Building an Azure Business Intelligence Solution End to End

Hands On Workshop

Paul Andrew | Senior Consultant

Terry McCann | Principal Consultant

Simon Whiteley | Cloud Architect

adatis

Microsoft Partner | Gold Data Analytics / Gold Data Platform / Gold Cloud Platform

Microsoft

https://github.com/Adatis

ModernDataWarehouseWorkshop

adatis

Agenda for the Day

Module 1

Microsoft Azure

Module 2

Storage
Uploading Data
Data Lake

Module 3

Real-time Data
Streaming
Power BI

Module 4

U-SQL - Data
Transformation
Basics

Module 5

USQL - Advanced
Analytics
Cognitive Services

Module 6

Data Factory
Orchestration
Dynamic Pipelines

Module 7

Data Presentation
& Consumption
Power BI Models

Module 8

Other Services
Q&A

Agenda

Data Factory Recap

Concepts
Components

Design Patterns
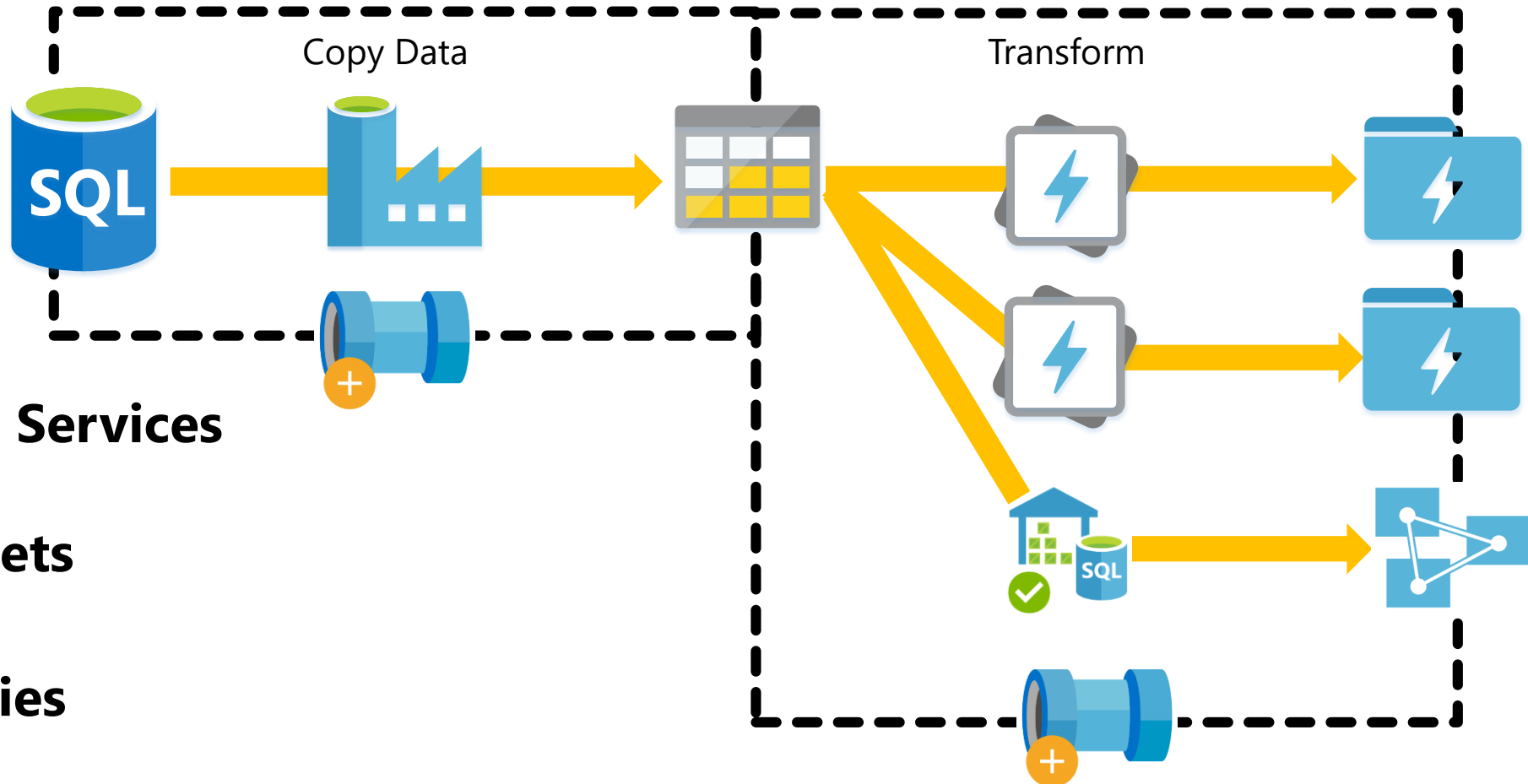
Meta Data Driven Pipelines

