**Azure Storage**

**Azure Blob Storage:**

**Description**: Blob Storage is designed for storing and serving large amounts of unstructured data, such as text, images, videos, and binary files.

**Key** **Features**: High scalability, durability, and availability; supports three types of blobs: block blobs, append blobs, and page blobs; provides tiered storage options (Hot, Cool, and Archive).

**Use** **Cases**: Storing files, media content, backups, logs, and serving static websites.

**Azure Files:**

**Description**: Azure Files provides fully managed file shares in the cloud, allowing you to create file shares accessible via the SMB protocol.

**Key** **Features**: Easy migration of file-based applications to the cloud; supports standard SMB file share features; can be mounted on Windows, Linux, and macOS.

**Use** **Cases**: Lift-and-shift applications with existing file dependencies, shared file storage for applications, document sharing, and cloud-based file servers.

**Azure Tables:**

**Description**: Azure Tables is a NoSQL key-value store for structured data, offering a schemaless storage solution.

**Key** **Features**: Highly scalable and cost-effective; suitable for simple querying and storage needs; provides automatic indexing of properties.

**Use** **Cases**: Storing large volumes of structured data, IoT telemetry, user data, metadata storage, and application configuration.

**Azure Queues:**

**Description**: Azure Queues is a simple message queuing service for decoupling components in distributed systems.

**Key** **Features**: Asynchronous communication between components; reliable message delivery; supports time-based visibility and automatic retries.

**Use** **Cases**: Background job processing, inter-process communication, buffering requests, task orchestration, and event-driven systems.

**Azure Disk Storage:**

**Description**: Azure Disk Storage offers durable and high-performance block storage for Azure VMs.

**Key** **Features**: Managed and unmanaged disk options; supports both Premium SSD and Standard HDD; offers encryption and snapshots.

**Use** **Cases**: Persistent storage for VMs, database storage, virtual machine disks, and application data storage.

**Azure Data Lake Storage:**

**Description**: Azure Data Lake Storage is a scalable and secure data lake solution for big data analytics.

**Key** **Features**: Designed for big data workloads; supports both structured and unstructured data; integrates with analytics frameworks like Hadoop and Spark.

**Use** **Cases**: Big data analytics, data exploration, machine learning, data warehousing, and data archiving

**Comparative study between AWS and Azure Services.**

|  |  |  |
| --- | --- | --- |
| Service | Azure | AWS |
| Compute | Virtual Machines, App Service, Azure Functions, Azure Container Instances, Azure Kubernetes Service | EC2, Elastic Beanstalk, Lambda, ECS, EKS |
| Storage | Blob Storage, File Storage, Queue Storage, Disk Storage, Azure Data Lake Storage | S3, EBS, EFS, Glacier, Storage Gateway, Snowball |
| Database | SQL Database, Cosmos DB, MySQL, PostgreSQL, Azure Database for MariaDB | RDS, DynamoDB, Aurora, Neptune, DocumentDB |
| Networking | Virtual Network, Load Balancer, Application Gateway, VPN Gateway, Azure DNS | VPC, ELB, CloudFront, Direct Connect, Route 53 |
| Identity and Access Management | Azure Active Directory | AWS Identity and Access Management (IAM) |
| AI/ML | Azure Machine Learning, Cognitive Services | Amazon SageMaker, AWS AI/ML services |
| Analytics and Big Data | Azure HDInsight, Data Lake Analytics, Stream Analytics, Power BI | Amazon EMR, Kinesis, Redshift, Athena, QuickSight |
| Internet of Things | Azure IoT Hub, IoT Central | AWS IoT Core, Greengrass |
| Serverless Computing | Azure Functions | AWS Lambda |
| Hybrid Cloud | Azure Arc | AWS Outposts, AWS Hybrid Cloud |

**Some Use cases for when to use which service:**

|  |  |
| --- | --- |
| **Storage Service** | **Use Case and Examples** |
| Azure Blob Storage | - Storing and serving unstructured data like images, videos, and files.  - Hosting static websites.  - Storing backups, logs, and archival data. |
| Azure Files | - Lift-and-shift applications with existing file dependencies.  - Shared file storage for applications.  - Cloud-based file servers. |
| Azure Tables | - Storing large volumes of structured data like IoT telemetry or user data.  - Storing metadata or application configuration. |
| Azure Queues | - Background job processing and task orchestration.  - Inter-process communication and event-driven systems. |
| Azure Disk Storage | - Persistent storage for Azure virtual machines.  - Storing databases and application da |
| Azure Data Lake | - Big data analytics, data exploration, and machine learning.  - Storing and processing large volumes of structured and unstructured data. |

**Azure Storage Account:**

**Blob Storage:** A storage account can have one or more containers to store blobs, which are used to store unstructured data such as images, videos, documents, and backups.

