Microsoft

Microsoft Fabric Architectures and Patterns

Delivering a federated data landscape

10th June 2025

Prepared by Andreas Bergstedt

Microsoft

About Me

Microsoft - Data and Analytics Global Black Belt (EMEA)
A lifetime in the IT Sector – 25 Years in Data

Industry solution experience: *FSI, FMCG, Healthcare, Pharma, Mining, Manufacturing, Transport and Logistics, Energy, Public Sector + NGO, Non-profit, Professional Services, Local and National Government, Defense, Aviation, Education and Retail.*

Agenda

1. Microsoft Fabric Overview
2. Patterns and Principles
3. Data Architecture
4. Cloud ODS
5. Federated Analytics
6. Capacity and Workloads

Microsoft

Microsoft Fabric

Microsoft Fabric

The unified data platform for AI transformation

AI-powered data platform

Microsoft Fabric

The unified data platform for AI transformation

AI-powered data platform

Open and AI-ready data lake

Microsoft Fabric

The unified data platform for AI transformation

AI-powered data platform

Open and AI-ready data lake

AI-enabled business users

Microsoft Fabric

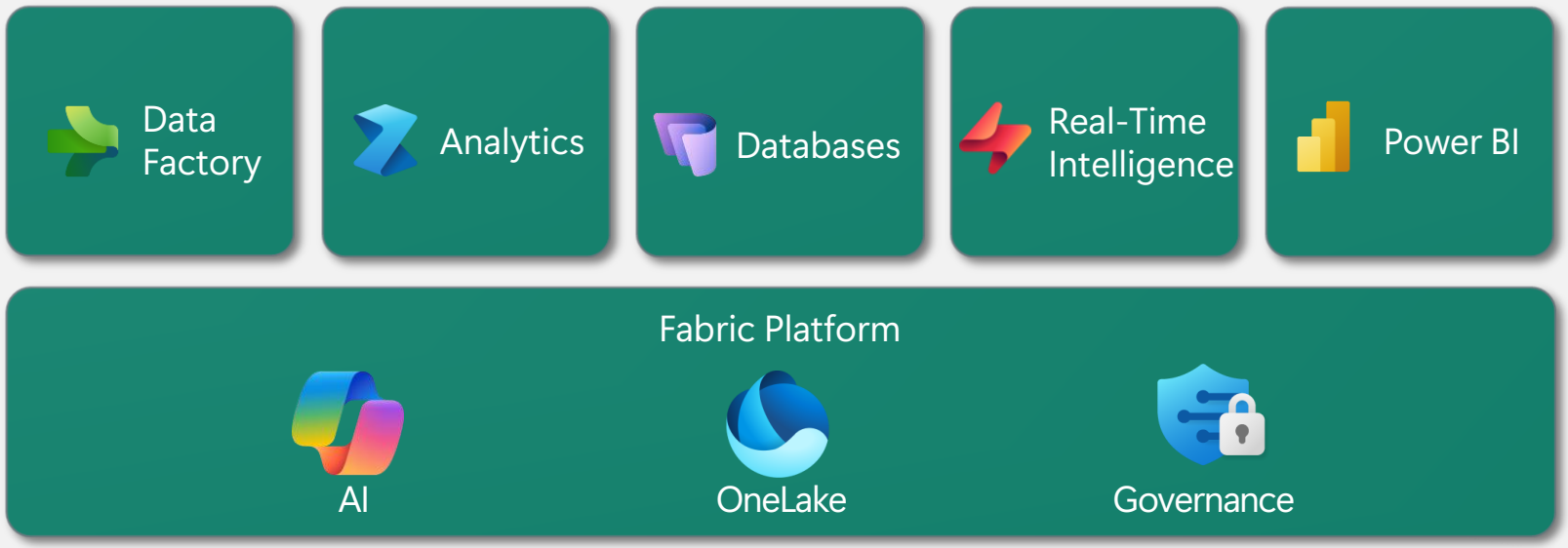
The unified data platform for AI transformation

AI-powered data platform

Open and AI-ready data lake

AI-enabled business users

Microsoft Fabric



The unified data platform for AI transformation

Single…

- Onboarding and trials
- Sign-on
- Navigation model
- UX model
- Workspace organization
- Collaboration experience
- Data Lake
- Storage format
- Data copy for all engines
- Security model
- CI/CD
- Monitoring hub
- Data hub
- Governance & compliance

Microsoft

Patterns and Principles

Data Lakehouse Pattern

Although the 3-layered design is common and well-known, there are many discussions on the scope, purpose, and best practices on each of these layers.

Bronze layer

Typically raw, "as-is"

- Maintains the raw state in the structure “as-is”
- Data is immutable (read-only)
- Delivery-based partitioned tables, i.e., YYYYMMDD
- Mostly Parquet. Sometimes other formats
- Can be any combination of streaming and batch transactions
- May include extra metadata (schema)
- May be fed from a “mediation layer”
- Used for debugging, testing

Silver layer

Matched and conformed

- Uses data quality rules for validation
- Usually only functional data
- Historization is merged (SCD2)
- Efficient storage format; Delta
- Versioning for rolling back
- Handles missing or incorrect data
- Usually enriched with reference data
- Source-oriented, although queryable and cluttered around subject areas
- Usually used by operational analytical teams

