

Question 1: **Correct**

A company has deployed a SAML 2.0 federated identity solution with their on-premises identity provider (IdP) to authenticate users' access to the AWS environment. A Solutions Architect ran authentication tests through the federated identity web portal and access to the AWS environment was granted. When a test user attempts to authenticate through the federated identity web portal, they are not able to access the AWS environment.

Which items should the solutions architect check to ensure identity federation is properly configured? (Select THREE.)

-

The IAM users are providing the time-based one-time password (TOTP) codes required for authenticated access.

-

The IAM users permissions policy has allowed the sts:AssumeRoleWithSAML API action allowed in their permissions policy.

-

The IAM roles created for the federated users' or federated groups' trust policy have set the SAML provider as the principal.

(Correct)

-

The company's IdP defines SAML assertions that properly map users or groups in the company to IAM roles with appropriate permissions.

(Correct)

-

The web portal calls the AWS STS AssumeRoleWithSAML API with the ARN of the SAML provider, the ARN of the IAM role, and the SAML assertion from IdP.

(Correct)

-

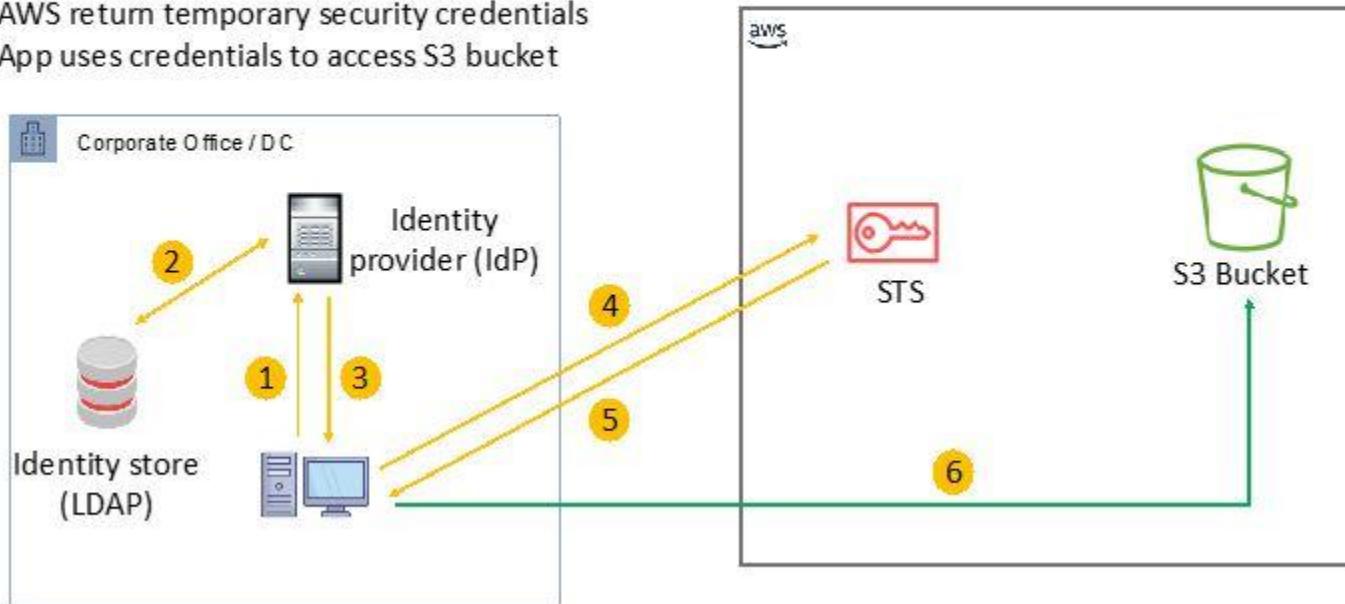
The AWS STS service has the on-premises IdP configured as an event source for authentication requests.

Explanation

AWS supports identity federation with SAML 2.0 (Security Assertion Markup Language 2.0), an open standard that many identity providers (IdPs) use. This feature enables federated single sign-on (SSO), so users can log into the AWS Management Console or call the AWS API operations without you having to create an IAM user for everyone in your organization.

In the diagram below the process of authentication is depicted in a situation where a client is authorized temporary access to an Amazon S3 bucket.

1. Client application attempts to authenticate using IdP
2. IdP authenticates the user
3. IdP sends client SAML assertion
4. App calls `sts:AssumeRoleWithSAML`
5. AWS return temporary security credentials
6. App uses credentials to access S3 bucket



The correct answers are validated based on the following facts:

- During step 4, the client app calls the AWS STS [AssumeRoleWithSAML](#) API, passing the ARN of the SAML provider, the ARN of the role to assume, and the SAML assertion from IdP.
- When configuring the IdP and AWS to trust each other, in IAM, you create one or more IAM roles. In the role's trust policy, you set the SAML provider as the principal, which establishes a trust relationship between your organization and AWS. And...

- In your organization's IdP, you define assertions that map users or groups in your organization to the IAM roles.

CORRECT: "The IAM roles created for the federated users' or federated groups' trust policy have set the SAML provider as the principal" is a correct answer.

CORRECT: "The web portal calls the AWS STS AssumeRoleWithSAML API with the ARN of the SAML provider, the ARN of the IAM role, and the SAML assertion from IdP" is also a correct answer.

CORRECT: "The company's IdP defines SAML assertions that properly map users or groups in the company to IAM roles with appropriate permissions" is also a correct answer.

INCORRECT: "The IAM users are providing the time-based one-time password (TOTP) codes required for authenticated access" is incorrect. TOTPs are used with multi-factor authentication which is not included in this solution.

INCORRECT: "The IAM users permissions policy has allowed the sts:AssumeRoleWithSAML API action allowed in their permissions policy" is incorrect. Users need to have permissions to access the role; the role that is being assumed must be allowed to federate using SAML as it is the role that performs the sts:AssumeRoleWithSAML action.

INCORRECT: "The AWS STS service has the on-premises IdP configured as an event source for authentication requests" is incorrect. It is not necessary to configure the AWS STS service with event sources.

References:

https://docs.aws.amazon.com/IAM/latest/UserGuide/id_roles_providers_saml.html

Save time with our AWS cheat sheets:

<https://digitalcloud.training/category/aws-cheat-sheets/aws-solutions-architect-professional/security-identity-compliance-sap/>

Question 2: **Correct**

A financial services company receives a data feed from a credit card service provider. The feed consists of approximately 2,500 records that are sent every 10 minutes in plaintext and delivered over HTTPS to an encrypted S3 bucket. The data includes credit card data that must be automatically masked before sending the data to another S3 bucket for additional internal processing. There is also a requirement to remove and merge specific fields, and then transform the record into JSON format.

Which solutions will meet these requirements?

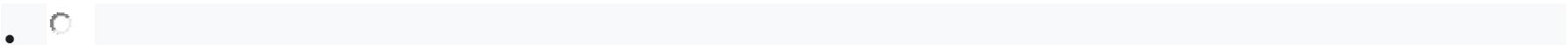
Trigger an AWS Lambda function on file delivery that extracts each record and writes it to an Amazon SQS queue. Configure an AWS Fargate container application to automatically scale to a single instance when the SQS queue contains messages. Have the application process each record and transform the record into JSON format. When the queue is empty, send the results to another S3 bucket for internal processing and scale down the AWS Fargate task.

Create an AWS Glue crawler and custom classifier based on the data feed formats and build a table definition to match. Trigger an AWS Lambda function on file delivery to start an AWS Glue ETL job to transform the entire record according to the processing and transformation requirements. Define the output format as JSON. Once complete, have the ETL job send the results to another S3 bucket for internal processing.

(Correct)



Trigger an AWS Lambda function on file delivery that extracts each record and writes it to an Amazon SQS queue. Trigger another Lambda function when new messages arrive in the SQS queue to process the records, writing the results to a temporary location in Amazon S3. Trigger a final Lambda function once the SQS queue is empty to transform the records into JSON format and send the results to another S3 bucket for internal processing.



Create an AWS Glue crawler and custom classifier based upon the data feed formats and build a table definition to match. Perform an Amazon Athena query on file delivery to start an Amazon EMR ETL job to transform the entire record according to the processing and transformation requirements. Define the output format as JSON. Once complete, send the results to another S3 bucket for internal processing and scale down the EMR cluster.

Explanation

You can use a Glue crawler to populate the AWS Glue Data Catalog with tables. The Lambda function can be triggered using S3 event notifications when object create events occur. The Lambda function will then trigger the Glue ETL job to transform the records masking the sensitive data and modifying the output format to JSON. This solution meets all requirements.

CORRECT: "Create an AWS Glue crawler and custom classifier based on the data feed formats and build a table definition to match. Trigger an AWS Lambda function on file delivery to start an AWS Glue ETL job to transform the entire record according to the processing and transformation requirements. Define the output format as JSON. Once complete, have the ETL job send the results to another S3 bucket for internal processing" is the correct answer.

INCORRECT: "Trigger an AWS Lambda function on file delivery that extracts each record and writes it to an Amazon SQS queue. Trigger another Lambda function when new messages arrive in the SQS queue to process the records, writing the results to a temporary location in Amazon S3. Trigger a final Lambda function once the SQS queue is empty to transform the records into JSON format and send the results to another S3 bucket for internal processing" is incorrect. AWS Glue is an ETL service and is therefore a better fit for processing the records as part of an ETL job rather than using Lambda.

INCORRECT: "Trigger an AWS Lambda function on file delivery that extracts each record and writes it to an Amazon SQS queue. Configure an AWS Fargate container application to automatically scale to a single instance when the SQS queue contains messages. Have the application process each record and transform the record into JSON format. When the queue is empty, send the results to another S3 bucket for internal processing and scale down the AWS Fargate task" is incorrect. AWS Glue is preferred for ETL work. Also, Lambda is more scalable and will be faster to respond than using a single Fargate task.

INCORRECT: "Create an AWS Glue crawler and custom classifier based upon the data feed formats and build a table definition to match. Perform an Amazon Athena query on file delivery to start an Amazon EMR ETL job to transform the entire record according to the processing and transformation requirements. Define the output format as JSON. Once complete, send the results to another S3 bucket for internal processing and scale down the EMR cluster" is incorrect. Amazon EMR is not a service that is used for ETL jobs, AWS Glue should be used instead.

References:

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

<https://docs.aws.amazon.com/glue/latest/dg/author-job.html>

Save time with our AWS cheat sheets:

<https://digitalcloud.training/category/aws-cheat-sheets/aws-solutions-architect-professional/aws-compute-sap/>

<https://digitalcloud.training/category/aws-cheat-sheets/aws-solutions-architect-professional/aws-analytics-sap/>

Question 3: **Correct**

A company has deployed an eCommerce application that is used by thousands of customers to place online orders. The application runs on Amazon ECS tasks behind an Application Load Balancer (ALB) and data is stored in an Amazon DynamoDB table.

The application has recently experienced attacks that caused application slowdowns and outages. The company must prevent attacks and ensure business continuity with minimal service interruptions.

Which combination of steps will meet these requirements MOST cost-effectively? (Select TWO.)

-

Create an Amazon CloudFront distribution with the ALB as the origin and configure a custom header and secret value. Configure the ALB to conditionally forward traffic only if the header and value match.

(Correct)

-

Deploy an AWS WAF web ACL that includes a rule group that blocks the attack traffic. Associate the web ACL with the Amazon CloudFront distribution.

(Correct)

-

Configure AWS Auto Scaling for Amazon ECS tasks. Create an Amazon DynamoDB Accelerator (DAX) cluster in front of the DynamoDB table.

-

Configure AWS Auto Scaling for Amazon ECS tasks. Configure an Amazon ElastiCache cluster in front of the DynamoDB table.

-

Deploy the application in two AWS Regions. Configure Amazon Route 53 to route to both Regions with equal weight.

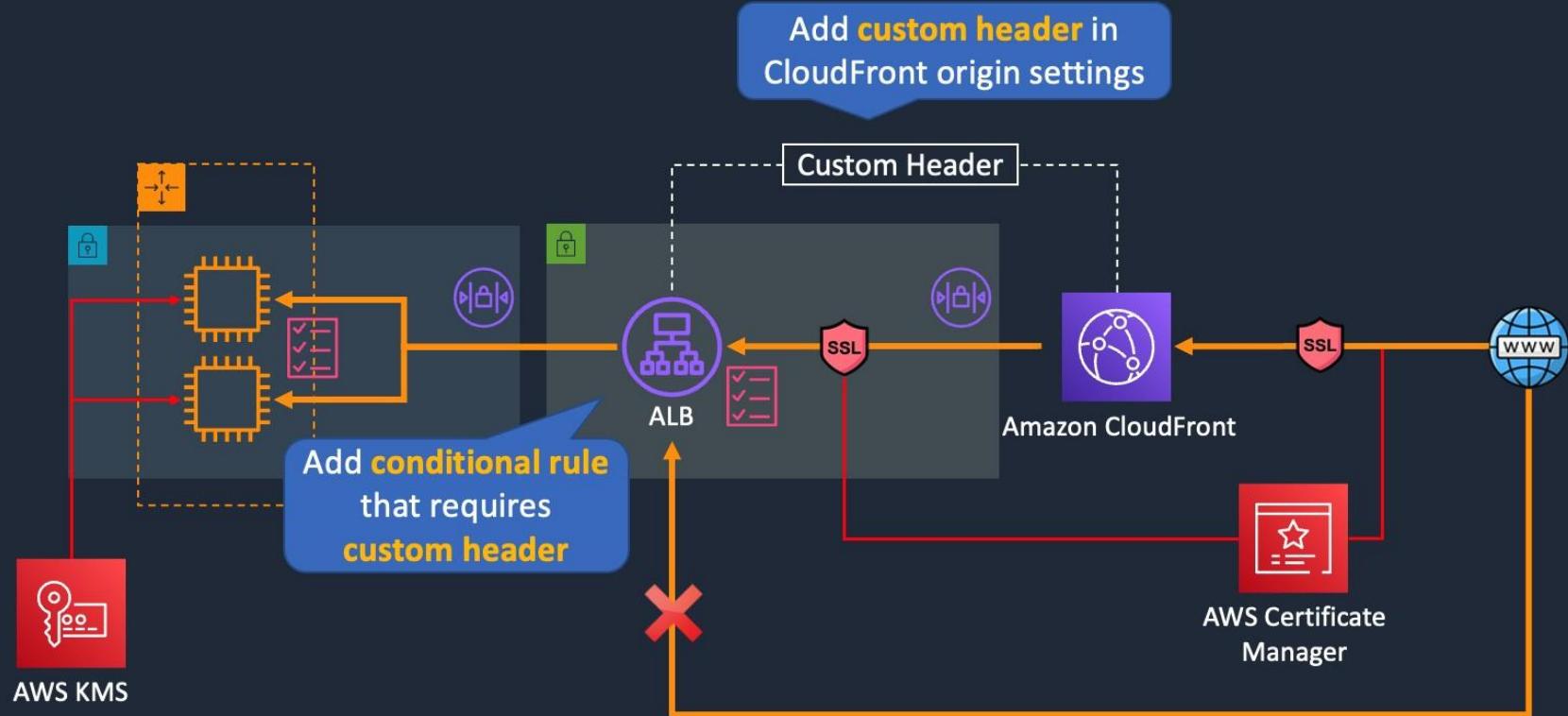
Explanation

Amazon CloudFront comes with AWS Shield standard by default which will provide some protection against DDoS attacks. For malicious web attacks an AWS WAF ACL should be associated with the distribution so that it can protect against the attacks using an appropriate rule group.

In this configuration it is important to ensure that the attacks cannot circumvent CloudFront and connect directly to the public ALB. For this, we can create a custom header and secret value in CloudFront. This will be forwarded in requests that originate from CloudFront. The ALB can conditionally forward only if this HTTP header information is present in the request.



Build a Secure Multi-Tier Architecture



© Digital Cloud Training | <https://digitalcloud.training>

CORRECT: "Create an Amazon CloudFront distribution with the ALB as the origin and configure a custom header and secret value. Configure the ALB to conditionally forward traffic only if the header and value match" is a correct answer (as explained above.)

CORRECT: "Deploy an AWS WAF web ACL that includes a rule group that blocks the attack traffic. Associate the web ACL with the Amazon CloudFront distribution" is also a correct answer (as explained above.)

INCORRECT: "Deploy the application in two AWS Regions. Configure Amazon Route 53 to route to both Regions with equal weight" is incorrect. This is not the most cost-effective option as the entire application stack is deployed in two Regions.

INCORRECT: "Configure AWS Auto Scaling for Amazon ECS tasks. Create an Amazon DynamoDB Accelerator (DAX) cluster in front of the DynamoDB table" is incorrect. DAX may assist with performance when caching requests but doesn't help with preventing web attacks from reaching the ALB or application servers.

INCORRECT: "Configure AWS Auto Scaling for Amazon ECS tasks. Configure an Amazon ElastiCache cluster in front of the DynamoDB table" is incorrect. As with the previous answer this solution does not assist with mitigating the impact of the attacks.

References:

<https://aws.amazon.com/premiumsupport/knowledge-center/elb-route-traffic-custom-http-header/>

Save time with our AWS cheat sheets:

<https://digitalcloud.training/aws-waf-shield/>

Question 4: **Correct**

A company has created several development accounts in an AWS Organizations organization. The company has defined a fixed budget for each development account and needs to ensure that developers cannot launch expensive services or exceed the fixed monthly budget.

Which combination of steps should a solutions architect take? (Select THREE.)

Create an SCP that defines a fixed monthly resource usage limit. Apply the SCP to an OU containing the development accounts.

Use the AWS Budgets service to define a fixed monthly budget for each development account.

(Correct)

Create an AWS Budgets alert action to send an Amazon SNS notification when the budgeted amount is reached. Invoke an AWS Lambda function to terminate all services.

(Correct)



Create an AWS Budgets alert action to terminate services when the budgeted amount is reached. Configure the action to terminate all services.



Create an SCP that denies access to expensive services. Apply the SCP to an OU containing the development accounts.

(Correct)



Create an IAM policy that denies access to expensive services. Apply the IAM policy to the development accounts.

Explanation

The solutions architect should use an SCP to deny access to expensive services as this will ensure that they cannot be launched in the first place. Then, an AWS Budget should be defined and configured with the fixed monthly cost allowance. The budget can trigger an SNS notification which in turn can invoke an AWS Lambda function to terminate the resources.

CORRECT: "Use the AWS Budgets service to define a fixed monthly budget for each development account" is a correct answer (as explained above.)

CORRECT: "Create an SCP that denies access to expensive services. Apply the SCP to an OU containing the development accounts" is also a correct answer (as explained above.)

CORRECT: "Create an AWS Budgets alert action to send an Amazon SNS notification when the budgeted amount is reached. Invoke an AWS Lambda function to terminate all services" is also a correct answer (as explained above.)

INCORRECT: "Create an SCP that defines a fixed monthly resource usage limit. Apply the SCP to an OU containing the development accounts" is incorrect. You cannot define resource limits in this manner using an SCP.

INCORRECT: "Create an IAM policy that denies access to expensive services. Apply the IAM policy to the development accounts" is incorrect. You cannot attach an IAM policy to an account, they must be attached to users, groups, or roles.

INCORRECT: "Create an AWS Budgets alert action to terminate services when the budgeted amount is reached. Configure the action to terminate all services" is incorrect. You cannot configure the alert action to terminate all services. You can stop EC2 and RDS instances or you can attach an IAM policy or SCP.

References:

<https://docs.aws.amazon.com/cost-management/latest/userguide/budgets-controls.html>

Save time with our AWS cheat sheets:

<https://digitalcloud.training/aws-organizations/>

Question 5: **Correct**

A company is migrating an order processing application to the AWS Cloud. The usage patterns vary significantly but the application must be available at all times. Orders must be processed immediately and in the order that they are received. Which actions should a Solutions Architect take to meet these requirements?



Use Amazon SNS with FIFO to send orders in the correct order. Use a single large Reserved Instance for processing.



Use Amazon SQS with FIFO to queue messages in the correct order. Use Spot Instances in multiple Availability Zones for processing.



Use Amazon SNS with FIFO to send orders in the correct order. Use Spot Instances in multiple Availability Zones for processing.



Use Amazon SQS with FIFO to queue messages in the correct order. Use Reserved Instances in multiple Availability Zones for processing.

(Correct)

Explanation

Amazon Simple Queue Service (SQS) with a first-in-first-out (FIFO) queue will ensure that messages are delivered to the processing layer in the correct order. An application component running on Amazon EC2 will then be configured to poll the queue and process the messages.

Reserved instances should be used for the processing layer as this is the best way to ensure that the application is available at all times at the best cost.

CORRECT: "Use Amazon SQS with FIFO to queue messages in the correct order. Use Reserved Instances in multiple Availability Zones for processing" is the correct answer.

INCORRECT: "Use Amazon SQS with FIFO to queue messages in the correct order. Use Spot Instances in multiple Availability Zones for processing" is incorrect. With Spot instances the application could become unavailable if the Spot price exceeds the default maximum price configured.

INCORRECT: "Use Amazon SNS with FIFO to send orders in the correct order. Use Spot Instances in multiple Availability Zones for processing" is incorrect. SNS is used for sending notifications, it is not the best service to use for this use case.

INCORRECT: "Use Amazon SNS with FIFO to send orders in the correct order. Use a single large Reserved Instance for processing" is incorrect. SNS is used for sending notifications, it is not the best service to use for this use case.

References:

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

Save time with our AWS cheat sheets:

<https://digitalcloud.training/category/aws-cheat-sheets/aws-solutions-architect-professional/aws-analytics-sap/>

Question 6: **Correct**

A Solutions Architect is developing a mechanism to gain security approval for Amazon EC2 images (AMIs) so that they can be used by developers. The AMIs must go through an automated assessment process (CVE assessment) and be marked as approved before developers can use them. The approved images must be scanned every 30 days to ensure compliance. Which combination of steps should the Solutions Architect take to meet these requirements while following best practices? (Select TWO.)

-
-

Use AWS GuardDuty to run the CVE assessment package on the EC2 instances launched from the approved AMIs.

-
-

Use AWS Lambda to write automatic approval rules. Store the approved AMI list in AWS Systems Manager Parameter Store. Use Amazon EventBridge to trigger an AWS Systems Manager Automation document on all EC2 instances every 30 days.

(Correct)

-
-

Use AWS Lambda to write automatic approval rules. Store the approved AMI list in AWS Systems Manager Parameter Store. Use a managed AWS Config rule for continuous scanning on all EC2 instances and use AWS Systems Manager Automation documents for remediation.

-
-

Use Amazon Inspector to run the CVE assessment package on the EC2 instances launched from the approved AMIs.

(Correct)



Use the AWS Systems Manager EC2 agent to run the CVE assessment on the EC2 instances launched from the approved AMIs.

Explanation

AWS Lambda can be used to run the approval process for the AMIs and then automatically store the results in AWS Systems Manager Parameter Store.

For the CVE assessment, Amazon Inspector can be used to perform security assessments of Amazon EC2 instances by using AWS managed rules packages such as the Common Vulnerabilities and Exposures (CVEs) package.

Amazon EventBridge (CloudWatch Events) can then be used to create scheduled triggers that run AWS Systems Manager Automation documents on a recurring schedule (30 days). AWS Systems Manager will update the running instances to ensure they are up to date with any security updates that need to be applied.

CORRECT: "Use AWS Lambda to write automatic approval rules. Store the approved AMI list in AWS Systems Manager Parameter Store. Use Amazon EventBridge to trigger an AWS Systems Manager Automation document on all EC2 instances every 30 days" is a correct answer.

CORRECT: "Use Amazon Inspector to run the CVE assessment package on the EC2 instances launched from the approved AMIs" is also a correct answer.

INCORRECT: "Use the AWS Systems Manager EC2 agent to run the CVE assessment on the EC2 instances launched from the approved AMIs" is incorrect. Systems Manager does not have a CVE assessment, use Amazon Inspector which is designed for this purpose and has a package preconfigured.

INCORRECT: "Use AWS Lambda to write automatic approval rules. Store the approved AMI list in AWS Systems Manager Parameter Store. Use a managed AWS Config rule for continuous scanning on all EC2 instances and use AWS Systems Manager Automation documents for remediation" is incorrect. Amazon Inspector is a better fit for a CVE assessment.

INCORRECT: "Use AWS GuardDuty to run the CVE assessment package on the EC2 instances launched from the approved AMIs" is incorrect. GuardDuty is an intelligent threat detection service. It is not suitable for a CVE assessment.

References:

<https://aws.amazon.com/blogs/security/how-to-set-up-continuous-golden-ami-vulnerability-assessments-with-amazon-inspector/>

Save time with our AWS cheat sheets:

<https://digitalcloud.training/category/aws-cheat-sheets/aws-solutions-architect-professional/aws-compute-sap/>

<https://digitalcloud.training/category/aws-cheat-sheets/aws-solutions-architect-professional/security-identity-compliance-sap/>

Question 7: **Correct**

A company runs a data processing application on-premises and plans to move it to the AWS Cloud. Files are uploaded by users to a web application which then stores the files on an NFS-based storage system and places a message on a queue. The files are then processed from the queue and the results are returned to the user (and stored in long-term storage). This process can take up to 30 minutes. The processing times vary significantly and can be much higher during business hours.

What is the MOST cost-effective migration recommendation?



Create a queue using Amazon MQ. Run the web application on Amazon EC2 and configure it to publish to the new queue. Use an AWS Lambda function to poll the queue, pull requests, and process the files. Store the processed files in Amazon EFS.



Create a queue using Amazon SQS. Run the web application on Amazon EC2 and configure it to publish to the new queue. Use an AWS Lambda function to poll the queue, pull requests, and process the files. Store the processed files in an Amazon S3 bucket.



Create a queue using Amazon MQ. Run the web application on Amazon EC2 and configure it to publish to the new queue. Launch an Amazon EC2 instance from a preconfigured AMI to poll the queue, pull requests, and process the files. Store the processed files in Amazon EFS. Terminate the EC2 instance after the task is complete.



Create a queue using Amazon SQS. Run the web application on Amazon EC2 and configure it to publish to the new queue. Use Amazon EC2 instances in an EC2 Auto Scaling group to pull requests from the queue and process the files. Scale the EC2 instances based on the SQS queue length. Store the processed files in an Amazon S3 bucket.

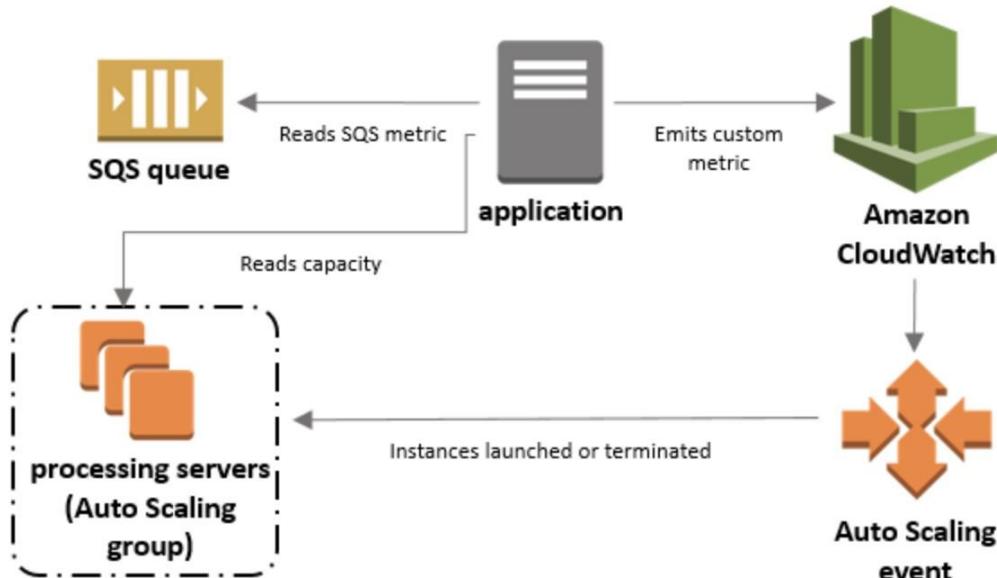
(Correct)

Explanation

The best solution is to use Amazon SQS for the message queue, Amazon EC2 Auto Scaling for the processing layer, and Amazon S3 for the storage layer. This solution meets all requirements and is the lowest cost option available.

The ASG should also be configured to scale based on the ApproximateNumberOfMessages queue attribute. This is used in combination with an acceptable backlog per instance metric which defines the number of messages in the queue to use as scaling criteria.

The following diagram illustrates the architecture of this configuration.



CORRECT: "Create a queue using Amazon SQS. Run the web application on Amazon EC2 and configure it to publish to the new queue. Use Amazon EC2 instances in an EC2 Auto Scaling group to pull requests from the queue and process the files. Scale the EC2 instances based on the SQS queue length. Store the processed files in an Amazon S3 bucket" is the correct answer.

INCORRECT: "Create a queue using Amazon SQS. Run the web application on Amazon EC2 and configure it to publish to the new queue. Use an AWS Lambda function to poll the queue, pull requests, and process the files. Store the processed files in an Amazon S3 bucket" is incorrect. Lambda cannot process files for 30 minutes; it has a maximum execution time of 15 minutes.

INCORRECT: "Create a queue using Amazon MQ. Run the web application on Amazon EC2 and configure it to publish to the new queue. Launch an Amazon EC2 instance from a preconfigured AMI to poll the queue, pull requests, and process the files. Store the processed files in Amazon EFS. Terminate the EC2 instance after the task is complete" is incorrect. If the instance is terminated each time then how can it poll the queue? This is not a workable solution.

INCORRECT: "Create a queue using Amazon MQ. Run the web application on Amazon EC2 and configure it to publish to the new queue. Use an AWS Lambda function to poll the queue, pull requests, and process the files. Store the processed files in Amazon EFS" is incorrect. Lambda cannot be used as mentioned previously (max execution limit).

References:

<https://docs.aws.amazon.com/autoscaling/ec2/userguide/as-using-sqs-queue.html>

Save time with our AWS cheat sheets:

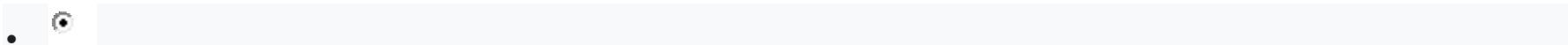
<https://digitalcloud.training/category/aws-cheat-sheets/aws-solutions-architect-professional/aws-compute-sap/>

<https://digitalcloud.training/category/aws-cheat-sheets/aws-solutions-architect-professional/aws-application-integration-sap/>

Question 8: **Incorrect**

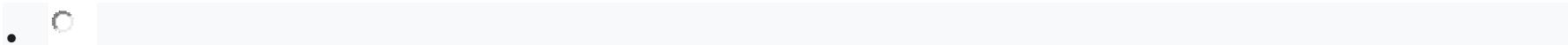
A company has established a 10 Gbps AWS Direct Connect (DX) connection to a single VPC in an AWS Region. A single private VIF has been created for the existing DX connection. The company requires redundancy for the existing DX connection and needs to connect to an additional VPC in a second Region.

Which solution meets these requirements?

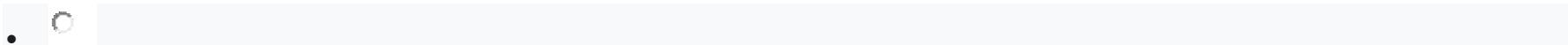


Create a new DX connection to the second Region. Provision a transit gateway and establish new private VIFs to a virtual private gateway in the VPCs in each Region.

(Incorrect)

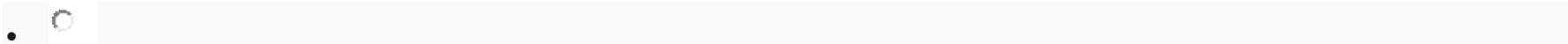


Create a new DX connection to the second Region. Create a new private VIF across the new DX connection to a virtual private gateway in the VPC in the second Region.



Create a new DX connection to the same Region. Provision a Direct Connect gateway and establish new private VIFs to a virtual private gateway in the VPCs in each Region.

(Correct)



Create a new DX connection to the same Region. Provision a Direct Connect gateway and establish new private VIFs to a transit gateway in the VPCs in each Region.

Explanation

Use AWS Direct Connect gateway to connect your VPCs. You associate an AWS Direct Connect gateway with either of the following gateways:

- A transit gateway when you have multiple VPCs in the same Region
- A virtual private gateway

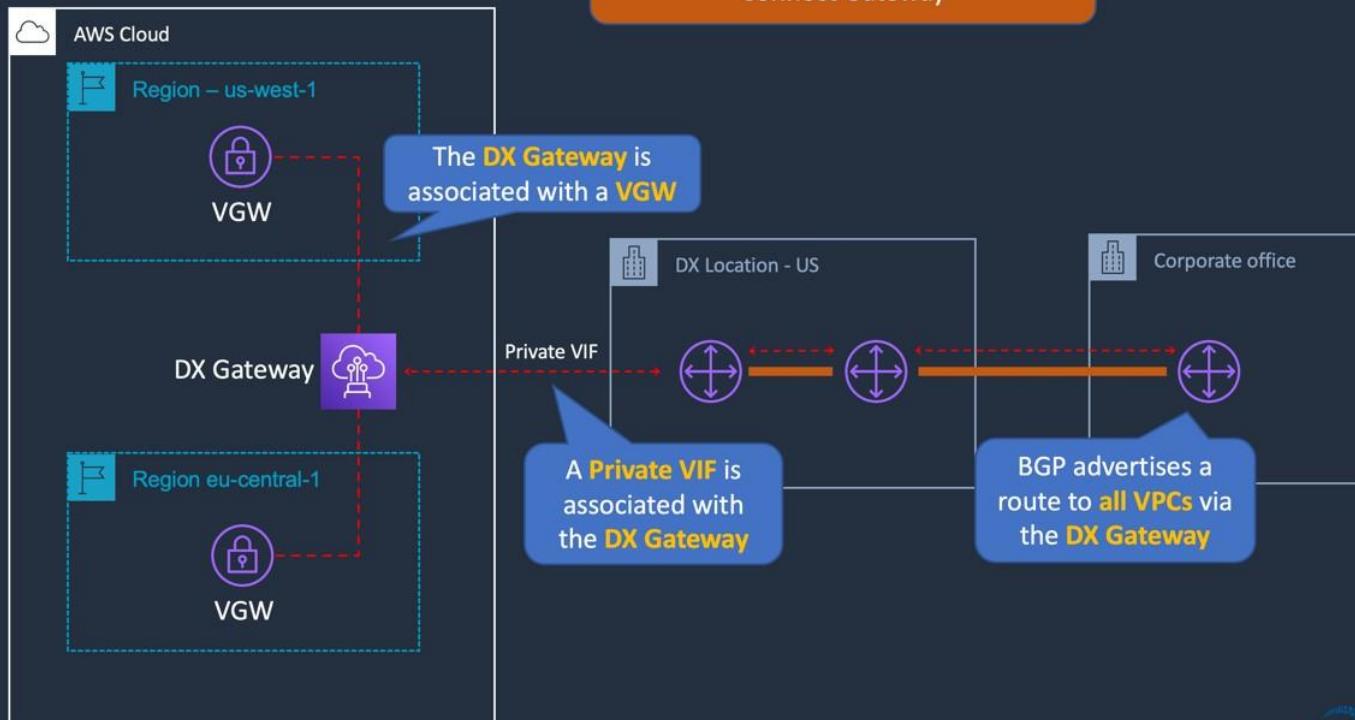
A Direct Connect gateway is a globally available resource. You can create the Direct Connect gateway in any Region and access it from all other Regions.

In this case the architecture should include a second DX connection from the on-premises network to the same Region as the existing DX connection to add redundancy. A DX GW can be provisioned and private VIFs can be created to virtual private gateways (VGWs) in the two AWS Regions the company wants to access.



Direct Connect - Multiple Regions

Example architecture **with** AWS Direct Connect Gateway



© Digital Cloud Training | <https://digitalcloud.training>

CORRECT: "Create a new DX connection to the same Region. Provision a Direct Connect gateway and establish new private VIFs to a virtual private gateway in the VPCs in each Region" is the correct answer (as explained above.)

INCORRECT: "Create a new DX connection to the same Region. Provision a Direct Connect gateway and establish new private VIFs to a transit gateway in the VPCs in each Region" is incorrect.

The transit gateway is not required in this scenario as the company only has a single VPC they need to connect to in each Region.

INCORRECT: "Create a new DX connection to the second Region. Create a new private VIF across the new DX connection to a virtual private gateway in the VPC in the second Region" is incorrect.

This does not provide the DX redundancy the company requires for the existing DX connection as the connection goes to a different Region and VPC.

INCORRECT: "Create a new DX connection to the second Region. Provision a transit gateway and establish new private VIFs to a virtual private gateway in the VPCs in each Region" is incorrect.

The company requires redundancy for the existing DX connection in the same Region and will then need a DX gateway to connect across Regions.

References:

<https://docs.aws.amazon.com/directconnect/latest/UserGuide/direct-connect-gateways-intro.html>

Save time with our AWS cheat sheets:

<https://digitalcloud.training/aws-direct-connect/>

Question 9: **Correct**

A company requires that only the master account in AWS Organizations is able to purchase Amazon EC2 Reserved Instances. Current and future member accounts should be blocked from purchasing Reserved Instances.

Which solution will meet these requirements?



Move all current member accounts to a new OU. Create an SCP with the Deny effect on the ec2:PurchaseReservedInstancesOffering action. Attach the SCP to the new OU.



Create an SCP with the Deny effect on the ec2:PurchaseReservedInstancesOffering action. Attach the SCP to the root of the organization.

(Correct)



Create an OU for the master account and each member account. Move the accounts into their respective OUs. Apply an SCP to the master accounts' OU with the Allow effect for the ec2:PurchaseReservedInstancesOffering.



Create an Amazon CloudWatch Events rule that triggers a Lambda function to terminate any Reserved Instances launched by member accounts.

Explanation

The only solution that works for both existing and future member accounts is to apply a Deny policy to the root of the organization. When you attach a policy to the organization root, all OUs and accounts in the organization inherit that policy which ensures that any new accounts that are added will inherit the policy automatically.

SCPs affect only **member** accounts in the organization. They have no effect on users or roles in the management account (also known as the master account). Therefore, the users in the management account are able to purchase reserved instances.

Note the following behavior in relation to policy inheritance:

You can attach policies to organization entities (organization root, organizational unit (OU), or account) in your organization:

When you attach a policy to the organization root, all OUs and accounts in the organization inherit that policy.

When you attach a policy to a specific OU, accounts that are directly under that OU or any child OU inherit the policy.

When you attach a policy to a specific account, it affects only that account.

CORRECT: "Create an SCP with the Deny effect on the ec2:PurchaseReservedInstancesOffering action. Attach the SCP to the root of the organization" is the correct answer.

INCORRECT: "Move all current member accounts to a new OU. Create an SCP with the Deny effect on the ec2:PurchaseReservedInstancesOffering action. Attach the SCP to the new OU" is incorrect. This will work for existing accounts but if new accounts are added and are not added to the same OU they will not inherit the policy.

INCORRECT: "Create an Amazon CloudWatch Events rule that triggers a Lambda function to terminate any Reserved Instances launched by member accounts" is incorrect. CloudWatch Events is not able to trigger based on EC2 Reserved Instance purchase actions.

INCORRECT: "Create an OU for the master account and each member account. Move the accounts into their respective OUs. Apply an SCP to the master accounts' OU with the Allow effect for the ec2:PurchaseReservedInstancesOffering" is incorrect. This is a complex setup and does not deny the relevant API actions from the member accounts.

References:

https://docs.aws.amazon.com/organizations/latest/userguide/orgs_manage_policies_inheritance.html

https://docs.aws.amazon.com/AWSEC2/latest/APIReference/API_PurchaseReservedInstancesOffering.html

Save time with our AWS cheat sheets:

<https://digitalcloud.training/category/aws-cheat-sheets/aws-solutions-architect-professional/aws-database-sap/>

Question 10: **Correct**

A company is migrating its on-premises systems to AWS. The computers consist of a combination of Windows and Linux virtual machines on VMware and physical servers.

The company wants to be able to identify dependencies between on-premises systems and group systems together into applications to build migration plans. The company also needs to understand the performance requirements for systems so they can be right-sized.

How can these requirements be met?



Install the AWS Application Discovery Service Discovery Connector in VMware vCenter. Install the AWS Application Discovery Service Discovery Agent on the physical on-premises servers. Allow the Discovery Agent to collect data for a period of time.

(Correct)



Install the AWS Application Discovery Service Discovery Connector in VMware vCenter. Allow the Discovery Connector to collect data for one week.



Extract system information from an on-premises configuration management database (CMDB). Import the data directly into the Application Discovery Service.



Install the AWS Application Discovery Service Discovery Agent on each of the on-premises systems. Allow the Discovery Agent to collect data for a period of time.

Explanation

The AWS Discovery Agent is AWS software that you install on on-premises servers and VMs targeted for discovery and migration. Agents capture system configuration, system performance, running processes and details of the network connections between systems. Agents support most Linux and Windows operating systems, and you can deploy them on physical on-premises servers, Amazon EC2 instances, and virtual machines.

Though you can use the agent on virtual machines, as described above, it is more efficient to use the Agentless Discovery connector for VMware virtual machines. This connector can be installed in VMware vCenter.

The AWS Discovery Connector is a VMware appliance that can collect information only about VMware virtual machines (VMs). You install the Discovery Connector as a VM in your VMware vCenter Server environment using an Open Virtualization Archive (OVA) file. Because the Discovery Connector relies on VMware metadata to gather server information regardless of operating system, it minimizes the time required for initial on-premises infrastructure assessment.

CORRECT: "Install the AWS Application Discovery Service Discovery Connector in VMware vCenter. Install the AWS Application Discovery Service Discovery Agent on the physical on-premises servers. Allow the Discovery Agent to collect data for a period of time" is the correct answer.

INCORRECT: "Extract system information from an on-premises configuration management database (CMDB). Import the data directly into the Application Discovery Service" is incorrect. It is possible to upload data to the Migration Hub but you must use a specially formatted CSV file. It is unlikely the CMDB export would be directly importable into Migration Hub so some work is likely required to format the data. Also, the CMDB will hold configuration data but not performance data so this solution does not satisfy all requirements.

INCORRECT: "Install the AWS Application Discovery Service Discovery Agent on each of the on-premises systems. Allow the Discovery Agent to collect data for a period of time" is incorrect. This is not the most efficient method of collecting data from the VMware virtual machines as the connector for vCenter is a better choice.

INCORRECT: "Install the AWS Application Discovery Service Discovery Connector in VMware vCenter. Allow the Discovery Connector to collect data for one week" is incorrect. This will only retrieve information about the virtual machines, not the physical servers.

References:

<https://docs.aws.amazon.com/application-discovery/latest/userguide/discovery-agent.html>

<https://docs.aws.amazon.com/application-discovery/latest/userguide/discovery-connector.html>

Save time with our AWS cheat sheets:

<https://digitalcloud.training/category/aws-cheat-sheets/aws-solutions-architect-professional/aws-migration-sap/>

Question 11: Correct

A company is designing an application that will require cross-Region disaster recovery with an RTO of less than 5 minutes and an RPO of less than 1 minute. The application tier DR solution has already been designed and a Solutions Architect must design the data recovery solution for the MySQL database tier.

How should the database tier be configured to meet the data recovery requirements?

- 

Use an Amazon Aurora global database with the primary in the active Region and the secondary in the failover Region.

(Correct)

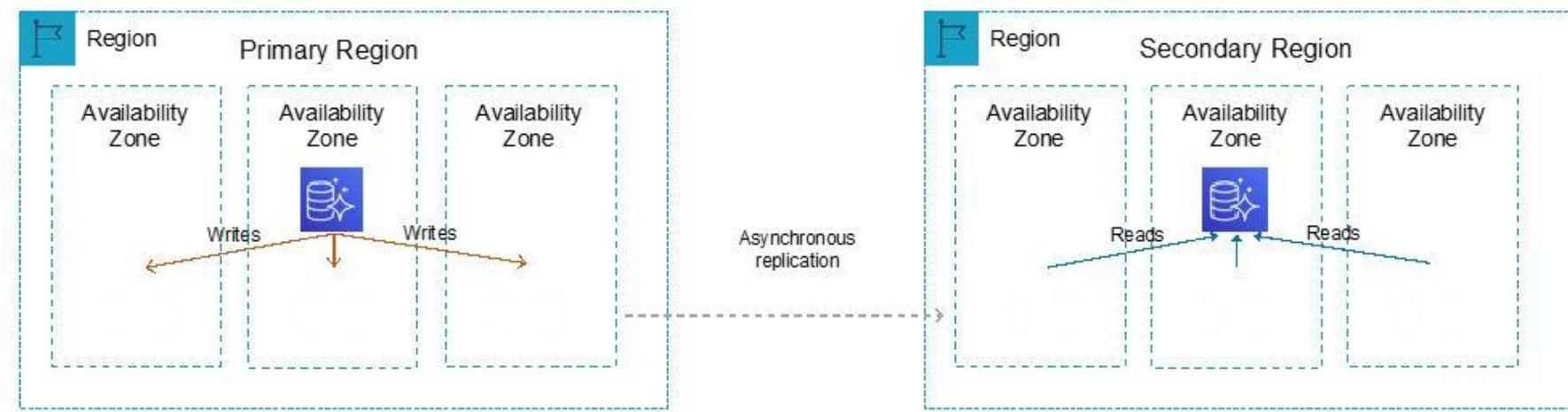
- Use an Amazon RDS for MySQL instance with a cross-Region read replica in the failover Region.

- Use an Amazon RDS for MySQL instance with a Multi-AZ deployment.

- Create an Amazon RDS instance in the active Region and use a MySQL standby database on an Amazon EC2 instance in the failover Region.

Explanation

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



This solution will meet the Recovery Time Objective (RTO) of less than 5 minutes and Recovery Point Objective (RPO) of less than 1 minute.

CORRECT: "Use an Amazon Aurora global database with the primary in the active Region and the secondary in the failover Region" is the correct answer.

INCORRECT: "Use an Amazon RDS for MySQL instance with a cross-Region read replica in the failover Region" is incorrect. A read replica cannot be used as a writable database. It is possible to promote a read replica but this may not be fast enough to meet the RTO requirement.

INCORRECT: "Use an Amazon RDS for MySQL instance with a Multi-AZ deployment" is incorrect. You cannot have a Multi-AZ deployment that spans across AWS Regions.

INCORRECT: "Create an Amazon RDS instance in the active Region and use a MySQL standby database on an Amazon EC2 instance in the failover Region" is incorrect. You cannot configure a MySQL DB on EC2 to be a standby for an Amazon RDS DB instance.

References:

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

Save time with our AWS cheat sheets:

<https://digitalcloud.training/category/aws-cheat-sheets/aws-solutions-architect-professional/aws-database-sap/>

Question 12: **Correct**

A company is testing an application that collects data from sensors fitted to vehicles. The application collects usage statistics data every 4 minutes. The data is sent to Amazon API Gateway, it is then processed by an AWS Lambda function and the results are stored in an Amazon DynamoDB table.

As the sensors have been fitted to more vehicles, and as more metrics have been configured for collection, the Lambda function execution time has increased from a few seconds to over 2 minutes. There are also many TooManyRequestsException errors being generated by Lambda.

Which combination of changes will resolve these issues? (Select TWO.)



Increase the memory available to the Lambda functions.

(Correct)



Increase the CPU units assigned to the Lambda functions.



Use Amazon EC2 instead of Lambda to process the data.



Stream the data into an Amazon Kinesis data stream from API Gateway and process the data in batches.

(Correct)



Collect data in an Amazon SQS FIFO queue, which triggers a Lambda function to process each message.

Explanation

To optimize the function execution time, the memory assigned to the Lambda functions can be increased. This will proportionally increase the amount of CPU assigned to each function execution.

The `TooManyRequestsException` error from Lambda can be resolved by configuring API Gateway to place incoming data into a Kinesis data stream. AWS Lambda can then process the data in batches which is more efficient.

CORRECT: "Increase the memory available to the Lambda functions" is the correct answer.

CORRECT: "Stream the data into an Amazon Kinesis data stream from API Gateway and process the data in batches" is also a correct answer.

INCORRECT: "Use Amazon EC2 instead of Lambda to process the data" is incorrect. There is no reason to use EC2 in place of Lambda. This would not be more cost-effective and does not offer any advantages to this solution.

INCORRECT: "Increase the CPU units assigned to the Lambda functions" is incorrect. You do not directly increase the CPU assigned to Lambda functions, you increase the memory assigned and that automatically adjust the amount of CPU assigned.

INCORRECT: "Collect data in an Amazon SQS FIFO queue, which triggers a Lambda function to process each message" is incorrect. This is less efficient in terms of the number of function executions as it does not include batching.

References:

<https://docs.aws.amazon.com/lambda/latest/dg/with-kinesis.html>

<https://docs.aws.amazon.com/lambda/latest/dg/configuration-memory.html>

Save time with our AWS cheat sheets:

<https://digitalcloud.training/category/aws-cheat-sheets/aws-solutions-architect-professional/aws-compute-sap/>

<https://digitalcloud.training/category/aws-cheat-sheets/aws-solutions-architect-professional/aws-analytics-sap/>

Question 13: **Correct**

A Solutions Architect is designing a web application that will serve static content in an Amazon S3 bucket and dynamic content hosted on Amazon EC2 instances behind an Application Load Balancer (ALB). The application will use Amazon CloudFront and the solution should require that the content is available through CloudFront only.

Which combination of steps should the Solutions Architect take to restrict direct content access to CloudFront? (Select THREE.)

-

Configure an S3 bucket policy to allow access from the CloudFront IP addresses only.

-

Create a web ACL in AWS WAF with a rule to validate the presence of a custom header and associate the web ACL with the ALB.

(Correct)

-

Configure the ALB to add a custom header to HTTP requests that are sent to the EC2 instances.

-

Create a web ACL in AWS WAF with a rule to validate the presence of a custom header and associate the web ACL with the CloudFront distribution.

-

Configure CloudFront to add a custom header to requests that it sends to the origin.

(Correct)

-

Create a CloudFront Origin Access Identity (OAI) and add it to the CloudFront distribution. Update the S3 bucket policy to allow access to the OAI only.

(Correct)

Explanation

If you use a custom origin, you can optionally set up custom headers to restrict access. For CloudFront to get your files from a custom origin, the files must be accessible by CloudFront using a standard HTTP (or HTTPS) request.

By using custom headers, you can further restrict access to your content so that users can access it only through CloudFront, not directly. In this case an AWS WAF web ACL can be used to filter the requests and validate the presence of the custom header.

For Amazon S3 an Origin Access Identity can be used (OAI). The OAI is a special CloudFront user that is associated with the distribution. After creating an OAI, the S3 bucket permissions can then be modified so that CloudFront can use the OAI to access the files in your bucket and serve them to your users (and also restrict any other access).

CORRECT: "Create a web ACL in AWS WAF with a rule to validate the presence of a custom header and associate the web ACL with the ALB" is a correct answer.

CORRECT: "Configure CloudFront to add a custom header to requests that it sends to the origin" is also a correct answer.

CORRECT: "Create a CloudFront Origin Access Identity (OAI) and add it to the CloudFront distribution. Update the S3 bucket policy to allow access to the OAI only" is also a correct answer.

INCORRECT: "Create a web ACL in AWS WAF with a rule to validate the presence of a custom header and associate the web ACL with the CloudFront distribution" is incorrect. The web ACL should be associated with the ALB, not the CloudFront distribution.

INCORRECT: "Configure the ALB to add a custom header to HTTP requests that are sent to the EC2 instances" is incorrect. ALBs cannot add custom headers to requests, this should be done by CloudFront and then validated using a web ACL that is applied to the ALB.

INCORRECT: "Configure an S3 bucket policy to allow access from the CloudFront IP addresses only" is incorrect. This is not the best solution (though it can be done). It is administratively easier to use an OAI rather than the CloudFront IP addresses as they change over time.

References:

<https://docs.aws.amazon.com/AmazonCloudFront/latest/DeveloperGuide/private-content-overview.html#forward-custom-headers-restrict-access>

<https://docs.aws.amazon.com/AmazonCloudFront/latest/DeveloperGuide/private-content-restricting-access-to-s3.html>

Save time with our AWS cheat sheets:

<https://digitalcloud.training/category/aws-cheat-sheets/aws-solutions-architect-professional/aws-networking-content-delivery-sap/>

<https://digitalcloud.training/category/aws-cheat-sheets/aws-solutions-architect-professional/security-identity-compliance-sap/>

Question 14: **Correct**

A new application that provides fitness and training advice has become extremely popular with thousands of new users from around the world. The web application is hosted on a fleet of Amazon EC2 instances in an Auto Scaling group behind an Application Load Balancer (ALB). The content consists of static media files and different resources must be loaded depending on the client operating system.

Users have reported increasing latency for loading web pages and Amazon CloudWatch is showing high utilization of the EC2 instances.

Which set actions should a solutions architect take to improve response times?



Move content to Amazon S3. Create an Amazon CloudFront distribution to serve content out of the S3 bucket. Use the User-Agent HTTP header to load different content.



Create a separate ALB for each client operating system. Create one Auto Scaling group behind each ALB. Use Amazon Route 53 to route to different ALBs depending on the User-Agent HTTP header.



Move content to Amazon S3. Create an Amazon CloudFront distribution to serve content out of the S3 bucket. Use Lambda@Edge to load different resources based on the User-Agent HTTP header.

(Correct)



Create separate Auto Scaling groups based on client operating systems. Switch to a Network Load Balancer (NLB). Use the User-Agent HTTP header in the NLB to route to a different set of EC2 instances.

Explanation

The load on the EC2 instances can be reduced by serving the static contents from Amazon CloudFront. This service will cache the content at Edge locations for faster delivery to clients.

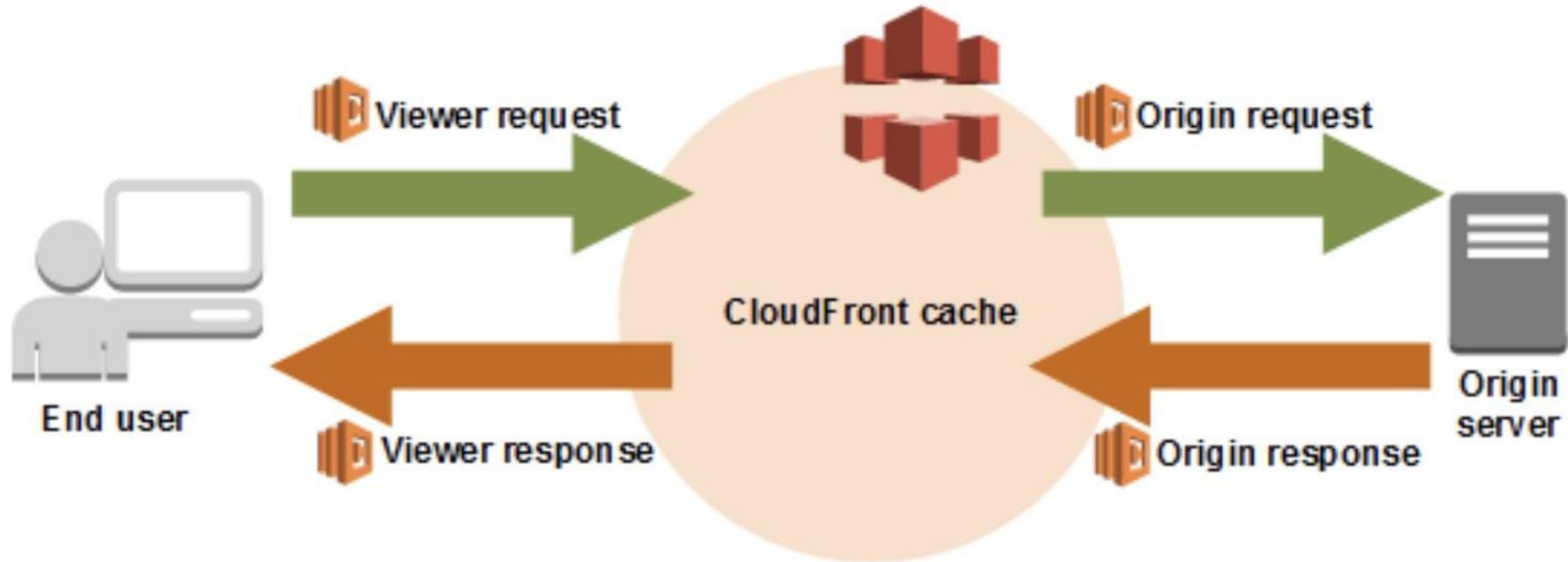
To load different content based on the client operating system Lambda@Edge can be used. Lambda@Edge lets you run Node.js and Python Lambda functions to customize the content that CloudFront delivers.

Lambda@Edge can be configured to inspect the viewer request and look for the user-agent HTTP header. This header is a string that can be used to identify the application, operating system, vendor, and/or version of the requesting user agent. Based on the operating system of the client, the function can then return different media assets from the CloudFront cache.

You can use Lambda functions to change CloudFront requests and responses at the following points:

- After CloudFront receives a request from a viewer (viewer request)
- Before CloudFront forwards the request to the origin (origin request)

- After CloudFront receives the response from the origin (origin response)
- Before CloudFront forwards the response to the viewer (viewer response)



CORRECT: "Move content to Amazon S3. Create an Amazon CloudFront distribution to serve content out of the S3 bucket. Use Lambda@Edge to load different resources based on the User-Agent HTTP header" is the correct answer.

INCORRECT: "Create separate Auto Scaling groups based on client operating systems. Switch to a Network Load Balancer (NLB). Use the User-Agent HTTP header in the NLB to route to a different set of EC2 instances" is incorrect. The user-agent HTTP header cannot be used by an NLB to route to a different target group (set of EC2 instances).

INCORRECT: "Create a separate ALB for each client operating system. Create one Auto Scaling group behind each ALB. Use Amazon Route 53 to route to different ALBs depending on the User-Agent HTTP header" is incorrect. Route 53 cannot be used to route traffic based on the user-agent HTTP header.

INCORRECT: "Move content to Amazon S3. Create an Amazon CloudFront distribution to serve content out of the S3 bucket. Use the User-Agent HTTP header to load different content" is incorrect. There is no solution here for how to process the user-agent HTTP header and load different content. This is not a native capability of CloudFront which is why the correct solution uses a Lambda function to perform this processing.

References:

<https://docs.aws.amazon.com/lambda/latest/dg/lambda-edge.html>

Save time with our AWS cheat sheets:

<https://digitalcloud.training/category/aws-cheat-sheets/aws-solutions-architect-professional/aws-storage-sap/>

<https://digitalcloud.training/category/aws-cheat-sheets/aws-solutions-architect-professional/aws-networking-content-delivery-sap/>

Question 15: **Correct**

A company includes several business units that each use a separate AWS account and a parent company AWS account. The company requires a single AWS bill across all AWS accounts with costs broken out for each business unit. The company also requires that services and features be restricted in the business unit accounts and this must be governed centrally.

Which combination of steps should a Solutions Architect take to meet these requirements? (Select TWO.)

-

Enable consolidated billing in the parent account's billing console and link the business unit AWS accounts.

-

Use AWS Organizations to create a separate organization for each AWS account with all features enabled. Then, create trust relationships between the AWS organizations.

-

Use AWS Organizations to create a single organization in the parent account with all features enabled. Then, invite each business unit's AWS account to join the organization.

(Correct)

-

Create an SCP that allows only approved services and features, then apply the policy to the business unit AWS accounts.

(Correct)

-

Use permissions boundaries applied to each business unit's AWS account to define the maximum permissions available for services and features.

Explanation

To enable the required features you simply need to setup a single AWS organization in the parent account with all features enabled. The existing business unit AWS accounts can then be invited to join the organization.

This setup will automatically enable consolidated billing which will ensure a single AWS bill is received in the parent account which has costs broken out by each AWS account.

Service Control Policies (SCPs) can then be used to restrict the maximum available permissions to services and features that the parent company wishes to apply to the member accounts. Once applied, all users will be affected in the member accounts.

CORRECT: "Use AWS Organizations to create a single organization in the parent account with all features enabled. Then, invite each business unit's AWS account to join the organization" is a correct answer.

CORRECT: "Create an SCP that allows only approved services and features, then apply the policy to the business unit AWS accounts" is also a correct answer.

INCORRECT: "Use AWS Organizations to create a separate organization for each AWS account with all features enabled. Then, create trust relationships between the AWS organizations" is incorrect. A single organization should be created; trust relationships are not a concept associated with AWS Organizations.

INCORRECT: "Use permissions boundaries applied to each business unit's AWS account to define the maximum permissions available for services and features" is incorrect. Permissions boundaries are applied to IAM entities (users or roles), not to AWS accounts.

INCORRECT: "Enable consolidated billing in the parent account's billing console and link the business unit AWS accounts" is incorrect. Consolidated billing is a feature of AWS Organizations, you must create an organization, invite the relevant accounts, and then a single bill will be generated in the management (parent) account.

References:

https://docs.aws.amazon.com/organizations/latest/userguide/orgs_manage_org_support-all-features.html

Save time with our AWS cheat sheets:

<https://digitalcloud.training/category/aws-cheat-sheets/aws-solutions-architect-professional/security-identity-compliance-sap/>

Question 16: **Correct**

A company has a requirement to store documents that will be accessed by a serverless application. The documents will be accessed frequently for the first 3 months, and rarely after that. The documents must be retained for 7 years.

What is the MOST cost-effective solution to meet these requirements?

Store the documents in an encrypted EBS volume and create a cron job to delete the documents after 7 years.

Store the documents in a secured Amazon S3 bucket with a lifecycle policy to move the documents that are older than 3 months to Amazon S3 Glacier. Create an AWS Lambda function to delete the documents in S3 Glacier that are older than 7 years.

Store the documents in Amazon EFS. Create a cron job to move the documents that are older than 3 months to Amazon S3 Glacier. Create an AWS Lambda function to delete the documents in S3 Glacier that are older than 7 years.

Store the documents in a secured Amazon S3 bucket with a lifecycle policy to move the documents that are older than 3 months to Amazon S3 Glacier, then expire the documents from Amazon S3 Glacier that are more than 7 years old.

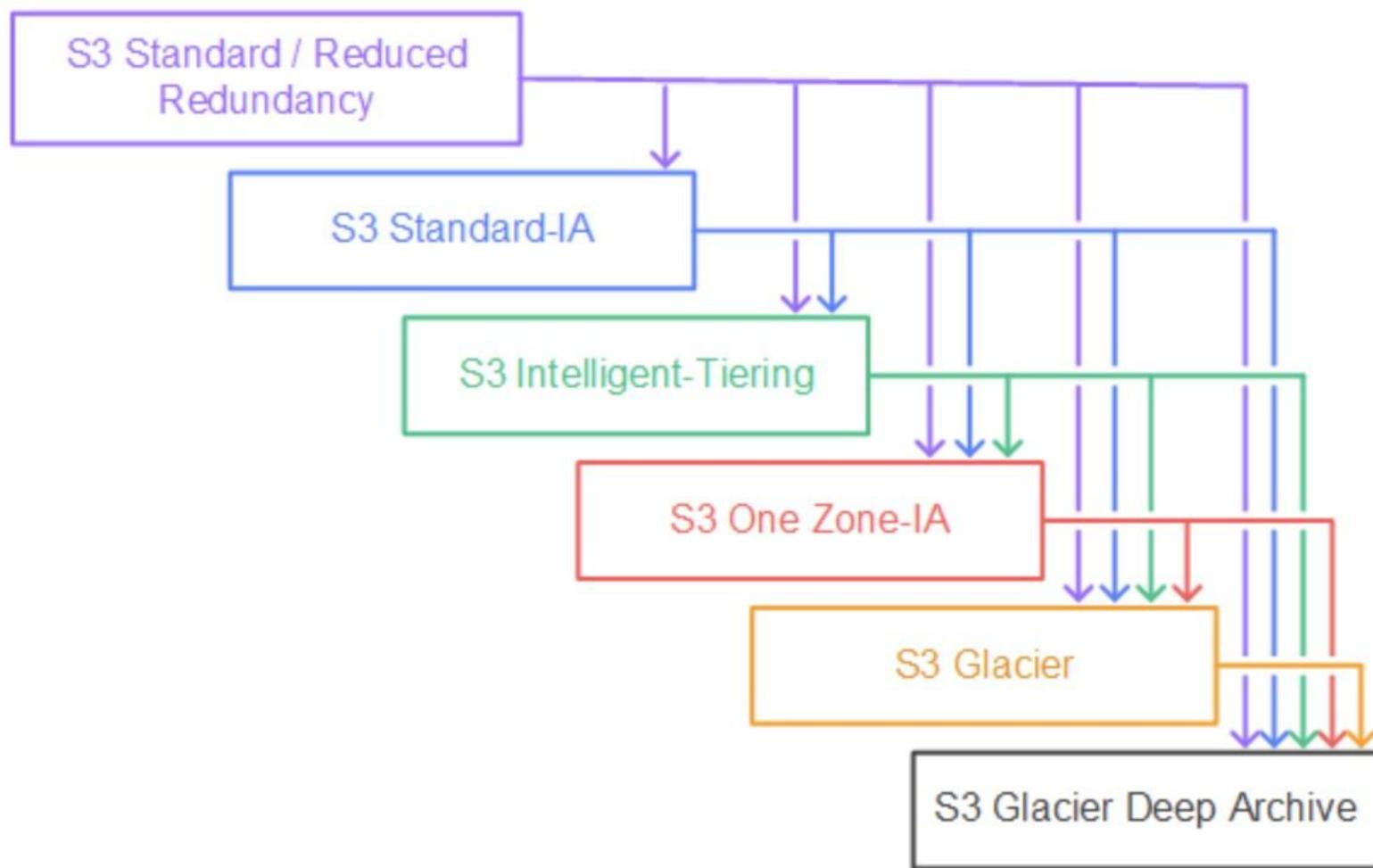
(Correct)

Explanation

An *S3 Lifecycle configuration* is a set of rules that define actions that Amazon S3 applies to a group of objects. Actions are to either transition objects to another storage class or expire (delete) the objects.

In this case the lifecycle policy can be created to move the objects to S3 Glacier (lower cost archival) when they are no longer frequently accessed, and then expire the objects when they no longer need to be retained.

The following image shows the waterfall model for support transitions between storage classes:



CORRECT: "Store the documents in a secured Amazon S3 bucket with a lifecycle policy to move the documents that are older than 3 months to Amazon S3 Glacier, then expire the documents from Amazon S3 Glacier that are more than 7 years old" is the correct answer.

INCORRECT: "Store the documents in an encrypted EBS volume and create a cron job to delete the documents after 7 years" is incorrect. Amazon EBS volumes must be mounted to EC2 instances and this is not a cost-effective solution.

INCORRECT: "Store the documents in Amazon EFS. Create a cron job to move the documents that are older than 3 months to Amazon S3 Glacier. Create an AWS Lambda function to delete the documents in S3 Glacier that are older than 7 years" is incorrect. Amazon EFS filesystems must be mounted to EC2 instances and this is not a cost-effective solution.

INCORRECT: "Store the documents in a secured Amazon S3 bucket with a lifecycle policy to move the documents that are older than 3 months to Amazon S3 Glacier. Create an AWS Lambda function to delete the documents in S3 Glacier that are older than 7 years" is incorrect. It is not necessary to use a Lambda function to delete the objects, a lifecycle policy can be used instead and is more cost-effective.

References:

<https://docs.aws.amazon.com/AmazonS3/latest/dev/object-lifecycle-mgmt.html>

Save time with our AWS cheat sheets:

<https://digitalcloud.training/category/aws-cheat-sheets/aws-solutions-architect-professional/aws-storage-sap/>

Question 17: **Correct**

An eCommerce company are running a promotional campaign and expect a large volume of user sign-ups on a web page that collects user information and preferences. The website runs on Amazon EC2 instances and uses an Amazon RDS for PostgreSQL DB instance. The volume of traffic is expected to be high and may be unpredictable with several spikes in activity. The traffic will result in a large number of database writes.

A solutions architect needs to build a solution that does not change the underlying data model and ensures that submissions are not dropped before they are committed to the database. Which solution meets these requirements?



Migrate to Amazon DynamoDB and manage throughput capacity with automatic scaling.



Use scheduled scaling to scale up the existing DB instance immediately before the event and then automatically scale down afterwards.



Create an Amazon SQS queue and decouple the application and database layers. Configure an AWS Lambda function to write items from the queue into the database.

(Correct)



Create an Amazon ElastiCache for Memcached cluster in front of the existing database instance to increase write performance.

Explanation

In order to avoid dropping records decoupling the application and database layers is the best solution for this specific scenario. This works as the application does not require synchronous responses (it's just writing the user information to the DB). The alternative is to increase write capacity on the database instance but as the traffic is unpredictable it's hard to know how much capacity to provision which could lead to underperformance or higher than necessary costs.

CORRECT: "Create an Amazon SQS queue and decouple the application and database layers. Configure an AWS Lambda function to write items from the queue into the database" is the correct answer.

INCORRECT: "Use scheduled scaling to scale up the existing DB instance immediately before the event and then automatically scale down afterwards" is incorrect. You cannot schedule RDS database instances to scale up or down.

INCORRECT: "Migrate to Amazon DynamoDB and manage throughput capacity with automatic scaling" is incorrect. DynamoDB is a NoSQL (non-relational) database whereas RDS is a relational database. This solution would change the underlying data model and is therefore not an option.

INCORRECT: "Create an Amazon ElastiCache for Memcached cluster in front of the existing database instance to increase write performance" is incorrect. ElastiCache is used for improving read performance, not write performance.

References:

<https://docs.aws.amazon.com/AWSSimpleQueueService/latest/SQSDeveloperGuide/welcome.html>

Save time with our AWS cheat sheets:

<https://digitalcloud.training/category/aws-cheat-sheets/aws-solutions-architect-professional/aws-compute-sap/>

<https://digitalcloud.training/category/aws-cheat-sheets/aws-solutions-architect-professional/aws-application-integration-sap/>

Question 18: **Correct**

A database for an eCommerce website was deployed on an Amazon RDS for MySQL DB instance with General Purpose SSD storage. The database was running performantly for several weeks until a peak shopping period when customers experienced slow performance and timeouts. Amazon CloudWatch metrics indicate that reads and writes to the DB instance were experiencing long response times. Metrics show that CPU utilization is <50%, plenty of available memory, and sufficient free storage space. There is no evidence of database connectivity issues in the application server logs.

What could be the root cause of database performance issues?

A large number of reads and writes exhausted the network bandwidth available to the RDS for MySQL DB instance.

The increased load caused the data in the tables to change frequently, requiring indexes to be rebuilt to optimize queries.

A large number of reads and writes exhausted the I/O credit balance due to provisioning low disk storage during the setup phase.

(Correct)

The increased load resulted in the maximum number of allowed connections to the database instance.

Explanation

Baseline I/O performance for General Purpose SSD storage is 3 IOPS for each GiB, with a minimum of 100 IOPS. This relationship means that larger volumes have better performance. In this case the volume is only 100 GB so it will only have 300 IOPS performance.

When using General Purpose SSD storage, a DB instance receives an initial I/O credit balance of 5.4 million I/O credits. This initial credit balance is enough to sustain a burst performance of 3,000 IOPS for 30 minutes. This balance is designed to provide a fast initial boot cycle for boot volumes and to provide a good bootstrapping experience for other applications.

Volumes earn I/O credits at the baseline performance rate of 3 IOPS for each GiB of volume size. For example, a 100-GiB SSD volume has a baseline performance of 300 IOPS.

It is clear that in this scenario the increased load has caused the I/O credit balance to become exhausted before the end of the peak shopping period. This means that performance will be limited until there is sufficient I/O credit.

CORRECT: "A large number of reads and writes exhausted the I/O credit balance due to provisioning low disk storage during the setup phase" is the correct answer.

INCORRECT: "The increased load caused the data in the tables to change frequently, requiring indexes to be rebuilt to optimize queries" is incorrect. Reading and writing items to the table should not result in indexes being rebuilt.

INCORRECT: "The increased load resulted in the maximum number of allowed connections to the database instance" is incorrect. MySQL on RDS can have up to 100,000 client connections. In this case the application servers are the clients and it is unlikely there are that many app servers.

INCORRECT: "A large number of reads and writes exhausted the network bandwidth available to the RDS for MySQL DB instance" is incorrect. Based on the storage configuration presented it is far more likely that storage performance is the issue as in this setup it will be exhausted long before the network bandwidth.

References:

https://docs.aws.amazon.com/AmazonRDS/latest/UserGuide/CHAP_Storage.html

Save time with our AWS cheat sheets:

<https://digitalcloud.training/category/aws-cheat-sheets/aws-solutions-architect-professional/aws-storage-sap/>

Question 19: **Correct**

A solution is required for updating user metadata and will be initiated by a fleet of front-end web servers. The solution must be capable of scaling rapidly from hundreds to tens of thousands of jobs in less than a minute. The solution must be asynchronous and minimize costs.

Which solution should a Solutions Architect use to meet these requirements?

-

Create an AWS Lambda function that will update user metadata. Create an Amazon SQS queue and configure it as an event source for the Lambda function. Update the web application to send jobs to the queue.

(Correct)

-

Create an AWS Lambda function that will update user metadata. Create AWS Step Functions that will trigger the Lambda function. Update the web application to initiate Step Functions for every job.

-

Create an Amazon EC2 Auto Scaling group of EC2 instances that pull messages from an Amazon SQS queue and process the user metadata updates. Configure the web application to send jobs to the queue.

-

Create an AWS CloudFormation stack that is updated by an AWS Lambda function. Configure the Lambda function to update the metadata.

Explanation

The most cost-effective solution to this requirement will be to use a fully serverless and decoupled architecture. With Amazon SQS the web application can asynchronously place jobs in the queue. The queue can be configured as an event source for AWS Lambda which means the Lambda function will be triggered each time a job is placed in the queue. Lambda processes the messages synchronously.

CORRECT: "Create an AWS Lambda function that will update user metadata. Create an Amazon SQS queue and configure it as an event source for the Lambda function. Update the web application to send jobs to the queue" is the correct answer.

INCORRECT: "Create an AWS Lambda function that will update user metadata. Create AWS Step Functions that will trigger the Lambda function. Update the web application to initiate Step Functions for every job" is incorrect. Step Functions is used to coordinate multiple serverless functions in a workflow and is not necessary for this use case.

INCORRECT: "Create an AWS CloudFormation stack that is updated by an AWS Lambda function. Configure the Lambda function to update the metadata" is incorrect. This solution doesn't make much sense. CloudFormation is used for deploying infrastructure and there is no mention of what is triggering Lambda.

INCORRECT: "Create an Amazon EC2 Auto Scaling group of EC2 instances that pull messages from an Amazon SQS queue and process the user metadata updates. Configure the web application to send jobs to the queue" is incorrect. This would work but would be less cost-effective compared to using AWS Lambda.

References:

<https://docs.aws.amazon.com/lambda/latest/dg/with-sqs.html>

Save time with our AWS cheat sheets:

<https://digitalcloud.training/category/aws-cheat-sheets/aws-solutions-architect-professional/aws-compute-sap/>

<https://digitalcloud.training/category/aws-cheat-sheets/aws-solutions-architect-professional/aws-application-integration-sap/>

Question 20: **Correct**

A company runs hundreds of applications across several data centers and office locations. The applications include Windows and Linux operating systems, physical installations as well as virtualized servers, and MySQL and Oracle databases. There is no central configuration management database (CMDB) and existing documentation is incomplete and outdated. A Solutions Architect needs to understand the current environment and estimate the cloud resource costs after the migration.

Which tools or services should the Solutions Architect use to plan the cloud migration (Select THREE.)

-

AWS Migration Hub

(Correct)

<input type="checkbox"/>	AWS Server Migration Service
<input checked="" type="checkbox"/>	AWS Cloud Adoption Readiness Tool (CART) (Correct)
<input type="checkbox"/>	AWS Config
<input type="checkbox"/>	AWS CloudWatch Logs
<input checked="" type="checkbox"/>	AWS Application Discovery Service (Correct)
Explanation	
A combination of tools and services will assist this organization with planning their cloud migration. The AWS Application Discovery Service helps enterprise customers plan migration projects by gathering information about their on-premises data centers.	
AWS Application Discovery Service performs server utilization data and dependency mapping and collects and presents configuration, usage, and behavior data from your servers to help you better understand your workloads.	
You can export this data as a CSV file and use it to estimate the Total Cost of Ownership (TCO) of running on AWS and to plan your migration to AWS. In addition, this data is also available in AWS Migration Hub, where you can migrate the discovered servers and track their progress as they get migrated to AWS.	
The AWS Cloud Adoption Readiness Tool (CART) helps organizations of all sizes develop efficient and effective plans for cloud adoption and enterprise cloud migrations. This 16-question online survey and assessment report details your cloud migration readiness across six perspectives including business, people, process, platform, operations, and security.	

CORRECT: "AWS Application Discovery Service" is a correct answer.

CORRECT: "AWS Cloud Adoption Readiness Tool (CART)" is also a correct answer.

CORRECT: "AWS Migration Hub" is also a correct answer.

INCORRECT: "AWS Server Migration Service" is incorrect. The SMS service may be used to implement the migrations of some servers but it is not used for the planning phase.

INCORRECT: "AWS Config" is incorrect. AWS Config is used to assess, audit, and evaluate the configurations of your AWS resources. It does not assess on-premises resources.

INCORRECT: "AWS CloudWatch Logs" is incorrect. This service collects application and system log files. It can collect log files from on-premises systems but these are not required for planning a migration.

References:

<https://aws.amazon.com/application-discovery/>

<https://cloudreadiness.amazonaws.com/#/cart>

<https://aws.amazon.com/migration-hub/>

Save time with our AWS cheat sheets:

<https://digitalcloud.training/category/aws-cheat-sheets/aws-solutions-architect-professional/aws-database-sap/>

Question 21: **Correct**

A company requires multi-Region availability for an application that runs on Amazon EC2 instances with an Amazon RDS for MySQL database. The solution must offer the highest availability.

Which solution should a solutions architect recommend?

-

Enable a multi-master cluster configuration across multiple Regions. Store the DB endpoint in AWS Secrets Manager. In the case of an outage, use AWS Lambda to update the endpoint address used by the applications.

-

Enable automated backups for the RDS database instance. In the case of an outage, promote the automated backup to be a standalone DB instance. Point applications to the new DB endpoint and create a read replica to maintain high availability.



Enable global tables for the RDS database instance across multiple Regions. Store the DB endpoint in AWS Secrets Manager. In the case of an outage, update the DB endpoint in Secrets Manager to the cross-Region table endpoint.



Enable a cross-Region read replica for the RDS database. In the case of an outage, promote the replica to be a standalone DB instance. Point applications to the new DB endpoint and create a read replica to maintain high availability.

(Correct)

Explanation

You can promote a read replica to a standalone instance as a disaster recovery solution if the primary DB instance fails. This is a fairly quick and easy process. Applications will then need to be redirected to point to the endpoint of the newly promoted DB instance. To maintain availability a new read replica can then be created.



Amazon RDS Multi-AZ and Read Replicas

Multi-AZ Deployments	Read Replicas
Synchronous replication – highly durable	Asynchronous replication – highly scalable
Only database engine on primary instance is active	All read replicas are accessible and can be used for read scaling
Automated backups are taken from standby	No backups configured by default
Always span two Availability Zones within a single Region	Can be within an Availability Zone, Cross-AZ, or Cross-Region
Database engine version upgrades happen on primary	Database engine version upgrade is independent from source instance
Automatic failover to standby when a problem is detected	Can be manually promoted to a standalone database instance

© Digital Cloud Training | <https://digitalcloud.training>



CORRECT: "Enable a cross-Region read replica for the RDS database. In the case of an outage, promote the replica to be a standalone DB instance. Point applications to the new DB endpoint and create a read replica to maintain high availability" is the correct answer (as explained above.)

INCORRECT: "Enable automated backups for the RDS database instance. In the case of an outage, promote the automated backup to be a standalone DB instance. Point applications to the new DB endpoint and create a read replica to maintain high availability" is incorrect.

You cannot promote automated backups. You can use them to restore databases to a specific point in time or create a new DB instance. This will not be useful if a regional outage occurs as the automated backup is stored within the same Region by default.

INCORRECT: "Enable global tables for the RDS database instance across multiple Regions. Store the DB endpoint in AWS Secrets Manager. In the case of an outage, update the DB endpoint in Secrets Manager to the cross-Region table endpoint" is incorrect.

Global tables is a feature of Amazon DynamoDB and is not available for Amazon RDS.

INCORRECT: "Enable a multi-master cluster configuration across multiple Regions. Store the DB endpoint in AWS Secrets Manager. In the case of an outage, use AWS Lambda to update the endpoint address used by the applications" is incorrect.

Multi-master configurations are only available for Amazon Aurora and only within a Region.

References:

https://docs.aws.amazon.com/AmazonRDS/latest/UserGuide/USER_ReadRepl.html

Save time with our AWS cheat sheets:

<https://digitalcloud.training/amazon-rds/>

Question 22: **Incorrect**

A company uses AWS Organizations. The company recently acquired a new business unit and invited the new unit's existing account to the company's organization. The organization uses a deny list SCP in the root of the organization and all accounts are members of a single OU named Production.

The administrators of the new business unit discovered that they are unable to access AWS Database Migration Service (DMS) to complete an in-progress migration.

Which option will temporarily allow administrators to access AWS DMS and complete the migration project?

-
-

Remove the organization's root SCPs that limit access to AWS DMS. Create an SCP that allows AWS DMS actions and apply the SCP to the Production OU.

-
-

Convert the organization's root SCPs from deny list SCPs to allow list SCPs to allow the required services only. Temporarily apply an SCP to the organization's root that allows AWS DMS actions for principals only in the new account.

-
-

Create a temporary OU named Staging for the new account. Apply an SCP to the Staging OU to allow AWS DMS actions. Move the new account to the Production OU when the migration project is complete.

(Incorrect)



Create a temporary OU named Staging for the new account. Apply an SCP to the Staging OU to allow AWS DMS actions. Move the organization's deny list SCP to the Production OU. Move the new account to the Production OU when adjustments to AWS DMS are complete.

(Correct)

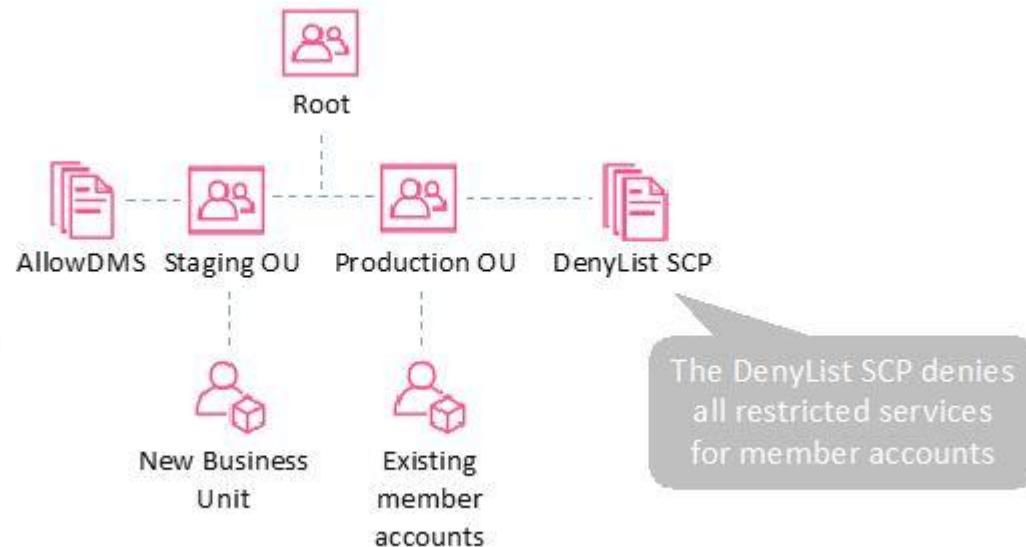
Explanation

A deny list strategy uses an implicit deny and has an SCP named AWSFullAccess applied at the root level (by default) which allows all actions. In this case the company has applied a deny list SCP at the root level which denies access to specific services.

In AWS Organizations any account has only those permissions permitted by **every** parent above it. If a permission is blocked at any level above the account, either implicitly (by not being included in an Allow policy statement) or explicitly (by being included in a Deny policy statement), a user or role in the affected account can't use that permission.

Therefore, it will not be possible to allow services in an OU that have been denied at the root level. The only solution is to move the deny list from the root level to the Production OU (which means it is still effective for all other accounts) and then create a temporary OU with an SCP that allows AWS DMS (the AWSFullAccess would do this if it has not been removed).

The diagram below depicts the temporary configuration after the DenyList SCP has been moved to the Production OU:



CORRECT: "Create a temporary OU named Staging for the new account. Apply an SCP to the Staging OU to allow AWS DMS actions. Move the organization's deny list SCP to the Production OU. Move the new account to the Production OU when adjustments to AWS DMS are complete" is the correct answer.

INCORRECT: "Remove the organization's root SCPs that limit access to AWS DMS. Create an SCP that allows AWS DMS actions and apply the SCP to the Production OU" is incorrect. This would enable AWS DMS for all member accounts which is more permissions than is required so this is not the best option.

INCORRECT: "Create a temporary OU named Staging for the new account. Apply an SCP to the Staging OU to allow AWS DMS actions. Move the new account to the Production OU when the migration project is complete" is incorrect. The deny list SCP at the root level will not allow the restricted actions to be allowed at any level beneath so this will not work.

INCORRECT: "Convert the organization's root SCPs from deny list SCPs to allow list SCPs to allow the required services only. Temporarily apply an SCP to the organization's root that allows AWS DMS actions for principals only in the new account" is incorrect. There is considerably more work involved with converting the SCPs, it would be much simpler to move the deny list SCP from the root to the Production OU to remove restrictions from higher in the hierarchy.

References:

https://docs.aws.amazon.com/organizations/latest/userguide/orgs_manage_policies_scps.html

Save time with our AWS cheat sheets:

<https://digitalcloud.training/category/aws-cheat-sheets/aws-solutions-architect-professional/security-identity-compliance-sap/>

Question 23: **Correct**

An application consists of three tiers within a single Region. A Solutions Architect is designing a disaster recovery strategy that includes an RTO of 30 minutes and an RPO of 5 minutes for the data tier. Application tiers use Amazon EC2 instances and are stateless. The data tier consists of a 30TB Amazon Aurora database.

Which combination of steps satisfies the RTO and RPO requirements while optimizing costs? (Select TWO.)

-

Create a cross-Region Aurora MySQL Replica of the database.

(Correct)

-

Create snapshots of the Aurora database every 5 minutes.

-

Create daily snapshots of the EC2 instances and replicate to another Region.

-

Deploy a hot standby of the application tiers to another Region.

(Correct)

-

Use AWS DMS to replicate the Aurora DB to an RDS database in another Region.

Explanation

The recovery time objective (RTO) defines how quickly a service must be restored and a recovery point objective (RPO) defines how much data it is acceptable to lose. For example an RTO of 30 minutes means the service must be running again within half an hour and an RPO of 5 minutes means no more than 5 minutes' worth of data can be lost.

To achieve these requirements in this scenario a host standby is required of the EC2 instances. With a hot standby a minimum of application/web servers should be running and can be scaled out as required.

For the data tier an Amazon Aurora cross-Region Replica is the best way to ensure that <5mins of data is lost. You can promote an Aurora Read Replica to a standalone DB cluster, and this would be performed in the event of a disaster affecting the source DB cluster.

CORRECT: "Deploy a hot standby of the application tiers to another Region" is a correct answer.

CORRECT: "Create a cross-Region Aurora MySQL Replica of the database" is also a correct answer.

INCORRECT: "Create daily snapshots of the EC2 instances and replicate to another Region" is incorrect. Snapshots could be used to create an AMI and launch EC2 instances in the second Region. However, depending on the specifics of the application this could take longer than 30 minutes.

INCORRECT: "Create snapshots of the Aurora database every 5 minutes" is incorrect. Aurora backs up your cluster volume automatically and retains restore data for the length of the backup retention period. Snapshots are used to retain data for longer than the retention period and cost extra.

INCORRECT: "Use AWS DMS to replicate the Aurora DB to an RDS database in another Region" is incorrect. There is no need to use AWS Database Migration Service (DMS) or to replicate data to an RDS database. Aurora can provide the required functionality natively.

References:

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

Save time with our AWS cheat sheets:

<https://digitalcloud.training/category/aws-cheat-sheets/aws-solutions-architect-professional/aws-storage-sap/>

<https://digitalcloud.training/category/aws-cheat-sheets/aws-solutions-architect-professional/aws-compute-sap/>

Question 24: **Correct**

A company wants to run an application on AWS. The company plans to provision its application in Docker containers running in an Amazon ECS cluster. The application requires a MySQL database and the company plans to use Amazon RDS.

What is the MOST cost-effective solution to meet these requirements?



Create an ECS cluster using a fleet of Spot Instances with Spot Instance draining enabled. Provision the database using On-Demand Instances.



Create an ECS cluster using On-Demand Instances. Provision the database using On-Demand Instances.

Create an ECS cluster using On-Demand Instances. Provision the database using Spot Instances.

Create an ECS cluster using a fleet of Spot Instances, with Spot Instance draining enabled. Provision the database using Reserved Instances.

(Correct)

Explanation

The most cost-effective solution is a combination of Spot instances for the ECS cluster and reserved instances for the database. If Amazon ECS Spot Instance draining is enabled on the instance, ECS receives the Spot Instance interruption notice and places the instance in DRAINING status.

Based on the facts provided in the question this is the best combination of options presented. All other options are less cost-effective. Note that if the application cannot terminate interruption (not specified), using Spot instances will not be an ideal solution.

CORRECT: "Create an ECS cluster using a fleet of Spot Instances, with Spot Instance draining enabled. Provision the database using Reserved Instances" is the correct answer.

INCORRECT: "Create an ECS cluster using On-Demand Instances. Provision the database using Spot Instances" is incorrect. On-demand instances do not provide any cost benefits so this is not a cost-effective solution.

INCORRECT: "Create an ECS cluster using On-Demand Instances. Provision the database using On-Demand Instances" is incorrect. On-demand instances do not provide any cost benefits so this is not a cost-effective solution.

INCORRECT: "Create an ECS cluster using a fleet of Spot Instances with Spot Instance draining enabled. Provision the database using On-Demand Instances" is incorrect. On-demand instances do not provide any cost benefits so this is not a cost-effective solution.

References:

<https://docs.aws.amazon.com/AmazonECS/latest/developerguide/container-instance-spot.html>

Save time with our AWS cheat sheets:

<https://digitalcloud.training/category/aws-cheat-sheets/aws-solutions-architect-professional/aws-compute-sap/>

Question 25: **Correct**

A new AWS Lambda function has been created to replicate objects that are received in an Amazon S3 bucket to several other S3 buckets in various AWS accounts. The Lambda function is triggered when an object create event occurs in the main S3 bucket. A Solutions Architect is concerned that the function may impact other critical functions due to Lambda's regional concurrency limit.

How can the solutions architect ensure the new Lambda function will not impact other critical Lambda functions?



Modify the execution timeout for the Lambda function to the maximum allowable value. Monitor existing critical Lambda functions with Amazon CloudWatch alarms for the Throttles Lambda metric.



Ensure the new Lambda function implements an exponential backoff algorithm. Monitor existing critical Lambda functions with Amazon CloudWatch alarms for the Throttles Lambda metric.



Configure the reserved concurrency limit for the new Lambda function. Monitor existing critical Lambda functions with Amazon CloudWatch alarms for the Throttles Lambda metric.

(Correct)



Configure Amazon S3 event notifications to publish events to an Amazon S3 queue in a different account. Create the Lambda function in the same account as the SQS queue and trigger the function when messages are published to the queue.

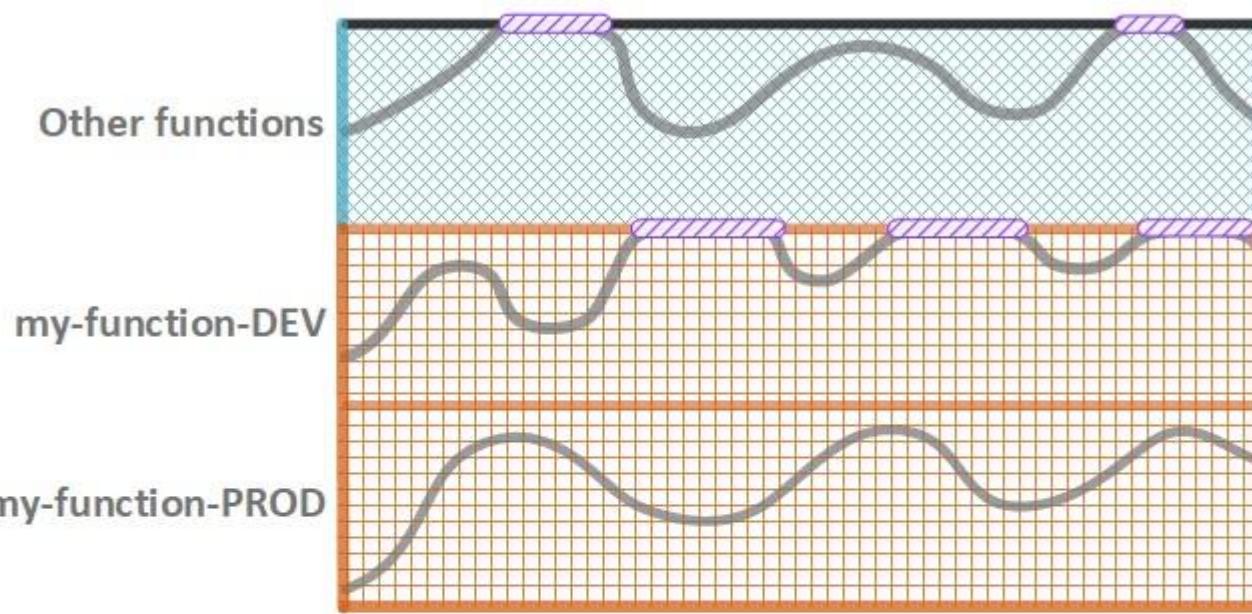
Explanation

Concurrency is the number of requests that your function is serving at any given time. When your function is invoked, Lambda allocates an instance of it to process the event. When the function code finishes running, it can handle another request. If the function is invoked again while a request is still being processed, another instance is allocated, which increases the function's concurrency.

Concurrency is subject to a Regional quota that is shared by all functions in a Region. To ensure that a function can always reach a certain level of concurrency, you can configure the function with reserved concurrency. When a function has reserved concurrency, no other function can use that concurrency. Reserved concurrency also limits the maximum concurrency for the function, and applies to the function as a whole, including versions and aliases.

Applying a reserved concurrency limit will ensure that the function does not use more than a specific maximum that is defined. To ensure other functions have adequate capacity, the Throttles Lambda metric can also be monitored which records the number of invocation requests that are throttled.

Reserved Concurrency



Legend

- Function concurrency
- Reserved concurrency
- Unreserved concurrency
- Throttling

CORRECT: "Configure the reserved concurrency limit for the new Lambda function. Monitor existing critical Lambda functions with Amazon CloudWatch alarms for the Throttles Lambda metric" is the correct answer.

INCORRECT: "Modify the execution timeout for the Lambda function to the maximum allowable value. Monitor existing critical Lambda functions with Amazon CloudWatch alarms for the Throttles Lambda metric" is incorrect. You can increase the execution time up to 15 minutes but this does not assist with ensuring that the other critical functions are not affected.

INCORRECT: "Configure Amazon S3 event notifications to publish events to an Amazon S3 queue in a different account. Create the Lambda function in the same account as the SQS queue and trigger the function when messages are published to the queue" is incorrect. You cannot publish S3 event notifications to an SQS queue in a different account.

INCORRECT: "Ensure the new Lambda function implements an exponential backoff algorithm. Monitor existing critical Lambda functions with Amazon CloudWatch alarms for the Throttles Lambda metric" is incorrect. The AWS SDK implements exponential backoff for better flow control. However, in this case setting the reserved concurrency limit will ensure that function leaves adequate capacity for other functions within the Region.

References:

<https://docs.aws.amazon.com/lambda/latest/dg/configuration-concurrency.html#configuration-concurrency-reserved>

Save time with our AWS cheat sheets:

<https://digitalcloud.training/category/aws-cheat-sheets/aws-solutions-architect-professional/aws-compute-sap/>

Question 26: **Correct**

A company has deployed an application that uses an Amazon DynamoDB table and the user base has increased significantly. Users have reported poor response times during busy periods but no error pages have been generated. The application uses Amazon DynamoDB in read-only mode. The operations team has determined that the issue relates to ProvisionedThroughputExceeded exceptions in the application logs when doing Scan and read operations.

A Solutions Architect has been tasked with improving application performance. Which solutions will meet these requirements whilst MINIMIZING changes to the application? (Select TWO.)

-

Enable adaptive capacity for the DynamoDB table to minimize throttling due to throughput exceptions.

-

Enable DynamoDB Auto Scaling to manage the throughput capacity as table traffic increases. Set the upper and lower limits to control costs and set a target utilization based on the peak usage.

(Correct)

-

Provision an Amazon ElastiCache for Redis cluster. The cluster should be provisioned with enough shards to handle the peak application load.



Provision a DynamoDB Accelerator (DAX) cluster with the correct number and type of nodes. Tune the item and query cache configuration for an optimal user experience.

(Correct)



Include error retries and exponential backoffs in the application code to handle throttling errors and reduce load during periods of high requests.

Explanation

When a ProvisionedThroughputExceeded error is generated it means that insufficient throughput has been enabled on the table. In this case that would be insufficient read capacity units (RCUs). Enabling DynamoDB Auto Scaling will ensure that the RCUs are adjusted based on load.

You configure Auto Scaling by specifying the minimum and maximum capacity units and the target utilization as you can see in the image below:

Auto scaling [Info](#)

Dynamically adjusts provisioned throughput capacity on your behalf in response to actual traffic patterns.

On

Off

Minimum capacity units

50

Maximum capacity units

100

Target utilization (%)

70

Another great addition to the solution is to create an Amazon DynamoDB Accelerator (DAX) cluster. DAX is a caching solution for DynamoDB that can be placed in front of the database. This will provide much improved read performance without any application changes.

CORRECT: "Enable DynamoDB Auto Scaling to manage the throughput capacity as table traffic increases. Set the upper and lower limits to control costs and set a target utilization based on the peak usage" is a correct answer.

CORRECT: "Provision a DynamoDB Accelerator (DAX) cluster with the correct number and type of nodes. Tune the item and query cache configuration for an optimal user experience" is also a correct answer.

INCORRECT: "Enable adaptive capacity for the DynamoDB table to minimize throttling due to throughput exceptions" is incorrect. Adaptive capacity is enabled automatically for every DynamoDB table, at no additional cost. You don't need to explicitly enable or disable it.

INCORRECT: "Provision an Amazon ElastiCache for Redis cluster. The cluster should be provisioned with enough shards to handle the peak application load" is incorrect. DAX is a better solution for a DynamoDB table as it works without any code changes which is preferred in this scenario.

INCORRECT: "Include error retries and exponential backoffs in the application code to handle throttling errors and reduce load during periods of high requests" is incorrect. This option also requires application code changes so should be avoided in this scenario.

References:

<https://docs.aws.amazon.com/amazondynamodb/latest/developerguide/AutoScaling.html>

<https://docs.aws.amazon.com/amazondynamodb/latest/developerguide/DAX.concepts.html>

Save time with our AWS cheat sheets:

<https://digitalcloud.training/category/aws-cheat-sheets/aws-solutions-architect-professional/aws-database-sap/>

Question 27: **Correct**

A global enterprise company is in the process of creating an infrastructure services platform for its users. The company has the following requirements:

- Centrally manage the creation of infrastructure services using a central AWS account.
- Distribute infrastructure services to multiple accounts in AWS Organizations.
- Follow the principle of least privilege to limit end users' permissions for launching and managing applications.

Which combination of actions using AWS services will meet these requirements? (Select TWO.)



Define the infrastructure services in AWS CloudFormation templates. Upload each template as an AWS Service Catalog product to portfolios created in a central AWS account. Share these portfolios with the AWS Organizations structure created for the company.

(Correct)

-

Allow IAM users to have `AWSServiceCatalogEndUserFullAccess` permissions. Assign the policy to a group called `Endusers`, add all users to the group. Apply launch constraints.

-

Define the infrastructure services in AWS CloudFormation templates. Add the templates to a central Amazon S3 bucket and add the IAM users that require access to the S3 bucket policy.

-

Allow IAM users to have `AWSServiceCatalogEndUserReadOnlyAccess` permissions only. Assign the policy to a group called `Endusers`, add all users to the group. Apply launch constraints.

(Correct)

-

Grant IAM users `AWSCloudFormationFullAccess` and `AmazonS3ReadOnlyAccess` permissions. Add an Organizations SCP at the AWS account root user level to deny all services except AWS CloudFormation and Amazon S3.

Explanation

There are three core requirements for this solution. The first two requirements are satisfied by adding each CloudFormation template to a product in AWS Service Catalog in a central AWS account and then sharing the portfolio with AWS Organizations.

In this model, the central AWS account hosts the organizationally approved infrastructure services and shares them to other AWS accounts in the company. AWS Service Catalog administrators can reference an existing organization in AWS Organizations when sharing a portfolio, and they can share the portfolio with any trusted organizational unit (OU) in the organization's tree structure.

The third requirement is satisfied by using a permissions policy with read only access to AWS Service Catalog combined with a launch constraint that will use a dedicated IAM role that ensures least privilege access.

Without a launch constraint, end users must launch and manage products using their own IAM credentials. To do so, they must have permissions for AWS CloudFormation, the AWS services used by the products, and AWS Service Catalog. By using a launch role, you can instead limit the end users' permissions to the minimum that they require for that product.

CORRECT: "Define the infrastructure services in AWS CloudFormation templates. Upload each template as an AWS Service Catalog product to portfolios created in a central AWS account. Share these portfolios with the AWS Organizations structure created for the company" is a correct answer.

CORRECT: "Allow IAM users to have `AWSServiceCatalogEndUserReadOnlyAccess` permissions only. Assign the policy to a group called `Endusers`, add all users to the group. Apply launch constraints" is also a correct answer.

INCORRECT: "Define the infrastructure services in AWS CloudFormation templates. Add the templates to a central Amazon S3 bucket and add the IAM users that require access to the S3 bucket policy" is incorrect. This uses a central account but doesn't have offer a mechanism to distribute the templates to accounts in AWS Organizations. It would also be very hard to manage access when adding users to bucket policies.

INCORRECT: "Grant IAM users AWSCloudFormationFullAccess and AmazonS3ReadOnlyAccess permissions. Add an Organizations SCP at the AWS account root user level to deny all services except AWS CloudFormation and Amazon S3" is incorrect. When launching services using CloudFormation, the principal used (user or role) must have permissions to the AWS services being launched through the template. This solution does not provide those permissions.

INCORRECT: "Allow IAM users to have AWSServiceCatalogEndUserFullAccess permissions. Assign the policy to a group called Endusers, add all users to the group. Apply launch constraints" is incorrect. Users do not need full access, read only is sufficient as it does not provide the ability for users to launch and manage products using their own accounts. The launch constraint provides the necessary permissions for launching products using an assigned role.

References:

https://docs.aws.amazon.com/servicecatalog/latest/adminguide/controlling_access.html

<https://docs.aws.amazon.com/organizations/latest/userguide/services-that-can-integrate-servicecatalog.html>

Save time with our AWS cheat sheets:

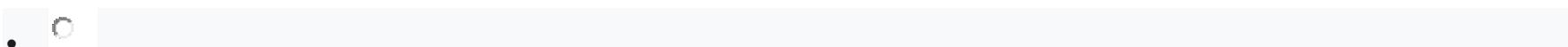
<https://digitalcloud.training/category/aws-cheat-sheets/aws-solutions-architect-professional/aws-networking-content-delivery-sap/>

<https://digitalcloud.training/category/aws-cheat-sheets/aws-solutions-architect-professional/security-identity-compliance-sap/>

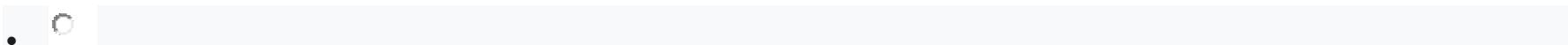
Question 28: **Correct**

An application stores user comment data in multiple Amazon DynamoDB tables. A solutions architect must use a serverless architecture to make the data accessible publicly through a simple and cost-effective API over HTTPS. The solution must scale automatically in response to demand.

Which solutions meet these requirements?



Create an Amazon API Gateway REST API. Configure this API with direct integrations to DynamoDB by using API Gateway's AWS Service integration type.



Create an Amazon API Gateway REST API. Configure this API with integrations to AWS Lambda functions that return data from the DynamoDB tables.

Create an Amazon API Gateway HTTP API. Configure this API with integrations to AWS Lambda functions that return data from the DynamoDB tables.

(Correct)

Create an Amazon API Gateway HTTP API. Configure this API with direct integrations to DynamoDB by using API Gateway's AWS Service integration type.

Explanation

A REST API must be used to provide a direct AWS Service integration to Amazon DynamoDB. An AWS Lambda function is not needed in that case as direct integration is possible and API Gateway will scale seamlessly.

However, in this case the requirement is for a simple and cost-effective API. This is therefore a good use case for an HTTP API which is lower cost. With an HTTP API direct integration to DynamoDB is not possible but you can connect to multiple Lambda functions and configure methods and paths. This solution meets all requirements.

CORRECT: "Create an Amazon API Gateway HTTP API. Configure this API with integrations to AWS Lambda functions that return data from the DynamoDB tables" is the correct answer (as explained above.)

INCORRECT: "Create an Amazon API Gateway HTTP API. Configure this API with direct integrations to DynamoDB by using API Gateway's AWS Service integration type" is incorrect.

You cannot create a direct integration to DynamoDB when using an HTTP API.

INCORRECT: "Create an Amazon API Gateway REST API. Configure this API with integrations to AWS Lambda functions that return data from the DynamoDB tables" is incorrect.

This is a less cost-effective solution as explained above.

INCORRECT: "Create an Amazon API Gateway REST API. Configure this API with direct integrations to DynamoDB by using API Gateway's AWS Service integration type" is incorrect.

This is a less cost-effective solution as explained above.

References:

<https://aws.amazon.com/blogs/compute/using-amazon-api-gateway-as-a-proxy-for-dynamodb/>

Save time with our AWS cheat sheets:

<https://digitalcloud.training/amazon-api-gateway/>

Question 29: **Correct**

A company has a mobile application that uses Amazon API Gateway, AWS Lambda, and Amazon DynamoDB. The application is write intensive and costs have recently increased significantly. The biggest increase in cost has been for the AWS Lambda functions. Application utilization is unpredictable but has been increasing steadily each month.

A Solutions Architect has noticed that the Lambda function execution time averages over 4 minutes. This is due to wait time for a high-latency network call to an on-premises MySQL database. A VPN is used to connect to the VPC.

How can the Solutions Architect reduce the cost of the current architecture?



- Migrate the MySQL database server into a Multi-AZ Amazon RDS for MySQL.
- Enable caching of the Amazon API Gateway results in Amazon CloudFront to reduce the number of Lambda function invocations.
- Enable DynamoDB Accelerator for frequently accessed records and enable the DynamoDB Auto Scaling feature.



- Replace the VPN with AWS Direct Connect to reduce the network latency to the on-premises MySQL database.
- Enable local caching in the mobile application to reduce the Lambda function invocation calls.
- Offload the frequently accessed records from DynamoDB to Amazon ElastiCache.



- Replace the VPN with AWS Direct Connect to reduce the network latency to the on-premises MySQL database.
- Cache the API Gateway results to Amazon CloudFront.
- Use Amazon EC2 Reserved Instances instead of Lambda.
- Enable Auto Scaling on EC2 and use Spot Instances during peak times.
- Enable DynamoDB Auto Scaling to manage target utilization.



- Migrate the MySQL database server into a Multi-AZ Amazon RDS for MySQL.
- Enable API caching on API Gateway to reduce the number of Lambda function invocations.
- Enable Auto Scaling in DynamoDB.

(Correct)

Explanation

The best way to reduce the latency of the network call to the on-premises database is to move the database to AWS using Amazon RDS. Additionally, API caching will cache the API responses for an API Gateway stage which further improves performance. Finally, enabling Auto Scaling in DynamoDB ensures that the read and write capacity of the table will adjust according to load which further increases cost efficiency.

CORRECT:

- Migrate the MySQL database server into a Multi-AZ Amazon RDS for MySQL.
- Enable API caching on API Gateway to reduce the number of Lambda function invocations.
- Enable Auto Scaling in DynamoDB.

INCORRECT:

- Replace the VPN with AWS Direct Connect to reduce the network latency to the on-premises MySQL database.
- Enable local caching in the mobile application to reduce the Lambda function invocation calls.
- Offload the frequently accessed records from DynamoDB to Amazon ElastiCache.

AWS Direct Connect will reduce latency however it comes at a significant cost. Using local caching in the mobile application may ensure some performance benefits but will not prevent the high-latency network calls from happening. ElastiCache can be used to cache DynamoDB table contents however DynamoDB Accelerator may be easier to implement in front of DDB.

INCORRECT:

- Replace the VPN with AWS Direct Connect to reduce the network latency to the on-premises MySQL database.
- Cache the API Gateway results to Amazon CloudFront.
- Use Amazon EC2 Reserved Instances instead of Lambda.

- Enable Auto Scaling on EC2 and use Spot Instances during peak times.
- Enable DynamoDB Auto Scaling to manage target utilization.

API Gateway results cannot be cached in CloudFront. EC2 RIs are unlikely to be more cost efficient compared to Lambda functions. The key is to prevent the high-latency network calls from occurring which will be the best resolution to the problem. Auto Scaling Spot instances and DynamoDB Auto Scaling are both valid options for cost and performance optimization.

INCORRECT:

- Migrate the MySQL database server into a Multi-AZ Amazon RDS for MySQL.
- Enable caching of the Amazon API Gateway results in Amazon CloudFront to reduce the number of Lambda function invocations.
- Enable DynamoDB Accelerator for frequently accessed records and enable the DynamoDB Auto Scaling feature.

You cannot cache API Gateway results in CloudFront. Otherwise this is a good solution.

References:

<https://docs.aws.amazon.com/apigateway/latest/developerguide/api-gateway-caching.html>

<https://docs.aws.amazon.com/amazondynamodb/latest/developerguide/AutoScaling.html>

<https://aws.amazon.com/getting-started/hands-on/create-mysql-db/>

Save time with our AWS cheat sheets:

<https://digitalcloud.training/category/aws-cheat-sheets/aws-solutions-architect-professional/aws-application-integration-sap/>

Question 30: **Correct**

A company is using multiple AWS accounts. The company's DNS records are stored in a private Amazon Route 53 hosted zone in the management account and their applications are running in a production account.

A Solutions Architect is attempting to deploy an application into the production account. The application must resolve a CNAME record set for an Amazon RDS endpoint. The CNAME record set was created in a private hosted zone in the management account.

The deployment failed to start and the Solutions Architect has discovered that the CNAME record is not resolvable on the application EC2 instance despite being correctly created in Route 53.

Which combination of steps should the Solutions Architect take to resolve this issue? (Select TWO.)

- Create a private hosted zone for the record set in the production account. Configure Route 53 replication between AWS accounts.
 - Associate a new VPC in the production account with a hosted zone in the management account. Delete the association authorization in the management account.

(Correct)
 - Deploy the database on a separate EC2 instance in the new VPC. Create a record set for the instance's private IP in the private hosted zone.
 - Hardcode the DNS name and IP address of the RDS database instance into the /etc/resolv.conf file on the application server.
 - Create an authorization to associate the private hosted zone in the management account with the new VPC in the production account.

(Correct)
- ## Explanation
- An application cannot resolve record sets created in the private hosted zone of another AWS account. The solution to this problem is to associate the Route 53 private hosted zone in the management account with the VPC in the production account.
- To associate a Route 53 private hosted zone in one AWS account (Account A) with a virtual private cloud that belongs to another AWS account (Account B), follow these steps using the AWS CLI:
1. From an instance in Account A, authorize the association between the private hosted zone in Account A and the virtual private cloud in Account B.
 2. From an instance in Account B, create the association between the private hosted zone in Account A and the virtual private cloud in Account B.
 3. Delete the association authorization after the association is created.

CORRECT: "Create an authorization to associate the private hosted zone in the management account with the new VPC in the production account" is a correct answer.

CORRECT: "Associate a new VPC in the production account with a hosted zone in the management account. Delete the association authorization in the management account" is also a correct answer.

INCORRECT: "Deploy the database on a separate EC2 instance in the new VPC. Create a record set for the instance's private IP in the private hosted zone" is incorrect. The record should be a CNAME record that points to the DNS endpoint of the RDS database, not to a private IP address.

INCORRECT: "Hardcode the DNS name and IP address of the RDS database instance into the /etc/resolv.conf file on the application server" is incorrect. This is not a best practice as the IP address of the RDS instance should not be used, a CNAME pointing to its DNS endpoint is preferred.

INCORRECT: "Create a private hosted zone for the record set in the production account. Configure Route 53 replication between AWS accounts" is incorrect. You cannot configure replication for hosted zones in Route 53.

References:

<https://aws.amazon.com/premiumsupport/knowledge-center/private-hosted-zone-different-account/>

Save time with our AWS cheat sheets:

<https://digitalcloud.training/category/aws-cheat-sheets/aws-solutions-architect-professional/aws-networking-content-delivery-sap/>

Question 1: **Correct**

A company is planning to migrate an application from an on-premises data center to the AWS Cloud. The application consists of a stateful servers and a separate MySQL database. The application is expected to receive significant traffic and must scale seamlessly. The solution design on AWS includes an Amazon Aurora MySQL database, Amazon EC2 Auto Scaling and Elastic Load Balancing.

A Solutions Architect needs to finalize the design for the solution. Which of the following configurations will ensure a consistent user experience and seamless scalability for both the application and database tiers?

-

Add Aurora Replicas and define a scaling policy. Use an Application Load Balancer and set the load balancing algorithm type to round_robin.

(Correct)

-

Add Aurora Replicas and define a scaling policy. Use a Network Load Balancer and set the load balancing algorithm type to round_robin.

-

Add Aurora Replicas and define a scaling policy. Use an Application Load Balancer and set the load balancing algorithm type to least_outstanding_requests.

-

Add Aurora Replicas and define a scaling policy. Use a Network Load Balancer and set the load balancing algorithm type to least_outstanding_requests.

Explanation

Aurora Auto Scaling dynamically adjusts the number of Aurora Replicas provisioned for an Aurora DB cluster using single-master replication. You define and apply a scaling policy to an Aurora DB cluster.

The *scaling policy* defines the minimum and maximum number of Aurora Replicas that Aurora Auto Scaling can manage. Based on the policy, Aurora Auto Scaling adjusts the number of Aurora Replicas up or down in response to actual workloads, determined by using Amazon CloudWatch metrics and target values.

By default, the round robin routing algorithm is used to route requests at the target group level. You can specify the least outstanding requests routing algorithm instead.

Consider using least outstanding requests when the requests for your application vary in complexity or your targets vary in processing capability. Round robin is a good choice when the requests and targets are similar, or if you need to distribute requests equally among targets.

In this case the round robin algorithm will be the best choice as the instances will have the same processing capability and requests should be routed evenly between them.

CORRECT: "Add Aurora Replicas and define a scaling policy. Use an Application Load Balancer and set the load balancing algorithm type to round_robin" is the correct answer.

INCORRECT: "Add Aurora Replicas and define a scaling policy. Use an Application Load Balancer and set the load balancing algorithm type to least_outstanding_requests" is incorrect. The least outstanding requests algorithm is not the best choice here as explained above.

INCORRECT: "Add Aurora Replicas and define a scaling policy. Use a Network Load Balancer and set the load balancing algorithm type to least_outstanding_requests" is incorrect. The NLB does not use this algorithm, it uses a flow hash algorithm.

INCORRECT: "Add Aurora Replicas and define a scaling policy. Use a Network Load Balancer and set the load balancing algorithm type to round_robin" is incorrect. The NLB does not use this algorithm, it uses a flow hash algorithm.

References:

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

<https://docs.aws.amazon.com/AmazonRDS/latest/AuroraUserGuide/Aurora.Integrating.AutoScale.html>

Save time with our AWS cheat sheets:

<https://digitalcloud.training/category/aws-cheat-sheets/aws-solutions-architect-professional/aws-database-sap/>

Question 2: Correct

A developer is attempting to access an Amazon S3 bucket in a member account in AWS Organizations. The developer is logged in to the account with user credentials and has received an access denied error with no bucket listed. The developer should have read-only access to all buckets in the account.

A Solutions Architect has reviewed the permissions and found that the developer's IAM user has been granted read-only access to all S3 buckets in the account.

Which additional steps should the Solutions Architect take to troubleshoot the issue? (Select TWO.)

-

Check if an appropriate IAM role is attached to the IAM user.

-

Check the SCPs set at the organizational units (OUs).

(Correct)

-

Check the bucket policies for all S3 buckets.

-

Check the ACLs for all S3 buckets.

-

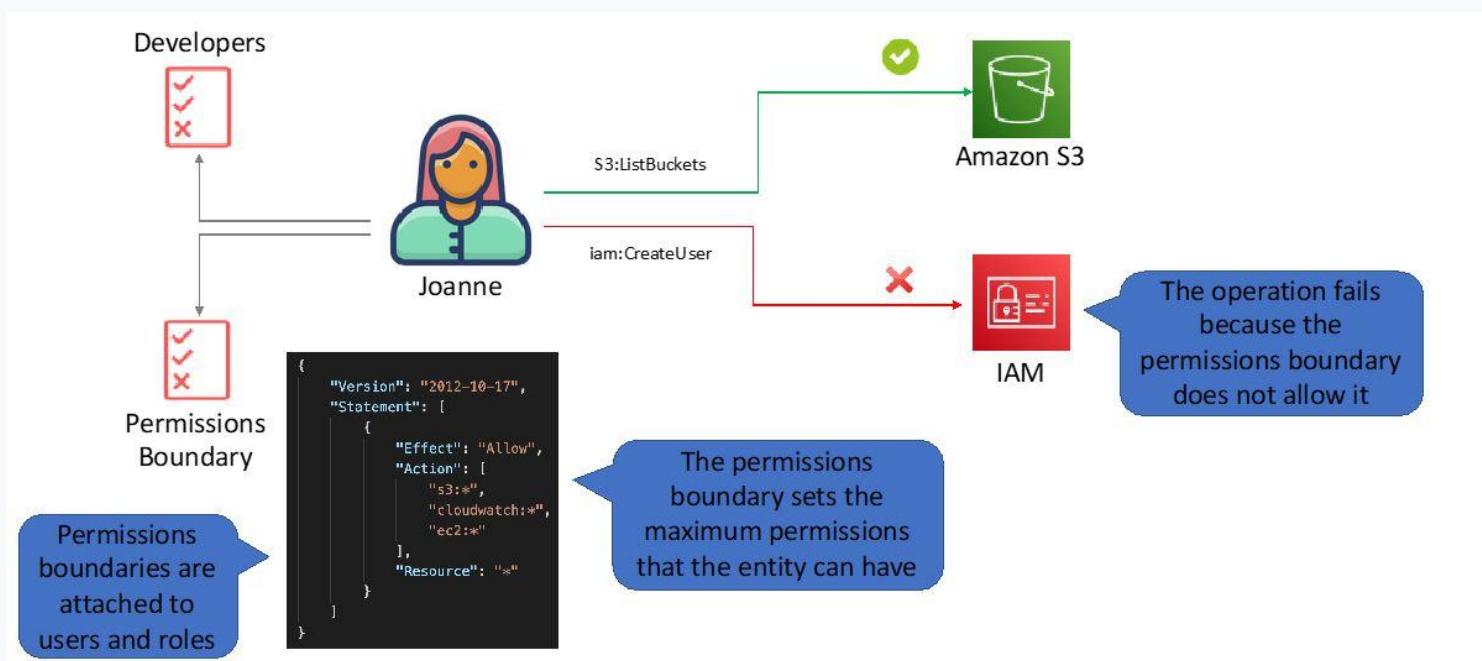
Check for the permissions boundaries set for the IAM user.

(Correct)

Explanation

A service control policy (SCP) may have been implemented that limits the API actions that are available for Amazon S3. This will apply to all users in the account regardless of the permissions they have assigned to their user account.

Another potential cause of the issue is that the permissions boundary for the user limits the S3 API actions available to the user. A permissions boundary is an advanced feature for using a managed policy to set the maximum permissions that an identity-based policy can grant to an IAM entity. An entity's permissions boundary allows it to perform only the actions that are allowed by both its identity-based policies and its permissions boundaries.



CORRECT: "Check the SCPs set at the organizational units (OUs)" is a correct answer.

CORRECT: "Check for the permissions boundaries set for the IAM user" is also a correct answer.

INCORRECT: "Check if an appropriate IAM role is attached to the IAM user" is incorrect. The question states that the user is logged in with a user account so is not assuming a role.

INCORRECT: "Check the bucket policies for all S3 buckets" is incorrect. The user has not been granted access to any buckets, and the error does not list access denied to any specific bucket. Therefore, it is more likely that the user is not been granted the API action to list the buckets.

INCORRECT: "Check the ACLs for all S3 buckets" is incorrect. With a bucket ACL the grantee is an AWS account or one of the predefined groups. With an ACL you can grant read/write at the bucket level but list is restricted to the object level so would not apply to the bucket itself. The user has been unable to list any buckets in this case so an ACL is unlikely to be the cause.

References:

https://docs.aws.amazon.com/IAM/latest/UserGuide/access_policies_boundaries.html

https://docs.aws.amazon.com/organizations/latest/userguide/orgs_manage_policies_scps.html

Save time with our AWS cheat sheets:

<https://digitalcloud.training/category/aws-cheat-sheets/aws-solutions-architect-professional/security-identity-compliance-sap/>

<https://digitalcloud.training/category/aws-cheat-sheets/aws-solutions-architect-professional/aws-management-tools-sap/>

Question 3: **Correct**

A company provides a service that allows users to upload high-resolution product images using an app on their phones for a price matching service. The service currently uses Amazon S3 in the us-west-1 Region. The company has expanded to Europe and users in European countries are experiencing significant delays when uploading images.

Which combination of changes can a Solutions Architect make to improve the upload times for the images? (Select TWO.)

- **Configure the S3 bucket to use S3 Transfer Acceleration.**
(Correct)
- **Redeploy the application to use Amazon S3 multipart upload.**
(Correct)
- **Modify the Amazon S3 bucket to use Intelligent Tiering.**
- **Configure the client application to use byte-range fetches.**
- **Create an Amazon CloudFront distribution with the S3 bucket as an origin.**

Explanation

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

Transfer Acceleration is a good solution for the following use cases:

- You have customers that upload to a centralized bucket from all over the world.
- You transfer gigabytes to terabytes of data on a regular basis across continents.
- You are unable to utilize all of your available bandwidth over the Internet when uploading to Amazon S3.

Multipart upload transfers parts of the file in parallel and can speed up performance. This should definitely be built into the application code. Multipart upload also handles the failure of any parts gracefully, allowing for those parts to be retransmitted.

Transfer Acceleration in combination with multipart upload will offer significant speed improvements when uploading data.

CORRECT: "Configure the S3 bucket to use S3 Transfer Acceleration" is the correct answer.

CORRECT: "Redeploy the application to use Amazon S3 multipart upload" is correct.

INCORRECT: "Create an Amazon CloudFront distribution with the S3 bucket as an origin" is incorrect. CloudFront can offer performance improvements for downloading data but to improve upload transfer times, Transfer Acceleration should be used.

INCORRECT: "Configure the client application to use byte-range fetches" is incorrect. This is a technique that is used when reading (not writing) data to fetch only the parts of the file that are required.

References:

<https://docs.aws.amazon.com/AmazonS3/latest/dev/optimizing-performance.html>

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

<https://docs.aws.amazon.com/AmazonS3/latest/dev/mpuoverview.html>

Save time with our AWS cheat sheets:

<https://digitalcloud.training/category/aws-cheat-sheets/aws-solutions-architect-professional/aws-migration-sap/>

Question 4: **Correct**

A company plans to build a gaming application in the AWS Cloud that will be used by Internet-based users. The application will run on a single instance and connections from users will be made over the UDP protocol. The company has requested that the service is implemented with a high level of security. A Solutions Architect has been asked to design a solution for the application on AWS.

Which combination of steps should the Solutions Architect take to meet these requirements? (Select THREE.)

-

Use AWS Global Accelerator with an Elastic Load Balancer as an endpoint.

-

Enable AWS Shield Advanced on all public-facing resources.

(Correct)

-

Configure a network ACL rule to block all non-UDP traffic. Associate the network ACL with the subnets that hold the load balancer instances.

(Correct)

-

Use an Application Load Balancer (ALB) in front of the application instance. Use a friendly DNS entry in Amazon Route 53 pointing to the ALB's internet-facing fully qualified domain name (FQDN).

-

Use a Network Load Balancer (NLB) in front of the application instance. Use a friendly DNS entry in Amazon Route 53 pointing to the NLB's Elastic IP address.

(Correct)



Define an AWS WAF rule to explicitly drop non-UDP traffic and associate the rule with the load balancer.

Explanation

The Network Load Balancer (NLB) supports the UDP protocol and can be placed in front of the application instance. This configuration may add some security if the instance is running in a private subnet.

An NLB can be configured with an Elastic IP in each subnet in which it has nodes. In this case it only has a single subnet (one instance) and so there will be 1 EIP.

Route 53 can be configured to resolve directly to the EIP rather than the DNS name of the NLB as there is only one IP address to return. To filter traffic the network ACL for the subnet can be configured to block all non-UDP traffic.

This solution meets all the stated requirements.

CORRECT: "Use a Network Load Balancer (NLB) in front of the application instance. Use a friendly DNS entry in Amazon Route 53 pointing to the NLB's Elastic IP address" is a correct answer.

CORRECT: "Configure a network ACL rule to block all non-UDP traffic. Associate the network ACL with the subnets that hold the load balancer instances" is also a correct answer.

CORRECT: "Enable AWS Shield Advanced on all public-facing resources" is also a correct answer.

INCORRECT: "Use an Application Load Balancer (ALB) in front of the application instance. Use a friendly DNS entry in Amazon Route 53 pointing to the ALB's internet-facing fully qualified domain name (FQDN)" is incorrect. An ALB only listens for HTTP and HTTPS traffic which uses the TCP protocol. It does not support UDP.

INCORRECT: "Define an AWS WAF rule to explicitly drop non-UDP traffic and associate the rule with the load balancer" is incorrect. WAF is unnecessary as a network ACL can filter the traffic.

INCORRECT: "Use AWS Global Accelerator with an Elastic Load Balancer as an endpoint" is incorrect. AWS Global Accelerator provides improved performance and high availability when you have copies of your application running in multiple AWS Regions. It is not required in this solution.

References:

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

Save time with our AWS cheat sheets:

<https://digitalcloud.training/category/aws-cheat-sheets/aws-solutions-architect-professional/aws-compute-sap/>

<https://digitalcloud.training/category/aws-cheat-sheets/aws-solutions-architect-professional/security-identity-compliance-sap/>

Question 5: Correct

A company is deploying a web service that will provide read and write access to structured data. The company expects there to be variable usage patterns with some short but significant spikes. The service must dynamically scale and must be fault tolerant across multiple AWS Regions.

Which actions should a Solutions Architect take to meet these requirements?

-

Store the data in Amazon Aurora global databases. Add Auto Scaling replicas to both Regions. Run the web service on Amazon EC2 instances in an Auto Scaling group behind an Application Load Balancer in each Region. In Amazon Route 53, configure an alias record and a multi-value routing policy.

-

Store the data in Amazon S3 buckets in two Regions and configure cross-Region replication. Create an Amazon CloudFront distribution that points to multiple origins. Use Amazon API Gateway and AWS Lambda for the web frontend and configure Amazon Route 53 with an alias record pointing to the REST API.



Store the data in an Amazon DynamoDB global table in two Regions using on-demand capacity mode. Run the web service in both Regions as Amazon ECS Fargate tasks in an Auto Scaling ECS service behind an Application Load Balancer (ALB). In Amazon Route 53, configure an alias record and a latency-based routing policy with health checks to distribute traffic between the two ALBs.

(Correct)



Store the data in Amazon DocumentDB in two Regions. Use AWS DMS to synchronize data between databases. Run the web service on Amazon EC2 instances in an Auto Scaling group behind an Application Load Balancer in each Region. In Amazon Route 53, configure an alias record and a failover routing policy.

Explanation

The DynamoDB global tables solution is the only database solution that will allow writes in both AWS Regions. Using Amazon ECS Fargate tasks with Auto Scaling will ensure the compute layer scales appropriately and an ALB will distribute connections across multiple tasks.

This solution is active-active and so a latency-based routing policy will direct users to the closest Region to improve performance. Health checks are enabled which means that if a Region outage occurs, traffic will be directed to the second Region.

CORRECT: "Store the data in an Amazon DynamoDB global table in two Regions using on-demand capacity mode. Run the web service in both Regions as Amazon ECS Fargate tasks in an Auto Scaling ECS service behind an Application Load Balancer (ALB). In Amazon Route 53, configure an alias record and a latency-based routing policy with health checks to distribute traffic between the two ALBs" is correct.

INCORRECT: "Store the data in Amazon Aurora global databases. Add Auto Scaling replicas to both Regions. Run the web service on Amazon EC2 instances in an Auto Scaling group behind an Application Load Balancer in each Region. In Amazon Route 53, configure an alias record and a multi-value routing policy" is incorrect. Aurora global databases provide read access in multiple Regions but

writes can only be made in one Region. In this solution the multi-value routing policy will direct connections to healthy ALBs in both Regions and so any attempts to write in a Region without that capability will fail.

INCORRECT: "Store the data in Amazon DocumentDB in two Regions. Use AWS DMS to synchronize data between databases. Run the web service on Amazon EC2 instances in an Auto Scaling group behind an Application Load Balancer in each Region. In Amazon Route 53, configure an alias record and a failover routing policy" is incorrect. DocumentDB is used for storing JSON data not structured data so is not suitable.

INCORRECT: "Store the data in Amazon S3 buckets in two Regions and configure cross-Region replication. Create an Amazon CloudFront distribution that points to multiple origins. Use Amazon API Gateway and AWS Lambda for the web frontend and configure Amazon Route 53 with an alias record pointing to the REST API" is incorrect. When using CloudFront with multiple origins, behaviors must be configured to direct traffic to a specific origin, this would not work when the content is the same – it does not function as a load balancer or provide automatic failover.

References:

<https://docs.aws.amazon.com/amazondynamodb/latest/developerguide/GlobalTables.html>

Save time with our AWS cheat sheets:

<https://digitalcloud.training/category/aws-cheat-sheets/aws-solutions-architect-professional/security-identity-compliance-sap/>

<https://digitalcloud.training/category/aws-cheat-sheets/aws-solutions-architect-professional/aws-database-sap/>

Question 6: **Correct**

A company uses Amazon RedShift for analytics. Several teams deploy and manage their own RedShift clusters and management has requested that the costs for these clusters is better managed. The management team has set budgets and once the budgetary thresholds have been reached a notification should be sent to a distribution list for managers. Teams should be able to view their RedShift cluster's expenses to date. A Solutions Architect needs to create a solution that ensures the policy is centrally enforced in a multi-account environment.

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

-

Create an AWS Service Catalog portfolio for each team. Add each team's Amazon RedShift cluster as an AWS CloudFormation template to their Service Catalog portfolio as a Product.

(Correct)

-

Create an Amazon CloudWatch metric for billing. Create a custom alert when costs exceed the budgetary threshold.

-

Update the AWS CloudFormation template to include the AWS::Budgets::Budget::resource with the NotificationsWithSubscribers property.

(Correct)

-

Install the unified CloudWatch Agent on the RedShift cluster hosts. Track the billing metric data in CloudWatch and trigger an alarm when a threshold is reached.

-

Create an AWS CloudTrail trail that tracks data events. Configure Amazon CloudWatch to monitor the trail and trigger an alarm when billing metrics exceed a certain threshold.

Explanation

You can use AWS Budgets to track your service costs and usage within AWS Service Catalog. You can associate budgets with AWS Service Catalog products and portfolios.

AWS Budgets gives you the ability to set custom budgets that alert you when your costs or usage exceed (or are forecasted to exceed) your budgeted amount.

If a budget is associated to a product, you can view information about the budget on the **Products** and **Product details** page. If a budget is associated to a portfolio, you can view information about the budget on the **Portfolios** and **Portfolio details** page.

When you click on a product or portfolio, you are taken to a detail page. These **Portfolio detail** and **Product detail** pages have a section with detailed information about the associated budget. You can see the budgeted amount, current spend, and forecasted spend. You also have the option to view budget details and edit the budget.

CORRECT: "Update the AWS CloudFormation template to include the AWS::Budgets::Budget::resource with the NotificationsWithSubscribers property" is a correct answer.

CORRECT: "Create an AWS Service Catalog portfolio for each team. Add each team's Amazon RedShift cluster as an AWS CloudFormation template to their Service Catalog portfolio as a Product" is also a correct answer.

INCORRECT: "Install the unified CloudWatch Agent on the RedShift cluster hosts. Track the billing metric data in CloudWatch and trigger an alarm when a threshold is reached" is incorrect. This agent is used on EC2 instances for sending additional metric data and logs to CloudWatch. However, it is not used for budgeting.

INCORRECT: "Create an AWS CloudTrail trail that tracks data events. Configure Amazon CloudWatch to monitor the trail and trigger an alarm when billing metrics exceed a certain threshold" is incorrect. CloudTrail tracks API calls, it cannot be used for tracking billing data.

INCORRECT: "Create an Amazon CloudWatch metric for billing. Create a custom alert when costs exceed the budgetary threshold" is incorrect. Billing data is automatically collected, you cannot create a metric for billing but you can create an alarm.

References:

https://docs.aws.amazon.com/servicecatalog/latest/adminguide/catalogs_budgets.html

<https://docs.aws.amazon.com/AWSCloudFormation/latest/UserGuide/aws-resource-budgets-budget.html>

Save time with our AWS cheat sheets:

<https://digitalcloud.training/category/aws-cheat-sheets/aws-solutions-architect-professional/aws-management-tools-sap/>

Question 7: **Correct**

A company has an NFS file server on-premises with 50 TB of data that is being migrated to Amazon S3. The data is made up of many millions of small files and a Snowball Edge device is being used for the migration. A shell script is being used to copy data using the file interface of the Snowball Edge device. Data transfer times are very slow and the Solutions Architect suspects this may be related to the overhead of encrypting all the small files and copying them over the network.

What change should be made to improve data transfer times?



Modify the shell script to ensure that individual files are being copied rather than directories.



Cluster two Snowball Edge devices together to increase the throughput of the devices.



Perform multiple copy operations at one time by running each command from a separate terminal window, in separate instances of the Snowball client.

(Correct)



Connect directly to the USB interface on the Snowball Edge device and copy the files locally.

Explanation

In general, you can improve the transfer speed from your data source to the Snowball in the following ways, ordered from largest to smallest positive impact on performance:

1. Use the latest Mac or Linux Snowball client
2. Batch small files together
3. Perform multiple copy operations at one time
4. Copy from multiple workstations
5. Transfer directories, not files

Option 3 entails performing multiple snowball cp commands at one time. You can do this by running each command from a separate terminal window, in separate instances of the Snowball client, all connected to the same Snowball.

CORRECT: "Perform multiple copy operations at one time by running each command from a separate terminal window, in separate instances of the Snowball client" is the correct answer.

INCORRECT: "Modify the shell script to ensure that individual files are being copied rather than directories" is incorrect. This is the opposite of what you should do to improve performance. Copy directories rather than individual files

INCORRECT: "Connect directly to the USB interface on the Snowball Edge device and copy the files locally" is incorrect. The Snowball Edge device does not come with a USB interface. It has SFP, QSFP, and RJ45 connections for Ethernet networking.

INCORRECT: "Cluster two Snowball Edge devices together to increase the throughput of the devices" is incorrect. You cannot cluster these devices for throughput.

References:

<https://docs.aws.amazon.com/snowball/latest/ug/performance.html>

<https://docs.aws.amazon.com/snowball/latest/developer-guide/specifications.html>

Save time with our AWS cheat sheets:

<https://digitalcloud.training/category/aws-cheat-sheets/aws-solutions-architect-professional/aws-migration-sap/>

Question 8: Incorrect

A company has deployed two Microsoft Active Directory Domain Controllers into an Amazon VPC with a default configuration. The DHCP options set associated with the VPC has been configured to assign the IP addresses of the Domain Controllers as DNS servers. A VPC interface endpoint has been created but EC2 instances within the VPC are unable to resolve the private endpoint addresses.

Which strategies could a Solutions Architect use to resolve the issue? (Select TWO.)

-

Define an inbound Amazon Route 53 Resolver. Set a conditional forwarding rule for the Active Directory domain to the Active Directory servers. Configure the DNS settings in the VPC DHCP options set to use the AmazonProvidedDNS servers.

-

Configure the DNS service on the EC2 instances in the VPC to use the VPC resolver server as the secondary DNS server.

(Incorrect)

-

Define an outbound Amazon Route 53 Resolver. Set a conditional forwarding rule for the Active Directory domain to the Active Directory servers. Configure the DNS settings in the VPC DHCP options set to use the AmazonProvidedDNS servers.

(Correct)

-

Update the DNS service on the Active Directory servers to forward all queries to the VPC Resolver.

-

Update the DNS service on the Active Directory servers to forward all non-authoritative queries to the VPC Resolver.

(Correct)

Explanation

The EC2 instances are unable to resolve the DNS name of the VPC interface endpoint to an IP address as they are configured to use the Domain Controllers for DNS and the DCs do not have a record for the VPC interface endpoint.

There are two solutions to this problem that both achieve the same outcome. The first involves modifying the DNS service on the DCs to forward non-authoritative queries to the VPC resolver. This simply means if the DNS service on the DC does not have the record in its zone file it will forward the query to another DNS service.

The second solution uses an outbound Route 53 resolver. With outbound resolvers (but not with inbound resolvers) you can configure forwarding rules. In this case you would need to modify the EC2 instances (via the DHCP options set) to use the Amazon provided DNS servers. These servers would be able to resolve the VPC interface endpoint. The forwarding rule will forward any traffic for the Domain Controllers to those servers.

CORRECT: "Update the DNS service on the Active Directory servers to forward all non-authoritative queries to the VPC Resolver" is a correct answer.

CORRECT: "Define an outbound Amazon Route 53 Resolver. Set a conditional forwarding rule for the Active Directory domain to the Active Directory servers. Configure the DNS settings in the VPC DHCP options set to use the AmazonProvidedDNS servers" is also a correct answer.

INCORRECT: "Define an inbound Amazon Route 53 Resolver. Set a conditional forwarding rule for the Active Directory domain to the Active Directory servers. Configure the DNS settings in the VPC DHCP options set to use the AmazonProvidedDNS servers" is incorrect. You cannot configure a forwarding rule on an inbound resolver.

INCORRECT: "Configure the DNS service on the EC2 instances in the VPC to use the VPC resolver server as the secondary DNS server" is incorrect. A secondary DNS server is used as a backup for the primary DNS server that is configured on a client. It is not meant for splitting queries.

INCORRECT: "Update the DNS service on the Active Directory servers to forward all queries to the VPC Resolver" is incorrect. This would mean the DNS server on the DCs does not respond to queries from its own zone file as it will forward ALL queries to the VPC resolver.

References:

<https://docs.aws.amazon.com/Route53/latest/DeveloperGuide/resolver-rules-managing.html>

Save time with our AWS cheat sheets:

<https://digitalcloud.training/category/aws-cheat-sheets/aws-solutions-architect-professional/aws-networking-content-delivery-sap/>

Question 9: Correct

A company is using AWS CloudFormation templates for infrastructure provisioning. The templates are hosted in the company's private GitHub repository. The company has experienced several issues with updates to the templates that have caused errors when executing the updates and creating the environment. A Solutions Architect must resolve these issues and implement automated testing of the CloudFormation template updates.

How can the Solutions Architect accomplish these requirements?

-

Use AWS CodePipeline to create and execute a change set when updates are made to the CloudFormation templates in GitHub. Include a CodePipeline action to test the deployment with testing scripts run using AWS CodeDeploy. Upon successful testing, configure CodePipeline to execute the change set and deploy to production.

-

Use AWS Lambda to synchronize the contents of the GitHub repository to AWS CodeCommit. Use AWS CodeBuild to create and execute a change set from the templates in GitHub. Configure CodeBuild to test the deployment with testing scripts.



Use AWS CodePipeline to a create a change set when updates are made to the CloudFormation templates in GitHub. Include a CodePipeline action to test the deployment with testing scripts run using AWS CodeBuild. Upon successful testing, configure CodePipeline to execute the change set and deploy to production.

(Correct)



Use AWS Lambda to synchronize the contents of the GitHub repository to AWS CodeCommit. Use AWS CodeDeploy to create and execute a change set. Configure CodeDeploy to test the environment using testing scripts run by AWS CodeBuild.

Explanation

You can apply continuous delivery practices to your AWS CloudFormation stacks using AWS CodePipeline. AWS CodePipeline is a continuous delivery service for fast and reliable application and infrastructure updates. CodePipeline builds, tests, and deploys your code every time there is a code change, based on the release process models you define.

When using CloudFormation change sets you first create the change set which allows you to preview how proposed changes to a stack might impact your running resources. This creates a separate stack that can be used to view the changes and ensure the updates apply successfully. Then, once you're happy with the changes the change set can be executed which will update the stack.

You can use AWS CodeBuild to both build and test code. CodeBuild can be configured with custom scripts to run tests and the result of the tests can determine the subsequent actions such as proceeding to deployment.

CORRECT: "Use AWS CodePipeline to a create a change set when updates are made to the CloudFormation templates in GitHub. Include a CodePipeline action to test the deployment with testing scripts run using AWS CodeBuild. Upon successful testing, configure CodePipeline to execute the change set and deploy to production" is the correct answer.

INCORRECT: "Use AWS CodePipeline to a create and execute a change set when updates are made to the CloudFormation templates in GitHub. Include a CodePipeline action to test the deployment with testing scripts run using AWS CodeDeploy. Upon successful

testing, configure CodePipeline to execute the change set and deploy to production" is incorrect. CodeDeploy is used for deploying the updates but it is not used for running testing scripts.

INCORRECT: "Use AWS Lambda to synchronize the contents of the GitHub repository to AWS CodeCommit. Use AWS CodeDeploy to create and execute a change set. Configure CodeDeploy to test the environment using testing scripts run by AWS CodeBuild" is incorrect. CodePipeline should be used for creating and executing the change set and it can use a private GitHub repository so the first step is unnecessary. Testing should also successfully complete before execution of the change set.

INCORRECT: "Use AWS Lambda to synchronize the contents of the GitHub repository to AWS CodeCommit. Use AWS CodeBuild to create and execute a change set from the templates in GitHub. Configure CodeBuild to test the deployment with testing scripts" is incorrect. CodePipeline should be used for creating and executing the change set and it can use a private GitHub repository so the first step is unnecessary. Testing should also successfully complete before execution of the change set.

References:

<https://aws.amazon.com/about-aws/whats-new/2016/11/continuously-deliver-changes-to-aws-cloudformation-stacks-with-aws-codepipeline/>

<https://docs.aws.amazon.com/codebuild/latest/userguide/how-to-create-pipeline.html>

Save time with our AWS cheat sheets:

<https://digitalcloud.training/category/aws-cheat-sheets/aws-solutions-architect-professional/aws-developer-tools-sap/>

Question 10: **Correct**

A company has deployed a new application into an Amazon VPC that does not have Internet access. The company has connected an AWS Direct Connection (DX) private VIF to the VPC and all communications will be over the DX connection. A new requirement states that all data in transit must be encrypted between users and the VPC.

Which strategy should a Solutions Architect use to maintain consistent network performance while meeting this new requirement?

-
-
-

Create a client VPN endpoint and configure the users' computers to use an AWS client VPN to connect to the VPC over the Internet.



Create a new Site-to-Site VPN that connects to the VPC over the internet.



Create a new public virtual interface for the existing DX connection, and create a new VPN that connects to the VPC over the DX public virtual interface.

(Correct)



Create a new private virtual interface for the existing DX connection, and create a new VPN that connects to the VPC over the DX private virtual interface.

