

Data Architectures in Azure for Analytics & Big Data

October 20, 2018



Melissa Coates

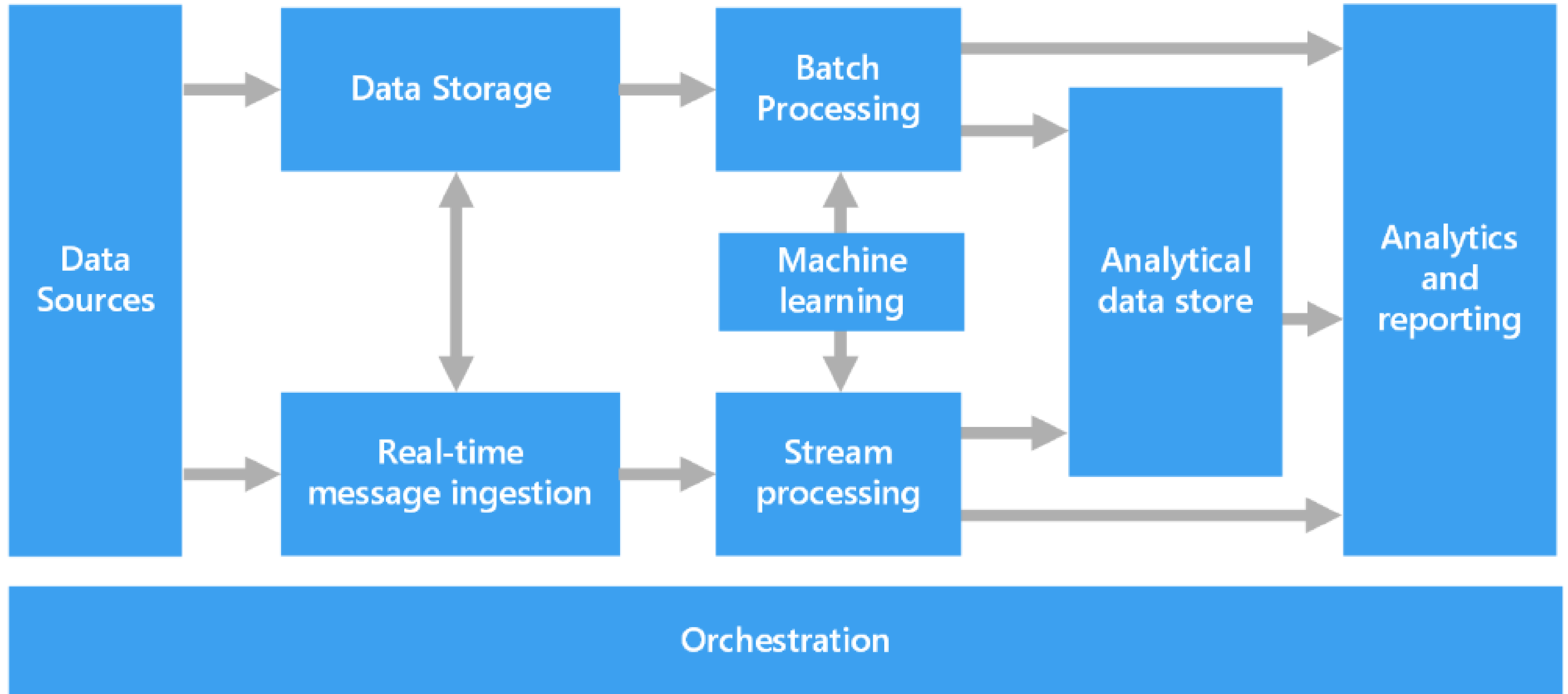
Solution Architect, BlueGranite
Microsoft Data Platform MVP

Blog: www.sqlchick.com
Twitter: @sqlchick

Data Architecture

A set of
rules, policies, standards, & models
that govern and define
the type of *data collected*
& how it is
used, stored, managed, & integrated
within an organization
& its database systems

Data Architecture Components











Azure









A public cloud computing platform and infrastructure for building, deploying, and managing MSFT-specific and third party software and services through a global network of Microsoft-managed datacenters

Source: <http://azureplatform.azurewebsites.net/>






Compute

 Virtual Machines	 Virtual Machine Scale Sets
 Azure Container Service	 Azure Container Registry
 Functions	 Batch
 Service Fabric	 Cloud Services








Networking

 Virtual Network	 Load Balancer
 Application Gateway	 VPN Gateway
 Azure DNS	 Traffic Manager
 ExpressRoute	 Network Watcher








Storage

 Storage: Blobs, Tables, Queues, Files, Disks	 Data Lake Store
 StorSimple	 Azure Backup
 Site Recovery	







Monitoring & Management

 Azure Portal	 Azure Resource Manager	 Azure Advisor	 Azure Monitor	 Log Analytics	 Automation	 Scheduler
--	--	---	---	---	--	---








Web & Mobile

 Web Apps	 Mobile Apps
 Logic Apps	 API Apps
 Content Delivery Network	 Media Services
 Search	








Databases

 SQL Database	 SQL Data Warehouse
 SQL Server Stretch Database	 DocumentDB
 Redis Cache	 Data Factory







Intelligence & Analytics

 HDInsight	 Machine Learning
 Cognitive Services	 Azure Bot Service*
 Data Lake Analytics	 Power BI Embedded
 Azure Analysis Services	








Internet of Things & Enterprise Integration

 Azure IoT Hub	 Event Hubs
 Stream Analytics	 Notification Hubs
 BizTalk Services	 Service Bus
 Data Catalog	

Security + Identity

 Security Center	 Key Vault
 Azure Active Directory	 B2C
 Domain Services	 Multi-Factor Authentication

Developer Services

 Visual Studio Team Services	 Azure DevTest Labs
 VS Application Insights	 API Management
 HockeyApp	 Developer Tools
 Service Profiler*	

Azure Technologies

Data Storage: Relational Databases

IaaS

Infrastructure as a Service

PaaS

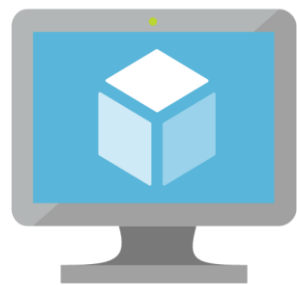
Platform as a Service

SMP

Symmetric Multi-Processing

MPP

Massively Parallel Processing



Relational database of your choice in a virtual machine



MPP

Massively Parallel Processing

TERADATA



Managed Database

Managed Instance

Azure SQL Database



Azure Database for MySQL



Azure Database for PostgreSQL



Azure Database for MariaDB



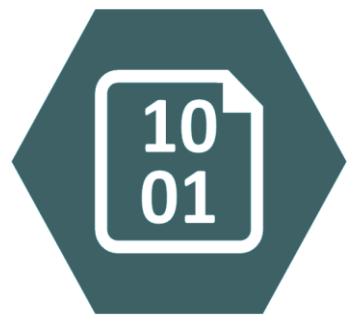
Azure SQL Data Warehouse

Data Storage: Big Data

PaaS

Platform as a Service

Object Storage (Flat)