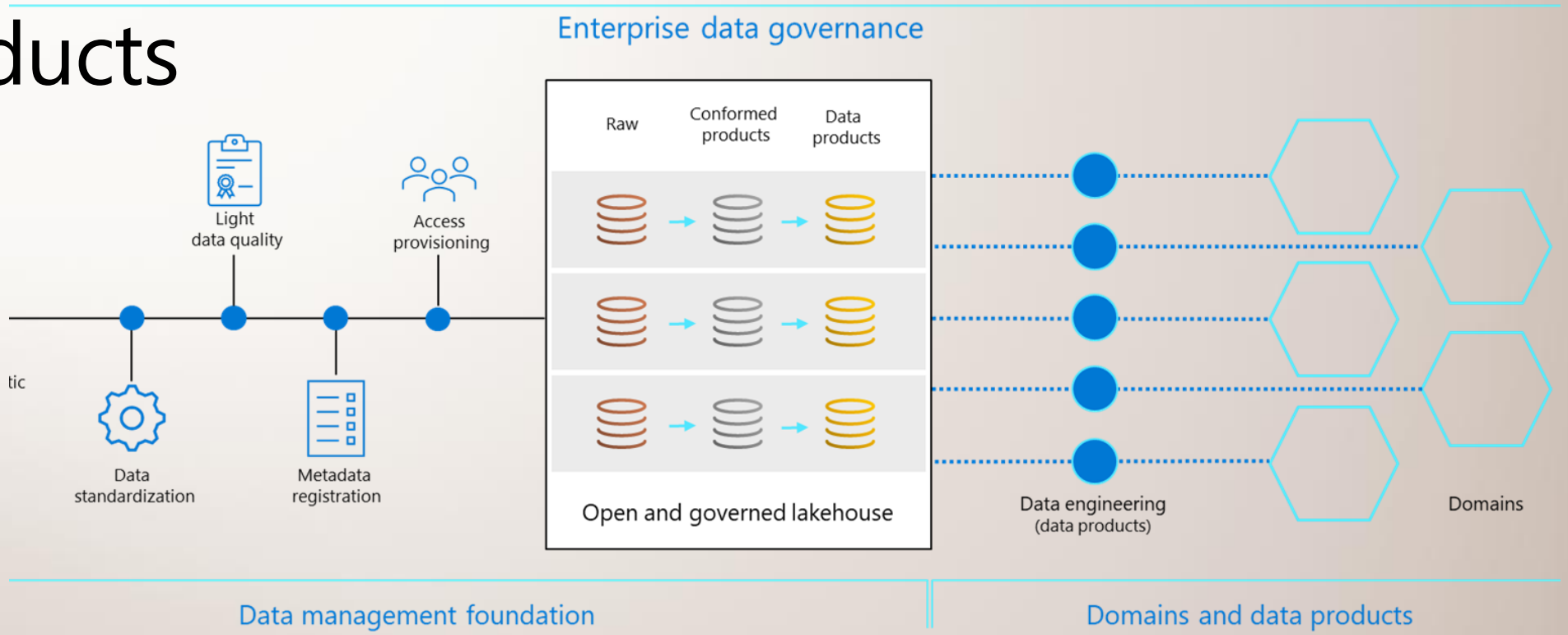
Gold layer

Refined business-level

- What enterprises call data products: consumer-ready / user-friendly data
- Data is highly governed and well-documented
- Historization is applied only for the set of use cases or consumers
- Contains complex business rules, such as calculations and enrichments
- Efficient storage format; Delta
- Versioning for rolling back
- Might contain additional sub layers for sharing or distributing data

