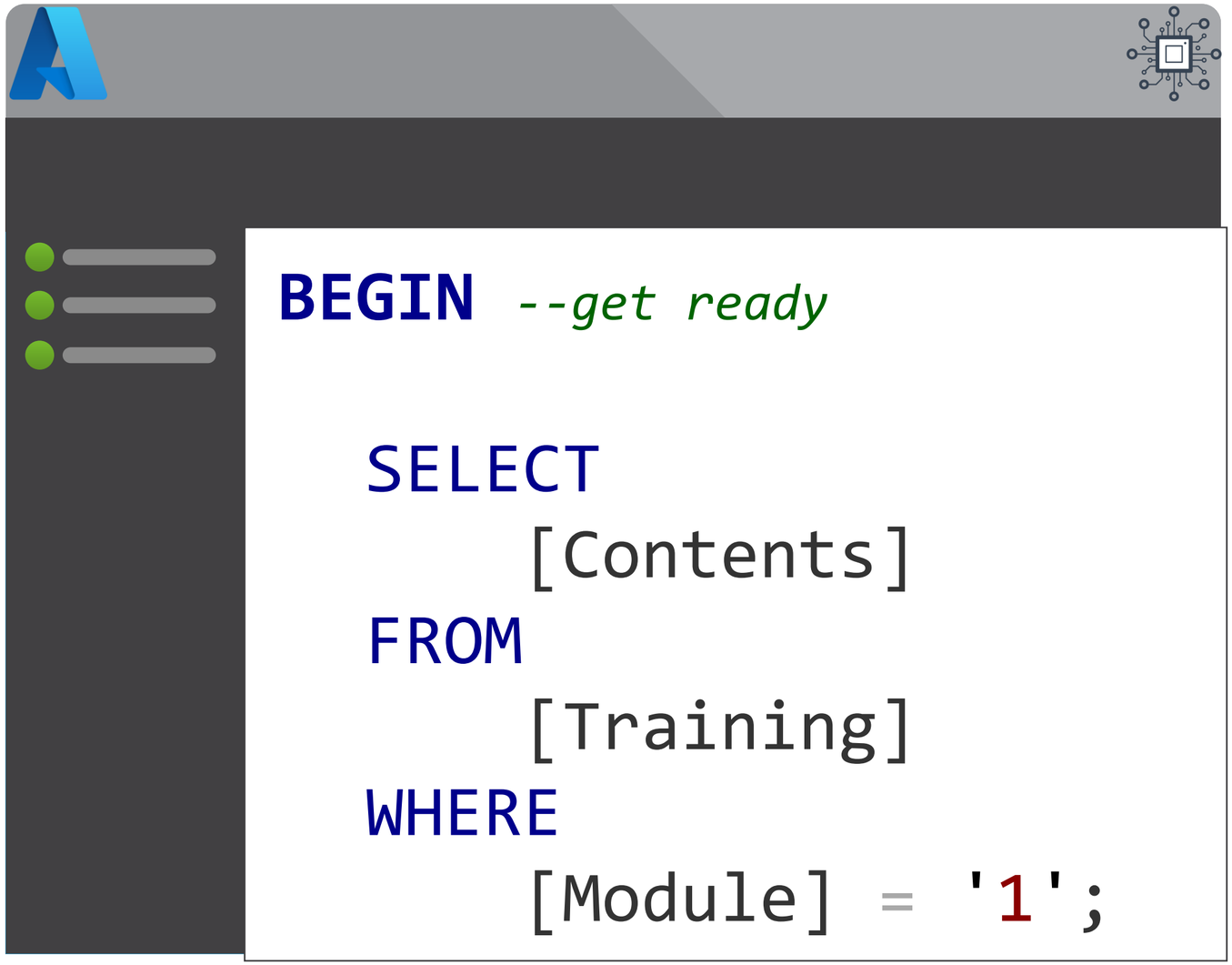
Module 1

Pipeline Fundamentals



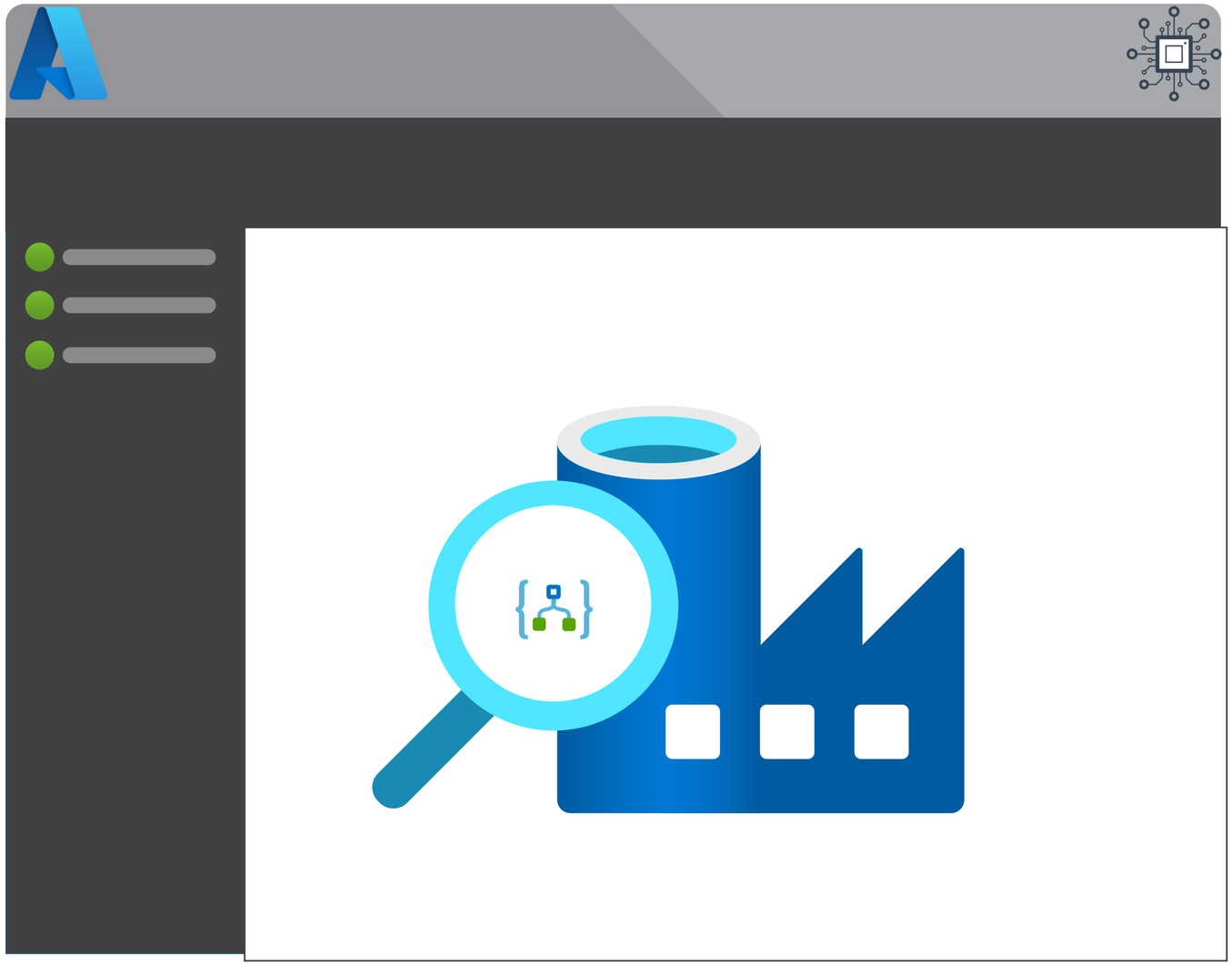
```
BEGIN --get ready

   SELECT
        [Contents]
   FROM
        [Training]
   WHERE
        [Module] = '1';
```

- The History of Azure Orchestration
- Synapse Analytics vs Data Factory
- Integration Components
- Common Activities
- Execution Dependencies

Module 1

Pipeline Fundamentals



- The History of Azure Orchestration
- Synapse Analytics vs Data Factory
- Integration Components
- Common Activities
- Execution Dependencies

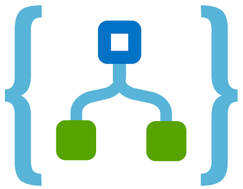
A Quick History Lesson

SQL Server
SQL Agent

SQLDB
(PaaS)

