

Traditional Modern Data Warehouse

dataMinds Connect

11th October 2022



ORANGEMAN

Thank you, sponsors



K[◇]HERA

element
experience & expertise

 Microsoft

inetum.
realdolmen
Positive digital flow

bmatix
Act informed

LACO / KNOWLEDGE,
AND MORE.

lytix

 DataSense
Because your data makes sense!

MICROPOL

KPMG

delaware
we commit. we deliver.



 AKABI

 ADVANCING
ANALYTICS

u2u

de
Adapt
and
Enable

 cegeka
IN CLOSE COOPERATION

 Business & Decision

 MONIN
Database Managed Services

ORDINA
Ahead of change

9Δ nine
altitudes
REACH NEW HEIGHTS

 Tabular Editor

Just Thorning Blindbæk

- ▶ Self-employed BI consultant in justB
- ▶ Trainer at Orange Man
- ▶ Founder
 - Danish Microsoft BI Community ([MsBIP.dk](https://msbip.dk))
 - Power BI UG Denmark ([PowerBI.dk](https://powerbi.dk))
- ▶ Strong focus on
 - Azure BI architecture
 - Analysis Services
 - Reporting Services
 - Power BI
- ▶ just@blindbaek.dk / blog.justB.dk / [@justblindbaek](https://twitter.com/justblindbaek) / youtube.com/c/justblindbaek



