AEROSPIKE USER SUMMIT 2018

AEROSPIKE

Aerospike Technical Overview

Brian Bulkowski
Founder and Chief Technology Officer
Aerospike

Srini Srinivasan
Founder and Chief Development Officer
Aerospike

Bharath Yadla
Vice President, Product Strategy, Ecosystems
Aerospike

Classic Distributed System Failures

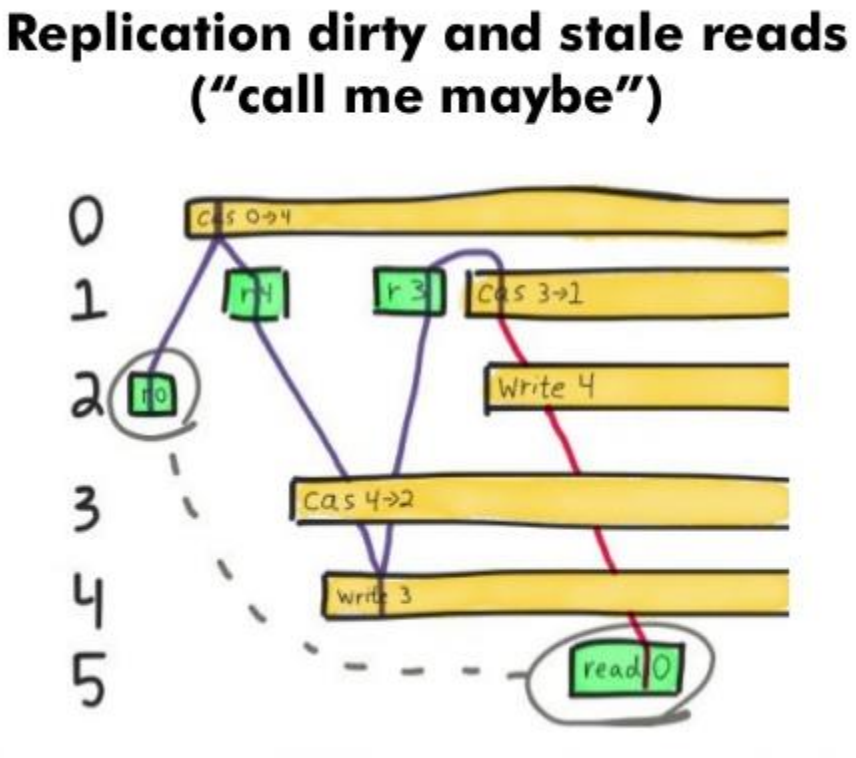
Data Location Updates: Reads or writes are applied to a wrong quorum of servers

Asynchronous Replication: Data is applied, but then crashes, other writes are applied, revived server overwrites

Buffered Writes: A crash occurs before data is written to persistent storage

Clock Problems: A subsequent update is applied to a server with a clock in the past

Bugs: A correct architecture, poorly implemented



AEROSPIKE

Aerospike 4.0: Strong Consistency with High Performance

Primary Key Consistency

Provide Strong Consistency on primary key, Linearizability and Session Consistency.

Hybrid Memory Architecture

Indexes in DRAM and data in Flash provide storage guarantees and unlock Flash's performance

Commit To Device

Data with highest durability requirements can be synchronously written to Flash storage with little performance loss

Advanced Cluster Management

New Aerospike cluster management enforces single-master but allows for predictable sub-second master handoff during failures

Transaction Model

Distributed master oriented approach creates two-phase commit semantics without transaction log

Hybrid Clock

High performance transaction clock tolerates up to 30 seconds of cluster clock skew and 1 million update per second per record granularity

Roster

List of nodes in cluster assigned to namespace

Easy to manage

Cluster is formed, administrator simply chooses from the list (or says “all”)

Theoretically required

“outside information” to create split-brain policies

Calculate the “designated masters and replicas”

The nodes which will contain data in “steady state”

Hybrid Clock

“Lamport Clock”
Best choice for master-based commit systems

Efficient
Less data than a vector clock

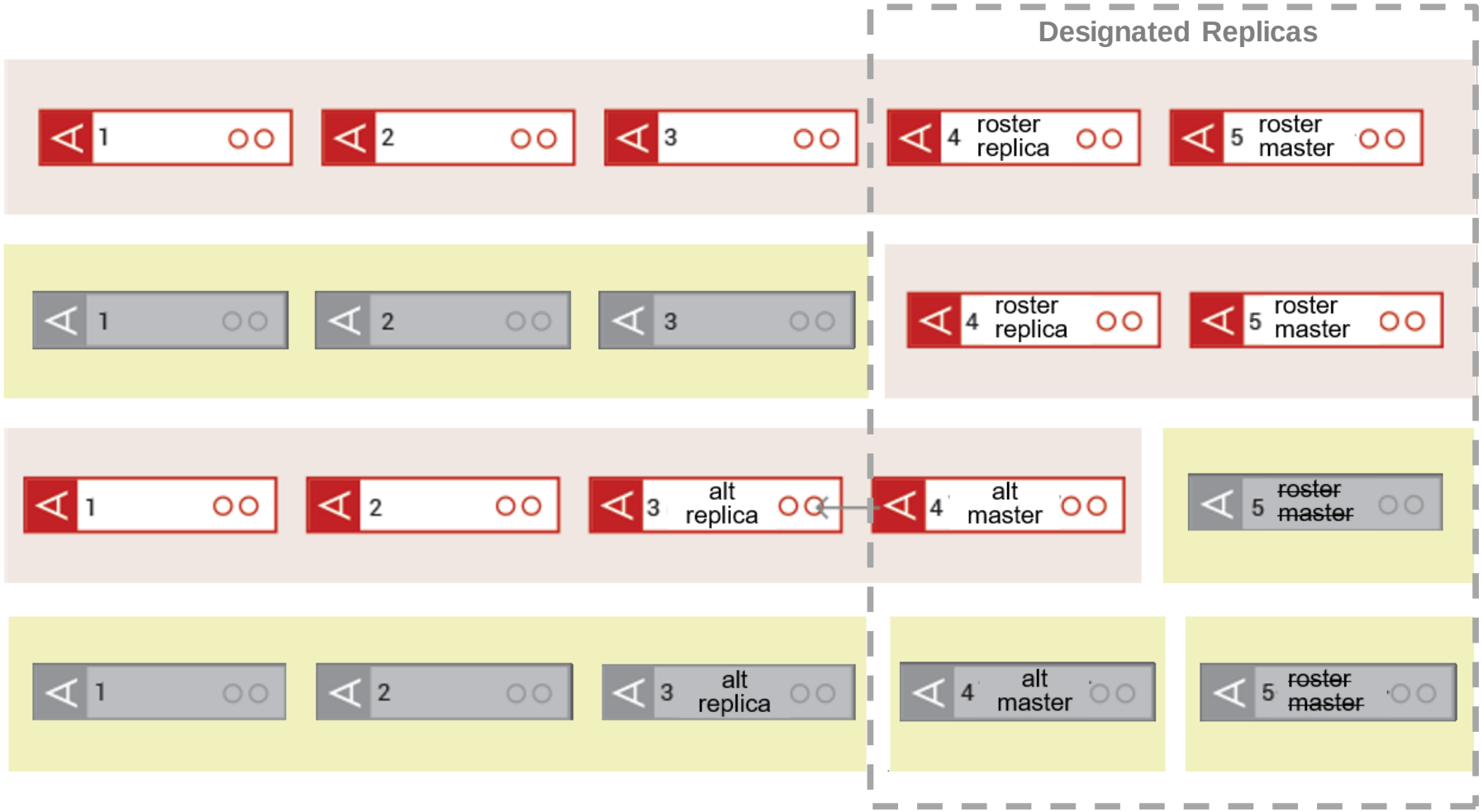
Combines three truths
“Regime” when master changes
Local timestamp (milliseconds)
Counter to gain 1M writes per sec