Explanation

Running an AWS VPN connection over a DX connection provides consistent levels of throughput and encryption algorithms that protect your data. Though a private VIF is typically used to connect to a VPC, in the case of running an IPSec VPN over the top of a DX connection it is necessary to use a public VIF (please check the AWS article linked below for instructions)

CORRECT: "Create a new public virtual interface for the existing DX connection, and create a new VPN that connects to the VPC over the DX public virtual interface" is the correct answer.

INCORRECT: "Create a new private virtual interface for the existing DX connection, and create a new VPN that connects to the VPC over the DX private virtual interface" is incorrect. A public VIF must be used when using an IPSec VPN over a DX connection.

INCORRECT: "Create a client VPN endpoint and configure the users' computers to use an AWS client VPN to connect to the VPC over the Internet" is incorrect. This does not maintain consistent network performance as the public internet offers variable performance.

INCORRECT: "Create a new Site-to-Site VPN that connects to the VPC over the internet" is incorrect. This does not maintain consistent network performance as the public internet offers variable performance. The DX connection should be utilized.

References:

<https://aws.amazon.com/premiumsupport/knowledge-center/create-vpn-direct-connect/>

Save time with our AWS cheat sheets:

<https://digitalcloud.training/category/aws-cheat-sheets/aws-solutions-architect-professional/aws-networking-content-delivery-sap/>

Question 11: **Correct**

A company runs an application that generates user activity reports and stores them in an Amazon S3 bucket. Users are able to download the reports using the application which generates a signed URL. A user recently reported that the reports of other users can be accessed directly from the S3 bucket. A Solutions Architect reviewed the bucket permissions and discovered that public access is currently enabled.

How can the documents be protected from unauthorized access without modifying the application workflow?

-
- Use the Block Public Access feature in Amazon S3 to set the BlockPublicPolicy option to TRUE on the bucket.**
-
- Modify the settings on the S3 bucket to enable default encryption for all objects.**
-
- Configure server access logging and monitor the log files to check for unauthorized access.**
-

Use the Block Public Access feature in Amazon S3 to set the IgnorePublicAcls option to TRUE on the bucket.

(Correct)

Explanation

The S3 bucket is allowing public access and this must be immediately disabled. Setting the IgnorePublicAcls option to TRUE causes Amazon S3 to ignore all public ACLs on a bucket and any objects that it contains.

The other settings you can configure with the Block Public Access Feature are:

- BlockPublicAcls – PUT bucket ACL and PUT objects requests are blocked if granting public access.
- BlockPublicPolicy – Rejects requests to PUT a bucket policy if granting public access.
- RestrictPublicBuckets – Restricts access to principles in the bucket owners' AWS account.

CORRECT: "Use the Block Public Access feature in Amazon S3 to set the IgnorePublicAcls option to TRUE on the bucket" is the correct answer.

INCORRECT: "Use the Block Public Access feature in Amazon S3 to set the BlockPublicPolicy option to TRUE on the bucket" is incorrect. This option will only reject requests to PUT a bucket policy that grants public access which is not relevant to the workflow in this scenario.

INCORRECT: "Configure server access logging and monitor the log files to check for unauthorized access" is incorrect. This will only identify unauthorized access; it does not block it.

INCORRECT: "Modify the settings on the S3 bucket to enable default encryption for all objects" is incorrect. Encryption will not prevent public access; it just encrypts the data at rest in the S3 bucket.

References:

<https://docs.aws.amazon.com/AmazonS3/latest/dev/access-control-block-public-access.html>

Save time with our AWS cheat sheets:

<https://digitalcloud.training/category/aws-cheat-sheets/aws-solutions-architect-professional/aws-storage-sap/>

Question 12: **Correct**

A company currently manages a fleet of Amazon EC2 instances running Windows and Linux in public and private subnets. The operations team currently connects over the Internet to manage the instances as there is no connection to the corporate network.

Security groups have been updated to allow the RDP and SSH protocols from any source IPv4 address. There have been reports of malicious attempts to access the resources as the company wishes to implement the most secure solution for managing the instances.

Which strategy should a Solutions Architect recommend?

-

Deploy the AWS Systems Manager Agent on the EC2 instances. Access the EC2 instances using Session Manager restricting access to users with permission to manage the instances.

(Correct)

-

Configure an IPSec Virtual Private Network (VPN) connecting the corporate network to the Amazon VPC. Update security groups to allow connections over SSH and RDP from the corporate network only.

-

Deploy a server on the corporate network that can be used for managing EC2 instances. Update the security groups to allow connections over SSH and RDP from the on-premises management server only.

-

Deploy a Linux bastion host with an Elastic IP address in the public subnet. Allow access to the bastion host from 0.0.0.0/0.

Explanation

The most secure option presented is to use AWS Systems Manager Session Manager. Session Manager is a fully managed AWS Systems Manager capability that lets you manage Amazon EC2 instances, on-premises instances, and virtual machines (VMs) through an interactive one-click browser-based shell or through the AWS Command Line Interface (AWS CLI).

Session Manager provides secure and auditable instance management without the need to open inbound ports, maintain bastion hosts, or manage SSH keys. Session Manager also makes it easy to comply with corporate policies that require controlled access to instances, strict security practices, and fully auditable logs with instance access details, while still providing end users with simple one-click cross-platform access to your managed instances.

CORRECT: "Deploy the AWS Systems Manager Agent on the EC2 instances. Access the EC2 instances using Session Manager restricting access to users with permission to manage the instances" is the correct answer.

INCORRECT: "Deploy a server on the corporate network that can be used for managing EC2 instances. Update the security groups to allow connections over SSH and RDP from the on-premises management server only" is incorrect. This is less secure compared to using session manager as the SSH and RDP ports must still be open on instances and it does not offer the robust controls offered by session manager.

INCORRECT: "Deploy a Linux bastion host with an Elastic IP address in the public subnet. Allow access to the bastion host from 0.0.0.0/0" is incorrect. This is less secure compared to using session manager as the SSH and RDP ports must still be open on instances and it does not offer the robust controls offered by session manager. This solution could be better secured by restricting access to the corporate IP ranges.

INCORRECT: "Configure an IPSec Virtual Private Network (VPN) connecting the corporate network to the Amazon VPC. Update security groups to allow connections over SSH and RDP from the corporate network only" is incorrect. This is less secure compared to using session manager as the SSH and RDP ports must still be open on instances and it does not offer the robust controls offered by session manager.

References:

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

Save time with our AWS cheat sheets:

<https://digitalcloud.training/category/aws-cheat-sheets/aws-solutions-architect-professional/aws-management-tools-sap/>

Question 13: **Correct**

A solutions architect developed a web application that includes an AWS Lambda function that queries an Amazon Aurora MySQL database. The database is configured with three read replicas. During periods of high demand, the application does not meet performance requirements. A solutions architect noticed that the application opens many database connections, and this causes latency in the application

Which actions should the solutions architect take to improve the performance? (Select TWO.)

- Configure the application to use the cluster endpoint of the Aurora database.
- Create a Gateway Load Balancer to distribute connections across the three Aurora Read Replicas.
- Connect an RDS Proxy connection pool to the reader endpoint of the Aurora database.
- (Correct) Configure an Amazon Aurora serverless database cluster and use automatic scaling.
-

Move Lambda function code for opening the database connection outside of the event handler.

(Correct)

Explanation

You can create and connect to read-only endpoints called *reader endpoints* when you use RDS Proxy with Aurora clusters. These reader endpoints help to improve the read scalability of your query-intensive applications. Reader endpoints also help to improve the availability of your connections if a reader DB instance in your cluster becomes unavailable.

It is a best practice to move the database connection outside the event handler so subsequent invocations of the Lambda function can reuse it.

CORRECT: "Connect an RDS Proxy connection pool to the reader endpoint of the Aurora database" is a correct answer (as explained above.)

CORRECT: "Move Lambda function code for opening the database connection outside of the event handler" is also a correct answer (as explained above.)

INCORRECT: "Configure the application to use the cluster endpoint of the Aurora database" is incorrect. The reader endpoint should be used when configuring Amazon RDS Proxy.

INCORRECT: "Configure an Amazon Aurora serverless database cluster and use automatic scaling" is incorrect. This will not help with connection management and may cause more latency.

INCORRECT: "Create a Gateway Load Balancer to distribute connections across the three Aurora Read Replicas" is incorrect. A GLB is used for routing traffic through managed virtual appliances such as IDS/IPS or firewalls.

References:

<https://aws.amazon.com/rds/proxy/>

Save time with our AWS cheat sheets:

<https://digitalcloud.training/amazon-aurora/>

Question 14: **Correct**

A company recently noticed an increase in costs associated with Amazon EC2 instances and Amazon RDS databases. The company needs to be able to track the costs. The company uses AWS Organizations for all of their accounts. AWS CloudFormation is used for deploying infrastructure and all resources are tagged. The management team has requested that cost center numbers and project ID numbers are added to all future EC2 instances and RDS databases.

What is the MOST efficient strategy a Solutions Architect should follow to meet these requirements?

-

Use an AWS Config rule to check for untagged resources. Create a centralized AWS Lambda based solution to tag untagged EC2 instances and RDS databases every hour using a cross-account role.

-

Use Tag Editor to tag existing resources. Create cost allocation tags to define the cost center and project ID and allow 24 hours for tags to activate.

-

Create cost allocation tags to define the cost center and project ID and allow 24 hours for tags to activate. Use permissions boundaries to restrict the creation of resources that do not have the cost center and project ID tags specified.

-

Use Tag Editor to tag existing resources. Create cost allocation tags to define the cost center and project ID. Use SCPs to restrict the creation of resources that do not have the cost center and project ID tags specified.

(Correct)

Explanation

You can use tags to organize your resources, and cost allocation tags to track your AWS costs on a detailed level. After you activate cost allocation tags, AWS uses the cost allocation tags to organize your resource costs on your cost allocation report, to make it easier for you to categorize and track your AWS costs.

By adding tags to all new resources, the management team will be better able to track costs and allocate costs to specific cost centers and projects.

Service Control Policies (SCPs) can be used to limit the maximum available permissions in an account in AWS Organizations. SCPs are policies and conditional statements can be added. In this case an SCP can be created with a conditional statement that only allows resources to be created if they have a tag specified.

CORRECT: "Use Tag Editor to tag existing resources. Create cost allocation tags to define the cost center and project ID. Use SCPs to restrict the creation of resources that do not have the cost center and project ID tags specified" is the correct answer.

INCORRECT: "Use an AWS Config rule to check for untagged resources. Create a centralized AWS Lambda based solution to tag untagged EC2 instances and RDS databases every hour using a cross-account role" is incorrect. AWS Config can be used for compliance but a better solution would be to enforce tags at creation time. Using Lambda to tag the resources would be complex in terms of identifying which tags to add to which resources.

INCORRECT: "Use Tag Editor to tag existing resources. Create cost allocation tags to define the cost center and project ID and allow 24 hours for tags to activate" is incorrect. There is no mechanism here to enforce application of tags.

INCORRECT: "Create cost allocation tags to define the cost center and project ID and allow 24 hours for tags to activate. Use permissions boundaries to restrict the creation of resources that do not have the cost center and project ID tags specified" is incorrect. Permissions boundaries apply to user accounts but SCPs apply to entire AWS accounts and will be easier to enforce for all users.

References:

<https://docs.aws.amazon.com/awsaccountbilling/latest/aboutv2/cost-alloc-tags.html>

https://docs.aws.amazon.com/organizations/latest/userguide/orgs_manage_policies_scps.html

Save time with our AWS cheat sheets:

<https://digitalcloud.training/category/aws-cheat-sheets/aws-solutions-architect-professional/aws-management-tools-sap/>

Question 15: **Correct**

A Solution Architect used the AWS Application Discovery Service to gather information about some on-premises database servers. The tool discovered an Oracle data warehouse and several MySQL databases. The company plans to migrate to AWS and the Solutions Architect must determine the best migration pattern for each database.

Which combination of migration patterns will reduce licensing costs and operational overhead? (Select TWO.)

-

Migrate the Oracle data warehouse to an Amazon ElastiCache for Redis cluster using AWS DMS.

-

Migrate the MySQL databases to Amazon RDS for MySQL using AWS DMS.

(Correct)

-

Lift and shift the Oracle data warehouse to Amazon EC2 using AWS Snowball.

-

Migrate the Oracle data warehouse to Amazon Redshift using AWS SCT and AWS DMS.

(Correct)

-

Lift and shift the MySQL databases to Amazon EC2 using AWS Snowball.

Explanation

In this scenario we must determine the best platform to run each database on and the best migration path to get there. Cost and operational overhead must be minimized.

The best solution for an Oracle data warehouse is to migrate it to Amazon RedShift which is a managed service that is designed to run data warehouses (relational DB for OLAP use cases). This will require the schema to be modified which means AWS SCT should be used, and AWS DMS can migrate the actual data.

For the MySQL databases these can be run on Amazon RDS for MySQL. This will provide a managed service and does not require modifications to the schema. Therefore, AWS DMS can be used without AWS SCT.

CORRECT: "Migrate the Oracle data warehouse to Amazon Redshift using AWS SCT and AWS DMS" is a correct answer.

CORRECT: "Migrate the MySQL databases to Amazon RDS for MySQL using AWS DMS" is also a correct answer.

INCORRECT: "Lift and shift the Oracle data warehouse to Amazon EC2 using AWS Snowball" is incorrect. There is no indication that bandwidth is an issue or the database is particularly large. Therefore, Snowball is not required. Also Amazon EC2 does not reduce operational overhead.

INCORRECT: "Lift and shift the MySQL databases to Amazon EC2 using AWS Snowball" is incorrect. Amazon EC2 does not reduce operational overhead so is not the best choice. As with the previous explanation, there's no indication that Snowball is required.

INCORRECT: "Migrate the Oracle data warehouse to an Amazon ElastiCache for Redis cluster using AWS DMS" is incorrect. ElastiCache is mainly used for caching data in-memory from other databases and is not the best choice for a data warehouse. Also, SCT would be needed to modify the schema.

References:

<https://aws.amazon.com/getting-started/hands-on/migrate-oracle-to-amazon-redshift/>

<https://docs.aws.amazon.com/prescriptive-guidance/latest/patterns/migrate-an-on-premises-mysql-database-to-amazon-rds-for-mysql.html>

Save time with our AWS cheat sheets:

<https://digitalcloud.training/category/aws-cheat-sheets/aws-solutions-architect-professional/aws-database-sap/>

<https://digitalcloud.training/category/aws-cheat-sheets/aws-solutions-architect-professional/aws-migration-sap/>

Question 16: **Correct**

A company is planning to build a high-performance computing (HPC) solution in the AWS Cloud. The solution will include a 10-node cluster running Linux. High speed and low latency inter-instance connectivity is required to optimize the performance of the cluster.

Which combination of steps will meet these requirements? (Choose two.)

-

Deploy Amazon EC2 instances in a cluster placement group.

(Correct)

-

Deploy instances across at least three Availability Zones.

-

Use Amazon EC2 instance types and AMIs that support EFA.

(Correct)

-

Use Amazon EC2 instances that support burstable performance.

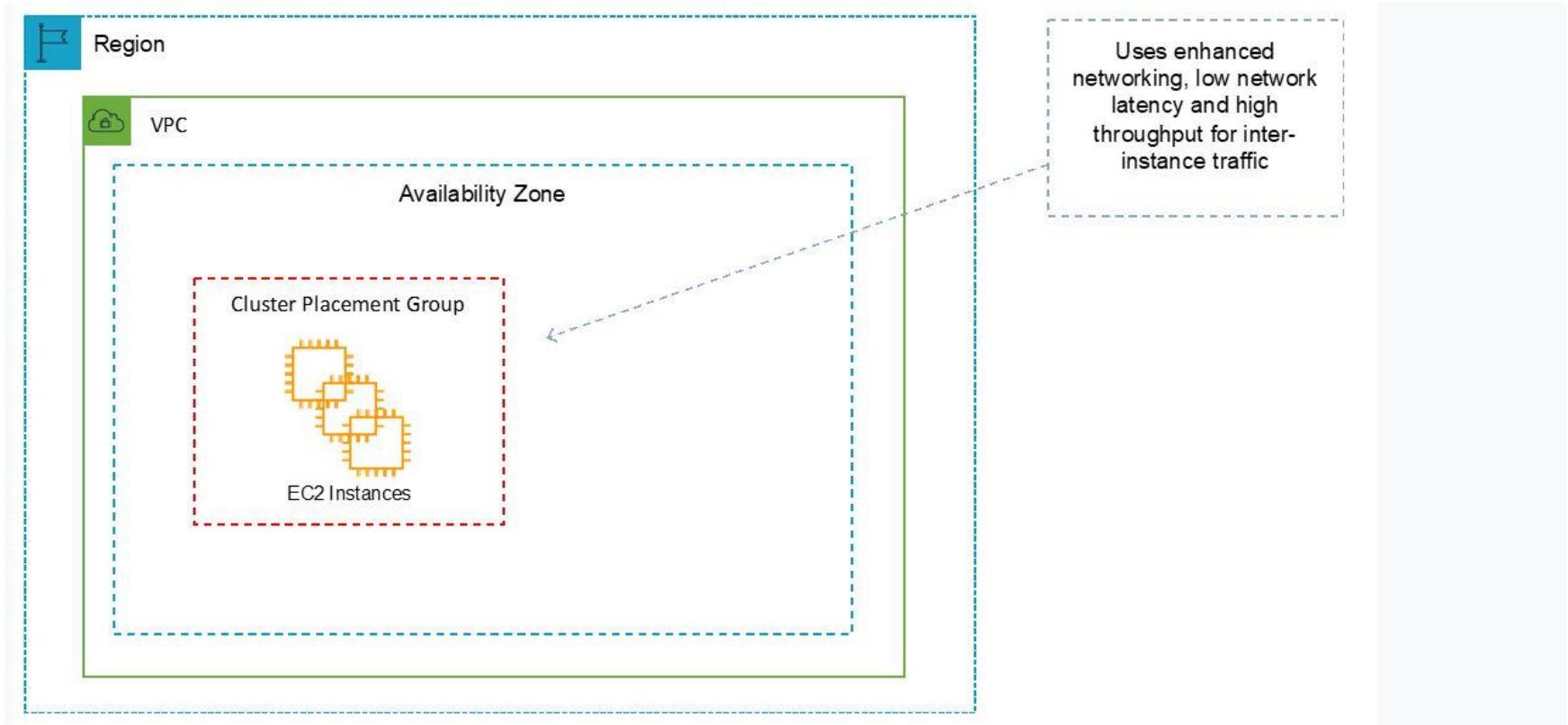


Deploy Amazon EC2 instances in a partition placement group.

Explanation

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

Cluster placement groups are recommended for applications that benefit from low network latency, high network throughput, or both. They are also recommended when the majority of the network traffic is between the instances in the group.



An Elastic Fabric Adapter (EFA) is a network device that you can attach to your Amazon EC2 instance to accelerate High Performance Computing (HPC) and machine learning applications. EFA enables you to achieve the application performance of an on-premises HPC cluster, with the scalability, flexibility, and elasticity provided by the AWS Cloud.

CORRECT: "Deploy Amazon EC2 instances in a cluster placement group" is a correct answer.

CORRECT: "Use Amazon EC2 instance types and AMIs that support EFA" is also a correct answer.

INCORRECT: "Deploy instances across at least three Availability Zones" is incorrect. This will increase latency between instances which is bad for HPC applications.

INCORRECT: "Deploy Amazon EC2 instances in a partition placement group" is incorrect. Partition placement groups are recommended for applications that are distributed and replicated. It does not provide the optimum latency and throughput required by an HPC application.

INCORRECT: "Use Amazon EC2 instances that support burstable performance" is incorrect. This feature provides bursts of CPU performance but in this scenario low latency and high throughput is required between instances and burstable CPUs do not assist with this.

References:

<https://docs.aws.amazon.com/AWSEC2/latest/UserGuide/placement-groups.html#placement-groups-cluster>

<https://docs.aws.amazon.com/AWSEC2/latest/UserGuide/efa.html>

Save time with our AWS cheat sheets:

<https://digitalcloud.training/category/aws-cheat-sheets/aws-solutions-architect-professional/aws-management-tools-sap/>

Question 17: **Correct**

A company has a large photo library stored on Amazon S3. They use AWS Lambda to extract metadata from the files according to various processing rules for different categories of photo. The output is then stored in an Amazon DynamoDB table.

The extraction process is performed whenever customer requests are submitted and can take up to 60 minutes to complete. The company wants to reduce the time taken to extract the metadata and has split the single Lambda function into separate Lambda functions for each category of photo.

Which additional steps should the Solutions Architect take to meet the requirements?

-

Create an AWS Step Functions workflow to run the Lambda functions in parallel. Create another Step Functions workflow that retrieves a list of files and executes a metadata extraction workflow for each one.

-

Create an AWS Step Functions workflow to run the Lambda functions in parallel. Create a Lambda function to retrieve a list of files and write each item to an Amazon SQS queue. Configure a Lambda function to retrieve messages from the SQS queue and call the StartExecution API.

(Correct)

-

Create an AWS Batch compute environment for each Lambda function. Configure an AWS Batch job queue for the compute environment. Create a Lambda function to retrieve a list of files and write each item to the job queue.

-

Create a Lambda function to retrieve a list of files and write each item to an Amazon SQS queue. Subscribe the metadata extraction Lambda functions to the SQS queue with a large batch size.

Explanation

The best solution presented is to use a combination of AWS Step Functions and Amazon SQS. This results in each Lambda function being able to run in parallel and use a queue for buffering the jobs. A Lambda function is required to retrieve messages from the queue. The information from the messages can then be used as input to the parallel functions that form the workflow using the StartExecution API.

CORRECT: "Create an AWS Step Functions workflow to run the Lambda functions in parallel. Create a Lambda function to retrieve a list of files and write each item to an Amazon SQS queue. Configure a Lambda function to retrieve messages from the SQS queue and call the StartExecution API" is the correct answer.

INCORRECT: "Create an AWS Step Functions workflow to run the Lambda functions in parallel. Create another Step Functions workflow that retrieves a list of files and executes a metadata extraction workflow for each one" is incorrect. There is no need for two Step Functions workflows, it is better to use one workflow and an SQS queue for storing the lists of files to be processed.

INCORRECT: "Create an AWS Batch compute environment for each Lambda function. Configure an AWS Batch job queue for the compute environment. Create a Lambda function to retrieve a list of files and write each item to the job queue" is incorrect. This would be a more costly way to process the files as Batch uses EC2 resources. It would be more cost-effective to use Lambda functions in a Step Functions workflow.

INCORRECT: "Create a Lambda function to retrieve a list of files and write each item to an Amazon SQS queue. Subscribe the metadata extraction Lambda functions to the SQS queue with a large batch size" is incorrect. The workflow is missing to coordinate the function execution. Using Step Functions can organize the execution of the metadata extraction process.

References:

<https://aws.amazon.com/step-functions/features/>

Save time with our AWS cheat sheets:

<https://digitalcloud.training/category/aws-cheat-sheets/aws-solutions-architect-professional/aws-application-integration-sap/>

Question 18: **Correct**

An application runs on an Amazon EC2 instance with an attached Amazon EBS Provisioned IOPS (PIOPS) volume. The volume is configured at 200-GB in size and has 3,000 IOPS provisioned. The application requires low latency and random access to the data. A Solutions Architect has been asked to consider options for lowering the cost of the storage without impacting performance and durability.

What should the Solutions Architect recommend?

-

Create an Amazon EFS file system with the performance mode set to Max I/O. Mount the EFS file system to the EC2 operating system.

-

Create an Amazon EFS file system with the throughput mode set to Provisioned. Mount the EFS file system to the EC2 operating system.



Change the PIOPS volume for a 1-TB Throughput Optimized HDD (st1) volume.



Change the PIOPS volume for a 1-TB EBS General Purpose SSD (gp2) volume.

(Correct)

Explanation

The most cost-effective solution is to use an Amazon EBS General Purpose SSD (gp2) volume. The volume should be configured with 1-TB as gp2 volumes provide 3 IOPS per GB, which will allow the full 3,000 IOPS to be achieved.

CORRECT: "Change the PIOPS volume for a 1-TB EBS General Purpose SSD (gp2) volume" is the correct answer.

INCORRECT: "Change the PIOPS volume for a 1-TB Throughput Optimized HDD (st1) volume" is incorrect. This volume type supports a maximum of 500 IOPS per volume.

INCORRECT: "Create an Amazon EFS file system with the performance mode set to Max I/O. Mount the EFS file system to the EC2 operating system" is incorrect. EFS will be much more expensive than using a gp2 volume.

INCORRECT: "Create an Amazon EFS file system with the throughput mode set to Provisioned. Mount the EFS file system to the EC2 operating system" is incorrect. EFS will be much more expensive than using a gp2 volume.

References:

<https://docs.aws.amazon.com/AWSEC2/latest/UserGuide/ebs-volume-types.html>

Save time with our AWS cheat sheets:

<https://digitalcloud.training/category/aws-cheat-sheets/aws-solutions-architect-professional/aws-storage-sap/>

Question 19: **Correct**

A Solutions Architect is migrating an application to AWS Fargate. The task runs in a private subnet and does not have direct connectivity to the internet. When the Fargate task is launched, it fails with the following error:

CannotPullContainerError: API error (500): Get https://11112222333.dkr.ecr.us-east-1.amazonaws.com/v2/: net/http: request canceled while waiting for connection"

What should the Solutions Architect do to correct the error?

-

Specify DISABLED for Auto-assign public IP when launching the task and configure a NAT gateway in a public subnet to route requests to the internet.

(Correct)

-

Specify DISABLED for Auto-assign public IP when launching the task and configure a NAT gateway in a private subnet to route requests to the internet.

-

Specify ENABLED for Auto-assign public IP when launching the task.

-

Enable dual-stack in the Amazon ECS account settings and configure the network for the task to use awsvpc.

Explanation

When a Fargate task is launched, its elastic network interface requires a route to the internet to pull container images. If you receive an error similar to the following when launching a task, it is because a route to the internet does not exist:

CannotPullContainerError: API error (500): Get https://111122223333.dkr.ecr.us-east-1.amazonaws.com/v2/: net/http: request canceled while waiting for connection"

To resolve this issue, you can:

- For tasks in public subnets, specify **ENABLED** for **Auto-assign public IP** when launching the task.
- For tasks in private subnets, specify **DISABLED** for **Auto-assign public IP** when launching the task, and configure a NAT gateway in your VPC to route requests to the internet.

CORRECT: "Specify **DISABLED** for **Auto-assign public IP** when launching the task and configure a NAT gateway in a public subnet to route requests to the internet" is the correct answer.

INCORRECT: "Specify **DISABLED** for **Auto-assign public IP** when launching the task and configure a NAT gateway in a private subnet to route requests to the internet" is incorrect. The NAT Gateway should be in a public subnet.

INCORRECT: "Specify **ENABLED** for **Auto-assign public IP** when launching the task" is incorrect. This will not work as the task is running in a private subnet and will not pick up a public IP.

INCORRECT: "Enable dual-stack in the Amazon ECS account settings and configure the network for the task to use awsvpc" is incorrect. This is used to enable IPv6 for a task but that is not required in this situation.

References:

https://docs.aws.amazon.com/AmazonECS/latest/developerguide/task_CANNOT_PULL_IMAGE.html

Save time with our AWS cheat sheets:

<https://digitalcloud.training/category/aws-cheat-sheets/aws-solutions-architect-professional/aws-compute-sap/>

Question 20: **Correct**

A company has deployed an application on Amazon EC2 instances behind an internet-facing Application Load Balancer (ALB). The ALB is configured as the origin in an Amazon CloudFront distribution. The company requires that the solution is secured against web-based attacks. An AWS WAF web ACL has been created and associated with the CloudFront distribution. The company must prevent anyone from circumventing the CloudFront distribution and connecting directly to the ALB.

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



Create a new web ACL that blocks access to the application. Associate the new web ACL with the ALB.



Add a security group rule to the ALB to allow traffic from the AWS managed prefix list for CloudFront only.

(Correct)



Create network ACL that denies access to all IP address blocks except the various CloudFront IP address ranges.



Add a security group rule to the ALB to allow only the various CloudFront IP address ranges.

Explanation

You can use the managed prefix list for CloudFront as a part of your inbound rules in security groups that you attach to your origin resources, such as your EC2 instances or Application Load Balancers. The AWS-managed prefix lists are created and maintained by AWS and are available to use at no additional cost.

You can reference the managed prefix list for CloudFront in your (Amazon VPC) security group rules. In this case the configuration would only allow access to the ALB with the source set to the prefix list for CloudFront. This solution requires the least operational overhead to maintain.

CORRECT: "Add a security group rule to the ALB to allow traffic from the AWS managed prefix list for CloudFront only" is the correct answer (as explained above.)

INCORRECT: "Add a security group rule to the ALB to allow only the various CloudFront IP address ranges" is incorrect. This would require a lot of operational overhead to maintain as the IP address ranges will change over time.

INCORRECT: "Create a new web ACL that blocks access the application. Associate the new web ACL with the ALB" is incorrect. This is not the most cost-effective solution and may also block access from the CloudFront distribution.

INCORRECT: "Create network ACL that denies access to all IP address blocks except the various CloudFront IP address ranges" is incorrect. This would require a lot of overhead to maintain as the IP ranges for CloudFront change over time.

References:

<https://aws.amazon.com/blogs/networking-and-content-delivery/limit-access-to-your-origins-using-the-aws-managed-prefix-list-for-amazon-cloudfront/>

Save time with our AWS cheat sheets:

<https://digitalcloud.training/amazon-cloudfront/>

Question 21: **Correct**

A company has hundreds of accounts in AWS Organizations. There are several OUs for development teams that each contain multiple accounts. **A manager requires that a report showing usage costs is generated for each development OU** that shows all costs accrued by accounts within the OU.

Which solution meets these requirements?

-

Create an AWS Cost and Usage Report (CUR) by using AWS OpsWorks. Allow each team to visualize the CUR through AWS OpsWorks Stacks.

- Create an AWS Cost and Usage Report (CUR) from the AWS Organizations management account. Allow each team to visualize the CUR through an Amazon QuickSight dashboard.

(Correct)

- Create an AWS Cost and Usage Report (CUR) in each AWS Organizations member account. Allow each team to visualize the CUR through an Amazon QuickSight dashboard.

- Create an AWS Cost and Usage Report (CUR) for each OU by using AWS Cost Explorer. Allow each team to visualize the CUR through an Amazon QuickSight dashboard.

Explanation

You can either generate a CUR from a management or member account. If you generate from member accounts then you must do this individually for each member account which will be a lot of work. In this case it would be better to generate the CUR from the management account of the organization and then use QuickSight to visualize it. Permissions can be granted to the manager of the development OUs to view data relating to the individual accounts in each OU.

CORRECT: "Create an AWS Cost and Usage Report (CUR) from the AWS Organizations management account. Allow each team to visualize the CUR through an Amazon QuickSight dashboard" is the correct answer (as explained above.)

INCORRECT: "Create an AWS Cost and Usage Report (CUR) in each AWS Organizations member account. Allow each team to visualize the CUR through an Amazon QuickSight dashboard" is incorrect. As explained above this would be a lot more work than using the management account to generate the CUR.

INCORRECT: "Create an AWS Cost and Usage Report (CUR) for each OU by using AWS Cost Explorer. Allow each team to visualize the CUR through an Amazon QuickSight dashboard" is incorrect. You cannot generate the CUR using AWS Cost Explorer.

INCORRECT: "Create an AWS Cost and Usage Report (CUR) by using AWS OpsWorks. Allow each team to visualize the CUR through AWS OpsWorks Stacks" is incorrect. OpsWorks has nothing to do with generating a cost and usage report.

References:

<https://aws.amazon.com/premiumsupport/knowledge-center/quicksight-cost-usage-report/>

<https://docs.aws.amazon.com/cur/latest/userguide/cur-consolidated-billing.html>

Save time with our AWS cheat sheets:

<https://digitalcloud.training/aws-cost-management/>

Question 22: **Correct**

A company is running a custom Java application on-premises and plans to migrate the application to the AWS Cloud. The application uses a MySQL database and the application servers maintain users' sessions locally. Which combination of architecture changes will be required to create a highly available solution on AWS? (Select THREE.)

-

Migrate the database to Amazon RDS for MySQL. Configure the RDS instance to use a Multi-AZ deployment.

(Correct)

-

Configure the application to store the user's session in Amazon ElastiCache. Use Application Load Balancers to distribute the load between application instances.

(Correct)

-

Migrate the database to Amazon EC2 instances in multiple Availability Zones. Configure Multi-AZ to synchronize the changes.



Move the Java content to an Amazon S3 bucket configured for static website hosting. Configure cross-Region replication for the S3 bucket contents.



Put the application instances in an Amazon EC2 Auto Scaling group. Configure the Auto Scaling group to create new instances if an instance becomes unhealthy.

(Correct)



Configure the application to run in multiple Regions. Use an Application Load Balancer to distribute the load between application instances.

Explanation

To create a highly available application architecture on AWS the Solutions Architect can use Amazon EC2 Auto Scaling across multiple availability zones to ensure the application instances are available. The instances can be placed behind an ALB to distribute incoming connection requests between them.

For the database layer the MySQL database can be directly migrated to Amazon RDS with Multi-AZ which will provide fault tolerance. Additionally, Amazon ElastiCache can be used for storing session state data so the failure of an instance does not cause data to be lost.

CORRECT: "Configure the application to store the user's session in Amazon ElastiCache. Use Application Load Balancers to distribute the load between application instances" is a correct answer.

CORRECT: "Migrate the database to Amazon RDS for MySQL. Configure the RDS instance to use a Multi-AZ deployment" is also a correct answer.

CORRECT: "Put the application instances in an Amazon EC2 Auto Scaling group. Configure the Auto Scaling group to create new instances if an instance becomes unhealthy" is also a correct answer.

INCORRECT: "Migrate the database to Amazon EC2 instances in multiple Availability Zones. Configure Multi-AZ to synchronize the changes" is incorrect. You cannot configure Multi-AZ for databases running on EC2 instances as this is an RDS feature.

INCORRECT: "Move the Java content to an Amazon S3 bucket configured for static website hosting. Configure cross-Region replication for the S3 bucket contents" is incorrect. Static website hosting cannot be used for dynamic content.

INCORRECT: "Configure the application to run in multiple Regions. Use an Application Load Balancer to distribute the load between application instances" is incorrect. You cannot use an ELB to load balance between Regions, multiple AZs should be used instead.

References:

<https://aws.amazon.com/getting-started/hands-on/building-fast-session-caching-with-amazon-elasticache-for-redis/>

<https://docs.aws.amazon.com/AmazonRDS/latest/UserGuide/Concepts.MultiAZ.html>

Save time with our AWS cheat sheets:

<https://digitalcloud.training/category/aws-cheat-sheets/aws-solutions-architect-professional/aws-management-tools-sap/>

Question 23: **Correct**

A Solutions Architect has deployed an application on Amazon EC2 instances in a private subnet behind a Network Load Balancer (NLB) in a public subnet. Customers have attempted to connect from their office location and are unable to access the application. The targets were registered by instance-id and are all healthy in the associated target group.

What step should the Solutions Architect take to resolve the issue and enable access for the customers?

-

Check the security group for the NLB to ensure it allows egress to the private subnet.

- Check the security group for the NLB to ensure it allows ingress from the customer office.
- Check the security group for the EC2 instances to ensure it allows ingress from the NLB subnets.
- Check the security group for the EC2 instances to ensure it allows ingress from the customer office.

(Correct)

Explanation

The Solutions Architect should check that the security group of the EC2 instances is allowing inbound connections from the customer office IP ranges. Note that NLBs do not have security groups configured and pass connections straight to EC2 instances with the source IP of the client preserved (when registered by instance-id).

With NLBs, when you register EC2 instances as targets, you must ensure that the security groups for these instances allow traffic on both the listener port and the health check port. We know that the health check port is already configured correctly as the targets are all healthy.

CORRECT: "Check the security group for the EC2 instances to ensure it allows ingress from the customer office" is the correct answer.

INCORRECT: "Check the security group for the EC2 instances to ensure it allows ingress from the NLB subnets" is incorrect. This is not necessary as the source IPs of clients are preserved.

INCORRECT: "Check the security group for the NLB to ensure it allows ingress from the customer office" is incorrect. There is no security group associated with an NLB.

INCORRECT: "Check the security group for the NLB to ensure it allows egress to the private subnet" is incorrect. There is no security group associated with an NLB.

References:

<https://docs.aws.amazon.com/elasticloadbalancing/latest/network/target-group-register-targets.html#target-security-groups>

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

Save time with our AWS cheat sheets:

<https://digitalcloud.training/category/aws-cheat-sheets/aws-solutions-architect-professional/security-identity-compliance-sap/>

Question 24: **Correct**

A company uses multiple AWS accounts. There are separate accounts for development, staging, and production environments. Some new requirements have been issued to control costs and improve the overall governance of the AWS accounts. The company must be able to calculate costs associated with each project and each environment. Commonly deployed IT services must be centrally managed and business units should be restricted to deploying pre-approved IT services only.

Which combination of actions should be taken to meet these requirements? (Select TWO.)

-

Use AWS Savings Plans to configure budget thresholds and send alerts to management.

-

Create an AWS Service Catalog portfolio for each business unit and add products to the portfolios using AWS CloudFormation templates.

(Correct)

-

Configure custom budgets and define thresholds using AWS Cost Explorer.

-

Use Amazon CloudWatch to create a billing alarm that notifies managers when a billing threshold is reached or exceeded.

-

Apply environment, cost center, and application name tags to all resources that accept tags.

(Correct)

Explanation

AWS Service Catalog enables organizations to create and manage catalogs of IT services that are approved for AWS. These IT services can include everything from virtual machine images, servers, software, and databases to complete multi-tier application architectures.

AWS Service Catalog allows organizations to centrally manage commonly deployed IT services, and helps organizations achieve consistent governance and meet compliance requirements. End users can quickly deploy only the approved IT services they need, following the constraints set by the organization.

To track the costs associated with projects and environments cost allocation tags should be applied to the relevant resources. Cost allocation tags are used to track AWS costs on a detailed level. After you activate cost allocation tags, AWS uses the cost allocation tags to organize your resource costs on your cost allocation report, to make it easier for you to categorize and track your AWS costs.

CORRECT: "Apply environment, cost center, and application name tags to all resources that accept tags" is a correct answer.

CORRECT: "Create an AWS Service Catalog portfolio for each business unit and add products to the portfolios using AWS CloudFormation templates" is also a correct answer.

INCORRECT: "Configure custom budgets and define thresholds using AWS Cost Explorer" is incorrect. Cost Explorer is used for viewing cost related information but not for creating budgets.

INCORRECT: "Use AWS Savings Plans to configure budget thresholds and send alerts to management" is incorrect as this is not a service but a pricing model and cannot be used for sending alerts.

INCORRECT: "Use Amazon CloudWatch to create a billing alarm that notifies managers when a billing threshold is reached or exceeded" is incorrect. There is no requirement to create billing alarms specified in the scenario.

References:

<https://docs.aws.amazon.com/servicecatalog/latest/adminguide/introduction.html>

<https://docs.aws.amazon.com/awsaccountbilling/latest/aboutv2/cost-alloc-tags.html>

Save time with our AWS cheat sheets:

<https://digitalcloud.training/category/aws-cheat-sheets/aws-solutions-architect-professional/aws-management-tools-sap/>

Question 25: **Correct**

A Solutions Architect needs to design the architecture for an application that requires high availability within and across AWS Regions. The design must support failover to the second Region within 1 minute and must minimize the impact on the user experience. The application will include three tiers, the web tier, application tier and NoSQL data tier.

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

-

Use an Amazon Aurora global database across both Regions so reads and writes can occur either location.

-

Run the web and application tiers in both Regions in an active/active configuration. Use Auto Scaling groups for the web and application layers across multiple Availability Zones in the Regions. Use Spot Instances for the required resources.

-

Use an Amazon Route 53 failover routing policy for failover from the primary Region to the disaster recovery Region. Set Time to Live (TTL) to 30 seconds.

(Correct)

-

Run the web and application tiers in both Regions in an active/passive configuration. Use Auto Scaling groups for the web and application layers across multiple Availability Zones in the Regions. Use zonal Reserved Instances for the minimum number of servers and On-Demand Instances for any additional resources.

(Correct)

-

Use an Amazon Route 53 weighted routing policy set to 100/0 across the two selected Regions. Set Time to Live (TTL) to 30 minutes.

-

Use Amazon DynamoDB with a global table across both Regions so reads and writes can occur in either location.

(Correct)

Explanation

The requirements can be achieved by using an Amazon DynamoDB database with a global table. DynamoDB is a NoSQL database so it fits the requirements. A global table also allows both reads and writes to occur in both Regions.

For the web and application tiers Auto Scaling groups should be configured. Due to the 1-minute RTO these must be configured in an active/passive state. The best pricing model to lower price but ensure resources are available when needed is to use a combination of zonal reserved instances and on-demand instances.

To failover between the Regions, a Route 53 failover routing policy can be configured with a TTL configured on the record of 30 seconds. This will mean clients must resolve against Route 53 every 30 seconds to get the latest record. In a failover scenario the clients would be redirected to the secondary site if the primary site is unhealthy.

CORRECT: "Use an Amazon Route 53 failover routing policy for failover from the primary Region to the disaster recovery Region. Set Time to Live (TTL) to 30 seconds" is a correct answer.

CORRECT: "Use Amazon DynamoDB with a global table across both Regions so reads and writes can occur in either location" is also a correct answer.

CORRECT: "Run the web and application tiers in both Regions in an active/passive configuration. Use Auto Scaling groups for the web and application layers across multiple Availability Zones in the Regions. Use zonal Reserved Instances for the minimum number of servers and On-Demand Instances for any additional resources" is also a correct answer.

INCORRECT: "Use an Amazon Route 53 weighted routing policy set to 100/0 across the two selected Regions. Set Time to Live (TTL) to 30 minutes" is incorrect. A weighted routing policy would need to be updated to change the weightings and the TTL here is too high as clients will cache the result for 30 minutes.

INCORRECT: "Use an Amazon Aurora global database across both Regions so reads and writes can occur either location" is incorrect. An Aurora database is a relational DB, not a NoSQL DB. Also, Aurora global database does not allow writes in multiple Regions, only reads.

INCORRECT: "Run the web and application tiers in both Regions in an active/active configuration. Use Auto Scaling groups for the web and application layers across multiple Availability Zones in the Regions. Use Spot Instances for the required resources" is incorrect. Spot instances may not be available if the maximum price configured is exceeded. This could result in instances not being available when needed.

References:

<https://docs.aws.amazon.com/amazondynamodb/latest/developerguide/GlobalTables.html>

<https://docs.aws.amazon.com/Route53/latest/DeveloperGuide/dns-failover-configuring.html>

Save time with our AWS cheat sheets:

<https://digitalcloud.training/category/aws-cheat-sheets/aws-solutions-architect-professional/aws-database-sap/>

<https://digitalcloud.training/category/aws-cheat-sheets/aws-solutions-architect-professional/aws-networking-content-delivery-sap/>

Question 26: Correct

A company is moving their IT infrastructure to the AWS Cloud and will have several Amazon VPCs within an AWS Region. The company requires centralized and controlled egress-only internet access. The solution must be highly available and horizontally scalable. The company is expecting to grow the number of VPCs to more than fifty.

A Solutions Architect is designing the network for the new cloud deployment. Which design pattern will meet the stated requirements?

-

Attach each VPC to a centralized transit VPC with a VPN connection to each standalone VPC. Outbound internet traffic will be controlled by firewall appliances.

-

Attach each VPC to a shared centralized VPC. Configure VPC peering between each VPC and the centralized VPC. Configure a NAT gateway in two AZs within the centralized VPC.

-

Attach each VPC to a shared transit gateway. Use an egress VPC with firewall appliances in two AZs and attach the transit gateway.



Attach each VPC to a shared transit gateway. Use an egress VPC with firewall appliances in two AZs and connect the transit gateway using IPSec VPNs with BGP.

(Correct)

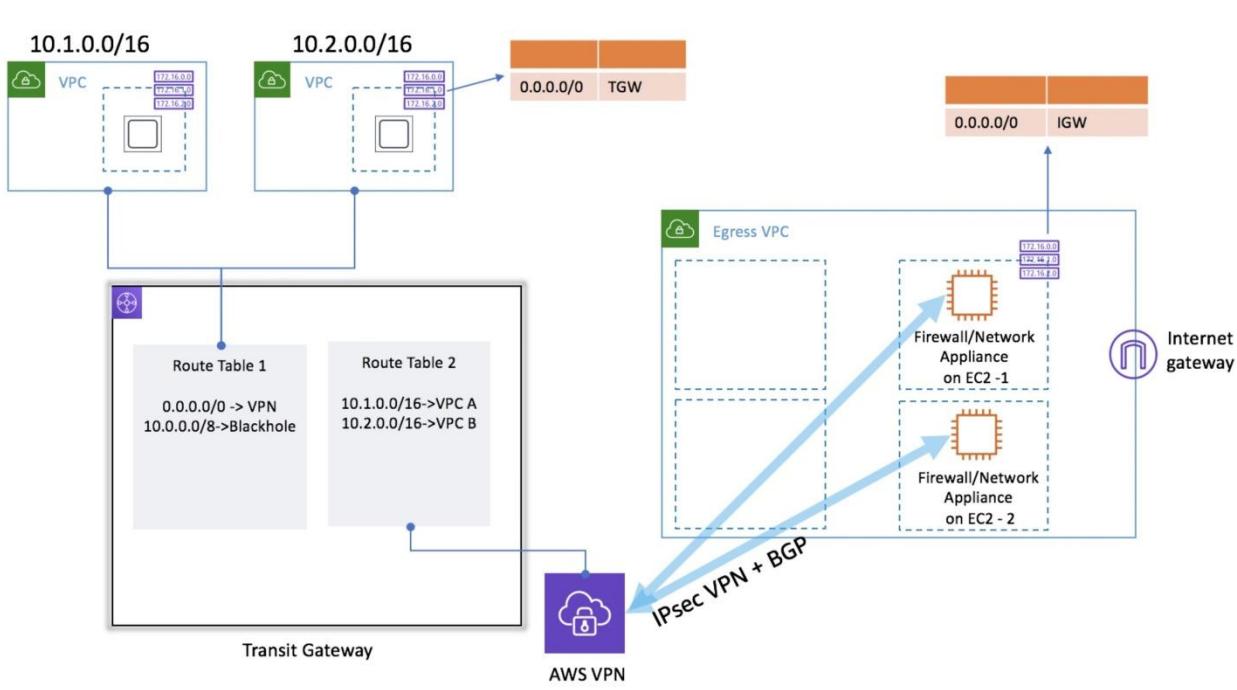
Explanation

A *transit gateway* is a network transit hub that you can use to interconnect your virtual private clouds (VPCs) and on-premises networks. You can attach the following to a transit gateway:

- One or more VPCs
- A Connect SD-WAN/third-party network appliance
- An AWS Direct Connect gateway
- A peering connection with another transit gateway
- A VPN connection to a transit gateway

The correct answer includes a VPN attachment with BGP for an AWS Transit Gateway. This allows BGP equal-cost multipathing (ECMP) to be used which can load balance traffic across multiple EC2 instances. This is the only solution that provides the ability to horizontally scale the outbound internet traffic across multiple appliances with HA across AZs.

The following diagram depicts this architecture:



CORRECT: "Attach each VPC to a shared transit gateway. Use an egress VPC with firewall appliances in two AZs and connect the transit gateway using IPsec VPNs with BGP" is the correct answer.

INCORRECT: "Attach each VPC to a shared transit gateway. Use an egress VPC with firewall appliances in two AZs and attach the transit gateway" is incorrect. Transit Gateway is not a load balancer and will not distribute your traffic evenly across instances in the two AZs. The traffic across the Transit Gateway will stay within an AZ, if possible. Therefore, you are limited by the bandwidth capabilities of a single EC2 instance.

INCORRECT: "Attach each VPC to a shared centralized VPC. Configure VPC peering between each VPC and the centralized VPC. Configure a NAT gateway in two AZs within the centralized VPC." is incorrect. Edge to edge routing is not supported for VPC peering so you cannot route across a VPC peering connection to VPC and then out via a NAT gateway.

INCORRECT: "Attach each VPC to a centralized transit VPC with a VPN connection to each standalone VPC. Outbound internet traffic will be controlled by firewall appliances" is incorrect. A transit VPC is a legacy design pattern, AWS would prefer you to use AWS Transit Gateway for all new requirements. There is also no mention of how scaling and HA is included.

References:

<https://docs.aws.amazon.com/whitepapers/latest/building-scalable-secure-multi-vpc-network-infrastructure/centralized-egress-to-internet.html>

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

Save time with our AWS cheat sheets:

<https://digitalcloud.training/category/aws-cheat-sheets/aws-solutions-architect-professional/aws-database-sap/>

Question 27: **Correct**

A company is updating their operating system patching processes. The company manages both on-premises servers and Amazon EC2 instances using multiple toolsets. A solutions architect wants to utilize a single tool for all servers and instances that can deploy patches and report on patch status.

Which set of actions should the solutions architect take to meet these requirements?

-

Use AWS OpsWorks to deploy patches on the on-premises servers and EC2 instances. Use Amazon Athena to generate patch compliance reports.

-

Use AWS Systems Manager Patch Manager to deploy patches on the on-premises servers and EC2 instances. Use Systems Manager to generate patch compliance reports.

(Correct)

-
-

Use AWS Systems Manager Patch Manager to deploy patches on the EC2 instances and use Run Command for the on-premises servers. Use Systems Manager to generate patch compliance reports.

-
-

Use an Amazon EventBridge rule to apply patches by scheduling an AWS Systems Manager patch remediation job. Use Amazon Inspector to generate patch compliance reports.

Explanation

AWS Systems Manager Patch Manager can be used for both on-premises servers and EC2 instances. Systems Manager can be configured for hybrid environments that include on-premises servers, edge devices, and virtual machines (VMs) that are configured for AWS Systems Manager, including VMs in other cloud environments.

There are several steps that must be taken to configure your on-premises nodes to be managed by AWS Systems Manager. Please refer to the link in the references below to learn more.

CORRECT: "Use AWS Systems Manager Patch Manager to deploy patches on the on-premises servers and EC2 instances. Use Systems Manager to generate patch compliance reports" is the correct answer (as explained above.)

INCORRECT: "Use AWS Systems Manager Patch Manager to deploy patches on the EC2 instances and use Run Command for the on-premises servers. Use Systems Manager to generate patch compliance reports" is incorrect. The patch manager capability of systems manager can be used for the on-premises servers and the EC2 instances, there is no need to use Run Command.

INCORRECT: "Use an Amazon EventBridge rule to apply patches by scheduling an AWS Systems Manager patch remediation job. Use Amazon Inspector to generate patch compliance reports" is incorrect. EventBridge is not needed as Systems Manager has its own capabilities for Patch Management. Systems Manager also has its own reporting, so Inspector is not needed here.

INCORRECT: "Use AWS OpsWorks to deploy patches on the on-premises servers and EC2 instances. Use Amazon Athena to generate patch compliance reports" is incorrect. OpsWorks is not the tool to use for deploying patches across servers and instances. Also, Athena runs queries against S3 buckets so the patch compliance data would need to be stored in S3 somehow.

References:

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

Save time with our AWS cheat sheets:

<https://digitalcloud.training/aws-systems-manager/>

Question 28: **Correct**

A company runs applications on Microsoft Windows servers in an on-premises data center. The servers access a file system shared from one of the Windows servers. Several gigabytes of new data are produced daily. The company is migrating to the cloud and requires the data to be accessible on a file system in the AWS cloud.

Which data migration strategy should the company use?

- Use an AWS Storage Gateway file gateway and point the existing on-premises application servers to the new file gateway.
- Use AWS DataPipeline to schedule a daily task to replicate data between the on-premises file share and Amazon EFS.
- Use an AWS Storage Gateway volume gateway and point the existing on-premises application servers to the new volume gateway.

Use AWS DataSync to schedule a daily task that replicates data between the on-premises file share and Amazon FSX.

(Correct)

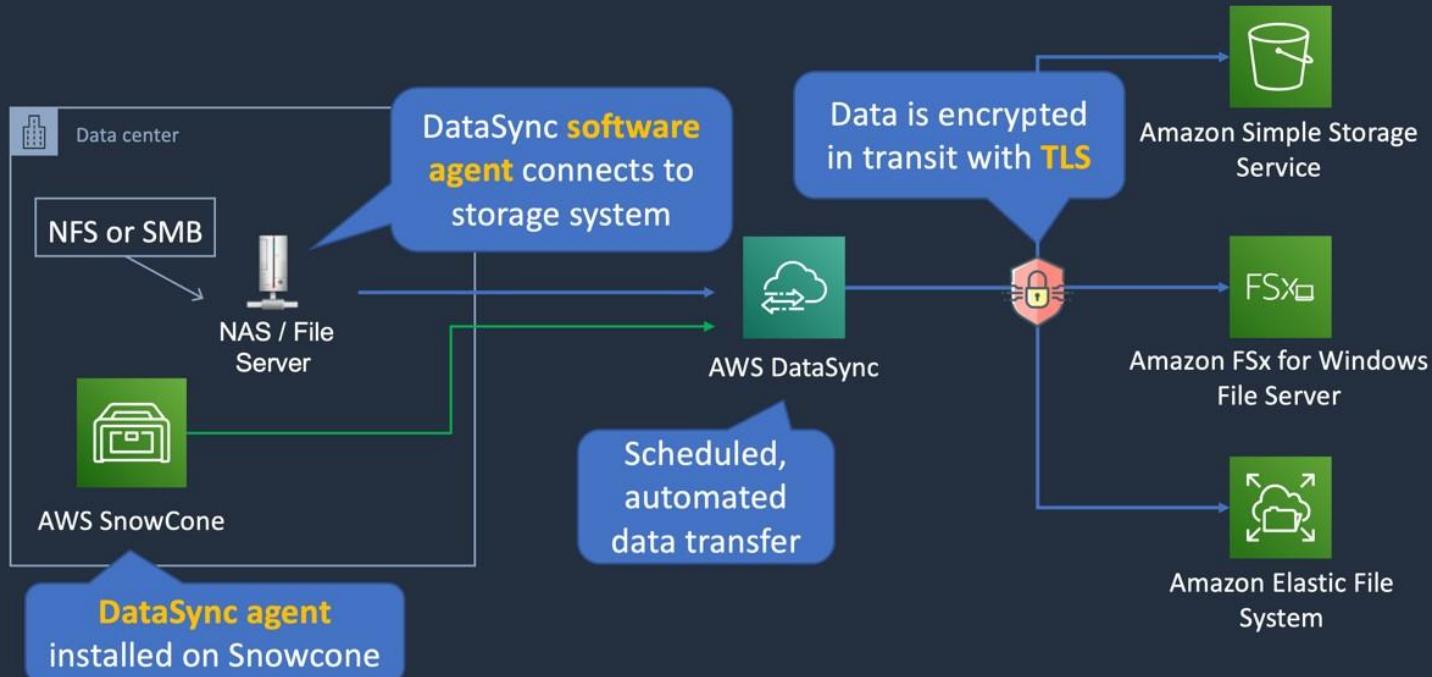
Explanation

AWS recommends using AWS DataSync to transfer data between FSx for Windows File Server file systems. DataSync is a data transfer service that simplifies, automates, and accelerates moving and replicating data between on-premises storage systems and other AWS storage services over the internet or AWS Direct Connect. DataSync can transfer your file system data and metadata, such as ownership, timestamps, and access permissions.

DataSync supports copying NTFS access control lists (ACLs), and also supports copying file audit control information, also known as NTFS system access control lists (SACLs), which are used by administrators to control audit logging of user attempts to access files.



AWS DataSync



© Digital Cloud Training | <https://digitalcloud.training>



CORRECT: "Use AWS DataSync to schedule a daily task that replicates data between the on-premises file share and Amazon FSX" is the correct answer (as explained above.)

INCORRECT: "Use an AWS Storage Gateway file gateway and point the existing on-premises application servers to the new file gateway" is incorrect. When using a file gateway the data is accessible via SMB locally in the on-premises data and is synchronized to Amazon S3. Therefore, there the data is not accessible in the AWS cloud on a file system.

INCORRECT: "Use AWS DataPipeline to schedule a daily task to replicate data between the on-premises file share and Amazon EFS" is incorrect. AWS DataSync is designed for this use case whereas DataPipeline is more suitable to ETL use cases.

INCORRECT: "Use an AWS Storage Gateway volume gateway and point the existing on-premises application servers to the new volume gateway" is incorrect. A volume gateway is used for block-based storage not file-based storage. You access a volume gateway over iSCSI rather than SMB.

References:

<https://docs.aws.amazon.com/fsx/latest/WindowsGuide/migrate-files-to-fsx-datasync.html>

Save time with our AWS cheat sheets:

<https://digitalcloud.training/aws-migration-services/>

Question 29: **Incorrect**

A serverless application is using AWS Lambda and Amazon DynamoDB and developers have finalized an update to the Lambda function code. AWS CodeDeploy will be used to deploy new versions of the function. Updates to the Lambda function should be delivered to a subset of users before deploying the changes to all users. The update process should also be easy to abort and rollback if necessary.

Which CodeDeploy configuration should the solutions architect use?



A blue/green deployment

(Incorrect)



A linear deployment

- An all-at-once deployment
- A canary deployment

(Correct)

Explanation

When using AWS CodeDeploy with AWS Lambda there are three ways traffic can be shifted during a deployment:

- Canary: Traffic is shifted in two increments. You can choose from predefined canary options that specify the percentage of traffic shifted to your updated Amazon ECS task set / Lambda function in the first increment and the interval, in minutes, before the remaining traffic is shifted in the second increment.
- Linear: Traffic is shifted in equal increments with an equal number of minutes between each increment. You can choose from predefined linear options that specify the percentage of traffic shifted in each increment and the number of minutes between each increment.
- All-at-once: All traffic is shifted from the original Lambda function to the updated Lambda function all at once.
- Blue/green: Traffic is shifted from one version of a Lambda function to a new version of the same Lambda function.

All AWS Lambda deployments are blue/green, as a new version of the function is created and traffic is shifted using one of the available options. Canary is the best choice if you just want to shift a percentage of traffic across to a subset of users and then shift the remainder.

CORRECT: "A canary deployment" is the correct answer.

INCORRECT: "A blue/green deployment" is incorrect. All Lambda deployments are blue/green as you cannot do in-place upgrades.

INCORRECT: "A linear deployment" is incorrect. Linear shifts in multiple increments using a percentage per interval. The scenario asks for a 2-step process.

INCORRECT: "An all-at-once deployment" is incorrect. This simply shifts all traffic at once and does not deploy to a subset of users first.

References:

<https://docs.aws.amazon.com/whitepapers/latest/modern-application-development-on-aws/canary-deployments-to-aws-lambda.html>

Save time with our AWS cheat sheets:

<https://digitalcloud.training/category/aws-cheat-sheets/aws-solutions-architect-professional/aws-developer-tools-sap/>

Question 30: Incorrect

A company has created a service that they would like a customer to access. The service runs in the company's AWS account and the customer has a separate AWS account. The company would like to enable the customer to establish least privilege security access using an API or command line tool to the customer account.

What is the MOST secure way to enable the customer to access the service?

-

The company should create an IAM role and assign the required permissions to the IAM role. The customer should then use the IAM role's Amazon Resource Name (ARN) when requesting access to perform the required tasks.

-

The company should provide the customer with their AWS account access keys to log in and perform the required tasks.

-

The company should create an IAM role and assign the required permissions to the IAM role. The customer should then use the IAM role's Amazon Resource Name (ARN), including the external ID in the IAM role's trust policy, when requesting access to perform the required tasks.

(Correct)

-

The company should create an IAM user and assign the required permissions to the IAM user. The company should then provide the credentials to the customer to log in and perform the required tasks.

(Incorrect)

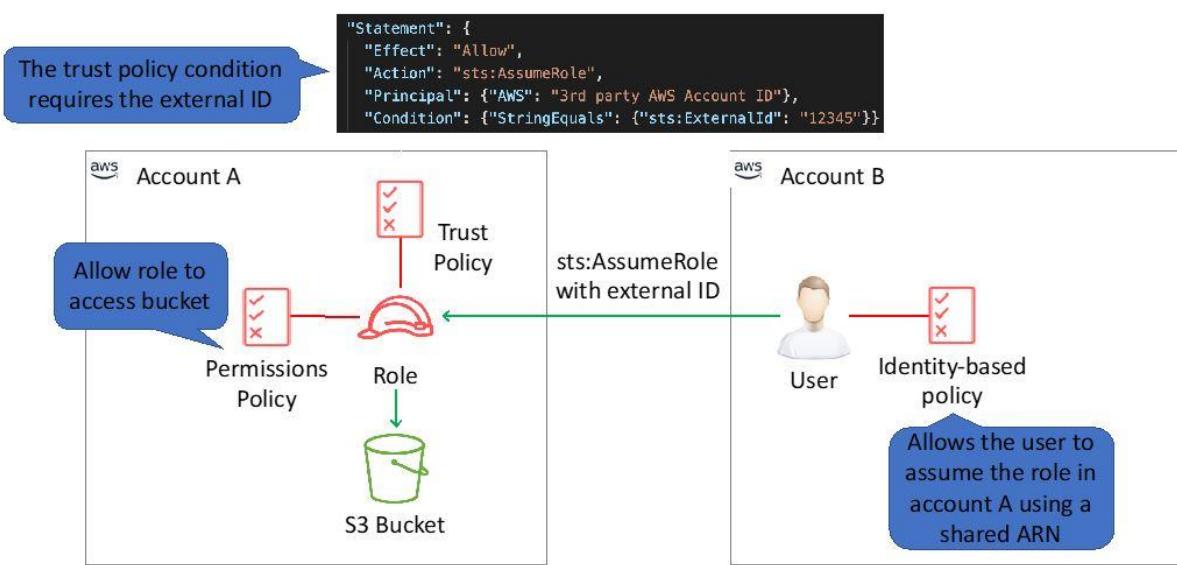
Explanation

At times, you need to give a third-party access to your AWS resources (delegate access). One important aspect of this scenario is the *External ID*, optional information that you can use in an IAM role trust policy to designate who can assume the role.

To require that the third party provides an external ID when assuming a role, update the role's trust policy with the external ID of your choice.

To provide an external ID when you assume a role, use the AWS CLI or AWS API to assume that role.

The following diagram depicts this configuration:



CORRECT: "The company should create an IAM role and assign the required permissions to the IAM role. The customer should then use the IAM role's Amazon Resource Name (ARN), including the external ID in the IAM role's trust policy, when requesting access to perform the required tasks" is the correct answer.

INCORRECT: "The company should create an IAM role and assign the required permissions to the IAM role. The customer should then use the IAM role's Amazon Resource Name (ARN) when requesting access to perform the required tasks" is incorrect. This answer is missing the external ID which is required for security.

INCORRECT: "The company should provide the customer with their AWS account access keys to log in and perform the required tasks" is incorrect. Access keys should never be shared!

INCORRECT: "The company should create an IAM user and assign the required permissions to the IAM user. The company should then provide the credentials to the customer to log in and perform the required tasks" is incorrect. A role should be used rather than a user.

References:

https://docs.aws.amazon.com/IAM/latest/UserGuide/id_roles_create_for-user_externalid.html

Save time with our AWS cheat sheets:

<https://digitalcloud.training/category/aws-cheat-sheets/aws-solutions-architect-professional/security-identity-compliance-sap/>

Question 1: **Correct**

A company has deployed a web application in an Amazon VPC. A CloudFront distribution is used for both scalability and performance. The operations team has noticed that the cache hit ratio has been dropping over time leading to a gradual degradation of the performance for the web application.

The cache metrics report indicates that query strings on some URLs are inconsistently ordered and are specified in a mixture of mixed-case letters.

Which actions can a Solutions Architect take to increase the cache hit ratio and resolve the performance issues on the web application?

-

Create a path pattern in the CloudFront distribution that forwards all requests to the origin with case-sensitivity turned off.

-

Update the CloudFront distribution to disable caching based on query string parameters.

-

Create a Lambda@Edge function to sort parameters by name and force them to be lowercase. Select the CloudFront viewer request trigger to invoke the function.

(Correct)

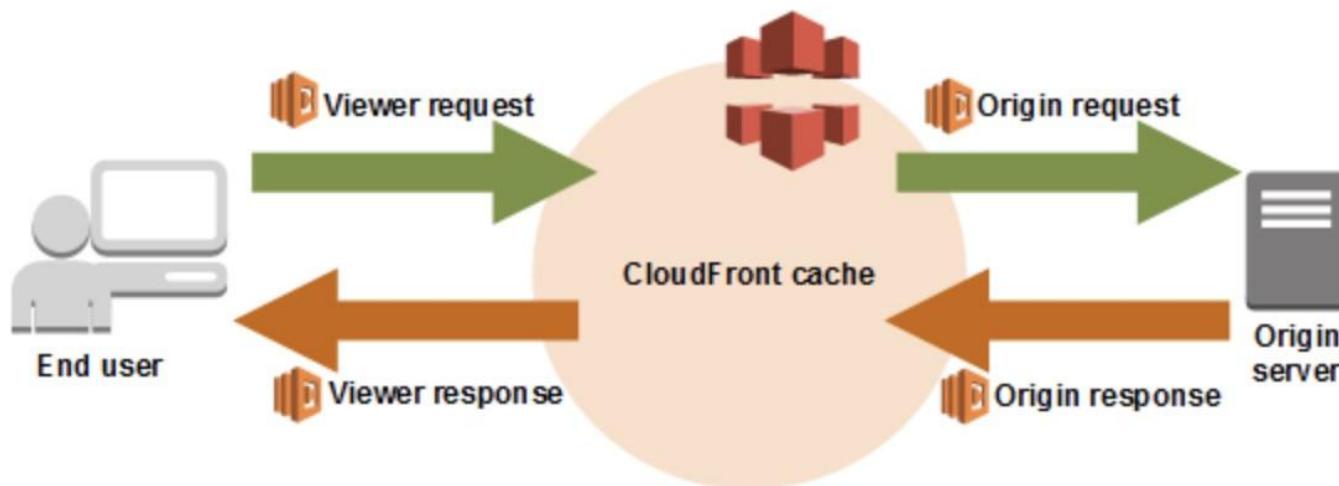
-

Use AWS WAF to create a WebACL and filter based on the case of the query strings in the URL. Configure WAF to trigger an AWS Lambda function that rewrites the URLs to lowercase.

Explanation

Lambda@Edge lets you run Node.js and Python Lambda functions to customize content that CloudFront delivers, executing the functions in AWS locations closer to the viewer. The functions run in response to CloudFront events, without provisioning or managing servers. You can use Lambda functions to change CloudFront requests and responses at the following points:

- After CloudFront receives a request from a viewer (viewer request)
- Before CloudFront forwards the request to the origin (origin request)
- After CloudFront receives the response from the origin (origin response)
- Before CloudFront forwards the response to the viewer (viewer response)



In this scenario the Lambda@Edge function can be written to modify the parameters and rewrite them in lowercase. The function should invoke at the viewer request so that the rewritten query strings can be used to match objects in the cache.

CORRECT: "Create a Lambda@Edge function to sort parameters by name and force them to be lowercase. Select the CloudFront viewer request trigger to invoke the function" is the correct answer.

INCORRECT: "Update the CloudFront distribution to disable caching based on query string parameters" is incorrect. This would not help as it would force all of these requests to the web application and reduce performance further.

INCORRECT: "Use AWS WAF to create a WebACL and filter based on the case of the query strings in the URL. Configure WAF to trigger an AWS Lambda function that rewrites the URLs to lowercase" is incorrect. WAF cannot be configured to directly trigger a Lambda function.

INCORRECT: "Create a path pattern in the CloudFront distribution that forwards all requests to the origin with case-sensitivity turned off" is incorrect. Path patterns are always case-sensitive and this cannot be turned off.

References:

<https://docs.aws.amazon.com/lambda/latest/dg/lambda-edge.html>

<https://aws.amazon.com/premiumsupport/knowledge-center/cloudfront-not-following-cache-behavior/>

Save time with our AWS cheat sheets:

<https://digitalcloud.training/category/aws-cheat-sheets/aws-solutions-architect-professional/aws-networking-content-delivery-sap/>

Question 2: **Incorrect**

A company is planning to migrate on-premises resources to AWS. The resources include over 150 virtual machines (VMs) that use around 50 TB of storage. Most VMs can be taken offline outside of business hours, however, a few are mission critical and downtime must be minimized. The company's internet bandwidth is fully utilized and cannot currently be increased. A Solutions Architect must design a migration strategy that can be completed within the next 3 months.

Which method would fulfill these requirements?

-
-

Migrate mission-critical VMs with AWS SMS. Export the other VMs locally and transfer them to Amazon S3 using AWS Snowball. Use VM Import/Export to import the VMs into Amazon EC2.

(Incorrect)

- Export the VMs locally, beginning with the most mission-critical servers first. Use Amazon S3 Transfer Acceleration to quickly upload each VM to Amazon S3 after they are exported. Use VM Import/Export to import the VMs into Amazon EC2.

- Use an AWS Storage Gateway file gateway. Mount the file gateway and synchronize the VM filesystems to cloud storage. Use the VM Import/Export to import from cloud storage to Amazon EC2.

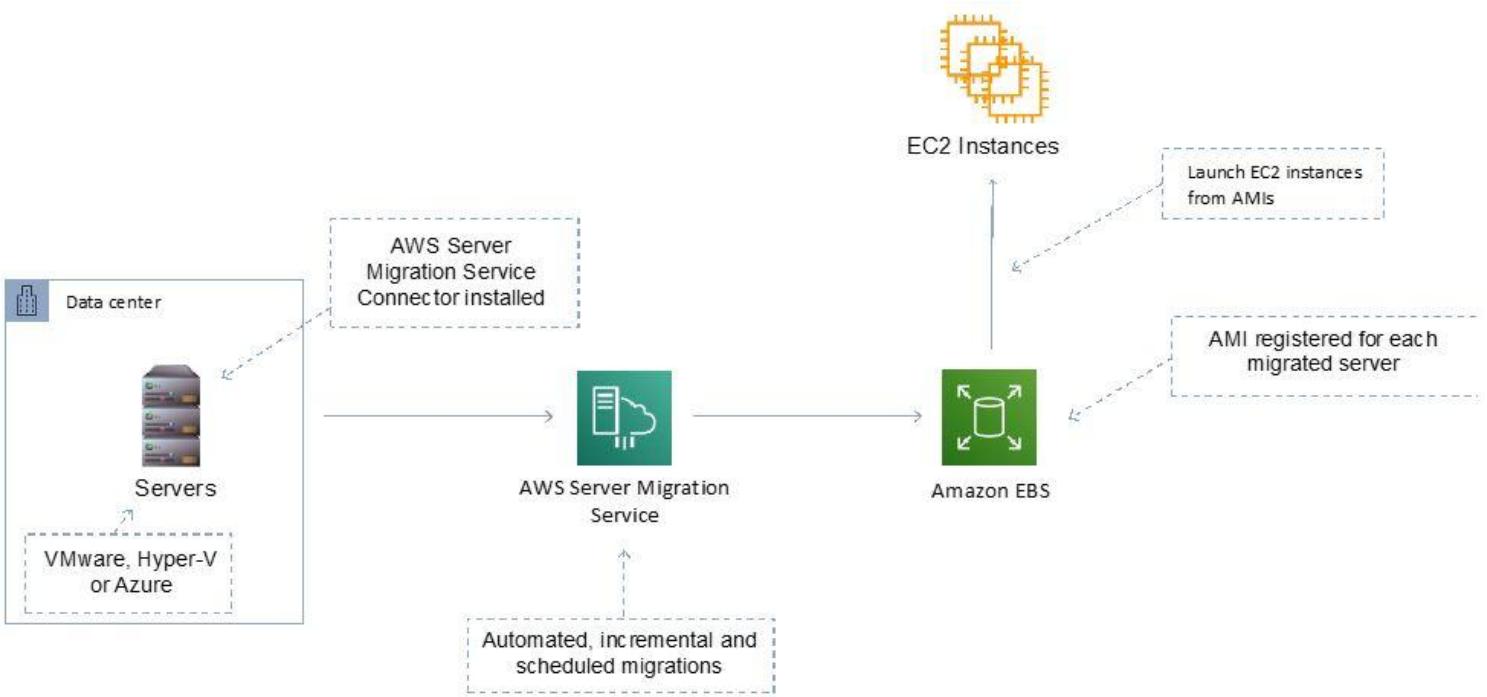
- Set up a 1 Gbps AWS Direct Connect connection. Then, provision a private virtual interface, and use AWS Server Migration Service (SMS) to migrate the VMs into Amazon EC2.

(Correct)

Explanation

The best way to avoid downtime is to provision an AWS Direct Connect connection and use AWS SMS to migrate the VMs into EC2. With support for incremental replication, AWS SMS allows fast, scalable testing of migrated servers. This can also be used to perform a final replication to synchronize the final changes before cutover.

The diagram below depicts the migration process when using AWS SMS:



CORRECT: "Set up a 1 Gbps AWS Direct Connect connection. Then, provision a private virtual interface, and use AWS Server Migration Service (SMS) to migrate the VMs into Amazon EC2" is the correct answer.

INCORRECT: "Migrate mission-critical VMs with AWS SMS. Export the other VMs locally and transfer them to Amazon S3 using AWS Snowball. Use VM Import/Export to import the VMs into Amazon EC2" is incorrect. The VMs that are exported and transported using Snowball will be offline for several days in this scenario which is not acceptable.

INCORRECT: "Use an AWS Storage Gateway file gateway. Mount the file gateway and synchronize the VM filesystems to cloud storage. Use the VM Import/Export to import from cloud storage to Amazon EC2" is incorrect. You cannot migrate VMs in this manner and you cannot mount block-based volumes and replicate the entire operating system volume using file-based storage systems.

INCORRECT: "Export the VMs locally, beginning with the most mission-critical servers first. Use Amazon S3 Transfer Acceleration to quickly upload each VM to Amazon S3 after they are exported. Use VM Import/Export to import the VMs into Amazon EC2" is incorrect. S3 has an object limit of 5 TB which could be an issue for some VMs (maybe). The key problem here is that there is no bandwidth to quickly upload these images, even using Transfer Acceleration will not help if the bottleneck is the saturated internet link at the data center.

References:

<https://docs.aws.amazon.com/server-migration-service/latest/userguide/server-migration.html>

Save time with our AWS cheat sheets:

<https://digitalcloud.training/category/aws-cheat-sheets/aws-solutions-architect-professional/aws-networking-content-delivery-sap/>

<https://digitalcloud.training/category/aws-cheat-sheets/aws-solutions-architect-professional/aws-migration-sap/>

Question 3: Correct

A Solutions Architect is working on refactoring a monolithic application into a modern application design that will be deployed in the AWS Cloud. A CI/CD pipeline should be used that supports the modern design and allows for multiple releases every hour. The pipeline should also ensure that changes can be quickly rolled back if required.

Which design will meet these requirements?



Use AWS CloudFormation StackSets to create production and staging stacks. Update the staging stack and use Amazon Route 53 weighted routing to point to the StackSet endpoint address.



Use AWS Elastic Beanstalk and create a secondary environment configured as a deployment target for the CI/CD pipeline. To deploy, swap the staging and production environment URLs.

(Correct)

-

Package updates into an Amazon EC2 AMI and update the Auto Scaling group to use the new AMI. Terminate existing instances in staged approach to cause launches using the new AMI.

-

Deploy a CI/CD pipeline that incorporates AMIs to contain the application and their configurations. Deploy the application by replacing Amazon EC2 instances.

Explanation

With AWS Elastic Beanstalk you can perform a blue/green deployment, where you deploy the new version to a separate environment, and then swap CNAMEs of the two environments to redirect traffic to the new version instantly.

This approach meets all requirements as it is possible to make multiple updates each hour and it's easy to swap the CNAMEs (URLs) back again in the case of issues occurring.

The image below shows the management console view where you can swap the environment URLs:

Swap environment URLs

When you swap an environment's URL with another environment's URL, you can deploy versions with no downtime. [Learn more](#)

⚠️ Swapping the environment URL will modify the Route 53 DNS configuration, which may take a few minutes. Your application will continue to run while the changes are propagated.

Environment details

Environment name:

staging-env

Environment URL:

staging-env.bx7dx222kw.us-east-2.elasticbeanstalk.com

Select an environment to swap

Environment name:

prod-env (e-2mwwbhpfc)

Environment URL:

prod-env.bx7dx222kw.us-east-2.elasticbeanstalk.com

Cancel

Swap

CORRECT: "Use AWS Elastic Beanstalk and create a secondary environment configured as a deployment target for the CI/CD pipeline. To deploy, swap the staging and production environment URLs" is the correct answer.

INCORRECT: "Use AWS CloudFormation StackSets to create production and staging stacks. Update the staging stack and use Amazon Route 53 weighted routing to point to the StackSet endpoint address" is incorrect. StackSets are used for deploying a stack to different accounts or Regions. There is no StackSet address that you can point users to.

INCORRECT: "Package updates into an Amazon EC2 AMI and update the Auto Scaling group to use the new AMI. Terminate existing instances in staged approach to cause launches using the new AMI" is incorrect. This is not an approach you can take without a lot of manual effort and it will also cause disruption and potential downtime. This approach also makes it hard to roll back.

INCORRECT: "Deploy a CI/CD pipeline that incorporates AMIs to contain the application and their configurations. Deploy the application by replacing Amazon EC2 instances" is incorrect. This approach also does not allow an easy method of rolling back as EC2 instances have been replaced.

References:

<https://docs.aws.amazon.com/elasticbeanstalk/latest/dg/using-features.managing.html>

Save time with our AWS cheat sheets:

<https://digitalcloud.training/category/aws-cheat-sheets/aws-solutions-architect-professional/aws-compute-sap/>

Question 4: **Incorrect**

A company runs Docker containers on Amazon ECS. A containerized application uses a custom tool that must be manually updated each time the container code is updated. The updated container image can then be used for new tasks. A Solutions Architect has been tasked with automating this process to eliminate the manual work and ensure a new container image is generated each time the tool code is updated.

Which combination of actions should the Solutions Architect take to meet these requirements? (Select THREE.)

-

Create an AWS CodeBuild project that pulls the latest container image from Amazon ECR, updates the container with code from the source AWS CodeCommit repository, and pushes the updated container image to Amazon ECR.

(Correct)



Create an AWS CodePipeline pipeline that sources the tool code from the AWS CodeCommit repository and initiates an AWS CodeBuild build.

(Correct)



Create an AWS CodeDeploy application that pulls the latest container image from Amazon ECR, updates the container with code from the source AWS CodeCommit repository, and pushes the updated container image to Amazon ECR.

(Incorrect)



Create an AWS CodePipeline pipeline that sources the tool code from the AWS CodeCommit repository and initiates an AWS CodeDeploy application update.



Create an Amazon EventBridge rule that triggers on commits to the AWS CodeCommit repository for the image. Configure the event to trigger an update to the image in Amazon ECR. Push the updated container image to Amazon ECR.



Create an Amazon ECR repository for the image. Create an AWS CodeCommit repository containing code for the tool being deployed to the container image in Amazon ECR.

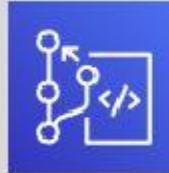
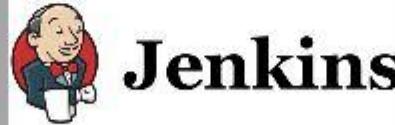
(Correct)

Explanation

Choosing the correct combination of answers relies on understanding which tool to use for which piece of the job. The basic steps and tools to use are as follows:

- 1) Store a container image in a repository – Amazon ECR can be used.
- 2) Store the tool code in the repository – AWS CodeCommit can be used.
- 3) Update the container image with the code and save back to repository – AWS CodeBuild can be used.
- 4) Automate the whole process – AWS CodePipeline can be used to create an automated CI/CD pipeline.

The diagram below depicts the AWS tools (and comparable popular third-party tools) and where they are used in the CI/CD pipeline:

	CODE	BUILD & TEST	DEPLOY	
	  AWS CodeCommit	AWS CodePipeline  AWS CodeBuild	 AWS CodeDeploy	
				

CORRECT: "Create an Amazon ECR repository for the image. Create an AWS CodeCommit repository containing code for the tool being deployed to the container image in Amazon ECR" is a correct answer.

CORRECT: "Create an AWS CodeBuild project that pulls the latest container image from Amazon ECR, updates the container with code from the source AWS CodeCommit repository, and pushes the updated container image to Amazon ECR" is also a correct answer.

CORRECT: "Create an AWS CodePipeline pipeline that sources the tool code from the AWS CodeCommit repository and initiates an AWS CodeBuild build" is also a correct answer.

INCORRECT: "Create an AWS CodeDeploy application that pulls the latest container image from Amazon ECR, updates the container with code from the source AWS CodeCommit repository, and pushes the updated container image to Amazon ECR" is incorrect. CodeDeploy is the wrong tool for this job, use CodeBuild instead.

INCORRECT: "Create an AWS CodePipeline pipeline that sources the tool code from the AWS CodeCommit repository and initiates an AWS CodeDeploy application update" is incorrect. CodeDeploy will deploy the container, but the application must first be updated into the image using CodeBuild.

INCORRECT: "Create an Amazon EventBridge rule that triggers on commits to the AWS CodeCommit repository for the image. Configure the event to trigger an update to the image in Amazon ECR. Push the updated container image to Amazon ECR" is incorrect. CodePipeline should be used rather than EventBridge.

References:

<https://docs.aws.amazon.com/codepipeline/latest/userguide/welcome.html>

<https://aws.amazon.com/ecr/getting-started/>

Save time with our AWS cheat sheets:

<https://digitalcloud.training/category/aws-cheat-sheets/aws-solutions-architect-professional/aws-compute-sap/>

<https://digitalcloud.training/category/aws-cheat-sheets/aws-solutions-architect-professional/aws-developer-tools-sap/>

Question 5: **Incorrect**

A company offers a photo sharing application to its users through a social networking app. To ensure images can be displayed with consistency, a single Amazon EC2 instance running JavaScript code processes the photos and stores the processed images in an Amazon S3 bucket. A front-end application runs from a static website in another S3 bucket and loads the processed images for display in the app.

The company has asked a Solutions Architect to make some recommendations for a cost-effective solution that offers massive scalability for a global user base.

Which combination of changes should the Solutions Architect recommend? (Select TWO.)

-

Deploy the applications in an Amazon ECS cluster and apply Service Auto Scaling.

(Incorrect)

-

Create an Amazon CloudFront distribution in front of the processed images bucket.

(Correct)

-

Replace the EC2 instance with AWS Lambda to run the image processing tasks.

(Correct)

-

Replace the EC2 instance with Amazon Rekognition for image processing.

-

Place the image processing EC2 instance into an Auto Scaling group.

Explanation

Though it is not clearly documented the image processing task is likely to take less than the maximum execution time for a Lambda function (15 minutes). Therefore, Lambda could be more cost-effective and will be highly scalable.

Amazon CloudFront can be used with an origin configured as the S3 bucket with the processed images. This will provide a better user experience as users of the app in multiple geographies can retrieve images cached at edge locations with great performance.

CORRECT: "Replace the EC2 instance with AWS Lambda to run the image processing tasks" is a correct answer.

CORRECT: "Create an Amazon CloudFront distribution in front of the processed images bucket" is also a correct answer.

INCORRECT: "Place the image processing EC2 instance into an Auto Scaling group" is incorrect. This will provide some improvements but it does not represent as much of an improvement as Lambda. Lambda is likely to be cheaper will scale very well for this application.

INCORRECT: "Replace the EC2 instance with Amazon Rekognition for image processing" is incorrect. The question states that the images are being processed for "consistency". Though this is vague, it is unlikely to mean identifying objects or performing the type of analysis Rekognition performs.

INCORRECT: "Deploy the applications in an Amazon ECS cluster and apply Service Auto Scaling" is incorrect. This is also a valid solution for the image processing but less cost-effective than Lambda.

References:

<https://aws.amazon.com/lambda/features/>

<https://docs.aws.amazon.com/AmazonCloudFront/latest/DeveloperGuide/DownloadDistS3AndCustomOrigins.html>

Save time with our AWS cheat sheets:

<https://digitalcloud.training/category/aws-cheat-sheets/aws-solutions-architect-professional/aws-compute-sap/>

<https://digitalcloud.training/category/aws-cheat-sheets/aws-solutions-architect-professional/aws-networking-content-delivery-sap/>

Question 6: **Correct**

A pharmaceutical company has deployed an application on their private Amazon VPC. They need to use a third-party software-as-a-service (SaaS) application which is hosted in another AWS account inside an Amazon VPC.

They need to connect applications to the third-party SaaS from private subnets in the company VPC. The company's security team has mandated policies that private network needs to be used without internet propagation. No resources that run in the company VPC are allowed to be accessed from outside the company's VPC. All permissions must conform to the principles of least privilege.

Which solution meets these requirements?

-

Create a VPC peering connection between the third-party SaaS application and the company VPC. Update route tables by adding the required routes for the peering connection.

-

Create an AWS PrivateLink interface VPC endpoint. Connect this endpoint to the endpoint service that the third-party SaaS application provides. Create a security group to limit the access to the endpoint and associate the security group with the endpoint.

(Correct)

-

Create an AWS Site-to-Site VPN connection between the third-party SaaS application and the company VPC. Configure network ACLs to limit access across the VPN tunnels.

-

Create an AWS PrivateLink endpoint service. Ask the third-party SaaS provider to create an interface VPC endpoint for this endpoint service. Grant permissions for the endpoint service to the specific account of the third-party SaaS provider.

Explanation

AWS PrivateLink provides you option to connect to SaaS products privately, as if it is running on customer's VPC itself.

A service consumer creates a *VPC endpoint* to connect their VPC to an endpoint service. A service consumer must specify the service name of the endpoint service when creating a VPC endpoint. There are multiple types of VPC endpoints. You must create the type of VPC endpoint that's required by the endpoint service.

- **Interface** - Create an *interface endpoint* to send traffic to endpoint services that use a Network Load Balancer to distribute traffic. Traffic destined for the endpoint service is resolved using DNS.
- **GatewayLoadBalancer** - Create a *Gateway Load Balancer endpoint* to send traffic to a fleet of virtual appliances using private IP addresses. You route traffic from your VPC to the Gateway Load Balancer endpoint using route tables. The Gateway Load Balancer distributes traffic to the virtual appliances and can scale with demand.
- **Gateway** - Create a *gateway endpoint* to send traffic to Amazon S3 or DynamoDB using private IP addresses. You route traffic from your VPC to the gateway endpoint using route tables. Gateway endpoints do not enable AWS PrivateLink.

In this case the AWS PrivateLink service must already be published by the service provider. The company then acts as the service consumer by creating an interface endpoint to connect to the service provider's service.

CORRECT: "Create an AWS PrivateLink interface VPC endpoint. Connect this endpoint to the endpoint service that the third-party SaaS application provides. Create a security group to limit the access to the endpoint. Associate the security group with the endpoint" is the correct answer (as explained above.)

INCORRECT: "Create an AWS Site-to-Site VPN connection between the third-party SaaS application and the company VPC. Configure network ACLs to limit access across the VPN tunnels" is incorrect.

You cannot create S2S VPNs between VPCs or applications using AWS services. This would need to be configured using 3rd party software.

INCORRECT: "Create a VPC peering connection between the third-party SaaS application and the company VPC. Update route tables by adding the required routes for the peering connection" is incorrect.

While VPC peering enables you to privately connect VPCs, AWS PrivateLink enables you to configure applications or services in VPCs as endpoints that your VPC peering connections can connect to. This is a more secure solution as there is less trust required.

INCORRECT: "Create an AWS PrivateLink endpoint service. Ask the third-party SaaS provider to create an interface VPC endpoint for this endpoint service. Grant permissions for the endpoint service to the specific account of the third-party SaaS provider" is incorrect.

The AWS PrivateLink endpoint service is published by the service provider which in this case is the third-party SaaS provider. The company is then acting as the service consumer and needs to create a VPC interface endpoint.

References:

<https://docs.aws.amazon.com/vpc/latest/privatelink/privatelink-access-saas.html>

<https://docs.aws.amazon.com/vpc/latest/privatelink/interface-endpoints.html>

Save time with our AWS cheat sheets:

<https://digitalcloud.training/amazon-vpc/>

Question 7: Incorrect

A fintech company runs an on-premises environment that ingests data feeds from financial services companies, transforms the data, and then sends it to an on-premises Apache Kafka cluster. The company plans to use AWS services to build a scalable, near real-time solution that offers consistent network performance to provide the data feeds to a web application. Which steps should a Solutions Architect take to build the solution? (Select THREE.)

-

Establish a Site-to-Site VPN from the on-premises data center to AWS.

(Incorrect)

-

Create a GraphQL API in AWS AppSync, create an AWS Lambda function to process the Amazon Kinesis data stream, and use the @connections command to send callback messages to connected clients.

-

Create a WebSocket API in Amazon API Gateway, create an AWS Lambda function to process an Amazon Kinesis data stream, and use the @connections command to send callback messages to connected clients.

(Correct)

-

Create an Amazon EC2 Auto Scaling group to pull the messages from the on-premises Kafka cluster and use the Amazon Kinesis Producer Library to put the data into a Kinesis data stream.

(Correct)

-

Establish an AWS Direct Connect connection from the on-premises data center to AWS.

(Correct)

-

Create an Amazon EC2 Auto Scaling group to pull the messages from the on-premises Kafka cluster and use the Amazon Consumer Library to put the data into an Amazon Kinesis data stream.

Explanation

The solution requires consistent network performance so we can eliminate the VPN as that would use the internet. Instead, we must use AWS Direct Connect to connect the on-premises environment to the Amazon VPC.

An Auto Scaling group of EC2 instances can then be used with the Kinesis Producer Library to put the data into a Kinesis data stream. The role here is producer NOT consumer.

Finally, a Lambda function can process the messages from the stream using the Kinesis Consumer Library and update the web application. An API Gateway with a WebSocket API is used for the backend and @connections commands can send callback messages to connected clients as data is updated.

CORRECT: "Establish an AWS Direct Connect connection from the on-premises data center to AWS" is a correct answer.

CORRECT: "Create an Amazon EC2 Auto Scaling group to pull the messages from the on-premises Kafka cluster and use the Amazon Kinesis Producer Library to put the data into a Kinesis data stream" is also a correct answer.

CORRECT: "Create a WebSocket API in Amazon API Gateway, create an AWS Lambda function to process an Amazon Kinesis data stream, and use the @connections command to send callback messages to connected clients" is also a correct answer.

INCORRECT: "Create an Amazon EC2 Auto Scaling group to pull the messages from the on-premises Kafka cluster and use the Amazon Consumer Library to put the data into an Amazon Kinesis data stream" is incorrect. The producer library should be used here rather than the consumer library.

INCORRECT: "Create a GraphQL API in AWS AppSync, create an AWS Lambda function to process the Amazon Kinesis data stream, and use the @connections command to send callback messages to connected clients" is incorrect. The @connections command is a feature of a WebSocket API and is not available for GraphQL.

INCORRECT: "Establish a Site-to-Site VPN from the on-premises data center to AWS" is incorrect. A VPN will typically use the public internet and therefore cannot offer consistent network performance.

References:

<https://docs.aws.amazon.com/streams/latest/dev/developing-producers-with-kpl.html>

<https://docs.aws.amazon.com/apigateway/latest/developerguide/apigateway-how-to-call-websocket-api.html>

Save time with our AWS cheat sheets:

<https://digitalcloud.training/category/aws-cheat-sheets/aws-solutions-architect-professional/aws-networking-content-delivery-sap/>

<https://digitalcloud.training/category/aws-cheat-sheets/aws-solutions-architect-professional/aws-compute-sap/>

Question 8: **Correct**

A media advertising company currently has resources hosted in two AWS accounts: Management and Production. DNS records are stored in a private hosted zone using Amazon Route 53 in the Management account. The Production account is used for applications and databases.

The company has deployed a two-tier application in a new VPC. To simplify the configuration, the database.company.com CNAME record set for the Amazon RDS endpoint was created in a private hosted zone for Amazon Route 53.

While deploying, the application failed to start. Troubleshooting revealed that database.company.com is not resolvable within the Amazon EC2 instance. The solutions architect confirmed that the record set was created correctly in Route 53.

Which combination of steps should the solutions architect take to resolve this issue? (Select TWO.)

-

Create an authorization to associate the private hosted zone in the Management account with the new VPC in the Production account.

(Correct)

-

Use SSH to connect to the application tier EC2 instance. Add an RDS endpoint IP address to the /etc/resolv.conf file.

-

Deploy the database on a separate EC2 instance in the new VPC. Create a record set for the instance's private IP in the private hosted zone.



Create a private hosted zone for the example.com domain in the Production account. Configure Route 53 replication between AWS accounts.



Associate a new VPC in the Production account with a hosted zone in the Management account. Delete the association authorization in the Management account.

(Correct)

Explanation

Below is the sequence of steps for achieving this:

- Connect to an Amazon Elastic Compute Cloud (Amazon EC2) instance in the Management account.
- Run the following commands:
 - pip3 install awscli --upgrade --user
 - aws route53 list-hosted-zones (to find the hosted zone to be linked)
 - aws route53 create-vpc-association-authorization --hosted-zone-id <hosted-zone-id> --vpc VPCRegion=<region>,VPCId=<vpc-id> --region <Region>
- Connect to an Amazon EC2 instance in target account

CORRECT: "Create an authorization to associate the private hosted zone in the Management account with the new VPC in the Production account" is a correct answer (as explained above.)

CORRECT: "Associate a new VPC in the Production account with a hosted zone in the Management account. Delete the association authorization in the Management account" is also a correct answer (as explained above.)

INCORRECT: "Deploy the database on a separate EC2 instance in the new VPC. Create a record set for the instance's private IP in the private hosted zone" is incorrect.

Deploying the database on a separate EC2 instance might be cost inefficient and might not be possible due to architectural constraints. Since associating these accounts is possible by VPC association, that is a better option.

INCORRECT: "Use SSH to connect to the application tier EC2 instance. Add an RDS endpoint IP address to the /etc/resolv.conf file" is incorrect.

Using SSH to login into instance would mean maintaining private keys for the instance which can lead to security issues hence this option is not correct.

INCORRECT: "Create a private hosted zone for the example.com domain in the Production account. Configure Route 53 replication between AWS accounts" is incorrect.

You cannot replicate the records on an ongoing basis between accounts / zones. This is not a native Route 53 feature.

References:

<https://aws.amazon.com/premiumsupport/knowledge-center/route53-private-hosted-zone/>

<https://docs.aws.amazon.com/Route53/latest/DeveloperGuide/hosted-zone-private-associate-vpcs-different-accounts.html>

Save time with our AWS cheat sheets:

<https://digitalcloud.training/amazon-route-53/>

Question 9: **Correct**

An application uses Amazon EC2 instances in an Auto Scaling group and an Amazon RDS MySQL database. The web application has occasional spikes of traffic during the day. The operations team have determined the most appropriate instances sizes for both the EC2 instances and the DB instance. All instances use On-Demand pricing.

What of the following steps can be taken to gain the most cost savings without impacting the reliability of the application?

-
- Use On-Demand pricing for the RDS database and use Spot pricing for the EC2 instances in the Auto Scaling group**
-
- Reserve capacity for all EC2 instances and leverage Spot Instance pricing for the RDS database.**
-
- Use Spot instance pricing for the RDS database and the EC2 instances in the Auto Scaling group.**
-
- Reserve capacity for the RDS database and the minimum number of EC2 instances that are constantly running.**

(Correct)

Explanation

The best cost saving measure is to reserve capacity for the RDS database as it should be adequately sized to handle any small bursts of traffic (RDS must have vertical capacity, or you must offload reads).

For Amazon EC2 instance, a combination of reserved instances and on-demand is the best option. The reserved instances should be used for the steady state usage requirement, and on-demand can be used to handle additional instances launched during busy periods.

CORRECT: "Reserve capacity for the RDS database and the minimum number of EC2 instances that are constantly running" is the correct answer.

INCORRECT: "Use Spot instance pricing for the RDS database and the EC2 instances in the Auto Scaling group" is incorrect. Spot instances can be terminated when AWS need the capacity back, this could impact the reliability of the application.

INCORRECT: "Use On-Demand pricing for the RDS database and use Spot pricing for the EC2 instances in the Auto Scaling group" is incorrect. This is not the most cost-effective choice and leaves the web application vulnerable to instance termination.

INCORRECT: "Reserve capacity for all EC2 instances and leverage Spot Instance pricing for the RDS database" is incorrect. This option does not cater for the spikes in load and leaves the DB vulnerable to instance termination.

References:

<https://aws.amazon.com/ec2/pricing/reserved-instances/>

<https://aws.amazon.com/ec2/spot/>

Save time with our AWS cheat sheets:

<https://digitalcloud.training/category/aws-cheat-sheets/aws-solutions-architect-professional/aws-compute-sap/>

<https://digitalcloud.training/category/aws-cheat-sheets/aws-solutions-architect-professional/aws-management-tools-sap/>

Question 10: **Correct**

A company requires federated access to AWS for users of a mobile application. The security team has mandated that the application must use a custom-built solution for authenticating users and use IAM roles for authorization.

Which of the following actions would enable authentication and authorization and satisfy the requirements? (Select TWO.)

-

Use a custom-built SAML-compatible solution that uses LDAP for authentication and uses a SAML assertion to perform authorization to the IAM identity provider.

(Correct)



Use a custom-built SAML-compatible solution for authentication and use AWS SSO for authorization.



Create a custom-built LDAP connector using Amazon API Gateway and AWS Lambda for authentication. Use a token-based Lambda authorizer that uses JWT.



Use a custom-built OpenID Connect-compatible solution for authentication and use Amazon Cognito for authorization.

(Correct)



Use a custom-built OpenID Connect-compatible solution with AWS SSO for authentication and authorization.

Explanation

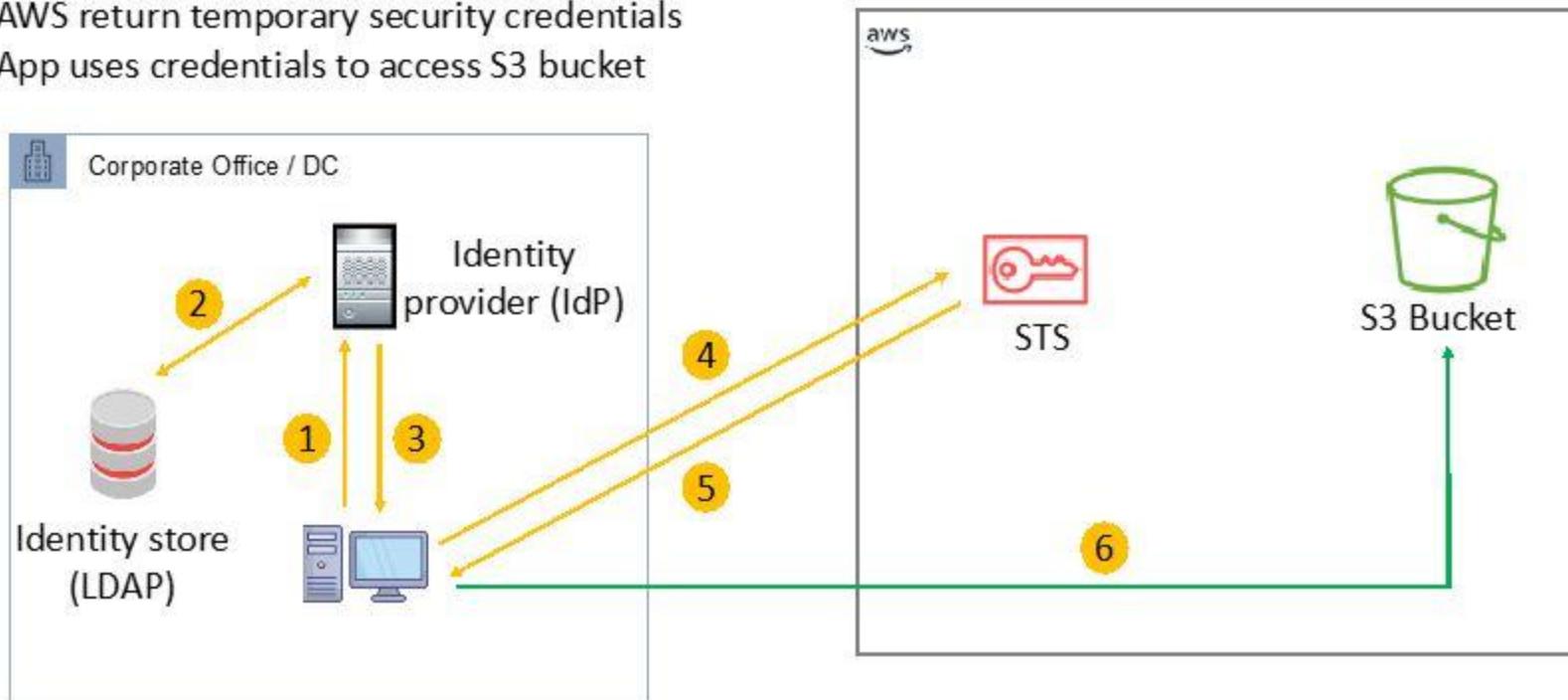
There are two possible solutions for this scenario:

- An OpenID Connect provider can be added in IAM to enable federated authentication. An Amazon Cognito identity pool can then be used for authorization. Amazon Cognito identity pools provide temporary AWS credentials for users who are guests (unauthenticated) and for users who have been authenticated and received a token.

- AWS supports identity federation with SAML 2.0. This feature enables federated single sign-on (SSO), so users can log into the AWS Management Console or call the AWS API operations without you having to create an IAM user for everyone in your organization. In your organization's IdP, you define assertions that map users or groups in your organization to the IAM roles.

The diagram below shows how the authentication and authorization process works when assuming a role using SAML:

1. Client application attempts to authenticate using IdP
2. IdP authenticates the user
3. IdP sends client SAML assertion
4. App calls sts:AssumeRoleWithSAML
5. AWS return temporary security credentials
6. App uses credentials to access S3 bucket



CORRECT: "Use a custom-built OpenID Connect-compatible solution for authentication and use Amazon Cognito for authorization" is a correct answer.

CORRECT: "Use a custom-built SAML-compatible solution that uses LDAP for authentication and uses a SAML assertion to perform authorization to the IAM identity provider" is also a correct answer.

INCORRECT: "Use a custom-built SAML-compatible solution for authentication and use AWS SSO for authorization" is incorrect. AWS SSO cannot be used for mobile applications.

INCORRECT: "Create a custom-built LDAP connector using Amazon API Gateway and AWS Lambda for authentication. Use a token-based Lambda authorizer that uses JWT" is incorrect. This is not a complete solution and API Gateway is not required for this solution.

INCORRECT: "Use a custom-built OpenID Connect-compatible solution with AWS SSO for authentication and authorization" is incorrect. AWS SSO cannot be used for mobile applications.

References:

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

https://docs.aws.amazon.com/IAM/latest/UserGuide/id_roles_providers_create_oidc.html

https://docs.aws.amazon.com/IAM/latest/UserGuide/id_roles_providers_saml.html

Save time with our AWS cheat sheets:

<https://digitalcloud.training/category/aws-cheat-sheets/aws-solutions-architect-professional/security-identity-compliance-sap/>

Question 11: **Incorrect**

A Solutions Architect is helping to standardize a company's method of deploying applications to AWS using AWS CodePipeline and AWS CloudFormation. A group of developers create applications using JavaScript and TypeScript and they are concerned about needing to learn new domain-specific languages. They are also reluctant to lose access to features of the existing languages such as looping.

How can the Solutions Architect address the developers concerns and quickly bring the applications up to deployment standards?



Define the AWS resources using JavaScript or TypeScript. Use the AWS Cloud Development Kit (AWS CDK) to create CloudFormation templates from the developers' code and use the AWS CDK to create CloudFormation stacks. Incorporate the AWS CDK as a CodeBuild job in CodePipeline.

(Correct)



Create CloudFormation templates and re-use parts of the JavaScript and TypeScript code as Instance user data. Use the AWS Cloud Development Kit (AWS CDK) to deploy the application using these templates. Incorporate the AWS CDK into CodePipeline and deploy the application to AWS using these templates.

(Incorrect)



Use a third-party resource provisioning engine inside AWS CodeBuild to standardize the deployment processes. Orchestrate the CodeBuild job using CodePipeline and use CloudFormation for deployment.

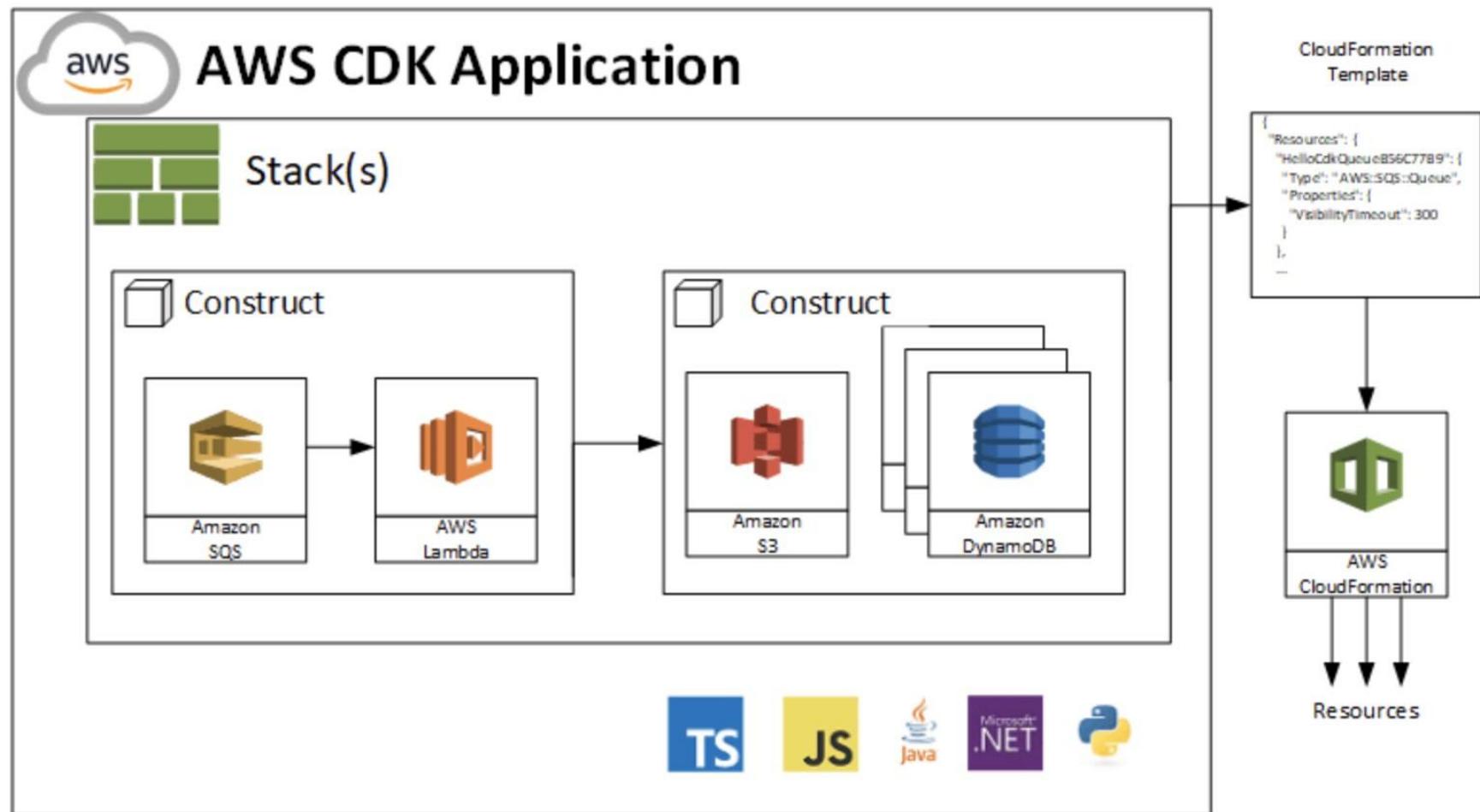


Use AWS SAM and specify a serverless transform. Add the JavaScript and TypeScript code as metadata to the template file. Use AWS CodeBuild to build the code and output a CloudFormation template.

Explanation

The AWS CDK is a software development framework for defining cloud infrastructure in code and provisioning it through AWS CloudFormation. You can use the AWS CDK to define your cloud resources in a familiar programming language. The AWS CDK supports TypeScript, JavaScript, Python, Java, and C#/.Net.

Developers can use one of the supported programming languages to define reusable cloud components known as [Constructs](#). You compose these together into [Stacks](#) and [Apps](#). The diagram below depicts how an AWS CDK application is constructed:



CORRECT: "Define the AWS resources using JavaScript or TypeScript. Use the AWS Cloud Development Kit (AWS CDK) to create CloudFormation templates from the developers' code and use the AWS CDK to create CloudFormation stacks. Incorporate the AWS CDK as a CodeBuild job in CodePipeline" is the correct answer.

INCORRECT: "Create CloudFormation templates and re-use parts of the JavaScript and TypeScript code as Instance user data. Use the AWS Cloud Development Kit (AWS CDK) to deploy the application using these templates. Incorporate the AWS CDK into CodePipeline and deploy the application to AWS using these templates" is incorrect. You cannot use instance user data to run JavaScript or TypeScript code. This is not a way to deploy one of these applications.

INCORRECT: "Use AWS SAM and specify a serverless transform. Add the JavaScript and TypeScript code as metadata to the template file. Use AWS CodeBuild to build the code and output a CloudFormation template" is incorrect. AWS SAM is used for deploying serverless applications using CloudFormation. You cannot run code using metadata in an AWS SAM app. You also cannot use CodeBuild to create a CloudFormation template.

INCORRECT: "Use a third-party resource provisioning engine inside AWS CodeBuild to standardize the deployment processes. Orchestrate the CodeBuild job using CodePipeline and use CloudFormation for deployment" is incorrect. CodeBuild does not do resources provisioning – CodeDeploy or CloudFormation does.

References:

<https://docs.aws.amazon.com/cdk/latest/guide/home.html>

Save time with our AWS cheat sheets:

<https://digitalcloud.training/category/aws-cheat-sheets/aws-solutions-architect-professional/aws-developer-tools-sap/>

Question 12: **Correct**

A company runs an application in an on-premises data center that uses an IBM Db2 database. The web application calls an API that runs stored procedures on the database to retrieve read-only data. The dataset is constantly updated. Users have reported significant latency when attempting to retrieve data. The company are concerned about Db2 CPU licensing costs and the performance of the database.

Which approach should a Solutions Architect take to migrate to AWS and resolve these concerns?



Use AWS DMS to migrate data to Amazon DynamoDB using a continuous replication task. Refactor the API to use the DynamoDB data. Implement the refactored API in Amazon API Gateway and enable API caching.

(Correct)



Export data on a daily basis and upload to Amazon S3. Refactor the API to use the S3 data. Implement Amazon API Gateway and enable API caching.



Use local storage to cache query output. Use S3 COPY commands to sync the dataset to Amazon S3. Refactor the API to use Amazon EFS. Implement Amazon API Gateway and enable API caching.



Rehost the Db2 database to an Amazon EC2 instance. Migrate all the data. Enable caching using an instance store. Refactor the API to use the Amazon EC2 Db2 database. Implement Amazon API Gateway and enable API caching.

Explanation

The AWS Database Migration Service (DMS) can be used to migrate from IBM Db2 to targets including relational databases (such as Oracle and Amazon Aurora), a data warehouse (Amazon Redshift), a NoSQL database (Amazon DynamoDB), or an Amazon S3 bucket.

You can create an AWS DMS task that captures ongoing changes to the source data store. You can do this capture while you are migrating your data. You can also create a task that captures ongoing changes after you complete your initial (full-load) migration to a supported target data store.

This process is called ongoing replication or change data capture (CDC). AWS DMS uses this process when replicating ongoing changes from a source data store. This process works by collecting changes to the database logs using the database engine's native API.

Create endpoint

AWS DMS accesses your data sources and targets using endpoints. A source endpoint allows AWS DMS to read data from a database (on-premise or in the cloud), or from a non-database source such as Amazon S3. A target endpoint allows AWS DMS to write data to a database, or to a non-database target.

We recommend that you choose "Run test" on this page, to verify that your endpoint is valid before using it in an AWS DMS task.

Endpoint type* Source Target ⓘ

Endpoint identifier* db2 ⓘ

Source engine* db2 ⓘ

Server name* db2.server.com

Port* 8192 ⓘ

SSL mode* none ⓘ

User name* admin ⓘ

Password* ⓘ

Database name* db2 ⓘ

CORRECT: "Use AWS DMS to migrate data to Amazon DynamoDB using a continuous replication task. Refactor the API to use the DynamoDB data. Implement the refactored API in Amazon API Gateway and enable API caching" is the correct answer.

INCORRECT: "Rehost the Db2 database to an Amazon EC2 instance. Migrate all the data. Enable caching using an instance store. Refactor the API to use the Amazon EC2 Db2 database. Implement Amazon API Gateway and enable API caching" is incorrect. This solution does not include a method of synchronizing the data changes.

INCORRECT: "Use local storage to cache query output. Use S3 COPY commands to sync the dataset to Amazon S3. Refactor the API to use Amazon EFS. Implement Amazon API Gateway and enable API caching" is incorrect. You cannot refactor an API to use EFS as EFS is a file store service and must be mounted to an EC2 instance.

INCORRECT: "Export data on a daily basis and upload to Amazon S3. Refactor the API to use the S3 data. Implement Amazon API Gateway and enable API caching" is incorrect. This solution does not use constant replication for the dataset so the dataset could be out of date when it is queried.

References:

<https://aws.amazon.com/blogs/database/aws-database-migration-service-and-aws-schema-conversion-tool-now-support-ibm-db2-as-a-source/>

https://docs.aws.amazon.com/dms/latest/userguide/CHAP_Task.CDC.html

Save time with our AWS cheat sheets:

<https://digitalcloud.training/category/aws-cheat-sheets/aws-solutions-architect-professional/aws-migration-sap/>

Question 13: **Incorrect**

A Solutions Architect has been asked to implement a disaster recovery (DR) site for an eCommerce platform that is growing at an increasing rate. The platform runs on Amazon EC2 web servers behind Elastic Load Balancers, images stored in Amazon S3 and Amazon DynamoDB tables that store product and customer data. The DR site should be located in a separate AWS Region.

Which combinations of actions should the Solutions Architect take to implement the DR site? (Select THREE.)

-

Enable DynamoDB Streams and use an event-source mapping to a Lambda function which populates a table in the second Region.

-

Enable multi-Region targets on the Elastic Load Balancer and target Amazon EC2 instances in both Regions.

(Incorrect)

-

Enable versioning on the Amazon S3 buckets and enable cross-Region snapshots.

-

Enable DynamoDB global tables to achieve multi-Region table replication.

(Correct)

-

Enable Amazon Route 53 health checks to determine if the primary site is down, and route traffic to the disaster recovery site if there is an issue.

(Correct)

-

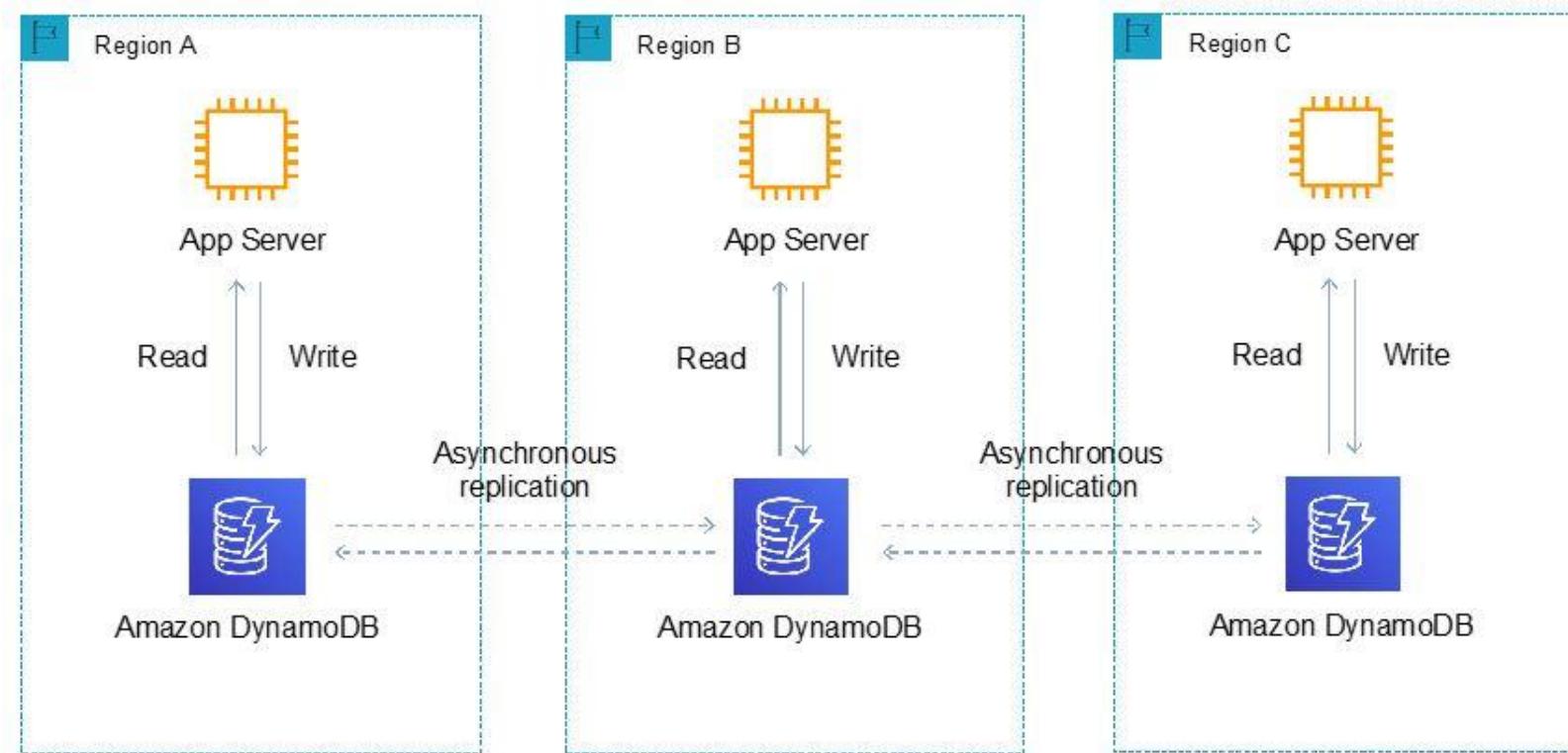
Enable Amazon S3 cross-Region replication on the buckets that contain images.

(Correct)

Explanation

To enable disaster recovery for this solution the Solutions Architect can create Amazon EC2 instances behind an ELB in a second Region and use DynamoDB Global Tables to create a multiregion, multi-active database. The failover can then be initiated by Amazon Route 53 using a failover routing policy configured for active-passive failover. This solutions meets all requirements.

A DynamoDB global table supports reads and writes in multiple Regions as can be seen in the diagram below:



CORRECT: "Enable Amazon Route 53 health checks to determine if the primary site is down, and route traffic to the disaster recovery site if there is an issue" is a correct answer.

CORRECT: "Enable Amazon S3 cross-Region replication on the buckets that contain images" is also a correct answer.

CORRECT: "Enable DynamoDB global tables to achieve multi-Region table replication" is also a correct answer.

INCORRECT: "Enable DynamoDB Streams and use an event-source mapping to a Lambda function which populates a table in the second Region" is incorrect. This is not a good method of synchronizing the data. DynamoDB global tables should be used instead.

INCORRECT: "Enable multi-Region targets on the Elastic Load Balancer and target Amazon EC2 instances in both Regions" is incorrect. You cannot have multi-Region targets with an ELB.

INCORRECT: "Enable versioning on the Amazon S3 buckets and enable cross-Region snapshots" is incorrect. There is no such thing as cross-Region snapshots with Amazon S3, snapshots are not an S3 feature.

References:

<https://docs.aws.amazon.com/amazondynamodb/latest/developerguide/GlobalTables.html>

<https://docs.aws.amazon.com/Route53/latest/DeveloperGuide/dns-failover-configuring.html>

Save time with our AWS cheat sheets:

<https://digitalcloud.training/category/aws-cheat-sheets/aws-solutions-architect-professional/aws-networking-content-delivery-sap/>

<https://digitalcloud.training/category/aws-cheat-sheets/aws-solutions-architect-professional/aws-storage-sap/>

<https://digitalcloud.training/category/aws-cheat-sheets/aws-solutions-architect-professional/aws-compute-sap/>

Question 14: **Correct**

An eCommerce company runs a successful website with a growing base of customers. The website is becoming popular internationally and demand is increasing quickly. The website is currently hosted in an on-premises data center with web servers and a MySQL database. The company plans to migrate the workloads to AWS. A Solutions Architect has been asked to create a solution that:

- Improves security

- Improves reliability
- Improves availability
- Reduces latency
- Reduces maintenance

Which combination of steps should the Solutions Architect take to meet these requirements? (Select THREE.)

-

Host static website content in Amazon S3. Use S3 Transfer Acceleration to reduce latency while serving webpages. Use AWS WAF to improve website security.

-

Migrate the database to a single-AZ Amazon RDS for MySQL DB instance.

-

Launch Amazon EC2 instances in two Availability Zones to host a highly available MySQL database cluster.

-

Create an Auto Scaling group of Amazon EC2 instances in two Availability Zones and attach an Application Load Balancer.

(Correct)

-

Host static website content in Amazon S3. Use Amazon CloudFront to reduce latency while serving webpages. Use AWS WAF to improve website security.

(Correct)



Migrate the database to an Amazon Aurora MySQL DB cluster configured for Multi-AZ.

(Correct)

Explanation

This is a simple migration to cloud that requires a standard set of security, performance, and reliability requirements. To meet these requirements an ASG should be created across multiple AZs for the web layer. This should be behind an ALB for distributing incoming connections.

For the database layer an Aurora MySQL DB cluster with an Aurora Replica in another AZ will provide Multi-AZ failover. This ensures the DB layer is highly available, and reduces maintenance.

Another way to improve performance for global users is to host static content in Amazon S3 and use the Amazon CloudFront CDN to cache the content in Edge Locations around the world. Adding AWS WAF adds additional security.

CORRECT: "Create an Auto Scaling group of Amazon EC2 instances in two Availability Zones and attach an Application Load Balancer" is a correct answer.

CORRECT: "Migrate the database to an Amazon Aurora MySQL DB cluster configured for Multi-AZ" is a correct answer.

CORRECT: "Host static website content in Amazon S3. Use Amazon CloudFront to reduce latency while serving webpages. Use AWS WAF to improve website security" is a correct answer.

INCORRECT: "Launch Amazon EC2 instances in two Availability Zones to host a highly available MySQL database cluster" is incorrect. This requires more maintenance to maintain so is not the best solution.

INCORRECT: "Host static website content in Amazon S3. Use S3 Transfer Acceleration to reduce latency while serving webpages. Use AWS WAF to improve website security" is incorrect. Transfer Acceleration is for uploading data using the CloudFront Edge network. For serving static assets use a CloudFront distribution.

INCORRECT: "Migrate the database to a single-AZ Amazon RDS for MySQL DB instance" is incorrect. This does not provide the availability required, deploying an Aurora Replica in another AZ provides high availability.

References:

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

<https://docs.aws.amazon.com/AmazonCloudFront/latest/DeveloperGuide/distribution-web-awswaf.html>

Save time with our AWS cheat sheets:

<https://digitalcloud.training/category/aws-cheat-sheets/aws-solutions-architect-professional/aws-compute-sap/>

<https://digitalcloud.training/category/aws-cheat-sheets/aws-solutions-architect-professional/aws-database-sap/>

Question 15: Incorrect

A new application will ingest millions of records per minute from user devices all over the world. Each record is less than 4 KB in size and must be stored durably and accessed with low latency. The data must be stored for 90 days after which it can be deleted. It has been estimated that storage requirements for a year will be 15-20TB.

Which storage strategy is the MOST cost-effective and meets the design requirements?

-

Store each incoming record in an Amazon DynamoDB table. Configure the DynamoDB Time to Live (TTL) feature to delete records older than 90 days.

(Correct)



Store the records in an Amazon Kinesis Data Stream. Configure the Time to Live (TTL) feature to delete records older than 90 days.



Store each incoming record in a single .csv file in an Amazon S3 bucket. Configure a lifecycle policy to delete data older than 90 days.

(Incorrect)



Store each incoming record in a single table in an Amazon RDS MySQL database. Run a nightly cron job that executes a query to delete any records older than 90 days.

Explanation

Amazon DynamoDB is a suitable data store as it can scale to the throughput required and offers low latency. The TTL feature can be used to expire data.

Amazon DynamoDB Time to Live (TTL) allows you to define a per-item timestamp to determine when an item is no longer needed. Shortly after the date and time of the specified timestamp, DynamoDB deletes the item from your table without consuming any write throughput. TTL is provided at no extra cost as a means to reduce stored data volumes by retaining only the items that remain current for your workload's needs.

The image below depicts a table with an "expiry" column that specifies the expiry date in Epoch format:

sessionid	data	date	expiry
23040210248	https://ama.	2020-04..	158415112
84324350122	https://ama.	2020-04..	158416214
923424325040	https://ama.	2020-04...	158418985

The item will be deleted when the expiry date (in Epoch format) expires

CORRECT: "Store each incoming record in an Amazon DynamoDB table. Configure the DynamoDB Time to Live (TTL) feature to delete records older than 90 days" is the correct answer.

INCORRECT: "Store each incoming record in a single .csv file in an Amazon S3 bucket. Configure a lifecycle policy to delete data older than 90 days" is incorrect. The maximum object size in S3 is 5TB. Also, you cannot expire entries within a file, only the entire file (object).

INCORRECT: "Store each incoming record in a single table in an Amazon RDS MySQL database. Run a nightly cron job that executes a query to delete any records older than 90 days" is incorrect. DynamoDB is more suitable as it has a native feature for expiring data and is better suited to this kind of workload than an SQL database.

INCORRECT: "Store the records in an Amazon Kinesis Data Stream. Configure the Time to Live (TTL) feature to delete records older than 90 days" is incorrect. There is no TTL feature in KDS so this is not possible.

References:

<https://docs.aws.amazon.com/amazondynamodb/latest/developerguide/TTL.html>

Save time with our AWS cheat sheets:

<https://digitalcloud.training/category/aws-cheat-sheets/aws-solutions-architect-professional/aws-database-sap/>

Question 16: Incorrect

A company is in the process of migrating applications to AWS using multiple accounts in AWS Organizations . The management account is at the root of the Organizations hierarchy. Business units each have different accounts and requirements for the services they need to use. The security team needs to implement controls across all accounts to prohibit many AWS services. In some cases a business unit may have a valid exception to these controls and this must be achievable.

Which solution will meet these requirements with minimal optional overhead?

-

Use an SCP in Organizations to implement an allow list of AWS services. Apply this SCP at the root level. Remove the default AWS managed SCP from the root level and all OU levels. For any specific exceptions for an OU, modify the SCP attached to that OU, and add the required AWS services to the allow list.

-

Use an SCP in Organizations to implement a deny list of AWS services. Apply this SCP at the root level and each OU. Remove the default AWS managed SCP from the root level and all OU levels. For any specific exceptions, modify the SCP attached to that OU, and add the required AWS services to the allow list.

-

Use an SCP in Organizations to implement a deny list of AWS services. Apply this SCP at each OU level. Leave the default AWS managed SCP attached to the root level and all OUs. For accounts that require specific exceptions, create an OU under root and attach an SCP that denies fewer services.

(Correct)

-

Use an SCP in Organizations to implement a deny list of AWS services. Apply this SCP at the root level. For any specific exceptions for an OU, create a new SCP for that OU and add the required AWS services to the allow list.

(Incorrect)

Explanation

AWS Organizations Service Control Policies (SCPs) can be used to control the maximum available permissions in an AWS account. The permissions flow through a hierarchy from the root to all entities beneath. With a deny list strategy a default SCP allows all services and deny lists must be implemented for any specific services that must be restricted.

In this scenario using a deny list strategy will mean the AWS managed SCP at the root level and all OUs will allow all services for all business units. An SCP can then be defined that prohibits AWS services and this can be applied to all OUs.

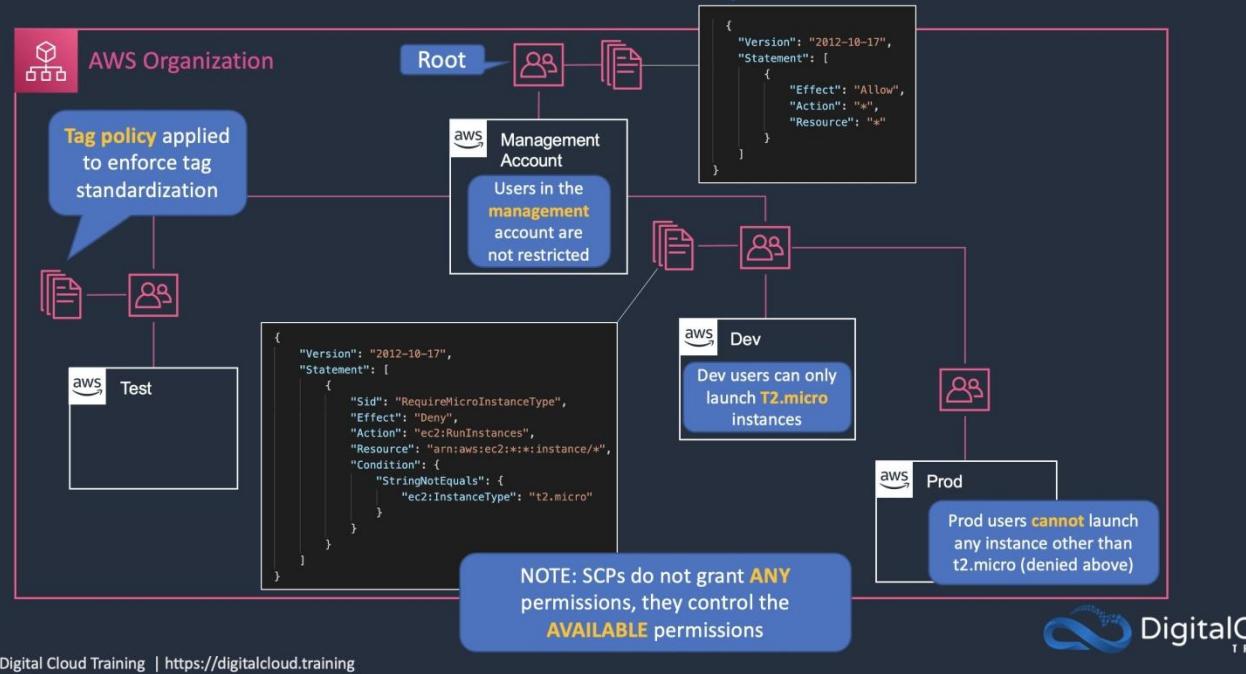
When an exception is required, an OU must be created under root with an SCP that has fewer restrictions. This OU must be under root so it is not part of the hierarchy of OUs that have denies, as a deny at a higher level always overrides an allow.

The diagram below depicts a scenario with a deny list strategy and SCPs applied to multiple OUs/accounts:



Service Control Policies

SCPs control the maximum available permissions



© Digital Cloud Training | <https://digitalcloud.training>



CORRECT: "Use an SCP in Organizations to implement a deny list of AWS services. Apply this SCP at each OU level. Leave the default AWS managed SCP at the root level. For any specific exceptions for an OU, remove the standard deny list SCP and add a new deny list SCP for that OU" is the correct answer.

INCORRECT: "Use an SCP in Organizations to implement an allow list of AWS services. Apply this SCP at the root level. Remove the default AWS managed SCP from the root level and all OU levels. For any specific exceptions for an OU, modify the SCP attached to that OU, and add the required AWS services to the allow list" is incorrect. You cannot modify an SCP that is applied to multiple OUs at one child OU. It must be edited in one place.

INCORRECT: "Use an SCP in Organizations to implement a deny list of AWS services. Apply this SCP at the root level. For any specific exceptions for an OU, create a new SCP for that OU and add the required AWS services to the allow list" is incorrect. This will not work as there is an explicit deny at a higher level and this will override an allow at any level beneath.

INCORRECT: "Use an SCP in Organizations to implement a deny list of AWS services. Apply this SCP at the root level and each OU. Remove the default AWS managed SCP from the root level and all OU levels. For any specific exceptions, modify the SCP attached to that OU, and add the required AWS services to the allow list" is incorrect. You cannot modify an SCP that is applied to multiple OUs at one child OU. It must be edited in one place. This will also not work as there is an explicit deny at a higher level and this will override an allow at any level beneath.

References:

https://docs.aws.amazon.com/organizations/latest/userguide/orgs_manage_policies_inheritance_auth.html

Save time with our AWS cheat sheets:

<https://digitalcloud.training/category/aws-cheat-sheets/aws-solutions-architect-professional/aws-management-tools-sap/>

Question 17: **Incorrect**

An advertising company hosts static content in an Amazon S3 bucket that is served by Amazon CloudFront. The static content is generated programmatically from a Development account, and the S3 bucket and CloudFront are in a Production account. The build pipeline uploads the files to Amazon S3 using an IAM role in the Development Account. The S3 bucket has a bucket policy that only allows CloudFront to read objects using an origin access identity (OAI). During testing all attempts to upload objects using the to the S3 bucket are denied..

How can a Solutions Architect resolve this issue and allow the objects to be uploaded to Amazon S3?

-

Modify the S3 upload process in the Development account to add the bucket-owner-full-control ACL to the objects at upload.

-

Create a new cross-account IAM role in the Production account with write access to the S3 bucket. Modify the build pipeline to assume this role to upload the files to the Production Account.

(Correct)



Create a new IAM role in the Development account with read access to the S3 bucket. Configure S3 to use this new role as its OAI. Modify the build pipeline to assume this role when uploading files from the Development Account.

(Incorrect)



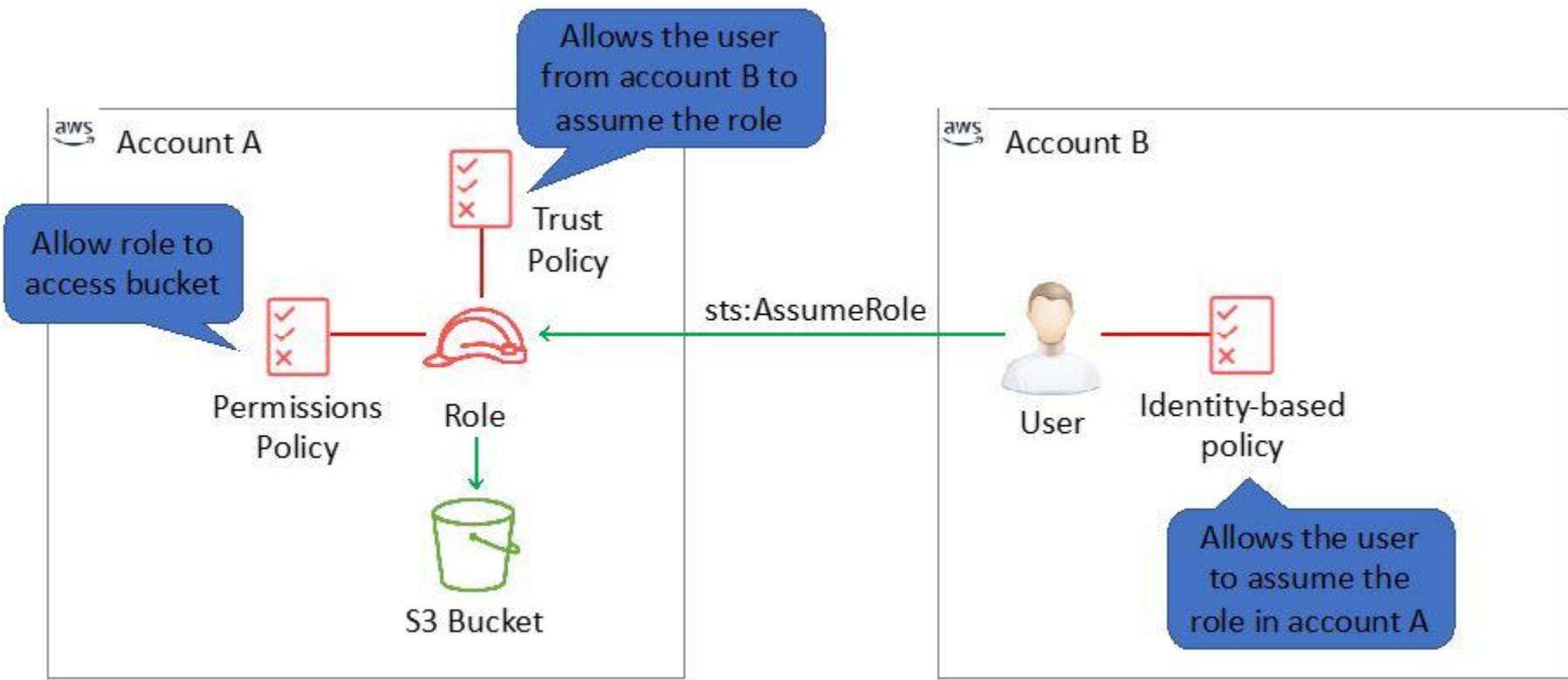
Modify the S3 upload process in the Development account to set the object owner to the Production Account.

Explanation

In the current configuration the S3 bucket will not allow any uploads. The only way to access the bucket is by using the OAI through CloudFront and this method only allows read access.

To be able to access the bucket directly and upload objects, a new IAM role can be created in the Production account with the necessary permissions to access the bucket. The role can then be assumed by the build pipeline using cross-account access.

The diagram below depicts how a user in one account can assume a role in a second account. In this case of the scenario in this question, rather than a user assuming the role a service principal can assume the role.



CORRECT: "Create a new cross-account IAM role in the Production account with write access to the S3 bucket. Modify the build pipeline to assume this role to upload the files to the Production Account" is the correct answer.

INCORRECT: "Create a new IAM role in the Development account with read access to the S3 bucket. Configure S3 to use this new role as its OAI. Modify the build pipeline to assume this role when uploading files from the Development Account" is incorrect. You cannot use a cross-account role for an OAI. An OAI is a special user account that is associated with a CloudFront distribution.

INCORRECT: "Modify the S3 upload process in the Development account to add the bucket-owner-full-control ACL to the objects at upload" is incorrect. You cannot add permissions when uploading an object. The permissions must be previously specified on the bucket.

INCORRECT: "Modify the S3 upload process in the Development account to set the object owner to the Production Account" is incorrect. You cannot modify bucket ownership as part of an object upload.

References:

https://docs.aws.amazon.com/IAM/latest/UserGuide/tutorial_cross-account-with-roles.html

Save time with our AWS cheat sheets:

<https://digitalcloud.training/category/aws-cheat-sheets/aws-solutions-architect-professional/security-identity-compliance-sap/>

Question 18: **Correct**

A company has experienced issues updating an AWS Lambda function that is deployed using an AWS CloudFormation stack. The issues have resulted in outages that affected large numbers of customers. A Solutions Architect must adjust the deployment process to support a canary release strategy. Invocation traffic should be routed based on specified weights.

Which solution will meet these requirements?

-

Deploy the application into a new CloudFormation stack. Use an Amazon Route 53 weighted routing policy to distribute the load.

-

Create a version for every new update to the Lambda function code. Use the AWS CLI update-function-configuration command with the routing-config parameter to distribute the load.

-

Create an alias for new versions of the Lambda function. Use the AWS CLI update-alias command with the routing-config parameter to distribute the load.

(Correct)



Use AWS CodeDeploy to deploy using the `CodeDeployDefault.HalfAtATime` deployment configuration to distribute the load.

Explanation

With the introduction of alias traffic shifting, it is now possible to trivially implement canary deployments of Lambda functions. By updating additional version weights on an alias, invocation traffic is routed to the new function versions based on the weight specified.

Detailed CloudWatch metrics for the alias and version can be analyzed during the deployment, or other health checks performed, to ensure that the new version is healthy before proceeding.

The following example AWS CLI command points an alias to a new version, weighted at 5% (original version at 95% of traffic):

```
aws lambda update-alias --function-name myfunction --name myalias --routing-config '{"AdditionalVersionWeights" : {"2" : 0.05} }'
```

CORRECT: "Create an alias for new versions of the Lambda function. Use the AWS CLI update-alias command with the routing-config parameter to distribute the load" is the correct answer.

INCORRECT: "Create a version for every new update to the Lambda function code. Use the AWS CLI update-function-configuration command with the routing-config parameter to distribute the load" is incorrect. A Lambda alias should be used to point to a new version. A Lambda alias is like a pointer to a specific function version. Users can access the function version using the alias Amazon Resource Name (ARN). The alias is used in the CLI command. This method is used as you can change the version that is associated with the alias easily.

INCORRECT: "Deploy the application into a new CloudFormation stack. Use an Amazon Route 53 weighted routing policy to distribute the load" is incorrect. Route 53 is not a method of doing a canary release, the solution should use Lambda aliases with the CLI.

INCORRECT: "Use AWS CodeDeploy to deploy using the `CodeDeployDefault.HalfAtATime` deployment configuration to distribute the load" is incorrect. This CodeDeploy configuration is used for in-place and blue/green migrations but not for canary releases.

References:

<https://aws.amazon.com/blogs/compute/implementing-canary-deployments-of-aws-lambda-functions-with-alias-traffic-shifting/>

Save time with our AWS cheat sheets:

<https://digitalcloud.training/category/aws-cheat-sheets/aws-solutions-architect-professional/aws-compute-sap/>

Question 19: **Correct**

A financial company processes transactions using on-premises application servers which save output to an Amazon DynamoDB table. The company's data center is connected to AWS using an AWS Direct Connect (DX) connection. Company managed has mandated that the solution should be available across multiple Regions. Consistent network performance must be maintained at all times.

What changes should the company make to meet these requirements?



Use an AWS managed VPN to connect to a second AWS Region. Create a copy of the DynamoDB table in the second Region. Enable DynamoDB streams in the primary Region and use AWS DMS to synchronize data to the copied table.



Create a DX connection to a second AWS Region. Use DynamoDB global tables to replicate data to the second Region. Modify the application to fail over to the second Region.

(Correct)



Create a DX connection to a second AWS Region. Create an identical DynamoDB table in the second Region. Enable DynamoDB auto scaling to manage throughput capacity. Modify the application to write to the second Region.



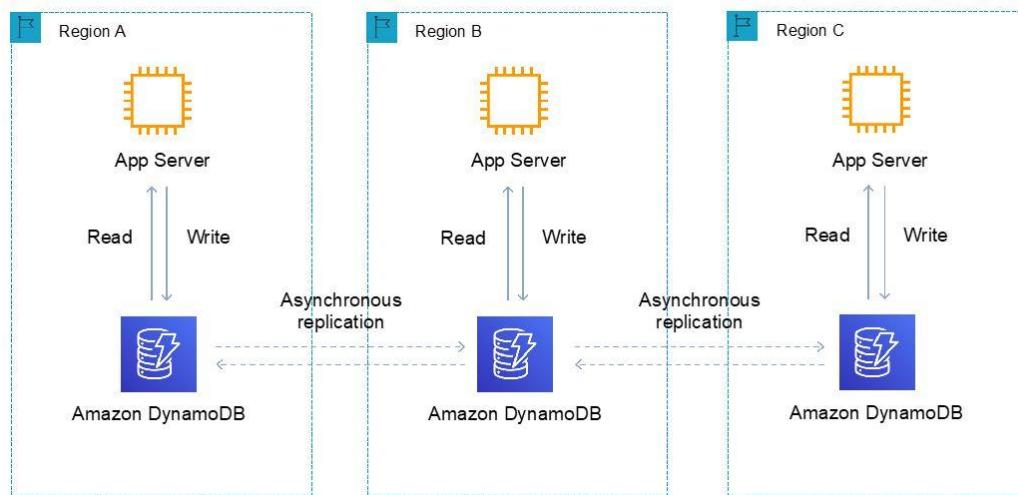
Use an AWS managed VPN to connect to a second AWS Region. Create a copy of the DynamoDB table in the second Region. Enable DynamoDB streams in the primary Region and use AWS Lambda to synchronize data to the copied table.

