



Microsoft Azure Data Services & IoT Overview

Martin Pöckl
Technical Solution Architect – Data & AI | Microsoft Austria

Platform Services

Security & Management



Compute



Cloud Services



Service Fabric



Batch



Remote App

Web and Mobile



Web Apps



API Apps



Mobile Apps



Logic Apps



Notification Hubs

Developer Services



Visual Studio



Azure SDK



Team Project



Application Insights

Hybrid Operations



Azure AD Connect Health



AD Privileged Identity Management



Backup



Operational Insights



Import/Export



Site Recovery



StorSimple

Integration



Storage Queues



Biztalk Services



Hybrid Connections



Service Bus

Media & CDN



Media Services



Content Delivery Network (CDN)

Analytics & IoT



HDInsight



Machine Learning



Data Factory



Event Hubs



Stream Analytics



Mobile Engagement

Data



SQL Database



SQL Data Warehouse



Redis Cache



Search



DocumentDB



Tables

Compute



Virtual Machines



Containers

Storage



BLOB Storage



Azure Files



Premium Storage

Networking



Virtual Network



Load Balancer



DNS



Express Route



Traffic Manager



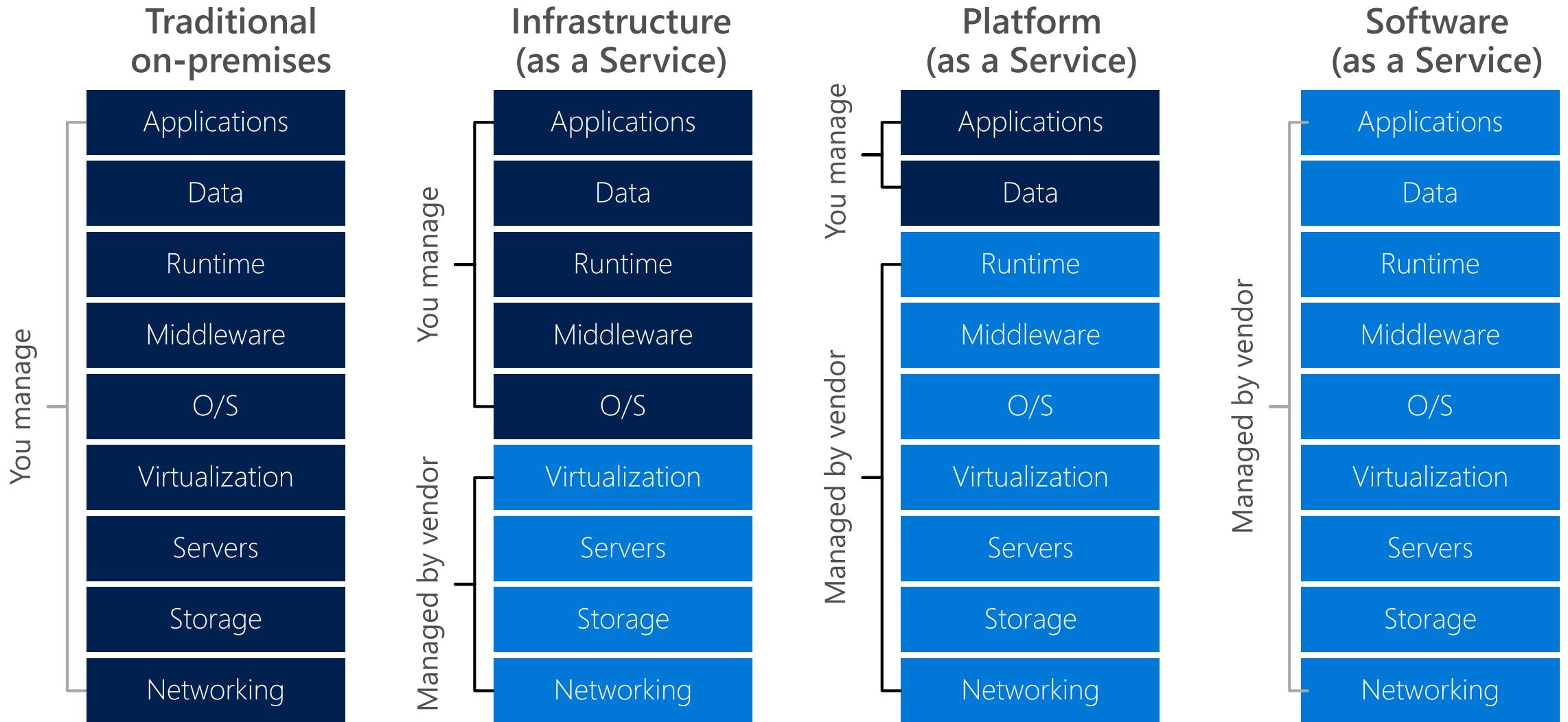
VPN Gateway



Application Gateway

Datacenter Infrastructure (42 Regions, 100+ Datacenters)

Cloud service models



Microsoft Azure data services

Data Store for Applications

(Transactional & mixed workloads, OLTP, operational analytics)

Relational, Non-relational (NoSQL, document-based, key-value, ...)

Data Store for Analytics

(Statistic & analytic queries, BI models, OLAP)

Relational (DWH), Non-relational ("Big Data", Data Lake)

BI & (Advanced) Analytics

BI models, (federated) queries, statistic & predictive models, in-stream analytics

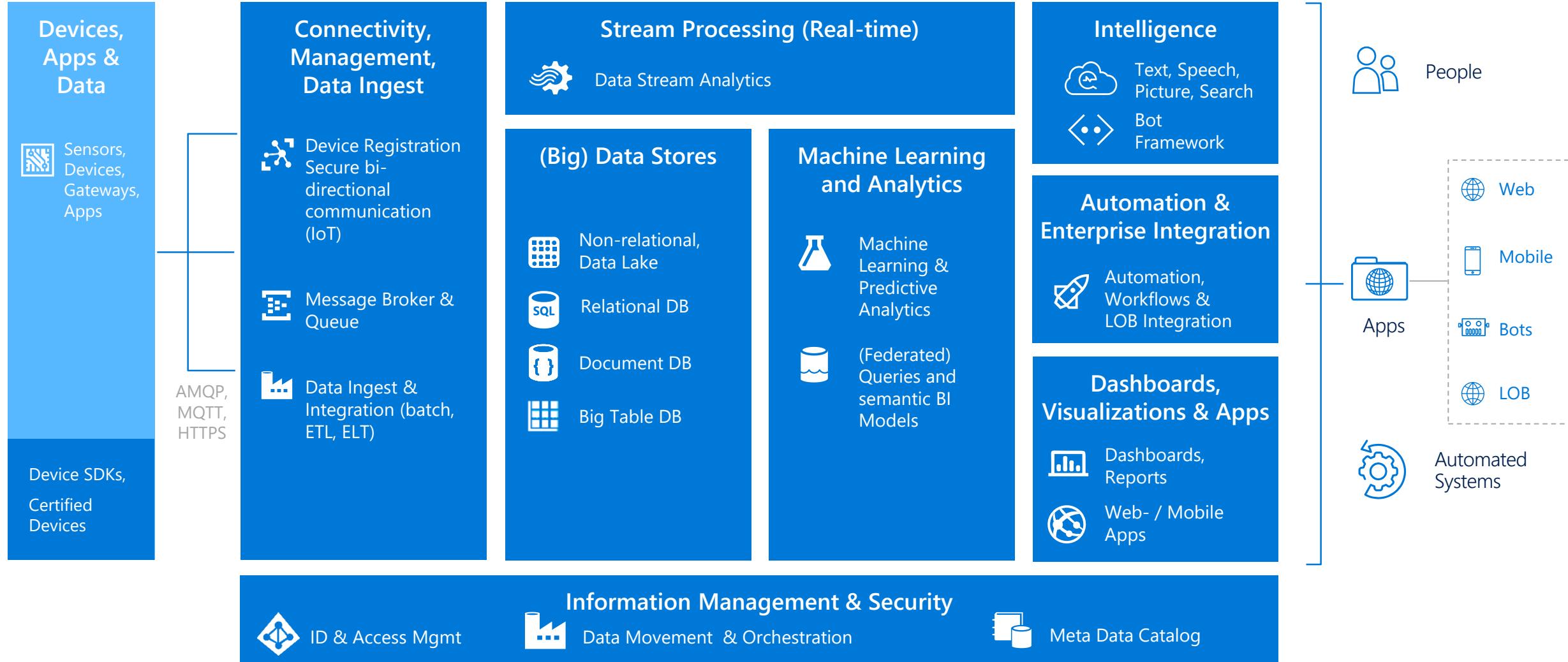
Data Engineering & Information Management

Data orchestration, data catalog, data integration/ingest

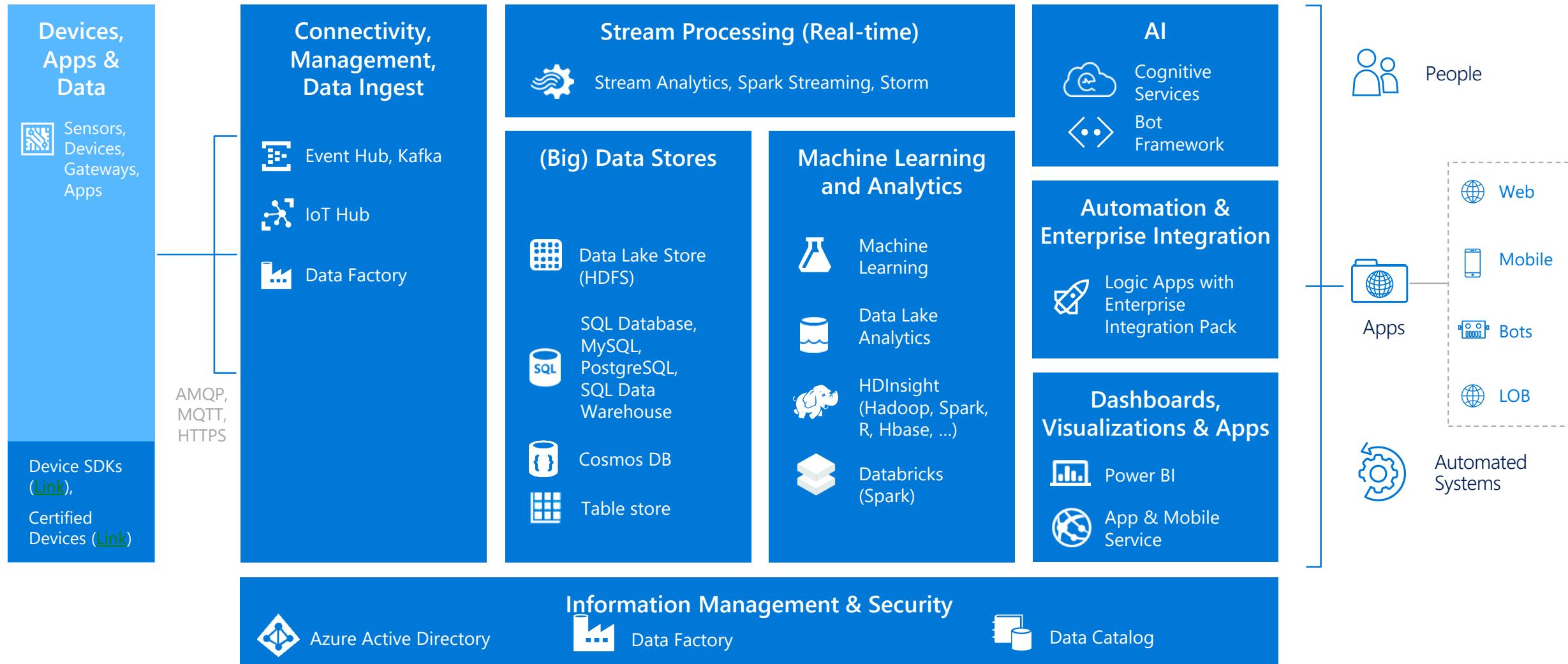
Data Movement & Connectivity

Gateways (Cloud – on-premises), Bus, Queues, Message Broker, IoT Hub

Microsoft Data, Analytics & IoT capabilities



Azure (Big) Data, Analytics & IoT services

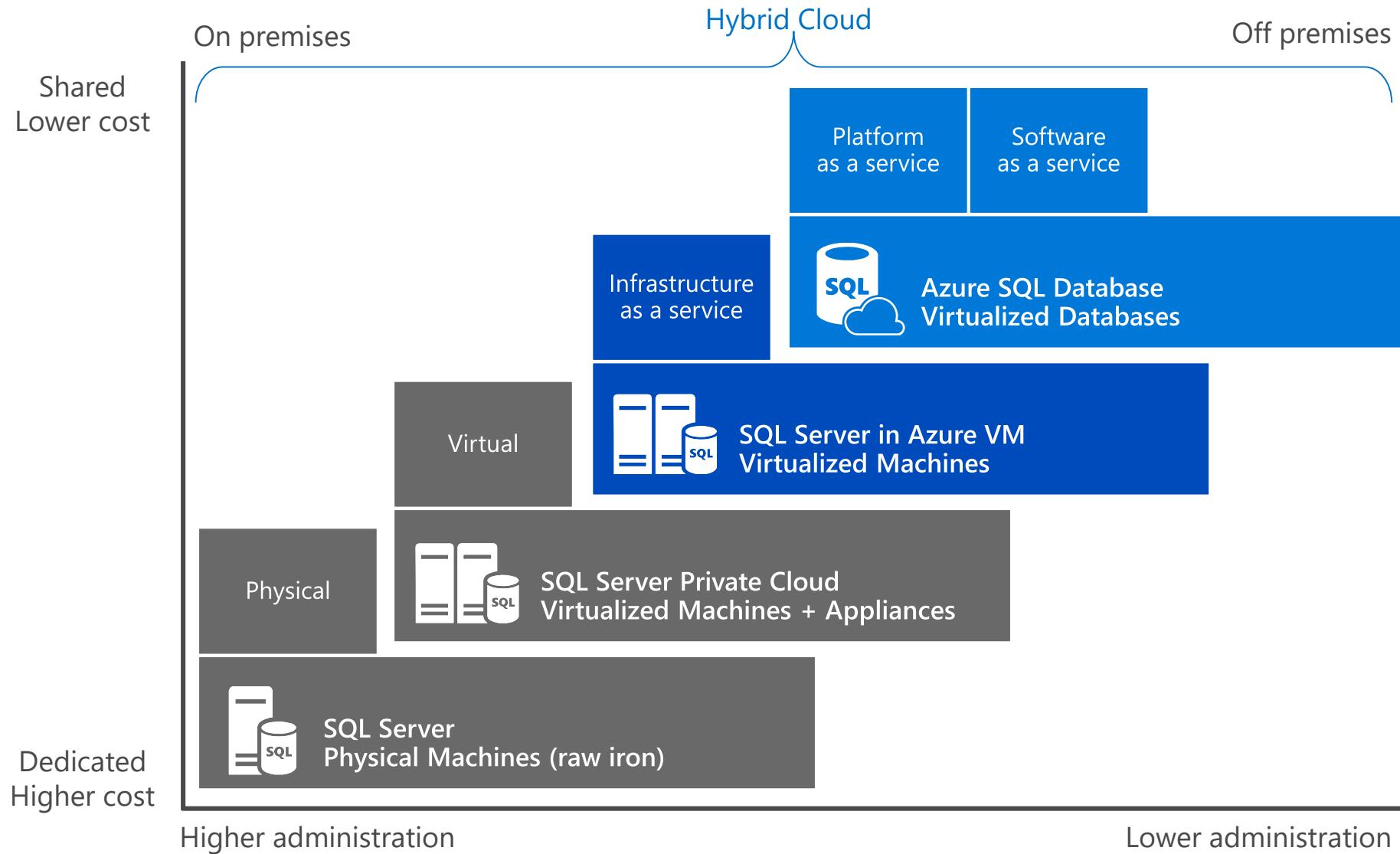


Data Stores in Azure

Database/stores in IaaS Unmanaged		Database/stores in PaaS Managed		
VMs	Containers	Relational Data	Non-relational, structured Data	Un-/semi- structured Data
Everything you can run in a VM Linux or Windows based 100s of pre-installed images	Everything database/store you can run in a container (SQL Server on Docker)	SQL Database, SQL Data Warehouse, PostgreSQL, MySQL	Cosmos DB (Document, Graph, Column-family, Key-value store)	Data Lake
+ Full range of technology + Full flexibility + Performance (128 Cores, 3.8 TB Memory, SSDs w 80.000 IOPS) + HA/DR, Data distribution - Installation & Configuration - Management & Operations	+ Full flexibility + Ease of use ~ Choice of technology - "Stateful" Containers (high availability, replicated data) - App Design (sharding) - Management & Operations	+ Ease of use + No Installation, no OS or Product Configuration + No Operations (very high SLAs), just Monitoring + Scalability ~ Choice of technology ~ App Design - Limited flexibility / control		

