

BlobEventTrigger along with dynamic folder path and timestamp in filename

Last updated by | Jackie Huang | Jan 4, 2022 at 12:24 AM PST

Contents

- [Issue](#)
- [Resolution](#)

Issue

The customer would like to copy files that arrive to a storage account container and copy them to a datalake folder. The files should be copied to the sink with the timestamp on each file for when they were copied.

In storage account there will be several subfolders (for each country) and when a file is collected from the source that is in any of these subfolders, the name of said subfolder should be in the output folder.

Example: COUNTRY can be (ES PL DE FR IT...)

- Source Storage Account
 - /mytest/year/month/day/"COUNTRY"/version/filename_with_timestamp.json
- Sink Data Lake Storage
 - /mytest/yyyy/mm/dd/"COUNTRY"/version/v1_esb_storesdayend_20200311125305799.json

In a nutshell, these are the goals that customer would like to achieve:

- Only new files are copied to the sink
- Dynamic folder path with country abbreviation
- Timestamp for each file copied

Resolution

The country abbreviation comes from only new files captured by a blob event trigger.

In other words, whenever a new file is uploaded to a container, it will start a pipeline that will capture the folder path and filename, will extract the country and set to a variable.

This variable is later used to define the destination folder path.

1. We need to create 2 parameters *in the pipeline*

General Parameters Variables

+ New | Delete

☐ NAME

sourceFolderPath

sourceFilename

- Then, we create a blob event trigger that will track a container in the blob storage and will start a pipeline when a file arrives

Edit trigger

Name *
triggerFileWithCountryAbbreviation

Description

Type *
BlobEventsTrigger

Account selection method *
☒ From Azure subscription ☐ Enter mail

Azure subscription
Select all

Storage account name *
/subscriptions/

Container name *
/testlztoraw/

Blob path begins with
mytest/

Blob path ends with

Event *
☒ Blob created ☐ Blob deleted

Ignore empty blobs *
☒ Yes ☐ No

Annotations
+ New

Activated *

- Associate this trigger with the pipeline

Trigger (1)

Trigger now

New/Edit (1)

4. Click on the link to the 2 parameters we have created

Add triggers		
Choose trigger...		
TRIGGER	TYPE	PARAMETERS
triggerFileWithCountryAbbreviation	Event	2

5. Populate these parameters with **@triggerBody().folderPath** and **@triggerBody().fileName**

Edit trigger

Trigger Run Parameters

NAME	TYPE	VALUE
sourceFolderPath	string	@triggerBody().folderPath
sourceFilename	string	@triggerBody().fileName

At this point, we are specifying that the folder and filename of all incoming files will be saved to 2 parameters in the pipeline

6. Then, now, we should use these parameters in the dataset, to dynamically choose the right folder and file in the source. So, create a dataset with the 2 same parameters



Json

fileBlobEventTrigger

General Connection Schema **Parameters**

+ New | Delete

<input type="checkbox"/>	NAME	TYPE	DEFAULT VALUE
	sourceFolderPath	String	Value
	sourceFilename	String	Value

7. Then, the connection should be defined as

General **Connection** Schema Parameters

Linked service * AzureBlobStorage1 Test connection Open + New

File path * @dataset().sourceFolderPath / Directory / @dataset().sourceFilename

Compression type none

Encoding Default(UTF-8)

8. Go back to the pipeline and create a copy activity using this dataset as source. You will notice that you will have to populate the dataset parameters created in (6).

Use **@pipeline().parameters.sourceFolderPath** and **@pipeline().parameters.sourceFilename**

General **Source** Sink Mapping Settings User properties

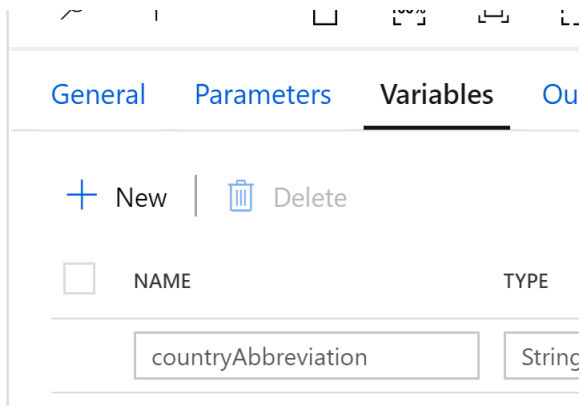
Source dataset * fileBlobEventTrigger Open + New

Dataset properties ⓘ

NAME	VALUE
sourceFolderPath	@pipeline().parameters.sourceFolderPath
sourceFilename	@pipeline().parameters.sourceFilename

