

#AzConfDev



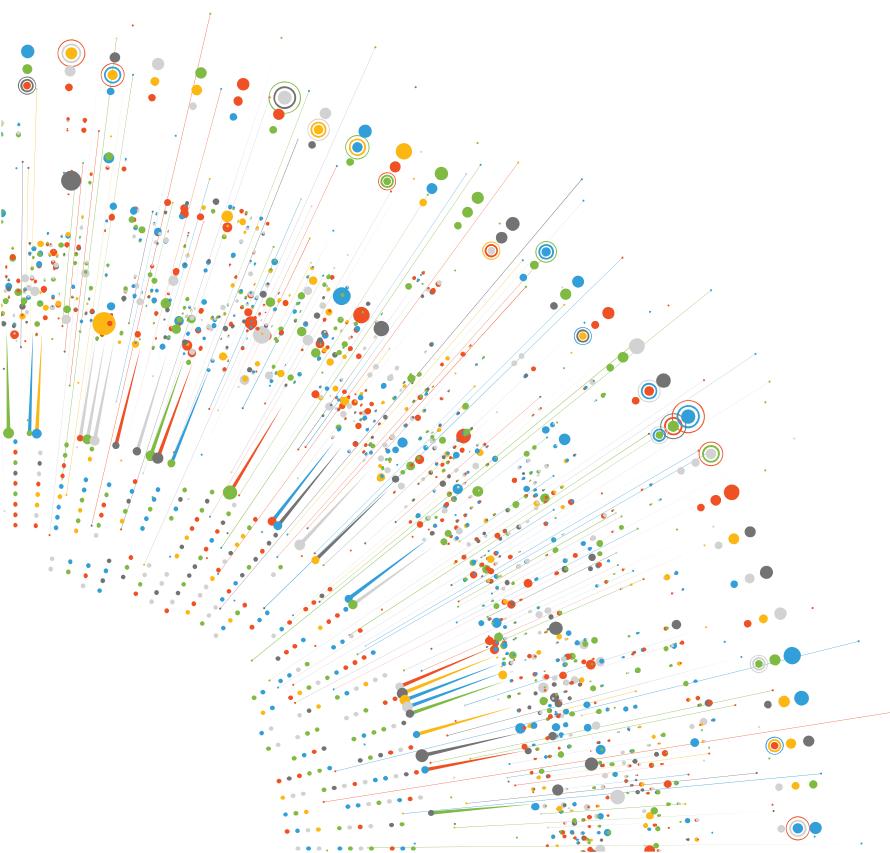
# Azure COMMUNITY CONFERENCE

Sponsored  
by  
 Microsoft



# Introduction to Azure Data Platform

---



#AzConfDev



# JAVIER VILLEGRAS

---

IT Director, Mediterranean Shipping Company



javier.ignacio.villegas@gmail.com



javier\_vill



javiervillegas



sql-javier-villegas.blogspot.com.ar

# Javier Villegas

IT Director – Data and BI at Mediterranean Shipping Company

Involved with the Microsoft SQL Server since early versions

Specialization in SQL Server Administration, Performance Tuning and High Availability

## Microsoft MVP Data Platform

MCP and MCT

## Technical Speaker

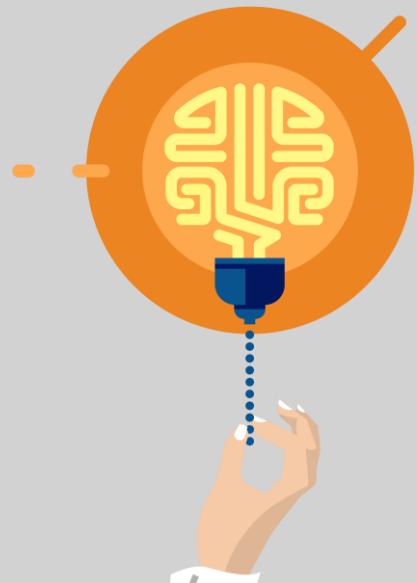
SQL PASS, 24 HOP, SQL Saturdays , PASS Marathon and PASS Virtual Groups,  
vOpen, Microsoft AI+ Tour, GroupBy and DataPlatformGeeks



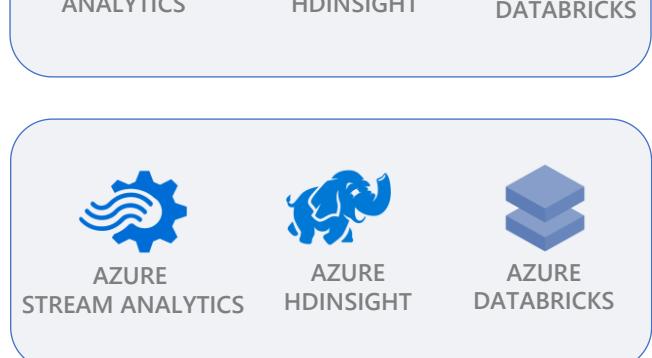
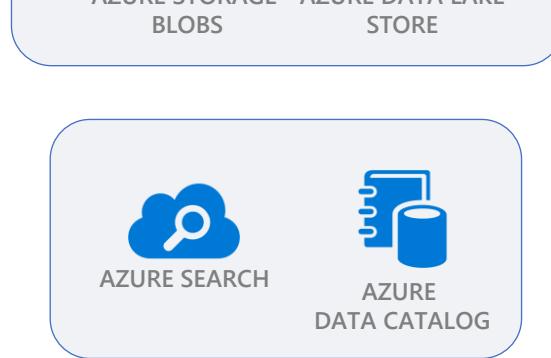
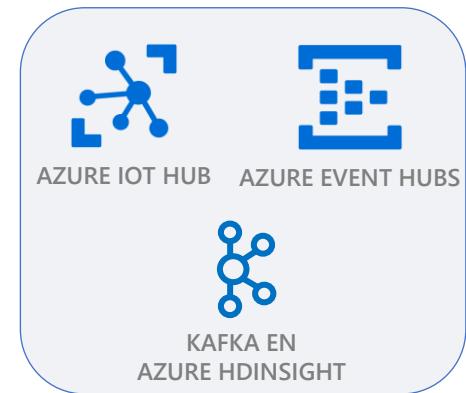
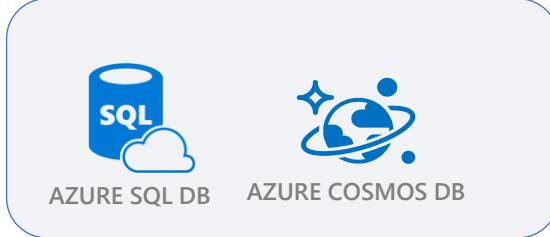


