Data Factory & Synapse Pipelines Best Practices

Orlando Gomes

Cloud Solution Architect (Engineering) - Microsoft

Datasaturdaybr

<> A SQL SATURDAY EVENT

Sp_help(palestrante)

Orlando Gomes

- Cloud Solution Architect (Engineering) - Microsoft
- + 8 Years Experience - Microsoft BI Stack
- Microsoft MVP Reconnect
- Microsoft Certified Professional

- Pokemon & Soccer Lover

Gratidão - Patrocinadores!

Data Factory

Code-Free ETL as a service

Ingest

- Multi-cloud and on-premise hybrid copy data
- 100+ native connectors
- Serverless and auto-scale
- Use wizard for quick copy jobs

Control Flow

- Design code-free data pipelines
- Generate pipelines via SDK
- Utilize workflow constructs: loops, branches, conditional execution, variables, parameters, ...

Data Flow

- Code-free data transformations that execute in Spark
- Scale-out with Azure Integration Runtimes
- Generate data flows via SDK
- Designers for data engineers and data analysts

Schedule

- Build and maintain operational schedules for your data pipelines
- Wall clock, event-based, tumbling windows, chained
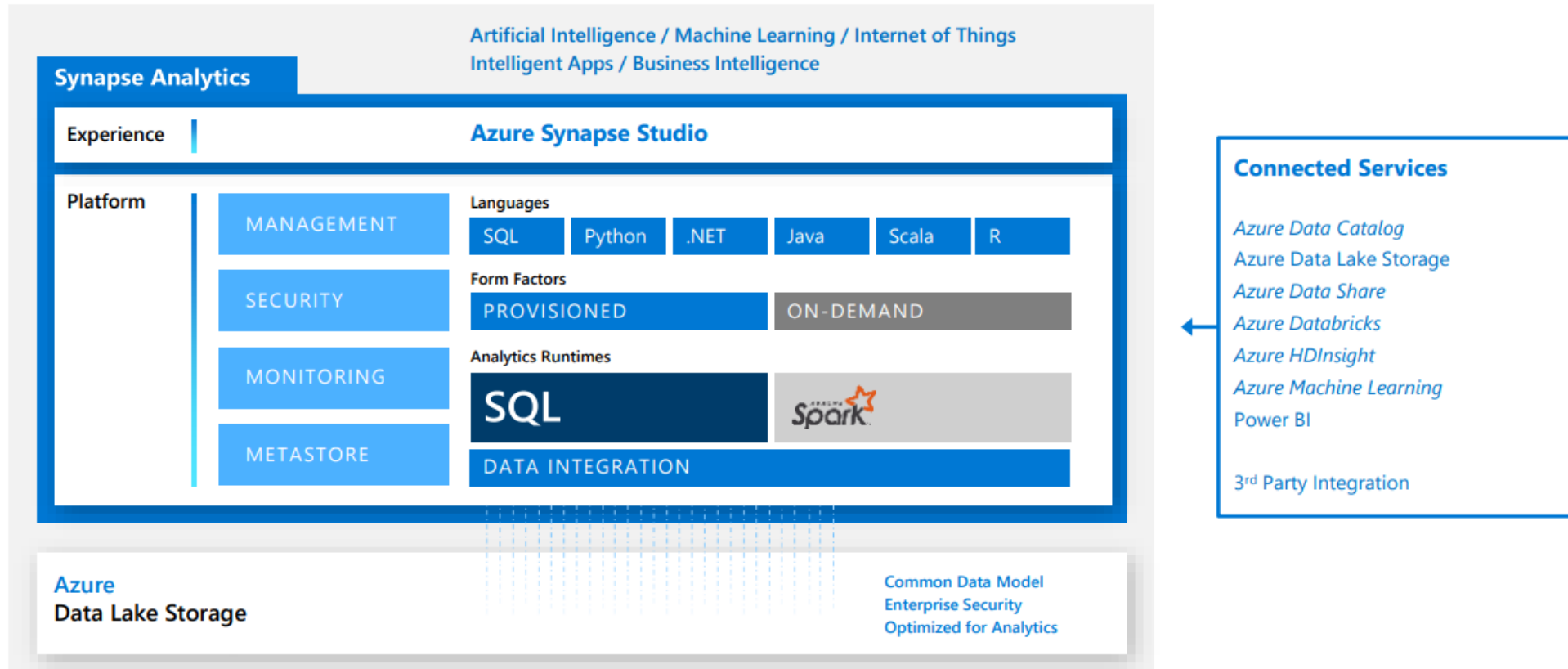
Monitor

