Practice Exam - AWS Certified Solutions Architect Associate - Results

Attempt 1

Question 1 Correct

A retail organization is moving some of its on-premises data to AWS Cloud. The DevOps team at the organization has set up an AWS Managed IPSec VPN Connection between their remote on-premises network and their Amazon VPC over the internet.

Which of the following represents the correct configuration for the IPSec VPN Connection?

Create a virtual private gateway (VGW) on both the AWS side of the VPN as well as the on-premises side of the VPN

Create a virtual private gateway (VGW) on the on-premises side of the VPN and a Customer Gateway on the AWS side of the VPN

Your answer is correct

Create a virtual private gateway (VGW) on the AWS side of the VPN and a Customer Gateway on the on-premises side of the VPN

Create a Customer Gateway on both the AWS side of the VPN as well as the on-premises side of the VPN

Overall explanation

Correct option:

Create a virtual private gateway (VGW) on the AWS side of the VPN and a Customer Gateway on the on-premises side of the VPN

Amazon VPC provides the facility to create an IPsec VPN connection (also known as AWS site-to-site VPN) between remote customer networks and their Amazon VPC over the internet. The following are the key concepts for a site-to-site VPN:

Virtual private gateway: A virtual private gateway (VGW), also known as a VPN Gateway is the endpoint on the AWS VPC side of your VPN connection.

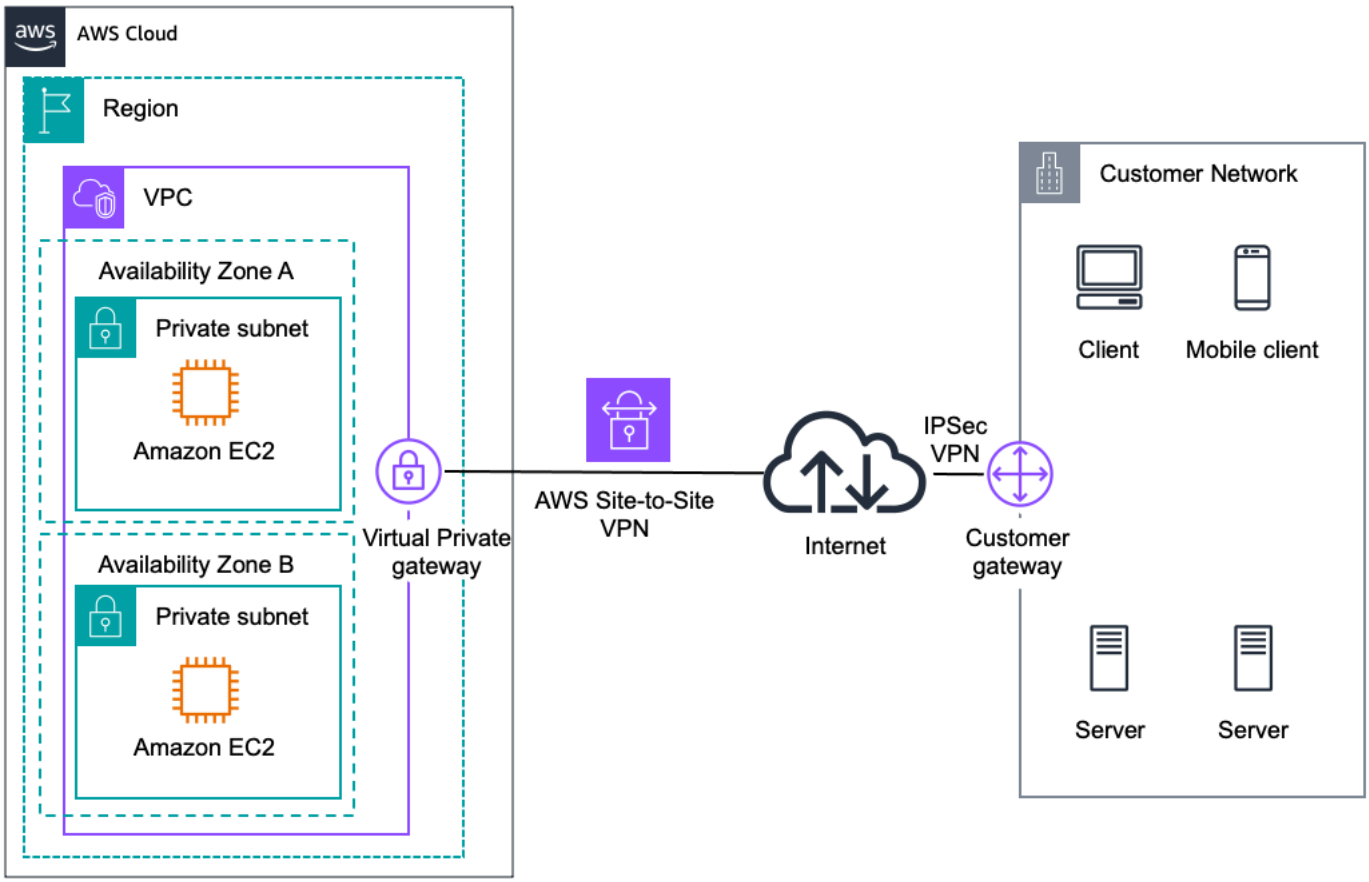
VPN connection: A secure connection between your on-premises equipment and your VPCs.

VPN tunnel: An encrypted link where data can pass from the customer network to or from AWS.

Customer Gateway: An AWS resource that provides information to AWS about your Customer Gateway device.

Customer Gateway device: A physical device or software application on the customer side of the Site-to-Site VPN connection.

AWS Managed IPSec VPN



via - https://docs.aws.amazon.com/whitepapers/latest/aws-vpc-connectivity-options/aws-managed-vpn-network-to-amazon.html

Incorrect options:

Create a virtual private gateway (VGW) on the on-premises side of the VPN and a Customer Gateway on the AWS side of the VPN - You need to create a virtual private gateway (VGW) on

the AWS side of the VPN and a Customer Gateway on the on-premises side of the VPN. Therefore, this option is wrong.

Create a Customer Gateway on both the AWS side of the VPN as well as the on-premises side of the VPN - You need to create a virtual private gateway (VGW) on the AWS side of the VPN and a Customer Gateway on the on-premises side of the VPN. Therefore, this option is wrong.

Create a virtual private gateway (VGW) on both the AWS side of the VPN as well as the on-premises side of the VPN - You need to create a virtual private gateway (VGW) on the AWS side of the VPN and a Customer Gateway on the on-premises side of the VPN. Therefore, this option is wrong.

References:

https://docs.aws.amazon.com/whitepapers/latest/aws-vpc-connectivity-options/aws-managed-vpn-network-to-amazon.html

https://docs.aws.amazon.com/vpn/latest/s2svpn/VPC_VPN.html

Domain

Design Secure Architectures

Question 2 Correct

A pharma company is working on developing a vaccine for the COVID-19 virus. The researchers at the company want to process the reference healthcare data in a highly available as well as HIPAA compliant in-memory database that supports caching results of SQL queries.

As a solutions architect, which of the following AWS services would you recommend for this task?

Amazon DynamoDB Accelerator (DAX)

Your answer is correct

Amazon ElastiCache for Redis/Memcached

Amazon DocumentDB

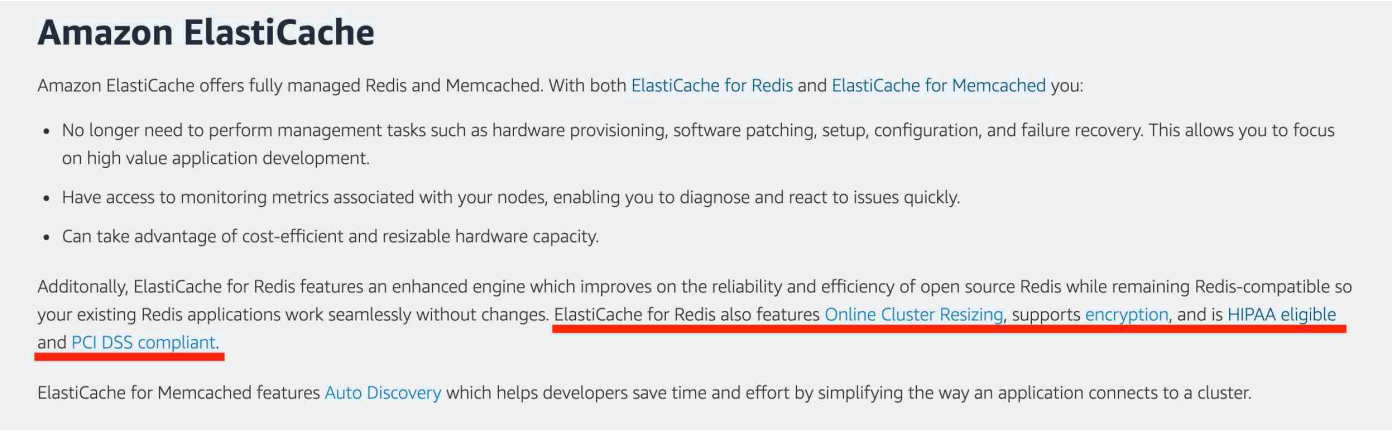
Amazon DynamoDB

Overall explanation

Correct option:

Amazon ElastiCache for Redis/Memcached

Amazon ElastiCache Overview:



Amazon ElastiCache

Amazon ElastiCache offers fully managed Redis and Memcached. With both ElastiCache for Redis and ElastiCache for Memcached you:

- No longer need to perform management tasks such as hardware provisioning, software patching, setup, configuration, and failure recovery. This allows you to focus on high value application development.
- Have access to monitoring metrics associated with your nodes, enabling you to diagnose and react to issues quickly.
- Can take advantage of cost-efficient and resizable hardware capacity.

Additonally, ElastiCache for Redis features an enhanced engine which improves on the reliability and efficiency of open source Redis while remaining Redis-compatible so your existing Redis applications work seamlessly without changes. ElastiCache for Redis also features Online Cluster Resizing, supports encryption, and is HIPAA eligible and PCI DSS compliant.

ElastiCache for Memcached features Auto Discovery which helps developers save time and effort by simplifying the way an application connects to a cluster.

via - https://aws.amazon.com/elasticache/redis-vs-memcached/

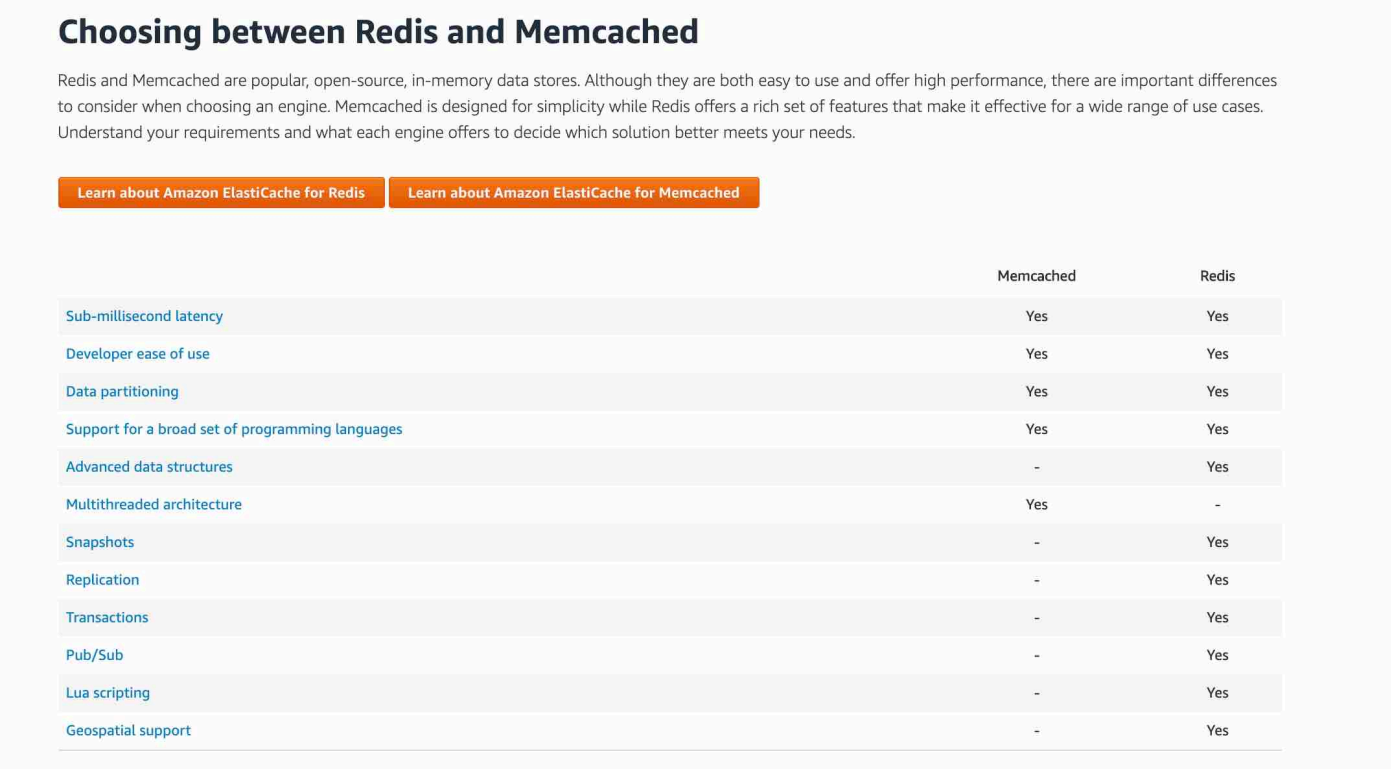
Amazon ElastiCache for Redis is a blazing fast in-memory data store that provides sub-millisecond latency to power internet-scale real-time applications. Amazon ElastiCache for Redis is a great choice for real-time transactional and analytical processing use cases such as caching, chat/messaging, gaming leaderboards, geospatial, machine learning, media streaming, queues, real-time analytics, and session store. ElastiCache for Redis supports replication, high availability, and cluster sharding right out of the box.

Amazon ElastiCache for Memcached is a Memcached-compatible in-memory key-value store service that can be used as a cache or a data store. Amazon ElastiCache for Memcached is a great choice for implementing an in-memory cache to decrease access latency, increase throughput, and ease the load off your relational or NoSQL database. Session stores are easy to create with Amazon ElastiCache for Memcached.

Both Amazon ElastiCache for Redis and Amazon ElastiCache for Memcached are HIPAA Eligible. Therefore, this is the correct option.

Exam Alert:

Please review this comparison sheet for Redis vs Memcached features:



Choosing between Redis and Memcached

Redis and Memcached are popular, open-source, in-memory data stores. Although they are both easy to use and offer high performance, there are important differences to consider when choosing an engine. Memcached is designed for simplicity while Redis offers a rich set of features that make it effective for a wide range of use cases. Understand your requirements and what each engine offers to decide which solution better meets your needs.

Learn about Amazon ElastiCache for Redis | Learn about Amazon ElastiCache for Memcached

	Memcached	Redis
Sub-millisecond latency	Yes	Yes
Developer ease of use	Yes	Yes
Data partitioning	Yes	Yes
Support for a broad set of programming languages	Yes	Yes
Advanced data structures	-	Yes
Multithreaded architecture	Yes	-
Snapshots	-	Yes
Replication	-	Yes
Transactions	-	Yes
Pub/Sub	-	Yes
Lua scripting	-	Yes
Geospatial support	-	Yes

via - https://aws.amazon.com/elasticache/redis-vs-memcached/

Incorrect Options:

Amazon DynamoDB Accelerator (DAX) - Amazon DynamoDB is a key-value and document database that delivers single-digit millisecond performance at any scale. It's a fully managed, multi-region, multi-master, durable database with built-in security, backup and restore, and in-memory caching for internet-scale applications. DAX is a DynamoDB-compatible caching service that enables you to benefit from fast in-memory performance for demanding applications. DAX does not support SQL query caching.

Amazon DynamoDB - Amazon DynamoDB is a key-value and document database that delivers single-digit millisecond performance at any scale. It's a fully managed, multi-region, multi-master, durable database with built-in security, backup and restore, and in-memory caching (via DAX) for internet-scale applications. Amazon DynamoDB is not an in-memory database, so this option is incorrect.

Amazon DocumentDB - Amazon DocumentDB is a fast, scalable, highly available, and fully managed document database service that supports MongoDB workloads. As a document

database, Amazon DocumentDB makes it easy to store, query, and index JSON data. Amazon DocumentDB is not an in-memory database, so this option is incorrect.

References:

https://aws.amazon.com/about-aws/whats-new/2017/11/amazon-elasticache-for-redis-is-now-hipaa-eligible-to-help-you-power-secure-healthcare-applications-with-sub-millisecond-latency/

https://aws.amazon.com/elasticache/redis/

https://aws.amazon.com/about-aws/whats-new/2022/08/amazon-elasticache-memcached-hipaa-eligible/

https://aws.amazon.com/blogs/database/automating-sql-caching-for-amazon-elasticache-and-amazon-rds/

Domain

Design Secure Architectures

Question 3 **Correct**

A Big Data processing company has created a distributed data processing framework that performs best if the network performance between the processing machines is high. The application has to be deployed on AWS, and the company is only looking at performance as the key measure.

As a Solutions Architect, which deployment do you recommend?

Optimize the Amazon EC2 kernel using EC2 User Data

Your answer is correct

Use a Cluster placement group

Use a Spread placement group

Use Spot Instances

Overall explanation

Correct option:

When you launch a new Amazon EC2 instance, the EC2 service attempts to place the instance in such a way that all of your instances are spread out across underlying hardware to minimize correlated failures. You can use placement groups to influence the placement of a group of interdependent instances to meet the needs of your workload. Depending on the type of workload, you can create a placement group using one of the following placement strategies:

Cluster – packs instances close together inside an Availability Zone (AZ). This strategy enables workloads to achieve the low-latency network performance necessary for tightly-coupled node-to-node communication that is typical of HPC applications.

Partition – spreads your instances across logical partitions such that groups of instances in one partition do not share the underlying hardware with groups of instances in different partitions. This strategy is typically used by large distributed and replicated workloads, such as Hadoop, Cassandra, and Kafka.

Spread – strictly places a small group of instances across distinct underlying hardware to reduce correlated failures.

There is no charge for creating a placement group.

Use a Cluster placement group

A cluster placement group is a logical grouping of instances within a single Availability Zone (AZ). A cluster placement group can span peered VPCs in the same Region. Instances in the same cluster placement group enjoy a higher per-flow throughput limit of up to 10 Gbps for TCP/IP traffic and are placed in the same high-bisection bandwidth segment of the network.

Cluster placement groups are recommended for applications that benefit from low network latency, high network throughput, or both. They are also recommended when the majority of the network traffic is between the instances in the group. To provide the lowest latency and the highest packet-per-second network performance for your placement group, choose an instance type that supports enhanced networking.

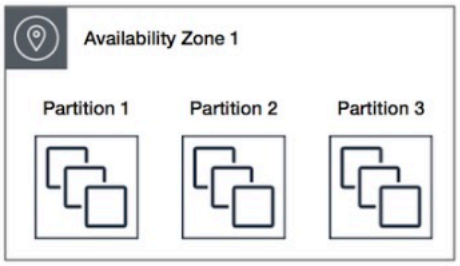
Image of Cluster placement group:

The following image shows instances that are placed into a cluster placement group.

via - https://docs.aws.amazon.com/AWSEC2/latest/UserGuide/placement-groups.html

Image of Partition placement group:

The following image is a simple visual representation of a partition placement group in a single Availability Zone. It shows instances that are placed into a partition placement group with three partitions—**Partition 1**, **Partition 2**, and **Partition 3**. Each partition comprises multiple instances. The instances in a partition do not share racks with the instances in the other partitions, allowing you to contain the impact of a single hardware failure to only the associated partition.



via - https://docs.aws.amazon.com/AWSEC2/latest/UserGuide/placement-groups.html

Image of Spread placement group:

The following image shows seven instances in a single Availability Zone that are placed into a spread placement group. The seven instances are placed on seven different racks.

via - https://docs.aws.amazon.com/AWSEC2/latest/UserGuide/placement-groups.html

Incorrect options:

Use Spot Instances - A Spot Instance is an unused Amazon EC2 instance that is available for less than the On-Demand price. Because Spot Instances enable you to request unused Amazon EC2 instances at steep discounts, you can lower your Amazon EC2 costs significantly. Spot Instances

are a cost-effective choice if you can be flexible about when your applications run and if your applications can be interrupted. Since performance is the key criteria, this is not the right choice.

Optimize the Amazon EC2 kernel using EC2 User Data - Optimizing the Amazon EC2 kernel won't help with network performance as it's bounded by the EC2 instance type mainly. Therefore, this option is incorrect.

Use a Spread placement group - A spread placement group is a group of instances that are each placed on distinct racks, with each rack having its own network and power source. The instances are placed across distinct underlying hardware to reduce correlated failures. A spread placement group can span multiple Availability Zones (AZs) in the same Region. You can have a maximum of seven running instances per Availability Zone (AZ) per group.

Reference:

https://docs.aws.amazon.com/AWSEC2/latest/UserGuide/placement-groups.html

Domain

Design High-Performing Architectures

Question 4 **Correct**

A leading social media analytics company is contemplating moving its dockerized application stack into AWS Cloud. The company is not sure about the pricing for using Amazon Elastic Container Service (Amazon ECS) with the EC2 launch type compared to the Amazon Elastic Container Service (Amazon ECS) with the Fargate launch type.

Which of the following is correct regarding the pricing for these two services?

Both Amazon ECS with EC2 launch type and Amazon ECS with Fargate launch type are charged based on vCPU and memory resources that the containerized application requests

Both Amazon ECS with EC2 launch type and Amazon ECS with Fargate launch type are charged based on Amazon EC2 instances and Amazon EBS Elastic Volumes used

Your answer is correct

Amazon ECS with EC2 launch type is charged based on EC2 instances and EBS volumes used. Amazon ECS with Fargate launch type is charged based on vCPU and memory resources that the containerized application requests

Both Amazon ECS with EC2 launch type and Amazon ECS with Fargate launch type are just charged based on Elastic Container Service used per hour

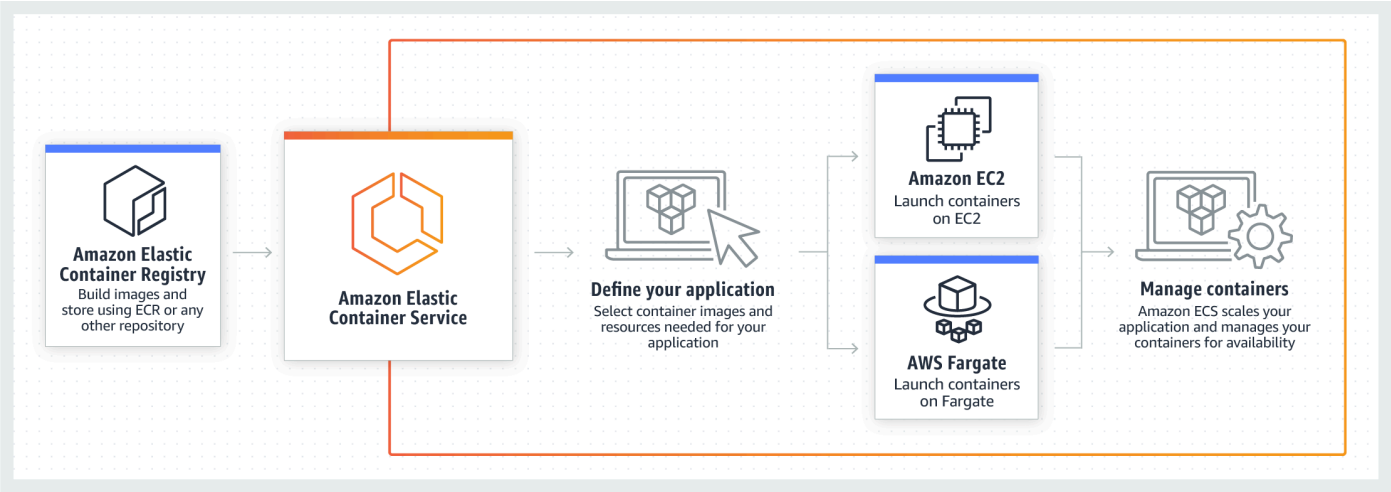
Overall explanation

Correct option:

Amazon ECS with EC2 launch type is charged based on EC2 instances and EBS volumes used. Amazon ECS with Fargate launch type is charged based on vCPU and memory resources that the containerized application requests

Amazon Elastic Container Service (Amazon ECS) is a fully managed container orchestration service. ECS allows you to easily run, scale, and secure Docker container applications on AWS.

Amazon ECS Overview:



via - https://aws.amazon.com/ecs/

With the Fargate launch type, you pay for the amount of vCPU and memory resources that your containerized application requests. vCPU and memory resources are calculated from the time your container images are pulled until the Amazon ECS Task terminates, rounded up to the nearest second. With the EC2 launch type, there is no additional charge for the EC2 launch type.

You pay for AWS resources (e.g. EC2 instances or EBS volumes) you create to store and run your application.

Incorrect options:

Both Amazon ECS with EC2 launch type and Amazon ECS with Fargate launch type are charged based on vCPU and memory resources that the containerized application requests

Both Amazon ECS with EC2 launch type and Amazon ECS with Fargate launch type are charged based on Amazon EC2 instances and Amazon EBS Elastic Volumes used

As mentioned above - with the Fargate launch type, you pay for the amount of vCPU and memory resources. With EC2 launch type, you pay for AWS resources (e.g. EC2 instances or EBS volumes). Hence both these options are incorrect.

Both Amazon ECS with EC2 launch type and Amazon ECS with Fargate launch type are just charged based on Elastic Container Service used per hour

This is a made-up option and has been added as a distractor.

References:

https://aws.amazon.com/ecs/pricing/

Domain

Design Cost-Optimized Architectures

Question 5 **Correct**

An Internet of Things (IoT) company would like to have a streaming system that performs real-time analytics on the ingested IoT data. Once the analytics is done, the company would like to send notifications back to the mobile applications of the IoT device owners.

As a solutions architect, which of the following AWS technologies would you recommend to send these notifications to the mobile applications?

Amazon Kinesis with Amazon Simple Email Service (Amazon SES)

Your answer is correct

Amazon Kinesis with Amazon Simple Notification Service (Amazon SNS)

Amazon Kinesis with Amazon Simple Queue Service (Amazon SQS)

Amazon Simple Queue Service (Amazon SQS) with Amazon Simple Notification Service (Amazon SNS)

Overall explanation

Correct option:

Amazon Kinesis with Amazon Simple Notification Service (Amazon SNS)

Amazon Kinesis makes it easy to collect, process, and analyze real-time, streaming data so you can get timely insights and react quickly to new information. Amazon Kinesis offers key capabilities to cost-effectively process streaming data at any scale, along with the flexibility to choose the tools that best suit the requirements of your application.

With Amazon Kinesis, you can ingest real-time data such as video, audio, application logs, website clickstreams, and IoT telemetry data for machine learning, analytics, and other applications. Amazon Kinesis enables you to process and analyze data as it arrives and respond instantly instead of having to wait until all your data is collected before the processing can begin.

Amazon Kinesis will be great for event streaming from the IoT devices, but not for sending notifications as it doesn't have such a feature.

Amazon Simple Notification Service (Amazon SNS) is a highly available, durable, secure, fully managed pub/sub messaging service that enables you to decouple microservices, distributed systems, and serverless applications. Amazon SNS provides topics for high-throughput, push-based, many-to-many messaging. Amazon SNS is a notification service and will be perfect for this use case.

Streaming data with Amazon Kinesis and using Amazon SNS to send the response notifications is the optimal solution for the current scenario.

Incorrect options:

Amazon Simple Queue Service (Amazon SQS) with Amazon Simple Notification Service (Amazon SNS) - Amazon Simple Queue Service (SQS) is a fully managed message queuing service that enables you to decouple and scale microservices, distributed systems, and serverless applications. Amazon SQS eliminates the complexity and overhead associated with managing and operating message-oriented middleware and empowers developers to focus on differentiating work. Using SQS, you can send, store, and receive messages between software components at any volume, without losing messages or requiring other services to be available. Kinesis is better for streaming data since queues aren't meant for real-time streaming of data.

Amazon Kinesis with Amazon Simple Email Service (Amazon SES) - Amazon Simple Email Service (Amazon SES) is a cloud-based email sending service designed to help digital marketers and application developers send marketing, notification, and transactional emails. It is a reliable, cost-effective service for businesses of all sizes that use email to keep in contact with their customers. It is an email service and not a notification service as is the requirement in the current use case.

Amazon Kinesis with Amazon Simple Queue Service (Amazon SQS) - As explained above, Amazon Kinesis works well for streaming real-time data. Amazon SQS is a queuing service that helps decouple system architecture by offering flexibility and ease of maintenance. It cannot send notifications. Amazon SQS is paired with SNS to provide this functionality.

References:

https://aws.amazon.com/sns/

https://aws.amazon.com/kinesis/

https://aws.amazon.com/ses/

https://aws.amazon.com/sqs/

Domain

Design Resilient Architectures

Question 6 **Correct**

The engineering team at a company wants to use Amazon Simple Queue Service (Amazon SQS) to decouple components of the underlying application architecture. However, the team is concerned about the VPC-bound components accessing Amazon Simple Queue Service (Amazon SQS) over the public internet.

As a solutions architect, which of the following solutions would you recommend to address this use-case?

Use Network Address Translation (NAT) instance to access Amazon SQS

Use Internet Gateway to access Amazon SQS

Use VPN connection to access Amazon SQS

Your answer is correct

Use VPC endpoint to access Amazon SQS

Overall explanation

Correct option:

Use VPC endpoint to access Amazon SQS

AWS customers can access Amazon Simple Queue Service (Amazon SQS) from their Amazon Virtual Private Cloud (Amazon VPC) using VPC endpoints, without using public IPs, and without needing to traverse the public internet. VPC endpoints for Amazon SQS are powered by AWS PrivateLink, a highly available, scalable technology that enables you to privately connect your VPC to supported AWS services.

Amazon VPC endpoints are easy to configure. They also provide reliable connectivity to Amazon SQS without requiring an internet gateway, Network Address Translation (NAT) instance, VPN connection, or AWS Direct Connect connection. With VPC endpoints, the data between your Amazon VPC and Amazon SQS queue is transferred within the Amazon network, helping protect your instances from internet traffic.

AWS PrivateLink simplifies the security of data shared with cloud-based applications by eliminating the exposure of data to the public Internet. AWS PrivateLink provides private connectivity between VPCs, AWS services, and on-premises applications, securely on the Amazon network. AWS PrivateLink makes it easy to connect services across different accounts and VPCs to significantly simplify the network architecture.

Incorrect options:

Use Internet Gateway to access Amazon SQS - An internet gateway is a horizontally scaled, redundant, and highly available VPC component that allows communication between instances in your VPC and the internet. It, therefore, imposes no availability risks or bandwidth constraints on your network traffic. This option is ruled out as the team does not want to use the public internet to access Amazon SQS.

Use VPN connection to access Amazon SQS - AWS Site-to-Site VPN (aka VPN Connection) enables you to securely connect your on-premises network or branch office site to your Amazon Virtual Private Cloud (Amazon VPC). You can securely extend your data center or branch office network to the cloud with an AWS Site-to-Site VPN connection. A VPC VPN Connection utilizes IPSec to establish encrypted network connectivity between your intranet and Amazon VPC over the Internet. VPN Connections can be configured in minutes and are a good solution if you have an immediate need, have low to modest bandwidth requirements, and can tolerate the inherent variability in Internet-based connectivity. As the existing infrastructure is within AWS Cloud, therefore a VPN connection is not required.

Use Network Address Translation (NAT) instance to access Amazon SQS - You can use a network address translation (NAT) instance in a public subnet in your VPC to enable instances in the private subnet to initiate outbound IPv4 traffic to the Internet or other AWS services, but prevent the instances from receiving inbound traffic initiated by someone on the Internet. Amazon provides Amazon Linux AMIs that are configured to run as NAT instances. These AMIs include the string amzn-ami-vpc-nat in their names, so you can search for them in the Amazon EC2 console. This option is ruled out because NAT instances are used to provide internet access to any instances in a private subnet.

References:

https://aws.amazon.com/privatelink/

https://aws.amazon.com/about-aws/whats-new/2018/12/amazon-sqs-vpc-endpoints-aws-privatelink/

Domain

Design Secure Architectures

Question 7 Incorrect

A media company has its corporate headquarters in Los Angeles with an on-premises data center using an AWS Direct Connect connection to the AWS VPC. The branch offices in San Francisco and Miami use AWS Site-to-Site VPN connections to connect to the AWS VPC. The company is looking for a solution to have the branch offices send and receive data with each other as well as with their corporate headquarters.

As a solutions architect, which of the following AWS services would you recommend addressing this use-case?

Your answer is incorrect

VPC Endpoint

Correct answer

AWS VPN CloudHub

VPC Peering connection

Software VPN

Overall explanation

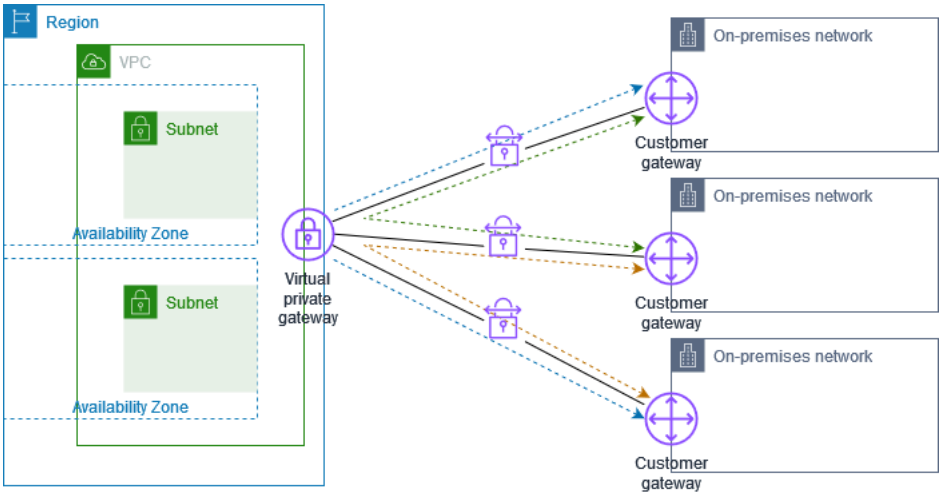
Correct option:

AWS VPN CloudHub

If you have multiple AWS Site-to-Site VPN connections, you can provide secure communication between sites using the AWS VPN CloudHub. This enables your remote sites to communicate with each other, and not just with the VPC. Sites that use AWS Direct Connect connections to the virtual private gateway can also be part of the AWS VPN CloudHub. The VPN CloudHub operates on a simple hub-and-spoke model that you can use with or without a VPC. This design is suitable if you have multiple branch offices and existing internet connections and would like to implement a convenient, potentially low-cost hub-and-spoke model for primary or backup connectivity between these remote offices.

Per the given use-case, the corporate headquarters has an AWS Direct Connect connection to the VPC and the branch offices have Site-to-Site VPN connections to the VPC. Therefore using the AWS VPN CloudHub, branch offices can send and receive data with each other as well as with their corporate headquarters.

AWS VPN CloudHub:



via - https://docs.aws.amazon.com/vpn/latest/s2svpn/VPN_CloudHub.html

Incorrect options:

VPC Endpoint - A VPC endpoint enables you to privately connect your VPC to supported AWS services and VPC endpoint services powered by AWS PrivateLink without requiring an internet gateway, NAT device, VPN connection, or AWS Direct Connect connection. Instances in your VPC

do not require public IP addresses to communicate with resources in the service. AWS PrivateLink simplifies the security of data shared with cloud-based applications by eliminating the exposure of data to the public Internet. When you use VPC endpoint, the traffic between your VPC and the other AWS service does not leave the Amazon network, therefore this option cannot be used to send and receive data between the remote branch offices of the company.

VPC Peering connection - A VPC peering connection is a networking connection between two VPCs that enables you to route traffic between them using private IPv4 addresses or IPv6 addresses. Instances in either VPC can communicate with each other as if they are within the same network. VPC peering facilitates a connection between two VPCs within the AWS network, therefore this option cannot be used to send and receive data between the remote branch offices of the company.

Software VPN - Amazon VPC offers you the flexibility to fully manage both sides of your Amazon VPC connectivity by creating a VPN connection between your remote network and a software VPN appliance running in your Amazon VPC network. Since Software VPN just handles connectivity between the remote network and Amazon VPC, therefore it cannot be used to send and receive data between the remote branch offices of the company.

References:

https://docs.aws.amazon.com/whitepapers/latest/aws-vpc-connectivity-options/aws-vpn-cloudhub-network-to-amazon.html

https://docs.aws.amazon.com/vpn/latest/s2svpn/VPN_CloudHub.html

Domain

Design Secure Architectures

Question 8 **Correct**

An engineering team wants to examine the feasibility of the `user data` feature of Amazon EC2 for an upcoming project.