## Azure Data Platform

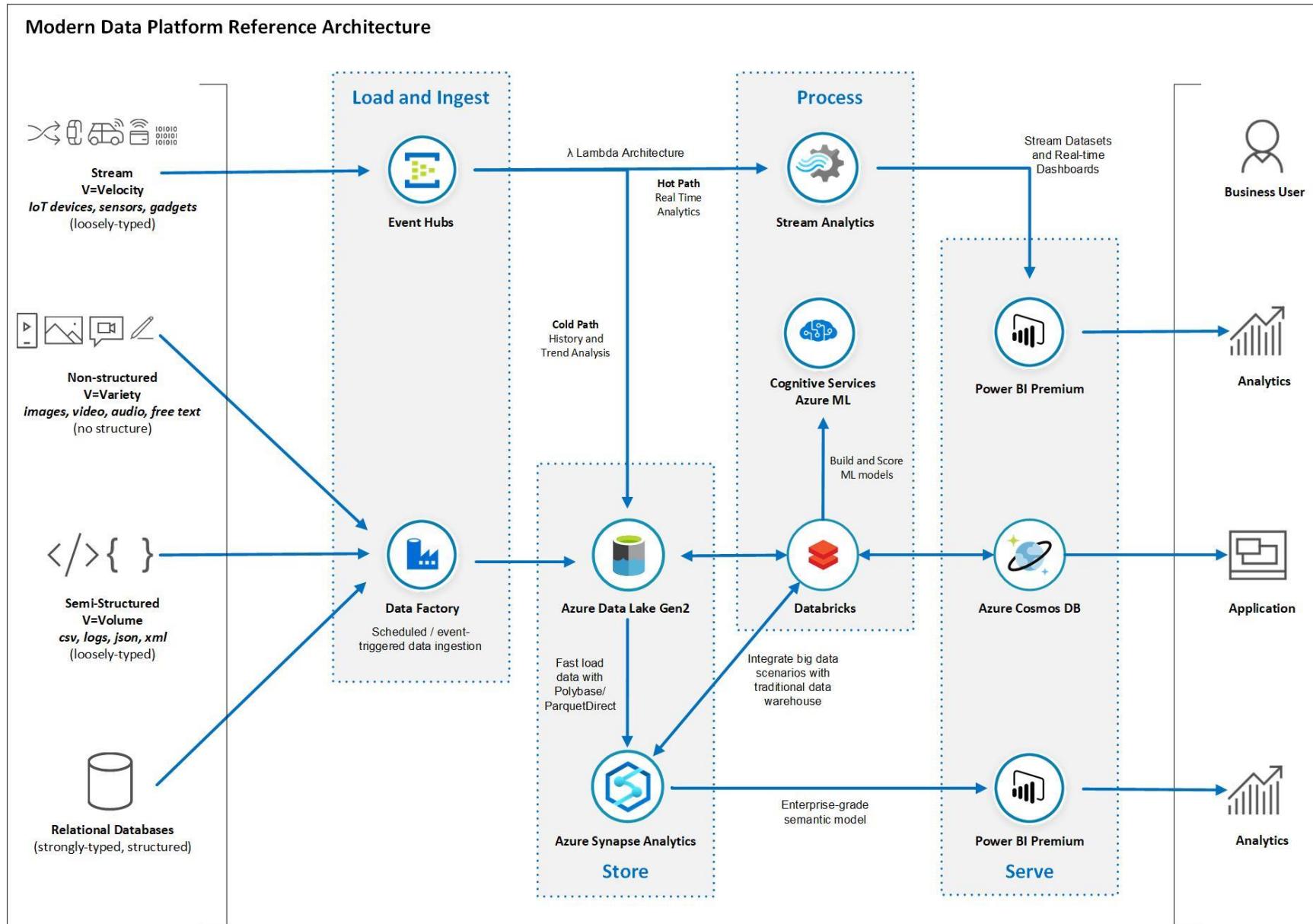
#AzConfDev



# Azure Data Platform



# Azure Data Platform - Architecture





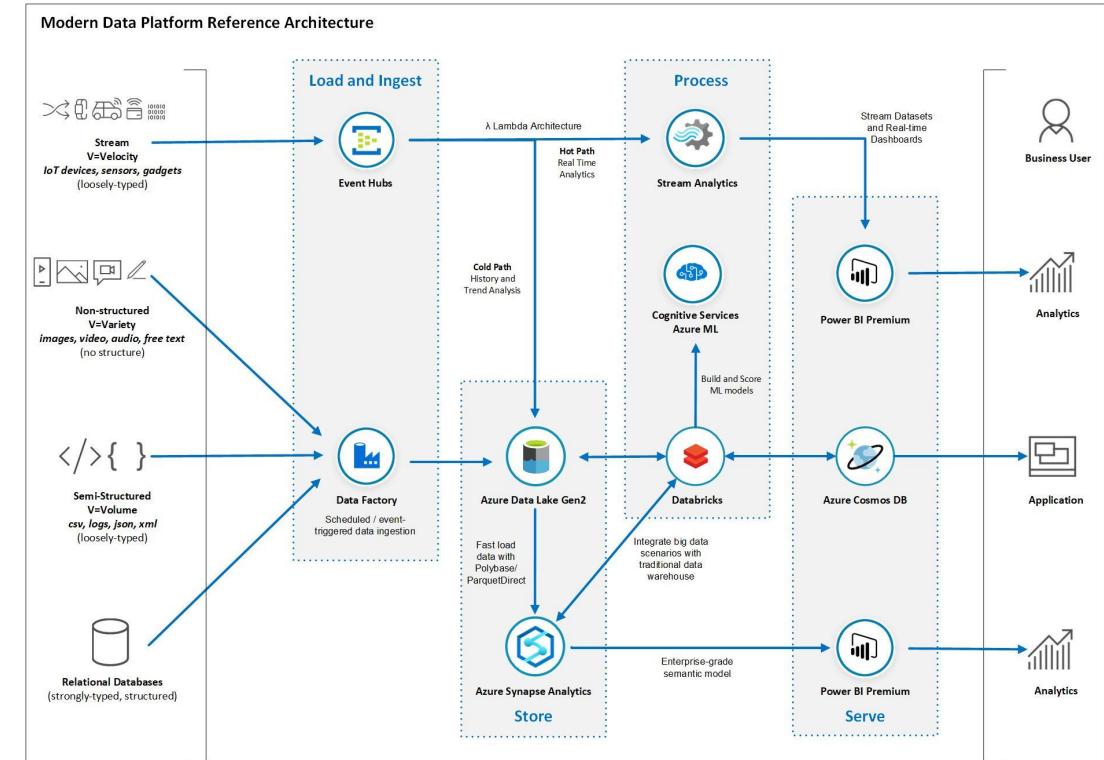
# Architecture components

- Azure Data Factory
- Azure Data Lake Gen2
- Azure Synapse Analytics
- Azure Databricks
- Azure Cosmos DB
- Azure Cognitive Services
- Azure Event Hubs
- Azure Stream Analytics
- Microsoft Power Bi

# Scenarios

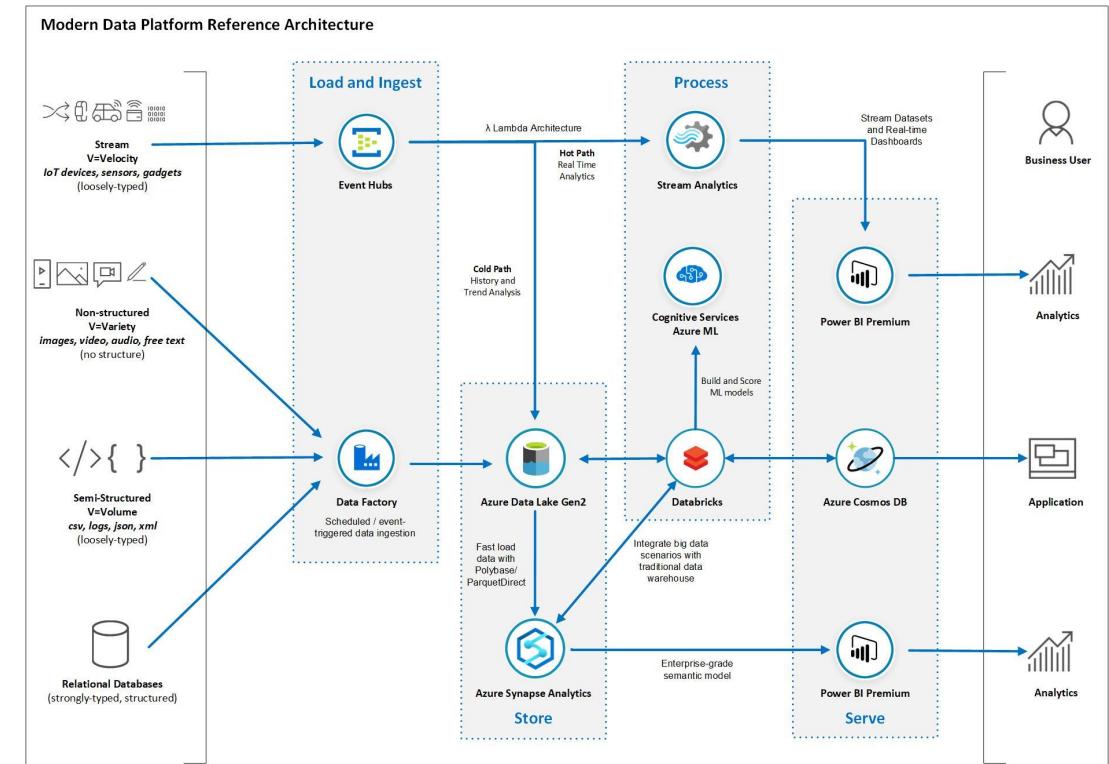


- Traditional relational data pipelines
  - Big data transformations
  - Unstructured data ingestion and enrichment with AI-based functions
  - Stream ingestion and processing following the Lambda architecture
  - Serving insights for data-driven applications and rich data visualization