**File Storage:** With a storage account, you can create one or more file shares that can be accessed using the SMB protocol, enabling you to store and share files across multiple machines and platforms.

**Table Storage:** Azure Storage accounts include the capability to store structured data in tables, offering a NoSQL key-value store for high scalability and simple querying.

**Queue** **Storage**: Storage accounts provide message queue functionality, allowing you to decouple components of your application and enable asynchronous communication.

**Disk** **Storage**: Azure Disk Storage provides durable and high-performance block storage for virtual machines, allowing you to attach disks to VMs and store data persistently.

**Redundancy** **Options**: Storage accounts offer different redundancy options to ensure data durability and availability, including locally redundant storage (LRS), zone-redundant storage (ZRS), geo-redundant storage (GRS), and read-access geo-redundant storage (RA-GRS).

**Access** **Control**: Storage accounts have built-in access control mechanisms, allowing you to define fine-grained access policies and manage authentication and authorization for accessing the stored data.

**Security** **and** **Encryption**: Azure Storage accounts support encryption at rest and in transit, ensuring the security of your data. Additionally, you can enable advanced security features like Azure Private Link and virtual network service endpoints.

**Scalability** **and** **Performance**: Azure Storage accounts are highly scalable, allowing you to store massive amounts of data and handle high-demand workloads. The performance can be optimized using features like Azure Premium Blob Storage and Azure Premium Files.

**Monitoring** **and** **Analytics**: Azure Storage accounts provide monitoring and analytics capabilities, allowing you to track storage usage, performance metrics, and generate insights using Azure Monitor, Azure Storage Analytics, and Azure Log Analytics.

**Overview of Storage account Console with some Options:**

**Overview Tab**

Displaying essential information about the storage account

Monitoring key metrics and health status

**Containers Tab (Blob Storage)**

Listing and managing blob containers

Creating new containers

Configuring container properties and access control

**File Shares Tab (File Storage)**

Listing and managing file shares

Creating new file shares

Setting file share properties and permissions

**Tables Tab (Table Storage)**

Listing and managing tables

Creating new tables

Managing table entities and properties

**Queues Tab (Queue Storage)**

Listing and managing queues

Creating new queues

Sending and receiving messages from queues

**Disks Tab (Disk Storage)**

Listing and managing disks

Creating new disks

Attaching and detaching disks to/from virtual machines

**Access Control (IAM) Tab**

Managing role-based access control (RBAC) assignments for the storage account

Adding and removing access to the storage account

**Networking Tab**

Configuring network access and security for the storage account

Setting up virtual network service endpoints

**Firewalls and Virtual Networks Tab**

Configuring network rules and restrictions for the storage account

Restricting access to specific IP addresses or virtual networks

**Data Protection Tab**

Configuring backup and restore settings for the storage account

Setting up geo-redundant storage (GRS) or read-access geo-redundant storage (RA-GRS)

**Monitoring Tab**

Configuring metrics and alerts for the storage account

Viewing and analyzing metrics data

**Configuration Tab**

Managing general configuration settings for the storage account

Configuring replication options and access tiers

**Access Keys Tab**

Viewing and managing access keys for the storage account

Regenerating access keys for enhanced security

**Shared Access Signature (SAS) Tab**

Generating and managing shared access signatures for the storage account

Configuring SAS permissions and expiration.

**Creating Storage Account:**

**Project Details:**

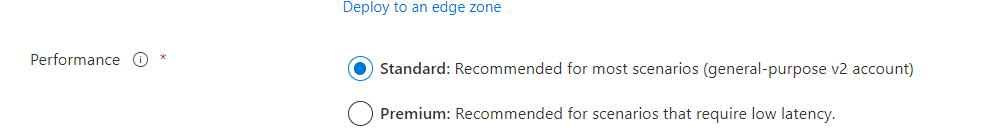
-Subscription & Resource Group

**Instance Details:**

-Account Name:

-Region:

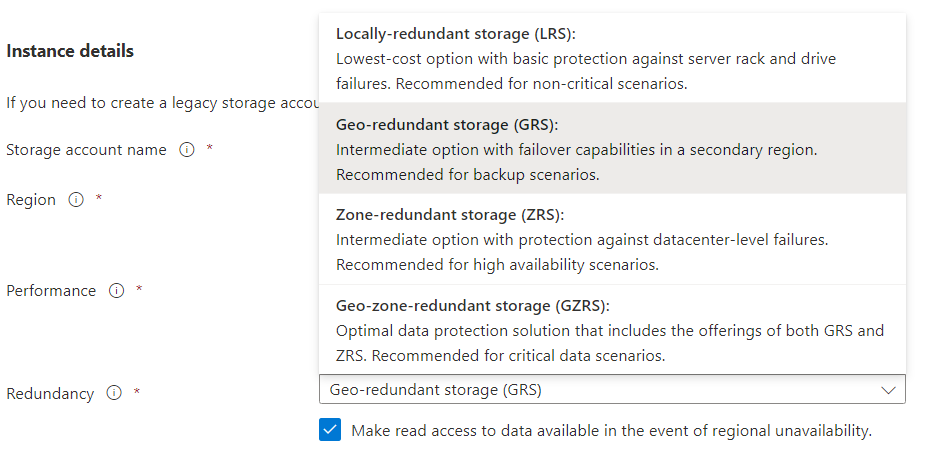
Performance:

****

-Standard (General Purpose (Where not only common use is there not frequent access or Latency))

-Premium (When there is a requirement of Low Latency like, Want to access it fast).

**Redundancy:**

****

**LRS (Locally Redundant Storage):**

Data is replicated synchronously three times within a single data center in a specific region.

Provides high durability and availability within a single data center.

Least expensive replication option.

Suitable for scenarios where cost is a primary consideration and data replication across regions is not required.

Does not provide protection against regional disasters.

**GRS (Geo-Redundant Storage):**

Data is replicated synchronously to a secondary region, located hundreds of miles away from the primary region.

Provides higher durability and availability compared to LRS.

In case of a regional outage, data can be accessed from the secondary region.

Offers a higher cost compared to LRS due to the additional replication and storage resources required in the secondary region.

**ZRS (Zone-Redundant Storage):**

Data is replicated synchronously across multiple availability zones within a single region.

Offers higher durability and availability compared to LRS.

Provides protection against data center failures within a region.

Suitable for scenarios that require high availability and durability within a region, but don't require data replication across regions.

Cost is higher compared to LRS.

**GZRS (Geo-Zone-Redundant Storage):**

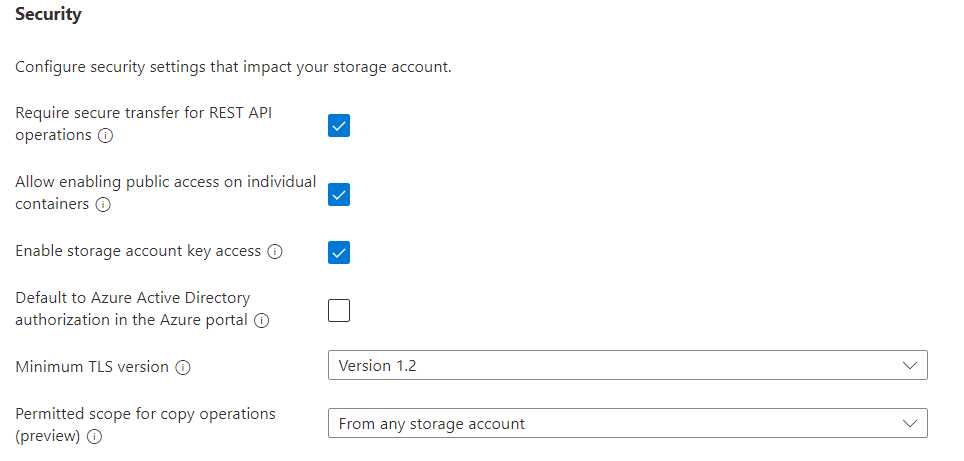
Data is replicated synchronously across multiple availability zones in the primary region and asynchronously to a secondary region.

Provides the highest level of durability and availability.

Offers protection against data center failures within the primary region and regional outages.

Suitable for mission-critical workloads that require maximum data resilience and availability.

Has the highest cost among the replication options due to additional replication and storage resources in both primary and secondary regions.



**REST API:**

it means that all communication between the client and the server must be encrypted using protocols such as HTTPS (HTTP over SSL/TLS) to ensure the confidentiality and integrity of the data being transmitted.

TLS Version (Transport Layer Security): Security version for deciding Client Server Security Rules and Protocols according to Documentation.

**Hierarchical Namespace (HNS)** is a feature available in Azure Data Lake Storage Gen2. It provides a hierarchical file system namespace for organizing and managing data within the storage account. HNS enables a folder-based structure similar to traditional file systems, allowing you to organize data into directories and subdirectories.

