

Create User Assigned Managed Identity ...



Basics Tags Review + create

Project details

Select the subscription to manage deployed resources and costs. Use resource groups like folders to organize and manage all your resources.

Subscription * ⓘ

Azure for Students

Resource group * ⓘ

covid-reporting-rg

[Create new](#)

Instance details

Region * ⓘ

East US

Name * ⓘ

covid_hd_mngidentity

Add role assignment



Role Members Conditions (optional) Review + assign

Selected role Storage Blob Data Owner

Assign access to ☒ User, group, or service principal
 ☐ Managed identity

Members [+ Select members](#)

Name	Object ID	Type	
covid_hd_mngidentity	e8d4c5dd-b847-4540-b80b-339355073...	App	

Description

Optional

Azure for Students | Resource providers

Subscription

Search



Register Unregister Refresh Feedback

- Overview
- Activity log
- Access control (IAM)
- Tags
- Diagnose and solve problems
- Security
- Events
- Cost Management**
 - Cost analysis
 - Cost alerts
 - Budgets
 - Advisor recommendations
- Billing**
 - Billing profile invoices
- Settings**
 - Programmatic deployment
 - Billing properties

hd

Provider	Status
Microsoft.HDInsight	Registered

Subscription Registering Resource Provider

Successfully registered resource provider:
'microsoft.hdinsight' , for subscription: '50381b31-77bb-4cc2-a6f1-156c0ef3f834'

Create HDInsight cluster ...



Select the subscription to manage deployed resources and costs. Use resource groups like folders to organize and manage all your resources.

Subscription *	<div>Azure for Students</div>
Resource group *	<div>covid-reporting-rg</div> <div>Create new</div>

Cluster details

Name your cluster, pick a region, and choose a cluster type and version. [Learn more](#)

Cluster name *	<div>ercovid-reporting-hdi</div>
Region *	<div>East US</div>
Availability zone ⓘ	<div></div>
Cluster type *	<div>Hadoop</div> <div>Change</div>
Version *	<div>Hadoop 3.1.0 (HDI 4.0)</div>

Cluster credentials

Enter new credentials that will be used to administer or access the cluster.

Cluster login username * ⓘ	<div>admin</div>
Cluster login password *	<div>.....</div>
Confirm cluster login password *	<div>.....</div>
Secure Shell (SSH) username * ⓘ	<div>sshuser</div>
Use cluster login password for SSH	<div><input checked="" type="checkbox"/></div>

[Review + create](#)

[« Previous](#)

[Next: Storage »](#)

Create HDInsight cluster ...



- Basics
- Storage**
- Security + networking
- Configuration + pricing
- Tags
- Review + create

Select or create storage accounts that will be used for the cluster's logs, job input, and job output. Configure the cluster's access to these accounts, if needed.

Primary storage

Select or create a storage account that will be the default location for cluster logs and other output.

Primary storage type *

Azure Data Lake Storage Gen2

Primary storage account *

ercovidreportingdatalake

└─

Filesystem * ⓘ

hdi-logs

☒ Enable secure channel ⓘ

Identity

Select a user-assigned managed identity to represent the cluster for Azure Data Lake Gen2 Storage account access. Only identities with access to the selected storage account are listed. Assign the managed identity to the 'Storage Blob Data Owner' role on the storage account. [Learn More](#)

User-assigned managed identity * ⓘ

covid_hd_mngidentity

Additional Azure Storage

Link additional Azure Storage accounts to the cluster.