- View active executions and pipeline history
- Detail activity and data flow executions
- Establish alerts and notifications

[Introduction to Azure Data Factory - Azure Data Factory | Microsoft Learn](#)

Datasaturdaybr
A SQL SATURDAY EVENT

Azure Synapse Analytics

[What is Azure Synapse Analytics? - Azure Synapse Analytics | Microsoft Learn](#)

Ok...But...

When use
Data Factory -
Synapse Pipelines ?

Depends…

Ask yourself, customer, team, manager...

- How mature is my orgaization mature about Cloud – DW Process ?
- What I have today to attend orchestration ?
- Where my Data Sources are located ?
- What resources I need – wait from a ETL Cloud Based Tool ?
- I have a Data Factory, should I migrate to Synapse Pipelines ?
  (Time, Cost, Compatibility, Resources...)
- If yes, what Synapse resources I will use ? (justify go only because Pipelines ?)
- I have my process on Synapse Pipelines, should I migrate to Data Factory ?
  Is there some specifc resource I need there ?
- How Much ? $

Synapse Pipelines Data Factory Features

Category	Feature	Azure Data Factory	Azure Synapse Analytics
Integration Runtime	Using SSIS and SSIS Integration Runtime	✓	✓ *Public preview*
	Support for Cross-region Integration Runtime (Data Flows)	✓	X
	Integration Runtime Sharing	✓ *Can be shared across different data factories*	X
Pipelines Activities	SSIS Package Activity	✓	✓ *Public preview*
	Support for Power Query Activity	✓	X
	Support for global parameters	✓	X
Template Gallery and Knowledge center	Solution Templates	✓ *Azure Data Factory Template Gallery*	✓ *Synapse Workspace Knowledge center*
GIT Repository Integration	GIT Integration	✓	✓
Monitoring	Monitoring of Spark Jobs for Data Flow	X	✓ *Leverage the Synapse Spark pools*