Access Protocols:

Enable SFTP

Enables the SSH File Transfer Protocol for your storage account that allows users to access blobs via an SFTP endpoint. Local users need to be created before the SFTP endpoint can be accessed.

Enable network file system v3

Enables the Network File System Protocol for your storage account that allows users to share files across a network. This option must be set during storage account creation

**Blob Storage**

**Allow cross-tenant replication**

Allow object replication to copy blobs to a destination account on a different Azure Active Directory (Azure AD) tenant. Not enabling cross-tenant replication will limit object replication within the same Azure AD tenant.

**Access Tier:**

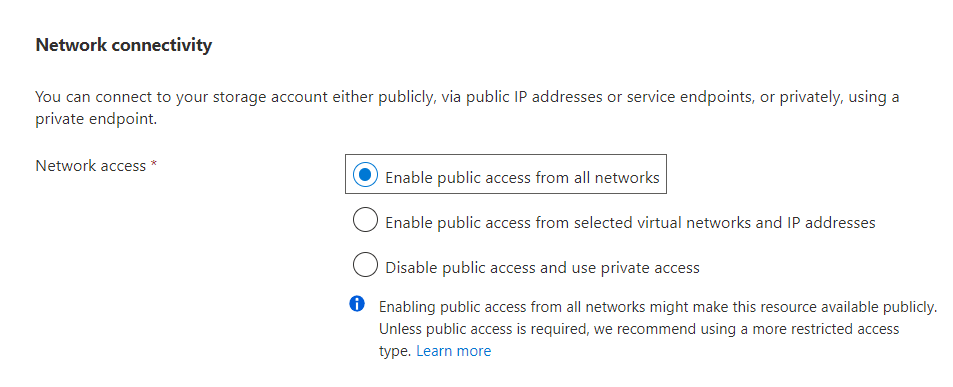
**Hot:** It means we need frequent access to the files which are stored in there.

**Cool:** It means the files which are stored there will not be needed frequently SO the IOPs can be LOW.

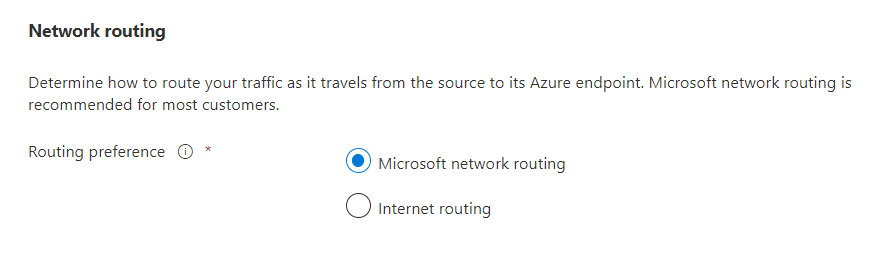
**Azure Flies:** as the name suggests it will enable large file share. SO that large files like up to 4TiB it will allow to store and up to 100Tib it will allow to share.

**Networking:**

**Network Connectivity:**

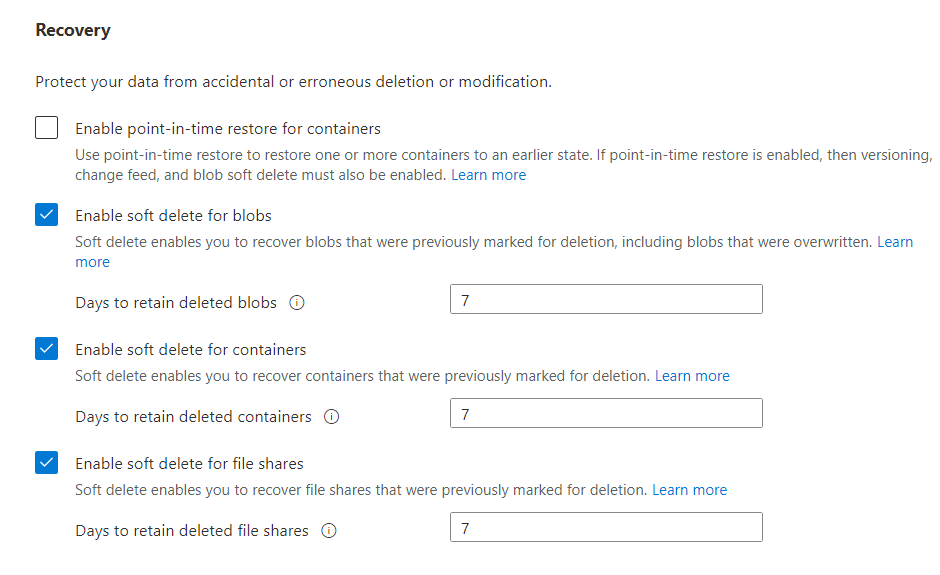


There are options like connect this account publicly, With some private ip addresses or using a private endpoint.