Explanation

To ensure consistent network performance and AWS Direct Connect connection must be used as AWS Managed VPN relies on the public internet which cannot offer consistent performance.

Amazon DynamoDB Global Tables is a fully managed, multi-region, multi-active database. This means you can read and write to multiple Regions. In the event of the failure of a Region the application logic must be set to fail to an endpoint in another Region where a replica table is running.

The diagram below depicts how DynamoDB replicates data between Regions and accepts reads and writes:



CORRECT: "Create a DX connection to a second AWS Region. Use DynamoDB global tables to replicate data to the second Region. Modify the application to fail over to the second Region" is the correct answer.

INCORRECT: "Use an AWS managed VPN to connect to a second AWS Region. Create a copy of the DynamoDB table in the second Region. Enable DynamoDB streams in the primary Region and use AWS Lambda to synchronize data to the copied table" is incorrect. A VPN does not offer consistent performance and DynamoDB global tables should be used for a multi-active database.

INCORRECT: "Create a DX connection to a second AWS Region. Create an identical DynamoDB table in the second Region. Enable DynamoDB auto scaling to manage throughput capacity. Modify the application to write to the second Region" is incorrect. This does not offer any solution for creating a synchronized copy of the database in a second Region.

INCORRECT: "Use an AWS managed VPN to connect to a second AWS Region. Create a copy of the DynamoDB table in the second Region. Enable DynamoDB streams in the primary Region and use AWS DMS to synchronize data to the copied table" is incorrect. A VPN does not offer consistent performance and AWS DMS should be replaced with DynamoDB global tables.

References:

https://docs.aws.amazon.com/amazondynamodb/latest/developerguide/V2globaltables_HowItWorks.html

Save time with our AWS cheat sheets:

<https://digitalcloud.training/category/aws-cheat-sheets/aws-solutions-architect-professional/aws-networking-content-delivery-sap/>

Question 20: **Correct**

A security team uses a ticketing system to capture suspicious events that require investigation. The security team has created a system where events are captured using CloudTrail Logs and saved to Amazon S3. A scheduled AWS Lambda function then uses Amazon Athena to query the logs for any API actions performed by the root user. The results are then submitted to the ticketing system by the Lambda function.

The ticketing system has a monthly 4-hour maintenance window when the system is offline and cannot log new tickets and an audit revealed that several tickets were not created due to the ticketing system being unavailable.

Which combination of steps should a solutions architect take to ensure that the incidents are reported to the ticketing system even during planned maintenance? (Select TWO.)

-

Create an Amazon EventBridge rule with a pattern that looks for AWS CloudTrail events where the API calls involve the root user account. Configure an Amazon SQS queue as a target for the rule.

(Correct)

-

Update the Lambda function to be triggered by messages published to an Amazon SNS topic. Update the existing application code to retry every 5 minutes if the ticketing system's API endpoint is unavailable.

-

Create an Amazon SNS topic to which Amazon CloudWatch alarms will be published. Configure a CloudWatch alarm to invoke the Lambda function.

-

Create an Amazon SQS queue to which Amazon CloudWatch alarms will be published. Configure a CloudWatch alarm to publish to the SQS queue.

-

Update the Lambda function to poll the Amazon SQS queue for messages and to return successfully when the ticketing system API has processed the request.

(Correct)

Explanation

The existing system can be modified to use Amazon EventBridge instead of using AWS CloudTrail with Amazon Athena. Eventbridge can be configured with a rule that checks all AWS API calls via CloudTrail. The rule can be configured to look for the usage or the root user account. Eventbridge can then be configured with an Amazon SQS queue as a target that puts a message in the queue waiting to be processed.

The Lambda function can then be configured to poll the queue for messages (event-source mapping), process the event synchronously and only return a successful result when the ticketing system has processed the request. The message will be deleted only if the result is successful, allowing for retries.

This system will ensure that the important events are not missed when the ticketing system is unavailable.

CORRECT: "Create an Amazon EventBridge rule with a pattern that looks for AWS CloudTrail events where the API calls involve the root user account. Configure an Amazon SQS queue as a target for the rule" is a correct answer.

CORRECT: "Update the Lambda function to poll the Amazon SQS queue for messages and to return successfully when the ticketing system API has processed the request" is also a correct answer.

INCORRECT: "Update the Lambda function to be triggered by messages published to an Amazon SNS topic. Update the existing application code to retry every 5 minutes if the ticketing system's API endpoint is unavailable" is incorrect. SNS does not store messages so if the Lambda function is unable to successfully process the message it will be lost.

INCORRECT: "Create an Amazon SNS topic to which Amazon CloudWatch alarms will be published. Configure a CloudWatch alarm to invoke the Lambda function" is incorrect. SQS should be used instead of SNS as SNS will not store messages for subsequent attempts at processing.

INCORRECT: "Create an Amazon SQS queue to which Amazon CloudWatch alarms will be published. Configure a CloudWatch alarm to publish to the SQS queue" is incorrect. EventBridge rules should be used (CloudWatch Events) as they can look for patterns in events. CloudWatch alarms will alarm based on metrics.

References:

<https://docs.aws.amazon.com/lambda/latest/dg/with-sqs.html>

Save time with our AWS cheat sheets:

<https://digitalcloud.training/category/aws-cheat-sheets/aws-solutions-architect-professional/aws-application-integration-sap/>

<https://digitalcloud.training/category/aws-cheat-sheets/aws-solutions-architect-professional/aws-management-tools-sap/>

Question 21: **Correct**

An agricultural company is rolling out thousands of devices that will send environmental data to a data platform. The platform will process and analyze the data and provide information back to researchers. The devices will send 8 KB of data every second and the solution must support near real-time analytics, provide durability for the data, and deliver results to a data warehouse.

Which strategy should a solutions architect use to meet these requirements?

-
- Use an Amazon API Gateway to put requests into an Amazon SQS queue, analyze the data with an AWS Lambda function, and save the results to an Amazon Redshift cluster using Amazon EMR.**
-
- Use Amazon Kinesis Data Streams to collect the inbound data, analyze the data with Kinesis clients, and save the results to an Amazon Redshift cluster using Amazon EMR.**

(Correct)

-
- Use Amazon S3 to collect the inbound device data, analyze the data from Amazon SQS with Kinesis, and save the results to an Amazon Redshift cluster.**
-

Use Amazon Kinesis Data Firehose to collect the inbound sensor data, analyze the data with Kinesis clients, and save the results to an Amazon RDS instance.

Explanation

The solution must support near real-time analytics. For this Amazon Kinesis data streams can be used with clients processing and analyzing the data using Amazon EMR. The solution must also deliver results to a data warehouse and Amazon RedShift is ideal for this purpose.

CORRECT: "Use Amazon Kinesis Data Streams to collect the inbound data, analyze the data with Kinesis clients, and save the results to an Amazon Redshift cluster using Amazon EMR" is the correct answer.

INCORRECT: "Use Amazon Kinesis Data Firehose to collect the inbound sensor data, analyze the data with Kinesis clients, and save the results to an Amazon RDS instance" is incorrect. Firehose does not use Kinesis clients; it loads data directly to a destination.

INCORRECT: "Use Amazon S3 to collect the inbound device data, analyze the data from Amazon SQS with Kinesis, and save the results to an Amazon Redshift cluster" is incorrect. Amazon S3 should not be used for near real-time ingestion of streaming data on this scale. Amazon Kinesis a better fit for this use case. Analyzing with Kinesis from SQS does not make sense either.

INCORRECT: "Use an Amazon API Gateway to put requests into an Amazon SQS queue, analyze the data with an AWS Lambda function, and save the results to an Amazon Redshift cluster using Amazon EMR" is incorrect. API Gateway should not be used for streaming data and cannot directly put data into an SQS queue.

References:

<https://aws.amazon.com/kinesis/data-streams/getting-started/>

<https://docs.aws.amazon.com/streams/latest/dev/shared-throughput-kcl-consumers.html>

<https://aws.amazon.com/emr/>

<https://aws.amazon.com/redshift/>

Save time with our AWS cheat sheets:

<https://digitalcloud.training/category/aws-cheat-sheets/aws-solutions-architect-professional/aws-analytics-sap/>

<https://digitalcloud.training/category/aws-cheat-sheets/aws-solutions-architect-professional/aws-database-sap/>

Question 22: **Incorrect**

A data hosting company has developed a new application which works on a custom TCP port. The service must use fixed address assignments so other companies can whitelist the addresses in their firewalls. The application will be hosted on the publicly accessible DNS domain name cloud.myservice.com. The solution must offer high availability and redundancy across Availability Zones in a single AWS Region.

Which solution will meet these requirements?



Create an Amazon ECS cluster and a service definition for the application. Create and assign public IP address for each host in the cluster. Create an Application Load Balancer (ALB) and expose the static TCP port. Create a target group and assign the ECS service definition name to the ALB. Create a new CNAME record set and associate the public IP addresses to the record set. Provide the Elastic IP addresses of the Amazon EC2 instances to the other companies to add to their allow lists.

(Incorrect)

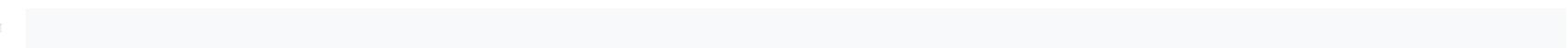


Create an Amazon ECS cluster and a service definition for the application. Create and assign public IP addresses for the ECS cluster. Create a Network Load Balancer (NLB) and expose the TCP port. Create a target group and assign the ECS cluster name to the NLB. Create a new A record set named cloud.myservice.com and assign the public IP addresses of the ECS cluster to the record set. Provide the public IP addresses of the ECS cluster to the other companies to add to their allow lists.



Create Amazon EC2 instances for the service. Create one Elastic IP address for each Availability Zone. Create a Network Load Balancer (NLB) and expose the assigned TCP port. Assign the Elastic IP addresses to the NLB for each Availability Zone. Create a target group and register the EC2 instances with the NLB. Create a new A (alias) record set named cloud.myservice.com and assign the NLB DNS name to the record set.

(Correct)



Create Amazon EC2 instances with an Elastic IP address for each instance. Create a Network Load Balancer (NLB) and expose the static TCP port. Register EC2 instances with the NLB. Create a new name server record set named cloud.myservice.com and assign the Elastic IP addresses of the EC2 instances to the record set. Provide the Elastic IP addresses of the EC2 instances to the other companies to add to their allow lists.

Explanation

Below is the sequence of steps to achieve the above:

- Create Amazon EC2 instances for the service.
- Create one Elastic IP address for each Availability Zone.
- Create a Network Load Balancer (NLB) and expose the assigned TCP port.
- Assign the Elastic IP addresses to the NLB for each Availability Zone.
- Create a target group and register the EC2 instances with the NLB.
- Create a new A (alias) record set named cloud.myservice.com and assign the NLB DNS name to the record set.

This solution exposes static public IP addresses that can be whitelisted in the firewalls of the clients. The solution also offers high availability across AZs within an AWS Region.

CORRECT: "Create Amazon EC2 instances for the service. Create one Elastic IP address for each Availability Zone. Create a Network Load Balancer (NLB) and expose the assigned TCP port. Assign the Elastic IP addresses to the NLB for each Availability Zone. Create a target group and register the EC2 instances with the NLB. Create a new A (alias) record set named cloud.myservice.com and assign the NLB DNS name to the record set" is the correct answer (as explained above.)

INCORRECT: "Create Amazon EC2 instances with an Elastic IP address for each instance. Create a Network Load Balancer (NLB) and expose the static TCP port. Register EC2 instances with the NLB. Create a new name server record set named cloud.myservice.com and assign the Elastic IP addresses of the EC2 instances to the record set. Provide the Elastic IP addresses of the EC2 instances to the other companies to add to their allow lists" is incorrect.

When creating NLB, you need to pick the available Elastic IP address which was initially allocated as mentioned in the question above. This can be one of three options:

- Amazon's pool of IPv4 addresses—if you want an IPv4 address to be allocated from Amazon's pool of IPv4 addresses.
- Public IPv4 address that you bring to your AWS account—if you want to allocate an IPv4 address from an IP address pool that you have brought to your AWS account. This option is disabled if you do not have any IP address pools.
- Customer owned pool of IPv4 addresses—if you want to allocate an IPv4 address from a pool created from your on-premises network for use with an AWS Outpost. This option is disabled if you do not have an AWS Outpost.

INCORRECT: "Create an Amazon ECS cluster and a service definition for the application. Create and assign public IP addresses for the ECS cluster. Create a Network Load Balancer (NLB) and expose the TCP port. Create a target group and assign the ECS cluster name to the NLB. Create a new A record set named my.service.com and assign the public IP addresses of the ECS cluster to the record set. Provide the public IP addresses of the ECS cluster to the other companies to add to their allow list" is incorrect.

To have an ECS cluster behind NLB as a target group, ECS has dynamic port mapping which binds instances dynamically. Binding a cluster name directly behind NLB is not an option.

INCORRECT: "Create an Amazon ECS cluster and a service definition for the application. Create and assign public IP address for each host in the cluster. Create an Application Load Balancer (ALB) and expose the static TCP port. Create a target group and assign the ECS service definition name to the ALB. Create a new CNAME record set and associate the public IP addresses to the record set. Provide the Elastic IP addresses of the Amazon EC2 instances to the other companies to add to their allow lists" is incorrect.

An ALB cannot be used as it only supports HTTP and HTTPS protocols and you cannot configure a listener with a custom TCP port.

References:

<https://docs.aws.amazon.com/elasticloadbalancing/latest/network/create-network-load-balancer.html>

<https://aws.amazon.com/premiumsupport/knowledge-center/elb-attach-elastic-ip-to-public-nlb>

Save time with our AWS cheat sheets:

<https://digitalcloud.training/aws-elastic-load-balancing-aws-elb/>

<https://digitalcloud.training/amazon-route-53/>

Question 23: **Correct**

A financial services company runs an application that allows traders to perform online simulations of market conditions. The backend runs on a fleet of virtual machines in an on-premises data center and the business logic is exposed using a REST API with multiple functions. The trader's session data is stored in a NAS file system in the on-premises data center. During busy periods of the day the server capacity is insufficient and latency issues have occurred when fetching the session data for a simulation.

A Solutions Architect must create a design for moving the application to AWS. The design must use the same API model but should be capable of scaling for the variable load and ensure access to session data is provided with low-latency.

Which solutions meets these requirements?

-

Implement the REST API using a Network Load Balancer (NLB). Run the business logic on an Amazon EC2 instance behind the NLB. Store trader session data in Amazon Aurora Serverless.

-

Implement the REST API using an Application Load Balancer (ALB). Run the business logic in AWS Lambda. Store trader session data in Amazon DynamoDB with on-demand capacity.

-

Implement the REST API using Amazon API Gateway. Run the business logic in AWS Lambda. Store trader session data in Amazon DynamoDB with on-demand capacity.

(Correct)

-

Implement the REST API using AWS AppSync. Run the business logic in AWS Lambda. Store trader session data in Amazon Aurora Serverless.

Explanation

Amazon API Gateway can be used to preserve the API model used in the current application while moving to a serverless technology. The business logic should be moved to AWS Lambda. This will ensure scalability for periods of high demand. Amazon DynamoDB is often used for storing session data. It is a low-latency NoSQL database that is well suited for this use case. This solution meets all of the stated requirements.

CORRECT: "Implement the REST API using Amazon API Gateway. Run the business logic in AWS Lambda. Store trader session data in Amazon DynamoDB with on-demand capacity" is the correct answer.

INCORRECT: "Implement the REST API using a Network Load Balancer (NLB). Run the business logic on an Amazon EC2 instance behind the NLB. Store trader session data in Amazon Aurora Serverless" is incorrect. NLBs do not provide REST APIs so you could not use an NLB to process API requests and forward them to the business logic instances.

INCORRECT: "Implement the REST API using an Application Load Balancer (ALB). Run the business logic in AWS Lambda. Store trader session data in Amazon DynamoDB with on-demand capacity" is incorrect. ALBs do not provide REST APIs so you could not use an ALB to process API requests and forward them to the business logic function.

INCORRECT: "Implement the REST API using AWS AppSync. Run the business logic in AWS Lambda. Store trader session data in Amazon Aurora Serverless" is incorrect. AppSync is a GraphQL service that is useful for use cases where real-time updates over WebSockets are required. API Gateway is a better fit for this use case to preserve the API model.

References:

<https://docs.aws.amazon.com/aws-sdk-php/v2/guide/feature-dynamodb-session-handler.html>

Save time with our AWS cheat sheets:

<https://digitalcloud.training/category/aws-cheat-sheets/aws-solutions-architect-professional/aws-networking-content-delivery-sap/>

Question 24: **Incorrect**

A company runs its IT services from an on-premises data center and is moving to AWS. The company wants to move their development and deployment processes to use managed services where possible. They would like to leverage their existing Chef tools and experience. The application must be deployed to a staging environment and then to production. The ability to roll back quickly must be available in case issues occur following a production deployment.

Which AWS service and deployment strategy should a Solutions Architect use to meet the company's requirements?



Use AWS OpsWorks and deploy the application using a blue/green deployment strategy.

(Correct)



Use AWS CodeDeploy and deploy the application using an in-place update deployment strategy.



Use AWS OpsWorks and deploy the application using a canary deployment strategy.



Use AWS Elastic Beanstalk and deploy the application using a rolling update deployment strategy.

(Incorrect)

Explanation

There is only one AWS service in the options presented that can allow the company to leverage their existing investments in Chef. AWS OpsWorks for Chef Automate is a fully managed configuration management service that hosts Chef Automate, a suite of automation tools from Chef for configuration management, compliance and security, and continuous deployment.

The next requirement is to choose the correct deployment strategy and two options are presented for AWS OpsWorks: canary and blue/green. A canary deployment is not supported by AWS OpsWorks so the only option that will work is blue/green. This option will support a fast rollback if issues occur.

CORRECT: "Use AWS OpsWorks and deploy the application using a blue/green deployment strategy" is the correct answer.

INCORRECT: "Use AWS OpsWorks and deploy the application using a canary deployment strategy" is incorrect. As mentioned above, a canary deployment is not supported for AWS OpsWorks.

INCORRECT: "Use AWS Elastic Beanstalk and deploy the application using a rolling update deployment strategy" is incorrect. Elastic Beanstalk will not enable the company to use their existing Chef recipes. Also, it is harder to quickly roll back using this deployment strategy (roll back is per batch).

INCORRECT: "Use AWS CodeDeploy and deploy the application using an in-place update deployment strategy" is incorrect. CodeDeploy is used for deploying updates, but does not provide a solution for hosting the application and using Chef recipes. The in-place update strategy also does not allow for quick rollback.

References:

<https://aws.amazon.com/opsworks/chefautomate/>

<https://docs.aws.amazon.com/opsworks/latest/userguide/best-deploy.html>

Save time with our AWS cheat sheets:

<https://digitalcloud.training/category/aws-cheat-sheets/aws-solutions-architect-professional/aws-management-tools-sap/>

Question 25: **Incorrect**

A university is running computational algorithms that require large amounts of compute power. The algorithms are being run using a high-performance compute cluster on Amazon EC2 Spot instances. Each time an instance launches a DNS record must be created in an Amazon Route 53 private hosted zone. When the instance is terminated the DNS record must be deleted.

The current configuration uses an Amazon CloudWatch Events rule that triggers an AWS Lambda function to create the DNS record. When scaling the solution to thousands of instances the university has experienced "HTTP 400 error (Bad request)" errors in the Lambda logs. The response header also includes a status code element with a value of "Throttling" and a status message element with a value of "Rate exceeded".

Which combination of steps should the Solutions Architect take to resolve these issues? (Select THREE.)

-

Configure an Amazon Kinesis data stream and configure a CloudWatch Events rule to use this queue as a target. Remove the Lambda target from the CloudWatch Events rule.

-

Update the CloudWatch Events rule to trigger on Amazon EC2 "Instance Launch Successful" and "Instance Terminate Successful" events for the Auto Scaling group used by the cluster.

(Correct)

-

Configure an Amazon SQS FIFO queue and configure a CloudWatch Events rule to use this queue as a target. Remove the Lambda target from the CloudWatch Events rule.

(Incorrect)



Configure a Lambda function to read data from the Amazon Kinesis data stream and configure the batch window to 5 minutes. Modify the function to make a single API call to Amazon Route 53 with all records read from the Kinesis data stream.

(Incorrect)



Configure a Lambda function to retrieve messages from an Amazon SQS queue. Modify the Lambda function to retrieve a maximum of 10 messages then batch the messages by Amazon Route 53 API call type and submit. Delete the messages from the SQS queue after successful API calls.

(Correct)



Configure an Amazon SQS standard queue and configure the existing CloudWatch Events rule to use this queue as a target. Remove the Lambda target from the CloudWatch Events rule.

(Correct)

Explanation

The errors in the Lambda logs indicate that throttling is occurring. Throttling is intended to protect your resources and downstream applications. Though Lambda automatically scales to accommodate incoming traffic, functions can still be throttled for various reasons.

In this case it is most likely that the throttling is not occurring in Lambda itself but in API calls made to Amazon Route 53. In Route 53 you are limited (by default) to five requests per second per AWS account. If you submit more than five requests per second, Amazon Route 53 returns an HTTP 400 error (Bad request). The response header also includes a Code element with a value of Throttling and a Message element with a value of Rate exceeded.

The resolution here is to place the data for the DNS records into an SQS queue where they can buffer. AWS Lambda can then poll the queue and process the messages, making sure to batch the messages to reduce the likelihood of receiving more errors.

CORRECT: "Update the CloudWatch Events rule to trigger on Amazon EC2 "Instance Launch Successful" and "Instance Terminate Successful" events for the Auto Scaling group used by the cluster" is a correct answer.

CORRECT: "Configure a Lambda function to retrieve messages from an Amazon SQS queue. Modify the Lambda function to retrieve a maximum of 10 messages then batch the messages by Amazon Route 53 API call type and submit. Delete the messages from the SQS queue after successful API calls" is also a correct answer.

CORRECT: "Configure an Amazon SQS standard queue and configure the existing CloudWatch Events rule to use this queue as a target. Remove the Lambda target from the CloudWatch Events rule" is also a correct answer.

INCORRECT: "Configure an Amazon SQS FIFO queue and configure a CloudWatch Events rule to use this queue as a target. Remove the Lambda target from the CloudWatch Events rule" is incorrect. A FIFO queue is not necessary and may not support the rate of messages (up to 3,000 messages per second).

INCORRECT: "Configure an Amazon Kinesis data stream and configure a CloudWatch Events rule to use this queue as a target. Remove the Lambda target from the CloudWatch Events rule" is incorrect. As this is a decoupling use case and SQS queue is a better fit.

INCORRECT: "Configure a Lambda function to read data from the Amazon Kinesis data stream and configure the batch window to 5 minutes. Modify the function to make a single API call to Amazon Route 53 with all records read from the Kinesis data stream" is incorrect. For decoupling use cases SQS is a better fit than using Kinesis.

References:

<https://docs.aws.amazon.com/Route53/latest/DeveloperGuide/DNSLimitations.html>

Save time with our AWS cheat sheets:

<https://digitalcloud.training/category/aws-cheat-sheets/aws-solutions-architect-professional/aws-compute-sap/>

<https://digitalcloud.training/category/aws-cheat-sheets/aws-solutions-architect-professional/aws-application-integration-sap/>

<https://digitalcloud.training/category/aws-cheat-sheets/aws-solutions-architect-professional/aws-networking-content-delivery-sap/>

Question 26: **Incorrect**

A company runs a traffic sensor related IoT platform on AWS. Applications are hosted on EC2 instances and receive sensor data containing traffic information in real time. Applications are written in Node.js and have an Application Load Balancer in front. The backend includes an Amazon RDS MySQL DB instance that uses a 4 TB General Purpose SSD volume.

The company want to deploy the application to a much larger number of sensors. During initial testing the API servers were consistently overloaded and RDS metrics showed high write latency.

Which of the following steps together will resolve the issues permanently and enable growth as new sensors are provisioned, while keeping this platform cost-efficient? (Select TWO.)

-

Leverage Amazon Kinesis Data Streams and AWS Lambda to ingest and process the raw data.

(Correct)

-

Increase the MySQL General Purpose SSD storage to 6 TB to improve the volume's IOPS.

(Incorrect)

-

Use AWS X-Ray to analyze and debug application issues and add more EC2 instances to match the load.

-

Re-architect the database tier to use Amazon DynamoDB instead of an RDS MySQL DB instance.

(Correct)



Re-architect the database tier to use Amazon Aurora instead of an RDS MySQL DB instance and add read replicas.

(Incorrect)

Explanation

Kinesis Data Streams is ideal for ingesting sensor data from IoT platforms and can use AWS Lambda functions for processing the data. This is the best way to ingest and analyze real time data.

In place of RDS, DynamoDB provides much higher throughput and scalability, though this involves one time effort to refactor, this modification can resolve the write latency for any future roll out since DynamoDB is able to scale for any volume of data.

CORRECT: "Leverage Amazon Kinesis Data Streams and AWS Lambda to ingest and process the raw data" is a correct answer (as explained above.)

CORRECT: "Re-architect the database tier to use Amazon DynamoDB instead of an RDS MySQL DB instance" is also a correct answer (as explained above.)

INCORRECT: "Increase the MySQL General Purpose SSD storage to 6 TB to improve the volume's IOPS" is incorrect.

Increasing the database size might temporarily fix the issue but does not ensure that issue won't reoccur in future. Hence, this option is incorrect.

INCORRECT: "Re-architect the database tier to use Amazon Aurora instead of an RDS MySQL DB instance and add read replicas" is incorrect.

Read replicas won't help in this scenario as the issues described relate to write latency, not read latency.

INCORRECT: "Use AWS X-Ray to analyze and debug application issues and add more EC2 instances to match the load" is incorrect.

X-Ray is used for tracing application performance which may be useful. However, adding EC2 instances is not the best solution it would be better to use KDS and Lambda which are better suited to ingesting and processing streaming data.

References:

<https://aws.amazon.com/kinesis/data-streams/>

<https://aws.amazon.com/blogs/database/how-to-determine-if-amazon-dynamodb-is-appropriate-for-your-needs-and-then-plan-your-migration/>

Save time with our AWS cheat sheets:

<https://digitalcloud.training/aws-application-integration-services/>

Question 27: **Correct**

A company uses Amazon DynamoDB as the backend for the development environment of a new serverless application. While benchmarking the load, they have configured the RCU and WCU for DynamoDB based on the maximum anticipated load for peak usage.

Peak usage runs over several hours each weekend and is twice the usual load across the week. Within this duration, write operations are significant and take up most of the traffic.

The company must optimize the cost of running the application before releasing to production. Which solution will meet these requirements?

-

Configure on-demand capacity mode for the table and configure DynamoDB Accelerator (DAX) in front of the table.

-

Configure on-demand capacity mode for the table to enable pay-per-request pricing for read and write requests.

(Correct)



Purchase reserved RCU and WCU for the DynamoDB table and use AWS Application Auto scaling to match the increased load during the peak usage period.



Reduce the provisioned read capacity to match the new peak load on the table and configure DynamoDB Accelerator (DAX) in front of the table.

Explanation

With a manually set scaling policy, you can set the upper and low limits of the scaling. So, for example, you could set it at a minimum of 1 and a max of 20 read capacity. This would mean that no matter what you would always have at least 1 read capacity but never more than 20.

So now if you receive that same spike but have the max throughput to 20, DynamoDB will throttle those requests and prevent you from going over your max auto scaled threshold.

With On-Demand Scaling you don't need to think about provisioning throughput. Instead, your table will scale all behind the scenes automatically. Extreme spikes in load can occur and be handled seamlessly by AWS. This also brings a change in the cost structure for DynamoDB. With On-Demand Scaling, instead of paying for provisioned throughput, you pay per read or write request.

Remember for Provisioned with Auto-Scaling you are basically paying for throughput 24/7. Whereas for On-Demand Scaling you pay per request. This means for applications still in development or low traffic applications, it might be more economical to use On-Demand Scaling and not worry about provisioning throughput. However, at scale, this can quickly shift once you have a more consistent usage pattern.

Provisioned with Auto-Scaling	On-Demand Scaling
Predictable, consistent traffic	Variable traffic with lots of traffic spikes
Predictable cost structure, while also setting limits	Don't want to think about provisioning throughput (just want it to work)
	Unpredictable/Limited traffic loads (e.g. application in development)

CORRECT: "Configure on-demand capacity mode for the table to enable pay-per-request pricing for read and write requests" is the correct answer (as explained above.)

INCORRECT: "Purchase reserved RCU and WCU for the DynamoDB table and use AWS Application Auto scaling to match the increased load during the peak usage period" is incorrect.

This is a close option however extremely rapid spikes in load could still occur faster than the table could scale.

INCORRECT: "Reduce the provisioned read capacity to match the new peak load on the table and configure DynamoDB Accelerator (DAX) in front of the table" is incorrect.

The question doesn't ask about repeated load, hence caching isn't appropriate in this case.

INCORRECT: " Configure on-demand capacity mode for the table and configure DynamoDB Accelerator (DAX) in front of the table" is incorrect.

As mentioned above, since the application doesn't have repeated load or hot data, caching isn't appropriate in this case.

References:

<https://aws.amazon.com/blogs/aws/amazon-dynamodb-on-demand-no-capacity-planning-and-pay-per-request-pricing/>

Save time with our AWS cheat sheets:

<https://digitalcloud.training/amazon-dynamodb/>

Question 28: **Incorrect**

A company plans to migrate a content management system (CMS) to AWS. The CMS will use Amazon CloudFront to ensure optimum performance for users from around the world. The CMS includes both static and dynamic content and has been placed behind an Application Load Balancer (ALB) which is the default origin for the CloudFront distribution. The static assets are served from an Amazon S3 bucket.

When users attempt to access the static assets HTTP status code 404 errors are generated. Which actions should a Solutions Architect take to resolve the issue? (Select TWO.)

-

Add a CachePolicyConfig to allow HTTP headers to be included in requests to the origin.

(Incorrect)

-

Add another origin to the CloudFront distribution for the static assets.

(Correct)

-

Add a rule to the distribution to forward GET method requests to Amazon S3.

(Incorrect)

-

Add a behavior to the CloudFront distribution for the path pattern and the origin of the static assets.

(Correct)

-

Add a host header condition to the ALB listener and forward the header from CloudFront to add traffic to the allow list.

Explanation

The configuration of Amazon CloudFront is such that all requests are being directed to the default origin, which in this case is the ALB. As the static assets do not exist on the ALB (more specifically, on its targets), the requests fail with a 404. A 404 HTTP status code indicates that the object has not been found.

The resolution to this issue is to add an additional origin that points to the S3 bucket and then use a path pattern behavior to direct requests for these URLs to the S3 origin. The pattern (for example, images/*.jpg) specifies which requests to apply the behavior to. When CloudFront receives a viewer request, the requested path is compared with path patterns in the order in which cache behaviors are listed in the distribution. This will resolve the 404 Not Found error messages.

CORRECT: "Add another origin to the CloudFront distribution for the static assets" is a correct answer.

CORRECT: "Add a behavior to the CloudFront distribution for the path pattern and the origin of the static assets" is a correct answer.

INCORRECT: "Add a rule to the distribution to forward GET method requests to Amazon S3" is incorrect. GET method requests also need to be directed to the ALB, and you cannot configure rules to direct methods to different origins. You can configure the AllowedMethods parameter which controls which methods are available.

INCORRECT: "Add a CachePolicyConfig to allow HTTP headers to be included in requests to the origin" is incorrect. This parameter specifies the values that CloudFront includes in the cache key. These values can include HTTP headers, cookies, and URL query strings. CloudFront uses the cache key to find an object in its cache that it can return to the viewer. This cannot be used to ensure that requests for the static assets will be directed to S3.

INCORRECT: "Add a host header condition to the ALB listener and forward the header from CloudFront to add traffic to the allow list" is incorrect. This will not resolve the 404 Not Found errors. The only way to resolve these errors is to direct traffic to the correct origin.

References:

https://docs.aws.amazon.com/cloudfront/latest/APIReference/API_CacheBehavior.html

Save time with our AWS cheat sheets:

<https://digitalcloud.training/category/aws-cheat-sheets/aws-solutions-architect-professional/aws-networking-content-delivery-sap/>

Question 29: **Incorrect**

A web application allows users to upload video clips of celebrities. The website consists of Amazon EC2 instances and static content. The videos are stored on Amazon EBS volumes and analyzed by custom recognition software for facial analysis. The image processing jobs are picked up from an Amazon SQS queue by an Auto Scaling layer of EC2 instances.

A Solutions Architect has been asked to re-architect the application to reduce operational overhead using AWS managed services where possible. Which of the following recommendations should the Solutions Architect make?



Store the uploaded videos in Amazon EFS and mount the file system to the EC2 instances for the web application. Process the queue with an AWS Lambda functions that calls the Amazon Rekognition API to perform facial analysis.



Use an Amazon S3 static website for the web application. Store uploaded videos in an S3 bucket. Use S3 event notification to publish events to the SQS queue. Process the queue with an AWS Lambda functions that calls the Amazon Rekognition API to perform facial analysis.

(Correct)



Use an Amazon S3 static website for the web application. Store uploaded videos in an S3 bucket. Use S3 event notification to publish events to the SQS queue. Process the queue with Amazon ECS tasks using the EC2 launch type for running the custom recognition software.



Use an Amazon S3 static website for the web application. Store uploaded videos in an S3 bucket. Use S3 event notification to publish events to the SQS queue. Process the queue with Amazon ECS tasks using the Fargate launch type for running the custom recognition software.

(Incorrect)

Explanation

For static websites an Amazon S3 bucket can be used. S3 event notifications can then trigger a Lambda function invocation each time a video is uploaded to the bucket. The Lambda function can then process the images and call the Rekognition API to perform the facial analysis. This solution uses managed AWS services and will reduce operational overhead.

CORRECT: "Use an Amazon S3 static website for the web application. Store uploaded videos in an S3 bucket. Use S3 event notification to publish events to the SQS queue. Process the queue with an AWS Lambda functions that calls the Amazon Rekognition API to perform facial analysis" is the correct answer.

INCORRECT: "Use an Amazon S3 static website for the web application. Store uploaded videos in an S3 bucket. Use S3 event notification to publish events to the SQS queue. Process the queue with Amazon ECS tasks using the EC2 launch type for running the custom recognition software" is incorrect. Lambda is better than ECS for ad-hoc tasks and with the EC2 launch type you must manage instances. Also, the custom recognition software can be replaced with Rekognition – an aws managed service.

INCORRECT: "Store the uploaded videos in Amazon EFS and mount the file system to the EC2 instances for the web application. Process the queue with an AWS Lambda functions that calls the Amazon Rekognition API to perform facial analysis" is incorrect. It will be more operationally efficient to use Lambda and S3 rather than EC2 and EFS.

INCORRECT: "Use an Amazon S3 static website for the web application. Store uploaded videos in an S3 bucket. Use S3 event notification to publish events to the SQS queue. Process the queue with Amazon ECS tasks using the Fargate launch type for running the custom recognition software" is incorrect. Lambda is be a better fit than ECS for ad-hoc tasks. Also, the custom recognition software can be replaced with Rekognition – an aws managed service.

References:

<https://aws.amazon.com/rekognition/>

<https://docs.aws.amazon.com/AmazonS3/latest/dev/NotificationHowTo.html>

Save time with our AWS cheat sheets:

<https://digitalcloud.training/category/aws-cheat-sheets/aws-solutions-architect-professional/aws-compute-sap/>

<https://digitalcloud.training/category/aws-cheat-sheets/aws-solutions-architect-professional/aws-application-integration-sap/>

<https://digitalcloud.training/category/aws-cheat-sheets/aws-solutions-architect-professional/aws-storage-sap/>

Question 30: Incorrect

A company has deployed a high performance computing (HPC) cluster in an Amazon VPC. The cluster runs a tightly coupled workload that generates a large number of shared files that are stored in an Amazon EFS file system. The cluster has grown to over 800 instances and the performance has degraded to a problematic level.

A Solutions Architect needs to make some changes to the design to improve the overall performance. Which of the following changes should the Solutions Architect make? (Select THREE.)

-

Replace Amazon EFS with Amazon FSx for Lustre.

(Correct)

Replace Amazon EFS with multiple FXs for Windows File Server.

Attach multiple elastic network interfaces (ENI) to reduce latency.

Ensure the HPC cluster is launched within a single Availability Zone.

(Correct)

Ensure the cluster is launched across multiple Availability Zones.

(Incorrect)

Enable an Elastic Fabric Adapter (EFA) on a supported EC2 instance type.

(Correct)

Explanation

With HPC clusters with tightly coupled workloads require inter-node communication that is high-performance and low-latency. 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.

In addition to using EFAs with supported EC2 instances, AWS recommend that you launch instances into a single Availability Zone to ensure the latency between nodes is low.

Another cause of latency in the design could be the EFS file system. Amazon FSx for Lustre is a fully managed service that provides cost-effective, high-performance, scalable storage for compute workloads. FSx for Lustre offers sub-millisecond latencies, up to hundreds of gigabytes per second of throughput, and millions of IOPS.

Replacing EFS with Lustre will reduce the latency for storage operations.

CORRECT: "Ensure the HPC cluster is launched within a single Availability Zone" is a correct answer.

CORRECT: "Enable an Elastic Fabric Adapter (EFA) on a supported EC2 instance type" is also a correct answer.

CORRECT: "Replace Amazon EFS with Amazon FSx for Lustre" is also a correct answer.

INCORRECT: "Attach multiple elastic network interfaces (ENI) to reduce latency" is incorrect. With HPC workloads latency is the biggest concern and it's better to use EFS adapters rather than ENIs to reduce latency.

INCORRECT: "Ensure the cluster is launched across multiple Availability Zones" is incorrect. The cluster should be launched in a single AZ for performance.

INCORRECT: "Replace Amazon EFS with multiple FXs for Windows File Server" is incorrect. This service is used for Microsoft file systems and is not a suitable replacement for EFS, nor will it offer performance advantages.

References:

<https://aws.amazon.com/fsx/lustre/>

<https://aws.amazon.com/hpc/efa/>

Save time with our AWS cheat sheets:

<https://digitalcloud.training/category/aws-cheat-sheets/aws-solutions-architect-professional/security-identity-compliance-sap/>

<https://digitalcloud.training/category/aws-cheat-sheets/aws-solutions-architect-professional/aws-storage-sap/>

Question 1: **Incorrect**

A finance company needs to implement a solution to share a common network across multiple AWS accounts which are a part of an AWS organization.

The company's operations team uses a dedicated operations account with a VPC, and this must be used for network management. Individual accounts cannot have the ability to manage their own networks. However, individual accounts must be able to create AWS resources within subnets.

Which combination of actions should be taken to meet these requirements? (Select TWO.)

-

Create a resource share in AWS Resource Access Manager in the operations account. Select the specific AWS Organizations OU that will use the shared network. Select each prefix list to associate with the resource share.

(Correct)

-

Create a transit gateway in the operations account and enable transitive routing.

-

Create a resource share in AWS Resource Access Manager in the operations account. Select the specific AWS Organizations OU that will use the shared network. Select each subnet to associate with the resource share.

(Incorrect)

-

Create VPCs in each AWS account within the organization in AWS Organizations. Configure the VPCs to share the same CIDR range and subnets as the VPC in the operations account. Peer the VPCs in each individual account with the VPC in the operations account.



- **Enable resource sharing from the AWS Organizations management account.**

(Correct)

Explanation

To share resources within an organization, you must first use the AWS RAM console or AWS Command Line Interface (AWS CLI) to enable sharing with AWS Organizations. When you share resources in your organization, AWS RAM doesn't send invitations to principals. Principals in your organization gain access to shared resources without exchanging invitations.

From the AWS CLI, command `enable-sharing-with-aws-organization` enables resource sharing within your organization in Organizations. Calling this operation enables RAM to retrieve information about the organization and its structure. This lets you share resources with all the accounts in an organization by specifying the organization's ID, or all the accounts in an organizational unit (OU) by specifying the OU's ID.

CORRECT: "Enable resource sharing from the AWS Organizations management account" is a correct answer (as explained above.)

CORRECT: "Create a resource share in AWS Resource Access Manager in the operations account. Select the specific AWS Organizations OU that will use the shared network. Select each prefix list to associate with the resource share" is also a correct answer (as explained above.)

INCORRECT: "Create a transit gateway in the operations account and enable transitive routing" is incorrect.

Transit gateway by default only allows VPCs from the same AWS account to be attached. For our cross-account scenario, we'll have to use another AWS service called the Resource Access Manager (RAM). RAM lets you share certain resources between AWS accounts. Sharing the transit gateway with another AWS account means that VPCs from that account can be attached to it. Since the question mentions use of AWS organizations, this is not an apt option.

INCORRECT: "Create VPCs in each AWS account within the organization in AWS Organizations. Configure the VPCs to share the same CIDR range and subnets as the VPC in the operations account. Peer the VPCs in each individual account with the VPC in the operations account" is incorrect.

Peering the individual accounts within a VPC can only be used when the accounts are static and small. You can have a maximum of 125 peering connections per VPC. AWS VPC best practices [recommend](#) you do not use more than 10 VPCs in a mesh to limit management complexity. To create a mesh network where every VPC is peered to every other VPC, it takes $n - 1$ connections per VPC where n is the number of VPCs.

INCORRECT: "Create a resource share in AWS Resource Access Manager in the operations account. Select the specific AWS Organizations OU that will use the shared network. Select each subnet to associate with the resource share" is incorrect.

This option is incorrect since prefix list needs to be selected rather than a specific subnet.

References:

<https://docs.aws.amazon.com/ram/latest/userguide/getting-started-sharing.html>

Save time with our AWS cheat sheets:

<https://digitalcloud.training/amazon-vpc/>

Question 2: **Correct**

A mobile app has become extremely popular with global usage increasing to millions of users. The app allows users to capture and upload funny images of animals and add captions. The current application runs on Amazon EC2 instances with Amazon EFS storage behind an Application Load Balancer. The data access patterns are unpredictable and during peak periods the application has experienced performance issues.

Which changes should a Solutions Architect make to the application architecture to control costs and improve performance?

-
-

Place AWS Global Accelerator in front of the ALB. Migrate the static content to Amazon FSx for Windows File Server. Use an AWS Lambda function to reduce image size during the migration process.

-
-

Use an Amazon S3 bucket for static images and use the Intelligent Tiering storage class. Use an Amazon CloudFront distribution in front of the S3 bucket and AWS Lambda for processing the images.

(Correct)

- Use an Amazon S3 bucket for static images and use the Intelligent Tiering storage class. Use an Amazon CloudFront distribution in front of the S3 bucket and the ALB.**

-

Create an Amazon CloudFront distribution and place the ALB behind the distribution. Store static content in Amazon S3 in an Infrequent Access storage class.

Explanation

The best option for reducing costs and improving performance would be to move to a fully serverless solution. Amazon S3 can store the image files and CloudFront can be used to improve performance for the global user base. AWS Lambda is ideal for processing the images. The solution will scale seamlessly and handle peak loads and is also low cost.

CORRECT: "Use an Amazon S3 bucket for static images and use the Intelligent Tiering storage class. Use an Amazon CloudFront distribution in front of the S3 bucket and AWS Lambda for processing the images" is the correct answer.

INCORRECT: "Use an Amazon S3 bucket for static images and use the Intelligent Tiering storage class. Use an Amazon CloudFront distribution in front of the S3 bucket and the ALB" is incorrect. This is a good solution but not quite as cost-effective as using Lambda in place of the ALB and EC2 instances.

INCORRECT: "Create an Amazon CloudFront distribution and place the ALB behind the distribution. Store static content in Amazon S3 in an Infrequent Access storage class" is incorrect. Infrequent access storage class incurs retrieval fees so the data costs could be more expensive compared to other storage classes.

INCORRECT: "Place AWS Global Accelerator in front of the ALB. Migrate the static content to Amazon FSx for Windows File Server. Use an AWS Lambda function to reduce image size during the migration process" is incorrect. GA is best suited for use cases where

you need to leverage the AWS global network to improve performance to your application endpoints across multiple Regions. In this case it represents a more costly solution.

References:

<https://aws.amazon.com/premiumsupport/knowledge-center/cloudfront-serve-static-website/>

<https://aws.amazon.com/premiumsupport/knowledge-center/lambda-configure-s3-event-notification/>

Save time with our AWS cheat sheets:

<https://digitalcloud.training/category/aws-cheat-sheets/aws-solutions-architect-professional/aws-networking-content-delivery-sap/>

Question 3: Incorrect

A Solutions Architect has deployed a REST API using an Amazon API Gateway Regional endpoint. The API will be consumed by a growing number of US-based companies. Each company will use the API twice each day to get the latest data.

Following the deployment of the API the operations team noticed thousands of requests coming from hundreds of IP addresses around the world. The traffic is believed to be originating from a botnet. The Solutions Architect must secure the API while minimizing cost.

Which approach should the company take to secure its API?

-

Create an Amazon CloudFront distribution with the API as the origin. Create an AWS WAF web ACL with a rule to block clients that submit more than ten requests per day. Associate the web ACL with the CloudFront distribution. Add a custom header to the CloudFront distribution populated with an API key. Configure the API to require an API key on the GET method.

-

Create an AWS WAF web ACL with a rule to allow access to the IP addresses used by the companies. Associate the web ACL with the API. Create a resource policy with a request limit and associate it with the API. Configure the API to require an API key on the POST method.



Create an AWS WAF web ACL with a rule to allow access from the IP addresses used by the companies. Associate the web ACL with the API. Create a usage plan with a request limit and associate it with the API. Create an API key and add it to the usage plan.

(Correct)



Create an Amazon CloudFront distribution with the API as the origin. Create an AWS WAF web ACL with a rule to block clients that submit more than ten requests per day. Associate the web ACL with the CloudFront distribution. Configure CloudFront with an origin access identity (OAI) and associate it with the distribution. Configure API Gateway to ensure only the OAI can execute the GET method.

(Incorrect)

Explanation

The rate-based rules associated with usage plans specify the number of web requests that are allowed by each client IP in a trailing, continuously updated, 5-minute period. The API key associated with the usage plan ensures that only clients who are using the API key in their requests are granted access. This solution requires that the IP addresses of clients are whitelisted and the API key is distributed to clients to use in their requests to the API.

CORRECT: "Create an AWS WAF web ACL with a rule to allow access from the IP addresses used by the companies. Associate the web ACL with the API. Create a usage plan with a request limit and associate it with the API. Create an API key and add it to the usage plan" is correct.

INCORRECT: "Create an Amazon CloudFront distribution with the API as the origin. Create an AWS WAF web ACL with a rule to block clients that submit more than ten requests per day. Associate the web ACL with the CloudFront distribution. Add a custom header to the CloudFront distribution populated with an API key. Configure the API to require an API key on the GET method" is incorrect answer.

A rate-based rule tracks the rate of requests for each originating IP address, and triggers the rule action on IPs with rates that go over a limit. You set the limit as the number of requests per 5-minute time span. You cannot configure an AWS WAF rate-based rule to limit request to 10 per day. Also, to minimize cost CloudFront is not required in this solution.

INCORRECT: "Create an Amazon CloudFront distribution with the API as the origin. Create an AWS WAF web ACL with a rule to block clients that submit more than ten requests per day. Associate the web ACL with the CloudFront distribution. Configure CloudFront with an origin access identity (OAI) and associate it with the distribution. Configure API Gateway to ensure only the OAI can execute the GET method" is incorrect.

As above for rate-based rules. An OAI is a special CloudFront user that is used with Amazon S3 buckets to prevent direct access using S3 URLs. It is usually used along with other protections such as signed URLs and signed cookies. It is not possible to use an OAI with API Gateway APIs.

INCORRECT: "Create an AWS WAF web ACL with a rule to allow access to the IP addresses used by the companies. Associate the web ACL with the API. Create a resource policy with a request limit and associate it with the API. Configure the API to require an API key on the POST method" is incorrect.

AWS WAF resource policies control whether or not a principal or source IP address/CIDR block is allowed to invoke the API. A resource policy does not have a request limit associated with it, use a Web ACL rate-based rule for that. The API key should be configured on the GET method as this API is being used to get data and not post it.

References:

<https://docs.aws.amazon.com/apigateway/latest/developerguide/apigateway-resource-policies.html>

<https://docs.aws.amazon.com/waf/latest/developerguide/waf-rule-statement-type-rate-based.html>

<https://docs.aws.amazon.com/apigateway/latest/developerguide/api-gateway-api-usage-plans.html>

<https://docs.aws.amazon.com/apigateway/latest/developerguide/apigateway-control-access-aws-waf.html>

Save time with our AWS cheat sheets:

<https://digitalcloud.training/category/aws-cheat-sheets/aws-solutions-architect-professional/aws-networking-content-delivery-sap/>

Question 4: **Incorrect**

A company is creating a multi-account structure using AWS Organizations. The accounts will include the Management account, Production account, and Development account. The company requires auditing for all API actions across accounts. A Solutions Architect is advising the company on how to configure the accounts. Which of the following recommendations should the Solutions Architect make? (Select TWO.)

-

Create all resources in the Management account and grant access to the Production and Development accounts.

-

Create user accounts in the Management account and use cross-account access to access resources.

(Incorrect)

-

Enable AWS CloudTrail and keep all CloudTrail trails and logs within each account.

(Incorrect)

-

Create user accounts in the Production and Development accounts.

(Correct)

-

Enable AWS CloudTrail and keep all CloudTrail trails and logs in the management account.

(Correct)

Explanation

AWS recommends that you use the management account and its users and roles only for tasks that can be performed only by that account. Store all of your AWS resources in *other* AWS accounts in the organization and keep them out of the management account. The one exception is that AWS does recommend that you enable AWS CloudTrail and keep relevant CloudTrail trails and logs in the management account.

One important reason to keep your resources in other accounts is because Organizations service control policies (SCPs) do not work to restrict any users or roles in the management account.

Separating your resources from your management account also help you to understand the charges on your invoices.

CORRECT: "Enable AWS CloudTrail and keep all CloudTrail trails and logs in the management account" is a correct answer.

CORRECT: "Create user accounts in the Production and Development accounts" is also a correct answer.

INCORRECT: "Enable AWS CloudTrail and keep all CloudTrail trails and logs within each account" is incorrect. AWS recommends that you centralize this data to the management account.

INCORRECT: "Create user accounts in the Management account and use cross-account access to access resources" is incorrect. Because user accounts in the management account are not restricted by SCPs, this is not a best practice.

INCORRECT: "Create all resources in the Management account and grant access to the Production and Development accounts" is incorrect. Resources should be created in the relevant accounts (Production/Development), you cannot just grant access and have them in the management account.

References:

https://docs.aws.amazon.com/organizations/latest/userguide/orgs_best-practices_mgmt-acct.html

Save time with our AWS cheat sheets:

<https://digitalcloud.training/category/aws-cheat-sheets/aws-solutions-architect-professional/aws-management-tools-sap/>

Question 5: Incorrect

An S3 endpoint has been created in an Amazon VPC. A staff member assumed an IAM role and attempted to download an object from a bucket using the endpoint. The staff member received the error message "403: Access Denied". The bucket is encrypted using an AWS KMS key. A Solutions Architect has verified that the staff member assumed the correct IAM role and the role does allow the object to be downloaded. The bucket policy and NACL are also valid.

Which additional step should the Solutions Architect take to troubleshoot this issue?

- Check that local firewall rules are not preventing access to the S3 endpoint.
- Ensure that blocking all public access has not been enabled in the S3 bucket.
- Verify that the IAM role has permission to decrypt the referenced KMS key.
(Correct)
- Verify that the IAM role has the correct trust relationship configured.
(Incorrect)

Explanation

If an IAM user can't access an object that the user has full permissions to, then check if the object is encrypted by AWS KMS. You can use the Amazon S3 console to view the object's properties, which include the object's encryption information.

If the object is KMS encrypted, then make sure that the KMS key policy grants permissions to the IAM user for the following actions:

- "kms:Encrypt"
- "kms:Decrypt"
- "kms:ReEncrypt"
- "kms:GenerateDataKey"
- "kms:DescribeKey"

CORRECT: "Verify that the IAM role has permission to decrypt the referenced KMS key" is the correct answer.

INCORRECT: "Verify that the IAM role has the correct trust relationship configured" is incorrect. If the IAM role trust relationship was not configured correctly the user would not be able to assume the role and the question states that the user did assume the role.

INCORRECT: "Ensure that blocking all public access has not been enabled in the S3 bucket" is incorrect. This is not a case of public access, the S3 bucket is being accessed using an IAM role with the permissions set correctly.

INCORRECT: "Check that local firewall rules are not preventing access to the S3 endpoint" is incorrect. The NACL is valid and an access denied error is being generated by S3, it would not be generated by a firewall.

References:

<https://aws.amazon.com/premiumsupport/knowledge-center/s3-troubleshoot-403/>

Save time with our AWS cheat sheets:

<https://digitalcloud.training/category/aws-cheat-sheets/aws-solutions-architect-professional/security-identity-compliance-sap/>

Question 6: **Correct**

A company recently migrated a high-traffic eCommerce website to the AWS Cloud. The website is experiencing strong growth. Developers use a private GitHub repository to manage code and the DevOps team use Jenkins for builds and unit testing.

The Developers need to receive notifications when a build does not work and ensure there is no downtime during deployments. It is also required that any changes to production are seamless for users and can be easily rolled back if a significant issue occurs.

A Solutions Architect is finalizing the design for the environment and will use AWS CodePipeline to manage the build and deployment process. What other steps should be taken to meet the requirements?

- Use GitHub websockets to trigger the CodePipeline pipeline. Use AWS X-Ray for unit testing and static code analysis. Send alerts to an Amazon SNS topic for any bad builds. Deploy in a blue/green deployment using AWS CodeDeploy.
- Use GitHub webhooks to trigger the CodePipeline pipeline. Use the Jenkins plugin for AWS CodeBuild to conduct unit testing. Send alerts to an Amazon SNS topic for any bad builds. Deploy in an in-place, all-at-once deployment configuration using AWS CodeDeploy.
- Use GitHub webhooks to trigger the CodePipeline pipeline. Use the Jenkins plugin for AWS CodeBuild to conduct unit testing. Send alerts to an Amazon SNS topic for any bad builds. Deploy in a blue/green deployment using AWS CodeDeploy.

(Correct)

Use GitHub webhooks to trigger the CodePipeline pipeline. Use AWS X-Ray for unit testing and static code analysis. Send alerts to an Amazon SNS topic for any bad builds. Deploy in an in-place, all-at-once deployment configuration using AWS CodeDeploy.

Explanation

AWS CodePipeline can receive a webhook from GitHub when a change is made to your GitHub repository. Webhooks can tell CodePipeline to initiate a pipeline execution.

You can use the Jenkins plugin for AWS CodeBuild to integrate CodeBuild with your Jenkins build jobs. Instead of sending your build jobs to Jenkins build nodes, you use the plugin to send your build jobs to CodeBuild. This eliminates the need for you to provision, configure, and manage Jenkins build nodes.

SNS is an obvious tool to use for sending notifications and CodeDeploy should use a blue/green deployment strategy. This strategy ensures that updates are seamless and can be easily rolled back if necessary.

CORRECT: "Use GitHub webhooks to trigger the CodePipeline pipeline. Use the Jenkins plugin for AWS CodeBuild to conduct unit testing. Send alerts to an Amazon SNS topic for any bad builds. Deploy in a blue/green deployment using AWS CodeDeploy" is the correct answer.

INCORRECT: "Use GitHub webhooks to trigger the CodePipeline pipeline. Use AWS X-Ray for unit testing and static code analysis. Send alerts to an Amazon SNS topic for any bad builds. Deploy in an in-place, all-at-once deployment configuration using AWS CodeDeploy" is incorrect. You cannot use X-Ray for unit testing, it is used for tracing. Also, an in-place, all-at-once deployment strategy makes it difficult to roll back.

INCORRECT: "Use GitHub websockets to trigger the CodePipeline pipeline. Use the Jenkins plugin for AWS CodeBuild to conduct unit testing. Send alerts to an Amazon SNS topic for any bad builds. Deploy in an in-place, all-at-once deployment configuration using AWS CodeDeploy" is incorrect. You cannot use websockets to trigger CodePipeline, webhooks must be used.

INCORRECT: "Use GitHub websockets to trigger the CodePipeline pipeline. Use AWS X-Ray for unit testing and static code analysis. Send alerts to an Amazon SNS topic for any bad builds. Deploy in a blue/green deployment using AWS CodeDeploy" is incorrect. You cannot use websockets to trigger CodePipeline, webhooks must be used.

References:

<https://aws.amazon.com/about-aws/whats-new/2018/05/aws-codepipeline-supports-push-events-from-github-via-webhooks/>

<https://docs.aws.amazon.com/codebuild/latest/userguide/jenkins-plugin.html>

Save time with our AWS cheat sheets:

<https://digitalcloud.training/category/aws-cheat-sheets/aws-solutions-architect-professional/aws-developer-tools-sap/>

Question 7: Incorrect

A company leases data center space in a co-location facility and needs to move out before the end of the financial year in 90 days. The company currently runs 150 virtual machines and a NAS device that holds over 50 TB of data. Access patterns for the data are infrequent but when access is required it must be immediate. The VM configurations are highly customized. The company has a 1 Gbps internet connection which is mostly idle and almost completely unused outside of business hours.

Which combination of steps should a Solutions Architect take to migrate the VMs to AWS with minimal downtime and operational impact? (Select TWO.)

-

Copy infrequently accessed data from the NAS using AWS SMS.

(Incorrect)

-

Migrate the virtual machines with AWS SMS.

(Correct)

-

Migrate the NAS data to AWS using AWS Snowball.

-

Migrate the NAS data to AWS using AWS Storage Gateway.

(Correct)

-

Launch new Amazon EC2 instances and reinstall all applications.

Explanation

Explanation:

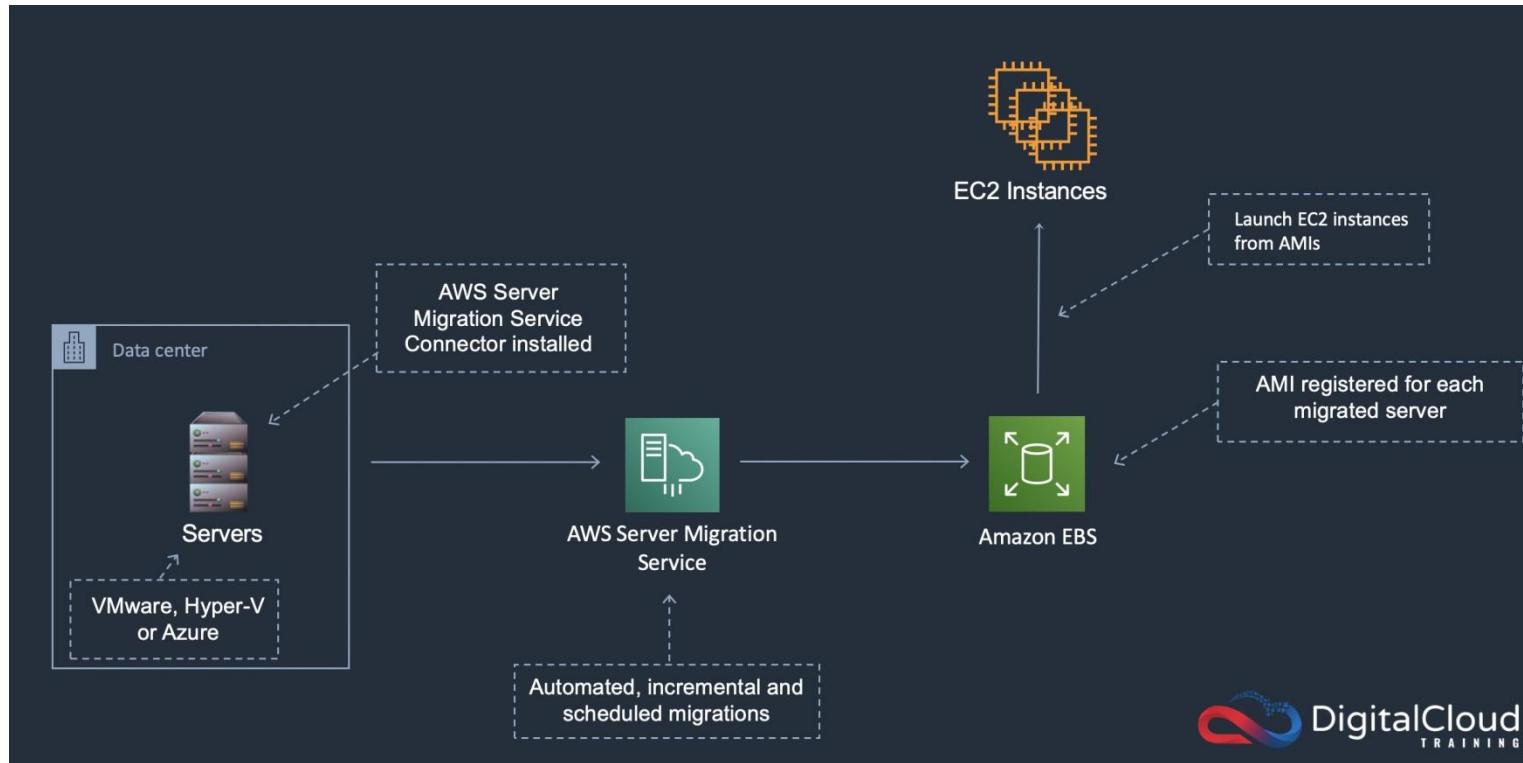
The key to answering this question correctly is to understand all of the different migration tools and the variables that determine which is best for the scenario. In this case there are two variables that are of the most importance:

- VMs have customized configurations.
- Data is 50 TB and there is an idle 1 Gbps connection.

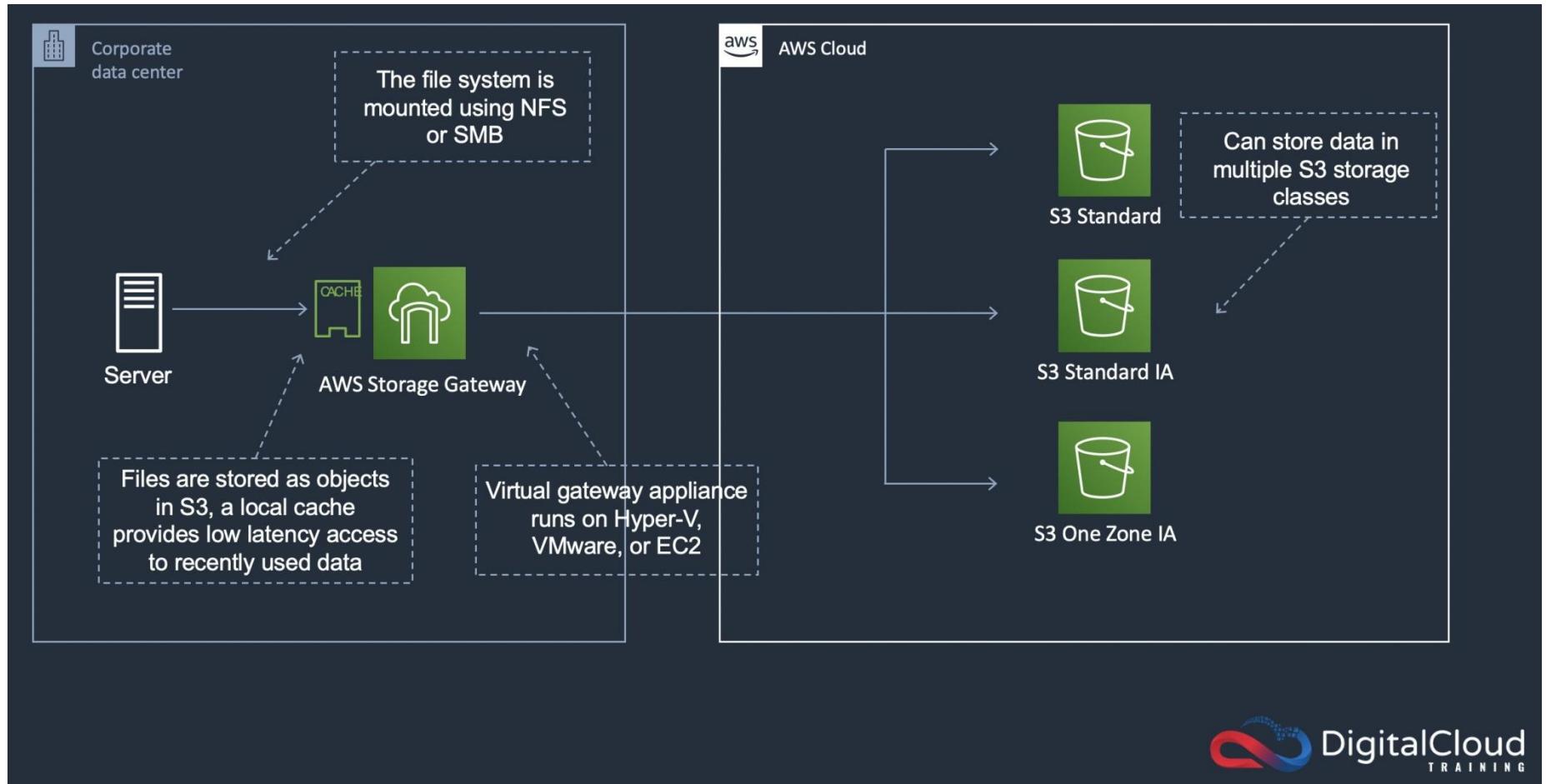
The customized configuration of the VMs means that creating new instances on EC2 will not work as customizations would need to be made individually. Therefore, the VMs should be migrated with their OS intact.

Moving 50 TB of data over an idle 1 Gbps connection should take around 5-6 days. Assuming there's some traffic on the connection during business hours this could slow down by a few days but there's still enough bandwidth to migrate all VMs and NAS data within the 90 day period.

Therefore, we can use online tools rather than migrating offline with Snowball. In this case AWS Server Migration Service (SMS) can be used to migrate VMs which will keep the OS intact and migrate the VMs to Amazon EC2.



AWS Storage Gateway can also be used which will synchronize the NAS data into Amazon where it can be stored in a storage class suited for infrequent access. This provides a method to synchronize data using the existing internet connection.



CORRECT: "Migrate the virtual machines with AWS SMS" is a correct answer.

CORRECT: "Migrate the NAS data to AWS using AWS Storage Gateway" is also a correct answer.

INCORRECT: "Launch new Amazon EC2 instances and reinstall all applications" is incorrect. The VMs are individually customized so there would too many configuration updates to make this a viable solution.

INCORRECT: "Migrate the NAS data to AWS using AWS Snowball" is incorrect. This would mean the data is offline for many days which is not ideal. It is also unnecessary as the data can be easily migrated across the existing internet connection within acceptable timeframes.

INCORRECT: "Copy infrequently accessed data from the NAS using AWS SMS" is incorrect. SMS is used for migrating servers not data from a NAS device.

References:

<https://aws.amazon.com/storagegateway/file/>

<https://docs.aws.amazon.com/server-migration-service/latest/userguide/server-migration.html>

Save time with our AWS cheat sheets:

<https://digitalcloud.training/category/aws-cheat-sheets/aws-solutions-architect-professional/aws-migration-sap/>

<https://digitalcloud.training/category/aws-cheat-sheets/aws-solutions-architect-professional/aws-storage-sap/>

Question 8: **Correct**

A company captures financial transactions in Amazon DynamoDB tables. The security team is concerned about identifying fraudulent behavior and has requested that all changes to items stored in DynamoDB tables must be logged within 30 minutes.

How can a Solutions Architect meet this requirement?

-

Use AWS CloudTrail to capture all the APIs that change the DynamoDB tables. Send SNS notifications when anomalous behaviors are detected using CloudTrail event filtering.

-

Use event patterns in Amazon CloudWatch Events to capture DynamoDB API call events with an AWS Lambda function as a target to analyze behavior. Send SNS notifications when anomalous behaviors are detected.



Use Amazon DynamoDB Streams to capture and send updates to AWS Lambda. Create a Lambda function to output records to Amazon Kinesis Data Streams. Analyze any anomalies with Amazon Kinesis Data Analytics. Send SNS notifications when anomalous behaviors are detected.

(Correct)



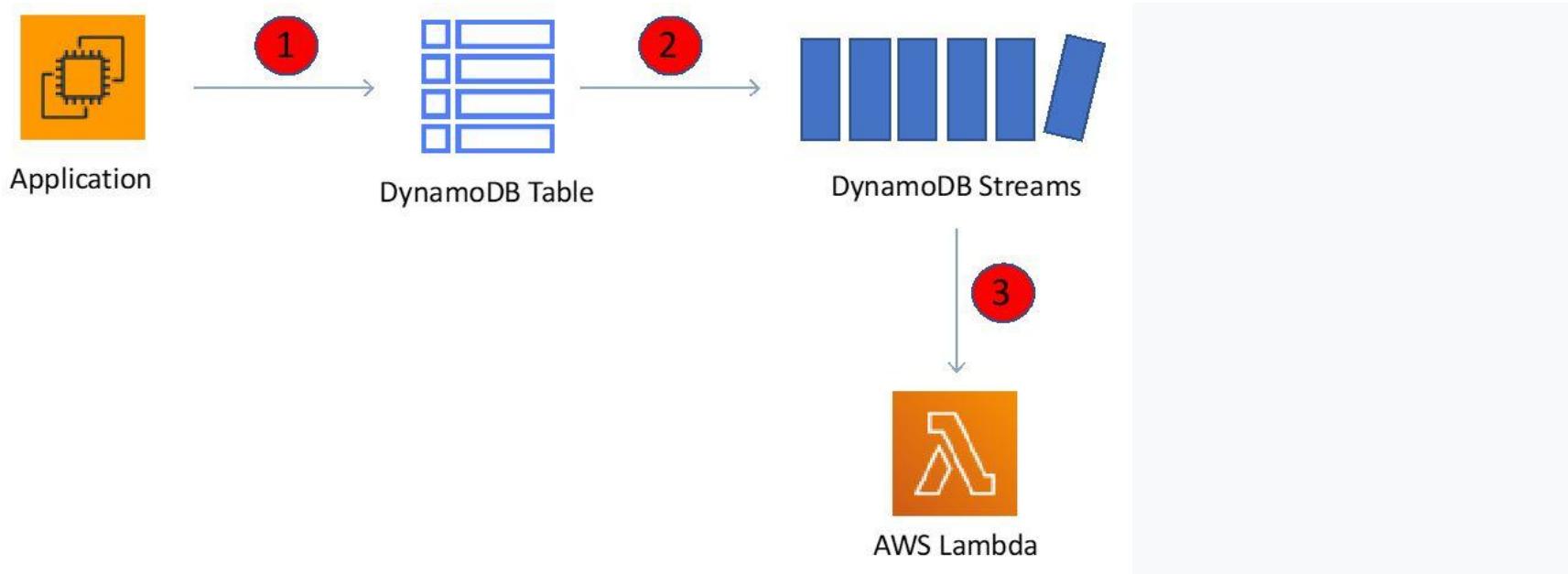
Copy the DynamoDB tables into Apache Hive tables on Amazon EMR every hour and analyze them for anomalous behaviors. Send Amazon SNS notifications when anomalous behaviors are detected.

Explanation

DynamoDB Streams captures a time-ordered sequence of item-level modifications in a DynamoDB table. DynamoDB is integrated with AWS Lambda so that you can create *triggers*—pieces of code that automatically respond to events in DynamoDB Streams.

With triggers, you can build applications that react to data modifications in DynamoDB tables. In this case the Lambda function can process the data and place it in a Kinesis Data Stream where Data Analytics can analyze the data and send an SNS notification if any fraudulent behavior is detected.

The diagram below depicts how data is updated in the table (1), the modification is added to the stream (2), and then AWS Lambda processes the record (3). The record of the update can then be analyzed using Kinesis (not shown).



CORRECT: "Use Amazon DynamoDB Streams to capture and send updates to AWS Lambda. Create a Lambda function to output records to Amazon Kinesis Data Streams. Analyze any anomalies with Amazon Kinesis Data Analytics. Send SNS notifications when anomalous behaviors are detected" is the correct answer.

INCORRECT: "Use event patterns in Amazon CloudWatch Events to capture DynamoDB API call events with an AWS Lambda function as a target to analyze behavior. Send SNS notifications when anomalous behaviors are detected" is incorrect. To capture item-level modification DynamoDB streams should be used, capturing API calls will not help.

INCORRECT: "Copy the DynamoDB tables into Apache Hive tables on Amazon EMR every hour and analyze them for anomalous behaviors. Send Amazon SNS notifications when anomalous behaviors are detected" is incorrect. This solution uses the wrong features and tools for the job and is not as automated as the correct answer.

INCORRECT: "Use AWS CloudTrail to capture all the APIs that change the DynamoDB tables. Send SNS notifications when anomalous behaviors are detected using CloudTrail event filtering" is incorrect. Capturing API calls does not give you the information you need at an item-level. To capture the changes to the items you must use DynamoDB streams.

References:

<https://docs.aws.amazon.com/amazondynamodb/latest/developerguide/Streams.Lambda.html>

Save time with our AWS cheat sheets:

<https://digitalcloud.training/category/aws-cheat-sheets/aws-solutions-architect-professional/aws-database-sap/>

Question 9: Incorrect

A company has created a fitness tracking mobile app that uses a serverless REST API. The app consists of an Amazon API Gateway API with a Regional endpoint, AWS Lambda functions and an Amazon Aurora MySQL database cluster. The company recently secured a deal with a sports company to promote the new app which resulted in a significant increase in the number of requests received.

Unfortunately, the increase in traffic resulted in sporadic database memory errors and performance degradation. The traffic included significant numbers of HTTP requests querying the same data in short bursts of traffic during weekends and holidays.

The company needs to improve its ability to support the additional usage while minimizing the increase in costs associated with the solution.

Which strategy meets these requirements?



Implement an Amazon ElastiCache for Redis cache to store the results of the database calls. Modify the Lambda functions to use the cache.

(Incorrect)



Create usage plans in API Gateway and distribute API keys to clients. Configure metered access to the production stage.

-

Modify the instance type of the Aurora database cluster to use an instance with more memory.

-

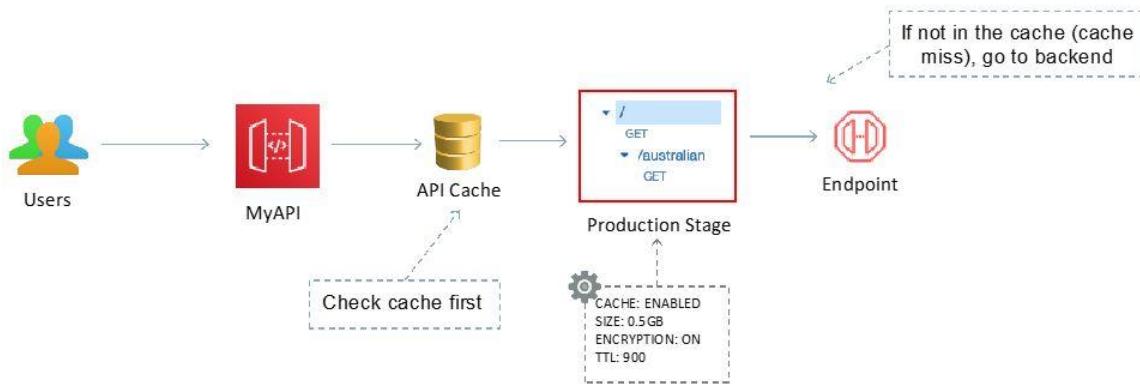
Convert the API Gateway Regional endpoint to an edge-optimized endpoint. Enable caching in the production stage.

(Correct)

Explanation

An *edge-optimized API endpoint* is best for geographically distributed clients. API requests are routed to the nearest CloudFront Point of Presence (POP). For mobile clients this is a good use case for this type of endpoint. The Regional endpoint is best suited to traffic coming from within the Region only.

You can enable API caching in Amazon API Gateway to cache your endpoint's responses. With caching, you can reduce the number of calls made to your endpoint and also improve the latency of requests to your API.



CORRECT: "Convert the API Gateway Regional endpoint to an edge-optimized endpoint. Enable caching in the production stage" is the correct answer.

INCORRECT: "Create usage plans in API Gateway and distribute API keys to clients. Configure metered access to the production stage" is incorrect. This does not support the additional usage; it limits additional usage.

INCORRECT: "Implement an Amazon ElastiCache for Redis cache to store the results of the database calls. Modify the Lambda functions to use the cache" is incorrect. This will increase costs associated with the solution as the ElastiCache cluster could be expensive.

INCORRECT: "Modify the instance type of the Aurora database cluster to use an instance with more memory" is incorrect. This would mean the database cluster cost more at all times, not just when the traffic increases.

References:

<https://docs.aws.amazon.com/apigateway/latest/developerguide/api-gateway-caching.html>

Save time with our AWS cheat sheets:

<https://digitalcloud.training/category/aws-cheat-sheets/aws-solutions-architect-professional/aws-networking-content-delivery-sap/>

Question 10: **Incorrect**

A company runs several IT services in an on-premises data center that is connected to AWS using an AWS Direct Connect (DX) connection. The service data is sensitive and the company uses an IPSec VPN over the DX connection to encrypt data. Security requirements mandate that the data cannot traverse the internet. The company wants to offer the IT services to other companies who use AWS.

Which solution will meet these requirements?

-

Create a VPC Endpoint Service that accepts TCP traffic and host it behind a Network Load Balancer. Enable access to the IT services over the DX connection.

(Correct)

-

Configure a mesh of AWS VPN CloudHub IPsec VPN connections between the customer AWS accounts and the service provider AWS account.

-

Create a VPC Endpoint Service that accepts HTTP or HTTPS traffic and host it behind an Application Load Balancer. Enable access to the IT services over the DX connection.

(Incorrect)

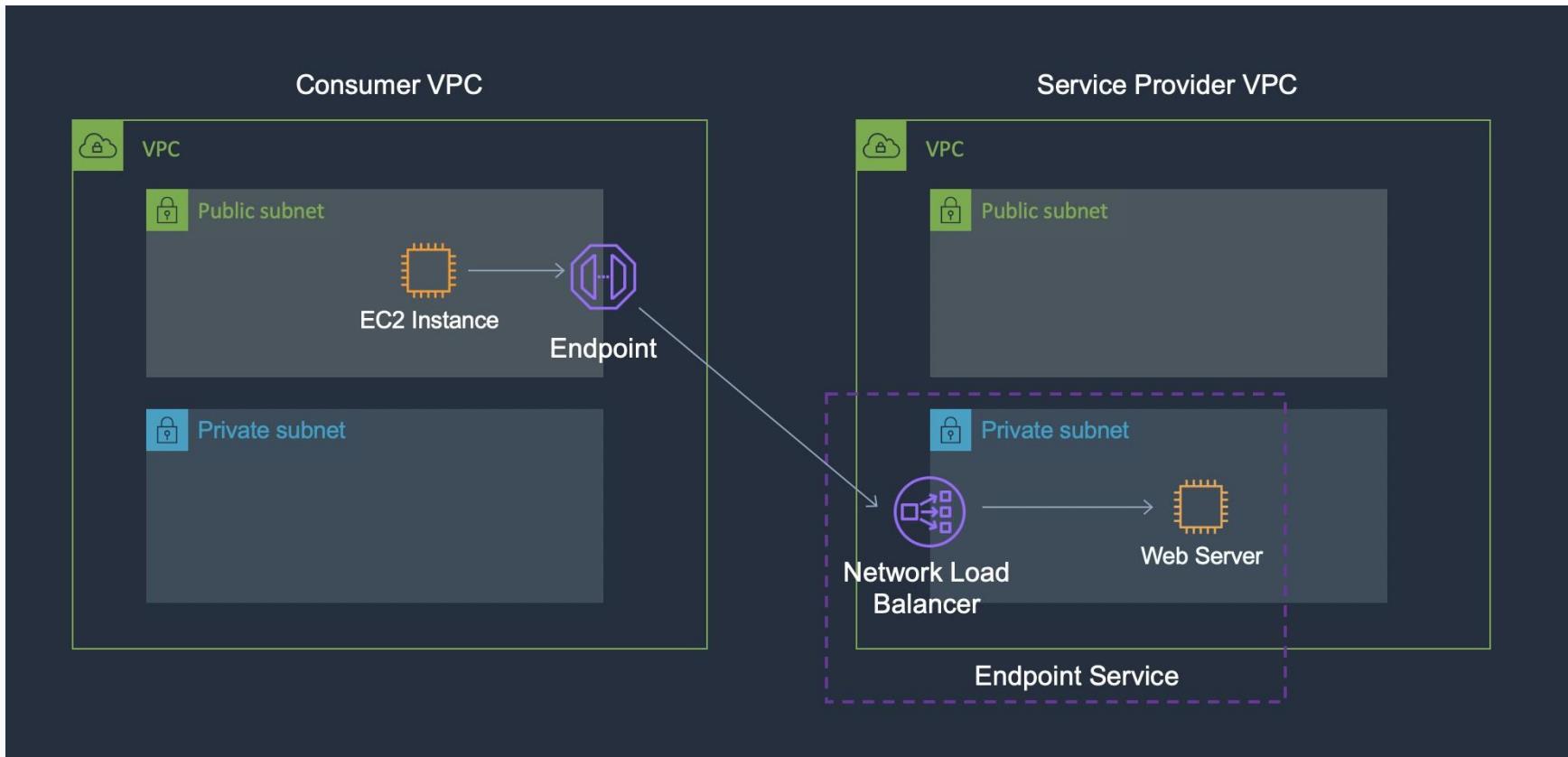
-

Attach an internet gateway to the VPC and ensure that network access control and security group rules allow the relevant inbound and outbound traffic.

Explanation

The solution is to use VPC endpoint services in a service provider model. In this model a network load balancer must be created in the service provider VPC in front of the application services. Remember that NLBs can use on-premises targets. A VPC endpoint is then created that uses the NLB.

A service consumer that has been granted permissions then creates an interface endpoint to your service, optionally in each Availability Zone in which you configured your service. This is depicted in the image below:



CORRECT: "Create a VPC Endpoint Service that accepts TCP traffic and host it behind a Network Load Balancer. Enable access to the IT services over the DX connection" is the correct answer.

INCORRECT: "Create a VPC Endpoint Service that accepts HTTP or HTTPS traffic and host it behind an Application Load Balancer. Enable access to the IT services over the DX connection" is incorrect. A NLB should be used for a VPC endpoint service.

INCORRECT: "Configure a mesh of AWS VPN CloudHub IPsec VPN connections between the customer AWS accounts and the service provider AWS account" is incorrect. VPNs use the internet and the internet must be avoided. Note that the VPN used by the company runs over DX, not over the internet, so the connection can be encrypted.

INCORRECT: "Attach an internet gateway to the VPC and ensure that network access control and security group rules allow the relevant inbound and outbound traffic" is incorrect. An internet gateway is used for internet-based connectivity which should be avoided. It is not needed for a VPC endpoint service.

References:

<https://docs.aws.amazon.com/vpc/latest/userguide/endpoint-service-overview.html>

Save time with our AWS cheat sheets:

<https://digitalcloud.training/category/aws-cheat-sheets/aws-solutions-architect-professional/aws-networking-content-delivery-sap/>

Question 11: **Correct**

An online retailer is updating its catalogue of products. The retailer has a dynamic website which uses EC2 instances for web and application servers. The web tier is behind an Application Load Balancer and the application tier stores data in an Amazon Aurora MySQL database. There is additionally a lot of static content and most website traffic is read-only.

The company is expecting a large spike in traffic to the website when the new catalogue is launched and optimal performance is a high priority.

Which combination of steps should a Solutions Architect take to reduce system response times for a global audience? (Select TWO.)

-

Configure an Aurora global database for storage-based cross-Region replication. Use Amazon S3 with cross-Region replication for static content and resources and create Amazon CloudFront distributions.

(Correct)

-

Use logical cross-Region replication to replicate the Aurora MySQL database to a secondary Region. Replace the web servers with Amazon S3. Configure cross-Region replication for the S3 buckets.

-

Use Amazon Route 53 with a latency-based routing policy. Create Auto Scaling groups for the web and application tiers and deploy them in multiple global Regions.

(Correct)

-

Migrate the database from Amazon Aurora to Amazon RDS for MySQL. Replace the web and application tiers with AWS Lambda functions, create an Amazon SQS queue.

-

Create Auto Scaling groups for the web and application tiers and deploy them in multiple global Regions. Setup an AWS Direct Connect connection.

Explanation

The website performance can be optimized for global users through a combination of using Amazon CloudFront to deliver static assets and EC2 instances launched in multiple Regions in Auto Scaling groups. To direct traffic to the correct instances latency-based routing policies can be created in Route 53 which will direct traffic to the closest (lowest-latency) AWS Region.

The database layer can be configured as an Aurora Global Database. This configuration replicates data with no impact on database performance, enables fast local reads with low latency in each region, and provides disaster recovery from region-wide outages.

This solution is optimized for providing high performance for reads. The application will need to be updated to write to the primary Aurora database and send reads to local endpoints.

CORRECT: "Configure an Aurora global database for storage-based cross-Region replication. Use Amazon S3 with cross-Region replication for static content and resources and create Amazon CloudFront distributions" is a correct answer.

CORRECT: "Use Amazon Route 53 with a latency-based routing policy. Create Auto Scaling groups for the web and application tiers and deploy them in multiple global Regions" is also a correct answer.

INCORRECT: "Use logical cross-Region replication to replicate the Aurora MySQL database to a secondary Region. Replace the web servers with Amazon S3. Configure cross-Region replication for the S3 buckets" is incorrect. S3 cannot be used as a replacement for the web servers as they are dynamic websites and S3 can only be used to host a static website.

INCORRECT: "Create Auto Scaling groups for the web and application tiers and deploy them in multiple global Regions. Setup an AWS Direct Connect connection" is incorrect. The AWS Direct Connect (DX) connection does not make sense here. DX connections are used for optimizing network performance between data centers and AWS.

INCORRECT: "Migrate the database from Amazon Aurora to Amazon RDS for MySQL. Replace the web and application tiers with AWS Lambda functions, create an Amazon SQS queue" is incorrect. There is no need to migrate database types or use Lambda functions in place of the EC2 instances. Using Auto Scaling the EC2 instances will provide adequate performance and we also don't know how long processes may run for and if they can be migrated to serverless functions. There is also no mention of lowering costs here and no requirement for decoupling tiers.

References:

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

<https://docs.aws.amazon.com/Route53/latest/DeveloperGuide/routing-policy.html>

Save time with our AWS cheat sheets:

<https://digitalcloud.training/category/aws-cheat-sheets/aws-solutions-architect-professional/aws-database-sap/>

<https://digitalcloud.training/category/aws-cheat-sheets/aws-solutions-architect-professional/aws-networking-content-delivery-sap/>

Question 12: **Incorrect**

A company is running several development projects. Developers are assigned to a single project but move between projects frequently. Each project team requires access to different AWS resources.

Currently, there are projects for serverless, analytics, and database development. The resources used within each project can change over time. Developers require full control over the project they are assigned to and no access to the other projects.

When developers are assigned to a different project or new AWS resources are added, the company wants to minimize policy maintenance.

What type of control policy should a Solutions Architect recommend?

-

Create an IAM role for each project that requires access to AWS resources. Attach an inline policy document to the role that specifies the IAM users that are allowed to assume the role, with full control of the resources that belong to the project. Update the policy document when the set of resources changes, or developers change projects.

-

Create a policy document for each project with specific project tags and allow full control of the resources with a matching tag. Attach the project-specific policy document to the IAM role for that project. Change the role assigned to the developer's IAM user when they change projects. Assign a specific project tag to new resources when they are created.

(Incorrect)

-

Create a customer managed policy document for each project that requires access to AWS resources. Specify full control of the resources that belong to the project. Attach the project-specific policy document to an IAM group. Change the group membership when developers change projects. Update the policy document when the set of resources changes.

(Correct)

-

Create a customer managed policy document for each project that requires access to AWS resources. Specify full control of the resources that belong to the project. Attach the project-specific policy document to the developer's IAM user when they change projects. Update the policy document when the set of resources changes.

Explanation

The correct answer follows the simple principle of using groups to assign permissions to users. A policy document specifying full control to resources for Developers in that group can be created. This represents the most administratively simple approach as group membership and policy updates are centralized to each group/policy document.

CORRECT: "Create a customer managed policy document for each project that requires access to AWS resources. Specify full control of the resources that belong to the project. Attach the project-specific policy document to an IAM group. Change the group membership when developers change projects. Update the policy document when the set of resources changes" is the correct answer.

INCORRECT: "Create a policy document for each project with specific project tags and allow full control of the resources with a matching tag. Attach the project-specific policy document to the IAM role for that project. Change the role assigned to the developer's IAM user when they change projects. Assign a specific project tag to new resources when they are created" is incorrect. This is not as simple as using group membership to control access and requires developers to assume a role rather than interacting directly.

INCORRECT: "Create an IAM role for each project that requires access to AWS resources. Attach an inline policy document to the role that specifies the IAM users that are allowed to assume the role, with full control of the resources that belong to the project. Update the policy document when the set of resources changes, or developers change projects" is incorrect. This solution requires IAM users to assume a role rather than interacting directly. It also requires that the role they assume changes each time they move between projects which would require new instructions to be provided to the users. It's simple just to change group membership and allow them direct access to resources.

INCORRECT: "Create a customer managed policy document for each project that requires access to AWS resources. Specify full control of the resources that belong to the project. Attach the project-specific policy document to the developer's IAM user when they change projects. Update the policy document when the set of resources changes" is incorrect. Inline policies should be avoided as they must be administered on each IAM user account. Attaching a policy to a group and moving users between groups is much simpler.

References:

https://docs.aws.amazon.com/IAM/latest/UserGuide/id_groups.html

Save time with our AWS cheat sheets:

<https://digitalcloud.training/category/aws-cheat-sheets/aws-solutions-architect-professional/security-identity-compliance-sap/>

Question 13: **Correct**

A company plans to migrate physical servers and VMs from an on-premises data center to the AWS Cloud using AWS Migration Hub. The VMs run on a combination of VMware and Hyper-V hypervisors. A Solutions Architect must determine the best services for data collection and discovery. The company has also requested the ability to generate reports from the collected data.

Which solution meets these requirements?



Use the AWS Application Discovery Service agent for data collection on physical servers and Hyper-V. Use the AWS Agentless Discovery Connector for data collection on VMware. Store the collected data in Amazon S3. Query the data with Amazon Athena. Generate reports by using Amazon QuickSight.

(Correct)



Use the AWS Application Discovery Service agent for data collection on physical servers and all VMs. Store the collected data in Amazon Elastic File System (Amazon EFS). Query the data and generate reports with Amazon Athena.



Use the AWS Systems Manager agent for data collection on physical servers. Use the AWS Agentless Discovery Connector for data collection on all VMs. Store, query, and generate reports from the collected data by using Amazon Redshift.

-

Use the AWS Agentless Discovery Connector for data collection on physical servers and all VMs. Store the collected data in Amazon S3. Query the data with S3 Select. Generate reports by using Kibana hosted on Amazon EC2.

Explanation

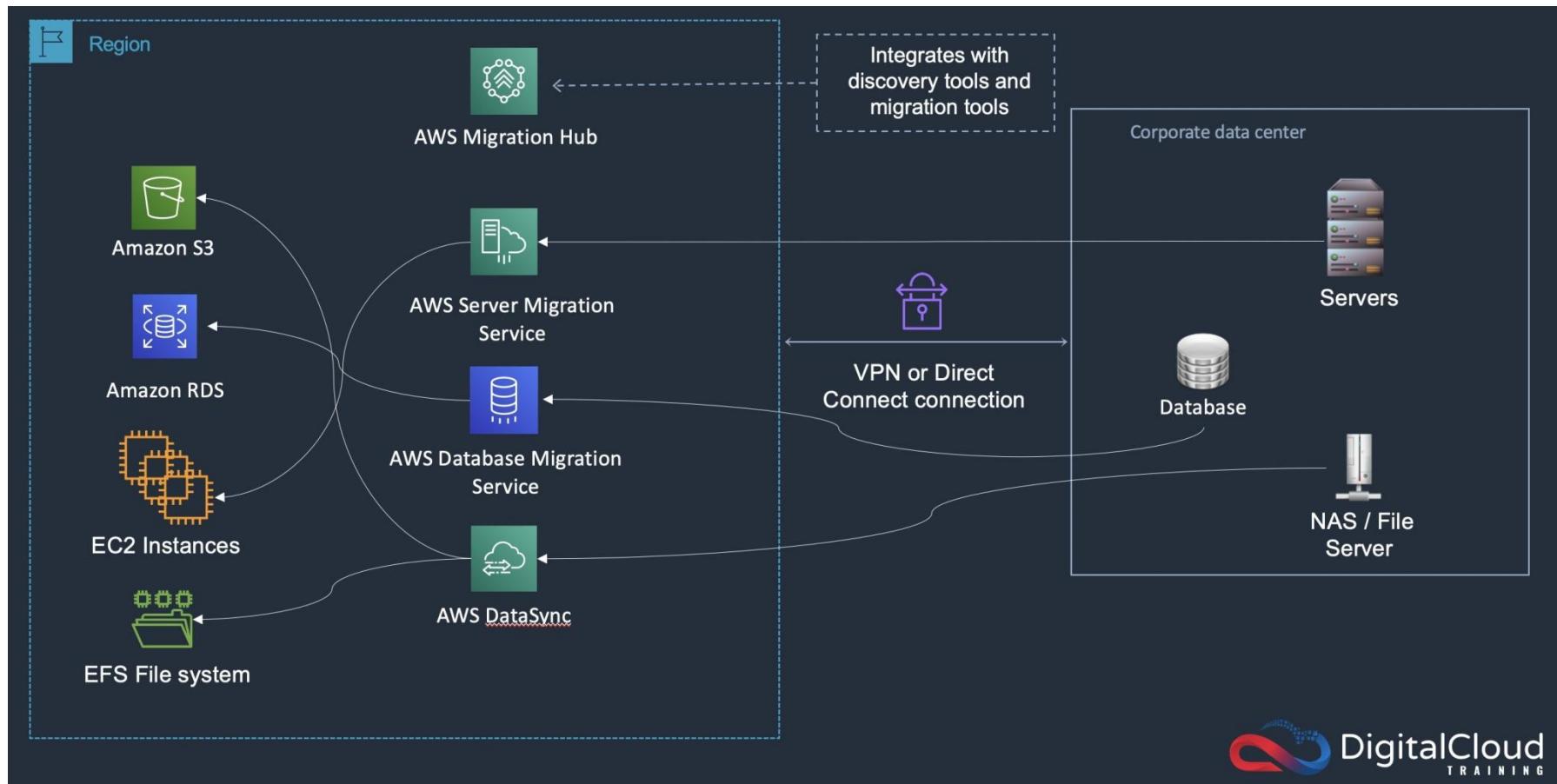
AWS Application Discovery Service helps you plan your migration to the AWS cloud by collecting usage and configuration data about your on-premises servers. Application Discovery Service is integrated with AWS Migration Hub, which simplifies your migration tracking as it aggregates your migration status information into a single console.

Application Discovery Service offers two ways of performing discovery and collecting data about your on-premises servers:

- **Agentless discovery** – Identifies VMs running on VMware.
- **Agent-based discovery** – Used for physical servers and VMs running on Hyper-V.

For reporting the data can be saved to Amazon S3 and queried using Amazon Athena and Amazon QuickSight.

Migration Hub can then be used to manage the migration process using AWS DMS.



CORRECT: "Use the AWS Application Discovery Service agent for data collection on physical servers and Hyper-V. Use the AWS Agentless Discovery Connector for data collection on VMware. Store the collected data in Amazon S3. Query the data with Amazon Athena. Generate reports by using Amazon QuickSight" is the correct answer.

INCORRECT: "Use the AWS Agentless Discovery Connector for data collection on physical servers and all VMs. Store the collected data in Amazon S3. Query the data with S3 Select. Generate reports by using Kibana hosted on Amazon EC2" is incorrect. You cannot use the agentless connector on physical servers or Hyper-V VMs.

INCORRECT: "Use the AWS Application Discovery Service agent for data collection on physical servers and all VMs. Store the collected data in Amazon Elastic File System (Amazon EFS). Query the data and generate reports with Amazon Athena" is incorrect. It is more efficient to use the agentless connector on VMware VMs. Athena cannot query data in EFS.

INCORRECT: "Use the AWS Systems Manager agent for data collection on physical servers. Use the AWS Agentless Discovery Connector for data collection on all VMs. Store, query, and generate reports from the collected data by using Amazon Redshift" is incorrect. You cannot use the agentless discovery connector on Hyper-V VMs. RedShift will be an expensive solution and requires application components to read/write data.

References:

<https://docs.aws.amazon.com/application-discovery/latest/userguide/what-is-appdiscovery.html>

<https://aws.amazon.com/quicksight/>

Save time with our AWS cheat sheets:

<https://digitalcloud.training/category/aws-cheat-sheets/aws-solutions-architect-professional/aws-migration-sap/>

<https://digitalcloud.training/category/aws-cheat-sheets/aws-solutions-architect-professional/aws-analytics-sap/>

Question 14: **Correct**

A company needs to host a highly available and secure image processing application in AWS. Their VPC architecture consists of a public and a private subnet within an Amazon VPC traversing two Availability Zones.

The application is hosted on Amazon EC2 instances in the private subnet. The application needs to communicate with the internet via two NAT gateways and uses an Application Load Balancer in the public subnet. Images are stored in an Amazon S3 bucket which average around 1 TB in new objects per day.

A solutions architect must reduce the associated cost of the solution and reduce manual effort while maintaining security.

How can this be accomplished?

Move the EC2 instances to the public subnets. Remove the NAT gateways.

Attach an Amazon Elastic File System (Amazon EFS) volume to the EC2 instances and host the images on this EFS volume.

Set up an S3 gateway VPC endpoint in the VPC. Attach an endpoint policy to the endpoint to allow the required actions on the S3 bucket.

(Correct)

Use NAT instances in place of the NAT gateways. In the VPC route table, create a route from the private subnets to the NAT instances.

Explanation

There are already two NAT gateways in place but Amazon S3 and DynamoDB come with the option to place gateway endpoints in the VPC which provide reliable connectivity to Amazon S3 without requiring an internet gateway or a NAT device for your VPC. There is no additional charge for using gateway endpoints and this is a secure method of connecting to these service endpoints without using public IP addresses. The cost of the solution can then be reduced as the NAT gateways would no longer be needed.

CORRECT: "Set up an S3 gateway VPC endpoint in the VPC. Attach an endpoint policy to the endpoint to allow the required actions on the S3 bucket" is the correct answer (as explained above.)

INCORRECT: " Use NAT instances in place of the NAT gateways. In the VPC route table, create a route from the private subnets to the NAT instances." is incorrect.

Including NAT instances would have negative impact as they require maintenance and cost is associated with them as well.

INCORRECT: "Move the EC2 instances to the public subnets. Remove the NAT gateways" is incorrect.

Moving an instance to public subnet poses additional security threats and is not recommended. Instances should run in private subnets when possible.

INCORRECT: "Attach an Amazon Elastic File System (Amazon EFS) volume to the EC2 instances and host the images on the EFS volume" is incorrect.

Since the question doesn't mention file shares, but rather S3 (object storage), EFS can be eliminated.

References:

<https://docs.aws.amazon.com/vpc/latest/privatelink/vpc-endpoints-s3.html>

Save time with our AWS cheat sheets:

<https://digitalcloud.training/amazon-vpc/>

Question 15: **Incorrect**

A Solutions Architect must enable an AWS CloudHSM M of N access control—also named a quorum authentication mechanism—to allow security officers to make administrative changes to a hardware security module (HSM). The new security policy states that at least two of the four security officers must authorize any administrative changes to CloudHSM. This is the first time this configuration has been setup. Which steps must be taken to enable quorum authentication (Select TWO.)

-

Using the cloudmgmt_util command line tool, enable encrypted communication, login as a CO, and set the Quorum minimum value to two using the setMValue command.

(Correct)

-

Edit the cloudmgmt_client.cfg document to import a key and register the key for signing.



- Using the `cloudhsm_mgmt_util` command line tool, enable encrypted communication, login as a CO, and get a Quorum token with the `getToken` command.

(Incorrect)



- Use AWS IAM to create a policy that requires a minimum of three crypto officers (COs) to configure the minimum number of approvals required to perform HSM user management operations.

(Incorrect)



- Using the `cloudhsm_mgmt_util` command line tool, enable encrypted communication, login as a CO, and register a key for signing with the `registerMofnPubKey` command.

(Correct)

Explanation

The first time setup for M of N authentication involves creating and registering a key for signing and setting the minimum value on the HSM. This involves the following high-level steps:

- To use quorum authentication, each CO must create an asymmetric key for signing (a *signing key*). This is done outside of the HSM. Keys can be personal keys or public keys.
- A CO must log in to the HSM and then set the *quorum minimum value*, also known as the *m value*. This is the minimum number of CO approvals that are required to perform HSM user management operations. Any CO on the HSM can set the quorum minimum value, including COs that have not registered a key for signing.

CORRECT: "Using the cloudhsm_mgmt_util command line tool, enable encrypted communication, login as a CO, and register a key for signing with the registerMofnPubKey command" is a correct answer.

CORRECT: "Using the cloudhsm_mgmt_util command line tool, enable encrypted communication, login as a CO, and set the Quorum minimum value to two using the setMValue command" is a correct answer.

INCORRECT: "Using the cloudhsm_mgmt_util command line tool, enable encrypted communication, login as a CO, and get a Quorum token with the getToken command" is incorrect. The getToken command is used by a CO to get a token after the quorum authentication has been setup successfully.

INCORRECT: "Use AWS IAM to create a policy that requires a minimum of three crypto officers (COs) to configure the minimum number of approvals required to perform HSM user management operations" is incorrect. IAM is not used to configure the quorum minimum value.

INCORRECT: "Edit the cloudhsm_client.cfg document to import a key and register the key for signing" is incorrect. This document is used for specifying client-side synchronization for keys and is not related to setting up quorum authentication.

References:

<https://docs.aws.amazon.com/clouhdsm/latest/userguide/quorum-authentication-crypto-officers.html#quorum-crypto-officers-use-token>

<https://docs.aws.amazon.com/clouhdsm/latest/userguide/quorum-authentication-crypto-officers-first-time-setup.html>

Save time with our AWS cheat sheets:

<https://digitalcloud.training/category/aws-cheat-sheets/aws-solutions-architect-professional/security-identity-compliance-sap/>

Question 16: **Correct**

A company runs an eCommerce web application on a pair of Amazon EC2 instances behind an Application Load Balancer. The application stores data in an Amazon DynamoDB table. Traffic has been increasing with some major sales events and read and write traffic has slowed down considerably over the busiest periods.

Which option provides a scalable application architecture to handle peak traffic loads with the LEAST development effort?

-
-

Use Auto Scaling groups for the web application and use DynamoDB auto scaling.

(Correct)

-
-

Use Auto Scaling groups for the web application and use Amazon Simple Queue Service (Amazon SQS) and an AWS Lambda function to write to DynamoDB.

-
-

Use AWS Lambda for the web application. Configure DynamoDB to use global tables.

-
-

Use AWS Lambda for the web application. Increase the read and write capacity of DynamoDB.

Explanation

This is a simple case of needing to add elasticity to the application. The question specifically states that the chosen option must incur the least development effort. Therefore, the best option is to simply use Amazon EC2 Auto Scaling for the web application and enable auto scaling for DynamoDB.

This solution provides a simple way to enable elasticity and does not require any refactoring of the application or updates to code.

CORRECT: "Use Auto Scaling groups for the web application and use DynamoDB auto scaling" is the correct answer.

INCORRECT: "Use Auto Scaling groups for the web application and use Amazon Simple Queue Service (Amazon SQS) and an AWS Lambda function to write to DynamoDB" is incorrect. In this scenario it would be simpler and require less development effort to use Auto Scaling for both layers.

INCORRECT: "Use AWS Lambda for the web application. Increase the read and write capacity of DynamoDB" is incorrect. This requires major updates to the application code which is more development effort.

INCORRECT: "Use AWS Lambda for the web application. Configure DynamoDB to use global tables" is incorrect. This requires major updates to the application code which is more development effort.

References:

<https://aws.amazon.com/ec2/autoscaling/>

<https://docs.aws.amazon.com/amazondynamodb/latest/developerguide/AutoScaling.html>

Save time with our AWS cheat sheets:

<https://digitalcloud.training/category/aws-cheat-sheets/aws-solutions-architect-professional/aws-database-sap/>

Question 17: **Correct**

A company is running a two-tier web-based application in an on-premises data center. The application layer consists of a single server running a stateless application. The application connects to a PostgreSQL database running on a separate server. A Solutions Architect is planning a migration to AWS. The company requires that the application and database layer must be highly available across three availability zones.

Which solution will meet the company's requirements?

-

Create an Auto Scaling group of Amazon EC2 instances across three availability zones behind a Network Load Balancer. Create an Amazon Aurora PostgreSQL database in one AZ with storage auto scaling enabled.

-

Create an Auto Scaling group of Amazon EC2 instances across three availability zones behind a Network Load Balancer. Create an Amazon RDS Multi-AZ PostgreSQL database in one AZ and standby instances in two more AZs.



Create an Auto Scaling group of Amazon EC2 instances across three availability zones behind an Application Load Balancer. Create an Amazon Aurora PostgreSQL database in one AZ and add Aurora Replicas in two more AZs.

(Correct)

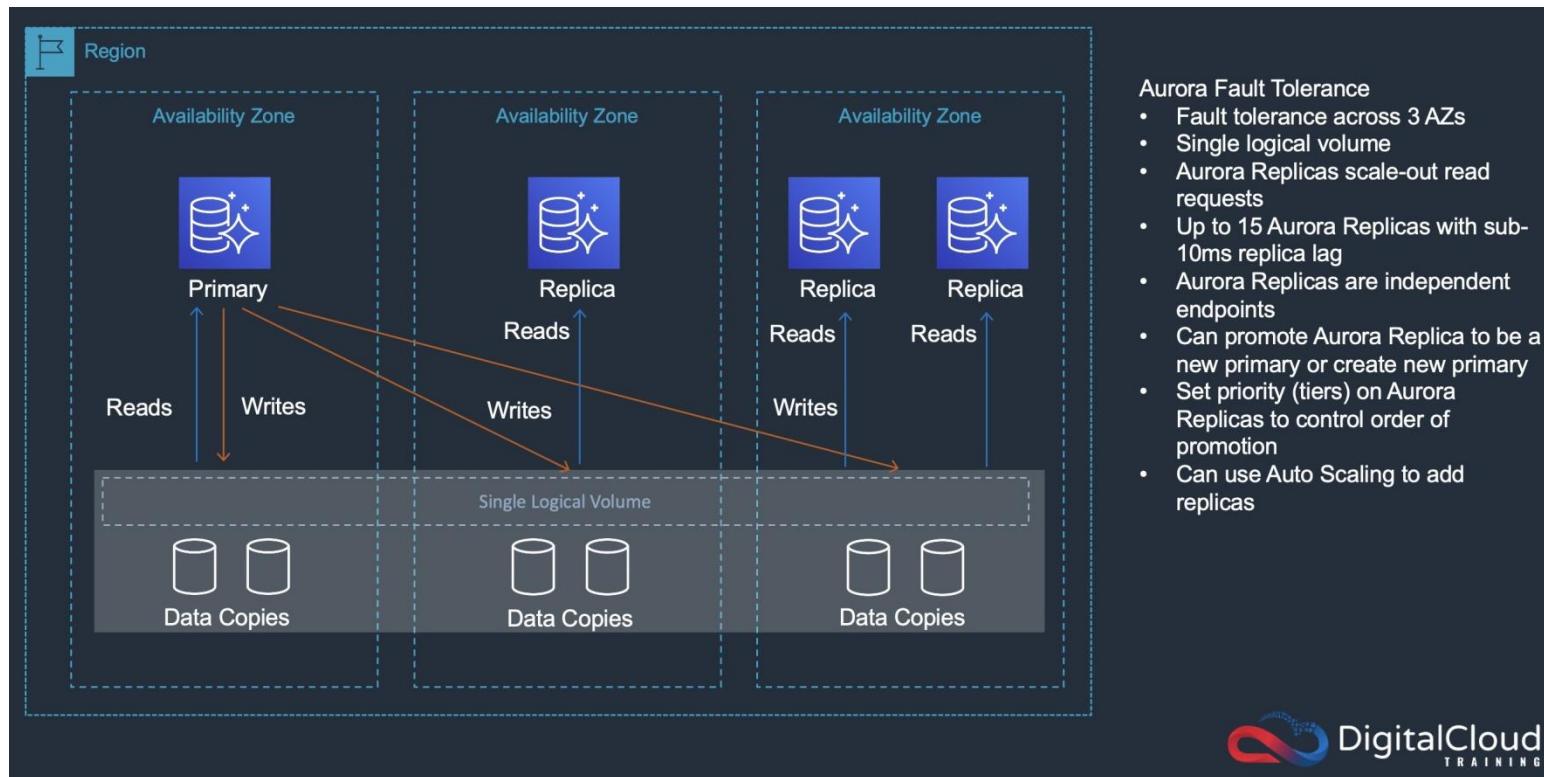


Create an Auto Scaling group of Amazon EC2 instances across three availability zones behind an Application Load Balancer. Create an Amazon Aurora Global database.

Explanation

All answers include an ASG with Auto Scaling across three AZs. An Application Load Balancer is better suited to this particular use case but a Network Load Balancer could also be used. The key to answering correctly is to choose a storage solution that meets the requirement of high availability across three Availability Zones (AZs).

With Amazon Aurora you can have one writer instance and up to 15 Aurora Replicas. Aurora Replicas are read-only but are automatically promoted to be writer instances in the event of a failure. Therefore, for this solution we can have a single Aurora writer instance and two Aurora Replicas, each in separate AZs.



CORRECT: "Create an Auto Scaling group of Amazon EC2 instances across three availability zones behind an Application Load Balancer. Create an Amazon Aurora PostgreSQL database in one AZ and add Aurora Replicas in two more AZs" is the correct answer.

INCORRECT: "Create an Auto Scaling group of Amazon EC2 instances across three availability zones behind a Network Load Balancer. Create an Amazon RDS Multi-AZ PostgreSQL database in one AZ and standby instances in two more AZs" is incorrect. With RDS you can only have one standby instance.

INCORRECT: "Create an Auto Scaling group of Amazon EC2 instances across three availability zones behind an Application Load Balancer. Create an Amazon Aurora Global database" is incorrect. Global database is used for cross-Region databases which is not required in this scenario.

INCORRECT: "Create an Auto Scaling group of Amazon EC2 instances across three availability zones behind a Network Load Balancer. Create an Amazon Aurora PostgreSQL database in one AZ with storage auto scaling enabled" is incorrect. Storage auto scaling is enabled for all Aurora databases and does not provide high availability for the database, it provides scaling for the storage.

References:

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

Save time with our AWS cheat sheets:

<https://digitalcloud.training/category/aws-cheat-sheets/aws-solutions-architect-professional/aws-storage-sap/>

Question 18: **Incorrect**

An Amazon RDS database was created with encryption enabled using an AWS managed CMK. The database has been reclassified and no longer requires encryption. How can a Solutions Architect unencrypt the database with the LEAST operational overhead?

-

Create an unencrypted snapshot of the DB instance and create a new unencrypted DB instance from the snapshot.

(Incorrect)

-

Create an unencrypted read replica of the encrypted DB instance and then promote the read replica to primary.

-

Disable encryption by running the CreateDBInstnace API operation and setting the StorageEncrypted parameter to false.

-

Export the data from the DB instance and import the data into an unencrypted DB instance.

(Correct)

Explanation

The only way to unencrypt an encrypted database is to export the data and import the data into another DB instance. You cannot create unencrypted snapshots of encrypted DB instances and you cannot create unencrypted read replicas of an encrypted DB instance.

You also cannot modify the encrypted status of an existing DB instance using the API, CLI, or AWS Management Console.

CORRECT: "Export the data from the DB instance and import the data into an unencrypted DB instance" is the correct answer.

INCORRECT: "Create an unencrypted snapshot of the DB instance and create a new unencrypted DB instance from the snapshot" is incorrect as explained above.

INCORRECT: "Create an unencrypted read replica of the encrypted DB instance and then promote the read replica to primary" is incorrect as explained above.

INCORRECT: "Disable encryption by running the CreateDBInstnace API operation and setting the StorageEncrypted parameter to false" is incorrect as explained above.

References:

<https://docs.aws.amazon.com/AmazonRDS/latest/UserGuide/Overview.Encryption.html>

Save time with our AWS cheat sheets:

<https://digitalcloud.training/category/aws-cheat-sheets/aws-solutions-architect-professional/aws-database-sap/>

Question 19: **Incorrect**

A company is closing an on-premises data center and needs to move some business applications to AWS. There are over 100 applications that run on virtual machines in the data center. The applications are simple PHP, Java, Ruby, and Node.js web applications. The applications are not developed and are not heavily utilized.

A Solutions Architect must determine the best approach to migrate these applications to AWS with the LOWEST operational overhead.

Which method best fits these requirements?



Use AWS SMS to create an AMI for each virtual machine, run the AMI on Amazon EC2.

(Incorrect)



Use Amazon EBS cross-Region replication to create an AMI for each application, run the AMI on Amazon EC2.



Deploy each application to a single-instance AWS Elastic Beanstalk environment without a load balancer.

(Correct)



Refactor the applications to Docker containers and deploy them to an Amazon ECS cluster behind an Application Load Balancer.
Explanation

The simplest option is to upload the application code to Elastic Beanstalk. This will result in a managed environment that runs on Amazon EC2 instances. Elastic Beanstalk is best suited for running web applications that are developed with Java, .NET, PHP, Node.js, Python, Ruby, Go, and Docker.

A load balancer should not be used as there is only a single instance of each application and a load balancer would not offer many advantages (and would increase the cost).

CORRECT: "Deploy each application to a single-instance AWS Elastic Beanstalk environment without a load balancer" is the correct answer.

INCORRECT: "Use AWS SMS to create an AMI for each virtual machine, run the AMI on Amazon EC2" is incorrect. This would work but an operationally simpler approach would be to take the application code and deploy it to Elastic Beanstalk.

INCORRECT: "Refactor the applications to Docker containers and deploy them to an Amazon ECS cluster behind an Application Load Balancer" is incorrect. This requires refactoring the application which entails operational overhead. Also, with over 100 single-container applications behind a single ALB, requests would be randomly distributed and not directed to the correct application. Complex path-based routing and target group configurations may be able to resolve this but it gets very complex with very little advantage. Better to use Route 53 to direct traffic to the correct containers.

INCORRECT: "Use Amazon EBS cross-Region replication to create an AMI for each application, run the AMI on Amazon EC2" is incorrect. You cannot use EBS to create an AMI from an on-premises virtual machine, use AWS SMS or VM Import/Export instead.

References:

<https://aws.amazon.com/elasticbeanstalk/details/>

Save time with our AWS cheat sheets:

<https://digitalcloud.training/category/aws-cheat-sheets/aws-solutions-architect-professional/aws-compute-sap/>

Question 20: **Correct**

A company runs a single application in an AWS account. The application uses an Auto Scaling Group of Amazon EC2 instances with a combination of Reserved Instances (RIs) and On-Demand instances. To maintain cost-effectiveness the RIs should cover 70% of the workload. The solution should include the ability to alert the DevOps team if coverage drops below the 70% threshold.

Which set of steps should a Solutions Architect take to create the report and alert the DevOps team?

• Use AWS Cost Explorer to configure a report for RI utilization and set the utilization target to 70%. Configure an alert that notifies the DevOps team.

• Use AWS Budgets to create a budget for RI coverage and set the threshold to 70%. Configure an alert that notifies the DevOps team.

(Correct)

• Use the AWS Billing and Cost Management console to create a reservation budget for RI utilization, set the utilization to 70%. Configure an alert that notifies the DevOps team.

• Use AWS Cost Explorer to create a budget for RI coverage and set the threshold to 70%. Configure an alert that notifies the DevOps team.

Explanation

AWS Budgets allows customers to monitor how much of their Amazon EC2 instance usage is covered by reservations and to receive alerts when coverage falls below a specified threshold.

Reserved Instance (RI) coverage tracks the number of running instance hours that are covered by RIs, and can be measured over a daily, monthly, quarterly or yearly cadence. For example, you can monitor your RI coverage either at an aggregate level (e.g., monthly coverage of your entire Amazon EC2 RI fleet) or at a more granular level of detail (e.g., monthly coverage of Amazon RDS db.r3.large instances running in US East region).

You can then define up to five notifications per budget. Each notification can be sent to ten email subscribers and broadcast to an Amazon SNS topic of your choice.

Note: Cost Allocation Tags are not required for monitoring RI coverage. However, they are useful when you need to include detailed cost reporting for your EC2 instances.

Specify your Reservation budget

Track the RI Utilization or RI Coverage associated with your reservations. This budget supports Amazon EC2, RDS, Redshift, ElastiCache and Elasticsearch reservation models.

Reservation budget type

- RI Utilization
- RI Coverage

Service

EC2-Instances (Elastic Com... ▾)

Coverage threshold

70

%

Last month's coverage 0%

CORRECT: "Use AWS Budgets to create a budget for RI coverage and set the threshold to 70%. Configure an alert that notifies the DevOps team" is the correct answer.

INCORRECT: "Use AWS Cost Explorer to configure a report for RI utilization and set the utilization target to 70%. Configure an alert that notifies the DevOps team" is incorrect. You should use AWS Budgets not Cost Explorer for configuring a threshold of RI usage (should also user *coverage* not *utilization*).

INCORRECT: "Use the AWS Billing and Cost Management console to create a reservation budget for RI utilization, set the utilization to 70%. Configure an alert that notifies the DevOps team" is incorrect. This answer misses the steps required to add cost allocation tags so there is no way to report on usage. RI coverage should be configured, not RI utilization. Utilization is simply the percentage of purchased RI hours that were used by matching instances. In this case we want to monitor coverage which is how much EC2 instance usage is covered by RIs.

INCORRECT: "Use AWS Cost Explorer to create a budget for RI coverage and set the threshold to 70%. Configure an alert that notifies the DevOps team" is incorrect. AWS Budgets should be used which can be found in the AWS Billing and Cost Management Console.

References:

<https://aws.amazon.com/about-aws/whats-new/2017/08/monitor-your-reserved-instance-utilization-by-receiving-alerts-via-aws-budgets>

<https://docs.aws.amazon.com/awsaccountbilling/latest/aboutv2/cost-alloc-tags.html>

Save time with our AWS cheat sheets:

<https://digitalcloud.training/category/aws-cheat-sheets/aws-solutions-architect-professional/aws-compute-sap/>

<https://digitalcloud.training/category/aws-cheat-sheets/aws-solutions-architect-professional/aws-management-tools-sap/>

Question 21: **Correct**

A Solutions Architect has been tasked with migrating an application to AWS. The application includes a desktop client application and web application. The web application has an uptime SLA of 99.5%. The Solutions Architect must re-architect the application to meet or exceed this SLA.

The application contains a MySQL database running on a single virtual machine. The web application uses multiple virtual machines with a load balancer. Remote users complain about slow load times while using this latency-sensitive application.

The Solutions Architect must minimize changes to the application whilst improving the user experience, minimizing costs, and ensuring the availability requirements are met. Which solutions best meets these requirements?



Migrate the database to a MySQL database in Amazon EC2. Host the web application on automatically scaled Amazon ECS containers behind an Application Load Balancer. Allocate an Amazon WorkSpaces WorkSpace for each end user to improve the user experience.



Migrate the database to an Amazon EMR cluster with at least two nodes. Deploy the web application on automatically scaled Amazon ECS containers behind an Application Load Balancer. Use Amazon CloudFront to improve the user experience.



Migrate the database to an Amazon RDS Aurora MySQL configuration. Host the web application on an Auto Scaling configuration of Amazon EC2 instances behind an Application Load Balancer. Use Amazon AppStream 2.0 to improve the user experience.

(Correct)



Migrate the database to an Amazon RDS MySQL Multi-AZ configuration. Host the web application on automatically scaled AWS Fargate containers behind a Network Load Balancer. Use Amazon ElastiCache to improve the user experience.

Explanation

The uptime SLA for Amazon RDS is 99.5%. Therefore, it is not necessary to add a Multi-AZ configuration which will increase the solution cost. For the compute layer this could be containers or EC2 instances. To minimize changes to the application using EC2 instances may be slightly easier, but could work. To solve the user experience issues Amazon AppStream 2.0 should be used.

Amazon AppStream 2.0 is a fully managed non-persistent application and desktop streaming service. You centrally manage your desktop applications on AppStream 2.0 and securely deliver them to any computer. Each end user has a fluid and responsive experience because applications run on virtual machines optimized for specific use cases and each streaming sessions automatically adjust to network conditions.

CORRECT: "Migrate the database to an Amazon RDS Aurora MySQL configuration. Host the web application on an Auto Scaling configuration of Amazon EC2 instances behind an Application Load Balancer. Use Amazon AppStream 2.0 to improve the user experience" is the correct answer.

INCORRECT: "Migrate the database to a MySQL database in Amazon EC2. Host the web application on automatically scaled Amazon ECS containers behind an Application Load Balancer. Allocate an Amazon WorkSpaces WorkSpace for each end user to improve the user experience" is incorrect. An RDS managed service would be better for the database and AppStream 2.0 is a better fit for an optimized desktop application.

INCORRECT: "Migrate the database to an Amazon EMR cluster with at least two nodes. Deploy the web application on automatically scaled Amazon ECS containers behind an Application Load Balancer. Use Amazon CloudFront to improve the user experience" is incorrect. EMR is a hosted Hadoop framework for running analytics on big data and is not suitable for this workload. CloudFront is not well suited to optimizing performance for desktop applications.

INCORRECT: "Migrate the database to an Amazon RDS MySQL Multi-AZ configuration. Host the web application on automatically scaled AWS Fargate containers behind a Network Load Balancer. Use Amazon ElastiCache to improve the user experience" is incorrect. AppStream 2.0 is a better fit for the desktop application. ElastiCache will just improve database query performance.

References:

<https://aws.amazon.com/rds/sla/>

<https://aws.amazon.com/appstream2/>

Save time with our AWS cheat sheets:

<https://digitalcloud.training/category/aws-cheat-sheets/aws-solutions-architect-professional/aws-front-end-web-and-mobile-sap/>

<https://digitalcloud.training/category/aws-cheat-sheets/aws-solutions-architect-professional/aws-database-sap/>

Question 22: **Correct**

An application currently runs on Amazon EC2 instances in a single Availability Zone. A Solutions Architect has been asked to re-architect the solution to make it highly available and secure. The security team has requested that all inbound requests are filtered for common vulnerability attacks and all rejected requests must be sent to a third-party auditing application.

Which solution meets the high availability and security requirements?



Configure a Multi-AZ Auto Scaling group using the application's AMI. Create an Application Load Balancer (ALB) and select the previously created Auto Scaling group as the target. Create an Amazon Kinesis Data Firehose with a destination of the third-party

auditing application. Create a web ACL in WAF. Create an AWS WAF using the WebACL and ALB then enable logging by selecting the Kinesis Data Firehose as the destination. Subscribe to AWS Managed Rules in AWS Marketplace, choosing the WAF as the subscriber.

(Correct)

-

Configure an Application Load Balancer (ALB) and add the EC2 instances as targets. Create a web ACL in WAF. Create an AWS WAF using the web ACL and ALB name and enable logging with Amazon CloudWatch Logs. Use an AWS Lambda function to frequently push the logs to the third-party auditing application.

-

Configure an Application Load Balancer (ALB) along with a target group adding the EC2 instances as targets. Create an Amazon Kinesis Data Firehose with the destination of the third-party auditing application. Create a web ACL in WAF. Create an AWS WAF using the web ACL and ALB then enable logging by selecting the Kinesis Data Firehose as the destination. Subscribe to AWS Managed Rules in AWS Marketplace, choosing the WAF as the subscriber.

-

Configure a Multi-AZ Auto Scaling group using the application's AMI. Create an Application Load Balancer (ALB) and select the previously created Auto Scaling group as the target. Use Amazon Inspector to monitor traffic to the ALB and EC2 instances. Create a web ACL in WAF. Create an AWS WAF using the web ACL and ALB. Use an AWS Lambda function to frequently push the Amazon Inspector report to the third-party auditing application.

Explanation

The high availability requirement can be easily met by ensuring the solution includes deploying EC2 instances across multiple AZs. Answers that don't mention multiple AZs can be quickly eliminated.

For the security requirements, we must use AWS WAF to filter based on common vulnerabilities. An easy way to set this up is to subscribe to the AWS Managed Rules via the marketplace.

The requirement to send rejected request data to a third-party auditing application can be met by configuring logging in AWS WAF to Kinesis Data Firehose. The destination in Firehose can be configured as the third-party auditing application. Kinesis Firehose supports HTTP destinations as well as Datadog, New Relic, and Splunk.

CORRECT: "Configure a Multi-AZ Auto Scaling group using the application's AMI. Create an Application Load Balancer (ALB) and select the previously created Auto Scaling group as the target. Create an Amazon Kinesis Data Firehose with a destination of the third-party auditing application. Create a web ACL in WAF. Create an AWS WAF using the WebACL and ALB then enable logging by selecting the Kinesis Data Firehose as the destination. Subscribe to AWS Managed Rules in AWS Marketplace, choosing the WAF as the subscriber" is the correct answer.

INCORRECT: "Configure a Multi-AZ Auto Scaling group using the application's AMI. Create an Application Load Balancer (ALB) and select the previously created Auto Scaling group as the target. Use Amazon Inspector to monitor traffic to the ALB and EC2 instances. Create a web ACL in WAF. Create an AWS WAF using the web ACL and ALB. Use an AWS Lambda function to frequently push the Amazon Inspector report to the third-party auditing application" is incorrect.

Amazon Inspector can be used to assess instances for exposure; however, it does not actually log request data. Therefore, inspector cannot be used to send data about rejected requests to the third-party auditing application.

INCORRECT: "Configure an Application Load Balancer (ALB) and add the EC2 instances as targets. Create a web ACL in WAF. Create an AWS WAF using the web ACL and ALB name and enable logging with Amazon CloudWatch Logs. Use an AWS Lambda function to frequently push the logs to the third-party auditing application" is incorrect.

This solution does not include high availability as there is no mention of adding the EC2 instances to multiple AZs. WAF sends log data to Kinesis Data Firehose, not CloudWatch Logs.

INCORRECT: "Configure an Application Load Balancer (ALB) along with a target group adding the EC2 instances as targets. Create an Amazon Kinesis Data Firehose with the destination of the third-party auditing application. Create a web ACL in WAF. Create an AWS WAF using the web ACL and ALB then enable logging by selecting the Kinesis Data Firehose as the destination. Subscribe to AWS Managed Rules in AWS Marketplace, choosing the WAF as the subscriber" is incorrect.

This solution does not include high availability as there is no mention of adding the EC2 instances to multiple AZs.

References:

<https://docs.aws.amazon.com/waf/latest/developerguide/logging.html>

<https://docs.aws.amazon.com/firehose/latest/dev/create-destination.html>

Save time with our AWS cheat sheets:

<https://digitalcloud.training/category/aws-cheat-sheets/aws-solutions-architect-professional/security-identity-compliance-sap/>

<https://digitalcloud.training/category/aws-cheat-sheets/aws-solutions-architect-professional/aws-analytics-sap/>

Question 23: **Correct**

A company uses AWS CodePipeline to manage an application that runs on Amazon EC2 instances in an Auto Scaling group. All AWS resources are defined in CloudFormation templates. Application code is stored in an Amazon S3 bucket and installed at launch time using lifecycle hooks with EventBridge and AWS Lambda. Recent changes in the CloudFormation templates have resulted in issues that have caused outages and management require a solution to ensure this situation is not repeated.

What should a Solutions Architect do to reduce the likelihood that future changes in the templates will cause downtime?

-

Use AWS CodeBuild for automated testing. Use CloudFormation changes sets to evaluate changes ahead of deployment. Use AWS CodeDeploy to leverage blue/green deployment patterns.

(Correct)

-

Use AWS CodeBuild to detect and report CloudFormation error conditions when performing deployments. Deploy updates to a separate stack in a test account and use manual test plans to validate the changes.

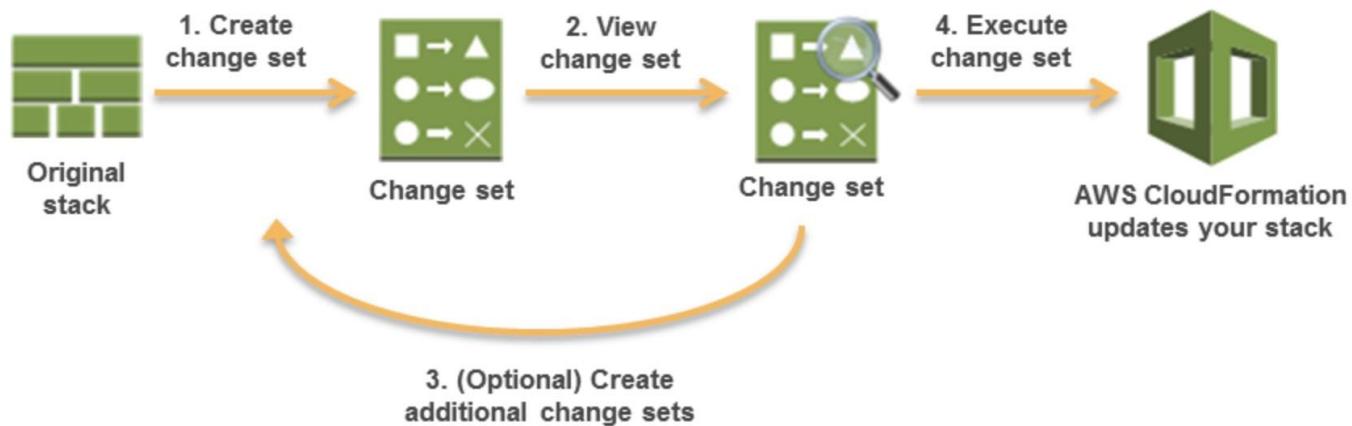
- Move the application code to AWS CodeCommit. Use CodeBuild to validate the application code and automate testing. Use CloudFormation StackSets to deploy updates to different environments to leverage a blue/green deployment pattern.

- Use AWS CodeDeploy and a blue/green deployment pattern with CloudFormation to replace the lifecycle hooks. Gather feedback from users to identify and issues that may require a rollback.

Explanation

AWS CodeBuild can be used to provide automated testing of application code. With CodeBuild you can use Amazon S3 as a source provider. CloudFormation change sets can be used to test updates to the templates before the infrastructure is updated.

Change sets allow you to preview how proposed changes to a stack might impact your running resources, for example, whether your changes will delete or replace any critical resources, AWS CloudFormation makes the changes to your stack only when you decide to execute the change set, allowing you to decide whether to proceed with your proposed changes or explore other changes by creating another change set.



Blue/green deployments mean that the updates are made to a separate application stack and validated before the user traffic is switched across. This provides an easy rollback path if there are any issues identified in the updated application.

CORRECT: "Use AWS CodeBuild for automated testing. Use CloudFormation changes sets to evaluate changes ahead of deployment. Use AWS CodeDeploy to leverage blue/green deployment patterns" is the correct answer.

INCORRECT: "Use AWS CodeBuild to detect and report CloudFormation error conditions when performing deployments. Deploy updates to a separate stack in a test account and use manual test plans to validate the changes" is incorrect. CodeBuild does not detect CloudFormation error conditions, it is used to build and test application code.

INCORRECT: "Move the application code to AWS CodeCommit. Use CodeBuild to validate the application code and automate testing. Use CloudFormation StackSets to deploy updates to different environments to leverage a blue/green deployment pattern" is incorrect. CloudFormation StackSets are used for deploying stacks into different Regions or accounts. Change sets should be used instead for this use case.

INCORRECT: "Use AWS CodeDeploy and a blue/green deployment pattern with CloudFormation to replace the lifecycle hooks. Gather feedback from users to identify and issues that may require a rollback" is incorrect. This doesn't offer a way of validating issues ahead of users being directed to the updated application. Testing should be automated and issues should be identified before users access the application as much as possible.

References:

<https://docs.aws.amazon.com/AWSCloudFormation/latest/UserGuide/using-cfn-updating-stacks-changesets.html>

Save time with our AWS cheat sheets:

<https://digitalcloud.training/category/aws-cheat-sheets/aws-solutions-architect-professional/aws-networking-content-delivery-sap/>

Question 24: **Correct**

A company runs a web application in an on-premises data center in Paris. The application includes stateless web servers behind a load balancer, shared files in a NAS device, and a MySQL database server. The company plans to migrate the solution to AWS and has the following requirements:

- Provide optimum performance for customers.
- Implement elastic scalability for the web tier.
- Optimize the database server performance for read-heavy workloads.
- Reduce latency for users across Europe and the US.
- Design the new architecture with a 99.9% availability SLA.

Which solution should a Solutions Architect propose to meet these requirements while optimizing operational efficiency?

- Use an Application Load Balancer (ALB) in front of an Auto Scaling group of Amazon EC2 instances in one AWS Region and three Availability Zones. Configure an Amazon ElastiCache cluster in front of a Multi-AZ Amazon Aurora MySQL DB cluster. Move the shared files to Amazon EFS. Configure Amazon CloudFront with the ALB as the origin and select a price class that includes the US and Europe.**

(Correct)
- Use an Application Load Balancer (ALB) in front of an Auto Scaling group of Amazon EC2 instances in two AWS Regions and two Availability Zones in each Region. Configure an Amazon ElastiCache cluster in front of a global Amazon Aurora MySQL database. Move the shared files to Amazon EFS. Configure Amazon CloudFront with the ALB as the origin and select a price class that includes the US and Europe. Configure EFS cross-Region replication.**
- Use an Application Load Balancer (ALB) in front of an Auto Scaling group of Amazon EC2 instances in one AWS Region and three Availability Zones. Configure an Amazon DocumentDB table in front of a Multi-AZ Amazon Aurora MySQL DB cluster. Move the**

shared files to Amazon EFS. Configure Amazon CloudFront with the ALB as the origin and select a price class that includes all global locations.



Use an Application Load Balancer (ALB) in front of an Auto Scaling group of Amazon EC2 instances in two AWS Regions and three Availability Zones in each Region. Configure an Amazon ElastiCache cluster in front of a global Amazon Aurora MySQL database. Move the shared files to Amazon FSx with cross-Region synchronization. Configure Amazon CloudFront with the ALB as the origin and a price class that includes the US and Europe.

Explanation

To meet the 99.9% availability SLA a solution in a single Region with Auto Scaling and Load Balancing across multiple AZs is sufficient. To optimize the DB for read-heavy workloads, Amazon ElastiCache can be placed in front of an Aurora MySQL DB. The shared files can be easily moved to an Amazon EFS file system. CloudFront can be used to reduce latency for users in different geographies. In this case US and Europe price classes can be selected in CloudFront and this will cache the content in those locations only which reduces cost.

CORRECT: "Use an Application Load Balancer (ALB) in front of an Auto Scaling group of Amazon EC2 instances in one AWS Region and three Availability Zones. Configure an Amazon ElastiCache cluster in front of a Multi-AZ Amazon Aurora MySQL DB cluster. Move the shared files to Amazon EFS. Configure Amazon CloudFront with the ALB as the origin and select a price class that includes the US and Europe" is the correct answer.

INCORRECT: "Use an Application Load Balancer (ALB) in front of an Auto Scaling group of Amazon EC2 instances in two AWS Regions and two Availability Zones in each Region. Configure an Amazon ElastiCache cluster in front of a global Amazon Aurora MySQL database. Move the shared files to Amazon EFS. Configure Amazon CloudFront with the ALB as the origin and select a price class that includes the US and Europe. Configure EFS cross-Region replication" is incorrect.

There's no such thing as EFS cross-Region replication so the shared files cannot be synchronized that way. There's also no need to have a cross-Region solution to meet a 99.9% availability SLA and there's no mechanism mentioned for directing traffic between Regions.

INCORRECT: "Use an Application Load Balancer (ALB) in front of an Auto Scaling group of Amazon EC2 instances in one AWS Region and three Availability Zones. Configure an Amazon DocumentDB table in front of a Multi-AZ Amazon Aurora MySQL DB cluster. Move

the shared files to Amazon EFS. Configure Amazon CloudFront with the ALB as the origin and select a price class that includes all global locations" is incorrect.

DocumentDB is not a caching engine and cannot be used in front of an Aurora DB. CloudFront should not use a price class that includes all global locations as this will be more costly and is not required in the solution.

INCORRECT: "Use an Application Load Balancer (ALB) in front of an Auto Scaling group of Amazon EC2 instances in two AWS Regions and three Availability Zones in each Region. Configure an Amazon ElastiCache cluster in front of a global Amazon Aurora MySQL database. Move the shared files to Amazon FSx with cross-Region synchronization. Configure Amazon CloudFront with the ALB as the origin and a price class that includes the US and Europe" is incorrect.

Amazon FSx also does not have a feature for cross-Region synchronization. There's also no need to have a cross-Region solution to meet a 99.9% availability SLA and there's no mechanism mentioned for directing traffic between Regions.

References:

<https://docs.aws.amazon.com/AmazonCloudFront/latest/DeveloperGuide/PriceClass.html>

Save time with our AWS cheat sheets:

<https://digitalcloud.training/category/aws-cheat-sheets/aws-solutions-architect-professional/aws-networking-content-delivery-sap/>

Question 25: **Correct**

An e-commerce company has developed a newer version of a shopping application with many new features. But before rolling it out to the public, they want to test the new version incrementally using small incremental deployments. The application is deployed using AWS CloudFormation and uses multiple AWS Lambda functions.

Which solution will meet these requirements?

-

Deploy the application using a new CloudFormation stack. Use an Amazon Route 53 weighted routing policy to distribute the load between the stacks.

-
- Enable versioning of Lambda function to identify each increment. Use the AWS CLI ‘update-function-configuration’ command with the ‘routing-config’ parameter to distribute the load.**
-
- Enable versioning for the AWS Lambda function and associate an alias for every new version. Use the AWS CLI ‘update-alias’ command with the ‘routing-config’ parameter to distribute the load.**

(Correct)

- - Configure AWS CodeDeploy and use CodeDeployDefault.AllAtOnce in the Deployment configuration to distribute the load.**
- Explanation**
Testing a new version of application in gradual traffic shift increments is a very common pattern also known as Canary deployment, where actual traffic is routed to newer version of application in constant increments over time and basis the feedback, new version is rolled out. AWS provides canary deployment for many components and the table below shows the prebuilt options for AWS Lambda:

Deployment configuration	Description
CodeDeployDefault.LambdaCanary10Percent5Minutes	Shifts 10 percent of traffic in the first increment. The remaining 90 percent is deployed five minutes later.
CodeDeployDefault.LambdaCanary10Percent10Minutes	Shifts 10 percent of traffic in the first increment. The remaining 90 percent is deployed 10 minutes later.
CodeDeployDefault.LambdaCanary10Percent15Minutes	Shifts 10 percent of traffic in the first increment. The remaining 90 percent is deployed 15 minutes later.
CodeDeployDefault.LambdaCanary10Percent30Minutes	Shifts 10 percent of traffic in the first increment. The remaining 90 percent is deployed 30 minutes later.
CodeDeployDefault.LambdaLinear10PercentEvery1Minute	Shifts 10 percent of traffic every minute until all traffic is shifted.
CodeDeployDefault.LambdaLinear10PercentEvery2Minutes	Shifts 10 percent of traffic every two minutes until all traffic is shifted.
CodeDeployDefault.LambdaLinear10PercentEvery3Minutes	Shifts 10 percent of traffic every three minutes until all traffic is shifted.
CodeDeployDefault.LambdaLinear10PercentEvery10Minutes	Shifts 10 percent of traffic every 10 minutes until all traffic is shifted.
CodeDeployDefault.LambdaAllAtOnce	Shifts all traffic to the updated Lambda functions at once.

Using these options, with every code base, a separate version of Lambda can be created, and those versions can be tested independently using canary deployments as mentioned above.

CORRECT: "Enable versioning for the AWS Lambda function and associate an alias for every new version. Use the AWS CLI 'update-alias' command with the 'routing-config' parameter to distribute the load" is the correct answer (as explained above.)

INCORRECT: "Deploy the application using a new CloudFormation stack. Use an Amazon Route 53 weighted routing policy to distribute the load between the stacks" is incorrect.

This can work but will need manual intervention. Two parallel stacks would also incur cost based on the resources they hold hence this is not the most optimal option.

INCORRECT: "Enable versioning of Lambda function to identify each increment. Use the AWS CLI 'update-function-configuration' command with the 'routing-config' parameter to distribute the load" is incorrect.

This command provides several parameters which can be passed to a newer version of the Lambda function. This again requires more manual intervention because with each route config, a version will need to be bound and then weight for each version would need to be controlled, hence this is incorrect.

INCORRECT: "Configure AWS CodeDeploy and use CodeDeployDefault.AllAtOnce in the Deployment configuration to distribute the load" is incorrect.

This option attempts to deploy an application revision to as many instances as possible at once. The status of the overall deployment is displayed as Succeeded if the application revision is deployed to one or more of the instances. The status of the overall deployment is displayed as Failed if the application revision is not deployed to any of the instances. Using an example of nine instances, CodeDeployDefault.AllAtOnce attempts to deploy to all nine instances at once. The overall deployment succeeds if deployment to even a single instance is successful. It fails only if deployments to all nine instances fail.

This is not an incremental deployment so is not a correct answer.

References:

<https://docs.aws.amazon.com/lambda/latest/dg/configuration-aliases.html>

<https://aws.amazon.com/blogs/compute/implementing-canary-deployments-of-aws-lambda-functions-with-alias-traffic-shifting/>

Save time with our AWS cheat sheets:

<https://digitalcloud.training/aws-lambda/>

Question 26: **Correct**

A company wants to host a web application on AWS. The application will be used by users around the world. A Solutions Architect has been given the following design requirements:

- Allow the retrieval of data from multiple data sources.
- Minimize the cost of API calls.
- Reduce latency for user access.
- Provide user authentication and authorization and implement role-based access control.
- Implement a fully serverless solution.

How can the Solutions Architect meet these requirements?

- Use Amazon CloudFront with Amazon EC2 to host the web application. Use Amazon API Gateway to build the application APIs. Use AWS Lambda for custom authentication and authorization. Authorize data access by leveraging IAM roles.

- Use Amazon CloudFront with Amazon S3 to host the web application. Use AWS AppSync to build the application APIs. Use Amazon Cognito groups for RBAC. Authorize data access by leveraging Cognito groups in AWS AppSync resolvers.

(Correct)

Use Amazon CloudFront with Amazon FSx to host the web application. Use AWS AppSync to build the application APIs. Use IAM groups for RBAC. Authorize data access by leveraging IAM groups in AWS AppSync resolvers.



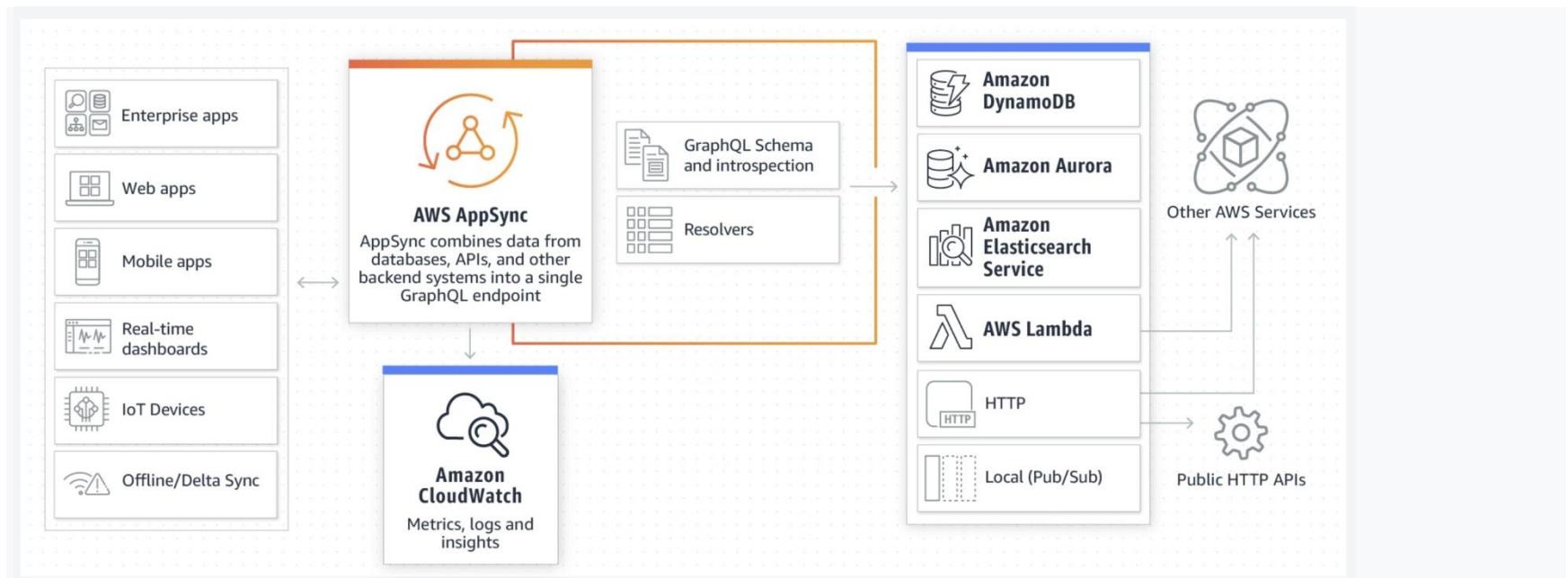
Use Amazon CloudFront with Amazon S3 to host the web application. Use Amazon API Gateway to build the application APIs with AWS Lambda for the custom authorizer. Authorize data access by performing user lookup in AWS Managed Microsoft AD.