Agenda

Data Factory Recap

Concepts
Components

Design Patterns

Meta Data Driven Pipelines

Data Factory Components

Copy Data

Transform

SQL

1 **Linked Services**
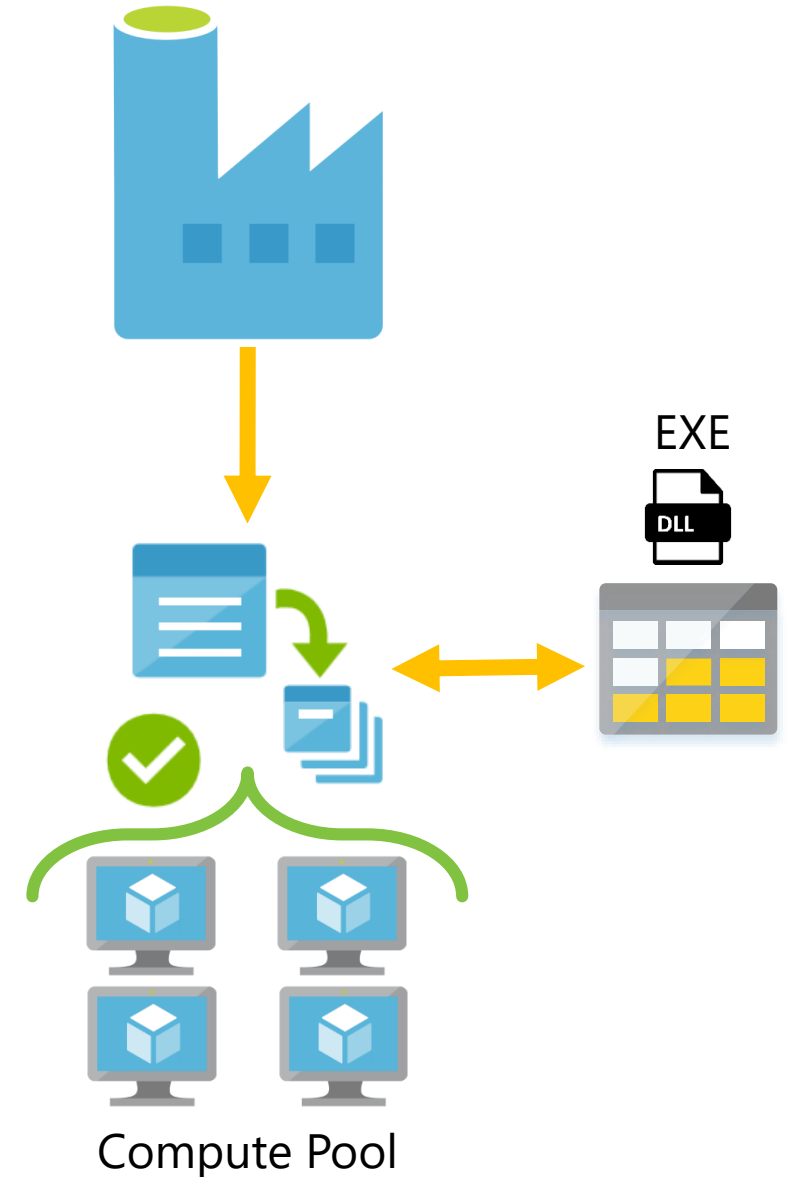
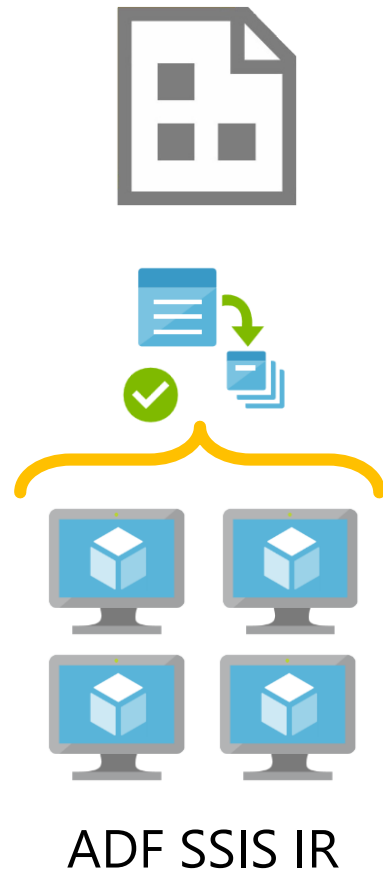2 **Data Sets**

3 **Activities**

4 **Pipelines** – What groups of work do I want to do?

