DBCLOUD
EXAMS

QUESTIONS & ANSWERS

verified



Technology Provider: Amazon
Exam: AWS Certified Data Analytics - Specialty (DAS-C01)
Number of questions in the database: 164
Exam Version: NOV.2026

Topic 1 - Single Topic

Question #1*Topic 1*

A financial services company needs to aggregate daily stock trade data from the exchanges into a data store. The company requires that data be streamed directly into the data store, but also occasionally allows data to be modified using SQL. The solution should integrate complex, analytic queries running with minimal latency. The solution must provide a business intelligence dashboard that enables viewing of the top contributors to anomalies in stock prices.
Which solution meets the company's requirements?

- A. Use Amazon Kinesis Data Firehose to stream data to Amazon S3. Use Amazon Athena as a data source for Amazon QuickSight to create a business intelligence dashboard.
- B. Use Amazon Kinesis Data Streams to stream data to Amazon Redshift. Use Amazon Redshift as a data source for Amazon QuickSight to create a business intelligence dashboard.
- C. Use Amazon Kinesis Data Firehose to stream data to Amazon Redshift. Use Amazon Redshift as a data source for Amazon QuickSight to create a business intelligence dashboard.
- D. Use Amazon Kinesis Data Streams to stream data to Amazon S3. Use Amazon Athena as a data source for Amazon QuickSight to create a business intelligence dashboard.

Correct Answer: *C*

Respuesta por mayor votación de la comunidad

C (83%)

B (17%)

Question #2*Topic 1*

A financial company hosts a data lake in Amazon S3 and a data warehouse on an Amazon Redshift cluster. The company uses Amazon QuickSight to build dashboards and wants to secure access from its on-premises Active Directory to Amazon QuickSight.
How should the data be secured?

- A. Use an Active Directory connector and single sign-on (SSO) in a corporate network environment.

https://dbcloud.academy/

- B. Use a VPC endpoint to connect to Amazon S3 from Amazon QuickSight and an IAM role to authenticate Amazon Redshift.
- C. Establish a secure connection by creating an S3 endpoint to connect Amazon QuickSight and a VPC endpoint to connect to Amazon Redshift.
- D. Place Amazon QuickSight and Amazon Redshift in the security group and use an Amazon S3 endpoint to connect Amazon QuickSight to Amazon S3.

Correct Answer: *A*

Respuesta por mayor votación de la comunidad

A (98%)

2%

Question #3*Topic 1*

A real estate company has a mission-critical application using Apache HBase in Amazon EMR. Amazon EMR is configured with a single master node. The company has over 5 TB of data stored on an Hadoop Distributed File System (HDFS). The company wants a cost-effective solution to make its HBase data highly available.
Which architectural pattern meets company's requirements?

- A. Use Spot Instances for core and task nodes and a Reserved Instance for the EMR master node. Configure the EMR cluster with multiple master nodes. Schedule automated snapshots using Amazon EventBridge.
- B. Store the data on an EMR File System (EMRFS) instead of HDFS. Enable EMRFS consistent view. Create an EMR HBase cluster with multiple master nodes. Point the HBase root directory to an Amazon S3 bucket.
- C. Store the data on an EMR File System (EMRFS) instead of HDFS and enable EMRFS consistent view. Run two separate EMR clusters in two different Availability Zones. Point both clusters to the same HBase root directory in the same Amazon S3 bucket.
- D. Store the data on an EMR File System (EMRFS) instead of HDFS and enable EMRFS consistent view. Create a primary EMR HBase cluster with multiple master nodes. Create a secondary EMR HBase read-replica cluster in a separate Availability Zone. Point both clusters to the same HBase root directory in the same Amazon S3 bucket.

Correct Answer: *D*
Reference:
https://docs.aws.amazon.com/emr/latest/ReleaseGuide/emr-hbase-s3.html

Respuesta por mayor votación de la comunidad

D (51%)

B (49%)

https://dbcloud.academy/

Question #4*Topic 1*

A software company hosts an application on AWS, and new features are released weekly. As part of the application testing process, a solution must be developed that analyzes logs from each Amazon EC2 instance to ensure that the application is working as expected after each deployment. The collection and analysis solution should be highly available with the ability to display new information with minimal delays.
Which method should the company use to collect and analyze the logs?

