

# Cloud Understanding

Dr. A. Suresh, INSOFE

# Azure Storage services

<a href="#"><u>Azure Disk Storage</u></a>	High-performance, durable block storage for Azure Virtual Machines
<a href="#"><u>Azure Blob Storage</u></a>	Massively scalable and secure object storage for cloud-native workloads, archives, data lakes, high-performance computing, and machine learning
<a href="#"><u>Azure Data Lake Storage</u></a>	Massively scalable and secure data lake for your high-performance analytics workloads
<a href="#"><u>Azure Files</u></a>	Simple, secure, and serverless enterprise-grade cloud file shares
<a href="#"><u>Azure NetApp Files</u></a>	Enterprise file storage, powered by NetApp
<a href="#"><u>Data Box</u></a>	Appliances and solutions for data transfer to Azure and edge compute
<a href="#"><u>Azure confidential ledger</u></a>	Store unstructured data that is completely tamper-proof and can be cryptographically verified

# Disk Storage Contd.,

	Ultra disk	Premium SSD	Standard SSD	Standard HDD
<b>Disk type</b>	SSD	SSD	SSD	HDD
<b>Scenario</b>	IO-intensive workloads such as <a href="#">SAP HANA</a> , top tier databases (for example, SQL, Oracle), and other transaction-heavy workloads.	Production and performance sensitive workloads	Web servers, lightly used enterprise applications and dev/test	Backup, non-critical, infrequent access
<b>Max disk size</b>	65,536 gibibyte (GiB)	32,767 GiB	32,767 GiB	32,767 GiB
<b>Max throughput</b>	4,000 MB/s	900 MB/s	750 MB/s	500 MB/s
<b>Max IOPS</b>	160,000	20,000	6,000	2,000

# Blob Storage

---

Blob: Binary Large Object

Store any file (Generally unstructured data)

Audio, video, binary, executables, source code or anything.

Storage Tiers:

- ☐ Hot : Very frequently accessed data
- ☐ Cool : Infrequently accessed data (lower availability, high durability)
- ☐ Archive : very rarely accessed data

# Blob Storage Contd.,

---

## Block blob

- ✓ Targeted at streaming workloads( eg. Files read from beginning to end like media files)
- ✓ Each blob consists of a sequence of blocks
- ✓ Each block is identified by a Block ID
- ✓ Each block can be a maximum of 64MB in size
- ✓ Size limit 200GB per blob

# Blob Storage Contd.,

---

## Page blob

- ✓ Targeted at random read/write workloads (eg. Backing storage for the VHDs used in Azure VMs)
- ✓ Each blob consists of an array of pages
- ✓ Each page is identified by its offset from the start of the blob
- ✓ Size limit 1 TB per blob

# Blob Storage Contd.,

---

## Append blobs

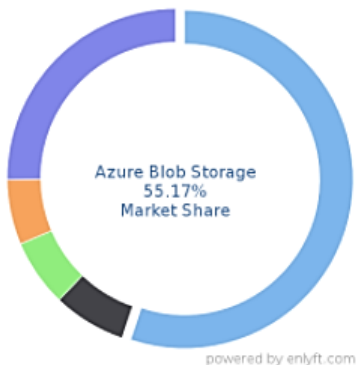
- ✓ Append blobs are made up of blocks like block blobs
- ✓ But they are optimized for append operations
- ✓ Frequently used for logging information
- ✓ A single append blob can be up to 195 GB

# Blob Storage Contd.,

Azure Blob Storage  
Data Storage Management

47,039

Companies using Azure Blob Storage



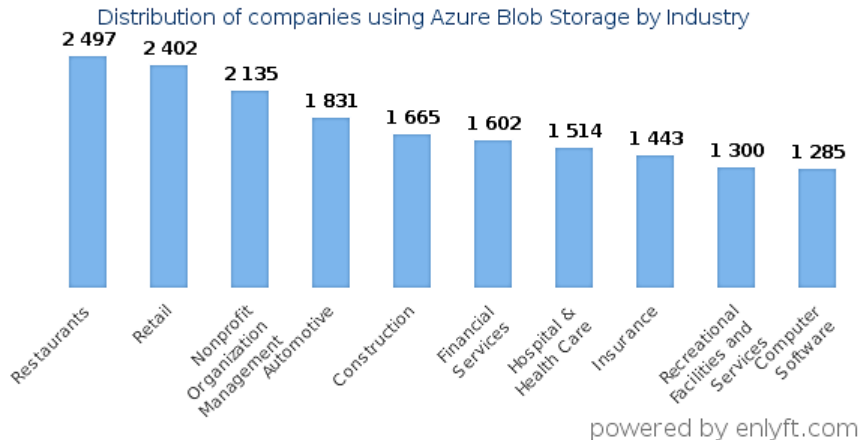
Data Storage Management

Azure Blob Storage (55.17%)  
IBM Tivoli Storage Manager (7.00%)  
Peak IO (6.44%)  
WinSCP (6.21%)  
[View other alternative products](#)



