1. What is cloud computing ? what are the advantages compare to on-premises?
2. What are all the aws service do you know?
3. What are all the services have you used in your project?
4. What is difference between cp and sync in aws s3 cli ?
5. What is aws cli ?
6. How to configure aws cli ?
7. What is lambda function in aws ? have you written any lambda funtions ?
8. What is auto-scalling ? types of autoscalling ?
9. When we will use autoscalling in your project ?
10. What is reagion and available zone?
11. How you can make it your application high available ?
12. What is ecr ?
13. What is eks in aws ? have you experience in aws eks?
14. What are the tools we can use to implement ci/cd in aws services?
15. If you configure load balancer with two ec2 instances one instance went down. What load balancer will do?
16. What type databases aws is providing?
17. Which type of RDS have you used in your project ?
18. What are the benefits we will get if we use RDS instead of manual database setup in servers ?
19. If my instance pem file missed how can I access the instance ?
20. What is elastic beanstalk
21. What is cross origin access in s3.
22. What is public ip and privateip and elastic ip

EC2 Interview theory –

1. What are the different types of EC2 instances, and How do you choose the right one for your workload ?

* EC2 Instance Types (100+ as of March 2025) :-
  + General Purpose → Balanced workloads (t3, m6i, a1)
  + Compute Optimized → High CPU (c6g, c7i)
  + Memory Optimized → Large RAM (r5, x2idn)
  + Storage Optimized → High-speed storage (i4i, d2)
  + Accelerated Computing → AI/ML, GPUs (p5, g5)
  + HPC → Parallel computing (hpc6id, c6gn)
* Choosing the Right Instance
  + Define Needs: CPU, RAM, storage, network, GPU.
* Match to Family:
  + Web/Dev/Test → t3, m6i
  + Compute-heavy → c6g, c7i
  + In-memory DB → r5, x2idn
  + Big Data → i4i, d2
  + AI/ML → p5, g5
  + HPC → hpc6id, c6gn

1. What is the difference between On-Demand, Reserved, and Spot Instances ?

A black screen with white text

AI-generated content may be incorrect.

1. How does EC2 Auto Scaling work, and what are its benefits ?

* Define a Scaling Group → Specifies instances, AMI, VPC, and subnets.
* Set Scaling Policies → Determines when to add/remove instances.
* Dynamic Scaling → Adjusts based on demand (CPU, network, requests).
* Predictive Scaling → Uses ML to forecast traffic and scale in advance.
* Monitor & Automate → Continuously tracks instance health and replaces unhealthy instances.

A screenshot of a computer

AI-generated content may be incorrect.

1. What is an EC2 Placement Group, and what types are available?

A screenshot of a computer program

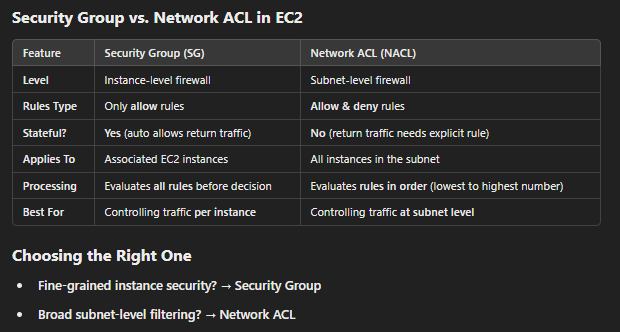
AI-generated content may be incorrect.

1. What are EC2 instance metadata and user data ? How can they be accessed ?

A screenshot of a computer

AI-generated content may be incorrect.

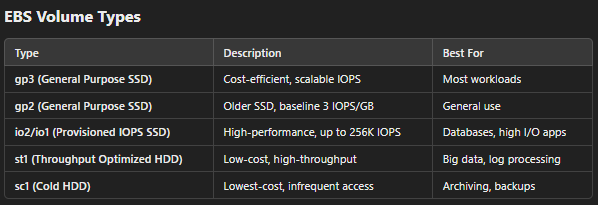
1. What is the difference between a security group and a network ACL in EC2?



1. How does an EC2 instance store data, and what are the differences between instance store and EBS volumes?

A black screen with white text

AI-generated content may be incorrect.



1. What is an EC2 Hibernate feature, and when would you use it ?

Saves the instance's RAM to EBS, allowing it to resume from the same state on restart.

