GIGASPACES
innovate with confidence

Converge Transactional and Predictive Analytics to Effectively Scale IoT

2018

11 billion

Is the number of connected devices…

Not including phones and computers

How can we leverage this information for context-sensitive insights that answer specific questions about specific things at the right point in time?

About GigaSpaces

We provide the leading in-memory computing platforms for real-time insight to action and extreme transactional processing. With GigaSpaces, enterprises can operationalize machine learning and transactional processing to gain real-time insights on their data and act upon them in the moment.

INSIGHTEDGE

InsightEdge is an in-memory real-time analytics platform for instant insights to action; analyzing data as it's born, enriching it with historical context, for smarter, faster decisions

XAP

In-Memory Computing Platform for microsecond scale transactional processing, data scalability, and powerful event-driven workflows

300+
Direct customers

50+ / 500+
Fortune / Organizations

5,000+
Large installations in production (OEM)

25+
ISVs

Our Customers Span Across Multiple Industries

FINANCIAL SERVICES

TELCO

RETAIL

TRANSPORTATION

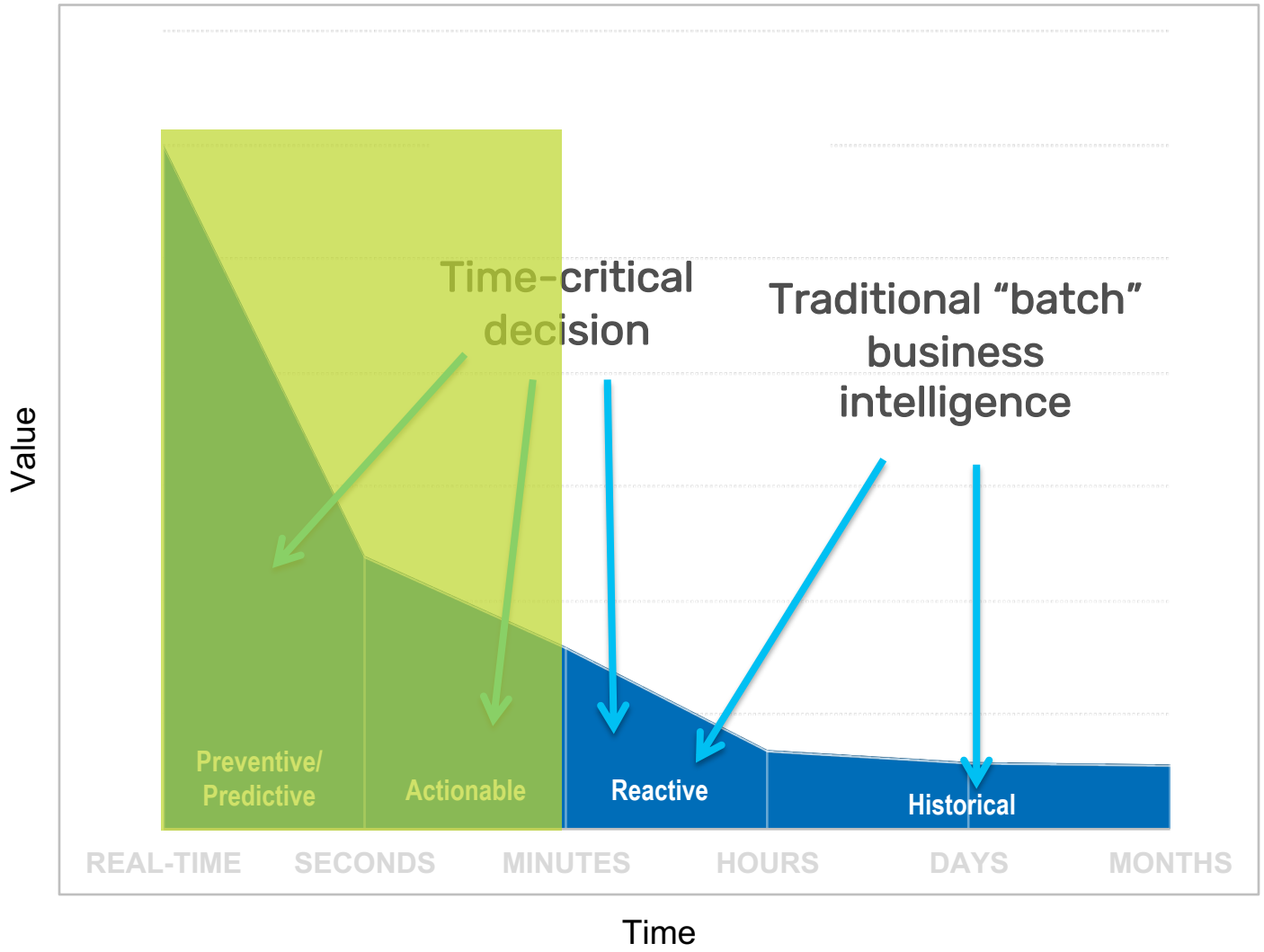
INSURANCE

Select Customers

OEMs / ISVs / Partners


74%
want to be
data driven
only 23%
are successful,
FORRESTER®

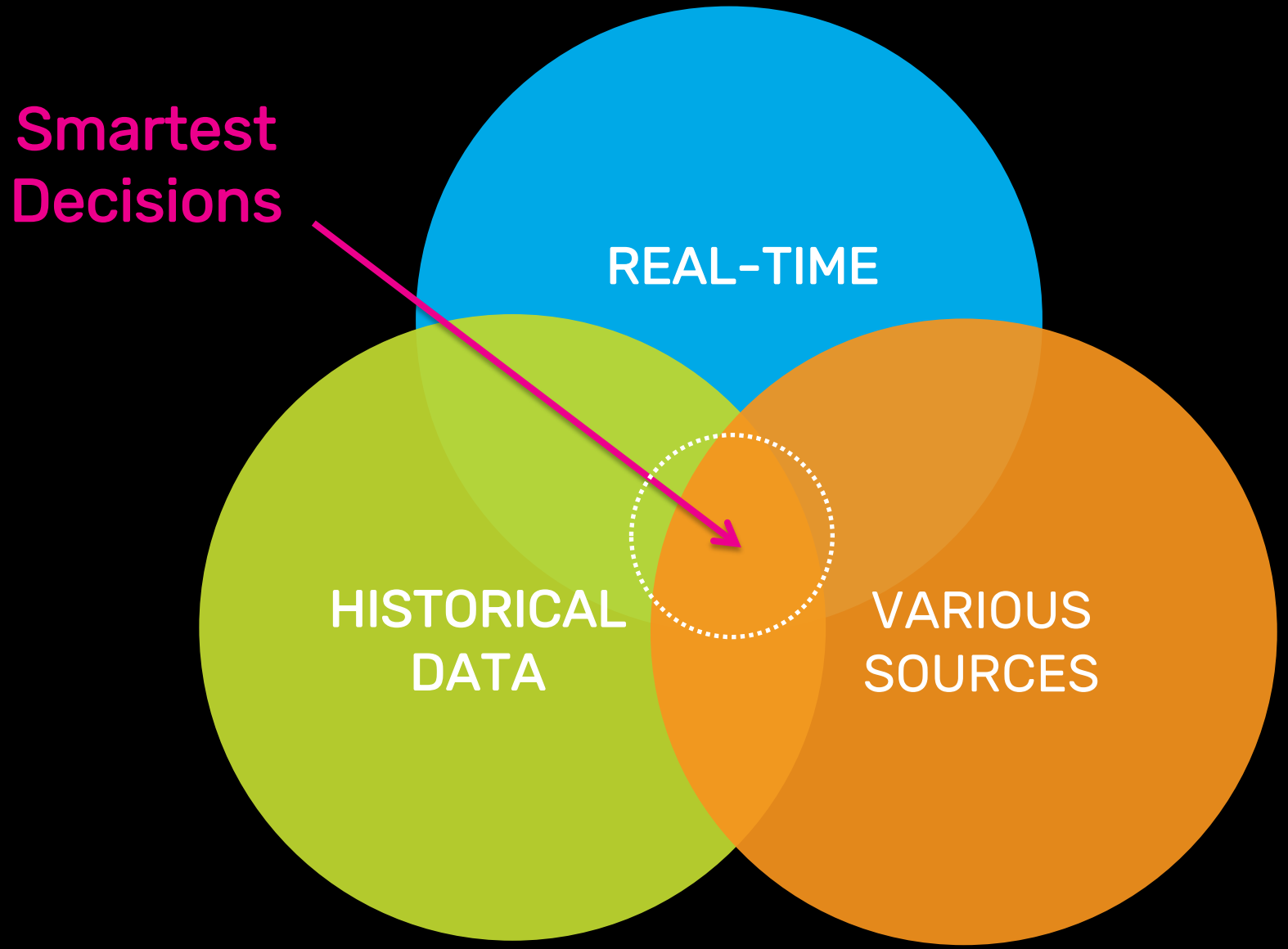
Value of Data to Decision

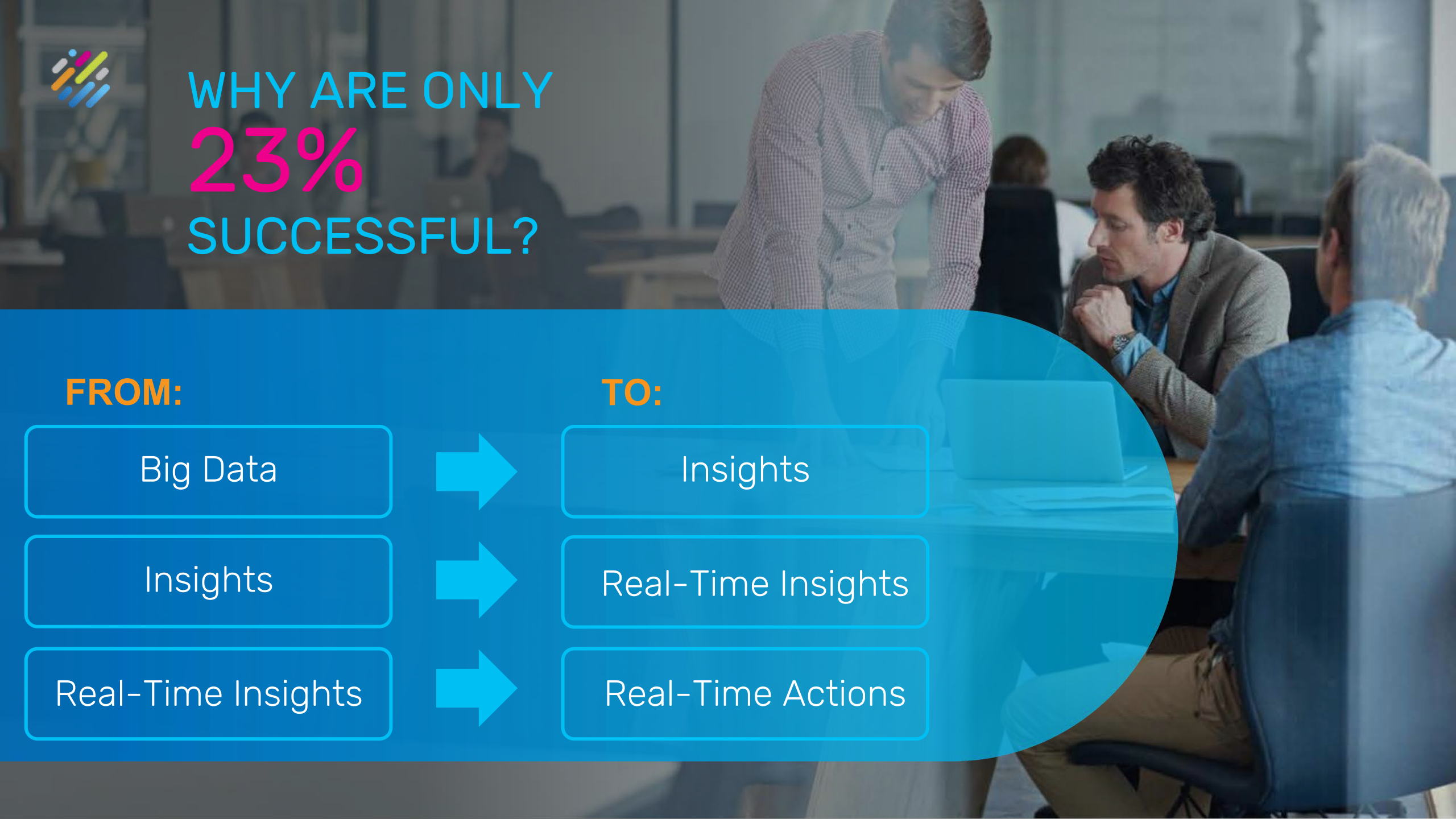


Real-time data is highly valuable if you act on it on time

Old + real-time data is more valuable if you have the means to combine them

Real-Time Applications





WHY ARE ONLY 23% SUCCESSFUL?

FROM:	TO:
Big Data	Insights
Insights	Real-Time Insights
Real-Time Insights	Real-Time Actions

Challenges of Internet of Things Application Enablement

Foundations of an IoT Analytics Platform

CENTER + EDGE LOW-LATENCY PROCESSING

Focus on low latency and event-driven, rather than high-throughput

BATCH + REAL-TIME DATA CONVERGENCE ON DATA FROM VARIOUS SOURCES

IoT data ingestion is heterogeneous: streaming, micro-batch, and batch on multiple data types

CLOSED LOOP ANALYTICS (INSIGHTS TRIGGER WORKFLOWS)

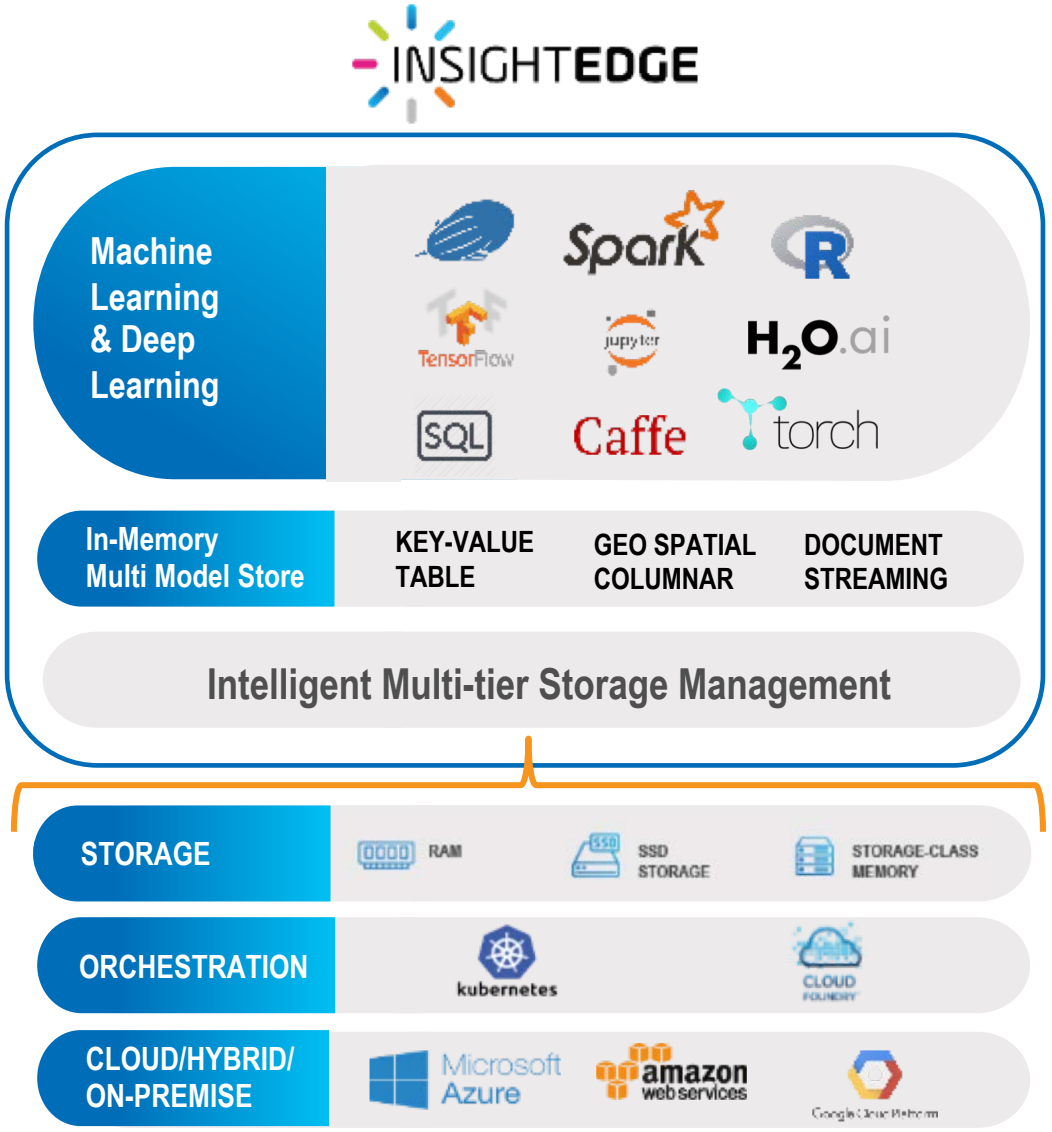
From sensors to actuators = from insights to action (at low latency)

MULTI-TENANT, GEO-FEDERATED, SCALE-OUT

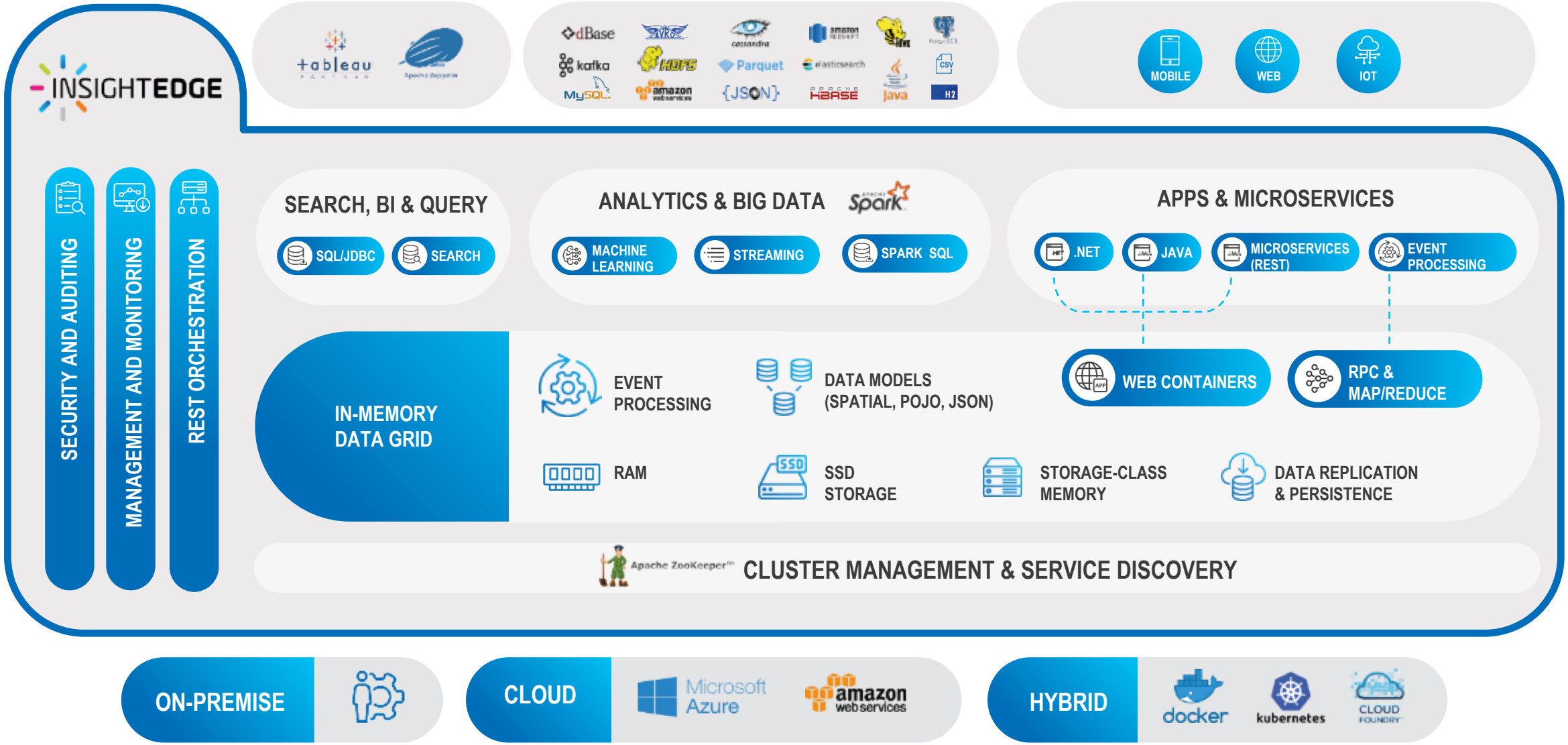
"Graphs and Topologies", rather than "Layers and Tiers"

