AWS Certified Solutions Architect Associate Practice Test 5 - Results

Return to review
Chart
Pie chart with 4 slices.
End of interactive chart.
Attempt 1
All knowledge areas
All questions
Question 1: **Incorrect**
A web application is hosted on an EC2 instance that processes sensitive financial information which is launched in a private subnet. All of the data are stored in an Amazon S3 bucket. Financial information is accessed by users over the Internet. The security team of the company is concerned that the Internet connectivity to Amazon S3 is a security risk.

In this scenario, what will you do to resolve this security vulnerability in the most cost-effective manner?

- **Change the web architecture to access the financial data in your S3 bucket through a VPN connection.**

- **Change the web architecture to access the financial data in S3 through an interface VPC endpoint, which is powered by AWS PrivateLink.**

- **Change the web architecture to access the financial data through a Gateway VPC Endpoint.**

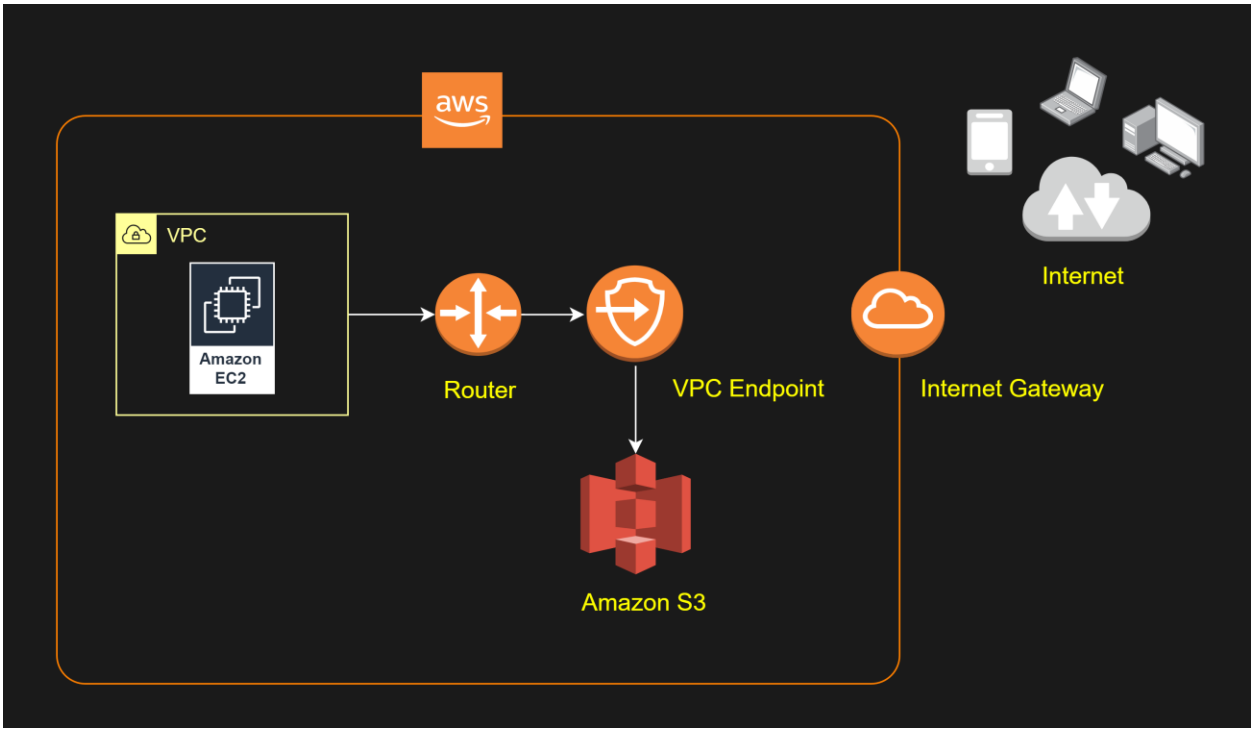
  (Correct)

- **Change the web architecture to access the financial data hosted in your S3 bucket by creating a custom VPC endpoint service.**

  (Incorrect)

Explanation
Take note that your VPC lives within a larger AWS network and the services, such as S3, DynamoDB, RDS, and many others, are located outside of your VPC, but still within the AWS network. By default, the connection that your VPC uses to connect to your S3 bucket or any other service traverses the public Internet via your Internet Gateway.

A **VPC endpoint** enables you to privately connect your VPC to supported AWS services and VPC endpoint services powered by 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.



There are two types of VPC endpoints: *interface endpoints* and *gateway endpoints*. You have to create the type of VPC endpoint required by the supported service.

An **interface endpoint** is an elastic network interface with a private IP address that serves as an entry point for traffic destined to a supported service. A **gateway endpoint** is a gateway that is a target for a specified route in your route table, used for traffic destined to a supported AWS service.

Gateway endpoints for Amazon S3	Interface endpoints for Amazon S3
In both cases, your network traffic remains on the AWS network.	
Use Amazon S3 public IP addresses	Use private IP addresses from your VPC to access Amazon S3
Does not allow access from on premises	Allow access from on premises
Does not allow access from another AWS Region	Allow access from a VPC in another AWS Region using VPC peering or AWS Transit Gateway
Not billed	Billed

Hence, the correct answer is: **Change the web architecture to access the financial data through a Gateway VPC Endpoint.**

The option that says: **Changing the web architecture to access the financial data in your S3 bucket through a VPN connection** is incorrect because a VPN connection still goes through the public Internet. You have to use a VPC Endpoint in this scenario and not VPN, to privately connect your VPC to supported AWS services such as S3.

The option that says: **Changing the web architecture to access the financial data hosted in your S3 bucket by creating a custom VPC endpoint service** is incorrect because a "VPC endpoint **service**" is quite different from a "VPC endpoint". With the VPC endpoint service, you are the service provider where you can create your own application in your VPC and configure it as an AWS PrivateLink-powered service (referred to as an endpoint service). Other AWS principals can create a connection from their VPC to your endpoint service using an interface VPC endpoint.

The option that says: **Changing the web architecture to access the financial data in S3 through an interface VPC endpoint, which is powered by AWS PrivateLink** is incorrect. Although you can use an Interface VPC Endpoint to satisfy the requirement, this type entails an associated cost, unlike a Gateway VPC Endpoint. Remember that you won't get billed if you use a Gateway VPC endpoint for your Amazon S3 bucket, unlike an Interface VPC endpoint that is billed for hourly usage and data processing charges. Take note that the scenario explicitly asks for the most cost-effective solution.

References:

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

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

Check out this Amazon VPC Cheat Sheet:

https://tutorialsdojo.com/amazon-vpc/

Question 2: **Correct**
A logistics company based in the USA runs its web application on a fleet of Amazon EC2 instances in an Auto Scaling group. It runs the same application in multiple AWS regions to cater to clients across several countries. A recent government policy has been enacted that prohibits the company from servicing a specific country.

Which of the following options is the recommended action to comply with the government requirement?

- **Update the route tables to forward all outbound traffic to AWS Network Firewall and configure a stateful domain list rule group to block the specified country**

- **Update the Network Access Control Lists of all subnets used by the Amazon EC2 instances to "deny" all IP addresses from the specific country.**

- **Update the Network Access Control Lists of all subnets used by the Application Load Balancers to "deny" all IP addresses from the specific country.**

- **Create a Web ACL rule in AWS WAF to block the specified country. Associate the rule to the Application Load Balancers.**

  (Correct)

Explanation
AWS WAF is a web application firewall that lets you monitor the HTTP(S) requests that are forwarded to an Amazon CloudFront distribution, an Amazon API Gateway REST API, an Application Load Balancer, or an AWS AppSync GraphQL API.

You use a **web access control list (ACL)** to protect a set of AWS resources. You create a web ACL and define its protection strategy by adding rules. Rules define criteria for inspecting web requests and specify how to handle requests that match the criteria. A web access control list (web ACL) gives you fine-grained control over all of the HTTP(S) web requests that your protected resource responds to.

You can use criteria like the following to allow or block requests:

-IP address origin of the request

-Country of origin of the request

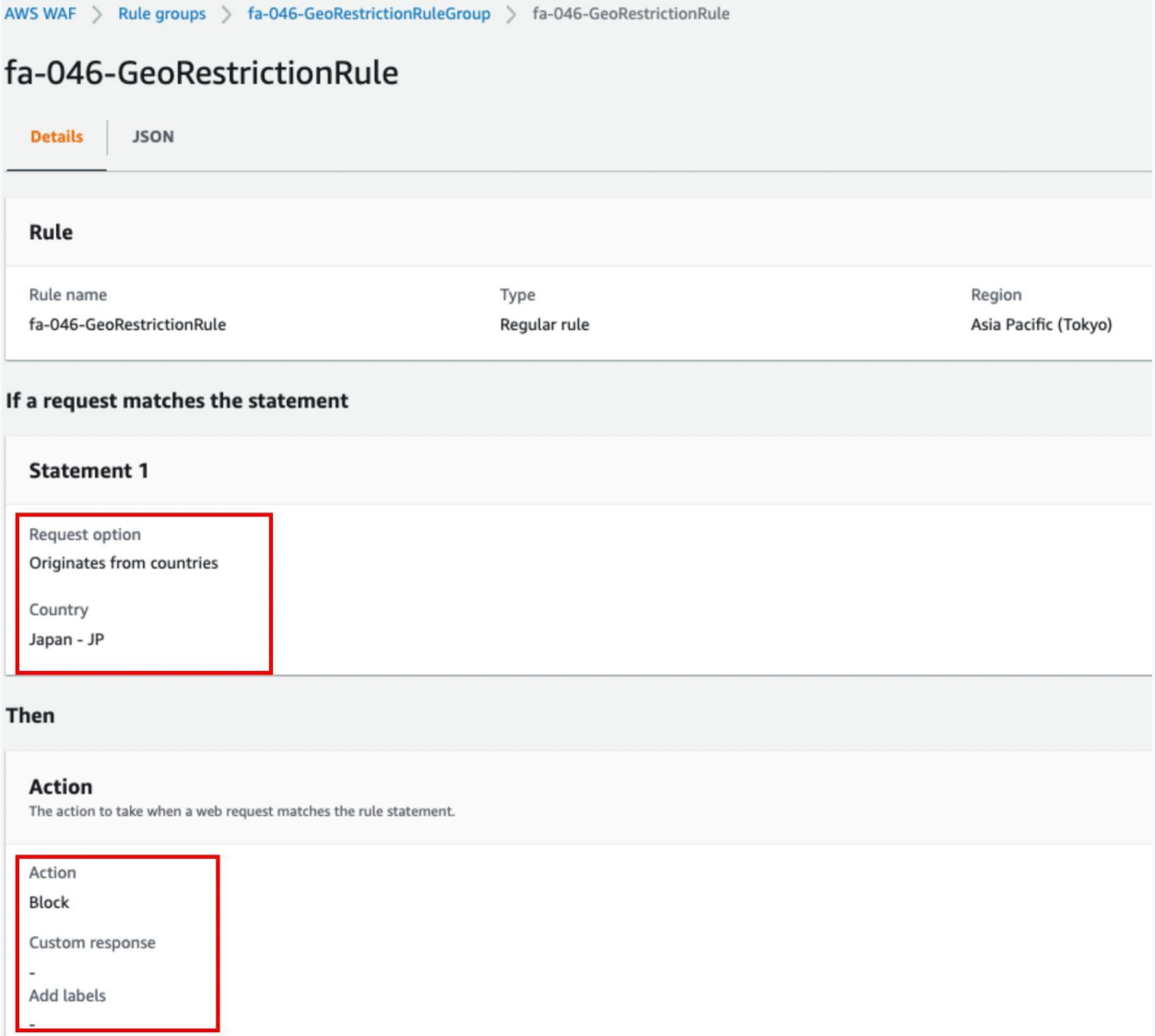
-String match or regular expression (regex) match in a part of the request

-Size of a particular part of the request

-Detection of malicious SQL code or scripting

You can also test for any combination of these conditions. You can block or count web requests that not only meet the specified conditions but also exceed a specified number of requests in any 5-minute period. You can combine conditions using logical operators. You can also run CAPTCHA controls against requests.

To allow or block web requests based on country of origin, create one or more geographical, or geo, match statements. You can use this to block access to your site from specific countries or to only allow access from specific countries.



Therefore, the correct answer is: **Create a Web ACL rule in AWS WAF to block the specified country. Associate this rule to the Application Load Balancers.** In AWS WAF, you can use the Geographic match rule statement to block access to your site from specific countries or to allow access only from specific countries.

The option that says: **Update the route tables to forward all outbound traffic to AWS Network Firewall and configure a stateful domain list rule group to block the specified country** is incorrect. Domain List Rules block HTTP or HTTPS traffic to domains identified as low-reputation, or that are known or suspected to be associated with malware or botnets. This can't be used to block a particular country.

The option that says: **Update the Network Access Control Lists of all subnets used by the Amazon EC2 instances to "deny" all IP addresses from the specific country** is

incorrect. The EC2 instances are behind the Application Load Balancers, thus, if you need to modify NACL for incoming requests you should modify the ALB subnet NACLs.

The option that says: **Update the Network Access Control Lists of all subnets used by the Application Load Balancers to "deny" all IP addresses from the specific country** is incorrect. This may be possible, however, this is not recommended as there will be a lot of IP addresses assigned to a country. The IP addresses list may also change regularly which is difficult to track.

References:

https://aws.amazon.com/premiumsupport/knowledge-center/waf-allow-block-country-geolocation/

https://docs.aws.amazon.com/waf/latest/developerguide/waf-rule-statement-type-geo-match.html

https://docs.aws.amazon.com/waf/latest/developerguide/web-acl.html

Check out this AWS WAF Sheet:

https://tutorialsdojo.com/aws-waf/

Question 3: **Correct**
A top IT Consultancy has a VPC with two On-Demand EC2 instances with Elastic IP addresses. You were notified that the EC2 instances are currently under SSH brute force attacks over the Internet. The IT Security team has identified the IP addresses where these attacks originated. You have to immediately implement a temporary fix to stop these attacks while the team is setting up AWS WAF, GuardDuty, and AWS Shield Advanced to permanently fix the security vulnerability.

Which of the following provides the quickest way to stop the attacks to the instances?

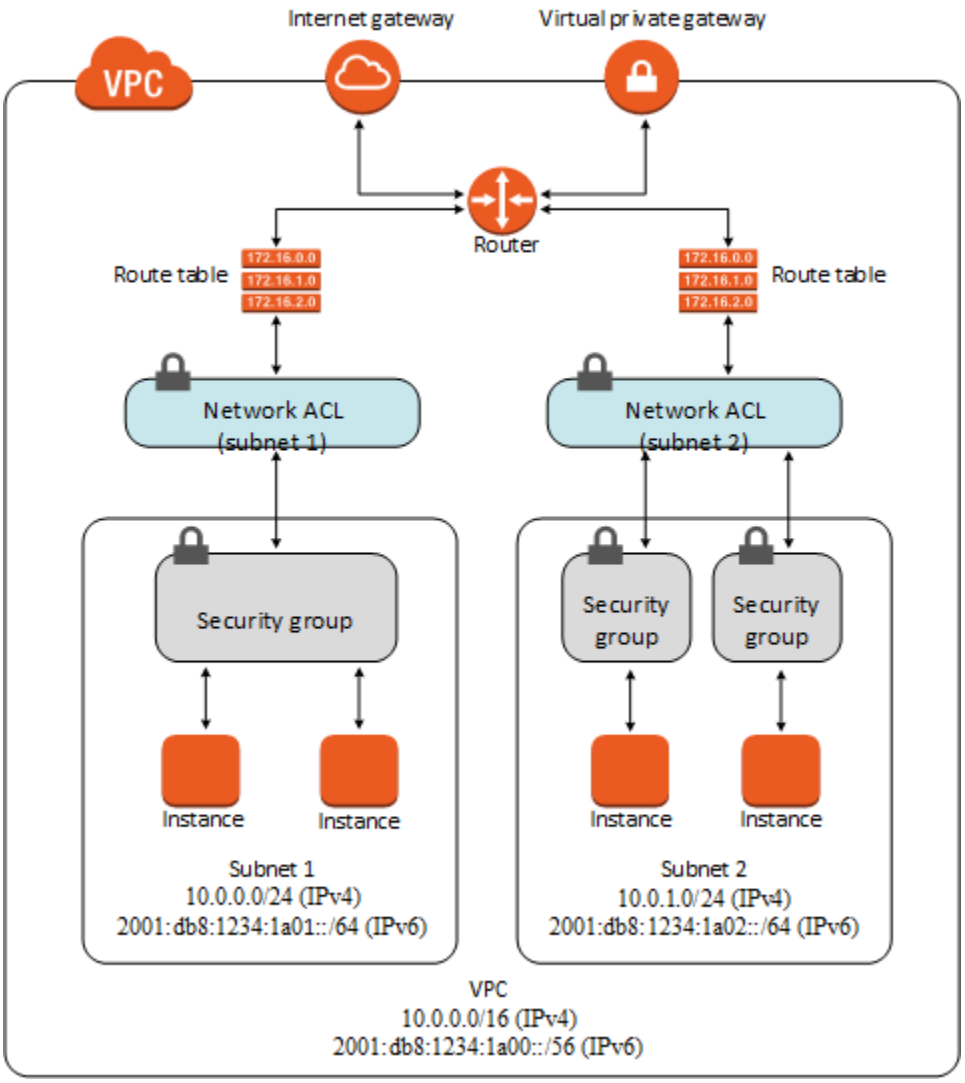
- **Assign a static Anycast IP address to each EC2 instance**
- **Place the EC2 instances into private subnets**

- **Block the IP addresses in the Network Access Control List**

  (Correct)

- **Remove the Internet Gateway from the VPC**

Explanation
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. You might set up network ACLs with rules similar to your security groups in order to add an additional layer of security to your VPC.



The following are the basic things that you need to know about network ACLs:

- Your VPC automatically comes with a modifiable default network ACL. By default, it allows all inbound and outbound IPv4 traffic and, if applicable, IPv6 traffic.

- You can create a custom network ACL and associate it with a subnet. By default, each custom network ACL denies all inbound and outbound traffic until you add rules.

- Each subnet in your VPC must be associated with a network ACL. If you don't explicitly associate a subnet with a network ACL, the subnet is automatically associated with the default network ACL.

- You can associate a network ACL with multiple subnets; however, a subnet can be associated with only one network ACL at a time. When you associate a network ACL with a subnet, the previous association is removed.

- A network ACL contains a numbered list of rules that we evaluate in order, starting with the lowest numbered rule, to determine whether traffic is allowed in or out of any subnet associated with the network ACL. The highest number that you can use for a rule is 32766. We recommend that you start by creating rules in increments (for example, increments of 10 or 100) so that you can insert new rules where you need to later on.

- A network ACL has separate inbound and outbound rules, and each rule can either allow or deny traffic.

- Network ACLs are stateless; responses to allowed inbound traffic are subject to the rules for outbound traffic (and vice versa).

The scenario clearly states that it requires the **quickest** way to fix the security vulnerability. In this situation, you can manually block the offending IP addresses using Network ACLs since the IT Security team has already identified the list of offending IP addresses. Alternatively, you can set up a bastion host, however, this option entails additional time to properly set up as you have to configure the security configurations of your bastion host.

Hence, **blocking the IP addresses in the Network Access Control List** is the best answer since it can quickly resolve the issue by blocking the IP addresses using Network ACL.

Placing the EC2 instances into private subnets is incorrect because if you deploy the EC2 instance in the private subnet without a public or EIP address, it would not be accessible over the Internet, even to you.

Removing the Internet Gateway from the VPC is incorrect because doing this will also make your EC2 instance inaccessible to you as it will cut down the connection to the Internet.

Assigning a static Anycast IP address to each EC2 instance is incorrect because a static Anycast IP address is primarily used by AWS Global Accelerator to enable organizations to route traffic seamlessly to multiple regions and improve availability and performance for their end-users.

References:

https://docs.aws.amazon.com/AmazonVPC/latest/UserGuide/VPC_ACLs.html

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

Security Group vs. NACL:

https://tutorialsdojo.com/security-group-vs-nacl/

Question 4: **Correct**
A company plans to implement a hybrid architecture. They need to create a dedicated connection from their Amazon Virtual Private Cloud (VPC) to their on-premises network. The connection must provide high bandwidth throughput and a more consistent network experience than Internet-based solutions.

Which of the following can be used to create a private connection between the VPC and the company's on-premises network?

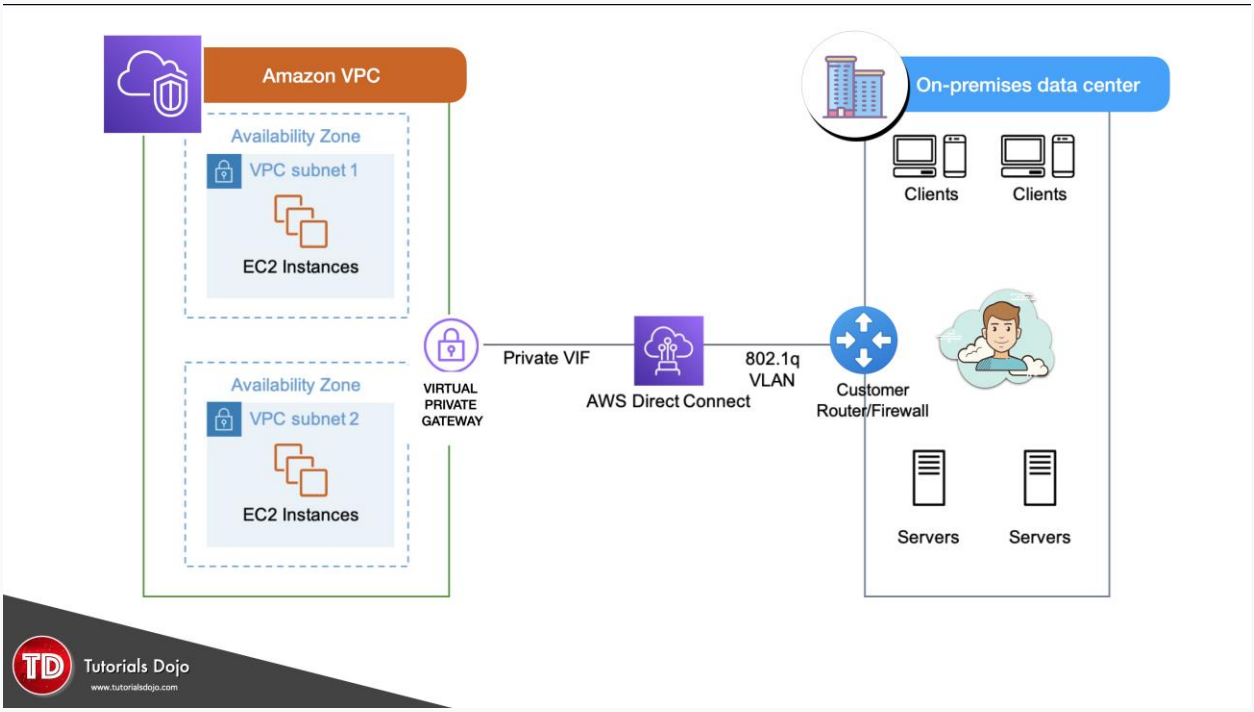
- **AWS Site-to-Site VPN**
- **Transit Gateway with equal-cost multipath routing (ECMP)**
- **Transit VPC**
-

AWS Direct Connect

(Correct)

Explanation

AWS Direct Connect links your internal network 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. You can use a single connection in a public Region or AWS GovCloud (US) to access public AWS services in all other public Regions

Hence, the correct answer is: **AWS Direct Connect.**

The option that says: **Transit VPC** is incorrect because this in itself is not enough to integrate your on-premises network to your VPC. You have to either use a VPN or a Direct Connect connection. A transit VPC is primarily used to connect multiple VPCs and remote networks in order to create a global network transit center and not for establishing a dedicated connection to your on-premises network.

The option that says: **Transit Gateway with equal-cost multipath routing (ECMP)** is incorrect because a transit gateway is commonly used to connect multiple VPCs and

on-premises networks through a central hub. Just like transit VPC, a transit gateway is not capable of establishing a direct and dedicated connection to your on-premises network.

The option that says: **AWS Site-to-Site VPN** is incorrect because this type of connection traverses the public Internet. Moreover, it doesn't provide a high bandwidth throughput and a more consistent network experience than Internet-based solutions.

References:

https://aws.amazon.com/premiumsupport/knowledge-center/connect-vpc/

https://docs.aws.amazon.com/directconnect/latest/UserGuide/Welcome.html

Check out this AWS Direct Connect Cheat Sheet:

https://tutorialsdojo.com/aws-direct-connect/

S3 Transfer Acceleration vs Direct Connect vs VPN vs Snowball vs Snowmobile:

https://tutorialsdojo.com/s3-transfer-acceleration-vs-direct-connect-vs-vpn-vs-snowball-vs-snowmobile/

Comparison of AWS Services Cheat Sheets:

https://tutorialsdojo.com/comparison-of-aws-services/

Question 5: **Correct**
A company has an application hosted in an Amazon ECS Cluster behind an Application Load Balancer. The Solutions Architect is building a sophisticated web filtering solution that allows or blocks web requests based on the country that the requests originate from. However, the solution should still allow specific IP addresses from that country.

Which combination of steps should the Architect implement to satisfy this requirement? (Select TWO.)

- ☑

Add another rule in the AWS WAF web ACL with a geo match condition that blocks requests that originate from a specific country.

(Correct)

- ☐ **In the Application Load Balancer, create a listener rule that explicitly allows requests from approved IP addresses.**

- ☐ **Place a Transit Gateway in front of the VPC where the application is hosted and set up Network ACLs that block requests that originate from a specific country.**

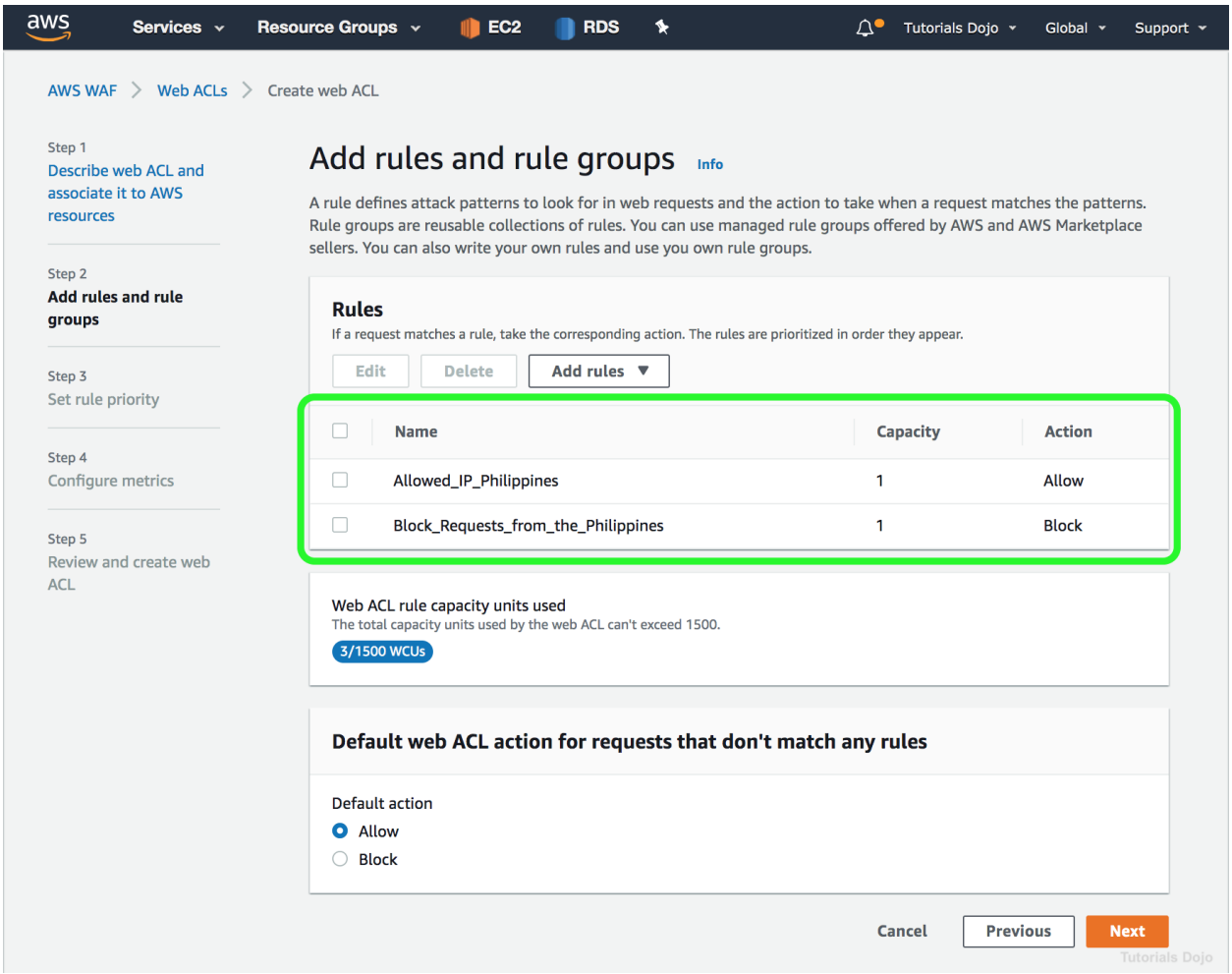
- ☐ **Set up a geo match condition in the Application Load Balancer that blocks requests from a specific country.**

- ☑ **Using AWS WAF, create a web ACL with a rule that explicitly allows requests from approved IP addresses declared in an IP Set.**

(Correct)

Explanation
If you want to allow or block web requests based on the country that the requests originate from, create one or more geo-match conditions. A geo match condition lists countries that your requests originate from. Later in the process, when you create a web ACL, you specify whether to allow or block requests from those countries.

You can use geo-match conditions with other **AWS WAF** Classic conditions or rules to build sophisticated filtering. For example, if you want to block certain countries but still allow specific IP addresses from that country, you could create a rule containing a geo match condition and an IP match condition. Configure the rule to block requests that originate from that country and do not match the approved IP addresses. As another example, if you want to prioritize resources for users in a particular country, you could include a geo-match condition in two different rate-based rules. Set a higher rate limit for users in the preferred country and set a lower rate limit for all other users.



If you are using 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 Classic. So if you want to allow or block requests based on geography plus other AWS WAF Classic conditions, you should *not* use the CloudFront geo restriction feature. Instead, you should use an AWS WAF Classic geo match condition.

Hence, the correct answers are:

- Using AWS WAF, create a web ACL with a rule that explicitly allows requests from approved IP addresses declared in an IP Set.

- Add another rule in the AWS WAF web ACL with a geo match condition that blocks requests that originate from a specific country.

The option that says: **In the Application Load Balancer, create a listener rule that explicitly allows requests from approved IP addresses** is incorrect because a listener rule just checks for connection requests using the protocol and port that you configure. It only determines how the load balancer routes the requests to its registered targets.

The option that says: **Set up a geo match condition in the Application Load Balancer that block requests that originate from a specific country** is incorrect because you can't configure a geo match condition in an Application Load Balancer. You have to use AWS WAF instead.

The option that says: **Place a Transit Gateway in front of the VPC where the application is hosted and set up Network ACLs that block requests that originate from a specific country** is incorrect because AWS Transit Gateway is simply a service that enables customers to connect their Amazon Virtual Private Clouds (VPCs) and their on-premises networks to a single gateway. Using this type of gateway is not warranted in this scenario. Moreover, Network ACLs are not suitable for blocking requests from a specific country. You have to use AWS WAF instead.

References:

https://docs.aws.amazon.com/waf/latest/developerguide/classic-web-acl-geo-conditions.html

https://docs.aws.amazon.com/waf/latest/developerguide/how-aws-waf-works.html

Check out this AWS WAF Cheat Sheet:

https://tutorialsdojo.com/aws-waf/

AWS Security Services Overview - WAF, Shield, CloudHSM, KMS:

https://youtu.be/-1S-RdeAmMo

Question 6: **Correct**
A top investment bank is in the process of building a new Forex trading platform. To ensure high availability and scalability, you designed the trading platform to use an Elastic Load Balancer in front of an Auto Scaling group of On-Demand EC2 instances across multiple Availability Zones. For its database tier, you chose to use a single Amazon Aurora instance to take advantage of its distributed, fault-tolerant, and self-healing storage system.

In the event of system failure on the primary database instance, what happens to Amazon Aurora during the failover?

- **Aurora will first attempt to create a new DB Instance in a different Availability Zone of the original instance. If unable to do so, Aurora will attempt to create a new DB Instance in the original Availability Zone in which the instance was first launched.**

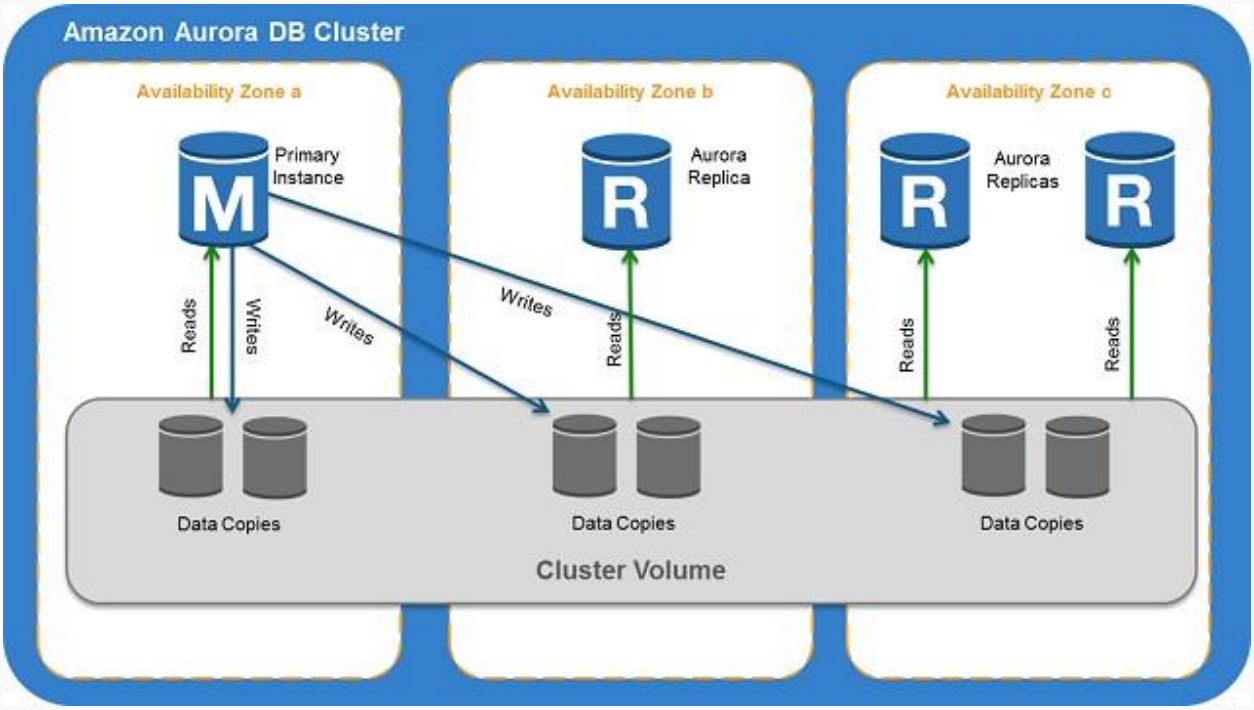
- **Amazon Aurora flips the canonical name record (CNAME) for your DB Instance to point at the healthy replica, which in turn is promoted to become the new primary.**

- **Amazon Aurora flips the A record of your DB Instance to point at the healthy replica, which in turn is promoted to become the new primary.**

- **Aurora will attempt to create a new DB Instance in the same Availability Zone as the original instance and is done on a best-effort basis.**

  (Correct)

Explanation
Failover is automatically handled by Amazon Aurora so that your applications can resume database operations as quickly as possible without manual administrative intervention.



If you have an Amazon Aurora Replica in the same or a different Availability Zone, when failing over, Amazon Aurora flips the canonical name record (CNAME) for your DB Instance to point at the healthy replica, which in turn is promoted to become the new primary. Start-to-finish failover typically completes within 30 seconds.

If you are running Aurora Serverless and the DB instance or AZ becomes unavailable, Aurora will automatically recreate the DB instance in a different AZ.

If you do not have an Amazon Aurora Replica (i.e., single instance) and are not running Aurora Serverless, Aurora will attempt to create a new DB Instance in the same Availability Zone as the original instance. This replacement of the original instance is done on a best-effort basis and may not succeed, for example, if there is an issue that is broadly affecting the Availability Zone.

Hence, the correct answer is the option that says: **Aurora will attempt to create a new DB Instance in the same Availability Zone as the original instance and is done on a best-effort basis.**

The options that say: **Amazon Aurora flips the canonical name record (CNAME) for your DB Instance to point at the healthy replica, which in turn is promoted to become the new primary** and **Amazon Aurora flips the A record of your DB Instance to point at the healthy replica, which in turn is promoted to become the new primary** are incorrect because this will only happen if you are using an Amazon Aurora Replica. In addition, Amazon Aurora flips the canonical name record (CNAME) and not the A record (IP address) of the instance.

The option that says: **Aurora will first attempt to create a new DB Instance in a different Availability Zone of the original instance. If unable to do so, Aurora will attempt to create a new DB Instance in the original Availability Zone in which the instance was first launched** is incorrect because Aurora will first attempt to create a new DB Instance in the same Availability Zone as the original instance. If unable to do so, Aurora will attempt to create a new DB Instance in a different Availability Zone and not the other way around.

References:

https://aws.amazon.com/rds/aurora/faqs/

https://docs.aws.amazon.com/AmazonRDS/latest/AuroraUserGuide/Concepts.AuroraHighAvailability.html

Amazon Aurora Overview:

https://youtu.be/iwS1h7rLNBQ

Check out this Amazon Aurora Cheat Sheet:

https://tutorialsdojo.com/amazon-aurora/

Question 7: **Correct**
A web application hosted in an Auto Scaling group of EC2 instances in AWS. The application receives a burst of traffic every morning, and a lot of users are complaining about request timeouts. The EC2 instance takes 1 minute to boot up before it can respond to user requests. The cloud architecture must be redesigned to better respond to the changing traffic of the application.

How should the Solutions Architect redesign the architecture?

- **Create a CloudFront distribution and set the EC2 instance as the origin.**
- **Create a new launch template and upgrade the size of the instance.**
-

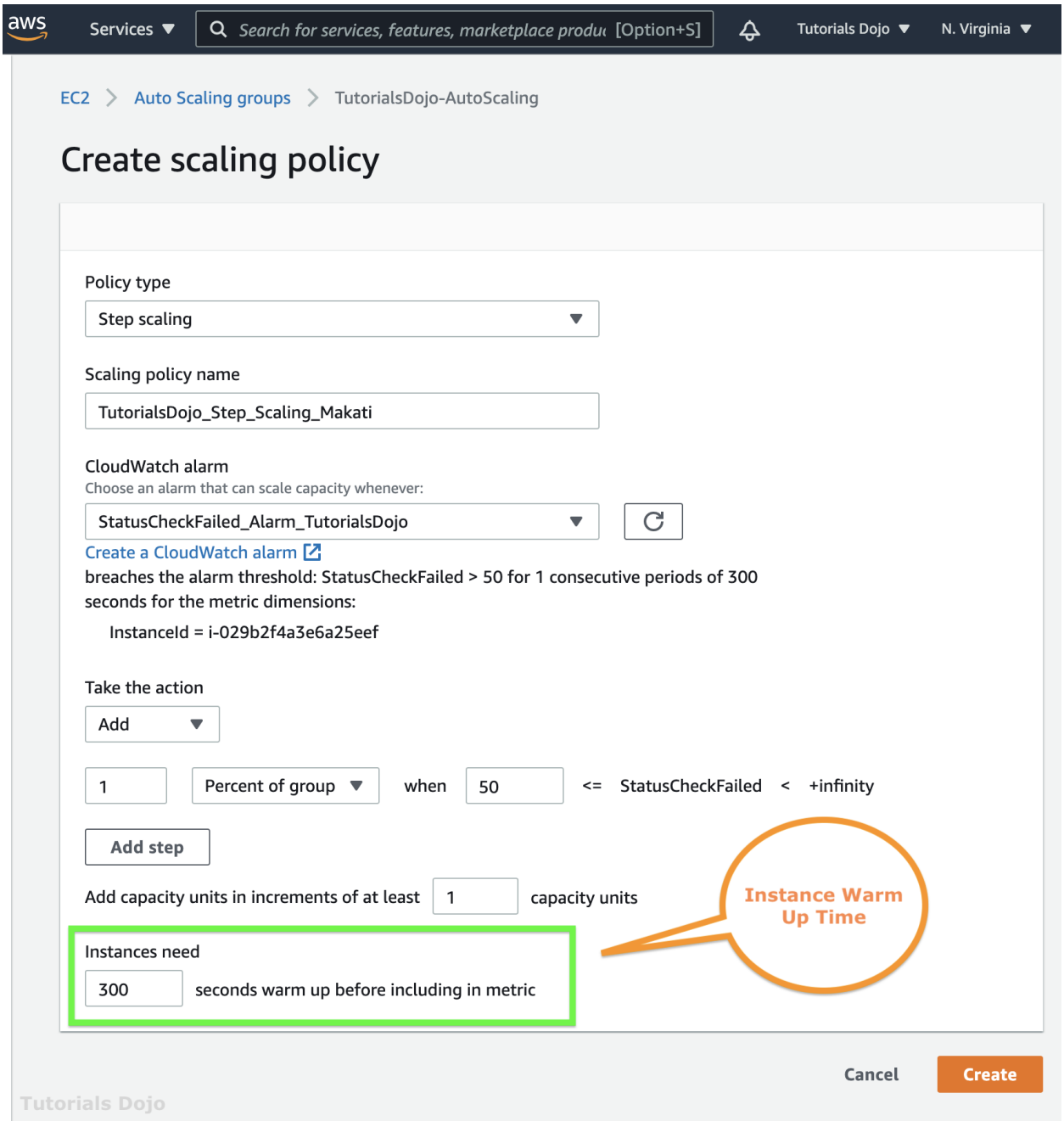
Create a Network Load Balancer with slow-start mode.