Explanation

CloudFront with S3 provides a low-latency solution for global users to access the web application. AWS AppSync can be used to provide a GraphQL API that can be used to query multiple databases, microservices, and APIs (allow the retrieval of data from multiple data sources).

Amazon Cognito Groups can be used to create collections of users to manage their permissions or to represent different types of users. You can assign an AWS Identity and Access Management (IAM) role to a group to define the permissions for members of a group.

AWS AppSync GraphQL resolvers connect the fields in a type's schema to a data source. Resolvers are the mechanism by which requests are fulfilled. Cognito groups can be used with resolvers to provide authorization based on identity.



CORRECT: "Use Amazon CloudFront with Amazon S3 to host the web application. Use AWS AppSync to build the application APIs. Use Amazon Cognito groups for RBAC. Authorize data access by leveraging Cognito groups in AWS AppSync resolvers" is the correct answer.

INCORRECT: "Use Amazon CloudFront with Amazon S3 to host the web application. Use Amazon API Gateway to build the application APIs with AWS Lambda for the custom authorizer. Authorize data access by performing user lookup in AWS Managed Microsoft AD" is incorrect. AppSync is a better fit than using API Gateway due to the requirement to retrieve data from multiple data sources.

INCORRECT: "Use Amazon CloudFront with Amazon FSx to host the web application. Use AWS AppSync to build the application APIs. Use IAM groups for RBAC. Authorize data access by leveraging IAM groups in AWS AppSync resolvers" is incorrect. You cannot point CloudFront at an Amazon FSx file system.

INCORRECT: "Use Amazon CloudFront with Amazon EC2 to host the web application. Use Amazon API Gateway to build the application APIs. Use AWS Lambda for custom authentication and authorization. Authorize data access by leveraging IAM roles" is incorrect. EC2 is not serverless so should not be used in this solution.

References:

<https://docs.aws.amazon.com/appsync/latest/devguide/security-authorization-use-cases.html>

Save time with our AWS cheat sheets:

<https://digitalcloud.training/category/aws-cheat-sheets/aws-solutions-architect-professional/aws-front-end-web-and-mobile-sap/>

Question 27: **Incorrect**

A company runs a two-tier application that uses EBS-backed Amazon EC2 instances in an Auto Scaling group and an Amazon Aurora PostgreSQL database. The company intends to use a pilot light approach for disaster recovery in a different AWS Region. The company has an RTO of 6 hours and an RPO of 24 hours.

Which solution would achieve the requirements with MINIMAL cost?



Use AWS Lambda to create daily EBS and RDS snapshots and copy them to the disaster recovery Region. Use Amazon Route 53 with an active-active failover configuration. Use Amazon EC2 in an Auto Scaling group configured the same as the primary Region.



Use EBS and RDS cross-Region snapshot copy capability to create snapshots in the disaster recovery (DR) Region. Use Amazon Route 53 with an active-active failover configuration. Use Amazon EC2 in an Auto Scaling group with the capacity set to 0 in the disaster recovery Region.



Use EBS cross-region snapshot copy capability to create snapshots in the disaster recovery (DR) Region. Implement an Aurora Replica in the DR Region. Use Amazon Route 53 with an active-passive failover configuration. Use Amazon EC2 in an Auto Scaling group configured the same as the primary Region.

(Incorrect)

-

Use AWS Lambda to create daily EBS snapshots and copy them to the disaster recovery Region. Implement an Aurora Replica in the DR Region. Use Amazon Route 53 with an active-passive failover configuration. Use Amazon EC2 in an Auto Scaling group with the capacity set to 0 in the disaster recovery Region.

(Correct)

Explanation

The correct answer must be the lowest cost option that delivers the RPO and RTO requirements and is in line with a pilot light strategy. Please note the AWS definition of a pilot light scenario:

"In this DR approach, you simply replicate part of your IT structure for a limited set of core services so that the AWS cloud environment seamlessly takes over in the event of a disaster. A small part of your infrastructure is always running simultaneously syncing mutable data (as databases or documents), while other parts of your infrastructure are switched off and used only during testing. Unlike a backup and recovery approach, you must ensure that your most critical core elements are already configured and running in AWS (the pilot light)."

The mutable data in this case is the Aurora database so this should be running in the DR Region. The best way to achieve that is to create an Aurora replica in the DR Region. This increases cost compared to replicating snapshots but is in line with the pilot light strategy.

For the EC2 instances AWS Lambda can be used to automate the replication of snapshots to the DR Region. These can be daily snapshots as the RPO is 24 hours. Auto Scaling can be configured with the capacity set to 0 and Route 53 failover records can be created in an active-passive configuration.

In the event of a disaster it would then be easy to create AMIs from the EC2 snapshots, add them to a launch config and then increase the capacity of the ASG. The Aurora Replica will automatically become a writable database and the entire configuration can be up and running well within the RTO.

CORRECT: "Use AWS Lambda to create daily EBS snapshots and copy them to the disaster recovery Region. Implement an Aurora Replica in the DR Region. Use Amazon Route 53 with an active-passive failover configuration. Use Amazon EC2 in an Auto Scaling group with the capacity set to 0 in the disaster recovery Region" is the correct answer.

INCORRECT: "Use EBS cross-region snapshot copy capability to create snapshots in the disaster recovery (DR) Region. Implement an Aurora Replica in the DR Region. Use Amazon Route 53 with an active-passive failover configuration. Use Amazon EC2 in an Auto Scaling group configured the same as the primary Region" is incorrect.

You cannot create snapshots of an instance in one Region directly into another Region as the wording suggests. You must create the snapshot in the same Region and then copy it across Regions. You would also not configure the ASG the same as in the primary Region as it would be more costly.

INCORRECT: "Use EBS and RDS cross-Region snapshot copy capability to create snapshots in the disaster recovery (DR) Region. Use Amazon Route 53 with an active-active failover configuration. Use Amazon EC2 in an Auto Scaling group with the capacity set to 0 in the disaster recovery Region" is incorrect.

As above, you cannot create snapshots of an instance in one Region directly into another Region as the wording suggests. You must create the snapshot in the same Region and then copy it across Regions.

INCORRECT: "Use AWS Lambda to create daily EBS and RDS snapshots and copy them to the disaster recovery Region. Use Amazon Route 53 with an active-active failover configuration. Use Amazon EC2 in an Auto Scaling group configured the same as the primary Region" is incorrect.

You would not configure the ASG the same as in the primary Region as it would be more costly.

References:

<https://aws.amazon.com/blogs/publicsector/rapidly-recover-mission-critical-systems-in-a-disaster/>

<https://docs.aws.amazon.com/Route53/latest/DeveloperGuide/dns-failover-types.html>

Save time with our AWS cheat sheets:

<https://digitalcloud.training/category/aws-cheat-sheets/aws-solutions-architect-professional/aws-storage-sap/>

<https://digitalcloud.training/category/aws-cheat-sheets/aws-solutions-architect-professional/aws-networking-content-delivery-sap/>

Question 28: **Correct**

A company requires an application in which employees can log expense claims for processing. The expense claims are typically submitted each week on a Friday. The application must store data in a format that will allow the finance team to be able to run end of month reports. The solution should be highly available and must scale seamlessly based on demand.

Which combination of solution options meets these requirements with the LEAST operational overhead? (Select TWO.)

-

Deploy the application front end to an Amazon S3 bucket served by Amazon CloudFront. Deploy the application backend using Amazon API Gateway with an AWS Lambda proxy integration.

(Correct)

-

Deploy the application in a container using Amazon ECS behind an Application Load Balancer. Use Service Auto Scaling and schedule additional capacity ahead of peak usage periods.

-

Store the expense claim data in Amazon EMR. Use Amazon QuickSight to generate the reports using Amazon EMR as the data source.

-

Store the expense claim data in Amazon S3. Use Amazon Athena and Amazon QuickSight to generate the reports using Amazon S3 as the data source.

(Correct)

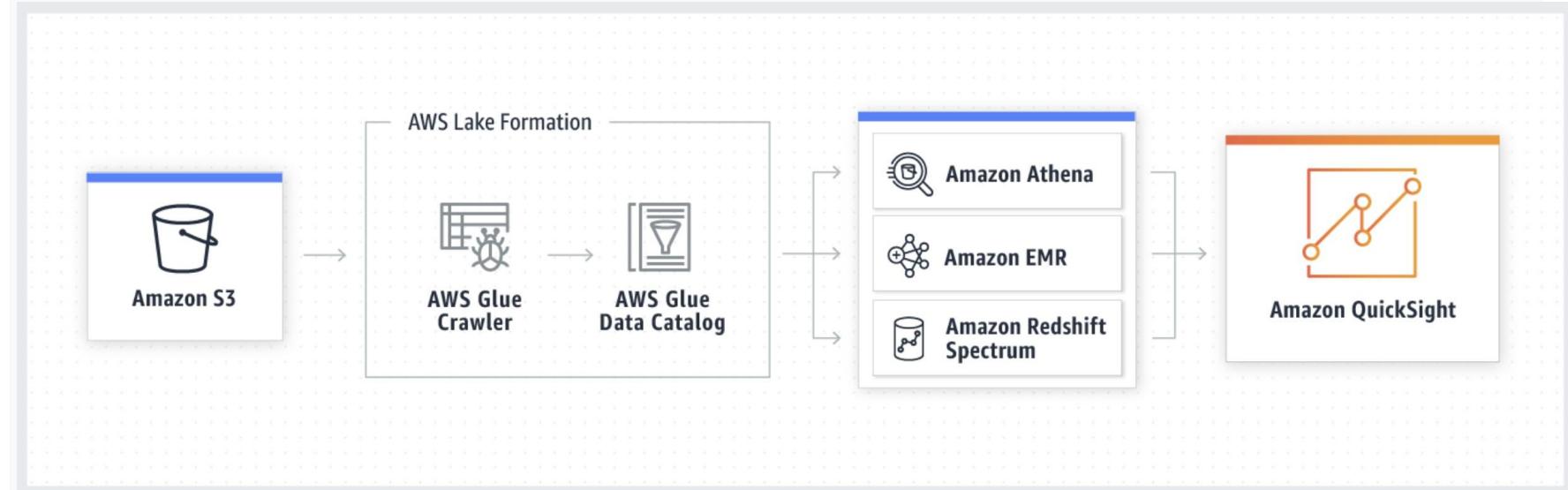
-

Deploy the application to Amazon EC2 On-Demand Instances behind an Application Load Balancer. Use Amazon EC2 Auto Scaling and schedule additional capacity ahead of peak usage periods.

Explanation

Using serverless technologies is the best option to reduce the operational overhead. Therefore, using S3 with CloudFront for the front-end and API Gateway with Lambda for the back-end offers seamless scalability with low maintenance.

For the data storage layer S3 can again be used in combination with Athena and QuickSight. This meets the requirement of storing the data in a format that can be used to generate reports.



CORRECT: "Deploy the application front end to an Amazon S3 bucket served by Amazon CloudFront. Deploy the application backend using Amazon API Gateway with an AWS Lambda proxy integration" is a correct answer.

CORRECT: "Store the expense claim data in Amazon S3. Use Amazon Athena and Amazon QuickSight to generate the reports using Amazon S3 as the data source" is a correct answer.

INCORRECT: "Deploy the application in a container using Amazon ECS behind an Application Load Balancer. Use Service Auto Scaling and schedule additional capacity ahead of peak usage periods" is incorrect. Amazon ECS requires instances to be managed (except when using Fargate which isn't specified), so this option requires more operational overhead.

INCORRECT: "Deploy the application to Amazon EC2 On-Demand Instances behind an Application Load Balancer. Use Amazon EC2 Auto Scaling and schedule additional capacity ahead of peak usage periods" is incorrect. Amazon EC2 requires instance maintenance so is another option that requires more operational overhead.

INCORRECT: "Store the expense claim data in Amazon EMR. Use Amazon QuickSight to generate the reports using Amazon EMR as the data source" is incorrect. Amazon EMR can be used with QuickSight but storing data in EMR is going to be much more expensive and require more maintenance than using S3 with Athena.

References:

<https://aws.amazon.com/quicksight/>

Save time with our AWS cheat sheets:

<https://digitalcloud.training/category/aws-cheat-sheets/aws-solutions-architect-professional/aws-analytics-sap/>

Question 29: **Incorrect**

A company is running an application on an on-premises VMware cluster that must be migrated to an Amazon EC2 instance. While migrating, they wish to preserve the software and configuration settings.

What is the best strategy to meet these requirements?

-

Use the VMware vSphere client to export the application as an image in Open Virtualization Format (OVF) format. Create an Amazon S3 bucket to store the image in the destination AWS Region. Create and apply an IAM role for VM Import. Use the AWS CLI to run the EC2 import command.

(Correct)

-

Configure the AWS DataSync agent to start replicating the data store to Amazon FSX for Windows File Server. Use the SMB share to host the VMware data store. Use VM Import/Export to move the VMs to Amazon EC2.

-

Configure AWS Storage Gateway for files service to export a Common Internet File System (CIFS) share. Create a backup copy to the shared folder. Sign into the AWS Management Console and create an AMI from the backup copy. Launch an EC2 instance that is based on the AMI.

-

Create a managed-instance activation for a hybrid environment in AWS Systems Manager. Download and install Systems Manager Agent on the on-premises VM. Register the VM with Systems Manager to be a managed instance. Use AWS Backup to create a snapshot of the VM and create an AMI. Launch an EC2 instance that is based on the AMI.

(Incorrect)

Explanation

You can use VM Import/Export to import virtual machine (VM) images from your virtualization environment to Amazon EC2 as Amazon Machine Images (AMI), which you can use to launch instances. Subsequently, you can export the VM images from an instance back to your virtualization environment. This enables you to leverage your investments in the VMs that you have built to meet your IT security, configuration management, and compliance requirements by bringing them into Amazon EC2. In this case the settings are preserved during the migration as per the requirement.

CORRECT: "Use the VMware vSphere client to export the application as an image in Open Virtualization Format (OVF) format. Create an Amazon S3 bucket to store the image in the destination AWS Region. Create and apply an IAM role for VM Import. Use the AWS CLI to run the EC2 import command" is the correct answer (as explained above.)

INCORRECT: "Configure the AWS DataSync agent to start replicating the data store to Amazon FSX for Windows File Server. Use the SMB share to host the VMware data store. Use VM Import/Export to move the VMs to Amazon EC2." is incorrect.

Since the question is specifically targeting a VMware instance, this option is incorrect, this would have been applicable only in case of a Windows based instance where the data on a file share must be migrated.

INCORRECT: "Configure AWS Storage Gateway for files service to export a Common Internet File System (CIFS) share. Create a backup copy to the shared folder. Sign into the AWS Management Console and create an AMI from the backup copy. Launch an EC2 instance that is based on the AMI" is incorrect.

Storage Gateway is used for creating hybrid storage solutions and is unsuitable for migrating a VM into Amazon EC2.

INCORRECT: "Create a managed-instance activation for a hybrid environment in AWS Systems Manager. Download and install Systems Manager Agent on the on-premises VM. Register the VM with Systems Manager to be a managed instance. Use AWS Backup to create a snapshot of the VM and create an AMI. Launch an EC2 instance that is based on the AMI." is incorrect.

This may be a suitable solution if the requirement is to maintain a hybrid environment where the instance running on-premises. As soon as the backup is complete and we spin up an instance in EC2, the entire setup before this step becomes redundant and might not be needed at a later stage.

References:

<https://docs.aws.amazon.com/vm-import/latest/userguide/vmimport-image-import.html>

Save time with our AWS cheat sheets:

<https://digitalcloud.training/aws-migration-services/>

Question 30: **Incorrect**

A media publishing company has created an online bookstore which gives users access to books and other reference material. These materials can be downloaded by users and new materials can also be uploaded on the portal. According to company requirements, all data must be encrypted in transit and at rest. A solutions architect is building the solution by using Amazon S3 and Amazon CloudFront.

Which combination of steps will meet the encryption requirements? (Select THREE.)

-

Configure redirection of HTTP requests to HTTPS requests in CloudFront.

(Correct)

-

For the read and write operations in the S3 ACLs, add condition "aws:SecureTransport": "true".

(Incorrect)

-

Create a bucket policy that denies any unencrypted operations in the S3 bucket that the web application uses.

(Correct)

-

Turn on the S3 server-side encryption for the S3 bucket in use.

(Correct)

-

Use the RequireSSL option in the creation of presigned URLs for the S3 bucket that the web application uses

(Incorrect)

-

Configure encryption at rest on CloudFront by using server-side encryption with AWS KMS keys (SSE-KMS).

Explanation

For object encryption at rest, you can set the default encryption behavior on an Amazon S3 bucket so that all objects are encrypted when they are stored in the bucket. The objects are encrypted using server-side encryption with either Amazon S3-managed keys (SSE-S3) or AWS Key Management Service (AWS KMS) keys.

To deny unencrypted objects, "s3:x-amz-server-side-encryption" can be added which allows only encrypted object upload and can restrict to a specific KMS key as well.

Amazon CloudFront can use 301 response code to redirect HTTP requests to HTTPS and allows only secured traffic.

CORRECT: "Turn on S3 server-side encryption for the S3 bucket that the web application uses" is a correct answer (as explained above.)

CORRECT: "Create a bucket policy that denies any unencrypted operations in the S3 bucket that the web application uses" is also a correct answer (as explained above.)

CORRECT: "Configure redirection of HTTP requests to HTTPS requests in CloudFront" is also a correct answer (as explained above.)

INCORRECT: "For the read and write operations in the S3 ACLs , add condition "aws:SecureTransport": "true"" is incorrect.

To only allow HTTPS requests, set "true" for key "aws:SecureTransport". When this key is true, the request is sent through HTTPS. This option configures encryption in transit and does not enforce encryption at rest.

INCORRECT: "Configure encryption at rest on CloudFront by using server-side encryption with AWS KMS keys (SSE-KMS)" is incorrect.

The data must be encrypted at rest in the Amazon S3 origin that is the source for the CloudFront distribution. This configuration does not ensure encryption at rest in the S3 bucket.

INCORRECT: "Use the RequireSSL option in the creation of presigned URLs for the S3 bucket that the web application uses" is incorrect.

Pre-signed URLs are used to provide short-term access to a private object in your S3 bucket. They work by appending an AWS Access Key, expiration time, and Sigv4 signature as query parameters to the S3 object. Encryption cannot be configured through the pre-signed URL.

References:

<https://aws.amazon.com/blogs/networking-and-content-delivery/serving-sse-kms-encrypted-content-from-s3-using-cloudfront/>

<https://docs.aws.amazon.com/AmazonCloudFront/latest/DeveloperGuide/data-protection-summary.html>

<https://aws.amazon.com/blogs/security/how-to-prevent-uploads-of-unencrypted-objects-to-amazon-s3>

Save time with our AWS cheat sheets:

<https://digitalcloud.training/amazon-s3-and-glacier/>

Practice 2

Question 1: **Correct**

A company has its flagship application fronted by an Application Load Balancer that is targeting several EC2 Linux instances running in an Auto Scaling group in a private subnet. AWS Systems Manager Agent is installed on all the EC2 instances. The company recently released a new version of the application, however, some of the EC2 instances are now being marked as unhealthy and are being terminated, thereby causing the application to run at reduced capacity. You have been tasked to ascertain the root cause by analyzing Amazon CloudWatch logs that are collected from the application, but you find that the logs are inconclusive.

Which of the following options would you propose to get access to an EC2 instance to troubleshoot the issue?



Enable EC2 instance termination protection. Use Session Manager to log In to an instance that is marked as unhealthy and analyze the system logs to figure out the root cause



Suspend the Auto Scaling group's **Terminate process. Use Session Manager to log in to an instance that is marked as unhealthy and analyze the system logs to figure out the root cause**

(Correct)



Suspend the Auto Scaling group's **Launch process. Use Session Manager to log in to an instance that is marked as unhealthy and analyze the system logs to figure out the root cause**



Suspend the Auto Scaling group's **HealthCheck process. Use EC2 instance connect to log in to an instance that is marked as unhealthy and analyze the system logs to figure out the root cause**

Explanation

Correct option:

Suspend the Auto Scaling group's [Terminate](#) process. Use Session Manager to log in to an instance that is marked as unhealthy and analyze the system logs to figure out the root cause

The [Terminate](#) process removes instances from the Auto Scaling group when the group scales in, or when Amazon EC2 Auto Scaling chooses to terminate instances for other reasons, such as when an instance is terminated for exceeding its maximum lifetime duration or failing a health check. You need to suspend the [Terminate](#) process which will allow you to get access to the instance without it being terminated even if it is marked as unhealthy. You should note that another way to prevent Amazon EC2 Auto Scaling from terminating unhealthy instances, is to suspend the [ReplaceUnhealthy](#) process. You can then leverage the Session Manager to log in to the instance that is marked as unhealthy and analyze the system logs to figure out the root cause.

Types of processes

The suspend-resume feature supports the following processes:

- **Launch**—Adds instances to the Auto Scaling group when the group scales out, or when Amazon EC2 Auto Scaling chooses to launch instances for other reasons, such as when it adds instances to a warm pool.
- **Terminate**—Removes instances from the Auto Scaling group when the group scales in, or when Amazon EC2 Auto Scaling chooses to terminate instances for other reasons, such as when an instance is terminated for exceeding its maximum lifetime duration or failing a health check.
- **AddToLoadBalancer**—Adds instances to the attached load balancer target group or Classic Load Balancer when they are launched. For more information, see [Use Elastic Load Balancing to distribute traffic across the instances in your Auto Scaling group](#).
- **AlarmNotification**—Accepts notifications from CloudWatch alarms that are associated with dynamic scaling policies. For more information, see [Dynamic scaling for Amazon EC2 Auto Scaling](#).
- **AZRebalance**—Balances the number of EC2 instances in the group evenly across all of the specified Availability Zones when the group becomes unbalanced, for example, when a previously unavailable Availability Zone returns to a healthy state. For more information, see [Rebalancing activities](#).
- **HealthCheck**—Checks the health of the instances and marks an instance as unhealthy if Amazon EC2 or Elastic Load Balancing tells Amazon EC2 Auto Scaling that the instance is unhealthy. This process can override the health status of an instance that you set manually. For more information, see [Health checks for Auto Scaling instances](#).
- **InstanceRefresh**—Terminates and replaces instances using the instance refresh feature. For more information, see [Replace Auto Scaling instances based on an instance refresh](#).
- **ReplaceUnhealthy**—Terminates instances that are marked as unhealthy and then creates new instances to replace them. For more information, see [Health checks for Auto Scaling instances](#).
- **ScheduledActions**—Performs the scheduled scaling actions that you create or that are created for you when you create an AWS Auto Scaling scaling plan and turn on predictive scaling. For more information, see [Scheduled scaling for Amazon EC2 Auto Scaling](#).

via -

https://docs.aws.amazon.com/AWSEC2/latest/WindowsGuide/terminating-instances.html#Using_ChangingDisableAPITermination

Enable termination protection

By default, you can terminate your instance using the Amazon EC2 console, command line interface, or API. To prevent your instance from being accidentally terminated using Amazon EC2, you can enable **termination protection** for the instance. The **DisableApiTermination** attribute controls whether the instance can be terminated using the console, CLI, or API. By default, termination protection is disabled for your instance. You can set the value of this attribute when you launch the instance, while the instance is running, or while the instance is stopped (for Amazon EBS-backed instances).

The **DisableApiTermination** attribute does not prevent you from terminating an instance by initiating shutdown from the instance (using an operating system command for system shutdown) when the **InstanceInitiatedShutdownBehavior** attribute is set. For more information, see [Change the instance initiated shutdown behavior](#).

Limitations

You can't enable termination protection for Spot Instances—a Spot Instance is terminated when the Spot price exceeds the amount you're willing to pay for Spot Instances. However, you can prepare your application to handle Spot Instance interruptions. For more information, see [Spot Instance interruptions](#).

The **DisableApiTermination** attribute does not prevent Amazon EC2 Auto Scaling from terminating an instance. For instances in an Auto Scaling group, use the following Amazon EC2 Auto Scaling features instead of Amazon EC2 termination protection:

- To prevent instances that are part of an Auto Scaling group from terminating on scale in, use instance scale-in protection. For more information, see [Using instance scale-in protection in the Amazon EC2 Auto Scaling User Guide](#).
- To prevent Amazon EC2 Auto Scaling from terminating unhealthy instances, suspend the **ReplaceUnhealthy** process. For more information, see [Suspending and Resuming Scaling Processes in the Amazon EC2 Auto Scaling User Guide](#).
- To specify which instances Amazon EC2 Auto Scaling should terminate first, choose a termination policy. For more information, see [Customizing the Termination Policy in the Amazon EC2 Auto Scaling User Guide](#).

via - https://docs.aws.amazon.com/AWSEC2/latest/WindowsGuide/terminating-instances.html#Using_ChangingDisableAPITermination

Incorrect options:

Suspend the Auto Scaling group's **HealthCheck process. Use EC2 instance connect to log in to an instance that is marked as unhealthy and analyze the system logs to figure out the root cause** - The **HealthCheck** process checks the health of the instances and marks an instance as unhealthy if Amazon EC2 or Elastic Load Balancing tells Amazon EC2 Auto Scaling that the instance is unhealthy. This process can override the health status of an instance that you set manually. If you suspend the **HealthCheck** process, then none of the instances would be marked as unhealthy. Therefore, you cannot suspend the **HealthCheck** process for the given use case, since you must identify the root cause behind some of the instances being marked as unhealthy.

Suspend the Auto Scaling group's **Launch process. Use Session Manager to log in to an instance that is marked as unhealthy and analyze the system logs to figure out the root cause** - The **Launch** process adds instances to the Auto Scaling group when the group scales out, or when Amazon EC2 Auto Scaling chooses to launch instances for other reasons, such as when it adds instances to

a warm pool. Suspending the **Launch** process will not help in identifying the root cause behind some instances being marked as unhealthy as those instances would still be terminated.

Enable EC2 instance termination protection. Use Session Manager to log In to an instance that is marked as unhealthy and analyze the system logs to figure out the root cause - Enabling EC2 instance termination (DisableApiTermination attribute) does not prevent Amazon EC2 Auto Scaling from terminating an instance.

References:

https://docs.aws.amazon.com/AWSEC2/latest/WindowsGuide/terminating-instances.html#Using_ChangingDisableAPITermination

<https://docs.aws.amazon.com/autoscaling/ec2/userguide/as-suspend-resume-processes.html>

Question 2: **Correct**

A multi-national company operates hundreds of AWS accounts and the CTO wants to rationalize the operational costs. The CTO has mandated a centralized process for purchasing new Reserved Instances (RIs) or modifying existing RIs. Whereas earlier the business units (BUs) would directly purchase or modify RIs in their own AWS accounts independently, now all BUs must be denied independent purchase and the BUs must submit requests to a dedicated central team for purchasing RIs.

As an AWS Certified Solutions Architect Professional, which of the following solutions would you combine to enforce the new process most efficiently? (Select two)



Leverage AWS Config to notify on the attachment of an IAM policy that allows access to the `ec2:PurchaseReservedInstancesOffering` and `ec2:ModifyReservedInstances` actions



Set up an IAM policy in each AWS account with a deny rule to the `ec2:PurchaseReservedInstancesOffering` and `ec2:ModifyReservedInstances` actions

-

Set up a Service Control Policy (SCP) that contains a deny rule to the ec2:PurchaseReservedInstancesOffering and ec2:ModifyReservedInstances actions. Attach the SCP to each organizational unit (OU) of the AWS Organizations structure

(Correct)

-

Make sure that all AWS accounts are assigned organizational units (OUs) within an AWS Organizations structure operating in the consolidated billing features mode

-

Make sure that all AWS accounts are assigned organizational units (OUs) within an AWS Organizations structure operating in all features mode

(Correct)

Explanation

Correct options:

Make sure that all AWS accounts are assigned organizational units (OUs) within an AWS Organizations structure operating in all features mode

AWS Organizations has two available feature sets:

All features – This feature set is the preferred way to work with AWS Organizations, and it includes Consolidating Billing features. When you create an organization, enabling all features is the default. With all features enabled, you can use the advanced account management features available in AWS Organizations such as integration with supported AWS services and organization management policies. Policies in AWS Organizations enable you to apply additional types of management to the AWS accounts in

your organization. You can use policies when all features are enabled in your organization. Service control policies (SCPs) offer central control over the maximum available permissions for all of the accounts in your organization.

Consolidated Billing features – All organizations support this subset of features, which provides basic management tools that you can use to centrally manage the accounts in your organization. You cannot leverage SCPs in this feature mode.

Set up a Service Control Policy (SCP) that contains a deny rule to the ec2:PurchaseReservedInstancesOffering and ec2:ModifyReservedInstances actions. Attach the SCP to each organizational unit (OU) of the AWS Organizations structure

Service control policies (SCPs) are a type of organizational 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 help you to ensure your accounts stay within your organization's access control guidelines. SCPs alone are not sufficient to grant permissions to the accounts in your organization. No permissions are granted by an SCP. An SCP defines a guardrail or sets limits, on the actions that the account's administrator can delegate to the IAM users and roles in the affected accounts. The administrator must still attach identity-based or resource-based policies to IAM users or roles, or the resources in your accounts to actually grant permissions. SCPs don't affect users or roles in the management account. They affect only the member accounts in your organization.

For the given use case, you can set up an SCP that contains a deny rule to the ec2:PurchaseReservedInstancesOffering and ec2:ModifyReservedInstances actions.

Incorrect options:

Make sure that all AWS accounts are assigned organizational units (OUs) within an AWS Organizations structure operating in the consolidated billing features mode - You cannot leverage SCPs in this feature mode, so this option is incorrect.

Leverage AWS Config to notify on the attachment of an IAM policy that allows access to the ec2:PurchaseReservedInstancesOffering and ec2:ModifyReservedInstances actions - AWS Config cannot prevent the attachment of an IAM policy that allows access to the ec2:PurchaseReservedInstancesOffering and ec2:ModifyReservedInstances actions. So this option is incorrect for the given requirements.

Set up an IAM policy in each AWS account with a deny rule to the ec2:PurchaseReservedInstancesOffering and ec2:ModifyReservedInstances actions - This is a cumbersome and inefficient solution to prevent each of the member AWS accounts from purchasing the RIs. This is not the best fit solution.

References:

https://docs.aws.amazon.com/organizations/latest/userguide/orgs_manage_org_support-all-features.html

https://docs.aws.amazon.com/organizations/latest/userguide/orgs_manage_policies.html

https://docs.aws.amazon.com/organizations/latest/userguide/orgs_manage_policies_scps.html

Question 3: **Correct**

The engineering team at a retail company wants to establish a dedicated, encrypted, low latency, and high throughput connection between its data center and AWS Cloud. The engineering team has set aside sufficient time to account for the operational overhead of establishing this connection.

Which of the following options represents the MOST optimal solution with the LEAST infrastructure set up required for provisioning the end to end connection?

-
- Use site-to-site VPN to establish a connection between the data center and AWS Cloud**
-
- Use AWS Direct Connect to establish a connection between the data center and AWS Cloud**
-
- Use VPC transit gateway to establish a connection between the data center and AWS Cloud**

-

Use AWS Direct Connect along with a site-to-site VPN to establish a connection between the data center and AWS Cloud

(Correct)

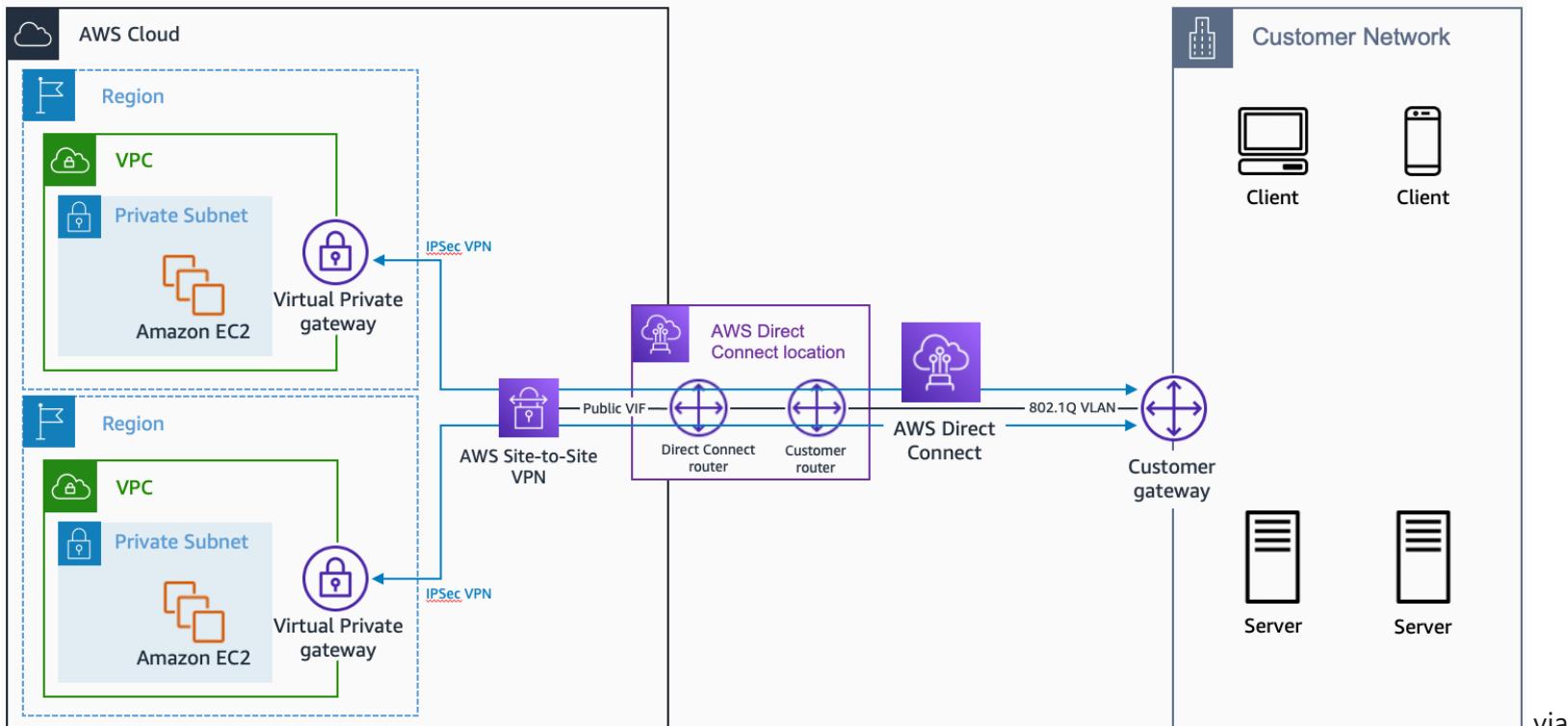
Explanation

Correct option: **Use AWS Direct Connect along with a site-to-site VPN to establish a connection between the data center and AWS Cloud**

AWS Direct Connect is a cloud service solution that makes it easy to establish a dedicated network connection from your premises to AWS. AWS Direct Connect lets you establish a dedicated network connection between your network and one of the AWS Direct Connect locations.

With AWS Direct Connect plus VPN, you can combine one or more AWS Direct Connect dedicated network connections with the Amazon VPC VPN. This combination provides an IPsec-encrypted private connection that also reduces network costs, increases bandwidth throughput, and provides a more consistent network experience than internet-based VPN connections. This solution combines the AWS managed benefits of the VPN solution with low latency, increased bandwidth, more consistent benefits of the AWS Direct Connect solution, and an end-to-end, secure IPsec connection. Therefore, AWS Direct Connect plus VPN is the correct solution for this use-case.

AWS Direct Connect Plus



VPN:

<https://docs.aws.amazon.com/whitepapers/latest/aws-vpc-connectivity-options/aws-direct-connect-vpn.html>

Incorrect options: **Use site-to-site VPN to establish a connection between the data center and AWS Cloud** - AWS Site-to-Site VPN enables you to securely connect your on-premises network or branch office site to your Amazon Virtual Private Cloud (Amazon VPC). A VPC VPN Connection utilizes IPSec to establish encrypted network connectivity between your intranet and Amazon VPC over the Internet. VPN Connections are a good solution if you have an immediate need, have low to modest bandwidth requirements, and can tolerate the inherent variability in Internet-based connectivity.

However, Site-to-site VPN cannot provide low latency and high throughput connection, therefore this option is ruled out.

Use VPC transit gateway to establish a connection between the data center and AWS Cloud - A transit gateway is a network transit hub that you can use to interconnect your virtual private clouds (VPC) and on-premises networks. A transit gateway by itself cannot establish a low latency and high throughput connection between a data center and AWS Cloud. Hence this option is incorrect.

Use AWS Direct Connect to establish a connection between the data center and AWS Cloud - AWS Direct Connect by itself cannot provide an encrypted connection between a data center and AWS Cloud, so this option is ruled out.

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

<https://docs.aws.amazon.com/directconnect/latest/UserGuide/encryption-in-transit.html>

<https://docs.aws.amazon.com/whitepapers/latest/aws-vpc-connectivity-options/aws-direct-connect-vpn.html>

Question 4: **Incorrect**

An Amazon Simple Storage Service (Amazon S3) bucket has been configured to host a static website. While using the S3 static website endpoint, the testing team has complained that they are receiving access denied error for this website.

What are the key points to consider while configuring an S3 bucket as a static website? (Select two)



Objects can't be encrypted by AWS Key Management Service (AWS KMS)

(Correct)



Amazon S3 Block Public Access must be disabled at the bucket level even though it is already disabled at the account level

(Incorrect)



Objects in the bucket must be publicly accessible. S3 bucket policy must allow access to the s3:GetObject and s3:Put Object actions

(Incorrect)

-

Amazon S3 static website endpoint needs to support both publicly and privately accessible content

-

The AWS account that owns the bucket must also own the object

(Correct)

Explanation

Correct options:

Objects can't be encrypted by AWS Key Management Service (AWS KMS) - AWS KMS doesn't support anonymous requests. As a result, any Amazon S3 bucket that allows anonymous or public access will not apply to objects that are encrypted with AWS KMS. You must remove KMS encryption from the objects that you want to serve using the Amazon S3 static website endpoint. Instead of using AWS KMS encryption, use AES-256 to encrypt your objects.

The AWS account that owns the bucket must also own the object - To allow public read access to objects, the AWS account that owns the bucket must also own the objects. A bucket or object is owned by the account of the AWS Identity and Access Management (IAM) identity that created the bucket or object. The object-ownership requirement applies to public read access granted by a bucket policy. It doesn't apply to public read access granted by the object's access control list (ACL).

Incorrect options:

Objects in the bucket must be publicly accessible. S3 bucket policy must allow access to the s3:GetObject and s3:Put Object actions - Objects in the bucket must indeed be publicly accessible. S3 bucket policy needs to just allow access to the s3:GetObject action for public access of objects in S3.

Amazon S3 Block Public Access must be disabled at the bucket level even though it is already disabled at the account level - Amazon S3 Block Public Access settings can apply to individual buckets or AWS accounts. Confirm that there aren't any Amazon S3 Block Public Access settings applied to either your S3 bucket or AWS account. These settings can override permissions that allow public read access. There is no need to disable the 'Block Public Access' feature at the bucket level even though it is already disabled at the account level.

Amazon S3 static website endpoint needs to support both publicly and privately accessible content - This statement is incorrect. Amazon S3 static website endpoint supports only publicly accessible content.

Reference:

<https://aws.amazon.com/premiumsupport/knowledge-center/s3-static-website-endpoint-error/>

Question 5: **Correct**

A team uses an Amazon S3 bucket to store the client data. After updating the S3 bucket with a few file deletes and some new file additions, the team has just realized that these changes have not been propagated to the AWS Storage Gateway file share.

What is the underlying issue? Which method can be used to resolve it?



Storage Gateway doesn't automatically update the cache when you upload a file directly to Amazon S3. Perform a `ResetCache` operation to see the changes on the file share



Uploading files from your file gateway to Amazon S3 when S3 Versioning is enabled results in cache update issues. Disable versioning on the S3 bucket



Storage Gateway doesn't automatically update the cache when you upload a file directly to Amazon S3. Perform a RefreshCache operation to see the changes on the file share

(Correct)

-

Configure correct permissions in Amazon S3 bucket policy to allow automatic refresh of cache

Explanation

Correct option:

Storage Gateway doesn't automatically update the cache when you upload a file directly to Amazon S3. Perform a RefreshCache operation to see the changes on the file share

Storage Gateway updates the file share cache automatically when you write files to the cache locally using the file share. However, Storage Gateway doesn't automatically update the cache when you upload a file directly to Amazon S3. When you do this, you must perform a RefreshCache operation to see the changes on the file share. If you have more than one file share, then you must run the RefreshCache operation on each file share.

You can refresh the cache using the Storage Gateway console and the AWS Command Line Interface (AWS CLI).

Incorrect options:

Uploading files from your file gateway to Amazon S3 when S3 Versioning is enabled results in cache update issues. Disable versioning on the S3 bucket - Carefully consider the use of S3 Versioning and Cross-Region Replication (CRR) in Amazon S3 when you're uploading data from your file gateway. Uploading files from your file gateway to Amazon S3 when S3 Versioning is enabled results in at least two versions of an S3 object. This option is not relevant to the given issue and has just been added as a distractor.

Storage Gateway doesn't automatically update the cache when you upload a file directly to Amazon S3. Perform a ResetCache operation to see the changes on the file share - 'ResetCache', resets all cache disks that have encountered an error, and make the disks available for reconfiguration as cache storage. When a cache is reset, the gateway loses its cache storage. At this point, you can reconfigure the disks as cache disks. This operation is only supported in the cached volume and tape gateway types.

Configure correct permissions in Amazon S3 bucket policy to allow automatic refresh of cache - This statement is incorrect and has just been added as a distractor.

References:

<https://docs.aws.amazon.com/filegateway/latest/files3/GettingStartedCreateFileShare.html>

<https://aws.amazon.com/premiumsupport/knowledge-center/storage-gateway-s3-changes-not-showing/>

<https://docs.aws.amazon.com/filegateway/latest/files3/refresh-cache.html>

https://docs.aws.amazon.com/storagegateway/latest/APIReference/API_RefreshCache.html

Question 6: Correct

A social media company manages a multi-AZ VPC environment consisting of public subnets and private subnets. Each public subnet contains a NAT Gateway as well as an Internet Gateway. Most of the company's applications are deployed in the private subnets and these applications read and write data to Kinesis Data Streams. The company has hired you as an AWS Certified Solutions Architect Professional to reduce costs and optimize the applications. Upon analysis in the AWS Cost Explorer, you notice that the cost in the EC2-Other category is consistently high due to the increasing NAT Gateway data transfer charges.

What do you recommend to address this requirement?

-

Set up a gateway VPC endpoint for Kinesis Data Streams in the VPC. Ensure that the applications have the required IAM permissions to use the gateway VPC endpoint

-

Set up an interface VPC endpoint for Kinesis Data Streams in the VPC. Ensure that the applications have the required IAM permissions to use the interface VPC endpoint

- Set up an interface VPC endpoint for Kinesis Data Streams in the VPC. Ensure that the VPC endpoint policy allows traffic from the applications
(Correct)

- Set up a gateway VPC endpoint for Kinesis Data Streams in the VPC. Ensure that the VPC endpoint policy allows traffic from the applications
- Explanation**
Correct option:
- Set up an interface VPC endpoint for Kinesis Data Streams in the VPC. Ensure that the VPC endpoint policy allows traffic from the applications**

You can use an interface VPC endpoint to keep traffic between your Amazon VPC and Kinesis Data Streams from leaving the Amazon network. Interface VPC endpoints don't require an internet gateway, NAT device, VPN connection, or AWS Direct Connect connection. Interface VPC endpoints are powered by AWS PrivateLink, an AWS technology that enables private communication between AWS services using an elastic network interface with private IPs in your Amazon VPC. You do not need to change the settings for your streams, producers, or consumers. Simply create an interface VPC endpoint for your Kinesis Data Streams traffic from and to your Amazon VPC-based applications to start flowing through the interface VPC endpoint.

VPC endpoint policies enable you to control access by either attaching a policy to a VPC endpoint or by using additional fields in a policy that is attached to an IAM user, group, or role to restrict access to only occur via the specified VPC endpoint. These policies can be used to restrict access to specific streams to a specified VPC endpoint when used in conjunction with the IAM policies to only grant access to Kinesis data stream actions via the specified VPC endpoint.

Controlling Access to VPCE Endpoints for Kinesis Data Streams

VPC endpoint policies enable you to control access by either attaching a policy to a VPC endpoint or by using additional fields in a policy that is attached to an IAM user, group, or role to restrict access to only occur via the specified VPC endpoint. These policies can be used to restrict access to specific streams to a specified VPC endpoint when used in conjunction with the IAM policies to only grant access to Kinesis data stream actions via the specified VPC endpoint.

The following are example endpoint policies for accessing Kinesis data streams.

- **VPC policy example: read-only access** - this sample policy can be attached to a VPC endpoint. (For more information, see [Controlling Access to Amazon VPC Resources](#)). It restricts actions to only listing and describing a Kinesis data stream through the VPC endpoint to which it is attached.

```
{  
  "Statement": [  
    {  
      "Sid": "ReadOnly",  
      "Principal": "*",  
      "Action": [  
        "kinesis>List*",  
        "kinesisDescribe*"  
      ],  
      "Effect": "Allow",  
      "Resource": "*"  
    }  
  ]  
}
```

- **VPC policy example: restrict access to a specific Kinesis data stream** - this sample policy can be attached to a VPC endpoint. It restricts access to a specific data stream through the VPC endpoint to which it is attached.

```
{  
  "Statement": [  
    {  
      "Sid": "AccessToSpecificDataStream",  
      "Principal": "*",  
      "Action": "kinesis:*",  
      "Effect": "Allow",  
      "Resource": "arn:aws:kinesis:us-east-1:123456789012:stream/MyStream"  
    }  
  ]  
}
```

- **IAM policy example: restrict access to a specific Stream from a specific VPC endpoint only** - this sample policy can be attached to an IAM user, role, or group. It restricts access to a specified Kinesis data stream to occur only from a specified VPC endpoint.

```
{  
  "Version": "2012-10-17",  
  "Statement": [  
    {  
      "Sid": "AccessFromSpecificEndpoint",  
      "Action": "kinesis:",  
      "Effect": "Deny",  
      "Resource": "arn:aws:kinesis:us-east-1:123456789012:stream/MyStream",  
      "Condition": { "StringNotEquals" : { "aws:sourceVpc": "vpce-11aa22bb" } }  
    }  
  ]  
}
```

via

- <https://docs.aws.amazon.com/streams/latest/dev/vpc.html>

Incorrect options:

Set up an interface VPC endpoint for Kinesis Data Streams in the VPC. Ensure that the applications have the required IAM permissions to use the interface VPC endpoint - Although you can use an IAM policy to restrict access to specific streams to a specified VPC endpoint by only granting access to Kinesis data stream actions via the specified VPC endpoint. However, you need to make changes in the code for the different applications to assume the relevant IAM role and then make changes in the permissions policy attached to the IAM role to grant access to Kinesis Data Streams via the VPC endpoint. This is not an elegant solution when compared to just using the VPC endpoint policy.

Set up a gateway VPC endpoint for Kinesis Data Streams in the VPC. Ensure that the VPC endpoint policy allows traffic from the applications

Set up a gateway VPC endpoint for Kinesis Data Streams in the VPC. Ensure that the applications have the required IAM permissions to use the gateway VPC endpoint

There are three types of VPC endpoints. You must create the type of VPC endpoint that's required by the endpoint service.

Interface - Create an interface endpoint to send traffic to endpoint services that use a Network Load Balancer to distribute traffic. Traffic destined for the endpoint service is resolved using DNS.

GatewayLoadBalancer - Create a Gateway Load Balancer endpoint to send traffic to a fleet of virtual appliances using private IP addresses. You route traffic from your VPC to the Gateway Load Balancer endpoint using route tables. The Gateway Load Balancer distributes traffic to the virtual appliances and can scale with demand.

Gateway - Create a gateway endpoint to send traffic to Amazon S3 or DynamoDB using private IP addresses. You route traffic from your VPC to the gateway endpoint using route tables. Gateway endpoints do not enable AWS PrivateLink.

You cannot set up a gateway VPC endpoint for Kinesis Data Streams. Gateway VPC endpoint is only supported for S3 and DynamoDB. Therefore, both these options are incorrect.

References:

<https://docs.aws.amazon.com/streams/latest/dev/vpc.html>

<https://docs.aws.amazon.com/vpc/latest/privatelink/vpc-endpoints-access.html>

Question 7: Incorrect

A company is migrating its two-tier legacy application (using MongoDB as a key-value database) from its on-premises data center to AWS. The company has mandated that the EC2 instances must be hosted in a private subnet with no internet access. In addition, all connectivity between the EC2 instance-hosted application and the database must be encrypted. The database must be able to scale to meet traffic spikes from any bursty or unpredictable workloads.

Which do you recommend?



Set up new Amazon DynamoDB tables for the application with on-demand capacity. Use a gateway VPC endpoint for DynamoDB so that the application can have a private and encrypted connection to the DynamoDB tables

(Correct)



Set up a new Amazon DocumentDB (with MongoDB compatibility) cluster for the application with provisioned capacity with auto-scaling enabled. Use an interface VPC endpoint for DocumentDB so that the application can have a private and encrypted connection to the DocumentDB tables

(Incorrect)



Set up new Amazon DynamoDB tables for the application with on-demand capacity. Use an interface VPC endpoint for DynamoDB so that the application can have a private and encrypted connection to the DynamoDB tables

-

Set up a new Amazon DocumentDB (with MongoDB compatibility) cluster for the application with on-demand capacity. Use a gateway VPC endpoint for DocumentDB so that the application can have a private and encrypted connection to the DocumentDB tables

Explanation

Correct option:

Set up new Amazon DynamoDB tables for the application with on-demand capacity. Use a gateway VPC endpoint for DynamoDB so that the application can have a private and encrypted connection to the DynamoDB tables

With provisioned capacity, you pay for the provision of read and write capacity units for your DynamoDB tables. Whereas with DynamoDB on-demand you pay per request for the data reads and writes that your application performs on your tables.

With on-demand capacity mode, DynamoDB charges you for the data reads and writes your application performs on your tables. You do not need to specify how much read and write throughput you expect your application to perform because DynamoDB instantly accommodates your workloads as they ramp up or down.

With provisioned capacity mode, you specify the number of reads and writes per second that you expect your application to require, and you are billed based on that. Furthermore, if you can forecast your capacity requirements you can also reserve a portion of DynamoDB provisioned capacity and optimize your costs even further. With provisioned capacity, you can also use auto-scaling to automatically adjust your table's capacity based on the specified utilization rate to ensure application performance, and also potentially reduce costs. To configure auto-scaling in DynamoDB, set the minimum and maximum levels of read and write capacity in addition to the target utilization percentage.

It is important to note that DynamoDB auto scaling modifies provisioned throughput settings only when the actual workload stays elevated or depressed for a sustained period of several minutes. This applies to scaling up or down the provisioned capacity of a DynamoDB table. In the case that you have an occasional usage spike, auto-scaling might not be able to react in time. This sometimes can be mitigated by DynamoDB burst capacity where DynamoDB reserves a portion of the unused provisioned capacity for later bursts of throughput. The burst capacity is limited though and these extra capacity units can be consumed quickly.

This means that provisioned capacity is probably best for you if you have relatively predictable application traffic, run applications whose traffic is consistent, and ramps up or down gradually.

Whereas on-demand capacity mode is probably best when you have new tables with unknown workloads, unpredictable application traffic, and also if you only want to pay exactly for what you use. The on-demand pricing model is ideal for bursty, new, or unpredictable workloads whose traffic can spike in seconds or minutes, and when under-provisioned capacity would impact the user experience.

Incorrect options:

Set up a new Amazon DocumentDB (with MongoDB compatibility) cluster for the application with provisioned capacity with auto-scaling enabled. Use an interface VPC endpoint for DocumentDB so that the application can have a private and encrypted connection to the DocumentDB tables

Set up a new Amazon DocumentDB (with MongoDB compatibility) cluster for the application with on-demand capacity. Use a gateway VPC endpoint for DocumentDB so that the application can have a private and encrypted connection to the DocumentDB tables

Amazon DocumentDB (with MongoDB compatibility) clusters are deployed within an Amazon Virtual Private Cloud (Amazon VPC). They can be accessed directly by Amazon EC2 instances or other AWS services that are deployed in the same Amazon VPC. Additionally, Amazon DocumentDB can be accessed by EC2 instances or other AWS services in different VPCs in the same AWS Region or other Regions via VPC peering. Therefore, neither the interface nor gateway VPC endpoint is supported for DocumentDB. So both these options are incorrect.

Set up new Amazon DynamoDB tables for the application with on-demand capacity. Use an interface VPC endpoint for DynamoDB so that the application can have a private and encrypted connection to the DynamoDB tables - Only gateway VPC endpoint is supported for DynamoDB, so this option is incorrect.

References:

<https://docs.aws.amazon.com/wellarchitected/latest/serverless-applications-lens/capacity.html>

<https://docs.aws.amazon.com/documentdb/latest/developerguide/connect-from-outside-a-vpc.html>

Question 8: **Incorrect**

The development team at a company needs to implement a client-side encryption mechanism for objects that will be stored in a new Amazon S3 bucket. The team created a CMK that is stored in AWS Key Management Service (AWS KMS) for this purpose. The team created the following IAM policy and attached it to an IAM role:

```
{  
    "Version": "2012-10-17",  
    "Id": "key-policy-1",  
    "Statement": [  
        {  
            "Sid": "GetPut",  
            "Effect": "Allow",  
            "Action": [  
                "s3:GetObject",  
                "s3:PutObject"  
            ],  
            "Resource": "arn:aws:s3:::ExampleBucket/*"  
        },  
        {  
            "Sid": "KMS",  
            "Effect": "Allow",  
            "Action": [  
                "kms:Decrypt",  
                "kms:Encrypt"  
            ],  
            "Resource": "arn:aws:kms:us-west-1:111122223333:key/keyid-12345"  
        }  
    ]  
}
```

The team was able to successfully get existing objects from the S3 bucket while testing. But any attempts to upload a new object resulted in an error. The error message stated that the action was forbidden.

Which IAM policy action should be added to the IAM policy to resolve the error?

kms:GetKeyPolicy

(Incorrect)



kms:GenerateDataKey

(Correct)



kms:GetPublicKey



kms:GetDataKey

Explanation

Correct option:

kms:GenerateDataKey

GenerateDataKey returns a unique symmetric data key for use outside of AWS KMS. This operation returns a plaintext copy of the data key and a copy that is encrypted under a symmetric encryption KMS key that you specify. The bytes in the plaintext key are random; they are not related to the caller or the KMS key. You can use the plaintext key to encrypt your data outside of AWS KMS and store the encrypted data key with the encrypted data.

How to use your data key

We recommend that you use the following pattern to encrypt data locally in your application. You can write your own code or use a client-side encryption library, such as the [AWS Encryption SDK](#), the [Amazon DynamoDB Encryption Client](#), or [Amazon S3 client-side encryption](#) to do these tasks for you.

To encrypt data outside of AWS KMS:

1. Use the `GenerateDataKey` operation to get a data key.
2. Use the plaintext data key (in the `Plaintext` field of the response) to encrypt your data outside of AWS KMS. Then erase the plaintext data key from memory.
3. Store the encrypted data key (in the `CiphertextBlob` field of the response) with the encrypted data.

To decrypt data outside of AWS KMS:

1. Use the `Decrypt` operation to decrypt the encrypted data key. The operation returns a plaintext copy of the data key.
2. Use the plaintext data key to decrypt data outside of AWS KMS, then erase the plaintext data key from memory.

Cross-account use: Yes. To perform this operation with a KMS key in a different AWS account, specify the key ARN or alias ARN in the value of the `KeyId` parameter.

Required permissions: `kms:GenerateDataKey` (key policy)

via -

https://docs.aws.amazon.com/kms/latest/APIReference/API_GenerateDataKey.html

Incorrect options:

kms:GetPublicKey - This option returns the public key of an asymmetric KMS key. Unlike the private key of an asymmetric KMS key, which never leaves AWS KMS unencrypted, callers with `kms:GetPublicKey` permission can download the public key of an asymmetric KMS key. It cannot be used for a client-side encryption mechanism.

kms:GetKeyPolicy - This option gets a key policy attached to the specified KMS key. It cannot be used for a client-side encryption mechanism.

kms:GetDataKey - This is a made-up option that serves as a distractor.

References:

https://docs.aws.amazon.com/kms/latest/APIReference/API_GenerateDataKey.html

https://docs.aws.amazon.com/kms/latest/APIReference/API_GetKeyPolicy.html

https://docs.aws.amazon.com/kms/latest/APIReference/API_GetPublicKey.html

Question 9: **Correct**

A company provides a web-based business-management platform for IT service companies across the globe to manage help desk, customer service, sales and marketing, and other critical business functions. More than 50,000 people use the company's platform, so the company must respond quickly to any reported problems. However, the company has issues with not having enough visibility into its systems to discover any issues. Multiple logs and monitoring systems are needed to understand the root cause of problems thereby taking hours to resolve. Even as the company is slowly moving towards serverless architecture using AWS Lambda/Amazon API Gateway/Amazon Elastic Container Service (Amazon ECS), the company wants to monitor the microservices and gain deeper insights into its serverless resources.

Which of the following will you recommend to address the given requirements?

-
- Use AWS X-Ray to analyze the microservices applications through request tracing. Configure Amazon EventBridge for monitoring containers, latency, web server requests, and incoming load-balancer requests and create alarms to send out notifications if system latency is increasing**
-
- Use AWS X-Ray to analyze the microservices applications through request tracing. Configure Amazon CloudWatch for monitoring containers, latency, web server requests, and incoming load-balancer requests and create CloudWatch alarms to send out notifications if system latency is increasing**

(Correct)

-

Configure Amazon CloudWatch to monitor and analyze all microservices through request tracing. Enable CloudTrail to log all user activity

-

Configure Amazon EventBridge for monitoring containers, latency, web server requests, and incoming load-balancer requests and create alarms to send out notifications if system latency is increasing. Use AWS Config to continually assesses, audit, and evaluate the configurations and relationships of your resources and trigger alarms when needed

Explanation

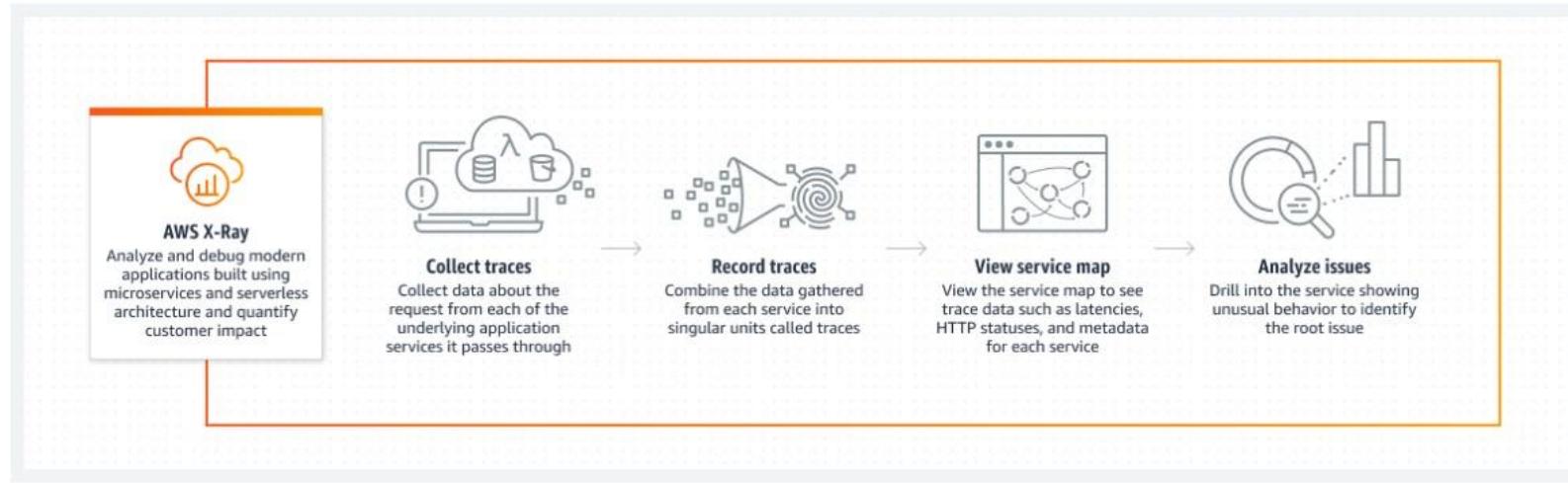
Correct option:

Use AWS X-Ray to analyze the microservices applications through request tracing. Configure Amazon CloudWatch for monitoring containers, latency, web server requests, and incoming load-balancer requests and create CloudWatch alarms to send out notifications if system latency is increasing

AWS X-Ray helps developers analyze and debug production, and distributed applications, such as those built using a microservices architecture.

Analyze and debug using X-Ray:

AWS X-Ray provides a complete view of requests as they travel through your application and filters visual data across payloads, functions, traces, services, APIs, and more with no-code and low-code motions.

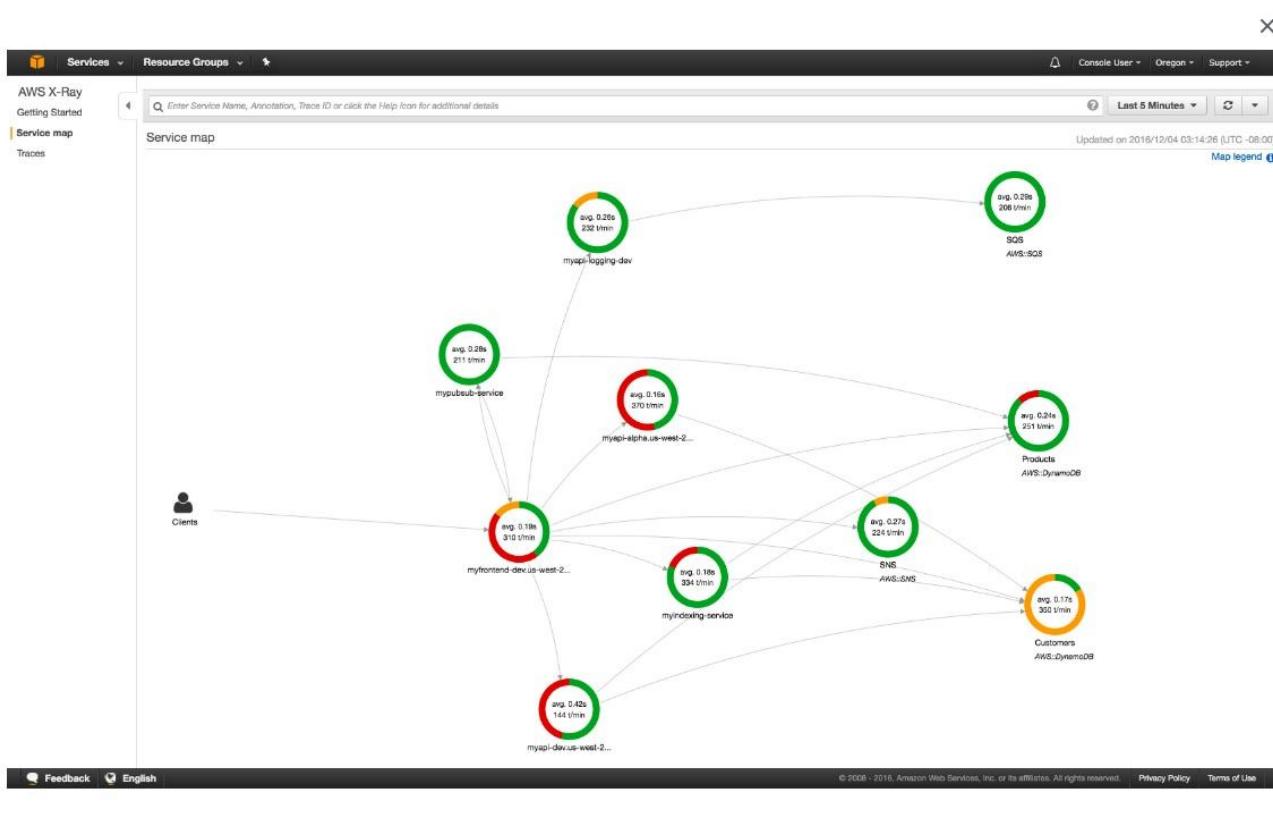


Click to enlarge

vi

a - <https://aws.amazon.com/xray/>

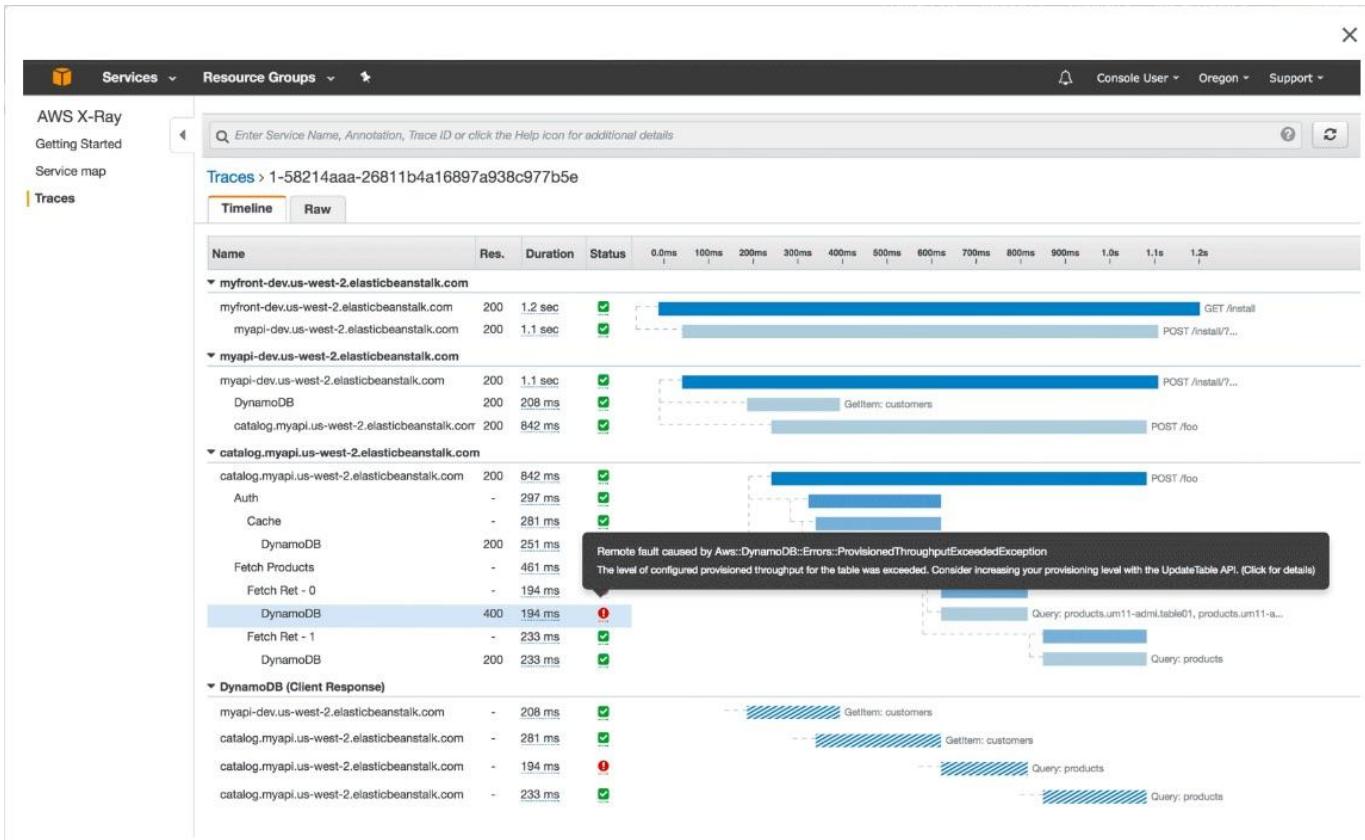
AWS X-Ray creates a map of services used by your application with trace data that you can use to drill into specific services or issues. This provides a view of connections between services in your application and aggregated data for each service, including average latency and failure rates.



X-Ray service map:

<https://aws.amazon.com/xray/features/>

via -



X-Ray Traces:

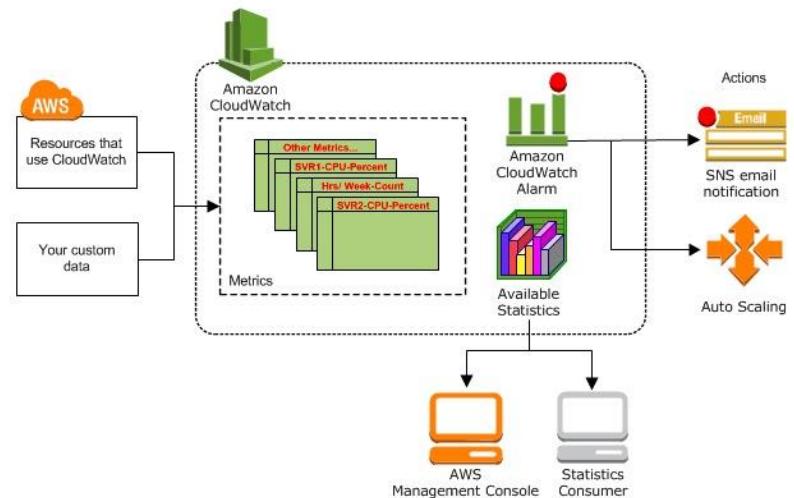
<https://aws.amazon.com/xray/features/>

via -

Amazon CloudWatch allows you to collect infrastructure metrics from more than 70 AWS services, such as Amazon Elastic Compute Cloud (Amazon EC2), Amazon DynamoDB, Amazon Simple Storage Service (Amazon S3), Amazon ECS, AWS Lambda, Amazon API Gateway, with no action on your part. For example, Amazon EC2 instances automatically publish CPU utilization, data transfer, and disk usage metrics to help you understand state changes. You can use built-in metrics for API Gateway to detect latency or use built-in metrics for AWS Lambda to detect errors or throttles. Likewise, Amazon CloudWatch also allows you to collect application metrics (such as user activity, error metrics or memory used) from your applications to monitor operational performance, troubleshoot issues, and spot trends.

Amazon CloudWatch for

Amazon CloudWatch is basically a metrics repository. An AWS service—such as Amazon EC2—puts metrics into the repository, and you retrieve statistics based on those metrics. If you put your own custom metrics into the repository, you can retrieve statistics on these metrics as well.



monitoring:

via -

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

Incorrect options:

Use AWS X-Ray to analyze the microservices applications through request tracing. Configure Amazon EventBridge for monitoring containers, latency, web server requests, and incoming load-balancer requests and create alarms to send out notifications if system latency is increasing

Configure Amazon EventBridge for monitoring containers, latency, web server requests, and incoming load-balancer requests and create alarms to send out notifications if system latency is increasing. Use AWS Config to continually assesses, audit, and evaluate the configurations and relationships of your resources and trigger alarms when needed

Amazon EventBridge is a serverless event bus service that uses the Amazon CloudWatch Events API, but also includes more functionality, like the ability to ingest events from SaaS apps. EventBridge is designed to extend the event model beyond AWS, bringing data from software-as-a-service (SaaS) providers into your AWS environment. This means you can consume events from popular providers such as Zendesk, PagerDuty, and Auth0. You can use these in your applications with the same ease as any AWS-generated event.

For the given use case, you can use Cloudwatch for monitoring containers, latency, web server requests, and incoming load-balancer requests and create CloudWatch alarms to send out notifications if system latency is increasing. Therefore, both these options are incorrect.

Configure Amazon CloudWatch to monitor and analyze all microservices through request tracing. Enable CloudTrail to log all user activity - X-Ray can be used to monitor and analyze AWS microservices through request tracing, so this option is incorrect.

References:

<https://aws.amazon.com/cloudwatch/features/>

<https://docs.aws.amazon.com/AmazonCloudWatch/latest/monitoring/ContainerInsights.html>

<https://aws.amazon.com/solutions/case-studies/connectwise/>

Question 10: **Correct**

A retail company offers its services to the customers via APIs that leverage Amazon API Gateway and Lambda functions. The company also has a legacy API hosted on an Amazon EC2 instance that is used by the company's supply chain partners. The security and audit team at the company has raised concerns over the use of these APIs and wants a solution to secure them all from any vulnerabilities, DDoS attacks, and malicious exploits.

Which of the following options would you use to address the security requirements of the company?

-

Enable AWS Network Firewall on API Gateway as well as the Amazon EC2 instances to check for vulnerabilities and protect against DDoS attacks as well as malicious exploits

- Configure Amazon CloudFront in front of the APIs to protect against malicious exploits and DDoS attacks. Install Amazon GuardDuty on the EC2 instances to assess any vulnerabilities**

- Use AWS Web Application Firewall (WAF) as the first line of defense to protect the API Gateway APIs against malicious exploits and DDoS attacks. Install Amazon Inspector on the EC2 instance to check for vulnerabilities. Configure Amazon GuardDuty to monitor any malicious attempts to access the APIs illegally**

(Correct)

- Use AWS Web Application Firewall (WAF) as the first line of defense to protect the API Gateway APIs against malicious exploits and DDoS attacks. Install Amazon Inspector on the EC2 instance to check for vulnerabilities. Configure Amazon GuardDuty to block any malicious attempts to access the APIs illegally**

Explanation

Correct option:

Use AWS Web Application Firewall (WAF) as the first line of defense to protect the API Gateway APIs against malicious exploits and DDoS attacks. Install Amazon Inspector on the EC2 instance to check for vulnerabilities. Configure Amazon GuardDuty to monitor any malicious attempts to access the APIs illegally

AWS WAF is a web application firewall that helps protect web applications and APIs from attacks. It enables you to configure a set of rules (called a web access control list (web ACL)) that allow, block, or count web requests based on customizable web security rules and conditions that you define. You can protect the following resource types:

1. Amazon CloudFront distribution

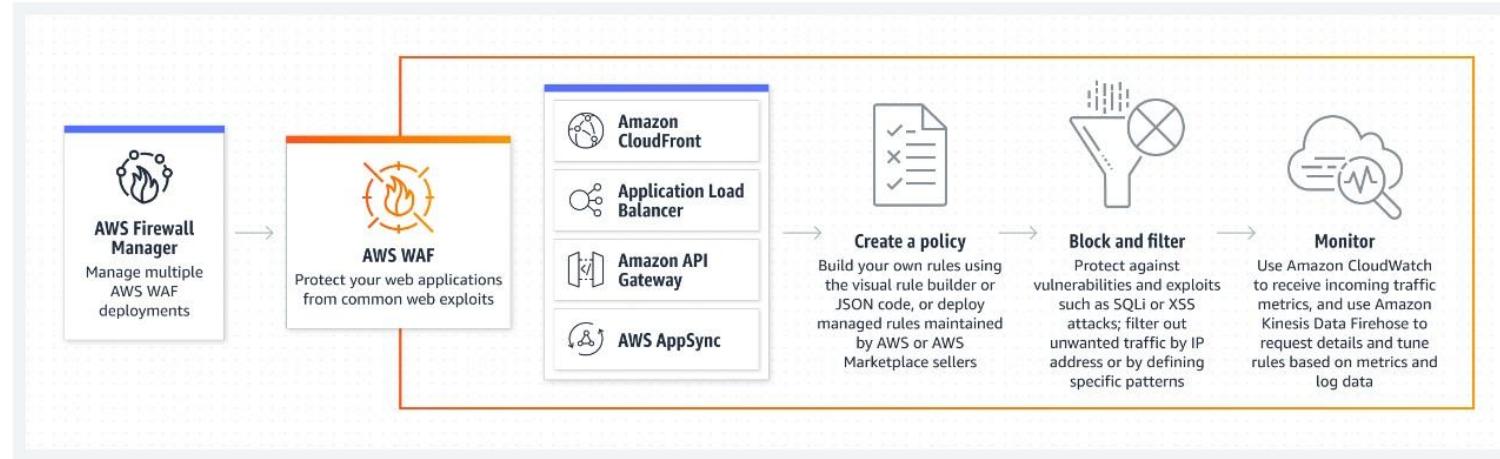
2. Amazon API Gateway REST API
3. Application Load Balancer
4. AWS AppSync GraphQL API
5. Amazon Cognito user pool

You can use AWS WAF to protect your API Gateway API from common web exploits, such as SQL injection and cross-site scripting (XSS) attacks. These could affect API availability and performance, compromise security, or consume excessive resources. For example, you can create rules to allow or block requests from specified IP address ranges, requests from CIDR blocks, requests that originate from a specific country or region, requests that contain malicious SQL code, or requests that contain malicious scripts.

DDoS attacks are attempts by an attacker to disrupt the availability of targeted systems. For infrastructure layer attacks, you can use AWS services such as Amazon CloudFront and Elastic Load Balancing (ELB) to provide automatic DDoS protection. For application layer attacks, you can use AWS WAF as the primary mitigation. AWS WAF web access control lists (web ACLs) minimize the effects of a DDoS attack at the application layer.

How WAF

AWS WAF helps you protect against common web exploits and bots that can affect availability, compromise security, or consume excessive resources.



Click to enlarge

works:

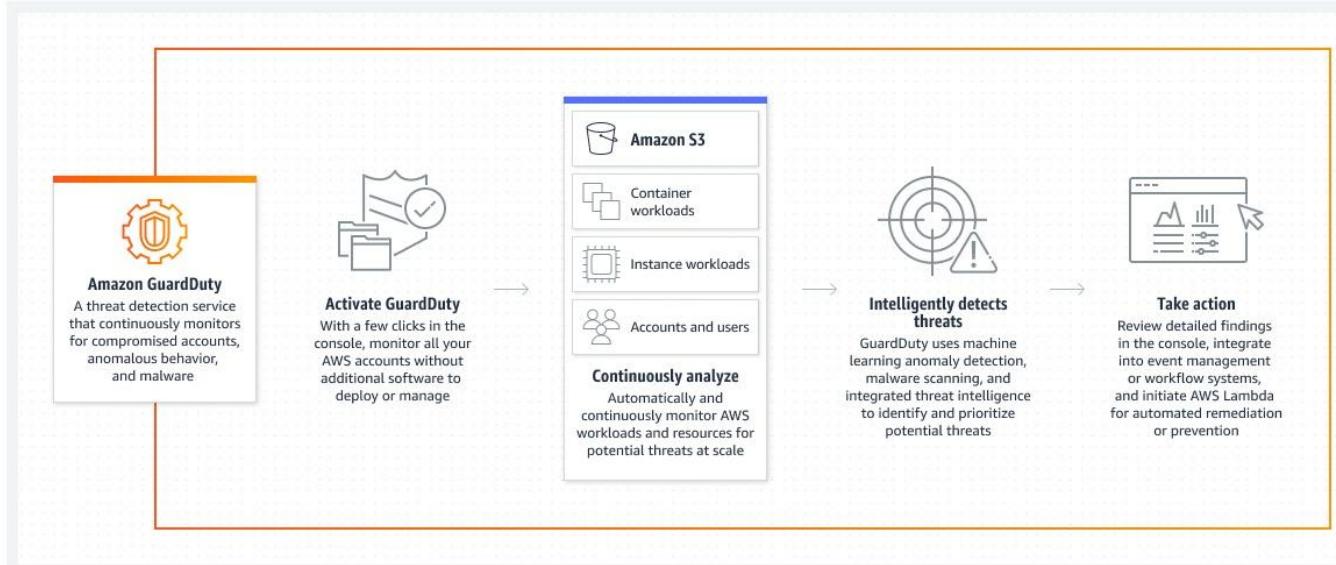
ia - <https://aws.amazon.com/waf/>

v

GuardDuty is an intelligent threat detection service that continuously monitors your AWS accounts, Amazon Elastic Compute Cloud (EC2) instances, Amazon Elastic Kubernetes Service (EKS) clusters, and data stored in Amazon Simple Storage Service (S3) for malicious activity without the use of security software or agents. If potential malicious activity, such as anomalous behavior, credential exfiltration, or command and control infrastructure (C2) communication is detected, GuardDuty generates detailed security findings that can be used for security visibility and assisting in remediation. GuardDuty can monitor reconnaissance activities by an attacker such as unusual API activity, intra-VPC port scanning, unusual patterns of failed login requests, or unblocked port probing from a known bad IP.

How GuardDuty

Amazon GuardDuty is a threat detection service that continuously monitors your AWS accounts and workloads for malicious activity and delivers detailed security findings for visibility and remediation.



works:

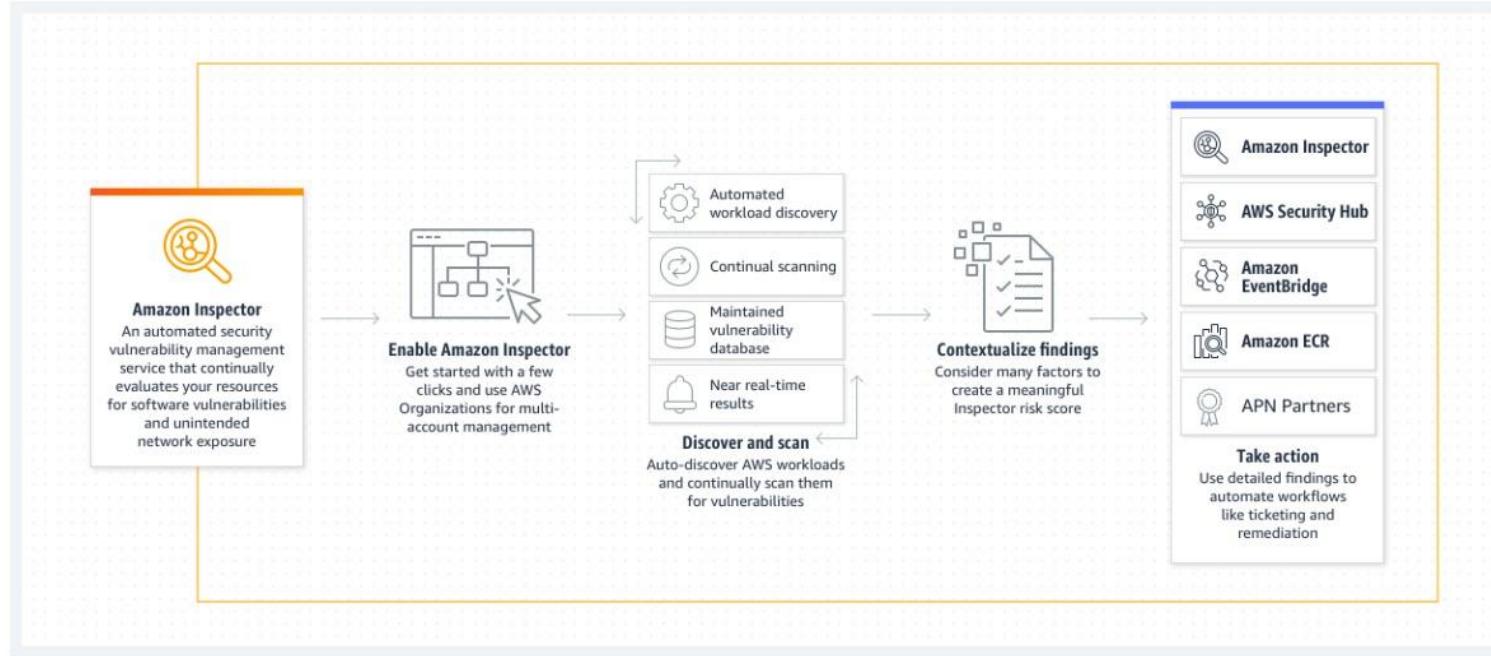
via -

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

Amazon Inspector is an automated vulnerability management service that continually scans Amazon Elastic Compute Cloud (EC2) and container workloads for software vulnerabilities and unintended network exposure.

How Amazon Inspector works:

Amazon Inspector is an automated vulnerability management service that continually scans AWS workloads for software vulnerabilities and unintended network exposure.



via -

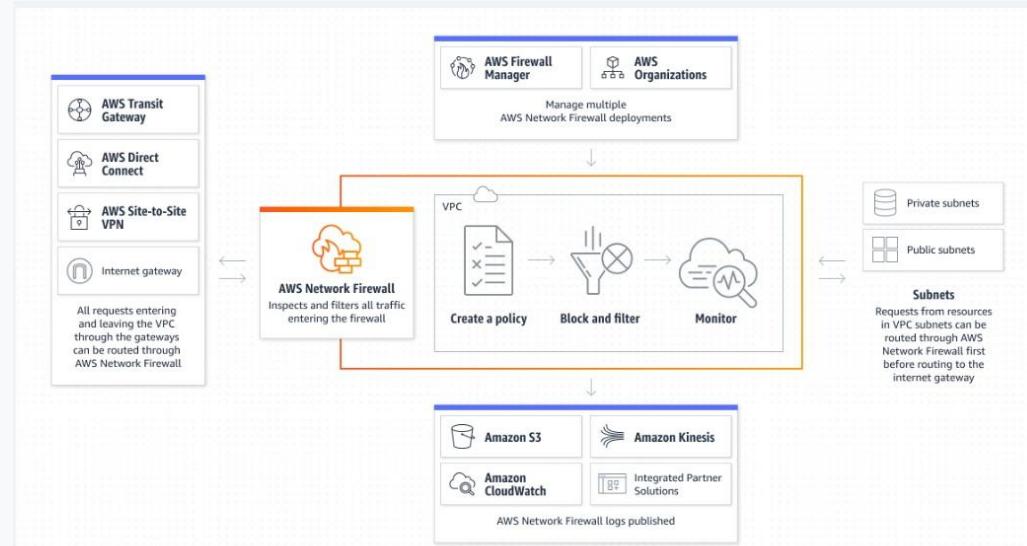
<https://aws.amazon.com/inspector>

Incorrect options:

Configure Amazon CloudFront in front of the APIs to protect against malicious exploits and DDoS attacks. Install Amazon GuardDuty on the EC2 instances to assess any vulnerabilities - This statement is incorrect. GuardDuty cannot assess vulnerabilities in the EC2 instances. Amazon Inspector is the automated vulnerability management service that continually scans Amazon Elastic Compute Cloud (EC2) and container workloads for software vulnerabilities and unintended network exposure.

Enable AWS Network Firewall on API Gateway as well as the Amazon EC2 instances to check for vulnerabilities and protect against DDoS attacks as well as malicious exploits - AWS Network Firewall is a managed service that makes it easy to deploy

essential network protections for all of your Amazon Virtual Private Clouds (VPCs). AWS Network Firewall works with AWS Firewall Manager to centrally manage security policies and automatically enforce mandatory security policies across existing and newly created accounts and VPCs. This service works at the VPC level and not at the API Gateway or the EC2 instance level.



How AWS Network Firewall works:

<https://aws.amazon.com/network-firewall/>

via -

Use AWS Web Application Firewall (WAF) as the first line of defense to protect the API Gateway APIs against malicious exploits and DDoS attacks. Install Amazon Inspector on the EC2 instance to check for vulnerabilities. Configure Amazon GuardDuty to block any malicious attempts to access the APIs illegally - GuardDuty cannot block any malicious attempts to access the APIs illegally. Rather, it can only monitor/detect such attempts.

References:

<https://aws.amazon.com/guardduty/faqs/>

<https://aws.amazon.com/premiumsupport/knowledge-center/waf-mitigate-ddos-attacks/>

Question 11: **Correct**

An ed-tech company needs to deliver its video-on-demand (VOD) content to approximately 1 million users in a cost-effective way. The learning material is in the form of videos with a maximum size of 10 GB each. The videos are highly watched when initially uploaded and subsequently have very less views after 6-8 months. While the old videos might not be accessed regularly, they need to be immediately accessible when needed. With trainers and material doubling every few months, the number of videos has exploded over the last few months, dramatically increasing the cost of storage for the company.

Which is the most cost-effective way of storing these videos to address the given use case?

- - Use Amazon Elastic File System (Amazon EFS) Intelligent-Tiering storage class to store the video files. Configure an Amazon EC2 instance to deliver this content from EFS to viewers through an Amazon CloudFront distribution**
 -
 - Use Amazon Elastic File System (Amazon EFS) Standard storage class to store the video files. Move these video files to EFS Standard–Infrequent Access (Standard-IA) through lifecycle management configuration. Configure a CloudFront custom distribution to deliver content from the EFS origin**
 -
 - Use Amazon S3 Intelligent-Tiering storage class to store the video files. Configure this S3 bucket as the origin of an Amazon CloudFront distribution for delivering the contents to the customers**
- (Correct)**
- - Use AWS Elemental MediaConvert and store the transcoded videos in S3. Configure an AWS Elemental MediaPackage endpoint to deliver the content from S3**

Explanation

Correct option:

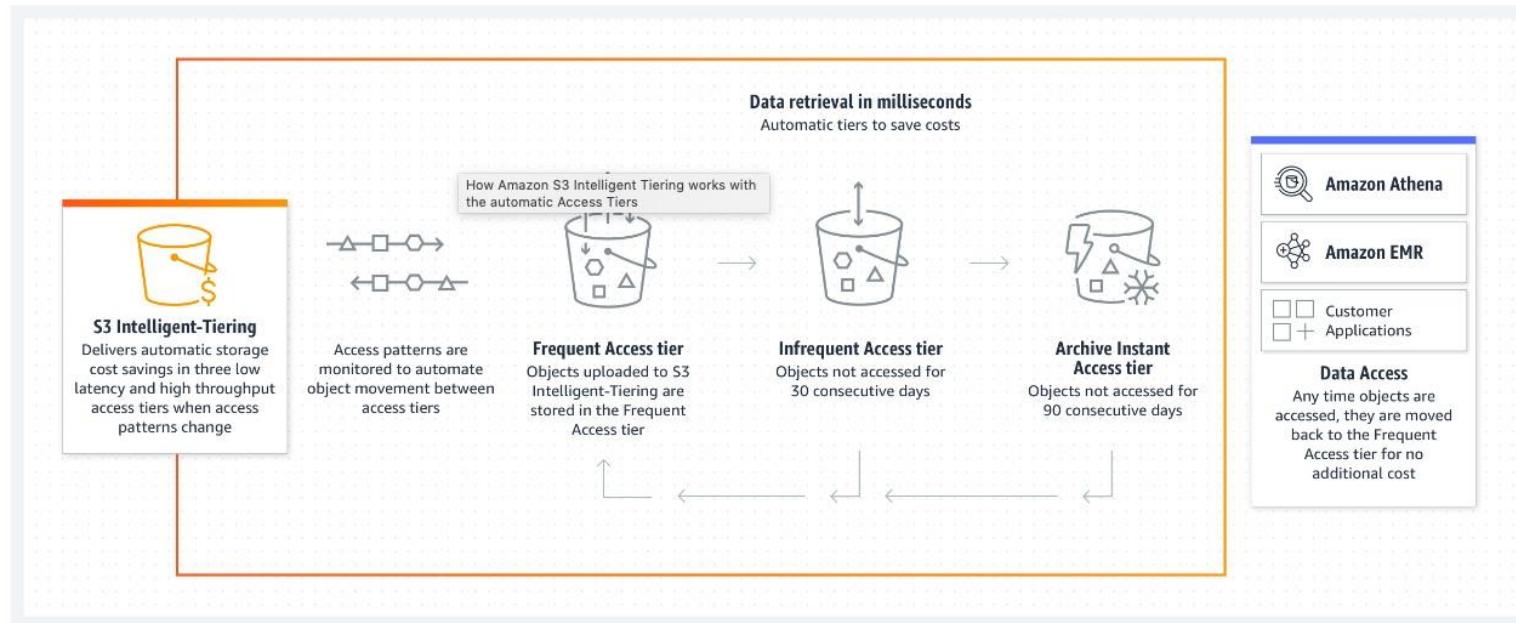
Use Amazon S3 Intelligent-Tiering storage class to store the video files. Configure this S3 bucket as the origin of an Amazon CloudFront distribution for delivering the contents to the customers

S3 Intelligent-Tiering is the only cloud storage class that delivers automatic storage cost savings when data access patterns change, without performance impact or operational overhead. The Amazon S3 Intelligent-Tiering storage class is designed to optimize storage costs by automatically moving data to the most cost-effective access tier when access patterns change. For a small monthly object monitoring and automation charge, S3 Intelligent-Tiering monitors access patterns and automatically moves objects that have not been accessed to lower-cost access tiers.

S3 Intelligent-Tiering is the ideal storage class for data with unknown, changing, or unpredictable access patterns, independent of object size or retention period. You can use S3 Intelligent-Tiering as the default storage class for virtually any workload, especially data lakes, data analytics, new applications, and user-generated content.

S3 Intelligent-Tiering storage

There are no retrieval charges in S3 Intelligent-Tiering. S3 Intelligent-Tiering has no minimum eligible object size, but objects smaller than 128 KB are not eligible for auto tiering. These smaller objects may be stored, but they'll always be charged at the Frequent Access tier rates and don't incur the monitoring and automation charge. See the [Amazon S3 Pricing](#) page for more information. To learn more, visit the [S3 Intelligent-Tiering user guide](#).



class:

<https://aws.amazon.com/s3/storage-classes/intelligent-tiering/>

When you store your objects in an Amazon S3 bucket, you can either have users get your objects directly from S3, or you can configure CloudFront to get your objects from S3 and then distribute them to your users. Using CloudFront can be more cost-effective if your users access your objects frequently because, at higher usage, the price for CloudFront data transfer is lower than the price for Amazon S3 data transfer. In addition, downloads are faster with CloudFront than with Amazon S3 alone since the cached content is served from the edge locations that are closer to the end users.

Incorrect options:

via -

Use Amazon Elastic File System (Amazon EFS) standard storage class to store the video files. Move these video files to EFS Standard-Infrequent Access (Standard-IA) through lifecycle management configuration. Configure a CloudFront custom distribution to deliver content from the EFS origin - Amazon Elastic File System (Amazon EFS) cannot be configured as an origin for a CloudFront distribution.

Use AWS Elemental MediaConvert and store the transcoded videos in S3. Configure an AWS Elemental MediaPackage endpoint to deliver the content from S3 - The main use case for AWS Media Services is to deliver live content to a global audience. For the given scenario, the content is primarily served via video on demand (VOD), so there is no need to incur extra costs for using the suite of AWS Media Services, so this option is incorrect.

Use Amazon Elastic File System (Amazon EFS) Intelligent-Tiering storage class to store the video files. Configure an Amazon EC2 instance to deliver this content from EFS to viewers through an Amazon CloudFront distribution - EFS is about three times costlier than S3. In addition, using EC2 as the origin for Cloudfront imposes additional costs for running the infrastructure. Therefore, this option is incorrect.

References:

<https://docs.aws.amazon.com/AmazonCloudFront/latest/DeveloperGuide/DownloadDistS3AndCustomOrigins.html>

<https://docs.aws.amazon.com/efs/latest/ug/storage-classes.html>

<https://docs.aws.amazon.com/AmazonCloudFront/latest/DeveloperGuide/live-streaming.html>

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

Question 12: **Incorrect**

An e-commerce company is investigating user reports of its Java-based web application errors on the day of the Thanksgiving sale. The development team recovered the logs created by the EC2 instance-hosted web servers and reviewed Aurora DB cluster performance metrics. Some of the web servers were terminated before logs could be collected and the Aurora metrics were inadequate for query performance analysis.

Which of the following steps would you recommend to make the monitoring process more reliable to troubleshoot any future events due to traffic spikes? (Select three)

-

Use CloudTrail and configure a trail to deliver Amazon Aurora query activity to an Amazon S3 bucket. Process and analyze these real-time log streams using Amazon Kinesis Data Streams

(Incorrect)

-

Install and configure an Amazon CloudWatch Logs agent on the EC2 instances to send the application logs to CloudWatch Logs

(Correct)

-

Enable `detailed monitoring` for Amazon EC2 instances to send data points to CloudWatch every minute. Track the metric 'CPUUtilization' to know when the auto-scaling process can kick in

-

Set up the AWS X-Ray SDK to trace incoming HTTP requests on the EC2 instances as well as set up tracing of SQL queries with the X-Ray SDK for Java

(Correct)

-

Enable `Aurora lab mode` which will then publish all logs and activity on Aurora DB to CloudWatch logs



Configure the Aurora MySQL DB cluster to publish slow query and error logs to Amazon CloudWatch Logs

(Correct)

Explanation

Correct options:

Configure the Aurora MySQL DB cluster to publish slow query and error logs to Amazon CloudWatch Logs

You can configure your Aurora MySQL DB cluster to publish general, slow, audit, and error log data to a log group in Amazon CloudWatch Logs. With CloudWatch Logs, you can perform real-time analysis of the log data, and use CloudWatch to create alarms and view metrics. You can use CloudWatch Logs to store your log records in highly durable storage.

To publish logs to CloudWatch Logs, the respective logs must be enabled. Error logs are enabled by default, but you must enable the other types of logs explicitly. The slow query logs and error logs can be used to identify the root cause behind the given issue.

Install and configure an Amazon CloudWatch Logs agent on the EC2 instances to send the application logs to CloudWatch Logs

You can collect metrics and logs from Amazon EC2 instances and on-premises servers with the CloudWatch agent. The unified CloudWatch agent enables you to collect internal system-level metrics from Amazon EC2 instances across operating systems. The metrics can include in-guest metrics, in addition to the metrics for EC2 instances. You can collect logs from Amazon EC2 instances and on-premises servers, running either Linux or Windows Server. The application logs (via the CloudWatch logs) can be used to identify the root cause behind the given issue.

Set up the AWS X-Ray SDK to trace incoming HTTP requests on the EC2 instances as well as set up tracing of SQL queries with the X-Ray SDK for Java

You can use the X-Ray SDK to trace incoming HTTP requests that your application serves on an EC2 instance. Use a Filter to instrument incoming HTTP requests. When you add the X-Ray servlet filter to your application, the X-Ray SDK for Java creates a

segment for each sampled request. This segment includes timing, method, and disposition of the HTTP request. You can also instrument your SQL database queries by adding the X-Ray SDK for Java JDBC interceptor to your data source configuration. X-Ray tracing for the HTTP requests as well as the SQL queries can help in identifying the root cause behind the given issue.

Incorrect options:

Use CloudTrail and configure a trail to deliver Amazon Aurora query activity to an Amazon S3 bucket. Process and analyze these real-time log streams using Amazon Kinesis Data Streams - You can use CloudTrail to view, search, download, archive, analyze, and respond to account activity across your AWS infrastructure. You can identify who or what took which action, what resources were acted upon, when the event occurred, and other details to help you analyze and respond to activity in your AWS account. CloudTrail provides a record of actions taken by a user, role, or AWS service in Amazon Aurora. However, CloudTrail does not capture any query activity in Aurora, so this option is incorrect.

Enable detailed monitoring for Amazon EC2 instances to send data points to CloudWatch every minute. Track the metric 'CPUUtilization' to know when the auto-scaling process can kick in - Tracking the 'CPUUtilization' parameter is irrelevant to the given use case as it would not point to the root cause behind the given issue.

Enable Aurora lab mode which will then publish all logs and activity on Aurora DB to CloudWatch logs - Aurora lab mode is used to enable Aurora features that are available in the current Aurora database version but are not enabled by default. These features are tested in development/test environments. **Aurora lab mode** is not relevant for capturing the log activity of Aurora DB. This option has been added as a distractor.

References:

<https://docs.aws.amazon.com/AmazonCloudWatch/latest/monitoring/Install-CloudWatch-Agent.html>

<https://docs.aws.amazon.com/AmazonRDS/latest/AuroraUserGuide/AuroraMySQL.Integrating.CloudWatch.html>

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

<https://docs.aws.amazon.com/xray/latest/devguide/xray-sdk-java-filters.html>

<https://docs.aws.amazon.com/xray/latest/devguide/xray-sdk-java-sqlclients.html>

Question 13: **Correct**

A data analytics company uses Amazon S3 as the data lake to store the input data that is ingested from the IoT field devices on an hourly basis. The ingested data has attributes such as the device type, ID of the device, the status of the device, the timestamp of the event, the source IP address, etc. The data runs into millions of records per day and the company wants to run complex analytical queries on this data daily for product improvements for each device type.

Which is the most optimal way to save this data to get the best performance from the millions of data points processed daily?



Store the data in Apache ORC, partitioned by date and sorted by device type of the device

(Correct)



Store the data in compressed .csv, partitioned by date and sorted by the status of the device



Store the data in compressed .csv, partitioned by date and sorted by device type



Store the data in Apache Parquet, partitioned by device type and sorted by date

Explanation

Correct option:

Store the data in Apache ORC, partitioned by date and sorted by device type of the device - Apache Parquet and ORC are columnar storage formats that are optimized for fast retrieval of data and used in AWS analytical applications.

By partitioning your data, you can restrict the amount of data scanned by each query, thus improving performance and reducing cost. You can partition your data by any key. A common practice is to partition the data based on time, often leading to a multi-level partitioning scheme. For example, a customer who has data coming in every hour might decide to partition by year, month, date, and hour. Another customer, who has data coming from many different sources but that is loaded only once per day, might partition by a data source identifier and date.

For the given use case, as the company does daily analysis, so it only needs to look at the data generated for a given date. Hence partitioning by date offers significant performance and cost advantages. Since the company also wants to analyze product improvements for each device type, it is better to keep the data sorted by device type, so it allows for faster query execution.

Incorrect options:

Store the data in Apache Parquet, partitioned by device type and sorted by date - Apache Parquet is a columnar storage format that is optimized for fast retrieval of data and used in AWS analytical applications. However, partitioning by device type is incorrect for this use case, and partitioning by date is optimal.

Store the data in compressed .csv, partitioned by date and sorted by the status of the device

Store the data in compressed .csv, partitioned by date and sorted by device type

Both the above options are not columnar storage formats, they are row-based formats that are not optimal for big data retrievals for complex analytical queries.

Reference:

<https://docs.aws.amazon.com/athena/latest/ug/partitions.html>

Question 14: **Incorrect**

An Amazon S3 bucket is shared by three different teams (managing their own separate AWS accounts) for document uploads. Initially, the S3 bucket settings were set to default. Later, the bucket sees the following updates:

After week 1, S3 Object Ownership bucket-level settings were used and all Access Control Lists (ACLs) were disabled. The three teams uploaded their documents to the shared bucket with this new setting.

After week 2, S3 bucket level settings were again set back to default and the ACLs were enabled once more

What is the outcome of these action(s) on the documents uploaded after week 1 and what are the key points of consideration for future S3 bucket configurations? (Select two)

-

You, as the bucket owner, still own any objects that were written to the bucket while the **bucket owner enforced** setting was applied. These objects are not owned by the **object writer**, even if you re-enable ACLs

(Correct)

-

To simplify permissions management and auditing, use the **Bucket owner preferred** S3 bucket setting

-

You, as the bucket owner, will not own the objects that were written to the bucket while the **bucket owner enforced setting** was applied. These objects will again be owned by the object writer when you re-enable the ACLs

-

If you used object ACLs for permissions management before you applied the **bucket owner enforced** setting and you didn't migrate these object ACL permissions to your bucket policy after you re-enable ACLs, these permissions are restored

(Correct)

-

If you used object ACLs for permissions management before you applied the **bucket owner enforced** setting and you didn't migrate these object ACL permissions to your bucket policy after you re-enable ACLs, these permissions are not restored

(Incorrect)

Explanation

Correct options:

You, as the bucket owner, still own any objects that were written to the bucket while the **bucket owner enforced** setting was applied. These objects are not owned by the **object writer**, even if you re-enable ACLs

If you used object ACLs for permissions management before you applied the **bucket owner enforced** setting and you didn't migrate these object ACL permissions to your bucket policy after you re-enable ACLs, these permissions are restored -

You can re-enable ACLs by changing from the bucket owner-enforced setting to another Object Ownership setting at any time. If you used object ACLs for permissions management before you applied the bucket owner-enforced setting and you didn't migrate these object ACL permissions to your bucket policy, after you re-enable ACLs, these permissions are restored. Additionally, objects written to the bucket while the bucket owner enforced setting was applied are still owned by the bucket owner.

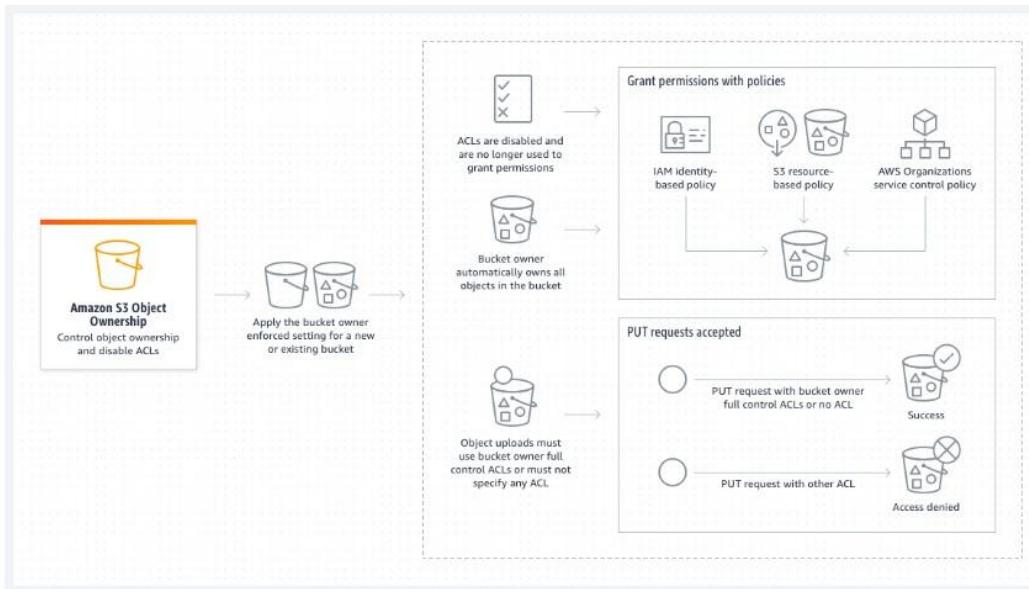
For example, if you change from the bucket owner-enforced setting back to object writer, you, as the bucket owner, no longer own and have full control over objects that were previously owned by other AWS accounts. Instead, the uploading accounts again own these objects. Objects owned by other accounts use ACLs for permissions, so you can't use policies to grant permissions to these objects. However, you, as the bucket owner, still own any objects that were written to the bucket while the bucket owner-enforced setting was applied. These objects are not owned by the object writer, even if you re-enable ACLs.

Changes introduced by disabling

Changes introduced by disabling ACLs

When you apply the bucket owner enforced setting for Object Ownership to disable ACLs, you automatically own and take full control over every object in the bucket without taking any additional actions. After you apply this setting, you will see three changes:

- All bucket ACLs and object ACLs are disabled, which gives full access to you, as the bucket owner. When you perform a read ACL request on your bucket or object, you will see that full access is given only to the bucket owner.
- You, as the bucket owner, automatically own and have full control over every object in your bucket.
- ACLs no longer affect access permissions to your bucket. As a result, access control for your data is based on policies, such as IAM policies, S3 bucket policies, VPC endpoint policies, and Organizations SCPs.



ACLs:

via -

<https://docs.aws.amazon.com/AmazonS3/latest/userguide/about-object-ownership.html>

Incorrect options:

If you used object ACLs for permissions management before you applied the bucket owner enforced setting and you didn't migrate these object ACL permissions to your bucket policy after you re-enable ACLs, these permissions are not restored - As explained above, the permissions are restored for this scenario. So this option is incorrect.

To simplify permissions management and auditing, use the Bucket owner preferred S3 bucket setting - This statement is incorrect. AWS recommends that you disable ACLs by choosing the **bucket owner enforced** setting and use your bucket policy to share data with users outside of your account as needed. This approach simplifies permissions management and auditing. You can disable ACLs on both newly created and already existing buckets.

You, as the bucket owner, will not own the objects that were written to the bucket while the bucket owner enforced setting was applied. These objects will again be owned by the object writer, when you re-enable the ACLs - This statement is incorrect. You, as the bucket owner, still own any objects that were written to the bucket while the bucket owner-enforced setting was applied. These objects are not owned by the object writer, even if you re-enable ACLs.

Reference:

<https://docs.aws.amazon.com/AmazonS3/latest/userguide/about-object-ownership.html>

Question 15: **Incorrect**

A leading pharmaceutical company has significant investments in running Oracle and PostgreSQL services on Amazon RDS which provide their scientists with near real-time analysis of millions of rows of manufacturing data generated by continuous manufacturing equipment with 1,600 data points per row. The business analytics team has been running ad-hoc queries on these databases to prepare daily reports for senior management. The engineering team has observed that the database performance takes a hit whenever these reports are run by the analytics team. To facilitate the business analytics reporting, the engineering team now wants to replicate this data with high availability and consolidate these databases into a petabyte-scale data warehouse by streaming data to Amazon Redshift.

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

-

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

-

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



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

(Correct)



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

(Incorrect)

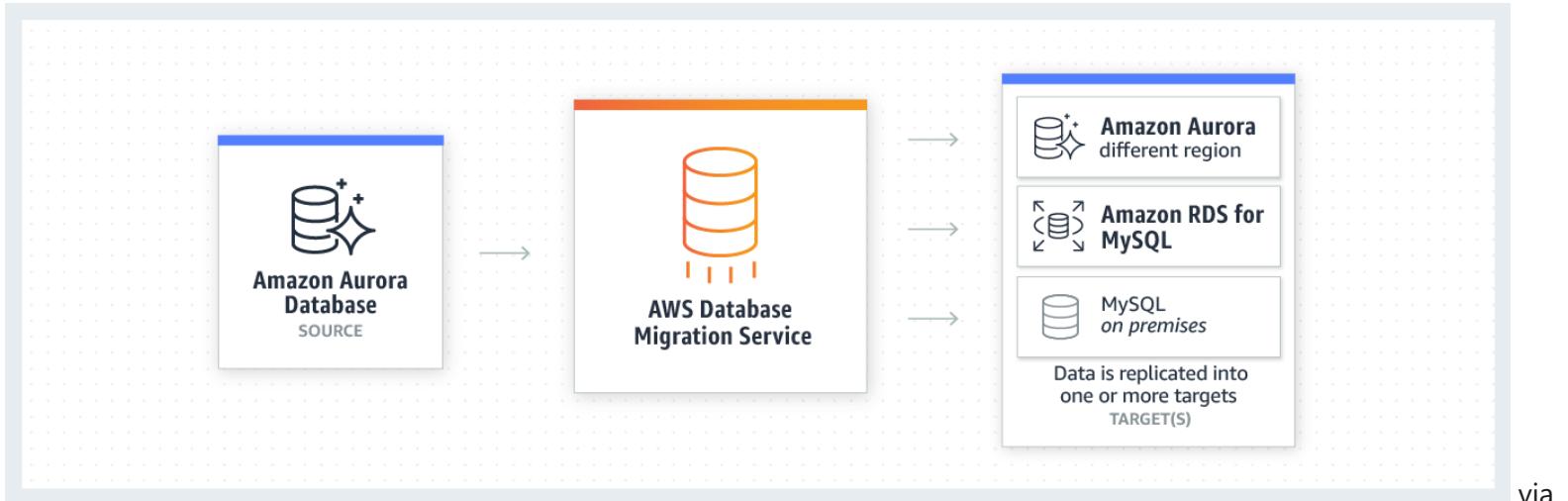
Explanation

Correct option:

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

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

Continuous Data



Replication

- <https://aws.amazon.com/dms/>

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

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

Incorrect options:

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

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

Use Amazon EMR to replicate the data from the databases into Amazon Redshift - Amazon EMR is the industry-leading cloud big data platform for processing vast amounts of data using open source tools such as Apache Spark, Apache Hive, Apache HBase, Apache Flink and Presto. With EMR you can run Petabyte-scale analysis at less than half of the cost of traditional on-premises solutions and over 3x faster than standard Apache Spark. For short-running jobs, you can spin up and spin down clusters and pay per second for the instances used. For long-running workloads, you can create highly available clusters that automatically scale to meet demand. Amazon EMR uses Hadoop, an open-source framework, to distribute your data and processing across a resizable cluster of Amazon EC2 instances.

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

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

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

References:

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

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

Question 16: **Incorrect**

A company manages a stateful web application that persists data on a MySQL database. The application stack is hosted in the company's on-premises data center using a single server. The company is looking at increasing its market presence through promotions and campaigns. While the user experience has been good so far, the current application architecture will not support the growth that the company envisages. The company has hired you as an AWS Certified Solutions Architect Professional to migrate the

current architecture to AWS which should continue to support SQL-based queries. The proposed solution should offer maximum reliability with better performance.

What would you recommend?



Set up database migration to Amazon Aurora MySQL. Deploy the application in an Auto Scaling group for Amazon EC2 instances that are fronted by an Application Load Balancer. Store sessions in an Amazon ElastiCache for Redis with replication group

(Correct)



Set up database migration to an Amazon RDS MySQL Multi-AZ DB instance. Deploy the application in an Auto Scaling group for Amazon EC2 instances that are fronted by an Application Load Balancer. Store sessions in Amazon ElastiCache for Memcached



Set up database migration to an Amazon RDS MySQL DB instance using read replicas. Deploy the application in an Auto Scaling group for Amazon EC2 instances that are fronted by an Application Load Balancer. Store sessions using Amazon Neptune

(Incorrect)



Set up database migration to an Amazon DocumentDB instance. Deploy the application in an Auto Scaling group for Amazon EC2 instances that are fronted by a Network Load Balancer. Store sessions in an Amazon ElastiCache for Redis with replication group

Explanation

Correct option:

Set up database migration to Amazon Aurora MySQL. Deploy the application in an Auto Scaling group for Amazon EC2 instances that are fronted by an Application Load Balancer. Store sessions in an Amazon ElastiCache for Redis with replication group

Amazon Aurora is designed for unparalleled high performance and availability at a global scale with full MySQL and PostgreSQL compatibility.

Amazon ElastiCache for Redis is a Redis-compatible in-memory service that delivers the ease of use and power of Redis along with the availability, reliability, and performance suitable for the most demanding applications. Both single-node and up to 15-shard clusters are available, enabling scalability to up to 3.55 TiB of in-memory data. Amazon ElastiCache for Redis is fully managed, scalable, and secure. This makes it an ideal candidate to power high-performance use cases such as web, mobile apps, gaming, ad tech, and IoT.

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

Redis lets you create multiple replicas of a Redis primary. This allows you to scale database reads and to have highly available clusters. The replication support makes Redis a more reliable solution than Memcached.

Choosing between Redis and Memcached

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

[Learn about Amazon ElastiCache for Redis](#)

[Learn about Amazon ElastiCache for Memcached](#)

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

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

Incorrect options:

Set up database migration to an Amazon RDS MySQL Multi-AZ DB instance. Deploy the application in an Auto Scaling group for Amazon EC2 instances that are fronted by an Application Load Balancer. Store sessions in Amazon ElastiCache for Memcached - As explained above, the replication support makes Redis a more reliable solution than Memcached. So this option is not the best fit.

Set up database migration to an Amazon RDS MySQL DB instance using read replicas. Deploy the application in an Auto Scaling group for Amazon EC2 instances that are fronted by an Application Load Balancer. Store sessions using Amazon Neptune - Amazon Neptune is a fast, fully managed database service powering graph use cases such as identity graphs, knowledge graphs, and fraud detection. You cannot use Neptune as a caching layer for storing user sessions.

Set up database migration to an Amazon DocumentDB instance. Deploy the application in an Auto Scaling group for Amazon EC2 instances that are fronted by a Network Load Balancer. Store sessions in an Amazon ElastiCache for Redis with replication group - Amazon DocumentDB is a scalable, highly durable, and fully managed database service for operating mission-critical MongoDB workloads. For the given use case, you cannot migrate the relational database (MySQL) into a document database (DocumentDB) as DocumentDB does not support SQL-based queries.

References:

<https://aws.amazon.com/elasticsearch/redis-vs-memcached/>

<https://docs.aws.amazon.com/whitepapers/latest/aws-overview/database.html>

<https://docs.aws.amazon.com/AmazonElastiCache/latest/red-ug/Replication.Redis.Groups.html>

Question 17: **Incorrect**

A financial services company wants to set up an AWS WAF-based solution to manage AWS WAF rules across multiple AWS accounts that are structured under different Organization Units (OUs) in AWS Organizations. The solution should automatically update and remediate noncompliant AWS WAF rules in all accounts. The solution should also facilitate adding or removing accounts or OUs from managed AWS WAF rule sets as needed.

Which of the following solutions is the most operationally efficient to address the given use case?

- Use AWS Firewall Manager to manage AWS WAF rules across accounts in the organization. Leverage AWS Systems Manager Parameter Store to store account numbers and OUs. Update AWS Systems Manager Parameter Store as needed to add or remove accounts or OUs. Create cross-account IAM roles in member accounts with permissions to create and update AWS WAF rules. Create a Lambda function to assume IAM roles in the management account to create and update AWS WAF rules in the member accounts

- Use AWS Control Tower to manage AWS WAF rules across accounts in the organization. Leverage AWS Secrets Manager to store account numbers and OUs. Update AWS Secrets Manager as needed to add or remove accounts or OUs. Create cross-account IAM roles in member accounts with permissions to create and update AWS WAF rules. Create a Lambda function to assume IAM roles in the management account to create and update AWS WAF rules in the member accounts

- Create an AWS Organizations organization-wide AWS Config rule that mandates all resources in the selected OUs to be associated with the AWS WAF rules. Configure automated remediation actions by using AWS Systems Manager Automation documents to fix non-compliant resources. Set up AWS WAF rules by using an AWS CloudFormation stack set to target the same OUs where the AWS Config rule is applied

(Correct)

- Use AWS Security Hub to manage AWS WAF rules across accounts in the organization. Leverage AWS KMS to store account numbers and OUs. Update AWS KMS as needed to add or remove accounts or OUs. Create IAM users in member accounts. Allow AWS Firewall Manager in the management account to use the access key and secret access key to create and update AWS WAF rules in the member accounts

(Incorrect)

Explanation

Correct option:

Create an AWS Organizations organization-wide AWS Config rule that mandates all resources in the selected OUs to be associated with the AWS WAF rules. Configure automated remediation actions by using AWS Systems Manager Automation documents to fix non-compliant resources. Set up AWS WAF rules by using an AWS CloudFormation stack set to target the same OUs where the AWS Config rule is applied

AWS Config allows you to manage AWS Config rules across all AWS accounts within an organization. You can:

Centrally create, update, and delete AWS Config rules across all accounts in your organization.

Deploy a common set of AWS Config rules across all accounts and specify accounts where AWS Config rules should not be created.

Use the APIs from the management account in AWS Organizations to enforce governance by ensuring that the underlying AWS Config rules are not modifiable by your organization's member accounts.

If a new account joins an organization, the rule or conformance pack is deployed to that account. When an account leaves an organization, the rule or conformance pack is removed.

Managing AWS Config Rules Across All Accounts in Your Organization

[PDF](#) | [RSS](#)

AWS Config allows you to manage AWS Config rules across all AWS accounts within an organization. You can:

- Centrally create, update, and delete AWS Config rules across all accounts in your organization.
- Deploy a common set of AWS Config rules across all accounts and **specify accounts where AWS Config rules should not be created**.
- Use the APIs from the management account in AWS Organizations to enforce governance by ensuring that the underlying AWS Config rules are not modifiable by your organization's member accounts.

Note

For deployments across different regions

The API call to deploy rules and conformance packs across accounts is region specific. At the organization level, you need to change the context of your API call to a different region if you want to deploy rules in other regions. For example, to deploy a rule in US East (N. Virginia), change the region to US East (N. Virginia) and then call `PutOrganizationConfigRule`.

For accounts within an organization

If a new account joins an organization, the rule or conformance pack is deployed to that account. When an account leaves an organization, the rule or conformance pack is removed.

If you deploy an organizational rule or conformance pack in an organization administrator account, and then establish a delegated administrator and deploy an organizational rule or conformance pack in the delegated administrator account, you won't be able to see the organizational rule or conformance pack in the organization administrator account from the delegated administrator account or see the organizational rule or conformance pack in the delegated administrator account from organization administrator account. The `DescribeOrganizationConfigRules` and `DescribeOrganizationConformancePacks` APIs can only see and interact with the organization-related resource that were deployed from within the account calling those APIs.

Retry mechanism for new accounts added to an organization

Deployment of existing organizational rules and conformance packs will only be retried for 7 hours after an account is added to your organization if a recorder is not available. You are expected to create a recorder if one doesn't exist within 7 hours of adding an account to your organization.

via -

<https://docs.aws.amazon.com/config/latest/developerguide/config-rule-multi-account-deployment.html>

AWS Config allows you to remediate noncompliant resources that are evaluated by AWS Config Rules. AWS Config applies remediation using AWS Systems Manager Automation documents. These documents define the actions to be performed on noncompliant AWS resources evaluated by AWS Config Rules. You can associate SSM documents by using AWS Management Console or by using APIs. To apply remediation on non-compliant resources, you can either choose the remediation action you want to associate from a prepopulated list or create your own custom remediation actions using SSM documents. AWS Config provides a recommended list of remediation actions in the AWS Management Console.

AWS CloudFormation StackSets extends the capability of CloudFormation stacks by enabling you to create, update, or delete stacks across multiple accounts and AWS Regions with a single operation. Using an administrator account, you define and manage an AWS CloudFormation template, and use the template as the basis for provisioning stacks into selected target accounts across specified AWS Regions.

Incorrect options:

Use AWS Firewall Manager to manage AWS WAF rules across accounts in the organization. Leverage AWS Systems Manager Parameter Store to store account numbers and OUs. Update AWS Systems Manager Parameter Store as needed to add or remove accounts or OUs. Create cross-account IAM roles in member accounts with permissions to create and update AWS WAF rules. Create a Lambda function to assume IAM roles in the management account to create and update AWS WAF rules in the member accounts - This option involves significant manual work every time an account is added/removed from the organization. You need to update the items in Systems Manager Parameter Store and further update the Lambda to assume the role for the new account. Hence this option is incorrect.

Use AWS Control Tower to manage AWS WAF rules across accounts in the organization. Leverage AWS Secrets Manager to store account numbers and OUs. Update AWS Secrets Manager as needed to add or remove accounts or OUs. Create cross-account IAM roles in member accounts with permissions to create and update AWS WAF rules. Create a Lambda function to assume IAM roles in the management account to create and update AWS WAF rules in the member accounts - This option involves significant manual work every time an account is added/removed from the organization. You need to update the items in Secrets Manager and further update the Lambda to assume the role for the new account. Hence this option is incorrect.

Use AWS Security Hub to manage AWS WAF rules across accounts in the organization. Leverage AWS KMS to store account numbers and OUs. Update AWS KMS as needed to add or remove accounts or OUs. Create IAM users in member accounts. Allow AWS Firewall Manager in the management account to use the access key and secret access key to create and update AWS WAF rules in the member accounts - This option has been added as a distractor. You cannot use AWS Security Hub to manage AWS WAF rules across accounts in the organization, rather you need to use AWS Firewall Manager to accomplish this. AWS KMS is a managed service that helps you more easily create and control the keys used for cryptographic operations. The service provides a highly available key generation, storage, management, and auditing solution for you to encrypt or digitally sign data within your own applications or control the encryption of data across AWS services. You cannot use AWS KMS to store account numbers and OUs.

References:

<https://docs.aws.amazon.com/config/latest/developerguide/config-rule-multi-account-deployment.html>

<https://aws.amazon.com/about-aws/whats-new/2019/12/aws-security-hub-integrates-with-aws-firewall-manager/>

<https://docs.aws.amazon.com/config/latest/developerguide/remediation.html>

<https://aws.amazon.com/premiumsupport/knowledge-center/lambda-function-assume-iam-role/>

Question 18: **Incorrect**

A company manages a healthcare diagnostics application that writes thousands of lab images to a mounted NFS file system each night from 10 PM - 5 AM. The company wants to migrate this application from its on-premises data center to AWS Cloud over a private network. The company has already established an AWS Direct Connect connection to AWS to facilitate this migration. This application is slated to be moved to Amazon EC2 instances with the Elastic File System (Amazon EFS) file system as the storage service.

Which of the following represents the MOST optimal way of replicating all images to the cloud before the application is fully migrated to the cloud?



Deploy an AWS DataSync agent to an on-premises server that has access to the NFS file system. Connect to AWS VPC endpoint for EFS over a public VIF of the Direct Connect connection. Configure a DataSync scheduled task to send the images to the EFS file system every night

(Incorrect)



Define a cron job on the on-premises system to run the AWS s3 sync command from the on-premises file system to Amazon S3. Use the Amazon S3 Event Notifications to call a Lambda function that will copy the images from the S3 bucket to the EFS file system



Deploy an AWS DataSync agent to an on-premises server that has access to the NFS file system. Send data over the Direct Connect connection to an AWS PrivateLink interface VPC endpoint for Amazon EFS by using a private VIF. Configure a DataSync scheduled task to send the images to the EFS file system every night

(Correct)

-
-

Create an NFS file share using AWS Storage Gateway file gateway. Mount your NFS file share on a drive on your client and map it to your Amazon S3 bucket. Configure an AWS Lambda function to process event notifications from Amazon S3 and copy the images from Amazon S3 to the EFS file system

Explanation

Correct option:

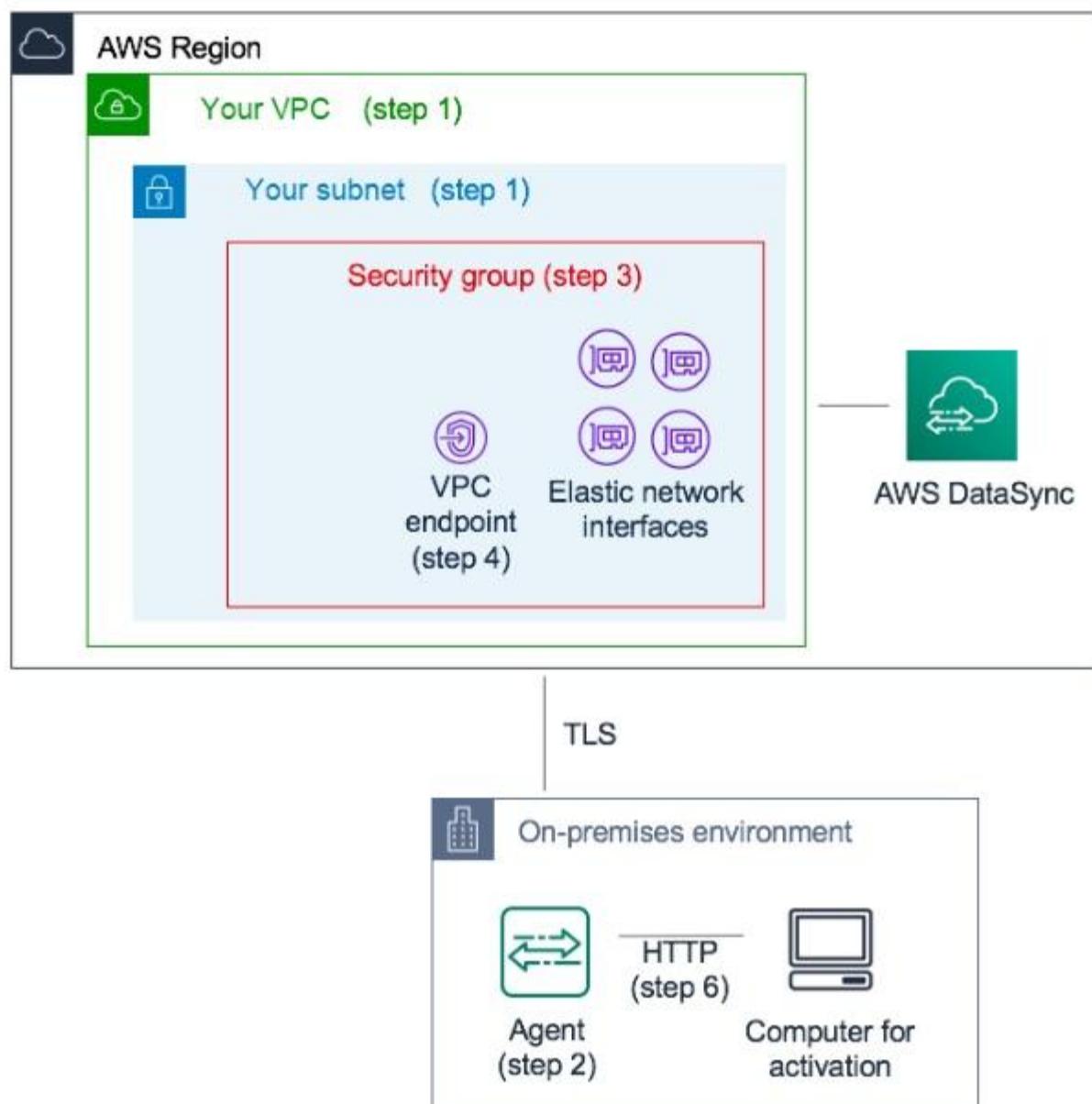
Deploy an AWS DataSync agent to an on-premises server that has access to the NFS file system. Send data over the Direct Connect connection to an AWS PrivateLink interface VPC endpoint for Amazon EFS by using a private VIF. Configure a DataSync scheduled task to send the images to the EFS file system every night

You can use VPC endpoints to ensure data transferred between your AWS DataSync agent, either deployed on-premises or in-cloud, doesn't traverse the public internet or need public IP addresses. Using VPC endpoints increases the security of your data by keeping network traffic within your Amazon Virtual Private Cloud (Amazon VPC). VPC endpoints for DataSync are powered by AWS PrivateLink, a highly available, scalable technology that enables you to privately connect your VPC to supported AWS services.

The DataSync agent transfers data between your storage and AWS. In most situations, you deploy the agent as a virtual machine in the same local network as your source storage. This approach minimizes network overhead associated with transferring data by using network protocols such as Network File System (NFS) and Server Message Block (SMB) or when accessing your object storage that's compatible with the Amazon S3 API. This setup is common regardless of the endpoint type you use to connect your agent to AWS.

When you use a VPC endpoint, your DataSync agent communicates directly with AWS without crossing the public internet. Data is transferred using AWS Direct Connect or a virtual private network (VPN). The private IP addresses that DataSync creates for the transfer are accessible only from inside your VPC.

Reference architecture of using DataSync with VPC



To configure a DataSync agent and task to communicate with AWS by using VPC endpoints
endpoints:

via

- <https://docs.aws.amazon.com/datasync/latest/userguide/datasync-in-vpc.html>

Incorrect options:

Create an NFS file share using the AWS Storage Gateway file gateway. Mount your NFS file share on a drive on your client and map it to your Amazon S3 bucket. Configure an AWS Lambda function to process event notifications from Amazon S3 and copy the images from Amazon S3 to the EFS file system - AWS Storage Gateway provides seamless access to data in hybrid architectures. When the entire application is being moved to AWS Cloud, AWS Storage Gateway is not the best fit for the given use case. In addition, the data is initially copied to S3 and then replicated into EFS, thereby making the process inefficient.

Define a cron job on the on-premises system to run the AWS s3 sync command from the on-premises file system to Amazon S3. Use the Amazon S3 Event Notifications to call a Lambda function that will copy the images from the S3 bucket to the EFS file system - The data is initially copied to S3 and then replicated into EFS, thereby making the process inefficient.

Deploy an AWS DataSync agent to an on-premises server that has access to the NFS file system. Connect to AWS VPC endpoint for EFS over a public VIF of the Direct Connect connection. Configure a DataSync scheduled task to send the images to the EFS file system every night - A VPC endpoint allows you to privately connect your VPC to supported AWS services without requiring an internet gateway or a NAT device, VPN connection, or AWS Direct Connect connection. A public virtual interface (VIF) can access all AWS public services using public IP addresses. You cannot leverage public VIF to access the VPC endpoint for EFS. Therefore this option is incorrect.

References:

<https://docs.aws.amazon.com/datasync/latest/userguide/datasync-in-vpc.html>

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

Question 19: **Correct**

An e-commerce company manages its flagship application on a load-balanced EC2 instance fleet for web hosting, database API services, and business logic. This tightly coupled architecture makes it inflexible for new feature additions while also making the architecture less scalable.

Which of the following options can be used to decouple the architecture, improve scalability and provide the ability to track the failed orders?



Use AWS Elastic Beanstalk for hosting the web application and Amazon API Gateway for database API services. Use Kinesis Data Streams for queuing orders and AWS Lambda to build business logic. Configure an Amazon S3 bucket for retaining failed orders on an hourly basis



Configure Amazon CloudFront for hosting the website and Amazon API Gateway for database API services. Use Amazon Simple Queue Service (Amazon SQS) for order queuing and AWS Lambda for business logic. Use Amazon SQS long polling for retaining failed orders



Configure Amazon S3 for hosting the web application while using AWS AppSync for database access services. Use Amazon Simple Queue Service (Amazon SQS) for queuing orders and AWS Lambda for business logic. Use Amazon SQS dead-letter queue for tracking and re-processing failed orders

(Correct)



Use Amazon Lightsail for web hosting with AWS AppSync for database API services. Use Simple Queue Service (Amazon SQS) for order queuing. Use Amazon Elastic Container Service (Amazon ECS) for business logic and use the **visibility timeout parameter of Amazon SQS to retain the failed orders**

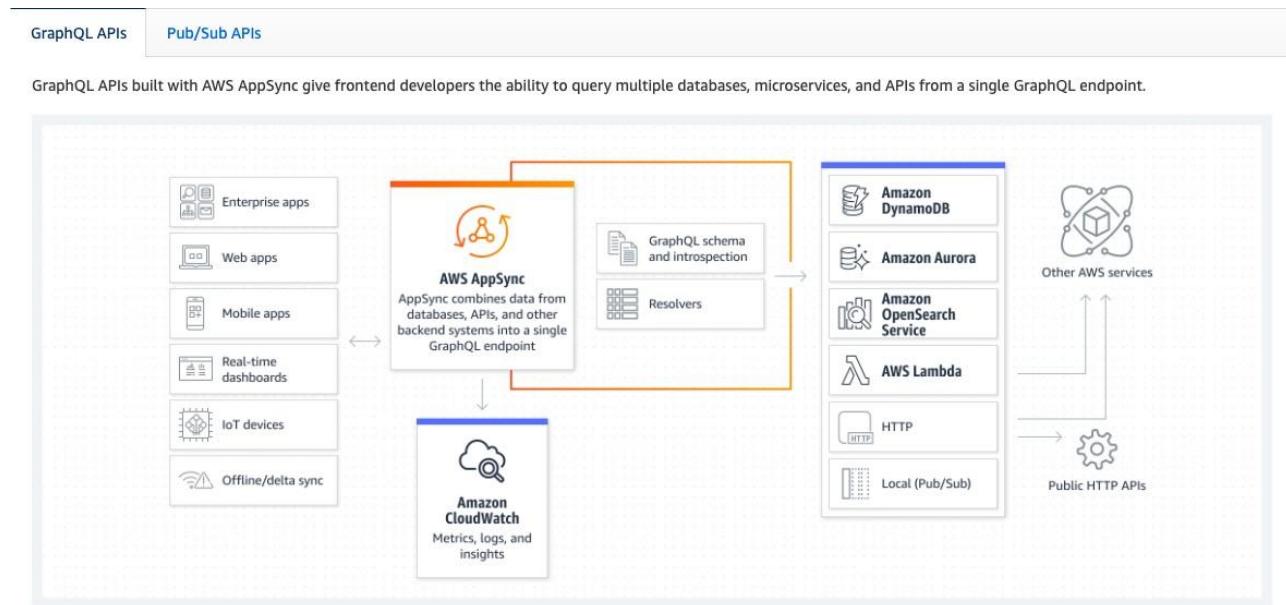
Explanation

Correct option:

Configure Amazon S3 for hosting the web application while using AWS AppSync for database access services. Use Amazon Simple Queue Service (Amazon SQS) for queuing orders and AWS Lambda for business logic. Use Amazon SQS dead-letter queue for tracking and re-processing failed orders

Amazon S3 can be configured to host a web application.

AWS AppSync creates serverless GraphQL and Pub/Sub APIs that simplify application development through a single endpoint to securely query, update, or publish data. AWS AppSync creates serverless GraphQL and Pub/Sub APIs that simplify application development through a single endpoint to securely query, update, or publish data.



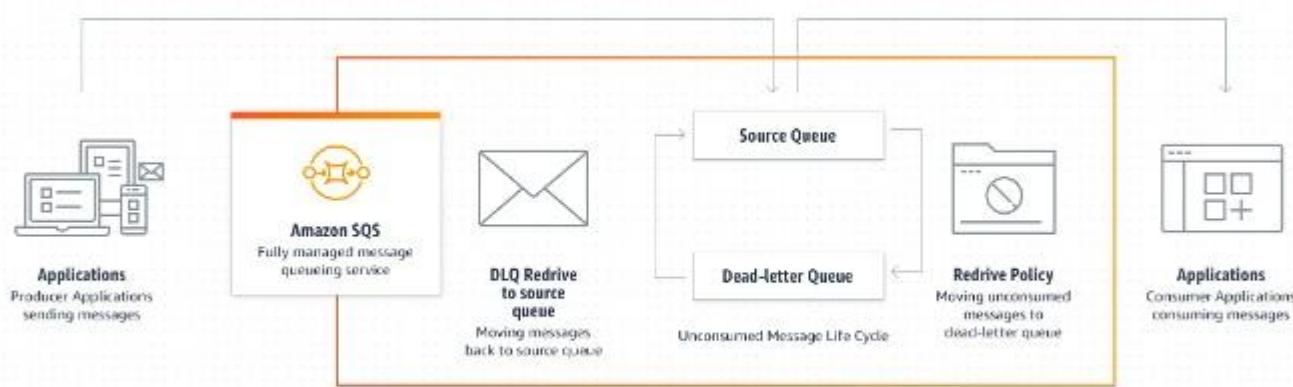
How AWS AppSync works:

<https://aws.amazon.com/appsync/>

via -

Amazon SQS supports dead-letter queues (DLQ), which other queues (source queues) can target for messages that can't be processed (consumed) successfully. Dead-letter queues are useful for debugging your application or messaging system because they let you isolate unconsumed messages to determine why their processing doesn't succeed.

SQS dead-letter



queues:

via - <https://docs.aws.amazon.com/AWSSimpleQueueService/latest/SQSDriverGuide/sqs-dead-letter-queues.html>

Incorrect options:

Configure Amazon CloudFront for hosting the website and Amazon API Gateway for database API services. Use Amazon Simple Queue Service (Amazon SQS) for order queuing and AWS Lambda for business logic. Use Amazon SQS long polling for retaining failed orders - You cannot use Amazon CloudFront for hosting a website as the website is hosted on the Cloudfront distribution's underlying origin (such as S3 or an EC2 instance). You cannot use Amazon SQS long polling for retaining failed orders. When the wait time for the `ReceiveMessage` API action is greater than 0, long polling is in effect. Long polling helps reduce the cost of using Amazon SQS by eliminating the number of empty responses and false empty responses. Long polling is a configurable parameter of SQS queues and not a temporary storage space to hold failed orders.

Use Amazon Lightsail for web hosting with AWS AppSync for database API services. Use Simple Queue Service (Amazon SQS) for order queuing. Use Amazon Elastic Container Service (Amazon ECS) for business logic and use the visibility timeout parameter of Amazon SQS to retain the failed orders - You cannot use the `visibility timeout` parameter of Amazon SQS to retain the failed orders. Immediately after a message is received in an SQS queue, it remains in the queue. To prevent other consumers from processing the message again, Amazon SQS sets a `visibility timeout`, a period during which Amazon SQS

prevents other consumers from receiving and processing the message. **Visibility timeout** is a configurable parameter of SQS queues and not a temporary storage space to hold failed orders.

Use AWS Elastic Beanstalk for hosting the web application and Amazon API Gateway for database API services. Use Kinesis Data Streams for queuing orders and AWS Lambda to build business logic. Configure an Amazon S3 bucket for retaining failed orders on an hourly basis - Amazon Kinesis Streams allows real-time processing of streaming big data and the ability to read and replay records to multiple Amazon Kinesis Applications. Amazon SQS offers a reliable, highly-scalable hosted queue for storing messages as they travel between applications or microservices. It moves data between distributed application components and helps you decouple these components.

You should not use S3 to retain failed orders on an hourly basis. This would result in too many small objects (1 object for each failed order) on S3 which need to be written and read multiple times. In addition, it would be cumbersome to keep track of the failed orders and do the root cause analysis. You could run SQL queries via Athena on this underlying data in S3. However, it would turn out to be costly and inefficient while querying small objects via Athena. Therefore, this use case is an anti-pattern for S3.

References:

<https://docs.aws.amazon.com/AmazonS3/latest/userguide/WebsiteHosting.html>

<https://aws.amazon.com/appsync/>

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

<https://aws.amazon.com/blogs/big-data/top-10-performance-tuning-tips-for-amazon-athena/>

<https://docs.aws.amazon.com/AmazonS3/latest/userguide/optimizing-performance-design-patterns.html>

Question 20: **Correct**

A company uses Elastic Load Balancing to distribute traffic across multiple Amazon EC2 instances. Auto Scaling groups start and stop Amazon EC2 machines based on the number of incoming requests. The company has recently started operations in a new AWS Region and is setting up an Application Load Balancer for its fleet of EC2 instances spread across two Availability Zones, with one instance as a target in Availability Zone X and four instances as targets in Availability Zone Y. The company is doing benchmarking for

server performance in the new Region for the case when cross-zone load balancing is enabled compared to the case when cross-zone load balancing is disabled.

As a Solutions Architect Professional, which of the following traffic distribution outcomes would you identify as correct?