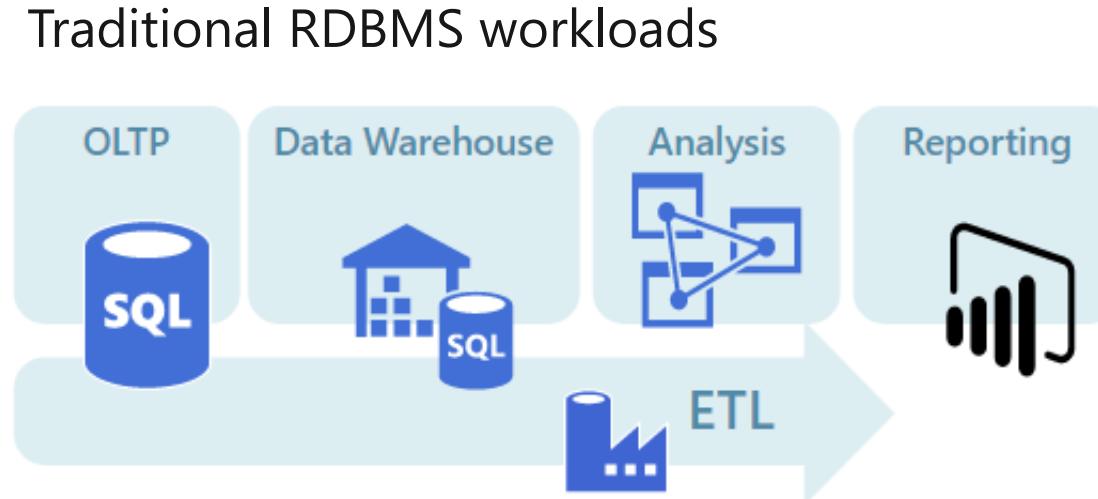
# Use cases

- Establish an enterprise-wide data hub consisting of a data warehouse for structured data and a data lake for semi-structured and unstructured data. This data hub becomes the single source of truth for your data.
- Integrate relational data sources with other unstructured datasets with the use of big data processing technologies
- Use semantic modeling and powerful visualization tools for simpler data analysis



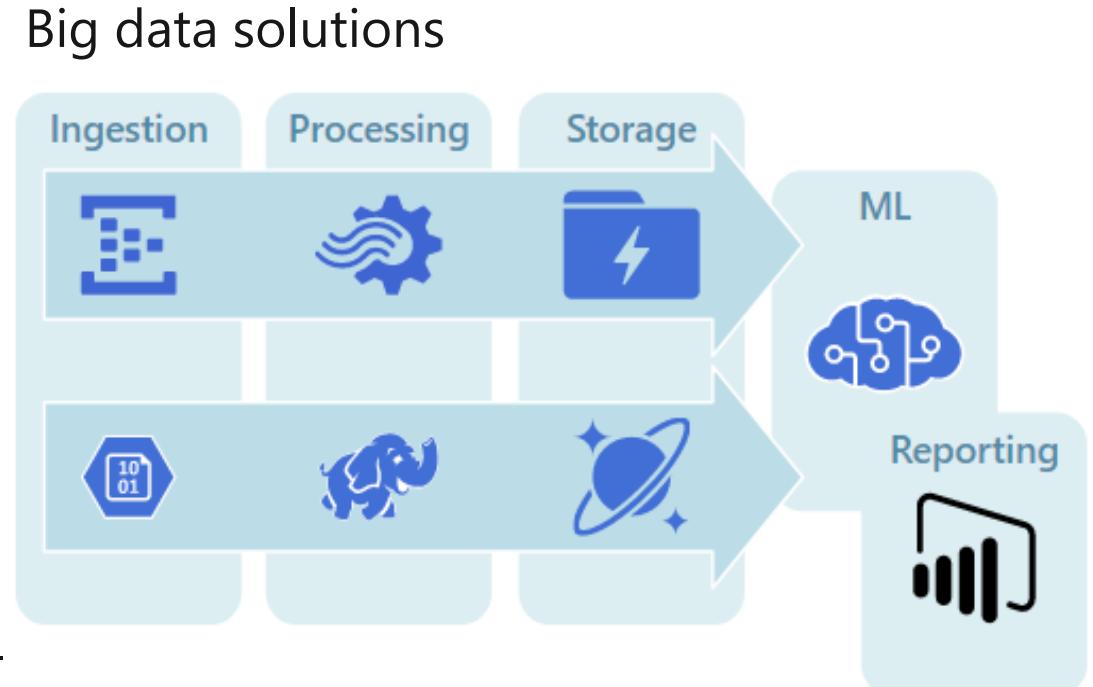
# Azure Data Architecture

- These workloads include online transaction processing (OLTP) and online analytical processing (OLAP)
- Data in OLTP systems is typically relational data with a predefined schema and a set of constraints to maintain referential integrity
- Often, data from multiple sources in the organization may be consolidated into a data warehouse, using an ETL process to move and transform the source data.



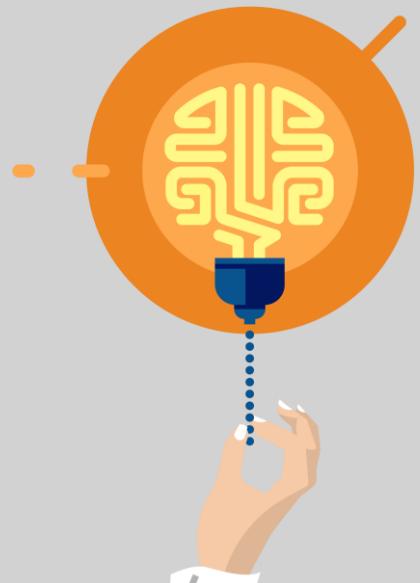
# Azure Data Architecture

- A big data architecture is designed to handle the ingestion, processing, and analysis of data that is too large or complex for traditional database systems. The data may be processed in batch or in real time.
- Big data solutions typically involve a large amount of non-relational data, such as key-value data, JSON documents, or time series data.
- Often traditional RDBMS systems are not well-suited to store this type of data.
- The term NoSQL refers to a family of databases designed to hold non-relational data. (The term isn't quite accurate, because many non-relational data stores support SQL compatible queries.)



## Components

#AzConfDev



# Azure SQL

## SQL virtual machines

Best for applications requiring OS-level access or specific SQL version



### SQL virtual machine

- SQL Server and OS server access
- Expansive SQL and OS version support
- Automated manageability features for SQL Server

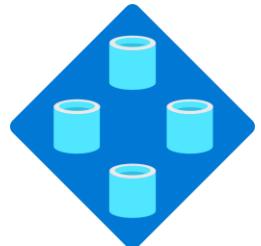
## Managed instances

Best for most migrations to the cloud and modernization along the way



## Databases

Best for modern cloud applications. Hyperscale and serverless options are available



### Single instance

- SQL Server surface area (vast majority)
- Native virtual network support
- Fully managed service

### Instance pool

- Resource sharing between multiple instances to price optimize
- Enables migration of many small instances at scale
- Fully managed service

