STRATEGIC PARTNER

Microsoft

GOLD SPONSOR

elitmind
think bright

Future Processing

SILVER SPONSOR

dbWatch
DATABASE CONTROL

nationale
nederlanden

BRONZE SPONSOR

predica.

SQLDay 2020

Just Blindbæk

Traditional Modern Data Warehouse

SQLDay 2020

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)
  - Power BI Quiz show
- Strong focus on
  - Architecture
  - Analysis Services
  - Reporting Services
  - Power BI
- just@blindbaek.dk / blog.justB.dk / @justblindbaek / youtube.com/c/justblindbaek

AGENDA

- Big data, big data, big data

  - Modern Data Warehouse
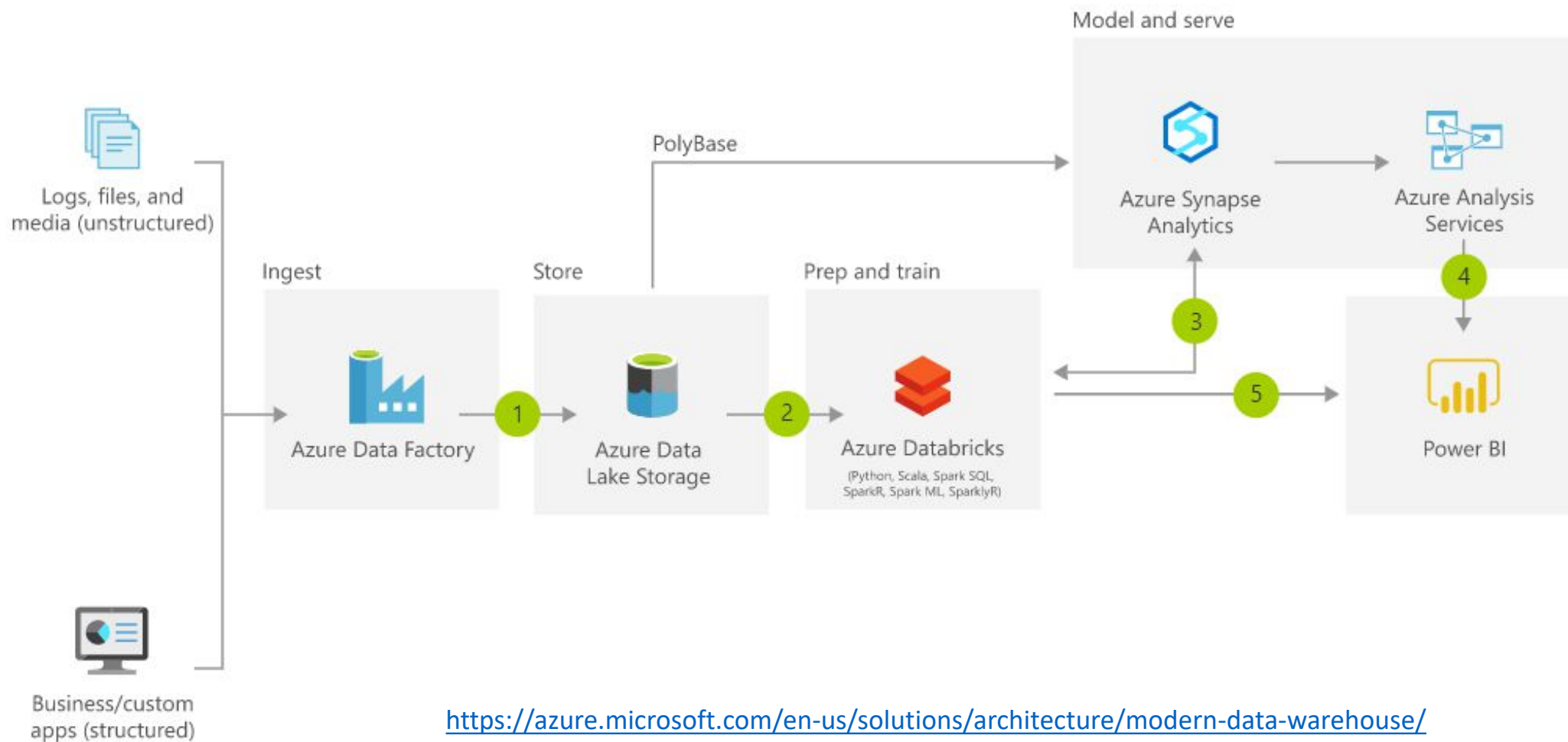
  - Synapse Analytics