27 seconds before theoretical roll-over
Keep your clocks loosly in sync

Aerospike 4.0: Master/Replica Promotion and Availability

Roster – This defines the list of nodes that are part of the cluster in steady state

*Example applies to an individual partition **p***



Cluster Healthy

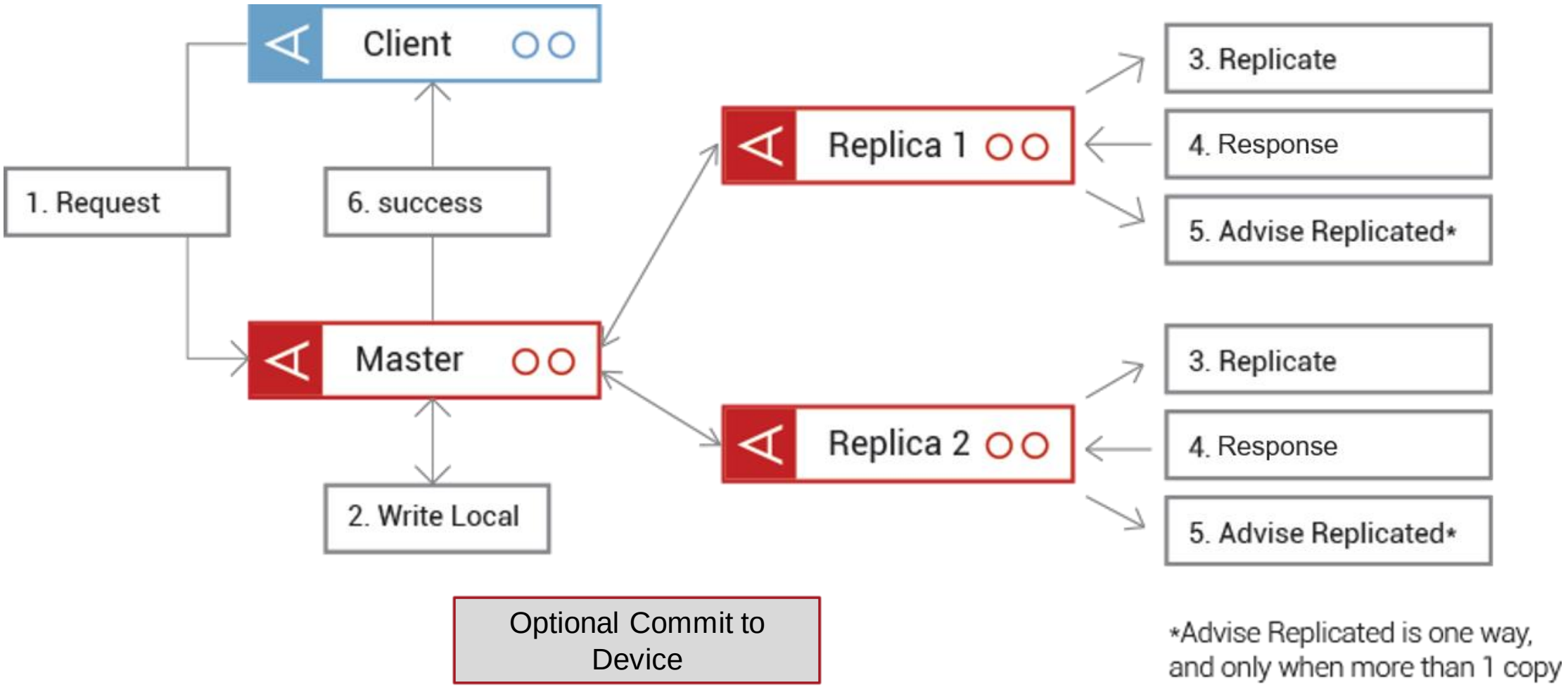
SPLIT – Rule 1; ***p is active***
All designated replicas in a sub-cluster, p active

SPLIT – Rule 2; ***p is active***
One designated replica, in a majority sub-cluster

SPLIT – Rule 3; ***p is inactive***
Majority has no designated replicas, minorities don't have all replicas