### Single database

- Hyperscale storage (up to 100TB)
- Serverless compute
- Fully managed service

### Elastic pool

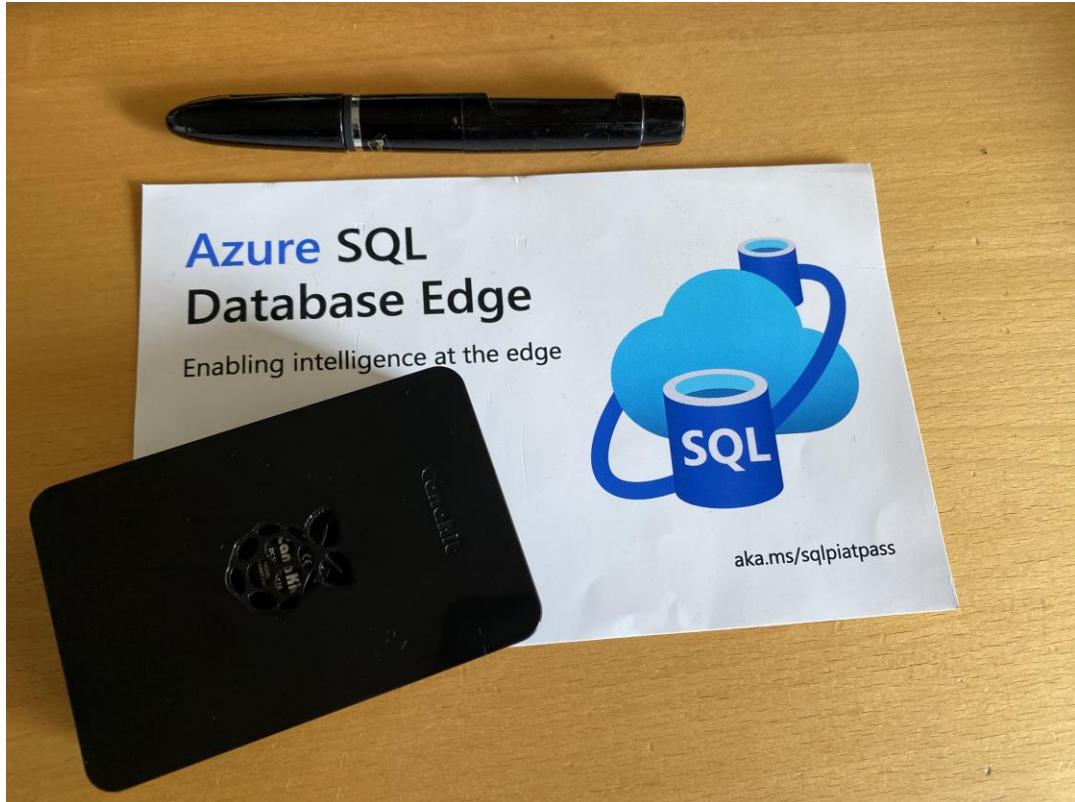
- Resource sharing between multiple databases to price optimize
- Simplified performance management for multiple databases
- Fully managed service

# Azure SQL Database Edge

Small-footprint, edge-optimized data engine  
with built-in AI



# Azure SQL Database Edge



# Azure Data Platform

Microsoft Azure

Home > New > Azure SQL

## Azure SQL

Microsoft



**Azure SQL** Save for later Microsoft

**Create**

[Overview](#) [Plans](#)

Azure SQL allows you to create and manage your SQL Server resources from a single view, ranging from fully managed PaaS databases to IaaS virtual machines with direct OS and database engine access. All deployment options enable you to bring your on-premises licenses to Azure using Azure Hybrid Benefit.

**Databases**  
Single databases are optimized for modern application development of new cloud-born applications. Databases provide a fully managed SQL experience with extensive and easy to use manageability features.  
**Includes:** single databases, elastic pools, and database servers

**Managed instances**  
Managed instances provide the PaaS benefits of SQL databases with added capabilities that were previously only available in SQL virtual machines. This includes a native virtual network and near 100% compatibility with on-premises SQL Server.  
**Includes:** single instances, instance pools

**SQL virtual machines**  
SQL virtual machines offer an IaaS architecture with extensive control over SQL Server and the underlying OS. Deployments include a management resource that focuses on SQL configuration and enables license updates with no server downtime.  
**Includes:** 60+ available images combining SQL Server 2008-2019 and a variety of available OS and license types

Useful Links  
[Documentation](#)  
[Product information](#)

# Azure Data Platform

Microsoft Azure ≡ Search resources, services, and docs (G+/-)