InsightEdge: Unifying Real-Time Analytics, AI and Transactional Processing in One Open Source Platform

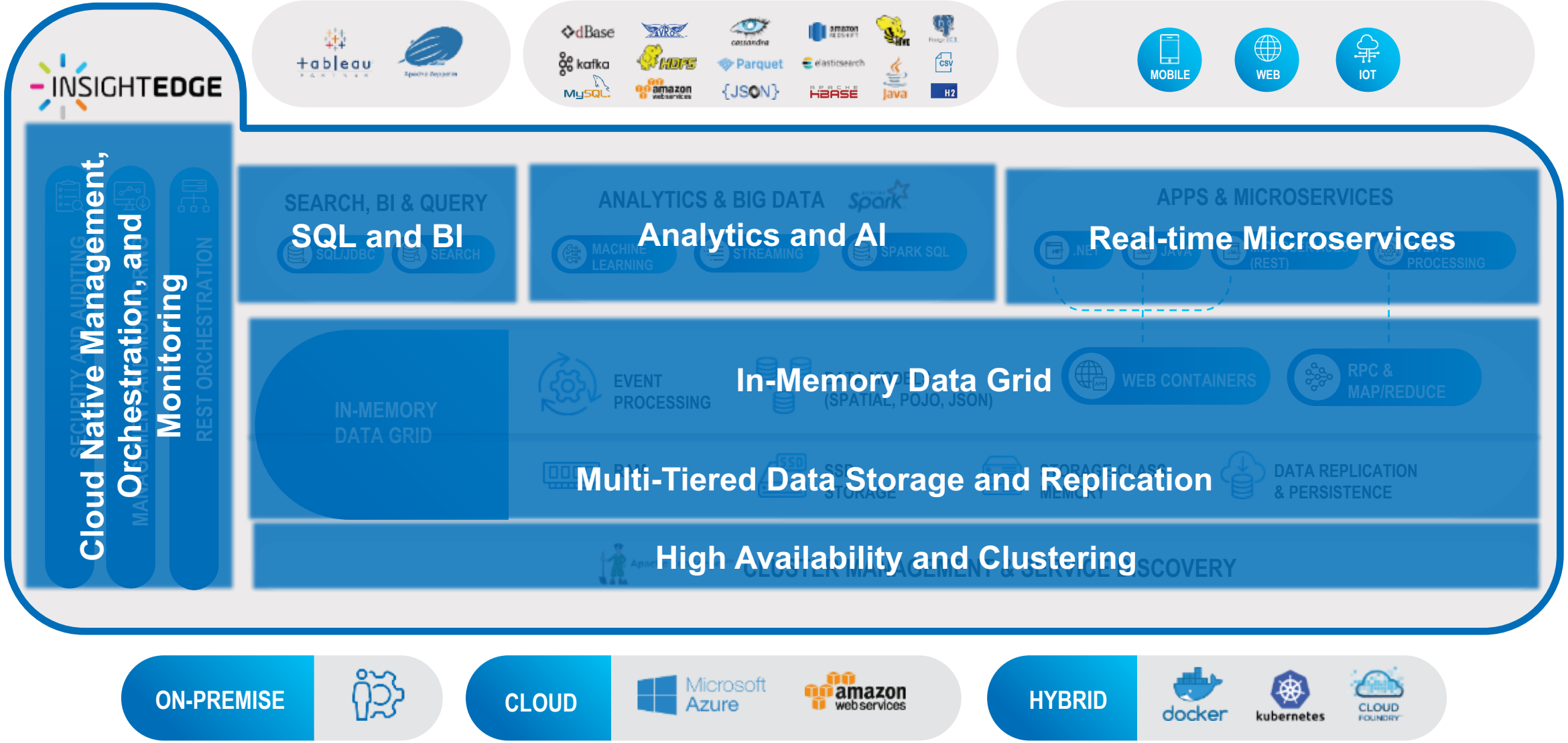
- Open Source & Open API
- Rich ML & DL support
- Extreme performance
- Fully Transactional
- ACID Compliance
- Enterprise-grade
  (Security, High Availability)
- Co-located Apps and Services
- Seamless integration with Big Data ecosystem
  - Data sources (Kafka/Nifi/Talend)
  - Data lakes (S3/Hadoop)
  - BI tools (Tableau/Looker/etc.)



InsightEdge Architecture Overview



InsightEdge
Unifying Fast Data Analytics, AI and Transactional Processing



Ultra-low latency and high throughput transactional processing IMDG

Partitioned In-Memory Grid
Shared-nothing, linear scalability, elastic capacity

Co-Location of Data and Business Logic
Co-located ops, event-driven, fast indexing

