hazelcast

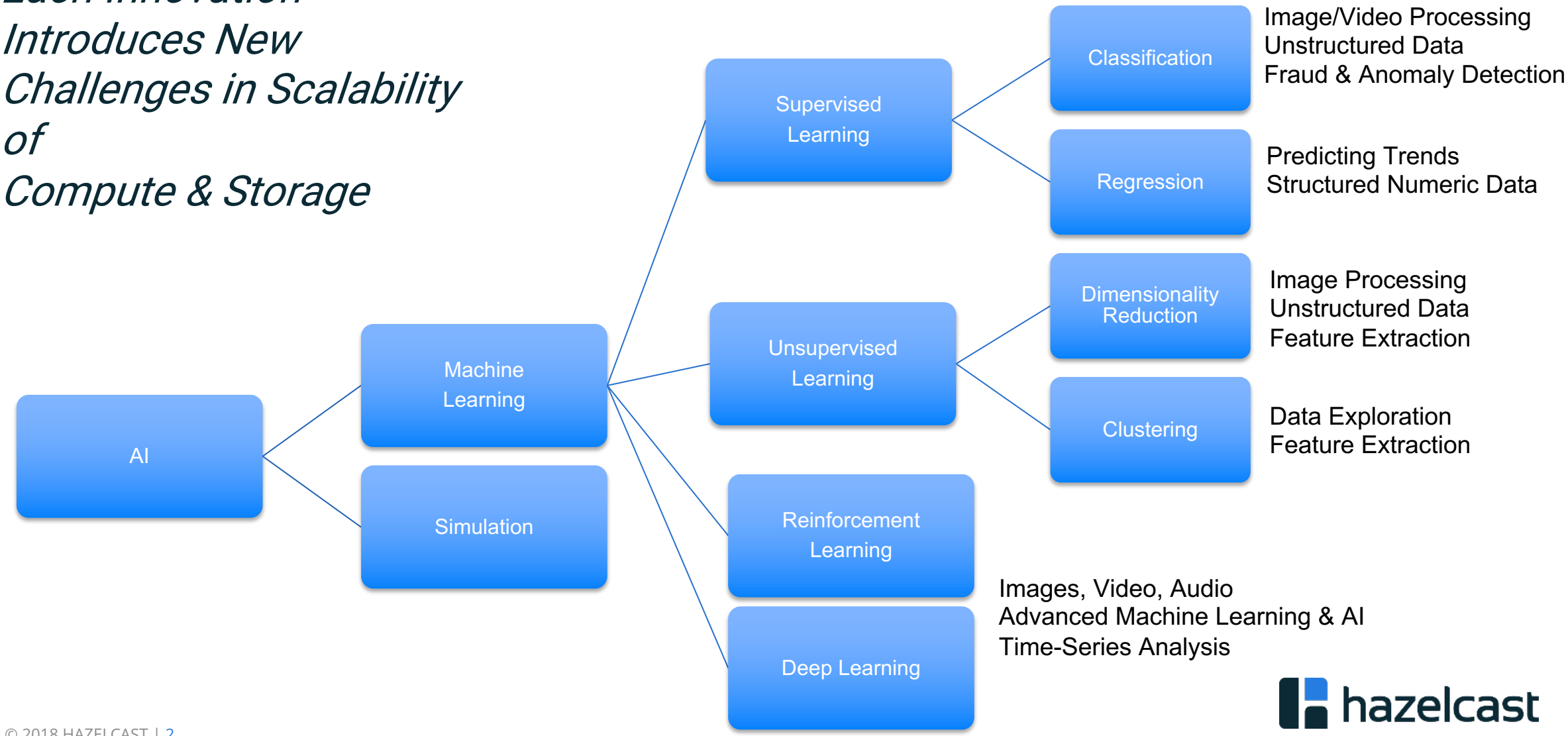
THE LEADING IN-MEMORY COMPUTING PLATFORM

# Low-Latency Machine Learning Applications

**John DesJardins – CTO N. America**

# AI Techniques Continue to Expand & Evolve

*Each Innovation Introduces New Challenges in Scalability of Compute & Storage*

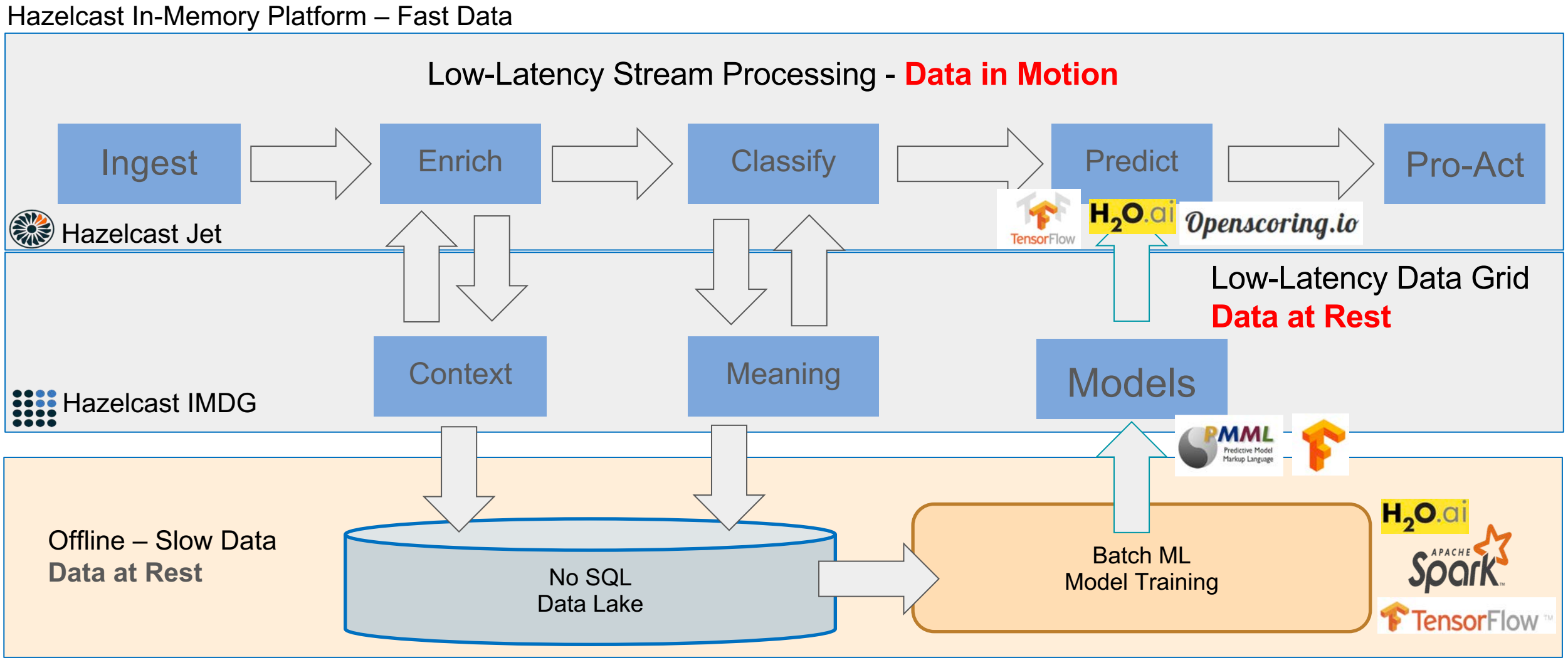



hazelcast

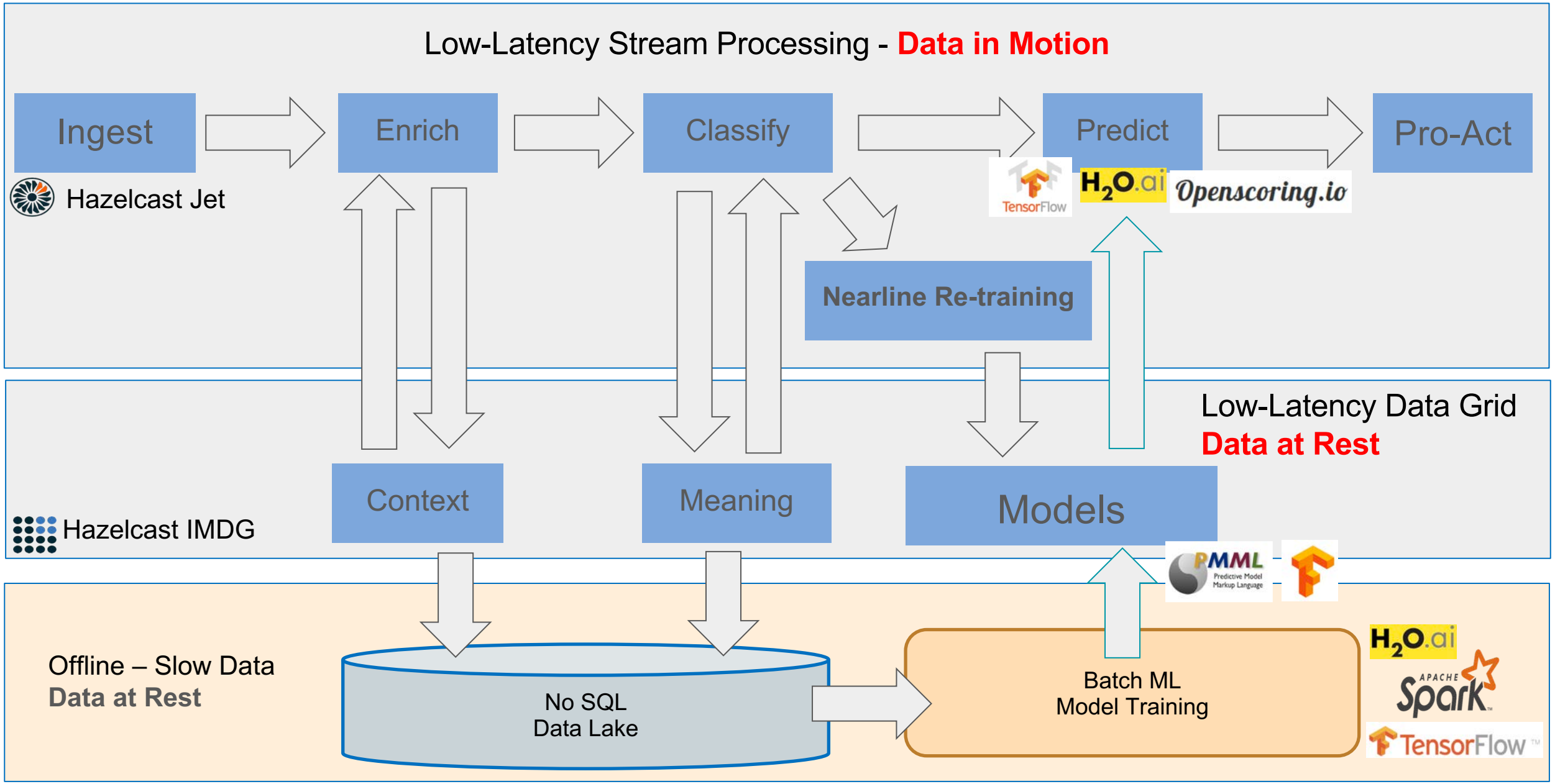
# Online Machine Learning within an In-Memory Platform