Relational Databases

Data platform continuum



AZURE SQL DATABASE

THE INTELLIGENT CLOUD DATABASE FOR APP DEVELOPERS

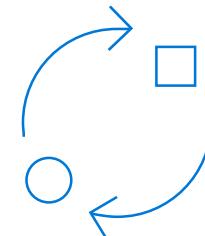
Intelligent
“Data Base as a Service”
DBaaS



Privacy
and trust



Seamless
and compatible



Competitive
TCO



Lift and shift
SQL Server apps

Migrate SQL Server
data marts

Build intelligent
apps with R built-in

Save up to 30% with Azure
Hybrid Benefit for SQL Server

AZURE SQL DATABASE



Intelligent DBaaS



Privacy and Trust



Seamless and compatible



Competitive TCO

AUTOMATIC PERFORMANCE ANALYSIS AND TUNING

Automatic database tuning
Automatic query plan correction
Performance Insight in OMS
Adaptive Query Processing

ACTIVITY & SECURITY MONITORING

Threat Detection
Centralized dashboard OMS

ACCESS CONTROL

SQL Firewall
AAD and MFA

DATA PROTECTION & SECURITY

Encrypt in motion (TLS)
TDE with BYK
Virtual Network
Data Discovery & Classification

DISCOVERY & ASSESSMENT

Vulnerability assessment

HA-DR BUILT-IN

99.99% SLA
Active Geo replicas (4)
Geo-restore
Multi-AZ

AUTOMATED BACKUP AND RESTORE

Backup with health check
35 days PITR
10 years data retention

DISTRIBUTED APPLICATION

Transaction replication
Data sync

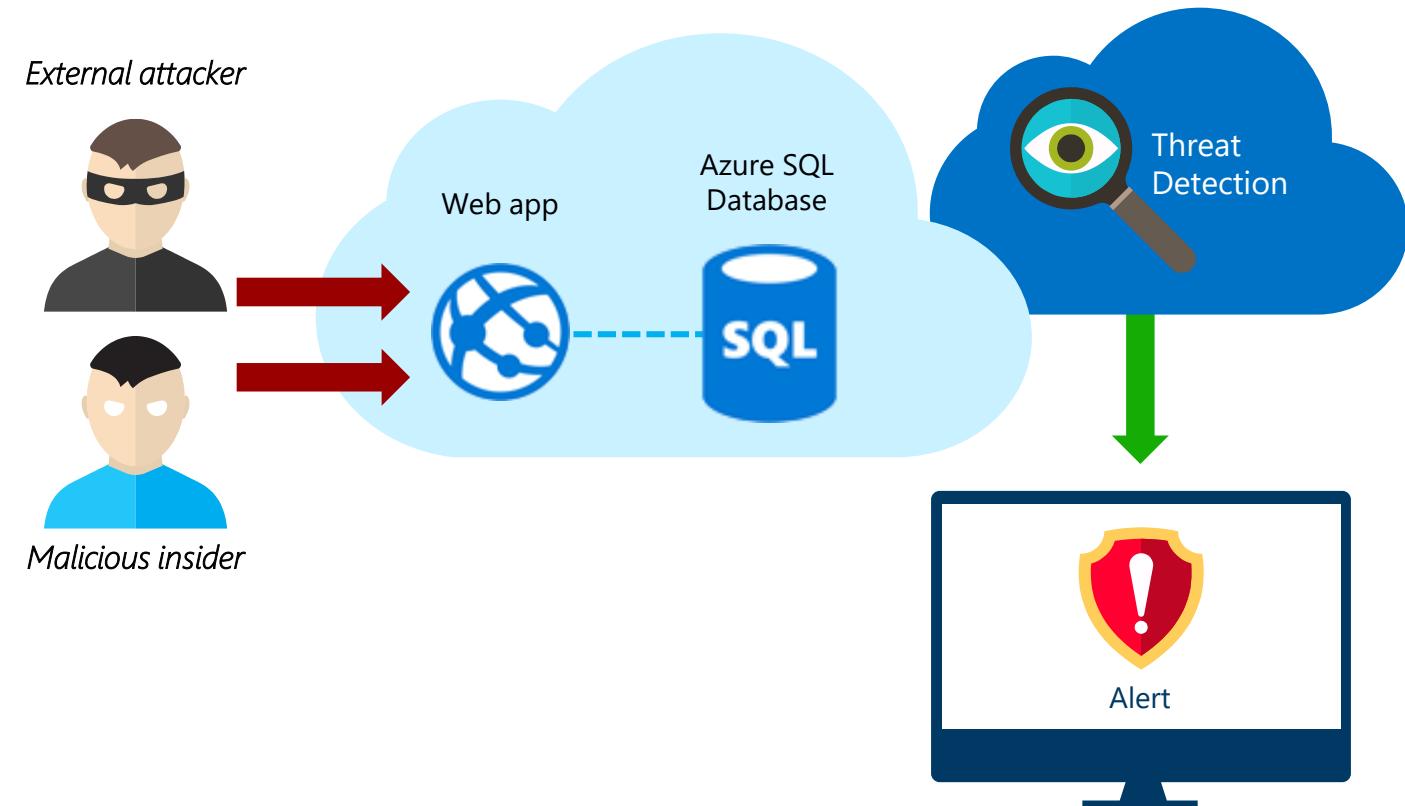
Threat detection

Detects anomalous database activities that could indicate potential threat

Configure threat detection policy in Azure Portal

Receive alerts from multiple database-threat detectors that identify anomalous activities

Explore audit log around the time of an event



Data Discovery & Classification

Secure the data, not just the database

Automated discovery of sensitive data

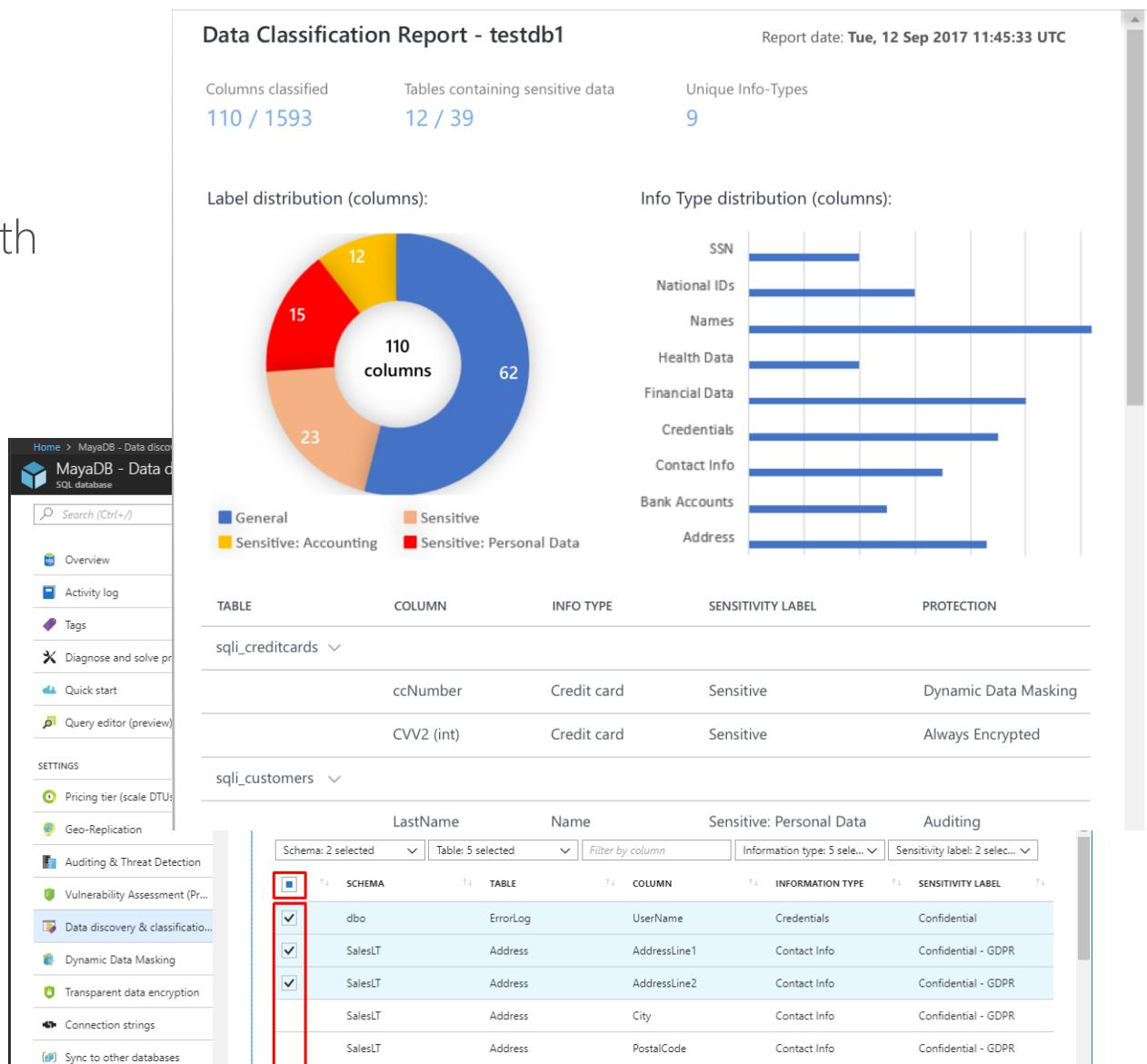
Labeling (tagging) sensitive data on column level with persistency

Classification as infrastructure for protection & compliance

Sensitivity metadata flows with data for protection outside database boundaries

Centralized policy management - ASC

Integration with unstructured data classification platforms (AIP/O365 - MIP)



AZURE DATABASE FOR MySQL, PostgreSQL AND MariaDB

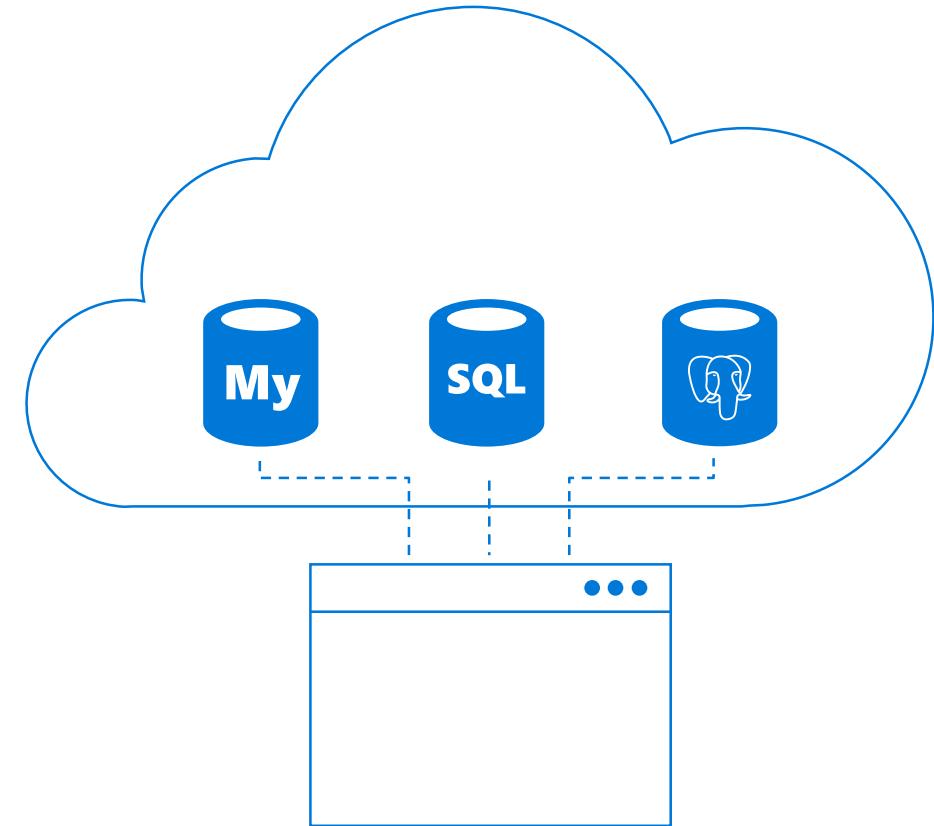
Work in the DB of your choice

Create with built-in high availability

Set up in minutes, scale on the fly

Sustain performance with adaptive improvements

Rest easy with unparalleled security



Lift and shift MySQL
and PostgreSQL apps

Migrate WordPress
and Drupal

Build apps with your
OSS DB of choice

Demo:
Work with Azure SQL Database

Non-relational Databases (NoSQL)

AZURE COSMOS DB

GLOBALLY DISTRIBUTED, MULTI-MODEL NOSQL DATABASE SERVICE

Multi-model data with your favorite API

Elastically scale storage and throughput

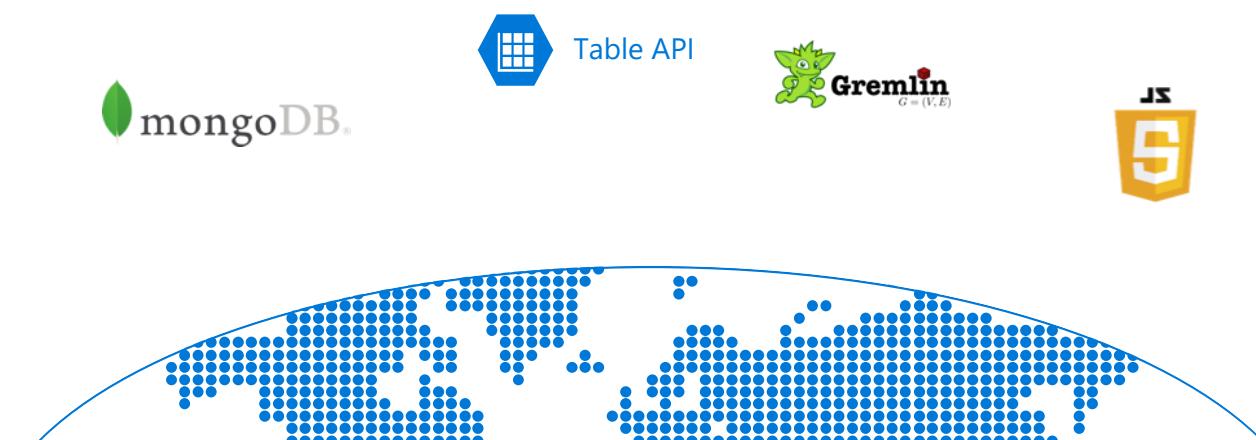
Multiple, well-defined consistency levels