Automation

Logic Apps



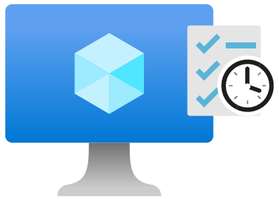
Functions

SQL Managed
Instance

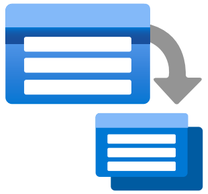
DTU Jobs

Elastic Job Agent

SQL Server
Virtual Machine

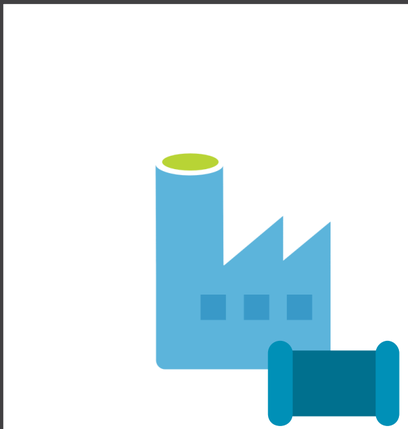


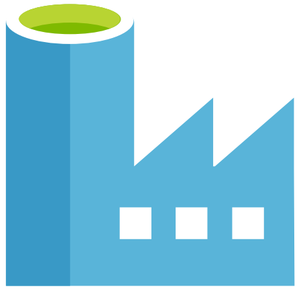
Virtual Machine
Job Schedule

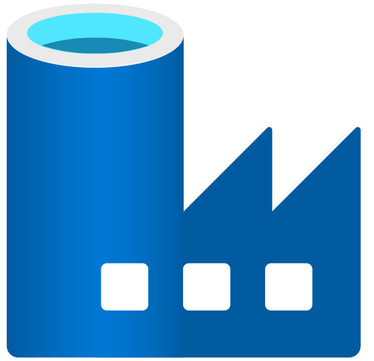


Batch

A Quick History Lesson



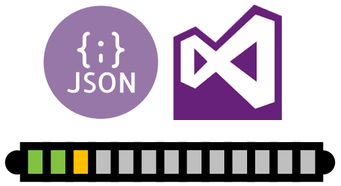




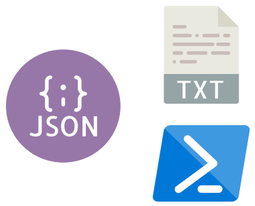
v1
July
2016

v2
Sept
2017

Oct
2020



Time Slices
VS2015 Projects



Notepad
Development

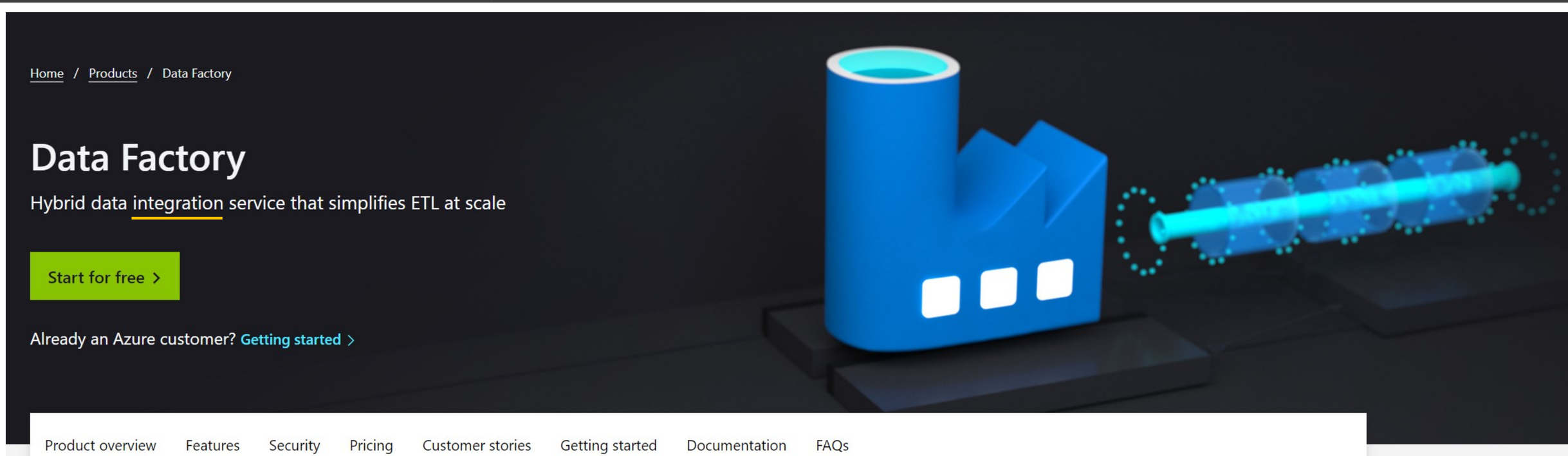
New UI

Source
Control

Data
Flows

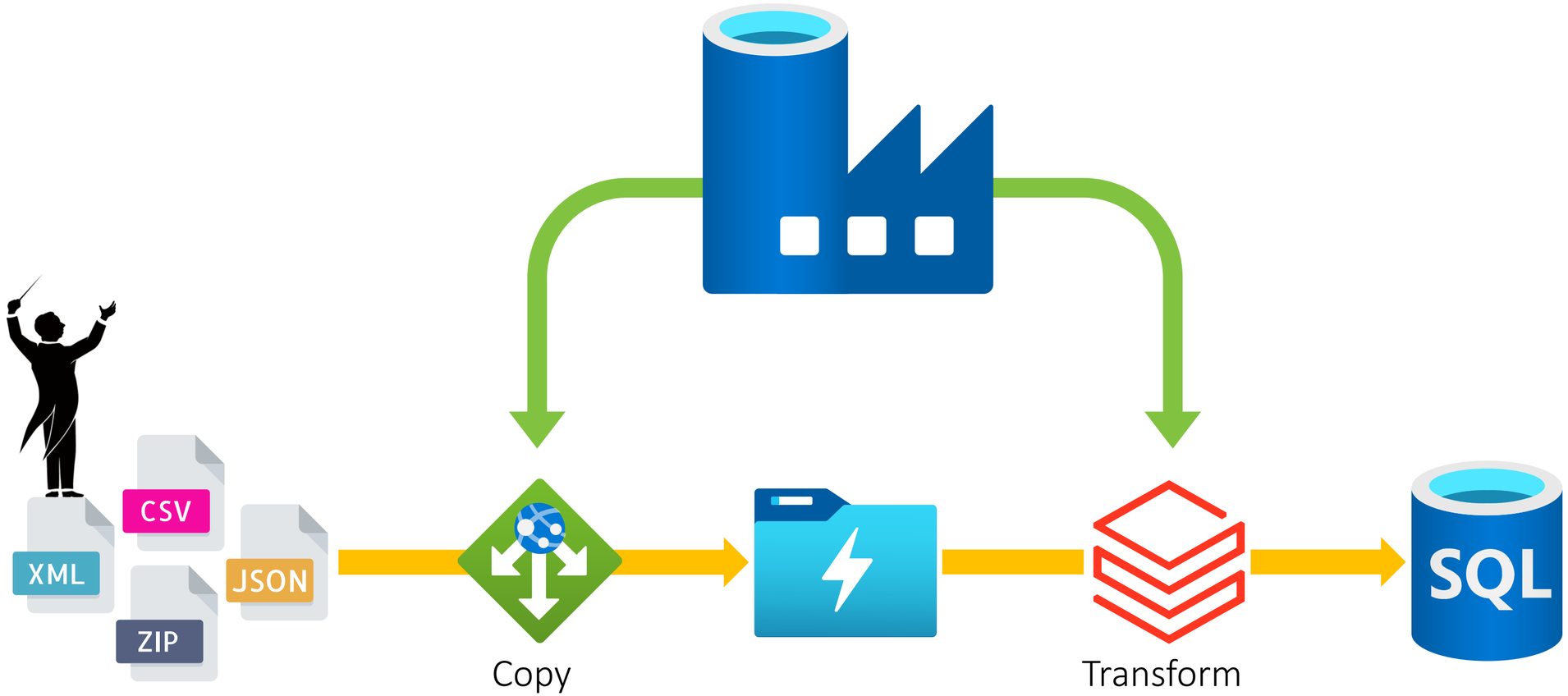
Private
Endpoints

What is Azure Data Factory (ADF)?



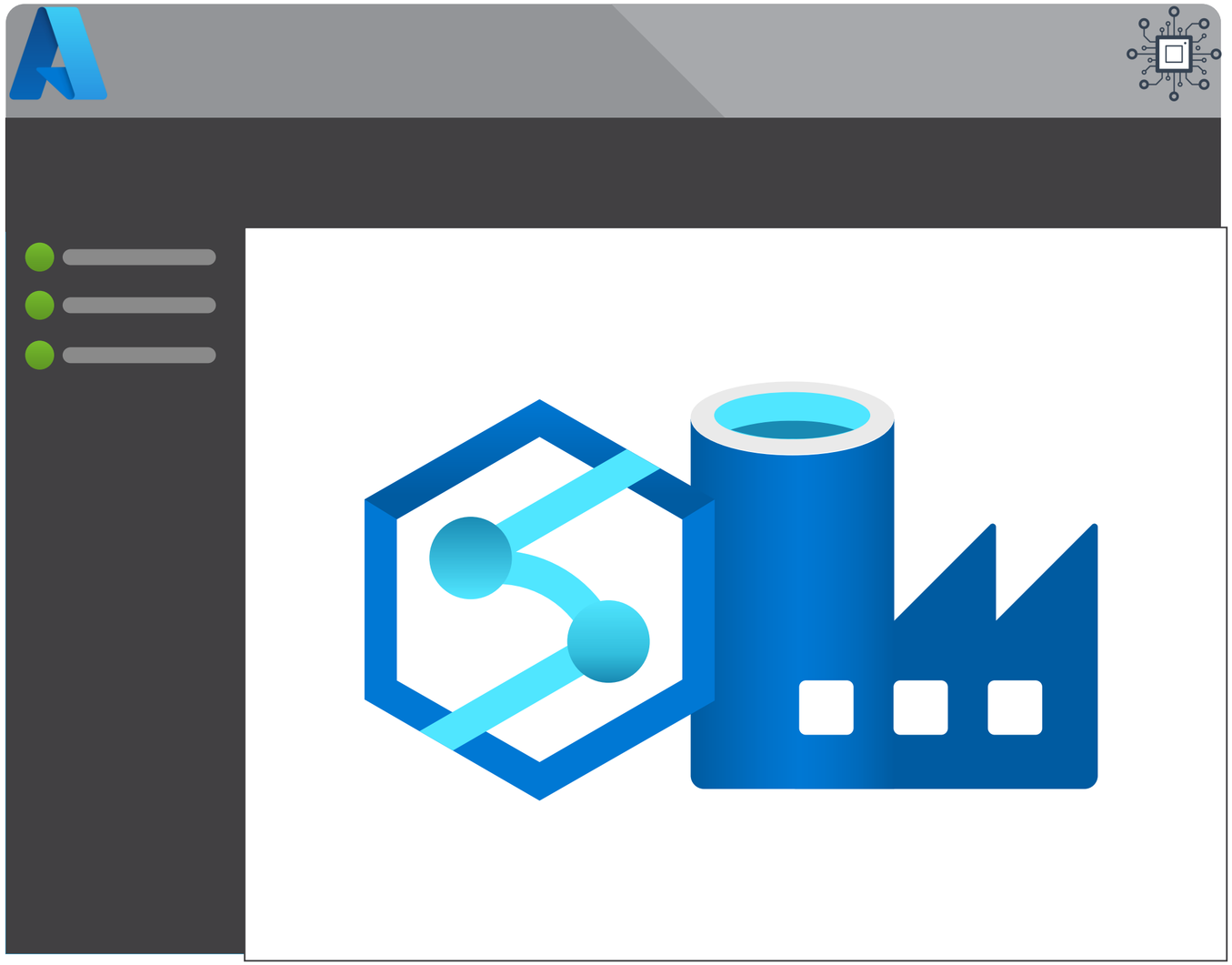
https://azure.microsoft.com/en-gb/services/data-factory

What is Azure Data Factory (ADF)?