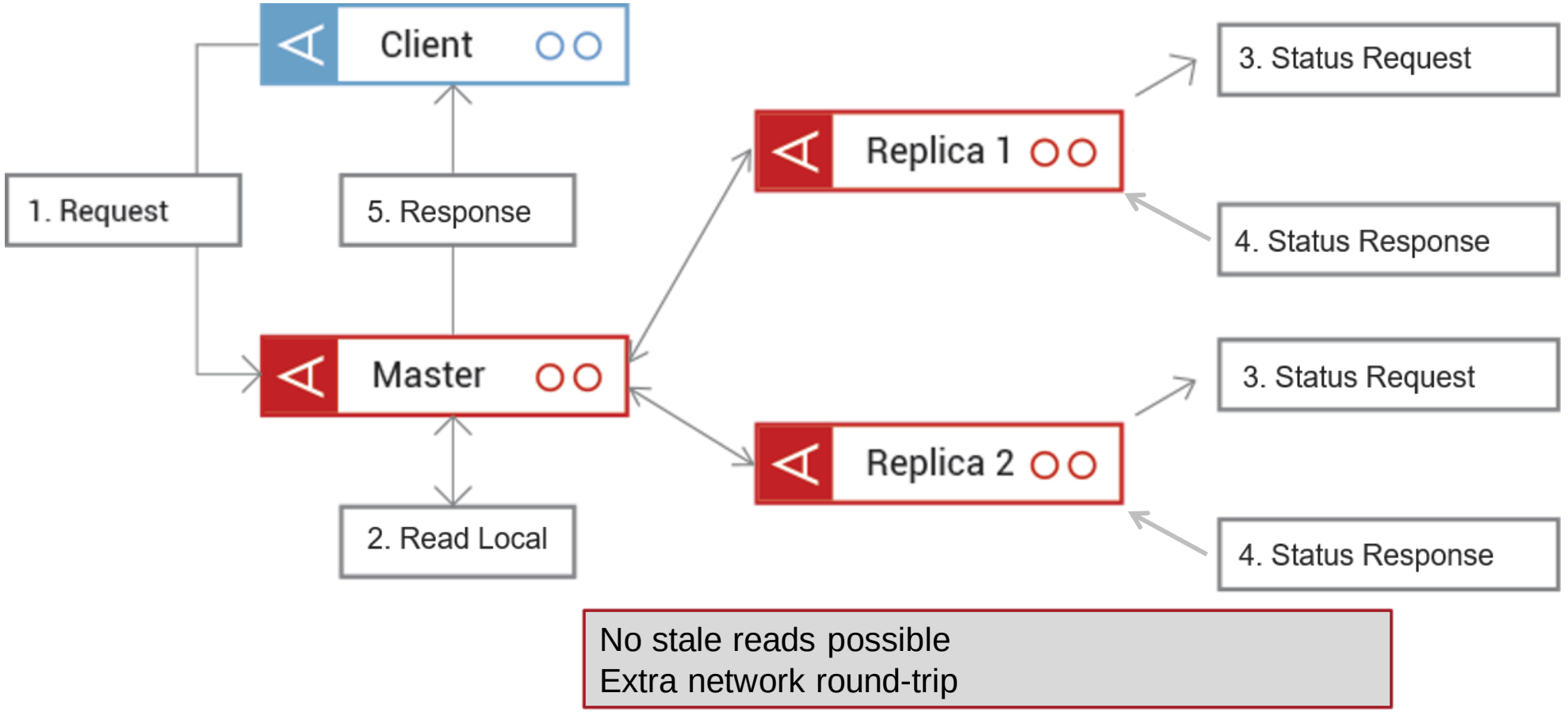
AEROSPIKE

Aerospike 4.0: Write Logic (2PC)



AEROSPIKE

Aerospike 4.0: Linearizing Reads

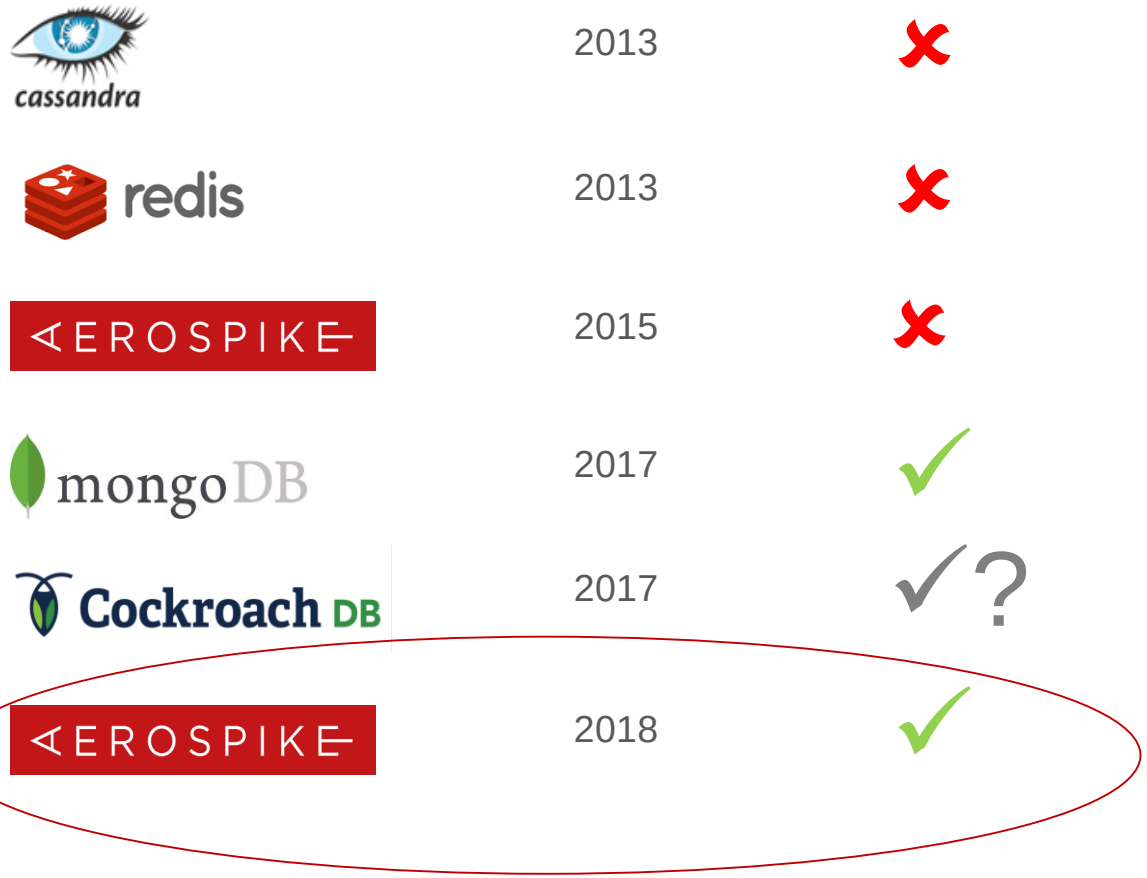


Aerospike 4.0: Jepsen Test Confirms Strong Consistency

"Aerospike does appear to provide linearizability through network partitions and process crashes"

--- Kyle Kingsbury, Jepsen.io

http://jepsen.io/analyses/aerospike-3-99-0-3



Aerospike 4.0: High Performance with Strong Consistency

Aerospike internal benchmark of Strong Consistency versus Availability

	Linearizable Consistency	Sequential Consistency	Availability
OPS	1.87 million	5.95 million	6 million
Read Latency	548 µs	225 µs	220 µs
Update Latency	630 µs	640 µs	640 µs

