

Reliable Data-Engineer-Associate Useful Dumps Offers Candidates 100% Pass-Rate Actual Amazon AWS Certified Data Engineer - Associate (DEA-C01) Exam Products



What's more, part of that TestSimulate Data-Engineer-Associate dumps now are free: https://drive.google.com/open?id=1wT_KHAcumSFCvEYo9GkJIg05wE5ygF0r

Our Data-Engineer-Associate practice materials are suitable for a variety of levels of users, no matter you are in a kind of cultural level, even if you only have high cultural level, you can find in our Data-Engineer-Associate study materials suitable for their own learning methods. So, for every user of our study materials are a great opportunity, a variety of types to choose from, more and more students also choose our Data-Engineer-Associate Study Materials, then why are you hesitating?

As is known to us, different people different understanding of learning, and also use different methods in different periods, and different learning activities suit different people, at different times of the day. Our Data-Engineer-Associate test questions are carefully designed by a lot of experts and professors in order to meet the needs of all customers. We can promise that our Data-Engineer-Associate Exam Question will be suitable for all people, including student, housewife, and worker and so on. No matter who you are, you must find that our Data-Engineer-Associate guide torrent will help you pass the Data-Engineer-Associate exam easily.

>> Data-Engineer-Associate Useful Dumps <<

New Data-Engineer-Associate Exam Objectives, Data-Engineer-Associate Certification

One of the best features of TestSimulate exam questions is free updates for up to 1 year. The TestSimulate has hired a team of experienced and qualified Amazon Data-Engineer-Associate exam trainers. They update the Data-Engineer-Associate exam questions as per the latest Data-Engineer-Associate Exam Syllabus. So rest assured that with the TestSimulate you will get the updated Data-Engineer-Associate exam practice questions all the time. Try a free demo if you to evaluate the features of our product. Best of luck!

Amazon AWS Certified Data Engineer - Associate (DEA-C01) Sample Questions (Q121-Q126):

NEW QUESTION # 121

A retail company uses an Amazon Redshift data warehouse and an Amazon S3 bucket. The company ingests retail order data into the S3 bucket every day.

The company stores all order data at a single path within the S3 bucket. The data has more than 100 columns.

The company ingests the order data from a third-party application that generates more than 30 files in CSV format every day. Each CSV file is between 50 and 70 MB in size.

The company uses Amazon Redshift Spectrum to run queries that select sets of columns. Users aggregate metrics based on daily orders. Recently, users have reported that the performance of the queries has degraded.

A data engineer must resolve the performance issues for the queries.

Which combination of steps will meet this requirement with LEAST developmental effort? (Select TWO.)

- A. Partition the order data in the S3 bucket based on order date.
- B. Configure the third-party application to create the files in a columnar format.
- C. Load the JSON data into the Amazon Redshift table in a SUPER type column.
- D. Develop an AWS Glue ETL job to convert the multiple daily CSV files to one file for each day.
- E. Configure the third-party application to create the files in JSON format.

Answer: A,B

Explanation:

The performance issue in Amazon Redshift Spectrum queries arises due to the nature of CSV files, which are row-based storage formats. Spectrum is more optimized for columnar formats, which significantly improve performance by reducing the amount of data scanned. Also, partitioning data based on relevant columns like order date can further reduce the amount of data scanned, as queries can focus only on the necessary partitions.

* A. Configure the third-party application to create the files in a columnar format:

* Columnar formats (like Parquet or ORC) store data in a way that is optimized for analytical queries because they allow queries to scan only the columns required, rather than scanning all columns in a row-based format like CSV.

* Amazon Redshift Spectrum works much more efficiently with columnar formats, reducing the amount of data that needs to be scanned, which improves query performance.

Reference: Amazon Redshift Spectrum and Columnar File Formats

C: Partition the order data in the S3 bucket based on order date:

Partitioning the data on columns like order date allows Redshift Spectrum to skip scanning unnecessary partitions, leading to improved query performance.

By organizing data into partitions, you minimize the number of files Spectrum has to read, further optimizing performance.

