AWS Solutions Architect SAA-C03 Exam Questions

Total Questions: 1100+
Demo Questions: 35
Version: Updated for 2025

Prepared and Verified by Cert Empire – Your Trusted IT Certification Partner

For Access to the full set of Updated Questions – Visit:
SAA-C03 Exam Dumps by Cert Empire

Question: 1

A retail company has several businesses. The IT team for each business manages its own AWS account. Each team account is part of an organization in AWS Organizations. Each team monitors its product inventory levels in an Amazon DynamoDB table in the team's own AWS account.

The company is deploying a central inventory reporting application into a shared AWS account. The application must be able to read items from all the teams' DynamoDB tables. Which authentication option will meet these requirements MOST securely?

A. Integrate DynamoDB with AWS Secrets Manager in the inventory application account. Configure
the application to use the correct secret from Secrets Manager to authenticate and read the DynamoDB table. Schedule secret rotation for every 30 days.

B. In every business account, create an IAM user that has programmatic access. Configure the application to use the correct IAM user access key ID and secret access key to authenticate and read
the DynamoDB table. Manually rotate IAM access keys every 30 days.

C. In every business account, create an IAM role named BUROLE with a policy that gives the role access to the DynamoDB table and a trust policy to trust a specific role in the inventory application
account. In the inventory account, create a role named APPROLE that allows access to the STS AssumeRole API operation. Configure the application to use APPROLE and assume the cross-account role BUROLE to read the DynamoDB table.

D. Integrate DynamoDB with AWS Certificate Manager (ACM). Generate identity certificates to authenticate DynamoDB. Configure the application to use the correct certificate to authenticate and
read the DynamoDB table.

Answer:

C

Explanation:

The most secure method for granting cross-account access is to use IAM roles with temporary security credentials. This approach avoids the use of long-lived credentials like IAM user access keys. In this solution, the application's role (APPROLE) in the central account is explicitly trusted by a role (BUROLE) in each business account. The application calls the AWS Security Token Service (STS) AssumeRole API to obtain short-term credentials for BUROLE. These temporary credentials are then used to access the DynamoDB table, adhering to the principle of least

privilege and AWS security best practices.

References:

1. AWS Identity and Access Management (IAM) User Guide: "IAM roles are the primary way to delegate permissions. You can use them to delegate permissions for users, applications, or services that don't normally have access to your AWS resources... The most common use case for roles is to grant cross-account access." (Reference: IAM User Guide, "IAM roles", "When to create an IAM role (instead of a user)").
2. AWS Identity and Access Management (IAM) User Guide: "We recommend that you use temporary security credentials instead of creating and distributing long-term access keys... When you assume a role, you give up your original permissions and take on the permissions that are assigned to the role." (Reference: IAM User Guide, "Security best practices in IAM", "Use temporary credentials").
3. AWS Identity and Access Management (IAM) User Guide: A detailed tutorial outlines the exact process described in the correct answer. It involves creating a role in the trusting account with a trust policy that specifies the trusted account, and then having a principal in the trusted account call sts:AssumeRole. (Reference: IAM User Guide, "Tutorial: Delegate access across AWS accounts using IAM roles").
4. AWS Certificate Manager User Guide: "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." (Reference: ACM User Guide, "What is AWS Certificate Manager?"). This confirms ACM is not used for IAM authentication.

Question: 2

A company uses an organization in AWS Organizations to manage AWS accounts that contain applications. The company sets up a dedicated monitoring member account in the organization.
The company wants to query and visualize observability data across the accounts by using Amazon CloudWatch.
Which solution will meet these requirements?

A. Enable CloudWatch cross-account observability for the monitoring account. Deploy an AWS CloudFormation template provided by the monitoring account in each AWS account to share the data with the monitoring account.

B. Set up service control policies (SCPs) to provide access to CloudWatch in the monitoring account
under the Organizations root organizational unit (OU).

C. Configure a new IAM user in the monitoring account. In each AWS account, configure an IAM policy to have access to query and visualize the CloudWatch data in the account. Attach the new IAM
policy to the new I AM user.

D. Create a new IAM user in the monitoring account. Create cross-account IAM policies in each AWS
account. Attach the IAM policies to the new IAM user.

Answer:

A

Explanation:

Amazon CloudWatch cross-account observability is the purpose-built feature for this scenario. It enables a central monitoring account to seamlessly search, query, and visualize metrics, logs, and traces from multiple source accounts. The standard setup process involves designating the monitoring account and then linking the source accounts to it. This linking process creates the necessary permissions and can be efficiently deployed across many accounts using an AWS CloudFormation template provided by the monitoring account's configuration wizard. This directly addresses the company's need to query and visualize observability data centrally.

References:

1. Amazon CloudWatch User Guide: "CloudWatch cross-account observability". This guide details the feature's purpose: "Amazon CloudWatch cross-account observability enables you to monitor and troubleshoot applications that span multiple AWS accounts within an AWS Region. You can seamlessly search, visualize, and analyze your metrics, logs, and traces as if you were operating

in a single account." The setup section confirms the process of designating a monitoring account and linking source accounts.
2. AWS Organizations User Guide: "Service control policies (SCPs)". The documentation states, "SCPs don't grant permissions. Instead, SCPs specify the maximum permissions for an organization, organizational unit (OU), or account. To grant permissions to the identities in your accounts, an administrator must attach identity-based or resource-based policies." This confirms that SCPs alone cannot provide access.
3. AWS Identity and Access Management User Guide: "How IAM roles differ from resource-based policies". This section clarifies how cross-account access works, stating, "For cross-account access, you can use a role or a resource-based policy." It does not mention attaching a policy from one account to a user in another, which is the flawed mechanism proposed in options C and D.

Question: 3

A company uses an organization in AWS Organizations to manage AWS accounts that contain applications. The company sets up a dedicated monitoring member account in the organization.
The company wants to query and visualize observability data across the accounts by using Amazon CloudWatch.
Which solution will meet these requirements?

A. Enable CloudWatch cross-account observability for the monitoring account. Deploy an AWS CloudFormation template provided by the monitoring account in each AWS account to share the data with the monitoring account.

B. Set up service control policies (SCPs) to provide access to CloudWatch in the monitoring account
under the Organizations root organizational unit (OU).

C. Configure a new IAM user in the monitoring account. In each AWS account, configure an IAM policy to have access to query and visualize the CloudWatch data in the account. Attach the new IAM
policy to the new IAM user.

D. Create a new IAM user in the monitoring account. Create cross-account IAM policies in each AWS
account. Attach the IAM policies to the new IAM user.

Answer:

A

Explanation:

Amazon CloudWatch cross-account observability is the feature specifically designed to enable a central monitoring account to view, query, and analyze metrics, logs, and traces from multiple source accounts within an AWS Organization. The standard setup process involves designating a monitoring account and then linking source accounts. This linking is accomplished by creating an IAM role and resource policies in each source account that grant the monitoring account the necessary permissions, a process which AWS simplifies by providing a CloudFormation template.

References:

1. AWS CloudWatch Documentation, "CloudWatch cross-account observability": "To set up CloudWatch cross-account observability, you choose one account in a Region to be a monitoring account... You then configure other individual accounts as source accounts... When you link a source account to a monitoring account, you create an IAM role that grants the monitoring account read-only access to the observability data in the source account."

2. AWS CloudWatch Documentation, "Set up CloudWatch cross-account observability": This page details the setup steps, including Step 2 for source accounts: "To configure a source account, you create an IAM role that gives your monitoring account permissions to view and access data from the source account... AWS provides an AWS CloudFormation template that you can use to create the role and attach the necessary policies."
3. AWS Organizations Documentation, "Service control policies (SCPs)": "SCPs are a type of organization policy that you can use to manage permissions in your organization. SCPs offer central control over the maximum available permissions for all accounts in your organization... SCPs don't grant permissions."
4. AWS Identity and Access Management (IAM) Documentation, "IAM roles": "An IAM role is an IAM identity that you can create in your account that has specific permissions... A role can be assumed by an IAM user in a different account." This highlights the correct mechanism for cross-account access (roles), which is abstracted by the CloudWatch feature in option A, and shows why the direct policy attachment in C and D is incorrect.

Question: 4

A company's web application that is hosted in the AWS Cloud recently increased in popularity.
The web application currently exists on a single Amazon EC2 instance in a single public subnet.
The web application has not been able to meet the demand of the increased web traffic.
The company needs a solution that will provide high availability and scalability to meet the increased user demand without rewriting the web application.

Which combination of steps will meet these requirements? (Select TWO.)

A. Replace the EC2 instance with a larger compute optimized instance.

B. Configure Amazon EC2 Auto Scaling with multiple Availability Zones in private subnets.

C. Configure a NAT gateway in a public subnet to handle web requests.

D. Replace the EC2 instance with a larger memory optimized instance.

E. Configure an Application Load Balancer in a public subnet to distribute web traffic

Answer:

B, E

Explanation:

To achieve high availability and scalability for a web application, the architecture must eliminate single points of failure and dynamically adjust to traffic demands. An Application Load Balancer (ALB) serves as a highly available entry point, distributing incoming requests across multiple targets in different Availability Zones (AZs). This prevents the load balancer itself from being a single point of failure. Configuring an Amazon EC2 Auto Scaling group to launch instances across multiple AZs ensures the application can scale horizontally by adding or removing instances based on demand. This multi-AZ deployment also provides high availability by ensuring the application remains operational even if one AZ fails.

References:

1. AWS Documentation - What is an Application Load Balancer?: "An Application Load Balancer distributes incoming application traffic across multiple targets, such as EC2 instances, in multiple Availability Zones. This increases the availability of your application." (Source: AWS Elastic Load Balancing User Guide, "What is an Application Load Balancer?" section).
2. AWS Documentation - What is Amazon EC2 Auto Scaling?: "To improve the availability of your application, you can configure your Auto Scaling group to launch instances in multiple Availability Zones. If one Availability Zone becomes unavailable, Amazon EC2 Auto Scaling launches

instances in another one to compensate." (Source: Amazon EC2 Auto Scaling User Guide, "Fault tolerance" section).
3. AWS Well-Architected Framework - Reliability Pillar: "To increase availability, architect your workload to use multiple Availability Zones (AZs)... Use a load balancer to distribute traffic to multiple instances. This prevents the load balancer from sending traffic to an unhealthy instance." (Source: AWS Well-Architected Framework, Reliability Pillar whitepaper, "Implement a highly available workload" section, p. 26).

Question: 5

A financial services company wants to shut down two data centers and migrate more than 100 TB of data to AWS. The data has an intricate directory structure with millions of small filesstored in deep hierarchies of subfolders. Most of the data is unstructured, and the company's file storage consists of SMB-based storage types from multiple vendors. The company does not want to change its applications to access the data after migration.
What should a solutions architect do to meet these requirements with the LEAST operational overhead?