- 

  Create a step scaling policy and configure an instance warm-up time condition.

  (Correct)

Explanation
Amazon EC2 Auto Scaling helps you maintain application availability and allows you to automatically add or remove EC2 instances according to conditions you define. You can use the fleet management features of EC2 Auto Scaling to maintain the health and availability of your fleet. You can also use the dynamic and predictive scaling features of EC2 Auto Scaling to add or remove EC2 instances. Dynamic scaling responds to changing demand and predictive scaling automatically schedules the right number of EC2 instances based on predicted demand. Dynamic scaling and predictive scaling can be used together to scale faster.



Step scaling applies “step adjustments” which means you can set multiple actions to vary the scaling depending on the size of the alarm breach. When you create a step scaling policy, you can also specify the number of seconds that it takes for a newly launched instance to warm up.

Hence, the correct answer is: **Create a step scaling policy and configure an instance warm-up time condition.**

The option that says: **Create a Network Load Balancer with slow start mode** is incorrect because Network Load Balancer does not support slow start mode. If you need to enable slow start mode, you should use Application Load Balancer.

The option that says: **Create a new launch template and upgrade the size of the instance** is incorrect because a larger instance does not always improve the boot time. Instead of upgrading the instance, you should create a step scaling policy and add a warm-up time.

The option that says: **Create a CloudFront distribution and set the EC2 instance as the origin** is incorrect because this approach only resolves the traffic latency. Take note that the requirement in the scenario is to resolve the timeout issue and not the traffic latency.

References:

https://docs.aws.amazon.com/autoscaling/ec2/userguide/as-scaling-simple-step.html

https://aws.amazon.com/ec2/autoscaling/faqs/

Check out these AWS Cheat Sheets:

https://tutorialsdojo.com/aws-auto-scaling/

https://tutorialsdojo.com/step-scaling-vs-simple-scaling-policies-in-amazon-ec2/

Question 8: **Correct**
A Solutions Architect is working for a multinational telecommunications company. The IT Manager wants to consolidate their log streams including the access, application, and security logs in one single system. Once consolidated, the company will analyze these logs in real-time based on heuristics. There will be some time in the future where the company will need to validate heuristics, which requires going back to data samples extracted from the last 12 hours.

What is the best approach to meet this requirement?

- **First, set up an Auto Scaling group of EC2 servers then store the logs on Amazon S3 then finally, use EMR to apply heuristics on the logs.**

- **First, configure Amazon Cloud Trail to receive custom logs and then use EMR to apply heuristics on the logs.**

- **First, send all of the log events to Amazon Kinesis then afterwards, develop a client process to apply heuristics on the logs.**

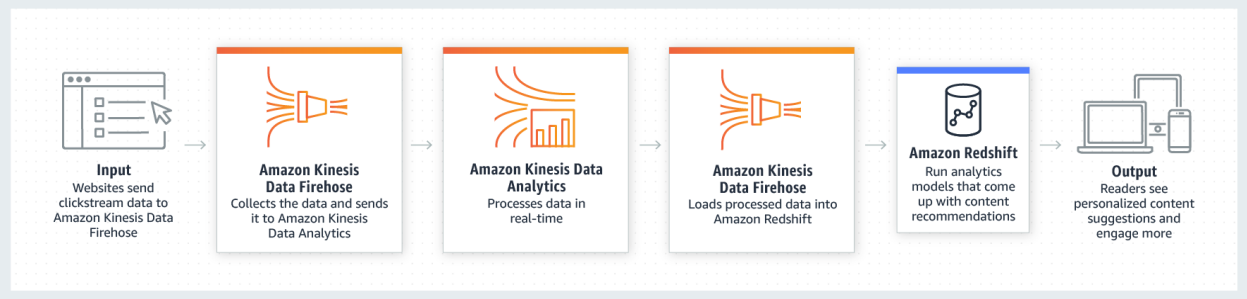
  (Correct)

- **First, send all the log events to Amazon SQS then set up an Auto Scaling group of EC2 servers to consume the logs and finally, apply the heuristics.**

Explanation

In this scenario, you need a service that can collect, process, and analyze data in real-time hence, the right service to use here is **Amazon Kinesis**.

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



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

All other options are incorrect since these services do not have real-time processing capability, unlike Amazon Kinesis.

Reference:

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

Check out this Amazon Kinesis Cheat Sheet:

https://tutorialsdojo.com/amazon-kinesis/

Question 9: **Incorrect**
A multimedia company needs to deploy web services to an AWS region that they have never used before. The company currently has an IAM role for its Amazon EC2 instance that permits the instance to access Amazon DynamoDB. They want their EC2 instances in the new region to have the exact same privileges.

What should be done to accomplish this?

- **Assign the existing IAM role to instances in the new region.**

  (Correct)

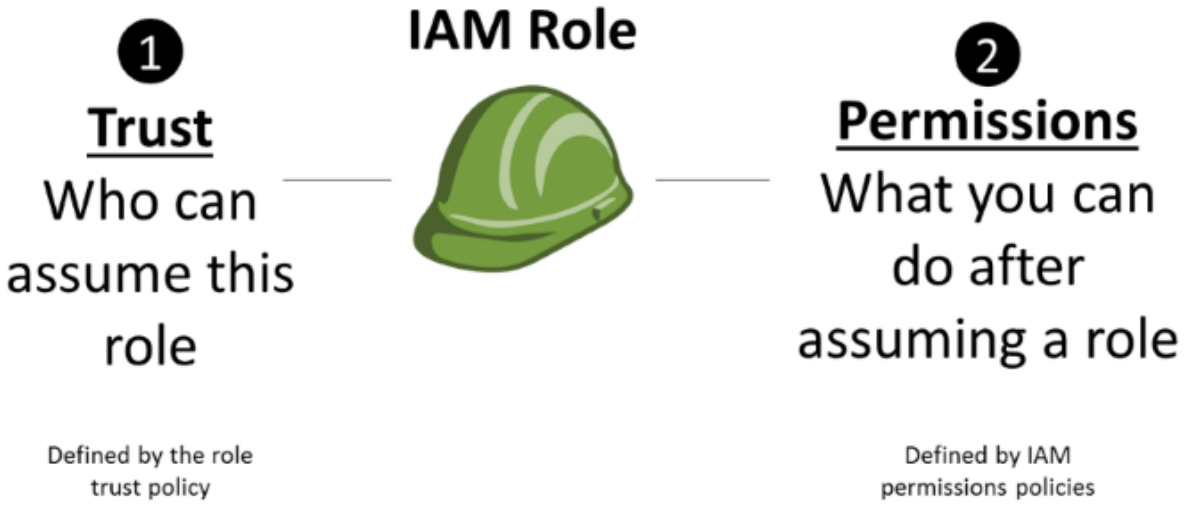
- **Duplicate the IAM role and associated policies to the new region and attach it to the instances.**

- **Create an Amazon Machine Image (AMI) of the instance and copy it to the new region.**

  (Incorrect)

- **In the new Region, create a new IAM role and associated policies then assign it to the new instance.**

Explanation
In this scenario, the company has an existing IAM role hence you don't need to create a new one. IAM roles are global services that are available to all regions hence, all you have to do is assign the existing IAM role to the instance in the new region.



The option that says: **In the new Region, create a new IAM role and associated policies then assign it to the new instance** is incorrect because you don't need to create another IAM role - there is already an existing one.

Duplicating the IAM role and associated policies to the new region and attaching it to the instances is incorrect as you don't need duplicate IAM roles for each region. One IAM role suffices for the instances on two regions.

Creating an Amazon Machine Image (AMI) of the instance and copying it to the new region is incorrect because creating an AMI image does not affect the IAM role of the instance.

Reference:

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

Check out this Amazon S3 Cheat Sheet:

https://tutorialsdojo.com/amazon-s3/

Question 10: **Correct**
A local bank has an in-house application that handles sensitive financial data in a private subnet. After the data is processed by the EC2 worker instances, they will be delivered to S3 for ingestion by other services.

How should you design this solution so that the data does not pass through the public Internet?

- **Provision a NAT gateway in the private subnet with a corresponding route entry that directs the data to S3.**
- **Configure a VPC Endpoint along with a corresponding route entry that directs the data to S3.**

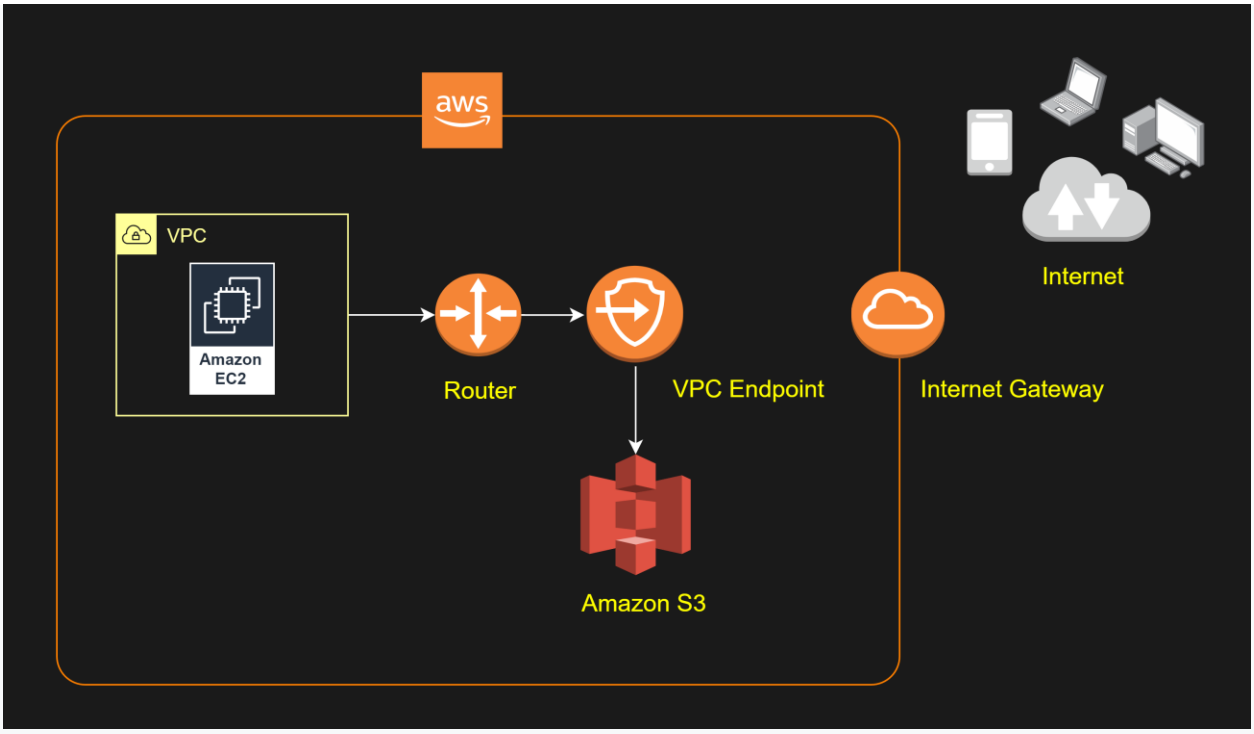
  (Correct)
- **Create an Internet gateway in the public subnet with a corresponding route entry that directs the data to S3.**
- **Configure a Transit gateway along with a corresponding route entry that directs the data to S3.**

Explanation

The important concept that you have to understand in this scenario is that your VPC and your S3 bucket are located within the larger AWS network. However, the traffic coming from your VPC to your S3 bucket is traversing the public Internet by default. To better protect your data in transit, you can set up a VPC endpoint so the incoming traffic from your VPC will not pass through the public Internet, but instead through the private AWS network.

A VPC endpoint enables you to privately connect your VPC to supported AWS services and VPC endpoint services powered by 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 services does not leave the Amazon network.

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



Hence, the correct answer is: **Configure a VPC Endpoint along with a corresponding route entry that directs the data to S3.**

The option that says: **Create an Internet gateway in the public subnet with a corresponding route entry that directs the data to S3** is incorrect because the Internet gateway is used for instances in the public subnet to have accessibility to the Internet.

The option that says: **Configure a Transit gateway along with a corresponding route entry that directs the data to S3** is incorrect because the Transit Gateway is used for interconnecting VPCs and on-premises networks through a central hub. Since Amazon S3 is outside of VPC, you still won't be able to connect to it privately.

The option that says: **Provision a NAT gateway in the private subnet with a corresponding route entry that directs the data to S3** is incorrect because NAT Gateway allows instances in the private subnet to gain access to the Internet, but not vice versa.

References:

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

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

Check out this Amazon VPC Cheat Sheet:

https://tutorialsdojo.com/amazon-vpc/

Question 11: **Correct**
A startup needs to use a shared file system for its .NET web application running on an Amazon EC2 Windows instance. The file system must provide a high level of throughput and IOPS that can also be integrated with Microsoft Active Directory.

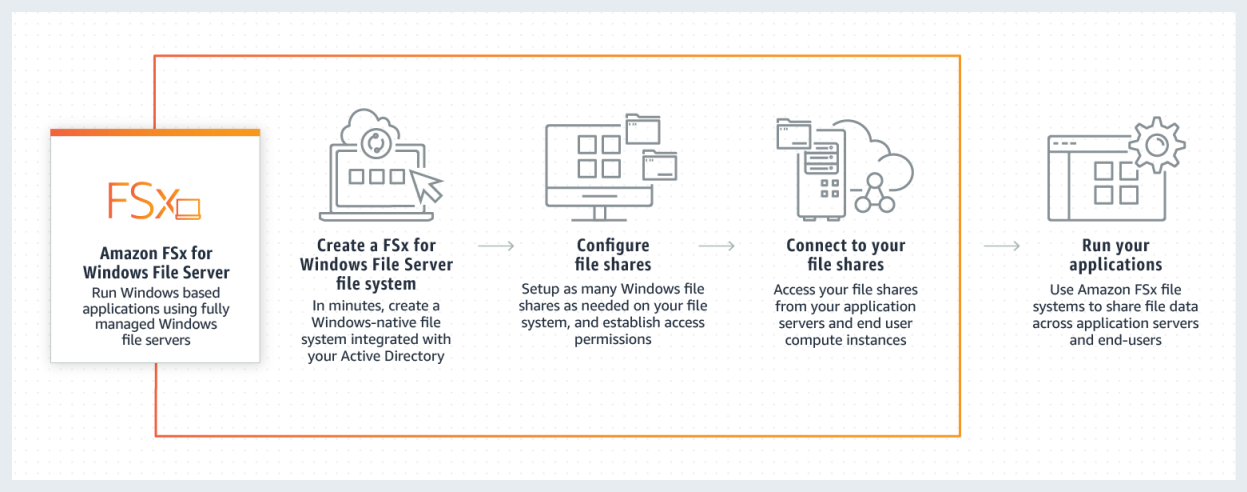
Which is the MOST suitable service that you should use to achieve this requirement?

- **Amazon EBS Provisioned IOPS SSD volumes**
- **Amazon FSx for Windows File Server**

  (Correct)
- **AWS Storage Gateway - File Gateway**
- **Amazon Elastic File System**

Explanation
Amazon FSx for Windows File Server provides fully managed, highly reliable, and scalable file storage accessible over the industry-standard Service Message Block (SMB) protocol. It is built on Windows Server, delivering a wide range of administrative features such as user quotas, end-user file restore, and Microsoft Active Directory (AD) integration.

Amazon FSx supports the use of Microsoft's Distributed File System (DFS) Namespaces to scale-out performance across multiple file systems in the same namespace up to tens of Gbps and millions of IOPS.



The key phrases in this scenario are "file system" and "Active Directory integration." You need to implement a solution that will meet these requirements. Among the options given, the possible answers are FSx Windows File Server and File Gateway. But you need to consider that the question also states that you need to provide a high level of throughput and IOPS. Amazon FSx Windows File Server can scale out storage to hundreds of petabytes of data with tens of GB/s of throughput performance and millions of IOPS.

Hence, the correct answer is: **Amazon FSx for Windows File Server**.

Amazon EBS Provisioned IOPS SSD volumes is incorrect because this is just a block storage volume and not a full-fledged file system. Amazon EBS is primarily used as persistent block storage for EC2 instances.

Amazon Elastic File System is incorrect because it is stated in the scenario that the startup uses an Amazon EC2 Windows instance. Remember that Amazon EFS can only handle Linux workloads.

AWS Storage Gateway - File Gateway is incorrect. Although it can be used as a shared file system for Windows and can also be integrated with Microsoft Active Directory, Amazon FSx still has a higher level of throughput and IOPS compared with AWS Storage Gateway. Amazon FSX is capable of providing hundreds of thousands (or even millions) of IOPS.

References:

https://aws.amazon.com/fsx/windows/faqs/

https://docs.aws.amazon.com/fsx/latest/WindowsGuide/what-is.html

Check out this Amazon FSx Cheat Sheet:

https://tutorialsdojo.com/amazon-fsx/

Question 12: **Correct**
An online shopping platform is hosted on an Auto Scaling group of On-Demand EC2 instances with a default Auto Scaling termination policy and no instance protection configured. The system is deployed across three Availability Zones in the US West region (us-west-1) with an Application Load Balancer in front to provide high availability and fault tolerance for the shopping platform. The us-west-1a, us-west-1b, and us-west-1c Availability Zones have 10, 8 and 7 running instances respectively. Due to the low number of incoming traffic, the scale-in operation has been triggered.

Which of the following will the Auto Scaling group do to determine which instance to terminate first in this scenario? (Select THREE.)

- [x] **Choose the Availability Zone with the most number of instances, which is the us-west-1a Availability Zone in this scenario.**

  (Correct)

- [x] **Select the instance that is closest to the next billing hour.**

  (Correct)

- [ ] **Choose the Availability Zone with the least number of instances, which is the us-west-1c Availability Zone in this scenario.**

- [ ] **Select the instance that is farthest to the next billing hour.**

- [x] **Select the instances with the oldest launch configuration.**

  (Correct)

- ☐

Select the instances with the most recent launch configuration.

Explanation

The default termination policy is designed to help ensure that your network architecture spans Availability Zones evenly. With the default termination policy, the behavior of the Auto Scaling group is as follows:

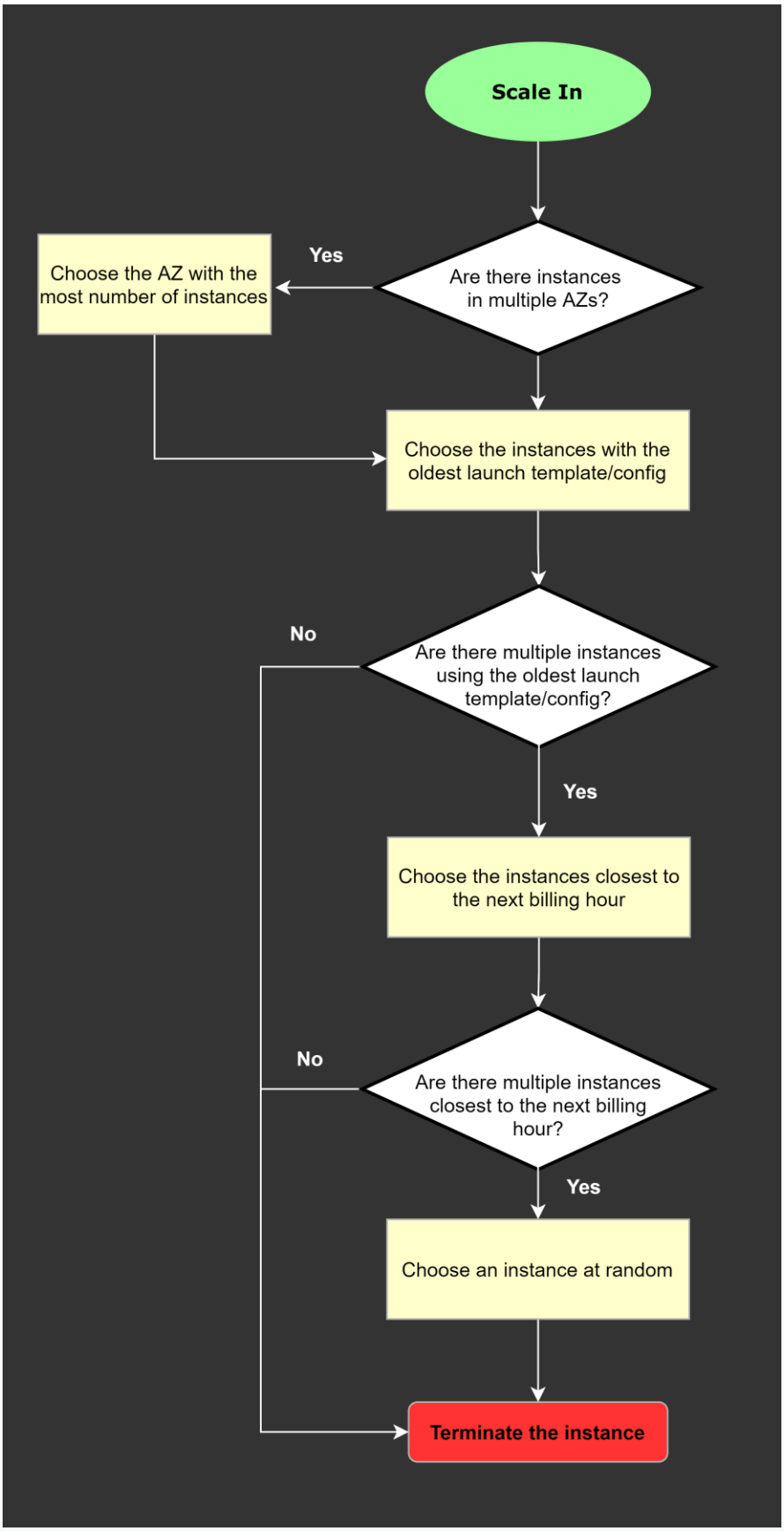
1. If there are instances in multiple Availability Zones, choose the Availability Zone with the most instances and at least one instance that is not protected from scale in. If there is more than one Availability Zone with this number of instances, choose the Availability Zone with the instances that use the oldest launch configuration.

2. Determine which unprotected instances in the selected Availability Zone use the oldest launch configuration. If there is one such instance, terminate it.

3. If there are multiple instances to terminate based on the above criteria, determine which unprotected instances are closest to the next billing hour. (This helps you maximize the use of your EC2 instances and manage your Amazon EC2 usage costs.) If there is one such instance, terminate it.

4. If there is more than one unprotected instance closest to the next billing hour, choose one of these instances at random.

The following flow diagram illustrates how the default termination policy works:


Scale In
Are there instances in multiple AZs?
Yes
Choose the AZ with the most number of instances
Choose the instances with the oldest launch template/config
Are there multiple instances using the oldest launch template/config?
No
Yes
Choose the instances closest to the next billing hour
Are there multiple instances closest to the next billing hour?
No
Yes
Choose an instance at random
Terminate the instance

Reference:

https://docs.aws.amazon.com/autoscaling/ec2/userguide/as-instance-termination.html#default-termination-policy

Check out this AWS Auto Scaling Cheat Sheet:

https://tutorialsdojo.com/aws-auto-scaling/

Question 13: **Correct**
A company has several microservices that send messages to an Amazon SQS queue and a backend application that poll the queue to process the messages. The company also has a Service Level Agreement (SLA) which defines the acceptable amount of time that can elapse from the point when the messages are received until a response is sent. The backend operations are I/O-intensive as the number of messages is constantly growing, causing the company to miss its SLA. The Solutions Architect must implement a new architecture that improves the application's processing time and load management.

Which of the following is the MOST effective solution that can satisfy the given requirement?

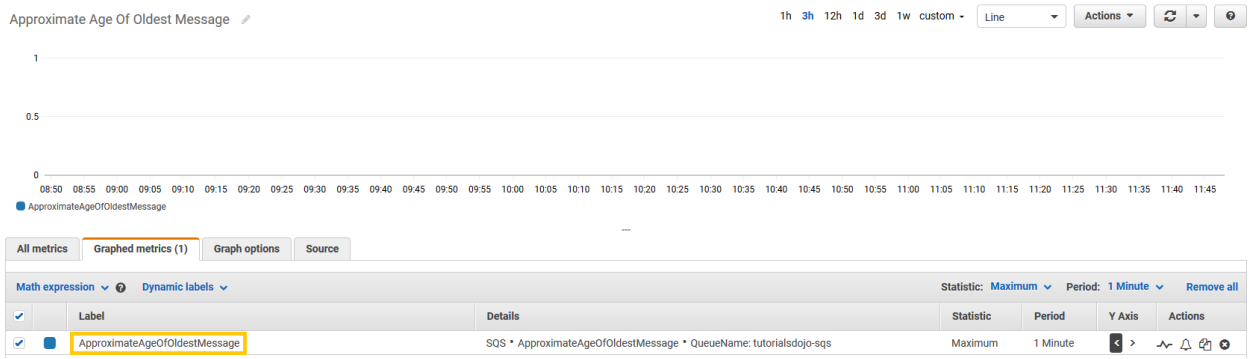
- **Create an AMI of the backend application's EC2 instance. Use the image to set up an Auto Scaling group and configure a target tracking scaling policy based on the `CPUUtilization` metric with a target value of 80%.**
- **Create an AMI of the backend application's EC2 instance and replace it with a larger instance size.**
- **Create an AMI of the backend application's EC2 instance. Use the image to set up an Auto Scaling group and configure a target tracking scaling policy based on the `ApproximateAgeOfOldestMessage` metric.**

  (Correct)
-

Create an AMI of the backend application's EC2 instance and launch it to a cluster placement group.

Explanation

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 eliminates the complexity and overhead associated with managing and operating message-oriented middleware and empowers developers to focus on differentiating work. Using SQS, you can send, store, and receive messages between software components at any volume, without losing messages or requiring other services to be available.



The `ApproximateAgeOfOldestMessage` metric is useful when applications have time-sensitive messages and you need to ensure that messages are processed within a specific time period. You can use this metric to set Amazon CloudWatch alarms that issue alerts when messages remain in the queue for extended periods of time. You can also use alerts to take action, such as increasing the number of consumers to process messages more quickly.

With a target tracking scaling policy, you can scale (increase or decrease capacity) a resource based on a target value for a specific CloudWatch metric. To create a custom metric for this policy, you need to use AWS CLI or AWS SDKs. Take note that you need to create an AMI from the instance first before you can create an Auto Scaling group to scale the instances based on the `ApproximateAgeOfOldestMessage` metric.

Hence, the correct answer is: **Create an AMI of the backend application's EC2 instance. Use the image to set up an Auto Scaling Group and configure a target tracking scaling policy based on the `ApproximateAgeOfOldestMessage` metric.**

The option that says: **Create an AMI of the backend application's EC2 instance. Use the image to set up an Auto Scaling Group and configure a target tracking scaling policy based on the `CPUUtilization` metric with a target value of 80%** is incorrect. Although this will improve the backend processing, the scaling policy based on the `CPUUtilization` metric is not meant for time-sensitive messages where you need to ensure that the messages are processed within a specific time period. It will only trigger

the scale-out activities based on the CPU Utilization of the current instances, and not based on the age of the message, which is a crucial factor in meeting the SLA. To satisfy the requirement in the scenario, you should use the `ApproximateAgeOfOldestMessage` metric.

The option that says: **Create an AMI of the backend application's EC2 instance and replace it with a larger instance size** is incorrect because replacing the instance with a large size won't be enough to dynamically handle workloads at any level. You need to implement an Auto Scaling group to automatically adjust the capacity of your computing resources.

The option that says: **Create an AMI of the backend application's EC2 instance and launch it to a cluster placement group** is incorrect because a cluster placement group is just a logical grouping of EC2 instances. Instead of launching the instance in a placement group, you must set up an Auto Scaling group for your EC2 instances and configure a target tracking scaling policy based on the `ApproximateAgeOfOldestMessage` metric.

References:

https://aws.amazon.com/about-aws/whats-new/2016/08/new-amazon-cloudwatch-metric-for-amazon-sqs-monitors-the-age-of-the-oldest-message/

https://docs.aws.amazon.com/AWSSimpleQueueService/latest/SQSDeveloperGuide/sqs-available-cloudwatch-metrics.html

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

Check out this Amazon SQS Cheat Sheet:

https://tutorialsdojo.com/amazon-sqs/

Question 14: **Correct**
A leading media company has recently adopted a hybrid cloud architecture which requires them to migrate their application servers and databases in AWS. One of their applications requires a heterogeneous database migration in which you need to transform your on-premises Oracle database to PostgreSQL in AWS. This entails a schema and code transformation before the proper data migration starts.

Which of the following options is the most suitable approach to migrate the database in AWS?

- **First, use the AWS Schema Conversion Tool to convert the source schema and application code to match that of the target database, and then use the AWS Database Migration Service to migrate data from the source database to the target database.**

  (Correct)

- **Use Amazon Neptune to convert the source schema and code to match that of the target database in RDS. Use the AWS Batch to effectively migrate the data from the source database to the target database in a batch process.**

- **Configure a Launch Template that automatically converts the source schema and code to match that of the target database. Then, use the AWS Database Migration Service to migrate data from the source database to the target database.**

- **Heterogeneous database migration is not supported in AWS. You have to transform your database first to PostgreSQL and then migrate it to RDS.**

Explanation

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. The AWS Database Migration Service can migrate your data to and from most widely used commercial and open-source databases.

AWS Database Migration Service can migrate your data to and from most of the widely used commercial and open source databases. It supports homogeneous migrations such as Oracle to Oracle, as well as heterogeneous migrations between different database platforms, such as Oracle to Amazon Aurora. Migrations can be from on-premises databases to Amazon RDS or Amazon EC2, databases running on EC2 to RDS, or vice versa, as well as from one RDS database to another RDS database. It can also move data between SQL, NoSQL, and text based targets.

In heterogeneous database migrations the source and target databases engines are different, like in the case of Oracle to Amazon Aurora, Oracle to PostgreSQL, or Microsoft SQL Server to MySQL migrations. In this case, the schema structure, data types, and database code of source and target databases can be quite different,

requiring a schema and code transformation before the data migration starts. That makes heterogeneous migrations a two step process. First use the AWS Schema Conversion Tool to convert the source schema and code to match that of the target database, and then use the AWS Database Migration Service to migrate data from the source database to the target database. All the required data type conversions will automatically be done by the AWS Database Migration Service during the migration. The source database can be located in your own premises outside of AWS, running on an Amazon EC2 instance, or it can be an Amazon RDS database. The target can be a database in Amazon EC2 or Amazon RDS.

The option that says: **Configure a Launch Template that automatically converts the source schema and code to match that of the target database. Then, use the AWS Database Migration Service to migrate data from the source database to the target database** is incorrect because Launch templates are primarily used in EC2 to enable you to store launch parameters so that you do not have to specify them every time you launch an instance.

The option that says: **Use Amazon Neptune to convert the source schema and code to match that of the target database in RDS. Use the AWS Batch to effectively migrate the data from the source database to the target database in a batch process** is incorrect because Amazon Neptune is a fully-managed graph database service and not a suitable service to use to convert the source schema. AWS Batch is not a database migration service and hence, it is not suitable to be used in this scenario. You should use the AWS Schema Conversion Tool and AWS Database Migration Service instead.

The option that says: **Heterogeneous database migration is not supported in AWS. You have to transform your database first to PostgreSQL and then migrate it to RDS** is incorrect because heterogeneous database migration is supported in AWS using the Database Migration Service.

References:

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

https://docs.aws.amazon.com/AWSEC2/latest/UserGuide/ec2-launch-templates.html

https://aws.amazon.com/batch/

Check out this AWS Database Migration Service Cheat Sheet:

https://tutorialsdojo.com/aws-database-migration-service/

AWS Migration Services Overview:

https://www.youtube.com/watch?v=yqNBkFMnsL8

Question 15: **Correct**
A company has an infrastructure that allows EC2 instances from a private subnet to fetch objects from Amazon S3 via a NAT Instance. The company's Solutions Architect was instructed to lower down the cost incurred by the current solution.

How should the Solutions Architect redesign the architecture in the most cost-efficient manner?

- **Use a smaller instance type for the NAT instance.**
- **Replace the NAT instance with NAT Gateway to access S3 objects.**
- **Remove the NAT instance and create an S3 interface endpoint to access S3 objects.**
- **Remove the NAT instance and create an S3 gateway endpoint to access S3 objects.**

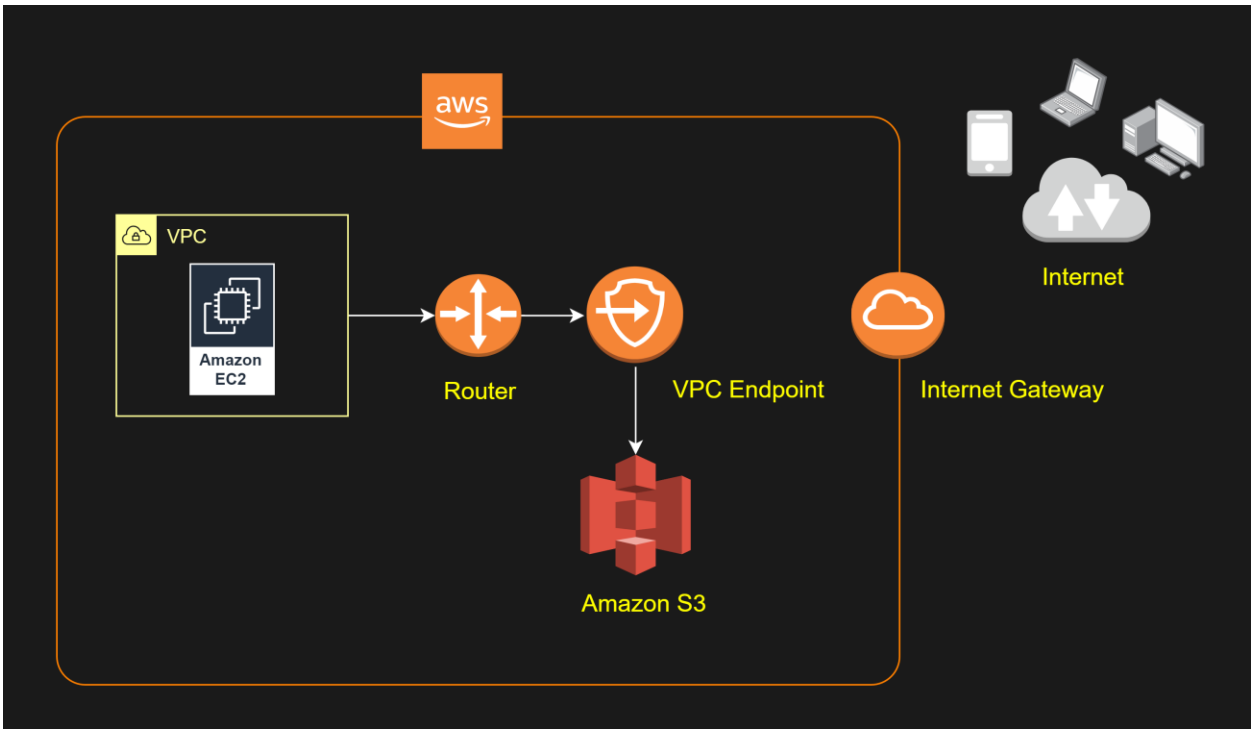
  (Correct)

Explanation
A **VPC endpoint** enables you to privately connect your VPC to supported AWS services and VPC endpoint services powered by 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 services does not leave the Amazon network.

Endpoints are virtual devices. They are horizontally scaled, redundant, and highly available VPC components that allow communication between instances in your VPC

and services without imposing availability risks or bandwidth constraints on your network traffic.

There are two types of VPC endpoints: *interface endpoints* and *gateway endpoints*. You should create the type of VPC endpoint required by the supported service. As a rule of thumb, most AWS services use VPC *Interface* Endpoint except for S3 and DynamoDB, which use VPC *Gateway* Endpoint.



There is no additional charge for using gateway endpoints. However, standard charges for data transfer and resource usage still apply.

Let's assume you created a NAT gateway and you have an EC2 instance routing to the Internet through the NAT gateway. Your EC2 instance behind the NAT gateway sends a 1 GB file to one of your S3 buckets. The EC2 instance, NAT gateway, and S3 Bucket are in the same region US East (Ohio), and the NAT gateway and EC2 instance are in the same availability zone.

Your cost will be calculated as follows:

- NAT Gateway Hourly Charge: NAT Gateway is charged on an hourly basis. For example, the rate is $0.045 per hour in this region.

- NAT Gateway Data Processing Charge: 1 GB data went through NAT gateway. The NAT Gateway Data Processing charge is applied and will result in a charge of $0.045.

- Data Transfer Charge: This is the standard EC2 Data Transfer charge. 1 GB data was transferred from the EC2 instance to S3 via the NAT gateway. There was no charge for the data transfer from the EC2 instance to S3 as it is Data Transfer Out to Amazon EC2 to S3 in the same region. There was also no charge for the data transfer between the NAT Gateway and the EC2 instance since the traffic stays in the same availability zone using private IP addresses. There will be a data transfer charge between your NAT Gateway and EC2 instance if they are in the different availability zone.

In summary, your charge will be $0.045 for 1 GB of data processed by the NAT gateway and a charge of $0.045 per hour will always apply once the NAT gateway is provisioned and available. The data transfer has no charge in this example. However, if you send the file to a non-AWS Internet location instead, there will be a data transfer charge as it is data transfer out from Amazon EC2 to the Internet.

To avoid the NAT Gateway Data Processing charge in this example, you could set up a Gateway Type VPC endpoint and route the traffic to/from S3 through the VPC endpoint instead of going through the NAT Gateway.

There is no data processing or hourly charges for using Gateway Type VPC endpoints.

Hence, the correct answer is the option that says: **Remove the NAT instance and create an S3 gateway endpoint to access S3 objects.**

The option that says: **Replace the NAT instance with NAT Gateway to access S3 objects** is incorrect. A NAT Gateway is just a NAT instance that is managed for you by AWS. It provides less operational management and you pay for the hour that your NAT Gateway is running. This is not the most effective solution since you will still pay for the idle time.

The option that says: **Use a smaller instance type for the NAT instance** is incorrect. Although this might reduce the cost, it still is not the most cost-efficient solution. An S3 Gateway endpoint is still the best solution because it comes with no additional charge.

The option that says: **Remove the NAT instance and create an S3 interface endpoint to access S3 objects** is incorrect. An interface endpoint is an elastic network interface with a private IP address from the IP address range of your subnet. Unlike a Gateway endpoint, you still get billed for the time your interface endpoint is running and the GB data it has processed. From a cost standpoint, using the S3 Gateway endpoint is the most favorable solution.

References:

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

https://aws.amazon.com/blogs/architecture/reduce-cost-and-increase-security-with-amazon-vpc-endpoints/

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

Check out this Amazon S3 Cheat Sheet:

https://tutorialsdojo.com/amazon-s3/

Question 16: **Correct**
A company has a UAT and production EC2 instances running on AWS. They want to ensure that employees who are responsible for the UAT instances don't have access to work on the production instances to minimize security risks.

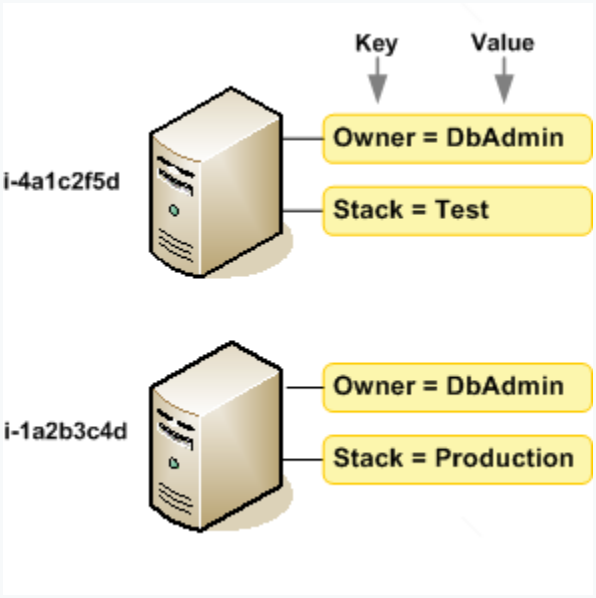
Which of the following would be the best way to achieve this?

- **Launch the UAT and production instances in different Availability Zones and use Multi Factor Authentication.**
- **Provide permissions to the users via the AWS Resource Access Manager (RAM) service to only access EC2 instances that are used for production or development.**
- **Define the tags on the UAT and production servers and add a condition to the IAM policy which allows access to specific tags.**

  (Correct)
- **Launch the UAT and production EC2 instances in separate VPC's connected by VPC peering.**

Explanation
For this scenario, the best way to achieve the required solution is to use a combination of Tags and IAM policies. You can define the tags on the UAT and production EC2 instances and add a condition to the IAM policy which allows access to specific tags.

Tags enable you to categorize your AWS resources in different ways, for example, by purpose, owner, or environment. This is useful when you have many resources of the same type — you can quickly identify a specific resource based on the tags you've assigned to it.



By default, IAM users don't have permission to create or modify Amazon EC2 resources or perform tasks using the Amazon EC2 API. (This means that they also can't do so using the Amazon EC2 console or CLI.) To allow IAM users to create or modify resources and perform tasks, you must create IAM policies that grant IAM users permission to use the specific resources and API actions they'll need and then attach those policies to the IAM users or groups that require those permissions.

Hence, the correct answer is: **Define the tags on the UAT and production servers and add a condition to the IAM policy which allows access to specific tags.**

The option that says: **Launch the UAT and production EC2 instances in separate VPC's connected by VPC peering** is incorrect because these are just network changes to your cloud architecture and don't have any effect on the security permissions of your users to access your EC2 instances.

The option that says: **Provide permissions to the users via the AWS Resource Access Manager (RAM) service to only access EC2 instances that are used for production or development** is incorrect because the AWS Resource Access Manager (RAM) is primarily used to securely share your resources across AWS accounts or within your Organization and not on a single AWS account. You also have to set up a custom IAM Policy in order for this to work.

The option that says: **Launch the UAT and production instances in different Availability Zones and use Multi Factor Authentication** is incorrect because placing the EC2 instances to different AZs will only improve the availability of the systems but won't have any significance in terms of security. You have to set up an IAM Policy that allows access to EC2 instances based on their tags. In addition, a Multi-Factor Authentication is not a suitable security feature to be implemented for this scenario.

References:

http://docs.aws.amazon.com/AWSEC2/latest/UserGuide/Using_Tags.html

https://docs.aws.amazon.com/AWSEC2/latest/UserGuide/iam-policies-for-amazon-ec2.html

Check out this Amazon EC2 Cheat Sheet:

https://tutorialsdojo.com/amazon-elastic-compute-cloud-amazon-ec2/

Question 17: **Correct**
A Solutions Architect joined a large tech company with an existing Amazon VPC. When reviewing the Auto Scaling events, the Architect noticed that their web application is scaling up and down multiple times within the hour.

What design change could the Architect make to optimize cost while preserving elasticity?

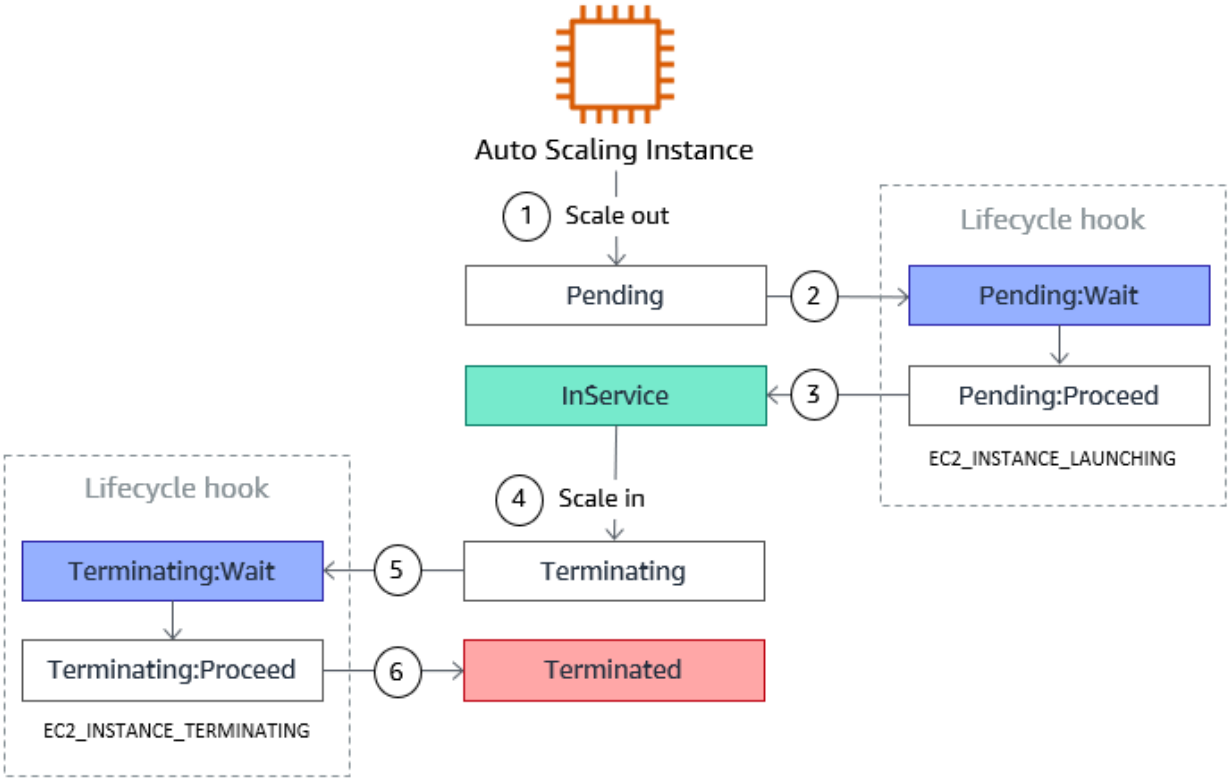
- **Add provisioned IOPS to the instances**
- **Change the cooldown period of the Auto Scaling group and set the CloudWatch metric to a higher threshold**

  (Correct)
- **Increase the instance type in the launch configuration**

-

Increase the base number of Auto Scaling instances for the Auto Scaling group

Explanation
Since the application is scaling up and down multiple times within the hour, the issue lies in the cooldown period of the Auto Scaling group.



The cooldown period is a configurable setting for your Auto Scaling group that helps to ensure that it doesn't launch or terminate additional instances before the previous scaling activity takes effect. After the Auto Scaling group dynamically scales using a simple scaling policy, it waits for the cooldown period to complete before resuming scaling activities.

When you manually scale your Auto Scaling group, the default is not to wait for the cooldown period, but you can override the default and honor the cooldown period. If an instance becomes unhealthy, the Auto Scaling group does not wait for the cooldown period to complete before replacing the unhealthy instance.

Reference:

http://docs.aws.amazon.com/autoscaling/latest/userguide/as-scale-based-on-demand.html

Check out this Amazon EC2 Cheat Sheet:

https://tutorialsdojo.com/amazon-elastic-compute-cloud-amazon-ec2/

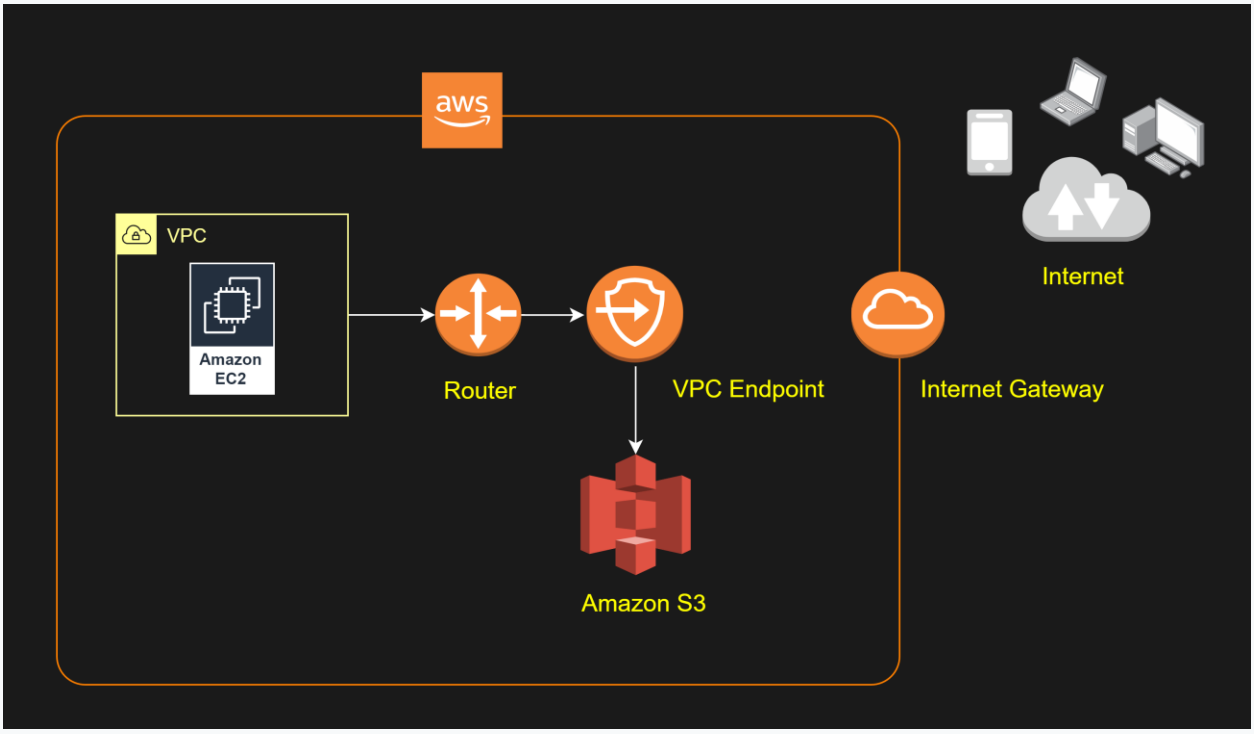
Question 18: **Correct**
A company is running a batch job on an EC2 instance inside a private subnet. The instance gathers input data from an S3 bucket in the same region through a NAT Gateway. The company is looking for a solution that will reduce costs without imposing risks on redundancy or availability.

Which solution will accomplish this?

- **Replace the NAT Gateway with a NAT instance hosted on a burstable instance type.**
- **Deploy a Transit Gateway to peer connection between the instance and the S3 bucket.**
- **Remove the NAT Gateway and use a Gateway VPC endpoint to access the S3 bucket from the instance.**

  (Correct)
- **Re-assign the NAT Gateway to a lower EC2 instance type.**

Explanation
A **gateway endpoint** is a gateway that you specify in your route table to access Amazon S3 from your VPC over the AWS network. Interface endpoints extend the functionality of gateway endpoints by using private IP addresses to route requests to Amazon S3 from within your VPC, on-premises, or from a different AWS Region. Interface endpoints are compatible with gateway endpoints. If you have an existing gateway endpoint in the VPC, you can use both types of endpoints in the same VPC.



There is no additional charge for using gateway endpoints. However, standard charges for data transfer and resource usage still apply.

Hence, the correct answer is: **Remove the NAT Gateway and use a Gateway VPC endpoint to access the S3 bucket from the instance.**

The option that says: **Replace the NAT Gateway with a NAT instance hosted on burstable instance type** is incorrect. This solution may possibly reduce costs, but the availability and redundancy will be compromised.

The option that says: **Deploy a Transit Gateway to peer connection between the instance and the S3 bucket** is incorrect. Transit Gateway is a service that is specifically used for connecting multiple VPCs through a central hub.

The option that says: **Re-assign the NAT Gateway to a lower EC2 instance type** is incorrect. NAT Gateways are fully managed resources. You cannot access nor modify the underlying instance that hosts it.

References:

https://docs.aws.amazon.com/AmazonS3/latest/userguide/privatelink-interface-endpoints.html

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

Amazon VPC Overview:

https://youtu.be/oIDHKeNxvQQ

Check out this Amazon VPC Cheat Sheet:

https://tutorialsdojo.com/amazon-vpc/

Question 19: **Correct**
A financial analytics application that collects, processes and analyzes stock data in real-time is using Kinesis Data Streams. The producers continually push data to Kinesis Data Streams while the consumers process the data in real time. In Amazon Kinesis, where can the consumers store their results? (Select TWO.)

- ☑ **Amazon Redshift**

  (Correct)

- ☐ **AWS Glue**

- ☑ **Amazon S3**

  (Correct)

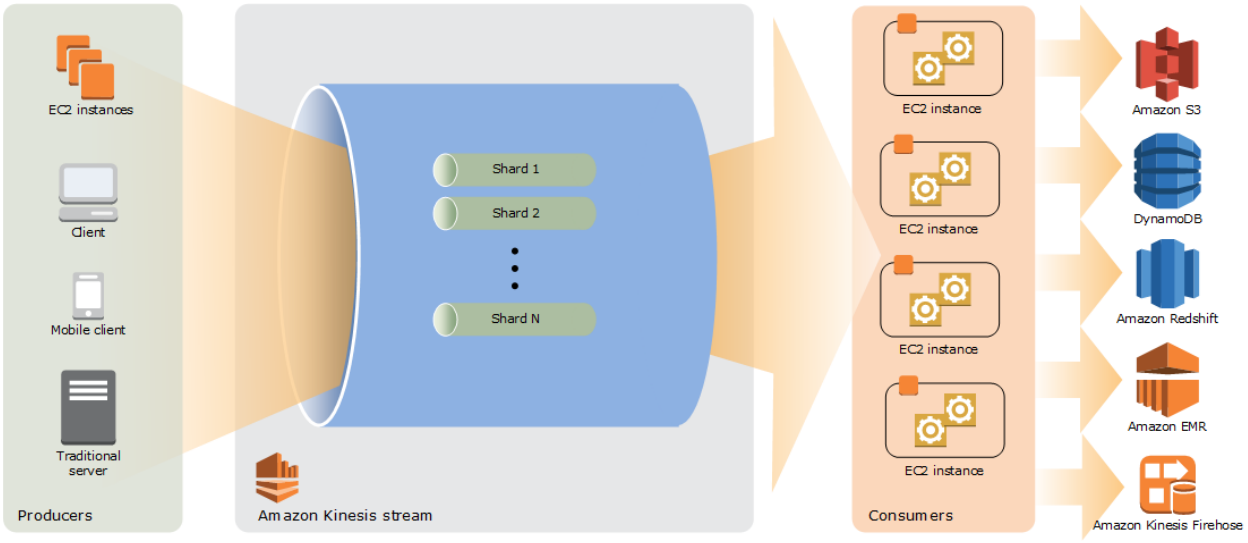
- ☐ **Glacier Select**

- ☐ **Amazon Athena**

Explanation

In Amazon Kinesis, the producers continually push data to Kinesis Data Streams and the consumers process the data in real-time. Consumers (*such as a custom application running on Amazon EC2, or an Amazon Kinesis Data Firehose delivery stream*) can store their results using an AWS service such as Amazon DynamoDB, Amazon Redshift, or Amazon S3.

Hence, **Amazon S3** and **Amazon Redshift** are the correct answers. The following diagram illustrates the high-level architecture of Kinesis Data Streams:



Glacier Select is incorrect because this is not a storage service. It is primarily used to run queries directly on data stored in Amazon Glacier, retrieving only the data you need out of your archives to use for analytics.

AWS Glue is incorrect because this is not a storage service. It is a fully managed extract, transform, and load (ETL) service that makes it easy for customers to prepare and load their data for analytics.

Amazon Athena is incorrect because this is just an interactive query service that makes it easy to analyze data in Amazon S3 using standard SQL. It is not a storage service where you can store the results processed by the consumers.

Reference:

http://docs.aws.amazon.com/streams/latest/dev/key-concepts.html

Amazon Redshift Overview:

https://youtu.be/jlLERNzhHOg

Check out this Amazon Kinesis Cheat Sheet:

https://tutorialsdojo.com/amazon-kinesis/

Question 20: **Correct**
A data analytics startup is collecting clickstream data and stores them in an S3 bucket. You need to launch an AWS Lambda function to trigger the ETL jobs to run as soon as new data becomes available in Amazon S3.

Which of the following services can you use as an extract, transform, and load (ETL) service in this scenario?

- **AWS Glue**

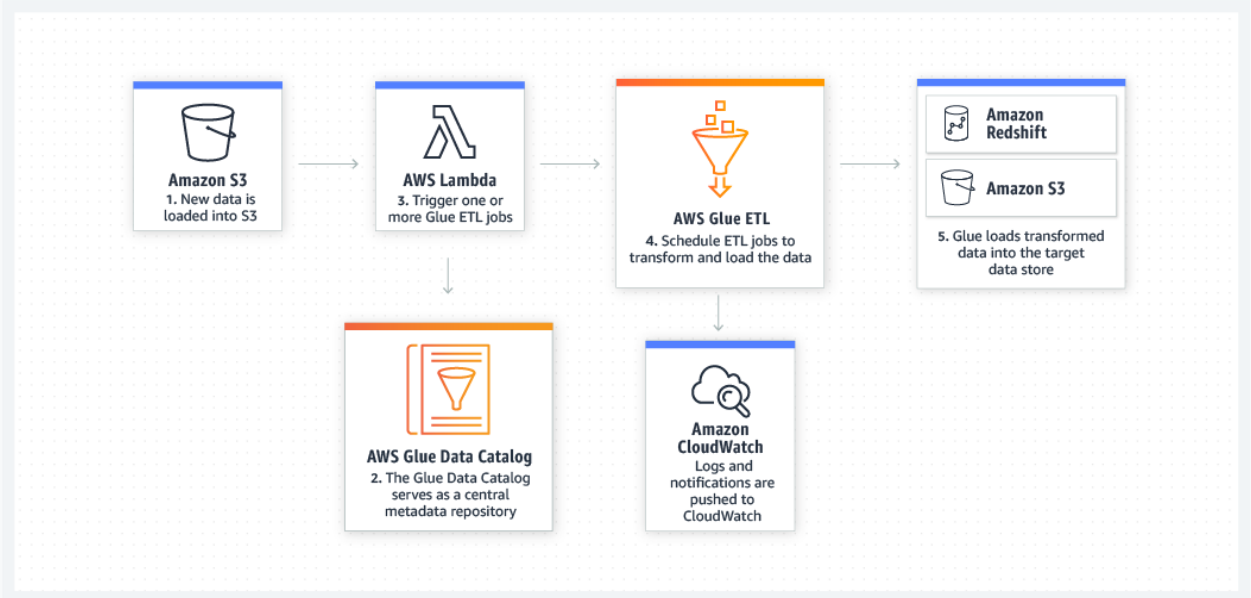
  (Correct)

- **Redshift Spectrum**

- **AWS Step Functions**

- **S3 Select**

Explanation
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 can create and run an ETL job with a few clicks in the AWS Management Console. You simply point AWS Glue to your data stored on AWS, and AWS Glue discovers your data and stores the associated metadata (e.g., table definition and schema) in the AWS Glue Data Catalog. Once cataloged, your data is immediately searchable, queryable, and available for ETL. AWS Glue generates the code to execute your data transformations and data loading processes.



Reference:

https://aws.amazon.com/glue/

Check out this AWS Glue Cheat Sheet:

https://tutorialsdojo.com/aws-glue/

Question 21: **Correct**
An On-Demand EC2 instance is launched into a VPC subnet with the Network ACL configured to allow all inbound traffic and deny all outbound traffic. The instance's security group has an inbound rule to allow SSH from any IP address and does not have any outbound rules.

In this scenario, what are the changes needed to allow SSH connection to the instance?

- **Both the outbound security group and outbound network ACL need to be modified to allow outbound traffic.**
- **The outbound security group needs to be modified to allow outbound traffic.**

- **The network ACL needs to be modified to allow outbound traffic.**

  (Correct)

- **No action needed. It can already be accessed from any IP address using SSH.**

Explanation
In order for you to establish an SSH connection from your home computer to your EC2 instance, you need to do the following:

- On the Security Group, add an Inbound Rule to allow SSH traffic to your EC2 instance.

- On the NACL, add both an Inbound and Outbound Rule to allow SSH traffic to your EC2 instance.

The reason why you have to add both Inbound and Outbound SSH rule is due to the fact that Network ACLs are stateless which means that responses to allow inbound traffic are subject to the rules for outbound traffic (and vice versa). In other words, if you only enabled an Inbound rule in NACL, the traffic can only go in but the SSH response will not go out since there is no Outbound rule.

Security groups are stateful which means that if an incoming request is granted, then the outgoing traffic will be automatically granted as well, regardless of the outbound rules.

References:

https://docs.aws.amazon.com/AmazonVPC/latest/UserGuide/VPC_ACLs.html

https://docs.aws.amazon.com/AWSEC2/latest/UserGuide/authorizing-access-to-an-instance.html

Question 22: **Incorrect**
A healthcare company stores sensitive patient health records in their on-premises storage systems. These records must be kept indefinitely and protected from any type of modifications once they are stored. Compliance regulations mandate that the records must have granular access control and each data access must be audited at all

levels. Currently, there are millions of obsolete records that are not accessed by their web application, and their on-premises storage is quickly running out of space. The Solutions Architect must design a solution to immediately move existing records to AWS and support the ever-growing number of new health records.

Which of the following is the most suitable solution that the Solutions Architect should implement to meet the above requirements?

- **Set up AWS DataSync to move the existing health records from the on-premises network to the AWS Cloud. Launch a new Amazon S3 bucket to store existing and new records. Enable AWS CloudTrail with Management Events and Amazon S3 Object Lock in the bucket.**

- **Set up AWS Storage Gateway to move the existing health records from the on-premises network to the AWS Cloud. Launch an Amazon EBS-backed EC2 instance to store both the existing and new records. Enable Amazon S3 server access logging and S3 Object Lock in the bucket.**

- **Set up AWS DataSync to move the existing health records from the on-premises network to the AWS Cloud. Launch a new Amazon S3 bucket to store existing and new records. Enable AWS CloudTrail with Data Events and Amazon S3 Object Lock in the bucket.**

  (Correct)

- **Set up AWS Storage Gateway to move the existing health records from the on-premises network to the AWS Cloud. Launch a new Amazon S3 bucket to store existing and new records. Enable AWS CloudTrail with Management Events and Amazon S3 Object Lock in the bucket.**

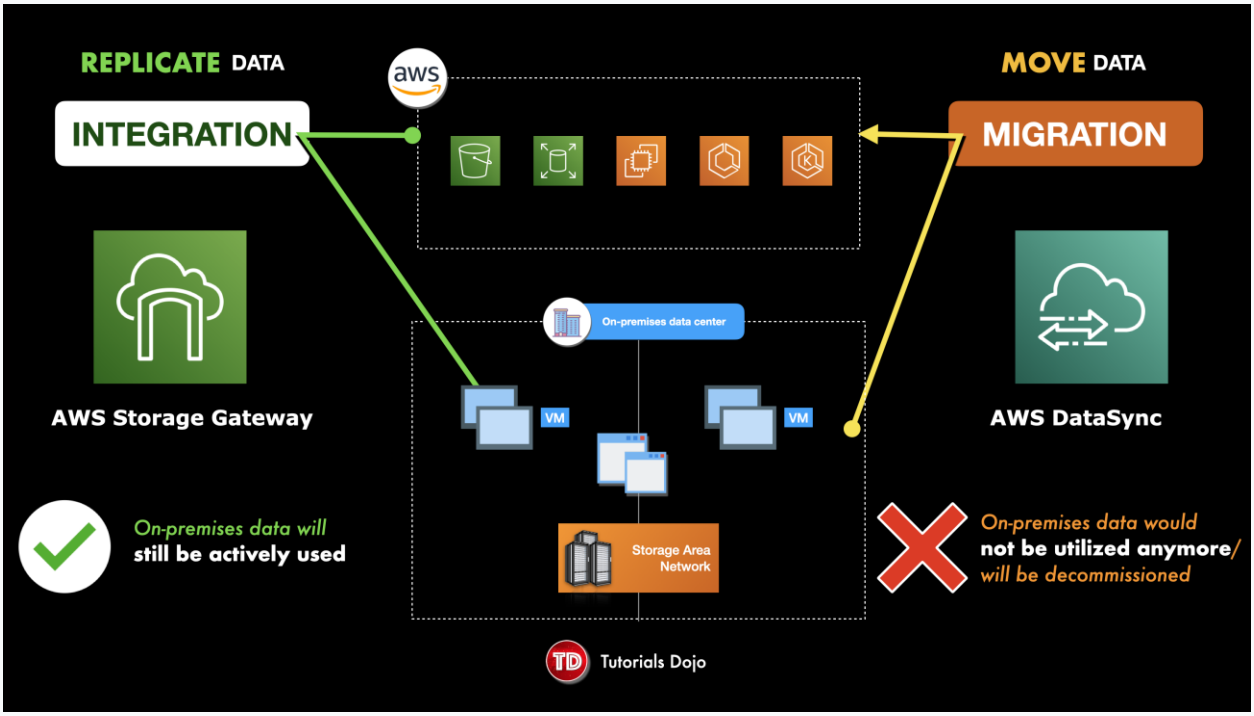
  (Incorrect)

Explanation
AWS Storage Gateway is a set of hybrid cloud services that gives you on-premises access to virtually unlimited cloud storage. Customers use Storage Gateway to **integrate** AWS Cloud storage with existing on-site workloads so they can simplify storage management and reduce costs for key hybrid cloud storage use cases. These

include moving backups to the cloud, using on-premises file shares backed by cloud storage, and providing low latency access to data in AWS for on-premises applications.

AWS DataSync is an online data transfer service that simplifies, automates, and accelerates **moving** data between on-premises storage systems and AWS Storage services, as well as between AWS Storage services. You can use DataSync to **migrate** active datasets 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.

Both AWS Storage Gateway and AWS DataSync can send data from your on-premises data center to AWS and vice versa. However, AWS Storage Gateway is more suitable to be used in integrating your storage services by replicating your data while AWS DataSync is better for workloads that require you to move or migrate your data.



You can also 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 is a fully managed solution that will automate and accelerate the replication of data between the on-premises storage systems and AWS storage services.

AWS CloudTrail is an AWS service that helps you enable governance, compliance, and operational and risk auditing of your AWS account. Actions taken by a user, role, or an

AWS service are recorded as events in CloudTrail. Events include actions taken in the AWS Management Console, AWS Command Line Interface, and AWS SDKs and APIs.

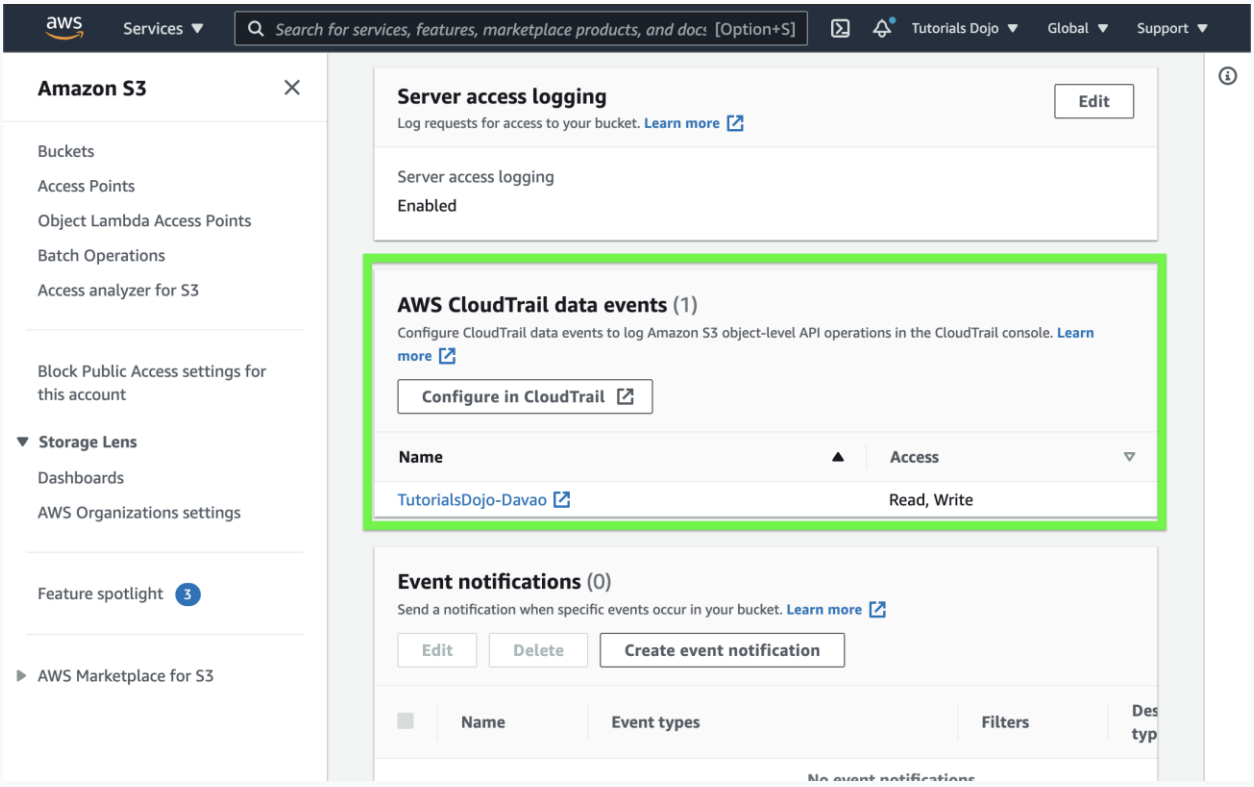
There are two types of events that you configure your CloudTrail for:

- Management Events

- Data Events

Management Events provide visibility into management operations that are performed on resources in your AWS account. These are also known as control plane operations. Management events can also include non-API events that occur in your account.

Data Events, on the other hand, provide visibility into the resource operations performed on or within a resource. These are also known as data plane operations. It allows granular control of data event logging with advanced event selectors. You can currently log data events on different resource types such as Amazon S3 object-level API activity (e.g. GetObject, DeleteObject, and PutObject API operations), AWS Lambda function execution activity (the Invoke API), DynamoDB Item actions, and many more.



With **S3 Object Lock**, you can store objects using a *write-once-read-many* (WORM) model. Object Lock can help prevent objects from being deleted or overwritten for a fixed amount of time or indefinitely. You can use Object Lock to help meet regulatory

requirements that require WORM storage or to simply add another layer of protection against object changes and deletion.

Log properties	AWS CloudTrail	Amazon S3 server logs
Can be forwarded to other systems (CloudWatch Logs, CloudWatch Events)	Yes	
Deliver logs to more than one destination (for example, send the same logs to two different buckets)	Yes	
Turn on logs for a subset of objects (prefix)	Yes	
Cross-account log delivery (target and source bucket owned by different accounts)	Yes	
Integrity validation of log file using digital signature/hashing	Yes	
Default/choice of encryption for log files	Yes	
Object operations (using Amazon S3 APIs)	Yes	Yes
Bucket operations (using Amazon S3 APIs)	Yes	Yes
Searchable UI for logs	Yes	
Fields for Object Lock parameters, Amazon S3 Select properties for log records	Yes	
Fields for `Object Size`, `Total Time`, `Turn-Around Time`, and `HTTP Referer` for log records		Yes
Lifecycle transitions, expirations, restores		Yes
Logging of keys in a batch delete operation		Yes
Authentication failures[1]		Yes
Accounts where logs get delivered	Bucket owner[2], and requester	Bucket owner only

You can record the actions that are taken by users, roles, or AWS services on Amazon S3 resources and maintain log records for auditing and compliance purposes. To do this, you can use server access logging, AWS CloudTrail logging, or a combination of both. AWS recommends that you use AWS CloudTrail for logging bucket and object-level actions for your Amazon S3 resources.

Hence, the correct answer is: **Set up AWS DataSync to move the existing health records from the on-premises network to the AWS Cloud. Launch a new Amazon S3 bucket to store existing and new records. Enable AWS CloudTrail with Data Events and Amazon S3 Object Lock in the bucket.**

The option that says: **Set up AWS Storage Gateway to move the existing health records from the on-premises network to the AWS Cloud. Launch a new Amazon S3 bucket to store existing and new records. Enable AWS CloudTrail with Management Events and Amazon S3 Object Lock in the bucket** is incorrect. The requirement explicitly says that the Solutions Architect must immediately move the existing records to AWS and not integrate or replicate the data. Using AWS DataSync is a more suitable service to use here since the primary objective is to migrate or move data. You also have to use Data Events here and not Management Events in CloudTrail, to properly track all the data access and changes to your objects.

The option that says: **Set up AWS Storage Gateway to move the existing health records from the on-premises network to the AWS Cloud. Launch an Amazon EBS-backed EC2 instance to store both the existing and new records. Enable Amazon S3 server access logging and S3 Object Lock in the bucket** is incorrect. Just as mentioned in the previous option, using AWS Storage Gateway is not a recommended service to use in

this situation since the objective is to move the obsolete data. Moreover, using Amazon EBS to store health records is not a scalable solution compared with Amazon S3. Enabling server access logging can help audit the stored objects. However, it is better to CloudTrail as it provides more granular access control and tracking.

The option that says: **Set up AWS DataSync to move the existing health records from the on-premises network to the AWS Cloud. Launch a new Amazon S3 bucket to store existing and new records. Enable AWS CloudTrail with Management Events and Amazon S3 Object Lock in the bucket** is incorrect. Although it is right to use AWS DataSync to move the health records, you still have to configure Data Events in AWS CloudTrail and not Management Events. This type of event only provides visibility into management operations that are performed on resources in your AWS account and not the data events that are happening in the individual objects in Amazon S3.

References:

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

https://aws.amazon.com/about-aws/whats-new/2020/12/aws-cloudtrail-provides-more-granular-control-of-data-event-logging/

https://docs.aws.amazon.com/AmazonS3/latest/userguide/object-lock.html

Check out this AWS DataSync Cheat Sheet:

https://tutorialsdojo.com/aws-datasync/

AWS Storage Gateway vs DataSync:

https://www.youtube.com/watch?v=tmfe1rO-AUs

Question 23: **Correct**
A financial company wants to store their data in Amazon S3 but at the same time, they want to store their frequently accessed data locally on their on-premises server. This is due to the fact that they do not have the option to extend their on-premises storage, which is why they are looking for a durable and scalable storage service to use in AWS.

What is the best solution for this scenario?

- **Use the Amazon Storage Gateway - Cached Volumes.**

  (Correct)

- **Use a fleet of EC2 instance with EBS volumes to store the commonly used data.**

- **Use both Elasticache and S3 for frequently accessed data.**

- **Use Amazon Glacier.**

Explanation
By using Cached volumes, you store your data in Amazon Simple Storage Service (Amazon S3) and retain a copy of frequently accessed data subsets locally in your on-premises network. Cached volumes offer substantial cost savings on primary storage and minimize the need to scale your storage on-premises. You also retain low-latency access to your frequently accessed data. This is the best solution for this scenario.

Using a fleet of EC2 instance with EBS volumes to store the commonly used data is incorrect because an EC2 instance is not a storage service and it does not provide the required durability and scalability.

Using both Elasticache and S3 for frequently accessed data is incorrect as this is not efficient. Moreover, the question explicitly said that the frequently accessed data should be stored locally on their on-premises server and not on AWS.

Using Amazon Glacier is incorrect as this is mainly used for data archiving.

Reference:

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

Check out this AWS Storage Gateway Cheat Sheet:

https://tutorialsdojo.com/aws-storage-gateway/

Question 24: **Correct**
An investment bank has a distributed batch processing application which is hosted in an Auto Scaling group of Spot EC2 instances with an SQS queue. You configured your components to use client-side buffering so that the calls made from the client will be buffered first and then sent as a batch request to SQS. What is a period of time during which the SQS queue prevents other consuming components from receiving and processing a message?

- **Processing Timeout**
- **Receiving Timeout**
- **Component Timeout**
- **Visibility Timeout**

  (Correct)

Explanation
The visibility timeout is a period of time during which Amazon SQS prevents other consuming components from receiving and processing a message.

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 the 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 maximum is 12 hours.

References:

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

https://docs.aws.amazon.com/AWSSimpleQueueService/latest/SQSDeveloperGuide/sqs-visibility-timeout.html

Check out this Amazon SQS Cheat Sheet:

https://tutorialsdojo.com/amazon-sqs/

Question 25: **Correct**
An e-commerce application is using a fanout messaging pattern for its order management system. For every order, it sends an Amazon SNS message to an SNS topic, and the message is replicated and pushed to multiple Amazon SQS queues for parallel asynchronous processing. A Spot EC2 instance retrieves the message from each SQS queue and processes the message. There was an incident that while an EC2 instance is currently processing a message, the instance was abruptly terminated, and the processing was not completed in time.

In this scenario, what happens to the SQS message?

- **The message will be sent to a Dead Letter Queue in AWS DataSync.**
- **When the message visibility timeout expires, the message becomes available for processing by other EC2 instances**

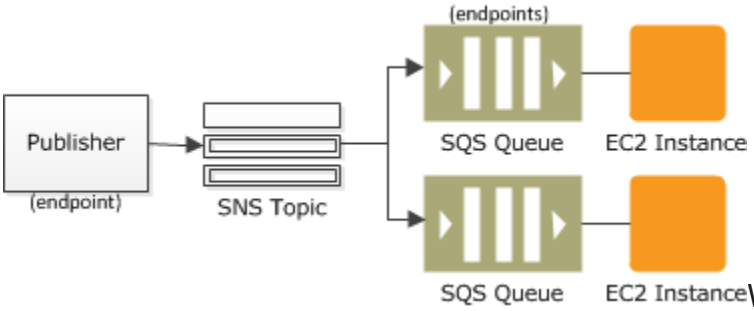
  (Correct)
- **The message is deleted and becomes duplicated in the SQS when the EC2 instance comes online.**

-

The message will automatically be assigned to the same EC2 instance when it comes back online within or after the visibility timeout.

Explanation

A "fanout" pattern is when an Amazon SNS message is sent to a topic and then replicated and pushed to multiple Amazon SQS queues, HTTP endpoints, or email addresses. This allows for parallel asynchronous processing. For example, you could develop an application that sends an Amazon SNS message to a topic whenever an order is placed for a product. Then, the Amazon SQS queues that are subscribed to that topic would receive identical notifications for the new order. The Amazon EC2 server instance attached to one of the queues could handle the processing or fulfillment of the order, while the other server instance could be attached to a data warehouse for analysis of all orders received.



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 the 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 maximum is 12 hours.

The option that says: **The message will automatically be assigned to the same EC2 instance when it comes back online within or after the visibility timeout** is incorrect because the message will not be automatically assigned to the same EC2 instance once it is abruptly terminated. When the message visibility timeout expires, the message becomes available for processing by other EC2 instances.

The option that says: **The message is deleted and becomes duplicated in the SQS when the EC2 instance comes online** is incorrect because the message will not be deleted and won't be duplicated in the SQS queue when the EC2 instance comes online.

The option that says: **The message will be sent to a Dead Letter Queue in AWS DataSync** is incorrect because although the message could be programmatically sent to a Dead Letter Queue (DLQ), it won't be handled by AWS DataSync but by Amazon SQS instead. AWS DataSync is primarily used to simplify your migration with AWS. It makes it simple and fast to move large amounts of data online between on-premises storage and Amazon S3 or Amazon Elastic File System (Amazon EFS).