# Blob Storage Contd.,

Looking at Azure Blob Storage customers by industry, we find that Restaurants (5%) and Retail (5%) are the largest segments.

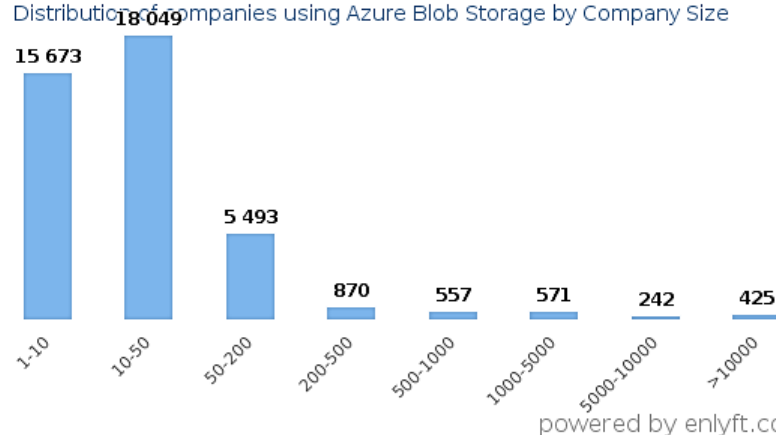


# Blob Storage Contd.,

Distribution of companies that use Azure Blob Storage based on company size (Employees)

Of all the customers that are using Azure Blob Storage, a majority (75%) are small (<50 employees), 7% are large (>1000 employees) and 19% are medium-sized.

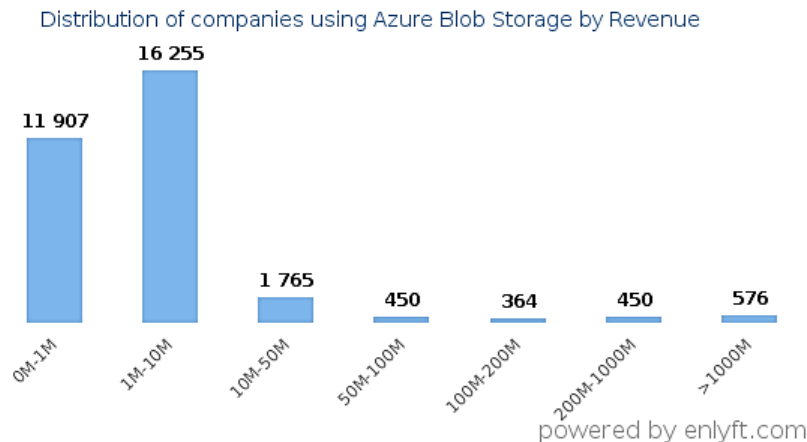
Distribution of companies using Azure Blob Storage by Company Size



