AWS Databases— CLF 02 MCQs with Interview prep

LinkedIn: [Amit Singh]

Medium: Amit Singh – Medium

Databases & Analytics Summary

- [ ] **Relational Databases** - OLTP: RDS & Aurora (SQL)
- [ ] **Differences between** Multi-AZ, Read Replicas, Multi-Region
- [ ] **In-memory Database:** ElastiCache
- [ ] **Key/Value Database:** DynamoDB (serverless) & DAX (cache for DynamoDB)
- [ ] **Warehouse - OLAP:** Redshift (SQL)
- [ ] **Hadoop Cluster:** EMR
- [ ] **Athena:** query data on Amazon S3 (serverless & SQL)
- [ ] **QuickSight:** dashboards on your data (serverless)
- [ ] **DocumentDB:** "Aurora for MongoDB" (JSON – NoSQL database)
- [ ] **Amazon QLDB:** Financial Transactions Ledger (immutable journal, cryptographically verifiable)
- [ ] **Amazon Managed Blockchain:** managed Hyperledger Fabric & Ethereum blockchains
- [ ] **Glue:** Managed ETL (Extract Transform Load) and Data Catalog service
- [ ] **Database Migration:** DMS
- [ ] **Neptune:** graph database

General Database Concepts - one word

Q: What are the benefits of using managed databases on AWS?
A: Reduced operational complexity, built-in high availability, disaster recovery, scalability, and enhanced security.

Q: What is the primary advantage of using a database over storing data on disk (like EBS or S3)?
A: Databases allow structured data, indexing, querying, and defining relationships between data.

Q: What are the main types of databases offered by AWS?
A: Relational (SQL), NoSQL, Data Warehousing, In-memory caching.

Relational Databases (SQL)

Q: What AWS service offers a fully managed SQL database solution?
A: Amazon RDS (Relational Database Service).

Q: Which SQL database engines are supported by Amazon RDS?
A: MySQL, PostgreSQL, Oracle, SQL Server, MariaDB, and Aurora.

Q: What is Amazon Aurora?
A: A high-performance, MySQL- and PostgreSQL-compatible relational database built for the cloud.

Q: How much faster is Aurora compared to standard MySQL and PostgreSQL?
A: 5x faster than MySQL and 3x faster than PostgreSQL.

Q: What storage does Aurora use and how does it scale?
A: Auto-scaling distributed storage up to 64 TB.

Amazon RDS Features & Deployments

Q: What RDS feature allows automatic failover between Availability Zones?
A: Multi-AZ deployments.

Q: What is the purpose of RDS Read Replicas?
A: To scale the read workload of your database.

Q: Can you SSH into an RDS database instance?
A: No, AWS manages the OS and underlying infrastructure.

Q: What is a key advantage of using RDS instead of installing a database on an EC2 instance?
A: RDS provides automated provisioning, patching, backups, scaling, and monitoring.

Q: What is the difference between Multi-AZ and Read Replicas in RDS?
A: Multi-AZ is for high availability and failover; Read Replicas are for scaling read workloads.

Q: What is the purpose of Multi-Region deployment in RDS?
A: Disaster recovery and local performance for global reads.

NoSQL & DynamoDB

Q: What is DynamoDB?
A: A fully managed, serverless NoSQL database that supports key-value and document models.

Q: What are the benefits of DynamoDB?
A: High availability, single-digit millisecond latency, auto-scaling, integrated with IAM, low cost.

Q: What is DAX (DynamoDB Accelerator)?
A: An in-memory cache for DynamoDB that provides microsecond latency for reads.

Q: What is a common format for NoSQL data?
A: JSON.

Q: How does DynamoDB ensure global availability?
A: Through Global Tables and replication across multiple regions.

In-Memory Databases (Caching)

Q: What is Amazon ElastiCache used for?
A: Caching frequently accessed data to reduce database load and improve latency.

Q: What engines does ElastiCache support?
A: Redis and Memcached.

