Boost Software and SaaS with In-Memory Computing

Matt Sarrel
Director of Technical Marketing
GridGain Systems
matt.sarrel@gridgain.com
@msarrel

Agenda

- Introduction
- In-Memory Computing
- GridGain / Apache Ignite Overview
- Survey Results
- Use Cases and Case Studies
- GridGain / Apache Ignite In-depth

Your Presenter

- Director of Technical Marketing at GridGain Systems
- 30 years in tech

- Matt.sarrel@gridgain.com
- @msarrel
- www.gridgain.com/resources/blog

Trends in Software and SaaS

- Open source
- Software as a critical component of business
- Businesses make data-driven decisions
- Data: open, streaming, data lakes
- Performance, stability, security, scalability
- Serverless architectures
- Distributed systems
- Containers
- Microservice based architectures
- Machine learning
- Self-service bots
- Augmented and virtual reality

IMC a key component of "real SaaS"

- http://blogs.workday.com/naomi-blooms-4-foundations-of-real-saas-no-4-in-memory/
  - Memory used to be the one thing that needed to be protected at all costs because memory was so expensive we had so little of it
  - Now we can get rid of the idea that we can only do reporting on off-cycles with separate data
  - IMC needs to be built in from the ground up - for example the analytics can be built into the tools that line of business people access
  - One source of truth, all the data is available all the time
  - Greater and more timely insight, operate independently to make decisions on the fly

Why In-Memory Now?

Digital Transformation is Driving Companies Closer to Their Customers

- Driving a need for real-time interactions

Internet Traffic, Data, and Connected Devices Continue to Grow

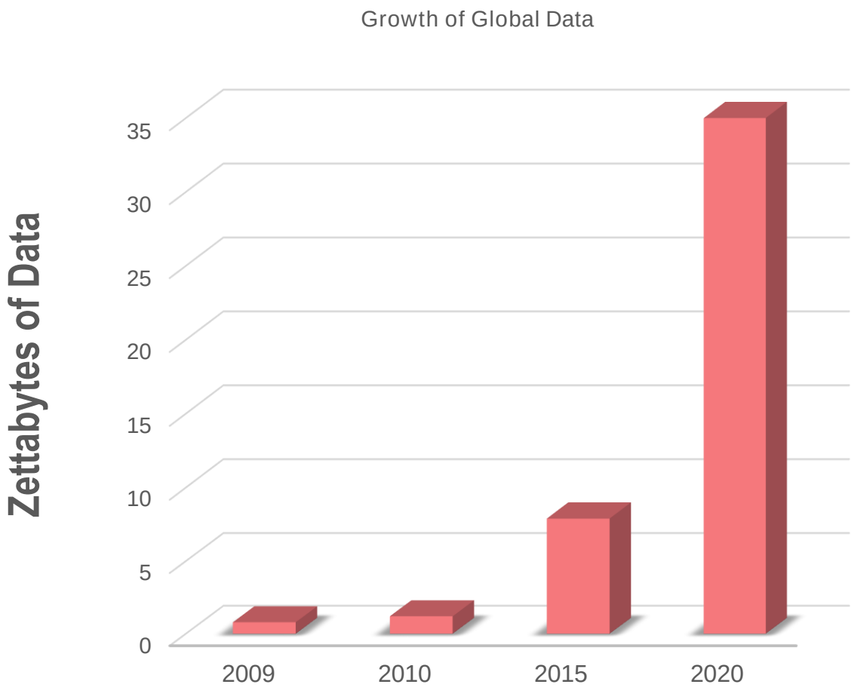
- Web-scale applications and massive datasets require in-memory computing to scale out and speed up to keep pace
- The Internet of Things generates huge amounts of data which require real-time analysis for real world uses

The Cost of RAM Continues to Fall

- In-memory solutions are increasingly cost effective versus disk-based storage for many use cases

