AWS Solution Architect Associate
Version : C03
Domain 4
Task 1

Design cost-optimized storage solutions

Access Options

Key access options and configurations :

➜ **Enabling Requester Pays** :
  - ◆ Navigate to the S3 console, choose the bucket you want to enable Requester Pays for
  - ◆ Go to the "Properties" tab and under "Requester pays," choose "Edit," and then select "Enable"

➜ **Accessing Data** :
  - ◆ When accessing a Requester Pays bucket, the requester must include the x-amz-request-payer parameter set to requester in their request headers
  - ◆ The requester must have the necessary permissions to access the bucket and objects

➜ **Billing and Cost Allocation** :
  - ◆ Costs incurred by the requester are billed to their AWS account
  - ◆ S3 object tagging can be used for organizing and managing access for billing and reporting purposes, which is particularly useful in Requester Pays scenarios

➜ **Use Cases** :
  - ◆ Ideal for sharing large datasets with other AWS accounts or public users without incurring costs for data access on the bucket owner
  - ◆ Used in scenarios where datasets are accessed by a wide range of users or entities who should bear the cost of access
  - ◆ By configuring requester Pays buckets, AWS account owners can manage costs effectively while providing flexible access to their data.

Cost Management Tool with usecases

Tool Name	Use Case
AWS Pricing Calculator	Create cost estimates for AWS services and model solutions before building them
AWS Cost Explorer	Analyze and visualize AWS costs and usage over time
AWS Budgets	Set custom cost and usage budgets and receive alerts when you exceed them
AWS Cost Anomaly Detection	Identify unusual spending patterns and potential cost overruns
AWS Cost and Usage Reports	Access comprehensive cost and usage data for detailed billing analysis
AWS Savings Plan	Save up to 72% on AWS usage in exchange for a commitment to use a specific amount
AWS Reserved Instances	Purchase Reserved Instances to save on long-term, steady-state usage
AWS Trusted Advisor	Receive recommendations for cost optimization, security, fault tolerance and performance
AWS Marketplace	Access third-party tools and solutions for cost management and optimization
CloudZero	Monitor and optimize cloud costs with advanced analytics and insights

Appropriate backup\archival solution

➔ **AWS Backup**
- ◆ Automates and centralizes the backup of AWS services such as Amazon RDS, Amazon EFS, Amazon DynamoDB and Amazon EBS
- ◆ It ensures data protection and compliance across AWS services

➔ **Amazon S3 Glacier and S3 Glacier Deep Archive**
- ◆ Provide low-cost, long-term data archiving solutions
- ◆ Ideal for infrequent access and long-term storage with Glacier offering retrieval times from minutes to hours and Deep Archive providing the lowest-cost storage with retrieval times in hours

➔ **AWS Storage Gateway**
- ◆ Enables seamless integration between on-premises environments and AWS, providing a bridge to securely backup data to the AWS cloud
- ◆ Supports hybrid cloud backup solutions and can replace traditional tape backups

➔ **AWS Data Lifecycle Manager**
- ◆ Automates the creation, retention and deletion of EBS snapshots and EBS-backed AMIs, making it easier to manage backups across different lifecycle stages

Block Storage Options

AWS provides several block storage options for use with their services, primarily through Amazon Elastic Block Store (EBS) and instance stores.

➜ **SSD-backed Storage**
- ◆ General Purpose SSD (gp3, gp2): Ideal for a wide range of workloads, including boot volumes, virtual desktops and small to medium databases. These offer a balance of price and performance
- ◆ Provisioned IOPS SSD (io1, io2): Designed for critical applications that require high performance, such as large databases or workloads that require consistent IOPS performance

➜ **HDD-backed Storage**
- ◆ Throughput Optimized HDD (st1), Best for frequently accessed, throughput-intensive workloads, like big data and log processing
- ◆ Cold HDD (sc1), Suitable for less frequently accessed data, providing a low-cost option for infrequently accessed data

➜ **Instance Store**
- ◆ Instance Store Volumes, Provide temporary block-level storage for instances.
- ◆ Physically attached to the host computer and offer high performance but data is lost when the instance is stopped or terminated

Data Lifecycles

AWS offers several tools for managing data lifecycles, each tailored for different types of data and storage needs. Key components :

➔ **Amazon Data Lifecycle Manager (DLM)**
   ◆ **Automates EBS Snapshots and AMIs** : DLM can automatically create, retain, and delete EBS snapshots and EBS-backed AMIs. This helps manage backups and recovery points efficiently
   ◆ **Lifecycle Policies** : Users can create policies to automate operations such as snapshot creation and retention across regions and accounts