Analytics & Big Data

Q: What is Amazon Redshift?
A: A fully managed data warehouse optimized for OLAP and big data analytics.

Q: What storage format does Redshift use for performance?
A: Columnar storage.

Q: What is Amazon Athena used for?
A: Serverless SQL querying of data stored in Amazon S3.

Q: What AWS service is best for visualizing data in dashboards?
A: Amazon QuickSight.

Q: What is Amazon EMR?
A: A managed cluster platform for processing big data using Hadoop, Spark, and Hive.

Q: What is AWS Glue used for?
A: Serverless ETL (Extract, Transform, Load) service and data cataloging.

Specialty Databases

Q: What is Amazon DocumentDB?
A: A managed document database service compatible with MongoDB.

Q: What is Amazon Neptune?
A: A managed graph database service optimized for highly connected data.

Q: What is Amazon QLDB best used for?
A: Immutable, cryptographically verifiable ledger use cases like financial transaction history.

Q: What is the difference between Amazon QLDB and Managed Blockchain?
A: QLDB is centralized with immutability; Managed Blockchain supports decentralized ledger technology.

Q: What is Amazon Managed Blockchain?
A: A managed service to create and manage scalable blockchain networks using Hyperledger Fabric or Ethereum.

Q: What is AWS DMS (Database Migration Service) used for?
A: To migrate databases to AWS with minimal downtime.

Q: What is the difference between homogeneous and heterogeneous migrations in AWS DMS?
A: Homogeneous: same engine (e.g., Oracle to Oracle); Heterogeneous: different engines (e.g., SQL Server to Aurora).

MCQ with each concept

1. RDS vs EC2-Based DB

Q: Your team needs to quickly deploy a production-grade PostgreSQL database with high availability and automatic backups, but you don't want to manage OS patching or backups manually. Which AWS service should you choose?
A. Amazon EC2 with PostgreSQL manually installed
B. Amazon RDS for PostgreSQL
C. Amazon DynamoDB
D. Amazon Redshift

2. NoSQL Use Case

Q: Your application needs to store user sessions and rapidly access them with sub-millisecond latency. The structure is key-value and changes often. What should you use?
A. Amazon RDS
B. Amazon ElastiCache
C. Amazon Neptune
D. Amazon QLDB

3. Global Database

Q: You are building a global e-commerce application and need a NoSQL database that replicates data across multiple AWS Regions automatically. Which service is best?
A. Amazon Aurora Global Database
B. Amazon DynamoDB Global Tables

C. Amazon RDS Multi-AZ
D. Amazon Redshift

4. RDS Read Replica

Q: You notice your RDS instance is under high read load. You want to improve read performance without affecting the write operations. What should you implement?
A. Multi-AZ deployment
B. Read Replica
C. Increase instance size
D. Use Amazon ElastiCache

5. Serverless SQL Queries

Q: Your data is stored in S3 in CSV and JSON format. You want to run ad-hoc SQL queries on this data without managing servers. Which service should you use?
A. Amazon Redshift
B. Amazon EMR
C. Amazon Athena
D. Amazon Glue

6. Analytics with Structured Data

Q: You are building a BI dashboard to analyze petabytes of structured data with high performance. What should you choose?
A. Amazon Redshift
B. Amazon Neptune
C. Amazon DynamoDB
D. Amazon QLDB

7. Cache for DynamoDB

Q: You want to reduce latency for frequently accessed data in DynamoDB. What should you implement?
A. Read Replicas
B. Amazon ElastiCache
C. DynamoDB Accelerator (DAX)
D. RDS Proxy

8. Migrate Oracle to Aurora

Q: You are migrating an on-premises Oracle DB to Amazon Aurora PostgreSQL. What service helps with schema conversion and continuous replication?
A. AWS Snowball
B. Amazon DMS

C. Amazon Glue
D. AWS Config

9. Document-Based NoSQL

Q: Your application stores complex JSON documents and requires MongoDB compatibility. Which AWS service should you choose?