A. Use AWS Direct Connect to migrate the data to Amazon S3.

B. Use AWS DataSync to migrate the data to Amazon FSx for Lustre.

C. Use AWS DataSync to migrate the data to Amazon FSx for Windows File Server.

D. Use AWS Direct Connect to migrate the data on-premises file storage to an AWS Storage Gateway
volume gateway.

Answer:

C

Explanation:

The core requirements are to migrate over 100 TB of SMB-based file data and ensure existing applications can access it without modification. AWS DataSync is a managed online data transfer service designed to simplify, automate, and accelerate moving large amounts of data between on-premises storage systems and AWS Storage services. Amazon FSx for Windows File Server is a fully managed native Windows file system that provides shared file storage with full support for the SMB protocol. Using DataSync to migrate the data to FSx for Windows File Server directly meets the requirements. This approach preserves the file structure and allows existing applications to connect to the new file shares via the SMB protocol, thus requiring no application changes and incurring the least operational overhead due to the managed nature of both services.

References:

1. AWS DataSync User Guide: In the "What is AWS DataSync?" section, it states, "DataSync can transfer your file data between... Server Message Block (SMB) file shares... and Amazon FSx for Windows File Server file systems." This confirms DataSync is the appropriate tool for migrating from on-premises SMB to FSx for Windows File Server. (Source: AWS DataSync User Guide,

"What is AWS DataSync?", Introduction)
2. Amazon FSx for Windows File Server User Guide: The "What is Amazon FSx for Windows File Server?" section explains, "Amazon FSx for Windows File Server provides fully managed, highly reliable, and scalable file storage that is accessible over the industry-standard Server Message Block (SMB) protocol." This validates it as the correct target for maintaining SMB access. (Source: Amazon FSx for Windows File Server User Guide, "What is Amazon FSx for Windows File Server?", Introduction)
3. AWS Documentation - Migrating storage with AWS DataSync: This documentation outlines a common use case: "You can use DataSync to migrate an active file data set from on-premises to AWS... For file data, you can migrate to Amazon FSx for Windows File Server..." This directly supports the chosen solution. (Source: AWS DataSync User Guide, "Use cases for AWS DataSync", Migrating an active data set)
4. AWS Direct Connect User Guide: The "What is AWS Direct Connect?" section clarifies its purpose: "AWS Direct Connect is a cloud service solution that makes it easy to establish a dedicated network connection from your premises to AWS." This shows it is a networking service, not a data migration service itself. (Source: AWS Direct Connect User Guide, "What is AWS Direct Connect?", Introduction)

Question: 6

A company has a multi-tier payment processing application that is based on virtual machines (VMs).
The communication between the tiers occurs asynchronously through a third-party middleware solution that guarantees exactly-once delivery.
The company needs a solution that requires the least amount of infrastructure management. The solution must guarantee exactly-once delivery for application messaging
Which combination of actions will meet these requirements? (Select TWO.)

A. Use AWS Lambda for the compute layers in the architecture.

B. Use Amazon EC2 instances for the compute layers in the architecture.

C. Use Amazon Simple Notification Service (Amazon SNS) as the messaging component between the
compute layers.

D. Use Amazon Simple Queue Service (Amazon SQS) FIFO queues as the messaging component
between the compute layers.

E. Use containers that are based on Amazon Elastic Kubemetes Service (Amazon EKS) for the compute layers in the architecture.

Answer:

A, D

Explanation:

The core requirements are to minimize infrastructure management and guarantee exactly-once message delivery. AWS Lambda is a serverless compute service, which completely abstracts the underlying infrastructure, thereby meeting the requirement for the least amount of management. Amazon SQS FIFO (First-In, First-Out) queues are specifically designed to prevent duplicate messages and ensure they are processed exactly once in the order they are sent. This combination creates a robust, serverless architecture that directly fulfills both of the company's requirements without the overhead of managing VMs or container clusters.

References:

1. AWS Lambda Documentation: "AWS Lambda is a serverless, event-driven compute service that lets you run code for virtually any type of application or backend service without provisioning or managing servers."
Source: AWS Lambda Developer Guide, "What is AWS Lambda?".
2. Amazon SQS FIFO Documentation: "FIFO (First-In-First-Out) queues are designed to enhance

messaging between applications when the order of operations and events is critical, or where duplicates can't be tolerated….. SQS FIFO queues provide the following features: Exactly-once processing. ."
Source: AWS Simple Queue Service Developer Guide, "Amazon SQS FIFO (First-In-First-Out) queues".
3. Amazon SNS Delivery Guarantees: "A message is delivered at least once. However, occasionally (in rare circumstances), more than one copy of a message is delivered."
Source: AWS Simple Notification Service Developer Guide, "Message delivery guarantees".

Question: 7

A company has established a new AWS account. The account is newly provisioned and no changes have been made to the default settings. The company is concerned about the security of the AWS account root user.

What should be done to secure the root user?

A. Create IAM users for daily administrative tasks. Disable the root user.

B. Create IAM users for daily administrative tasks. Enable multi-factor authentication on the root user.

C. Generate an access key for the root user Use the access key for daily administration tasks instead
of the AWS Management Console.

D. Provide the root user credentials to the most senior solutions architect. Have the solutions architect use the root user for daily administration tasks.

Answer:

B

Explanation:

The most critical security best practice for a new AWS account is to protect the root user. This is achieved by first enabling multi-factor authentication (MFA) on the root user to add a crucial layer of security against unauthorized access. Secondly, for all routine and administrative tasks, dedicated IAM users should be created with only the necessary permissions, adhering to the principle of least privilege. The root user should not be used for daily operations and its credentials should be securely stored.

References:

1. AWS Identity and Access Management (IAM) User Guide: In the "Security best practices in IAM" section, it explicitly states, "We strongly recommend that you do not use the root user for your everyday tasks... For everyday tasks, including administrative tasks, use an IAM user."
Source: AWS IAM User Guide, "Security best practices in IAM", Section: "Create individual IAM users".
2. AWS Identity and Access Management (IAM) User Guide: The documentation emphasizes the importance of MFA for the root user: "For extra security, we recommend that you require multi-factor authentication (MFA) for the AWS account root user... When you use MFA, users have to provide a one-time use code from a second device to sign in."
Source: AWS IAM User Guide, "Signing in to the AWS Management Console as the root user",

Section: "Enable MFA for your AWS account root user".
3. AWS Identity and Access Management (IAM) User Guide: The guide also warns against creating access keys for the root user: "If you don't have an access key for your AWS account root user, don't create one unless you absolutely need to... Instead, create an administrative user in IAM Identity Center... and use that user's credentials".
Source: AWS IAM User Guide, "Security best practices in IAM", Section: "Manage access keys for the root user".

Question: 8

A solutions architect creates a VPC that includes two public subnets and two private subnets. A corporate security mandate requires the solutions architect to launch all Amazon EC2 instances in a private subnet. However, when the solutions architect launches an EC2 instance that runs a web server on ports 80 and 443 in a private subnet, no external internet traffic can connect to the server. What should the solutions architect do to resolve this issue?

A. Attach the EC2 instance to an Auto Scaling group in a private subnet. Ensure that the DNS record
for the website resolves to the Auto Scaling group identifier.

B. Provision an internet-facing Application Load Balancer (ALB) in a public subnet. Add the EC2 instance to the target group that is associated with the ALB. Ensure that the DNS record for the website resolves to the ALB.

C. Launch a NAT gateway in a private subnet. Update the route table for the private subnets to add a
default route to the NAT gateway. Attach a public Elastic IP address to the NAT gateway.

D. Ensure that the security group that is attached to the EC2 instance allows HTTP traffic on port 80
and HTTPS traffic on port 443. Ensure that the DNS record for the website resolves to the public IP
address of the EC2 instance.

Answer:

B

Explanation:

The core requirement is to allow inbound internet traffic to an EC2 instance located in a private subnet, which by definition, is not directly reachable from the internet. An internet-facing Application Load Balancer (ALB) is the standard architectural component for this scenario. The ALB is placed in public subnets, where it receives traffic from the internet via its public DNS name. It then forwards this traffic to its registered targets (the EC2 instances) in the private subnets. This design exposes the web application to users while keeping the underlying compute instances secure and isolated from direct internet access, fulfilling the corporate security mandate.

References:

1. AWS Documentation - Elastic Load Balancing User Guide: "An internet-facing load balancer has a publicly resolvable DNS name, so it can route requests from clients over the internet to the EC2 instances that are registered with the load balancer... We recommend that you place your EC2 instances in private subnets. This ensures that they can't be reached directly from the internet."
Source: AWS Documentation, What is an Application Load Balancer?, Section: "Load balancer scheme".
2. AWS Documentation - Amazon VPC User Guide: "To allow inbound internet traffic to reach the instances in your private subnets, you can use a public-facing load balancer. The load balancer receives traffic from the internet and routes it to your instances."
Source: AWS Documentation, VPC with public and private subnets (NAT), Section: "Overview".
3. AWS Documentation - Amazon VPC User Guide: "If a subnet's traffic is routed to an internet gateway, the subnet is known as a public subnet... If a subnet doesn't have a route to an internet gateway, the subnet is known as a private subnet." This confirms why the instance is unreachable directly.
Source: AWS Documentation, Subnets for your VPC, Section: "Subnet routing".
4. AWS Documentation - Amazon VPC User Guide: "You can use a network address translation (NAT) gateway to enable instances in a private subnet to connect to the internet or other AWS services, but prevent the internet from initiating a connection with those instances."

Source: AWS Documentation, NAT gateways, Introduction paragraph.

Question: 9

A company has an application that uses Docker containers in its local data center The application runs on a container host that stores persistent data in a volume on the host. The containe instances use the stored persistent data.

The company wants to move the application to a fully managed service because the company does not want to manage any servers or storage infrastructure. Which solution will meet these requirements?

A. Use Amazon Elastic Kubernetes Service (Amazon EKS) with self-managed nodes. Create an Amazon Elastic Block Store (Amazon EBS) volume attached to an Amazon EC2 instance. Use the EBS
volume as a persistent volume mounted in the containers.

B. Use Amazon Elastic Container Service (Amazon ECS) with an AWS Fargate launch type. Create an
Amazon Elastic File System (Amazon EFS) volume. Add the EFS volume as a persistent storage volume
mounted in the containers.

C. Use Amazon Elastic Container Service (Amazon ECS) with an AWS Fargate launch type. Create an
Amazon S3 bucket. Map the S3 bucket as a persistent storage volume mounted in the containers.

D. Use Amazon Elastic Container Service (Amazon ECS) with an Amazon EC2 launch type. Create an
Amazon Elastic File System (Amazon EFS) volume. Add the EFS volume as a persistent storage volume
mounted in the containers.

Answer:

B

Explanation:

The core requirements are to use a fully managed service and to avoid managing servers or storage infrastructure. AWS Fargate is a serverless compute engine for containers that allows you to run Amazon Elastic Container Service (ECS) tasks without managing the underlying EC2 instances. This satisfies the "no server management" requirement. For persistent storage, Amazon Elastic File System (EFS) is a fully managed, serverless file storage service. EFS file systems can be mounted directly onto Fargate tasks, providing the necessary persistent storage without requiring the management of storage servers or volumes. This combination directly meets

all the company's requirements.

