Data-Engineer-Associate dumps torrent: AWS Certified Data Engineer - Associate (DEA-C01) & Data-Engineer-Associate valid test

P.S. Free 2026 Amazon Data-Engineer-Associate dumps are available on Google Drive shared by ITExamSimulator: https://drive.google.com/open?id=1HvSQ7jDrfPPqQyM6XU7gV57NQ92mPNKE

Unlike those impotent practice materials, our Data-Engineer-Associate study questions have salient advantages that you cannot ignore. They are abundant and effective enough to supply your needs of the Data-Engineer-Associate exam. Since we have the same ultimate goals, which is successfully pass the Data-Engineer-Associate Exam. So during your formative process of preparation, we are willing be your side all the time. As long as you have questions on the Data-Engineer-Associate learning braindumps, just contact us!

One of the best features of ITExamSimulator exam questions is free updates for up to 1 year. The ITExamSimulator 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 ITExamSimulator 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!

>> Prep Data-Engineer-Associate Guide <<

New Prep Data-Engineer-Associate Guide | Efficient Data-Engineer-Associate Actual Tests: AWS Certified Data Engineer - Associate (DEA-C01) 100% Pass

Are you ready to take your career to the next level with the AWS Certified Data Engineer - Associate (DEA-C01) (Data-Engineer-Associate)? Look no further than ITExamSimulator for all of your AWS Certified Data Engineer - Associate (DEA-C01) (Data-Engineer-Associate) exam needs. Our comprehensive and cost-effective solution includes regularly updated Amazon Data-Engineer-Associate Exam Questions, available in a convenient PDF format that can be downloaded on any device, including PC, laptop, mac, tablet, and smartphone.

Amazon AWS Certified Data Engineer - Associate (DEA-C01) Sample Questions (Q164-Q169):

NEW QUESTION # 164
A company receives test results from testing facilities that are located around the world. The company stores the test results in millions of 1 KB JSON files in an Amazon S3 bucket. A data engineer needs to process the files, convert them into Apache Parquet format, and load them into Amazon Redshift tables. The data engineer uses AWS Glue to process the files, AWS Step Functions to orchestrate the processes, and Amazon EventBridge to schedule jobs.
The company recently added more testing facilities. The time required to process files is increasing. The data engineer must reduce the data processing time.
Which solution will MOST reduce the data processing time?

- A. Use the AWS Glue dynamic frame file-grouping option to ingest the raw input files. Process the files. Load the files into the Amazon Redshift tables.
- B. Use the Amazon Redshift COPY command to move the raw input files from Amazon S3 directly into the Amazon Redshift tables. Process the files in Amazon Redshift.
- C. Use AWS Lambda to group the raw input files into larger files. Write the larger files back to Amazon S3. Use AWS Glue to process the files. Load the files into the Amazon Redshift tables.
- D. Use Amazon EMR instead of AWS Glue to group the raw input files. Process the files in Amazon EMR. Load the files into the Amazon Redshift tables.

Answer: A

Explanation:
Problem Analysis:
Millions of 1 KB JSON files in S3 are being processed and converted to Apache Parquet format using AWS Glue.
Processing time is increasing due to the additional testing facilities.
The goal is to reduce processing time while using the existing AWS Glue framework.
Key Considerations:
AWS Glue offers the dynamic frame file-grouping feature, which consolidates small files into larger, more efficient datasets during processing.
Grouping smaller files reduces overhead and speeds up processing.
Solution Analysis:
Option A: Lambda for File Grouping
Using Lambda to group files would add complexity and operational overhead. Glue already offers built-in grouping functionality.
Option B: AWS Glue Dynamic Frame File-Grouping
This option directly addresses the issue by grouping small files during Glue job execution.
Minimizes data processing time with no extra overhead.
Option C: Redshift COPY Command
COPY directly loads raw files but is not designed for pre-processing (conversion to Parquet).
Option D: Amazon EMR
While EMR is powerful, replacing Glue with EMR increases operational complexity.
Final Recommendation:
Use AWS Glue dynamic frame file-grouping for optimized data ingestion and processing.
Reference:
AWS Glue Dynamic Frames
Optimizing Glue Performance

NEW QUESTION # 165
A data engineer has a one-time task to read data from objects that are in Apache Parquet format in an Amazon S3 bucket. The data engineer needs to query only one column of the data.
Which solution will meet these requirements with the LEAST operational overhead?