Microsoft
CERTIFIED
Trainer

Microsoft
CERTIFIED
IT Professional

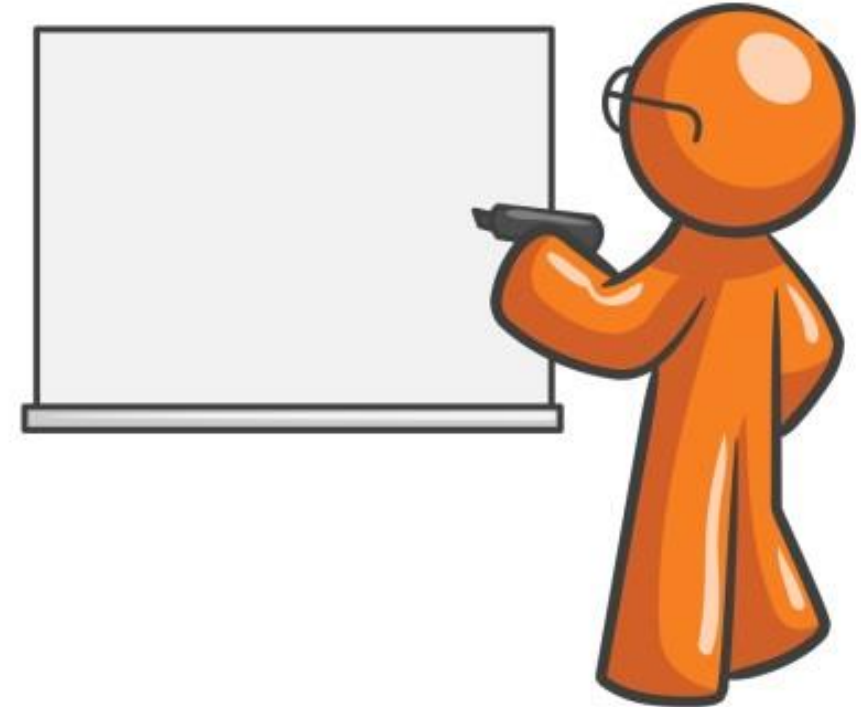
SSAS
MAESTRO
by Microsoft

MVP **Microsoft**
Most Valuable
Professional

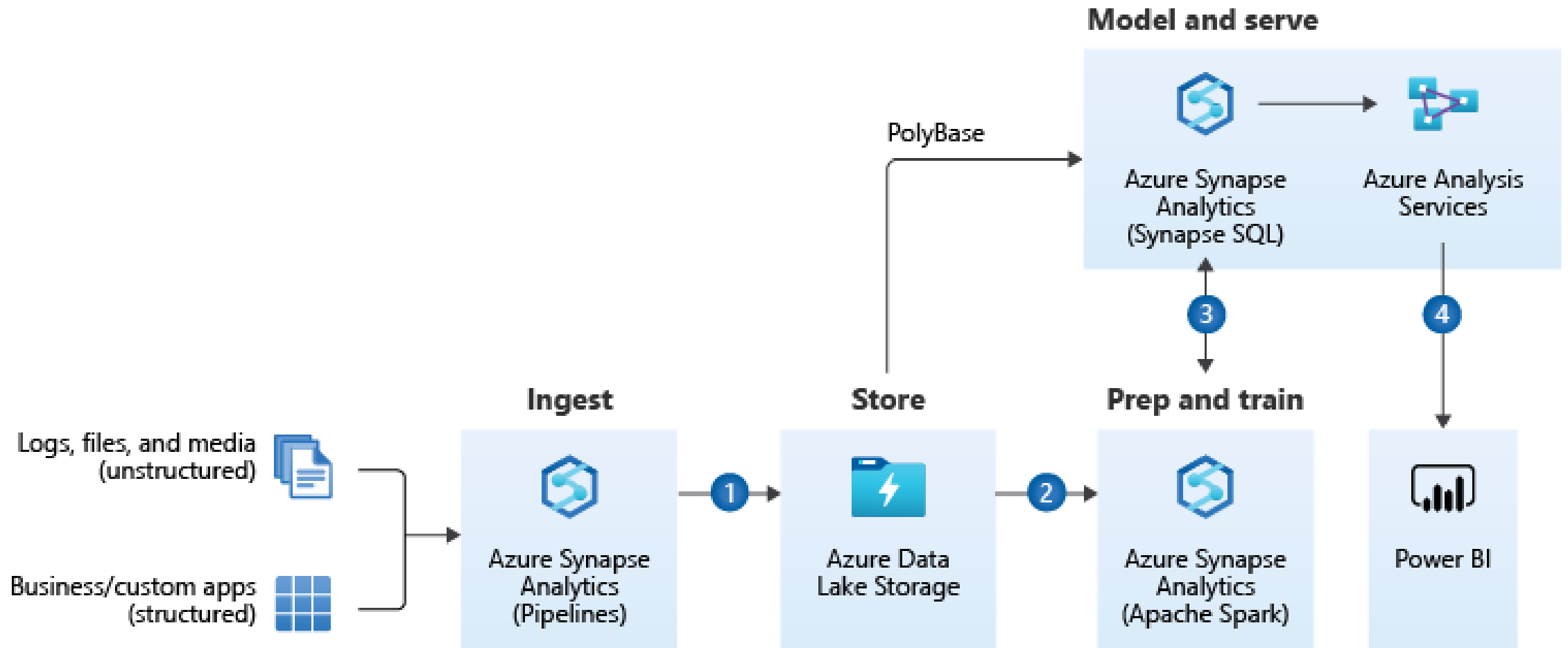
ORANGEMAN

Agenda

- ▶ Big data, big data, big data
 - Modern Data Warehouse
 - Synapse Analytics
- ▶ Small data – SMB market
 - Reuse SQL competencies
 - Simple and affordable E(L)TL architecture
 - Build to scale !
- ▶ Core components
- ▶ Pricing estimate



Enterprise Data Warehouse in Azure



Enterprise Data Warehouse in Azure

Model and serve

Enterprise Data Warehouse

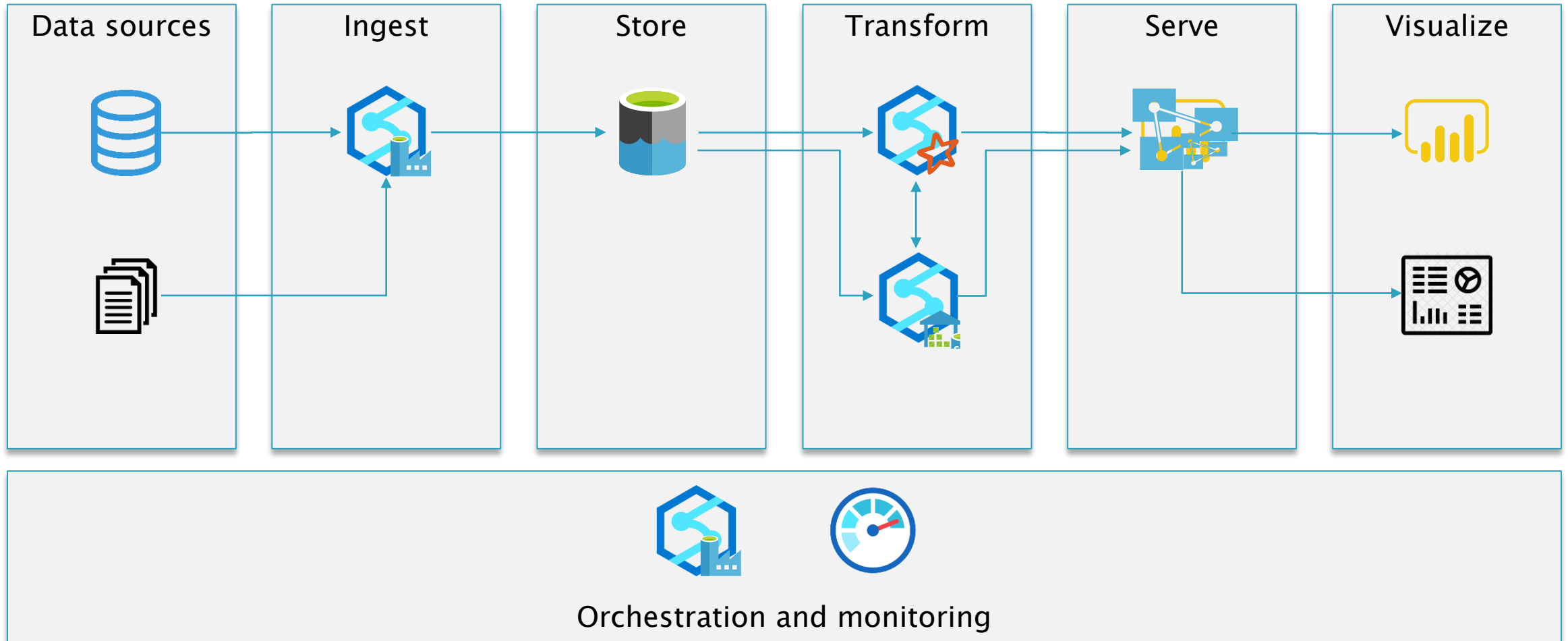
✓ Azure Synapse Analytics	Tier: Compute Optimized Gen2, Dedicated SQL Pool...	Monthly: \$1,755.90
✓ Azure Analysis Services	Developer (Hours), 5 Instance(s), 720 Hours	Monthly: \$475.20
✓ Storage Accounts	Data Lake Storage Gen2, Standard, LRS Redundancy...	Monthly: \$71.88
Estimated upfront cost		\$0.00
Estimated monthly cost		\$2,302.98