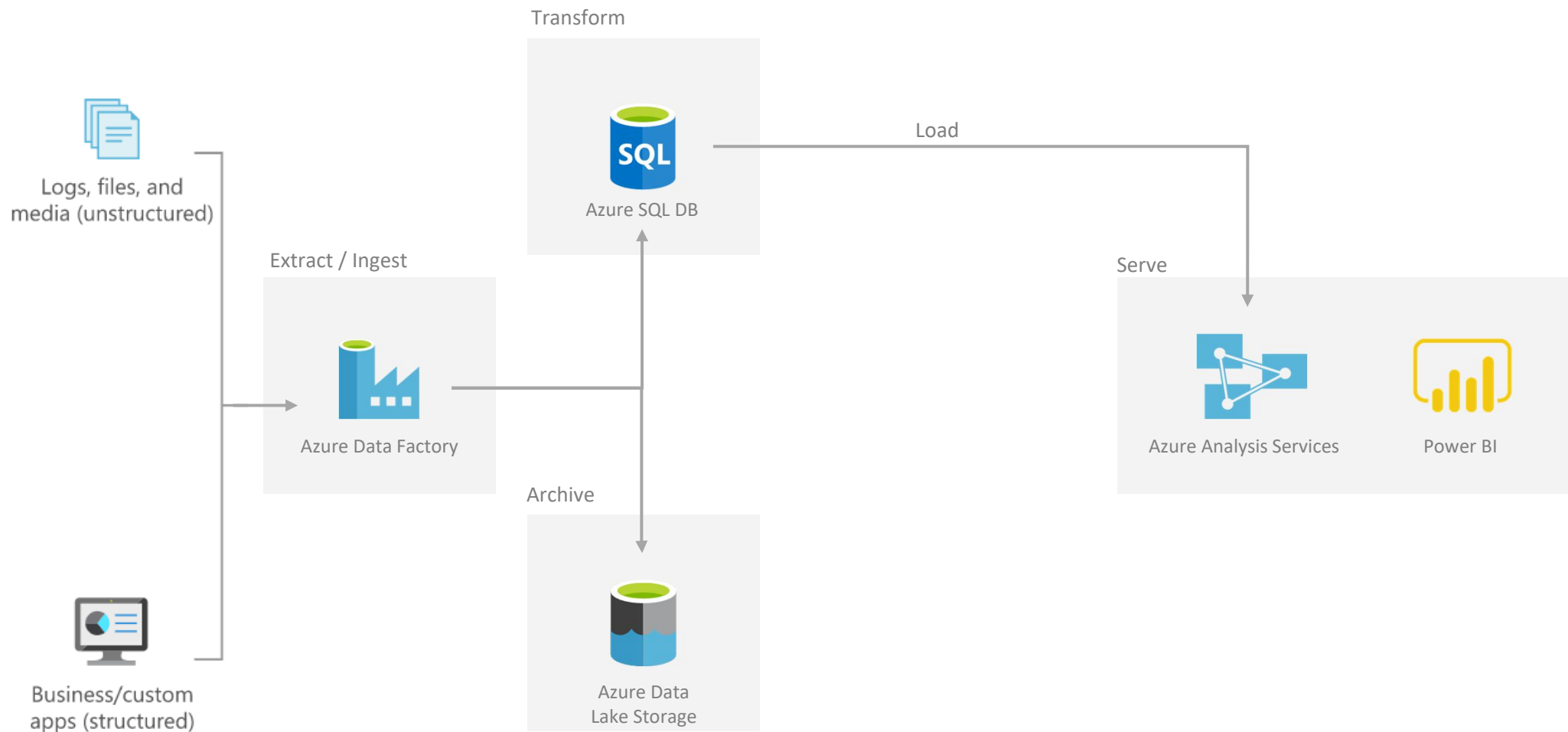- Small data – SMB market

  - ?