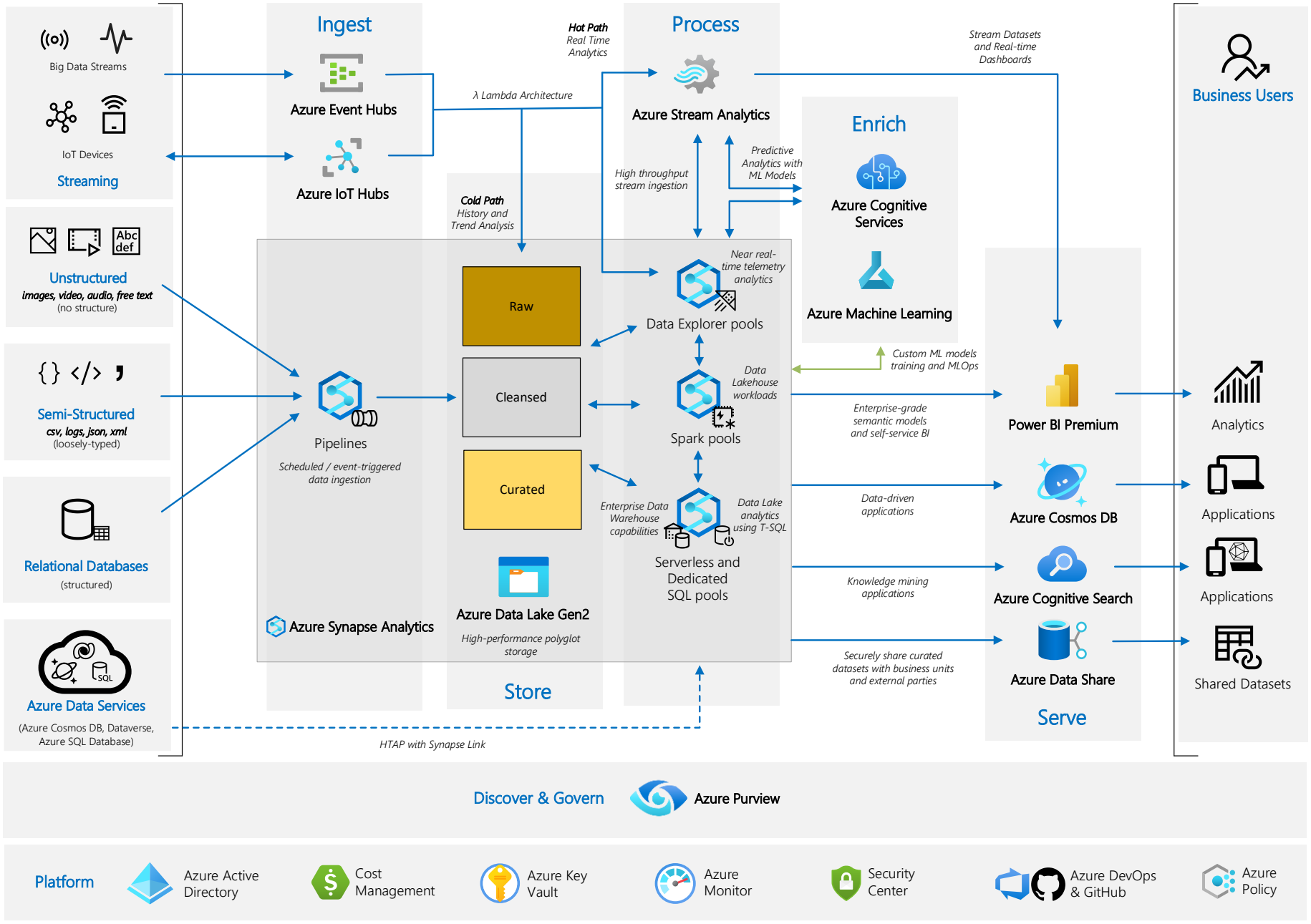
Microsoft

Data Products

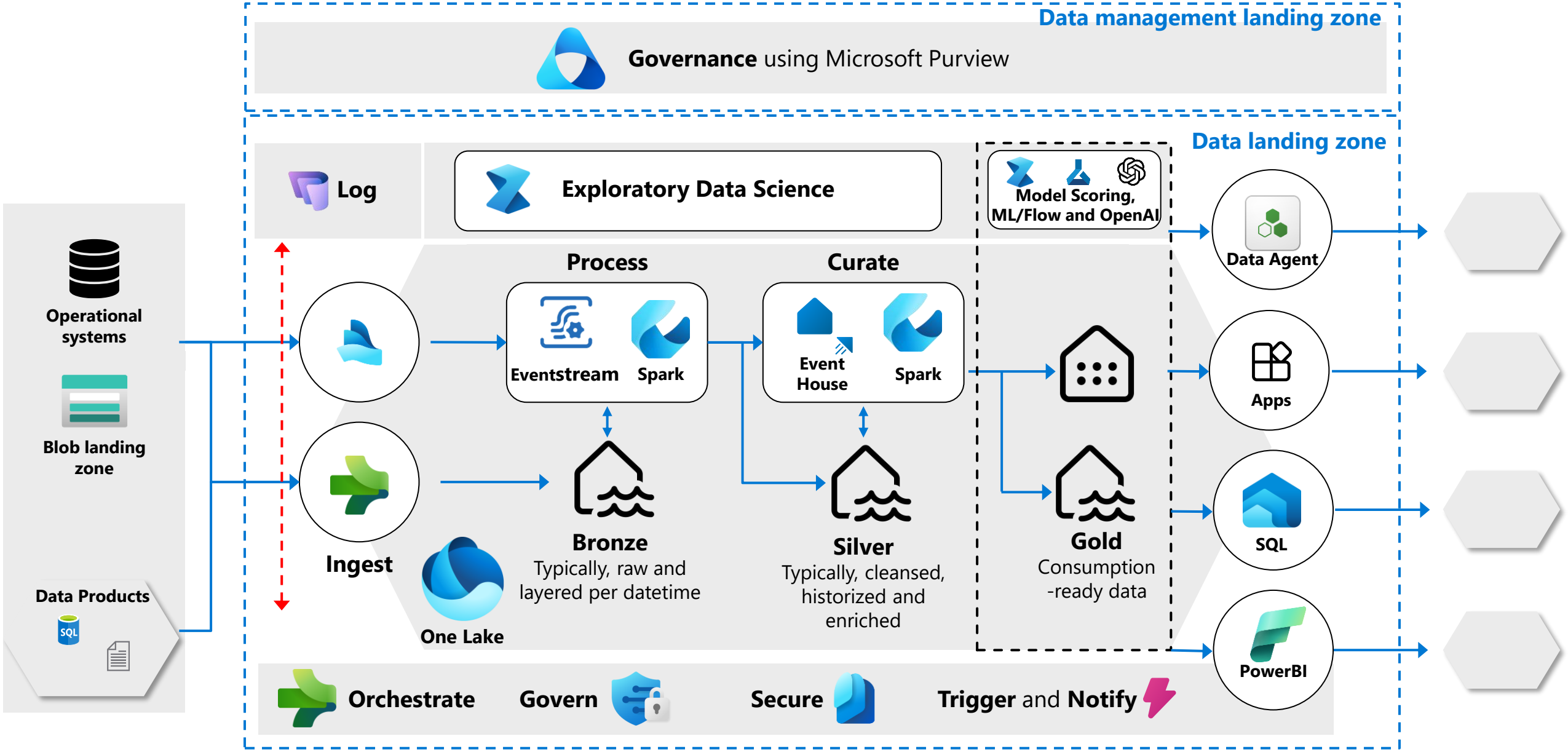


Microsoft

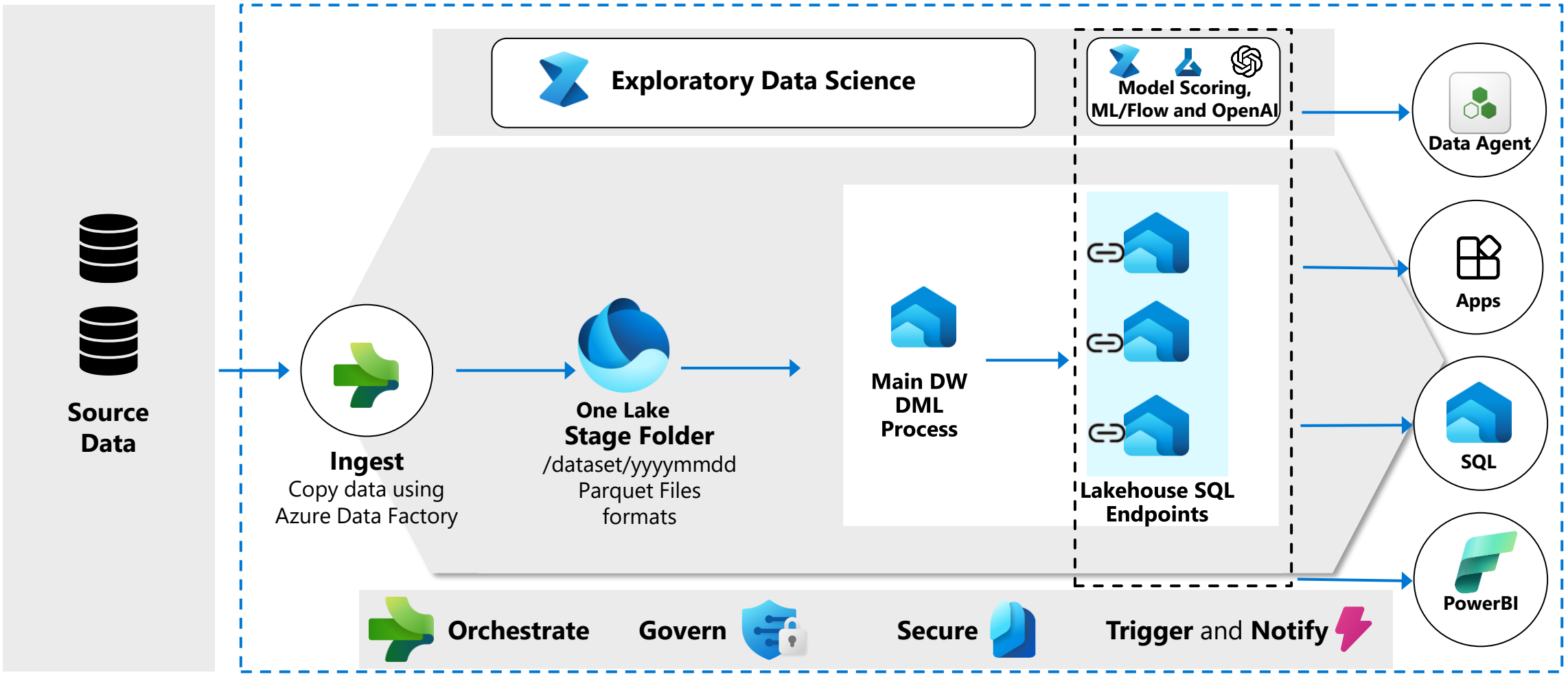
Data Architecture


Big Data Streams
IoT Devices
Streaming
Unstructured
images, video, audio, free text
(no structure)
Semi-Structured
csv, logs, json, xml
(loosely-typed)
Relational Databases
(structured)
Azure Data Services
(Azure Cosmos DB, Dataverse, Azure SQL Database)
Ingest
Azure Event Hubs
Azure IoT Hubs
λ Lambda Architecture
Hot Path
Real Time Analytics
Cold Path
History and Trend Analysis
Process
Azure Stream Analytics
High throughput stream ingestion
Predictive Analytics with ML Models
Enrich
Azure Cognitive Services
Azure Machine Learning
Stream Datasets and Real-time Dashboards
Business Users
Pipelines
Scheduled / event-triggered data ingestion
Raw
Cleansed
Curated
Near real-time telemetry analytics
Data Explorer pools
Data Lakehouse workloads
Spark pools
Enterprise Data Warehouse capabilities
Data Lake analytics using T-SQL
Serverless and Dedicated SQL pools
Azure Synapse Analytics
Azure Data Lake Gen2
High-performance polyglot storage
Store
Custom ML models training and MLOps
Enterprise-grade semantic models and self-service BI
Data-driven applications
Knowledge mining applications
Securely share curated datasets with business units and external parties
Power BI Premium
Azure Cosmos DB
Azure Cognitive Search
Azure Data Share
Serve
Analytics
Applications
Applications
Shared Datasets
HTAP with Synapse Link
Discover & Govern
Azure Purview
Platform
Azure Active Directory
Cost Management
Azure Key Vault
Azure Monitor
Security Center
Azure DevOps & GitHub
Azure Policy

