Masterclass Analytics & AI

David Aas Correia

inmeta
Part of Crayon

About me

David Aas Correia works as a consultant in Inmeta. His focus is on Azure, data integration, data storage, business intelligence, big data, and advanced analytics. He leads an Azure Data User Group in Inmeta and is a board member of the Norwegian Computer Association, BI & Analytics.

David Aas Correia

Consultant, Data Insight & BI

Inmeta Consulting AS

David.Aas.Correia@inmeta.no

inmeta

Who are you?

https://www.menti.com/

Code: 3124481

inmeta

Practical information

- Whiteboard session
- I want your opinions and experiences and I'll ask questions along the way
- I will most likely not be able to do error handling or follow-up during the session
- I will not rush. We will have some breaks

inmeta

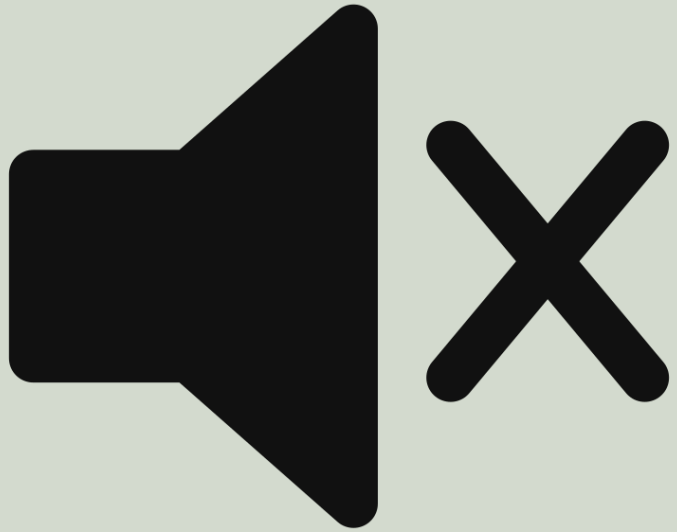
Prerequisites

- Azure *pay-as-you-go* subscription
- *Owner* role
- Power BI Desktop
- Optional: Power BI Pro license

inmeta

Before we start...

