Free Questions for DAS-C01 by certsinside

Shared by Fields on 29-01-2024

For More Free Questions and Preparation Resources

Question 1

A company uses Amazon Connect to manage its contact center. The company uses Salesforce to manage its customer relationship management (CRM) dat

a. The company must build a pipeline to ingest data from Amazon Connect and Salesforce into a data lake that is built on Amazon S3.

Which solution will meet this requirement with the LEAST operational overhead?

Options:

A- Use Amazon Kinesis Data Streams to ingest the Amazon Connect data. Use Amazon AppFlow to ingest the Salesforce data.

B- Use Amazon Kinesis Data Firehose to ingest the Amazon Connect data. Use Amazon Kinesis Data Streams to ingest the Salesforce data.

C- Use Amazon Kinesis Data Firehose to ingest the Amazon Connect data. Use Amazon AppFlow to ingest the Salesforce data.

D- Use Amazon AppFlow to ingest the Amazon Connect data. Use Amazon Kinesis Data Firehose to ingest the Salesforce data.

Answer:

B

Question 2

A company uses an Amazon Redshift provisioned cluster for data analysis. The data is not encrypted at rest. A data analytics specialist must implement a solution to encrypt the data at rest.

Which solution will meet this requirement with the LEAST operational overhead?

Options:

A- Use the ALTER TABLE command with the ENCODE option to update existing columns of the Redshift tables to use LZO encoding.

B- Export data from the existing Redshift cluster to Amazon S3 by using the UNLOAD command with the ENCRYPTED option. Create a new Redshift cluster with encryption configured. Load data into the new cluster by using the COPY command.

C- Create a manual snapshot of the existing Redshift cluster. Restore the snapshot into a new Redshift cluster with encryption configured.

D- Modify the existing Redshift cluster to use AWS Key Management Service (AWS KMS) encryption. Wait for the cluster to finish resizing.

Answer:

D

Question 3

A company wants to ingest clickstream data from its website into an Amazon S3 bucket. The streaming data is in JSON format. The data in the S3 bucket must be partitioned by product_id.

Which solution will meet these requirements MOST cost-effectively?

Options:

A- Create an Amazon Kinesis Data Firehose delivery stream to ingest the streaming data into the S3 bucket. Enable dynamic partitioning. Specify the data field of productjd as one partitioning key.

B- Create an AWS Glue streaming job to partition the data by productjd before delivering the data to the S3 bucket. Create an Amazon Kinesis Data Firehose delivery stream. Specify the AWS Glue job as the destination of the delivery stream.

C- Create an Amazon Kinesis Data Firehose delivery stream to ingest the streaming data into the S3 bucket. Create an AWS Glue ETL job to read the data stream in the S3 bucket, partition the data by productjd, and write the data into another S3 bucket.

D- Create an Amazon Kinesis Data Firehose delivery stream to ingest the streaming data into the S3 bucket. Create an Amazon EMR cluster that includes a job to read the data stream in the S3 bucket, partition the data by productjd, and write the data into another S3 bucket.

Answer:

Question 4

A network administrator needs to create a dashboard to visualize continuous network patterns over time in a company's AWS account. Currently, the company has VPC Flow Logs enabled and is publishing this data to Amazon CloudWatch Logs. To troubleshoot networking issues quickly, the dashboard needs to display the new data in near-real time.

Which solution meets these requirements?

Options:

A- Create a CloudWatch Logs subscription to stream CloudWatch Logs data to an AWS Lambda function that writes the data to an Amazon S3 bucket. Create an Amazon QuickSight dashboard to visualize the data.

B- Create an export task from CloudWatch Logs to an Amazon S3 bucket. Create an Amazon QuickSight dashboard to visualize the data.

C- Create a CloudWatch Logs subscription that uses an AWS Lambda function to stream the CloudWatch Logs data directly into an Amazon OpenSearch Service cluster. Use OpenSearch Dashboards to create the dashboard.

D- Create a CloudWatch Logs subscription to stream CloudWatch Logs data to an AWS Lambda function that writes to an Amazon Kinesis data stream to deliver the data into an Amazon OpenSearch Service cluster. Use OpenSearch Dashboards to create the

dashboard.

Question 5

Question Type: **MultipleChoice**

A company ingests a large set of sensor data in nested JSON format from different sources and stores it in an Amazon S3 bucket. The sensor data must be joined with performance data currently stored in an Amazon Redshift cluster.

A business analyst with basic SQL skills must build dashboards and analyze this data in Amazon QuickSight. A data engineer needs to build a solution to prepare the data for use by the business analyst. The data engineer does not know the structure of the JSON file. The company requires a solution with the least possible implementation effort.

Which combination of steps will create a solution that meets these requirements? (Select THREE.)

Options:

A- Use an AWS Glue ETL job to convert the data into Apache Parquet format and write to Amazon S3.

B- Use an AWS Glue crawler to catalog the data.

C- Use an AWS Glue ETL job with the ApplyMapping class to un-nest the data and write to Amazon Redshift tables.

D- Use an AWS Glue ETL job with the Regionalize class to un-nest the data and write to Amazon Redshift tables.

E- Use QuickSight to create an Amazon Athena data source to read the Apache Parquet files in Amazon S3.

F- Use QuickSight to create an Amazon Redshift data source to read the native Amazon Redshift tables.

Answer:

B, D, F

Question 6

Question Type: MultipleChoice

A financial services firm is processing a stream of real-time data from an application by using Apache Kafka and Kafka MirrorMaker. These tools run on premises and stream data to Amazon Managed Streaming for Apache Kafka (Amazon MSK) in the us-east-1 Region. An Apache Flink consumer running on Amazon EMR enriches the data in real time and transfers the output files to an Amazon S3 bucket. The company wants to ensure that the streaming application is highly available across AWS Regions with an RTO of less than 2 minutes.