Home > New > Azure SQL > Select SQL deployment option

## Select SQL deployment option

Microsoft Feedback

### How do you plan to use the service?

 **SQL databases**  
Best for modern cloud applications. Hyperscale and serverless options are available.

Resource type

Create Show details

 **SQL managed instances**  
Best for most migrations to the cloud. Lift-and-shift ready.

Resource type

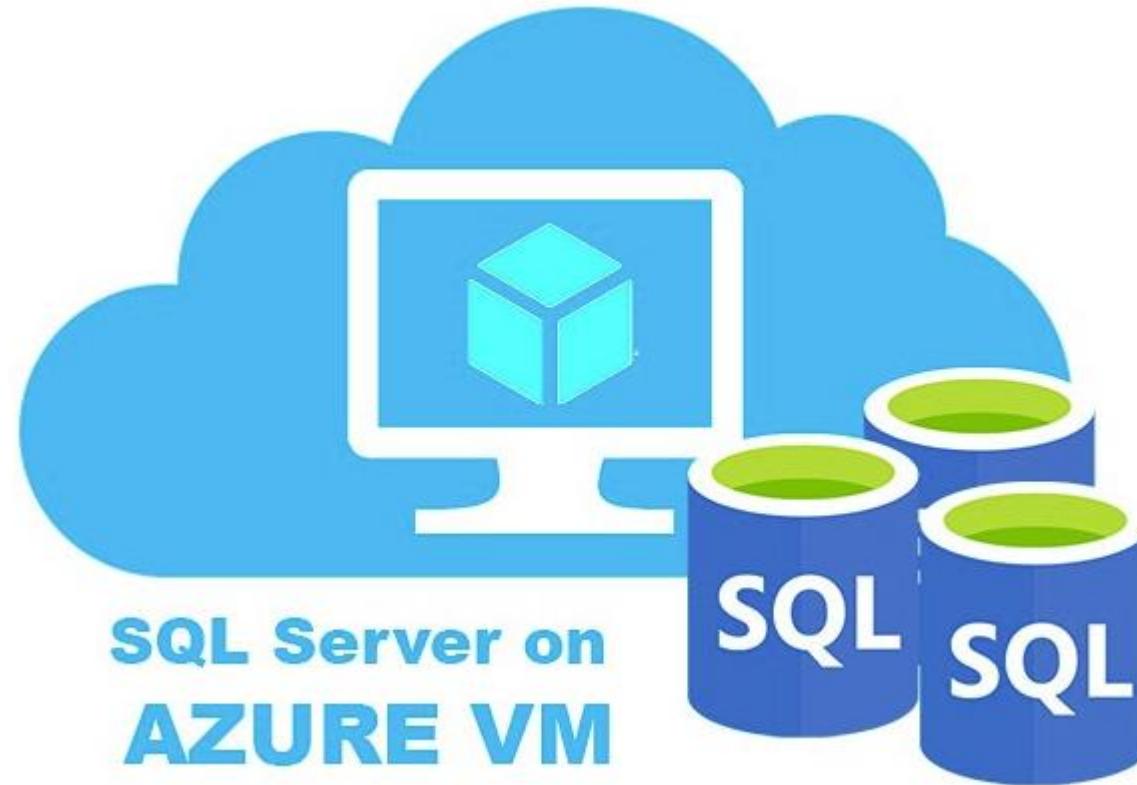
Create Show details

 **SQL virtual machines**  
Best for migrations and applications requiring OS-level access. Lift-and-shift ready.

Image

Create Show details

# SQL Server on Azure VM (IaaS)



# Azure SQL Database (PaaS)



# Azure SQL Database Managed Instance



# Azure Database for PostgreSQL



# Azure Database for MariaDB

Enterprise-ready, fully managed community  
MariaDB



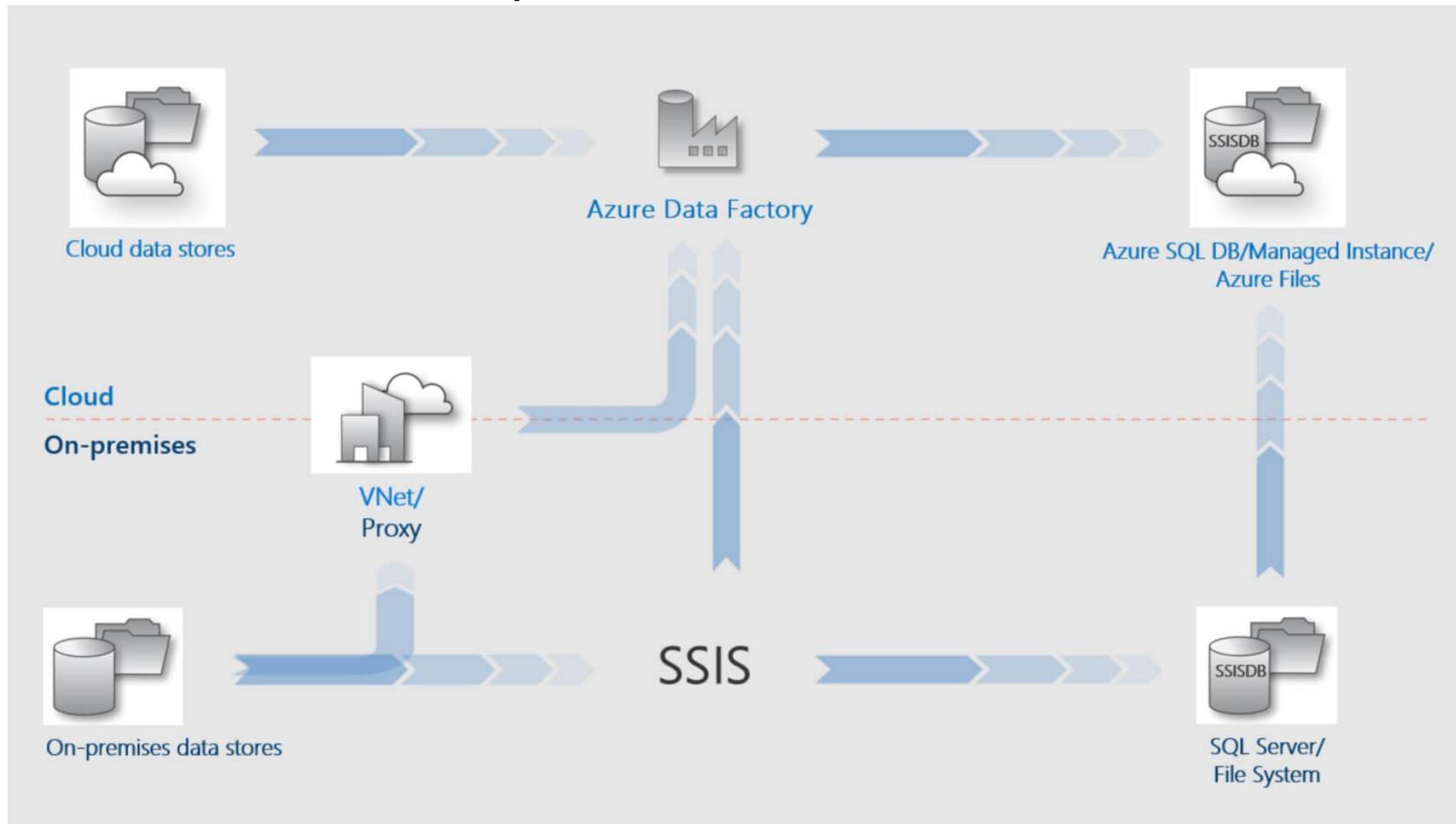
# Azure Database for MySQL



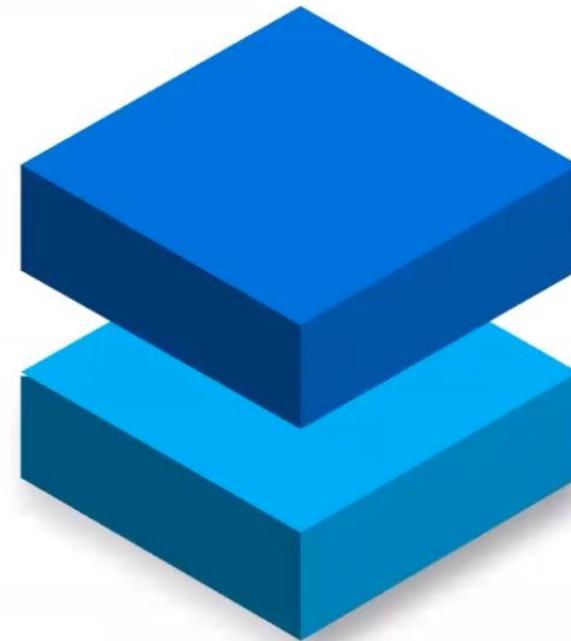
# Azure Data Factory



# Azure Data Factory



# Azure Databricks



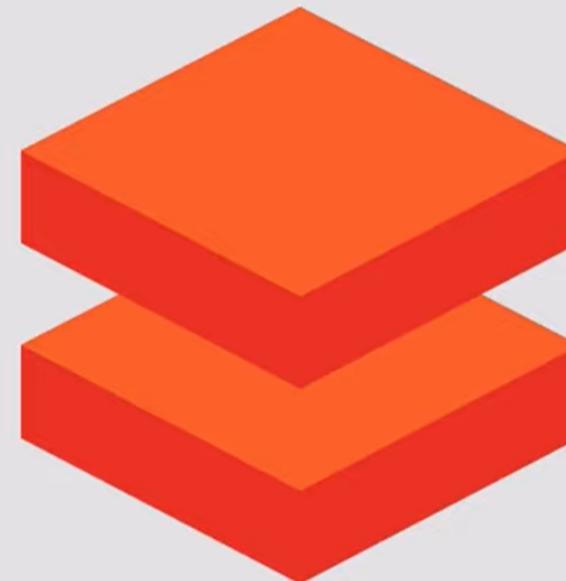
Fast, Easy, and Collaborative

# Azure Databricks



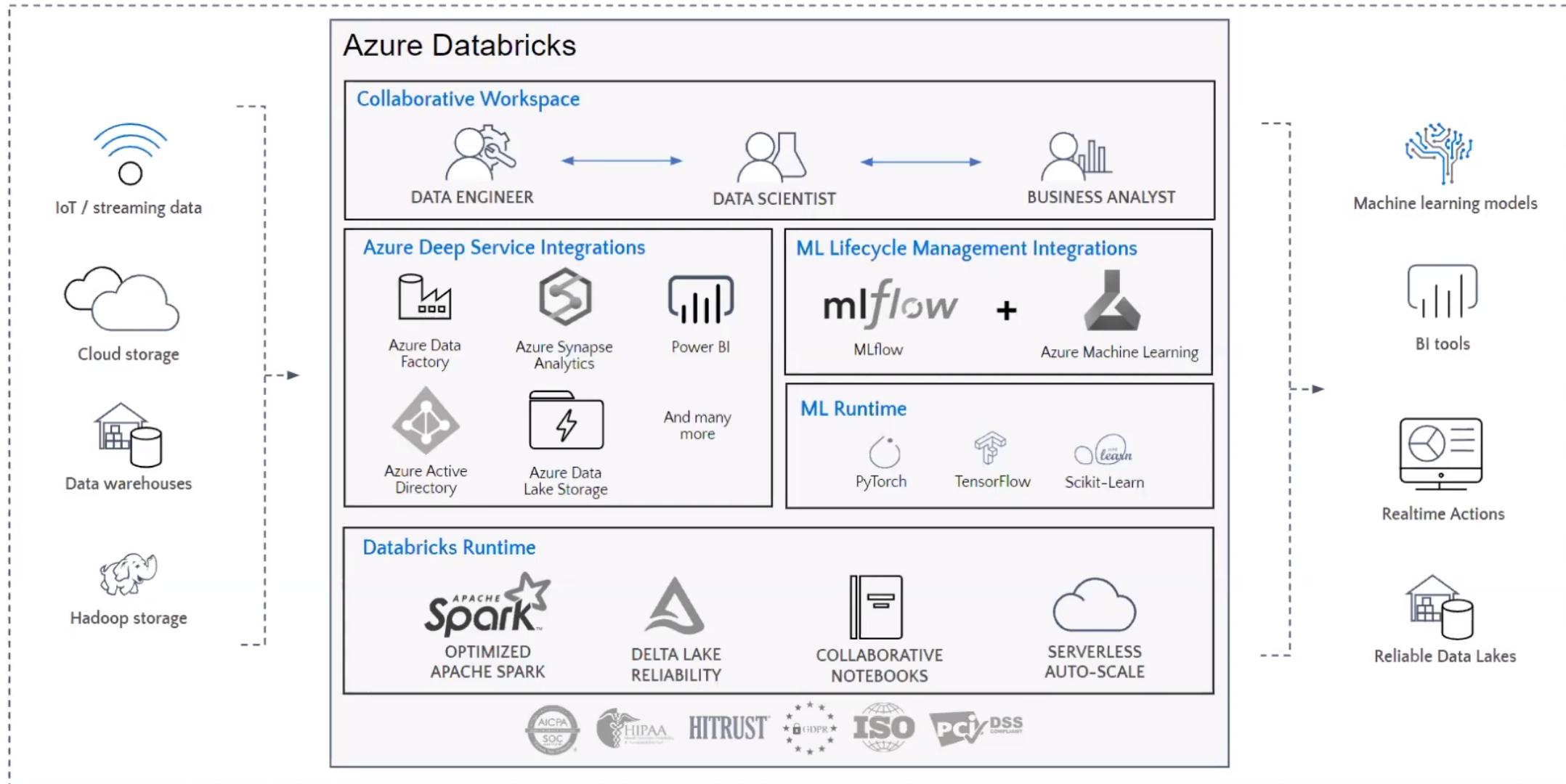
Azure Databricks is an Apache Spark-based analytics platform optimized for the Microsoft Azure cloud services platform.

