AWS Practise Paper- 4

Question 1:

**A data analytics company has been building its new generation big data and analytics platform on their AWS cloud infrastructure. They need a storage service that provides the scale and performance that their big data applications require such as high throughput to compute nodes coupled with read-after-write consistency and low-latency file operations. In addition, their data needs to be stored redundantly across multiple AZs and allows concurrent connections from multiple EC2 instances hosted on multiple AZs.**

**Which of the following AWS storage services will you use to meet this requirement?**

- **A. EFS(Correct)**
- B. EBS
- C. S3
- D. Glacier

EXPLANATION

In this question, you should take note of the two keywords/phrases: "file operation" and "allows concurrent connections from multiple EC2 instances". There are various AWS storage options that you can choose but whenever these criteria show up, always consider using EFS instead of using EBS Volumes which is mainly used as a "block" storage and can only have one connection to one EC2 instance at a time. Amazon EFS provides the scale and performance required for big data applications that require high throughput to compute nodes coupled with read-after-write consistency and low-latency file operations.

Amazon EFS is a fully-managed service that makes it easy to set up and scale file storage in the Amazon Cloud. With a few clicks in the AWS Management Console, you can create file systems that are accessible to Amazon EC2 instances via a file system interface (using standard operating system file I/O APIs) and supports full file system access semantics (such as strong consistency and file locking).

Amazon EFS file systems can automatically scale from gigabytes to petabytes of data without needing to provision storage. Tens, hundreds, or even thousands of Amazon EC2 instances can access an Amazon EFS file system at the same time, and Amazon EFS provides consistent performance to each Amazon EC2 instance. Amazon EFS is designed to be highly durable and highly available.

Option 2 is  because EBS does not allow concurrent connections from multiple EC2 instances hosted on multiple AZs and it does not store data redundantly across multiple AZs by default, unlike EFS.

Option 3 is  because although S3 can handle concurrent connections from multiple EC2 instances, it does not have the ability to provide low-latency file operations, which is required in this scenario.

Option 4 is  because Glacier is an archiving storage solution and is not applicable in this scenario.

Question 2:

**You are employed by a large electronics company that uses Amazon Simple Storage Service. For reporting purposes, they want to track and log every request access to their S3 buckets including the requester, bucket name, request time, request action, response status, and error code information. They also use this information for their internal security and access audits.**

**Which is the best solution among the following options that can satisfy the company requirement?**

A. Enable AWS CloudTrail to audit all Amazon S3 bucket access.
**B. Enable server access logging for all required Amazon S3 buckets. (Correct)**
C. Enable the Requester Pays option to track access via AWS Billing.
D. Enable Amazon S3 Event Notifications for PUT and POST.

EXPLANATION

For this scenario, you can use CloudTrail and the Server Access Logging feature of Amazon S3. However, the question mentioned that it needs detailed information about every access request sent to the S3 bucket such as requestor, bucket name, request time, request action, response status, and error code information. CloudTrail can only log the API calls and provides less information compared with the Server Access Logging feature in S3. Hence, the correct answer is Option 2.

Option 3 is  because this action refers to AWS billing and not for logging.

Option 4 is  because we are looking for a logging solution and not event notification.

Question 3:

**You are working for a tech company which currently has an on-premises infrastructure. They are currently running low on storage and want to have the ability to extend their storage using AWS cloud.**

**Which AWS service can help you achieve this requirement?**

A. Amazon EC2
**B. Amazon Storage Gateway (Correct)**
C. Amazon Elastic Block Storage
D. Amazon SQS

EXPLANATION

AWS Storage Gateway connects an on-premises software appliance with cloud-based storage to provide seamless integration with data security features between your on-premises IT environment and the AWS storage infrastructure. You can use the service to store data in the AWS Cloud for scalable and cost-effective storage that helps maintain data security.

Option 1 is  since EC2 is a compute service, not a storage service.

Option 3 is  since EBS is primarily used as a storage of your EC2 instances.

Option 4 is  since SQS is a message queuing service, and does not extend your on-premises storage capacity.

Question 4:

**You recently launched a new FTP server using an On-Demand EC2 instance in a newly created VPC with default settings. The server should not be accessible publicly but only through your IP address 175.45.116.100 and nowhere else.**

**Which of the following is the most suitable way to implement this requirement?**

A. Create a new inbound rule in the security group of the EC2 instance with the following details:

Protocol: TCP

Port Range: 20 - 21

Source: 175.45.116.100/32(Correct)

B. Create a new inbound rule in the security group of the EC2 instance with the following details:

Protocol: UDP

Port Range: 20 - 21

Source: 175.45.116.100/32

C. Create a new Network ACL inbound rule in the subnet of the EC2 instance with the following details:

Protocol: TCP

Port Range: 20 - 21

Source: 175.45.116.100/0

Allow/Deny: ALLOW

D. Create a new Network ACL inbound rule in the subnet of the EC2 instance with the following details:

   Protocol: UDP

   Port Range: 20 - 21

   Source: 175.45.116.100/0

   Allow/Deny: ALLOW

EXPLANATION

The FTP protocol uses TCP via ports 21 and 22. This should be configured in your security groups or in your Network ACL inbound rules. As required by the scenario, you should only allow the individual IP of the client and not the entire network. Therefore, in the Source, the proper CIDR notation should be used. The /32 denotes one IP address and the /0 refers to the entire network.

Notice that the scenario says that you launched the EC2 instances in a newly created VPC with default settings. 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. Hence, you actually don't need to explicitly add inbound rules to your Network ACL to allow inbound traffic, if your VPC has a default setting.

Option 2 is  because although the configuration of the Security Group is valid, the provided Protocol is . Take note that FTP uses TCP and not UDP.

Option 3 is  because although setting up an inbound Network ACL is valid, the source is invalid since it must be an IPv4 or IPv6 CIDR block. In the provided IP address, the /0 refers to the entire network and not a specific IP address. In addition, the scenario says that the newly created VPC has default settings and by default, the Network ACL allows all traffic. This means that there is actually no need to configure your Network ACL.

Option 4 is  because, just like Option 3, the source is also invalid. Take note that FTP uses TCP and not UDP, which is one of the reasons why this option is wrong. In addition, the scenario says that the newly created VPC has default settings and by default, the Network ACL allows all traffic. This means that there is actually no need to configure your Network ACL.

Question 5:

**The IT Operations team of your company wants to retrieve all of the Public IP addresses assigned to a running EC2 instance via the Instance metadata.**

**Which of the following URLs will you use?**

**A. http://169.254.169.254/latest/meta-data/public-ipv4(Correct)**
B. http://169.255.169.255/latest/meta-data/public-ipv4
C. (http://254.169.254.169/metadata/public-ipv4
D. http://255.169.255.169/latest/public-ipv4

EXPLANATION

http://169.254.169.254/latest/meta-data/ is the URL that you can use to retrieve the Instance Metadata of your EC2 instance, including the public-hostname, public-ipv4, public-keys et cetera.

This can be helpful when you're writing scripts to run from your instance as it enables you to access the local IP address of your instance from the instance metadata to manage a connection to an external application. Remember that you are not billed for HTTP requests used to retrieve instance metadata and user data.

Question 6:

**A customer is transitioning their ActiveMQ messaging broker service onto the AWS cloud in which they require an alternative asynchronous service that supports NMS and MQTT messaging protocol. The customer does not have the time and resources needed to recreate their messaging service in the cloud. The service has to be highly available and should require almost no management overhead.**

**Which of the following is the most suitable service to use to meet the above requirement?**

A. Amazon SNS
**B. Amazon MQ(Correct)**
C. Amazon SQS
D. Amazon SWF

EXPLANATION

Amazon MQ is a managed message broker service for Apache ActiveMQ that makes it easy to set up and operate message brokers in the cloud. Connecting your current applications to Amazon MQ is easy because it uses industry-standard APIs and protocols for messaging, including JMS, NMS, AMQP, STOMP, MQTT, and WebSocket. Using standards means that in most cases, there’s no need to rewrite any messaging code when you migrate to AWS.

Amazon MQ, Amazon SQS, and Amazon SNS are messaging services that are suitable for anyone from startups to enterprises. If you're using messaging with existing applications and want to move your messaging service to the cloud quickly

and easily, it is recommended that you consider Amazon MQ. It supports industry-standard APIs and protocols so you can switch from any standards-based message broker to Amazon MQ without rewriting the messaging code in your applications. Hence, Option 2 is the correct answer.

If you are building brand new applications in the cloud, then it is highly recommended that you consider Amazon SQS and Amazon SNS. Amazon SQS and SNS are lightweight, fully managed message queue and topic services that scale almost infinitely and provide simple, easy-to-use APIs. You can use Amazon SQS and SNS to decouple and scale microservices, distributed systems, and serverless applications, and improve reliability.

Option 1 is  because SNS is more suitable as a pub/sub messaging service instead of a message broker service.

Option 3 is  because although Amazon SQS is a fully managed message queuing service, it does not support an extensive list of industry-standard messaging APIs and protocol, unlike Amazon MQ. Moreover, using Amazon SQS requires you to do additional changes in the messaging code of applications to make it compatible.

Option 4 is  because SWF is a fully-managed state tracker and task coordinator service and not a messaging service, unlike Amazon MQ, AmazonSQS, and Amazon SNS.

Question 7:

**The company you are working for has a set of AWS resources hosted in ap-northeast-1 region. You have been requested by your IT Manager to create a shell script which could create duplicate resources in another region in the event that ap-northeast-1 region fails.**

**Which of the following AWS services could help fulfill this task?**

A. AWS Elastic Beanstalk
B. AWS SQS
**C. AWS CloudFormation(Correct)**
D. AWS SNS

EXPLANATION

AWS CloudFormation is a service that helps you model and set up your Amazon Web Services resources so that you can spend less time managing those resources and more time focusing on your applications that run in AWS.

You can create a template that describes all the AWS resources that you want (like Amazon EC2 instances or Amazon RDS DB instances), and AWS CloudFormation takes care of provisioning and configuring those resources for you. With this, you

can deploy an exact copy of your AWS architecture, along with all of the AWS resources which are hosted in one region to another.

Question 8:

**You have an Auto Scaling group which is configured to launch new t2.micro EC2 instances when there is a significant load increase in the application. To cope with the demand, you now need to replace those instances with a larger t2.2xlargeinstance type. How would you implement this change?**

A. Just change the instance type to t2.2xlarge in the current launch configuration
B. Create another Auto Scaling Group and attach the new instance type.
**C. Create a new launch configuration with the new instance type and update the Auto Scaling Group.(Correct)**
D. Change the instance type of each EC2 instance manually.

EXPLANATION

You can only specify one launch configuration for an Auto Scaling group at a time, and you can't modify a launch configuration after you've created it. Therefore, if you want to change the launch configuration for an Auto Scaling group, you must create a launch configuration and then update your Auto Scaling group with the new launch configuration.

Question 9:

**An auto-scaling group of Linux EC2 instances is created with basic monitoring enabled in CloudWatch. You noticed that your application is slow so you asked one of your engineers to check all of your EC2 instances. After checking your instances, you noticed that the auto scaling group is not launching more instances as it should be, even though the servers already have high memory usage.**

**What is the best solution that will fix this issue?**

A. Install AWS SDK in the EC2 instances. Create a script that will trigger the Auto Scaling event if there is a high memory usage.
**B. Install CloudWatch monitoring scripts in the instances. Send custom metrics to CloudWatch which will trigger your Auto Scaling group to scale up.(Correct)**
C. Enable detailed monitoring on the instances.
D. Modify the scaling policy to increase the threshold to scale up the number of instances.

EXPLANATION

The Amazon CloudWatch Monitoring Scripts for Amazon Elastic Compute Cloud (Amazon EC2) Linux-based instances demonstrate how to produce and consume Amazon CloudWatch custom metrics. These sample Perl scripts comprise a fully functional example that reports memory, swap, and disk space utilization metrics for a Linux instance.

Option 2 is correct because CloudWatch does not monitor EC2 memory usage as well as disk space utilization. You would have to send custom metrics to CloudWatch.

Option 1 is  because AWS SDK is a set of programming tools that allow you to create applications that run using Amazon cloud services. You would have to program the alert which is not the best strategy for this scenario.

