Azure Data Factory Deep Dive

Jason Horner & Cathrine Wilhelmsen

SQLBits · February 27th 2019

Abstract

This immersive full day hands on session will arm you with all the skills you need to build and orchestrate pipelines using Azure Data Factory v2 that solve both traditional batch data integration problems and also adapt to newer challenges such as real-time and complex event processing.

We will start with a conceptual overview, and then get into how you can use metadata to build scalable serverless pipelines to move data from disparate data sources including On-Premises and Platform As A Service. We will look at how to leverage SSIS Lift and shift to maximize reuse for your existing code base and review the pros and cons of this approach. We will dive deep into the newly released Data Flow capabilities and cover the design patterns and architectures for modern Data Warehouses and big data pipelines. We will also review best practices for operationalizing solutions including configuring monitoring, logging and alerting. And did I mention DevOps? We will implement DevOps best practices like CI/CD within the data factory.

This day is everything you need to do amazing things with Azure Data Factory.

Jason Horner

International Man of Leisure

@jasonhorner jasonhorner.com

/jasonhorner jason@jasonhorner.com

I Do Stuff

SQL Server DW / OLTP, Azure, MDM, PowerShell, Geospatial Data, Advanced Data and Analytics, SSIS, SSAS

I Know Stuff

MVP

Microsoft CERTIFIED Master SQL Server® 2008

Microsoft CERTIFIED Solutions Expert Business Intelligence Data Platform

Papa Needs Brand New Shoes

Always looking for new meaningful and interesting work: jason@jasonhorner.com

http://www.jasonhorner.com/hireme.html

Cathrine Wilhelmsen

Master Yak Shaver

@cathrinew

hi@cathrinew.net

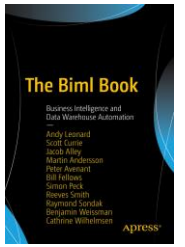
/cathrinewilhelmsen

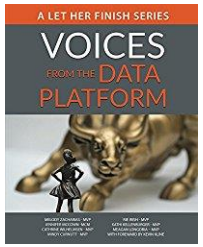
cathrinew.net

Work Things

Business Intelligence Consultant in Inmeta, focusing on the Microsoft Data Platform.
Specialties: ETL, SSIS, Biml and T-SQL

Geeky Things





Fun Things

Active in the PASS and Biml communities as an author, speaker, blogger and chronic volunteer.
Likes: coffee, chocolate and cat gifs :)

@jasonhorner
@cathrinew

#SQLBits
#AzureDataFactory

Part 1: Azure Data Factory Overview

Concepts and Objects

Naming Rules and Conventions

Security

Pricing

Part 2: Basic Azure Data Factory

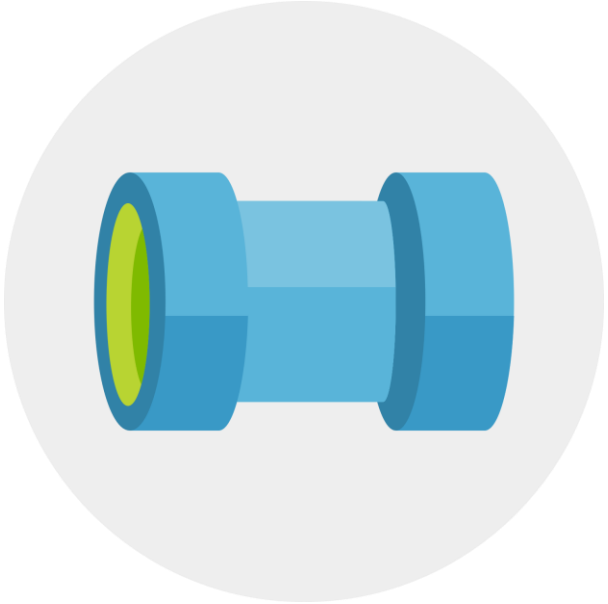
User Interface and Navigation

Pipelines

Annotations

Basic Activities

Part 3: Advanced Azure Data Factory



Dynamic Content

Variables

Iteration and Conditionals

Advanced Activities

Part 4: Design Patterns

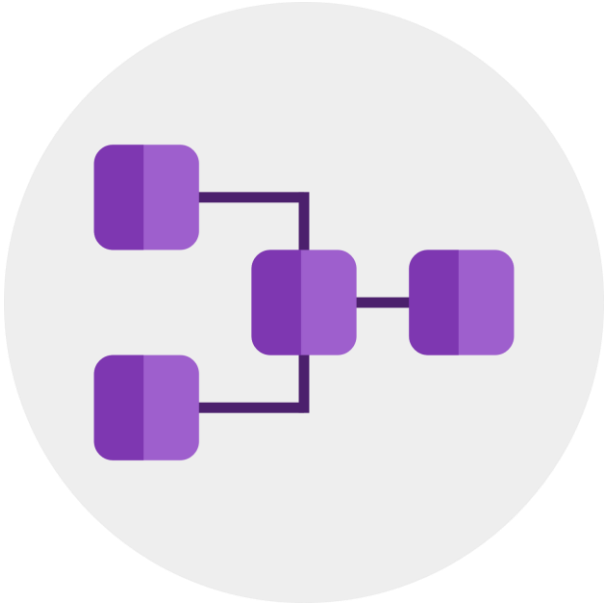
Pipeline Templates

Architectural Patterns

Load Patterns

Auditing and Logging Patterns

Part 5: Mapping Data Flows



Transformations

Schema Drift

Performance

Test and Debug

Expression Language

Part 6: SSIS in Azure Data Factory

Integration Runtime

Deployment and Execution

Scheduling

Part 7: Operationalizing Solutions

Source Control and ARM Templates

CI / CD and DevOps

Logging, Monitoring and Alerting

09:00 - 09:15	Welcome
09:15 - 10:15	Precon
10:15 - 11:00	Lab & Coffee Break
11:00 - 12:30	Precon
12:30 - 13:30	Lab & Lunch
13:30 - 15:00	Precon
15:00 - 15:45	Lab & Coffee Break
15:45 - 16:45	Precon
16:45 - 17:00	Wrap Up

Jason Horner (jason@jasonhorner.com) & Cathrine Wilhelmsen (contact@cathrinewilhelmsen.net)

Resources

Slide Deck:

https://github.com/jasonhorner/adfdeepdive/tree/master/Powerpoint

Labs:

https://github.com/jasonhorner/adfdeepdive/tree/master/Labs

Thank you!

Jason Horner, *Attunix*

jason@jasonhorner.com

@jasonhorner

Cathrine Wilhelmsen, *Inmeta*

hi@cathrinew.net

@cathrinew