S3TA%20shortens%20the%20distance%20between,Edge%20Locations%20to%20your%20application>.

S3TA shortens the distance between client applications and AWS servers that acknowledge PUTS and GETS to Amazon S3 using our global network of hundreds of CloudFront Edge Locations. We automatically route your uploads and downloads through the closest Edge Locations to your application.

## 11 What is Amazon Elastic File System (EFS)

<https://docs.aws.amazon.com/efs/latest/ug/whatisefs.html#:~:text=Amazon%20Elastic%20File%20System%20(Amazon,managing%20storage%20capacity%20and%20performance>.

Amazon Elastic File System (Amazon EFS) provides serverless, fully elastic file storage so that you can share file data without provisioning or managing storage capacity and performance.

**What is Amazon Elastic File System (EFS) and how is it different from Amazon Elastic Block Store (EBS)?**

 a.

EFS is a fully managed, scalable file storage for AWS Cloud services and on-premises resources, while EBS provides persistent block-level storage volumes for use with Amazon EC2 instances.

## 12 type of Storage

<https://www.futurelearn.com/info/courses/aws-cloud-technical-essentials/0/steps/287929#:~:text=AWS%20storage%20services%20are%20grouped,file%20storage%2C%20and%20object%20storage.&text=You%20may%20be%20familiar%20with,Explorer%20or%20Finder%20on%20MacOS>.

## 13 what is AWS Data Pipeline

<https://docs.aws.amazon.com/datapipeline/latest/DeveloperGuide/what-is-datapipeline.html>

AWS Data Pipeline is a web service that you can use to automate the movement and transformation of data. With AWS Data Pipeline, you can define data-driven workflows, so that tasks can be dependent on the successful completion of previous tasks.

## 14 what is Amazon Kinesis Data Firehose

<https://docs.aws.amazon.com/firehose/latest/dev/what-is-this-service.html>

Within this overall flow, Kinesis Data Firehose covers stream ingestion, storage, processing, and destination delivery. It provides near real-time streaming extract, transform, and load (ETL) processing for any data, at any scale, and at low cost

## 15. What is the difference between a public IP address and an Elastic IP address in Amazon VPC?

A public IP address is assigned to an instance when it's launched and cannot be associated with a different instance. An Elastic IP address is a static IP address that can be associated with and disassociated from instances.

## 16. What is the difference between strongly consistent and eventually consistent read?

Strong consistency guarantees that all reads reflect the latest write, while eventual consistency may sometimes return stale data but offers cost savings by using half the capacity units

## 17. Batch Operations (dynamoDB)

Amazon DynamoDB also provides several operations designed for working with large batches of items, including BatchGetItem and BatchWriteItem. Using the BatchWriteItem action, you can perform up to 25 item creates or updates with a single operation. This allows you to minimize the overhead of each individual call when processing large numbers of items.

## 18. Amazon DynamoDB table can scale horizontally through the use of partitions

to meet the storage and performance requirements of your application. Each individual partition represents a unit of compute and storage capacity

* One single partition can hold about 10GB of data and supports a maximum of 3,000 read capacity units or 1,000 write capacity units.

1B,2D,3C,4A,5B,6A,7B,8B(A),9D,10A,11A(B),12B,13ACD,14CD,15ABC,16ABC,17A(C),18A,19BD,20CD(BC)

## 19. SQS (note)

- Each message can have up to 10 attributes.

- The maximum length of a message ID is 100 characters. (name | id is identify in message-queue)

- The maximum length of a receipt handle is 1,024 characters.

- long poll :max 20s

- delay queue max : 15 min (900s)

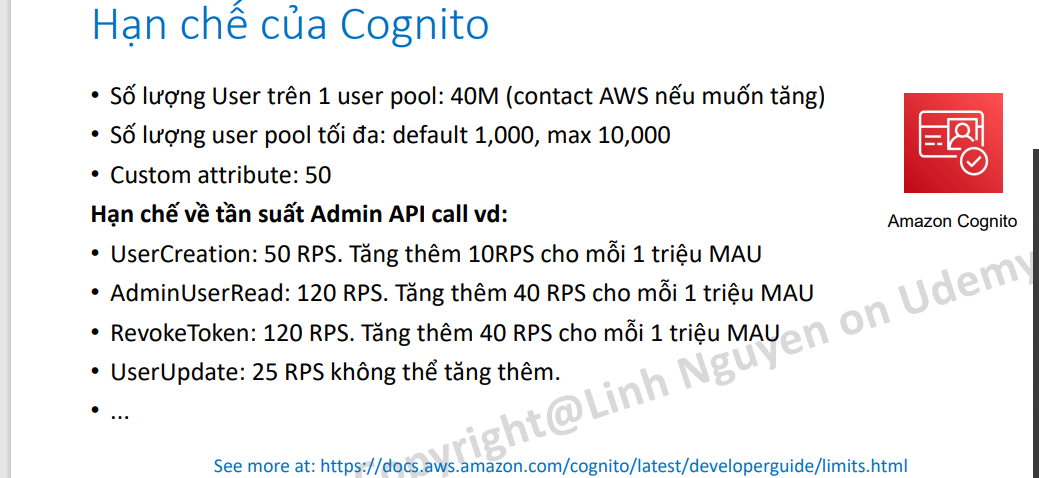
- visibility queue min-max: 30s-12 hours

-dead letter queue retain message in: default= 4day, max=14 days

## 20 SNS (note)

* Số lượng topic/account 100000, FIFO: 1000
* Subcription: 12500000/topic, Fifo: 100/topic
* Delivery rate for email: 10message/s (extendable)
* Message size : 256 KB
* You can use the following protocols with Amazon SNS: HTTP, HTTPS, SMS, email, email-JSON, Amazon SQS, and AWS Lambda
* When you create a new Amazon SNS topic, an Amazon ARN is created automatically.
* Topic names should typically be available for reuse approximately 30–60 seconds after the previous topic with the same name has been deleted

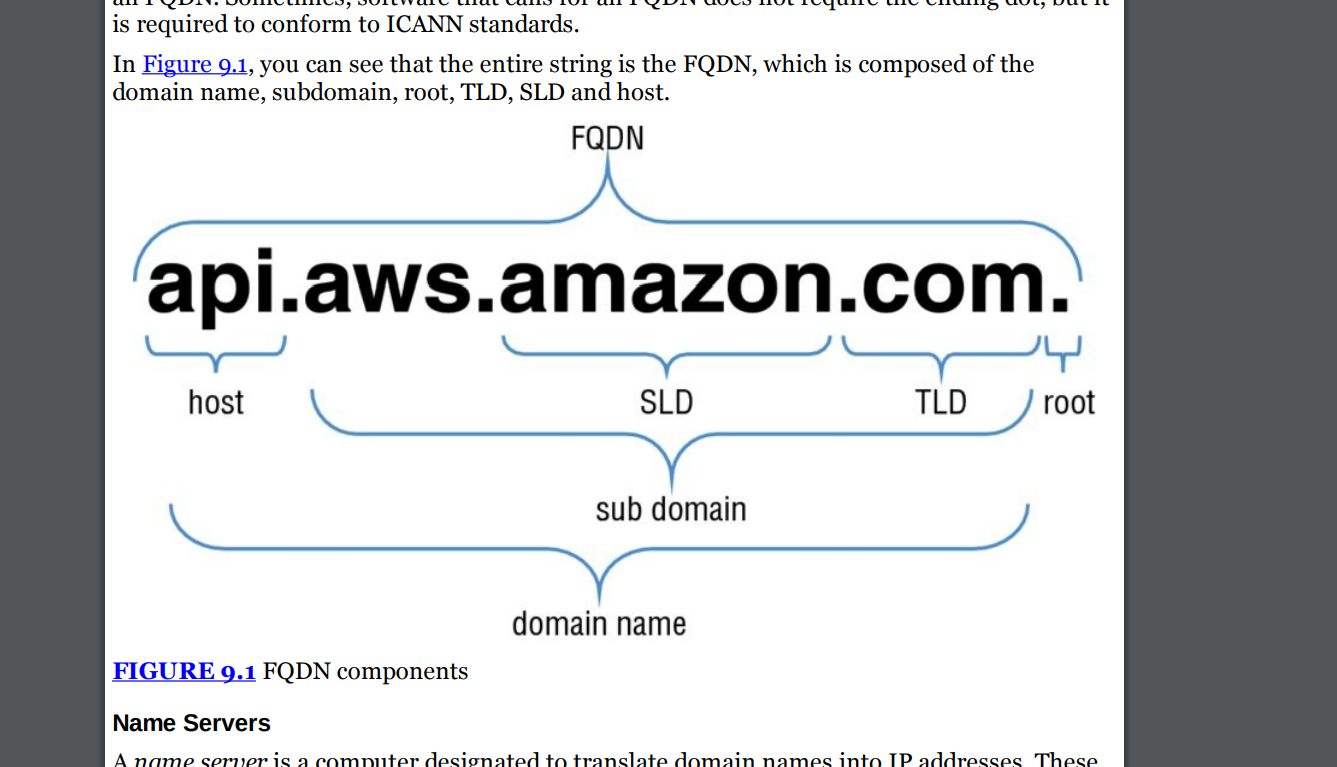
## 21 cognito (limit)



22 SWF (note)

* Actors can be workflow starters, deciders, or activity workers.
* Amazon SWF provides activity workers and deciders with work assignments, given as one of three types of tasks: activity tasks, AWS Lambda tasks, and decision tasks.
* It is possible to have more than one workflow in a domain; however, workflows in different domains cannot interact with one another.
* The decider schedules the activity tasks and provides input data to the activity workers. The decider also processes events that arrive while the workflow is in progress and closes the workflow when the objective has been completed

## 23. Khái niệm về thành phần của 1 domain (full)



* .com = TLD (top level domain)
* .amazon = SLD ( second level domain)
* .amazon.com is domain
* .Aws = sub domain
* Api = host
* Diff hostname ( chir trong1 amsy tinhs)
* Domain name (khái niệm trên internet

## 24. Cần nhớ các nhiệm vụ của record in file zone

• SOA (SOA) Record A Start of Authority (SOA) record is mandatory in all zone files, and it identifies the base DNS information about the domain. Each zone contains a single SOA record.