<10ms latency guarantees at the 99th percentile

Industry-leading SLAs across performance,
latency, availability and throughput

Lift and shift
MongoDB apps

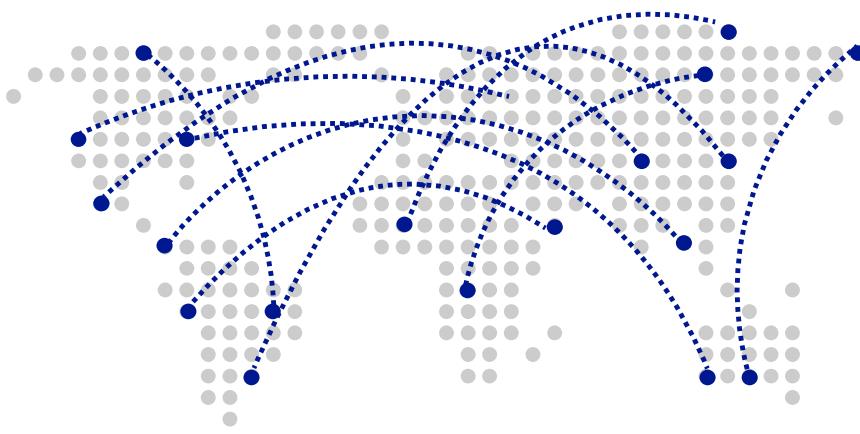
Run Spark over
operational data



Build real-time
customer experiences

Ideal for IoT, gaming
and eCommerce





Turnkey Global Distribution

On-demand global distribution

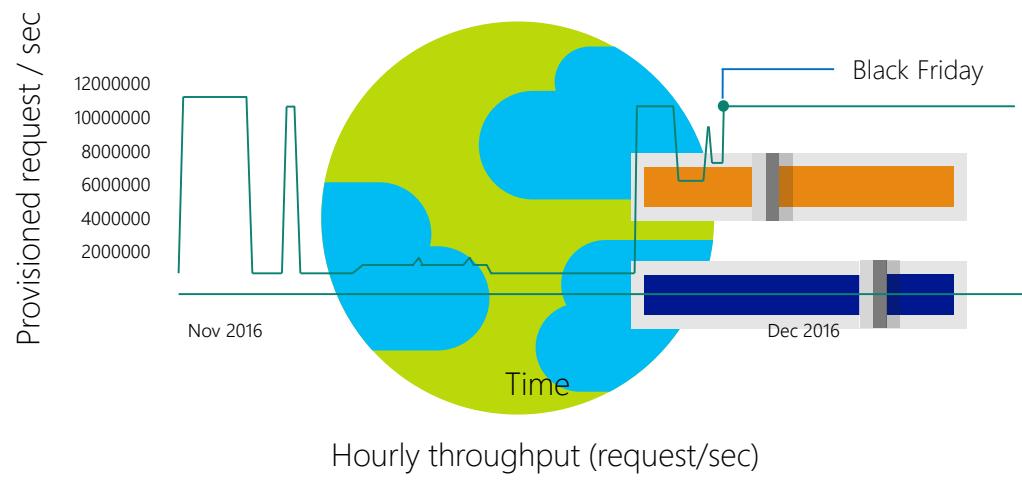
Worldwide presence as a Foundational Azure service

Automatic multi-region replication

Multi-homing APIs

Manual and automatic failovers

Designed for High Availability



Elastically scalable storage and throughput

Single machine is never a bottle neck

Transparent server-side partition management

Elastically scale storage (GB to PB) and throughput (100 to 100M req/sec) across many machines and multiple regions

Automatic expiration via policy based TTL

Pay by the hour, change throughput at any time for only what you need



Industry-leading, enterprise-grade SLAs

99.99% availability – even with a single region

Made possible with highly-redundant storage architecture

Guaranteed durability – writes are majority quorum committed

First and only service to offer SLAs on:

- Low-latency
- Consistency
- Throughput



Security & Compliance

Always encrypted at rest and in motion

Fine grained "row level" authorization

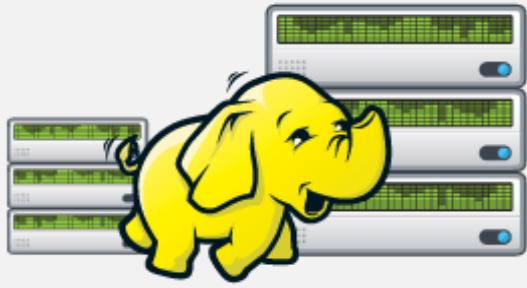
Network security with IP firewall rules

Comprehensive Azure compliance certification:

- ISO 27001
- ISO 27018
- EUMC
- HIPAA
- PCI
- SOC1 and SOC2

Big Data Stores

The big data “problems”



Huge volumes of data

Different formats of data

Analyze data in realtime

Data volumes



Data variety



Data velocity



Data volume

Hadoop stores files in a distributed file system ("HDFS")

Storage and computation is distributed across many servers

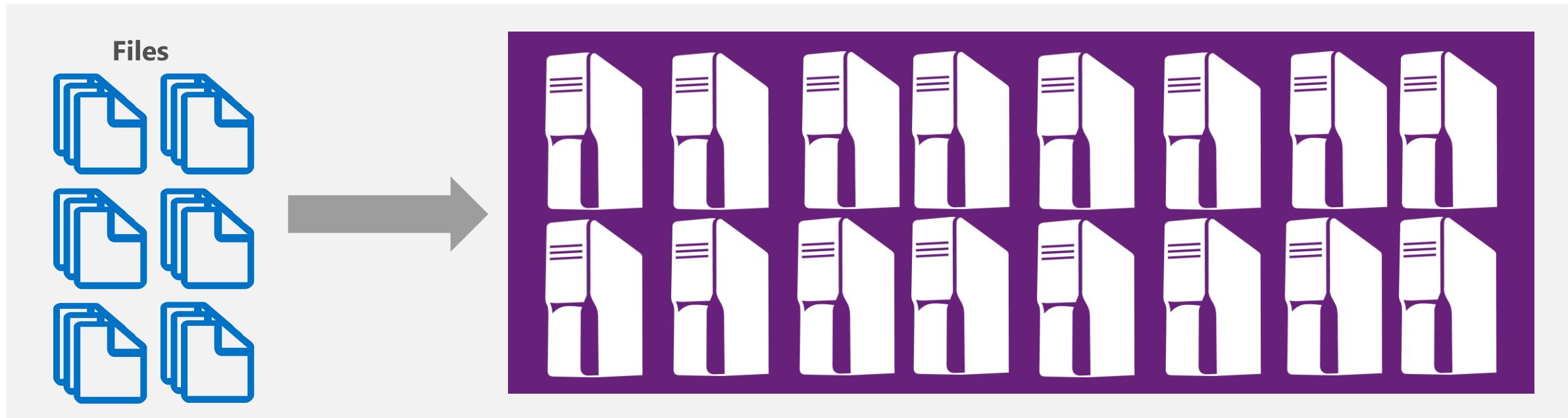
Files can be spread out over multiple nodes

Hadoop can store very large amounts of data

Combined storage resource can grow with demand from a few nodes to thousands of nodes

Scales out linearly

Very large files supported including those larger than the capacity of a single node



Data variety

File-based data store (non-relational store)

Files could have a variety of semi-structured or unstructured data

Previously, these files may not have been seen as providing value or insights

Today, new business questions and insights are being uncovered through data science



Sentiment
Understand how your customers feel about your brand and products—right now



Clickstream
Capture and analyze website visitors' data trails and optimize your website



Sensors
Discover patterns in data streaming automatically from remote sensors and machines



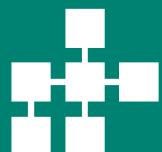
Geographic
Analyze location-based data to manage operations where they occur



Server logs
Research logs to diagnose process failures and prevent security breaches



Unstructured
Understand patterns in files across millions of web pages, emails, and documents

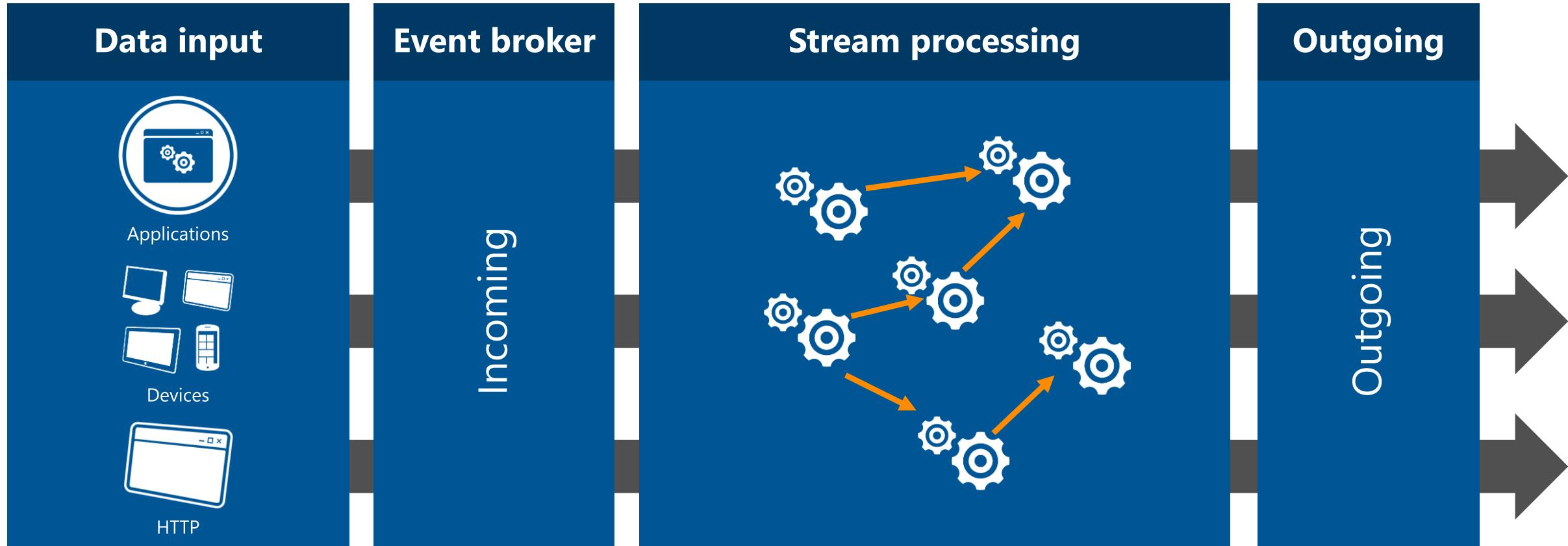


Data velocity

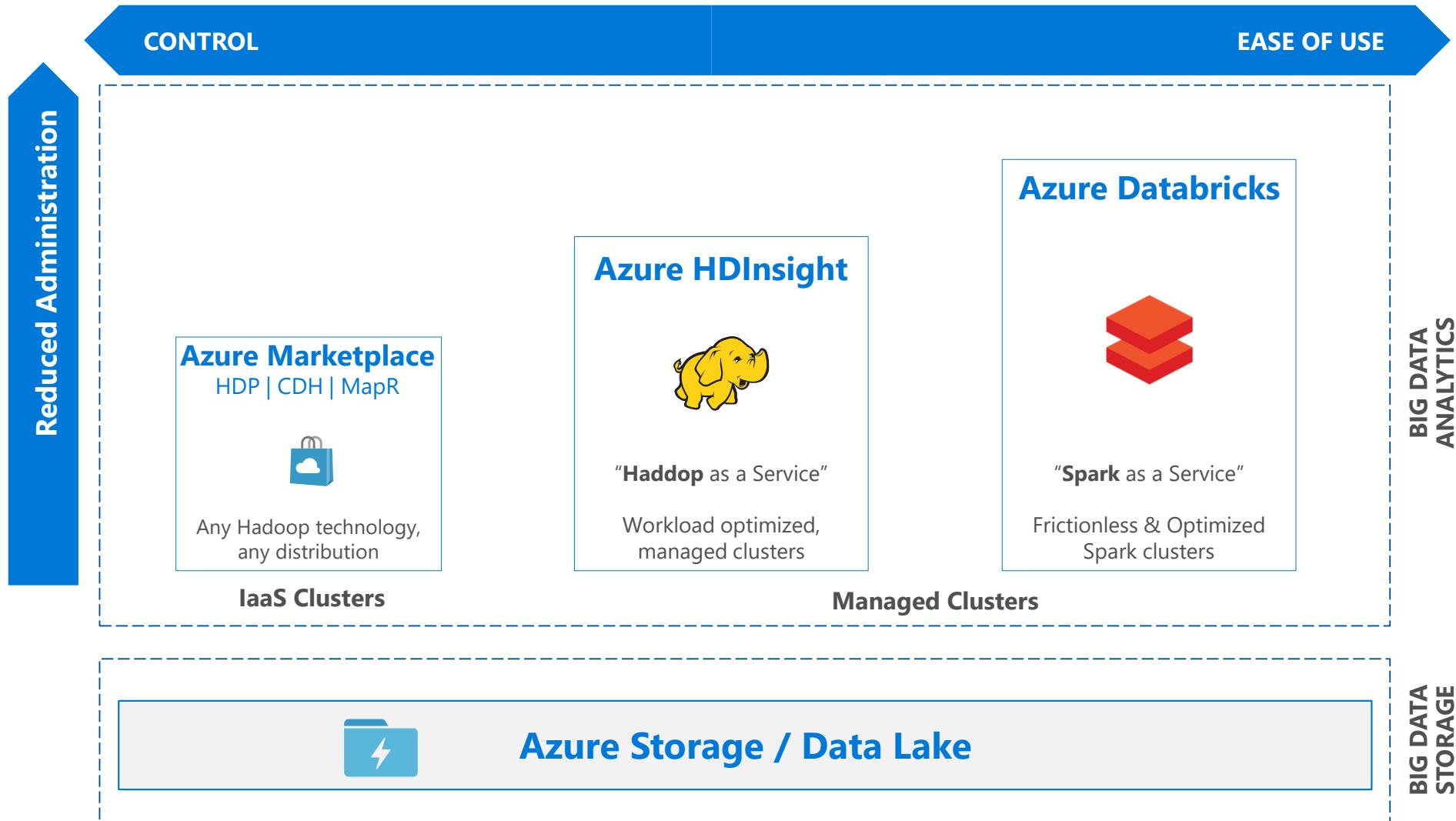
Stream live data and process it in real-time

Hadoop/Spark can act as scalable event stream ingestion

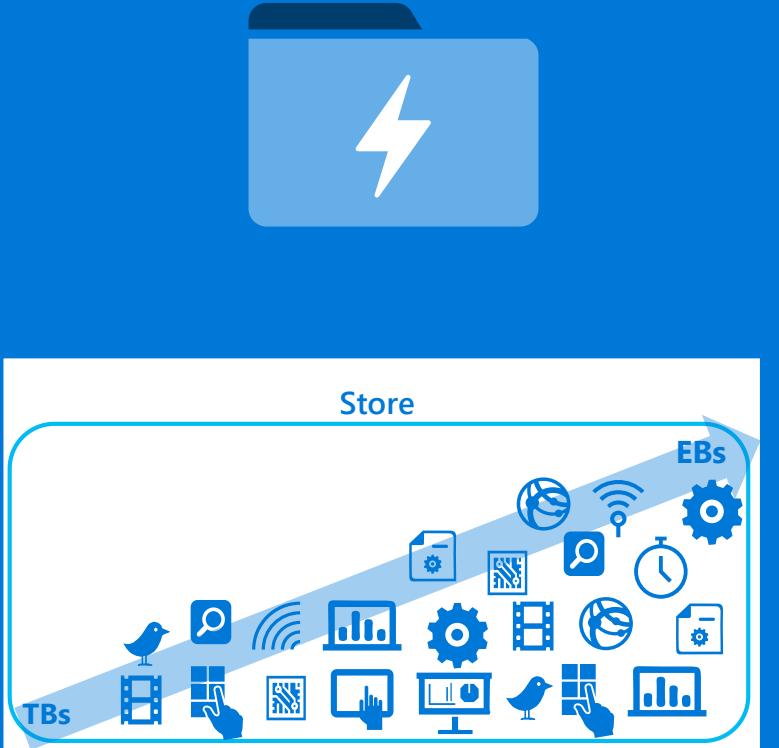
Hadoop/Spark can do near real-time in-stream processing



AZURE BIG DATA SOLUTIONS



Azure Data Lake



Petabyte size files and Trillions of objects

Scalable throughput for massively parallel analytics

Hadoop Distr. File System (HDFS) for the cloud

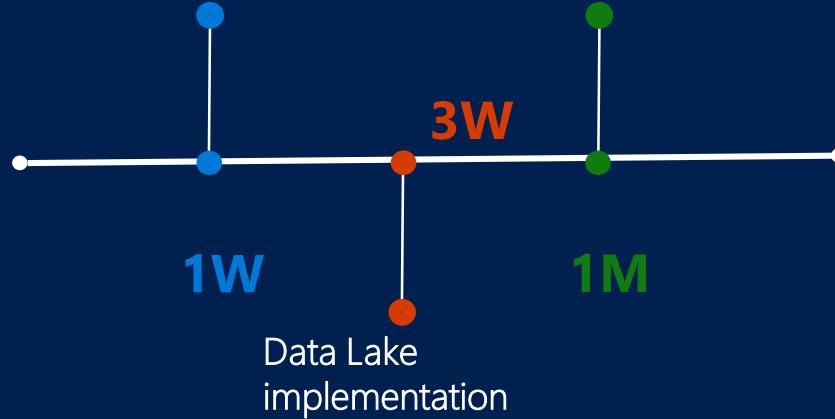
Always encrypted, role-based security & auditing

Enterprise-grade support

Integration with Azure Active Directory

PaaS Hadoop

Planning & Architecture Data Science Projects



On-Premise Hadoop

Planning & Architecture

Hardware Delivery

Data Lake Implementation

Data Science Projects

2M

6M

4M

2M

6M

2M

1.5 – 2 years

Project approval

HW Implementation

Test Phase

572.000€

Upfront Cost

280.000€

Annual Fees

Demo:
Work with Azure Databricks

Data Integration in Azure

Data-centric

Move & Transform data
Orchestrate & Schedule
Monitor & Manage the data pipeline
Bulk/Batch oriented

Data Pipelines
70+ Connectors
SSIS Integration Runtime

Azure Data Factory

Application-centric

Integrate internal & external applications/services
Business processes & workflows
Transaction-based

Event Trigger
130+ Connectors
Server less

Azure Logic Apps

Data Factory

Fully-managed data integration and orchestration service in the cloud

Microsoft Azure



CustomerChurnPipeline < CallDataRecords

Save ✓ Validate

Toolbox

Search

Source

- File
 - Azure Blob Storage
 - Amazon S3
- Relational
 - Azure SQL Database
 - Azure SQL Data Ware...