Which of the following are true about the Amazon EC2 user data configuration? (Select two)

By default, user data is executed every time an Amazon EC2 instance is re-started

By default, scripts entered as user data do not have root user privileges for executing

Your selection is correct

By default, user data runs only during the boot cycle when you first launch an instance

When an instance is running, you can update user data by using root user credentials

Your selection is correct

By default, scripts entered as user data are executed with root user privileges

Overall explanation

Correct options:

User Data is generally used to perform common automated configuration tasks and even run scripts after the instance starts. When you launch an instance in Amazon EC2, you can pass two types of user data - shell scripts and cloud-init directives. You can also pass this data into the launch wizard as plain text or as a file.

By default, scripts entered as user data are executed with root user privileges

Scripts entered as user data are executed as the root user, hence do not need the sudo command in the script. Any files you create will be owned by root; if you need non-root users to have file access, you should modify the permissions accordingly in the script.

By default, user data runs only during the boot cycle when you first launch an instance

By default, user data scripts and cloud-init directives run only during the boot cycle when you first launch an instance. You can update your configuration to ensure that your user data scripts and cloud-init directives run every time you restart your instance.

Incorrect options:

By default, user data is executed every time an Amazon EC2 instance is re-started - As discussed above, this is not a default configuration of the system. But, can be achieved by explicitly configuring the instance.

When an instance is running, you can update user data by using root user credentials - You can't change the user data if the instance is running (even by using root user credentials), but you can view it.

By default, scripts entered as user data do not have root user privileges for executing - Scripts entered as user data are executed as the root user, hence do not need the sudo command in the script.

Reference:

https://docs.aws.amazon.com/AWSEC2/latest/UserGuide/user-data.html

Domain

Design High-Performing Architectures

Question 9 **Correct**

A developer needs to implement an AWS Lambda function in AWS account A that accesses an Amazon Simple Storage Service (Amazon S3) bucket in AWS account B.

As a Solutions Architect, which of the following will you recommend to meet this requirement?

AWS Lambda cannot access resources across AWS accounts. Use Identity federation to work around this limitation of Lambda

Create an IAM role for the AWS Lambda function that grants access to the Amazon S3 bucket. Set the IAM role as the Lambda function's execution role and that would give the

AWS Lambda function cross-account access to the Amazon S3 bucket

The Amazon S3 bucket owner should make the bucket public so that it can be accessed by the AWS Lambda function in the other AWS account

Your answer is correct

Create an IAM role for the AWS Lambda function that grants access to the Amazon S3 bucket. Set the IAM role as the AWS Lambda function's execution role. Make sure that the bucket policy also grants access to the AWS Lambda function's execution role

Overall explanation

Correct option:

Create an IAM role for the AWS Lambda function that grants access to the Amazon S3 bucket. Set the IAM role as the AWS Lambda function's execution role. Make sure that the bucket policy also grants access to the AWS Lambda function's execution role

If the IAM role that you create for the Lambda function is in the same AWS account as the bucket, then you don't need to grant Amazon S3 permissions on both the IAM role and the bucket policy. Instead, you can grant the permissions on the IAM role and then verify that the bucket policy doesn't explicitly deny access to the Lambda function role. If the IAM role and the bucket are in different accounts, then you need to grant Amazon S3 permissions on both the IAM role and the bucket policy. Therefore, this is the right way of giving access to AWS Lambda for the given use-case.

Complete list of steps to be followed:

I want my AWS Lambda function to be able to access my Amazon Simple Storage Service (Amazon S3) bucket. How can I do that?

Short Description

Follow these steps:

1. Create an AWS Identity and Access Management (IAM) role for the Lambda function that also grants access to the S3 bucket.
2. Set the IAM role as the Lambda function's execution role.
3. Verify that the bucket policy grants access to the Lambda function's execution role.

Important: If the IAM role that you create for the Lambda function is in the same AWS account as the bucket, then you don't need to grant Amazon S3 permissions on both the IAM role and the bucket policy. Instead, you can grant the permissions on the IAM role and then verify that the bucket policy doesn't explicitly deny access to the Lambda function role. As an example, the following procedure grants Amazon S3 permissions on the IAM role. If the IAM role and the bucket are in different accounts, then you need to grant Amazon S3 permissions on both the IAM role and the bucket policy.

via - https://aws.amazon.com/premiumsupport/knowledge-center/lambda-execution-role-s3-bucket/

Incorrect options:

AWS Lambda cannot access resources across AWS accounts. Use Identity federation to work around this limitation of Lambda - This is an incorrect statement, used only as a distractor.

Create an IAM role for the AWS Lambda function that grants access to the Amazon S3 bucket. Set the IAM role as the Lambda function's execution role and that would give the AWS Lambda function cross-account access to the Amazon S3 bucket - When the execution role of AWS Lambda and Amazon S3 bucket to be accessed are from different accounts, then you need to grant Amazon S3 bucket access permissions to the IAM role and also ensure that the bucket policy grants access to the AWS Lambda function's execution role.

The Amazon S3 bucket owner should make the bucket public so that it can be accessed by the AWS Lambda function in the other AWS account - Making the Amazon S3 bucket public for the given use-case will be considered as a security bad practice. It's usually done for very few use-cases such as hosting a website on Amazon S3. Therefore this option is incorrect.

Reference:

https://aws.amazon.com/premiumsupport/knowledge-center/lambda-execution-role-s3-bucket/

Domain

Design Secure Architectures

Question 10 Incorrect

The DevOps team at an IT company is provisioning a two-tier application in a VPC with a public subnet and a private subnet. The team wants to use either a Network Address Translation (NAT) instance or a Network Address Translation (NAT) gateway in the public subnet to enable instances in the private subnet to initiate outbound IPv4 traffic to the internet but needs some technical assistance in terms of the configuration options available for the Network Address Translation (NAT) instance and the Network Address Translation (NAT) gateway.

As a solutions architect, which of the following options would you identify as CORRECT? (Select three)

Your selection is correct

Security Groups can be associated with a NAT instance

Correct selection

NAT instance supports port forwarding

Your selection is incorrect

NAT gateway supports port forwarding

Your selection is correct

NAT instance can be used as a bastion server

Security Groups can be associated with a NAT gateway

NAT gateway can be used as a bastion server

Overall explanation

Correct options:

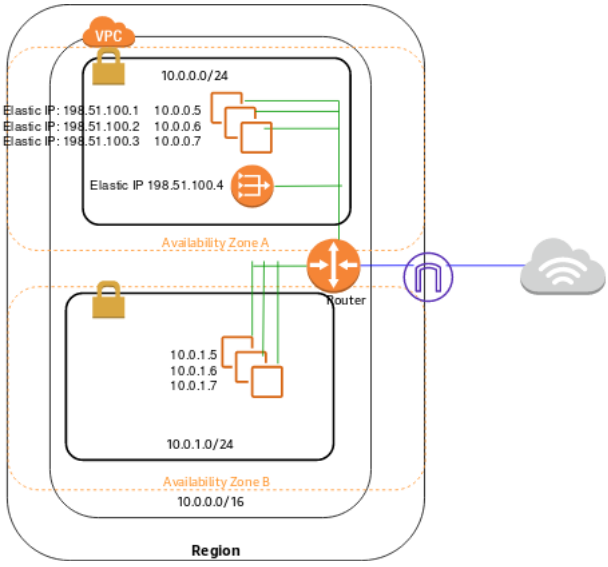
NAT instance can be used as a bastion server

Security Groups can be associated with a NAT instance

NAT instance supports port forwarding

A NAT instance or a NAT Gateway can be used in a public subnet in your VPC to enable instances in the private subnet to initiate outbound IPv4 traffic to the Internet.

How NAT Gateway works:



via - https://docs.aws.amazon.com/vpc/latest/userguide/vpc-nat-gateway.html

How NAT Instance works:

via - https://docs.aws.amazon.com/vpc/latest/userguide/VPC_NAT_Instance.html

Please see this high-level summary of the differences between NAT instances and NAT gateways relevant to the options described in the question:

The following is a high-level summary of the differences between NAT instances and NAT gateways.

Attribute	NAT gateway	NAT instance
Security groups	Cannot be associated with a NAT gateway. You can associate security groups with your resources behind the NAT gateway to control inbound and outbound traffic.	Associate with your NAT instance and the resources behind your NAT instance to control inbound and outbound traffic.
Network ACLs	Use a network ACL to control the traffic to and from the subnet in which your NAT gateway resides.	Use a network ACL to control the traffic to and from the subnet in which your NAT instance resides.
Flow logs	Use flow logs to capture the traffic.	Use flow logs to capture the traffic.
Port forwarding	Not supported.	Manually customize the configuration to support port forwarding.
Bastion servers	Not supported.	Use as a bastion server.
Traffic metrics	View CloudWatch metrics for the NAT gateway.	View CloudWatch metrics for the instance.
Timeout behavior	When a connection times out, a NAT gateway returns an RST packet to any resources behind the	When a connection times out, a NAT instance sends a FIN packet to resources behind the NAT

via - https://docs.aws.amazon.com/vpc/latest/userguide/vpc-nat-comparison.html

Incorrect options:

NAT gateway supports port forwarding

Security Groups can be associated with a NAT gateway

NAT gateway can be used as a bastion server

These three options contradict the details provided in the explanation above, so these options are incorrect.

Reference:

https://docs.aws.amazon.com/vpc/latest/userguide/vpc-nat-comparison.html

Domain

Design High-Performing Architectures

Question 11 Correct

A healthcare startup needs to enforce compliance and regulatory guidelines for objects stored in Amazon S3. One of the key requirements is to provide adequate protection against accidental deletion of objects.

As a solutions architect, what are your recommendations to address these guidelines? (Select two) ?

Change the configuration on Amazon S3 console so that the user needs to provide additional confirmation while deleting any Amazon S3 object

Your selection is correct

Enable versioning on the Amazon S3 bucket

Establish a process to get managerial approval for deleting Amazon S3 objects

Your selection is correct

Enable multi-factor authentication (MFA) delete on the Amazon S3 bucket

Create an event trigger on deleting any Amazon S3 object. The event invokes an Amazon Simple Notification Service (Amazon SNS) notification via email to the IT manager

Overall explanation

Correct options:

Enable versioning on the Amazon S3 bucket

Versioning is a means of keeping multiple variants of an object in the same bucket. You can use versioning to preserve, retrieve, and restore every version of every object stored in your Amazon S3 bucket. Versioning-enabled buckets enable you to recover objects from accidental deletion or overwrite.

For example:

If you overwrite an object, it results in a new object version in the bucket. You can always restore the previous version. If you delete an object, instead of removing it permanently, Amazon S3 inserts a delete marker, which becomes the current object version. You can always restore the previous version. Hence, this is the correct option.

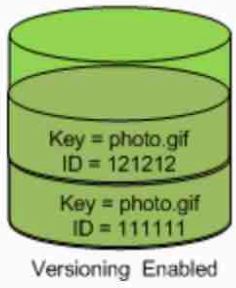
Versioning Overview:

Using versioning

PDF | Kindle | RSS

Versioning is a means of keeping multiple variants of an object in the same bucket. You can use versioning to preserve, retrieve, and restore every version of every object stored in your Amazon S3 bucket. With versioning, you can easily recover from both unintended user actions and application failures. When you enable versioning for a bucket, if Amazon S3 receives multiple write requests for the same object simultaneously, it stores all of the objects.

If you enable versioning for a bucket, Amazon S3 automatically generates a unique version ID for the object being stored. In one bucket, for example, you can have two objects with the same key, but different version IDs, such as `photo.gif` (version 111111) and `photo.gif` (version 121212).



via - https://docs.aws.amazon.com/AmazonS3/latest/dev/Versioning.html

Enable multi-factor authentication (MFA) delete on the Amazon S3 bucket

To provide additional protection, multi-factor authentication (MFA) delete can be enabled. MFA delete requires secondary authentication to take place before objects can be permanently deleted from an Amazon S3 bucket. Hence, this is the correct option.

Incorrect options:

Create an event trigger on deleting any Amazon S3 object. The event invokes an Amazon Simple Notification Service (Amazon SNS) notification via email to the IT manager - Sending

an event trigger after object deletion does not meet the objective of preventing object deletion by mistake because the object has already been deleted. So, this option is incorrect.

Establish a process to get managerial approval for deleting Amazon S3 objects - This option for getting managerial approval is just a distractor.

Change the configuration on Amazon S3 console so that the user needs to provide additional confirmation while deleting any Amazon S3 object - There is no provision to set up Amazon S3 configuration to ask for additional confirmation before deleting an object. This option is incorrect.

References:

https://docs.aws.amazon.com/AmazonS3/latest/dev/Versioning.html

https://docs.aws.amazon.com/AmazonS3/latest/dev/UsingMFADelete.html

Domain

Design Resilient Architectures

Question 12 Correct

A medium-sized business has a taxi dispatch application deployed on an Amazon EC2 instance. Because of an unknown bug, the application causes the instance to freeze regularly. Then, the instance has to be manually restarted via the AWS management console.

Which of the following is the MOST cost-optimal and resource-efficient way to implement an automated solution until a permanent fix is delivered by the development team?

Use Amazon EventBridge events to trigger an AWS Lambda function to check the instance status every 5 minutes. In the case of Instance Health Check failure, the AWS lambda function can use Amazon EC2 API to reboot the instance

Setup an Amazon CloudWatch alarm to monitor the health status of the instance. In case of an Instance Health Check failure, Amazon CloudWatch Alarm can publish to an Amazon Simple Notification Service (Amazon SNS) event which can then trigger an AWS lambda function. The AWS lambda function can use Amazon EC2 API to reboot the instance

Your answer is correct

Setup an Amazon CloudWatch alarm to monitor the health status of the instance. In case of an Instance Health Check failure, an EC2 Reboot CloudWatch Alarm Action can be used to reboot the instance

Use Amazon EventBridge events to trigger an AWS Lambda function to reboot the instance status every 5 minutes

Overall explanation

Correct option:

Setup an Amazon CloudWatch alarm to monitor the health status of the instance. In case of an Instance Health Check failure, an EC2 Reboot CloudWatch Alarm Action can be used to reboot the instance

Using Amazon CloudWatch alarm actions, you can create alarms that automatically stop, terminate, reboot, or recover your Amazon EC2 instances. You can use the stop or terminate actions to help you save money when you no longer need an instance to be running. You can use the reboot and recover actions to automatically reboot those instances or recover them onto new hardware if a system impairment occurs.

You can create an Amazon CloudWatch alarm that monitors an Amazon EC2 instance and automatically reboots the instance. The reboot alarm action is recommended for Instance Health Check failures (as opposed to the recover alarm action, which is suited for System Health Check failures).

Incorrect options:

Setup an Amazon CloudWatch alarm to monitor the health status of the instance. In case of an Instance Health Check failure, Amazon CloudWatch Alarm can publish to an Amazon Simple Notification Service (Amazon SNS) event which can then trigger an AWS lambda function. The AWS lambda function can use Amazon EC2 API to reboot the instance

Use Amazon EventBridge events to trigger an AWS Lambda function to check the instance status every 5 minutes. In the case of Instance Health Check failure, the AWS lambda function can use Amazon EC2 API to reboot the instance

Use Amazon EventBridge events to trigger an AWS Lambda function to reboot the instance status every 5 minutes

Using Amazon EventBridge event or Amazon CloudWatch alarm to trigger an AWS lambda function, directly or indirectly, is wasteful of resources. You should just use the EC2 Reboot CloudWatch Alarm Action to reboot the instance. So all the options that trigger the AWS lambda function are incorrect.

Reference:

https://docs.aws.amazon.com/AmazonCloudWatch/latest/monitoring/UsingAlarmActions.html

Domain

Design Cost-Optimized Architectures

Question 13 Incorrect

A retail company wants to rollout and test a blue-green deployment for its global application in the next 48 hours. Most of the customers use mobile phones which are prone to Domain Name System (DNS) caching. The company has only two days left for the annual Thanksgiving sale to commence.

As a Solutions Architect, which of the following options would you recommend to test the deployment on as many users as possible in the given time frame?

Use AWS CodeDeploy deployment options to choose the right deployment

Your answer is incorrect

Use Elastic Load Balancing (ELB) to distribute traffic across deployments

Correct answer

Use AWS Global Accelerator to distribute a portion of traffic to a particular deployment

Use Amazon Route 53 weighted routing to spread traffic across different deployments

Overall explanation

Correct option:

Blue/green deployment is a technique for releasing applications by shifting traffic between two identical environments running different versions of the application: "Blue" is the currently running version and "green" the new version. This type of deployment allows you to test features in the green environment without impacting the currently running version of your application. When you're satisfied that the green version is working properly, you can gradually reroute the traffic from the old blue environment to the new green environment. Blue/green deployments can mitigate common risks associated with deploying software, such as downtime and rollback capability.

Use AWS Global Accelerator to distribute a portion of traffic to a particular deployment

AWS Global Accelerator is a network layer service that directs traffic to optimal endpoints over the AWS global network, this improves the availability and performance of your internet applications. It provides two static anycast IP addresses that act as a fixed entry point to your application endpoints in a single or multiple AWS Regions, such as your Application Load Balancers, Network Load Balancers, Elastic IP addresses or Amazon EC2 instances, in a single or in multiple AWS regions.

AWS Global Accelerator uses endpoint weights to determine the proportion of traffic that is directed to endpoints in an endpoint group, and traffic dials to control the percentage of traffic that is directed to an endpoint group (an AWS region where your application is deployed).

While relying on the DNS service is a great option for blue/green deployments, it may not fit use-cases that require a fast and controlled transition of the traffic. Some client devices and internet resolvers cache DNS answers for long periods; this DNS feature improves the efficiency of the DNS service as it reduces the DNS traffic across the Internet, and serves as a resiliency technique by preventing authoritative name-server overloads. The downside of this in blue/green

deployments is that you don't know how long it will take before all of your users receive updated IP addresses when you update a record, change your routing preference or when there is an application failure.

With AWS Global Accelerator, you can shift traffic gradually or all at once between the blue and the green environment and vice-versa without being subject to DNS caching on client devices and internet resolvers, traffic dials and endpoint weights changes are effective within seconds.

Incorrect options:

Use Amazon Route 53 weighted routing to spread traffic across different deployments - Weighted routing lets you associate multiple resources with a single domain name (example.com) or subdomain name (acme.example.com) and choose how much traffic is routed to each resource. This can be useful for a variety of purposes, including load balancing and testing new versions of the software. As discussed earlier, DNS caching is a negative behavior for this use case and hence Amazon Route 53 is not a good option.

Use Elastic Load Balancing (ELB) to distribute traffic across deployments - Elastic Load Balancing (ELB) can distribute traffic across healthy instances. You can also use the Application Load Balancers weighted target groups feature for blue/green deployments as it does not rely on the DNS service. In addition you don't need to create new ALBs for the green environment. As the use-case refers to a global application, so this option cannot be used for a multi-Region solution which is needed for the given requirement.

Use AWS CodeDeploy deployment options to choose the right deployment - In AWS CodeDeploy, a deployment is the process, and the components involved in the process, of installing content on one or more instances. This content can consist of code, web and configuration files, executables, packages, scripts, and so on. AWS CodeDeploy deploys content that is stored in a source repository, according to the configuration rules you specify. Blue/Green deployment is one of the deployment types that CodeDeploy supports. CodeDeploy is not meant to distribute traffic across instances, so this option is incorrect.

References:

https://aws.amazon.com/blogs/networking-and-content-delivery/using-aws-global-accelerator-to-achieve-blue-green-deployments

https://docs.aws.amazon.com/codedeploy/latest/userguide/deployments.html

https://docs.aws.amazon.com/Route53/latest/DeveloperGuide/routing-policy.html#routing-policy-weighted

Domain

Design Resilient Architectures

Question 14 Correct

A media company has created an AWS Direct Connect connection for migrating its flagship application to the AWS Cloud. The on-premises application writes hundreds of video files into a mounted NFS file system daily. Post-migration, the company will host the application on an Amazon EC2 instance with a mounted Amazon Elastic File System (Amazon EFS) file system. Before the migration cutover, the company must build a process that will replicate the newly created on-premises video files to the Amazon EFS file system.

Which of the following represents the MOST operationally efficient way to meet this requirement?

Configure an AWS DataSync agent on the on-premises server that has access to the NFS file system. Transfer data over the AWS Direct Connect connection to an Amazon S3 bucket by using a VPC gateway endpoint for Amazon S3. Set up an AWS Lambda function to process event notifications from Amazon S3 and copy the video files from Amazon S3 to the Amazon EFS file system

Configure an AWS DataSync agent on the on-premises server that has access to the NFS file system. Transfer data over the AWS Direct Connect connection to an AWS VPC peering endpoint for Amazon EFS by using a private VIF. Set up an AWS DataSync scheduled task to send the video files to the Amazon EFS file system every 24 hours

Your answer is correct

Configure an AWS DataSync agent on the on-premises server that has access to the NFS file system. Transfer data over the AWS Direct Connect connection to an AWS PrivateLink interface VPC endpoint for Amazon EFS by using a private VIF. Set up an AWS DataSync scheduled task to send the video files to the Amazon EFS file system every 24 hours

Configure an AWS DataSync agent on the on-premises server that has access to the NFS file system. Transfer data over the AWS Direct Connect connection to an Amazon S3 bucket by using public VIF. Set up an AWS Lambda function to process event notifications from Amazon S3 and copy the video files from Amazon S3 to the Amazon EFS file system

Overall explanation

Correct option:

Configure an AWS DataSync agent on the on-premises server that has access to the NFS file system. Transfer data over the AWS Direct Connect connection to an AWS PrivateLink interface VPC endpoint for Amazon EFS by using a private VIF. Set up an AWS DataSync scheduled task to send the video files to the Amazon EFS file system every 24 hours

AWS DataSync is an online data transfer service that simplifies, automates, and accelerates copying large amounts of data between on-premises storage systems and AWS Storage services, as well as between AWS Storage services.

You can use AWS DataSync to migrate data located on-premises, at the edge, or in other clouds to Amazon S3, Amazon EFS, Amazon FSx for Windows File Server, Amazon FSx for Lustre, Amazon FSx for OpenZFS, and Amazon FSx for NetApp ONTAP.

AWS DataSync:

via - https://aws.amazon.com/datasync/

To establish a private connection between your virtual private cloud (VPC) and the Amazon EFS API, you can create an interface VPC endpoint. You can also access the interface VPC endpoint from on-premises environments or other VPCs using AWS VPN, AWS Direct Connect, or VPC peering.

AWS Direct Connect provides three types of virtual interfaces: public, private, and transit.

AWS Direct Connect VIFs:

Which type of Direct Connect virtual interface should I use to connect different AWS resources?

Last updated: 2022-01-21

AWS Direct Connect provides three types of virtual interfaces: public, private, and transit. How do I determine which type I should use to connect different AWS resources?

Resolution

Use the following Direct Connect virtual interface based on your use case.

Public virtual interface

To connect to AWS resources that are reachable by a public IP address such as an Amazon Simple Storage Service (Amazon S3) bucket or AWS public endpoints, use a public virtual interface. With a public virtual interface, you can:

- Connect to all AWS public IP addresses globally.
- Create public virtual interfaces in any Direct Connect location to receive Amazon's global IP routes.
- Access publicly routable Amazon services in any AWS Region (except the AWS China Region).

Private virtual interface

To connect to your resources hosted in an Amazon Virtual Private Cloud (Amazon VPC) using their private IP addresses, use a private virtual interface. With a private virtual interface, you can:

- Connect VPC resources such as Amazon Elastic Compute Cloud (Amazon EC2) instances or load balancers on your private IP address or endpoint.
- Connect a private virtual interface to a Direct Connect gateway. Then, associate the Direct Connect gateway with one or more virtual private gateways in any AWS Region (except the AWS China Region).
- Connect to multiple Amazon VPCs in any AWS Region (except the AWS China Region), because a virtual private gateway is associated with a single VPC.

Note: For a private virtual interface, AWS advertises the VPC CIDR only over the Border Gateway Protocol (BGP) neighbor. AWS can't advertise or suppress specific subnet blocks in the Amazon VPC for a private virtual interface.

Transit virtual interface

To connect to your resources hosted in an Amazon VPC (using their private IP addresses) through a transit gateway, use a transit virtual interface. With a transit virtual interface, you can:

- Connect multiple Amazon VPCs in the same or different AWS account using Direct Connect.
- Associate up to three transit gateways in any AWS Region when you use a transit virtual interface to connect to a Direct Connect gateway.
- Attach Amazon VPCs in the same AWS Region to the transit gateway. Then, access multiple VPCs in different AWS accounts in the same AWS Region using a transit virtual interface.

Note: For transit virtual interface, AWS advertises only routes that you specify in the allowed prefixes list on the Direct Connect gateway. For a list of all AWS Regions that offer Direct Connect support for AWS Transit Gateway, see AWS Transit Gateway support.

via - https://aws.amazon.com/premiumsupport/knowledge-center/public-private-interface-dx/

For the given use case, you can send data over the Direct Connect connection to an AWS PrivateLink interface VPC endpoint for Amazon EFS by using a private VIF.

Using task scheduling in AWS DataSync, you can periodically execute a transfer task from your source storage system to the destination. You can use the DataSync scheduled task to send the video files to the Amazon EFS file system every 24 hours.

Incorrect options:

Configure an AWS DataSync agent on the on-premises server that has access to the NFS file system. Transfer data over the AWS Direct Connect connection to an AWS VPC peering endpoint for Amazon EFS by using a private VIF. Set up an AWS DataSync scheduled task to send the video files to the Amazon EFS file system every 24 hours - A VPC peering connection is a networking connection between two VPCs that enables you to route traffic between them privately. You cannot use VPC peering to transfer data over the Direct Connect connection from the on-premises systems to AWS. So this option is incorrect.

Configure an AWS DataSync agent on the on-premises server that has access to the NFS file system. Transfer data over the AWS Direct Connect connection to an Amazon S3 bucket by

using public VIF. Set up an AWS Lambda function to process event notifications from Amazon S3 and copy the video files from Amazon S3 to the Amazon EFS file system - You can use a public virtual interface to connect to AWS resources that are reachable by a public IP address such as an Amazon Simple Storage Service (Amazon S3) bucket or AWS public endpoints. Although it is theoretically possible to set up this solution, however, it is not the most operationally efficient solution, since it involves sending data via AWS DataSync to Amazon S3 and then in turn using an AWS Lambda function to finally send data to Amazon EFS.

Configure an AWS DataSync agent on the on-premises server that has access to the NFS file system. Transfer data over the AWS Direct Connect connection to an Amazon S3 bucket by using a VPC gateway endpoint for Amazon S3. Set up an AWS Lambda function to process event notifications from Amazon S3 and copy the video files from Amazon S3 to the Amazon EFS file system - You can access Amazon S3 from your VPC using gateway VPC endpoints. You cannot use the Amazon S3 gateway endpoint to transfer data over the AWS Direct Connect connection from the on-premises systems to Amazon S3. So this option is incorrect.

References:

https://aws.amazon.com/datasync/

https://aws.amazon.com/blogs/storage/transferring-files-from-on-premises-to-aws-and-back-without-leaving-your-vpc-using-aws-datasync/

https://docs.aws.amazon.com/efs/latest/ug/efs-vpc-endpoints.html

https://aws.amazon.com/datasync/faqs/

https://aws.amazon.com/premiumsupport/knowledge-center/public-private-interface-dx/

https://docs.aws.amazon.com/datasync/latest/userguide/task-scheduling.html

Domain

Design High-Performing Architectures

Question 15 **Correct**

An organization wants to delegate access to a set of users from the development environment so that they can access some resources in the production environment which is managed under another AWS account.

As a solutions architect, which of the following steps would you recommend?

Your answer is correct

Create a new IAM role with the required permissions to access the resources in the production environment. The users can then assume this IAM role while accessing the resources from the production environment

Create new IAM user credentials for the production environment and share these credentials with the set of users from the development environment

It is not possible to access cross-account resources

Both IAM roles and IAM users can be used interchangeably for cross-account access

Overall explanation

Correct option:

Create a new IAM role with the required permissions to access the resources in the production environment. The users can then assume this IAM role while accessing the resources from the production environment

IAM roles allow you to delegate access to users or services that normally don't have access to your organization's AWS resources. IAM users or AWS services can assume a role to obtain temporary security credentials that can be used to make AWS API calls. Consequently, you don't have to share long-term credentials for access to a resource. Using IAM roles, it is possible to access cross-account resources.

Incorrect options:

Create new IAM user credentials for the production environment and share these credentials with the set of users from the development environment - There is no need to create new

IAM user credentials for the production environment, as you can use IAM roles to access cross-account resources.

It is not possible to access cross-account resources - You can use IAM roles to access cross-account resources.

Both IAM roles and IAM users can be used interchangeably for cross-account access - IAM roles and IAM users are separate IAM entities and should not be mixed. Only IAM roles can be used to access cross-account resources.

Reference:

https://aws.amazon.com/iam/features/manage-roles/

Domain

Design Secure Architectures

Question 16 Correct

A company has noticed that its application performance has deteriorated after a new Auto Scaling group was deployed a few days back. Upon investigation, the team found out that the Launch Configuration selected for the Auto Scaling group is using the incorrect instance type that is not optimized to handle the application workflow.

As a solutions architect, what would you recommend to provide a long term resolution for this issue?

Modify the launch configuration to use the correct instance type and continue to use the existing Auto Scaling group

No need to modify the launch configuration. Just modify the Auto Scaling group to use the correct instance type

Your answer is correct

Create a new launch configuration to use the correct instance type. Modify the Auto Scaling group to use this new launch configuration. Delete the old launch configuration as it is no longer needed

No need to modify the launch configuration. Just modify the Auto Scaling group to use more number of existing instance types. More instances may offset the loss of performance

Overall explanation

Correct option:

Create a new launch configuration to use the correct instance type. Modify the Auto Scaling group to use this new launch configuration. Delete the old launch configuration as it is no longer needed

A launch configuration is an instance configuration template that an Auto Scaling group uses to launch Amazon EC2 instances. When you create a launch configuration, you specify information for the instances. Include the ID of the Amazon Machine Image (AMI), the instance type, a key pair, one or more security groups, and a block device mapping.

It is not possible to modify a launch configuration once it is created. The correct option is to create a new launch configuration to use the correct instance type. Then modify the Auto Scaling group to use this new launch configuration. Lastly to clean-up, just delete the old launch configuration as it is no longer needed.

Incorrect options:

Modify the launch configuration to use the correct instance type and continue to use the existing Auto Scaling group - As mentioned earlier, it is not possible to modify a launch configuration once it is created. Hence, this option is incorrect.

No need to modify the launch configuration. Just modify the Auto Scaling group to use the correct instance type - You cannot use an Auto Scaling group to directly modify the instance type of the underlying instances. Hence, this option is incorrect.

No need to modify the launch configuration. Just modify the Auto Scaling group to use more number of existing instance types. More instances may offset the loss of performance - Using the Auto Scaling group to increase the number of instances to cover up for the performance loss is not recommended as it does not address the root cause of the problem.

The Machine Learning workflow requires a certain instance type that is optimized to handle Machine Learning computations. Hence, this option is incorrect.

Reference:

https://docs.aws.amazon.com/autoscaling/ec2/userguide/LaunchConfiguration.html

Domain

Design Resilient Architectures

Question 17 Correct

A Big Data analytics company writes data and log files in Amazon S3 buckets. The company now wants to stream the existing data files as well as any ongoing file updates from Amazon S3 to Amazon Kinesis Data Streams.

As a Solutions Architect, which of the following would you suggest as the fastest possible way of building a solution for this requirement?

Configure Amazon EventBridge events for the bucket actions on Amazon S3. An AWS Lambda function can then be triggered from the Amazon EventBridge event that will send the necessary data to Amazon Kinesis Data Streams

Leverage Amazon S3 event notification to trigger an AWS Lambda function for the file create event. The AWS Lambda function will then send the necessary data to Amazon Kinesis Data Streams

Amazon S3 bucket actions can be directly configured to write data into Amazon Simple Notification Service (Amazon SNS). Amazon SNS can then be used to send the updates to Amazon Kinesis Data Streams

Your answer is correct

Leverage AWS Database Migration Service (AWS DMS) as a bridge between Amazon S3 and Amazon Kinesis Data Streams

Overall explanation

Correct option:

Leverage AWS Database Migration Service (AWS DMS) as a bridge between Amazon S3 and Amazon Kinesis Data Streams

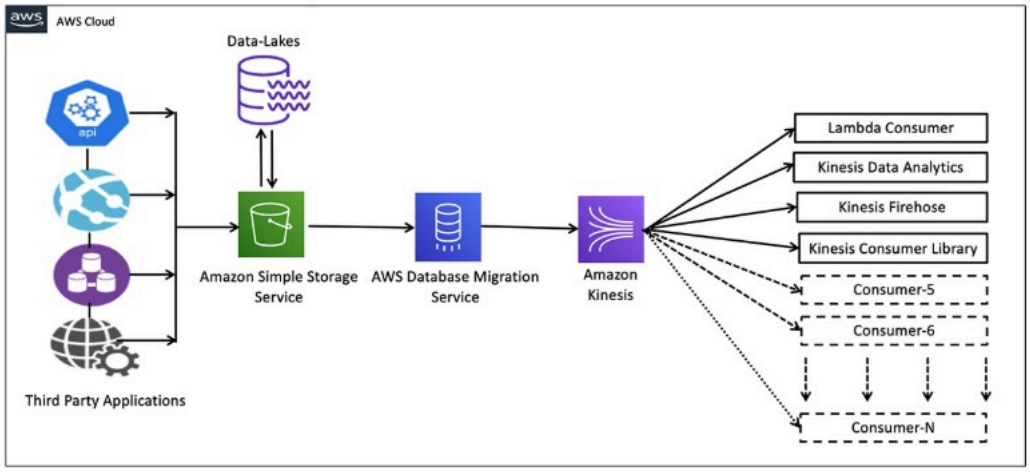
You can achieve this by using AWS Database Migration Service (AWS DMS). AWS DMS enables you to seamlessly migrate data from supported sources to relational databases, data warehouses, streaming platforms, and other data stores in AWS cloud.

The given requirement needs the functionality to be implemented in the least possible time. You can use AWS DMS for such data-processing requirements. AWS DMS lets you expand the existing application to stream data from Amazon S3 into Amazon Kinesis Data Streams for real-time analytics without writing and maintaining new code. AWS DMS supports specifying Amazon S3 as the source and streaming services like Kinesis and Amazon Managed Streaming of Kafka (Amazon MSK) as the target. AWS DMS allows migration of full and change data capture (CDC) files to these services. AWS DMS performs this task out of box without any complex configuration or code development. You can also configure an AWS DMS replication instance to scale up or down depending on the workload.

AWS DMS supports Amazon S3 as the source and Kinesis as the target, so data stored in an S3 bucket is streamed to Kinesis. Several consumers, such as AWS Lambda, Amazon Kinesis Data Firehose, Amazon Kinesis Data Analytics, and the Kinesis Consumer Library (KCL), can consume the data concurrently to perform real-time analytics on the dataset. Each AWS service in this architecture can scale independently as needed.

Architecture of the proposed solution:

The following diagram shows the architecture of this solution.



via - https://aws.amazon.com/blogs/big-data/streaming-data-from-amazon-s3-to-amazon-kinesis-data-streams-using-aws-dms/

Incorrect options:

Configure Amazon EventBridge events for the bucket actions on Amazon S3. An AWS Lambda function can then be triggered from the Amazon EventBridge event that will send the necessary data to Amazon Kinesis Data Streams - You will need to enable AWS Cloudtrail trail to use object-level actions as a trigger for Amazon EventBridge events. Also, using AWS Lambda functions would require significant custom development to write the data into Amazon Kinesis Data Streams, so this option is not the right fit.

Leverage Amazon S3 event notification to trigger an AWS Lambda function for the file create event. The AWS Lambda function will then send the necessary data to Amazon Kinesis Data Streams - Using AWS Lambda functions would require significant custom development to write the data into Amazon Kinesis Data Streams, so this option is not the right fit.

Amazon S3 bucket actions can be directly configured to write data into Amazon Simple Notification Service (Amazon SNS). Amazon SNS can then be used to send the updates to Amazon Kinesis Data Streams - Amazon S3 cannot directly write data into Amazon SNS, although it can certainly use Amazon S3 event notifications to send an event to Amazon SNS. Also, Amazon SNS cannot directly send messages to Amazon Kinesis Data Streams. So this option is incorrect.

Reference:

https://aws.amazon.com/blogs/big-data/streaming-data-from-amazon-s3-to-amazon-kinesis-data-streams-using-aws-dms/

Domain

Design High-Performing Architectures

Question 18 **Correct**

A company manages a multi-tier social media application that runs on Amazon Elastic Compute Cloud (Amazon EC2) instances behind an Application Load Balancer. The instances run in an Amazon EC2 Auto Scaling group across multiple Availability Zones (AZs) and use an Amazon Aurora database. As an AWS Certified Solutions Architect – Associate, you have been tasked to make the application more resilient to periodic spikes in request rates.