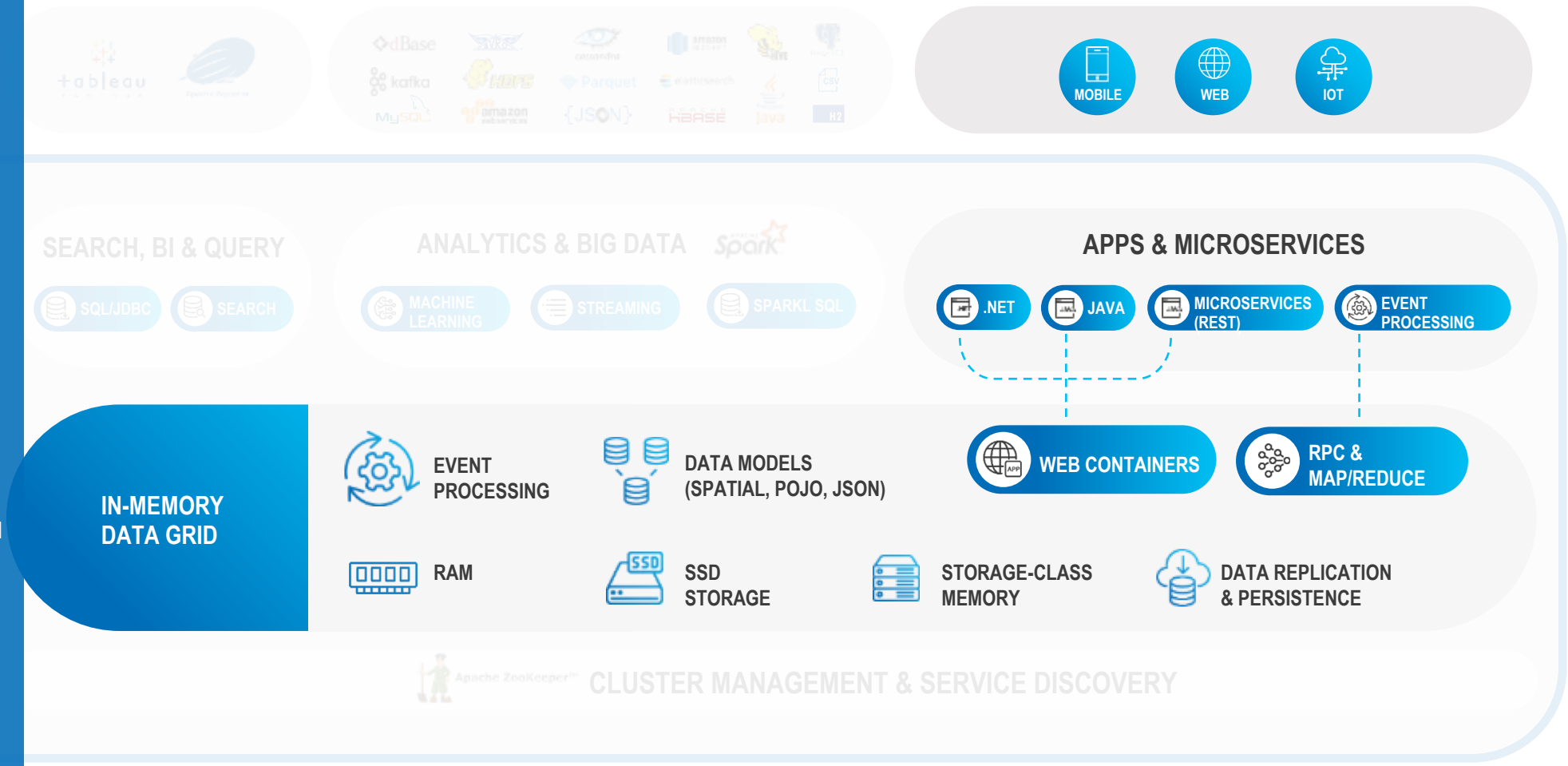
Event-Driven Processing and Map/Reduce

No Downtime
Auto-healing, multi-data center replication, fault tolerance

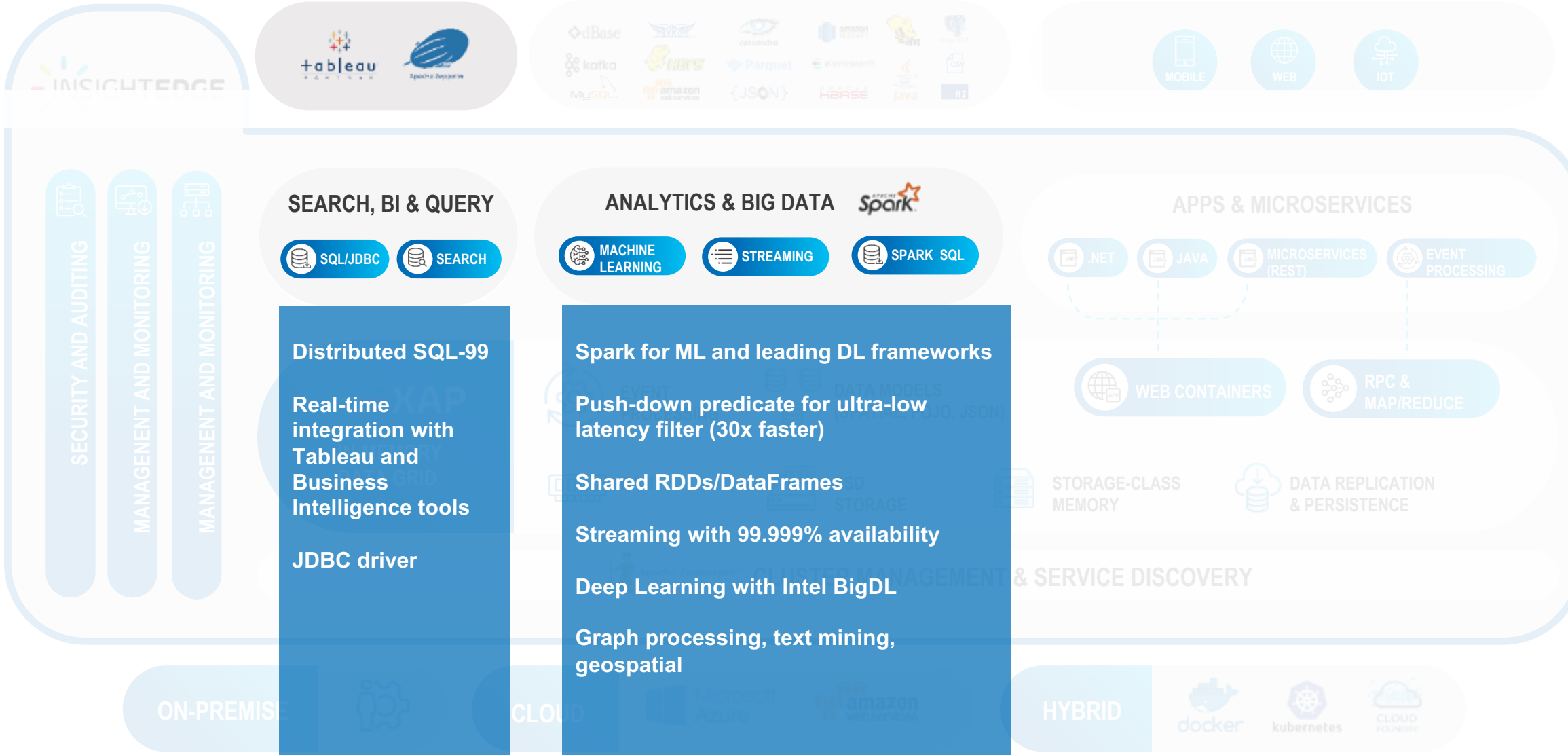
Fast Indexing Multi-Data Model
POJO, .NET, Document/JSON, Geospatial, Time-series

Seamless Integration wih Java/Scala ecosystem

Cloud Native



Co-located Analytics and AI with Transactional Processing



High Availability & Clustering

- ZooKeeper-based clustering for 1000s of nodes
- Back-up and auto-healing for each grid container
- N + 1 redundancy
- Unicast or Multicast discovery

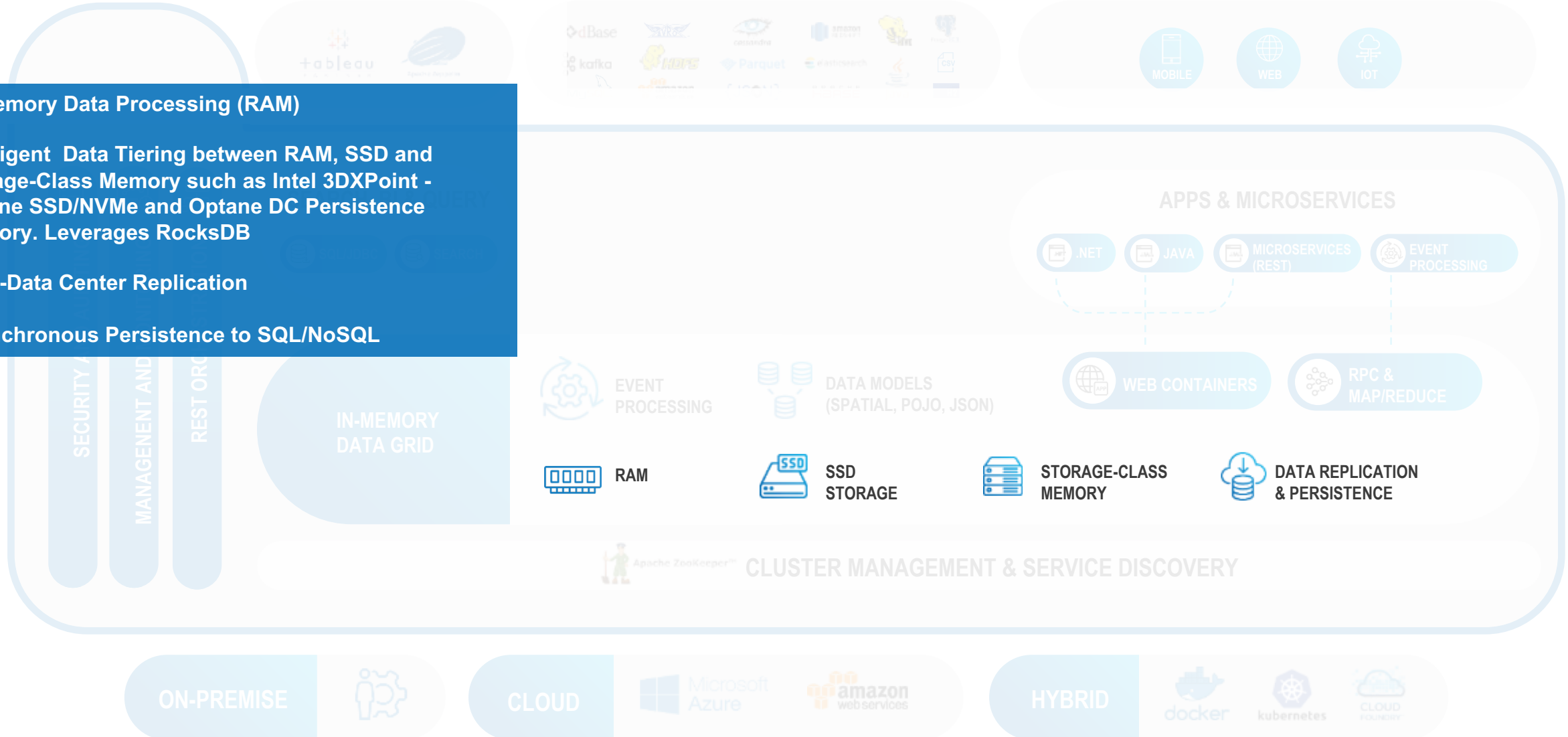
Multi-Tiered Data Storage and Replication for Optimized TCO

In-Memory Data Processing (RAM)

Intelligent Data Tiering between RAM, SSD and Storage-Class Memory such as Intel 3DXPoint - Optane SSD/NVMe and Optane DC Persistence memory. Leverages RocksDB

