

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.passapply.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 company has an application that uses the Amazon Kinesis Client Library (KCL) to read records from a Kinesis data stream.

After a successful marketing campaign, the application experienced a significant increase in usage. As a result, a data analyst had to split some shards in the data stream. When the shards were split, the application started throwing an

ExpiredIteratorExceptions error sporadically.

What should the data analyst do to resolve this?

A. Increase the number of threads that process the stream records.

B. Increase the provisioned read capacity units assigned to the stream\\'s Amazon DynamoDB table.

C. Increase the provisioned write capacity units assigned to the stream\\'s Amazon DynamoDB table.

D. Decrease the provisioned write capacity units assigned to the stream\\'s Amazon DynamoDB table.

Correct Answer: C

Reference: https://docs.aws.amazon.com/streams/latest/dev/troubleshooting-consumers.html#shard-iterator-expires-unexpectedly

QUESTION 2

A company is running Apache Spark on an Amazon EMR cluster. The Spark job writes data to an Amazon S3 bucket and generates a large number of PUT requests. The number of objects has increased over time.

After a recent increase in traffic, the Spark job started failing and returned an HTTP 503 Slow Down AmazonS3Exception error.

Which combination of actions will resolve this error? (Choose two.)

- A. Increase the number of S3 key prefixes for the S3 bucket.
- B. Increase the EMR File System (EMRFS) retry limit.
- C. Disable dynamic partition pruning in the Spark configuration for the cluster.
- D. Increase the repartitioning number for the Spark job.
- E. Increase the executor memory size on Spark.

Correct Answer: AC

QUESTION 3



A technology company has an application with millions of active users every day. The company queries daily usage data with Amazon Athena to understand how users interact with the application. The data includes the date and time, the location ID, and the services used. The company wants to use Athena to run queries to analyze the data with the lowest latency possible.

Which solution meets these requirements?

A. Store the data in Apache Avro format with the date and time as the partition, with the data sorted by the location ID.

B. Store the data in Apache Parquet format with the date and time as the partition, with the data sorted by the location ID.

C. Store the data in Apache ORC format with the location ID as the partition, with the data sorted by the date and time.

D. Store the data in .csv format with the location ID as the partition, with the data sorted by the date and time.

Correct Answer: B

Reference: https://cwiki.apache.org/confluence/display/hive/languagemanual+orc

QUESTION 4

A transport company wants to track vehicular movements by capturing geolocation records. The records are 10 B in size and up to 10,000 records are captured each second. Data transmission delays of a few minutes are acceptable, considering unreliable network conditions. The transport company decided to use Amazon Kinesis Data Streams to ingest the data. The company is looking for a reliable mechanism to send data to Kinesis Data Streams while maximizing the throughput efficiency of the Kinesis shards.

Which solution will meet the company\\'s requirements?

- A. Kinesis Agent
- B. Kinesis Producer Library (KPL)
- C. Kinesis Data Firehose
- D. Kinesis SDK
- Correct Answer: B

Reference: https://docs.aws.amazon.com/streams/latest/dev/developing-producers-with-sdk.htmls

QUESTION 5

A reseller that has thousands of AWS accounts receives AWS Cost and Usage Reports in an Amazon S3 bucket. The reports are delivered to the S3 bucket in the following format: //yyyymmdd-yyyymmdd/.parquet An AWS Glue crawler crawls the S3 bucket and populates an AWS Glue Data Catalog with a table. Business analysts use Amazon Athena to query the table and create monthly summary reports for the AWS accounts. The business analysts

are experiencing slow queries because of the accumulation of reports from the last 5 years. The business analysts want the operations team to make changes to improve query performance. Which action should the operations team take to meet these requirements?



- A. Change the file format to .csv.zip
- B. Partition the data by date and account ID
- C. Partition the data by month and account ID
- D. Partition the data by account ID, year, and month
- Correct Answer: A

Reference: https://docs.aws.amazon.com/cur/latest/userguide/access-cur-s3.html

Latest DAS-C01 Dumps

DAS-C01 Practice Test

DAS-C01 Study Guide