End to End Analytics with Microsoft Fabric



Modern Data Warehouse Pattern in Microsoft Fabric

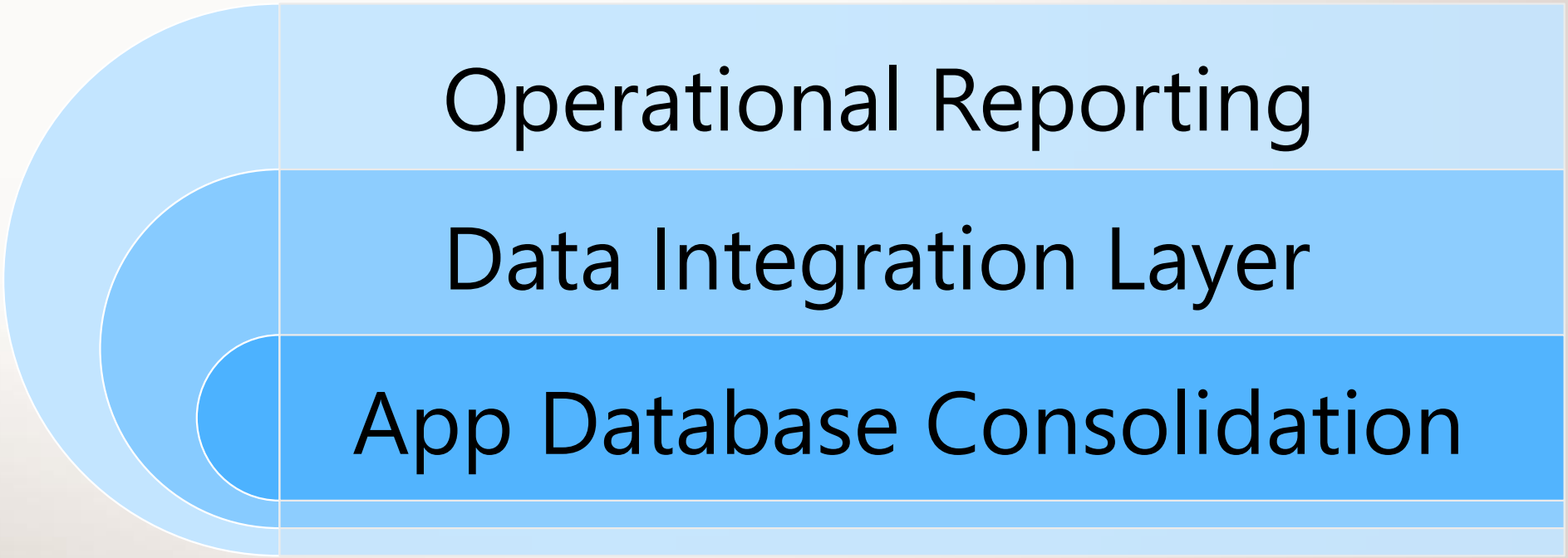


Microsoft

Cloud ODS

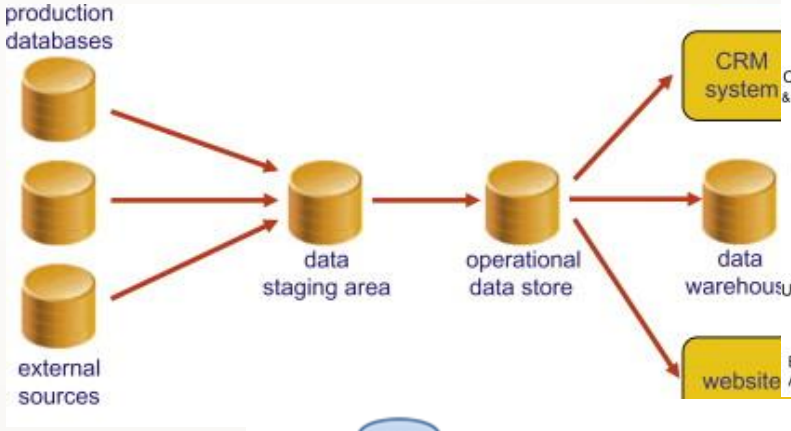
Microsoft

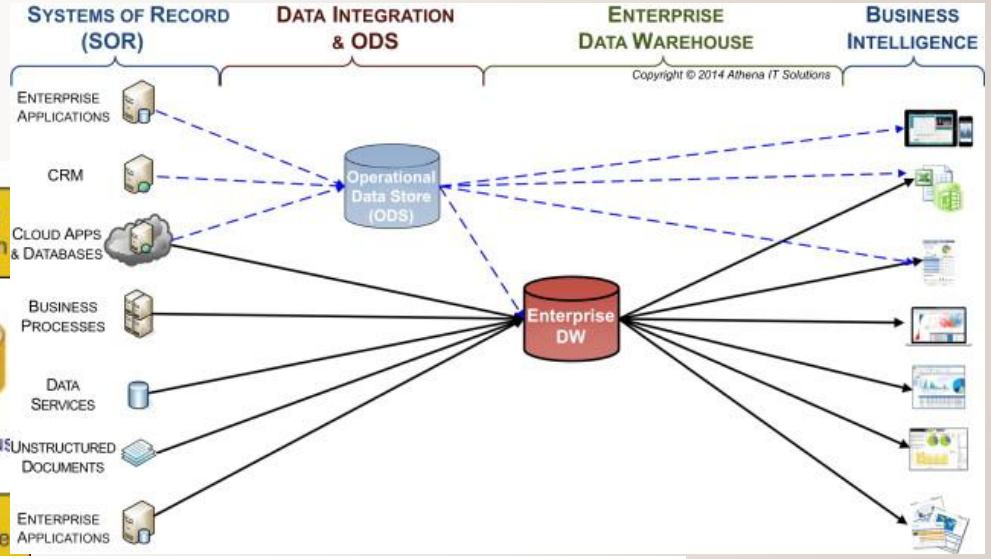
Cloud ODS

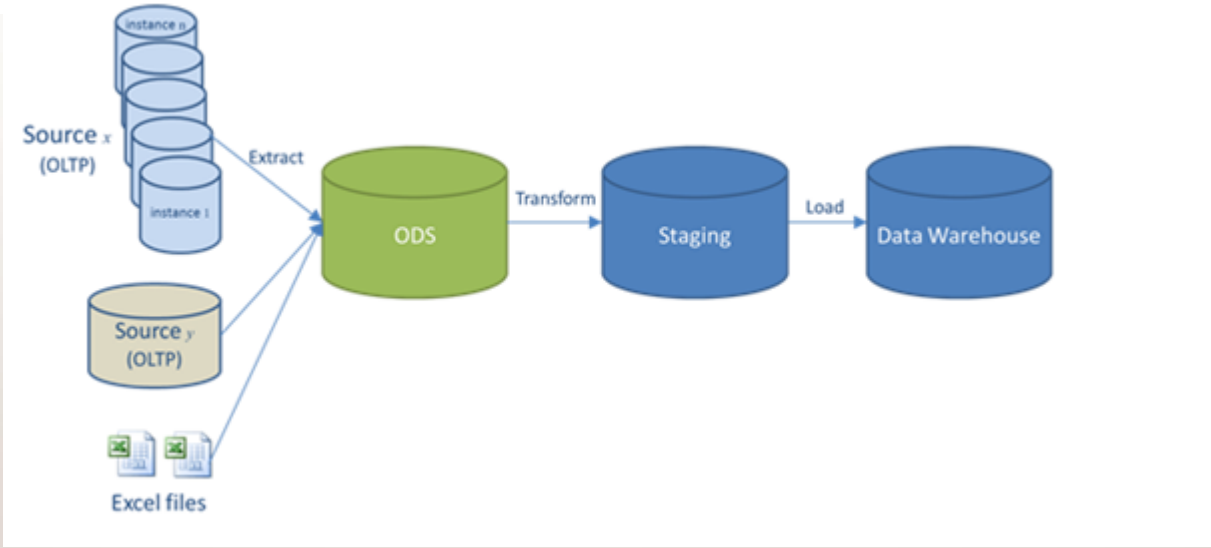


Microsoft

Cloud ODS

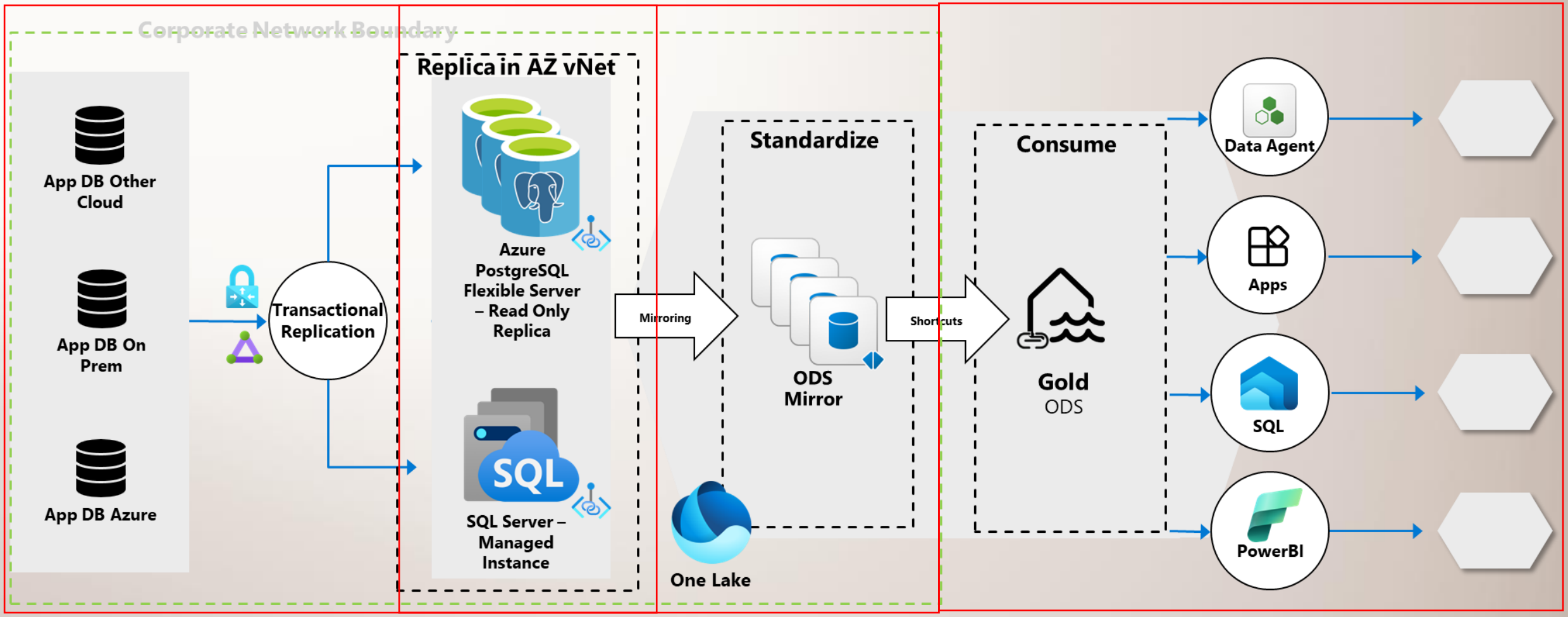






Microsoft

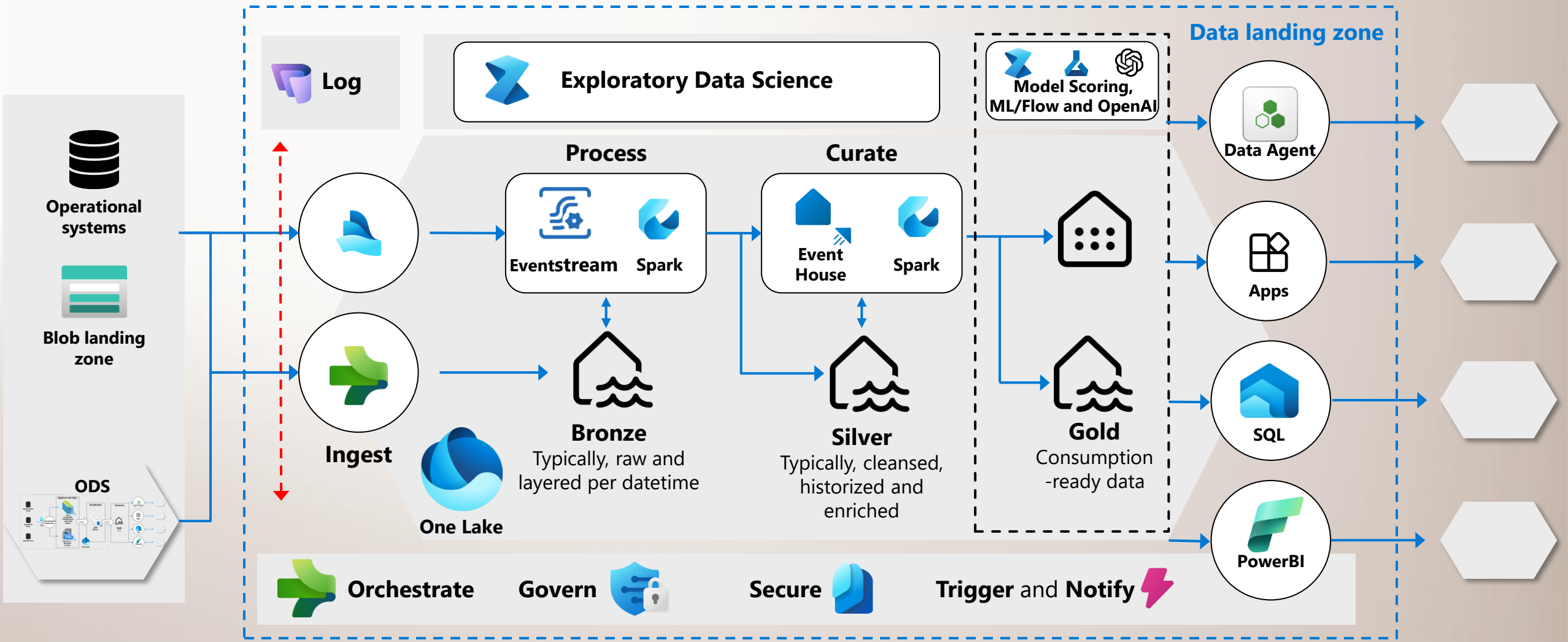
Cloud ODS



Microsoft

Unified Data Governance

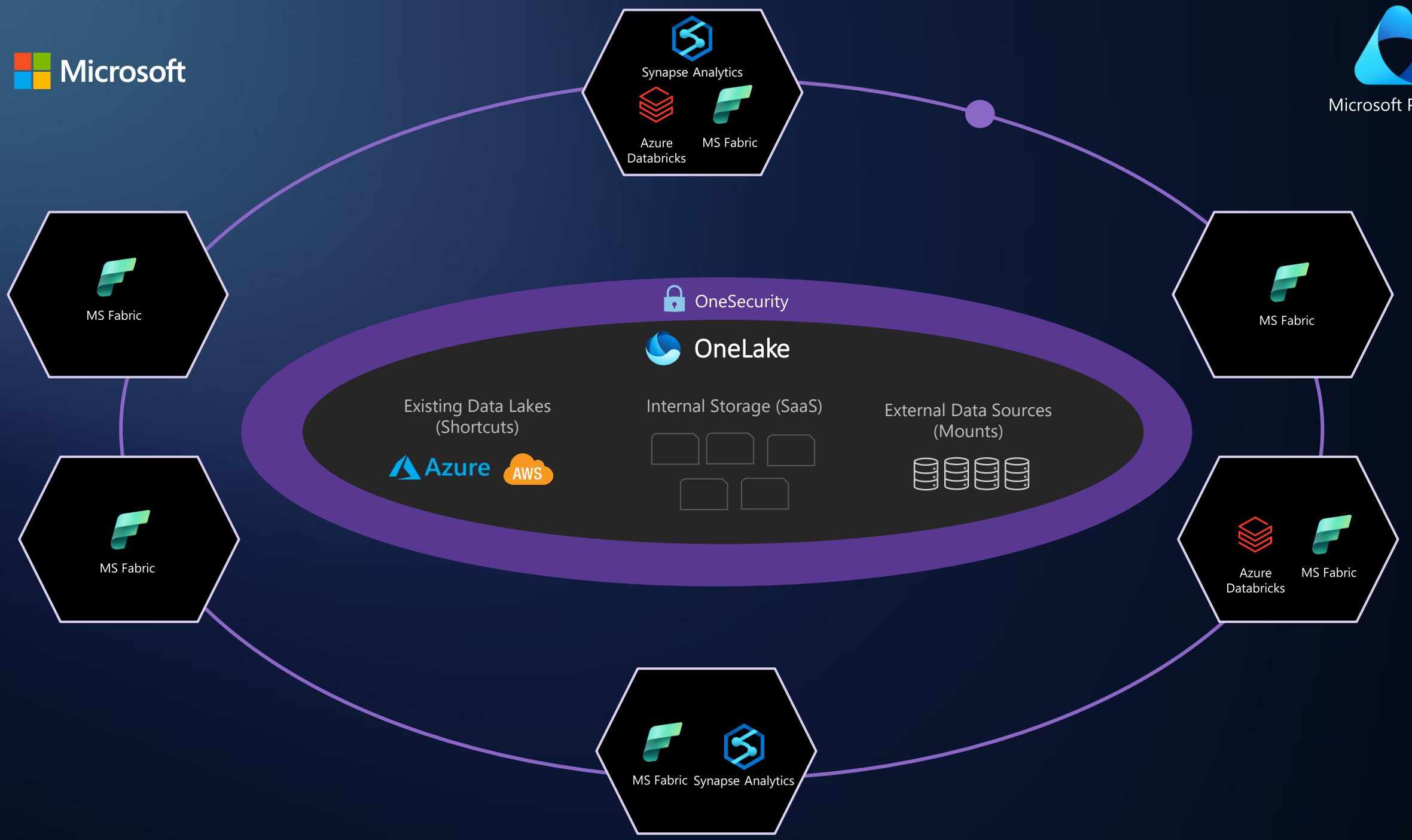
End to End Analytics with Microsoft Fabric



Microsoft

Federated Analytics


Microsoft
Microsoft Purview
Finance
Fabric Domain
Data Factory
Analytics
Fabric Databases
Real-time Intelligence
Power BI
Workspace 1
Workspace 2
Marketing
Fabric Domain
Data Factory
Analytics
Fabric Databases
Real-time Intelligence
Power BI
Workspace 1
Workspace 2
OneLake


Microsoft
Microsoft Purview
Synapse Analytics
Azure Databricks
MS Fabric
MS Fabric
MS Fabric
OneSecurity
OneLake
Existing Data Lakes (Shortcuts)
Azure
AWS
Internal Storage (SaaS)
External Data Sources (Mounts)
MS Fabric
Azure Databricks
MS Fabric
MS Fabric
Synapse Analytics