Reference: Best Practices for Amazon Redshift Spectrum Performance

Alternatives Considered:

B (Develop an AWS Glue ETL job): While consolidating files can improve performance by reducing the number of small files (which can be inefficient to process), it adds additional ETL complexity. Switching to a columnar format (Option A) and partitioning (Option C) provides more significant performance improvements with less development effort.

D and E (JSON-related options): Using JSON format or the SUPER type in Redshift introduces complexity and isn't as efficient as the proposed solutions, especially since JSON is not a columnar format.

References:

Amazon Redshift Spectrum Documentation

Columnar Formats and Data Partitioning in S3

NEW QUESTION # 122

A data engineer must manage the ingestion of real-time streaming data into AWS. The data engineer wants to perform real-time analytics on the incoming streaming data by using time-based aggregations over a window of up to 30 minutes. The data engineer needs a solution that is highly fault tolerant.

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

- A. Use an AWS Lambda function that includes both the business and the analytics logic to perform time-based aggregations over a window of up to 30 minutes for the data in Amazon Kinesis Data Streams.
- B. Use Amazon Managed Service for Apache Flink (previously known as Amazon Kinesis Data Analytics) to analyze the data by using multiple types of aggregations to perform time-based analytics over a window of up to 30 minutes.
- C. Use Amazon Managed Service for Apache Flink (previously known as Amazon Kinesis Data Analytics) to analyze the

data that might occasionally contain duplicates by using multiple types of aggregations.

- D. Use an AWS Lambda function that includes both the business and the analytics logic to perform aggregations for a tumbling window of up to 30 minutes, based on the event timestamp.

Answer: A

Explanation:

This solution meets the requirements of managing the ingestion of real-time streaming data into AWS and performing real-time analytics on the incoming streaming data with the least operational overhead. Amazon Managed Service for Apache Flink is a fully managed service that allows you to run Apache Flink applications without having to manage any infrastructure or clusters. Apache Flink is a framework for stateful stream processing that supports various types of aggregations, such as tumbling, sliding, and session windows, over streaming data. By using Amazon Managed Service for Apache Flink, you can easily connect to Amazon Kinesis Data Streams as the source and sink of your streaming data, and perform time-based analytics over a window of up to 30 minutes. This solution is also highly fault tolerant, as Amazon Managed Service for Apache Flink automatically scales, monitors, and restarts your Flink applications in case of failures.

References:

Amazon Managed Service for Apache Flink

Apache Flink

Window Aggregations in Flink

NEW QUESTION # 123

A data engineer needs to build an extract, transform, and load (ETL) job. The ETL job will process daily incoming .csv files that users upload to an Amazon S3 bucket. The size of each S3 object is less than 100 MB.

Which solution will meet these requirements MOST cost-effectively?

- A. Write a PySpark ETL script. Host the script on an Amazon EMR cluster.
- B. Write a custom Python application. Host the application on an Amazon Elastic Kubernetes Service (Amazon EKS) cluster.
- C. Write an AWS Glue PySpark job. Use Apache Spark to transform the data.
- D. Write an AWS Glue Python shell job. Use pandas to transform the data.

Answer: D

Explanation:

AWS Glue is a fully managed serverless ETL service that can handle various data sources and formats, including .csv files in Amazon S3. AWS Glue provides two types of jobs: PySpark and Python shell. PySpark jobs use Apache Spark to process large-scale data in parallel, while Python shell jobs use Python scripts to process small-scale data in a single execution environment. For this requirement, a Python shell job is more suitable and cost-effective, as the size of each S3 object is less than 100 MB, which does not require distributed processing. A Python shell job can use pandas, a popular Python library for data analysis, to transform the .csv data as needed. The other solutions are not optimal or relevant for this requirement. Writing a custom Python application and hosting it on an Amazon EKS cluster would require more effort and resources to set up and manage the Kubernetes environment, as well as to handle the data ingestion and transformation logic. Writing a PySpark ETL script and hosting it on an Amazon EMR cluster would also incur more costs and complexity to provision and configure the EMR cluster, as well as to use Apache Spark for processing small data files. Writing an AWS Glue PySpark job would also be less efficient and economical than a Python shell job, as it would involve unnecessary overhead and charges for using Apache Spark for small data files. Reference:

AWS Glue

Working with Python Shell Jobs

pandas

[AWS Certified Data Engineer - Associate DEA-C01 Complete Study Guide]

NEW QUESTION # 124

The company stores a large volume of customer records in Amazon S3. To comply with regulations, the company must be able to access new customer records immediately for the first 30 days after the records are created. The company accesses records that are older than 30 days infrequently.