- A. Use S3 Select to write a SQL SELECT statement to retrieve the required column from the S3 objects.
- B. Prepare an AWS Glue DataBrew project to consume the S3 objects and to query the required column.
- C. Run an AWS Glue crawler on the S3 objects. Use a SQL SELECT statement in Amazon Athena to query the required column.
- D. Configure an AWS Lambda function to load data from the S3 bucket into a pandas dataframe- Write a SQL SELECT statement on the dataframe to query the required column.

Answer: A

Explanation:
Option B is the best solution to meet the requirements with the least operational overhead because S3 Select is a feature that allows you to retrieve only a subset of data from an S3 object by using simple SQL expressions. S3 Select works on objects stored in CSV, JSON, or Parquet format. By using S3 Select, you can avoid the need to download and process the entire S3 object, which reduces the amount of data transferred and the computation time. S3 Select is also easy to use and does not require any additional services or resources.
Option A is not a good solution because it involves writing custom code and configuring an AWS Lambda function to load data from the S3 bucket into a pandas dataframe and query the required column. This option adds complexity and latency to the data retrieval process and requires additional resources and configuration. Moreover, AWS Lambda has limitations on the execution time, memory, and concurrency, which may affect the performance and reliability of the data retrieval process.
Option C is not a good solution because it involves creating and running an AWS Glue DataBrew project to consume the S3 objects and query the required column. AWS Glue DataBrew is a visual data preparation tool that allows you to clean, normalize, and transform data without writing code. However, in this scenario, the data is already in Parquet format, which is a columnar storage format that is optimized for analytics. Therefore, there is no need to use AWS Glue DataBrew to prepare the data. Moreover, AWS Glue DataBrew adds extra time and cost to the data retrieval process and requires additional resources and configuration.
Option D is not a good solution because it involves running an AWS Glue crawler on the S3 objects and using a SQL SELECT statement in Amazon Athena to query the required column. An AWS Glue crawler is a service that can scan data sources and create metadata tables in the AWS Glue Data Catalog. The Data Catalog is a central repository that stores information about the data sources, such as schema, format, and location. Amazon Athena is a serverless interactive query service that allows you to analyze data in S3 using standard SQL. However, in this scenario, the schema and format of the data are already known and fixed, so there is no need to run a crawler to discover them. Moreover, running a crawler and using Amazon Athena adds extra time and cost to the data retrieval process and requires additional services and configuration.
Reference:
AWS Certified Data Engineer - Associate DEA-C01 Complete Study Guide
S3 Select and Glacier Select - Amazon Simple Storage Service
AWS Lambda - FAQs
What Is AWS Glue DataBrew? - AWS Glue DataBrew
Populating the AWS Glue Data Catalog - AWS Glue
What is Amazon Athena? - Amazon Athena

NEW QUESTION # 166
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. Apply a lifecycle policy to transition records to S3 Standard Infrequent-Access (S3 Standard-IA) storage after 30 days.
- B. Use S3 Intelligent-Tiering storage.
- C. Use S3 Standard-Infrequent Access (S3 Standard-IA) storage for all customer records.
- D. Transition records to S3 Glacier Deep Archive storage after 30 days.

Answer: A

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

NEW QUESTION # 167
A company uses an Amazon Redshift provisioned cluster as its database. The Redshift cluster has five reserved ra3.4xlarge nodes and uses key distribution.
A data engineer notices that one of the nodes frequently has a CPU load over 90%. SQL Queries that run on the node are queued. The other four nodes usually have a CPU load under 15% during daily operations.
The data engineer wants to maintain the current number of compute nodes. The data engineer also wants to balance the load more evenly across all five compute nodes.
Which solution will meet these requirements?

