

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.pass2lead.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 ingests streaming data. The company needs to analyze this stream over a 5-minute timeframe to evaluate the stream for anomalies with Random Cut Forest (RCF) and summarize the current count of status codes. The source and summarized data should be persisted for future use.

Which approach would enable the desired outcome while keeping data persistence costs low?

- A. Ingest the data stream with Amazon Kinesis Data Streams. Have an AWS Lambda consumer evaluate the stream, collect the number status codes, and evaluate the data against a previously trained RCF model. Persist the source and results as a time series to Amazon DynamoDB.
- B. Ingest the data stream with Amazon Kinesis Data Streams. Have a Kinesis Data Analytics application evaluate the stream over a 5-minute window using the RCF function and summarize the count of status codes. Persist the source and results to Amazon S3 through output delivery to Kinesis Data Firehouse.
- C. Ingest the data stream with Amazon Kinesis Data Firehose with a delivery frequency of 1 minute or 1 MB in Amazon S3. Ensure Amazon S3 triggers an event to invoke an AWS Lambda consumer that evaluates the batch data, collects the number status codes, and evaluates the data against a previously trained RCF model. Persist the source and results as a time series to Amazon DynamoDB.
- D. Ingest the data stream with Amazon Kinesis Data Firehose with a delivery frequency of 5 minutes or 1 MB into Amazon S3. Have a Kinesis Data Analytics application evaluate the stream over a 1-minute window using the RCF function and summarize the count of status codes. Persist the results to Amazon S3 through a Kinesis Data Analytics output to an AWS Lambda integration.

Correct Answer: B

QUESTION 2

A company uses Amazon Redshift as its data warehouse. The Redshift cluster is not encrypted. A data analytics specialist needs to use hardware security module (HSM) managed encryption keys to encrypt the data that is stored in the Redshift cluster.

Which combination of steps will meet these requirements? (Choose three.)

- A. Stop all write operations on the source cluster. Unload data from the source cluster.
- B. Copy the data to a new target cluster that is encrypted with AWS Key Management Service (AWS KMS).
- C. Modify the source cluster by activating AWS CloudHSM encryption. Configure Amazon Redshift to automatically migrate data to a new encrypted cluster.
- D. Modify the source cluster by activating encryption from an external HSM. Configure Amazon Redshift to automatically migrate data to a new encrypted cluster.
- E. Copy the data to a new target cluster that is encrypted with an HSM from AWS CloudHSM.
- F. Rename the source cluster and the target cluster after the migration so that the target cluster is using the original endpoint.

Correct Answer: AEF

QUESTION 3

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 4

A financial institution is building an Amazon QuickSight business intelligence (BI) dashboard to show financial performance and analyze trends. The development team is using an Amazon Redshift database in the development environment and is having difficulty with validating the accuracy of the metrics calculation algorithm due to the lack of quality data. The Redshift production environment database is 500 TB and is in a different AWS account in the same AWS Region as the development environment account. The company needs to use up-to-date production environment data for development purposes.

Which solution MOST cost-effectively meets these requirements?

- A. Setup data streaming with Amazon Kinesis Data Streams from the production environment Redshift database to replicate the data to the development environment Redshift database.
- B. Create a Redshift datashare to share the production environment data with the development team.
- C. Upload the data from Amazon Redshift to Amazon S3. Then load the data directly from Amazon S3 to the development environment Redshift cluster using the `COPY` command.
- D. Create Redshift views that are configured to share all the data between the production and development clusters.

Correct Answer: D

QUESTION 5

An online retail company with millions of users around the globe wants to improve its ecommerce analytics capabilities. Currently, clickstream data is uploaded directly to Amazon S3 as compressed files. Several times each day, an application running on Amazon EC2 processes the data and makes search options and reports available for visualization by editors and marketers. The company wants to make website clicks and aggregated data available to editors and marketers in minutes to enable them to connect with users more effectively.

Which options will help meet these requirements in the MOST efficient way? (Choose two.)

- A. Use Amazon Kinesis Data Firehose to upload compressed and batched clickstream records to Amazon OpenSearch Service (Amazon Elasticsearch Service).
- B. Upload clickstream records to Amazon S3 as compressed files. Then use AWS Lambda to send data to Amazon OpenSearch Service (Amazon Elasticsearch Service) from Amazon S3.
- C. Use Amazon OpenSearch Service (Amazon Elasticsearch Service) deployed on Amazon EC2 to aggregate, filter, and process the data. Refresh content performance dashboards in near-real time.
- D. Use OpenSearch Dashboards (Kibana) to aggregate, filter, and visualize the data stored in Amazon OpenSearch Service (Amazon Elasticsearch Service). Refresh content performance dashboards in near-real time.
- E. Upload clickstream records from Amazon S3 to Amazon Kinesis Data Streams and use a Kinesis Data Streams consumer to send records to Amazon OpenSearch Service (Amazon Elasticsearch Service).

Correct Answer: CE

[DAS-C01 PDF Dumps](#)

[DAS-C01 Exam Questions](#)

[DAS-C01 Braindumps](#)