

## DATA ENGINEERING DATABRICKS

### Running Job on sample ingestion (sample job – ingestion task)

The screenshot shows the Databricks Jobs interface. On the left is a sidebar with navigation options: New, Workspace, Recents, Catalog, Workflows, Compute, Data Engineering, Job Runs, Machine Learning, Playground, Experiments, Features, Models, and Serving. The main area displays the 'Example job' details. The 'Runs' tab is active, showing a table of job runs. The table has columns: Start time, Run ID, Launched, Duration, Spark, Status, Error code, and Run paramet... The table shows a single run on Dec 02, 2024, at 0... with Run ID 43750820... and a status of 'Succ...'. The 'Job details' panel on the right shows the Job ID (641321982697134), Creator (azuser2367\_mml.local), Run as (azuser2367\_mml.local), Tags (Add tag), Description (Add description), Git (Not configured), Schedule (None), and Compute (azuser2367\_mml.local's Cluster).

### Creating azure blob storage:

The screenshot shows the 'Create a storage account' page in the Azure portal. The 'Basics' tab is selected. The page includes a description of Azure Storage and a 'Project details' section. The 'Subscription' dropdown is set to 'MML Learners'. The 'Resource group' dropdown is set to 'rg-azuser2367\_mml.local-jeNUz'. The 'Storage account name' field is filled with 'azdbblobij'. The 'Region' dropdown is set to '(Asia Pacific) Central India'. The 'Primary service' dropdown is set to 'Azure Blob Storage or Azure Data Lake Storage Gen 2'. The 'Review + create' button is visible at the bottom.

## Create a storage account

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  ☐  
 SFTP can only be enabled for hierarchical namespace accounts

Enable network file system v3  ☐  
 To enable NFS v3 'hierarchical namespace' must be enabled. [Learn more about NFS v3](#) 

### Blob storage

Allow cross-tenant replication  ☐

Access tier  ☒ Hot: Optimized for frequently accessed data and everyday usage scenarios

[Previous](#) [Next](#) [Review + create](#)

 Give feedback

## Create a storage account

Basics [Advanced](#) [Networking](#) [Data protection](#) [Encryption](#) [Tags](#) [Review + create](#)

 [View automation template](#)


### Basics

Subscription	MML Learners
Resource group	rg-azuser2367_mml.local-jeNUz
Location	Central India
Storage account name	azdbblobjj
Primary service	Azure Blob Storage or Azure Data Lake Storage Gen 2
Performance	Standard
Replication	Locally-redundant storage (LRS)

### Advanced







Enable hierarchical namespace	Disabled
Enable SFTP	Disabled
Enable network file system v3	Disabled
Allow cross-tenant replication	Disabled
Access tier	Hot
Enable larva file shares	Enabled

[Previous](#) [Next](#) [Create](#)

 Give feedback

## azdbblobjj

Storage account

 Upload  Open in Explorer  Delete  Move  Refresh  Open in mobile  CLI / PS  Feedback

### Overview

- Activity log
- Tags
- Diagnose and solve problems
- Access Control (IAM)
- Data migration
- Events
- Storage browser
- Storage Mover
- Partner solutions
- > Data storage
- > Security + networking
- > Data management
- > Settings
- > Monitoring
- > Monitoring (classic)
- > Automation

### Essentials

[JSON View](#)

Resource group <a href="#">(move)</a>	: <a href="#">rg-azuser2367_mml.local-jeNUz</a>	Performance	: Standard
Location	: centralindia	Replication	: Locally-redundant storage (LRS)
Subscription <a href="#">(move)</a>	: <a href="#">MML Learners</a>	Account kind	: StorageV2 (general purpose v2)
Subscription ID	: 2a3c6418-97b9-4d96-a24b-2c2d7633d375	Provisioning state	: Succeeded
Disk state	: Available	Created	: 12/2/2024, 7:44:29 PM
Tags <a href="#">(edit)</a>	: <a href="#">Add tags</a>		

[Properties](#) [Monitoring](#) [Capabilities \(7\)](#) [Recommendations \(0\)](#) [Tutorials](#) [Tools + SDKs](#)

### Blob service

Hierarchical namespace	Disabled
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

### 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
Private endpoint connections	0

jjcontainers

Container

Search

Upload

Change access level

Refresh

Delete

Change tier

Acquire lease

Break lease

View snapshots

Create snapshot

Give feedback

Overview

Diagnose and solve problems

Access Control (IAM)