Azure Blob
Storage

Hierarchical Storage



Azure Data Lake
Store (Gen1)

Multi-Modal



Azure Data Lake
Storage (Gen2)
(PREVIEW)

Data Storage: NoSQL

PaaS
Platform as a Service

Multi-Model



HDInsight
HBase
Cluster



Azure
Table Storage



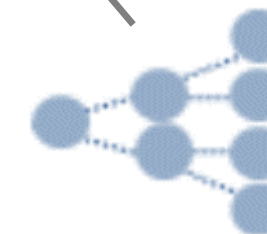
Azure
CosmosDB



Key
Value



Column
Family



JSON
Documents

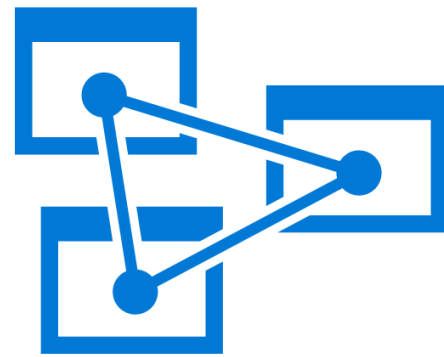


Graph

Data Storage: Analytical & OLAP

IaaS

Infrastructure as a Service

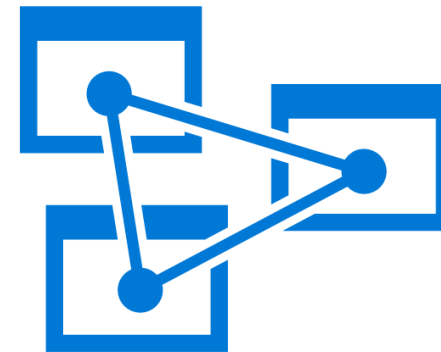


SQL Server
Analysis Services



PaaS

Platform as a Service



Azure
Analysis Services

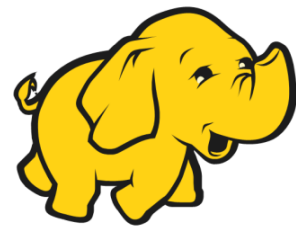


Power BI

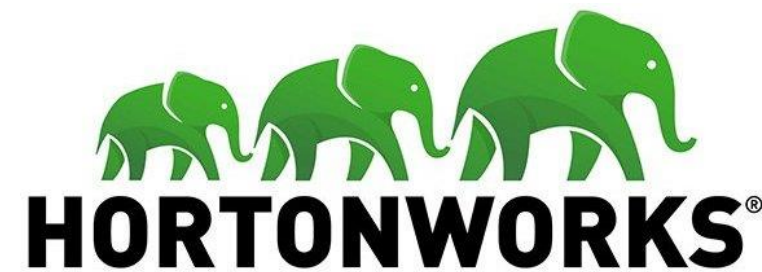
Compute: Big Data

IaaS

Infrastructure as a Service



HDInsight
in a VM



cloudera



PaaS

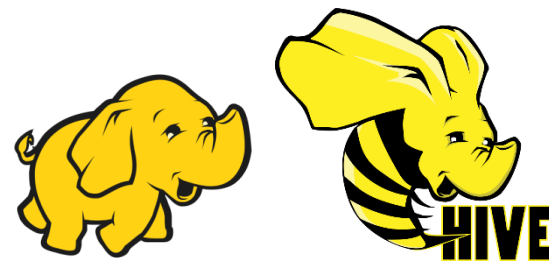
Platform as a Service



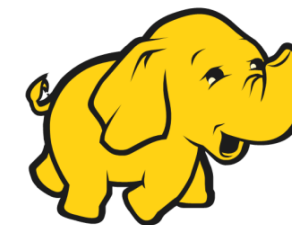
HDInsight
Spark Cluster



Azure
Databricks



HDInsight
Interactive
Query Cluster
(Hive LLAP)