Module 1

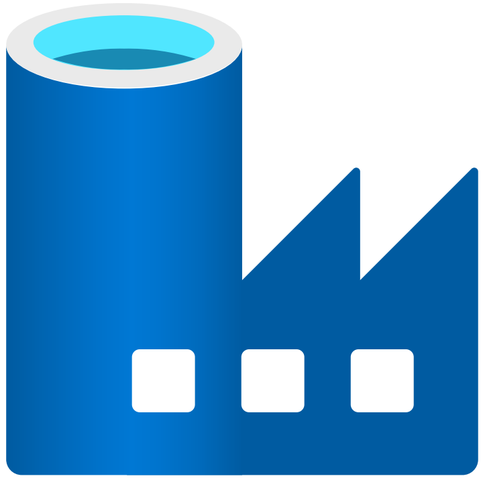
Pipeline Fundamentals

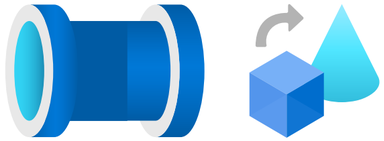


- The History of Azure Orchestration
- Synapse Analytics vs Data Factory
- Integration Components
- Common Activities
- Execution Dependencies

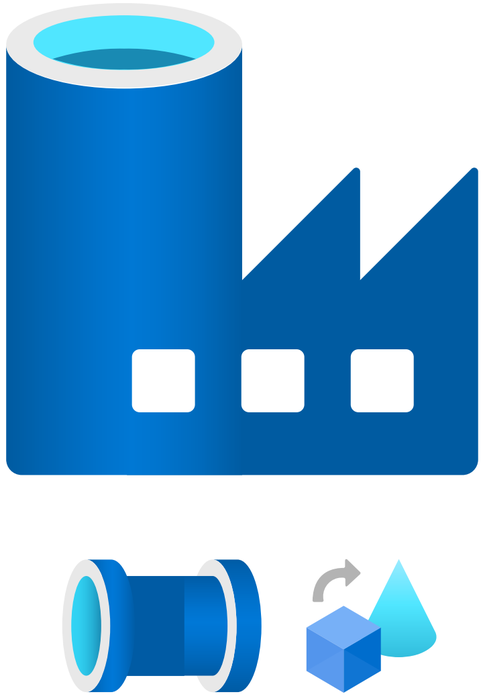
Synapse Analytics vs Data Factory

https://docs.microsoft.com/en-us/azure/synapse-analytics/data-integration/concepts-data-factory-differences



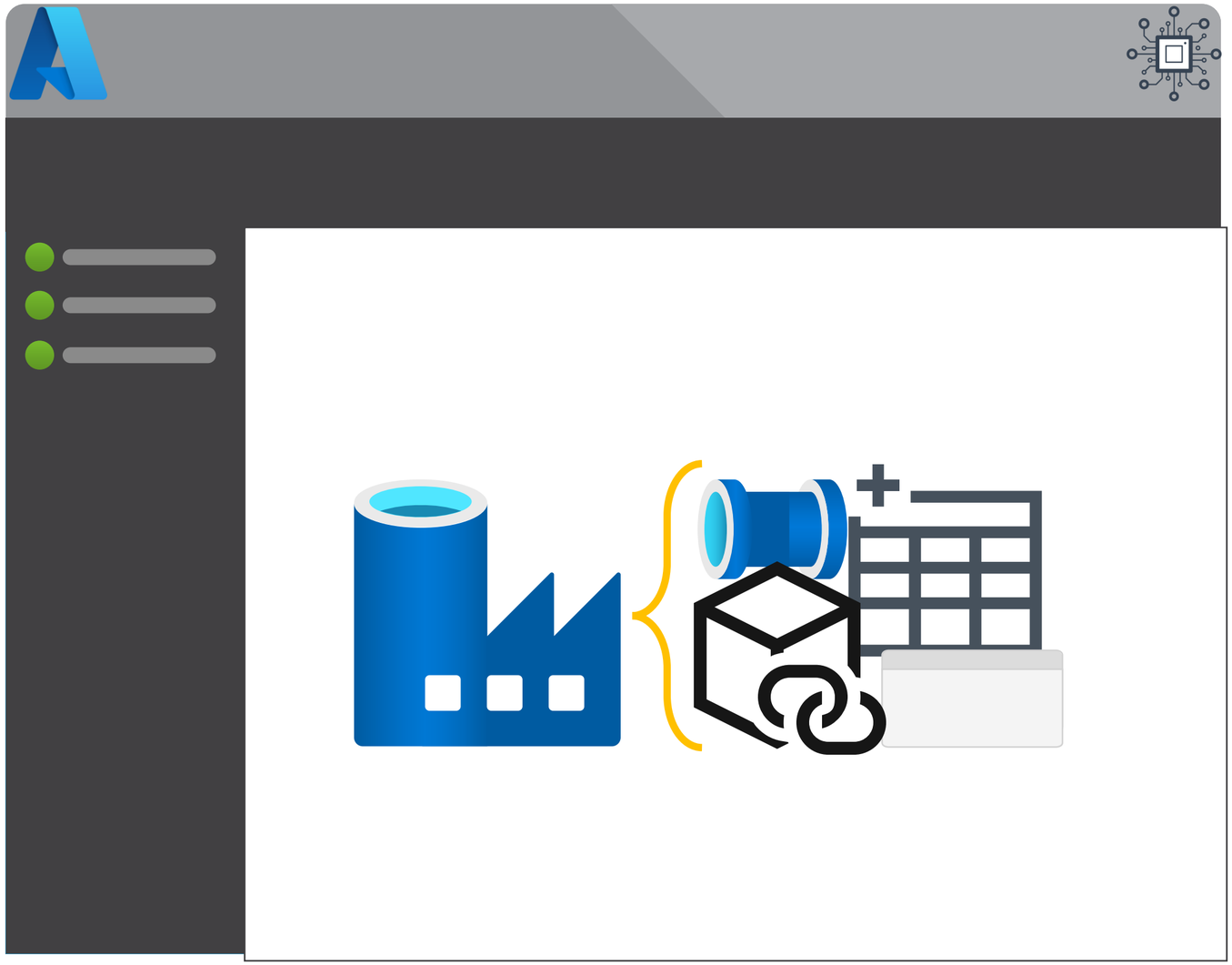


Synapse Analytics vs Data Factory



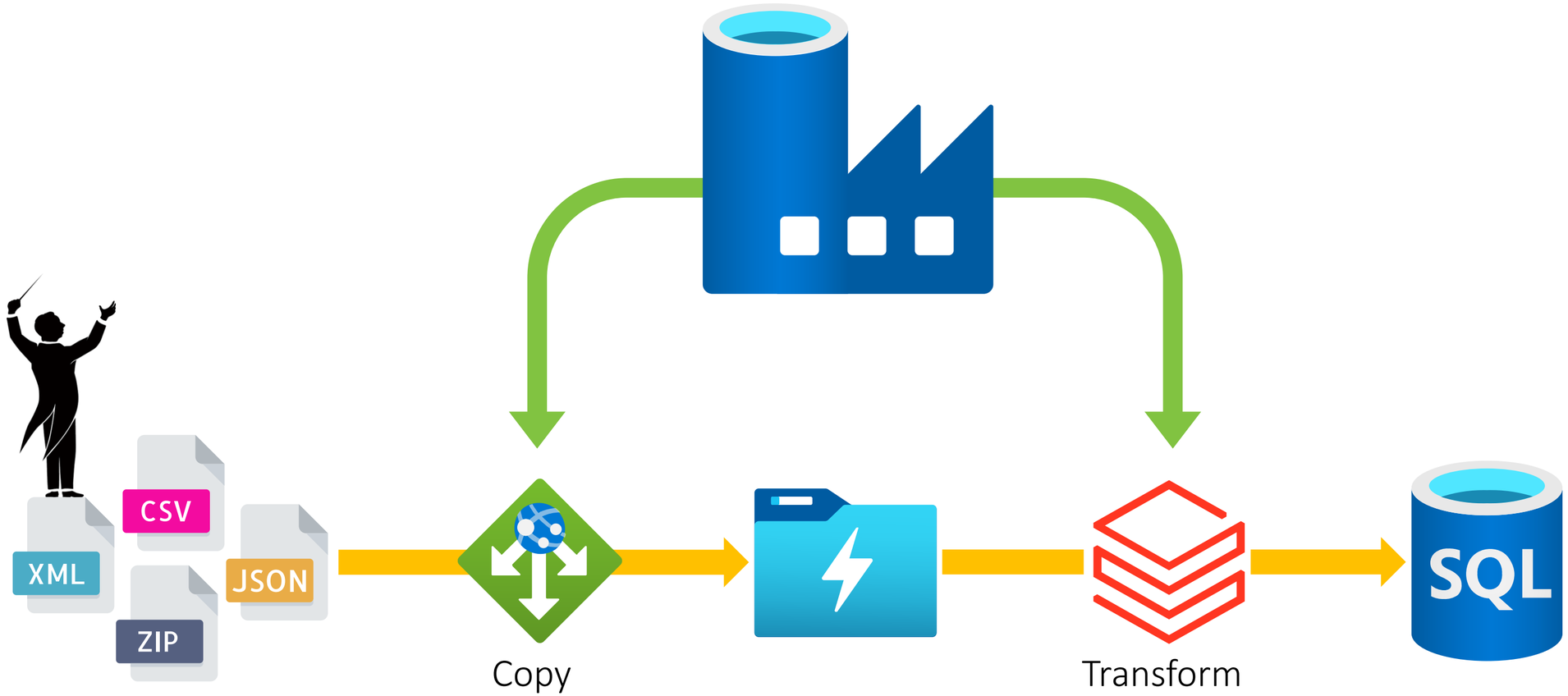
Module 1

Pipeline Fundamentals

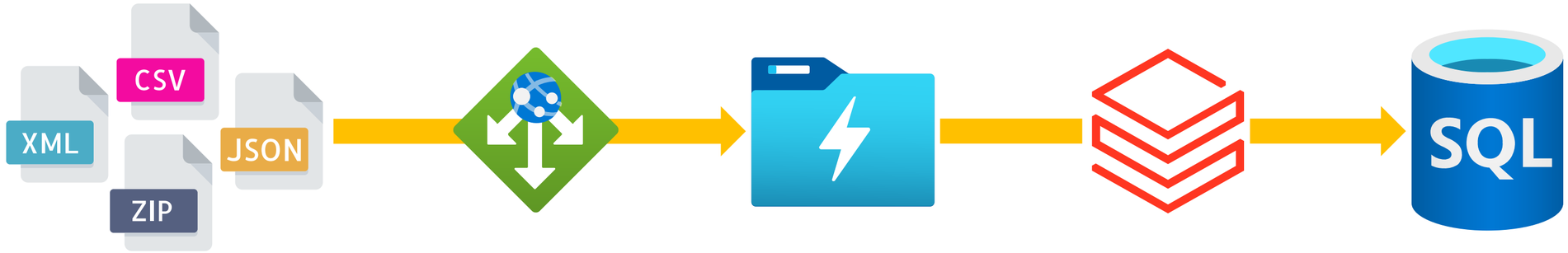


- The History of Azure Orchestration
- Synapse Analytics vs Data Factory
- Integration Components
- Common Activities
- Execution Dependencies