(structured) ■■■

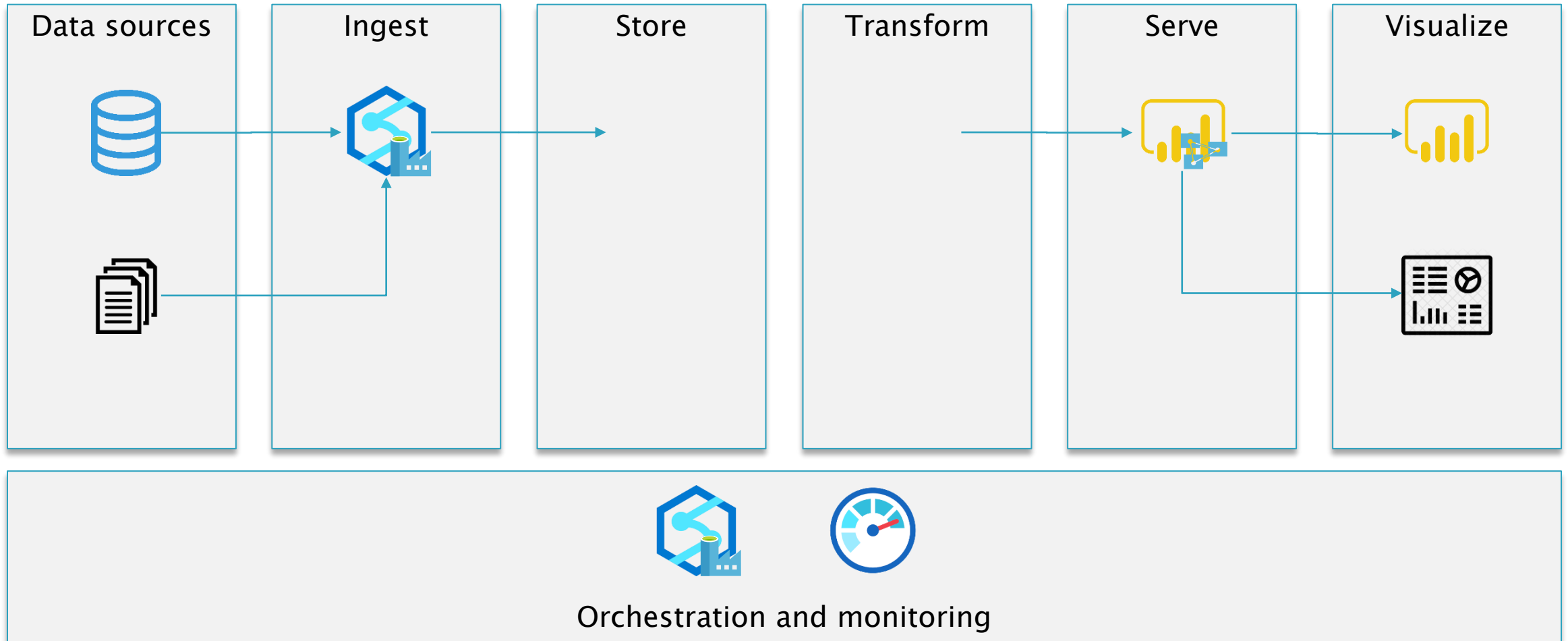
(Pipelines)

(Apache Spark)

