

CSCI 5902 - Fall 23 - Azure Tutorial

Designed under guidance of Dr. Lu Yang

Harmit Narula
©2023, Faculty of Computer Science

Recap

- Introduction to Azure
- Certification Pathways
- Course Outline for tutorials
- Azure Well Architected Framework
- Azure Global Infrastructure

T2 - Azure Storage

Azure Storage - Features

- Azure Storage is Microsoft's cloud storage solution.
- Supports high scalability, availability, durability and security.
- Objects stored can be accessed using variety of protocols.
- Offers libraries for interaction using most common languages/frameworks - .Net, Java, Python, JavaScript, C++, Go, etc.
- Supports scripting for data management and configuration.

Services Available in Azure Storage

- Azure Blob
- Azure Files
- Azure Queues
- Azure Tables

- Azure managed Disks
- Azure NetApp Files

Storage
Account

Azure Blob

What is a BLOB?

Binary Large Object

Binary data stored in the form of single entity in a database system. BLOBs primarily hold multimedia objects.

Azure Blob

- Microsoft's Blob Storage solution for cloud.
- Optimized for storing massive amount of unstructured data.
- `https://<storage_account_name>.blob.core.windows.net/<container_name>/<blob_name>`
- Ideal use:
 - Serving images or documents on browser
 - Storing files for distributed access
 - Streaming video and audio
 - Backup, restore, disaster recovery, archiving

Azure Files

- Helps setup highly available network files share.
- Supports SMB, NFS, REST API for access.
- Multiple VMs can share same files with read and write access.
- `https://<storage_account_name>.file.core.windows.net/<share_name>/<directory_path>/<file_name>`
- Ideal use:
 - Lift and Shift on-premise file share.
 - Storing configuration files.
 - Storing tools and utilities used by a team.

Azure Queue Storage

- Service to store large number of messages.
- Supports each message upto 64 KB.
- A queue can have millions of messages, upto total capacity limit of storage account.
- <https://<storage account>.queue.core.windows.net/<queue>>
- Ideal use:
 - Create a backlog of work to process asynchronously

Azure Table Storage

- Service to store large amount of non-relational structured data.
- Provides key/attribute store with a schemaless design.
- Table storage is cost efficient than traditional SQL for same volume of data.
- `http://<storage account>.table.core.windows.net/<table`
- Ideal use:
 - Store flexible datasets - user data for web applications, address books, device information etc.
 - Datasets that don't require complex joins, foreign keys, stored procedures and can be denormalized for fast access. Provides key/attribute store with a schemaless design.

Azure Managed Disks

- Virtual Hard Disk
- Block level storage volumes managed by Azure and used with Azure virtual machines.
- Available Types - Ultra disks, premium SSD, standard SSD, standard HDD.
- Support 5 9's availability.
- LRS disk has 11 9's durability ZRS disk has 12 9's durability over a year.

**Azure Storage account is unique namespace
within Azure**

What is the SLA of Azure Storage?

Writes upto 99.9%
Reads upto 99.99%

Maximum file size for Azure Storage?

Service version	Maximum block size (via Put Block)	Maximum blob size (via Put Block List)	Maximum blob size via single write operation (via Put Blob)
Version 2019-12-12 and later	4000 MiB	Approximately 190.7 TiB (4000 MiB X 50,000 blocks)	5000 MiB
Version 2016-05-31 through version 2019-07-07	100 MiB	Approximately 4.75 TiB (100 MiB X 50,000 blocks)	256 MiB
Versions prior to 2016-05-31	4 MiB	Approximately 195 GiB (4 MiB X 50,000 blocks)	64 MiB

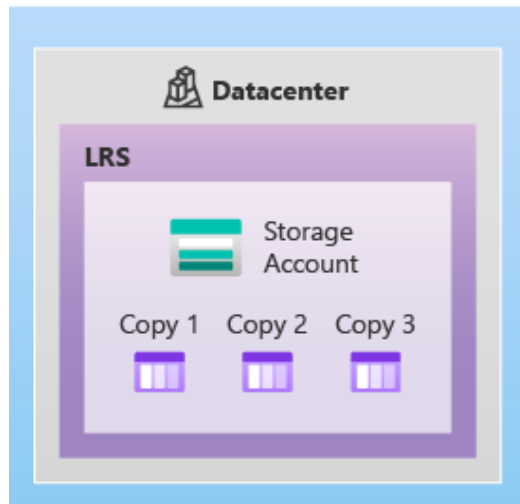
[1]

Redundancy

Storage Redundancy

- Azure Storage account is replicated 3 times in the primary region.
 - Locally Redundant Storage(LRS) - 3 copies of data is stored in a single physical location. Provides 11 9's durability over a year.