Sink

- File
 - Azure Blob Storage
 - Amazon S3
- Relational
 - Azure SQL Database
 - Azure SQL Data Ware...

Settings

General Mapping

Mapping Options

Automatic

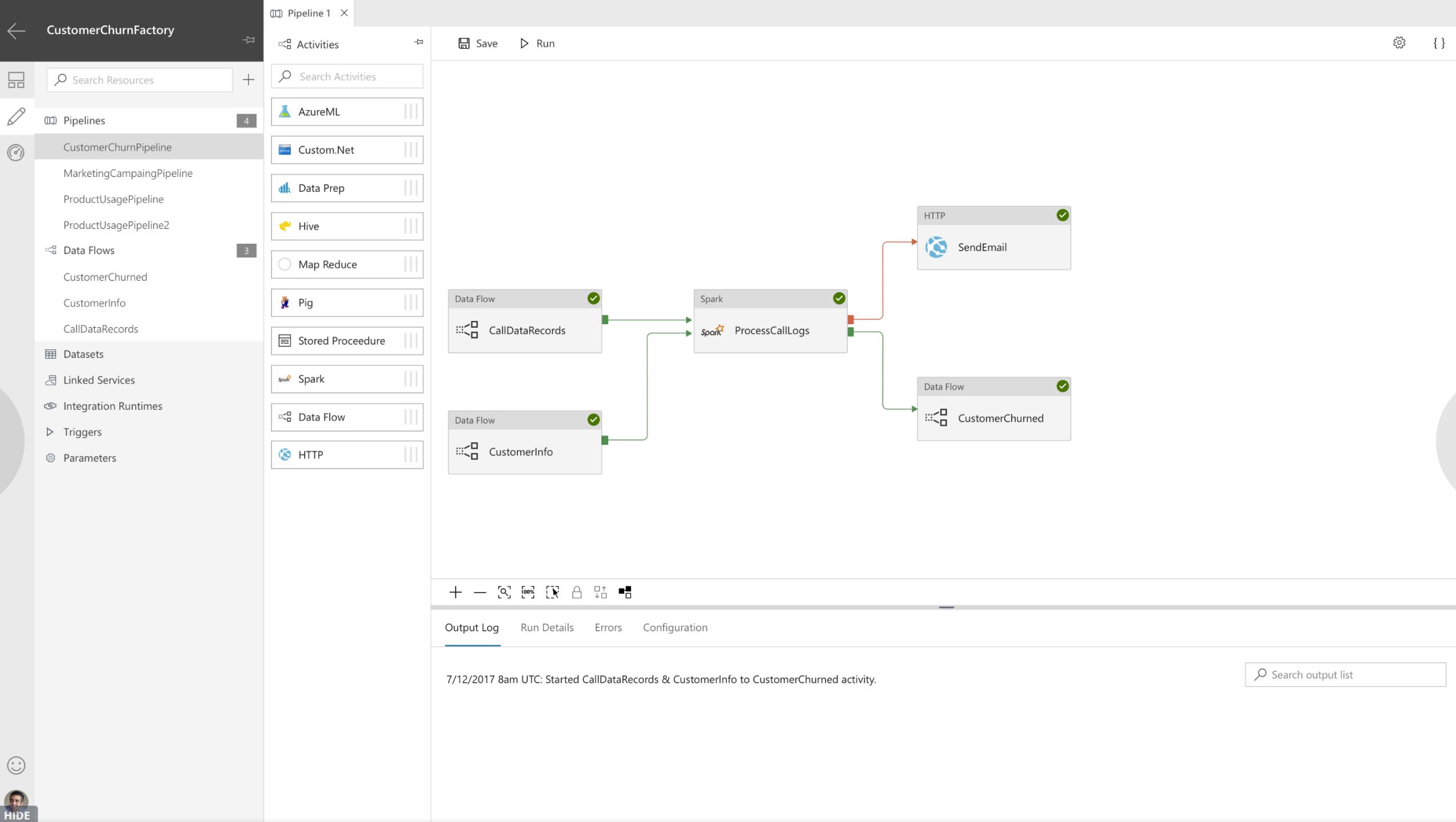
Auto Map

Source fields: 25 / 25 mapped

Sink fields: 25 / 25 mapped

FIELD	TYPE	FIELD	TYPE
Age	int	Age	int
AnnualIncome	BigInt	AnnualIncome	BigInt
CallDropRate	Double	CallDropRate	Double
CallFailureRate	Double	CallFailureRate	Double
CallingNum	String	CallingNum	String
CustomerID	Int	CustomerID	Int
CustomerSuspended	String	CustomerSuspended	String
Education	String	Education	String
Gender	String	Gender	String
HomeOwner	String	HomeOwner	String
MaritalStatus	String	MaritalStatus	String
MonthlyBilledAmount	Int	MonthlyBilledAmount	Int
NoAdditionalLines	Int	NoAdditionalLines	Int

Amazon S3 AWS 3 → Azure Blob Storage AzBlob



Logic Apps

Workflow-based App integration

- Data Analyst
- BI Database
- Identity
- Workday
- MyQ
- Logins
- Data Lake
- Logic App
- Power BI
- Synapse Analytics
- Machine Learning
- Function App
- API Management
- Velo
- ADF

Microsoft Azure



Logic Apps

Cloud APIs and platform functionality

Over 130 built-in connectors

Hosted and managed within the platform

Scales to meet your needs

First class designer experience

Rapid development

API connections

Authenticate once and reuse

Differentiate connection configuration

Simple to deploy

Portal experience for managing API Connections

SaaS

- 10to8
- Act!
- appFigures
- Asana
- Azure Active Directory
- Azure API Management
- Azure App Services
- Azure Automation
- Azure Cognitive Face API
- Azure Cognitive LUIS
- Azure Cognitive Text Analytics
- Azure Cognitive Vision
- Azure Data Lake Store
- Azure Document DB
- Azure Event Hubs
- Azure Functions
- Azure Machine Learning
- Azure Resource Manager
- Azure Service Bus
- Azure SQL
- Azure Storage Blob
- Azure Storage Queues
- Basecamp2&3
- Benchmark Email
- Bing Search
- BitBucket
- Bitly
- Blogger
- Box
- Buffer
- Campfire
- CapsuleCRM
- Chatter
- Common Data Service
- Disqus
- DocuSign
- Dropbox

- Dynamics AX Online
- Dynamics CRM Online
- Dynamics CRM Service Bus
- Dynamics Financials
- Dynamics Operations
- Easy Redmine
- Eventbrite
- Facebook
- FreshBooks
- Freshdesk
- GitHub
- Gmail
- Google Calendar
- Google Contacts
- Google Drive
- Google Sheets
- Google Tasks
- GoTo Meeting
- GoTo Training
- GoTo Webinar
- Harvest
- HelloSign
- HipChat
- Infusionsoft
- Inoreader
- Insightly
- Instagram
- Instapaper
- Intercom
- JIRA
- LeanKit
- LiveChat
- MailChimp
- Mandrill
- Microsoft Translator
- Medium
- MSN Weather
- Muhimbi PDF

- Nexmo
- Office 365
- Office 365 Users
- Office 365 Video
- OneDrive
- OneDrive for Business
- OneNote
- Outlook.com
- Outlook Customer Manager
- Outlook Tasks
- Paylocity
- PagerDuty
- Pinterest
- Pipedrive
- Pivotal Tracker
- Power BI
- Project Online
- Redmine
- Salesforce
- Salesforce Chatter
- SendGrid
- SharePoint Online
- Slack
- SmartSheet
- SparkPost
- Stripe
- Survey Monkey
- Teradata
- Todoist
- Toodledo
- Trello
- Twilio
- Twitter
- Typeform
- UserVoice
- Vimeo
- VS Team Services
- Webmerge

XML & EDI

- XML Validation
- Transform XML (+Mapper)
- Flat File Encode
- Flat File Decode
- X12
- EDIFACT
- AS2
- Int Account Artifact Lookup

Protocols/Native

- HTTP, HTTPS
- HTTP Webhook
- FTP, SFTP
- SMTP
- RSS
- Compose, Parse JSON
- Query, Join, Table, Select
- Schedule, Wait
- Terminate
- Workflow

Hybrid

- BizTalk Server
- File System
- IBM DB2
- Informix
- Oracle DB
- SharePoint Server
- SQL Server
- SAP
- Websphere MQ

Logic Apps Workflow Designer

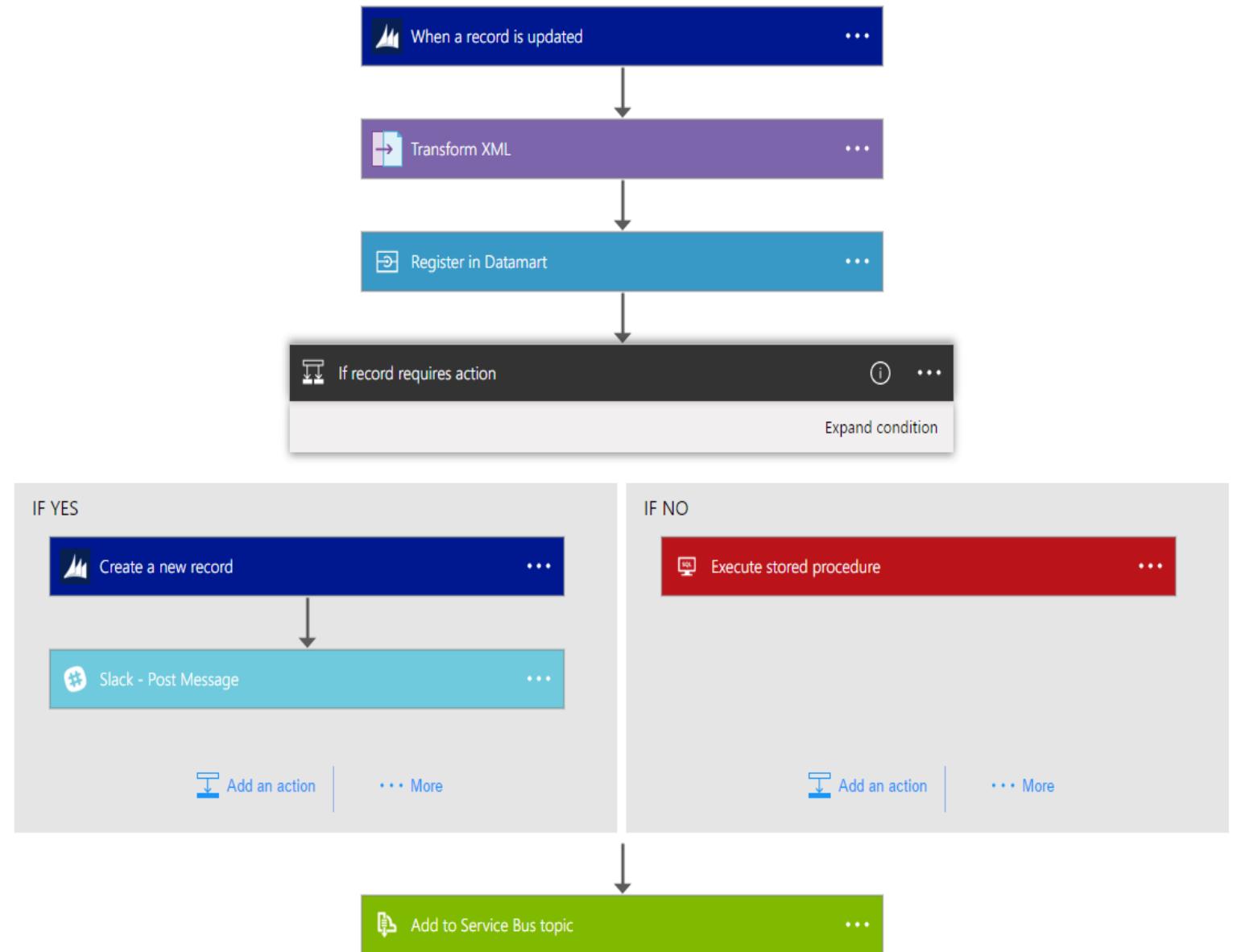
Workflow in the cloud

Powerful control flow

Connect disparate
applications

No code designer for
rapid creation

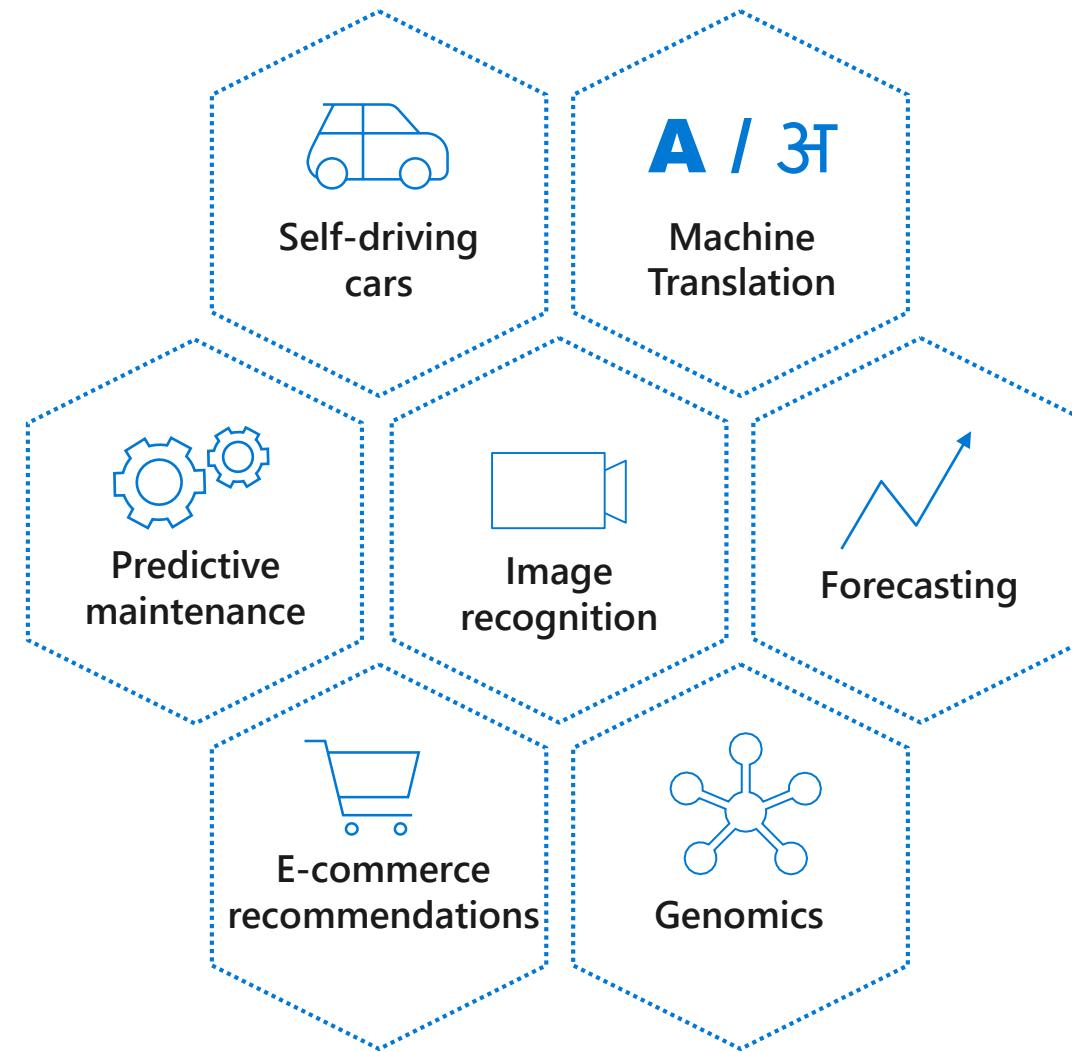
Also works within Visual
Studio for added CI/CD