ADF Extensibility Continued

1. **Custom Activities** – Run an Console App

2. **Rest API Calls** – Web Activities

3. **SSIS** – Packages
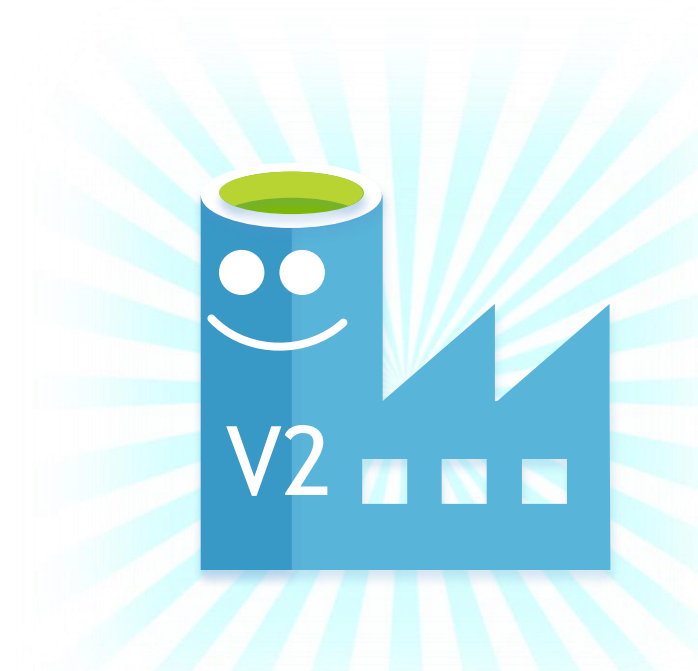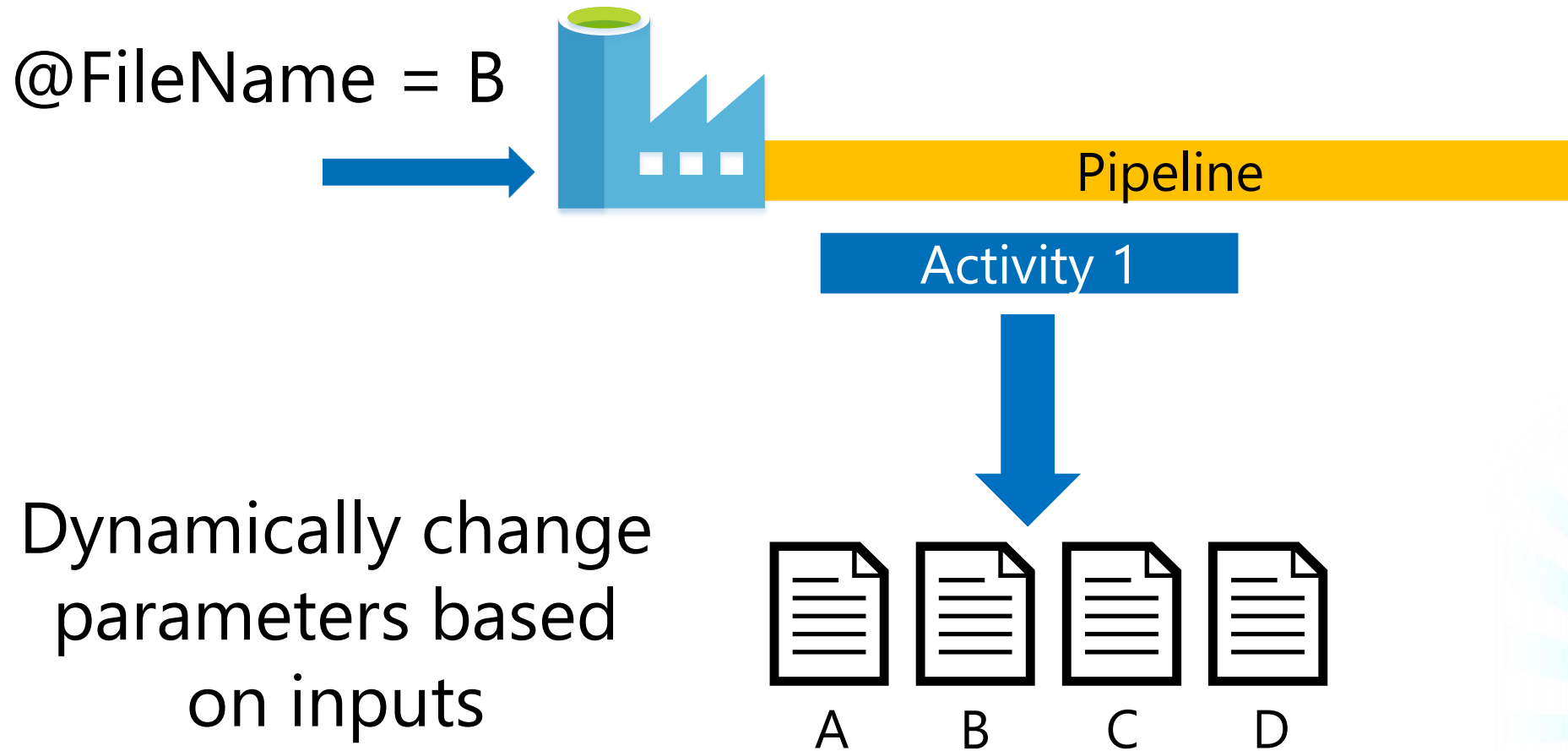
ADF SSIS IR