Primary region

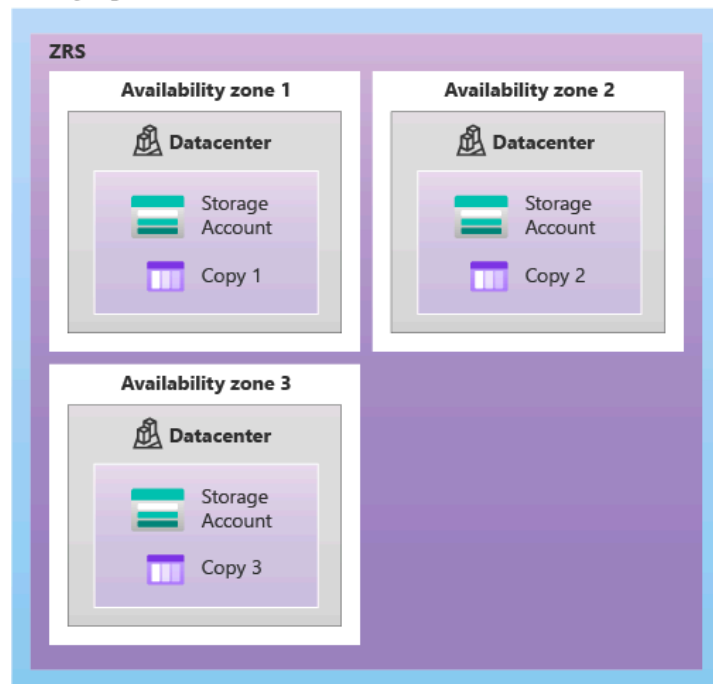


[2]

Storage Redundancy

- Zone Redundant Storage(ZRS) - Data stored across 3 availability zones. Provides 12 9's durability over a year.

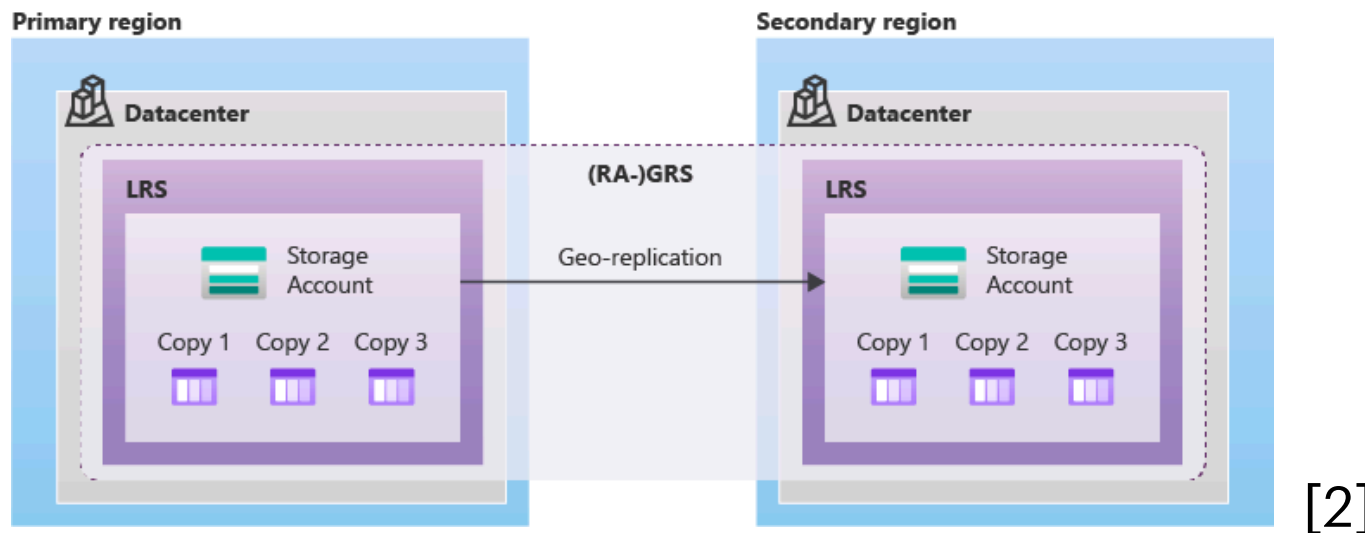
Primary region



[2]