References:

http://docs.aws.amazon.com/AWSSimpleQueueService/latest/SQSDeveloperGuide/sqs-visibility-timeout.html

https://docs.aws.amazon.com/sns/latest/dg/sns-common-scenarios.html

Check out this Amazon SQS Cheat Sheet:

https://tutorialsdojo.com/amazon-sqs/

Question 26: **Correct**
A leading e-commerce company is in need of a storage solution that can be simultaneously accessed by 1000 Linux servers in multiple availability zones. The servers are hosted in EC2 instances that use a hierarchical directory structure via the NFSv4 protocol. The service should be able to handle the rapidly changing data at scale while still maintaining high performance. It should also be highly durable and highly available whenever the servers will pull data from it, with little need for management.

As the Solutions Architect, which of the following services is the most cost-effective choice that you should use to meet the above requirement?

- **Storage Gateway**
- **S3**

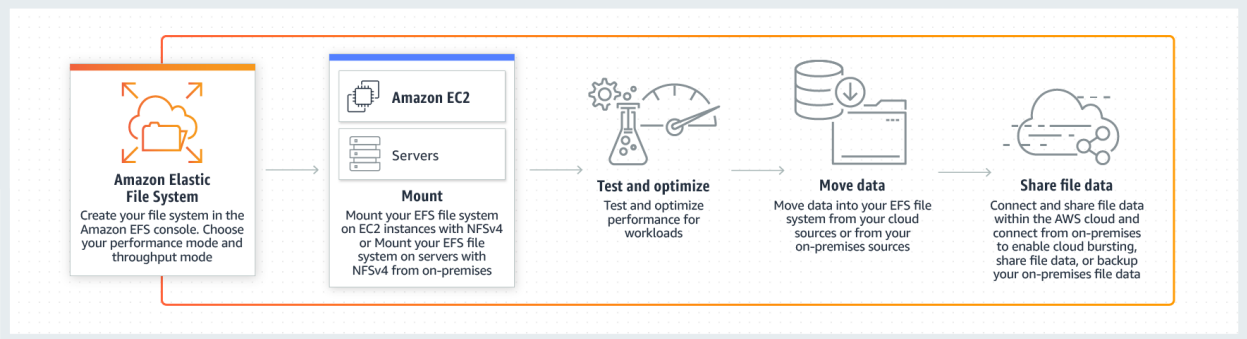
- **EBS**
- **EFS**

(Correct)

Explanation

Amazon Web Services (AWS) offers cloud storage services to support a wide range of storage workloads such as EFS, S3 and EBS. You have to understand when you should use Amazon EFS, Amazon S3 and Amazon Elastic Block Store (EBS) based on the specific workloads. In this scenario, the keywords are ***rapidly changing data*** and 1000 Linux servers.

Amazon EFS is a file storage service for use with Amazon EC2. Amazon EFS provides a file system interface, file system access semantics (such as strong consistency and file locking), and concurrently-accessible storage for **up to thousands of Amazon EC2 instances**. EFS provides the same level of high availability and high scalability like S3 however, this service is more suitable for scenarios where it is required to have a POSIX-compatible file system or if you are storing rapidly changing data.



Data that must be updated very frequently might be better served by storage solutions that take into account read and write latencies, such as Amazon EBS volumes, Amazon RDS, Amazon DynamoDB, Amazon EFS, or relational databases running on Amazon EC2.

Amazon EBS is a block-level storage service for use with Amazon EC2. Amazon EBS can deliver performance for workloads that require the lowest-latency access to data from a single EC2 instance.

Amazon S3 is an object storage service. Amazon S3 makes data available through an Internet API that can be accessed anywhere.

In this scenario, **EFS** is the best answer. As stated above, Amazon EFS provides a file system interface, file system access semantics (such as strong consistency and file locking), and concurrently-accessible storage for **up to thousands of Amazon EC2 instances**. EFS provides the performance, durability, high availability, and storage capacity needed by the 1000 Linux servers in the scenario.

S3 is incorrect. Although this provides the same level of high availability and high scalability like EFS, this service is not suitable for storing data that is rapidly changing, just as mentioned in the above explanation. It is still more effective to use EFS as it offers strong consistency and file locking which the S3 service lacks.

EBS is incorrect because an EBS Volume cannot be shared by multiple instances.

Storage Gateway is incorrect because this is primarily used to extend the storage of your on-premises data center to your AWS Cloud.

References:

https://docs.aws.amazon.com/efs/latest/ug/how-it-works.html

https://aws.amazon.com/efs/features/

https://d1.awsstatic.com/whitepapers/AWS%20Storage%20Services%20Whitepaper-v9.pdf#page=9

Check out this Amazon EFS Cheat Sheet:

https://tutorialsdojo.com/amazon-efs/

Question 27: **Correct**
A Solutions Architect is migrating several Windows-based applications to AWS that require a scalable file system storage for high-performance computing (HPC). The storage service must have full support for the SMB protocol and Windows NTFS, Active Directory (AD) integration, and Distributed File System (DFS).

Which of the following is the MOST suitable storage service that the Architect should use to fulfill this scenario?

- **AWS DataSync**

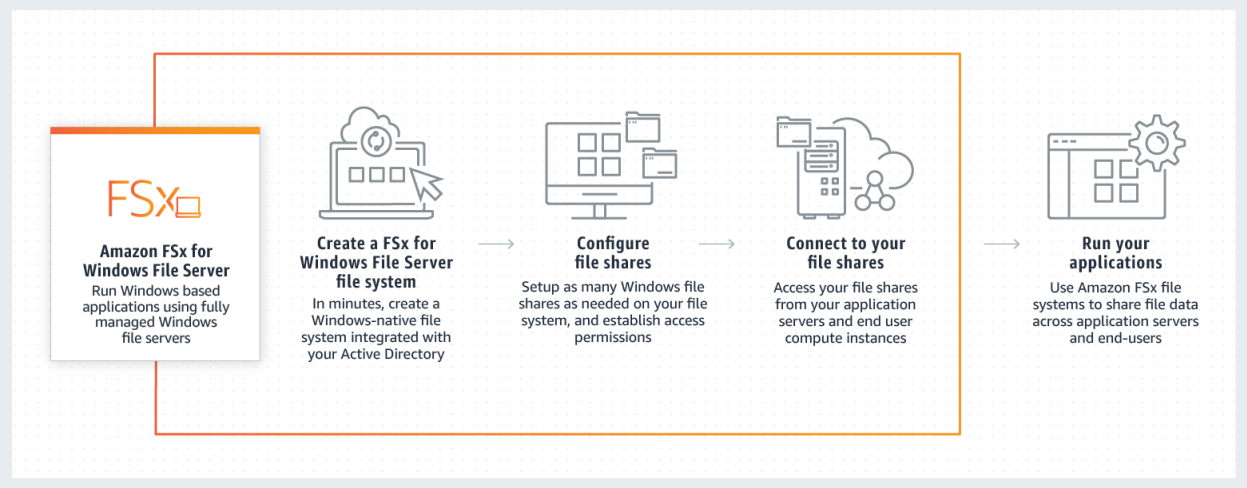
- **Amazon FSx for Lustre**
- **Amazon FSx for Windows File Server**

  (Correct)
- **Amazon S3 Glacier Deep Archive**

Explanation

Amazon FSx provides fully managed third-party file systems. Amazon FSx provides you with the native compatibility of third-party file systems with feature sets for workloads such as Windows-based storage, high-performance computing (HPC), machine learning, and electronic design automation (EDA). You don't have to worry about managing file servers and storage, as Amazon FSx automates time-consuming administration tasks such as hardware provisioning, software configuration, patching, and backups. Amazon FSx integrates the file systems with cloud-native AWS services, making them even more useful for a broader set of workloads.

Amazon FSx provides you with two file systems to choose from: Amazon FSx for Windows File Server for Windows-based applications and Amazon FSx for Lustre for compute-intensive workloads.



For Windows-based applications, Amazon FSx provides fully managed Windows file servers with features and performance optimized for "lift-and-shift" business-critical application workloads including home directories (user shares), media workflows, and ERP applications. It is accessible from Windows and Linux instances via the SMB

protocol. If you have Linux-based applications, Amazon EFS is a cloud-native fully managed file system that provides simple, scalable, elastic file storage accessible from Linux instances via the NFS protocol.

For compute-intensive and fast processing workloads, like high-performance computing (HPC), machine learning, EDA, and media processing, Amazon FSx for Lustre, provides a file system that's optimized for performance, with input and output stored on Amazon S3.

Hence, the correct answer is: **Amazon FSx for Windows File Server.**

Amazon S3 Glacier Deep Archive is incorrect because this service is primarily used as a secure, durable, and extremely low-cost cloud storage for data archiving and long-term backup.

AWS DataSync is incorrect because this service simply provides a fast way to move large amounts of data online between on-premises storage and Amazon S3 or Amazon Elastic File System (Amazon EFS).

Amazon FSx for Lustre is incorrect because this service doesn't support Windows-based applications as well as Windows servers.

References:

https://aws.amazon.com/fsx/

https://aws.amazon.com/getting-started/use-cases/hpc/3/

Check out this Amazon FSx Cheat Sheet:

https://tutorialsdojo.com/amazon-fsx/

Question 28: **Correct**
A company plans to use Route 53 instead of an ELB to load balance the incoming request to the web application. The system is deployed to two EC2 instances to which the traffic needs to be distributed. You want to set a specific percentage of traffic to go to each instance.

Which routing policy would you use?

- 

Latency

-

Geolocation

-

Weighted

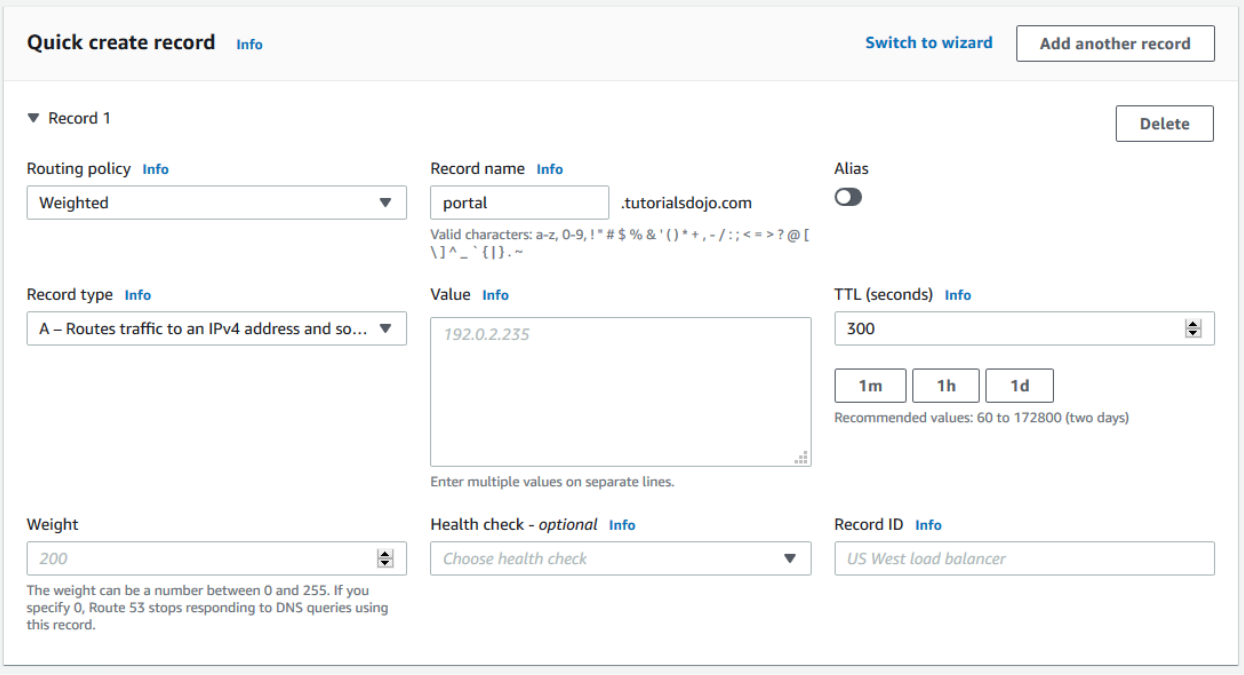
(Correct)

-

Failover

Explanation

Weighted routing lets you associate multiple resources with a single domain name (tutorialsdojo.com) or subdomain name (portal.tutorialsdojo.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 software. You can set a specific percentage of how much traffic will be allocated to the resource by specifying the weights.



For example, if you want to send a tiny portion of your traffic to one resource and the rest to another resource, you might specify weights of 1 and 255. The resource with a

weight of 1 gets 1/256th of the traffic (1/1+255), and the other resource gets 255/256ths (255/1+255).

You can gradually change the balance by changing the weights. If you want to stop sending traffic to a resource, you can change the weight for that record to 0.

Hence, the correct answer is **Weighted.**

Latency is incorrect because you cannot set a specific percentage of traffic for the 2 EC2 instances with this routing policy. Latency routing policy is primarily used when you have resources in multiple AWS Regions and if you need to automatically route traffic to a specific AWS Region that provides the best latency with less round-trip time.

Failover is incorrect because this type is commonly used if you want to set up an active-passive failover configuration for your web application.

Geolocation is incorrect because this is more suitable for routing traffic based on the location of your users, and not for distributing a specific percentage of traffic to two AWS resources.

Reference:

http://docs.aws.amazon.com/Route53/latest/DeveloperGuide/routing-policy.html

Amazon Route 53 Overview:

https://youtu.be/Su308t19ubY

Check out this Amazon Route 53 Cheat Sheet:

https://tutorialsdojo.com/amazon-route-53/

Question 29: **Correct**
A health organization is using a large Dedicated EC2 instance with multiple EBS volumes to host its health records web application. The EBS volumes must be encrypted due to the confidentiality of the data that they are handling and also to comply with the HIPAA (Health Insurance Portability and Accountability Act) standard.

In EBS encryption, what service does AWS use to secure the volume's data at rest? (Select TWO.)

- ☑ **By using your own keys in AWS Key Management Service (KMS).**

  (Correct)

- ☐ **By using a password stored in CloudHSM.**

- ☐ **By using S3 Server-Side Encryption.**

- ☐ **By using the SSL certificates provided by the AWS Certificate Manager (ACM).**

- ☑ **By using Amazon-managed keys in AWS Key Management Service (KMS).**

  (Correct)

- ☐ **By using S3 Client-Side Encryption.**

Explanation

Amazon EBS encryption offers seamless encryption of EBS data volumes, boot volumes, and snapshots, eliminating the need to build and maintain a secure key management infrastructure. EBS encryption enables data at rest security by encrypting your data using Amazon-managed keys, or keys you create and manage using the AWS Key Management Service (KMS). The encryption occurs on the servers that host EC2 instances, providing encryption of data as it moves between EC2 instances and EBS storage.

Hence, the correct answers are: **using your own keys in AWS Key Management Service (KMS)** and **using Amazon-managed keys in AWS Key Management Service (KMS).**

Using S3 Server-Side Encryption and **using S3 Client-Side Encryption** are both incorrect as these relate only to S3.

Using a password stored in CloudHSM is incorrect as you only store keys in CloudHSM and not passwords.

Using the SSL certificates provided by the AWS Certificate Manager (ACM) is incorrect as ACM only provides SSL certificates and not data encryption of EBS Volumes.

Reference:

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

Check out this Amazon EBS Cheat Sheet:

https://tutorialsdojo.com/amazon-ebs/

Question 30: **Correct**
A company needs secure access to its Amazon RDS for MySQL database that is used by multiple applications. Each IAM user must use a short-lived authentication token to connect to the database.

Which of the following is the most suitable solution in this scenario?

- **Use IAM DB Authentication and create database accounts using the AWS-provided `AWSAuthenticationPlugin` plugin in MySQL.**

  (Correct)

- **Use AWS SSO to access the RDS database.**

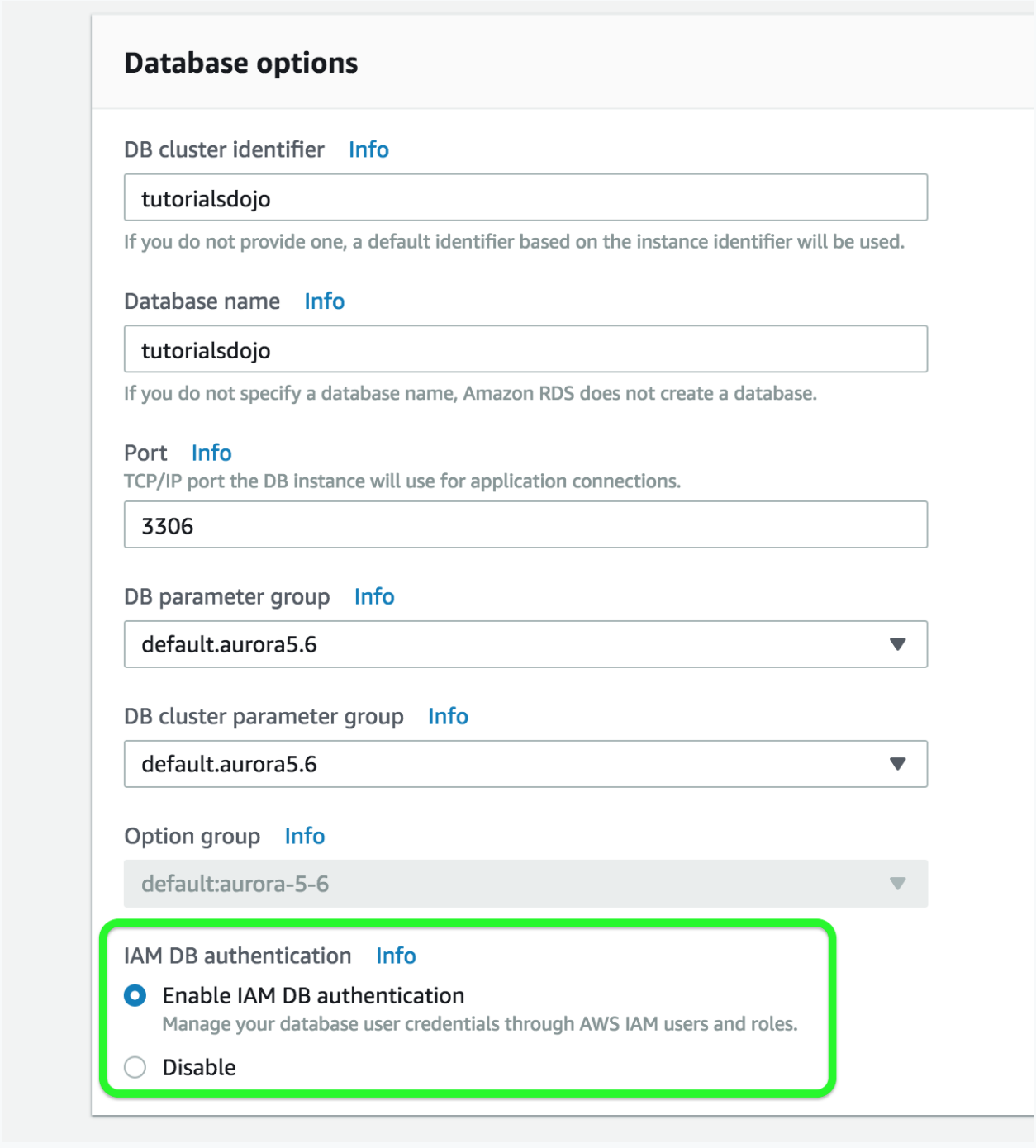
- **Use an MFA token to access and connect to a database.**

- **Use AWS Secrets Manager to generate and store short-lived authentication tokens.**

Explanation

You can authenticate to your DB instance using AWS Identity and Access Management (IAM) database authentication. IAM database authentication works with MySQL and PostgreSQL. With this authentication method, you don't need to use a password when you connect to a DB instance.

An **authentication token** is a string of characters that you use instead of a password. After you generate an authentication token, it's valid for 15 minutes before it expires. If you try to connect using an expired token, the connection request is denied.



Since the scenario asks you to create a short-lived authentication token to access an Amazon RDS database, you can use an IAM database authentication when connecting to a database instance. Authentication is handled by `AWSAuthenticationPlugin`—an AWS-provided plugin that works seamlessly with IAM to authenticate your IAM users.

IAM database authentication provides the following benefits:

1. Network traffic to and from the database is encrypted using Secure Sockets Layer (SSL).
2. You can use IAM to centrally manage access to your database resources instead of managing access individually on each DB instance.
3. For applications running on Amazon EC2, you can use profile credentials specific to your EC2 instance to access your database instead of a password for greater security

Hence, the correct answer is the option that says: **Use IAM DB Authentication and create database accounts using the AWS-provided `AWSAuthenticationPlugin` plugin in MySQL.**

The options that say: **Use AWS SSO to access the RDS database** is incorrect because AWS SSO just enables you to centrally manage SSO access and user permissions for all of your AWS accounts managed through AWS Organizations.

The option that says: **Use AWS Secrets Manager to generate and store short-lived authentication tokens** is incorrect because AWS Secrets Manager is not a suitable service to create an authentication token to access an Amazon RDS database. It's primarily used to store passwords, secrets, and other sensitive credentials. It can't generate a short-lived token either. You have to use IAM DB Authentication instead.

The option that says: **Use an MFA token to access and connect to a database** is incorrect because you can't use an MFA token to connect to your database. You have to set up IAM DB Authentication instead.

References:

https://docs.aws.amazon.com/AmazonRDS/latest/AuroraUserGuide/UsingWithRDS.IAMDBAuth.Connecting.html

https://docs.aws.amazon.com/AmazonRDS/latest/UserGuide/UsingWithRDS.IAMDBAuth.DBAccounts.html

Check out this AWS IAM Cheat Sheet:

https://tutorialsdojo.com/aws-identity-and-access-management-iam/

Question 31: **Correct**
A company is planning to launch a High Performance Computing (HPC) cluster in AWS that does Computational Fluid Dynamics (CFD) simulations. The solution should scale-out their simulation jobs to experiment with more tunable parameters for faster and more accurate results. The cluster is composed of Windows servers hosted on t3a.medium EC2 instances. As the Solutions Architect, you should ensure that the architecture provides higher bandwidth, higher packet per second (PPS) performance, and consistently lower inter-instance latencies.

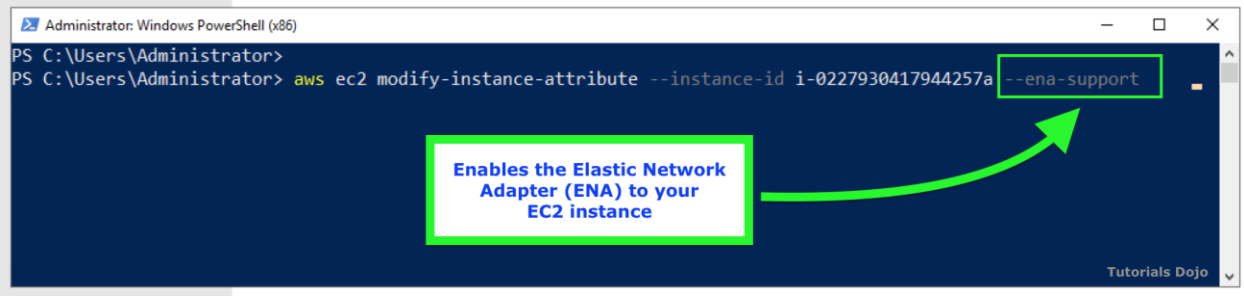
Which is the MOST suitable and cost-effective solution that the Architect should implement to achieve the above requirements?

- **Enable Enhanced Networking with Elastic Fabric Adapter (EFA) on the Windows EC2 Instances.**
- **Enable Enhanced Networking with Elastic Network Adapter (ENA) on the Windows EC2 Instances.**

  (Correct)
- **Use AWS ParallelCluster to deploy and manage the HPC cluster to provide higher bandwidth, higher packet per second (PPS) performance, and lower inter-instance latencies.**
- **Enable Enhanced Networking with Intel 82599 Virtual Function (VF) interface on the Windows EC2 Instances.**

Explanation
Enhanced networking uses single root I/O virtualization (SR-IOV) to provide high-performance networking capabilities on supported instance types. SR-IOV is a method of device virtualization that provides higher I/O performance and lower CPU utilization when compared to traditional virtualized network interfaces. Enhanced networking provides higher bandwidth, higher packet per second (PPS) performance, and

consistently lower inter-instance latencies. There is no additional charge for using enhanced networking.



Amazon EC2 provides enhanced networking capabilities through the Elastic Network Adapter (ENA). It supports network speeds of up to 100 Gbps for supported instance types. Elastic Network Adapters (ENAs) provide traditional IP networking features that are required to support VPC networking.

An Elastic Fabric Adapter (EFA) is simply an Elastic Network Adapter (ENA) with added capabilities. It provides all of the functionality of an ENA, with additional OS-bypass functionality. OS-bypass is an access model that allows HPC and machine learning applications to communicate directly with the network interface hardware to provide low-latency, reliable transport functionality.

The OS-bypass capabilities of EFAs are not supported on Windows instances. If you attach an EFA to a Windows instance, the instance functions as an Elastic Network Adapter without the added EFA capabilities.

Hence, the correct answer is to **Enable Enhanced Networking with Elastic Network Adapter (ENA) on the Windows EC2 Instances**.

Enabling Enhanced Networking with Elastic Fabric Adapter (EFA) on the Windows EC2 Instances is incorrect because the OS-bypass capabilities of the Elastic Fabric Adapter (EFA) are not supported on Windows instances. Although you can attach EFA to your Windows instances, this will just act as a regular Elastic Network Adapter without the added EFA capabilities. Moreover, it doesn't support the t3a.medium instance type that is being used in the HPC cluster.

Enabling Enhanced Networking with Intel 82599 Virtual Function (VF) interface on the Windows EC2 Instances is incorrect. Although you can attach an Intel 82599 Virtual Function (VF) interface to your Windows EC2 Instances to improve its networking capabilities, it doesn't support the t3a.medium instance type that is being used in the HPC cluster.

Using AWS ParallelCluster to deploy and manage the HPC cluster to provide higher bandwidth, higher packet per second (PPS) performance, and lower inter-instance

latencies is incorrect because an AWS ParallelCluster is just an AWS-supported open-source cluster management tool that makes it easy for you to deploy and manage High-Performance Computing (HPC) clusters on AWS. It does not provide higher bandwidth, higher packet per second (PPS) performance, and lower inter-instance latencies, unlike ENA or EFA.

References:

https://docs.aws.amazon.com/AWSEC2/latest/UserGuide/enhanced-networking.html

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

Question 32: **Correct**
A company troubleshoots the operational issues of their cloud architecture by logging the AWS API call history of all AWS resources. The Solutions Architect must implement a solution to quickly identify the most recent changes made to resources in their environment, including creation, modification, and deletion of AWS resources. One of the requirements is that the generated log files should be encrypted to avoid any security issues.

Which of the following is the most suitable approach to implement the encryption?

- **Use CloudTrail and configure the destination S3 bucket to use Server Side Encryption (SSE) with AES-128 encryption algorithm.**
- **Use CloudTrail with its default settings**

  (Correct)
- **Use CloudTrail and configure the destination Amazon Glacier archive to use Server-Side Encryption (SSE).**
- **Use CloudTrail and configure the destination S3 bucket to use Server-Side Encryption (SSE).**

Explanation

By default, CloudTrail event log files are encrypted using Amazon S3 server-side encryption (SSE). You can also choose to encrypt your log files with an AWS Key Management Service (AWS KMS) key. You can store your log files in your bucket for as long as you want. You can also define Amazon S3 lifecycle rules to archive or delete log files automatically. If you want notifications about log file delivery and validation, you can set up Amazon SNS notifications.

Using CloudTrail and configuring the destination Amazon Glacier archive to use Server-Side Encryption (SSE) is incorrect because CloudTrail stores the log files to S3 and not in Glacier. Take note that by default, CloudTrail event log files are already encrypted using Amazon S3 server-side encryption (SSE).

Using CloudTrail and configuring the destination S3 bucket to use Server-Side Encryption (SSE) is incorrect because CloudTrail event log files are already encrypted using the Amazon S3 server-side encryption (SSE) which is why you do not have to do this anymore.

Use CloudTrail and configure the destination S3 bucket to use Server Side Encryption (SSE) with AES-128 encryption algorithm is incorrect because Cloudtrail event log files are already encrypted using the Amazon S3 server-side encryption (SSE) by default. Additionally, SSE-S3 only uses the AES-256 encryption algorithm and not the AES-128.

References:

https://docs.aws.amazon.com/awscloudtrail/latest/userguide/how-cloudtrail-works.html

https://aws.amazon.com/blogs/aws/category/cloud-trail/

Check out this AWS CloudTrail Cheat Sheet:

https://tutorialsdojo.com/aws-cloudtrail/

Question 33: **Correct**
A startup launched a fleet of on-demand EC2 instances to host a massively multiplayer online role-playing game (MMORPG). The EC2 instances are configured with Auto Scaling and AWS Systems Manager.

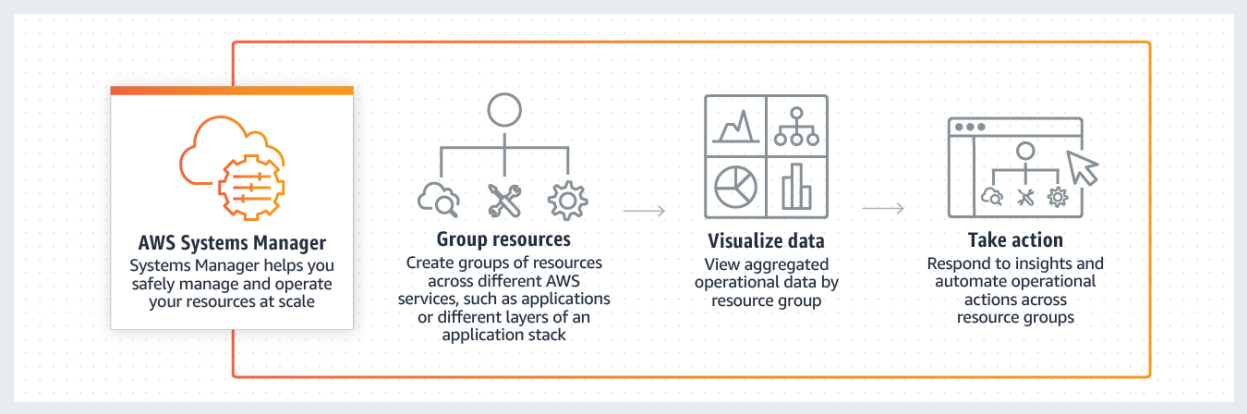
What can be used to configure the EC2 instances without having to establish an RDP or SSH connection to each instance?

- **AWS Config**
- **EC2Config**
- **Run Command**

(Correct)

- **AWS CodePipeline**

Explanation
You can use Run Command from the console to configure instances without having to login to each instance.



AWS Systems Manager Run Command lets you remotely and securely manage the configuration of your managed instances. A *managed instance* is any Amazon EC2 instance or on-premises machine in your hybrid environment that has been configured for Systems Manager. Run Command enables you to automate common administrative tasks and perform ad hoc configuration changes at scale. You can use Run Command from the AWS console, the AWS Command Line Interface, AWS Tools for Windows PowerShell, or the AWS SDKs. Run Command is offered at no additional cost.

Hence, the correct answer is: **Run Command.**

Reference:

https://docs.aws.amazon.com/systems-manager/latest/userguide/execute-remote-commands.html

AWS Systems Manager Overview:

https://youtu.be/KVFKyMAHxqY

Check out this AWS Systems Manager Cheat Sheet:

https://tutorialsdojo.com/aws-systems-manager/

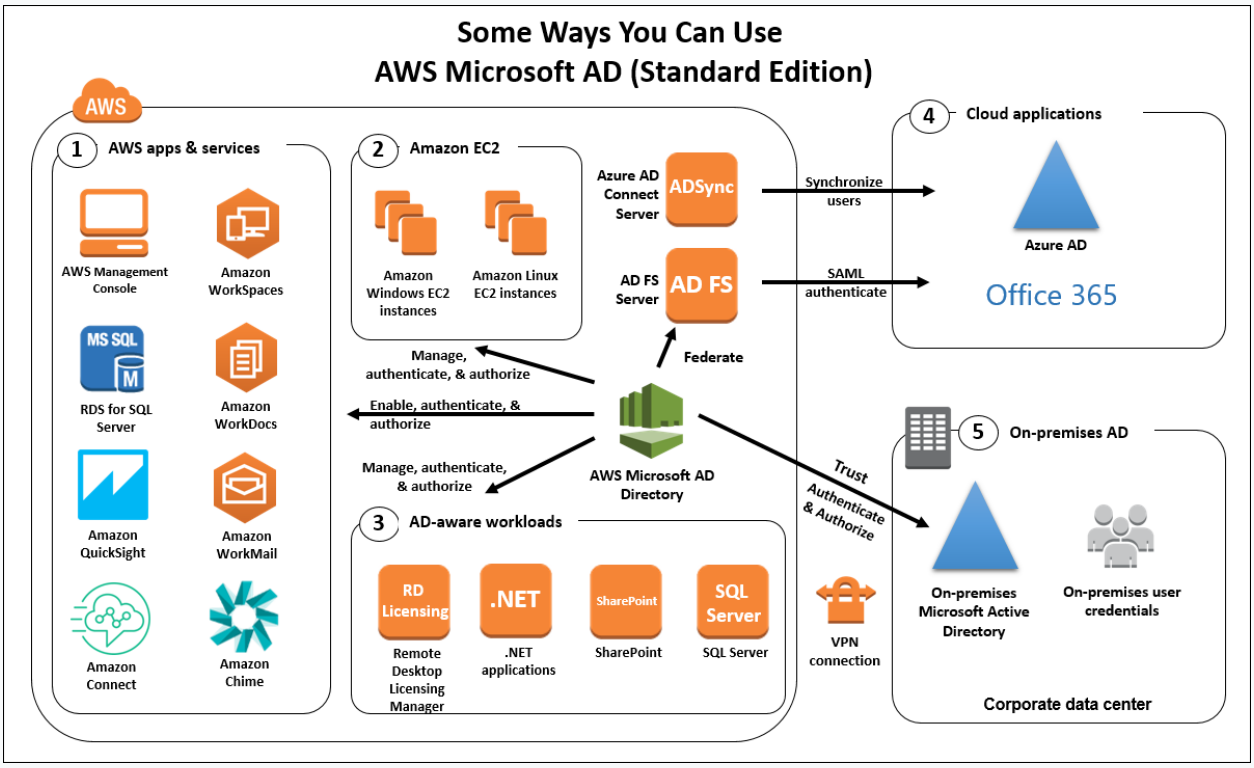
Question 34: **Correct**
A Solutions Architect is working for a fast-growing startup that just started operations during the past 3 months. They currently have an on-premises Active Directory and 10 computers. To save costs in procuring physical workstations, they decided to deploy virtual desktops for their new employees in a virtual private cloud in AWS. The new cloud infrastructure should leverage the existing security controls in AWS but can still communicate with their on-premises network.

Which set of AWS services will the Architect use to meet these requirements?

- **AWS Directory Services, VPN connection, and Amazon S3**
- **AWS Directory Services, VPN connection, and ClassicLink**
- **AWS Directory Services, VPN connection, and AWS Identity and Access Management**
- **AWS Directory Services, VPN connection, and Amazon Workspaces**

  (Correct)

Explanation
For this scenario, the best answer is: **AWS Directory Services, VPN connection, and Amazon Workspaces**.



First, you need a VPN connection to connect the VPC and your on-premises network. Second, you need AWS Directory Services to integrate with your on-premises Active Directory and lastly, you need to use Amazon Workspace to create the needed virtual desktops in your VPC.

References:

https://aws.amazon.com/directoryservice/

https://docs.aws.amazon.com/AmazonVPC/latest/UserGuide/vpn-connections.html

https://aws.amazon.com/workspaces/

AWS Identity Services Overview:

https://youtu.be/AIdUw0i8rr0

Check out these cheat sheets on AWS Directory Service, Amazon VPC, and Amazon WorkSpaces:

https://tutorialsdojo.com/aws-directory-service/

https://tutorialsdojo.com/amazon-vpc/

https://tutorialsdojo.com/amazon-workspaces/

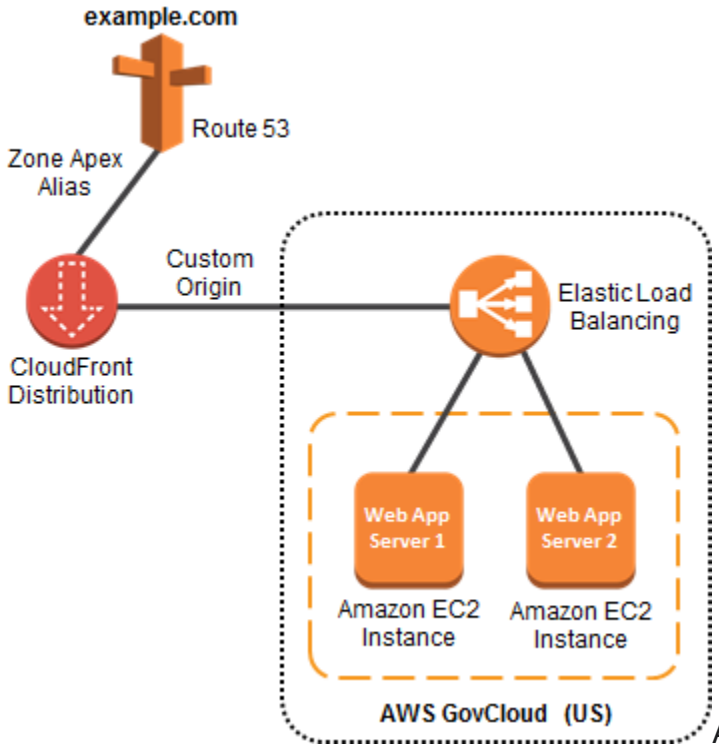
Question 35: **Correct**
A client is hosting their company website on a cluster of web servers that are behind a public-facing load balancer. The client also uses Amazon Route 53 to manage their public DNS.

How should the client configure the DNS zone apex record to point to the load balancer?

- **Create a CNAME record pointing to the load balancer DNS name.**
- **Create an A record aliased to the load balancer DNS name.**

  (Correct)
- **Create an A record pointing to the IP address of the load balancer.**
- **Create an alias for CNAME record to the load balancer DNS name.**

Explanation
Route 53's DNS implementation connects user requests to infrastructure running inside (and outside) of Amazon Web Services (AWS). For example, if you have multiple web servers running on EC2 instances behind an Elastic Load Balancing load balancer, Route 53 will route all traffic addressed to your website (e.g. `www.tutorialsdojo.com`) to the load balancer DNS name (e.g. `elbtutorialsdojo123.elb.amazonaws.com`).



Additionally, Route 53 supports the alias resource record set, which lets you map your **zone apex** (e.g. `tutorialsdojo.com`) DNS name to your load balancer DNS name. IP addresses associated with Elastic Load Balancing can change at any time due to scaling or software updates. Route 53 responds to each request for an Alias resource record set with one IP address for the load balancer.

Creating an A record pointing to the IP address of the load balancer is incorrect. You should be using an Alias record pointing to the DNS name of the load balancer since the IP address of the load balancer can change at any time.

Creating a CNAME record pointing to the load balancer DNS name and **creating an alias for CNAME record to the load balancer DNS name** are incorrect because CNAME records cannot be created for your **zone** apex. You should create an alias record at the top node of a DNS namespace which is also known as the *zone apex*. For example, if you register the DNS name tutorialsdojo.com, the zone apex is tutorialsdojo.com. You can't create a CNAME record directly for tutorialsdojo.com, but you can create an alias record for tutorialsdojo.com that routes traffic to **www**.tutorialsdojo.com.