- Components

- Pricing estimate
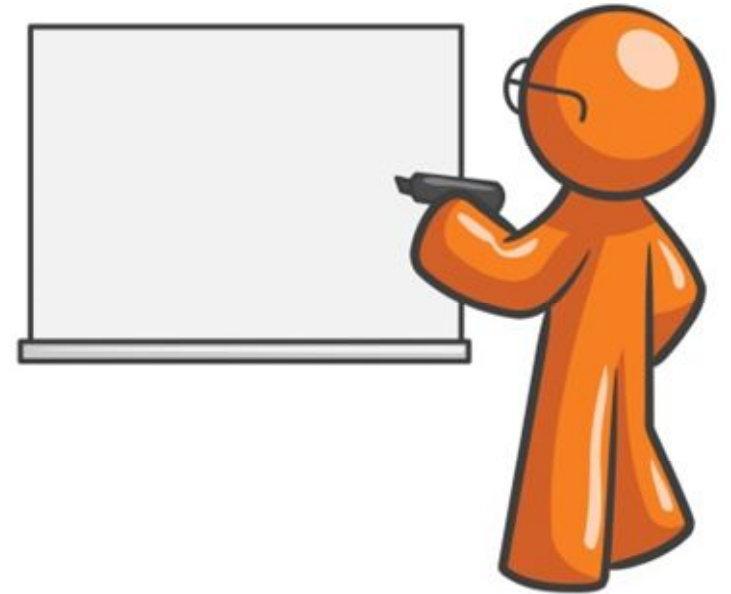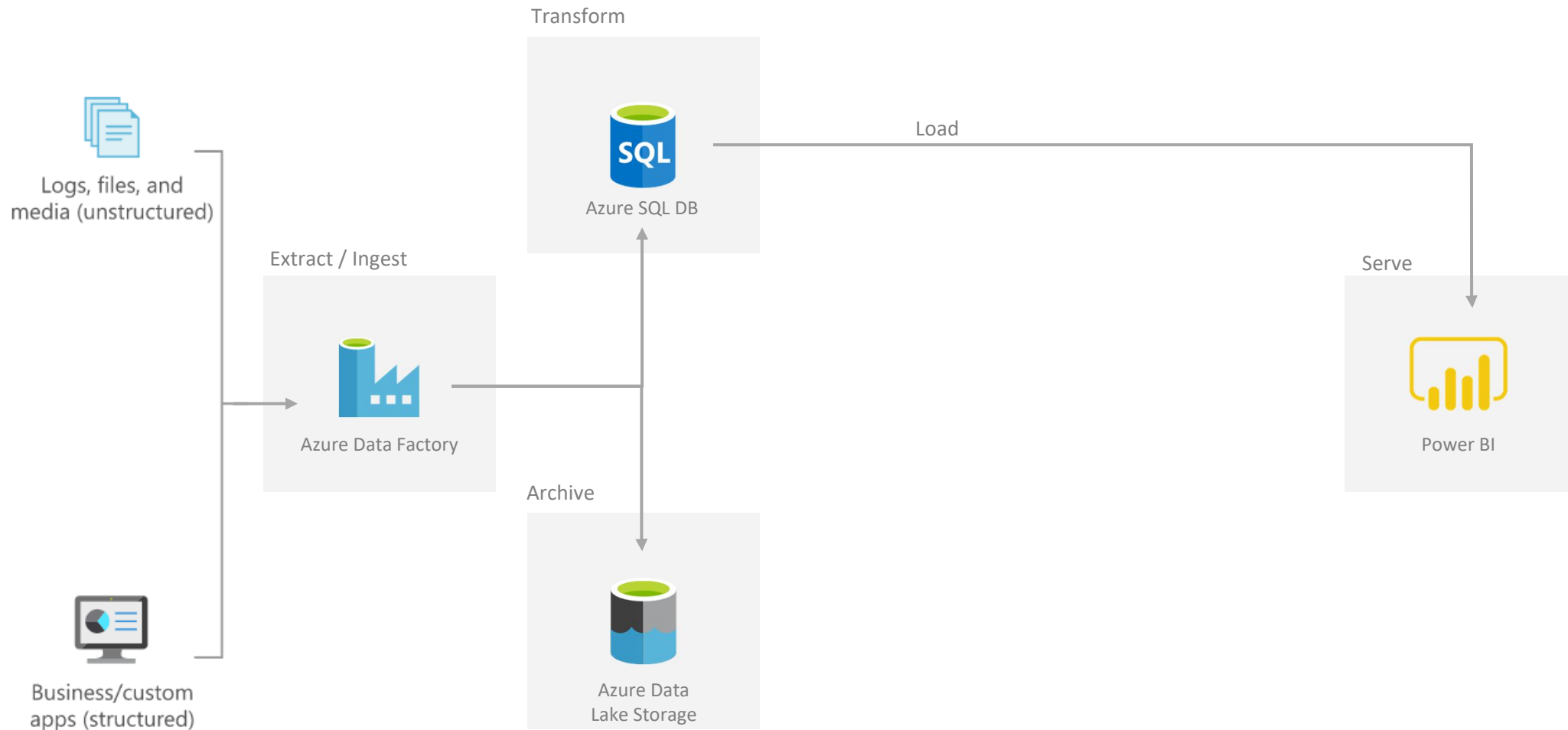