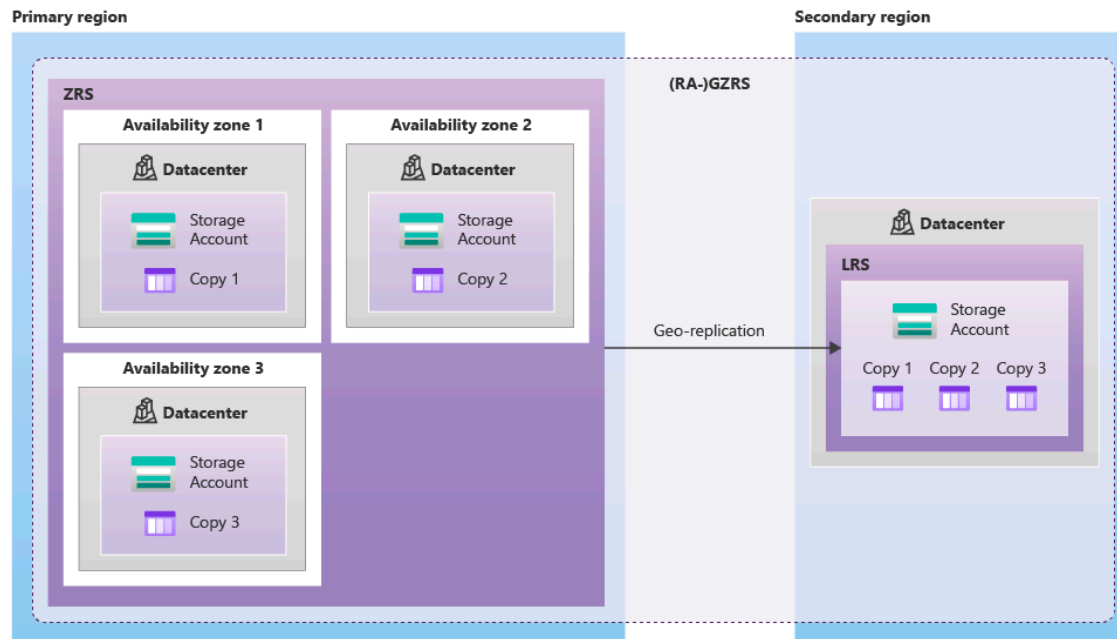
Storage Redundancy

- Geo Redundant Storage(GRS)/Read Access Geo Redundant Storage(RA-GRS) - Data is stored synchronously 3 times within primary region and asynchronously 3 times in secondary region. Provides 16 9's durability over a year.



Storage Redundancy

- Geo Zone Redundant Storage(GZRS) - Data is stored synchronously 3 times within primary region across AZs and asynchronously 3 times in secondary region(LRS). Provides 16 9's durability over a year.



[2]

Summary

Parameter	LRS	ZRS	GRS/RA-GRS	GZRS/RA-GZRS
Percent durability of objects over a given year	at least 99.999999999% (11 9's)	at least 99.999999999% (12 9's)	at least 99.9999999999999% (16 9's)	at least 99.9999999999999% (16 9's)
Availability for read requests	At least 99.9% (99% for Cool or Archive access tiers)	At least 99.9% (99% for Cool access tier)	At least 99.9% (99% for Cool or Archive access tiers) for GRS At least 99.99% (99.9% for Cool or Archive access tiers) for RA-GRS	At least 99.9% (99% for Cool access tier) for GZRS At least 99.99% (99.9% for Cool access tier) for RA-GZRS
Availability for write requests	At least 99.9% (99% for Cool or Archive access tiers)	At least 99.9% (99% for Cool access tier)	At least 99.9% (99% for Cool or Archive access tiers)	At least 99.9% (99% for Cool access tier)
Number of copies of data maintained on separate nodes	Three copies within a single region	Three copies across separate availability zones within a single region	Six copies total, including three in the primary region and three in the secondary region	Six copies total, including three across separate availability zones in the primary region and three locally redundant copies in the secondary region

[2]

How Azure ensures Data Integrity?

Azure regularly verifies data integrity using CRCs

Access Tiers

Access Tiers

- Access tiers help in storing blob data in most cost-effective manner.
 - Hot Tier - Data that is accessed or modified frequently.
 - Cool Tier - Data that is infrequently accessed, min 30 days
 - Cold Tier - Same as cool but min 90 days storage
 - Archive Tier - Offline tier for storing rarely accessed data.
Flexible data latency(hours), min 180 days storage.

**Storage Cost
Decreases**

**Access Cost
Increases**



Lifecycle Policies

Storage Lifecycle Policies

- Lifecycle policies help in adjusting access tier of blobs based on defined set of rules.
- Every rule definition includes
 - Filter Set: Action to certain set of objects within a container/object name.
 - Action Set: Applies tier or delete action to the filtered set of objects.
- Sample rule:
 - Tier blob to cool tier 30 days after last modification
 - Tier blob to archive tier 90 days after last modification
 - Delete blob 2,555 days (seven years) after last modification
 - Delete previous versions 90 days after creation

**How can I interact with
my Storage?**

Storage Interaction Options

- Azure Portal
- Storage Explorer/Browser
- PowerShell
- Azure CLI
- SDKs

**Is my data secure with
Azure?**

- Azure uses service-side encryption(SSE) to automatically encrypt data once it is in cloud. It uses 256 bit AES encryption.
- Azure Storage encryption cannot be disabled.
- Default encryption uses Microsoft managed keys. However, you can also provide customer-managed/customer-provided key for encryption and decryption.
- Customer managed key is stored in Azure Key Vault.

References

[1] <https://learn.microsoft.com/en-us/rest/api/storageservices/understanding-block-blobs--append-blobs--and-page-blobs>

[2] [https://learn.microsoft.com/en-us/azure/storage/common/storage-redundancy?
toc=%2Fazure%2Fstorage%2Fblobs%2Ftoc.json&bc=%2Fazure%2Fstorage%2Fblobs%2Fbreadcrumb%2Ftoc.json](https://learn.microsoft.com/en-us/azure/storage/common/storage-redundancy?toc=%2Fazure%2Fstorage%2Fblobs%2Ftoc.json&bc=%2Fazure%2Fstorage%2Fblobs%2Fbreadcrumb%2Ftoc.json)