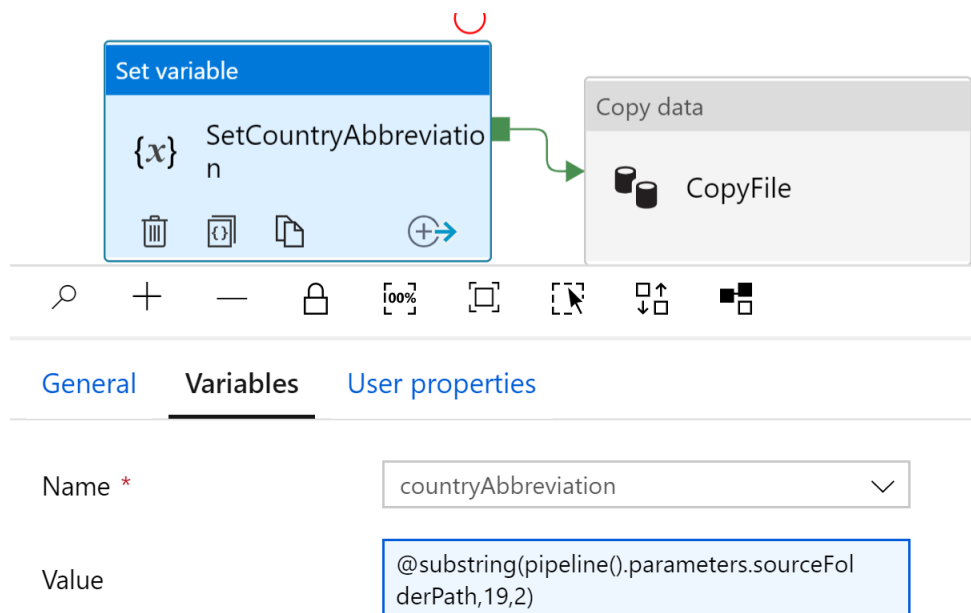
At this point, we are dynamically copying any file that is uploaded to a tracked container in our Azure Blob Storage from the source to any sink. Now, let's work on the sink to dynamically set the country abbreviation.

9. Go to the pipeline again and create a variable called countryAbbreviation



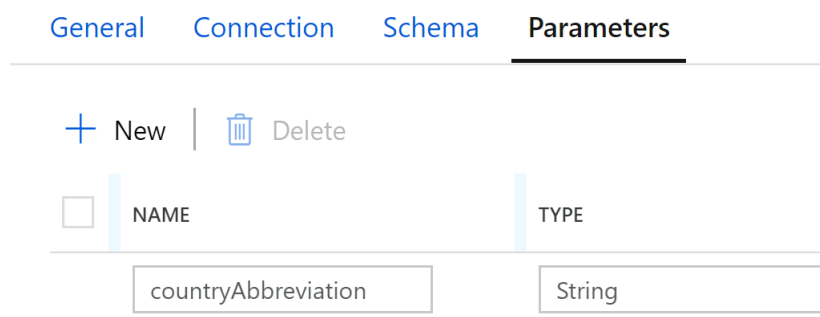
10. Right before the copy activity, add the Set variable activity:

It will use countryAbbreviation created above and the value will be **@substring(pipeline().parameters.sourceFolderPath,<any initial length>,2)** because the folder structure you mentioned is <any folder path length/><country abbreviation with 2 characters>/version/filename.json



Now, let's use this variable to set dynamically the folder in the sink

11. Create a sink dataset in your ADLS Gen1 with a parameter countryAbbreviation



The connection should be like this, with the filename as:

```
@concat(formatDateTime(utcnow(), 'yyyy'),'/',formatDateTime(utcnow(),
'MM'),'/',formatDateTime(utcnow(),
'dd'),'/',dataset().countryAbbreviation,'/v1/',v1_esb_storesdayend_',formatDateTime(utcnow(),
'yyyyMMddHHmmssfff'),'json')
```

General Connection Schema Parameters

Linked service * AzureDataLakeStore2 Test connection Open +

File path * mytest / @concat(formatDateTime(utcnow(), 'yyyy'),'/',formatDateTime(utcnow(), 'MM'),'/',formatDateTime(utcnow(), 'dd'),'/',dataset().countryAbbreviation,'/v1/',v1_esb_storesdayend_',formatDateTime(utcnow(), 'yyyyMMddHHmmssfff'),'json')

12. Go back to the pipeline and add this dataset to the sink. Populate the dataset property with the variable value, as

General Source Sink Mapping Settings User properties

Sink dataset * insertCountryAbbreviationAndTimestamp Open New

Dataset properties ⓘ

NAME	VALUE
countryAbbreviation	@variables('countryAbbreviation')

Now we are dynamically using the year, month, day and country in the sink folder path, and timestamp in the sink filename.

Additional Information:

- **Icm References:**
- **Author:** ganapol
- **Reviewer:** vimals
- **Keywords:**

How good have you found this content?

