

Updated Data-Engineer-Associate Testkings & Data-Engineer-Associate Valid Test Tutorial



P.S. Free & New Data-Engineer-Associate dumps are available on Google Drive shared by GuideTorrent:
<https://drive.google.com/open?id=1-7v6OIM65ygBfDEHhFyZ2YSC7FoHGEB>

GuideTorrent is the only one able to provide you the best and fastest updating information about Amazon Certification Data-Engineer-Associate Exam. Other websites may also provide information about Amazon certification Data-Engineer-Associate exam, but if you compare with each other, you will find that GuideTorrent provide the most comprehensive and highest quality information. And most of the information of other websites comes mainly from GuideTorrent.

Amazon Data-Engineer-Associate Exam Syllabus Topics:

Section	Weight	Objectives
Topic 1: Data Operations and Support	22%	- Ensure reliability and scalability - Backup, restore, and disaster recovery - Automate operational tasks - Monitor and troubleshoot data pipelines • 1. CloudWatch, X-Ray, logging and metrics
Topic 2: Data Security and Governance	18%	- Protect sensitive data - Encrypt data at rest and in transit - Implement access control and authentication • 1. IAM, Lake Formation, resource policies - Enforce compliance and data governance • 1. Data lineage, audit logging, regulatory requirements
Topic 3: Data Store Management	26%	- Optimize storage performance and cost - Design and implement data storage solutions • 1. Data lakes, data warehouses, databases • 2. S3, Redshift, DynamoDB, RDS, Lake Formation - Manage data lifecycle and storage tiers

Topic 4: Data Ingestion and Transformation	34%	- Transform and enrich data • 1. Use Spark, EMR, Step Functions • 2. Orchestrate data pipelines • 3. Apply data processing logic - Implement data quality and validation - Ingest data from various sources • 1. Batch and streaming data ingestion • 2. Use services like Kinesis, DMS, Glue, S3
--	-----	--

>> Updated Data-Engineer-Associate Testkings <<

Avail Pass-Sure Updated Data-Engineer-Associate Testkings to Pass Data-Engineer-Associate on the First Attempt

The Amazon Data-Engineer-Associate certification exam also enables you to stay updated and competitive in the market which will help you to gain more career opportunities. Do you want to gain all these Data-Engineer-Associate certification exam benefits? Looking for the quick and complete AWS Certified Data Engineer - Associate (DEA-C01) (Data-Engineer-Associate) exam dumps preparation way that enables you to pass the AWS Certified Data Engineer - Associate (DEA-C01) in Data-Engineer-Associate certification exam with good scores?

Amazon AWS Certified Data Engineer - Associate (DEA-C01) Sample Questions (Q273-Q278):

NEW QUESTION # 273

A data engineer must orchestrate a series of Amazon Athena queries that will run every day. Each query can run for more than 15 minutes.

Which combination of steps will meet these requirements MOST cost-effectively? (Choose two.)

- A. Use Amazon Managed Workflows for Apache Airflow (Amazon MWAA) to orchestrate the Athena queries in AWS Batch.
- B. Use an AWS Glue Python shell job and the Athena Boto3 client `start_query_execution` API call to invoke the Athena queries programmatically.
- C. Use an AWS Lambda function and the Athena Boto3 client `start_query_execution` API call to invoke the Athena queries programmatically.
- D. Create an AWS Step Functions workflow and add two states. Add the first state before the Lambda function. Configure the second state as a Wait state to periodically check whether the Athena query has finished using the Athena Boto3 `get_query_execution` API call. Configure the workflow to invoke the next query when the current query has finished running.
- E. Use an AWS Glue Python shell script to run a sleep timer that checks every 5 minutes to determine whether the current Athena query has finished running successfully. Configure the Python shell script to invoke the next query when the current query has finished running.

Answer: C,D

Explanation:

Option A and B are the correct answers because they meet the requirements most cost-effectively. Using an AWS Lambda function and the Athena Boto3 client `start_query_execution` API call to invoke the Athena queries programmatically is a simple and scalable way to orchestrate the queries. Creating an AWS Step Functions workflow and adding two states to check the query status and invoke the next query is a reliable and efficient way to handle the long-running queries.

Option C is incorrect because using an AWS Glue Python shell job to invoke the Athena queries programmatically is more expensive than using a Lambda function, as it requires provisioning and running a Glue job for each query.

Option D is incorrect because using an AWS Glue Python shell script to run a sleep timer that checks every 5 minutes to determine whether the current Athena query has finished running successfully is not a cost-effective or reliable way to orchestrate the queries, as it wastes resources and time.

Option E is incorrect because using Amazon Managed Workflows for Apache Airflow (Amazon MWAA) to orchestrate the Athena queries in AWS Batch is an overkill solution that introduces unnecessary complexity and cost, as it requires setting up and managing an Airflow environment and an AWS Batch compute environment.