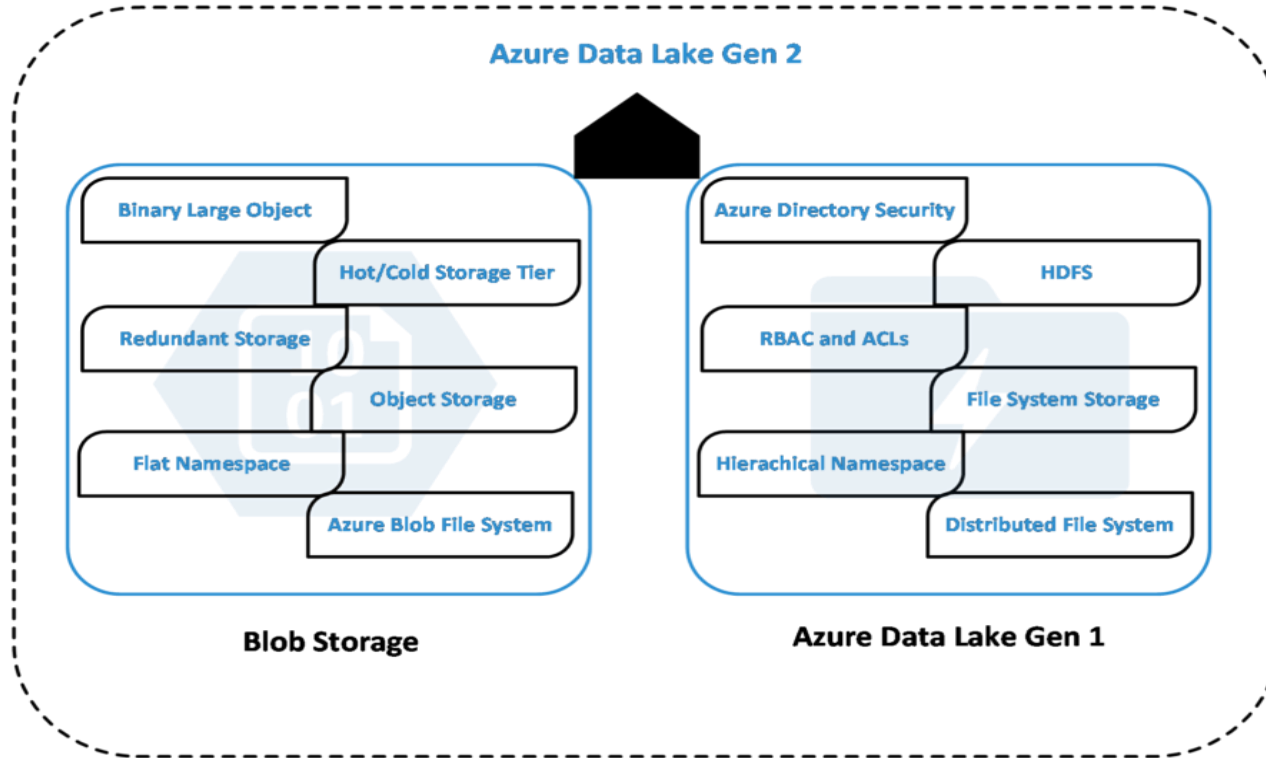
# Blob Storage Contd.,

Distribution of companies that use Azure Blob Storage based on company size (Revenue)

Of all the customers that are using Azure Blob Storage, a majority (77%) are small (<\$50M), 12% are large (>\$1000M) and 6% are medium-sized.