1. Did you mute your microphone?
2. Questions/discussions? Write in the chat and let's look into it between sections



inmeta

Agenda

1. Review customer case
2. Environment setup
3. Configuring environment
4. Adding other Azure Services
5. New features

inmeta

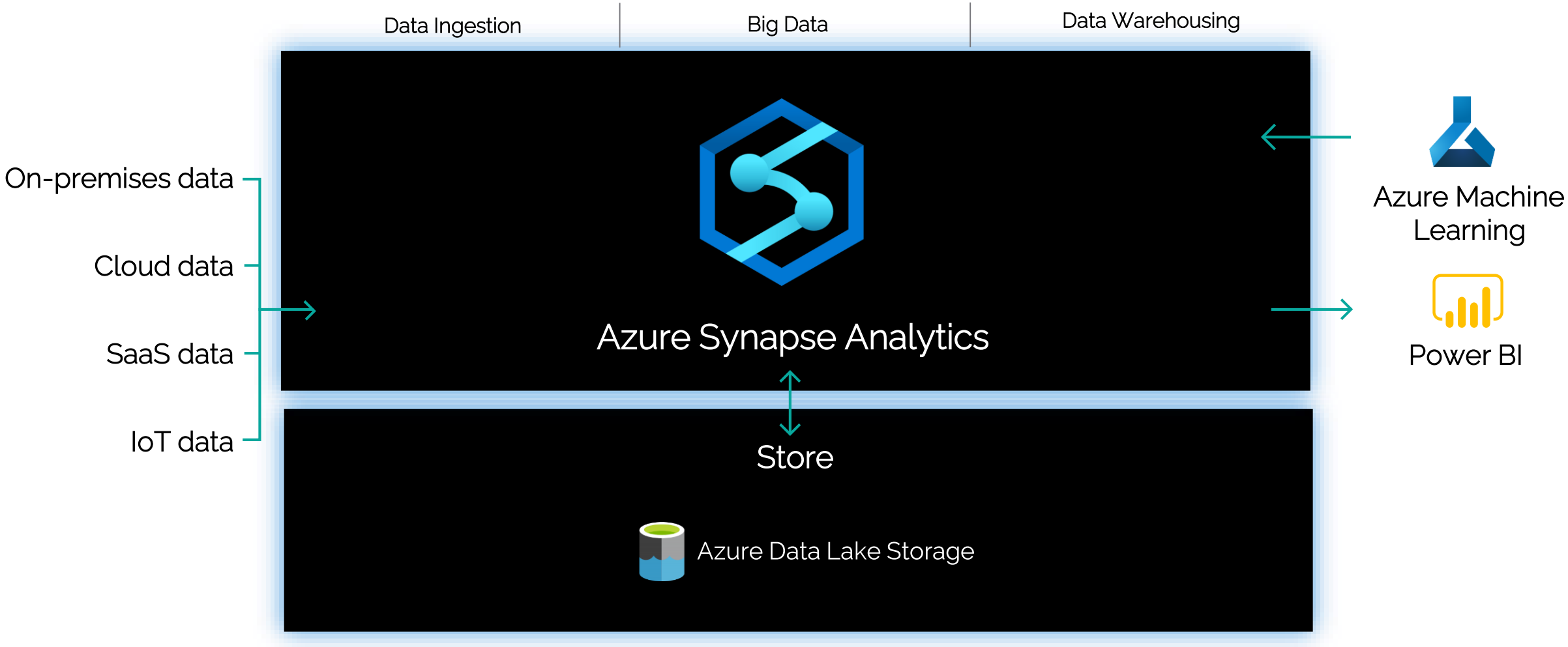
Step 1 – review customer case

inmeta

Challenge scenario

Customer wants to build a new enterprise data platform on Azure Synapse Analytics

Customer is asking you to help them with this process


Data Ingestion
Big Data
Data Warehousing
On-premises data
Cloud data
SaaS data
IoT data
Azure Synapse Analytics
Store
Azure Data Lake Storage
Azure Machine Learning
Power BI

inmeta

Step 2 – environment setup

Let’s go to Azure and deploy..

1. Azure storage accounts
    a) Normal
    b) Data Lake

2. Azure Synapse Workspace
    a) SQL Pool (DW500c)
    b) Spark Pool (Small)

3. Azure Key Vault

inmeta

inmeta

Step 3 – configuring environment

3. Configuring environment

- 3.1 Azure Key Vault
- 3.2 Populating our data lake
- 3.3 Populating our SQL Pool
- 3.4 Setting up our BI solution
- 3.5 Security
- 3.6 Monitoring
- 3.7 Optimizing environment

inmeta

3.1 Azure Key Vault

The customer wants to use Azure Key Vault as the central repository for all connection strings and passwords

Challenge: How to create an integration between Azure Key Vault and Azure Synapse Analytics?

3.2 Populating our data lake

The cust0mer wants to upload historical data to the data lake

Challenges:

1. Create a container level structure in our datalake
2. Create necessary integrations in Synapse
3. Create pipelines for data ingestion and deploy it

inmeta

3.3 Populating our SQL Pool

The cust0mer wants to upload to be able to report on data from 2018 and onwards

Challenges:

1. Create necessary objects in SQL Pool
2. Create integrations in Synapse
3. Create pipeline for data ingestion and deploy it

inmeta

3.4 Setting up our BI solution

The cust0mer wants have a Power BI dashboards that gives insights into taxi activity in NYC

Challenges:

1. Create Power BI report
2. Publish Power BI report
3. Share Power BI report

3.5 Security

Principal of least privilege

Use Azure AD

Use vNets and firewalls to restrict network access

Use Azure Key Vault

Audit, log and monitor resources

3.5.1 Security – Synapse Analytics

The customer wants to mask data and restrict access to columns and rows for certain groups in their organization

Challenge:

Which kind of built-in features can we use to achieve this goal?