A. Amazon DynamoDB
B. Amazon DocumentDB
C. Amazon Aurora
D. Amazon Neptune

10. Graph Database

Q: You're building a social media app and need to model relationships between users and their connections. What AWS service fits best?

A. Amazon Neptune
B. Amazon DynamoDB
C. Amazon Redshift
D. Amazon RDS

#	✅ Answer	Explanation
1	B. Amazon RDS for PostgreSQL	RDS handles backups, patching, monitoring, and failover automatically.
2	B. Amazon ElastiCache	ElastiCache provides in-memory caching with very low latency, ideal for session data.
3	B. Amazon DynamoDB Global Tables	DynamoDB Global Tables provide multi-region, fully active-active replication.
4	B. Read Replica	Read Replicas help scale read-heavy workloads by offloading reads from the main DB.
5	C. Amazon Athena	Athena lets you query S3 data directly using SQL. It's serverless and cost-effective.
6	A. Amazon Redshift	Redshift is optimized for analytics at scale on structured data.
7	C. DynamoDB Accelerator (DAX)	DAX provides an in-memory cache specifically for DynamoDB with microsecond latency.
8	B. Amazon DMS	AWS Database Migration Service (DMS) supports live migrations and replication.
9	B. Amazon DocumentDB	DocumentDB is fully managed and MongoDB-compatible.
10	A. Amazon Neptune	Neptune is a graph database ideal for connected data and relationship queries.

AWS Databases & Analytics – MCQs (CLF-C02 & Interview Level)

Covers Relational, NoSQL, Serverless, Caching, Analytics, and Migration Services.

1. Which of the following is an example of a relational database service in AWS?

A) Amazon DynamoDB
B) Amazon Redshift
C) Amazon RDS
D) Amazon S3

2. What is a key benefit of using Amazon RDS over deploying your own DB on EC2?

A) You get full root access to the OS
B) Automated backups, patching, and high availability
C) Supports all NoSQL databases
D) You can host websites from RDS

3. Which of the following is a NoSQL database offered by AWS?

A) Amazon Neptune
B) Amazon Aurora
C) Amazon DynamoDB
D) Amazon Redshift

4. What type of data structure is commonly used in NoSQL databases like DynamoDB?

A) Tables with rows and columns
B) JSON documents or key-value pairs
C) CSV files
D) SQL scripts

5. Who is responsible for patching and maintaining the database engine when using Amazon RDS?

A) The customer only
B) The database vendor
C) AWS (as per shared responsibility model)
D) Nobody

6. Which RDS deployment enhances read performance?

A) Multi-AZ deployment
B) Read replicas

C) Multi-Region failover
D) EC2 backup

7. Which RDS deployment improves disaster recovery and global performance?

A) Read Replica
B) Multi-AZ
C) Multi-Region
D) IAM integration

8. Which AWS database offers MySQL and PostgreSQL compatibility but with better performance?

A) Amazon DynamoDB
B) Amazon ElastiCache
C) Amazon Aurora
D) Amazon Redshift

9. What is the main purpose of Amazon ElastiCache?

A) Store videos and files
B) Run large-scale analytics
C) Provide in-memory caching to reduce DB load
D) Stream media

10. Which feature of DynamoDB provides microsecond read performance using in-memory cache?

A) DynamoDB Streams
B) DynamoDB Global Tables
C) DynamoDB Accelerator (DAX)
D) DynamoCache

11. Which service allows you to replicate DynamoDB tables across AWS regions?

A) DynamoDB Streams
B) Global Tables
C) RDS Multi-AZ
D) DAX

12. Which AWS service is a fully managed data warehouse?

A) Amazon DynamoDB
B) Amazon Aurora
C) Amazon Redshift
D) Amazon Athena

13. What does Amazon EMR primarily support?

A) Stream live video
B) Process big data using open-source tools like Hadoop, Spark
C) Migrate websites
D) Backup RDS

14. What is the primary function of Amazon Athena?

A) Create RDS instances
B) Analyze data in S3 using SQL
C) Encrypt EBS volumes
D) Stream Kinesis data

