

**Q:** API Gateway + traffic spike =

**A:** Throttling

**Q:** What service lets you trace and analyze user requests as they travel through your API gateway to underlying services?

**A:** AWS X-Ray

**Q:** Memory Utilization, disk space, and disk swap are...

**A:** Custom CloudWatch metrics and you need to install CloudWatch on the EC2 instance to use them

**Q:** Small, random I/O operations vs. Large sequential I/O operations

**A:** SSD vs. HDD (magnetic)

**Q:** What can't HDD volumes do?

**A:** HDD volumes can't be used as boot volumes

**Q:** When you hear "parallel file system, high performance", think:

**A:** FSx for Lustre

**Q:** When you see SSH, think:

**A:** TCP

**Q:** Application Load Balancer vs. Network Load Balancer

**A:** HTTP/HTTPS and OSI level 7 vs. TCP/UDP, SSH and OSI level 4

**Q:** When you hear "path-based routing" or "host-based routing" think:

**A:** Application Load Balancer

**Q:** How to avoid accidental delete from S3?

**A:** Versioning and MFA delete

**Q:** When you see DDoS attack, think:

**A:** AWS Shield Advanced (backend/infrastructure level) and WAF (application level)

**Q:** When you hear "flexible schema", think:

**A:** NoSQL > DynamoDB

**Q:** When you hear "messaging without rewriting code" think:

**A:** Amazon MQ

**Q:** When you hear "API credentials" think:

**A:** IAM role

**Q:** Ongoing large data transfer vs. one-time large data transfer

**A:** DirectConnect vs. Snowball Edge

**Q:** What is the auto-scaling cooldown period?

**A:** 300 seconds by default, waits for previous scaling activity to take effect

**Q:** What is the SQS message retention period?

**A:** 1 minute to 14 days, 4 days by default

**Q:** Why might you get duplicate SQS messages?

**A:** They weren't deleted from the queue after processing

**Q:** Geolocation vs. Geoproximity Routing

**A:** Geolocation – route traffic based on the location of your users

Geoproximity – route traffic based on the location of resources/shift traffic from resources in one location to resources in another location

**Q:** What do you use to perform processing geographically close to users?

**A:** Lambda@Edge

**Q:** Container + open source/cloud agnostic =

**A:** Kubernetes

**Q:** When you need “temporary, limited-privilege tokens” think:

**A:** Amazon STS

**Q:** What is SES used for?

**A:** Transactional marketing emails, NOT notification emails

**Q:** What are the pre-requisites for routing traffic to a website hosted in an S3 bucket?

**A:** 1) the bucket has to have the same name as your domain name

2) a registered domain name

3) Route53 as your DNS service

**Q:** When you hear “big data processing frameworks” think:

**A:** Elastic Map Reduce (EMR)

**Q:** Can you use IP addresses for multi-AZ failover?

**A:** No, because IP addresses are per subnet. Use CNAME instead.

**Q:** What Load Balancer do you use with Elastic IP addresses?

**A:** Network Load Balancer

**Q:** Storage Gateway vs. DataSync

**A:** Storage Gateway is for hybrid storage, DataSync is for migrating existing data

**Q:** When you hear “serverless compute engine for containers” think:

**A:** Fargate

**Q:** What do you use to encrypt EBS volumes?

**A:** KMS

**Q:** How do you launch encrypted EBS storage from unencrypted snapshots?

**A:** EBS Encryption By Default

**Q:** What is a launch configuration?

**A:** A template used by an Auto Scaling group to launch instances and that can be used to change the AMI the EC2 instances are generated from

**Q:** What is the largest file size you can upload to S3 in one PUT request?

**A:** 5GB

**Q:** How can you speed up the upload of a large file in S3?

**A:** Multipart upload

**Q:** How do you speed up upload to S3 for users around the world?

**A:** Transfer Acceleration

**Q:** How many requests per second does S3 support?

**A:** 3,500 requests/second to put data and 5,500 requests/second to retrieve data

**Q:** What is the similarity and differentiator between RDS and DynamoDB?

**A:** Both are fully managed; the differentiator is scaling. DynamoDB scales automatically, RDS is scaled manually.

**Q:** What does ElastiCache do?

**A:** Caches query results

**Q:** What is the max retention period for an automated backup?

**A:** 35 days

**Q:** When you hear “high-performance computing (HPC)” think:

**A:** Elastic Fabric Adapter (EFA) **UNLESS** Windows is involved, then think Elastic Network Adapter (ENA) because EFA is NOT compatible with Windows

**Q:** Global client base + AnyCast Static IP =

**A:** Global Accelerator

**Q:** For detailed logging info for S3 object-level access you need:

**A:** Server access logging

**Q:** Simple Workflow (SWF) vs. Step Functions

**A:** SWF is fully managed, Step Functions are serverless

**Q:** What do you do if a newly launched EC2 instance doesn't have a DNS host name associated with it?

**A:** Enable DNS hostname and DNS resolution in the VPC configuration

**Q:** What is RAID 0?

**A:** Configuration to improve performance of instance store volumes

**Q:** What is RAID 1?

**A:** Configuration for traffic mirroring

**Q:** When you need high IOPs and high sequential read/write capability, what EC2 instance type do you need?

**A:** Storage optimized

**Q:** What are the enhanced monitoring metrics CloudWatch gathers from RDS databases?

**A:** OS processes and RDS child processes

**Q:** What is the maximum ratio of provisioned IOPs to volume size?

**A:** 50:1 in GiB, for example a 10GiB volume can have a max of 500 IOPs

**Q:** Best practices + real time guidance =

**A:** Trusted Advisor

**Q:** What is AWS Glue?

**A:** A fully managed ETL (extract, transform, load) tool

**Q:** How many days do you need to wait to transition data from S3 to IA or One-Zone IA?

**A:** 30 days

**Q:** When you hear "zone apex" think:

**A:** a record alias

**Q:** VPC Gateway Endpoint vs. VPC Interface Endpoint

**A:** S3 and DynamoDB use VPC Gateway Endpoint, all other AWS services use VPC Interface Endpoint

**Q:** What lets you remotely and securely manage configuration for instances or other service configured with AWS Systems Manager?

**A:** Run Command

**Q:** What is Visibility Timeout?

**A:** A period of time during which Amazon SQS prevents other consuming components from receiving and processing a message. Default is 30 seconds and maximum is 12 hours.

**Q:** If you have an encrypted EBS volume...

**A:** All data moving between the instance and the volume are encrypted and snapshots are automatically encrypted

**Q:** What does long polling do?

**A:** Reduces cost by reducing the number of empty responses when there are no messages to return and eliminates the number of empty responses by querying all the servers. SQS uses short polling by default.

**Q:** When don't EIPs incur charges?

**A:** When attached to a running instance that has only one EIP associated with it

**Q:** What is the max execution time for a Lambda function?

**A:** 15 minutes

**Q:** How do you improve performance for Kinesis Data Streams?

**A:** Increase the number of shards using the UpdateShardCount command

**Q:** When you need a DB w/ single-digit millisecond response times you need...

**A:** DynamoDB

**Q:** Which port is used for Windows RDP?

**A:** Port 3389

**Q:** What is the max time data stays in Kinesis Data Streams?

**A:** 24 hours

**Q:** When you need a storage service for HPC that also works natively with S3 you need...

**A:** FSx for Lustre

**Q:** What do you need to do to an S3 bucket to use Cross-Region Replication?

**A:** Enable versioning

**Q:** What is the only required section of a CloudFront template?

**A:** Resources

**Q:** If you want your instances to talk to the internet without being reachable from the internet...

**A:** IPv4 = NAT Gateway

IPv6 = egress-only internet gateway

**Q:** Where are log files from Elastic Beanstalk stored?

**A:** Application files in S3, server log files optionally in S3 or CloudWatch Logs

**Q:** What should you not use as the storage engine for MySQL?

**A:** MyISAM

**Q:** What do you do to restrict access to your RDS database?

**A:** Enable IAM DB Authentication

**Q:** What do you need to be very wary of in a policy?

**A:** A wildcard (\*) because it's probably overwriting another permission

**Q:** When DB credentials should be supplied using environment variables, what do you use? And what is the deciding factor if both are options?

**A:** AWS Systems Manager Parameter Store or AWS Secrets Manager secrets encrypted with KMS. Deciding factors I to use IAM roles for ECS task execution

**Q:** Relational DB + sporadic, unpredictable workloads + scaling =

**A:** Aurora Serverless

**Q:** How do you send notifications based on changes in an RDS DB?

**A:** Have an event subscription send a notif to SNS and have SNS fan out to multiple SQS queues

**Q:** How do you improve the performance of Dynamo DB?

**A:** Use partition keys with high-cardinality attributes which have a large number of distinct values for each item

**Q:** What do you use with a corporate Active Directory?

**A:** AWS Directory Service AD Connector

**Q:** What do you need to have a successful site-to-site VPN connection between the VPC and your data center?

**A:** A customer gateway with a static internet routable IP address of its external interface for the on-premises network