References:

1. AWS Fargate Documentation, "What is AWS Fargate?": "AWS Fargate is a serverless, pay-as-you-go compute engine... Fargate removes the need to provision and manage servers, lets you specify and pay for resources per application, and improves security through application isolation by design." This supports the choice of Fargate to meet the "no server management" requirement.
2. AWS Documentation, Amazon ECS Developer Guide, "Amazon EFS volumes": "With Amazon EFS, storage capacity is elastic, growing and shrinking automatically as you add and remove files. Your applications can have the storage they need, when they need it... You can use Amazon EFS file systems with Amazon ECS to access file systems for your tasks. You can use this feature with tasks that are hosted on Fargate or Amazon EC2 instances." This confirms EFS is a managed, persistent storage solution compatible with Fargate.
3. AWS Documentation, Amazon ECS Developer Guide, "Amazon ECS launch types": "The EC2 launch type requires that you manage the cluster of Amazon EC2 instances that your tasks are run on... With the Fargate launch type, you only have to package your application in containers." This reference clarifies why options using the EC2 launch type (A and D) are incorrect.
4. AWS Documentation, "Persistent storage for Amazon ECS tasks": This page outlines storage options. It specifies that "Amazon EFS is recommended for applications that require shared storage" and is compatible with Fargate. It also notes that Amazon EBS volumes can be used with ECS on EC2, but not with Fargate, making option A incorrect.

Question: 10

An ecommerce application uses a PostgreSQL database that runs on an Amazon EC2 instance.

During a monthly sales event, database usage increases and causes database connection issues for the application. The traffic is unpredictable for subsequent monthly sales events, which impacts the sales forecast. The company needs to maintain performance when there is an unpredictable increase in
traffic.

Which solution resolves this issue in the MOST cost-effective way?

A. Migrate the PostgreSQL database to Amazon Aurora Serverless v2.

B. Enable auto scaling for the PostgreSQL database on the EC2 instance to accommodate increased
usage.

C. Migrate the PostgreSQL database to Amazon RDS for PostgreSQL with a larger instance type

D. Migrate the PostgreSQL database to Amazon Redshift to accommodate increased usage

Answer:

A

Explanation:

Amazon Aurora Serverless v2 is specifically designed for workloads with unpredictable or intermittent traffic patterns. It automatically and granularly scales database compute and memory capacity in real-time based on application demand, ensuring performance is maintained during traffic spikes. This on-demand scaling model is highly cost-effective because the company pays only for the resources consumed, avoiding the need to provision for peak capacity at all times. This directly addresses the scenario's requirements for handling unpredictable traffic while optimizing costs.

References:

1. Amazon Aurora User Guide - Amazon Aurora Serverless v2: "Aurora Serverless v2 is ideal for a broad set of applications. For example, it's well-suited for enterprises that have applications with infrequent, intermittent, or unpredictable workloads... With Aurora Serverless v2, you pay only for the capacity that you consume."
2. AWS Documentation - What is Amazon Aurora?: "Amazon Aurora Serverless is an on-demand, auto-scaling configuration for Amazon Aurora. It automatically starts up, shuts down, and scales capacity up or down based on your application's needs." (This general principle is refined in v2 for more granular scaling).

3. AWS Documentation - Amazon Redshift system overview: "Amazon Redshift is a fully managed, petabyte-scale data warehouse service in the cloud. You can start with just a few hundred gigabytes of data and scale to a petabyte or more. This enables you to use your data to acquire new insights for your business and customers." (This highlights its OLAP, not OLTP, purpose).
4. AWS Documentation - Modifying a DB instance - Amazon Relational Database Service: The process of changing an instance class (vertical scaling) is a manual modification that can result in an outage. It is not an automatic, on-demand scaling mechanism suitable for unpredictable traffic.

Question: 11

A company uses AWS Organizations. The company wants to operate some of its AWS accounts with different budgets. The company wants to receive alerts and automatically prevent provisioning of additional resources on AWS accounts when the allocated budget threshold is met during a specific period.

Which combination of solutions will meet these requirements? (Select THREE.)

A. Use AWS Budgets to create a budget. Set the budget amount under the Cost and Usage Reports
section of the required AWS accounts.

B. Use AWS Budgets to create a budget. Set the budget amount under the Billing dashboards of the
required AWS accounts.

C. Create an IAM user for AWS Budgets to run budget actions with the required permissions.

D. Create an IAM role for AWS Budgets to run budget actions with the required permissions.

E. Add an alert to notify the company when each account meets its budget threshold. Add a budget

action that selects the IAM identity created with the appropriate config rule to prevent provisioning of additional resources.

F. Add an alert to notify the company when each account meets its budget threshold. Add a budget
action that selects the IAM identity created with the appropriate service control policy (SCP) to prevent provisioning of additional resources.

Answer:

B, D, F

Explanation:

To meet the requirements, the company must first use AWS Budgets, which is accessed via the AWS Billing and Cost Management console, to create a budget for the specific accounts (B). To automatically prevent resource provisioning, an AWS Budgets action must be configured. This action requires permissions to modify policies within the target account. The secure and correct method for granting an AWS service these permissions is by creating an IAM role for it to assume (D). Finally, the budget action should be configured to apply a restrictive Service Control Policy (SCP) to the target account when the budget threshold is met. SCPs are the AWS Organizations feature designed to enforce permission guardrails and prevent actions like resource provisioning

(F).

References:

1. AWS Cost Management User Guide, "Creating a budget": This guide details the process of setting up a budget within the AWS Budgets service, which is part of the Billing and Cost Management console. This supports option B. (AWS Documentation, AWS Cost Management, User Guide, Managing your costs with AWS Budgets, Creating a budget).
2. AWS Cost Management User Guide, "Permissions and security for AWS Budgets actions": This document explicitly states, "To allow Budgets to run actions on your behalf, you create an AWS Identity and Access Management (IAM) role." This confirms the requirement for an IAM role (D) over an IAM user (C).
3. AWS Cost Management User Guide, "Applying IAM policies and service control policies": This section describes how to configure a budget action to "Attach a service control policy (SCP)" to a target account or OU. This directly supports using an SCP for preventative control (F).
4. AWS Organizations User Guide, "Service control policies (SCPs)": This documentation defines SCPs as a type of organization policy that you can use to manage permissions in your organization. It states, "SCPs offer central control over the maximum available permissions for all accounts in your organization," confirming their role in preventing actions. This validates why SCPs (F) are correct and Config rules (E) are not for this purpose.

Question: 12

A company is building a shopping application on AWS. The application offers a catalog that changes once each month and needs to scale with traffic volume. The company wants the lowest possible latency from the application. Data from each user's shopping carl needs to be highly available.
User session data must be available even if the user is disconnected and reconnects.
What should a solutions architect do to ensure that the shopping cart data is preserved at all times?

A. Configure an Application Load Balancer to enable the sticky sessions feature (session affinity) for
access to the catalog in Amazon Aurora.

B. Configure Amazon ElastiCacJie for Redis to cache catalog data from Amazon DynamoDB and shopping carl data from the user's session.

C. Configure Amazon OpenSearch Service to cache catalog data from Amazon DynamoDB and shopping cart data from the user's session.

D. Configure an Amazon EC2 instance with Amazon Elastic Block Store (Amazon EBS) storage for the
catalog and shopping cart. Configure automated snapshots.

Answer:

B

Explanation:

This solution effectively addresses all requirements. Amazon ElastiCache for Redis is an in-memory data store that provides extremely low latency, making it ideal for caching the infrequently changing product catalog. For the shopping cart and session data, using a centralized and durable store like ElastiCache for Redis decouples the session state from the individual application servers. This ensures that user data is highly available and preserved even if an application instance fails or the user is routed to a different instance upon reconnecting, fulfilling the core requirements of the question.

References:

1. Amazon ElastiCache for Redis User Guide, "Common use cases for Redis": This document explicitly lists "Caching" and "Session store" as primary use cases. For session stores, it states, "Redis is a popular choice for managing session data because it provides the low latency, high scalability, and availability that modern applications require." This directly supports using Redis for both the catalog cache and the shopping cart session data.

2. AWS Documentation, Elastic Load Balancing, "Sticky sessions for your Application Load Balancer": This guide explains that sticky sessions route requests from the same client to the same target. This design pattern makes the session state dependent on the health of a single target, which is contrary to the high availability requirement for the shopping cart data.
3. AWS Well-Architected Framework, "Reliability Pillar" whitepaper, Design Principles section: This paper emphasizes the principle to "Scale horizontally to increase aggregate workload availability." The solution in option D, using a single EC2 instance, directly contradicts this fundamental principle for building reliable and scalable systems on AWS.
4. Amazon OpenSearch Service Developer Guide, "What Is Amazon OpenSearch Service?": This document describes the service's purpose for use cases like "log analytics, real-time application monitoring, and clickstream analytics," confirming it is not the intended or optimal service for low-latency session management.

Question: 13

A company is designing a new web service that will run on Amazon EC2 instances behind an Elastic Load Balancing (ELB) load balancer. However, many of the web service clients can only reach IP addresses authorized on their firewalls.
What should a solutions architect recommend to meet the clients' needs?

A. A Network Load Balancer with an associated Elastic IP address.

B. An Application Load Balancer with an associated Elastic IP address.

C. An A record in an Amazon Route 53 hosted zone pointing to an Elastic IP address.

D. An EC2 instance with a public IP address running as a proxy in front of the load balancer.

Answer:

A

Explanation:

The primary requirement is to provide a static IP address for a load-balanced service that clients can whitelist in their firewalls. A Network Load Balancer (NLB) is the only type of Elastic Load Balancer that can be directly associated with a static Elastic IP address for each Availability Zone it is enabled in. This provides a fixed, predictable entry point for client traffic, fulfilling the requirement while distributing traffic across the backend EC2 instances. Application Load Balancers do not have static IP addresses, making the NLB the correct choice for this use case.

References:

1. AWS Documentation: Elastic Load Balancing User Guide. In the section for Network Load Balancers, it states: "For an internet-facing load balancer, you can optionally associate one Elastic IP address per subnet. This provides your load balancer with a static IP address in each Availability Zone." (Source: Elastic Load Balancing User Guide, "Network Load Balancers", section "Elastic IP addresses").
2. AWS Documentation: Elastic Load Balancing User Guide. In contrast, for Application Load Balancers, the documentation clarifies their dynamic nature: "The IP addresses for Application Load Balancers change over time and are not static." (Source: Elastic Load Balancing User Guide, "Application Load Balancers", section "Load balancer IP addresses").
3. AWS Documentation: What is AWS Global Accelerator? This service is mentioned as the solution for getting static IPs for Application Load Balancers, confirming that ALBs do not have this capability natively. "AWS Global Accelerator provides you with a set of two static IP addresses that are anycast from the AWS edge network... You can associate these addresses with regional AWS resources, such as Application Load Balancers..." (Source: AWS Global Accelerator Developer Guide, "What is AWS Global Accelerator?"). This highlights why option B

is incorrect and why a different service (NLB) is needed if Global Accelerator is not used.

Question: 14

A company runs a Java-based job on an Amazon EC2 instance. The job runs every hour and takes 10 seconds to run. The job runs on a scheduled interval and consumes 1 GB of memory. The CPU utilization of the instance is low except for short surges during which the job uses the maximum CPU available. The company wants to optimize the costs to run the job. Which solution will meet these requirements?

A. Use AWS App2Container (A2C) to containerize the job. Run the job as an Amazon Elastic Container
Service (Amazon ECS) task on AWS Fargate with 0.5 virtual CPU (vCPU) and 1 GB of memory.

B. Copy the code into an AWS Lambda function that has 1 GB of memory. Create an Amazon EventBridge scheduled rule to run the code each hour.