EXE

Compute Pool

Agenda

Data Factory Recap

Concepts
Components
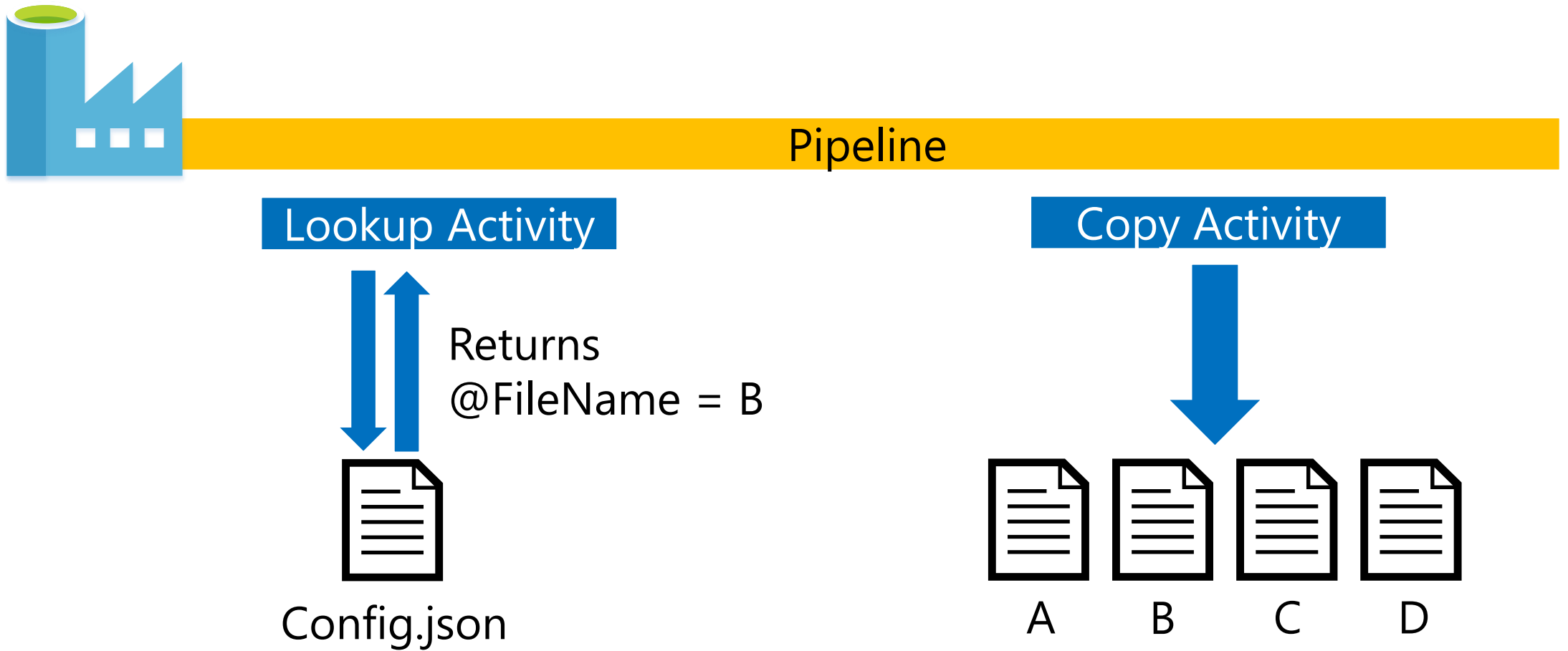
Design Patterns

Meta Data Driven Pipelines

Dynamic Pipelines using Parameters & Expressions

@FileName = B

Pipeline

Activity 1

Dynamically change parameters based on inputs

A   B   C   D

V2

Dynamic Pipelines using Lookup Activity

For Each Pipelines

@Files = [A,B,C]

Pipeline

ForEach Activity

Run activity for each item in an array

(max DOP 20)