HDInsight
Hadoop Cluster

SaaS

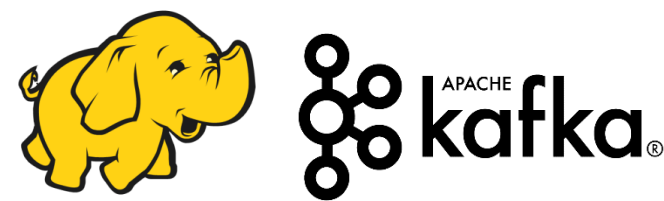
Software as a Service



Azure Data Lake
Analytics

Compute: Streaming & Event Processing

PaaS
Platform as a Service



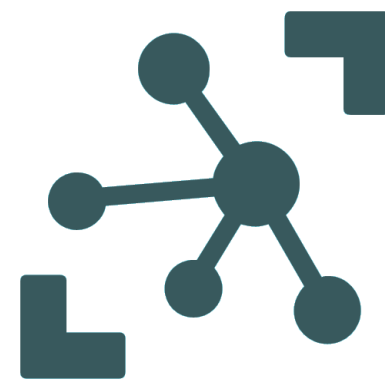
HDInsight
Kafka Cluster



HDInsight
Storm Cluster



HDInsight
Spark Streaming
Cluster



Azure
IoT Hub



Azure
Event Hub



Azure
Stream
Analytics

Compute: Data Integration

IaaS

Infrastructure as a Service



Tool of your choice
in a virtual machine



ORACLE®



Informatica™



PaaS

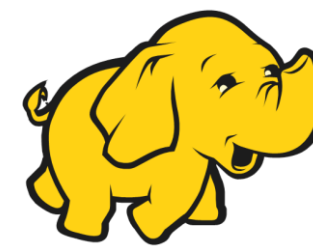
Platform as a Service



Azure Data
Factory



Azure
Databricks

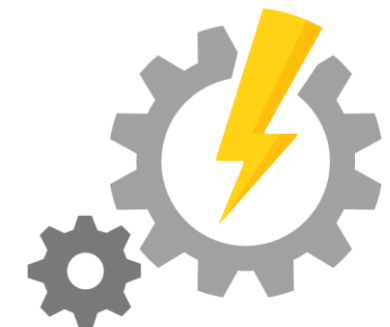


HDInsight:
Spark, Hive, Pig,
Scoop, Oozie

Serverless



Azure
Functions



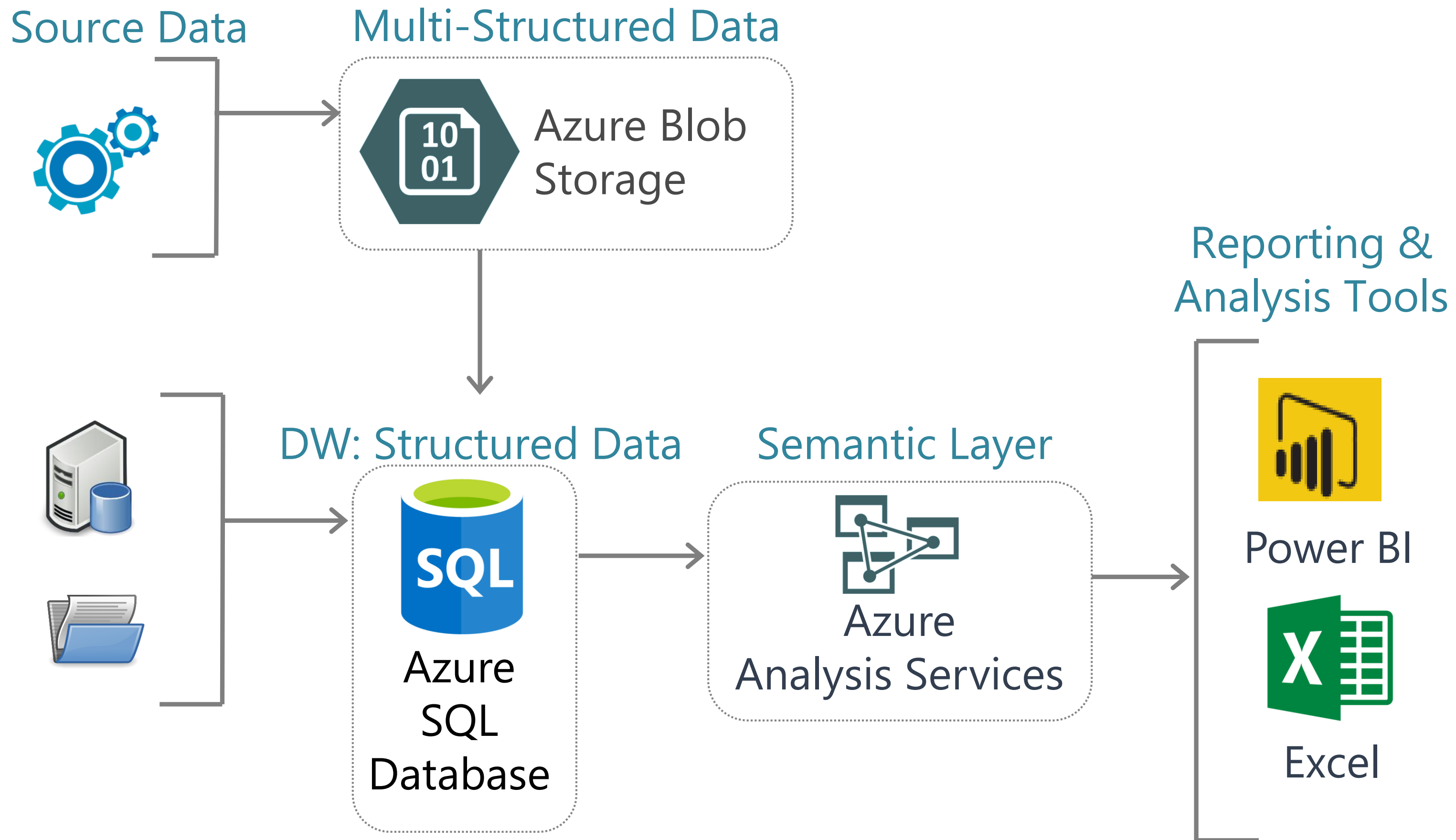
Azure
Automation

Reference Architectures

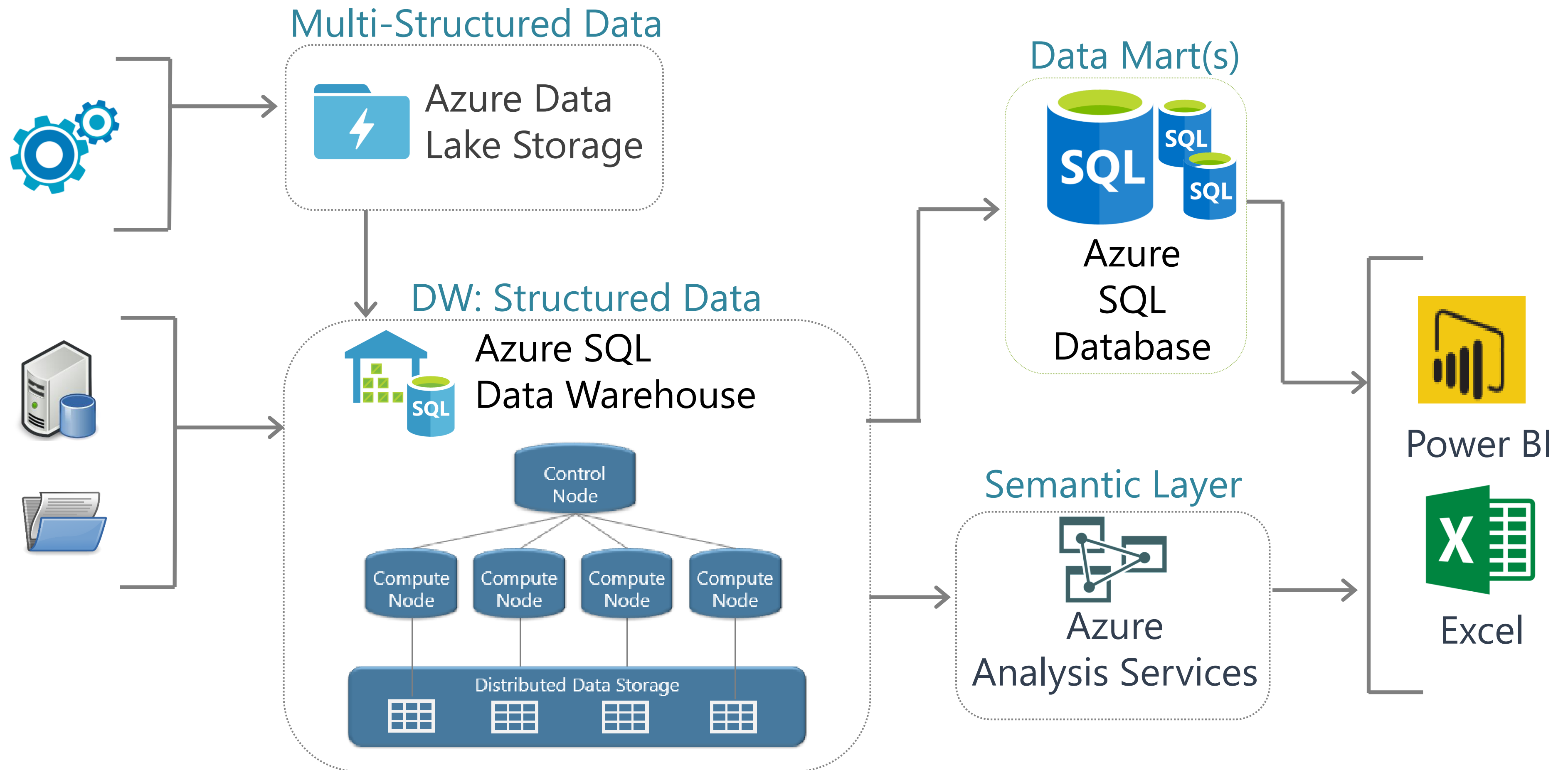
Following are examples only!

There are many variations & opportunities to
exchange one Azure service for another.