15. What is Amazon QuickSight used for?

A) Monitoring AWS services
B) Generating visual dashboards and reports
C) Encrypting data
D) Managing IAM

16. Which AWS database service is compatible with MongoDB?

A) DocumentDB
B) Neptune
C) Aurora
D) DynamoDB

17. Which AWS service supports graph databases and relationships?

A) Amazon Aurora
B) Amazon Neptune
C) DynamoDB
D) Amazon QLDB

18. What is Amazon QLDB primarily used for?

A) Caching data
B) Blockchain-based immutable ledger
C) BI dashboards
D) NoSQL storage

19. What AWS service lets you build your own blockchain network?

A) Amazon QLDB
B) AWS Blockchain Builder
C) Amazon Managed Blockchain
D) AWS ChainDB

20. What is AWS Glue used for?

A) Deploy containers
B) Manage IAM users
C) Prepare and transform data for analytics
D) Create AMIs

21. Which AWS service helps migrate on-premises databases to AWS?

A) CloudWatch
B) EC2
C) AWS DMS (Database Migration Service)
D) Route 53

22. What is a key advantage of using DynamoDB?

A) Runs only in one region
B) Manual scaling required
C) Fully managed and scales automatically
D) Fixed schema like RDS

23. Which database should you choose for ledger applications with audit trail requirements?

A) DynamoDB
B) Amazon QLDB
C) Redshift
D) Aurora

24. What AWS analytics service would be best for querying S3 logs using SQL?

A) EMR
B) Redshift
C) Athena
D) Glue

25. In AWS's shared responsibility model, who is responsible for patching the database engine in RDS?

A) Customer
B) AWS
C) Third-party
D) Both

#	✅ Answer	Explanation
1	C	Amazon RDS is a fully managed relational database service.
2	B	RDS provides managed services like backups, patching, and HA without managing OS.
3	C	DynamoDB is a fully managed NoSQL database optimized for key-value and document data.
4	B	NoSQL databases like DynamoDB store unstructured data like JSON documents.
5	C	AWS manages DB software patching and maintenance for RDS.
6	B	Read replicas allow offloading read queries to improve performance.
7	C	Multi-Region deployments support DR and low-latency global access.
8	C	Aurora is MySQL/PostgreSQL-compatible and offers 3–5x better performance.
9	C	ElastiCache (Redis or Memcached) is used to cache frequently accessed data.
10	C	DAX is an in-memory caching layer that improves DynamoDB read performance.
11	B	DynamoDB Global Tables replicate data across regions automatically.
12	C	Redshift is a scalable, managed data warehouse for OLAP and analytics.
13	B	EMR (Elastic MapReduce) is for processing massive amounts of data using Spark, Hive, etc.
14	B	Athena is a serverless query service to analyze S3 data using SQL.

15	B	QuickSight is a BI tool used to visualize and share insights from data.
16	A	DocumentDB is MongoDB-compatible and supports document-oriented data.
17	B	Neptune is a graph DB service that supports RDF and Property Graph models.
18	B	QLDB is a ledger DB that stores data with cryptographic verifiability.
19	C	Managed Blockchain supports Hyperledger Fabric and Ethereum for blockchain apps.
20	C	AWS Glue is a serverless ETL (Extract, Transform, Load) service for data pipelines.
21	C	AWS DMS helps move databases securely and reliably to AWS.
22	C	DynamoDB is serverless, highly scalable, and fully managed.
23	B	QLDB is purpose-built for ledgers and provides immutable transaction history.
24	C	Athena queries structured and semi-structured data in S3 using SQL.
25	B	AWS handles the underlying infrastructure and DB engine for managed services like RDS.