• A Record (Address Record): Xác định một địa chỉ IPv4 cho tên miền. Nó ánh xạ một tên miền vào một địa chỉ IP v4

• AAAA Record (IPv6 Address Record): Tương tự như A Record, nhưng sử dụng để xác định một địa chỉ IPv6 cho tên miền.

• CNAME Record (Canonical Name Record): Nó được sử dụng để tạo đường dẫn từ một tên miền thứ cấp (subdomain) đến một tên miền ở bất cứ đâu trên internet.

• MX Record (Mail Exchanger Record): Xác định các máy chủ chịu trách nhiệm nhận và xử lý thư điện tử cho một tên miền. Nó được sử dụng để định vị máy chủ email. AWS Cloud for beginner Các loại DNS Record Copyright Linh Nguyen – All right reserved 10

• TXT Record (Text Record): Cho phép bạn lưu trữ các dữ liệu văn bản tùy ý cho tên miền. Nó thường được sử dụng để xác thực tên miền và cung cấp thông tin khác nhau cho các dịch vụ khác.

• SRV Record (Service Record): Xác định vị trí và cấu hình dịch vụ cụ thể trên mạng. Nó được sử dụng chủ yếu trong việc xác định các máy chủ chịu trách nhiệm cho các dịch vụ như VoIP (Voice over IP) và IM (Instant Messaging).

• NS Record (Name Server Record): Xác định máy chủ tên miền (name server) chịu trách nhiệm quản lý các bản ghi DNS cho tên miền cụ thể. Nó cho phép bạn chỉ định máy chủ DNS mà bạn muốn sử dụng cho tên miền của mình.

• PTR Record (Pointer Record): Sử dụng để thực hiện ánh xạ địa chỉ IP thành tên miền. Nó được sử dụng chủ yếu trong việc xác định tên miền từ một địa chỉ IP cụ thể

## 25. Amazon Route 53 allows you to have several different routing policies, including the following:

* Simple—Most commonly used when you have a single resource that performs a given function for your domain
* Weighted—Used when you want to route a percentage of your traffic to one particular resource or resources
* Latency-Based—Used to route your traffic based on the lowest latency so that your users get the fastest response times
* Failover—Used for DR and to route your traffic from your resources in a primary location to a standby location
* Geolocation—Used to route your traffic based on your end user’s location

## 26 . DNS note

* uses port 53.
* DNS primarily uses UDP to serve requests.
* A. The TCP protocol is used by DNS server when the response data size exceeds 512 bytes or for tasks such as zone transfers.

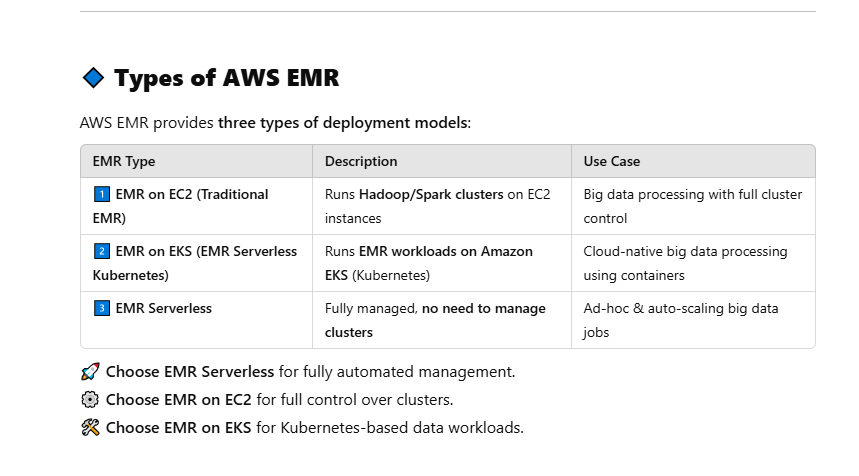
## 27 what is OpsWork

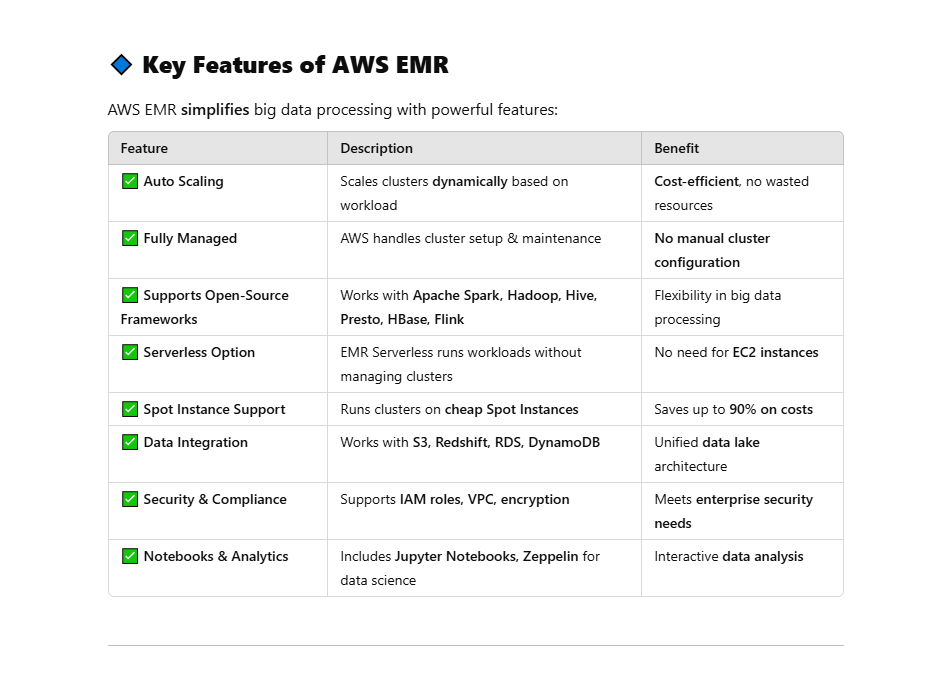
<https://docs.aws.amazon.com/opsworks-cm/latest/APIReference/Welcome.html#:~:text=AWS%20OpsWorks%20CM%20is%20a,for%20the%20servers%20to%20manage>.

AWS OpsWorks CM is a service that runs and manages configuration management servers. You can use AWS OpsWorks CM to create and manage AWS OpsWorks for Chef Automate and OpsWorks for Puppet Enterprise servers, and add or remove nodes for the servers to manage.

## 28 Amazon EMR

Link: <https://docs.aws.amazon.com/emr/latest/ManagementGuide/emr-what-is-emr.html>





Amazon EMR (previously called Amazon Elastic MapReduce) is a managed cluster platform that simplifies running big data frameworks, such as [Apache Hadoop](https://aws.amazon.com/elasticmapreduce/details/hadoop) and [Apache Spark](https://aws.amazon.com/elasticmapreduce/details/spark), on AWS to process and analyze vast amounts of data. Using these frameworks and related open-source projects, you can process data for analytics purposes and business intelligence workloads. Amazon EMR also lets you transform and move large amounts of data into and out of other AWS data stores and databases, such as Amazon Simple Storage Service (Amazon S3) and Amazon DynamoDB.

* Hadoop Distributed File System (HDFS) HDFS is the standard file system that comes with Hadoop. All data is replicated across multiple instances to ensure durability. Amazon EMR can use Amazon EC2 instance storage or Amazon EBS for HDFS. When a cluster is shut down, instance storage is lost and the data does not persist. HDFS can also make use of Amazon EBS storage, trading in the cost effectiveness of instance storage for the ability to shut down a cluster without losing data.
* EMR File System (EMRFS) EMRFS is an implementation of HDFS that allows clusters to store data on Amazon S3. EMRFS allows you to get the durability and low cost of Amazon S3 while preserving your data even if the cluster is shut down. A key factor driving the type of storage a cluster uses is whether the cluster is persistent or transient. A persistent cluster continues to run 24×7 after it is launched.