Option 3 is  because detailed monitoring does not provide metrics for memory usage. Cloudwatch does not monitor memory usage in its default set of EC2 metrics and detailed monitoring just provides higher frequency of metrics (1-minute frequency).

Option 4 is  because you are already maxing out your usage, which should in effect cause an auto-scaling event.

Question 10:

**You have a set of linux servers running on multiple On-Demand EC2 Instances. The Audit team wants to collect and process the application log files generated from these servers for their report.**

**Which of the following services is the best to use in this case?**

- **A. Amazon S3 for storing the application log files and Amazon Elastic MapReduce for processing the log files. (Correct)**
- B. Amazon Glacier for storing the application log files and Spot EC2 Instances for processing them.
- C. A single On-Demand Amazon EC2 instance for both storing and processing the log files
- D. Amazon RedShift to store the logs and Amazon Lambda for running custom log analysis scripts

EXPLANATION

Amazon EMR is a managed cluster platform that simplifies running big data frameworks, such as Apache Hadoop and Apache Spark, on AWS to process and analyze vast amounts of data. By using these frameworks and related open-source projects such as Apache Hive and Apache Pig, you can process data for analytics purposes and business intelligence workloads. Additionally, you can use Amazon EMR to transform and move large amounts of data into and out of other AWS data

stores and databases such as Amazon Simple Storage Service (Amazon S3) and Amazon DynamoDB.

Option 2 is wrong as Amazon Glacier is used for data archive only.

Option 3 is wrong as an EC2 instance is not a recommended storage service. In addition, Amazon EC2 does not have a built-in data processing engine to process large amounts of data.

Option 4 is wrong as Amazon RedShift is mainly used as a data warehouse service.

Question 11:

**To save costs, your manager instructed you to analyze and review the setup of your AWS cloud infrastructure. You should also provide an estimate of how much your company will pay for all of the AWS resources that they are using. In this scenario, which of the following will incur costs?**

- **A. A running EC2 Instance(Correct)**
- B. A stopped On-Demand EC2 Instance
- **C. EBS Volumes attached to stopped EC2 Instances(Correct)**
- D. Using an Amazon VPC
- E. Public Data Set

EXPLANATION

Billing commences when Amazon EC2 initiates the boot sequence of an AMI instance. Billing ends when the instance terminates, which could occur through a web services command, by running "shutdown -h", or through instance failure. When you stop an instance, AWS shuts it down but don't charge hourly usage for a stopped instance or data transfer fees, but AWS does charge for the storage of any Amazon EBS volumes. Hence, options 1 and 3 are the right answers and conversely, options 2 and 6 are  as there is no charge for a terminated EC2 instance that you have shut down.

Option 4 is  because there are no additional charges for creating and using the VPC itself. Usage charges for other Amazon Web Services, including Amazon EC2, still apply at published rates for those resources, including data transfer charges.

Option 5 is  due to the fact that Amazon stores the data sets at no charge to the community and, as with all AWS services, you pay only for the compute and storage you use for your own applications.

Question 12:

**A financial company instructed you to automate the recurring tasks in your department such as patch management, infrastructure selection, and data synchronization to improve their current processes. You need to have a**

**service which can coordinate multiple AWS services into serverless workflows.**

**Which of the following is the most cost-effective service to use in this scenario?**

A. SWF
B. AWS Lambda
**C. AWS Step Functions(Correct)**
D. AWS Batch

EXPLANATION

AWS Step Functions provides serverless orchestration for modern applications. Orchestration centrally manages a workflow by breaking it into multiple steps, adding flow logic, and tracking the inputs and outputs between the steps. As your applications execute, Step Functions maintains application state, tracking exactly which workflow step your application is in, and stores an event log of data that is passed between application components. That means that if networks fail or components hang, your application can pick up right where it left off.

Application development is faster and more intuitive with Step Functions, because you can define and manage the workflow of your application independently from its business logic. Making changes to one does not affect the other. You can easily update and modify workflows in one place, without having to struggle with managing, monitoring and maintaining multiple point-to-point integrations. Step Functions frees your functions and containers from excess code, so your applications are faster to write, more resilient, and easier to maintain.

Option 1 is  because SWF is a fully-managed state tracker and task coordinator service. It does not provide serverless orchestration to multiple AWS resources.

Option 2 is  because although Lambda is used for serverless computing, it does not provide a direct way to coordinate multiple AWS services into serverless workflows.

Option 4 is  because AWS Batch is primarily used to efficiently run hundreds of thousands of batch computing jobs in AWS.

Question 13:

**The game development company that you are working for has an Amazon VPC with a public subnet. It has 4 EC2 instances that are deployed in the public subnet. These 4 instances can successfully communicate with other hosts on the Internet. You launch a fifth instance in the same public subnet, using the same AMI and security group configuration that you used for the others. However, this new instance cannot be accessed from the internet unlike the other instance.**

**What should you do to enable access to the fifth instance over the Internet?**

A. Deploy a NAT instance into the public subnet.
**B. Assign an Elastic IP address to the fifth instance. (Correct)**
C. Configure a publicly routable IP Address in the host OS of the fifth instance.
D. Modify the routing table for the public subnet.

EXPLANATION

An Elastic IP address is a static IPv4 address designed for dynamic cloud computing. An Elastic IP address is associated with your AWS account. With an Elastic IP address, you can mask the failure of an instance or software by rapidly remapping the address to another instance in your account.

An Elastic IP address is a public IPv4 address, which is reachable from the Internet. If your instance does not have a public IPv4 address, you can associate an Elastic IP address with your instance to enable communication with the Internet; for example, to connect to your instance from your local computer.

Option 1 is  because it is already mentioned that your instances are in a public subnet. You only have to configure a NAT instance when your instances are on a private subnet.

Option 2 is the correct answer because you need to either add a public address or add an EIP for this EC2 instance for it to be able to access the internet.

Option 3 is  because the public IP address has to be configured in the Elastic Network Interface (ENI) of the EC2 instance and not on its Operating System (OS).

Option 4 is  because if the routing table was wrong then you would have an issue with the other 4 instances.

Question 14:

**A technology company is building a new cryptocurrency trading platform that allows buying and selling of Bitcoin, Ethereum, XRP, Ripple and many others. You were hired as a Cloud Engineer to build the required infrastructure needed for this new trading platform. On your first week at work, you started to create CloudFormation YAML scripts that defines all of the needed AWS resources for the application. Your manager was shocked that you haven't created the EC2 instances, S3 buckets and other AWS resources straight away. He does not understand the text-based scripts that you have done and was disappointed that you are just slacking off at your job.**

**In this scenario, what are the benefits of using the Amazon CloudFormation service that you should tell your manager to clarify his concerns?**

A. Provides highly durable and scalable data storage
B. A storage location for the code of your application

C. **Enables modeling, provisioning, and version-controlling of your entire AWS infrastructure(Correct)**
D. **Allows you to model your entire infrastructure in a text file(Correct)**
E. Using CloudFormation itself is free, including the AWS resources that have been created.

EXPLANATION

AWS CloudFormation provides a common language for you to describe and provision all the infrastructure resources in your cloud environment. CloudFormation allows you to use a simple text file to model and provision, in an automated and secure manner, all the resources needed for your applications across all regions and accounts. This file serves as the single source of truth for your cloud environment. AWS CloudFormation is available at no additional charge, and you pay only for the AWS resources needed to run your applications.

Question 15:

**You have a web application hosted in AWS cloud where the application logs are sent to Amazon CloudWatch. Lately, the web application has recently been encountering some errors which can be resolved simply by restarting the instance.**

**What will you do to automatically restart the EC2 instances whenever the same application error occurs?**

A. **First, look at the existing CloudWatch logs for keywords related to the application error to create a custom metric. Then, create a CloudWatch alarm for that custom metric which invokes an action to restart the EC2 instance.(Correct)**
B. First, look at the existing CloudWatch logs for keywords related to the application error to create a custom metric. Then, create an alarm in Amazon SNS for that custom metric which invokes an action to restart the EC2 instance.
C. First, look at the existing Flow logs for keywords related to the application error to create a custom metric. Then, create a CloudWatch alarm for that custom metric which invokes an action to restart the EC2 instance.
D. First, look at the existing Flow logs for keywords related to the application error to create a custom metric. Then, create a CloudWatch alarm for that custom metric which calls a Lambda function that invokes an action to restart the EC2 instance.

EXPLANATION

In this scenario, you can look at the existing CloudWatch logs for keywords related to the application error to create a custom metric. Then, create a CloudWatch alarm for that custom metric which invokes an action to restart the EC2 instance.

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

Option 2 is  because you can't create an alarm in Amazon SNS.

Options 3 and 4 are  because Flow Logs are used in VPC and not on specific EC2 instance.

Question 16:

**A real-time data analytics application is using AWS Lambda to process data and store results in JSON format to an S3 bucket. To speed up the existing workflow, you have to use a service where you can run sophisticated Big Data analytics on your data without moving them into a separate analytics system.**

**Which of the following group of services can you use to meet this requirement?**

A. S3 Select, Amazon Neptune, DynamoDB DAX
B. Amazon X-Ray, Amazon Neptune, DynamoDB
C. Amazon Glue, Glacier Select, Amazon Redshift
**D. S3 Select, Amazon Athena, Amazon Redshift Spectrum (Correct)**

EXPLANATION

Amazon S3 allows you to run sophisticated Big Data analytics on your data without moving the data into a separate analytics system. In AWS, there is a suite of tools that make analyzing and processing large amounts of data in the cloud faster, including ways to optimize and integrate existing workflows with Amazon S3:

1. S3 Select

Amazon S3 Select is designed to help analyze and process data within an object in Amazon S3 buckets, faster and cheaper. It works by providing the ability to retrieve a subset of data from an object in Amazon S3 using simple SQL expressions. Your applications no longer have to use compute resources to scan and filter the data from an object, potentially increasing query performance by up to 400%, and reducing query costs as much as 80%. You simply change your application to use SELECT instead of GET to take advantage of S3 Select.

2. Amazon Athena

Amazon Athena is an interactive query service that makes it easy to analyze data in Amazon S3 using standard SQL expressions. Athena is serverless, so there is no infrastructure to manage, and you pay only for the queries you run. Athena is easy to use. Simply point to your data in Amazon S3, define the schema, and start querying using standard SQL expressions. Most results are delivered within seconds. With Athena, there's no need for complex ETL jobs to prepare your data for analysis. This makes it easy for anyone with SQL skills to quickly analyze large-scale datasets.

3. Amazon Redshift Spectrum

Amazon Redshift also includes Redshift Spectrum, allowing you to directly run SQL queries against exabytes of unstructured data in Amazon S3. No loading or transformation is required, and you can use open data formats, including Avro, CSV, Grok, ORC, Parquet, RCFile, RegexSerDe, SequenceFile, TextFile, and TSV. Redshift Spectrum automatically scales query compute capacity based on the data being retrieved, so queries against Amazon S3 run fast, regardless of data set size.

Question 17:

**You have a web application hosted in an On-Demand EC2 instance in your VPC. You are creating a shell script that needs the instance's public and private IP addresses.**

**What is the best way to get the instance's associated IP addresses which your shell script can use?**

A. By using IAM.
B. By using a CloudWatch metric.
**C. By using a Curl or Get Command to get the latest metadata information from http://169.254.169.254/latest/meta-data/(Correct)**
D. By using a Curl or Get Command to get the latest user data information from http://169.254.169.254/latest/user-data/

EXPLANATION

Instance metadata is data about your EC2 instance that you can use to configure or manage the running instance. Because your instance metadata is available from your running instance, you do not need to use the Amazon EC2 console or the AWS CLI. This can be helpful when you're writing scripts to run from your instance. For example, you can access the local IP address of your instance from instance metadata to manage a connection to an external application.

