Aerospike ecosystem in the real-time enterprise

Edge (ms)

Core (ms)

Warehouse (sec-to-mins)

Predictive Analytics

Spark Streaming

AI/ML Engines

XDR

Spark SQL

Location A (SOE)

Location B (SOE)

Location C (SOE)

Real-time System of Record

Single Source of Truth

Query & Reporting Store

kafka

Legacy Data Store

TB's

PB's

100's PB's
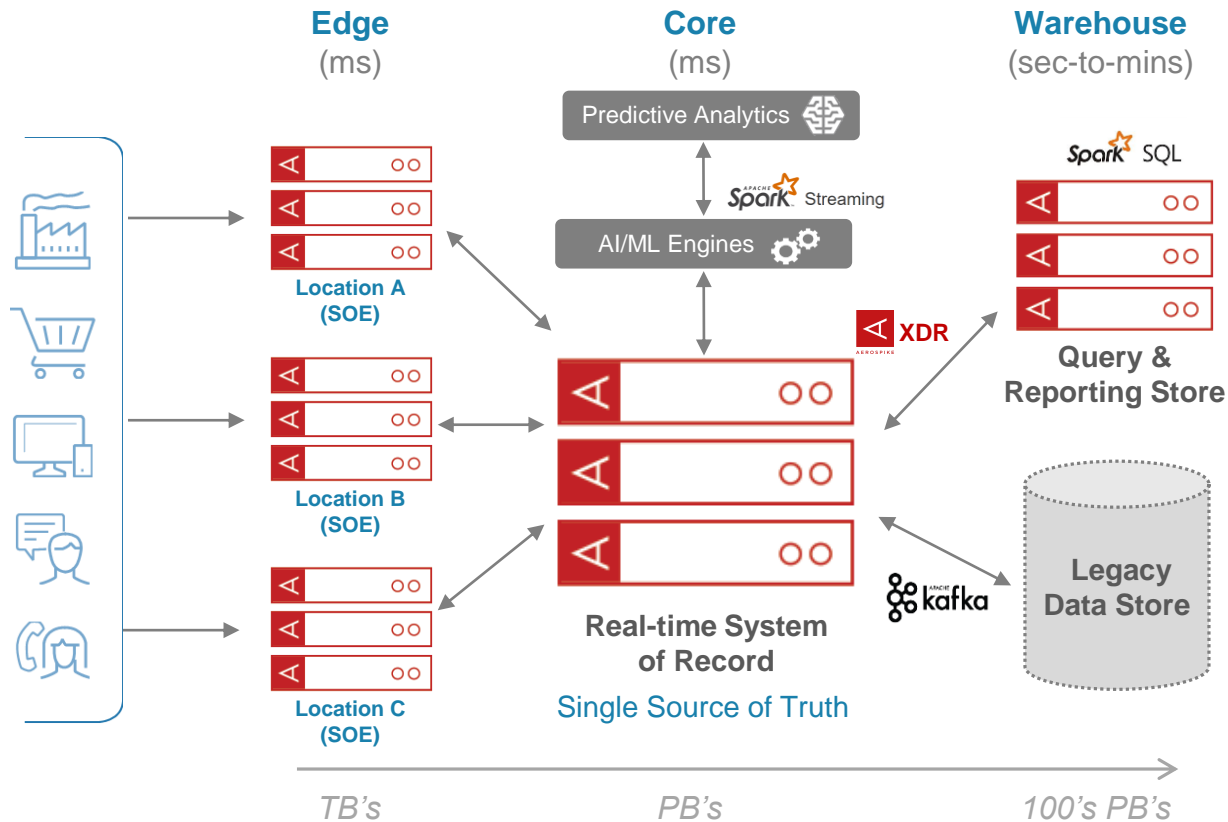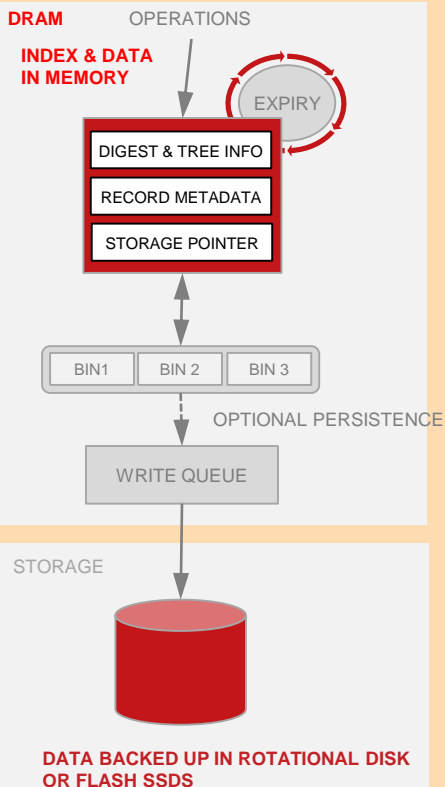
– **Consumers require customized real-time user experience**