C. Use AWS App2Container (A2C) to containerize the job. Install the container in the existing Amazon
Machine Image (AMI). Ensure that the schedule stops the container when the task finishes.

D. Configure the existing schedule to stop the EC2 instance at the completion of the job and restart
the EC2 instance when the next job starts.

Answer:

B

Explanation:

The most cost-effective solution is to migrate the job to AWS Lambda. Lambda is a serverless compute service that runs code in response to triggers and automatically manages the underlying compute resources. The billing model is based on the number of requests and the duration of execution, measured in milliseconds. For a job that runs for only 10 seconds every hour, this pay-per-use model eliminates the cost of an idle EC2 instance, which incurs charges even when the job is not running. An Amazon EventBridge rule can be configured to trigger the Lambda function on the required hourly schedule, providing a highly cost-optimized and efficient solution for this workload.

References:

1. AWS Lambda Pricing: "AWS Lambda counts a request each time it starts executing in response to an event notification trigger... Duration is calculated from the time your code begins executing until it returns or otherwise terminates, rounded up to the nearest 1ms."
Source: AWS Lambda Pricing documentation. (aws.amazon.com/lambda/pricing/)

2. Amazon EC2 On-Demand Pricing: "Per-second billing (with a 60-second minimum) for instances launched in On-Demand, Reserved, and Spot form."
Source: AWS News Blog, "Per-Second Billing for EC2 Instances and EBS Volumes". (aws.amazon.com/blogs/aws/new-per-second-billing-for-ec2-instances-and-ebs-volumes/)
3. AWS Well-Architected Framework - Cost Optimization Pillar: This framework recommends adopting a consumption model. "Pay-as-you-go allows you to easily adapt to changing business needs without overcommitting budgets... For example, serverless services like AWS Lambda let you pay only for what you use."
Source: AWS Well-Architected Framework, Cost Optimization Pillar, "Adopt a consumption model" section. (docs.aws.amazon.com/wellarchitected/latest/cost-optimization-pillar/adopt-a-consumption-model.html)
4. Amazon EventBridge (CloudWatch Events) Documentation: "You can create a rule that self-triggers on a schedule in Amazon EventBridge." This confirms the ability to schedule the Lambda invocation.
Source: Amazon EventBridge User Guide, "Schedule expressions for rules". (docs.aws.amazon.com/eventbridge/latest/userguide/eb-schedule-expressions.html)

Question: 15

The DNS provider that hosts a company's domain name records is experiencing outages that cause service disruption for a website running on AWS. The company needs to migrate to a more resilient managed DNS service and wants the service to run on AWS.

What should a solutions architect do to rapidly migrate the DNS hosting service?

A. Create an Amazon Route 53 public hosted zone for the domain name. Import the zone file containing the domain records hosted by the previous provider

B. Create an Amazon Route 53 private hosted zone for the domain name Import the zone file containing the domain records hosted by the previous provider.

C. Create a Simple AD directory in AWS. Enable zone transfer between the DNS provider and AWS
Directory Service for Microsoft Active Directory for the domain records.

D. Create an Amazon Route 53 Resolver inbound endpomt in the VPC. Specify the IP addresses that
the provider's DNS will forward DNS queries to. Configure the provider's DNS to forward DNS queries

for the domain to the IP addresses that are specified in the inbound endpoint.

Answer:

A

Explanation:

The goal is to rapidly migrate an unreliable external DNS provider to a resilient, AWS-managed service for a public website. Amazon Route 53 is AWS's highly available and scalable managed Domain Name System (DNS) web service. To make a website's domain name resolvable on the public internet, a public hosted zone is required in Route 53. The most efficient way to migrate existing DNS records is to export them from the current provider into a standard zone file format and then import this file directly into the newly created Route 53 public hosted zone. This process minimizes manual configuration and accelerates the migration. After importing, the domain's name server (NS) records at the registrar must be updated to point to the Route 53 name servers.

References:

1. Amazon Route 53 Developer Guide: "To make your domain's content available to users on the internet, you create a public hosted zone." This confirms the need for a public hosted zone for a website.
Source: AWS Documentation, Amazon Route 53 Developer Guide, "Working with hosted zones,"

"Creating a public hosted zone."
2. Amazon Route 53 Developer Guide: "If you have a lot of existing records, you can save time and reduce the likelihood of typographical errors by importing a zone file into your Amazon Route 53 hosted zone." This supports the "rapid migration" requirement by using a zone file import.
Source: AWS Documentation, Amazon Route 53 Developer Guide, "Creating records by importing a zone file."
3. Amazon Route 53 Developer Guide: "A private hosted zone is a container that holds information about how you want to route traffic for a domain and its subdomains within one or more Amazon Virtual Private Clouds (Amazon VPCs)." This explains why option B is incorrect.
Source: AWS Documentation, Amazon Route 53 Developer Guide, "Working with private hosted zones."
4. Amazon Route 53 Developer Guide: "With Route 53 Resolver, you can create endpoints in your VPCs that allow you to connect your VPCs and your on-premises network in a hybrid cloud environment." This confirms that Resolver endpoints are for hybrid DNS, not public hosting, making option D incorrect.
Source: AWS Documentation, Amazon Route 53 Developer Guide, "What is Amazon Route 53 Resolver?."

Question: 16

A company is building a microservices-based application that will be deployed on Amazon Elastic Kubernetes Service (Amazon EKS). The microservices will interact with each other. The company wants to ensure that the application is observable to identify performance issues in the future. Which solution will meet these requirements?

A. Configure the application to use Amazon ElastiCache to reduce the number of requests that are
sent to the microservices.

B. Configure Amazon CloudWatch Container Insights to collect metrics from the EKS clusters Configure AWS X-Ray to trace the requests between the microservices.

C. Configure AWS CloudTrail to review the API calls. Build an Amazon QuickSight dashboard to observe the microservice interactions.

D. Use AWS Trusted Advisor to understand the performance of the application.

Answer:

B

Explanation:

To achieve observability in a microservices architecture on Amazon EKS, a combination of metrics collection and distributed tracing is required. Amazon CloudWatch Container Insights is specifically designed to collect, aggregate, and summarize metrics and logs from containerized applications, providing visibility into the performance of the EKS cluster and its pods. AWS X-Ray complements this by providing distributed tracing capabilities. It traces user requests as they travel through the different microservices, creating a service map that helps identify performance bottlenecks, latency issues, and errors between services, directly fulfilling the requirement to identify performance issues.

References:

1. AWS Documentation - CloudWatch Container Insights: "CloudWatch Container Insights collects, aggregates, and summarizes metrics and logs from your containerized applications and microservices. Container Insights is available for Amazon Elastic Container Service (Amazon ECS), Amazon Elastic Kubernetes Service (Amazon EKS)..." (Source: AWS CloudWatch User Guide, "Using Container Insights", Introduction section).
2. AWS Documentation - AWS X-Ray: "AWS X-Ray helps developers analyze and debug production, distributed applications, such as those built using a microservices architecture. With X-Ray, you can understand how your application and its underlying services are performing to identify and troubleshoot the root cause of performance issues and errors." (Source: AWS X-Ray Developer Guide, "What Is AWS X-Ray?", Introduction section).

3. AWS Documentation - AWS CloudTrail: "AWS CloudTrail is an AWS service that helps you enable governance, compliance, and operational and risk auditing of your AWS account. Actions taken by a user, role, or an AWS service are recorded as events in CloudTrail." (Source: AWS CloudTrail User Guide, "What Is AWS CloudTrail?", Introduction section).
4. AWS Documentation - AWS Trusted Advisor: "AWS Trusted Advisor provides recommendations that help you follow AWS best practices. Trusted Advisor evaluates your account by using checks. These checks identify ways to optimize your AWS infrastructure, improve security and performance, reduce costs, and monitor service quotas." (Source: AWS Support User Guide, "AWS Trusted Advisor", Introduction section).

Question: 17

A company hosts a three-tier web application in the AWS Cloud. A Multi-AZ Amazon RDS for MySQL server forms the database layer. Amazon ElastiCache forms the cache layer. The company wants a caching strategy that adds or updates data in the cache when a customer adds an item to the database. The data in the cache must always match the data in the database. Which solution will meet these requirements?

A. Implement the lazy loading caching strategy

B. Implement the write-through caching strategy.

C. Implement the adding TTL caching strategy.

D. Implement the AWS AppConfig caching strategy.

Answer:

B

Explanation:

The core requirement is to ensure that data in the cache always matches the data in the database immediately after a write operation. The write-through caching strategy is designed for this purpose. With write-through, the application writes data to the cache and the database simultaneously. The write operation is considered complete only after both systems are updated. This approach maintains strong data consistency between the cache and the database, fulfilling the requirement at the cost of slightly increased write latency.

References:

1. AWS ElastiCache for Redis User Guide: In the "Caching strategies" section, it states, "The write-through strategy adds data or updates data in the cache whenever data is written to the database. This strategy keeps the cache and the database synchronized."
Source: AWS ElastiCache for Redis User Guide, Section: "Caching strategies", Sub-section: "Write-through".
2. AWS ElastiCache for Redis User Guide: The guide describes lazy loading: "Lazy loading is a caching strategy that loads data into the cache only when necessary... This can result in stale data if the data is updated in the database directly."
Source: AWS ElastiCache for Redis User Guide, Section: "Caching strategies", Sub-section: "Lazy loading".
3. AWS AppConfig User Guide: The documentation defines the service's purpose: "Use AWS AppConfig... to create, manage, and quickly deploy application configurations." This confirms it is not a general-purpose data caching service.

Source: AWS AppConfig User Guide, Section: "What Is AWS AppConfig?".

Question: 18

A company is running its production and nonproduction environment workloads in multiple AWS accounts. The accounts are in an organization in AWS Organizations. The company needs to design a solution that will prevent the modification of cost usage tags. Which solution will meet these requirements?

A. Create a custom AWS Config rule to prevent tag modification except by authorized principals.

B. Create a custom trail in AWS CloudTrail to prevent tag modification

C. Create a service control policy (SCP) to prevent tag modification except by authonzed principals.

D. Create custom Amazon CloudWatch logs to prevent tag modification.

Answer:

C

Explanation:

Service Control Policies (SCPs) are a feature of AWS Organizations used to manage permissions and enforce preventive guardrails across multiple AWS accounts. An SCP can be crafted with a Deny statement for tag modification actions (e.g., CreateTags, DeleteTags for various services). By applying a condition to this statement, specific IAM principals (users or roles) can be exempted, allowing them to manage tags while preventing all other users. This directly meets the requirement for a centrally managed, preventive control.

References:

1. AWS Organizations User Guide, "Service control policies (SCPs)": "SCPs are a type of organization policy that you can use to manage permissions in your organization. SCPs offer central control over the maximum available permissions for all accounts in your organization... SCPs are powerful because they affect all users, including the root user of the account." This establishes SCPs as the tool for central, preventive control.
2. AWS Organizations User Guide, "Strategies for using SCPs": The section on "Deny access based on a condition" explains how to use Deny statements with Condition elements. An example policy could deny actions like ec2:CreateTags unless the request comes from an authorized principal (aws:PrincipalArn).
3. AWS Config Developer Guide, "What Is AWS Config?": "AWS Config is a service that enables you to assess, audit, and evaluate the configurations of your AWS resources." This confirms its role as a detective, not preventive, service.
4. AWS CloudTrail User Guide, "What Is AWS CloudTrail?": "CloudTrail is an AWS service that helps you enable operational and risk auditing, governance, and compliance of your AWS

account. Actions taken by a user, role, or an AWS service are recorded as events in CloudTrail." This highlights its function as a logging and auditing tool.