To view the private IPv4 address, public IPv4 address, and all other categories of instance metadata from within a running instance, use the following URL:

http://169.254.169.254/latest/meta-data/

Question 18:

**You are the technical lead of the Cloud Infrastructure team in your company and you were consulted by a software developer regarding the required AWS resources of the web application that he is building. He knows that an Instance Store only provides ephemeral storage where the data is automatically deleted when the instance is terminated. To ensure that the data of his web application persists, the app should be launched in an EC2 instance that has a durable, block-level storage volume attached. He knows that they need to use an EBS volume, but they are not sure what type they need to use.**

**In this scenario, which of the following is true about Amazon EBS volume types and their respective usage?**

A. Spot volumes provide the lowest cost per gigabyte of all EBS volume types and are ideal for workloads where data is accessed infrequently, and applications where the lowest storage cost is important.
**B. Provisioned IOPS volumes offer storage with consistent and low-latency performance, and are designed for I/O intensive applications such as large relational or NoSQL databases.(Correct)**
C. Magnetic volumes provide the lowest cost per gigabyte of all EBS volume types and are ideal for workloads where data is accessed infrequently, and applications where the lowest storage cost is important.
D. Reduced Redundancy Storage volumes offer consistent and low-latency performance, and are designed for I/O intensive applications such as large relational or NoSQL databases.
E. Single root I/O virtualization (SR-IOV) volumes are suitable for a broad range of workloads, including small to medium sized databases, development and test environments, and boot volumes.

EXPLANATION

Amazon EBS provides three volume types to best meet the needs of your workloads: General Purpose (SSD), Provisioned IOPS (SSD), and Magnetic.

General Purpose (SSD) is the new, SSD-backed, general purpose EBS volume type that we recommend as the default choice for customers. General Purpose (SSD) volumes are suitable for a broad range of workloads, including small to medium sized databases, development, and test environments, and boot volumes.

Provisioned IOPS (SSD) volumes offer storage with consistent and low-latency performance and are designed for I/O intensive applications such as large relational or NoSQL databases. Magnetic volumes provide the lowest cost per gigabyte of all EBS volume types.

Magnetic volumes are ideal for workloads where data is accessed infrequently, and applications where the lowest storage cost is important.

Question 19:

**A company is using hundreds of AWS resources in multiple AWS regions. They require a way to uniquely identify all of their AWS resources that will allow them to specify a resource unambiguously across all of AWS, such as in IAM policies, Amazon Relational Database Service (Amazon RDS) tags, and API calls.**

**Which of the following is the most suitable option to use in this scenario?**

A. AWS Resource ID
B. AWS Service Namespaces
**C. Amazon Resource Name(Correct)**
D. Tags

EXPLANATION

Amazon Resource Names (ARNs) uniquely identify AWS resources. We require an ARN when you need to specify a resource unambiguously across all of AWS, such as in IAM policies, Amazon Relational Database Service (Amazon RDS) tags, and API calls.

Option 1 is  because an AWS Resource ID is primarily used to find your resources in the Amazon EC2 console only and not your entire VPC or AWS account.

Option 2 is  because AWS Service Namespaces only helps you identify an AWS service and not a unique resource. For example, the namespace for Amazon S3 is s3, and the namespace for Amazon EC2 is ec2.

Option 4 is  because although Tags can enable you to categorize your AWS resources by purpose, owner, or environment, it is still limited because you cannot tag all of your AWS resources. Take note that you cannot tag Egress-only internet gateway, VPC flow log, VPC endpoint, and many others. Amazon Resource Names (ARNs) uniquely identify all of your AWS resources which is a more suitable option for this scenario.

Question 20:

**You are a Solutions Architect for a large London-based software company. You are assigned to improve the performance and current processes of supporting the AWS resources in your VPC. Upon checking, you noticed that the Operations team does not have an automated way to monitor and resolve issues with their on-demand EC2 instances.**

**What can be used to automatically monitor your EC2 instances and notify the Operations team for any incidents?**

A. AWS Cloudtrail
**B. AWS Cloudwatch(Correct)**
C. AWS SWF
D. AWS SQS

EXPLANATION

Amazon CloudWatch is a monitoring service for AWS cloud resources and the applications you run on AWS. You can use Amazon CloudWatch to collect and track metrics, collect and monitor log files, and set alarms.

Amazon CloudWatch can monitor AWS resources such as Amazon EC2 instances, Amazon DynamoDB tables, and Amazon RDS DB instances, as well as custom metrics generated by your applications and services, and any log files your applications generate. You can use Amazon CloudWatch to gain system-wide visibility into resource utilization, application performance, and operational health. You can use these insights to react and keep your application running smoothly.

Option 1 is  as CloudTrail is mainly used for logging and not for monitoring.

Options 3 and 4 are  as SWF and SQS are used for creating distributed application with decoupled components and not for monitoring.

Question 21:

**You are responsible for running a global news website hosted in a fleet of EC2 Instances. Lately, the load on the website has increased which resulted to slower response time for the site visitors. This issue impacts the revenue of the company as some readers tend to leave the site if it does not load after 10 seconds.**

**Which of the below services in AWS can be used to solve this problem?**

**A. Use AWS CloudFront with website as the custom origin.(Correct)**
B. For better read throughput, use AWS Storage Gateway to distribute the content across multiple regions.
**C. Use Amazon ElastiCache for the website's in-memory data store or cache.(Correct)**
D. Deploy the website to all regions in different VPCs for faster processing.

EXPLANATION

The global news website has a problem with latency considering that there are a lot of readers of the site from all parts of the globe. In this scenario, you can use a content delivery network (CDN) which is a geographically distributed group of servers which work together to provide fast delivery of Internet content. And since

this is a news website, most of its data are read-only, which can be cached to improve the read throughput and avoid the repetitive requests from the server.

In AWS, Amazon CloudFront is the global content delivery network (CDN) service that you can use and for web caching, Amazon ElastiCache is the suitable service. Hence, the answers here are options 1 and 3.

Option 2 is  as AWS Storage Gateway is used for storage.

Option 4 is  as this would be costly and totally unnecessary considering that you can use Amazon CloudFront and ElastiCache to improve the performance of the website.

Question 22:

**You are creating a Provisioned IOPS volume in AWS. The size of the volume is 10 GiB.**

**Which of the following is the correct value that should be put for the IOPS of the volume?**

A. 400
**B. 500(Correct)**
C. 600
D. 800

EXPLANATION

50:1 is the maximum ratio of provisioned IOPS to requested volume size in Gibibyte (GiB).

So for instance, a 10 GiB volume can be provisioned with up to 500 IOPS. Any volume 640 GiB in size or greater allows provisioning up to the 32,000 IOPS maximum (50 × 640 GiB = 32,000).

Question 23:

**You are managing an online platform which allows people to easily buy, sell, spend, and manage their cryptocurrency. To meet the strict IT audit requirements, each of the API calls on all of your AWS resources should be properly captured and recorded. You used CloudTrail in your VPC to help you in the compliance, operational auditing, and risk auditing of your AWS account.**

**In this scenario, where does CloudTrail store all of the logs that it creates?**

A. DynamoDB
B. A RDS instance
C. Amazon Redshift
**D. Amazon S3(Correct)**

EXPLANATION

CloudTrail is enabled on your AWS account when you create it. When activity occurs in your AWS account, that activity is recorded in a CloudTrail event. You can easily view events in the CloudTrail console by going to Event history.

Event history allows you to view, search, and download the past 90 days of supported activity in your AWS account. In addition, you can create a CloudTrail trail to further archive, analyze, and respond to changes in your AWS resources. A trail is a configuration that enables delivery of events to an Amazon S3 bucket that you specify. You can also deliver and analyze events in a trail with Amazon CloudWatch Logs and Amazon CloudWatch Events. You can create a trail with the CloudTrail console, the AWS CLI, or the CloudTrail API.

The rest of the answers are . Options 1 and 2 are for database, option 3 is used for data warehouse that scales horizontally and allows you to store terabytes and petabytes of data.

Question 24:

**Your company has a two-tier environment in their on-premises data center which is composed of an application tier and database tier. You are instructed to migrate their environment to the AWS cloud, and to design the subnets in their VPC with the following requirements:**

a) There is an application load balancer that would distribute the incoming traffic among the servers in the application tier.

**b) The application tier and the database tier must not be accessible from the public Internet. The application tier should only accept traffic coming from the load balancer.**

c) The database tier contains very sensitive data. It must not share the same subnet with other AWS resources and its custom route table with other instances in the environment.

d) The environment must be highly available and scalable to handle a surge of incoming traffic over the Internet.

Question 25

**How many subnets should you create to meet the above requirements?**

A. 2

B. 3
C. 4
**D. 6(Correct)**

EXPLANATION

In the given scenario, it is evident that only the load balancer is accessible from the public. Therefore, a public subnet is required. Take note that if you have more than one private subnet in the same Availability Zone that contains instances that need to be registered with the load balancer, you only need to create one public subnet. You only need one public subnet per Availability Zone; you can add the private instances in all the private subnets that reside in that particular Availability Zone.

Since the application tier and database tier should not be accessible from the Internet, they should both be in a private subnet of the VPC. The issue here is that the database servers should not be in the same subnet as the application servers. Therefore, they will each have their own private subnet. So in total that makes 3 subnets already.

The catch though is found on the final requirement. The environment should be highly available and in the event that one Availability Zone goes down, there is another AZ available to handle the incoming traffic. Redundancy would solve the matter by deploying the environment in two Availability Zones. For the database tier, you can set up a master-slave replication between the two instances across the two AZs.

Hence, the correct answer is 6 subnets.

Question 25:

**You are working for a major financial firm in Wall Street where you are tasked to design an application architecture for their online trading platform which should have high availability and fault tolerance. The application is using an Amazon S3 bucket located in the us-east-1 region to store large amounts of intraday financial data.**

**To avoid any costly service disruptions, what will you do to ensure that the stored financial data in the S3 bucket would not be affected even if there is an outage in one of the Availability Zones or a regional service failure in us-east-1?**

A. Copy the S3 bucket to an EBS-backed EC2 instance.
B. Create a Lifecycle Policy to regularly backup the S3 bucket to Amazon Glacier.
C. Use AWS Storage Gateway to keep a backup of the data.
D. Do nothing since the S3 bucket can withstand an outage in one of the Availability Zones and even regional service failures.
**E. Enable Cross-Region Replication.(Correct)**

EXPLANATION

In this scenario, you need to enable Cross-Region Replication to ensure that your S3 bucket would not be affected even if there is an outage in one of the Availability Zones or a regional service failure in us-east-1. When you upload your data in S3, your objects are redundantly stored on multiple devices across multiple facilities within the region only, where you created the bucket. Hence, if there is an outage on the entire region, your S3 bucket will be unavailable if you do not enable Cross-Region Replication, which should make your data available to another region.

Note that an Availability Zone (AZ) is more related with Amazon EC2 instances rather than Amazon S3 so if there is any outage in the AZ, the S3 bucket is usually not affected but only the EC2 instances deployed on that zone.

Question 26:

**A popular augmented reality (AR) mobile game is heavily using a RESTful API which is hosted in AWS. The API uses Amazon API Gateway and a DynamoDB table with a preconfigured read and write capacity. Based on your systems monitoring, the DynamoDB table begins to throttle requests during high peak loads which causes the slow performance of the game.**

**Which of the following can you do to improve the performance of your app?**

A. Integrate an Application Load Balancer with your DynamoDB table.
B. Add the DynamoDB table to an Auto Scaling Group.
**C. Use DynamoDB Auto Scaling (Correct)**
D. Create an SQS queue in front of the DynamoDB table.

EXPLANATION

DynamoDB auto scaling uses the AWS Application Auto Scaling service to dynamically adjust provisioned throughput capacity on your behalf, in response to actual traffic patterns. This enables a table or a global secondary index to increase its provisioned read and write capacity to handle sudden increases in traffic, without throttling. When the workload decreases, Application Auto Scaling decreases the throughput so that you don't pay for unused provisioned capacity.

Option 3 is the best answer. DynamoDB Auto Scaling uses the AWS Application Auto Scaling service to dynamically adjust provisioned throughput capacity on your behalf.

Option 1 is  because an Application Load Balancer is not suitable to be used with DynamoDB and in addition, this will not increase the throughput of your DynamoDB table.

Option 2 is  because you usually put EC2 instances on an Auto Scaling Group, and not a DynamoDB table.

Option 4 is  because this is not a design principle for high throughput DynamoDB table. Using SQS is for handling queuing and polling the request. This will not increase the throughput of DynamoDB which is required in this situation.