References:

http://docs.aws.amazon.com/govcloud-us/latest/UserGuide/setting-up-route53-zoneapex-elb.html

https://docs.aws.amazon.com/Route53/latest/DeveloperGuide/resource-record-sets-choosing-alias-non-alias.html

Check out this Amazon Route 53 Cheat Sheet:

https://tutorialsdojo.com/amazon-route-53/

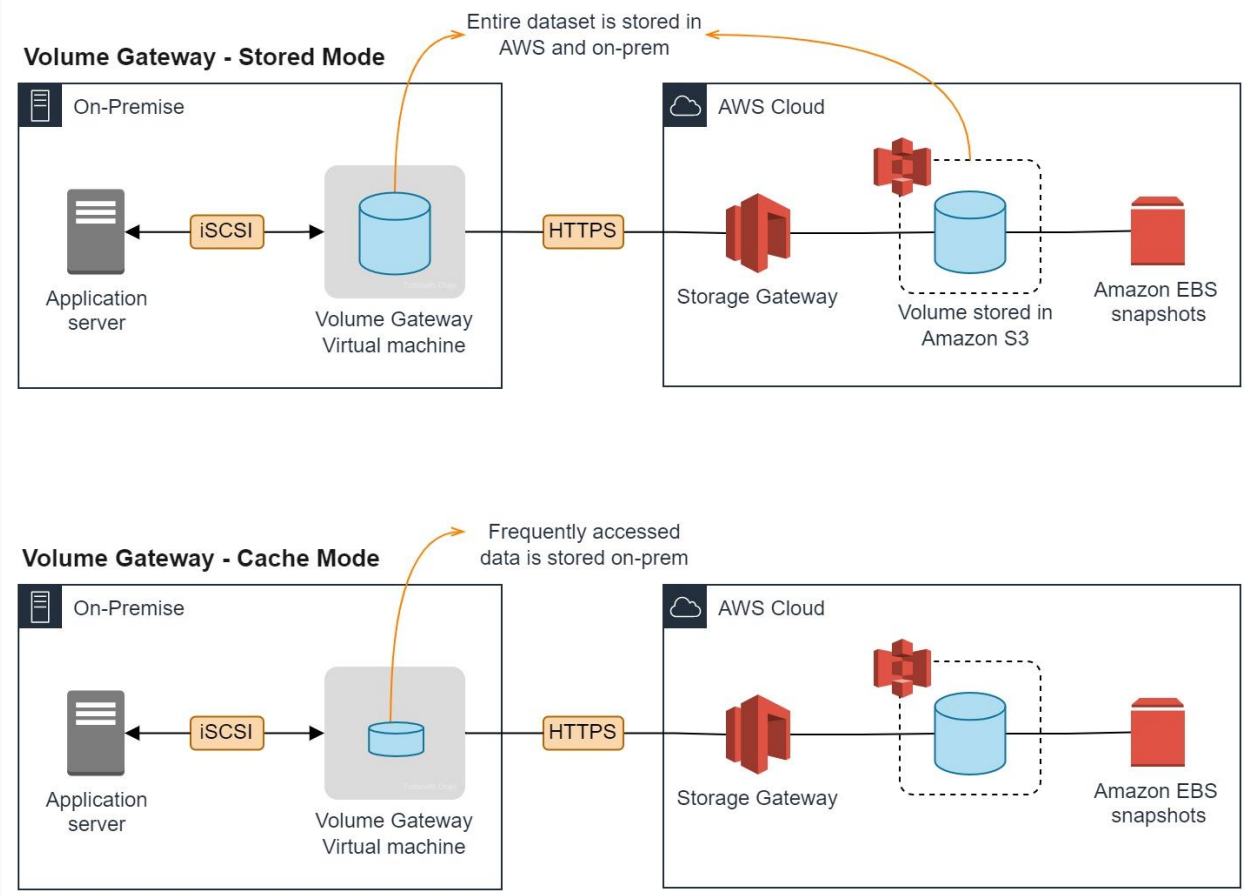
Question 36: **Correct**
A data analytics company keeps a massive volume of data that they store in their on-premises data center. To scale their storage systems, they are looking for cloud-backed storage volumes that they can mount using Internet Small Computer System Interface (iSCSI) devices from their on-premises application servers. They have an on-site data analytics application that frequently accesses the latest data subsets locally while the older data are rarely accessed. You are required to minimize the need to scale the on-premises storage infrastructure while still providing their web application with low-latency access to the data.

Which type of AWS Storage Gateway service will you use to meet the above requirements?

- **Tape Gateway**
- **File Gateway**
- **Volume Gateway in cached mode**

  (Correct)
- **Volume Gateway in stored mode**

Explanation
The **Volume Gateway** is a cloud-based iSCSI block storage volume for your on-premises applications. The Volume Gateway provides either a local cache or full volumes on-premises while also storing full copies of your volumes in the AWS cloud.



There are two options for Volume Gateway:

Cached Volumes - you store volume data in AWS, with a small portion of recently accessed data in the cache on-premises.

Stored Volumes - you store the entire set of volume data on-premises and store periodic point-in-time backups (snapshots) in AWS.

In this scenario, the technology company is looking for a storage service that will enable their analytics application to frequently access the latest data subsets and not the entire data set (as it was mentioned that the old data are rarely being used). This requirement can be fulfilled by setting up a Cached Volume Gateway in AWS Storage Gateway.

By using cached volumes, you can use Amazon S3 as your primary data storage while retaining frequently accessed data locally in your storage gateway. Cached volumes minimize the need to scale your on-premises storage infrastructure while still providing your applications with low-latency access to frequently accessed data. You can create storage volumes up to 32 TiB in size and afterward, attach these volumes as iSCSI devices to your on-premises application servers. When you write to these volumes, your

gateway stores the data in Amazon S3. It retains the recently read data in your on-premises storage gateway's cache and uploads buffer storage.

Cached volumes can range from 1 GiB to 32 TiB in size and must be rounded to the nearest GiB. Each gateway configured for cached volumes can support up to 32 volumes for a total maximum storage volume of 1,024 TiB (1 PiB).

In the cached volumes solution, AWS Storage Gateway stores all your on-premises application data in a storage volume in Amazon S3. Hence, the correct answer is: **Volume Gateway in cached mode.**

Volume Gateway in stored mode is incorrect because the requirement is to provide low latency access to the frequently accessed data subsets locally. Stored Volumes are used if you need low-latency access to your entire dataset.

Tape Gateway is incorrect because this is just a cost-effective, durable, long-term offsite alternative for data archiving, which is not needed in this scenario.

File Gateway is incorrect because the scenario requires you to mount volumes as iSCSI devices. File Gateway is used to store and retrieve Amazon S3 objects through NFS and SMB protocols.

References:

https://docs.aws.amazon.com/storagegateway/latest/userguide/StorageGatewayConcepts.html#volume-gateway-concepts

https://docs.aws.amazon.com/storagegateway/latest/userguide/WhatIsStorageGateway.html

AWS Storage Gateway Overview:

https://youtu.be/pNb7xOBJjHE

Check out this AWS Storage Gateway Cheat Sheet:

https://tutorialsdojo.com/aws-storage-gateway/

Tutorials Dojo's AWS Certified Solutions Architect Associate Exam Study Guide:

https://tutorialsdojo.com/aws-certified-solutions-architect-associate-saa-c03/

Question 37: **Correct**
A company needs to use Amazon S3 to store irreproducible financial documents. For their quarterly reporting, the files are required to be retrieved after a period of 3 months. There will be some occasions when a surprise audit will be held, which requires access to the archived data that they need to present immediately.

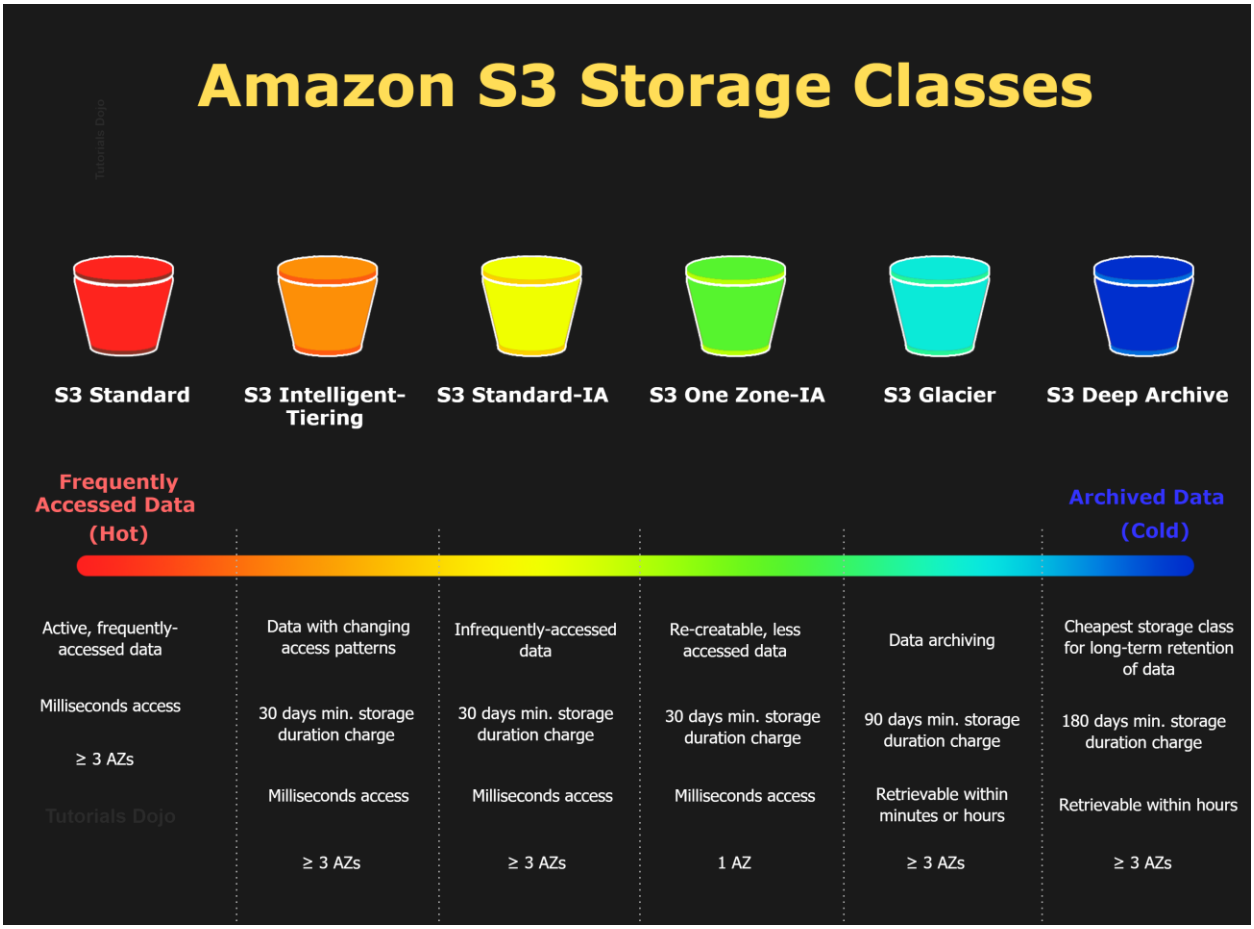
What will you do to satisfy this requirement in a cost-effective way?

- **Use Amazon S3 -Intelligent Tiering**
- **Use Amazon Glacier Deep Archive**
- **Use Amazon S3 Standard - Infrequent Access**

  (Correct)
- **Use Amazon S3 Standard**

Explanation
In this scenario, the requirement is to have a storage option that is cost-effective and has the ability to access or retrieve the archived data immediately. The cost-effective options are Amazon Glacier Deep Archive and Amazon S3 Standard- Infrequent Access (Standard - IA). However, the former option is not designed for rapid retrieval of data which is required for the surprise audit.

Hence, **using Amazon Glacier Deep Archive** is incorrect and the best answer is to **use Amazon S3 Standard - Infrequent Access**.



Using Amazon S3 Standard is incorrect because the standard storage class is not cost-efficient in this scenario. It costs more than Glacier Deep Archive and S3 Standard - Infrequent Access.

Using Amazon S3 -Intelligent Tiering is incorrect because the Intelligent Tiering storage class entails an additional fee for monitoring and automation of each object in your S3 bucket vs. the Standard storage class and S3 Standard - Infrequent Access.

Amazon S3 Standard - Infrequent Access is an Amazon S3 storage class for data that is accessed less frequently but requires rapid access when needed. Standard - IA offers the high durability, throughput, and low latency of Amazon S3 Standard, with a low per GB storage price and per GB retrieval fee.

This combination of low cost and high performance makes Standard - IA ideal for long-term storage, backups, and as a data store for disaster recovery. The Standard - IA storage class is set at the object level and can exist in the same bucket as Standard, allowing you to use lifecycle policies to automatically transition objects between storage classes without any application changes.

References:

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

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

Check out this Amazon S3 Cheat Sheet:

https://tutorialsdojo.com/amazon-s3/

S3 Standard vs S3 Standard-IA vs S3 One Zone IA vs S3 Intelligent Tiering:

https://tutorialsdojo.com/s3-standard-vs-s3-standard-ia-vs-s3-one-zone-ia/

Question 38: **Correct**
A company has a web application hosted on a fleet of EC2 instances located in two Availability Zones that are all placed behind an Application Load Balancer. As a Solutions Architect, you have to add a health check configuration to ensure your application is highly-available.

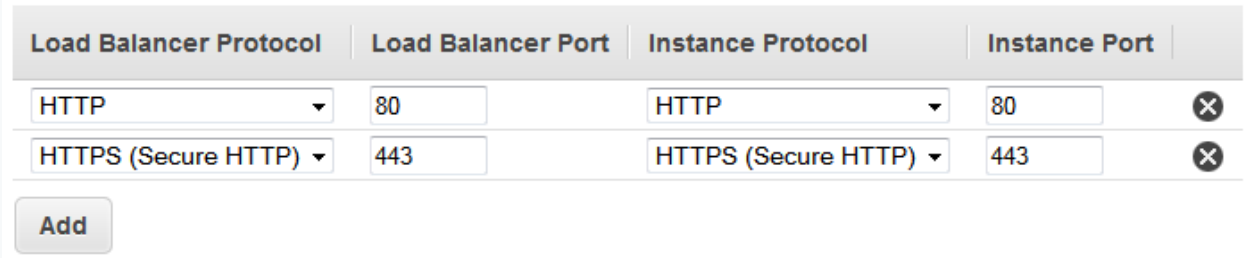
Which health checks will you implement?

- **TCP health check**
- **ICMP health check**
- **HTTP or HTTPS health check**

  (Correct)
- **FTP health check**

Explanation

A load balancer takes requests from clients and distributes them across the EC2 instances that are registered with the load balancer. You can create a load balancer that listens to both the HTTP (80) and HTTPS (443) ports. If you specify that the HTTPS listener sends requests to the instances on port 80, the load balancer terminates the requests, and communication from the load balancer to the instances is not encrypted. If the HTTPS listener sends requests to the instances on port 443, communication from the load balancer to the instances is encrypted.



If your load balancer uses an encrypted connection to communicate with the instances, you can optionally enable authentication of the instances. This ensures that the load balancer communicates with an instance only if its public key matches the key that you specified to the load balancer for this purpose.

The type of ELB that is mentioned in this scenario is an Application Elastic Load Balancer. This is used if you want a flexible feature set for your web applications with HTTP and HTTPS traffic. Conversely, it only allows 2 types of health check: HTTP and HTTPS.

Hence, the correct answer is: **HTTP or HTTPS health check.**

ICMP health check and **FTP health check** are incorrect as these are not supported.

TCP health check is incorrect. A TCP health check is only offered in Network Load Balancers and Classic Load Balancers.

References:

http://docs.aws.amazon.com/elasticloadbalancing/latest/classic/elb-healthchecks.html

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

Check out this AWS Elastic Load Balancing (ELB) Cheat Sheet:

https://tutorialsdojo.com/aws-elastic-load-balancing-elb/

EC2 Instance Health Check vs. ELB Health Check vs. Auto Scaling and Custom Health Check:

https://tutorialsdojo.com/ec2-instance-health-check-vs-elb-health-check-vs-auto-scaling-and-custom-health-check/

Comparison of AWS Services Cheat Sheets:

https://tutorialsdojo.com/comparison-of-aws-services/

Question 39: **Incorrect**
A company is using AWS IAM to manage access to AWS services. The Solutions Architect of the company created the following IAM policy for AWS Lambda:

```
{
  "Version": "2012-10-17",
  "Statement": [
   {
    "Effect": "Allow",
    "Action": [
     "lambda:CreateFunction",
     "lambda:DeleteFunction"
    ],
   "Resource": "*"
},
{
  "Effect": "Deny",
  "Action": [
   "lambda:CreateFunction",
   "lambda:DeleteFunction",
   "lambda:InvokeFunction",
   "lambda:TagResource"
],
  "Resource": "*",
  "Condition": {
    "IpAddress": {
     "aws:SourceIp": "187.5.104.11/32"
    }
   }
  }
 ]
}
```

Which of the following options are allowed by this policy?

- **Create an AWS Lambda function using the `100.220.0.11/32` address.**

  (Correct)

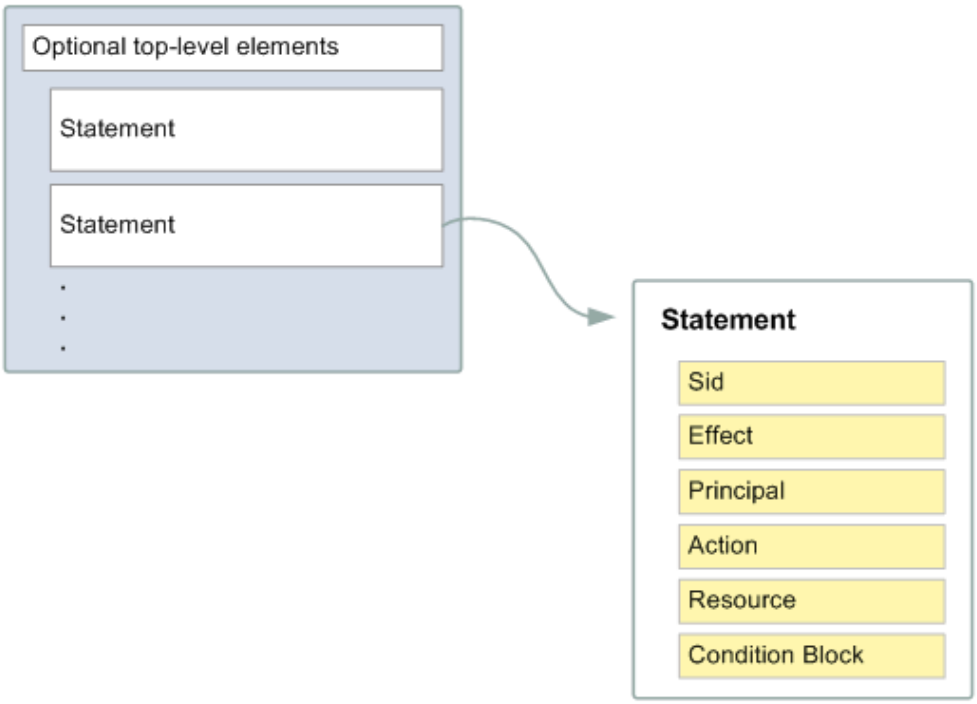
- **Delete an AWS Lambda function using the `187.5.104.11/32` address.**

- **Create an AWS Lambda function using the `187.5.104.11/32` address.**

- **Delete an AWS Lambda function from any network address.**

  (Incorrect)

Explanation
You manage access in AWS by creating policies and attaching them to IAM identities (users, groups of users, or roles) or AWS resources. A policy is an object in AWS that, when associated with an identity or resource, defines their permissions. AWS evaluates these policies when an IAM principal (user or role) makes a request. Permissions in the policies determine whether the request is allowed or denied. Most policies are stored in AWS as JSON documents.



You can use AWS Identity and Access Management (IAM) to manage access to the Lambda API and resources like functions and layers. Based on the given IAM policy, you can create and delete a Lambda function from any network address except for the IP address 187.5.104.11/32. Since the IP address 100.220.0.11/32 is not denied in the policy, you can use this address to create a Lambda function.

Hence, the correct answer is: **Create an AWS Lambda function using the `100.220.0.11/32` address**.

The option that says: **Delete an AWS Lambda function using the `187.5.104.11/32` address** is incorrect because the source IP used in this option is denied by the IAM policy.

The option that says: **Delete an AWS Lambda function from any network address** is incorrect. You can't delete a Lambda function from any network address because the address `187.5.104.11/32` is denied by the policy.

The option that says: **Create an AWS Lambda function using the `187.5.104.11/32` address** is incorrect. Just like the option above, the IAM policy denied the IP address 187.5.104.11/32.

References:

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

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

Check out this AWS IAM Cheat Sheet:

https://tutorialsdojo.com/aws-identity-and-access-management-iam/

Question 40: **Correct**
A company has a web application hosted in their on-premises infrastructure that they want to migrate to AWS cloud. Your manager has instructed you to ensure that there is no downtime while the migration process is on-going. In order to achieve this, your team decided to divert 50% of the traffic to the new application in AWS and the other 50% to the application hosted in their on-premises infrastructure. Once the migration is over and the application works with no issues, a full diversion to AWS will be implemented. The company's VPC is connected to its on-premises network via an AWS Direct Connect connection.

Which of the following are the possible solutions that you can implement to satisfy the above requirement? (Select TWO.)

- ☐ **Use a Network Load balancer with Weighted Target Groups to divert the traffic between the on-premises and AWS-hosted application. Divert 50% of the traffic to the new application in AWS and the other 50% to the application hosted in their on-premises infrastructure.**

- ☐ **Use Route 53 with Failover routing policy to divert and proportion the traffic between the on-premises and AWS-hosted application. Divert 50% of the traffic to the new application in AWS and the other 50% to the application hosted in their on-premises infrastructure.**

- ☐ **Use AWS Global Accelerator to divert and proportion the HTTP and HTTPS traffic between the on-premises and AWS-hosted application. Ensure that the on-premises network has an AnyCast static IP address and is connected to your VPC via a Direct Connect Gateway.**

- ☑

Use an Application Elastic Load balancer with Weighted Target Groups to divert and proportion the traffic between the on-premises and AWS-hosted application. Divert 50% of the traffic to the new application in AWS and the other 50% to the application hosted in their on-premises infrastructure.

(Correct)

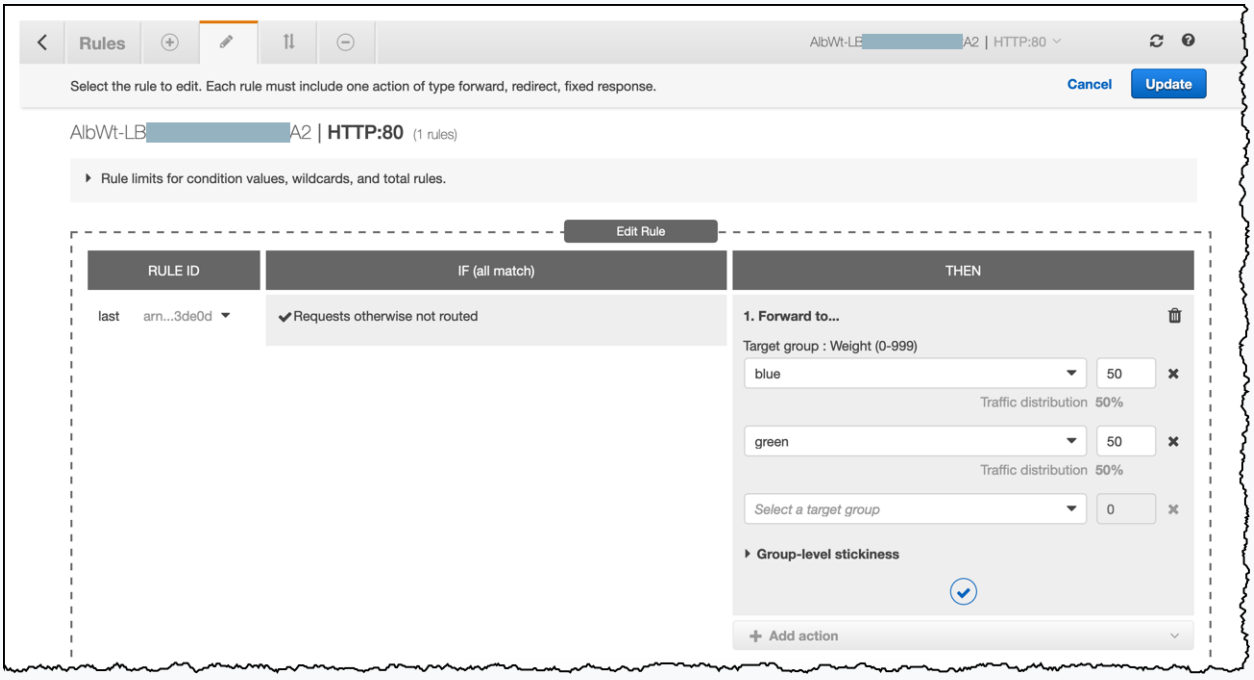
- ☑

Use Route 53 with Weighted routing policy to divert the traffic between the on-premises and AWS-hosted application. Divert 50% of the traffic to the new application in AWS and the other 50% to the application hosted in their on-premises infrastructure.

(Correct)

Explanation

Application Load Balancers support Weighted Target Groups routing. With this feature, you will be able to do weighted routing of the traffic forwarded by a rule to multiple target groups. This enables various use cases like blue-green, canary, and hybrid deployments without the need for multiple load balancers. It even enables zero-downtime migration between on-premises and cloud or between different compute types like EC2 and Lambda.



To divert 50% of the traffic to the new application in AWS and the other 50% to the application, you can also use Route 53 with Weighted routing policy. This will divert the traffic between the on-premises and AWS-hosted applications accordingly.

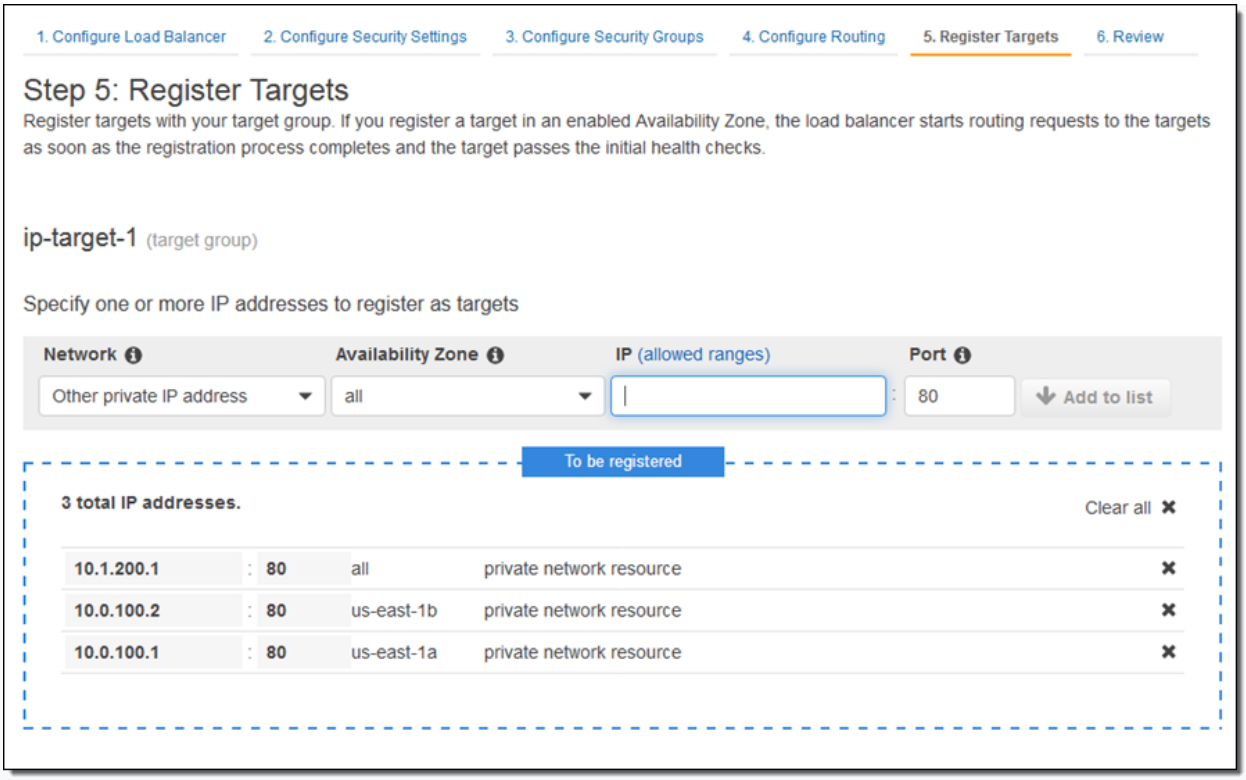
Weighted routing lets you associate multiple resources with a single domain name (tutorialsdojo.com) or subdomain name (portal.tutorialsdojo.com) and choose how much traffic is routed to each resource. This can be useful for a variety of purposes, including load balancing and testing new versions of software. You can set a specific percentage of how much traffic will be allocated to the resource by specifying the weights.

For example, if you want to send a tiny portion of your traffic to one resource and the rest to another resource, you might specify weights of 1 and 255. The resource with a weight of 1 gets 1/256th of the traffic (1/1+255), and the other resource gets 255/256ths (255/1+255).

You can gradually change the balance by changing the weights. If you want to stop sending traffic to a resource, you can change the weight for that record to 0.

When you create a target group in your Application Load Balancer, you specify its target type. This determines the type of target you specify when registering with this target group. You can select the following target types:

1. **instance** - The targets are specified by instance ID.
2. **ip** - The targets are IP addresses.
3. **Lambda** - The target is a Lambda function.



When the target type is **ip**, you can specify IP addresses from one of the following CIDR blocks:

- `10.0.0.0/8` (RFC 1918)
- `100.64.0.0/10` (RFC 6598)
- `172.16.0.0/12` (RFC 1918)
- `192.168.0.0/16` (RFC 1918)
- The subnets of the VPC for the target group

These supported CIDR blocks enable you to register the following with a target group: ClassicLink instances, instances in a VPC that is peered to the load balancer VPC, AWS resources that are addressable by IP address and port (for example, databases), and on-premises resources linked to AWS through AWS Direct Connect or a VPN connection.

Take note that you can not specify publicly routable IP addresses. If you specify targets using an instance ID, traffic is routed to instances using the primary private IP address specified in the primary network interface for the instance. If you specify targets using IP addresses, you can route traffic to an instance using any private IP address from one or more network interfaces. This enables multiple applications on an instance to use the same port. Each network interface can have its own security group.

Hence, the correct answers are the following options:

- Use an Application Elastic Load balancer with Weighted Target Groups to divert and proportion the traffic between the on-premises and AWS-hosted application. Divert 50% of the traffic to the new application in AWS and the other 50% to the application hosted in their on-premises infrastructure.

- Use Route 53 with Weighted routing policy to divert the traffic between the on-premises and AWS-hosted application. Divert 50% of the traffic to the new application in AWS and the other 50% to the application hosted in their on-premises infrastructure.

The option that says: **Use a Network Load balancer with Weighted Target Groups to divert the traffic between the on-premises and AWS-hosted application. Divert 50% of the traffic to the new application in AWS and the other 50% to the application hosted in their on-premises infrastructure** is incorrect because a Network Load balancer doesn't have Weighted Target Groups to divert the traffic between the on-premises and AWS-hosted application.

The option that says: **Use Route 53 with Failover routing policy to divert and proportion the traffic between the on-premises and AWS-hosted application. Divert 50% of the traffic to the new application in AWS and the other 50% to the application hosted in their on-premises infrastructure** is incorrect because you cannot divert and proportion the traffic between the on-premises and AWS-hosted application using Route 53 with Failover routing policy. This is primarily used if you want to configure active-passive failover to your application architecture.

The option that says: **Use AWS Global Accelerator to divert and proportion the HTTP and HTTPS traffic between the on-premises and AWS-hosted application. Ensure that the on-premises network has an AnyCast static IP address and is connected to your VPC via a Direct Connect Gateway** is incorrect. Although you can control the proportion of traffic directed to each endpoint using AWS Global Accelerator by assigning weights across the endpoints, it is still wrong to use a Direct Connect Gateway and an AnyCast IP address since these are not required at all. You can only associate static IP addresses provided by AWS Global Accelerator to regional AWS resources or endpoints, such as Network Load Balancers, Application Load Balancers, EC2 Instances, and Elastic IP addresses. Take note that a Direct Connect Gateway, per se, doesn't establish a connection from your on-premises network to your Amazon VPCs. It simply enables you to use your AWS Direct Connect connection to connect to two or more VPCs that are located in different AWS Regions.

References:

http://docs.aws.amazon.com/Route53/latest/DeveloperGuide/routing-policy.html

https://aws.amazon.com/blogs/aws/new-application-load-balancer-simplifies-deployment-with-weighted-target-groups/

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

Check out this Amazon Route 53 Cheat Sheet:

https://tutorialsdojo.com/amazon-route-53/

Question 41: **Correct**
A company has an application hosted in an Auto Scaling group of Amazon EC2 instances across multiple Availability Zones behind an Application Load Balancer. There are several occasions where some instances are automatically terminated after failing the HTTPS health checks in the ALB and then purges all the ephemeral logs stored in the instance. A Solutions Architect must implement a solution that collects all of the application and server logs effectively. She should be able to perform a root cause analysis based on the logs, even if the Auto Scaling group immediately terminated the instance.

What is the EASIEST way for the Architect to automate the log collection from the Amazon EC2 instances?

- **Add a lifecycle hook to your Auto Scaling group to move instances in the `Terminating` state to the `Terminating:Wait` state to delay the termination of the unhealthy Amazon EC2 instances. Set up AWS Step Functions to collect the application logs and send them to a CloudWatch Log group. Configure the solution to resume the instance termination as soon as all the logs were successfully sent to CloudWatch Logs.**

- **Add a lifecycle hook to your Auto Scaling group to move instances in the `Terminating` state to the `Terminating:Wait` state to delay the termination of the unhealthy Amazon EC2 instances. Configure a CloudWatch Events rule for the `EC2 Instance Terminate Successful` Auto Scaling Event with an associated Lambda function. Set up the AWS Systems Manager Run Command service to run a script that collects and uploads the application logs from the instance to a CloudWatch Logs group. Resume the instance termination once all the logs are sent.**

- **Add a lifecycle hook to your Auto Scaling group to move instances in the `Terminating` state to the `Pending:Wait` state to delay the termination of the unhealthy Amazon EC2 instances. Configure a CloudWatch Events rule for the `EC2 Instance-terminate Lifecycle Action` Auto Scaling Event with an associated Lambda function. Set up an AWS Systems Manager Automation script that collects and uploads the application logs from the instance to a CloudWatch Logs group. Configure the solution to only resume the instance termination once all the logs were successfully sent.**

- **Add a lifecycle hook to your Auto Scaling group to move instances in the `Terminating` state to the `Terminating:Wait` state to delay the termination of unhealthy Amazon EC2 instances. Configure a CloudWatch Events rule for the `EC2 Instance-terminate Lifecycle Action` Auto Scaling Event with an associated Lambda function. Trigger the CloudWatch agent to push the application logs and then resume the instance termination once all the logs are sent to CloudWatch Logs.**

  (Correct)

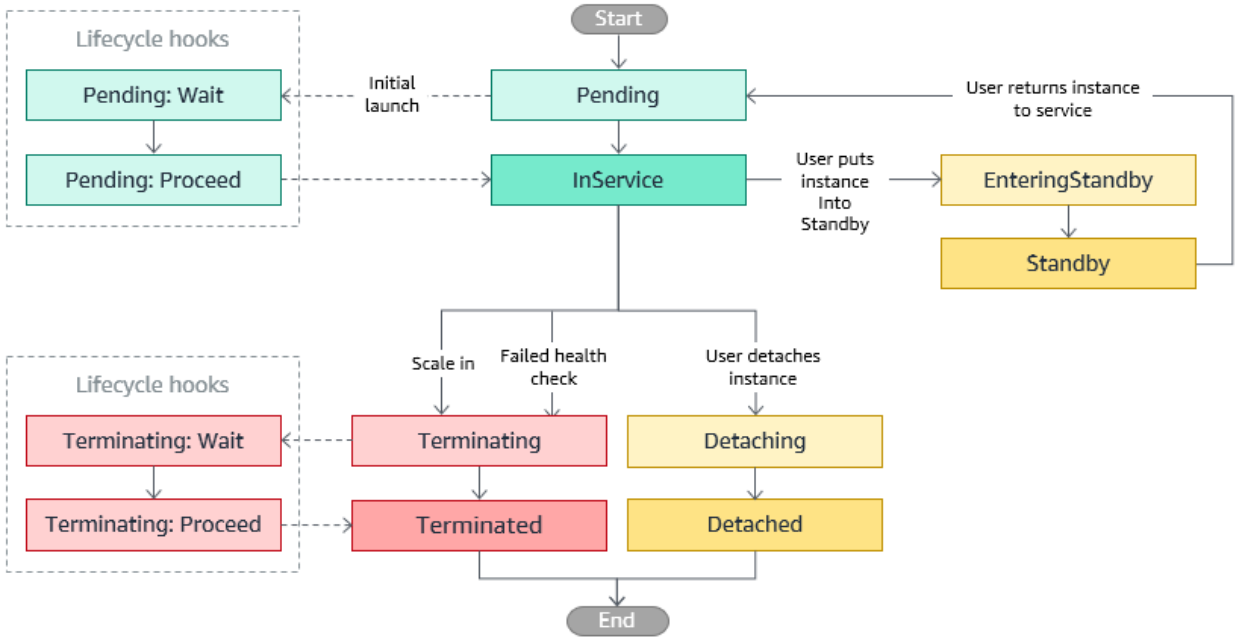
Explanation
The EC2 instances in an Auto Scaling group have a path, or lifecycle, that differs from that of other EC2 instances. The lifecycle starts when the Auto Scaling group launches an instance and puts it into service. The lifecycle ends when you terminate the instance, or the Auto Scaling group takes the instance out of service and terminates it.

You can add a lifecycle hook to your Auto Scaling group so that you can perform custom actions when instances launch or terminate.

When Amazon EC2 Auto Scaling responds to a scale-out event, it launches one or more instances. These instances start in the `Pending` state. If you added an `autoscaling:EC2_INSTANCE_LAUNCHING` lifecycle hook to your Auto Scaling group, the instances move from the `Pending` state to the `Pending:Wait` state. After you complete the lifecycle action, the instances enter the `Pending:Proceed` state. When the instances are fully configured, they are attached to the Auto Scaling group and they enter the `InService` state.

When Amazon EC2 Auto Scaling responds to a scale-in event, it terminates one or more instances. These instances are detached from the Auto Scaling group and enter the `Terminating` state. If you added an `autoscaling:EC2_INSTANCE_TERMINATING` lifecycle hook to your Auto Scaling group,

the instances move from the `Terminating` state to the `Terminating:Wait` state. After you complete the lifecycle action, the instances enter the `Terminating:Proceed` state. When the instances are fully terminated, they enter the `Terminated` state.