In-memory configuration with persistence enabled

5 node cluster
500M keys

Replication factor 2
Objects were a 8 byte integers

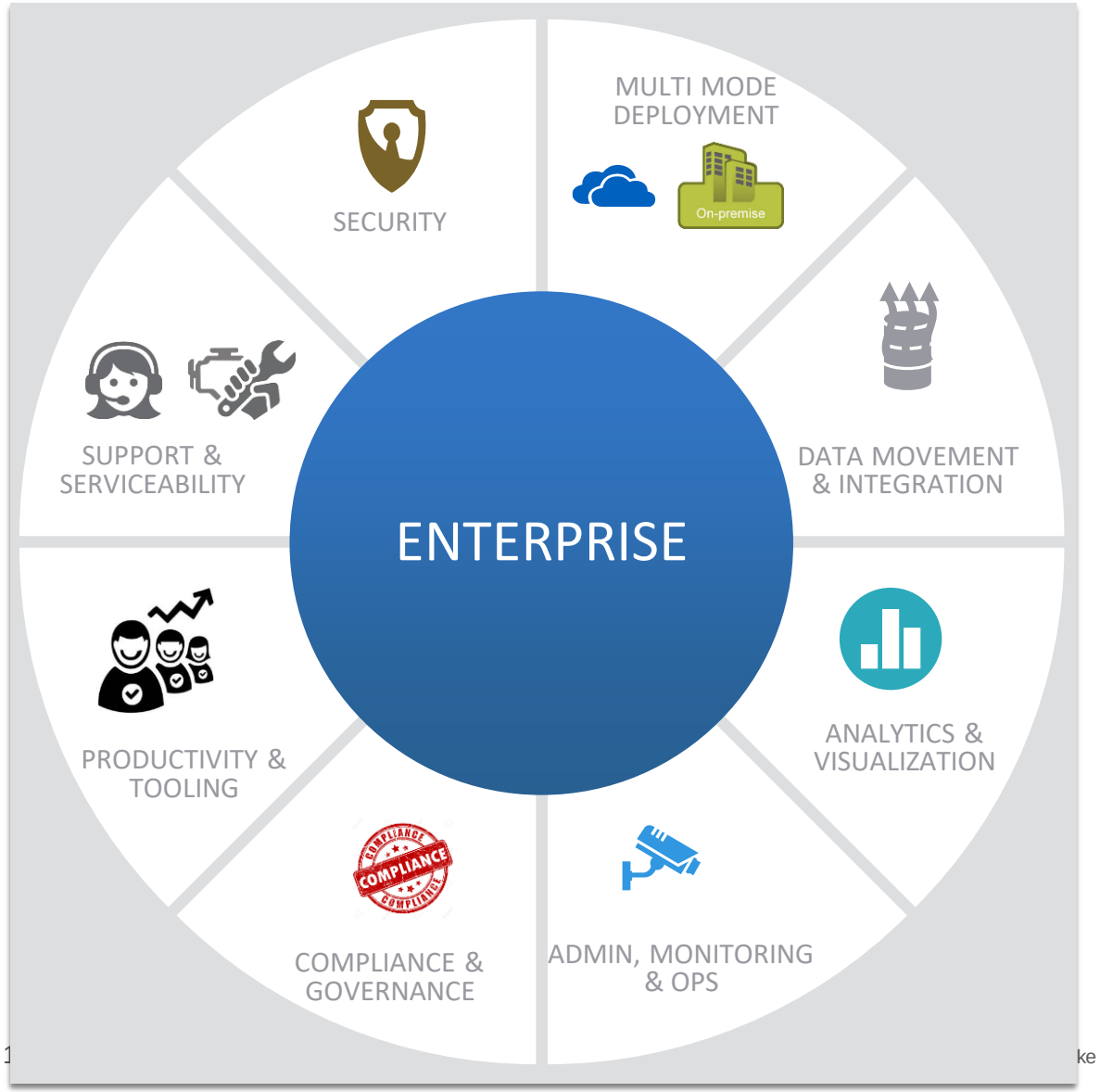
Aerospike Ecosystems

May 23, 2018

Overview

- Aerospike for Enterprises
- Ecosystems
- Real-time Insights
  - Trends
  - Challenges
  - Components
- Aerospike Real-Time Analysis Framework
  - Reference Architecture
  - Use Cases

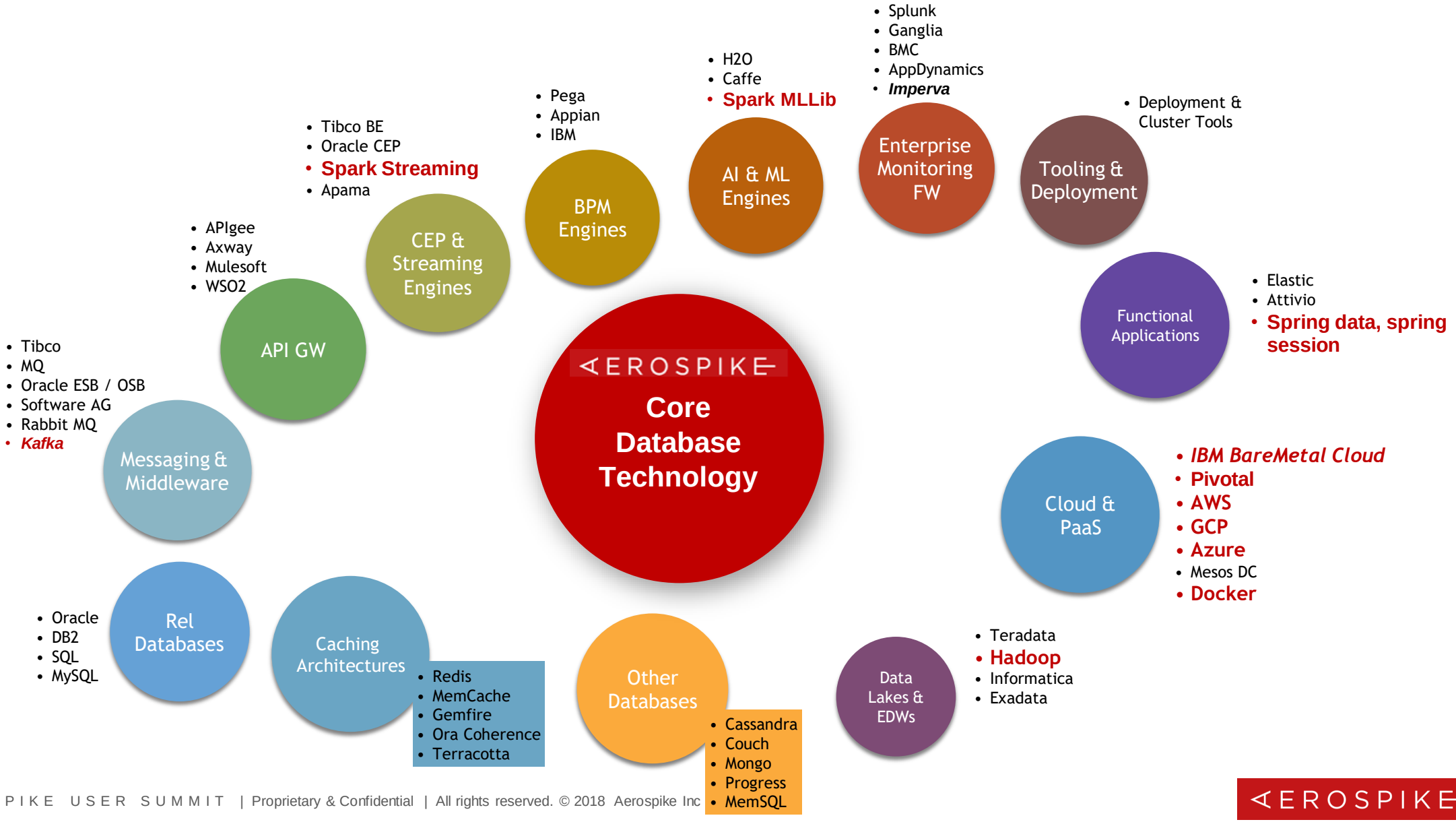
Frictionless experience in Enterprises - What we believe it would take?