*Move from Reactive to Pro-Active*

*Take Action before negative impact or ahead of opportunity*

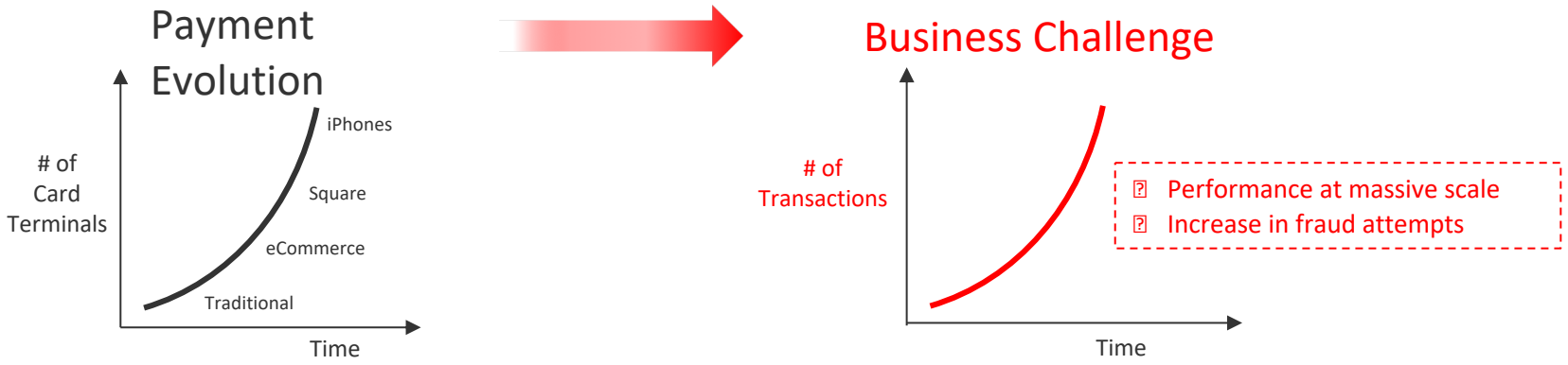


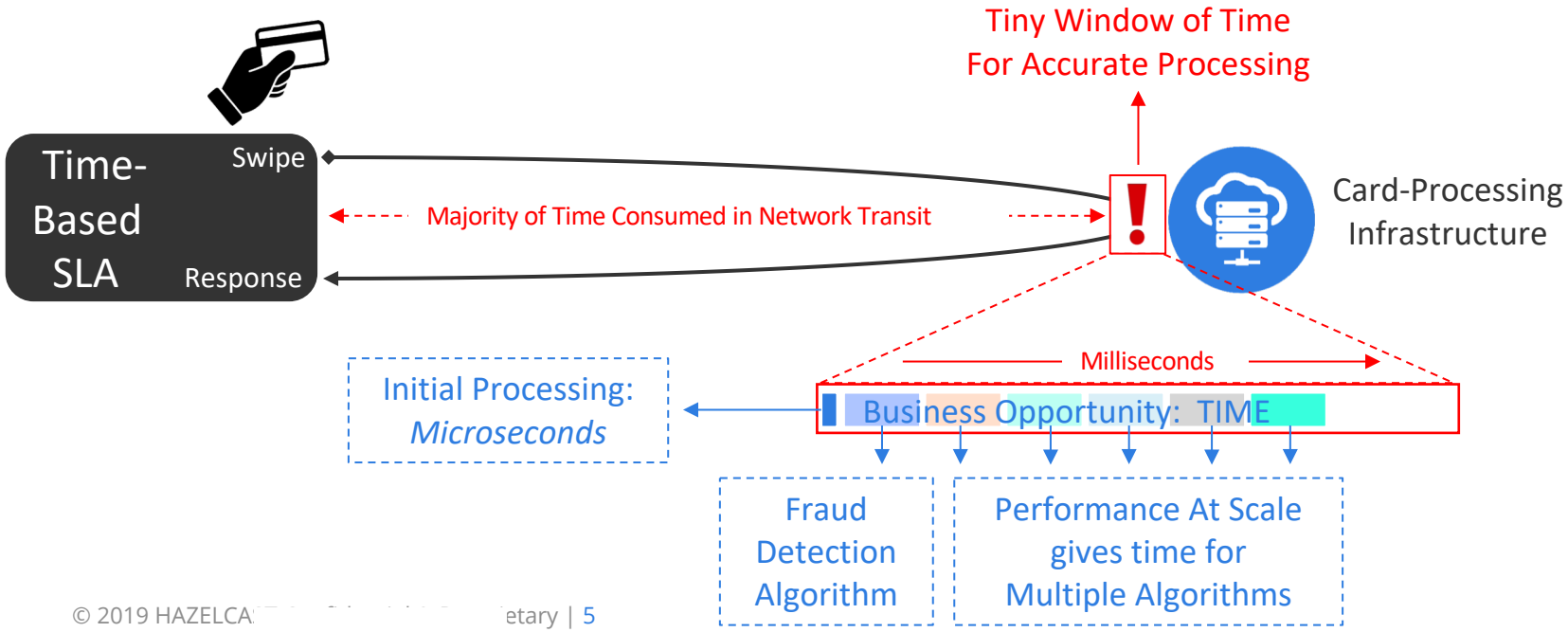
# Online Inference & Nearline Machine Learning