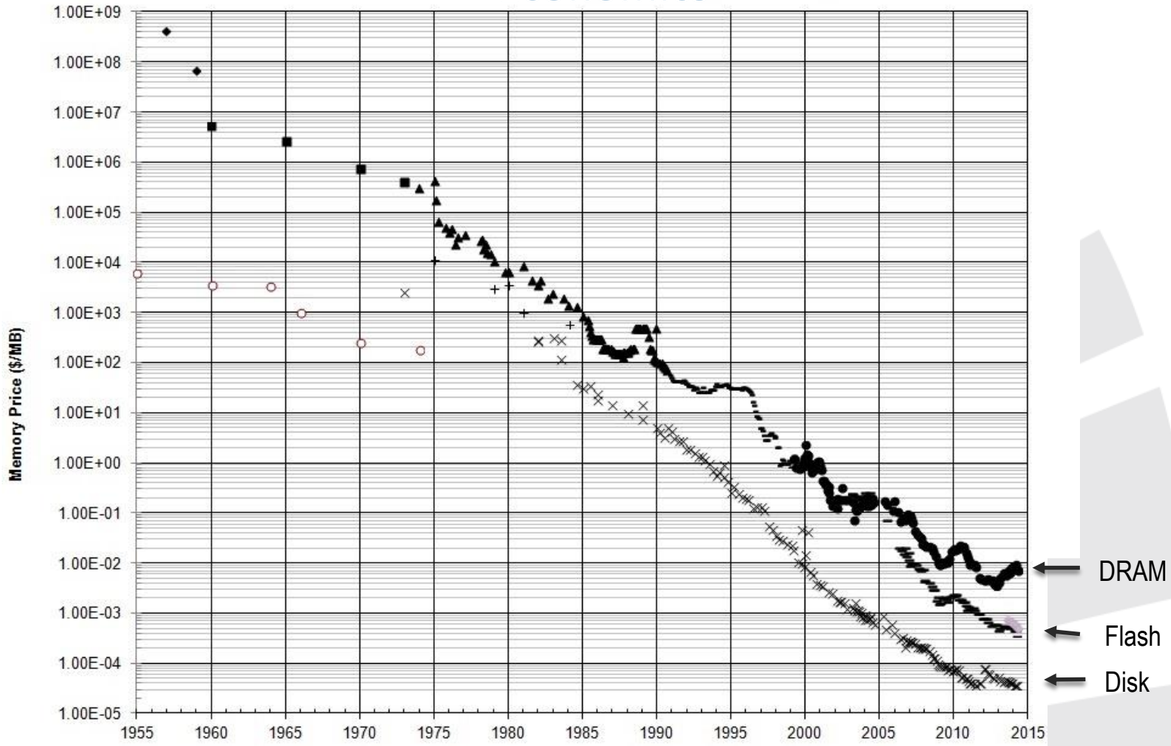
Why Now?