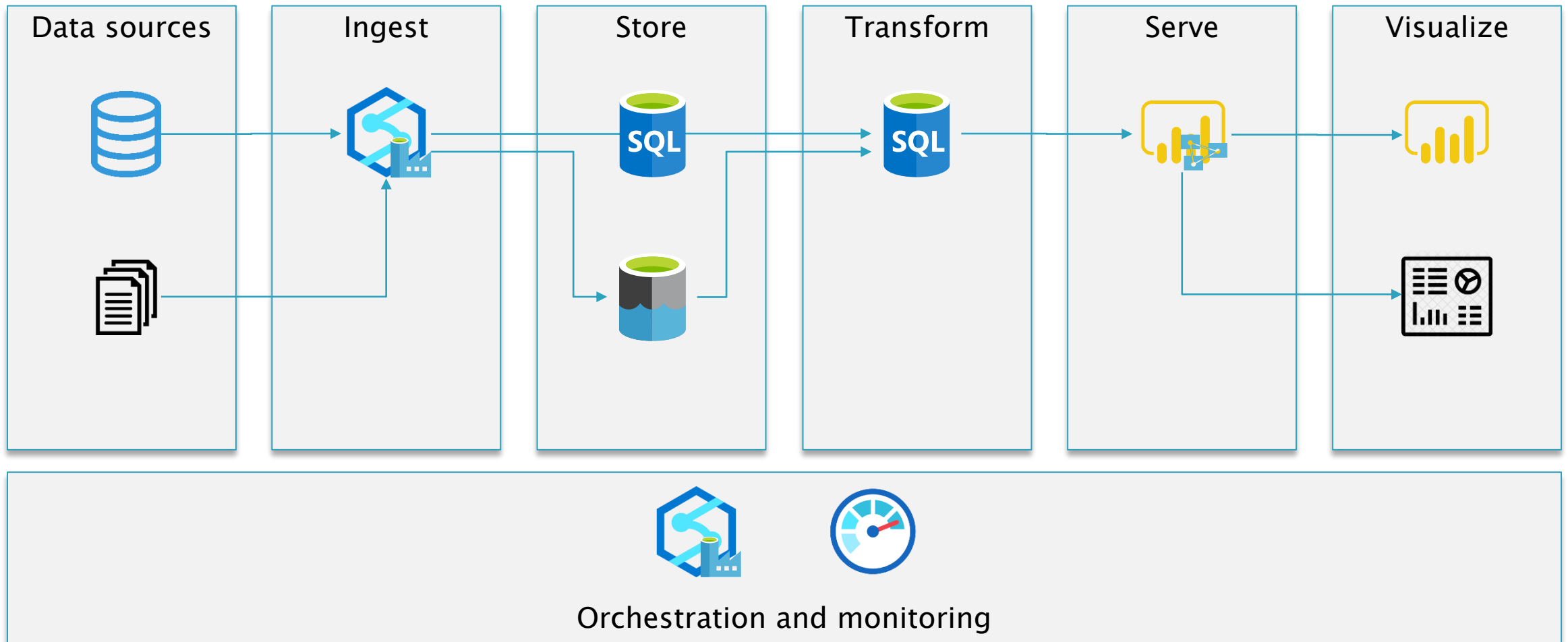
Enterprise Data Warehouse in Azure



Where to Store and Transform?

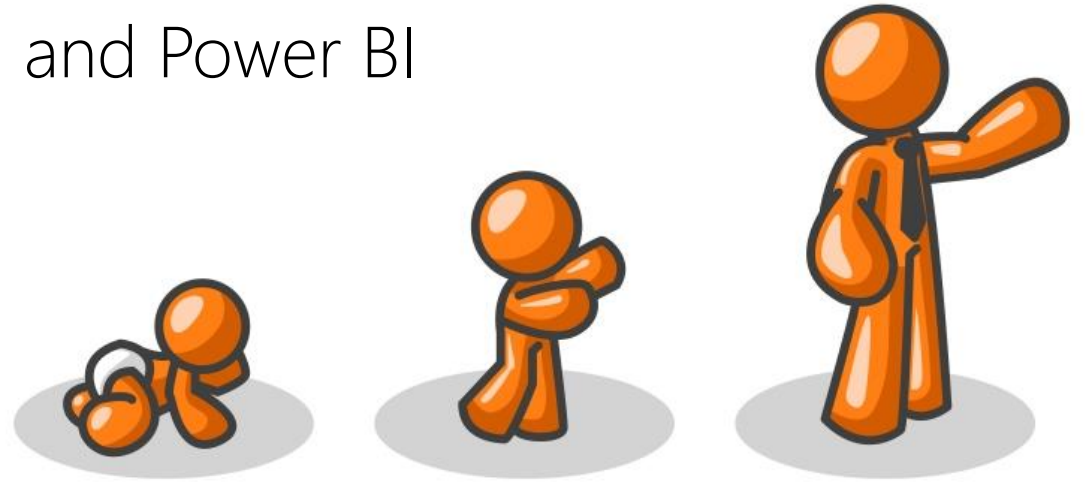


Traditional Modern Data Warehouse



Components

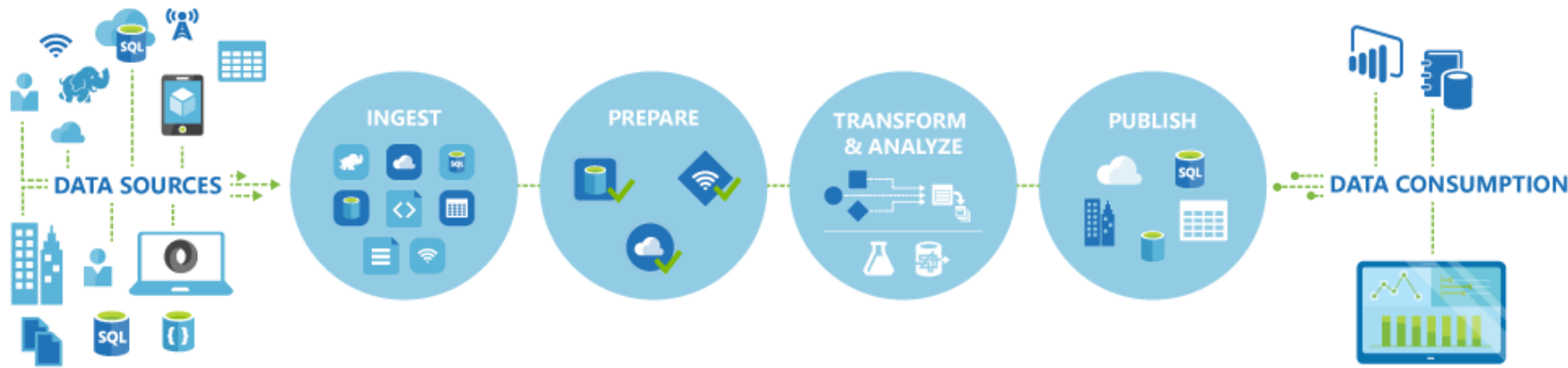
- ▶ Data Factory, Data Lake Store, SQL DB and Power BI