Question 27:

**A data analytics company, which uses machine learning to collect and analyze consumer data, is using Redshift cluster as their data warehouse. You are instructed to implement a disaster recovery plan for their systems to ensure business continuity even in the event of an AWS region outage.**

**Which of the following is the best approach to meet this requirement?**

A. Create a scheduled job that will automatically take the snapshot of your Redshift Cluster and store it to an S3 bucket. Restore the snapshot in case of an AWS region outage.
B. Do nothing because Amazon Redshift is a highly available, fully managed data warehouse which can withstand an outage of an entire AWS region.
C. Use Automated snapshots of your Redshift Cluster.
**D. Enable Cross-Region Snapshots Copy in your Amazon Redshift Cluster. (Correct)**

EXPLANATION

You can configure Amazon Redshift to copy snapshots for a cluster to another region. To configure cross-region snapshot copy, you need to enable this copy feature for each cluster and configure where to copy snapshots and how long to keep copied automated snapshots in the destination region. When cross-region copy is enabled for a cluster, all new manual and automatic snapshots are copied to the specified region.

Option 1 is  because although this option is possible, this entails a lot of manual work and hence, not the best option. You should configure cross-region snapshot copy instead.

Option 2 is  because although Amazon Redshift is a fully-managed data warehouse, you will still need to configure cross-region snapshot copy to ensure that your data is properly replicated to another region.

Option 3 is  because using automated snapshots is not enough and will not be available in case the entire AWS region is down.

Question 28:

**There is a new compliance rule in your company that audits every Windows and Linux EC2 instances each month to view any performance issues. They have more than a hundred EC2 instances running in production, and each must have a logging function that collects various system details regarding**

**that instance. The SysOps team will periodically review these logs and analyze their contents using AWS Analytics tools, and the result will need to be retained in an S3 bucket.**

**In this scenario, what is the most efficient way to collect and analyze logs from the instances with minimal effort?**

A. **Install the unified CloudWatch Logs agent in each instance which will automatically collect and push data to CloudWatch Logs. Analyze the log data with CloudWatch Logs Insights. (Correct)**
B. Install AWS SDK in each instance and create a custom daemon script that would collect and push data to CloudWatch Logs periodically. Enable CloudWatch detailed monitoring and use CloudWatch Logs Insights to analyze the log data of all instances.
C. Install the AWS Systems Manager Agent (SSM Agent) in each instance which will automatically collect and push data to CloudWatch Logs. Analyze the log data with CloudWatch Logs Insights.
D. Install AWS Inspector Agent in each instance which will collect and push data to CloudWatch Logs periodically. Set up a CloudWatch dashboard to properly analyze the log data of all instances.

EXPLANATION

To collect logs from your Amazon EC2 instances and on-premises servers into CloudWatch Logs, AWS offers both a new unified CloudWatch agent, and an older CloudWatch Logs agent. It is recommended to use the unified CloudWatch agent which has the following advantages:

- You can collect both logs and advanced metrics with the installation and configuration of just one agent.

- The unified agent enables the collection of logs from servers running Windows Server.

- If you are using the agent to collect CloudWatch metrics, the unified agent also enables the collection of additional system metrics, for in-guest visibility.

- The unified agent provides better performance.

CloudWatch Logs Insights enables you to interactively search and analyze your log data in Amazon CloudWatch Logs. You can perform queries to help you quickly and effectively respond to operational issues. If an issue occurs, you can use CloudWatch Logs Insights to identify potential causes and validate deployed fixes.

CloudWatch Logs Insights includes a purpose-built query language with a few simple but powerful commands. CloudWatch Logs Insights provides sample queries, command descriptions, query autocompletion, and log field discovery to help you get started quickly. Sample queries are included for several types of AWS service logs.

Option 2 is . Although this is a valid solution, this entails a lot of effort to implement as you have to allocate time to install the AWS SDK to each instance and develop a custom monitoring solution. Remember that the question is specifically looking for a

solution that can be implemented with minimal effort. In addition, it is unnecessary and not cost-efficient to enable detailed monitoring in CloudWatch in order to meet the requirements of this scenario since this can be done using CloudWatch Logs.

Option 3 is  as although this is also a valid solution, it is more efficient to use CloudWatch agent than an SSM agent. Manually connecting to an instance to view log files and troubleshoot an issue with SSM Agent is time-consuming hence, for more efficient instance monitoring, you can use the CloudWatch Agent instead to send the log data to Amazon CloudWatch Logs.

Option 4 is  because AWS Inspector is simply a security assessments service which only helps you in checking for unintended network accessibility of your EC2 instances and for vulnerabilities on those EC2 instances. Furthermore, setting up an Amazon CloudWatch dashboard is not suitable since its primarily used for scenarios where you have to monitor your resources in a single view, even those resources that are spread across different AWS Regions. It is better to use CloudWatch Logs Insights instead since it enables you to interactively search and analyze your log data.

Question 29:

**You are a Solutions Architect working for a large multinational investment bank. They have a web application that requires a minimum of 4 EC2 instances to run to ensure that it can cater to its users across the globe. You are instructed to ensure fault tolerance of this system.**

**Which of the following is the best option?**

A. **Deploy an Auto Scaling group with 2 instances in each of 3 Availability Zones behind an Application Load Balancer. (Correct)**
B. Deploy an Auto Scaling group with 2 instances in each of 2 Availability Zones behind an Application Load Balancer.
C. Deploy an Auto Scaling group with 4 instances in one Availability Zone behind an Application Load Balancer.
D. Deploy an Auto Scaling group with 1 instance in each of 4 Availability Zones behind an Application Load Balancer.

EXPLANATION

Fault Tolerance is the ability of a system to remain in operation even if some of the components used to build the system fail. In AWS, this means that in the event of server fault or system failures, the number of running EC2 instances should not fall below the minimum number of instances required by the system for it to work properly. So if the the application requires a minimum of 4 instances, there should be at least 4 instances running in case there is an outage in one of the Availability Zones or if there are server issues.

One of the differences between Fault Tolerance and High Availability is that, the former refers to the minimum number of running instances. For example, you have a system that requires a minimum of 4 running instances and currently has 6 running instances deployed in two Availability Zones. There was a component failure in one of the Availability Zones which knocks out 3 instances. In this case, the system can still be regarded as Highly Available since there are still instances running that can accomodate the requests. However, it is not Fault Tolerant since the required minimum of four instances have not been met.

As such, Option 1 is the correct answer because even if there was an outage in one of the Availability Zones, the system still satisfies the requirement of a minimum of 4 running instances.

Option 2 is  because if one Availability Zone went out, there will only be 2 running instances available out of the required 4 minimum instances. Although the Auto Scaling group can spin up another 2 instances, the fault tolerance of the web application has already been compromised.

Option 3 is  because if the Availability Zone went out, there will be no running instance available to accommodate the request.

Option 4 is  because if one Availability Zone went out, there will only be 3 instances available to accommodate the request.

Question 30:

You are a Cloud Migration Engineer in a media company which uses EC2, ELB, and S3 for its video-sharing portal for filmmakers. They are using a standard S3 storage class to store all high-quality videos that are frequently accessed only during the first three months of posting. What should you do if the company needs to automatically transfer or archive media data from an S3 bucket to Glacier?

A. Use a custom shell script that transfers data from the S3 bucket to Glacier
**B. Use Lifecycle Policies(Correct)**
C. Use AWS SQS
D. Use AWS SWF

EXPLANATION

You can create a lifecycle policy in S3 to automatically transfer your data to Glacier.

Lifecycle configuration enables you to specify the lifecycle management of objects in a bucket. The configuration is a set of one or more rules, where each rule defines an action for Amazon S3 to apply to a group of objects.

These actions can be classified as follows:

•Transition actions – In which you define when objects transition to another storage class. For example, you may choose to transition objects to the STANDARD_IA (IA,

for infrequent access) storage class 30 days after creation, or archive objects to the GLACIER storage class one year after creation.

•Expiration actions – In which you specify when the objects expire. Then Amazon S3 deletes the expired objects on your behalf.

Question 31:

**You are working as a Solutions Architect for a major supermarket store chain. They have an e-commerce application which is running in eu-east-2 region that strictly requires six EC2 instances running at all times. In that region, there are 3 Availability Zones (AZ) - eu-east-2a, eu-east-2b, and eu-east-2c that you can use.**

**Which of the following deployments provide 100% fault tolerance if any single AZ in the region becomes unavailable?**

A. eu-east-2a with two EC2 instances, eu-east-2b with four EC2 instances, and eu-east-2c with two EC2 instances
B. eu-east-2a with two EC2 instances, eu-east-2b with two EC2 instances, and eu-east-2c with two EC2 instances
C. eu-east-2a with four EC2 instances, eu-east-2b with two EC2 instances, and eu-east-2c with two EC2 instances
**D. eu-east-2a with six EC2 instances, eu-east-2b with six EC2 instances, and eu-east-2c with no EC2 instances(Correct)**
**E. eu-east-2a with three EC2 instances, eu-east-2b with three EC2 instances, and eu-east-2c with three EC2 instances(Correct)**

EXPLANATION

Fault Tolerance is the ability of a system to remain in operation even if some of the components used to build the system fail. In AWS, this means that in the event of server fault or system failures, the number of running EC2 instance should not fall below the minimum number of instances required by the system for it to work properly. So if the the application requires a minimum of 4 instances, there should be at least 4 instances running in case there is an outage in one of the Availability Zones or server issues.

In this scenario, you have to simulate a situation where one Availability Zone became unavailable for each option and check whether it still has 6 running instances. Hence, the correct answers are Options 4 and 5 because even if there is an outage in one of the Availability Zones, there are still 6 running instances:

1.eu-east-2a with six EC2 instances, eu-east-2b with six EC2 instances, and eu-east-2c with no EC2 instances

2.eu-east-2a with three EC2 instances, eu-east-2b with three EC2 instances, and eu-east-2c with three EC2 instances

Question 32:

**You are working as a Solutions Architect for a startup in which you are tasked to develop a custom messaging service that will also be used to train their AI for an automatic response feature which they plan to implement in the future. Based on their research and tests, the service can receive up to thousands of messages a day, and all of these data are to be sent to Amazon EMR for further processing. It is crucial that none of the messages will be lost, no duplicates will be produced and that they are processed in EMR in the same order as their arrival.**

**Which of the following options should you implement to meet the startup's requirements?**

A. **Create an Amazon Kinesis Data Stream to collect the messages.(Correct)**
B. Set up a default Amazon SQS queue to handle the messages.
C. Set up an Amazon SNS Topic to handle the messages.
D. Create a pipeline using AWS Data Pipeline to handle the messages.

EXPLANATION

Two important requirements that the chosen AWS service should fulfill is that data should not go missing, is durable, and streams data in the sequence of arrival. Kinesis can do the job just fine because of its architecture. A Kinesis data stream is a set of shards that has a sequence of data records, and each data record has a sequence number that is assigned by Kinesis Data Streams. Kinesis can also easily handle the high volume of messages being sent to the service.

Amazon Kinesis Data Streams enables real-time processing of streaming big data. It provides ordering of records, as well as the ability to read and/or replay records in the same order to multiple Amazon Kinesis Applications. The Amazon Kinesis Client Library (KCL) delivers all records for a given partition key to the same record processor, making it easier to build multiple applications reading from the same Amazon Kinesis data stream (for example, to perform counting, aggregation, and filtering).

Option 2 is  because although SQS is a valid messaging service, it is not suitable for scenarios where you need to process the data based on the order they were received. Take note that a default queue in SQS is just a standard queue and not a FIFO (First-In-First-Out) queue. In addition, SQS does not guarantee that no duplicates will be sent.

Option 3 is  because SNS is a pub-sub messaging service in AWS. SNS might not be capable of handling such a large volume of messages being received and sent at a time. It does not also guarantee that the data will be transmitted in the same order they were received.

Option 4 is  because Data pipeline is primarily used as a cloud-based data workflow service that helps you process and move data between different AWS services and on-premises data sources. It is not suitable for collecting data from distributed sources such as users, IoT devices, or clickstreams.