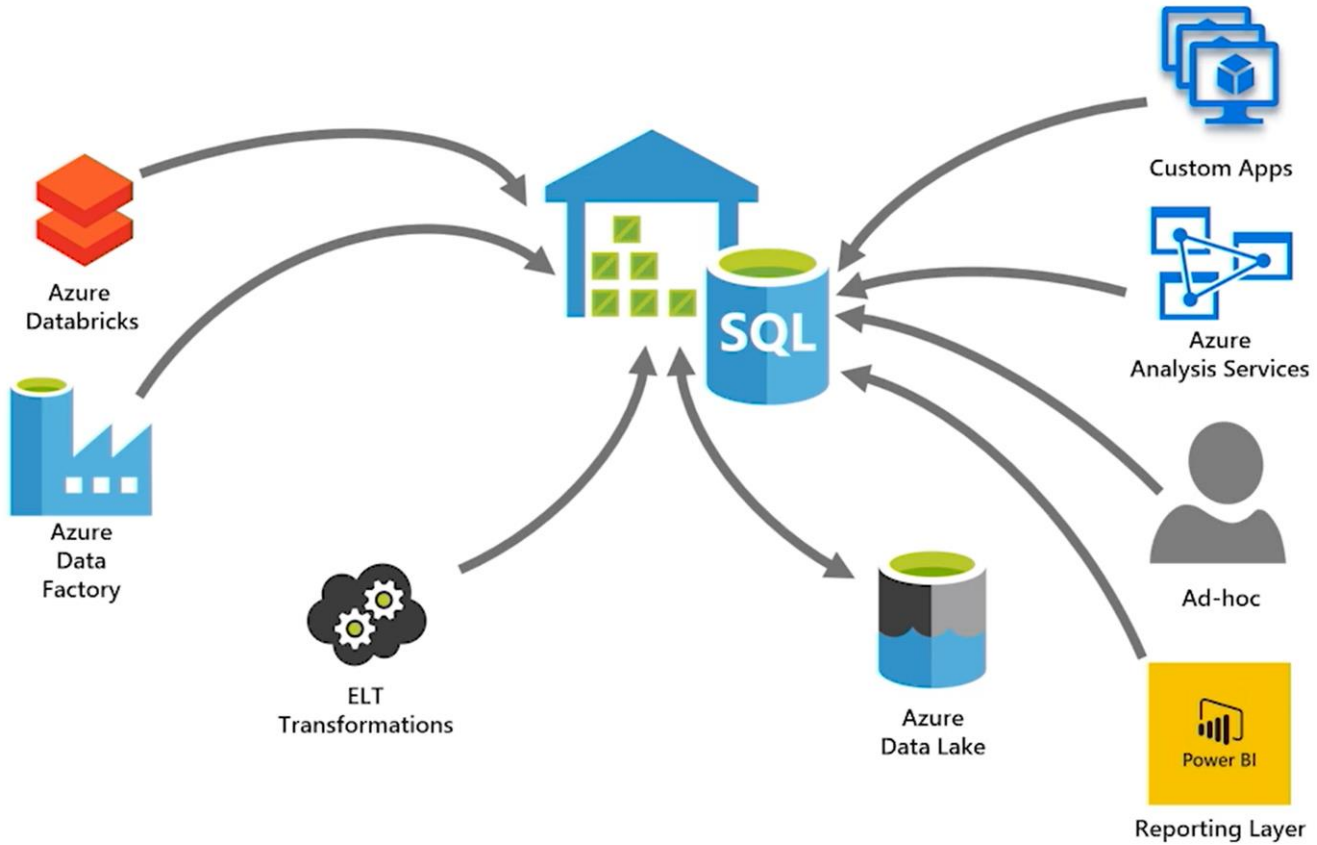
3.6 Monitoring

The customer wants to know how to apply workload management and monitoring

Challenge:

Which features can we use to achieve this goal? And what information can we get from our workloads?

3.6 Monitoring



inmeta

3.6 Monitoring

Workload classification: assigning a request to a workload group

Workload importance: influences the order in which a request gets access to resources

Workload isolation: reserves resources for a workload group

Resource class: determines performance capacity, static or dynamic

3.7 Optimizing environment

Customer wants you to assess how you can optimize the Synapse environment

Challenge:

Which processes and tools can you apply?

inmeta

Step 4 – adding services

IoT and Machine Learning

inmeta

4.1 IoT

The customer wants to have real-time and near real-time updates for NYSE trade activity.

Challenge:

How can we achieve this goal?

inmeta

4.2 Machine learning

The customer wants to predict what trips will give tip or not

Challenge:

How can we achieve this goals?

inmeta

Step 5 – new features

Integration with DevOps

New studio interface

inmeta

inmeta

Remember to delete your resources

Documentation and inspiration

- https://docs.microsoft.com/en-us/azure/synapse-analytics/sql-data-warehouse/sql-data-warehouse-workload-management
- Synapse gallery – "Data Exploration and ML Modeling - NYC taxi predict using Spark Mllib"
- Azure Data Platform End2End (V2) - https://github.com/fabragaMS/ADPE2E
- ThisOldDataWarehouse - https://github.com/microsoft/WhatTheHack/tree/master/019-ThisOldDataWarehouse
- Azure Synapse Analytics and AI - https://github.com/microsoft/MCW-Azure-Synapse-Analytics-and-AI

Thank you!

For questions, email: David.Aas.Correia@inmeta.no

inmeta