Frictionless experience for Enterprise demands interoperability, coexistence, manageability

AEROSPIKE

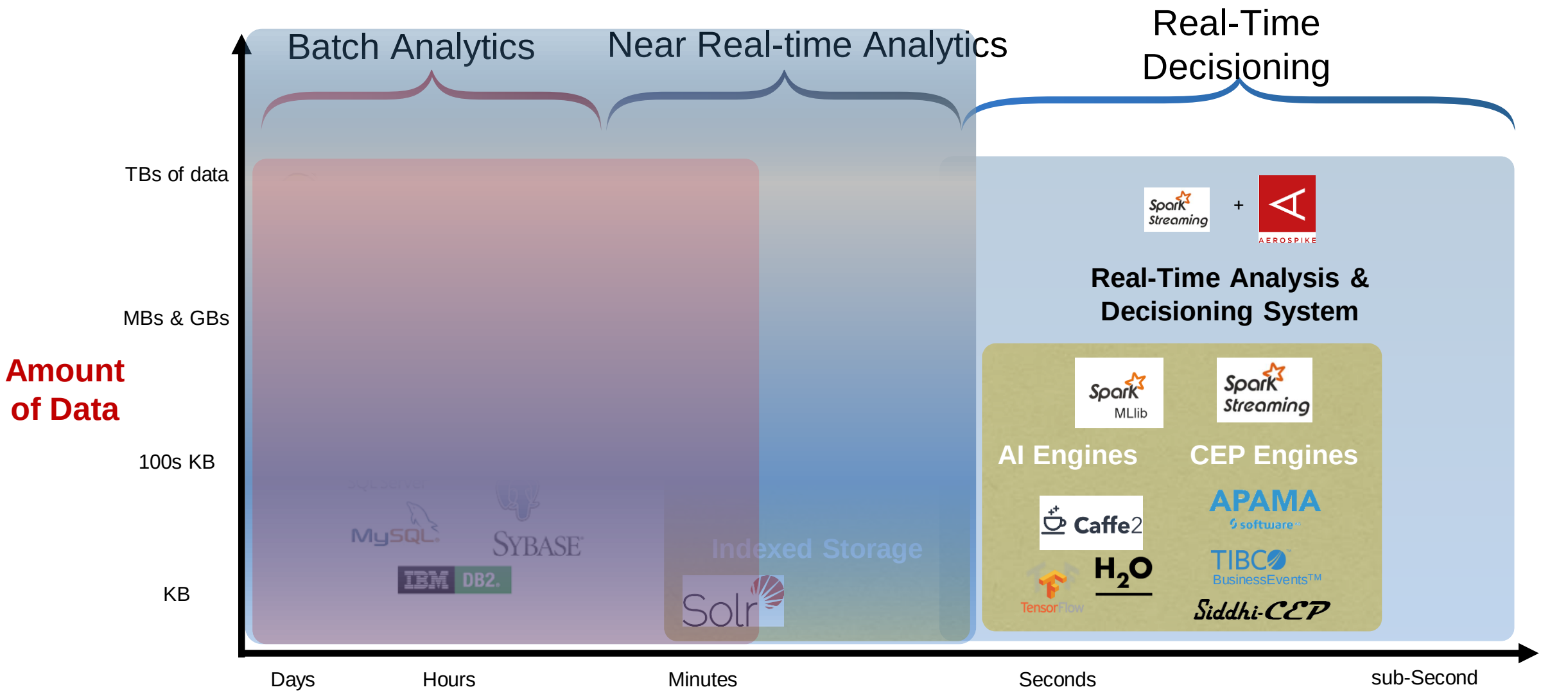
Aerospike related ecosystems in an Enterprise



Announcing Availability of

Aerospike Real-time Analysis Framework [RAF] 1.0

Analytics to Analysis

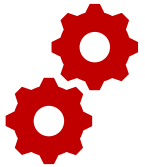


AEROSPIKE

Real-time Business Insights - Challenges

Need to Combine Data

- Analysis needed on combined data (stream + transactional)
- Transactional data enhances stream data = more insights
- Storage needed to complement both stream & transactional data
- Difficult to scale systems (designed for low scale)