A Copy Activity

B Copy Activity

C Copy Activity
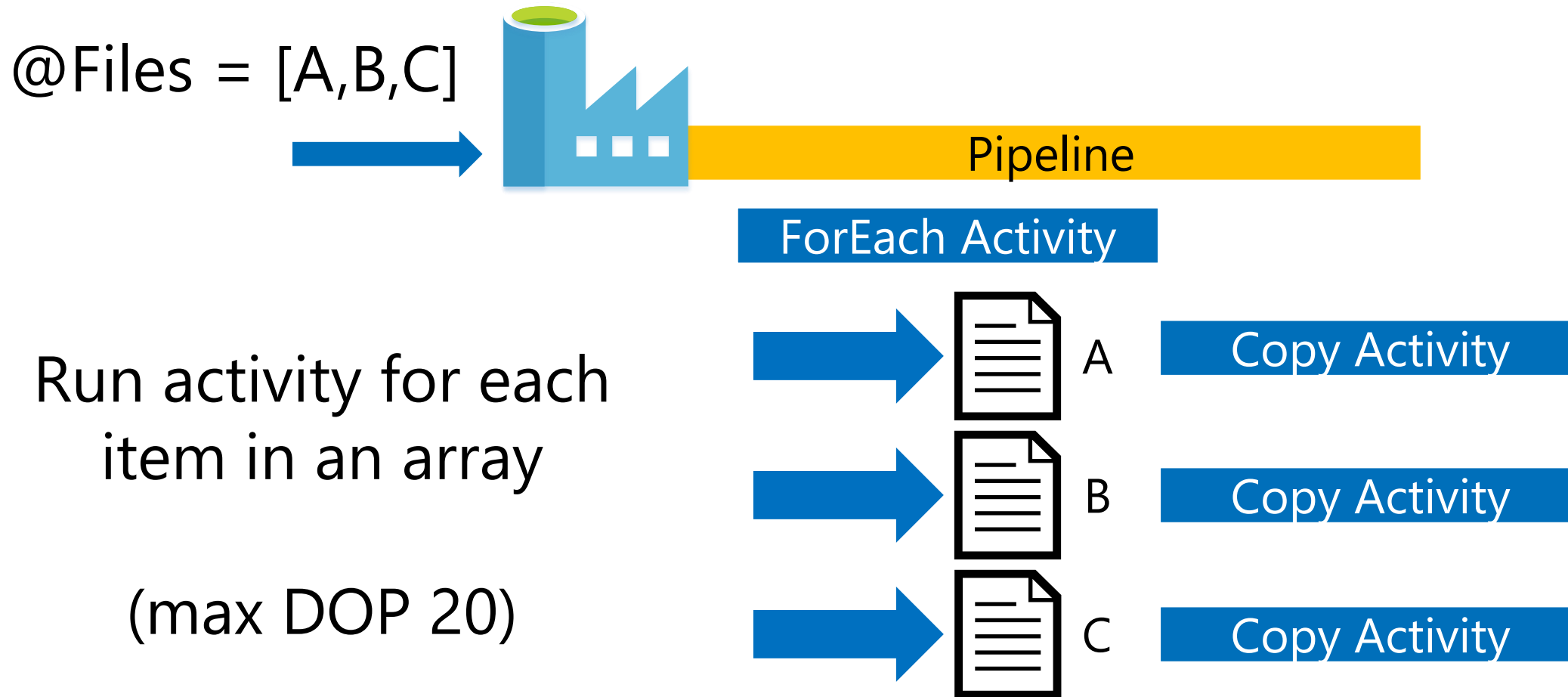
Parent/Child Pipelines & Triggering