Question 33:

**A Fortune 500 company which has numerous offices and customers around the globe has hired you as their Principal Architect. You have staff and customers that upload gigabytes to terabytes of data to a centralized S3 bucket from the regional data centers, across continents, all over the world on a regular basis. At the end of the financial year, there are thousands of data being uploaded to the central S3 bucket which is in ap-southeast-2 (Sydney) region and a lot of employees are starting to complain about the slow upload times. You were instructed by the CTO to resolve this issue as soon as possible to avoid any delays in processing their global end of financial year (EOFY) reports.**

**Which feature in Amazon S3 enables fast, easy, and secure transfer of your files over long distances between your client and your Amazon S3 bucket?**

A. Reduced Redundancy Storage (RRS)
B. Cross-Region Replication
**C. Transfer Acceleration(Correct)**
D. Multipart Upload

EXPLANATION

Amazon S3 Transfer Acceleration enables fast, easy, and secure transfer of files over long distances between your client and your Amazon S3 bucket. Transfer Acceleration leverages Amazon CloudFront's globally distributed AWS Edge Locations. As data arrives at an AWS Edge Location, data is routed to your Amazon S3 bucket over an optimized network path.

Question 34:

**A leading IT consulting company has an application which processes a large stream of financial data by an Amazon ECS Cluster then stores the result to a DynamoDB table. You have to design a solution to detect new entries in the DynamoDB table then automatically trigger a Lambda function to run some tests to verify the processed data.**

**What solution can be easily implemented to alert the Lambda function of new entries while requiring minimal configuration change to your architecture?**

A. Use CloudWatch Alarms to trigger the Lambda function whenever a new entry is created in the DynamoDB table.
B. Invoke the Lambda functions using SNS each time that the ECS Cluster successfully processed financial data.
**C. Enable DynamoDB Streams to capture table activity and automatically trigger the Lambda function.(Correct)**
D. Use Systems Manager Automation to detect new entries in the DynamoDB table then automatically invoke the Lambda function for processing.

EXPLANATION

Amazon DynamoDB is integrated with AWS Lambda so that you can create triggers—pieces of code that automatically respond to events in DynamoDB Streams. With triggers, you can build applications that react to data modifications in DynamoDB tables.

If you enable DynamoDB Streams on a table, you can associate the stream ARN with a Lambda function that you write. Immediately after an item in the table is modified, a new record appears in the table's stream. AWS Lambda polls the stream and invokes your Lambda function synchronously when it detects new stream records.

You can create a Lambda function which can perform a specific action that you specify, such as sending a notification or initiating a workflow. For instance, you can set up a Lambda function to simply copy each stream record to persistent storage, such as EFS or S3, to create a permanent audit trail of write activity in your table.

Suppose you have a mobile gaming app that writes to a TutorialsDojoCourses table. Whenever the TopCourse attribute of the TutorialsDojoScores table is updated, a corresponding stream record is written to the table's stream. This event could then trigger a Lambda function that posts a congratulatory message on a social media network. (The function would simply ignore any stream records that are not updates to TutorialsDojoCourses or that do not modify the TopCourse attribute.)

Hence, Option 3 is the correct answer because the requirement can be met with minimal configuration change using DynamoDB streams which can automatically trigger Lambda functions whenever there is a new entry.

Option 1 is  because CloudWatch Alarms only monitor service metrics, not changes in DynamoDB table data.

Option 2 is  because you don't need to create an SNS topic just to invoke Lambda functions. You can enable DynamoDB streams instead to meet the requirement with less configuration.

Option 4 is  because the Systems Manager Automation service is primarily used to simplify common maintenance and deployment tasks of Amazon EC2 instances and

other AWS resources. It does not have the capability to detect new entries in a DynamoDB table.

Question 35:

**You have a web-based order processing system which is currently using a queue in Amazon SQS. The support team noticed that there are a lot of cases where an order was processed twice. This issue has caused a lot of trouble in your processing and made your customers very unhappy. Your IT Manager has asked you to ensure that this issue does not happen again.**

**What can you do to prevent this from happening again in the future?**

A. Alter the retention period in Amazon SQS.
B. Alter the visibility timeout of SQS.
**C. Replace Amazon SQS and instead, use Amazon Simple Workflow service.(Correct)**
D. Change the message size in SQS.

EXPLANATION

The main issue here is that the order management system produces duplicate orders at times. Since the company is using SQS, there is a possibility that a message can have a duplicate in case an EC2 instance failed to delete the already processed message. To prevent this issue from happening, you have to use Amazon Simple Workflow service instead of SQS.

For standard queues, the visibility timeout isn't a guarantee against receiving a message twice. Hence, Option 2 is . To avoid duplicate SQS messages, it is better to design your applications to be idempotent (they should not be affected adversely when processing the same message more than once).

Amazon SWF helps developers build, run, and scale background jobs that have parallel or sequential steps. You can think of Amazon SWF as a fully-managed state tracker and task coordinator in the Cloud. If your app's steps take more than 500 milliseconds to complete, you need to track the state of processing, and you need to recover or retry if a task fails.

Question 36:

**You are working as a Junior Solutions Architect where you are responsible in enhancing the availability and durability of the database instances in your VPC. Your company has a Multi-AZ RDS instance in the ap-northeast-1 region. If a storage volume on the primary instance fails in a Multi-AZ deployment, Amazon RDS automatically initiates a failover to the up-to-date standby instance.**

**In case of a failover, which record in Route 53 is changed?**

A. CAA
**B. CNAME(Correct)**
C. TXT
D. MX

EXPLANATION

Failover is automatically handled by Amazon RDS so that you can resume database operations as quickly as possible without administrative intervention. When failing over, Amazon RDS simply flips the canonical name record (CNAME) in Route53 for your DB instance to point at the standby, which in turn is promoted to become the new primary.

Question 37:

**You are working as an IT Consultant for a transportation agency of the government where you were hired to design and build their online portal. There would be thousands of contracts, permits, and other financial documents that would be submitted to and processed by the portal 24 hours a day, 7 days a week, which is why you have to ensure the reliability of your cloud architecture in case of any infrastructure issues.**

**Which AWS services should you use to build a fault-tolerant and highly available architecture?**

A. Amazon DynamoDB
**B. Amazon Elastic Compute Cloud (EC2)(Correct)**
**C. Amazon Elastic Load Balancing(Correct)**
D. Amazon Simple Notification Service (SNS)
E. Amazon Simple Storage Service (S3)
F. Amazon Certificate Manager

EXPLANATION

EC2 instances placed in different Availability Zones are both logically and physically separated, and they provide an easy-to-use model for deploying your application across data centers for both high availability and reliability.

Elastic Load Balancers (ELB) allow you to spread the load across multiple Availability Zones and Amazon EC2 Auto Scaling groups for redundancy and decoupling of services. It provides high availability such that if one of its Availability Zones failed, it can direct the request to another healthy Availability Zone to avoid any downtime. Hence, Options 2 and 3 are the correct answers.

The scenario says that you are building an online portal of a government agency. Although it is true that S3 and DynamoDB are both fault-tolerant and highly

available, the requirement says that you have to select two services where you can build a highly available and fault-tolerant architecture.

Since the government agency will be having an online portal that will accept and process various documents, then we could conclude that they need a server to host the API and the public facing website. S3 can be used as a data storage as well as a static website, but since the requirement says that it is for an online portal which should be dynamic, then using S3 might not be the best option since it needs to accept documents and generate output, which can't be handled by a static website hosted in S3.

DynamoDB can certainly be used as well, but this service is only limited to provide a NoSQL database. Using a combination of S3 and DynamoDB may be valid but this has certain restrictions unlike EC2 + ELB where you can support more database types, dynamic apps and many others. And even without Auto Scaling, you can still attach multiple EC2 instances to your ELB to ensure high availability. Don't forget as well that EC2 Spot Fleet has an automatic scaling feature as well, which removes the need for an Auto Scaling / Launch configuration.

Question 38:

**You are working as a Solutions Architect for a leading data analytics company in which you are tasked to process real-time streaming data of your users across the globe. This will enable you to track and analyse globally distributed user activity on your website and mobile applications, including click stream analysis. Your cloud architecture should process the data in close geographical proximity to your users and to respond to user requests at low latencies.**

**Which of the following options is the most ideal solution that you should implement?**

A. Use a CloudFront web distribution and Route 53 with a latency-based routing policy, in order to process the data in close geographical proximity to users and respond to user requests at low latencies. Process real-time streaming data using Kinesis and durably store the results to an Amazon S3 bucket.
B. Integrate CloudFront with Lambda@Edge in order to process the data in close geographical proximity to users and respond to user requests at low latencies. Process real-time streaming data using Amazon Athena and durably store the results to an Amazon S3 bucket.
C. Use a CloudFront web distribution and Route 53 with a Geoproximity routing policy in order to process the data in close geographical proximity to users and respond to user requests at low latencies. Process real-time streaming data using Kinesis and durably store the results to an Amazon S3 bucket.
**D. Integrate CloudFront with Lambda@Edge in order to process the data in close geographical proximity to users and respond to user requests at low latencies. Process real-time streaming data using Kinesis and durably store the results to an Amazon S3 bucket.(Correct)**

EXPLANATION

Lambda@Edge is a feature of Amazon CloudFront that lets you run code closer to users of your application, which improves performance and reduces latency. With Lambda@Edge, you don't have to provision or manage infrastructure in multiple locations around the world. You pay only for the compute time you consume - there is no charge when your code is not running.

With Lambda@Edge, you can enrich your web applications by making them globally distributed and improving their performance — all with zero server administration. Lambda@Edge runs your code in response to events generated by the Amazon CloudFront content delivery network (CDN). Just upload your code to AWS Lambda, which takes care of everything required to run and scale your code with high availability at an AWS location closest to your end user.

By using Lambda@Edge and Kinesis together, you can process real-time streaming data so that you can track and analyze globally-distributed user activity on your website and mobile applications, including clickstream analysis. Hence, Option 4 is the correct answer in this scenario.

Options 1 and 3 are both  because you can only route traffic using Route 53 since it does not have any computing capability. This solution would not be able to process and return the data in close geographical proximity to your users since it is not using Lambda@Edge.

Option 2 is  because although using Lambda@Edge is correct, Amazon Athena is just an interactive query service that enables you to easily analyze data in Amazon S3 using standard SQL. Kinesis should be used to process the streaming data in real-time.

Question 39:

**You are managing a global news website which is deployed to AWS and is using MySQL RDS. The website has millions of viewers from all over the world which means that the website has read-heavy database workloads.**

**In this scenario, which of the following is the best option to use to increase the read throughput on the MySQL database?**

A. Enable Multi-AZ deployments
B. Enable Amazon RDS Standby Replicas
**C. Enable Amazon RDS Read Replicas(Correct)**
D. Use SQS to queue up the requests

EXPLANATION

Amazon RDS Read Replicas provide enhanced performance and durability for database (DB) instances. This feature makes it easy to elastically scale out beyond

the capacity constraints of a single DB instance for read-heavy database workloads. You can create one or more replicas of a given source DB Instance and serve high-volume application read traffic from multiple copies of your data, thereby increasing aggregate read throughput. Read replicas can also be promoted when needed to become standalone DB instances. Read replicas are available in Amazon RDS for MySQL, MariaDB, and PostgreSQL as well as Amazon Aurora.

Option 1 is  because the Multi-AZ deployments feature is mainly used to achieve high availability and failover support for your database.

Option 2 is  because a Standby replica is used in Multi-AZ deployments and hence, it is not a solution to reduce read-heavy database workloads.

Option 4 is  because although an SQS queue can effectively manage the requests, it won't be able to entirely improve the read-throughput of the database by itself.

Question 40:

**You are a Big Data Engineer who is assigned to handle the online enrollment system database of a prestigious university, which is hosted in RDS. You are required to monitor the database metrics in Amazon CloudWatch to ensure the availability of the enrollment system.**

**What are the enhanced monitoring metrics that Amazon CloudWatch provides for Amazon RDS DB instances?**

A. The amount of available random access memory.
B. The average number of disk I/O operations per second during the polling period.
C. The percentage of CPU utilization.
**D. RDS child processes.(Correct)**
**E. OS processes(Correct)**

EXPLANATION

Amazon RDS provides metrics in real time for the operating system (OS) that your DB instance runs on. You can view the metrics for your DB instance using the console, or consume the Enhanced Monitoring JSON output from CloudWatch Logs in a monitoring system of your choice.