Data Growth and Internet Scale Driving Demand



8 zettabytes in 2015 growing to 35 in 2020

Declining DRAM Cost Driving Attractive Economics

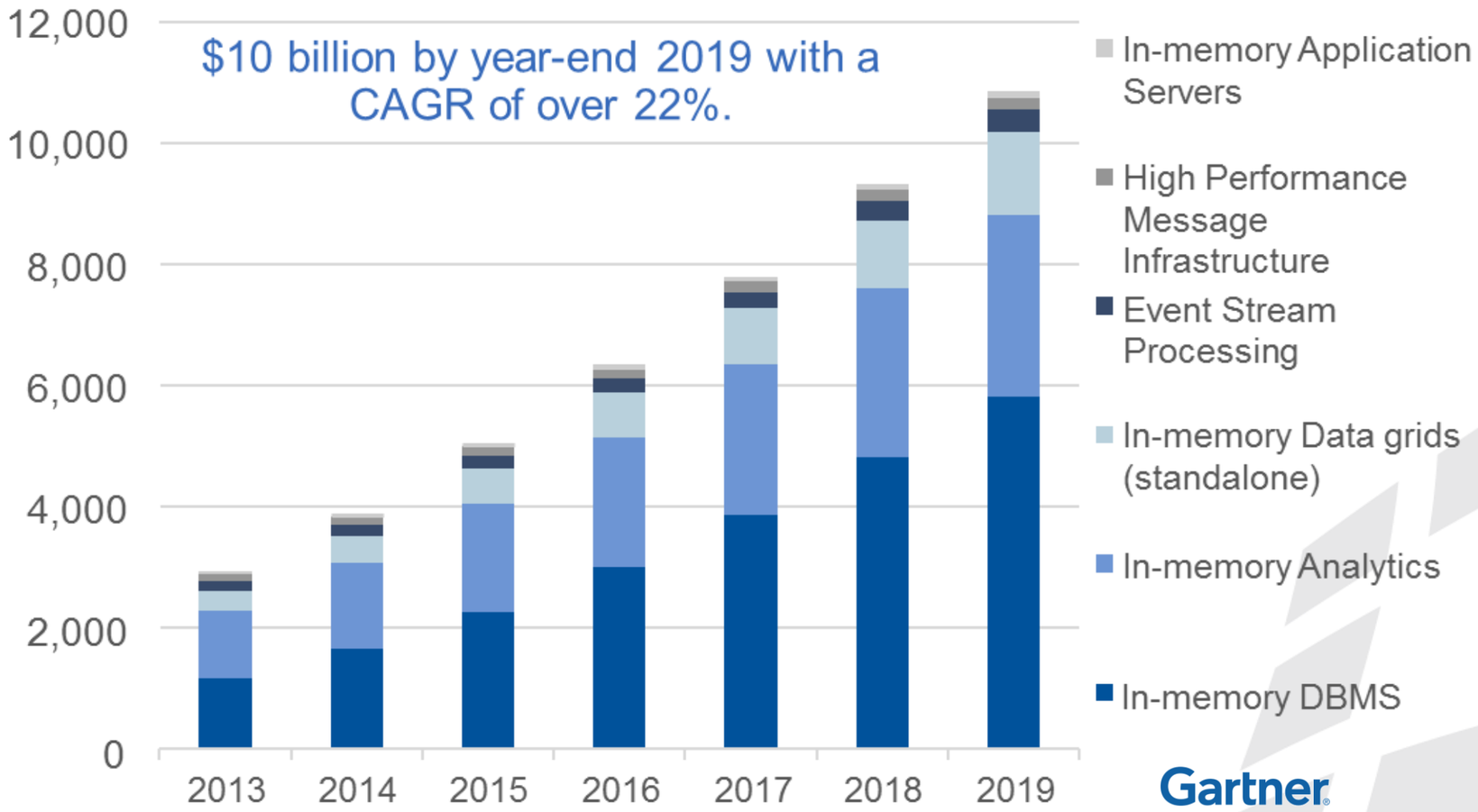


Cost drops 30% every 12 months

Apache Ignite GridGain

The In-Memory Computing Technology Market Is Big — And Growing Rapidly

IMC-Enabling Application Infrastructure ($M)



What is an In-Memory Computing Platform?

Multi-Featured Solution

- Supports data caching, massive parallel processing, in-memory SQL, streaming and much more