The company needs to cost-optimize its Amazon S3 storage.

Which solution will meet these requirements MOST cost-effectively?

- A. Transition records to S3 Glacier Deep Archive storage after 30 days.
- B. Use S3 Intelligent-Tiering storage.
- C. Apply a lifecycle policy to transition records to S3 Standard Infrequent-Access (S3 Standard-IA) storage after 30 days.

- D. Use S3 Standard-Infrequent Access (S3 Standard-IA) storage for all customer records.

Answer: C

Explanation:

The most cost-effective solution in this case is to apply a lifecycle policy to transition records to Amazon S3 Standard-IA storage after 30 days. Here's why:

- * **Amazon S3 Lifecycle Policies:** Amazon S3 offers lifecycle policies that allow you to automatically transition objects between different storage classes to optimize costs. For data that is frequently accessed in the first 30 days and infrequently accessed after that, transitioning from the S3 Standard storage class to S3 Standard-Infrequent Access (S3 Standard-IA) after 30 days makes the most sense. S3 Standard-IA is designed for data that is accessed less frequently but still needs to be retained, offering lower storage costs than S3 Standard with a retrieval cost for access.
- * **Cost Optimization:** S3 Standard-IA offers a lower price per GB than S3 Standard. Since the data will be accessed infrequently after 30 days, using S3 Standard-IA will lower storage costs while still allowing for immediate retrieval when necessary.
- * **Compliance with Regulations:** Since the records need to be immediately accessible for the first 30 days, the use of S3 Standard for that period ensures compliance with regulatory requirements. After 30 days, transitioning to S3 Standard-IA continues to meet access requirements for infrequent access while reducing storage costs.
- * **Alternatives Considered:**
 - * **Option B (S3 Intelligent-Tiering):** While S3 Intelligent-Tiering automatically moves data between access tiers based on access patterns, it incurs a small monthly monitoring and automation charge per object. It could be a viable option, but transitioning data to S3 Standard-IA directly would be more cost-effective since the pattern of access is well-known (frequent for 30 days, infrequent thereafter).
 - * **Option C (S3 Glacier Deep Archive):** Glacier Deep Archive is the lowest-cost storage class, but it is not suitable in this case because the data needs to be accessed immediately within 30 days and on an infrequent basis thereafter. Glacier Deep Archive requires hours for data retrieval, which is not acceptable for infrequent access needs.
 - * **Option D (S3 Standard-IA for all records):** Using S3 Standard-IA for all records would result in higher costs for the first 30 days, as the data is frequently accessed. S3 Standard-IA incurs retrieval charges, making it less suitable for frequently accessed data.

References:

- * Amazon S3 Lifecycle Policies
- * S3 Storage Classes
- * Cost Management and Data Optimization Using Lifecycle Policies
- * AWS Data Engineering Documentation

NEW QUESTION # 125

A data engineer wants to orchestrate a set of extract, transform, and load (ETL) jobs that run on AWS. The ETL jobs contain tasks that must run Apache Spark jobs on Amazon EMR, make API calls to Salesforce, and load data into Amazon Redshift. The ETL jobs need to handle failures and retries automatically. The data engineer needs to use Python to orchestrate the jobs. Which service will meet these requirements?