Question: 19

A research company runs experiments that are powered by a simu-lation application and a visualization application. The simu-lation application runs on Linux and outputs intermediate data to an NFS share every 5 minutes. The visualization application is a Windows desktop application that displays the simu-lation output and requires an SMB file system.

The company maintains two synchronized file systems. This strategy is causing data duplication and inefficient resource usage. The company needs to migrate the applications to AWS without making code changes to either application.

Which solution will meet these requirements?

A. Migrate both applications to AWS Lambda. Create an Amazon S3 bucket to exchange data between the applications.

B. Migrate both applications to Amazon Elastic Container Service (Amazon ECS). Configure Amazon
FSx File Gateway for storage.

C. Migrate the simulation application to Linux Amazon EC2 instances. Migrate the visualization application to Windows EC2 instances. Configure Amazon Simple Queue Service (Amazon SQS) to
exchange data between the applications.

D. Migrate the simulation application to Linux Amazon EC2 instances. Migrate the visualization application to Windows EC2 instances. Configure Amazon FSx for NetApp ONTAP for storage.

Answer:

D

Explanation:

The primary requirement is to migrate two applications, one Linux-based using NFS and one Windows-based using SMB, to a single, efficient storage solution on AWS without code changes. Migrating the applications to their respective EC2 instance types (Linux and Windows) is a standard lift-and-shift approach. Amazon FSx for NetApp ONTAP is a fully managed file storage service that uniquely provides simultaneous, multi-protocol access to the same data over both NFS and SMB. This allows the Linux EC2 instance to connect via NFS and the Windows EC2 instance to connect via SMB to a single, shared data repository, directly solving the problem of data duplication and inefficiency.

References:

1. Amazon FSx for NetApp ONTAP Documentation: "Amazon FSx for NetApp ONTAP provides fully managed shared storage in the AWS Cloud with the popular data access and management capabilities of ONTAP. It provides multi-protocol access to data over NFS, SMB, and iSCSI protocols."
Source: AWS Documentation, "What is Amazon FSx for NetApp ONTAP?", Introduction.
2. Accessing data across multiple protocols - FSx for NetApp ONTAP: "You can access your data in an Amazon FSx for NetApp ONTAP file system from Linux, Windows, and macOS clients over the Network File System (NFS) protocol... and the Server Message Block (SMB) protocol.. "
Source: AWS Documentation, "FSx for NetApp ONTAP User Guide", section "Accessing data across multiple protocols".
3. Lift-and-shift migration to AWS: "Rehosting-otherwise known as "lift and shift"-is a common strategy. .. In this scenario, you are moving your application to the cloud with little to no changes." Migrating to EC2 is a primary example of this strategy.
Source: AWS Cloud Enterprise Strategy Blog, "6 Strategies for Migrating Applications to the Cloud", section "Rehosting ("lift and shift")".
4. Amazon SQS Documentation: "Amazon Simple Queue Service (SQS) is a fully managed message queuing service.. " This confirms it is not a file system and would require application changes to use.
Source: AWS Documentation, "Amazon SQS Developer Guide", section "What is Amazon SQS?".

Question: 20

A company has a web application that includes an embedded NoSQL database. The application runs on Amazon EC2 instances behind an Application Load Balancer (ALB). The instances run in an Amazon EC2 Auto Scaling group in a single Availability Zone.

A recent increase in traffic requires the application to be highly available and for the database to be eventually consistent. Which solution will meet these requirements with the LEAST operational overhead?

A. Replace the ALB with a Network Load Balancer Maintain the embedded NoSQL database with its
replication service on the EC2 instances.

B. Replace the ALB with a Network Load Balancer Migrate the embedded NoSQL database to Amazon
DynamoDB by using AWS Database Migration Service (AWS DMS).

C. Modify the Auto Scaling group to use EC2 instances across three Availability Zones. Maintain the
embedded NoSQL database with its replication service on the EC2 instances.

D. Modify the Auto Scaling group to use EC2 instances across three Availability Zones. Migrate the
embedded NoSQL database to Amazon DynamoDB by using AWS Database Migration Service (AWS
DMS).

Answer:

D

Explanation:

This solution comprehensively addresses all requirements with the least operational overhead. Modifying the Auto Scaling group to span multiple Availability Zones (AZs) resolves the single point of failure for the application tier, making it highly available. The Application Load Balancer (ALB) is designed to distribute traffic across multiple AZs. Migrating the embedded NoSQL database to Amazon DynamoDB, a fully managed service, fulfills the database requirements. DynamoDB is inherently highly available, replicating data across multiple AZs by default, and provides eventually consistent reads, which aligns with the requirements while offloading the operational burden of managing, scaling, and replicating the database.

References:

1. Amazon EC2 Auto Scaling User Guide: "To improve the availability of your applications, you can distribute your instances across multiple Availability Zones. When one Availability Zone becomes unhealthy or unavailable, Amazon EC2 Auto Scaling launches new instances in an unaffected Availability Zone." (Source: AWS Documentation, Auto Scaling groups with multiple instance types and purchase options, Section: "Availability Zone balance").
2. Amazon DynamoDB Developer Guide: "Amazon DynamoDB is a fully managed NoSQL database service... DynamoDB automatically spreads the data and traffic for your tables over a sufficient number of servers to handle your throughput and storage requirements, while maintaining consistent and fast performance. All of your data is stored on solid-state disks (SSDs) and is automatically replicated across multiple Availability Zones in an AWS Region, providing built-in high availability and data durability." (Source: AWS Documentation, What Is Amazon DynamoDB?, Introduction).
3. AWS Database Migration Service User Guide: "You can use AWS DMS to migrate data to and from most widely used commercial and open-source databases... For example, you can migrate your data from an on-premises Oracle database to an Amazon Aurora MySQL-Compatible Edition database. You can also migrate data between the same database engines, such as from an Oracle database on-premises to an Oracle database on Amazon EC2." (This confirms DMS can be used for migrations, including to DynamoDB). (Source: AWS Documentation, What is AWS Database Migration Service?).

4. Elastic Load Balancing User Guide: "When you enable an Availability Zone for your load balancer, Elastic Load Balancing creates a load balancer node in the Availability Zone. If you register targets in an Availability Zone but do not enable it, these registered targets do not receive traffic." (This implies ALBs are designed to work across AZs for high availability). (Source: AWS Documentation, Load balancer Availability Zones).

Question: 21

A media company stores movies in Amazon S3. Each movie is stored in a single video file that ranges from 1 GB to 10 GB in size.

The company must be able to provide the streaming content of a movie within 5 minutes of a user purchase. There is higher demand for movies that are less than 20 years old than for movies that are more than 20 years old. The company wants to minimize hosting service costs based on demand.
Which solution will meet these requirements?

A. Store all media content in Amazon S3. Use S3 Lifecycle policies to move media data into the Infrequent Access tier when the demand for a movie decreases.

B. Store newer movie video files in S3 Standard Store older movie video files in S3 Standard-Infrequent Access (S3 Standard-IA). When a user orders an older movie, retrieve the video file by using standard retrieval.

C. Store newer movie video files in S3 Intelligent-Tiering. Store older movie video files in S3 Glacier
Flexible Retrieval. When a user orders an older movie, retrieve the video file by using expedited retrieval.

D. Store newer movie video files in S3 Standard. Store older movie video files in S3 Glacier Flexible
Retrieval. When a user orders an older movie, retrieve the video file by using bulk retrieval.

Answer:

C

Explanation:

This solution provides the most cost-effective storage strategy while meeting the strict retrieval time requirement. S3 Intelligent-Tiering is ideal for newer movies, as it automatically moves objects between frequent and infrequent access tiers based on changing access patterns, optimizing costs without performance impact. For older, rarely accessed movies, S3 Glacier Flexible Retrieval offers significantly lower storage costs than other tiers. The key to meeting the performance requirement is using expedited retrieval, which typically makes objects available within 1-5 minutes, satisfying the "within 5 minutes" service level agreement (SLA) for user purchases.

References:

1. Amazon S3 User Guide, "Amazon S3 storage classes": This document details the use cases for different S3 storage classes. It describes S3 Intelligent-Tiering as designed for "data with unknown, changing, or unpredictable access patterns" and S3 Glacier Flexible Retrieval for "archive data that is accessed 1-2 times per year and is retrieved asynchronously."
2. Amazon S3 User Guide, "Restoring an archived object": Under the section "Restore options," the documentation specifies the retrieval times for S3 Glacier Flexible Retrieval. It states, "Expedited retrievals are typically made available within 1-5 minutes." This directly supports the feasibility of option C meeting the 5-minute requirement.
3. Amazon S3 User Guide, "Comparing the Amazon S3 storage classes": The comparison table in this section shows that S3 Glacier Flexible Retrieval has a lower per-GB storage cost than S3 Standard-Infrequent Access (S3 Standard-IA), confirming that option C is more cost-effective for storing the older movies than option B.

Question: 22

A company copies 200 TB of data from a recent ocean survey onto AWS Snowball Edge Storage Optimized devices. The company has a high performance computing (HPC) cluster that is hosted on AWS to look for oil and gas deposits. A solutions architect must provide the cluster with consistent sub-millisecond latency and high-throughput access to the data on the Snowball Edge Storage Optimized devices. The company is sending the devices back to AWS.
Which solution will meet these requirements?

A. Create an Amazon S3 bucket. Import the data into the S3 bucket. Configure an AWS Storage Gateway file gateway to use the S3 bucket. Access the file gateway from the HPC cluster instances.

B. Create an Amazon S3 bucket. Import the data into the S3 bucket. Configure an Amazon FSx for Lustre file system, and integrate it with the S3 bucket. Access the FSx for Lustre file system from the HPC cluster instances

C. Create an Amazon S3 bucket and an Amazon Elastic File System (Amazon EFS) file system. Import

the data into the S3 bucket. Copy the data from the S3 bucket to the EFS file system. Access the EFS file system from the HPC cluster instances

D. Create an Amazon FSx for Lustre file system. Import the data directly into the FSx for Lustre file system Access the FSx for Lustre file system from the HPC cluster instances

Answer:

B

Explanation:

The scenario requires a storage solution for a High-Performance Computing (HPC) cluster with consistent sub-millisecond latency and high throughput. Amazon FSx for Lustre is a fully managed file system specifically designed and optimized for HPC and other compute-intensive workloads. It provides sub-millisecond latencies and up to hundreds of gigabytes per second of throughput. The standard workflow for importing data from an AWS Snowball device is to an Amazon S3 bucket. FSx for Lustre can be seamlessly integrated with an S3 bucket, presenting the S3 objects as a high-performance POSIX-compliant file system to the HPC cluster. This architecture directly meets all the specified performance and operational requirements.

