Question 1

A company uses an Amazon Redshift cluster to run its analytical workloads. Corporate policy requires that the company's data be encrypted at rest with customer managed keys. The company's disaster recovery plan requires that backups of the cluster be copied into another AWS Region on a regular basis.

How should a database specialist automate the process of backing up the cluster data in compliance with these policies?

Options:

A- Copy the AWS Key Management Service (AWS KMS) customer managed key from the source Region to the destination Region. Set up an AWS Glue job in the source Region to copy the latest snapshot of the Amazon Redshift cluster from the source Region to the destination Region. Use a time-based schedule in AWS Glue to run the job on a daily basis.

B- Create a new AWS Key Management Service (AWS KMS) customer managed key in the destination Region. Create a snapshot copy grant in the destination Region specifying the new key. In the source Region, configure cross-Region snapshots for the Amazon Redshift cluster specifying the destination Region, the snapshot copy grant, and retention periods for the snapshot.

C- Copy the AWS Key Management Service (AWS KMS) customer-managed key from the source Region to the destination Region. Create Amazon S3 buckets in each Region using the keys from their respective Regions. Use Amazon EventBridge (Amazon CloudWatch Events) to schedule an AWS Lambda function in the source Region to copy the latest snapshot to the S3 bucket in that Region. Configure S3 Cross-Region Replication to copy the snapshots to the destination Region, specifying the source and destination KMS key IDs in the replication configuration.

D- Use the same customer-supplied key materials to create a CMK with the same private key in the destination Region. Configure cross-Region snapshots in the source Region targeting the destination Region. Specify the corresponding CMK in the destination Region to encrypt the snapshot.

Answer:

B

Explanation:

According to the Amazon Redshift documentation1, you can enable database encryption for your clusters to help protect data at rest. You can use either AWS Key Management Service (AWS KMS) or a hardware security module (HSM) to manage the top-level encryption keys in this hierarchy. The process that Amazon Redshift uses for encryption differs depending on how you manage keys.

To copy encrypted snapshots across Regions, you need to create a snapshot copy grant in the destination Region and specify a CMK in that Region. You also need to configure cross-Region snapshots in the source Region and provide the destination Region, the snapshot copy grant, and retention periods for the snapshots. This way, you can automate the process of backing up the cluster data in compliance with the corporate policies.

Question 2

Question Type: **MultipleChoice**

A retail company uses Amazon Redshift for its 1 PB data warehouse. Several analytical workloads run on a Redshift cluster. The tables within the cluster have grown rapidly. End users are reporting poor performance of daily reports that run on the transaction fact tables.

A database specialist must change the design of the tables to improve the reporting performance. All the changes must be applied dynamically. The changes must have the least possible impact on users and must optimize the overall table size.

Which solution will meet these requirements?

Options:

A- Use the STL SCAN view to understand how the tables are getting scanned. Identify the columns that are used in filter and group by conditions. Create a temporary table with the identified columns as sort keys and compression as Zstandard (ZSTD) by copying the data from the original table. Drop the original table. Give the temporary table the same name that the original table had.

B- Run an explain plan to analyze the queries on the tables. Consider recommendations from Amazon Redshift Advisor. Identify the columns that are used in filter and group by conditions. Convert the recommended columns from Redshift Advisor into sort keys with compression encoding set to RAW. Set the rest of the column compression encoding to AZ64.

C- Run an explain plan to analyze the queries on the tables. Consider recommendations from Amazon Redshift Advisor. Identify the columns that are used in filter and group by conditions. Convert the recommended columns from Redshift Advisor into sort keys with compression encoding set to I_ZO. Set the rest of the column compression encoding to Zstandard (ZSTD).

D- Run an explain plan to analyze the queries on the tables. Consider recommendations from Amazon Redshift Advisor. Identify the columns that are used in filter and group by conditions. Create a deep copy of the table with the identified columns as sort keys and compression for all columns as Zstandard (ZSTD) by using a bulk insert. Drop the original table. Give the copy table the same name that the original table had.

Question 3

Question Type: MultipleChoice

A gaming company is building a mobile game that will have as many as 25,000 active concurrent users in the first 2 weeks after launch. The game has a leaderboard that shows the 10 highest scoring players over the last 24 hours. The leaderboard calculations are processed by an AWS Lambda function, which takes about 10 seconds. The company wants the data on the leaderboard to be no more than 1 minute old.

Which architecture will meet these requirements in the MOST operationally efficient way?

Options:

A- Deliver the player data to an Amazon Timestream database. Create an Amazon ElastiCache for Redis cluster. Configure the Lambda function to store the results in Redis. Create a scheduled event with Amazon EventBridge to invoke the Lambda function once every minute. Reconfigure the game server to query the Redis cluster for the leaderboard data.