Azure Data Factory in Synapse



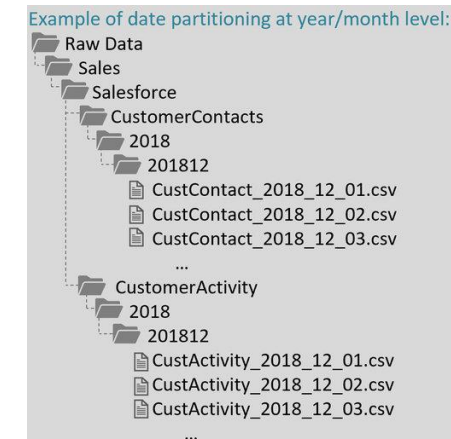
- ▶ Hybrid data integration at scale
- ▶ 90+ connectors provided to Cloud & Hybrid
- ▶ Create, plan, administer and monitor data pipelines
- ▶ Execute activities (copy, transform or orchestration)
- ▶ Parameters, control flow and triggers



Azure Data Lake Store

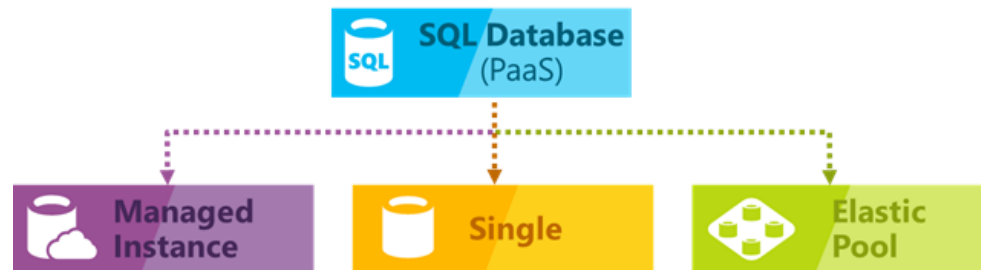


- ▶ Storage Account with “hierarchical namespace” enabled
- ▶ Unlimited cheap, cheap, cheap storage
- ▶ Can handle structured, semi-structured, and unstructured data
- ▶ Apache Hadoop file system compatible (HDFS)