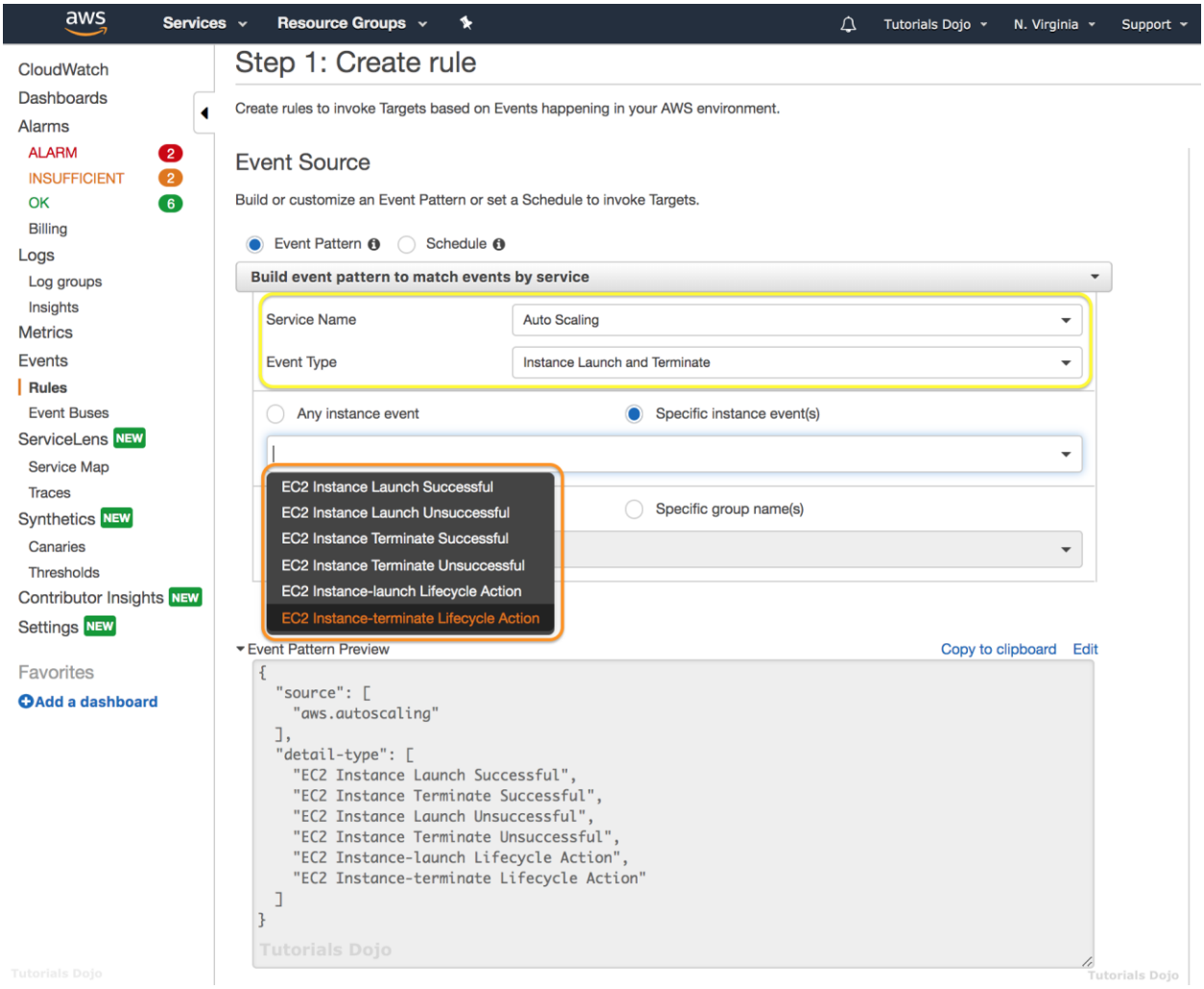
Using CloudWatch agent is the most suitable tool to use to collect the logs. The unified CloudWatch agent enables you to do the following:

- Collect more 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. The additional metrics that can be collected are listed in Metrics Collected by the CloudWatch Agent.

- Collect system-level metrics from on-premises servers. These can include servers in a hybrid environment as well as servers not managed by AWS.

- Retrieve custom metrics from your applications or services using the `StatsD` and `collectd` protocols. `StatsD` is supported on both Linux servers and servers running Windows Server. `collectd` is supported only on Linux servers.

- Collect logs from Amazon EC2 instances and on-premises servers, running either Linux or Windows Server.



You can store and view the metrics that you collect with the CloudWatch agent in CloudWatch just as you can with any other CloudWatch metrics. The default namespace for metrics collected by the CloudWatch agent is `CWAgent`, although you can specify a different namespace when you configure the agent.

Hence, the correct answer is: **Add a lifecycle hook to your Auto Scaling group to move instances in the `Terminating` state to the `Terminating:Wait` state to delay the termination of unhealthy Amazon EC2 instances. Configure a CloudWatch Events rule for the `EC2 Instance-terminate Lifecycle Action` Auto Scaling Event with an associated Lambda function. Trigger the CloudWatch agent to push the application logs and then resume the instance termination once all the logs are sent to CloudWatch Logs.**

The option that says: **Add a lifecycle hook to your Auto Scaling group to move instances in the `Terminating` state to the `Pending:Wait` state to delay the termination of the unhealthy Amazon EC2 instances. Configure a CloudWatch Events rule for the `EC2 Instance-terminate Lifecycle Action` Auto Scaling Event with an associated Lambda function. Set up an AWS Systems Manager Automation script that collects**

and uploads the application logs from the instance to a CloudWatch Logs group. Configure the solution to only resume the instance termination once all the logs were successfully sent is incorrect because the `Pending:Wait` state refers to the scale-out action in Amazon EC2 Auto Scaling and not for scale-in or for terminating the instances.

The option that says: **Add a lifecycle hook to your Auto Scaling group to move instances in the `Terminating` state to the `Terminating:Wait` state to delay the termination of the unhealthy Amazon EC2 instances. Set up AWS Step Functions to collect the application logs and send them to a CloudWatch Log group. Configure the solution to resume the instance termination as soon as all the logs were successfully sent to CloudWatch Logs** is incorrect because using AWS Step Functions is inappropriate in collecting the logs from your EC2 instances. You should use a CloudWatch agent instead.

The option that says: **Add a lifecycle hook to your Auto Scaling group to move instances in the `Terminating` state to the `Terminating:Wait` state to delay the termination of the unhealthy Amazon EC2 instances. Configure a CloudWatch Events rule for the `EC2 Instance Terminate Successful` Auto Scaling Event with an associated Lambda function. Set up the AWS Systems Manager Run Command service to run a script that collects and uploads the application logs from the instance to a CloudWatch Logs group. Resume the instance termination once all the logs are sent** is incorrect. Although this solution could work, it entails a lot of effort to write a custom script that the AWS Systems Manager Run Command will run. Remember that the scenario asks for a solution that you can implement with the least amount of effort. This solution can be simplified by automatically uploading the logs using a CloudWatch Agent. You have to use the `EC2 Instance-terminate Lifecycle Action` event instead.

References:

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

https://docs.aws.amazon.com/autoscaling/ec2/userguide/cloud-watch-events.html#terminate-successful

https://aws.amazon.com/premiumsupport/knowledge-center/auto-scaling-delay-termination/

Check out this AWS Auto Scaling Cheat Sheet:

https://tutorialsdojo.com/aws-auto-scaling/

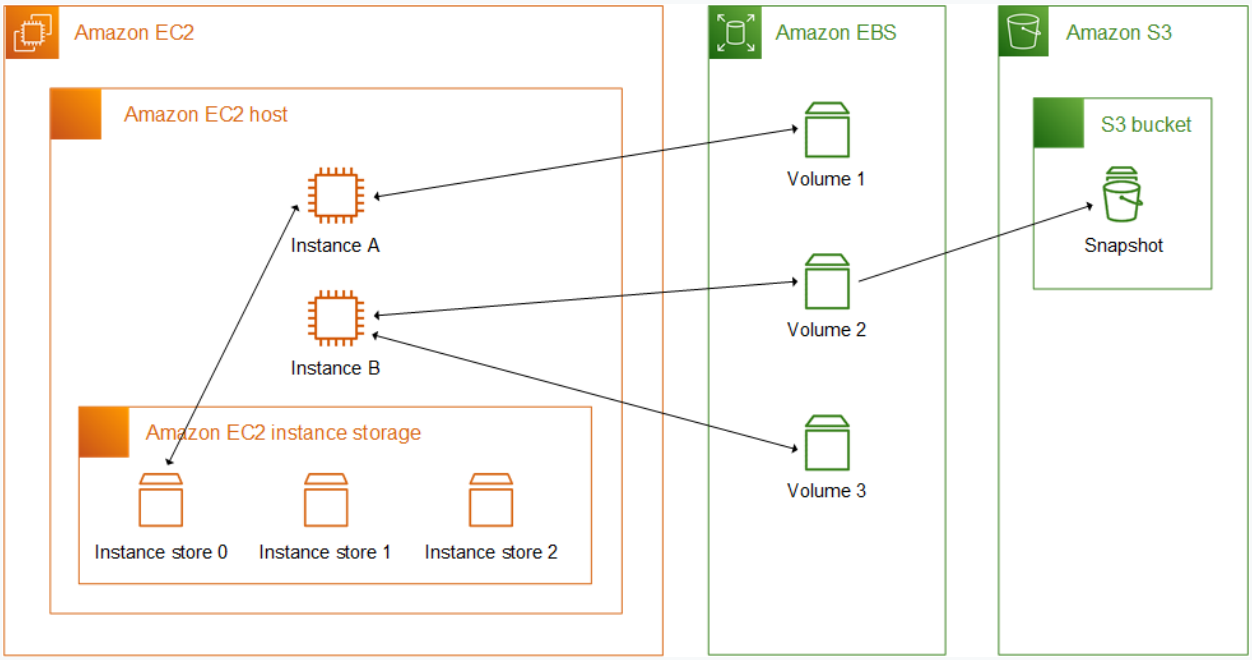
Question 42: **Correct**
A company has a running m5ad.large EC2 instance with a default attached 75 GB SSD instance-store backed volume. You shut it down and then start the instance. You noticed that the data which you have saved earlier on the attached volume is no longer available.

What might be the cause of this?

- **The instance was hit by a virus that wipes out all data.**
- **The EC2 instance was using instance store volumes, which are ephemeral and only live for the life of the instance.**

  (Correct)
- **The volume of the instance was not big enough to handle all of the processing data.**
- **The EC2 instance was using EBS backed root volumes, which are ephemeral and only live for the life of the instance.**

Explanation
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. While an instance store is dedicated to a particular instance, the disk subsystem is shared among instances on a host computer.

The data in an instance store persists only during the lifetime of its associated instance. If an instance reboots (intentionally or unintentionally), data in the instance store persists. However, data in the instance store is lost under the following circumstances:

- The underlying disk drive fails

- The instance stops

- The instance terminates

Reference:

http://docs.aws.amazon.com/AWSEC2/latest/UserGuide/InstanceStorage.html

Amazon EC2 Overview:

https://www.youtube.com/watch?v=7VsGIHT_jQE

Check out this Amazon EC2 Cheat Sheet:

https://tutorialsdojo.com/amazon-elastic-compute-cloud-amazon-ec2/

Question 43: **Correct**

The start-up company that you are working for has a batch job application that is currently hosted on an EC2 instance. It is set to process messages from a queue created in SQS with default settings. You configured the application to process the messages once a week. After 2 weeks, you noticed that not all messages are being processed by the application.

What is the root cause of this issue?

- **Missing permissions in SQS.**
- **The batch job application is configured to long polling.**
- **Amazon SQS has automatically deleted the messages that have been in a queue for more than the maximum message retention period.**

  (Correct)
- **The SQS queue is set to short-polling.**

Explanation

Amazon SQS automatically deletes messages that have been in a queue for more than the maximum message retention period. The default message retention period is 4 days. Since the queue is configured to the default settings and the batch job application only processes the messages once a week, the messages that are in the queue for more than 4 days are deleted. This is the root cause of the issue.

To fix this, you can increase the message retention period to a maximum of 14 days using the SetQueueAttributes action.

References:

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

https://docs.aws.amazon.com/AWSSimpleQueueService/latest/SQSDeveloperGuide/sqs-message-lifecycle.html

Check out this Amazon SQS Cheat Sheet:

https://tutorialsdojo.com/amazon-sqs/

Question 44: **Correct**
A social media company needs to capture the detailed information of all HTTP requests that went through their public-facing Application Load Balancer every five minutes. The client's IP address and network latencies must also be tracked. They want to use this data for analyzing traffic patterns and for troubleshooting their Docker applications orchestrated by the Amazon ECS Anywhere service.

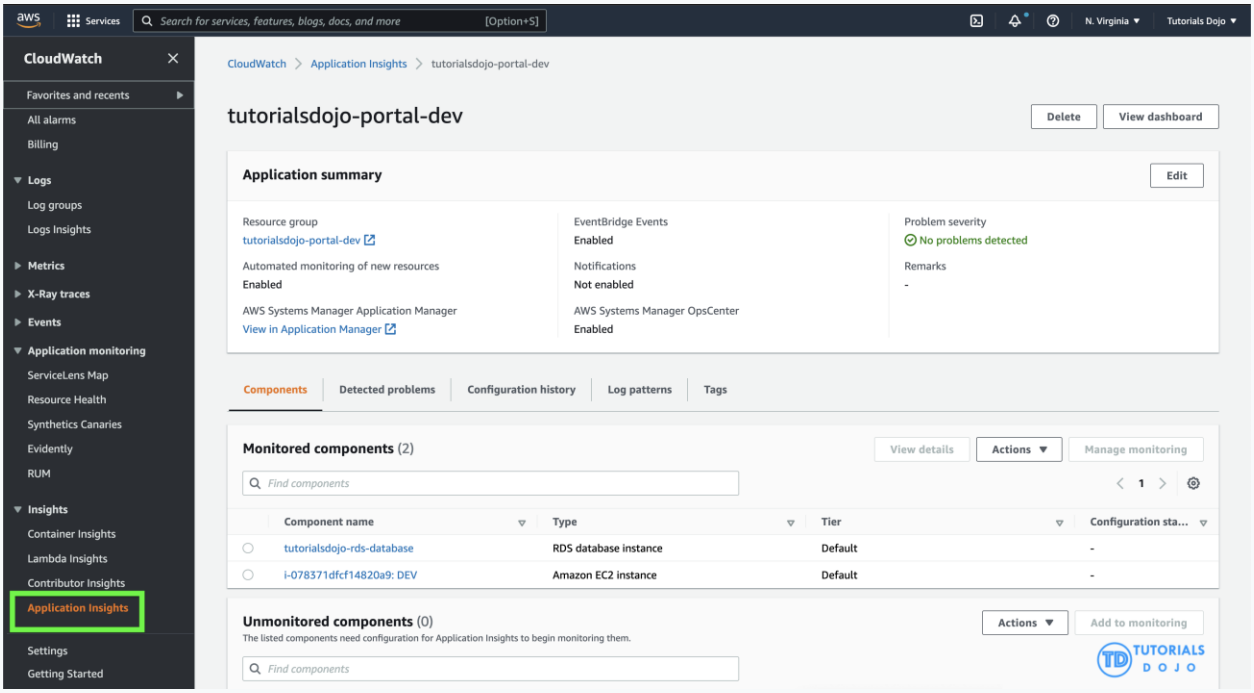
Which of the following options meets the customer requirements with the LEAST amount of overhead?

- **Integrate Amazon EventBridge (Amazon CloudWatch Events) metrics on the Application Load Balancer to capture the client IP address. Use Amazon CloudWatch Container Insights to analyze traffic patterns.**
- **Install and run the AWS X-Ray daemon on the Amazon ECS cluster. Use the Amazon CloudWatch ServiceLens to analyze the traffic that goes through the application.**
- **Enable AWS CloudTrail for their Application Load Balancer. Use the AWS CloudTrail Lake to analyze and troubleshoot the application traffic.**
- **Enable access logs on the Application Load Balancer. Integrate the Amazon ECS cluster with Amazon CloudWatch Application Insights to analyze traffic patterns and simplify troubleshooting.**

(Correct)

Explanation

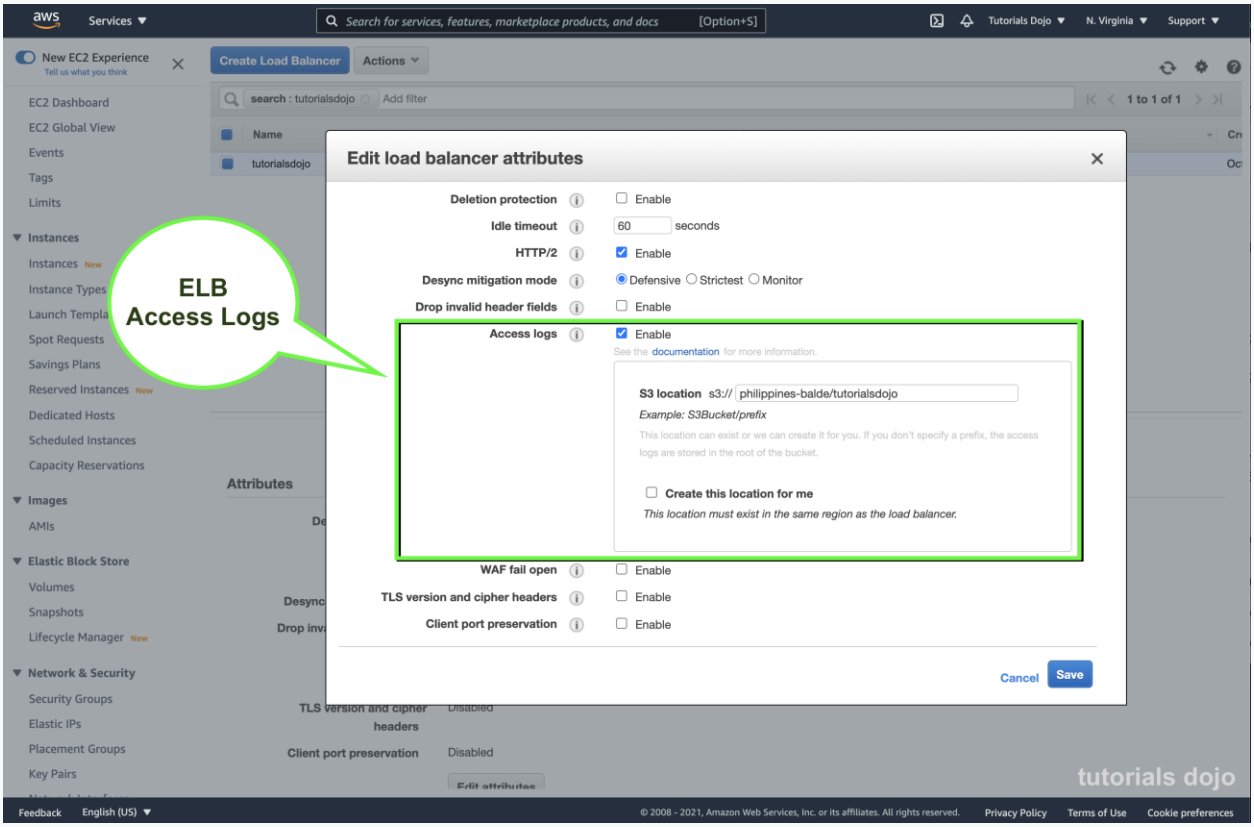
Amazon CloudWatch Application Insights facilitates observability for your applications and underlying AWS resources. It helps you set up the best monitors for your application resources to continuously analyze data for signs of problems with your applications. Application Insights, which is powered by SageMaker and other AWS technologies, provides automated dashboards that show potential problems with monitored applications, which help you to quickly isolate ongoing issues with your applications and infrastructure. The enhanced visibility into the health of your applications that Application Insights provides helps reduce the "mean time to repair" (MTTR) to troubleshoot your application issues.



When you add your applications to Amazon CloudWatch Application Insights, it scans the resources in the applications and recommends and configures metrics and logs on CloudWatch for application components. Example application components include SQL Server backend databases and Microsoft IIS/Web tiers. Application Insights analyzes metric patterns using historical data to detect anomalies and continuously detects errors and exceptions from your application, operating system, and infrastructure logs. It correlates these observations using a combination of classification algorithms and built-in rules. Then, it automatically creates dashboards that show the relevant observations and problem severity information to help you prioritize your actions.

Elastic Load Balancing provides access logs that capture detailed information about requests sent to your load balancer. Each log contains information such as the time the request was received, the client's IP address, latencies, request paths, and server

responses. You can use these access logs to analyze traffic patterns and troubleshoot issues.



Access logging is an optional feature of Elastic Load Balancing that is disabled by default. After you enable access logging for your load balancer, Elastic Load Balancing captures the logs and stores them in the Amazon S3 bucket that you specify as compressed files. You can disable access logging at any time.

Hence, the correct answer is: **Enable access logs on the Application Load Balancer. Integrate the Amazon ECS cluster with Amazon CloudWatch Application Insights to analyze traffic patterns and simplify troubleshooting.**

The option that says: **Enable AWS CloudTrail for their Application Load Balancer. Use the AWS CloudTrail Lake to analyze and troubleshoot the application traffic** is incorrect because AWS CloudTrail is primarily used to monitor and record the account activity across your AWS resources and not your web applications. You cannot use CloudTrail to capture the detailed information of all HTTP requests that go through your public-facing Application Load Balancer (ALB). CloudTrail can only track the resource changes made to your ALB, but not the actual IP traffic that goes through it. For this use case, you have to enable the access logs feature instead. In addition, the AWS CloudTrail Lake feature is more suitable for running SQL-based queries on your API events and not for analyzing application traffic.

The option that says: **Install and run the AWS X-Ray daemon on the Amazon ECS cluster. Use the Amazon CloudWatch ServiceLens to analyze the traffic that goes through the application** is incorrect. Although this solution is possible, this won't track the client's IP address since the access log feature in the ALB is not enabled. Take note that the scenario explicitly mentioned that the client's IP address and network latencies must also be tracked.

The option that says: **Integrate Amazon EventBridge (Amazon CloudWatch Events) metrics on the Application Load Balancer to capture the client IP address. Use Amazon CloudWatch Container Insights to analyze traffic patterns** is incorrect because Amazon EventBridge doesn't track the actual traffic to your ALB. It is the Amazon CloudWatch service that monitors the changes to your ALB itself and the actual IP traffic that it distributes to the target groups. The primary function of CloudWatch Container Insights is to collect, aggregate, and summarize metrics and logs from your containerized applications and microservices.

References:

https://docs.aws.amazon.com/AmazonCloudWatch/latest/monitoring/cloudwatch-application-insights.html

http://docs.aws.amazon.com/elasticloadbalancing/latest/application/load-balancer-access-logs.html

https://docs.aws.amazon.com/elasticloadbalancing/latest/application/load-balancer-monitoring.html

AWS Elastic Load Balancing Overview:

https://youtu.be/UBl5dw59DO8

Check out this AWS Elastic Load Balancing (ELB) Cheat Sheet:

https://tutorialsdojo.com/aws-elastic-load-balancing-elb/

Application Load Balancer vs Network Load Balancer vs Gateway Load Balancer:

https://tutorialsdojo.com/application-load-balancer-vs-network-load-balancer-vs-classic-load-balancer/

Question 45: **Correct**

A company deployed a web application that stores static assets in an Amazon Simple Storage Service (S3) bucket. The Solutions Architect expects the S3 bucket to immediately receive over 2000 PUT requests and 3500 GET requests per second at peak hour.

What should the Solutions Architect do to ensure optimal performance?

- **Do nothing. Amazon S3 will automatically manage performance at this scale.**

  (Correct)

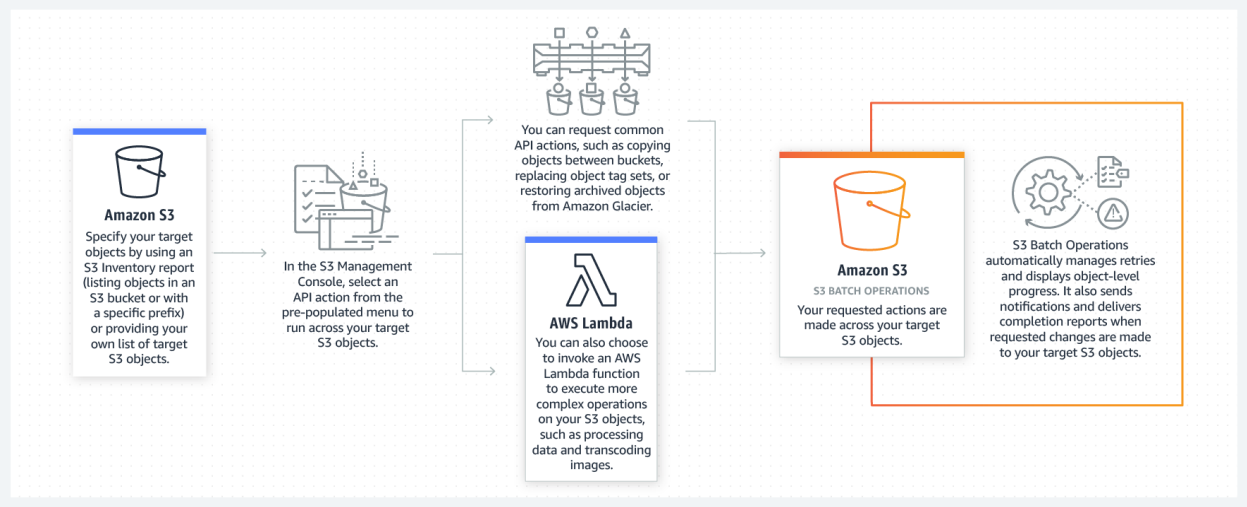
- **Use Byte-Range Fetches to retrieve multiple ranges of an object data per GET request.**

- **Use a predictable naming scheme in the key names such as sequential numbers or date time sequences.**

- **Add a random prefix to the key names.**

Explanation
Amazon S3 now provides increased performance to support at least 3,500 requests per second to add data and 5,500 requests per second to retrieve data, which can save significant processing time for no additional charge. Each S3 prefix can support these request rates, making it simple to increase performance significantly.

Applications running on Amazon S3 today will enjoy this performance improvement with no changes, and customers building new applications on S3 do not have to make any application customizations to achieve this performance. Amazon S3's support for parallel requests means you can scale your S3 performance by the factor of your compute cluster without making any customizations to your application. Performance scales per prefix, so you can use as many prefixes as you need in parallel to achieve the required throughput. There are no limits to the number of prefixes.



This S3 request rate performance increase removes any previous guidance to randomize object prefixes to achieve faster performance. That means you can now use logical or sequential naming patterns in S3 object naming without any performance implications. This improvement is now available in all AWS Regions.

Using Byte-Range Fetches to retrieve multiple ranges of an object data per GET request is incorrect. Although a Byte-Range Fetch helps you achieve higher aggregate throughput, Amazon S3 does **not** support retrieving multiple ranges of data per GET request. 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. Fetching smaller ranges of a large object also allows your application to improve retry times when requests are interrupted.

Adding a random prefix to the key names is incorrect. Adding a random prefix is not required in this scenario because S3 can now scale automatically to adjust performance. You do not need to add a random prefix anymore for this purpose since S3 has increased performance to support at least 3,500 requests per second to add data and 5,500 requests per second to retrieve data, which covers the workload in the scenario.

Using a predictable naming scheme in the key names such as sequential numbers or date time sequences is incorrect because Amazon S3 already maintains an index of object key names in each AWS region. S3 stores key names in alphabetical order. The key name dictates which partition the key is stored in. Using a sequential prefix increases the likelihood that Amazon S3 will target a specific partition for a large number of your keys, overwhelming the I/O capacity of the partition.

References:

https://docs.aws.amazon.com/AmazonS3/latest/dev/request-rate-perf-considerations.html

https://d1.awsstatic.com/whitepapers/AmazonS3BestPractices.pdf

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

Check out this Amazon S3 Cheat Sheet:

https://tutorialsdojo.com/amazon-s3/

Question 46: **Incorrect**
A company needs to accelerate the performance of its AI-powered medical diagnostic application by running its machine learning workloads on the edge of telecommunication carriers' 5G networks. The application must be deployed to a Kubernetes cluster and have role-based access control (RBAC) access to IAM users and roles for cluster authentication.

Which of the following should the Solutions Architect implement to ensure single-digit millisecond latency for the application?

- **Launch the application to an Amazon Elastic Kubernetes Service (Amazon EKS) cluster. Create node groups in Wavelength Zones for the Amazon EKS cluster via the AWS Wavelength service. Apply the AWS authenticator configuration map (`aws-auth ConfigMap`) to your cluster.**

  (Correct)

- **Host the application to an Amazon EKS cluster and run the Kubernetes pods on AWS Fargate. Create node groups in AWS Wavelength Zones for the Amazon EKS cluster. Add the EKS pod execution IAM role (`AmazonEKSFargatePodExecutionRole`) to your cluster and ensure that the Fargate profile has the same IAM role as your Amazon EC2 node groups.**

- **Launch the application to an Amazon Elastic Kubernetes Service (Amazon EKS) cluster. Create VPC endpoints for the AWS Wavelength Zones and apply**

them to the Amazon EKS cluster. Install the AWS IAM Authenticator for Kubernetes (`aws-iam-authenticator`) to your cluster.

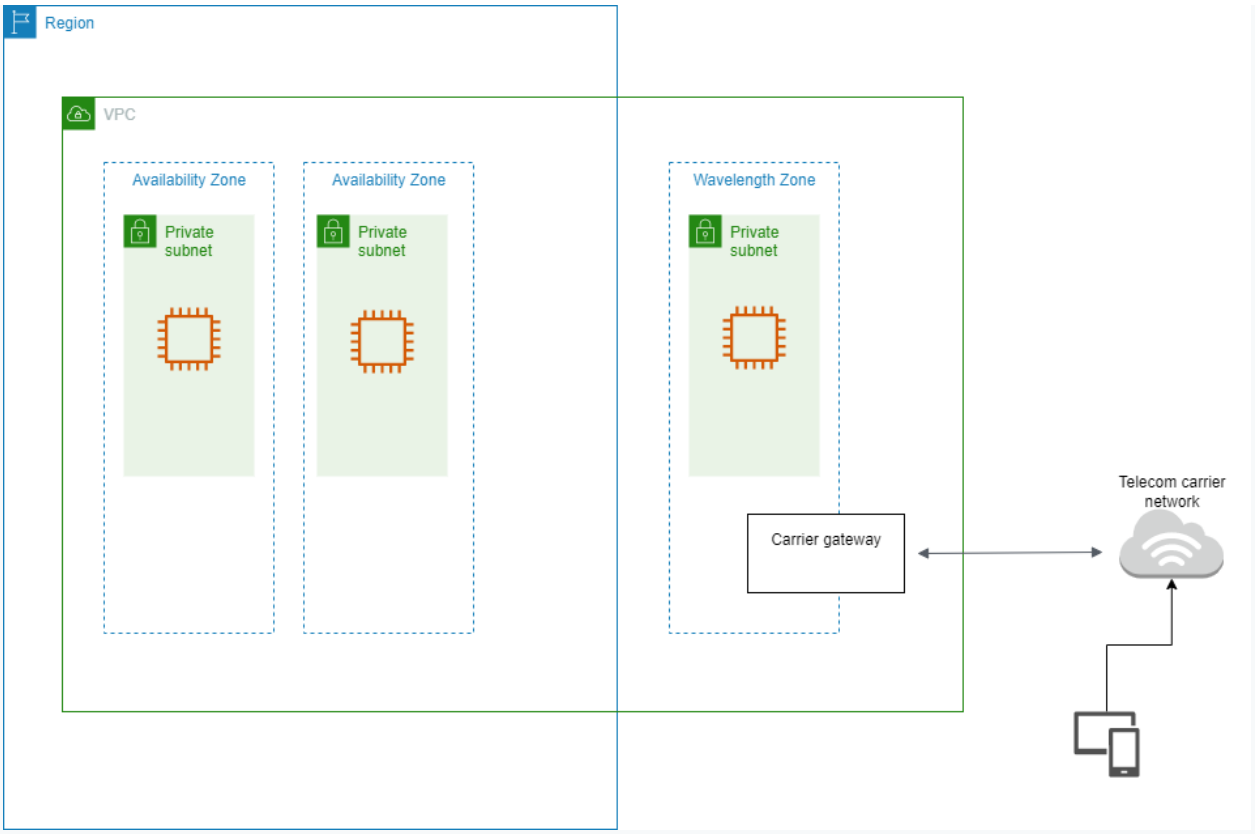
- 

Host the application to an Amazon Elastic Kubernetes Service (Amazon EKS) cluster. Set up node groups in AWS Wavelength Zones for the Amazon EKS cluster. Attach the Amazon EKS connector agent role (`AmazonECSConnectorAgentRole`) to your cluster and use AWS Control Tower for RBAC access.

(Incorrect)

Explanation
AWS Wavelength combines the high bandwidth and ultralow latency of 5G networks with AWS compute and storage services so that developers can innovate and build a new class of applications.

Wavelength Zones are AWS infrastructure deployments that embed AWS compute and storage services within telecommunications providers' data centers at the edge of the 5G network, so application traffic can reach application servers running in Wavelength Zones without leaving the mobile providers' network. This prevents the latency that would result from multiple hops to the internet and enables customers to take full advantage of 5G networks. Wavelength Zones extend AWS to the 5G edge, delivering a consistent developer experience across multiple 5G networks around the world. Wavelength Zones also allow developers to build the next generation of ultra-low latency applications using the same familiar AWS services, APIs, tools, and functionality they already use today.



Amazon EKS uses IAM to provide authentication to your Kubernetes cluster, but it still relies on native Kubernetes Role-Based Access Control (RBAC) for authorization. This means that IAM is only used for the authentication of valid IAM entities. All permissions for interacting with your Amazon EKS cluster's Kubernetes API are managed through the native Kubernetes RBAC system.

Access to your cluster using AWS Identity and Access Management (IAM) entities is enabled by the AWS IAM Authenticator for Kubernetes, which runs on the Amazon EKS control plane. The authenticator gets its configuration information from the `aws-auth` `ConfigMap` (AWS authenticator configuration map).

The `aws-auth` `ConfigMap` is automatically created and applied to your cluster when you create a managed node group or when you create a node group using `eksctl`. It is initially created to allow nodes to join your cluster, but you also use this `ConfigMap` to add role-based access control (RBAC) access to IAM users and roles.

Hence, the correct answer is: **Launch the application to an Amazon Elastic Kubernetes Service (Amazon EKS) cluster. Create node groups in Wavelength Zones for the Amazon EKS cluster via the AWS Wavelength service. Apply the AWS authenticator configuration map (`aws-auth ConfigMap`) to your cluster.**

The option that says: **Host the application to an Amazon Elastic Kubernetes Service (Amazon EKS) cluster. Set up node groups in AWS Wavelength Zones for the Amazon EKS cluster. Attach the Amazon EKS connector agent role (`AmazonECSConnectorAgentRole`) to your cluster and use AWS Control Tower for RBAC access** is incorrect. An Amazon EKS connector agent is only used to connect your externally hosted Kubernetes clusters and to allow them to be viewed in your AWS Management Console. The AWS Control Tower doesn't provide RBAC access too to your EKS cluster. This service is commonly used for setting up a secure multi-account AWS environment and not for providing cluster authentication using IAM users and roles.

The option that says: **Launch the application to an Amazon Elastic Kubernetes Service (Amazon EKS) cluster. Create VPC endpoints for the AWS Wavelength Zones and apply them to the Amazon EKS cluster. Install the AWS IAM Authenticator for Kubernetes (`aws-iam-authenticator`) to your cluster** is incorrect because you cannot create VPC Endpoints in AWS Wavelength Zones. In addition, it is more appropriate to apply the AWS authenticator configuration map (`aws-auth ConfigMap`) to your Amazon EKS cluster to enable RBAC access.

The option that says: **Host the application to an Amazon EKS cluster and run the Kubernetes pods on AWS Fargate. Create node groups in AWS Wavelength Zones for the Amazon EKS cluster. Add the EKS pod execution IAM role (`AmazonEKSFargatePodExecutionRole`) to your cluster and ensure that the Fargate profile has the same IAM role as your Amazon EC2 node groups** is incorrect. Although this solution is possible, the security configuration of the Amazon EKS control plane is wrong. You have to ensure that the Fargate profile has a different IAM role as your Amazon EC2 node groups and not the other way around.

References:

https://aws.amazon.com/wavelength/

https://docs.aws.amazon.com/eks/latest/userguide/add-user-role.html#aws-auth-configmap

https://docs.aws.amazon.com/eks/latest/userguide/cluster-auth.html

Question 47: **Incorrect**
A company has 10 TB of infrequently accessed financial data files that would need to be stored in AWS. These data would be accessed infrequently during specific weeks when they are retrieved for auditing purposes. The retrieval time is not strict as long as it does not exceed 24 hours.

Which of the following would be a secure, durable, and cost-effective solution for this scenario?

- **Upload the data to S3 then use a lifecycle policy to transfer data to S3 One Zone-IA.**

- **Upload the data to S3 then use a lifecycle policy to transfer data to S3-IA.**

  (Incorrect)

- **Upload the data to S3 and set a lifecycle policy to transition data to Glacier after `0` days.**

  (Correct)

- **Upload the data to Amazon FSx for Windows File Server using the Server Message Block (SMB) protocol.**

Explanation
Glacier is a cost-effective archival solution for large amounts of data. Bulk retrievals are S3 Glacier's lowest-cost retrieval option, enabling you to retrieve large amounts, even petabytes, of data inexpensively in a day. Bulk retrievals typically complete within 5 – 12 hours. You can specify an absolute or relative time period (including 0 days) after which the specified Amazon S3 objects should be transitioned to Amazon Glacier.

Hence, the correct answer is the option that says: **Upload the data to S3 and set a lifecycle policy to transition data to Glacier after `0` days.**

Glacier has a management console that you can use to create and delete vaults. However, you cannot directly upload archives to Glacier by using the management console. To upload data such as photos, videos, and other documents, you must either use the AWS CLI or write code to make requests by using either the REST API directly or by using the AWS SDKs.

Take note that uploading data to the S3 Console and setting its storage class of "Glacier" is a different story as the proper way to upload data to Glacier is still via its API or CLI. In this way, you can set up your vaults and configure your retrieval options. If you

uploaded your data using the S3 console then it will be managed via S3 even though it is internally using a Glacier storage class.

Uploading the data to S3 then using a lifecycle policy to transfer data to S3-IA is incorrect because using Glacier would be a more cost-effective solution than using S3-IA. Since the required retrieval period should not exceed more than a day, Glacier would be the best choice.

Uploading the data to Amazon FSx for Windows File Server using the Server Message Block (SMB) protocol is incorrect because this option costs more than Amazon Glacier, which is more suitable for storing infrequently accessed data. Amazon FSx for Windows File Server provides fully managed, highly reliable, and scalable file storage that is accessible over the industry-standard Server Message Block (SMB) protocol.

Uploading the data to S3 then using a lifecycle policy to transfer data to S3 One Zone-IA is incorrect because with S3 One Zone-IA, the data will only be stored in a single availability zone and thus, this storage solution is not durable. It also costs more compared to Glacier.

References:

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

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

https://docs.aws.amazon.com/amazonglacier/latest/dev/uploading-an-archive.html

Amazon S3 and S3 Glacier Overview:

https://www.youtube.com/watch?v=1ymyeN2tki4

Check out this Amazon S3 Glacier Cheat Sheet:

https://tutorialsdojo.com/amazon-glacier/

Question 48: **Correct**
A company plans to design a highly available architecture in AWS. They have two target groups with three EC2 instances each, which are added to an Application Load Balancer. In the security group of the EC2 instance, you have verified that port 80 for

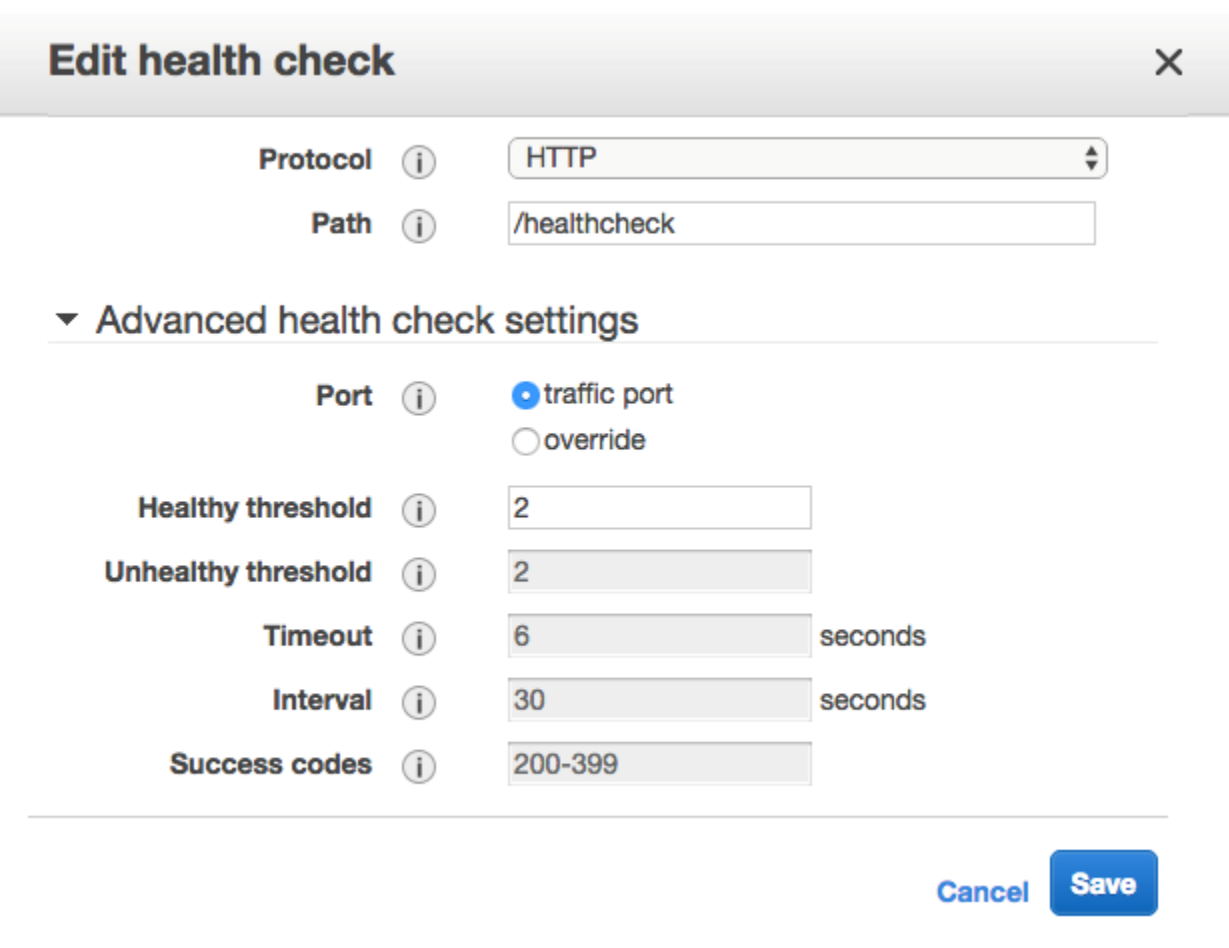
HTTP is allowed. However, the instances are still showing out of service from the load balancer.

