wortell

Lab 7 - Global Parameters

Requirements

In order to start the lab, it is important that Lab5 is completed and that the Virtual Machine is booted up and the IR is running.

Objective

Yesterday, we set up the ADF from A to Z and ran pipelines. As mentioned in Lab2, it can happen that another team also has an ADF and there might be a dependency between both ADFs or to other services. In the lab, we will tackle this by executing the pipeline in adf-linked using an API-call.

Follow the instructions step by step.

Assignment 1 - Global parameters

1. Go to the ADF (adf.azure.com) and choose the **Non** linked ADF.
2. Click on the **toolbox** (Manage) on the left. Then click on **Global parameters** on the left.
3. Click on **New**, a new screen will appear. Fill in the following for **Name**: **SubscriptionID** and your subscriptionID for **Value**. You can find this in the URL of the ADF.
   Example: `https://adf.azure.com/en-us/management/globalparameters?factory=%2Fsubscriptions%2Ffae3cd10-ede1-4e32-b796-362b72f8e236%2FresourceGroups%2Frg-adf-training%2F`

   It is important not to copy the %2F. Based on the example, the SubscriptionID would be: **fae3cd10-ede1-4e32-b796-362b72f8e236**
4. Repeat step 3, but now create a Global parameter named: **Resourcegroup** with as **Value** the resource group name, which you can also copy from the URL.
5. Repeat step 3, but now create a Global parameter named: **DataFactory** with as **Value** the name of the adf-linked, which you can find in your resource group.
6. Repeat step 3, but now create a Global parameter named: **Pipeline** with as **Value** the name of the pipeline in the adf-linked, `PL_Wait`.
7. Click on the **Blue button** with the text **Publish all** and then on the button **Publish**.

We have now created four parameters at *factory* level. These are global constants that can be used throughout the entire Data Factory. In this way, you can easily manage settings centrally.

Assignment 2 - API caller

1. Click on the **Pencil** (Author) on the left and create a new pipeline.
2. Give the pipeline a clear name.
3. From the list of **Activities**, click on the **General** option. Click and drag **Web** onto the canvas.

4. Give the **Web block** a clear name and then click on the **Settings** tab.
5. Click on the field next to **URL** and then on **Add dynamic content**.
6. Paste or type the following code into the field:
   `https://management.azure.com/subscriptions/@{pipeline().globalParameters.SubscriptionID}/resourceGroups/@{pipeline().globalParameters.Resourcegroup}/providers/Microsoft.DataFactory/factories/@{pipeline().globalParameters.DataFactory}/pipelines/@{pipeline().globalParameters.Pipeline}/createRun?api-version=2018-06-01`

   > If you want to know more about the possibilities of the Data Factory REST API, you can read about it in the Microsoft documentation.

7. Choose **POST** for **Method**.
8. Fill in the field next to **body

** with: **{}**.

9. Choose **System Assigned Managed Identity** for **Authentication** and fill in the following for **Resource**: `https://management.core.windows.net/`.
10. Click on **New** next to **Headers** and fill in the following for **name**: **Content-Type** and for **value**: **application/json**.
11. Choose the self-named **Azure IR** under **Advanced** for **Integration runtime**.
12. Click on the **Blue button** with the text **Publish all** and then on the button **Publish**.

Assignment 3 - Linked ADF pipeline

1. At the top right of the screen, you will see a row of icons. Click on the second from the left, the icon with **the 2 screens and arrows** (Switch Data Factory).
2. A new screen will appear, and most of it will already be filled in. Choose the adf-Linked at **Data Factory Name** and then click on **OK**.
3. Create a new pipeline named: `PL_Wait`.
4. From the list of **Activities**, click on the **General** option. Click and drag **Wait** onto the canvas.
5. Give the **Wait block** a clear name.
6. Then click on the **Settings** tab and change the **Wait time in seconds** to 10.
7. Click on the **Blue button** with the text **Publish all** and then on the button **Publish**.
8. Switch back to the **non** linked adf in the same way as steps 1 and 2.
9. Click on the pipeline you created in Assignment 2, then click on **Debug**.
10. Once the pipeline is completed, switch back to the adf-linked via steps 1 and 2.

11. Click on **Monitor** on the left side. You will now immediately land on **Pipeline runs** and `PL_wait` should be among the list of executed pipelines.

12. Finally, you could take a look at the run time of the two pipelines (in the linked and the non-linked factory). What do you notice? How would you explain this?

Table of Contents