Microsoft

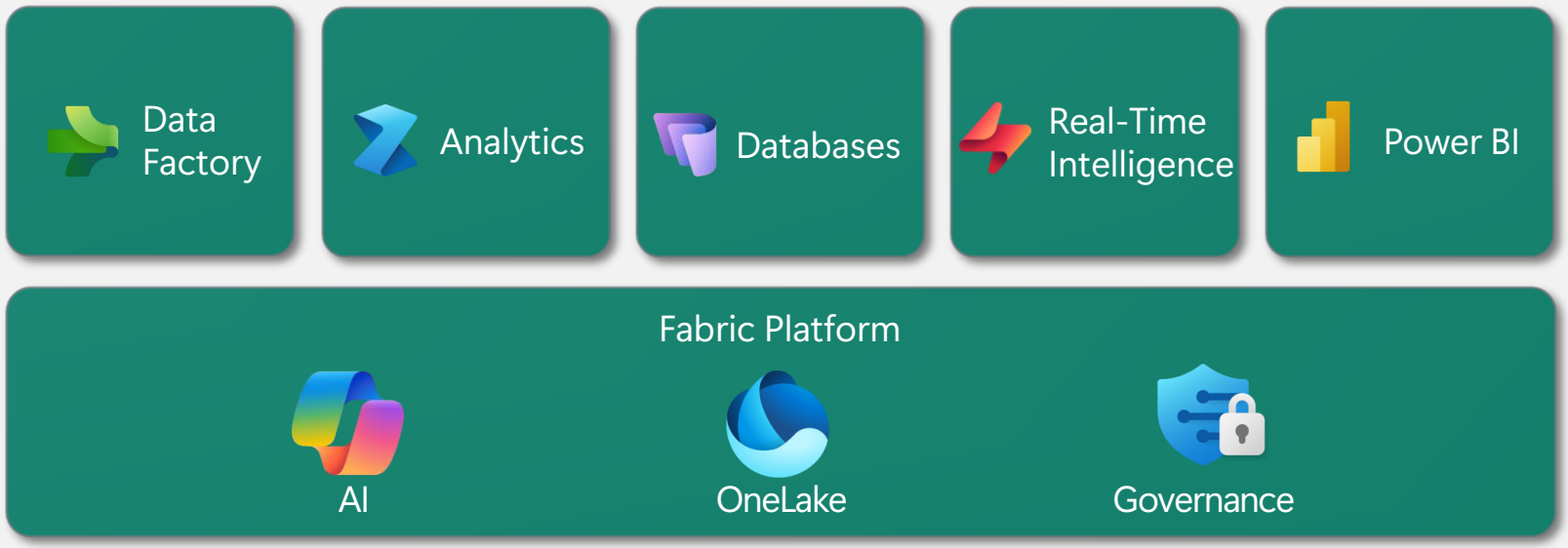
Key Take Aways for Fabric Architecture

- Fabric addresses the modern data paradigms and principles independent of workloads eliminating compute and data silos.
- Fabric's OneCopy promise removes data duplication by standardising data consumption within OneLake with open standards.
- Fabric caters for multi disciplined data and analytics teams at any skill level.

Microsoft

Capacity and Workloads

Microsoft Fabric

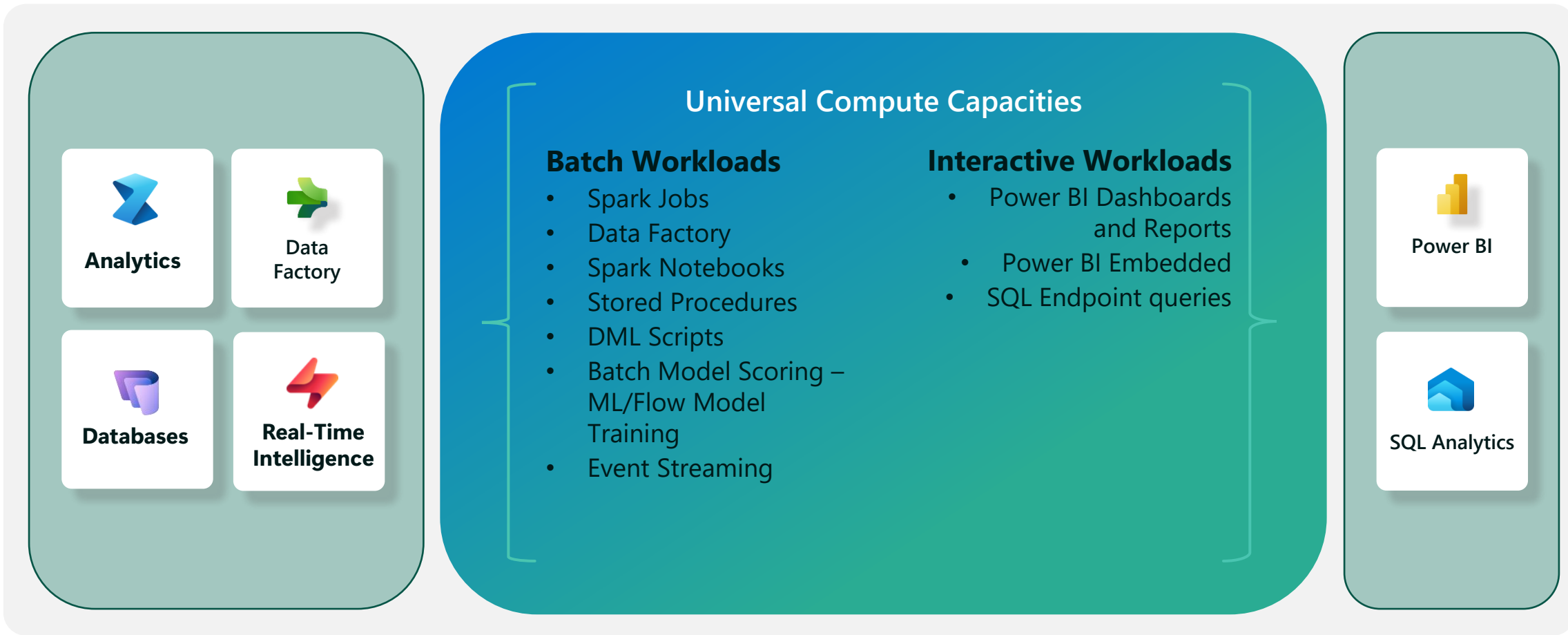


The unified data platform for AI transformation

Single…

- Onboarding and trials
- Sign-on
- Navigation model
- UX model
- Workspace organization
- Collaboration experience
- Data Lake
- Storage format
- Data copy for all engines
- Security model
- CI/CD
- Monitoring hub
- Data hub
- Governance & compliance

Microsoft Fabric



Microsoft Fabric

Universal Compute Capacities

Batch Workloads

- Spark Jobs
- Data Factory
- Spark Notebooks
- Stored Procedures
- DML Scripts
- Batch Model Scoring – ML/Flow Model Training
- Event Streaming