2 type:

- Hadoop Distributed File System (HDFS) HDFS is the standard file system that comes with Hadoop. All data is replicated across multiple instances to ensure durability. Amazon EMR can use Amazon EC2 instance storage or Amazon EBS for HDFS.

- EMR File System (EMRFS) EMRFS is an implementation of HDFS that allows clusters to store data on Amazon S3. EMRFS allows you to get the durability and low cost of Amazon S3 while preserving your data even if the cluster is shut down.

-- note

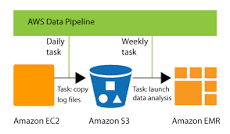
- Know what service Amazon EMR provides. Amazon EMR provides a managed Hadoop service on AWS that allows you to spin up large Hadoop clusters in minutes.

- Know the difference between persistent and transient clusters. Persistent clusters run continuously, so they do not lose data stored on instance-based HDFS. Transient clusters are launched for a specific task, then terminated, so they access data on Amazon S3 via EMRFS.

- Know the use cases for Amazon EMR. Amazon EMR is useful for big data analytics in virtually any industry, including, but not limited to, log processing, clickstream analysis, and genomics and life sciences

-

## 29 What is a data pipeline in AWS?



AWS Data Pipeline is a web service that you can use to automate the movement and transformation of data. With AWS Data Pipeline, you can define data-driven workflows, so that tasks can be dependent on the successful completion of previous tasks.

## 30 Amazon elastic cache

Cache Engines

* Memcache

+ A single Memcached cluster can contain up to 20 nodes

* Redis :
* + no support cluster but support multi AZ a up to 5 replica

+ max 5 replica

+ Cache clusters running Redis support the concept of replication groups. A replication group consists of up to six clusters, with five of them designated as read replicas

## 31 AWS CloudTrail

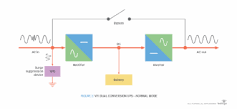
* A Trail That Applies to All Regions

When you create a trail that applies to all AWS regions, AWS CloudTrail creates the same trail in each region, records the log files in each region, and delivers the log files to the single Amazon S3 bucket (and optionally to the Amazon CloudWatch Logs log group) that you specify

* A Trail That Applies to One Region :

You specify a bucket that receives events only from that region. The bucket can be in any region that you specify. If you create additional individual trails that apply to specific regions, you can have those trails deliver event logs to a single Amazon S3 bucket

## 32 What is UPS used for?



An uninterruptible power supply (UPS) is a device that allows a computer to keep running for at least a short time when incoming power is interrupted. Provided utility power is flowing, it also replenishes and maintains energy storage.

## 33 AWS Import/Export (SnowBall)

### AWS Snowball AWS

AWS Snowball AWS Snowball uses Amazon-provided shippable storage appliances shipped through UPS. Each AWS Snowball is protected by AWS KMS and made physically rugged to secure and protect your data while the device is in transit. At the time of this writing, AWS Snowballs come in two sizes: 50TB and 80TB, and the availability of each varies by region. AWS Snowball provides the following features: - + You can import and export data between your on-premises data storage locations and Amazon S3.

+ Encryption is enforced, protecting your data at rest and in physical transit.

+ You don’t have to buy or maintain your own hardware devices.

+ You can manage your jobs through the AWS Snowball console.

+ The AWS Snowball is its own shipping container, and the shipping label is an E Ink display that automatically shows the correct address when the AWS Snowball is ready to ship. You can drop it off with UPS, no box required.

### AWS Import/Export Disk

AWS Import/Export Disk supports transfers data directly onto and off of storage devices you own using the Amazon high-speed internal network.

Important things to understand about AWS Import/Export Disk include:

+ You can import your data into Amazon Glacier and Amazon EBS, in addition to Amazon S3.

+ You can export data from Amazon S3.

+ Encryption is optional and not enforced.

+ You buy and maintain your own hardware devices.

+ You can’t manage your jobs through the AWS Snowball console.

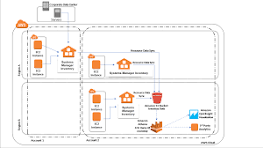
+ Unlike AWS Snowball, AWS Import/Export Disk has an upper limit of 16TB.

## 34 AWS Directory Service

AWS Directory Service is a managed service offering that provides directories that contain information about your organization, including users, groups, computers, and other resources.

* AWS Directory Service for Microsoft Active Directory (Enterprise Edition) This Directory Service is your best choice if you have more than 5,000 users and need a trust relationship set up between an AWS-hosted directory and your on-premises directories.
* Simple AD In most cases, Simple AD is the least expensive option and your best choice if you have 5,000 or fewer users and don’t need the more advanced Microsoft Active Directory features. (Note that you cannot set up trust relationships between Simple AD and other Active Directory domains.)
* AD Connector AD Connector is your best choice when you want to use your existing onpremises directory with AWS cloud service

## 35 What is inventory in AWS?



Link <https://docs.aws.amazon.com/systems-manager/latest/userguide/systems-manager-inventory.html>

AWS Systems Manager Inventory provides visibility into your Amazon EC2 and on-premises computing environment. You can use Inventory to collect metadata from your managed instances.

## 36 Amazon CloudFront?

* What origin servers are supported by Amazon CloudFront
* Amazon CloudFront can use an Amazon S3 bucket or any HTTP server

### Whole Website

Using cache behaviors and multiple origins, you can easily use Amazon CloudFront to serve your whole website and to support different behaviors for different client devices.

### Private Content

* Signed URLs Use URLs that are valid only between certain times and optionally from certain IP addresses.
* Signed Cookies Require authentication via public and private key pairs.
* Origin Access Identities (OAI) Restrict access to an Amazon S3 bucket only to a special Amazon CloudFront user associated with your distribution. This is the easiest way to ensure that content in a bucket is only accessed by Amazon CloudFront.

##### **What is the HTTP response status code for a request that is blocked by AWS CloudFront?**

: 403

## 37 AWS Storage Gateway

AWS Storage Gateway is a service connecting an on-premises software appliance with cloud- based storage to provide seamless and secure integration between an organization’s onpremises IT environment and AWS storage infrastructure. The service enables you to store data securely on the AWS cloud in a scalable and cost-effective manner

### - Gateway-Cached Volumes

Gateway-Cached volumes allow you to expand your local storage capacity into Amazon S3. All data stored on a Gateway-Cached volume is moved to Amazon S3, while recently read data is retained in local storage to provide low-latency access. While each volume is limited to a maximum size of 32TB, a single gateway can support up to 32 volumes for a maximum storage of 1 PB

### Gateway-Stored Volumes

Gateway-Stored volumes allow you to store your data on your on-premises storage and asynchronously back up that data to Amazon S3. This provides lowlatency access to all data, while also providing off-site backups taking advantage of the durability of Amazon S3. The data is backed up in the form of Amazon Elastic Block Store (Amazon EBS) snapshots. While each volume is limited to a maximum size of 16TB, a single gateway can support up to 32 volumes for a maximum storage of 512TB.

## 38 AWS Key Management Service (KMS)

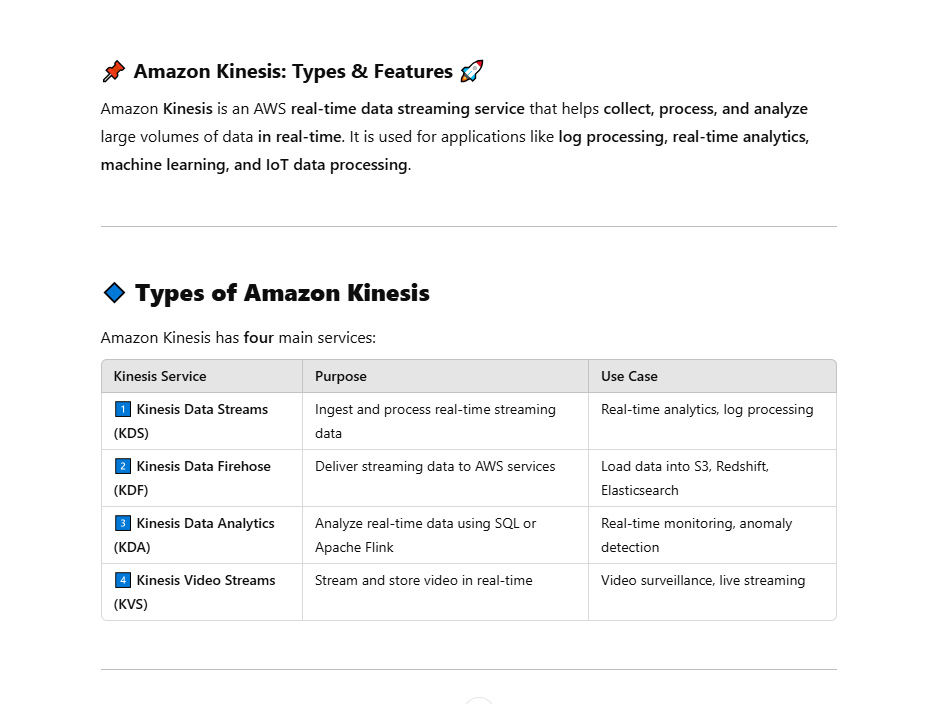
AWS Key Management Service (AWS KMS) AWS KMS is a managed service that makes it easy for you to create and control the encryption keys used to encrypt your data.