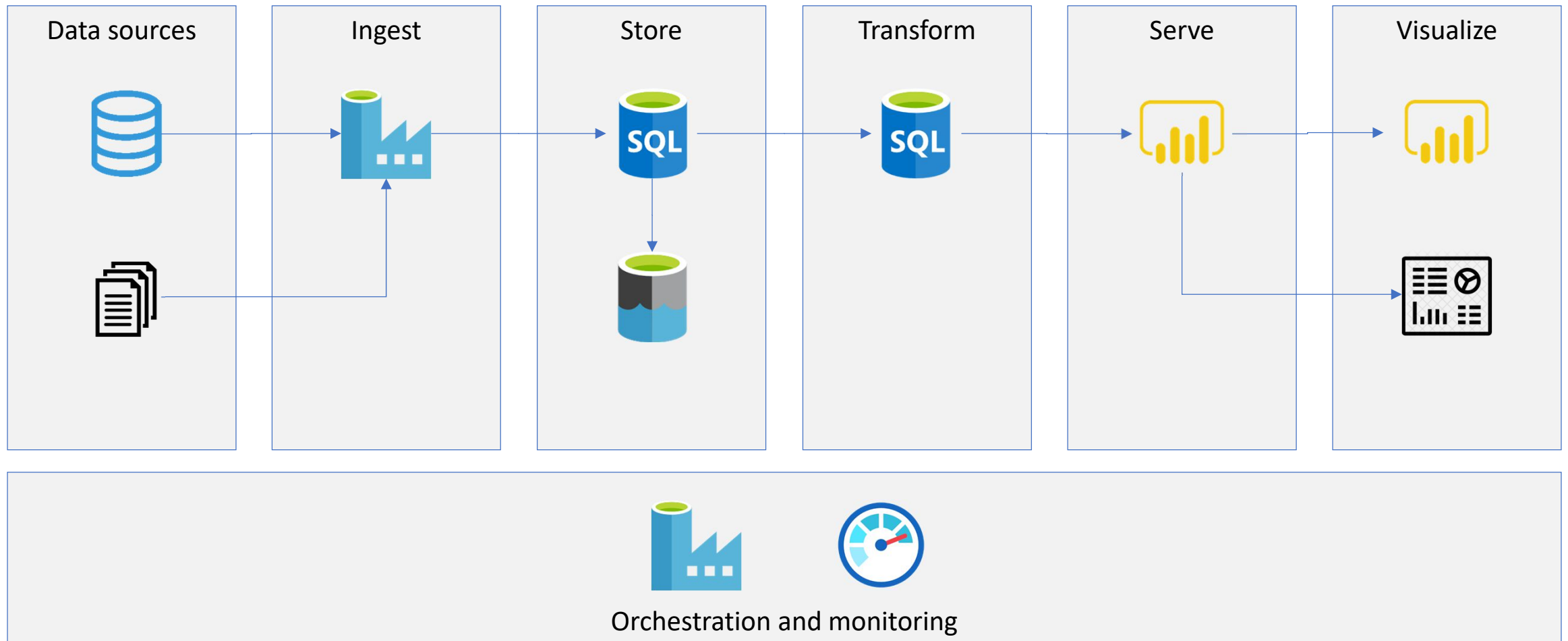# Reference architecture V1
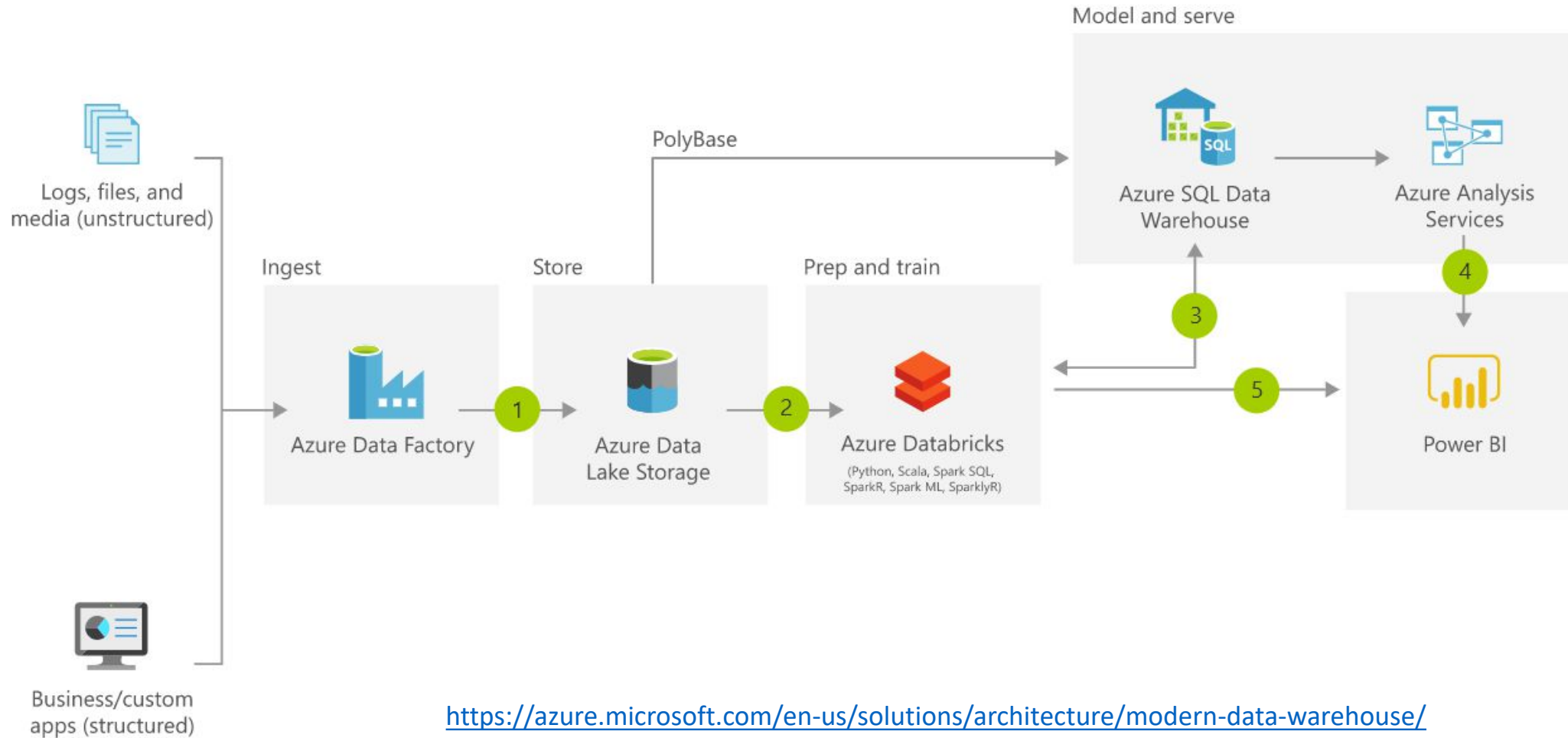
Reference architecture V2