– **DB must deliver predictable performance at scale**

✓ Throughput > 1M TPS

✓ Latency < 1ms

✓ Reliability > five 9s

✓ Scale up to petabytes

✓ Strong consistency

✓ TCO < 2-5X of other DBs

AEROSPIKE

SUMMIT '19

Hybrid-Memory Architecture Delivers Scale
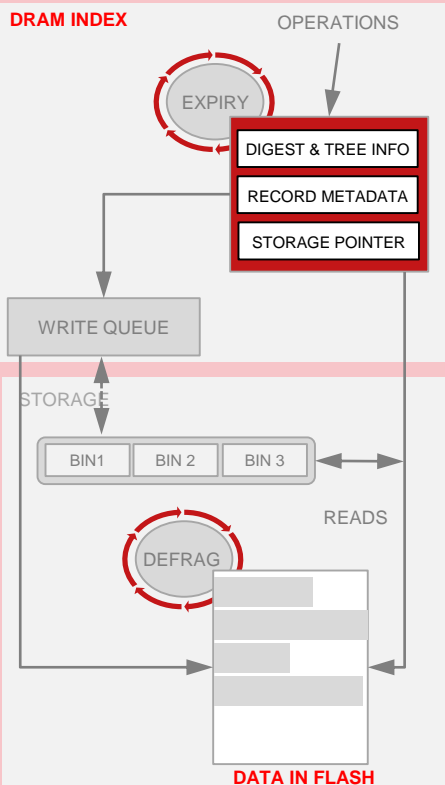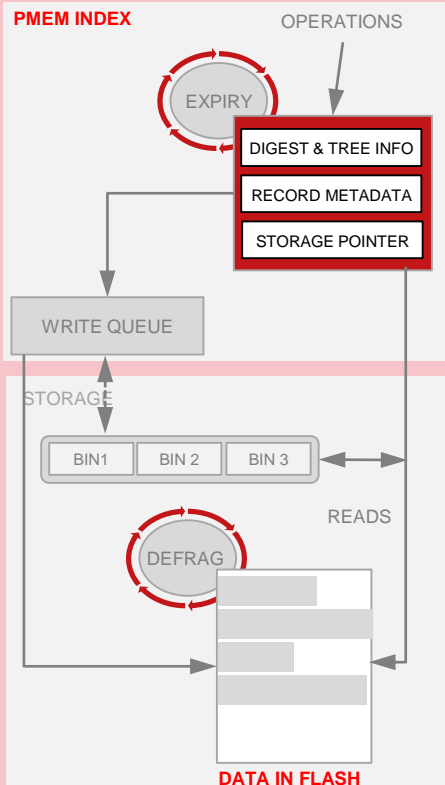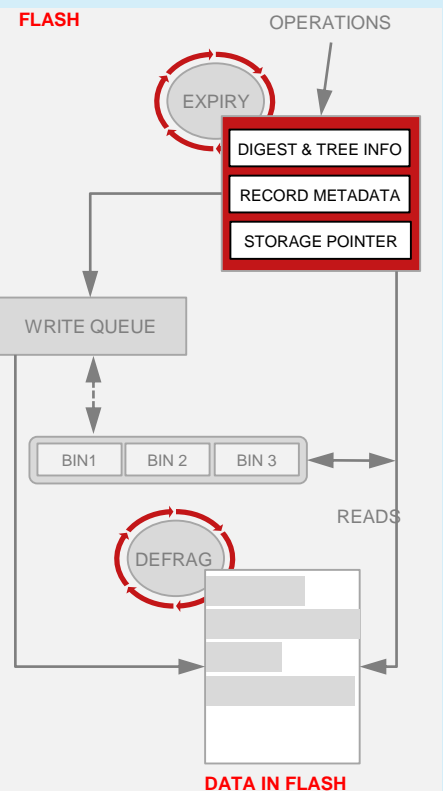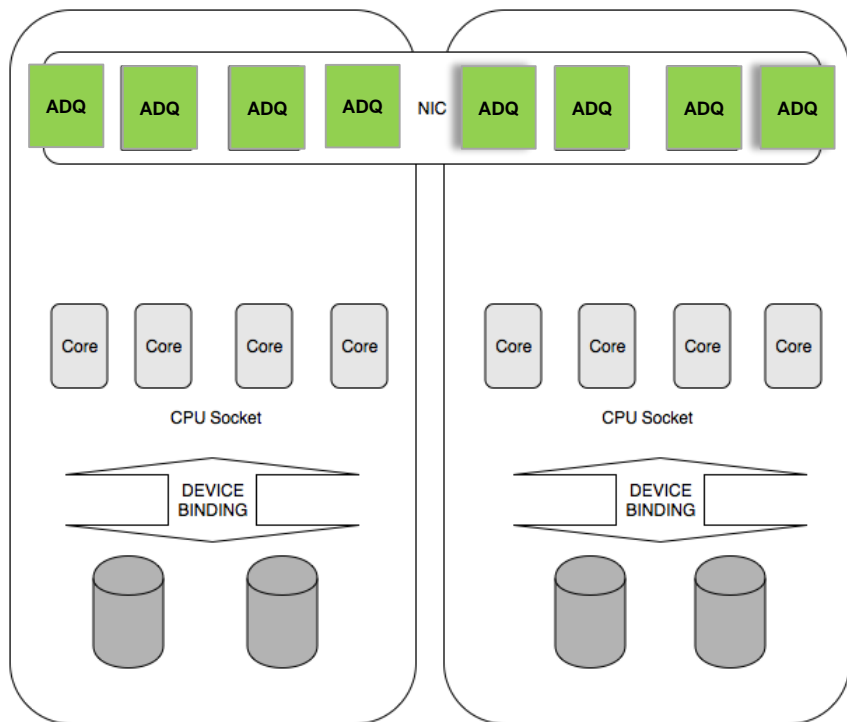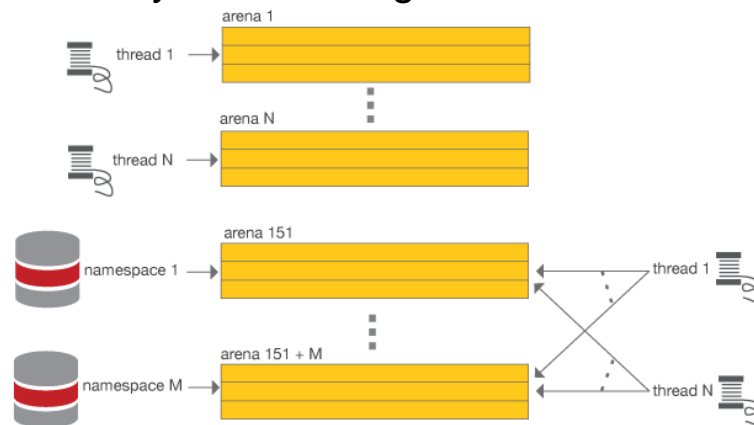
Designed for wire-line speed