1. How does EC2 support high availability and fault tolerance?

A black screen with white text

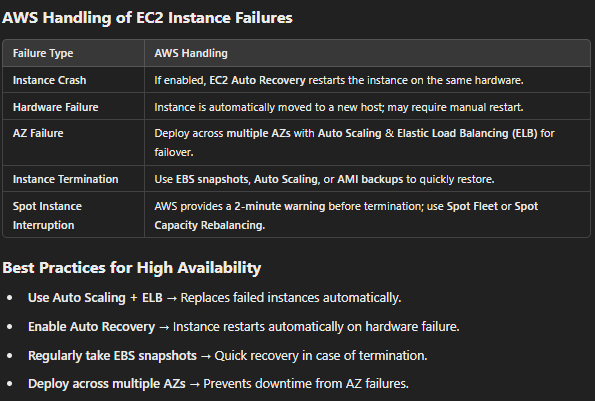
AI-generated content may be incorrect.

1. What is a dedicated host in EC2, and why would you use it ?

A screenshot of a computer program

AI-generated content may be incorrect.

1. How does AWS handle EC2 instance failures ?



1. What are the different tenancy options available for EC2 instances?

A black and white screen with white text

AI-generated content may be incorrect.

S3 :

1. What is Amazon S3, and how does it work?

Amazon S3 (Simple Storage Service) is a **scalable object storage** service for storing and retrieving data over the internet. It works by storing data as **objects inside buckets**, ensuring **high durability (11 9’s)** and **availability** across multiple Availability Zones.

1. What are the different S3 storage classes, and how do you decide which one to use ?

A screenshot of a computer

AI-generated content may be incorrect.

1. How does S3 ensure high availability and durability ?

S3 ensures high durability (99.999999999% or 11 9’s) by automatically replicating data across multiple Availability Zones (AZs). This redundancy prevents data loss due to hardware failures. For high availability (99.99%), S3 uses a distributed architecture that allows access even if one AZ goes down. It also performs automatic integrity checks and repairs corrupted data. Additional features like versioning, Cross-Region Replication (CRR), and failover mechanisms further enhance reliability.

1. What are S3 Access Points, and how do they improve access control?

Custom endpoints for managing access to an S3 bucket.

Each access point has its own permissions instead of relying on a single bucket policy.

How They Improve Access Control:

Simplified Permissions – Assign different rules for different users or applications.

Better Security – Restrict access to a specific VPC for private data access.

Scalability – Allow multiple teams or apps to use the same bucket with custom rules.

Reduced Complexity – No need to constantly edit the main bucket policy.

1. Explain the difference between S3 Object Lock and S3 Versioning.

What is Amazon S3, and how does it work ?

Prevents deletion or modification of objects for compliance.

Keeps multiple versions of an object to recover accidental changes

1. Explain about S3 objects and events ?

S3 Objects:

Basic storage units in S3, consisting of data, metadata, and a unique key (name).

Stored inside buckets and accessed via URLs, API, or CLI.

Can have different versions if S3 Versioning is enabled.

Metadata includes details like size, last modified date, and encryption settings.

S3 Events:

Triggered when specific actions happen in an S3 bucket (e.g., object creation, deletion).

Can notify AWS services like SNS, SQS, or Lambda for automation.

Processing newly uploaded files with Lambda.

Sending alerts when objects are deleted.

Automating data replication or backups.

1. How S3 Event Notifications Work ?

Event Triggers – S3 detects changes like object creation, deletion, or updates.

Notification Configuration – You set up event rules at the bucket level.

Destination Services – S3 sends notifications to SNS, SQS, or Lambda for further processing.

Common Use Cases:

Image/Video Processing – Trigger Lambda to process files when uploaded.

Data Backup & Replication – Notify another S3 bucket for replication.

Security Alerts – Send SNS alerts if a sensitive file is deleted.

Workflow Automation – Push events to SQS for queue-based processing.

Logging & Auditing – Record events in CloudWatch for monitoring.

1. What are S3 bucket policies and access control lists (ACLs) ?

A screenshot of a computer screen

AI-generated content may be incorrect.

1. What are the different ways to secure an S3 bucket ? How does AWS prevent public access to sensitive data in S3?

Block Public Access – Use "Block Public Access" settings to prevent accidental public exposure.

Bucket Policies – Define fine-grained access control rules using JSON-based policies.

IAM Policies – Restrict access using IAM roles, users, and groups with the least privilege principle.

ACLs (Access Control Lists) – Manage access at the object level (less recommended than bucket policies).

S3 Encryption – Encrypt data using SSE-S3, SSE-KMS, SSE-C, or client-side encryption.