* Customer Managed Keys AWS KMS uses a type of key called a Customer Master Key (CMK) to encrypt and decrypt data. CMKs are the fundamental resources that AWS KMS manages. They can be used inside of AWS KMS to encrypt or decrypt up to 4 KB of data directly. They can also be used to encrypt generated data keys that are then used to encrypt or decrypt larger amounts of data outside of the service. CMKs can never leave AWS KMS unencrypted, but data keys can leave the service unencrypted.
* Data Keys You use data keys to encrypt large data objects within your own application outside AWS KMS. When you call GenerateDataKey, AWS KMS returns a plaintext version of the key and ciphertext that contains the key encrypted under the specified CMK. AWS KMS tracks which CMK was used to encrypt the data key. You use the plaintext data key in your application to encrypt data, and you typically store the encrypted key alongside your encrypted data. Security best practices suggest that you should remove the plaintext key from memory as soon as is practical after use. To decrypt data in your application, pass the encrypted data key to the Decrypt function. AWS KMS uses the associated CMK to decrypt and retrieve your plaintext data key. Use the plaintext key to decrypt your data, and then remove the key from memory.
* Envelope Encryption AWS KMS uses envelope encryption to protect data. AWS KMS creates a data key, encrypts it under a CMK, and returns plaintext and encrypted versions of the data key to you. You use the plaintext key to encrypt data and store the encrypted key alongside the encrypted data. The key should be removed from memory as soon as is practical after use. You can retrieve a plaintext data key only if you have the encrypted data key and you have permission to use the corresponding master key.
* Encryption Context All AWS KMS cryptographic operations accept an optional key/value map of additional contextual information called an encryption context. The specified context must be the same for both the encrypt and decrypt operations or decryption will not succeed. The encryption context is logged, can be used for additional auditing, and is available as context in the AWS policy language for fine-grained policybased authorization.

## 39 Cloud Trail

AWS CloudTrail captures AWS API calls and related events made by or on behalf of an AWS account and delivers log files to an Amazon S3 bucket that you specify

## 40 Amazon Kinesis



* Kinesis data analytic only analytic when have data from kinesis data firehoue or kineis stream. Summary kinesis cant get data from source

Amazon Kinesis is a platform for handling massive streaming data on AWS, offering powerful services to make it easy to load and analyze streaming data and also providing the ability for you to build custom streaming data applications for specialized needs.

Overview Amazon Kinesis is a streaming data platform consisting of three services addressing different real-time streaming data challenges:

* Amazon Kinesis Firehose: A service enabling you to load massive volumes of streaming data into AWS (can store to Amazon S3, Amazon Redshift, or Amazon Elasticsearch.)
* Amazon Kinesis Streams: A service enabling you to build custom applications for more complex analysis of streaming data in real time

. Amazon Kinesis Streams can scale to support nearly limitless data streams by distributing incoming data across a number of shards

* Amazon Kinesis Analytics: A service enabling you to easily analyze streaming data real time with standard SQL
* **Kinesis Video Streams (KVS)**

🔸 **Purpose:** Capture, process, and analyze **real-time video streams**.  
🔸 **Use Case:** **Live video streaming, surveillance, IoT cameras, AI/ML applications**.  
🔸 **Key Features:**

* **Supports multiple formats**: H.264, WebRTC.
* **Low latency streaming**: Stream to AWS AI services for analysis.
* **Integration with AI/ML**: Works with Amazon Rekognition for **video analysis**.

## 42 S3 Inventory

Link: <https://docs.aws.amazon.com/AmazonS3/latest/userguide/storage-inventory.html>

You can use Amazon S3 Inventory to help manage your storage. For example, you can use it to audit and report on the replication and encryption status of your objects for business, compliance, and regulatory needs. You can also simplify and speed up business workflows and big data jobs by using Amazon S3 Inventory, which provides a scheduled alternative to the Amazon S3 synchronous List API operations. Amazon S3 Inventory does not use the List API operations to audit your objects and does not affect the request rate of your bucket.

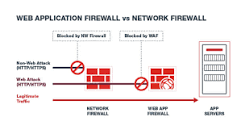
You can query Amazon S3 Inventory with standard SQL queries by using [Amazon Athena](https://docs.aws.amazon.com/athena/latest/ug/what-is.html), [Amazon Redshift Spectrum](https://docs.aws.amazon.com/redshift/latest/dg/c-getting-started-using-spectrum.html), and other tools, such as [Presto](https://prestodb.io/), [Apache Hive](https://hive.apache.org/), and [Apache Spark](https://databricks.com/spark/about/). For more information about using Athena to query your inventory files

## 43 What is the use of Amazon Athena?

Link: <https://docs.aws.amazon.com/athena/latest/ug/what-is.html>

Amazon Athena is a service that enables data analysts to perform interactive queries in the web-based cloud storage service, Amazon Simple Storage Service (S3). Athena is used with large-scale data sets. Amazon S3 is designed for online backup and archiving of data and applications on Amazon Web Services (AWS).

## 44 What is the difference between AWS firewall and WAF?



A WAF is essential when you need specialized protection for web applications against web-specific threats. Firewalls provide broader network security but may not catch application-layer attacks.

## 45 what is  AWS Global Accelerator

## 46 What is the purpose of VPC Flow Logs?

## 47 Which Amazon RDS engine supports encryption at rest as a default option?

## 48 Amazon CloudWatch

CloudWatch is the AWS service designed to monitor and troubleshoot applications running across hybrid environments. It provides a unified view of operational health and collects metrics, logs, and events from various AWS resources, on-premises servers, and applications

 CloudWatch enables you to set alarms, visualize logs and metrics side by side, take automated actions, and gain system-wide visibility into resource utilization, application performance, and operational health. This makes it an ideal tool for monitoring and troubleshooting applications in hybrid environments, offering insights that span both cloud and on-premises infrastructure.

## 49 \* Access instance metadata for an EC2 instance

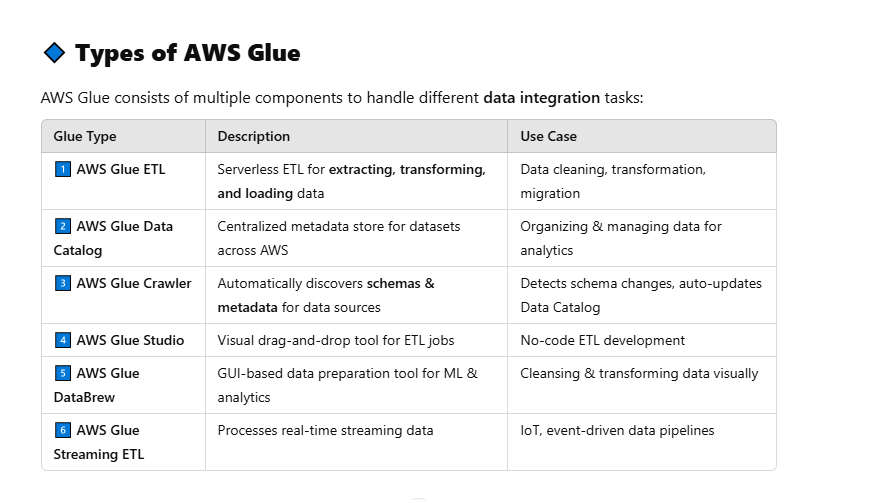
Link: <https://docs.aws.amazon.com/AWSEC2/latest/UserGuide/instancedata-data-retrieval.html>

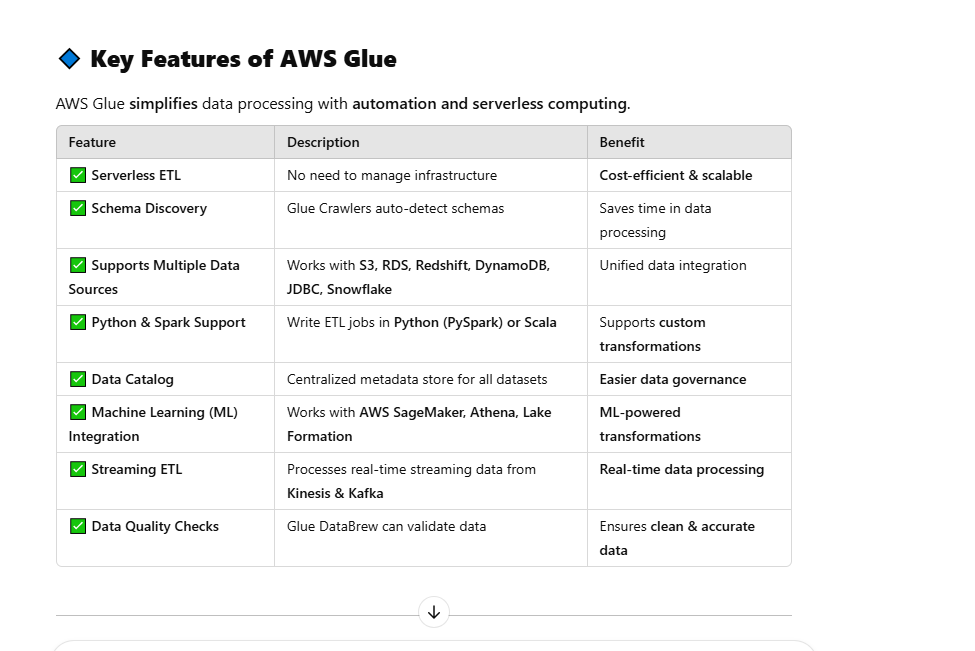
The correct answer, 'curl http://169.254.169.254/latest/user-data', allows you to view the user data script directly from the instance without accessing the log file.