Multi-core Architecture

Optimized C based DB kernel

1. Multi-threaded data structures (NUMA pinned)
2. Nested locking model for synchronization
3. Lockless data structures
4. Partitioned single threaded data structures
5. Index entries are aligned to cache line (64 bytes)
6. Custom memory management (arenas)

Memory Arena Assignment

Delivered over the past year

$ = TCO Savings

Q2 2018	Q3 2018	Q4 2018	Q1 2019
LDAP	All flash $	Node quiescence & delay fill migrations	Relaxed strong consistency reads
Storage optimization $	Uniform partition balance $	Rack aware reads $	Improved data expiry & system metadata
8MB objects		Change notification	Aerospike Connect for Spark
		Record storage $ compression	Aerospike Connect for Kafka
		Index in persistent memory $	Aerospike REST Client

AEROSPIKE

SUMMIT '19

Benefits (a few highlights)

- **No application level SLA impact during maintenance (no timeouts)**
  - Allows operators to protect their running applications

- **Storage went down by 30% in production**
  - Several customer deployments benefited

- **Uniform balance**
  - Increased effective capacity of existing large cluster by > 20%

- **Ease of integration**
  - Aerospike as a SOR can now send real-time changes to other enterprise databases

Use Case: Target Instant Payment Settlement (TIPS)

BANCA D'ITALIA
EUROSISTEMA

- **Needs**
  - True end-to-end user payments via their bank within a matter of seconds
  - All Euro Consumer Transactions
  - 24x7x365 availability
  - Costs under €0.01 per payment
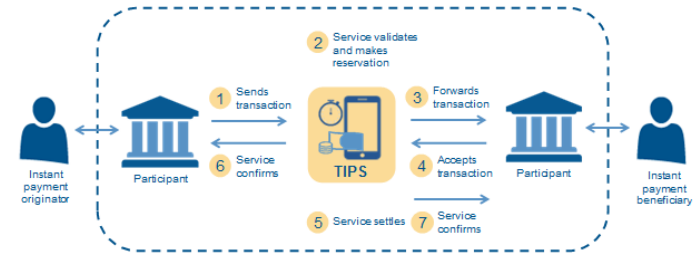
TARGET instant payment settlement (TIPS)