CloudWatch gathers metrics about CPU utilization from the hypervisor for a DB instance, and Enhanced Monitoring gathers its metrics from an agent on the instance. As a result, you might find differences between the measurements, because the hypervisor layer performs a small amount of work. The differences can be greater if your DB instances use smaller instance classes, because then there are likely more virtual machines (VMs) that are managed by the hypervisor layer on a single physical instance. Enhanced Monitoring metrics are useful when you want to see how different processes or threads on a DB instance use the CPU.

In RDS, the Enhanced Monitoring metrics shown in the Process List view are organized as follows:

-RDS child processes – Shows a summary of the RDS processes that support the DB instance, for example aurora for Amazon Aurora DB clusters and mysqld for MySQL DB instances. Process threads appear nested beneath the parent process. Process threads show CPU utilization only as other metrics are the same for all threads for the process. The console displays a maximum of 100 processes and threads. The results are a combination of the top CPU consuming and memory consuming processes and threads. If there are more than 50 processes and more than 50 threads, the console displays the top 50 consumers in each category. This display helps you identify which processes are having the greatest impact on performance.

-RDS processes – Shows a summary of the resources used by the RDS management agent, diagnostics monitoring processes, and other AWS processes that are required to support RDS DB instances.

-OS processes – Shows a summary of the kernel and system processes, which generally have minimal impact on performance.

Question 41:

**You are a Solutions Architect working for a large multi-national bank in the Asia-Pacific region. You designed an application architecture that is deployed to AWS, which has four Reserved EC2 instances. To be able to securely and easily manage these instances, you created a bastion host in your VPC. When your CTO found out, he was concerned and asked you about what you have done.**

**How will you describe what a bastion host is to your boss?**

A. A bastion host is an EC2 instance in a private subnet of your VPC and is typically accessed using SSH or RDP. Once remote connectivity has been established with the bastion host, it then acts as a 'jump' server that allows you to use SSH or RDP to log into other EC2 instances deployed in public subnets.
B. A bastion host is an EC2 instance in a public subnet of your VPC and is typically accessed using SSH or RDP. Once remote connectivity has been established with the bastion host, it then acts as a 'jump' server, allowing you to use HTTPS to log into other EC2 instances deployed in private subnets.
**C. A bastion host is an EC2 instance in a public subnet of your VPC and is typically accessed using SSH or RDP. Once remote connectivity has been established with a bastion host, it then acts as a 'jump' server, allowing you to use SSH or RDP to log into other EC2 instances deployed in private subnets.(Correct)**
D. A bastion host is an EC2 instance in a private subnet of your VPC and is typically accessed using SSH or RDP. Once remote connectivity has been

established with the bastion host, it then acts as a 'jump' server, allowing you to use HTTPS to log into other EC2 instances deployed in public subnets.

EXPLANATION

A bastion host is basically an EC2 instance in the public subnet of your VPC and is typically accessed using SSH or RDP, which are used as a jump server to other EC2 instances and other AWS resources within other subnets. A bastion is a special purpose server instance that is designed to be the primary access point from the Internet and acts as a proxy to your other EC2 instances which are preferably deployed in a private subnet.

Question 42:

**Your company recently decided to adopt a hybrid cloud infrastructure with AWS to take advantage of its global infrastructure of Availability Zones, Regions, and Edge Locations. What is the difference between an Availability Zone and an Edge location?**

A. Edge locations are used as central control stations for your AWS resources deployed on all regions.
B. An edge location is used as a link when building load balancing between regions
**C. Availability Zones are collections of data centres that run on physically distinct, independent infrastructure within an AWS region while an Edge location is used to deliver cached content to the closest location to reduce latency.(Correct)**
D. An availability zone is a grouping of AWS resources in a specific region while an edge location is a specific resource within the AWS region

EXPLANATION

Availability Zones are collections of data centres that run on physically distinct, independent infrastructure. Availability Zones are engineered to be highly reliable. Common points of failure such as generators and cooling equipment are not shared between Availability Zones. Availability Zones are also physically separate so that even an extreme disaster such as a fire, tornado, or flood will affect only the single Availability Zone where it occurred.

Amazon EC2 is hosted in multiple locations worldwide. These locations are composed of regions and Availability Zones. Each region is a separate geographic area. Each region has multiple, isolated locations known as Availability Zones.

Edge location is used to deliver cached content to the closest location to reduce latency. It delivers your content through a worldwide network of data centers called edge locations. These are primarily used by the Amazon CloudFront service.

Question 43:

**A web application is hosted in an Auto Scaling group of EC2 instances deployed across multiple Availability Zones in front of an Application Load Balancer. You need to implement an SSL solution for your system to improve its security which is why you requested an SSL/TLS certificate from a third-party certificate authority (CA).**

**Where can you safely import the SSL/TLS certificate of your application?**

- **A. AWS Certificate Manager (Correct)**
- B. **IAM certificate store (Correct**)
- C. A private S3 bucket with versioning enabled
- D. An S3 bucket configured with server-side encryption with customer-provided encryption keys (SSE-C)
- E. CloudFront

EXPLANATION

If you got your certificate from a third-party CA, import the certificate into ACM or upload it to the IAM certificate store. Hence, Options 1 and 2 are the correct answers.

ACM lets you import third-party certificates from the ACM console, as well as programmatically. If ACM is not available in your region, use AWS CLI to upload your third-party certificate to the IAM certificate store.

Options 3 and 4 are  as S3 is not a suitable service to store the SSL certificate.

Option 5 is  because although you can upload certificates to CloudFront, it doesn't mean that you can import SSL certificates on it. You would not be able to export the certificate that you have loaded in CloudFront nor assign them to your EC2 or ELB instances as it would be tied to a single CloudFront distribution.

Question 44:

**You are working for a large media company that has a single 3-TB volume storage on their on-premises network that is used to hold their digital footages, films, and other files. The storage is growing at 500 GB a year and must be presented as a single logical volume. The company is becoming increasingly constrained with their local storage capacity and wants an off-site backup of this data, while maintaining low-latency access to their frequently accessed data.**

**Which AWS Storage Gateway configuration meets the customer requirements?**

A. **AWS Storage Gateway - Cached volumes with snapshots scheduled to Amazon S3(Correct)**
B. AWS Storage Gateway - Stored volumes with snapshots scheduled to Amazon S3
C. AWS Storage Gateway - Virtual Tape Library with snapshots to Amazon S3
D. AWS Storage Gateway - Virtual Tape Library with snapshots to Amazon Glacier

EXPLANATION

In Cached volumes, you store your data in Amazon Simple Storage Service (Amazon S3) and retain a copy of frequently accessed data subsets locally. 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.

Question 45:

**You are working for a large global media company with multiple office locations all around the world. You are instructed to build a system to distribute training videos to all employees. Using CloudFront, what method would be used to serve content that is stored in S3, but not publicly accessible from S3 directly?**

A. **Create an Origin Access Identity (OAI) for CloudFront and grant access to the objects in your S3 bucket to that OAI. (Correct)**
B. Create an Identity and Access Management (IAM) user for CloudFront and grant access to the objects in your S3 bucket to that IAM user.
C. Create an S3 bucket policy that lists the CloudFront distribution ID as the principal and the target bucket as the Amazon Resource Name (ARN).
D. Add the CloudFront account security group.

EXPLANATION

When you create or update a distribution in CloudFront, you can add an origin access identity (OAI) and automatically update the bucket policy to give the origin access identity permission to access your bucket. Alternatively, you can choose to manually change the bucket policy or change ACLs, which control permissions on individual objects in your bucket.

You can update the Amazon S3 bucket policy using either the AWS Management Console or the Amazon S3 API:

• Grant the CloudFront origin access identity the applicable permissions on the bucket.

• Deny access to anyone that you don't want to have access using Amazon S3 URLs.

Question 46:

**You are working as a Solutions Architect for a leading financial firm where you are responsible in ensuring that their applications are highly available and safe from common web security vulnerabilities. Which is the most suitable AWS service to use to mitigate Distributed Denial of Service (DDoS) attacks from hitting your back-end EC2 instances?**

A. AWS WAF
**B. AWS Shield (Correct)**
C. AWS Firewall Manager
D. Amazon GuardDuty

EXPLANATION

AWS Shield is a managed Distributed Denial of Service (DDoS) protection service that safeguards applications running on AWS. AWS Shield provides always-on detection and automatic inline mitigations that minimize application downtime and latency, so there is no need to engage AWS Support to benefit from DDoS protection. There are two tiers of AWS Shield - Standard and Advanced.

All AWS customers benefit from the automatic protections of AWS Shield Standard, at no additional charge. AWS Shield Standard defends against most common, frequently occurring network and transport layer DDoS attacks that target your web site or applications. When you use AWS Shield Standard with Amazon CloudFront and Amazon Route 53, you receive comprehensive availability protection against all known infrastructure (Layer 3 and 4) attacks.

Option 1 is  because AWS WAF is a web application firewall service that helps protect your web apps from common exploits that could affect app availability, compromise security, or consume excessive resources. Although this can help you against DDoS attacks, AWS WAF alone is not enough to fully protect your VPC. You still need to use AWS Shield in this scenario.

Option 3 is  because AWS Firewall Manager just simplifies your AWS WAF administration and maintenance tasks across multiple accounts and resources.

Option 4 is  because Amazon GuardDuty is an intelligent threat detection service to protect your AWS accounts and workloads. Using this alone will not fully protect your AWS resources against DDoS attacks.

Question 47:

**Your company wants to host a static website on Amazon S3 using a bucket named "tutorialsdojo" in the Asia Pacific (Sydney) region. What website URL will be assigned to the S3 bucket?**

- **A. tutorialsdojo.s3-website-ap-southeast-2.amazonaws.com(Correct)**
- B. ap-southeast-2.s3-website-tutorialsdojo.amazonaws.com
- C. tutorialsdojo.s3-website-ap-southeast-2.amazon.aws.com
- D. ap-southeast-2.s3-website-tutorialsdojo.amazon.aws.com

EXPLANATION

To host a static website, you configure an Amazon S3 bucket for website hosting, and then upload your website content to the bucket. The website is then available at the AWS Region-specific website endpoint of the bucket, which is in one of the following formats:

<bucket-name>.s3-website-<AWS-region>.amazonaws.com

Hence, the correct answer is option A:

Question 48:

**An application is using a RESTful API hosted in AWS which uses Amazon API Gateway and AWS Lambda. There is a requirement to trace and analyze user requests as they travel through your Amazon API Gateway APIs to the underlying services.**

**Which of the following is the most suitable service to use to meet this requirement?**

- A. VPC Flow Logs
- B. CloudWatch
- C. CloudTrail
- **D. AWS X-Ray (Correct)**

EXPLANATION

You can use AWS X-Ray to trace and analyse user requests as they travel through your Amazon API Gateway APIs to the underlying services. API Gateway supports AWS X-Ray tracing for all API Gateway endpoint types: regional, edge-optimized, and private. You can use AWS X-Ray with Amazon API Gateway in all regions where X-Ray is available.

X-Ray gives you an end-to-end view of an entire request, so you can analyse latencies in your APIs and their backend services. You can use an X-Ray service map to view the latency of an entire request and that of the downstream services that are integrated with X-Ray. And you can configure sampling rules to tell X-Ray which requests to record, at what sampling rates, according to criteria that you specify. If you call an API Gateway API from a service that's already being traced, API Gateway passes the trace through, even if X-Ray tracing is not enabled on the API.

You can enable X-Ray for an API stage by using the API Gateway management console, or by using the API Gateway API or CLI.

Option 1 is  because VPC Flow Logs is a feature that enables you to capture information about the IP traffic going to and from network interfaces in your entire VPC. Although it can capture some details about the incoming user requests, it is still better to use AWS X-Ray as it provides a better way to debug and analyze your microservices applications with request tracing so you can find the root cause of your issues and performance.

Option 2 is  because CloudWatch is a monitoring and management service. It does not have the capability to trace and analyze user requests as they travel through your Amazon API Gateway APIs.

