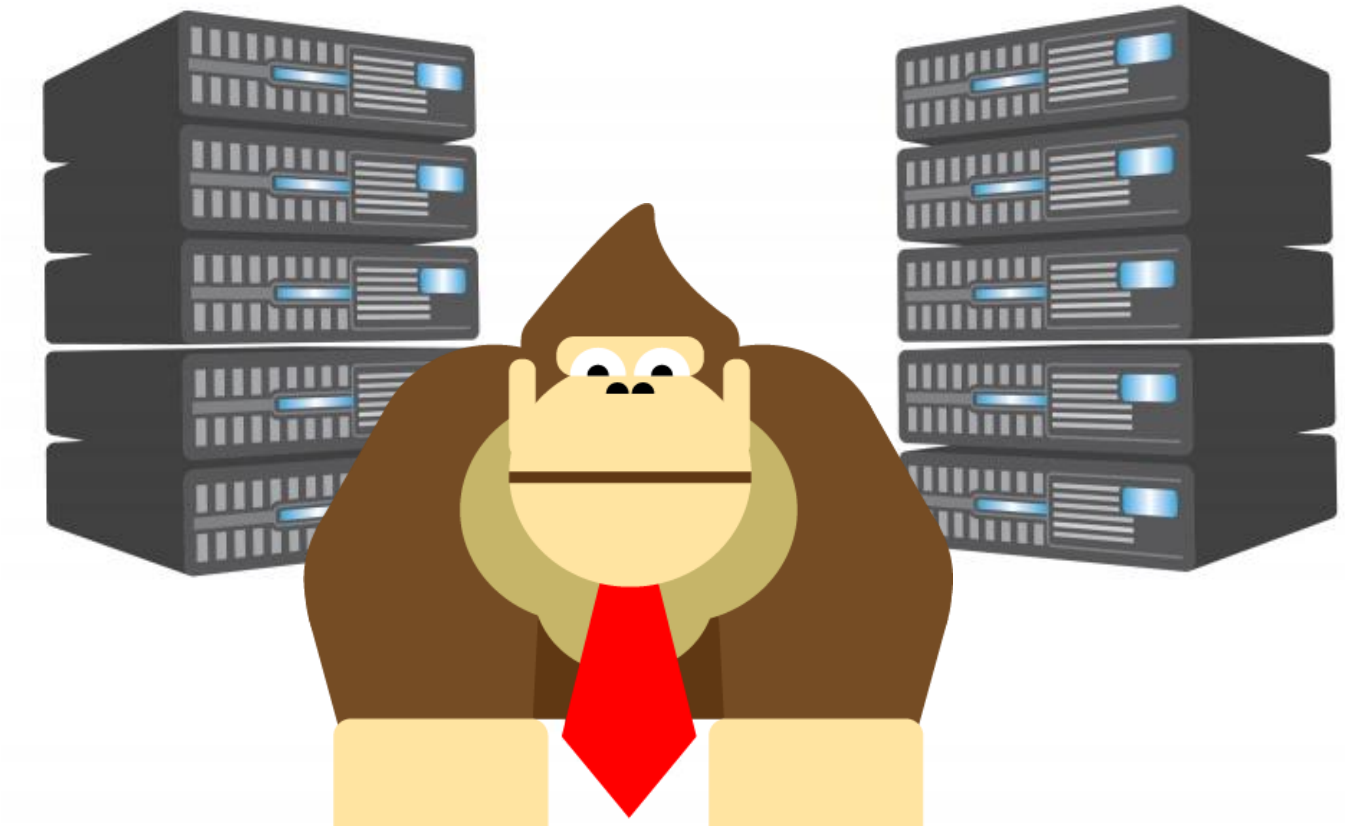
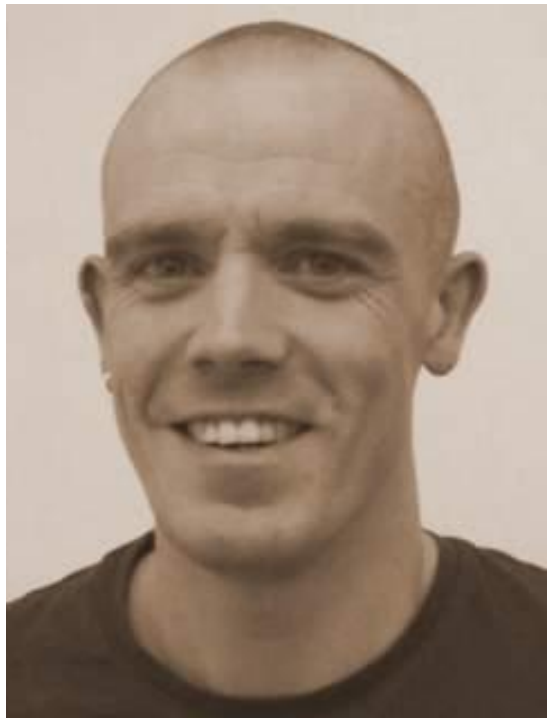
Architectural blueprints for the Modern Data Warehouse