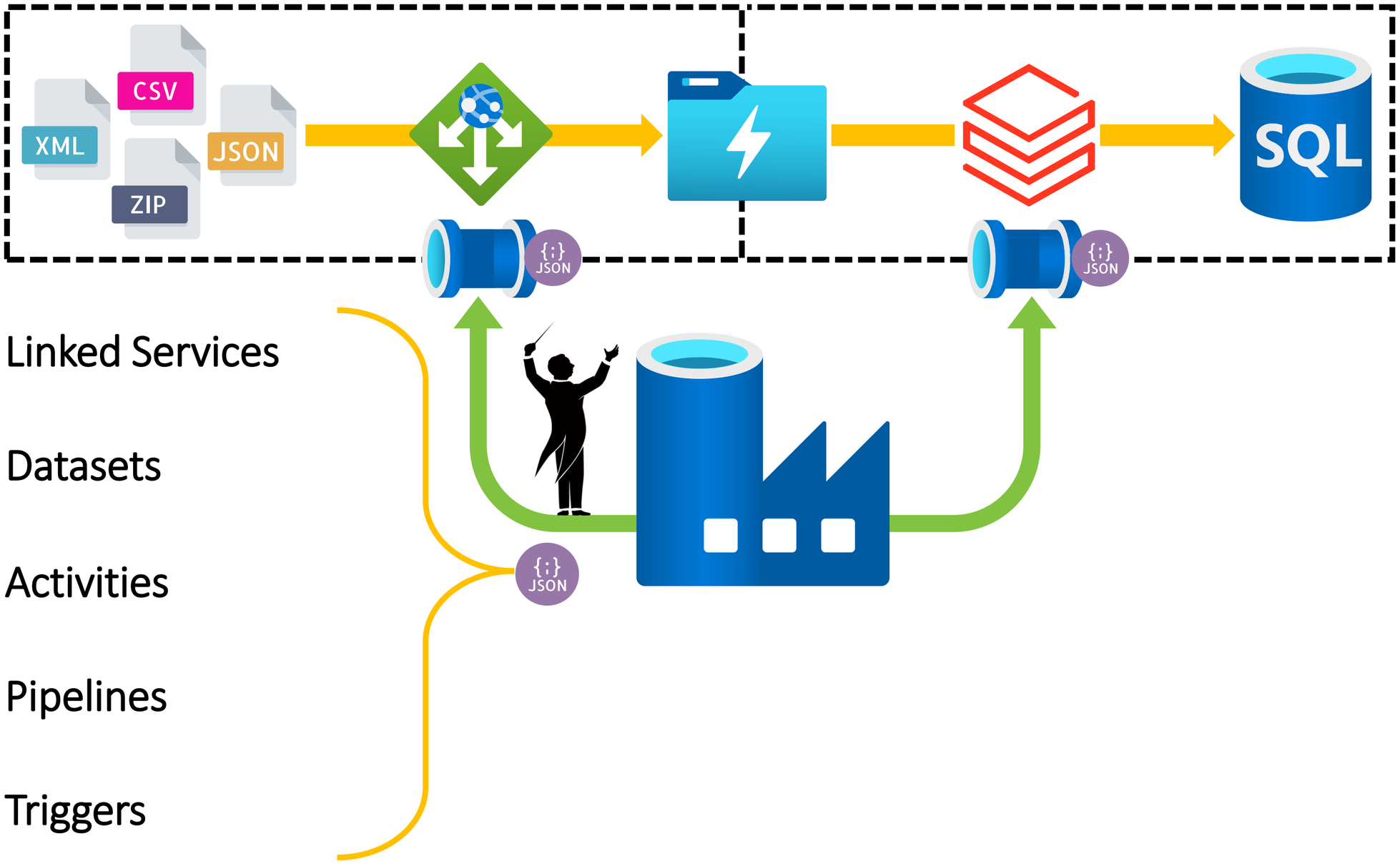
Data Factory Components



Data Factory Components



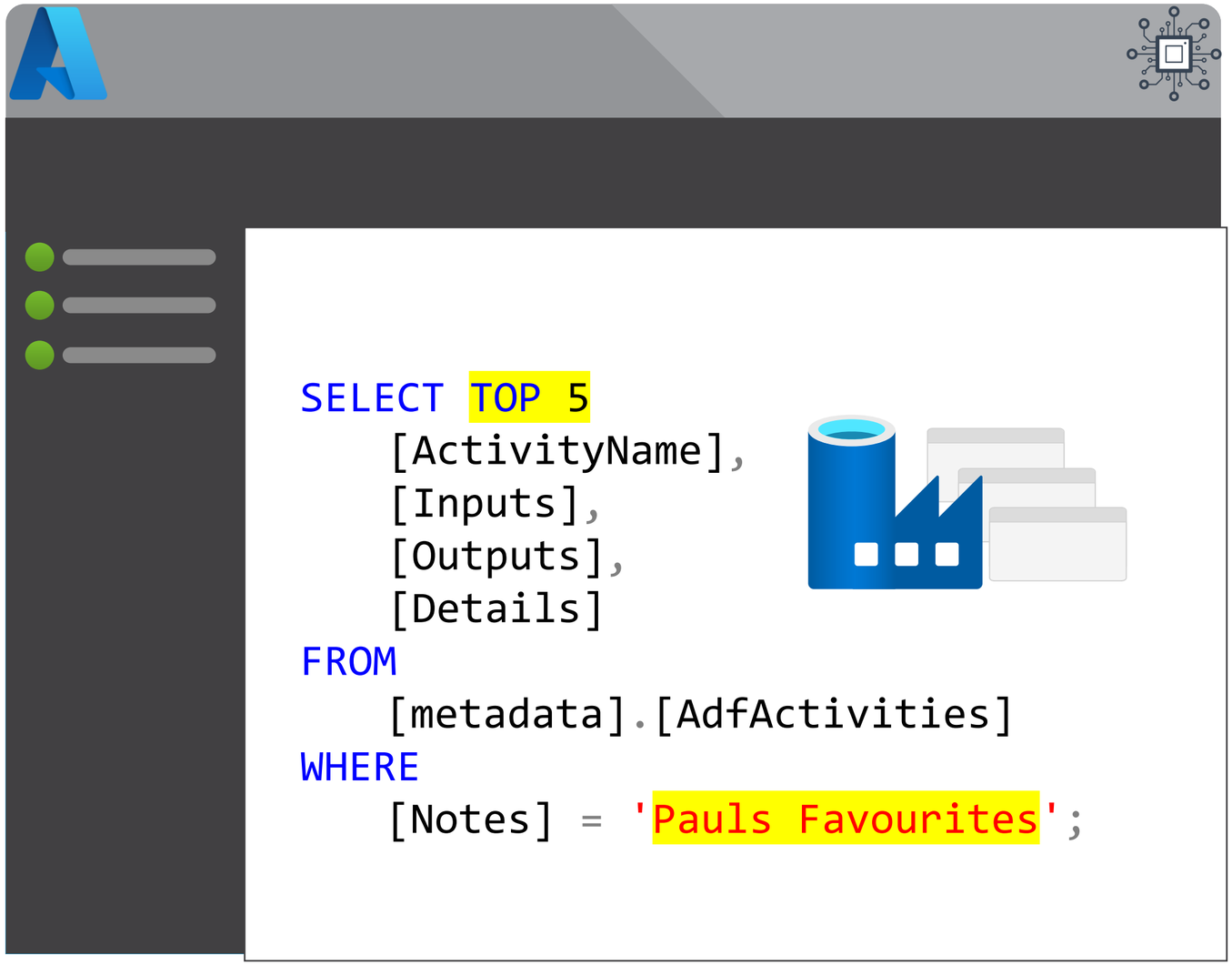
Data Factory Components



1. Linked Services
2. Datasets
3. Activities
4. Pipelines
5. Triggers

Module 1

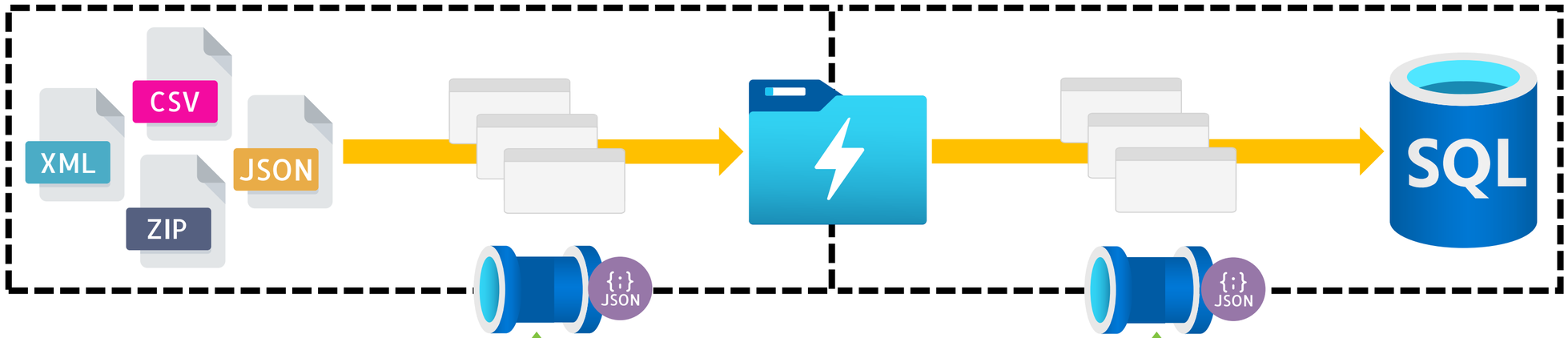
Pipeline Fundamentals



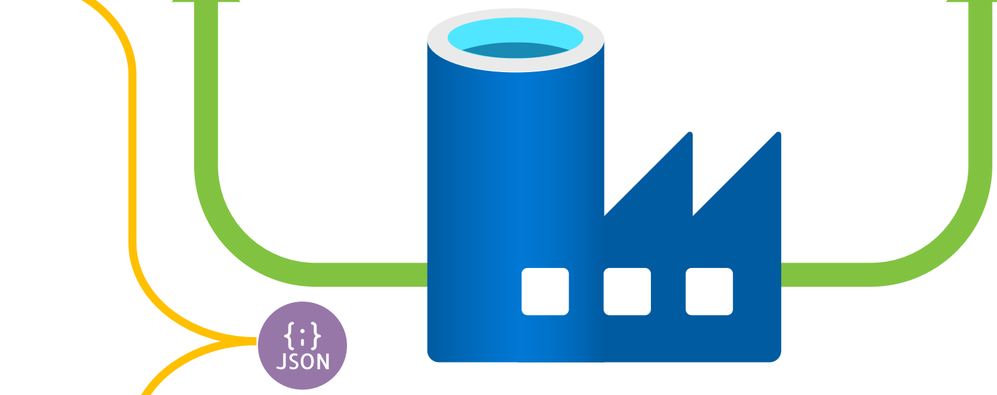
```
SELECT TOP 5
    [ActivityName],
    [Inputs],
    [Outputs],
    [Details]
FROM
    [metadata].[AdfActivities]
WHERE
    [Notes] = 'Pauls Favourites';
```