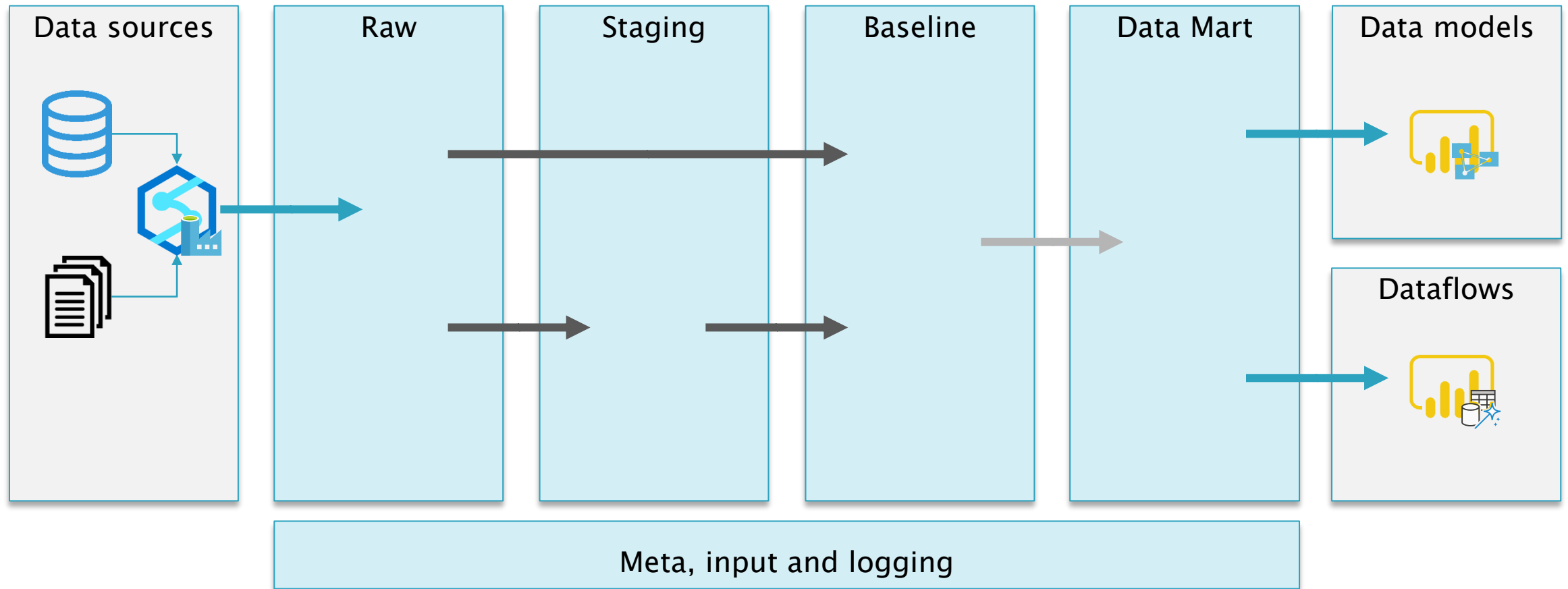


Azure SQL Database

- ▶ Relational database-as-a service
- ▶ Scale to your needs
- ▶ DTU- and vCore-based models