Which of the following solutions would you recommend for the given use-case? (Select two)

Use AWS Direct Connect

Use AWS Shield

Use AWS Global Accelerator

Your selection is correct

Use Amazon CloudFront distribution in front of the Application Load Balancer

Your selection is correct

Use Amazon Aurora Replica

Overall explanation

Correct options:

You can use Amazon Aurora replicas and Amazon CloudFront distribution to make the application more resilient to spikes in request rates.

Use Amazon Aurora Replica

Amazon Aurora Replicas have two main purposes. You can issue queries to them to scale the read operations for your application. You typically do so by connecting to the reader endpoint of the cluster. That way, Aurora can spread the load for read-only connections across as many Aurora Replicas as you have in the cluster. Amazon Aurora Replicas also help to increase availability. If the writer instance in a cluster becomes unavailable, Aurora automatically promotes one of the reader instances to take its place as the new writer. Up to 15 Aurora Replicas can be distributed across the Availability Zones (AZs) that a DB cluster spans within an AWS Region.

Use Amazon CloudFront distribution in front of the Application Load Balancer

Amazon CloudFront is a fast content delivery network (CDN) service that securely delivers data, videos, applications, and APIs to customers globally with low latency, high transfer speeds, all within a developer-friendly environment. CloudFront points of presence (POPs) (edge locations) make sure that popular content can be served quickly to your viewers. Amazon CloudFront also has regional edge caches that bring more of your content closer to your viewers, even when the content is not popular enough to stay at a POP, to help improve performance for that content.

Amazon CloudFront offers an origin failover feature to help support your data resiliency needs. Amazon CloudFront is a global service that delivers your content through a worldwide network of data centers called edge locations or points of presence (POPs). If your content is not already cached in an edge location, Amazon CloudFront retrieves it from an origin that you've identified as the source for the definitive version of the content.

Incorrect options:

Use AWS Shield - AWS Shield is a managed Distributed Denial of Service (DDoS) protection service that safeguards applications running on AWS. AWS Shield provides always-on detection and automatic inline mitigations that minimize application downtime and latency. There are two tiers of AWS Shield - Standard and Advanced. AWS Shield cannot be used to improve application resiliency to handle spikes in traffic.

Use AWS Global Accelerator - AWS Global Accelerator is a service that improves the availability and performance of your applications with local or global users. It provides static IP addresses that act as a fixed entry point to your application endpoints in a single or multiple AWS Regions, such as your Application Load Balancers, Network Load Balancers or Amazon EC2 instances. Amazon Global Accelerator is a good fit for non-HTTP use cases, such as gaming (UDP), IoT (MQTT), or Voice over IP, as well as for HTTP use cases that specifically require static IP addresses or deterministic, fast regional failover. Since Amazon CloudFront is better for improving application resiliency to handle spikes in traffic, so this option is ruled out.

Use AWS Direct Connect - AWS Direct Connect lets you establish a dedicated network connection between your network and one of the AWS Direct Connect locations. Using industry-standard 802.1q VLANs, this dedicated connection can be partitioned into multiple virtual interfaces. AWS Direct Connect does not involve the Internet; instead, it uses dedicated, private network connections between your intranet and Amazon VPC. AWS Direct Connect cannot be used to improve application resiliency to handle spikes in traffic.

References:

https://docs.aws.amazon.com/AmazonRDS/latest/AuroraUserGuide/disaster-recovery-resiliency.html

https://docs.aws.amazon.com/AmazonRDS/latest/AuroraUserGuide/Aurora.Replication.html

https://aws.amazon.com/global-accelerator/faqs/

https://docs.aws.amazon.com/global-accelerator/latest/dg/disaster-recovery-resiliency.html

Domain

Design Resilient Architectures

Question 19 **Correct**

A retail company uses AWS Cloud to manage its technology infrastructure. The company has deployed its consumer-focused web application on Amazon EC2-based web servers and uses Amazon RDS PostgreSQL database as the data store. The PostgreSQL database is set up in a private subnet that allows inbound traffic from selected Amazon EC2 instances. The database also uses AWS Key Management Service (AWS KMS) for encrypting data at rest.

Which of the following steps would you recommend to facilitate end-to-end security for the data-in-transit while accessing the database?

Create a new security group that blocks SSH from the selected Amazon EC2 instances into the database

Use IAM authentication to access the database instead of the database user's access credentials

Your answer is correct

Configure Amazon RDS to use SSL for data in transit

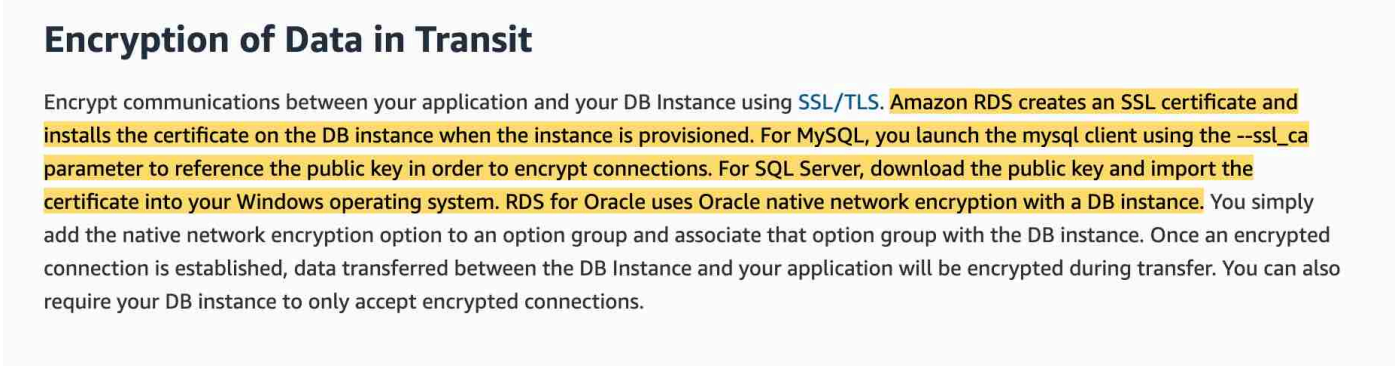
Create a new network access control list (network ACL) that blocks SSH from the entire Amazon EC2 subnet into the database

Overall explanation

Correct option:

Configure Amazon RDS to use SSL for data in transit

You can use Secure Socket Layer / Transport Layer Security (SSL/TLS) connections to encrypt data in transit. Amazon RDS creates an SSL certificate and installs the certificate on the DB instance when the instance is provisioned. For MySQL, you launch the MySQL client using the --ssl_ca parameter to reference the public key to encrypt connections. Using SSL, you can encrypt a PostgreSQL connection between your applications and your PostgreSQL DB instances. You can also force all connections to your PostgreSQL DB instance to use SSL.



via - https://aws.amazon.com/rds/features/security/

Incorrect options:

Use IAM authentication to access the database instead of the database user's access credentials - You can authenticate to your database instance using AWS Identity and Access Management (IAM) database authentication. IAM database authentication works with MySQL and PostgreSQL. With this authentication method, you don't need to use a password when you connect to a database instance. Instead, you use an authentication token.

IAM authentication is just another way to authenticate the user's credentials while accessing the database. It would not significantly enhance the security in a way that enabling SSL does by facilitating the in-transit encryption for the database.

Create a new security group that blocks SSH from the selected Amazon EC2 instances into the database

Create a new network access control list (network ACL) that blocks SSH from the entire Amazon EC2 subnet into the database

Both these options are added as distractors. You cannot SSH into an Amazon RDS database instance.

References:

https://aws.amazon.com/rds/features/security/

https://docs.aws.amazon.com/AmazonRDS/latest/UserGuide/CHAP_MySQL.html#MySQL.Concepts.SSLS

https://docs.aws.amazon.com/AmazonRDS/latest/UserGuide/CHAP_PostgreSQL.html#PostgreSQL.Conce

https://aws.amazon.com/blogs/database/using-iam-authentication-to-connect-with-pgadmin-amazon-aurora-postgresql-or-amazon-rds-for-postgresql/

Domain

Design Secure Architectures

Question 20 **Incorrect**

A media company wants to get out of the business of owning and maintaining its own IT infrastructure. As part of this digital transformation, the media company wants to archive about 5 petabytes of data in its on-premises data center to durable long term storage.

As a solutions architect, what is your recommendation to migrate this data in the MOST cost-optimal way?

Your answer is incorrect

Transfer the on-premises data into multiple AWS Snowball Edge Storage Optimized devices. Copy the AWS Snowball Edge data into Amazon S3 Glacier

Correct answer

Transfer the on-premises data into multiple AWS Snowball Edge Storage Optimized devices. Copy the AWS Snowball Edge data into Amazon S3 and create a lifecycle policy to transition the data into Amazon S3 Glacier

Setup AWS direct connect between the on-premises data center and AWS Cloud. Use this connection to transfer the data into Amazon S3 Glacier

Setup AWS Site-to-Site VPN connection between the on-premises data center and AWS Cloud. Use this connection to transfer the data into Amazon S3 Glacier

Overall explanation

Correct option:

Transfer the on-premises data into multiple AWS Snowball Edge Storage Optimized devices. Copy the AWS Snowball Edge data into Amazon S3 and create a lifecycle policy to transition

the data into Amazon S3 Glacier

AWS Snowball Edge Storage Optimized is the optimal choice if you need to securely and quickly transfer dozens of terabytes to petabytes of data to AWS. It provides up to 80 TB of usable HDD storage, 40 vCPUs, 1 TB of SATA SSD storage, and up to 40 Gb network connectivity to address large scale data transfer and pre-processing use cases. The data stored on AWS Snowball Edge device can be copied into Amazon S3 bucket and later transitioned into Amazon S3 Glacier via a lifecycle policy. You can't directly copy data from AWS Snowball Edge devices into Amazon S3 Glacier.

Incorrect options:

Transfer the on-premises data into multiple AWS Snowball Edge Storage Optimized devices. Copy the AWS Snowball Edge data into Amazon S3 Glacier - As mentioned earlier, you can't directly copy data from AWS Snowball Edge devices into Amazon S3 Glacier. Hence, this option is incorrect.

Setup AWS direct connect between the on-premises data center and AWS Cloud. Use this connection to transfer the data into Amazon S3 Glacier - AWS Direct Connect lets you establish a dedicated network connection between your network and one of the AWS Direct Connect locations. Using industry-standard 802.1q VLANs, this dedicated connection can be partitioned into multiple virtual interfaces. Direct Connect involves significant monetary investment and takes more than a month to set up, therefore it's not the correct fit for this use-case where just a one-time data transfer has to be done.

Setup AWS Site-to-Site VPN connection between the on-premises data center and AWS Cloud. Use this connection to transfer the data into Amazon S3 Glacier - AWS Site-to-Site VPN enables you to securely connect your on-premises network or branch office site to your Amazon Virtual Private Cloud (Amazon VPC). VPN Connections are a good solution if you have an immediate need, and have low to modest bandwidth requirements. Because of the high data volume for the given use-case, Site-to-Site VPN is not the correct choice.

Reference:

https://aws.amazon.com/snowball/

Domain

Design Cost-Optimized Architectures

Question 21 **Correct**

A retail company maintains an AWS Direct Connect connection to AWS and has recently migrated its data warehouse to AWS. The data analysts at the company query the data warehouse using a visualization tool. The average size of a query returned by the data warehouse is 60 megabytes and

the query responses returned by the data warehouse are not cached in the visualization tool. Each webpage returned by the visualization tool is approximately 600 kilobytes.

Which of the following options offers the LOWEST data transfer egress cost for the company?

Your answer is correct

Deploy the visualization tool in the same AWS region as the data warehouse. Access the visualization tool over a Direct Connect connection at a location in the same region

Deploy the visualization tool in the same AWS region as the data warehouse. Access the visualization tool over the internet at a location in the same region

Deploy the visualization tool on-premises. Query the data warehouse directly over an AWS Direct Connect connection at a location in the same AWS region

Deploy the visualization tool on-premises. Query the data warehouse over the internet at a location in the same AWS region

Overall explanation

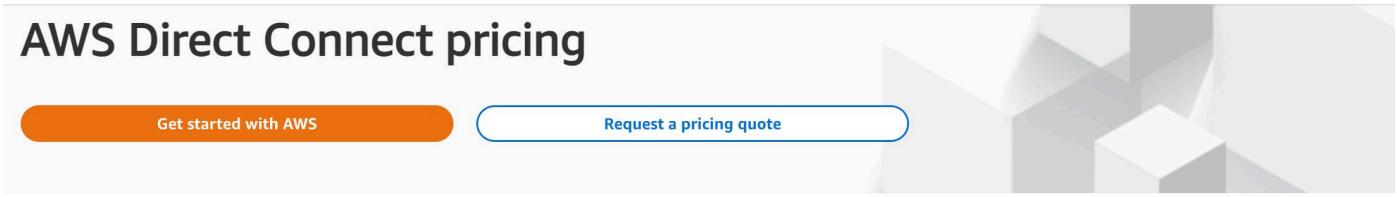
Correct option:

Deploy the visualization tool in the same AWS region as the data warehouse. Access the visualization tool over a Direct Connect connection at a location in the same region

AWS Direct Connect is a networking service that provides an alternative to using the internet to connect to AWS. Using AWS Direct Connect, data that would have previously been transported

over the internet is delivered through a private network connection between your on-premises data center and AWS.

For the given use case, the main pricing parameter while using the AWS Direct Connect connection is the Data Transfer Out (DTO) from AWS to the on-premises data center. DTO refers to the cumulative network traffic that is sent through AWS Direct Connect to destinations outside of AWS. This is charged per gigabyte (GB), and unlike capacity measurements, DTO refers to the amount of data transferred, not the speed.



AWS Direct Connect is a cloud service that links your network directly to AWS to deliver consistent, low-latency performance. With AWS Direct Connect, you pay only for what you use and there is no minimum fee. There are no setup charges, and you may cancel at any time. However, services provided by your AWS Direct Connect Delivery Partners or other local service provider may have other terms that apply.

Once you have linked your locations to AWS Direct Connect, you can send data between them using SiteLink. When using SiteLink, data travels over the shortest path between locations. The SiteLink feature is off by default and can be turned on or off at any time.

Pricing components

When connecting to resources running in any AWS Region (such as an Amazon Virtual Private Cloud or AWS Transit Gateway), there are three factors that determine pricing: capacity, port hours, and data transfer out (DTO).

Capacity is the maximum rate that data can be transferred through a network connection. The capacity of AWS Direct Connect connections are measured in megabit per second (Mbps) or gigabit per second (Gbps). One gigabit per second, or 1 Gbps, is equal to 1,000 megabits per second (1,000 Mbps).

Port hours measure the time that a port is provisioned for your use with AWS, or an AWS Direct Connect Delivery Partner's, networking equipment inside an AWS Direct Connect location. Even when no data is passing through the port, you are charged for port hours. Port hour pricing is determined by the connection type: dedicated or hosted.

> **Dedicated connections** are physical connections between your network port and an AWS network port inside an AWS Direct Connect location. Dedicated port hours are billed as long as that port is provisioned for your use. You request a dedicated connection through the AWS Direct Connect section of the AWS Management Console.
>
> **Hosted connections** are logical connections that an AWS Direct Connect Delivery Partner provisions on your behalf. When using hosted connections, you connect to the AWS network using one of the partner's ports. You request a hosted connection by contacting an AWS Direct Connect Delivery Partner directly.

Data transfer out (DTO) refers to the cumulative network traffic that is sent through AWS Direct Connect to destinations outside of AWS. This is charged per gigabyte (GB), and unlike capacity measurements, DTO refers to the amount of data transferred, not the speed. When calculating DTO, exact pricing depends on the AWS Region or AWS Local Zone, and the AWS Direct Connect location, you are using (see tables below).

Data transfer in refers to network traffic that is sent into AWS from outside, over AWS Direct Connect. AWS Direct Connect data transfer in is charged at 0.00 USD per GB in all locations.

Except as otherwise noted, our prices are exclusive of applicable taxes and duties, including VAT and applicable sales tax. For customers with a Japanese billing address, use of the Asia Pacific (Tokyo) Region is subject to Japanese Consumption Tax. Learn more.

via - https://aws.amazon.com/directconnect/pricing/

Each query response is 60 megabytes in size and each webpage for the visualization tool is 600 kilobytes in size. If you deploy the visualization tool in the same AWS region as the data warehouse, then you only need to pay for the 600 kilobytes of DTO charges for the webpage. Therefore this option is correct.

However, if you deploy the visualization tool on-premises, then you need to pay for the 60 MB of DTO charges for the query response from the data warehouse to the visualization tool.

Incorrect options:

Deploy the visualization tool in the same AWS region as the data warehouse. Access the visualization tool over the internet at a location in the same region

Deploy the visualization tool on-premises. Query the data warehouse over the internet at a location in the same AWS region

Data transfer pricing over AWS Direct Connect is lower than data transfer pricing over the internet, so both of these options are incorrect.

Deploy the visualization tool on-premises. Query the data warehouse directly over an AWS Direct Connect connection at a location in the same AWS region - As mentioned in the explanation above, if you deploy the visualization tool on-premises, then you need to pay for the 60 megabytes of DTO charges for the query response from the data warehouse to the visualization tool. So this option is incorrect.

References:

https://aws.amazon.com/directconnect/pricing/

https://aws.amazon.com/getting-started/hands-on/connect-data-center-to-aws/services-costs/

https://aws.amazon.com/directconnect/faqs/

Domain

Design Cost-Optimized Architectures

Question 22 **Correct**

A retail company uses AWS Cloud to manage its IT infrastructure. The company has set up AWS Organizations to manage several departments running their AWS accounts and using resources such as Amazon EC2 instances and Amazon RDS databases. The company wants to provide shared and centrally-managed VPCs to all departments using applications that need a high degree of interconnectivity.

As a solutions architect, which of the following options would you choose to facilitate this use-case?

Use VPC sharing to share a VPC with other AWS accounts belonging to the same parent organization from AWS Organizations

Use VPC peering to share one or more subnets with other AWS accounts belonging to the same parent organization from AWS Organizations

Your answer is correct

Use VPC sharing to share one or more subnets with other AWS accounts belonging to the same parent organization from AWS Organizations

Use VPC peering to share a VPC with other AWS accounts belonging to the same parent organization from AWS Organizations

Overall explanation

Correct option:

Use VPC sharing to share one or more subnets with other AWS accounts belonging to the same parent organization from AWS Organizations

VPC sharing (part of Resource Access Manager) allows multiple AWS accounts to create their application resources such as Amazon EC2 instances, Amazon RDS databases, Amazon Redshift clusters, and AWS Lambda functions, into shared and centrally-managed Amazon Virtual Private Clouds (VPCs). To set this up, the account that owns the VPC (owner) shares one or more subnets with other accounts (participants) that belong to the same organization from AWS Organizations. After a subnet is shared, the participants can view, create, modify, and delete their application resources in the subnets shared with them. Participants cannot view, modify, or delete resources that belong to other participants or the VPC owner.

You can share Amazon VPCs to leverage the implicit routing within a VPC for applications that require a high degree of interconnectivity and are within the same trust boundaries. This reduces the number of VPCs that you create and manage while using separate accounts for billing and access control.

Incorrect options:

Use VPC sharing to share a VPC with other AWS accounts belonging to the same parent organization from AWS Organizations - Using VPC sharing, an account that owns the VPC (owner) shares one or more subnets with other accounts (participants) that belong to the same organization from AWS Organizations. The owner account cannot share the VPC itself. Therefore this option is incorrect.

Use VPC peering to share a VPC with other AWS accounts belonging to the same parent organization from AWS Organizations - A VPC peering connection is a networking connection between two VPCs that enables you to route traffic between them using private IPv4 addresses or IPv6 addresses. Instances in either VPC can communicate with each other as if they are within the same network. VPC peering does not facilitate centrally managed VPCs. Therefore this option is incorrect.

Use VPC peering to share one or more subnets with other AWS accounts belonging to the same parent organization from AWS Organizations - A VPC peering connection is a networking connection between two VPCs that enables you to route traffic between them using private IPv4 addresses or IPv6 addresses. Instances in either VPC can communicate with each other as if they are within the same network. VPC peering does not facilitate centrally managed VPCs. Moreover, an AWS owner account cannot share the VPC itself with another AWS account. Therefore this option is incorrect.

References:

https://docs.aws.amazon.com/vpc/latest/userguide/vpc-sharing.html

https://docs.aws.amazon.com/vpc/latest/peering/what-is-vpc-peering.html

Domain

Design Secure Architectures

Question 23 **Correct**

A big data analytics company is using Amazon Kinesis Data Streams (KDS) to process IoT data from the field devices of an agricultural sciences company. Multiple consumer applications are using the incoming data streams and the engineers have noticed a performance lag for the data delivery speed between producers and consumers of the data streams.

As a solutions architect, which of the following would you recommend for improving the performance for the given use-case?

Swap out Amazon Kinesis Data Streams with Amazon SQS FIFO queues

Swap out Amazon Kinesis Data Streams with Amazon SQS Standard queues

Your answer is correct

Use Enhanced Fanout feature of Amazon Kinesis Data Streams

Swap out Amazon Kinesis Data Streams with Amazon Kinesis Data Firehose

Overall explanation

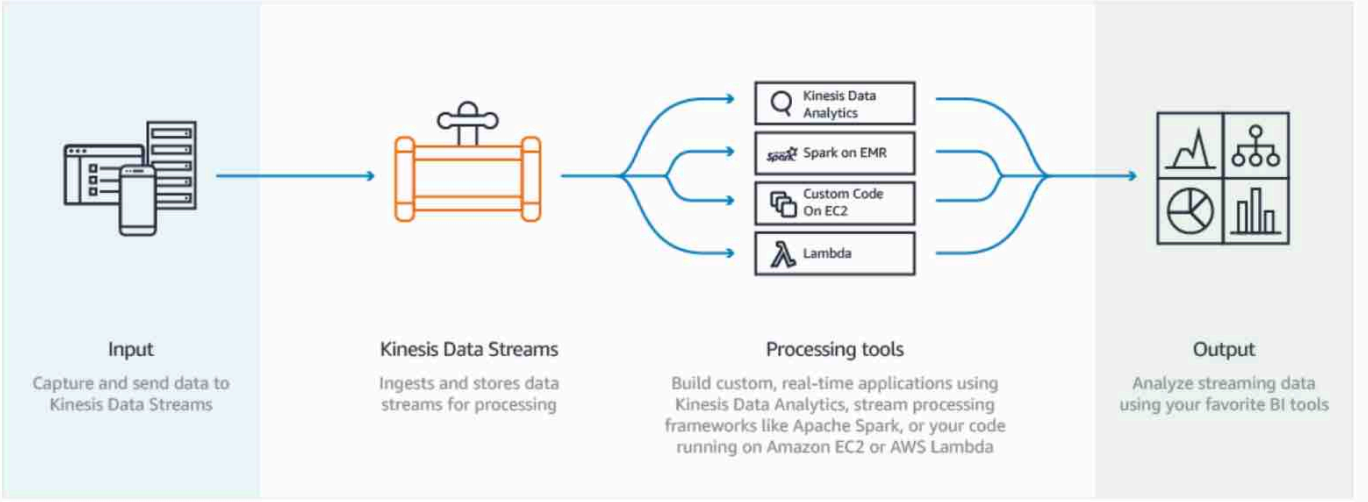
Correct option:

Use Enhanced Fanout feature of Amazon Kinesis Data Streams

Amazon Kinesis Data Streams (KDS) is a massively scalable and durable real-time data streaming service. KDS can continuously capture gigabytes of data per second from hundreds of thousands of sources such as website clickstreams, database event streams, financial transactions, social media feeds, IT logs, and location-tracking events.

By default, the 2MB/second/shard output is shared between all of the applications consuming data from the stream. You should use enhanced fan-out if you have multiple consumers retrieving data from a stream in parallel. With enhanced fan-out developers can register stream consumers to use enhanced fan-out and receive their own 2MB/second pipe of read throughput per shard, and this throughput automatically scales with the number of shards in a stream.

Amazon Kinesis Data Streams Fanout:



Kinesis Data Streams are scaled using the concept of a **shard**. One shard provides an ingest capacity of 1MB/second or 1000 records/second and an output capacity of 2MB/second. It's not uncommon for customers to have thousands or tens of thousands of shards supporting 10s of GB/sec of ingest and egress. Before the enhanced fan-out capability, that 2MB/second/shard output was shared between all of the applications consuming data from the stream. With enhanced fan-out developers can register stream consumers to use enhanced fan-out and receive their own 2MB/second pipe of read throughput per shard, and this throughput automatically scales with the number of shards in a stream. Prior to the launch of Enhanced Fan-out customers would frequently fan-out their data out to multiple streams to support their desired read throughput for their downstream applications. That sounds like undifferentiated heavy lifting to us, and that's something we decided our customers shouldn't need to worry about. Customers pay for enhanced fan-out based on the amount of data retrieved from the stream using enhanced fan-out and the number of consumers registered per-shard. You can find additional info on the pricing page.

via - https://aws.amazon.com/blogs/aws/kds-enhanced-fanout/

Incorrect options:

Swap out Amazon Kinesis Data Streams with Amazon Kinesis Data Firehose - Amazon Kinesis Data Firehose is the easiest way to reliably load streaming data into data lakes, data stores, and analytics tools. It is a fully managed service that automatically scales to match the throughput of your data and requires no ongoing administration. It can also batch, compress, transform, and encrypt the data before loading it, minimizing the amount of storage used at the destination and increasing security. Amazon Kinesis Data Firehose can only write to Amazon S3, Amazon Redshift, Amazon Elasticsearch or Splunk. You can't have applications consuming data streams from Amazon Kinesis Data Firehose, that's the job of Amazon Kinesis Data Streams. Therefore this option is not correct.

Swap out Amazon Kinesis Data Streams with Amazon SQS Standard queues

Swap out Amazon Kinesis Data Streams with Amazon SQS FIFO queues

Amazon Simple Queue Service (Amazon SQS) is a fully managed message queuing service that enables you to decouple and scale microservices, distributed systems, and serverless applications. Amazon SQS offers two types of message queues. Standard queues offer maximum throughput, best-effort ordering, and at-least-once delivery. Amazon SQS FIFO queues are designed to guarantee that messages are processed exactly once, in the exact order that they are sent. As multiple applications are consuming the same stream concurrently, both Amazon SQS Standard and Amazon SQS FIFO are not the right fit for the given use-case.

Exam Alert:

Please understand the differences between the capabilities of Amazon Kinesis Data Streams vs Amazon SQS, as you may be asked scenario-based questions on this topic in the exam.

Q: When should I use Amazon Kinesis Data Streams, and when should I use Amazon SQS?

We recommend Amazon Kinesis Data Streams for use cases with requirements that are similar to the following:

- Routing related records to the same record processor (as in streaming MapReduce). For example, counting and aggregation are simpler when all records for a given key are routed to the same record processor.
- Ordering of records. For example, you want to transfer log data from the application host to the processing/archival host while maintaining the order of log statements.
- Ability for multiple applications to consume the same stream concurrently. For example, you have one application that updates a real-time dashboard and another that archives data to Amazon Redshift. You want both applications to consume data from the same stream concurrently and independently.
- Ability to consume records in the same order a few hours later. For example, you have a billing application and an audit application that runs a few hours behind the billing application. Because Amazon Kinesis Data Streams stores data for up to 7 days, you can run the audit application up to 7 days behind the billing application.

We recommend Amazon SQS for use cases with requirements that are similar to the following:

- Messaging semantics (such as message-level ack/fail) and visibility timeout. For example, you have a queue of work items and want to track the successful completion of each item independently. Amazon SQS tracks the ack/fail, so the application does not have to maintain a persistent checkpoint/cursor. Amazon SQS will delete acked messages and redeliver failed messages after a configured visibility timeout.
- Individual message delay. For example, you have a job queue and need to schedule individual jobs with a delay. With Amazon SQS, you can configure individual messages to have a delay of up to 15 minutes.
- Dynamically increasing concurrency/throughput at read time. For example, you have a work queue and want to add more readers until the backlog is cleared. With Amazon Kinesis Data Streams, you can scale up to a sufficient number of shards (note, however, that you'll need to provision enough shards ahead of time).
- Leveraging Amazon SQS's ability to scale transparently. For example, you buffer requests and the load changes as a result of occasional load spikes or the natural growth of your business. Because each buffered request can be processed independently, Amazon SQS can scale transparently to handle the load without any provisioning instructions from you.

via - https://aws.amazon.com/kinesis/data-streams/faqs/

References:

https://aws.amazon.com/blogs/aws/kds-enhanced-fanout/

https://aws.amazon.com/kinesis/data-streams/faqs/

Domain

Design High-Performing Architectures

Question 24 Incorrect

A media agency stores its re-creatable assets on Amazon Simple Storage Service (Amazon S3) buckets. The assets are accessed by a large number of users for the first few days and the frequency of access falls down drastically after a week. Although the assets would be accessed occasionally after the first week, but they must continue to be immediately accessible when required. The cost of

maintaining all the assets on Amazon S3 storage is turning out to be very expensive and the agency is looking at reducing costs as much as possible.

As an AWS Certified Solutions Architect – Associate, can you suggest a way to lower the storage costs while fulfilling the business requirements?

Configure a lifecycle policy to transition the objects to Amazon S3 Standard-Infrequent Access (S3 Standard-IA) after 7 days

Your answer is incorrect

Configure a lifecycle policy to transition the objects to Amazon S3 One Zone-Infrequent Access (S3 One Zone-IA) after 7 days

Configure a lifecycle policy to transition the objects to Amazon S3 Standard-Infrequent Access (S3 Standard-IA) after 30 days

Correct answer

Configure a lifecycle policy to transition the objects to Amazon S3 One Zone-Infrequent Access (S3 One Zone-IA) after 30 days

Overall explanation

Correct option:

Configure a lifecycle policy to transition the objects to Amazon S3 One Zone-Infrequent Access (S3 One Zone-IA) after 30 days

Amazon S3 One Zone-IA is for data that is accessed less frequently, but requires rapid access when needed. Unlike other S3 Storage Classes which store data in a minimum of three Availability Zones (AZs), Amazon S3 One Zone-IA stores data in a single Availability Zone (AZ) and costs 20% less than Amazon S3 Standard-IA. Amazon S3 One Zone-IA is ideal for customers who want a lower-cost option for infrequently accessed and re-creatable data but do not require the availability and resilience of Amazon S3 Standard or Amazon S3 Standard-IA. The minimum storage duration is 30 days before you can transition objects from Amazon S3 Standard to Amazon S3 One Zone-IA.

Amazon S3 One Zone-IA offers the same high durability, high throughput, and low latency of Amazon S3 Standard, with a low per GB storage price and per GB retrieval fee. S3 Storage Classes can be configured at the object level, and a single bucket can contain objects stored across Amazon S3 Standard, Amazon S3 Intelligent-Tiering, Amazon S3 Standard-IA, and Amazon S3 One Zone-IA. You can also use S3 Lifecycle policies to automatically transition objects between storage classes without any application changes.

Constraints for Lifecycle storage class transitions:

Constraints

Lifecycle storage class transitions have the following constraints:

Object size and transitions from S3 Standard or S3 Standard-IA to S3 Intelligent-Tiering, S3 Standard-IA, or S3 One Zone-IA

When you transition objects from the S3 Standard or S3 Standard-IA storage classes to S3 Intelligent-Tiering, S3 Standard-IA, or S3 One Zone-IA, the following object size constraints apply:

- **Larger objects** - For the following transitions, there is a cost benefit to transitioning larger objects:
  - From the S3 Standard or S3 Standard-IA storage classes to S3 Intelligent-Tiering.
  - From the S3 Standard storage class to S3 Standard-IA or S3 One Zone-IA.
- **Objects smaller than 128 KB** - For the following transitions, Amazon S3 does not transition objects that are smaller than 128 KB because it's not cost effective:
  - From the S3 Standard or S3 Standard-IA storage classes to S3 Intelligent-Tiering.
  - From the S3 Standard storage class to S3 Standard-IA or S3 One Zone-IA.

Minimum days for transition from S3 Standard or S3 Standard-IA to S3 Standard-IA or S3 One Zone-IA

Before you transition objects from the S3 Standard or S3 Standard-IA storage classes to S3 Standard-IA or S3 One Zone-IA, you must store them at least 30 days in the S3 Standard storage class. For example, you cannot create a Lifecycle rule to transition objects to the S3 Standard-IA storage class one day after you create them. Amazon S3 doesn't transition objects within the first 30 days because newer objects are often accessed more frequently or deleted sooner than is suitable for S3 Standard-IA or S3 One Zone-IA storage.

Similarly, if you are transitioning noncurrent objects (in versioned buckets), you can transition only objects that are at least 30 days noncurrent to S3 Standard-IA or S3 One Zone-IA storage.

Minimum 30-Day storage charge for S3 Intelligent-Tiering, S3 Standard-IA, and S3 One Zone-IA

The S3 Intelligent-Tiering, S3 Standard-IA, and S3 One Zone-IA storage classes have a minimum 30-day storage charge. Therefore, you can't specify a single Lifecycle rule for both an S3 Intelligent-Tiering, S3 Standard-IA, or S3 One Zone-IA transition and a S3 Glacier or S3 Glacier Deep Archive transition when the S3 Glacier or S3 Glacier Deep Archive transition occurs less than 30 days after the S3 Intelligent-Tiering, S3 Standard-IA, or S3 One Zone-IA transition.

The same 30-day minimum applies when you specify a transition from S3 Standard-IA storage to S3 One Zone-IA or S3 Intelligent-Tiering storage. You can specify two rules to accomplish this, but you pay minimum storage charges. For more information about cost considerations, see Amazon S3 pricing.

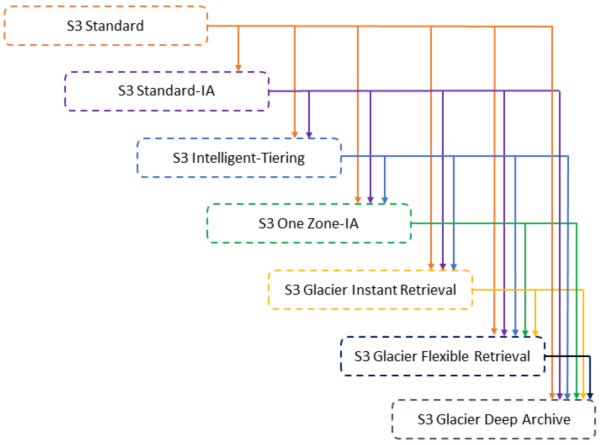
via - https://docs.aws.amazon.com/AmazonS3/latest/dev/lifecycle-transition-general-considerations.html

Supported Amazon S3 lifecycle transitions:

Supported transitions and related constraints

In an S3 Lifecycle configuration, you can define rules to transition objects from one storage class to another to save on storage costs. When you don't know the access patterns of your objects, or if your access patterns are changing over time, you can transition the objects to the S3 Intelligent-Tiering storage class for automatic cost savings. For information about storage classes, see Using Amazon S3 storage classes.

Amazon S3 supports a waterfall model for transitioning between storage classes, as shown in the following diagram.



via - https://docs.aws.amazon.com/AmazonS3/latest/dev/lifecycle-transition-general-considerations.html

Incorrect options:

Configure a lifecycle policy to transition the objects to Amazon S3 Standard-Infrequent Access (S3 Standard-IA) after 7 days

Configure a lifecycle policy to transition the objects to Amazon S3 One Zone-Infrequent Access (S3 One Zone-IA) after 7 days

As mentioned earlier, the minimum storage duration is 30 days before you can transition objects from Amazon S3 Standard to Amazon S3 One Zone-IA or Amazon S3 Standard-IA, so both these options are added as distractors.

Configure a lifecycle policy to transition the objects to Amazon S3 Standard-Infrequent Access (S3 Standard-IA) after 30 days - Amazon S3 Standard-IA is for data that is accessed less frequently, but requires rapid access when needed. S3 Standard-IA offers the high durability, high throughput, and low latency of Amazon S3 Standard, with a low per GB storage price and per GB retrieval fee. This combination of low cost and high performance makes Amazon S3 Standard-IA ideal for long-term storage, backups, and as a data store for disaster recovery files. But, it costs more than Amazon S3 One Zone-IA because of the redundant storage across Availability Zones (AZs). As the data is re-creatable, so you don't need to incur this additional cost.

References:

https://aws.amazon.com/s3/storage-classes/

https://docs.aws.amazon.com/AmazonS3/latest/dev/lifecycle-transition-general-considerations.html

Domain

Design Cost-Optimized Architectures

Question 25 **Correct**

A company wants to store business-critical data on Amazon Elastic Block Store (Amazon EBS) volumes which provide persistent storage independent of Amazon EC2 instances. During a test run, the development team found that on terminating an Amazon EC2 instance, the attached Amazon EBS volume was also lost, which was contrary to their assumptions.