- The History of Azure Orchestration
- Synapse Analytics vs Data Factory
- Integration Components
- Common Activities
- Execution Dependencies

Data Factory Common Activities

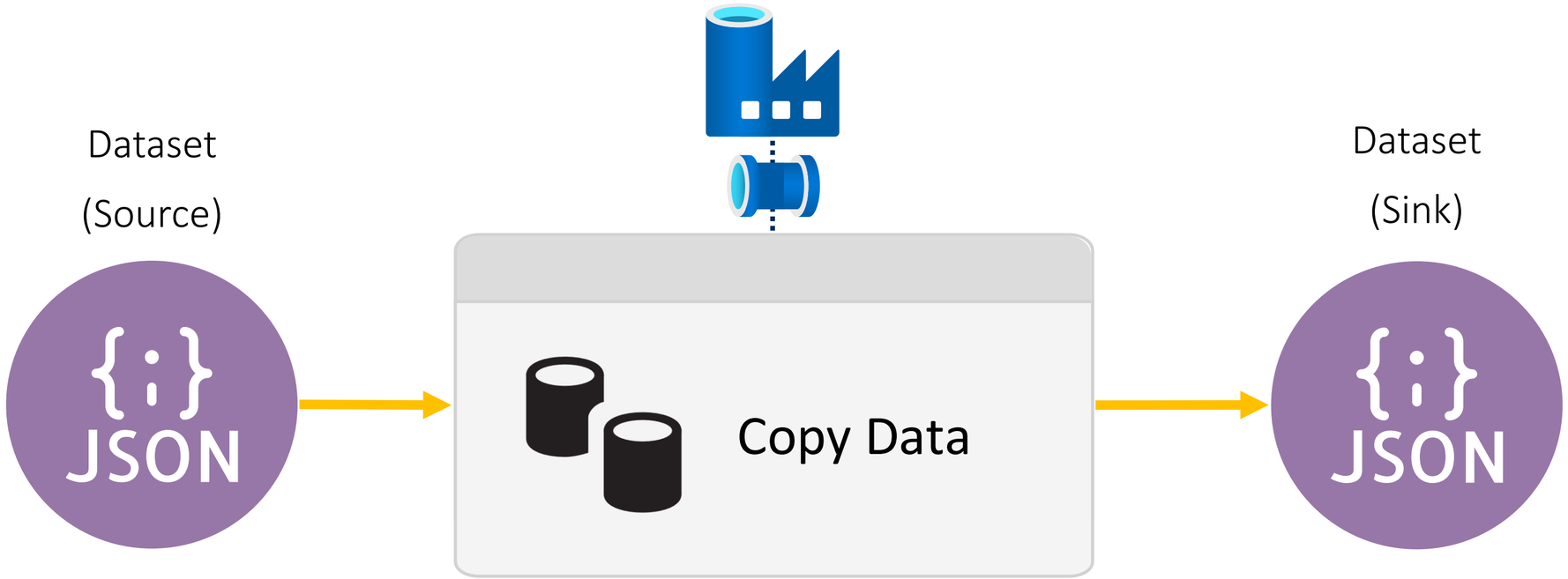


1. Linked Services
2. Datasets
3. Activities
4. Pipelines
5. Triggers



Copy

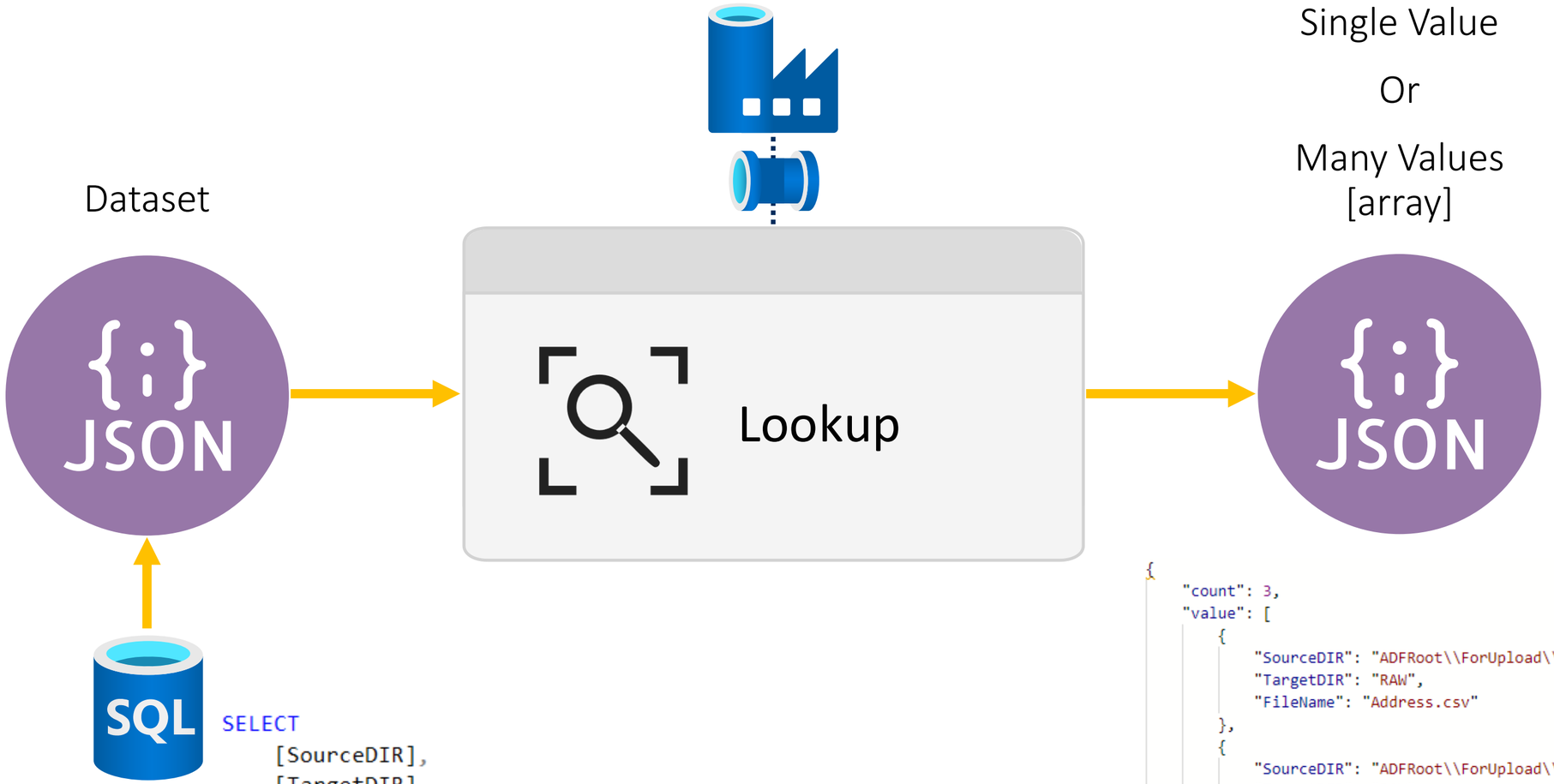
Getting your data from A to B (not a Move operation)



- Auto Scaling
- Transactional Restarts
- Handle Zip Compression
- Attribute Mapping and Schema Drift
- Handle Failed Rows
- Add Custom Attributes
- Parse Excel & JSON Files

Lookup

Get value(s) to support other control flow activities



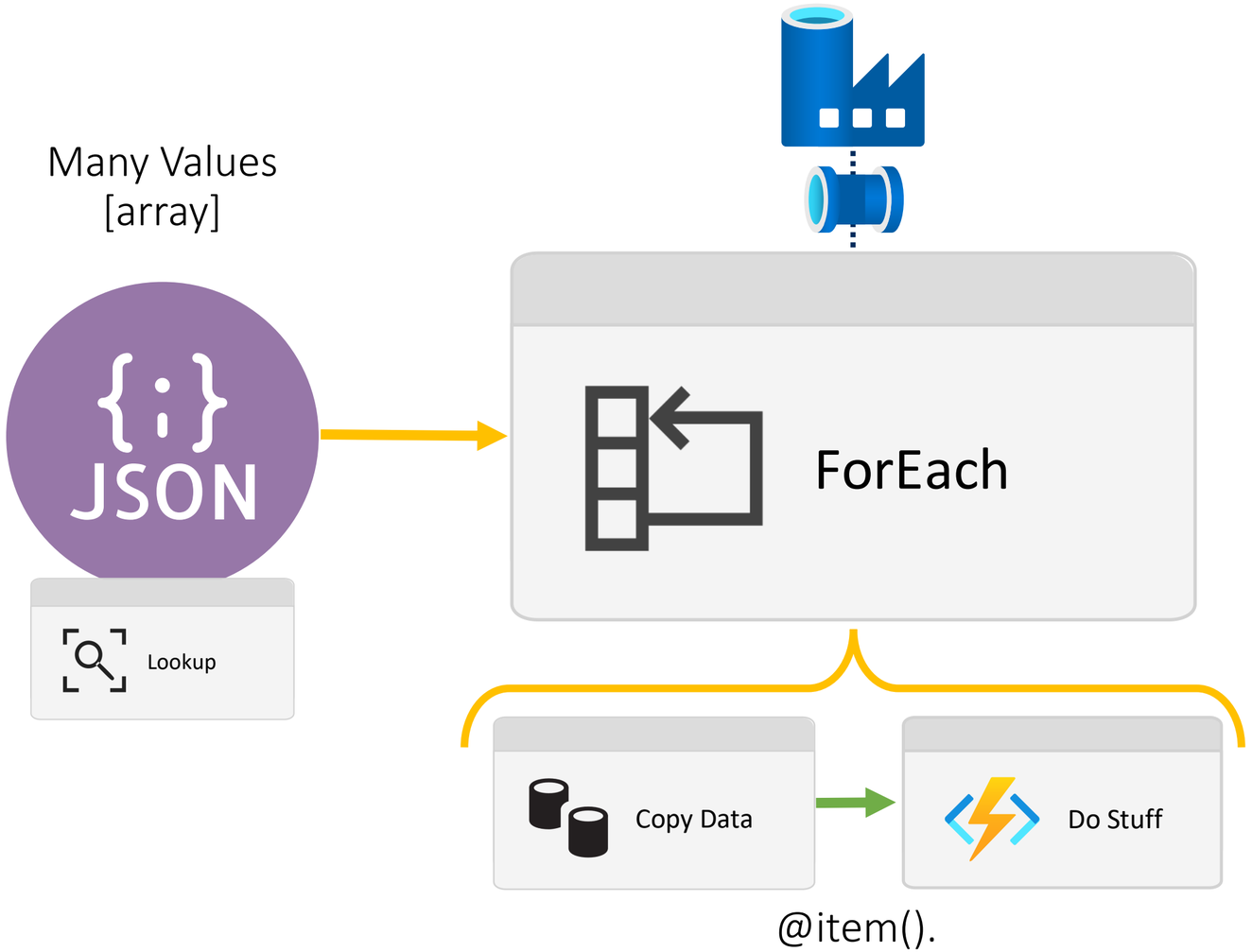

```
SELECT
    [SourceDIR],
    [TargetDIR],
    [FileName]
FROM
    [dbo].[FileList]
```

```
{
    "count": 3,
    "value": [
        {
            "SourceDIR": "ADFRoot\\ForUpload\\People\\",
            "TargetDIR": "RAW",
            "FileName": "Address.csv"
        },
        {
            "SourceDIR": "ADFRoot\\ForUpload\\People\\",
            "TargetDIR": "RAW",
            "FileName": "Gender.csv"
        },
        {
            "SourceDIR": "ADFRoot\\ForUpload\\People\\",
            "TargetDIR": "RAW",
            "FileName": "Ids.csv"
        }
    ]
}
```

