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.pass4itsure.com/das-c01.html

100% Passing Guarantee
100% Money Back Assurance

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

QUESTION 1

A social media company is using business intelligence tools to analyze its data for forecasting. The company is using Apache Kafka to ingest the low-velocity data in near-real time. The company wants to build dynamic dashboards with machine learning (ML) insights to forecast key business trends. The dashboards must provide hourly updates from data in Amazon S3. Various teams at the company want to view the dashboards by using Amazon QuickSight with ML insights. The solution also must correct the scalability problems that the company experiences when it uses its current architecture to ingest data.

Which solution will MOST cost-effectively meet these requirements?

A. Replace Kafka with Amazon Managed Streaming for Apache Kafka. Ingest the data by using AWS Lambda, and store the data in Amazon S3. Use QuickSight Standard edition to refresh the data in SPICE from Amazon S3 hourly and create a dynamic dashboard with forecasting and ML insights.

B. Replace Kafka with an Amazon Kinesis data stream. Use an Amazon Kinesis Data Firehose delivery stream to consume the data and store the data in Amazon S3. Use QuickSight Enterprise edition to refresh the data in SPICE from Amazon S3 hourly and create a dynamic dashboard with forecasting and ML insights.

C. Configure the Kafka-Kinesis-Connector to publish the data to an Amazon Kinesis Data Firehose delivery stream that is configured to store the data in Amazon S3. Use QuickSight Enterprise edition to refresh the data in SPICE from Amazon S3 hourly and create a dynamic dashboard with forecasting and ML insights.

D. Configure the Kafka-Kinesis-Connector to publish the data to an Amazon Kinesis Data Firehose delivery stream that is configured to store the data in Amazon S3. Configure an AWS Glue crawler to crawl the data. Use an Amazon Athena data source with QuickSight Standard edition to refresh the data in SPICE hourly and create a dynamic dashboard with forecasting and ML insights.

Correct Answer: B

Reference: https://noise.getoto.net/tag/amazon-kinesis-data-firehose/

QUESTION 2

A financial services company hosts its data warehouse on a Hadoop cluster located in an on-premises data center. The data is 300 TB in size and grows by 1 TB each day. The data is generated in real time from the company\\'s trading system. The raw data is transformed at the end of the trading day using a custom tool running on the Hadoop cluster.

The company is migrating its data warehouse to AWS using a managed data warehouse product provided by a third party that can ingest data from Amazon S3. The company has already established a 10 Gbps connection with an AWS Region using AWS Direct Connect. The company is required by its security and regulatory compliance policies to not transfer data over the public internet. The company wants to minimize changes to its custom tool for data transformation. The company also plans to eliminate the on-premises Hadoop cluster after the migration.

Which solution MOST cost-effectively meets these requirements?

A. Create a VPC endpoint for Amazon S3. Run a one-time copy job using the DistCp tool to copy existing files from Hadoop to a target S3 bucket over the VPC endpoint Schedule a nightly DistCp job on the Hadoop cluster to copy the incremental files produced by the custom tool to the target S3 bucket

B. Create a VPC endpoint for Amazon S3. Run a one-time copy job using the DistCp tool to copy existing files from Hadoop to a target S3 bucket over the VPC endpoint. Schedule a nightly job on the trading system servers that produces raw data to copy the incremental raw files to the target S3 bucket. Run the data transformation tool on a

transient Amazon EMR cluster to output files to Amazon S3.

C. Create a VPC endpoint for Amazon S3. Run a one-time copy job using the DistCp tool to copy existing files from Hadoop to a target S3 bucket over the VPC endpoint. Set up an Amazon Kinesis data stream to ingest raw data from the trading system in real time. Use Amazon Kinesis Data Analytics to transform the raw data and output files to Amazon S3.

D. Complete a one-time transfer of the data using AWS Snowball Edge devices transferring to a target S3 bucket. Schedule a nightly job on the trading system servers that produces raw data to copy the incremental raw files to the target S3 bucket Run the data transformation tool on a transient Amazon EMR cluster to output files to Amazon S3.

Correct Answer: A

QUESTION 3

A company has a producer application that collects device log data. The producer application writes to an Amazon Kinesis Data Firehose delivery stream that delivers data to an Amazon S3 bucket. The company needs to build a series of dashboards to display real-time trends of the metrics in the log data.

Which solution will meet these requirements?

A. Update the Kinesis Data Firehose delivery stream to add an Amazon OpenSearch Service (Amazon Elasticsearch Service) cluster as another destination. Use OpenSearch Dashboards (Kibana) for log data visualization.

B. Update the Kinesis Data Firehose delivery stream to add an Amazon Kinesis Data Analytics application as an additional destination. Use Amazon QuickSight to display the output of the Kinesis Data Analytics application.