As a solutions architect, could you explain this issue?

The Amazon EBS volumes were not backed up on Amazon S3 storage, resulting in the loss of volume

The Amazon EBS volumes were not backed up on Amazon EFS file system storage, resulting in the loss of volume

Your answer is correct

The Amazon EBS volume was configured as the root volume of Amazon EC2 instance. On termination of the instance, the default behavior is to also terminate the attached root volume

On termination of an Amazon EC2 instance, all the attached Amazon EBS volumes are always terminated

Overall explanation

Correct option:

The Amazon EBS volume was configured as the root volume of Amazon EC2 instance. On termination of the instance, the default behavior is to also terminate the attached root volume

Amazon Elastic Block Store (EBS) is an easy to use, high-performance block storage service designed for use with Amazon Elastic Compute Cloud (EC2) for both throughput and transaction-intensive workloads at any scale.

When you launch an instance, the root device volume contains the image used to boot the instance. You can choose between AMIs backed by Amazon EC2 instance store and AMIs backed by Amazon EBS.

By default, the root volume for an AMI backed by Amazon EBS is deleted when the instance terminates. You can change the default behavior to ensure that the volume persists after the instance terminates. Non-root EBS volumes remain available even after you terminate an instance to which the volumes were attached. Therefore, this option is correct.

Incorrect options:

The Amazon EBS volumes were not backed up on Amazon S3 storage, resulting in the loss of volume

The Amazon EBS volumes were not backed up on Amazon EFS file system storage, resulting in the loss of volume

Amazon EBS volumes do not need to back up the data on Amazon S3 or Amazon EFS filesystem. Both these options are added as distractors.

On termination of an Amazon EC2 instance, all the attached Amazon EBS volumes are always terminated - As mentioned earlier, non-root Amazon EBS volumes remain available even after you terminate an instance to which the volumes were attached. Hence this option is incorrect.

Reference:

https://docs.aws.amazon.com/AWSEC2/latest/UserGuide/RootDeviceStorage.html

Domain

Design High-Performing Architectures

Question 26 **Incorrect**

For security purposes, a development team has decided to deploy the Amazon EC2 instances in a private subnet. The team plans to use VPC endpoints so that the instances can access some AWS

services securely. The members of the team would like to know about the two AWS services that support Gateway Endpoints.

As a solutions architect, which of the following services would you suggest for this requirement? (Select two)

Correct selection

Amazon DynamoDB

Amazon Simple Notification Service (Amazon SNS)

Amazon Kinesis

Your selection is correct

Amazon S3

Your selection is incorrect

Amazon Simple Queue Service (Amazon SQS)

Overall explanation

Correct options:

Amazon S3

Amazon DynamoDB

A VPC endpoint enables you to privately connect your VPC to supported AWS services and VPC endpoint services powered by AWS PrivateLink without requiring an internet gateway, NAT device, VPN connection, or AWS Direct Connect connection. Instances in your VPC do not require public IP addresses to communicate with resources in the service. Traffic between your VPC and the other service does not leave the Amazon network.

Endpoints are virtual devices. They are horizontally scaled, redundant, and highly available VPC components. They allow communication between instances in your VPC and services without imposing availability risks or bandwidth constraints on your network traffic.

There are two types of VPC endpoints: Interface Endpoints and Gateway Endpoints. An Interface Endpoint is an Elastic Network Interface with a private IP address from the IP address range of your subnet that serves as an entry point for traffic destined to a supported service.

A Gateway Endpoint is a gateway that you specify as a target for a route in your route table for traffic destined to a supported AWS service. The following AWS services are supported: Amazon S3 and Amazon DynamoDB.

You can use two types of VPC endpoints to access Amazon S3: gateway endpoints and interface endpoints. A gateway endpoint is a gateway that you specify in your route table to access Amazon S3 from your VPC over the AWS network. Interface endpoints extend the functionality of gateway endpoints by using private IP addresses to route requests to Amazon S3 from within your VPC, on premises, or from a VPC in another AWS Region using VPC peering or AWS Transit Gateway.

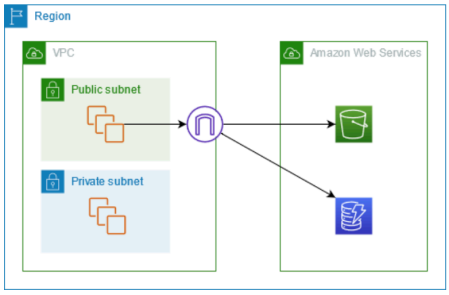
You must remember that these two services use a VPC gateway endpoint. The rest of the AWS services use VPC interface endpoints.

Gateway VPC endpoints:

You can access Amazon S3 and DynamoDB through their public service endpoints or through gateway endpoints. This overview compares these methods.

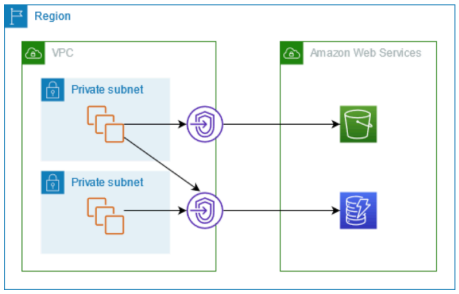
Access through an internet gateway

The following diagram shows how instances access Amazon S3 and DynamoDB through their public service endpoints. Traffic to Amazon S3 or DynamoDB from an instance in a public subnet is routed to the internet gateway for the VPC and then to the service. Instances in a private subnet can't send traffic to Amazon S3 or DynamoDB, because by definition private subnets do not have routes to an internet gateway. To enable instances in the private subnet to send traffic to Amazon S3 or DynamoDB, you would add a NAT device to the public subnet and route traffic in the private subnet to the NAT device. While traffic to Amazon S3 or DynamoDB traverses the internet gateway, it does not leave the AWS network.



Access through a gateway endpoint

The following diagram shows how instances access Amazon S3 and DynamoDB through a gateway endpoint. Traffic from your VPC to Amazon S3 or DynamoDB is routed to the gateway endpoint. Each subnet route table must have a route that sends traffic destined for the service to the gateway endpoint using the prefix list for the service. For more information, see AWS-managed prefix lists in the *Amazon VPC User Guide*.



via - https://docs.aws.amazon.com/vpc/latest/userguide/vpce-gateway.html

Incorrect options:

Amazon Simple Queue Service (Amazon SQS)

Amazon Simple Notification Service (Amazon SNS)

Amazon Kinesis

As mentioned in the description above, these three options use interface endpoints, so these are incorrect.

Reference:

https://docs.aws.amazon.com/vpc/latest/userguide/vpc-endpoints.html

Domain

Design Secure Architectures

Question 27 Incorrect

The business analytics team at a company has been running ad-hoc queries on Oracle and PostgreSQL services on Amazon RDS to prepare daily reports for senior management. To facilitate the business analytics reporting, the engineering team now wants to continuously replicate this data and

consolidate these databases into a petabyte-scale data warehouse by streaming data to Amazon Redshift.

As a solutions architect, which of the following would you recommend as the MOST resource-efficient solution that requires the LEAST amount of development time without the need to manage the underlying infrastructure?

Your answer is incorrect

Use AWS EMR to replicate the data from the databases into Amazon Redshift

Use AWS Glue to replicate the data from the databases into Amazon Redshift

Use Amazon Kinesis Data Streams to replicate the data from the databases into Amazon Redshift

Correct answer

Use AWS Database Migration Service (AWS DMS) to replicate the data from the databases into Amazon Redshift

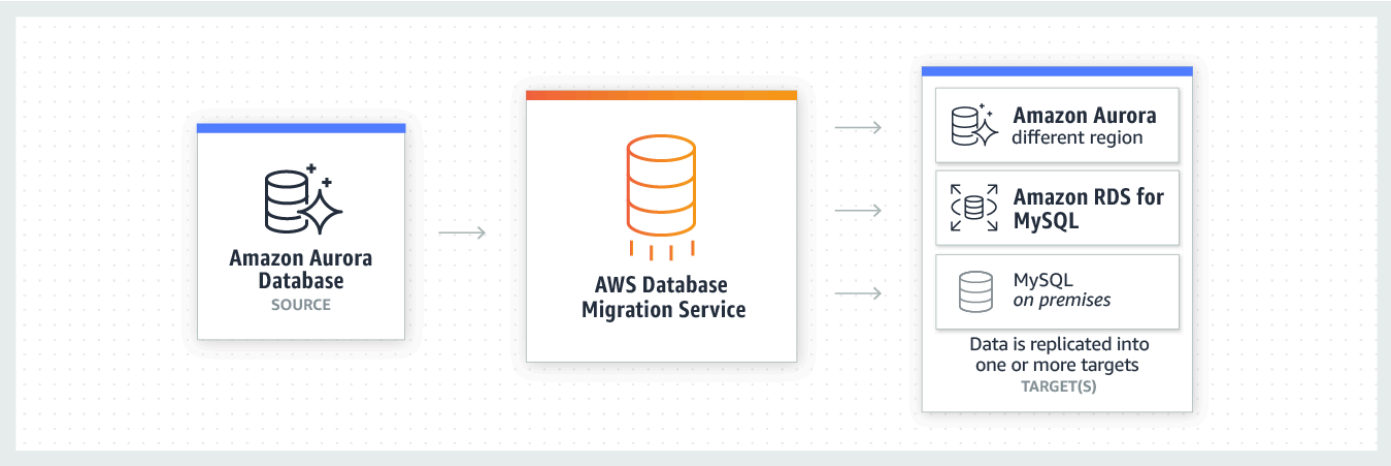
Overall explanation

Correct option:

Use AWS Database Migration Service (AWS DMS) to replicate the data from the databases into Amazon Redshift

AWS Database Migration Service helps you migrate databases to AWS quickly and securely. The source database remains fully operational during the migration, minimizing downtime to applications that rely on the database. With AWS Database Migration Service, you can continuously replicate your data with high availability and consolidate databases into a petabyte-scale data warehouse by streaming data to Amazon Redshift and Amazon S3.

Continuous Data Replication



via - https://aws.amazon.com/dms/

You can migrate data to Amazon Redshift databases using AWS Database Migration Service. Amazon Redshift is a fully managed, petabyte-scale data warehouse service in the cloud. With an Amazon Redshift database as a target, you can migrate data from all of the other supported source databases.

The Amazon Redshift cluster must be in the same AWS account and the same AWS Region as the replication instance. During a database migration to Amazon Redshift, AWS DMS first moves data to an Amazon S3 bucket. When the files reside in an Amazon S3 bucket, AWS DMS then transfers them to the proper tables in the Amazon Redshift data warehouse. AWS DMS creates the S3 bucket in the same AWS Region as the Amazon Redshift database. The AWS DMS replication instance must be located in that same region.

Incorrect options:

Use AWS Glue to replicate the data from the databases into Amazon Redshift - AWS Glue is a fully managed extract, transform, and load (ETL) service that makes it easy for customers to prepare and load their data for analytics. AWS Glue job is meant to be used for batch ETL data processing.

Using AWS Glue involves significant development efforts to write custom migration scripts to copy the database data into Redshift.

Use AWS EMR to replicate the data from the databases into Amazon Redshift - Amazon EMR is the industry-leading cloud big data platform for processing vast amounts of data using open source tools such as Apache Spark, Apache Hive, Apache HBase, Apache Flink, Apache Hudi, and Presto. With EMR you can run Petabyte-scale analysis at less than half of the cost of traditional on-premises solutions and over 3x faster than standard Apache Spark. For short-running jobs, you can spin up and spin down clusters and pay per second for the instances used. For long-

running workloads, you can create highly available clusters that automatically scale to meet demand. Amazon EMR uses Hadoop, an open-source framework, to distribute your data and processing across a resizable cluster of Amazon EC2 instances.

Using EMR involves significant infrastructure management efforts to set up and maintain the EMR cluster. Additionally this option involves a major development effort to write custom migration jobs to copy the database data into Redshift.

Use Amazon Kinesis Data Streams to replicate the data from the databases into Amazon Redshift - Amazon Kinesis Data Streams (KDS) is a massively scalable and durable real-time data streaming service. KDS can continuously capture gigabytes of data per second from hundreds of thousands of sources such as website clickstreams, database event streams, financial transactions, social media feeds, IT logs, and location-tracking events.

However, the user is expected to manually provision an appropriate number of shards to process the expected volume of the incoming data stream. The throughput of an Amazon Kinesis data stream is designed to scale without limits via increasing the number of shards within a data stream. Therefore Kinesis Data Streams is not the right fit for this use-case.

References:

https://docs.aws.amazon.com/dms/latest/userguide/CHAP_Target.Redshift.html

https://aws.amazon.com/dms/

Domain

Design Resilient Architectures

Question 28 **Correct**

A cybersecurity company uses a fleet of Amazon EC2 instances to run a proprietary application. The infrastructure maintenance group at the company wants to be notified via an email whenever the CPU utilization for any of the Amazon EC2 instances breaches a certain threshold.

Which of the following services would you use for building a solution with the LEAST amount of development effort? (Select two)

AWS Step Functions

Your selection is correct

Amazon Simple Notification Service (Amazon SNS)

Amazon Simple Queue Service (Amazon SQS)

Your selection is correct

Amazon CloudWatch

AWS Lambda

Overall explanation

Correct options:

Amazon Simple Notification Service (Amazon SNS)

Amazon Simple Notification Service (Amazon SNS) is a highly available, durable, secure, fully managed pub/sub messaging service that enables you to decouple microservices, distributed systems, and serverless applications. Amazon SNS provides topics for high-throughput, push-based, many-to-many messaging.

Amazon CloudWatch

Amazon CloudWatch is a monitoring and observability service built for DevOps engineers, developers, site reliability engineers (SREs), and IT managers. Amazon CloudWatch provides you with data and actionable insights to monitor your applications, respond to system-wide performance changes, optimize resource utilization, and get a unified view of operational health. Amazon CloudWatch allows you to monitor AWS cloud resources and the applications you run on AWS.

You can use Amazon CloudWatch Alarms to send an email via Amazon SNS whenever any of the Amazon EC2 instances breaches a certain threshold. Hence both these options are correct.

Incorrect options:

AWS Lambda - With AWS Lambda, you can run code without provisioning or managing servers. You pay only for the compute time that you consume—there's no charge when your code isn't running. You can run code for virtually any type of application or backend service—all with zero administration. You cannot use AWS Lambda to monitor CPU utilization of Amazon EC2 instances or send notification emails, hence this option is incorrect.

Amazon Simple Queue Service (Amazon SQS) - Amazon SQS Standard offers a reliable, highly scalable hosted queue for storing messages as they travel between computers. Amazon SQS lets you easily move data between distributed application components and helps you build applications in which messages are processed independently (with message-level ack/fail semantics), such as automated workflows. You cannot use Amazon SQS to monitor CPU utilization of Amazon EC2 instances or send notification emails, hence this option is incorrect.

AWS Step Functions - AWS Step Functions lets you coordinate multiple AWS services into serverless workflows so you can build and update apps quickly. Using Step Functions, you can design and run workflows that stitch together services, such as AWS Lambda, AWS Fargate, and Amazon SageMaker, into feature-rich applications. You cannot use Step Functions to monitor CPU utilization of Amazon EC2 instances or send notification emails, hence this option is incorrect.

References:

https://aws.amazon.com/cloudwatch/faqs/

https://aws.amazon.com/sns/

Domain

Design Resilient Architectures

Question 29 **Incorrect**

A financial services firm uses a high-frequency trading system and wants to write the log files into Amazon S3. The system will also read these log files in parallel on a near real-time basis. The engineering team wants to address any data discrepancies that might arise when the trading system overwrites an existing log file and then tries to read that specific log file.

Which of the following options BEST describes the capabilities of Amazon S3 relevant to this scenario?

A process replaces an existing object and immediately tries to read it. Until the change is fully propagated, Amazon S3 does not return any data

Correct answer

A process replaces an existing object and immediately tries to read it. Amazon S3 always returns the latest version of the object

A process replaces an existing object and immediately tries to read it. Until the change is fully propagated, Amazon S3 might return the new data

Your answer is incorrect

A process replaces an existing object and immediately tries to read it. Until the change is fully propagated, Amazon S3 might return the previous data

Overall explanation

Correct option:

A process replaces an existing object and immediately tries to read it. Amazon S3 always returns the latest version of the object

Amazon S3 delivers strong read-after-write consistency automatically, without changes to performance or availability, without sacrificing regional isolation for applications, and at no additional cost.

After a successful write of a new object or an overwrite of an existing object, any subsequent read request immediately receives the latest version of the object. Amazon S3 also provides strong consistency for list operations, so after a write, you can immediately perform a listing of the objects in a bucket with any changes reflected.

Strong read-after-write consistency helps when you need to immediately read an object after a write. For example, strong read-after-write consistency when you often read and list immediately after writing objects.

To summarize, all Amazon S3 GET, PUT, and LIST operations, as well as operations that change object tags, ACLs, or metadata, are strongly consistent. What you write is what you will read, and the results of a LIST will be an accurate reflection of what's in the bucket.

Incorrect options:

A process replaces an existing object and immediately tries to read it. Until the change is fully propagated, Amazon S3 might return the previous data

A process replaces an existing object and immediately tries to read it. Until the change is fully propagated, Amazon S3 does not return any data

A process replaces an existing object and immediately tries to read it. Until the change is fully propagated, Amazon S3 might return the new data

These three options contradict the earlier details provided in the explanation.

References:

https://docs.aws.amazon.com/AmazonS3/latest/dev/Introduction.html#ConsistencyModel

https://aws.amazon.com/s3/faqs/

Domain

Design Resilient Architectures

Question 30 **Correct**

The engineering team at an e-commerce company has been tasked with migrating to a serverless architecture. The team wants to focus on the key points of consideration when using AWS Lambda as a backbone for this architecture.

As a Solutions Architect, which of the following options would you identify as correct for the given requirement? (Select three)

Your selection is correct

By default, AWS Lambda functions always operate from an AWS-owned VPC and hence have access to any public internet address or public AWS APIs. Once an AWS Lambda function is VPC-enabled, it will need a route through a Network Address Translation gateway (NAT gateway) in a public subnet to access public resources

The bigger your deployment package, the slower your AWS Lambda function will cold-start. Hence, AWS suggests packaging dependencies as a separate package from the actual AWS Lambda package

Your selection is correct

Since AWS Lambda functions can scale extremely quickly, it's a good idea to deploy a Amazon CloudWatch Alarm that notifies your team when function metrics such as `ConcurrentExecutions` or `Invocations exceeds the expected threshold`

AWS Lambda allocates compute power in proportion to the memory you allocate to your function. AWS, thus recommends to over provision your function time out settings for the proper performance of AWS Lambda functions

Serverless architecture and containers complement each other but you cannot package and deploy AWS Lambda functions as container images

Your selection is correct

If you intend to reuse code in more than one AWS Lambda function, you should consider creating an AWS Lambda Layer for the reusable code

Overall explanation

Correct options:

By default, AWS Lambda functions always operate from an AWS-owned VPC and hence have access to any public internet address or public AWS APIs. Once an AWS Lambda function is VPC-enabled, it will need a route through a Network Address Translation gateway (NAT gateway) in a public subnet to access public resources

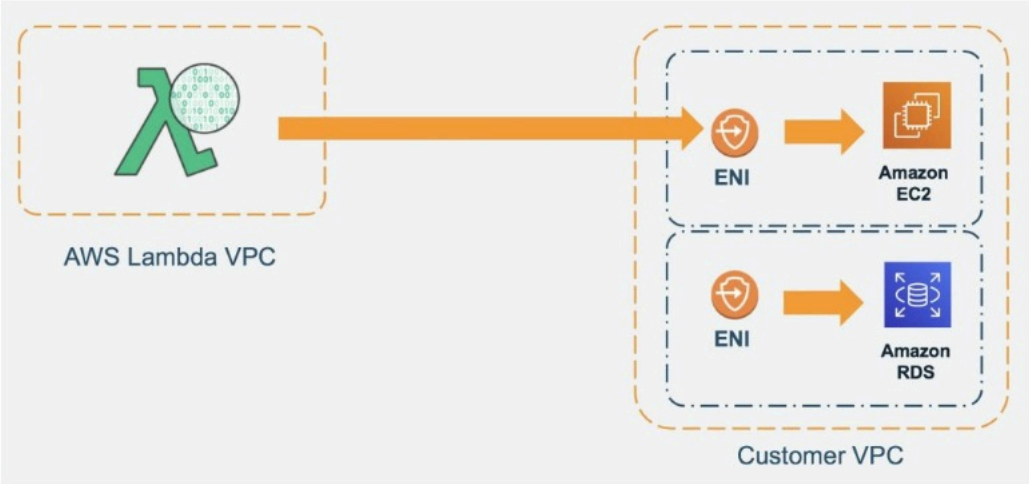
AWS Lambda functions always operate from an AWS-owned VPC. By default, your function has the full ability to make network requests to any public internet address — this includes access to any of the public AWS APIs. For example, your function can interact with AWS DynamoDB APIs to `PutItem` or Query for records. You should only enable your functions for VPC access when you need to interact with a private resource located in a private subnet. An Amazon RDS instance is a good example.

Once your function is VPC-enabled, all network traffic from your function is subject to the routing rules of your VPC/Subnet. If your function needs to interact with a public resource, you will need a route through a NAT gateway in a public subnet.

When to VPC-Enable an AWS Lambda Function:

Tip #1: When to VPC-Enable a Lambda Function

Lambda functions always operate from an AWS-owned VPC. By default, your function has full ability to make network requests to any public internet address — this includes access to any of the public AWS APIs. For example, your function can interact with AWS DynamoDB APIs to PutItem or Query for records. You should only enable your functions for VPC access when you need to interact with a private resource located in a private subnet. An RDS instance is a good example.



Once your function is VPC-enabled, all network traffic from your function is subject to the routing rules of your VPC/Subnet. If your function needs to interact with a public resource, you will need a route through a NAT gateway in a public subnet.

via - https://aws.amazon.com/blogs/architecture/best-practices-for-developing-on-aws-lambda/

Since AWS Lambda functions can scale extremely quickly, it's a good idea to deploy a Amazon CloudWatch Alarm that notifies your team when function metrics such as `ConcurrentExecutions` or `Invocations exceeds the expected threshold`

Since AWS Lambda functions can scale extremely quickly, this means you should have controls in place to notify you when you have a spike in concurrency. A good idea is to deploy an Amazon CloudWatch Alarm that notifies your team when function metrics such as `ConcurrentExecutions` or `Invocations exceeds your threshold`. You should create an AWS Budget so you can monitor costs on a daily basis.

If you intend to reuse code in more than one AWS Lambda function, you should consider creating an AWS Lambda Layer for the reusable code

You can configure your AWS Lambda function to pull in additional code and content in the form of layers. A layer is a ZIP archive that contains libraries, a custom runtime, or other dependencies. With layers, you can use libraries in your function without needing to include them in your deployment package. Layers let you keep your deployment package small, which makes development easier. A function can use up to 5 layers at a time.

You can create layers, or use layers published by AWS and other AWS customers. Layers support resource-based policies for granting layer usage permissions to specific AWS accounts, AWS Organizations, or all accounts. The total unzipped size of the function and all layers can't exceed the unzipped deployment package size limit of 250 megabytes.

Incorrect options:

AWS Lambda allocates compute power in proportion to the memory you allocate to your function. AWS, thus recommends to over provision your function time out settings for the

proper performance of AWS Lambda functions - AWS Lambda allocates compute power in proportion to the memory you allocate to your function. This means you can over-provision memory to run your functions faster and potentially reduce your costs. However, AWS recommends that you should not over provision your function time out settings. Always understand your code performance and set a function time out accordingly. Overprovisioning function timeout often results in Lambda functions running longer than expected and unexpected costs.

The bigger your deployment package, the slower your AWS Lambda function will cold-start. Hence, AWS suggests packaging dependencies as a separate package from the actual AWS Lambda package - This statement is incorrect and acts as a distractor. All the dependencies are also packaged into the single Lambda deployment package.

Serverless architecture and containers complement each other but you cannot package and deploy AWS Lambda functions as container images - This statement is incorrect. You can now package and deploy AWS Lambda functions as container images.

References:

https://aws.amazon.com/blogs/architecture/best-practices-for-developing-on-aws-lambda/

https://docs.aws.amazon.com/lambda/latest/dg/configuration-layers.html

https://aws.amazon.com/blogs/aws/new-for-aws-lambda-container-image-support/

Domain

Design High-Performing Architectures

Question 31 **Correct**

A financial services company recently launched an initiative to improve the security of its AWS resources and it had enabled AWS Shield Advanced across multiple AWS accounts owned by the company. Upon analysis, the company has found that the costs incurred are much higher than expected.

Which of the following would you attribute as the underlying reason for the unexpectedly high costs for AWS Shield Advanced service?

Savings Plans has not been enabled for the AWS Shield Advanced service across all the AWS accounts

AWS Shield Advanced is being used for custom servers, that are not part of AWS Cloud, thereby resulting in increased costs

AWS Shield Advanced also covers AWS Shield Standard plan, thereby resulting in increased costs

Your answer is correct

Consolidated billing has not been enabled. All the AWS accounts should fall under a single consolidated billing for the monthly fee to be charged only once

Overall explanation

Correct option:

Consolidated billing has not been enabled. All the AWS accounts should fall under a single consolidated billing for the monthly fee to be charged only once

If your organization has multiple AWS accounts, then you can subscribe multiple AWS Accounts to AWS Shield Advanced by individually enabling it on each account using the AWS Management Console or API. You will pay the monthly fee once as long as the AWS accounts are all under a single consolidated billing, and you own all the AWS accounts and resources in those accounts.

Incorrect options:

AWS Shield Advanced is being used for custom servers, that are not part of AWS Cloud, thereby resulting in increased costs - AWS Shield Advanced does offer protection to resources outside of AWS. This should not cause unexpected spike in billing costs.

AWS Shield Advanced also covers AWS Shield Standard plan, thereby resulting in increased costs - AWS Shield Standard is automatically enabled for all AWS customers at no additional cost. AWS Shield Advanced is an optional paid service.

Savings Plans has not been enabled for the AWS Shield Advanced service across all the AWS accounts - This option has been added as a distractor. Savings Plans is a flexible pricing model that offers low prices on Amazon EC2 instances, AWS Lambda, and AWS Fargate usage, in exchange for a commitment to a consistent amount of usage (measured in $/hour) for a 1 or 3 year term. Savings Plans is not applicable for the AWS Shield Advanced service.

References:

https://aws.amazon.com/shield/faqs/

https://aws.amazon.com/savingsplans/faq/

Domain

Design Cost-Optimized Architectures

Question 32 Incorrect

A retail company uses Amazon Elastic Compute Cloud (Amazon EC2) instances, Amazon API Gateway, Amazon RDS, Elastic Load Balancer and Amazon CloudFront services. To improve the security of these services, the Risk Advisory group has suggested a feasibility check for using the Amazon GuardDuty service.

Which of the following would you identify as data sources supported by Amazon GuardDuty?

Your answer is incorrect

Amazon CloudFront logs, Amazon API Gateway logs, AWS CloudTrail events

VPC Flow Logs, Amazon API Gateway logs, Amazon S3 access logs

Correct answer

VPC Flow Logs, Domain Name System (DNS) logs, AWS CloudTrail events

Elastic Load Balancing logs, Domain Name System (DNS) logs, AWS CloudTrail events

Overall explanation

Correct option:

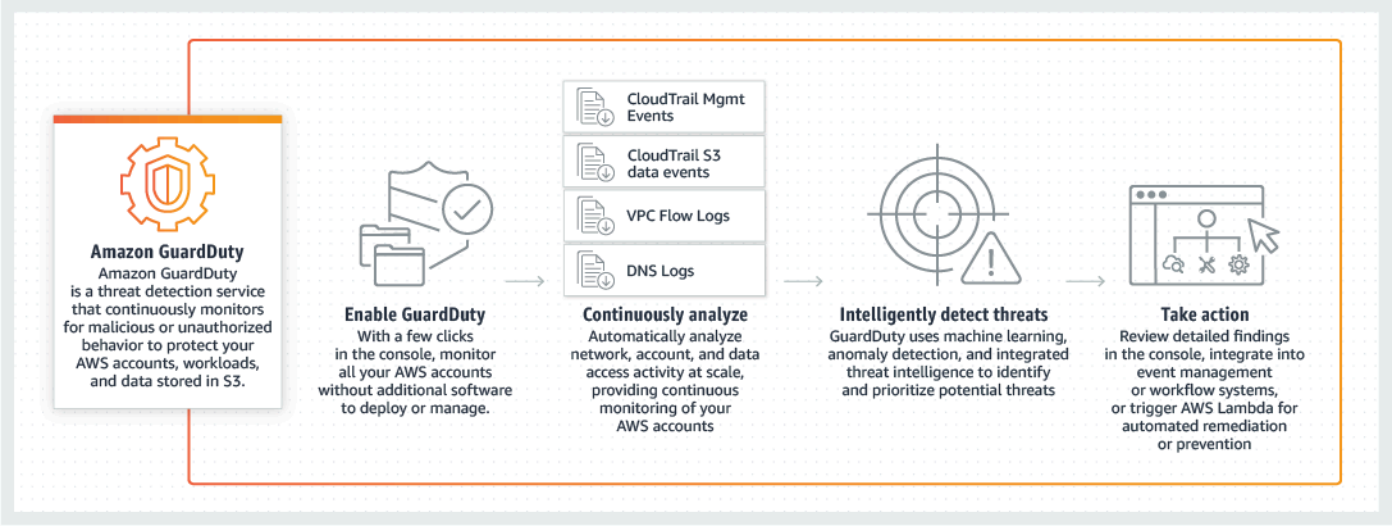
VPC Flow Logs, Domain Name System (DNS) logs, AWS CloudTrail events

Amazon GuardDuty is a threat detection service that continuously monitors for malicious activity and unauthorized behavior to protect your AWS accounts, workloads, and data stored in Amazon S3. With the cloud, the collection and aggregation of account and network activities is simplified, but it can be time-consuming for security teams to continuously analyze event log data for potential threats. With GuardDuty, you now have an intelligent and cost-effective option for continuous threat detection in AWS. The service uses machine learning, anomaly detection, and integrated threat intelligence to identify and prioritize potential threats.

Amazon GuardDuty analyzes tens of billions of events across multiple AWS data sources, such as AWS CloudTrail events, Amazon VPC Flow Logs, and DNS logs.

With a few clicks in the AWS Management Console, GuardDuty can be enabled with no software or hardware to deploy or maintain. By integrating with Amazon EventBridge Events, GuardDuty alerts are actionable, easy to aggregate across multiple accounts, and straightforward to push into existing event management and workflow systems.

How Amazon GuardDuty works:



via - https://aws.amazon.com/guardduty/

Incorrect options:

VPC Flow Logs, Amazon API Gateway logs, Amazon S3 access logs

Elastic Load Balancing logs, Domain Name System (DNS) logs, AWS CloudTrail events

Amazon CloudFront logs, Amazon API Gateway logs, AWS CloudTrail events

These three options contradict the explanation provided above, so these options are incorrect.

Reference:

https://aws.amazon.com/guardduty/

Domain

Design Secure Architectures

Question 33 Correct

The infrastructure team at a company maintains 5 different VPCs (let's call these VPCs A, B, C, D, E) for resource isolation. Due to the changed organizational structure, the team wants to interconnect all VPCs together. To facilitate this, the team has set up VPC peering connection between VPC A and all other VPCs in a hub and spoke model with VPC A at the center. However, the team has still failed to establish connectivity between all VPCs.

As a solutions architect, which of the following would you recommend as the MOST resource-efficient and scalable solution?

Your answer is correct

Use AWS transit gateway to interconnect the VPCs

Use a VPC endpoint to interconnect the VPCs

Use an internet gateway to interconnect the VPCs

Establish VPC peering connections between all VPCs

Overall explanation

Correct option:

Use AWS transit gateway to interconnect the VPCs

An AWS transit gateway is a network transit hub that you can use to interconnect your virtual private clouds (VPC) and on-premises networks.

AWS Transit Gateway Overview:

What is a transit gateway?

PDF

A *transit gateway* is a network transit hub that you can use to interconnect your virtual private clouds (VPC) and on-premises networks.

For more information, see AWS Transit Gateway.

Transit gateway concepts

The following are the key concepts for transit gateways:

- **attachment** — You can attach a VPC, an AWS Direct Connect gateway, a peering connection with another transit gateway, or a VPN connection to a transit gateway.
- **transit gateway Maximum Transmission Unit (MTU)** — The maximum transmission unit (MTU) of a network connection is the size, in bytes, of the largest permissible packet that can be passed over the connection. The larger the MTU of a connection, the more data can be passed in a single packet. A transit gateway supports an MTU of 8500 bytes for traffic between VPCs, Direct Connect and peering attachments. Traffic over VPN connections can have an MTU of 1500 bytes.
- **transit gateway route table** — A transit gateway has a default route table and can optionally have additional route tables. A route table includes dynamic and static routes that decide the next hop based on the destination IP address of the packet. The target of these routes could be a VPC or a VPN connection. By default, transit gateway attachments are associated with the default transit gateway route table.
- **associations** — Each attachment is associated with exactly one route table. Each route table can be associated with zero to many attachments.
- **route propagation** — A VPC or VPN connection can dynamically propagate routes to a transit gateway route table. With a VPC, you must create static routes to send traffic to the transit gateway. With a VPN connection, routes are propagated from the transit gateway to your on-premises router using Border Gateway Protocol (BGP). With a peering attachment, you must create a static route in the transit gateway route table to point to the peering attachment.

via - https://docs.aws.amazon.com/vpc/latest/tgw/what-is-transit-gateway.html

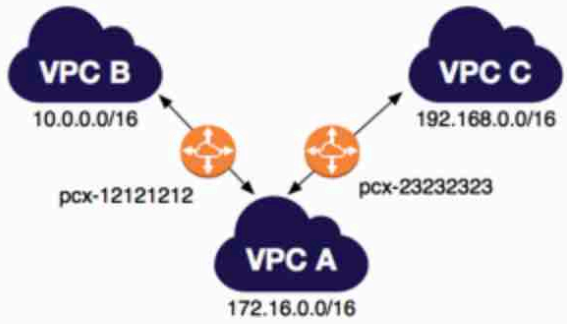
A VPC peering connection is a networking connection between two VPCs that enables you to route traffic between them using private IPv4 addresses or IPv6 addresses. Transitive Peering does not work for VPC peering connections. So, if you have a VPC peering connection between VPC A and VPC B (pcx-aaaabbbb), and between VPC A and VPC C (pcx-aaaacccc). Then, there is no VPC peering connection between VPC B and VPC C. Instead of using VPC peering, you can use an AWS Transit Gateway that acts as a network transit hub, to interconnect your VPCs or connect your VPCs with on-premises networks. Therefore this is the correct option.

VPC Peering Connections Overview:

Multiple VPC peering connections

A VPC peering connection is a one to one relationship between two VPCs. You can create multiple VPC peering connections for each VPC that you own, but transitive peering relationships are not supported. You do not have any peering relationship with VPCs that your VPC is not directly peered with.

The following diagram is an example of one VPC peered to two different VPCs. There are two VPC peering connections: VPC A is peered with both VPC B and VPC C. VPC B and VPC C are not peered, and you cannot use VPC A as a transit point for peering between VPC B and VPC C. If you want to enable routing of traffic between VPC B and VPC C, you must create a unique VPC peering connection between them.



via - https://docs.aws.amazon.com/vpc/latest/peering/vpc-peering-basics.html

Incorrect options:

Use an internet gateway to interconnect the VPCs - An internet gateway is a horizontally scaled, redundant, and highly available VPC component that allows communication between instances in your VPC and the internet. It, therefore, imposes no availability risks or bandwidth constraints on your network traffic. You cannot use an internet gateway to interconnect your VPCs and on-premises networks, hence this option is incorrect.

Use a VPC endpoint to interconnect the VPCs - A VPC endpoint enables you to privately connect your VPC to supported AWS services and VPC endpoint services powered by AWS PrivateLink without requiring an internet gateway, NAT device, VPN connection, or AWS Direct Connect connection. You cannot use a VPC endpoint to interconnect your VPCs and on-premises networks, hence this option is incorrect.

Establish VPC peering connections between all VPCs - Establishing VPC peering between all VPCs is an inelegant and clumsy way to establish connectivity between all VPCs. Instead, you should use a Transit Gateway that acts as a network transit hub to interconnect your VPCs and on-premises networks.

References:

https://docs.aws.amazon.com/vpc/latest/peering/what-is-vpc-peering.html

https://docs.aws.amazon.com/vpc/latest/tgw/what-is-transit-gateway.html

Domain

Design Secure Architectures

Question 34 **Correct**

An IT company provides Amazon Simple Storage Service (Amazon S3) bucket access to specific users within the same account for completing project specific work. With changing business requirements, cross-account S3 access requests are also growing every month. The company is looking for a solution that can offer user level as well as account-level access permissions for the data stored in Amazon S3 buckets.