Does Not Replace Existing Databases

- Slides in between the existing application and data layers

Supports OLTP and OLAP Use Cases

- Offers ACID compliant transactions as well as analytics support

Multi-Platform Integration

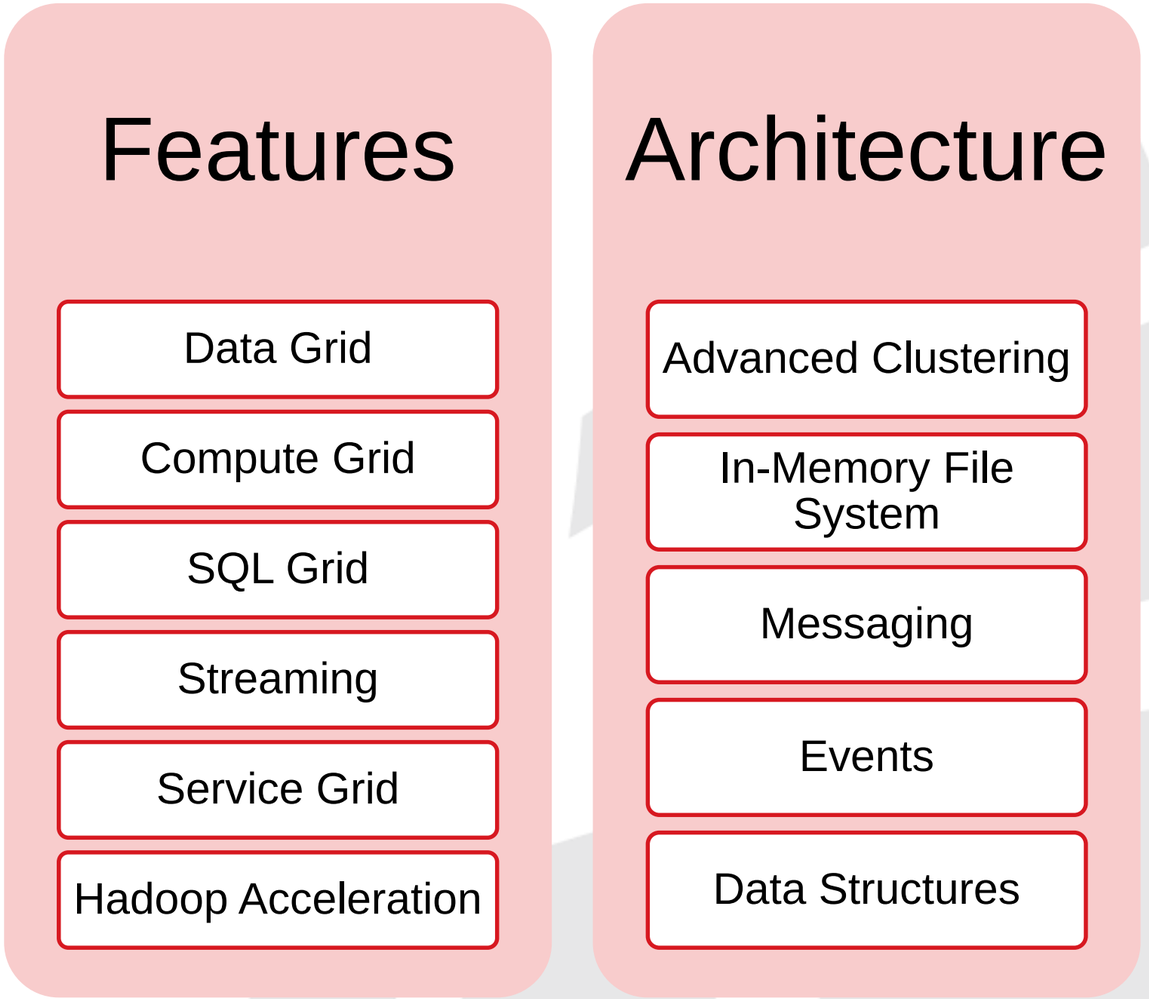
- Works with all popular RDBMS, NoSQL and Hadoop databases and offers a Unified API with support for a wide range of languages