What could be the root cause of this issue?

- **The wrong subnet was used in your VPC**
- **The health check configuration is not properly defined.**

  (Correct)
- **The wrong instance type was used for the EC2 instance.**
- **The instances are using the wrong AMI.**

Explanation
Since the security group is properly configured, the issue may be caused by a wrong health check configuration in the Target Group.



Your **Application Load Balancer** periodically sends requests to its registered targets to test their status. These tests are called *health checks*. Each load balancer node routes requests only to the healthy targets in the enabled Availability Zones for the load balancer. Each load balancer node checks the health of each target, using the health check settings for the target group with which the target is registered. After your target is registered, it must pass one health check to be considered healthy. After each health check is completed, the load balancer node closes the connection that was established for the health check.

Reference:

http://docs.aws.amazon.com/elasticloadbalancing/latest/classic/elb-healthchecks.html

AWS Elastic Load Balancing Overview:

https://www.youtube.com/watch?v=UBI5dw59DO8

Check out this AWS Elastic Load Balancing (ELB) Cheat Sheet:

https://tutorialsdojo.com/aws-elastic-load-balancing-elb/

ELB Health Checks vs Route 53 Health Checks For Target Health Monitoring:

https://tutorialsdojo.com/elb-health-checks-vs-route-53-health-checks-for-target-health-monitoring/

Question 49: **Correct**
An organization plans to run an application in a dedicated physical server that doesn't use virtualization. The application data will be stored in a storage solution that uses an NFS protocol. To prevent data loss, you need to use a durable cloud storage service to store a copy of your data.

Which of the following is the most suitable solution to meet the requirement?

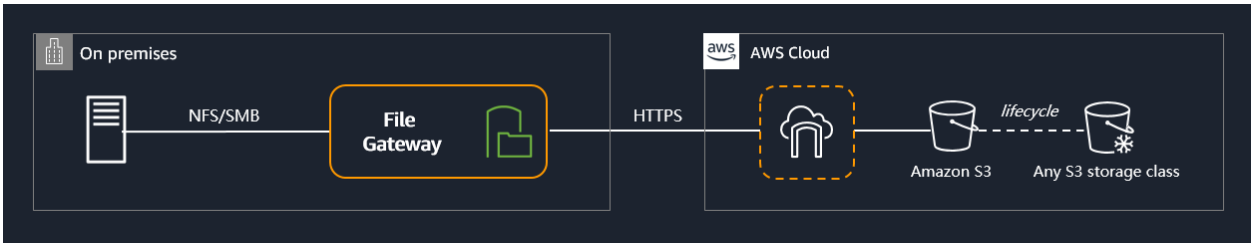
- **Use AWS Storage Gateway with a gateway VM appliance for your compute resources. Configure File Gateway to store the application data and backup data.**
- **Use an AWS Storage Gateway hardware appliance for your compute resources. Configure File Gateway to store the application data and create an Amazon S3 bucket to store a backup of your data.**

  (Correct)
- **Use an AWS Storage Gateway hardware appliance for your compute resources. Configure Volume Gateway to store the application data and create an Amazon S3 bucket to store a backup of your data.**

Use an AWS Storage Gateway hardware appliance for your compute resources. Configure Volume Gateway to store the application data and backup data.

Explanation

AWS Storage Gateway is a hybrid cloud storage service that gives you on-premises access to virtually unlimited cloud storage by linking it to S3. Storage Gateway provides 3 types of storage solutions for your on-premises applications: file, volume, and tape gateways. The AWS Storage Gateway Hardware Appliance is a physical, standalone, validated server configuration for on-premises deployments.



The AWS Storage Gateway Hardware Appliance is a physical hardware appliance with the Storage Gateway software preinstalled on a validated server configuration. The hardware appliance is a high-performance 1U server that you can deploy in your data center or on-premises inside your corporate firewall. When you buy and activate your hardware appliance, the activation process associates your hardware appliance with your AWS account. After activation, your hardware appliance appears in the console as a gateway on the *Hardware* page. You can configure your hardware appliance as a file gateway, tape gateway, or volume gateway type. The procedure that you use to deploy and activate these gateway types on a hardware appliance is the same as on a virtual platform.

Since the company needs to run a dedicated physical appliance, you can use an AWS Storage Gateway Hardware Appliance. It comes pre-loaded with Storage Gateway software and provides all the required resources to create a file gateway. A file gateway can be configured to store and retrieve objects in Amazon S3 using the protocols NFS and SMB.

Hence, the correct answer in this scenario is: **Use an AWS Storage Gateway hardware appliance for your compute resources. Configure File Gateway to store the application data and create an Amazon S3 bucket to store a backup of your data**.

The option that says: **Use AWS Storage Gateway with a gateway VM appliance for your compute resources. Configure File Gateway to store the application data and backup data** is incorrect because as per the scenario, the company needs to use an on-premises hardware appliance and not just a Virtual Machine (VM).

The options that say: **Use an AWS Storage Gateway hardware appliance for your compute resources. Configure Volume Gateway to store the application data and backup data** and **Use an AWS Storage Gateway hardware appliance for your compute**

resources. Configure Volume Gateway to store the application data and create an Amazon S3 bucket to store a backup of your data are both incorrect. As per the scenario, the requirement is a file system that uses an NFS protocol and not iSCSI devices. Among the AWS Storage Gateway storage solutions, only file gateway can store and retrieve objects in Amazon S3 using the protocols NFS and SMB.

References:

https://docs.aws.amazon.com/storagegateway/latest/userguide/hardware-appliance.html

https://docs.aws.amazon.com/storagegateway/latest/userguide/WhatIsStorageGateway.html

AWS Storage Gateway Overview:

https://youtu.be/pNb7xOBJjHE

Check out this AWS Storage Gateway Cheat Sheet:

https://tutorialsdojo.com/aws-storage-gateway/

Question 50: **Correct**
A company has recently adopted a hybrid cloud architecture and is planning to migrate a database hosted on-premises to AWS. The database currently has over 50 TB of consumer data, handles highly transactional (OLTP) workloads, and is expected to grow. The Solutions Architect should ensure that the database is ACID-compliant and can handle complex queries of the application.

Which type of database service should the Architect use?

- **Amazon Aurora**

  (Correct)

- **Amazon RDS**

- **Amazon Redshift**
- **Amazon DynamoDB**

Explanation

Amazon Aurora (Aurora) is a fully managed relational database engine that's compatible with MySQL and PostgreSQL. You already know how MySQL and PostgreSQL combine the speed and reliability of high-end commercial databases with the simplicity and cost-effectiveness of open-source databases. The code, tools, and applications you use today with your existing MySQL and PostgreSQL databases can be used with Aurora. With some workloads, Aurora can deliver up to five times the throughput of MySQL and up to three times the throughput of PostgreSQL without requiring changes to most of your existing applications.

Aurora includes a high-performance storage subsystem. Its MySQL- and PostgreSQL-compatible database engines are customized to take advantage of that fast distributed storage. The underlying storage grows automatically as needed, up to 64 tebibytes (TiB). Aurora also automates and standardizes database clustering and replication, which are typically among the most challenging aspects of database configuration and administration.

	Relational databases	NoSQL databases
Optimal workloads	Relational databases are designed for transactional and strongly consistent online transaction processing (OLTP) applications and are good for online analytical processing (OLAP).	NoSQL key-value, document, graph, and in-memory databases are designed for OLTP for a number of data access patterns that include low-latency applications. NoSQL search databases are designed for analytics over semi-structured data.
Data model	The relational model normalizes data into tables that are composed of rows and columns. A schema strictly defines the tables, rows, columns, indexes, relationships between tables, and other database elements. The database enforces the referential integrity in relationships between tables.	NoSQL databases provide a variety of data models that includes document, graph, key-value, in-memory, and search.
ACID properties	Relational databases provide atomicity, consistency, isolation, and durability (ACID) properties: • Atomicity requires a transaction to execute completely or not at all. • Consistency requires that when a transaction has been committed, the data must conform to the database schema. • Isolation requires that concurrent transactions execute separately from each other. • Durability requires the ability to recover from an unexpected system failure or power outage to the last known state.	NoSQL databases often make tradeoffs by relaxing some of the ACID properties of relational databases for a more flexible data model that can scale horizontally. This makes NoSQL databases an excellent choice for high throughput, low-latency use cases that need to scale horizontally beyond the limitations of a single instance.
Performance	Performance is generally dependent on the disk subsystem. The optimization of queries, indexes, and table structure is often required to achieve peak performance.	Performance is generally a function of the underlying hardware cluster size, network latency, and the calling application.
Scale	Relational databases typically scale up by increasing the compute capabilities of the hardware or scale-out by adding replicas for read-only workloads.	NoSQL databases typically are partitionable because key-value access patterns are able to scale out by using distributed architecture to increase throughput that provides consistent performance at near boundless scale.
APIs	Requests to store and retrieve data are communicated using queries that conform to a structured query language (SQL). These queries are parsed and executed by the relational database.	Object-based APIs allow app developers to easily store and retrieve in-memory data structures. Partition keys let apps look up key-value pairs, column sets, or semistructured documents that contain serialized app objects and attributes.

For Amazon RDS MariaDB DB instances, the maximum provisioned storage limit constrains the size of a table to a maximum size of 64 TB when using InnoDB file-per-table tablespaces. This limit also constrains the system tablespace to a maximum size of 16 TB. InnoDB file-per-table tablespaces (with tables each in their own tablespace) is set by default for Amazon RDS MariaDB DB instances.

Hence, the correct answer is **Amazon Aurora.**

Amazon Redshift is incorrect because this is primarily used for OLAP applications and not for OLTP. Moreover, it doesn't scale automatically to handle the exponential growth of the database.

Amazon DynamoDB is incorrect. Although you can use this to have an ACID-compliant database, it is not capable of handling complex queries and highly transactional (OLTP) workloads.

Amazon RDS is incorrect. Although this service can host an ACID-compliant relational database that can handle complex queries and transactional (OLTP) workloads, it is still not scalable to handle the growth of the database. Amazon Aurora is the better choice as its underlying storage can grow automatically as needed.

References:

https://aws.amazon.com/rds/aurora/

https://docs.aws.amazon.com/amazondynamodb/latest/developerguide/SQLtoNoSQL.html

https://aws.amazon.com/nosql/

Amazon Aurora Overview:

https://youtu.be/iwS1h7rLNBQ

Check out this Amazon Aurora Cheat Sheet:

https://tutorialsdojo.com/amazon-aurora/

Question 51: **Correct**
A company needs to accelerate the development of its GraphQL APIs for its new customer service portal. The solution must be serverless to lower the monthly operating cost of the business. Their GraphQL APIs must be accessible via HTTPS and have a custom domain.

What solution should the Solutions Architect implement to meet the above requirements?

- **Develop the application using the AWS AppSync service and use its built-in custom domain feature. Associate an SSL certificate to the AWS AppSync API using the AWS Certificate Manager (ACM) service to enable HTTPS communication.**

  (Correct)

- **Launch an AWS Elastic Beanstalk environment and use Amazon Route 53 for the custom domain. Configure Domain Name System Security Extensions (DNSSEC) in the Route 53 hosted zone to enable HTTPS communication.**

-

Host the application in the VMware Cloud on AWS service. Associate a custom domain to the GraphSQL APIs via the AWS Directory Service for Microsoft Active Directory and provide multiple domain controllers to enable HTTPS communication.

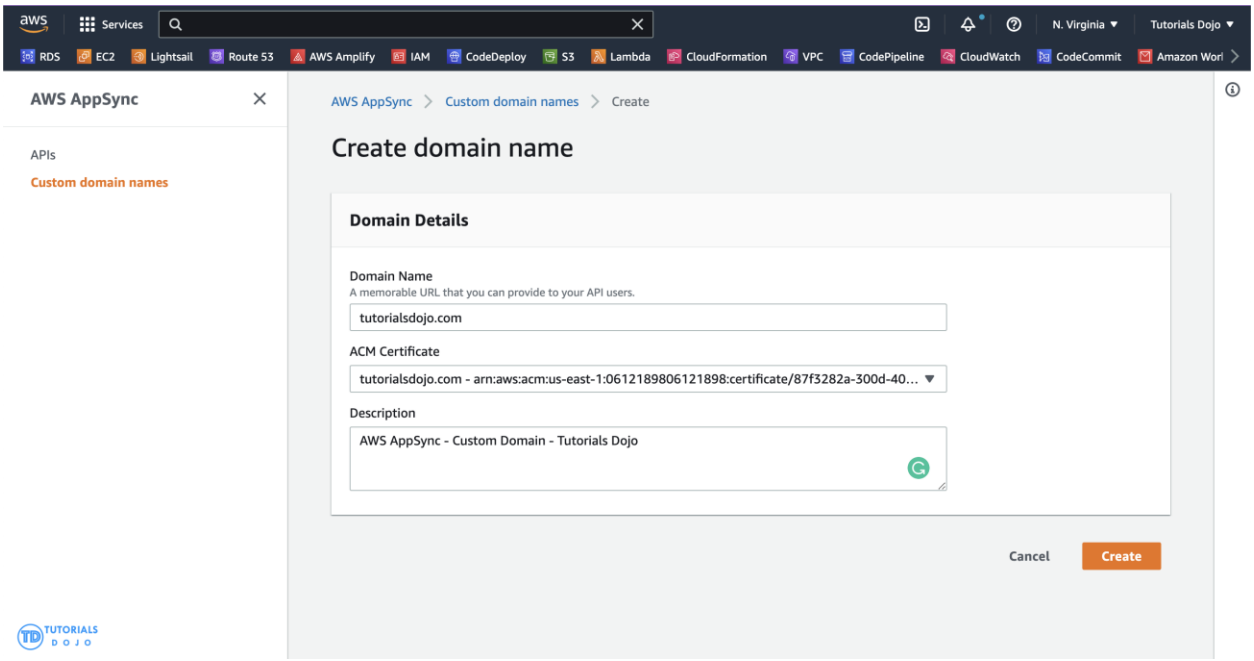
-

Deploy the GraphQL APIs as Kubernetes pods to AWS Fargate and AWS Outposts using Amazon EKS Anywhere for deployment. Create a custom domain using Amazon CloudFront and enable the Origin Shield feature to allow HTTPS communication to the GraphQL APIs.

Explanation

AWS AppSync is a serverless GraphQL and Pub/Sub API service that simplifies building modern web and mobile applications. It provides a robust, scalable GraphQL interface for application developers to combine data from multiple sources, including Amazon DynamoDB, AWS Lambda, and HTTP APIs.

GraphQL is a data language to enable client apps to fetch, change and subscribe to data from servers. In a GraphQL query, the client specifies how the data is to be structured when it is returned by the server. This makes it possible for the client to query only for the data it needs, in the format that it needs it in.



With AWS AppSync, you can use custom domain names to configure a single, memorable domain that works for both your GraphQL and real-time APIs.

In other words, you can utilize simple and memorable endpoint URLs with domain names of your choice by creating custom domain names that you associate with the AWS AppSync APIs in your account.

When you configure an AWS AppSync API, two endpoints are provisioned:

AWS AppSync GraphQL endpoint:`https://example1234567890000.appsync-api.us-east-1.amazonaws.com/graphql`**AWS AppSync real-time endpoint:**`wss://example1234567890000.appsync-realtime-api.us-east-1.amazonaws.com/graphql`

Hence, the correct answer is: **Develop the application using the AWS AppSync service and use its built-in custom domain feature. Associate an SSL certificate to the AWS AppSync API using the AWS Certificate Manager (ACM) service to enable HTTPS communication.**

The option that says: **Launch an AWS Elastic Beanstalk environment and use Amazon Route 53 for the custom domain. Configure Domain Name System Security Extensions (DNSSEC) in the Route 53 hosted zone to enable HTTPS communication** is incorrect because the AWS Elastic Beanstalk service is not a serverless solution. This will launch Amazon EC2 instances in your AWS account for your application. Take note that the requirements explicitly mentioned that the solution should be serverless. In addition, the primary function of the DNSSEC feature is to authenticate the responses of domain name lookups and not for HTTPS communication.

The option that says: **Host the application in the VMware Cloud on AWS service. Associate a custom domain to the GraphSQL APIs via the AWS Directory Service for Microsoft Active Directory and provide multiple domain controllers to enable HTTPS communication** is incorrect. The VMware Cloud on AWS is only a service for vSphere-based workloads and not for GraphQL use cases. Moreover, the main use case for AWS Directory Service is to enable your directory-aware workloads and AWS resources to use managed Active Directory (AD) in AWS and not for HTTPS communication.

The option that says: **Deploy the GraphQL APIs as Kubernetes pods to AWS Fargate and AWS Outposts using Amazon EKS Anywhere for deployment. Create a custom domain using Amazon CloudFront and enable the Origin Shield feature to allow HTTPS communication to the GraphQL APIs** is incorrect. Although the AWS Fargate service is serverless, the AWS Outposts service is not. Furthermore, the Origin Shield feature in Amazon CloudFront is simply a centralized caching layer that helps increase your cache hit ratio which effectively reduces the load on your origin. A better solution is to use AWS AppSync and use its built-in custom domain.

References:

https://docs.aws.amazon.com/appsync/latest/devguide/custom-domain-name.html

https://docs.aws.amazon.com/appsync/latest/devguide/what-is-appsync.html

https://aws.amazon.com/appsync/

Question 52: **Correct**
A business plans to deploy an application on EC2 instances within an Amazon VPC and is considering adopting a Network Load Balancer to distribute incoming traffic among the instances. A solutions architect needs to suggest a solution that will enable the security team to inspect traffic entering and exiting their VPC.

Which approach satisfies the requirements?

- **Create a firewall using the AWS Network Firewall service at the VPC level then add custom rule groups for inspecting ingress and egress traffic. Update the necessary VPC route tables.**

  (Correct)

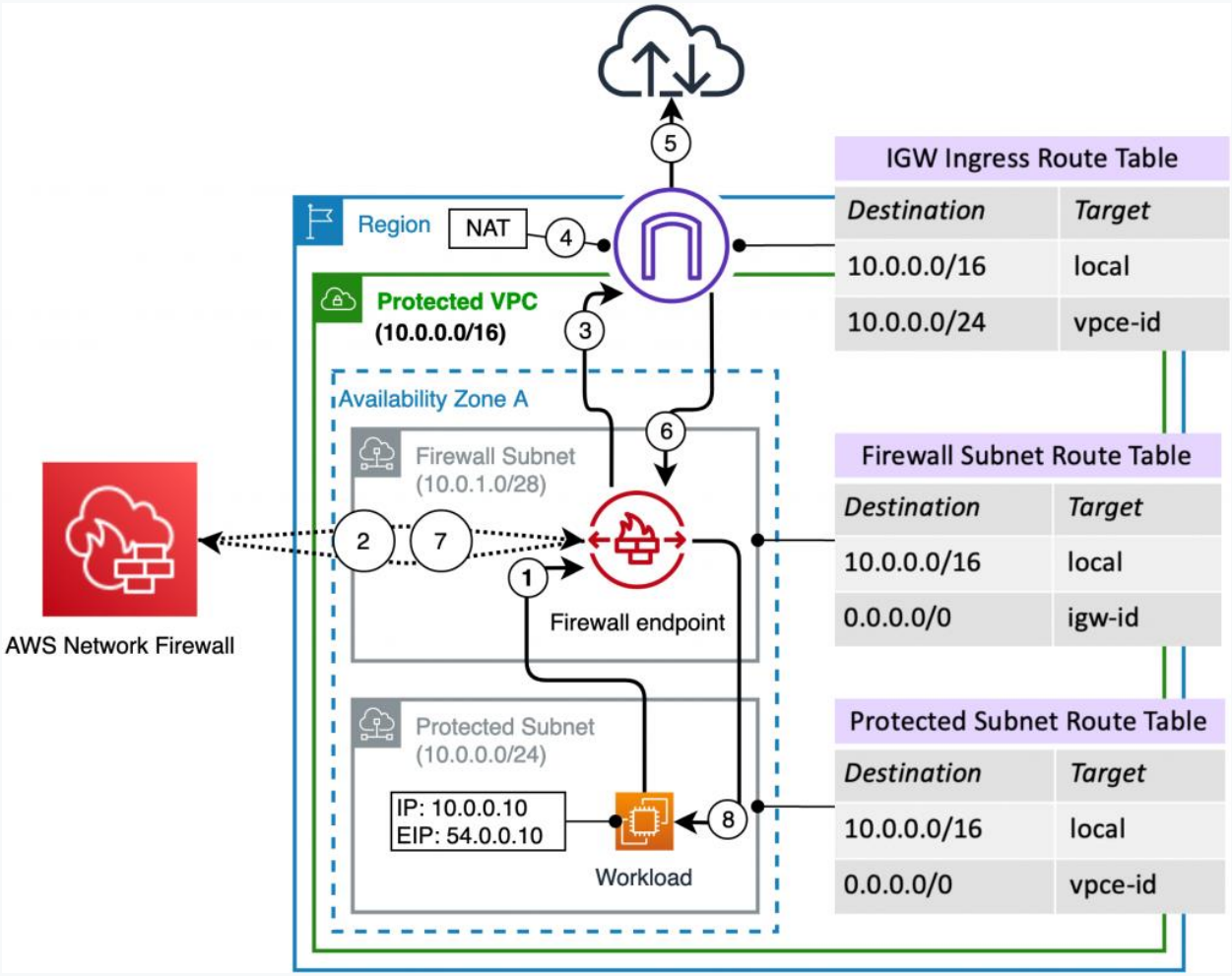
- **Create a firewall at the subnet level using the Amazon Detective service. Inspect the ingress and egress traffic using the VPC Reachability Analyzer.**

- **Enable Traffic Mirroring on the Network Load Balancer and forward traffic to the instances. Create a traffic mirror filter to inspect the ingress and egress of data that traverses your Amazon VPC.**

- **Use the Network Access Analyzer service on the application's VPC for inspecting ingress and egress traffic. Create a new Network Access Scope to filter and analyze all incoming and outgoing requests.**

Explanation
AWS Network Firewall is a stateful, managed, network firewall, and intrusion detection and prevention service for your virtual private cloud (VPC). With Network Firewall, you can filter traffic at the perimeter of your VPC. This includes traffic going to and coming from an internet gateway, NAT gateway, or over VPN or AWS Direct Connect. Network Firewall uses Suricata — an open-source intrusion prevention system (IPS) for stateful inspection.

The diagram below shows an AWS Network firewall deployed in a single availability zone and traffic flow for a workload in a public subnet:



You can use Network Firewall to monitor and protect your Amazon VPC traffic in a number of ways, including the following:

- Pass traffic through only from known AWS service domains or IP address endpoints, such as Amazon S3.

- Use custom lists of known bad domains to limit the types of domain names that your applications can access.

- Perform deep packet inspection on traffic entering or leaving your VPC.

- Use stateful protocol detection to filter protocols like HTTPS, independent of the port used.

Therefore, the correct answer is: **Create a firewall using the AWS Network Firewall service at the VPC level then add custom rule groups for inspecting ingress and egress traffic. Update the necessary VPC route tables.**

The option that says: **Use the Network Access Analyzer service on the application's VPC for inspecting ingress and egress traffic. Create a new Network Access Scope to filter and analyze all incoming and outgoing requests** is incorrect. Network Access Analyzer is a feature of VPC that reports on unintended access to your AWS resources based on the security and compliance that you set. This service is not capable of performing deep packet inspection on traffic entering or leaving your VPC, unlike AWS Network Firewall.

The option that says: **Create a firewall at the subnet level using the Amazon Detective service. Inspect the ingress and egress traffic using the VPC Reachability Analyzer** is incorrect because a firewall must be created at the VPC level and not at the subnet level. Moreover, Amazon Detective can't be used to create a firewall. This service just automatically collects log data from your AWS resources to analyze, investigate, and quickly identify the root cause of potential security issues or suspicious activities in your AWS account. For this scenario, you have to use the AWS Network Firewall instead.

The option that says: **Enable Traffic Mirroring on the Network Load Balancer and forward traffic to the instances. Create a traffic mirror filter to inspect the ingress and egress of data that traverses your Amazon VPC** is incorrect as this alone accomplishes nothing. It would make more sense if you redirect the traffic to an EC2 instance where an Intrusion Detection System (IDS) is running. Remember that Traffic Mirroring is simply an Amazon VPC feature that you can use to copy network traffic from an elastic network interface. Traffic mirror filters can't inspect the actual packet of the incoming and outgoing traffic.

References:

https://aws.amazon.com/blogs/networking-and-content-delivery/deployment-models-for-aws-network-firewall/

https://docs.aws.amazon.com/network-firewall/latest/developerguide/what-is-aws-network-firewall.html

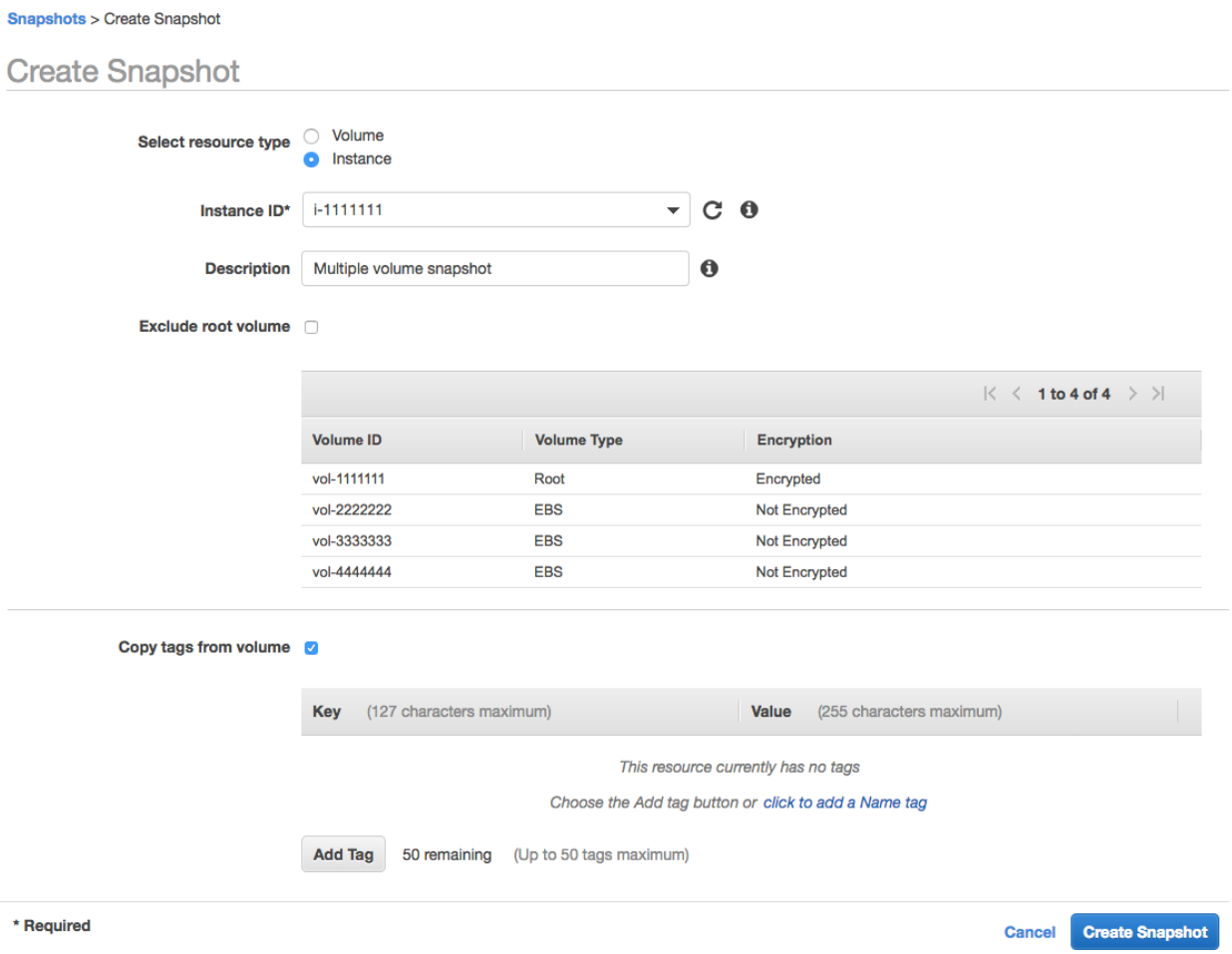
Question 53: **Correct**
A company has an On-Demand EC2 instance with an attached EBS volume. There is a scheduled job that creates a snapshot of this EBS volume every midnight at 12 AM when the instance is not used. One night, there has been a production incident where you need to perform a change on both the instance and on the EBS volume at the same time when the snapshot is currently taking place.

Which of the following scenario is true when it comes to the usage of an EBS volume while the snapshot is in progress?

- **The EBS volume can be used in read-only mode while the snapshot is in progress.**
- **The EBS volume can be used while the snapshot is in progress.**

  (Correct)
- **The EBS volume cannot be used until the snapshot completes.**
- **The EBS volume cannot be detached or attached to an EC2 instance until the snapshot completes**

Explanation
Snapshots occur asynchronously; the point-in-time snapshot is created immediately, but the status of the snapshot is `pending` until the snapshot is complete (when all of the modified blocks have been transferred to Amazon S3), which can take several hours for large initial snapshots or subsequent snapshots where many blocks have changed.



While it is completing, an in-progress snapshot is not affected by ongoing reads and writes to the volume hence, you can still use the EBS volume normally.

When you create an EBS volume based on a snapshot, the new volume begins as an exact replica of the original volume that was used to create the snapshot. The replicated volume loads data lazily in the background so that you can begin using it immediately. If you access data that hasn't been loaded yet, the volume immediately downloads the requested data from Amazon S3 and then continues loading the rest of the volume's data in the background.

References:

https://docs.aws.amazon.com/AWSEC2/latest/UserGuide/ebs-creating-snapshot.html

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

Check out this Amazon EBS Cheat Sheet:

https://tutorialsdojo.com/amazon-ebs/

Question 54: **Correct**
A Solutions Architect needs to ensure that all of the AWS resources in Amazon VPC don't go beyond their respective service limits. The Architect should prepare a system that provides real-time guidance in provisioning resources that adheres to the AWS best practices.

Which of the following is the MOST appropriate service to use to satisfy this task?

- **AWS Cost Explorer**
- **AWS Budgets**
- **Amazon Inspector**
- **AWS Trusted Advisor**

  (Correct)

Explanation
AWS Trusted Advisor is an online tool that provides you with real-time guidance to help you provision your resources following AWS best practices. It inspects your AWS environment and makes recommendations for saving money, improving system performance and reliability, or closing security gaps.

Whether establishing new workflows, developing applications, or as part of ongoing improvement, take advantage of the recommendations provided by Trusted Advisor on a regular basis to help keep your solutions provisioned optimally.

Trusted Advisor includes an ever-expanding list of checks in the following five categories:

Cost Optimization – recommendations that can potentially save you money by highlighting unused resources and opportunities to reduce your bill.

Security – identification of security settings that could make your AWS solution less secure.

Fault Tolerance – recommendations that help increase the resiliency of your AWS solution by highlighting redundancy shortfalls, current service limits, and over-utilized resources.

Performance – recommendations that can help to improve the speed and responsiveness of your applications.

Service Limits – recommendations that will tell you when service usage is more than 80% of the service limit.

Hence, the correct answer in this scenario is **AWS Trusted Advisor.**

AWS Cost Explorer is incorrect because this is just a tool that enables you to view and analyze your costs and usage. You can explore your usage and costs using the main graph, the Cost Explorer cost and usage reports, or the Cost Explorer RI reports. It has an easy-to-use interface that lets you visualize, understand, and manage your AWS costs and usage over time.

AWS Budgets is incorrect because it simply 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. You can also use AWS Budgets to set reservation utilization or

coverage targets and receive alerts when your utilization drops below the threshold you define.

Amazon Inspector is incorrect because it is just an automated security assessment service that helps improve the security and compliance of applications deployed on AWS. Amazon Inspector automatically assesses applications for exposure, vulnerabilities, and deviations from best practices.

References:

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

https://aws.amazon.com/premiumsupport/technology/trusted-advisor/faqs/

Check out this AWS Trusted Advisor Cheat Sheet:

https://tutorialsdojo.com/aws-trusted-advisor/

Question 55: **Correct**
A company has both on-premises data center as well as AWS cloud infrastructure. They store their graphics, audios, videos, and other multimedia assets primarily in their on-premises storage server and use an S3 Standard storage class bucket as a backup. Their data is heavily used for only a week (7 days) but after that period, it will only be infrequently used by their customers. The Solutions Architect is instructed to save storage costs in AWS yet maintain the ability to fetch a subset of their media assets in a matter of minutes for a surprise annual data audit, which will be conducted on their cloud storage.

Which of the following are valid options that the Solutions Architect can implement to meet the above requirement? (Select TWO.)

- ☐ **Set a lifecycle policy in the bucket to transition the data to S3 - One Zone-Infrequent Access storage class after one week (7 days).**
- ☑ **Set a lifecycle policy in the bucket to transition to S3 - Standard IA after 30 days**

(Correct)

- ☑

  Set a lifecycle policy in the bucket to transition the data to Glacier after one week (7 days).

  (Correct)

- ☐

  Set a lifecycle policy in the bucket to transition the data to S3 - Standard IA storage class after one week (7 days).

- ☐

  Set a lifecycle policy in the bucket to transition the data to S3 Glacier Deep Archive storage class after one week (7 days).

Explanation

You can add rules in a lifecycle configuration to tell Amazon S3 to transition objects to another Amazon S3 storage class. For example: When you know that objects are infrequently accessed, you might transition them to the STANDARD_IA storage class. Or transition your data to the GLACIER storage class in case you want to archive objects that you don't need to access in real-time.

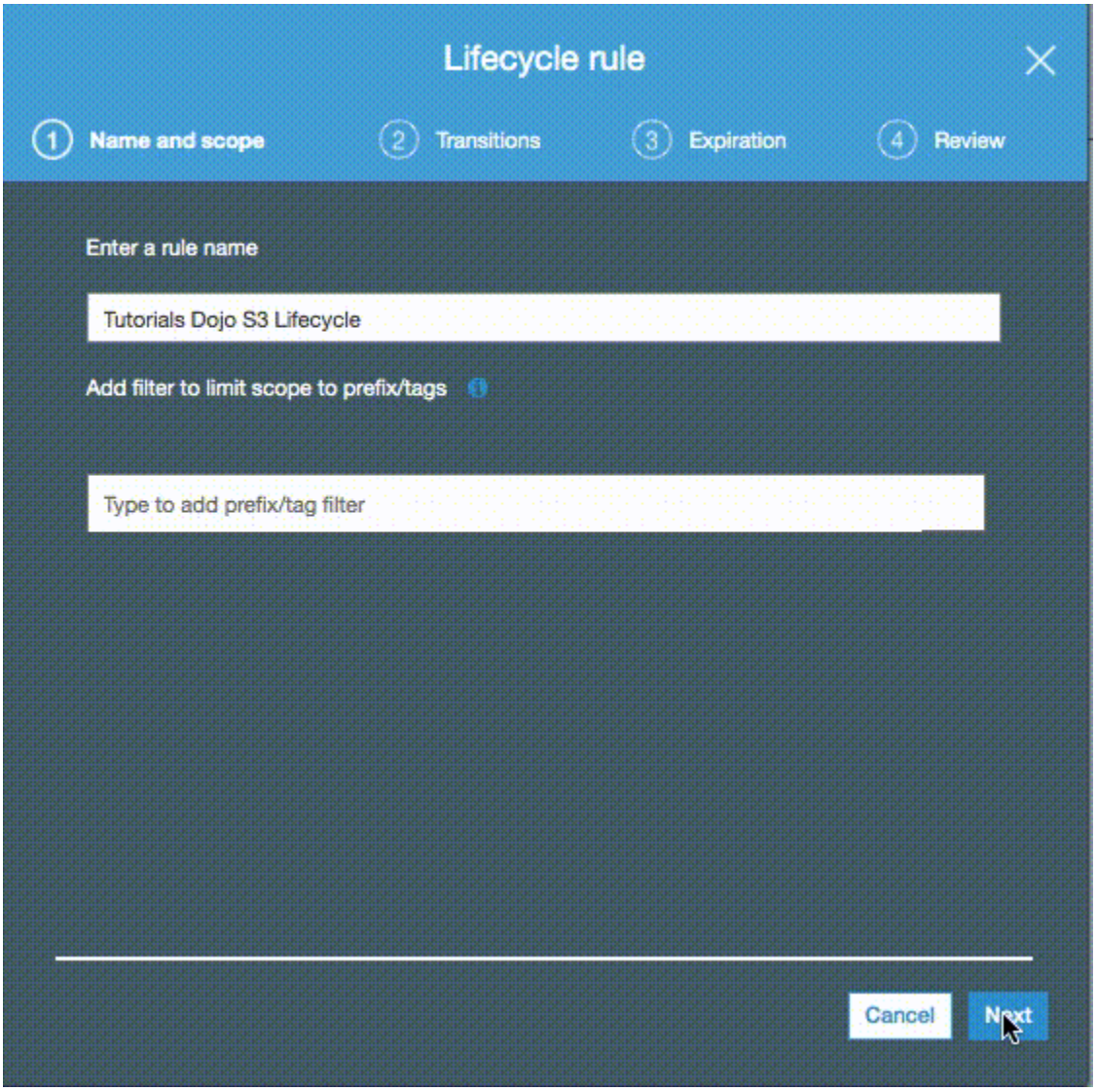
In a lifecycle configuration, you can define rules to transition objects from one storage class to another to save on storage costs. When you don't know the access patterns of your objects or your access patterns are changing over time, you can transition the objects to the INTELLIGENT_TIERING storage class for automatic cost savings.

The lifecycle storage class transitions have a constraint when you want to transition from the STANDARD storage classes to either STANDARD_IA or ONEZONE_IA. The following constraints apply:

- For larger objects, there is a cost-benefit for transitioning to STANDARD_IA or ONEZONE_IA. Amazon S3 does not transition objects that are smaller than 128 KB to the STANDARD_IA or ONEZONE_IA storage classes because it's not cost-effective.

- Objects must be stored for **at least 30 days** in the current storage class before you can transition them to STANDARD_IA or ONEZONE_IA. For example, you cannot create a lifecycle rule to transition objects to the 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 STANDARD_IA or ONEZONE_IA storage.

- If you are transitioning noncurrent objects (in versioned buckets), you can transition only objects that are at least 30 days noncurrent to STANDARD_IA or ONEZONE_IA storage.



Since there is a time constraint in transitioning objects in S3, you can only change the storage class of your objects from S3 Standard storage class to STANDARD_IA or ONEZONE_IA storage after 30 days. This limitation does not apply to INTELLIGENT_TIERING, GLACIER, and DEEP_ARCHIVE storage class.

In addition, the requirement says that the media assets should be fetched in a matter of minutes for a surprise **annual** data audit. This means that the retrieval will only happen once a year. You can use expedited retrievals in Glacier which will allow you to quickly access your data (within 1–5 minutes) when occasional urgent requests for a subset of archives are required.

In this scenario, you can set a lifecycle policy in the bucket to transition to S3 - Standard IA after 30 days or alternatively, you can directly transition your data to Glacier after one week (7 days).

Hence, the following are the correct answers:

- Set a lifecycle policy in the bucket to transition the data from Standard storage class to Glacier after one week (7 days).

- Set a lifecycle policy in the bucket to transition to S3 - Standard IA after 30 days.