As a Solutions Architect, which of the following would you suggest as the MOST optimized way of controlling access for this use-case?

Use Identity and Access Management (IAM) policies

Use Access Control Lists (ACLs)

Your answer is correct

Use Amazon S3 Bucket Policies

Use Security Groups

Overall explanation

Correct option:

Use Amazon S3 Bucket Policies

Bucket policies in Amazon S3 can be used to add or deny permissions across some or all of the objects within a single bucket. Policies can be attached to users, groups, or Amazon S3 buckets, enabling centralized management of permissions. With bucket policies, you can grant users within your AWS Account or other AWS Accounts access to your Amazon S3 resources.

You can further restrict access to specific resources based on certain conditions. For example, you can restrict access based on request time (Date Condition), whether the request was sent using SSL (Boolean Conditions), a requester's IP address (IP Address Condition), or based on the requester's client application (String Conditions). To identify these conditions, you use policy keys.

Types of access control in Amazon S3:

- **Bucket Policies.** Bucket policies in Amazon S3 can be used to add or deny permissions across some or all of the objects within a single bucket. Policies can be attached to users, groups, or Amazon S3 buckets, enabling centralized management of permissions. With bucket policies, you can grant users within your AWS Account or other AWS Accounts access to your Amazon S3 resources.

Table 3: Types of access control

Type of Access Control	AWS Account Level Control	User Level Control
IAM Policies	No	Yes
ACLs	Yes	No
Bucket Policies	Yes	Yes

You can further restrict access to specific resources based on certain conditions. For example, you can restrict access based on request time (Date Condition), whether the request was sent using SSL (Boolean Conditions), a requester's IP address (IP Address Condition), or based on the requester's client application (String Conditions). To identify these conditions, you use policy keys. For more information about action-specific policy keys available within Amazon S3, see the Amazon Simple Storage Service Developer Guide.

via - https://d1.awsstatic.com/whitepapers/aws-security-whitepaper.pdf

Incorrect options:

Use Identity and Access Management (IAM) policies - AWS IAM enables organizations with many employees to create and manage multiple users under a single AWS account. IAM policies are attached to the users, enabling centralized control of permissions for users under your AWS Account to access buckets or objects. With IAM policies, you can only grant users within your own AWS account permission to access your Amazon S3 resources. So, this is not the right choice for the current requirement.

Use Access Control Lists (ACLs) - Within Amazon S3, you can use ACLs to give read or write access on buckets or objects to groups of users. With ACLs, you can only grant other AWS accounts (not specific users) access to your Amazon S3 resources. So, this is not the right choice for the current requirement.

Use Security Groups - A security group acts as a virtual firewall for Amazon EC2 instances to control incoming and outgoing traffic. Amazon S3 does not support Security Groups, this option just acts as a distractor.

Reference:

https://d1.awsstatic.com/whitepapers/aws-security-whitepaper.pdf

Domain

Design Secure Architectures

Question 35 **Correct**

An IT company wants to review its security best-practices after an incident was reported where a new developer on the team was assigned full access to Amazon DynamoDB. The developer accidentally deleted a couple of tables from the production environment while building out a new feature.

Which is the MOST effective way to address this issue so that such incidents do not recur?

Your answer is correct

Use permissions boundary to control the maximum permissions employees can grant to the IAM principals

Remove full database access for all IAM users in the organization

Only root user should have full database access in the organization

The CTO should review the permissions for each new developer's IAM user so that such incidents don't recur

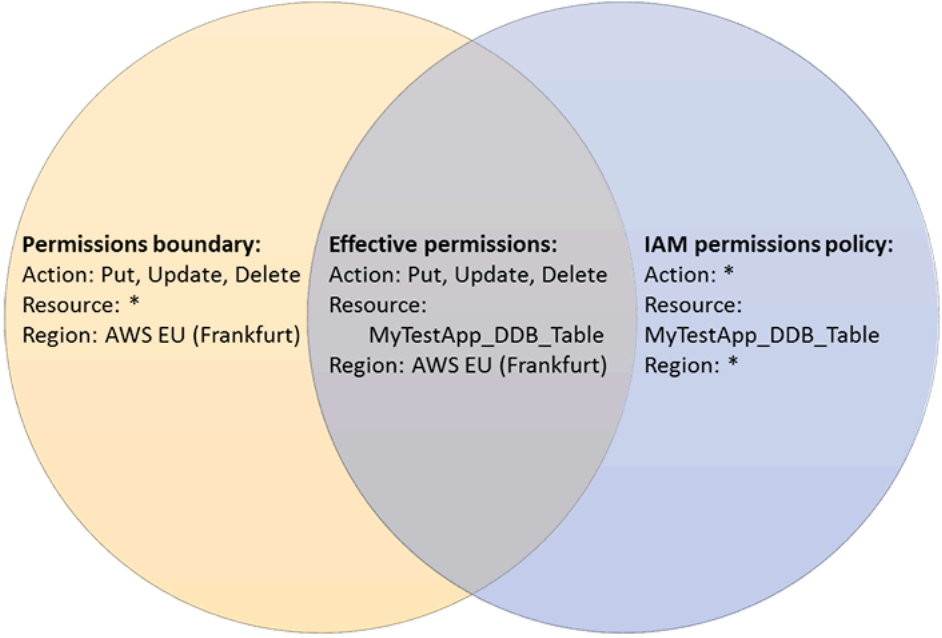
Overall explanation

Correct option:

Use permissions boundary to control the maximum permissions employees can grant to the IAM principals

A permissions boundary can be used to control the maximum permissions employees can grant to the IAM principals (that is, users and roles) that they create and manage. As the IAM administrator, you can define one or more permissions boundaries using managed policies and allow your employee to create a principal with this boundary. The employee can then attach a permissions policy to this principal. However, the effective permissions of the principal are the intersection of the permissions boundary and permissions policy. As a result, the new principal cannot exceed the boundary that you defined. Therefore, using the permissions boundary offers the right solution for this use-case.

Permission Boundary Example:



via

- https://aws.amazon.com/blogs/security/delegate-permission-management-to-developers-using-iam-permissions-boundaries/

Incorrect options:

Remove full database access for all IAM users in the organization - It is not practical to remove full access for all IAM users in the organization because a select set of users need this access for database administration. So this option is not correct.

The CTO should review the permissions for each new developer's IAM user so that such incidents don't recur - Likewise the CTO is not expected to review the permissions for each new developer's IAM user, as this is best done via an automated procedure. This option has been added as a distractor.

Only root user should have full database access in the organization - As a best practice, the root user should not access the AWS account to carry out any administrative procedures. So this option is not correct.

Reference:

https://aws.amazon.com/blogs/security/delegate-permission-management-to-developers-using-iam-permissions-boundaries/

Domain

Design Secure Architectures

Question 36 **Correct**

An IT company has an Access Control Management (ACM) application that uses Amazon RDS for MySQL but is running into performance issues despite using Read Replicas. The company has hired you as a solutions architect to address these performance-related challenges without moving away

from the underlying relational database schema. The company has branch offices across the world, and it needs the solution to work on a global scale.

Which of the following will you recommend as the MOST cost-effective and high-performance solution?

Spin up Amazon EC2 instances in each AWS region, install MySQL databases and migrate the existing data into these new databases

Spin up a Amazon Redshift cluster in each AWS region. Migrate the existing data into Redshift clusters

Use Amazon DynamoDB Global Tables to provide fast, local, read and write performance in each region

Your answer is correct

Use Amazon Aurora Global Database to enable fast local reads with low latency in each region

Overall explanation

Correct option:

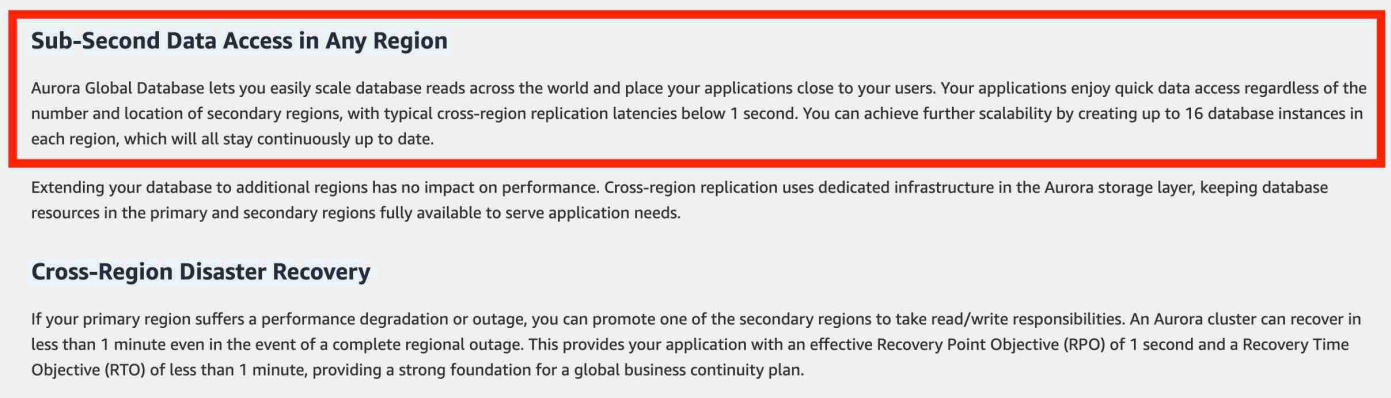
Use Amazon Aurora Global Database to enable fast local reads with low latency in each region

Amazon Aurora is a MySQL and PostgreSQL-compatible relational database built for the cloud, that combines the performance and availability of traditional enterprise databases with the

simplicity and cost-effectiveness of open source databases. Amazon Aurora features a distributed, fault-tolerant, self-healing storage system that auto-scales up to 64TB per database instance. Aurora is not an in-memory database.

Amazon Aurora Global Database is designed for globally distributed applications, allowing a single Amazon Aurora database to span multiple AWS regions. It replicates your data with no impact on database performance, enables fast local reads with low latency in each region, and provides disaster recovery from region-wide outages. Amazon Aurora Global Database is the correct choice for the given use-case.

Amazon Aurora Global Database Features:



Sub-Second Data Access in Any Region

Aurora Global Database lets you easily scale database reads across the world and place your applications close to your users. Your applications enjoy quick data access regardless of the number and location of secondary regions, with typical cross-region replication latencies below 1 second. You can achieve further scalability by creating up to 16 database instances in each region, which will all stay continuously up to date.

Extending your database to additional regions has no impact on performance. Cross-region replication uses dedicated infrastructure in the Aurora storage layer, keeping database resources in the primary and secondary regions fully available to serve application needs.

Cross-Region Disaster Recovery

If your primary region suffers a performance degradation or outage, you can promote one of the secondary regions to take read/write responsibilities. An Aurora cluster can recover in less than 1 minute even in the event of a complete regional outage. This provides your application with an effective Recovery Point Objective (RPO) of 1 second and a Recovery Time Objective (RTO) of less than 1 minute, providing a strong foundation for a global business continuity plan.

via - https://aws.amazon.com/rds/aurora/global-database/

Incorrect options:

Use Amazon DynamoDB Global Tables to provide fast, local, read and write performance in each region - Amazon DynamoDB is a key-value and document database that delivers single-digit millisecond performance at any scale. It's a fully managed, multi-region, multi-master, durable database with built-in security, backup and restore, and in-memory caching for internet-scale applications.

Global Tables builds upon DynamoDB's global footprint to provide you with a fully managed, multi-region, and multi-master database that provides fast, local, read, and write performance for massively scaled, global applications. Global Tables replicates your Amazon DynamoDB tables automatically across your choice of AWS regions. Given that the use-case wants you to continue with the underlying schema of the relational database, DynamoDB is not the right choice as it's a NoSQL database.

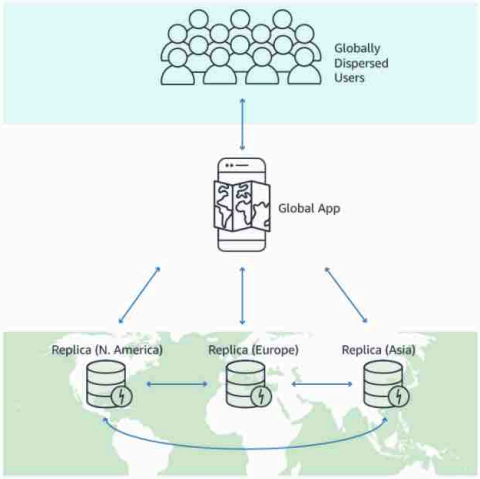
Amazon DynamoDB Global Tables Overview:

Global Tables

Multi-Region, Multi-Master tables for fast local performance for globally distributed apps

Global Tables builds upon DynamoDB's global footprint to provide you with a fully managed, multi-region, and multi-master database that provides fast, local, read and write performance for massively scaled, global applications. Global Tables replicates your Amazon DynamoDB tables automatically across your choice of AWS regions.

Global Tables eliminates the difficult work of replicating data between regions and resolving update conflicts, enabling you to focus on your application's business logic. In addition, Global Tables enables your applications to stay highly available even in the unlikely event of isolation or degradation of an entire region.



Global Tables

You can setup Global Tables with just a few clicks in the AWS Management Console. No application changes are required because Global Tables use existing DynamoDB APIs. There are no upfront costs or commitments for using Global Tables, and you pay only for the resources provisioned. Learn more about setting up Global Tables in the DynamoDB Developer Guide.

via - https://aws.amazon.com/dynamodb/global-tables/

Spin up a Amazon Redshift cluster in each AWS region. Migrate the existing data into Redshift clusters - Amazon Redshift is a fully-managed petabyte-scale cloud-based data warehouse product designed for large scale data set storage and analysis. Amazon Redshift is not suited to be used as a transactional relational database, so this option is not correct.

Spin up Amazon EC2 instances in each AWS region, install MySQL databases and migrate the existing data into these new databases - Setting up Amazon EC2 instances in multiple regions with manually managed MySQL databases represents a maintenance nightmare and is not the correct choice for this use-case.

References:

https://aws.amazon.com/rds/aurora/global-database/

https://aws.amazon.com/dynamodb/global-tables/

Domain

Design High-Performing Architectures

Question 37 **Incorrect**

An e-commerce company has copied 1 petabyte of data from its on-premises data center to an Amazon S3 bucket in the `us-west-1` Region using an AWS Direct Connect link. The company now wants to set up a one-time copy of the data to another Amazon S3 bucket in the `us-east-1` Region. The on-premises data center does not allow the use of AWS Snowball.

As a Solutions Architect, which of the following options can be used to accomplish this goal? (Select two)

Your selection is correct

Set up Amazon S3 batch replication to copy objects across Amazon S3 buckets in another Region using S3 console and then delete the replication configuration

Use AWS Snowball Edge device to copy the data from one Region to another Region

Your selection is incorrect

Copy data from the source Amazon S3 bucket to a target Amazon S3 bucket using the S3 console

Correct selection

Copy data from the source bucket to the destination bucket using the aws S3 sync command

Set up Amazon S3 Transfer Acceleration (Amazon S3TA) to copy objects across Amazon S3 buckets in different Regions using S3 console

Overall explanation

Correct options:

Copy data from the source bucket to the destination bucket using the aws S3 sync command

The aws S3 sync command uses the CopyObject APIs to copy objects between Amazon S3 buckets. The sync command lists the source and target buckets to identify objects that are in the source bucket but that aren't in the target bucket. The command also identifies objects in the source bucket that have different LastModified dates than the objects that are in the target bucket. The sync command on a versioned bucket copies only the current version of the object—previous versions aren't copied. By default, this preserves object metadata, but the access control lists (ACLs) are set to FULL_CONTROL for your AWS account, which removes any additional ACLs. If the operation fails, you can run the sync command again without duplicating previously copied objects.

You can use the command like so:

```
aws s3 sync s3://DOC-EXAMPLE-BUCKET-SOURCE s3://DOC-EXAMPLE-BUCKET-TARGET
```

Set up Amazon S3 batch replication to copy objects across Amazon S3 buckets in another Region using S3 console and then delete the replication configuration

Amazon S3 Batch Replication provides you a way to replicate objects that existed before a replication configuration was in place, objects that have previously been replicated, and objects that have failed replication. This is done through the use of a Batch Operations job.

You should note that batch replication differs from live replication which continuously and automatically replicates new objects across Amazon S3 buckets. You cannot directly use the AWS S3 console to configure cross-Region replication for existing objects. By default, replication only supports copying new Amazon S3 objects after it is enabled using the AWS S3 console. Replication enables automatic, asynchronous copying of objects across Amazon S3 buckets. Buckets that are configured for object replication can be owned by the same AWS account or by different accounts. Object may be replicated to a single destination bucket or multiple destination buckets. Destination buckets can be in different AWS Regions or within the same Region as the source bucket. Once done, you can delete the replication configuration, as it ensures that batch replication is only used for this one-time data copy operation.

If you want to enable live replication for existing objects for your bucket, you must contact AWS Support and raise a support ticket. This is required to ensure that replication is configured correctly.

Incorrect options:

Use AWS Snowball Edge device to copy the data from one Region to another Region - As the given requirement is about copying the data from one AWS Region to another AWS Region, so AWS Snowball Edge cannot be used here. AWS Snowball Edge Storage Optimized is the optimal data transfer choice if you need to securely and quickly transfer terabytes to petabytes of data to AWS. You can use AWS Snowball Edge Storage Optimized if you have a large backlog of data to transfer or if you frequently collect data that needs to be transferred to AWS and your storage is in an area where high-bandwidth internet connections are not available or cost-prohibitive. AWS Snowball Edge can operate in remote locations or harsh operating environments, such as factory floors, oil and gas rigs, mining sites, hospitals, and on moving vehicles.

Copy data from the source Amazon S3 bucket to a target Amazon S3 bucket using the S3 console - AWS S3 console cannot be used to copy 1 petabytes of data from one bucket to another as it's not feasible. You should note that this option is different from using the replication options on the AWS console, since here you are using the copy and paste options provided on the AWS console, which is suggested for small or medium data volume. You should use S3 sync for the requirement of one-time copy of data.

Set up Amazon S3 Transfer Acceleration (Amazon S3TA) to copy objects across Amazon S3 buckets in different Regions using S3 console - Amazon S3 Transfer Acceleration (Amazon S3TA) is a bucket-level feature that enables fast, easy, and secure transfers of files over long distances between your client and an Amazon S3 bucket. You cannot use Transfer Acceleration to copy objects across Amazon S3 buckets in different Regions using Amazon S3 console.

References:

https://aws.amazon.com/premiumsupport/knowledge-center/move-objects-s3-bucket/

https://aws.amazon.com/snowball/faqs/

https://docs.aws.amazon.com/AmazonS3/latest/userguide/replication.html

Domain

Design Resilient Architectures

Question 38 **Incorrect**

A startup's cloud infrastructure consists of a few Amazon EC2 instances, Amazon RDS instances and Amazon S3 storage. A year into their business operations, the startup is incurring costs that seem too high for their business requirements.

Which of the following options represents a valid cost-optimization solution?

Use Amazon S3 Storage class analysis to get recommendations for transitions of objects to Amazon S3 Glacier storage classes to reduce storage costs. You can also automate moving these objects into lower-cost storage tier using Lifecycle Policies

Use AWS Compute Optimizer recommendations to help you choose the optimal Amazon EC2 purchasing options and help reserve your instance capacities at reduced costs

Correct answer

Use AWS Cost Explorer Resource Optimization to get a report of Amazon EC2 instances that are either idle or have low utilization and use AWS Compute Optimizer to look at instance type recommendations

Your answer is incorrect

Use AWS Trusted Advisor checks on Amazon EC2 Reserved Instances to automatically renew reserved instances (RI). AWS Trusted advisor also suggests Amazon RDS idle database instances

Overall explanation

Correct option:

Use AWS Cost Explorer Resource Optimization to get a report of Amazon EC2 instances that are either idle or have low utilization and use AWS Compute Optimizer to look at instance type recommendations

AWS Cost Explorer helps you identify under-utilized Amazon EC2 instances that may be downsized on an instance by instance basis within the same instance family, and also understand the potential impact on your AWS bill by taking into account your Reserved Instances and Savings Plans.

AWS Compute Optimizer recommends optimal AWS Compute resources for your workloads to reduce costs and improve performance by using machine learning to analyze historical utilization metrics. Compute Optimizer helps you choose the optimal Amazon EC2 instance types, including those that are part of an Amazon EC2 Auto Scaling group, based on your utilization data.

Incorrect options:

Use Amazon S3 Storage class analysis to get recommendations for transitions of objects to Amazon S3 Glacier storage classes to reduce storage costs. You can also automate moving these objects into lower-cost storage tier using Lifecycle Policies - By using Amazon S3 Analytics Storage Class analysis you can analyze storage access patterns to help you decide when to transition the right data to the right storage class. This new Amazon S3 analytics feature observes data access patterns to help you determine when to transition less frequently accessed STANDARD storage to the STANDARD_IA (IA, for infrequent access) storage class. Storage class analysis does not give recommendations for transitions to the ONEZONE_IA or S3 Glacier storage classes.

Use AWS Trusted Advisor checks on Amazon EC2 Reserved Instances to automatically renew reserved instances (RI). AWS Trusted advisor also suggests Amazon RDS idle database instances - AWS Trusted Advisor checks for Amazon EC2 Reserved Instances that are scheduled to expire within the next 30 days or have expired in the preceding 30 days. Reserved Instances do not renew automatically; you can continue using an Amazon EC2 instance covered by the reservation without interruption, but you will be charged On-Demand rates. AWS Trusted advisor does not have a feature to auto-renew Reserved Instances.

Use AWS Compute Optimizer recommendations to help you choose the optimal Amazon EC2 purchasing options and help reserve your instance capacities at reduced costs - AWS Compute Optimizer recommends optimal AWS Compute resources for your workloads to reduce costs and improve performance by using machine learning to analyze historical utilization metrics. Over-provisioning compute can lead to unnecessary infrastructure cost and under-provisioning compute can lead to poor application performance. Compute Optimizer helps you choose the optimal Amazon EC2 instance types, including those that are part of an Amazon EC2 Auto Scaling group, based on your utilization data. It does not recommend instance purchase options.

References:

https://aws.amazon.com/compute-optimizer/

https://aws.amazon.com/premiumsupport/technology/trusted-advisor/best-practice-checklist/

https://docs.aws.amazon.com/AmazonS3/latest/dev/analytics-storage-class.html

Domain

Design Cost-Optimized Architectures

Question 39 Correct

A financial services company wants to identify any sensitive data stored on its Amazon S3 buckets. The company also wants to monitor and protect all data stored on Amazon S3 against any malicious activity.

As a solutions architect, which of the following solutions would you recommend to help address the given requirements?

Use Amazon GuardDuty to monitor any malicious activity on data stored in Amazon S3 as well as to identify any sensitive data stored on Amazon S3

Use Amazon Macie to monitor any malicious activity on data stored in Amazon S3. Use Amazon GuardDuty to identify any sensitive data stored on Amazon S3

Your answer is correct

Use Amazon GuardDuty to monitor any malicious activity on data stored in Amazon S3. Use Amazon Macie to identify any sensitive data stored on Amazon S3

Use Amazon Macie to monitor any malicious activity on data stored in Amazon S3 as well as to identify any sensitive data stored on Amazon S3

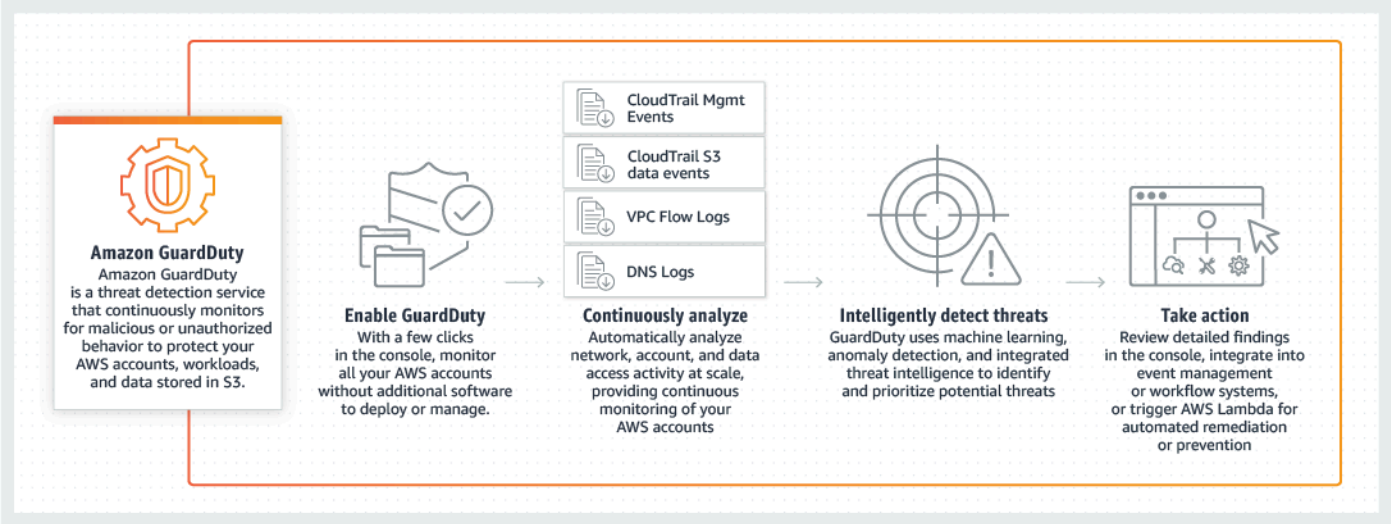
Overall explanation

Correct option:

Use Amazon GuardDuty to monitor any malicious activity on data stored in Amazon S3. Use Amazon Macie to identify any sensitive data stored on Amazon S3

Amazon GuardDuty offers threat detection that enables you to continuously monitor and protect your AWS accounts, workloads, and data stored in Amazon S3. GuardDuty analyzes continuous streams of meta-data generated from your account and network activity found in AWS CloudTrail Events, Amazon VPC Flow Logs, and DNS Logs. It also uses integrated threat intelligence such as known malicious IP addresses, anomaly detection, and machine learning to identify threats more accurately.

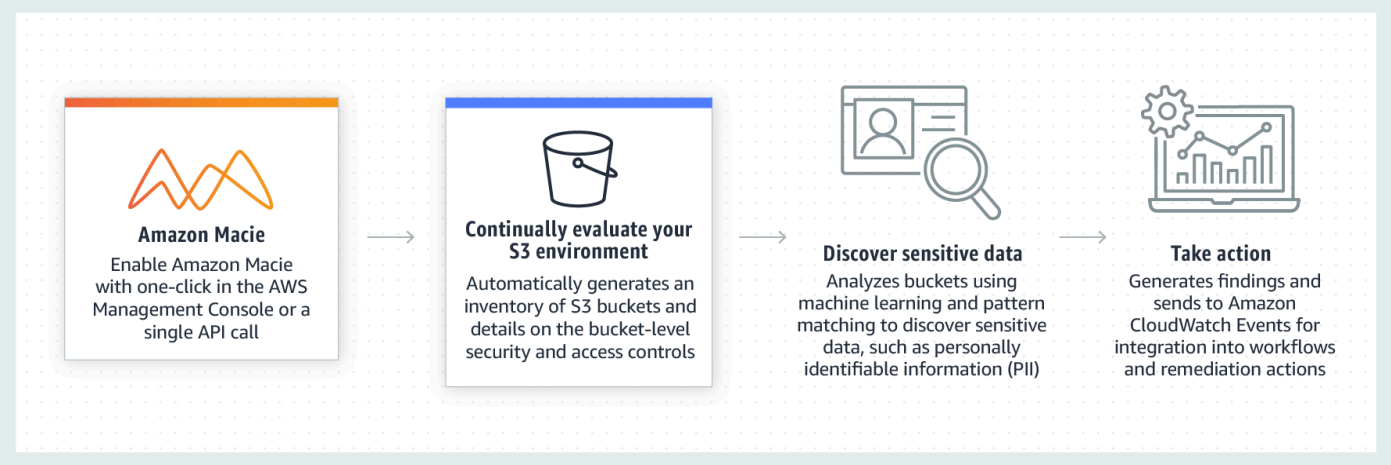
How Amazon GuardDuty works:



via - https://aws.amazon.com/guardduty/

Amazon Macie is a fully managed data security and data privacy service that uses machine learning and pattern matching to discover and protect your sensitive data on Amazon S3. Macie automatically detects a large and growing list of sensitive data types, including personally identifiable information (PII) such as names, addresses, and credit card numbers. It also gives you constant visibility of the data security and data privacy of your data stored in Amazon S3.

How Amazon Macie works:



via - https://aws.amazon.com/macie/

Incorrect options:

Use Amazon GuardDuty to monitor any malicious activity on data stored in Amazon S3 as well as to identify any sensitive data stored on Amazon S3

Use Amazon Macie to monitor any malicious activity on data stored in Amazon S3 as well as to identify any sensitive data stored on Amazon S3

Use Amazon Macie to monitor any malicious activity on data stored in Amazon S3. Use Amazon GuardDuty to identify any sensitive data stored on Amazon S3

These three options contradict the explanation provided above, so these options are incorrect.

References:

https://aws.amazon.com/guardduty/

https://aws.amazon.com/macie/

Domain

Design Secure Architectures

Question 40 **Incorrect**

The engineering team at an e-commerce company is working on cost optimizations for Amazon Elastic Compute Cloud (Amazon EC2) instances. The team wants to manage the workload using a mix of on-demand and spot instances across multiple instance types. They would like to create an Auto Scaling group with a mix of these instances.

Which of the following options would allow the engineering team to provision the instances for this use-case?

You can only use a launch configuration to provision capacity across multiple instance types using both On-Demand Instances and Spot Instances to achieve the desired scale, performance, and cost

You can neither use a launch configuration nor a launch template to provision capacity across multiple instance types using both On-Demand Instances and Spot Instances to achieve the desired scale, performance, and cost

Your answer is incorrect

You can use a launch configuration or a launch template to provision capacity across multiple instance types using both On-Demand Instances and Spot Instances to achieve the desired scale, performance, and cost

Correct answer

You can only use a launch template to provision capacity across multiple instance types using both On-Demand Instances and Spot Instances to achieve the desired scale, performance, and cost

Overall explanation

Correct option:

You can only use a launch template to provision capacity across multiple instance types using both On-Demand Instances and Spot Instances to achieve the desired scale, performance, and cost

A launch template is similar to a launch configuration, in that it specifies instance configuration information such as the ID of the Amazon Machine Image (AMI), the instance type, a key pair, security groups, and the other parameters that you use to launch EC2 instances. Also, defining a launch template instead of a launch configuration allows you to have multiple versions of a template.

With launch templates, you can provision capacity across multiple instance types using both On-Demand Instances and Spot Instances to achieve the desired scale, performance, and cost. Hence this is the correct option.

Incorrect options:

You can only use a launch configuration to provision capacity across multiple instance types using both On-Demand Instances and Spot Instances to achieve the desired scale, performance, and cost

You can use a launch configuration or a launch template to provision capacity across multiple instance types using both On-Demand Instances and Spot Instances to achieve the desired scale, performance, and cost

A launch configuration is an instance configuration template that an Auto Scaling group uses to launch EC2 instances. When you create a launch configuration, you specify information for the instances such as the ID of the Amazon Machine Image (AMI), the instance type, a key pair, one or more security groups, and a block device mapping.

You cannot use a launch configuration to provision capacity across multiple instance types using both On-Demand Instances and Spot Instances. Therefore both these options are incorrect.

You can neither use a launch configuration nor a launch template to provision capacity across multiple instance types using both On-Demand Instances and Spot Instances to achieve the desired scale, performance, and cost - You can use a launch template to provision capacity across multiple instance types using both On-Demand Instances and Spot Instances. So this option is incorrect.

References:

https://docs.aws.amazon.com/autoscaling/ec2/userguide/LaunchTemplates.html

https://docs.aws.amazon.com/autoscaling/ec2/userguide/LaunchConfiguration.html

Domain

Design Cost-Optimized Architectures

Question 41 **Incorrect**

You have been hired as a Solutions Architect to advise a company on the various authentication/authorization mechanisms that AWS offers to authorize an API call within the Amazon API Gateway. The company would prefer a solution that offers built-in user management.

Which of the following solutions would you suggest as the best fit for the given use-case?

Use AWS_IAM authorization

Correct answer

Use Amazon Cognito User Pools

Use Amazon Cognito Identity Pools

Your answer is incorrect

Use AWS Lambda authorizer for Amazon API Gateway

Overall explanation

Correct option:

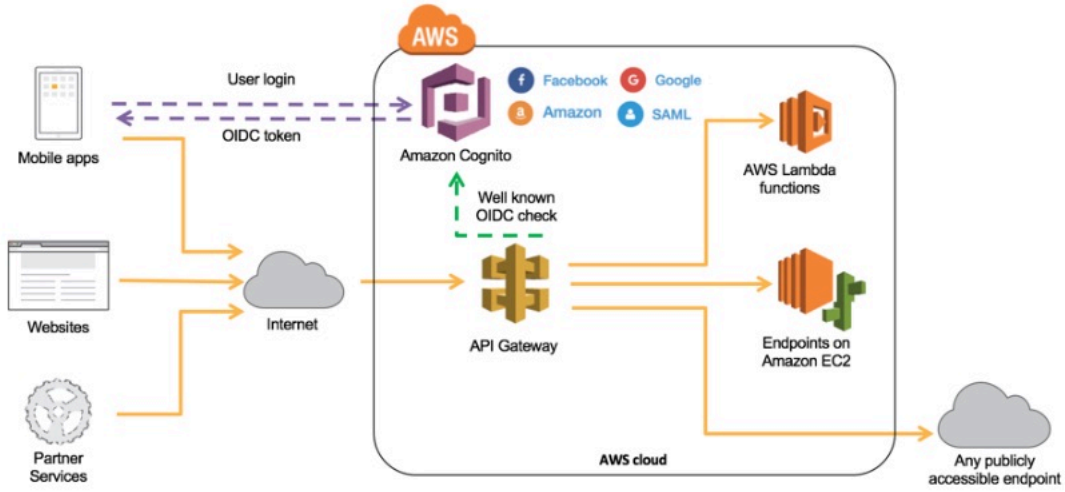
Use Amazon Cognito User Pools

A user pool is a user directory in Amazon Cognito. You can leverage Amazon Cognito User Pools to either provide built-in user management or integrate with external identity providers, such as Facebook, Twitter, Google+, and Amazon. Whether your users sign-in directly or through a third party, all members of the user pool have a directory profile that you can access through a Software Development Kit (SDK).

User pools provide: 1. Sign-up and sign-in services. 2. A built-in, customizable web UI to sign in users. 3. Social sign-in with Facebook, Google, Login with Amazon, and Sign in with Apple, as well as sign-in with SAML identity providers from your user pool. 4. User directory management and user profiles. 5. Security features such as multi-factor authentication (MFA), checks for compromised credentials, account takeover protection, and phone and email verification. 6. Customized workflows and user migration through AWS Lambda triggers.

After creating an Amazon Cognito user pool, in API Gateway, you must then create a COGNITO_USER_POOLS authorizer that uses the user pool.

Amazon Cognito User Pools:



via - https://docs.aws.amazon.com/wellarchitected/latest/serverless-applications-lens/identity-and-access-management.html

Incorrect options:

Use AWS_IAM authorization - For consumers who currently are located within your AWS environment or have the means to retrieve AWS Identity and Access Management (IAM) temporary credentials to access your environment, you can use AWS_IAM authorization and add least-privileged permissions to the respective IAM role to securely invoke your API. API Gateway API Keys is not a security mechanism and should not be used for authorization unless it's a public API. It should be used primarily to track a consumer's usage across your API.

Use AWS Lambda authorizer for Amazon API Gateway - If you have an existing Identity Provider (IdP), you can use an AWS Lambda authorizer for Amazon API Gateway to invoke a Lambda function to authenticate/validate a given user against your Identity Provider. You can use a Lambda authorizer for custom validation logic based on identity metadata.

A Lambda authorizer can send additional information derived from a bearer token or request context values to your backend service. For example, the authorizer can return a map containing user IDs, user names, and scope. By using Lambda authorizers, your backend does not need to map authorization tokens to user-centric data, allowing you to limit the exposure of such information to just the authorization function.

When using Lambda authorizers, AWS strictly advises against passing credentials or any sort of sensitive data via query string parameters or headers, so this is not as secure as using Amazon Cognito User Pools.

In addition, both these options do not offer built-in user management.

Use Amazon Cognito Identity Pools - The two main components of Amazon Cognito are user pools and identity pools. Identity pools provide AWS credentials to grant your users access to other AWS services. To enable users in your user pool to access AWS resources, you can configure an identity pool to exchange user pool tokens for AWS credentials. So, identity pools aren't an authentication mechanism in themselves and hence aren't a choice for this use case.

