



Free Questions for [DAS-C01](#) by [dumpshq](#)

Shared by [Colon](#) on [20-10-2022](#)

For More Free Questions and Preparation Resources

[Check the Links on Last Page](#)

Question 1

Question Type: MultipleChoice

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?

Options:

- 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

Answer:

D

Question 2

Question Type: MultipleChoice

A company stores Apache Parquet-formatted files in Amazon S3 The company uses an AWS Glue Data Catalog to store the table metadata and Amazon Athena to query and analyze the data The tables have a large number of partitions The queries are only run on small subsets of data in the table A data analyst adds new time partitions into the table as new data arrives The data analyst has been asked to reduce the query runtime

Which solution will provide the MOST reduction in the query runtime?

Options:

- A- Convert the Parquet files to the csv file format..Then attempt to query the data again
- B- Convert the Parquet files to the Apache ORC file format. Then attempt to query the data again
- C- Use partition projection to speed up the processing of the partitioned table
- D- Add more partitions to be used over the table. Then filter over two partitions and put all columns in the WHERE clause

Answer:

C

Question 3

Question Type: MultipleChoice

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 quenes The solution must also provide a way to look up pre-calculated reference data while making recommendations

Which solution meets these requirements?

Options:

- 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

Answer:

D

Question 4

Question Type: MultipleChoice

A company has several Amazon EC2 instances sitting behind an Application Load Balancer (ALB) The company wants its IT Infrastructure team to analyze the IP addresses coming into the company's ALB The ALB is configured to store access logs in Amazon S3 The access logs create about 1 TB of data each day, and access to the data will be infrequent The company needs a solution that is scalable, cost-effective and has minimal maintenance requirements

Which solution meets these requirements?

Options:

- A- Copy the data into Amazon Redshift and query the data
- B- Use Amazon EMR and Apache Hive to query the S3 data
- C- Use Amazon Athena to query the S3 data
- D- Use Amazon Redshift Spectrum to query the S3 data

Answer:

D

Question 5

Question Type: MultipleChoice

A bank wants to migrate a Teradata data warehouse to the AWS Cloud. The bank needs a solution for reading large amounts of data and requires the highest possible performance. The solution also must maintain the separation of storage and compute.

Which solution meets these requirements?

Options:

- A- Use Amazon Athena to query the data in Amazon S3
- B- Use Amazon Redshift with dense compute nodes to query the data in Amazon Redshift managed storage
- C- Use Amazon Redshift with RA3 nodes to query the data in Amazon Redshift managed storage
- D- Use PrestoDB on Amazon EMR to query the data in Amazon S3

Answer:

C

Question 6

Question Type: MultipleChoice

A company recently created a test AWS account to use for a development environment. The company also created a production AWS account in another AWS Region. As part of its security testing, the company wants to send log data from Amazon CloudWatch Logs in its production account to an Amazon Kinesis data stream in its test account.

Which solution will allow the company to accomplish this goal?

Options:

- A-** Create a subscription filter in the production accounts CloudWatch Logs to target the Kinesis data stream in the test account as its destination In the test account create an IAM role that grants access to the Kinesis data stream and the CloudWatch Logs resources in the production account
- B-** In the test account create an IAM role that grants access to the Kinesis data stream and the CloudWatch Logs resources in the production account Create a destination data stream in Kinesis Data Streams in the test account with an IAM role and a trust policy that allow CloudWatch Logs in the production account to write to the test account
- C-** In the test account, create an IAM role that grants access to the Kinesis data stream and the CloudWatch Logs resources in the production account Create a destination data stream in Kinesis Data Streams in the test account with an IAM role and a trust policy that allow CloudWatch Logs in the production account to write to the test account
- D-** Create a destination data stream in Kinesis Data Streams in the test account with an IAM role and a trust policy that allow CloudWatch Logs in the production account to write to the test account Create a subscription filter in the production accounts CloudWatch Logs to target the Kinesis data stream in the test account as its destination

Answer:

D

Question 7

Question Type: MultipleChoice

A company uses an Amazon EMR cluster with 50 nodes to process operational data and make the data available for data analysts. These jobs run nightly use Apache Hive with the Apache Tez framework as a processing model and write results to Hadoop Distributed File System (HDFS). In the last few weeks, jobs are failing and are producing the following error message:

"File could only be replicated to 0 nodes instead of 1"

A data analytics specialist checks the DataNode logs, the NameNode logs, and network connectivity for potential issues that could have prevented HDFS from replicating data. The data analytics specialist rules out these factors as causes for the issue.

Which solution will prevent the jobs from failing?

Options:

- A-** Monitor the HDFSUtilization metric. If the value crosses a user-defined threshold, add task nodes to the EMR cluster.
- B-** Monitor the HDFSUtilization metric. If the value crosses a user-defined threshold, add core nodes to the EMR cluster.
- C-** Monitor the MemoryAllocatedMB metric. If the value crosses a user-defined threshold, add task nodes to the EMR cluster.
- D-** Monitor the MemoryAllocatedMB metric. If the value crosses a user-defined threshold, add core nodes to the EMR cluster.

Answer:

C

Question 8

Question Type: MultipleChoice

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)

Options:

- 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

Answer:

B, D

Question 9

Question Type: MultipleChoice

A large energy company is using Amazon QuickSight to build dashboards and report the historical usage data of its customers. This data is hosted in Amazon Redshift. The reports need access to all the fact tables' billions of records to create aggregation in real time grouping by multiple dimensions.

A data analyst created the dataset in QuickSight by using a SQL query and not SPICE. Business users have noted that the response time is not fast enough to meet their needs.

Which action would speed up the response time for the reports with the LEAST implementation effort?

Options:

- A- Use QuickSight to modify the current dataset to use SPICE
- B- Use AWS Glue to create an Apache Spark job that joins the fact table with the dimensions. Load the data into a new table
- C- Use Amazon Redshift to create a materialized view that joins the fact table with the dimensions
- D- Use Amazon Redshift to create a stored procedure that joins the fact table with the dimensions. Load the data into a new table

Answer:

A

To Get Premium Files for DAS-C01 Visit

<https://www.p2pexams.com/products/das-c01>

For More Free Questions Visit

<https://www.p2pexams.com/amazon/pdf/das-c01>