- A. AWS Glue
- B. Amazon EventBridge
- C. AWS Step Functions
- **D. Amazon Managed Workflows for Apache Airflow (Amazon MWAA)**

Answer: D

Explanation:

The data engineer needs to orchestrate ETL jobs that include Spark jobs on Amazon EMR, API calls to Salesforce, and loading data into Redshift. They also need automatic failure handling and retries. Amazon Managed Workflows for Apache Airflow (Amazon MWAA) is the best solution for this requirement.

- * **Option A: Amazon Managed Workflows for Apache Airflow (Amazon MWAA)** Apache Airflow is designed for complex job orchestration, allowing users to define workflows (DAGs) in Python. MWAA manages Airflow and its integrations with other AWS services, including Amazon EMR, Redshift, and external APIs like Salesforce. It provides automatic retry handling, failure detection, and detailed monitoring, which fits the use case perfectly.
- * **Option B (AWS Step Functions)** can orchestrate tasks but doesn't natively support complex workflow definitions with Python like Airflow does.
- * **Option C (AWS Glue)** is more focused on ETL and doesn't handle the orchestration of external systems like Salesforce as well as Airflow.
- * **Option D (Amazon EventBridge)** is more suited for event-driven architectures rather than complex workflow orchestration.

References:

- * Amazon Managed Workflows for Apache Airflow
- * Apache Airflow on AWS

NEW QUESTION # 126

.....

The software keeps track of the previous AWS Certified Data Engineer - Associate (DEA-C01) (Data-Engineer-Associate) practice exam attempts and shows the changes of each attempt. You don't need to wait days or weeks to get your performance report. The software displays the result of the Amazon Data-Engineer-Associate Practice Test immediately, which is an excellent way to understand which area needs more attention.

New Data-Engineer-Associate Exam Objectives: <https://www.testsimulate.com/Data-Engineer-Associate-study-materials.html>

Amazon Data-Engineer-Associate Useful Dumps We strive towards continuous improvement of our products and service, Luckily for all the Amazon Data-Engineer-Associate experts, TestSimulate is now here to help you with your Amazon IT certification problems, as we are the best AWS Certified Data Engineer exam questions training material providing vendor, Through the good reputation of word of mouth, more and more people choose to use Data-Engineer-Associate study materials to prepare for the Data-Engineer-Associate exam, which makes us very gratified.

Perhaps this desire to help others is best expressed by Ceballos, as he tells Data-Engineer-Associate how he came to join the faculty: One day I received a call from the technology coordinator offering me a full time position as IT Instructor.

Pass Guaranteed Quiz 2025 Authoritative Amazon Data-Engineer-Associate Useful Dumps

In the pursuit of a good work/life balance, Michael spends his spare time Data-Engineer-Associate Pass4sure Study Materials diving, riding motorcycles, and enjoying the outdoors when possible, We strive towards continuous improvement of our products and service.

Luckily for all the Amazon Data-Engineer-Associate experts, TestSimulate is now here to help you with your Amazon IT certification problems, as we are the best AWS Certified Data Engineer exam questions training material providing vendor.

Through the good reputation of word of mouth, more and more people choose to use Data-Engineer-Associate study materials to prepare for the Data-Engineer-Associate exam, which makes us very gratified.

Life is too short to wake up in the morning with Data-Engineer-Associate Pass4sure Study Materials regrets, Not only will you be able to pass any Amazon AWS Certified Data Engineer test, but it gets better!