B- Deliver the player data to an Amazon Timestream database. Create an Amazon DynamoDB table. Configure the Lambda function to store the results in DynamoDB. Create a scheduled event with Amazon EventBridge to invoke the Lambda function once every minute.

Reconfigure the game server to query the DynamoDB table for the leaderboard data.

C- Deliver the player data to an Amazon Aurora MySQL database. Create an Amazon DynamoDB table. Configure the Lambda function to store the results in MySQL. Create a scheduled event with Amazon EventBridge to invoke the Lambda function once every minute. Reconfigure the game server to query the DynamoDB table for the leaderboard data.

D- Deliver the player data to an Amazon Neptune database. Create an Amazon ElastiCache for Redis cluster. Configure the Lambda function to store the results in Redis. Create a scheduled event with Amazon EventBridge to invoke the Lambda function once every minute. Reconfigure the game server to query the Redis cluster for the leaderboard data.

Answer:

A

Explanation:

Amazon Timestream is a serverless time series database service that allows you to store and analyze time series data at any scale1.It is well suited for gaming applications that generate high volumes of data from player events, such as scores, achievements, and actions2.Amazon ElastiCache for Redis is a fully managed in-memory data store that provides fast and scalable performance for applications that need sub-millisecond latency3. It can be used as a cache layer to store frequently accessed data, such as leaderboard results, and reduce the load on the database. AWS Lambda is a serverless compute service that lets you run code without provisioning or managing servers. It can be used to process the data from Amazon Timestream and store the leaderboard results in Amazon ElastiCache for Redis. Amazon EventBridge is a serverless event bus service that makes it easy to connect your applications with data from a variety of sources. It can be used to create a scheduled event that triggers the Lambda function once every minute, ensuring that the leaderboard data is updated regularly. The game server can then query the Redis cluster for the leaderboard data, which will be no more than 1 minute old.

Option B is incorrect because Amazon DynamoDB is a key-value and document database that delivers single-digit millisecond performance at any scale. It is not designed for time series data, which requires efficient ingestion, compression, and querying of high-volume data streams. Option C is incorrect because Amazon Aurora is a relational database that combines the performance and availability of traditional enterprise databases with the simplicity and cost-effectiveness of open source databases. It is not optimized for time series data, which requires specialized indexing and partitioning techniques. Option D is incorrect because Amazon Neptune is a graph database that supports property graph and RDF models. It is not suitable for time series data, which requires high ingestion rates and temporal queries.

Question 4

Question Type: **MultipleChoice**

A database specialist needs to replace the encryption key for an Amazon RDS DB instance. The database specialist needs to take immediate action to ensure security of the database.

Which solution will meet these requirements?

Options:

A- Modify the DB instance to update the encryption key. Perform this update immediately without waiting for the next scheduled maintenance window.

B- Export the database to an Amazon S3 bucket. Import the data to an existing DB instance by using the export file. Specify a new encryption key during the import process.

C- Create a manual snapshot of the DB instance. Create an encrypted copy of the snapshot by using a new encryption key. Create a new DB instance from the encrypted snapshot.

D- Create a manual snapshot of the DB instance. Restore the snapshot to a new DB instance. Specify a new encryption key during the restoration process.

Answer:

D

Question 5

Question Type: MultipleChoice

A coffee machine manufacturer is equipping all of its coffee machines with 10T sensors. The 10T core application is writing measurements for each record to Amazon Timestream. The records have multiple dimensions and measures. The measures include multiple measure names and values.

An analysis application is running queries against the Timestream database and is focusing on data from the current week. A database specialist needs to optimize the query costs of the analysis application.

Which solution will meet these requirements?

Options:

A- Ensure that queries contain whole records over the relevant time range.

B- Use time range, measure name, and dimensions in the WHERE clause of the query.

C- Avoid canceling any query after the query starts running.

D- Implement exponential backoff in the application.

Answer:

B

Explanation:

Correct Answer: B. Use time range, measure name, and dimensions in the WHERE

Explanation from Amazon documents:

Amazon Timestream is a serverless time series database service that allows you to store and analyze time series data at any scale. To optimize the cost of queries, you should use the following best practices1:

Include only the measure and dimension names essential to query. Adding extraneous columns will increase data scans and therefore will also increase the query cost.

Include a time range in the WHERE clause of your query. For example, if you only need the last one hour of data in your dataset, include a time predicate such as time > ago (1h).

Include the measure names in the WHERE clause of the query when a query accesses a subset of measures in a table.

Option B follows these best practices, while option A does not. Option C is incorrect because canceling a query can save on cost if the query will not return the desired results1. Option D is irrelevant because exponential backoff is a technique to handle throttling errors, not to optimize query costs2.