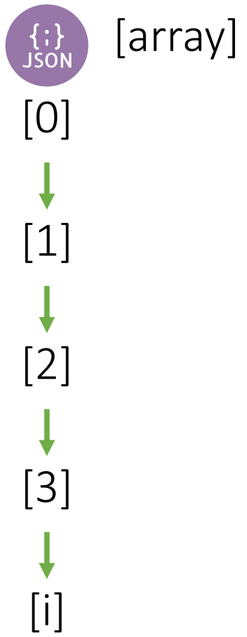
https://docs.microsoft.com/en-us/azure/data-factory/control-flow-lookup-activity

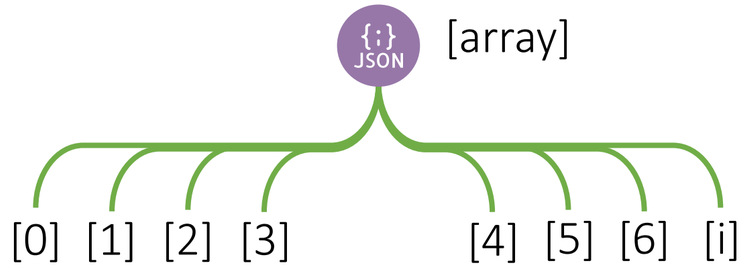
For Each

Iterating over other control flow activities



IsSequential: true





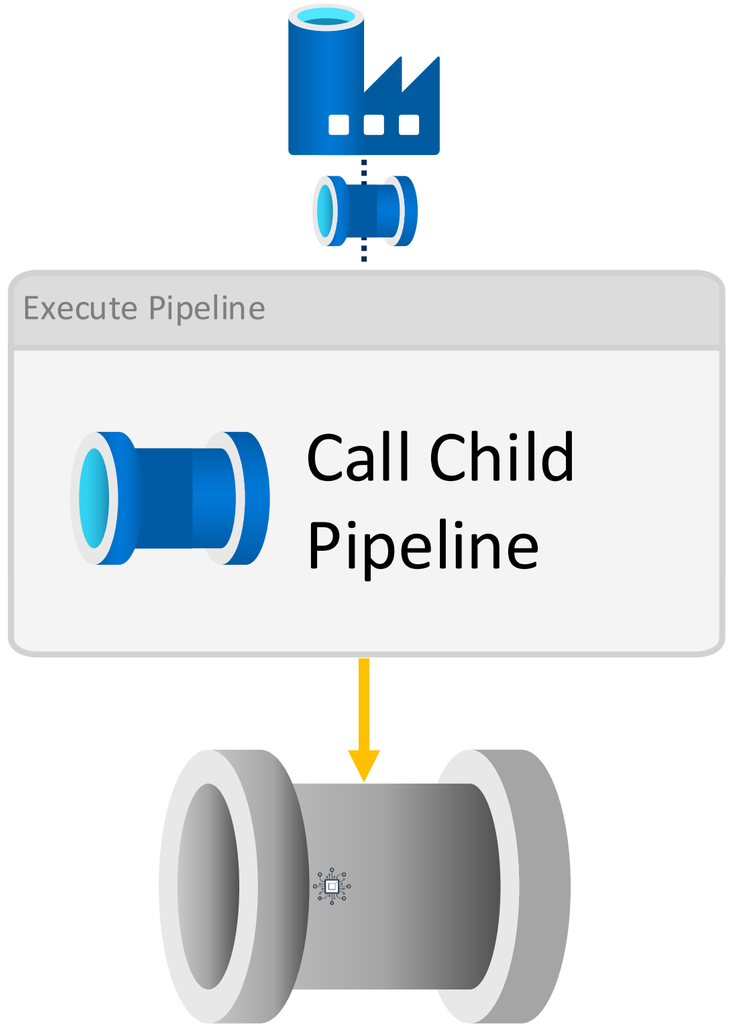
Batch Count Default: 20

Batch Count Max: 50

https://docs.microsoft.com/en-us/azure/data-factory/control-flow-for-each-activity

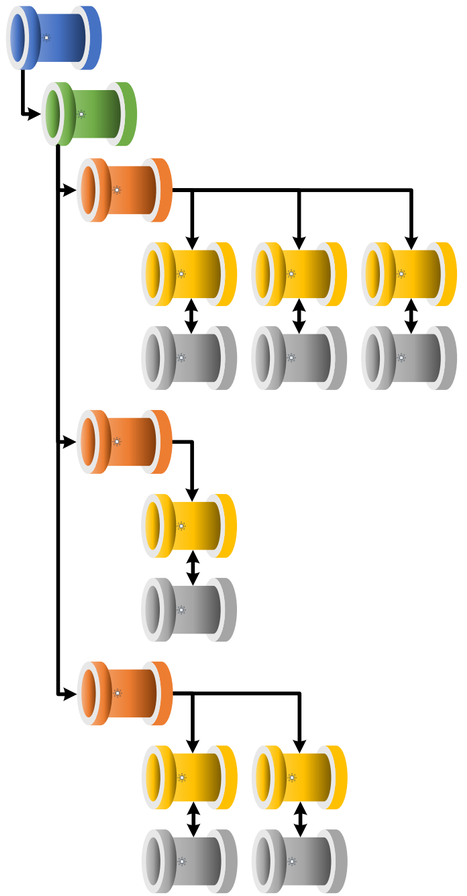
Execute Pipeline

Chaining pipeline executions via an activity



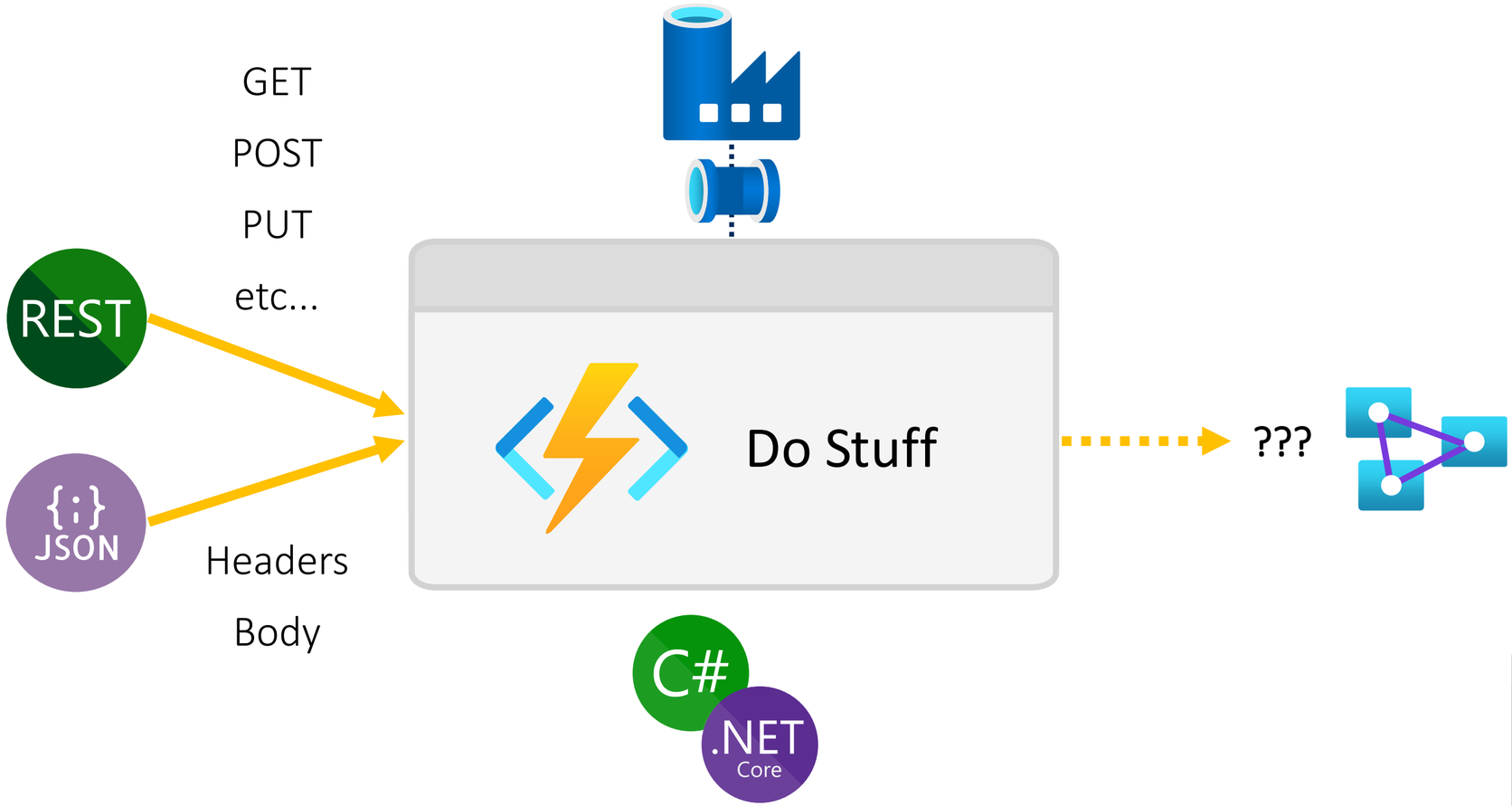
Pipeline Hierarchies Generation Control

https://mrpaulandrew.com/2019/09/25/azure-data-factory-pipeline-hierarchies-generation-control