# Data Lake Gen2



# Advantage's of Gen2

---

## Blob storage

- ✓ Storage types (Hot, Cold , Archive)
- ✓ Economical
- ✓ Availability

## Data Lake

- ✓ Big data storage
- ✓ Hierarchical File system
- ✓ POSIX policy

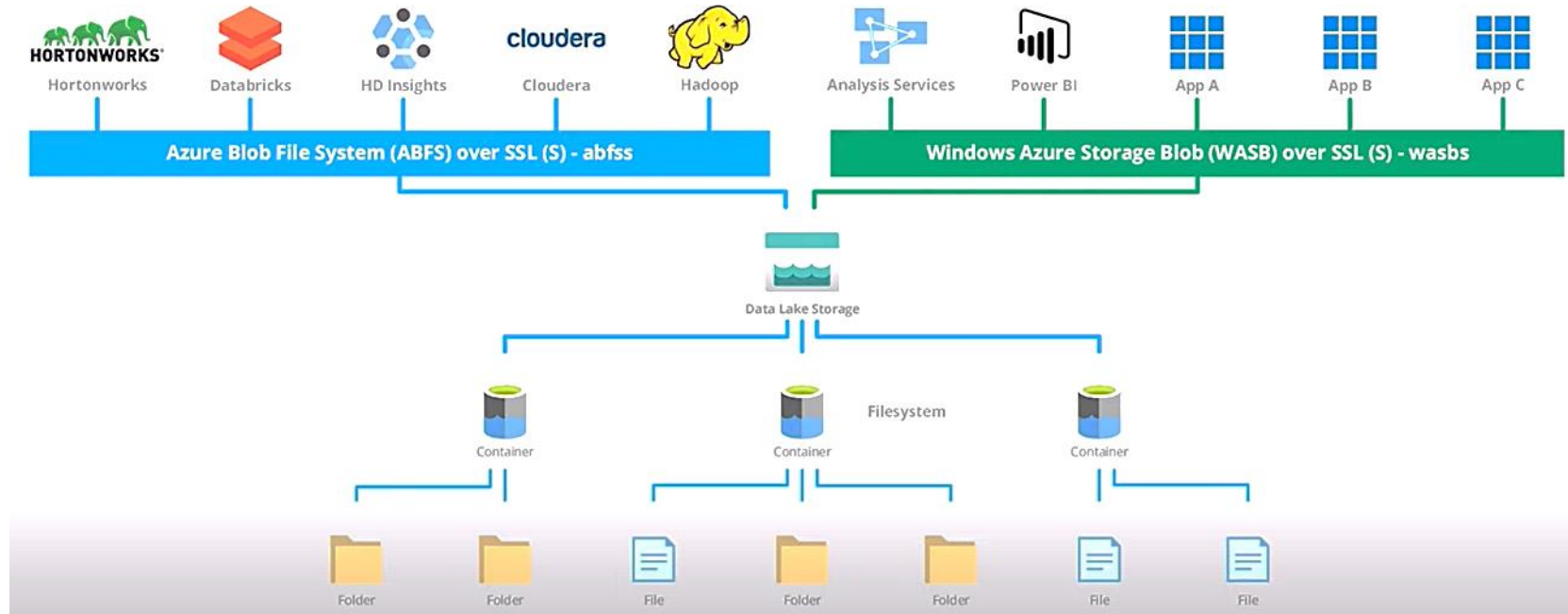
# Data Lake1 vs Gen2

Azure Data Lake Gen1	Azure Data Lake Gen2
Azure Data Lake Gen 1 is file system storage in which data is distributed in blocks in a hierarchical file system.	Azure Data Lake Gen 2 contains both file system storage for performance & security and object storage for scalability.
Hot/Cold storage tier not supported	Supports Hot/Cold Storage tier
Redundant Storage not supported	Supports Redundant Storage
Azure data lake Analytics support available	Azure Data Lake analytics support is not available (till date 2nd July 2019).

# Blob vs Gen2

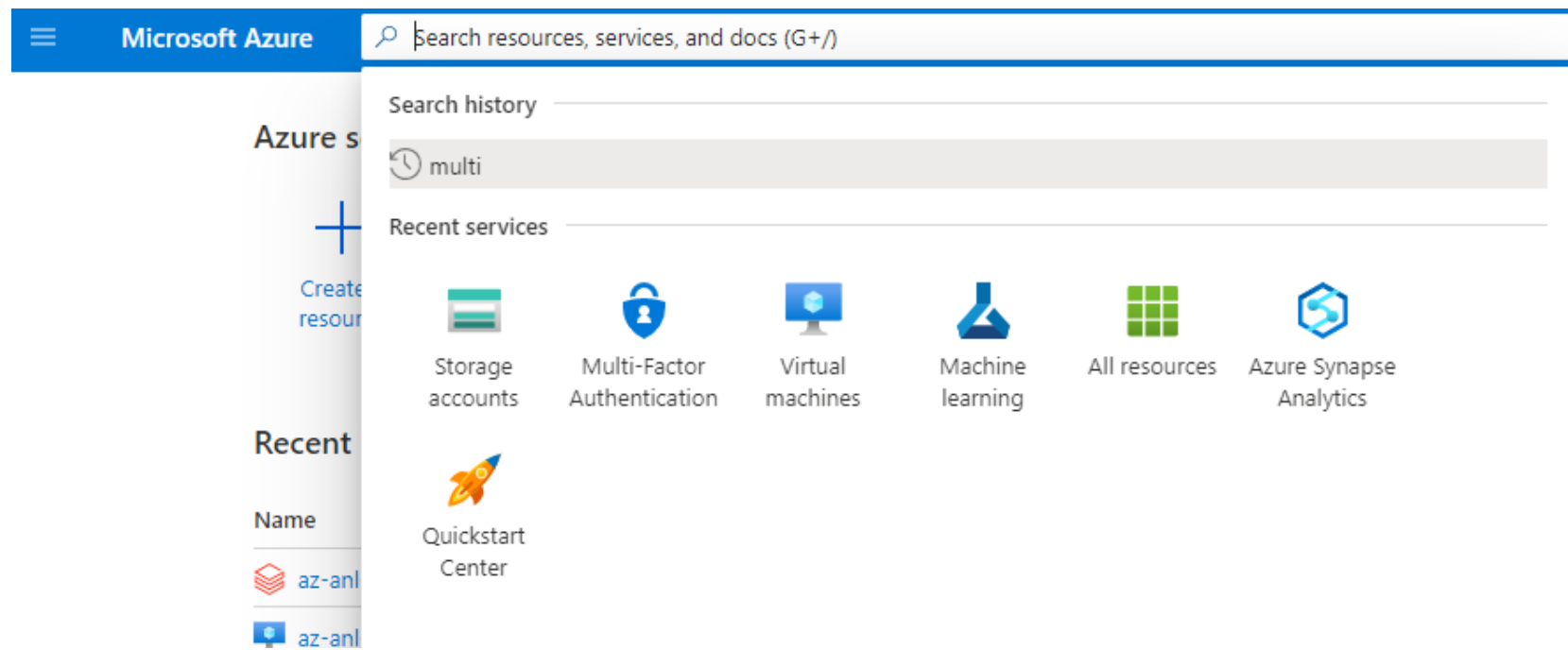
Concept	Top Level Organization	Lower Level Organization	Data Container
<b>Blobs - General purpose object storage</b>	Container	Virtual directory (SDK only - does not provide atomic manipulation)	Blob
<b>Azure Data Lake Storage Gen2 - Analytics Storage</b>	Container	Directory	File