Deployable Anywhere

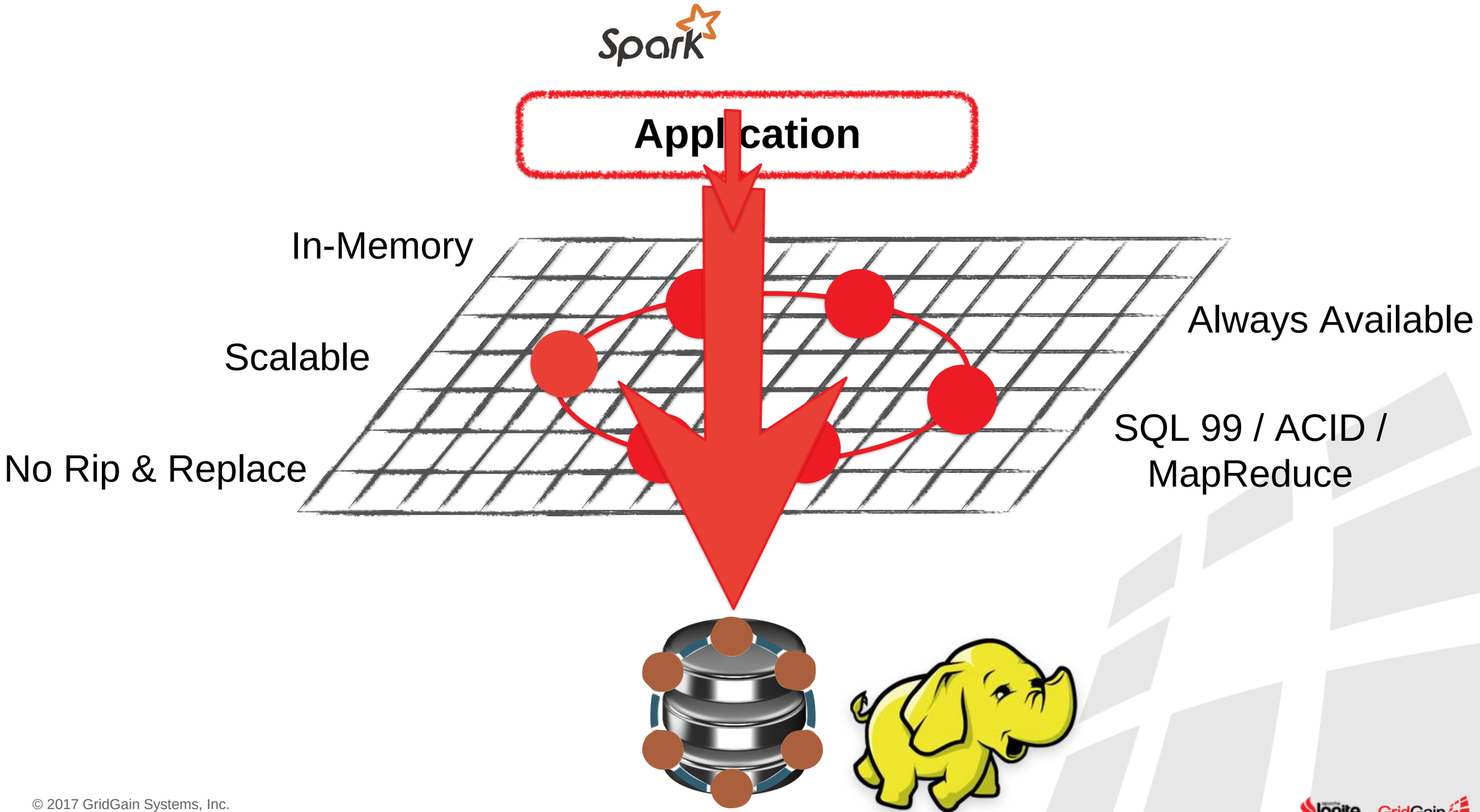
- Can be deployed on premise, in the cloud, or in hybrid environments

The GridGain In-Memory Computing Platform

- *A high-performance, distributed, in-memory platform for computing and transacting on large-scale data sets in real-time*
- *Built on Apache® Ignite™*

