AWS Solution Architect Associate

Version : C03

Domain 4

Task 3

Design cost-optimized database solutions

Appropriate backup and retention policies

Backup Policies :

Identify Critical Data and Resources

- ➔ Determine which data and resources are critical to your business operations

Use AWS Backup

- ➔ Centralized backup service that allows to automate and manage backups across AWS services
- ➔ Create backup plans with AWS Backup to automate backup schedules

Automated Backups

- ➔ EC2 Instances : Use Amazon Data Lifecycle Manager (DLM) to automate the creation, retention and deletion of EBS snapshots
- ➔ RDS Databases : Enable automated backups and specify a backup window and retention period
- ➔ S3 Buckets : Use versioning and lifecycle policies to automatically backup and archive objects

Custom Scripts and AWS Lambda

- ➔ Custom backup solutions if AWS Backup does not support a particular resource or if you need more control over the backup process

Appropriate backup and retention policies

Retention Policies :

Define Retention Requirements

➔ Retention requirements based on regulatory, compliance, and business needs. Retention periods varies depending on these factors

AWS Backup Retention Rules

➔ Retention rules within AWS Backup to specify how long backups should be retained. Create rules for different types of backups and apply appropriate retention periods to each

Lifecycle Policies for S3

➔ S3 lifecycle policies transition objects to cheaper storage classes (e.g., S3 Glacier) after a certain period and automatically delete objects after they have been retained for the required duration

Compliance and Auditing

➔ Retention policies should comply with industry regulations such as GDPR, HIPAA or SOX. AWS Config and AWS CloudTrail can be used to audit and monitor compliance and retention policies

Appropriate Database Engine

DB Service	Use Case
RDS	General Purpose Relational DB, Transactional Applications, Enterprise Applications
Aurora	HA and High Performance applications, SAS applications, Large scale enterprise systems
DynamoDB	Web,Mobile,Gaming applications,IoT,Real time bidding
Redshift	Data Warehousing, Big Data Analytics, Business Intelligence and reporting
ElasticCache	Caching , Session Stores, Real Time analytics,Gaming leaderboards and queuing
DocumentDB	JSON document storage,Content management system,catalogs,User profiles
Neptune	Graph-based applications,social networking,recommendation engines,fraud detection
Timestream	IoT applications,Operational monitoring, Time series data analysis
QLDB	Applications requiring an authoritative data source,supply chain,finance,manufacturing,registrations

AWS database services and use cases

- **Amazon RDS (Relational Database Service)** : Provides managed relational databases, supporting engines like MySQL, PostgreSQL, MariaDB, Oracle, and SQL Server. Ideal for traditional applications requiring ACID transactions and complex queries
- **Amazon Aurora** : A MySQL- and PostgreSQL-compatible relational database with improved performance and availability. Suitable for enterprise applications needing high scalability and reliability
- **Amazon DynamoDB** : A managed NoSQL database offering low-latency performance for key-value and document data models. Best for mobile, web, gaming, ad tech, and IoT applications
- **Amazon Redshift** : A fully managed data warehouse service optimized for complex queries on large datasets. Ideal for business intelligence and analytics workloads
- **Amazon ElastiCache** : Provides managed Redis or Memcached for in-memory caching to improve application performance. Useful for real-time applications needing quick data retrieval

AWS database services and use cases

- **Amazon Neptune :** A managed graph database service for applications needing graph queries, such as social networking, recommendation engines and fraud detection
- **Amazon DocumentDB :** A managed MongoDB-compatible document database service. Suitable for content management, cataloging and user profiles
- **Amazon Timestream :** A time series database service for IoT and operational applications. Optimized for storing and analyzing time-stamped data
- **Amazon Quantum Ledger Database (QLDB) :** A fully managed ledger database that provides a transparent, immutable, and cryptographically verifiable transaction log. Suitable for systems of record and audit trails

Migrating database schemas

DB schema is a skeleton structure that represents the logical view of the entire db and typically includes -

- **DB Storage Objects**
- **DB Code Objects**

Two types of schema migration :

- **Homogenous Schema Migration** - If the source DB is MySQL 5.6 compliant and is running on RDS,EC2 or outside AWS, Native MySQL tools can be used to export and import the schema.
  - **Exporting DB schema** - mysqldump client utility can be used to export the DB schema
  - **Importing DB schema in Aurora** - By connecting the Aurora instance to the Aurora DB from a MySQL command line client
- **Heterogeneous Schema Migration** - If the source DB is not MySQL compatible , the schema has to be converted to a format compatible with Amazon Aurora. Schema conversion from one DB engine to another is a nontrivial task and may involve rewriting certain parts of th DB and application code.Schema can be converted and migrated to Amazon Aurora by using -
  - **AWS Schema Conversion Tool** - Automatically converts the source database schema
  - **Manual Schema Migration and Third party tools** - Third party tools can be used to migrate the schema to a format that is compatible with MySQL 5.6

Cost-comparison of AWS database types

Database Type	Pricing Model	Key Features
RDS - MySQL	Pay as you go	Fully Managed, Easy to operate and scale
RDS - PostgreSQL	Pay as you go , Free Tier	Fully Managed , Supports High Availability
Aurora	Pay as you go , Pay only for use	MySQL and PostgreSQL compatible , built for the cloud
Redshift	Pay as you go	Scalable data warehouse , Optimised for analytics
DynamoDB	Pay as you go	NoSQL database , Scalable and High performance

AWS Database types and services

Database Type	Service Name	Description
Relational DB's	**RDS**	Managed relational DB.Supported Engines- MySQL,PostgreSQL etc
	Aurora	High Performance managed relational DB.Supported Engines- MySQL,PostgreSQL etc
Key-Value DB's	**DynamoDB**	NoSQL DB service for Key-Value and document data models
Data Warehousing	**Redshift**	Fully Managed data warehouse service for large scale data set analysis
In-Memory DB's	**ElasticCache**	In-Memory caching service supporting Redis and Memcached
Graph DB's	**Neptune**	Managed graph DB supporting Apache TinkerPop and RDF/SPARQL
Time Series DB's	**Timestream**	Fully Managed Time series DB for IoT and operational applications
Ledger DB's	**Quantum Ledger DB**	Fully Managed Ledger DB to track changes in application data over time
Analytics DB's	**EMR**	Managed Cluster platform that simplifies running big data frameworks

AWS Solution Architect Associate

Version : C03

Domain 4

Task 3

The END