



# DAS-C01<sup>Q&As</sup>

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 marketing company collects clickstream data. The company sends the data to Amazon Kinesis Data Firehose and stores the data in Amazon S3. The company wants to build a series of dashboards that will be used by hundreds of users across different departments. The company will use Amazon QuickSight to develop these dashboards. The company has limited resources and wants a solution that could scale and provide daily updates about clickstream activity.

Which combination of options will provide the MOST cost-effective solution? (Select TWO )

- A. Use Amazon Redshift to store and query the clickstream data
- B. Use QuickSight with a direct SQL query
- C. Use Amazon Athena to query the clickstream data in Amazon S3
- D. Use S3 analytics to query the clickstream data
- E. Use the QuickSight SPICE engine with a daily refresh

Correct Answer: BD

Reference: <https://aws.amazon.com/blogs/big-data/create-real-time-clickstream-sessions-and-run-analytics-with-amazon-kinesis-data-analytics-aws-glue-and-amazon-athena/>

---

### QUESTION 2

A company uses Amazon OpenSearch Service (Amazon Elasticsearch Service) to store and analyze its website clickstream data. The company ingests 1 TB of data daily using Amazon Kinesis Data Firehose and stores one day's worth of data in an Amazon OpenSearch Service (Amazon Elasticsearch Service) cluster.

The company has very slow query performance on the Amazon OpenSearch Service (Amazon Elasticsearch Service) index and occasionally sees errors from Kinesis Data Firehose when attempting to write to the index. The Amazon OpenSearch Service (Amazon Elasticsearch Service) cluster has 10 nodes running a single index and 3 dedicated master nodes. Each data node has 1.5 TB of Amazon EBS storage attached and the cluster is configured with 1,000 shards. Occasionally, JVMMemoryPressure errors are found in the cluster logs.

Which solution will improve the performance of Amazon OpenSearch Service (Amazon Elasticsearch Service)?

- A. Increase the memory of the Amazon OpenSearch Service (Amazon Elasticsearch Service) master nodes.
- B. Decrease the number of Amazon OpenSearch Service (Amazon Elasticsearch Service) data nodes.
- C. Decrease the number of Amazon OpenSearch Service (Amazon Elasticsearch Service) shards for the index.
- D. Increase the number of Amazon OpenSearch Service (Amazon Elasticsearch Service) shards for the index.

Correct Answer: C

Reference: <https://www.bluematador.com/docs/troubleshooting/aws-elasticsearch-jvm-pressure#:~:text=Amazon%20recommends%20keeping%20JVM%20pressure,getting%20into%20a%20red%20state>

---



### QUESTION 3

A company with a video streaming website wants to analyze user behavior to make recommendations to users in real time. Clickstream data is being sent to Amazon Kinesis Data Streams and reference data is stored in Amazon S3. The company wants a solution that can use standard SQL queries. The solution must also provide a way to look up pre-calculated reference data while making recommendations.

Which solution meets these requirements?

- A. Use an AWS Glue Python shell job to process incoming data from Kinesis Data Streams. Use the Boto3 library to write data to Amazon Redshift.
- B. Use AWS Glue streaming and Scale to process incoming data from Kinesis Data Streams. Use the AWS Glue connector to write data to Amazon Redshift.
- C. Use Amazon Kinesis Data Analytics to create an in-application table based upon the reference data. Process incoming data from Kinesis Data Streams. Use a data stream to write results to Amazon Redshift.
- D. Use Amazon Kinesis Data Analytics to create an in-application table based upon the reference data. Process incoming data from Kinesis Data Streams. Use an Amazon Kinesis Data Firehose delivery stream to write results to Amazon Redshift. Pass Your Certification With

Correct Answer: D

Reference: <https://aws.amazon.com/blogs/big-data/real-time-analytics-with-amazon-redshift-streaming-ingestion/>

---

### QUESTION 4

A company hosts an Apache Flink application on premises. The application processes data from several Apache Kafka clusters. The data originates from a variety of sources, such as web applications, mobile apps, and operational databases. The company has migrated some of these sources to AWS and now wants to migrate the Flink application. The company must ensure that data that resides in databases within the VPC does not traverse the internet. The application must be able to process all the data that comes from the company's AWS solution, on-premises resources, and the public internet.

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

- A. Implement Flink on Amazon EC2 within the company's VPC. Create Amazon Managed Streaming for Apache Kafka (Amazon MSK) clusters in the VPC to collect data that comes from applications and databases within the VPC. Use Amazon Kinesis Data Streams to collect data that comes from the public internet. Configure Flink to have sources from Kinesis Data Streams, Amazon MSK, and any on-premises Kafka clusters by using AWS Client VPN or AWS Direct Connect.
- B. Implement Flink on Amazon EC2 within the company's VPC. Use Amazon Kinesis Data Streams to collect data that comes from applications and databases within the VPC and the public internet. Configure Flink to have sources from Kinesis Data Streams and any on-premises Kafka clusters by using AWS Client VPN or AWS Direct Connect.
- C. Create an Amazon Kinesis Data Analytics application by uploading the compiled Flink .jar file. Use Amazon Kinesis Data Streams to collect data that comes from applications and databases within the VPC and the public internet. Configure the Kinesis Data Analytics application to have sources from Kinesis Data Streams and any on-premises Kafka clusters by using AWS Client VPN or AWS Direct Connect.
- D. Create an Amazon Kinesis Data Analytics application by uploading the compiled Flink .jar file. Create Amazon Managed Streaming for Apache Kafka (Amazon MSK) clusters in the company's VPC to collect data that comes from applications and databases within the VPC. Use Amazon Kinesis Data Streams to collect data that comes from the



public internet. Configure the Kinesis Data Analytics application to have sources from Kinesis Data Streams, Amazon MSK, and any on-premises Kafka clusters by using AWS Client VPN or AWS Direct Connect.

Correct Answer: C

Reference: <https://aws.amazon.com/blogs/big-data/streaming-etl-with-apache-flink-and-amazon-kinesis-data-analytics/>  
<https://docs.aws.amazon.com/streams/latest/dev/vpc.html>

---

## QUESTION 5

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

[DAS-C01 PDF Dumps](#)

[DAS-C01 VCE Dumps](#)

[DAS-C01 Study Guide](#)