It is for traffic, which traffic we want to allow for accessing the stored files.

**Data Protection:**

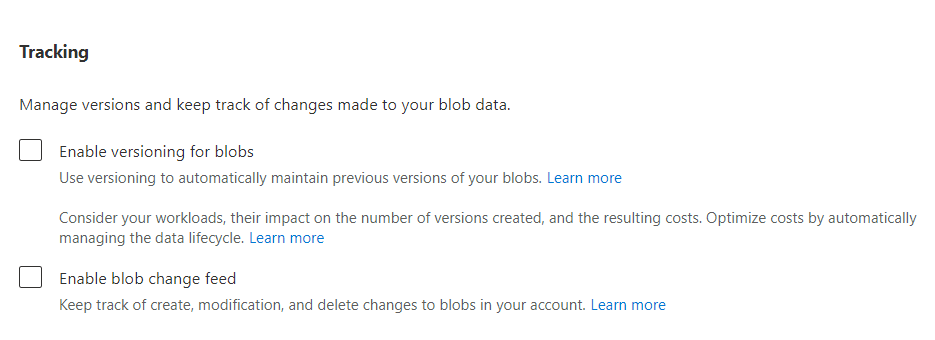


Point-in-time restore provides protection against accidental deletion or corruption by enabling you to restore block blob data to an earlier state.

It helps to recover lost data in accidental situations.

**Soft Delete**: It is like delete marker in AWS, It will not delete any object directly, so that we can recover it if needed. We can set no. of days as buffer before it will be deleted permanently.

It will go for Blobs, Containers and file shares.

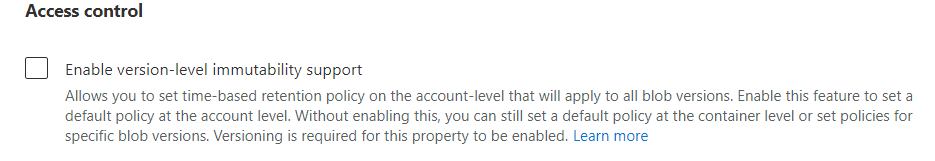


Tracking allows you to make versions of data stored in There.

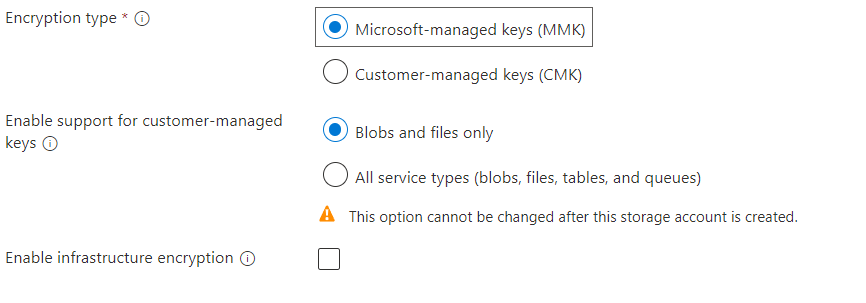
It will have the record as you upload the same files with modifications.

So that we can easily find out which file we need exactly.

**Enable Blob change feed:** it will make activity log of everything we do in Account.



Using this setting we can define retention period for blob storage, during this period that files can’t be modified or deleted.



MMK: Main meaning of this is **MASTER MANAGEMENT KEY** but here it stands for **Microsoft Managed Key.** It will make sure that data which are stored in storage account will be encrypted using MMK so that customer will not have to take care about key and security of data.

There I also option for BYOK (By Your Own Key) or CMK (Customer Managed Key). This will allow user to use their own key for encryption.

**Enable support for CMK:** Here if customer uses CMK for encryption than there is also option for support from Microsoft Azure side.

But it can only select one time, once account is created than it can’t be changed.

This will help user in case of having issues with key like accidental deletion corrupted file.

**Blob Storage:**

Block blobs store text and binary data. Block blobs are made up of blocks of data that can be managed individually. Block blobs can store up to about 190.7 TiB.

Append blobs are made up of blocks like block blobs, but are optimized for append operations. Append blobs are ideal for scenarios such as logging data from virtual machines.

Page blobs store random access files up to 8 TiB in size. Page blobs store virtual hard drive (VHD) files and serve as disks for Azure virtual machines. For more information about page blobs.

Blob storage is same as Containers: