Azure Data Factory

Lab: Trigger a pipeline using storage event trigger

Pre-requisites:

- Azure Pass subscription
- Azure Data Lake Storage Gen2 storage account
- Azure SQL Database

Lab Objective:

After completing this lab, you will be able to:

- Create the trigger in Azure Data Factory.
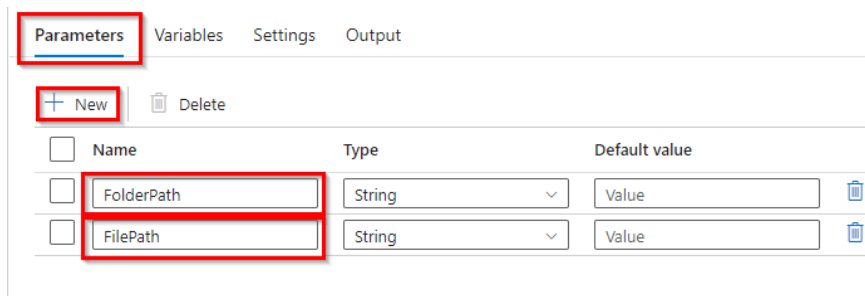- Create use the triggers in the pipeline

Exercise 1: Create a trigger to execute the pipeline based on storage event.

The main task for this exercise are as follows:

- Create a copy activity to copy data.
- Create a storage event trigger to run the pipeline.

Task 1: Create a copy activity to copy data

1. Open data factory adfxx.

2. **Open the authoring canvas** in ADF homepage, click on the **pencil icon** on the left sidebar and select the **+ pipeline button** to open the authoring canvas and create a pipeline.

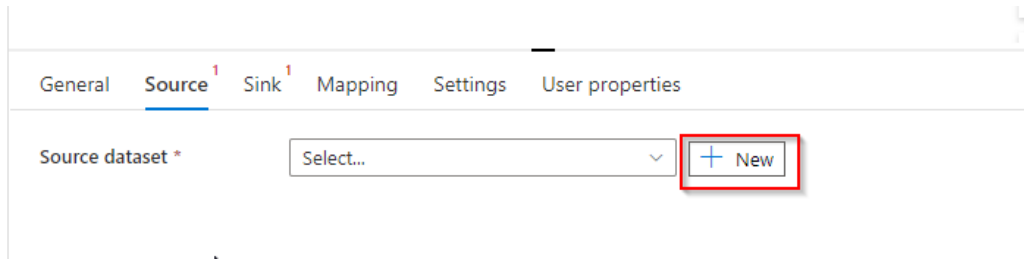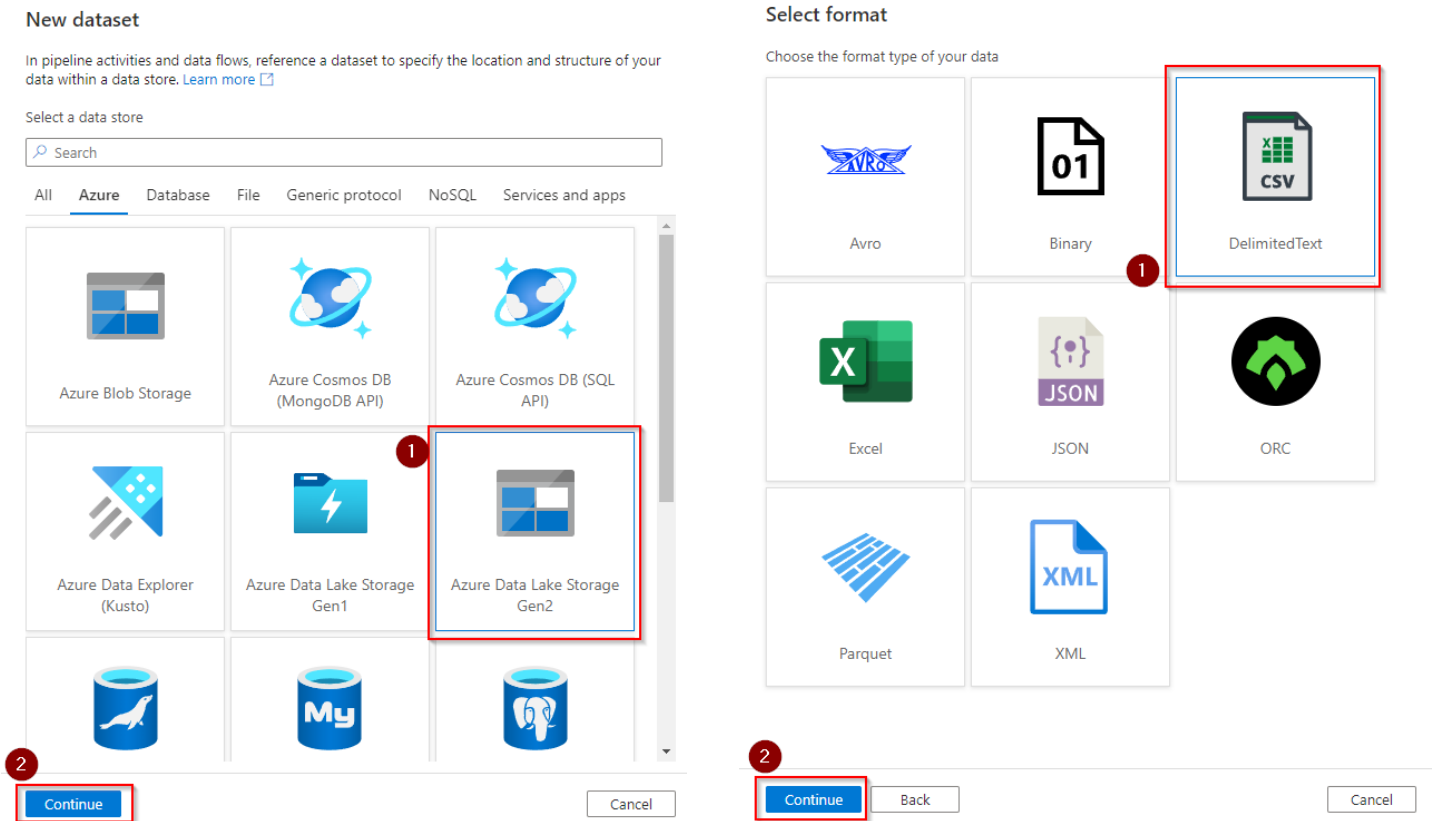3. Create two pipeline parameters FolderPath and FilePath.

4. **Add a copy activity** In the Activities pane, open the **Move and Transform** accordion and drag the **Copy data** activity onto the pipeline canvas.
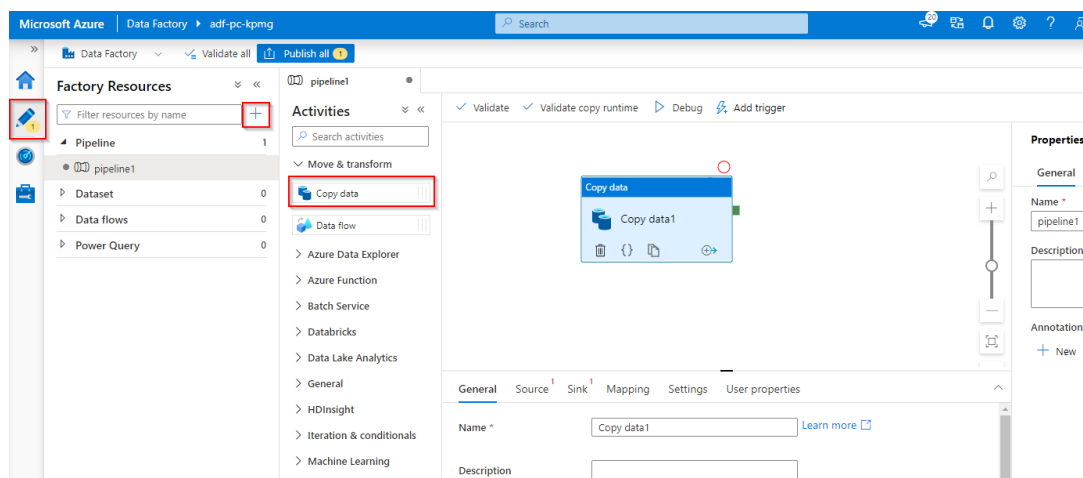
5. Create a new Dataset

6. Select Azure Data Lake Storage Gen 2 and click continue and the select Delimited Text and continue

7. In set properties enter name ADLSGen2Source, Select Existing Linked service for the Storage account otherwise create a new.

Set properties

Name

ADLSGen2Source **1**

Linked service *

Select... **2** ⌄

Filter...

Select...

+ New

ADLSGen2 **3**

8. Do not select the folder path or any file.

Set properties

Name

ADLSGen2Source **1**

Linked service *

ADLSGen2 ⌄ ✎

File path

File system / Directory / File 📁 | ⌄

First row as header ☐ Do notSelect any file

Import schema

○ From connection/store ○ From sample file ● None

> Advanced

2

OK   Back   Cancel

9. Open ADLSGen2Source and create two dataset parameters, FolderPath and FilePath

10. Under connection set the file path using add dynamic content as mention in the image below:

Set the file path as show below:

11. Select the copy activity, under source setting add dynamic content for Dataset properties using pipeline parameters.

12. Create another dataset for sink of type ADLSGen2 and create the parameter FileName

13. Under connection setting of sink dataset browse the file path **to data->output** and for the filename select **add dynamic content** and select the Filename @dataset parameter

14. In the sink of copy activity add the sink data set and pass the pipeline parameter value to dataset parameter filename.

15. In Azure portal, enable the Microsoft.EventGrid Resource Provider from Azure Pass subscription as shown below.
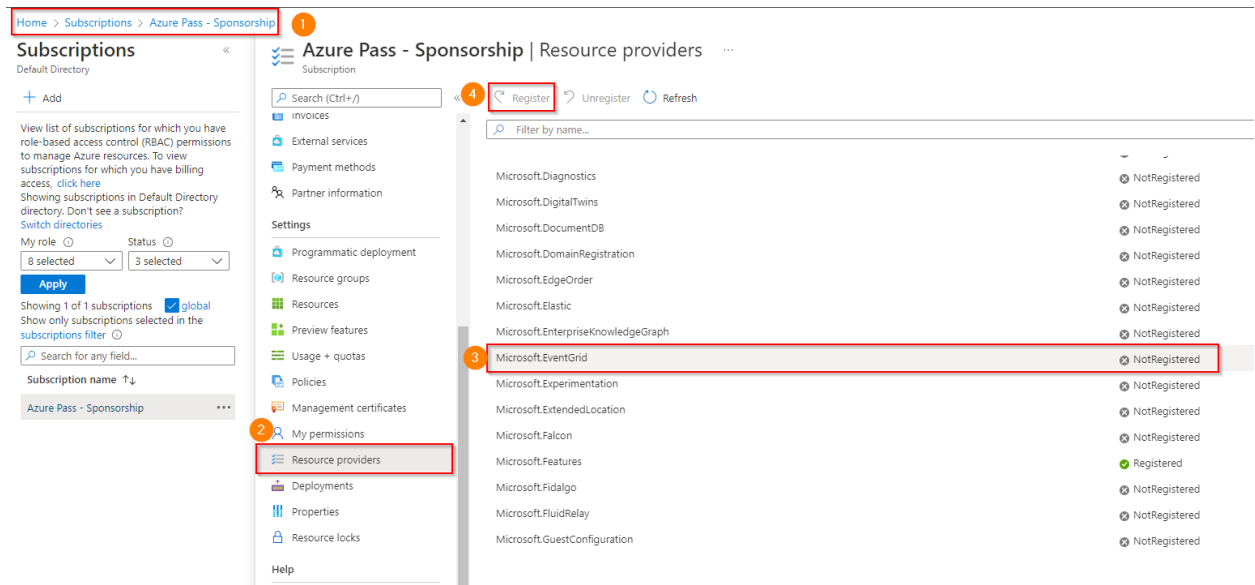
16. Add trigger to the pipeline:
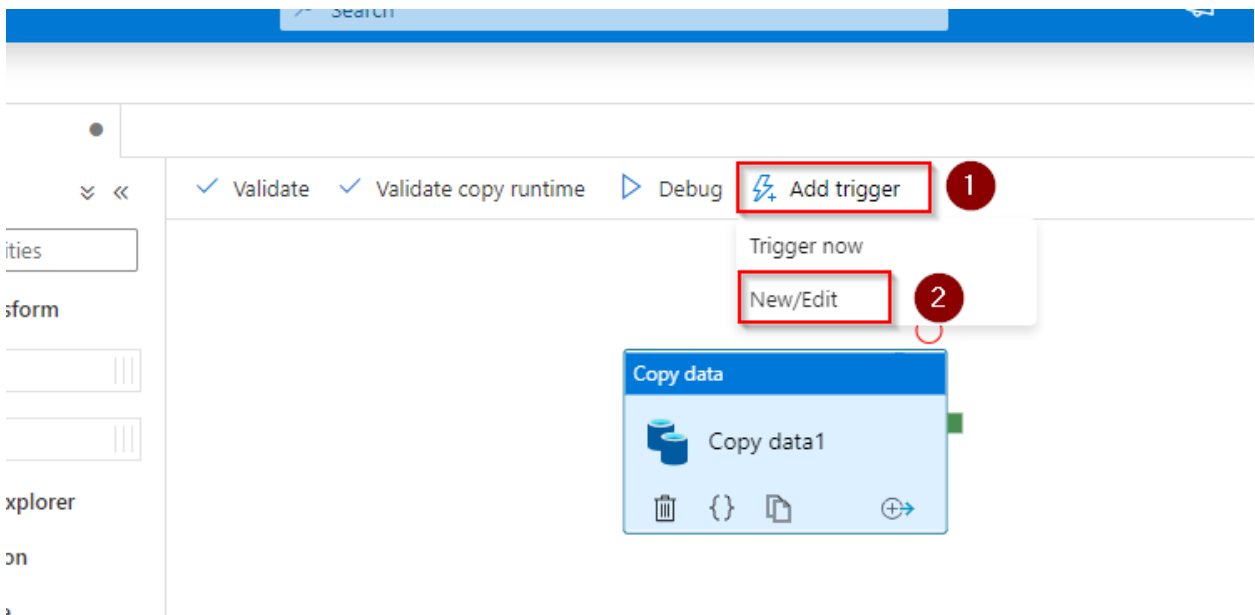
17. Configure the Trigger as show below:

Add triggers

Choose trigger...

Search

+ New

New trigger

Name *

StorageEventTrigger **1**

Description

Type *

Storage events **2**

Account selection method * ⓘ

3 ● From Azure subscription ○ Enter manually

Azure subscription ⓘ

4 Azure Pass - Sponsorship (c8fd5e96-3329-4700-b2ef-455eb9a984e4)

Storage account name * ⓘ

5 datalakepc1

Container name * ⓘ

data **6**

Blob path begins with ⓘ

input/ **7**

Blob path ends with ⓘ

.csv **8**

Click Continue:

Next initialize the following:

> FolderPath: @triggerBody().folderPath
>
> FilePath: @triggerBody().fileName

Then click OK and Publish the pipeline

Edit trigger

Trigger Run Parameters

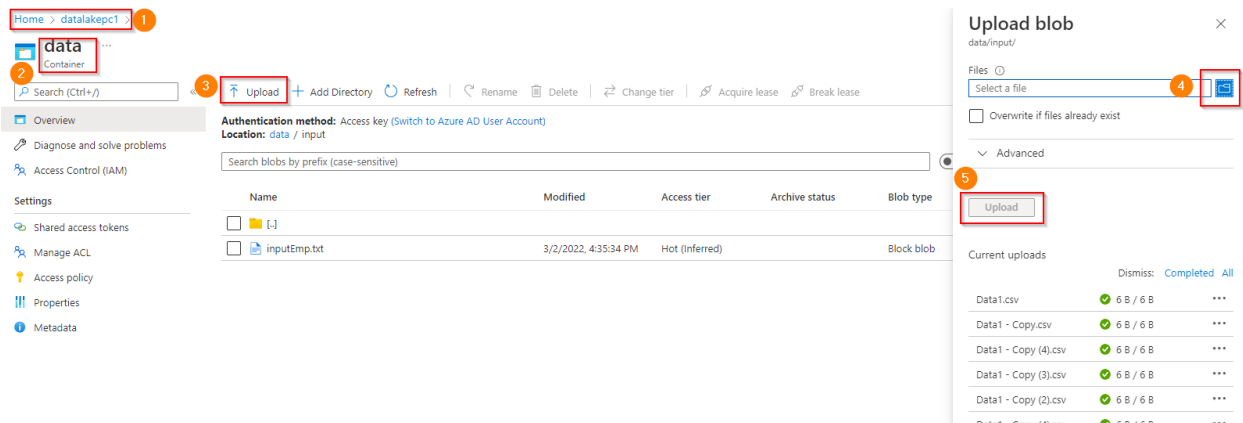ℹ️ Parameters that are not provided a value will not be included in the trigger.

Name	Type	Value
FolderPath	string	@triggerBody().folderPath
FilePath	string	@triggerBody().fileName

Make sure to "Publish" for trigger to be activated after clicking "OK"

OK   Cancel

18. Create some dummy csv file and upload them from storage account, this will trigger the pipeline you can go to monitor hub under pipeline run you can see the pipeline is executing and you can also monitor the trigger run under monitor hub.