➔ **Amazon S3 Lifecycle Configuration**
   ◆ **Rule-based Actions** : S3 Lifecycle configurations consist of rules that specify actions (e.g., transitioning objects to a different storage class, expiration of objects) based on the age or size of the objects
   ◆ **XML Configuration** : These rules are defined in an XML file, and actions can include transitioning data to colder storage like Glacier or automatically deleting objects

Hybrid Storage Options

AWS offers various services for hybrid storage, enabling seamless data transfer and integration between on-premises environments and the cloud

➜ **AWS DataSync**
   ◆ Automates and accelerates moving large amounts of data between on-premises storage and AWS
   ◆ Supports data transfers for migrations, workflows and ongoing replication

➜ **AWS Transfer Family**
   ◆ Provides fully managed support for Secure File Transfer Protocol (SFTP), File Transfer Protocol over SSL (FTPS), and File Transfer Protocol (FTP)
   ◆ Facilitates secure and reliable data transfer into Amazon S3 or Amazon EFS, simplifying file exchange with partners or internal teams

➜ **AWS Storage Gateway**- Bridges on-premises environments with cloud storage
   ◆ **File Gateway** : Integrates with Amazon S3, providing a seamless way to store and retrieve files using standard protocols
   ◆ **Volume Gateway** : Presents cloud-backed iSCSI block storage volumes to your on-premises applications
   ◆ **Tape Gateway** : Provides a virtual tape library, enabling backup and archival to the cloud

Storage data lifecycle

The storage data lifecycle in AWS can be managed using several tools and practices - Amazon S3 and Amazon EBS

➔ **Amazon S3 Lifecycle Management**
  ◆ Involves setting up lifecycle policies to manage objects in S3
  ◆ Rules can be created in an XML file to automatically transition objects between different storage classes (like S3 Standard, S3 Standard-IA, S3 Glacier) or to delete them after a specified period

➔ **Amazon Data Lifecycle Manager (DLM)**
  ◆ Helps manage the lifecycle of Amazon EBS snapshots and AMIs.
  ◆ Creation, retention and deletion of snapshots can be automated ensuring efficient use of storage and compliance with data retention policies

Storage costs can be optimized and compliance with data retention requirements can be maintained by using these tools

Storage Tiering

Amazon S3 Intelligent-Tiering is a storage class in AWS designed to optimize storage costs by automatically moving data between different access tiers based on usage patterns.
It includes three access tiers :

➔ **Frequent Access Tier** : For data that is accessed often

➔ **Infrequent Access Tier** : For data that is accessed less frequently, offering a lower storage cost than the Frequent Access Tier

➔ **Archive Access Tier** : For data that is rarely accessed, providing the lowest storage cost

This service is particularly beneficial for data with unpredictable, unknown or dynamic access patterns, ensuring cost-efficiency without requiring manual intervention.

Storage Tiering

To implement cold tiering for object storage in AWS, Amazon S3's storage classes can be used , particularly the "S3 Intelligent-Tiering" class -

➜ **S3 Intelligent-Tiering** : This storage class automatically moves objects between two access tiers:
  ◆ **Frequent Access Tier** : Optimized for objects accessed frequently.
  ◆ **Infrequent Access Tier** : Designed for objects accessed less frequently but still require immediate availability

➜ **Configuration** : Enable S3 Intelligent-Tiering for your bucket and set up lifecycle policies to transition objects to the Infrequent Access Tier for cold data

➜ **Cost Savings** : This tiering strategy can reduce storage costs by automatically moving objects to a lower-cost tier when they're accessed infrequently, helping optimize costs without sacrificing performance

By leveraging S3 Intelligent-Tiering, Cold data in AWS object storage can be efficiently managed while maintaining accessibility and cost-effectiveness.

Storage types with associated characteristics

➔ **Object Storage :**
  ◆ Description : Manages data as distinct units (objects), suitable for large amounts of unstructured data
  ◆ Use Cases : Ideal for storing photos, videos, backups, and big data analytics
  ◆ Characteristics : Scalable, metadata-rich, accessed via HTTP APIs
  ◆ Examples : Amazon S3

➔ **File Storage** :
  ◆ Description : Stores data in a hierarchical structure with files and directories.
  ◆ Use Cases : Suitable for shared file systems, content management, and home directories
  ◆ Characteristics : Accessible via traditional file protocols (NFS, SMB), good for collaborative environments
  ◆ Examples : Amazon EFS

➔ **Block Storage :**
  ◆ Description : Divides data into fixed-size blocks and manages them individually
  ◆ Use Cases : Best for transactional databases, virtual machines, and applications requiring low latency
  ◆ Characteristics : High performance, low latency, used for structured data.
  ◆ Examples : Amazon EBS

AWS Solution Architect Associate
Version : C03
Domain 4
Task 1

The END