Question 6

An online retailer uses Amazon DynamoDB for its product catalog and order dat

a. Some popular items have led to frequently accessed keys in the data, and the company is using DynamoDB Accelerator (DAX) as the caching solution to cater to the frequently accessed keys. As the number of popular products is growing, the company realizes that more items need to be cached. The company observes a high cache miss rate and needs a solution to address this issue.

What should a database specialist do to accommodate the changing requirements for DAX?

Options:

A- Increase the number of nodes in the existing DAX cluster.

B- Create a new DAX cluster with more nodes. Change the DAX endpoint in the application to point to the new cluster.

C- Create a new DAX cluster using a larger node type. Change the DAX endpoint in the application to point to the new cluster.

D- Modify the node type in the existing DAX cluster.

Answer:

C

Explanation:

Correct Answer: C. Create a new DAX cluster using a larger node type. Change

Explanation from Amazon documents:

The cache miss rate is the percentage of read requests that are not satisfied by the DAX cache and have to be forwarded to DynamoDB1. A high cache miss rate indicates that the DAX cluster does not have enough memory to store all the frequently accessed items. Increasing the number of nodes in the existing DAX cluster (option A) or creating a new DAX cluster with more nodes (option B) will not increase the total memory available for caching, because DAX uses a partitioned cache model, where each node is responsible for caching a subset of the data2. Modifying the node type in the existing DAX cluster (option D) will cause downtime and data loss, because DAX does not support online resizing of clusters3. Therefore, the best option is to create a new DAX cluster using a larger node type (option C), which will provide more memory per node and allow more items to be cached. The application will need to change the DAX endpoint to point to the new cluster, which can be done with minimal disruption by using DNS aliasing or load balancing3.

Question 7

A database specialist needs to enable IAM authentication on an existing Amazon Aurora PostgreSQL DB cluster. The database specialist already has modified the DB cluster settings, has created IAM and database credentials, and has distributed the credentials to the appropriate users.

What should the database specialist do next to establish the credentials for the users to use to log in to the DB cluster?

Options:

A- Add the users' IAM credentials to the Aurora cluster parameter group.

B- Run the generate-db-auth-token command with the user names to generate a temporary password for the users.

C- Add the users' IAM credentials to the default credential profile, Use the AWS Management Console to access the DB cluster.

D- Use an AWS Security Token Service (AWS STS) token by sending the IAM access key and secret key as headers to the DB cluster API endpoint.

Answer:

B

Explanation:

Correct Answer: B

Explanation from Amazon documents:

Amazon Aurora PostgreSQL supports IAM authentication, which is a method of using AWS Identity and Access Management (IAM) to manage database access. IAM authentication allows you to use IAM users and roles to control who can access your Aurora PostgreSQL DB cluster, instead of using a traditional database username and password. IAM authentication also provides more security by using temporary credentials that are automatically rotated.

To enable IAM authentication on an existing Aurora PostgreSQL DB cluster, the database specialist needs to do the following :

Modify the DB cluster settings to enable IAM database authentication. This can be done using the AWS Management Console, the AWS CLI, or the RDS API.

Create IAM and database credentials for each user who needs access to the DB cluster. The IAM credentials consist of an access key ID and a secret access key. The database credentials consist of a database username and an optional password. The IAM credentials and the database username must match.

Distribute the IAM and database credentials to the appropriate users. The users must keep their credentials secure and not share them with anyone else.

Run the generate-db-auth-token command with the user names to generate a temporary password for the users. This command is part of the AWS CLI and it generates an authentication token that is valid for 15 minutes. The authentication token is a string that has the

same format as a password. The users can use this token as their password when they connect to the DB cluster using a SQL client.

Therefore, option B is the correct solution to establish the credentials for the users to use to log in to the DB cluster. Option A is incorrect because adding the users' IAM credentials to the Aurora cluster parameter group is not necessary or possible. A cluster parameter group is a collection of DB engine configuration values that define how a DB cluster operates. Option C is incorrect because adding the users' IAM credentials to the default credential profile and using the AWS Management Console to access the DB cluster is not supported or secure. The default credential profile is a file that stores your AWS credentials for use by AWS CLI or SDKs. The AWS Management Console does not allow you to connect to an Aurora PostgreSQL DB cluster using IAM authentication. Option D is incorrect because using an AWS Security Token Service (AWS STS) token by sending the IAM access key and secret key as headers to the DB cluster API endpoint is not supported or secure. AWS STS is a service that enables you to request temporary, limited-privilege credentials for IAM users or federated users. The DB cluster API endpoint is an endpoint that allows you to perform administrative actions on your DB cluster using RDS API calls.