Azure Function

Extend Data Factory with custom serverless code executions via REST calls



Web

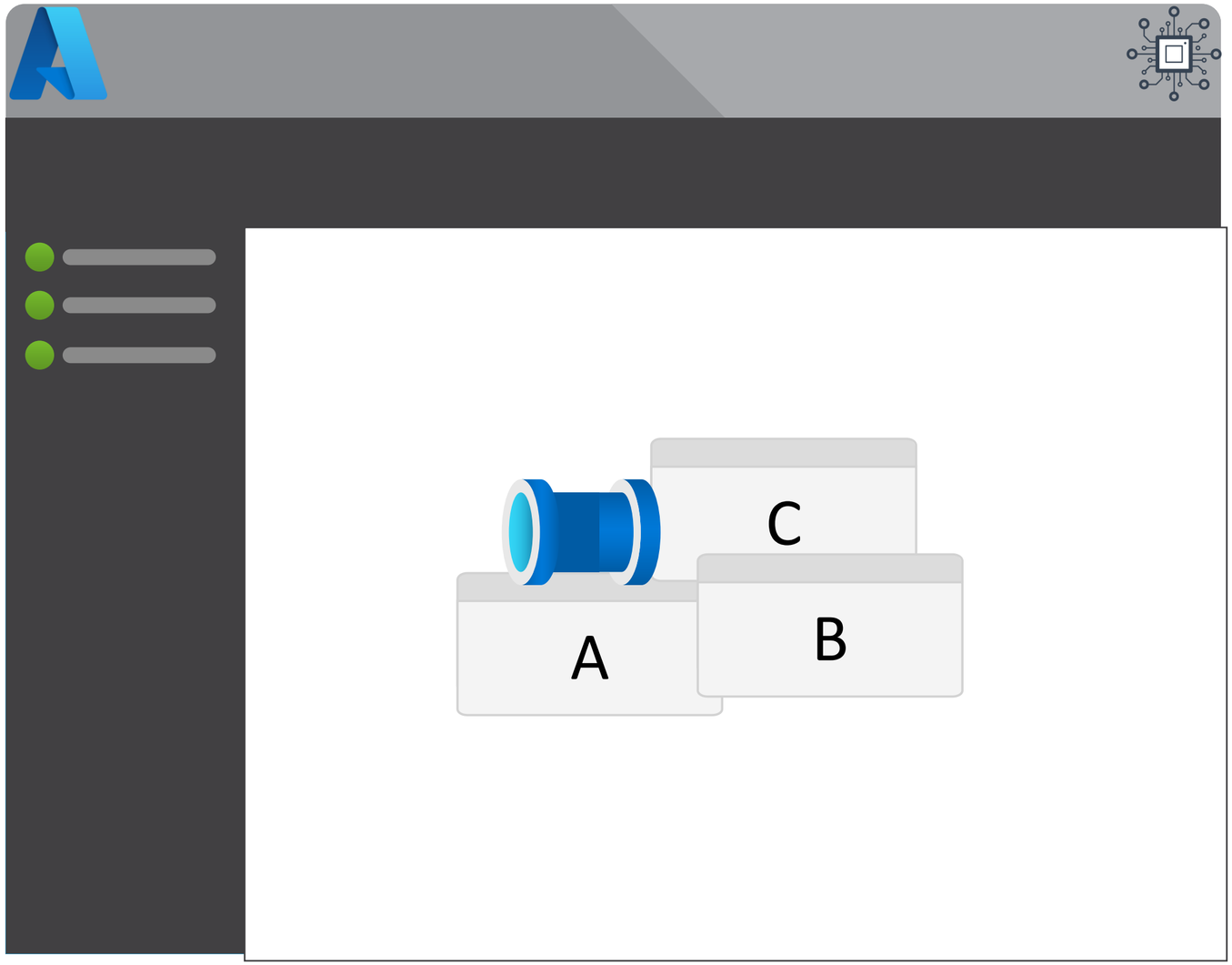
HTTP Do Stuff

Web Hook

Do Stuff

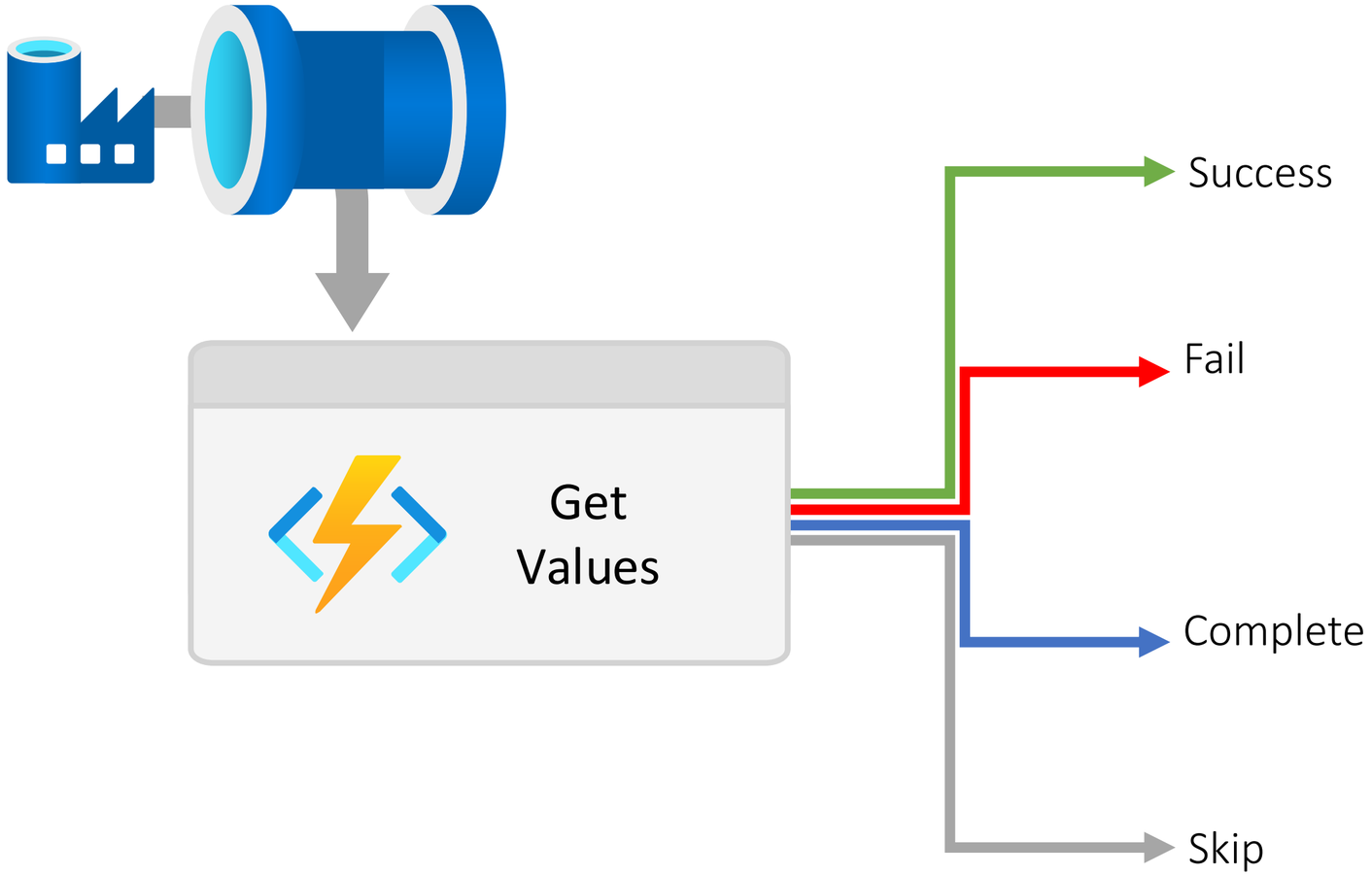
Module 1

Pipeline Fundamentals

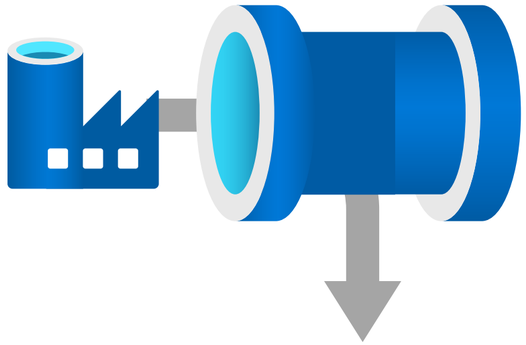


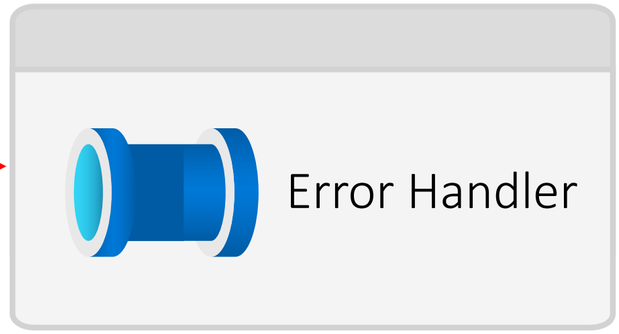
- The History of Azure Orchestration
- Synapse Analytics vs Data Factory
- Integration Components
- Common Activities
- Execution Dependencies