References:

https://docs.aws.amazon.com/wellarchitected/latest/serverless-applications-lens/identity-and-access-management.html

https://docs.aws.amazon.com/apigateway/latest/developerguide/apigateway-enable-cognito-user-pool.html

https://docs.aws.amazon.com/cognito/latest/developerguide/cognito-user-identity-pools.html

Domain

Design Secure Architectures

Question 42 **Correct**

A news network uses Amazon Simple Storage Service (Amazon S3) to aggregate the raw video footage from its reporting teams across the US. The news network has recently expanded into new geographies in Europe and Asia. The technical teams at the overseas branch offices have reported huge delays in uploading large video files to the destination Amazon S3 bucket.

Which of the following are the MOST cost-effective options to improve the file upload speed into Amazon S3 (Select two)

Create multiple AWS Direct Connect connections between the AWS Cloud and branch offices in Europe and Asia. Use the direct connect connections for faster file uploads into Amazon S3

Use AWS Global Accelerator for faster file uploads into the destination Amazon S3 bucket

Create multiple AWS Site-to-Site VPN connections between the AWS Cloud and branch offices in Europe and Asia. Use these VPN connections for faster file uploads into Amazon S3

Your selection is correct

Use multipart uploads for faster file uploads into the destination Amazon S3 bucket

Your selection is correct

Use Amazon S3 Transfer Acceleration (Amazon S3TA) to enable faster file uploads into the destination S3 bucket

Overall explanation

Correct options:

Use Amazon S3 Transfer Acceleration (Amazon S3TA) to enable faster file uploads into the destination S3 bucket

Amazon S3 Transfer Acceleration enables fast, easy, and secure transfers of files over long distances between your client and an S3 bucket. Amazon S3TA takes advantage of Amazon CloudFront's globally distributed edge locations. As the data arrives at an edge location, data is routed to Amazon S3 over an optimized network path.

Use multipart uploads for faster file uploads into the destination Amazon S3 bucket

Multipart upload allows you to upload a single object as a set of parts. Each part is a contiguous portion of the object's data. You can upload these object parts independently and in any order. If transmission of any part fails, you can retransmit that part without affecting other parts. After all parts of your object are uploaded, Amazon S3 assembles these parts and creates the object. In general, when your object size reaches 100 MB, you should consider using multipart uploads instead of uploading the object in a single operation. Multipart upload provides improved throughput, therefore it facilitates faster file uploads.

Incorrect options:

Create multiple AWS Direct Connect connections between the AWS Cloud and branch offices in Europe and Asia. Use the direct connect connections for faster file uploads into Amazon S3 - AWS Direct Connect is a cloud service solution that makes it easy to establish a dedicated network connection from your premises to AWS. AWS Direct Connect lets you establish

a dedicated network connection between your network and one of the AWS Direct Connect locations. Direct connect takes significant time (several months) to be provisioned and is an overkill for the given use-case.

Create multiple AWS Site-to-Site VPN connections between the AWS Cloud and branch offices in Europe and Asia. Use these VPN connections for faster file uploads into Amazon S3 - AWS Site-to-Site VPN enables you to securely connect your on-premises network or branch office site to your Amazon Virtual Private Cloud (Amazon VPC). You can securely extend your data center or branch office network to the cloud with an AWS Site-to-Site VPN connection. A VPC VPN Connection utilizes IPSec to establish encrypted network connectivity between your intranet and Amazon VPC over the Internet. VPN Connections are a good solution if you have low to modest bandwidth requirements and can tolerate the inherent variability in Internet-based connectivity. Site-to-site VPN will not help in accelerating the file transfer speeds into S3 for the given use-case.

Use AWS Global Accelerator for faster file uploads into the destination Amazon S3 bucket - AWS Global Accelerator is a service that improves the availability and performance of your applications with local or global users. It provides static IP addresses that act as a fixed entry point to your application endpoints in a single or multiple AWS Regions, such as your Application Load Balancers, Network Load Balancers or Amazon EC2 instances. AWS Global Accelerator will not help in accelerating the file transfer speeds into S3 for the given use-case.

References:

https://docs.aws.amazon.com/AmazonS3/latest/dev/transfer-acceleration.html

https://docs.aws.amazon.com/AmazonS3/latest/dev/uploadobjusingmpu.html

Domain

Design Cost-Optimized Architectures

Question 43 **Correct**

A leading online gaming company is migrating its flagship application to AWS Cloud for delivering its online games to users across the world. The company would like to use a Network Load Balancer to handle millions of requests per second. The engineering team has provisioned multiple instances in a public subnet and specified these instance IDs as the targets for the NLB.

As a solutions architect, can you help the engineering team understand the correct routing mechanism for these target instances?

Traffic is routed to instances using the primary elastic IP address specified in the primary network interface for the instance

Traffic is routed to instances using the instance ID specified in the primary network interface for the instance

Your answer is correct

Traffic is routed to instances using the primary private IP address specified in the primary network interface for the instance

Traffic is routed to instances using the primary public IP address specified in the primary network interface for the instance

Overall explanation

Correct option:

Traffic is routed to instances using the primary private IP address specified in the primary network interface for the instance

A Network Load Balancer functions at the fourth layer of the Open Systems Interconnection (OSI) model. It can handle millions of requests per second. After the load balancer receives a connection request, it selects a target from the target group for the default rule. It attempts to open a TCP connection to the selected target on the port specified in the listener configuration.

Request Routing and IP Addresses -

If you specify targets using an instance ID, traffic is routed to instances using the primary private IP address specified in the primary network interface for the instance. The load balancer rewrites the destination IP address from the data packet before forwarding it to the target instance.

If you specify targets using IP addresses, you can route traffic to an instance using any private IP address from one or more network interfaces. This enables multiple applications on an instance to use the same port. Note that each network interface can have its security group. The load balancer rewrites the destination IP address before forwarding it to the target.

Incorrect options:

Traffic is routed to instances using the primary public IP address specified in the primary network interface for the instance - If you specify targets using an instance ID, traffic is routed to instances using the primary private IP address specified in the primary network interface for the instance. So public IP address cannot be used to route the traffic to the instance.

Traffic is routed to instances using the primary elastic IP address specified in the primary network interface for the instance - If you specify targets using an instance ID, traffic is routed to instances using the primary private IP address specified in the primary network interface for the instance. So elastic IP address cannot be used to route the traffic to the instance.

Traffic is routed to instances using the instance ID specified in the primary network interface for the instance - You cannot use instance ID to route traffic to the instance. This option is just added as a distractor.

References:

https://docs.aws.amazon.com/elasticloadbalancing/latest/network/introduction.html

https://docs.aws.amazon.com/elasticloadbalancing/latest/network/load-balancer-target-groups.html

Domain

Design High-Performing Architectures

Question 44 **Correct**

A financial services company has deployed its flagship application on Amazon EC2 instances. Since the application handles sensitive customer data, the security team at the company wants to ensure that any third-party Secure Sockets Layer certificate (SSL certificate) SSL/Transport Layer Security (TLS) certificates configured on Amazon EC2 instances via the AWS Certificate Manager (ACM) are renewed before their expiry date. The company has hired you as an AWS Certified Solutions Architect Associate to build a solution that notifies the security team 30 days before the certificate expiration. The solution should require the least amount of scripting and maintenance effort.

What will you recommend?

Monitor the `days to expiry` Amazon CloudWatch metric for certificates created via ACM. Create a CloudWatch alarm to monitor such certificates based on the `days to expiry` metric and then trigger a custom action of notifying the security team

Leverage AWS Config managed rule to check if any SSL/TLS certificates created via ACM are marked for expiration within 30 days. Configure the rule to trigger an Amazon SNS notification to the security team if any certificate expires within 30 days

Monitor the `days to expiry` Amazon CloudWatch metric for certificates imported into ACM. Create a CloudWatch alarm to monitor such certificates based on the `days to expiry` metric and then trigger a custom action of notifying the security team

Your answer is correct

Leverage AWS Config managed rule to check if any third-party SSL/TLS certificates imported into ACM are marked for expiration within 30 days. Configure the rule to trigger an Amazon SNS notification to the security team if any certificate expires within 30 days

Overall explanation

Correct option:

Leverage AWS Config managed rule to check if any third-party SSL/TLS certificates imported into ACM are marked for expiration within 30 days. Configure the rule to trigger an Amazon SNS notification to the security team if any certificate expires within 30 days

AWS Certificate Manager is a service that lets you easily provision, manage, and deploy public and private Secure Sockets Layer/Transport Layer Security (SSL/TLS) certificates for use with AWS services and your internal connected resources. SSL/TLS certificates are used to secure network

communications and establish the identity of websites over the Internet as well as resources on private networks.

AWS Config provides a detailed view of the configuration of AWS resources in your AWS account. This includes how the resources are related to one another and how they were configured in the past so that you can see how the configurations and relationships change over time.

How AWS Config Works

PDF | RSS

When you turn on AWS Config, it first discovers the supported AWS resources that exist in your account and generates a configuration item for each resource.

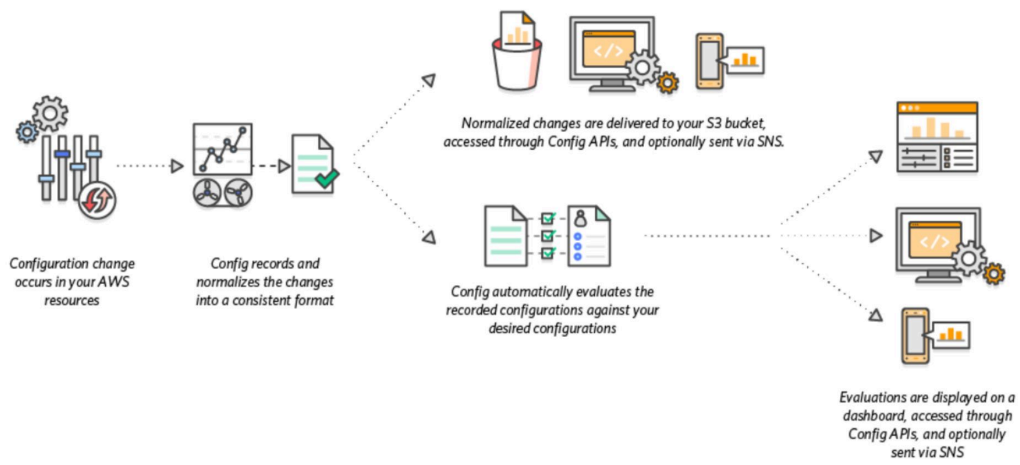
AWS Config also generates configuration items when the configuration of a resource changes, and it maintains historical records of the configuration items of your resources from the time you start the configuration recorder. By default, AWS Config creates configuration items for every supported resource in the region. If you don't want AWS Config to create configuration items for all supported resources, you can specify the resource types that you want it to track.

AWS Config keeps track of all changes to your resources by invoking the Describe or the List API call for each resource in your account. The service uses those same API calls to capture configuration details for all related resources.

For example, removing an egress rule from a VPC security group causes AWS Config to invoke a Describe API call on the security group. AWS Config then invokes a Describe API call on all of the instances associated with the security group. The updated configurations of the security group (the resource) and of each instance (the related resources) are recorded as configuration items and delivered in a configuration stream to an Amazon Simple Storage Service (Amazon S3) bucket.

AWS Config also tracks the configuration changes that were not initiated by the API. AWS Config examines the resource configurations periodically and generates configuration items for the configurations that have changed.

If you are using AWS Config rules, AWS Config continuously evaluates your AWS resource configurations for desired settings. Depending on the rule, AWS Config will evaluate your resources either in response to configuration changes or periodically. Each rule is associated with an AWS Lambda function, which contains the evaluation logic for the rule. When AWS Config evaluates your resources, it invokes the rule's AWS Lambda function. The function returns the compliance status of the evaluated resources. If a resource violates the conditions of a rule, AWS Config flags the resource and the rule as noncompliant. When the compliance status of a resource changes, AWS Config sends a notification to your Amazon SNS topic.



via - https://docs.aws.amazon.com/config/latest/developerguide/how-does-config-work.html

AWS Config provides AWS-managed rules, which are predefined, customizable rules that AWS Config uses to evaluate whether your AWS resources comply with common best practices. You can leverage an AWS Config managed rule to check if any ACM certificates in your account are marked for expiration within the specified number of days. Certificates provided by ACM are automatically renewed. ACM does not automatically renew the certificates that you import. The rule is NON_COMPLIANT if your certificates are about to expire.

acm-certificate-expiration-check

PDF | RSS

Checks if AWS Certificate Manager Certificates in your account are marked for expiration within the specified number of days. Certificates provided by ACM are automatically renewed. ACM does not automatically renew certificates that you import. The rule is NON_COMPLIANT if your certificates are about to expire.

Identifier: ACM_CERTIFICATE_EXPIRATION_CHECK

Trigger type: Configuration changes

AWS Region: All supported AWS regions except China (Beijing), China (Ningxia), Asia Pacific (Osaka), Europe (Milan) Region

Parameters:

daysToExpiration (Optional)
Type: int
Default: 14
Specify the number of days before the rule flags the ACM Certificate as noncompliant.

via - https://docs.aws.amazon.com/config/latest/developerguide/how-does-config-work.html

You can configure AWS Config to stream configuration changes and notifications to an Amazon SNS topic. For example, when a resource is updated, you can get a notification sent to your email, so that you can view the changes. You can also be notified when AWS Config evaluates your custom or managed rules against your resources.

Incorrect options:

Monitor the `days to expiry` Amazon CloudWatch metric for certificates imported into ACM. Create a CloudWatch alarm to monitor such certificates based on the `days to expiry` metric and then trigger a custom action of notifying the security team - AWS Certificate Manager (ACM) does not attempt to renew third-party certificates that are imported. Also, an administrator needs to reconfigure missing DNS records for certificates that use DNS validation if the record was removed for any reason after the certificate was issued. Metrics and events provide you visibility into such certificates that require intervention to continue the renewal process. Amazon CloudWatch metrics and Amazon EventBridge events are enabled for all certificates that are managed by ACM. Users can monitor `days to expiry` as a metric for ACM certificates through Amazon CloudWatch. An Amazon EventBridge expiry event is published for any certificate that is at least 45 days away from expiry by default. Users can build alarms to monitor certificates based on days to expiry and also trigger custom actions such as calling a Lambda function or paging an administrator.

It is certainly possible to use the `days to expiry` CloudWatch metric to build a CloudWatch alarm to monitor the imported ACM certificates. The alarm will, in turn, trigger a notification to the security team. But this option needs more configuration effort than directly using the AWS Config managed rule that is available off-the-shelf.

Leverage AWS Config managed rule to check if any SSL/TLS certificates created via ACM are marked for expiration within 30 days. Configure the rule to trigger an Amazon SNS notification to the security team if any certificate expires within 30 days

Monitor the `days to expiry` Amazon CloudWatch metric for certificates created via ACM. Create a CloudWatch alarm to monitor such certificates based on the `days to expiry` metric and then trigger a custom action of notifying the security team

Any SSL/TLS certificates created via ACM do not need any monitoring/intervention for expiration. ACM automatically renews such certificates. Hence both these options are incorrect.

References:

https://docs.aws.amazon.com/config/latest/developerguide/WhatIsConfig.html

https://docs.aws.amazon.com/config/latest/developerguide/how-does-config-work.html

https://docs.aws.amazon.com/config/latest/developerguide/evaluate-config.html

https://docs.aws.amazon.com/config/latest/developerguide/acm-certificate-expiration-check.html

https://aws.amazon.com/blogs/security/how-to-monitor-expirations-of-imported-certificates-in-aws-certificate-manager-acm/

Domain

Design Secure Architectures

Question 45 Correct

A weather forecast agency collects key weather metrics across multiple cities in the US and sends this data in the form of key-value pairs to AWS Cloud at a one-minute frequency.

As a solutions architect, which of the following AWS services would you use to build a solution for processing and then reliably storing this data with high availability? (Select two)

Your selection is correct

AWS Lambda

Amazon RDS

Amazon Redshift

Amazon ElastiCache

Your selection is correct

Amazon DynamoDB

Overall explanation

Correct options:

AWS Lambda

With AWS Lambda, you can run code without provisioning or managing servers. You pay only for the compute time that you consume—there's no charge when your code isn't running. You can run code for virtually any type of application or backend service—all with zero administration.

Amazon DynamoDB

Amazon DynamoDB is a key-value and document database that delivers single-digit millisecond performance at any scale. It's a fully managed, multi-region, multi-master, durable database with built-in security, backup and restore, and in-memory caching for internet-scale applications. Amazon DynamoDB is a NoSQL database and it's best suited to store data in key-value pairs.

AWS Lambda can be combined with DynamoDB to process and capture the key-value data from the IoT sources described in the use-case. So both these options are correct.

Incorrect options:

Amazon Redshift - Amazon Redshift is a fully-managed petabyte-scale cloud-based data warehouse product designed for large scale data set storage and analysis. You cannot use

Redshift to capture data in key-value pairs from the IoT sources, so this option is not correct.

Amazon ElastiCache - Amazon ElastiCache allows you to seamlessly set up, run, and scale popular open-Source compatible in-memory data stores in the cloud. Build data-intensive apps or boost the performance of your existing databases by retrieving data from high throughput and low latency in-memory data stores. Amazon ElastiCache is a popular choice for real-time use cases like Caching, Session Stores, Gaming, Geospatial Services, Real-Time Analytics, and Queuing. Elasticache is used as a caching layer in front of relational databases. It is not a good fit to store data in key-value pairs from the IoT sources, so this option is not correct.

Amazon RDS - Amazon Relational Database Service (Amazon RDS) makes it easy to set up, operate, and scale a relational database in the cloud. It provides cost-efficient and resizable capacity while automating time-consuming administration tasks such as hardware provisioning, database setup, patching, and backups. Relational databases are not a good fit to store data in key-value pairs, so this option is not correct.

References:

https://aws.amazon.com/dynamodb/

https://aws.amazon.com/lambda/faqs/

Domain

Design Resilient Architectures

Question 46 **Correct**

The engineering team at an e-commerce company wants to migrate from Amazon Simple Queue Service (Amazon SQS) Standard queues to FIFO (First-In-First-Out) queues with batching.

As a solutions architect, which of the following steps would you have in the migration checklist? (Select three)

Make sure that the throughput for the target FIFO (First-In-First-Out) queue does not exceed 300 messages per second

Your selection is correct

Delete the existing standard queue and recreate it as a FIFO (First-In-First-Out) queue

Make sure that the name of the FIFO (First-In-First-Out) queue is the same as the standard queue

Convert the existing standard queue into a FIFO (First-In-First-Out) queue

Your selection is correct

Make sure that the name of the FIFO (First-In-First-Out) queue ends with the .fifo suffix

Your selection is correct

Make sure that the throughput for the target FIFO (First-In-First-Out) queue does not exceed 3,000 messages per second

Overall explanation

Correct options:

Delete the existing standard queue and recreate it as a FIFO (First-In-First-Out) queue

Make sure that the name of the FIFO (First-In-First-Out) queue ends with the .fifo suffix

Make sure that the throughput for the target FIFO (First-In-First-Out) queue does not exceed 3,000 messages per second

Amazon Simple Queue Service (SQS) is a fully managed message queuing service that enables you to decouple and scale microservices, distributed systems, and serverless applications. Amazon SQS eliminates the complexity and overhead associated with managing and operating message oriented middleware, and empowers developers to focus on differentiating work. Using Amazon SQS, you can send, store, and receive messages between software components at any volume, without losing messages or requiring other services to be available.

Amazon SQS offers two types of message queues. Standard queues offer maximum throughput, best-effort ordering, and at-least-once delivery. SQS FIFO queues are designed to guarantee that messages are processed exactly once, in the exact order that they are sent.

By default, FIFO queues support up to 3,000 messages per second with batching, or up to 300 messages per second (300 send, receive, or delete operations per second) without batching. Therefore, using batching you can meet a throughput requirement of upto 3,000 messages per second.

The name of a FIFO queue must end with the .fifo suffix. The suffix counts towards the 80-character queue name limit. To determine whether a queue is FIFO, you can check whether the queue name ends with the suffix.

If you have an existing application that uses standard queues and you want to take advantage of the ordering or exactly-once processing features of FIFO queues, you need to configure the queue and your application correctly. You can't convert an existing standard queue into a FIFO queue. To make the move, you must either create a new FIFO queue for your application or delete your existing standard queue and recreate it as a FIFO queue.

Incorrect options:

Convert the existing standard queue into a FIFO (First-In-First-Out) queue

Make sure that the name of the FIFO (First-In-First-Out) queue is the same as the standard queue - The name of a FIFO queue must end with the .fifo suffix.

Make sure that the throughput for the target FIFO (First-In-First-Out) queue does not exceed 300 messages per second - By default, FIFO queues support up to 3,000 messages per second with batching.

References:

https://aws.amazon.com/sqs/faqs/

https://docs.aws.amazon.com/AWSSimpleQueueService/latest/SQSDeveloperGuide/FIFO-queues.html

Domain

Design Resilient Architectures

Question 47 **Incorrect**

A cyber security company is running a mission critical application using a single Spread placement group of Amazon EC2 instances. The company needs 15 Amazon EC2 instances for optimal performance.

How many Availability Zones (AZs) will the company need to deploy these Amazon EC2 instances per the given use-case?

15

Correct answer

3

Your answer is incorrect

14

7

Overall explanation

Correct option:

3

When you launch a new Amazon EC2 instance, the EC2 service attempts to place the instance in such a way that all of your instances are spread out across underlying hardware to minimize correlated failures. You can use placement groups to influence the placement of a group of interdependent instances to meet the needs of your workload. Depending on the type of workload, you can create a placement group using one of the following placement strategies:

Cluster placement group

Partition placement group

Spread placement group.

A Spread placement group is a group of instances that are each placed on distinct racks, with each rack having its own network and power source.

Spread placement groups are recommended for applications that have a small number of critical instances that should be kept separate from each other. Launching instances in a spread placement group reduces the risk of simultaneous failures that might occur when instances share the same racks.

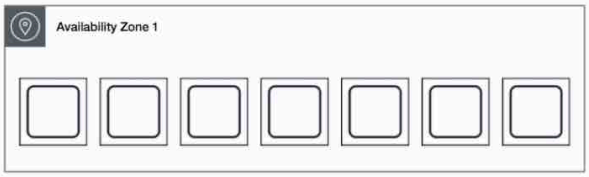
A spread placement group can span multiple Availability Zones in the same Region. You can have a maximum of seven running instances per Availability Zone per group. Therefore, to deploy 15 Amazon EC2 instances in a single Spread placement group, the company needs to use 3 Availability Zones.

Spread placement group overview:

Spread placement groups

A spread placement group is a group of instances that are each placed on distinct racks, with each rack having its own network and power source.

The following image shows seven instances in a single Availability Zone that are placed into a spread placement group. The seven instances are placed on seven different racks.



Spread placement groups are recommended for applications that have a small number of critical instances that should be kept separate from each other. Launching instances in a spread placement group reduces the risk of simultaneous failures that might occoccur when instances share the same racks. Spread placement groups provide access to distinct racks, and are therefore suitable for mixing instance types or launching instances over time.

A spread placement group can span multiple Availability Zones in the same Region. You can have a maximum of seven running instances per Availability Zone per group.

If you start or launch an instance in a spread placement group and there is insufficient unique hardware to fulfill the request, the request fails. Amazon EC2 makes more distinct hardware available over time, so you can try your request again later.

via - https://docs.aws.amazon.com/AWSEC2/latest/UserGuide/placement-groups.html

Incorrect options:

7

14

15

These three options contradict the details provided in the explanation above, so these options are incorrect.

Reference:

https://docs.aws.amazon.com/AWSEC2/latest/UserGuide/placement-groups.html

Domain

Design Resilient Architectures

Question 48 Correct

A Big Data analytics company wants to set up an AWS cloud architecture that throttles requests in case of sudden traffic spikes. The company is looking for AWS services that can be used for buffering or throttling to handle such traffic variations.

Which of the following services can be used to support this requirement?

Elastic Load Balancer, Amazon Simple Queue Service (Amazon SQS), AWS Lambda

Your answer is correct

Amazon API Gateway, Amazon Simple Queue Service (Amazon SQS) and Amazon Kinesis

Amazon Gateway Endpoints, Amazon Simple Queue Service (Amazon SQS) and Amazon Kinesis

Amazon Simple Queue Service (Amazon SQS), Amazon Simple Notification Service (Amazon SNS) and AWS Lambda

Overall explanation

Correct option:

Throttling is the process of limiting the number of requests an authorized program can submit to a given operation in a given amount of time.

Amazon API Gateway, Amazon Simple Queue Service (Amazon SQS) and Amazon Kinesis

To prevent your API from being overwhelmed by too many requests, Amazon API Gateway throttles requests to your API using the token bucket algorithm, where a token counts for a request. Specifically, API Gateway sets a limit on a steady-state rate and a burst of request

submissions against all APIs in your account. In the token bucket algorithm, the burst is the maximum bucket size.

Amazon Simple Queue Service (Amazon SQS) - Amazon Simple Queue Service (SQS) is a fully managed message queuing service that enables you to decouple and scale microservices, distributed systems, and serverless applications. Amazon SQS offers buffer capabilities to smooth out temporary volume spikes without losing messages or increasing latency.

Amazon Kinesis - Amazon Kinesis is a fully managed, scalable service that can ingest, buffer, and process streaming data in real-time.

Incorrect options:

Amazon Simple Queue Service (Amazon SQS), Amazon Simple Notification Service (Amazon SNS) and AWS Lambda - Amazon SQS has the ability to buffer its messages. Amazon Simple Notification Service (SNS) cannot buffer messages and is generally used with SQS to provide the buffering facility. When requests come in faster than your Lambda function can scale, or when your function is at maximum concurrency, additional requests fail as the Lambda throttles those requests with error code 429 status code. So, this combination of services is incorrect.

Amazon Gateway Endpoints, Amazon Simple Queue Service (Amazon SQS) and Amazon Kinesis - A Gateway Endpoint is a gateway that you specify as a target for a route in your route table for traffic destined to a supported AWS service. This cannot help in throttling or buffering of requests. Amazon SQS and Kinesis can buffer incoming data. Since Gateway Endpoint is an incorrect service for throttling or buffering, this option is incorrect.

Elastic Load Balancer, Amazon Simple Queue Service (Amazon SQS), AWS Lambda - Elastic Load Balancer cannot throttle requests. Amazon SQS can be used to buffer messages. When requests come in faster than your Lambda function can scale, or when your function is at maximum concurrency, additional requests fail as the Lambda throttles those requests with error code 429 status code. So, this combination of services is incorrect.

References:

https://docs.aws.amazon.com/apigateway/latest/developerguide/api-gateway-request-throttling.html

https://aws.amazon.com/sqs/features/

Domain

Design Resilient Architectures

Question 49 Correct

The development team at a social media company wants to handle some complicated queries such as "What are the number of likes on the videos that have been posted by friends of a user A?".

As a solutions architect, which of the following AWS database services would you suggest as the BEST fit to handle such use cases?

Amazon Aurora

Amazon Redshift

Your answer is correct

Amazon Neptune

Amazon OpenSearch Service

Overall explanation

Correct option:

Amazon Neptune

Amazon Neptune is a fast, reliable, fully managed graph database service that makes it easy to build and run applications that work with highly connected datasets. The core of Amazon Neptune is a purpose-built, high-performance graph database engine optimized for storing

billions of relationships and querying the graph with milliseconds latency. Neptune powers graph use cases such as recommendation engines, fraud detection, knowledge graphs, drug discovery, and network security.

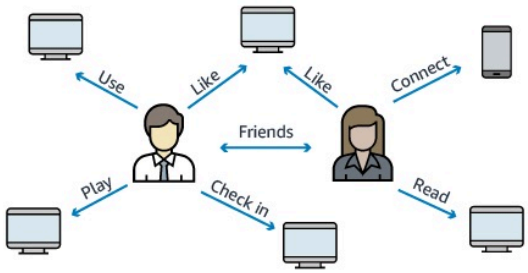
Amazon Neptune is highly available, with read replicas, point-in-time recovery, continuous backup to Amazon S3, and replication across Availability Zones. Neptune is secure with support for HTTPS encrypted client connections and encryption at rest. Neptune is fully managed, so you no longer need to worry about database management tasks such as hardware provisioning, software patching, setup, configuration, or backups.

Amazon Neptune can quickly and easily process large sets of user-profiles and interactions to build social networking applications. Neptune enables highly interactive graph queries with high throughput to bring social features into your applications. For example, if you are building a social feed into your application, you can use Neptune to provide results that prioritize showing your users the latest updates from their family, from friends whose updates they 'Like,' and from friends who live close to them.

Social Networking example with Amazon Neptune:

Social Networking

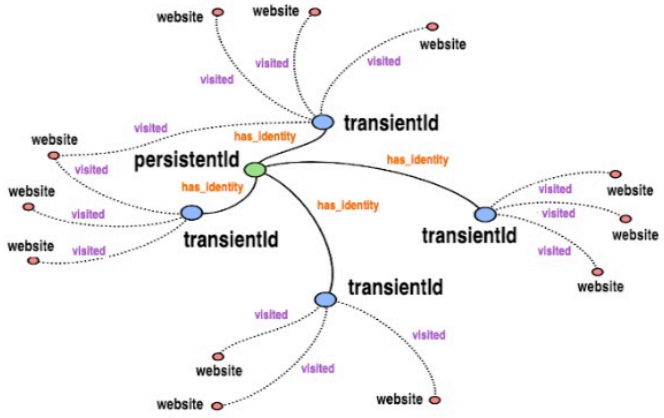
Amazon Neptune can quickly and easily process large sets of user profiles and interactions to build social networking applications. Neptune enables highly interactive graph queries with high throughput to bring social features into your applications. For example, if you are building a social feed into your application, you can use Neptune to provide results that prioritize showing your users the latest updates from their family, from friends whose updates they 'Like,' and from friends who live close to them.



via - https://aws.amazon.com/neptune/

Identity graphs example with Amazon Neptune:

Identity Graphs



You can use Neptune to build identity graphs for any identity resolution solutions, including device and social graphs, personalization and recommendations, and pattern detection. Using a graph database for an identity graph enables you to link identifiers and update profiles easily and query at ultra-low latency — enabling faster updates and up-to-date profile data for ad targeting, personalization, analytics, and ad attribution. Learn more about Identity Graphs on AWS.

via - https://aws.amazon.com/neptune/

Incorrect options:

Amazon OpenSearch Service - Amazon OpenSearch Service is a managed service that makes it easy for you to perform interactive log analytics, real-time application monitoring, website search, and more. OpenSearch is an open source, distributed search and analytics suite derived from

Elasticsearch. Amazon OpenSearch Service offers the latest versions of OpenSearch, support for 19 versions of Elasticsearch (1.5 to 7.10 versions), as well as visualization capabilities powered by OpenSearch Dashboards and Kibana (1.5 to 7.10 versions). Amazon OpenSearch Service currently has tens of thousands of active customers with hundreds of thousands of clusters under management processing trillions of requests per month.

Amazon Redshift - Amazon Redshift is a fully-managed petabyte-scale cloud-based data warehouse product designed for large scale data set storage and analysis. The given use-case is not about data warehousing, so this is not a correct option.

Amazon Aurora - Amazon Aurora is a MySQL and PostgreSQL-compatible relational database built for the cloud, that combines the performance and availability of traditional enterprise databases with the simplicity and cost-effectiveness of open source databases. Amazon Aurora features a distributed, fault-tolerant, self-healing storage system that auto-scales up to 64 terabytes per database instance. Aurora is not an in-memory database. Here, we need a graph database due to the highly connected datasets and queries, therefore Neptune is the best answer.

Reference:

https://aws.amazon.com/neptune/

Domain

Design High-Performing Architectures

Question 50 **Correct**

A company has grown from a small startup to an enterprise employing over 1000 people. As the team size has grown, the company has recently observed some strange behavior, with Amazon S3 buckets settings being changed regularly.

How can you figure out what's happening without restricting the rights of the users?

Your answer is correct

Use AWS CloudTrail to analyze API calls

Use Amazon S3 access logs to analyze user access using Athena

Implement an IAM policy to forbid users to change Amazon S3 bucket settings

Implement a bucket policy requiring AWS Multi-Factor Authentication (AWS MFA) for all operations

Overall explanation

Correct option:

Use AWS CloudTrail to analyze API calls

AWS CloudTrail is a service that enables governance, compliance, operational auditing, and risk auditing of your AWS account. With AWS CloudTrail, you can log, continuously monitor, and retain account activity related to actions across your AWS infrastructure. AWS CloudTrail provides event history of your AWS account activity, including actions taken through the AWS Management Console, AWS SDKs, command-line tools, and other AWS services.

In general, to analyze any API calls made within an AWS account, AWS CloudTrail is used. You can record the actions that are taken by users, roles, or AWS services on Amazon S3 resources and maintain log records for auditing and compliance purposes. To do this, you can use server access logging, AWS CloudTrail logging, or a combination of both. AWS recommends that you use AWS CloudTrail for logging bucket and object-level actions for your Amazon S3 resources.

Incorrect options:

Implement an IAM policy to forbid users to change Amazon S3 bucket settings - You manage access in AWS by creating policies and attaching them to IAM identities (users, groups of users, or roles) or AWS resources. A policy is an object in AWS that, when associated with an identity or resource, defines their permissions. AWS evaluates these policies when an IAM

principal (user or role) makes a request. Permissions in the policies determine whether the request is allowed or denied. Most policies are stored in AWS as JSON documents. AWS supports six types of policies: identity-based policies, resource-based policies, permissions boundaries, AWS Organizations service control policy (SCP), access control list (ACL), and session policies.

Implementing an IAM policy to forbid users would be disruptive and wouldn't go unnoticed.

Use Amazon S3 access logs to analyze user access using Athena - Amazon S3 server access logging provides detailed records for the requests that are made to a bucket. Server access logs are useful for many applications. For example, access log information can be useful in security and access audits. It can also help you learn about your customer base and understand your Amazon S3 bill. AWS recommends that you use AWS CloudTrail for logging bucket and object-level actions for your Amazon S3 resources, as it provides more options to store, analyze and act on the log information.

Implement a bucket policy requiring AWS Multi-Factor Authentication (AWS MFA) for all operations - Amazon S3 supports MFA-protected API access, a feature that can enforce multi-factor authentication (MFA) for access to your Amazon S3 resources. Multi-factor authentication provides an extra level of security that you can apply to your AWS environment. It is a security feature that requires users to prove the physical possession of an MFA device by providing a valid MFA code. Changing the bucket policy to require MFA would not go unnoticed.

References:

https://docs.aws.amazon.com/AmazonS3/latest/userguide/logging-with-S3.html

https://aws.amazon.com/cloudtrail/

https://docs.aws.amazon.com/IAM/latest/UserGuide/access_policies.html

https://docs.aws.amazon.com/AmazonS3/latest/dev/ServerLogs.html

https://docs.aws.amazon.com/AmazonS3/latest/dev/example-bucket-policies.html#example-bucket-policies-use-case-7

Domain

Design Resilient Architectures

Question 51 **Correct**

An IT company has built a custom data warehousing solution for a retail organization by using Amazon Redshift. As part of the cost optimizations, the company wants to move any historical data (any data older than a year) into Amazon S3, as the daily analytical reports consume data for just the last one year. However the analysts want to retain the ability to cross-reference this historical data along with the daily reports.

The company wants to develop a solution with the LEAST amount of effort and MINIMUM cost. As a solutions architect, which option would you recommend to facilitate this use-case?

