

Azure Storage Account

An Azure Storage Account is a cloud-based storage service in Microsoft Azure that provides scalable, durable, and highly available storage for various data types. It supports multiple storage services under one account, including:

Key Storage Services Included:

Blob Storage – For unstructured data (e.g., documents, images, videos).

File Storage – Managed file shares (SMB/NFS) for cloud or hybrid use.

Table Storage – NoSQL key-value store for semi-structured data.

Queue Storage – Messaging for decoupling application components.

Key Features:

Scalable & Durable – Replicated for high availability (e.g., LRS, ZRS, GRS).

Secure – Encryption at rest, role-based access control (RBAC), and shared access signatures (SAS).

Cost-Effective – Pay-as-you-go pricing with different tiers (Hot, Cool, Archive).

Multi-Protocol Access – Supports REST API, SDKs, and Azure services.

Use Cases: Backup, media storage, app data, logs, and more.

Types of Storage Accounts:

1. Azure Blob Storage

(BLOB - BINARY LARGE OBJECTS)

Purpose: General-purpose object storage for unstructured data (images, videos, backups, etc.).

Key Features:

- No hierarchical namespace (flat structure).
- Supports Hot/Cool/Archive tiers.
- Access via REST API, SDKs, or SFTP (premium only).

Use Case: Simple storage with low cost, no analytics optimizations.

3. Azure Data Lake Storage Gen2 (ADLS Gen2) (Recommended)

Purpose: Unified storage for analytics + Blob Storage capabilities.

Key Features:

- Blob Storage + Hierarchical Namespace (best of both).
- Optimized for Spark, Databricks, Synapse, Hive.
- Supports ACLs and RBAC.
- Multi-protocol access (REST, HDFS, SMB, NFS).


Use Case: Modern data lakes, AI/ML, and high-scale analytics.


Access Tiers

Available Access Tiers

Hot - Optimized for frequently accessed data.

- Lowest access cost, higher storage cost. Active data (e.g., real-time analytics, live app data).


 High storage


 Low access cost

Cool - For infrequently accessed data (stored ≥ 30 days).

- Lower storage cost, higher access cost than Hot.

- Short-term backups, older logs, historical data.

 Lower storage


 Higher access cost

Cold - Similar to Cool, but with lower storage cost and higher retrieval cost.

- Data must remain ≥ 90 days.

- Rarely accessed backups.

 Cheapest storage

 Highest access cost

Redundancy Levels in Azure Data Lake Storage (ADLS Gen2)

Azure Data Lake Storage (ADLS) Gen2 supports multiple redundancy (replication) options to ensure data durability and high availability.

Supported Redundancy Levels

LRS (Locally Redundant Storage) - 3 copies in a single datacenter.

- Lowest cost, but no regional failure protection.
- Non-critical data, dev/test.
- SLA -99.9%

ZRS (Zone-Redundant Storage) - 3 copies across availability zones (AZs) in one region.

- Protects against datacenter failures.
- Production workloads needing high availability.
- SLA - 99.9%

GRS (Geo-Redundant Storage) - 6 copies total (3 in primary region + 3 in secondary region).

- Read access in secondary region (RA-GRS).
- Disaster recovery (DR), backup.
- SLA -99.9%

Summary

LRS: Cheap, single datacenter.

ZRS: Best for high availability in one region.

GRS/GZRS: Best for disaster recovery.

For ADLS Gen2 analytics workloads, ZRS or GZRS is recommended for production.

What is SLA (Service Level Agreement) in Azure?

An SLA (Service Level Agreement) is a formal commitment by Microsoft Azure regarding the uptime, performance, and reliability of its cloud services. It defines:

Guaranteed availability (e.g., 99.9%).

Compensation policies (service credits) if Azure fails to meet the SLA.

Steps to create a storage account:

1. Navigate to the Azure Portal.
2. On the homepage, click on the + Create a resource option.
3. Search for Storage Account and select it.
4. The storage accounts pane appears.
5. Select Create option.
6. Then, select your subscription and the resource group where you want to create this storage account.
7. Enter the name and location of your storage account.
8. Select account kind, replication, and Access Tier and enable hierarchial namespace if you want to create adls gen2 account or disable hierarchial namespace if you want to create blob storage account.

9. Hit Review+Create.

10. Wait for the validation to be successful and then click on Create.

Data Security in Azure Data Lake Storage (ADLS Gen2)

Azure Data Lake Storage (ADLS) Gen2 provides enterprise-grade security for data at rest and in transit. Below are the key security features and best practices:

1. Authentication & Access Control

A. Identity-Based Access

- **Azure Active Directory (Azure AD) Integration**
 - Users, apps, and services authenticate via Azure AD (OAuth 2.0).
 - Supports Multi-Factor Authentication (MFA) for added security.
- **Role-Based Access Control (RBAC)**
 - Assign predefined roles (e.g., Storage Blob Data Owner/Contributor/Reader) at subscription, storage account, or container level.

B. POSIX-Style ACLs (Access Control Lists)






- Fine-grained permissions at file/folder level (e.g., rwx for users/groups).

How Data Files Are Stored in ADLS Gen2

1. Inside the storage account go to data management or storage browser pane and click on container

2. Make a "Container" (Like a Folder Bucket)

- Inside your storage, create a container (e.g., raw-data).
- This acts like the main folder where files will go
- Click on add directories to create folders
- Click on upload to add files in the folder

-  raw-data (container)
-   sales (Directory/folder)
-   data.csv (Uploaded file)

POSSIBLE INTERVIEW QUESTIONS FROM STORAGE ACCOUNT

-Types of storage accounts ?

> Blob storage account and adls gen2

-Difference between BLOB storage and adls gen2 ?

> Hierarchical name spacing (folder like structure) will be in disable state whereas hierarchical name spacing (folder like structure) will be in enable state for ADLS GEN2 account

> Raw data like videos, audios, images will be stored in blob storage whereas raw files in file formats will be stored in adls gen2

-what are different types of Access tiers?

> Hot tier – for frequently accessible data

> Cool tier – for moderately accessible data

> Cold tier – for rarely accessible data

-How data security maintained in adls gen2?

> ROLE BASED ACCESS (RBAC) – role assignment permissions

> Access control list – file level permissions

-Types of Redundancy levels in adls gen2?

LRS: Cheap, single datacenter.

ZRS: Best for high availability in one region.

GRS/GZRS: Best for disaster recovery.