With cross-zone load balancing enabled, one instance in Availability Zone X receives 50% traffic and four instances in Availability Zone Y receive 12.5% traffic each. With cross-zone load balancing disabled, one instance in Availability Zone X receives 20% traffic and four instances in Availability Zone Y receive 20% traffic each



With cross-zone load balancing enabled, one instance in Availability Zone X receives 20% traffic and four instances in Availability Zone Y receive 20% traffic each. With cross-zone load balancing disabled, one instance in Availability Zone X receives no traffic and four instances in Availability Zone Y receive 25% traffic each



With cross-zone load balancing enabled, one instance in Availability Zone X receives no traffic and four instances in Availability Zone Y receive 25% traffic each. With cross-zone load balancing disabled, one instance in Availability Zone X receives 50% traffic and four instances in Availability Zone Y receive 12.5% traffic each



With cross-zone load balancing enabled, one instance in Availability Zone X receives 20% traffic and four instances in Availability Zone Y receive 20% traffic each. With cross-zone load balancing disabled, one instance in Availability Zone X receives 50% traffic and four instances in Availability Zone Y receive 12.5% traffic each

(Correct)

Explanation

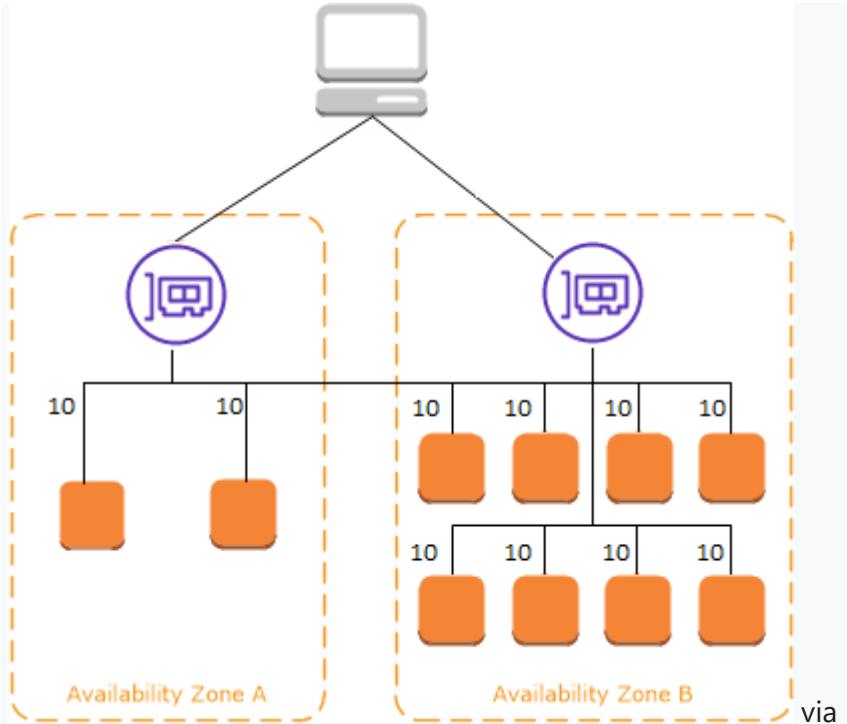
Correct option:

With cross-zone load balancing enabled, one instance in Availability Zone X receives 20% traffic and four instances in Availability Zone Y receive 20% traffic each. With cross-zone load balancing disabled, one instance in Availability Zone X receives 50% traffic and four instances in Availability Zone Y receive 12.5% traffic each

The nodes for your load balancer distribute requests from clients to registered targets. When cross-zone load balancing is enabled, each load balancer node distributes traffic across the registered targets in all enabled Availability Zones. Therefore, one instance in Availability Zone X receives 20% traffic and four instances in Availability Zone Y receive 20% traffic each. When cross-zone load balancing is disabled, each load balancer node distributes traffic only across the registered targets in its Availability Zone. Therefore, one instance in Availability Zone X receives 50% traffic and four instances in Availability Zone Y receive 12.5% traffic each.

Consider the following diagrams (the scenario illustrated in the diagrams involves 10 target instances split across 2 AZs) to understand the effect of cross-zone load balancing.

If cross-zone load balancing is enabled, each of the 10 targets receives 10% of the traffic. This is because each load balancer node can route its 50% of the client traffic to all 10 targets.



via -

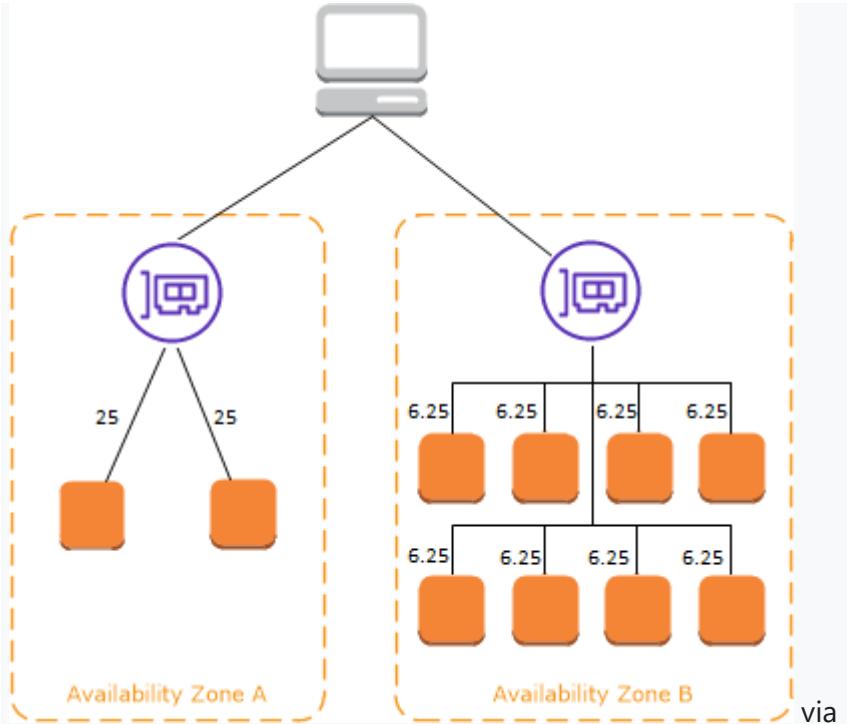
<https://docs.aws.amazon.com/elasticloadbalancing/latest/userguide/how-elastic-load-balancing-works.html>

If cross-zone load balancing is disabled:

Each of the two targets in Availability Zone X receives 25% of the traffic.

Each of the eight targets in Availability Zone Y receives 6.25% of the traffic.

This is because each load balancer node can route its 50% of the client traffic only to targets in its Availability Zone



<https://docs.aws.amazon.com/elasticloadbalancing/latest/userguide/how-elastic-load-balancing-works.html>

Incorrect options:

With cross-zone load balancing enabled, one instance in Availability Zone X receives 50% traffic and four instances in Availability Zone Y receive 12.5% traffic each. With cross-zone load balancing disabled, one instance in Availability Zone X receives 20% traffic and four instances in Availability Zone Y receive 20% traffic each

With cross-zone load balancing enabled, one instance in Availability Zone X receives no traffic and four instances in Availability Zone Y receive 25% traffic each. With cross-zone load balancing disabled, one instance in Availability Zone X receives 50% traffic and four instances in Availability Zone Y receive 12.5% traffic each

With cross-zone load balancing enabled, one instance in Availability Zone X receives 20% traffic and four instances in Availability Zone Y receive 20% traffic each. With cross-zone load balancing disabled, one instance in Availability Zone X receives no traffic and four instances in Availability Zone Y receive 25% traffic each

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

Reference:

<https://docs.aws.amazon.com/elasticloadbalancing/latest/userguide/how-elastic-load-balancing-works.html>

Question 21: **Incorrect**

A financial services company has multiple AWS accounts hosting its portfolio of IT applications that serve the company's retail and enterprise customers. A CloudWatch Logs agent is installed on each of the EC2 instances running these IT applications. The company wants to aggregate all security events in a centralized AWS account dedicated to log storage. The centralized operations team at the company needs to perform near-real-time gathering and collating events across multiple AWS accounts.

As a Solutions Architect Professional, which of the following solutions would you suggest to meet these requirements?

-

Set up Kinesis Data Firehose in the logging account and then subscribe the delivery stream to CloudWatch Logs streams in each application AWS account via subscription filters. Persist the log data in an Amazon S3 bucket inside the logging AWS account

(Correct)

-

Set up CloudWatch Logs streams in each application AWS account to forward events to CloudWatch Logs in the centralized logging AWS account. In the centralized logging AWS account, subscribe a Kinesis Data Firehose stream to Amazon EventBridge events and further use the Firehose stream to store the log data in S3

(Incorrect)

- Set up CloudWatch Logs agents to publish data to a Kinesis Data Firehose stream in the centralized logging AWS account. Create a Lambda function to read messages from the stream and push messages to Kinesis Data Firehose and then store the data in S3
- Set up a new IAM role in each application AWS account with permissions to view CloudWatch Logs. Create a Lambda function to assume this new role and perform an hourly export of each AWS account's CloudWatch Logs data to an S3 bucket in the centralized logging AWS account

Explanation

Correct option:

Set up Kinesis Data Firehose in the logging account and then subscribe the delivery stream to CloudWatch Logs streams in each application AWS account via subscription filters. Persist the log data in an Amazon S3 bucket inside the logging AWS account

You can configure Amazon Kinesis Data Firehose to aggregate and collate CloudWatch Logs from different AWS accounts and receive their log events in a centralized logging AWS Account (this is known as cross-account data sharing) by using a CloudWatch Logs destination and then creating a Subscription Filter. This log event data can be read from a centralized Amazon Kinesis Firehose delivery stream to perform downstream processing and analysis.

You can collaborate with an owner of a different AWS account and receive their log events on your AWS resources, such as an Amazon Kinesis or Amazon Kinesis Data Firehose stream (this is known as cross-account data sharing). You can use a subscription filter with Kinesis Streams, Lambda, or Kinesis Data Firehose. Logs that are sent to a receiving service through a subscription filter are Base64 encoded and compressed with the gzip format.

Incorrect options:

Set up a new IAM role in each application AWS account with permissions to view CloudWatch Logs. Create a Lambda function to assume this new role and perform an hourly export of each AWS account's CloudWatch Logs data to an S3 bucket in the centralized logging AWS account - As the Lambda function is performing an hourly export, so it's not a near-real

time solution. In addition, Lambda is not the right choice to build a high volume and high-velocity streaming solution which is better handled by using the Kinesis Family of services.

Set up CloudWatch Logs agents to publish data to a Kinesis Data Firehose stream in the centralized logging AWS account. Create a Lambda function to read messages from the stream and push messages to Kinesis Data Firehose and then store the data in S3 - The CloudWatch Logs agent (on the path to deprecation) supports the collection of logs from only servers running Linux. It is recommended to use the unified CloudWatch agent. It enables you to collect both logs and advanced metrics with one agent. It offers support across operating systems, including servers running Windows Server. This agent also provides better performance. CloudWatch Logs agent cannot publish data to a Kinesis Data Firehose stream, so this option is incorrect.

Set up CloudWatch Logs streams in each application AWS account to forward events to CloudWatch Logs in the centralized logging AWS account. In the centralized logging AWS account, subscribe a Kinesis Data Firehose stream to Amazon EventBridge events and further use the Firehose stream to store the log data in S3 - You can use a subscription filter with Kinesis Streams, Lambda, or Kinesis Data Firehose. So you cannot just forward events directly to CloudWatch Logs in another account. In addition, Kinesis Data Firehose stream cannot subscribe to EventBridge events, so this option is incorrect.

References:

<https://docs.aws.amazon.com/AmazonCloudWatch/latest/logs/CrossAccountSubscriptions.html>

<https://docs.aws.amazon.com/AmazonCloudWatch/latest/logs/SubscriptionFilters.html>

<https://aws.amazon.com/blogs/architecture/stream-amazon-cloudwatch-logs-to-a-centralized-account-for-audit-and-analysis/>

Question 22: Incorrect

A retail company is introducing multiple business units as part of its expansion plans. To implement this change, the company will be building several new business-unit-specific workloads by leveraging a variety of AWS services. The company wants to track the expenses of each business unit and limit the spending to a pre-defined threshold. In addition, the solution should allow the security team to identify and respond to threats as quickly as possible for all the workloads across the business units. Also, workload accounts may need to be pulled off into a temporary holding area due to resource audit reasons.

Which of the following can be combined to build a solution for the given requirements? (Select three)

- Configure an AWS Budget alert to move an AWS account to Exceptions OU if the account reaches a predefined budget threshold. Use Service Control Policies (SCPs) to limit/block resource usage in the Exceptions OU. Configure a Suspended OU to hold workload accounts with retired resources. Use Service Control Policies (SCPs) to limit/block resource usage in the Suspended OU

(Correct)
- Configure GuardDuty in all member accounts within the AWS Organizations organization. Create an SNS topic in each account. Subscribe the security team to the topic so that the security team can receive alerts from GuardDuty via SNS

(Incorrect)
- Configure an AWS Cost Explorer alert to move an AWS account to Exceptions OU if the account reaches a predefined budget threshold. Use Service Control Policies (SCPs) to limit/block resource usage in the Exceptions OU. Configure a Suspended OU to hold workload accounts with retired resources. Use Service Control Policies (SCPs) to limit/block resource usage in the Suspended OU
- Designate an account within the AWS Organizations organization to be the GuardDuty delegated administrator. Create an SNS topic in this account. Subscribe the security team to the topic so that the security team can receive alerts from GuardDuty via SNS

(Correct)
- Use AWS Organizations to set up a multi-account environment. Organize the accounts into the following Service Control Policies (SCPs): Security, Infrastructure, Workloads, Suspended, and Exceptions. Grant necessary permissions to the accounts by using the SCP guardrails

(Incorrect)

-

Use AWS Organizations to set up a multi-account environment. Organize the accounts into the following Organizational Units (OUs): **Security, Infrastructure, Workloads, Suspended and Exceptions**

(Correct)

Explanation

Correct options:

Use AWS Organizations to set up a multi-account environment. Organize the accounts into the following Organizational Units (OUs): Security, Infrastructure, Workloads, Suspended and Exceptions - AWS categorizes the Security OU and the Infrastructure OU as foundational. The foundational OUs contain accounts, workloads, and other AWS resources that provide common security and infrastructure capabilities to secure and support your overall AWS environment.

The Suspended OU is used as a temporary holding area for accounts that are required to have their use suspended either temporarily or permanently.

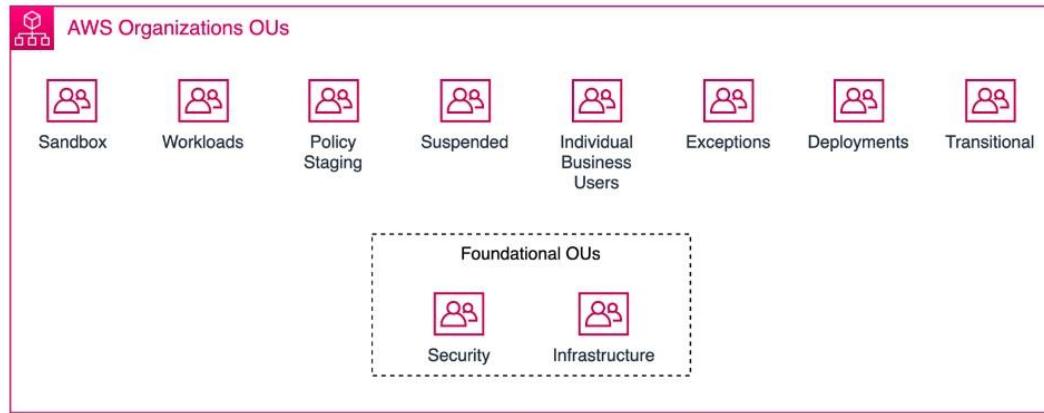
The Exceptions OU houses an account that requires an exception to the security policies that are applied to your Workloads OU.

AWS recommended OUs and

Recommended OUs and accounts

[PDF](#) | [RSS](#)

This section provides details on the recommended OUs and, in the case of the [Security OU and accounts](#), a set of recommended AWS accounts.



accounts:

via -

<https://docs.aws.amazon.com/whitepapers/latest/organizing-your-aws-environment/recommended-ous-and-accounts.html>

Configure an AWS Budget alert to move an AWS account to Exceptions OU if the account reaches a predefined budget threshold. Use Service Control Policies (SCPs) to limit/block resource usage in the Exceptions OU. Configure a Suspended OU to hold workload accounts with retired resources. Use Service Control Policies (SCPs) to limit/block resource usage in the Suspended OU - AWS Budgets provides the capability to configure cost-saving controls, or actions, that run either automatically on your behalf or by using a workflow approval process. You can use actions to define an explicit response that you want to take when a budget exceeds its action threshold. You can trigger these alerts on actual or forecasted cost and usage budgets.

For the given scenario, the management account can move the member account to restrictive OU (Exceptions OU) after the budget threshold for the member account is met.

Using AWS Budgets actions to move an AWS account to an
OU:

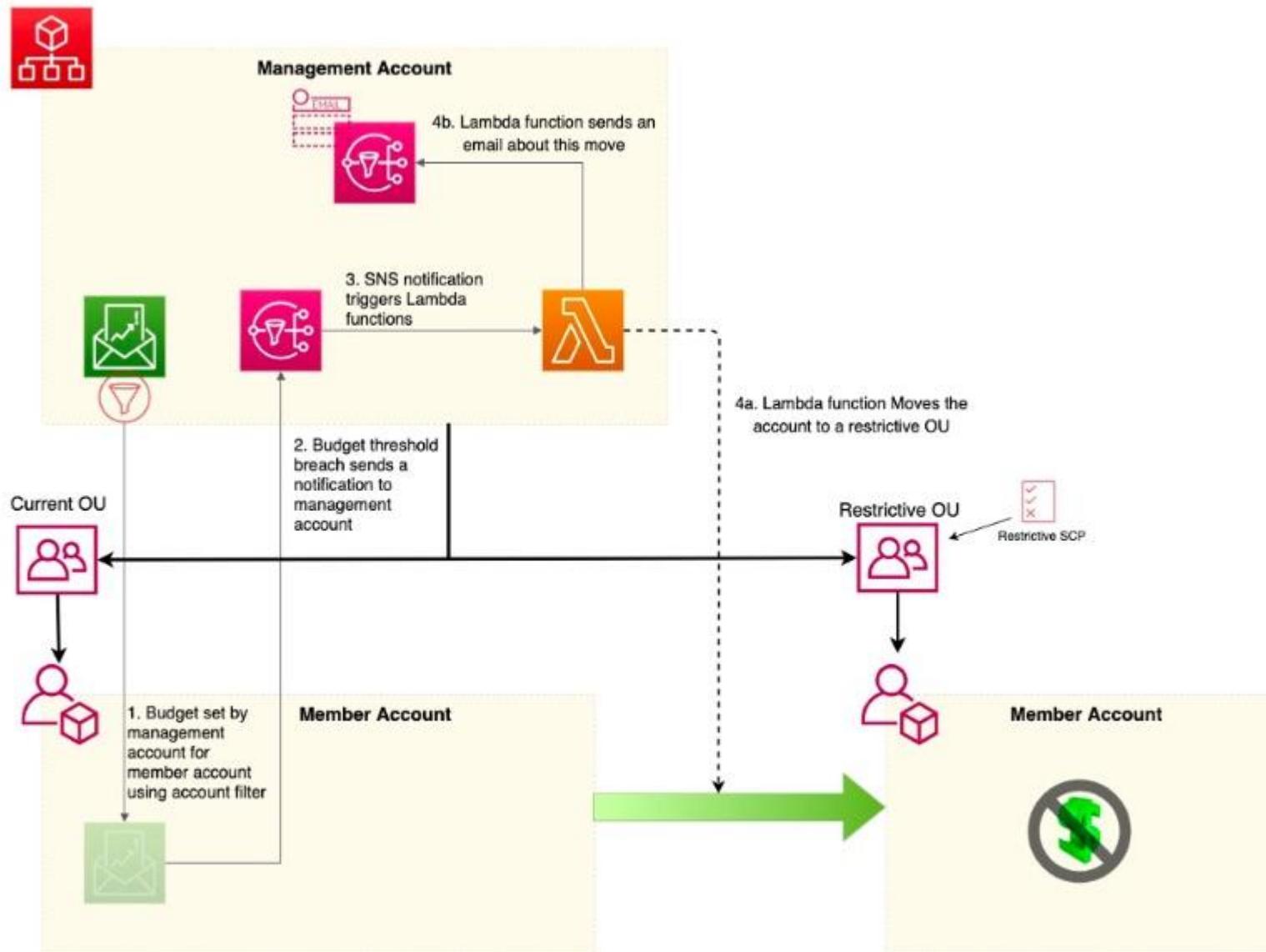


Figure 5: Using AWS Budgets actions to move an AWS account to an OU

via - <https://aws.amazon.com/blogs/mt/manage-cost-overruns-part-1/>

Designate an account within the AWS Organizations organization to be the GuardDuty delegated administrator. Create an SNS topic in this account. Subscribe the security team to the topic so that the security team can receive alerts from GuardDuty via SNS

When you use GuardDuty with an AWS Organizations organization, you can designate any account within the organization to be the GuardDuty delegated administrator. Only the organization management account can designate GuardDuty delegated administrators.

An account that is designated as a delegated administrator becomes a GuardDuty administrator account, has GuardDuty automatically enabled in the designated Region and is granted permission to enable and manage GuardDuty for all accounts in the organization within that Region. The other accounts in the organization can be viewed and added as GuardDuty member accounts associated with the delegated administrator account.

For the given use case, you can set up an SNS topic in this account and then subscribe the security team to the topic so that the security team can receive alerts from GuardDuty.

Incorrect options:

Use AWS Organizations to set up a multi-account environment. Organize the accounts into the following Service Control Policies (SCPs): Security, Infrastructure, Workloads, Suspended, and Exceptions. Grant necessary permissions to the accounts by using the SCP guardrails - Service control policies (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 alone are not sufficient to grant permissions to the accounts in your organization. No permissions are granted by an SCP. You cannot organize AWS accounts into Service Control Policies (SCPs). You organize AWS accounts into Organization Units by using AWS Organizations.

Configure an AWS Cost Explorer alert to move an AWS account to Exceptions OU if the account reaches a predefined budget threshold. Use Service Control Policies (SCPs) to limit/block resource usage in the Exceptions OU. Configure a Suspended OU to hold workload accounts with retired resources. Use Service Control Policies (SCPs) to limit/block resource usage in the Suspended OU - AWS Cost Explorer lets you explore your AWS costs and usage at both a high level and at a detailed level of

analysis, empowering you to dive deeper using several filtering dimensions (e.g., AWS Service, Region, Member Account, etc.). It is not possible to define actionable alerts in AWS Cost Explorer.

Configure GuardDuty in all member accounts within the AWS Organizations organization. Create an SNS topic in each account. Subscribe the security team to the topic so that the security team can receive alerts from GuardDuty via SNS - It is inefficient and cumbersome to configure GuardDuty in all member accounts within the AWS Organizations organization. It's better to centrally manage GuardDuty for all AWS accounts by using a GuardDuty delegated administrator account within the AWS Organizations organization.

References:

<https://docs.aws.amazon.com/whitepapers/latest/organizing-your-aws-environment/suspended-ou.html>

https://docs.aws.amazon.com/organizations/latest/userguide/orgs_manage_policies_scps.html

<https://docs.aws.amazon.com/whitepapers/latest/organizing-your-aws-environment/benefits-of-using-ous.html>

https://docs.aws.amazon.com/guardduty/latest/ug/guardduty_organizations.html

Question 23: **Incorrect**

The security team at a company has put forth a requirement to track the external IP address when a customer or a third party uploads files to the Amazon Simple Storage Service (Amazon S3) bucket owned by the company.

How will you track the external IP address used for each upload? (Select two)

-

Enable Amazon S3 server access logging to capture all bucket-level and object-level events

(Correct)

-

CloudWatch Logs centrally maintain the logs from all of your systems, applications, and AWS services that you use. Use these logs to capture the IP address at the object level for the S3 bucket

- **Enable VPC Flow Logs to capture all object-level events occurring on the S3 bucket**
(Incorrect)
- **Enable AWS Systems Manager Agent (SSM Agent) that writes information about executions, commands, scheduled actions on all AWS resources**
- **Enable AWS CloudTrail data events to enable object-level logging for S3 bucket**
(Correct)

Explanation

Correct options:

Enable Amazon S3 server access logging to capture all bucket-level and object-level events

Enable AWS CloudTrail data events to enable object-level logging for S3 bucket

To find the IP addresses for object-level requests to Amazon S3 (uploads and downloads), you must first enable one of the following logging methods:

1. Amazon S3 server access logging captures all bucket-level and object-level events. These logs use a format similar to Apache web server logs. After you enable server access logging, review the logs to find the IP addresses used with each upload to your bucket.
2. AWS CloudTrail data events capture the last 90 days of bucket-level events (for example, PutBucketPolicy and DeleteBucketPolicy), and you can enable object-level logging. These logs use a JSON format. After you enable object-level logging with data events, review the logs to find the IP addresses used with each upload to your bucket. It might take a few hours for AWS CloudTrail to start creating logs.

Incorrect options:

Enable VPC Flow Logs to capture all object-level events occurring on the S3 bucket - VPC Flow Logs is a feature that enables you to capture information about the IP traffic going to and from network interfaces in your VPC, so it does not apply to S3. Flow log data can be published to Amazon CloudWatch Logs or Amazon S3.

Enable AWS Systems Manager Agent (SSM Agent) that writes information about executions, commands, scheduled actions on all AWS resources - AWS Systems Manager Agent (SSM Agent) writes information about executions, commands, scheduled actions, errors, and health statuses to log files on each managed node. You can view log files by manually connecting to a managed node, or you can automatically send logs to Amazon CloudWatch Logs. AWS Systems Manager Agent (SSM Agent) is Amazon software that runs on Amazon Elastic Compute Cloud (Amazon EC2) instances, edge devices, and on-premises servers and virtual machines (VMs), so it does not apply to S3.

CloudWatch Logs centrally maintain the logs from all of your systems, applications, and AWS services that you use. Use these logs to capture the IP address at the object level for the S3 bucket - You can use Amazon CloudWatch Logs to monitor, store, and access your log files from Amazon Elastic Compute Cloud (Amazon EC2) instances, AWS CloudTrail, Route 53, and other sources. CloudWatch Logs enables you to see all of your logs, regardless of their source, as a single and consistent flow of events ordered by time, and you can query them and sort them based on other dimensions, group them by specific fields, create custom computations with a powerful query language, and visualize log data in dashboards. CloudWatch Logs cannot be used to track the external IP address used for uploads to S3.

Reference:

<https://aws.amazon.com/premiumsupport/knowledge-center/external-ip-address-s3-bucket/>

Question 24: **Correct**

A bioinformatics company leverages multiple open source tools to manage data analysis workflows running on its on-premises servers to process biological data which is generated and stored on a Network Attached Storage (NAS). The existing workflow receives around 100 GB of input biological data for each job run and individual jobs can take several hours to process the data. The CTO at the company wants to re-architect its proprietary analytics workflow on AWS to meet the workload demands and reduce the turnaround time from months to days. The company has provisioned a high-speed AWS Direct Connect connection. The final result needs to be stored in Amazon S3. The company is expecting approximately 20 job requests each day.

Which of the following options would you recommend for the given use case?



Leverage AWS Data Pipeline to transfer the biological data to Amazon S3. Use S3 events to trigger an AWS Step Functions workflow for orchestrating an AWS Batch job that processes the biological data



Leverage AWS Data Pipeline to transfer the biological data to Amazon S3. Use S3 events to trigger an Amazon EC2 Auto Scaling group to launch custom-AMI EC2 instances to process the biological data



Leverage AWS DataSync to transfer the biological data to Amazon S3. Use S3 events to trigger an AWS Lambda function that starts an AWS Step Functions workflow for orchestrating an AWS Batch job that processes the biological data

(Correct)



Leverage AWS Storage Gateway file gateway to transfer the biological data to Amazon S3. Use S3 events to trigger an AWS Lambda function that starts an AWS Step Functions workflow for orchestrating an AWS Batch job that processes the biological data

Explanation

Correct option:

Leverage AWS DataSync to transfer the biological data to Amazon S3. Use S3 events to trigger an AWS Lambda function that starts an AWS Step Functions workflow for orchestrating an AWS Batch job that processes the biological data

AWS DataSync is an online data movement and discovery service that simplifies and accelerates data migrations to AWS as well as moving data between on-premises storage, edge locations, other clouds, and AWS Storage. You can use DataSync to migrate active data to AWS, archive data to free up on-premises storage capacity, replicate data to AWS for business continuity, or transfer data to the cloud for analysis and processing.

For data transfer between on-premises and AWS Storage services, a single DataSync task is capable of fully utilizing a 10 Gbps network link. Since each workflow job consumes around 100GB of data and the company sees approximately 20 runs every day, DataSync can easily handle such active data transfer workloads. For the given use case, you can then configure an S3 event to trigger an AWS Lambda function that starts an AWS Step Functions workflow for orchestrating an AWS Batch job that processes the biological data.

When to choose AWS DataSync

Q: How is AWS DataSync different from using command line tools such as rsync or the Amazon S3 command line interface?

A: AWS DataSync fully automates and accelerates moving large active datasets to AWS. It is natively integrated with Amazon S3, Amazon EFS, Amazon FSx, [Amazon CloudWatch](#), and [AWS CloudTrail](#), which provides seamless and secure access to your storage services, as well as detailed monitoring of the transfer.

DataSync uses a purpose-built network protocol and scale-out architecture to transfer data. For data transfer between on-premises and AWS Storage services, a single DataSync task is capable of fully utilizing a 10 Gbps network link.

DataSync fully automates the data transfer. It comes with retry and network resiliency mechanisms, network optimizations, built-in task scheduling, monitoring via the DataSync API and Console, and CloudWatch metrics, events and logs that provide granular visibility into the transfer process. DataSync performs data integrity verification both during the transfer and at the end of the transfer.

DataSync provides end-to-end security, and integrates directly with AWS storage services. All data transferred between the source and destination is encrypted via TLS, and access to your AWS storage is enabled via built-in AWS security mechanisms such as IAM roles. DataSync with VPC endpoints are enabled to ensure that data transferred between an organization and AWS does not traverse the public internet, further increasing the security of data as it is copied over the network.

Q: To transfer objects between my buckets, when do I use AWS DataSync, when do I use S3 Replication, and when do I use S3 Batch Operations?

A: AWS provides multiple tools to copy objects between your buckets.

Use AWS DataSync for ongoing data distribution, data pipelines, and data lake ingest, as well as for consolidating or splitting data between multiple buckets.

Use [S3 Replication](#) for continuous replication of data to a specific destination bucket.

Use [S3 Batch Operations](#) for large-scale batch operations on S3 objects, such as to copy objects, set object tags or access control lists (ACLs), initiate object restores from Amazon S3 Glacier Flexible Retrieval (formerly S3 Glacier), invoke an AWS Lambda function to perform custom actions using your objects, manage S3 Object Lock legal hold, or manage S3 Object Lock retention dates.

via -

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

Q: When do I use AWS DataSync and when do I use AWS Snowball Edge?

A: AWS DataSync is ideal for online data transfers. You can use DataSync to migrate active data to AWS, transfer data to the cloud for analysis and processing, archive data to free up on-premises storage capacity, or replicate data to AWS for business continuity.

[AWS Snowball Edge](#) is ideal for offline data transfers, for customers who are bandwidth constrained, or transferring data from remote, disconnected, or austere environments.

Q: When do I use AWS DataSync and when do I use AWS Storage Gateway?

A: Use AWS DataSync to migrate existing data to Amazon S3, and subsequently use the File Gateway configuration of [AWS Storage Gateway](#) to retain access to the migrated data and for ongoing updates from your on-premises file-based applications.

You can use a combination of DataSync and File Gateway to minimize your on-premises infrastructure while seamlessly connecting on-premises applications to your cloud storage. AWS DataSync enables you to automate and accelerate online data transfers to AWS Storage services. After the initial data transfer phase using AWS DataSync, File Gateway provides your on-premises applications with low latency access to the migrated data. When using DataSync with NFS shares, POSIX metadata from your source on-premises storage is preserved, and permissions from the source storage apply when accessing your files using File Gateway.

Q: When do I use AWS DataSync, and when do I use Amazon S3 Transfer Acceleration?

A: If your applications are already integrated with the Amazon S3 API, and you want higher throughput for transferring large files to S3, you can use [S3 Transfer Acceleration](#). If you want to transfer data from existing storage systems (e.g., Network Attached Storage), or from instruments that cannot be changed (e.g., DNA sequencers, video cameras), or if you want multiple destinations, you use AWS DataSync. DataSync also automates and simplifies the data transfer by providing additional functionality, such as built-in retry and network resiliency mechanisms, data integrity verification, and flexible configuration to suit your specific needs, including bandwidth throttling, etc.

Q: When do I use AWS DataSync and when do I use AWS Transfer Family?

via - <https://aws.amazon.com/datasync/faqs/>

Incorrect options:

Leverage AWS Data Pipeline to transfer the biological data to Amazon S3. Use S3 events to trigger an AWS Step Functions workflow for orchestrating an AWS Batch job that processes the biological data - You cannot trigger an AWS Step Function directly from an S3 event, so this option is incorrect.

Leverage AWS Data Pipeline to transfer the biological data to Amazon S3. Use S3 events to trigger an Amazon EC2 Auto Scaling group to launch custom-AMI EC2 instances to process the biological data - You cannot trigger an Amazon EC2 Auto Scaling group directly from an S3 event, so this option is incorrect.

Leverage AWS Storage Gateway file gateway to transfer the biological data to Amazon S3. Use S3 events to trigger an AWS Lambda function that starts an AWS Step Functions workflow for orchestrating an AWS Batch job that processes the biological data - You should use AWS DataSync to migrate existing or active data to Amazon S3 and use the File Gateway configuration of AWS Storage Gateway to retain access to the migrated data and for ongoing updates from your on-premises file-based applications. Since the data processing workflow/application is being migrated from on-premises to AWS Cloud, you no longer have any on-premises applications that need to access the processed data from AWS Cloud. So this option is incorrect.

References:

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

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

Question 25: **Correct**

A team has recently created a secret using AWS Secrets Manager to access their private Amazon Relational Database Service (Amazon RDS) instance. When the team tried to rotate the AWS Secrets Manager secret in an Amazon Virtual Private Cloud (Amazon VPC), the operation failed. On analyzing the Amazon CloudWatch Logs, the team realized that the AWS Lambda task timed out.

Which of the following solutions needs to be implemented for rotating the secret successfully?

- Configure an Amazon VPC interface endpoint for the Lambda service to enable access for your Secrets Manager Lambda rotation function and private Amazon Relational Database Service (Amazon RDS) instance
 - Configure an Amazon VPC interface endpoint for the Secrets Manager service to enable access for your Secrets Manager Lambda rotation function and private Amazon Relational Database Service (Amazon RDS) instance
- (Correct)
- Interface VPC endpoints support traffic only over HTTP. If this is incorrectly configured, the AWS Lambda function can timeout
 - Your Lambda rotation function might be based on an older template that doesn't support SSL/TLS. To support connections that use SSL/TLS, you must recreate your Lambda rotation function

Explanation

Correct option:

Configure an Amazon VPC interface endpoint to access your Secrets Manager Lambda rotation function and private Amazon Relational Database Service (Amazon RDS) instance

Secrets Manager can't rotate secrets for AWS services running in Amazon VPC private subnets because these subnets don't have internet access. To rotate the keys successfully you need to configure an Amazon VPC interface endpoint to access your Secrets Manager Lambda function and private Amazon Relational Database Service (Amazon RDS) instance.

Steps that need to be followed: 1. Create security groups for the Secrets Manager VPC endpoint, Amazon RDS instance, and the Lambda rotation function 2. Add rules to Amazon VPC endpoint and Amazon RDS instance security groups 3. Attach security groups

to AWS resources 4. Create an Amazon VPC interface endpoint for the Secrets Manager service and associate it with a security group
5. Verify that the Secrets Manager can rotate the secret

Incorrect options:

Configure an Amazon VPC interface endpoint for the Lambda service to enable access for your Secrets Manager Lambda rotation function and private Amazon Relational Database Service (Amazon RDS) instance - As explained above, you need to create an Amazon VPC interface endpoint for the Secrets Manager and not for the Lambda service. This option has been added as a distractor.

Interface VPC endpoints support traffic only over HTTP. If this is incorrectly configured, the AWS Lambda function can timeout - This statement is incorrect. Interface VPC endpoints support traffic only over TCP.

Your Lambda rotation function might be based on an older template that doesn't support SSL/TLS. To support connections that use SSL/TLS, you must recreate your Lambda rotation function - Rotation functions for Amazon RDS (except Amazon RDS for Oracle) and Amazon DocumentDB automatically use SSL/TLS to connect to your database if it's available. If you set up secret rotation before December 20, 2021, then your rotation function might be based on an older template that doesn't support SSL/TLS. To support connections that use SSL/TLS, you must recreate your rotation function. If this is the issue then the following error crops up ": setSecret: Unable to log into the database with previous, current, or the pending secret of secret".

References:

<https://aws.amazon.com/premiumsupport/knowledge-center/rotate-secrets-manager-secret-vpc/>

<https://docs.aws.amazon.com/vpc/latest/privatelink/create-interface-endpoint.html#vpce-interface-limitations>

<https://aws.amazon.com/premiumsupport/knowledge-center/rotate-secret-db-ssl/>

Question 26: **Correct**

A social learning platform allows students to connect with other students as well as experts and professionals from academic, research institutes and industry. The engineering team at the company manages 5 Amazon EC2 instances that make read-heavy

database requests to the Amazon RDS for PostgreSQL DB cluster. As an AWS Certified Solutions Architect Professional, you have been asked to make the database cluster resilient from a disaster recovery perspective.

Which of the following features will help you prepare for database disaster recovery? (Select two)

- Use database cloning feature of the RDS DB cluster
- Use RAID 1 configuration for the RDS DB cluster
- Use RDS Provisioned IOPS (SSD) Storage in place of General Purpose (SSD) Storage
- Use cross-Region Read Replicas
(Correct)
- Enable the automated backup feature of Amazon RDS in a multi-AZ deployment that creates backups in a single or multiple AWS Region(s)
(Correct)

Explanation
Correct option:

Use cross-Region Read Replicas

In addition to using Read Replicas to reduce the load on your source DB instance, you can also use Read Replicas to implement a DR solution for your production DB environment. If the source DB instance fails, you can promote your Read Replica to a standalone source server. Read Replicas can also be created in a different Region than the source database. Using a cross-Region Read Replica can help ensure that you get back up and running if you experience a regional availability issue.

Enable the automated backup feature of Amazon RDS in a multi-AZ deployment that creates backups in a single or multiple AWS Region(s)

Amazon RDS provides high availability and failover support for DB instances using Multi-AZ deployments. Amazon RDS uses several different technologies to provide failover support. Multi-AZ deployments for MariaDB, MySQL, Oracle, and PostgreSQL DB instances use Amazon's failover technology.

The automated backup feature of Amazon RDS enables point-in-time recovery for your database instance. Amazon RDS will backup your database and transaction logs and store both for a user-specified retention period. If it's a Multi-AZ configuration, backups occur on the standby to reduce I/O impact on the primary. Amazon RDS supports single Region or cross-Region automated backups.

Incorrect options:

Use RAID 1 configuration for the RDS DB cluster - This option has been added as a distractor. RAID configuration options can only be used for EC2 instance-hosted databases. By using EBS storage volumes with EC2 instances, you can configure volumes with any RAID levels. For example, for greater I/O performance, you can opt for RAID 0, which can stripe multiple volumes together. RAID 1 can be used for data redundancy because it mirrors two volumes together.

Use RDS Provisioned IOPS (SSD) Storage in place of General Purpose (SSD) Storage - Amazon RDS Provisioned IOPS Storage is an SSD-backed storage option designed to deliver fast, predictable, and consistent I/O performance. This storage type enhances the performance of the RDS database, but this isn't a disaster recovery option.

Use database cloning feature of the RDS DB cluster - This option has been added as a distractor. Database cloning is only available for Aurora and not for RDS.

References:

<https://aws.amazon.com/rds/features/>

<https://aws.amazon.com/blogs/database/implementing-a-disaster-recovery-strategy-with-amazon-rds/>

<https://aws.amazon.com/about-aws/whats-new/2021/07/amazon-rds-cross-region-automated-backups-regional-expansion/>

<https://aws.amazon.com/blogs/database/best-storage-practices-for-running-production-workloads-on-hosted-databases-with-amazon-rds-or-amazon-ec2/>

Question 27: **Incorrect**

A company uses Amazon FSx for Windows File Server with deployment type of Single-AZ 2 as its file storage service for its non-core functions. With a change in the company's policy that mandates high availability of data for all its functions, the company needs to change the existing configuration. The company also needs to monitor the file system activity as well as the end-user actions on the Amazon FSx file server.

Which solutions will you combine to implement these requirements? (Select two)

-

Configure a new Amazon FSx for Windows file system with a deployment type of Multi-AZ. Transfer data to the newly created file system using the AWS DataSync service. Point all the file system users to the new location. You can test the failover of your Multi-AZ file system by modifying its throughput capacity

(Correct)

-

Configure a new Amazon FSx for Windows file system with a deployment type of Multi-AZ. Transfer data to the newly created file system using the AWS DataSync service. Point all the file system users to the new location. You can test the failover of your Multi-AZ file system by modifying the elastic network interfaces associated with your file system

(Incorrect)

-

You can monitor the file system activity using AWS CloudTrail and monitor end-user actions with file access auditing using Amazon CloudWatch Logs

(Incorrect)

-

You can monitor storage capacity and file system activity using Amazon CloudWatch, and monitor end-user actions with file access auditing using Amazon CloudWatch Logs and Amazon Kinesis Data Firehose

(Correct)

-

Configure a new Amazon FSx for Windows file system with a deployment type of Single-AZ 1. Transfer data to the newly created file system using the AWS DataSync service. Point all the file system users to the new location

Explanation

Correct options:

Configure a new Amazon FSx for Windows file system with a deployment type of Multi-AZ. Transfer data to the newly created file system using the AWS DataSync service. Point all the file system users to the new location. You can test the failover of your Multi-AZ file system by modifying its throughput capacity

In a Multi-AZ deployment, Amazon FSx automatically provisions and maintains a standby file server in a different Availability Zone. Any changes written to disk in your file system are synchronously replicated across Availability Zones to the standby. If there is planned file system maintenance or unplanned service disruption, Amazon FSx automatically fails over to the secondary file server, allowing you to continue accessing your data without manual intervention. Multi-AZ file systems are recommended for most production workloads that require high availability of shared Windows file data.

To migrate your existing files to FSx for Windows File Server file systems, AWS recommends using AWS DataSync, an online data transfer service designed to simplify, automate, and accelerate copying large amounts of data to and from AWS storage services. DataSync copies data over the internet or AWS Direct Connect.

You can test the failover of your Multi-AZ file system by modifying its throughput capacity. When you modify your file system's throughput capacity, Amazon FSx switches out the file system's file server. Multi-AZ file systems automatically fail over to the secondary server while Amazon FSx replaces the preferred server file server first. Then the file system automatically fails back to the new primary server and Amazon FSx replaces the secondary file server.

Choosing Single-AZ or Multi-AZ file system deployment

With Single-AZ file systems, Amazon FSx automatically replicates your data within an Availability Zone (AZ) to protect it from component failure. It continuously monitors for hardware failures and automatically replaces infrastructure components in the event of a failure. *Single-AZ 2* is the latest generation of Single-AZ file systems, and it supports both SSD and HDD storage. *Single-AZ 1* file systems support SSD storage, Microsoft Distributed File System Replication (DFSR), and the use of custom DNS names. Single-AZ file systems will experience unavailability during file system maintenance, infrastructure component replacement, and when an Availability Zone is unavailable.

Multi-AZ file systems support all the availability and durability features of Single-AZ file systems. In addition, they are designed to provide continuous availability to data, even during file system maintenance, infrastructure component replacement, and when an Availability Zone is unavailable. In a Multi-AZ deployment, Amazon FSx automatically provisions and maintains a standby file server in a different Availability Zone. Any changes written to disk in your file system are synchronously replicated across Availability Zones to the standby. If there is planned file system maintenance or unplanned service disruption, Amazon FSx automatically fails over to the secondary file server, allowing you to continue accessing your data without manual intervention.

Multi-AZ file systems are recommended for most production workloads that require high availability to shared Windows file data. Single-AZ file systems offer a lower price point for workloads that don't require the high availability of a Multi-AZ solution and that can recover from the most recent file system backup if data is lost. Amazon FSx takes automatic daily backups of all file systems by default.

Feature support by deployment types

The following table summarizes features supported by the FSx for Windows File Server file system deployment types:

Deployment type	SSD storage	HDD storage	DFS namespaces	DFS replication	Custom DNS names	CA shares
Single-AZ 1	✓		✓	✓	✓	
Single-AZ 2	✓	✓	✓		✓	✓*
Multi-AZ	✓	✓	✓		✓	✓*

via -

<https://docs.aws.amazon.com/fsx/latest/WindowsGuide/high-availability-multiAZ.html>

You can monitor storage capacity and file system activity using Amazon CloudWatch, and monitor end-user actions with file access auditing using Amazon CloudWatch Logs and Amazon Kinesis Data Firehose

You can monitor storage capacity and file system activity using Amazon CloudWatch. You can monitor end-user actions with file access auditing at any time (during or after the creation of a file system) via the AWS Management Console or the Amazon FSx CLI or API. You can also change the destination for publishing user access events by logging these events to CloudWatch Logs or streaming to Kinesis Data Firehose.

Incorrect options:

Configure a new Amazon FSx for Windows file system with a deployment type of Single-AZ 1. Transfer data to the newly created file system using the AWS DataSync service. Point all the file system users to the new location - With Single-AZ file systems, Amazon FSx automatically replicates your data within an Availability Zone (AZ) to protect it from component failure. It continuously monitors for hardware failures and automatically replaces infrastructure components in the event of a failure. Single-AZ 2 is the latest generation of Single-AZ file systems, and it supports both SSD and HDD storage. Single-AZ 1 file systems support SSD storage, Microsoft Distributed File System Replication (DFSR), and the use of custom DNS names. Single-AZ file systems will experience unavailability during file system maintenance, infrastructure component replacement, and when an Availability Zone is unavailable.

Multi-AZ file systems are recommended for most production workloads that require high availability of shared Windows file data. Single-AZ file systems offer a lower price point for workloads that don't require the high availability of a Multi-AZ solution and that can recover from the most recent file system backup if data is lost.

You can monitor the file system activity using AWS CloudTrail and monitor end-user actions with file access auditing using Amazon CloudWatch Logs - This statement is incorrect. You can monitor storage capacity and file system activity using Amazon CloudWatch, monitor all Amazon FSx API calls using AWS CloudTrail, and monitor end-user actions with file access auditing using Amazon CloudWatch Logs and Amazon Kinesis Data Firehose.

Configure a new Amazon FSx for Windows file system with a deployment type of Multi-AZ. Transfer data to the newly created file system using the AWS DataSync service. Point all the file system users to the new location. You can test the failover of your Multi-AZ file system by modifying the elastic network interfaces associated with your file system - You must not modify or delete the elastic network interfaces associated with your file system. Modifying or deleting the network interface can cause a permanent loss of connection between your VPC and your file system. Therefore this option is incorrect.

References:

<https://docs.aws.amazon.com/fsx/latest/WindowsGuide/high-availability-multiAZ.html>

<https://docs.aws.amazon.com/fsx/latest/WindowsGuide/migrate-files-fsx.html>

<https://aws.amazon.com/fsx/windows/faqs/>

<https://docs.aws.amazon.com/fsx/latest/WindowsGuide/limit-access-security-groups.html>

<https://aws.amazon.com/blogs/aws/file-access-auditing-is-now-available-for-amazon-fsx-for-windows-file-server/>

Question 28: **Incorrect**

Recently, an Amazon CloudFront distribution has been configured with an Amazon S3 bucket as the origin. However, users are getting an HTTP 307 Temporary Redirect response from Amazon S3.

What could be the reason for this behavior and how will you resolve the issue? (Select two)

- Configure CloudFront **Cache-Control** and **Expires** headers to a value of zero, to fetch new objects immediately from the S3 bucket

- Enable Amazon S3 Transfer Acceleration to help CloudFront access data faster over long distances from the S3 bucket

- CloudFront by default, forwards the requests to the default S3 endpoint. Change the origin domain name of the distribution to include the Regional endpoint of the bucket

(Correct)

- Enable Cross-Region replication for the S3 bucket so that CloudFront can retrieve the data immediately after the creation of the bucket

(Incorrect)

-

When a new Amazon S3 bucket is created, it takes up to 24 hours before the bucket name propagates across all AWS Regions

(Correct)

Explanation

Correct options:

When a new Amazon S3 bucket is created, it takes up to 24 hours before the bucket name propagates across all AWS Regions

CloudFront by default, forwards the requests to the default S3 endpoint. Change the origin domain name of the distribution to include the Regional endpoint of the bucket

After you create an Amazon S3 bucket, up to 24 hours can pass before the bucket name propagates across all AWS Regions. During this time, you might receive the 307 Temporary Redirect response for requests to Regional endpoints that aren't in the same Region as your bucket.

To avoid the 307 Temporary Redirect response, send requests only to the Regional endpoint in the same Region as your S3 bucket. If you're using an Amazon CloudFront distribution with an Amazon S3 origin, CloudFront forwards requests to the default S3 endpoint (s3.amazonaws.com). The default S3 endpoint is in the us-east-1 Region. If you must access Amazon S3 within the first 24 hours of creating the bucket, you can change the origin domain name of the distribution. The domain name must include the Regional endpoint of the bucket. For example, if the bucket is in us-west-2, you can change the origin domain name from awsexamplebucketname.s3.amazonaws.com to awsexamplebucket.s3.us-west-2.amazonaws.com.

Incorrect options:

Enable Cross-Region replication for the S3 bucket so that CloudFront can retrieve the data immediately after the creation of the bucket - S3 Cross-Region Replication (CRR) is used to copy objects across Amazon S3 buckets in different AWS Regions. CRR can help you do the following - meet compliance requirements, minimize latency and increase operational efficiency. CRR however, cannot resolve the HTTP 307 error.

Configure CloudFront Cache-Control and Expires headers to a value of zero, to fetch new objects immediately from the S3 bucket - You can use the Cache-Control and Expires headers to control how long objects stay in the CloudFront cache. This option has been added as a distractor and is unrelated to the HTTP 307 error.

Enable Amazon S3 Transfer Acceleration to help CloudFront access data faster over long distances from the S3 bucket - Amazon S3 Transfer Acceleration is a bucket-level feature that enables fast, easy, and secure transfers of files over long distances between your client and an S3 bucket. Transfer Acceleration is designed to optimize transfer speeds from across the world into S3 buckets. This option has been added as a distractor and is unrelated to the HTTP 307 error.

Reference:

<https://aws.amazon.com/premiumsupport/knowledge-center/s3-http-307-response/>

Question 29: **Correct**

A social media company has VPC Flow Logs enabled for its NAT gateway. The security team is seeing Action = ACCEPT for inbound traffic that comes from the public IP address 198.21.200.1 destined for a private EC2 instance. The team must determine whether the traffic represents unsolicited inbound connections from the internet. The first two octets of the VPC CIDR block are 205.1.

Which of the following options can address this requirement?

-

Inspect the VPC Flow Logs using the CloudTrail console and select the log group that contains the NAT gateway's ENI and the EC2 instance's ENI. Leverage a query filter with the destination address set as like 205.1 and the source address set as like 198.21.200.1. Execute the stats command to filter the sum of bytes transferred by the source address and the destination address

-

Inspect the VPC Flow Logs using the CloudTrail console and select the log group that contains the NAT gateway's ENI and the EC2 instance's ENI. Leverage a query filter with the source address set as like 205.1 and the destination address set as like 198.21.200.1. Execute the stats command to filter the sum of bytes transferred by the source address and the destination address

-

Inspect the VPC Flow Logs using the CloudWatch console and select the log group that contains the NAT gateway's ENI and the EC2 instance's ENI. Leverage a query filter with the destination address set as **like 205.1** and the source address set as **like 198.21.200.1**. Execute the stats command to filter the sum of bytes transferred by the source address and the destination address

(Correct)

-

Inspect the VPC Flow Logs using the CloudWatch console and select the log group that contains the NAT gateway's ENI and the EC2 instance's ENI. Leverage a query filter with the source address set as **like 205.1** and the destination address set as **like 198.21.200.1**. Execute the stats command to filter the sum of bytes transferred by the source address and the destination address

Explanation

Correct option:

Inspect the VPC Flow Logs using the CloudWatch console and select the log group that contains the NAT gateway's ENI and the EC2 instance's ENI. Leverage a query filter with the destination address set as **like 205.1 and the source address set as **like 198.21.200.1**. Execute the stats command to filter the sum of bytes transferred by the source address and the destination address**

NAT gateways managed by AWS don't accept traffic initiated from the internet. However, there are two reasons why the information in your VPC flow logs might appear to indicate that inbound traffic is being accepted from the internet.

1: Inbound internet traffic is permitted by your security group or network access control lists (ACL)

VPC flow logs show inbound internet traffic as accepted if the traffic is permitted by your security group or network ACLs. If network ACLs attached to a NAT gateway don't explicitly deny traffic from the internet, then the traffic to the NAT gateway appears accepted. However, the traffic isn't actually accepted by the NAT gateway and is dropped. You can use just the first two octets in the search filter to analyze all network interfaces in the VPC, like so:

```
filter (dstAddr like 'xxx.xxx' and srcAddr like 'public IP')
| stats sum(bytes) as bytesTransferred by srcAddr, dstAddr
```

```
| limit 10
```

If the query results show traffic on the NAT gateway private IP from the public IP, but not traffic on other private IPs in the VPC. These results confirm that the incoming traffic was unsolicited.

I'm seeing inbound traffic in the VPC flow logs for my public NAT gateway. Is my public NAT gateway accepting inbound traffic from the internet?

Last updated: 2022-10-31

My Virtual Private Cloud (VPC) flow logs show Action = ACCEPT for inbound traffic coming from public IP addresses. However, my understanding of network address translation (NAT) gateways was that they don't accept traffic from the internet. Is my NAT gateway accepting inbound traffic from the internet?

Resolution

NAT gateways managed by AWS don't accept traffic initiated from the internet. However, there are two reasons why information in your VPC flow logs might appear to indicate that inbound traffic is being accepted from the internet.

Reason #1: Inbound internet traffic is permitted by your security group or network access control lists (ACL)

VPC flow logs show inbound internet traffic as accepted if the traffic is permitted by your security group or network ACLs. If network ACLs attached to a NAT gateway don't explicitly deny traffic from the internet, then the traffic to the NAT gateway appears accepted. However, the traffic isn't actually accepted by the NAT gateway and is dropped.

To confirm this, do the following:

1. Open the [Amazon CloudWatch console](#).
2. In the navigation pane, choose **Insights**.
3. From the dropdown, select the log group that contains the NAT gateway's elastic network interface and the private instance's elastic network interface.
4. Run the query below.

```
filter (dstAddr like 'xxx.xxx' and srcAddr like 'public IP')
| stats sum(bytes) as bytesTransferred by srcAddr, dstAddr
| limit 10
```

Note: You can use just the first two octets in the search filter to analyze all network interfaces in the VPC. In the example above, replace **xxx.xxx** with the first two octets of your VPC classless inter-domain routing (CIDR). Also, replace **public IP** with the public IP that you're seeing in the VPC flow log entry.

Query results show traffic on the NAT gateway private IP from the public IP, but not traffic on other private IPs in the VPC. These results confirm that the incoming traffic was unsolicited. However, if you do see traffic on the private instance's IP, then follow the steps under Reason #2.

via

- <https://aws.amazon.com/premiumsupport/knowledge-center/vpc-analyze-inbound-traffic-nat-gateway/>

Reason #2: Traffic to the public IP was initiated from a private instance

If there are private instances using NAT gateway for internet access, then your VPC flow log includes the response traffic from the public IP address. To confirm that traffic to the public IP was initiated from your private instance, run the query below:

Note: Before running the query, do the following:

- Select the time frame that corresponds with the time frame when you observed traffic in the VPC flow logs.
- If you have multiple log groups in your VPC, then select the appropriate one.

```
filter (dstAddr like 'public IP' and srcAddr like 'xxx.xxx') or (srcAddr like 'public IP' and dstAddr like 'xxx.xxx')
| limit 10
```

Be sure to replace **xxx.xxx** with the first two octets of your VPC CIDR. Replace **public IP** with the public IP you're seeing in the VPC flow log entry. Increase the **limit** if more than 10 resources in your VPC have initiated traffic to the public IP.

The query results show bi-directional traffic between the private instance and public IP addresses. To determine whether the private instance or external public IP address is the initiator, see the following example:

```
2022-09-28T12:05:18.000+10:00 eni-023466675b6xxxxxx 10.0.101.222 8.8.8.8 53218 53 6(17) 4 221 1664330718 1664330746 ACCEPT OK
2022-09-28T12:05:18.000+10:00 eni-023466675b6xxxxxx 8.8.8.8 10.0.101.222 53 53218 6(17) 4 216 1664330718 1664330746 ACCEPT OK
```

In this TCP/UDP traffic (protocol ID = 6 or 17) example, the private IP address **10.0.101.222** with the source port **53218** (ephemeral port) is the initiator. The IP address **8.8.8.8** with destination port **53** is the receiver and responder.

Note: It's a best practice to turn on the [TCP flag field](#) for your VPC flow log.

For example, the following entries have TCP flag field in the last column:

```
2022-09-28T12:05:52.000+10:00 eni-023466675b6xxxxxx 10.0.1.231 8.8.8.8 50691 53 6(17) 3 4 221 1664330752 1664330776 ACCEPT OK 2
2022-09-28T12:05:21.000+10:00 eni-023466675b6xxxxxx 8.8.8.8 10.0.1.231 53 50691 6(17) 19 4 216 1664330721 1664330742 ACCEPT OK 18
```

The private IP address **10.0.101.222** is the initiator with TCP flag **2**, which represents a TCP SYN packet.

In the following ICMP protocol example, there isn't enough information to determine which side is the initiator because there is no port information or TCP flag:

```
2022-09-27T17:57:39.000+10:00 eni-023466675b6xxxxxx 10.0.1.231 8.8.8.8 0 0 1 17 1428 1664265459 1664265483 ACCEPT OK
2022-09-27T17:57:39.000+10:00 eni-023466675b6xxxxxx 8.8.8.8 10.0.1.231 0 0 1 0 17 1428 1664265459 1664265483 ACCEPT OK
```

via - <https://aws.amazon.com/premiumsupport/knowledge-center/vpc-analyze-inbound-traffic-nat-gateway/>

Incorrect options:

Inspect the VPC Flow Logs using the CloudTrail console and select the log group that contains the NAT gateway's ENI and the EC2 instance's ENI. Leverage a query filter with the destination address set as like 205.1 and the source address set as like 198.21.200.1. Execute the stats command to filter the sum of bytes transferred by the source address and the destination address

Inspect the VPC Flow Logs using the CloudTrail console and select the log group that contains the NAT gateway's ENI and the EC2 instance's ENI. Leverage a query filter with the source address set as like 205.1 and the destination address set as like 198.21.200.1. Execute the stats command to filter the sum of bytes transferred by the source address and the destination address

You cannot use the CloudTrail console to analyze the VPC Flow Logs. You need to use the CloudWatch console for this analysis. Therefore both these options are incorrect.

Inspect the VPC Flow Logs using the CloudWatch console and select the log group that contains the NAT gateway's ENI and the EC2 instance's ENI. Leverage a query filter with the source address set as like 205.1 and the destination address set as like 198.21.200.1. Execute the stats command to filter the sum of bytes transferred by the source address and the destination address - For the query filter, you need to use the destination address set as like 205.1 with the source address set as like 198.21.200.1 and NOT vice-versa.

References:

<https://docs.aws.amazon.com/vpc/latest/userguide/flow-logs.html>

<https://aws.amazon.com/premiumsupport/knowledge-center/vpc-analyze-inbound-traffic-nat-gateway/>

Question 30: **Correct**

A financial services company had a security incident recently and wants to review the security of its two-tier server architecture. The company wants to ensure that it follows the principle of least privilege while configuring the security groups for access between the

EC2 instance-based app servers and RDS MySQL database servers. The security group for the EC2 instances as well as the security group for the MySQL database servers has no inbound and outbound rules configured currently.

As an AWS Certified Solutions Architect Professional, which of the following options would you recommend to adhere to the given requirements? (Select two)



Create an outbound rule in the security group for the MySQL DB servers using TCP protocol on port 3306. Set the destination as the security group for the EC2 instance app servers



Create an outbound rule in the security group for the EC2 instance app servers using TCP protocol on port 3306. Set the destination as the security group for the MySQL DB servers

(Correct)



Create an outbound rule in the security group for the EC2 instance app servers using TCP protocol on the ephemeral port range. Set the destination as the security group for the MySQL DB servers



Create an outbound rule in the security group for the MySQL DB servers using TCP protocol on the ephemeral port range. Set the destination as the security group for the EC2 instance app servers



Create an inbound rule in the security group for the MySQL DB servers using TCP protocol on port 3306. Set the source as the security group for the EC2 instance app servers

(Correct)

Explanation

Correct options:

Create an outbound rule in the security group for the EC2 instance app servers using TCP protocol on port 3306. Set the destination as the security group for the MySQL DB servers

Create an inbound rule in the security group for the MySQL DB servers using TCP protocol on port 3306. Set the source as the security group for the EC2 instance app servers

A security group controls the traffic that is allowed to reach and leave the resources that it is associated with. For example, after you associate a security group with an EC2 instance, it controls the inbound and outbound traffic for the instance. Security groups are stateful. For example, if you send a request from an instance, the response traffic for that request is allowed to reach the instance regardless of the inbound security group rules. Responses to allowed inbound traffic are allowed to leave the instance, regardless of the outbound rules.

When you first create a security group, it has no inbound rules. Therefore, no inbound traffic is allowed until you add inbound rules to the security group. When you first create a security group, it has an outbound rule that allows all outbound traffic from the resource. You can remove the rule and add outbound rules that allow specific outbound traffic only. If your security group has no outbound rules, no outbound traffic is allowed.

Security group basics

Characteristics of security groups

- When you create a security group, you must provide it with a name and a description. The following rules apply:
 - A security group name must be unique within the VPC.
 - Names and descriptions can be up to 255 characters in length.
 - Names and descriptions are limited to the following characters: a-z, A-Z, 0-9, spaces, and ._-:/()#@[]+=;&:;"\$*.
 - When the name contains trailing spaces, we trim the space at the end of the name. For example, if you enter "Test Security Group " for the name, we store it as "Test Security Group".
 - A security group name cannot start with sg-.
- Security groups are stateful.** For example, if you send a request from an instance, the response traffic for that request is allowed to reach the instance regardless of the inbound security group rules. Responses to allowed inbound traffic are allowed to leave the instance, regardless of the outbound rules.
- There are quotas on the number of security groups that you can create per VPC, the number of rules that you can add to each security group, and the number of security groups that you can associate with a network interface. For more information, see [Amazon VPC quotas](#).

Characteristics of security group rules

- You can specify allow rules, but not deny rules.
- When you first create a security group, it has no inbound rules. Therefore, no inbound traffic is allowed until you add inbound rules to the security group.**
- When you first create a security group, it has an outbound rule that allows all outbound traffic from the resource. You can remove the rule and add outbound rules that allow specific outbound traffic only. If your security group has no outbound rules, no outbound traffic is allowed.**
- When you associate multiple security groups with a resource, the rules from each security group are aggregated to form a single set of rules that are used to determine whether to allow access.
- When you add, update, or remove rules, your changes are automatically applied to all resources associated with the security group. The effect of some rule changes can depend on how the traffic is tracked. For more information, see [Connection tracking](#) in the *Amazon EC2 User Guide for Linux Instances*.
- When you create a security group rule, AWS assigns a unique ID to the rule. You can use the ID of a rule when you use the API or CLI to modify or delete the rule.

Best practices

- Authorize only specific IAM principals to create and modify security groups.
- Create the minimum number of security groups that you need, to decrease the risk of error. Use each security group to manage access to resources that have similar functions and security requirements.
- When you add inbound rules for ports 22 (SSH) or 3389 (RDP) so that you can access your EC2 instances, authorize only specific IP address ranges. If you specify 0.0.0.0/0 (IPv4) and ::/ (IPv6), this enables anyone to access your instances from any IP address using the specified protocol.
- Do not open large port ranges. Ensure that access through each port is restricted to the sources or destinations that require it.
- Consider creating network ACLs with rules similar to your security groups, to add an additional layer of security to your VPC. For more information about the differences between security groups and network ACLs, see [Compare security groups and network ACLs](#).

via -

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

For the given use case, you need to set up an outbound rule in the security group for the EC2 instance app servers using TCP protocol on port 3306 and then select the destination as the security group for the MySQL DB servers. Further, you need to set up an inbound rule in the security group for the MySQL DB servers using TCP protocol on port 3306 and then select the source as the security group for the EC2 instance app servers. This combination would let the request be initiated from the EC2 instances and allowed into the DB servers. Since the security groups are stateful, the response from the DB servers would be allowed out of the DB

servers (even though no outbound rules are configured in the DB security group) and further into the EC2 instances (even though no inbound rules are configured in the EC2 instance security group)

Incorrect options:

Create an outbound rule in the security group for the EC2 instance app servers using TCP protocol on the ephemeral port range. Set the destination as the security group for the MySQL DB servers - As explained above, you need to set up an outbound rule in the security group for the EC2 instance app servers using TCP protocol on port 3306 and NOT on the ephemeral port range because the MySQL DB is configured to process requests on port 3306. A common use-case for ephemeral ports: these are used in NACLs to handle response traffic. Consider a custom network ACL for a VPC that supports IPv4 only. It includes rules that allow HTTP and HTTPS traffic in (inbound rules 100 and 110). There's a corresponding outbound rule that enables responses to that inbound traffic (outbound rule 140, which covers ephemeral ports 32768-65535). So this option is incorrect.

Ephemeral ports

The example network ACL in the preceding section uses an ephemeral port range of 32768-65535. However, you might want to use a different range for your network ACLs depending on the type of client that you're using or with which you're communicating.

The client that initiates the request chooses the ephemeral port range. The range varies depending on the client's operating system.

- Many Linux kernels (including the Amazon Linux kernel) use ports 32768-61000.
- Requests originating from Elastic Load Balancing use ports 1024-65535.
- Windows operating systems through Windows Server 2003 use ports 1025-5000.
- Windows Server 2008 and later versions use ports 49152-65535.
- A NAT gateway uses ports 1024-65535.
- AWS Lambda functions use ports 1024-65535.

For example, if a request comes into a web server in your VPC from a Windows 10 client on the internet, your network ACL must have an outbound rule to enable traffic destined for ports 49152-65535.

If an instance in your VPC is the client initiating a request, your network ACL must have an inbound rule to enable traffic destined for the ephemeral ports specific to the type of instance (Amazon Linux, Windows Server 2008, and so on).

In practice, to cover the different types of clients that might initiate traffic to public-facing instances in your VPC, you can open ephemeral ports 1024-65535. However, you can also add rules to the ACL to deny traffic on any malicious ports within that range. Ensure that you place the *deny* rules earlier in the table than the *allow* rules that open the wide range of ephemeral ports.

via -

<https://docs.aws.amazon.com/vpc/latest/userguide/vpc-network-acls.html#nacl-ephemeral-ports>

Create an outbound rule in the security group for the MySQL DB servers using TCP protocol on the ephemeral port range. Set the destination as the security group for the EC2 instance app servers

Create an outbound rule in the security group for the MySQL DB servers using TCP protocol on port 3306. Set the destination as the security group for the EC2 instance app servers

There is no need to create an outbound rule in the security group for the MySQL DB servers either on the ephemeral port range or port 3306 since the security groups are stateful therefore the response from the DB servers would be allowed out of the DB servers. Hence both these options are incorrect.

References:

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

<https://docs.aws.amazon.com/vpc/latest/userguide/vpc-network-acls.html#nacl-ephemeral-ports>

Question 1: **Correct**

A leading internet television network company uses AWS Cloud for analytics, recommendation engines and video transcoding. To monitor and optimize this network, the engineering team at the company has developed a solution for ingesting, augmenting, and analyzing the multiple terabytes of data its network generates daily in the form of virtual private cloud (VPC) flow logs. This would enable the company to identify performance-improvement opportunities such as identifying apps that are communicating across regions and collocating them. The VPC flow logs data is funneled into Kinesis Data Streams which further acts as the source of a delivery stream for Kinesis Firehose. The engineering team has now configured a Kinesis Agent to send the VPC flow logs data from another set of network devices to the same Firehose delivery stream. They noticed that data is not reaching Firehose as expected.

As a Solutions Architect Professional, which of the following options would you identify as the MOST plausible root cause behind this issue?

-
- Kinesis Firehose delivery stream has reached its limit and needs to be scaled manually**
-
- Kinesis Agent cannot write to a Kinesis Firehose for which the delivery stream source is already set as Kinesis Data Streams**
- (Correct)**
-
- Kinesis Agent can only write to Kinesis Data Streams, not to Kinesis Firehose**
-
- The data sent by Kinesis Agent is lost because of a configuration error**

Explanation

Correct option:

Kinesis Agent cannot write to a Kinesis Firehose for which the delivery stream source is already set as Kinesis Data Streams

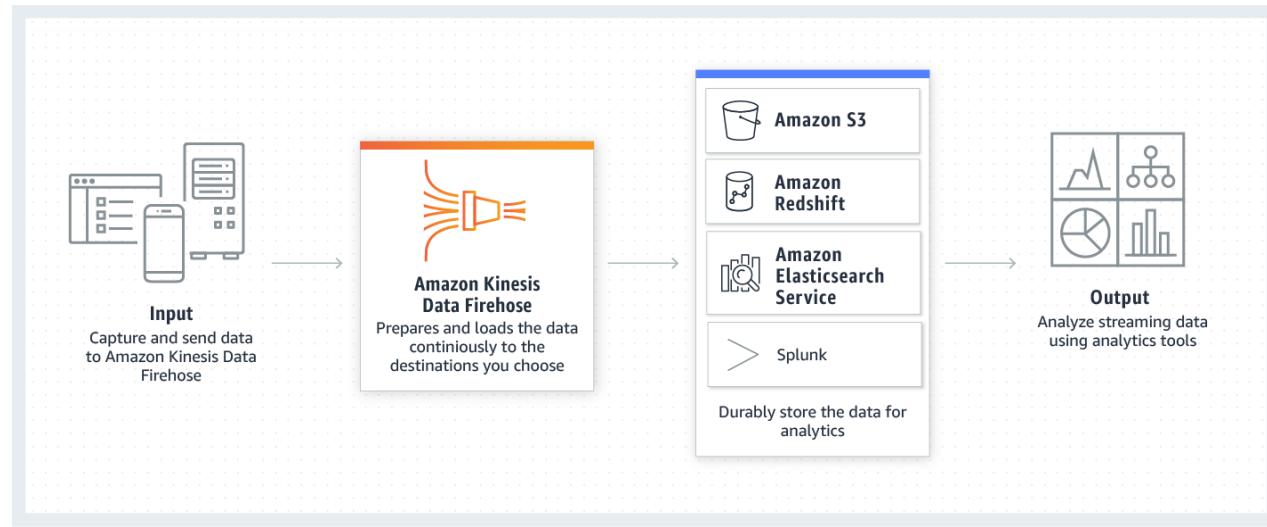
Amazon Kinesis Data Firehose is the easiest way to reliably load streaming data into data lakes, data stores, and analytics tools. It is a fully managed service that automatically scales to match the throughput of your data and requires no ongoing administration. It can also batch, compress, transform, and encrypt the data before loading it, minimizing the amount of storage used at the destination and increasing security.

When a Kinesis data stream is configured as the source of a Firehose delivery stream, Firehose's PutRecord and PutRecordBatch operations are disabled and Kinesis Agent cannot write to Firehose delivery stream directly. Data needs to be added to the Kinesis data stream through the Kinesis Data Streams PutRecord and PutRecords operations instead. Therefore, this option is correct.

Incorrect options:

Kinesis Agent can only write to Kinesis Data Streams, not to Kinesis Firehose - Kinesis Agent is a stand-alone Java software application that offers an easy way to collect and send data to Kinesis Data Streams or Kinesis Firehose. So this option is incorrect.

Kinesis Firehose delivery stream has reached its limit and needs to be scaled manually - Kinesis Firehose is a fully managed service that automatically scales to match the throughput of your data and requires no ongoing administration. Therefore this option is not correct.



How Kinesis Firehose works:

<https://aws.amazon.com/kinesis/data-firehose/>

The data sent by Kinesis Agent is lost because of a configuration error - This is a made-up option and has been added as a distractor.

References:

<https://aws.amazon.com/kinesis/data-firehose/>

<https://aws.amazon.com/kinesis/data-firehose/faqs/>

Question 2: **Incorrect**

A solo entrepreneur is working on a new digital media startup and wants to have a hands-on understanding of the comparative pricing for various storage types available on AWS Cloud. The entrepreneur has created a test file of size 5 GB with some random data. Next, he uploads this test file into AWS S3 Standard storage class, provisions an EBS volume (General Purpose SSD (gp2)) with 50 GB of provisioned storage and copies the test file into the EBS volume, and lastly copies the test file into an EFS Standard Storage filesystem. At the end of the month, he analyses the bill for costs incurred on the respective storage types for the test file.

What of the following represents the correct order of the storage charges incurred for the test file on these three storage types?



Cost of test file storage on S3 Standard < Cost of test file storage on EFS < Cost of test file storage on EBS

(Correct)



Cost of test file storage on EBS < Cost of test file storage on S3 Standard < Cost of test file storage on EFS



Cost of test file storage on S3 Standard < Cost of test file storage on EBS < Cost of test file storage on EFS

(Incorrect)



Cost of test file storage on EFS < Cost of test file storage on S3 Standard < Cost of test file storage on EBS

Explanation

Correct option: **Cost of test file storage on S3 Standard < Cost of test file storage on EFS < Cost of test file storage on EBS**

With Amazon EFS, you pay only for the resources that you use. The EFS Standard Storage pricing is \$0.30 per GB per month.

Therefore the cost for storing the test file of 5 GB on EFS is $5 * \$0.30 = \1.5 for the month.

Pricing Table

Region: [US East \(Ohio\)](#) ▾

Standard Storage (GB-Month)	\$0.30
Infrequent Access Storage (GB-Month)	\$0.025
Infrequent Access Requests (per GB transferred)	\$0.01
Provisioned Throughput (MB/s-Month)	\$6.00

via -

<https://aws.amazon.com/efs/pricing/>

For EBS General Purpose SSD (gp2) volumes, the charges are \$0.10 per GB-month of provisioned storage. Therefore, for a provisioned storage of 50 GB for this use-case, the monthly cost on EBS is $\$0.10 \times 50 = \5 . This cost is irrespective of how much storage is actually consumed by the test file.

Amazon EBS Volumes

With Amazon EBS, you pay only for what you use. The pricing for Amazon EBS volumes is listed below

General Purpose SSD (gp2) Volumes	\$0.10 per GB-month of provisioned storage
Provisioned IOPS SSD (io2) Volumes	\$0.125 per GB-month of provisioned storage AND \$0.065 per provisioned IOPS-month
Provisioned IOPS SSD (io1) Volumes	\$0.125 per GB-month of provisioned storage AND \$0.065 per provisioned IOPS-month
Throughput Optimized HDD (st1) Volumes	\$0.045 per GB-month of provisioned storage
Cold HDD (sc1) Volumes	\$0.025 per GB-month of provisioned storage

via -

<https://aws.amazon.com/ebs/pricing/>

For S3 Standard storage, the pricing is \$0.023 per GB per month. Therefore, the monthly storage cost on S3 for the test file of 5 GB is $\$0.023 * 5 = \0.115

Storage pricing	
S3 Standard - General purpose storage for any type of data, typically used for frequently accessed data	
First 50 TB / Month	\$0.023 per GB
Next 450 TB / Month	\$0.022 per GB
Over 500 TB / Month	\$0.021 per GB
S3 Intelligent - Tiering * - Automatic cost savings for data with unknown or changing access patterns	
Frequent Access Tier, First 50 TB / Month	\$0.023 per GB
Frequent Access Tier, Next 450 TB / Month	\$0.022 per GB
Frequent Access Tier, Over 500 TB / Month	\$0.021 per GB
Infrequent Access Tier, All Storage / Month	\$0.0125 per GB
Monitoring and Automation, All Storage / Month	\$0.0025 per 1,000 objects
S3 Standard - Infrequent Access * - For long lived but infrequently accessed data that needs millisecond access	
All Storage / Month	\$0.0125 per GB
S3 One Zone - Infrequent Access * - For re-createable infrequently accessed data that needs millisecond access	
All Storage / Month	\$0.01 per GB
S3 Glacier ** - For long-term backups and archives with retrieval option from 1 minute to 12 hours	
All Storage / Month	\$0.004 per GB
S3 Glacier Deep Archive ** - For long-term data archiving that is accessed once or twice in a year and can be restored within 12 hours	
All Storage / Month	\$0.00099 per GB

via -

<https://aws.amazon.com/s3/pricing/>

Therefore this is the correct option.

Incorrect options: **Cost of test file storage on S3 Standard < Cost of test file storage on EBS < Cost of test file storage on EFS**

Cost of test file storage on EFS < Cost of test file storage on S3 Standard < Cost of test file storage on EBS

Cost of test file storage on EBS < Cost of test file storage on S3 Standard < Cost of test file storage on EFS

Following the computations shown earlier in the explanation, these three options are incorrect.

References: <https://aws.amazon.com/ebs/pricing/>

<https://aws.amazon.com/s3/pricing/> (<https://aws.amazon.com/s3/pricing/>)

<https://aws.amazon.com/efs/pricing/>

Question 3: Incorrect

A multi-national company uses Amazon S3 as its data lake to store the data that flows into its business. This data is both structured and semi-structured and is organized under different buckets in the company's AWS account in the same Region. Hundreds of applications in the company's AWS account use structured data for running data analytics, event monitoring, report generation, event creation, and many more. While the semi-structured data runs through several transformations and is sent to downstream applications for further processing. While the company's security policy restricts S3 bucket access over the internet, the internal security team has requested tighter access rules for the applications using the S3 data lake.

Which combination of steps will you undertake to implement this requirement in the most efficient way? (Select three)

-

From each application VPC, create a gateway endpoint for Amazon S3. Configure the endpoint policy to allow access to an S3 access point. Specify the route table that is used to access the access point

(Incorrect)

-

Create an interface endpoint for Amazon S3 in each application's VPC. Configure the endpoint policy to allow access to an S3 access point. Create a VPC gateway attachment for the S3 endpoint

-

Create an S3 access point for each application in each AWS account and attach the access points to the S3 bucket. Configure each access point to be accessible only from the application's VPC. Update the bucket policy to require access from an access point

-

In the AWS account that owns the S3 buckets, create an S3 access point for each bucket that the applications must use to access the data. Set up all applications in a single data lake VPC

(Correct)

-

Add a bucket policy on the buckets to deny access from applications outside the data lake VPC

(Correct)

-

Create a gateway endpoint for Amazon S3 in the data lake VPC. Attach an endpoint policy to allow access to the S3 bucket only via the access points. Specify the route table that is used to access the bucket

(Correct)

Explanation

Correct options:

In the AWS account that owns the S3 buckets, create an S3 access point for each bucket that the applications must use to access the data. Set up all applications in a single data lake VPC

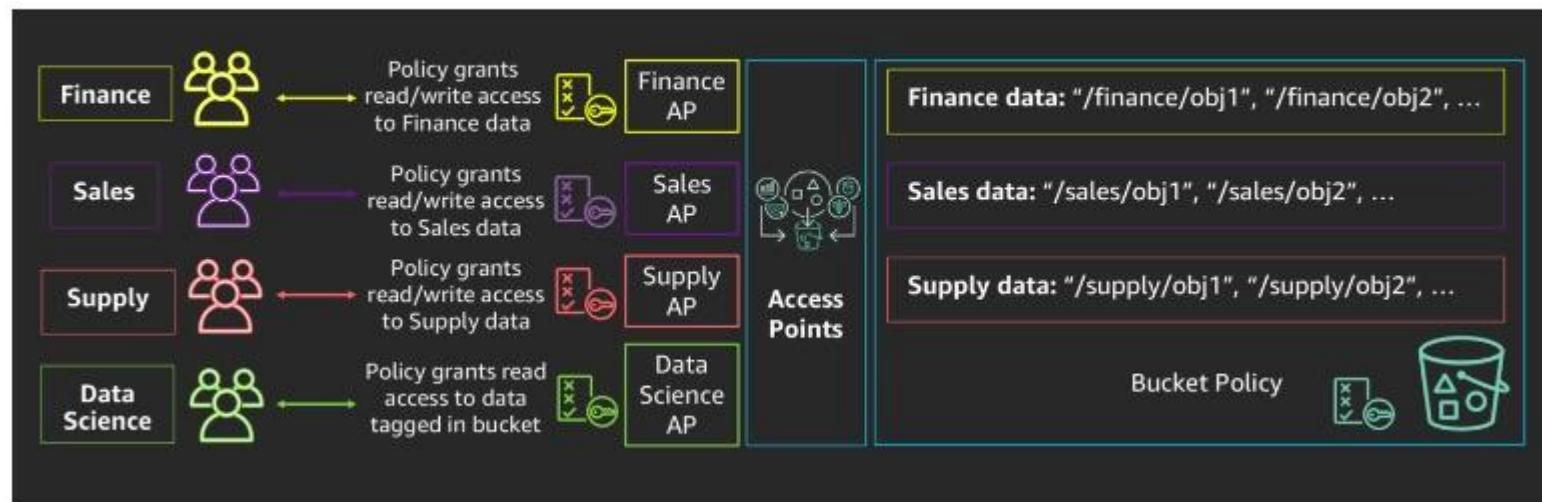
Here is the approach that uses S3 Access Points in combination with VPC endpoint policies to make it easy to manage access to shared datasets on Amazon S3. The idea is to create an Amazon S3 VPC-Only Access Point, and then use it in the VPC endpoint policy to control access to the S3 bucket. You also have the option to use bucket policies to firewall S3 bucket access to VPCs only.

S3 Access Points are unique hostnames that you can create to enforce distinct permissions and network controls for any request made through the Access Point.

Some key features of S3 Access Points: 1. Access Points contain a hostname, an AWS ARN, and an AWS IAM resource policy. 2. Access Points by default have a specific setting to Block Public Access. 3. Access Points are unique to an account and Region. 4. Access Points can have custom IAM permissions for a user or application. 5. Access Points can have custom IAM permissions to specific objects in a bucket via a prefix to precisely control access. 6. Access Points can be configured to accept requests only from a virtual private cloud (VPC) to restrict Amazon S3 data access to a private network.

Use S3 Access Points to manage access to shared datasets on Amazon

The following image shows one example of how you can use S3 Access Points to manage access to shared datasets on Amazon S3.

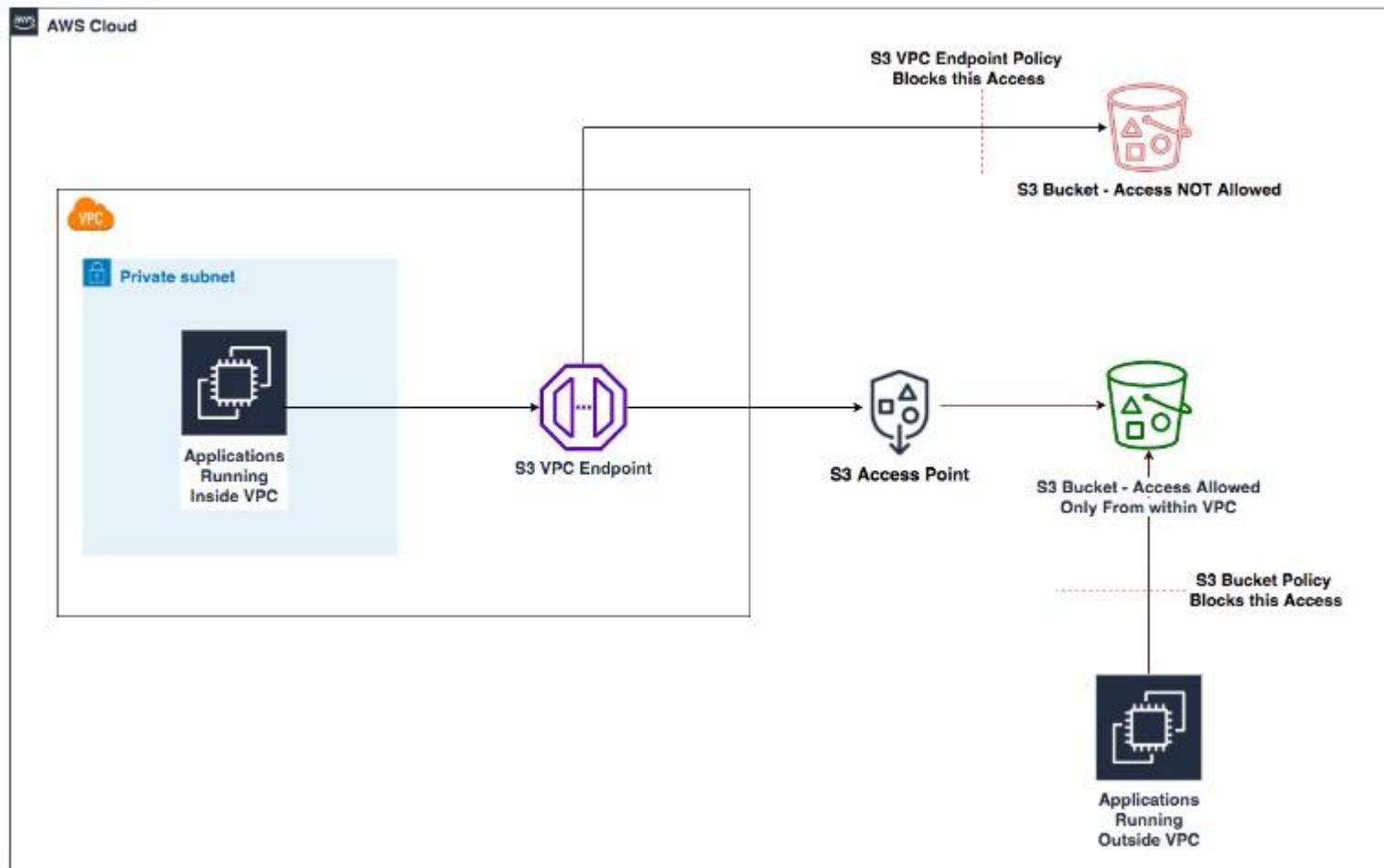


S3:

via - <https://aws.amazon.com/blogs/storage/managing-amazon-s3-access-with-vpc-endpoints-and-s3-access-points/>

How to set up S3 Access Points for an Amazon S3 bucket and use it with VPC endpoint:

We now look at how to set up S3 Access Points for an Amazon S3 bucket and use it with VPC endpoints. The following diagram shows the setup in full:



via -

<https://aws.amazon.com/blogs/storage/managing-amazon-s3-access-with-vpc-endpoints-and-s3-access-points/>

Create a gateway endpoint for Amazon S3 in the data lake VPC. Attach an endpoint policy to allow access to the S3 bucket only via the access points. Specify the route table that is used to access the bucket

You can access Amazon S3 from your VPC using gateway VPC endpoints. After you create the gateway endpoint, you can add it as a target in your route table for traffic destined from your VPC to Amazon S3. Leverage the following condition with a deny effect in the gateway endpoint policy:

```
"Condition": {  
    "ArnNotLikeIfExists": {  
        "s3:DataAccessPointArn": "arn:aws:s3:us-east-1:<Account ID>:accesspoint/*"  
    }  
}
```

When a new Amazon S3 bucket is created, to allow access from the VPC, you can create an S3 Access Point on the S3 bucket. The preceding condition in the VPC endpoint policy would automatically allow access to this new S3 bucket via the Access Point, without having to edit the VPC endpoint policy.

Add a bucket policy on the buckets to deny access from applications outside the data lake VPC

Broadly, the steps involved are: 1. Create a VPC-only Access Point for the Amazon S3 bucket. This makes sure that this Access Point can only be accessed by resources in a specific VPC.

1. Create Amazon S3 gateway endpoint in the VPC and add a VPC endpoint policy. This VPC endpoint policy will have a statement that allows S3 access only via access points owned by the organization. We take advantage of the account ID in the Access Point ARN to make this possible.
2. Add a bucket policy on the bucket to allow access only from the VPC: This prevents any access from outside the VPC.

Incorrect options:

Create an interface endpoint for Amazon S3 in each application's VPC. Configure the endpoint policy to allow access to an S3 access point. Create a VPC gateway attachment for the S3 endpoint - You need to create an S3 access point for Amazon S3 in each application's VPC and not an interface endpoint. Gateway attachments are used with Transit Gateways and not with S3 endpoints.

Create an S3 access point for each application in each AWS account and attach the access points to the S3 bucket. Configure each access point to be accessible only from the application's VPC. Update the bucket policy to require access from an access point - This statement is incorrect. Amazon S3 access point can only be created from the AWS account that owns the S3 bucket.

From each application VPC, create a gateway endpoint for Amazon S3. Configure the endpoint policy to allow access to an S3 access point. Specify the route table that is used to access the access point - There is no need to create separate VPCs for each application, as just a single data lake VPC can house all applications, which allows you to configure a single S3 gateway endpoint having a policy with a condition to limit access via a common prefix for the access points of all the S3 buckets for the data lake. So this option is not the best fit.

References:

<https://aws.amazon.com/s3/features/access-points/>

<https://aws.amazon.com/blogs/storage/managing-amazon-s3-access-with-vpc-endpoints-and-s3-access-points/>

Question 4: Correct

The DevOps team at a leading SaaS company is planning to release the major upgrade of its flagship CRM application in a week. The team is testing the alpha release of the application running on 20 EC2 instances managed by an Auto Scaling group in subnet 172.20.0.0/24 within VPC X with CIDR block 172.20.0.0/16. The team has noticed connection timeout errors in the application logs while connecting to a MySQL database running on an EC2 instance in the same region in subnet 172.30.0.0/24 within VPC Y with CIDR block 172.30.0.0/16. The IP of the database instance is hard-coded in the application instances.

As a Solutions Architect Professional, which of the following solutions would you recommend to the DevOps team to solve the problem in a secure way with minimal maintenance and overhead? (Select two)

-

Create and attach internet gateways for both VPCs and set up default routes to the Internet gateways for both VPCs. Assign an Elastic IP for the EC2 instance running MySQL database in VPC Y. Update the application instances to connect to this Elastic IP

-

Create and attach virtual private gateways for both VPCs and set up default routes to the customer gateways for both VPCs. Assign an Elastic IP for the EC2 instance running MySQL database in VPC Y. Update the application instances to connect to this Elastic IP

-

Set up a VPC peering connection between the two VPCs and add a route to the routing table of VPC Y that points to the IP address range of 172.20.0.0/16

(Correct)

-

Create and attach NAT gateways for both VPCs and set up routes to the NAT gateways for both VPCs. Assign an Elastic IP for the EC2 instance running MySQL database in VPC Y. Update the application instances to connect to this Elastic IP

-

Set up a VPC peering connection between the two VPCs and add a route to the routing table of VPC X that points to the IP address range of 172.30.0.0/16

(Correct)

Explanation

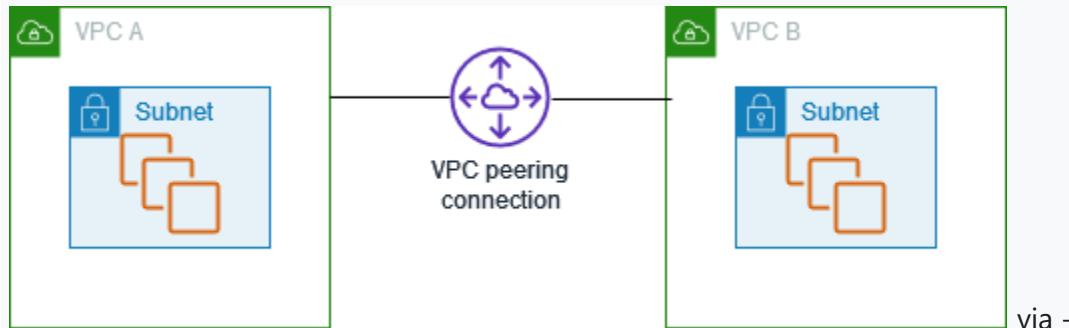
Correct option:

Set up a VPC peering connection between the two VPCs and add a route to the routing table of VPC X that points to the IP address range of 172.30.0.0/16

Set up a VPC peering connection between the two VPCs and add a route to the routing table of VPC Y that points to the IP address range of 172.20.0.0/16

A VPC peering connection is a networking connection between two VPCs that enables you to route traffic between them using private IPv4 addresses or IPv6 addresses. Instances in either VPC can communicate with each other as if they are within the same

network. You can create a VPC peering connection between your own VPCs, or with a VPC in another AWS account. The VPCs can be in different regions (also known as an inter-region VPC peering connection).



via -

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

For the given use-case, you must set up a VPC peering connection between the two VPCs and then add a route to the route table of VPC X that's associated with your subnet in which the 20 EC2 instances running the application reside. The route points to the CIDR block of the peer VPC Y (172.30.0.0/16) or it can also point to just a portion of the CIDR block such as the subnet 172.30.0.0/24 in the VPC peering connection and lastly specify the VPC peering connection as the target. You should also add a route to the routing table of VPC Y that points back to VPC X via the IP address range of 172.20.0.0/16.

Updating your Route tables for a VPC peering connection

PDF

To send private IPv4 traffic from your instance to an instance in a peer VPC, you must add a route to the route table that's associated with your subnet in which your instance resides. The route points to the CIDR block (or portion of the CIDR block) of the peer VPC in the VPC peering connection, and specifies the VPC peering connection as the target.

Similarly, if the VPCs in the VPC peering connection have associated IPv6 CIDR blocks, you can add a route to your route table to enable communication with the peer VPC over IPv6.

If a subnet is not explicitly associated with a route table, it uses the main route table by default. For more information, see [Route Tables](#) in the *Amazon VPC User Guide*.

You have a [quota](#) on the number of entries you can add per route table. If the number of VPC peering connections in your VPC exceeds the route table entry quota for a single route table, consider using multiple subnets that are each associated with a custom route table.

For more information about supported route table configurations for VPC peering connections, see [VPC peering configurations](#).

You can add a route for a VPC peering connection that's in the pending-acceptance state. However, the route will have a state of `blackhole` and have no effect until the VPC peering connection is in the active state.

⚠ Warning

If you have a VPC peered with multiple VPCs that have overlapping or matching IPv4 CIDR blocks, ensure that your route tables are configured to avoid sending response traffic from your VPC to the incorrect VPC. AWS currently does not support unicast reverse path forwarding in VPC peering connections that checks the source IP of packets and routes reply packets back to the source. For more information, see [Routing for response traffic](#).

To add an IPv4 route for a VPC peering connection

1. Open the Amazon VPC console at <https://console.aws.amazon.com/vpc/>.
 2. In the navigation pane, choose [Route Tables](#).
 3. Select the route table that's associated with the subnet in which your instance resides.
- ### ⓘ Note
- If you do not have a route table associated with that subnet, select the main route table for the VPC, as the subnet then uses this route table by default.
4. Choose [Routes](#), [Edit](#), [Add Route](#).
 5. For [Destination](#), enter the IPv4 address range to which the network traffic in the VPC peering connection must be directed. You can specify the entire IPv4 CIDR block of the peer VPC, a specific range, or an individual IPv4 address, such as the IP address of the instance with which to communicate. For example, if the CIDR block of the peer VPC is `10.0.0.0/16`, you can specify a portion `10.0.0.0/28`, or a specific IP address `10.0.0.7/32`.
 6. Select the VPC peering connection from [Target](#), and then choose [Save](#).

Destination	Target	Status	Propagated	Remove
192.168.0.0/28	local	Active	No	
10.0.0.0/28	pcx-c379f9aa	Active	No	X

Add another route

The owner of the peer VPC must also complete these steps to add a route to direct traffic back to your VPC through the VPC peering connection.

If both VPCs in the VPC peering connection are in the same region, have IPv6 CIDR blocks, and the resources in the VPC are enabled to use IPv6, you can also add a route for IPv6 communication.

via -

<https://docs.aws.amazon.com/vpc/latest/peering/vpc-peering-routing.html>

Incorrect options:

Create and attach NAT gateways for both VPCs and set up routes to the NAT gateways for both VPCs. Assign an Elastic IP for the EC2 instance running MySQL database in VPC Y. Update the application instances to connect to this Elastic IP - You can attach NAT gateways to enable instances in a private subnet to connect to the internet (for example, for software updates) or other AWS services, but prevent the internet from initiating connections with the instances. A NAT gateway forwards traffic from the instances in the private subnet to the internet or other AWS services, and then sends the response back to the instances. NAT gateways are attached in the public subnets. Additionally, adding Elastic IP to the EC2 instance running MySQL database would necessitate updating configurations in the application instances. So, this option is incorrect.

Create and attach internet gateways for both VPCs and set up default routes to the Internet gateways for both VPCs. Assign an Elastic IP for the EC2 instance running MySQL database in VPC Y. Update the application instances to connect to this Elastic IP - Attaching internet gateway to both VPCs would take away the security considerations for the given use-case as in that case the application instances would connect to the EC2 instance running MySQL database over the public internet. Additionally, adding Elastic IP to the EC2 instance running MySQL database would necessitate updating configurations in the application instances. So, this option is incorrect.

Create and attach VPC Gateway endpoints for both VPCs and set up default routes to the Gateway endpoints for both VPCs. Assign an Elastic IP for the EC2 instance running MySQL database in VPC Y. Update the application instances to connect to this Elastic IP - A VPC endpoint enables you to privately connect your VPC to supported AWS services. Traffic between your VPC and the other service does not leave the Amazon network. A gateway endpoint is a gateway that you specify as a target for a route in your route table for traffic destined to a supported AWS service such as Amazon S3. Gateway endpoints cannot be used for establishing routes between VPCs. Additionally, adding Elastic IP to the EC2 instance running MySQL database would necessitate updating configurations in the application instances. So, this option is incorrect.

References:

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

<https://docs.aws.amazon.com/vpc/latest/peering/vpc-peering-routing.html>

Question 5: **Incorrect**

A big data analytics company is leveraging AWS Cloud to process Internet of Things (IoT) sensor data from the field devices of an agricultural sciences company. The analytics company stores the IoT sensor data in Amazon DynamoDB tables. To detect anomalous behaviors and respond quickly, all changes to the items stored in the DynamoDB tables must be logged in near real-time.

As an AWS Certified Solutions Architect Professional, which of the following solutions would you recommend to meet the requirements of the given use-case so that it requires minimal custom development and infrastructure maintenance?

-
- Set up DynamoDB Streams to capture and send updates to a Lambda function that outputs records directly to Kinesis Data Analytics (KDA). Detect and analyze anomalies in KDA and send notifications via SNS**
-
- Configure event patterns in EventBridge events to capture DynamoDB API call events and set up Lambda function as a target to analyze anomalous behavior. Send SNS notifications when anomalous behaviors are detected**
-
- Set up CloudTrail to capture all API calls that update the DynamoDB tables. Leverage CloudTrail event filtering to analyze anomalous behaviors and send SNS notifications in case anomalies are detected**

(Incorrect)

Set up DynamoDB Streams to capture and send updates to a Lambda function that outputs records to Kinesis Data Analytics (KDA) via Kinesis Data Streams (KDS). Detect and analyze anomalies in KDA and send notifications via SNS

(Correct)

Explanation

Correct option:

Set up DynamoDB Streams to capture and send updates to a Lambda function that outputs records to Kinesis Data Analytics (KDA) via Kinesis Data Streams (KDS). Detect and analyze anomalies in KDA and send notifications via SNS

A DynamoDB stream is an ordered flow of information about changes to items in a DynamoDB table. When you enable a stream on a table, DynamoDB captures information about every modification to data items in the table for up to 24 hours.

Whenever an application creates, updates, or deletes items in the table, DynamoDB Streams writes a stream record with the primary key attributes of the items that were modified. A stream record contains information about a data modification to a single item in a DynamoDB table.

DynamoDB Streams supports the following stream record views:

KEYS_ONLY — Only the key attributes of the modified item

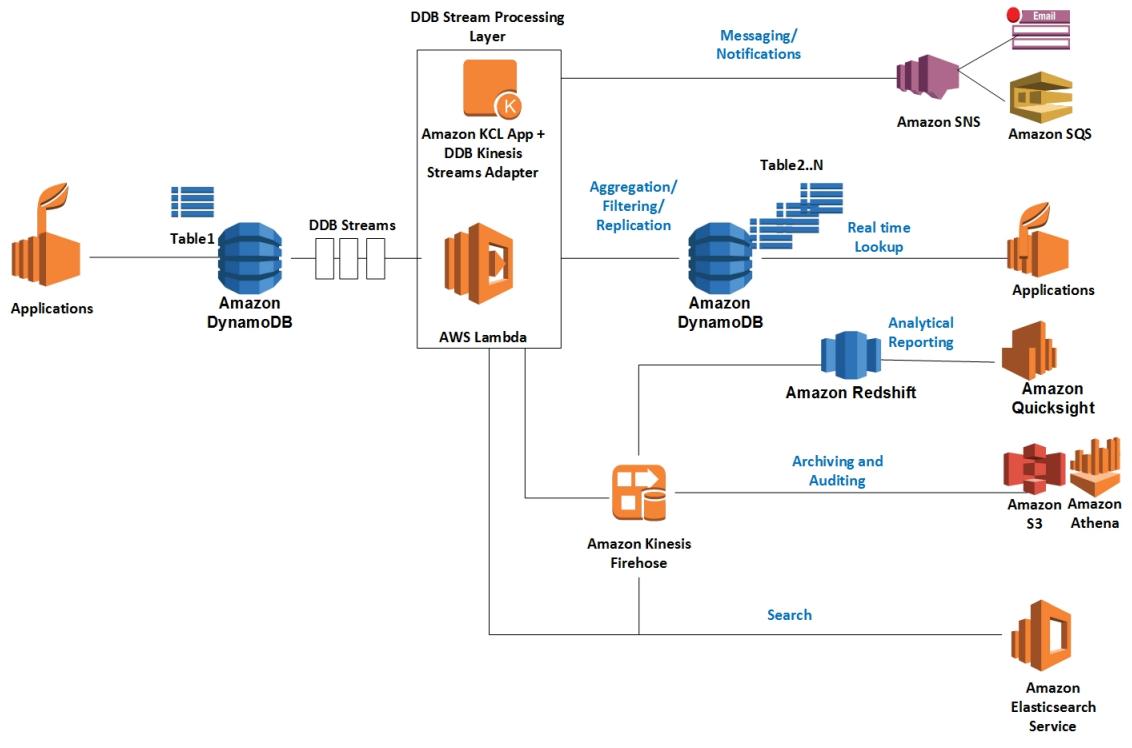
NEW_IMAGE — The entire item, as it appears after it was modified

OLD_IMAGE — The entire item, as it appears before it was modified

NEW_AND_OLD_IMAGES — Both the new and the old images of the item

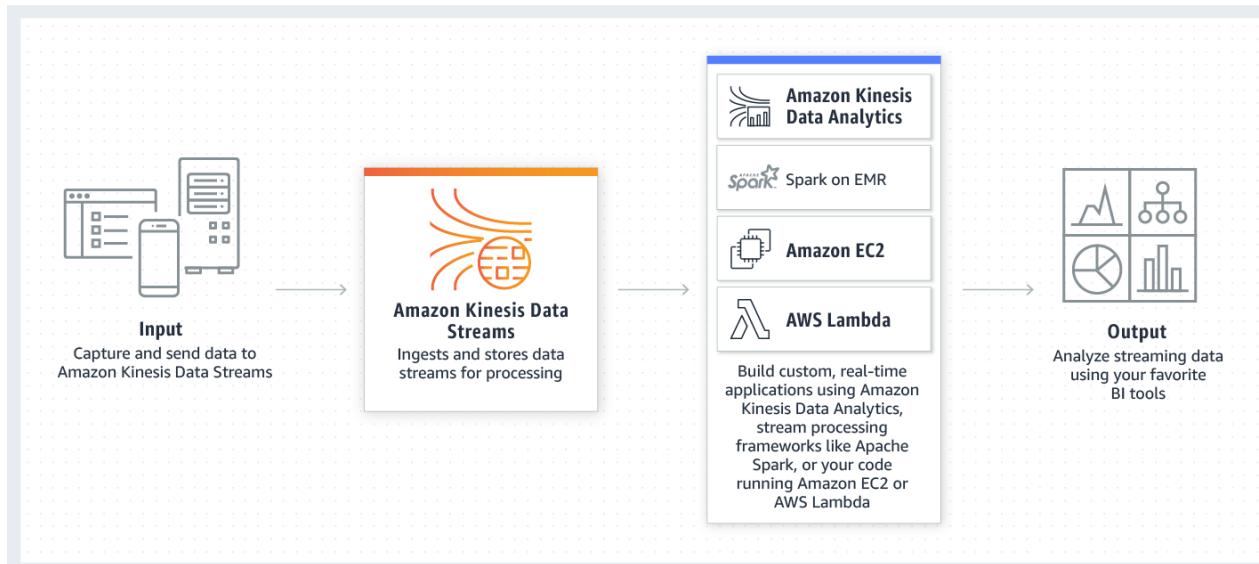
You can process DynamoDB streams in multiple ways. The most common approaches use AWS Lambda or a standalone application that uses the Kinesis Client Library (KCL) with the DynamoDB Streams Kinesis Adapter. The KCL is a client-side library that provides an interface to process DynamoDB stream changes. If you enable DynamoDB Streams on a table, you can associate the stream Amazon Resource Name (ARN) with an AWS Lambda function that you write. Immediately after an item in the table is modified, a new record appears in the table's stream. AWS Lambda polls the stream and invokes your Lambda function synchronously when it detects new stream records.

Please review this excellent reference architecture for DynamoDB streams design patterns:



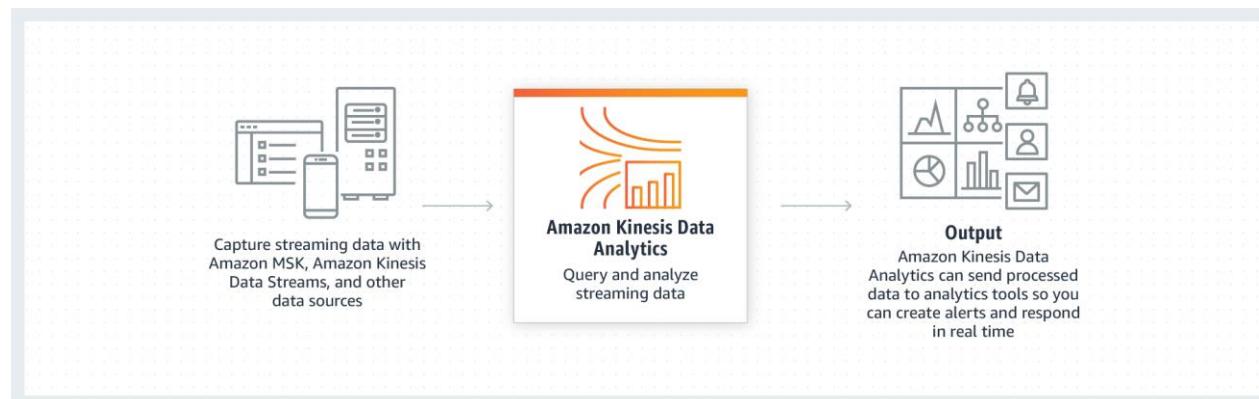
<https://aws.amazon.com/blogs/database/dynamodb-streams-use-cases-and-design-patterns/>

For the given use-case, you can use a Lambda function to capture updates from DynamoDB Streams and send those records to KDA via KDS. You can then detect and analyze anomalies in KDA and send notifications via SNS.



How KDS Works:

<https://aws.amazon.com/kinesis/data-streams/>



How KDA Works:

<https://aws.amazon.com/kinesis/data-analytics/>

It is important to note that Kinesis Data Analytics (KDA) only supports the following streaming sources for an application:

A Kinesis data stream (KDS)

A Kinesis Data Firehose (KDF) delivery stream

Therefore, you cannot directly write the output of the records from a Lambda function to KDA, although you can certainly use a Lambda function to pre-process the incoming stream from either KDS or KDF.

Incorrect options:

Set up CloudTrail to capture all API calls that update the DynamoDB tables. Leverage CloudTrail event filtering to analyze anomalous behaviors and send SNS notifications in case anomalies are detected - You can use CloudTrail to capture API calls for DynamoDB as events. The calls captured include calls from the DynamoDB console and code calls to the DynamoDB API operations. If you create a trail, you can enable continuous delivery of CloudTrail events to an Amazon S3 bucket, including events for DynamoDB. The CloudTrail does not support the GetRecords API for DynamoDB Streams so you cannot use it to capture the actual records. Moreover, you cannot use CloudTrail event filtering to analyze anomalous behaviors as it is just a simple filtering mechanism based on certain event attributes such as Read Only, Event Source, Event Time etc.

Configure event patterns in EventBridge events to capture DynamoDB API call events and set up Lambda function as a target to analyze anomalous behavior. Send SNS notifications when anomalous behaviors are detected - EventBridge events service does not offer event type for DynamoDB as it's dependent on CloudTrail to get the relevant API call information. As explained above, CloudTrail itself cannot capture the DynamoDB streams records as CloudTrail does not support the GetRecords API for DynamoDB Streams. Therefore this option is incorrect.

CloudWatch
Dashboards
Alarms
ALARM 0
INSUFFICIENT 0
OK 1
Billing
Logs
Log groups
Insights
Metrics
Events
Rules
Event Buses
ServiceLens
Service Map
Traces
Container Insights NEW

Step 1: Create rule

Create rules to invoke Targets based on Events happening in your AWS environment.

Event Source

Build or customize an Event Pattern or set a Schedule to invoke Targets.

Event Pattern ⓘ Schedule ⓘ

Build event pattern to match events by service

Service Name

Event Type

For AWS API call events, CloudWatch Events supports the same read/write APIs as CloudTrail does. Read-only APIs, such as those that begin with **List**, **Get**, or **Describe** are not supported by CloudWatch Events. [See more details](#) about which services are supported by CloudTrail.

Any operation

Specific operation(s)

Set up DynamoDB Streams to capture and send updates to a Lambda function that outputs records directly to Kinesis Data Analytics (KDA). Detect and analyze anomalies in KDA and send notifications via SNS - As mentioned earlier, KDA only supports KDS and KDF as the streaming sources for an application, so this option is incorrect.

References:

<https://aws.amazon.com/blogs/database/dynamodb-streams-use-cases-and-design-patterns/>

<https://docs.aws.amazon.com/amazondynamodb/latest/developerguide/Streams.Lambda.html>

<https://docs.aws.amazon.com/kinesisanalytics/latest/dev/how-it-works-input.html>

<https://docs.aws.amazon.com/amazondynamodb/latest/developerguide/logging-using-cloudtrail.html>

Question 6: Incorrect

A retail company recently saw a huge spike in its monthly AWS spend. Upon further investigation, it was found that some developers had accidentally launched Amazon RDS instances in unexpected Regions. The company has hired you as an AWS Certified Solutions Architect Professional to establish best practices around least privileges for developers and control access to on-premises as well as AWS Cloud resources using Active Directory. The company has mandated you to institute a mechanism to control costs by restricting the level of access that developers have to the AWS Management Console without impacting their productivity. The company would also like to allow developers to launch RDS instances only in us-east-1 Region without limiting access to other services in any Region.

How can you help the company achieve the new security mandate while minimizing the operational burden on the DevOps team?



Set up an IAM user for each developer and add them to the developer IAM group that has the PowerUserAccess managed policy attached to it. Attach a customer-managed policy that allows the developers access to RDS only in us-east-1 Region



Configure SAML-based authentication tied to an IAM role that has the PowerUserAccess managed policy attached to it. Attach a customer-managed policy that denies access to RDS in any AWS Region except us-east-1

(Correct)



Configure SAML-based authentication tied to an IAM role that has a PowerUserAccess managed policy and a customer-managed policy that denies all the developers access to any AWS services except AWS Service Catalog. Within AWS Service Catalog, create a product containing only RDS service in us-east-1 region

(Incorrect)



Configure SAML-based authentication tied to an IAM role that has the AdministrativeAccess managed policy attached to it. Attach a customer-managed policy that denies access to RDS in any AWS Region except us-east-1

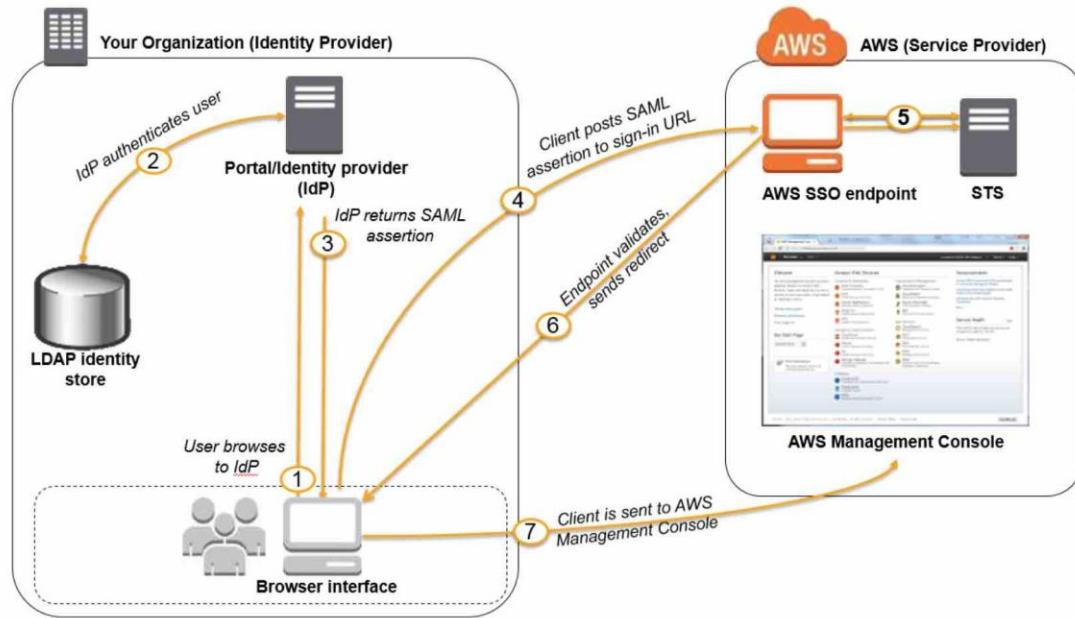
Explanation

Correct option:

Configure SAML-based authentication tied to an IAM role that has the PowerUserAccess managed policy attached to it. Attach a customer-managed policy that denies access to RDS in any AWS Region except us-east-1

Security Assertion Markup Language 2.0 (SAML) is an open federation standard that allows an identity provider (IdP) to authenticate users and pass identity and security information about them to a service provider which is an AWS application or service for the current use-case. With SAML, you can enable a single sign-on experience for your users across many SAML-enabled applications and services. Users authenticate with the IdP once using a single set of credentials, and then get access to multiple applications and services without additional sign-ins.

For the given scenario, the company wants to control access to on-premises as well as AWS Cloud resources (specifically via the AWS Management Console) using Active Directory, so it should use SAML 2.0 federated users to access the AWS Management Console. You also create an IAM role with a trust policy that sets the SAML provider as the principal, which establishes a trust relationship between your organization and AWS. The role's permission policy establishes what users from your organization are allowed to do in AWS. In this case, the role will have a PowerUserAccess managed policy attached. As the PowerUserAccess managed policy will allow the developers to create RDS instances in any Region, therefore, you also need to attach a customer-managed policy that denies access to RDS in any AWS Region except us-east-1.



The diagram illustrates the following steps:

1. The user browses to your organization's portal and selects the option to go to the AWS Management Console. In your organization, the portal is typically a function of your IdP that handles the exchange of trust between your organization and AWS. For example, in Active Directory Federation Services, the portal URL is:
<https://ADFSServiceName/adfs/ls/IdpInitiatedSignOn.aspx>
2. The portal verifies the user's identity in your organization.
3. The portal generates a SAML authentication response that includes assertions that identify the user and include attributes about the user. You can also configure your IdP to include a SAML assertion attribute called SessionDuration that specifies how long the console session is valid. You can also configure the IdP to pass attributes as session tags. The portal sends this response to the client browser.
4. The client browser is redirected to the AWS single sign-on endpoint and posts the SAML assertion.
5. The endpoint requests temporary security credentials on behalf of the user and creates a console sign-in URL that uses those credentials.
6. AWS sends the sign-in URL back to the client as a redirect.
7. The client browser is redirected to the AWS Management Console. If the SAML authentication response includes attributes that map to multiple IAM roles, the user is first prompted to select the role for accessing the console.

From the user's perspective, the process happens transparently: The user starts at your organization's internal portal and ends up at the AWS Management Console, without ever having to supply any AWS credentials.

https://docs.aws.amazon.com/IAM/latest/UserGuide/id_roles_providers_enable-console-saml.html

At a high-level, it is useful to think of these access privileges in the form of this equation:

via -

PowerUserAccess = AdministrativeAccess - IAM

Incorrect options:

Configure SAML-based authentication tied to an IAM role that has the AdministrativeAccess managed policy attached to it.
Attach a customer-managed policy that denies access to RDS in any AWS Region except us-east-1 - Using an IAM role with an AdministrativeAccess managed policy attached to it would violate the key requirement of providing the least privileges for developers. PowerUserAccess provides full access to AWS services and resources but does not allow management of users and groups.

At a high-level, it is useful to think of these access privileges in the form of this equation:

PowerUserAccess = AdministrativeAccess - IAM

So, PowerUserAccess provides just the right access privileges required for the given use-case.

Identity and Access Management (IAM)

Policies > PowerUserAccess

Summary

Policy ARN: arn:aws:iam::aws:policy/PowerUserAccess [Edit](#)

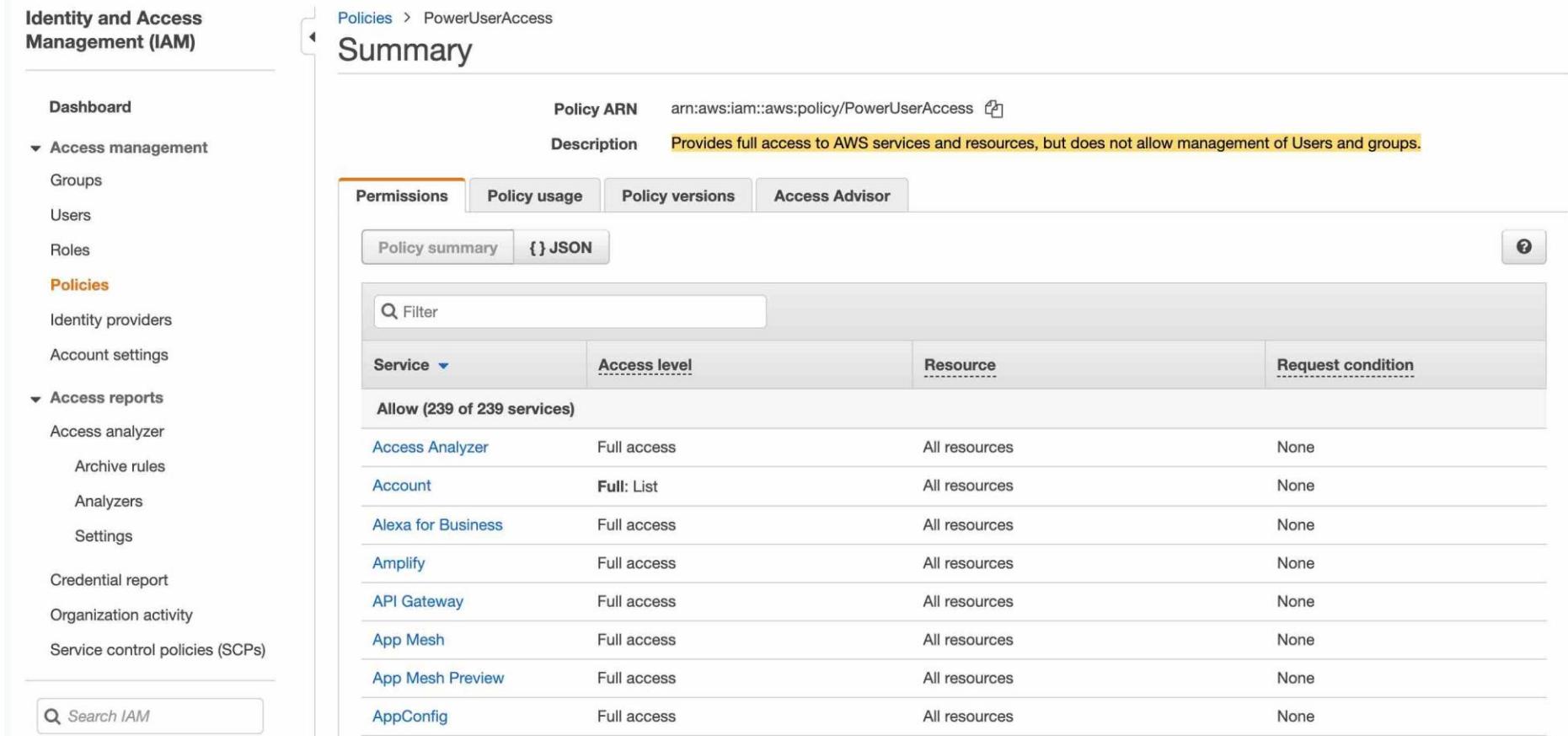
Description: Provides full access to AWS services and resources, but does not allow management of Users and groups.

Permissions Policy usage Policy versions Access Advisor

Policy summary { } JSON [?](#)

Service	Access level	Resource	Request condition
Allow (239 of 239 services)			
Access Analyzer	Full access	All resources	None
Account	Full: List	All resources	None
Alexa for Business	Full access	All resources	None
Amplify	Full access	All resources	None
API Gateway	Full access	All resources	None
App Mesh	Full access	All resources	None
App Mesh Preview	Full access	All resources	None
AppConfig	Full access	All resources	None

Search IAM



Set up an IAM user for each developer and add them to the developer IAM group that has the PowerUserAccess managed policy attached to it. Attach a customer-managed policy that allows the developers access to RDS only in us-east-1 Region -

Setting up an IAM user for each developer and add them to the developer IAM group goes against the requirement of minimizing the operational burden on the DevOps team because this solution does not take advantage of the existing Active Directory that supports SAML-based authentication.

Configure SAML-based authentication tied to an IAM role that has a PowerUserAccess managed policy and a customer-managed policy that denies all the developers access to any AWS services except AWS Service Catalog. Within AWS Service

Catalog, create a product containing only RDS service in us-east-1 region - This option is a distractor as it's too restrictive. As the customer-managed policy denies the developers access to any AWS services except AWS Service Catalog, therefore it would limit access to all other services in any Region, hence this option is incorrect.

References:

https://docs.aws.amazon.com/IAM/latest/UserGuide/id_roles_providers_enable-console-saml.html

https://docs.aws.amazon.com/IAM/latest/UserGuide/id_roles_providers_saml.html

Question 7: **Correct**

The product team at a global IoT technology company is looking to build features to facilitate better collaboration with the company's customers. As part of its research, the product team has figured out a market need to support both stateful and stateless client-server communications via the APIs developed using its platform.

You have been hired by the company as an AWS Certified Solutions Architect Professional to build a solution to fulfill this market need using AWS API Gateway. Which of the following would you recommend to the company?

-

API Gateway creates RESTful APIs that enable stateless client-server communication and API Gateway also creates WebSocket APIs that adhere to the WebSocket protocol, which enables stateful, full-duplex communication between client and server

(Correct)

-

API Gateway creates RESTful APIs that enable stateless client-server communication and API Gateway also creates WebSocket APIs that adhere to the WebSocket protocol, which enables stateless, full-duplex communication between client and server

-

API Gateway creates RESTful APIs that enable stateful client-server communication and API Gateway also creates WebSocket APIs that adhere to the WebSocket protocol, which enables stateless, full-duplex communication between client and server

-

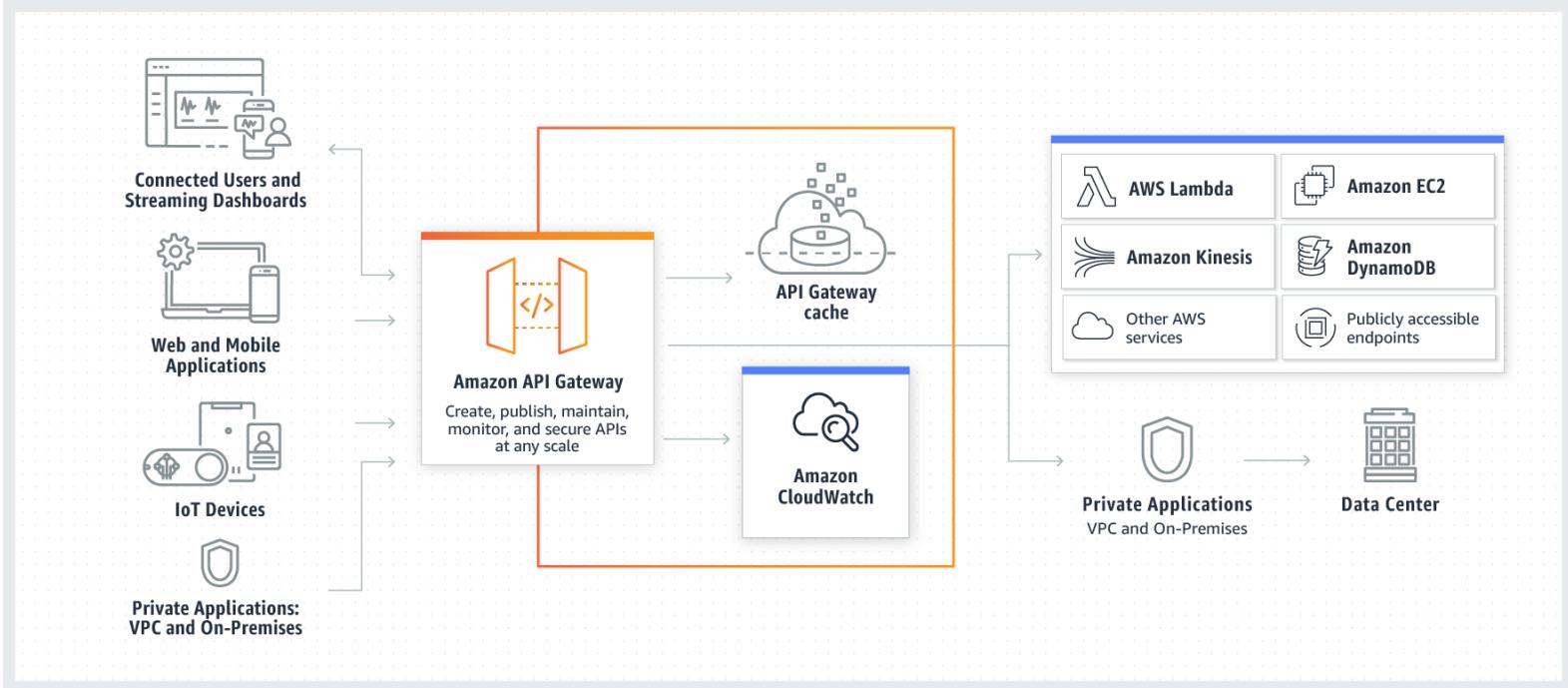
API Gateway creates RESTful APIs that enable stateful client-server communication and API Gateway also creates WebSocket APIs that adhere to the WebSocket protocol, which enables stateful, full-duplex communication between client and server

Explanation

Correct option: **API Gateway creates RESTful APIs that enable stateless client-server communication and API Gateway also creates WebSocket APIs that adhere to the WebSocket protocol, which enables stateful, full-duplex communication between client and server**

Amazon API Gateway is a fully managed service that makes it easy for developers to create, publish, maintain, monitor, and secure APIs at any scale. APIs act as the front door for applications to access data, business logic, or functionality from your backend services. Using API Gateway, you can create RESTful APIs and WebSocket APIs that enable real-time two-way communication applications.

How API Gateway



Works:

- <https://aws.amazon.com/api-gateway/>

API Gateway creates RESTful APIs that: Are HTTP-based. Enable stateless client-server communication. Implement standard HTTP methods such as GET, POST, PUT, PATCH, and DELETE.

API Gateway creates WebSocket APIs that: Adhere to the WebSocket protocol, which enables stateful, full-duplex communication between client and server. Route incoming messages based on message content.

So API Gateway supports stateless RESTful APIs as well as stateful WebSocket APIs. Therefore this option is correct.

Incorrect options: **API Gateway creates RESTful APIs that enable stateful client-server communication and API Gateway also creates WebSocket APIs that adhere to the WebSocket protocol, which enables stateful, full-duplex communication between client and server**

API Gateway creates RESTful APIs that enable stateless client-server communication and API Gateway also creates WebSocket APIs that adhere to the WebSocket protocol, which enables stateless, full-duplex communication between client and server

API Gateway creates RESTful APIs that enable stateful client-server communication and API Gateway also creates WebSocket APIs that adhere to the WebSocket protocol, which enables stateless, full-duplex communication between client and server

These three options contradict the earlier details provided in the explanation. To summarize, API Gateway supports stateless RESTful APIs and stateful WebSocket APIs. Hence these options are incorrect.

What is Amazon API Gateway?

[PDF](#) | [Kindle](#) | [RSS](#)

Amazon API Gateway is an AWS service for creating, publishing, maintaining, monitoring, and securing REST, HTTP, and WebSocket APIs at any scale. API developers can create APIs that access AWS or other web services, as well as data stored in the [AWS Cloud](#). As an API Gateway API developer, you can create APIs for use in your own client applications. Or you can make your APIs available to third-party app developers. For more information, see [Who uses API Gateway?](#).

API Gateway creates RESTful APIs that:

- Are HTTP-based.
- [Enable stateless client-server communication](#).
- Implement standard HTTP methods such as GET, POST, PUT, PATCH, and DELETE.

For more information about API Gateway REST APIs and HTTP APIs, see [Choosing between HTTP APIs and REST APIs](#), [Working with HTTP APIs](#), [Use API Gateway to create REST APIs](#), and [Creating a REST API in Amazon API Gateway](#).

API Gateway creates WebSocket APIs that:

- [Adhere to the WebSocket protocol, which enables stateful, full-duplex communication between client and server](#).
- Route incoming messages based on message content.

via -

<https://docs.aws.amazon.com/apigateway/latest/developerguide/welcome.html>

Reference:

<https://docs.aws.amazon.com/apigateway/latest/developerguide/welcome.html>

Question 8: **Correct**

A digital media company has hired you as an AWS Certified Solutions Architect Professional to optimize the architecture for its backup solution for applications running on the AWS Cloud. Currently, all of the applications running on AWS use at least two Availability Zones (AZs). The updated backup policy at the company mandates that all nightly backups for its data are durably stored

in at least two geographically distinct Regions for Production and Disaster Recovery (DR) and the backup processes for both Regions must be fully automated. The new backup solution must ensure that the backup is available to be restored immediately for the Production Region and should be restored within 24 hours in the DR Region.

Which of the following represents the MOST cost-effective solution that will address the given use-case?

- Create a backup process to persist all the data to a large Amazon EBS volume attached to the backup server in the Production Region. Run nightly cron jobs to snapshot these volumes and then copy these snapshots to the DR Region
- Create a backup process to persist all the data to Amazon Glacier in the Production Region. Set up cross-Region replication of this data to Amazon Glacier in the DR Region to ensure minimum possible costs in both Regions
- Create a backup process to persist all the data to an S3 bucket A using S3 standard storage class in the Production Region. Set up cross-Region replication of this S3 bucket A to an S3 bucket B using S3 standard-IA storage class in the DR Region and set up a lifecycle policy in the DR Region to immediately move this data to Amazon Glacier
- Create a backup process to persist all the data to an S3 bucket A using S3 standard storage class in the Production Region. Set up cross-Region replication of this S3 bucket A to an S3 bucket B using S3 standard storage class in the DR Region and set up a lifecycle policy in the DR Region to immediately move this data to Amazon Glacier

(Correct)

Explanation

Correct option:

Create a backup process to persist all the data to an S3 bucket A using S3 standard storage class in the Production Region. Set up cross-Region replication of this S3 bucket A to an S3 bucket B using S3 standard storage class in the DR Region and set up a lifecycle policy in the DR Region to immediately move this data to Amazon Glacier

S3 Replication enables automatic, asynchronous copying of objects across Amazon S3 buckets. There are two types of Replications:

Cross-Region replication (CRR) is used to copy objects across Amazon S3 buckets in different AWS Regions.

Same-Region replication (SRR) is used to copy objects across Amazon S3 buckets in the same AWS Region.

When to use CRR

Cross-Region replication can help you do the following:

- **Meet compliance requirements** — Although Amazon S3 stores your data across multiple geographically distant Availability Zones by default, compliance requirements might dictate that you store data at even greater distances. Cross-Region replication allows you to replicate data between distant AWS Regions to satisfy these requirements.
- **Minimize latency** — If your customers are in two geographic locations, you can minimize latency in accessing objects by maintaining object copies in AWS Regions that are geographically closer to your users.
- **Increase operational efficiency** — If you have compute clusters in two different AWS Regions that analyze the same set of objects, you might choose to maintain object copies in those Regions.

When to use SRR

Same-Region replication can help you do the following:

- **Aggregate logs into a single bucket** — If you store logs in multiple buckets or across multiple accounts, you can easily replicate logs into a single, in-Region bucket. This allows for simpler processing of logs in a single location.
- **Configure live replication between production and test accounts** — If you or your customers have production and test accounts that use the same data, you can replicate objects between those multiple accounts, while maintaining object metadata, by implementing SRR rules.
- **Abide by data sovereignty laws** — You might be required to store multiple copies of your data in separate AWS accounts within a certain Region. Same-Region replication can help you automatically replicate critical data when compliance regulations don't allow the data to leave your country.

via -

<https://docs.aws.amazon.com/AmazonS3/latest/dev/replication.html>

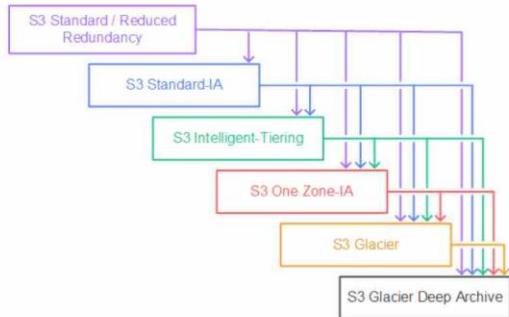
For the given use-case, you can set up cross-Region replication from S3 bucket A using S3 standard storage class in the production Region to S3 bucket B using S3 standard storage class in the DR Region and further create a lifecycle policy to transition this data in bucket B from standard storage class to Amazon Glacier.

Please note the allowed transitions for S3 Lifecycle

Supported transitions and related constraints

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

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



Supported lifecycle transitions

Amazon S3 supports the following lifecycle transitions between storage classes using an S3 Lifecycle configuration.

You *can* transition from the following:

- The S3 Standard storage class to any other storage class.
- Any storage class to the S3 Glacier or S3 Glacier Deep Archive storage classes.
- The S3 Standard-IA storage class to the S3 Intelligent-Tiering or S3 One Zone-IA storage classes.
- The S3 Intelligent-Tiering storage class to the S3 One Zone-IA storage class.
- The S3 Glacier storage class to the S3 Glacier Deep Archive storage class.

Unsupported lifecycle transitions

Amazon S3 does not support any of the following lifecycle transitions.

You *can't* transition from the following:

- Any storage class to the S3 Standard storage class.
- Any storage class to the Reduced Redundancy storage class.
- The S3 Intelligent-Tiering storage class to the S3 Standard-IA storage class.
- The S3 One Zone-IA storage class to the S3 Standard-IA or S3 Intelligent-Tiering storage classes.

Policy:

via -

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

By default, Amazon S3 stores object replicas using the same storage class as the source object. You can also specify a different storage class for the replicas. This allows you to use something like an S3 Standard-IA for the replica bucket, however, S3 standard IA has a minimum storage duration charge of 30 days thereby making it costlier than using S3 Standard storage class for the given scenario because the data would be moved to Glacier via a Lifecycle policy immediately.

Performance across the S3 Storage Classes

	S3 Standard	S3 Intelligent-Tiering*	S3 Standard-IA	S3 One Zone-IA†	S3 Glacier	S3 Glacier Deep Archive
Designed for durability	99.999999999% (11 9's)					
Designed for availability	99.99%	99.9%	99.9%	99.5%	99.99%	99.99%
Availability SLA	99.9%	99%	99%	99%	99.9%	99.9%
Availability Zones	≥3	≥3	≥3	1	≥3	≥3
Minimum capacity charge per object	N/A	N/A	128KB	128KB	40KB	40KB
Minimum storage duration charge	N/A	30 days	30 days	30 days	90 days	180 days
Retrieval fee	N/A	N/A	per GB retrieved	per GB retrieved	per GB retrieved	per GB retrieved
First byte latency	milliseconds	milliseconds	milliseconds	milliseconds	select minutes or hours	select hours
Storage type	Object	Object	Object	Object	Object	Object
Lifecycle transitions	Yes	Yes	Yes	Yes	Yes	Yes

via - <https://aws.amazon.com/s3/storage-classes/>

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

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

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

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

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

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

via -

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

Incorrect options:

Create a backup process to persist all the data to an S3 bucket A using S3 standard storage class in the Production Region. Set up cross-Region replication of this S3 bucket A to an S3 bucket B using S3 standard-IA storage class in the DR Region and set up a lifecycle policy in the DR Region to immediately move this data to Amazon Glacier - As mentioned in the explanation above, S3 standard IA has a minimum storage duration charge of 30 days thereby making it costlier than using S3 Standard storage class for the given scenario, so this option is incorrect.

Create a backup process to persist all the data to Amazon Glacier in the Production Region. Set up cross-Region replication of this data to Amazon Glacier in the DR Region to ensure minimum possible costs in both Regions - One of the key requirements of the given scenario is to ensure that the backup is available to be restored immediately for the Production Region. However, Glacier has a first byte latency of minutes to hours while restoring data, hence this option is not correct for the given use-case.

Create a backup process to persist all the data to a large Amazon EBS volume attached to the backup server in the Production Region. Run nightly cron jobs to snapshot these volumes and then copy these snapshots to the DR Region - One of the key requirements of the given scenario is to ensure that the backup is durable but the data in an EBS volume is only replicated within its Availability Zone so it is not highly durable. However, the EBS snapshots are stored on S3 which are durable. The issue with this option is that it introduces additional cost of an EBS volume and also does not optimize the storage cost in the DR Region as it does not leverage Glacier for the backup data storage.

References:

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

<https://docs.aws.amazon.com/AmazonS3/latest/dev/replication.html>

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

Question 9: **Incorrect**

The DevOps team at a leading social media company uses Chef to automate the configurations of servers in the on-premises data center. The CTO at the company now wants to migrate the IT infrastructure to AWS Cloud with minimal changes to the server configuration workflows and at the same time account for less operational overhead post-migration to AWS. The company has hired you as an AWS Certified Solutions Architect Professional to recommend a solution for this migration.

Which of the following solutions would you recommend to address the given use-case?

-

Rehost the IT infrastructure to AWS Cloud by leveraging AWS Elastic Beanstalk as a configuration management service to automate the configurations of servers on AWS



Rehost the IT infrastructure to AWS Cloud by leveraging AWS OpsWorks as a configuration management service to automate the configurations of servers on AWS

(Incorrect)



Replatform the IT infrastructure to AWS Cloud by leveraging AWS Config as a configuration management service to automate the configurations of servers on AWS



Replatform the IT infrastructure to AWS Cloud by leveraging AWS OpsWorks as a configuration management service to automate the configurations of servers on AWS

(Correct)

Explanation

Correct option:

Replatform the IT infrastructure to AWS Cloud by leveraging AWS OpsWorks as a configuration management service to automate the configurations of servers on AWS

Replatforming is a migration strategy where you don't change the core architecture but leverage some cloud optimizations. Here are the six most common application migration strategies:

via - <https://aws.amazon.com/blogs/enterprise-strategy/6-strategies-for-migrating-applications-to-the-cloud/>

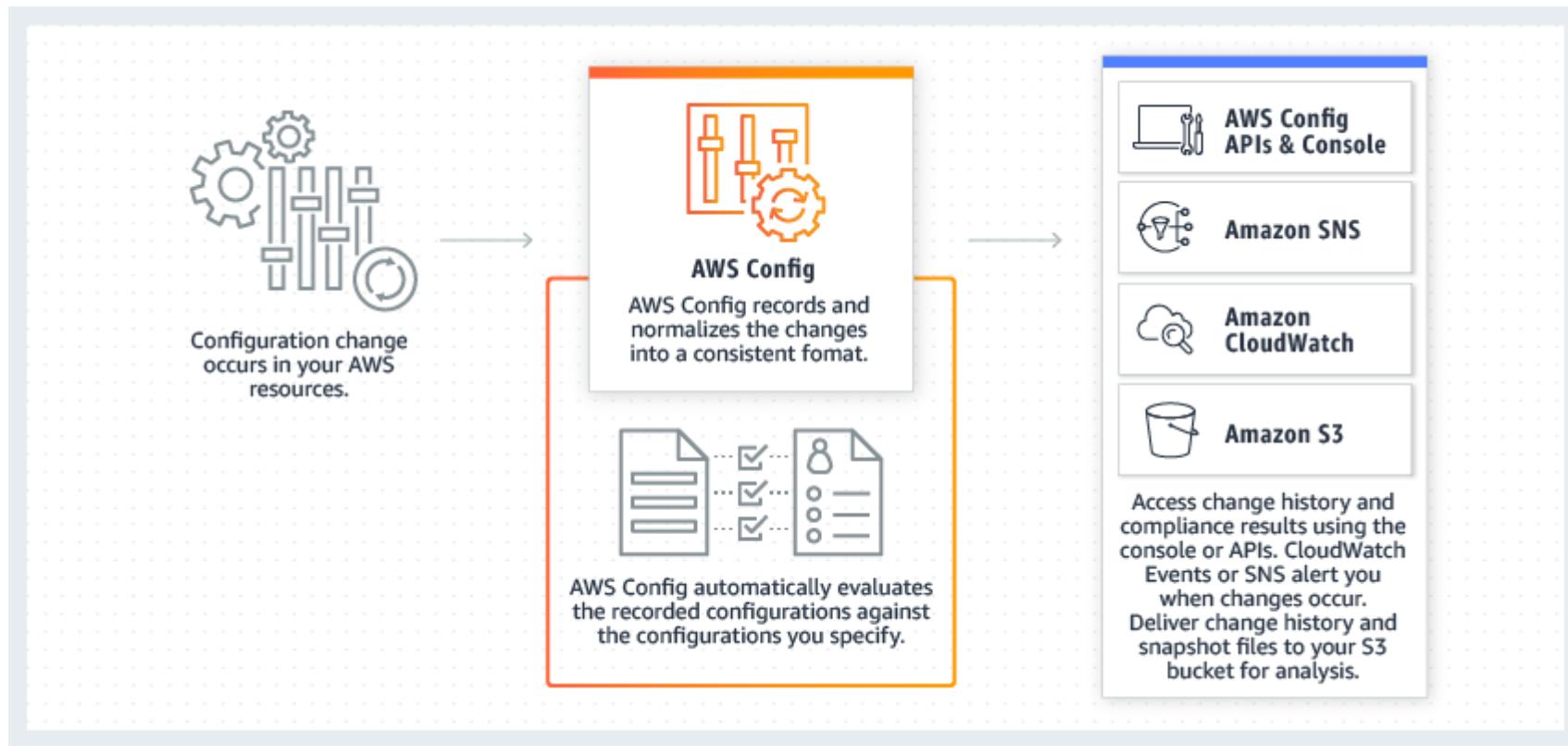
For the given use-case, you can leverage AWS OpsWorks to automate the configurations of your servers. OpsWorks lets you use Chef and Puppet to automate how servers are configured, deployed, and managed across your Amazon EC2 instances or on-premises compute environments. Specifically, OpsWorks for Chef Automate provides a fully managed Chef server and suite of automation tools that give you workflow automation for continuous deployment, automated testing for compliance and security, and a user interface that gives you visibility into your nodes and their status.

The migration results in an optimized IT infrastructure as OpsWorks for Chef Automate provides a configuration management experience that is fully compatible with Chef, including all community scripts and tooling, but without the operational overhead of managing the underlying Chef server.

Incorrect options:

Replatform the IT infrastructure to AWS Cloud by leveraging AWS Config as a configuration management service to automate the configurations of servers on AWS - AWS Config is a service that enables you to assess, audit, and evaluate the configurations of your AWS resources. With Config, you can review changes in configurations and relationships between AWS resources, dive into detailed resource configuration histories, and determine your overall compliance against the configurations specified in your internal guidelines. You can use Config to answer questions such as - "What did my AWS resource look like at xyz point in time?". You cannot use AWS Config as a configuration management service to automate the configurations of servers on AWS.

How Config
Works:



via - <https://aws.amazon.com/config/>

Rehost the IT infrastructure to AWS Cloud by leveraging AWS OpsWorks as a configuration management service to automate the configurations of servers on AWS - This option is incorrect because the correct migration strategy used for this use-case is Replatforming, as described in the explanation above. Rehosting refers to a migration strategy where no cloud optimizations are done and the application is migrated as-is.