ORANGEMAN

Just Blindbæk

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



Microsoft CERTIFIED Trainer

Microsoft CERTIFIED IT Professional

SSAS MAESTRO by Microsoft

MVP Microsoft® Most Valuable Professional

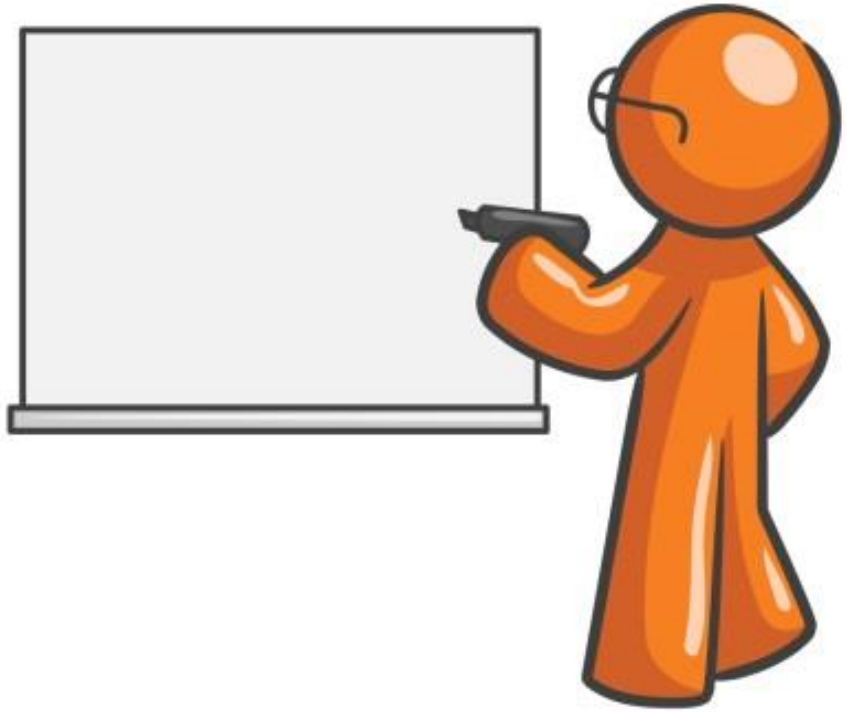
ORANGEMAN

Agenda

- What is a Modern Data Warehouse?

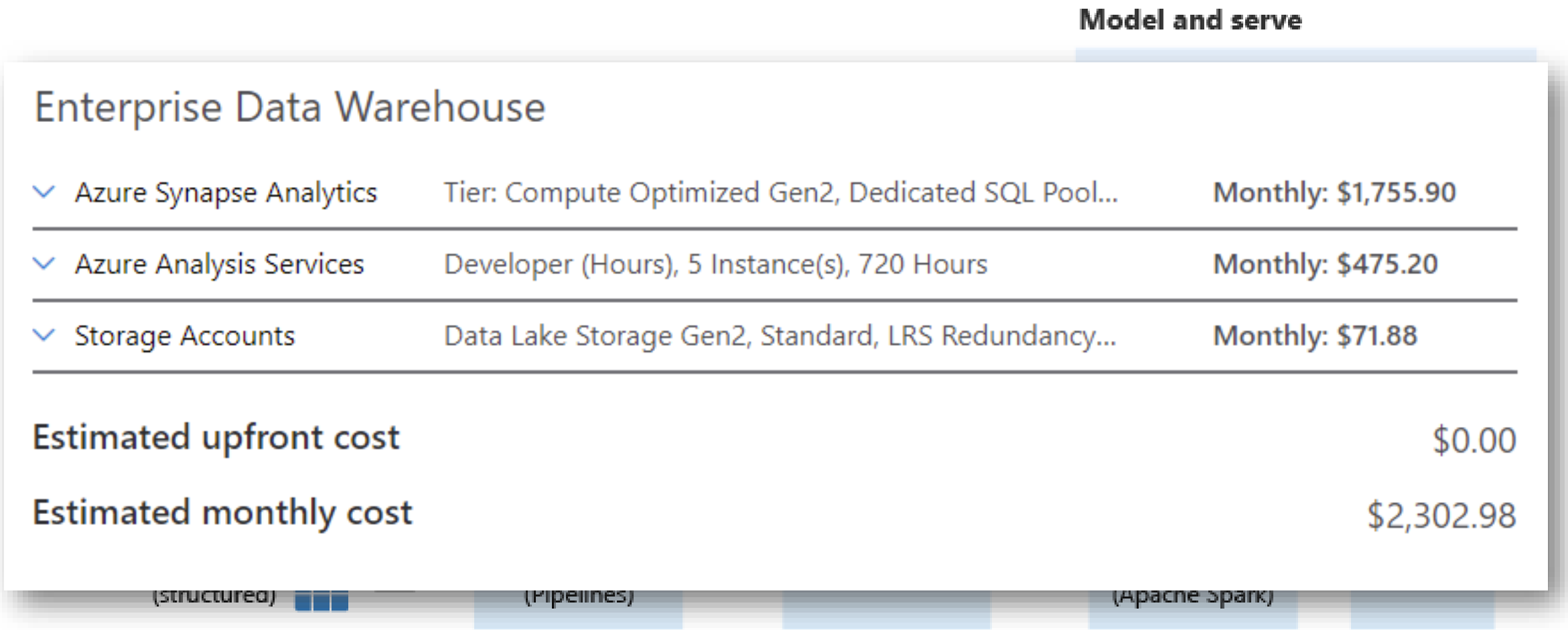
- Traditional Modern Data Warehouse
- Data Lakehouse with Spark
- Self-service with Power BI Dataflows