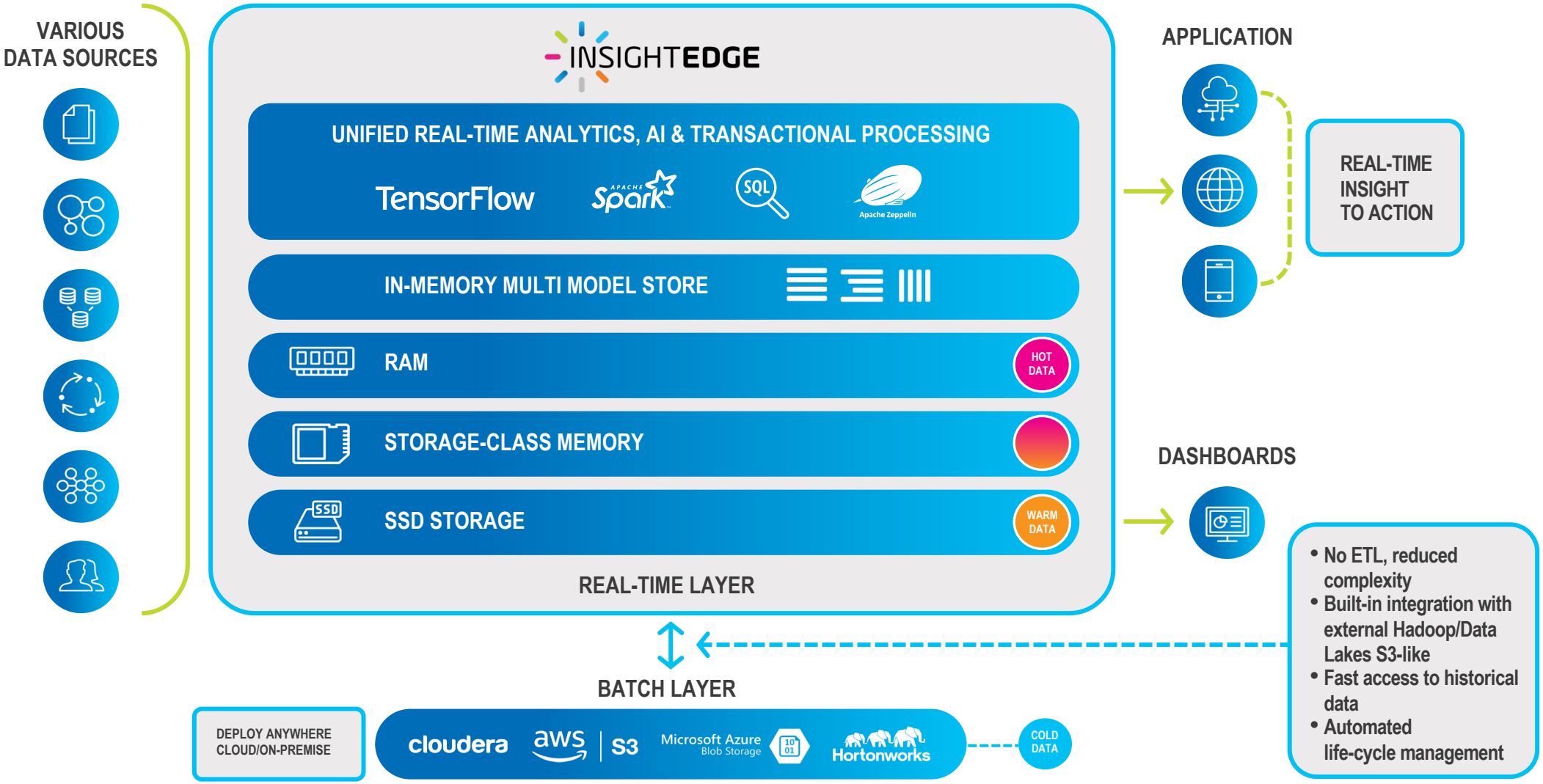
Multi-Data Center Replication

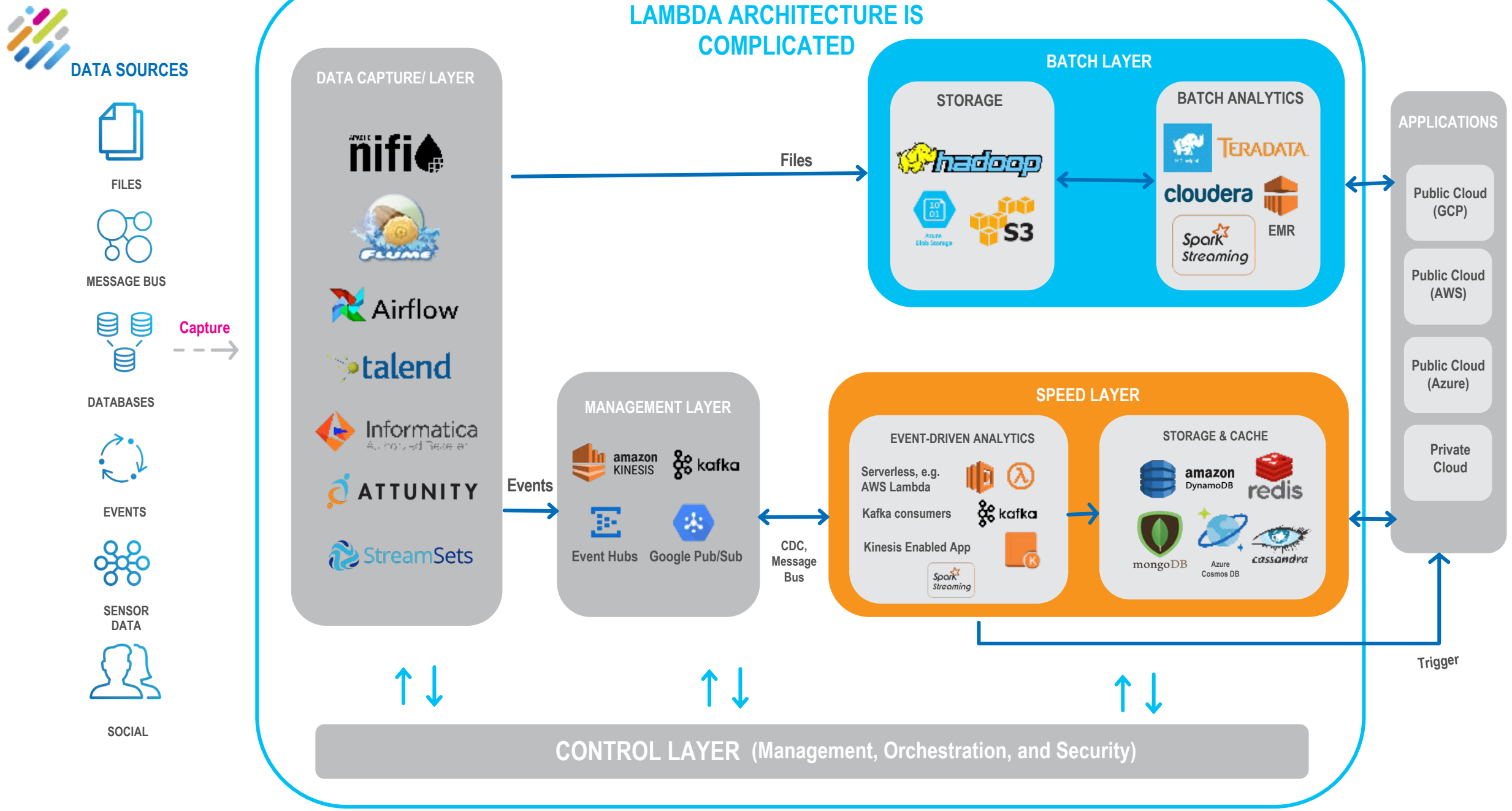
Asynchronous Persistence to SQL/NoSQL

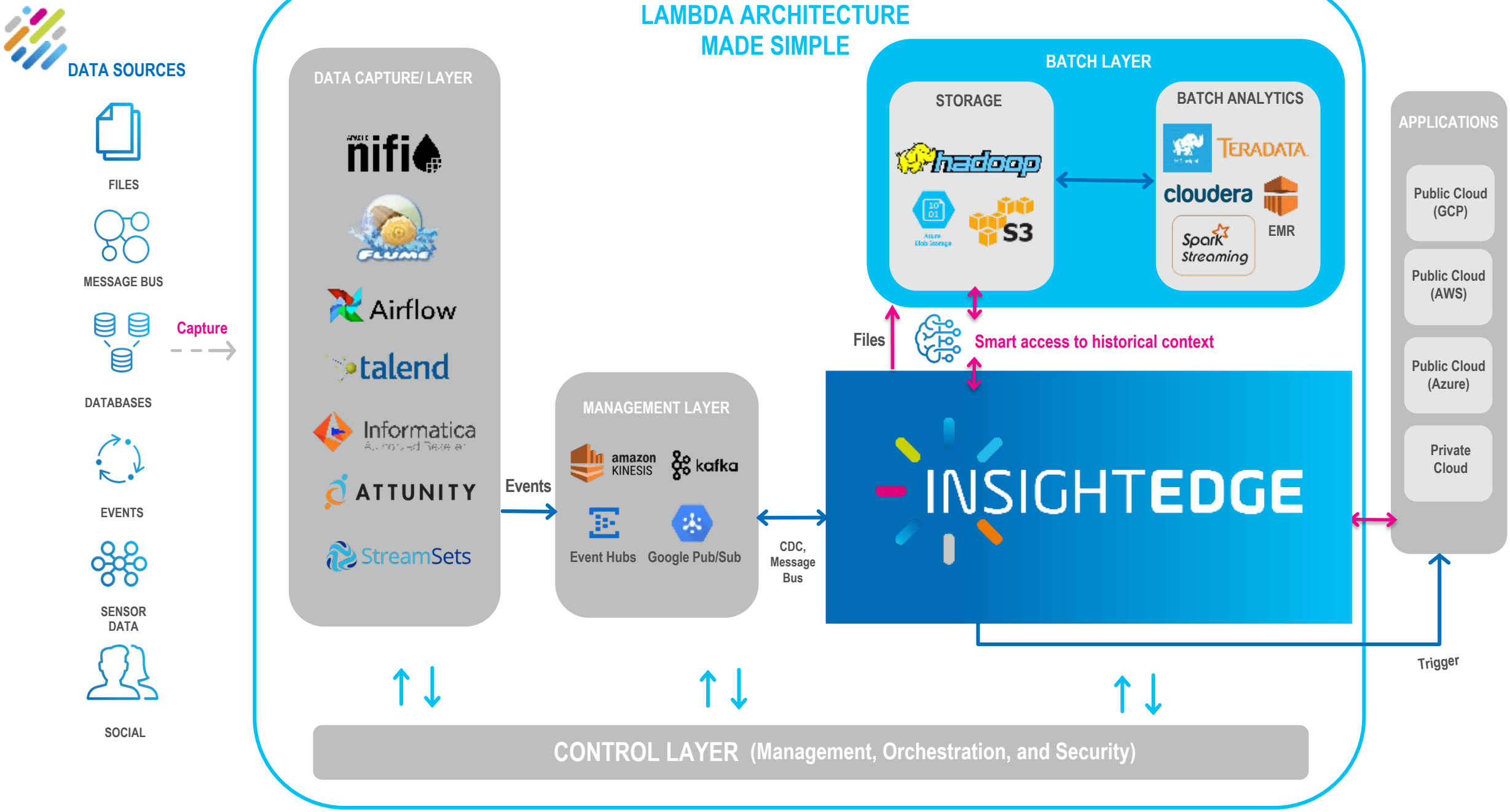


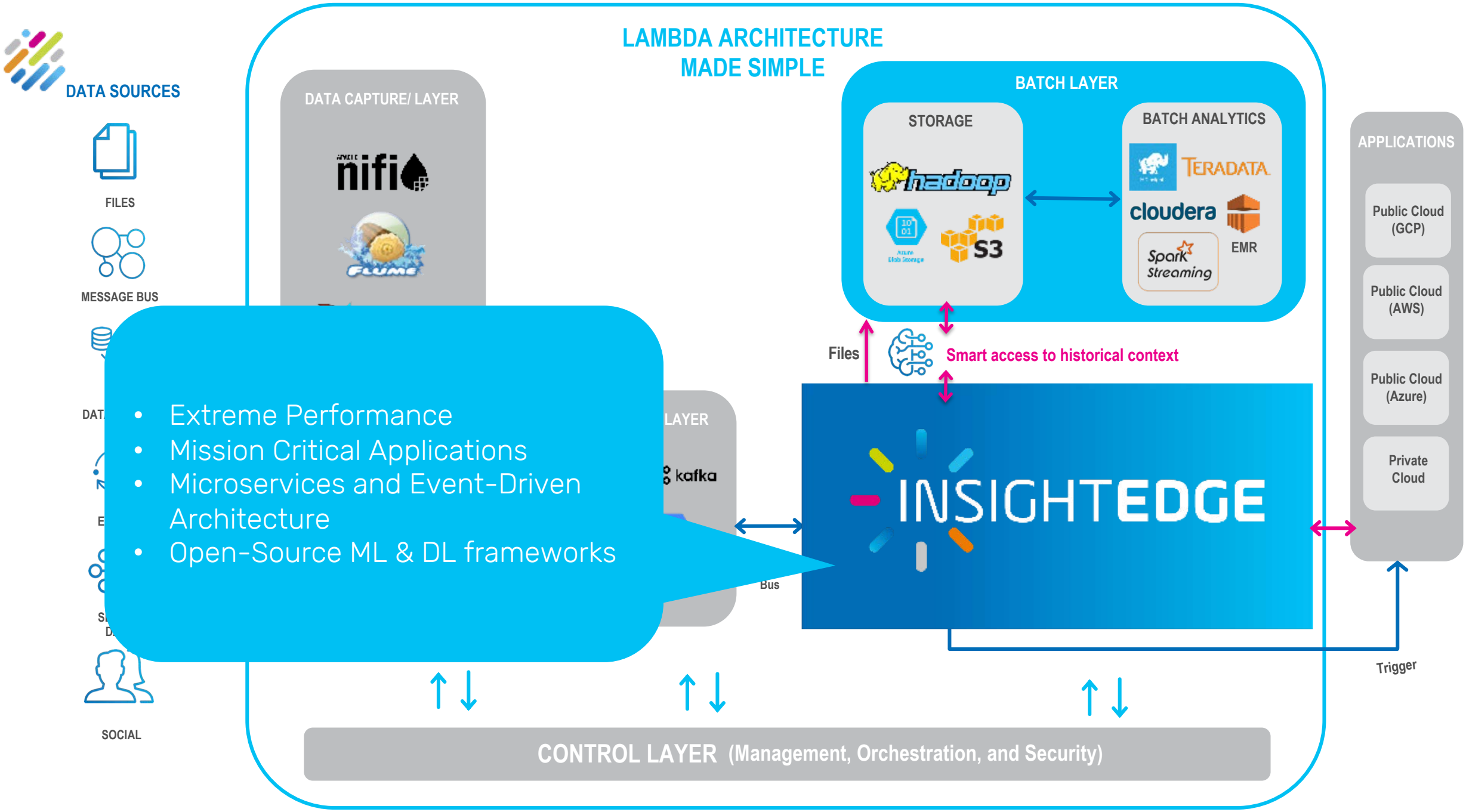
* Apache Pass support in Q4 2018

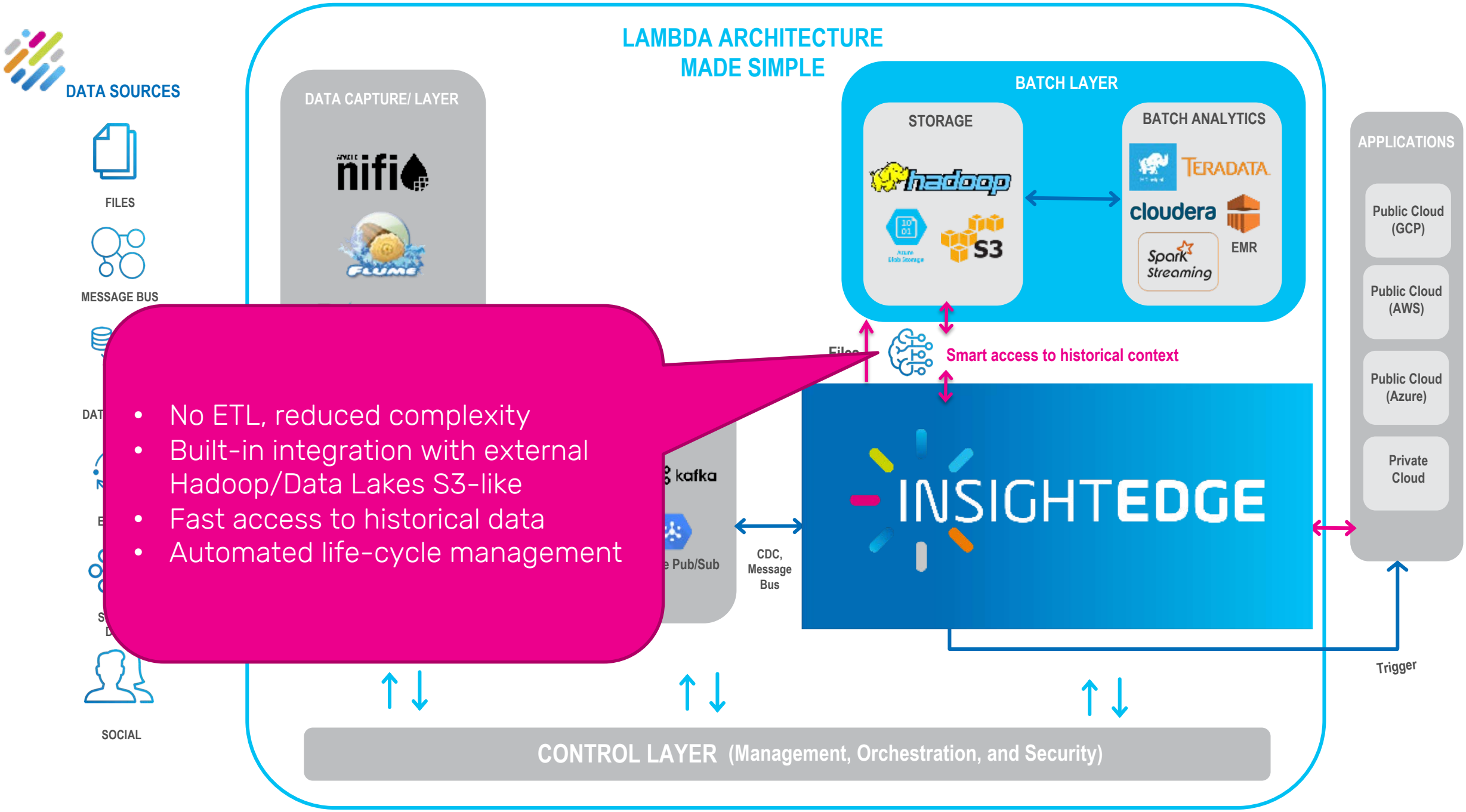
Effectively Scale IoT: Real-time Analytics for Instant Insights To Action

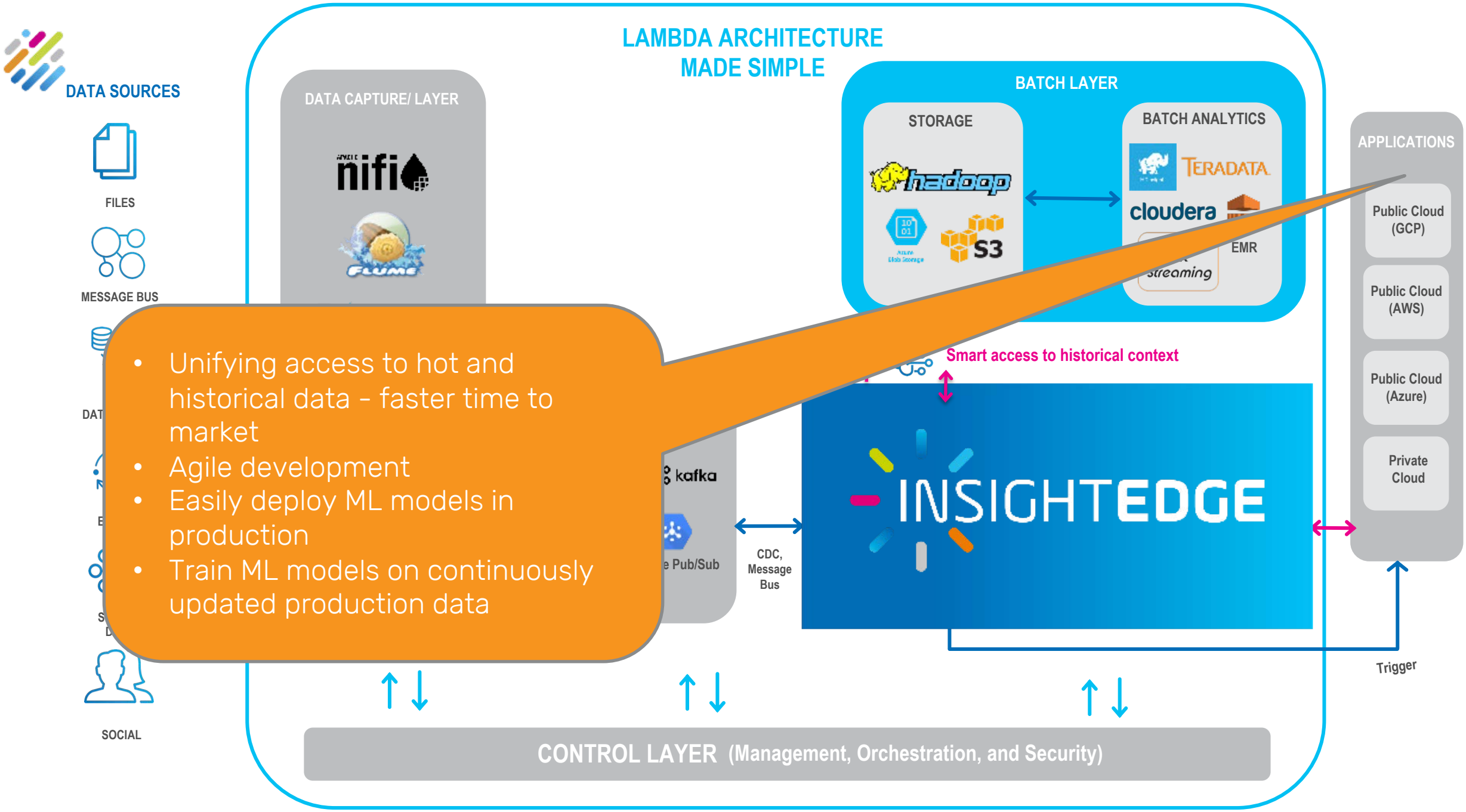



DATA SOURCES
FILES
MESSAGE BUS
DATABASES
Capture
EVENTS
SENSOR DATA
SOCIAL
LAMBDA ARCHITECTURE IS COMPLICATED
DATA CAPTURE/ LAYER
nifi
FLUME
Airflow
talend
Informatica
ATTUNITY
StreamSets
Files
BATCH LAYER
STORAGE
hadoop
S3
BATCH ANALYTICS
TERADATA
cloudera
Spark Streaming
EMR
APPLICATIONS
Public Cloud (GCP)
Public Cloud (AWS)
Public Cloud (Azure)
Private Cloud
Events
MANAGEMENT LAYER
amazon KINESIS
kafka
Event Hubs
