**Data Polling and Extraction Logic In ADF**

**A screenshot of a computer

AI-generated content may be incorrect.**

A screenshot of a computer

AI-generated content may be incorrect.

1. Use Lookup task to check for the fileMapping.json file in adls folder. This is to keep the list of all source files names.
2. Use ForEach activity and get output of lookup activity  
     
   **@activity('Lookup1').output.value**
3. Create empty Pipeline Variable ExpectedFiles (Array datatype). Inside ForEach, keep a Append variable activity, transform and append each output from foreach to ExpectedFiles variable.

**@concat(item().filenamePrefix, formatDateTime(utcNow(),'yyyyMMdd'), '.csv')**

1. Use Until activity with timeout 2 hours and it should keep checking for the two files with latest dates in S3.

**@equals(length(variables('FoundFiles')), length(variables('ExpectedFiles')))**

Inside until activity, use pipeline execution activity. Inside Pipeline activity, have metadata activity. That metadata activity should extract files names from S3 bucket. Send that output to Foreach activity.

@activity('Get Metadata1\_copy1').output.childItems

A screenshot of a computer

AI-generated content may be incorrect.

Create empty Pipeline Variable FoundFiles (Array datatype). Add Append variable activity in foreach and insert the filenames from s3 to that variable FoundFiles.

**@item().name**

Create FoundFiles Parameter in that pipeline. Then Add Set Variable activity to copy FoundFiles variable value to FoundFiles parameter to pass it to parent pipeline.

A screenshot of a computer

AI-generated content may be incorrect.

Now come outside the child pipeline and in Until activity (parent pipeline), next to Execute pipeline1, add set variable activity. This activity should update the FoundFiles variable which have file names that matches in s3 and expected list.

A screenshot of a computer

AI-generated content may be incorrect.

**@intersection(**

**variables('ExpectedFiles'),**

**activity('Execute Pipeline1').output.pipelineReturnValue.FoundFiles**

**)**

Add wait activity for 1 minute, the add if condition activity. Inside that IF, check if both expected and found file variable count are matching. IF matches go to another execute pipeline activity else loop until again.

**@equals(length(variables('FoundFiles')), length(variables('ExpectedFiles')))**

A screenshot of a computer

AI-generated content may be incorrect.

When adding execution activity, pass the FoundFiles value to that child pipeline. Inside that pipeline, add foreach activity to loop the filenames and extract all \*.csv files from s3 to adls container. In adls, each file should be under a directory with its name.

A screenshot of a computer

AI-generated content may be incorrect.

A screenshot of a computer

AI-generated content may be incorrect.

SINK

A screenshot of a computer

AI-generated content may be incorrect.

Now, it will keep on checking for the latest csv files in s3 with current date for maximum two hours. After two hours, it will fail and be alert. If files are there it will extract that files and load it in ADLS container in respective folder.

**Archive Batch and streaming files**

A screenshot of a computer

AI-generated content may be incorrect.

Add Metadata activity to read adls folder of streaming data. Add Foreach to read the values of metadata.

@activity('Get Metadata1').output.childItems

In foreach, add IF to check if the folder name date is older than 3 days, if yes copy those folders and it files to archive container under streaming folder using copy activity. Later, using delete activity delete it in bronze layer. In sink, keep copy behaviour as preserve hierarchy.

@less(item().name, formatDateTime(addDays(utcNow(), -3), 'yyyyMMdd'))

A screenshot of a computer

AI-generated content may be incorrect.

A screenshot of a computer

AI-generated content may be incorrect.

A screenshot of a computer

AI-generated content may be incorrect.

Add Metadata activity to read adls folder of batch data. Add Foreach to read the values of metadata.

**@activity('Get Metadata2').output.childItems**

In foreach, add copy the files to archive container under batch folder using copy activity. Later, using delete activity to delete it in bronze layer.

A screenshot of a computer

AI-generated content may be incorrect.