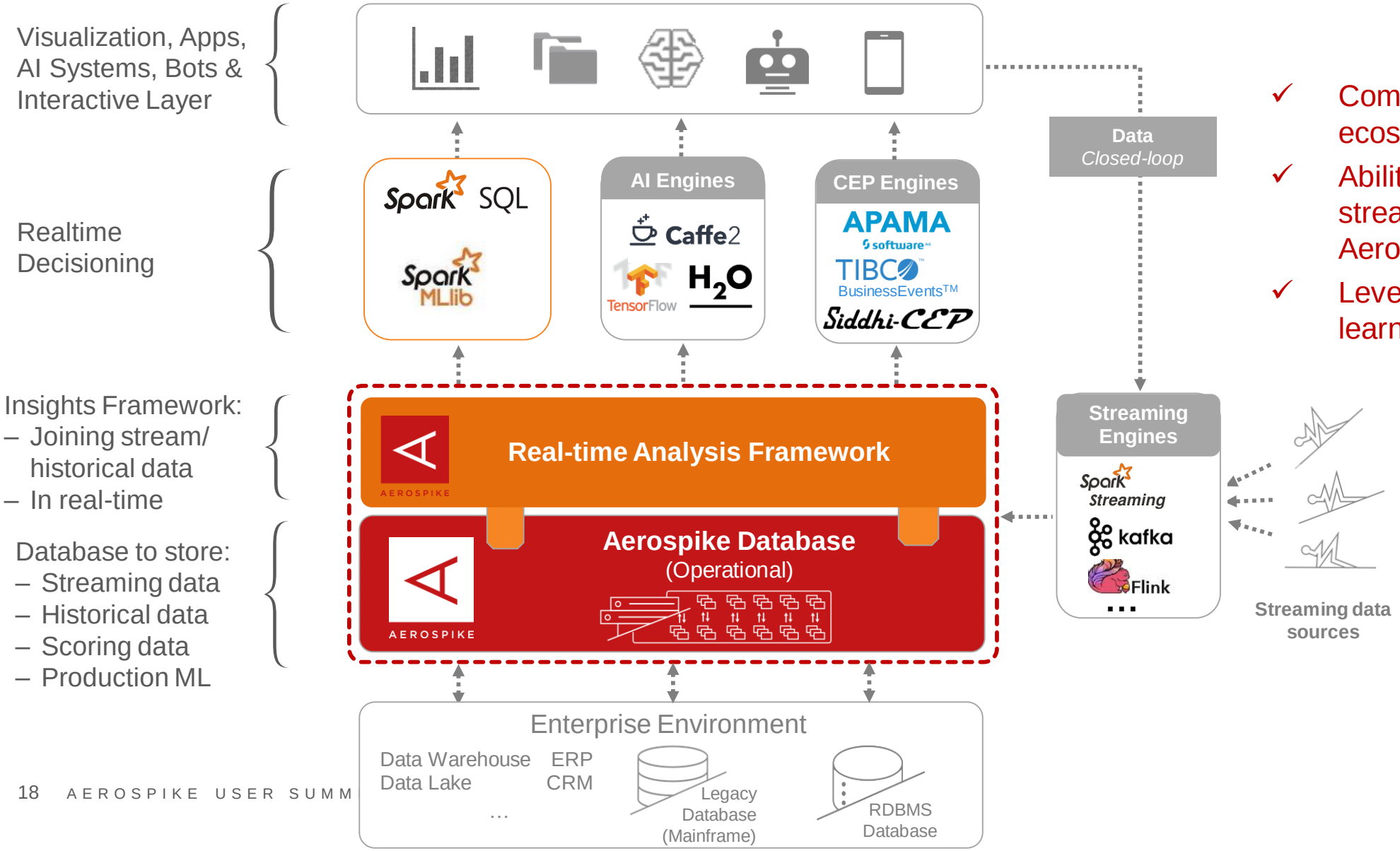


ML needed for Stream + Transactional data

- Complex server/cluster deployments
- Complex Ops mgmt.
- Larger in-memory foot print leads to less actionable data

Introducing a new approach for Real-time Insights:

Aerospike Realtime Analysis Framework (RAF)



- ✓ Compatible with Spark ecosystem
- ✓ Ability to join spark streaming data with Aerospike data
- ✓ Leverage spark machine-learning for analysis

AEROSPIKE

Features and Benefits of RAF 1.0

Key Features

- Ability to join stream and transactional data in near real time
- Perform join & intersect operations on Aerospike datasets with streaming data
- Reduced friction to developers to develop analytics applications using Spark
- SQL interface via spark SQLLib for Aerospike datasets
- ML-driven operations using SparkML [support for other ML libraries]

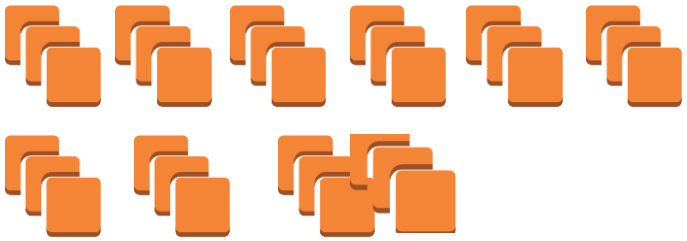
Customer Benefits

- Expanding the “frame of data” that is processed in real-time by joining stream data from Spark with transactional data from Aerospike
- Lower TCO for real-time analytics by operating on larger datasets yet with a smaller cluster footprint
- Gain closed-loop business insights by operating on both transactional and stream datasets
- Rapidly develop using Spark libraries - no additional skill set required
- Accelerate business insights by enabling decisions in seconds as opposed to hours or days

A Sample Deployment Scenario

- ✓ Total of 1Bn Records with record size of 1.5K : Total 1.4TB data
- ✓ Processing 5% of total records for every 30min window : 50 mn events, ~28KTPS
- ✓ CEP Engine with all in-memory processing ability vs RAF+Aerospike

CEP data In Memory



32 Nodes required for processing : r4.2xl

$ 199,000 with 3-yr upfront reserved instances

RAF add-on with Aerospike

4 Nodes required for processing : i3.2xl

$ 32,933 with 3-yr upfront reserved instances

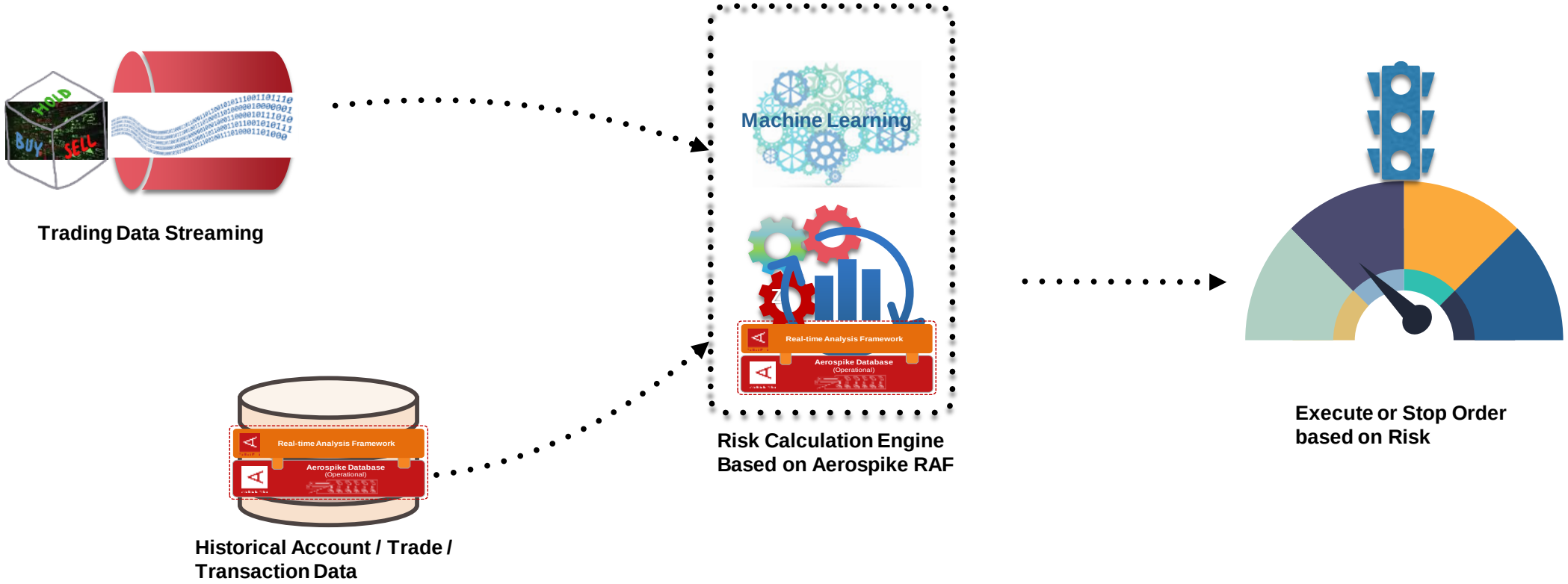
5x Operative cost savings on infra [$167K]

Let us say the case is not for 1.4TB but for 10TB, that is about $1.2mn in infra cost savings

