

PRACTICE QUESTIONS

AWS

DEA-C01

AWS Certified Data Engineer - Associate

SHARED BY [u/hopeful_c01](#) on aiexamcerts

EDITION Free Set 2

QUESTIONS 10

SHARED October 15, 2026

EXAM AWS DEA-C01, AWS Certified Data Engineer - Associate

Shared by [u/hopeful_c01](#) on aiexamcerts

10 practice questions for the AWS DEA-C01 exam, AWS Certified Data Engineer - Associate, each followed by its answer and a short explanation. More free sets and the full practice exam: www.aiexamcerts.com/aws/dea-c01. Spotted a mistake? Email support@aiexamcerts.com with the exam code DEA-C01 and the question number.

1. Which Kinesis Data Streams feature gives each registered consumer its own dedicated read throughput on every shard?

- A Enhanced fan-out
- B Shard iterators with GetRecords
- C Kinesis Producer Library aggregation
- D Server-side encryption with KMS

Answer: A

Enhanced fan-out gives each registered consumer a dedicated 2 MB per second of read throughput per shard, pushed over HTTP/2. Shared GetRecords consumers split the 2 MB per second per shard limit. KPL aggregation affects how producers pack records, and encryption protects data without changing read throughput.

2. AWS Lake Formation can grant permissions on Data Catalog resources to IAM principals and enforce them in query engines such as Amazon Athena.

- A True
- B False

Answer: A (True)

Lake Formation stores grants on databases, tables, and columns and applies them when Athena, Redshift Spectrum, or EMR read the data. The IAM-only approach does not enforce column or row restrictions in those engines, which is why Lake Formation exists.

3. Which Amazon Redshift feature lets one cluster read live data from another cluster without copying it?

- A Redshift Spectrum
- B Redshift data sharing
- C Redshift Serverless snapshots
- D Concurrency scaling

Answer: B

Data sharing exposes live tables from a producer cluster or serverless namespace to consumers, which read the data without copying it. Spectrum reads files in S3, not another cluster's tables. Snapshots are point-in-time backups, and concurrency scaling adds temporary capacity for queries.

4. Aurora keeps six copies of each data block across three Availability Zones in a shared, distributed _____ layer.

Answer: storage (also accepted: storage layer)

Aurora separates compute from a distributed storage layer that replicates data across three Availability Zones, which is what lets it fail over quickly. Standard RDS Multi-AZ instances keep a separate standby copy rather than a shared storage layer.

5. Which two AWS services can run serverless data transformation jobs without you managing servers? (Choose two.)

- A AWS Glue
- B Amazon EMR on a fixed EC2 cluster
- C Amazon EMR Serverless
- D Amazon Kinesis Video Streams
- E Amazon Textract

Answer: A, C

AWS Glue and EMR Serverless both run Spark-based transformations without you provisioning or patching servers. An EMR cluster on EC2 requires you to size and manage the nodes. Kinesis Video Streams ingests video, and Textract extracts text from documents, so neither transforms tabular data.

6. A transformation must join two streams of events arriving in near real time and produce results continuously. Which AWS service is designed for stream processing with SQL or Apache Flink?

- A Amazon Athena with SQL over stored data
- B AWS Glue DataBrew for batch data preparation
- C Amazon Managed Service for Apache Flink
- D Amazon QuickSight dashboards and visualizations

Answer: C

Amazon Managed Service for Apache Flink runs continuous stream processing jobs, including joins over windows, with Apache Flink or SQL. Athena runs batch-style queries over stored data. DataBrew is a visual preparation tool for batch datasets, and QuickSight is a visualization service.

7. A team wants a notification when failed Glue job runs exceed a threshold within a day. Which combination fits best?
- A CloudTrail log file validation with Athena queries
 - B Glue crawler schedule change with S3 event notifications
 - C Macie findings exported to Redshift
 - D A CloudWatch alarm on Glue failure metrics with SNS

Answer: D

Glue publishes job metrics to CloudWatch, and an alarm on the failure count can notify an SNS topic when the threshold is crossed. CloudTrail validation proves log integrity and does not alert on job outcomes. Crawler schedules and Macie findings do not report on job failures.

8. Amazon EventBridge is the service that stores and queries the Redshift query history.
- A True
 - B False

Answer: B (False)

Redshift keeps query history in system tables such as the STL and SYS views, which you query with SQL. EventBridge is an event bus that routes events between services, so it does not store query history. The statement confuses event routing with system logging.

9. Which AWS service lets you create and control encryption keys, with key policies and grants, and records key usage in AWS CloudTrail?
- A AWS Key Management Service
 - B AWS Certificate Manager
 - C AWS Secrets Manager rotation schedules
 - D Amazon Cognito user pools

Answer: A

AWS KMS manages customer managed keys, controls their use through key policies and grants, and logs cryptographic operations to CloudTrail. Certificate Manager issues TLS certificates. Secrets Manager stores and rotates secrets but does not hold the keys that encrypt your data, and Cognito handles user sign-in.

10. Which S3 storage class suits data that is accessed rarely but must be retrieved in milliseconds, with a 30-day minimum storage duration?
- A S3 Glacier Deep Archive
 - B S3 Standard-Infrequent Access
 - C S3 Standard
 - D S3 Glacier Flexible Retrieval class

Answer: B

S3 Standard-IA gives millisecond access at a lower storage price, with a 30-day minimum storage duration and a retrieval fee. Glacier Deep Archive and Flexible Retrieval have hours-long or minutes-to-hours retrieval, so they do not meet the millisecond requirement. S3 Standard has no minimum duration and is priced for frequent access.