- Comparison of the solutions



ORANGEMAN

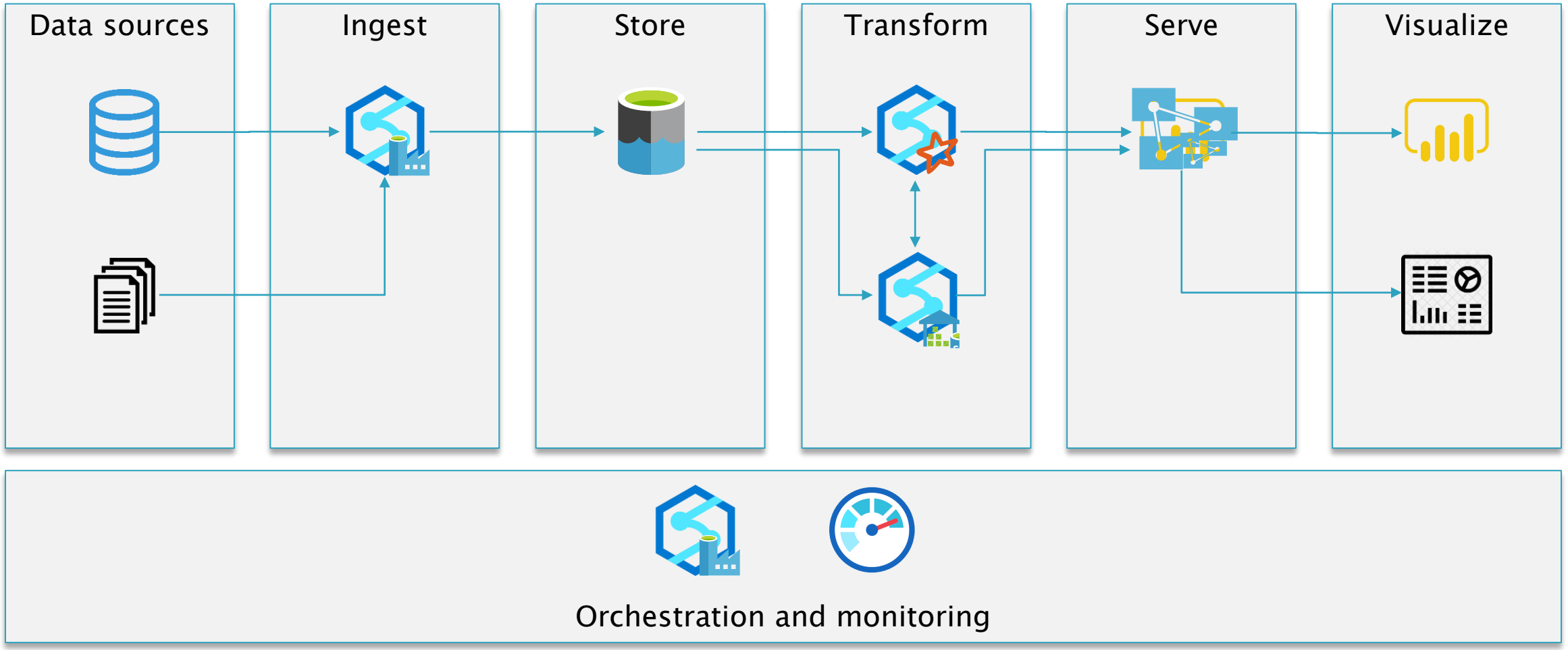
Enterprise Data Warehouse in Azure



ORANGEMAN

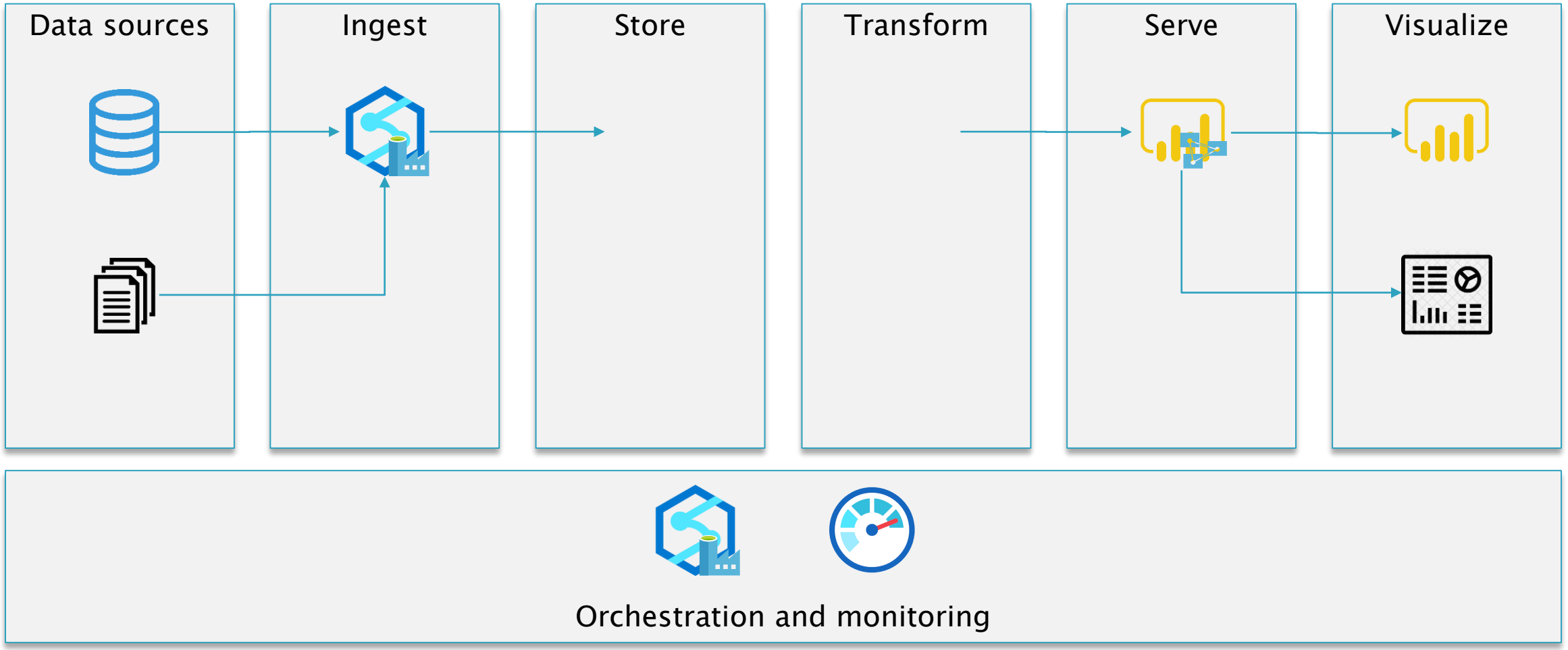