Advanced Analytics + Machine Learning

Machine Learning

Applications



Machine Learning on Azure

Domain specific pretrained models

To simplify solution development



Vision



Speech



Language



Search

Familiar Data Science tools

To simplify model development



Visual Studio Code



Azure Notebooks



Jupyter



Command line

Popular frameworks

To build advanced deep learning solutions



PyTorch



TensorFlow



Scikit-Learn



ONNX

Productive services

To empower data science and development teams



Azure
Databricks



Azure
Machine
Learning



Machine
Learning VMs

Powerful infrastructure

To accelerate deep learning



CPU



GPU

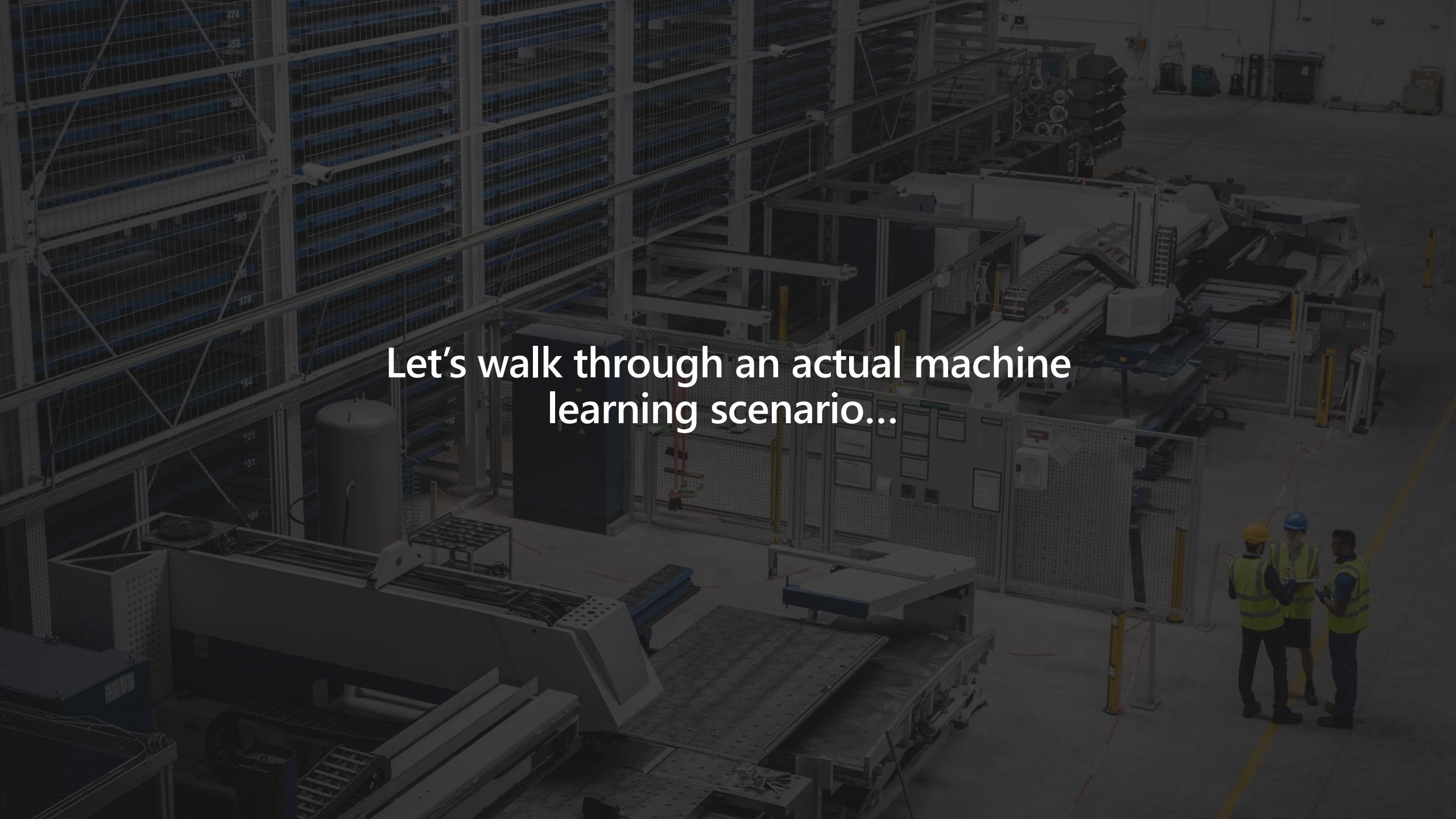


FPGA



From the Intelligent Cloud to the Intelligent Edge





Let's walk through an actual machine learning scenario...

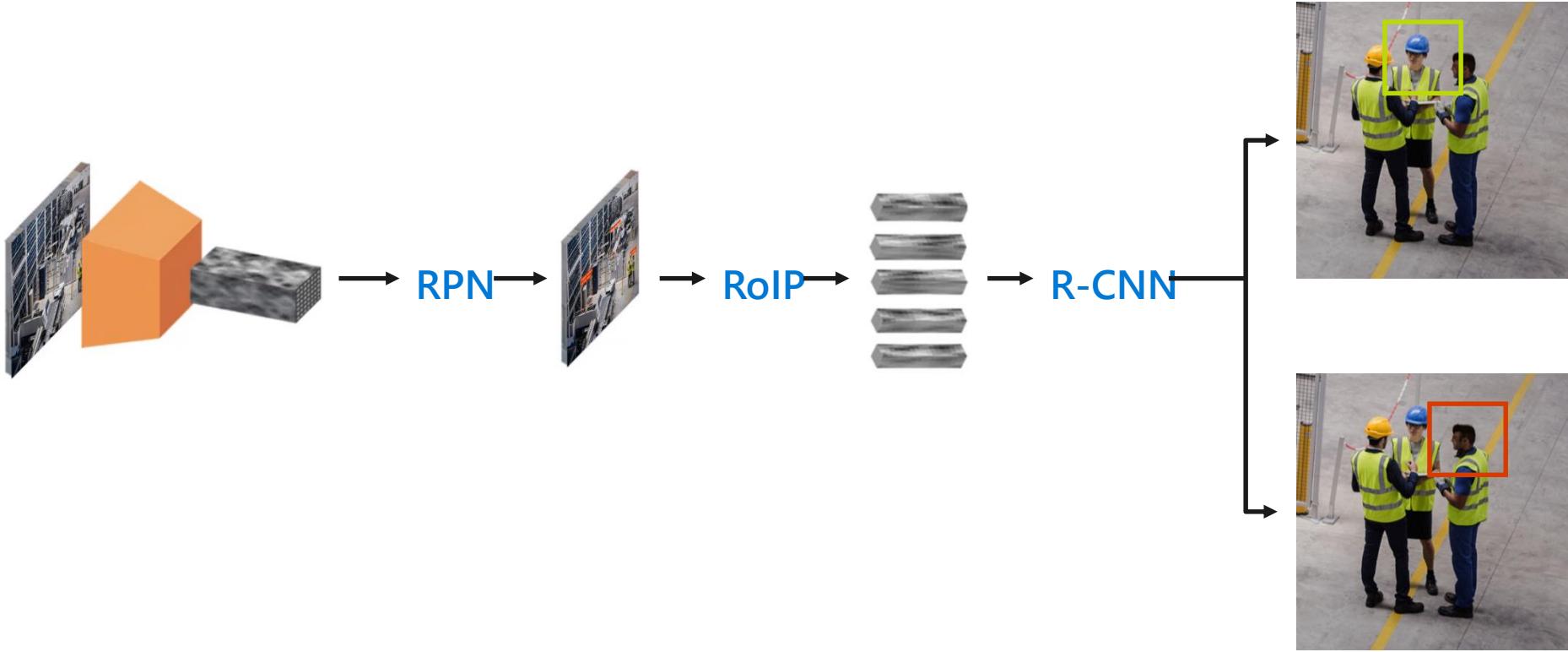






No hardhat

Next challenge is to build a model



Time-consuming



Needs specialized knowledge



Complex

Sophisticated pretrained models

Infuse apps with powerful, pre-trained AI models

Customize easily and tailor to your needs



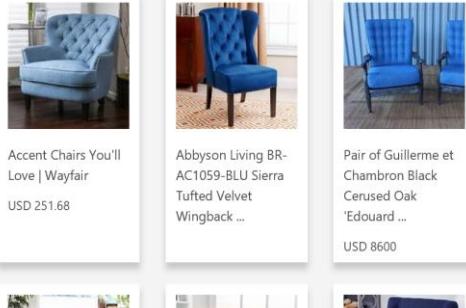
Vision



Computer Vision | Video Indexer | Face | Content Moderator



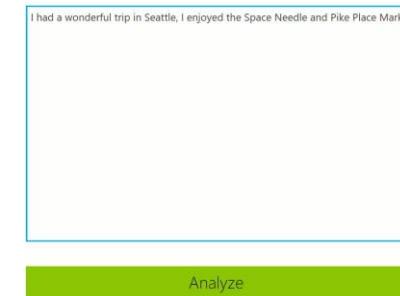
Bing
Search



Big Web Search | Video Search | Image Search | Visual Search | Entity Search |
News Search | Autosuggest



Language



Text Analytics | Spell Check | Language Understanding | Text Translation | QnA Maker

Analyzed text JSON

LANGUAGES: English (confidence: 100 %)

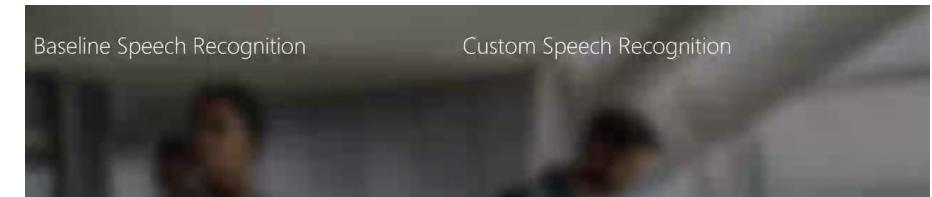
KEY PHRASES:

SENTIMENT: 73 %

LINKED ENTITIES (PREVIEW): a



Speech



Speech to Text | Text to Speech | Speech Translation | Speaker Recognition

Demo:
AI For Earth

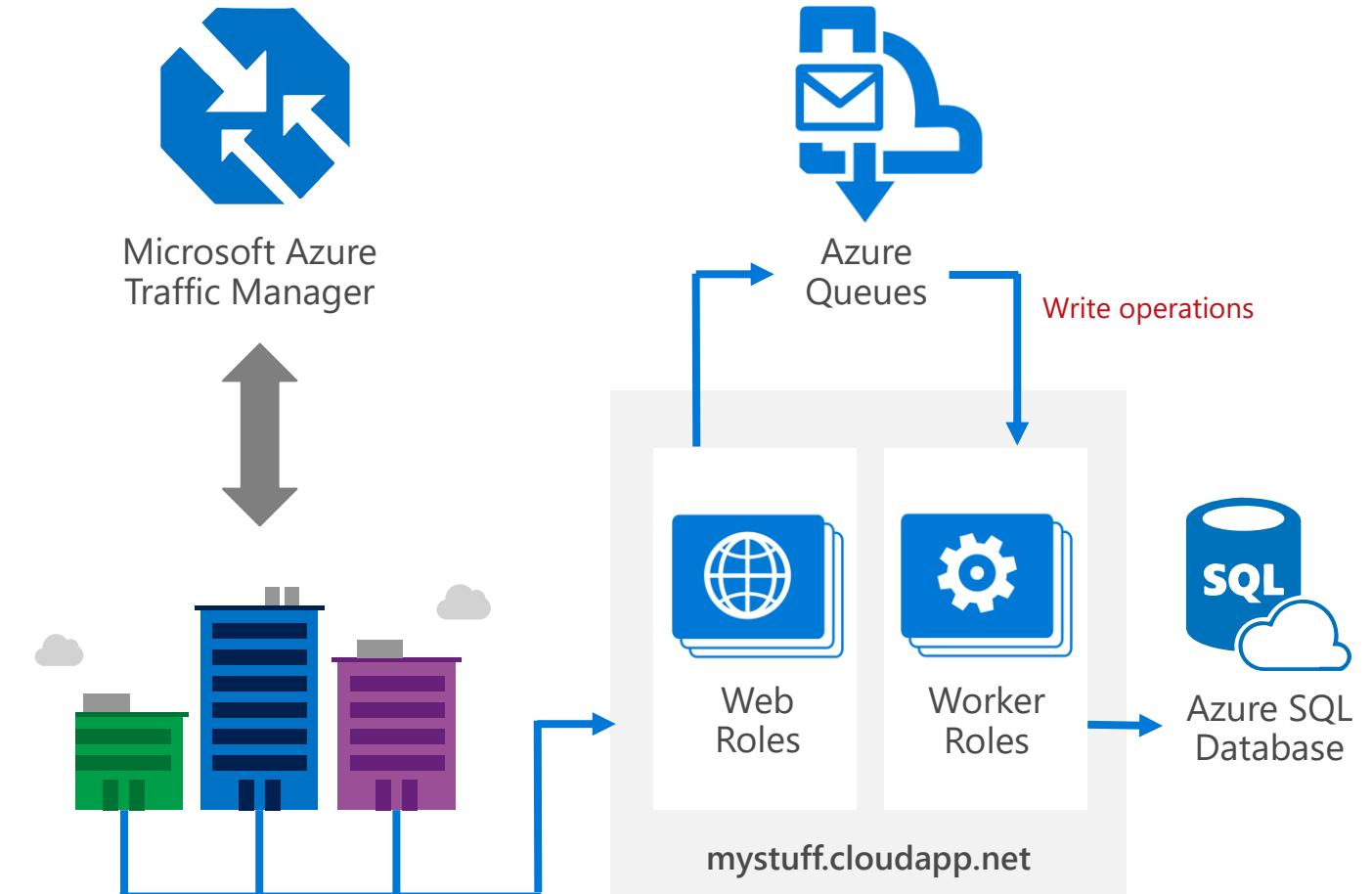
Combine Azure Services to Solutions – Data Patterns

Sample cloud app architecture

Classic 3-tier enterprise architecture

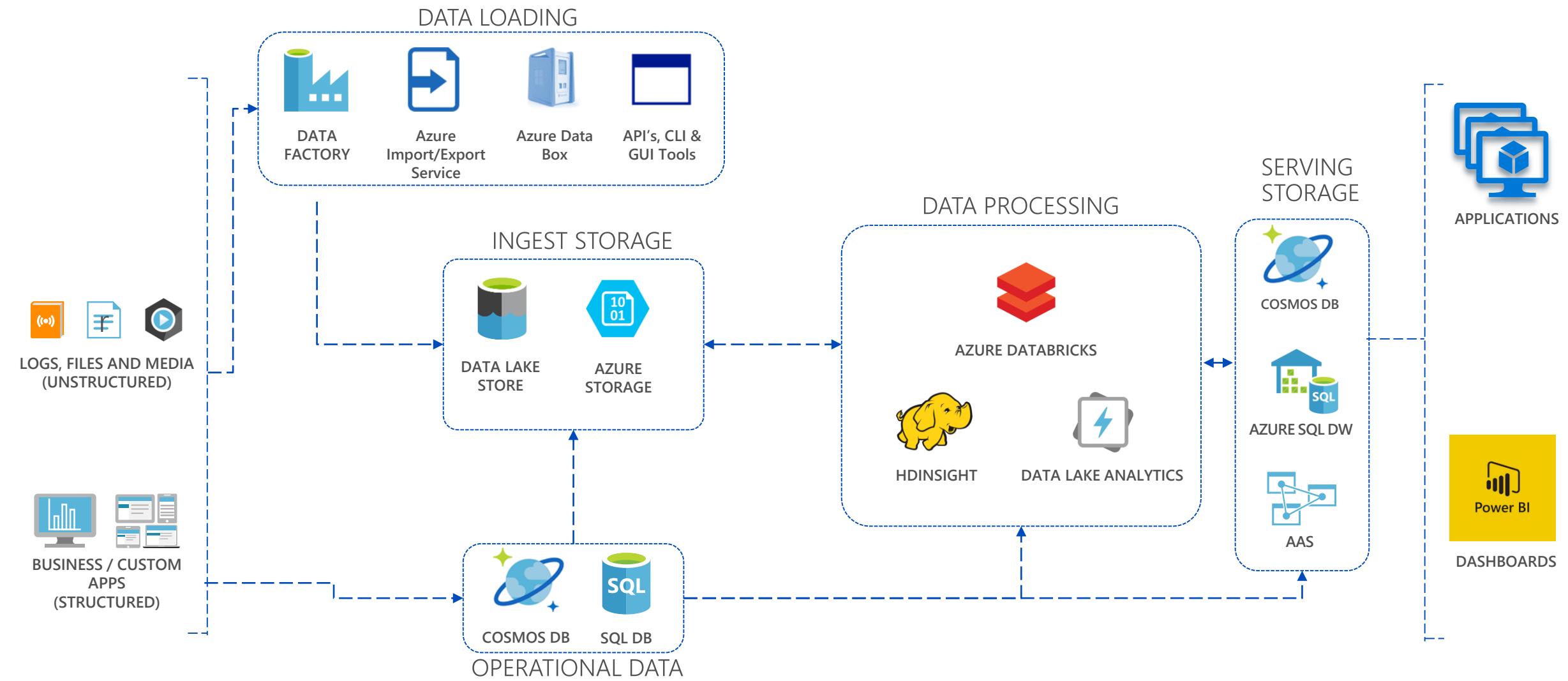
Required to scale to 10,000s of users and process terabytes of relational data