AEROSPIKE

RAF Use Cases

Risk Management in Capital Markets

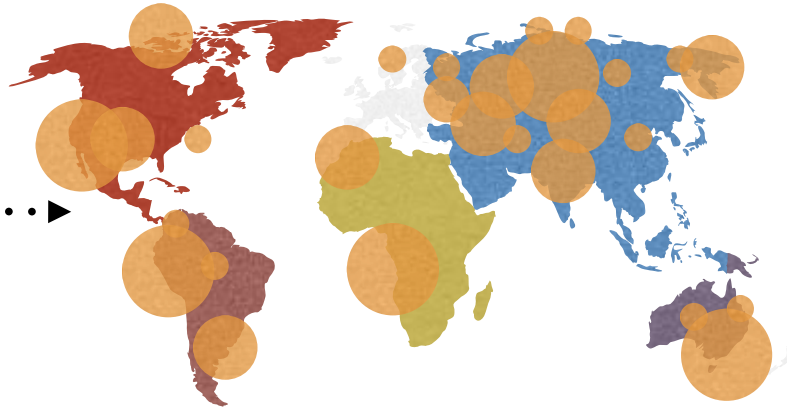


Risk calculation in real-time during market hours, processing insights on combined trade stream data and historical customer data for risk monitoring in real-time

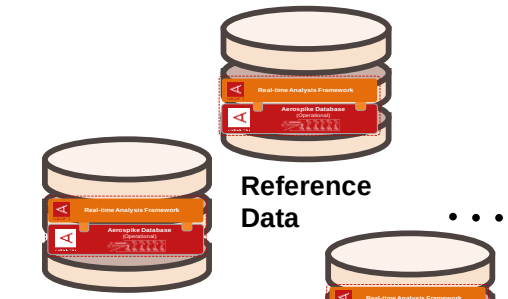
Fraud Detection

Card Processing Data Stream

Fraud Detection Engine



Fraud O Meter
process or stop payment



Reference Data

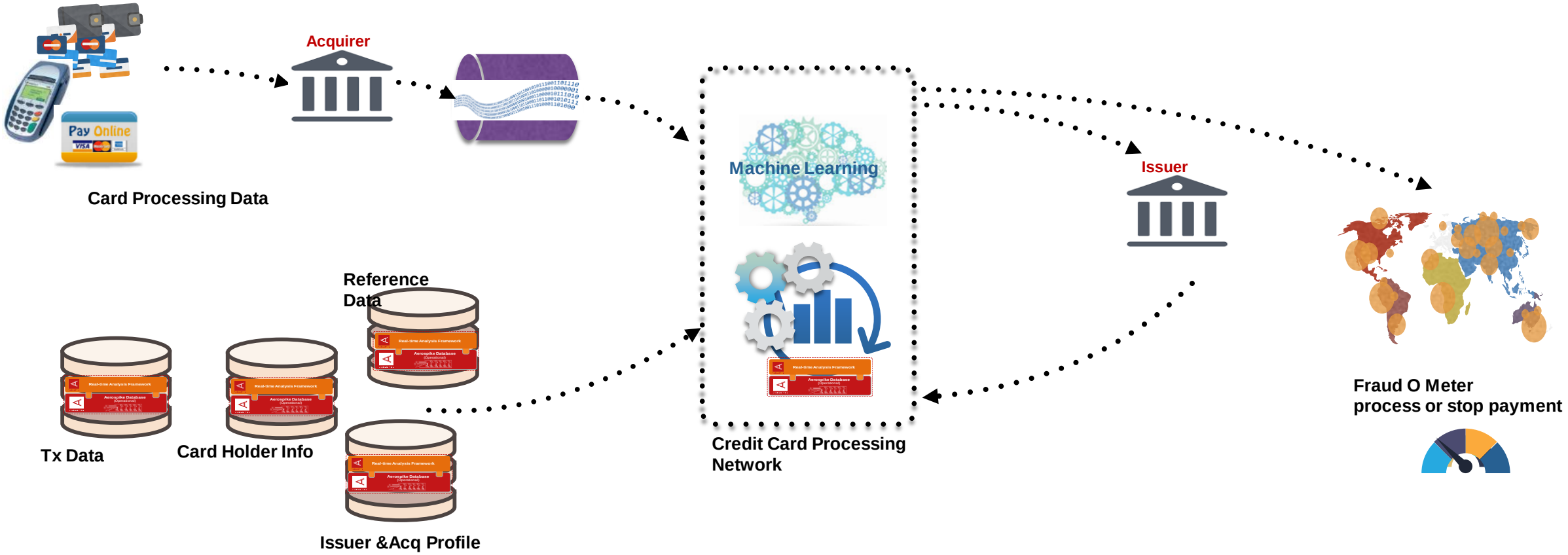
Historical Transaction Data

TP Data

Fraud and false positive identification in real-time for every payment event under ~750ms, monitoring fraud & false positives in real-time

AEROSPIKE

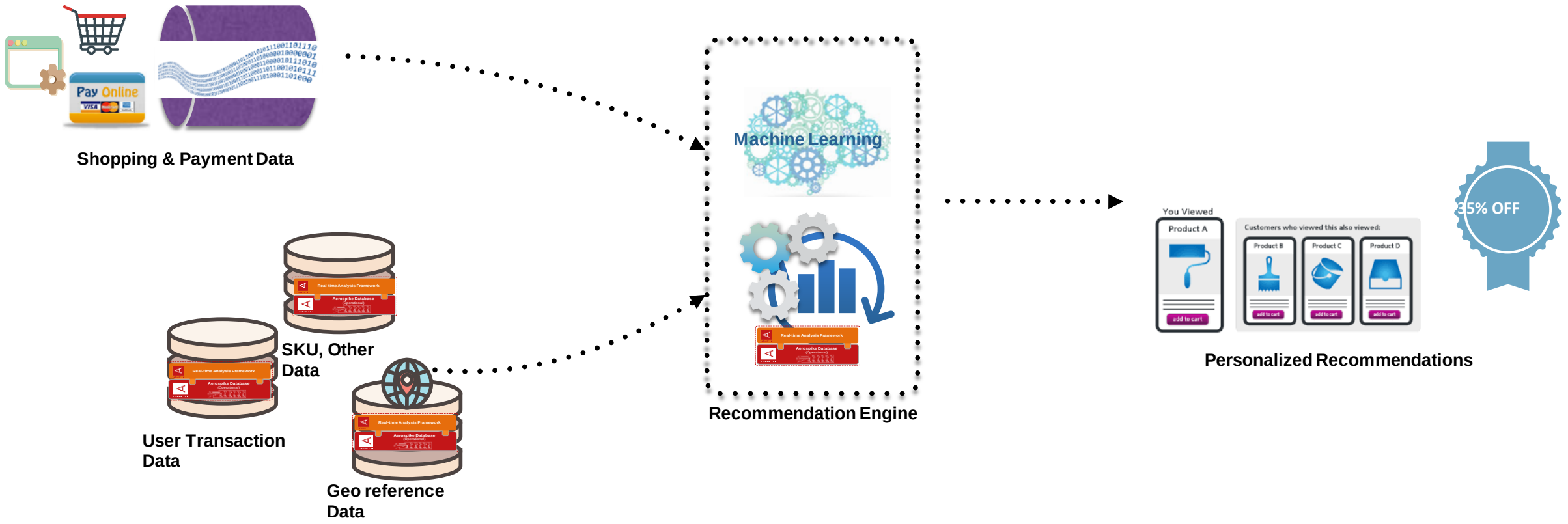
Credit Card Processing Network



Authorization of CC processing with mean latency of ~13ms, enabling processing of transactions through Network