## 50 what is  AWS Glue?

AWS Glue is a fully managed ETL service that makes it easy for customers to prepare and load their data for analytics. You can create and run an ETL job with a few clicks in the AWS Management Console .

AWS **Glue** is a **serverless data integration service** that helps **extract, transform, and load (ETL) data** from multiple sources. It automates **data discovery, transformation, and cataloging** for **data lakes, analytics, and machine learning (ML) workflows**.





Link: <https://docs.aws.amazon.com/glue/latest/dg/add-classifier.html>

* A classifier is a metadata table used to infer schema of data

## 51 What is the difference between a public IP address and an Elastic IP address in Amazon VPC?

A public IP address is assigned to an instance when it's launched and cannot be associated with a different instance. An Elastic IP address is a static IP address that can be associated with and disassociated from instances.

## 52 what is ENI (VPC)

Elastic Network Interfaces (ENIs)

An Elastic Network Interface (ENI) is a virtual network interface that you can attach to an instance in an Amazon VPC. ENIs are only available within an Amazon VPC, and they are associated with a subnet upon creation. They can have one public IP address and multiple private IP addresses. If there are multiple private IP addresses, one of them is primary. Assigning a second network interface to an instance via an ENI allows it to be dual-homed (have network presence in different subnets).

## 53 Network File System (NFS)

Link : <https://cmccloud.vn/tin-tuc/nfs-la-gi-197>

The Network File System (NFS) is a mechanism for storing files on a network. It is a distributed file system that allows users to access files and directories located on remote computers and treat those files and directories as if they were local.

## 54 What is Amazon Athena?

<https://docs.aws.amazon.com/athena/latest/ug/what-is.html>

is ServerLess (no Server => no config and managed)

Amazon Athena is an interactive query service that makes it easy to analyze data directly in Amazon Simple Storage Service (Amazon S3) using standard [SQL](https://docs.aws.amazon.com/athena/latest/ug/ddl-sql-reference.html). With a few actions in the AWS Management Console, you can point Athena at your data stored in Amazon S3 and begin using standard SQL to run ad-hoc queries and get results in seconds.

## 55 what is AWS Direct Connect

<https://aws.amazon.com/vi/directconnect/>

Direct Connect connections are provisioned on standalone hardware devices that enables you to create highly resilient network connections between Amazon Virtual Private Cloud and your on-premises infrastructure. This capability enables you to access your AWS resources in a reliable, scalable, and cost-effective way.

Dịch vụ đám mây AWS Direct Connect là đường dẫn ngắn nhất đến các tài nguyên AWS của bạn. Trong khi chuyển tiếp, lưu lượng mạng của bạn vẫn ở trên mạng toàn cầu AWS và không dùng đến Internet công cộng. Điều này giúp giảm khả năng gặp phải sự cố tắc nghẽn hoặc tăng độ trễ bất ngờ. Khi tạo kết nối mới, bạn có thể chọn một kết nối lưu trữ do Đối tác phân phối AWS Direct Connect cung cấp hoặc chọn một kết nối chuyên dụng từ AWS và triển khai tại các vị trí AWS Direct Connect trên toàn thế giới. Với AWS Direct Connect SiteLink, bạn có thể gửi dữ liệu giữa các vị trí AWS Direct Connect để tạo kết nối mạng riêng giữa các văn phòng và trung tâm dữ liệu trong mạng toàn cầu của mình.

## 56 AWS Config

AWS Config is a fully managed service that provides you with an AWS resource inventory, configuration history, and configuration change notifications to enable security and governance. With AWS Config, you can discover existing and deleted AWS resources, determine your overall compliance against rules, and dive into configuration details of a resource at any point in time. These capabilities enable compliance auditing, security analysis, resource change tracking, and troubleshooting.

<https://aws.amazon.com/vi/config/>

* Liên tục đánh giá, giám sát và ghi lại các thay đổi cấu hình tài nguyên để đơn giản hóa hoạt động quản lý thay đổi.
* Liên tục kiểm tra và đánh giá sự tuân thủ với các chính sách do tổ chức ban hành của cấu hình tài nguyên của bạn.
* Đơn giản hóa hoạt động khắc phục sự cố vận hành bằng cách tạo mối tương quan giữa các thay đổi cấu hình với các sự kiện cụ thể trong tài khoản của bạn.

### Hợp lý hóa việc khắc phục sự cố hoạt động và quản lý thay đổi

Khám phá các tài nguyên tồn tại trong tài khoản của bạn hoặc phát hành dữ liệu cấu hình của các tài nguyên bên thứ ba vào AWS Config, ghi lại cấu hình của chúng và nắm bắt mọi thay đổi để nhanh chóng khắc phục các sự cố hoạt động.

## 57 Amazon Inspector

<https://aws.amazon.com/vi/inspector/pricing/#:~:text=Amazon%20Inspector%20l%C3%A0%20m%E1%BB%99t%20d%E1%BB%8Bch,tu%C3%A2n%20th%E1%BB%A7%20cho%20kh%E1%BB%91i%20l%C6%B0%E1%BB%A3ng>

là một dịch vụ quét lỗ hổng bảo mật tự động và liên tục nhằm đánh giá các phiên bản Đám mây điện toán linh hoạt (EC2) của Amazon, hàm AWS Lambda và hình ảnh bộ chứa trong Amazon ECR và trong các công cụ tích hợp liên tục và phân phối liên tục (CI/CD) để cải thiện tính bảo mật và tuân thủ cho khối lượng ...

## 58 what is AWS Secrets Manager

<https://aws.amazon.com/vi/secrets-manager/faqs/#:~:text=AWS%20Secrets%20Manager%20enables%20you,unauthorized%20users%20viewing%20sensitive%20information>.

AWS Secrets Manager enables you to store, retrieve, control access to, rotate, audit, and monitor secrets centrally. You can encrypt secrets at rest to reduce the likelihood of unauthorized users viewing sensitive information.

* Must encrypt

 **Automatic Secret Rotation**: Allows you to automatically rotate secrets without disrupting your applications.

 **Cross-Region Replication**: Enables replication of secrets across multiple AWS regions to support disaster recovery and multi-region applications.

 **Higher Storage Limits**: Supports secrets up to 64KB in size.

## 59 what is AWS Systems Manager Parameter Store

<https://docs.aws.amazon.com/managedservices/latest/userguide/sys-man-param-store.html#:~:text=AWS%20Systems%20Manager%20Parameter%20Store%20provides%20secure%2C%20hierarchical%20storage%20for,license%20codes%20as%20parameter%20values>.

AWS Systems Manager Parameter Store provides secure, hierarchical storage for configuration data management and secrets management. You can store data such as passwords, database strings, and license codes as parameter values.

Not have rote key param, encrypt is optional

 **Standard Parameters**: Allow storage of values up to 4KB in size. The first 10,000 standard parameters are free.

 **Advanced Parameters**: Support values up to 8KB and offer additional features like parameter policies, but they incur additional costs.

## 60 ****What is AWS CloudFormation?****

AWS CloudFormation enables you to automate the provisioning of AWS infrastructure using code. Instead of manually creating resources (like EC2, S3, or RDS), you define them in a CloudFormation **template**, and AWS handles the provisioning.

### **Key Concepts**

* **Template**: A JSON or YAML file that describes the AWS resources.
* **Stack**: A collection of resources created from a template.
* **StackSet**: Used to deploy stacks across multiple AWS accounts and regions.
* **Change Set**: Shows the differences before updating a stack.
* Stack updates: First the changes are submitted and compared with the current state of the stack and only the changed resources get updated. There are two methods for updating stacks:

● Direct update - when there is a need to quickly deploy the updates.

● Creating and executing change sets - they are JSON files, providing a preview option for the changes to be applied. StackSets are responsible for safely provisioning, updating, or deleting stacks.