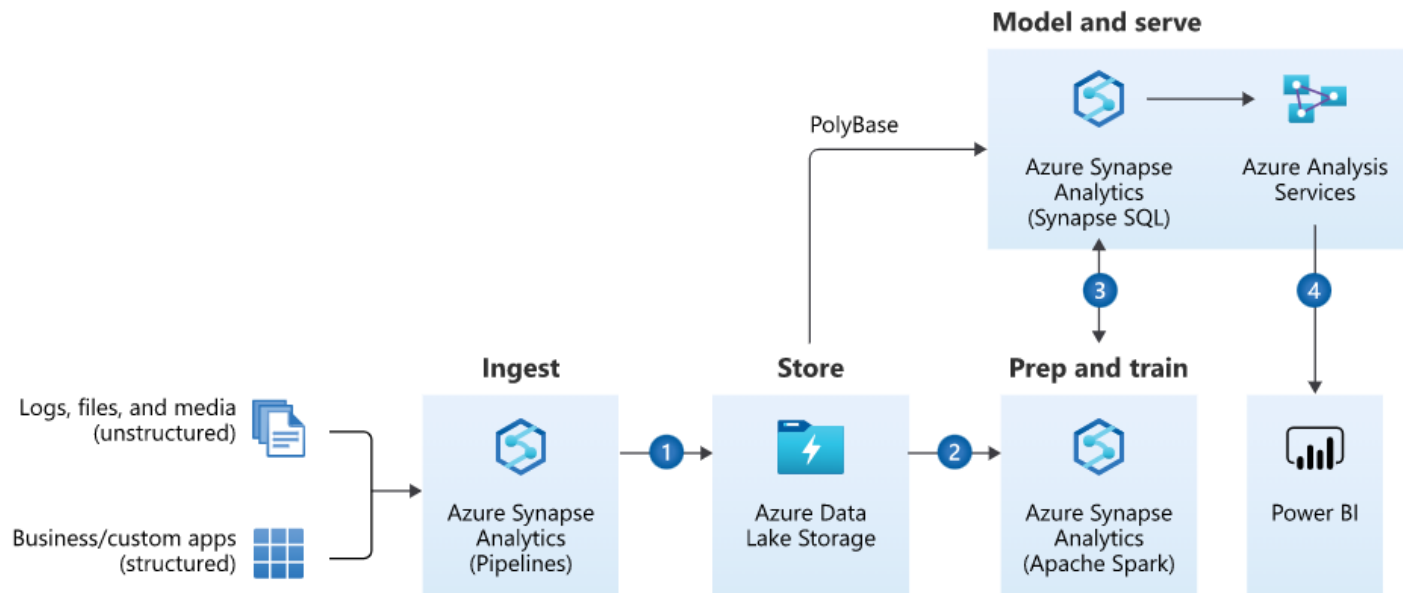[Differences from Azure Data Factory - Azure Synapse Analytics | Microsoft Learn](#)

(ETL Process) What I should care...

- Security;
- Performance;
- Network (Bandwith, Firewall);
- Identity;
- Productivity;
- Naming Standard (Datasets, Linked Services);
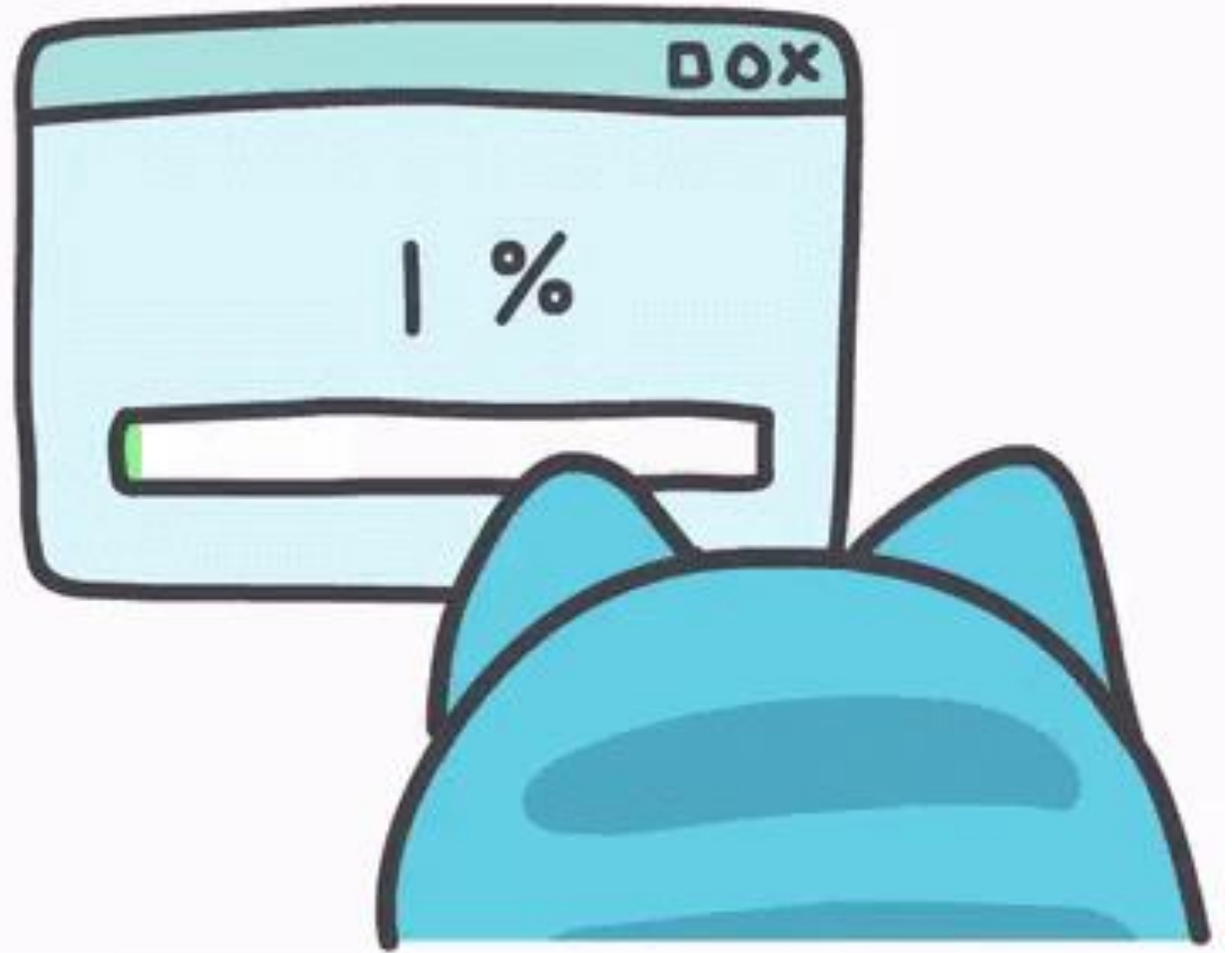- Collaboration.

