

Question 1 : How to rerun a pipe line from Data Factory Monitor.

we have to navigate to the path

Monitor >>>> Pipeline Runs >>>>

Triggered , then click on the rerun icon.

pop up will appear to confirm ,

click to OK to proceed.

Question 2 : If you have 15 activities in pipeline, if we want to debug only first 10 activities how do we do that?

On the top right corner of each activity we have red circle , just click on the activity till you want to debug.

Question 3 : How to restart failed pipe line jobs in Azure data factory.

we have 3 options to restart the failed pipeline

i) Rerun

ii) Rerun from failed activity

iii) Rerun from activity

Question 4 : How to make activity dependency in Pipeline Multiple Activities and those types.

Activity Dependency defines how subsequent activities depend on previous activities,

The different dependency conditions are: Succeeded, Failed, Skipped, and Completed.

For example, if a pipeline has Activity A -> Activity B,

succeeded : Activity B has dependency condition on Activity A with succeeded: Activity B only runs if Activity A has a final status of succeeded

failed : Activity B only runs if Activity A has a final status of failed

completed : Activity B runs if Activity A has a final status of succeeded or failed

skipped: Activity B runs if Activity A has a final status of skipped. Skipped occurs in the scenario of Activity X -> Activity Y -> Activity Z, where each activity runs only if the previous activity succeeds.

Question 5 : What is shared self-hosted Integration Runtime or How to Share Integration Runtime from One Data Factory to another Data Factory?

Shared IR: An original self-hosted IR that runs on a physical infrastructure.

Linked IR: An IR that references another shared IR. The linked IR is a logical IR and uses the infrastructure of another shared self-hosted IR.

While Creating Self-Hosted Integration Runtime we

will have separate option to Create Shared

In the self-hosted IR to be shared,

i) select Grant permission to

another Data factory

ii) in the "Integration runtime

setup" page, select the Data factory in which you want to create the linked IR.

In the data factory to which the permissions were granted, create a new self-hosted IR (linked) and enter the resource ID.

Question 6 : What is Logic App? Or How to send an email notification in Azure Data Factory? Or what is WEB Activity and when we can use this activity?

Logic app: A logic app is the Azure resource you create when you want to develop a workflow. There are multiple logic app resource types that run in different environments.

The communication between these two Azure parts is done with a JSON message via an HTTP request (post). The JSON message contains the name of the Data Factory

and the pipeline that failed, an error message and an email address.

Next, we will add a new step to our Logic App, called "Send an email". If it is the first then we have to connect your Gmail account to Azure by signing in

After creation of Azure Logic App and saving the Logic App, Azure created an endpoint URL for our Logic Apps, you'll find in the first step. Copy this URL to a

notepad, we'll need this later.

Now add an Web activity to the pipeline and rename it

Question 7 : What is Get Metadata Activity? When we can use this? OR How to get folders and filenames at dynamically?

You can use the Get Metadata activity to retrieve the metadata of any data in Azure Data Factory.

you can use get Metadata for :-

i) Validation of schema /validate the metadata of any data

ii) Trigger pipeline when data is ready/available

The Get Metadata activity takes a dataset as an input and returns metadata information as output.

child Items : List of subfolders and files in the given folder. Applicable only to folders. Returned value is a list of the name and type of each child item.

Question 8 : What is Lookup Activity and when we can use this activity?

Lookup activity can retrieve a dataset from any of the Azure Data Factory-supported data sources.

Dynamically determine which objects to operate on in a subsequent activity, instead of hard coding the object name. Some object examples are files and tables.

Question 9 : If you are running more no of pipelines and its taking longer time to execute. How to resolve this type of issues?

We can go with splitting pipelines batch wise, and create multiple integration runtime to share the load balance. by which load performance will improve

Question 10 : What is auto resolve integration runtime in azure data factory? AutoResolveIntegrationRuntions is the default integration runtime.

And the region is set to auto resolve. that means azure data factory will decide the physical location were to run/execute activity based on source , sink, and activity type.

Question 11 : Data Factory supports three types of triggers. Mention these types briefly?

- i) Schedule Trigger
- ii) Tumbling window
- iii) Event Based trigger

Question 12 : Any Data Factory pipeline can be executed using three methods. Mention these methods

Under Debug mode
Manual execution using Trigger now
Using an added scheduled, tumbling window or event trigger

Question 13 : How to load data whenever we receive a file in azure data factory? Or How to run a pipeline if we receive a file or if we delete a file?

Using event based trigger we can acheive this requirement.

Event based trigger runs pipeline only when any event occurs. such as arrival of the file, deletion of the file in azure blob storage.

steps : Triggers >> newtrigger >> event trigger(select from the option)

Fill below required option

container name

Blob path begins with

blob path ends with

event 1) Blob create 2) blob deleted 3 both select

Question 14 : Difference between Scheduled Trigger and Tumblingwindow trigger?

i) Create dependency

we have 2 pipelines and we want second pipeline to run only when first pipeline is completed successfully. so we can create dependency

using advance option for Tumblingwindow trigger

dependency size,

dependency A --> B , offset : -1 hour

size : unspecified

dependency A --> B , offset : -1 hour

size : 2 hours

self dependency

Trigger A

dependency A --> A , offset : -1 hour

size : unspecified

dependency A --> A , offset : -1 hour

size : 2 hours

ii) accessing the data in the past (when the start time and end time in the past) in this we will go for tumbling window trigger. We can not go with scheduled trigger

iii) having concurrency

Workflow. A workflow is a series of steps that defines a task or process. Each workflow starts with a single trigger, after which you must add one or more actions.

Trigger. A trigger is always the first step in any workflow and specifies the condition for running any further steps in that workflow.

Action. An action is each step in a workflow after the trigger. Every action runs some operation in a workflow.

Built-in operations. A built-in trigger or action is an operation that runs natively in Azure Logic Apps. ...