Scaling out (and in) web and worker tiers is relatively easy

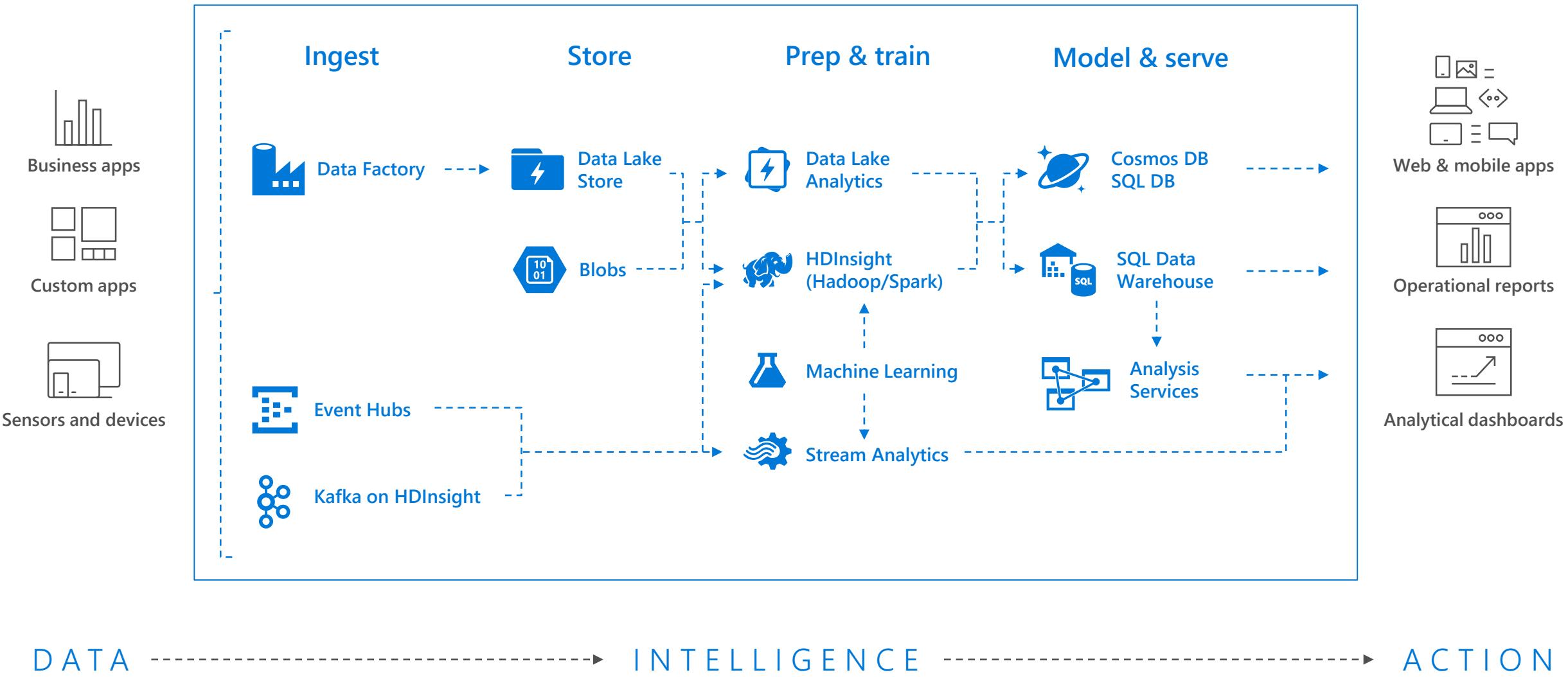


DATA WAREHOUSING PATTERN IN AZURE

Loading and preparing data for analysis with a data warehouse



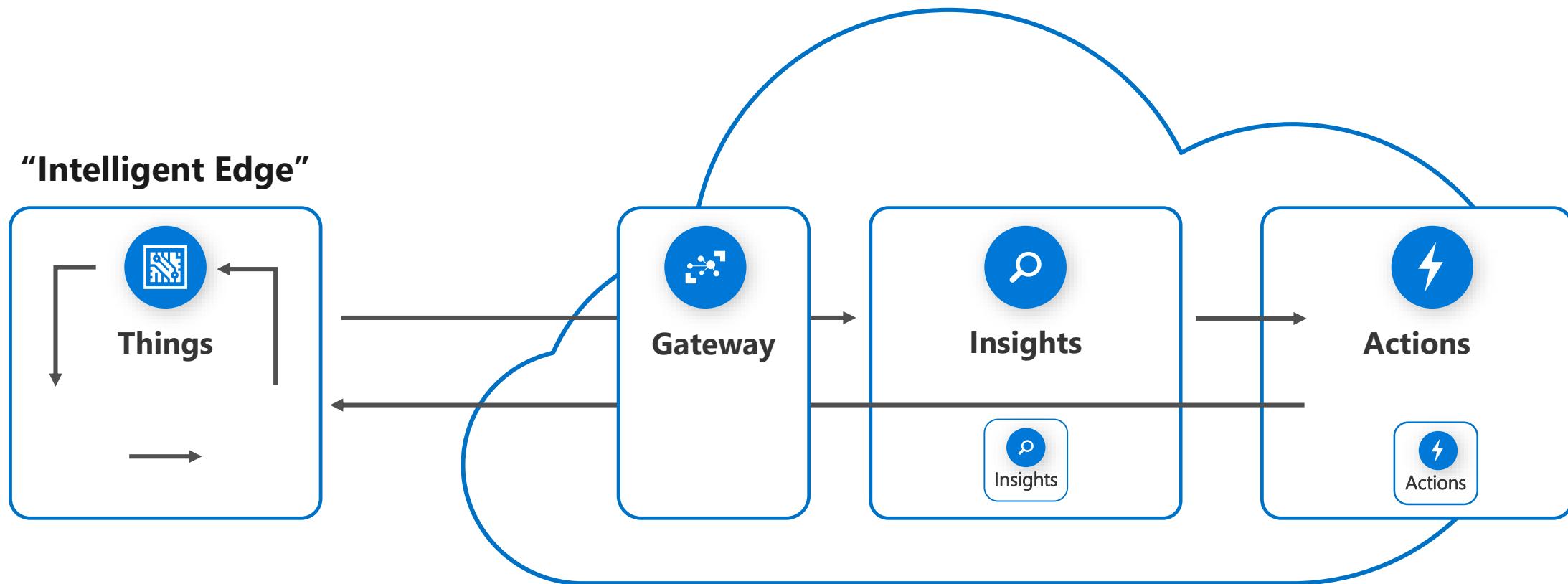
BIG DATA & ADVANCED ANALYTICS



(Industrial) IoT & Analytics

IoT Basic Concept

From Things To Actions



Benefit from a comprehensive solution

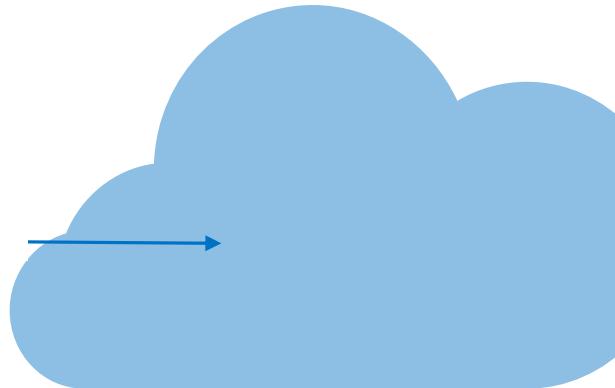
Securely connect and scale with efficiency

Analyze and act on new data

Integrate and transform business processes



10101010110100011010001



10101001110100101



Real-time operating systems

IoT Edge

Analytics

Device Registry

Rules and Actions

Dashboards & Visualization

SAP

ORACLE®

IBM DB2

Informix software

Microsoft Dynamics

salesforce

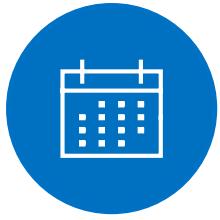
WebSphere

Office 365

However, IoT projects are complex



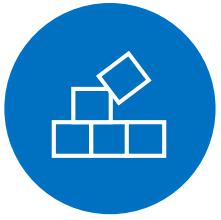
Security
is a **challenge**



Time-consuming
to get started



Incompatible with
existing infrastructure



Challenging
to scale

Cloud



Globally available,
unlimited compute
resources

IoT



Harnessing signals from
sensors and devices,
managed centrally
by the cloud

AI



Breakthrough intelligence
capabilities, in the cloud
and at the edge

Intelligent Edge



Intelligence offloaded from
the cloud to IoT devices

Azure IoT & Analytics capabilities



Device Connectivity & Management



Data Ingestion and Command & Control



Stream Processing & Predictive Analytics



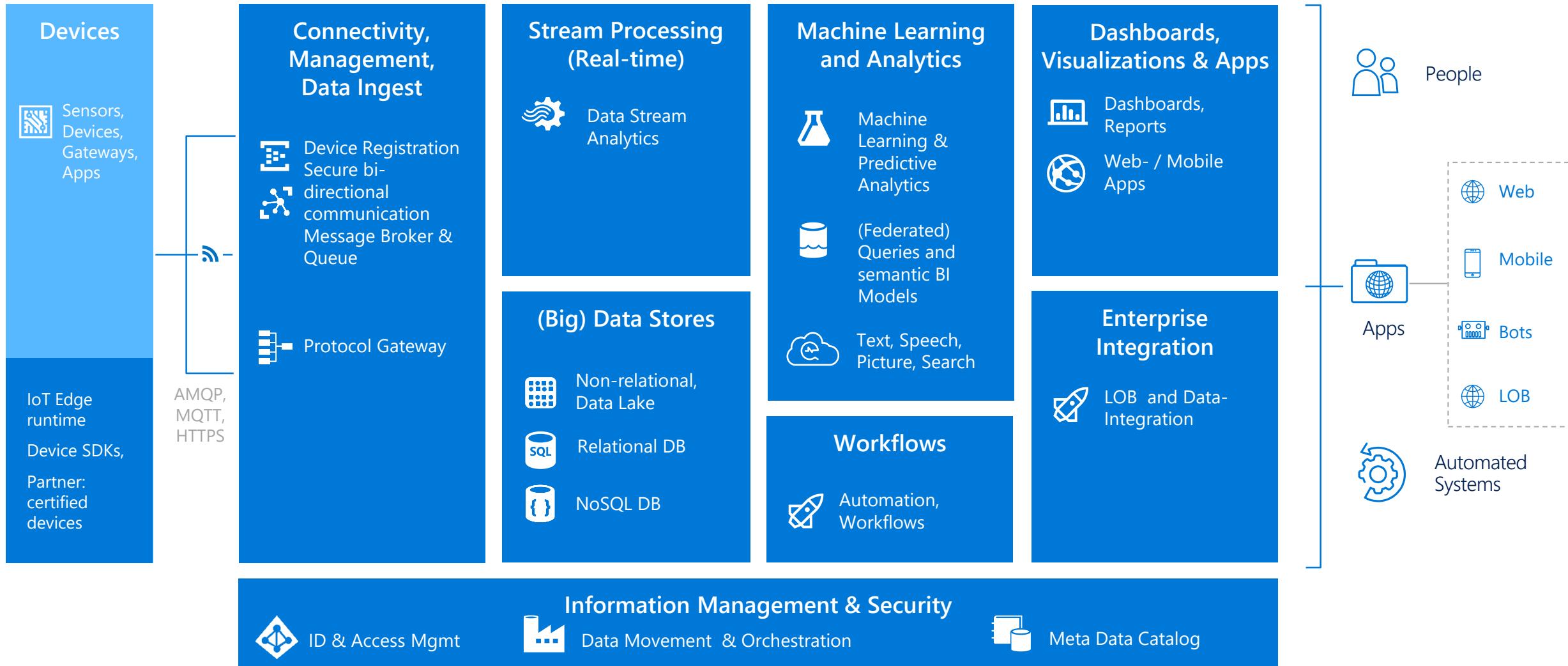
Workflow Automation and Enterprise Integration



Dashboards, Visualization and Apps



IoT and Analytics building blocks



Microsoft is simplifying IoT

Azure IoT Solution Accelerators

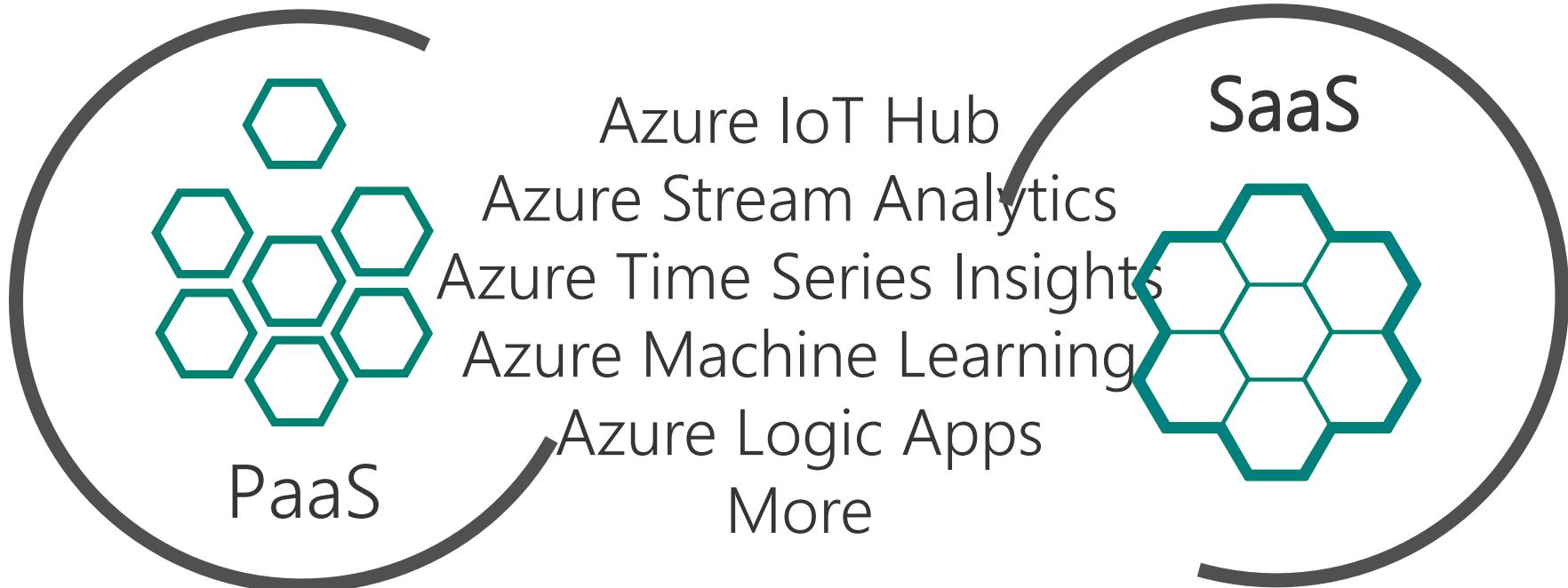
Preconfigured solutions for common IoT scenarios