References:

1. Amazon FSx for Lustre - Use cases: "High performance computing (HPC) is used to solve complex, data-intensive problems... These applications often require high-performance storage with sub-millisecond latencies and high levels of throughput. Amazon FSx for Lustre provides a high-performance file system optimized for these workloads." (AWS Documentation, Amazon FSx for Lustre, "Use cases for Amazon FSx for Lustre").
2. Amazon FSx for Lustre - S3 Integration: "You can link your FSx for Lustre file system to your Amazon S3 data lake. When linked to an S3 bucket, an FSx for Lustre file system transparently presents S3 objects as files..." (AWS Documentation, Amazon FSx for Lustre, "Working with data repositories").
3. AWS Snowball Edge - Job Completion: "After the device arrives at the AWS facility, the data is moved from the device into your Amazon S3 buckets. The time it takes to transfer the data varies depending on the amount of data and the terms of your job." (AWS Documentation, AWS Snowball Edge Developer Guide, "How AWS Snowball Edge Works", Section: "Completing a Job").
4. Amazon EFS - Performance: "Amazon EFS provides low, consistent latencies... For latency-sensitive applications, use the General Purpose performance mode... EFS offers single-digit millisecond latencies for file system operations." (AWS Documentation, Amazon EFS User Guide, "Amazon EFS performance"). This contrasts with the sub-millisecond latency offered by FSx for Lustre.

Question: 23

A company is running a photo hosting service in the us-east-1 Region. The service enables users across multiple countries to upload and view photos. Some photos are heavily viewed for months, and others are viewed for less than a week. The application allows uploads of up to 20 MB for each photo. The service uses the photo metadata to determine which photos to display to each user. Which solution provides the appropriate user access MOST cost-effectively?

A. Store the photos in Amazon DynamoDB. Turn on DynamoDB Accelerator (DAX) to cache frequently
viewed items.

B. Store the photos in the Amazon S3 Intelligent-Tiering storage class. Store the photo metadata and
its S3 location in DynamoDB.

C. Store the photos in the Amazon S3 Standard storage class. Set up an S3 Lifecycle policy to move
photos older than 30 days to the S3 Standard-Infrequent Access (S3 Standard-IA) storage class. Use
the object tags to keep track of metadata.

D. Store the photos in the Amazon S3 Glacier storage class. Set up an S3 Lifecycle policy to move
photos older than 30 days to the S3 Glacier Deep Archive storage class. Store the photo metadata
and its S3 location in Amazon OpenSearch Service.

Answer:

B

Explanation:

This solution is the most cost-effective and technically sound. Amazon S3 is the ideal service for storing large binary objects like photos. The Amazon S3 Intelligent-Tiering storage class is specifically designed to optimize costs for data with unknown or changing access patterns, as described in the scenario ("Some photos are heavily viewed for months, and others are viewed for less than a week"). It automatically moves objects between frequent and infrequent access tiers without performance impact or operational overhead. Storing the photo metadata and its S3 location in Amazon DynamoDB provides a highly scalable, low-latency database for the fast lookups required by the application to display photos to users.

References:

1. Amazon S3 Intelligent-Tiering: AWS Documentation states, "The S3 Intelligent-Tiering storage class is designed to optimize storage costs by automatically moving data to the most cost-effective access tier, without performance impact or operational overhead.....S3 Intelligent-Tiering is the ideal storage class for data with unknown, changing, or unpredictable access patterns. ."
Source: AWS Documentation, Amazon S3 User Guide, "Storage classes for frequently and infrequently accessed objects," section "S3 Intelligent-Tiering."
2. DynamoDB for Metadata: AWS recommends DynamoDB for use cases requiring low-latency data retrieval, such as storing metadata for objects located in Amazon S3. This is a common architectural pattern.
Source: AWS Documentation, Amazon DynamoDB Developer Guide, "Use Cases and Design Patterns," section "Adjacency Lists." (This pattern is often used for metadata).
3. DynamoDB Item Size Limit: The official service quotas for DynamoDB confirm the size limitation that makes option A invalid.
Source: AWS Documentation, Amazon DynamoDB Developer Guide, "Service, account, and table quotas in Amazon DynamoDB," section "Item size."
4. S3 Glacier Retrieval Times: The S3 documentation specifies that S3 Glacier is for archiving, with retrieval options that take minutes to hours, making it unsuitable for this use case.
Source: AWS Documentation, Amazon S3 User Guide, "Archiving objects," section "Retrieval options."

Question: 24

A company is deploying an application in three AWS Regions using an Application Load Balancer Amazon Route 53 will be used to distribute traffic between these Regions. Which Route 53 configuration should a solutions architect use to provide the MOST high-performing experience?

A. Create an A record with a latency policy.

B. Create an A record with a geolocation policy.

C. Create a CNAME record with a failover policy.

D. Create a CNAME record with a geoproximity policy.

Answer:

A

Explanation:

To provide the most high-performing experience for a multi-region application, the primary goal is to minimize latency for end-users. Amazon Route 53's latency-based routing policy is specifically designed for this purpose. It evaluates the network latency between the user and the AWS Regions where the application is deployed. Route 53 then responds to DNS queries with the IP address (via an Alias A record pointing to the Application Load Balancer) for the region that offers the lowest latency to that user. This ensures traffic is directed to the regional endpoint that can serve the user the fastest, directly optimizing for performance.

References:

1. AWS Documentation - Route 53 Developer Guide: "Choosing a routing policy".
Section: Latency-based routing.
Quote: "Use latency-based routing when you have resources in multiple AWS Regions and you want to route traffic to the region that provides the best latency... Route 53 responds with the resource for the region that provides the best latency." This directly supports the choice of latency policy for the best performance.
Reference: https://docs.aws.amazon.com/Route53/latest/DeveloperGuide/routing-policy.html#routing-policy-latency
2. AWS Documentation - Route 53 Developer Guide: "Choosing between alias and non-alias records".
Section: "Application and Classic Load Balancers".
Quote: "We recommend that you use an alias record to route traffic to an ELB load balancer." This confirms that an A record (specifically, an Alias A record) is the correct record type for an Application Load Balancer, not a CNAME.
Reference: https://docs.aws.amazon.com/Route53/latest/DeveloperGuide/resource-record-sets-choosing-alias-non-alias.html

Question: 25

A company has an application that delivers on-demand training videos to students around the world.
The application also allows authorized content developers to upload videos. The data is stored in an Amazon S3 bucket in the us-east-2 Region.

The company has created an S3 bucket in the eu-west-2 Region and an S3 bucket in the ap-southeast-1 Region. The company wants to replicate the data to the new S3 buckets.
The company needs to minimize latency for developers who upload videos and students who stream videos near eu-west-2 and ap-southeast-1.

Which combination of steps will meet these requirements with the FEWEST changes to the application? (Select TWO.)

A. Configure one-way replication from the us-east-2 S3 bucket to the eu-west-2 S3 bucket. Configure
one-way replication from the us-east-2 S3 bucket to the ap-southeast-1 S3 bucket.

B. Configure one-way replication from the us-east-2 S3 bucket to the eu-west-2 S3 bucket. Configure
one-way replication from the eu-west-2 S3 bucket to the ap-southeast-1 S3 bucket.

C. Configure two-way (bidirectional) replication among the S3 buckets that are in all three Regions.

D. Create an S3 Multi-Region Access Point. Modify the application to use the Amazon Resource Name
(ARN) of the Multi-Region Access Point for video streaming. Do not modify the application for video
uploads.

E. Create an S3 Multi-Region Access Point Modify the application to use the Amazon Resource Name
(ARN) of the Multi-Region Access Point for video streaming and uploads.

Answer:

C, E

Explanation:

To meet the requirements of minimizing latency for both global video streaming (downloads) and developer uploads (writes) with minimal application changes, a combination of S3 Multi-Region Access Points (MRAP) and bidirectional replication is required.

An S3 Multi-Region Access Point provides a single global endpoint. It automatically routes client requests to the S3 bucket with the lowest network latency, addressing the core requirement for both students and developers. Using the MRAP for both uploads and streaming requires only a single endpoint change in the application.

For an MRAP to function correctly when writes can occur in any region, the underlying S3 buckets must be kept synchronized. This necessitates a two-way (bidirectional) replication configuration among all three buckets to ensure that a video uploaded in one region is replicated to all other regions, maintaining data consistency.

References:

1. AWS S3 User Guide, "Multi-Region Access Points in Amazon S3": This guide explains that MRAPs provide a global endpoint to route requests to the lowest-latency bucket. It states, "When you create a Multi-Region Access Point, you can also have S3 create the replication rules for you. S3 creates a replication rule that replicates all objects between your buckets." This implies a bidirectional or multi-way replication setup is the standard for this use case.
2. AWS S3 User Guide, "Creating Multi-Region Access Points": The documentation details the process, noting: "To keep the contents of your buckets synchronized, we recommend that you configure S3 Replication." This confirms that replication is a foundational component for a functional MRAP that serves consistent data.
3. AWS S3 User Guide, "Replication": This section describes replication configurations. For an active-active architecture where writes can occur in any region (as enabled by an MRAP for uploads), a configuration where each bucket is both a source and a destination for others is required to maintain synchronization. This is functionally equivalent to two-way or bidirectional replication.

Question: 26

An analytics company uses Amazon VPC to run its multi-tier services. The company wants to use RESTful APIs to offer a web analytics service to millions of users. Users must be verified by using an authentication service to access the APIs.

Which solution will meet these requirements with the MOST operational efficiency?

A. Configure an Amazon Cognito user pool for user authentication. Implement Amazon API Gateway
REST APIs with a Cognito authorizer.

B. Configure an Amazon Cognito identity pool for user authentication. Implement Amazon API Gateway HTTP APIs with a Cognito authorizer.

C. Configure an AWS Lambda function to handle user authentication. Implement Amazon API Gateway REST APIs with a Lambda authorizer.

D. Configure an IAM user to handle user authentication. Implement Amazon API Gateway HTTP APIs
with an IAM authorizer.

Answer:

A

Explanation:

This solution is the most operationally efficient for managing and authenticating millions of users for a RESTful API. Amazon Cognito User Pools are specifically designed as a scalable, fully managed user directory and authentication service, handling user sign-up, sign-in, and token management. Amazon API Gateway REST APIs integrate natively with Cognito User Pools through a "Cognito authorizer." This built-in integration requires minimal configuration and code, offloading the complex and critical work of authentication to managed AWS services, thereby maximizing operational efficiency as required by the scenario.

References:

1. Amazon Cognito Developer Guide, "What is Amazon Cognito?": This guide distinguishes between User Pools and Identity Pools. It states, "Amazon Cognito user pools are user directories that provide sign-up and sign-in options for your app users... With an identity pool, your users can obtain temporary AWS credentials to access other AWS services." This supports why option A (User Pools for authentication) is correct and B (Identity Pools for authentication) is incorrect.
2. Amazon API Gateway Developer Guide, "Control access to a REST API using Amazon Cognito user pools": This document explicitly describes the target use case. It states, "You can use an Amazon Cognito user pool to control who can access your API in Amazon API Gateway...

After a user signs in, your web or mobile app sends the resulting identity or access token to your API Gateway REST API in an authorization header. API Gateway uses the token to authenticate the call." This directly validates the architecture in option A.
3. Amazon API Gateway Developer Guide, "Controlling and managing access to a REST API in API Gateway": This section details the different authorizer types. It describes the "Amazon Cognito user pool authorizer" as a managed feature for REST APIs, contrasting it with the custom logic required for a "Lambda authorizer" (Option C) and the specific use case of the "IAM authorizer" for AWS credentials (Option D). This highlights the operational efficiency of the Cognito authorizer.

Question: 27

A company is designing a tightly coupled high performance computing (HPC) environment in the AWS Cloud The company needs to include features that will optimize the HPC environment for networking and storage.
Which combination of solutions will meet these requirements? (Select TWO)

A. Create an accelerator in AWS Global Accelerator. Configure custom routing for the accelerator.

B. Create an Amazon FSx for Lustre file system. Configure the file system with scratch storage.

C. Create an Amazon CloudFront distribution. Configure the viewer protocol policy to be HTTP and
HTTPS.

D. Launch Amazon EC2 instances. Attach an Elastic Fabric Adapter (EFA) to the instances.

E. Create an AWS Elastic Beanstalk deployment to manage the environment.

Answer:

B, D

Explanation:

A tightly coupled High-Performance Computing (HPC) environment requires two key optimizations: low-latency, high-throughput networking for inter-node communication, and a high-performance parallel storage system for massive I/O operations.
An Elastic Fabric Adapter (EFA) is a specialized network interface for Amazon EC2 instances designed specifically for HPC. It uses a custom OS-bypass protocol to provide lower, more consistent latency and higher throughput, which is essential for tightly coupled workloads that rely on Message Passing Interface (MPI).
Amazon FSx for Lustre is a fully managed, high-performance file system based on the Lustre parallel file system, which is a standard in HPC. It provides the massive throughput and low-latency access to shared data required by compute-intensive workloads.

References:

1. Elastic Fabric Adapter (EFA): AWS Documentation, Amazon EC2 User Guide for Linux Instances, Section: "Elastic Fabric Adapter". The documentation states, "Elastic Fabric Adapter (EFA) is a network interface for Amazon EC2 instances that enables customers to run applications requiring high levels of inter-node communications at scale on AWS... It is optimized to work on the current generation of Amazon EC2 infrastructure and can scale depending on the instance type."
2. Amazon FSx for Lustre: AWS Documentation, Amazon FSx for Lustre User Guide, Section: "What Is Amazon FSx for Lustre?". The guide explains, "Amazon FSx for Lustre is a fully

managed service that provides cost-effective, high-performance, scalable storage for compute workloads... It's ideal for high-performance computing (HPC), machine learning, media and entertainment workflows, and financial analytics."
3. HPC on AWS Overview: AWS Documentation, High Performance Computing on AWS, "Networking" and "Storage" sections. This page explicitly lists EFA for "tightly coupled workloads" under Networking and Amazon FSx for Lustre as a key "high performance storage" solution for HPC.

Question: 28

A gaming company wants to launch a new internet-facing application in multiple AWS Regions The application will use the TCP and UDP protocols for communication. The company needs to provide high availability and minimum latency for global users.

Which combination of actions should a solutions architect take to meet these requirements? (Select
TWO.)

A. Create internal Network Load Balancers in front of the application in each Region.

B. Create external Application Load Balancers in front of the application in each Region.

C. Create an AWS Global Accelerator accelerator to route traffic to the load balancers in each Region.

D. Configure Amazon Route 53 to use a geolocation routing policy to distribute the traffic.

E. Configure Amazon CloudFront to handle the traffic and route requests to the application in each
Region.

Answer:

A, C

Explanation:

This scenario requires a global, low-latency, and highly available solution for a gaming application that uses both TCP and UDP protocols.
AWS Global Accelerator is the ideal service for the global routing layer. It provides static anycast IP addresses that act as a fixed entry point and routes user traffic over the AWS global network to the nearest healthy regional endpoint. This significantly reduces latency and improves performance for global users.
In each region, a Network Load Balancer (NLB) is required. NLBs operate at the transport layer (Layer 4) and are designed to handle high-throughput TCP and UDP traffic, which is essential for the gaming application. Global Accelerator can direct traffic to internal NLBs, which then distribute it to the application instances, providing a secure and scalable architecture.

References:

1. AWS Global Accelerator, Features: "AWS Global Accelerator is a good fit for non-HTTP use cases, such as gaming (UDP), IoT (MQTT), or Voice over IP... It uses the vast, congestion-free AWS global network to route TCP and UDP traffic to a healthy application endpoint in the closest

AWS Region to your user."
Source: AWS Global Accelerator Official Product Page, "Features" section.
2. Elastic Load Balancing, Listeners for your Network Load Balancers: "A listener is a process that checks for connection requests, using the protocol and port that you configure... Network Load Balancers support the following protocols: TCP, UDP, TLS, and TCPUDP."
Source: AWS Documentation, User Guide for Network Load Balancers, "Listeners for your Network Load Balancers" section.
3. Elastic Load Balancing, Listeners for your Application Load Balancers: "An Application Load Balancer functions at the application layer, the seventh layer of the Open Systems Interconnection (OSI) model... You can configure listeners for the HTTP and HTTPS protocols..."
Source: AWS Documentation, User Guide for Application Load Balancers, "Listeners for your Application Load Balancers" section.
4. AWS Global Accelerator, Endpoints: "For a standard accelerator, you can add Network Load Balancers, Application Load Balancers, EC2 instances, or Elastic IP addresses as endpoints." This confirms that NLBs (both internal and internet-facing) are valid endpoints for Global Accelerator.
Source: AWS Documentation, AWS Global Accelerator Developer Guide, "Endpoints in AWS Global Accelerator" section.

Question: 29

A company is running a legacy system on an Amazon EC2 instance. The application code cannot be modified, and the system cannot run on more than one instance. A solutions architect must design a resilient solution that can improve the recovery time for the system.

What should the solutions architect recommend to meet these requirements?

A. Enable termination protection for the EC2 instance.

B. Configure the EC2 instance for Multi-AZ deployment.

C. Create an Amazon CloudWatch alarm to recover the EC2 instance in case of failure.

D. Launch the EC2 instance with two Amazon Elastic Block Store (Amazon EBS) volumes that use RAID
configurations for storage redundancy.

Answer:

C

Explanation:

The most effective way to improve recovery time for a single EC2 instance, without modifying the application or using multiple instances, is to automate the recovery process. An Amazon CloudWatch alarm can be configured to monitor the instance's system status check (StatusCheckFailedSystem). If this check fails, indicating an issue with the underlying hardware, the alarm can trigger an EC2 "recover" action. This action automatically migrates the instance to new, healthy hardware while preserving its instance ID, private IP address, Elastic IP address, and all attached EBS volume data. This directly addresses the need for a resilient solution that minimizes recovery time.

References:

1. Amazon EC2 User Guide for Linux Instances, "Recover your instance": "You can create an Amazon CloudWatch alarm that monitors an Amazon EC2 instance and automatically recovers the instance if it becomes impaired due to an underlying hardware failure or a problem that requires AWS involvement to repair. A recovered instance is identical to the original instance, including the instance ID, private IP addresses, Elastic IP addresses, and all instance metadata."
2. Amazon EC2 User Guide for Linux Instances, "Types of status checks": This document explains that a "System status check" monitors the AWS systems on which the instance runs. A failure indicates a problem with the underlying host. The recover action is the recommended automated solution for this type of failure.
3. Amazon EC2 User Guide for Linux Instances, "Enable termination protection": This section

clarifies that termination protection's sole purpose is to prevent an instance from being terminated accidentally. It does not contribute to recovery from operational failures.
4. Amazon EC2 User Guide for Linux Instances, "RAID on Amazon EBS volumes": This documentation describes how RAID configurations can be used to improve the performance or reliability of the storage subsystem, but it does not address the recovery of the compute instance itself.

Question: 30

A solutions architect wants to use the following JSON text as an identity-based policy to grant specific permissions:

Which IAM principals can the solutions architect attach this policy to? (Select TWO.)

A. Role

B. Group

C. Organization

D. Amazon Elastic Container Service (Amazon ECS) resource

E. Amazon EC2 resource

Answer:

A, B

Explanation:

Identity-based policies in AWS IAM are JSON documents that define permissions for an identity. These policies are attached directly to IAM principals (identities). The primary IAM principals are users, user groups, and roles. When a policy is attached to a group, all users within that group inherit the permissions. When attached to a role, any entity that assumes the role receives the defined permissions. The provided JSON is a standard identity-based policy, which can be attached to IAM groups and IAM roles.

References:

1. AWS IAM User Guide: "Identity-based policies and resource-based policies". Under the "Identity-based policies" section, it states, "You can attach identity-based policies to IAM users, groups, and roles." This directly confirms that roles and groups are valid targets.
2. AWS IAM User Guide: "IAM identities (users, user groups, and roles)". This document defines users, groups, and roles as the IAM identities to which you attach policies to provide them with permissions to access AWS resources.
3. AWS Organizations User Guide: "Service control policies (SCPs)". It clarifies, "SCPs are a type of organization policy... SCPs don't grant permissions. Instead, SCPs are JSON policies that specify the maximum permissions for the affected accounts." This distinguishes SCPs from the identity-based policy in the question.
4. Amazon EC2 User Guide for Linux Instances: "IAM roles for Amazon EC2". The documentation states, "To grant permissions to applications running on EC2 instances, you use an IAM role." This shows that policies are not attached directly to the EC2 resource itself.

Question: 31

A company hosts a database that runs on an Amazon RDS instance that is deployed to multiple Availability Zones. The company periodically runs a script against the database to report new entries that are added to the database. The script that runs against the database negatively affects the performance of a critical application. The company needs to improve application performance with minimal costs.

Which solution will meet these requirements with the LEAST operational overhead?

A. Add functionality to the script to identify the instance that has the fewest active connections. Configure the script to read from that instance to report the total new entries.

B. Create a read replica of the database. Configure the script to query only the read replica to report
the total new entries.

C. Instruct the development team to manually export the new entries for the day in the database at
the end of each day.

D. Use Amazon ElastiCache to cache the common queries that the script runs against the database.

Answer:

B

Explanation:

The core issue is that a reporting script's read-heavy workload is degrading the performance of a critical application using the same primary database. The most effective and operationally simple solution is to offload this read traffic. An Amazon RDS Read Replica is a separate, read-only copy of the primary database instance created specifically for this purpose. By directing the reporting script to the read replica's endpoint, the read load is isolated from the primary instance, which immediately improves the performance of the critical application. Creating a read replica and updating a connection string in a script represents the least operational overhead among the viable options.

References:

1. Amazon RDS User Guide, "High availability (Multi-AZ) for Amazon RDS": This document clarifies the role of the standby instance. It states, "The standby replica can't serve read traffic. Multi-AZ deployments are designed for high availability and durability, rather than read scaling." This directly invalidates option A.

2. Amazon RDS User Guide, "Working with read replicas": This source explains the primary use case for read replicas. It states, "You can reduce the load on your source DB instance by routing read queries from your applications to the read replica." This directly supports option B as the intended solution for offloading read-intensive workloads like reporting.
3. AWS Well-Architected Framework, Performance Efficiency Pillar, "PERF 5: How do you select the best performing architecture?": The documentation discusses database scaling strategies. Under "Offload the database," it recommends using "read replicas to serve read-only traffic." This aligns with the best practice of using read replicas (Option B) to improve performance.
4. Amazon ElastiCache for Redis User Guide, "Caching strategies": This guide details caching patterns. The scenario of querying for new, non-existent data is the opposite of a cache-hit scenario, making caching an inappropriate solution. This supports the reasoning for invalidating option D.