MFA Delete – Enable Multi-Factor Authentication (MFA) to prevent accidental or malicious deletions.

VPC Endpoints – Restrict S3 access to private networks (no internet exposure).

S3 Access Logs – Enable logging and monitoring using AWS CloudTrail & S3 Access Logs.

Object Lock & Versioning – Protect against accidental deletions and ransomware attacks.

Restrict Cross-Origin Requests (CORS) – Limit which domains can request S3 resources.

1. How does AWS prevent public access to sensitive data in S3?

An S3 pre-signed URL is a temporary, secure link that allows users to upload or download objects from an S3 bucket without needing direct S3 permissions. The URL includes a signature and expiration time, ensuring controlled access.

Secure File Sharing – Allow users to download private files without making the bucket public.

Temporary Uploads – Let users upload files without giving them direct S3 write permissions.

Time-Limited Access – Grant short-term access (e.g., URL expires in 10 minutes).

API-Based Access – Use in serverless applications where S3 access is controlled via signed URLs.

1. How does AWS Key Management Service (KMS) help encrypt S3 data ?

Encrypts Data at Rest – Uses KMS-managed keys to encrypt S3 objects before storing them.

Fine-Grained Access Control – Restricts who can encrypt/decrypt data using IAM policies.

Automatic Key Rotation – Regularly updates encryption keys for better security.

Audit & Monitoring – Logs all encryption/decryption actions in AWS CloudTrail for tracking.

Stronger Security – Provides an extra security layer beyond standard SSE-S3 encryption.

1. How to restrict people access aws services ?
2. If new employee join in your organization how you will provide aws account access to him ?
3. What is IAM?
4. What is difference between user and role ?

Roles used to grant access to AWS resources for a specific task or service. Instead of creating a user and assigning permissions, a role is created with specific permissions and is assumed by the service or resource that needs to access those permissions. This allows for more secure access management and helps to reduce the risk of unauthorized access.

Users in AWS are typically associated with an AWS account and are authenticated using a username and password. Users can be assigned permissions and access to AWS resources using AWS Identity and Access Management (IAM) policies. IAM policies define what actions users are allowed or denied to perform on AWS resources.

1. What is iam policy ?
2. Have you used aws cloud watch? What are the metric

monitor using cloud watch ?

* CPU Utilation , Network In/Out, Disk I/O, Request Count, Error Count,
* Latecy(This metric tracks the amount of time it takes for an application or a resource to respond to requests.)

1. How to monitor RAM utilization in aws cloud watch ?

Install and configure the CloudWatch agent on your Amazon EC2 instance

Install the CloudWatch agent is installed and configured

Once the CloudWatch agent is installed and configured,

you can use the Agent Metrics feature to monitor RAM utilization. Agent Metrics allow you to

collect and monitor custom metrics that are not available through CloudWatch natively.

To monitor RAM utilization using the CloudWatch agent, you will need to

create a custom metric filter that collects RAM utilization data. You can do this by editing the agent configuration file. (opt/aws/amazon-cloudwatch-agent/bin/config.json)

1. Your S3 costs are increasing. What steps would you take to optimize costs ?

Use the Right Storage Class – Move infrequently accessed data to S3 Intelligent-Tiering, S3 Standard-IA, or Glacier.

Enable Lifecycle Policies – Automatically delete or transition old data to lower-cost storage.

Delete Unused Objects – Identify and remove stale or unnecessary objects.

Compress and Optimize Files – Store data in compressed formats (e.g., GZIP, Parquet).

Use S3 Intelligent-Tiering – Automatically move data between storage classes based on usage.

Minimize API Requests – Batch operations and use S3 Transfer Acceleration only when needed.

Use S3 Replication Wisely – Avoid unnecessary Cross-Region Replication unless required for disaster recovery.

Enable Object Expiration – Set auto-deletion policies for temporary files.

Monitor Usage with AWS Cost Explorer – Identify high-cost buckets and optimize usage.

Leverage S3 Requester Pays – If applicable, make data consumers pay for access.

1. How do S3 Lifecycle Policies help in managing storage costs?

How S3 Lifecycle Policies Help Manage Storage Costs

Automate Data Movement – Moves objects to cheaper storage classes like S3 Standard-IA, Glacier after a set time.

Auto-Delete Old Data – Removes unneeded objects, reducing storage costs.

Optimize Infrequent Access – Moves less-used files to S3 Intelligent-Tiering for cost savings.

Reduce Manual Work – Eliminates the need for manual monitoring and data transfers.