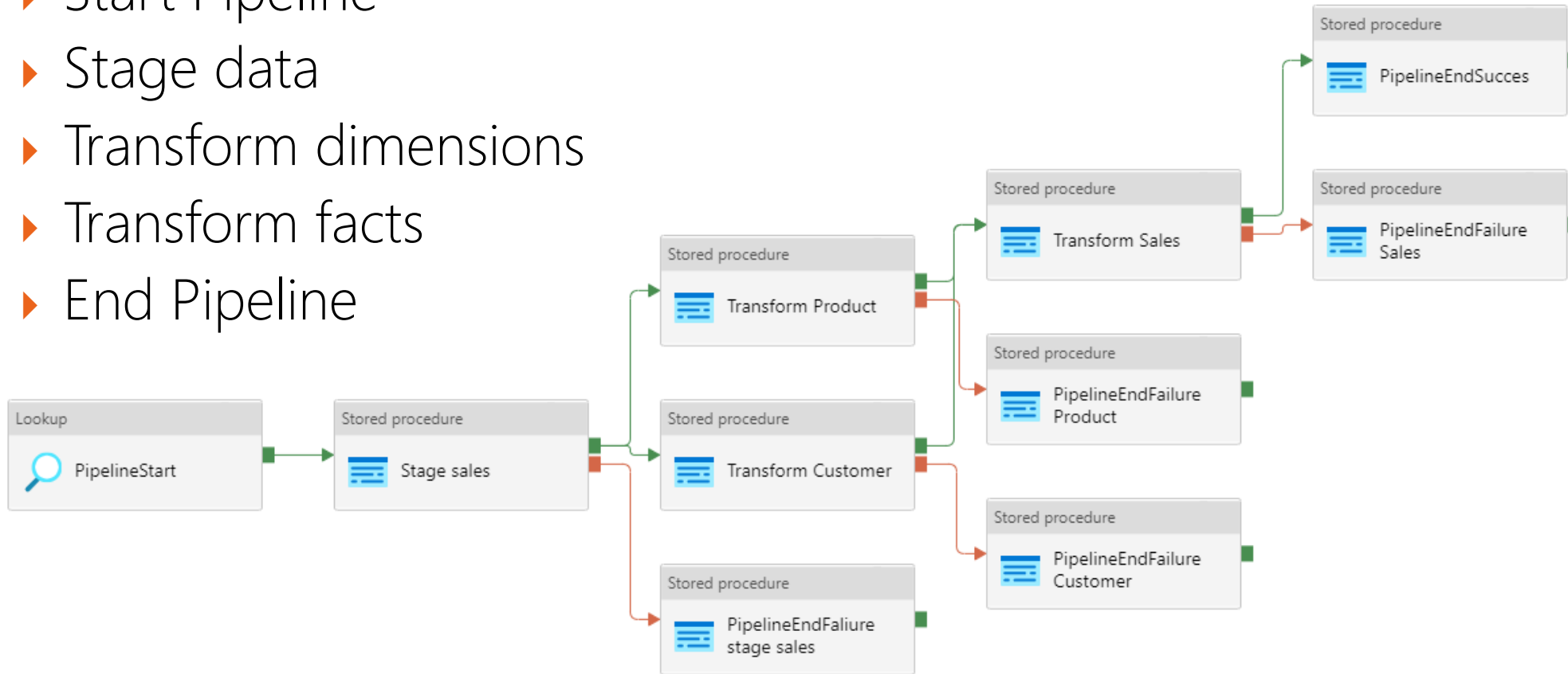
Transform "framework" in SQL Database



Transform pipeline



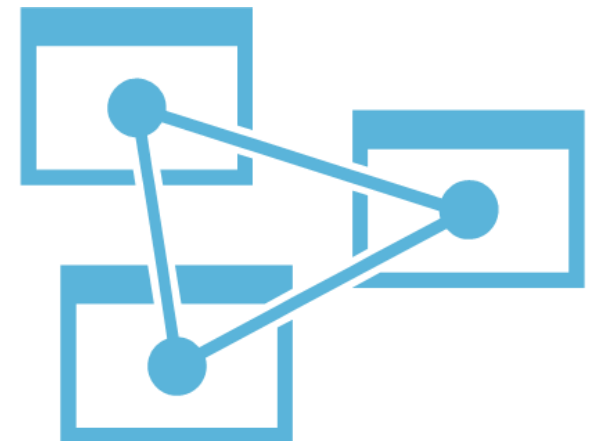
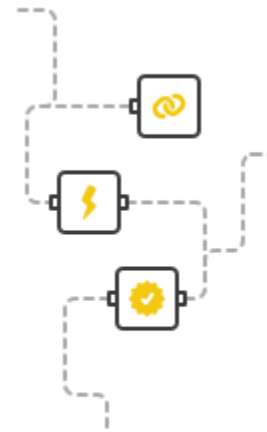
- ▶ Start Pipeline
- ▶ Stage data
- ▶ Transform dimensions
- ▶ Transform facts
- ▶ End Pipeline



Analysis Services in Power BI



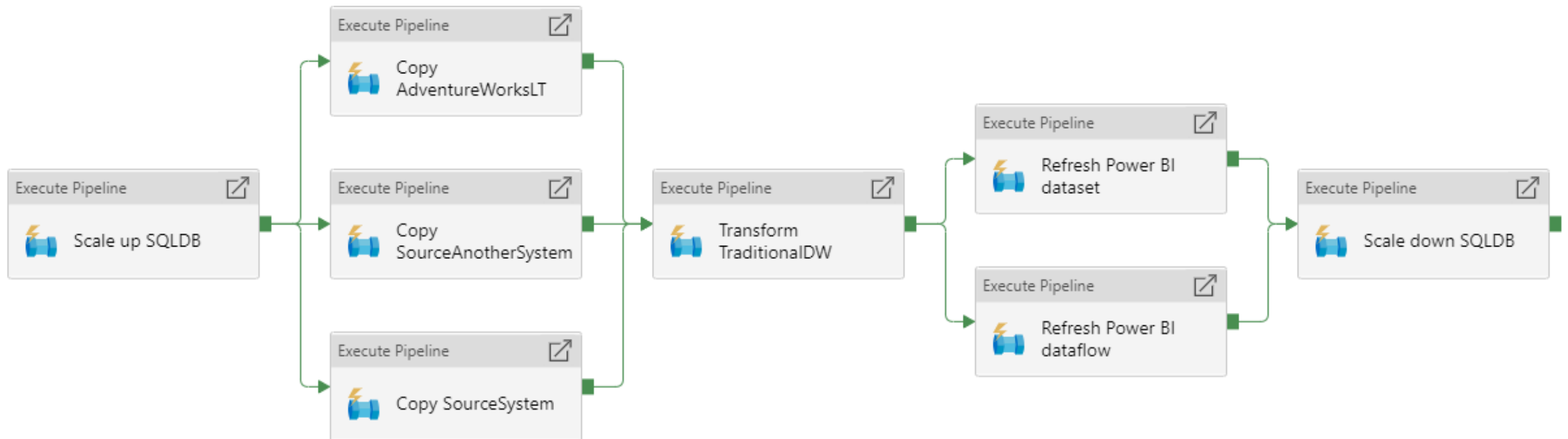
- ▶ Analysis Services engine inside
 - XMLA endpoint to manage the Semantic data models
- ▶ Also have stunning visuals, reports and dashboards
- ▶ And dataflows to expose “self-service” data to business analysts
- ▶ Recommended license options
 - Power BI Embedded capacity
 - Power BI Premium Per User