Question 8

Question Type: **MultipleChoice**

A company migrated an on-premises Oracle database to Amazon RDS for Oracle. A database specialist needs to monitor the latency of the database.

Which solution will meet this requirement with the LEAST operational overhead?

Options:

A- Publish RDS Performance insights metrics to Amazon CloudWatch. Add AWS CloudTrail filters to monitor database performance

B- Install Oracle Statspack. Enable the performance statistics feature to collect, store, and display performance data to monitor database performance.

C- Enable RDS Performance Insights to visualize the database load. Enable Enhanced Monitoring to view how different threads use the CPU

D- Create a new DB parameter group that includes the AllocatedStorage, DBInstanceClassMemory, and DBInstanceVCPU variables. Enable RDS Performance Insights

Answer:

C

Explanation:

Correct Answer: C

Explanation from Amazon documents:

Amazon RDS for Oracle is a fully managed relational database service that supports Oracle Database. Amazon RDS for Oracle provides several features to monitor the performance and health of your database, such as RDS Performance Insights, Enhanced Monitoring, Amazon CloudWatch, and AWS CloudTrail.

RDS Performance Insights is a feature that helps you quickly assess the load on your database and determine when and where to take action. RDS Performance Insights displays a dashboard that shows the database load in terms of average active sessions (AAS), which is the average number of sessions that are actively running SQL statements at any given time. RDS Performance Insights also shows the top SQL statements, waits, hosts, and users that are contributing to the database load.

Enhanced Monitoring is a feature that provides metrics in real time for the operating system (OS) that your DB instance runs on. Enhanced Monitoring metrics include CPU utilization, memory, file system, disk I/O, network I/O, process list, and thread count. Enhanced Monitoring allows you to view how different threads use the CPU and how much memory each thread consumes.

By enabling RDS Performance Insights and Enhanced Monitoring for the RDS for Oracle DB instance, the database specialist can monitor the latency of the database with the least operational overhead. This solution will allow the database specialist to use the RDS console or API to enable these features and view the metrics and dashboards without installing any additional software or tools. This solution will also provide comprehensive and granular information about the database load and resource utilization.

Therefore, option C is the correct solution to meet the requirement. Option A is not optimal because publishing RDS Performance Insights metrics to Amazon CloudWatch and adding AWS CloudTrail filters to monitor database performance will incur additional operational overhead and cost. Amazon CloudWatch is a service that collects monitoring and operational data in the form of logs, metrics, and events. AWS CloudTrail is a service that records AWS API calls for your account and delivers log files to you. These services are useful for monitoring performance trends and auditing activities, but they are not necessary for monitoring latency in real time. Option B is not optimal because installing Oracle Statspack and enabling the performance statistics feature will require manual intervention and configuration on the RDS for Oracle DB instance. Oracle Statspack is a tool that collects, stores, and displays performance data for Oracle Database. The performance statistics feature is an option that enables Statspack to collect additional statistics such as wait events, latches, SQL statements, segments, rollback segments, etc. These tools are useful for performance tuning and troubleshooting, but they are not as easy to use as RDS Performance Insights and Enhanced Monitoring. Option D is not relevant because creating a new DB parameter group that includes the AllocatedStorage, DBInstanceClassMemory, and DBInstanceVCPU variables will not help monitor the latency of the database. A DB parameter group is a collection of DB engine configuration values that

define how a DB instance operates. The AllocatedStorage parameter specifies the allocated storage size in gibibytes (GiB). The DBInstanceClassMemory parameter specifies the amount of memory available to an instance class in bytes. The DBInstanceVCPU parameter specifies the number of virtual CPUs available to an instance class. These parameters are used to configure the capacity and performance of a DB instance, but they do not provide any monitoring or metrics information. Enabling RDS Performance Insights alone will not provide enough information about the OS-level metrics such as CPU utilization or memory usage.

Question 9

A company performs an audit on various data stores and discovers that an Amazon S3 bucket is storing a credit card number. The S3 bucket is the target of an AWS Database Migration Service (AWS DMS) continuous replication task that uses change data capture (CDC). The company determines that this field is not needed by anyone who uses the target dat

a. The company has manually removed the existing credit card data from the S3 bucket.

What is the MOST operationally efficient way to prevent new credit card data from being written to the S3 bucket?

Options:

A- Add a transformation rule to the DMS task to ignore the column from the source data endpoint.

B- Add a transformation rule to the DMS task to mask the column by using a simple SQL query.

C- Configure the target S3 bucket to use server-side encryption with AWS KMS keys (SSE-KMS).

D- Remove the credit card number column from the data source so that the DMS task does not need to be altered.

Answer:

A