Control Versioning Costs – Deletes older versions of objects, preventing excess storage use.

1. What is S3 Intelligent-Tiering ?

Unpredictable Access Patterns – When you don’t know how often data will be accessed.

Long-Term Data Storage – Keeps data cost-efficient while ensuring quick access.

Reducing Storage Costs – Avoids manual data transfers while optimizing expenses.

Workloads with Changing Usage – Great for machine learning, logs, and backups.

Auto Cost Optimization – Moves data between storage tiers based on access frequency.

No Retrieval Fees – Unlike S3 Standard-IA, you pay only for storage, not access.

Four Tiers – Frequent, Infrequent, Archive, and Deep Archive Access tiers.

Low Latency – Same performance as S3 Standard for frequently accessed data.

No Manual Intervention – AWS automatically moves objects based on usage.

1. Ways to Reduce API Request Costs in S3

Batch Operations – Use S3 Batch Operations instead of making multiple API requests.

Use Larger Object Sizes – Reduce the number of PUT/GET requests by storing larger files

Enable S3 Transfer Acceleration Wisely – Use it only when needed to avoid extra costs.

Cache Data with CloudFront – Reduce direct requests to S3 by using Amazon CloudFront as a CDN.

Use S3 Inventory – Instead of frequent LIST requests, use S3 Inventory Reports to track objects.

Avoid Unnecessary HEAD/LIST Requests – Minimize API calls by keeping track of object metadata.

Use Object Lifecycle Policies – Automatically delete old or infrequently accessed data to prevent excessive requests.

Leverage S3 Intelligent-Tiering – AWS moves data across tiers without frequent LIST or GET requests.

1. What is S3 Intelligent-Tiering ?

Large File Uploads – Speeds up transfers for big data, media files, and backups.

Global Users – Improves upload speed for users far from the S3 bucket’s region.

Latency-Sensitive Applications – Ideal for real-time data processing and media streaming.

Remote Teams & Distributed Workloads – Helps globally distributed teams upload faster.

Unstable Network Conditions – Improves performance in high-latency or unreliable networks.

1. How does S3 Multi-Part Upload improve upload performance ?

Faster Uploads – Splits large files into smaller parts that are uploaded simultaneously.

Resumable Uploads – If a part fails, only that part is retried instead of restarting the entire upload.

Handles Large Files Efficiently – Required for files larger than 5GB and recommended for 100MB+.

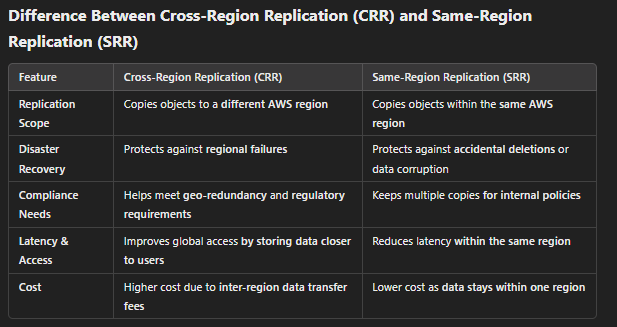
Better Network Utilization – Uses parallel connections, reducing the impact of network latency.

Efficient Data Transfer – After all parts are uploaded, S3 reassembles them into a single object.

1. What is S3 replication, and how does it help in disaster recovery ?

S3 Replication is a feature that automatically copies objects from one S3 bucket to another, either in the same AWS region (SRR) or a different region (CRR).

1. What is S3 replication, and how does it help in disaster recovery ?



1. How do you back up and restore S3 data ?

**Backup Methods:**

**S3 Replication (CRR/SRR)** – Automatically copies data to another bucket for redundancy.

**S3 Versioning** – Stores multiple versions of an object to prevent accidental deletion.

**AWS Backup** – Centrally manages and automates backups across AWS services, including S3.

**Manual Copy** – Use AWS CLI, SDKs, or S3 Batch Operations to duplicate data to another bucket.

**Glacier for Archival** – Move infrequently accessed data to S3 Glacier for cost-efficient backups.

**Restore Methods:**

**Recover from Versioning** – Restore an older version if a file is accidentally modified/deleted.

**Retrieve from Replicated Bucket** – Access backup copies from a secondary S3 bucket.

**Restore from Glacier** – Use S3 Restore API to retrieve archived objects.

**AWS Backup Restore** – Recover data from automated AWS Backup snapshots.

**Re-upload from Local Storage** – If backed up externally, upload files back to S3.

1. How do you back up and restore S3 data ?

S3 Lifecycle Policies – Automatically delete objects after a specified period.