- A. Change the sort key to be the data column that is most often used in a WHERE clause of the SQL SELECT statement.
- B. Upgrade the reserved node from ra3.4xlarqe to ra3.16xlarqe.
- C. Change the primary key to be the data column that is most often used in a WHERE clause of the SQL SELECT statement.
- D. Change the distribution key to the table column that has the largest dimension.

Answer: D

Explanation:
Changing the distribution key to the table column that has the largest dimension will help to balance the load more evenly across all five compute nodes. The distribution key determines how the rows of a table are distributed among the slices of the cluster. If the distribution key is not chosen wisely, it can cause data skew, meaning some slices will have more data than others, resulting in uneven CPU load and query performance.
By choosing the table column that has the largest dimension, meaning the column that has the most distinct values, as the distribution key, the data engineer can ensure that the rows are distributed more uniformly across the slices, reducing data skew and improving query performance.
The other options are not solutions that will meet the requirements. Option A, changing the sort key to be the data column that is most often used in a WHERE clause of the SQL SELECT statement, will not affect the data distribution or the CPU load. The sort key determines the order in which the rows of a table are stored on disk, which can improve the performance of range-restricted queries, but not the load balancing. Option C, upgrading the reserved node from ra3.4xlarge to ra3.16xlarge, will not maintain the current number of compute nodes, as it will increase the cost and the capacity of the cluster. Option D, changing the primary key to be the data column that is most often used in a WHERE clause of the SQL SELECT statement, will not affect the data distribution or the CPU load either. The primary key is a constraint that enforces the uniqueness of the rows in a table, but it does not influence the data layout or the query optimization.
References:
* Choosing a data distribution style
* Choosing a data sort key
* Working with primary keys

NEW QUESTION # 168
A company stores data from an application in an Amazon DynamoDB table that operates in provisioned capacity mode. The workloads of the application have predictable throughput load on a regular schedule.
Every Monday, there is an immediate increase in activity early in the morning. The application has very low usage during weekends. The company must ensure that the application performs consistently during peak usage times.
Which solution will meet these requirements in the MOST cost-effective way?

- A. Change the capacity mode from provisioned to on-demand. Configure the table to scale up and scale down based on the load on the table.
- B. Divide the table into two tables. Provision each table with half of the provisioned capacity of the original table. Spread queries evenly across both tables.
- C. Use AWS Application Auto Scaling to schedule higher provisioned capacity for peak usage times. Schedule lower capacity during off-peak times.
- D. Increase the provisioned capacity to the maximum capacity that is currently present during peak load times.