Event Trigger

Time Slices

Wall Clock

Parent Pipeline

Exec Pipeline

Child Pipeline

Child Pipeline
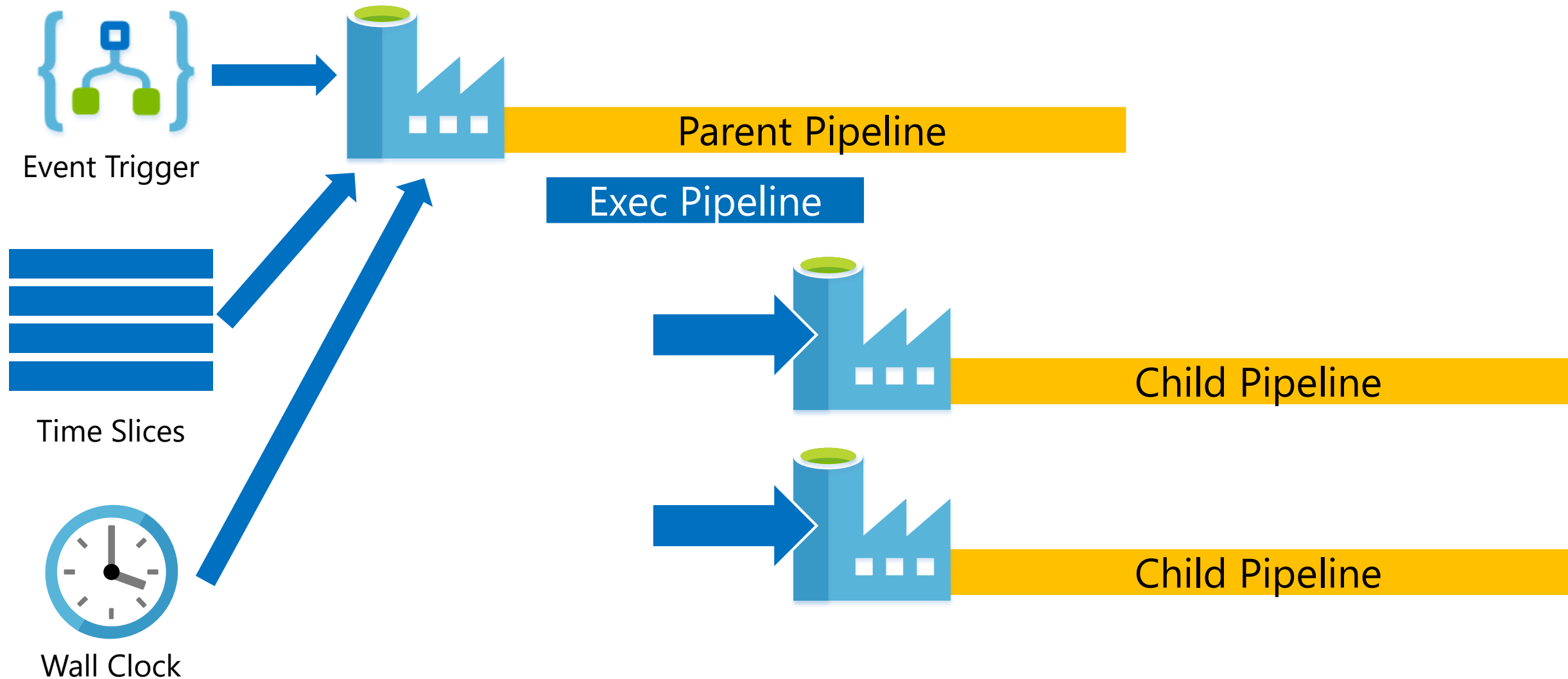
The SSIS IR with Azure V-Net Access