- **Nested Stacks** are stacks created within another stack by using the AWS::CloudFormation::Stack resource

**- AWS CloudFormation Registry** helps to provision third-party application resources alongside AWS resources. Examples of third-party resources are incident management, version control tools.

## 61 AWS Organizations What are AWS Organizations?\

AWS Organizations is a global service that enables users to consolidate and manage multiple AWS accounts into an organization. It includes account management and combined billing capabilities that help to meet the budgetary, and security needs of the business better.

● The main account is the management account – it cannot be changed.

● Other accounts are member accounts that can only be part of a single organization

## 62 AWS Systems Manager What is AWS Systems manager?

AWS Systems Manager is a service which helps users to manage EC2 and on-premises systems at scale. It not only detects the insights about the state of the infrastructure but also easily detects problems as well. Additionally, we can patch automation for enhanced compliance. This AWS service works for both Windows and Linux operating systems

Features:

● Easily integrated with CloudWatch metrics/dashboards and AWS Config.

● It helps to discover and audit the software installed.

● Compliance management

● We can group more than 100 resource types into applications, business units, and environments.

● It helps to view instance information such as operating system patch levels, install software and see the compliance with the desired state.

● Act associate and configurations with resources and find out the discrepancies.

● Distribute multiple software versions safely across the instances.

● Increase the security area by running a command or maintaining scripts.

● Patch your instances of schedule to keep them compliant.

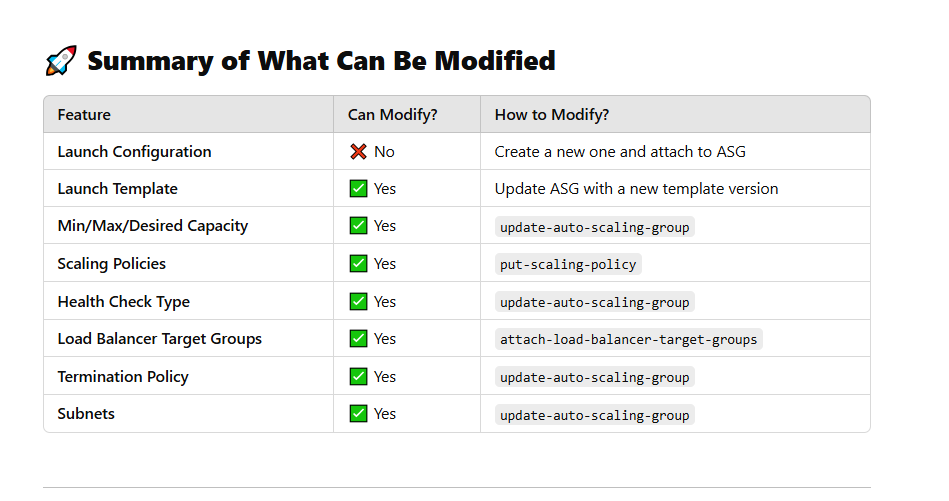
● Helps managers to automate workflows.

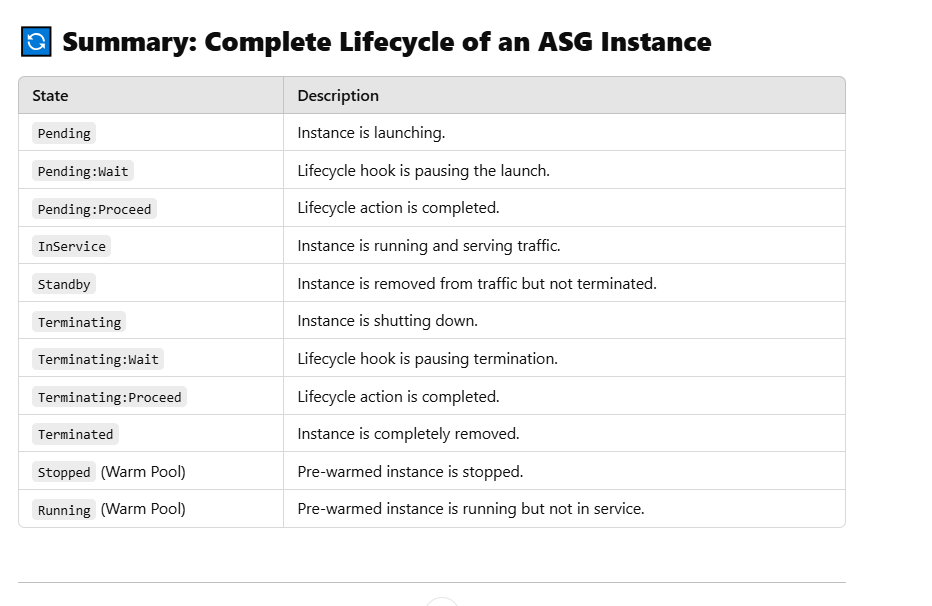
● It helps to reduce errors by securely applying configurable parameters into centralized service

**How does the System Manager work?**

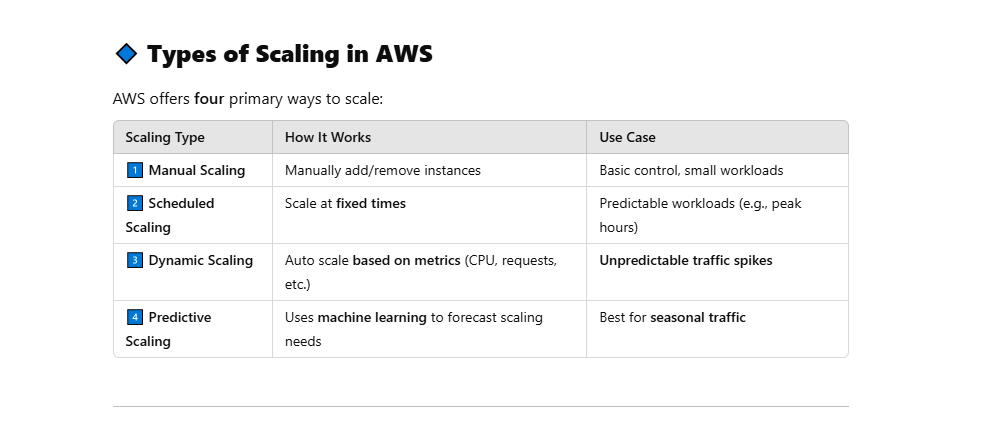
Firstly, User needs to install the SSM agent on the system they control. If an instance can’t be controlled with SSM, it’s probably an issue with the SSM agent. Also, we ne

## \*63 Auto scaling group

* It is not possible to modify a launch configuration once it is created
* 



## 64 Types of Scaling in AWS



# Một số note trong sách

## S3

- create bucket is private deafult

- . Bucket names can contain up to 63 lowercase letters, numbers, hyphens, and periods

- Objects can range in size from 0 bytes up to 5TB, and a single bucket can store an unlimited number of objects

- Each object consists of data (the file itself) and metadata (data about the file). The data portion of an Amazon S3 object is opaque to Amazon S3. This means that an object’s data is treated as simply a stream of bytes—Amazon S3 doesn’t know or care what type of data you are storing, and the service doesn’t act differently for text data versus binary data. The metadata associated with an Amazon S3 object is a set of name/value pairs that describe the object.

- Amazon S3 scales automatically, but for request rates over 100 GETS per second, it helps to make sure there is some randomness in the key space. Replication and logging will not affect performance or scalability. Using sequential key names could have a negative effect on performance or scalability.

- Amazon Glacier Is frequently erased within 90 days. Data is stored in encrypted archives that can be as large as 40TB. Archives typically contain TAR or ZIP files (a archives = 40TB), need 3-5 hours to restore

- Amazon Glacier can be used as a standalone service and as an Amazon S3 storage class.

## EC2

Max size of a volumn EBS is 16TB

#### Addressing an Instance:

- Public Domain Name System (DNS) Name—When you launch an instance, AWS creates a DNS name that can be used to access the instance. This DNS name is generated automatically and cannot be specified by the customer

- Public IP—A launched instance may also have a public IP address assigned. This IP address is assigned from the addresses reserved by AWS and cannot be specified

- Elastic IP—An elastic IP address is an address unique on the Internet that you reserve independently and associate with an Amazon EC2 instance

- Can modify type of instand :

Instance Type The ability to change the instance type of an instance contributes greatly to the agility of running workloads in the cloud

#### Security Groups:

If an instance is running in an Amazon VPC (discussed in Chapter 4), you can change which security groups are associated with an instance while the instance is running. For instances outside of an Amazon VPC (called EC2-Classic), the association of the security groups cannot be changed after launch.

* While termination instance but cli still can access instance. Turn on termination protection will disable termination in cli

. In order to prevent termination via the AWS Management Console, CLI, or API, termination protection can be enabled for an instance. While enabled, calls to terminate the instance will fail until termination protection is disabled. This helps to prevent accidental termination through human error.