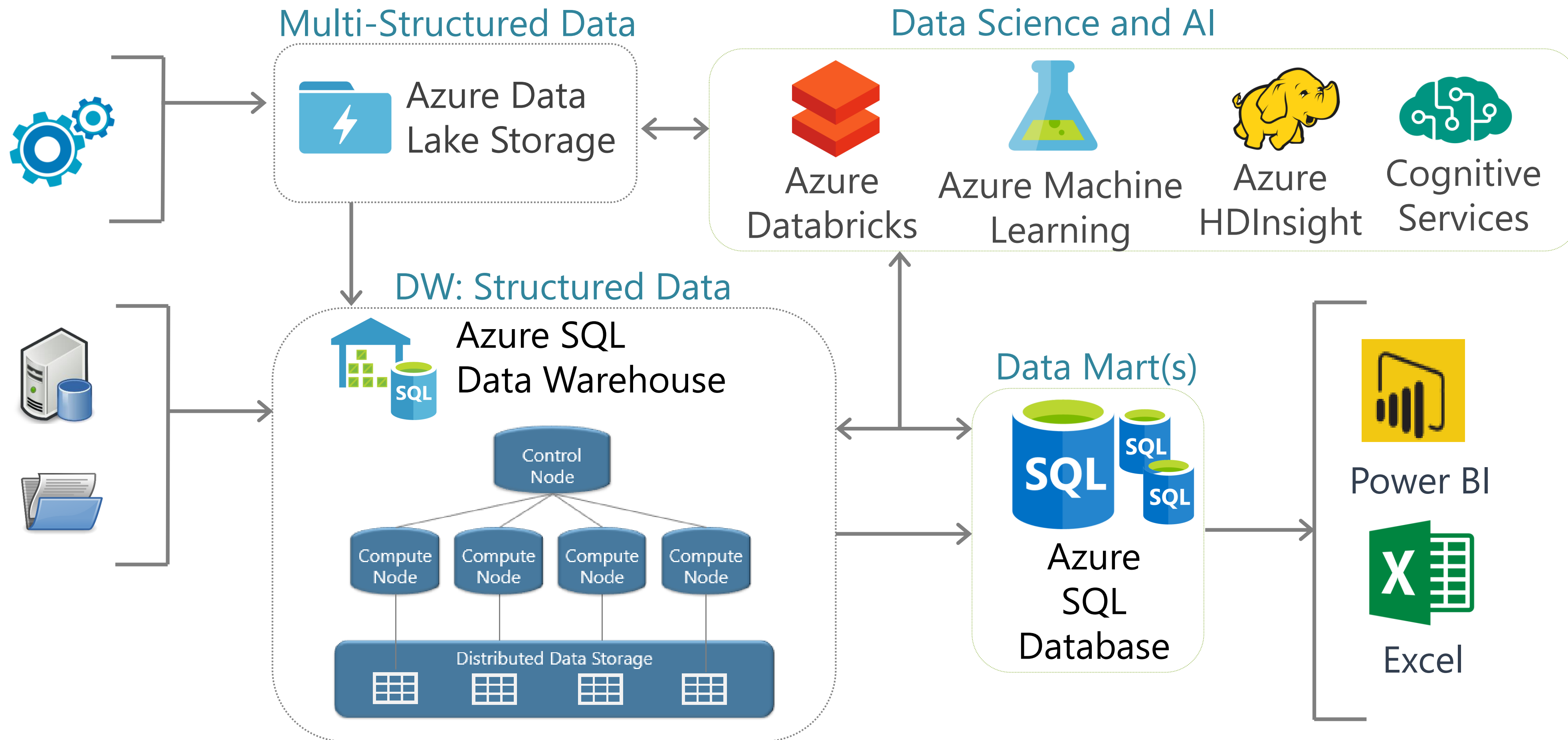
Small/Medium Data Warehousing



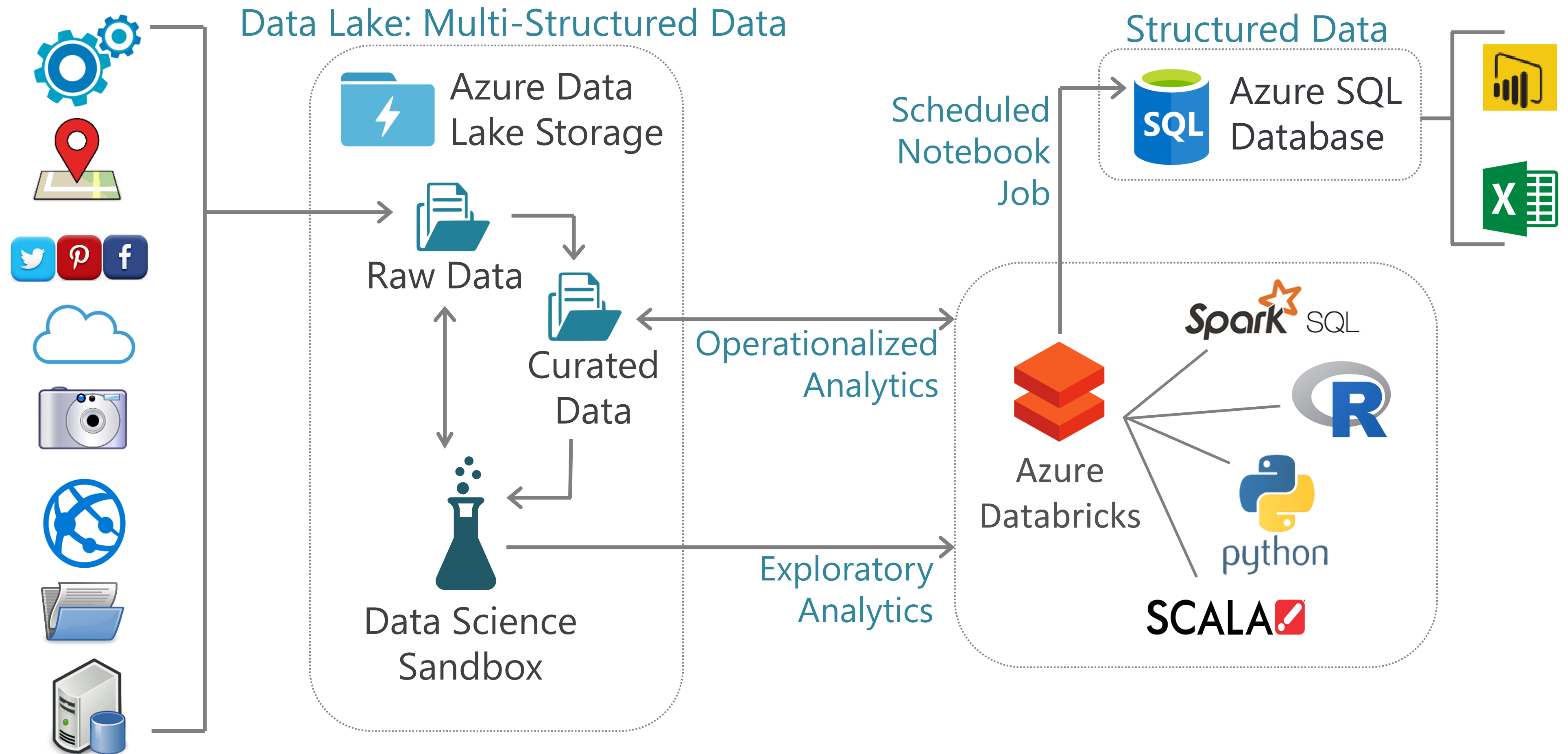