# Data Lake Gen2





# Create Data lake Gen2



# Create Data lake Gen2 Contd.,

Microsoft Azure

Search resources, services, and docs (G+)

BIGA1P7K@fundaciongr...  
FUNDACIONGRUPOSOCIAL.CO [...]

Home >

## Storage accounts

fundaciongruposocial.co (fundaciongruposocialco.onmicrosoft.com)

[+ Create](#) [Manage view](#) [Refresh](#) [Export to CSV](#) [Open query](#) [Assign tags](#) [Delete](#) [Feedback](#)

[Subscription == all](#) [Resource group == all](#) [Location == all](#) [+ Add filter](#)

Showing 1 to 5 of 5 records. No grouping List view

<input type="checkbox"/> Name ↑↓	Type ↑↓	Kind ↑↓	Resource group ↑↓	Location ↑↓	Subscription ↑↓
<input type="checkbox"/> azanldesdlsdave2000	Storage account	StorageV2	az-trv-rg-dbstorage-e2-000	East US 2	az-devqa-sub-analitica
<input type="checkbox"/> azanldesdlsyndave2000	Storage account	StorageV2	az-anl-dev-rg-syn-e2-000	East US 2	az-devqa-sub-analitica
<input type="checkbox"/> azanldqadlsdave2000	Storage account	StorageV2	az-anl-qa-rg-dbstorage-e2-0...	East US 2	az-devqa-sub-analitica
<input type="checkbox"/> azanldqadlsyndave2000	Storage account	StorageV2	az-anl-qa-rg-syn-e2-000	East US 2	az-devqa-sub-analitica
<input type="checkbox"/> sqlvaxkyxqiwpypyrw	Storage account	StorageV2	az-anl-dev-rg-syn-e2-000	East US 2	az-devqa-sub-analitica

# Create Data lake Gen2 Contd.,

## Create a storage account

Basics Advanced Networking Data protection Tags Review + create

Azure Storage is a Microsoft-managed service providing cloud storage that is highly available, secure, durable, scalable, and redundant. Azure Storage includes Azure Blobs (objects), Azure Data Lake Storage Gen2, Azure Files, Azure Queues, and Azure Tables. The cost of your storage account depends on the usage and the options you choose below. [Learn more about Azure storage accounts](#)

### Project details

Select the subscription in which to create the new storage account. Choose a new or existing resource group to organize and manage your storage account together with other resources.

Subscription \* az-devqa-sub-analitica

Resource group \* az-anl-dev-rg-adf-e2-000

[Create new](#)

## Create a storage account

Basics Advanced Networking Data protection Tags Review + create

### Instance details

If you need to create a legacy storage account type, please click [here](#).

Storage account name \* azureadlsgen2demo

Region \* (US) East US

Performance \* ☒ Standard: Recommended for most scenarios (general-purpose v2 account)

☐ Premium: Recommended for scenarios that require low latency.

Redundancy \* Geo-redundant storage (GRS)

☒ Make read access to data available in the event of regional unavailability.

Review + create

< Previous

Next : Advanced >

# Create Data lake Gen2 Contd.,

## Create a storage account ...

Basics Advanced Networking Data protection Tags Review + create

Resource group \*

az-anl-dev-rg-adf-e2-000

[Create new](#)

### Instance details

If you need to create a legacy storage account

Storage account name ⓘ \*

Region ⓘ \*

A resource group is a container that holds related resources for an Azure solution.