Rehost the IT infrastructure to AWS Cloud by leveraging AWS Elastic Beanstalk as a configuration management service to automate the configurations of servers on AWS - As explained above, Rehosting is not the correct migration strategy for the given use-case. Elastic Beanstalk makes it easier for developers to quickly deploy and manage applications in the AWS Cloud. Developers simply upload their application, and Elastic Beanstalk automatically handles the deployment details of capacity provisioning, load

balancing, auto-scaling, and application health monitoring. You cannot use Elastic Beanstalk as a configuration management service to automate the configurations of servers on AWS.

References:

<https://aws.amazon.com/blogs/enterprise-strategy/6-strategies-for-migrating-applications-to-the-cloud/>

<https://aws.amazon.com/opsworks/>

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

<https://aws.amazon.com/elasticbeanstalk/faqs/>

Question 10: **Incorrect**

A Big Data Analytics company has built a custom data warehousing solution for a large airline by using Amazon Redshift. The solution helps the airline to analyze the international and domestic flight reservations, ticket issuing and boarding information, aircraft operation records, and cargo transportation records. As part of the cost optimizations, the airline now wants to move any historical data (any data older than a year) into S3, as the daily analytical reports consume data for just the last one year. However, the analysts at multiple divisions of the airline want to retain the ability to cross-reference this historical data along with the daily reports. The airline wants to develop a solution with the LEAST amount of effort and MINIMUM cost.

As a Solutions Architect Professional, which option would you recommend to address this use-case?

-

Use Redshift Spectrum to create Redshift cluster tables pointing to the underlying historical data in S3. The analytics team can then query this historical data to cross-reference with the daily reports from Redshift

(Correct)

-

Use the Redshift COPY command to load the S3 based historical data into Redshift. Once the ad-hoc queries are run for the historic data, it can be removed from Redshift



Set up access to the historical data via Athena. The analytics team can run historical data queries on Athena and continue the daily reporting on Redshift. In case the reports need to be cross-referenced, the analytics team needs to export these in flat files and then do further analysis



Use Glue ETL job to load the S3 based historical data into Redshift. Once the ad-hoc queries are run for the historic data, it can be removed from Redshift

(Incorrect)

Explanation

Correct option:

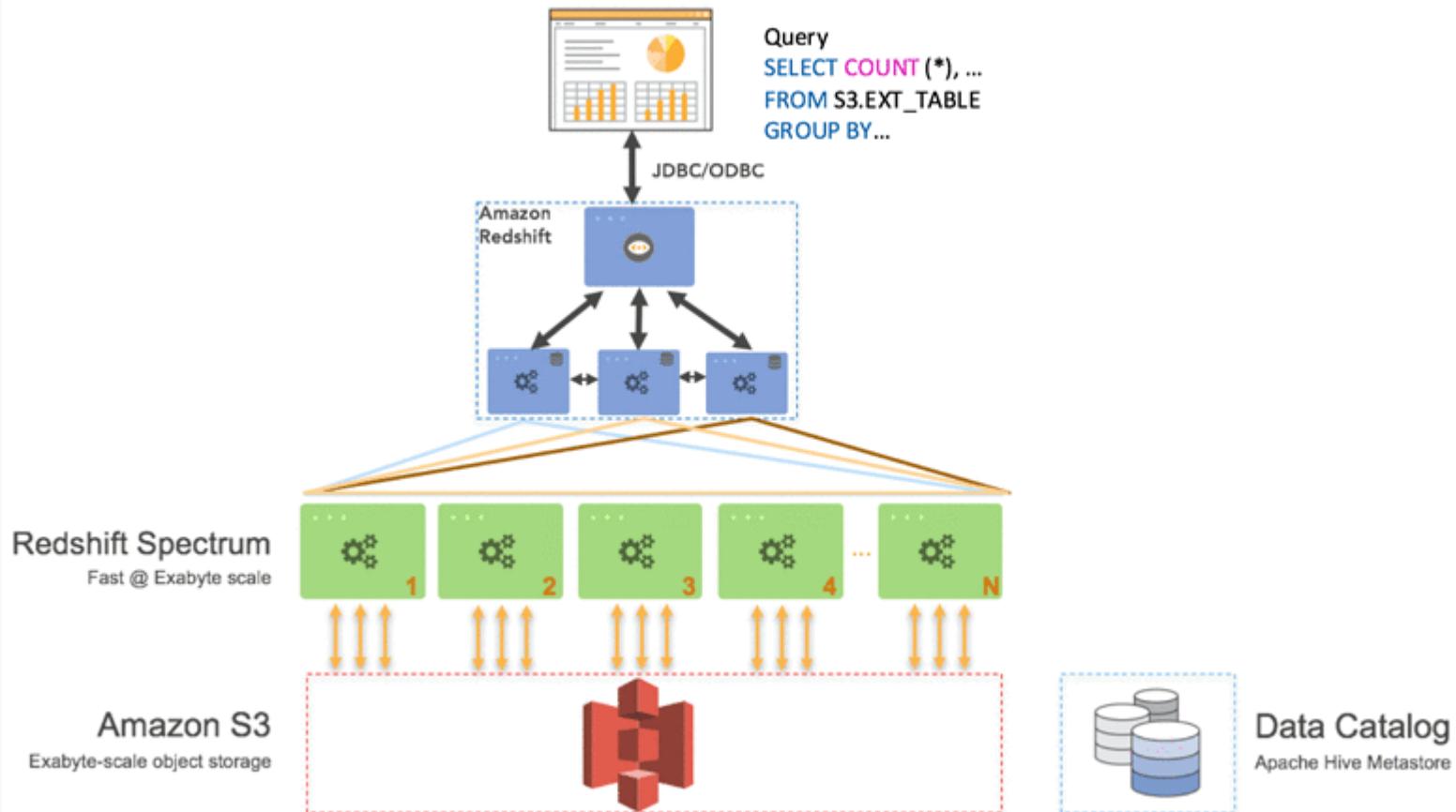
Use Redshift Spectrum to create Redshift cluster tables pointing to the underlying historical data in S3. The analytics team can then query this historical data to cross-reference with the daily reports from Redshift

Amazon Redshift is a fully-managed petabyte-scale cloud-based data warehouse product designed for large scale data set storage and analysis.

Using Amazon Redshift Spectrum, you can efficiently query and retrieve structured and semistructured data from files in Amazon S3 without having to load the data into Amazon Redshift tables.

Amazon Redshift Spectrum resides on dedicated Amazon Redshift servers that are independent of your cluster. Redshift Spectrum pushes many compute-intensive tasks, such as predicate filtering and aggregation, down to the Redshift Spectrum layer. Thus, Redshift Spectrum queries use less of your cluster's processing capacity than other queries.

Redshift Spectrum



Overview

a - <https://aws.amazon.com/blogs/big-data/amazon-redshift-spectrum-extends-data-warehousing-out-to-exabytes-no-loading-required/>

vi

Incorrect options:

Setup access to the historical data via Athena. The analytics team can run historical data queries on Athena and continue the daily reporting on Redshift. In case the reports need to be cross-referenced, the analytics team needs to export these in flat

files and then do further analysis - Amazon Athena is an interactive query service that makes it easy to analyze data directly in Amazon S3 using standard SQL. Athena is serverless, so there is no infrastructure to set up or manage, and customers pay only for the queries they run. You can use Athena to process logs, perform ad-hoc analysis, and run interactive queries.

Providing access to historical data via Athena would mean that historical data reconciliation would become difficult as the daily report would still be produced via Redshift. Such a setup is cumbersome to maintain on a day to day basis. Hence the option to use Athena is ruled out.

Use the Redshift COPY command to load the S3 based historical data into Redshift. Once the ad-hoc queries are run for the historic data, it can be removed from Redshift

Use Glue ETL job to load the S3 based historical data into Redshift. Once the ad-hoc queries are run for the historic data, it can be removed from Redshift

Loading historical data into Redshift via COPY command or via Glue ETL job would be cost-heavy for a one-time ad-hoc process. The same result can be achieved more cost-efficiently by using Redshift Spectrum. Therefore both these options to load historical data into Redshift are also incorrect for the given use-case.

References:

<https://aws.amazon.com/blogs/big-data/amazon-redshift-spectrum-extends-data warehousing-out-to-exabytes-no-loading-required/>

<https://aws.amazon.com/blogs/big-data/10-best-practices-for-amazon-redshift-spectrum/>

Question 11: **Correct**

A leading telecommunications company has developed its cloud storage solution on Amazon RDS for MySQL but it's running into performance issues despite using Read Replicas. The company has hired you as an AWS Certified Solutions Architect Professional to address these performance-related challenges on an urgent basis without moving away from the underlying relational database schema. The company has branch offices across the world, and it needs the solution to work on a global scale.

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

-

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

-

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

(Correct)

-

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

-

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

Explanation

Correct option:

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

Amazon Aurora is a MySQL and PostgreSQL-compatible relational database built for the cloud, that combines the performance and availability of traditional enterprise databases with the simplicity and cost-effectiveness of open source databases. Amazon Aurora features a distributed, fault-tolerant, self-healing storage system that auto-scales up to 64TB per database instance. Aurora is not an in-memory database.

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

Amazon Aurora Global Database

Features:

Features

Sub-Second Data Access in Any Region

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

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

Cross-Region Disaster Recovery

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

via -

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

Incorrect options:

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

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

Global Tables

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

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

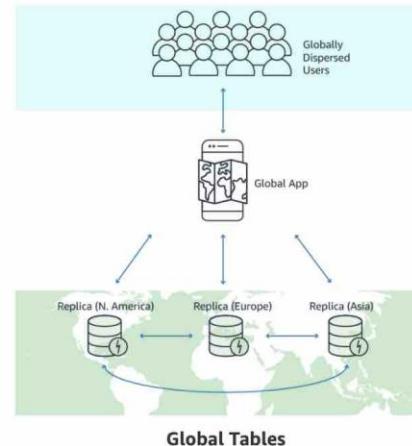
Global Tables eliminates the difficult work of replicating data between regions and resolving update conflicts, enabling you to focus on your application's business logic. In addition, Global Tables enables your applications to stay highly available even in the unlikely event of isolation or degradation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

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

DynamoDB Global Tables Overview:

via

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



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

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

References:

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

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

Question 12: **Correct**

A silicon valley based unicorn startup recently launched a video-sharing social networking service called KitKot. The startup uses AWS Cloud to manage the IT infrastructure. Users upload video files up to 1 GB in size to a single EC2 instance based application server which stores them on a shared EFS file system. Another set of EC2 instances managed via an Auto Scaling group, periodically scans the EFS share directory for new files to process and generate new videos (for thumbnails and composite visual effects) according to the video processing instructions that are uploaded alongside the raw video files. Post-processing, the raw video files are deleted from the EFS file system and the results are stored in an S3 bucket. Links to the processed video files are sent via in-app notifications to the users. The startup has recently found that even as more instances are added to the Auto Scaling Group, many files are processed twice, therefore image processing speed is not improved.

As an AWS Certified Solutions Architect Professional, what would you recommend to improve the reliability of the solution as well as eliminate the redundant processing of video files?

-
-

Create an hourly cron job on the application server to synchronize the contents of the EFS share with S3. Trigger a Lambda function every time a file is uploaded to S3 and process the video file to store the results in another S3 bucket. Once the file is processed, leverage EventBridge events to trigger an SNS notification to send an in-app notification to the user containing the links to the processed files

-
-

Refactor the application to run from S3 instead of EFS and upload the video files directly to an S3 bucket. Set up an EventBridge event to trigger a Lambda function on each file upload that puts a message in an SQS queue containing the link and the video processing instructions. Change the video processing application to read from SQS queue for new files and configure the queue depth metric to scale instances in the video processing Auto Scaling group. Leverage EventBridge events to trigger an SNS notification to the user containing the links to the processed files



Refactor the application to run from Amazon S3 instead of the EFS file system and upload the video files directly to an S3 bucket via an API Gateway based REST API. Configure an S3 trigger to invoke a Lambda function each time a file is uploaded and the Lambda in turn processes the video and stores the processed files in another bucket. Leverage EventBridge events to trigger an SNS notification to send an in-app notification to the user containing the links to the processed files



Refactor the application to run from S3 instead of EFS and upload the video files directly to an S3 bucket. Configure an S3 trigger to invoke a Lambda function on each video file upload to S3 that puts a message in an SQS queue containing the link and the video processing instructions. Change the video processing application to read from the SQS queue and the S3 bucket. Configure the queue depth metric to scale the size of the Auto Scaling group for video processing instances. Leverage EventBridge events to trigger an SNS notification to the user containing the links to the processed files

(Correct)

Explanation

Correct option:

Refactor the application to run from S3 instead of EFS and upload the video files directly to an S3 bucket. Configure an S3 trigger to invoke a Lambda function on each video file upload to S3 that puts a message in an SQS queue containing the link and the video processing instructions. Change the video processing application to read from the SQS queue and the S3 bucket. Configure the queue depth metric to scale the size of the Auto Scaling group for video processing instances. Leverage EventBridge events to trigger an SNS notification to the user containing the links to the processed files

For the given use-case, the primary way to address the issues related to reliability, as well as redundant processing of video files, is by introducing SQS into the solution stack. SQS offers a secure, durable, and available hosted queue that lets you integrate and decouple distributed software systems and components. SQS locks your messages during processing, so that multiple producers can send and multiple consumers can receive messages at the same time. Using the right combination of delay queues and visibility timeout, you can optimize the solution to address use-cases where the consumer application needs additional time to process messages such as the one in this scenario. Messages are put into the SQS queue via a Lambda function that is triggered when a new video file is uploaded to S3 for processing.

Amazon SQS delay queues

[PDF](#) | [Kindle](#) | [RSS](#)

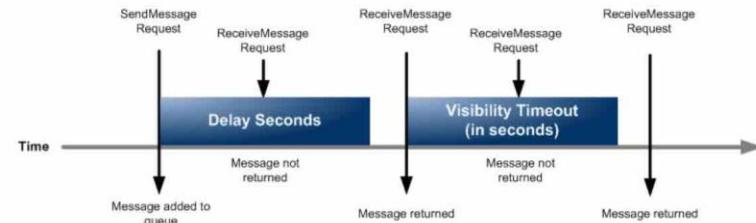
Delay queues let you postpone the delivery of new messages to a queue for a number of seconds, for example, [when your consumer application needs additional time to process messages](#). If you create a delay queue, any messages that you send to the queue remain invisible to consumers for the duration of the delay period. The default (minimum) delay for a queue is 0 seconds. The maximum is 15 minutes. For information about configuring delay queues using the console see [Configuring queue parameters \(console\)](#).

Note

For standard queues, the per-queue delay setting is *not retroactive*—changing the setting doesn't affect the delay of messages already in the queue.

For FIFO queues, the per-queue delay setting is *retroactive*—changing the setting affects the delay of messages already in the queue.

Delay queues are similar to [visibility timeouts](#) because both features make messages unavailable to consumers for a specific period of time. The difference between the two is that, for delay queues, a message is hidden *when it is first added to queue*, whereas for visibility timeouts a message is hidden *only after it is consumed from the queue*. The following diagram illustrates the relationship between delay queues and visibility timeouts.



To set delay seconds on *individual messages*, rather than on an entire queue, use [message timers](#) to allow Amazon SQS to use the message timer's `DelaySeconds` value instead of the delay queue's `DelaySeconds` value.

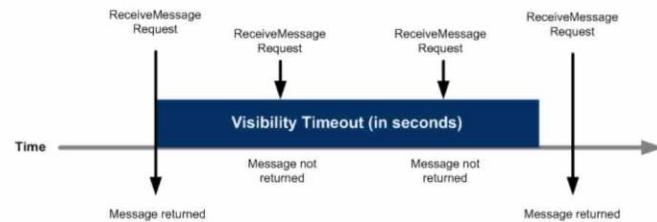
via -

<https://docs.aws.amazon.com/AWSSimpleQueueService/latest/SQSDriverGuide/sqs-delay-queues.html>

Amazon SQS visibility timeout

[PDF](#) | [Kindle](#) | [RSS](#)

When a consumer receives and processes a message from a queue, the message remains in the queue. Amazon SQS doesn't automatically delete the message. Because Amazon SQS is a distributed system, there's no guarantee that the consumer actually receives the message (for example, due to a connectivity issue, or due to an issue in the consumer application). Thus, the consumer must delete the message from the queue after receiving and processing it.



Immediately after a message is received, it remains in the queue. To prevent other consumers from processing the message again, Amazon SQS sets a *visibility timeout*, a period of time during which Amazon SQS prevents other consumers from receiving and processing the message. The default visibility timeout for a message is 30 seconds. The minimum is 0 seconds. The maximum is 12 hours. For information about configuring visibility timeout for a queue using the console, see [Configuring queue parameters \(console\)](#).

Note

For standard queues, the visibility timeout isn't a guarantee against receiving a message twice. For more information, see [At-least-once delivery](#).

FIFO queues allow the producer or consumer to attempt multiple retries:

- If the producer detects a failed SendMessage action, it can retry sending as many times as necessary, using the same message deduplication ID. Assuming that the producer receives at least one acknowledgement before the deduplication interval expires, multiple retries neither affect the ordering of messages nor introduce duplicates.
- If the consumer detects a failed ReceiveMessage action, it can retry as many times as necessary, using the same receive request attempt ID. Assuming that the consumer receives at least one acknowledgement before the visibility timeout expires, multiple retries don't affect the ordering of messages.
- When you receive a message with a message group ID, no more messages for the same message group ID are returned unless you delete the message or it becomes visible.

via -

<https://docs.aws.amazon.com/AWSSimpleQueueService/latest/SQSDeveloperGuide/sqs-visibility-timeout.html>

To ensure that the consumer applications running on the video processing instances can scale via an Auto Scaling group, you could use the SQS queue depth (known as the CloudWatch Amazon SQS metric - ApproximateNumberOfMessages) as the underlying metric. However, the issue with using a CloudWatch Amazon SQS metric like ApproximateNumberOfMessagesVisible for target tracking is that the number of messages in the queue might not change proportionally to the size of the Auto Scaling group that

processes messages from the queue. An optimized solution would be to use a backlog per instance metric with the target value being the acceptable backlog per instance to maintain.

Scaling based on Amazon SQS

[PDF](#) | [Kindle](#) | [RSS](#)

This section shows you how to scale your Auto Scaling group in response to changes in system load in an Amazon Simple Queue Service (Amazon SQS) queue. To learn more about how you can use Amazon SQS, see the [Amazon Simple Queue Service Developer Guide](#).

There are some scenarios where you might think about scaling in response to activity in an Amazon SQS queue. For example, suppose that you have a web app that lets users upload images and use them online. In this scenario, each image requires resizing and encoding before it can be published. The app runs on EC2 instances in an Auto Scaling group, and it's configured to handle your typical upload rates. Unhealthy instances are terminated and replaced to maintain current instance levels at all times. The app places the raw bitmap data of the images in an SQS queue for processing. It processes the images and then publishes the processed images where they can be viewed by users. The architecture for this scenario works well if the number of image uploads doesn't vary over time. But if the number of uploads changes over time, you might consider using dynamic scaling to scale the capacity of your Auto Scaling group.

Using target tracking with the right metric

If you use a target tracking scaling policy based on a custom Amazon SQS queue metric, dynamic scaling can adjust to the demand curve of your application more effectively. For more information about choosing metrics for target tracking, see [Choosing metrics](#).

The issue with using a CloudWatch Amazon SQS metric like `ApproximateNumberOfMessagesVisible` for target tracking is that the number of messages in the queue might not change proportionally to the size of the Auto Scaling group that processes messages from the queue. That's because the number of messages in your SQS queue does not solely define the number of instances needed. The number of instances in your Auto Scaling group can be driven by multiple factors, including how long it takes to process a message and the acceptable amount of latency (queue delay).

The solution is to use a *backlog per instance* metric with the target value being the *acceptable backlog per instance* to maintain. You can calculate these numbers as follows:

- **Backlog per instance:** To calculate your backlog per instance, start with the `ApproximateNumberOfMessages` queue attribute to determine the length of the SQS queue (number of messages available for retrieval from the queue). Divide that number by the fleet's running capacity, which for an Auto Scaling group is the number of instances in the `InService` state, to get the backlog per instance.
- **Acceptable backlog per instance:** To calculate your target value, first determine what your application can accept in terms of latency. Then, take the acceptable latency value and divide it by the average time that an EC2 instance takes to process a message.

To illustrate with an example, let's say that the current `ApproximateNumberOfMessages` is 1500 and the fleet's running capacity is 10. If the average processing time is 0.1 seconds for each message and the longest acceptable latency is 10 seconds, then the acceptable backlog per instance is $10 / 0.1$, which equals 100. This means that 100 is the target value for your target tracking policy. If the backlog per instance is currently at 150 ($1500 / 10$), your fleet scales out, and it scales out by five instances to maintain proportion to the target value.

via -

<https://docs.aws.amazon.com/autoscaling/ec2/userguide/as-using-sqs-queue.html>

Incorrect options:

Refactor the application to run from Amazon S3 instead of the EFS file system and upload the video files directly to an S3 bucket via an API Gateway based REST API. Configure an S3 trigger to invoke a Lambda function each time a file is uploaded and the Lambda, in turn, processes the video and stores the processed files in another bucket. Leverage EventBridge events to trigger an SNS notification to send an in-app notification to the user containing the links to the processed files - API Gateway supports payload size of only up to 10 MB therefore this option is incorrect for the given use-case since you need to support file sizes of up to 1GB for video processing.

API Gateway quotas for configuring and running a REST API

The following quotas apply to configuring and running a REST API in Amazon API Gateway.

Resource or operation	Default quota	Can be increased
Length, in characters, of the key in a stage variable	64	No
Length, in characters, of the value in a stage variable	512	No
Usage plans per account per Region	300	Yes
Usage plans per API key	10	Yes
VPC links per account per Region	20	Yes
API caching TTL	300 seconds by default and configurable between 0 and 3600 by an API owner.	Not for the upper bound (3600)
Cached response size	1048576 Bytes. Cache data encryption may increase the size of the item that is being cached.	No
Integration timeout	50 milliseconds - 29 seconds for all integration types, including Lambda, Lambda proxy, HTTP, HTTP proxy, and AWS integrations.	Not for the lower or upper bounds.
Total combined size of all header values	10240 Bytes	No
Payload size	10 MB	No
Tags per stage	50	No
Number of iterations in a #foreach ... #end loop in mapping templates	1000	No
ARN length of a method with authorization	1600 bytes	No

via -

<https://docs.aws.amazon.com/apigateway/latest/developerguide/limits.html>

Refactor the application to run from S3 instead of EFS and upload the video files directly to an S3 bucket. Set up an EventBridge event to trigger a Lambda function on each file upload that puts a message in an SQS queue containing the link and the video processing instructions. Change the video processing application to read from SQS queue for new files and

configure the queue depth metric to scale instances in the video processing Auto Scaling group. Leverage EventBridge events to trigger an SNS notification to the user containing the links to the processed files - You can certainly configure an EventBridge event to handle a new object upload on S3, which in turn triggers a lambda function. However, this is a roundabout way of propagating the object upload event to the Lambda function. So this is not the best fit option.

Create an hourly cron job on the application server to synchronize the contents of the EFS share with S3. Trigger a Lambda function every time a file is uploaded to S3 and process the video file to store the results in another S3 bucket. Once the file is processed, leverage EventBridge events to trigger an SNS notification to send an in-app notification to the user containing the links to the processed files - The issue with this option is lack of reliability. In case the Lambda function (which is triggered when a video file is uploaded to S3) fails to process a given video file, then the source video file would always remain unprocessed as there is no queue-based mechanism to re-process failed events. So this option is incorrect.

References:

<https://docs.aws.amazon.com/AWSSimpleQueueService/latest/SQSDeveloperGuide/sqs-delay-queues.html>

<https://docs.aws.amazon.com/AWSSimpleQueueService/latest/SQSDeveloperGuide/sqs-visibility-timeout.html>

<https://docs.aws.amazon.com/autoscaling/ec2/userguide/as-using-sqs-queue.html>

<https://docs.aws.amazon.com/apigateway/latest/developerguide/limits.html>

Question 13: **Incorrect**

The CTO at a multi-national retail company is pursuing an IT re-engineering effort to set up a hybrid network architecture that would facilitate the company's envisaged long-term data center migration from multiple on-premises data centers to the AWS Cloud. The current on-premises data centers are in different locations and are inter-linked via a private fiber. Due to the unique constraints of the existing legacy applications, using NAT is not an option. During the migration period, many critical applications will need access to other applications deployed in both the on-premises data centers and AWS Cloud.

As a Solutions Architect Professional, which of the following options would you suggest to set up a hybrid network architecture that is highly available and supports high bandwidth for a multi-Region deployment post-migration?

- - Set up multiple hardware VPN connections between AWS cloud and the on-premises data centers. Configure each subnet's traffic through different VPN connections for redundancy. Make sure that no VPC CIDR blocks overlap one another or the on-premises network
 - - Set up multiple software VPN connections between AWS cloud and the on-premises data centers. Configure each subnet's traffic through different VPN connections for redundancy. Make sure that no VPC CIDR blocks overlap one another or the on-premises network
 - - Set up a Direct Connect as primary connection for all on-premises data centers with another VPN as backup. Configure both connections to use the same virtual private gateway and BGP. Make sure that no VPC CIDR blocks overlap one another or the on-premises network
- (Incorrect)
- - Set up a Direct Connect to each on-premises data center from different service providers and configure routing to failover to the other on-premises data center's Direct Connect in case one connection fails. Make sure that no VPC CIDR blocks overlap one another or the on-premises network
- (Correct)
- Explanation**
Correct option:
- Set up a Direct Connect to each on-premises data center from different service providers and configure routing to failover to the other on-premises data center's Direct Connect in case one connection fails. Make sure that no VPC CIDR blocks overlap one another or the on-premises network**

AWS Direct Connect links your on-premises data center to an AWS Direct Connect location over a standard Ethernet fiber-optic cable. One end of the cable is connected to your router, the other to an AWS Direct Connect router. With this connection, you can create virtual interfaces directly to public AWS services (for example, to Amazon S3) or to Amazon VPC, bypassing internet service providers in your network path. An AWS Direct Connect location provides access to AWS in the Region with which it is associated.

There are two types of Direct Connect connections:

Dedicated Connection: A physical Ethernet connection associated with a single customer. Customers can request a dedicated connection through the AWS Direct Connect console, the CLI, or the API. This supports speed of 1Gbps and 10Gbps.

Hosted Connection: A physical Ethernet connection that an AWS Direct Connect Partner provisions on behalf of a customer. Customers request a hosted connection by contacting a partner in the AWS Direct Connect Partner Program, who provisions the connection. This supports speed of 50Mbps, 100Mbps, 200Mbps, 300Mbps, 400Mbps, 500Mbps, 1Gbps, 2Gbps, 5Gbps, and 10Gbps.

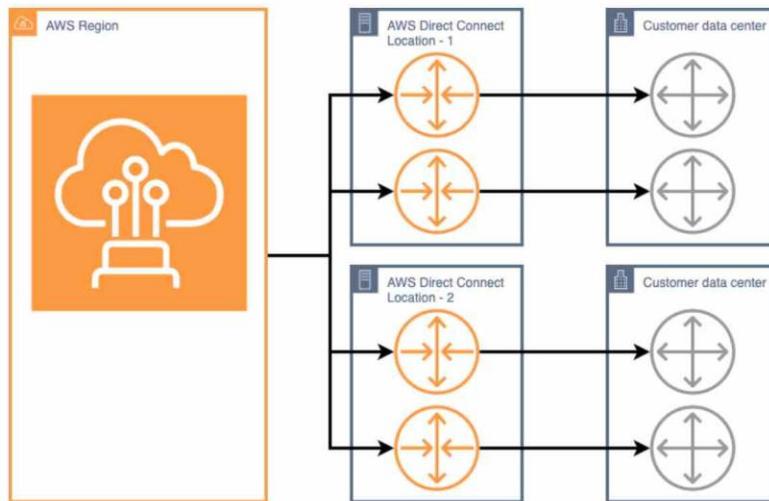
via - <https://docs.aws.amazon.com/directconnect/latest/UserGuide/Welcome.html>

As the use-case requires a hybrid network architecture that is highly available and supports high bandwidth, therefore you should configure the Direct Connect based hybrid network to achieve maximum resiliency for critical workloads by using separate connections from different service providers that terminate on separate devices in more than one location.

Maximum resiliency

[PDF](#) | [Kindle](#) | [RSS](#)

You can achieve maximum resiliency for critical workloads by using separate connections that terminate on separate devices in more than one location (as shown in the following figure). This model provides resiliency against device, connectivity, and complete location failures. The following figure shows both connections from each customer data center going to the same AWS Direct Connect locations. You can optionally have each connection from a customer data center going to different Direct Connect locations.



via -

https://docs.aws.amazon.com/directconnect/latest/UserGuide/maximum_resiliency.html

To configure a high resiliency model

1. Open the AWS Direct Connect console at <https://console.aws.amazon.com/directconnect/v2/home>.
2. On the **AWS Direct Connect** screen, under **Get started**, choose **Create a connection**.
3. Under **Connection ordering type**, choose **Connection wizard**.
4. Under **Resiliency level**, choose **High Resiliency**, and then choose **Next**.
5. On the **Configure connections** pane, under **Connection settings**, do the following:
 - a. For **bandwidth**, choose the connection bandwidth.

This bandwidth applies to all of the created connections.
 - b. For **First location service provider**, select the appropriate AWS Direct Connect location.
 - c. If applicable, for **First Sub location**, choose the floor closest to you or your network provider. This option is only available if the location has meet-me rooms (MMRs) on multiple floors of the building.
 - d. If you selected **Other** for **First location service provider**, for **Name of other provider**, enter the name of the partner that you use.
 - e. For **Second location service provider**, select the appropriate AWS Direct Connect location.
 - f. If applicable, for **Second Sub location**, choose the floor closest to you or your network provider. This option is only available if the location has meet-me rooms (MMRs) on multiple floors of the building.
 - g. If you selected **Other** for **Second location service provider**, for **Name of other provider**, enter the name of the partner that you use.
 - h. (Optional) Add or remove a tag.

[Add a tag] Choose **Add tag** and do the following:
 - For **Key**, enter the key name.
 - For **Value**, enter the key value.

[Remove a tag] Next to the tag, choose **Remove tag**.
6. Choose **Next**.
7. Review your connections, and then choose **Continue**.

If your LOAs are ready, you can choose **Download LOA**, and then click **Continue**.

It can take up to 72 hours for AWS to review your request and provision a port for your connection. During this time, you might receive an email with a request for more information about your use case or the specified location. The email is sent to the email address that you used when you signed up for AWS. You must respond within 7 days or the connection is deleted.

via -

https://docs.aws.amazon.com/directconnect/latest/UserGuide/high_resiliency.html

Incorrect options:

Set up multiple hardware VPN connections between AWS cloud and the on-premises data centers. Configure each subnet's traffic through different VPN connections for redundancy. Make sure that no VPC CIDR blocks overlap one another or the on-premises network

Set up multiple software VPN connections between AWS cloud and the on-premises data centers. Configure each subnet's traffic through different VPN connections for redundancy. Make sure that no VPC CIDR blocks overlap one another or the on-premises network

A VPN connection refers to the connection between your VPC and your own on-premises network. Site-to-Site VPN supports Internet Protocol security (IPsec) VPN connections. VPNs on AWS come in three flavours: hardware only, software only and a mix of hardware/software. The hardware only VPN uses a hardware VPN device to connect the virtual private gateway on the AWS end to a customer VPN gateway on the customers end, via IPsec VPN tunnels.

Hardware only VPNs include both the AWS managed AWS Site-to-Site VPN solution and the AWS VPN CloudHub.

You can also create a VPN connection to your remote network by using an Amazon EC2 instance in your VPC that's running a third party software VPN appliance.

The limitation with both options is that VPNs do not support high bandwidth data transfer as these operate over the public internet infrastructure. VPN Connections are a good solution if you have an immediate need, and have low to modest bandwidth requirements.

VPN connections

[PDF](#) | [Kindle](#) | [RSS](#)

You can connect your Amazon VPC to remote networks and users using the following VPN connectivity options.

VPN connectivity option	Description
AWS Site-to-Site VPN	You can create an IPsec VPN connection between your VPC and your remote network. On the AWS side of the Site-to-Site VPN connection, a virtual private gateway or transit gateway provides two VPN endpoints (tunnels) for automatic failover. You configure your <i>customer gateway device</i> on the remote side of the Site-to-Site VPN connection. For more information, see the AWS Site-to-Site VPN User Guide .
AWS Client VPN	AWS Client VPN is a managed client-based VPN service that enables you to securely access your AWS resources or your on-premises network. With AWS Client VPN, you configure an endpoint to which your users can connect to establish a secure TLS VPN session. This enables clients to access resources in AWS or an on-premises from any location using an OpenVPN-based VPN client. For more information, see the AWS Client VPN Administrator Guide .
AWS VPN CloudHub	If you have more than one remote network (for example, multiple branch offices), you can create multiple AWS Site-to-Site VPN connections via your virtual private gateway to enable communication between these networks. For more information, see Providing secure communication between sites using VPN CloudHub in the AWS Site-to-Site VPN User Guide .
Third party software VPN appliance	You can create a VPN connection to your remote network by using an Amazon EC2 instance in your VPC that's running a third party software VPN appliance. AWS does not provide or maintain third party software VPN appliances; however, you can choose from a range of products provided by partners and open source communities. Find third party software VPN appliances on the AWS Marketplace .

via -

<https://docs.aws.amazon.com/vpc/latest/userguide/vpn-connections.html>

Set up a Direct Connect as primary connection for all on-premises data centers with another VPN as backup. Configure both connections to use the same virtual private gateway and BGP. Make sure that no VPC CIDR blocks overlap one another or the on-premises network - This option has been added as a distractor as you cannot have just one Direct Connect connection for multiple on-premises data centers that are in different locations. Also having a VPN as a backup does not provide a high-bandwidth and high-availability fallback option.

References:

<https://docs.aws.amazon.com/directconnect/latest/UserGuide/Welcome.html>

https://docs.aws.amazon.com/directconnect/latest/UserGuide/maximum_resiliency.html

<https://docs.aws.amazon.com/vpc/latest/userguide/vpn-connections.html>

https://medium.com/@datapath_io/aws-direct-connect-vs-vpn-vs-direct-connect-gateway-97900cdf7d04

Question 14: **Incorrect**

A company wants to migrate its on-premises Oracle database to Aurora MySQL. The company has hired an AWS Certified Solutions Architect Professional to carry out the migration with minimal downtime using AWS DMS. The company has mandated that the migration must have minimal impact on the performance of the source database and the Solutions Architect must validate that the data was migrated accurately from the source to the target before the cutover.

Which of the following solutions will MOST effectively address this use-case?

- Use AWS Schema Conversion Tool for the migration task so it can compare the source and target data and report any mismatches**
(Incorrect)
- Use the table metrics of the DMS task to verify the statistics for tables being migrated including the DDL statements completed**
- Configure DMS premigration assessment on the migration task so the assessment can compare the source and target data and report any mismatches**
- Configure DMS data validation on the migration task so it can compare the source and target data for the DMS task and report any mismatches**

(Correct)

Explanation

Correct option:

Configure DMS data validation on the migration task so it can compare the source and target data for the DMS task and report any mismatches

You can use AWS DMS data validation to ensure that your data has migrated accurately from the source to the target. DMS compares the source and target records and then reports any mismatches. In addition, for a CDC-enabled task, AWS DMS compares the incremental changes and reports any mismatches. As part of data validation, DMS compares each row in the source with its corresponding row at the target and verifies that those rows contain the same data. For this comparison, DMS issues appropriate queries to retrieve the data. These queries consume additional resources at the source and the target as well as additional network resources.

AWS DMS provides support for data validation, to ensure that your data was migrated accurately from the source to the target. If you enable it for a task, then AWS DMS begins comparing the source and target data immediately after a full load is performed for a table.

Data validation is optional. AWS DMS compares the source and target records, and reports any mismatches. In addition, for a CDC-enabled task, AWS DMS compares the incremental changes and reports any mismatches.

During data validation, AWS DMS compares each row in the source with its corresponding row at the target, and verifies that those rows contain the same data. To accomplish this, AWS DMS issues appropriate queries to retrieve the data. Note that these queries will consume additional resources at the source and the target, as well as additional network resources.

Data validation works with the following databases wherever AWS DMS supports them as source and target endpoints:

- Oracle
- PostgreSQL
- MySQL
- MariaDB
- Microsoft SQL Server
- Amazon Aurora (MySQL)
- Amazon Aurora (PostgreSQL)
- IBM Db2 LUW

For more information about the supported endpoints, see [Working with AWS DMS endpoints](#).

Data validation requires additional time, beyond the amount required for the migration itself. The extra time required depends on how much data was migrated.

Data validation settings include the following:

- `EnableValidation` – Enables or disables data validation.
- `FailureMaxCount` – Specifies the maximum number of records that can fail validation before validation is suspended for the task.
- `HandleCollationDiff` – Accounts for column collation differences in PostgreSQL endpoints when identifying source and target records to compare.
- `RecordFailureDelayLimitInMinutes` – Specifies the delay before reporting any validation failure details.
- `TableFailureMaxCount` – Specifies the maximum number of tables that can fail validation before validation is suspended for the task.
- `ThreadCount` – Adjusts the number of execution threads that AWS DMS uses during validation.
- `ValidationOnly` – Previews the validation for the task without performing any migration or replication of data. To use this option, set the task migration type to `Replicate data changes only` in the AWS DMS console, or set the migration type to `cdc` in the AWS DMS API. In addition, set the target table task setting, `TargetTablePrepMode`, to `DO_NOTHING`.

DMS data validation overview:

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

via -

Incorrect options:

Use the table metrics of the DMS task to verify the statistics for tables being migrated including the DDL statements completed

- You can use table metrics to capture statistics such as insert, update, delete, and DDL statements completed for the tables being migrated. This option will not help you compare the source and target data for the DMS task and report any mismatches.

AWS Database Migration Service metrics

AWS DMS provides statistics for the following:

- **Host Metrics** – Performance and utilization statistics for the replication host, provided by Amazon CloudWatch. For a complete list of the available metrics, see [Replication instance metrics](#).
- **Replication Task Metrics** – Statistics for replication tasks including incoming and committed changes, and latency between the replication host and both the source and target databases. For a complete list of the available metrics, see [Replication task metrics](#).
- **Table Metrics** – Statistics for tables that are in the process of being migrated, including the number of insert, update, delete, and DDL statements completed.

Task metrics are divided into statistics between the replication host and the source endpoint, and statistics between the replication host and the target endpoint. You can determine the total statistic for a task by adding two related statistics together. For example, you can determine the total latency, or replica lag, for a task by combining the **CDCLatencySource** and **CDCLatencyTarget** values.

Task metric values can be influenced by current activity on your source database. For example, if a transaction has begun, but has not been committed, then the **CDCLatencySource** metric continues to grow until that transaction has been committed.

For the replication instance, the **FreeableMemory** metric requires clarification. Freeable memory is not a indication of the actual free memory available. It is the memory that is currently in use that can be freed and used for other uses; it's a combination of buffers and cache in use on the replication instance.

While the **FreeableMemory** metric does not reflect actual free memory available, the combination of the **FreeableMemory** and **SwapUsage** metrics can indicate if the replication instance is overloaded.

Monitor these two metrics for the following conditions:

- The **FreeableMemory** metric approaching zero.
- The **SwapUsage** metric increases or fluctuates.

If you see either of these two conditions, they indicate that you should consider moving to a larger replication instance. You should also consider reducing the number and type of tasks running on the replication instance. Full Load tasks require more memory than tasks that just replicate changes.

via -

https://docs.aws.amazon.com/dms/latest/userguide/CHAP_Monitoring.html#CHAP_Monitoring.Metrics

Configure DMS premigration assessment on the migration task so the assessment can compare the source and target data and report any mismatches - A premigration assessment evaluates specified components of a database migration task to help identify any problems that might prevent a migration task from running as expected. This assessment gives you a chance to identify issues before you run a new or modified task. You can then fix problems before they occur while running the migration task itself. This can avoid delays in completing a given database migration needed to repair data and your database environment. This option will not help you compare the source and target data for the DMS task and report any mismatches.

Use AWS Schema Conversion Tool for the migration task so it can compare the source and target data and report any mismatches - You can use the AWS Schema Conversion Tool (AWS SCT) to convert your existing database schema from one database engine to another. You can convert relational OLTP schema or data warehouse schema. Your converted schema is suitable for an Amazon Relational Database Service (Amazon RDS) MySQL, MariaDB, Oracle, SQL Server, PostgreSQL DB, an Amazon Aurora

DB cluster, or an Amazon Redshift cluster. This option will not help you compare the source and target data for the DMS task and report any mismatches.

References:

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

https://docs.aws.amazon.com/dms/latest/userguide/CHAP_Monitoring.html#CHAP_Monitoring.Metrics

Question 15: **Incorrect**

A healthcare technology solutions company recently faced a security event resulting in an S3 bucket with sensitive data containing Personally Identifiable Information (PII) for patients being made public. The company policy mandates never to have public S3 objects so the Governance and Compliance team must be notified immediately as soon as any public objects are identified. The company has hired you as an AWS Certified Solutions Architect Professional to help build a solution that detects the presence of a public S3 object, which in turn sets off an alarm to trigger notifications and then automatically remediates the said object.

Which of the following solutions would you implement in tandem to meet the requirements of the given use-case? (Select two)



Leverage AWS Trusted Advisor to check for S3 bucket public-read permissions and invoke a Lambda function to send a notification via SNS as soon as a public object is uploaded

(Incorrect)



Enable object-level logging for S3. When a PutObject API call is made with a public-read permission, use S3 event notifications to trigger a Lambda that sends a notification via SNS

(Incorrect)

- **Leverage AWS Access Analyzer to check for S3 bucket public-read permissions and invoke a Lambda function to send a notification via SNS as soon as a public object is uploaded**
- **Configure a Lambda function as one of the SNS topic subscribers, which is invoked to secure the objects in the S3 bucket**
(Correct)
- **Enable object-level logging for S3. Set up a EventBridge event pattern when a PutObject API call with public-read permission is detected in the AWS CloudTrail logs and set the target as an SNS topic for downstream notifications**

(Correct)

Explanation

Correct options:

Configure a Lambda function as one of the SNS topic subscribers, which is invoked to secure the objects in the S3 bucket

Enable object-level logging for S3. Set up a EventBridge event pattern when a PutObject API call with public-read permission is detected in the AWS CloudTrail logs and set the target as an SNS topic for downstream notifications

You can enable object-level logging for an S3 bucket to send logs to CloudTrail for object-level API operations such as GetObject, DeleteObject, and PutObject. These events are called data events. By default, CloudTrail trails don't log data events, but you can configure trails to log data events for S3 buckets that you specify, or to log data events for all the Amazon S3 buckets in your AWS account.

How do I enable object-level logging for an S3 bucket with AWS CloudTrail data events?

[PDF](#) | [Kindle](#) | [RSS](#)

This section describes how to enable an AWS CloudTrail trail to log data events for objects in an S3 bucket by using the Amazon S3 console. CloudTrail supports logging Amazon S3 object-level API operations such as `GetObject`, `DeleteObject`, and `PutObject`. These events are called data events. By default, CloudTrail trails don't log data events, but you can configure trails to log data events for S3 buckets that you specify, or to log data events for all the Amazon S3 buckets in your AWS account. For more information, see [Logging Amazon S3 API Calls Using AWS CloudTrail](#). CloudTrail does not populate data events in the CloudTrail event history. Additionally, not all bucket-level actions are populated in the CloudTrail event history. For more information, see [Using Amazon CloudWatch Logs filter patterns and Amazon Athena to query CloudTrail logs](#).

To configure a trail to log data events for an S3 bucket, you can use either the AWS CloudTrail console or the Amazon S3 console. If you are configuring a trail to log data events for all the Amazon S3 buckets in your AWS account, it's easier to use the CloudTrail console. For information about using the CloudTrail console to configure a trail to log S3 data events, see [Data Events](#) in the *AWS CloudTrail User Guide*.

via -

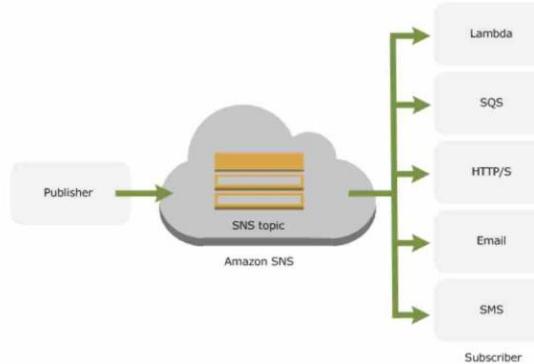
<https://docs.aws.amazon.com/AmazonS3/latest/user-guide/enable-cloudtrail-events.html>

You need to further configure a EventBridge event-pattern based rule to analyze the CloudTrail logs for S3 `PutObject` API call with public-read permissions. The target for this rule can be set as an SNS topic. The SNS would send the notification via an email or SMS as soon as a public object is detected. Moreover, the SNS topic is also subscribed by a Lambda function which runs custom code to secure the objects in the S3 bucket.

What is Amazon SNS?

[PDF](#) | [Kindle](#) | [RSS](#)

Amazon Simple Notification Service (Amazon SNS) is a web service that coordinates and manages the delivery or sending of messages to subscribing endpoints or clients. In Amazon SNS, there are two types of clients—publishers and subscribers—also referred to as producers and consumers. Publishers communicate asynchronously with subscribers by producing and sending a message to a topic, which is a logical access point and communication channel. Subscribers (that is, web servers, email addresses, Amazon SQS queues, AWS Lambda functions) consume or receive the message or notification over one of the supported protocols (that is, Amazon SQS, HTTP/S, email, SMS, Lambda) when they are subscribed to the topic.



When using Amazon SNS, you (as the owner) create a topic and control access to it by defining policies that determine which publishers and subscribers can communicate with the topic. A publisher sends messages to topics that they have created or to topics they have permission to publish to. Instead of including a specific destination address in each message, a publisher sends a message to the topic. Amazon SNS matches the topic to a list of subscribers who have subscribed to that topic, and delivers the message to each of those subscribers. Each topic has a unique name that identifies the Amazon SNS endpoint for publishers to post messages and subscribers to register for notifications. Subscribers receive all messages published to the topics to which they subscribe, and all subscribers to a topic receive the same messages.

via -

<https://docs.aws.amazon.com/sns/latest/dg/welcome.html>

Incorrect options:

Enable object-level logging for S3. When a PutObject API call is made with a public-read permission, use S3 event notifications to trigger a Lambda that sends a notification via SNS - S3 event notification allows you to receive notifications when certain events happen in your bucket. To enable notifications, you must first add a notification configuration that identifies the events you want Amazon S3 to publish and the destinations where you want Amazon S3 to send the notifications. S3 can publish notifications for the new object create events.

You can request notification when only a specific API is used (for example, s3:ObjectCreated:Put), or you can use a wildcard (for example, s3:ObjectCreated:*), however, you cannot check if the API call was made with a public-read permission. So, this option is incorrect.

Overview of notifications

Currently, Amazon S3 can publish notifications for the following events:

- **New object created events** — Amazon S3 supports multiple APIs to create objects. You can request notification when only a specific API is used (for example, s3:ObjectCreated:Put), or you can use a wildcard (for example, s3:ObjectCreated:*) to request notification when an object is created regardless of the API used.
- **Object removal events** — Amazon S3 supports deletes of versioned and unversioned objects. For information about object versioning, see [Object Versioning](#) and [Using versioning](#).
You can request notification when an object is deleted or a versioned object is permanently deleted by using the s3:ObjectRemoved:Delete event type. Or you can request notification when a delete marker is created for a versioned object by using s3:ObjectRemoved:DeleteMarkerCreated. You can also use a wildcard s3:ObjectRemoved:* to request notification anytime an object is deleted. For information about deleting versioned objects, see [Deleting object versions](#).
- **Restore object events** — Amazon S3 supports the restoration of objects archived to the S3 Glacier storage classes. You request to be notified of object restoration completion by using s3:ObjectRestore:Completed. You use s3:ObjectRestore:Post to request notification of the initiation of a restore.
- **Reduced Redundancy Storage (RRS) object lost events** — Amazon S3 sends a notification message when it detects that an object of the RRS storage class has been lost.
- **Replication events** — Amazon S3 sends event notifications for replication configurations that have S3 Replication Time Control (S3 RTC) enabled. It sends these notifications when an object fails replication, when an object exceeds the 15-minute threshold, when an object is replicated after the 15-minute threshold, and when an object is no longer tracked by replication metrics. It publishes a second event when that object replicates to the destination Region.

For a list of supported event types, see [Supported event types](#).

Amazon S3 supports the following destinations where it can publish events:

- **Amazon Simple Notification Service (Amazon SNS) topic**

Amazon SNS is a flexible, fully managed push messaging service. Using this service, you can push messages to mobile devices or distributed services. With SNS you can publish a message once, and deliver it one or more times. For more information about SNS, see the [Amazon SNS](#) product detail page.

- **Amazon Simple Queue Service (Amazon SQS) queue**

Amazon SQS is a scalable and fully managed message queuing service. You can use SQS to transmit any volume of data without requiring other services to be always available. In your notification configuration, you can request that Amazon S3 publish events to an SQS queue.

Currently, Standard SQS queue is only allowed as an Amazon S3 event notification destination, whereas FIFO SQS queue is not allowed. For more information about Amazon SQS, see the [Amazon SQS](#) product detail page.

- **AWS Lambda**

AWS Lambda is a compute service that makes it easy for you to build applications that respond quickly to new information. AWS Lambda runs your code in response to events such as image uploads, in-app activity, website clicks, or outputs from connected devices.

You can use AWS Lambda to extend other AWS services with custom logic, or create your own backend that operates at AWS scale, performance, and security. With AWS Lambda, you can easily create discrete, event-driven applications that run only when needed and scale automatically from a few requests per day to thousands per second.

AWS Lambda can run custom code in response to Amazon S3 bucket events. You upload your custom code to AWS Lambda and create what is called a Lambda function. When Amazon S3 detects an event of a specific type (for example, an object created event), it can publish the event to AWS Lambda and invoke your function in Lambda. In response, AWS Lambda runs your function.

via -

<https://docs.aws.amazon.com/AmazonS3/latest/dev/NotificationHowTo.html>

Leverage AWS Trusted Advisor to check for S3 bucket public-read permissions and invoke a Lambda function to send a notification via SNS as soon as a public object is uploaded - Trusted Advisor is an application that inspects your AWS environment and makes recommendations for saving money, improving system performance, or closing security gaps. The Trusted Advisor notification feature helps you stay up-to-date with your AWS resource deployment. However, you will only be notified by weekly email when you opt-in for this service, so this does not meet the key requirement for the use-case wherein the notification should be sent as soon as a public object is uploaded. Also, Trusted Advisor just checks buckets in Amazon Simple Storage Service (Amazon S3) that have open access permissions. It cannot be used for near real-time detection of a new public object uploaded on S3.

Q: What is AWS Trusted Advisor?

AWS Trusted Advisor is an application that draws upon best practices learned from AWS's aggregated operational history of serving hundreds of thousands of AWS customers. Trusted Advisor inspects your AWS environment and makes recommendations for saving money, improving system performance, or closing security gaps.

Q: How do I access Trusted Advisor?

Trusted Advisor is available in the [AWS Management Console](#). You can access the Trusted Advisor console directly at <https://console.aws.amazon.com/trustedadvisor/>.

Q: What does Trusted Advisor check?

Trusted Advisor includes an expanding list of checks in the categories of cost optimization, security, fault tolerance, performance, and service limits. For a complete list of checks and descriptions, explore [AWS Trusted Advisor Best Practices](#).

Q: How does the Trusted Advisor notification feature work?

The Trusted Advisor notification feature helps you stay up-to-date with your AWS resource deployment. You will be notified by weekly email when you opt in for this service.

- **What is in the notification?** The notification email includes the summary of saving estimates and your check status, especially highlighting changes of check status.
- **How do I sign up for the notification?** This is an opt-in service, so do make sure to set up the notification in your dashboard. You can choose which contacts receive notification on the Preferences pane of the Trusted Advisor console.
- **Who can get this notification?** You can indicate up to three recipients for the weekly status updates and savings estimates.
- **What language will the notification be in?** The notification is available in English and Japanese.
- **How often will I get notified, and when?** You will receive a weekly notification email, typically on Thursday or Friday, and it will reflect your resource configuration over the past week (7 days).
- **Can I unsubscribe from the notifications if I do not want to receive the email anymore?** Yes. You can change the setting in your dashboard by clearing all the check boxes and then selecting "Save Preferences".

via -

<https://aws.amazon.com/premiumsupport/faqs/>

Leverage AWS Access Analyzer to check for S3 bucket public-read permissions and invoke a Lambda function to send a notification via SNS as soon as a public object is uploaded - You can use AWS Access Analyzer to receive findings into the source and level of public or shared access for each public or shared bucket. For example, Access Analyzer for S3 might show that a bucket has read or write access provided through a bucket access control list (ACL), a bucket policy, or an access point policy. It cannot be

used for near real-time detection of a new public object uploaded on S3. Additionally, you cannot invoke a Lambda function from Access Analyzer. The findings for Access Analyzer are available within the AWS Console or they can be downloaded in a CSV report.

Using Access Analyzer for S3

[PDF](#) | [Kindle](#) | [RSS](#)

Access Analyzer for S3 alerts you to S3 buckets that are configured to allow access to anyone on the internet or other AWS accounts, including AWS accounts outside of your organization. For each public or shared bucket, you receive findings into the source and level of public or shared access. For example, Access Analyzer for S3 might show that a bucket has read or write access provided through a bucket access control list (ACL), a bucket policy, or an access point policy. Armed with this knowledge, you can take immediate and precise corrective action to restore your bucket access to what you intended.

When reviewing an at-risk bucket in Access Analyzer for S3, you can block all public access to the bucket with a single click. We recommend that you block all access to your buckets unless you require public access to support a specific use case. Before you block all public access, ensure that your applications will continue to work correctly without public access. For more information, see [Using Amazon S3 Block Public Access](#) in the *Amazon Simple Storage Service Developer Guide*.

You can also drill down into bucket-level permission settings to configure granular levels of access. For specific and verified use cases that require public access, such as static website hosting, public downloads, or cross-account sharing, you can acknowledge and record your intent for the bucket to remain public or shared by archiving the findings for the bucket. You can revisit and modify these bucket configurations at any time. You can also download your findings as a CSV report for auditing purposes.

Access Analyzer for S3 is available at no extra cost on the Amazon S3 console. Access Analyzer for S3 is powered by AWS Identity and Access Management (IAM) Access Analyzer. To use Access Analyzer for S3 in the Amazon S3 console, you must visit the IAM console and enable IAM Access Analyzer on a per-Region basis.

For more information about IAM Access Analyzer, see [What is Access Analyzer?](#) in the *IAM User Guide*. For more information about Access Analyzer for S3, review the following sections.

 **Important**

- Access Analyzer for S3 requires an account-level analyzer. To use Access Analyzer for S3, you must visit IAM Access Analyzer and create an analyzer that has an account as the zone of trust. For more information, see [Enabling Access Analyzer](#) in *IAM User Guide*.
- When a bucket policy or bucket ACL is added or modified, Access Analyzer generates and updates findings based on the change within 30 minutes. Findings related to account level block public access settings may not be generated or updated for up to 6 hours after you change the settings.

via -

<https://docs.aws.amazon.com/AmazonS3/latest/user-guide/access-analyzer.html>

References:

<https://docs.aws.amazon.com/AmazonS3/latest/user-guide/enable-cloudtrail-events.html>

<https://docs.aws.amazon.com/sns/latest/dg/welcome.html>

<https://docs.aws.amazon.com/AmazonS3/latest/dev/NotificationHowTo.html>

<https://aws.amazon.com/premiumsupport/faqs/>

<https://docs.aws.amazon.com/AmazonS3/latest/user-guide/access-analyzer.html>

Question 16: **Incorrect**

A Wall Street based trading firm is modernizing its message queuing system by migrating from self-managed message-oriented middleware systems to Amazon SQS. The firm is using SQS to migrate several trading applications to the cloud to ensure high availability and cost efficiency while simplifying administrative complexity and overhead. The development team at the firm expects a peak rate of about 2,400 messages per second to be processed via SQS. It is important that the messages are processed in the order they are received.

Which of the following options can be used to implement this system in the most cost-effective way?

-

Use Amazon SQS FIFO queue in batch mode of 4 messages per operation to process the messages at the peak rate

-

Use Amazon SQS FIFO queue in batch mode of 12 messages per operation to process the messages at the peak rate

(Incorrect)

-

Use Amazon SQS standard queue to process the messages

-

Use Amazon SQS FIFO queue in batch mode of 8 messages per operation to process the messages at the peak rate

(Correct)

Explanation

Correct option:

Use Amazon SQS FIFO queue in batch mode of 8 messages per operation to process the messages at the peak rate

Amazon Simple Queue Service (SQS) is a fully managed message queuing service that enables you to decouple and scale microservices, distributed systems, and serverless applications. SQS offers two types of message queues - Standard queues vs FIFO queues.

For FIFO queues, the order in which messages are sent and received is strictly preserved (i.e. First-In-First-Out). On the other hand, the standard SQS queues offer best-effort ordering. This means that occasionally, messages might be delivered in an order different from which they were sent.

By default, FIFO queues support up to 300 messages (API calls) per second (300 send, receive, or delete operations per second). When you batch 10 messages per operation (maximum), FIFO queues can support up to 3,000 ($300*10$) messages per second. Therefore, you need to process 8 messages per operation so that the FIFO queue can support up to 2,400 ($300*8$) messages per second, which satisfies the peak rate constraint.

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, for example:

- Ensure that user-entered commands are executed in the right order.
- Display the correct product price by sending price modifications in the right order.
- Prevent a student from enrolling in a course before registering for an account.

FIFO queues also provide exactly-once processing but have a limited number of transactions per second (TPS):

- If you use batching, FIFO queues support up to 3,000 transactions per second, per API method (`SendMessageBatch`, `ReceiveMessage`, or `DeleteMessageBatch`). The 3000 transactions represent 300 API calls, each with a batch of 10 messages. To request a quota increase, submit a support request [↗](#).
- Without batching, FIFO queues support up to 300 API calls per second, per API method (`SendMessage`, `ReceiveMessage`, or `DeleteMessage`).

 **Note**

- Amazon SNS isn't currently compatible with FIFO queues.
- The name of a FIFO queue must end with the `.fifo` suffix. The suffix counts towards the 80-character queue name quota. To determine whether a queue is FIFO, you can check whether the queue name ends with the suffix.

Amazon SQS FIFO queues are available in all Regions where Amazon SQS is available, except in the Asia Pacific (Osaka-Local) Region.

via -

FIFO Queues Overview:

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

Incorrect options:

Use Amazon SQS standard queue to process the messages - As messages need to be processed in order, therefore standard queues are ruled out.

Use Amazon SQS FIFO queue in batch mode of 12 messages per operation to process the messages at the peak rate - This option has been added as a distractor, as SQS FIFO only supports a maximum of 10 messages per operation in batch mode.

Use Amazon SQS FIFO queue in batch mode of 4 messages per operation to process the messages at the peak rate - As mentioned earlier in the explanation, you need to use FIFO queues in batch mode and process 8 messages per operation, so that the FIFO queue can support up to 2,400 messages per second. With 4 messages per operation, you can only support up to 1,200 messages per second.

References:

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

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

<https://aws.amazon.com/sqs/features/>

Question 17: **Incorrect**

A global healthcare company wants to develop a solution called Health Information Systems (HIS) on AWS Cloud that would allow the providers, payers, and government agencies to collaborate, anticipate and navigate the changing healthcare landscape. While pursuing this endeavor, the company would like to decrease its IT operational overhead so it could focus more intently on its core business - healthcare analytics. The solution should help the company eliminate the bottleneck created by manual provisioning of development pipelines while adhering to crucial governance and control requirements. As a means to this end, the company has set up "AWS Organizations" to manage several of these scenarios and would like to use Service Control Policies (SCP) for central control over the maximum available permissions for the various accounts in their organization. This allows the organization to ensure that all accounts stay within the organization's access control guidelines.

As a Solutions Architect Professional, which of the following scenarios would you identify as correct regarding the given use-case?
(Select three)

-

SCPs do not affect service-linked role

(Correct)

-

If a user or role has an IAM permission policy that grants access to an action that is either not allowed or explicitly denied by the applicable SCPs, the user or role can still perform that action

-

SCPs affect all users and roles in attached accounts, including the root user

(Correct)

-

SCPs affect all users and roles in attached accounts, excluding the root user

(Incorrect)

-

If a user or role has an IAM permission policy that grants access to an action that is either not allowed or explicitly denied by the applicable SCPs, the user or role can't perform that action

(Correct)

-

SCPs affect service-linked roles

Explanation

Correct options:

If a user or role has an IAM permission policy that grants access to an action that is either not allowed or explicitly denied by the applicable SCPs, the user or role can't perform that action

SCPs affect all users and roles in attached accounts, including the root user

SCPs do not affect service-linked role

Service control policies (SCPs) are one type of policy that can be used to manage your organization. SCPs offer central control over the maximum available permissions for all accounts in your organization, allowing you to ensure your accounts stay within your organization's access control guidelines.

In SCPs, you can restrict which AWS services, resources, and individual API actions the users and roles in each member account can access. You can also define conditions for when to restrict access to AWS services, resources, and API actions. These restrictions even override the administrators of member accounts in the organization.

Please note the following effects on permissions vis-a-vis the SCPs:

If a user or role has an IAM permission policy that grants access to an action that is either not allowed or explicitly denied by the applicable SCPs, the user or role can't perform that action.

SCPs affect all users and roles in the attached accounts, including the root user.

SCPs do not affect any service-linked role.

Effects on permissions

SCPs are similar to AWS Identity and Access Management (IAM) permission policies and use almost the same syntax. However, an SCP never grants permissions. Instead, SCPs are JSON policies that specify the maximum permissions for the affected accounts. For more information, see [Policy Evaluation Logic](#) in the *IAM User Guide*.

- SCPs affect only IAM users and roles that are managed by accounts that are part of the organization. SCPs don't affect resource-based policies directly. They also don't affect users or roles from accounts outside the organization. For example, consider an Amazon S3 bucket that's owned by account A in an organization. The bucket policy (a resource-based policy) grants access to users from account B outside the organization. Account A has an SCP attached. That SCP doesn't apply to those outside users in account B. The SCP applies only to users that are managed by account A in the organization.
- An SCP restricts permissions for IAM users and roles in member accounts, including the member account's root user. Any account has only those permissions permitted by every parent above it. If a permission is blocked at any level above the account, either implicitly (by not being included in an Allow policy statement) or explicitly (by being included in a Deny policy statement), a user or role in the affected account can't use that permission, even if the account administrator attaches the AdministratorAccess IAM policy with /* permissions to the user.
- SCPs affect only member accounts in the organization. They have no effect on users or roles in the master account.
- Users and roles must still be granted permissions with appropriate IAM permission policies. A user without any IAM permission policies has no access, even if the applicable SCPs allow all services and all actions.
- If a user or role has an IAM permission policy that grants access to an action that is also allowed by the applicable SCPs, the user or role can perform that action.
- If a user or role has an IAM permission policy that grants access to an action that is either not allowed or explicitly denied by the applicable SCPs, the user or role can't perform that action.
- SCPs affect all users and roles in attached accounts, *including the root user*. The only exceptions are those described in [Tasks and entities not restricted by SCPs](#).
- SCPs do not affect any service-linked role. Service-linked roles enable other AWS services to integrate with AWS Organizations and can't be restricted by SCPs.
- When you disable the SCP policy type in a root, all SCPs are automatically detached from all AWS Organizations entities in that root. AWS Organizations entities include organizational units, organizations, and accounts. If you reenable SCPs in a root, that root reverts to only the default FullAWSAccess policy automatically attached to all entities in the root. Any attachments of SCPs to AWS Organizations entities from before SCPs were disabled are lost and aren't automatically recoverable, although you can manually reattach them.
- If both a permissions boundary (an advanced IAM feature) and an SCP are present, then the boundary, the SCP, and the identity-based policy must all allow the action.

via -

https://docs.aws.amazon.com/organizations/latest/userguide/orgs_manage_policies_scp.html

Incorrect options:

If a user or role has an IAM permission policy that grants access to an action that is either not allowed or explicitly denied by the applicable SCPs, the user or role can still perform that action

SCPs affect all users and roles in attached accounts, excluding the root user

SCPs affect service-linked roles

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

Reference:

https://docs.aws.amazon.com/organizations/latest/userguide/orgs_manage_policies_scp.html

Question 18: **Correct**

A project uses two AWS accounts for accessing various AWS services. The engineering team has just configured an Amazon S3 bucket in the first AWS account for writing data from the Amazon Redshift cluster provisioned in the second AWS account. The team has noticed that the files created in the S3 bucket using UNLOAD command from the Redshift cluster are not accessible to the users present in the same AWS account as the S3 bucket.

What could be the reason for this denial of permission for resources belonging to the same AWS account?

- By default, an S3 object is owned by the AWS account that uploaded it. So the S3 bucket owner will not implicitly have access to the objects written by Redshift cluster

(Correct)

- The owner of an S3 bucket has implicit access to all objects in his bucket. Permissions are set on objects after they are completely copied to the target location. Since the owner is unable to access the uploaded files, it is possible that the write operation is still in progress

- When objects are uploaded to S3 bucket from a different AWS account, the S3 bucket owner will get implicit permissions to access these objects. It is an upload error that can be fixed by providing manual access from AWS console



When two different AWS accounts are accessing an S3 bucket, both the accounts need to share the bucket policies, explicitly defining the actions possible for each account. An erroneous policy can lead to such permission failures

Explanation

Correct option:

By default, an S3 object is owned by the AWS account that uploaded it. So the S3 bucket owner will not implicitly have access to the objects written by Redshift cluster - By default, an S3 object is owned by the AWS account that uploaded it. This is true even when the bucket is owned by another account. Because the Amazon Redshift data files from the UNLOAD command were put into your bucket by another account, you (the bucket owner) don't have default permission to access those files.

To get access to the data files, an AWS Identity and Access Management (IAM) role with cross-account permissions must run the UNLOAD command again. Follow these steps to set up the Amazon Redshift cluster with cross-account permissions to the bucket:

1. From the account of the S3 bucket, create an IAM role (Bucket Role) with permissions to the bucket.
2. From the account of the Amazon Redshift cluster, create another IAM role (Cluster Role) with permissions to assume the Bucket Role.
3. Update the Bucket Role to grant bucket access and create a trust relationship with the Cluster Role.
4. From the Amazon Redshift cluster, run the UNLOAD command using the Cluster Role and Bucket Role.

This resolution doesn't apply to Amazon Redshift clusters or S3 buckets that use server-side encryption with AWS Key Management Service (AWS KMS).

Incorrect options:

When objects are uploaded to S3 bucket from a different AWS account, the S3 bucket owner will get implicit permissions to access these objects. It is an upload error that can be fixed by providing manual access from AWS console - By default, an S3 object is owned by the AWS account that uploaded it. So, the bucket owner will not have any default permissions on the objects.

The owner of an S3 bucket has implicit access to all objects in his bucket. Permissions are set on objects after they are completely copied to the target location. Since the owner is unable to access the uploaded files, it is possible that the write operation is still in progress - This is an incorrect statement, given only as a distractor.

When two different AWS accounts are accessing an S3 bucket, both the accounts need to share the bucket policies, explicitly defining the actions possible for each account. An erroneous policy can lead to such permission failures - This is an incorrect statement, given only as a distractor.

Reference:

<https://aws.amazon.com/premiumsupport/knowledge-center/s3-access-denied-redshift-unload/>

Question 19: **Correct**

A leading mobility company wants to use AWS for its connected cab application that would collect sensor data from its electric cab fleet to give drivers dynamically updated map information. The company would like to build its new sensor service by leveraging fully serverless components that are provisioned and managed automatically by AWS. The development team at the company does not want an option that requires the capacity to be manually provisioned, as it does not want to respond manually to changing volumes of sensor data. The company has hired you as an AWS Certified Solutions Architect Professional to provide consultancy for this strategic initiative.

Given these constraints, which of the following solutions would you suggest as the BEST fit to develop this service?



Ingest the sensor data in an Amazon SQS standard queue, which is polled by an application running on an EC2 instance and the data is written into an auto-scaled DynamoDB table for downstream processing



Ingest the sensor data in an Amazon SQS standard queue, which is polled by a Lambda function in batches and the data is written into an auto-scaled DynamoDB table for downstream processing

(Correct)

-

Ingest the sensor data in Kinesis Data Firehose, which directly writes the data into an auto-scaled DynamoDB table for downstream processing

-

Ingest the sensor data in a Kinesis Data Stream, which is polled by an application running on an EC2 instance and the data is written into an auto-scaled DynamoDB table for downstream processing

Explanation

Correct option: **Ingest the sensor data in an Amazon SQS standard queue, which is polled by a Lambda function in batches and the data is written into an auto-scaled DynamoDB table for downstream processing**

AWS Lambda lets you run code without provisioning or managing servers. You pay only for the compute time you consume. Amazon Simple Queue Service (SQS) is a fully managed message queuing service that enables you to decouple and scale microservices, distributed systems, and serverless applications. SQS offers two types of message queues. Standard queues offer maximum throughput, best-effort ordering, and at-least-once delivery. SQS FIFO queues are designed to guarantee that messages are processed exactly once, in the exact order that they are sent.

You can use an AWS Lambda function to process messages in an Amazon Simple Queue Service (Amazon SQS) queue. Lambda event source mappings support standard queues and first-in, first-out (FIFO) queues. With Amazon SQS, you can offload tasks from one component of your application by sending them to a queue and processing them asynchronously.

Scaling and processing

For standard queues, Lambda uses long polling to poll a queue until it becomes active. When messages are available, Lambda reads up to 5 batches and sends them to your function. If messages are still available, Lambda increases the number of processes that are reading batches by up to 60 more instances per minute. The maximum number of batches that can be processed simultaneously by an event source mapping is 1000.

For FIFO queues, Lambda sends messages to your function in the order that it receives them. When you send a message to a FIFO queue, you specify a [message group ID](#). Amazon SQS ensures that messages in the same group are delivered to Lambda in order. Lambda sorts the messages into groups and sends only one batch at a time for a group. If the function returns an error, all retries are attempted on the affected messages before Lambda receives additional messages from the same group.

Your function can scale in concurrency to the number of active message groups. For more information, see [SQS FIFO as an event source](#) on the AWS Compute Blog.

Configuring a queue for use with Lambda

Create an SQS queue to serve as an event source for your Lambda function. Then configure the queue to allow time for your Lambda function to process each batch of events—and for Lambda to retry in response to throttling errors as it scales up.

To allow your function time to process each batch of records, set the source queue's visibility timeout to at least 6 times the [timeout](#) that you configure on your function. The extra time allows for Lambda to retry if your function execution is throttled while your function is processing a previous batch.

If a message fails to be processed multiple times, Amazon SQS can send it to a [dead-letter queue](#). When your function returns an error, Lambda leaves it in the queue. After the visibility timeout occurs, Lambda receives the message again. To send messages to a second queue after a number of receives, configure a dead-letter queue on your source queue.

via -

<https://docs.aws.amazon.com/lambda/latest/dg/with-sqs.html>

AWS manages all ongoing operations and underlying infrastructure needed to provide a highly available and scalable message queuing service. With SQS, there is no upfront cost, no need to acquire, install, and configure messaging software, and no time-consuming build-out and maintenance of supporting infrastructure. SQS queues are dynamically created and scale automatically so you can build and grow applications quickly and efficiently. As there is no need to manually provision the capacity, so this is the correct option.

Incorrect options:

Ingest the sensor data in Kinesis Data Firehose, which directly writes the data into an auto-scaled DynamoDB table for downstream processing

Amazon Kinesis Data Firehose is a fully managed service for delivering real-time streaming data to destinations such as Amazon Simple Storage Service (Amazon S3), Amazon Redshift, Amazon OpenSearch Service, Splunk, and any custom HTTP endpoint or HTTP endpoints owned by supported third-party service providers, including Datadog, Dynatrace, LogicMonitor, MongoDB, New Relic, and Sumo Logic.

Firehose cannot directly write into a DynamoDB table, so this option is incorrect.

Ingest the sensor data in an Amazon SQS standard queue, which is polled by an application running on an EC2 instance and the data is written into an auto-scaled DynamoDB table for downstream processing

Ingest the sensor data in a Kinesis Data Stream, which is polled by an application running on an EC2 instance and the data is written into an auto-scaled DynamoDB table for downstream processing

Using an application on an EC2 instance is ruled out as the company wants to use fully serverless components. So both these options are incorrect.

References: <https://aws.amazon.com/sqs/>

<https://docs.aws.amazon.com/lambda/latest/dg/with-kinesis.html>

<https://docs.aws.amazon.com/lambda/latest/dg/with-sqs.html>

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

Question 20: **Incorrect**

A company has built its serverless solution using Amazon API Gateway REST API and AWS Lambda across multiple AWS Regions configured into a single AWS account. During peak hours, customers began to receive 429 Too Many Requests errors from multiple API methods. While troubleshooting the issue, the team realized that AWS Lambda function(s) have not been invoked for these API methods. Also, the company wants to provide a separate quota for its premium customers to access the APIs.

Which solution will you offer to meet this requirement?

-

The error is the outcome of the company reaching its API Gateway account limit for calls per second, set Lambda-level throttling targets in the API Gateway usage plan, and configure customers to use a particular API method when the client identifier is set

-

The error is the outcome of the company reaching its API Gateway limits for the steady-state requests per second (RPS) across all APIs within an AWS account per Region. These limits can be overwritten by configuring the AWS Regional throttling parameters to a greater value. However, based on the AWS account type, a limit is set to the overwritten throttling values

(Incorrect)

-

The error is the outcome of the company reaching its API Gateway account per-method limit for calls per second, configure API keys as client identifiers using usage plans to define the per-client throttling limits for premium customers

-

The error is the outcome of the company reaching its API Gateway account limit for calls per second, configure API keys as client identifiers using usage plans to define the per-client throttling limits for premium customers

(Correct)

Explanation

Correct options:

The error is the outcome of the company reaching its API Gateway account limit for calls per second, configure API keys as client identifiers using usage plans to define the per-client throttling limits for premium customers

After you create, test, and deploy your APIs, you can use API Gateway usage plans to make them available as product offerings for your customers. You can configure usage plans and API keys to allow customers to access selected APIs, and begin throttling requests to those APIs based on defined limits and quotas. These can be set at the API, or API method level.

Per-client throttling limits are applied to clients that use API keys associated with your usage plan as a client identifier. Note that these limits can't be higher than the per-account limits.

When request submissions exceed the steady-state request rate and burst limits, API Gateway begins to throttle requests. Clients may receive 429 Too Many Requests error responses at this point. Since the error is at API Gateway, the Lambda functions configured are not invoked at all.

How throttling limit settings are applied in API Gateway:

How throttling limit settings are applied in API Gateway

Before you configure throttle and quota settings for your API, it's useful to understand how they are applied by Amazon API Gateway.

Amazon API Gateway provides four basic types of throttling-related settings:

- *AWS throttling limits* are applied across all accounts and clients in a region. These limit settings exist to prevent your API—and your account—from being overwhelmed by too many requests. These limits are set by AWS and can't be changed by a customer.
- Per-account limits are applied to all APIs in an account in a specified Region. The account-level rate limit can be increased upon request - higher limits are possible with APIs that have shorter timeouts and smaller payloads. To request an increase of account-level throttling limits per Region, contact the [AWS Support Center](#). For more information, see [Amazon API Gateway quotas and important notes](#). Note that these limits can't be higher than the AWS throttling limits.
- Per-API, per-stage throttling limits are applied at the API method level for a stage. You can configure the same settings for all methods, or configure different throttle settings for each method. Note that these limits can't be higher than the AWS throttling limits.
- *Per-client throttling limits* are applied to clients that use API keys associated with your usage plan as client identifier. Note that these limits can't be higher than the per-account limits.

API Gateway throttling-related settings are applied in the following order:

1. [Per-client or per-method throttling limits](#) that you set for an API stage in a [usage plan](#)
2. Per-method throttling limits that you set for an API stage.
3. [Account-level throttling per Region](#)
4. AWS Regional throttling

via - <https://docs.aws.amazon.com/apigateway/latest/developerguide/api-gateway-request-throttling.html>

Incorrect options:

"The error is the outcome of the company reaching its API Gateway account per-method limit for calls per second, configure API keys as client identifiers using usage plans to define the per-client throttling limits for premium customers - You should note that a throttling limit sets the target point at which request throttling should start. This can be set at the API or API method level. The use case states that the 429 Too Many Requests errors were received from multiple API methods, so the API Gateway reached its limit at the API level since none of the methods invoked the downstream Lambda function. So this option is incorrect.

The error is the outcome of the company reaching its API Gateway account limit for calls per second, set Lambda-level throttling targets in the API Gateway usage plan, and configure customers to use a particular API method when the client identifier is set - This is incorrect. You cannot define Lambda-level throttling targets in the API Gateway usage plan.

The error is the outcome of the company reaching its API Gateway limits for the steady-state requests per second (RPS) across all APIs within an AWS account per Region. These limits can be overwritten by configuring the AWS Regional throttling parameters to a greater value. However, based on the AWS account type, a limit is set to the overwritten throttling values - Per-account limits are applied to all APIs in an account in a specified Region. The account-level rate limit can be increased upon request - higher limits are possible with APIs that have shorter timeouts and smaller payloads. To request an increase in account-level throttling limits per Region, contact the AWS Support Center. It cannot be done from the AWS account directly.

References:

<https://docs.aws.amazon.com/apigateway/latest/developerguide/api-gateway-api-usage-plans.html>

<https://docs.aws.amazon.com/apigateway/latest/developerguide/api-gateway-request-throttling.html>

Question 21: **Incorrect**

A US-based retailer wants to ensure website availability as the company's traditional infrastructure hasn't been easy to scale. By moving its e-commerce platform to AWS, the company wants to scale with demand and ensure better availability. Last year, the company handled record Black Friday sale orders at a rate of nearly 10,000 orders/hour. The engineering team at the company now

wants to finetune the disaster recovery strategy for its database tier. As an AWS Certified Solutions Architect Professional, you have been asked to implement a disaster recovery strategy for all the Amazon RDS databases that the company owns.

Which of the following points do you need to consider for creating a robust recovery plan? (Select three)

-

Recovery time objective (RTO) represents the number of hours it takes, to return the Amazon RDS database to a working state after a disaster

(Correct)

-

You can share automated Amazon RDS snapshots with up to 20 AWS accounts

-

Automated backups, manual snapshots and Read Replicas are supported across multiple Regions

(Correct)

-

Similar to an Amazon RDS Multi-AZ configuration, failover to a Read Replica is an automated process that requires no manual intervention after initial configurations

(Incorrect)

-

Database snapshots are user-initiated backups of your complete DB instance that serve as full backups. These snapshots can be copied and shared to different Regions and accounts

(Correct)

-

Recovery time objective (RTO), expressed in hours, represents how much data you could lose when a disaster happens

(Incorrect)

Explanation

Correct options:

Recovery time objective (RTO) represents the number of hours it takes, to return the Amazon RDS database to a working state after a disaster - Recovery time objective (RTO) and recovery point objective (RPO) are two key metrics to consider when developing a DR plan. RTO represents how many hours it takes you to return to a working state after a disaster.

More info on RTO and

Understanding RTO and RPO

Recovery time objective (RTO) and recovery point objective (RPO) are two key metrics to consider when developing a DR plan. RTO represents how many hours it takes you to return to a working state after a disaster. RPO, which is also expressed in hours, represents how much data you could lose when a disaster happens. For example, an RPO of 1 hour means that you could lose up to 1 hour's worth of data when a disaster occurs.

Different features of Amazon RDS support different RTOs and RPOs at different cost points:

Feature	RTO	RPO	Cost	Scope
Automated backups	Good	Better	Low	Single Region
Manual snapshots	Better	Good	Medium	Cross-Region
Read replicas	Best	Best	High	Cross-Region

As you can see, automated backups are limited to a single AWS Region while manual snapshots and Read Replicas are supported across multiple Regions.

RPO:

via - <https://aws.amazon.com/blogs/database/implementing-a-disaster-recovery-strategy-with-amazon-rds/>

Automated backups, manual snapshots and Read Replicas are supported across multiple Regions - The automated backup feature of Amazon RDS enables point-in-time recovery for your database instance. Amazon RDS will backup your database and transaction logs and store both for a user-specified retention period. If it's a Multi-AZ configuration, backups occur on the standby to reduce I/O impact on the primary. Amazon RDS supports Cross-Region Automated Backups. Manual snapshots and Read Replicas are also supported across multiple Regions.

Database snapshots are user-initiated backups of your complete DB instance that serve as full backups. These snapshots can be copied and shared to different Regions and accounts - Database snapshots are manual (user-initiated) backups of your complete DB instance that serve as full backups. They're stored in Amazon S3 and are retained until you explicitly delete them. These snapshots can be copied and shared to different Regions and accounts. Because DB snapshots include the entire DB instance, including data files and temporary files, the size of the instance affects the amount of time it takes to create the snapshot.

Incorrect options:

Recovery time objective (RTO), expressed in hours, represents how much data you could lose when a disaster happens - RTO represents how many hours it takes you to return to a working state after a disaster. RPO, which is also expressed in hours, represents how much data you could lose when a disaster happens.

You can share automated Amazon RDS snapshots with up to 20 AWS accounts - This statement is incorrect. Amazon RDS enables you to share DB snapshots or cluster snapshots with other AWS accounts. You can share manual DB snapshots with up to 20 AWS accounts. Automated Amazon RDS snapshots cannot be shared directly with other AWS accounts.

Similar to an Amazon RDS Multi-AZ configuration, failover to a Read Replica is an automated process that requires no manual intervention after initial configurations - Unlike an Amazon RDS Multi-AZ configuration, failover to a Read Replica is not an automated process. If you are using cross-Region Read Replicas, you should be certain that you want to switch your AWS resources between Regions. Cross-Region traffic can experience latency, and reconfiguring applications can be complicated.

Reference:

<https://aws.amazon.com/blogs/database/implementing-a-disaster-recovery-strategy-with-amazon-rds/>

Question 22: **Incorrect**

A digital marketing company uses S3 to store artifacts that may only be accessible to EC2 instances running in a private VPC. The security team at the company is apprehensive about an attack vector wherein any team member with access to this instance could also set up an EC2 instance in another VPC to access these artifacts.

As an AWS Certified Solutions Architect Professional, which of the following solutions will you recommend to prevent such unauthorized access to the artifacts in S3?

-

Configure an S3 VPC endpoint and create an S3 bucket policy to allow access only from this VPC endpoint

(Correct)

-

Attach an Elastic IP to the EC2 instance and create an S3 bucket policy to allow access only from this Elastic IP

-

Set up a highly restricted Security Group for the EC2 instance and create an S3 bucket policy to allow access only from this Security Group

-

Set up an IAM role that allows access to the artifacts in S3 and create an S3 bucket policy to allow access only from this role attached to the instance profile

(Incorrect)

Explanation

Correct option:

Configure an S3 VPC endpoint and create an S3 bucket policy to allow access only from this VPC endpoint

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

A gateway endpoint is a gateway that you specify as a target for a route in your route table for traffic destined to a supported AWS service. One of the ways of letting EC2 instances running in private subnets of a VPC access S3 based resources is by setting up NAT instances in a public subnet and then access those S3 based resources. However, there is a more efficient and secure way. The EC2 instances running in private subnets of a VPC can control access to S3 buckets, objects, and API functions that are in the same Region as the VPC by using the S3 gateway endpoints.

Here are the steps to set up a gateway endpoint:

Step 1: Configure Endpoint

A VPC Endpoint allows you to securely connect your Amazon VPC to another AWS service.

VPC*

Select a VPC



Service

com.amazonaws.us-west-2.s3



Policy*

Full Access - Allow access by any user or service within the VPC using credentials from any AWS accounts to any S3 resources



Custom

Use the [policy creation tool](#) to generate a policy, then paste the generated policy below.

```
{  
  "Statement": [  
    {  
      "Action": "",  
      "Effect": "Allow",  
      "Resource": "",  
      "Principal": ""  
    }  
  ]  
}
```

Cancel and Exit

Next Step

Step 2: Configure Route Tables

A rule with destination pl-68a54001 (com.amazonaws.us-west-2.s3) and a target with this endpoints' ID (e.g. vpce-12345678) will be added to the route tables you select below.

Subnets associated with selected route tables will be able to access this endpoint.

Route Table ID	Name	Main	Associated With
<input type="checkbox"/> rtb-2536b340		Yes	subnet-2f903b58 (54.200.1.0/24) Private s...
<input type="checkbox"/> rtb-2436b341		No	subnet-2e903b59 (54.200.0.0/24) Public s...



When you use an S3 endpoint, the source IP addresses from your instances in your affected subnets for S3 access in the same region will be private IP addresses, not public IP addresses. Existing connections from your affected subnets to S3 that use public IP addresses may be dropped. Ensure that you don't have critical tasks running when you create or modify an endpoint



[Cancel and Exit](#)

[Previous Step](#)

[Create Endpoint](#)

via -

<https://aws.amazon.com/blogs/aws/new-vpc-endpoint-for-amazon-s3/>

Important Characteristics for S3 Gateway Endpoints:

Endpoints for Amazon S3

[PDF](#) | [Kindle](#) | [RSS](#)

If you've already set up access to your Amazon S3 resources from your VPC, you can continue to use Amazon S3 DNS names to access those resources after you've set up an endpoint. However, take note of the following:

- Your endpoint has a policy that controls the use of the endpoint to access Amazon S3 resources. The default policy allows access by any user or service within the VPC, using credentials from any AWS account, to any Amazon S3 resource; including Amazon S3 resources for an AWS account other than the account with which the VPC is associated. For more information, see [Controlling access to services with VPC endpoints](#).
- The source IPv4 addresses from instances in your affected subnets as received by Amazon S3 change from public IPv4 addresses to the private IPv4 addresses from your VPC. An endpoint switches network routes, and disconnects open TCP connections. The previous connections that used public IPv4 addresses are not resumed. We recommend that you do not have any critical tasks running when you create or modify an endpoint; or that you test to ensure that your software can automatically reconnect to Amazon S3 after the connection break.
- You cannot use an IAM policy or bucket policy to allow access from a VPC IPv4 CIDR range (the private IPv4 address range). VPC CIDR blocks can be overlapping or identical, which may lead to unexpected results. Therefore, you cannot use the `aws:SourceIp` condition in your IAM policies for requests to Amazon S3 through a VPC endpoint. This applies to IAM policies for users and roles, and any bucket policies. If a statement includes the `aws:SourceIp` condition, the value fails to match any provided IP address or range. Instead, you can do the following:
 - Use your route tables to control which instances can access resources in Amazon S3 via the endpoint.
 - For bucket policies, you can restrict access to a specific endpoint or to a specific VPC. For more information, see [Using Amazon S3 bucket policies](#).
- Endpoints currently do not support cross-Region requests—ensure that you create your endpoint in the same Region as your bucket. You can find the location of your bucket by using the Amazon S3 console, or by using the `get-bucket-location` command. Use a Region-specific Amazon S3 endpoint to access your bucket; for example, `mybucket.s3-us-west-2.amazonaws.com`. For more information about Region-specific endpoints for Amazon S3, see [Amazon Simple Storage Service \(S3\)](#) in [Amazon Web Services General Reference](#). If you use the AWS CLI to make requests to Amazon S3, set your default Region to the same Region as your bucket, or use the `--region` parameter in your requests.

 **Note**

Treat Amazon S3's US Standard Region as mapped to the us-east-1 Region.

via -

<https://docs.aws.amazon.com/vpc/latest/userguide/vpc-endpoints-s3.html>

You can further use an S3 bucket policy to indicate which VPCs and which VPC Endpoints have access to your S3 buckets.

Restricting Access to a Specific VPC Endpoint

The following is an example of an Amazon S3 bucket policy that restricts access to a specific bucket, `awsexamplebucket1`, only from the VPC endpoint with the ID `vpce-1a2b3c4d`. The policy denies all access to the bucket if the specified endpoint is not being used. The `aws:SourceVpce` condition is used to specify the endpoint. The `aws:SourceVpce` condition does not require an Amazon Resource Name (ARN) for the VPC endpoint resource, only the VPC endpoint ID. For more information about using conditions in a policy, see [Amazon S3 Condition Keys](#).

Important

- Before using the following example policy, replace the VPC endpoint ID with an appropriate value for your use case. Otherwise, you won't be able to access your bucket.
- This policy disables console access to the specified bucket, because console requests don't originate from the specified VPC endpoint.

```
{  
    "Version": "2012-10-17",  
    "Id": "Policy1415115909152",  
    "Statement": [  
        {  
            "Sid": "Access-to-specific-VPCE-only",  
            "Principal": "*",  
            "Action": "s3:*",  
            "Effect": "Deny",  
            "Resource": ["arn:aws:s3:::awsexamplebucket1",  
                        "arn:aws:s3:::awsexamplebucket1/*"],  
            "Condition": {  
                "StringNotEquals": {  
                    "aws:SourceVpce": "vpce-1a2b3c4d"  
                }  
            }  
        }  
    ]  
}
```

Restricting Access to a Specific VPC

You can create a bucket policy that restricts access to a specific VPC by using the `aws:SourceVpc` condition. This is useful if you have multiple VPC endpoints configured in the same VPC, and you want to manage access to your Amazon S3 buckets for all of your endpoints. The following is an example of a policy that allows VPC `vpc-111bbb22` to access `awsexamplebucket1` and its objects. The policy denies all access to the bucket if the specified VPC is not being used. The `vpc-111bbb22` condition key does not require an ARN for the VPC resource, only the VPC ID.

Important

- Before using the following example policy, replace the VPC ID with an appropriate value for your use case. Otherwise, you won't be able to access your bucket.
- This policy disables console access to the specified bucket, because console requests don't originate from the specified VPC.

```
{  
    "Version": "2012-10-17",  
    "Id": "Policy1415115909153",  
    "Statement": [  
        {  
            "Sid": "Access-to-specific-VPC-only",  
            "Principal": "*",  
            "Action": "s3:*",  
            "Effect": "Deny",  
            "Resource": ["arn:aws:s3:::awsexamplebucket1",  
                        "arn:aws:s3:::awsexamplebucket1/*"],  
            "Condition": {  
                "StringNotEquals": {  
                    "aws:SourceVpc": "vpc-111bbb22"  
                }  
            }  
        }  
    ]  
}
```

via - <https://docs.aws.amazon.com/vpc/latest/userguide/vpc-endpoints-s3.html>

Incorrect options:

Set up an IAM role that allows access to the artifacts in S3 and create an S3 bucket policy to allow access only from this role attached to the instance profile - This allows the possibility to attach the given role to multiple EC2 instance profiles and therefore opens up doors for unauthorized access from different EC2 instances. Hence this option is incorrect.

Attach an Elastic IP to the EC2 instance and create an S3 bucket policy to allow access only from this Elastic IP - As described in the explanation above, you cannot use the aws:SourceIp condition in your IAM policies for requests to Amazon S3 through a VPC endpoint. This applies to IAM policies for users and roles, and any bucket policies. Hence this option is incorrect.

Set up a highly restricted Security Group for the EC2 instance and create an S3 bucket policy to allow access only from this Security Group - This option has been added as a distractor as a Security Group is not a valid Principal to be used in an S3 bucket policy. Security Group also cannot be used in a valid Condition statement in the bucket policy.

References:

<https://aws.amazon.com/blogs/aws/new-vpc-endpoint-for-amazon-s3/>

<https://docs.aws.amazon.com/vpc/latest/userguide/vpc-endpoints-s3.html>

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

Question 23: **Incorrect**

The engineering team at a retail company has deployed a fleet of EC2 instances under an Auto Scaling group (ASG). The instances under the ASG span two Availability Zones (AZ) within the eu-west-1 region. All the incoming requests are handled by an Application Load Balancer (ALB) that routes the requests to the EC2 instances under the ASG. A planned migration went wrong last week when two instances (belonging to AZ 1) were manually terminated and desired capacity was reduced causing the Availability Zones to become unbalanced. Later that day, another instance (belonging to AZ 2) was detected as unhealthy by the Application Load Balancer's health check.

Which of the following options represent the correct outcomes for the aforesaid events? (Select two)



- As the Availability Zones got unbalanced, Amazon EC2 Auto Scaling will compensate by rebalancing the Availability Zones. When rebalancing, Amazon EC2 Auto Scaling launches new instances before terminating the old ones, so that rebalancing does not compromise the performance or availability of your application

(Correct)



- Amazon EC2 Auto Scaling creates a new scaling activity for terminating the unhealthy instance and then terminates it. Later, another scaling activity launches a new instance to replace the terminated instance

(Correct)



- As the Availability Zones got unbalanced, Amazon EC2 Auto Scaling will compensate by rebalancing the Availability Zones. When rebalancing, Amazon EC2 Auto Scaling terminates old instances before launching new instances, so that rebalancing does not cause extra instances to be launched



- Amazon EC2 Auto Scaling creates a new scaling activity to terminate the unhealthy instance and launch the new instance simultaneously



- Amazon EC2 Auto Scaling creates a new scaling activity for launching a new instance to replace the unhealthy instance. Later, EC2 Auto Scaling creates a new scaling activity for terminating the unhealthy instance and then terminates it

(Incorrect)

Explanation

Correct options:

As the Availability Zones got unbalanced, Amazon EC2 Auto Scaling will compensate by rebalancing the Availability Zones. When rebalancing, Amazon EC2 Auto Scaling launches new instances before terminating the old ones, so that rebalancing does not compromise the performance or availability of your application

Amazon EC2 Auto Scaling helps you ensure that you have the correct number of Amazon EC2 instances available to handle the load for your application. You create collections of EC2 instances, called Auto Scaling groups. You can specify the minimum number of instances in each Auto Scaling group, and Amazon EC2 Auto Scaling ensures that your group never goes below this size.

Actions such as changing the Availability Zones for your group or explicitly terminating or detaching instances and decrement the desired capacity at the same time, it can lead to the Auto Scaling group becoming unbalanced between Availability Zones. Amazon EC2 Auto Scaling compensates by rebalancing the Availability Zones.

When rebalancing, Amazon EC2 Auto Scaling launches new instances before terminating the old ones, so that rebalancing does not compromise the performance or availability of your application. Therefore, this option is correct.

Rebalancing activities

After certain actions occur, your Auto Scaling group can become unbalanced between Availability Zones. Amazon EC2 Auto Scaling compensates by rebalancing the Availability Zones. The following actions can lead to rebalancing activity:

- You change the Availability Zones for your group.
- You explicitly terminate or detach instances and the group becomes unbalanced.
- An Availability Zone that previously had insufficient capacity recovers and has additional capacity available.
- An Availability Zone that previously had a Spot price above your maximum price now has a Spot price below your maximum price.

When rebalancing, Amazon EC2 Auto Scaling launches new instances before terminating the old ones, so that rebalancing does not compromise the performance or availability of your application.

Because Amazon EC2 Auto Scaling attempts to launch new instances before terminating the old ones, being at or near the specified maximum capacity could impede or completely halt rebalancing activities. To avoid this problem, the system can temporarily exceed the specified maximum capacity of a group by a 10 percent margin (or by a 1-instance margin, whichever is greater) during a rebalancing activity. The margin is extended only if the group is at or near maximum capacity and needs rebalancing, either because of user-requested rezoning or to compensate for zone availability issues. The extension lasts only as long as needed to rebalance the group typically a few minutes.

Availability Zone Rebalancing Overview:

<https://docs.aws.amazon.com/autoscaling/ec2/userguide/auto-scaling-benefits.html>

via -

Amazon EC2 Auto Scaling creates a new scaling activity for terminating the unhealthy instance and then terminates it. Later, another scaling activity launches a new instance to replace the terminated instance

However, the scaling activity of Auto Scaling works in a different sequence compared to the rebalancing activity. Auto Scaling creates a new scaling activity for terminating the unhealthy instance and then terminates it. Later, another scaling activity launches a new instance to replace the terminated instance.

Incorrect options:

Amazon EC2 Auto Scaling creates a new scaling activity for launching a new instance to replace the unhealthy instance. Later, EC2 Auto Scaling creates a new scaling activity for terminating the unhealthy instance and then terminates it - This option contradicts the correct sequence of events outlined earlier for scaling activity created by EC2 Auto Scaling. Actually, Auto Scaling first terminates the unhealthy instance and then launches a new instance. Hence this is incorrect.

As the Availability Zones got unbalanced, Amazon EC2 Auto Scaling will compensate by rebalancing the Availability Zones. When rebalancing, Amazon EC2 Auto Scaling terminates old instances before launching new instances, so that rebalancing does not cause extra instances to be launched - This option contradicts the correct sequence of events outlined earlier for rebalancing activity. When rebalancing, Amazon EC2 Auto Scaling launches new instances before terminating the old ones. Hence this is incorrect.

Amazon EC2 Auto Scaling creates a new scaling activity to terminate the unhealthy instance and launch the new instance simultaneously - This is a made-up option as both the terminate and launch activities can't happen simultaneously. This option has been added as a distractor.

References:

<https://docs.aws.amazon.com/autoscaling/ec2/userguide/what-is-amazon-ec2-auto-scaling.html>

<https://docs.aws.amazon.com/autoscaling/ec2/userguide/auto-scaling-benefits.html>

<https://docs.aws.amazon.com/autoscaling/ec2/userguide/as-instance-termination.html>

<https://docs.aws.amazon.com/autoscaling/ec2/userguide/healthcheck.html>

<https://docs.aws.amazon.com/autoscaling/ec2/userguide/common-scenarios-termination.html#common-scenarios-termination-rebalancing>

Question 24: **Incorrect**

A multi-national digital media company wants to exit out of the business of owning and maintaining its own IT infrastructure so it can redeploy resources toward innovation in Artificial Intelligence and related areas to create a better customer experience. As part of this

digital transformation, the media company wants to archive about 9 PB of data in its on-premises data center to durable long term storage.

As a Solutions Architect Professional, what is your recommendation to migrate and store this data in the quickest and MOST cost-optimal way?



Transfer the on-premises data into a Snowmobile device. Copy the Snowmobile data into Amazon S3 and create a lifecycle policy to transition the data into AWS Glacier

(Incorrect)



Transfer the on-premises data into a Snowmobile device. Copy the Snowmobile data directly into AWS Glacier



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

(Correct)



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

Explanation

Correct option:

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

Snowball Edge Storage Optimized is the optimal choice if you need to securely and quickly transfer dozens of terabytes to petabytes of data to AWS. It provides up to 80 TB of usable HDD storage, 40 vCPUs, 1 TB of SATA SSD storage, and up to 40 Gb network connectivity to address large scale data transfer and pre-processing use cases.

The data stored on the Snowball Edge device can be copied into the S3 bucket and later transitioned into AWS Glacier via a lifecycle policy. You can't directly copy data from Snowball Edge devices into AWS Glacier.

AWS Snowball Edge

AWS Snowball Edge [\[\]](#) is a data migration and edge computing device that comes in two options. Snowball Edge Storage Optimized provides 100 TB of capacity and 24 vCPUs and is well suited for local storage and large scale data transfer. Snowball Edge Compute Optimized provides 52 vCPUs and an optional GPU for use cases such as advanced machine learning and full motion video analysis in disconnected environments. Customers can use these two options for data collection, machine learning and processing, and storage in environments with intermittent connectivity (such as manufacturing, industrial, and transportation) or in extremely remote locations (such as military or maritime operations) before shipping it back to AWS. These devices may also be rack mounted and clustered together to build larger, temporary installations.

Snowball Edge supports specific Amazon EC2 instance types as well as AWS Lambda functions, so customers may develop and test in AWS then deploy applications on devices in remote locations to collect, pre-process, and return the data. Common use cases include data migration, data transport, image collation, IoT sensor stream capture, and machine learning.

AWS Snowmobile

AWS Snowmobile [\[\]](#) is an exabyte-scale data transfer service used to move extremely large amounts of data to AWS. You can transfer up to 100 PB per Snowmobile, a 45-foot long ruggedized shipping container, pulled by a semi-trailer truck. Snowmobile makes it easy to move massive volumes of data to the cloud, including video libraries, image repositories, or even a complete data center migration. Transferring data with Snowmobile is secure, fast, and cost effective.

After an initial assessment, a Snowmobile will be transported to your data center, and AWS personnel will configure it for you so it can be accessed as a network storage target. When your Snowmobile is on site, AWS personnel will work with your team to connect a removable, high-speed network switch from the Snowmobile to your local network. Then you can begin your high-speed data transfer from any number of sources within your data center to the Snowmobile. After your data is loaded, the Snowmobile is driven back to AWS where your data is imported into Amazon S3 or S3 Glacier.

AWS Snowmobile uses multiple layers of security designed to protect your data including dedicated security personnel, GPS tracking, alarm monitoring, 24/7 video surveillance, and an optional escort security vehicle while in transit. All data is encrypted with 256-bit encryption keys managed through AWS KMS and designed to ensure both security and full chain of custody of your data.

via -

<https://docs.aws.amazon.com/whitepapers/latest/aws-overview/migration-services.html>

Incorrect options:

Transfer the on-premises data into a Snowmobile device. Copy the Snowmobile data into Amazon S3 and create a lifecycle policy to transition the data into AWS Glacier - AWS Snowmobile is an exabyte-scale data transfer service used to move extremely large amounts of data to AWS. You can transfer up to 100 PB per Snowmobile. Snowmobile makes it easy to move massive volumes of data to the cloud, including video libraries, image repositories, or even a complete data center migration.

AWS recommends that you should use Snowmobile to migrate large datasets of 10PB or more in a single location. For datasets less than 10PB or distributed in multiple locations, you should use Snowball. So, this option is not the best fit for the given use-case.

Q: How should I choose between Snowmobile and Snowball?

To migrate large datasets of 10PB or more in a single location, you should use Snowmobile. For datasets less than 10PB or distributed in multiple locations, you should use Snowball. In addition, you should evaluate the amount of available bandwidth in your network backbone. If you have a high speed backbone with hundreds of Gb/s of spare throughput, then you can use Snowmobile to migrate the large datasets all at once. If you have limited bandwidth on your backbone, you should consider using multiple Snowballs to migrate the data incrementally.

via -

<https://aws.amazon.com/snowmobile/faqs/>

Transfer the on-premises data into a Snowmobile device. Copy the Snowmobile data directly into AWS Glacier - As mentioned above, AWS recommends that you should use Snowmobile to migrate large datasets of 10PB or more in a single location. For datasets less than 10PB or distributed in multiple locations, you should use Snowball. So, this option is not the best fit for the given use-case.

Although you should note that for Snowmobile, you can import your data directly into Glacier.

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

References:

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

<https://docs.aws.amazon.com/snowball/latest/ug/how-it-works.html>

<https://aws.amazon.com/snowmobile/faqs/>

<https://docs.aws.amazon.com/whitepapers/latest/aws-overview/migration-services.html>

Question 25: **Correct**

A gaming company runs its flagship application with an SLA of 99.99%. Global users access the application 24/7. The application is currently hosted on the on-premises data centers and it routinely fails to meet its SLA, especially when hundreds of thousands of users access the application concurrently. The engineering team has also received complaints from some users about high latency.

As a Solutions Architect Professional, how would you redesign this application for scalability and also allow for automatic failover at the lowest possible cost?



Configure Route 53 latency-based routing to route to the nearest Region and activate the health checks. Host the website on S3 in each Region and use API Gateway with AWS Lambda for the application layer. Set up the data layer using DynamoDB global tables with DAX for caching

(Correct)



Configure Route 53 round-robin routing policy to distribute load evenly across all Regions and activate the health checks. Host the website behind a Network Load Balancer (NLB) with targets as ECS containers using Fargate. Repeat this configuration of NLB with ECS containers using Fargate in multiple Regions. Use Aurora Global database as the data layer



Configure Route 53 geolocation-based routing to route to the nearest Region and activate the health checks. Host the website behind a Network Load Balancer (NLB) with targets as ECS containers using Fargate. Repeat this configuration of NLB with ECS containers using Fargate in multiple Regions. Use Aurora Global database as the data layer

-

Configure a combination of Route 53 failover routing with geolocation-based routing. Host the website behind an Application Load Balancer (ALB) with targets as EC2 instances that are automatically scaled via Auto-Scaling Group (ASG). Repeat this configuration of ALB with EC2 instances as targets that are scaled via ASG in multiple Regions. Use a Multi-AZ deployment with RDS MySQL as the data layer

Explanation

Correct option:

Configure Route 53 latency-based routing to route to the nearest Region and activate the health checks. Host the website on S3 in each Region and use API Gateway with AWS Lambda for the application layer. Set up the data layer using DynamoDB global tables with DAX for caching

You can use Route 53 routing policies for a hosted zone record to determine how Amazon Route 53 responds to queries.

Choosing a routing policy

[PDF](#) | [Kindle](#) | [RSS](#)

When you create a record, you choose a routing policy, which determines how Amazon Route 53 responds to queries:

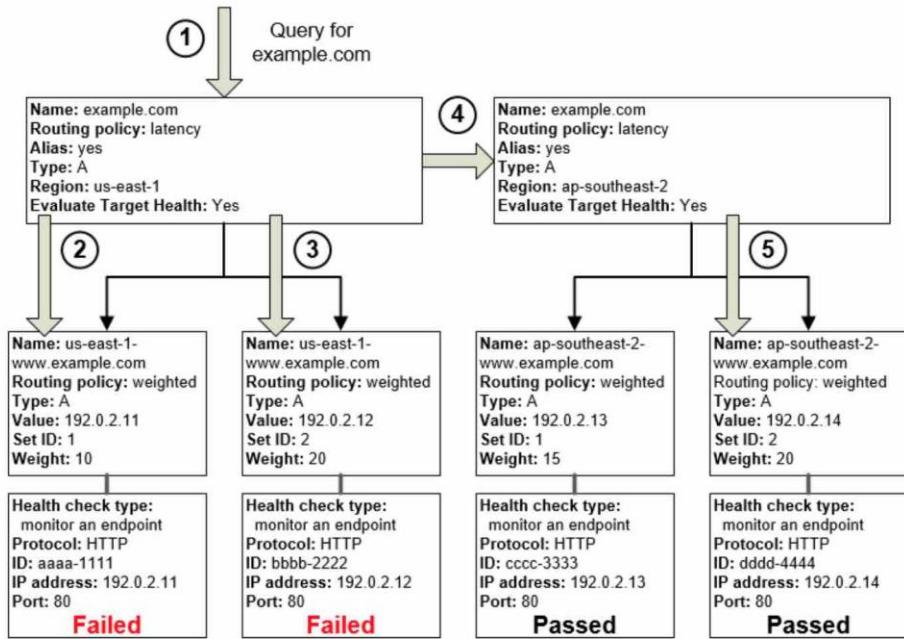
- **Simple routing policy** – Use for a single resource that performs a given function for your domain, for example, a web server that serves content for the example.com website.
- **Failover routing policy** – Use when you want to configure active-passive failover.
- **Geolocation routing policy** – Use when you want to route traffic based on the location of your users.
- **Geoproximity routing policy** – Use when you want to route traffic based on the location of your resources and, optionally, shift traffic from resources in one location to resources in another.
- **Latency routing policy** – Use when you have resources in multiple AWS Regions and you want to route traffic to the region that provides the best latency.
- **Multivalue answer routing policy** – Use when you want Route 53 to respond to DNS queries with up to eight healthy records selected at random.
- **Weighted routing policy** – Use to route traffic to multiple resources in proportions that you specify.

via -

<https://docs.aws.amazon.com/Route53/latest/DeveloperGuide/routing-policy.html>

You can use a combination of alias records (such as latency alias, weighted alias or failover alias) and non-alias records to build a decision tree that gives you greater control over how Route 53 responds to requests. For example, you might use latency alias records to select a Region close to a user and use weighted records for two or more resources within each Region to protect against the failure of a single endpoint or an Availability Zone.

The following schematic shows how you can use a combination of latency routing policy (that has health checks activated) with a weighted routing policy to manage a multi-Region routing infrastructure. In case a Region goes down, Route 53 would look for the latency alias record with the next-best latency and choose the record for the other Region.



The preceding diagram illustrates the following sequence of events:

1. Route 53 receives a query for `example.com`. Based on the latency for the user making the request, Route 53 selects the latency alias record for the `us-east-1` region.
2. Route 53 selects a weighted record based on weight. Evaluate Target Health is Yes for the latency alias record, so Route 53 checks the health of the selected weighted record.
3. The health check failed, so Route 53 chooses another weighted record based on weight and checks its health. That record also is unhealthy.
4. Route 53 backs out of that branch of the tree, looks for the latency alias record with the next-best latency, and chooses the record for `ap-southeast-2`.
5. Route 53 again selects a record based on weight, and then checks the health of the selected resource. The resource is healthy, so Route 53 returns the applicable value in response to the query.

via -

<https://docs.aws.amazon.com/Route53/latest/DeveloperGuide/dns-failover-complex-configs.html>

Therefore, for the given use-case, this solution allows for automatic failover while minimizing the latency. As the downstream components for this solution such as the API Gateway, Lambda functions and DynamoDB global tables (with DAX) are serverless, so it allows the solution to scale easily.

Incorrect options:

Configure Route 53 geolocation-based routing to route to the nearest Region and activate the health checks. Host the website behind a Network Load Balancer (NLB) with targets as ECS containers using Fargate. Repeat this configuration of NLB with ECS containers using Fargate in multiple Regions. Use Aurora Global database as the data layer - You cannot use geolocation-based routing to route to the nearest Region as you need to use latency based routing to accomplish that.

Geolocation routing is used when you want to route traffic based on the location of your users. When you use geolocation routing, you can localize your content and present some or all of your website in the language of your users. You can also use geolocation routing to restrict the distribution of content to only the locations in which you have distribution rights. Other components mentioned in the stack are not a deal-breaker, but using geolocation-based routing makes this option incorrect.

Configure Route 53 round-robin routing policy to distribute load evenly across all Regions and activate the health checks. Host the website behind a Network Load Balancer (NLB) with targets as ECS containers using Fargate. Repeat this configuration of NLB with ECS containers using Fargate in multiple Regions. Use Aurora Global database as the data layer - This option has been added as a distractor as there is no such thing as a Route 53 round-robin routing policy.

Configure a combination of Route 53 failover routing with geolocation-based routing. Host the website behind an Application Load Balancer (ALB) with targets as EC2 instances that are automatically scaled via Auto-Scaling Group (ASG). Repeat this configuration of ALB with EC2 instances as targets that are scaled via ASG in multiple Regions. Use a Multi-AZ deployment with RDS MySQL as the data layer - As mentioned earlier, you cannot use geolocation-based routing to route to the nearest Region as you need to use latency based routing to accomplish that. Failover routing could help with ensuring failover across AWS Regions, but using geolocation-based routing is a deal-breaker. Another red-flag is using Multi-AZ RDS MySQL as data layer, which can be deployed in just one AWS Region. You would need to develop and maintain custom data sync scripts/jobs to maintain data consistency across Regions.

References:

<https://docs.aws.amazon.com/Route53/latest/DeveloperGuide/routing-policy.html>

<https://docs.aws.amazon.com/Route53/latest/DeveloperGuide/dns-failover-complex-configs.html>

<https://aws.amazon.com/premiumsupport/knowledge-center/route-53-dns-health-checks/>

Question 26: **Correct**

The world's largest cable company uses AWS in a hybrid environment to innovate and deploy features for its flagship video product, XFINITY X1, several times a week. The company uses AWS products such as Amazon Virtual Private Cloud (Amazon VPC) and Amazon Direct Connect to deliver the scalability and security needed for rapidly innovating in a hybrid environment. As part of an internal product roadmap, the engineering team at the company has created a private hosted zone and associated it with a virtual private cloud (VPC). However, the domain names remain unresolved, resulting in errors.

As a Solutions Architect Professional, which of the following Amazon VPC configuration options would you use to get the private hosted zone to work?

-

Name server (NS) record and Start Of Authority (SOA) records should have the correct configurations

-

The private and public hosted zones should not have overlapping namespaces

-

There is a private hosted zone and a Resolver rule that routes traffic to your network for the same domain name resulting in an ambiguous routing rule

-

To use private hosted zones, DNS hostnames and DNS resolution should be enabled for the VPC

(Correct)

Explanation

Correct option:

To use private hosted zones, DNS hostnames and DNS resolution should be enabled for the VPC - DNS hostnames and DNS resolution are required settings for private hosted zones. DNS queries for private hosted zones can be resolved by the Amazon-provided VPC DNS server only. As a result, these options must be enabled for your private hosted zone to work.

DNS hostnames: For non-default virtual private clouds that aren't created using the Amazon VPC wizard, this option is disabled by default. If you create a private hosted zone for a domain and create records in the zone without enabling DNS hostnames, private hosted zones aren't enabled. To use a private hosted zone, this option must be enabled.

DNS resolution: Private hosted zones accept DNS queries only from a VPC DNS server. The IP address of the VPC DNS server is the reserved IP address at the base of the VPC IPv4 network range plus two. Enabling DNS resolution allows you to use the VPC DNS server as a Resolver for performing DNS resolution. Keep this option disabled if you're using a custom DNS server in the DHCP Options set, and you're not using a private hosted zone.

Incorrect options:

The private and public hosted zones should not have overlapping namespaces - If you have private and public hosted zones that have overlapping namespaces, such as example.com and accounting.example.com, Resolver routes traffic based on the most specific match. It won't result in an error.

Name server (NS) record and Start Of Authority (SOA) records should have the correct configurations - When you create a hosted zone, Amazon Route 53 automatically creates a name server (NS) record and a start of authority (SOA) record for the zone for public hosted zone. The current requirement is about the private hosted zone, hence this is a wrong choice.

There is a private hosted zone and a Resolver rule that routes traffic to your network for the same domain name resulting in an ambiguous routing rule - If you have a private hosted zone (example.com) and a Resolver rule that routes traffic to your network for the same domain name, the Resolver rule takes precedence. It doesn't result in any error.

References:

<https://aws.amazon.com/premiumsupport/knowledge-center/vpc-enable-private-hosted-zone/>

<https://docs.aws.amazon.com/Route53/latest/DeveloperGuide/hosted-zone-private-considerations.html>

<https://docs.aws.amazon.com/Route53/latest/DeveloperGuide/hosted-zone-public-considerations.html>

Question 27: **Correct**

A global biomedicine company has built a Genomics Solution on AWS Cloud. The company's labs generate hundreds of terabytes of research data daily. To further accelerate the innovation process, the engineering team at the company wants to move most of the on-premises data into Amazon S3, Amazon EFS, and Amazon FSx for Windows File Server easily, quickly, and cost-effectively. The team would like to automate and accelerate online data transfers to these AWS storage services.

As a Solutions Architect Professional, which of the following solutions would you recommend as the BEST fit?



Use AWS Transfer Family to automate and accelerate online data transfers to the given AWS storage services



Use AWS Snowball Edge Storage Optimized device to automate and accelerate online data transfers to the given AWS storage services



Use AWS DataSync to automate and accelerate online data transfers to the given AWS storage services

(Correct)



Use File Gateway to automate and accelerate online data transfers to the given AWS storage services

Explanation

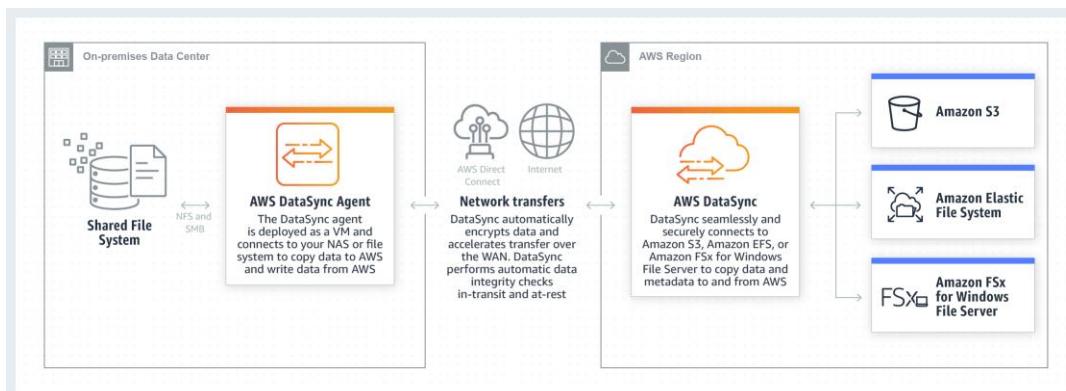
Correct option:

Use AWS DataSync to automate and accelerate online data transfers to the given AWS storage services

AWS DataSync is an online data transfer service that simplifies, automates, and accelerates copying large amounts of data to and from AWS storage services over the internet or AWS Direct Connect.

AWS DataSync fully automates and accelerates moving large active datasets to AWS, up to 10 times faster than command-line tools. It is natively integrated with Amazon S3, Amazon EFS, Amazon FSx for Windows File Server, Amazon CloudWatch, and AWS CloudTrail, which provides seamless and secure access to your storage services, as well as detailed monitoring of the transfer.

DataSync uses a purpose-built network protocol and scale-out architecture to transfer data. A single DataSync agent is capable of saturating a 10 Gbps network link. DataSync fully automates the data transfer. It comes with retry and network resiliency mechanisms, network optimizations, built-in task scheduling, monitoring via the DataSync API and Console, and CloudWatch metrics, events, and logs that provide granular visibility into the transfer process. DataSync performs data integrity verification both during the transfer and at the end of the transfer.



How DataSync Works

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

via -

Incorrect options:

Use AWS Snowball Edge Storage Optimized device to automate and accelerate online data transfers to the given AWS storage services - Snowball Edge Storage Optimized is the optimal choice if you need to securely and quickly transfer dozens of terabytes to petabytes of data to AWS. It provides up to 80 TB of usable HDD storage, 40 vCPUs, 1 TB of SATA SSD storage, and up to 40 Gb network connectivity to address large scale data transfer and pre-processing use cases. As each Snowball Edge Storage Optimized device can handle 80TB of data, you can order 10 such devices to take care of the data transfer for all applications. The original Snowball devices were transitioned out of service and Snowball Edge Storage Optimized are now the primary devices used for data transfer. You may see the Snowball device on the exam, just remember that the original Snowball device had 80TB of storage space.

AWS Snowball Edge is suitable for offline data transfers, for customers who are bandwidth constrained or transferring data from remote, disconnected, or austere environments. Therefore, it cannot support automated and accelerated online data transfers.

Use AWS Transfer Family to automate and accelerate online data transfers to the given AWS storage services - The AWS Transfer Family provides fully managed support for file transfers directly into and out of Amazon S3. Therefore, it cannot support migration into the other AWS storage services mentioned in the given use-case (such as EFS and Amazon FSx for Windows File Server).

Use File Gateway to automate and accelerate online data transfers to the given AWS storage services - AWS Storage Gateway's file interface, or file gateway, offers you a seamless way to connect to the cloud to store application data files and backup images as durable objects on Amazon S3 cloud storage. File gateway offers SMB or NFS-based access to data in Amazon S3 with local caching. It can be used for on-premises applications, and for Amazon EC2-based applications that need file protocol access to S3 object storage. Therefore, it cannot support migration into the other AWS storage services mentioned in the given use-case (such as EFS and Amazon FSx for Windows File Server).

References:

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

<https://aws.amazon.com/storagegateway/file/>

<https://aws.amazon.com/aws-transfer-family/>

Question 28: **Correct**

The engineering team at a company is evaluating the Multi-AZ and Read Replica capabilities of RDS MySQL vs Aurora MySQL before they implement the solution in their production environment. The company has hired you as an AWS Certified Solutions Architect Professional to provide a detailed report on this technical requirement.

Which of the following would you identify as correct regarding the given use-case? (Select three)

-

Multi-AZ deployments for both RDS MySQL and Aurora MySQL follow synchronous replication

(Correct)

-

Multi-AZ deployments for Aurora MySQL follow synchronous replication whereas Multi-AZ deployments for RDS MySQL follow asynchronous replication

-

Read Replicas can be manually promoted to a standalone database instance for Aurora MySQL whereas Read Replicas for RDS MySQL can be promoted to the primary instance

-

The primary and standby DB instances are upgraded at the same time for RDS MySQL Multi-AZ. All instances are upgraded at the same time for Aurora MySQL

(Correct)

-

Read Replicas can be manually promoted to a standalone database instance for RDS MySQL whereas Read Replicas for Aurora MySQL can be promoted to the primary instance

(Correct)

-

Database engine version upgrades happen on primary for Aurora MySQL whereas all instances are updated together for RDS MySQL
Explanation

Correct options:

Multi-AZ deployments for both RDS MySQL and Aurora MySQL follow synchronous replication

Read Replicas can be manually promoted to a standalone database instance for RDS MySQL whereas Read Replicas for Aurora MySQL can be promoted to the primary instance

The primary and standby DB instances are upgraded at the same time for RDS MySQL Multi-AZ. All instances are upgraded at the same time for Aurora MySQL

RDS Read replicas are a special type of DB instances that make use of the built-in replication functionality for RDS. The source DB instance becomes the primary DB instance. Updates made to the primary DB instance are asynchronously copied to the read replica. When you create a read replica, you first specify an existing DB instance as the source. Then Amazon RDS takes a snapshot of the source instance and creates a read-only instance from the snapshot. Amazon RDS then uses the asynchronous replication method for the DB engine to update the read replica whenever there is a change to the primary DB instance.

Aurora Replicas are independent endpoints in an Aurora DB cluster, best used for scaling read operations and increasing availability. Up to 15 Aurora Replicas can be distributed across the Availability Zones that a DB cluster spans within an AWS Region. The DB cluster volume is made up of multiple copies of the data for the DB cluster. However, the data in the cluster volume is represented as a single, logical volume to the primary instance and to Aurora Replicas in the DB cluster.

In a Multi-AZ deployment, Amazon RDS automatically provisions and maintains a synchronous standby replica in a different Availability Zone. The primary DB instance is synchronously replicated across Availability Zones to a standby replica to provide data

redundancy, eliminate I/O freezes, and minimize latency spikes during system backups. Aurora stores copies of the data in a DB cluster across multiple Availability Zones in a single AWS Region. When data is written to the primary DB instance, Aurora synchronously replicates the data across Availability Zones to six storage nodes associated with your cluster volume.

To increase availability, you can use Aurora Replicas as failover targets. That is, if the primary instance fails, an Aurora Replica is promoted to the primary instance. There is a brief interruption during which read and write requests made to the primary instance fail with an exception, and the Aurora Replicas are rebooted.

For RDS MySQL in Multi-AZ configuration, database engine version upgrades happen on both the primary and standby DB instances at the same time. For Aurora MySQL, all instances are upgraded at the same time.

Overview of Upgrading

Amazon RDS takes two DB snapshots during the upgrade process. The first DB snapshot is of the DB instance before any upgrade changes have been made. If the upgrade doesn't work for your databases, you can restore this snapshot to create a DB instance running the old version. The second DB snapshot is taken when the upgrade completes.

 **Note**

Amazon RDS only takes DB snapshots if you have set the backup retention period for your DB instance to a number greater than 0. To change your backup retention period, see [Modifying an Amazon RDS DB Instance](#).

After the upgrade is complete, you can't revert to the previous version of the database engine. If you want to return to the previous version, restore the first DB snapshot taken to create a new DB instance.

You control when to upgrade your DB instance to a new version supported by Amazon RDS. This level of control helps you maintain compatibility with specific database versions and test new versions with your application before deploying in production. When you are ready, you can perform version upgrades at the times that best fit your schedule.

If your DB instance is using read replication, upgrade all of the read replicas before upgrading the source instance.

If your DB instance is in a Multi-AZ deployment, both the primary and standby DB instances are upgraded. The primary and standby DB instances are upgraded at the same time and you experience an outage until the upgrade is complete. The time for the outage varies based on the size of your DB instance.

via -

https://docs.aws.amazon.com/AmazonRDS/latest/UserGuide/USER_UpgradeDBInstance.MySQL.html

Incorrect options:

Database engine version upgrades happen on primary for Aurora MySQL whereas all instances are updated together for RDS MySQL

Read Replicas can be manually promoted to a standalone database instance for Aurora MySQL whereas Read Replicas for RDS MySQL can be promoted to the primary instance

Multi-AZ deployments for Aurora MySQL follow synchronous replication whereas Multi-AZ deployments for RDS MySQL follow asynchronous replication

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

References:

<https://aws.amazon.com/rds/features/read-replicas/>

https://docs.aws.amazon.com/AmazonRDS/latest/UserGuide/USER_UpgradeDBInstance.MySQL.html

<https://docs.aws.amazon.com/AmazonRDS/latest/AuroraUserGuide/Concepts.AuroraHighAvailability.html>

Question 29: **Incorrect**

An e-commerce company wants to rollout and test a blue-green deployment for its global application in the next couple of days. Most of the customers use mobile phones which are prone to DNS caching. The company has only two days left before the big sale will be launched.

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



Use AWS CodeDeploy deployment options to choose the right deployment



Use Elastic Load Balancer to distribute traffic across deployments



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

(Incorrect)



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

(Correct)

Explanation

Correct option:

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

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

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

While relying on the DNS service is a great option for blue/green deployments, it may not fit use-cases that require a fast and controlled transition of the traffic. Some client devices and internet resolvers cache DNS answers for long periods; this DNS feature improves the efficiency of the DNS service as it reduces the DNS traffic across the Internet, and serves as a resiliency technique by preventing authoritative name-server overloads. The downside of this in blue/green deployments is that you don't know how long it will take before all of your users receive updated IP addresses when you update a record, change your routing preference or when there is an application failure.

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

Incorrect options:

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

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

Use AWS CodeDeploy deployment options to choose the right deployment - In CodeDeploy, a deployment is the process, and the components involved in the process, of installing content on one or more instances. This content can consist of code, web and configuration files, executables, packages, scripts, and so on. CodeDeploy deploys content that is stored in a source repository, according to the configuration rules you specify. Blue/Green deployment is one of the deployment types that CodeDeploy supports. AWS CodeDeploy performs deployments with AWS resources located in the same region, so this option is ruled out.

References:

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

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

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

Question 30: **Correct**

An e-commerce company has hired an AWS Certified Solutions Architect Professional to transform a standard three-tier web application architecture in AWS. Currently, the web and application tiers run on EC2 instances and the database tier runs on RDS MySQL. The company wants to redesign the web and application tiers to use API Gateway with Lambda Functions with the final goal of deploying the new application within 6 months. As an immediate short-term task, the Engineering Manager has mandated the Solutions Architect to reduce costs for the existing stack.

Which of the following options should the Solutions Architect recommend as the MOST cost-effective and reliable solution?

-

Provision On-Demand Instances for the web and application tiers and Reserved Instances for the database tier

(Correct)

-

Provision Reserved Instances for the web and application tiers and On-Demand Instances for the database tier

-

Provision Spot Instances for the web and application tiers and Reserved Instances for the database tier

-

Provision Reserved Instances for the web, application and database tiers

Explanation

Correct option:

Provision On-Demand Instances for the web and application tiers and Reserved Instances for the database tier

EC2 Instances support five different ways to pay for provisioning the servers: On-Demand, Savings Plans, Reserved Instances, Spot Instances and Dedicated Hosts.

On-Demand

With On-Demand instances, you pay for compute capacity by the hour or the second depending on which instances you run. No longer-term commitments or upfront payments are needed. You can increase or decrease your compute capacity depending on the demands of your application and only pay the specified per hourly rates for the instance you use.

On-Demand instances are recommended for:

- Users that prefer the low cost and flexibility of Amazon EC2 without any up-front payment or long-term commitment
- Applications with short-term, spiky, or unpredictable workloads that cannot be interrupted
- Applications being developed or tested on Amazon EC2 for the first time

[See On-Demand pricing »](#)

Spot instances

Amazon EC2 Spot instances allow you to request spare Amazon EC2 computing capacity for up to 90% off the On-Demand price. [Learn More](#).

Spot instances are recommended for:

- Applications that have flexible start and end times
- Applications that are only feasible at very low compute prices
- Users with urgent computing needs for large amounts of additional capacity

[See Spot pricing »](#)

Savings Plans

Savings Plans are a flexible pricing model that offer low prices on EC2 and Fargate usage, in exchange for a commitment to a consistent amount of usage (measured in \$/hour) for a 1 or 3 year term.

Dedicated Hosts

A Dedicated Host is a physical EC2 server dedicated for your use. Dedicated Hosts can help you reduce costs by allowing you to use your existing server-bound software licenses, including Windows Server, SQL Server, and SUSE Linux Enterprise Server (subject to your license terms), and can also help you meet compliance requirements. [Learn more](#).

- Can be purchased On-Demand (hourly).
- Can be purchased as a Reservation for up to 70% off the On-Demand price.

[See Dedicated pricing »](#)

Reserved Instances

Reserved Instances provide you with a significant discount (up to 75%) compared to On-Demand instance pricing. In addition, when Reserved Instances are assigned to a specific Availability Zone, they provide a capacity reservation, giving you additional confidence in your ability to launch instances when you need them.

For applications that have steady state or predictable usage, Reserved Instances can provide significant savings compared to using On-Demand instances. See [How to Purchase Reserved Instances](#) for more information.

Reserved Instances are recommended for:

- Applications with steady state usage
- Applications that may require reserved capacity
- Customers that can commit to using EC2 over a 1 or 3 year term to reduce their total computing costs

via -

<https://aws.amazon.com/ec2/pricing/>

An On-Demand Instance is an instance that you use on-demand. You have full control over its lifecycle — you decide when to launch, stop, hibernate, start, reboot, or terminate it. There is no long-term commitment required when you purchase On-Demand Instances. There is no upfront payment and you pay only for the seconds that your On-Demand Instances are running. The price per second for

running an On-Demand Instance is fixed. On-demand instances cannot be interrupted. However, On-demand instances are not as cost-effective as Reserved instances.

A Spot Instance is an unused EC2 instance that is available for less than the On-Demand price. Because Spot Instances enable you to request unused EC2 instances at steep discounts (up to 90%), you can lower your Amazon EC2 costs significantly. Spot Instances are well-suited for data analysis, batch jobs, background processing, and optional tasks. These can be terminated at short notice, so these are not suitable for critical workloads that need to run at a specific point in time.

Reserved Instances provide you with significant savings (up to 75%) on your Amazon EC2 costs compared to On-Demand Instance pricing. Reserved Instances are not physical instances, but rather a billing discount applied to the use of On-Demand Instances in your account. You can purchase a Reserved Instance for a one-year or three-year commitment, with the three-year commitment offering a bigger discount. Reserved instances cannot be interrupted.

For the given use-case, only the web and application tiers would be re-engineered using API Gateway and Lambda within a duration of 6 months, so you cannot use Reserved Instances for these tiers as the minimum duration to purchase a Reserved Instance is 1 year. Additionally, using Spot Instances for these tiers is also ruled out because these can be terminated at short notice and would not be able to offer reliability for the web and application tiers. Therefore On-Demand is the best option for the web and application tiers. As the proposed transformation would not impact the database tier running on RDS MySQL, therefore you can purchase Reserved Instances for the database tier as the most cost-effective solution.

Incorrect options:

Provision Reserved Instances for the web, application and database tiers - As explained above, Reserved Instances are not a good fit for running the web and application tiers, so this option is not correct.

Provision Spot Instances for the web and application tiers and Reserved Instances for the database tier - As explained above, Spot Instances are not a good fit for running the web and application tiers, so this option is not correct.

Provision Reserved Instances for the web and application tiers and On-Demand Instances for the database tier - As explained above, Reserved Instances are not a good fit for running the web and application tiers, so this option is not correct. Also using On-Demand Instances for the database tier is not the most cost-effective option as you should use Reserved Instances for the database tier.

Reference:

<https://aws.amazon.com/ec2/pricing/>

Question 31: **Incorrect**

A multi-national retail company has built a hub-and-spoke network with AWS Transit Gateway. VPCs have been provisioned into multiple AWS accounts to facilitate network isolation and to enable delegated network administration. The organization is looking at a cost-effective, quick and secure way of maintaining this distributed architecture so that it provides access to services required by workloads in each of the VPCs.

As a Solutions Architect Professional, which of the following options would you recommend for the given use-case?

- Use Transit VPC to reduce cost and share the resources across VPCs**
-
- Use Fully meshed VPC Peers**
-
- Use VPCs connected with AWS Direct Connect**
(Incorrect)
-
- Use Centralized VPC Endpoints for connecting with multiple VPCs, also known as shared services VPC**
(Correct)

Explanation

Correct option:

Use Centralized VPC Endpoints for connecting with multiple VPCs, also known as shared services VPC - A VPC endpoint allows you to privately connect your VPC to supported AWS services without requiring an Internet gateway, NAT device, VPN connection, or AWS Direct Connect connection. Endpoints are virtual devices that are horizontally scaled, redundant, and highly available VPC components. They allow communication between instances in your VPC and services without imposing availability risks or bandwidth constraints on your network traffic.

VPC endpoints enable you to reduce data transfer charges resulting from network communication between private VPC resources (such as Amazon Elastic Cloud Compute—or EC2—instances) and AWS Services (such as Amazon Quantum Ledger Database, or QLDB). Without VPC endpoints configured, communications that originate from within a VPC destined for public AWS services must egress AWS to the public Internet in order to access AWS services. This network path incurs outbound data transfer charges. Data transfer charges for traffic egressing from Amazon EC2 to the Internet vary based on volume. With VPC endpoints configured, communication between your VPC and the associated AWS service does not leave the Amazon network. If your workload requires you to transfer significant volumes of data between your VPC and AWS, you can reduce costs by leveraging VPC endpoints.

In larger multi-account AWS environments, network design can vary considerably. Consider an organization that has built a hub-and-spoke network with AWS Transit Gateway. VPCs have been provisioned into multiple AWS accounts, perhaps to facilitate network isolation or to enable delegated network administration. When deploying distributed architectures such as this, a popular approach is to build a "shared services VPC, which provides access to services required by workloads in each of the VPCs. This might include directory services or VPC endpoints. Sharing resources from a central location instead of building them in each VPC may reduce administrative overhead and cost.

Centralized VPC Endpoints (multiple VPCs):

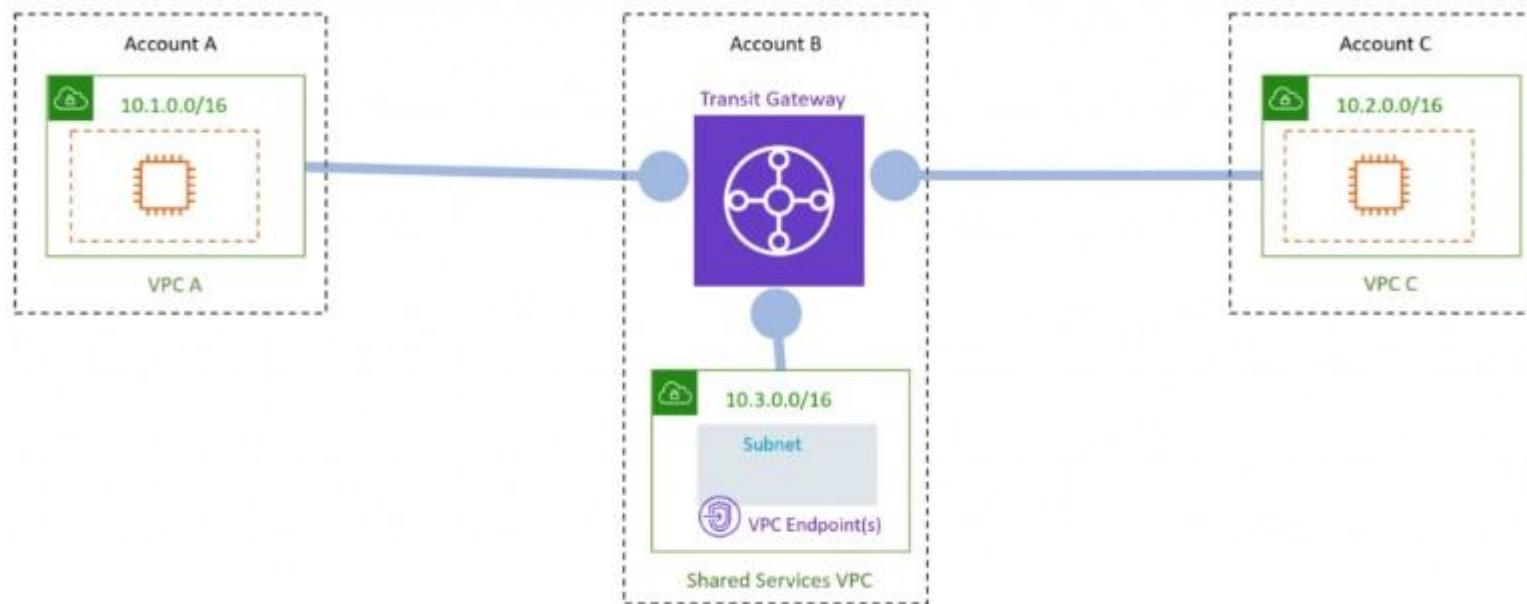


Figure 3: Centralized VPC Endpoints (multiple VPCs)

via -

<https://aws.amazon.com/blogs/architecture/reduce-cost-and-increase-security-with-amazon-vpc-endpoints/>

Incorrect options:

Use Transit VPC to reduce cost and share the resources across VPCs - Transit VPC uses customer-managed Amazon Elastic Compute Cloud (Amazon EC2) VPN instances in a dedicated transit VPC with an Internet gateway. This design requires the customer to deploy, configure, and manage EC2-based VPN appliances, which will result in additional EC2 instances, and potentially third-party product and licensing charges. Note that this design will generate additional data transfer charges for traffic traversing the transit VPC: data is charged when it is sent from a spoke VPC to the transit VPC, and again from the transit VPC to the on-premises network or to a different AWS Region. Transit VPC is not the right choice here because it's not cost-optimal for the given use-case.

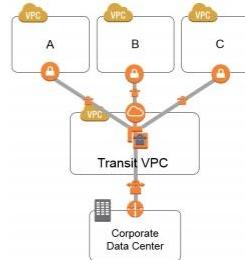
Transit VPC

This approach uses customer-managed Amazon Elastic Compute Cloud (Amazon EC2) VPN instances in a dedicated transit VPC with an Internet gateway. The EC2 instances initiate the VPN connections and route traffic between multiple VPCs and shared-services VPCs. The spoke VPCs can leverage VPC peering to circumvent the transit VPC, providing more scalable, direct access between VPCs.

This design requires the customer to deploy, configure, and manage EC2-based VPN appliances, which will result in additional EC2, and potentially third-party product and licensing charges. Therefore, it is best suited for customers who have already implemented a transit VPC and want to leverage it to manage more advanced connection types, such as inter-region connectivity, or multi-VPC connectivity to on-premises resources.

Note that this design will generate additional data transfer charges for traffic traversing the transit VPC: data is charged when it is sent from a spoke VPC to the transit VPC, and again from the transit VPC to the on-premises network or a different AWS Region.

For additional details on this solution, see the [Multiple-VPC VPN Connection Sharing Solution Brief](#).



via -

More on Transit VPC:

https://d0.awsstatic.com/aws-answers/AWS_Single_Region_Multi_VPC_Connectivity.pdf

Use Fully meshed VPC Peers - This approach creates multiple peering connections to facilitate the sharing of information between resources in different VPCs. This design connects multiple VPCs in a fully meshed configuration, with peering connections between each pair of VPCs. With this configuration, each VPC has access to the resources in all other VPCs. Each peering connection requires modifications to all the other VPCs' route tables and, as the number of VPCs grows, this can be difficult to maintain. And keep in mind that AWS recommends a maximum of 125 peering connections per VPC. It's complex to manage and isn't the right fit for the current scenario.

Configuration Details

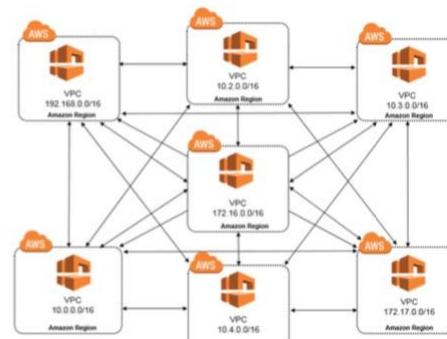
This design connects multiple VPCs in a fully meshed configuration, with peering connections between each pair of VPCs. With this configuration, each VPC has access to the resources in all other VPCs.

To enable the flow of traffic between VPCs, each VPC route table must contain entries that point to the IP address ranges of all the other VPCs in the fully meshed configuration. This design is more complicated to set up than a partially meshed configuration, but it enables communication across all VPCs in the system.

Considerations

Each peering connection requires modifications to all the other VPCs' route tables and, as the number of VPCs grows, this can be difficult to maintain. And keep in mind that AWS recommends a maximum of 125 peering connections per VPC.

Customers can create VPC peering connections between VPCs in the same account, or with VPCs in a different AWS account, as long as the VPCs are in the same region.



via -

More on Fully meshed VPC Peers:

https://d0.awsstatic.com/aws-answers/AWS_Single_Region_Multi_VPC_Connectivity.pdf

Use VPCs connected with AWS Direct Connect - This approach is a good alternative for customers who need to connect a high number of VPCs to a central VPC or to on-premises resources, or who already have an AWS Direct Connect connection in place. This design also offers customers the ability to incorporate transitive routing into their network design. For example, if VPC A and VPC B

are both connected to an on-premises network using AWS Direct Connect connections, then the two VPCs can be connected to each other via AWS Direct Connect. Direct Connect requires physical cables and takes about a month for setting up. This option is not the best fit for the current scenario as there is no on-premises component for the given IT infrastructure.

References:

<https://aws.amazon.com/blogs/architecture/reduce-cost-and-increase-security-with-amazon-vpc-endpoints/>

https://d0.awsstatic.com/aws-answers/AWS_Single_Region_Multi_VPC_Connectivity.pdf

Question 32: **Correct**

A stock trading firm uses AWS Cloud for its IT infrastructure. The firm runs several trading-risk simulation applications, developing complex algorithms to simulate diverse scenarios in order to evaluate the financial health of its customers. The firm stores customers' financial records on Amazon S3. The engineering team needs to implement an archival solution based on Amazon S3 Glacier to enforce regulatory and compliance controls on the archived data.

As a Solutions Architect Professional, which of the following solutions would you recommend?

-

Use S3 Glacier to store the sensitive archived data and then use an S3 Access Control List to enforce compliance controls

-

Use S3 Glacier to store the sensitive archived data and then use an S3 lifecycle policy to enforce compliance controls

-

Use S3 Glacier vault to store the sensitive archived data and then use an S3 Access Control List to enforce compliance controls

-

Use S3 Glacier vault to store the sensitive archived data and then use a vault lock policy to enforce compliance controls

(Correct)

Explanation

Correct option:

Use S3 Glacier vault to store the sensitive archived data and then use a vault lock policy to enforce compliance controls

Amazon S3 Glacier is a secure, durable, and extremely low-cost Amazon S3 cloud storage class for data archiving and long-term backup. It is designed to deliver 99.99999999% durability, and provide comprehensive security and compliance capabilities that can help meet even the most stringent regulatory requirements.

An S3 Glacier vault is a container for storing archives. When you create a vault, you specify a vault name and the AWS Region in which you want to create the vault. S3 Glacier Vault Lock allows you to easily deploy and enforce compliance controls for individual S3 Glacier vaults with a vault lock policy. You can specify controls such as "write once read many" (WORM) in a vault lock policy and lock the policy from future edits. Therefore, this is the correct option.

Incorrect options:

Use S3 Glacier to store the sensitive archived data and then use an S3 lifecycle policy to enforce compliance controls - You can use lifecycle policy to define actions you want Amazon S3 to take during an object's lifetime. For example, use a lifecycle policy to transition objects to another storage class, archive them, or delete them after a specified period. It cannot be used to enforce compliance controls. Therefore, this option is incorrect.

Use S3 Glacier vault to store the sensitive archived data and then use an S3 Access Control List to enforce compliance controls - Amazon S3 access control lists (ACLs) enable you to manage access to buckets and objects. It cannot be used to enforce compliance controls. Therefore, this option is incorrect.

Use S3 Glacier to store the sensitive archived data and then use an S3 Access Control List to enforce compliance controls - Amazon S3 access control lists (ACLs) enable you to manage access to buckets and objects. It cannot be used to enforce compliance controls. Therefore, this option is incorrect.