Use AWS Glue ETL job to load the Amazon S3 based historical data into Redshift. Once the ad-hoc queries are run for the historic data, it can be removed from Amazon Redshift

Use the Amazon Redshift COPY command to load the Amazon S3 based historical data into Amazon Redshift. Once the ad-hoc queries are run for the historic data, it can be removed from Amazon Redshift

Your answer is correct

Use Amazon Redshift Spectrum to create Amazon Redshift cluster tables pointing to the underlying historical data in Amazon S3. The analytics team can then query this historical data to cross-reference with the daily reports from Redshift

Setup access to the historical data via Amazon Athena. The analytics team can run historical data queries on Amazon Athena and continue the daily reporting on Amazon Redshift. In case the reports need to be cross-referenced, the analytics team need to export these in flat files and then do further analysis

Overall explanation

Correct option:

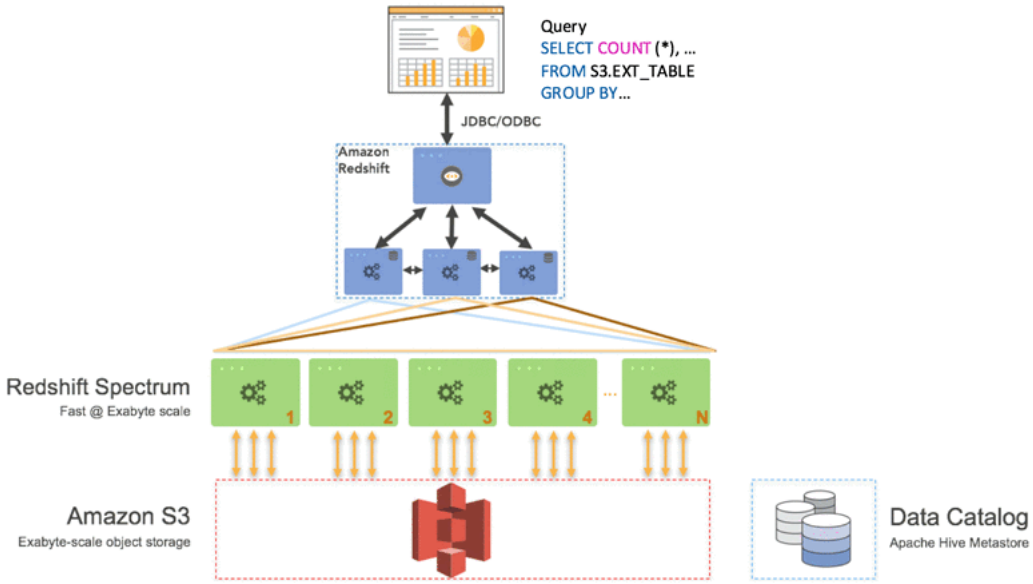
Use Amazon Redshift Spectrum to create Amazon Redshift cluster tables pointing to the underlying historical data in Amazon S3. The analytics team can then query this historical data to cross-reference with the daily reports from Redshift

Amazon Redshift is a fully-managed petabyte-scale cloud-based data warehouse product designed for large scale data set storage and analysis.

Using Amazon Redshift Spectrum, you can efficiently query and retrieve structured and semistructured data from files in Amazon S3 without having to load the data into Amazon Redshift tables.

Amazon Redshift Spectrum resides on dedicated Amazon Redshift servers that are independent of your cluster. Redshift Spectrum pushes many compute-intensive tasks, such as predicate filtering and aggregation, down to the Redshift Spectrum layer. Thus, Amazon Redshift Spectrum queries use much less of your cluster's processing capacity than other queries.

Redshift Spectrum Overview:



via - https://aws.amazon.com/blogs/big-data/amazon-redshift-spectrum-extends-data-warehousing-out-to-exabytes-no-loading-required/

Incorrect options:

Setup access to the historical data via Amazon Athena. The analytics team can run historical data queries on Amazon Athena and continue the daily reporting on Amazon Redshift. In case the reports need to be cross-referenced, the analytics team need to export these in flat files and then do further analysis - Amazon Athena is an interactive query service that makes it easy to analyze data directly in Amazon S3 using standard SQL. Athena is serverless, so there is no infrastructure to set up or manage, and customers pay only for the queries they run. You can use Athena to process logs, perform ad-hoc analysis, and run interactive queries. Providing access to historical data via Athena would mean that historical data reconciliation would become difficult as the daily report would still be produced via Redshift. Such a setup is cumbersome to maintain on a day to day basis. Hence the option to use Athena is ruled out.

Use the Amazon Redshift COPY command to load the Amazon S3 based historical data into Amazon Redshift. Once the ad-hoc queries are run for the historic data, it can be removed from Amazon Redshift

Use AWS Glue ETL job to load the Amazon S3 based historical data into Redshift. Once the ad-hoc queries are run for the historic data, it can be removed from Amazon Redshift

Loading historical data into Amazon Redshift via COPY command or AWS Glue ETL job would cost heavy for a one-time ad-hoc process. The same result can be achieved more cost-efficiently by using Amazon Redshift Spectrum. Therefore both these options to load historical data into Redshift are also incorrect for the given use-case.

References:

https://docs.aws.amazon.com/redshift/latest/dg/c-using-spectrum.html#c-spectrum-overview
https://aws.amazon.com/blogs/big-data/

amazon-redshift-spectrum-extends-data-warehousing-out-to-exabytes-no-loading-required/

Domain

Design High-Performing Architectures

Question 52 Incorrect

A financial services company is looking to move its on-premises IT infrastructure to AWS Cloud. The company has multiple long-term server bound licenses across the application stack and the CTO wants to continue to utilize those licenses while moving to AWS.

As a solutions architect, which of the following would you recommend as the MOST cost-effective solution?

Your answer is incorrect

Use Amazon EC2 on-demand instances

Correct answer

Use Amazon EC2 dedicated hosts

Use Amazon EC2 dedicated instances

Use Amazon EC2 reserved instances (RI)

Overall explanation

Correct option:

Use Amazon EC2 dedicated hosts

You can use Dedicated Hosts to launch Amazon EC2 instances on physical servers that are dedicated for your use. Dedicated Hosts give you additional visibility and control over how instances are placed on a physical server, and you can reliably use the same physical server over time. As a result, Dedicated Hosts enable you to use your existing server-bound software licenses like Windows Server and address corporate compliance and regulatory requirements.

Incorrect options:

Use Amazon EC2 dedicated instances - Dedicated instances are Amazon EC2 instances that run in a VPC on hardware that's dedicated to a single customer. Your dedicated instances are physically isolated at the host hardware level from instances that belong to other AWS accounts. Dedicated instances may share hardware with other instances from the same AWS account that are not dedicated instances. Dedicated instances cannot be used for existing server-bound software licenses.

Use Amazon EC2 on-demand instances

Use Amazon EC2 reserved instances (RI)

Amazon EC2 presents a virtual computing environment, allowing you to use web service interfaces to launch instances with a variety of operating systems, load them with your custom application environment, manage your network's access permissions, and run your image using as many or few systems as you desire.

Amazon EC2 provides the following purchasing options to enable you to optimize your costs based on your needs:

On-Demand Instances – Pay, by the second, for the instances that you launch.

Reserved Instances (RI) – Reduce your Amazon EC2 costs by making a commitment to a consistent instance configuration, including instance type and Region, for a term of 1 or 3 years.

Neither on-demand instances nor reserved instances can be used for existing server-bound software licenses.

References:

https://aws.amazon.com/ec2/dedicated-hosts/

https://aws.amazon.com/ec2/dedicated-hosts/faqs/

https://aws.amazon.com/ec2/pricing/dedicated-instances/

Domain

Design Cost-Optimized Architectures

Question 53 **Correct**

An Electronic Design Automation (EDA) application produces massive volumes of data that can be divided into two categories. The 'hot data' needs to be both processed and stored quickly in a parallel and distributed fashion. The 'cold data' needs to be kept for reference with quick access for reads and updates at a low cost.

Which of the following AWS services is BEST suited to accelerate the aforementioned chip design process?

Your answer is correct

Amazon FSx for Lustre

Amazon EMR

AWS Glue

Amazon FSx for Windows File Server

Overall explanation

Correct option:

Amazon FSx for Lustre

Amazon FSx for Lustre makes it easy and cost-effective to launch and run the world's most popular high-performance file system. It is used for workloads such as machine learning, high-performance computing (HPC), video processing, and financial modeling. The open-source Lustre file system is designed for applications that require fast storage – where you want your storage to keep up with your compute. FSx for Lustre integrates with Amazon S3, making it easy to process data sets with the Lustre file system. When linked to an S3 bucket, an FSx for Lustre file system transparently presents S3 objects as files and allows you to write changed data back to S3.

FSx for Lustre provides the ability to both process the 'hot data' in a parallel and distributed fashion as well as easily store the 'cold data' on Amazon S3. Therefore this option is the BEST fit for the given problem statement.

Incorrect options:

Amazon FSx for Windows File Server - Amazon FSx for Windows File Server provides fully managed, highly reliable file storage that is accessible over the industry-standard Service Message Block (SMB) protocol. It is built on Windows Server, delivering a wide range of administrative features such as user quotas, end-user file restore, and Microsoft Active Directory (AD) integration. FSx for Windows does not allow you to present S3 objects as files and does not allow you to write changed data back to S3. Therefore you cannot reference the "cold data" with quick access for reads and updates at low cost. Hence this option is not correct.

Amazon EMR - Amazon EMR is the industry-leading cloud big data platform for processing vast amounts of data using open source tools such as Apache Spark, Apache Hive, Apache HBase, Apache Flink, Apache Hudi, and Presto. Amazon EMR uses Hadoop, an open-source framework,

to distribute your data and processing across a resizable cluster of Amazon EC2 instances. EMR does not offer the same storage and processing speed as FSx for Lustre. So it is not the right fit for the given high-performance workflow scenario.

AWS Glue - AWS Glue is a fully managed extract, transform, and load (ETL) service that makes it easy for customers to prepare and load their data for analytics. AWS Glue job is meant to be used for batch ETL data processing. AWS Glue does not offer the same storage and processing speed as FSx for Lustre. So it is not the right fit for the given high-performance workflow scenario.

References:

https://aws.amazon.com/fsx/lustre/

https://aws.amazon.com/fsx/windows/faqs/

Domain

Design High-Performing Architectures

Question 54 **Correct**

A social photo-sharing web application is hosted on Amazon Elastic Compute Cloud (Amazon EC2) instances behind an Elastic Load Balancer. The app gives the users the ability to upload their photos and also shows a leaderboard on the homepage of the app. The uploaded photos are stored in Amazon Simple Storage Service (Amazon S3) and the leaderboard data is maintained in Amazon DynamoDB. The Amazon EC2 instances need to access both Amazon S3 and Amazon DynamoDB for these features.

As a solutions architect, which of the following solutions would you recommend as the MOST secure option?

Encrypt the AWS credentials via a custom encryption library and save it in a secret directory on the Amazon EC2 instances. The application code can then safely decrypt the AWS credentials to make the API calls to Amazon S3 and Amazon DynamoDB

Save the AWS credentials (access key Id and secret access token) in a configuration file within the application code on the Amazon EC2 instances. Amazon EC2 instances can use these credentials to access Amazon S3 and Amazon DynamoDB

Your answer is correct

Attach the appropriate IAM role to the Amazon EC2 instance profile so that the instance can access Amazon S3 and Amazon DynamoDB

Configure AWS CLI on the Amazon EC2 instances using a valid IAM user's credentials. The application code can then invoke shell scripts to access Amazon S3 and Amazon DynamoDB via AWS CLI

Overall explanation

Correct option:

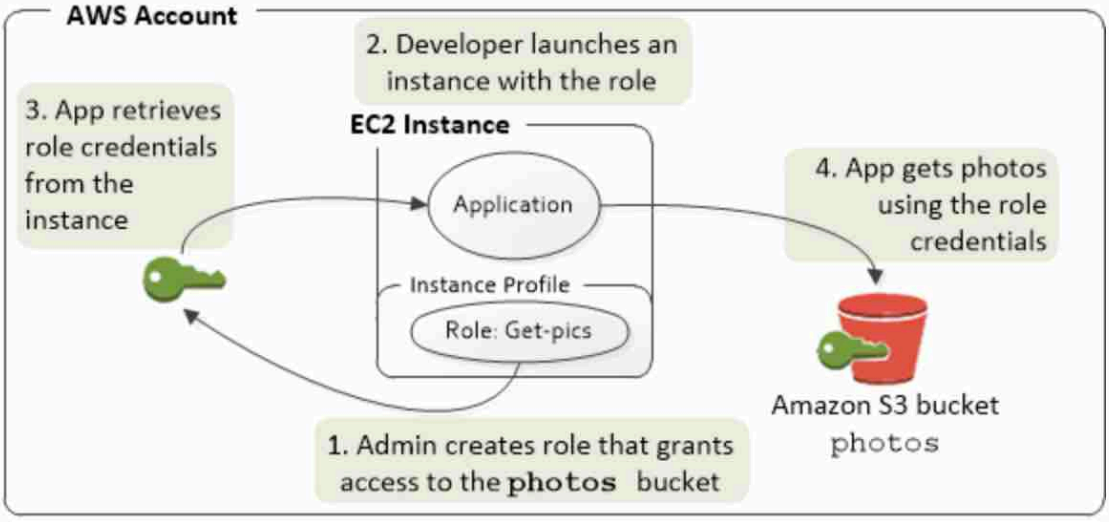
Attach the appropriate IAM role to the Amazon EC2 instance profile so that the instance can access Amazon S3 and Amazon DynamoDB

Applications that run on an Amazon EC2 instance must include AWS credentials in their AWS API requests. You could have your developers store AWS credentials directly within the Amazon EC2 instance and allow applications in that instance to use those credentials. But developers would then have to manage the credentials and ensure that they securely pass the credentials to each instance and update each Amazon EC2 instance when it's time to rotate the credentials.

Instead, you should use an IAM role to manage temporary credentials for applications that run on an Amazon EC2 instance. When you use a role, you don't have to distribute long-term credentials (such as a username and password or access keys) to an Amazon EC2 instance. The role supplies temporary permissions that applications can use when they make calls to other AWS resources. When you launch an Amazon EC2 instance, you specify an IAM role to associate with the instance. Applications that run on the instance can then use the role-supplied temporary credentials to sign API requests. Therefore, this option is correct.

How Do Roles for EC2 Instances Work?

In the following figure, a developer runs an application on an EC2 instance that requires access to the S3 bucket named `photos`. An administrator creates the `Get-pics` service role and attaches the role to the EC2 instance. The role includes a permissions policy that grants read-only access to the specified S3 bucket. It also includes a trust policy that allows the EC2 instance to assume the role and retrieve the temporary credentials. When the application runs on the instance, it can use the role's temporary credentials to access the photos bucket. The administrator doesn't have to grant the developer permission to access the photos bucket, and the developer never has to share or manage credentials.



via - https://docs.aws.amazon.com/IAM/latest/UserGuide/id_roles_use_switch-role-ec2.html

Incorrect options:

Save the AWS credentials (access key Id and secret access token) in a configuration file within the application code on the Amazon EC2 instances. Amazon EC2 instances can use these credentials to access Amazon S3 and Amazon DynamoDB

Configure AWS CLI on the Amazon EC2 instances using a valid IAM user's credentials. The application code can then invoke shell scripts to access Amazon S3 and Amazon DynamoDB via AWS CLI

Encrypt the AWS credentials via a custom encryption library and save it in a secret directory on the Amazon EC2 instances. The application code can then safely decrypt the AWS credentials to make the API calls to Amazon S3 and Amazon DynamoDB

Keeping the AWS credentials (encrypted or plain text) on the Amazon EC2 instance is a bad security practice, therefore these three options using the AWS credentials are incorrect.

Reference:

https://docs.aws.amazon.com/IAM/latest/UserGuide/id_roles_use_switch-role-ec2.html

Domain

Design Secure Architectures

Question 55 **Correct**

A company has a license-based, expensive, legacy commercial database solution deployed at its on-premises data center. The company wants to migrate this database to a more efficient, open-source, and cost-effective option on AWS Cloud. The CTO at the company wants a solution that can handle complex database configurations such as secondary indexes, foreign keys, and stored procedures.

As a solutions architect, which of the following AWS services should be combined to handle this use-case? (Select two)

Your selection is correct

AWS Schema Conversion Tool (AWS SCT)

AWS Snowball Edge

Your selection is correct

AWS Database Migration Service (AWS DMS)

Basic Schema Copy

AWS Glue

Overall explanation

Correct options:

AWS Schema Conversion Tool (AWS SCT)

AWS Database Migration Service (AWS DMS)

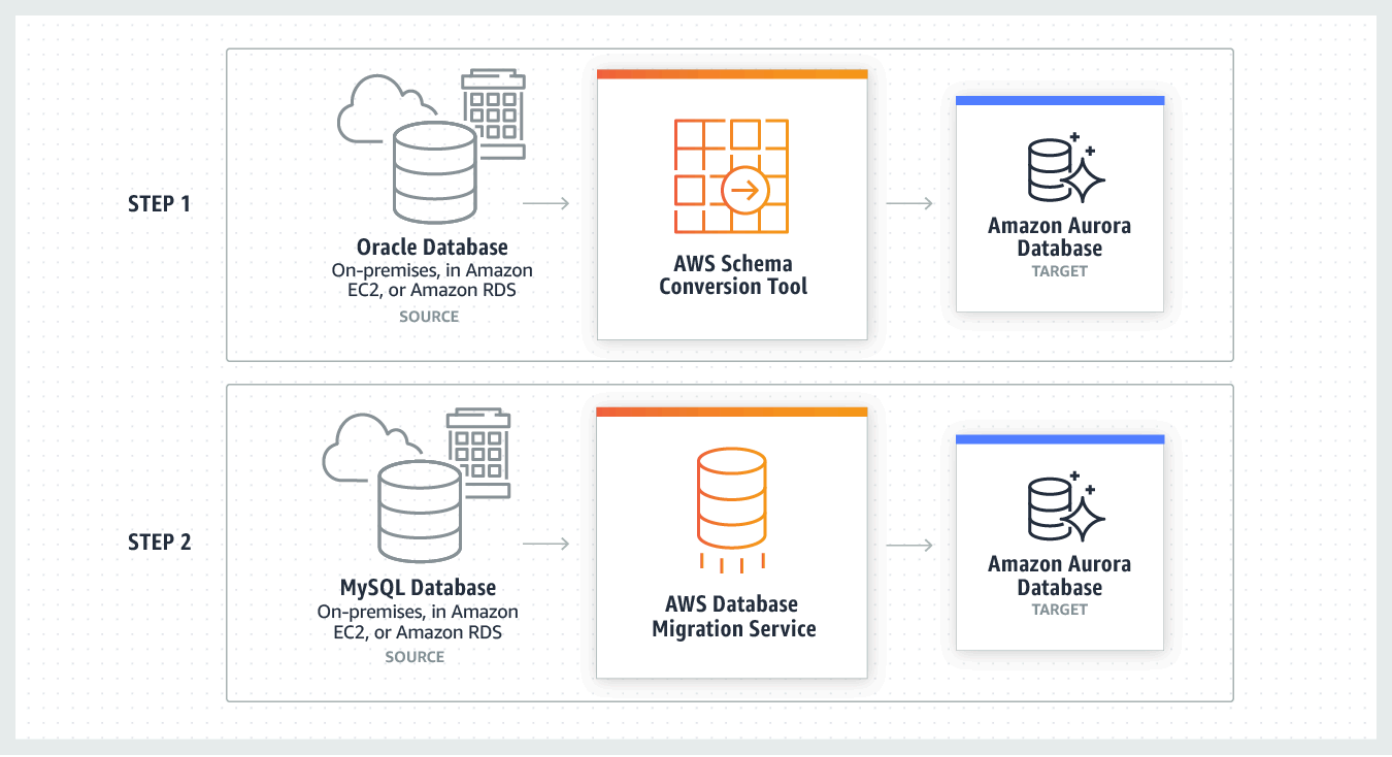
AWS Database Migration Service helps you migrate databases to AWS quickly and securely. The source database remains fully operational during the migration, minimizing downtime to applications that rely on the database. AWS Database Migration Service supports homogeneous migrations such as Oracle to Oracle, as well as heterogeneous migrations between different database platforms, such as Oracle or Microsoft SQL Server to Amazon Aurora.

Given the use-case where the CTO at the company wants to move away from license-based, expensive, legacy commercial database solutions deployed at the on-premises data center to more efficient, open-source, and cost-effective options on AWS Cloud, this is an example of heterogeneous database migrations.

For such a scenario, the source and target databases engines are different, like in the case of Oracle to Amazon Aurora, Oracle to PostgreSQL, or Microsoft SQL Server to MySQL migrations. In this case, the schema structure, data types, and database code of source and target databases can be quite different, requiring a schema and code transformation before the data migration starts.

That makes heterogeneous migrations a two-step process. First use the AWS Schema Conversion Tool to convert the source schema and code to match that of the target database, and then use the AWS Database Migration Service to migrate data from the source database to the target database. All the required data type conversions will automatically be done by the AWS Database Migration Service during the migration. The source database can be located on your on-premises environment outside of AWS, running on an Amazon EC2 instance, or it can be an Amazon RDS database. The target can be a database in Amazon EC2 or Amazon RDS.

Heterogeneous Database Migrations:



via - https://aws.amazon.com/dms/

Incorrect options:

AWS Snowball Edge - AWS Snowball Edge Storage Optimized is the optimal choice if you need to securely and quickly transfer dozens of terabytes to petabytes of data to AWS. It provides up to 80 TB of usable HDD storage, 40 vCPUs, 1 TB of SATA SSD storage, and up to 40 Gb network connectivity to address large scale data transfer and pre-processing use cases. As each Snowball Edge Storage Optimized device can handle 80TB of data, you can order 10 such devices to take care of the data transfer for all applications. The original Snowball devices were transitioned out of service and AWS Snowball Edge Storage Optimized are now the primary devices used for data transfer. You may see the Snowball device on the exam, just remember that the original Snowball device had 80TB of storage space. AWS Snowball Edge cannot be used for database migrations.

AWS Glue - AWS Glue is a fully managed extract, transform, and load (ETL) service that makes it easy for customers to prepare and load their data for analytics. AWS Glue job is meant to be used for batch ETL data processing. Therefore, it cannot be used for database migrations.

Basic Schema Copy - To quickly migrate a database schema to your target instance you can rely on the Basic Schema Copy feature of AWS Database Migration Service. Basic Schema Copy will automatically create tables and primary keys in the target instance if the target does not already contain tables with the same names. Basic Schema Copy is great for doing a test migration, or when you are migrating databases heterogeneously e.g. Oracle to MySQL or SQL Server to Oracle. Basic Schema Copy will not migrate secondary indexes, foreign keys or stored procedures. When you need to use a more customizable schema migration process (e.g. when you are migrating your production database and need to move your stored procedures and secondary database objects), you must use the AWS Schema Conversion Tool.

References:

https://aws.amazon.com/dms/

https://aws.amazon.com/dms/faqs/

https://aws.amazon.com/dms/schema-conversion-tool/

Domain

Design Cost-Optimized Architectures

Question 56 **Correct**

A mobile gaming company is experiencing heavy read traffic to its Amazon Relational Database Service (Amazon RDS) database that retrieves player's scores and stats. The company is using an Amazon RDS database instance type that is not cost-effective for their budget. The company would like to implement a strategy to deal with the high volume of read traffic, reduce latency, and also downsize the instance size to cut costs.

Which of the following solutions do you recommend?

Switch application code to AWS Lambda for better performance

Setup Amazon RDS Read Replicas

Move to Amazon Redshift

Your answer is correct

Setup Amazon ElastiCache in front of Amazon RDS

Overall explanation

Correct option:

Setup Amazon ElastiCache in front of Amazon RDS

Amazon ElastiCache is an ideal front-end for data stores such as Amazon RDS, providing a high-performance middle tier for applications with extremely high request rates and/or low latency requirements. The best part of caching is that it's minimally invasive to implement and by doing so, your application performance regarding both scale and speed is dramatically improved.

Incorrect options:

Setup Amazon RDS Read Replicas - Adding read replicas would further add to the database costs and will not help in reducing latency when compared to a caching solution. So this option is ruled out.

Move to Amazon Redshift - Amazon Redshift is optimized for datasets ranging from a few hundred gigabytes to a petabyte or more. If the company is looking at cost-cutting, moving to Amazon Redshift from Amazon RDS is not an option.

Switch application code to AWS Lambda for better performance - AWS Lambda can help in running data processing workflows. But, data still needs to be read from RDS and hence we need a solution to speed up the data reads and not before/after processing.

Reference:

https://aws.amazon.com/caching/database-caching/

Domain

Design Cost-Optimized Architectures

Question 57 Incorrect

An e-commerce application uses an Amazon Aurora Multi-AZ deployment for its database. While analyzing the performance metrics, the engineering team has found that the database reads are causing high input/output (I/O) and adding latency to the write requests against the database.

As an AWS Certified Solutions Architect Associate, what would you recommend to separate the read requests from the write requests?

Configure the application to read from the Multi-AZ standby instance

Correct answer

Set up a read replica and modify the application to use the appropriate endpoint

Provision another Amazon Aurora database and link it to the primary database as a read replica

Your answer is incorrect

Activate read-through caching on the Amazon Aurora database

Overall explanation

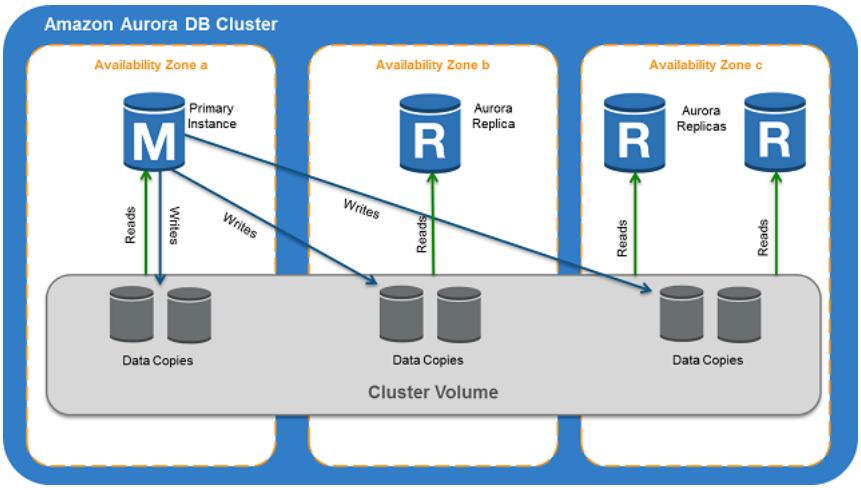
Correct option:

Set up a read replica and modify the application to use the appropriate endpoint

An Amazon Aurora DB cluster consists of one or more DB instances and a cluster volume that manages the data for those DB instances. An Aurora cluster volume is a virtual database storage volume that spans multiple Availability Zones (AZs), with each Availability Zone (AZ) having a copy of the DB cluster data. Two types of DB instances make up an Aurora DB cluster:

Primary DB instance – Supports read and write operations, and performs all of the data modifications to the cluster volume. Each Aurora DB cluster has one primary DB instance.

Aurora Replica – Connects to the same storage volume as the primary DB instance and supports only read operations. Each Aurora DB cluster can have up to 15 Aurora Replicas in addition to the primary DB instance. Aurora automatically fails over to an Aurora Replica in case the primary DB instance becomes unavailable. You can specify the failover priority for Aurora Replicas. Aurora Replicas can also offload read workloads from the primary DB instance.

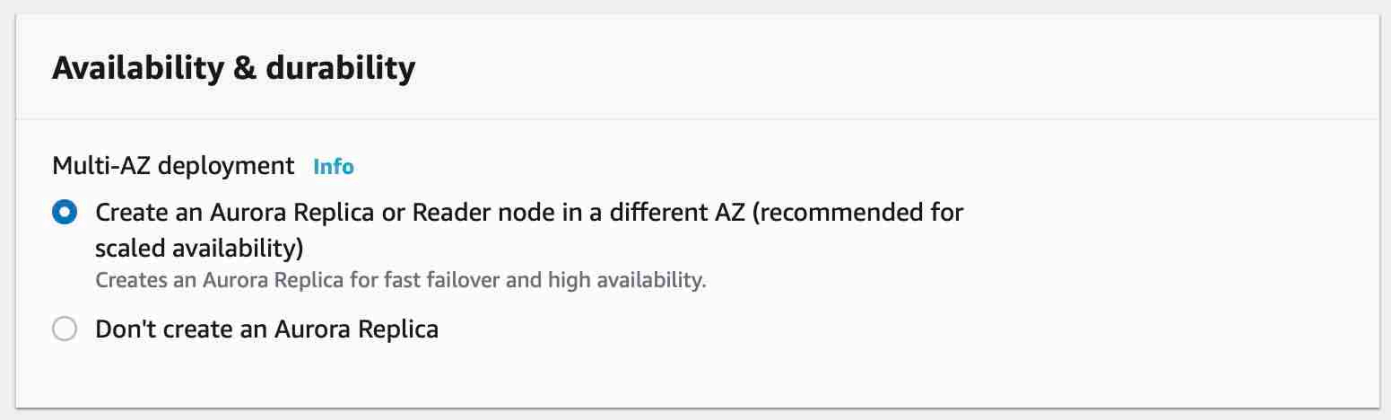


via - https://docs.aws.amazon.com/AmazonRDS/latest/AuroraUserGuide/Aurora.Overview.html

Aurora Replicas have two main purposes. You can issue queries to them to scale the read operations for your application. You typically do so by connecting to the reader endpoint of the cluster. That way, Aurora can spread the load for read-only connections across as many Aurora Replicas as you have in the cluster. Aurora Replicas also help to increase availability. If the writer instance in a cluster becomes unavailable, Aurora automatically promotes one of the reader instances to take its place as the new writer.

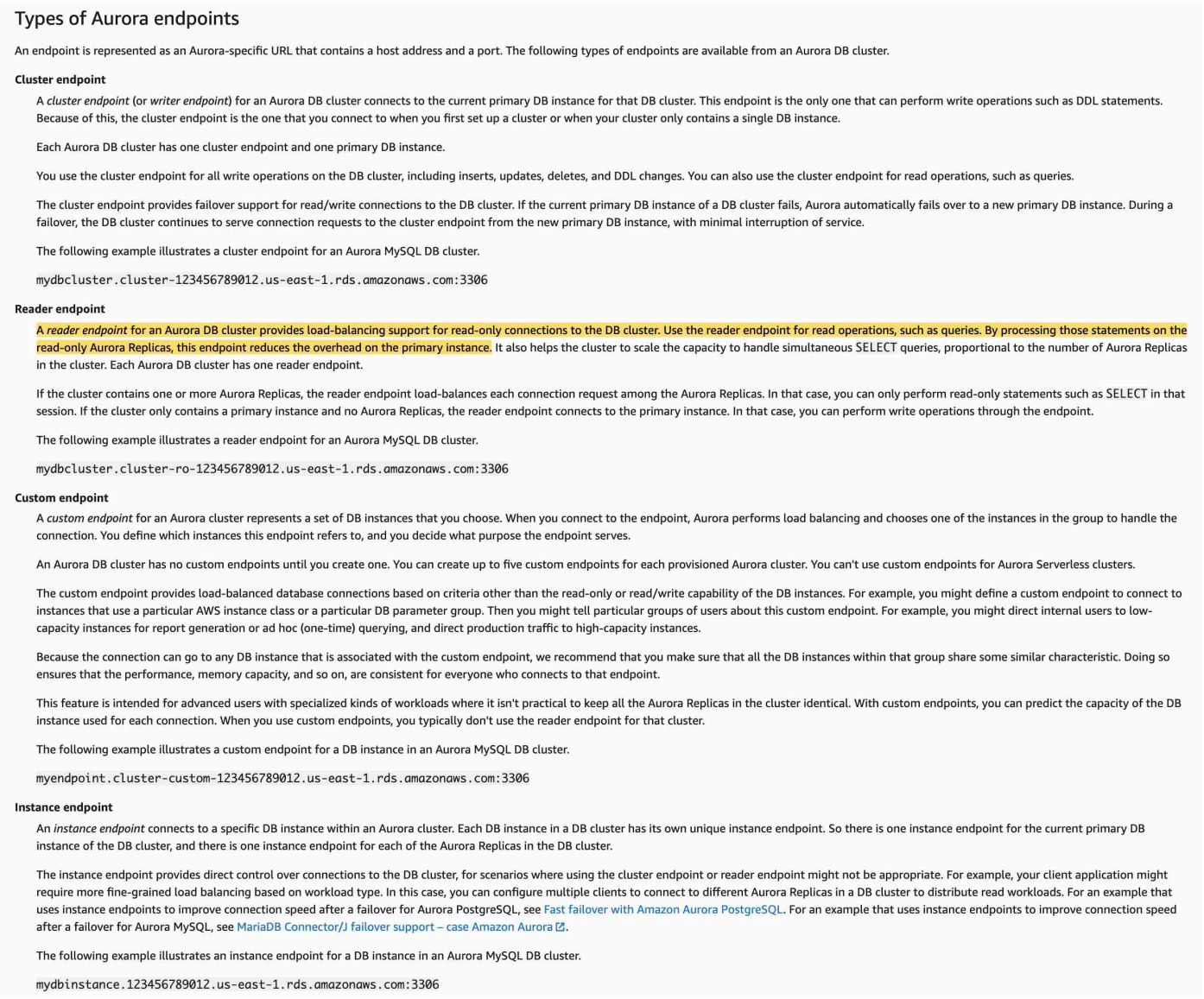
While setting up a Multi-AZ deployment for Aurora, you create an Aurora replica or reader node in a different Availability Zone (AZ).

Multi-AZ for Aurora:



You use the reader endpoint for read-only connections for your Aurora cluster. This endpoint uses a load-balancing mechanism to help your cluster handle a query-intensive workload. The reader endpoint is the endpoint that you supply to applications that do reporting or other read-

only operations on the cluster. The reader endpoint load-balances connections to available Aurora Replicas in an Aurora DB cluster.



Types of Aurora endpoints

An endpoint is represented as an Aurora-specific URL that contains a host address and a port. The following types of endpoints are available from an Aurora DB cluster.

Cluster endpoint

A *cluster endpoint* (or *writer endpoint*) for an Aurora DB cluster connects to the current primary DB instance for that DB cluster. This endpoint is the only one that can perform write operations such as DDL statements. Because of this, the cluster endpoint is the one that you connect to when you first set up a cluster or when your cluster only contains a single DB instance.

Each Aurora DB cluster has one cluster endpoint and one primary DB instance.

You use the cluster endpoint for all write operations on the DB cluster, including inserts, updates, deletes, and DDL changes. You can also use the cluster endpoint for read operations, such as queries.

The cluster endpoint provides failover support for read/write connections to the DB cluster. If the current primary DB instance of a DB cluster fails, Aurora automatically fails over to a new primary DB instance. During a failover, the DB cluster continues to serve connection requests to the cluster endpoint from the new primary DB instance, with minimal interruption of service.

The following example illustrates a cluster endpoint for an Aurora MySQL DB cluster.

`mydbcluster.cluster-123456789012.us-east-1.rds.amazonaws.com:3306`

Reader endpoint

A *reader endpoint* for an Aurora DB cluster provides load-balancing support for read-only connections to the DB cluster. Use the reader endpoint for read operations, such as queries. By processing those statements on the read-only Aurora Replicas, this endpoint reduces the overhead on the primary instance. It also helps the cluster to scale the capacity to handle simultaneous `SELECT` queries, proportional to the number of Aurora Replicas in the cluster. Each Aurora DB cluster has one reader endpoint.

If the cluster contains one or more Aurora Replicas, the reader endpoint load-balances each connection request among the Aurora Replicas. In that case, you can only perform read-only statements such as `SELECT` in that session. If the cluster only contains a primary instance and no Aurora Replicas, the reader endpoint connects to the primary instance. In that case, you can perform write operations through the endpoint.

The following example illustrates a reader endpoint for an Aurora MySQL DB cluster.

`mydbcluster.cluster-ro-123456789012.us-east-1.rds.amazonaws.com:3306`

Custom endpoint

A *custom endpoint* for an Aurora cluster represents a set of DB instances that you choose. When you connect to the endpoint, Aurora performs load balancing and chooses one of the instances in the group to handle the connection. You define which instances this endpoint refers to, and you decide what purpose the endpoint serves.

An Aurora DB cluster has no custom endpoints until you create one. You can create up to five custom endpoints for each provisioned Aurora cluster. You can't use custom endpoints for Aurora Serverless clusters.

The custom endpoint provides load-balanced database connections based on criteria other than the read-only or read/write capability of the DB instances. For example, you might define a custom endpoint to connect to instances that use a particular AWS instance class or a particular DB parameter group. Then you might tell particular groups of users about this custom endpoint. For example, you might direct internal users to low-capacity instances for report generation or ad hoc (one-time) querying, and direct production traffic to high-capacity instances.

Because the connection can go to any DB instance that is associated with the custom endpoint, we recommend that you make sure that all the DB instances within that group share some similar characteristic. Doing so ensures that the performance, memory capacity, and so on, are consistent for everyone who connects to that endpoint.

This feature is intended for advanced users with specialized kinds of workloads where it isn't practical to keep all the Aurora Replicas in the cluster identical. With custom endpoints, you can predict the capacity of the DB instance used for each connection. When you use custom endpoints, you typically don't use the reader endpoint for that cluster.

The following example illustrates a custom endpoint for a DB instance in an Aurora MySQL DB cluster.

`myendpoint.cluster-custom-123456789012.us-east-1.rds.amazonaws.com:3306`

Instance endpoint

An *instance endpoint* connects to a specific DB instance within an Aurora cluster. Each DB instance in a DB cluster has its own unique instance endpoint. So there is one instance endpoint for the current primary DB instance of the DB cluster, and there is one instance endpoint for each of the Aurora Replicas in the DB cluster.

The instance endpoint provides direct control over connections to the DB cluster, for scenarios where using the cluster endpoint or reader endpoint might not be appropriate. For example, your client application might require more fine-grained load balancing based on workload type. In this case, you can configure multiple clients to connect to different Aurora Replicas in a DB cluster to distribute read workloads. For an example that uses instance endpoints to improve connection speed after a failover for Aurora PostgreSQL, see Fast failover with Amazon Aurora PostgreSQL. For an example that uses instance endpoints to improve connection speed after a failover for Aurora MySQL, see MariaDB Connector/J failover support – case Amazon Aurora.