- **Challenges**
  - Existing solutions are either expensive and/or slow:
    - VISA charges between 2% and 3% $15 Billion in profits per year
    - Blockchain is dollars & 10 minutes per trans
  - Private solutions will cherry-pick profitable regions
  - Private solutions sell data and leak information

- **Why Aerospike**
  - Prevents bottlenecks in persistence layer
  - Integrates with Oracle for long-term storage
  - 24/7 availability with High Availability built-in
  - 10k writes/sec with lowest possible latency
  - Cross-datacenter replication
  - Strong support
  - Scalability for projected 1000+ European banks
  - Able to keep track of failures (history of payments)/ replay if in case of failure

"In hours I had my dev environment"

- Vitangelo Lasorella, Technical Officer, IT Innovation and Development Department.

AEROSPIKE

AEROSPIKE SUMMIT '19

Aerospike Core

Speed and Scale (1 of 2)

- **Relaxed consistency (for reads)**
  - Reads allowed from master or replica nodes
  - Reads allowed from partitions not available for writes
  - Therefore, some reads can violate session/sequential consistency
  - No Dirty Reads

- **Efficient storage/network data formats**
  - Use the storage format on the wire
  - Use compressed data on the wire
  - Improves CPU, memory and network utilizations

Speed and Scale (2 of 2)

- **Parallel Scans**
  - Allow large and small scan jobs to be processed in a pipelined manner so all scans make progress
  - Enables small scan jobs queued behind a large scan job (e.g., database backup) to make progress without waiting for the larger scan to complete

- **Managing quotas**
  - Cap transaction and scan/query resources for individual users/applications
  - Allows multi-tenancy of applications on a single Aerospike service
  - Prevents a new application from interfering with existing applications

- **All PMEM**
  - Allows Aerospike index and data to both be stored in PMEM
  - PMEM is used as the persistent store and no additional Flash is needed

AEROSPIKE

AEROSPIKE SUMMIT '19

XDR & support for tiers

- **Phase I**
  - Integrate into core database and eliminate separate digest log
  - Enable independent shipping from single source to target data centers
  - Dynamic configuration of shipping information
  - Improved correctness in failure situations
  - Record/Bin level filters
  - Support for shipping from past timestamp for reliable data sync across clusters

- **Phase II**
  - Support for strong consistency in active/passive setup
  - Optimizations to improve shipping performance
  - Support seamless data movement across clusters

◁EROSPIKE◁

◁EROSPIKE◁
SUMMIT '19

Query & secondary index

- **Core efficiency improvements**
  - 3X reduction in index size
  - 4X improvement in performance
  - Better garbage collection

- **Reliable Queries**
  - Ability to query accurately in the presence of cluster changes
  - Support for restarting queries from a specific point

- **Secondary index operational improvements**
  - SLAB style allocation
  - Warm start

- **Secondary index storage improvements**
  - All flash support
  - PMEM support

Advances in programming model

- **Multi level complex data types (CDT)**
  - Extends the map and list operations API to support arbitrarily nested maps and lists
  - Define a way to identify the target list and map within the nested structure

- **Bitwise operations**
  - Atomic bitwise operations on the BLOB data type
  - Modeling with BLOBs should be highly compressible