Answer: C

Explanation:
Amazon DynamoDB is a fully managed NoSQL database service that provides fast and predictable performance with seamless scalability. DynamoDB offers two capacity modes for throughput capacity:
provisioned and on-demand. In provisioned capacity mode, you specify the number of read and write capacity units per second that you expect your application to require. DynamoDB reserves the resources to meet your throughput needs with consistent performance. In on-demand capacity mode, you pay per request and DynamoDB scales the resources up and down automatically based on the actual workload. On-demand capacity mode is suitable for unpredictable workloads that can vary significantly over time1.
The solution that meets the requirements in the most cost-effective way is to use AWS Application Auto Scaling to schedule higher provisioned capacity for peak usage times and lower capacity during off-peak times. This solution has the following advantages:
It allows you to optimize the cost and performance of your DynamoDB table by adjusting the provisioned capacity according to your predictable workload patterns. You can use scheduled scaling to specify the date and time for the scaling actions, and the new minimum and maximum capacity limits. For example, you can schedule higher capacity for every Monday morning and lower capacity for weekends2.
It enables you to take advantage of the lower cost per unit of provisioned capacity mode compared to on- demand capacity mode. Provisioned capacity mode charges a flat hourly rate for the capacity you reserve, regardless of how much you use. On-demand capacity mode charges for each read and write request you consume, with no minimum capacity required. For predictable workloads, provisioned capacity mode can be more cost-effective than on-demand capacity mode1.
It ensures that your application performs consistently during peak usage times by having enough capacity to handle the increased load. You can also use auto scaling to automatically adjust the provisioned capacity based on the actual utilization of your table, and set a target utilization percentage for your table or global secondary index. This way, you can avoid under-provisioning or over-provisioning your table2.
Option A is incorrect because it suggests increasing the provisioned capacity to the maximum capacity that is currently present during peak load times. This solution has the following disadvantages:
It wastes money by paying for unused capacity during off-peak times. If you provision the same high capacity for all times, regardless of the actual workload, you are over-provisioning your table and paying for resources that you don't need1.
It does not account for possible changes in the workload patterns over time. If your peak load times increase or decrease in the future, you may need to manually adjust the provisioned capacity to match the new demand. This adds operational overhead and complexity to your application2.
Option B is incorrect because it suggests dividing the table into two tables and provisioning each table with half of the provisioned capacity of the original table. This solution has the following disadvantages:
It complicates the data model and the application logic by splitting the data into two separate tables. You need to ensure that the queries are evenly distributed across both tables, and that the data is consistent and synchronized between them. This adds extra development and maintenance effort to your application3.
It does not solve the problem of adjusting the provisioned capacity according to the workload patterns. You still need to manually or automatically scale the capacity of each table based on the actual utilization and demand. This may result in under-provisioning or over-provisioning your tables2.
Option D is incorrect because it suggests changing the capacity mode from provisioned to on-demand. This solution has the following disadvantages:
It may incur higher costs than provisioned capacity mode for predictable workloads. On-demand capacity mode charges for each read and write request you consume, with no minimum capacity required. For predictable workloads, provisioned capacity mode can be more cost-effective than on-demand capacity mode, as you can reserve the capacity you need at a lower rate1.
It may not provide consistent performance during peak usage times, as on-demand capacity mode may take some time to scale up the resources to meet the sudden increase in demand. On-demand capacity mode uses adaptive capacity to handle bursts of traffic, but it may not be able to handle very large spikes or sustained high throughput. In such cases, you may experience throttling or increased latency.
1: Choosing the right DynamoDB capacity mode - Amazon DynamoDB
2: Managing throughput capacity automatically with DynamoDB auto scaling - Amazon DynamoDB
3: Best practices for designing and using partition keys effectively - Amazon DynamoDB
[4]: On-demand mode guidelines - Amazon DynamoDB
[5]: How to optimize Amazon DynamoDB costs - AWS Database Blog

[6]: DynamoDB adaptive capacity: How it works and how it helps - AWS Database Blog
[7]: Amazon DynamoDB pricing - Amazon Web Services (AWS)

NEW QUESTION # 169
.....

Keep making progress is a very good thing for all people. If you try your best to improve yourself continuously, you will that you will harvest a lot, including money, happiness and a good job and so on. The Data-Engineer-Associate preparation exam from our company will help you keep making progress. Choosing our Data-Engineer-Associate Study Material, you will find that it will be very easy for you to overcome your shortcomings and become a persistent person. Just come and buy our Data-Engineer-Associate learning guide!

Data-Engineer-Associate Actual Tests: https://www.itexamsimulator.com/Data-Engineer-Associate-brain-dumps.html

The quality & service of Data-Engineer-Associate exam dumps will above your expectations, Amazon Prep Data-Engineer-Associate Guide To fit oneself with the market need, they will choose to make progress in one specific developing direction, The AWS Certified Data Engineer Data-Engineer-Associate 100% pass test helps you to get twice the result with half the effort in learning with its elite study material, By gathering, analyzing, filing essential contents into our Data-Engineer-Associate training quiz, our professional experts have helped more than 98 percent of exam candidates pass the exam effortlessly and efficiently.

Creating a running footer, Welcome to Google Voicemail, The quality & service of Data-Engineer-Associate exam dumps will above your expectations, To fit oneself with the market Data-Engineer-Associate need, they will choose to make progress in one specific developing direction.

Pass Guaranteed Quiz 2026 High Pass-Rate Amazon Prep Data-Engineer-Associate Guide

The AWS Certified Data Engineer Data-Engineer-Associate 100% pass test helps you to get twice the result with half the effort in learning with its elite study material, By gathering, analyzing, filing essential contents into our Data-Engineer-Associate training quiz, our professional experts have helped more than 98 percent of exam candidates pass the exam effortlessly and efficiently.