S3 Object Expiration – Removes objects based on age to save storage costs.

S3 Versioning with Lifecycle Rules – Deletes old versions of objects to prevent unnecessary storage use.

AWS Lambda with Event Triggers – Triggers automatic deletion based on custom conditions.

S3 Intelligent-Tiering Auto-Deletion – Moves objects to cheaper storage and eventually deletes them.

S3 Batch Operations – Runs bulk delete jobs on selected objects.

1. If an S3 bucket accidentally gets deleted, how would you recover the lost data ?

**If Versioning Was Enabled:-**

Restore Object Versions – If S3 Versioning was enabled, you can recover older versions of deleted objects.

Check Replicated Buckets – If S3 Replication (CRR/SRR) was set up, retrieve data from the replicated bucket.

**If Backups Were Taken:-**

Restore from AWS Backup – If AWS Backup was used, recover the bucket and its contents.

Check Glacier Storage – If data was archived in S3 Glacier, restore it using S3 Restore API.

**If No Backups Exist:-**

Bucket Itself Cannot Be Recovered – Once deleted, the bucket name is permanently lost, and a new bucket must be created.

Contact AWS Support – AWS cannot recover deleted data unless backups, versioning, or replication were enabled.

NAT (Network Address Translation) gateway is a managed service that

allows instances in a private subnet to access the internet or other

AWS services while also providing a level of security by hiding the private IP addresses from the public internet.

AWS services while also providing a level of security by hiding the private IP addresses from the public internet.

1. What is difference between NAT instance vs NAT gateway ?

NAT gateway is a fully managed service provided by AWS,

while NAT instance requires you to launch and manage an EC2 instance yourself.

Availability: NAT gateway is highly available across multiple availability zones within a region,

while NAT instance requires you to set up redundancy and failover across multiple instances or availability zones manually.

Scalability: NAT gateway is automatically scaled by AWS based on the traffic needs,

while NAT instance requires manual scaling and configuration.

Performance: NAT gateway is designed to provide higher performance than

NAT instance and can handle higher network throughput.

Security: NAT gateway uses an Elastic IP address and is automatically configured to be more secure than

NAT instance, which requires manual configuration of security groups, network ACLs, and firewall rules.

1. What is Internet gateway ?

communication between instances in your VPC (Virtual Private Cloud) and the internet.

It is important to note that an internet gateway is a one-way communication channel:

it allows instances in your VPC to connect to the internet,

but it does not allow the internet to initiate connections to instances in your VPC.

For incoming traffic, you can use other AWS services, such as Elastic Load Balancing or Amazon Route 53, to direct traffic to instances in your VPC.

1. What is route table ?

In AWS, a route table is a virtual table that is associated with a subnet in a

VPC (Virtual Private Cloud) and is used to determine where network traffic is directed.

Route tables are an important component of AWS networking that allow you to control how traffic flows within your VPC

and between your VPC and other AWS services or the internet. By creating and managing route tables, you can direct traffic to the appropriate destinations

and ensure that your applications and services are accessible and performant.

1. How to access private subnet instances ?

* Need to attach (release in public public) NAT gateway in private subnets

🡪 then conn via public instance

1. What is bastion host ?

In AWS, a bastion host, also known as a jump box or a jump server,

is a specially configured EC2 instance that is designed

to provide secure access to instances in private subnets that are not directly accessible from the internet.

1. What elb ?

In AWS, ELB stands for Elastic Load Balancer, which is a managed service that automatically distributes incoming application traffic across multiple targets, such as EC2 instances, containers, and IP addresses, in one or more Availability Zones.

1. How many types of load balancer are there in aws ?

3(CLB, NLB, ALB)

1. What is difference between Application load balancer and network load balancer & Classic load Blancer ?

30) How to transfer data from one S3 bucket to another?

**a) Copy All Files (Recursive) -**

**b) Sync Buckets -**

|  |
| --- |
| aws s3 sync s3://source-bucket/ s3://destination-bucket/ |

**c) Copy All Files (Recursive) -**

* Using Boto3 (Python SDK)

AWS Lambda –

*AWS Lambda is like a magic box that runs your code only when something happens (like a new file uploaded to S3 or a new request to an API). You don’t have to worry about setting up or managing servers. When an event triggers Lambda, it quickly sets up an environment to run your code, executes it, and then cleans up everything once done. It can handle lots of requests at once, automatically scaling, and you only pay for the time your code runs. AWS handles everything on the backend, including scaling, error handling, and logging, so you don’t need to think about infrastructure!*

Let me know if you want more details!