Designed with the founders of Apache Spark, Databricks is integrated with Azure to provide one-click setup, streamlined workflows, and an interactive workspace that enables collaboration between data scientists, data engineers, and business analysts.

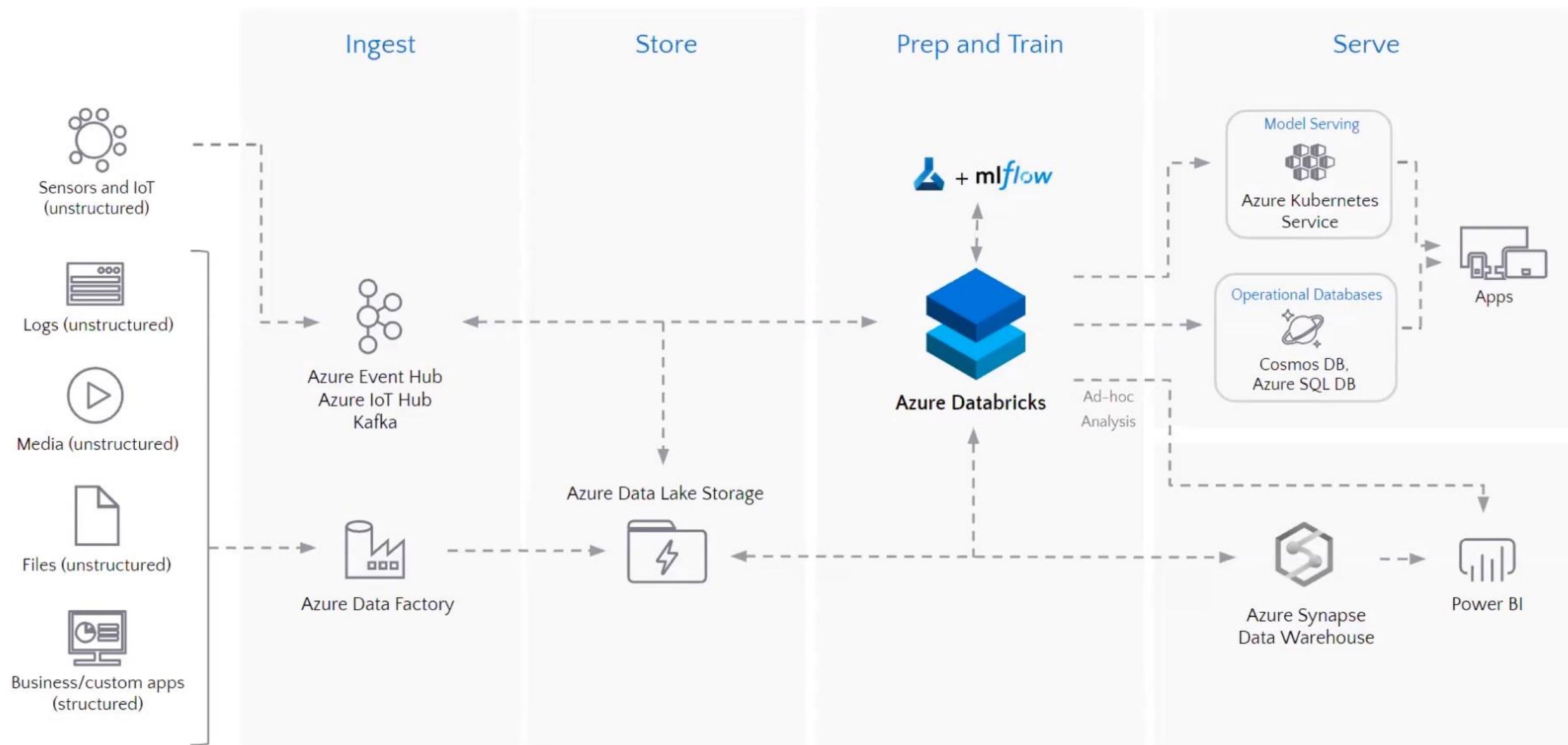




# Azure Databricks – Introduction



# Azure Databricks – Architecture Example



# Apache Spark Ecosystem



## Apache Spark Ecosystem

Spark SQL  
+ DataFrames

Streaming

MLib  
Machine Learning

GraphX  
Graph Computation

Spark Core API

R

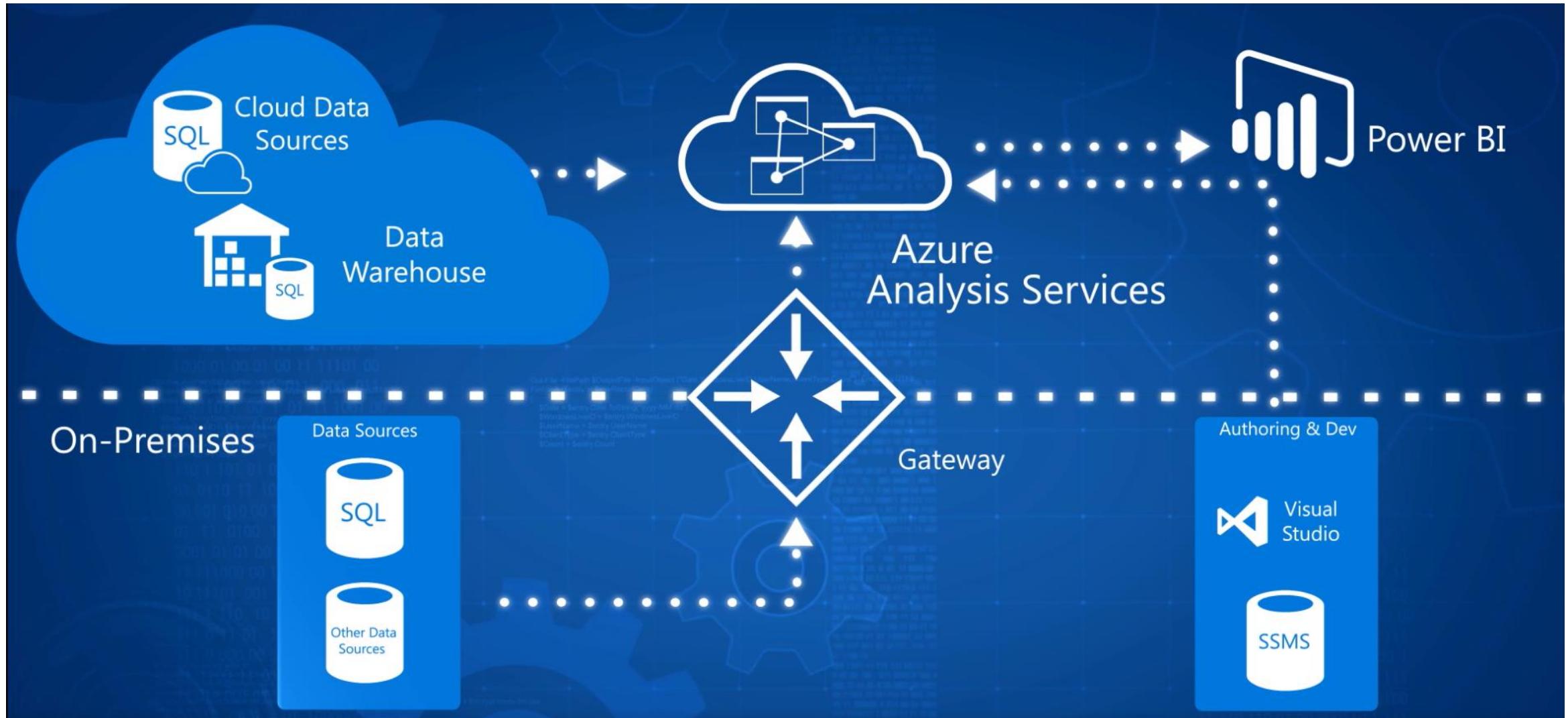
SQL

Python

Scala

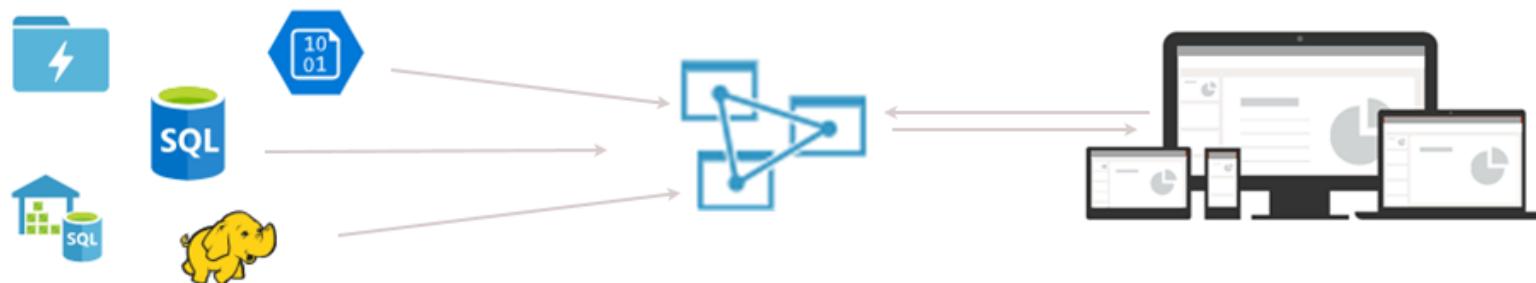