C. Create another Kinesis Data Firehose delivery stream. Update the producer application to write a copy of the log data into the new delivery stream. Set the new delivery stream to deliver data into an Amazon QuickSight dashboard.

D. Update the producer application to write the log data to an Amazon Kinesis data stream. Deliver this data stream to the original Kinesis Data Firehose delivery stream and a new Kinesis Data Firehose delivery stream. Set the new delivery stream to deliver data into an Amazon OpenSearch Service (Amazon Elasticsearch Service) cluster. Use OpenSearch Dashboards (Kibana) for log data visualization.

Correct Answer: B

Reference: https://docs.aws.amazon.com/firehose/latest/dev/what-is-this-service.html

QUESTION 4

An online retail company is migrating its reporting system to AWS. The company\\'s legacy system runs data processing on online transactions using a complex series of nested Apache Hive queries. Transactional data is exported from the online system to the reporting system several times a day. Schemas in the files are stable between updates.

A data analyst wants to quickly migrate the data processing to AWS, so any code changes should be minimized. To keep storage costs low, the data analyst decides to store the data in Amazon S3. It is vital that the data from the reports and associated analytics is completely up to date based on the data in Amazon S3.

Which solution meets these requirements?

A. Create an AWS Glue Data Catalog to manage the Hive metadata. Create an AWS Glue crawler over Amazon S3 that runs when data is refreshed to ensure that data changes are updated. Create an Amazon EMR cluster and use the

metadata in the AWS Glue Data Catalog to run Hive processing queries in Amazon EMR.

B. Create an AWS Glue Data Catalog to manage the Hive metadata. Create an Amazon EMR cluster with consistent view enabled. Run emrfs sync before each analytics step to ensure data changes are updated. Create an EMR cluster and use the metadata in the AWS Glue Data Catalog to run Hive processing queries in Amazon EMR.

C. Create an Amazon Athena table with CREATE TABLE AS SELECT (CTAS) to ensure data is refreshed from underlying queries against the raw dataset. Create an AWS Glue Data Catalog to manage the Hive metadata over the CTAS table. Create an Amazon EMR cluster and use the metadata in the AWS Glue Data Catalog to run Hive processing queries in Amazon EMR.

D. Use an S3 Select query to ensure that the data is properly updated. Create an AWS Glue Data Catalog to manage the Hive metadata over the S3 Select table. Create an Amazon EMR cluster and use the metadata in the AWS Glue Data Catalog to run Hive processing queries in Amazon EMR.

Correct Answer: A

QUESTION 5

A gaming company is collecting cllckstream data into multiple Amazon Kinesis data streams. The company uses Amazon Kinesis Data Firehose delivery streams to store the data in JSON format in Amazon S3 Data scientists use Amazon Athena to query the most recent data and derive business insights. The company wants to reduce its Athena costs without having to recreate the data pipeline. The company prefers a solution that will require less management effort.

Which set of actions can the data scientists take immediately to reduce costs?

A. Change the Kinesis Data Firehose output format to Apache Parquet Provide a custom S3 object YYYYMMDD prefix expression and specify a large buffer size For the existing data, run an AWS Glue ETL job to combine and convert small JSON files to large Parquet files and add the YYYYMMDD prefix Use ALTER TABLE ADD PARTITION to reflect the partition on the existing Athena table.

B. Create an Apache Spark Job that combines and converts JSON files to Apache Parquet files Launch an Amazon EMR ephemeral cluster daily to run the Spark job to create new Parquet files in a different S3 location Use ALTER TABLE SET LOCATION to reflect the new S3 location on the existing Athena table.

C. Create a Kinesis data stream as a delivery target for Kinesis Data Firehose Run Apache Flink on Amazon Kinesis Data Analytics on the stream to read the streaming data, aggregate ikand save it to Amazon S3 in Apache Parquet format with a custom S3 object YYYYMMDD prefix Use ALTER TABLE ADD PARTITION to reflect the partition on the existing Athena table

D. Integrate an AWS Lambda function with Kinesis Data Firehose to convert source records to Apache Parquet and write them to Amazon S3 In parallel, run an AWS Glue ETL job to combine and convert existing JSON files to large Parquet

files Create a custom S3 object YYYYMMDD prefix Use ALTER TABLE ADD PARTITION to reflect the partition on the existing Athena table.

Correct Answer: D

Reference: https://docs.aws.amazon.com/firehose/latest/dev/record-format-conversion.html
https://docs.aws.amazon.com/firehose/latest/dev/data-analysis.html

DAS-C01 PDF Dumps DAS-C01 Practice Test DAS-C01 Exam Questions