https://azure.microsoft.com/en-us/solutions/architecture/modern-data-warehouse/

Traditional Data Warehouse

SQLDay 2020

Traditional Data Warehouse

Traditional Data Warehouse

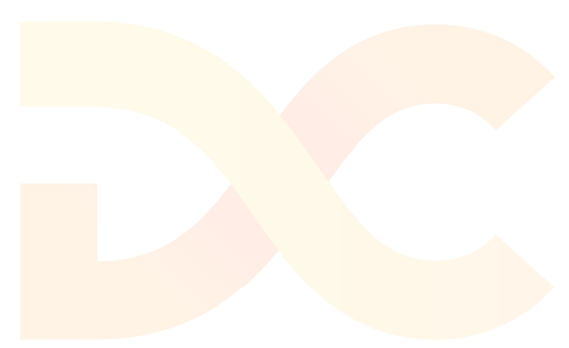| Data sources	Ingest	Store	Transform	Serve	Visualize

Orchestration and monitoring
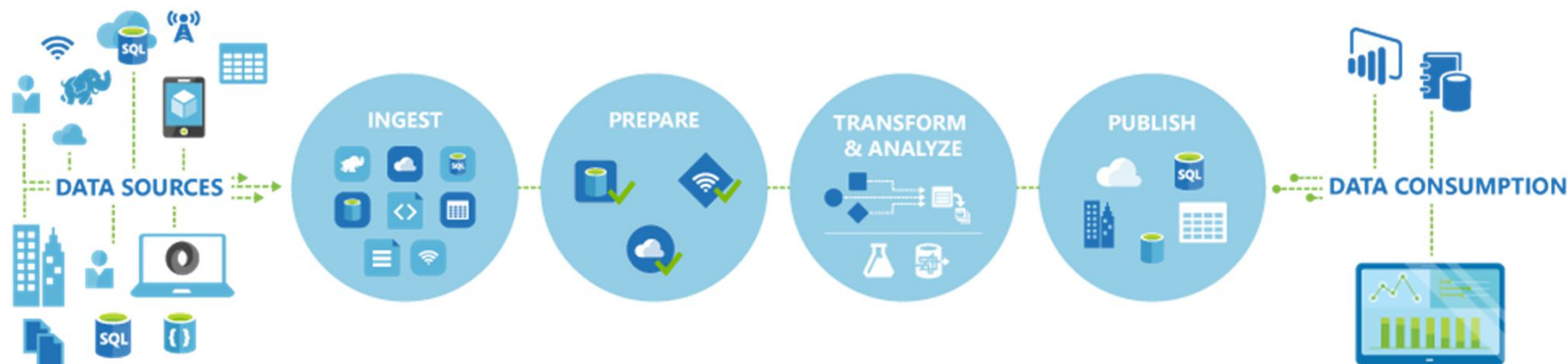
Traditional Data Warehouse
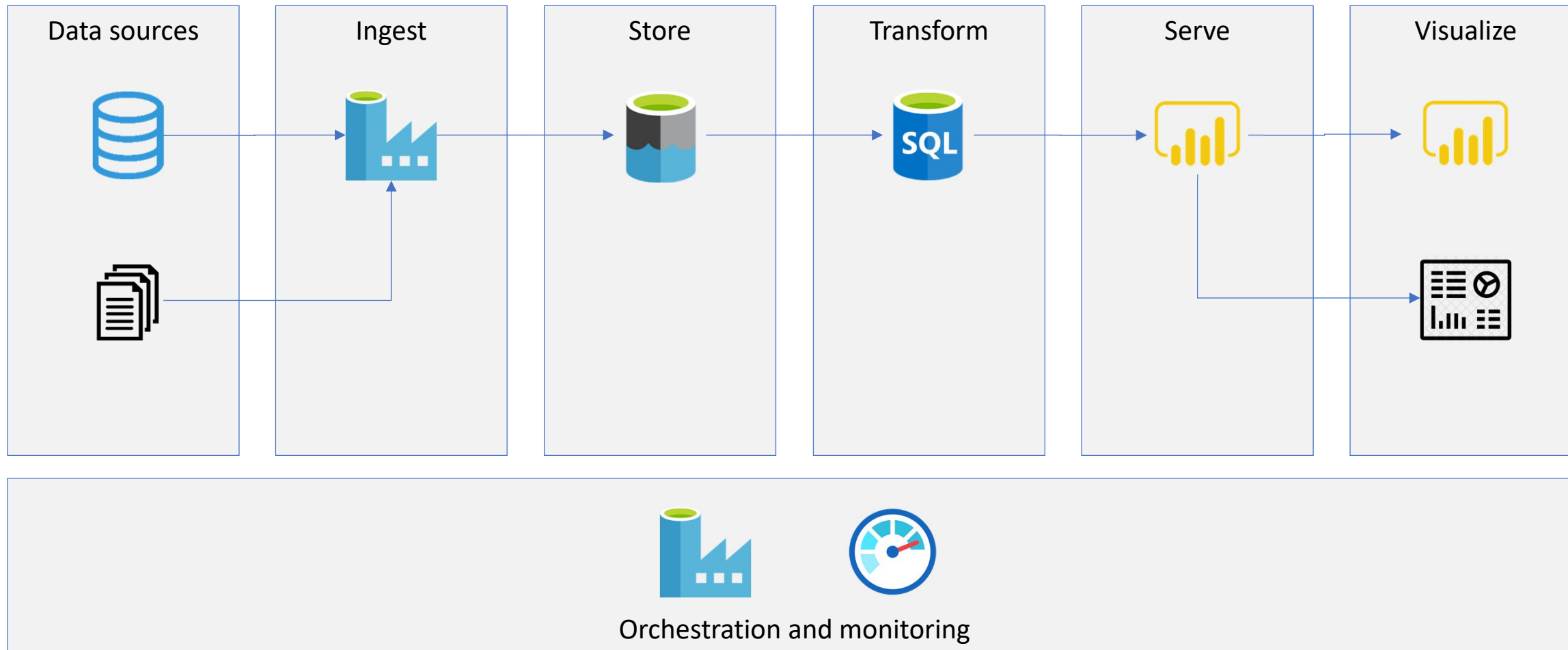
COMPONENTS

Data Factory, Data Lake Store, SQL DB and Power BI
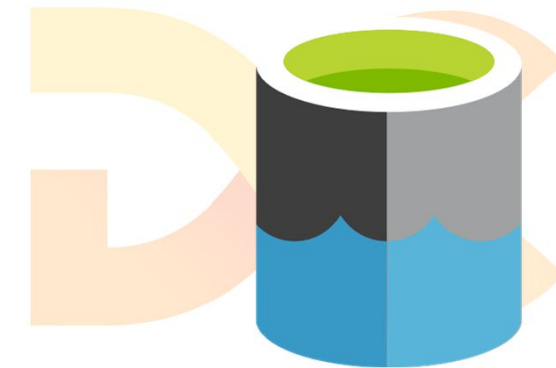
Azure Data Factory

- Hybrid data integration at scale

- Execute activities (copy, transform or orchestration)

