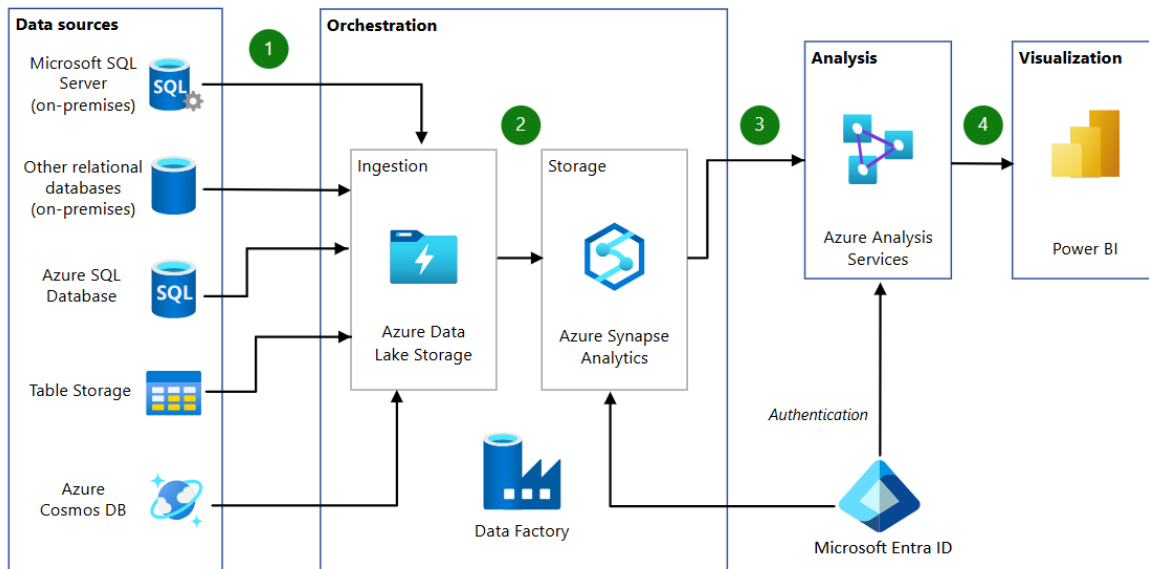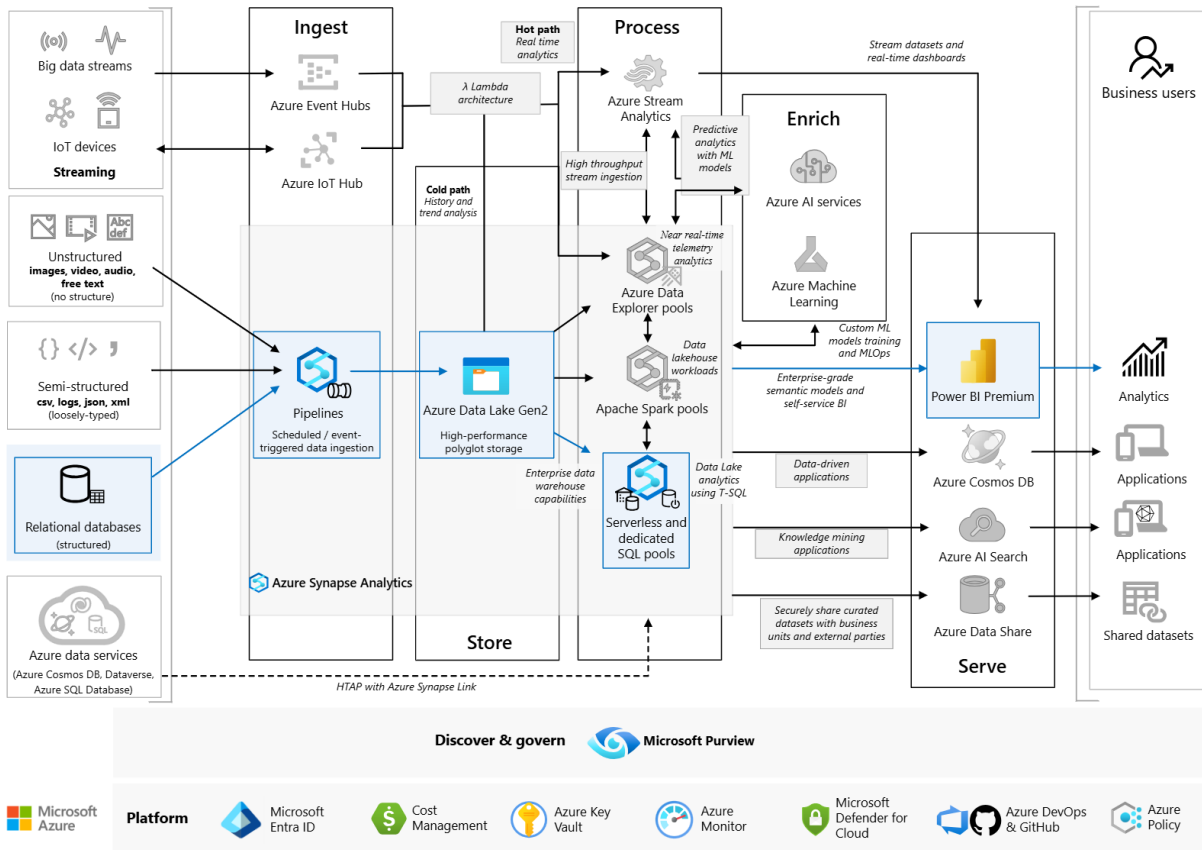# End-to-End Analytics Platform Patterns on Azure