Execution Dependency Options

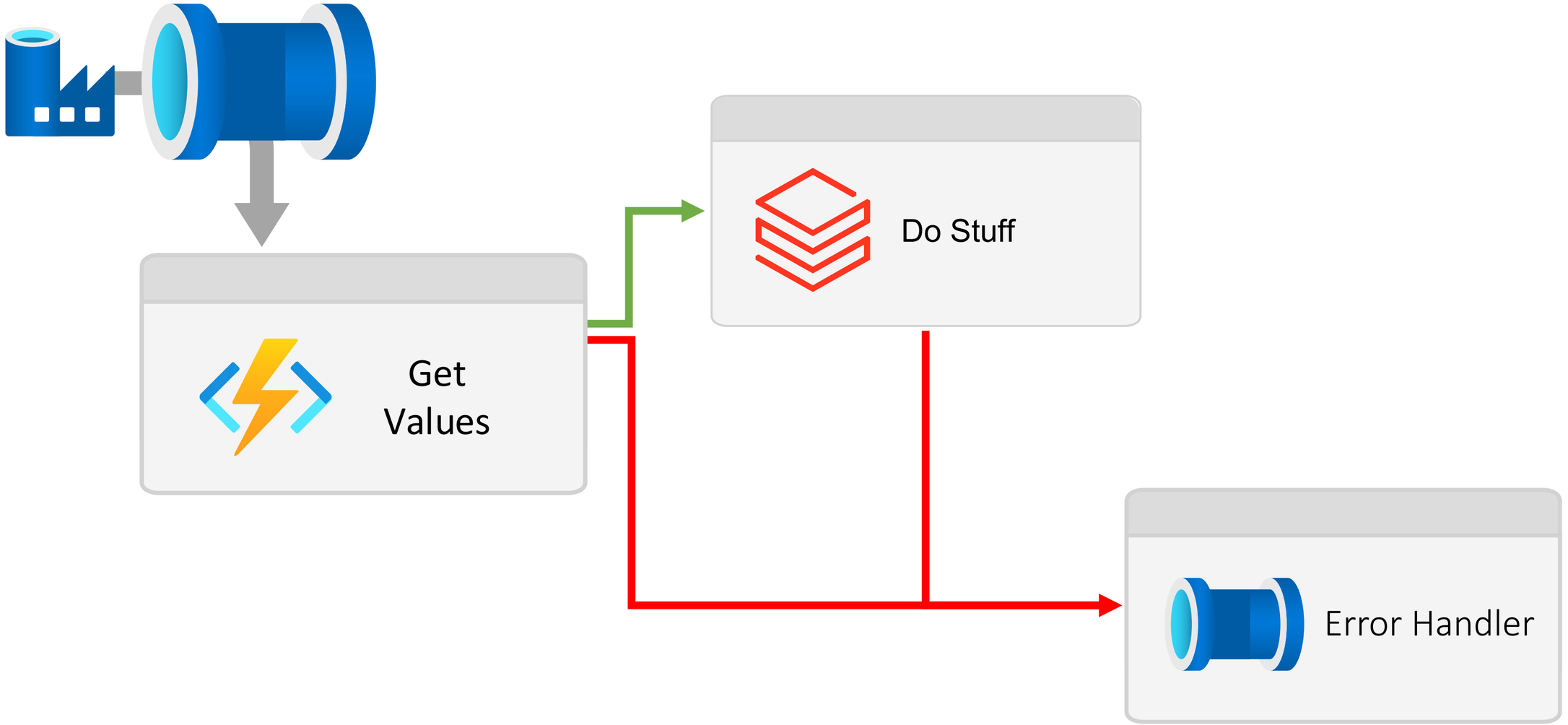


Execution On Failure

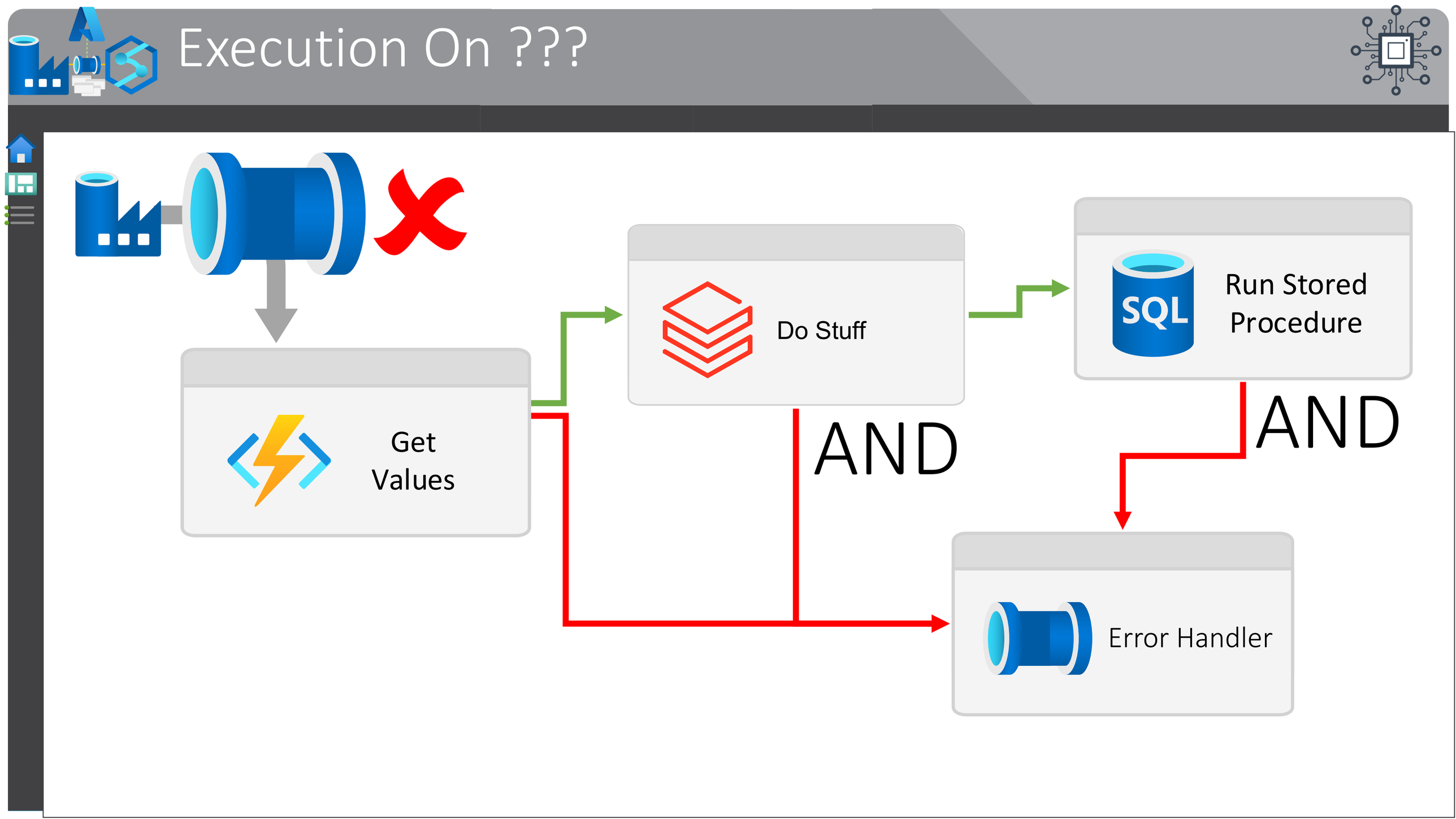




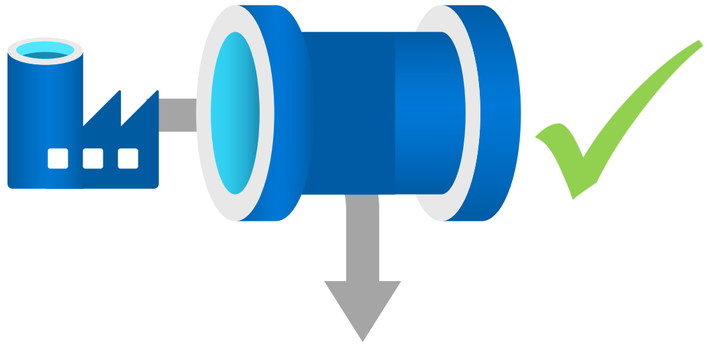
Execution On Failure or On Success



Execution On ???



Execution On Failure or On Success



Get Values

OR

Error Handler

Do Stuff

OR

Error Handler

SQL

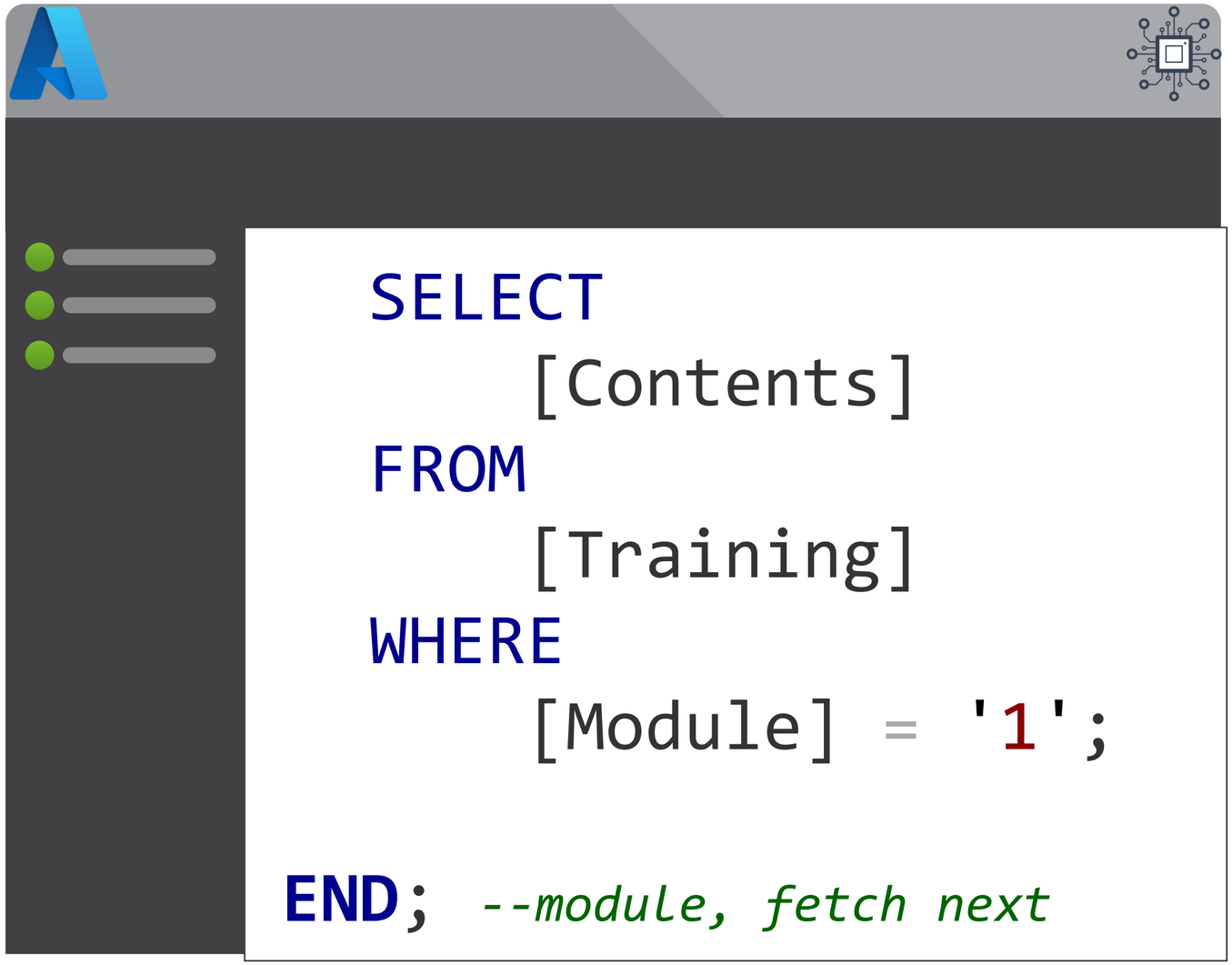
Run Stored Procedure

OR

Error Handler

Module 1

Pipeline Fundamentals



```
    SELECT
        [Contents]
    FROM
        [Training]
    WHERE
        [Module] = '1';

END; --module, fetch next
```

- The History of Azure Orchestration
- Synapse Analytics vs Data Factory
- Integration Components
- Common Activities
- Execution Dependencies