AWS Certified Data Engineer - Associate DEA-C01 Complete Study Guide, Chapter 5: Data Orchestration, Section 5.2: AWS Lambda, Section 5.3: AWS Step Functions, Pages 125-135 Building Batch Data Analytics Solutions on AWS, Module 5: Data Orchestration, Lesson 5.1: AWS Lambda, Lesson 5.2: AWS Step Functions, Pages 1-15 AWS Documentation Overview, AWS Lambda Developer Guide, Working with AWS Lambda Functions, Configuring Function Triggers, Using AWS Lambda with Amazon Athena, Pages 1-4 AWS Documentation Overview, AWS Step Functions Developer Guide, Getting Started, Tutorial: Create a Hello World Workflow, Pages 1-8

NEW QUESTION # 274

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

Answer: A

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 # 275

A data engineer needs to securely transfer 5 TB of data from an on-premises data center to an Amazon S3 bucket. Approximately 5% of the data changes every day. Updates to the data need to be regularly proliferated to the S3 bucket. The data includes files that are in multiple formats. The data engineer needs to automate the transfer process and must schedule the process to run periodically. Which AWS service should the data engineer use to transfer the data in the MOST operationally efficient way?

- **A. AWS DataSync**
- B. AWS Direct Connect
- C. AWS Glue
- D. Amazon S3 Transfer Acceleration

Answer: A

Explanation:

AWS DataSync is an online data movement and discovery service that simplifies and accelerates data migrations to AWS as well as moving data to and from on-premises storage, edge locations, other cloud providers, and AWS Storage services¹. AWS DataSync can copy data to and from various sources and targets, including Amazon S3, and handle files in multiple formats. AWS DataSync also supports incremental transfers, meaning it can detect and copy only the changes to the data, reducing the amount of data transferred and improving the performance. AWS DataSync can automate and schedule the transfer process using triggers, and monitor the progress and status of the transfers using CloudWatch metrics and events¹.

AWS DataSync is the most operationally efficient way to transfer the data in this scenario, as it meets all the requirements and offers a serverless and scalable solution. AWS Glue, AWS Direct Connect, and Amazon S3 Transfer Acceleration are not the best

options for this scenario, as they have some limitations or drawbacks compared to AWS DataSync. AWS Glue is a serverless ETL service that can extract, transform, and load data from various sources to various targets, including Amazon S3². However, AWS Glue is not designed for large-scale data transfers, as it has some quotas and limits on the number and size of files it can process³. AWS Glue also does not support incremental transfers, meaning it would have to copy the entire data set every time, which would be inefficient and costly.

AWS Direct Connect is a service that establishes a dedicated network connection between your on-premises data center and AWS, bypassing the public internet and improving the bandwidth and performance of the data transfer. However, AWS Direct Connect is not a data transfer service by itself, as it requires additional services or tools to copy the data, such as AWS DataSync, AWS Storage Gateway, or AWS CLI. AWS Direct Connect also has some hardware and location requirements, and charges you for the port hours and data transfer out of AWS.

Amazon S3 Transfer Acceleration is a feature that enables faster data transfers to Amazon S3 over long distances, using the AWS edge locations and optimized network paths. However, Amazon S3 Transfer Acceleration is not a data transfer service by itself, as it requires additional services or tools to copy the data, such as AWS CLI, AWS SDK, or third-party software. Amazon S3 Transfer Acceleration also charges you for the data transferred over the accelerated endpoints, and does not guarantee a performance improvement for every transfer, as it depends on various factors such as the network conditions, the distance, and the object size.

References:

AWS DataSync

AWS Glue

AWS Glue quotas and limits

[AWS Direct Connect]

[Data transfer options for AWS Direct Connect]

[Amazon S3 Transfer Acceleration]

[Using Amazon S3 Transfer Acceleration]

NEW QUESTION # 276

A financial company wants to use Amazon Athena to run on-demand SQL queries on a petabyte-scale dataset to support a business intelligence (BI) application. An AWS Glue job that runs during non-business hours updates the dataset once every day. The BI application has a standard data refresh frequency of 1 hour to comply with company policies.

A data engineer wants to cost optimize the company's use of Amazon Athena without adding any additional infrastructure costs. Which solution will meet these requirements with the LEAST operational overhead?

- A. Configure an Amazon S3 Lifecycle policy to move data to the S3 Glacier Deep Archive storage class after 1 day
- B. Add an Amazon ElastiCache cluster between the BI application and Athena.
- C. Change the format of the files that are in the dataset to Apache Parquet.
- **D. Use the query result reuse feature of Amazon Athena for the SQL queries.**

Answer: D

Explanation:

The best solution to cost optimize the company's use of Amazon Athena without adding any additional infrastructure costs is to use the query result reuse feature of Amazon Athena for the SQL queries. This feature allows you to run the same query multiple times without incurring additional charges, as long as the underlying data has not changed and the query results are still in the query result location in Amazon S3¹. This feature is useful for scenarios where you have a petabyte-scale dataset that is updated infrequently, such as once a day, and you have a BI application that runs the same queries repeatedly, such as every hour. By using the query result reuse feature, you can reduce the amount of data scanned by your queries and save on the cost of running Athena. You can enable or disable this feature at the workgroup level or at the individual query level¹.