- **Azure Synapse patterns for analytics**
  - Data Pipeline Design for a Unified Analytics Platform on Azure
  - Enterprise BI Solution Design on Azure
  - Near Realtime Processing for Syncing Lakehouse Data with OLTP
- **Azure Databricks Patterns**
  - Stream processing with Azure Databricks
- **Azure Fabric Patterns**
  - Modern Data Warehouse with MS Fabric
- **HA/DR for Analytics**
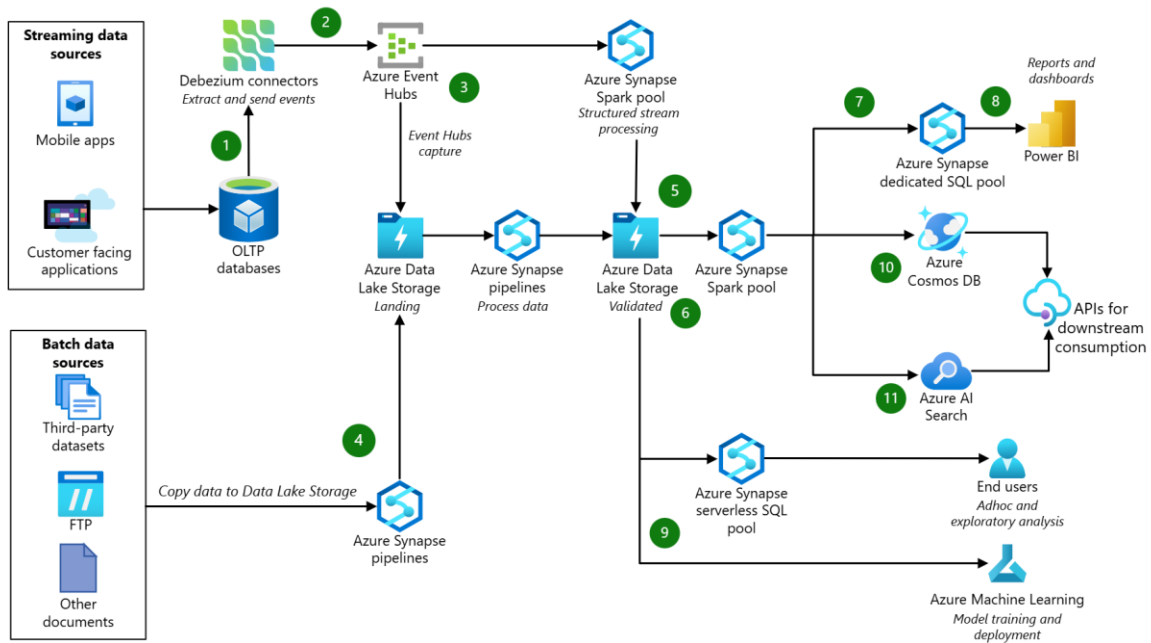 - HA/DR Design for Azure DataFactory and Azure Synapse