- 80+ connectors provided to Cloud & Hybrid

- Create, plan, administer and monitor data pipelines

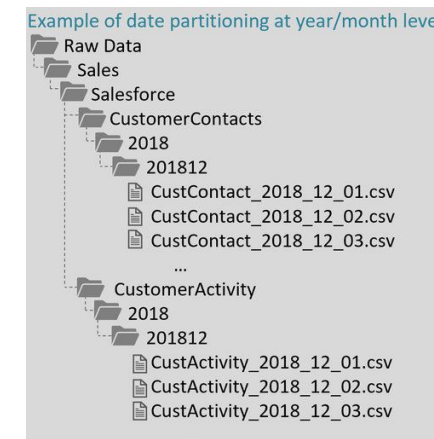- Parameters, control flow and triggers

Azure Data Lake Storage

- Unlimited cheap, cheap, cheap storage

- Can handle structured, semi-structured, and unstructured data

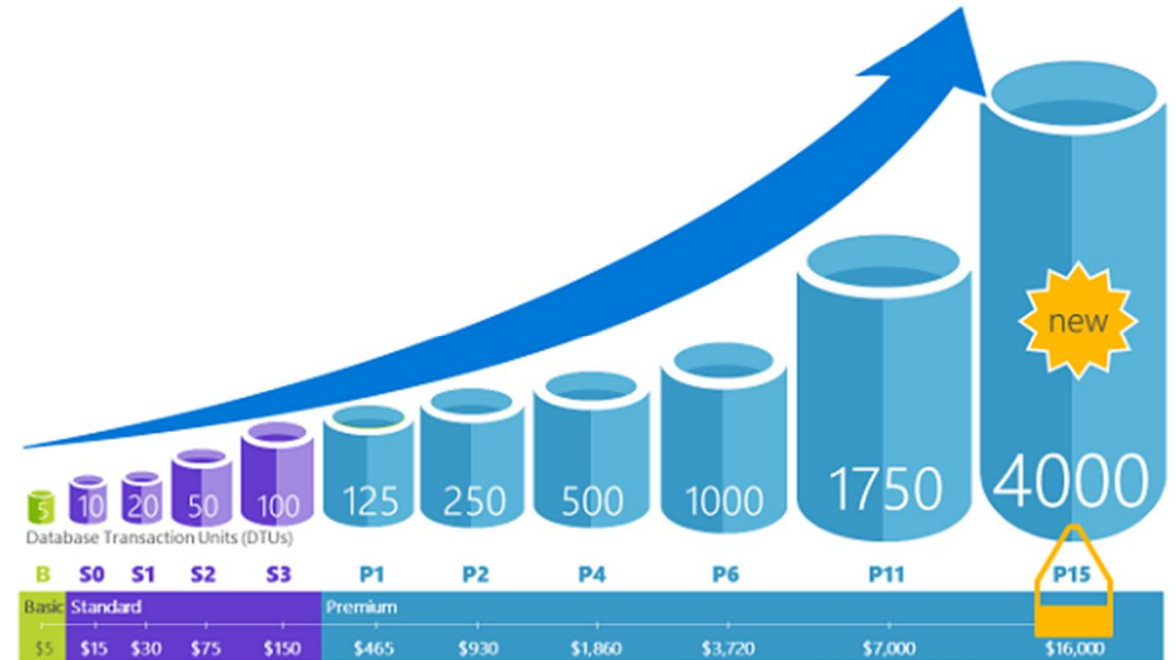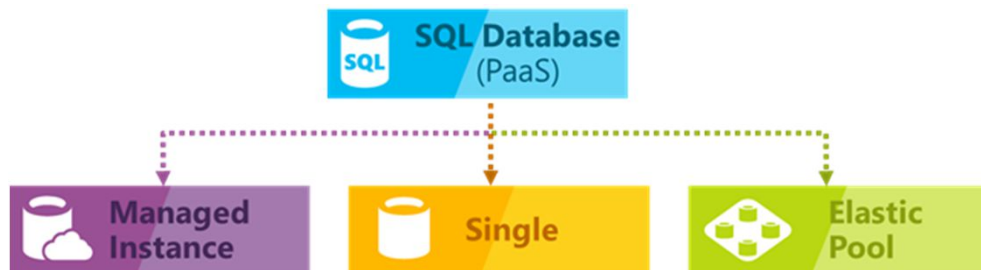- Apache Hadoop file system compatible (HDFS)