https://docs.microsoft.com/en-us/azure/architecture/solution-ideas/articles/enterprise-data-warehouse

Enterprise Data Warehouse in Azure



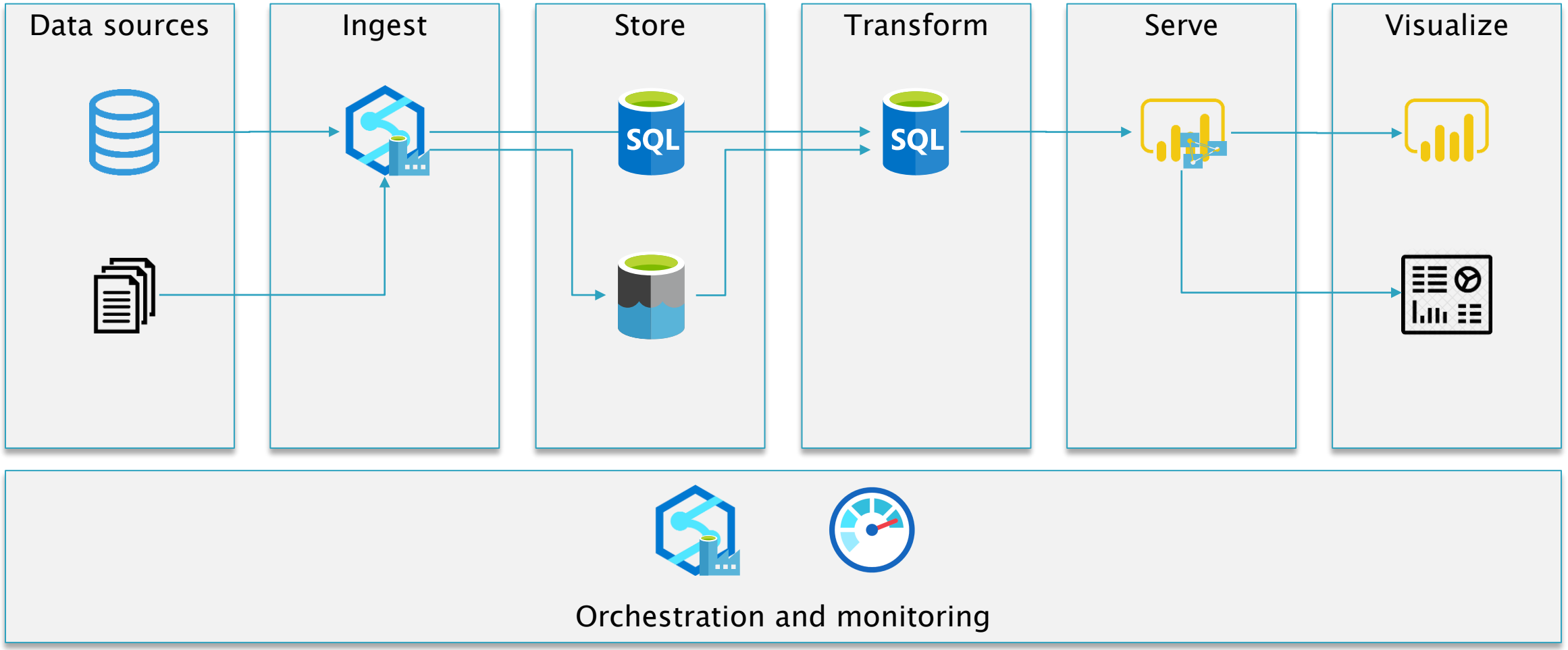
ORANGEMAN

Where to Store and Transform?