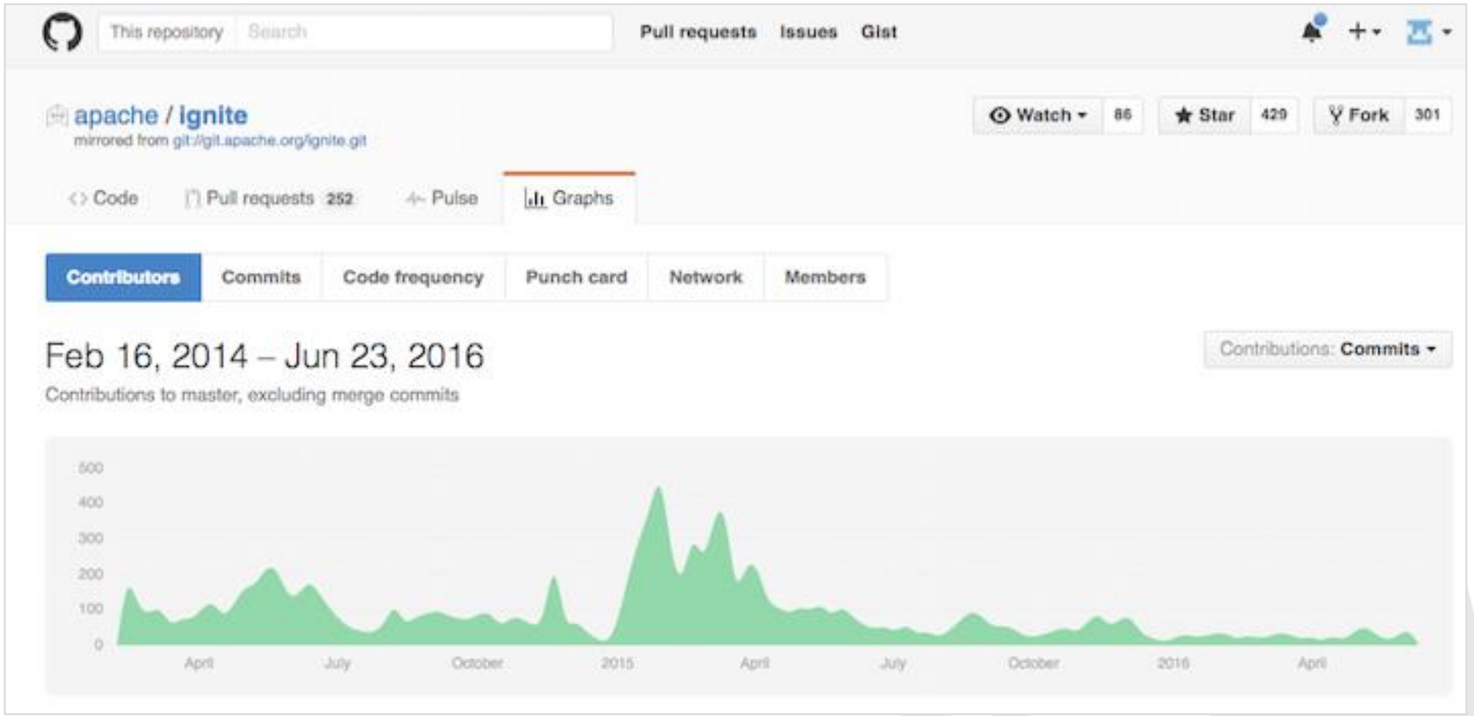


GridGain



Apache Ignite Project

- 2007: First version of GridGain
- Oct. 2014: GridGain contributes Ignite to ASF
- Aug. 2015: Ignite is the second fastest project to | graduate after Spark
- Today:
  - 60+ contributors and rapidly growing
  - Huge development momentum - Estimated 192 years of effort since the first commit in February, 2014 [Openhub]
  - Mature codebase: 1M+ lines of code



GridGain's Open Core Business Model
Apache Ignite vs. GridGain Enterprise

GridGain Enterprise Subscriptions include:

- Right to use GridGain Enterprise Edition
- Bug fixes, patches, updates and upgrades
- 9x5 or 24x7 Support
- Ability to procure Training and Consulting Services from GridGain
- Confidence and protection, not provided under Open Source licensing, that only a commercial vendor can provide, such as indemnification

Features	Apache Ignite	GridGain Enterprise
In-Memory Data Grid	√	√
In-Memory Compute Grid	√	√
In-Memory Service Grid	√	√
In-Memory Streaming	√	√
In-Memory Hadoop Acceleration	√	√
Distributed In-Memory File System	√	√
Advanced Clustering	√	√
Distributed Messaging	√	√
Distributed Events	√	√
Distributed Data Structures	√	√
Portable Binary Objects	√	√
Management & Monitoring GUI		√
Enterprise-Grade Security		√
Network Segmentation Protection		√
Recoverable Local Store		√
Rolling Production Updates		√
Data Center Replication		√
Integration with Oracle GoldenGate		√
Basic Support (9×5)	√	√
Enterprise Support (9×5 and 24×7)		√
Security Updates		√
Maintenance Releases & Patches	**Free w/ optional Paid Support**	**Annual License Subscription** √

GridGain In-Memory Computing Use Cases

Data Grid

- Web session clustering
- Distributed caching
- Scalable SaaS

Compute Grid

- High performance computing
- Machine learning
- Risk analysis
- Grid computing

SQL Grid

- In-memory SQL
- Distributed SQL processing
- Real-time analytics

Streaming

- Real-time analytics
- Streaming Big Data analysis
- Monitoring tools

Hadoop Acceleration

- Faster Big Data insights
- Real-time analytics
- Batch processing

Events

- Complex event processing (CEP)
- Event driven design

1000's of Deployments

Automated Trading Systems

- *Real time analysis of trading positions*
- *Real time market risk assessment*
- *High volume transactions*
- *Ultra low latencies trading*