Name \*

OK

Cancel

# Create Data lake Gen2 Contd.,

Basics Advanced Networking Data protection Tags Review + create

## Instance details

If you need to create a legacy storage account type, please click [here](#).

Storage account name ⓘ \*

azureadlsgen2demo

Region ⓘ \*

(US) East US

Performance ⓘ \*



**Standard:** Recommended for most scenarios (general-purpose v2 account)



**Premium:** Recommended for scenarios that require low latency.

Review + create

< Previous

Next : Advanced >

(Europe) Sweden Central

(Europe) Switzerland North

(Europe) UK South

(Europe) West Europe

(Middle East) UAE North

(South America) Brazil South

(US) Central US

(US) East US

(US) East US 2

(US) North Central US

(US) South Central US

(US) West US

(US) West US 2

(US) West US 3

# Create Data lake Gen2 Contd.,

## Create a storage account ...

Basics Advanced Networking Data protection Tags Review + create

### Instance details

If you need to create a legacy storage account, select **Legacy storage account**.

Storage account name ⓘ \*

Region ⓘ \*

Performance ⓘ \*

Redundancy ⓘ \*

#### Locally-redundant storage (LRS):

Lowest-cost option with basic protection against server rack and drive failures. Recommended for non-critical scenarios.

#### Geo-redundant storage (GRS):

Intermediate option with failover capabilities in a secondary region. Recommended for backup scenarios.

#### Zone-redundant storage (ZRS):

Intermediate option with protection against datacenter-level failures. Recommended for high availability scenarios.

#### Geo-zone-redundant storage (GZRS):

Optimal data protection solution that includes the offerings of both GRS and ZRS. Recommended for critical data scenarios.

Geo-redundant storage (GRS) ▼

☒ Make read access to data available in the event of regional unavailability.

Basics Advanced Networking Data protection Tags Review + create

### Data Lake Storage Gen2

The Data Lake Storage Gen2 hierarchical namespace accelerates big data analytics workloads and enables file-level access control lists (ACLs). [Learn more](#)

Enable hierarchical namespace ☒

### SSH File Transfer Protocol (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. [Learn more](#)

Enable SFTP ⓘ ☐

ⓘ You must opt-in on a per subscription basis to enable SFTP for hierarchical namespace accounts.

### Blob storage

Enable network file system v3 ⓘ ☐

Allow cross-tenant replication ⓘ ☐

ⓘ Cross-tenant replication and hierarchical namespace cannot be enabled simultaneously.

Access tier ⓘ

☒ Hot: Frequently accessed data and day-to-day usage scenarios

☐ Cool: Infrequently accessed data and backup scenarios

# Create Data lake Gen2 Contd.,

Scenario	LRS	ZRS	GRS	RA-GRS
Node unavailability within a data center	Yes	Yes	Yes	Yes
An entire data center (zonal or non-zonal) becomes unavailable	No	Yes	Yes	Yes
A region-wide outage	No	No	Yes	Yes
Read access to your data (in a remote, geo-replicated region) in the event of region-wide unavailability	No	No	No	Yes
Designed to provide _ durability of objects over a given year	at least 99.9999999999% (11 9's)	at least 99.9999999999% (12 9's)	at least 99.99999999999999% (16 9's)	at least 99.99999999999999% (16 9's)
Available in _ storage account types	GPv1, GPv2, Blob	GPv2	GPv1, GPv2, Blob	GPv1, GPv2, Blob

# Create Data lake Gen2 Contd.,

Basics   Advanced   Networking   Data protection   Tags   Review + create

### Network connectivity


You can connect to your storage account either publicly, via public IP addresses or service endpoints, or privately, using a private endpoint.

Connectivity method \*

☒ Public endpoint (all networks)

☐ Public endpoint (selected networks)

☐ Private endpoint

 All networks will be able to access this storage account. We recommend using Private endpoint for accessing this resource privately from your network. [Learn more](#)

### Network routing

Determine how to route your traffic as it travels from the source to its Azure endpoint. Microsoft network routing is recommended for most customers.

Routing preference ⓘ \*

☒ Microsoft network routing

☐ Internet routing

[Review + create](#)   [< Previous](#)   [Next : Data protection >](#)



# Create Data lake Gen2 Contd.,

Basics Advanced Networking Data protection Tags Review + create

## Recovery

Protect your data from accidental or erroneous deletion or modification.