#### Pricing Options

- On-Demand Instances

- Reserved Instances

+ All Upfront—Pay for the entire reservation up front. There is no monthly charge for the customer during the term.

+ Partial Upfront—Pay a portion of the reservation charge up front and the rest in monthly installments for the duration of the term.

+ No Upfront—Pay the entire reservation charge in monthly installments for the duration of the term.

- Spot Instances

#### Tenancy Options

- Shared Tenancy Shared tenancy is the default tenancy model for all Amazon EC2 instances, regardless of instance type, pricing model, and so forth

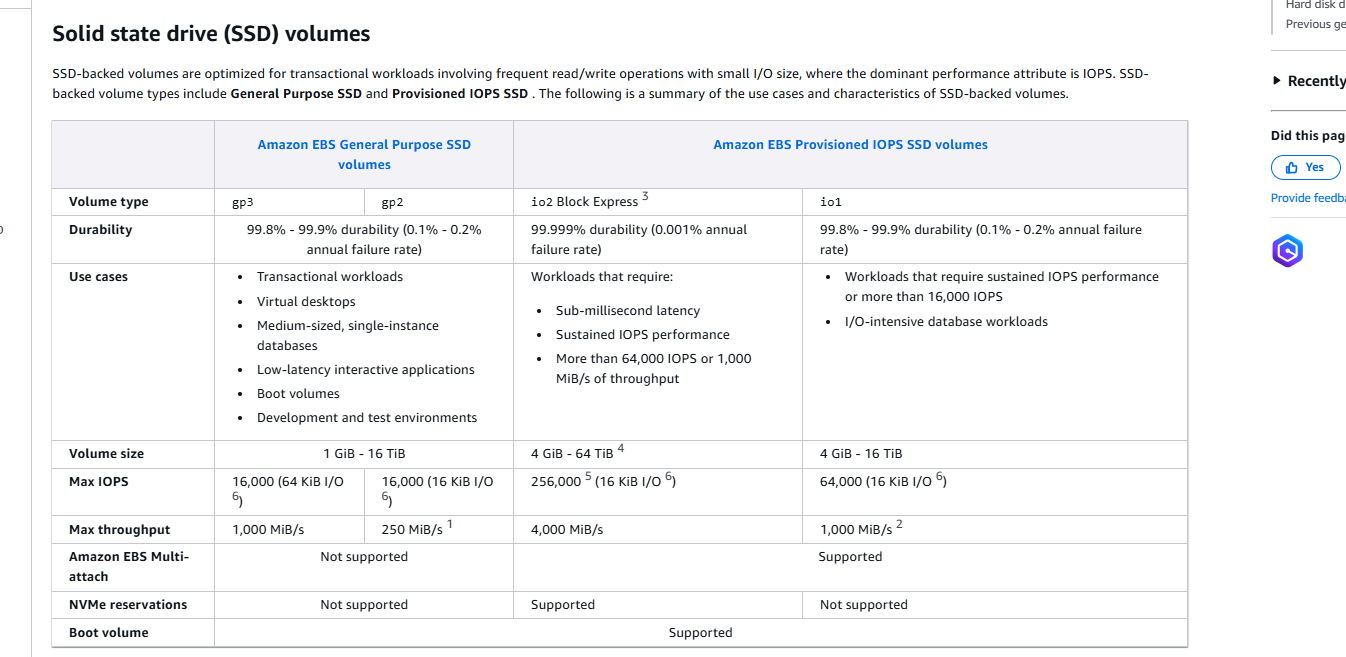
- Dedicated Instances Dedicated Instances run on hardware that’s dedicated to a single customer

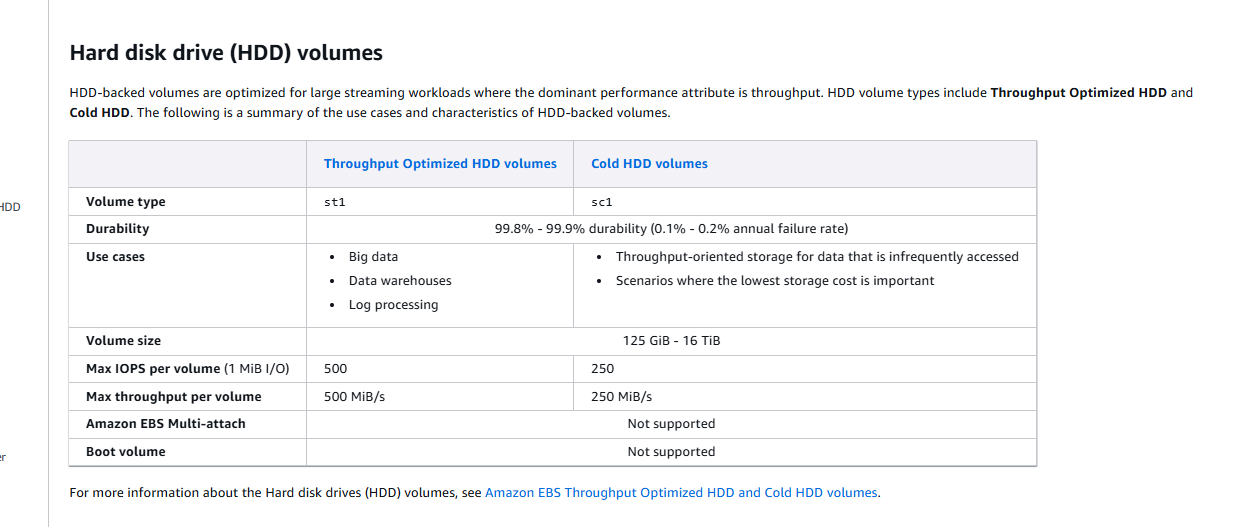
- Dedicated Host An Amazon EC2 Dedicated Host is a physical server with Amazon EC2 instance capacity fully dedicated to a single customer’s use.

#### Amazon Elastic Block Store (Amazon EBS)

- Magnetic Volumes

A magnetic Amazon EBS volume can range in size from 1 GB to 1 TB and will average 100 IOPS, but has the ability to burst to hundreds of IOPS. They are best suited for: Workloads where data is accessed infrequently Sequential reads Situations where low-cost storage is a requirement

* General-Purpose SSD 1GB- 16TB max 10000 IOPS, 1TB =3000IOPS ⬄ 3TB=10000IOPS
* Provisioned IOPS SSD Provisioned IOPS SSD volumes are designed to meet the needs of I/O-intensive workloads
* 



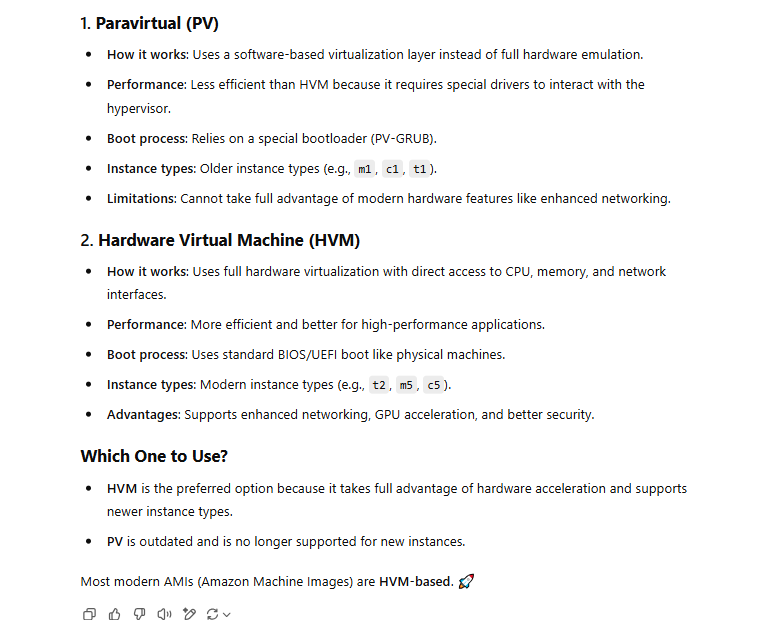
### Enhanced networking on Amazon EC2 instances

Link: <https://docs.aws.amazon.com/AWSEC2/latest/UserGuide/enhanced-networking.html>

Enhanced networking uses single root I/O virtualization (SR-IOV) to provide high-performance networking capabilities on supported instance types. SR-IOV is a method of device virtualization that provides higher I/O performance and lower CPU utilization when compared to traditional virtualized network interfaces. Enhanced networking provides higher bandwidth, higher packet per second (PPS) performance, and consistently lower latency between instances. There is no additional charge for using enhanced networking.

* Instances be launched from an HVM AMI (not PV)
* Is only supported in a VPC

### EC2 PV and HVM



### EC2 - Bastion Host

A Bastion Host is a secure instance that is used to access instances in a private subnet. The Bastion Host acts as a gateway to the private subnet and usually has minimal security exposure.