Remote Monitoring | Predictive Maintenance | Connected Factory

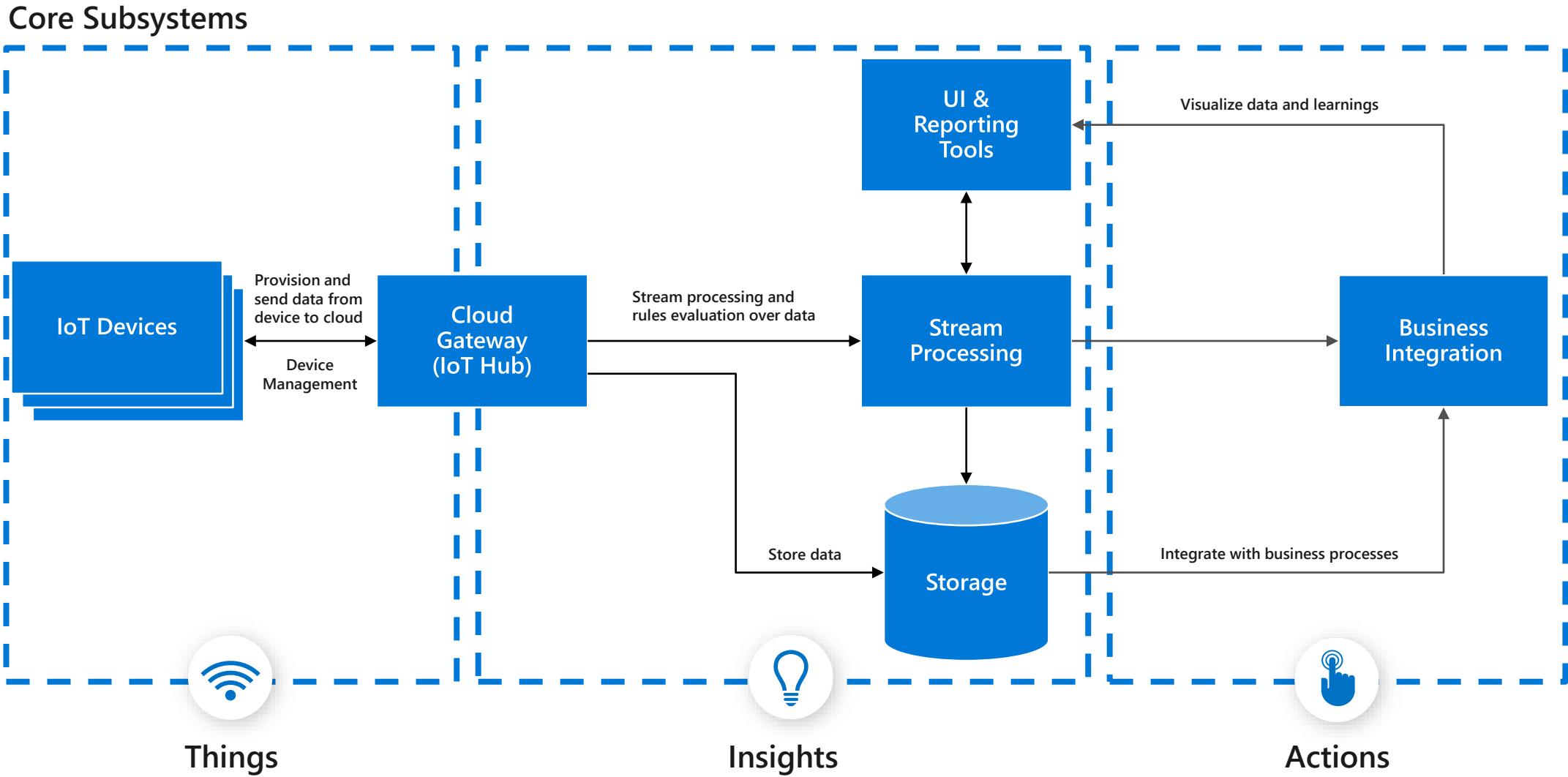
Microsoft IoT Central

Fully managed IoT SaaS
No cloud solution expertise required

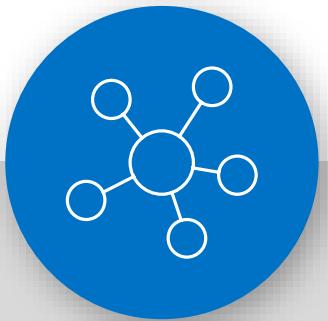


Demo: Microsoft IoT Central

Azure IoT reference architecture



Azure core IoT portfolio



Solutions

Azure IoT Central

Azure IoT Solution Accelerators

Azure IoT Remote Monitoring

Azure IoT Predictive Maintenance

Azure IoT Device Simulation

Azure IoT Connected Factory



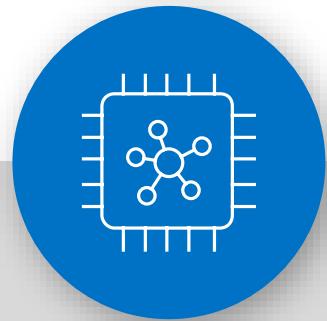
Services

Azure Digital Twins

Azure IoT Hub

Azure IoT Security

Azure Time Series Insights



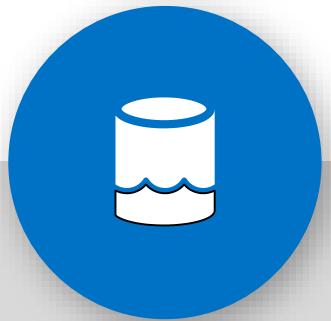
Edge

Azure IoT Edge

Azure Sphere

Windows IoT

Azure Data & AI Services



Storage

Azure Data Lake

Azure Cosmos DB (NoSQL)

Azure SQL Database

Azure SQL Data Warehouse

Azure Database for PostgreSQL,
Maria DB, MySQL



Analytics

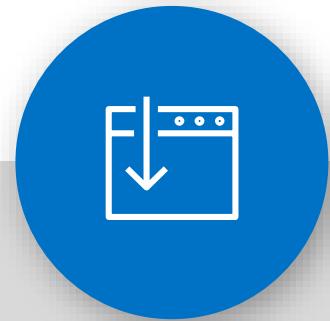
Azure Stream Analytics

Azure Databricks (Spark)

Azure HDInsight (Hadoop)

Azure Machine Learning

Cognitive Services



Integration & Apps

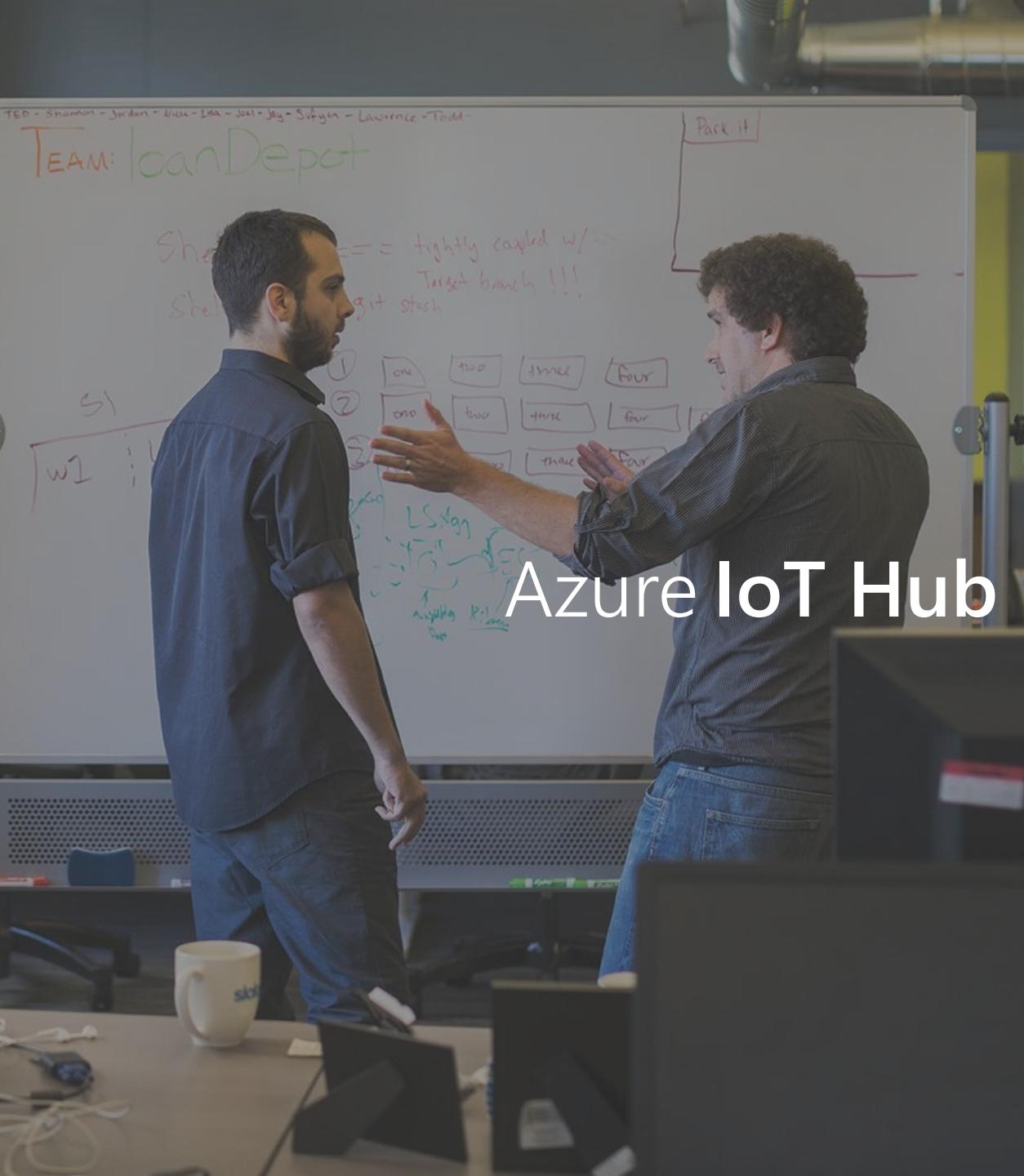
Azure Logic Apps

Azure Data Factory

Azure API Apps

Azure Web/Mobile Apps

Power BI



Azure IoT Hub



Establish bi-directional communication with billions of IoT devices



Enhance security with per device authentication



Provision devices at scale w/IoT Hub Device Provisioning Service

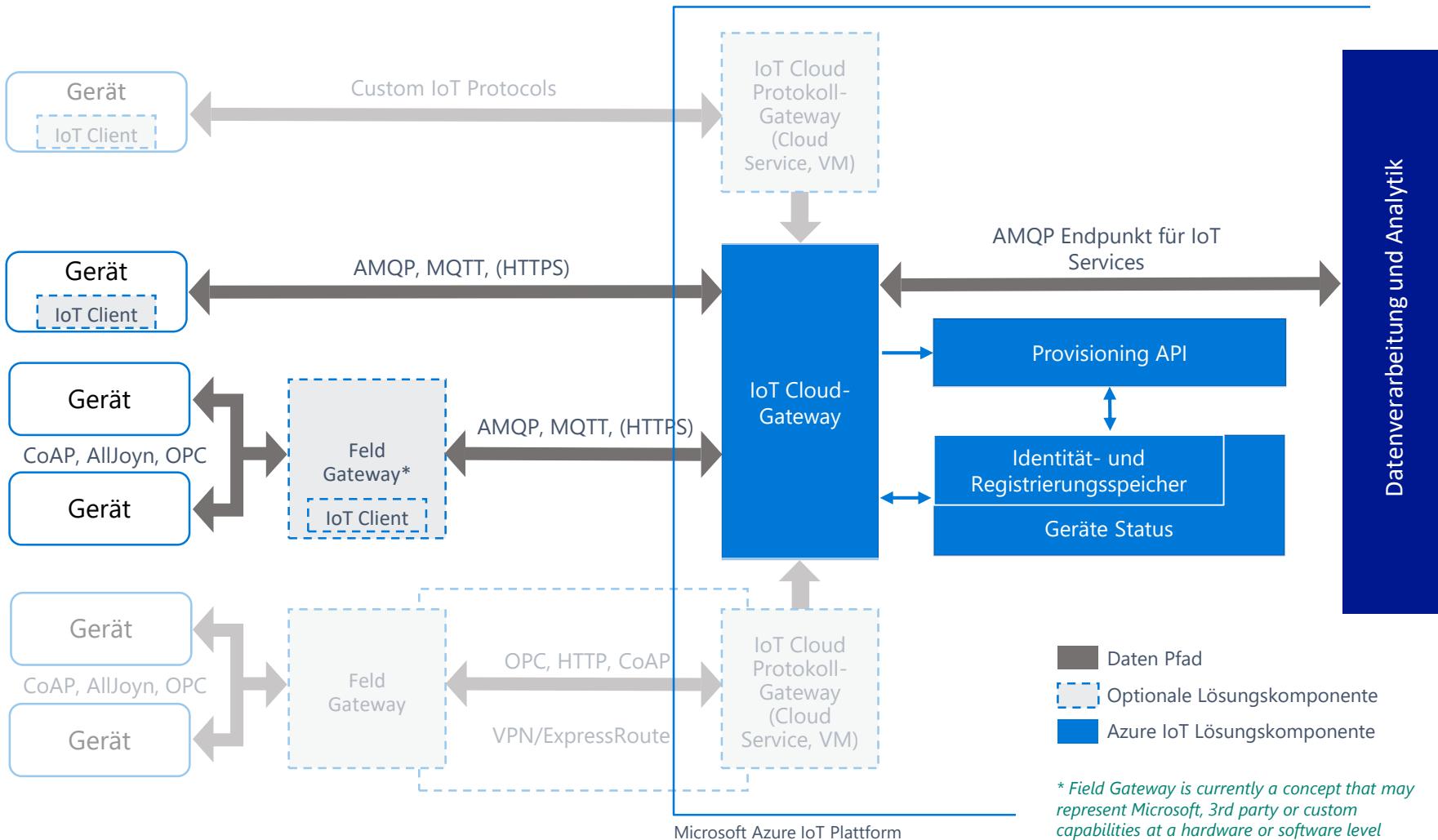


Manage devices at scale with device management



Multi-language and open source SDKs

IoT: Gerätekonnektivität und -Verwaltung



Azure IoT Hub



Bi-directional communication

Millions of devices
Multi-language, open source SDKs
HTTPS/AMQPS/MQTT-S
Send Telemetry
Receive commands
Device management
Device Twins
Queries and jobs



Enterprise scale and integration

Billions of messages
Scale up and down
Declarative message routes
File Upload
WebSockets and Multiplexing
Azure Monitor
Azure Resource Health
Configuration management