Enterprise Data Warehousing and BI



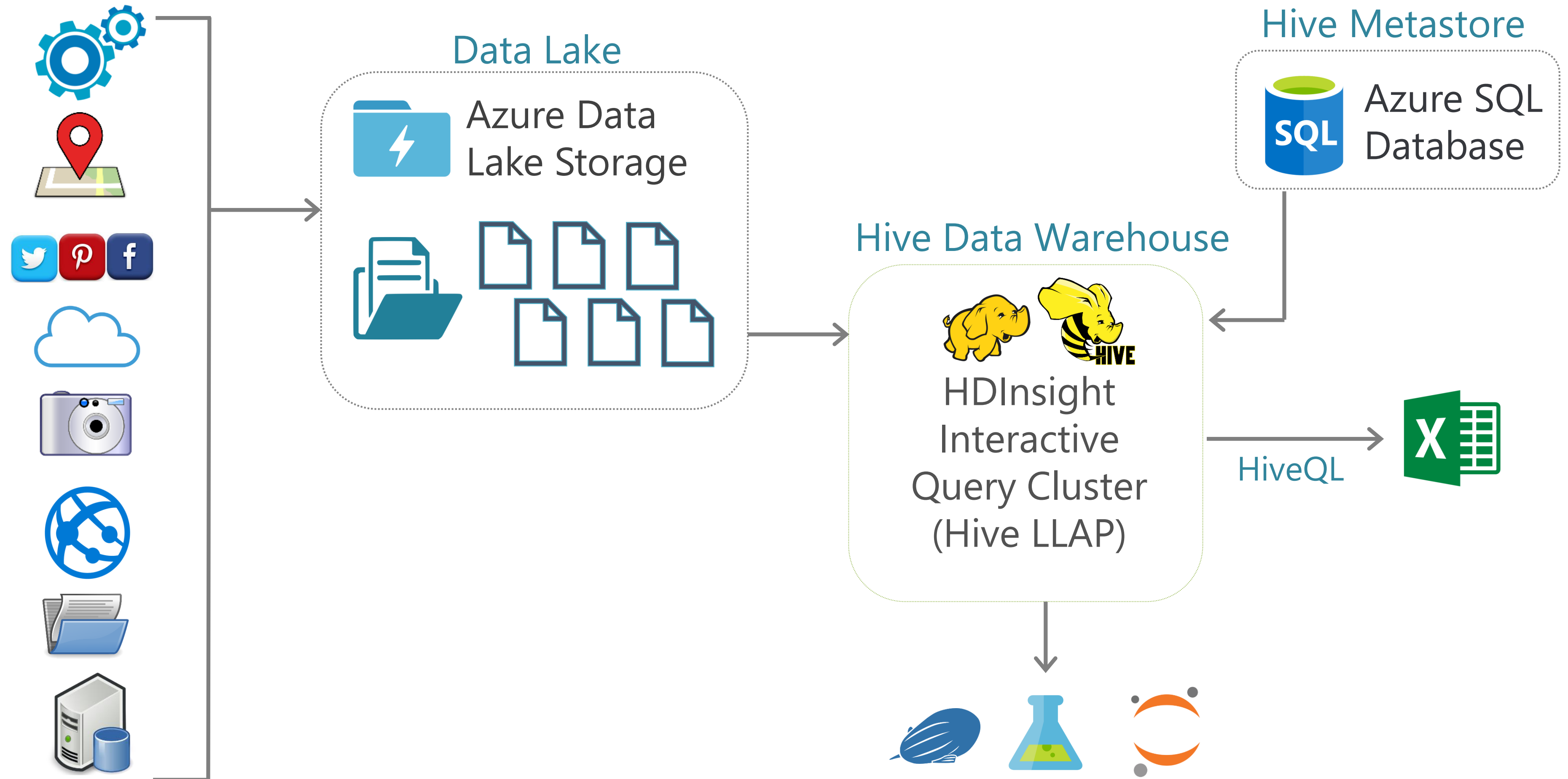
Data Science and Artificial Intelligence