Settings

Authentication method: Access key (Switch to Microsoft Entra user account)

Location: jjcontainers

Search blobs by prefix (case-sensitive)

Show deleted blobs

Add filter

Name	Modified	Access tier	Archive status	Blob type	Size	Lease state	
<input type="checkbox"/> credit card.csv	12/2/2024, 7:51:04 PM	Hot (Inferred)		Block blob	649.27 KiB	Available	...
<input type="checkbox"/> bin.csv	12/2/2024, 7:51:06 PM	Hot (Inferred)		Block blob	7.52 MiB	Available	...

## Creating azure data lake storage

Home > Storage accounts (classic) >

Create a storage account

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 \*

MML Learners

Resource group \*

rg-azuser2367\_mml.local-jeNUz

Create new

Instance details

Storage account name \*

azdatalakejj

Region \*

(Asia Pacific) Central India

Deploy to an Azure Extended Zone

Primary service

Azure Blob Storage or Azure Data Lake Storage Gen 2

Performance \*

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

Premium: Recommended for scenarios that require low latency.

Redundancy \*

Locally-redundant storage (LRS)

Previous

Next

Review + create

Give feedback

Home > Storage accounts (classic) >

Create a storage account

Enable storage account key access ☒

Default to Microsoft Entra authorization in the Azure portal ☐

Minimum TLS version

Permitted scope for copy operations (preview)

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 ☐

Enable network file system v3 ☐

Blob storage

Allow cross-tenant replication ☐

Previous

Next

Review + create

# Create a storage account ...

- Basics
- Advanced
- Networking
- Data protection
- Encryption
- Tags
- Review + create

[View automation template](#)

### Basics

Subscription	MML Learners
Resource group	rg-azuser2367_mml.local-jeNUz
Location	Central India
Storage account name	azdatalakejj
Primary service	Azure Blob Storage or Azure Data Lake Storage Gen 2
Performance	Standard
Replication	Locally-redundant storage (LRS)

### Advanced

Enable hierarchical namespace	Enabled
Enable SFTP	Disabled
Enable network file system v3	Disabled
Allow cross-tenant replication	Disabled
Access tier	Hot
Enable large file shares	Enabled

- Previous
- Next
- Create

Home > [azdatalakejj\\_1733149495671](#) | Overview >

azdatalakejj

Storage account

Search

UploadOpen in ExplorerDeleteMoveRefreshOpen in mobileCLI / PSFeedback

Overview

Activity logTagsDiagnose and solve problemsAccess Control (IAM)Data migrationEventsStorage browserPartner solutionsData storageSecurity + networkingData managementSettingsMonitoringMonitoring (classic)AutomationHelp

Essentials

Resource group (move)

: rg-azuser2367\_mml.local-jeNUz

Location

: centralindia

Subscription (move)

: MML\_Learners

Subscription ID

: 2a3c6418-97b9-4d96-a24b-2c2d7633d375

Disk state

: Available

Tags (edit)

: Add tags

Performance

: Standard

Replication

: Locally-redundant storage (LRS)

Account kind

: StorageV2 (general purpose v2)

Provisioning state

: Succeeded

Created

: 12/2/2024, 7:55:23 PM

Properties

MonitoringCapabilities (5)Recommendations (0)TutorialsTools + 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

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

Private endpoint connections

0

Microsoft Azure

Search resources, services, and docs (G+I)

Copilot

azuser2367\_mml.local@TECHADEMY LEARNING SOLUT...

Home > [azdatalakejj\\_1733149495671](#) | Overview > [azdatalakejj](#) | Containers >

datajj

Container

Search

UploadAdd DirectoryRefreshRenameDeleteChange tierAcquire leaseBreak leaseGive feedback

Overview

Diagnose and solve problemsAccess Control (IAM)Settings

Authentication method: Access key (Switch to Microsoft Entra user account)

Location: datajj

Search blobs by prefix (case-sensitive)

Show deleted objects

Name	Modified	Access tier	Archive status	Blob type	Size	Lease state	
<div>orders.csv</div>	12/2/2024, 7:57:49 PM	Hot (Inferred)		Block blob	583 B	Available	...

Successfully uploaded blob(s)  
Successfully uploaded 1 blob(s).