There is another important reason about why our Data-Engineer-Associate test preparation: AWS Certified Data Engineer - Associate (DEA-C01) can sell like hot cakes in the international market is our considerate after sale service.

- Data-Engineer-Associate Reliable Real Test □ Answers Data-Engineer-Associate Free □ Data-Engineer-Associate Latest Test Discount □ Search on ➡ www.vce4dumps.com □□□ for { Data-Engineer-Associate } to obtain exam materials for free download □Most Data-Engineer-Associate Reliable Questions
- AWS Certified Data Engineer - Associate (DEA-C01) dumps torrent - Data-Engineer-Associate exam pdf - AWS Certified Data Engineer - Associate (DEA-C01) study practice □ Search for （ Data-Engineer-Associate ） and download it for free on 《 www.pdfvce.com 》 website □Data-Engineer-Associate Latest Materials
- Reliable Data-Engineer-Associate Study Guide □ Data-Engineer-Associate Reliable Real Test □ Data-Engineer-Associate Latest Materials □ Search for ☀ Data-Engineer-Associate □☀□ on ➽ www.troytecdumps.com □ immediately to obtain a free download ↖Answers Data-Engineer-Associate Free
- Data-Engineer-Associate test valid questions - Data-Engineer-Associate exam latest torrent - Data-Engineer-Associate test review dumps □ Search for □ Data-Engineer-Associate □ and obtain a free download on ➽ www.pdfvce.com □ □ □Data-Engineer-Associate Preparation Store
- Reliable Data-Engineer-Associate Study Guide □ Data-Engineer-Associate Preparation Store □ Trustworthy Data-Engineer-Associate Exam Torrent □ Search on 「 www.dumpsmaterials.com 」 for □ Data-Engineer-Associate □ to obtain exam materials for free download □Data-Engineer-Associate Exam Torrent
- Desktop-Based Amazon Data-Engineer-Associate Practice Test Software □ Easily obtain free download of [Data-Engineer-Associate] by searching on □ www.pdfvce.com □ ☯Data-Engineer-Associate Exam Registration
- Answers Data-Engineer-Associate Free □ Data-Engineer-Associate Test Quiz □ Data-Engineer-Associate Reliable Real Test □ Search for □ Data-Engineer-Associate □ and download exam materials for free through 【 www.examdiscuss.com 】 □Data-Engineer-Associate Updated Dumps
- Data-Engineer-Associate Latest Test Discount □ Answers Data-Engineer-Associate Free □ Most Data-Engineer-Associate Reliable Questions □ Search on ➡ www.pdfvce.com □□□ for □ Data-Engineer-Associate □ to obtain exam materials for free download □Reliable Data-Engineer-Associate Study Guide
- High Pass-Rate Prep Data-Engineer-Associate Guide - Leading Provider in Qualification Exams - Fast Download Data-Engineer-Associate Actual Tests ❣ Search for ➽ Data-Engineer-Associate □ and obtain a free download on ➥ www.troytecdumps.com □ □Verified Data-Engineer-Associate Answers

- Desktop-Based Amazon Data-Engineer-Associate Practice Test Software □ Simply search for ▸ Data-Engineer-Associate ◂ for free download on □ www.pdfvce.com □ □Data-Engineer-Associate Preparation Store
- Answers Data-Engineer-Associate Free □ Exam Data-Engineer-Associate Sample □ Most Data-Engineer-Associate Reliable Questions □ Simply search for 「 Data-Engineer-Associate 」 for free download on { www.examcollectionpass.com } □Practice Data-Engineer-Associate Exam Online
- alyssatilt695400.blogaritma.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, bookmarksden.com, www.stes.tyc.edu.tw, www.stes.tyc.edu.tw, larissafumq393648.thenerdsblog.com, www.stes.tyc.edu.tw, lancejolw424898.elbloglibre.com, miriamsiqt555838.wikiworldstock.com, www.dkcomposite.com, Disposable vapes

P.S. Free 2026 Amazon Data-Engineer-Associate dumps are available on Google Drive shared by ITExamSimulator: https://drive.google.com/open?id=1HvSQ7jDrfPPqQyM6XU7gV57NQ92mPNKE