Java

# Azure Analysis Services

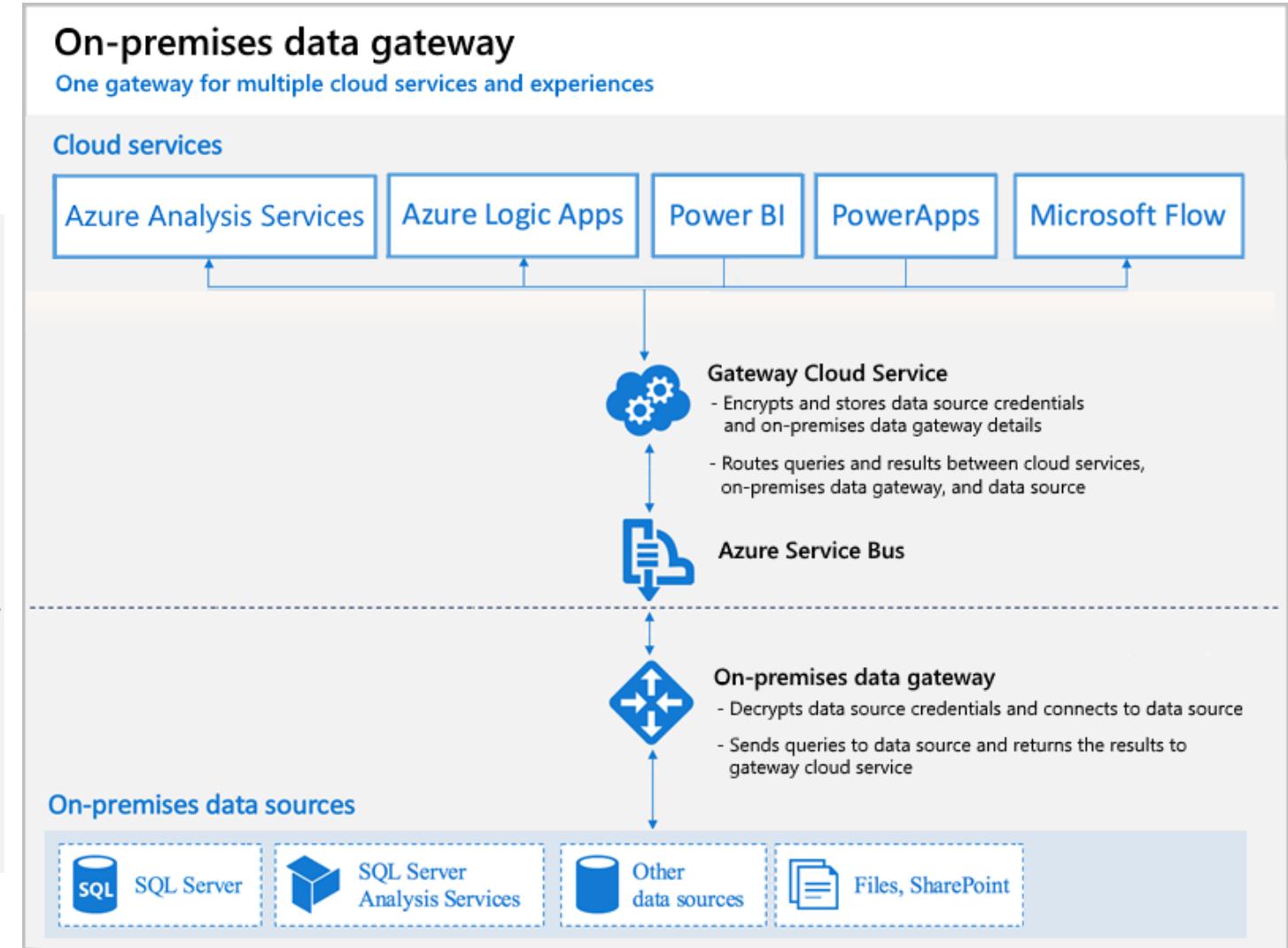
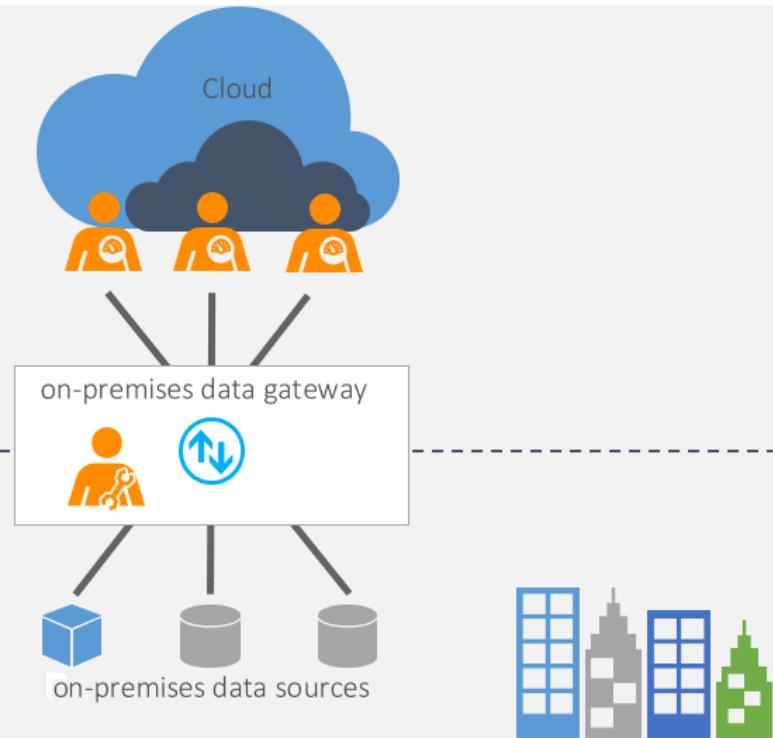


# On Premise Data Gateway

The on-premises data gateway provides secure data transfer between on-premises data sources and your Azure Analysis Services servers in the cloud. In addition to working with multiple Azure Analysis Services servers in the same region, the latest version of the gateway also works with Azure Logic Apps, Power BI, Power Apps, and Power Automate. While the gateway you install is the same across all of these services, Azure Analysis Services and Logic Apps have some additional steps.