Unified Data Science & Data Engineering

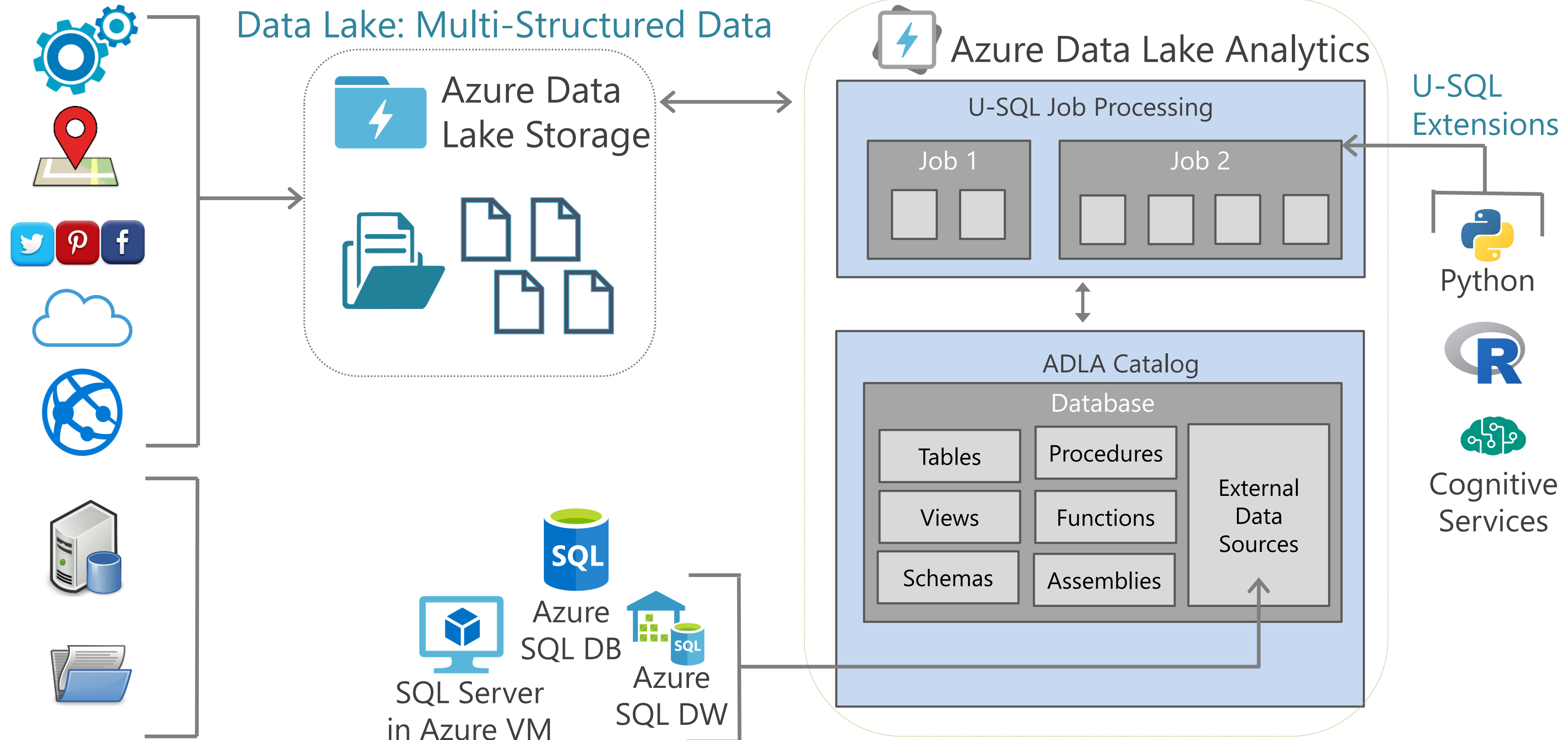


Big Data Interactive Querying (SQL on Hadoop)