- ☐ Enable point-in-time restore for containers  
Use point-in-time restore to restore one or more containers to an earlier state. If point-in-time restore is enabled, then vers change feed, and blob soft delete must also be enabled. [Learn more](#)

**i** Point-in-time restore and hierarchical namespace cannot be enabled simultaneously.

- ☒ Enable soft delete for blobs  
Soft delete enables you to recover blobs and directories that were previously marked for deletion. [Learn more](#)

Days to retain deleted blobs ⓘ

- ☒ Enable soft delete for containers  
Soft delete enables you to recover containers that were previously marked for deletion. [Learn more](#)

Days to retain deleted containers ⓘ

- ☒ Enable soft delete for file shares  
Soft delete enables you to recover file shares that were previously marked for deletion. [Learn more](#)

Days to retain deleted file shares ⓘ

Basics Advanced Networking Data protection Tags Review + create

Soft delete enables you to recover the shares that were previously marked for deletion. [Learn more](#)

Days to retain deleted file shares ⓘ

## Tracking

Manage versions and keep track of changes made to your blob data.

- ☐ Enable versioning for blobs  
Use versioning to automatically maintain previous versions of your blobs for recovery and restoration. [Learn more](#)

**i** Versioning and hierarchical namespace cannot be enabled simultaneously.

- ☐ Enable blob change feed  
Keep track of create, modification, and delete changes to blobs in your account. [Learn more](#)

**i** Blob change feed and hierarchical namespace cannot be enabled simultaneously.

## Access control

- ☐ Enable version-level immutability support  
Allows you to set time-based retention policy on the account-level that will apply to all blob versions. Enable this feature to set default policy at the account level. Without enabling this, you can still set a default policy at the container level or set policies for specific blob versions. Versioning is required for this property to be enabled. [Learn more](#)

[Review + create](#)

[< Previous](#)

[Next : Tags >](#)

# Create Data lake Gen2 Contd.,

Basics Advanced Networking Data protection **Tags** Review + create

Tags are name/value pairs that enable you to categorize resources and view consolidated billing by applying the same tag to multiple resources and resource groups. [Learn more about tags](#)

Note that if you create tags and then change resource settings on other tabs, your tags will be automatically updated.

Name	Value	Resource
<input type="text"/>	Ambiente	<input type="text"/>
	Dominio o entidad	
	Impacto Empresarial	
	Proyecto	
	Unidad de negocio	
	Dominio o entidad	
	application	
	databricks-environment	

## Create a storage account ...

✓ Validation passed

Basics Advanced Networking Data protection Tags **Review + create**

Point-in-time restore	Disabled
Blob soft delete	Enabled
Blob retainment period in days	7
Container soft delete	Enabled
Container retainment period in days	7
File share soft delete	Enabled
File share retainment period in days	7
Versioning	Disabled
Blob change feed	Disabled
Version-level immutability support	Disabled

Create

< Previous

Next >

[Download a template for automation](#)

# Create Data lake Gen2 Contd.,

The screenshot displays the Microsoft Azure portal interface for a resource group named 'myResourceGroup'. The left sidebar shows the navigation menu with options like Overview, Activity log, Access control (IAM), Tags, Events, Settings, Quickstart, Deployments, Policies, Properties, Locks, Export template, Cost Management, Cost analysis, and Cost alerts. The main content area shows the resource group details, including the subscription (mysubscription), subscription ID (01234567-89ab-cdef-0123-456789abcdef), and tags. The 'Export template' button is highlighted in the top right. Below the details, there is a table of resources with columns for Name, Type, and Location. The 'mystoregwkf52ihhxyi' Storage account is selected.

Name	Type	Location
exampleplan	App Service plan	Central US
mystoregwkf52ihhxyi	Storage account	Central US
mywebappgwkf52ihhxyi	App Service	Central US

# Create Data lake Gen2 Contd.,

The screenshot shows the Microsoft Azure portal interface. At the top, the header includes the 'Microsoft Azure' logo, a search bar, and user information for 'BIGA1P7K@fundaciongr...' with the organization 'FUNDACIONGRUPOSOCIALCO'. The main content area displays the 'Overview' page for a deployment named 'azureadlsgen2demo\_1640440206371'. A notification banner at the top right states '\*\*\* Deployment in progress...' with a close button. Below this, a feedback message says 'We'd love your feedback!'. The central status message reads '\*\*\* Deployment is in progress'. Underneath, there is a link for 'Deployment details (Download)'. At the bottom, a table with columns 'Resource', 'Type', 'Status', and 'Operation details' is shown, containing the text 'No results.'.