Google Pub/Sub
CDC, Message Bus
SPEED LAYER
EVENT-DRIVEN ANALYTICS
Serverless, e.g. AWS Lambda
Kafka consumers
kafka
Kinesis Enabled App
Spark Streaming
STORAGE & CACHE
amazon DynamoDB
redis
mongoDB
Azure Cosmos DB
cassandra
Trigger
CONTROL LAYER (Management, Orchestration, and Security)


LAMBDA ARCHITECTURE
MADE SIMPLE
DATA SOURCES
FILES
MESSAGE BUS
DATABASES
EVENTS
SENSOR DATA
SOCIAL
Capture
DATA CAPTURE/ LAYER
nifi
FLUME
Airflow
talend
Informatica
ATTUNITY
StreamSets
Events
MANAGEMENT LAYER
amazon KINESIS
kafka
Event Hubs
Google Pub/Sub
CDC, Message Bus
BATCH LAYER
STORAGE
hadoop
S3
BATCH ANALYTICS
TERADATA
cloudera
EMR
Spark Streaming
Files
Smart access to historical context
INSIGHTEDGE
APPLICATIONS
Public Cloud (GCP)
Public Cloud (AWS)
Public Cloud (Azure)
Private Cloud
Trigger
CONTROL LAYER (Management, Orchestration, and Security)