Setting a lifecycle policy in the bucket to transition the data to S3 - Standard IA storage class after one week (7 days) and **setting a lifecycle policy in the bucket to transition the data to S3 - One Zone-Infrequent Access storage class after one week (7 days)** are both incorrect because there is a constraint in S3 that objects must be stored at least 30 days in the current storage class before you can transition them to STANDARD_IA or ONEZONE_IA. You cannot create a lifecycle rule to transition objects to either STANDARD_IA or ONEZONE_IA storage class 7 days after you create them because you can only do this after the 30-day period has elapsed. Hence, these options are incorrect.

Setting a lifecycle policy in the bucket to transition the data to S3 Glacier Deep Archive storage class after one week (7 days) is incorrect. Although DEEP_ARCHIVE storage class provides the most cost-effective storage option, it does not have the ability to do expedited retrievals, unlike Glacier. In the event that the surprise annual data audit happens, it may take several hours before you can retrieve your data.

References:

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

https://docs.aws.amazon.com/AmazonS3/latest/dev/restoring-objects.html

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

Check out this Amazon S3 Cheat Sheet:

https://tutorialsdojo.com/amazon-s3/

Question 56: **Incorrect**

A company has multiple research departments that have deployed several resources to the AWS cloud. The departments are free to provision their own resources as they are needed. To ensure normal operations, the company wants to track its AWS resource usage so that it is not reaching the AWS service quotas unexpectedly.

Which combination of actions should the Solutions Architect implement to meet the company requirements? (Select TWO.)

- [x] **Query the AWS Trusted Advisor Service Limits check every 24 hours by calling the `DescribeTrustedAdvisorChecks` API operation. Ensure that your AWS account has a Developer support plan.**

  (Incorrect)

- [ ] **Write an AWS Lambda function that refreshes the AWS Trusted Advisor Service Limits checks and set it to run every 24 hours.**

  (Correct)

- [ ] **Capture the events using Amazon EventBridge (Amazon CloudWatch Events) and use an Amazon Simple Notification Service (Amazon SNS) topic as the target for notifications.**

  (Correct)

- [x] **Create an Amazon Simple Notification Service (Amazon SNS) topic and configure it as a target for notifications.**

  (Incorrect)

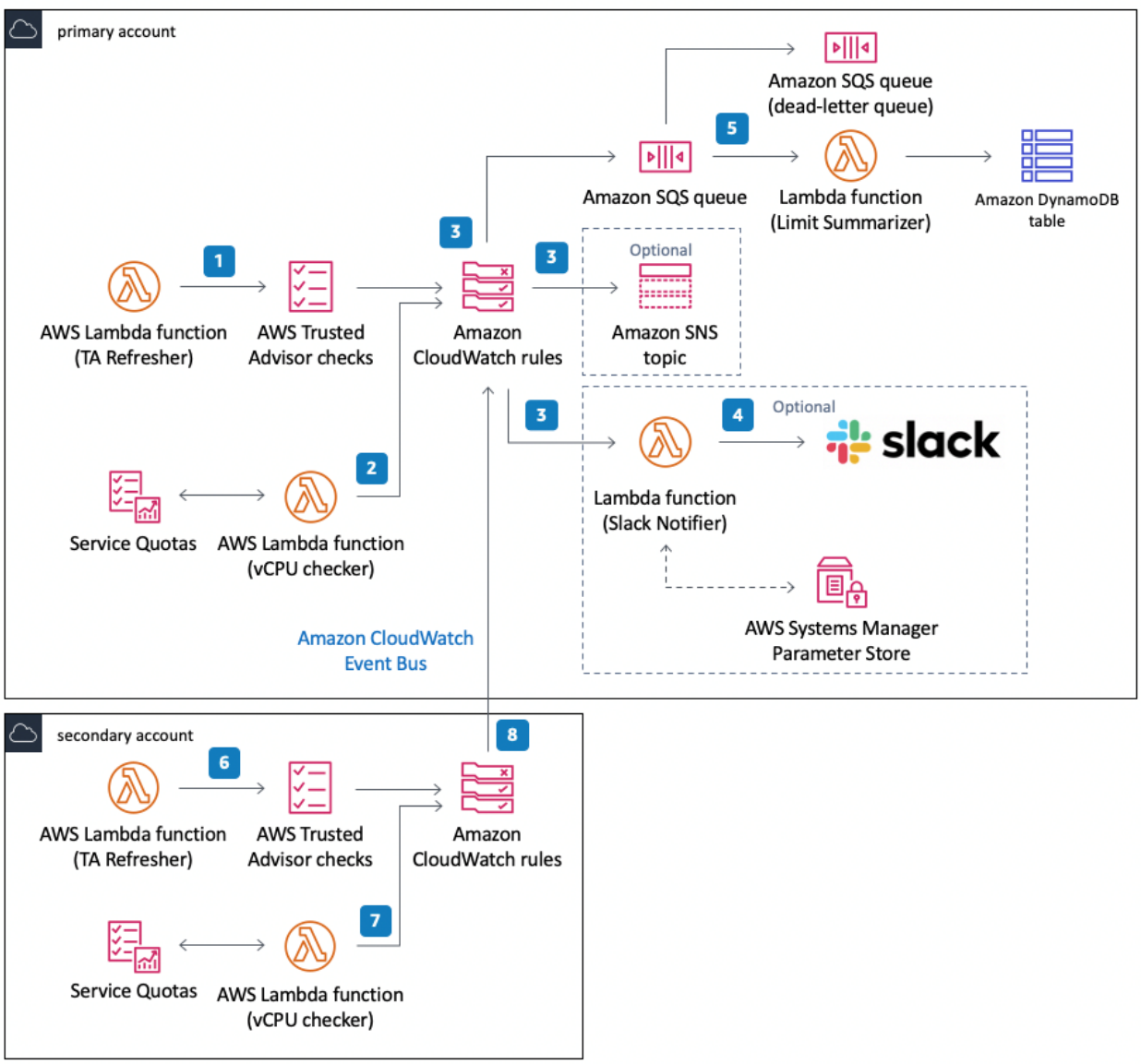
- [ ] **Utilize the AWS managed rule on AWS Config to monitor AWS resource service quotas. Schedule this checking using an AWS Lambda function.**

Explanation

AWS Trusted Advisor draws upon best practices learned from serving hundreds of thousands of AWS customers. Trusted Advisor inspects your AWS environment, and

then makes recommendations when opportunities exist to save money, improve system availability and performance, or help close security gaps. If you have a Basic or Developer Support plan, you can use the Trusted Advisor console to access all checks in the Service Limits category and six checks in the Security category.

AWS has an example of the implementation of Quota Monitor CloudFormation template that you can deploy on your AWS account. The template uses an AWS Lambda function that runs once every 24 hours. The Lambda function refreshes the AWS Trusted Advisor Service Limits checks to retrieve the most current utilization and quota data through API calls. Amazon CloudWatch Events captures the status events from Trusted Advisor. It uses a set of CloudWatch Events rules to send the status events to all the targets you choose during initial deployment of the solution: an Amazon Simple Queue Service (Amazon SQS) queue, an Amazon Simple Notification Service (Amazon SNS) topic or a Lambda function for Slack notifications.



The AWS Trusted Advisor Service limit publishes service limits metric to CloudWatch; thus, you can configure an alarm and send a notification to Amazon SNS. You can also create an AWS Lambda function to read data from specific Trusted Advisor checks. A Lambda function invocation can be scheduled using AWS EventBridge (Amazon CloudWatch Events) to automated the process.

Hence, the following options are correct:

-Capture the events using Amazon EventBridge (Amazon CloudWatch Events) and use an Amazon Simple Notification Service (Amazon SNS) topic as the target for notifications

-Write an AWS Lambda function that refreshes the AWS Trusted Advisor Service Limits checks and set it to run every 24 hours

The option that says: **Create an Amazon Simple Notification Service (Amazon SNS) topic and configure it as a target for notifications** is incorrect. This option is incomplete as it doesn't specify where the notification comes from such as from EventBridge, Lambda functions, etc.

The option that says: **Query the AWS Trusted Advisor Service Limits check every 24 hours by calling the `DescribeTrustedAdvisorChecks` API operation. Ensure that your AWS account has a Developer support plan** is incorrect. This API returns information about all available AWS Trusted Advisor checks, so it will be difficult to extract only "service limits" information from this API call. Moreover, the Trusted Advisor APIs (AWS Support APIs) are only available for Business, Enterprise On-Ramp, or Enterprise Support plans.

The option that says: **Utilize the AWS managed rule on AWS Config to monitor AWS resource service quotas. Schedule this checking using an AWS Lambda function** is incorrect. There is no AWS config "managed rule" that checks for service quotas.

References:

https://aws.amazon.com/solutions/implementations/quota-monitor/

https://aws.amazon.com/blogs/mt/monitoring-service-limits-with-trusted-advisor-and-amazon-cloudwatch/

Check out these Amazon CloudWatch and AWS Trusted Advisor Cheat Sheets:

https://tutorialsdojo.com/amazon-cloudwatch/

https://tutorialsdojo.com/aws-trusted-advisor/

Question 57: **Correct**
In a startup company you are working for, you are asked to design a web application that requires a NoSQL database that has no limit on the storage size for a given table. The startup is still new in the market and it has very limited human resources who can take care of the database infrastructure.

Which is the most suitable service that you can implement that provides a fully managed, scalable and highly available NoSQL service?

- **Amazon Aurora**
- **Amazon Neptune**
- **DynamoDB**

  (Correct)
- **SimpleDB**

Explanation
The term "**fully managed**" means that Amazon will manage the underlying infrastructure of the service hence, you don't need an additional human resource to support or maintain the service. Therefore, Amazon DynamoDB is the right answer. Remember that Amazon RDS is a managed service but not "fully managed" as you still have the option to maintain and configure the underlying server of the database.

Amazon DynamoDB is a fast and flexible NoSQL database service for all applications that need consistent, single-digit millisecond latency at any scale. It is a fully managed cloud database and supports both document and key-value store models. Its flexible data model, reliable performance, and automatic scaling of throughput capacity make it a great fit for mobile, web, gaming, ad tech, IoT, and many other applications.

Amazon Neptune is incorrect because this is primarily used as a graph database.

Amazon Aurora is incorrect because this is a relational database and not a NoSQL database.

SimpleDB is incorrect. Although SimpleDB is also a highly available and scalable NoSQL database, it has a limit on the request capacity or storage size for a given table, unlike DynamoDB.

Reference:

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

Check out this Amazon DynamoDB Cheat Sheet:

https://tutorialsdojo.com/amazon-dynamodb/

Amazon DynamoDB Overview:

https://www.youtube.com/watch?v=3ZOyUNIeorU

Question 58: **Correct**
An application is hosted in an On-Demand EC2 instance and is using Amazon SDK to communicate to other AWS services such as S3, DynamoDB, and many others. As part of the upcoming IT audit, you need to ensure that all API calls to your AWS resources are logged and durably stored.

Which is the most suitable service that you should use to meet this requirement?

- **Amazon CloudWatch**
- **AWS CloudTrail**

  (Correct)
- **AWS X-Ray**
- **Amazon API Gateway**

Explanation
AWS CloudTrail increases visibility into your user and resource activity by recording AWS Management Console actions and API calls. You can identify which users and accounts called AWS, the source IP address from which the calls were made, and when the calls occurred.

Amazon CloudWatch is incorrect because this is primarily used for systems monitoring based on the server metrics. It does not have the capability to track API calls to your AWS resources.

AWS X-Ray is incorrect because this is usually used to debug and analyze your microservices applications with request tracing so you can find the root cause of issues and performance. Unlike CloudTrail, it does not record the API calls that were made to your AWS resources.

Amazon API Gateway is incorrect because this is not used for logging each and every API call to your AWS resources. It is a fully managed service that makes it easy for developers to create, publish, maintain, monitor, and secure APIs at any scale.

Reference:

https://aws.amazon.com/cloudtrail/

Check out this AWS CloudTrail Cheat Sheet:

https://tutorialsdojo.com/aws-cloudtrail/

Question 59: **Correct**
An application is hosted in an Auto Scaling group of EC2 instances. To improve the monitoring process, you have to configure the current capacity to increase or decrease based on a set of scaling adjustments. This should be done by specifying the scaling metrics and threshold values for the CloudWatch alarms that trigger the scaling process.

Which of the following is the most suitable type of scaling policy that you should use?

- **Target tracking scaling**
- **Simple scaling**
- **Scheduled Scaling**

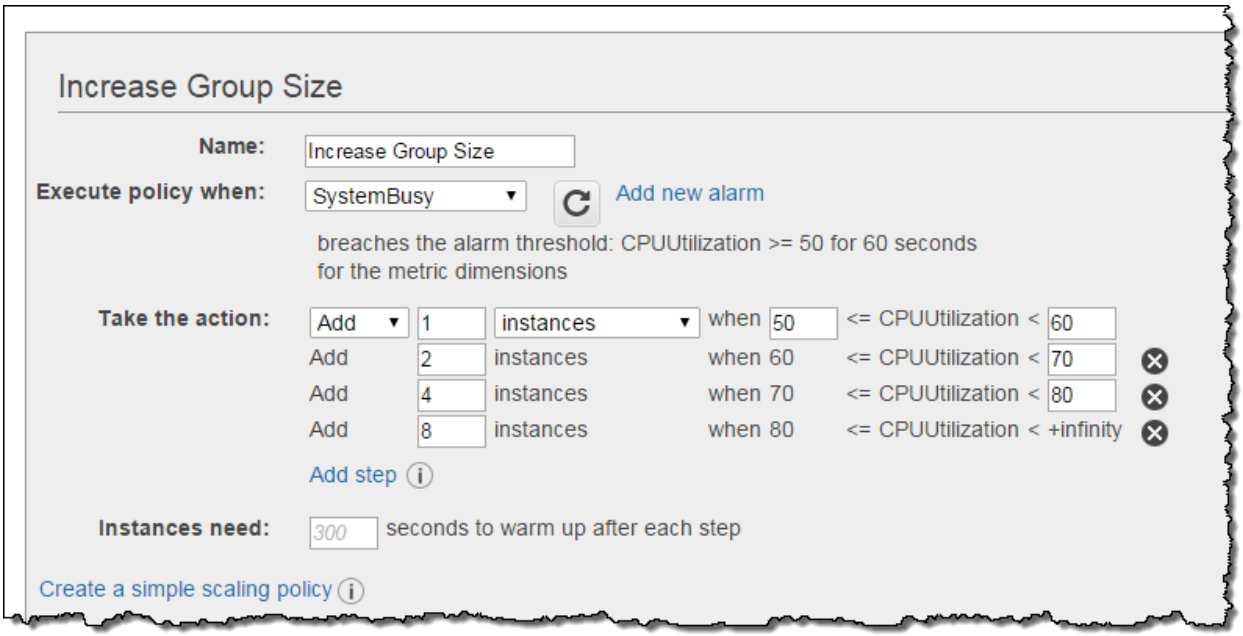
- **Step scaling**

  (Correct)

Explanation

With step scaling, you choose scaling metrics and threshold values for the CloudWatch alarms that trigger the scaling process as well as define how your scalable target should be scaled when a threshold is in breach for a specified number of evaluation periods. Step scaling policies increase or decrease the current capacity of a scalable target based on a set of scaling adjustments, known as step adjustments. The adjustments vary based on the size of the alarm breach. After a scaling activity is started, the policy continues to respond to additional alarms, even while a scaling activity is in progress. Therefore, all alarms that are breached are evaluated by Application Auto Scaling as it receives the alarm messages.

When you configure dynamic scaling, you must define how to scale in response to changing demand. For example, you have a web application that currently runs on two instances and you want the CPU utilization of the Auto Scaling group to stay at around 50 percent when the load on the application changes. This gives you extra capacity to handle traffic spikes without maintaining an excessive amount of idle resources. You can configure your Auto Scaling group to scale automatically to meet this need. The policy type determines how the scaling action is performed.



Amazon EC2 Auto Scaling supports the following types of scaling policies:

Target tracking scaling - Increase or decrease the current capacity of the group based on a target value for a specific metric. This is similar to the way that your thermostat maintains the temperature of your home – you select a temperature and the thermostat does the rest.

Step scaling - Increase or decrease the current capacity of the group based on a set of scaling adjustments, known as *step adjustments*, that vary based on the size of the alarm breach.

Simple scaling - Increase or decrease the current capacity of the group based on a single scaling adjustment.

If you are scaling based on a utilization metric that increases or decreases proportionally to the number of instances in an Auto Scaling group, then it is recommended that you use target tracking scaling policies. Otherwise, it is better to use step scaling policies instead.

Hence, the correct answer in this scenario is **Step Scaling**.

Target tracking scaling is incorrect because the target tracking scaling policy increases or decreases the current capacity of the group based on a **target value for a specific metric** instead of a set of scaling adjustments.

Simple scaling is incorrect because the simple scaling policy increases or decreases the current capacity of the group based on a **single** scaling adjustment instead of a set of scaling adjustments.

Scheduled Scaling is incorrect because the scheduled scaling policy is based on a schedule that allows you to set your own scaling schedule for **predictable** load changes. This is not considered as one of the types of dynamic scaling.

References:

https://docs.aws.amazon.com/autoscaling/ec2/userguide/as-scale-based-on-demand.html

https://docs.aws.amazon.com/autoscaling/application/userguide/application-auto-scaling-step-scaling-policies.html

Question 60: **Correct**
An application is hosted on an EC2 instance with multiple EBS Volumes attached and uses Amazon Neptune as its database. To improve data security, you encrypted all of

the EBS volumes attached to the instance to protect the confidential data stored in the volumes.

Which of the following statements are true about encrypted Amazon Elastic Block Store volumes? (Select TWO.)

- [x] **Snapshots are automatically encrypted.**

  (Correct)

- [ ] **Only the data in the volume is encrypted and not all the data moving between the volume and the instance.**

- [x] **All data moving between the volume and the instance are encrypted.**

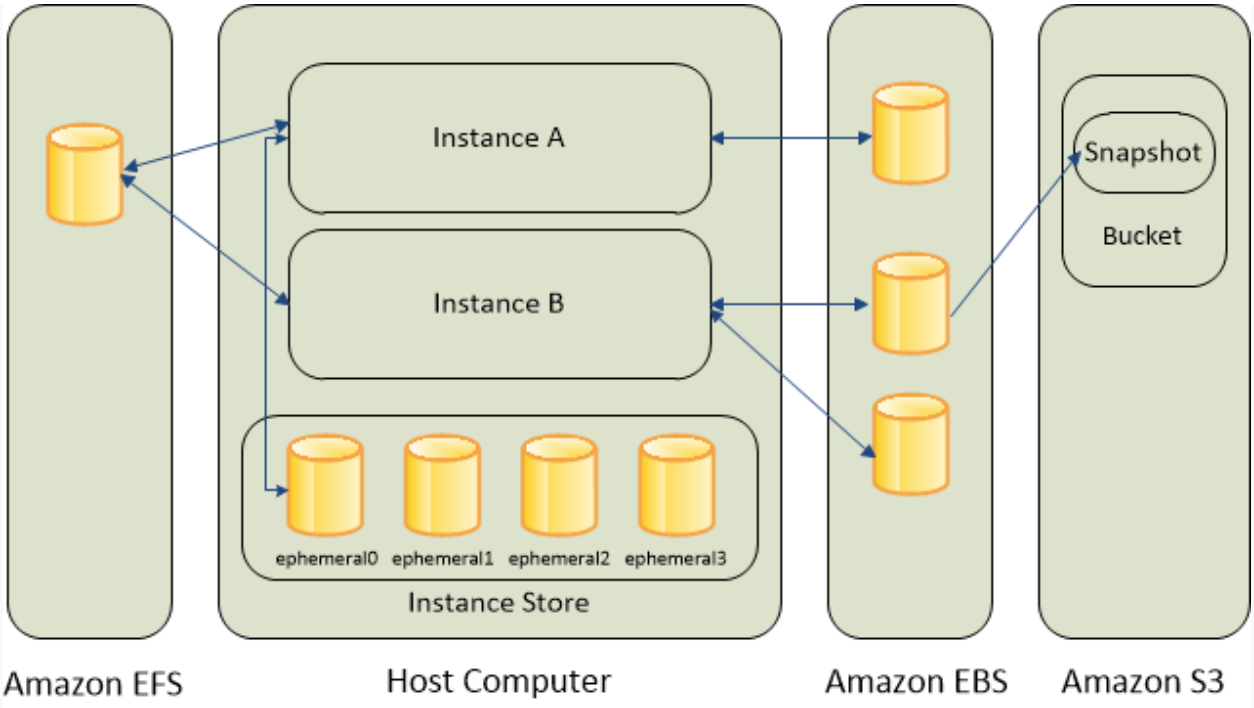
  (Correct)

- [ ] **Snapshots are not automatically encrypted.**

- [ ] **The volumes created from the encrypted snapshot are not encrypted.**

Explanation

Amazon Elastic Block Store (Amazon EBS) provides block-level storage volumes for use with EC2 instances. EBS volumes are highly available and reliable storage volumes that can be attached to any running instance that is in the same Availability Zone. EBS volumes that are attached to an EC2 instance are exposed as storage volumes that persist independently from the life of the instance.



When you create an encrypted EBS volume and attach it to a supported instance type, the following types of data are encrypted:

- Data at rest inside the volume

- All data moving between the volume and the instance

- All snapshots created from the volume

- All volumes created from those snapshots

Encryption operations occur on the servers that host EC2 instances, ensuring the security of both data-at-rest and data-in-transit between an instance and its attached EBS storage. You can encrypt both the boot and data volumes of an EC2 instance.

References:

http://docs.aws.amazon.com/AWSEC2/latest/UserGuide/AmazonEBS.html

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

Check out this Amazon EBS Cheat Sheet:

https://tutorialsdojo.com/amazon-ebs/

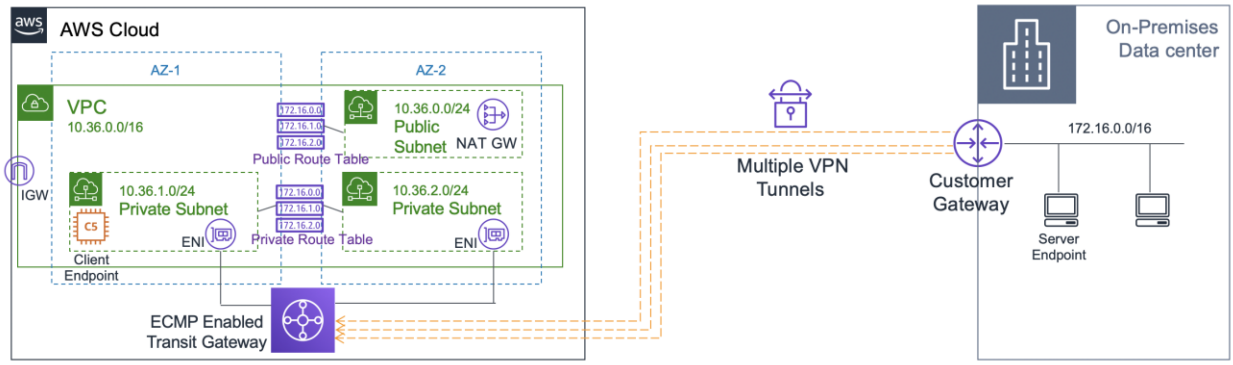
Question 61: **Correct**
A company has multiple AWS Site-to-Site VPN connections placed between their VPCs and their remote network. During peak hours, many employees are experiencing slow connectivity issues, which limits their productivity. The company has asked a solutions architect to scale the throughput of the VPN connections.

Which solution should the architect carry out?

- **Add more virtual private gateways to a VPC and enable Equal Cost Multipath Routing (ECMR) to get higher VPN bandwidth.**
- **Modify the VPN configuration by increasing the number of tunnels to scale the throughput.**
- **Associate the VPCs to an Equal Cost Multipath Routing (ECMR)-enabled transit gateway and attach additional VPN tunnels.**

  (Correct)
- **Re-route some of the VPN connections to a secondary customer gateway device on the remote network's end.**

Explanation
With AWS Transit Gateway, you can simplify the connectivity between multiple VPCs and also connect to any VPC attached to AWS Transit Gateway with a single VPN connection.



AWS Transit Gateway also enables you to scale the IPsec VPN throughput with equal-cost multi-path (ECMP) routing support over multiple VPN tunnels. A single VPN tunnel still has a maximum throughput of 1.25 Gbps. If you establish multiple VPN tunnels to an ECMP-enabled transit gateway, it can scale beyond the default limit of 1.25 Gbps.

Hence, the correct answer is: **Associate the VPCs to an Equal Cost Multipath Routing (ECMR)-enabled transit gateway and attach additional VPN tunnels.**

The option that says: **Add more virtual private gateways to a VPC and enable Equal Cost Multipath Routing (ECMR) to get higher VPN bandwidth** is incorrect because a VPC can only have a single virtual private gateway attached to it one at a time. Also, there is no option to enable ECMR in a virtual private gateway.

The option that says: **Modify the VPN configuration by increasing the number of tunnels to scale the throughput** is incorrect. The maximum tunnel for a VPN connection is two. You cannot increase this beyond its limit.

The option that says: **Re-route some of the VPN connections to a secondary customer gateway device on the remote network's end** is incorrect. This would only increase connection redundancy and won't increase throughput. For example, connections can fail over to the secondary customer gateway device in case the primary customer gateway device becomes unavailable.

References:

https://aws.amazon.com/premiumsupport/knowledge-center/transit-gateway-ecmp-multiple-tunnels/

https://aws.amazon.com/blogs/networking-and-content-delivery/scaling-vpn-throughput-using-aws-transit-gateway/

Check out this AWS Transit Gateway Cheat Sheet:

https://tutorialsdojo.com/aws-transit-gateway/

Question 62: **Correct**
A company has a web-based ticketing service that utilizes Amazon SQS and a fleet of EC2 instances. The EC2 instances that consume messages from the SQS queue are configured to poll the queue as often as possible to keep end-to-end throughput as high as possible. The Solutions Architect noticed that polling the queue in tight loops is using unnecessary CPU cycles, resulting in increased operational costs due to empty responses.

In this scenario, what should the Solutions Architect do to make the system more cost-effective?

- **Configure Amazon SQS to use short polling by setting the ReceiveMessageWaitTimeSeconds to zero.**
- **Configure Amazon SQS to use long polling by setting the ReceiveMessageWaitTimeSeconds to a number greater than zero.**

  (Correct)
- **Configure Amazon SQS to use short polling by setting the ReceiveMessageWaitTimeSeconds to a number greater than zero.**
- **Configure Amazon SQS to use long polling by setting the ReceiveMessageWaitTimeSeconds to zero.**

Explanation
In this scenario, the application is deployed in a fleet of EC2 instances that are polling messages from a single SQS queue. Amazon SQS uses short polling by default, querying only a subset of the servers (based on a weighted random distribution) to determine whether any messages are available for inclusion in the response. Short polling works for scenarios that require higher throughput. However, you can also configure the queue to use Long polling instead, to reduce cost.

The ReceiveMessageWaitTimeSeconds is the queue attribute that determines whether you are using Short or Long polling. By default, its value is zero which means it is using Short polling. If it is set to a value greater than zero, then it is Long polling.

Hence, **configuring Amazon SQS to use long polling by setting the ReceiveMessageWaitTimeSeconds to a number greater than zero is the correct answer.**

Quick facts about SQS Long Polling:

- Long polling helps reduce your cost of using Amazon SQS by reducing the number of empty responses when there are no messages available to return in reply to a *ReceiveMessage* request sent to an Amazon SQS queue and eliminating false empty responses when messages are available in the queue but aren't included in the response.

- Long polling reduces the number of empty responses by allowing Amazon SQS to wait until a message is available in the queue before sending a response. Unless the connection times out, the response to the `ReceiveMessage` request contains at least one of the available messages, up to the maximum number of messages specified in the `ReceiveMessage` action.

- Long polling eliminates false empty responses by querying all (rather than a limited number) of the servers. Long polling returns messages as soon any message becomes available.

Reference:

https://docs.aws.amazon.com/AWSSimpleQueueService/latest/SQSDeveloperGuide/sqs-long-polling.html

Check out this Amazon SQS Cheat Sheet:

https://tutorialsdojo.com/amazon-sqs/

Question 63: **Correct**
A news company is planning to use a Hardware Security Module (CloudHSM) in AWS for secure key storage of their web applications. You have launched the CloudHSM cluster but after just a few hours, a support staff mistakenly attempted to log in as the administrator three times using an invalid password in the Hardware Security Module. This has caused the HSM to be zeroized, which means that the encryption keys on it

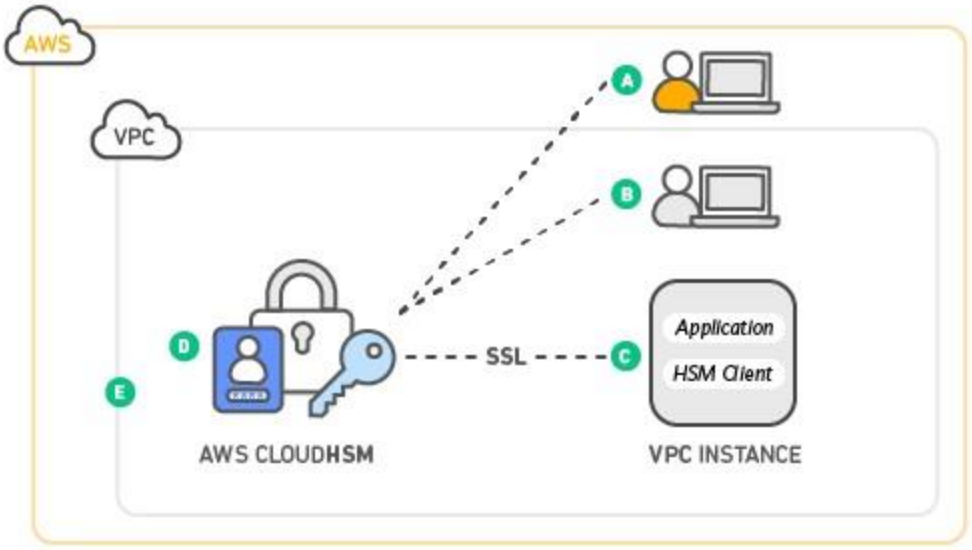
have been wiped. Unfortunately, you did not have a copy of the keys stored anywhere else.

How can you obtain a new copy of the keys that you have stored on Hardware Security Module?

- **Restore a snapshot of the Hardware Security Module.**
- **Contact AWS Support and they will provide you a copy of the keys.**
- **The keys are lost permanently if you did not have a copy.**

  (Correct)
- **Use the Amazon CLI to get a copy of the keys.**

Explanation
Attempting to log in as the administrator more than twice with the wrong password zeroizes your HSM appliance. When an HSM is zeroized, all keys, certificates, and other data on the HSM is destroyed. You can use your cluster's security group to prevent an unauthenticated user from zeroizing your HSM.



Amazon does not have access to your keys nor to the credentials of your Hardware Security Module (HSM) and therefore has no way to recover your keys if you lose your credentials. Amazon strongly recommends that you use two or more HSMs in separate Availability Zones in any production CloudHSM Cluster to avoid loss of cryptographic keys.

Refer to the CloudHSM FAQs for reference:

Q: Could I lose my keys if a single HSM instance fails?

Yes. It is possible to lose keys that were created since the most recent daily backup if the CloudHSM cluster that you are using fails and you are not using two or more HSMs. Amazon strongly recommends that you use two or more HSMs, in separate Availability Zones, in any production CloudHSM Cluster to avoid loss of cryptographic keys.

Q: Can Amazon recover my keys if I lose my credentials to my HSM?

No. Amazon does not have access to your keys or credentials and therefore has no way to recover your keys if you lose your credentials.

References:

https://aws.amazon.com/premiumsupport/knowledge-center/stop-cloudhsm/

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

https://d1.awsstatic.com/whitepapers/Security/security-of-aws-cloudhsm-backups.pdf

Question 64: **Correct**
A company plans to deploy a Docker-based batch application in AWS. The application will be used to process both mission-critical data as well as non-essential batch jobs.

Which of the following is the most cost-effective option to use in implementing this architecture?

- **Use ECS as the container management service then set up a combination of Reserved and Spot EC2 Instances for processing mission-critical and non-essential batch jobs respectively.**

  (Correct)

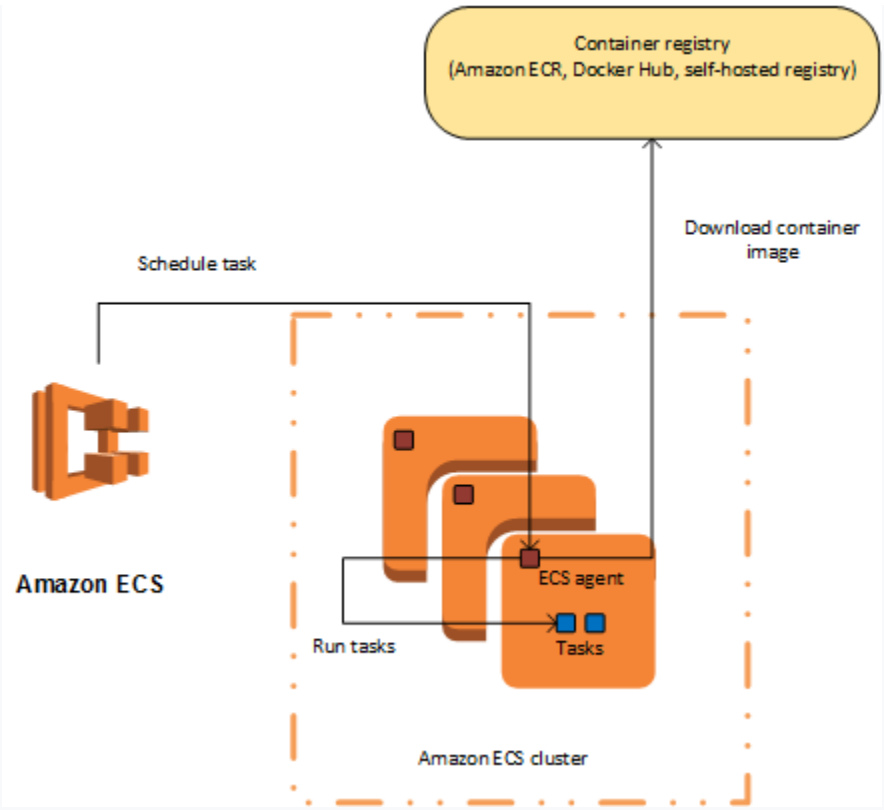
- **Use ECS as the container management service then set up On-Demand EC2 Instances for processing both mission-critical and non-essential batch jobs.**

- **Use ECS as the container management service then set up Reserved EC2 Instances for processing both mission-critical and non-essential batch jobs.**

- **Use ECS as the container management service then set up Spot EC2 Instances for processing both mission-critical and non-essential batch jobs.**

Explanation
Amazon ECS lets you run batch workloads with managed or custom schedulers on Amazon EC2 On-Demand Instances, Reserved Instances, or Spot Instances. You can launch a combination of EC2 instances to set up a cost-effective architecture depending on your workload. You can launch Reserved EC2 instances to process the mission-critical data and Spot EC2 instances for processing non-essential batch jobs.

There are two different charge models for Amazon Elastic Container Service (ECS): Fargate Launch Type Model and EC2 Launch Type Model. With Fargate, you pay for the amount of vCPU and memory resources that your containerized application requests while for EC2 launch type model, there is no additional charge. You pay for AWS resources (e.g., EC2 instances or EBS volumes) you create to store and run your application. You only pay for what you use, as you use it; there are no minimum fees and no upfront commitments.



In this scenario, the most cost-effective solution is to use ECS as the container management service then set up a combination of Reserved and Spot EC2 Instances for processing mission-critical and non-essential batch jobs respectively. You can use Scheduled Reserved Instances (Scheduled Instances) which enables you to purchase capacity reservations that recur on a daily, weekly, or monthly basis, with a specified start time and duration, for a one-year term. This will ensure that you have an uninterrupted compute capacity to process your mission-critical batch jobs.

Hence, the correct answer is the option that says: **Use ECS as the container management service then set up a combination of Reserved and Spot EC2 Instances for processing mission-critical and non-essential batch jobs respectively.**

Using ECS as the container management service then setting up Reserved EC2 Instances for processing both mission-critical and non-essential batch jobs is incorrect because processing the non-essential batch jobs can be handled much cheaper by using Spot EC2 instances instead of Reserved Instances.

Using ECS as the container management service then setting up On-Demand EC2 Instances for processing both mission-critical and non-essential batch jobs is incorrect because an On-Demand instance costs more compared to Reserved and Spot EC2 instances. Processing the non-essential batch jobs can be handled much cheaper by using Spot EC2 instances instead of On-Demand instances.

Using ECS as the container management service then setting up Spot EC2 Instances for processing both mission-critical and non-essential batch jobs is incorrect. Although this setup provides the cheapest solution among other options, it will not be able to meet the required workload. Using Spot instances to process mission-critical workloads is not suitable since these types of instances can be terminated by AWS at any time, which can affect critical processing.

References:

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

https://aws.amazon.com/ec2/spot/containers-for-less/get-started/

Check out this Amazon ECS Cheat Sheet:

https://tutorialsdojo.com/amazon-elastic-container-service-amazon-ecs/

AWS Container Services Overview:

https://youtu.be/5QBgDX7O7pw

Question 65: **Correct**
An operations team has an application running on EC2 instances inside two custom VPCs. The VPCs are located in the Ohio and N.Virginia Region respectively. The team wants to transfer data between the instances without traversing the public internet.

Which combination of steps will achieve this? (Select TWO.)

- ☑ **Re-configure the route table's target and destination of the instances' subnet.**

  (Correct)

- ☐ **Launch a NAT Gateway in the public subnet of each VPC.**

- ☐

Deploy a VPC endpoint on each region to enable a private connection.

- ☑

Set up a VPC peering connection between the VPCs.

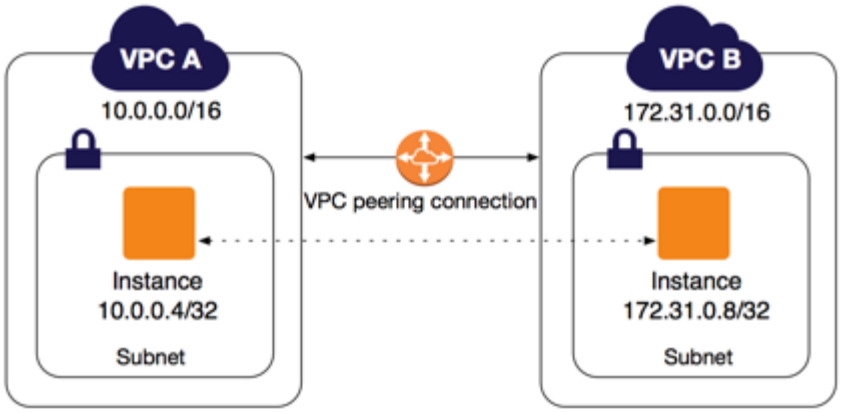
(Correct)

- ☐

Create an Egress-only Internet Gateway.

Explanation

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).



Inter-Region VPC Peering provides a simple and cost-effective way to share resources between regions or replicate data for geographic redundancy. Built on the same horizontally scaled, redundant, and highly available technology that powers VPC today, Inter-Region VPC Peering encrypts inter-region traffic with no single point of failure or bandwidth bottleneck. Traffic using Inter-Region VPC Peering always stays on the global AWS backbone and never traverses the public internet, thereby reducing threat vectors, such as common exploits and DDoS attacks.

Hence, the correct answers are:

- Set up a VPC peering connection between the VPCs.

- Re-configure the route table's target and destination of the instances' subnet.

The option that says: **Create an Egress only Internet Gateway** is incorrect because this will just enable outbound IPv6 communication from instances in a VPC to the internet. Take note that the scenario requires private communication to be enabled between VPCs from two different regions.

The option that says: **Launch a NAT Gateway in the public subnet of each VPC** is incorrect because NAT Gateways are used to allow instances in private subnets to access the public internet. Note that the requirement is to make sure that communication between instances will not traverse the internet.

The option that says: **Deploy a VPC endpoint on each region to enable private connection** is incorrect. VPC endpoints are region-specific only and do not support inter-region communication.

References:

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

https://aws.amazon.com/about-aws/whats-new/2017/11/announcing-support-for-inter-region-vpc-peering/

Check out this Amazon VPC Cheat Sheet:

https://tutorialsdojo.com/amazon-vpc/