



Journey PPT

Shell Training
Week 2 – 04/9/2023 to 08/9/2023

Gaurav Singh
Associate Data Engineer



Day 1 -

- Introduction to Azure Storage
- Benefits of Azure Storage
- Azure Storage data services
- Review options for storing data in Azure
- Creating a Container
- Blob Configuration
- Azure Storage Security





Day 2 -

- Azure File Storage
- Table & Queue Storage
- Azure Disk Storage
- Storage Monitoring
- Storage Resource Tools





InPrivate (2) C X n B S S W S S C d T SQL S G d N N n T h A +

https://portal.azure.com/#create/Microsoft.StorageAccount-ARM

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

Shellunext_1693422731... UNEXT

Home > Storage accounts >

Create a storage account

Basics Advanced Networking Data protection Encryption Tags Review

Create new

Instance details

Storage account name ⓘ * shellzrs007

Region ⓘ * (US) East US

Deploy to an edge zone

Performance ⓘ *

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

☐ Premium: Recommended for scenarios that require low latency.

Redundancy ⓘ *

Zone-redundant storage (ZRS)

Review < Previous Next : Advanced >

Give feedback

Search ENG IN 3:20 PM 9/4/2023



Day 3 -

- Azure Data Lake
- Input Datasets
- Pipeline
- Output Datasets
- Linked Services





Day 4 -

- Control Flow Activity
- Get Metadata Activity
- Filter Activity
- Copy Data
- ForEach Loop
- If Condition
- Parameterization
- Derived Column
- Join
- Conditional Split
- Assessment / MCQ Quiz





Microsoft Azure | Data Factory | shelldemodatafa | Search factory and documentation | Shellunext_1693422731647@npunext.onmicrosoft.com UNEXT

» «

- Dashboards
- Runs
- Pipeline runs
- Trigger runs
- Change Data Capture (previ...
- Runtimes & sessions
- Integration runtimes
- Data flow debug
- Notifications
- Alerts & metrics

All pipeline runs > employee_blob_2_adls - Activity runs

Rerun Cancel Refresh Update pipeline List Gantt

```
graph LR; A[Copy data: employee_blob_2_adls] --> B[Get Metadata: Get Metadata1]; B --> C[Filter: Filter1]; C --> D[Copy data: Copy data1];
```

Activity runs

Pipeline run ID ac12a9b4-3c36-41cf-8899-77bf7876f3aa

All status List Monitor in Azure Metrics Export to CSV

Showing 1 - 7 items

Activity name	Activity status	Activity type	Run start	Duration	Log	Integ
If Condition1	Inactive	If Condition	9/6/2023, 3:16:38 PM	Less than 1s		
Copy data1	Inactive	Copy data	9/6/2023, 3:16:38 PM	Less than 1s		Unkn
Filter1	Inactive	Filter	9/6/2023, 3:16:38 PM	Less than 1s		
Get Metadata1	Inactive	Get Metadata	9/6/2023, 3:16:38 PM	Less than 1s		Unkn

Windows Search ENG IN 3:22 PM 9/6/2023



Microsoft Azure | Data Factory | shelldemodatafa | Search factory and documentation | Shellunext_1693422731647@npunext.onmicrosoft.com UNEXT

Preview experience Off

Factory Resources

- Pipelines 1
 - employee_blob_2_adls
- Change Data Capture (preview) 0
- Datasets 5
 - AzureTable1
 - DelimitedText1
 - DelimitedText2
 - DelimitedText3
 - nation_src_blob
- Data flows 1
 - dataflow1
- Power Query 0

Activities

- data flow
- Move and transform
 - Data flow

Run Succeeded
Successfully ran employee_blob_2_adls (Pipeline).
[View pipeline run](#)

Data flow

Data flow1

General Settings **Parameters** 1 User properties

Data flow parameters ⓘ

Name	Value	Type	Expression ⓘ
No data flow parameters			

Windows taskbar: Search, File Explorer, Edge, Teams, Outlook, OneDrive, 3:22 PM 9/6/2023



Microsoft Azure | Data Factory | shelldemodatafa | Search factory and documentation

Shellunext_1693422731647@npunext.onmicrosoft.com UNEXT

» «

- Dashboards
- Runs
 - Pipeline runs
 - Trigger runs
 - Change Data Capture (previ...
- Runtimes & sessions
 - Integration runtimes
 - Data flow debug
- Notifications
 - Alerts & metrics

All pipeline runs > ✓ employee_blob_2_adls - Activity runs

Rerun Cancel Refresh Update pipeline

Copy data (employee_blob_2_adls) → Get Metadata (Get Metadata1) → Filter (Filter1) → Copy data (Copy data1)

Activity runs

Pipeline run ID: ac12a9b4-3c36-41cf-8899-77bf7876f3aa

All status List [Monitor in Azure Metrics](#) [Export to CSV](#)