LAMBDA ARCHITECTURE
MADE SIMPLE
DATA SOURCES
FILES
MESSAGE BUS
SOCIAL
DATA CAPTURE/ LAYER
nifi
FLUME
LAYER
kafka
Bus
BATCH LAYER
STORAGE
hadoop
S3
BATCH ANALYTICS
TERADATA
cloudera
EMR
Spark Streaming
Files
Smart access to historical context
INSIGHTEDGE
APPLICATIONS
Public Cloud (GCP)
Public Cloud (AWS)
Public Cloud (Azure)
Private Cloud
Trigger
CONTROL LAYER (Management, Orchestration, and Security)
• Extreme Performance
• Mission Critical Applications
• Microservices and Event-Driven Architecture
• Open-Source ML & DL frameworks


LAMBDA ARCHITECTURE
MADE SIMPLE
DATA SOURCES
FILES
MESSAGE BUS
SOCIAL
DATA CAPTURE/ LAYER
nifi
FLUME
kafka
Pub/Sub
BATCH LAYER
STORAGE
hadoop
S3
BATCH ANALYTICS
TERADATA
cloudera
EMR
Spark Streaming
Smart access to historical context
INSIGHTEDGE
CDC, Message Bus
APPLICATIONS
Public Cloud (GCP)
Public Cloud (AWS)
Public Cloud (Azure)
Private Cloud
Trigger
CONTROL LAYER (Management, Orchestration, and Security)
• No ETL, reduced complexity
• Built-in integration with external Hadoop/Data Lakes S3-like
• Fast access to historical data
• Automated life-cycle management


LAMBDA ARCHITECTURE
MADE SIMPLE
DATA SOURCES
FILES
MESSAGE BUS
SOCIAL
DATA CAPTURE/ LAYER
nifi
FLUME
kafka
Pub/Sub
CDC,
Message
Bus
BATCH LAYER
STORAGE
hadoop
S3
BATCH ANALYTICS
TERADATA
cloudera
EMR
streaming
Smart access to historical context
INSIGHTEDGE
APPLICATIONS
Public Cloud
(GCP)
Public Cloud
(AWS)
Public Cloud
(Azure)
Private
Cloud
Trigger
CONTROL LAYER (Management, Orchestration, and Security)
• Unifying access to hot and historical data - faster time to market
• Agile development
• Easily deploy ML models in production
• Train ML models on continuously updated production data

451 Research

"GigaSpaces is now focused on in-memory data processing… The combination of Spark and XAP will enable GigaSpaces to target the new breed of real-time analytics and hybrid operational and analytic workloads."

FROST & SULLIVAN

"InsightEdge contains all the necessary SQL, Spark, Streaming, and Deep Learning toolkits for scalable data-driven solutions… our preferred solution components: the three-tier Kappa model, including Spark and Kafka, as implemented by GigaSpaces, in combination with its commercial InsightEdge platform."

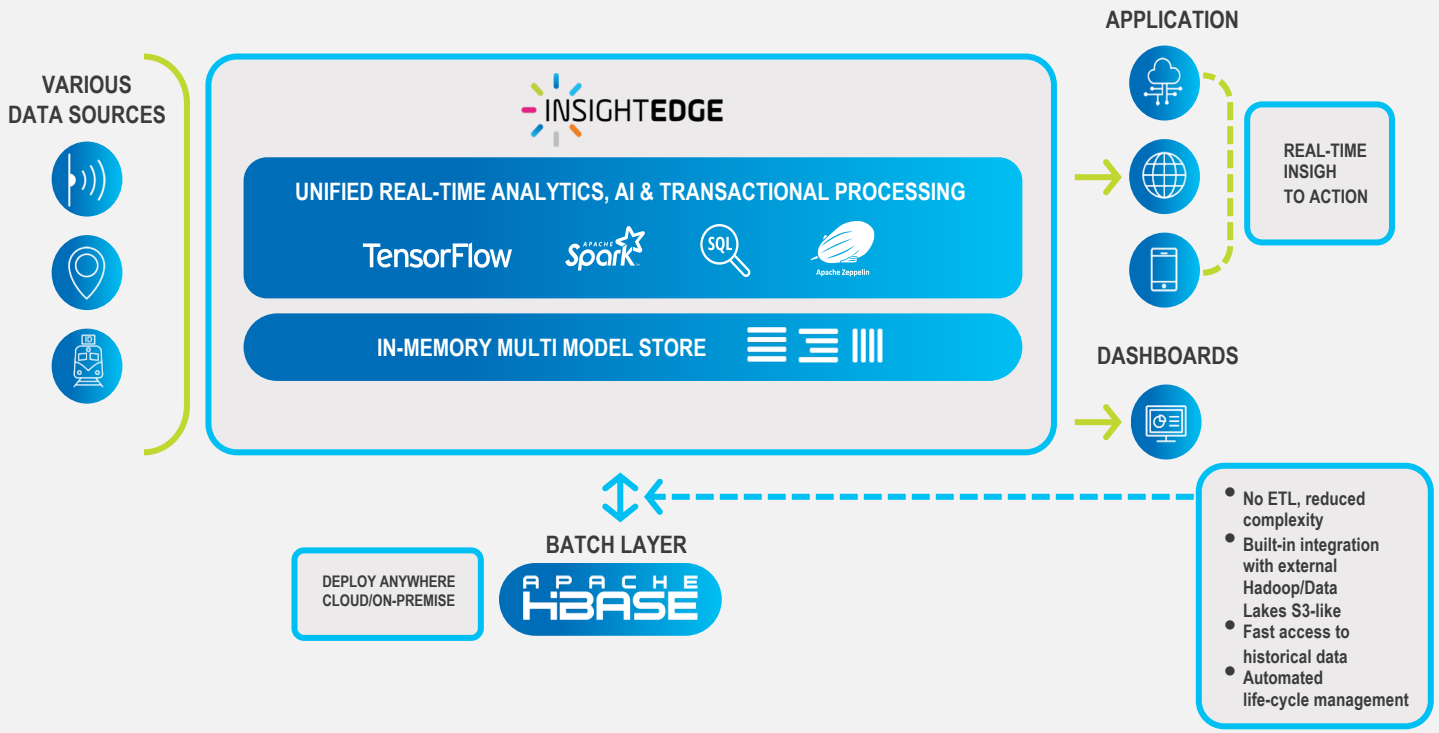
Everyone Wants "Real-time Analytic Insights"
But Which Architecture Will Get You There?

CASE STUDY: Predictive Maintenance for Leading Rail-Based Transportation Company

TRANSPORTATION

BUSINESS CHALLENGE:

- Predictive maintenance of equipment, field data ingestion and stream processing
- Ability to redirect trains in a timely manner



TECHNICAL CHALLENGE:

- Process streaming data at scale and query from a live data mart
- Event driven analytics and business logic
- Many small low-volume streams that require correlation and statefulness (the IoT streaming problem)
- Real-time analytics leveraging GPS, train sensor data with reference to historical data

RESULTS:

- Simplified big data pipeline
- High performance stream processing with High Availability
- Real-time analytics on relevant data from train events, fence events and GPS
- Event-based triggers to direct the output to a operational workflows and live dashboards for timely maintenance and redirecting of fast moving trains in time

USE CASE: Magic Software

IOT

BUSINESS CHALLENGE:

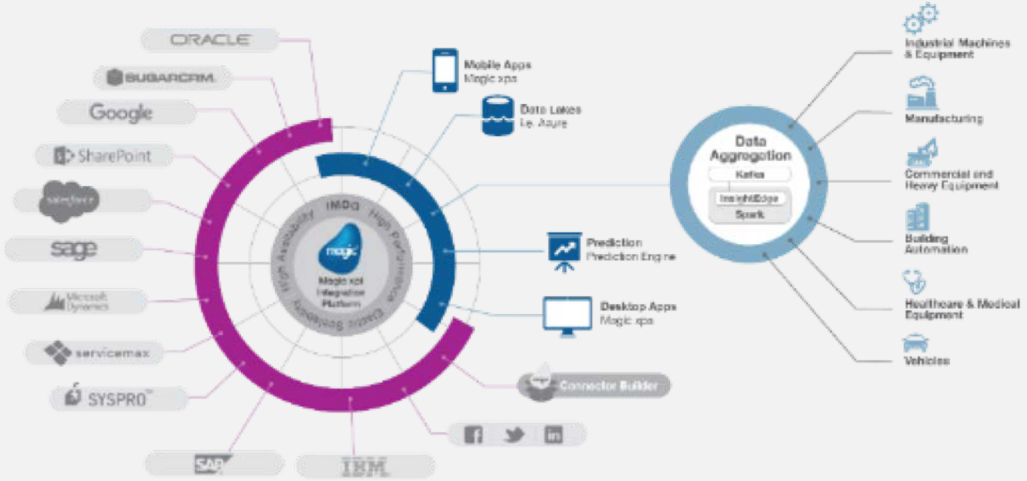
- IoT Hub + Predictive Analytics

TECHNICAL CHALLENGE

- Implement predictive analytics and anomaly detection
- Expand insight context through customer/data-360 integration
- Trigger transactional workflows based on prediction criteria

RESULTS:

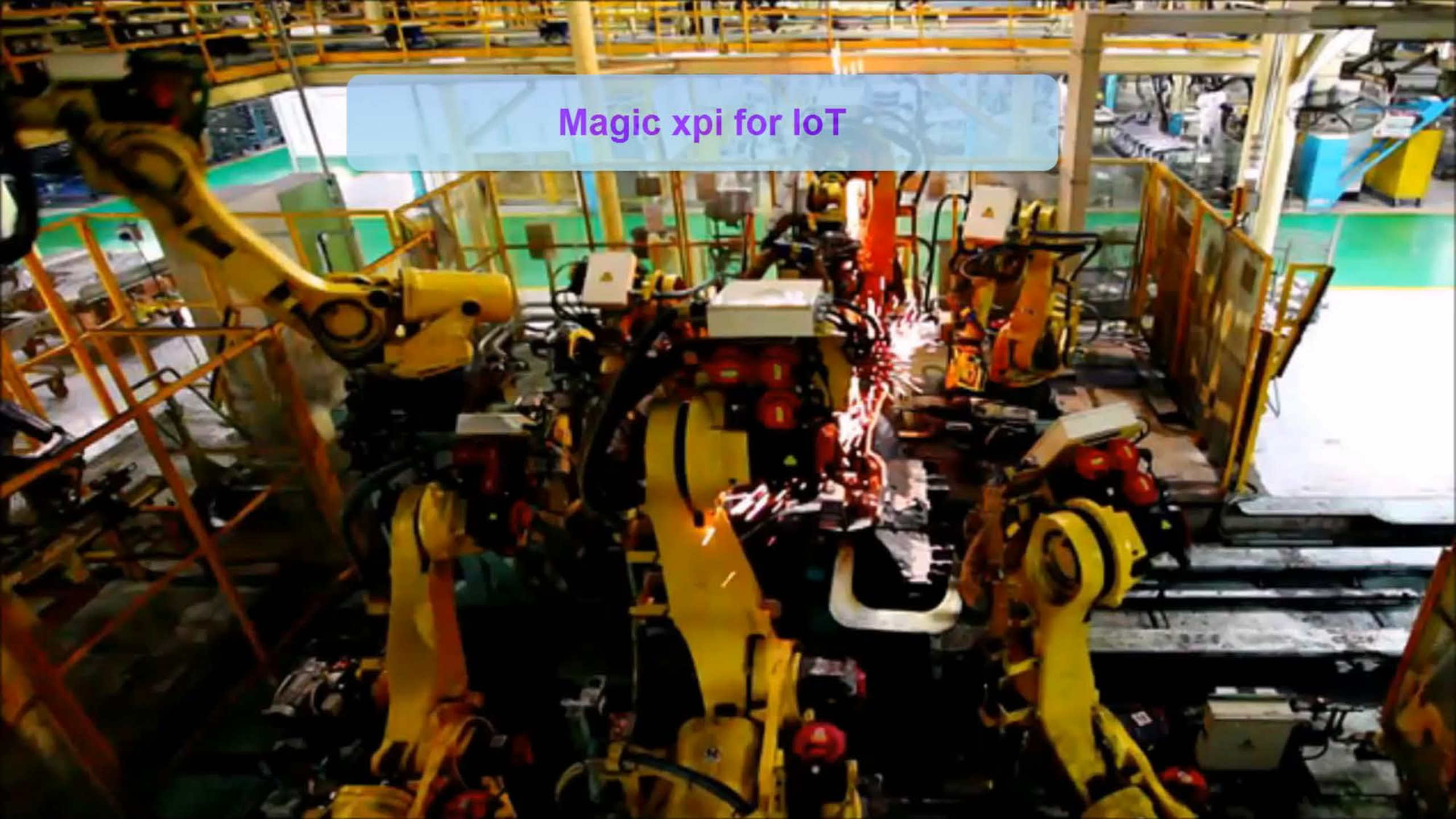
- Simplified HTAP with Streaming data pipeline (3 tiers)
- IoT streaming analytics with 9s high availability



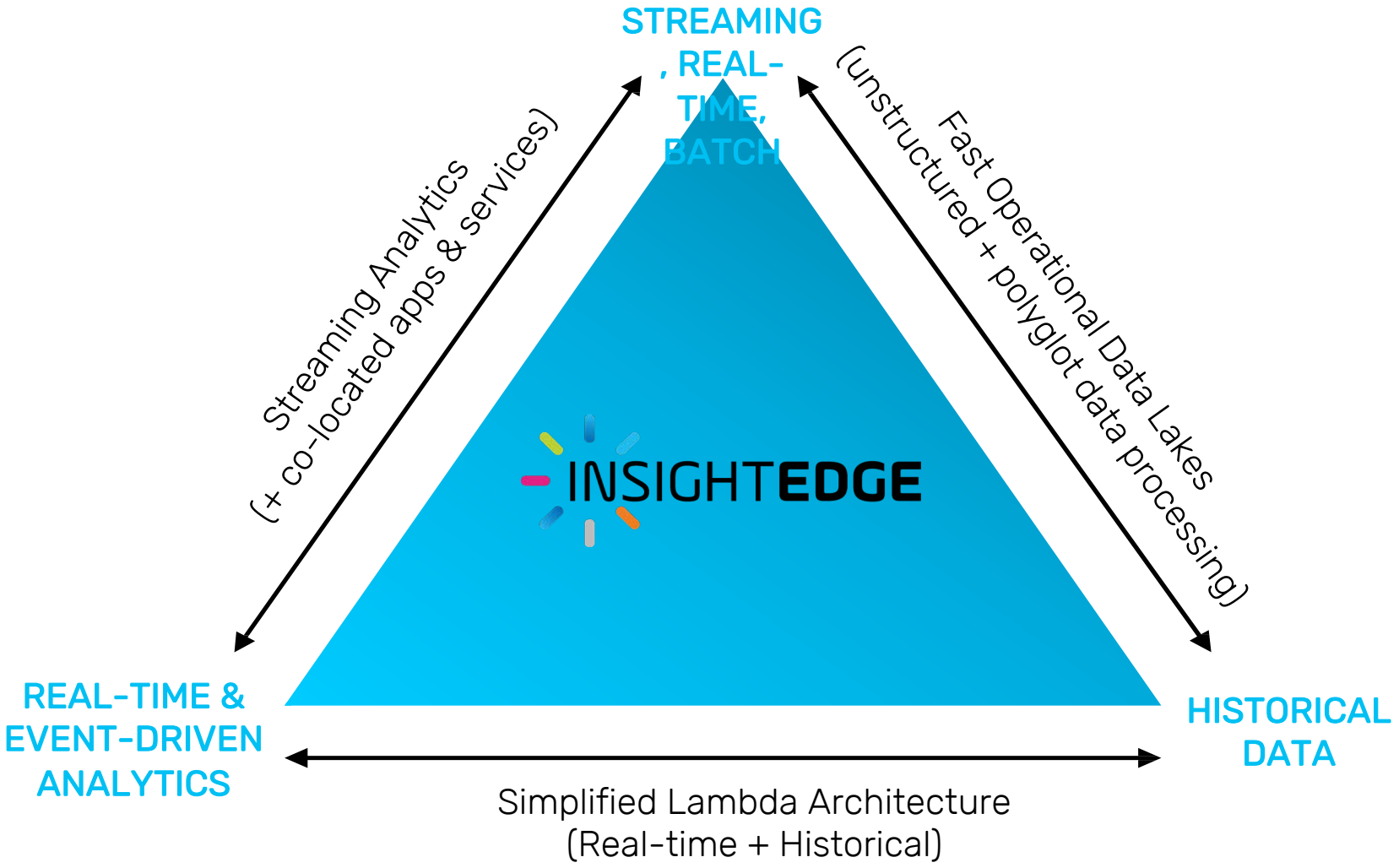
"GigaSpaces enables our customers to simplify and accelerate telemetry ingestion, to gain full business value from IoT adoption."

Yuval Lavi, Vice President of Innovation and Strategy, Magic Software.

Magic xpi for IoT



Faster, Smarter Insights and Actions



INSTANT INSIGHTS TO ACTION

EXTREME PERFORMANCE

TCO OPTIMIZATION

MISSION CRITICAL AVAILABILITY

<1 sec from data to insight to action

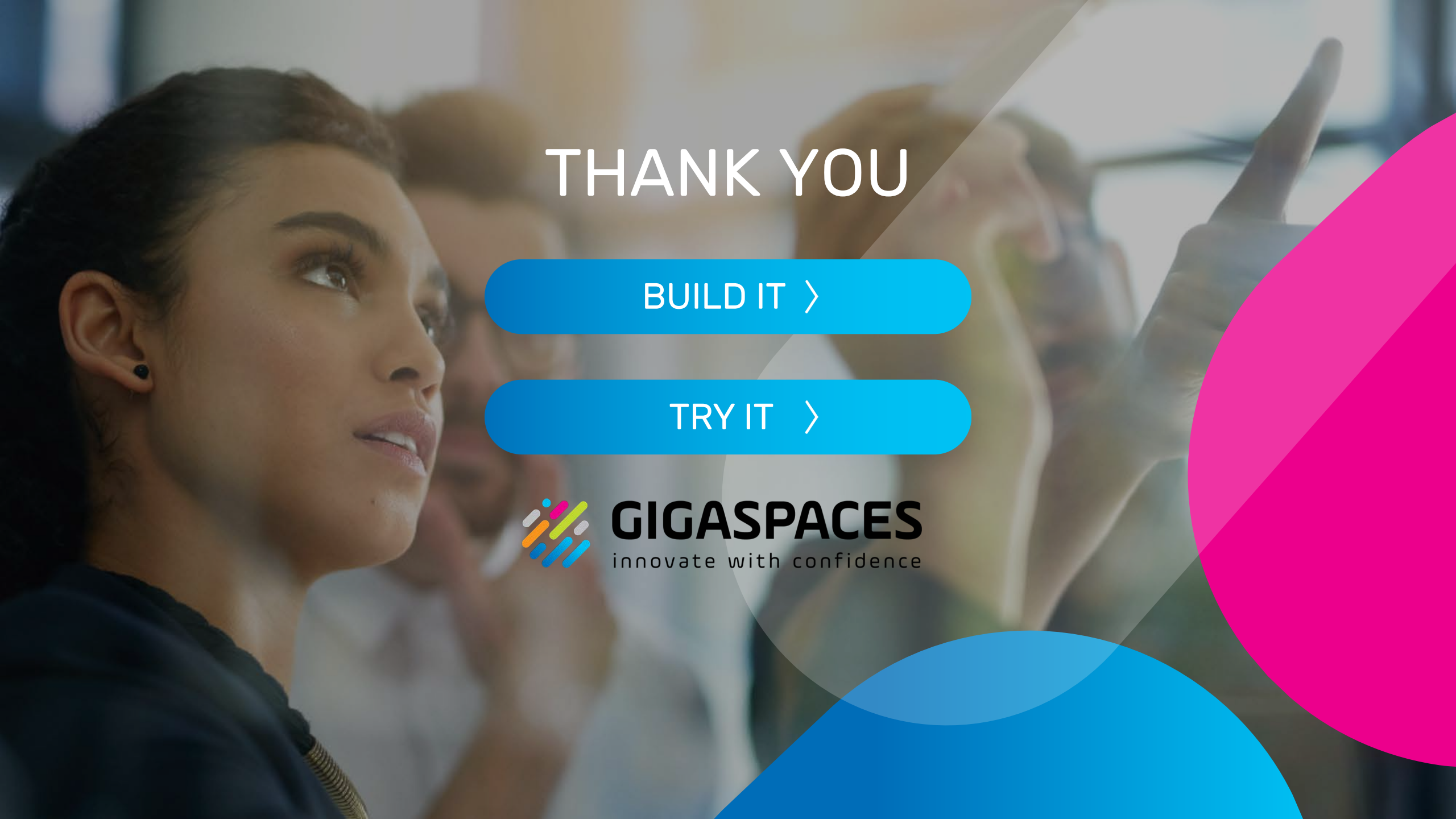
millions of IOPS

10X less expensive than only RAM with In-memory performance

No Downtime at leading enterprise customers for

YEARS

And still counting


THANK YOU
BUILD IT 〉
TRY IT 〉
GIGASPACES
innovate with confidence