Which solution meets these requirements?

A- Launch another Amazon MSK and Apache Flink cluster in the us-west-1 Region that is the same size as the original cluster in the us-east-1 Region. Simultaneously publish and process the data in both Regions. In the event of a disaster that impacts one of the Regions, switch to the other Region.

B- Set up Cross-Region Replication from the Amazon S3 bucket in the us-east-1 Region to the us-west-1 Region. In the event of a disaster, immediately create Amazon MSK and Apache Flink clusters in the us-west-1 Region and start publishing data to this Region.

C- Add an AWS Lambda function in the us-east-1 Region to read from Amazon MSK and write to a global Amazon

D- DynamoDB table in on-demand capacity mode. Export the data from DynamoDB to Amazon S3 in the us-west-1 Region. In the event of a disaster that impacts the us-east-1 Region, immediately create Amazon MSK and Apache Flink clusters in the us-west-1 Region and start publishing data to this Region.

E- Set up Cross-Region Replication from the Amazon S3 bucket in the us-east-1 Region to the us-west-1 Region. In the event of a disaster, immediately create Amazon MSK and Apache Flink clusters in the us-west-1 Region and start publishing data to this Region. Store 7 days of data in on-premises Kafka clusters and recover the data missed during the recovery time from the on-premises cluster.

Answer:

A

Question 7

Question Type: **MultipleChoice**

A company wants to use a data lake that is hosted on Amazon S3 to provide analytics services for historical dat

a. The data lake consists of 800 tables but is expected to grow to thousands of tables. More than 50 departments use the tables, and each department has hundreds of users. Different departments need access to specific tables and columns.

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

Options:

A- Create an 1AM role for each department. Use AWS Lake Formation based access control to grant each 1AM role access to specific tables and columns. Use Amazon Athena to analyze the data.

B- Create an Amazon Redshift cluster for each department. Use AWS Glue to ingest into the Redshift cluster only the tables and columns that are relevant to that department. Create Redshift database users. Grant the users access to the relevant department's Redshift cluster. Use Amazon Redshift to analyze the data.

C- Create an 1AM role for each department. Use AWS Lake Formation tag-based access control to grant each 1AM role access to only the relevant resources. Create LF-tags that are attached to tables and columns. Use Amazon Athena to analyze the data.

D- Create an Amazon EMR cluster for each department. Configure an 1AM service role for each EMR cluster to access

E- relevant S3 files. For each department's users, create an 1AM role that provides access to the relevant EMR cluster. Use Amazon EMR to analyze the data.

Answer:
C

Question 8

A company uses Amazon EC2 instances to receive files from external vendors throughout each day. At the end of each day, the EC2 instances combine the files into a single file, perform gzip compression, and upload the single file to an Amazon S3 bucket. The total size of all the files is approximately 100 GB each day.

When the files are uploaded to Amazon S3, an AWS Batch job runs a COPY command to load the files into an Amazon Redshift cluster.

Which solution will MOST accelerate the COPY process?

Options:

A- Upload the individual files to Amazon S3. Run the COPY command as soon as the files become available.

B- Split the files so that the number of files is equal to a multiple of the number of slices in the Redshift cluster. Compress and upload the files to Amazon S3. Run the COPY command on the files.

C- Split the files so that each file uses 50% of the free storage on each compute node in the Redshift cluster. Compress and upload the files to Amazon S3. Run the COPY command on the files.

D- pply sharding by breaking up the files so that the DISTKEY columns with the same values go to the same file. Compress and upload the sharded files to Amazon S3. Run the COPY command on the files.

Answer:

B

Question 9

Question Type: MultipleChoice

A company wants to use automatic machine learning (ML) to create and visualize forecasts of complex scenarios and trends.

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

Options:

A- Use an AWS Glue ML job to transform the data and create forecasts. Use Amazon QuickSight to visualize the data.

B- Use Amazon QuickSight to visualize the data. Use ML-powered forecasting in QuickSight to create forecasts.

C- Use a prebuilt ML AMI from the AWS Marketplace to create forecasts. Use Amazon QuickSight to visualize the data.

D- Use Amazon SageMaker inference pipelines to create and update forecasts. Use Amazon QuickSight to visualize the combined data.

Answer:

B

Question 10

A marketing company has an application that stores event data in an Amazon RDS database. The company is replicating this data to Amazon Redshift for reporting and

business intelligence (BI) purposes. New event data is continuously generated and ingested into the RDS database throughout the day and captured by a change data

capture (CDC) replication task in AWS Database Migration Service (AWS DMS). The company requires that the new data be replicated to Amazon Redshift in near-real

time.

Which solution meets these requirements?

Options:

A- Use Amazon Kinesis Data Streams as the destination of the CDC replication task in AWS DMS. Use an AWS Glue streaming job to read changed records from Kinesis Data Streams and perform an upsert into the Redshift cluster.

B- Use Amazon S3 as the destination of the CDC replication task in AWS DMS. Use the COPY command to load data into the Redshift cluster.

C- Use Amazon DynamoDB as the destination of the CDC replication task in AWS DMS. Use the COPY command to load data into the Redshift cluster.

D- Use Amazon Kinesis Data Firehose as the destination of the CDC replication task in AWS DMS. Use an AWS Glue streaming job to read changed records from Kinesis Data Firehose and perform an upsert into the Redshift cluster.

Answer:

A