Option 3 is  because CloudTrail is primarily used for API logging of all of your AWS resources.

Question 49:

**You are using an On-Demand EC2 instance to host a legacy web application that uses an Amazon Instance Store-Backed AMI. The web application should be decommissioned as soon as possible and hence, you need to terminate the EC2 instance.**

**When the instance is terminated, what happens to the data on the root volume?**

A. Data is automatically saved as an EBS snapshot.
B. Data is automatically saved as an EBS volume.
C. Data is unavailable until the instance is restarted.
**D. Data is automatically deleted.(Correct)**

EXPLANATION

AMIs are categorized as either backed by Amazon EBS or backed by instance store. The former means that the root device for an instance launched from the AMI is an Amazon EBS volume created from an Amazon EBS snapshot. The latter means that the root device for an instance launched from the AMI is an instance store volume created from a template stored in Amazon S3.

Option 4 is the correct answer because the data on instance store volumes persist only during the life of the instance which means that if the instance is terminated, the data will be automatically deleted.

Question 50:

**You are working for a large bank that is developing a web application that receives large amounts of object data. They are using the data to generate a report for their stockbrokers to use on a daily basis. Unfortunately, a recent financial crisis has left the bank short on cash and cannot afford to purchase expensive storage hardware. They had resorted to use AWS instead.**

**Which is the best service to use in order to store a virtually unlimited amount of object data without any effort to scale when demand unexpectedly increases?**

A. **Amazon S3(Correct)**
B. Amazon Glacier
C. Amazon Import/Export
D. Amazon EC2
E. DynamoDB

EXPLANATION

In this scenario, you can use Amazon S3 and Amazon Glacier as a storage service. And since we are looking for the best option, we have to consider that the object data being stored by the bank is used on a daily basis as well. Hence, Amazon S3 is the better choice as it provides frequent access to your object data.

Amazon S3 is a durable, secure, simple, and fast storage service designed to make web-scale computing easier for developers. Use Amazon S3 if you need low latency or frequent access to your data. Use Amazon Glacier if low storage cost is paramount, and you do not require millisecond access to your data.

Question 51:

**To save cost, a company decided to change their third-party data analytics tool to a cheaper solution. They sent a full data export on a CSV file which contains all of their analytics information. You then save the CSV file to an S3 bucket for storage. Your manager asked you to do some validation on the provided data export.**

**In this scenario, what is the most cost-effective and easiest way to analyze export data using a standard SQL?**

A. Create a migration tool to load the CSV export file from S3 to a DynamoDB instance. Once the data has been loaded, run queries using DynamoDB.

B. Use mysqldump client utility to load the CSV export file from S3 to a MySQL RDS instance. Run some SQL queries once the data has been loaded to complete your validation.
**C. To be able to run SQL queries, use AWS Athena to analyze the export data file in S3. (Correct)**
D. Use a migration tool to load the CSV export file from S3 to a database which is designed for online analytic processing (OLAP) such as AWS RedShift. Run some queries once the data has been loaded to complete your validation.

EXPLANATION

Amazon Athena is an interactive query service that makes it easy to analyze data directly in Amazon Simple Storage Service (Amazon S3) using standard SQL. With a few actions in the AWS Management Console, you can point Athena at your data stored in Amazon S3 and begin using standard SQL to run ad-hoc queries and get results in seconds.

Athena is serverless, so there is no infrastructure to set up or manage, and you pay only for the queries you run. Athena scales automatically—executing queries in parallel—so results are fast, even with large datasets and complex queries.

Athena helps you analyze unstructured, semi-structured, and structured data stored in Amazon S3. Examples include CSV, JSON, or columnar data formats such as Apache Parquet and Apache ORC. You can use Athena to run ad-hoc queries using ANSI SQL, without the need to aggregate or load the data into Athena.

Hence, the most cost-effective and appropriate answer in this scenario is Option 3: Using AWS Athena.

Options 1, 2 and 4 are all  because it is not necessary to set up a database to be able to analyze the CSV export file. You can use a cost-effective option (AWS Athena), which is a serverless service that enables you to pay only for the queries you run.

Question 52:

**You have a prototype web application that uses one Spot EC2 instance. What will happen to the instance by default if it gets interrupted by Amazon EC2 for capacity requirements?**

**A. The instance will be terminated (Correct)**
B. The instance will be stopped
C. The instance will be restarted
D. This is not possible as only On-Demand instances can be interrupted by Amazon EC2

EXPLANATION

The main differences are that:

1. Spot instances typically offer a significant discount off the On-Demand prices

2. Your instances can be interrupted by Amazon EC2 for capacity requirements with a 2-minute notification

3. Spot prices adjust gradually based on long term supply and demand for spare EC2 capacity.

You can choose to have your Spot instances terminated, stopped, or hibernated upon interruption. Stop and hibernate options are available for persistent Spot requests and Spot Fleets with the maintain option enabled. By default, your instances are terminated hence, option 1 is the correct answer.

Question 53:

**You are planning to reduce the amount of data that Amazon S3 transfers to your servers in order to lower your operating costs as well as to lower the latency of retrieving the data. To accomplish this, you need to use simple structured query language (SQL) statements to filter the contents of Amazon S3 objects and retrieve just the subset of data that you need.**

**Which of the following services will help you accomplish this requirement?**

A. RDS
B. Redshift Spectrum
**C. S3 Select (Correct)**
D. AWS Step Functions

EXPLANATION

With Amazon S3 Select, you can use simple structured query language (SQL) statements to filter the contents of Amazon S3 objects and retrieve just the subset of data that you need. By using Amazon S3 Select to filter this data, you can reduce the amount of data that Amazon S3 transfers, which reduces the cost and latency to retrieve this data.

Amazon S3 Select works on objects stored in CSV, JSON, or Apache Parquet format. It also works with objects that are compressed with GZIP or BZIP2 (for CSV and JSON objects only), and server-side encrypted objects. You can specify the format of the results as either CSV or JSON, and you can determine how the records in the result are delimited.

Option 1 is  because although RDS is an SQL database where you can perform SQL operations, it is still not valid because you want to apply SQL transactions on S3 itself, and not on the database, which RDS cannot do.

Option 2 is  because although Amazon Redshift Spectrum provides a similar in-query functionality like S3 Select, this service is more suitable for querying your Redshift clusters and not your S3 buckets. The Redshift queries are run on your cluster resources against local disk. Redshift Spectrum queries run using per-query scale-out resources against data in S3 which can entail additional costs compared with S3 Select.

Option 4 is  because Step functions only let you coordinate multiple AWS services into serverless workflows so you can build and update apps quickly.

Question 54:

**A tech company is currently using Amazon Simple Workflow (SWF) service with a default configuration for their order processing system. The system works fine but you noticed that some of the orders seem to be stuck for almost 4 weeks.**

**What could be the possible reason for this?**

- **A. It is because SWF is waiting human input from an activity task. (Correct)**
- B. The workflow has exceeded SWF's 15-day maximum workflow execution time.
- C. The workflow has exceeded SWF's 14-day maximum workflow execution time.
- D. SWF should be restarted.

EXPLANATION

By default, each workflow execution can run for a maximum of 1 year in Amazon SWF. This means that it is possible that in your workflow, there are some tasks which require manual action that renders it idle. As a result, some orders get stuck for almost 4 weeks.

Amazon SWF does not take any special action if a workflow execution is idle for an extended period of time. Idle executions are subject to the timeouts that you configure. For example, if you have set the maximum duration for an execution to be 1 day, then an idle execution will be timed out if it exceeds the 1 day limit. Idle executions are also subject to the Amazon SWF limit on how long an execution can run (1 year).

Options 2 and 3 are  as the maximum execution time is 1 year.

Option 4 is  as there is no problem with SWF and you can't manually restart this service.

Question 55:

**You had recently set up a CloudWatch Alarm that performs status checks on your EBS volume. However, you noticed that the volume check has a status of insufficient-data. What does this status mean?**

A. All EBS Volume checks have failed.
B. The EBS Volume has been abruptly stopped.
C. All EBS Volume checks have been completed.
**D. The check on the EBS volume is still in progress.(Correct)**

EXPLANATION

Volume status checks are automated tests that run every 5 minutes and return a pass or fail status.

If all checks pass, the status of the volume is ok. Option 3 is, therefore, .

If a check fails, the status of the volume is impaired. Option 1 is, therefore, .

If the status is insufficient-data, the checks may still be in progress on the volume. Option 4 is, therefore, correct.

There is no status code for option 2, which is also an  choice.

You can view the results of volume status checks to identify any impaired volumes and take any necessary actions.

Question 56:

**A data analytics application requires a service that can collect, process, and analyze clickstream data from various websites in real-time. Which of the following is the most suitable service to use for the application?**

**A. Kinesis (Correct)**
B. Redshift Spectrum
C. AWS Glue
D. Amazon EMR

EXPLANATION

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.

Option 2 is  because Redshift Spectrum is primarily used to directly query open data formats stored in Amazon S3 without the need for unnecessary data movement, which enables you to analyze data across your data warehouse and data lake, together, with a single service. It does not provide the ability to process your data in real-time, unlike Kinesis.

Option 3 is  because 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. It does not provide the ability to process your data in real-time, unlike Kinesis.

Option 4 is  because Amazon EMR is a web service that uses an open-source Hadoop framework, to quickly & cost-effectively process vast amounts of data. It does not provide the ability to process your data in real-time, unlike Kinesis.

Question 57:

**You are working for a media company and you need to configure an Amazon S3 bucket to serve static assets for your public-facing web application. Which methods ensure that all of the objects uploaded to the S3 bucket can be read publicly all over the Internet?**

- **A. In S3, set the permissions of the object to public read during upload. (Correct)**
- B. Configure the ACL of the S3 bucket to set all objects to be publicly readable and writeable.
- **C. Configure the S3 bucket policy to set all objects to public read. (Correct)**
- D. Create an IAM role to set the objects inside the S3 bucket to public read.
- E. Do nothing. Amazon S3 objects are already public by default.

EXPLANATION

By default, all Amazon S3 resources such as buckets, objects, and related subresources are private which means that only the AWS account holder (resource owner) that created it has access to the resource. The resource owner can optionally grant access permissions to others by writing an access policy. In S3, you also set the permissions of the object during upload to make it public.

Amazon S3 offers access policy options broadly categorized as resource-based policies and user policies. Access policies you attach to your resources (buckets and objects) are referred to as resource-based policies.

For example, bucket policies and access control lists (ACLs) are resource-based policies. You can also attach access policies to users in your account. These are called user policies. You may choose to use resource-based policies, user policies, or some combination of these to manage permissions to your Amazon S3 resources.

Option 2 is  as ACLs are primarily used to grant basic read/write permissions to AWS accounts and not suitable for providing public access over the Internet.

Option 4 is . Although with IAM, you can create a user, group, or role that has certain permissions to the S3 bucket, it does not control the individual objects that are hosted in the bucket.

Option 5 is  because by default, all the S3 resources are private, so only the AWS account that created the resources can access them.

Question 58:

**You have a requirement to integrate the Lightweight Directory Access Protocol (LDAP) directory service of your on-premises data center to your AWS VPC using IAM. The identity store which is currently being used is not compatible with SAML.**

**Which of the following provides the most valid approach to implement the integration?**

A. Use an IAM policy that references the LDAP identifiers and AWS credentials.
B. Use AWS Single Sign-On (SSO) service to enable single sign-on between AWS and your LDAP.
**C. Develop an on-premises custom identity broker application and use STS to issue short-lived AWS credentials. (Correct)**
D. Use IAM roles to rotate the IAM credentials whenever LDAP credentials are updated.

EXPLANATION

If your identity store is not compatible with SAML 2.0, then you can build a custom identity broker application to perform a similar function. The broker application authenticates users, requests temporary credentials for users from AWS, and then provides them to the user to access AWS resources.

The application verifies that employees are signed into the existing corporate network's identity and authentication system, which might use LDAP, Active Directory, or another system. The identity broker application then obtains temporary security credentials for the employees.