Option A is not the best solution, as configuring an Amazon S3 Lifecycle policy to move data to the S3 Glacier Deep Archive storage class after 1 day would not cost optimize the company's use of Amazon Athena, but rather increase the cost and complexity. Amazon S3 Lifecycle policies are rules that you can define to automatically transition objects between different storage classes based on specified criteria, such as the age of the object². S3 Glacier Deep Archive is the lowest-cost storage class in Amazon S3, designed for long-term data archiving that is accessed once or twice in a year³. While moving data to S3 Glacier Deep Archive can reduce the storage cost, it would also increase the retrieval cost and latency, as it takes up to 12 hours to restore the data from S3 Glacier Deep Archive³. Moreover, Athena does not support querying data that is in S3 Glacier or S3 Glacier Deep Archive storage classes⁴. Therefore, using this option would not meet the requirements of running on-demand SQL queries on the dataset.

Option C is not the best solution, as adding an Amazon ElastiCache cluster between the BI application and Athena would not cost optimize the company's use of Amazon Athena, but rather increase the cost and complexity. Amazon ElastiCache is a service that offers fully managed in-memory data stores, such as Redis and Memcached, that can improve the performance and scalability of web applications by caching frequently accessed data. While using ElastiCache can reduce the latency and load on the BI application, it would not reduce the amount of data scanned by Athena, which is the main factor that determines the cost of running Athena. Moreover, using ElastiCache would introduce additional infrastructure costs and operational overhead, as you would have

to provision, manage, and scale the ElastiCache cluster, and integrate it with the BI application and Athena.

Option D is not the best solution, as changing the format of the files that are in the dataset to Apache Parquet would not cost optimize the company's use of Amazon Athena without adding any additional infrastructure costs, but rather increase the complexity. Apache Parquet is a columnar storage format that can improve the performance of analytical queries by reducing the amount of data that needs to be scanned and providing efficient compression and encoding schemes. However, changing the format of the files that are in the dataset to Apache Parquet would require additional processing and transformation steps, such as using AWS Glue or Amazon EMR to convert the files from their original format to Parquet, and storing the converted files in a separate location in Amazon S3. This would increase the complexity and the operational overhead of the data pipeline, and also incur additional costs for using AWS Glue or Amazon EMR. References:

Query result reuse

Amazon S3 Lifecycle

S3 Glacier Deep Archive

Storage classes supported by Athena

[What is Amazon ElastiCache?]

[Amazon Athena pricing]

[Columnar Storage Formats]

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

NEW QUESTION # 277

A healthcare company stores patient records in an on-premises MySQL database. The company creates an application to access the MySQL database. The company must enforce security protocols to protect the patient records. The company currently rotates database credentials every 30 days to minimize the risk of unauthorized access.

The company wants a solution that does not require the company to modify the application code for each credential rotation.

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

- **A. Use AWS Secrets Manager to automatically rotate credentials. Allow the application to retrieve the credentials by using API calls.**
- B. Assign an IAM role access permissions to the database. Configure the application to obtain temporary credentials through the IAM role.
- C. Use AWS Key Management Service (AWS KMS) to generate encryption keys. Configure automatic key rotation. Store the encrypted credentials in an Amazon DynamoDB table.
- D. Store credentials in an encrypted Amazon S3 bucket. Rotate the credentials every month by using an S3 Lifecycle policy. Use bucket policies to control access.

Answer: A

Explanation:

The correct solution is C: AWS Secrets Manager.

* Why? AWS Secrets Manager is a fully managed service that helps you protect access to your applications, services, and IT resources without the upfront cost and complexity of managing your own hardware security module (HSM) infrastructure.

* It allows for automatic rotation of secrets without requiring changes to the application code, meeting the requirement of minimal operational overhead.

* Applications can securely retrieve credentials using Secrets Manager APIs, and the service integrates with AWS Identity and Access Management (IAM) to control access to secrets.

"You can use Secrets Manager to store credentials and to configure automatic rotation." Reference: AWS Certified Data Engineer - Associate Study Guide, Chapter 7 - Data Security and Governance Also verified in AWS Documentation:

<https://docs.aws.amazon.com/secretsmanager/latest/userguide/rotating-secrets.html>

NEW QUESTION # 278

.....