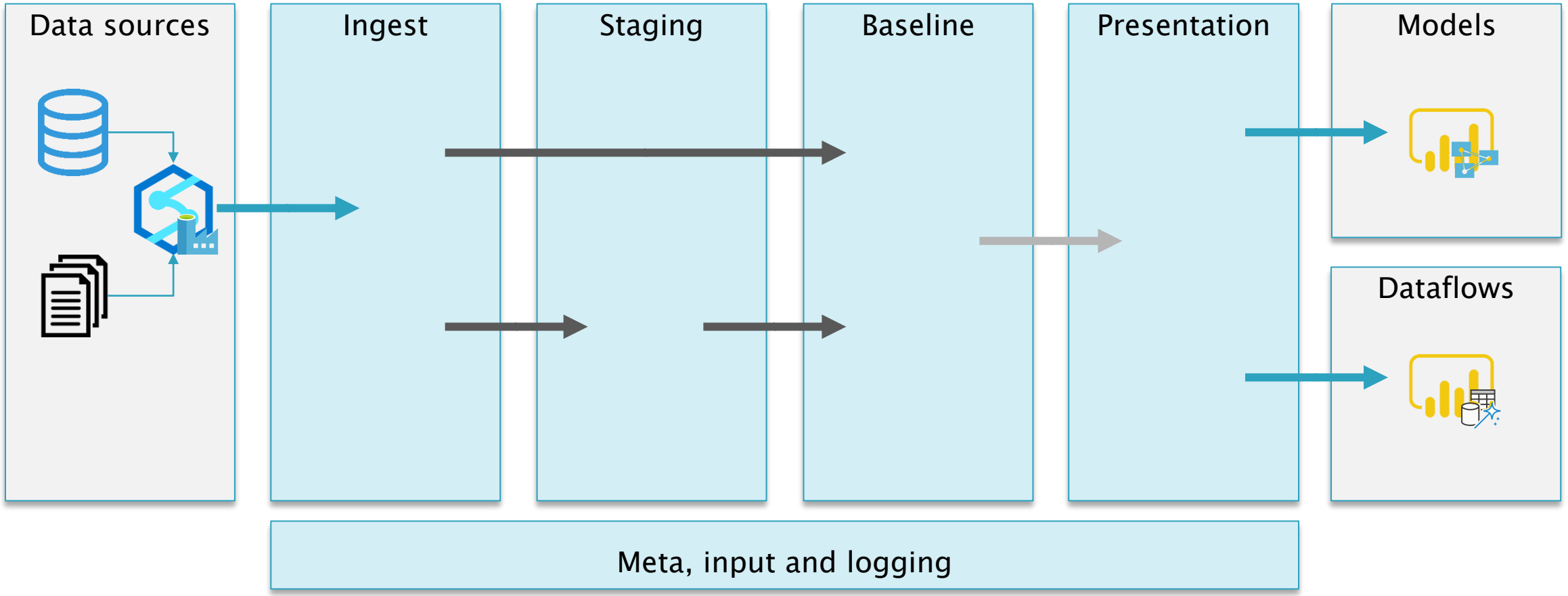
ORANGEMAN

Traditional Modern Data Warehouse



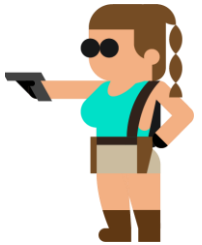
ORANGEMAN

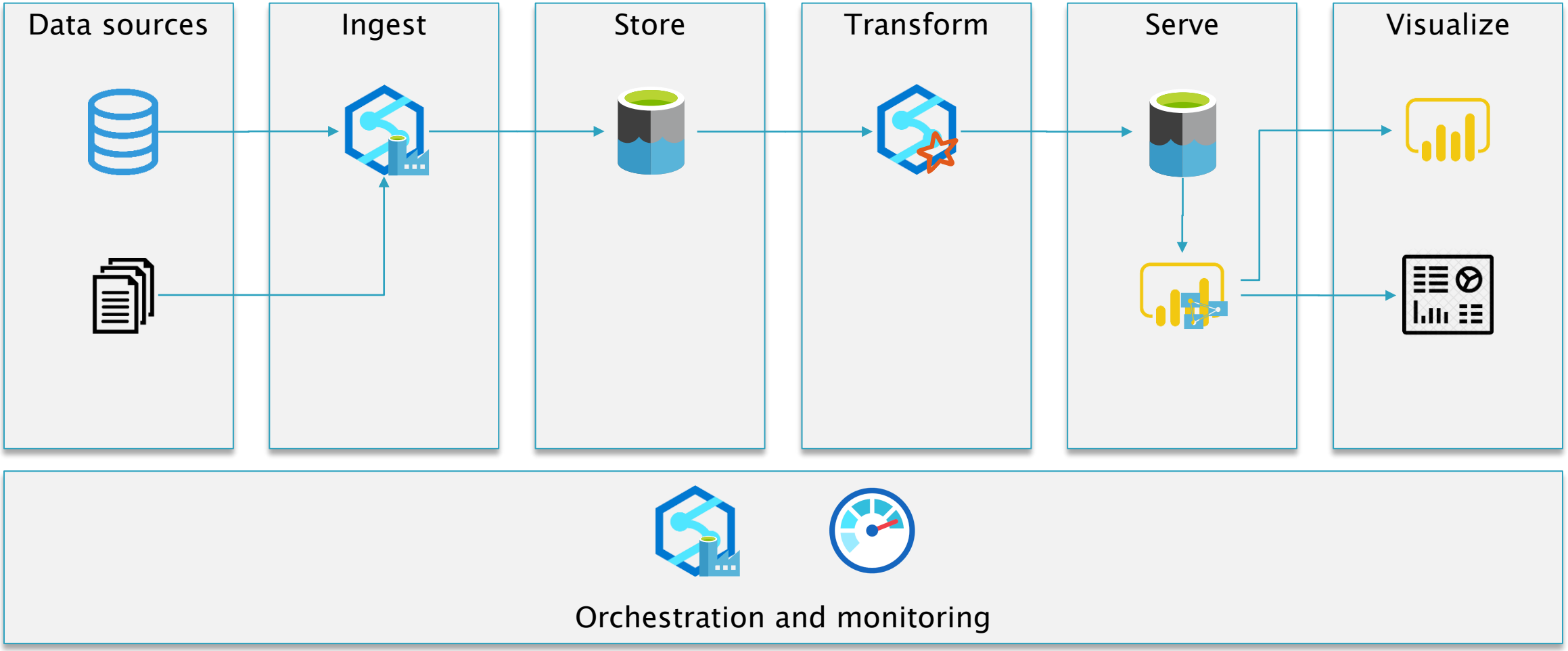
Transform "framework" in SQL Database



ORANGEMAN

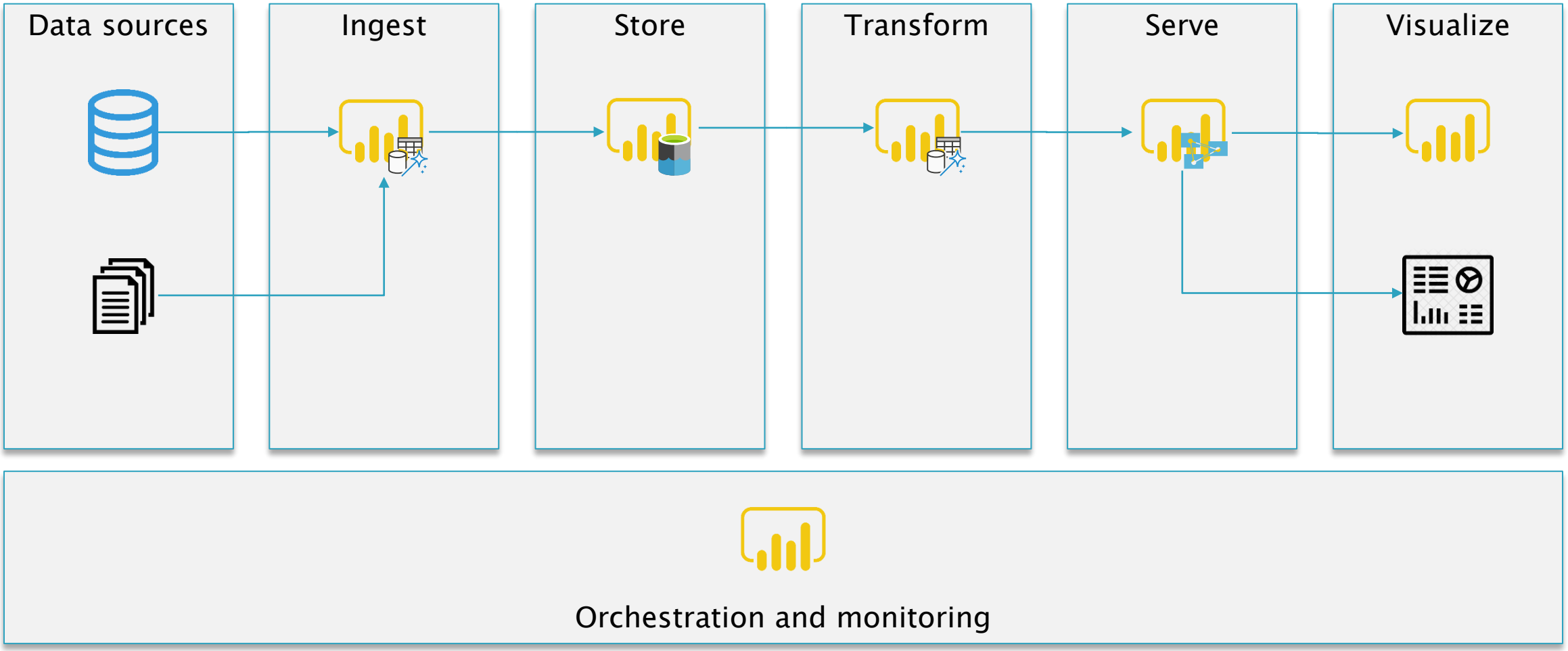
Data Lakehouse with Spark





ORANGEMAN

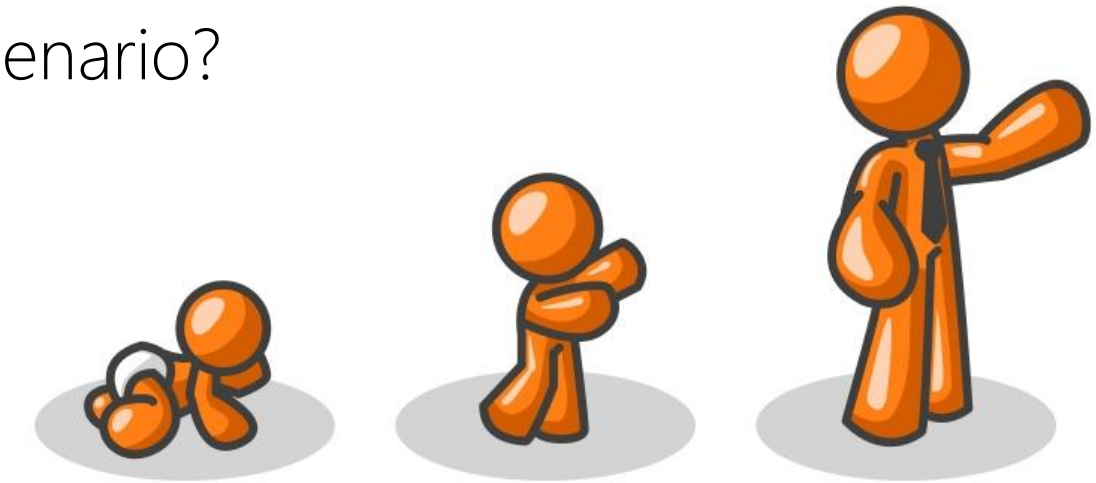
Self-service with Power BI Dataflows



ORANGEMAN

Comparison of the solutions

- The conclusion – what's best in your senario?



ORANGEMAN

Comparison of the solutions

	Traditional Database 	Data Lakehouse 	Power BU Dataflows 
Transformation language	SQL	SQL or PySpark	M or GUI
Developer persona	Pro	Pro	Citizen
Scalability in data volume	Medium	High	Low
Portability to other services	High	High	Low
Pricing model	Compute	Compute	Free?
Monthly cost	Low	Low	Very low
Extensibility with AI, ML etc.	Limited	Many options	Limited
Data processing	Schema on write	Schema on read	Schema on write
DevOps possibilities	Yes	Yes	No