**Q:** What do you do if a shard iterator expires unexpectedly?

**A:** Increase the write capacity assigned to the shard table

**Q:** If the question doesn't mention serverless, the answer is NOT

**A:** Step Functions

**Q:** When do you use throughput optimized HDD vs. cold HDD?

**A:** Throughput optimized HDD is for frequently accessed data and cold HDD is used for infrequently accessed data.

**Q:** If an EC2 instance is in a public subnet but unreachable from the internet, it may not have...

**A:** A public IP address associated with it

**Q:** If the question asks for the storage solution with the lowest-latency access to the data, the answer is...

**A:** EBS because it's attached to the EC2 instance

**Q:** When you need to connect multiple on-premises networks, VPNs, and VPCs into a single gateway, you need...

**A:** AWS Transit Gateway

**Q:** If only a specific IP address should have access to something, what is the WRONG answer?

**A:** Anything with /0 because that refers to the entire network

**Q:** When you need to use a second authentication method with Cognito, add...

**A:** MFA (you can still use SMS as a method but you don't need another service involved)

**Q:** HPC + OS-bypass =

**A:** Elastic Fabric Adapter (EFA)

**Q:** "Automate and accelerate the replication of data" =

**A:** DataSync

**Q:** What is the most efficient way to collect and analyze logs from EC2 instances with minimal effort?

**A:** Install the CloudWatch Logs agent in each instance

**Q:** If you have a 3<sup>rd</sup> party SSL/TLS certificate, where can you import and store it?

**A:** AWS Certificate Manager and/or IAM Certificate Store

**Q:** When you need short-lived tokens for an RDS DB you need...

**A:** IAM DB Authentication

**Q:** What section of a CloudFormation template describes the values associated with your stack's properties?

**A:** Outputs

**Q:** When deciding between File Gateway vs. Volume Gateway...

**A:** NFS or SMB = File Gateway, iSCSI = Volume Gateway

**Q:** What are the three benefits of IAM DB Authentication?

**A:** 1) network traffic to and from the DB is SSL encrypted, 2) centrally manage DB access in IAM rather than in each instance individually, 3) can use profile credentials specific to your EC2 instance instead of a password

**Q:** For a company whose divisions want to maintain discrete control of their resources but also centrally manage billing, you need...

**A:** Separate accounts, not just separate VPCs in one account, and to use AWS Organizations

**Q:** Kinesis Data Firehose only supports...

**A:** S3, Redshift, ElasticSearch, and an HTTP endpoint as destinations

**Q:** When you need millisecond response times (aka DynamoDB as a destination) and processing for streaming data, you need...

**A:** Kinesis Data Streams

**Q:** What do you need for storing user session data with sub-millisecond latency for users?

**A:** ElastiCache

**Q:** When viewing a Route Table with a route for an internet gateway, what kind of entry should you see?

**A:** A target of igw-xxxx with a destination of 0.0.0.0/0 for IPv4 or a destination of ::/0 for IPv6

**Q:** What three things do you get with RDS Multi-AZ deployments?

**A:** 1) synchronously replicate data to a standby instance in a different AZ (but the same region), 2) increase availability in the event of system upgrades like OS patching or db scaling, 3) provide enhanced durability in case of DB instance failure or AZ outage

**Q:** What can you do with Scheduled Reserved Instances?

**A:** Purchase capacity on a recurring daily, weekly, or monthly basis with a specified start time and duration for a one-year term

**Q:** What do you need to remember about...

- Tape Gateway – doesn't give you fast access to files
- DynamoDB streams – it's not enabled by default
- SSL – only transit encryption
- IPv4 and IPv6 – 4 is required, 6 is optional
- Instance limits – there are limits to how many of each type of EC2 instance you can have, and you need to submit a form to AWS to get approved for more
- VPC peering – not supported in a Direct Connect connection
- SQS FIFO queue – has exactly-once processing
- CloudTrail logs – are encrypted by default

- EBS volumes when a snapshot is in progress – the EBS volume can still be used while a snapshot is in progress
- IAM – it's global (not restricted by region/AZ)
- AWS Elastic Beanstalk – it can launch applications from containers
- EFS vs. FSx – when a question specifies the need for flexible managed storage that can be accessed by instances in multiple AZs it's EFS **UNLESS** Windows/Active Directory is specified in which case it's FSx for Windows **OR** native integration with S3 is specified in which case it's FSx for Lustre
- SNS vs. SQS – SNS is subscription, SQS is polling
- Aurora Endpoints – you can use endpoints to map connections to specific instances or clusters of instances, and you can use custom endpoints for load balancing based on criteria other than read-only or read-write capability
- NACL by default – allow all traffic