References:

<https://docs.aws.amazon.com/amazonglacier/latest/dev/working-with-vaults.html>

<https://docs.aws.amazon.com/amazonglacier/latest/dev/vault-lock.html>

<https://docs.aws.amazon.com/AmazonS3/latest/user-guide/create-lifecycle.html>

Question 33: **Incorrect**

The engineering team at a data analytics company is currently optimizing a production workload on AWS that is I/O intensive with frequent read/write/update operations and it's currently constrained on the IOPS. This workload consists of a single-tier with 15 r6g.8xlarge instances, each with 3 TB gp2 volume. The number of processing jobs has increased recently, resulting in an increase in latency as well. The team has concluded that they need to increase the IOPS by 3,000 for each of the instances for the application to perform efficiently.

As an AWS Certified Solutions Architect Professional, which of the following solutions will you suggest to meet the performance goal in the MOST cost-efficient way?

-

Modify the size of the gp2 volume for each instance from 3 TB to 4 TB

(Correct)

-

Set up a new Amazon S3 bucket and migrate all the data to this new bucket. Configure each instance to access this S3 bucket and use it for storage

-

Provision a new EFS file system and migrate all the data to this new file system. Mount this file system on all 15 instances

-

Modify the type of Amazon EBS volume on each instance from gp2 to io1 and set provisioned IOPS to 12,000

(Incorrect)

Explanation

Correct option:

Modify the size of the gp2 volume for each instance from 3 TB to 4 TB

EBS provides block-level storage volumes for use with EC2 instances. EBS is well suited to both database-style applications that rely on random reads and writes, and to throughput-intensive applications that perform long, continuous reads and writes. EBS provides the following volume types: General Purpose SSD (gp2), Provisioned IOPS SSD (io1 and io2), Throughput Optimized HDD (st1), and Cold HDD (sc1).

Volume characteristics

The following table describes the use cases and performance characteristics for each volume type. The default volume type is General Purpose SSD (gp2).

	Solid-state drives (SSD)			Hard disk drives (HDD)	
Volume type	General Purpose SSD (gp2)	Provisioned IOPS SSD		Throughput Optimized HDD (st1)	Cold HDD (sc1)
Description	General purpose SSD volume that balances price and performance for a wide variety of workloads	Highest-performance SSD volume for mission-critical low-latency or high-throughput workloads		Low-cost HDD volume designed for frequently accessed, throughput-intensive workloads	Lowest cost HDD volume designed for less frequently accessed workloads
Durability	99.8% - 99.9% durability (0.1% - 0.2% annual failure rate)	99.999% durability (0.001% annual failure rate)	99.8% - 99.9% durability (0.1% - 0.2% annual failure rate)	99.8% - 99.9% durability (0.1% - 0.2% annual failure rate)	99.8% - 99.9% durability (0.1% - 0.2% annual failure rate)
Use cases	<ul style="list-style-type: none"> Recommended for most workloads System boot volumes Virtual desktops Low-latency interactive apps Development and test environments 	<ul style="list-style-type: none"> Critical business applications that require sustained IOPS performance, or more than 16,000 IOPS or 250 MiB/s of throughput per volume Large database workloads, such as: <ul style="list-style-type: none"> MongoDB Cassandra Microsoft SQL Server MySQL PostgreSQL Oracle 		<ul style="list-style-type: none"> Streaming workloads requiring consistent, fast throughput at a low price Big data Data warehouses Log processing Cannot be a boot volume 	<ul style="list-style-type: none"> Throughput-oriented storage for large volumes of data that is infrequently accessed Scenarios where the lowest storage cost is important Cannot be a boot volume
Amazon EBS Multi-attach	Not supported	Not Supported	Supported	Not supported	Not supported
API name	gp2	io2	io1	st1	sc1
Volume size	1 GiB - 16 TiB	4 GiB - 16 TiB		500 GiB - 16 TiB	500 GiB - 16 TiB
Dominant performance attribute	IOPS	IOPS		MiB/s	MiB/s
Max IOPS per volume	16,000 (16 KiB I/O)*	64,000 (16 KiB I/O) †		500 (1 MiB I/O)	250 (1 MiB I/O)
Max throughput per volume	250 MiB/s *	1,000 MiB/s †		500 MiB/s	250 MiB/s
Max IOPS per instance ‡‡	160,000				
Max throughput per instance ‡‡	4,750 MB/s				

* The throughput limit is between 128 MiB/s and 250 MiB/s, depending on the volume size. Volumes smaller than or equal to 170 GiB deliver a maximum throughput of 128 MiB/s. Volumes larger than 170 GiB but smaller than 334 GiB deliver a maximum throughput of 250 MiB/s if burst credits are available. Volumes larger than or equal to 334 GiB deliver 250 MiB/s regardless of burst credits. Older gp2 volumes might not reach full performance unless you modify the volume. For more information, see [Amazon EBS Elastic Volumes](#).

† Maximum IOPS and throughput are guaranteed only on instances built on the Nitro System provisioned with more than 32,000 IOPS. Other instances guarantee up to 32,000 IOPS and 500 MiB/s. Older io1 volumes might not reach full performance unless you modify the volume. For more information, see [Amazon EBS Elastic Volumes](#).

‡‡ To achieve this throughput, you must have an instance that supports [EBS optimization](#).

via - <https://docs.aws.amazon.com/AWSEC2/latest/UserGuide/ebs-volume-types.html>

gp2 volumes deliver single-digit millisecond latencies and the ability to burst to 3,000 IOPS for extended periods of time. Between a minimum of 100 IOPS (at 33.33 GiB and below) and a maximum of 16,000 IOPS (at 5,334 GiB and above), baseline performance scales linearly at 3 IOPS per GiB of volume size. Therefore, for the given use-case, the engineering team can address the shortfall of 3,000

IOPS by increasing the EBS volume size by 1 TB which will add 3,000 IOPS (3 IOPS per GB * 1000 GB) to the EBS volume on each instance.

General Purpose SSD (gp2) volumes

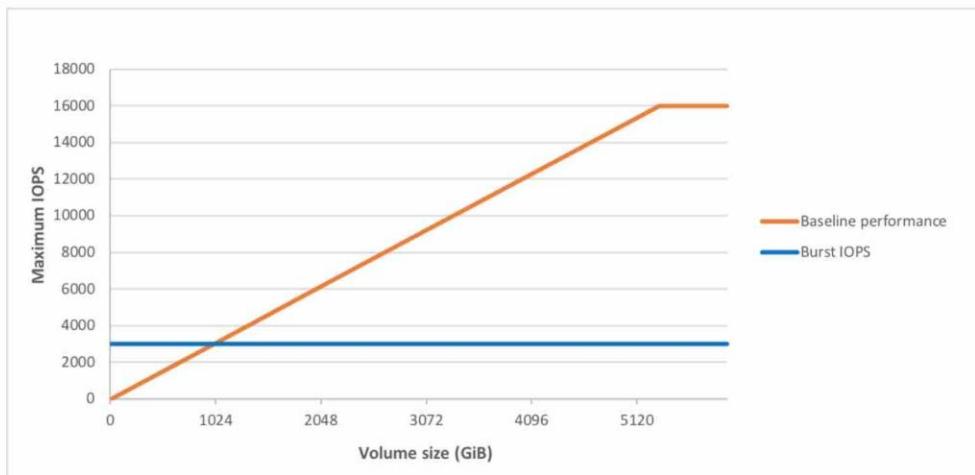
General Purpose SSD (gp2) volumes offer cost-effective storage that is ideal for a broad range of workloads. These volumes deliver single-digit millisecond latencies and the ability to burst to 3,000 IOPS for extended periods of time. Between a minimum of 100 IOPS (at 33.33 GiB and below) and a maximum of 16,000 IOPS (at 5,334 GiB and above), baseline performance scales linearly at 3 IOPS per GiB of volume size. AWS designs gp2 volumes to deliver their provisioned performance 99% of the time. A gp2 volume can range in size from 1 GiB to 16 TiB.

I/O Credits and burst performance

The performance of gp2 volumes is tied to volume size, which determines the baseline performance level of the volume and how quickly it accumulates I/O credits; larger volumes have higher baseline performance levels and accumulate I/O credits faster. I/O credits represent the available bandwidth that your gp2 volume can use to burst large amounts of I/O when more than the baseline performance is needed. The more credits your volume has for I/O, the more time it can burst beyond its baseline performance level and the better it performs when more performance is needed. The following diagram shows the burst-bucket behavior for gp2.



Each volume receives an initial I/O credit balance of 5.4 million I/O credits, which is enough to sustain the maximum burst performance of 3,000 IOPS for 30 minutes. This initial credit balance is designed to provide a fast initial boot cycle for boot volumes and to provide a good bootstrapping experience for other applications. Volumes earn I/O credits at the baseline performance rate of 3 IOPS per GiB of volume size. For example, a 100 GiB gp2 volume has a baseline performance of 300 IOPS.



via - <https://docs.aws.amazon.com/AWSEC2/latest/UserGuide/ebs-volume-types.html>

Incorrect options:

Modify the type of Amazon EBS volume on each instance from gp2 to io1 and set provisioned IOPS to 12,000 - Provisioned IOPS SSD (io1 and io2) volumes are designed to meet the needs of I/O-intensive workloads, particularly database workloads, that are sensitive to storage performance and consistency. You can provision from 100 IOPS up to 64,000 IOPS per volume on Instances built on the Nitro System and up to 32,000 on other instances. The maximum ratio of provisioned IOPS to requested volume size (in GiB) is 50:1 for io1 volumes, and 500:1 for io2 volumes.

Provisioned IOPS SSD (io1 and io2) volumes

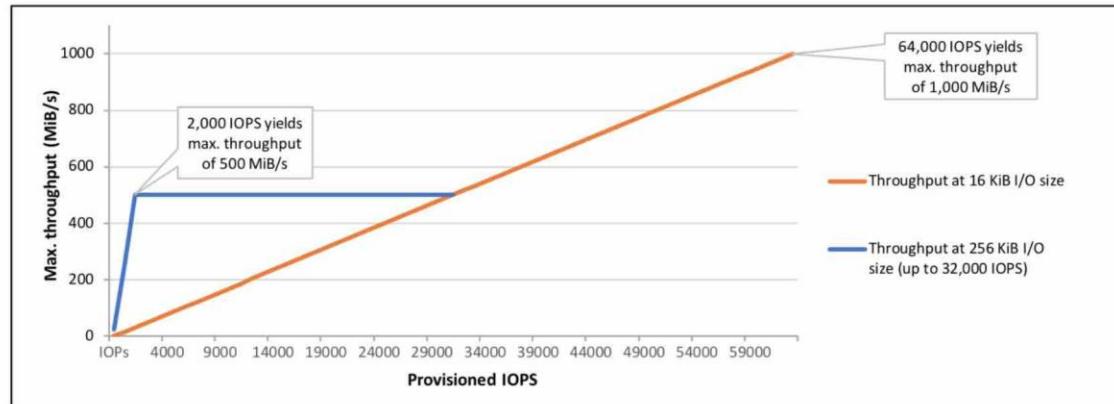
Provisioned IOPS SSD (io1 and io2) volumes are designed to meet the needs of I/O-intensive workloads, particularly database workloads, that are sensitive to storage performance and consistency. Unlike gp2, which uses a bucket and credit model to calculate performance, io1 and io2 volumes allow you to specify a consistent IOPS rate when you create volumes, and Amazon EBS delivers the provisioned performance 99.9 percent of the time.

io1 volumes are designed to provide 99.8 to 99.9 percent volume durability with an annual failure rate (AFR) no higher than 0.2 percent, which translates to a maximum of two volume failures per 1,000 running volumes over a one-year period. io2 volumes are designed to provide 99.999 percent volume durability with an AFR no higher than 0.001 percent, which translates to a single volume failure per 100,000 running volumes over a one-year period.

io1 and io2 volumes can range in size from 4 GiB to 16 TiB. You can provision from 100 IOPS up to 64,000 IOPS per volume on Instances built on the Nitro System and up to 32,000 on other instances. The maximum ratio of provisioned IOPS to requested volume size (in GiB) is 50:1 for io1 volumes, and 500:1 for io2 volumes. For example, a 100 GiB io1 volume can be provisioned with up to 5,000 IOPS, while a 100 GiB io2 volume can be provisioned with up to 50,000 IOPS. On a supported instance type, the following volume sizes allow provisioning up to the 64,000 IOPS maximum:

- io1 volume 1,280 GiB in size or greater ($50 \times 1,280 \text{ GiB} = 64,000 \text{ IOPS}$)
- io2 volume 128 GiB in size or greater ($500 \times 128 \text{ GiB} = 64,000 \text{ IOPS}$)

io1 and io2 volumes provisioned with up to 32,000 IOPS support a maximum I/O size of 256 KiB and yield as much as 500 MiB/s of throughput. With the I/O size at the maximum, peak throughput is reached at 2,000 IOPS. A volume provisioned with more than 32,000 IOPS (up to the cap of 64,000 IOPS) supports a maximum I/O size of 16 KiB and yields as much as 1,000 MiB/s of throughput. The following graph illustrates these performance characteristics:



via - <https://docs.aws.amazon.com/AWSEC2/latest/UserGuide/ebs-volume-types.html>

For the given use-case, io1 does not represent the most cost-optimal solution as it is at least 25% more expensive compared to gp2 for per GB-month of provisioned storage. As identified above, a better solution is to just increase the existing gp2 EBS volume size by 1 TB to account for the additional 3,000 IOPS required for the production workload.

Amazon EBS Volumes

With Amazon EBS, you pay only for what you use. The pricing for Amazon EBS volumes is listed below

General Purpose SSD (gp2) Volumes	\$0.10 per GB-month of provisioned storage
Provisioned IOPS SSD (io2) Volumes	\$0.125 per GB-month of provisioned storage AND \$0.065 per provisioned IOPS-month
Provisioned IOPS SSD (io1) Volumes	\$0.125 per GB-month of provisioned storage AND \$0.065 per provisioned IOPS-month
Throughput Optimized HDD (st1) Volumes	\$0.045 per GB-month of provisioned storage
Cold HDD (sc1) Volumes	\$0.025 per GB-month of provisioned storage

via -

<https://aws.amazon.com/ebs/pricing/>

Provision a new EFS file system and migrate all the data to this new file system. Mount this file system on all 15 instances -
For the given use-case, EFS does not represent the most cost-optimal solution as it is 3 times more expensive compared to gp2 on the basis of per GB storage cost. As identified above, a better solution is to just increase the existing gp2 EBS volume size by 1 TB to account for the additional 3,000 IOPS required for the production workload.

Pricing Table

Region: [US East \(Ohio\) ▾](#)

Standard Storage (GB-Month)	\$0.30
Infrequent Access Storage (GB-Month)	\$0.025
Infrequent Access Requests (per GB transferred)	\$0.01
Provisioned Throughput (MB/s-Month)	\$6.00

via -

<https://aws.amazon.com/efs/pricing/>

Set up a new Amazon S3 bucket and migrate all the data to this new bucket. Configure each instance to access this S3 bucket and use it for storage - Although S3 provides cheap object-based storage, you cannot use S3 for the given high IOPS use-case as the instances need to be able to perform frequent read/write/update operations which cannot be supported out-of-the-box in S3.

Storage pricing

S3 Standard - General purpose storage for any type of data, typically used for frequently accessed data

First 50 TB / Month	\$0.023 per GB
Next 450 TB / Month	\$0.022 per GB
Over 500 TB / Month	\$0.021 per GB

S3 Intelligent - Tiering * - Automatic cost savings for data with unknown or changing access patterns

Frequent Access Tier, First 50 TB / Month	\$0.023 per GB
Frequent Access Tier, Next 450 TB / Month	\$0.022 per GB
Frequent Access Tier, Over 500 TB / Month	\$0.021 per GB
Infrequent Access Tier, All Storage / Month	\$0.0125 per GB
Monitoring and Automation, All Storage / Month	\$0.0025 per 1,000 objects

S3 Standard - Infrequent Access * - For long lived but infrequently accessed data that needs millisecond access

All Storage / Month	\$0.0125 per GB
---------------------	-----------------

S3 One Zone - Infrequent Access * - For re-createable infrequently accessed data that needs millisecond access

All Storage / Month	\$0.01 per GB
---------------------	---------------

S3 Glacier ** - For long-term backups and archives with retrieval option from 1 minute to 12 hours

All Storage / Month	\$0.004 per GB
---------------------	----------------

S3 Glacier Deep Archive ** - For long-term data archiving that is accessed once or twice in a year and can be restored within 12 hours

All Storage / Month	\$0.00099 per GB
---------------------	------------------

via -

<https://aws.amazon.com/s3/pricing/>

References:

<https://docs.aws.amazon.com/AWSEC2/latest/UserGuide/ebs-volume-types.html>

<https://aws.amazon.com/ebs/pricing/>

<https://aws.amazon.com/efs/pricing/>

<https://aws.amazon.com/s3/pricing/>

Question 34: **Correct**

The DevOps team for a CRM SaaS company wants to implement a patching plan on AWS Cloud for a large mixed fleet of Windows and Linux servers. The patching plan has to be auditable and must be implemented securely to ensure compliance with the company's business requirements.

As a Solutions Architect Professional, which of the following options would you recommend to address these requirements with MINIMAL effort? (Select two)

-

Configure OpsWorks automatic patching support for all applications which will keep the OS up-to-date following the initial installation. Set up AWS Config to provide audit and compliance reporting

-

Apply patch baselines using the AWS-ApplyPatchBaseline SSM document

-

Set up an OS-native patching service to manage the update frequency and release approval for all instances. Set up AWS Config to provide audit and compliance reporting

-

Set up Systems Manager Agent on all instances to manage patching. Test patches in pre-production and then deploy as a maintenance window task with the appropriate approval

(Correct)

-

Apply patch baselines using the AWS-RunPatchBaseline SSM document

(Correct)

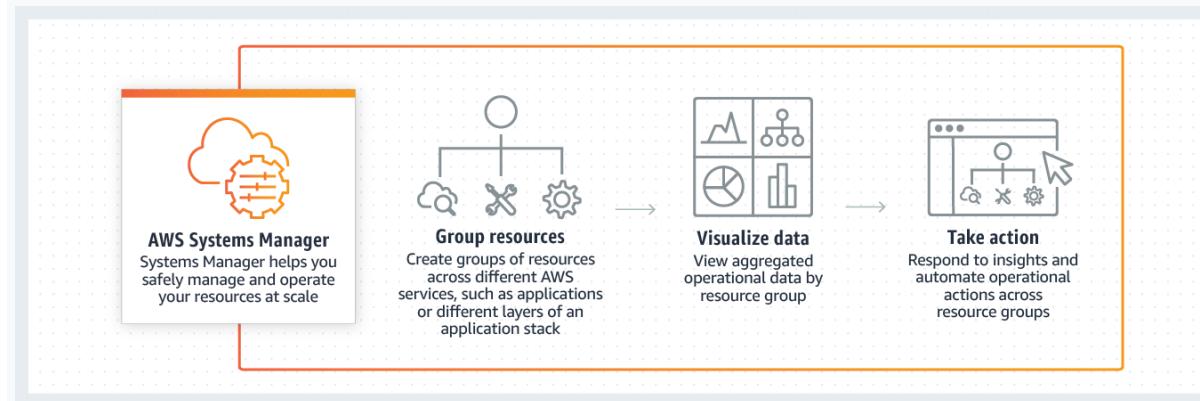
Explanation

Correct options:

Set up Systems Manager Agent on all instances to manage patching. Test patches in pre-production and then deploy as a maintenance window task with the appropriate approval

Apply patch baselines using the AWS-RunPatchBaseline SSM document

Systems Manager is an AWS service that you can use to view and control your infrastructure on AWS. Using the Systems Manager console, you can view operational data from multiple AWS services and automate operational tasks across your AWS resources. Systems Manager helps you maintain security and compliance by scanning your managed instances and reporting on (or taking corrective action on) any policy violations it detects. AWS Systems Manager Agent (SSM Agent) is Amazon software that can be installed and configured on an EC2 instance, an on-premises server, or a virtual machine (VM). SSM Agent makes it possible for Systems Manager to update, manage, and configure these resources.



via - <https://aws.amazon.com/systems-manager/>

You can use Patch Manager to apply patches for both operating systems and applications. (On Windows Server, application support is limited to updates for Microsoft applications). Patch Manager uses patch baselines, which include rules for auto-approving patches within days of their release, as well as a list of approved and rejected patches. You can install patches individually or to large groups of instances by using Amazon EC2 tags.

For the given use-case, you can install patches on a regular basis by scheduling patching to run as a Systems Manager maintenance window task.

Systems Manager supports an SSM document for Patch Manager, AWS-RunPatchBaseline, which performs patching operations on instances for both security-related and other types of updates. When the document is run, it uses the patch baseline currently specified as the "default" for an operating system type.

The AWS-ApplyPatchBaseline SSM document supports patching on Windows instances only and doesn't support Linux instances. For applying patch baselines to both Windows Server and Linux instances, the recommended SSM document is AWS-RunPatchBaseline.

AWS Systems Manager Patch Manager

[PDF](#) | [Kindle](#) | [RSS](#)

AWS Systems Manager Patch Manager automates the process of patching managed instances with both security related and other types of updates. You can use Patch Manager to apply patches for both operating systems and applications. (On Windows Server, application support is limited to updates for Microsoft applications.) You can use Patch Manager to install Service Packs on Windows instances and perform minor version upgrades on Linux instances. You can patch fleets of EC2 instances or your on-premises servers and virtual machines (VMs) by operating system type. This includes supported versions of Windows Server, Amazon Linux, Amazon Linux 2, CentOS, Debian, Oracle Linux, Red Hat Enterprise Linux (RHEL), SUSE Linux Enterprise Server (SLES), and Ubuntu Server. You can scan instances to see only a report of missing patches, or you can scan and automatically install all missing patches.

Important

AWS does not test patches for Windows Server or Linux before making them available in Patch Manager. Also, Patch Manager doesn't support upgrading major versions of operating systems, such as Windows Server 2016 to Windows Server 2019, or SUSE Linux Enterprise Server (SLES) 12.0 to SLES 15.0.

Patch Manager uses *patch baselines*, which include rules for auto-approving patches within days of their release, as well as a list of approved and rejected patches. You can install patches on a regular basis by scheduling patching to run as a Systems Manager maintenance window task. You can also install patches individually or to large groups of instances by using Amazon EC2 tags. (Tags are keys that help identify and sort your resources within your organization.) You can add tags to your patch baselines themselves when you create or update them.

Patch Manager provides options to scan your instances and report compliance on a schedule, install available patches on a schedule, and patch or scan instances on demand whenever you need to.

Patch Manager integrates with AWS Identity and Access Management (IAM), AWS CloudTrail, and Amazon EventBridge to provide a secure patching experience that includes event notifications and the ability to audit usage.

via -

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

Difference 1: Patch evaluation

Linux

For Linux patching, Systems Manager evaluates patch baseline rules and the list of approved and rejected patches on *each* managed instance. Systems Manager must evaluate patching on each instance because the service retrieves the list of known patches and updates from the repositories that are configured on the instance.

Windows

Patch Manager uses different processes on Windows managed instances and Linux managed instances in order to evaluate which patches should be present. For Windows patching, Systems Manager evaluates patch baseline rules and the list of approved and rejected patches *directly in the service*. It can do this because Windows patches are pulled from a single repository (Windows Update).

Difference 2: Not Applicable patches

Due to the large number of available packages for Linux operating systems, Systems Manager does not report details about patches in the *Not Applicable* state. A *Not Applicable* patch is, for example, a patch for Apache software when the instance does not have Apache installed. Systems Manager does report the number of *Not Applicable* patches in the summary, but if you call the [DescribeInstancePatches API](#) for an instance, the returned data does not include patches with a state of *Not Applicable*. This behavior is different from Windows.

Difference 3: SSM document support

The AWS-ApplyPatchBaseline SSM document doesn't support Linux instances. For applying patch baselines to both Windows Server and Linux instances, the recommended SSM document is AWS-RunPatchBaseline. For more information, see [About SSM documents for patching instances](#) and [About the SSM document AWS-RunPatchBaseline](#).

Difference 4: Application patches

Patch Manager's primary focus is applying patches to operating systems. However, you can also use Patch Manager to apply patches to some applications on your instances.

Linux

On Linux operating systems, Patch Manager uses the configured repositories for updates, and does not differentiate between operating systems and application patches. You can use Patch Manager to define which repositories to fetch updates from. For more information, see [How to specify an alternative patch source repository \(Linux\)](#).

Windows

On Windows Server instances, you can apply approval rules, as well as *Approved* and *Rejected* patch exceptions, for applications released by Microsoft, such as Microsoft Word 2011 and Microsoft Exchange Server 2016. For more information, see [Working with custom patch baselines](#).

via -

<https://docs.aws.amazon.com/systems-manager/latest/userguide/sysman-patch-differences.html>

Incorrect options:

Configure OpsWorks automatic patching support for all applications which will keep the OS up-to-date following the initial installation. Set up AWS Config to provide audit and compliance reporting - AWS OpsWorks is a configuration management service that provides managed instances of Chef and Puppet. Chef and Puppet are automation platforms that allow you to use code to automate the configurations of your servers. OpsWorks lets you use Chef and Puppet to automate how servers are configured, deployed, and managed across your Amazon EC2 instances or on-premises compute environments. You cannot use OpsWorks for automatic patch management, so this option is incorrect.

The part about using AWS Config is a distraction. AWS Config is a service that enables you to assess, audit, and evaluate the configurations of your AWS resources. With Config, you can review changes in configurations and relationships between AWS resources, dive into detailed resource configuration histories, and determine your overall compliance against the configurations specified in your internal guidelines. You can use Config to answer questions such as - "What did my AWS resource look like at xyz point in time?".

Set up an OS-native patching service to manage the update frequency and release approval for all instances. Set up AWS Config to provide audit and compliance reporting - You could use an OS-native patching service to manage the update frequency and release approval for all instances but it would take considerable effort to set up and configure this solution. This violates the minimal effort requirement of the given use-case.

Apply patch baselines using the AWS-ApplyPatchBaseline SSM document - As mentioned in the explanation above, the AWS-ApplyPatchBaseline SSM document supports patching on Windows instances only and doesn't support Linux instances. For applying patch baselines to both Windows Server and Linux instances, the recommended SSM document is AWS-RunPatchBaseline.

References:

<https://aws.amazon.com/systems-manager/>

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

<https://docs.aws.amazon.com/systems-manager/latest/userguide/sysman-patch-differences.html>

Question 35: **Correct**

A health and beauty products company processes thousands of orders each day from 100 countries and its website is localized in 15 languages. The company's website faces continual security threats and challenges in the form of HTTP flood attacks, distributed denial of service (DDoS) attacks, rogue robots that flood its website with traffic, SQL-injection attacks designed to extract data and cross-site scripting attacks (XSS). Most of these attacks originate from certain countries. Therefore, the company wants to block access to its application from specific countries; however, the company wants to allow its remote development team (from one of the blocked countries) to have access to the application. The application is deployed on EC2 instances running under an Application Load Balancer (ALB) with AWS WAF.

As a Solutions Architect Professional, which of the following solutions would you suggest as the BEST fit for the given use-case?
(Select two)

- **Use ALB geo match statement listing the countries that you want to block**
- **Use WAF geo match statement listing the countries that you want to block**
(Correct)
- **Create a deny rule for the blocked countries in the NACL associated with each of the EC2 instances**
- **Use WAF IP set statement that specifies the IP addresses that you want to allow through**
(Correct)
-

Use ALB IP set statement that specifies the IP addresses that you want to allow through

Explanation

Correct options:

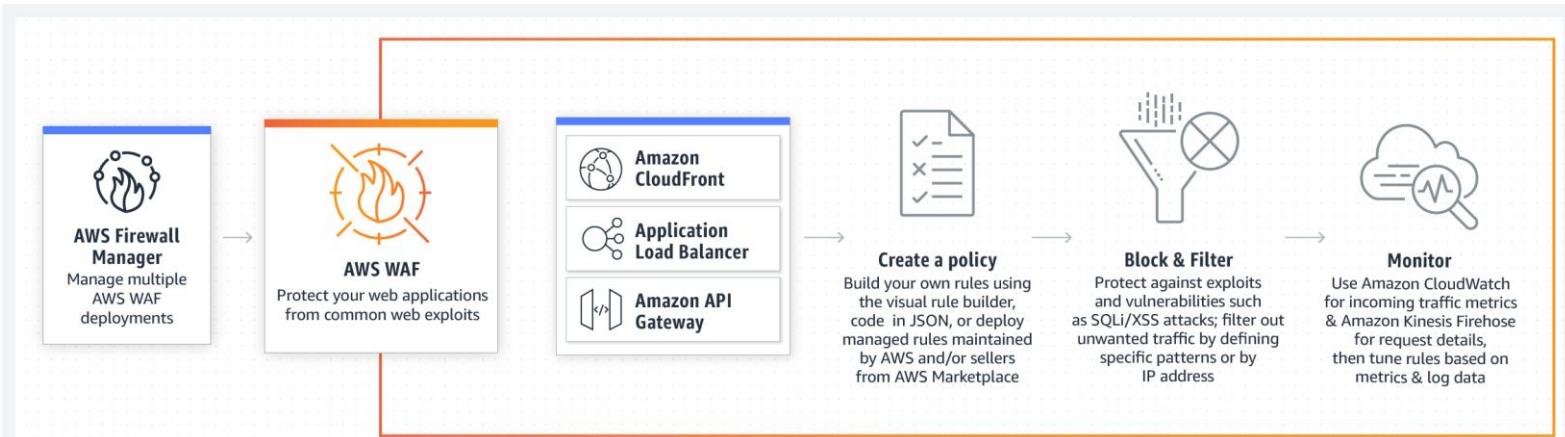
Use WAF geo match statement listing the countries that you want to block

Use WAF IP set statement that specifies the IP addresses that you want to allow through

AWS WAF is a web application firewall that helps protect your web applications or APIs against common web exploits that may affect availability, compromise security, or consume excessive resources. AWS WAF gives you control over how traffic reaches your applications by enabling you to create security rules that block common attack patterns and rules that filter out specific traffic patterns you define.

You can deploy AWS WAF on Amazon CloudFront as part of your CDN solution, the Application Load Balancer that fronts your web servers or origin servers running on EC2, or Amazon API Gateway for your APIs.

AWS WAF - How it



Works

via

- <https://aws.amazon.com/waf/>

To block specific countries, you can create a WAF geo match statement listing the countries that you want to block, and to allow traffic from IPs of the remote development team, you can create a WAF IP set statement that specifies the IP addresses that you want to allow through. You can combine the two rules as shown below:

Geographic match rule statement

[PDF](#) | [Kindle](#) | [RSS](#)

To allow or block web requests based on country of origin, create one or more geographical, or geo, match statements.

Note

If you use the CloudFront geo restriction feature to block a country from accessing your content, any request from that country is blocked and is not forwarded to AWS WAF. So if you want to allow or block requests based on geography plus other AWS WAF criteria, you should *not* use the CloudFront geo restriction feature. Instead, you should use an AWS WAF geo match condition.

You can use this to block access to your site from specific countries or to only allow access from specific countries. If you want to allow some web requests and block others based on country of origin, add a geo match statement for the countries that you want to allow and add a second one for the countries that you want to block.

You can use geo match statements with other AWS WAF statements to build sophisticated filtering. For example, to block certain countries, but still allow requests from a specific set of IP addresses in that country, you could create a rule with the action set to Block and the following nested statements:

- AND statement
 - Geo match statement listing the countries that you want to block
 - NOT statement
 - IP set statement that specifies the IP addresses that you want to allow through

As another example, if you want to prioritize resources for users in a particular country, you could create a different rate-based rules statement for each geo match condition. Set a higher rate limit for users in the preferred country and set a lower rate limit for all other users.

AWS WAF determines the country of origin by resolving the IP address of the web request's origin. If you want to instead use an IP address from an alternate header, like X-Forwarded-For, enable forwarded IP configuration.

via -

<https://docs.aws.amazon.com/waf/latest/developerguide/waf-rule-statement-type-geo-match.html>

Incorrect options:

Create a deny rule for the blocked countries in the NACL associated to each of the EC2 instances - A network access control list (ACL) is an optional layer of security for your VPC that acts as a firewall for controlling traffic in and out of one or more subnets. NACL does not have the capability to block traffic based on geographic match conditions.

Use ALB geo match statement listing the countries that you want to block

Use ALB IP set statement that specifies the IP addresses that you want to allow through

An Application Load Balancer (ALB) operates at the request level (layer 7), routing traffic to targets – EC2 instances, containers, IP addresses, and Lambda functions based on the content of the request. Ideal for advanced load balancing of HTTP and HTTPS traffic, Application Load Balancer provides advanced request routing targeted at the delivery of modern application architectures, including microservices and container-based applications.

An ALB cannot block or allow traffic based on geographic match conditions or IP based conditions. Both these options have been added as distractors.

References:

<https://docs.aws.amazon.com/waf/latest/developerguide/waf-rule-statement-type-geo-match.html>

<https://aws.amazon.com/blogs/security/how-to-use-aws-waf-to-filter-incoming-traffic-from-embargoed-countries/>

Question 36: Correct

A company wants to use SharePoint to deploy a content and collaboration platform with document and records management functionality. The company wants to establish an AWS Direct Connect link to connect the AWS Cloud with the internal corporate network using AWS Storage Gateway. Using AWS Direct Connect would enable the company to deliver on its performance benchmark requirements including a three second or less response time for sending small documents across the internal network. To facilitate this goal, the company wants to be able to resolve DNS queries for any resources in the on-premises network from the AWS VPC and also resolve any DNS queries for resources in the AWS VPC from the on-premises network.

As a Solutions Architect Professional, which of the following solutions would you recommend for this use-case? (Select two)

-

Create an outbound endpoint on Route 53 Resolver and then DNS resolvers on the on-premises network can forward DNS queries to Route 53 Resolver via this endpoint

-

Create an outbound endpoint on Route 53 Resolver and then Route 53 Resolver can conditionally forward queries to resolvers on the on-premises network via this endpoint

(Correct)

- Create an universal endpoint on Route 53 Resolver and then Route 53 Resolver can receive and forward queries to resolvers on the on-premises network via this endpoint

Create a universal endpoint on Route 53 Resolver and then Route 53 Resolver can receive and forward queries to resolvers on the on-premises network via this endpoint

- Create an inbound endpoint on Route 53 Resolver and then Route 53 Resolver can conditionally forward queries to resolvers on the on-premises network via this endpoint

Create an inbound endpoint on Route 53 Resolver and then Route 53 Resolver can conditionally forward queries to resolvers on the on-premises network via this endpoint

- Create an inbound endpoint on Route 53 Resolver and then DNS resolvers on the on-premises network can forward DNS queries to Route 53 Resolver via this endpoint

(Correct)

Explanation

Correct options:

Create an inbound endpoint on Route 53 Resolver and then DNS resolvers on the on-premises network can forward DNS queries to Route 53 Resolver via this endpoint

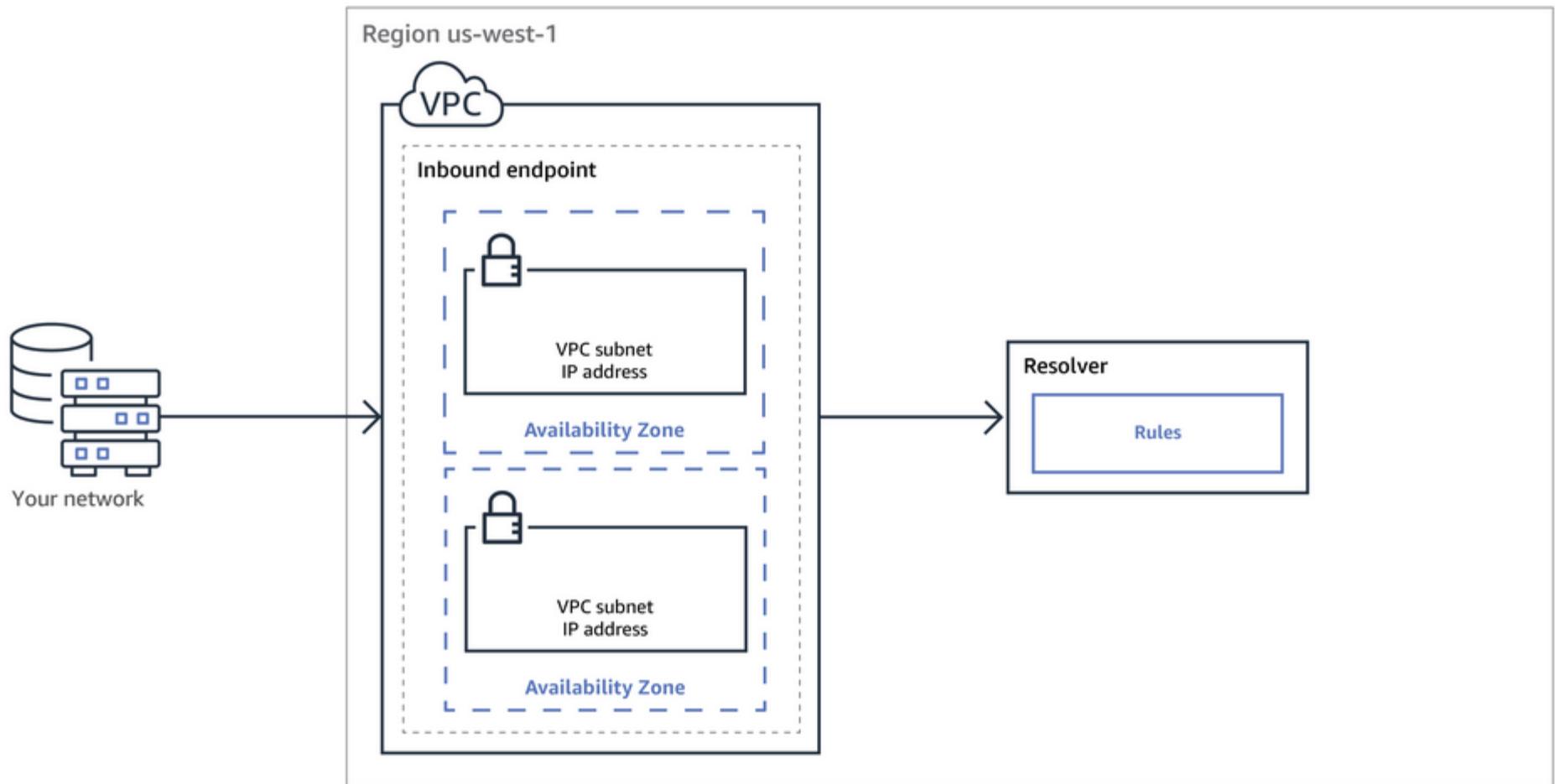
Create an outbound endpoint on Route 53 Resolver and then Route 53 Resolver can conditionally forward queries to resolvers on the on-premises network via this endpoint

Amazon Route 53 is a highly available and scalable cloud Domain Name System (DNS) web service. Amazon Route 53 effectively connects user requests to infrastructure running in AWS – such as Amazon EC2 instances – and can also be used to route users to

infrastructure outside of AWS. By default, Route 53 Resolver automatically answers DNS queries for local VPC domain names for EC2 instances. You can integrate DNS resolution between Resolver and DNS resolvers on your on-premises network by configuring forwarding rules.

To resolve any DNS queries for resources in the AWS VPC from the on-premises network, you can create an inbound endpoint on Route 53 Resolver and then DNS resolvers on the on-premises network can forward DNS queries to Route 53 Resolver via this endpoint.

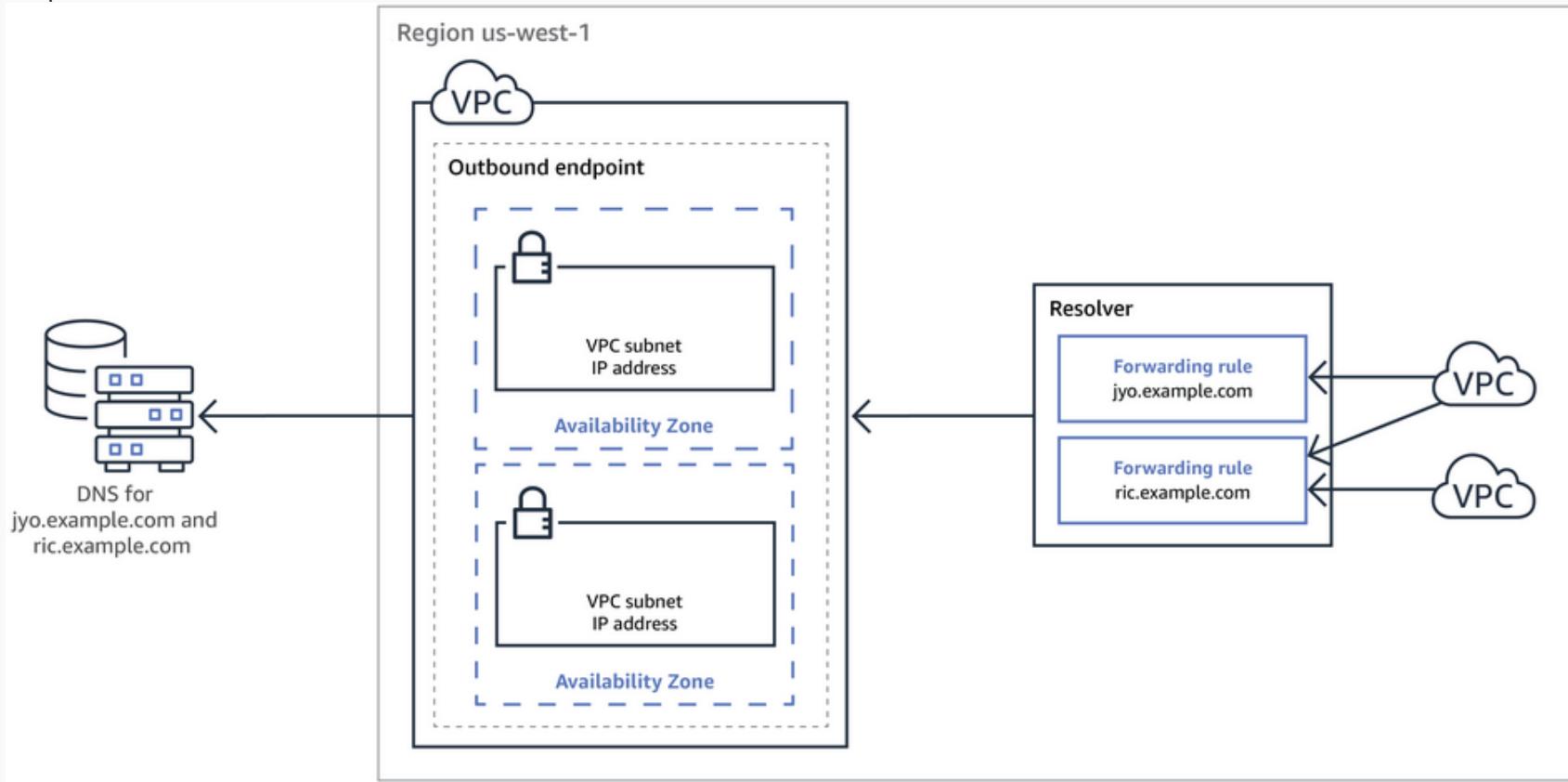
Resolver Inbound Endpoint



via - <https://docs.aws.amazon.com/Route53/latest/DeveloperGuide/resolver.html>

To resolve DNS queries for any resources in the on-premises network from the AWS VPC, you can create an outbound endpoint on Route 53 Resolver and then Route 53 Resolver can conditionally forward queries to resolvers on the on-premises network via this endpoint. To conditionally forward queries, you need to create Resolver rules that specify the domain names for the DNS queries that you want to forward (such as example.com) and the IP addresses of the DNS resolvers on the on-premises network that you want to forward the queries to.

Resolver Outbound Endpoint



ia - <https://docs.aws.amazon.com/Route53/latest/DeveloperGuide/resolver.html>

Incorrect options:

Create an outbound endpoint on Route 53 Resolver and then DNS resolvers on the on-premises network can forward DNS queries to Route 53 Resolver via this endpoint - DNS resolvers on the on-premises network can forward DNS queries to Route 53 Resolver via an inbound endpoint. Hence, this option is incorrect.

Create an inbound endpoint on Route 53 Resolver and then Route 53 Resolver can conditionally forward queries to resolvers on the on-premises network via this endpoint - Route 53 Resolver can conditionally forward queries to resolvers on the on-premises network via an outbound endpoint. Hence, this option is incorrect.

Create a universal endpoint on Route 53 Resolver and then Route 53 Resolver can receive and forward queries to resolvers on the on-premises network via this endpoint - There is no such thing as a universal endpoint on Route 53 Resolver. This option has been added as a distractor.

References:

<https://docs.aws.amazon.com/Route53/latest/DeveloperGuide/resolver-getting-started.html>

<https://docs.aws.amazon.com/Route53/latest/DeveloperGuide/resolver.html>

<https://aws.amazon.com/blogs/networking-and-content-delivery/automating-dns-infrastructure-using-route-53-resolver-endpoints/>

Question 37: **Incorrect**

A multi-national bank has recently migrated to AWS Cloud to utilize dedicated instances that are physically isolated at the host hardware level from instances that belong to other AWS accounts. The bank's flagship application is hosted on a fleet of EC2 instances which are part of an Auto Scaling group (ASG). The ASG uses a Launch Configuration (LC-A) with "dedicated" instance placement tenancy but the VPC (VPC-A) used by the Launch Configuration LC-A has the instance tenancy set to default. Later the engineering team creates a new Launch Configuration (LC-B) with "default" instance placement tenancy but the VPC (VPC-B) used by the Launch Configuration LC-B has the instance tenancy set to dedicated.

As a Solutions Architect Professional, which of the following options would you identify as correct regarding the instances launched via Launch Configuration LC-A and Launch Configuration LC-B?



The instances launched by Launch Configuration LC-A will have dedicated instance tenancy while the instances launched by the Launch Configuration LC-B will have default instance tenancy

(Incorrect)

-
- The instances launched by Launch Configuration LC-A will have default instance tenancy while the instances launched by the Launch Configuration LC-B will have dedicated instance tenancy**
-
- The instances launched by both Launch Configuration LC-A and Launch Configuration LC-B will have default instance tenancy**
-
- The instances launched by both Launch Configuration LC-A and Launch Configuration LC-B will have dedicated instance tenancy**

(Correct)

Explanation

Correct option:

The instances launched by both Launch Configuration LC-A and Launch Configuration LC-B will have dedicated instance tenancy

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

When you create a launch configuration, the default value for the instance placement tenancy is null and the instance tenancy is controlled by the tenancy attribute of the VPC. If you set the Launch Configuration Tenancy to default and the VPC Tenancy is set to dedicated, then the instances have dedicated tenancy. If you set the Launch Configuration Tenancy to dedicated and the VPC Tenancy is set to default, then again the instances have dedicated tenancy.

Launch Configuration Tenancy	VPC Tenancy = default	VPC Tenancy = dedicated
not specified	shared-tenancy instance	dedicated instance
default	shared-tenancy instance	dedicated instance
dedicated	dedicated instance	dedicated instance

Launch Configuration Tenancy vs VPC Tenancy

<https://docs.aws.amazon.com/autoscaling/ec2/userguide/asg-in-vpc.html#as-vpc-tenancy>

via -

Incorrect options:

The instances launched by Launch Configuration LC-A will have dedicated instance tenancy while the instances launched by the Launch Configuration LC-B will have default instance tenancy - If either Launch Configuration Tenancy or VPC Tenancy is set to dedicated, then the instance tenancy is also dedicated. Therefore, this option is incorrect.

The instances launched by Launch Configuration LC-A will have default instance tenancy while the instances launched by the Launch Configuration LC-B will have dedicated instance tenancy - If either Launch Configuration Tenancy or VPC Tenancy is set to dedicated, then the instance tenancy is also dedicated. Therefore, this option is incorrect.

The instances launched by both Launch Configuration LC-A and Launch Configuration LC-B will have default instance tenancy - If either Launch Configuration Tenancy or VPC Tenancy is set to dedicated, then the instance tenancy is also dedicated. Therefore, this option is incorrect.

References:

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

<https://docs.aws.amazon.com/autoscaling/ec2/userguide/asg-in-vpc.html#as-vpc-tenancy>

Question 38: **Incorrect**

The engineering team at a social media company is building an ElasticSearch based index for all the existing files in S3. To build this index, it only needs to read the first 250 bytes of each object in S3, which contains some metadata about the content of the file itself. There are over 100,000 files in your S3 bucket, adding up to 50TB of data.

As a Solutions Architect Professional, which of the following solutions can be used to build this index MOST efficiently? (Select two)

-

Create an application that will traverse the S3 bucket, read the entire files one by one, extract the first 250 bytes, and store that information in ElasticSearch

(Incorrect)

-

Create an application that will traverse the S3 bucket, issue a Byte Range Fetch for the first 250 bytes, and store that information in ElasticSearch

(Correct)

-

Use the Database Migration Service to load the entire data from S3 to ElasticSearch and then ElasticSearch would automatically build the index

-

Create an application that will use the S3 Select ScanRange parameter to get the first 250 bytes and store that information in ElasticSearch

(Correct)

-

Use the ElasticSearch Import feature to load the entire data from S3 to ElasticSearch and then ElasticSearch would automatically build the index

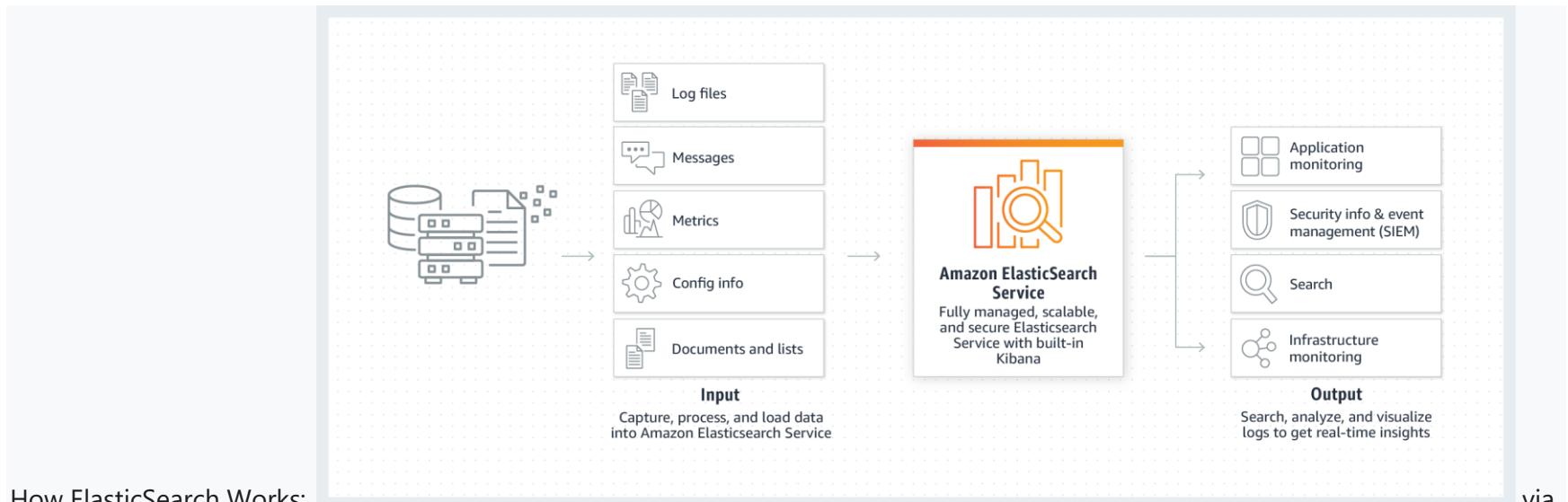
Explanation

Correct options:

Create an application that will traverse the S3 bucket, issue a Byte Range Fetch for the first 250 bytes, and store that information in ElasticSearch

Amazon Simple Storage Service (Amazon S3) is an object storage service that offers industry-leading scalability, data availability, security, and performance.

Amazon Elasticsearch Service (Amazon ES) is a managed service that makes it easy to deploy, operate, and scale Elasticsearch clusters in the AWS Cloud. Elasticsearch is a popular open-source search and analytics engine for use cases such as log analytics, real-time application monitoring, and clickstream analysis. With Amazon ES, you get direct access to the Elasticsearch APIs; existing code and applications work seamlessly with the service.



How ElasticSearch Works:

- <https://aws.amazon.com/elasticsearch-service/>

Using the Range HTTP header in a GET Object request, you can fetch a byte-range from an object, transferring only the specified portion. You can use concurrent connections to Amazon S3 to fetch different byte ranges from within the same object. This helps you achieve higher aggregate throughput versus a single whole-object request. Fetching smaller ranges of a large object also allows your application to improve retry times when requests are interrupted.

A byte-range request is a perfect way to get the beginning of a file and ensuring we remain efficient during our scan of our S3 bucket. You can then store the relevant information in the form of a JSON document in ElasticSearch.

Create an application that will use the S3 Select ScanRange parameter to get the first 250 bytes and store that information in ElasticSearch

With Amazon S3 Select, you can scan a subset of an object by specifying a range of bytes to query using the ScanRange parameter. This capability lets you parallelize scanning the whole object by splitting the work into separate Amazon S3 Select requests for a series of non-overlapping scan ranges. Use the Amazon S3 Select ScanRange parameter and Start at (Byte) and End at (Byte). You can then store the relevant information in the form of a JSON document in ElasticSearch.

Requests using scan ranges

With Amazon S3 Select, you can scan a subset of an object by specifying a range of bytes to query. This capability lets you parallelize scanning the whole object by splitting the work into separate Amazon S3 Select requests for a series of non-overlapping scan ranges. Scan ranges don't need to be aligned with record boundaries. An Amazon S3 Select scan range request runs across the byte range that you specify. A record that starts within the scan range specified but extends beyond the scan range will be processed by the query. For example; the following shows an Amazon S3 object containing a series of records in a line-delimited CSV format:

A,B
C,D
D,E
E,F
G,H
I,J

Use the Amazon S3 Select ScanRange parameter and Start at (Byte) 1 and End at (Byte) 4. So the scan range would start at "," and scan till the end of record starting at "C" and return the result C, D because that is the end of the record.

Amazon S3 Select scan range requests support Parquet, CSV (without quoted delimiters), and JSON objects (in LINES mode only). CSV and JSON objects must be uncompressed. For line-based CSV and JSON objects, when a scan range is specified as part of the Amazon S3 Select request, all records that start within the scan range are processed. For Parquet objects, all of the row groups that start within the scan range requested are processed.

Amazon S3 Select scan range requests are available to use on the Amazon S3 CLI, API and SDK. You can use the ScanRange parameter in the Amazon S3 Select request for this feature. For more information, see the [Amazon S3 SELECT Object Content](#) in the *Amazon Simple Storage Service API Reference*.

via -

<https://docs.aws.amazon.com/AmazonS3/latest/dev/selecting-content-from-objects.html>

Incorrect options:

Use the ElasticSearch Import feature to load the entire data from S3 to ElasticSearch and then ElasticSearch would automatically build the index - This option has been added as a distractor as there is no ElasticSearch Import feature to load data from S3.

Create an application that will traverse the S3 bucket, read the entire files one by one, extract the first 250 bytes, and store that information in ElasticSearch - If you build an application that loads all the files from S3, that would work, but you would read 50TB of data and that may be very expensive and slow. So this option is incorrect.

Use the Database Migration Service to load the entire data from S3 to ElasticSearch and then ElasticSearch would automatically build the index - Although you could use Database Migration Service to load the entire data from S3 to ElasticSearch, but you would read 50TB of data and that may be very expensive and slow. So this option is incorrect.

References:

<https://docs.aws.amazon.com/AmazonS3/latest/dev/optimizing-performance-guidelines.html#optimizing-performance-guidelines-get-range>

<https://docs.aws.amazon.com/AmazonS3/latest/dev/selecting-content-from-objects.html>

<https://aws.amazon.com/elasticsearch-service/>

Question 39: **Incorrect**

A leading hotel reviews website has a repository of more than one million high-quality digital images. When this massive volume of images became too cumbersome to handle in-house, the company decided to offload the content to a central repository on Amazon S3 as part of its hybrid cloud strategy. The company now wants to reprocess its entire collection of photographic images to change the watermarks. The company wants to use Amazon EC2 instances and Amazon SQS in an integrated workflow to generate the sizes they need for each photo. The team wants to process a few thousand photos each night, using Amazon EC2 Spot Instances. The team uses Amazon SQS to communicate the photos that need to be processed and the status of the jobs. To handle certain sensitive photos, the team wants to postpone the delivery of certain messages to the queue by one minute while all other messages need to be delivered immediately to the queue.

As a Solutions Architect Professional, which of the following solutions would you suggest to the company to handle the workflow for sensitive photos?

-

Use message timers to postpone the delivery of certain messages to the queue by one minute

(Correct)



Use dead-letter queues to postpone the delivery of certain messages to the queue by one minute

(Incorrect)



Use visibility timeout to postpone the delivery of certain messages to the queue by one minute



Use delay queues to postpone the delivery of certain messages to the queue by one minute

Explanation

Correct option:

Use message timers to postpone the delivery of certain messages to the queue by one minute

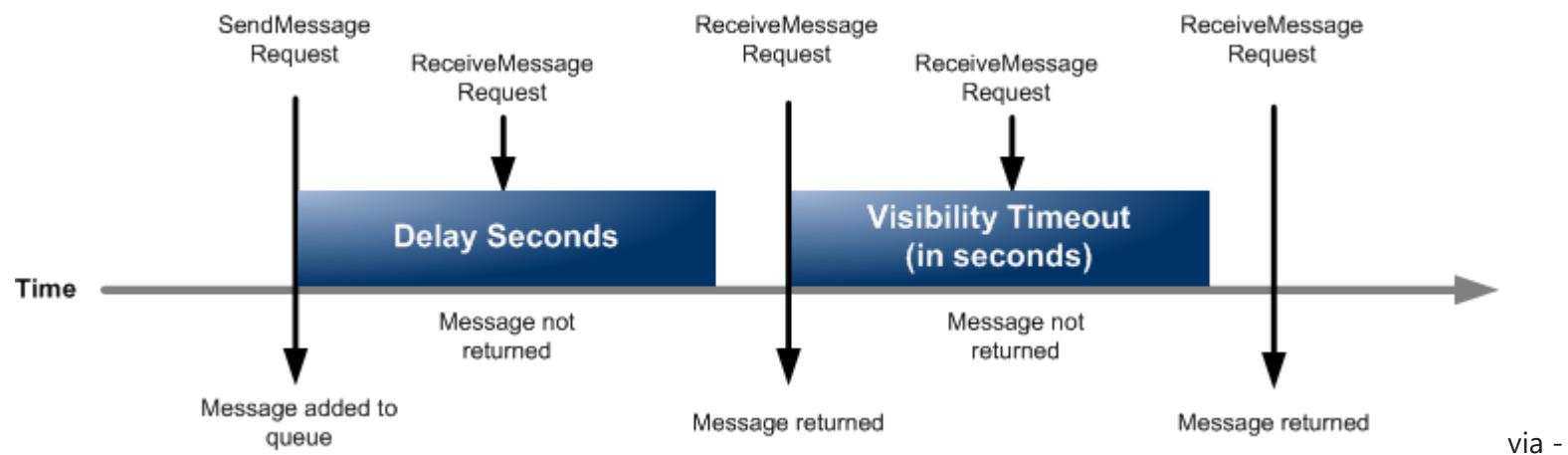
You can use message timers to set an initial invisibility period for a message added to a queue. So, if you send a message with a 60-second timer, the message isn't visible to consumers for its first 60 seconds in the queue. The default (minimum) delay for a message is 0 seconds. The maximum is 15 minutes. Therefore, you should use message timers to postpone the delivery of certain messages to the queue by one minute.

Incorrect options:

Use dead-letter queues to postpone the delivery of certain messages to the queue by one minute - Dead-letter queues can be used by other queues (source queues) as a target for messages that can't be processed (consumed) successfully. Dead-letter queues are useful for debugging your application or messaging system because they let you isolate problematic messages to determine why their processing doesn't succeed. You cannot use dead-letter queues to postpone the delivery of certain messages to the queue by one minute.

Use visibility timeout to postpone the delivery of certain messages to the queue by one minute - Visibility timeout is a period during which Amazon SQS prevents other consumers from receiving and processing a given message. The default visibility timeout for a message is 30 seconds. The minimum is 0 seconds. The maximum is 12 hours. You cannot use visibility timeout to postpone the delivery of certain messages to the queue by one minute.

Use delay queues to postpone the delivery of certain messages to the queue by one minute - Delay queues let you postpone the delivery of all new messages to a queue for several seconds, for example, when your consumer application needs additional time to process messages. If you create a delay queue, any messages that you send to the queue remain invisible to consumers for the duration of the delay period. The default (minimum) delay for a queue is 0 seconds. The maximum is 15 minutes. You cannot use delay queues to postpone the delivery of only certain messages to the queue by one minute.



<https://docs.aws.amazon.com/AWSSimpleQueueService/latest/SQSDriverGuide/sqs-delay-queues.html>

References:

<https://docs.aws.amazon.com/AWSSimpleQueueService/latest/SQSDriverGuide/sqs-message-timers.html>

<https://docs.aws.amazon.com/AWSSimpleQueueService/latest/SQSDriverGuide/sqs-delay-queues.html>

<https://docs.aws.amazon.com/AWSSimpleQueueService/latest/SQSDriverGuide/sqs-dead-letter-queues.html>

<https://docs.aws.amazon.com/AWSSimpleQueueService/latest/SQSDeveloperGuide/sqs-visibility-timeout.html>

Question 40: **Correct**

A global apparel, footwear, and accessories retailer uses Amazon S3 for centralized storage of the static media assets such as images and videos for its products. The product planning specialists typically upload and download video files (about 100MB each) to the same S3 bucket as part of their day to day work. Initially, the product planning specialists were based out of a single region and there were no performance issues. However, as the company grew and started running offices from multiple countries, it resulted in poor latency while accessing data from S3 and uploading data to S3. The company wants to continue with the serverless solution for its storage requirements but wants to improve its performance.

As a solutions architect, which of the following solutions do you propose to address this issue? (Select two)

- **Use Amazon CloudFront distribution with origin as the S3 bucket. This would speed up uploads as well as downloads for the video files**
(Correct)
- **Create new S3 buckets in every region where the company has an office, so that each office can maintain its storage for the media assets**
- **Spin up EC2 instances in each region where the company has an office. Create a daily job to transfer S3 data into EBS volumes attached to the EC2 instances**
- **Move S3 data into EFS file system created in a US region, connect to EFS file system from EC2 instances in other AWS regions using an inter-region VPC peering connection**

-

Enable Amazon S3 Transfer Acceleration for the S3 bucket. This would speed up uploads as well as downloads for the video files

(Correct)

Explanation

Correct options:

Use Amazon CloudFront distribution with origin as the S3 bucket. This would speed up uploads as well as downloads for the video files

Amazon CloudFront is a fast content delivery network (CDN) service that securely delivers data, videos, applications, and APIs to customers globally with low latency, high transfer speeds, within a developer-friendly environment.

When an object from S3 that is set up with CloudFront CDN is requested, the request would come through the Edge Location transfer paths only for the first request. Thereafter, it would be served from the nearest edge location to the users until it expires. For uploads, you can use the POST and PUT methods for your CloudFront distribution to accelerate content uploads to the origin, which is S3 for the given use-case. So in this way, you can speed up uploads as well as downloads for the video files.

Please review this excellent reference blog on optimizing uploads via CloudFront: <https://aws.amazon.com/blogs/aws/amazon-cloudfront-content-uploads-post-put-other-methods/>

Enable Amazon S3 Transfer Acceleration for the S3 bucket. This would speed up uploads as well as downloads for the video files

Amazon S3 Transfer Acceleration can speed up content transfers to and from Amazon S3 by as much as 50-500% for long-distance transfer of larger objects. Transfer Acceleration takes advantage of Amazon CloudFront's globally distributed edge locations. As the data arrives at an edge location, data is routed to Amazon S3 over an optimized network path. So this option is also correct.

Amazon S3 Transfer Acceleration can speed up content transfers to and from Amazon S3 by as much as 50-500% for long-distance transfer of larger objects. Customers who have either web or mobile applications with widespread users or applications hosted far away from their S3 bucket can experience long and variable upload and download speeds over the Internet. S3 Transfer Acceleration (S3TA) reduces the variability in Internet routing, congestion and speeds that can affect transfers, and logically shortens the distance to S3 for remote applications. **S3TA improves transfer performance by routing traffic through Amazon CloudFront's globally distributed Edge Locations and over AWS backbone networks, and by using network protocol optimizations.** You can turn on S3TA with a few clicks in the S3 console, and test its benefits from your location with [a speed comparison tool](#). **With S3TA, you pay only for transfers that are accelerated.**

Benefits

Move data faster over long distances

S3TA can accelerate long-distance transfers to and from your Amazon S3 buckets. The longer the distance between your client application (mobile, web application, or upload tool) and the target S3 bucket, the more S3TA can help. And if S3TA would not accelerate a transfer, you are not charged.

Reduce network variability

For applications interacting with your S3 buckets through the S3 API from outside of your bucket's region, S3TA helps avoid the variability in Internet routing and congestion. It does this by routing your uploads and downloads over the AWS global network infrastructure, so you get the benefit of our network optimizations.

Shorten the distance to S3

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

Maximize bandwidth utilization

S3TA on average fully utilizes your bandwidth for transfers, and minimizes the effect of distance on throughput. This helps to ensure consistently fast performance to Amazon S3 regardless of your client's location.

via -

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

Incorrect options:

Create new S3 buckets in every region where the company has an office, so that each office can maintain its storage for the media assets - Creating new S3 buckets in every region is not an option, since the company maintains centralized storage. Hence this option is incorrect.

Move S3 data into EFS file system created in a US region, connect to EFS file system from EC2 instances in other AWS regions using an inter-region VPC peering connection

Spin up EC2 instances in each region where the company has an office. Create a daily job to transfer S3 data into EBS volumes attached to the EC2 instances

Both these options using EC2 instances are not correct for the given use-case, as the company wants a serverless storage solution.

References:

<https://aws.amazon.com/blogs/aws/amazon-cloudfront-content-uploads-post-put-other-methods/>

<https://docs.aws.amazon.com/AmazonCloudFront/latest/DeveloperGuide/DownloadDistS3AndCustomOrigins.html>

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

<https://aws.amazon.com/s3/transfer-acceleration/>

Question 41: **Incorrect**

An e-commerce company runs a data archival workflow once a month for its on-premises data center which is connected to the AWS Cloud over a minimally used 10-Gbps Direct Connect connection using a private virtual interface to its virtual private cloud (VPC). The company internet connection is 200 Mbps, and the usual archive size is around 140 TB that is created on the first Friday of a month. The archive must be transferred and available in Amazon S3 by the next Monday morning.

As a Solutions Architect Professional, which of the following options would you recommend as the LEAST expensive way to address the given use-case?



Configure a VPC endpoint for S3 and then leverage the Direct Connect connection for data transfer with VPC endpoint as the target

(Incorrect)

-
- Configure a private virtual interface on the 10-Gbps Direct Connect connection and then copy the data securely to S3 over the connection**
-
- Order multiple AWS Snowball Edge appliances, transfer the data in parallel to these appliances and ship them to AWS which will then copy the data from the Snowball Edge appliances to S3**
-
- Configure a public virtual interface on the 10-Gbps Direct Connect connection and then copy the data to S3 over the connection**

(Correct)

Explanation

Correct option:

Configure a public virtual interface on the 10-Gbps Direct Connect connection and then copy the data to S3 over the connection

AWS Direct Connect links your on-premises data center to an AWS Direct Connect location over a standard Ethernet fiber-optic cable. One end of the cable is connected to your router, the other to an AWS Direct Connect router. With this connection, you can create virtual interfaces directly to public AWS services (for example, to Amazon S3) or to Amazon VPC, bypassing internet service providers in your network path. An AWS Direct Connect location provides access to AWS in the Region with which it is associated.

There are two types of Direct Connect connections:

Dedicated Connection: A physical Ethernet connection associated with a single customer. Customers can request a dedicated connection through the AWS Direct Connect console, the CLI, or the API. This supports speed of 1Gbps and 10Gbps.

Hosted Connection: A physical Ethernet connection that an AWS Direct Connect Partner provisions on behalf of a customer. Customers request a hosted connection by contacting a partner in the AWS Direct Connect Partner Program, who provisions the connection. This supports speed of 50Mbps, 100Mbps, 200Mbps, 300Mbps, 400Mbps, 500Mbps, 1Gbps, 2Gbps, 5Gbps, and 10Gbps.

via - <https://docs.aws.amazon.com/directconnect/latest/UserGuide/Welcome.html>

The given use-case requires transferring 140 TB of archived data within a duration of approximately two days (48 hours). Now, the Direct Connect connection assures a speed of 10Gbps (Gigabits per second) as it is minimally used. As 1 Byte = 8 bits, therefore you can transfer approximately 1GBps (Gigabytes per second).

So the hourly data transfer is $1 * 60 * 60 = 3,600 \text{ GB}$ or approximately 3.6 TB.

So the approximate daily data transfer is $3.6 * 24 = \sim 86 \text{ TB}$

Therefore, the entire archived dataset of 140 TB can be transferred in less than two days (48 hours)

Additionally, you should configure a public virtual interface from the Direct Connect connection to connect to AWS resources that are reachable by a public IP address (such as an S3 bucket).

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

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

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

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

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

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

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

Note: For transit virtual interface, AWS advertises only routes that you specify in the allowed prefixes list on the DX gateway. For a list of all AWS Regions that offer DX support for AWS Transit Gateway, see [AWS Transit Gateway Support](#) under Direct Connect FAQs.

via -

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

Incorrect options:

Configure a private virtual interface on the 10-Gbps Direct Connect connection and then copy the data securely to S3 over the connection - You can only use a private virtual interface to connect to your resources hosted in an Amazon Virtual Private Cloud (Amazon VPC) using their private IP addresses. You must use a public virtual interface from the Direct Connect connection to connect to an S3 bucket.

Configure a VPC endpoint for S3 and then leverage the Direct Connect connection for data transfer with VPC endpoint as the target - This option has been added as a distractor. A VPC endpoint is only meant to be used from within a VPC to connect to an S3 bucket, like so:

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

Order multiple AWS Snowball Edge appliances, transfer the data in parallel to these appliances and ship them to AWS which will then copy the data from the Snowball Edge appliances to S3 - The end-to-end time to transfer up to 80 TB of data into AWS with Snowball Edge is approximately one week, including the usual shipping and handling time in AWS data centers. Therefore this option is ruled out.

References:

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

<https://docs.aws.amazon.com/directconnect/latest/UserGuide/WorkingWithVirtualInterfaces.html>

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

Question 42: **Incorrect**

A data analytics company needs to set up a data lake on Amazon S3 for a financial services client. The data lake is split in raw and curated zones. For compliance reasons, the source data needs to be kept for a minimum of 5 years. The source data arrives in the raw zone and is then processed via an AWS Glue based ETL job into the curated zone. The business analysts run ad-hoc queries only on the data in the curated zone using Athena. The team is concerned about the cost of data storage in both the raw and curated zones as the data is increasing at a rate of 2 TB daily in each zone.

Which of the following options would you implement together as the MOST cost-optimal solution? (Select two)



Create a Lambda function based job to delete the raw zone data after 1 day



Use Glue ETL job to write the transformed data in the curated zone using CSV format

(Incorrect)

-

Use Glue ETL job to write the transformed data in the curated zone using a compressed file format

(Correct)

-

Setup a lifecycle policy to transition the raw zone data into Glacier Deep Archive after 1 day of object creation

(Correct)

-

Setup a lifecycle policy to transition the curated zone data into Glacier Deep Archive after 1 day of object creation

Explanation

Correct options:

Setup a lifecycle policy to transition the raw zone data into Glacier Deep Archive after 1 day of object creation

You can manage your objects so that they are stored cost-effectively throughout their lifecycle by configuring their Amazon S3 Lifecycle. An S3 Lifecycle configuration is a set of rules that define actions that Amazon S3 applies to a group of objects. For example, you might choose to transition objects to the S3 Standard-IA storage class 30 days after you created them, or archive objects to the S3 Glacier storage class one year after creating them.

For the given use-case, the raw zone consists of the source data, so it cannot be deleted due to compliance reasons. Therefore, you should use a lifecycle policy to transition the raw zone data into Glacier Deep Archive after 1 day of object creation.

Please read more about S3 Object Lifecycle

Object lifecycle management

[PDF](#) | [Kindle](#) | [RSS](#)

To manage your objects so that they are stored cost effectively throughout their lifecycle, configure their *Amazon S3 Lifecycle*. An *S3 Lifecycle configuration* is a set of rules that define actions that Amazon S3 applies to a group of objects. There are two types of actions:

- **Transition actions**—Define when objects transition to another [storage class](#). For example, you might choose to transition objects to the S3 Standard-IA storage class 30 days after you created them, or archive objects to the S3 Glacier storage class one year after creating them.

There are costs associated with the lifecycle transition requests. For pricing information, see [Amazon S3 pricing](#).

- **Expiration actions**—Define when objects expire. Amazon S3 deletes expired objects on your behalf.

The lifecycle expiration costs depend on when you choose to expire objects. For more information, see [Understanding object expiration](#).

For more information about S3 Lifecycle rules, see [Lifecycle configuration elements](#).

When should I use lifecycle configuration?

Define S3 Lifecycle configuration rules for objects that have a well-defined lifecycle. For example:

- If you upload periodic logs to a bucket, your application might need them for a week or a month. After that, you might want to delete them.
- Some documents are frequently accessed for a limited period of time. After that, they are infrequently accessed. At some point, you might not need real-time access to them, but your organization or regulations might require you to archive them for a specific period. After that, you can delete them.
- You might upload some types of data to Amazon S3 primarily for archival purposes. For example, you might archive digital media, financial and healthcare records, raw genomics sequence data, long-term database backups, and data that must be retained for regulatory compliance.

Management:

With S3 Lifecycle configuration rules, you can tell Amazon S3 to transition objects to less expensive storage classes, or archive or delete them.

via -

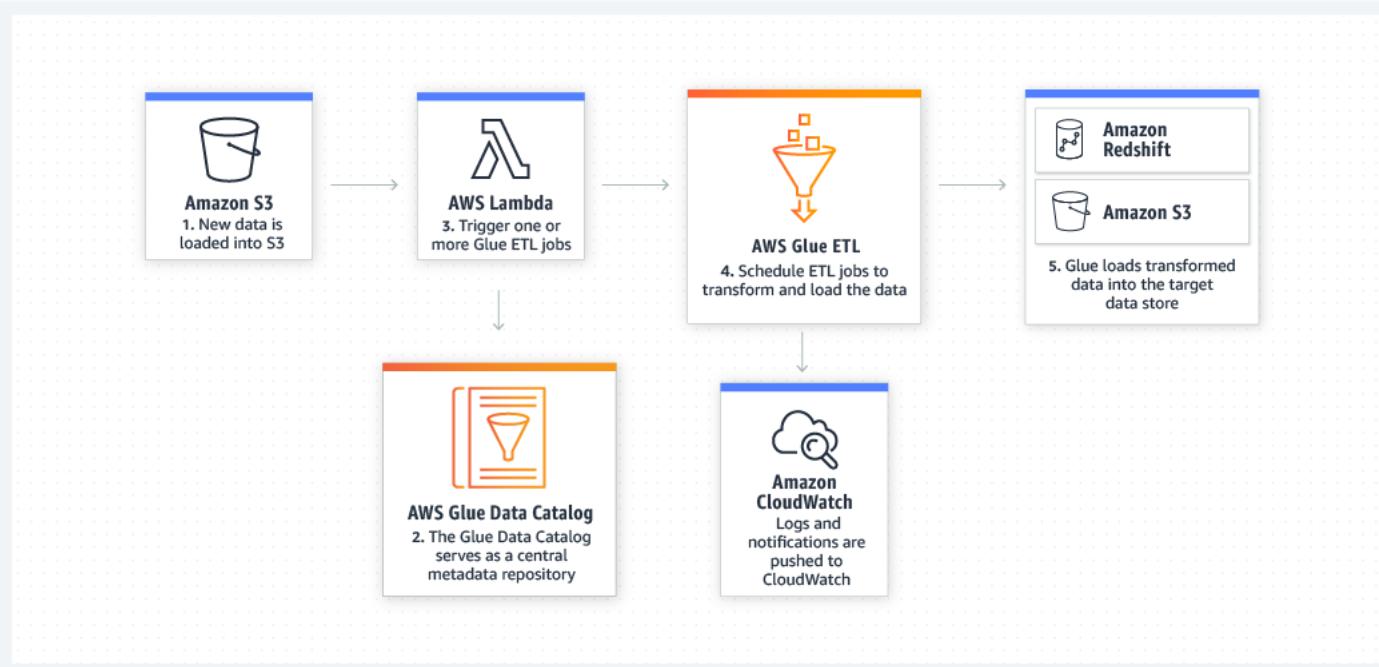
<https://docs.aws.amazon.com/AmazonS3/latest/dev/object-lifecycle-mgmt.html>

Use Glue ETL job to write the transformed data in the curated zone using a compressed file format

AWS Glue is a fully managed extract, transform, and load (ETL) service that makes it easy for customers to prepare and load their data for analytics.

You cannot transition the curated zone data into Glacier Deep Archive because it is used by the business analysts for ad-hoc querying. Therefore, the best optimization is to have the curated zone data stored in a compressed format via the Glue job. The compressed data would reduce the storage cost incurred on the data in the curated zone.

Please see this example for a Glue ETL



Pipeline:

<https://aws.amazon.com/glue/>

via -

Incorrect options:

Create a Lambda function based job to delete the raw zone data after 1 day - As mentioned in the use-case, the source data needs to be kept for a minimum of 5 years for compliance reasons. Therefore the data in the raw zone cannot be deleted after 1 day.

Setup a lifecycle policy to transition the curated zone data into Glacier Deep Archive after 1 day of object creation - You cannot transition the curated zone data into Glacier Deep Archive because it is used by the business analysts for ad-hoc querying. Hence this option is incorrect.

Use Glue ETL job to write the transformed data in the curated zone using CSV format - It is cost-optimal to write the data in the curated zone using a compressed format instead of CSV format. The compressed data would reduce the storage cost incurred on the data in the curated zone. So, this option is incorrect.

References:

<https://docs.aws.amazon.com/AmazonS3/latest/dev/object-lifecycle-mgmt.html>

<https://aws.amazon.com/glue/>

Question 43: **Correct**

A company allows property owners and travelers to connect with each other for the purpose of renting unique vacation spaces around the world. The engineering team at the company uses Amazon MySQL RDS DB cluster because it simplifies much of the time-consuming administrative tasks typically associated with databases. The team uses Multi-Availability Zone (Multi-AZ) deployment to further automate its database replication and augment data durability. The current cluster configuration also uses Read Replicas. An intern has joined the team and wants to understand the replication capabilities for Multi-AZ as well as Read Replicas for the given RDS cluster.

As a Solutions Architect Professional, which of the following capabilities would you identify as correct for the given database?



Multi-AZ follows asynchronous replication and spans at least two Availability Zones within a single region. Read Replicas follow asynchronous replication and can be within an Availability Zone, Cross-AZ, or Cross-Region



Multi-AZ follows asynchronous replication and spans at least two Availability Zones within a single region. Read Replicas follow synchronous replication and can be within an Availability Zone, Cross-AZ, or Cross-Region

- - Multi-AZ follows asynchronous replication and spans one Availability Zone within a single region. Read Replicas follow synchronous replication and can be within an Availability Zone, Cross-AZ, or Cross-Region**
 -
- Multi-AZ follows synchronous replication and spans at least two Availability Zones within a single region. Read Replicas follow asynchronous replication and can be within an Availability Zone, Cross-AZ, or Cross-Region**

(Correct)

Explanation

Correct option:

Multi-AZ follows synchronous replication and spans at least two Availability Zones within a single region. Read Replicas follow asynchronous replication and can be within an Availability Zone, Cross-AZ, or Cross-Region

Amazon RDS Multi-AZ deployments provide enhanced availability and durability for RDS database (DB) instances, making them a natural fit for production database workloads. When you provision a Multi-AZ DB Instance, Amazon RDS automatically creates a primary DB Instance and synchronously replicates the data to a standby instance in a different Availability Zone (AZ). Multi-AZ spans at least two Availability Zones within a single region.

Amazon RDS Read Replicas provide enhanced performance and durability for RDS database (DB) instances. They make it easy to elastically scale out beyond the capacity constraints of a single DB instance for read-heavy database workloads. For the MySQL, MariaDB, PostgreSQL, Oracle, and SQL Server database engines, Amazon RDS creates a second DB instance using a snapshot of the source DB instance. It then uses the engines' native asynchronous replication to update the read replica whenever there is a change to the source DB instance.

Amazon RDS replicates all databases in the source DB instance. Read Replicas can be within an Availability Zone, Cross-AZ, or Cross-Region.

Read replicas, Multi-AZ deployments, and multi-region deployments

Amazon RDS read replicas complement Multi-AZ deployments. While both features maintain a second copy of your data, there are differences between the two:

Multi-AZ deployments	Multi-Region deployments	Read replicas
Main purpose is high availability	Main purpose is disaster recovery and local performance	Main purpose is scalability
Non-Aurora: synchronous replication; Aurora: asynchronous replication	Asynchronous replication	Asynchronous replication
Non-Aurora: only the primary instance is active; Aurora: all instances are active	All regions are accessible and can be used for reads	All read replicas are accessible and can be used for readscaling
Non-Aurora: automated backups are taken from standby; Aurora: automated backups are taken from shared storage layer	Automated backups can be taken in each region	No backups configured by default
Always span at least two Availability Zones within a single region	Each region can have a Multi-AZ deployment	Can be within an Availability Zone, Cross-AZ, or Cross-Region
Non-Aurora: database engine version upgrades happen on primary; Aurora: all instances are updated together	Non-Aurora: database engine version upgrade is independent in each region; Aurora: all instances are updated together	Non-Aurora: database engine version upgrade is independent from source instance; Aurora: all instances are updated together
Automatic failover to standby (non-Aurora) or read replica (Aurora) when a problem is detected	Aurora allows promotion of a secondary region to be the master	Can be manually promoted to a standalone database instance (non-Aurora) or to be the primary instance (Aurora)

via -

<https://aws.amazon.com/rds/features/multi-az/>

Incorrect Options:

Multi-AZ follows asynchronous replication and spans one Availability Zone within a single region. Read Replicas follow synchronous replication and can be within an Availability Zone, Cross-AZ, or Cross-Region

Multi-AZ follows asynchronous replication and spans at least two Availability Zones within a single region. Read Replicas follow synchronous replication and can be within an Availability Zone, Cross-AZ, or Cross-Region

Multi-AZ follows asynchronous replication and spans at least two Availability Zones within a single region. Read Replicas follow asynchronous replication and can be within an Availability Zone, Cross-AZ, or Cross-Region

These three options contradict the earlier details provided in the explanation. Hence these options are incorrect.

References:

<https://aws.amazon.com/rds/features/multi-az/>

<https://aws.amazon.com/rds/features/read-replicas/>

Question 44: **Correct**

A financial services company runs more than 400 core-banking microservices on AWS, using services including Amazon Elastic Compute Cloud (Amazon EC2), Amazon Elastic Block Store (Amazon EBS), and Amazon Simple Storage Service (Amazon S3). The company also segregates parts of its infrastructure using separate AWS accounts, so if one account is compromised, critical parts of the infrastructure in other accounts remain unaffected. The company uses one account for production, one for non-production, and one for storing and managing users' login information and roles within AWS. The privileges that are assigned in the user account then allow users to read or write to production and non-production accounts. The company has set up "AWS Organizations" to manage several of these scenarios. The company wants to provide shared and centrally-managed VPCs to all business units for certain applications that need a high degree of interconnectivity.

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



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

(Correct)



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



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



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

Explanation

Correct option:

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

VPC sharing (part of Resource Access Manager) allows multiple AWS accounts to create their application resources such as EC2 instances, RDS databases, Redshift clusters, and Lambda functions, into shared and centrally-managed Amazon Virtual Private Clouds (VPCs).

To set this up, the account that owns the VPC (owner) shares one or more subnets with other accounts (participants) that belong to the same organization from AWS Organizations. After a subnet is shared, the participants can view, create, modify, and delete their application resources in the subnets shared with them. Participants cannot view, modify, or delete resources that belong to other participants or the VPC owner.

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

Working with shared VPCs

[PDF](#) | [Kindle](#) | [RSS](#)

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

You can share Amazon VPCs to leverage the implicit routing within a VPC for applications that require a high degree of interconnectivity and are within the same trust boundaries. This reduces the number of VPCs that you create and manage, while using separate accounts for billing and access control. You can simplify network topologies by interconnecting shared Amazon VPCs using connectivity features, such as AWS PrivateLink, AWS Transit Gateway, and Amazon VPC peering. For more information about VPC sharing benefits, see [VPC sharing: A new approach to multiple accounts and VPC management](#).

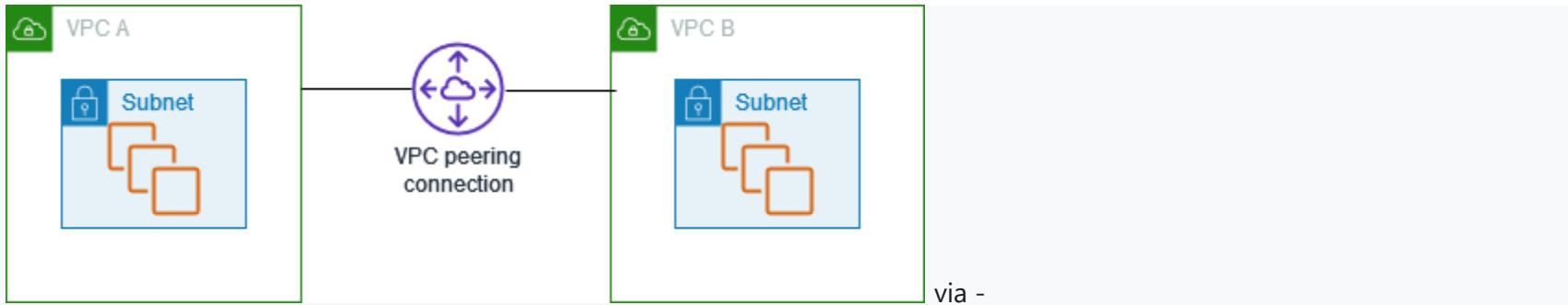
via -

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

Incorrect options:

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

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



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

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

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

References:

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

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

Question 45: **Correct**

A company has built a serverless electronic document management system for users to upload their documents. The system also has a web application that connects to an Amazon API Gateway with Regional endpoints which in turn invokes AWS Lambda functions. The Lambda functions write the metadata of the documents to the Amazon Aurora Serverless database before uploading the actual documents to the Amazon S3 bucket. While the serverless architecture has been tested in the US East (N. Virginia) Region, the solution should be scalable for other AWS Regions too.

As an AWS Certified Solutions Architect Professional, which options would you recommend to make the architecture scalable while offering low latency service to customers of any AWS region? (Select two)

- Change the API Gateway Regional endpoints to edge-optimized endpoints

(Correct)
- Configure CloudFront to use signed URLs for providing low latency access to customers of all AWS regions
- Enable S3 Transfer Acceleration on the S3 bucket and configure the web application to use the Transfer Acceleration endpoints

(Correct)
- Configure AWS Global Accelerator to front the CloudFront distribution for providing low latency access to customers of all AWS regions
- Change the API Gateway Regional endpoints to private API endpoints

Explanation
Correct options:

Enable S3 Transfer Acceleration on the S3 bucket and configure the web application to use the Transfer Acceleration endpoints

Amazon S3 Transfer Acceleration can speed up content transfers to and from Amazon S3 by as much as 50-500% for long-distance transfer of larger objects. Customers who have either web or mobile applications with widespread users or applications hosted far away from their S3 bucket can experience long and variable upload and download speeds over the Internet. S3 Transfer Acceleration

(S3TA) reduces the variability in Internet routing, congestion, and speeds that can affect transfers, and logically shortens the distance to S3 for remote applications. S3TA improves transfer performance by routing traffic through Amazon CloudFront's globally distributed Edge Locations and over AWS backbone networks, and by using network protocol optimizations.

To access the bucket that is enabled for Transfer Acceleration, you must use the endpoint `bucketname.s3-accelerate.amazonaws.com`.

Change the API Gateway Regional endpoints to edge-optimized endpoints

An API endpoint type refers to the hostname of the API. The API endpoint type can be edge-optimized, regional, or private, depending on where the majority of your API traffic originates from.

A regional API endpoint is intended for clients in the same region. An edge-optimized API endpoint is best for geographically distributed clients. API requests are routed to the nearest CloudFront Point of Presence (POP). This is the default endpoint type for API Gateway REST APIs.

Incorrect options:

Configure AWS Global Accelerator to direct traffic to the CloudFront distribution for providing low latency access to customers of all AWS regions - AWS Global Accelerator has the following types of endpoints only - Network Load Balancers, Application Load Balancers, Amazon EC2 instances, or Elastic IP addresses. Hence, this option is incorrect.

Global Accelerator

Types:

Accelerator

An accelerator directs traffic to endpoints over the AWS global network to improve the performance of your internet applications. Each accelerator includes one or more listeners.

There are two types of accelerators:

- A *standard* accelerator directs traffic to the optimal AWS endpoint based on several factors, including the user's location, the health of the endpoint, and the endpoint weights that you configure. This improves the availability and performance of your applications. Endpoints can be Network Load Balancers, Application Load Balancers, Amazon EC2 instances, or Elastic IP addresses.
- A *custom routing* accelerator lets you deterministically route multiple users to a specific EC2 destination behind your accelerator, as is required for some use cases. You do this by directing users to a unique IP address and port on your accelerator, which Global Accelerator has mapped to the destination. Note that custom routing accelerators do not support dual-stack for IP addresses.

via - <https://docs.aws.amazon.com/global-accelerator/latest/dg/introduction-components.html>

Q: How is AWS Global Accelerator different from Amazon CloudFront?

A: AWS Global Accelerator and Amazon CloudFront are separate services that use the AWS global network and its edge locations around the world. CloudFront improves performance for both cacheable content (such as images and videos) and dynamic content (such as API acceleration and dynamic site delivery). Global Accelerator improves performance for a wide range of applications over TCP or UDP by proxying packets at the edge to applications running in one or more AWS Regions. Global Accelerator is a good fit for non-HTTP use cases, such as gaming (UDP), IoT (MQTT), or Voice over IP, as well as for HTTP use cases that specifically require static IP addresses or deterministic, fast regional failover. Both services integrate with AWS Shield for DDoS protection.

Global Accelerator vs CloudFront:

via

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

Configure CloudFront to use signed URLs for providing low latency access to customers of all AWS regions - A signed URL includes additional information, for example, expiration date and time, that gives you more control over access to your content. You can distribute private content using a signed URL. However, latency issues cannot be fixed by using signed URLs.

Change the API Gateway Regional endpoints to private API endpoints - A private API endpoint is an API endpoint that can only be accessed from your Amazon Virtual Private Cloud (VPC) using an interface VPC endpoint, which is an endpoint network interface (ENI) that you create in your VPC. This option is incorrect for the given use case.

References:

<https://docs.aws.amazon.com/AmazonS3/latest/userguide/transfer-acceleration.html>

<https://docs.aws.amazon.com/apigateway/latest/developerguide/api-gateway-api-endpoint-types.html>

<https://docs.aws.amazon.com/global-accelerator/latest/dg/introduction-components.html>

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

Question 46: **Correct**

An e-commerce company is planning to migrate its IT infrastructure from the on-premises data center to AWS Cloud to ramp up its capabilities well in time for the upcoming Holiday Sale season. The company's CTO has hired you as an AWS Certified Solutions Architect Professional to design a distributed, highly available and loosely coupled order processing application. The application is responsible for receiving and processing orders before storing them in a DynamoDB table. The application has seen sporadic traffic spikes in the past and the CTO wants the application to be able to scale during marketing campaigns to process the orders with minimal disruption.

Which of the following options would you recommend as the MOST reliable solution to address these requirements?



Ingest the orders in an SQS queue and trigger a Lambda function to process them

(Correct)



Ingest the orders via a Step Function state machine and trigger an ECS container to process them



Push the orders to an SNS topic and subscribe a Lambda function to process them

-

Push the orders to Kinesis Data Streams and use Amazon EC2 instances to process them

Explanation

Correct option:

Ingest the orders in an SQS queue and trigger a Lambda function to process them

AWS Lambda lets you run code without provisioning or managing servers. You pay only for the compute time you consume. Amazon Simple Queue Service (SQS) is a fully managed message queuing service that enables you to decouple and scale microservices, distributed systems, and serverless applications. SQS offers two types of message queues. Standard queues offer maximum throughput, best-effort ordering, and at-least-once delivery. SQS FIFO queues are designed to guarantee that messages are processed exactly once, in the exact order that they are sent.

You can use an AWS Lambda function to process messages in an Amazon Simple Queue Service (Amazon SQS) queue. Lambda event source mappings support standard queues and first-in, first-out (FIFO) queues. With Amazon SQS, you can offload tasks from one component of your application by sending them to a queue and processing them asynchronously.

Scaling and processing

For standard queues, Lambda uses long polling to poll a queue until it becomes active. When messages are available, Lambda reads up to 5 batches and sends them to your function. If messages are still available, Lambda increases the number of processes that are reading batches by up to 60 more instances per minute. The maximum number of batches that can be processed simultaneously by an event source mapping is 1000.

For FIFO queues, Lambda sends messages to your function in the order that it receives them. When you send a message to a FIFO queue, you specify a [message group ID](#). Amazon SQS ensures that messages in the same group are delivered to Lambda in order. Lambda sorts the messages into groups and sends only one batch at a time for a group. If the function returns an error, all retries are attempted on the affected messages before Lambda receives additional messages from the same group.

Your function can scale in concurrency to the number of active message groups. For more information, see [SQS FIFO as an event source](#) on the AWS Compute Blog.

Configuring a queue for use with Lambda

Create an SQS queue to serve as an event source for your Lambda function. Then configure the queue to allow time for your Lambda function to process each batch of events—and for Lambda to retry in response to throttling errors as it scales up.

To allow your function time to process each batch of records, set the source queue's visibility timeout to at least 6 times the [timeout](#) that you configure on your function. The extra time allows for Lambda to retry if your function execution is throttled while your function is processing a previous batch.

If a message fails to be processed multiple times, Amazon SQS can send it to a [dead-letter queue](#). When your function returns an error, Lambda leaves it in the queue. After the visibility timeout occurs, Lambda receives the message again. To send messages to a second queue after a number of receives, configure a dead-letter queue on your source queue.

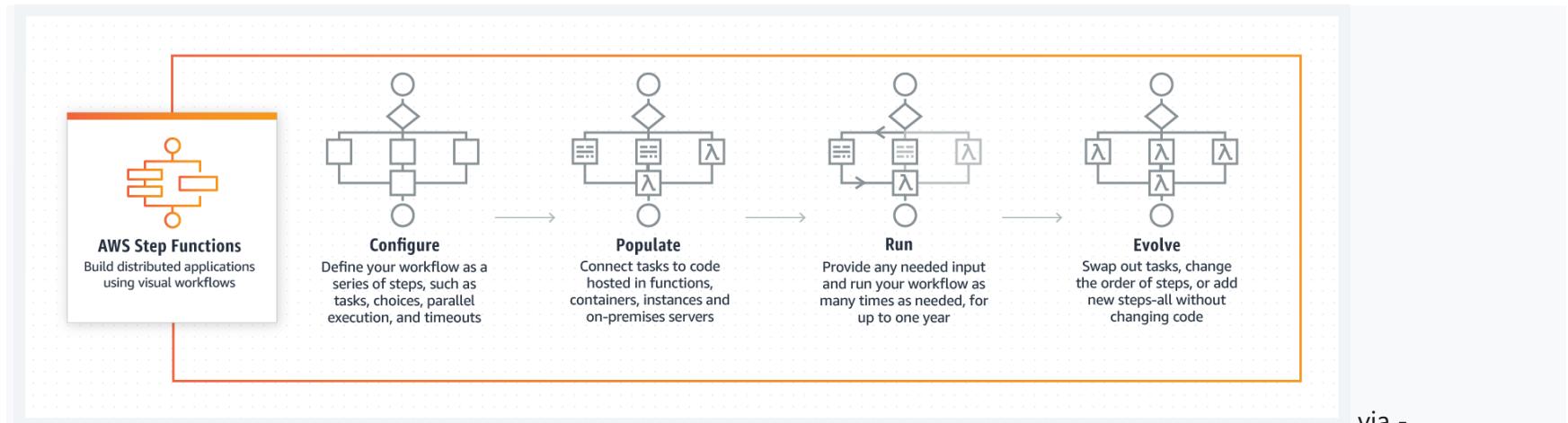
via -

<https://docs.aws.amazon.com/lambda/latest/dg/with-sqs.html>

AWS manages all ongoing operations and underlying infrastructure needed to provide a highly available and scalable message queuing service. With SQS, there is no upfront cost, no need to acquire, install, and configure messaging software, and no time-consuming build-out and maintenance of supporting infrastructure. SQS queues are dynamically created and scale automatically so you can build and grow applications quickly and efficiently.

Incorrect options:

Ingest the orders via a Step Function state machine and trigger an ECS container to process them - AWS Step Functions lets you coordinate multiple AWS services into serverless workflows so you can build and update apps quickly. Using Step Functions, you can design and run workflows that stitch together services such as AWS Lambda into feature-rich applications.



via -

<https://aws.amazon.com/step-functions/>

You cannot use a Step Functions state machine to directly ingest incoming orders, so this option is incorrect.

Push the orders to Kinesis Data Streams and use Amazon EC2 instances to process them - You cannot use EC2 instances to process the orders because if the EC2 instances become unhealthy then the application would be unable to process the orders, thereby making the architecture unreliable.

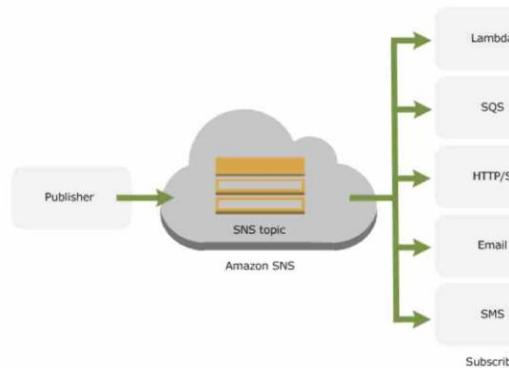
Push the orders to an SNS topic and subscribe a Lambda function to process them - Amazon Simple Notification Service (SNS) is a highly available, durable, secure, fully managed pub/sub messaging service that enables you to decouple microservices, distributed systems, and serverless applications. You cannot use SNS to facilitate reliable order processing workflow for the given use-case because once the Lambda function picks an order's message from the SNS topic, the order would remain unprocessed if the Lambda fails to process it successfully. In the case of SQS, the message will be available for re-processing after visibility timeout.

How SNS

What is Amazon SNS?

[PDF](#) | [Kindle](#) | [RSS](#)

Amazon Simple Notification Service (Amazon SNS) is a web service that coordinates and manages the delivery or sending of messages to subscribing endpoints or clients. In Amazon SNS, there are two types of clients—publishers and subscribers—also referred to as producers and consumers. Publishers communicate asynchronously with subscribers by producing and sending a message to a topic, which is a logical access point and communication channel. Subscribers (that is, web servers, email addresses, Amazon SQS queues, AWS Lambda functions) consume or receive the message or notification over one of the supported protocols (that is, Amazon SQS, HTTP/S, email, SMS, Lambda) when they are subscribed to the topic.



When using Amazon SNS, you (as the owner) create a topic and control access to it by defining policies that determine which publishers and subscribers can communicate with the topic. A publisher sends messages to topics that they have created or to topics they have permission to publish to. Instead of including a specific destination address in each message, a publisher sends a message to the topic. Amazon SNS matches the topic to a list of subscribers who have subscribed to that topic, and delivers the message to each of those subscribers. Each topic has a unique name that identifies the Amazon SNS endpoint for publishers to post messages and subscribers to register for notifications. Subscribers receive all messages published to the topics to which they subscribe, and all subscribers to a topic receive the same messages.

Works:

via -

<https://docs.aws.amazon.com/sns/latest/dg/welcome.html>

References:

<https://docs.aws.amazon.com/lambda/latest/dg/with-sqs.html>

<https://docs.aws.amazon.com/lambda/latest/dg/invocation-eventsourcemapping.html>

<https://docs.aws.amazon.com/sns/latest/dg/welcome.html>

Question 47: **Correct**

A medical technology company has recently set up a hybrid cloud between its on-premises data centers and AWS Cloud. The engineering team at the company has developed a Media Archiving and Communication System application that runs on AWS to support real-time collaboration among radiologists and other specialists. The company uses Amazon S3 to aggregate the raw medical images and video footage from its research teams across the world to discover tremendous medical insights. The technical teams at the overseas research facilities have reported huge delays in uploading large video files to the destination S3 bucket.

As a Solutions Architect Professional, which of the following would you recommend as the MOST cost-effective solutions to improve the file upload speed into S3? (Select two)

-

Create multiple AWS direct connect connections between the AWS Cloud and research facilities running in the on-premises data centers. Use the direct connect connections for faster file uploads into S3

-

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

(Correct)

-

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

-

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

(Correct)

-

Create multiple site-to-site VPN connections between the AWS Cloud and research facilities running in the on-premises data centers. Use these VPN connections for faster file uploads into S3

Explanation

Correct options:

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

Amazon S3 Transfer Acceleration

[PDF](#) | [Kindle](#) | [RSS](#)

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

When using Transfer Acceleration, additional data transfer charges may apply. For more information about pricing, see [Amazon S3 Pricing](#).

Topics

- [Why Use Amazon S3 Transfer Acceleration?](#)
- [Getting Started with Amazon S3 Transfer Acceleration](#)
- [Requirements for using Amazon S3 Transfer Acceleration](#)
- [Amazon S3 Transfer Acceleration Examples](#)

Why Use Amazon S3 Transfer Acceleration?

You might want to use Transfer Acceleration on a bucket for various reasons, including the following:

- You have customers that upload to a centralized bucket from all over the world.
- You transfer gigabytes to terabytes of data on a regular basis across continents.
- You are unable to utilize all of your available bandwidth over the Internet when uploading to Amazon S3.

via -

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

Use multipart uploads for faster file uploads into the destination S3 bucket - Multipart upload allows you to upload a single object as a set of parts. Each part is a contiguous portion of the object's data. You can upload these object parts independently and in any order. If transmission of any part fails, you can retransmit that part without affecting other parts. After all parts of your object are

uploaded, Amazon S3 assembles these parts and creates the object. In general, when your object size reaches 100 MB, you should consider using multipart uploads instead of uploading the object in a single operation. Multipart upload provides improved throughput, therefore it facilitates faster file uploads.

Multipart upload initiation

When you send a request to initiate a multipart upload, Amazon S3 returns a response with an upload ID, which is a unique identifier for your multipart upload. You must include this upload ID whenever you upload parts, list the parts, complete an upload, or stop an upload. If you want to provide any metadata describing the object being uploaded, you must provide it in the request to initiate multipart upload.

Parts upload

When uploading a part, in addition to the upload ID, you must specify a part number. You can choose any part number between 1 and 10,000. A part number uniquely identifies a part and its position in the object you are uploading. The part number that you choose doesn't need to be in a consecutive sequence (for example, it can be 1, 5, and 14). If you upload a new part using the same part number as a previously uploaded part, the previously uploaded part is overwritten. Whenever you upload a part, Amazon S3 returns an *ETag* header in its response. For each part upload, you must record the part number and the ETag value. You need to include these values in the subsequent request to complete the multipart upload.

Note

After you initiate a multipart upload and upload one or more parts, you must either complete or stop the multipart upload in order to stop getting charged for storage of the uploaded parts. Only after you either complete or stop a multipart upload will Amazon S3 free up the parts storage and stop charging you for the parts storage.

Multipart upload completion

When you complete a multipart upload, Amazon S3 creates an object by concatenating the parts in ascending order based on the part number. If any object metadata was provided in the *initiate multipart upload* request, Amazon S3 associates that metadata with the object. After a successful *complete* request, the parts no longer exist. Your *complete multipart upload* request must include the upload ID and a list of both part numbers and corresponding ETag values. Amazon S3 response includes an ETag that uniquely identifies the combined object data. This ETag will not necessarily be an MD5 hash of the object data. You can optionally stop the multipart upload. After stopping a multipart upload, you cannot upload any part using that upload ID again. All storage from any part of the cancelled multipart upload is then freed. If any part uploads were in-progress, they can still succeed or fail even after you stop. To free all storage consumed by all parts, you must stop a multipart upload only after all part uploads have completed.

Multipart upload listings

You can list the parts of a specific multipart upload or all in-progress multipart uploads. The *list parts* operation returns the parts information that you have uploaded for a specific multipart upload. For each *list parts* request, Amazon S3 returns the parts information for the specified multipart upload, up to a maximum of 1,000 parts. If there are more than 1,000 parts in the multipart upload, you must send a series of *list part* requests to retrieve all the parts. Note that the returned list of parts doesn't include parts that haven't completed uploading. Using the *list multipart uploads* operation, you can obtain a list of multipart uploads in progress. An in-progress multipart upload is an upload that you have initiated, but have not yet completed or stopped. Each request returns at most 1000 multipart uploads. If there are more than 1,000 multipart uploads in progress, you need to send additional requests to retrieve the remaining multipart uploads. Only use the returned listing for verification. You should not use the result of this listing when sending a *complete multipart upload* request. Instead, maintain your own list of the part numbers you specified when uploading parts and the corresponding ETag values that Amazon S3 returns.

via -

<https://docs.aws.amazon.com/AmazonS3/latest/dev/mpuoverview.html>

Incorrect options:

Create multiple AWS direct connect connections between the AWS Cloud and research facilities running in the on-premises data centers. Use the direct connect connections for faster file uploads into S3 - AWS Direct Connect is a cloud service solution that makes it easy to establish a dedicated network connection from your premises to AWS. AWS Direct Connect lets you establish a dedicated network connection between your network and one of the AWS Direct Connect locations. Direct connect takes significant time (several months) to be provisioned and is an overkill for the given use-case.

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

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

References:

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

<https://docs.aws.amazon.com/AmazonS3/latest/dev/mpuoverview.html>

Question 48: **Incorrect**

A web development studio runs hundreds of Proof-of-Concept (PoC) and demo applications on virtual machines running on an on-premises server. Many of the applications are simple PHP, JavaScript or Python web applications which are no longer actively developed and serve little traffic.

As a Solutions Architect Professional, which of the following approaches would you suggest to migrate these applications to AWS with the lowest infrastructure cost and least development effort?



Migrate the application code to use a serverless stack comprising of Lambda functions and DynamoDB



Dockerize each application and then deploy to an ECS cluster running behind an Application Load Balancer

(Correct)



Leverage VM Import/Export to create AMIs for each virtual machine and run them in single-instance AWS Elastic Beanstalk environments by configuring a custom image



Leverage AWS Server Migration Service (SMS) to create AMIs for each virtual machine and run each application on a dedicated EC2 instance

(Incorrect)

Explanation

Correct option:

Dockerize each application and then deploy to an ECS cluster running behind an Application Load Balancer

Amazon Elastic Container Service (ECS) is a highly scalable, high performance container management service that supports Docker containers and allows you to easily run applications on a managed cluster of Amazon EC2 instances.



How ECS Works:

<https://aws.amazon.com/ecs/>

Using an ECS cluster running behind an Application Load Balancer offers advantages such as Application Load Balancers allow containers to use dynamic host port mapping so that multiple tasks from the same service are allowed per container instance. This reduces the number of instances required for migration and therefore reduces the overall costs.

via -

Service load balancing

[PDF](#) | [Kindle](#) | [RSS](#)

Your Amazon ECS service can optionally be configured to use Elastic Load Balancing to distribute traffic evenly across the tasks in your service.

Amazon ECS services support the Application Load Balancer, Network Load Balancer, and Classic Load Balancer load balancer types. Application Load Balancers are used to route HTTP/HTTPS (or layer 7) traffic. Network Load Balancers are used to route TCP or UDP (or layer 4) traffic. Classic Load Balancers are used to route TCP traffic. For more information, see [Load balancer types](#).

Application Load Balancers offer several features that make them attractive for use with Amazon ECS services:

- Each service can serve traffic from multiple load balancers and expose multiple load balanced ports by specifying multiple target groups.
- They are supported by tasks using both the Fargate and EC2 launch types.
- Application Load Balancers allow containers to use dynamic host port mapping (so that multiple tasks from the same service are allowed per container instance).
- Application Load Balancers support path-based routing and priority rules (so that multiple services can use the same listener port on a single Application Load Balancer).

We recommend that you use Application Load Balancers for your Amazon ECS services so that you can take advantage of these latest features, unless your service requires a feature that is only available with Network Load Balancers or Classic Load Balancers. For more information about Elastic Load Balancing and the differences between the load balancer types, see the [Elastic Load Balancing User Guide](#).

via -

<https://docs.aws.amazon.com/AmazonECS/latest/developerguide/service-load-balancing.html>

Incorrect options:

Migrate the application code to use a serverless stack comprising of Lambda functions and DynamoDB - It would take significant development effort to migrate the existing applications to a serverless stack. In addition, some of the web applications, such as those built using PHP are not supported by Lambda functions (unless you use custom runtime). So, this option is incorrect.

Leverage AWS Server Migration Service (SMS) to create AMIs for each virtual machine and run each application on a separate EC2 instance - AWS Server Migration Service automates the migration of your on-premises VMware vSphere, Microsoft Hyper-V/SCVMM, and Azure virtual machines to the AWS Cloud. AWS SMS incrementally replicates your server VMs as cloud-hosted Amazon Machine Images (AMIs) ready for deployment on Amazon EC2. As this solution provisions a separate EC2 instance for each application, so it is a costly migration solution.

Leverage VM Import/Export to create AMIs for each virtual machine and run them in single-instance AWS Elastic Beanstalk environments by configuring a custom image - VM Import/Export enables you to easily import virtual machine images from your existing environment to Amazon EC2 instances. You can use Elastic Beanstalk to quickly deploy and manage applications in the AWS Cloud without having to learn about the infrastructure that runs those applications. 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.

A single-instance environment doesn't have a load balancer, which can help you reduce costs compared to a load-balanced, scalable environment. Use a single-instance environment if you expect your production application to have low traffic or if you are doing remote development. For the given use-case, you will end up creating as many instances as the number of applications, which will turn out to be a really costly solution.

References:

<https://aws.amazon.com/ecs/>

<https://docs.aws.amazon.com/AmazonECS/latest/developerguide/service-load-balancing.html>

<https://docs.aws.amazon.com/elasticbeanstalk/latest/dg/Welcome.html>

Question 49: **Correct**

A big data analytics company leverages its proprietary analytics workflow (built using Redshift) to correlate traffic with marketing campaigns and to help retailers optimize hours for peak traffic, among other activities. The company has hired you as an AWS Certified Solutions Architect Professional to review the company's Redshift cluster, which has now become an integral part of its technology solutions. You have been asked to improve the reliability and availability of the cluster in case of a disaster and provide options to ensure that if an issue arises, the cluster can either operate or be restored within five hours.

Which of the following would you suggest as the BEST solution to meet the business needs in the most cost-effective way?



Set up a CloudFormation stack set for Redshift cluster creation so it can be launched in another Region and configure Amazon Redshift to automatically copy snapshots for the cluster to the other AWS Region. In case of a disaster, restore the cluster in the other AWS Region from that Region's snapshot

(Correct)



Set up two identical Amazon Redshift clusters in different regions in a primary-secondary configuration. Develop a solution using the Kinesis Data Streams to collect the data prior to ingestion into the primary Redshift cluster and stream the data to the secondary cluster



Set up two identical Amazon Redshift clusters in different regions in a primary-secondary configuration. Create a cron job to run the UNLOAD command every five hours to export data for all tables in primary cluster to S3. Use cross-region replication from the primary region to secondary region. Create another cron job to ingest the data for all tables from S3 into the secondary cluster using the LOAD command



Configure the Amazon Redshift cluster to make use of Auto Scaling groups with the nodes in the cluster spread across multiple Availability Zones (AZs). In case of a disaster, the nodes in the other AZs will ensure reliability and availability

Explanation

Correct option:

"Set up a CloudFormation stack set for Redshift cluster creation so it can be launched in another Region and configure Amazon Redshift to automatically copy snapshots for the cluster to the other AWS Region. In case of a disaster, restore the cluster in the other AWS Region from that Region's snapshot"

A CloudFormation stack set lets you create stacks in AWS accounts across regions by using a single CloudFormation template. All the resources included in each stack are defined by the stack set's CloudFormation template. As you create the stack set, you specify the template to use, as well as any parameters and capabilities that the template requires. For the given use-case, you can set up a CloudFormation stack set for Redshift cluster creation in another Region.

When automated snapshots are enabled for a cluster, Amazon Redshift periodically takes snapshots of that cluster. By default, Amazon Redshift takes a snapshot about every eight hours or following every 5 GB per node of data changes, or whichever comes first. For the given use-case, you can configure Amazon Redshift to automatically copy snapshots for the cluster to another AWS Region. When a snapshot is created in the cluster's primary AWS Region, it's copied to a secondary AWS Region. The two AWS Regions are known respectively as the source AWS Region and destination AWS Region.

In case of a disaster, you can restore your cluster from the snapshot in the destination Region. Amazon Redshift uses the cluster information to create a new cluster. Then it restores all the databases from the snapshot data. The cluster is restored in the destination AWS Region and a random, system-chosen Availability Zone, unless you specify another Availability Zone in your request.

Automated snapshots

When automated snapshots are enabled for a cluster, Amazon Redshift periodically takes snapshots of that cluster. By default Amazon Redshift takes a snapshot about every eight hours or following every 5 GB per node of data changes, or whichever comes first. Alternatively, you can create a snapshot schedule to control when automated snapshots are taken. Automated snapshots are enabled by default when you create a cluster.

Automated snapshots are deleted at the end of a retention period. The default retention period is one day, but you can modify it by using the Amazon Redshift console or programmatically by using the Amazon Redshift API or CLI.

To disable automated snapshots, set the retention period to zero. If you disable automated snapshots, Amazon Redshift stops taking snapshots and deletes any existing automated snapshots for the cluster.

Only Amazon Redshift can delete an automated snapshot; you cannot delete them manually. Amazon Redshift deletes automated snapshots at the end of a snapshot's retention period, when you disable automated snapshots for the cluster, or when you delete the cluster. Amazon Redshift retains the latest automated snapshot until you disable automated snapshots or delete the cluster.

If you want to keep an automated snapshot for a longer period, you can create a copy of it as a manual snapshot. The automated snapshot is retained until the end of the retention period, but the corresponding manual snapshot is retained until you manually delete it or until the end of the retention period.

Copying snapshots to another AWS Region

You can configure Amazon Redshift to automatically copy snapshots (automated or manual) for a cluster to another AWS Region. When a snapshot is created in the cluster's primary AWS Region, it's copied to a secondary AWS Region. The two AWS Regions are known respectively as the *source AWS Region* and *destination AWS Region*. If you store a copy of your snapshots in another AWS Region, you can restore your cluster from recent data if anything affects the primary AWS Region. You can configure your cluster to copy snapshots to only one destination AWS Region at a time. For a list of Amazon Redshift Regions, see [Regions and endpoints](#) in the *Amazon Web Services General Reference*.

When you enable Amazon Redshift to automatically copy snapshots to another AWS Region, you specify the destination AWS Region to copy the snapshots to. For automated snapshots, you can also specify the retention period to keep them in the destination AWS Region. After an automated snapshot is copied to the destination AWS Region and it reaches the retention time period there, it's deleted from the destination AWS Region. Doing this keeps your snapshot usage low. To keep the automated snapshots for a shorter or longer time in the destination AWS Region, change this retention period.

The retention period that you set for automated snapshots that are copied to the destination AWS Region is separate from the retention period for automated snapshots in the source AWS Region. The default retention period for copied snapshots is seven days. That seven-day period applies only to automated snapshots. In both the source and destination AWS Regions, manual snapshots are deleted at the end of the snapshot retention period or when you manually delete them.

You can disable automatic snapshot copy for a cluster at any time. When you disable this feature, snapshots are no longer copied from the source AWS Region to the destination AWS Region. Any automated snapshots copied to the destination AWS Region are deleted as they reach the retention period limit, unless you create manual snapshot copies of them. These manual snapshots, and any manual snapshots that were copied from the destination AWS Region, are kept in the destination AWS Region until you manually delete them.

To change the destination AWS Region that you copy snapshots to, first disable the automatic copy feature. Then re-enable it, specifying the new destination AWS Region.

After a snapshot is copied to the destination AWS Region, it becomes active and available for restoration purposes.

To copy snapshots for AWS KMS–encrypted clusters to another AWS Region, create a grant for Amazon Redshift to use a KMS customer master key (CMK) in the destination AWS Region. Then choose that grant when you enable copying of snapshots in the source AWS Region. For more information about configuring snapshot copy grants, see [Copying AWS KMS–encrypted snapshots to another AWS Region](#).

Restoring a cluster from a snapshot

A snapshot contains data from any databases that are running on your cluster. It also contains information about your cluster, including the number of nodes, node type, and master user name. If you restore your cluster from a snapshot, Amazon Redshift uses the cluster information to create a new cluster. Then it restores all the databases from the snapshot data.

For the new cluster created from the original snapshot, you can choose the configuration, such as node type and number of nodes. The cluster is restored in the same AWS Region and a random, system-chosen Availability Zone, unless you specify another Availability Zone in your request. When you restore a cluster from a snapshot, you can optionally choose a compatible maintenance track for the new cluster.

 **Note**

When you restore a snapshot to a cluster with a different configuration, the snapshot must have been taken on a cluster with cluster version 1.0.10013, or later.

When a restore is in progress, events are typically emitted in the following order:

1. RESTORE_STARTED – REDSHIFT-EVENT-2008 sent when the restore process begins.
2. RESTORE_SUCCEEDED – REDSHIFT-EVENT-3003 sent when the new cluster has been created.
The cluster is available for queries.
3. DATA_TRANSFER_COMPLETED – REDSHIFT-EVENT-3537 sent when data transfer complete.

 **Note**

RA3 clusters only emit RESTORE_STARTED and RESTORE_SUCCEEDED events. There is no explicit data transfer to be done after a RESTORE succeeds because RA3 node types store data in Amazon Redshift managed storage. With RA3 nodes, data is continuously transferred between RA3 nodes and Amazon Redshift managed storage as part of normal query processing. RA3 nodes cache hot data locally and keep less frequently queried blocks in Amazon Redshift managed storage automatically.

You can monitor the progress of a restore by either calling the [DescribeClusters](#) API operation, or viewing the cluster details in the AWS Management Console. For an in-progress restore, these display information such as the size of the snapshot data, the transfer rate, the elapsed time, and the estimated time remaining. For a description of these metrics, see [RestoreStatus](#).

You can't use a snapshot to revert an active cluster to a previous state.

 **Note**

When you restore a snapshot into a new cluster, the default security group and parameter group are used unless you specify different values.

You might want to restore a snapshot to a cluster with a different configuration for these reasons:

- When a cluster is made up of smaller node types and you want to consolidate it into a larger node type with fewer nodes.
- When you have monitored your workload and determined the need to move to a node type with more CPU and storage.
- When you want to measure performance of test workloads with different node types.

Restore has the following constraints:

- The new node configuration must have enough storage for existing data. Even when you add nodes, your new configuration might not have enough storage because of the way that data is redistributed.
- The restore operation checks if the snapshot was created on a cluster version that is compatible with the cluster version of the new cluster. If the new cluster has a version level that is too early, then the restore operation fails and reports more information in an error message.
- The possible configurations (number of nodes and node type) you can restore to is determined by the number of nodes in the original cluster and the target node type of the new cluster. To determine the possible configurations available, you can use the Amazon Redshift console or the `describe-node-configuration-options` AWS CLI command with `action-type restore-cluster`. For more information about the restoring using the Amazon Redshift console, see [Restoring a cluster from a snapshot](#).

via - <https://docs.aws.amazon.com/redshift/latest/mgmt/working-with-snapshots.html>

Incorrect options:

"Set up two identical Amazon Redshift clusters in different regions in a primary-secondary configuration. Develop a solution using the Kinesis Data Streams to collect the data prior to ingestion into the primary Redshift cluster and stream the data to the secondary cluster" - This option is not cost-effective as you need to keep two Redshift clusters in operation for each AWS Region. In addition, Kinesis Data Streams is meant to be used for streaming use-cases and it also adds another layer of costs to this proposed solution. To load or unload data in Redshift, you would rather use the LOAD and UNLOAD command and use S3 as the underlying data repository.

"Set up two identical Amazon Redshift clusters in different regions in a primary-secondary configuration. Create a cron job to run the UNLOAD command every five hours to export data for all tables in primary cluster to S3. Use cross-region replication from the primary region to secondary region. Create another cron job to ingest the data for all tables from S3 into the secondary cluster using the LOAD command" - This option is not cost effective as you need to keep two Redshift clusters in operation for each AWS Region. Moreover, using the LOAD and UNLOAD commands in combination with leveraging cross-region replication for S3 makes this solution complex and difficult to maintain.

"Configure the Amazon Redshift cluster to make use of Auto Scaling groups with the nodes in the cluster spread across multiple Availability Zones (AZs). In case of a disaster, the nodes in the other AZs will ensure reliability and availability" - This option has been added as a distractor as you cannot configure a Redshift cluster to make use of Auto Scaling groups.

References:

<https://docs.aws.amazon.com/redshift/latest/mgmt/working-with-snapshots.html>

<https://docs.aws.amazon.com/AWSCloudFormation/latest/UserGuide/stacksets-concepts.html>

Question 50: **Incorrect**

A web-hosting startup manages more than 500 public web applications on AWS Cloud which are deployed in a single AWS Region. The fully qualified domain names (FQDNs) of all of the applications are configured to use HTTPS and are served via Application Load Balancers (ALBs). These ALBs are configured to use public SSL/TLS certificates. The startup has hired you as an AWS Certified Solutions Architect Professional to migrate the web applications to a multi-Region architecture. You must ensure that all HTTPS services continue to work without interruption.

Which of the following solutions would you suggest to address these requirements?



Set up the key pairs and then generate the certificate for each FQDN via AWS KMS. Associate the same FQDN certificate with the ALBs in the relevant AWS Regions



Generate a separate certificate for each FQDN in each AWS Region using AWS Certificate Manager. Associate the certificates with the corresponding ALBs in the relevant AWS Region

(Correct)



Generate a separate certificate for each FQDN in each AWS Region using AWS KMS. Associate the certificates with the corresponding ALBs in the relevant AWS Region



Generate a certificate for each FQDN via AWS Certificate Manager. Associate the same FQDN certificate with the ALBs in the relevant AWS Regions

(Incorrect)

Explanation

Correct option:

Generate a separate certificate for each FQDN in each AWS Region using AWS Certificate Manager. Associate the certificates with the corresponding ALBs in the relevant AWS Region

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

certificates are used to secure network communications and establish the identity of websites over the Internet as well as resources on private networks. AWS Certificate Manager removes the time-consuming manual process of purchasing, uploading, and renewing SSL/TLS certificates.

A fully qualified domain name (FQDN) is the complete DNS name for a computer, website, or other resource connected to a network or to the internet. For example, aws.amazon.com is the FQDN for Amazon Web Services. An FQDN includes all domains up to the top-level domain. For example, [subdomain1].[subdomain2]...[subdomainn].[apex domain].[top-level domain] represents the general format of an FQDN.

To use a certificate with an Application Load Balancer for the same site (the same fully qualified domain name, or FQDN, or set of FQDNs) in a different Region, you must request a new certificate for each Region in which you plan to use it. To use an ACM certificate with Amazon CloudFront, you must request the certificate in the US East (N. Virginia) Region.

Therefore, to migrate the web applications to a multi-Region architecture, you must request a separate certificate for each FQDN in each AWS Region using AWS Certificate Manager and then associate the certificates with the corresponding ALBs in the relevant AWS Region.

Supported Regions

[PDF](#) | [Kindle](#) | [RSS](#)

Visit [AWS Regions and Endpoints](#) in the [AWS General Reference](#) or the [AWS Region Table](#) to see the regional availability for ACM.

Certificates in ACM are regional resources. To use a certificate with Elastic Load Balancing for the same fully qualified domain name (FQDN) or set of FQDNs in more than one AWS region, you must request or import a certificate for each region. For certificates provided by ACM, this means you must revalidate each domain name in the certificate for each region. You cannot copy a certificate between regions.

To use an ACM certificate with Amazon CloudFront, you must request or import the certificate in the US East (N. Virginia) region. ACM certificates in this region that are associated with a CloudFront distribution are distributed to all the geographic locations configured for that distribution.

via -

<https://docs.aws.amazon.com/acm/latest/userguide/acm-regions.html>

Incorrect options:

Generate a certificate for each FQDN via AWS Certificate Manager. Associate the same FQDN certificate with the ALBs in the relevant AWS Regions - As explained above, you cannot use the same certificate for a given FQDN across multiple AWS Regions, so this option is incorrect.

Generate a new certificate for each FQDN in the relevant AWS Region using AWS KMS. Associate the certificate with the corresponding ALBs in the relevant AWS Region

Generate a separate certificate for each FQDN in each AWS Region using AWS KMS. Associate the certificates with the corresponding ALBs in the relevant AWS Region

AWS KMS is a managed service that enables you to easily create and control the keys used for cryptographic operations. You can use KMS to centrally manage the encryption keys that control access to your data so that you can secure your data across AWS services. You cannot use KMS to provision or manage SSL/TLS certificates for an FQDN, so both these options are incorrect.

Reference:

<https://docs.aws.amazon.com/acm/latest/userguide/acm-Regions.html>

Question 51: **Correct**

A healthcare company has migrated some of its IT infrastructure to AWS Cloud and is looking for a solution to enable real-time data transfer between AWS and its data centers to reduce the turnaround time to generate the patients' diagnostic reports. The company wants to build a patient results archival solution such that only the most frequently accessed results are available as cached data locally while backing up all results on Amazon S3.

As a Solutions Architect Professional, which of the following solutions would you recommend for this use-case?



Use AWS direct connect to store the most frequently accessed results locally for low-latency access while storing the full backup of results in an Amazon S3 bucket



Use AWS Volume Gateway - Cached Volume - to store the most frequently accessed results locally for low-latency access while storing the full volume with all results in its Amazon S3 service bucket

(Correct)

-
- Use AWS Volume Gateway - Stored Volume - to store the most frequently accessed results locally for low-latency access while storing the full volume with all results in its Amazon S3 service bucket**
-
- Use AWS Snowball Edge Storage Optimized device to store the most frequently accessed results locally for low-latency access while storing the full backup of results in an Amazon S3 bucket**

Explanation

Correct option:

Use AWS Volume Gateway - Cached Volume - to store the most frequently accessed results locally for low-latency access while storing the full volume with all results in its Amazon S3 service bucket

AWS Storage Gateway is a hybrid cloud storage service that gives you on-premises access to virtually unlimited cloud storage. The service provides three different types of gateways – Tape Gateway, File Gateway, and Volume Gateway – that seamlessly connect on-premises applications to cloud storage, caching data locally for low-latency access.

With cached volumes, the AWS Volume Gateway stores the full volume in its Amazon S3 service bucket, and just the recently accessed data is retained in the gateway's local cache for low-latency access.

Volume Gateway presents cloud-backed iSCSI block storage volumes to your on-premises applications. Volume Gateway stores and manages on-premises data in Amazon S3 on your behalf and operates in either cache mode or stored mode. In the cached Volume Gateway mode, your primary data is stored in Amazon S3, while retaining your frequently accessed data locally in the cache for low latency access. In the stored Volume Gateway mode, your primary data is stored locally and your entire dataset is available for low latency access on premises while also asynchronously getting backed up to Amazon S3. In either mode, you can take point-in-time copies of your volumes using AWS Backup, which are stored in AWS as Amazon EBS snapshots. Using Amazon EBS Snapshots enables you to make space-efficient versioned copies of your volumes for data protection, recovery, migration, and various other copy data needs.

How it works



Benefits

Integrates seamlessly with on-premises applications

Volume Gateway offers cloud-backed storage to your on-premises applications using industry standard iSCSI connectivity. You don't need to rewrite your on-premises applications to use cloud storage. You can deploy Volume Gateway as a virtual machine or on the Storage Gateway Hardware Appliance at your premises.

Provides low latency access to cloud-backed storage

Volume Gateway maintains on-premises either a cache of recently accessed data, or a full volume copy, so your applications get the benefit of fast access to data. Concurrently, all of your volume data is compressed and stored durably and cost-effectively in AWS, with petabyte scalability.

Offers flexible data protection and recovery

With Amazon EBS snapshots, Storage Gateway volume clones, and AWS Backup, you have several options to restore the application data stored in your volumes - back to the existing Volume Gateway onsite, to EBS for recovery of your application into EC2, or even to a new Volume Gateway running at another on-premises location.

via -

<https://aws.amazon.com/storagegateway/volume/>

Incorrect options:

Use AWS direct connect to store the most frequently accessed results locally for low-latency access while storing the full backup of results in an Amazon S3 bucket - AWS Direct Connect lets you establish a dedicated network connection between your network and one of the AWS Direct Connect locations. Direct connect cannot be used to store the most frequently accessed results locally for low-latency access.

Use AWS Volume Gateway - Stored Volume - to store the most frequently accessed results locally for low-latency access while storing the full volume with all results in its Amazon S3 service bucket - With stored volumes, your entire data volume is available locally in the gateway, for fast read access. Volume Gateway also maintains an asynchronous copy of your stored volume in the service's Amazon S3 bucket. This does not fit the requirements per the given use-case, hence this option is not correct.

Use AWS Snowball Edge Storage Optimized device to store the most frequently accessed results locally for low-latency access while storing the full backup of results in an Amazon S3 bucket - You can use Snowball Edge Storage Optimized device to securely and quickly transfer dozens of terabytes to petabytes of data to AWS. Snowball Edge Storage Optimized device cannot be used to store the most frequently accessed results locally for low-latency access.

Reference:

<https://aws.amazon.com/storagegateway/volume/>

Question 52: **Incorrect**

A data analytics company stores event data in its on-premises PostgreSQL database. With the increase in the number of clients, the company is spending a lot of resources managing and maintaining the infrastructure while performance seems to be dwindling. The company has established connectivity between its on-premises systems and AWS Cloud already and wants a hybrid solution that can automatically buffer and transform event data in a scalable way and create visualizations to track and monitor events in real time. The transformed event data would be in semi-structured JSON format and have dynamic schemas.

Which combination of services/technologies will you suggest to implement the requirements?

-
-

Set up Amazon Kinesis Data Firehose to buffer events and an AWS Lambda function to process and transform the events. Provision an Amazon Aurora PostgreSQL DB cluster to receive the transformed events from Firehose and use QuickSight to create near-real-time visualizations and dashboards

(Incorrect)

-
-

Set up Amazon Kinesis Data Firehose to buffer events and an AWS Lambda function to process and transform the events. Provision an Amazon Aurora Neptune DB cluster to receive the transformed events from Firehose and use QuickSight to create near-real-time visualizations and dashboards



Set up Amazon Kinesis data stream to buffer events and an AWS Lambda function to process and transform the events. Use AWS Athena to create real-time visualizations of the events



Set up Amazon Kinesis Data Firehose to buffer events and an AWS Lambda function to process and transform the events. Set up Amazon OpenSearch to receive the transformed events. Use the Kibana endpoint that is deployed with OpenSearch to create near-real-time visualizations and dashboards

(Correct)

Explanation

Correct option:

Set up Amazon Kinesis Data Firehose to buffer events and an AWS Lambda function to process and transform the events. Set up Amazon OpenSearch to receive the transformed events. Use the Kibana endpoint that is deployed with OpenSearch to create near-real-time visualizations and dashboards - Amazon OpenSearch Service makes it easy for you to perform interactive log analytics, real-time application monitoring, a website search, and more. OpenSearch is an open-source, distributed search and analytics suite derived from Elasticsearch. Amazon OpenSearch Service is the successor to Amazon Elasticsearch Service. It offers visualization capabilities powered by OpenSearch Dashboards and Kibana.

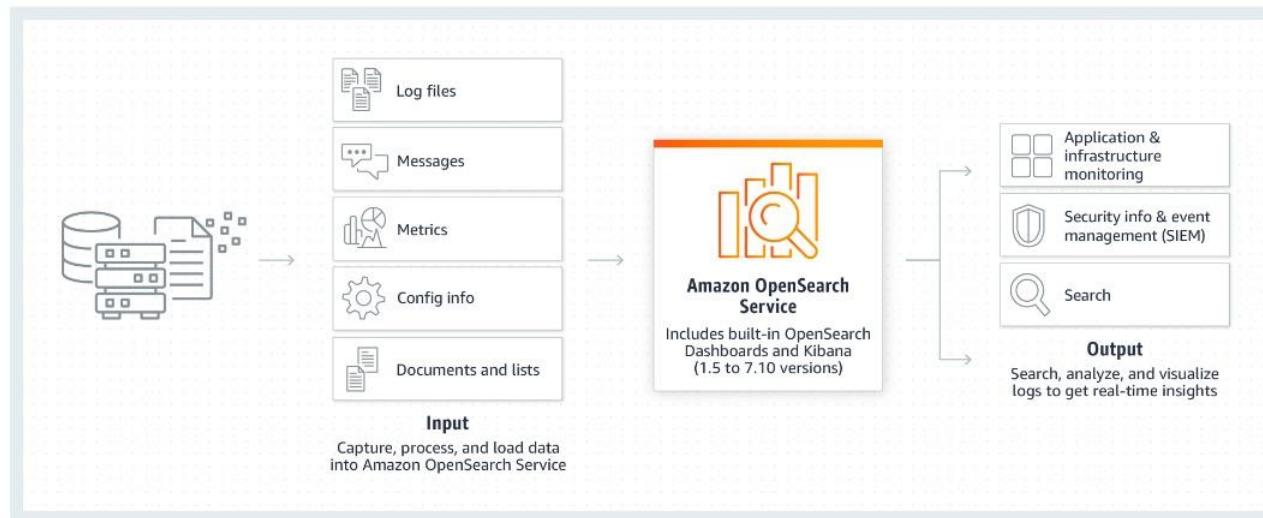
You can search and analyze large amounts of semi-structured data with native capabilities. You can visualize, monitor, and alert with anomaly detection observability features of OpenSearch Dashboards, and conduct interactive analysis and visualization on data with Piped Processing Language (PPL), a query interface.

For large data volumes, AWS recommends Amazon Kinesis Data Firehose as a data ingestion service. Kinesis Data Firehose is a fully managed service that automatically scales to match the throughput of your data and requires no ongoing administration. It can also transform, batch, and compress the data before loading it.

Amazon OpenSearch

How it works

Amazon OpenSearch Service makes it easy for you to perform interactive log analytics, real-time application monitoring, website search, and more. OpenSearch is an open source, distributed search and analytics suite derived from Elasticsearch. Amazon OpenSearch Service is the successor to Amazon Elasticsearch Service, and offers the latest versions of OpenSearch, support for 19 versions of Elasticsearch (1.5 to 7.10 versions), as well as visualization capabilities powered by OpenSearch Dashboards and Kibana (1.5 to 7.10 versions). Amazon OpenSearch Service currently has tens of thousands of active customers with hundreds of thousands of clusters under management processing hundreds of trillions of requests per month.



Service:

<https://aws.amazon.com/opensearch-service/>

via -

Incorrect options:

Set up Amazon Kinesis Data Firehose to buffer events and an AWS Lambda function to process and transform the events. Provision an Amazon Aurora PostgreSQL DB cluster to receive the transformed events from Firehose and use QuickSight to

create near-real-time visualizations and dashboards - This option is incorrect. Firehose does not support PostgreSQL as a destination. In addition, PostgreSQL DB is not the best fit to store semi-structured data that has a dynamic schema.

Set up Amazon Kinesis Data Firehose to buffer events and an AWS Lambda function to process and transform the events. Provision an Amazon Aurora Neptune DB cluster to receive the transformed events from Firehose and use QuickSight to create near-real-time visualizations and dashboards - Amazon Neptune is a fast, reliable, fully-managed graph database service that makes it easy to build and run applications that work with highly connected datasets. SQL queries for highly connected data are complex and hard to tune for performance. Instead, with Amazon Neptune, you can use open and popular graph query languages to execute powerful queries that are easy to write and perform well on connected data. Firehose does not support Neptune DB as a destination. This option is incorrect.

Set up Amazon Kinesis data stream to buffer events and an AWS Lambda function to process and transform the events. Use AWS Athena to create real-time visualizations of the events - Athena is not a visualization service, so this option is incorrect.

References:

<https://aws.amazon.com/opensearch-service/>

<https://aws.amazon.com/blogs/aws/amazon-elasticsearch-service-is-now-amazon-opensearch-service-and-supports-opensearch-10/>

Question 53: **Incorrect**

The engineering team at a healthcare company is working on the Disaster Recovery (DR) plans for its Redshift cluster deployed in the eu-west-1 Region. The existing cluster is encrypted via AWS KMS and the team wants to copy the Redshift snapshots to another Region to meet the DR requirements.

As a Solutions Architect Professional, which of the following solutions would you suggest to address the given use-case?



Create a snapshot copy grant in the destination Region for a KMS key in the destination Region. Configure Redshift cross-Region snapshots in the source Region

(Correct)

-

Create an IAM role in destination Region with access to the KMS key in the source Region. Create a snapshot copy grant in the destination Region for this KMS key in the source Region. Configure Redshift cross-Region snapshots in the source Region

(Incorrect)

-

Create a snapshot copy grant in the destination Region for a KMS key in the destination Region. Configure Redshift cross-Region replication in the source Region

-

Create a snapshot copy grant in the source Region for a KMS key in the source Region. Configure Redshift cross-Region snapshots in the destination Region

Explanation

Correct option:

Create a snapshot copy grant in the destination Region for a KMS key in the destination Region. Configure Redshift cross-Region snapshots in the source Region

To copy snapshots for AWS KMS–encrypted clusters to another AWS Region, you need to create a grant for Redshift to use a KMS customer master key (CMK) in the destination AWS Region. Then choose that grant when you enable copying of snapshots in the source AWS Region. You cannot use a KMS key from the source Region as AWS KMS keys are specific to an AWS Region.

Copying AWS KMS–encrypted snapshots to another AWS Region

AWS KMS keys are specific to an AWS Region. If you enable copying of Amazon Redshift snapshots to another AWS Region, and the source cluster and its snapshots are encrypted using a master key from AWS KMS, you need to configure a grant for Amazon Redshift to use a master key in the destination AWS Region. This grant enables Amazon Redshift to encrypt snapshots in the destination AWS Region. For more information about cross-Region snapshot copy, see [Copying snapshots to another AWS Region](#).

Note

If you enable copying of snapshots from an encrypted cluster and use AWS KMS for your master key, you cannot rename your cluster because the cluster name is part of the encryption context. If you must rename your cluster, you can disable copying of snapshots in the source AWS Region, rename the cluster, and then configure and enable copying of snapshots again.

The process to configure the grant for copying snapshots is as follows.

1. In the destination AWS Region, create a snapshot copy grant by doing the following:
 - If you do not already have an AWS KMS key to use, create one. For more information about creating AWS KMS keys, see [Creating Keys](#) in the *AWS Key Management Service Developer Guide*.
 - Specify a name for the snapshot copy grant. This name must be unique in that AWS Region for your AWS account.
 - Specify the AWS KMS key ID for which you are creating the grant. If you do not specify a key ID, the grant applies to your default key.
2. In the source AWS Region, enable copying of snapshots and specify the name of the snapshot copy grant that you created in the destination AWS Region.

This preceding process is only necessary if you enable copying of snapshots using the AWS CLI, the Amazon Redshift API, or SDKs. If you use the console, Amazon Redshift provides the proper workflow to configure the grant when you enable cross-Region snapshot copy. For more information about configuring cross-Region snapshot copy for AWS KMS-encrypted clusters by using the console, see [Configure cross-Region snapshot copy for an AWS KMS–encrypted cluster](#).

via -

<https://docs.aws.amazon.com/redshift/latest/mgmt/working-with-db-encryption.html#configure-snapshot-copy-grant>

Incorrect options:

Create a snapshot copy grant in the source Region for a KMS key in the source Region. Configure Redshift cross-Region snapshots in the destination Region - As described above, you need to configure the Redshift cross-Region snapshot in the source Region and not the destination Region. Also, the snapshot copy grant must be set up in the destination Region for a KMS key in the destination Region.

Create an IAM role in destination Region with access to the KMS key in the source Region. Create a snapshot copy grant in the destination Region for this KMS key in the source Region. Configure Redshift cross-Region snapshots in the source Region - This has been added as a distractor as AWS KMS keys are specific to an AWS Region. You cannot create a snapshot copy grant in the destination Region for a KMS key in the source Region.

Create a snapshot copy grant in the destination Region for a KMS key in the destination Region. Configure Redshift cross-Region replication in the source Region - This has been added as a distractor as there is no such thing as cross-Region replication for Redshift. The concept of cross-Region replication (CRR) applies to Amazon S3.

Reference:

<https://docs.aws.amazon.com/redshift/latest/mgmt/working-with-db-encryption.html#configure-snapshot-copy-grant>

Question 54: **Incorrect**

An e-commerce company has hired an AWS Certified Solutions Architect Professional to design a dual-tier storage layer for its flagship application running on EC2 instances. One of the tiers of this storage layer is a data tier that should support a POSIX file system shared across many systems. The other tier of this storage layer is a service tier that supports static file content that requires block storage with more than a million IOPS.

Which of the following solutions represent the BEST combination of AWS services for this use-case? (Select two)

-

Use EBS volumes with Provisioned IOPS as the service tier of the storage layer

(Incorrect)

-

Use EFS as the data tier of the storage layer

(Correct)

- Use Amazon S3 as the data tier of the storage layer
- Use EC2 Instance Store as the service tier of the storage layer
(Correct)
- Use EC2 Instance Store as the data tier of the storage layer

Use EFS as the data tier of the storage layer

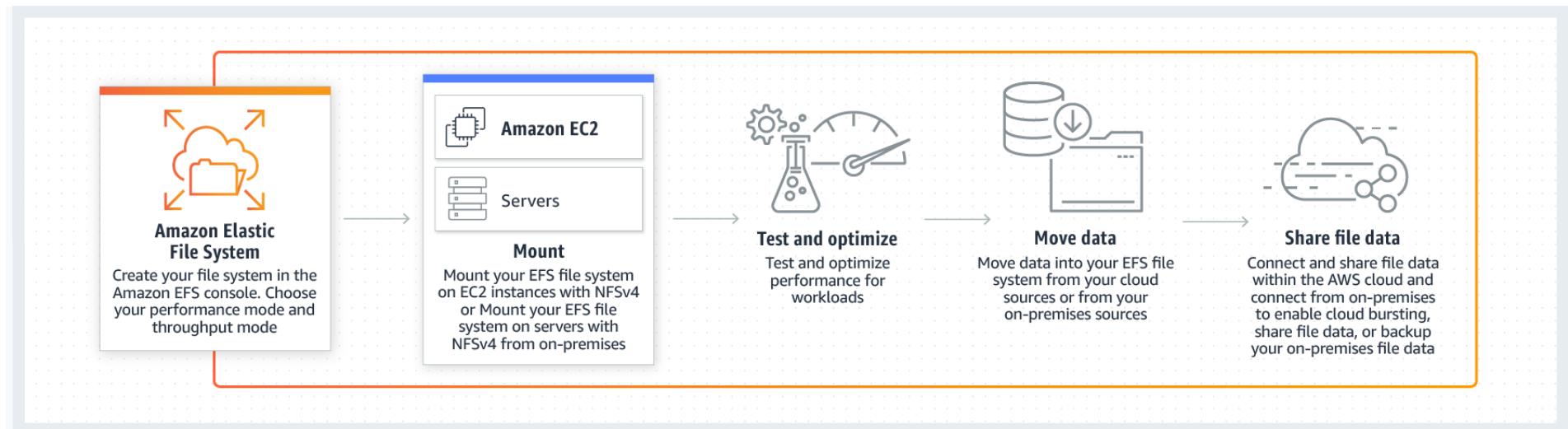
Explanation

Correct options:

Use EFS as the data tier of the storage layer

Amazon Elastic File System (Amazon EFS) provides a simple, scalable, fully managed elastic NFS file system for use with AWS Cloud services and on-premises resources.