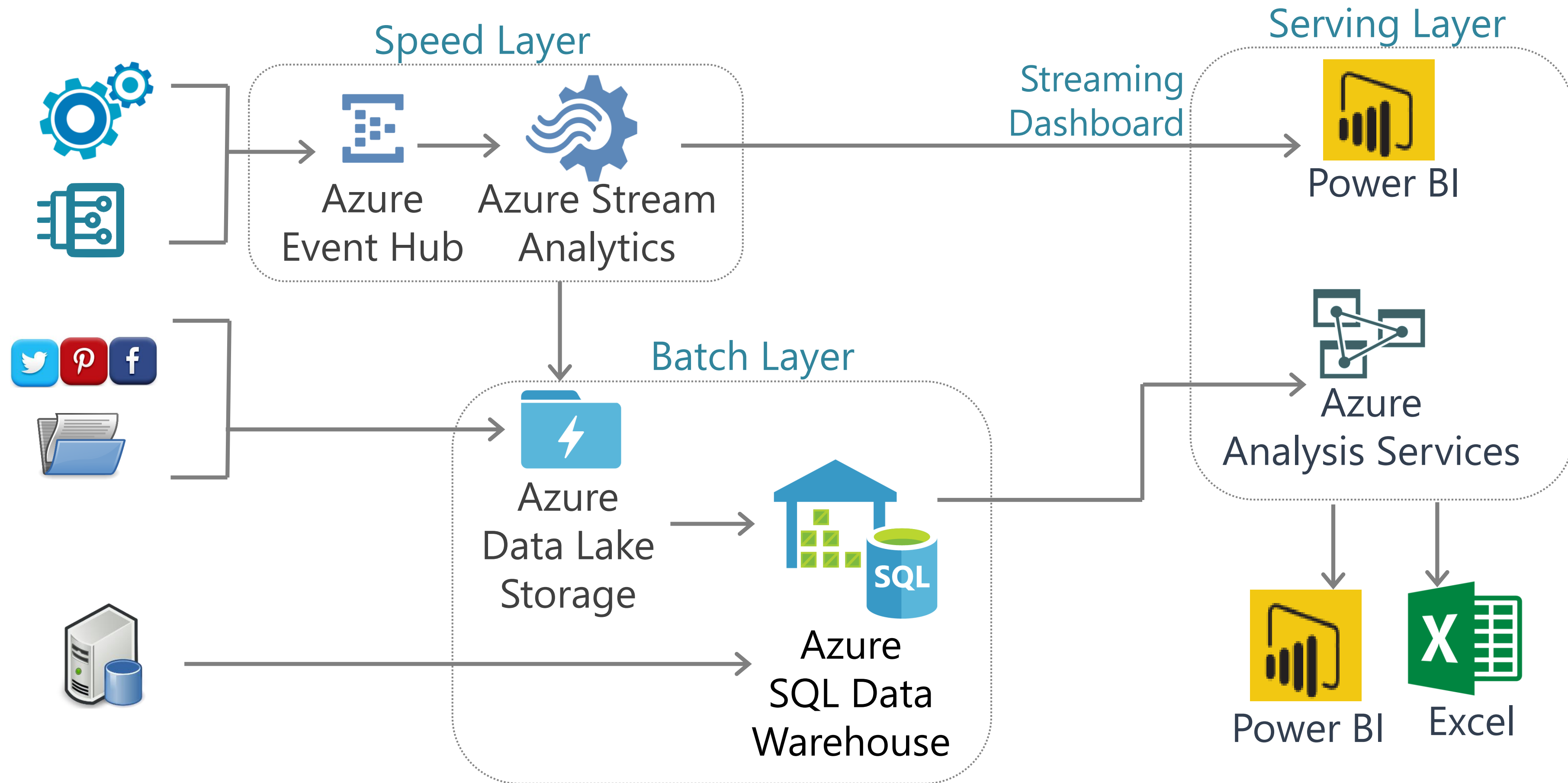


Big Data Batch Processing

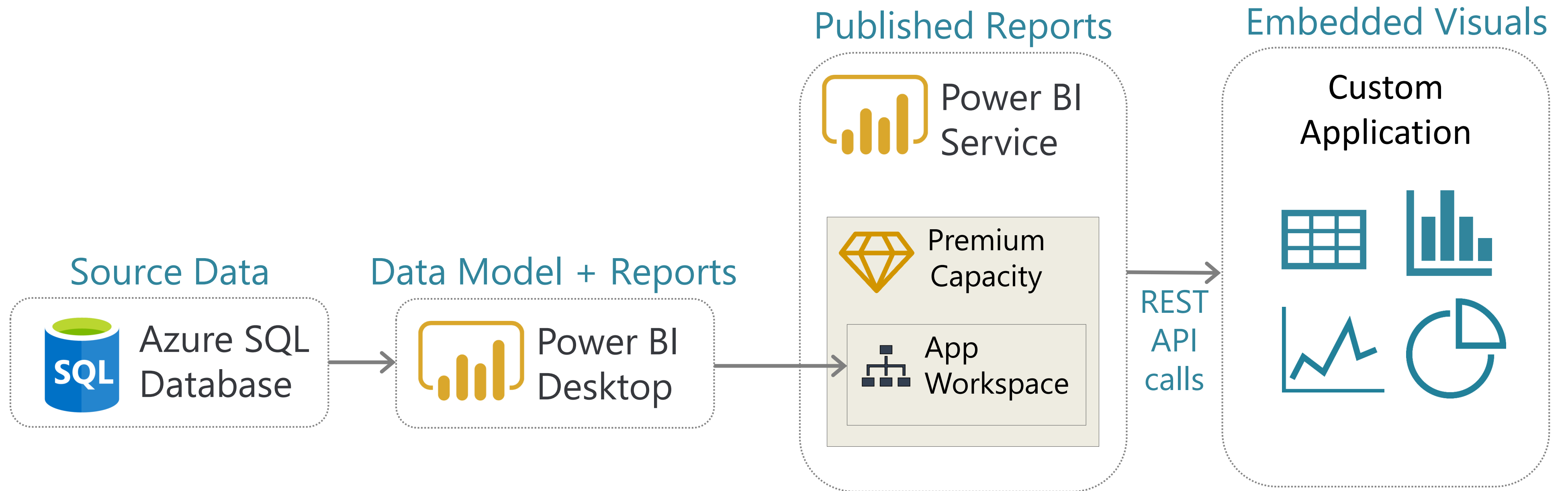
Big Data Job Processing



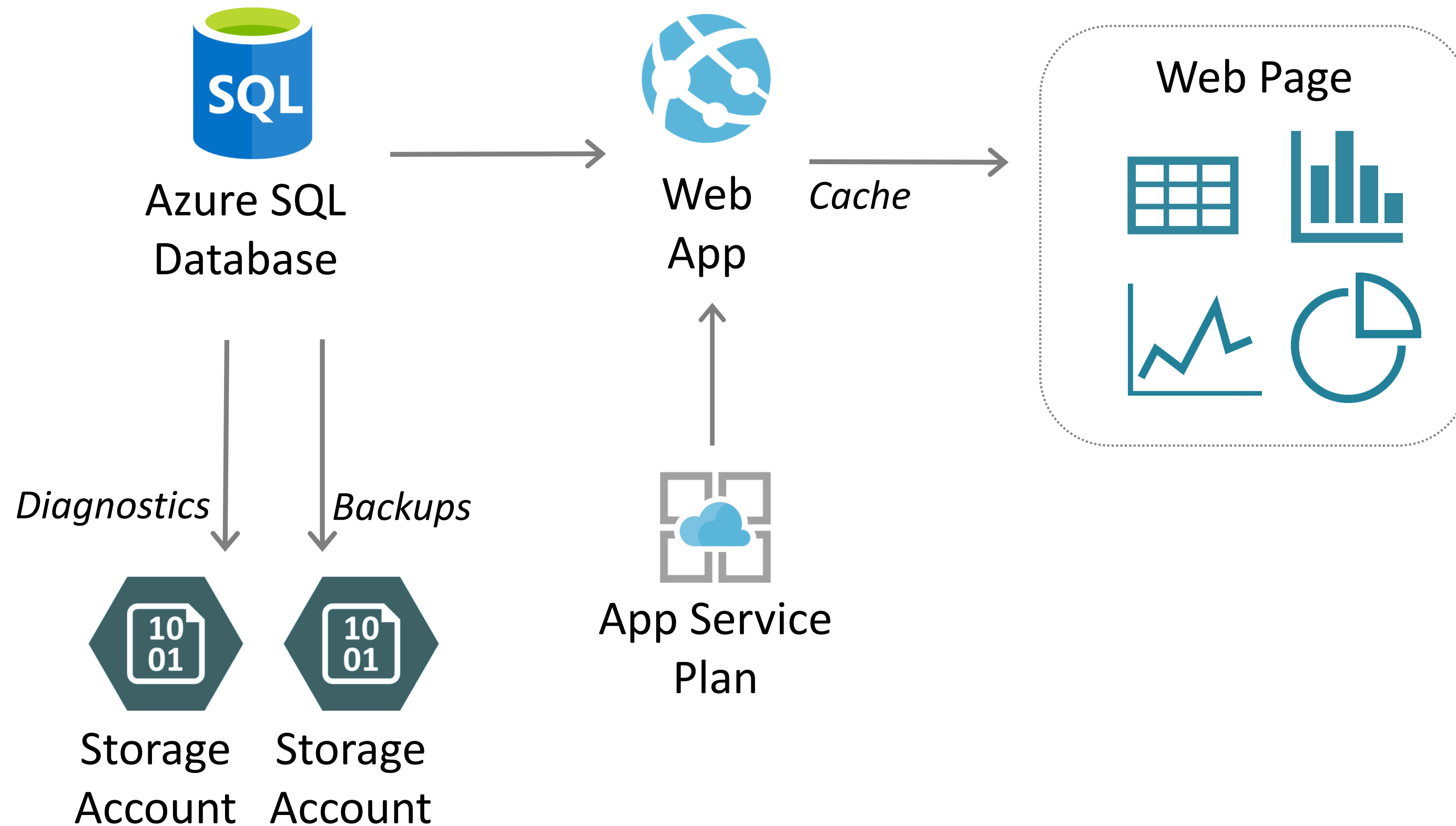
IoT + Batch Data (Lambda Architecture)



Operational BI (Embedded BI)