# The Hazelcast Difference

## Example: Credit Card Processing



# Hazelcast - High Performance Platform

Mobile

Apps

Social

Commerce

Communities

## IMDG In-Memory Data Grid

| | |
|---|---|
| **Integrate**<br>APIs, Microservices, Notifications | **Communicate**<br>Serialization, Protocols |
| **Store/Update**<br>Caching, CRUD Persistence | **Compute**<br>Query, Process, Execute |
| **Scale**<br>Clustering & Cloud, High Density | **Replicate**<br>WAN Replication, Partitioning |
| **Secure**<br>Privacy, Authentication, Authorization | **Available**<br>Rolling Upgrades, Hot Restart |

## Jet In-Memory Streams

| | |
|---|---|
| **Ingest & Transform**<br>Events, Connectors, Filtering | **Combine**<br>Join, Enrich, Group, Aggregate |
| **Stream**<br>Windowing, Event-Time Processing | **Compute & Act**<br>Distributed & Parallel Computations |
| **Secure**<br>Privacy, Authentication, Authorization | **Available**<br>Job Elasticity, Graceful shutdown |

## Management Center

**Secure | Manage | Operate**
Embeddable | Scalable | Low-Latency
Secure | Resilient | Distributed

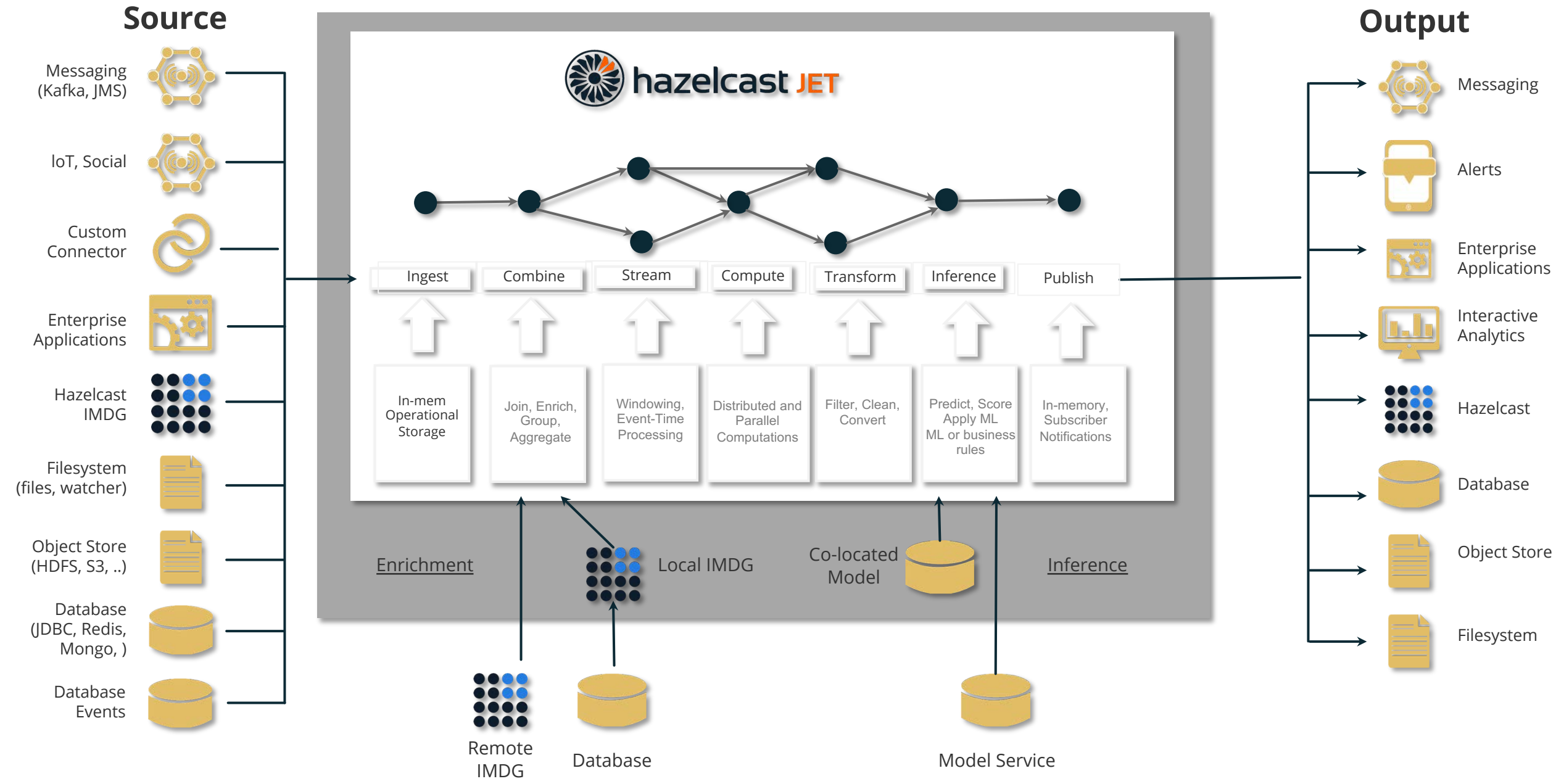
**Analytics**
Visualization
Data Lake

**Live Streams**
Kafka, JMS,