The following example illustrates an instance endpoint for a DB instance in an Aurora MySQL DB cluster.

`mydbinstance.123456789012.us-east-1.rds.amazonaws.com:3306`

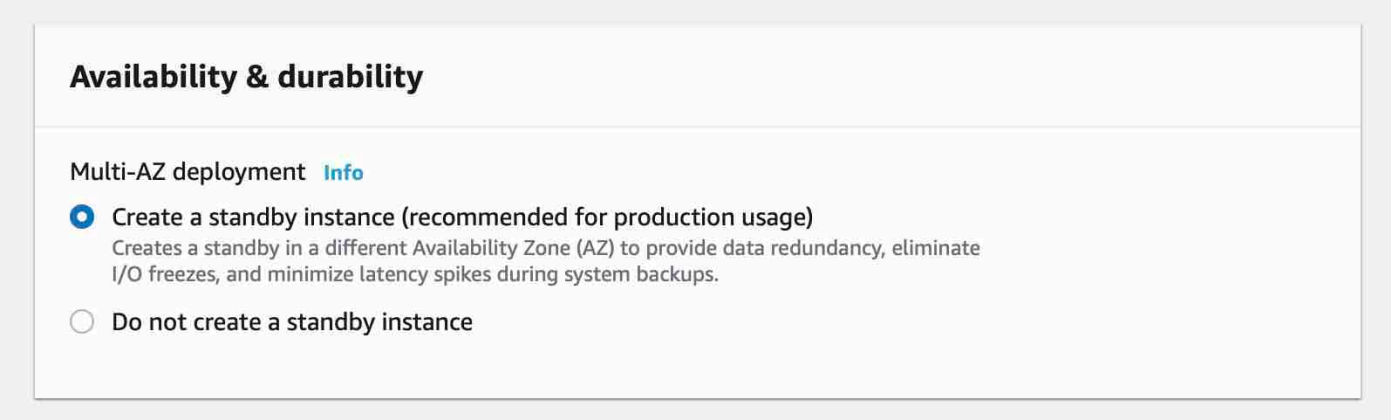
via - https://docs.aws.amazon.com/AmazonRDS/latest/AuroraUserGuide/Aurora.Overview.Endpoints.html

Incorrect options:

Provision another Amazon Aurora database and link it to the primary database as a read replica - You cannot provision another Aurora database and then link it as a read-replica for the primary database. This option is ruled out.

Configure the application to read from the Multi-AZ standby instance - This option has been added as a distractor as Aurora does not have any entity called standby instance. You create a standby instance while setting up a Multi-AZ deployment for Amazon RDS and NOT for Aurora.

Multi-AZ for Amazon RDS:



Activate read-through caching on the Amazon Aurora database - Amazon Aurora does not have built-in support for read-through caching, so this option just serves as a distractor. To implement caching, you will need to integrate something like Amazon ElastiCache and that would need code changes for the application.

References:

https://docs.aws.amazon.com/AmazonRDS/latest/AuroraUserGuide/Aurora.Overview.html

https://docs.aws.amazon.com/AmazonRDS/latest/AuroraUserGuide/Concepts.AuroraHighAvailability.htm

https://docs.aws.amazon.com/AmazonRDS/latest/AuroraUserGuide/Aurora.Overview.Endpoints.html

Domain

Design Resilient Architectures

Question 58 Correct

A pharmaceutical company is considering moving to AWS Cloud to accelerate the research and development process. Most of the daily workflows would be centered around running batch jobs on Amazon EC2 instances with storage on Amazon Elastic Block Store (Amazon EBS) volumes. The CTO is concerned about meeting HIPAA compliance norms for sensitive data stored on Amazon EBS.

Which of the following options outline the correct capabilities of an encrypted Amazon EBS volume? (Select three)

Your selection is correct

Data at rest inside the volume is encrypted

Your selection is correct

Data moving between the volume and the instance is encrypted

Data at rest inside the volume is NOT encrypted

Data moving between the volume and the instance is NOT encrypted

Your selection is correct

Any snapshot created from the volume is encrypted

Any snapshot created from the volume is NOT encrypted

Overall explanation

Correct options:

Data at rest inside the volume is encrypted

Any snapshot created from the volume is encrypted

Data moving between the volume and the instance is encrypted

Amazon Elastic Block Store (Amazon EBS) provides block-level storage volumes for use with Amazon EC2 instances. When you create an encrypted Amazon EBS volume and attach it to a supported instance type, data stored at rest on the volume, data moving between the volume and the instance, snapshots created from the volume and volumes created from those snapshots are all encrypted. It uses AWS Key Management Service (AWS KMS) customer master keys (CMK) when creating encrypted volumes and snapshots. Encryption operations occur on the servers that host Amazon EC2 instances, ensuring the security of both data-at-rest and data-in-transit between an instance and its attached Amazon EBS storage.

Therefore, the incorrect options are:

Data moving between the volume and the instance is NOT encrypted

Any snapshot created from the volume is NOT encrypted

Data at rest inside the volume is NOT encrypted

Reference:

https://docs.aws.amazon.com/AWSEC2/latest/UserGuide/EBSEncryption.html

Domain

Design Secure Architectures

Question 59 **Correct**

A development team has deployed a microservice to the Amazon Elastic Container Service (Amazon ECS). The application layer is in a Docker container that provides both static and dynamic content through an Application Load Balancer. With increasing load, the Amazon ECS cluster is experiencing higher network usage. The development team has looked into the network usage and found that 90% of it is due to distributing static content of the application.

As a Solutions Architect, what do you recommend to improve the application's network usage and decrease costs?

Distribute the dynamic content through Amazon S3

Distribute the static content through Amazon EFS

Distribute the dynamic content through Amazon EFS

Your answer is correct

Distribute the static content through Amazon S3

Overall explanation

Correct option:

Distribute the static content through Amazon S3

You can use Amazon S3 to host a static website. On a static website, individual web pages include static content. They might also contain client-side scripts. To host a static website on Amazon S3, you configure an Amazon S3 bucket for website hosting and then upload your website content to the bucket. When you configure a bucket as a static website, you must enable website hosting, set permissions, and create and add an index document. Depending on your website requirements, you can also configure redirects, web traffic logging, and a custom error document.

Distributing the static content through Amazon S3 allows us to offload most of the network usage to Amazon S3 and free up our applications running on Amazon ECS.

Incorrect options:

Distribute the dynamic content through Amazon S3 - By contrast, a dynamic website relies on server-side processing, including server-side scripts such as PHP, JSP, or ASP.NET. Amazon S3 does not support server-side scripting, but AWS has other resources for hosting dynamic websites.

Distribute the static content through Amazon EFS

Distribute the dynamic content through Amazon EFS

Amazon Elastic File System (Amazon EFS) provides a simple, scalable, fully managed elastic NFS file system for use with AWS Cloud services and on-premises resources. Using Amazon EFS for static or dynamic content will not change anything as static content on EFS would still have to be distributed by the Amazon ECS instances.

References:

https://docs.aws.amazon.com/AmazonS3/latest/dev/WebsiteHosting.html

Domain

Design Cost-Optimized Architectures

Question 60 Incorrect

A media company wants a low-latency way to distribute live sports results which are delivered via a proprietary application using UDP protocol.

As a solutions architect, which of the following solutions would you recommend such that it offers the BEST performance for this use case?

Your answer is incorrect

Use Elastic Load Balancing (ELB) to provide a low latency way to distribute live sports results

Correct answer

Use AWS Global Accelerator to provide a low latency way to distribute live sports results

Use Amazon CloudFront to provide a low latency way to distribute live sports results

Use Auto Scaling group to provide a low latency way to distribute live sports results

Overall explanation

Correct option:

Use AWS Global Accelerator to provide a low latency way to distribute live sports results

AWS Global Accelerator is a networking service that helps you improve the availability and performance of the applications that you offer to your global users. AWS Global Accelerator is easy to set up, configure, and manage. It provides static IP addresses that provide a fixed entry point to your applications and eliminate the complexity of managing specific IP addresses for different AWS Regions and Availability Zones (AZs). AWS Global Accelerator always routes user traffic to the optimal endpoint based on performance, reacting instantly to changes in application health, your user's location, and policies that you configure. Global Accelerator is a good fit for non-HTTP use cases, such as gaming (UDP), IoT (MQTT), or Voice over IP. Therefore, this option is correct.

How AWS Global Accelerator Works:

via - https://aws.amazon.com/global-accelerator/

Incorrect options:

Use Amazon CloudFront to provide a low latency way to distribute live sports results - Amazon CloudFront is a fast content delivery network (CDN) service that securely delivers data, videos, applications, and APIs to customers globally with low latency, high transfer speeds, all within a developer-friendly environment.

Amazon CloudFront points of presence (POPs) (edge locations) make sure that popular content can be served quickly to your viewers. Amazon CloudFront also has regional edge caches that bring more of your content closer to your viewers, even when the content is not popular enough to stay at a POP, to help improve performance for that content. Regional edge caches help with all types of content, particularly content that tends to become less popular over time. Examples include user-generated content, such as video, photos, or artwork; e-commerce assets such as product photos and videos; and news and event-related content that might suddenly find new popularity. CloudFront supports HTTP/RTMP protocol based requests, therefore this option is incorrect.

Use Elastic Load Balancing (ELB) to provide a low latency way to distribute live sports results - Elastic Load Balancing (ELB) automatically distributes incoming application traffic across multiple targets, such as Amazon EC2 instances, containers, IP addresses, and AWS Lambda functions. It can handle the varying load of your application traffic in a single Availability Zone or across multiple Availability Zones. Elastic Load Balancer cannot help with decreasing latency of incoming traffic from the source.

Use Auto Scaling group to provide a low latency way to distribute live sports results - Amazon EC2 Auto Scaling helps you ensure that you have the correct number of Amazon EC2 instances available to handle the load for your application. You create collections of Amazon EC2 instances, called Auto Scaling groups. You can specify the minimum number of instances in each Auto Scaling group, and Amazon EC2 Auto Scaling ensures that your group never goes below this size. Auto Scaling group cannot help with decreasing latency of incoming traffic from the source.

Exam Alert:

Please note the differences between the capabilities of AWS Global Accelerator and Amazon CloudFront -

AWS Global Accelerator and Amazon CloudFront are separate services that use the AWS global network and its edge locations around the world. Amazon CloudFront improves performance for both cacheable content (such as images and videos) and dynamic content (such as API acceleration and dynamic site delivery). AWS Global Accelerator improves performance for a wide range of applications over TCP or UDP by proxying packets at the edge to applications running in one or more AWS Regions.

AWS Global Accelerator is a good fit for non-HTTP use cases, such as gaming (UDP), IoT (MQTT), or Voice over IP, as well as for HTTP use cases that specifically require static IP addresses or deterministic, fast regional failover. Both services integrate with AWS Shield for DDoS protection.

References:

https://aws.amazon.com/global-accelerator/

https://aws.amazon.com/cloudfront/faqs/

Domain

Design High-Performing Architectures

Question 61 Correct

An Elastic Load Balancer has marked all the Amazon EC2 instances in the target group as unhealthy. Surprisingly, when a developer enters the IP address of the Amazon EC2 instances in the web browser, he can access the website.

What could be the reason the instances are being marked as unhealthy? (Select two)

The Amazon Elastic Block Store (Amazon EBS) volumes have been improperly mounted

Your selection is correct

The route for the health check is misconfigured

Your selection is correct

The security group of the Amazon EC2 instance does not allow for traffic from the security group of the Application Load Balancer

You need to attach elastic IP address (EIP) to the Amazon EC2 instances

Your web-app has a runtime that is not supported by the Application Load Balancer

Overall explanation

Correct options:

The security group of the Amazon EC2 instance does not allow for traffic from the security group of the Application Load Balancer

The route for the health check is misconfigured

An Application Load Balancer periodically sends requests to its registered targets to test their status. These tests are called health checks.

Each load balancer node routes requests only to the healthy targets in the enabled Availability Zones (AZs) for the load balancer. Each load balancer node checks the health of each target, using the health check settings for the target groups with which the target is registered. If a target group contains only unhealthy registered targets, the load balancer nodes route requests across its unhealthy targets.

You must ensure that your load balancer can communicate with registered targets on both the listener port and the health check port. Whenever you add a listener to your load balancer or update the health check port for a target group used by the load balancer to route requests, you must verify that the security groups associated with the load balancer allow traffic on the new port in both directions.

Application Load Balancer Configuration for Security Groups and Health Check Routes:

Security groups for your Application Load Balancer

PDF | Kindle | RSS

You must ensure that your load balancer can communicate with registered targets on both the listener port and the health check port. Whenever you add a listener to your load balancer or update the health check port for a target group used by the load balancer to route requests, you must verify that the security groups associated with the load balancer allow traffic on the new port in both directions. If they do not, you can edit the rules for the currently associated security groups or associate different security groups with the load balancer.

Recommended rules

The following are the recommended rules for an Internet-facing load balancer.

Inbound		
Source	**Port Range**	**Comment**
0.0.0.0/0	*listener*	Allow all inbound traffic on the load balancer listener port
Outbound		
Destination	**Port Range**	**Comment**
instance security group	*instance listener*	Allow outbound traffic to instances on the instance listener port
instance security group	*health check*	Allow outbound traffic to instances on the health check port

The following are the recommended rules for an internal load balancer.

Inbound		
Source	**Port Range**	**Comment**
VPC CIDR	*listener*	Allow inbound traffic from the VPC CIDR on the load balancer listener port
Outbound		
Destination	**Port Range**	**Comment**
instance security group	*instance listener*	Allow outbound traffic to instances on the instance listener port
instance security group	*health check*	Allow outbound traffic to instances on the health check port

via - https://docs.aws.amazon.com/elasticloadbalancing/latest/application/load-balancer-update-security-groups.html

Incorrect options:

The Amazon Elastic Block Store (Amazon EBS) volumes have been improperly mounted - You can access the website using the IP address which means there is no issue with the Amazon EBS volumes. So this option is not correct.

Your web-app has a runtime that is not supported by the Application Load Balancer - There is no connection between a web app runtime and the application load balancer. This option has been added as a distractor.

You need to attach elastic IP address (EIP) to the Amazon EC2 instances - This option is a distractor as Elastic IPs do not need to be assigned to Amazon EC2 instances while using an Application Load Balancer.

References:

https://docs.aws.amazon.com/elasticloadbalancing/latest/application/load-balancer-update-security-groups.html

https://docs.aws.amazon.com/elasticloadbalancing/latest/application/target-group-health-checks.html

Domain

Design Secure Architectures

Question 62 **Correct**

A company has hired you as an AWS Certified Solutions Architect – Associate to help with redesigning a real-time data processor. The company wants to build custom applications that process and analyze the streaming data for its specialized needs.

Which solution will you recommend to address this use-case?

Use Amazon Kinesis Data Firehose to process the data streams as well as decouple the producers and consumers for the real-time data processor

Your answer is correct

Use Amazon Kinesis Data Streams to process the data streams as well as decouple the producers and consumers for the real-time data processor

Use Amazon Simple Notification Service (Amazon SNS) to process the data streams as well as decouple the producers and consumers for the real-time data processor

Use Amazon Simple Queue Service (Amazon SQS) to process the data streams as well as decouple the producers and consumers for the real-time data processor

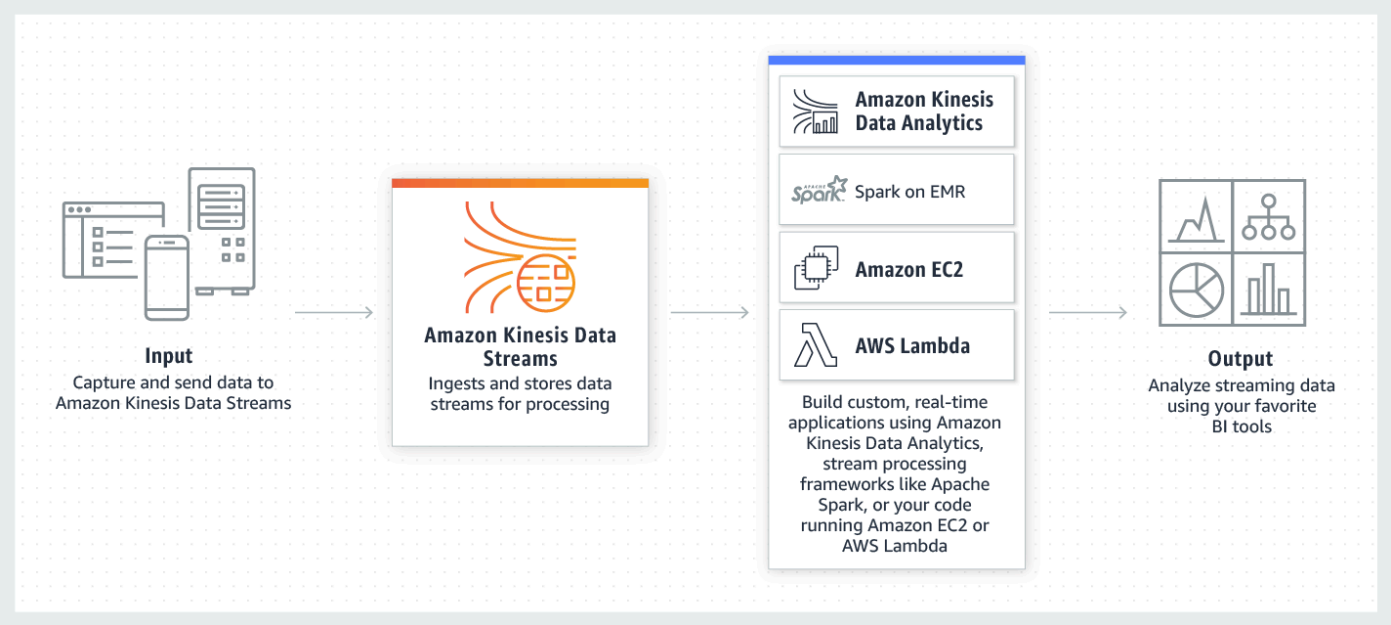
Overall explanation

Correct option:

Use Amazon Kinesis Data Streams to process the data streams as well as decouple the producers and consumers for the real-time data processor

Amazon Kinesis Data Streams is useful for rapidly moving data off data producers and then continuously processing the data, be it to transform the data before emitting to a data store, run real-time metrics and analytics, or derive more complex data streams for further processing. Kinesis data streams can continuously capture gigabytes of data per second from hundreds of thousands of sources such as website clickstreams, database event streams, financial transactions, social media feeds, IT logs, and location-tracking events.

Kinesis Data Streams Overview:



via - https://aws.amazon.com/kinesis/data-streams/

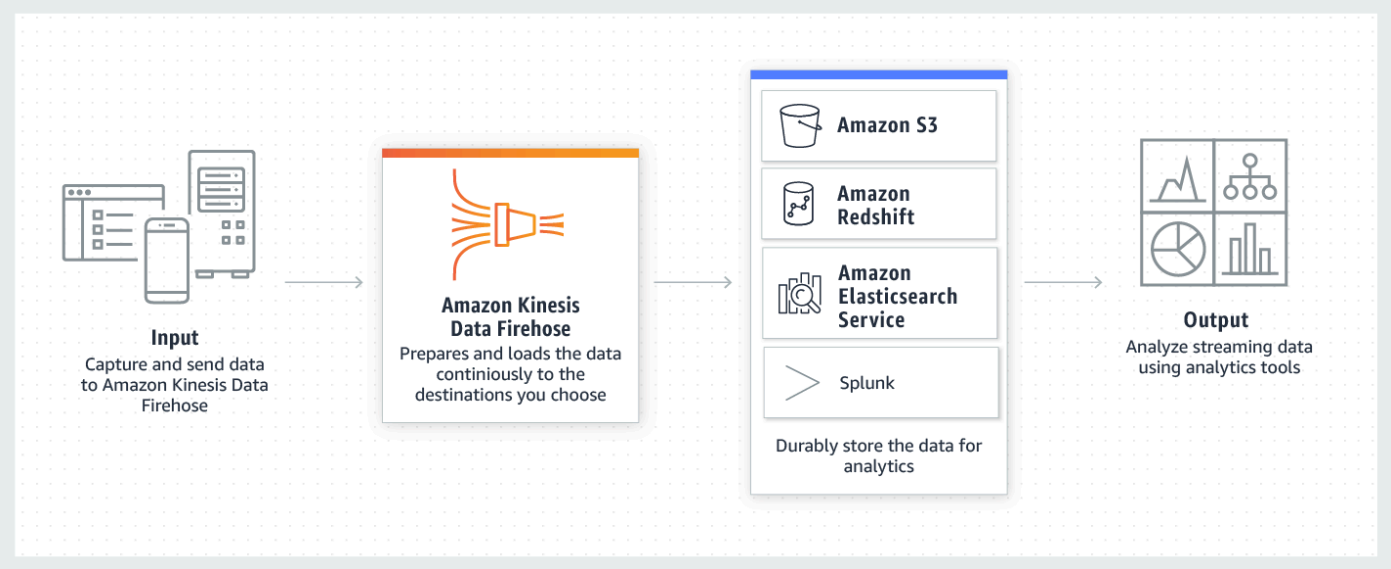
Incorrect options:

Use Amazon Simple Notification Service (Amazon SNS) to process the data streams as well as decouple the producers and consumers for the real-time data processor - Amazon Simple Notification Service (Amazon SNS) is a highly available, durable, secure, fully managed pub/sub messaging service that enables you to decouple microservices, distributed systems, and serverless applications. SNS cannot be used to decouple the producers and consumers for the real-time data processor as described in the given use-case.

Use Amazon Simple Queue Service (Amazon SQS) to process the data streams as well as decouple the producers and consumers for the real-time data processor - Amazon Simple Queue Service (Amazon SQS) offers a secure, durable, and available hosted queue that lets you integrate and decouple distributed software systems and components. SQS cannot be used to decouple the producers and consumers for the real-time data processor as described in the given use-case.

Use Amazon Kinesis Data Firehose to process the data streams as well as decouple the producers and consumers for the real-time data processor - Amazon Kinesis Data Firehose is the easiest way to load streaming data into data stores and analytics tools. Kinesis Firehose cannot be used to process and analyze the streaming data in custom applications. It can capture, transform, and load streaming data into Amazon S3, Amazon Redshift, Amazon Elasticsearch Service, and Splunk, enabling near real-time analytics.

Amazon Kinesis Data Firehose Overview



via - https://aws.amazon.com/kinesis/data-firehose/

References:

https://aws.amazon.com/kinesis/data-streams/

https://aws.amazon.com/kinesis/data-firehose/

Domain

Design Resilient Architectures

Question 63 Incorrect

Your company is deploying a website running on AWS Elastic Beanstalk. The website takes over 45 minutes for the installation and contains both static as well as dynamic files that must be generated during the installation process.

As a Solutions Architect, you would like to bring the time to create a new instance in your AWS Elastic Beanstalk deployment to be less than 2 minutes. Which of the following options should be combined to build a solution for this requirement? (Select two)

Your selection is incorrect

Use AWS Elastic Beanstalk deployment caching feature

Your selection is correct

Create a Golden Amazon Machine Image (AMI) with the static installation components already setup

Store the installation files in Amazon S3 so they can be quickly retrieved

Use Amazon EC2 user data to install the application at boot time

Correct selection

Use Amazon EC2 user data to customize the dynamic installation parts at boot time

Overall explanation

Correct options:

AWS Elastic Beanstalk is an easy-to-use service for deploying and scaling web applications and services developed with Java, .NET, PHP, Node.js, Python, Ruby, Go, and Docker on familiar servers such as Apache, Nginx, Passenger, and IIS.

You can simply upload your code and Elastic Beanstalk automatically handles the deployment, from capacity provisioning, load balancing, auto-scaling to application health monitoring. At the same time, you retain full control over the AWS resources powering your application and can access the underlying resources at any time.

When you create an AWS Elastic Beanstalk environment, you can specify an Amazon Machine Image (AMI) to use instead of the standard Elastic Beanstalk AMI included in your platform version. A custom AMI can improve provisioning times when instances are launched in your environment if you need to install a lot of software that isn't included in the standard AMIs.

Create a Golden Amazon Machine Image (AMI) with the static installation components already setup

A Golden AMI is an AMI that you standardize through configuration, consistent security patching, and hardening. It also contains agents you approve for logging, security, performance monitoring, etc. For the given use-case, you can have the static installation components already setup via the golden AMI.

Use Amazon EC2 user data to customize the dynamic installation parts at boot time

Amazon EC2 instance user data is the data that you specified in the form of a configuration script while launching your instance. You can use Amazon EC2 user data to customize the dynamic installation parts at boot time, rather than installing the application itself at boot time.

Incorrect options:

Store the installation files in Amazon S3 so they can be quickly retrieved - Amazon S3 bucket can be used as a storage location for your source code, logs, and other artifacts that are created when you use AWS Elastic Beanstalk. It cannot be used to run or generate dynamic files since Amazon S3 is not an environment but a storage service.

Use Amazon EC2 user data to install the application at boot time - User data of an instance can be used to perform common automated configuration tasks or run scripts after the instance starts. User data, cannot, however, be used to install the application since it takes over 45 minutes for the installation which contains static as well as dynamic files that must be generated during the installation process.

Use AWS Elastic Beanstalk deployment caching feature - AWS Elastic Beanstalk deployment caching is a made-up option. It is just added as a distractor.

References:

https://aws.amazon.com/elasticbeanstalk/

https://aws.amazon.com/blogs/awsmarketplace/announcing-the-golden-ami-pipeline/

https://docs.aws.amazon.com/elasticbeanstalk/latest/dg/AWSHowTo.S3.html

https://docs.aws.amazon.com/elasticbeanstalk/latest/dg/using-features.customenv.html

https://docs.aws.amazon.com/AWSEC2/latest/UserGuide/instancedata-add-user-data.html

Domain

Design Resilient Architectures

Question 64 Correct

A retail company has developed a REST API which is deployed in an Auto Scaling group behind an Application Load Balancer. The REST API stores the user data in Amazon DynamoDB and any static content, such as images, are served via Amazon Simple Storage Service (Amazon S3). On analyzing the usage trends, it is found that 90% of the read requests are for commonly accessed data across all users.

As a Solutions Architect, which of the following would you suggest as the MOST efficient solution to improve the application performance?

Enable ElastiCache Redis for DynamoDB and ElastiCache Memcached for Amazon S3

Your answer is correct

Enable Amazon DynamoDB Accelerator (DAX) for Amazon DynamoDB and Amazon CloudFront for Amazon S3

Enable Amazon DynamoDB Accelerator (DAX) for Amazon DynamoDB and ElastiCache Memcached for Amazon S3

Enable ElastiCache Redis for DynamoDB and Amazon CloudFront for Amazon S3

Overall explanation

Correct option:

Enable Amazon DynamoDB Accelerator (DAX) for Amazon DynamoDB and Amazon CloudFront for Amazon S3

Amazon DynamoDB Accelerator (DAX) is a fully managed, highly available, in-memory cache for Amazon DynamoDB that delivers up to a 10 times performance improvement—from milliseconds to microseconds—even at millions of requests per second.

Amazon DynamoDB Accelerator (DAX) is tightly integrated with Amazon DynamoDB—you simply provision a DAX cluster, use the DAX client SDK to point your existing Amazon DynamoDB API calls at the DAX cluster, and let DAX handle the rest. Because DAX is API-compatible with Amazon DynamoDB, you don't have to make any functional application code changes. DAX is used to natively cache Amazon DynamoDB reads.

Amazon CloudFront is a content delivery network (CDN) service that delivers static and dynamic web content, video streams, and APIs around the world, securely and at scale. By design, delivering data out of Amazon CloudFront can be more cost-effective than delivering it from S3 directly to your users.

When a user requests content that you serve with CloudFront, their request is routed to a nearby Edge Location. If CloudFront has a cached copy of the requested file, CloudFront delivers it to the user, providing a fast (low-latency) response. If the file they've requested isn't yet cached, CloudFront retrieves it from your origin – for example, the Amazon S3 bucket where you've stored your content.

So, you can use Amazon CloudFront to improve application performance to serve static content from Amazon S3.

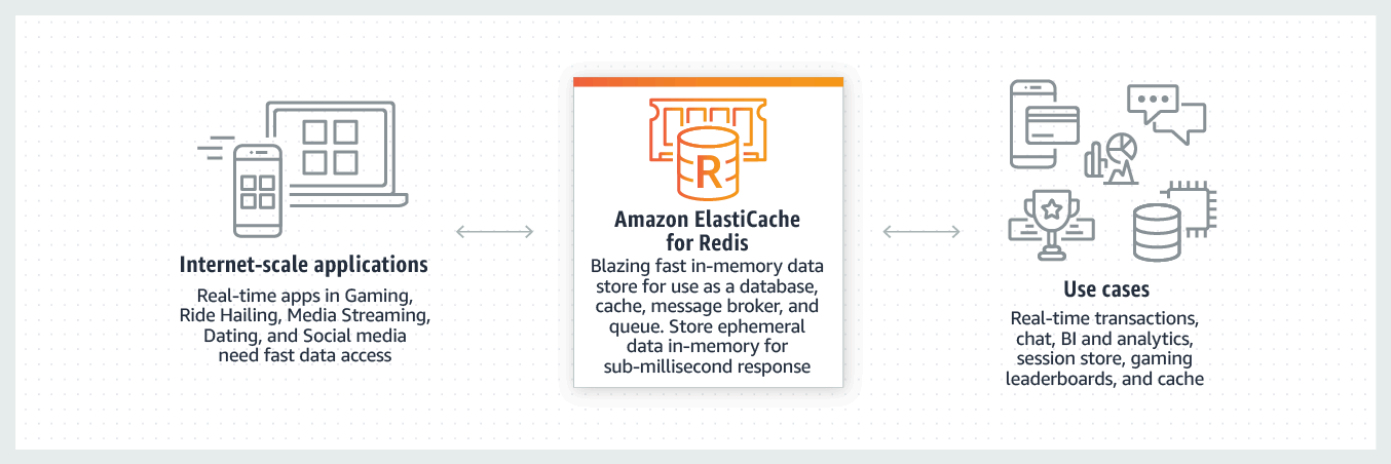
Incorrect options:

Enable ElastiCache Redis for DynamoDB and Amazon CloudFront for Amazon S3

Amazon ElastiCache for Redis is a blazing fast in-memory data store that provides sub-millisecond latency to power internet-scale real-time applications. Amazon ElastiCache for Redis is a great choice for real-time transactional and analytical processing use cases such as caching,

chat/messaging, gaming leaderboards, geospatial, machine learning, media streaming, queues, real-time analytics, and session store.

Amazon ElastiCache for Redis Overview:



via - https://aws.amazon.com/elasticache/redis/

Although you can integrate Redis with DynamoDB, it's much more involved than using DAX which is a much better fit.

Enable Amazon DynamoDB Accelerator (DAX) for Amazon DynamoDB and ElastiCache Memcached for Amazon S3

Enable ElastiCache Redis for DynamoDB and ElastiCache Memcached for Amazon S3

Amazon ElastiCache for Memcached is a Memcached-compatible in-memory key-value store service that can be used as a cache or a data store. Amazon ElastiCache for Memcached is a great choice for implementing an in-memory cache to decrease access latency, increase throughput, and ease the load off your relational or NoSQL database.

Amazon ElastiCache Memcached cannot be used as a cache to serve static content from Amazon S3, so both these options are incorrect.

References:

https://aws.amazon.com/dynamodb/dax/

https://aws.amazon.com/blogs/networking-and-content-delivery/amazon-s3-amazon-cloudfront-a-match-made-in-the-cloud/

https://aws.amazon.com/elasticache/redis/

Domain

Design Cost-Optimized Architectures

Question 65 Correct

A healthcare company uses its on-premises infrastructure to run legacy applications that require specialized customizations to the underlying Oracle database as well as its host operating system (OS). The company also wants to improve the availability of the Oracle database layer. The company has hired you as an AWS Certified Solutions Architect – Associate to build a solution on AWS that meets these requirements while minimizing the underlying infrastructure maintenance effort.

Which of the following options represents the best solution for this use case?

Deploy the Oracle database layer on multiple Amazon EC2 instances spread across two Availability Zones (AZs). This deployment configuration guarantees high availability and also allows the Database Administrator (DBA) to access and customize the database environment and the underlying operating system

Leverage cross AZ read-replica configuration of Amazon RDS for Oracle that allows the Database Administrator (DBA) to access and customize the database environment and the underlying operating system

Leverage multi-AZ configuration of Amazon RDS for Oracle that allows the Database Administrator (DBA) to access and customize the database environment and the underlying operating system

Your answer is correct

Leverage multi-AZ configuration of Amazon RDS Custom for Oracle that allows the Database Administrator (DBA) to access and customize the database environment and the underlying operating system

Overall explanation

Correct option:

Leverage multi-AZ configuration of Amazon RDS Custom for Oracle that allows the Database Administrator (DBA) to access and customize the database environment and the

underlying operating system

Amazon RDS is a managed service that makes it easy to set up, operate, and scale a relational database in the cloud. It provides cost-efficient and resizable capacity while managing time-consuming database administration tasks. Amazon RDS can automatically back up your database and keep your database software up to date with the latest version. However, RDS does not allow you to access the host OS of the database.

For the given use-case, you need to use Amazon RDS Custom for Oracle as it allows you to access and customize your database server host and operating system, for example by applying special patches and changing the database software settings to support third-party applications that require privileged access. Amazon RDS Custom for Oracle facilitates these functionalities with minimum infrastructure maintenance effort. You need to set up the RDS Custom for Oracle in multi-AZ configuration for high availability.

Amazon RDS Custom for Oracle:

Amazon RDS Custom for Oracle – New Control Capabilities in Database Environment

by Channy Yun | on 26 OCT 2021 | in Amazon RDS, Database, Launch, News, RDS For Oracle | Permalink | Share

▶ 0:00 / 0:00

Voiced by Amazon Polly

Managing databases in self-managed environments such as on premises or Amazon Elastic Compute Cloud (Amazon EC2) requires customers to spend time and resources doing database administration tasks such as provisioning, scaling, patching, backups, and configuring for high availability. So, hundreds of thousands of AWS customers use Amazon Relational Database Service (Amazon RDS) because it automates these undifferentiated administration tasks.

However, there are some legacy and packaged applications that require customers to make specialized customizations to the underlying database and the operating system (OS), such as Oracle industry specialized applications for healthcare and life sciences, telecom, retail, banking, and hospitality. Customers with these specific customization requirements cannot get the benefits of a fully managed database service like Amazon RDS, and they end up deploying their databases on premises or on EC2 instances.

Today, I am happy to announce the general availability of Amazon RDS Custom for Oracle, new capabilities that enable database administrators to access and customize the database environment and operating system. With RDS Custom for Oracle, you can now access and customize your database server host and operating system, for example by applying special patches and changing the database software settings to support third-party applications that require privileged access.

You can easily move your existing self-managed database for these applications to Amazon RDS and automate time-consuming database management tasks, such as software installation, patching, and backups. Here is a simple comparison of features and responsibilities between Amazon EC2, RDS Custom for Oracle, and RDS.

Features and Responsibilities	Amazon EC2	RDS Custom for Oracle	Amazon RDS
Application optimization	Customer	Customer	Customer
Scaling/high availability	Customer	Shared	AWS
DB backups	Customer	Shared	AWS
DB software maintenance	Customer	Shared	AWS
OS maintenance	Customer	Shared	AWS
Server maintenance	AWS	AWS	AWS

via - https://aws.amazon.com/blogs/aws/amazon-rds-custom-for-oracle-new-control-capabilities-in-database-environment/

Incorrect options:

Leverage multi-AZ configuration of Amazon RDS for Oracle that allows the Database Administrator (DBA) to access and customize the database environment and the underlying operating system

Leverage cross AZ read-replica configuration of Amazon RDS for Oracle that allows the Database Administrator (DBA) to access and customize the database environment and the underlying operating system

Amazon RDS for Oracle does not allow you to access and customize your database server host and operating system. Therefore, both these options are incorrect.

Deploy the Oracle database layer on multiple Amazon EC2 instances spread across two Availability Zones (AZs). This deployment configuration guarantees high availability and also allows the Database Administrator (DBA) to access and customize the database environment and the underlying operating system - The use case requires that the best solution should involve minimum infrastructure maintenance effort. When you use Amazon EC2 instances to host the databases, you need to manage the server health, server maintenance, server patching, and database maintenance tasks yourself. In addition, you will also need to manage the multi-AZ configuration by deploying Amazon EC2 instances across two Availability Zones (AZs), perhaps by using an Auto Scaling group. These steps entail significant maintenance effort. Hence this option is incorrect.

References:

https://aws.amazon.com/blogs/aws/amazon-rds-custom-for-oracle-new-control-capabilities-in-database-environment/

https://aws.amazon.com/rds/faqs/

Domain

Design Resilient Architectures