- **Conditional key-value operations**
  - Conditional writes
    - Only write/execute a multi-operation transaction if a predicate filter is satisfied
  - Conditional reads
    - Only read a record if a predicate filter is satisfied (e.g., last updated under N seconds ago)
    - Filter records of a batch-read using a specified filter (aligns batch-reads with scan and query)
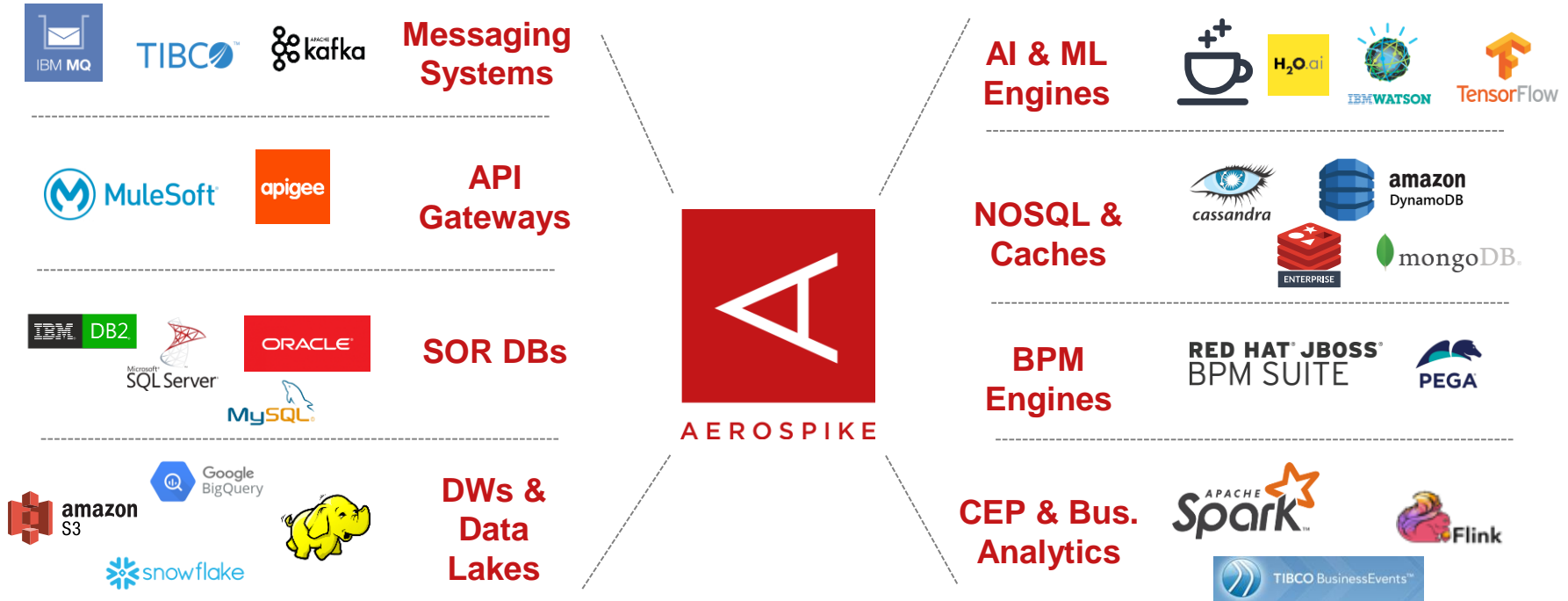
Aerospike Connect & Cloud

AEROSPIKE

AEROSPIKE
SUMMIT '19

Aerospike Connect – support for your data pipelines

Interoperate with Datastores

Power Existing Apps

Messaging Systems

API Gateways

SOR DBs

DWs & Data Lakes

AI & ML Engines

NOSQL & Caches

BPM Engines

CEP & Bus. Analytics

Reduce friction with interoperability & coexistence - with Aerospike Connect product line based GTM

Aerospike Connect

- **Real-time focus**
  - Connect for JMS
    - Solace and IBM MQ

  - Connect for Flink
    - Support for Java, Python and Scala

  - Connect for Spark
    - Support for PySpark

  - Connect for Elastic Search

- **Connect for AI/ML**
  - Connect for AI/ML: PyTorch
  - Connect for AI/ML: TensorFlow
  - Connect for AI/ML: SciKit

- **Easy integration**
  - REST Client
  - Connect for Hadoop (refresh)
  - Spring boot cache manager

AEROSPIKE
SUMMIT '19

Aerospike Cloud

"Make Aerospike as easy to operate as cloud native databases"

- **Aerospike Command Center API**
  - Cluster management
  - Monitoring
  - Alerting
  - Scale up & scale down
  - Rolling upgrade
  - Backup & restore

- **Aerospike Command Center UI**
  - For fast developer adoption and production use

- **Aerospike Cloud Backup**
  - S3, Google cloud storage

- **Kubernetes**
  - Aerospike Kubernetes operator
  - Update Kubernetes deployment for Google & Azure

- **Tools/Plugins**
  - Data browser
  - Prometheus exporter

Longer term

Next set of candidates

- **Aerospike Enterprise Edition**
  - CRDT support between active/active clusters
  - Native support for aggregate functions (min, max, avg, etc.)
  - UDF (WebAssembly, Java, …)
  - Probabilistic data types
  - Distributed global multi-record transactions
  - CoreDNS integration for updating node IP list, finding seed node, etc.
  - Multi-level clustering for 100X, 1000X scale
  - Time series support
  - Kerberos support

- **Aerospike Connect**
  - Generic integration with graph database
  - Fluentd logging integration
  - Other CNCF integration

- **Aerospike Cloud**
  - Aerospike as a service
  - Deeper integration with Kubernetes engines

Questions