1. Basic Questions

- What is AWS Lambda, and how does it work?

AWS Lambda is a serverless compute service that runs code in response to events, without the need for provisioning or managing servers.

- What are the key benefits of using AWS Lambda?

Lambda provides auto-scaling, high availability, and pay-per-use pricing, allowing developers to focus on code without managing infrastructure.

- What are the supported runtimes for AWS Lambda?

AWS Lambda supports Python, Node.js, Java, Go, .NET Core, Ruby, and custom runtimes.

- How does AWS Lambda scale?

Lambda automatically scales by running code in parallel for each event, handling multiple requests simultaneously without manual intervention.

- What is the maximum execution timeout for a Lambda function?

The maximum execution timeout for a Lambda function is 15 minutes.

---

2. Architecture & Performance

- How does AWS Lambda handle concurrency?

Lambda functions are invoked concurrently, and AWS manages the scaling to handle multiple requests in parallel.

- What is a cold start, and how can you minimize it?

A cold start occurs when a new Lambda instance is created, introducing latency; you can minimize it by using provisioned concurrency or optimizing the function code.

- What are Lambda execution environments?

Lambda execution environments are isolated environments where Lambda functions run, including the runtime, libraries, and temporary storage.

- How does Lambda pricing work?

Pricing is based on the number of requests and the duration of the code execution, measured in 100ms increments.

- How do you optimize AWS Lambda performance?

Optimize by minimizing package size, reducing cold starts, and using provisioned concurrency for critical functions.

---

3. Integrations & Event Sources

- How do you trigger Lambda from an S3 bucket?

Lambda can be triggered by an S3 event (e.g., object creation) through an S3 bucket notification.

- How do you use AWS API Gateway with Lambda?

API Gateway routes HTTP requests to Lambda, allowing you to build RESTful APIs with Lambda functions.

- How do you integrate Lambda with DynamoDB streams?

Lambda can process DynamoDB stream records to react to data changes automatically.

- How does Lambda work with AWS Step Functions?

Step Functions allow Lambda functions to be orchestrated into workflows with error handling and retries.

- What is the difference between synchronous and asynchronous invocation?

Synchronous invocation waits for the function to complete and returns the result; asynchronous invocation immediately returns and processes in the background.

---

4. Security & Permissions

- How do you secure an AWS Lambda function?

Use IAM roles with the least privilege, encrypt environment variables, and ensure network security with VPCs and security groups.

- What IAM permissions are required for Lambda execution?

Lambda requires an IAM execution role with permissions to interact with other AWS services, like S3, DynamoDB, etc.

- How do you restrict access to a Lambda function?

You can restrict access using IAM policies, resource-based policies, and VPC-based access control.

- How do you manage secrets in AWS Lambda?

Use AWS Secrets Manager or AWS Systems Manager Parameter Store to securely manage and retrieve secrets.

- What is a VPC-enabled Lambda function?

A VPC-enabled Lambda function runs inside a VPC, allowing it to access resources like RDS and EC2 instances within the VPC.

---

5. Deployment & Monitoring

- How do you deploy a Lambda function using AWS CLI?

Use the `aws lambda create-function` command along with the ZIP file containing the function code and necessary permissions.

- What are AWS Lambda layers, and how do you use them?

Lambda layers allow you to package and share common libraries and dependencies across multiple Lambda functions.

- What monitoring tools can you use for AWS Lambda?

AWS CloudWatch provides logs, metrics, and alarms to monitor Lambda function performance and behavior.

- How do you enable logging in AWS Lambda?

Enable logging by using the `logging` module in Python or the `console.log` in Node.js, and the logs will be sent to CloudWatch.

- What is AWS X-Ray, and how does it help with Lambda debugging?

AWS X-Ray helps trace and analyze requests, identify performance bottlenecks, and troubleshoot Lambda function execution.

---

6. Advanced Topics

- What are the limitations of AWS Lambda (memory, timeout, execution)?

Lambda functions can have a maximum memory of 10GB, a timeout of 15 minutes, and an execution environment with limited disk space (512MB).

- How do you handle retries and error handling in AWS Lambda?

Lambda automatically retries failed asynchronous invocations, and you can configure dead-letter queues (DLQ) or use Step Functions for error handling.

- How do you use Provisioned Concurrency?

Provisioned concurrency pre-warms Lambda instances, ensuring predictable start-up times and avoiding cold starts for critical functions.