Microsoft Azure Search resources, services, and docs (G+)

Home >

azureadlsgen2demo\_1640440206371 | Overview

Deployment

Search (Ctrl+/) Delete Cancel Redeploy Refresh

Overview

Inputs

Outputs

Template

\*\*\* Deployment in progress...  
Deployment to resource group 'az-anl-dev-rg-adf-e2-000' is in progress.

We'd love your feedback! →

\*\*\* Deployment is in progress

Deployment details (Download)

Resource	Type	Status	Operation details
No results.			

# Create Data lake Gen2 Contd.,

Microsoft Azure Search resources, services, and docs (G+/)

Home > azureadlsgen2demo

azureadlsgen2demo | Containers Storage account

Search (Ctrl+/)

Overview  
Activity log  
Tags  
Diagnose and solve problems  
Access Control (IAM)  
Data migration  
Events  
Storage browser (preview)

Data storage  
Containers  
File shares  
Queues

+ Container Change access level Restore containers Refresh Delete

Search containers by prefix

Name	Last modified	Public access level
<input type="checkbox"/> \$logs	12/25/2021, 7:20:57 PM	Private

**New container**

Name \*  
adlsgen2folder ✓

Public access level ⓘ  
Private (no anonymous access) ▼

- Private (no anonymous access)
- Blob (anonymous read access for blobs only)
- Container (anonymous read access for containers and blobs)

Create Discard

# Create Data lake Gen2 Contd.,

Home > azureadlsgen2demo >

**adlsgen2folder** ...  
Container

Search (Ctrl+/) «

Upload Add Directory Refresh Rename Delete Change tier Acquire lease Break lease

**Authentication method:** Access key (Switch to Azure AD User Account)  
**Location:** adlsgen2folder

Search blobs by prefix (case-sensitive)

Name	Modified	Access tier	Archive status	Blob type	Size
No results					

Settings

- Shared access tokens
- Manage ACL
- Access policy
- Properties
- Metadata

select-all-resource....png

**Upload blob** adlsgen2folder/

Files ⓘ  
Select a file

☐ Overwrite if files already exist

Advanced

**Open**

This ... > Windows OS ...

Search Windows OS SSD (C:)

Organize New folder

Name	Date modified
SGetCurrent	6/4/2021 11:37 AM
SWINDOWS.-BT	10/19/2021 10:30 AM
SWinREAgent	12/18/2021 2:40 PM
DF_Files	3/19/2021 12:28 AM
...	...

File name: All Files

Open Cancel

Show all

# Create Data lake Gen2 Contd.,

Home > azureadlsgen2demo >



Search (Ctrl+/)

Overview

Diagnose and solve problems

Access Control (IAM)

Settings

Shared access tokens

Manage ACL

Access policy

Properties

Metadata

Upload Add Directory Refresh Rename Delete Change tier Acquire lease

Authentication method: Access key (Switch to Azure AD User Account)

Location: adlsgen2folder

Search blobs by prefix (case-sensitive)

	Name	Modified	Access tier	Archive status	Blob type
<input type="checkbox"/>	Oracle_Academy_rgb...	12/25/2021, 7:29:01 ...	Hot (Inferred)		Block blob

## Upload blob

adlsgen2folder/

Files

Select a file

☐ Overwrite if files already exist

Advanced

Upload

Current uploads

Dismiss: Completed All

Oracle\_Academy\_rgb.png 4 KiB / 4 KiB ...

# Create Data lake Gen2 Contd.,

Home > adlsgen2folder >



Search (Ctrl+/)

Overview

Diagnose and solve problems

Access Control (IAM)

Settings

Shared access tokens

Manage ACL

Access policy

Properties

Metadata

Upload + Add Directory ...

Authentication method: Access key ([Switch to Azure AD User Account](#))

Location: adlsgen2folder

Search blobs by prefix (case-...

Show deleted objects

Name

☒ Oracle\_Academy\_rgb.png

## Manage ACL

container: adlsgen2folder (storage account: azureadlsgen2demo)

Set and manage permissions for:  
/Oracle\_Academy\_rgb.png

[Learn more about access control lists \(ACLs\)](#)

### Access permissions

+ Add principal + Add mask

Security principal	Read	Write	Execute	
Owner: \$superuser	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<a href="#">Edit</a>
Owning group: \$superuser	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<a href="#">Edit</a>
Other	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	

**i** Read and write permissions will only work for a security principal if the security principal also has execute permissions on all parent directories, including the container (root directory).

Save

Discard



---

*Q&A ?*

*Thank  
You*