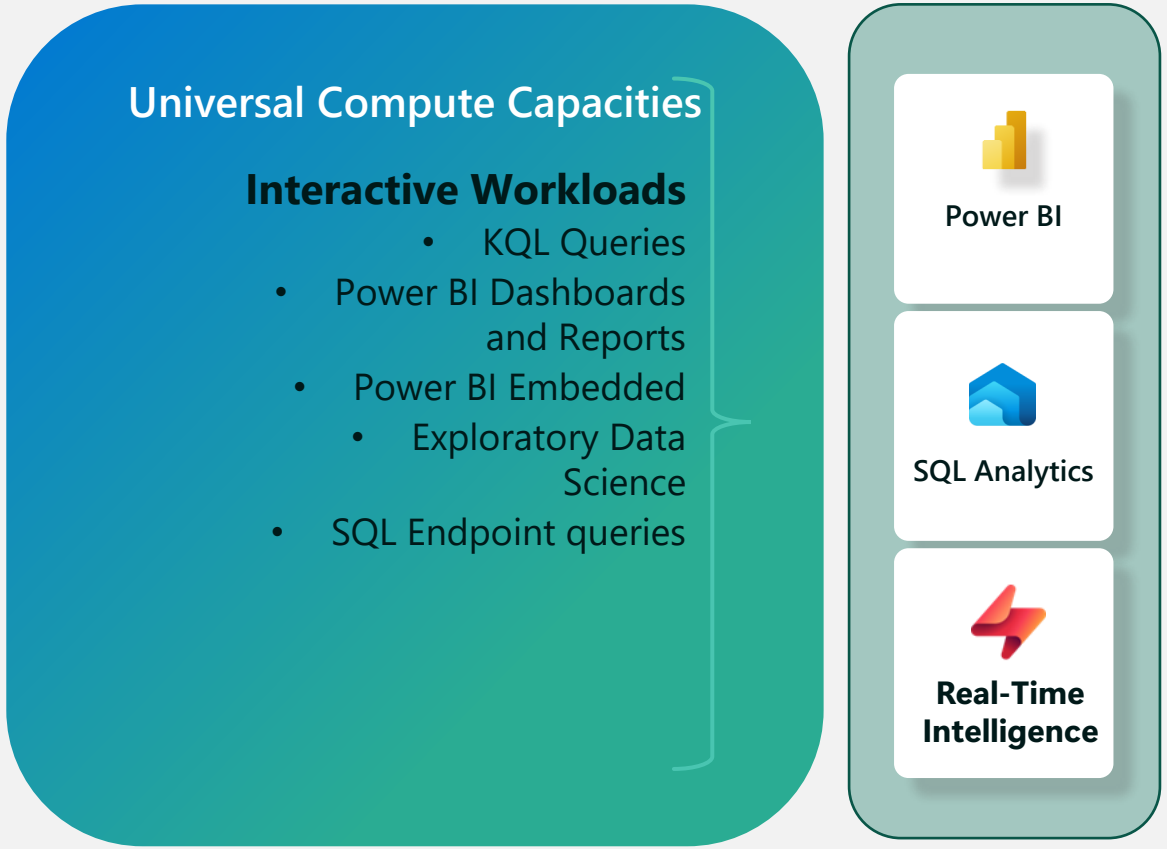
Smoothing/Bursting:

- Background operations deducted over up to 24h from execution
- Burstable workloads can scale many times the deployed capacity.
- Some workloads have guardrails limiting bursting to 12x of the deployed capacity

Microsoft Fabric

Smoothing/Bursting:

- Interactive operations deducted over 5minutes from execution
- Burstable workloads can scale many times the deployed capacity.

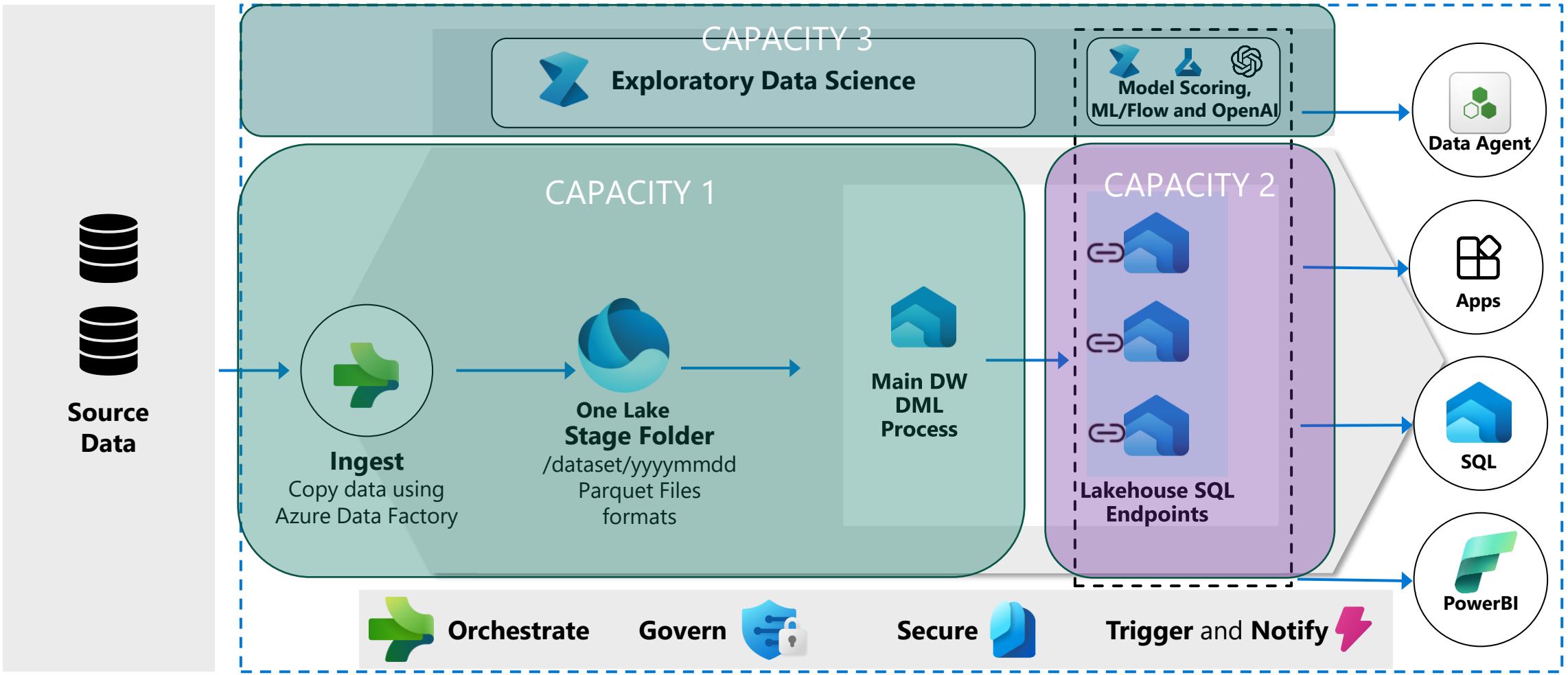


Microsoft Fabric

Smoothing/Bursting:

- All Workloads in a shared capacity impact each other
- Interactive , Batch and Realtime Analytics workloads should be isolated.
- Sandpits and Rouge Data engineering workloads could “Sink” all interactive workloads.

Modern Data Warehouse Pattern in Microsoft Fabric




Microsoft
APAC
F 64
Data Factory
Analytics
Fabric Databases
Real-time Intelligence
Power BI
Workspace 1
Workspace 2
EMEA
F 128
Data Factory
Analytics
Fabric Databases
Real-time Intelligence
Power BI
Workspace 1
Workspace 2
Microsoft Purview
OneLake
Group IT
F 256
Data Factory
Analytics
Fabric Databases
Real-time Intelligence
Power BI
Workspace 1
Workspace 2


Microsoft
Dev
F 64
Data Factory
Analytics
Fabric Databases
Real-time Intelligence
Power BI
Workspace 1
Workspace 2
Test
F 128
Data Factory
Analytics
Fabric Databases
Real-time Intelligence
Power BI
Workspace 1
Workspace 2
Microsoft Purview
OneLake
Production
F 256
Data Factory
Analytics
Fabric Databases
Real-time Intelligence
Power BI
Workspace 1
Workspace 2

Microsoft

Key Take Aways for Capacity

- Workloads Should be isolated based on Type (Batch, Interactive, RTA)
- Capacity can be split by domain
- Capacity design is a balance between capability, cost, workload isolation, and practicality

Microsoft

Thank You