DAS-C01 Q&As

AWS Certified Data Analytics - Specialty (DAS-C01)

Pass Amazon DAS-C01 Exam with 100% Guarantee

Free Download Real Questions & Answers **PDF** and **VCE** file from:

https://www.leads4pass.com/das-c01.html

100% Passing Guarantee
100% Money Back Assurance

Following Questions and Answers are all new published by Amazon
Official Exam Center

⚙ **Instant Download** After Purchase

⚙ **100% Money Back** Guarantee

⚙ **365 Days** Free Update

⚙ **800,000+** Satisfied Customers

QUESTION 1

A central government organization is collecting events from various internal applications using Amazon Managed Streaming for Apache Kafka (Amazon MSK). The organization has configured a separate Kafka topic for each application to separate the data. For security reasons, the Kafka cluster has been configured to only allow TLS encrypted data and it encrypts the data at rest.

A recent application update showed that one of the applications was configured incorrectly, resulting in writing data to a Kafka topic that belongs to another application. This resulted in multiple errors in the analytics pipeline as data from different applications appeared on the same topic. After this incident, the organization wants to prevent applications from writing to a topic different than the one they should write to.

Which solution meets these requirements with the least amount of effort?

A. Create a different Amazon EC2 security group for each application. Configure each security group to have access to a specific topic in the Amazon MSK cluster. Attach the security group to each application based on the topic that the applications should read and write to.

B. Install Kafka Connect on each application instance and configure each Kafka Connect instance to write to a specific topic only.

C. Use Kafka ACLs and configure read and write permissions for each topic. Use the distinguished name of the clients\\' TLS certificates as the principal of the ACL.

D. Create a different Amazon EC2 security group for each application. Create an Amazon MSK cluster and Kafka topic for each application. Configure each security group to have access to the specific cluster.

Correct Answer: C

Reference: https://docs.aws.amazon.com/msk/latest/developerguide/msk-acls.html

QUESTION 2

A company creates daily and monthly business metrics from data that partners provide. Each day, the partners deliver JSON data files to an Amazon S3 bucket that the company owns. The S3 object keys use Apache Hive style date

partitions. The company uses an Amazon EventBridge rule to invoke an AWS Lambda function that reads all objects in the S3 bucket to aggregate the daily and monthly metrics.

The company performs occasional analysis that requires access to historical data. As more data has accumulated, the Lambda function is timing out frequently. A data analytics specialist must prevent the Lambda function timeouts.

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

A. Update the EventBridge rule to invoke AWS Step Functions to retry the Lambda function if the function fails.

B. Modify the Lambda function to delete older S3 objects during the daily processing.

C. Modify the Lambda function to query the S3 objects by using Amazon Athena with date filters.

D. Create an AWS Glue job to invoke the Lambda function. Update the EventBridge rule to invoke the AWS Glue job.

Correct Answer: B

QUESTION 3

A company is building an analytical solution that includes Amazon S3 as data lake storage and Amazon Redshift for data warehousing. The company wants to use Amazon Redshift Spectrum to query the data that is stored in Amazon S3. Which steps should the company take to improve performance when the company uses Amazon Redshift Spectrum to query the S3 data files? (Select THREE)

A. Use gzip compression with individual file sizes of 1-5 GB

B. Use a columnar storage file format

C. Partition the data based on the most common query predicates

D. Split the data into KB-sized files.

E. Keep all files about the same size.

F. Use file formats that are not splittable

Correct Answer: CDE

Reference: https://aws.amazon.com/blogs/big-data/10-best-practices-for-amazon-redshift-spectrum/

QUESTION 4

A company is reading data from various customer databases that run on Amazon RDS. The databases contain many inconsistent fields For example, a customer record field that is place_id in one database is location_id in another database. The company wants to link customer records across different databases, even when many customer record fields do not match exactly.

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

A. Create an Amazon EMR cluster to process and analyze data in the databases Connect to the Apache Zeppelin notebook, and use the FindMatches transform to find duplicate records in the data.

B. Create an AWS Glue crawler to crawl the databases. Use the FindMatches transform to find duplicate records in the data Evaluate and tune the transform by evaluating performance and results of finding matches

C. Create an AWS Glue crawler to crawl the data in the databases Use Amazon SageMaker to construct Apache Spark ML pipelines to find duplicate records in the data

D. Create an Amazon EMR cluster to process and analyze data in the databases. Connect to the Apache Zeppelin notebook, and use Apache Spark ML to find duplicate records in the data. Evaluate and tune the model by evaluating performance and results of finding duplicates

Correct Answer: B

QUESTION 5

An IoT company wants to release a new device that will collect data to track overnight sleep on an intelligent mattress. Sensors will send data that will be uploaded to an Amazon S3 bucket. Each mattress generates about 2 MB of data each night.

An application must process the data and summarize the data for each user. The application must make the results available as soon as possible. Every invocation of the application will require about 1 GB of memory and will finish running within 30 seconds.

Which solution will run the application MOST cost-effectively?

A. AWS Lambda with a Python script

B. AWS Glue with a Scala job

C. Amazon EMR with an Apache Spark script

D. AWS Glue with a PySpark job

Correct Answer: A

[DAS-C01 PDF Dumps](https://www.leads4pass.com)   [DAS-C01 VCE Dumps](https://www.leads4pass.com)   [DAS-C01 Braindumps](https://www.leads4pass.com)