Account name
ercovidreportingsa <div>✕</div>

EXPLORER

Search for resources

Collapse all Refresh all

Quick Access

- Emulator & Attached
 - Storage Accounts
- Azure for Students (AvinashChinta@my.unt.edu)
 - Storage Accounts
 - cs71003200143e43ad6
 - dbstorageocb4jzhwww7hy (ADLS Gen2)
 - dlspractise (ADLS Gen2)
 - ercovidreportingdatalake (ADLS Gen2)
 - Blob Containers
 - \$logs
 - lookup
 - processed
 - raw
 - File Shares
 - Queues
 - Tables
 - ercovidreportingsa
 - Blob Containers
 - \$logs
 - configs
 - population
 - scripts
 - File Shares
 - Queues
 - Tables
 - Disks
 - cloud-shell-storage-southcentralus
 - covid-reporting-rg
 - databricks-rg-dbs-handson-bj47lln3b
 - NetworkWatcherRG
 - rg-avinash-handson

processed raw \$logs population configs lookup scripts

Upload Download Open Preview New Folder Select All Copy Paste Clone Delete Undo Manage History Folder Statistics Refresh

Active blobs (default) scripts > hql

Name	Access Tier	Access Tier Last Modified	Last Modified	Blob Type	Content Type	Size	Status	Remaining Days	Deleted Time	Lease State	Disk Name
covid_transform_testing.hql	Hot (inferred)		7/1/2023 7:32 PM	Block Blob	text/plain	3.41 KiB	Active				

Showing 1 to 1 of 1 cached items

Activities

Clear completed Clear successful

- ✓ Transfer of 'C:\Users\Dheeraji Bajjuri\Downloads\testing\testing\000000_0' to 'processed/ecdc/testing/' complete: 1 item transferred (used SAS, discovery completed) Started at: 7/1/2023 7:35 PM, Duration: 5 seconds [Copy AzCopy Command to Clipboard](#)
- ✓ Successfully created folder 'ecdc/testing'
- ✓ Transfer of 'C:\Users\Dheeraji Bajjuri\Downloads\covid_transform_testing.hql' to 'scripts/hql/' complete: 1 item transferred (used SAS, discovery completed) Started at: 7/1/2023 7:32 PM, Duration: 7 seconds [Copy AzCopy Command to Clipboard](#)
- ✓ Successfully created blob container 'scripts'

EXPLORES

Search for resources

[Collapse all](#)

[Refresh all](#)

Quick Access

Emulator & Attached

Storage Accounts

Azure for Students (AvinashChinta@my.unt.edu)

Storage Accounts

cs71003200143e43ad6

dbstorageocb4jzhwww7hy (ADLS Gen2)

dlspractise (ADLS Gen2)

ercovidreportingdatalake (ADLS Gen2)

Blob Containers

\$logs

lookup

processed

raw

File Shares

Queues

Tables

ercovidreportingsa

Blob Containers

\$logs

configs

population

scripts

File Shares

Queues

Tables

Disks

cloud-shell-storage-southcentralus

covid-reporting-rg

databricks-rg-dbs-handson-bj47lln3b

NetworkWatcherRG

rg-avinash-handson

processed raw \$logs population configs lookup scripts

Upload Download Open Preview New Folder Select All Copy Paste Clone Rename Move Manage ACLs Properties Delete Undo Folder Statistics Refresh

← → ∨ ↑

Active blobs (default)

processed > ecdc > testing

Filter by prefix (case-sensitive)

Name	Access Tier	Access Tier Last Modified	Last Modified	Blob Type	Content Type	Size	Status	Deletion ID	Remaining Days	Deleted Time	Lease State
000000_0	Hot (inferred)		7/1/2023 7:35 PM	Block Blob	text/plain	114.10 KiB	Active				

Showing 1 to 1 of 1 cached items

Activities

Clear completed Clear successful

✓

Transfer of 'C:\Users\Dheeraji Bajjuri\Downloads\testing\testing\000000_0' to 'processed/ecdc/testing/' complete: 1 item transferred (used SAS, discovery completed)
Started at: 7/1/2023 7:35 PM, Duration: 5 seconds

Copy AzCopy Command to Clipboard

✓

Successfully created folder 'ecdc/testing'

✓

Transfer of 'C:\Users\Dheeraji Bajjuri\Downloads\covid_transform_testing.hql' to 'scripts/hql/' complete: 1 item transferred (used SAS, discovery completed)
Started at: 7/1/2023 7:32 PM, Duration: 7 seconds

Copy AzCopy Command to Clipboard