Amazon EFS is a Regional service storing data within and across multiple Availability Zones (AZs) for high availability and durability. Amazon EC2 instances can access your file system across AZs, Regions, and VPCs, while on-premises servers can access using AWS Direct Connect or AWS VPN. You can connect to Amazon EFS file systems from EC2 instances in other AWS Regions using an inter-Region VPC peering connection, and from on-premises servers using an AWS VPN connection. EFS is also POSIX compliant and can be shared across many systems, so it fits the given use-case.



via - <https://aws.amazon.com/efs/>

Use EC2 Instance Store as the service tier of the storage layer

An instance store (also known as ephemeral storage) provides temporary block-level storage for your instance. This storage is located on disks that are physically attached to the host computer. Instance store is ideal for the temporary storage of information that changes frequently such as buffers, caches, scratch data, and other temporary content, or for data that is replicated across a fleet of instances, such as a load-balanced pool of web servers. Instance store volumes are included as part of the instance's usage cost.

As Instance Store based volumes provide high random I/O performance at low cost (as the storage is part of the instance's usage cost) and the fault-tolerant architecture can adjust for the loss of any instance, therefore you should use Instance Store based EC2 instances for this use-case.

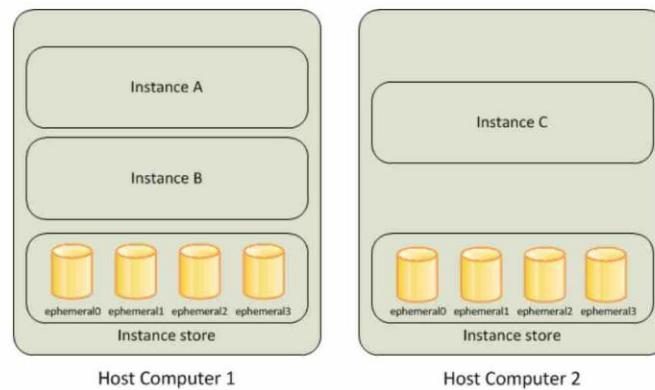
Amazon EC2 Instance Store

[PDF](#) | [Kindle](#) | [RSS](#)

An *instance store* provides temporary block-level storage for your instance. This storage is located on disks that are physically attached to the host computer. Instance store is ideal for temporary storage of information that changes frequently, such as buffers, caches, scratch data, and other temporary content, or for data that is replicated across a fleet of instances, such as a load-balanced pool of web servers.

An instance store consists of one or more instance store volumes exposed as block devices. The size of an instance store as well as the number of devices available varies by instance type.

The virtual devices for instance store volumes are `ephemeral[0-23]`. Instance types that support one instance store volume have `ephemeral0`. Instance types that support two instance store volumes have `ephemeral0` and `ephemeral1`, and so on.



EC2 Instance Store Overview:

<https://docs.aws.amazon.com/AWSEC2/latest/UserGuide/InstanceStorage.html>

via -

Per the given use-case, the key requirement for the service tier of the storage layer is to support block storage with more than a million IOPS. The Max IOPS per volume supported by EBS is only 256K for provisioned IOPS SSD (io2 block express). On the other hand, SSD-based instance store volumes support more than a million IOPS for random reads. So, this option is correct.

EBS Volume

Summary:

Volume characteristics

The following table describes the use cases and performance characteristics for each volume type. The default volume type is General Purpose SSD (gp2).

	Solid-state drives (SSD)			Hard disk drives (HDD)	
Volume type	General Purpose SSD (gp2)	Provisioned IOPS SSD		Throughput Optimized HDD (st1)	Cold HDD (sc1)
Description	General purpose SSD volume that balances price and performance for a wide variety of workloads	Highest-performance SSD volume for mission-critical low-latency or high-throughput workloads		Low-cost HDD volume designed for frequently accessed, throughput-intensive workloads	Lowest cost HDD volume designed for less frequently accessed workloads
Durability	99.8% - 99.9% durability (0.1% - 0.2% annual failure rate)	99.999% durability (0.001% annual failure rate)	99.8% - 99.9% durability (0.1% - 0.2% annual failure rate)	99.8% - 99.9% durability (0.1% - 0.2% annual failure rate)	99.8% - 99.9% durability (0.1% - 0.2% annual failure rate)
Use cases	<ul style="list-style-type: none"> Recommended for most workloads System boot volumes Virtual desktops Low-latency interactive apps Development and test environments 	<ul style="list-style-type: none"> Critical business applications that require sustained IOPS performance, or more than 16,000 IOPS or 250 MiB/s of throughput per volume Large database workloads, such as: <ul style="list-style-type: none"> MongoDB Cassandra Microsoft SQL Server MySQL PostgreSQL Oracle 	<ul style="list-style-type: none"> Streaming workloads requiring consistent, fast throughput at a low price Big data Data warehouses Log processing Cannot be a boot volume 	<ul style="list-style-type: none"> Throughput-oriented storage for large volumes of data that is infrequently accessed Scenarios where the lowest storage cost is important Cannot be a boot volume 	
Amazon EBS Multi-attach	Not supported	Not Supported	Supported	Not supported	Not supported
API name	gp2	io2	io1	st1	sc1
Volume size	1 GiB - 16 TiB	4 GiB - 16 TiB		500 GiB - 16 TiB	500 GiB - 16 TiB
Dominant performance attribute	IOPS	IOPS		MiB/s	MiB/s
Max IOPS per volume	16,000 (16 KiB I/O) *	64,000 (16 KiB I/O) †		500 (1 MiB I/O)	250 (1 MiB I/O)
Max throughput per volume	250 MiB/s *	1,000 MiB/s †		500 MiB/s	250 MiB/s
Max IOPS per instance ‡‡	160,000				
Max throughput per instance ‡‡	4,750 MB/s				

* The throughput limit is between 128 MiB/s and 250 MiB/s, depending on the volume size. Volumes smaller than or equal to 170 GiB deliver a maximum throughput of 128 MiB/s. Volumes larger than 170 GiB but smaller than 334 GiB deliver a maximum throughput of 250 MiB/s if burst credits are available. Volumes larger than or equal to 334 GiB deliver 250 MiB/s regardless of burst credits. Older gp2 volumes might not reach full performance unless you modify the volume. For more information, see [Amazon EBS Elastic Volumes](#).

† Maximum IOPS and throughput are guaranteed only on Instances built on the Nitro System provisioned with more than 32,000 IOPS. Other instances guarantee up to 32,000 IOPS and 500 MiB/s. Older io1 volumes might not reach full performance unless you modify the volume. For more information, see [Amazon EBS Elastic Volumes](#).

‡‡ To achieve this throughput, you must have an instance that supports [EBS optimization](#).

via - <https://docs.aws.amazon.com/AWSEC2/latest/UserGuide/ebs-volume-types.html>

SSD I/O performance

If you use a Linux AMI with kernel version 4.4 or later and use all the SSD-based instance store volumes available to your instance, you get the IOPS (4,096 byte block size) performance listed in the following table (at queue depth saturation). Otherwise, you get lower IOPS performance.

Instance Size	100% Random Read IOPS	Write IOPS
i3.large *	100,125	35,000
i3.xlarge *	206,250	70,000
i3.2xlarge	412,500	180,000
i3.4xlarge	825,000	360,000
i3.8xlarge	1.65 million	720,000
i3.16xlarge	3.3 million	1.4 million
i3.metal	3.3 million	1.4 million
i3en.large *	42,500	32,500
i3en.xlarge *	85,000	65,000
i3en.2xlarge *	170,000	130,000
i3en.3xlarge	250,000	200,000
i3en.6xlarge	500,000	400,000
i3en.12xlarge	1 million	800,000
i3en.24xlarge	2 million	1.6 million
i3en.metal	2 million	1.6 million

<https://docs.aws.amazon.com/AWSEC2/latest/UserGuide/storage-optimized-instances.html>

Incorrect options:

Use EBS volumes with Provisioned IOPS as the service tier of the storage layer - As mentioned in the explanation above, the Max IOPS per volume supported by EBS is 256K for provisioned IOPS SSD (io2 block express), so this option is incorrect.

Use EC2 Instance Store as the data tier of the storage layer - This option is incorrect as Instance Store cannot be used as data tier for the given use-case because it cannot be shared across many systems at the same time. This capability is only offered by EFS.

Use Amazon S3 as the data tier of the storage layer - This option is incorrect as S3 is not POSIX compliant.

References:

<https://docs.aws.amazon.com/AWSEC2/latest/UserGuide/InstanceStorage.html>

<https://docs.aws.amazon.com/AWSEC2/latest/UserGuide/ebs-volume-types.html>

<https://docs.aws.amazon.com/AWSEC2/latest/UserGuide/storage-optimized-instances.html>

Question 55: **Correct**

A blog hosting company has an existing SaaS product architected as an on-premises three-tier web application. The blog content is posted and updated several times a day by multiple authors, so the Linux web servers serve content from a centralized file share on a NAS server. The CTO at the company has done an extensive technical review and highlighted to the company management that the existing infrastructure is not optimized. The company would like to migrate to AWS so that the resources can be dynamically scaled in response to load. The on-premises infrastructure and AWS Cloud are connected using Direct Connect.

As a Solutions Architect Professional, which of the following solutions would you recommend to the company so that it can migrate the web infrastructure to AWS without delaying the content updation process?



Set up an on-premises file gateway using Storage Gateway to replace the NAS server and then replicate the existing content to AWS. On the AWS Cloud, mount the same Storage Gateway bucket to the EC2 instance based web servers to serve the content



Attach an EFS file system to the on-premises servers to act as the NAS server. Mount the same EFS file system to the AWS based web servers running on EC2 instances to serve the content

(Correct)



Provision a cluster of 20 EC2 instances based web servers running behind an Application Load Balancer on AWS across multiple Availability Zones. Share an EBS volume among all instances for accessing the content. Develop custom code to periodically synchronize this volume with the NAS server



Provision EC2 instances based web servers with an Auto Scaling group. Create a nightly data transfer batch job to update the web server instances from the NAS server

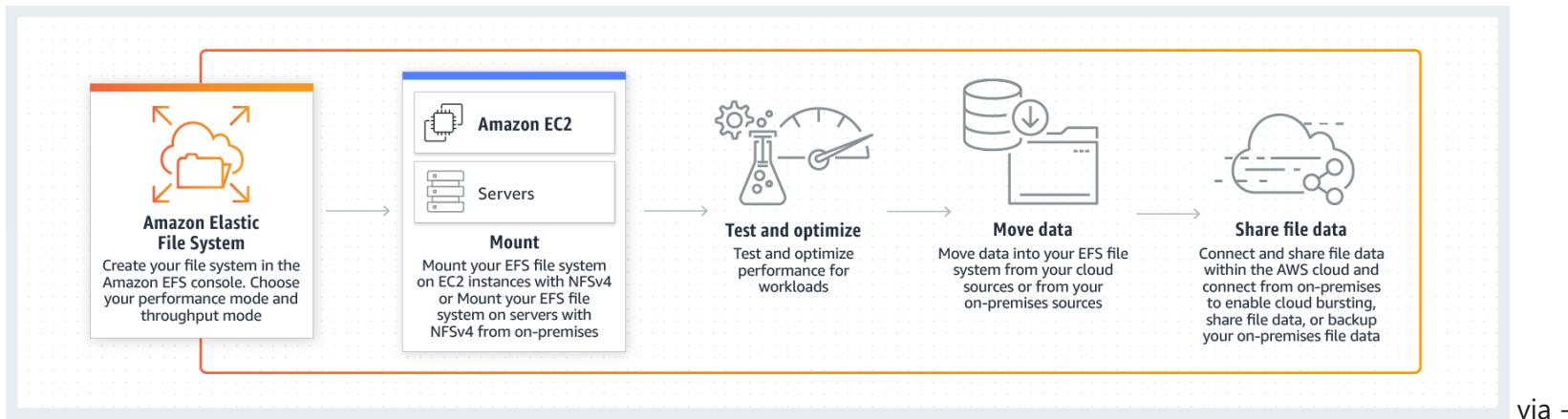
Explanation

Correct option:

Attach an EFS file system to the on-premises servers to act as the NAS server. Mount the same EFS file system to the AWS based web servers running on EC2 instances to serve the content

Amazon Elastic File System (Amazon EFS) provides a simple, scalable, fully managed elastic NFS file system for use with AWS Cloud services and on-premises resources.

Amazon EFS is a Regional service storing data within and across multiple Availability Zones (AZs) for high availability and durability. Amazon EC2 instances can access your file system across AZs, Regions, and VPCs, while on-premises servers can access using AWS Direct Connect or AWS VPN. You can connect to Amazon EFS file systems from EC2 instances in other AWS Regions using an inter-Region VPC peering connection, and from on-premises servers using an AWS VPN connection. EFS is also POSIX compliant.



via -

<https://aws.amazon.com/efs/>

For the given use-case, you can attach an EFS file system to your on-premises servers, copy your data to it, and then process it in the cloud as desired, leaving your data in AWS for the long term. Further, you can mount the EFS file system from the EC2 instances for a concurrent access. Connecting to EFS is similar to connecting to your network drive since it supports NFS protocols, which are standard for network attached storage (NAS) devices. This ensures that the company can migrate the web infrastructure to AWS Cloud without delaying the content updation process as the underlying workflows do not need to be modified.

Amazon EFS Update – On-Premises Access via Direct Connect

by Jeff Barr | on 20 DEC 2016 | in Amazon EC2, Amazon Elastic File System (EFS), AWS Direct Connect, AWS Re:Invent | Permalink |  Share

▶ 0:00 / 0:00

Voice by [Amazon Polly](#)

I introduced you to [Amazon Elastic File System \(EFS\)](#) last year ([Amazon Elastic File System – Shared File Storage for Amazon EC2](#)) and announced production readiness earlier this year ([Amazon Elastic File System – Production-Ready in Three Regions](#)). Since the launch earlier this year, thousands of AWS customers have used it to set up, scale, and operate shared file storage in the cloud.



Elastic File System
Fully Managed File System for EC2

Today we are making EFS even more useful with the introduction of simple and reliable on-premises access via [AWS Direct Connect](#). This has been a much-requested feature and I know that it will be useful for migration, cloudbursting, and backup. **To use this feature for migration, you simply attach an EFS file system to your on-premises servers, copy your data to it, and then process it in the cloud as desired, leaving your data in AWS for the long term.** For cloudbursting, you would copy on-premises data to an EFS file system, analyze it at high speed using a fleet of [Amazon Elastic Compute Cloud \(EC2\)](#) instances, and then copy the results back on-premises or visualize them in [Amazon QuickSight](#).

You'll get the same file system access semantics including strong consistency and file locking, whether you access your EFS file systems from your on-premises servers or from your EC2 instances (of course, you can do both concurrently). You will also be able to enjoy the same multi-AZ availability and durability that is part-and-parcel of EFS.

via - <https://aws.amazon.com/blogs/aws/amazon-efs-update-on-premises-access-via-direct-connect-vpc/>

Incorrect options:

Set up an on-premises file gateway using Storage Gateway to replace the NAS server and then replicate the existing content to AWS. On the AWS Cloud, mount the same Storage Gateway bucket to the EC2 instance based web servers to serve the content

AWS Storage Gateway is a hybrid cloud storage service that gives you on-premises access to virtually unlimited cloud storage. The service provides three different types of gateways – Tape Gateway, File Gateway, and Volume Gateway – that seamlessly connect on-premises applications to cloud storage, caching data locally for low-latency access.

AWS Storage Gateway's file interface, or file gateway, offers you a seamless way to connect to the cloud in order to store application data files and backup images as durable objects on Amazon S3 cloud storage. File gateway offers SMB or NFS-based access to data in Amazon S3 with local caching.

File Gateway Overview: via - <https://docs.aws.amazon.com/storagegateway/latest/userguide/StorageGatewayConcepts.html>

The issue with transitioning to Storage Gateway is that you would run into performance issues once the local cache fills up and then the application has to source the data from S3 which is the underlying object based storage. Moreover, S3 is not POSIX compliant and it does not support operations such as file append. The file gateway takes care of these abstractions but it also adds up to making this architecture not as scalable as just mounting EFS on both the on-premises servers as well as EC2 instances. So this option is incorrect.

Provision a cluster of 20 EC2 instances based web servers running behind an Application Load Balancer on AWS across multiple Availability Zones. Share an EBS volume among all instances for accessing the content. Develop custom code to periodically synchronize this volume with the NAS server - You cannot share an EBS volume with multiple instances (unless it's a nitro based instance. Even for nitro based instances, you can only share an EBS volume with up to 16 instances in the same Availability Zone). So this option is incorrect.

Provision EC2 instances based web servers with an Auto Scaling group. Create a nightly data transfer batch job to update the web server instances from the NAS server - Using a nightly data transfer batch job to update the web server instances from the NAS server implies that the solution would delay the content updation process, which is a key requirement of the use-case. So this option is incorrect.

References:

<https://aws.amazon.com/efs/>

<https://aws.amazon.com/blogs/aws/amazon-efs-update-on-premises-access-via-direct-connect-vpc/>

<https://aws.amazon.com/about-aws/whats-new/2017/02/aws-storage-gateway-supports-running-file-gateway-in-ec2-and-adds-file-share-security-options/>

Question 56: **Incorrect**

A global SaaS company has recently migrated its technology infrastructure from its on-premises data center to AWS Cloud. The engineering team has provisioned an RDS MySQL DB cluster for the company's flagship application. An analytics workload also runs on the same database which publishes near real-time reports for the management of the company. When the analytics workload runs, it slows down the SaaS application as well, resulting in bad user experience.

As a Solutions Architect Professional, which of the following would you recommend as the MOST cost-optimal solution to fix this issue?



For Disaster Recovery purposes, create a Read Replica in another Region as the Master database and point the analytics workload there
(Incorrect)



Migrate the analytics application to AWS Lambda



Enable Multi-AZ for the RDS database and run the analytics workload on the standby database



Create a Read Replica in the same Region as the Master database and point the analytics workload there

(Correct)

Explanation

Correct option:

Create a Read Replica in the same Region as the Master database and point the analytics workload there

Amazon RDS Read Replicas provide enhanced performance and durability for RDS database (DB) instances. They make it easy to elastically scale out beyond the capacity constraints of a single DB instance for read-heavy database workloads. For the MySQL, MariaDB, PostgreSQL, Oracle, and SQL Server database engines, Amazon RDS creates a second DB instance using a snapshot of the source DB instance. It then uses the engines' native asynchronous replication to update the read replica whenever there is a change to the source DB instance. Read replicas can be within an Availability Zone, Cross-AZ, or Cross-Region.

Creating a Read Replica within the same Region is the correct answer. As we want to minimize the costs, we need to launch the Read Replica in the same Region, because we have to pay for inter-Region data transfer, whereas the transfer of data within a single Region is free.

Incorrect options:

Enable Multi-AZ for the RDS database and run the analytics workload on the standby database - Amazon RDS Multi-AZ deployments provide enhanced availability and durability for RDS database (DB) instances, making them a natural fit for production database workloads. When you provision a Multi-AZ DB Instance, Amazon RDS automatically creates a primary DB Instance and synchronously replicates the data to a standby instance in a different Availability Zone (AZ). Multi-AZ spans at least two Availability Zones within a single region.

Enabling Multi-AZ helps make our database highly-available, but the standby database is not accessible and cannot be used for reads or write. It's just a database that will become primary when the other database encounters a failure. So this option is not correct.

Migrate the analytics application to AWS Lambda- AWS Lambda lets you run code without provisioning or managing servers. You pay only for the compute time you consume. Running the application on AWS Lambda will not help, as it will still run against the main database and slow down the application.

For Disaster Recovery purposes, create a Read Replica in another Region as the Master database and point the analytics workload there - This option is not correct because we have to pay for inter-Region data transfer for the Read Replica, whereas the transfer of data within a single Region is free. Disaster Recover is not within the scope of the requirements mentioned for the given-use. The correct solution needs to optimize costs.

References:

<https://aws.amazon.com/rds/faqs/>

<https://aws.amazon.com/rds/features/multi-az/>

<https://aws.amazon.com/rds/features/read-replicas/>

Question 57: **Correct**

A leading gaming company runs multiple game platforms that need to store game state, player data, session history, and leaderboards. The company is looking to move to AWS Cloud to scale reliably to millions of concurrent users and requests while ensuring consistently low latency measured in single-digit milliseconds. The engineering team at the company is evaluating multiple in-memory data stores with the ability to power its on-demand, live leaderboard. The company's leaderboard requires high availability, low latency, and real-time processing to deliver customizable user data for the community of its users.

As an AWS Certified Solutions Architect Professional, which of the following solutions would you recommend? (Select two)

- **Develop the leaderboard using DynamoDB as it meets the in-memory, high availability, low latency requirements**
 - **Develop the leaderboard using AWS Neptune as it meets the in-memory, high availability, low latency requirements**
 - **Develop the leaderboard using DynamoDB with DynamoDB Accelerator (DAX) as it meets the in-memory, high availability, low latency requirements**
- (Correct)

-

Develop the leaderboard using RDS Aurora as it meets the in-memory, high availability, low latency requirements

-

Develop the leaderboard using ElastiCache Redis as it meets the in-memory, high availability, low latency requirements

(Correct)

Explanation

Correct options:

Develop the leaderboard using ElastiCache Redis as it meets the in-memory, high availability, low latency requirements

Amazon ElastiCache for Redis is a blazing fast in-memory data store that provides sub-millisecond latency to power internet-scale real-time applications. Amazon ElastiCache for Redis is a great choice for real-time transactional and analytical processing use cases such as caching, chat/messaging, gaming leaderboards, geospatial, machine learning, media streaming, queues, real-time analytics, and session store. ElastiCache for Redis can be used to power the live leaderboard, so this option is correct.

ElastiCache for Redis

Overview:

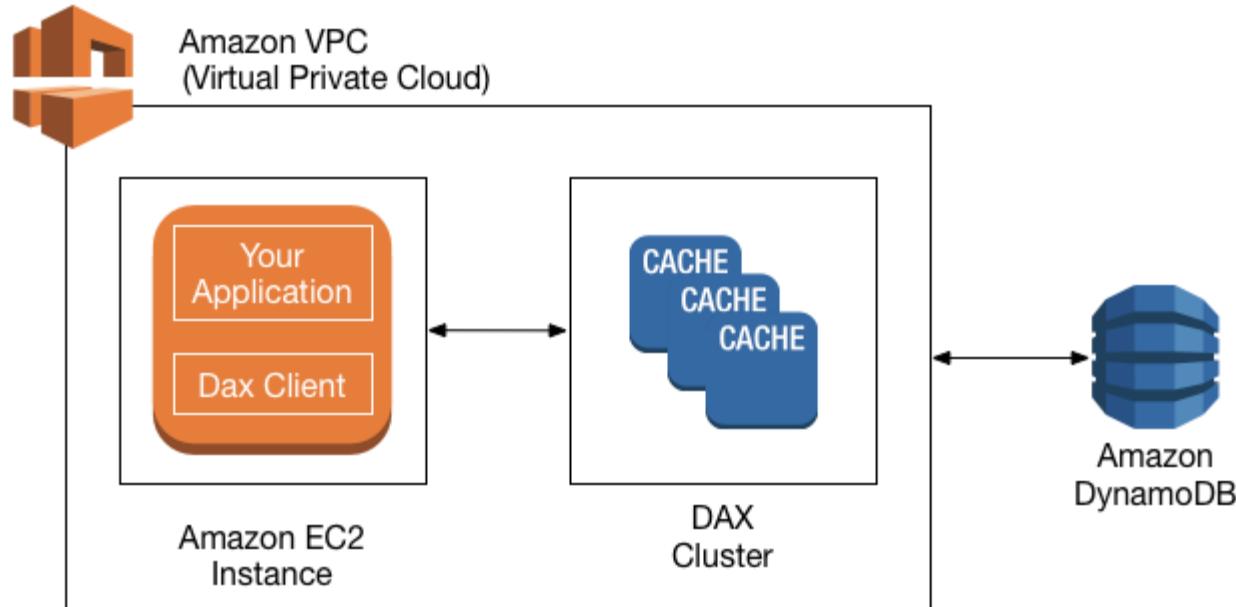


Develop the leaderboard using DynamoDB with DynamoDB Accelerator (DAX) as it meets the in-memory, high availability, low latency requirements

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

DAX is a DynamoDB-compatible caching service that enables you to benefit from fast in-memory performance for demanding applications. So DynamoDB with DAX can be used to power the live leaderboard.

DAX



Overview:

via -

<https://docs.aws.amazon.com/amazondynamodb/latest/developerguide/DAX.concepts.html>

Incorrect options:

Develop the leaderboard using AWS Neptune as it meets the in-memory, high availability, low latency requirements -

Amazon Neptune is a fast, reliable, fully-managed graph database service that makes it easy to build and run applications that work with highly connected datasets. Neptune is not an in-memory database, so this option is not correct.

Develop the leaderboard using DynamoDB as it meets the in-memory, high availability, low latency requirements -

DynamoDB is not an in-memory database, so this option is not correct.

Develop the leaderboard using RDS Aurora as it meets the in-memory, high availability, low latency requirements - Amazon Aurora is a MySQL and PostgreSQL-compatible relational database built for the cloud, that combines the performance and availability of traditional enterprise databases with the simplicity and cost-effectiveness of open source databases. Amazon Aurora features a distributed, fault-tolerant, self-healing storage system that auto-scales up to 64TB per database instance. Aurora is not an in-memory database, so this option is not correct.

References:

<https://aws.amazon.com/elasticache/>

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

<https://aws.amazon.com/dynamodb/dax/>

Question 58: **Incorrect**

After a recent DDoS assault, the IT security team of a media company has asked the Security Engineer to revamp the security of the application to prevent future attacks. The website is hosted on an Amazon EC2 instance and data is maintained on Amazon RDS. A large part of the application data is static and this data is in the form of images.

Which of the following steps can be combined to constitute the revamped security model? (Select two)

-

Use Amazon Route 53 to distribute traffic

(Correct)

-

Use Global Accelerator to distribute traffic

-

Move the static content to Amazon S3, and front this with an Amazon CloudFront distribution. Configure another layer of protection by adding AWS Web Application Firewall (AWS WAF) to the CloudFront distribution

(Correct)



Configure the Amazon EC2 instance with an Auto Scaling Group (ASG) to scale in case of a DDoS assault. Front the ASG with AWS Web Application Firewall (AWS WAF) for another layer of security

(Incorrect)



Configure Amazon Inspector with AWS Security Hub to mitigate DDoS attacks by continual scanning that delivers near real-time vulnerability findings

Explanation

Correct options:

Use Amazon Route 53 to distribute traffic

Move the static content to Amazon S3, and front this with an Amazon CloudFront distribution. Configure another layer of protection by adding AWS Web Application Firewall (AWS WAF) to the CloudFront distribution

AWS WAF is a web application firewall that helps protect web applications from attacks by allowing you to configure rules that allow, block, or monitor (count) web requests based on conditions that you define. These conditions include IP addresses, HTTP headers, HTTP body, URI strings, SQL injection, and cross-site scripting.

AWS WAF is tightly integrated with Amazon CloudFront, the Application Load Balancer (ALB), Amazon API Gateway, and AWS AppSync – services that AWS customers commonly use to deliver content for their websites and applications. When you use AWS WAF on Amazon CloudFront, your rules run in all AWS Edge Locations, located around the world close to your end users. Blocked requests are stopped before they reach your web servers.

Route 53 DNS requests and subsequent application traffic routed through CloudFront are inspected inline. Always-on monitoring, anomaly detection, and mitigation against common infrastructure DDoS attacks such as SYN/ACK floods, UDP floods, and reflection attacks are built into both Route 53 and CloudFront.

Route 53 is also designed to withstand DNS query floods, which are real DNS requests that can continue for hours and attempt to exhaust DNS server resources. Route 53 uses shuffle sharding and anycast striping to spread DNS traffic across edge locations and help protect the availability of the service.

When used with Amazon CloudFront distribution, AWS Shield adds security against DDoS attacks.

AWS Shield is a managed Distributed Denial of Service (DDoS) protection service that safeguards applications running on AWS. AWS Shield provides always-on detection and automatic inline mitigations that minimize application downtime and latency, so there is no need to engage AWS Support to benefit from DDoS protection.

All AWS customers benefit from the automatic protections of AWS Shield Standard, at no additional charge. AWS Shield Standard defends against the most common, frequently occurring network and transport layer DDoS attacks that target your website or applications. When you use AWS Shield Standard with Amazon CloudFront and Amazon Route 53, you receive comprehensive availability protection against all known infrastructure (Layer 3 and 4) attacks.

Incorrect options:

Configure the Amazon EC2 instance with an Auto Scaling Group (ASG) to scale in case of a DDoS assault. Front the ASG with AWS Web Application Firewall (AWS WAF) for another layer of security - AWS WAF is tightly integrated with Amazon CloudFront, the Application Load Balancer (ALB), Amazon API Gateway, and AWS AppSync – services that AWS customers commonly use to deliver content for their websites and applications. WAF cannot be directly configured in front of an ASG, so this option is incorrect.

Use Global Accelerator to distribute traffic - Global Accelerator is effective in traffic distribution across AWS Regions. However, the given use case needs services that can help mitigate DDoS attacks.

Configure Amazon Inspector with AWS Security Hub to mitigate DDoS attacks by continual scanning that delivers near real-time vulnerability findings - Amazon Inspector is an automated vulnerability management service that continually scans Amazon

Elastic Compute Cloud (EC2) and container workloads for software vulnerabilities and unintended network exposure. It cannot be used to mitigate DDoS attacks.

References:

<https://aws.amazon.com/shield/>

<https://aws.amazon.com/blogs/security/how-to-protect-dynamic-web-applications-against-ddos-attacks-by-using-amazon-cloudfront-and-amazon-route-53/>

<https://aws.amazon.com/waf/faqs/>

Question 59: **Incorrect**

A social media company is transitioning its IT infrastructure from its on-premises data center to the AWS Cloud. The company wants to move its data artifacts, 200 TB in total size, to Amazon S3 on the AWS Cloud in the shortest possible time. The company has hired you as an AWS Certified Solutions Architect Professional to provide consultancy for this data migration. In terms of the networking infrastructure, the company has a 500 Mbps Direct Connect connection to the AWS Cloud as well as an IPSec based AWS VPN connection using the public internet that supports a bandwidth of 1 Gbps.

Which of the following solutions would you recommend to address the given use-case?

-

Leverage S3 Transfer Acceleration to transfer the data to S3

-

Order three AWS Snowball Edge appliances, split and transfer the data to these three appliances and ship them to AWS which will then copy the data from the Snowball Edge appliances to S3

(Correct)

- Leverage the 1Gbps IPSec based AWS VPN connection to transfer the data to S3 over the public internet
- Leverage the 500 Mbps Direct Connect connection to transfer the data to S3 over the dedicated connection

(Incorrect)

Explanation

Correct option:

Order three AWS Snowball Edge appliances, split and transfer the data to these three appliances and ship them to AWS which will then copy the data from the Snowball Edge appliances to S3

The AWS Snowball Edge is a type of Snowball device with on-board storage and compute power for select AWS capabilities. Each Snowball Edge device can transport data at speeds faster than the internet. This transport is done by shipping the data in the appliances through a Regional carrier. The appliances are rugged shipping containers, complete with E Ink shipping labels. Snowball Edge devices have three options for device configurations – storage optimized, compute optimized, and with GPU.

Snowball Edge is the optimal choice if you need to securely and quickly transfer dozens of terabytes to petabytes of data to AWS. The AWS Snow Family is ideal for customers moving large batches of data at once. The AWS Snowball has a typical 5-7 days turnaround time. As each Snowball Edge device can handle 80TB of data, you can order 3 such devices to take care of the data transfer for the given use-case.

Incorrect options:

Leverage the 1Gbps IPSec based AWS VPN connection to transfer the data to S3 over the public internet

The given use-case requires transferring 200 TB of archived data within the shortest possible duration. Now, the IPSec based AWS VPN connection assures a speed of 1 Gbps (Gigabits per second). As 1 Byte = 8 bits, therefore you can transfer 0.125 GBps (Gigabytes per second).

So the hourly data transfer is $0.125 \times 60 \times 60 = 450$ GB or approximately 0.45 TB.

So the approximate daily data transfer is $0.45 \times 24 = \sim 10$ TB

Therefore, the entire archived dataset of 200 TB can be transferred in 20 days. As explained above, using Snowball appliances, you can complete this data transfer in just 5-7 days. So this option is incorrect.

Leverage the 500 Mbps Direct Connect connection to transfer the data to S3 over the dedicated connection - We established above that a 1GBps connection would take 20 days to transfer the dataset, so the 500 Mbps Direct Connect connection would take 40 days to transfer the dataset. So this option is incorrect.

Exam Alert:

The exam would probe you on multiple scenarios around competing options to migrate data from on-premises data centers to AWS Cloud. *You should remember that a 1Gbps connection at full utilization can transfer approximately 10 TB of data in a day.* You will see questions that tweak the available connection bandwidth or change the available utilization, so you just need to factor these changes into the above-mentioned unit rate to identify the correct answer.

Leverage S3 Transfer Acceleration to transfer the data to S3 - S3 Transfer Acceleration enables fast, easy, and secure transfers of files over long distances between your client and your Amazon S3 bucket. S3 Transfer Acceleration leverages Amazon CloudFront's globally distributed AWS Edge Locations.

You should note that S3 Transfer Acceleration over a fully-utilized 1 Gbps line can only transfer up to 75 TBs in a week's duration. So this option is not correct for the given use-case.

Q: How should I choose between S3 Transfer Acceleration and AWS Snow Family (Snowball, Snowball Edge, and Snowmobile)?

The AWS Snow Family is ideal for customers moving large batches of data at once. The AWS Snowball has a typical 5-7 days turnaround time. As a rule of thumb, S3 Transfer Acceleration over a fully-utilized 1 Gbps line can transfer up to 75 TBs in the same time period. In general, if it will take more than a week to transfer over the Internet, or there are recurring transfer jobs and there is more than 25Mbps of available bandwidth, S3 Transfer Acceleration is a good option. Another option is to use both: perform initial heavy lift moves with an AWS Snowball (or series of AWS Snowballs) and then transfer incremental ongoing changes with S3 Transfer Acceleration.

via -

https://aws.amazon.com/s3/faqs/#Amazon_S3_Transfer_Acceleration

References:

https://aws.amazon.com/s3/faqs/#Amazon_S3_Transfer_Acceleration

Question 60: **Incorrect**

A leading Internet-of-Things (IoT) solutions company needs to develop a platform that would analyze real-time clickstream events from embedded sensors in consumer electronic devices. The company has hired you as an AWS Certified Solutions Architect Professional to consult the engineering team and develop a solution using the AWS Cloud. The company wants to use clickstream data to perform data science, develop algorithms, and create visualizations and dashboards to support the business stakeholders. Each of these groups would work independently and would need real-time access to this clickstream data for their applications.

Which of the following options would provide a highly available and fault-tolerant solution to capture the clickstream events from the source and also provide a simultaneous feed of the data stream to the downstream applications?



Use AWS Kinesis Data Analytics to facilitate multiple applications consume and analyze same streaming data concurrently and independently

(Incorrect)

-
-

Use AWS Kinesis Data Firehose to allow applications to consume the same streaming data concurrently and independently

-
-

Use AWS Kinesis Data Streams to facilitate multiple applications consume same streaming data concurrently and independently

(Correct)

-
-

Use Amazon SQS to facilitate multiple applications process same streaming data concurrently and independently

Explanation

Correct option:

Use AWS Kinesis Data Streams to facilitate multiple applications consume the same streaming data concurrently and independently

Amazon Kinesis Data Streams (KDS) is a massively scalable and durable real-time data streaming service. KDS can continuously capture gigabytes of data per second from hundreds of thousands of sources such as website clickstreams, database event streams, financial transactions, social media feeds, IT logs, and location-tracking events. The data collected is available in milliseconds to enable real-time analytics use cases such as real-time dashboards, real-time anomaly detection, dynamic pricing, and more.

Amazon Kinesis Data Streams enables real-time processing of streaming big data. It provides ordering of records, as well as the ability to read and/or replay records in the same order to multiple Amazon Kinesis Applications. The Amazon Kinesis Client Library (KCL) delivers all records for a given partition key to the same record processor, making it easier to build multiple applications reading from the same Amazon Kinesis data stream (for example, to perform counting, aggregation, and filtering).

Amazon Kinesis Data Streams is recommended when you need the ability for multiple applications to consume the same stream concurrently. For example, you have one application that updates a real-time dashboard and another application that archives data to Amazon Redshift. You want both applications to consume data from the same stream concurrently and independently.

KDS provides the ability for multiple applications to consume the same stream

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

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

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

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

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

concurrently
streams/faqs/

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

Incorrect options:

Use AWS Kinesis Data Firehose to allow applications to consume same streaming data concurrently and independently -

Amazon Kinesis Data Firehose is the easiest way to load streaming data into data stores and analytics tools. It can capture, transform, and load streaming data into Amazon S3, Amazon Redshift, Amazon Elasticsearch Service, and Splunk, enabling near real-time analytics with existing business intelligence tools and dashboards you're already using today. It is a fully managed service that automatically scales to match the throughput of your data and requires no ongoing administration. It can also batch, compress, and encrypt the data before loading it, minimizing the amount of storage used at the destination and increasing security. As Kinesis Data Firehose is used to load streaming data into data stores, therefore this option is incorrect.

Use AWS Kinesis Data Analytics to facilitate multiple applications consume and analyze same streaming data concurrently and independently -

Amazon Kinesis Data Analytics is the easiest way to analyze streaming data in real-time. You can quickly build SQL queries and sophisticated Java applications using built-in templates and operators for common processing functions to organize, transform, aggregate, and analyze data at any scale. Kinesis Data Analytics enables you to easily and quickly build queries and sophisticated streaming applications in three simple steps: setup your streaming data sources, write your queries or streaming applications and set up your destination for processed data. As Kinesis Data Analytics is used to build SQL queries on streaming data, therefore this option is incorrect.

Use Amazon SQS to facilitate multiple applications process same streaming data concurrently and independently -

Amazon Simple Queue Service (SQS) is a fully managed message queuing service that enables you to decouple and scale microservices, distributed systems, and serverless applications. SQS offers two types of message queues. Standard queues offer maximum throughput, best-effort ordering, and at-least-once delivery. SQS FIFO queues are designed to guarantee that messages are processed exactly once, in the exact order that they are sent. For SQS, you cannot have the same message being consumed by multiple consumers at the same time, therefore this option is incorrect.

References:

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

<https://aws.amazon.com/kinesis/data-firehose/faqs/>

<https://aws.amazon.com/kinesis/data-analytics/faqs/>

Question 61: **Correct**

A company runs its two-tier web application from an on-premises data center. The web servers connect to a PostgreSQL database running on a different server. With the consistent increase in users, both the web servers and the database are underperforming leading to a bad user experience. The company has decided to migrate to AWS Cloud and has chosen Amazon Aurora PostgreSQL as its database solution. The company needs a solution that can scale the web servers and the database layer based on user traffic.

Which of the following options will you combine to improve the application scalability and improve the user experience? (Select two)

- **Configure EC2 instances behind an Application Load Balancer with Round Robin routing algorithm and sticky sessions enabled**
(Correct)
- **Configure EC2 instances behind an Application Load Balancer with flow hash routing algorithm and sticky sessions enabled**
- **Configure EC2 instances behind a Network Load Balancer with Least Outstanding Requests routing algorithm and sticky sessions enabled**
- **Enable Aurora Auto Scaling for Aurora Writes. Deploy the application on Amazon EC2 instances configured behind an Auto Scaling Group**
- **Enable Aurora Auto Scaling for Aurora Replicas. Deploy the application on Amazon EC2 instances configured behind an Auto Scaling Group**

(Correct)

Explanation

Correct options:

Enable Aurora Auto Scaling for Aurora Replicas. Deploy the application on Amazon EC2 instances configured behind an Auto Scaling Group - To meet your connectivity and workload requirements, Aurora Auto Scaling dynamically adjusts the number of Aurora Replicas provisioned for an Aurora DB cluster using single-master replication. Aurora Auto Scaling enables your Aurora DB cluster to handle sudden increases in connectivity or workload. When the connectivity or workload decreases, Aurora Auto Scaling removes unnecessary Aurora Replicas so that you don't pay for unused provisioned DB instances.

You define and apply a scaling policy to an Aurora DB cluster. The scaling policy defines the minimum and maximum number of Aurora Replicas that Aurora Auto Scaling can manage. Based on the policy, Aurora Auto Scaling adjusts the number of Aurora Replicas up or down in response to actual workloads, determined by using Amazon CloudWatch metrics and target values.

Configure EC2 instances behind an Application Load Balancer with Round Robin routing algorithm and sticky sessions enabled - Your load balancer serves as a single point of contact for clients and distributes incoming traffic across its healthy registered targets. You can register each target with one or more target groups.

By default, the round-robin routing algorithm is used to route requests at the target group level. Round robin is a good choice when the requests and targets are similar, or if you need to distribute requests equally among targets.

By default, an Application Load Balancer (ALB) routes each request independently to a registered target based on the chosen load-balancing algorithm. However, you can use the sticky session feature (also known as session affinity) to enable the load balancer to bind a user's session to a specific target. This ensures that all requests from the user during the session are sent to the same target. This feature is useful for servers that maintain state information to provide a continuous experience to clients. To use sticky sessions, the client must support cookies.

Incorrect options:

Enable Aurora Auto Scaling for Aurora Writes. Deploy the application on Amazon EC2 instances configured behind an Auto Scaling Group - Aurora Auto Scaling is possible for Aurora replicas and not for Aurora writer instances. Multi-master Aurora Cluster architecture is needed if multiple writers are needed for any use case.

Configure EC2 instances behind an Application Load Balancer with flow hash routing algorithm and sticky sessions enabled - The flow hash routing algorithm can only be used with Network Load Balancers. So this option is incorrect.

Configure EC2 instances behind a Network Load Balancer with Least Outstanding Requests routing algorithm and sticky sessions enabled - This statement is incorrect. Network Load Balancer does not support Least Outstanding Requests routing algorithm. AWS suggests using the Least Outstanding Requests with an ALB when the requests for your application vary in complexity or your targets vary in processing capability.

For TCP traffic, the Network Load Balancer selects a target using a flow hash algorithm based on the protocol, source IP address, source port, destination IP address, destination port, and TCP sequence number. The TCP connections from a client have different source ports and sequence numbers and can be routed to different targets. Each TCP connection is routed to a single target for the life of the connection.

For UDP traffic, the Network Load Balancer selects a target using a flow hash algorithm based on the protocol, source IP address, source port, destination IP address, and destination port. A UDP flow has the same source and destination, so it is consistently routed to a single target throughout its lifetime. Different UDP flows have different source IP addresses and ports, so they can be routed to different targets.

References:

<https://docs.aws.amazon.com/AmazonRDS/latest/AuroraUserGuide/Aurora.Integrating.AutoScaling.html>

https://docs.amazonaws.cn/en_us/elasticloadbalancing/latest/application/load-balancer-target-groups.html#modify-routing-algorithm

https://docs.amazonaws.cn/en_us/elasticloadbalancing/latest/application/sticky-sessions.html

<https://aws.amazon.com/about-aws/whats-new/2019/11/application-load-balancer-now-supports-least-outstanding-requests-algorithm-for-load-balancing-requests/>

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

Question 62: **Incorrect**

A leading medical imaging equipment and diagnostic imaging solutions provider uses AWS Cloud to run its healthcare data flows through more than 500,000 medical imaging devices globally. The solutions provider stores close to one petabyte of medical imaging data on Amazon S3 to provide the durability and reliability needed for their critical data. A research assistant working with the radiology department is trying to upload a high-resolution image into S3 via the public internet. The image size is approximately 5GB. The research assistant is using S3 Transfer Acceleration (S3TA) for faster image upload. It turns out that S3TA did not result in an accelerated transfer.

Given this scenario, which of the following is correct regarding the charges for this image transfer?

-

The research assistant does not need to pay any transfer charges for the image upload

(Correct)

-

The research assistant only needs to pay S3 transfer charges for the image upload

-

The research assistant only needs to pay S3TA transfer charges for the image upload

-

The research assistant needs to pay both S3 transfer charges and S3TA transfer charges for the image upload

(Incorrect)

Explanation

Correct option: **The research assistant does not need to pay any transfer charges for the image upload**

There are no S3 data transfer charges when data is transferred in from the internet.

S3 Data Transfer

Pricing:

Storage	Requests and data retrievals	Data transfer	Management and replication
You pay for all bandwidth into and out of Amazon S3, except for the following:			
<ul style="list-style-type: none">• Data transferred in from the internet.• Data transferred out to an Amazon Elastic Compute Cloud (Amazon EC2) instance, when the instance is in the same AWS Region as the S3 bucket.• Data transferred out to Amazon CloudFront (CloudFront).			
The pricing below is based on data transferred "in" and "out" of Amazon S3 (over the public Internet)†††. Transfers between S3 buckets or from Amazon S3 to any service(s) within the same AWS Region are free. You also pay a fee for any data transferred using Amazon S3 Transfer Acceleration. Learn more about AWS Direct Connect pricing .			
Region:	US East (Ohio) 		Price
Data Transfer IN To Amazon S3 From Internet			
All data transfer in		\$0.00 per GB	

via -

<https://aws.amazon.com/s3/pricing/>

Also with S3TA, you pay only for transfers that are accelerated. Therefore the research assistant does not need to pay any transfer charges for the image upload because S3TA did not result in an accelerated transfer.

S3 Transfer Acceleration (S3TA)

Overview:

Amazon S3 Transfer Acceleration can speed up content transfers to and from Amazon S3 by as much as 50-500% for long-distance transfer of larger objects. Customers who have either web or mobile applications with widespread users or applications hosted far away from their S3 bucket can experience long and variable upload and download speeds over the Internet. S3 Transfer Acceleration (S3TA) reduces the variability in Internet routing, congestion and speeds that can affect transfers, and logically shortens the distance to S3 for remote applications. **S3TA improves transfer performance by routing traffic through Amazon CloudFront's globally distributed Edge Locations and over AWS backbone networks, and by using network protocol optimizations.** You can turn on S3TA with a few clicks in the S3 console, and test its benefits from your location with [a speed comparison tool](#). **With S3TA, you pay only for transfers that are accelerated.**

Benefits

Move data faster over long distances

S3TA can accelerate long-distance transfers to and from your Amazon S3 buckets. The longer the distance between your client application (mobile, web application, or upload tool) and the target S3 bucket, the more S3TA can help. And if S3TA would not accelerate a transfer, you are not charged.

Reduce network variability

For applications interacting with your S3 buckets through the S3 API from outside of your bucket's region, S3TA helps avoid the variability in Internet routing and congestion. It does this by routing your uploads and downloads over the AWS global network infrastructure, so you get the benefit of our network optimizations.

Shorten the distance to S3

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

Maximize bandwidth utilization

S3TA on average fully utilizes your bandwidth for transfers, and minimizes the effect of distance on throughput. This helps to ensure consistently fast performance to Amazon S3 regardless of your client's location.

via -

<https://aws.amazon.com/s3/transfer-acceleration/>

Incorrect options: **The research assistant only needs to pay S3TA transfer charges for the image upload** - Since S3TA did not result in an accelerated transfer, there are no S3TA transfer charges to be paid.

The research assistant only needs to pay S3 transfer charges for the image upload - There are no S3 data transfer charges when data is transferred in from the internet. So this option is incorrect.

The research assistant needs to pay both S3 transfer charges and S3TA transfer charges for the image upload - There are no S3 data transfer charges when data is transferred in from the internet. Since S3TA did not result in an accelerated transfer, there are no S3TA transfer charges to be paid.

References: <https://aws.amazon.com/s3/transfer-acceleration/>

<https://aws.amazon.com/s3/pricing/>

Question 63: **Correct**

An Internet-of-Things (IoT) company is using Kinesis Data Streams (KDS) to process IoT data from field devices. Multiple consumer applications are using the incoming data streams and the engineers have noticed a performance lag for the data delivery speed between producers and consumers of the data streams.

As a Solutions Architect Professional, which of the following would you recommend to improve the performance for the given use-case?



Swap out Kinesis Data Streams with SQS FIFO queues to support the desired read throughput for the downstream applications



Swap out Kinesis Data Streams with Kinesis Data Firehose to support the desired read throughput for the downstream applications



Swap out Kinesis Data Streams with SQS Standard queues to support the desired read throughput for the downstream applications

-

Use Enhanced Fanout feature of Kinesis Data Streams to support the desired read throughput for the downstream applications

(Correct)

Explanation

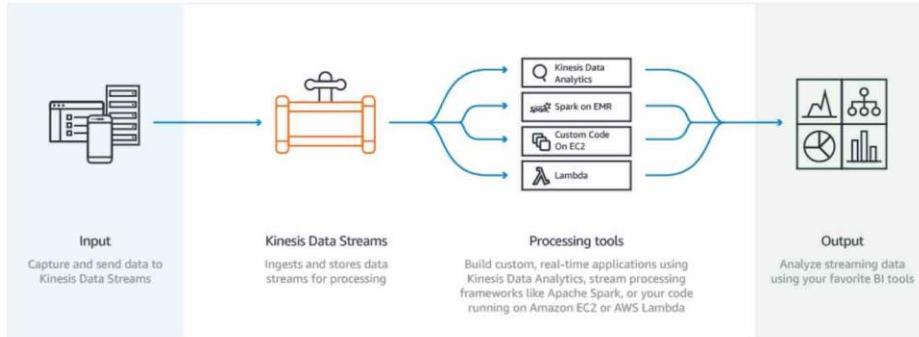
Correct option:

Use Enhanced Fanout feature of Kinesis Data Streams to support the desired read throughput for the downstream applications

Amazon Kinesis Data Streams (KDS) is a massively scalable and durable real-time data streaming service. KDS can continuously capture gigabytes of data per second from hundreds of thousands of sources such as website clickstreams, database event streams, financial transactions, social media feeds, IT logs, and location-tracking events. By default, the 2MB/second/shard output is shared between all of the applications consuming data from the stream.

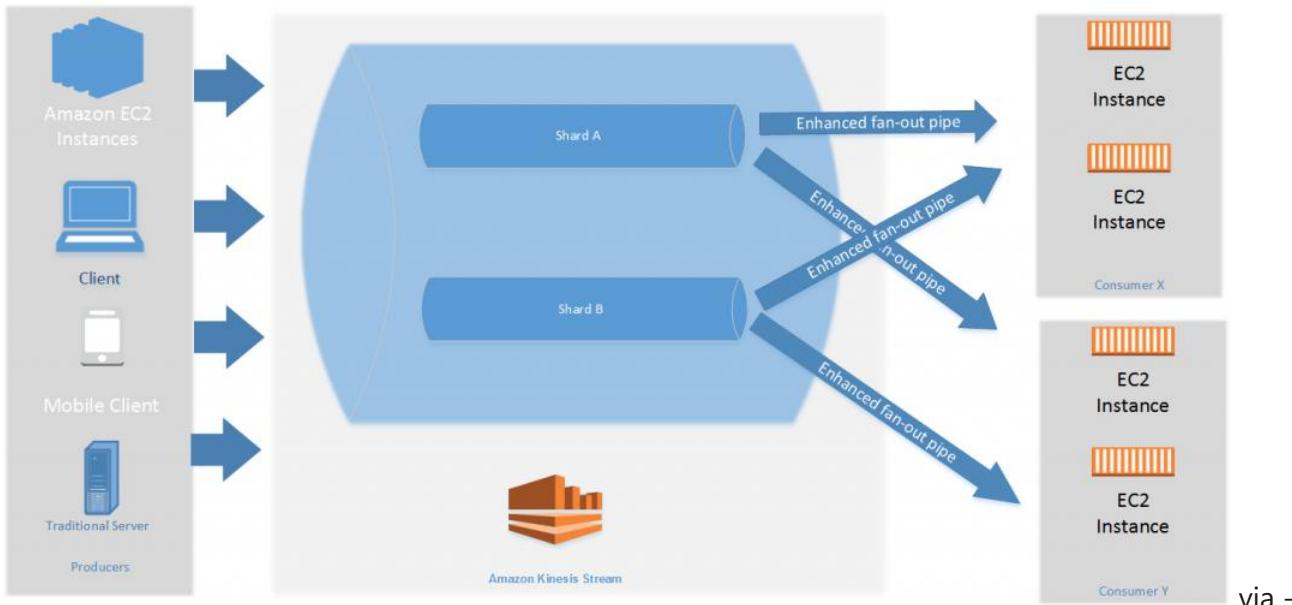
You should use enhanced fan-out if you have multiple consumers retrieving data from a stream in parallel. With enhanced fan-out developers can register stream consumers to use enhanced fan-out and receive their own 2MB/second pipe of read throughput per shard, and this throughput automatically scales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

Kinesis actually refers to a family of streaming services: [Kinesis Video Streams](#), [Kinesis Data Firehose](#), [Kinesis Data Analytics](#), and the topic of today's blog post, Kinesis Data Streams (KDS). Kinesis Data Streams allows developers to easily and continuously collect, process, and analyze streaming data in real-time with a fully-managed and massively scalable service. KDS can capture gigabytes of data per second from hundreds of thousands of sources – everything from website clickstreams and social media feeds to financial transactions and location-tracking events.



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

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



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

Incorrect options:

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

Swap out Kinesis Data Streams with SQS Standard queues to support the desired read throughput for the downstream applications

Swap out Kinesis Data Streams with SQS FIFO queues to support the desired read throughput for the downstream applications

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

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

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

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

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

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

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

References:

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

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

Question 64: **Correct**

A mobile app based social media company is using Amazon CloudFront to deliver media-rich content to its audience across the world. The Content Delivery Network (CDN) offers a multi-tier cache by default, with regional edge caches that improve latency and lower the load on the origin servers when the object is not already cached at the edge. However, there are certain content types that bypass the regional edge cache and go directly to the origin.

Which of the following content types skip the regional edge cache? (Select two)

-

E-commerce assets such as product photos

-

Dynamic content, as determined at request time (cache-behavior configured to forward all headers)

(Correct)

-

Proxy methods PUT/POST/PATCH/OPTIONS/DELETE go directly to the origin

(Correct)

-

Static content such as style sheets, JavaScript files

-

User-generated videos

Explanation

Correct options:

Dynamic content, as determined at request time (cache-behavior configured to forward all headers)

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

Dynamic content, as determined at request time (cache-behavior configured to forward all headers), does not flow through regional edge caches, but goes directly to the origin. So this option is correct.

Proxy methods PUT/POST/PATCH/OPTIONS/DELETE go directly to the origin

Proxy methods PUT/POST/PATCH/OPTIONS/DELETE go directly to the origin from the POPs and do not proxy through the regional edge caches. So this option is also correct.

How CloudFront works with regional edge

How CloudFront works with regional edge caches

CloudFront points of presence (POPs) (edge locations) make sure that popular content can be served quickly to your viewers. CloudFront also has *regional edge caches* that bring more of your content closer to your viewers, even when the content is not popular enough to stay at a POP, to help improve performance for that content.

Regional edge caches help with all types of content, particularly content that tends to become less popular over time. Examples include user-generated content, such as video, photos, or artwork; e-commerce assets such as product photos and videos; and news and event-related content that might suddenly find new popularity.

How regional caches work

Regional edge caches are CloudFront locations that are deployed globally, close to your viewers. They're located between your origin server and the POPs—global edge locations that serve content directly to viewers. As objects become less popular, individual POPs might remove those objects to make room for more popular content. Regional edge caches have a larger cache than an individual POP, so objects remain in the cache longer at the nearest regional edge cache location. This helps keep more of your content closer to your viewers, reducing the need for CloudFront to go back to your origin server, and improving overall performance for viewers.

When a viewer makes a request on your website or through your application, DNS routes the request to the POP that can best serve the user's request. This location is typically the nearest CloudFront edge location in terms of latency. In the POP, CloudFront checks its cache for the requested files. If the files are in the cache, CloudFront returns them to the user. If the files are not in the cache, the POPs go to the nearest regional edge cache to fetch the object.

In the regional edge cache location, CloudFront again checks its cache for the requested files. If the files are in the cache, CloudFront forwards the files to the POP that requested them. As soon as the first byte arrives from regional edge cache location, CloudFront begins to forward the files to the user. CloudFront also adds the files to the cache in the POP for the next time someone requests those files.

For files not cached at either the POP or the regional edge cache location, CloudFront compares the request with the specifications in your distributions and forwards the request for your files to the origin server. After your origin server sends the files back to the regional edge cache location, they are forwarded to the POP, and CloudFront forwards the files to the user. In this case, CloudFront also adds the files to the cache in the regional edge cache location in addition to the POP for the next time a viewer requests those files. This makes sure that all of the POPs in a region share a local cache, eliminating multiple requests to origin servers. CloudFront also keeps persistent connections with origin servers so files are fetched from the origins as quickly as possible.

Note

- Regional edge caches have feature parity with POPs. For example, a cache invalidation request removes an object from both POP caches and regional edge caches before it expires. The next time a viewer requests the object, CloudFront returns to the origin to fetch the latest version of the object.
- Proxy methods PUT/POST/PATCH/OPTIONS/DELETE go directly to the origin from the POPs and do not proxy through the regional edge caches.
- Dynamic requests, as determined at request time, do not flow through regional edge caches, but go directly to the origin.

caches:

via -

<https://docs.aws.amazon.com/AmazonCloudFront/latest/DeveloperGuide/HowCloudFrontWorks.html>

Incorrect Options:

E-commerce assets such as product photos

User-generated videos

Static content such as style sheets, JavaScript files

The following type of content flows through the regional edge caches - user-generated content, such as video, photos, or artwork; e-commerce assets such as product photos and videos and static content such as style sheets, JavaScript files. Hence these three options are not correct.

Reference: <https://docs.aws.amazon.com/AmazonCloudFront/latest/DeveloperGuide/HowCloudFrontWorks.html>

Question 65: **Correct**

A retail company has hired you as an AWS Certified Solutions Architect Professional to provide consultancy for managing a serverless application that consists of multiple API gateways, Lambda functions, S3 buckets and DynamoDB tables. The company is getting reports from customers that some of the application components seem to be lagging while loading dynamic images and some are timing out with the "504 Gateway Timeout" error. As part of your investigations to identify the root cause behind this issue, you can confirm that DynamoDB monitoring metrics are at acceptable levels.

Which of the following steps would you recommend to address these application issues? (Select two)

-

Process and analyze the VPC Flow Logs to determine if there is packet loss between the Lambda function and S3

-

Process and analyze the Amazon CloudWatch Logs for Lambda function to determine processing times for requested images at pre-configured intervals

(Correct)

-

Enable execution logging for the API Gateway. Process and analyze the execution logs in the API Gateway for HTTP errors to determine the root cause of the errors

-

Process and analyze the AWS X-Ray traces and analyze HTTP methods to determine the root cause of the HTTP errors

(Correct)

-

Enable access logging for the API Gateway. Process and analyze the access logs in the API Gateway for HTTP errors to determine the root cause of the errors

Explanation

Correct options:

Process and analyze the Amazon CloudWatch Logs for Lambda function to determine processing times for requested images at pre-configured intervals

To help you troubleshoot failures in a function, the Lambda service logs all requests handled by a Lambda function and also automatically stores logs generated by your code through Amazon CloudWatch Logs. You can insert logging statements into your code to determine processing times for requested images. These logs can then be processed at certain pre-configured intervals for further analysis.

Accessing Amazon CloudWatch logs for AWS Lambda

[PDF](#) | [Kindle](#) | [RSS](#)

AWS Lambda automatically monitors Lambda functions on your behalf, reporting metrics through Amazon CloudWatch. To help you troubleshoot failures in a function, Lambda logs all requests handled by your function and also automatically stores logs generated by your code through Amazon CloudWatch Logs.

You can insert logging statements into your code to help you validate that your code is working as expected. Lambda automatically integrates with CloudWatch Logs and pushes all logs from your code to a CloudWatch Logs group associated with a Lambda function, which is named /aws/lambda/<function name>. To learn more about log groups and accessing them through the CloudWatch console, see the [Monitoring system, application, and custom log files](#) in the *Amazon CloudWatch User Guide*.

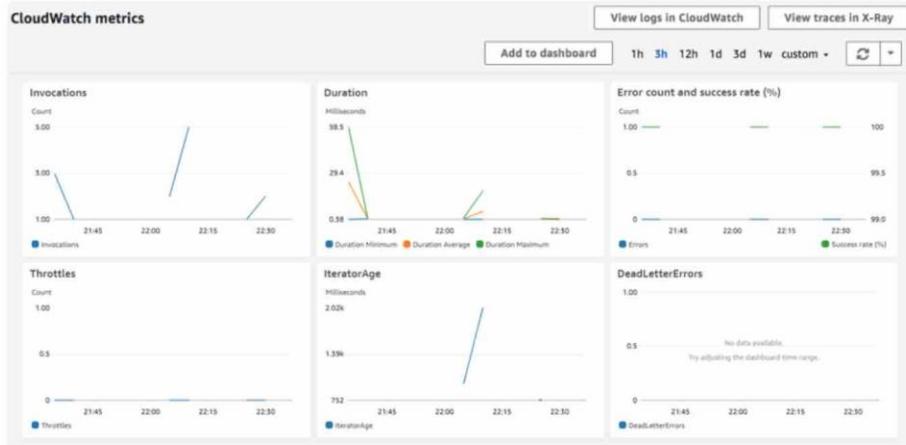
You can view logs for Lambda by using the Lambda console, the CloudWatch console, the AWS CLI, or the CloudWatch API. The following procedure show you how to view the logs by using the Lambda console.

Note

There is no additional charge for using Lambda logs; however, standard CloudWatch Logs charges apply. For more information, see [CloudWatch pricing](#).

To view logs using the Lambda console

1. Open the Lambda console [Functions page](#).
2. Choose a function.
3. Choose **Monitoring**.



via -

<https://docs.aws.amazon.com/lambda/latest/dg/monitoring-cloudwatchlogs.html>

Process and analyze the AWS X-Ray traces and analyze HTTP methods to determine the root cause of the HTTP errors

You can use AWS X-Ray to visualize the components of your application, identify performance bottlenecks such as the one described in the use-case for processing images and troubleshoot those requests that resulted in an error. Your Lambda functions send trace data to X-Ray, and X-Ray processes the data to generate a service map and searchable trace summaries.

Using AWS Lambda with AWS X-Ray

[PDF](#) | [Kindle](#) | [RSS](#)

You can use AWS X-Ray to visualize the components of your application, identify performance bottlenecks, and troubleshoot requests that resulted in an error. Your Lambda functions send trace data to X-Ray, and X-Ray processes the data to generate a service map and searchable trace summaries.



If you've enabled X-Ray tracing in a service that invokes your function, Lambda sends traces to X-Ray automatically. The upstream service, such as Amazon API Gateway, or an application hosted on Amazon EC2 that is instrumented with the X-Ray SDK, samples incoming requests and adds a tracing header that tells Lambda to send traces or not.

To trace requests that don't have a tracing header, enable active tracing in your function's configuration.

To enable active tracing

1. Open the Lambda console [Functions page](#).
2. Choose a function.
3. Under **AWS X-Ray**, choose **Active tracing**.

<https://docs.aws.amazon.com/lambda/latest/dg/services-xray.html>

Incorrect options:

Enable execution logging for the API Gateway. Process and analyze the execution logs in the API Gateway for HTTP errors to determine the root cause of the errors

Enable access logging for the API Gateway. Process and analyze the access logs in the API Gateway for HTTP errors to determine the root cause of the errors

For an API Gateway, a "504 Gateway Timeout" error implies an "Endpoint Request Timed-out Exception".

Error Codes (Client and Server Errors)

HTTP status codes indicate whether an operation is successful or not.

A response code of 2xx indicates the operation was successful. Other error codes indicate either a client error (4xx) or a server error (5xx).

The following table lists the errors returned by Amazon API Gateway. Some errors are resolved if you simply retry the same request. The table indicates which errors are likely to be resolved with successive retries. If the value of the Retry column is:

- **Yes:** Submit the same request again.
- **Yes if idempotent:** Submit the same request again, but only if the given method has been implemented with idempotence.
- **No:** Fix the problem on the client side before submitting a new request.

For more information about retrying requests, see Error Retries and Exponential Backoff.

HTTP Status Code	Error code	Retry
400	Bad Request Exception	No
403	Access Denied Exception	No
404	Not Found Exception	No
409	Conflict Exception	No
429	Limit Exceeded Exception	No
429	Too Many Requests Exception	Yes
502	Bad Gateway Exception, usually for an incompatible output returned from a Lambda proxy integration backend and occasionally for out-of-order invocations due to heavy loads.	Yes if idempotent
503	Service Unavailable Exception	Yes
504	Endpoint Request Timed-out Exception	Yes if idempotent

via -

<https://docs.aws.amazon.com/apigateway/api-reference/handling-errors/>

To troubleshoot an API Gateway REST API or WebSocket API that you're developing, enable execution logging and access logging to Amazon CloudWatch Logs. Execution logs contain helpful information that you can use to identify and fix most errors with your APIs. Access logs contain details about who accessed your API and how they accessed it, which you can also use for troubleshooting.

To troubleshoot an API Gateway REST API or WebSocket API that you're developing, [enable execution logging and access logging to Amazon CloudWatch Logs](#).

Note: [HTTP APIs currently support access logging only](#), and logging setup is different for these APIs. For more information, see [Configuring Logging for an HTTP API](#).

Execution logs contain helpful information that you can use to identify and fix most errors with your APIs. This information includes:

- The requests that your API receives.
- Your API's [integration backend responses](#).
- The response provided by [Lambda authorizers](#).
- The [requestId](#) for AWS integration endpoints.
- Whether a provided [API key](#) was authorized.

[Access logs contain details about who accessed your API and how they accessed it](#), which you can also use for troubleshooting. For more information about each type of logging, see [CloudWatch Log Formats for API Gateway](#).

via -

<https://aws.amazon.com/premiumsupport/knowledge-center/api-gateway-cloudwatch-logs/>

However, neither execution logs nor access logs at the API Gateway level will provide information to identify the root cause for the "504 Gateway Timeout" error as it needs to be analyzed at the source system level which is Lambda function for the given use-case, as that's where the images are being processing and the application is lagging or timing out for some of those images. Another thing to note is that only access logs are available for HTTP APIs, so you do not have access to execution logs for the given use-case.

Therefore, both of these options are incorrect.

Process and analyze the VPC Flow Logs to determine if there is packet loss between the Lambda function and S3

VPC Flow Logs allow you to capture information about the IP traffic going to and from network interfaces in your VPC. You can create a flow log for a VPC, a subnet, or a network interface. If you create a flow log for a subnet or VPC, each network interface in that subnet or VPC is monitored.

Flow logs basics

You can create a flow log for a VPC, a subnet, or a network interface. If you create a flow log for a subnet or VPC, each network interface in that subnet or VPC is monitored.

Flow log data for a monitored network interface is recorded as *flow log records*, which are log events consisting of fields that describe the traffic flow. For more information, see [Flow log records](#).

To create a flow log, you specify:

- The resource for which to create the flow log
- The type of traffic to capture (accepted traffic, rejected traffic, or all traffic)
- The destinations to which you want to publish the flow log data

In the following example, you create a flow log (fl-aaa) that captures accepted traffic for the network interface for instance A1 and publishes the flow log records to an Amazon S3 bucket. You create a second flow log that captures all traffic for subnet B and publishes the flow log records to Amazon CloudWatch Logs. The flow log (fl-bbb) captures traffic for all network interfaces in subnet B. There are no flow logs that capture traffic for instance A2's network interface.



via -

<https://docs.aws.amazon.com/vpc/latest/userguide/flow-logs.html>

This option has been added as a distractor as you cannot use VPC Flow Logs to determine packet loss. AWS has a built-in tool called AWS Support-SetupIPMonitoringFromVPC that you can use to monitor metrics such as latency and the percentage of packet loss across a network path. It monitors the selected target IP addresses by continuously running ping, MTR, TCP traceroute, and tracepath network diagnostic tests.

References:

<https://docs.aws.amazon.com/lambda/latest/dg/monitoring-cloudwatchlogs.html>

<https://docs.aws.amazon.com/lambda/latest/dg/services-xray.html>

<https://docs.aws.amazon.com/apigateway/api-reference/handling-errors/>

<https://aws.amazon.com/premiumsupport/knowledge-center/api-gateway-cloudwatch-logs/>

<https://aws.amazon.com/blogs/networking-and-content-delivery/debugging-tool-for-network-connectivity-from-amazon-vpc/>

<https://docs.aws.amazon.com/vpc/latest/userguide/flow-logs.html>

Question 66: **Incorrect**

An analytics company wants to leverage ElastiCache for Redis in cluster mode to enhance the performance and scalability of its existing two-tier application architecture. The ElastiCache cluster is configured to listen on port 6379. The company has hired you as an AWS Certified Solutions Architect Professional to build a secure solution so that the cache data is secure and protected from unauthorized access.

Which of the following steps would address the given use-case? (Select three)

-

Create the cluster with auth-token parameter and make sure that the parameter is included in all subsequent commands to the cluster

(Correct)

-

Configure the ElastiCache cluster to have both in-transit as well as at-rest encryption

(Correct)

-

Configure the security group for the ElastiCache cluster with the required rules to allow inbound traffic from the cluster itself as well as from the cluster's clients on port 6379

(Correct)

-

Enable CloudWatch Logs to monitor the security credentials for the ElastiCache cluster

-

Configure the security group for the ElastiCache cluster with the required rules to allow outbound traffic to the cluster's clients on port 6379

-

Enable CloudTrail to monitor the API Calls for the ElastiCache cluster

(Incorrect)

Explanation

Correct options:

Configure the ElastiCache cluster to have both in-transit as well as at-rest encryption

You can use both in-transit as well as at-rest encryption to guard against unauthorized access of your data on the server. In-transit encryption encrypts your data whenever it is moving from one place to another, such as between nodes in your cluster or between your cluster and your application. At-rest encryption encrypts your on-disk data during sync and backup operations.

Data Security in Amazon ElastiCache

[PDF](#) | [Kindle](#) | [RSS](#)

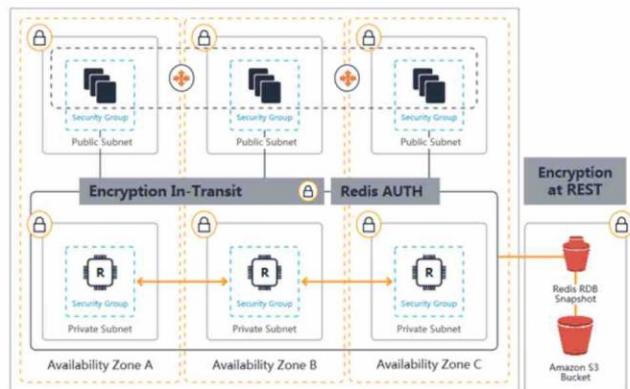
To help keep your data secure, Amazon ElastiCache and Amazon EC2 provide mechanisms to guard against unauthorized access of your data on the server.

Amazon ElastiCache for Redis also provides optional encryption features for data on clusters running Redis versions 3.2.6, 4.0.10 or later:

- In-transit encryption encrypts your data whenever it is moving from one place to another, such as between nodes in your cluster or between your cluster and your application.
- At-rest encryption encrypts your on-disk data during sync and backup operations.

If you want to enable in-transit or at-rest encryption, you must meet the following conditions.

- Your cluster or replication group must be running Redis 3.2.6, 4.0.10 or later.
- Your cluster or replication group must be created in a VPC based on Amazon VPC.
- Optionally, you can also use AUTH and the AUTH token (password) needed to perform operations on this cluster or replication group.



via -

<https://docs.aws.amazon.com/AmazonElastiCache/latest/red-ug/encryption.html>

Create the cluster with auth-token parameter and make sure that the parameter is included in all subsequent commands to the cluster

Redis authentication tokens enable Redis to require a token (password) before allowing clients to run commands, thereby improving data security. You can require that users enter a token on a token-protected Redis server. You also need to include it in all subsequent commands to the replication group or cluster.

Overview of AUTH in ElastiCache for Redis

When you use Redis **AUTH** with your ElastiCache for Redis cluster, there are some refinements.

In particular, be aware of these **AUTH** token constraints when using **AUTH** with ElastiCache for Redis:

- Tokens must be 16–128 printable characters.
- Nonalphanumeric characters are restricted to (!, &, #, \$, ^, <, >, -).
- AUTH can only be enabled for encryption in-transit enabled ElastiCache for Redis clusters.

To set up a strong token, we recommend that you follow a strict token policy, such as requiring the following:

- Tokens must include at least three of the following character types:
 - Uppercase characters
 - Lowercase characters
 - Digits
 - Nonalphanumeric characters (!, &, #, \$, ^, <, >, -)
- Tokens must not contain a dictionary word or a slightly modified dictionary word.
- Tokens must not be the same as or similar to a recently used token.

Applying Authentication to an ElastiCache for Redis Cluster

You can require that users enter a token on a token-protected Redis server. To do this, include the parameter `--auth-token` (API: `AuthToken`) with the correct token when you create your replication group or cluster. Also include it in all subsequent commands to the replication group or cluster.

The following AWS CLI operation creates a replication group with encryption in transit (TLS) enabled and the **AUTH** token *This-is-a-sample-token*. Replace the subnet group `sng-test` with a subnet group that exists.

Key Parameters

- `--engine` – Must be `redis`.
- `--engine-version` – Must be 3.2.6, 4.0.10, or later.
- `--transit-encryption-enabled` – Required for authentication and HIPAA eligibility.
- `--auth-token` – Required for HIPAA eligibility. This value must be the correct token for this token-protected Redis server.
- `--cache-subnet-group` – Required for HIPAA eligibility.

For Linux, macOS, or Unix:

```
aws elasticache create-replication-group \
--replication-group-id authtestgroup \
--replication-group-description authtest \
--engine redis \
--engine-version 4.0.10 \
--cache-node-type cache.m4.large \
--num-node-groups 1 \
--replicas-per-node-group 2 \
--cache-parameter-group default.redis3.2.cluster.on \
--transit-encryption-enabled \
--auth-token This-is-a-sample-token \
--cache-subnet-group sng-test
```

via -

<https://docs.aws.amazon.com/AmazonElastiCache/latest/red-ug/auth.html>

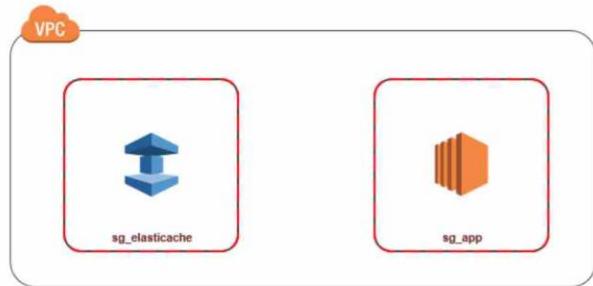
Configure the security group for the ElastiCache cluster with the required rules to allow inbound traffic from the cluster itself as well as from the cluster's clients on port 6379

You can create a VPC security group to restrict access to the cluster instances. Configure rules that only allow inbound traffic from the cluster itself as well as from the cluster's clients on port 6379. Typically the ElastiCache cluster is accessed from the web servers running on EC2 instances. You can configure the security groups like so:

Accessing an ElastiCache Cluster when it and the Amazon EC2 Instance are in the Same Amazon VPC

The most common use case is when an application deployed on an EC2 instance needs to connect to a Cluster in the same VPC.

The following diagram illustrates this scenario



The simplest way to manage access between EC2 instances and DB instances in the same VPC is to do the following:

1. Create a VPC security group for your cluster. This security group can be used to restrict access to the cluster instances. For example, you can create a custom rule for this security group that allows TCP access using the port you assigned to the cluster when you created it and an IP address you will use to access the cluster.
The default port for Redis clusters and replication groups is 6379.
2. Create a VPC security group for your EC2 instances (web and application servers). This security group can, if needed, allow access to the EC2 instance from the Internet via the VPC's routing table. For example, you can set rules on this security group to allow TCP access to the EC2 instance over port 22.
3. Create custom rules in the security group for your Cluster that allow connections from the security group you created for your EC2 instances. This would allow any member of the security group to access the DB instances.

via -

<https://docs.aws.amazon.com/AmazonElastiCache/latest/red-ug/elasticsearch-vpc-accessing.html>

Incorrect options:

Configure the security group for the ElastiCache cluster with the required rules to allow outbound traffic to the cluster's clients on port 6379 - As mentioned in the explanation above, you need to create a security group that allows inbound traffic from the cluster itself as well as from the cluster's clients on port 6379. Creating a security group rule that allows outbound traffic from the cluster on port 6379 is not relevant to the use-case.

Enable CloudWatch Logs to monitor the security credentials for the ElastiCache cluster

Enable CloudTrail to monitor the API Calls for the ElastiCache cluster

Both these options are added as distractors since both CloudWatch Logs and CloudTrail can be used for post-facto analysis to ascertain the series of access events relevant to the cluster. These options will not prevent unauthorized access.

References:

<https://docs.aws.amazon.com/AmazonElastiCache/latest/red-ug/encryption.html>

<https://docs.aws.amazon.com/AmazonElastiCache/latest/red-ug/auth.html>

<https://docs.aws.amazon.com/AmazonElastiCache/latest/red-ug/elasticache-vpc-accessing.html>

Question 67: **Correct**

A web hosting company's CFO recently analyzed the company's monthly bill for the AWS account for the development environment and identified an opportunity to reduce the cost for AWS Elastic Beanstalk infrastructure in use. The CFO in consultation with the CTO has hired you as an AWS Certified Solutions Architect Professional to design a highly available solution that will provision an Elastic Beanstalk environment in the morning and terminate it at the end of the day. The solution should be designed with minimal operational overhead with a focus on minimizing costs. The solution should also facilitate the increased use of Elastic Beanstalk environments among different development teams and must provide a one-stop scheduler solution for all teams to keep the operational costs as low as possible.

Which of the following solution designs will you suggest to address these requirements?

- Set up separate Lambda functions to provision and terminate the Elastic Beanstalk environment. Configure a Lambda execution role granting the required Elastic Beanstalk environment permissions and assign the role to the Lambda functions. Configure cron expression based Amazon EventBridge events rules to trigger the Lambda functions

(Correct)

- Leverage the activity task of an AWS Step Function to provision and terminate the Elastic Beanstalk environment. Create a role for the Step Function to allow it to provision and terminate the Elastic Beanstalk environment. Execute the Step Function daily and use the "wait state" to control the start and stop time
- Configure the Elastic Beanstalk environment to use custom commands in the EC2 instance user data. Leverage the scheduled action for an Auto Scaling group to scale-out EC2 instances in the morning and scale-in the instance count to 0 to terminate the EC2 instances at the end of the day
- Provision an EC2 Micro instance. Configure an IAM role with the required Elastic Beanstalk environment permissions and attach it to the instance profile. Create scripts on the instance to provision and terminate the Elastic Beanstalk environment. Set up cron jobs on the instance to execute the scripts

Explanation

Correct option:

Set up separate Lambda functions to provision and terminate the Elastic Beanstalk environment. Configure a Lambda execution role granting the required Elastic Beanstalk environment permissions and assign the role to the Lambda functions. Configure cron expression based Amazon EventBridge events rules to trigger the Lambda functions

You can configure Lambda functions to make the Elastic Beanstalk API calls for provisioning and terminating the Elastic Beanstalk environment. To perform these API calls on a schedule, you can configure events in Amazon EventBridge events to trigger these Lambda functions at a specific time each day via cron expressions. Make sure that you attach the policy with the required Elastic Beanstalk environment permissions to the Lambda execution role.

How do I stop and restart my Elastic Beanstalk environment on a schedule?

Last updated: 2019-11-8

How can I terminate and rebuild my test or non-critical AWS Elastic Beanstalk environment at a scheduled time?

Short Description

You can stop and restart your Elastic Beanstalk environment with the API calls `terminate-environment` and `rebuild-environment`. You can only rebuild terminated environments within six weeks (42 days) of their termination.

To perform these calls on a schedule, configure events in Amazon CloudWatch Events to trigger AWS Lambda functions at a specific time each day. Then, configure those Lambda functions to make the Elastic Beanstalk API calls.

Important: Any out-of-band changes that you make to an Elastic Beanstalk environment or its instances don't persist after the environment is terminated. Be sure to consider this factor when changing your environment. Also, note the termination time, and complete any work that's using the instance before that time. The instance terminates at the scheduled time even if a user isn't connected to that instance.

via -

<https://aws.amazon.com/premiumsupport/knowledge-center/schedule-elastic-beanstalk-stop-restart/>

Cron Expressions

Cron expressions have six required fields, which are separated by white space.

Syntax

cron(<i>fields</i>)		
Field	Values	Wildcards
Minutes	0-59	, - * /
Hours	0-23	, - * /
Day-of-month	1-31	, - * ? / L W
Month	1-12 or JAN-DEC	, - * /
Day-of-week	1-7 or SUN-SAT	, - * ? L #
Year	1970-2199	, - * /

Wildcards

- The , (comma) wildcard includes additional values. In the Month field, JAN,FEB,MAR would include January, February, and March.
- The - (dash) wildcard specifies ranges. In the Day field, 1-15 would include days 1 through 15 of the specified month.
- The * (asterisk) wildcard includes all values in the field. In the Hours field, * would include every hour. You cannot use * in both the Day-of-month and Day-of-week fields. If you use it in one, you must use ? in the other.
- The / (forward slash) wildcard specifies increments. In the Minutes field, you could enter 1/10 to specify every tenth minute, starting from the first minute of the hour (for example, the 11th, 21st, and 31st minute, and so on).
- The ? (question mark) wildcard specifies one or another. In the Day-of-month field you could enter ? and if you didn't care what day of the week the 7th was, you could enter ? in the Day-of-week field.
- The L wildcard in the Day-of-month or Day-of-week fields specifies the last day of the month or week.
- The W wildcard in the Day-of-month field specifies a weekday. In the Day-of-month field, 3W specifies the weekday closest to the third day of the month.
- The # wildcard in the Day-of-week field specifies a certain instance of the specified day of the week within a month. For example, 3#2 would be the second Tuesday of the month: the 3 refers to Tuesday because it is the third day of each week, and the 2 refers to the second day of that type within the month.

Restrictions

- You can't specify the Day-of-month and Day-of-week fields in the same cron expression. If you specify a value (or a *) in one of the fields, you must use a ? (question mark) in the other.
- Cron expressions that lead to rates faster than 1 minute are not supported.

Examples

You can use the following sample cron strings when creating a rule with schedule.

Minutes	Hours	Day of month	Month	Day of week	Year	Meaning
0	10	*	*	?	*	Run at 10:00 am (UTC) every day
15	12	*	*	?	*	Run at 12:15 pm (UTC) every day
0	18	?	*	MON-FRI	*	Run at 6:00 pm (UTC) every Monday through Friday
0	8	1	*	?	*	Run at 8:00 am (UTC) every 1st day of the month
0/15	*	*	*	?	*	Run every 15 minutes
0/10	*	?	*	MON-FRI	*	Run every 10 minutes Monday through Friday
0/5	8-17	?	*	MON-FRI	*	Run every 5 minutes Monday through Friday between 8:00 am and 5:55 pm (UTC)

The following examples show how to use Cron expressions with the AWS CLI put-rule command. The first example creates a rule that is triggered every day at 12:00pm UTC.

```
aws events put-rule --schedule-expression "cron(@ 12 * * ? *)" --name MyRule1
```

The next example creates a rule that is triggered every day, at 5 and 35 minutes past 2:00pm UTC.

```
aws events put-rule --schedule-expression "cron(5,35 14 * * ? *)" --name MyRule2
```

The next example creates a rule that is triggered at 10:15am UTC on the last Friday of each month during the years 2002 to 2005.

```
aws events put-rule --schedule-expression "cron(15 10 ? * 6L 2002-2005)" --name MyRule3
```

via -

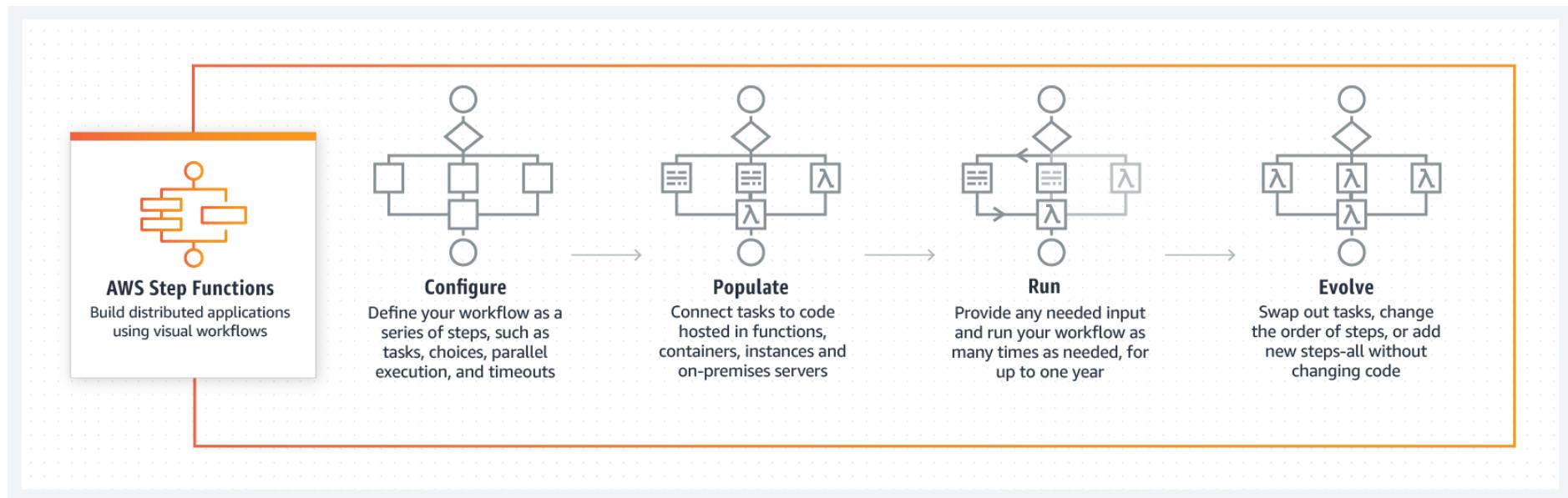
<https://docs.aws.amazon.com/AmazonCloudWatch/latest/events/ScheduledEvents.html>

Please review this excellent reference material for a deep-dive on how to stop and restart an Elastic Beanstalk environment on a schedule:

<https://aws.amazon.com/premiumsupport/knowledge-center/schedule-elastic-beanstalk-stop-restart/>

Incorrect options:

Leverage the activity task of an AWS Step Function to provision and terminate the Elastic Beanstalk environment. Create a role for the Step Function to allow it to provision and terminate the Elastic Beanstalk environment. Execute the Step Function daily and use the "wait state" to control the start and stop time - AWS Step Functions lets you coordinate multiple AWS services into serverless workflows so you can build and update apps quickly. Using Step Functions, you can design and run workflows that stitch together services such as AWS Lambda into feature-rich applications.



via - <https://aws.amazon.com/step-functions/>

This option involves cost components that are not needed. There is no need to keep the Step Function running in a "wait state" for the most part of the day just to control the provisioning and termination of Elastic Beanstalk environment. This is better handled via a "serverless cron" type of solution that can be invoked twice a day for provisioning and termination of Elastic Beanstalk environment.

Configure the Elastic Beanstalk environment to use custom commands in the EC2 instance user data. Leverage the scheduled action for an Auto Scaling group to scale-out EC2 instances in the morning and scale-in the instance count to 0 to terminate the EC2 instances at the end of the day - Scheduled action allows you to set your own scaling schedule for an Auto Scaling group. For example, let's say that every week the traffic to your web application starts to increase on Wednesday, remains high on Thursday, and starts to decrease on Friday. You can plan your scaling actions based on the predictable traffic patterns of your web application. Scaling actions are performed automatically as a function of time and date. A scheduled action sets the minimum, maximum, and desired sizes to what is specified by the scheduled action at the time specified by the scheduled action.

The issue with this option is that it's costly. Firstly, there is a cost element of running the EC2 instances for the day hours. Secondly, Elastic Beanstalk environment is provisioned via custom commands in the EC2 instance user data (it should also be emphasized that EC2 instance user data is not the best place to trigger the creation of an Elastic Beanstalk environment) however the environment itself is not terminated at the end of the day. So the costs for the resources created by Elastic Beanstalk keep accumulating. Hence this option is incorrect.

Provision an EC2 Micro instance. Configure an IAM role with the required Elastic Beanstalk environment permissions and attach it to the instance profile. Create scripts on the instance to provision and terminate the Elastic Beanstalk environment. Set up cron jobs on the instance to execute the scripts - This option involves cost components that are not needed. There is no need to provision an EC2 Micro instance and keep it running just to control the provisioning and termination of Elastic Beanstalk environment. This is better handled via a "serverless cron" type of solution that can be invoked twice a day for provisioning and termination of Elastic Beanstalk environment.

References:

<https://aws.amazon.com/premiumsupport/knowledge-center/schedule-elastic-beanstalk-stop-restart/>

<https://docs.aws.amazon.com/AmazonCloudWatch/latest/events/ScheduledEvents.html>

<https://aws.amazon.com/step-functions/>

Question 68: **Incorrect**

A digital media company wants to use AWS Cloudfront to manage its content. Firstly, it would like to allow only those new users who have paid the annual subscription fee the ability to download the application installation file. Secondly, only the subscribers should be able to view the files in the members' area.

As a Solutions Architect Professional, which of the following would you recommend as the MOST optimal solutions to deliver restricted content to the bona fide end users? (Select two)

-

Use CloudFront signed cookies to restrict access to all the files in the members' area of the website

(Correct)

-

Use CloudFront signed URLs to restrict access to the application installation file

(Correct)

-

Use CloudFront signed URLs to restrict access to all the files in the members' area of the website

-

Require HTTPS for communication between CloudFront and your S3 origin

-

Use CloudFront signed cookies to restrict access to the application installation file

(Incorrect)

Explanation

Correct options:

Use CloudFront signed URLs to restrict access to the application installation file

Use CloudFront signed cookies to restrict access to all the files in the members' area of the website

Many companies that distribute content over the internet want to restrict access to documents, business data, media streams, or content that is intended for selected users, for example, users who have paid a fee.

To securely serve this private content by using CloudFront, you can do the following:

Require that your users access your private content by using special CloudFront signed URLs or signed cookies.

You should use a signed URL if you want to restrict access to individual files, for example, an installation download for your application. A signed URL includes additional information, for example, expiration date and time, that gives you more control over access to your content.

On the other hand, CloudFront signed cookies allow you to control who can access your content when you don't want to change your current URLs or when you want to provide access to multiple restricted files, for example, all of the files in the members' area of a website.

Choosing Between Signed URLs and Signed Cookies

[PDF](#) | [Kindle](#) | [RSS](#)

CloudFront signed URLs and signed cookies provide the same basic functionality: they allow you to control who can access your content. If you want to serve private content through CloudFront and you're trying to decide whether to use signed URLs or signed cookies, consider the following.

Use signed URLs in the following cases:

- You want to use an RTMP distribution. Signed cookies aren't supported for RTMP distributions.
- You want to restrict access to individual files, for example, an installation download for your application.
- Your users are using a client (for example, a custom HTTP client) that doesn't support cookies.

Use signed cookies in the following cases:

- You want to provide access to multiple restricted files, for example, all of the files for a video in HLS format or all of the files in the subscribers' area of website.
- You don't want to change your current URLs.

If you are not currently using signed URLs, and if your (unsigned) URLs contain any of the following query string parameters, you cannot use either signed URLs or signed cookies:

- Expires
- Policy
- Signature
- Key-Pair-Id

CloudFront assumes that URLs that contain any of those query string parameters are signed URLs, and therefore won't look at signed cookies.

via -

<https://docs.aws.amazon.com/AmazonCloudFront/latest/DeveloperGuide/private-content-choosing-signed-urls-cookies.html>

Incorrect options:

Use CloudFront signed cookies to restrict access to the application installation file

Use CloudFront signed URLs to restrict access to all the files in the members' area of the website

These two options contradict the description provided in the explanation above, so these options are incorrect.

Require HTTPS for communication between CloudFront and your S3 origin

Requiring HTTPS for communication between CloudFront and your custom origin (or S3 origin) only enables secure access to the underlying content. You cannot use HTTPS to restrict access to your private content. So this option is incorrect.

References:

<https://docs.aws.amazon.com/AmazonCloudFront/latest/DeveloperGuide/private-content-choosing-signed-urls-cookies.html>

<https://docs.aws.amazon.com/AmazonCloudFront/latest/DeveloperGuide/private-content-signed-urls.html>

<https://docs.aws.amazon.com/AmazonCloudFront/latest/DeveloperGuide/private-content-signed-cookies.html>

Question 69: **Incorrect**

An automobile company helps more than 20 million web and mobile users browse automobile dealer inventory, read vehicle reviews, and consume other automobile-related content by leveraging its library of 50 million vehicle photos uploaded by auto dealers. The company is planning a key update with even better image quality and faster load times on the company's website as well as mobile apps but the existing image-handling solution based on Cloudera MapReduce clusters is not the right tool for the job. The company now wants to switch to a serverless solution on AWS Cloud. As part of this process, the engineering team has been studying various best practices for serverless solutions. They intend to use AWS Lambda extensively and are looking at the salient features to consider when using Lambda as the backbone for the serverless architecture.

As a Solutions Architect Professional, which of the following would you identify as key considerations for a serverless architecture? (Select three)

-

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

-

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

(Correct)

-

Serverless databases and Lambda complement each other and you should install databases on the Lambda functions

-

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

(Incorrect)

-

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

(Correct)

-

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

(Correct)

Explanation

Correct options:

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

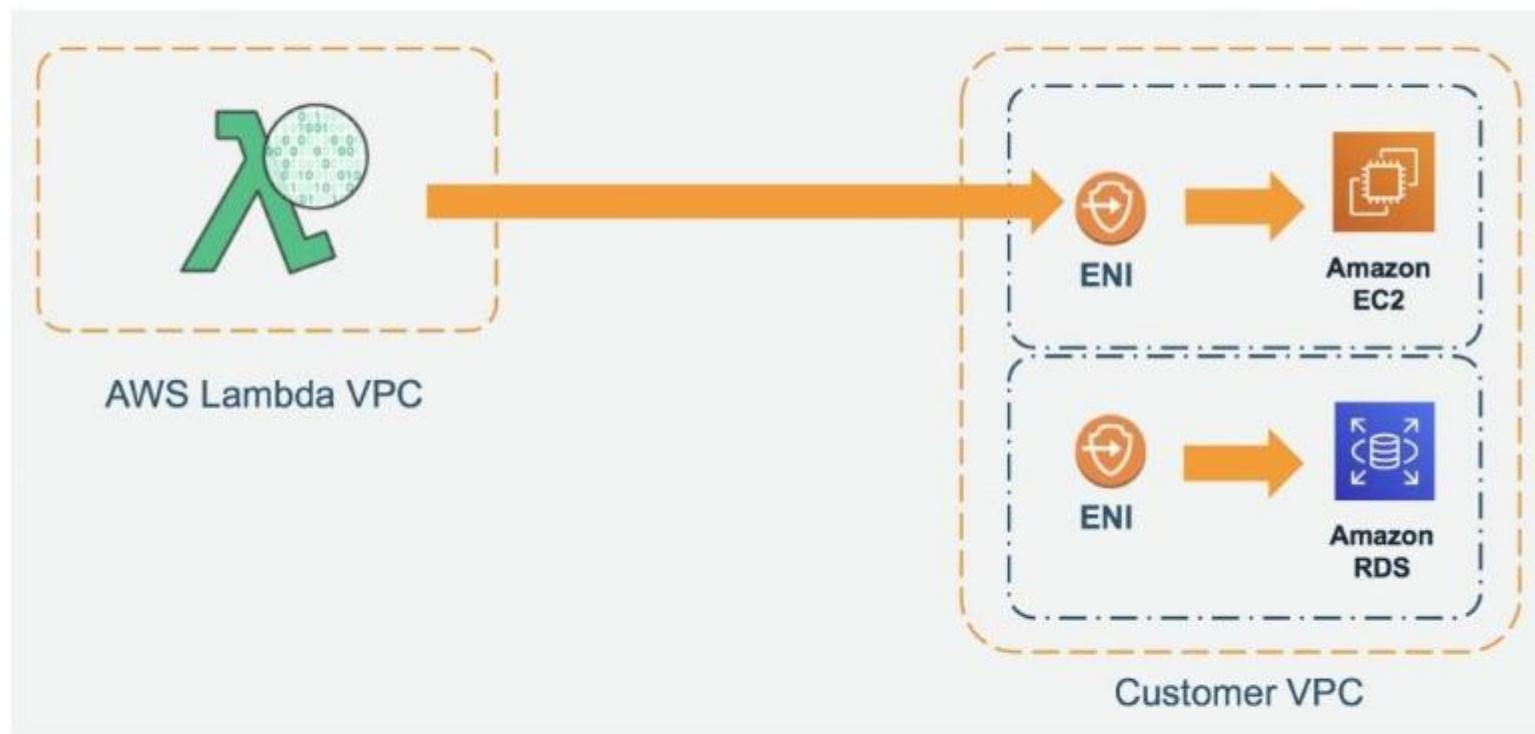
Once your function is VPC-enabled, all network traffic from your function is subject to the routing rules of your VPC/Subnet. If your function needs to interact with a public resource, you will need a route through a NAT gateway in a public subnet.

When to VPC-Enable a Lambda

Function:

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

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



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

via -

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

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

If you intend to reuse code in more than one Lambda function, you should consider creating a Lambda Layer for the reusable code - You can configure your Lambda function to pull in additional code and content in the form of layers. A layer is a ZIP archive that contains libraries, a custom runtime, or other dependencies. With layers, you can use libraries in your function without needing to include them in your deployment package. Layers let you keep your deployment package small, which makes development easier. A function can use up to 5 layers at a time.

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

Incorrect options:

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

The bigger your deployment package, the slower your Lambda function will cold-start. Hence, AWS suggests packaging dependencies as a separate package from the actual Lambda package - This statement is added as a distractor. All the dependencies can be packaged into the single Lambda deployment package without any performance impact.

Serverless databases and Lambda complement each other and you should install databases on the Lambda functions - This statement is incorrect. AWS Lambda does not support installation of databases.

References:

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

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

Question 70: **Incorrect**

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

As a Solutions Architect Professional, which of the following solutions would you recommend to meet these requirements?

-

Configure VPC Endpoints between branch offices and corporate headquarters which will enable branch offices to send and receive data with each other as well as with their corporate headquarters

-

Configure Public Virtual Interfaces (VIFs) between branch offices and corporate headquarters which will enable branch offices to send and receive data with each other as well as with their corporate headquarters

(Incorrect)

-

Set up VPN CloudHub between branch offices and corporate headquarters which will enable branch offices to send and receive data with each other as well as with their corporate headquarters

(Correct)



Set up VPC Peering between branch offices and corporate headquarters which will enable branch offices to send and receive data with each other as well as with their corporate headquarters

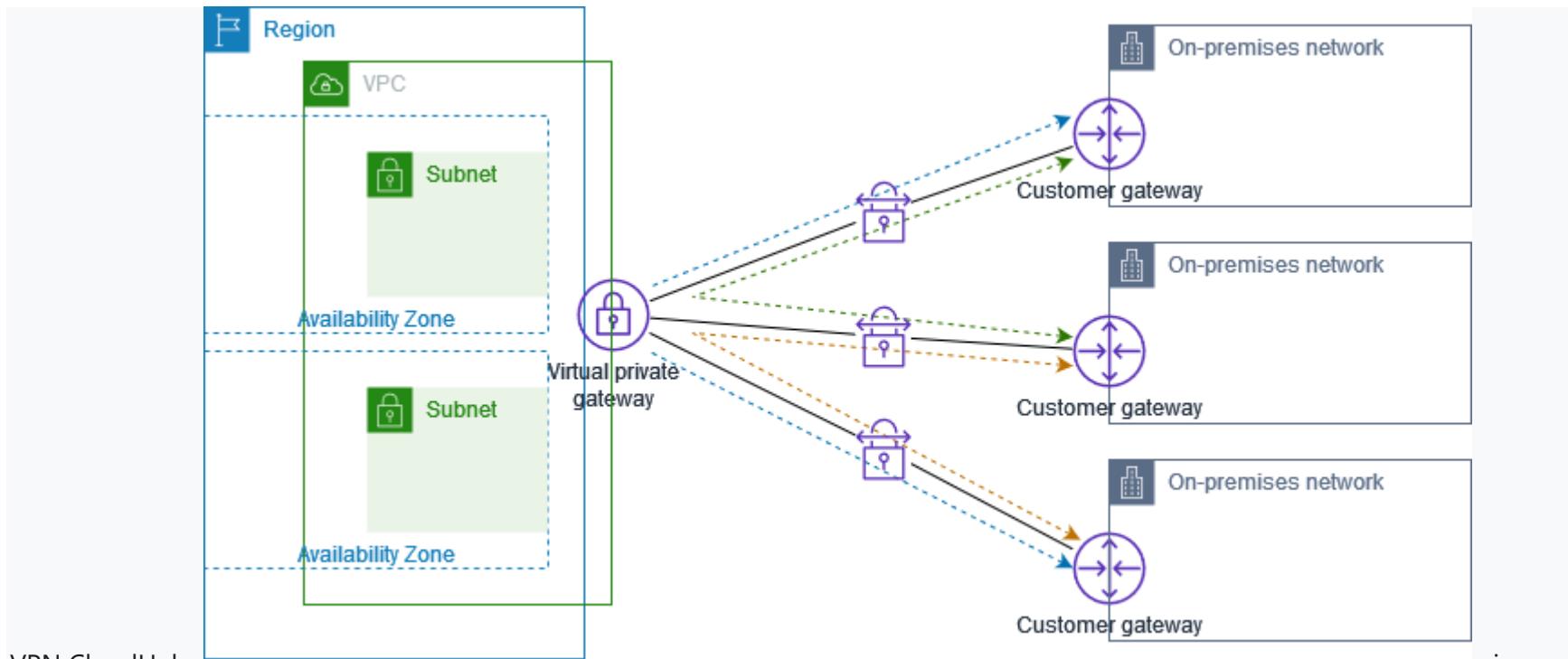
Explanation

Correct option:

Set up VPN CloudHub between branch offices and corporate headquarters which will enable branch offices to send and receive data with each other as well as with their corporate headquarters

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

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



VPN CloudHub:

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

via -

Incorrect options:

Configure VPC Endpoints between branch offices and corporate headquarters which will enable branch offices to send and receive data with each other as well as with their corporate headquarters - A VPC endpoint enables you to privately connect your VPC to supported AWS services and VPC endpoint services powered by AWS PrivateLink without requiring an internet gateway, NAT device, VPN connection, or AWS Direct Connect connection. Instances in your VPC do not require public IP addresses to communicate with resources in the service. AWS PrivateLink simplifies the security of data shared with cloud-based applications by eliminating the exposure of data to the public Internet.

When you use VPC endpoint, the traffic between your VPC and the other AWS service does not leave the Amazon network, therefore this option cannot be used to send and receive data between the remote branch offices of the company.

Set up VPC Peering between branch offices and corporate headquarters which will enable branch offices to send and receive data with each other as well as with their corporate headquarters - A VPC peering connection is a networking connection between two VPCs that enables you to route traffic between them using private IPv4 addresses or IPv6 addresses. Instances in either VPC can communicate with each other as if they are within the same network.

VPC peering facilitates a connection between two VPCs within the AWS network, therefore this option cannot be used to send and receive data between the remote branch offices of the company.

Configure Public Virtual Interfaces (VIFs) between branch offices and corporate headquarters which will enable branch offices to send and receive data with each other as well as with their corporate headquarters - AWS Direct Connect (DX) provides three types of virtual interfaces: public, private, and transit. To connect to AWS resources that are reachable by a public IP address (such as an Amazon Simple Storage Service bucket) or AWS public endpoints, use a public virtual interface. Therefore this option cannot be used to send and receive data between the remote branch offices of the company.

AWS Direct Connect virtual interfaces

[PDF](#) | [Kindle](#) | [RSS](#)

You must create one of the following virtual interfaces to begin using your AWS Direct Connect connection.

- Private virtual interface: A private virtual interface should be used to access an Amazon VPC using private IP addresses.
- Public virtual interface: A public virtual interface can access all AWS public services using public IP addresses.
- Transit virtual interface: A transit virtual interface should be used to access one or more Amazon VPC Transit Gateways associated with Direct Connect gateways. You can use transit virtual interfaces with 1/2/5/10 Gbps AWS Direct Connect connections. For information about Direct Connect gateway configurations, see [Direct Connect gateways](#).

via -

<https://docs.aws.amazon.com/directconnect/latest/UserGuide/WorkingWithVirtualInterfaces.html>

References:

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

https://docs.aws.amazon.com/vpn/latest/s2vpn/VPN_CloudHub.html

<https://docs.aws.amazon.com/directconnect/latest/UserGuide/WorkingWithVirtualInterfaces.html>

Question 71: **Correct**

A leading club in the Major League Baseball runs a web platform that boasts over 50,000 pages and over 100 million digitized photographs. It is available in six languages and maintains up-to-date information for the season. The engineering team has built a notification system on the web platform using SNS notifications which are then handled by a Lambda function for end-user delivery. During the off-season, the notification systems need to handle about 100 requests per second. During the peak baseball season, the rate touches about 5000 requests per second and it is noticed that a significant number of the notifications are not being delivered to the end-users on the web platform.

As a Solutions Architect Professional, which of the following would you suggest as the BEST fit solution to address this issue?



The engineering team needs to provision more servers running the SNS service



The engineering team needs to provision more servers running the Lambda service



Amazon SNS message deliveries to AWS Lambda have crossed the account concurrency quota for Lambda, so the team needs to contact AWS support to raise the account limit

(Correct)



Amazon SNS has hit a concurrency limit, so the team needs to contact AWS support to raise the account limit

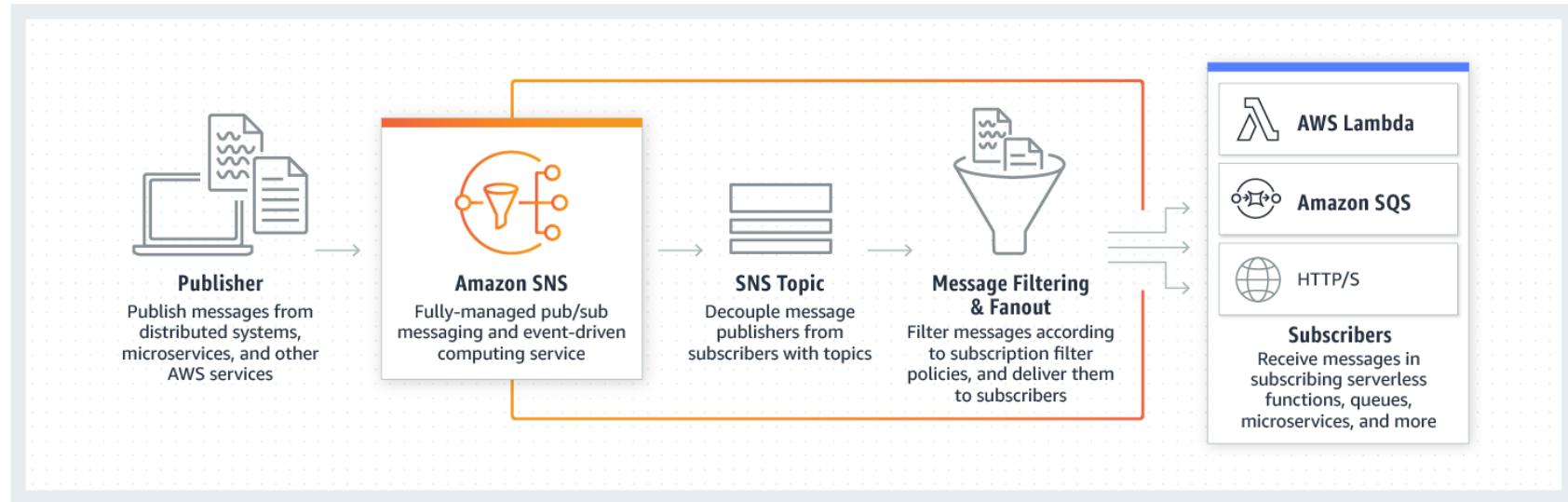
Explanation

Correct option: **Amazon SNS message deliveries to AWS Lambda have crossed the account concurrency quota for Lambda, so the team needs to contact AWS support to raise the account limit**

Amazon Simple Notification Service (SNS) is a highly available, durable, secure, fully managed pub/sub messaging service that enables you to decouple microservices, distributed systems, and serverless applications.

How SNS

Works:



vi

a - <https://aws.amazon.com/sns/>

With AWS Lambda, you can run code without provisioning or managing servers. You pay only for the compute time that you consume—there's no charge when your code isn't running.

AWS Lambda currently supports 1000 concurrent executions per AWS account per region. If your Amazon SNS message deliveries to AWS Lambda contribute to crossing these concurrency quotas, your Amazon SNS message deliveries will be throttled. You need to contact AWS support to raise the account limit. Therefore this option is correct.

AWS Lambda quotas

[PDF](#) | [Kindle](#) | [RSS](#)

AWS Lambda sets quotas for the amount of compute and storage resources that you can use to run and store functions. The following quotas apply per-region and can be increased. To request an increase, use the [Support Center console](#).

Resource	Default quota	Can Be Increased Up To
Concurrent executions	1,000	Hundreds of thousands
Function and layer storage	75 GB	Terabytes
Elastic network interfaces per VPC	250	Hundreds

For details on concurrency and how Lambda scales your function concurrency in response to traffic, see [AWS Lambda function scaling](#).

The following quotas apply to function configuration, deployments, and execution. They cannot be changed.

Resource	Quota
Function memory allocation	128 MB to 3,008 MB, in 64 MB increments.
Function timeout	900 seconds (15 minutes)
Function environment variables	4 KB
Function resource-based policy	20 KB
Function layers	5 layers
Function burst concurrency	500 - 3000 (varies per region)
Invocation payload (request and response)	6 MB (synchronous) 256 KB (asynchronous)
Deployment package size	50 MB (zipped, for direct upload) 250 MB (unzipped, including layers) 3 MB (console editor)
Test events (console editor)	10
/tmp directory storage	512 MB
File descriptors	1,024
Execution processes/threads	1,024

via -

<https://docs.aws.amazon.com/lambda/latest/dg/gettingstarted-limits.html>

Incorrect options: **Amazon SNS has hit a concurrency limit, so the team needs to contact AWS support to raise the account limit** - Amazon SNS leverages the proven AWS cloud to dynamically scale with your application. You don't need to contact AWS support, as SNS is a fully managed service, taking care of the heavy lifting related to capacity planning, provisioning, monitoring, and patching. Therefore, this option is incorrect.

The engineering team needs to provision more servers running the SNS service

The engineering team needs to provision more servers running the Lambda service

As both Lambda and SNS are serverless and fully managed services, the engineering team cannot provision more servers. Both of these options are incorrect.

Reference: <https://aws.amazon.com/sns/>

<https://aws.amazon.com/sns/faqs/>

<https://docs.aws.amazon.com/lambda/latest/dg/gettingstarted-limits.html>

Question 72: **Incorrect**

An IT company wants to move all its clients belonging to the regulated and security-sensitive industries such as financial services and healthcare to the AWS Cloud as it wants to leverage the out-of-box security-specific capabilities offered by AWS. The Security team at the company is developing a framework to validate the adoption of AWS best practices and industry-recognized compliance standards. The AWS Management Console is the preferred method for the in-house teams wanting to provision resources. You have been hired as an AWS Certified Solutions Architect Professional to spearhead this strategic initiative.

Which of the following strategies would you adopt to address these business requirements for continuously assessing, auditing and monitoring the configurations of AWS resources? (Select two)



Leverage Config rules to audit changes to AWS resources and monitor the compliance of the configuration by running the evaluations for the rule at a frequency that you choose. Develop AWS Config custom rules to establish a test-driven development approach by triggering the evaluation when any resource that matches the rule's scope changes in configuration

(Correct)



Leverage CloudWatch Logs agent to collect all the AWS SDK logs. Search the log data using a pre-defined set of filter patterns that match mutating API calls. Use CloudWatch alarms to send notifications via SNS when unintended changes are performed. Archive log data by using a batch export to Amazon S3 and analyze via Athena



Enable trails and set up CloudTrail events to review and monitor management activities of all AWS accounts by logging these activities into CloudWatch Logs using a KMS key. Ensure that CloudTrail is enabled for all accounts as well as all available AWS services

(Correct)



Leverage CloudTrail integration with SNS to automatically notify unauthorized API activities. Ensure that CloudTrail is enabled for all accounts as well as all available AWS services. Use Lambda functions to automatically revert non-authorized changes in AWS resources

(Incorrect)



Leverage EventBridge events near-real-time capabilities to monitor system events patterns to trigger Lambda functions to automatically revert non-authorized changes in AWS resources. Send notifications via SNS topics to improve the incidence response time

Explanation

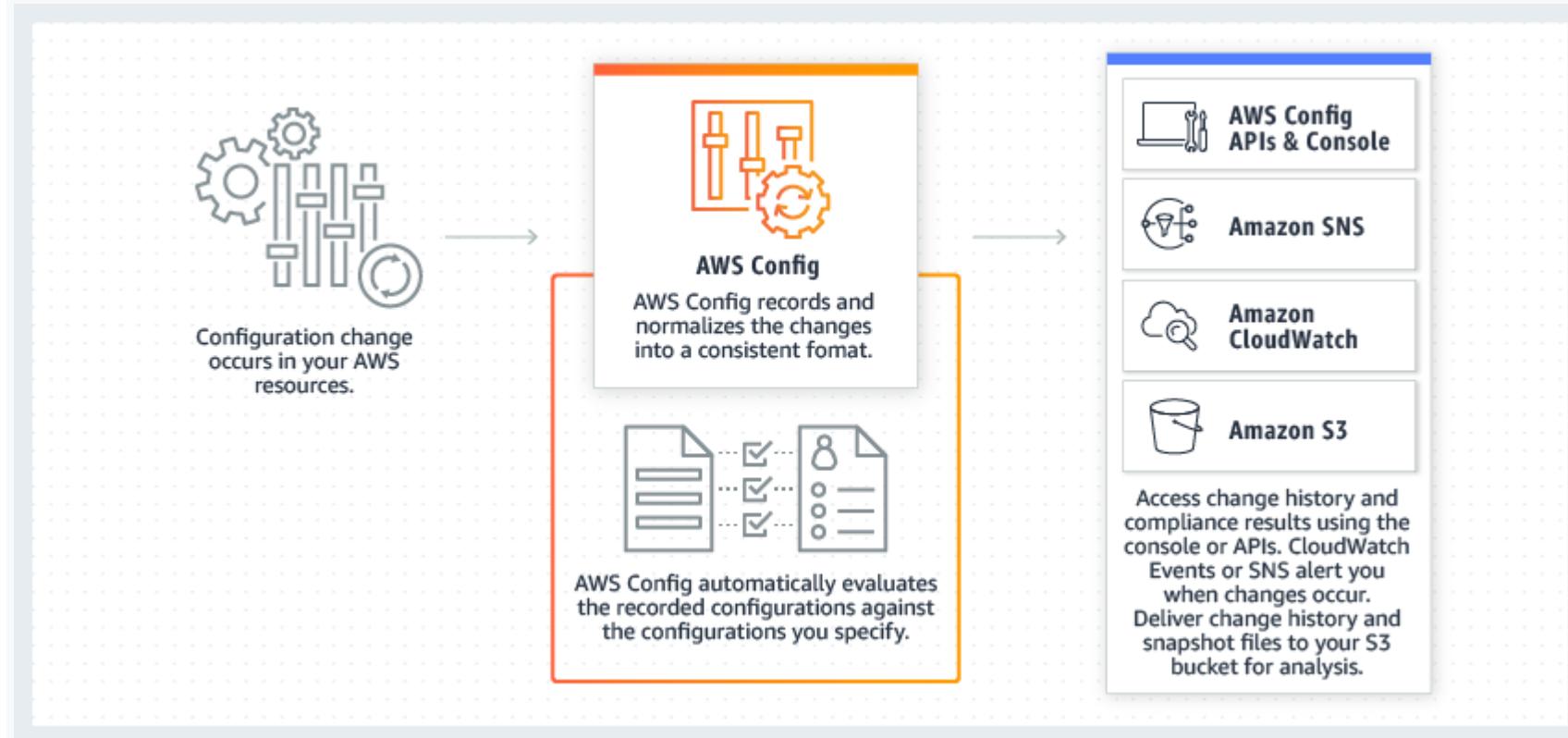
Correct options:

Leverage Config rules to audit changes to AWS resources and monitor the compliance of the configuration by running the evaluations for the rule at a frequency that you choose. Develop AWS Config custom rules to establish a test-driven development approach by triggering the evaluation when any resource that matches the rule's scope changes in configuration

AWS Config is a service that enables you to assess, audit, and evaluate the configurations of your AWS resources. With Config, you can review changes in configurations and relationships between AWS resources, dive into detailed resource configuration histories,

and determine your overall compliance against the configurations specified in your internal guidelines. You can use Config to answer questions such as - "What did my AWS resource look like at xyz point in time?".

How AWS Config Works:



via - <https://aws.amazon.com/config/>

For the given use-case, you can use AWS Config to evaluate the configuration settings of your AWS resources. You do this by creating AWS Config rules, which represent your ideal configuration settings. AWS Config provides customizable, predefined rules called managed rules to help you get started. You can also create your own custom rules. While AWS Config continuously tracks the

configuration changes that occur among your resources, it checks whether these changes violate any of the conditions in your rules. If a resource violates a rule, AWS Config flags the resource and marks the rule as noncompliant.

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

There are two types of evaluation trigger types for Config rules:

Configuration changes – AWS Config triggers the evaluation when any resource that matches the rule's scope changes in configuration. The evaluation runs after AWS Config sends a configuration item change notification.

Periodic – AWS Config runs evaluations for the rule at a frequency that you choose (for example, every 24 hours).

Enable trails and set up CloudTrail events to review and monitor management activities of all AWS accounts by logging these activities into CloudWatch Logs using a KMS key. Ensure that CloudTrail is enabled for all accounts as well as all available AWS services

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. An event in CloudTrail is the record of activity in an AWS account. This activity can be an action taken by a user, role, or service that is monitorable by CloudTrail. CloudTrail events provide a history of both API and non-API account activity made through the AWS Management Console, AWS SDKs, command line tools, and other AWS services.

CloudTrail data events are disabled by default. You can enable logging at an additional cost. Data events are also known as data plane operations and are often high-volume activities. Data events aren't viewable in CloudTrail event history and are charged for all copies at a reduced rate compared to management events.

CloudTrail records management events for the last 90 days free of charge, and are viewable in the Event History with the CloudTrail console. For Amazon S3 delivery of CloudTrail events, the first copy delivered is free. Additional copies of management events are charged.

What Are CloudTrail Events?

An event in CloudTrail is the record of an activity in an AWS account. This activity can be an action taken by a user, role, or service that is monitorable by CloudTrail. CloudTrail events provide a history of both API and non-API account activity made through the AWS Management Console, AWS SDKs, command line tools, and other AWS services. There are two types of events that can be logged in CloudTrail: management events and data events. By default, trails log management events, but not data events.

Both management events and data events use the same CloudTrail JSON log format.

 **Note**

CloudTrail does not log all AWS services. Some AWS services do not enable logging of all APIs and events. Even if you configure logging all management and data events in a trail, you will not create a log with all possible AWS events. For details about which APIs are logged for a specific service, see documentation for that service in [CloudTrail Supported Services and Integrations](#).

What Are Management Events?

Management events provide information about management operations that are performed on resources in your AWS account. These are also known as *control plane operations*. Example management events include:

- Configuring security (for example, IAM AttachRolePolicy API operations).
- Registering devices (for example, Amazon EC2 CreateDefaultVpc API operations).
- Configuring rules for routing data (for example, Amazon EC2 CreateSubnet API operations).
- Setting up logging (for example, AWS CloudTrail CreateTrail API operations).

Management events can also include non-API events that occur in your account. For example, when a user signs in to your account, CloudTrail logs the ConsoleLogin event. For more information, see [Non-API Events Captured by CloudTrail](#). For a list of management events that CloudTrail logs for AWS services, see [CloudTrail Supported Services and Integrations](#).

What Are Data Events?

Data events provide information about the resource operations performed on or in a resource. These are also known as *data plane operations*. Data events are often high-volume activities. The following two data types are recorded:

- Amazon S3 object-level API activity (for example, GetObject, DeleteObject, and PutObject API operations).
- AWS Lambda function execution activity (the Invoke API).

Data events are disabled by default when you create a trail. To record CloudTrail data events, you must explicitly add to a trail the supported resources or resource types for which you want to collect activity. For more information, see [Creating a Trail and Data Events](#).

<https://docs.aws.amazon.com/awscloudtrail/latest/userguide/cloudtrail-concepts.html#cloudtrail-concepts-management-events>

Incorrect options:

Leverage CloudWatch Logs agent to collect all the AWS SDK logs. Search the log data using a pre-defined set of filter patterns that match mutating API calls. Use CloudWatch alarms to send notifications via SNS when unintended changes are performed. Archive log data by using a batch export to Amazon S3 and analyze via Athena - One of the key constraints for the given scenario is that the AWS Management Console is the preferred method for the in-house teams wanting to provision resources. Although this option is technically feasible, but it focuses on using CloudWatch Logs agent to collect all the AWS SDK logs. The given use-case has no specific requirements for AWS SDKs or AWS APIs because AWS Management Console is the preferred method to provision resources. So this option is not the best fit solution.

Leverage CloudTrail integration with SNS to automatically notify unauthorized API activities. Ensure that CloudTrail is enabled for all accounts as well as all available AWS services. Use Lambda functions to automatically revert non-authorized changes in AWS resources - One of the key constraints for the given scenario is that the AWS Management Console is the preferred method for the in-house teams wanting to provision resources. Although this option is technically feasible, but it focuses on capturing unauthorized API activities. The given use-case has no specific requirements for AWS SDKs or AWS APIs because AWS Management Console is the preferred method to provision resources. In addition, the use-case just talks about assessing, auditing and monitoring the configurations of AWS resources. Reverting non-authorized changes in AWS resources is not part of the mandate. So this option is not correct.

Leverage EventBridge events near-real-time capabilities to monitor system events patterns to trigger Lambda functions to automatically revert non-authorized changes in AWS resources. Send notifications via SNS topics to improve the incidence response time - The use-case just talks about assessing, auditing and monitoring the configurations of AWS resources. Reverting non-authorized changes in AWS resources is not part of the mandate. So this option is not correct.

References:

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

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

<https://docs.aws.amazon.com/awscloudtrail/latest/userguide/cloudtrail-concepts.html#cloudtrail-concepts-management-events>

Question 73: **Incorrect**

A social media company has a serverless application stack that consists of CloudFront, API Gateway and Lambda functions. The company has hired you as an AWS Certified Solutions Architect Professional to improve the current deployment process which creates a new version of the Lambda function and then runs an AWS CLI script for deployment. In case the new version errors out, then another CLI script is invoked to deploy the previous working version of the Lambda function. The company has mandated you to decrease the time to deploy new versions of the Lambda functions and also reduce the time to detect and rollback when errors are identified.

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

- Set up and deploy nested CloudFormation stacks with the CloudFront distribution as well as the API Gateway in the parent stack. Create and deploy a child stack containing the Lambda functions. To address any changes in a Lambda function, create a CloudFormation change set and deploy. Use pre-traffic and post-traffic test functions of the change set to verify the deployment. Rollback in case CloudWatch alarms are triggered
-
- Set up and deploy a CloudFormation stack containing a new API Gateway endpoint that points to the new Lambda version. Test the updated CloudFront origin that points to this new API Gateway endpoint and in case errors are detected then revert the CloudFront origin to the previous working API Gateway endpoint
-
- Set up and deploy nested CloudFormation stacks with the CloudFront distribution as well as the API Gateway in the parent stack. Create and deploy a child stack containing the Lambda functions. To address any changes in a Lambda function, create a CloudFormation change set and deploy. In case the Lambda function errors out, rollback the CloudFormation change set to the previous version

(Incorrect)

-

Use Serverless Application Model (SAM) and leverage the built-in traffic-shifting feature of SAM to deploy the new Lambda version via CodeDeploy and use pre-traffic and post-traffic test functions to verify code. Rollback in case CloudWatch alarms are triggered

(Correct)

Explanation

Correct option:

Use Serverless Application Model (SAM) and leverage the built-in traffic-shifting feature of SAM to deploy the new Lambda version via CodeDeploy and use pre-traffic and post-traffic test functions to verify code. Rollback in case CloudWatch alarms are triggered

The AWS Serverless Application Model (SAM) is an open source framework for building serverless applications. It provides shorthand syntax to express functions, APIs, databases, and event source mappings. You define the application you want with just a few lines per resource and model it using YAML. During deployment, SAM transforms and expands the SAM syntax into AWS CloudFormation syntax. Then, CloudFormation provisions your resources with reliable deployment capabilities.

To address the given use-case, you can use the traffic shifting feature of SAM to easily test the new version of the Lambda function without having to manually move 100% of the traffic to the new version in one shot.

You can use CodeDeploy to create a deployment process that publishes the new Lambda version but does not send any traffic to it. Then it executes a PreTraffic test to ensure that your new function works as expected. After the test succeeds, CodeDeploy automatically shifts traffic gradually to the new version of the Lambda function. This workflow address one of the key requirements of reducing the time to detect errors. You can roll back to the previous version in case the new version errors out.

Implementing safe AWS Lambda deployments with AWS CodeDeploy

by Chris Munns | on 19 APR 2018 | in Amazon API Gateway, AWS CodeDeploy, AWS Lambda, Serverless | Permalink | [Comments](#) | [Share](#)

This post courtesy of George Mao, AWS Senior Serverless Specialist – Solutions Architect

AWS Lambda and AWS CodeDeploy recently made it possible to automatically [shift incoming traffic](#) between two function versions based on a preconfigured rollout strategy. This new feature allows you to gradually shift traffic to the new function. If there are any issues with the new code, you can quickly rollback and control the impact to your application.

Previously, you had to manually move 100% of traffic from the old version to the new version. Now, you can have CodeDeploy automatically execute pre- or post-deployment tests and automate a gradual rollout strategy. Traffic shifting is built right into the AWS Serverless Application Model (SAM), making it easy to define and deploy your traffic shifting capabilities. SAM is an extension of AWS CloudFormation that provides a simplified way of defining serverless applications.

In this post, I show you how to use SAM, CloudFormation, and CodeDeploy to accomplish an automated rollout strategy for safe Lambda deployments.

Scenario

For this walkthrough, you write a Lambda application that returns a count of the S3 buckets that you own. You deploy it and use it in production. Later on, you receive requirements that tell you that you need to change your Lambda application to count only buckets that begin with the letter "a".

Before you make the change, you need to be sure that your new Lambda application works as expected. If it does have issues, you want to minimize the number of impacted users and roll back easily. To accomplish this, you create a deployment process that publishes the new Lambda function, but does not send any traffic to it. You use CodeDeploy to execute a PreTraffic test to ensure that your new function works as expected. After the test succeeds, CodeDeploy automatically shifts traffic gradually to the new version of the Lambda function.

Your Lambda function is exposed as a REST service via an Amazon API Gateway deployment. This makes it easy to test and integrate.

via -

<https://aws.amazon.com/blogs/compute/implementing-safe-aws-lambda-deployments-with-aws-codedeploy/>

Incorrect options:

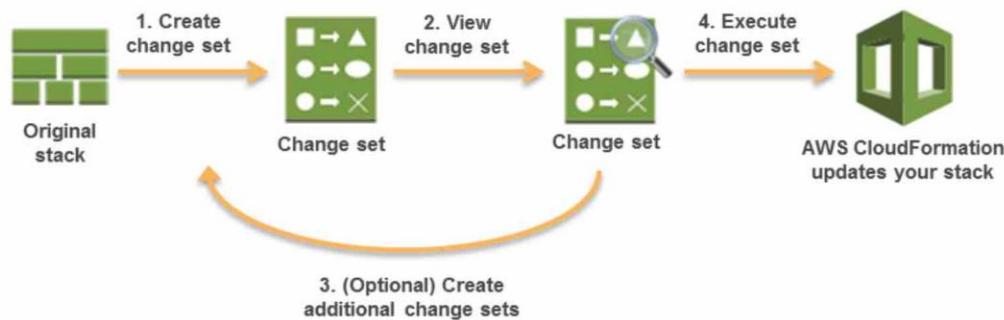
Set up and deploy nested CloudFormation stacks with the CloudFront distribution as well as the API Gateway in the parent stack. Create and deploy a child stack containing the Lambda functions. To address any changes in a Lambda function, create

a CloudFormation change set and deploy. In case the Lambda function errors out, rollback the CloudFormation change set to the previous version - You can use CloudFormation change sets to preview how proposed changes to a stack might impact your running resources, for example, whether your changes will delete or replace any critical resources, AWS CloudFormation makes the changes to your stack only when you decide to execute the change set, allowing you to decide whether to proceed with your proposed changes or explore other changes by creating another change set.

This option does not help in reducing the time to detect any potential deployment errors as you would not know about any potential failures until you actually deploy the stack.

Change Set Overview

The following diagram summarizes how you use change sets to update a stack:



1. Create a change set by submitting changes for the stack that you want to update. You can submit a modified stack template or modified input parameter values. AWS CloudFormation compares your stack with the changes that you submitted to generate the change set; it doesn't make changes to your stack at this point.
2. View the change set to see which stack settings and resources will change. For example, you can see which resources AWS CloudFormation will add, modify, or delete.
3. Optional: If you want to consider other changes before you decide which changes to make, create additional change sets. Creating multiple change sets helps you understand and evaluate how different changes will affect your resources. You can create as many change sets as you need.
4. Execute the change set that contains the changes that you want to apply to your stack. AWS CloudFormation updates your stack with those changes.

Note

After you execute a change, AWS CloudFormation removes all change sets that are associated with the stack because they aren't applicable to the updated stack.

via -

<https://docs.aws.amazon.com/AWSCloudFormation/latest/UserGuide/using-cfn-updating-stacks-changesets.html>

Instead, you should use SAM to create your serverless application as it comes built-in with CodeDeploy to provide gradual Lambda deployments. Also, you can define pre-traffic and post-traffic test functions to verify that the newly deployed code is configured correctly and your application operates as expected. You can roll back the deployment if CloudWatch alarms are triggered.

Set up and deploy a CloudFormation stack containing a new API Gateway endpoint that points to the new Lambda version. Test the updated CloudFront origin that points to this new API Gateway endpoint and in case errors are detected then revert

the CloudFront origin to the previous working API Gateway endpoint - This option does not help in reducing the time to detect any potential deployment errors as you would not know about any potential failures until you actually deploy the stack and point to the new endpoint.

Instead, you should use SAM to create your serverless application as it comes built-in with CodeDeploy to provide gradual Lambda deployments. Also, you can define pre-traffic and post-traffic test functions to verify that the newly deployed code is configured correctly and your application operates as expected. You can roll back the deployment if CloudWatch alarms are triggered.

Set up and deploy nested CloudFormation stacks with the CloudFront distribution as well as the API Gateway in the parent stack. Create and deploy a child stack containing the Lambda functions. To address any changes in a Lambda function, create a CloudFormation change set and deploy. Use pre-traffic and post-traffic test functions of the change set to verify the deployment. Rollback in case CloudWatch alarms are triggered - This option has been added as a distractor, since CloudFormation change sets do not have pre-traffic and post-traffic test functions. Therefore this option is incorrect.

References:

<https://aws.amazon.com/blogs/compute/implementing-safe-aws-lambda-deployments-with-aws-codedeploy/>

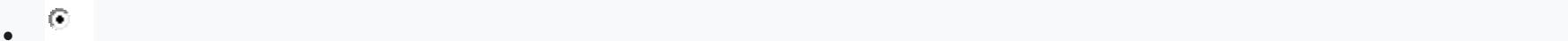
<https://docs.aws.amazon.com/AWSCloudFormation/latest/UserGuide/using-cfn-updating-stacks-changesets.html>

<https://docs.aws.amazon.com/serverless-application-model/latest/developerguide/automating-updates-to-serverless-apps.html>

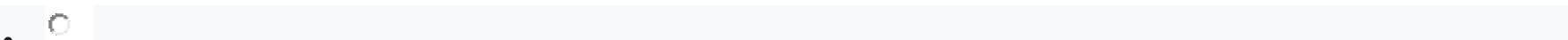
Question 74: **Incorrect**

A leading video creation and distribution company has recently migrated to AWS Cloud for digitally transforming its movie business. The company wants to speed up its media distribution process and improve data security while also reducing costs and eliminating errors. The company wants to set up a Digital Cinema Network that would allow it to store content in Amazon S3 as well as to accelerate the online distribution of movies and advertising to theaters in 38 key media markets worldwide. The company also wants to do an accelerated online migration of hundreds of terabytes of files from their on-premises data center to Amazon S3 and then establish a mechanism for low-latency access of the migrated data for ongoing updates from the on-premises applications.

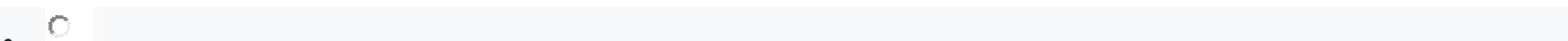
As a Solutions Architect Professional, which of the following would you select as the MOST performant solution for the given use-case?



(Incorrect)



(Correct)



Explanation

Correct options:

Use AWS DataSync to migrate existing data to Amazon S3 and then use File Gateway for low latency access to the migrated data for ongoing updates from the on-premises applications

AWS DataSync is an online data transfer service that simplifies, automates, and accelerates copying large amounts of data to and from AWS storage services over the internet or AWS Direct Connect. AWS DataSync fully automates and accelerates moving large active datasets to AWS, up to 10 times faster than command-line tools. It is natively integrated with Amazon S3, Amazon EFS,

Amazon FSx for Windows File Server, Amazon CloudWatch, and AWS CloudTrail, which provides seamless and secure access to your storage services, as well as detailed monitoring of the transfer.

DataSync uses a purpose-built network protocol and scale-out architecture to transfer data. A single DataSync agent is capable of saturating a 10 Gbps network link. DataSync fully automates the data transfer. It comes with retry and network resiliency mechanisms, network optimizations, built-in task scheduling, monitoring via the DataSync API and Console, and CloudWatch metrics, events, and logs that provide granular visibility into the transfer process. DataSync performs data integrity verification both during the transfer and at the end of the transfer.

How DataSync



Works

via -

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

AWS Storage Gateway is a hybrid cloud storage service that gives you on-premises access to virtually unlimited cloud storage. The service provides three different types of gateways – Tape Gateway, File Gateway, and Volume Gateway – that seamlessly connect on-premises applications to cloud storage, caching data locally for low-latency access. File gateway offers SMB or NFS-based access to data in Amazon S3 with local caching.

The combination of DataSync and File Gateway is the correct solution. AWS DataSync enables you to automate and accelerate online data transfers to AWS storage services. File Gateway then provides your on-premises applications with low latency access to the migrated

Nearly all enterprises, regardless of industry, have to store files, whether they are backups, media content, or files generated by specialized industry applications. Managing and scaling on-premises infrastructure to provide online storage and distribution of such backup or content files is often burdensome and costly, requiring expensive hardware refreshes, data center expansion, and software licensing. These large file data repositories can be siloed in specialized file servers, NAS units, or backup systems, limiting access for big data analytics or media processing applications.

File Gateway provides a seamless way to connect to the cloud in order to store application data files and backup images as durable objects in Amazon S3 cloud storage. File Gateway offers SMB or NFS-based access to data in Amazon S3 with local caching. It can be used for on-premises applications, and for Amazon EC2-based applications that need file protocol access to S3 object storage.

How it works



Benefits

Provides on-premises access to virtually unlimited cloud storage on-demand

Extends the AWS Cloud to your data center

Requires no changes to existing applications

Reduces physical infrastructure, cost, and complexity

Gets your data into Amazon S3 for use with other AWS services

Utilizes standard storage protocols with NFS & SMB

data.Gateway:

via - <https://aws.amazon.com/storagegateway/file/>

Incorrect options:

Use AWS DataSync to first migrate existing data to Amazon S3 and then configure low latency access to the migrated data for ongoing updates from the on-premises applications - AWS DataSync is used to easily transfer data to and from AWS with up to 10x faster speeds. It is used to transfer data and should not be used for low latency access to the migrated data for ongoing updates from the on-premises applications.

Use File Gateway configuration of AWS Storage Gateway to migrate data to Amazon S3 and then use S3 Transfer Acceleration for low latency access to the migrated data for ongoing updates from the on-premises applications - File Gateway can be used to move on-premises data to AWS Cloud, but it is not an optimal solution for high volumes. Migration services such as DataSync are best suited for this purpose. S3 Transfer Acceleration cannot facilitate low latency access to the migrated data for ongoing updates from the on-premises applications.

Use S3 Transfer Acceleration to migrate existing data to Amazon S3 and then use DataSync for low latency access to the migrated data for ongoing updates from the on-premises applications - If your application is already integrated with the Amazon S3 API, and you want higher throughput for transferring large files to S3, S3 Transfer Acceleration can be used. However, DataSync should not be used for low latency access to the migrated data for ongoing updates from the on-premises applications.

Q: When do I use AWS DataSync and when do I use AWS Snowball Edge?

A: AWS DataSync is ideal for online data transfers. You can use DataSync to migrate active data to AWS, transfer data to the cloud for analysis and processing, archive data to free up on-premises storage capacity, or replicate data to AWS for business continuity.

[AWS Snowball Edge](#) is suitable for offline data transfers, for customers who are bandwidth constrained, or transferring data from remote, disconnected, or austere environments.

Q: When do I use AWS DataSync and when do I use AWS Storage Gateway?

A: Use AWS DataSync to migrate existing data to Amazon S3, and then use the File Gateway configuration of AWS Storage Gateway to retain access to the migrated data and for ongoing updates from your on-premises file-based applications.

You can use a combination of DataSync and File Gateway to minimize your on-premises infrastructure while seamlessly connecting on-premises applications to your cloud storage. AWS DataSync enables you to automate and accelerate online data transfers to AWS storage services. File Gateway then provides your on-premises applications with low latency access to the migrated data.

Q: When do I use AWS DataSync, and when do I use Amazon S3 Transfer Acceleration?

A: If your applications are already integrated with the Amazon S3 API, and you want higher throughput for transferring large files to S3, you can use [S3 Transfer Acceleration](#). If you want to transfer data from existing storage systems (e.g. Network Attached Storage), or from instruments that cannot be changed (e.g. DNA sequencers, video cameras), or if you want multiple destinations, you use AWS DataSync. DataSync also automates and simplifies the data transfer by providing additional functionality, such as built-in retry and network resiliency mechanisms, data integrity verification, and flexible configuration to suit your specific needs, including bandwidth throttling, etc.

Q: When do I use AWS DataSync and when do I use AWS Transfer for SFTP?

A: If you currently use SFTP to exchange data with third parties, [AWS Transfer for SFTP](#) provides a fully managed SFTP transfer directly into and out of Amazon S3, while reducing your operational burden.

If you want an accelerated and automated data transfer between NFS servers, SMB file shares, self-managed object storage, AWS Snowcone, Amazon S3, Amazon EFS, and Amazon FSx for Windows File Server, you can use AWS DataSync.

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

References:

<https://aws.amazon.com/datasync/features/>

<https://aws.amazon.com/storagegateway/file/>

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

Question 75: **Correct**

The DevOps team at a financial services company has provisioned a new GPU optimized EC2 instance X by choosing the default security group of the default VPC. The team can ping instance X from other instances in the VPC. The other instances were also created using the default security group. The next day, the team launches another GPU optimized instance Y by creating a new security group and attaching it to instance Y. All other configuration options for instance Y are chosen as default. However, the team is not able to ping instance Y from other instances in the VPC.

As a Solutions Architect Professional, which of the following would you identify as the root cause of the issue?

-

Instance X is in the default security group. The default rules for the default security group allow inbound traffic from network interfaces (and their associated instances) that are assigned to the same security group. Instance Y is in a new security group. The default rules for a security group that you create allow no inbound traffic

(Correct)

-

Instance X is in the default security group. The default rules for the default security group allow inbound traffic from all sources. Instance Y is in a new security group. The default rules for a security group that you create allow no inbound traffic

-

**Instance X is in the default security group. The default rules for the default security group allow no inbound traffic from all sources.
Instance Y is in a new security group. The default rules for a security group that you create allow inbound traffic from all sources**

-

Instance X is in the default security group. The default rules for the default security group allow no inbound traffic from network interfaces (and their associated instances) that are assigned to the same security group. Instance Y is in a new security group. The default rules for a security group that you create allow inbound traffic from all sources

Explanation

Correct option:

Instance X is in the default security group. The default rules for the default security group allow inbound traffic from network interfaces (and their associated instances) that are assigned to the same security group. Instance Y is in a new security group. The default rules for a security group that you create allow no inbound traffic

A security group acts as a virtual firewall that controls the traffic for one or more instances. When you launch an instance, you can specify one or more security groups; otherwise, AWS uses the default security group. You can add rules to each security group that allows traffic to or from its associated instances. You can modify the rules for a security group at any time; the new rules are automatically applied to all instances that are associated with the security group. To decide whether to allow traffic to reach an instance, AWS evaluates all the rules from all the security groups that are associated with the instance.

The following are the default rules for a default security group:

Allow inbound traffic from network interfaces (and their associated instances) that are assigned to the same security group.

Allows all outbound traffic

So instance X can be pinged from other instances in the default security group.

The following are the default rules for a security group that you create:

Allows no inbound traffic

Allows all outbound traffic

So instance Y cannot be pinged from other instances in the new security group created by the DevOps team because any new security group allows no inbound traffic by default.

Please note that once you've created a security group, you can change its inbound rules to reflect the type of inbound traffic that you want to reach the associated instances. You can also change its outbound rules.

Default security groups

Your AWS account automatically has a *default security group* for the default VPC in each Region. If you don't specify a security group when you launch an instance, the instance is automatically associated with the default security group for the VPC.

A default security group is named `default`, and it has an ID assigned by AWS. The following are the default rules for each default security group:

- Allows all inbound traffic from other instances associated with the default security group. The security group specifies itself as a source security group in its inbound rules.
- Allows all outbound traffic from the instance.

You can add or remove inbound and outbound rules for any default security group.

You can't delete a default security group. If you try to delete a default security group, you see the following error: `Client.CannotDelete: the specified group: "sg-51530134" name: "default" cannot be deleted by a user.`

Custom security groups

If you don't want your instances to use the default security group, you can create your own security groups and specify them when you launch your instances. You can create multiple security groups to reflect the different roles that your instances play; for example, a web server or a database server.

When you create a security group, you must provide it with a name and a description. Security group names and descriptions can be up to 255 characters in length, and are limited to the following characters:

a-z, A-Z, 0-9, spaces, and `._:/()#@[]+=;&;!$*`

A security group name cannot start with `sg-`. A security group name must be unique for the VPC.

The following are the default rules for a security group that you create:

- Allows no inbound traffic
- Allows all outbound traffic

After you've created a security group, you can change its inbound rules to reflect the type of inbound traffic that you want to reach the associated instances. You can also change its outbound rules.

via -

<https://docs.aws.amazon.com/AWSEC2/latest/UserGuide/ec2-security-groups.html>

Incorrect options:

Instance X is in the default security group. The default rules for the default security group allow inbound traffic from all sources. Instance Y is in a new security group. The default rules for a security group that you create allow no inbound traffic - The default security group allows inbound traffic only from network interfaces (and their associated instances) that are assigned to the same security group. The default security group does not allow inbound traffic from all sources. So this option is incorrect.

Instance X is in the default security group. The default rules for the default security group allow no inbound traffic from network interfaces (and their associated instances) that are assigned to the same security group. Instance Y is in a new security group. The default rules for a security group that you create allow inbound traffic from all sources - The default security group allows inbound traffic from network interfaces (and their associated instances) that are assigned to the same security group. So this option is incorrect.

Instance X is in the default security group. The default rules for the default security group allow no inbound traffic from all sources. Instance Y is in a new security group. The default rules for a security group that you create allow inbound traffic from all sources - The default security group allows inbound traffic from network interfaces (and their associated instances) that are assigned to the same security group. It's wrong to say that the default security group allows no inbound traffic from all sources. So this option is incorrect.

References:

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

<https://docs.aws.amazon.com/AWSEC2/latest/UserGuide/ec2-security-groups.html>