- A. Enable detailed monitoring on Amazon EC2, use Amazon CloudWatch agent to store logs in Amazon S3, and use Amazon Athena for fast, interactive log analytics.
- B. Use the Amazon Kinesis Producer Library (KPL) agent on Amazon EC2 to collect and send data to Kinesis Data Streams to further push the data to Amazon OpenSearch Service (Amazon Elasticsearch Service) and visualize using Amazon QuickSight.
- C. Use the Amazon Kinesis Producer Library (KPL) agent on Amazon EC2 to collect and send data to Kinesis Data Firehose to further push the data to Amazon OpenSearch Service (Amazon Elasticsearch Service) and OpenSearch Dashboards (Kibana).
- D. Use Amazon CloudWatch subscriptions to get access to a real-time feed of logs and have the logs delivered to Amazon Kinesis Data Streams to further push the data to Amazon OpenSearch Service (Amazon Elasticsearch Service) and OpenSearch Dashboards (Kibana).

Correct Answer: *C*
Reference:
https://docs.aws.amazon.com/AmazonCloudWatch/latest/logs/Subscriptions.html

Respuesta por mayor votación de la comunidad

C (72%)

D (24%)

4%

Question #5*Topic 1*

A data analyst is using AWS Glue to organize, cleanse, validate, and format a 200 GB dataset. The data analyst triggered the job to run with the Standard worker type. After 3 hours, the AWS Glue job status is still RUNNING. Logs from the job run show no error codes. The data analyst wants to improve the job execution time without overprovisioning.
Which actions should the data analyst take?

- A. Enable job bookmarks in AWS Glue to estimate the number of data processing units (DPUs). Based on the profiled metrics, increase the value of the executor- cores job parameter.
- B. Enable job metrics in AWS Glue to estimate the number of data processing units (DPUs). Based on the profiled metrics, increase the value of the maximum capacity job parameter.

https://dbcloud.academy/

- C. Enable job metrics in AWS Glue to estimate the number of data processing units (DPUs). Based on the profiled metrics, increase the value of the spark.yarn.executor.memoryOverhead job parameter.
- D. Enable job bookmarks in AWS Glue to estimate the number of data processing units (DPUs). Based on the profiled metrics, increase the value of the num- executors job parameter.

Correct Answer: *B*
Reference:
https://docs.aws.amazon.com/glue/latest/dg/monitor-debug-capacity.html

Respuesta por mayor votación de la comunidad

B (100%)

Question #6*Topic 1*

A company has a business unit uploading .csv files to an Amazon S3 bucket. The company's data platform team has set up an AWS Glue crawler to do discovery, and create tables and schemas. An AWS Glue job writes processed data from the created tables to an Amazon Redshift database. The AWS Glue job handles column mapping and creating the Amazon Redshift table appropriately. When the AWS Glue job is rerun for any reason in a day, duplicate records are introduced into the Amazon Redshift table.
Which solution will update the Redshift table without duplicates when jobs are rerun?

- A. Modify the AWS Glue job to copy the rows into a staging table. Add SQL commands to replace the existing rows in the main table as postactions in the DynamicFrameWriter class.
- B. Load the previously inserted data into a MySQL database in the AWS Glue job. Perform an upsert operation in MySQL, and copy the results to the Amazon Redshift table.
- C. Use Apache Spark's DataFrame dropDuplicates() API to eliminate duplicates and then write the data to Amazon Redshift.
- D. Use the AWS Glue ResolveChoice built-in transform to select the most recent value of the column.

Correct Answer: *A*

Respuesta por mayor votación de la comunidad

A (100%)

Question #7*Topic 1*

A streaming application is reading data from Amazon Kinesis Data Streams and immediately writing the data to an Amazon S3 bucket every 10 seconds. The application is reading data from hundreds of shards. The batch interval cannot be changed due to a separate requirement. The data is being accessed by Amazon

https://dbcloud.academy/