

## Azure Databricks Assignment - 5

### Mitushi Vishwakarma

Week - 4.

Azure Databricks  
Day - 5

ADLS

Azure Data Lake Storage

→ Create a Data Lake Storage Gen 1.

Create resource > Storage > DL Storage Gen 1

ADLS

Data Lake Storage Gen 1

Azure storage  
~~Blob containers~~

Data Lake Storage  
Gen 2

Gen 1 is not in service for new users.

→ creating normal Storage account.

In order to work in Gen 2.

we need to enable hierarchical namespace.

It enables file and directory semantics, ACLs  
access control lists

9. Hierarchical namespace is disabled  
↳ Azure Blob Storage account.

22 - 11 - 2023

22

NOVEMBER

WK 47 • DAY 326-039

WEDNESDAY

2023

DECEMBER 2023

WK	M	T	W	T	F	S	S
48	1	2	3				
49	4	5	6	7	8	9	10
50	11	12	13	14	15	16	17
51	18	19	20	21	22	23	24
52	25	26	27	28	29	30	31

If enabled, to perform analytics on data  
↳ Azure Data Lake Storage Gen 2.

→ In azure blob storage (hierarchical disabled)  
we can store large amounts of  
unstructured data in flat namespace  
within blob containers.

→ accessible by HTTPS or HTTP

→ ADLS Gen 2 builds on blob storage  
and optimizes I/O of high volume data  
by using hierarchical namespace

→ ~~Data~~ Organized blob data into  
directories and metadata about it.

Four stages for processing big data  
solutions.

Ingest →

→ it identifies the technology and  
processes that are used to acquire the  
source data.

→ This data comes from files, logs,  
other unstructured data that must be  
put into data lake.

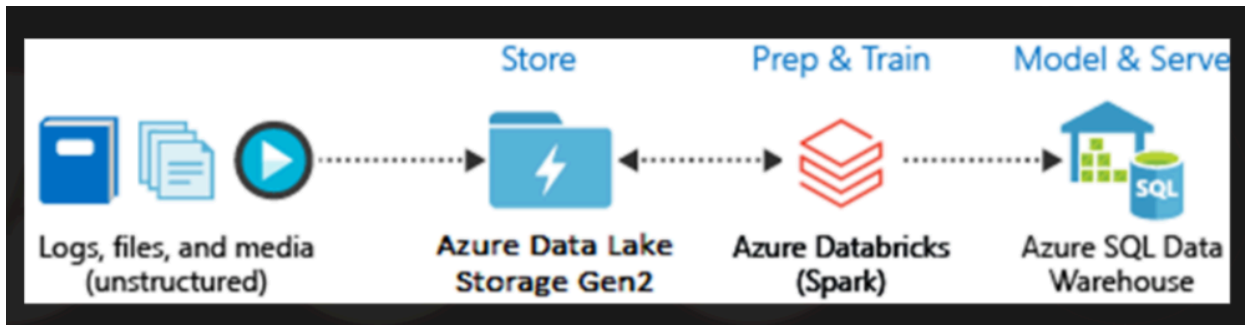
45	8	14	15	16	17	18	19
46	13	14	15	16	17	18	19
47	20	21	22	23	24	25	26
48	27	28	29	30			

NOVEMBER  
WK 47 • DAY 325-040  
TUESDAY

- 11 - 2023

21





→ Azure Synapse Analytics is appropriate technology to use.

### Store

→ This phase where ingested data should be placed.

→ ADLS Gen2 is secure solution.

### Prep & Train

→ This phase identifies technologies that are used to perform data preparation.

→ Azure Synapse Analytics, Azure Databricks, are used in this phase.

### Model & Serve

→ this phase involves the tech. that will present data to user.

→ MS Power BI,

20 - 11 - 2023

20

NOVEMBER

WK 47 • DAY 324-041

MONDAY

2023

NOVEMBER 2023  
S M T W T F S S  
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31

## Created ADLS Gen2 :

The screenshot shows the 'Create a storage account' page in the Microsoft Azure portal. The 'Advanced' tab is selected, showing settings for 'Permitted scope for copy operations' (set to 'From any storage account'), 'Hierarchical Namespace' (enabled), and 'Access protocols' (SFTP is disabled). The 'Review' button is visible at the bottom.

Home > Storage accounts >

### Create a storage account

Basics **Advanced** Networking Data protection Encryption Tags Review

Permitted scope for copy operations (preview) From any storage account

**Hierarchical Namespace**

Hierarchical namespace, complemented by Data Lake Storage Gen2 endpoint, enables file and directory semantics, accelerates big data analytics workloads, and enables access control lists (ACLs) [Learn more](#)

Enable hierarchical namespace ☒

**Access protocols**

Blob and Data Lake Gen2 endpoints are provisioned by default [Learn more](#)

Enable SFTP ☐

**Review** < Previous Next : Networking > [Give feedback](#)

## Enabled Hierarchical namespace :

The screenshot shows the 'adlshexa1046' storage account overview page in the Microsoft Azure portal. The 'Overview' tab is selected, displaying various properties and settings for the storage account.

Home > adlshexa1046\_170806008194 | Overview >

### adlshexa1046

Storage account

Search

Location : centralindia

Primary/Secondary Location : Primary: Central India, Secondary: South India

Subscription (move) : [Azure subscription 1](#)

Subscription ID : 984f097c-963c-4eb6-a20d-839457ae9f08

Disk state : Primary: Available, Secondary: Available

Replication : Read-access geo-redundant storage (RA-GRS)

Account kind : StorageV2 (general purpose v2)

Provisioning state : Succeeded

Created : 2/16/2024, 12:16:53 PM

Tags (edit) : Add tags

**Properties** Monitoring Capabilities (5) Recommendations (0) Tutorials Tools + SDKs

**Data Lake Storage**

Hierarchical namespace	Enabled
Default access tier	Hot
Blob anonymous access	Disabled
Blob soft delete	Enabled (7 days)
Container soft delete	Enabled (7 days)
Versioning	Disabled
Change feed	Disabled
NFS v3	Disabled
SFTP	Disabled

**Security**

Require secure transfer for REST API operations	Enabled
Storage account key access	Enabled
Minimum TLS version	Version 1.2
Infrastructure encryption	Disabled

**Networking**

Allow access from	All networks
Number of private endpoint connections	0
Network routing	Microsoft network routing

**Overview**

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

**Data storage**

- Containers
- File shares
- Queues
- Tables

**Security + networking**

- Networking
- Access keys

# Attached ADLS Gen2 in Storage Explorer in local :