- What is the difference between AWS Lambda and Fargate?

Lambda is for serverless, event-driven compute, whereas Fargate is a serverless compute engine for containers that offers more control over the environment.

- How do you implement blue/green deployment for Lambda?

Use Lambda aliases and versions to manage blue/green deployments, allowing for gradual traffic shifting between versions.

---

Let me know if you'd like to dive deeper into any of these topics!

Protocols & Layers & Key Feature

Classic Load Balancer:

This is the original AWS load balancer and is ideal for simple use cases that require balancing of traffic across multiple EC2 instances. It supports both HTTP and HTTPS protocols, as well as TCP and SSL connections, and can

distribute traffic using round-robin or least-connections algorithms. It can also automatically scale to handle increases in traffic. This is a layer (4) (Transport Layer) load balancer

Application Load Balancer (ALB):

This is a layer 7 (APPLICATION LAYER) load balancer that is designed to route traffic to different targets based on the content of the request. It supports HTTP and HTTPS protocols and can distribute traffic using URL-based routing, host-based routing, or path-based routing. It also provides advanced features such as content-based routing, routing based on HTTP headers, and support for WebSockets.

Network Load Balancer (NLB):

This is a layer 4 load balancer that is designed to handle high levels of traffic and provide ultra-low latency performance. It supports TCP, TLS, and UDP protocols and can distribute traffic using source IP address affinity or destination IP address affinity. It also provides advanced features such as support for static IP addresses, integration with AWS PrivateLink, and support for proxy protocol v2.

1. What is route53 ?

Overall, Route 53 is a reliable and scalable DNS service that can help improve the availability and performance of applications and services hosted on AWS or other infrastructure.

1. What is recordset ?

record sets are a fundamental part of managing DNS in Amazon Route 53 and are *used to map domain names to various types of resources*, such as IP addresses, load balancers, and other AWS resources.

A (Address) Record: This maps a domain name to an IPv4 address.

AAAA (IPv6 Address) Record: This maps a domain name to an IPv6 address.

CNAME (Canonical Name) Record: This creates an alias for a domain name, allowing it to resolve to another domain name.

1. What is ttl ?

"Time to Live" and refers to the length of time that a DNS resolver is allowed to cache or store a DNS query before it must request the information again from the authoritative DNS server.

1. How may type of routing policies are there?

6 (Simple Routing Policy, Weighted, Latency, Failover, Geolocation)

Simple Routing Policy: This sends all traffic to a single resource, such as a web server or an IP address.

Weighted Routing Policy: This sends traffic to multiple resources in proportions that are specified by assigning different weights to each resource.

Latency Routing Policy: This sends traffic to the resource that has the lowest network latency (i.e., the shortest round-trip time) to the user.

Failover Routing Policy: This sends traffic to a primary resource, and if the resource becomes unavailable, it automatically switches to a secondary resource.

Geolocation Routing Policy: This sends traffic to resources based on the geographic location of the user.

Multivalue Answer Routing Policy: This returns multiple values in response to a DNS query and allows Route 53 to respond with multiple IP addresses for a domain name.

1. Advantages of route53 ?

High Availability and Scalability, Fast and Reliable DNS Resolution, Flexible Routing Policies, Health Checks and Failover, Integration with Other AWS Services.

1. What is s3 ?
2. What are the types of storage classes ?

S3 Standard-Infrequent Access (S3 Standard-IA),S3 One Zone-IA,

S3 Glacier,S3 Glacier Deep Archive, Amazon EBS (Elastic Block Store)

VPC:

1. What are the differences between a Security Group and a Network ACL?

Answer:- Security Groups are stateful (response traffic is automatically allowed), operate at instance level. NACLs are stateless (response rules must be explicitly allowed), operate at subnet level. SG is easier for fine-grained access; NACLs handle broader subnet-wide control.

2. Can a subnet span multiple Availability Zones? Why or why not?

Answer:- No, a subnet is restricted to a single AZ. AWS designs it this way to maintain fault isolation and high availability across AZs. You must create separate subnets in each AZ for redundancy.

3. What is the maximum number of VPCs per region? How can it be increased?

Answer: Default is 5 VPCs per region. You can request a limit increase by raising a service quota request in AWS Support Center. Limits vary depending on the AWS account and service agreements.

4. Explain the use cases of Transit Gateway vs VPC Peering.

Answer: VPC Peering is best for 1-to-1 VPC connections (simple environments). Transit Gateway simplifies many-to-many VPC and on-premises connectivity, highly scalable. TGW is better for hub-and-spoke architectures.