AEROSPIKE

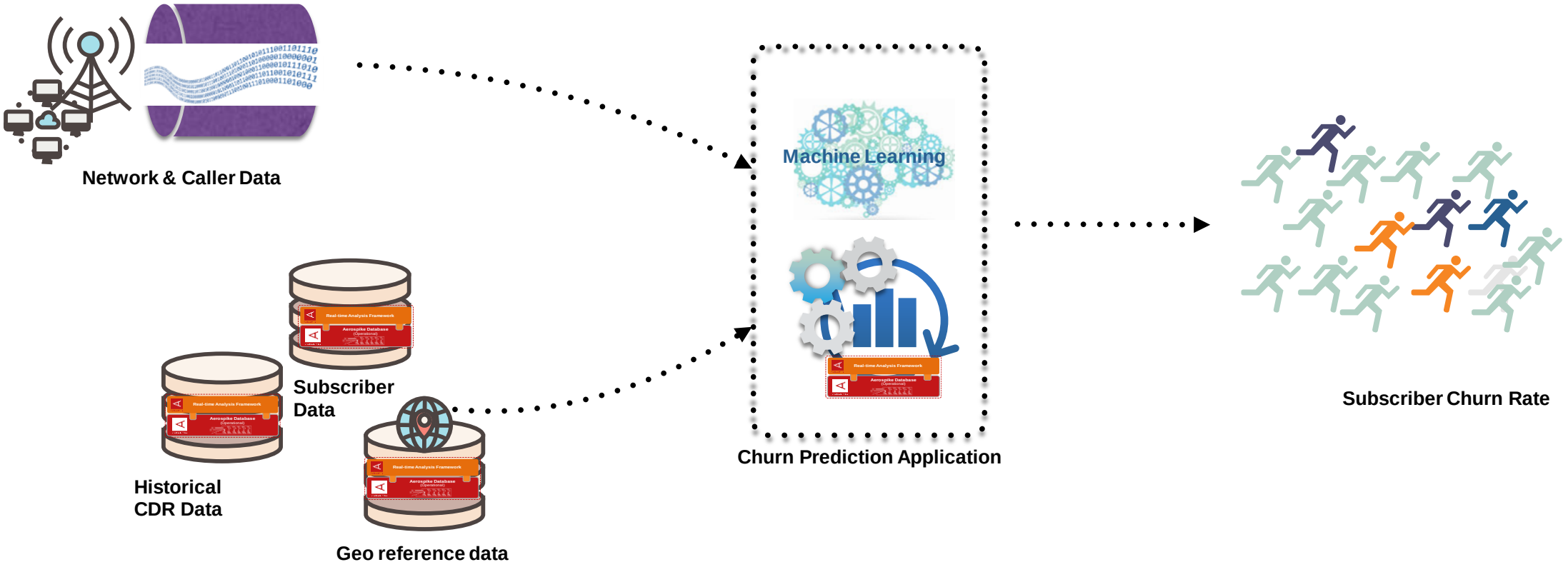
Personalization & Recommendation Engines



Product recommendations and personalized pricing, based on shopper behavior data, buying patterns, product movement, transaction data etc

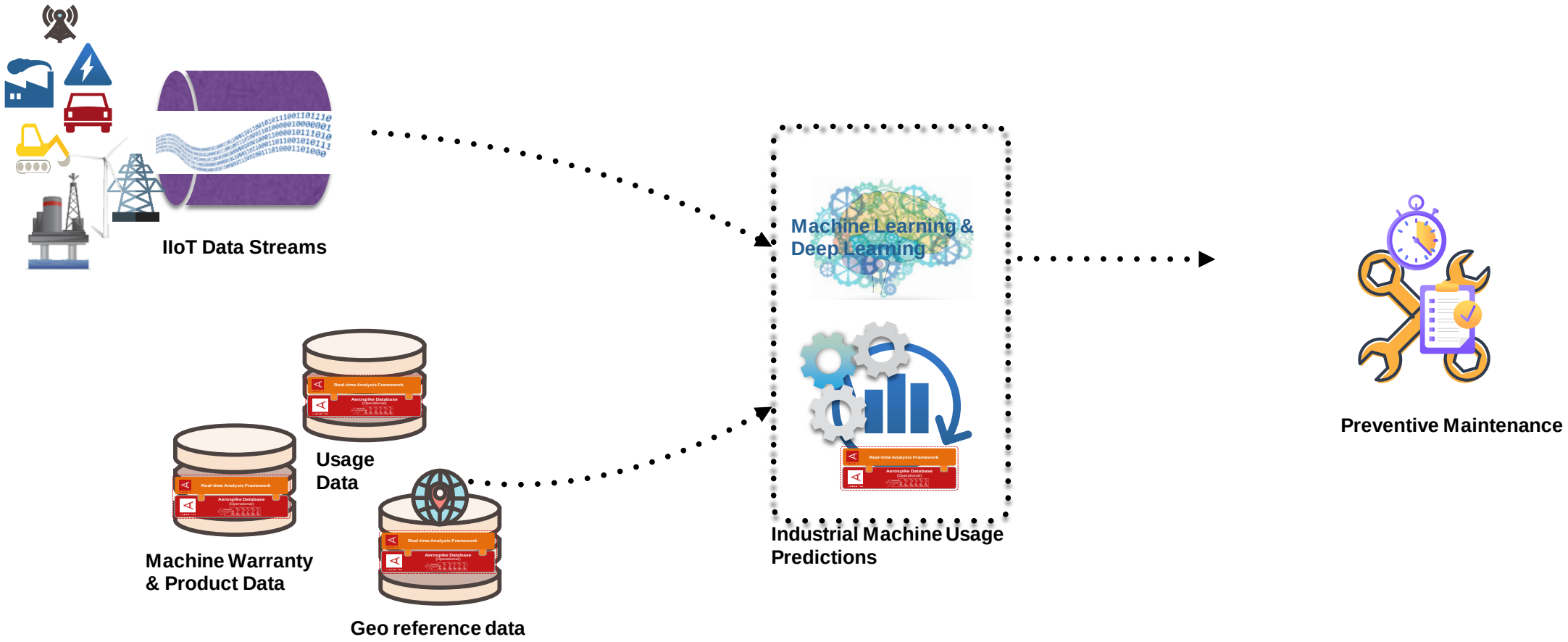
AEROSPIKE

Real-Time Churn prediction in Telco



Predicting churn in customer base using, data like CDR, stream, network, geo reference

Industrial IoT



Predicting when to service & maintain the industrial equipment, based on usage, machine data, warranty, consumption etc

AEROSPIKE

Thank You