MCQs: SAA & CLF-C02

1. Which of the following best describes Amazon RDS?

A) A NoSQL database service

B) A fully managed relational database service
C) A file storage system
D) A web application firewall

2. Which database engines are supported by Amazon RDS?
A) MongoDB and Cassandra
B) MySQL, PostgreSQL, MariaDB, Oracle, SQL Server, and Aurora
C) DynamoDB and Redis
D) SQLite and IBM Db2

3. What is a key feature of Amazon Aurora compared to standard MySQL?
A) Cheaper pricing
B) Server access to OS
C) 5x better performance
D) Works only in on-premise systems

4. Which feature in RDS provides automatic failover in case of instance failure?
A) Read Replica
B) Multi-AZ Deployment
C) Elastic Load Balancer
D) CloudFront Distribution

5. What does Amazon RDS NOT allow users to do?
A) Create read replicas
B) Manually access the OS and DB software
C) Enable automated backups
D) Choose the database engine

6. What is the purpose of Amazon RDS Read Replicas?
A) To store static files
B) To support backup recovery
C) To offload read traffic and scale reads
D) To encrypt the database

7. Which of the following is true about Amazon Aurora Serverless?
A) Requires constant manual scaling
B) Only supports Oracle
C) Automatically scales based on demand
D) Works only for non-relational databases

#	✅ Answer	Explanation
1	B	Amazon RDS is a managed service for relational databases like MySQL, PostgreSQL, Oracle, and more.
2	B	RDS supports major relational engines including Aurora, a high-performance cloud-native DB.
3	C	Aurora is up to 5x faster than standard MySQL and 3x faster than PostgreSQL.
4	B	Multi-AZ deployment ensures high availability and automatic failover in the event of a failure.
5	B	RDS is managed by AWS, and users cannot access the underlying operating system.
6	C	Read replicas help distribute and scale read workloads across multiple instances.
7	C	Aurora Serverless automatically adjusts capacity based on workload.

SAA - Level

1. You're building a high-performance web app that needs microsecond latency caching for session data. Which AWS service should you use?

A. Amazon Redshift
B. Amazon ElastiCache using Redis
C. Amazon Athena
D. AWS Glue

2. You need to migrate a 5TB on-premise Oracle DB to AWS with minimal downtime. Which service helps most?

A. AWS Glue
B. AWS DMS with CDC (Change Data Capture)

C. Amazon Athena
D. Amazon RDS Read Replica

3. Your app needs to serve global users with ultra-low latency from local regions. Which database architecture is best?

A. Amazon RDS with Multi-AZ
B. DynamoDB with DAX
C. DynamoDB Global Tables
D. Amazon Redshift

4. Your team runs large-scale log analytics using Spark. Which AWS service fits best?

A. Amazon Athena
B. Amazon EMR
C. Amazon RDS
D. Amazon QuickSight

5. A company wants a fully-managed, immutable, cryptographically verifiable ledger. Which AWS service do you recommend?

A. Amazon QLDB
B. Amazon Neptune
C. Amazon Redshift
D. AWS Glue

#	✅ Correct Answer	Explanation
1	B. Amazon ElastiCache using Redis	Redis provides sub-millisecond latency, ideal for real-time caching like sessions.
2	B. AWS DMS with CDC	AWS DMS with CDC replicates changes while the database remains online, minimizing downtime.
3	C. DynamoDB Global Tables	Global Tables replicate across regions automatically, providing low latency worldwide.
4	B. Amazon EMR	EMR supports Apache Spark and is optimized for big data processing.
5	A. Amazon QLDB	QLDB provides an immutable ledger with cryptographic verification, suitable for compliance and audit.

1. Relational Databases (SQL)

Q: What is a relational database?
A: A database with structured data stored in tables and queried using SQL.

Q: Use cases for relational databases?
A: Transactional systems, financial apps.

Q: Examples of relational DB engines supported by AWS RDS?
A: MySQL, PostgreSQL, Oracle, SQL Server, MariaDB.

2. NoSQL Databases

Q: Key features of NoSQL?
A: Flexible schema, scalable, high-performance.

Q: Use cases for NoSQL?
A: IoT, mobile, real-time apps.

Q: Example of NoSQL data format?
A: JSON

Q: AWS NoSQL services?
A: DynamoDB, DocumentDB, Neptune

3. AWS RDS (Relational Database Service)

Q: What is AWS RDS?
A: A managed service for relational databases.