Amazon will provide you with all the Amazon Data-Engineer-Associate exam dumps, practice exams, and other necessary documentation that will help you understand the Amazon Data-Engineer-Associate exam questions and pass the Amazon Data-Engineer-Associate Exam. You will find it easy to adjust to this new thing and get complete support from the Amazon Data-Engineer-Associate exam questions and practice exams for the Amazon Data-Engineer-Associate certification exam.

Data-Engineer-Associate Valid Test Tutorial: <https://www.guidetorrent.com/Data-Engineer-Associate-pdf-free-download.html>

- Valid Data-Engineer-Associate Vce ☐ Dumps Data-Engineer-Associate Guide ☉ New Data-Engineer-Associate Exam Experience ☐ Search on ▷ www.practicevce.com ◁ for ☐ Data-Engineer-Associate ☐ to obtain exam materials for free download ☐ Data-Engineer-Associate Pass Test Guide
- Data-Engineer-Associate Actual Braindumps ☐ Questions Data-Engineer-Associate Pdf ☐ Data-Engineer-Associate Free Sample Questions ✓ ☐ Easily obtain ▷ Data-Engineer-Associate ◁ for free download through ☼ www.pdfvce.com ☐☼ ☐ Data-Engineer-Associate Reliable Exam Vce
- 100% Pass 2026 Data-Engineer-Associate: AWS Certified Data Engineer - Associate (DEA-C01) Perfect Updated Testkings ☐ Search for ➡ Data-Engineer-Associate ☐ and download it for free immediately on 《 www.pass4test.com 》 ☐ Data-Engineer-Associate Reliable Exam Vce
- 100% Pass 2026 Data-Engineer-Associate: AWS Certified Data Engineer - Associate (DEA-C01) Perfect Updated Testkings ☐ Open ➡ www.pdfvce.com ☐ enter ⇒ Data-Engineer-Associate ⇐ and obtain a free download ☐ New Data-Engineer-Associate Exam Experience
- Test Data-Engineer-Associate Dates ☐ Data-Engineer-Associate Actual Braindumps ☐ Data-Engineer-Associate Reliable Exam Vce ☐ Open website { www.dumpsmaterials.com } and search for ➤ Data-Engineer-Associate ☐ for free download ☐ New Data-Engineer-Associate Exam Experience
- Authorized Data-Engineer-Associate Exam Dumps ☐ Authorized Data-Engineer-Associate Exam Dumps ☐ Data-Engineer-Associate Practice Test Engine ☐ Go to website ✓ www.pdfvce.com ☐✓ ☐ open and search for ☼ Data-Engineer-Associate ☐☼ ☐ to download for free ☐ Data-Engineer-Associate Reliable Exam Vce
- Data-Engineer-Associate Pass Test Guide ☐ Latest Data-Engineer-Associate Test Blueprint ☐ Updated Data-Engineer-Associate Test Cram ☐ Search for ➡ Data-Engineer-Associate ☐ and download it for free on ⇒ www.troytecdumps.com ⇐ website ☐ Data-Engineer-Associate Reliable Test Cram
- Data-Engineer-Associate Pass Test Guide ☐ Latest Data-Engineer-Associate Test Blueprint ☐ Data-Engineer-Associate Practice Test Engine ☐ Search for “Data-Engineer-Associate” on ✓ www.pdfvce.com ☐✓ ☐ immediately to obtain a free download ☐ Data-Engineer-Associate Practice Test Engine
- Get Fantastic Updated Data-Engineer-Associate Testkings and Pass Exam in First Attempt ☐ Go to website (www.practicevce.com) open and search for 「 Data-Engineer-Associate 」 to download for free ☐ Latest Data-Engineer-Associate Test Blueprint
- Easy to use Formats of Pdfvce Amazon Data-Engineer-Associate Practice Exam Material ☐ Open website ➡ www.pdfvce.com ☐☐☐ and search for ➡ Data-Engineer-Associate ☐ for free download ☐ Data-Engineer-Associate Test Preparation
- Get Fantastic Updated Data-Engineer-Associate Testkings and Pass Exam in First Attempt ☐ Search for (Data-Engineer-Associate) and download it for free on ☐ www.troytecdumps.com ☐ website ☐ Latest Data-Engineer-Associate Test Blueprint
- www.stes.tyc.edu.tw, www.stes.tyc.edu.tw, myportal.utt.edu.tt, myportal.utt.edu.tt, myportal.utt.edu.tt, myportal.utt.edu.tt, www.stes.tyc.edu.tw, dorahacks.io, www.stes.tyc.edu.tw, www.stes.tyc.edu.tw, 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, www.stes.tyc.edu.tw, www.stes.tyc.edu.tw, Disposable vapes

What's more, part of that GuideTorrent Data-Engineer-Associate dumps now are free: <https://drive.google.com/open?id=1-7v6OIM65ygBfFdEHhFyZ2YSC7FoHGEB>