Sensors, Feeds

**Databases**
JDBC,
Relational, NoSQL,
Change Events

**Files**
HDFS, Flat Files,
Logs, File watcher

**Applications**
Sockets

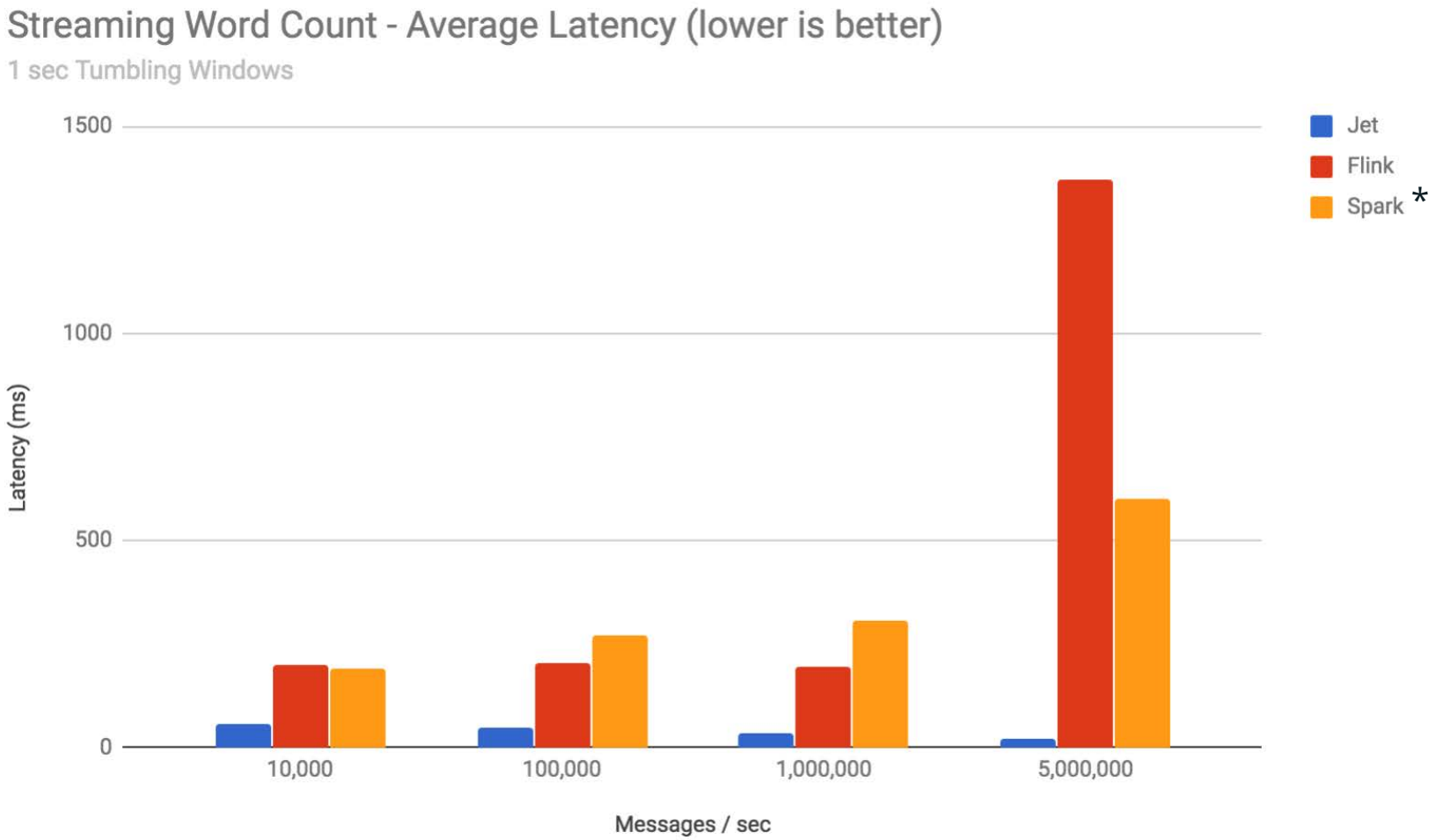

Source
Messaging (Kafka, JMS)
IoT, Social
Custom Connector
Enterprise Applications
Hazelcast IMDG
Filesystem (files, watcher)
Object Store (HDFS, S3, ..)
Database (JDBC, Redis, Mongo, )
Database Events
hazelcast JET
Ingest
Combine
Stream
Compute
Transform
Inference
Publish
In-mem Operational Storage
Join, Enrich, Group, Aggregate
Windowing, Event-Time Processing
Distributed and Parallel Computations
Filter, Clean, Convert
Predict, Score Apply ML ML or business rules
In-memory, Subscriber Notifications
Enrichment
Local IMDG
Co-located Model
Inference
Remote IMDG
Database
Model Service
Output
Messaging
Alerts
Enterprise Applications
Interactive Analytics
Hazelcast
Database
Object Store
Filesystem

# Evolution of Stream Processing

| | |
|---|---|
|  <br>1st Gen (2000s)<br>Hadoop(batch) or Apama(CEP)<br>***hard choices*** | **Distributed Batch Compute** – MapReduce – scaled, parallelized, distributed, resilient, - not real-time<br>or<br>**Siloed, Real-time** – Complex Event Processing – specialized languages, not resilient, not distributed(single instance), hard to scale, fast, but brittle, proprietary |
|  <br>2nd Gen (2014)<br>Spark<br>***hard to manage*** | **Micro-batch distributed** – heavy weight, complex to manage, not elastic, require large dedicated environments with many moving parts,<br>not Cloud-friendly, not low-latency |
|  <br>**3rd Gen (2017 Jet & Flink)**<br>**flexible & scalable**<br>**True "Fast Data"** | **Distributed, real-time streaming** – highly parallel, true streams, advanced techniques (Directed Acyclic Graph) enabling reliable distributed job execution<br>Flexible deployment - Cloud-native, elastic, embeddable, light-weight, supports serverless, fog & edge.<br>Low-latency Streaming, ETL, and fast-batch processing, built on proven data grid |

hazelcast

# Streaming Performance




* Spark had all performance options, including Tungsten, turned on

# Give Hazelcast a try

Hazelcast.org – Data Grid

https://github.com/hazelcast

Jet.hazelcast.org – Streaming Analytics

Demos: https://jet.hazelcast.org/demos/

My info:

- John DesJardins
- john.desjardins@hazelcast.com - email
- @johnmdesjardins - Twitter


Thank You
hazelcast