Web Application

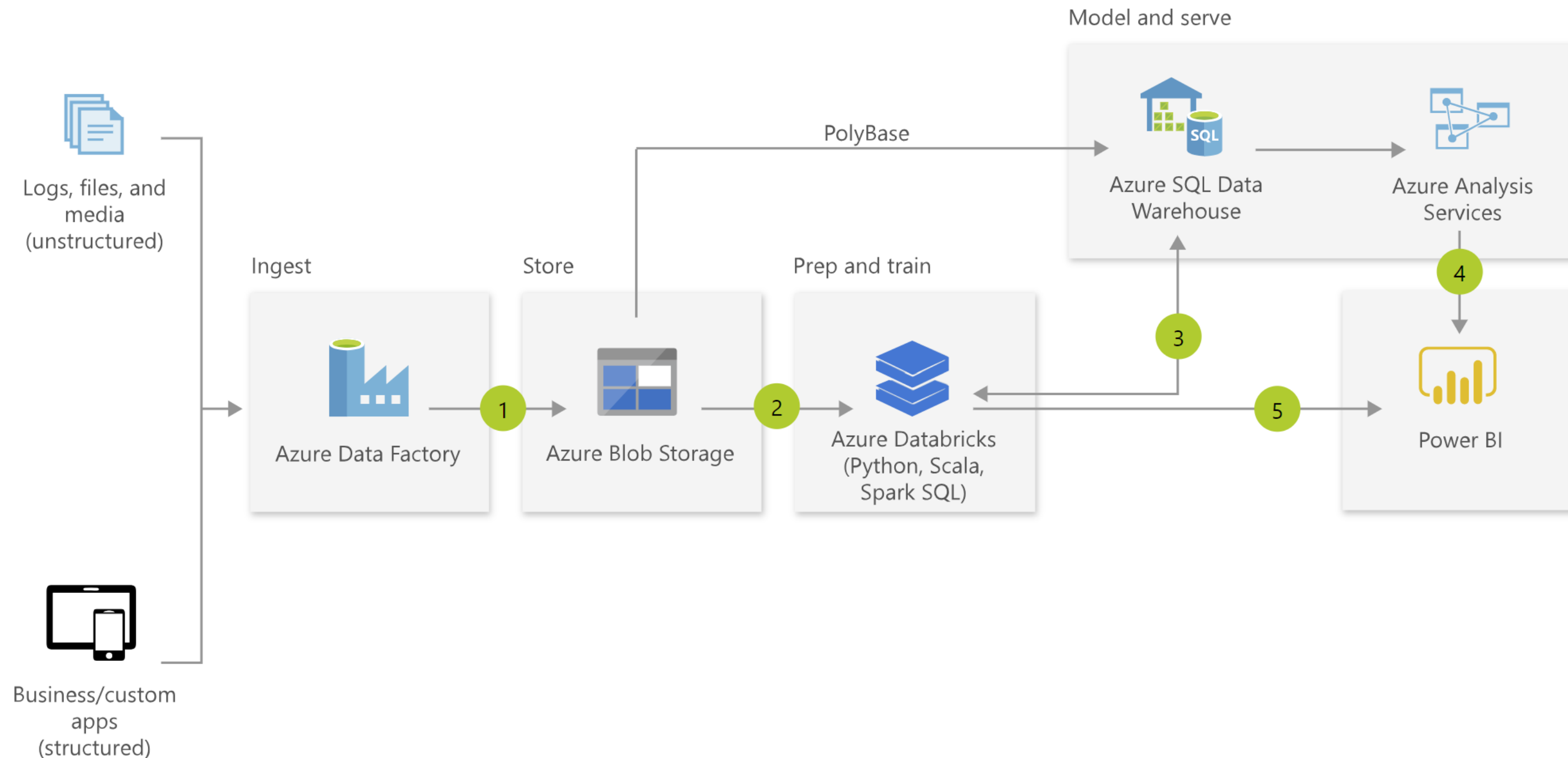


Wrap-Up

More Info

Azure Solution Architectures:

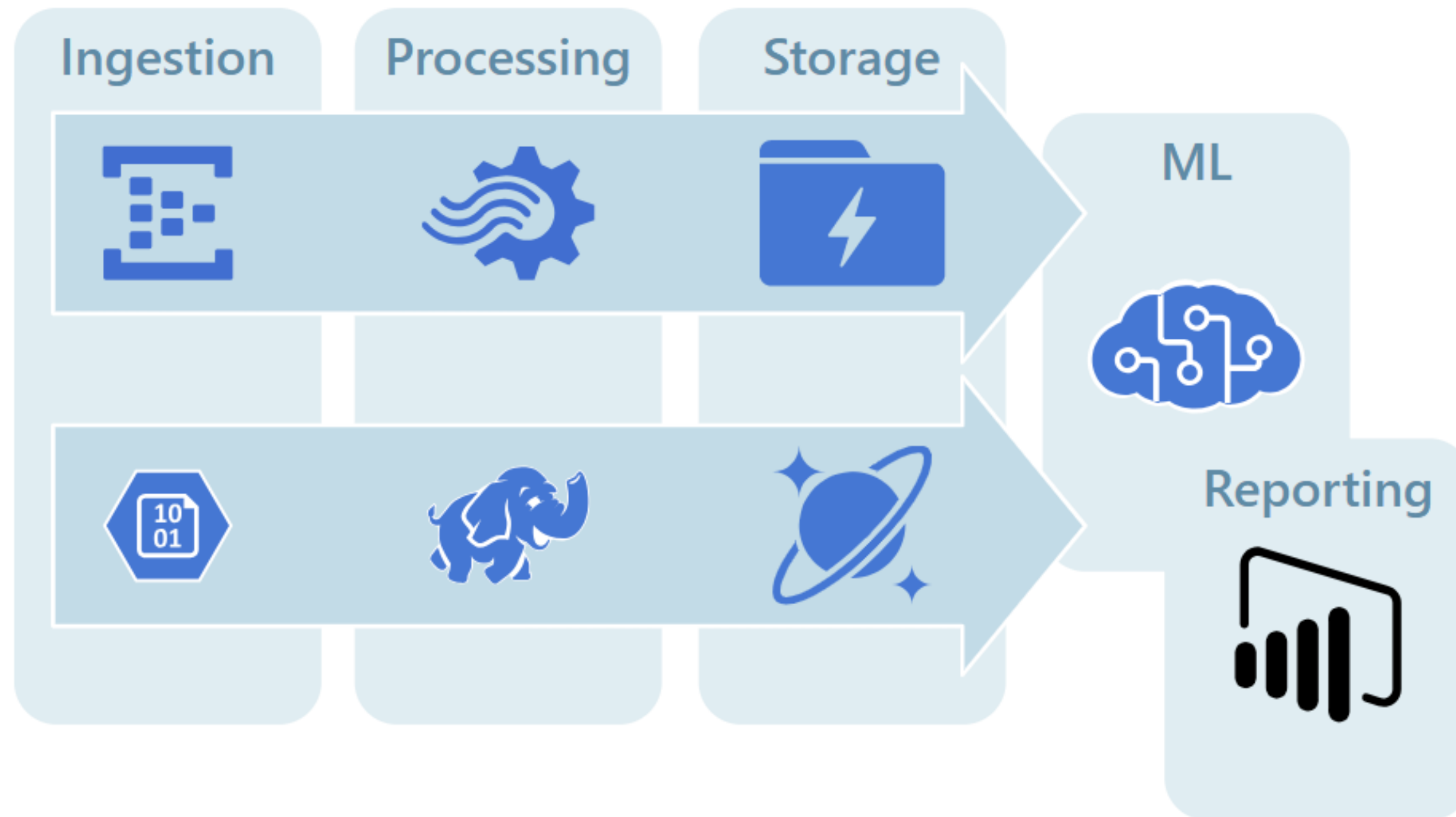
<https://azure.microsoft.com/en-us/solutions/architecture/>



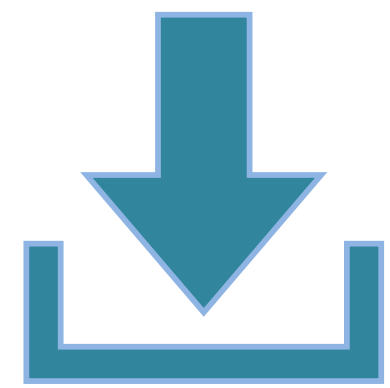
More Info

Azure Data Architecture Guide:

<https://docs.microsoft.com/en-us/azure/architecture/data-guide/>



Thanks!



Download latest version of slides:

[SQLChick.com > Presentations & Downloads page](#)



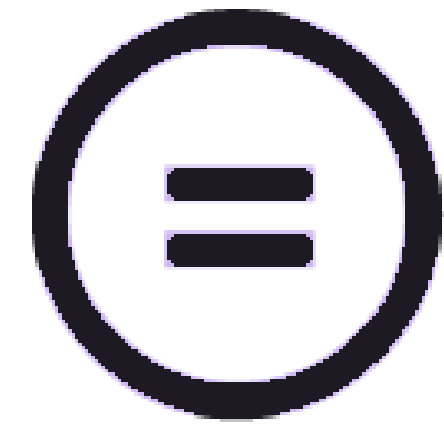
[Creative Commons](#)
[License 3.0](#)



*Attribute to me as original
author if you
share this material*



*No usage of this
material for
commercial purposes*



*No derivatives or
changes to this material*