* 1 Computer được cấu hình đặc biệt, thuộc miền external/ public (DMZ) hoặc bên ngoài firewall, hoạt động như một server trung gian, cho phép bạn connect vào các Instance nằm trong Private Subnet
* Trường hợp Instance bị terminated, nhưng Auto Scaling Group đang launches, thì Elastic IP sẽ được đính lên cho instances mới

## VPC

### VPC Peering

### is used to enable private connectivity between VPCs within the same region or across different regions. It allows instances in one VPC to communicate with instances in another VPC using private IP addresses.

## EBL (ELATIC LOAD\_BALANLING)

### **Types of AWS Load Balancers**

AWS offers different types of load balancers based on the use case:

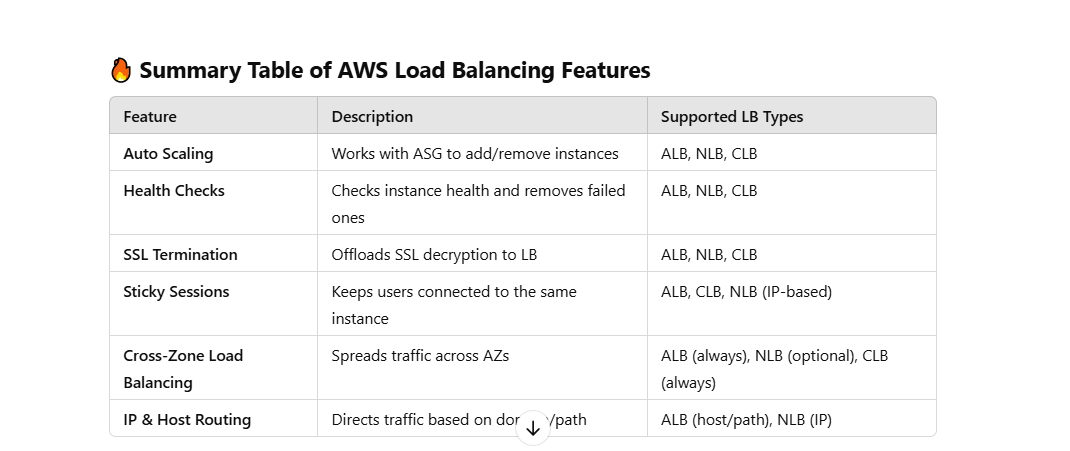
| **Load Balancer** | **Protocol** | **Best For** |
| --- | --- | --- |
| **Application Load Balancer (ALB)** | HTTP, HTTPS | Web applications (Layer 7 - Application layer) |
| **Network Load Balancer (NLB)** | TCP, UDP, TLS | High-performance, low-latency traffic (Layer 4 - Transport layer) |
| **Gateway Load Balancer (GWLB)** | IP | Third-party security and network appliances |
| **Classic Load Balancer (CLB)** (Legacy, not recommended) | HTTP, HTTPS, TCP | Basic load balancing |

### **2️⃣ How Load Balancers Work**

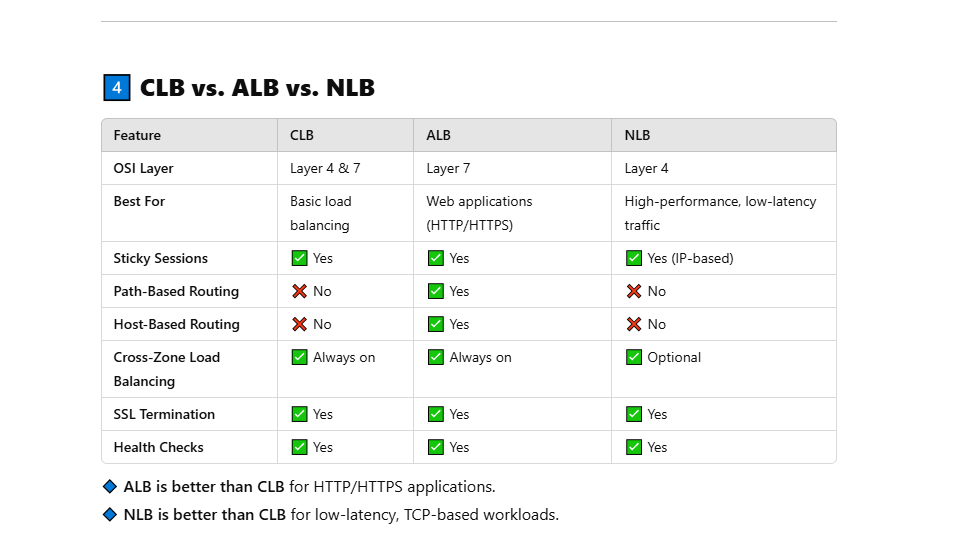
1. **Receives Requests**: A load balancer receives incoming traffic.
2. **Distributes Traffic**: It distributes requests across multiple backend instances.
3. **Performs Health Checks**: Ensures only healthy instances receive traffic.
4. **Handles Scaling**: Works with **Auto Scaling Groups** to dynamically add/remove instances.

### **3️⃣ Features of AWS Load Balancing**

✅ **Automatic Scaling**: Works with Auto Scaling Groups to handle traffic spikes.  
✅ **Health Checks**: Detects unhealthy instances and routes traffic only to healthy ones.  
✅ **SSL Termination**: Offloads SSL/TLS processing to the load balancer (ALB & NLB).  
✅ **Sticky Sessions**: Ensures a user session is always routed to the same instance.  
✅ **Cross-Zone Load Balancing**: Distributes traffic evenly across AZs.  
✅ **IP-based or Hostname-based Routing**: ALB allows routing based on host/path rules.



### Type of load balanling



## RedShift (fully managed)

### **📌 Ad-Hoc Queries in Amazon Redshift**

🔹 **Ad-hoc queries** in Amazon Redshift are **one-time, exploratory SQL queries** that users run **on-demand** to analyze data **without predefined reports or dashboards**. These queries allow users to investigate trends, perform quick data checks, and gain insights without extensive ETL or pre-processing.

### **📌 Amazon Redshift Spectrum:**

#### **Query Data in S3 Without Loading It**

**Amazon Redshift Spectrum** is a feature that allows you to **query data directly from Amazon S3** **without loading it into Redshift**. It extends Redshift’s capabilities by allowing you to run **SQL queries on structured and semi-structured data** stored in S3, **without needing to store the data inside Redshift**.

🚀 **Main Benefit:** You can analyze vast amounts of data in S3 **without using Redshift storage**, reducing costs while maintaining performance.

#### **Key Features of Redshift Spectrum**

✅ **Query data in S3 without moving it into Redshift**  
✅ **Supports various file formats** (CSV, JSON, Parquet, ORC, Avro)  
✅ **Works with Athena and Glue** for metadata management  
✅ **Joins S3 data with Redshift tables**  
✅ **Pay only for the amount of data scanned**

#### **🚀 When to Use Redshift Spectrum**

✅ You have **large datasets** in S3 that you don't want to move into Redshift.  
✅ You need to **analyze infrequent or historical data** without increasing Redshift storage costs.  
✅ You want to **combine S3 and Redshift data** in a single SQL query.  
✅ You prefer **serverless data analysis** rather than managing a large Redshift cluster.

### **Redshift Data Sharing (Share Data Across Clusters)**

✅ **Share data across multiple Redshift clusters** **without duplication**.  
✅ Allows **real-time analytics** across departments, teams, or businesses.  
✅ Supports **cross-account data sharing** in AWS Organizations.

### **Amazon Redshift Serverless (Auto-Scaling, No Cluster Management)**

✅ **Fully managed serverless Redshift** → No need to provision clusters.  
✅ **Automatically scales compute resources** based on query demand.  
✅ **Pay only for the compute used** (no idle cluster costs).  
✅ Ideal for **event-driven analytics and unpredictable workloads**.

### **IAM**

An IAM user with full administrator access can perform almost all AWS tasks except a few tasks designated only for the root account user.

Some of the AWS tasks that only a root account user can do are as follows :

+ change account name or root password or root email address

+ change AWS support plan

+ close AWS account

+ enable MFA on S3 bucket delete

+ create Cloudfront key pair

+ register for GovCloud.

# II.Questions

## 2.1. Elastic Load Balancing supports which of the following types of load balancers? (Choose 3 answers)

* Internet-facing
* Internal
* Hypertext Transfer Protocol Secure (HTTPS) using Secure Sockets Layer (SSL)

##### **What is a network ACL in Amazon VPC?**

##### A network ACL is a firewall that controls traffic to and from subnets in a VPC. network ACL is a firewall that controls tr

## How can you speed up the propagation of a DNS change in Amazon Route 53?

##### affic to and Reduce the TTL value of the DNS record.

##### rom subnets in a VPC.

The TTL (Time to Live) value determines how long a DNS record can be cached by resolvers. By reducing the TTL value, you can speed up the propagation of a DNS change. However, decreasing the TTL value also increases the load on the DNS servers and increases the risk of DNS amplification attacks.

## What is the purpose of an alias record in Amazon Route 53?

To map a DNS name to an AWS resource, such as an ELB load balancer or an S3 bucket