5. How do Route Tables work inside a VPC?

Answer: Route Tables control how traffic is directed within a VPC. Each subnet must be associated with a Route Table. Routes determine traffic flow, e.g., local, NAT, Internet Gateway.

6. What is a VPC Endpoint? Difference between Interface and Gateway Endpoints?

Answer: VPC Endpoint allows private connections to AWS services without traversing the internet. Interface Endpoint uses ENIs for services like SSM, Secrets Manager. Gateway Endpoint is used for S3 and DynamoDB with route table entries.

7. What happens when you associate an Elastic IP to an instance in a private subnet?

Answer: It won't work directly. A private subnet lacks direct internet routing, even with an Elastic IP. Instance must be in public subnet + route via IGW for EIP to work.

8. What is the default behavior of a newly created Security Group?

Answer: All inbound traffic is denied by default. All outbound traffic is allowed by default. You must explicitly allow inbound rules for communication.

9. Can two EC2 instances in different subnets communicate without a NAT Gateway?

Answer: Yes, if both subnets are in the same VPC and security groups + route tables allow it. NAT Gateway is only needed for internet access, not internal VPC communication.

10. How would you secure VPC resources from external and internal threats?

Answer: Use Security Groups, NACLs, and VPC Flow Logs. Enable private subnets for sensitive workloads. Use AWS WAF, Shield, and GuardDuty for advanced threat detection and protection.

Scenario-

1. How would you design a 3-tier app (Web, App, DB) using VPC, subnets, and security groups?

Answer: Create a VPC across 2+ AZs. Public subnet for Web tier (Load Balancer + EC2), private subnet for App tier, isolated private subnet for DB tier. Security Groups allow traffic only from specific layers (e.g., App tier can talk to DB tier, not public internet).

2. How do you connect your on-premises data center securely to your VPC?

Answer: Set up a VPN connection using Virtual Private Gateway or Transit Gateway. Alternatively, use AWS Direct Connect for low-latency, dedicated links. Ensure routes are updated in Route Tables and on-prem devices.

3. Two VPCs need to communicate securely. How would you set this up without using the public internet?

Answer: Create a VPC Peering Connection between the two VPCs. Update Route Tables to allow traffic through the peering link. Ensure there are no overlapping CIDRs between the VPCs.

4. You have private subnets that need internet access for software updates. How do you design this?

Answer: Deploy a NAT Gateway in a public subnet. Route private subnet traffic (0.0.0.0/0) to the NAT Gateway via the Route Table.Ensure Security Groups allow necessary outbound traffic.

5. How would you design a VPC for high availability across AZs?

Answer: Create subnets in at least two AZs. Deploy resources (EC2, RDS, etc.) with Multi-AZ or Auto Scaling Groups enabled. Use Elastic Load Balancers (ALB/NLB) across subnets for fault tolerance.

6. Your EC2 instances need to access S3 privately. How would you configure this?

Answer: Create a VPC Gateway Endpoint for S3. Update Route Tables in the private subnets to route S3 traffic through the Endpoint. No need for NAT or public internet exposure.

7. How would you extend your VPC to support both AWS and on-premises workloads?

Answer: Use AWS Site-to-Site VPN or Direct Connect. If multiple VPCs are involved, use Transit Gateway to manage connectivity centrally.Use shared services VPC patterns if needed.

8. How would you monitor and analyze traffic inside a VPC for security purposes?

Answer: Enable VPC Flow Logs to capture traffic metadata. Use CloudWatch Logs or S3 for storage and analysis. Integrate with GuardDuty and Security Hub for automated threat detection.

9. When and why would you separate resources into public vs private subnets?

Answer: Public subnet for resources needing internet access (e.g., ALB, bastion hosts). Private subnets for internal resources (e.g., App servers, databases). It enhances security and network segregation following best practices.

10. How would you enable IPv6 inside your VPC? What are the challenges?

Answer: Associate an IPv6 CIDR block with your VPC and subnets. Modify Route Tables to route traffic to the Internet Gateway. Challenges include managing dual-stack apps and updating security policies for IPv6.

Internet Gateway vs NAT Gateway ---

"IGW is for public-facing servers, NAT Gateway is for private servers needing only outbound access."

A screenshot of a computer

AI-generated content may be incorrect.

NAT Gateway vs NAT Instance –

NAT Gateway = production, managed.

NAT Instance = dev/test, manual management.

A screenshot of a computer

AI-generated content may be incorrect.