ORANGEMAN

Just Blindbæk

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

https://sqlb.it/?7005

Microsoft CERTIFIED Trainer

Microsoft CERTIFIED IT Professional

SSAS MAESTRO by Microsoft

MVP Microsoft® Most Valuable Professional

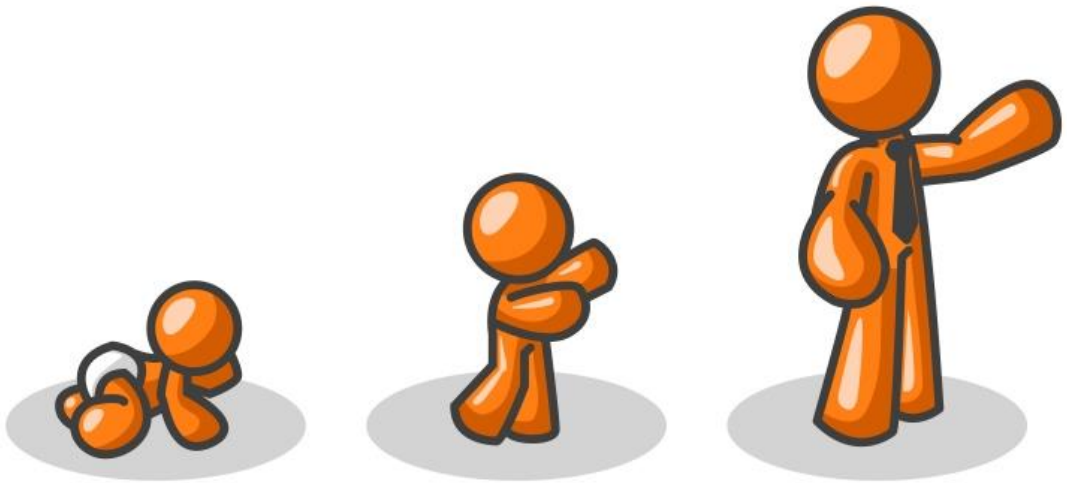
ORANGEMAN

Questions

ORANGEMAN

Pricing

- Estimate



ORANGEMAN

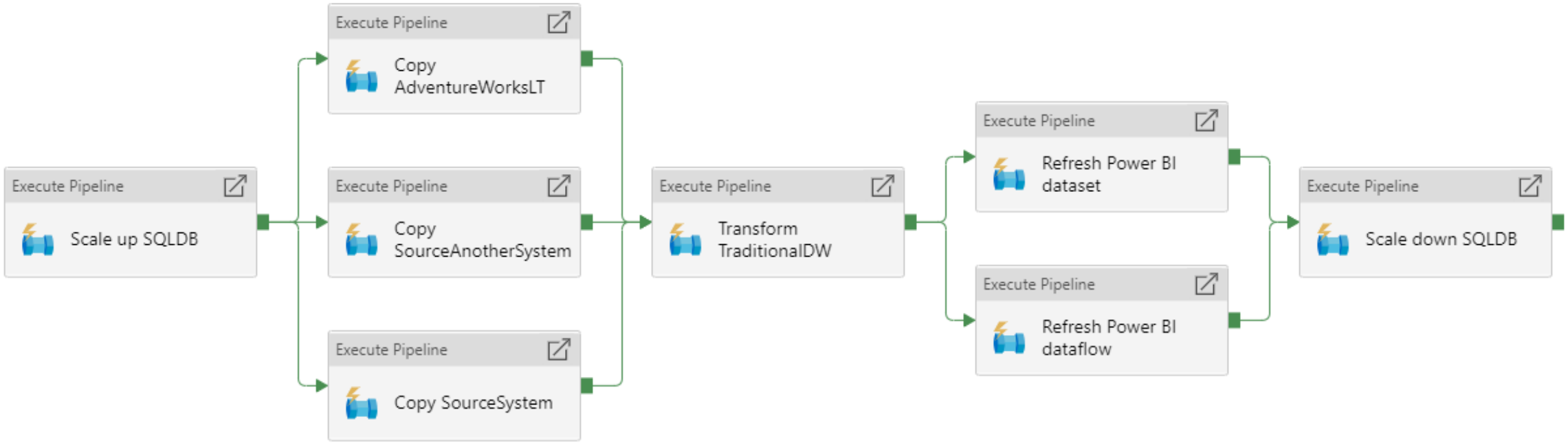
Pricing estimate: Traditional Modern DW