Event Driven Loading

1 File added to Network Folder

3 Logic App triggers Data Factory event. Provides filename as parameter

Event
Driven
Loading

Pipeline 1

2 Logic App watches folder, triggering at the next polling interval

4 Pipeline runs, processing file as soon as it is available. Only one ADF Pipeline required
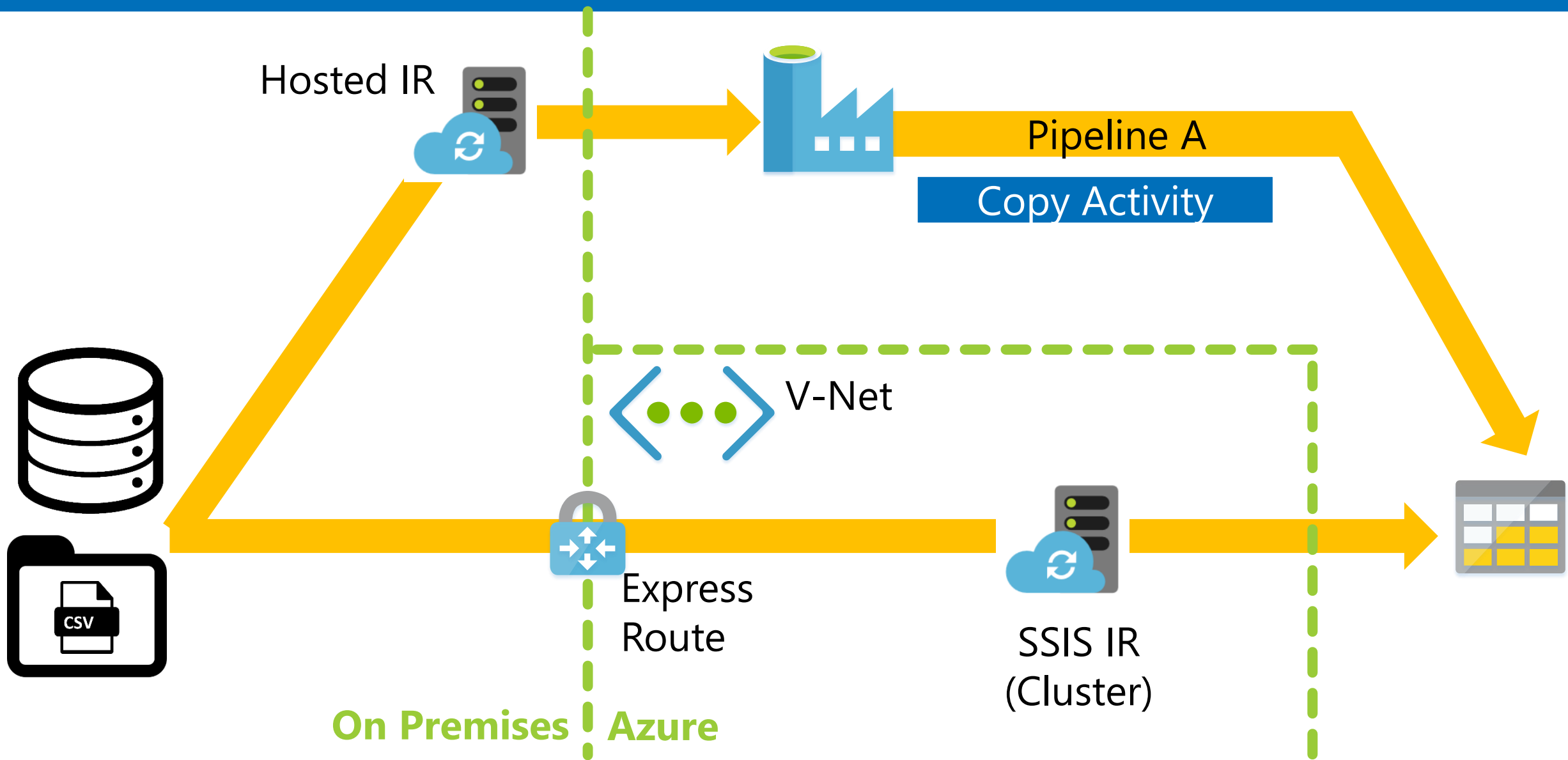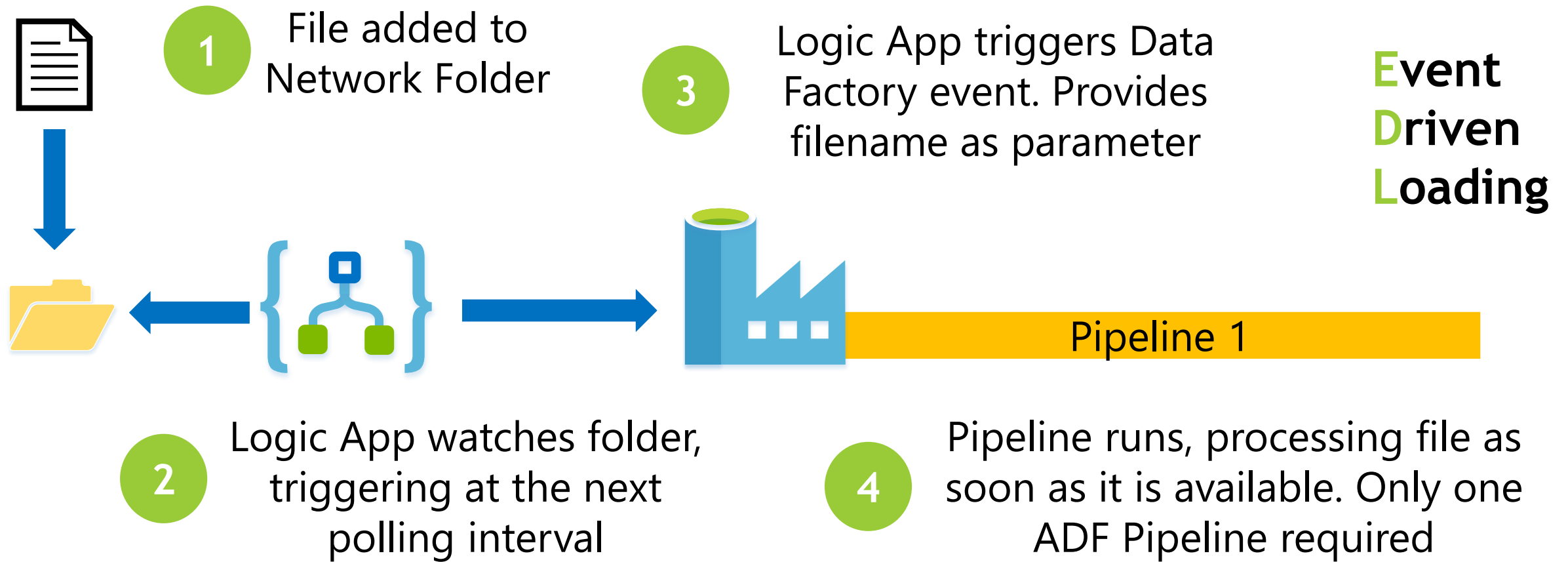
Reusable Pipelines with Conditional Logic
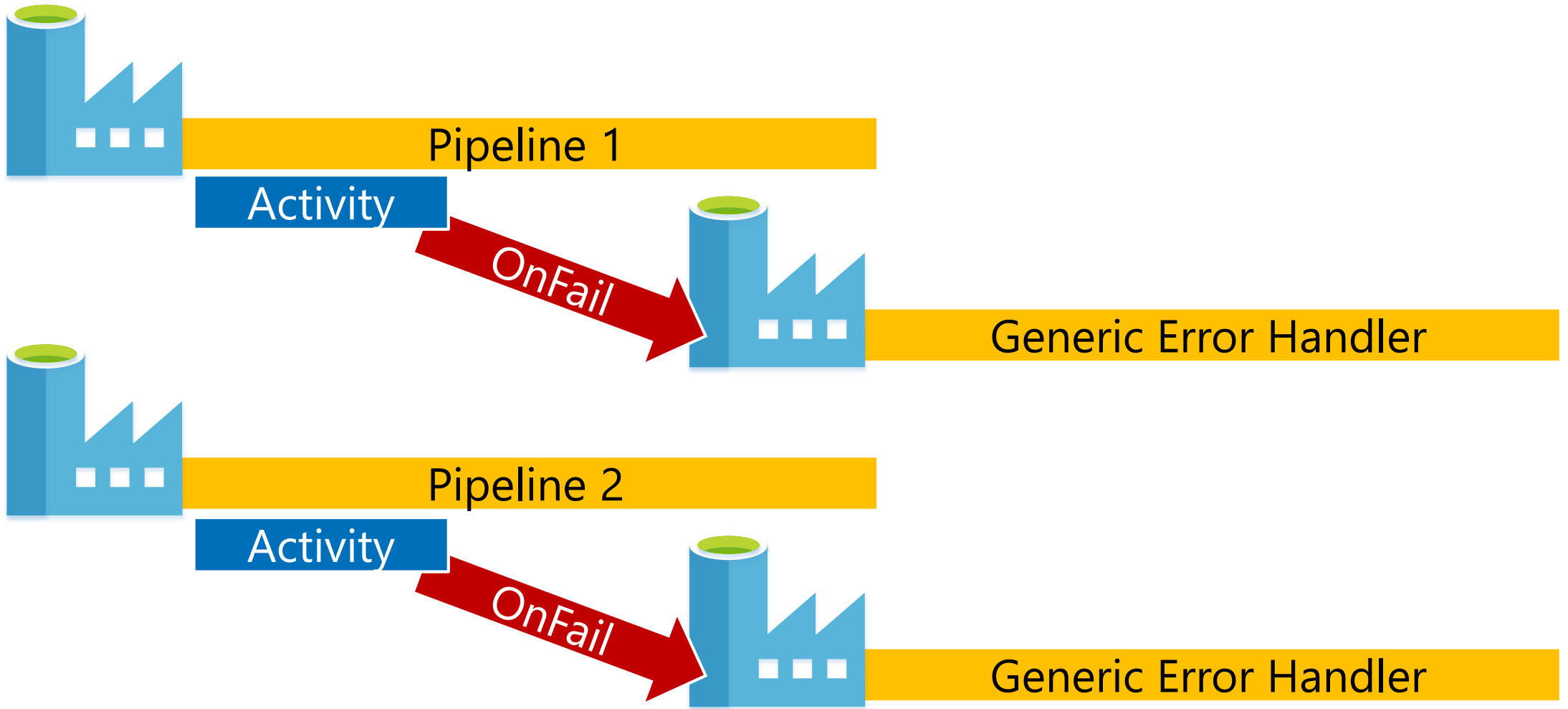
Design Pattern Combinations

2 Pipeline Triggered

Pipeline A

Lookup Activity

ForEach Activity

3

Lookup reads list of files into parameter

Pipeline B

Pipeline B

1 List of files to process written to json doc

Pipeline B

4 Child pipeline triggered for each file in lookup parameter. Up to 20 concurrent executions

Agenda for the Day

Module 1 Microsoft Azure	**Module 2** Storage Uploading Data Data Lake	**Module 3** Real-time Data Streaming Power BI	**Module 4** U-SQL - Data Transformation Basics
Module 5 USQL - Advanced Analytics Cognitive Services	**Module 6** Data Factory Orchestration Dynamic Pipelines	**Module 7** Data Presentation & Consumption Power BI Models	**Module 8** Other Services Q&A