Showing 1 - 7 items

Activity name	Activity status	Activity type	Run start	Duration	Log	Integ
If Condition1	⊖ Inactive	If Condition	9/6/2023, 3:16:38 PM	Less than 1s		
Copy data1	⊖ Inactive	Copy data	9/6/2023, 3:16:38 PM	Less than 1s		Unkn
Filter1	⊖ Inactive	Filter	9/6/2023, 3:16:38 PM	Less than 1s		
Get Metadata1	⊖ Inactive	Get Metadata	9/6/2023, 3:16:38 PM	Less than 1s		Unkn

Windows taskbar: Search, 3:22 PM 9/6/2023



Microsoft Azure | Data Factory | lumenlabdatafactory007 | Search factory and documentation

Microsoft recently announced the public preview of Microsoft Fabric, a brand new and exciting way to build cloud-first data analytics. Click [here](#) to get started with Fabric Data Factory!

Preview experience: Off

Factory Resources

- Pipelines: 1
 - pipeline1
- Change Data Capture (preview): 0
- Datasets: 4
 - DelimitedText1
 - DelimitedText2
 - DelimitedText3
 - lumenlaboutput
- Data flows: 1
 - dataflow1
- Power Query: 0

Activities

- Move and transform
- Synapse
- Azure Data Explorer
- Azure Function
- Batch Service
- Databricks
- Data Lake Analytics
- General
- HDInsight
- Iteration & conditionals
- Machine Learning
- Power Query

pipeline1

Validate Debug Add trigger

Get Metadata

Get Metadata1

ForEach

ForEach2

Activities

Copy data1

ForEach

ForEach1

Parameters Variables Settings Output

+ New

Windows taskbar: Search, File Explorer, Edge, Teams, Settings, Task View, 2:51 PM 9/7/2023



Microsoft Azure | Data Factory | lumenlabdatafactory007 - Azure Data Factory | Search factory and documentation

Microsoft recently announced the public preview of Microsoft Fabric, a brand new and exciting way to build cloud-first data analytics. Click [here](#) to get started with Fabric Data Factory!

Data Factory | Validate all | Publish all | Preview experience | Off

Factory Resources

- Pipelines 1
 - pipeline1
- Change Data Capture (preview) 0
- Datasets 4
 - DelimitedText1
 - DelimitedText2
 - DelimitedText3
 - lumenlaboutput
- Data flows 1
 - dataflow1
 - Power Query 0

DelimitedText1 | pipeline1 | dataflow1 | lumenlaboutput | DelimitedText2

Validate | Data flow debug

source1 → sales → select1 → Truck → sink1

Air → sink2

Others → sink3

Parameters | Settings

+ New

2:51 PM 9/7/2023



Microsoft Azure | Data Factory | lumenlabdatafactory007 - Azure Data Factory | Search factory and documentation | Shellunext_1693422731647@npunext.onmicrosoft.com

Microsoft recently announced the public preview of Microsoft Fabric, a brand new and exciting way to build cloud-first data analytics. Click [here](#) to get started with Fabric Data Factory!

Data Factory | Validate all | Publish all | Preview experience | Off

Factory Resources

- Pipelines 1
 - pipeline1
- Change Data Capture (preview) 0
- Datasets 4
 - DelimitedText1
 - DelimitedText2
 - DelimitedText3
 - lumenlaboutput**
- Data flows 1
 - dataflow1
- Power Query 0

DelimitedText1 | pipeline1 | dataflow1 | **lumenlaboutput** | DelimitedText2

DelimitedText
lumenlaboutput

Connection | Schema | Parameters

Linked service * | AzureDataLakeStorage1 | Test connection | Edit | New | Learn more

File path * | output1 | Directory | File name | Browse | Preview

Compression type | Select...

Column delimiter | Tab (\t)

Row delimiter | Default (\r\n, or \r\n)

Encoding | Default(UTF-8)

Quote character | Double quote (")

Windows Taskbar: Search, File Explorer, Edge, Office, Settings, Task View, 2:51 PM 9/7/2023



Microsoft Azure | Data Factory | lumenlabdf | Search factory and documentation | Shellunext_1693422731647@npunext.onmicrosoft.com UNEXT

CDC > adfcdc2

✓ adfcdc2
Status Running Last processed 5 minutes ago

Refresh Edit CDC Delete CDC Stop Chennai, Kolkata, Mu... : Last 24 hours

CHANGES READ	CHANGES WRITTEN
0	0

Captured changes count

Sep 8 02:25 pm

2:29 PM 9/8/2023



Day 5 -

- Mapping data flows in Azure Data Factory
- Change data capture in Azure Data Factory
- Mapping data flow Debug Mode



**Thank you
for your
time.**