Q: Key features of RDS?
A: Backups, patching, Multi-AZ, Read Replicas, Point-in-Time Restore.

Q: Read Replica vs Multi-AZ?
A: Read Replica = scalability; Multi-AZ = high availability.

Q: Can you SSH into RDS?
A: ❌ No

4. Amazon Aurora

Q: What is Aurora?
A: A high-performance managed relational DB (MySQL/PostgreSQL compatible).

Q: How much faster is Aurora than standard MySQL?
A: Up to 5x

Q: Max number of read replicas in Aurora?
A: Up to 15

5. DynamoDB

Q: What is DynamoDB?
A: A fully managed, serverless NoSQL database.

Q: Key features of DynamoDB?
A: Key-value/document model, auto scaling, single-digit ms latency, multi-AZ replication.

Q: What is DynamoDB Accelerator (DAX)?
A: In-memory caching for DynamoDB, 10x faster reads.

Q: What are DynamoDB Global Tables?
A: Multi-region replication with low-latency reads/writes.

6. ElastiCache

Q: What is Amazon ElastiCache?
A: In-memory caching service for databases.

Q: Engines supported by ElastiCache?
A: Redis, Memcached.

Q: Purpose of ElastiCache?
A: Reduce DB load, improve response times.

7. Amazon Redshift

Q: What is Amazon Redshift?
A: Managed data warehouse for OLAP and big data analytics.

Q: Redshift storage format?
A: Columnar

Q: What makes Redshift fast?
A: Columnar storage + Massively Parallel Processing (MPP)

8. Amazon EMR

Q: What is Amazon EMR?
A: Managed Hadoop/Spark-based big data processing service.

Q: Use cases for EMR?
A: Machine learning, data transformation, ETL.

Q: Integration with?
A: S3, Redshift, DynamoDB.

9. Amazon Athena

Q: What is Amazon Athena?
A: Serverless SQL query engine for data in S3.

Q: Pricing for Athena?
A: $5 per TB scanned

Q: Best practice for cost savings?
A: Use compressed/columnar formats like Parquet, ORC.

10. Amazon QuickSight

Q: What is Amazon QuickSight?
A: Serverless BI tool to create dashboards and visualizations.

Q: Data sources supported?
A: S3, Redshift, RDS, others.

Q: Pricing model?
A: Per-session

11. DocumentDB

Q: What is DocumentDB?
A: Managed NoSQL document DB, MongoDB-compatible.

Q: Optimized for?
A: JSON document storage

Q: Use cases?
A: Content management, mobile apps

12. Amazon Neptune

Q: What is Amazon Neptune?
A: Managed graph database.

Q: Use cases for Neptune?
A: Social networks, fraud detection, recommendation engines.

Q: Performance?
A: Supports billions of relationships, low-latency queries.

13. Amazon QLDB

Q: What is Amazon QLDB?
A: Immutable ledger database, cryptographically verifiable.

Q: Use cases?
A: Financial transactions, compliance auditing.

Q: Difference from blockchain?
A: Centralized, no decentralization.

14. Amazon Managed Blockchain

Q: What is Amazon Managed Blockchain?
A: Managed blockchain network creation and joining.

Q: Supported frameworks?
A: Ethereum, Hyperledger Fabric

Q: Use case?
A: Multi-party business transactions without a central authority

15. AWS Glue

Q: What is AWS Glue?
A: Serverless ETL and Data Catalog service.

Q: Use cases?
A: Data transformation, cataloging, preparing for analytics.

Q: Integrates with?
A: Athena, Redshift, EMR

16. DMS – Database Migration Service

Q: What is DMS?
A: Fully managed DB migration service.

Q: Supports which types of migration?
A: Homogeneous (e.g., Oracle → Oracle)
Heterogeneous (e.g., SQL Server → Aurora)

Q: Is the source DB available during migration?
A: Yes

Thanks Everyone!

Connect with me: Amit Singh