- Data-Engineer-Associate Training Pdf ☐ Valid Test Data-Engineer-Associate Vce Free ☐ Dump Data-Engineer-Associate Collection ☐ Search on “ www.pass4leader.com ” for ➡ Data-Engineer-Associate ☐ to obtain exam materials for free download ☐ Data-Engineer-Associate Valid Dumps Sheet
- Data-Engineer-Associate Valid Real Test ☐ Data-Engineer-Associate Valid Dumps Sheet ☐ Data-Engineer-Associate Actualtest ☐ 「 www.pdfvce.com 」 is best website to obtain ☐ Data-Engineer-Associate ☐ for free download ☐ Data-Engineer-Associate Trustworthy Exam Content
- TOP Data-Engineer-Associate Useful Dumps 100% Pass | Valid Amazon New AWS Certified Data Engineer - Associate (DEA-C01) Exam Objectives Pass for sure ☐ Search on 「 www.pass4test.com 」 for ☐ Data-Engineer-Associate ☐ to obtain exam materials for free download ☞ Data-Engineer-Associate Reliable Test Cost
- Hot Data-Engineer-Associate Useful Dumps - 100% Pass-Rate New Data-Engineer-Associate Exam Objectives - Useful Data-Engineer-Associate Certification ☐ Open “ www.pdfvce.com ” enter ➤ Data-Engineer-Associate ☐ and obtain a free download ☐ Valid Data-Engineer-Associate Exam Tips
- Valid Test Data-Engineer-Associate Vce Free ☐ Dump Data-Engineer-Associate Collection ☐ Data-Engineer-Associate Reliable Exam Practice ☐ Search on ✓ www.actual4labs.com ☐ ✓ ☐ for 《 Data-Engineer-Associate 》 to obtain exam materials for free download ☐ Data-Engineer-Associate Valid Dumps Sheet
- Valid Data-Engineer-Associate Exam Tips ☐ Data-Engineer-Associate Valid Vce Dumps ☐ Data-Engineer-Associate Exam Revision Plan ☐ Open ⇒ www.pdfvce.com ⇐ and search for 【 Data-Engineer-Associate 】 to download exam materials for free ☐ Data-Engineer-Associate Valid Real Test
- Data-Engineer-Associate Exam Revision Plan ☐ Data-Engineer-Associate Real Question ☐ Latest Test Data-Engineer-Associate Discount ☐ Search for ➡ Data-Engineer-Associate ☐ and download it for free on ➡ www.vceengine.com ☐ website ☐ Data-Engineer-Associate Exam Revision Plan
- Valid Data-Engineer-Associate Exam Tips ☐ Data-Engineer-Associate Valid Dumps Sheet ☐ Reliable Data-Engineer-

Associate Exam Sample ☐ ⇒ www.pdfvce.com ⇐ is best website to obtain ☐ Data-Engineer-Associate ☐ for free download ☐ Test Data-Engineer-Associate Simulator Free

- High Effective AWS Certified Data Engineer - Associate (DEA-C01) Test Torrent Make the Most of Your Free Time ☐ Open ⇒ www.actual4labs.com ☐ enter ✓ Data-Engineer-Associate ☐ ✓ ☐ and obtain a free download ☐ Exam Data-Engineer-Associate Practice
- 100% Pass Updated Data-Engineer-Associate - AWS Certified Data Engineer - Associate (DEA-C01) Useful Dumps ☐ The page for free download of ☐ Data-Engineer-Associate ☐ on 《 www.pdfvce.com 》 will open immediately ☐ Reliable Data-Engineer-Associate Exam Sample
- Quiz Perfect Amazon - Data-Engineer-Associate - AWS Certified Data Engineer - Associate (DEA-C01) Useful Dumps ☐ ☐ Search for [Data-Engineer-Associate] and download it for free on ☐ www.dumpsquestion.com ☐ website ☐ Data-Engineer-Associate Real Question
- taqaddm.com, myportal.utt.edu.tt, myportal.utt.edu.tt, myportal.utt.edu.tt, myportal.utt.edu.tt, myportal.utt.edu.tt, ncon.edu.sa, benward394.blogvivi.com, learn.anantraad.in, mindlybody.com, cerfindia.com, chartered-eng.com, benward394.atualblog.com, myportal.utt.edu.tt, myportal.utt.edu.tt, myportal.utt.edu.tt, myportal.utt.edu.tt, myportal.utt.edu.tt, myportal.utt.edu.tt, myportal.utt.edu.tt, myportal.utt.edu.tt, myportal.utt.edu.tt, myportal.utt.edu.tt, Disposable vapes

2025 Latest TestSimulate Data-Engineer-Associate PDF Dumps and Data-Engineer-Associate Exam Engine Free Share:
https://drive.google.com/open?id=1wT_KHAcumSFCvEYo9GkJIg05wE5ygF0r