Question: 32

A company has an organization in AWS Organizations. The company runs Amazon EC2 instances across four AWS accounts in the root organizational unit (OU). There are three nonproduction accounts and one production account. The company wants to prohibit users from launching EC2 instances of a certain size in the nonproduction accounts. The company has created a service control policy (SCP) to deny access to launch instances that use the prohibited types. Which solutions to deploy the SCP will meet these requirements? (Select TWO.)

A. Attach the SCP to the root OU for the organization.

B. Attach the SCP to the three nonproduction Organizations member accounts.

C. Attach the SCP to the Organizations management account.

D. Create an OU for the production account. Attach the SCP to the OU. Move the production member
account into the new OU.

E. Create an OU for the required accounts. Attach the SCP to the OU. Move the nonproduction member accounts into the new OU.

Answer:

B, E

Explanation:

The objective is to apply a restrictive Service Control Policy (SCP) exclusively to the three non-production accounts, ensuring the production account remains unaffected. This can be accomplished in two ways. The first method is to attach the SCP directly to each of the three non-production member accounts. The second, and more scalable, method is to create a new Organizational Unit (OU), move the three non-production accounts into this OU, and then attach the SCP to the OU. All accounts within the OU inherit the policy, which is a best practice for managing groups of accounts with common governance requirements. Both solutions correctly isolate the policy's effect to the intended accounts.

References:

1. AWS Organizations User Guide, "Attaching and detaching service control policies": This document states, "You can attach an SCP to any of the following: The organization root, An organizational unit (OU), An account". This directly supports the validity of attaching the policy to individual accounts (Option B) or to an OU (Option E).
2. AWS Organizations User Guide, "Inheritance for service control policies": This section explains,

"When you attach an SCP to a specific root or OU, it's inherited by all the OUs and accounts in that root or OU." This principle confirms that attaching the SCP to the root (Option A) would impact the production account, while attaching it to a dedicated OU containing only non-production accounts (Option E) would correctly limit its scope.
3. AWS Organizations User Guide, "Example service control policies": The section "Prevent member accounts from leaving the organization" notes that "SCPs don't affect users or roles in the management account. They affect only the member accounts in your organization." This clarifies why attaching the policy to the management account (Option C) is ineffective for controlling member accounts.
4. AWS Whitepaper, "Organizing Your AWS Environment Using Multiple Accounts", Page 11, "Organizational Units": This paper recommends grouping accounts into OUs based on function or environment (e.g., "Workloads OU" with "Prod" and "Dev" child OUs). This endorses the strategy in Option E as a best practice for applying targeted policies.

Question: 33

A company wants to migrate an on-premises legacy application to AWS. The application ingests customer order files from an on-premises enterprise resource planning (ERP) system. The application then uploads the files to an SFTP server. The application uses a scheduled job that checks for order files every hour.

The company already has an AWS account that has connectivity to the on-premises network.
The new application on AWS must support integration with the existing ERP system. The new application must be secure and resilient and must use the SFTP protocol to process orders from the ERP system immediately.
Which solution will meet these requirements?

A. Create an AWS Transfer Family SFTP internet-facing server in two Availability Zones. Use Amazon
S3 storage. Create an AWS Lambda function to process order files. Use S3 Event Notifications to send

s3: ObjectCreated: * events to the Lambda function.

B. Create an AWS Transfer Family SFTP internet-facing server in one Availability Zone. Use Amazon
Elastic File System (Amazon EFS) storage. Create an AWS Lambda function to process order files. Use
a Transfer Family managed workflow to invoke the Lambda function.

C. Create an AWS Transfer Family SFTP internal server in two Availability Zones. Use Amazon Elastic
File System (Amazon EFS) storage. Create an AWS Step Functions state machine to processorder files.
Use Amazon EventBridge Scheduler to invoke the state machine to periodically check Amazon EFS for
order files.

D. Create an AWS Transfer Family SFTP internal server in two Availability Zones. Use Amazon S3
storage. Create an AWS Lambda function to process order files. Use a Transfer Family managed workflow to invoke the Lambda function.

Answer:

D

Explanation:

This solution correctly addresses all the stated requirements. An AWS Transfer Family server with an internal endpoint type is secure, as it is only accessible from within the VPC or connected on-premises networks, not the public internet. Deploying the server endpoint in two Availability Zones provides the required resilience. Amazon S3 is a durable and scalable storage backend for the ingested files. A Transfer Family managed workflow is the ideal mechanism to process files immediately upon successful upload, triggering an AWS Lambda function to handle the order processing logic in an event-driven manner.

References:

1. AWS Transfer Family User Guide - Endpoint types: "For SFTP and FTPS servers, you can host your server's endpoint within your virtual private cloud (VPC). In this case, the server is accessible only to clients within your VPC, or clients that can access your VPC, for example, through AWS Direct Connect or AWS VPN." This supports the use of an internal server for security. (Source: AWS Transfer Family User Guide, Section: "Choose an endpoint type").
2. AWS Transfer Family User Guide - High Availability: "For VPC-hosted endpoints, you can achieve high availability by creating your server with multiple hostnames (one in each of your chosen Availability Zones)." This supports the multi-AZ deployment for resilience. (Source: AWS Transfer Family User Guide, Section: "High Availability for AWS Transfer Family").
3. AWS Transfer Family User Guide - Managed workflows: "You can use managed workflows in AWS Transfer Family to process files that are uploaded using the service... A workflow can contain steps to copy, tag, and delete files... You can also create a workflow step that invokes a specific AWS Lambda function." This confirms the ability to process files immediately upon successful upload. (Source: AWS Transfer Family User Guide, Section: "Working with managed workflows").
4. AWS Documentation - Amazon S3: "Amazon S3 is designed for 99.999999999% (11 9's) of durability, and stores data for millions of applications for companies all around the world." This validates S3 as a durable and resilient storage choice. (Source: Amazon S3 product page, "Benefits").

Question: 34

A company needs to provide customers with secure access to its data. The company processes customer data and stores the results in an Amazon S3 bucket. All the data is subject to strong regulations and security requirements. The data must be encrypted at rest. Each customer must be able to access only their data from their AWS account. Company employees must not be able to access the data.
Which solution will meet these requirements?

A. Provision an AWS Certificate Manager (ACM) certificate for each customer. Encrypt the data client-
side. In the private certificate policy, deny access to the certificate for all principals except an IAM role that the customer provides.

B. Provision a separate AWS Key Management Service (AWS KMS) key for each customer. Encrypt the
data server-side. In the S3 bucket policy, deny decryption of data for all principals except an IAM role
that the customer provides.

C. Provision a separate AWS Key Management Service (AWS KMS) key for each customer. Encrypt the

data server-side. In each KMS key policy, deny decryption of data for all principals except an IAM role
that the customer provides.

D. Provision an AWS Certificate Manager (ACM) certificate for each customer. Encrypt the data client-
side. In the public certificate policy, deny access to the certificate for all principals except an IAM role
that the customer provides.

Answer:

C

Explanation:

This solution correctly implements a multi-tenant data isolation strategy using AWS Key Management Service (AWS KMS). Provisioning a separate KMS key for each customer ensures strong cryptographic separation. The KMS key policy is the primary access control mechanism for the key itself. By creating a key policy that explicitly denies all principals except for the customer's specific cross-account IAM role the kms:Decrypt permission, the solution meets all requirements: data is encrypted at rest (via SSE-KMS), each customer can only access their data, and the

company's own employees are prevented from decrypting the sensitive information.

References:

1. AWS KMS Developer Guide, "Key policies in AWS KMS": "The key policy is the primary way to control access to your KMS keys... You must use the key policy to give other AWS accounts and their users permission to use a KMS key." This confirms that the KMS key policy is the correct mechanism for managing cross-account permissions.
2. AWS KMS Developer Guide, "Allowing users in other accounts to use a KMS key": This section details the exact process described in the correct answer. It states, "To allow users in a different AWS account to use a KMS key, you can use the key policy... In the key policy, add a policy statement that allows an external account or its users and roles to use the key."
3. Amazon S3 User Guide, "Protecting data using server-side encryption with AWS Key Management Service keys (SSE-KMS)": "When you use SSE-KMS, you can use the default AWS managed key... or you can specify a customer managed key that you have already created... If you use a customer managed key, you must grant the user permissions for the key." This establishes that permissions to the KMS key are separate from S3 permissions and must be managed accordingly.

Question: 35

A company wants to rearchitect a large-scale web application to a serverless microservices architecture. The application uses Amazon EC2 instances and is written in Python.
The company selected one component of the web application to test as a microservice. The component supports hundreds of requests each second. The company wants to create and test the microservice on an AWS solution that supports Python. The solution must also scale automatically and require minimal infrastructure and minimal operational support. Which solution will meet these requirements?

A. Use a Spot Fleet with auto scaling of EC2 instances that run the most recent Amazon Linux operating system.

B. Use an AWS Elastic Beanstalk web server environment that has high availability configured.

C. Use Amazon Elastic Kubernetes Service (Amazon EKS). Launch Auto Scaling groups of self-managed EC2 instances.

D. Use an AWS Lambda function that runs custom developed code.

Answer:

D

Explanation:

The core requirements are to build a serverless microservice that supports Python, scales automatically, and requires minimal infrastructure and operational support. AWS Lambda is a serverless, event-driven compute service that perfectly aligns with these needs. It natively supports Python, automatically manages all compute resources, and scales precisely with the number of requests, from a few to thousands per second. This eliminates the need for provisioning or managing servers, patching operating systems, or configuring scaling policies, thus fulfilling the "minimal operational support" requirement for a microservices architecture.

References:

1. AWS Lambda - Developer Guide: "AWS Lambda is a serverless, event-driven compute service that lets you run code for virtually any type of application or backend service without provisioning or managing servers. You can trigger Lambda from over 200 AWS services and software as a service (SaaS) applications, and only pay for what you use." (Source: AWS Lambda Developer Guide, "What is AWS Lambda?", Introduction section).
2. AWS Lambda - Features: "With Lambda, you can run code for virtually any type of application or backend service with zero administration. Just upload your code and Lambda handles everything required to run and scale your code with high availability." (Source: AWS Lambda

product page, "Features" section, "Run code with zero administration" subsection).
3. AWS Elastic Beanstalk - Developer Guide: "With Elastic Beanstalk, you can quickly deploy and manage applications in the AWS Cloud without having to learn about the infrastructure that runs those applications. AWS Elastic Beanstalk reduces management complexity without restricting choice or control. You simply upload your application, and Elastic Beanstalk automatically handles the details of capacity provisioning, load balancing, scaling, and application health monitoring." (Source: AWS Elastic Beanstalk Developer Guide, "What is AWS Elastic Beanstalk?", Introduction. This highlights it as a PaaS managing infrastructure, not a serverless compute service like Lambda).
4. Amazon EKS - User Guide: "Amazon EKS is a managed service that you can use to run Kubernetes on AWS without needing to install, operate, and maintain your own Kubernetes control plane or nodes." The option specifies "self-managed EC2 instances," which requires significant operational effort, unlike a serverless model. (Source: Amazon EKS User Guide, "What is Amazon EKS?", Introduction).