End-to-end security

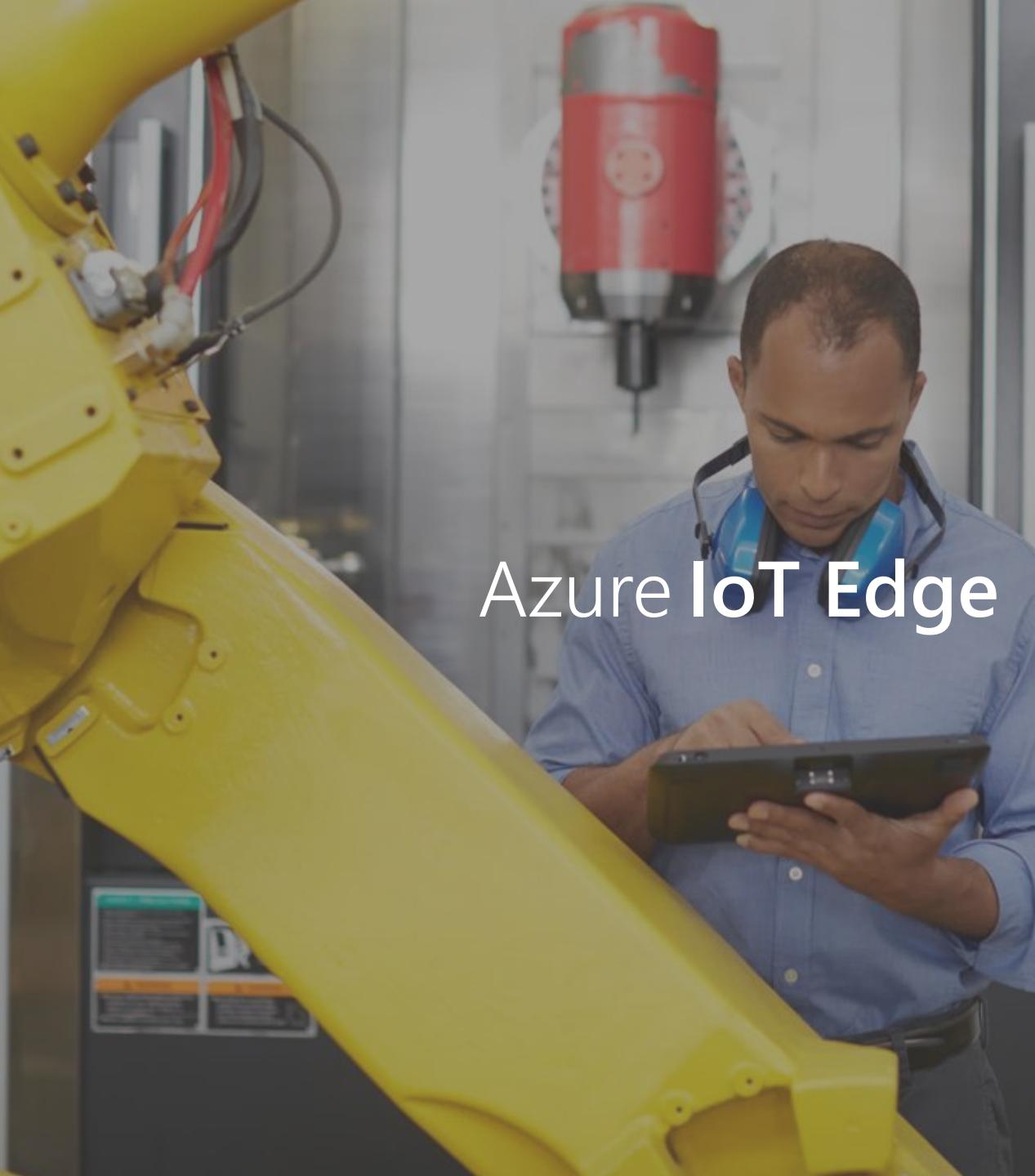
Per device certificates
Per device enable/disable
TLS security
X.509 support
IP whitelisting/blacklisting
Shared access policies
Firmware/software updates
Azure security center support

Azure IoT Hub Device Provisioning Service



IoT-scale automated provisioning

Zero-touch provisioning
Centralize your provisioning workflow
Load balance across multiple IoT Hubs
Re-provisioning support
Supports TPM + X.509



Azure IoT Edge



Move cloud and custom workloads to the edge, securely



Seamless deployment of AI and advanced analytics



Configure, update and monitor from the cloud



Compatible with popular operating systems

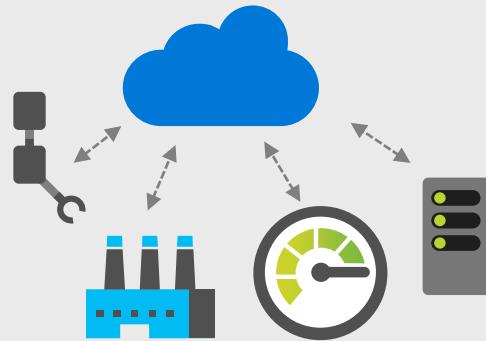


Code symmetry between cloud and edge for easy development and testing



Secure solution from chipset to cloud

IoT on the edge



IoT in the Cloud

Remote monitoring and control

Merging remote data from across multiple IoT devices

Near infinite compute and storage to train machine learning and other advanced AI tools



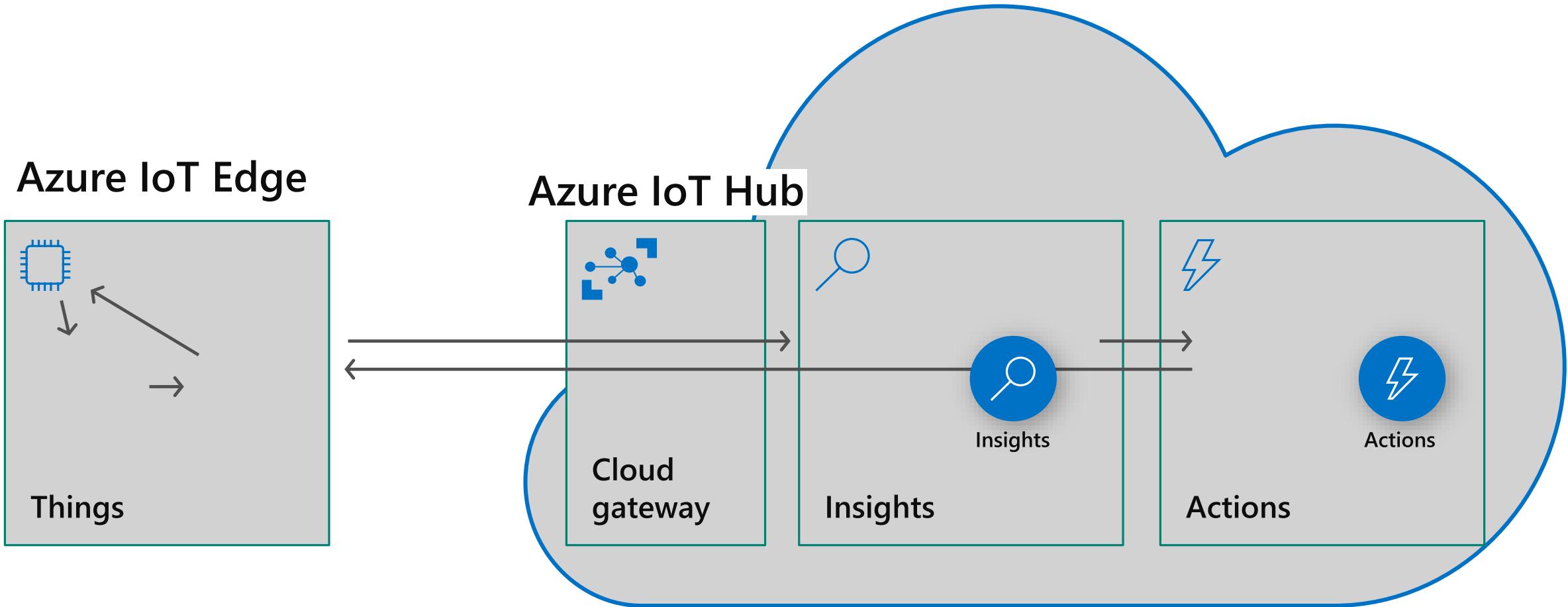
IoT on the Edge

Low latency tight control loops require near real-time response

Public internet inherently unpredictable

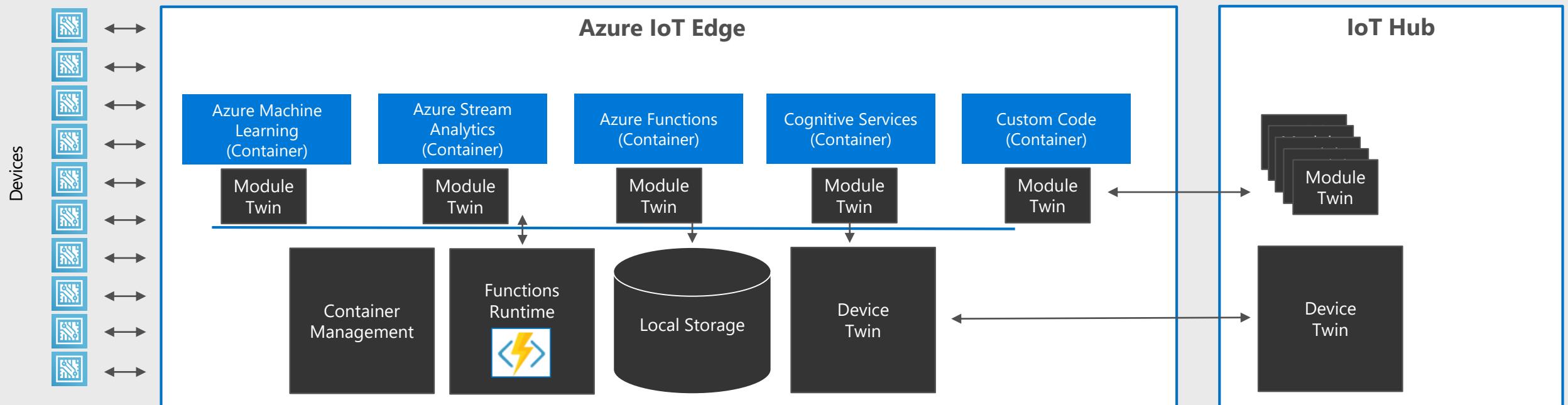
Privacy of data and protection of IP

IoT pattern + edge



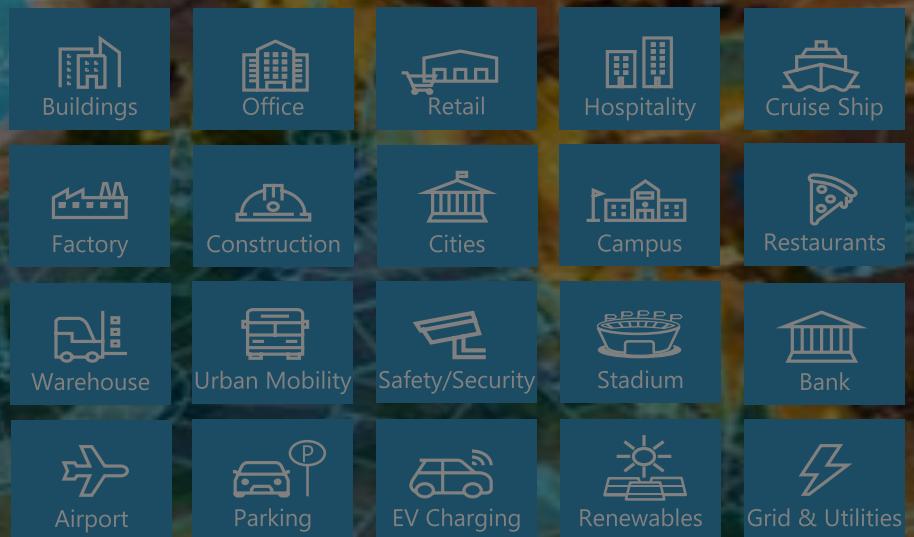
Azure IoT Edge

- Container based modules
 - Azure Functions
 - Azure Stream Analytics
 - Azure Machine Learning
 - Cognitive Services
- Offline / Synchronized Device Twins
 - Local Storage
 - Cloud Management & Deployment
 - High Availability / Fault Tolerance
 - Cloud Dev/Test Support



Public preview

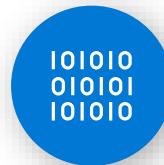
Azure Digital Twins



Build next generation IoT solutions with Azure Digital Twins



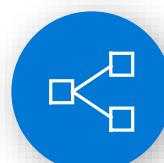
Virtually represent the physical world with a digital twin that models the relationships between people, places and devices



Leverage predefined and extensible Twin Object Models to build contextually-aware solutions uniquely attuned to your industry domain

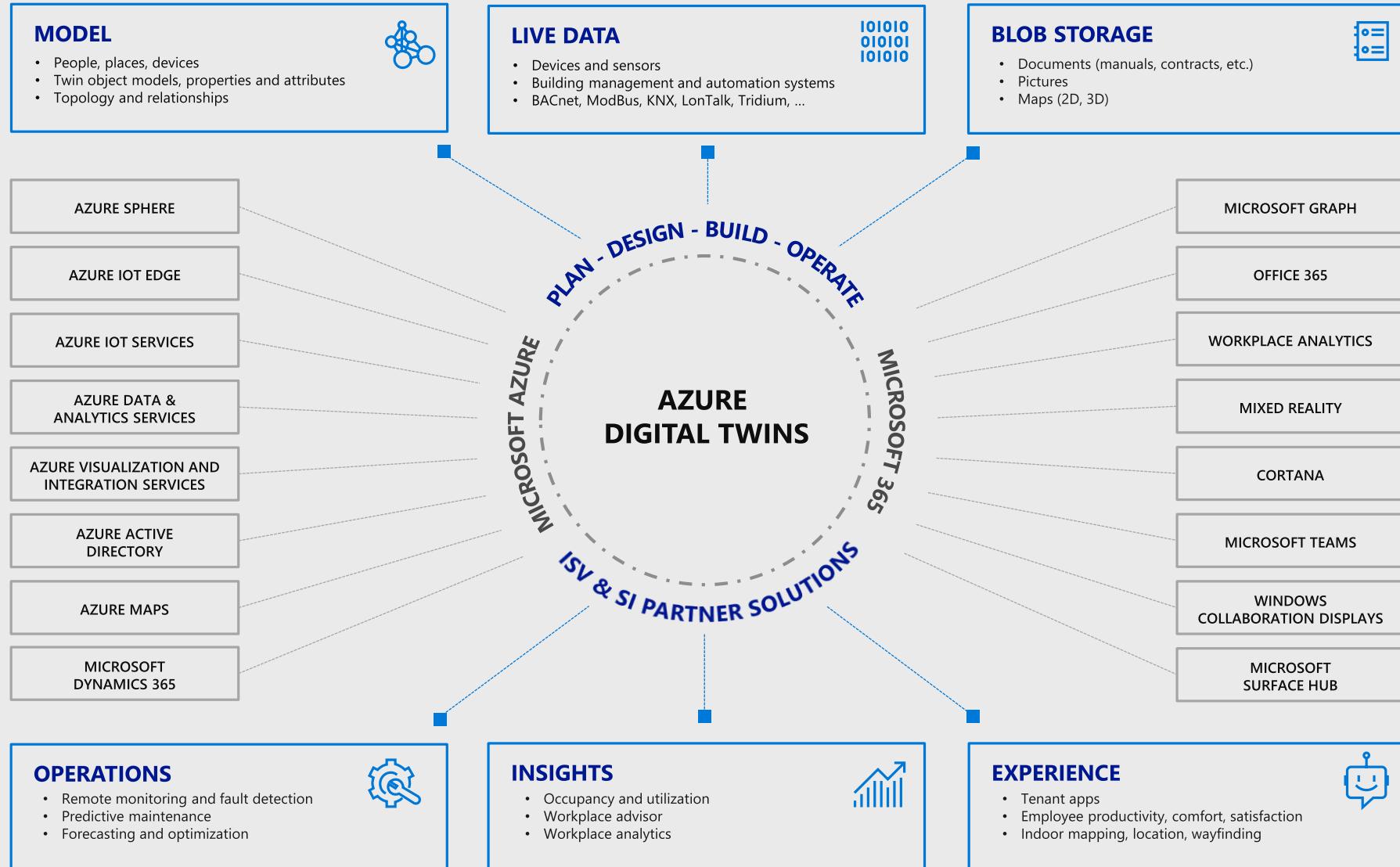


Automate actions in a space with custom functions that send events and/or notifications to endpoints based on incoming telemetry



Securely replicate solutions across multiple tenants through built-in multi and nested-tenancy

Azure Digital Twins for Smart Spaces



A woman with blonde hair, wearing a flight suit and glasses, is looking at a tablet device. The tablet screen displays a complex interface with various charts, graphs, and data tables, likely related to aircraft maintenance or performance monitoring. The background shows the interior of a cockpit with numerous control panels, switches, and gauges.

Azure IoT solution accelerators

-  End-to-end implementation
-  Completely customizable
-  Open-source microservices based architecture
-  Device connectivity and management
-  Dashboards, visualization, and insights
-  Workflow automation and integration
-  Command and control
-  Preconfigured solutions
-  Remote monitoring
-  Connected factory
-  Predictive maintenance
-  Device simulation

Demo: IoT Solution Accelerators

Azure Hands-on Labs in Wien

<https://aka.ms/azurelabs-vienna>



© 2019 Microsoft Corporation. All rights reserved. Microsoft, Windows, and other product names are or may be registered trademarks and/or trademarks in the U.S. and/or other countries.
The information herein is for informational purposes only and represents the current view of Microsoft Corporation as of the date of this presentation. Because Microsoft must respond to changing market conditions, it should not be interpreted to be a commitment on the part of Microsoft, and Microsoft cannot guarantee the accuracy of any information provided after the date of this presentation. MICROSOFT MAKES NO WARRANTIES, EXPRESS, IMPLIED OR STATUTORY, AS TO THE INFORMATION IN THIS PRESENTATION.