Service type	Description	Estimated Cost
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
	Monthly Total	$ 110.60

ORANGEMAN

Orchestration

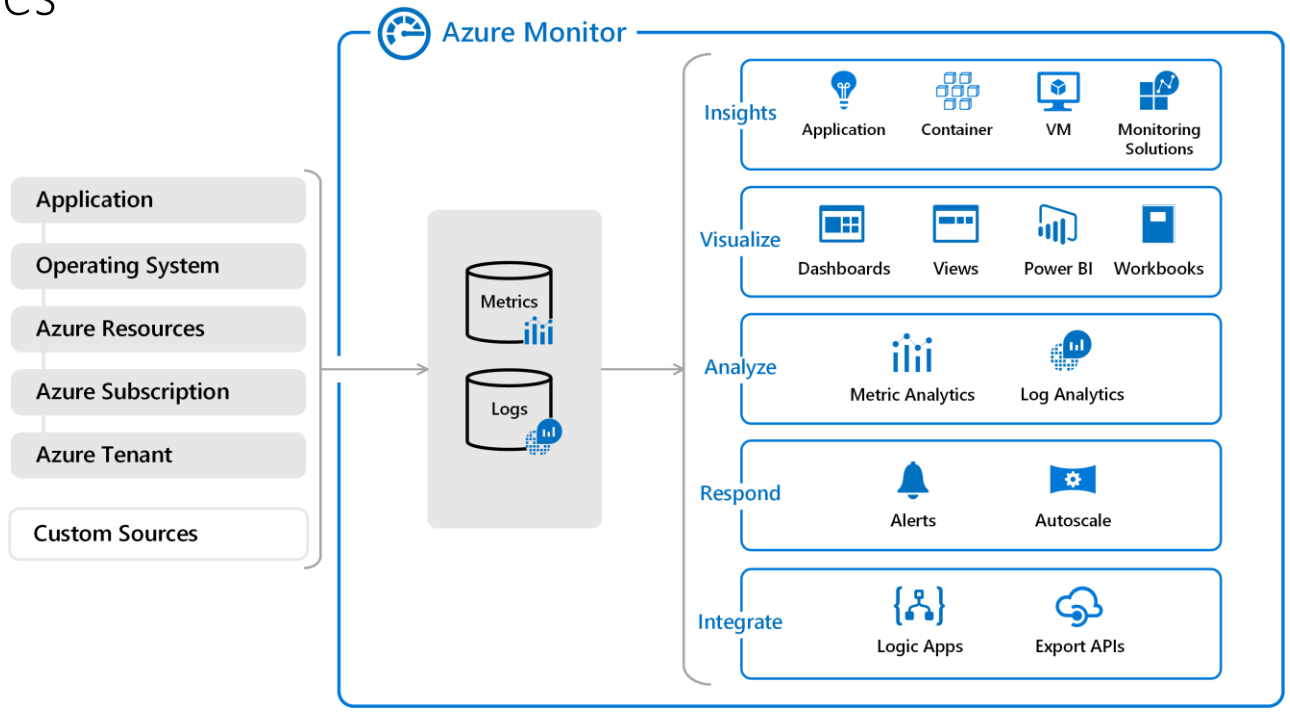
- Master Pipeline in Azure Data Factory



ORANGEMAN

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



ORANGEMAN

Usefull ressources

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

ORANGEMAN