Enterprise Data Warehouse

Demo

16:15 → 45 min

Extrair conteúdo Dinamicamente com SQL no Azure Data Factory
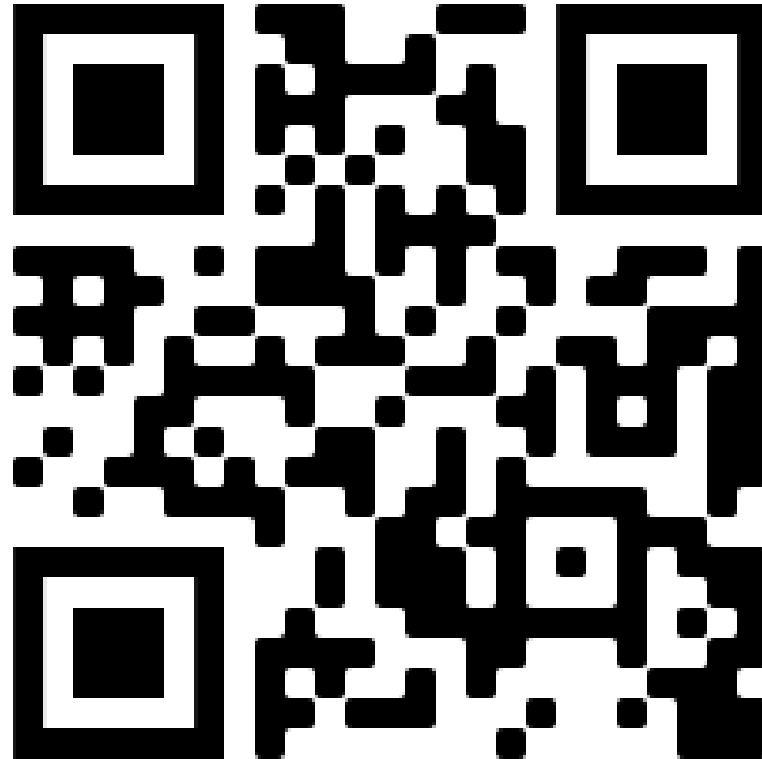
Fabio Santos

Recommended

13:35 → 45 min

Azure Synapse Analytics - Designing for Performance SQL Pool

Sidney Cirqueira

Recommended

Thanks! =D

https://linktr.ee/orlandogomes