Orchestration

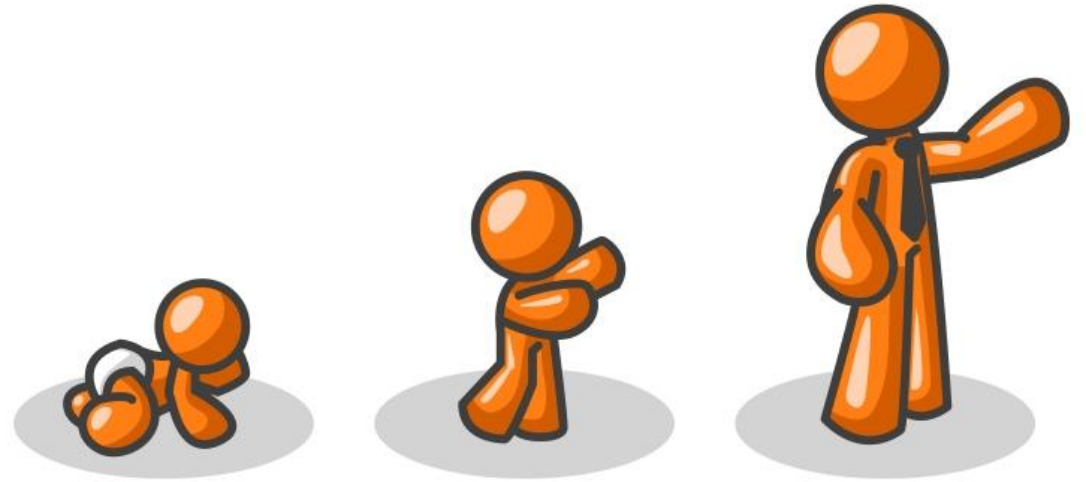


► Master Pipeline in Azure Data Factory



Pricing

- ▶ Estimate



ORANGEMAN

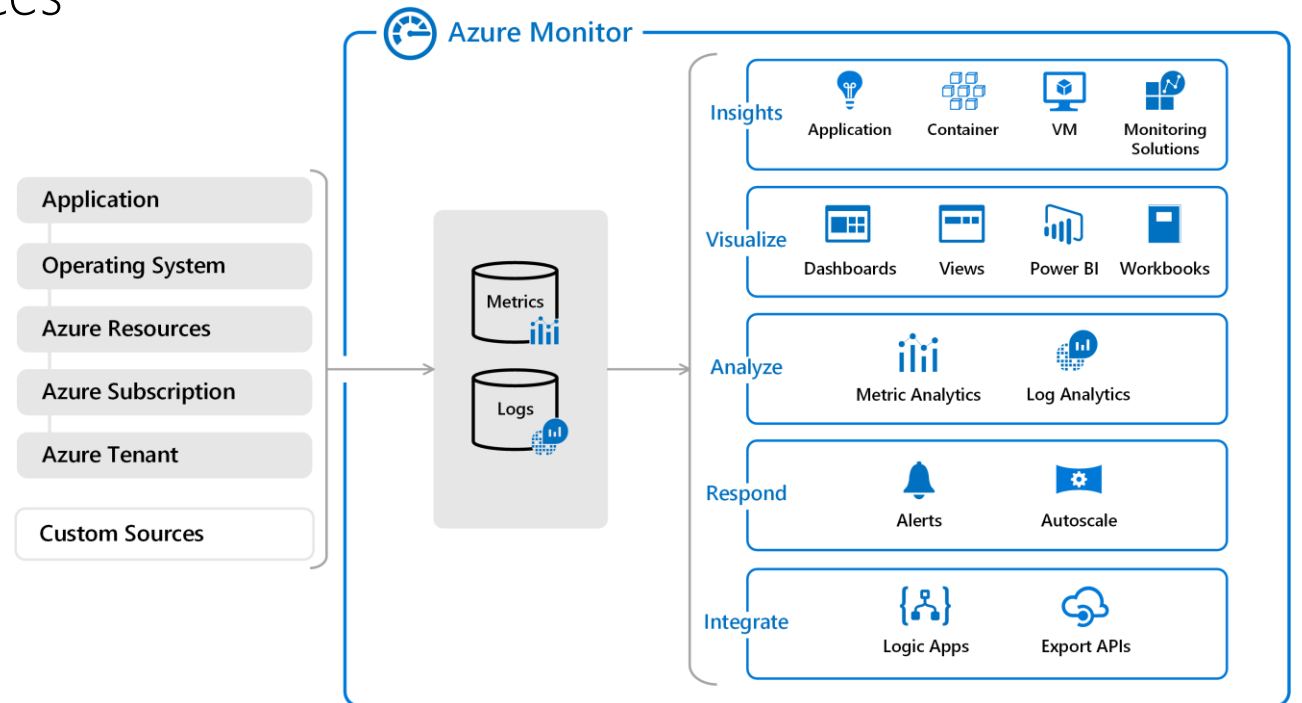
Pricing estimate: Traditional Modern DW

Service type	Description	Monthly
Storage Account	Data Lake Storage Gen2, Standard, LRS Redundancy, Hot Access Tier, Hierarchical Namespace File Structure, 100 GB Capacity	\$ 3.71
Azure SQL Database	Single Database, DTU Purchase Model, Standard Tier, S0: 10 DTUs, 250 GB included storage per DB, 1 Database(s) x 640 Hours	\$ 12.90
Azure SQL Database	Single Database, DTU Purchase Model, Standard Tier, S6: 400 DTUs, 250 GB included storage per DB, 1 Database(s) x 90 Hours	\$ 72.59
Azure Data Factory	Azure Data Factory V2 Type, Data Pipeline Service Type, Azure Integration Runtime: 1 Activity Run(s), 30 Data movement unit(s), 90 Pipeline activities	\$ 9.72
Power BI Embedded	Power BI Embedded, Node type: A1, 3GB Memory, 10 Hours	\$ 10.08
Azure Monitor	Log analytics	\$ 1.60
Estimated Monthly Cost		\$ 110.60

Monitoring



- ▶ Custom logging in SQL Database
 - Pipeline and Task level
- ▶ Azure Monitor
 - Gather diagnostic from Azure Services
 - Analyze with Log Analytics
 - Setup alerts
 - Export with Power Query



Questions



ORANGEMAN

Useful resources

- ▶ Enterprise Data Warehouse:
<https://docs.microsoft.com/en-us/azure/architecture/solution-ideas/articles/enterprise-data-warehouse>
- ▶ Pricing estimate:
<https://azure.com/e/9d56929b959546b384594e1074a1f506>
- ▶ Code on GitHub :
<https://github.com/justBlindbaek/TraditionalModernDW>