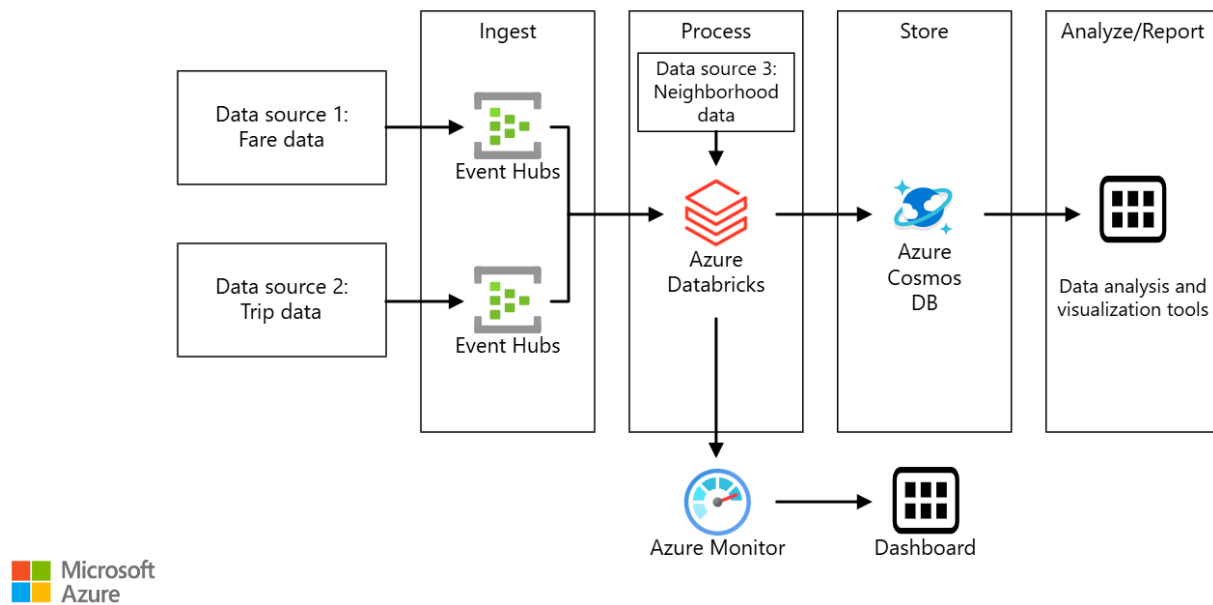## Data Pipeline Design For Building a Unified Analytics Platform on Azure

Enterprise BI Solution Design on Azure

Ingest
- Azure Event Hubs
- Azure IoT Hub

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)

λ Lambda architecture

Hot path
Real time analytics

Cold path
History and trend analysis

Process
- Azure Stream Analytics
- Azure Data Explorer pools
- Apache Spark pools
- Serverless and dedicated SQL pools

High throughput stream ingestion

Near real-time telemetry analytics

Data lakehouse workloads

Data Lake analytics using T-SQL

Enrich
- Azure AI services
- Azure Machine Learning

Predictive analytics with ML models

Custom ML models training and MLOps

Pipelines
Scheduled / event-triggered data ingestion

Azure Data Lake Gen2
High-performance polyglot storage

Enterprise data warehouse capabilities

Store

Azure Synapse Analytics

HTAP with Azure Synapse Link

Stream datasets and real-time dashboards

Enterprise-grade semantic models and self-service BI

Power BI Premium

Data-driven applications

Knowledge mining applications

Securely share curated datasets with business units and external parties

Serve
- Azure Cosmos DB
- Azure AI Search
- Azure Data Share

Business users

Analytics

Applications

Applications

Shared datasets

Discover & govern — Microsoft Purview

Microsoft Azure | **Platform** | Microsoft Entra ID | Cost Management | Azure Key Vault | Azure Monitor | Microsoft Defender for Cloud | Azure DevOps & GitHub | Azure Policy

Near Realtime Processing for Syncing Lakehouse Data with OLTP

Streaming data sources
- Mobile apps
- Customer facing applications

Batch data sources
- Third-party datasets
- FTP
- Other documents

Debezium connectors
Extract and send events

Azure Event Hubs

Event Hubs capture

OLTP databases

Azure Data Lake Storage
Landing

Copy data to Data Lake Storage

Azure Synapse pipelines

Azure Synapse pipelines
Process data

Azure Data Lake Storage
Validated

Azure Synapse Spark pool
Structured stream processing

Azure Synapse Spark pool

Azure Synapse dedicated SQL pool

Reports and dashboards

Power BI

Azure Cosmos DB

Azure AI Search

APIs for downstream consumption

Azure Synapse serverless SQL pool

End users
Adhoc and exploratory analysis

Azure Machine Learning
Model training and deployment

Stream Processing With Azure Databricks

Ingest

Data source 1: Fare data → Event Hubs

Data source 2: Trip data → Event Hubs

Process

Data source 3: Neighborhood data → Azure Databricks

Store

Azure Cosmos DB

Analyze/Report

Data analysis and visualization tools

Azure Monitor → Dashboard

Microsoft Azure

Modern Data Warehouse with Microsoft Fabric

Stream
3 → Event Hubs

Microsoft Dynamics 365
2

Non-structured data

Semi-structured data

Relational databases

Load and ingest

ADLS
4

Data Factory Pipeline

1 → SQL Database

Store

Fabric capacity or Premium capacity or Premium Per User

Real-Time Analytics

Serverless analysis

OneLake

SQL Analytics endpoint
5

Process/manipulate

Spark pool

Pipelines

Collaborate/consume

Consume/serve

Power Apps

Microsoft Dynamics 365

Microsoft Dynamics CRM

Power BI

Functions apps

Logic Apps

Web apps

Microsoft Azure

Disaster Recovery Design for Azure DataFactory and Azure Synapse