# On Premise Data Gateway



# Azure Cosmos DB



# Azure Cosmos DB

A globally distributed, massively scalable, multi-model database service

**SQL**



MongoDB



Table API



Key-value



Column-family



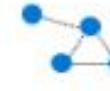
**Gremlin**  
 $G = (V, E)$



*cassandra*



Document



Graph

Elastic scale out  
of storage & throughput

Guaranteed low latency at the 99<sup>th</sup> percentile

Five well-defined consistency models

Turnkey global distribution

Comprehensive SLAs



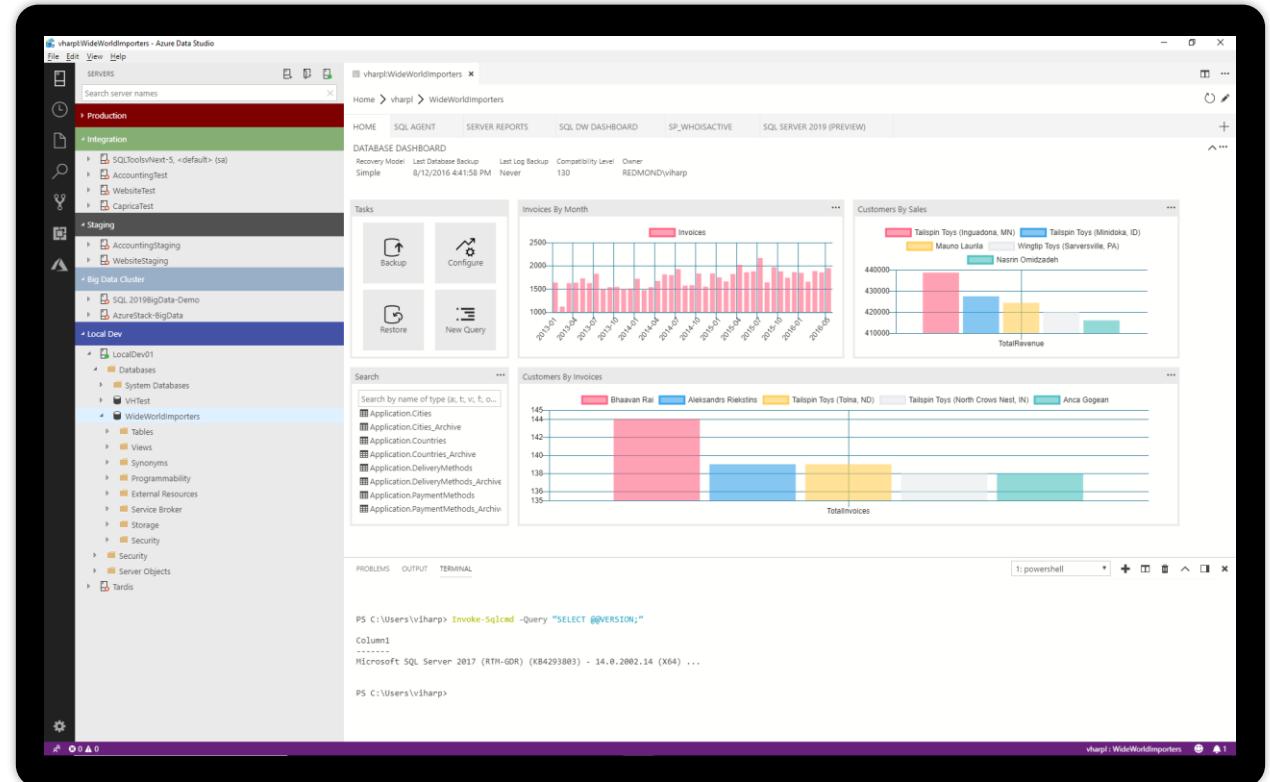
# The Azure Data Studio tools experience

**Azure Data Studio** is a lightweight, open source, cross-platform graphical management tool and code editor

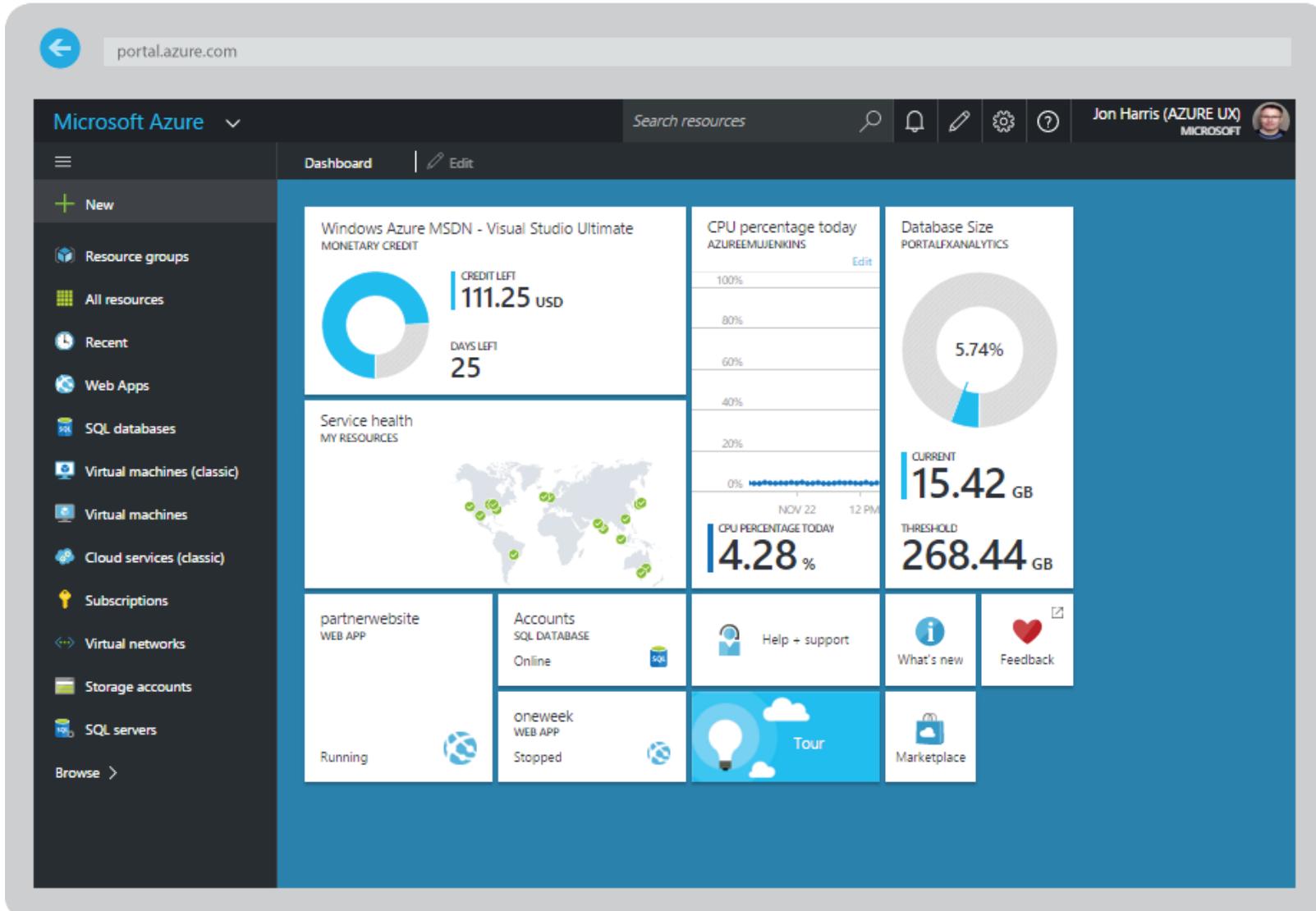
Enable a modern DevOps experience for database developers and DBAs on their platform of choice

Simplify development, configuration, management, monitoring and troubleshooting for SQL databases on-premises and in the cloud

Use **SQL Server Management Studio 18** to access, configure, manage, and administer all SQL Server components



# Azure Portal

A screenshot of the Microsoft Azure Portal dashboard. The left sidebar contains navigation links such as New, Resource groups, All resources, Recent, Web Apps, SQL databases, Virtual machines (classic), Virtual machines, Cloud services (classic), Subscriptions, Virtual networks, Storage accounts, and SQL servers. The main dashboard features several cards: 1) Windows Azure MSDN - Visual Studio Ultimate MONETARY CREDIT showing 111.25 USD credit left and 25 days left. 2) CPU percentage today AZUREEMJENKINS showing a chart from 0% to 100% with a current value of 4.28%. 3) Database Size PORTALFXANALYTICS showing a chart for PORTALFXANALYTICS with a current value of 15.42 GB and a threshold of 268.44 GB. 4) Service health MY RESOURCES showing a world map with green dots indicating healthy resources. 5) A grid section with two rows: Row 1 contains "partnerwebsite WEB APP" (Running) and "Accounts SQL DATABASE Online". Row 2 contains "oneweek WEB APP Stopped" and "Tour" (with a lightbulb icon). Other buttons in this section include Help + support, What's new, Feedback, and Marketplace.

# Questions ?



@javier\_vill



<https://ar.linkedin.com/in/javiervillegas>



<http://sql-javier-villegas.blogspot.com.ar>



javier.ignacio.villegas@gmail.com

## Introduction to Azure Data Platform

# Thank you!!

# Gracias !!