To get temporary security credentials, the identity broker application calls either AssumeRole or GetFederationToken to obtain temporary security credentials, depending on how you want to manage the policies for users and when the temporary credentials should expire. The call returns temporary security credentials consisting of an AWS access key ID, a secret access key, and a session token. The identity broker application makes these temporary security credentials available to the internal company application. The app can then use the temporary credentials to

make calls to AWS directly. The app caches the credentials until they expire, and then requests a new set of temporary credentials.

Option 1 is  because using an IAM policy is not enough to integrate your LDAP service to IAM. You need to use SAML, STS or a custom identity broker.

Option 2 is  because the scenario did not require SSO and in addition, the identity store that you are using is not SAML-compatible.

Option 4 is  because manually rotating the IAM credentials is not an optimal solution to integrate your on-premises and VPC network. You need to use SAML, STS or a custom identity broker.

Question 59:

**You are working as a Solutions Architect for a financial firm which is building an internal application that processes loans, accruals, and interest rates for their clients. They require a durable storage service that is able to handle future increases in storage capacity and can provide the lowest-latency access to their data. Their web application will be hosted in a single m5ad.24xlarge Reserved EC2 instance which will process and store data to the storage service.**

**Which of the following would be the most suitable storage service that you should use to meet this requirement?**

- **A. EBS(Correct)**
- B. Storage Gateway
- C. S3
- D. EFS

EXPLANATION

Amazon Web Services (AWS) offers cloud storage services to support a wide range of storage workloads such as Amazon S3, EFS and EBS. 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. Amazon S3 is an object storage service. Amazon S3 makes data available through an Internet API that can be accessed anywhere. 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.

You can also increase EBS storage for up to 16TB or add new volumes for additional storage.

In this scenario, the company is looking for a storage service which can provide the lowest-latency access to their data which will be fetched by a single m5ad.24xlarge Reserved EC2 instance. This type of workloads can be supported better by using either EFS or EBS but in this case, the latter is the most suitable storage service. As mentioned above, EBS provides the lowest-latency access to the data for your EC2 instance since the volume is directly attached to the instance. In addition, the scenario does not require concurrently-accessible storage since they only have one instance. Hence, the correct answer is Option 1.

Option 2 is  since Storage Gateway is primarily used to extend your on-premises storage to your AWS Cloud.

Option 3 is  because although S3 is also highly available and highly scalable, it still does not provide the lowest-latency access to the data, unlike EBS. Remember that S3 does not reside within your VPC by default, which means the data will traverse the public Internet that may result to higher latency. You can set up a VPC Endpoint for S3 yet still, its latency is greater than that of EBS.

Option 4 is  because the scenario does not require concurrently-accessible storage since the internal application is only hosted in one instance. Although EFS can provide low latency data access to the EC2 instance as compared with S3, the storage service that can provide the lowest latency access is still EBS.

Question 60:

**You launched an EC2 instance in your newly created VPC. You have noticed that the generated instance does not have an associated DNS hostname.**

**Which of the following options could be a valid reason for this issue?**

A. The newly created VPC has an invalid CIDR block.
B. Amazon Route53 is not enabled.
**C. The DNS resolution and DNS hostname of the VPC configuration should be enabled. (Correct)**
D. The security group of the EC2 instance needs to be modified.

EXPLANATION

When you launch an EC2 instance into a default VPC, AWS provides it with public and private DNS hostnames that correspond to the public IPv4 and private IPv4 addresses for the instance.

However, when you launch an instance into a non-default VPC, AWS provides the instance with a private DNS hostname only. New instances will only be provided with public DNS hostname depending on these two DNS attributes: the DNS resolutionand DNS hostnames, that you have specified for your VPC, and if your instance has a public IPv4 address.

In this case, the new EC2 instance does not automatically get a DNS hostname because the DNS resolution and DNS hostnames attributes are disabled in the newly created VPC.

Option 1 is  since it's very unlikely that a VPC has an invalid CIDR block because of AWS validation schemes.

Option 2 is  since Route 53 does not need to be enabled. Route 53 is the DNS service of AWS, but the VPC is the one that enables assigning of instance hostnames.

Option 4 is  since security groups are just firewalls for your instances. They filter traffic based on a set of security group rules.

Question 61:

**You have a distributed application in AWS that periodically processes large volumes of data across multiple instances. You designed the application to recover gracefully from any instance failures. You are required to launch the application in the most cost-effective way.**

**Which type of EC2 instance will meet your requirements?**

- **A. Spot Instances (Correct)**
- B. Reserved instances
- C. Dedicated instances
- D. On-Demand instances

EXPLANATION

You require an EC2 instance that is the most cost-effective among other types. In addition, the application it will host is designed to gracefully recover in case of instance failures.

In terms of cost-effectiveness, Spot and Reserved instances are the top options. And since the application can gracefully recover from instance failures, the Spot instance is the best option for this case as it is the cheapest type of EC2 instance. Remember

that when you use Spot Instances, there will be interruptions. Amazon EC2 can interrupt your Spot Instance when the Spot price exceeds your maximum price, when the demand for Spot Instances rise, or when the supply of Spot Instances decreases. This makes Option 1 the correct answer.

Option 2 is  because although you could also use reserved instances to save costs, it entails a commitment of 1-year or 3-year terms of usage. Since your processes only run periodically, you won't be able to maximize the discounted price of using reserved instances.

Options 3 and 4 are also  because Dedicated and on-demand instances are not a cost-effective solution to use for your application.

Question 62:

**A tech startup has recently received a Series A round of funding to continue building their mobile forex trading application. You are hired to set up their cloud architecture in AWS and to implement a highly available, fault tolerant system. For their database, they are using DynamoDB and for authentication, they have chosen to use Cognito. Since the mobile application contains confidential financial transactions, there is a requirement to add a second authentication method that doesn't rely solely on user name and password.**

**How can you implement this in AWS?**

A. **Add multi-factor authentication (MFA) to a user pool in Cognito to protect the identity of your users.(Correct)**
B. Add a new IAM policy to a user pool in Cognito.
C. Integrate Cognito with Amazon SNS Mobile Push to allow additional authentication via SMS.
D. Develop a custom application that integrates with Cognito that implements a second layer of authentication.

EXPLANATION

You can add multi-factor authentication (MFA) to a user pool to protect the identity of your users. MFA adds a second authentication method that doesn't rely solely on user name and password. You can choose to use SMS text messages, or time-based one-time (TOTP) passwords as second factors in signing in your users. You can also use adaptive authentication with its risk-based model to predict when you might need another authentication factor. It's part of the user pool advanced security features, which also include protections against compromised credentials.

Question 63:

**There are a few, easily reproducible but confidential files that your client wants to store in AWS without worrying about storage capacity. For the first month, all of these files will be accessed frequently but after that, they will rarely be accessed at all. The old files will only be accessed by developers so there is no set retrieval time requirement. However, the files under a specific tutorialsdojo-finance prefix in the S3 bucket will be used for post-processing that requires millisecond retrieval time.**

**Given these conditions, which of the following options would be the most cost-effective solution for your client's storage needs?**

A. Store the files in S3 then after a month, change the storage class of the bucket to S3-IA using lifecycle policy.
B. Store the files in S3 then after a month, change the storage class of the bucket to Intelligent-Tiering using lifecycle policy.
**C. Store the files in S3 then after a month, change the storage class of the tutorialsdojo-finance prefix to One Zone-IA while the remaining go to Glacier using lifecycle policy. (Correct)**
D. Store the files in S3 then after a month, change the storage class of the tutorialsdojo-finance prefix to S3-IA while the remaining go to Glacier using lifecycle policy.

EXPLANATION

Initially, the files will be accessed frequently, and S3 is a durable and highly available storage solution for that. After a month has passed, the files won't be accessed frequently anymore, so it is a good idea to use lifecycle policies to move them to a storage class that would have a lower cost for storing them.

Since the files are easily reproducible and some of them are needed to be retrieved quickly based on a specific prefix filter (tutorialsdojo-finance), S3-One Zone IA would be a good choice for storing them. The other files that do not contain such prefix would then be moved to Glacier for low cost archival. This setup would also be the most cost-effective for the client. Hence, the correct answer is Option 3.

Option 1 is  because although it is valid to move the files to S3-IA, this solution still costs more compared with using a combination of S3-One Zone IA and Glacier.

Option 2 is  because while S3 Intelligent-Tiering can automatically move data between two access tiers (frequent access and infrequent access) when access patterns change, it is more suitable for scenarios where you don't know the access patterns of your data. It may take some time for S3 Intellgent-Tiering to analyze the access patterns before it moves the data to a cheaper storage class like S3-IA which means you may still end up paying more in the beginning. In addition, you already

know the access patterns of the files which means you can directly change the storage class immediately and save cost right away.

Option 4 is  because although S3-IA costs less than S3 Standard storage class, it is still more expensive than S3-One Zone IA. Remember that the files are easily reproducible so you can safely move the data to S3-One Zone IA and in case there is an outage, you can simply generate the missing data again.

Question 64:

**You are working as a Network Engineer for an electronics and communications company in Japan. You are told to implement a NAT instance in your VPC to allow certain EC2 instances to initiate connections to the Internet but restrict any requests coming from the Internet.**

**In this scenario, what is the best way to configure a fault-tolerant NAT instance in your VPC?**

A. Launch a NAT Gateway in a public subnet. Alternatively, you can also create a NAT instance in one private subnet.
B. Launch a NAT instance in the public subnet and add a route from the private subnet to that NAT instance.
**C. Launch two NAT instances in two separate public subnets and add a route from the private subnet to each NAT instance to make it more fault tolerant. (Correct)**
D. Launch two NAT instances in a public subnet and add a route from the private subnet to each NAT instance to make it more fault tolerant.

EXPLANATION

You can use a NAT device to enable instances in a private subnet to connect to the Internet (for example, for software updates) or other AWS services, but prevent the Internet from initiating connections with the instances. A NAT device forwards traffic from the instances in the private subnet to the Internet or other AWS services, and then sends the response back to the instances. When traffic goes to the Internet, the source IPv4 address is replaced with the NAT device's address and similarly, when the response traffic goes to those instances, the NAT device translates the address back to those instances' private IPv4 addresses.

Option 1 is  because you should not be putting the NAT instances in private subnet as they need to communicate with the Internet. They should be in public subnet.

Option 2 is  because you would need at least two NAT instances for fault tolerance.

Option 3 is correct because you should place two NAT instances in two separate public subnets, and create route from instances via each NAT instance for achieving fault tolerance.

Option 4 is  because if you put both NAT instances in a single public subnet and that subnet becomes unavailable or unreachable to the other instances, the architecture would not be fault tolerant.

Question 65:

**A bank portal application is hosted in an Auto Scaling group of EC2 instances behind a Classic Load Balancer (CLB). You are required to set up the architecture so that any back-end EC2 instances that you de-register should complete the in-progress requests first before the de-registration process takes effect. Conversely, if a back-end instance fails health checks, the load balancer should not send any new requests to the unhealthy instance but should allow existing requests to complete.**

**How will you configure your load balancer to satisfy the above requirement?**

A. Configure Sticky Sessions
B. Configure both Cross-Zone Load Balancing and Sticky Sessions
**C. Configure Connection Draining (Correct)**
D. Configure Proxy Protocol

EXPLANATION

To ensure that a Classic Load Balancer stops sending requests to instances that are de-registering or unhealthy while keeping the existing connections open, use connection draining. This enables the load balancer to complete in-flight requests made to instances that are de-registering or unhealthy. Hence, Option 3 is correct.

When you enable connection draining, you can specify a maximum time for the load balancer to keep connections alive before reporting the instance as de-registered. The maximum timeout value can be set between 1 and 3,600 seconds (the default is 300 seconds). When the maximum time limit is reached, the load balancer forcibly closes connections to the de-registering instance.

Option 1 is  because the sticky sessions feature is mainly used to ensure that all requests from the user during the session are sent to the same instance.

Option 2 is  because configuring both Cross-Zone Load Balancing and Sticky Sessions will still not satisfy the requirement. Cross-Zone load balancing is mainly used to distribute requests evenly across the registered instances in all enabled Availability Zones. You have to enable Connection Draining.

Option 4 is  because Proxy Protocol is an Internet protocol used to carry connection information from the source requesting the connection to the destination for which the connection was requested.