- Now we have Gen2 – combining Azure Blob Storage
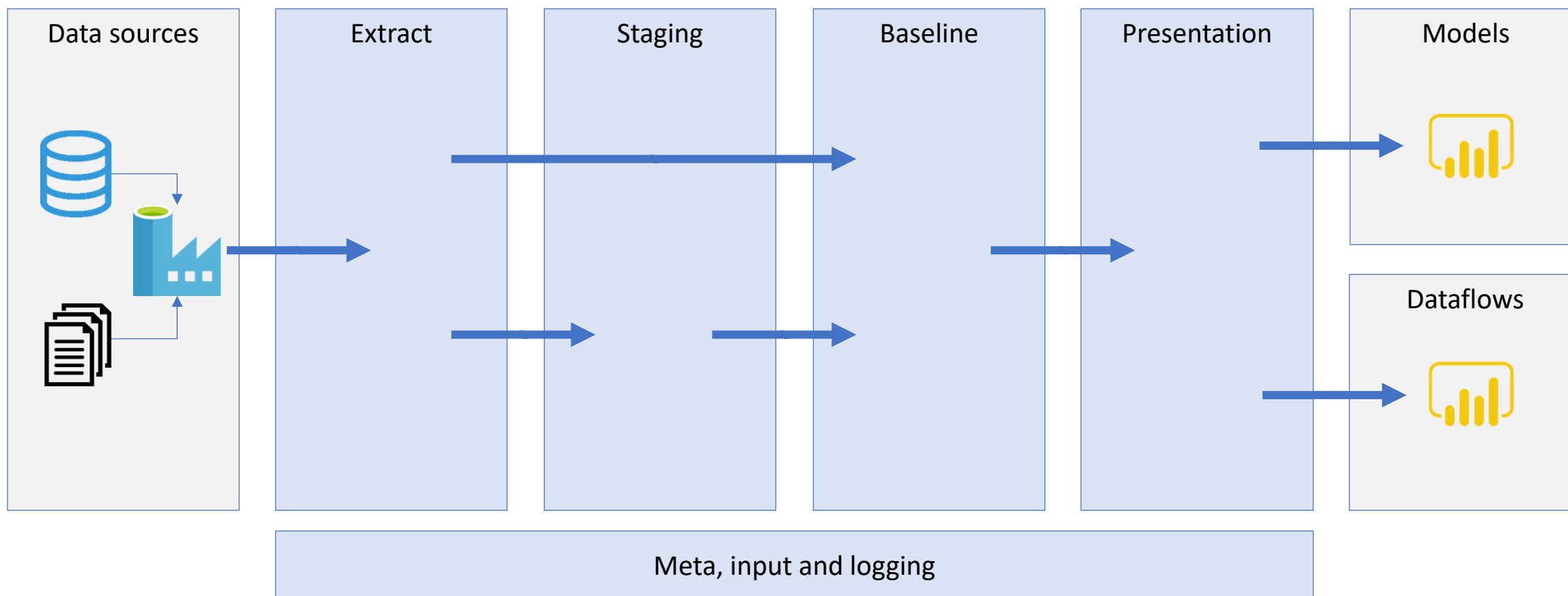
Azure SQL Database

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

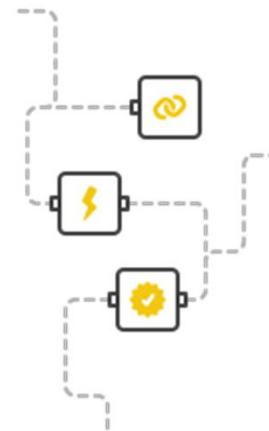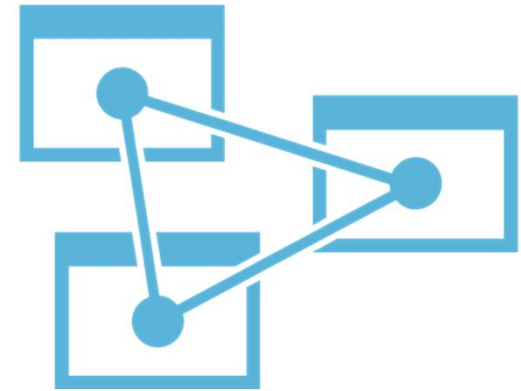# Transform in SQL DB

Serve with Power BI

- Stunning visuals, reports and dashboards
- Analysis Services engine inside
  - XMLA endpoint to manage the Semantic models
  - Security handled with RLS
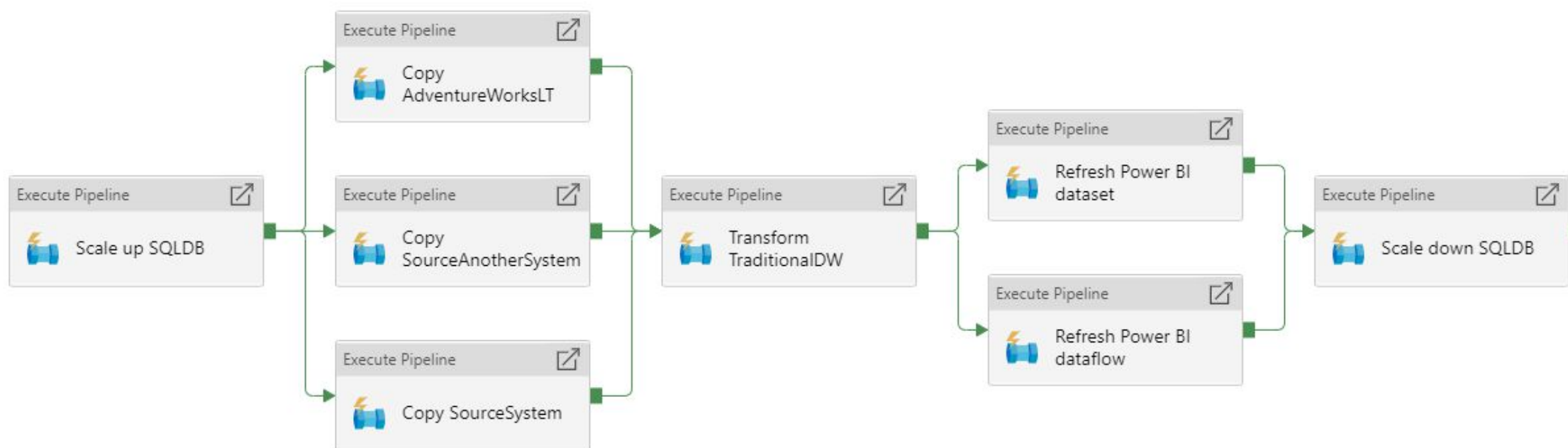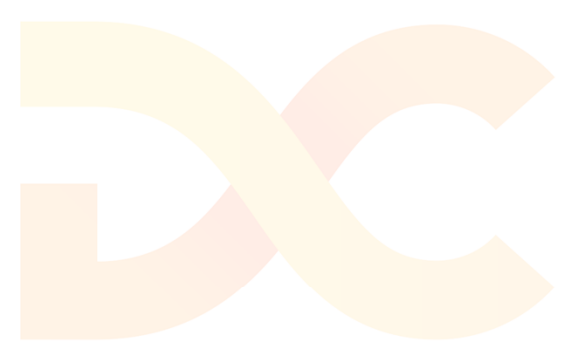- Dataflows
  - Expose data to business analysts

Orchestration

- Master Pipeline in Azure Data Factory
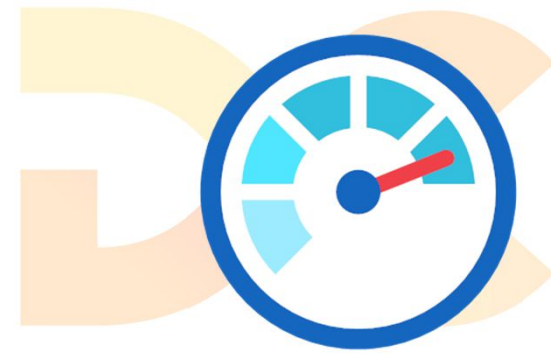
PRICING

Estimate

Pricing estimate

Service type	Description	Estimated Cost
Storage Accounts	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
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, 10 Hours	$ 10.08
Azure Monitor		$ 1.60
	Monthly Total	$ 110.60

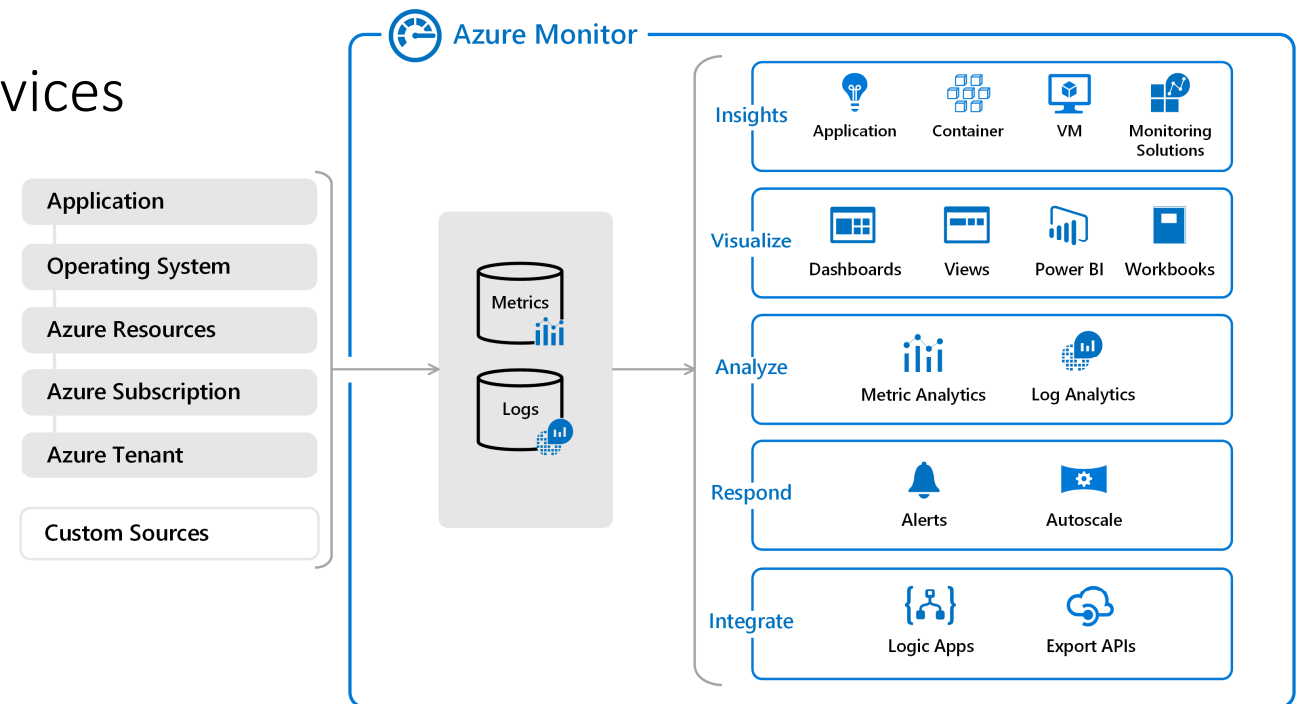Azure pricing calculator: https://azure.microsoft.com/en-us/pricing/calculator/
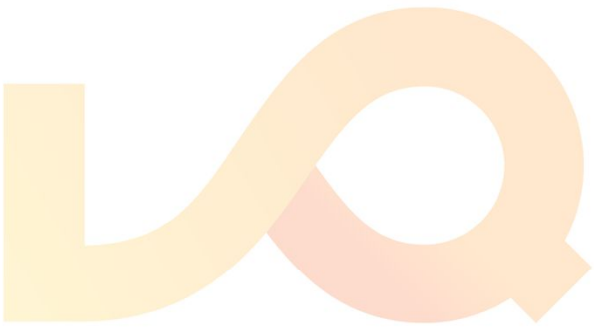
Monitoring

- 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

Usefull ressources

- Modern Data Warehouse reference: https://azure.microsoft.com/en-us/solutions/architecture/modern-data-warehouse/

- Pricing estimate: https://azure.com/e/7787ad5038584b12b62292abf51352ec

- Code on GitHub : https://github.com/justBlindbaek/TraditionalModernDW

STRATEGIC PARTNER

Microsoft

GOLD SPONSOR

SILVER SPONSOR

BRONZE SPONSOR

SQLDay 2020