Financial Services

- *Fraud Detection*
- *Risk Analysis*
- *Insurance rating and modeling*

Big Data Analytics

- *Real time analysis of inventory*
- *Operational up-to-the-second BI*

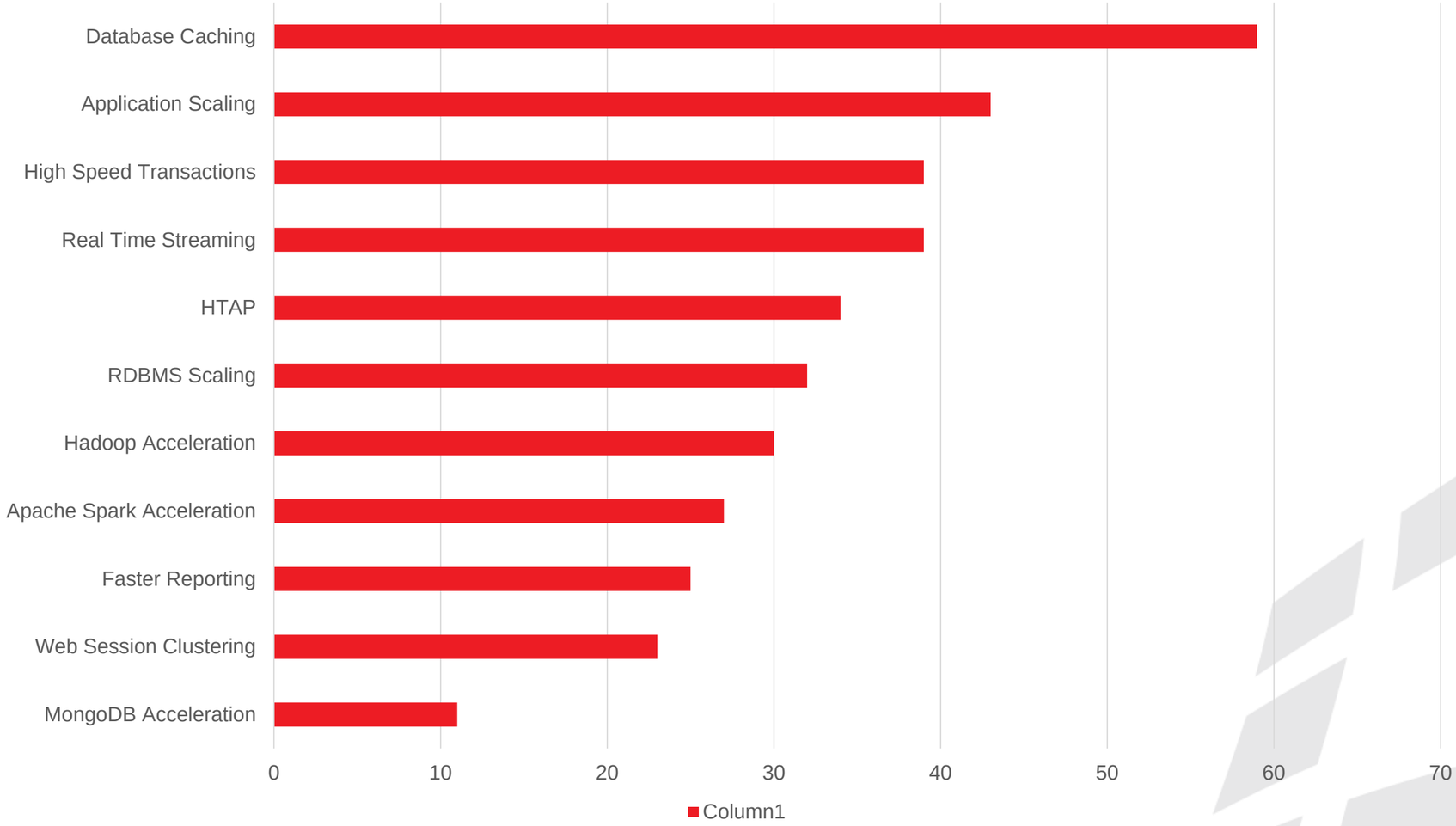
Mobile & IoT

- *Real-time streaming processing*
- *Complex event processing*

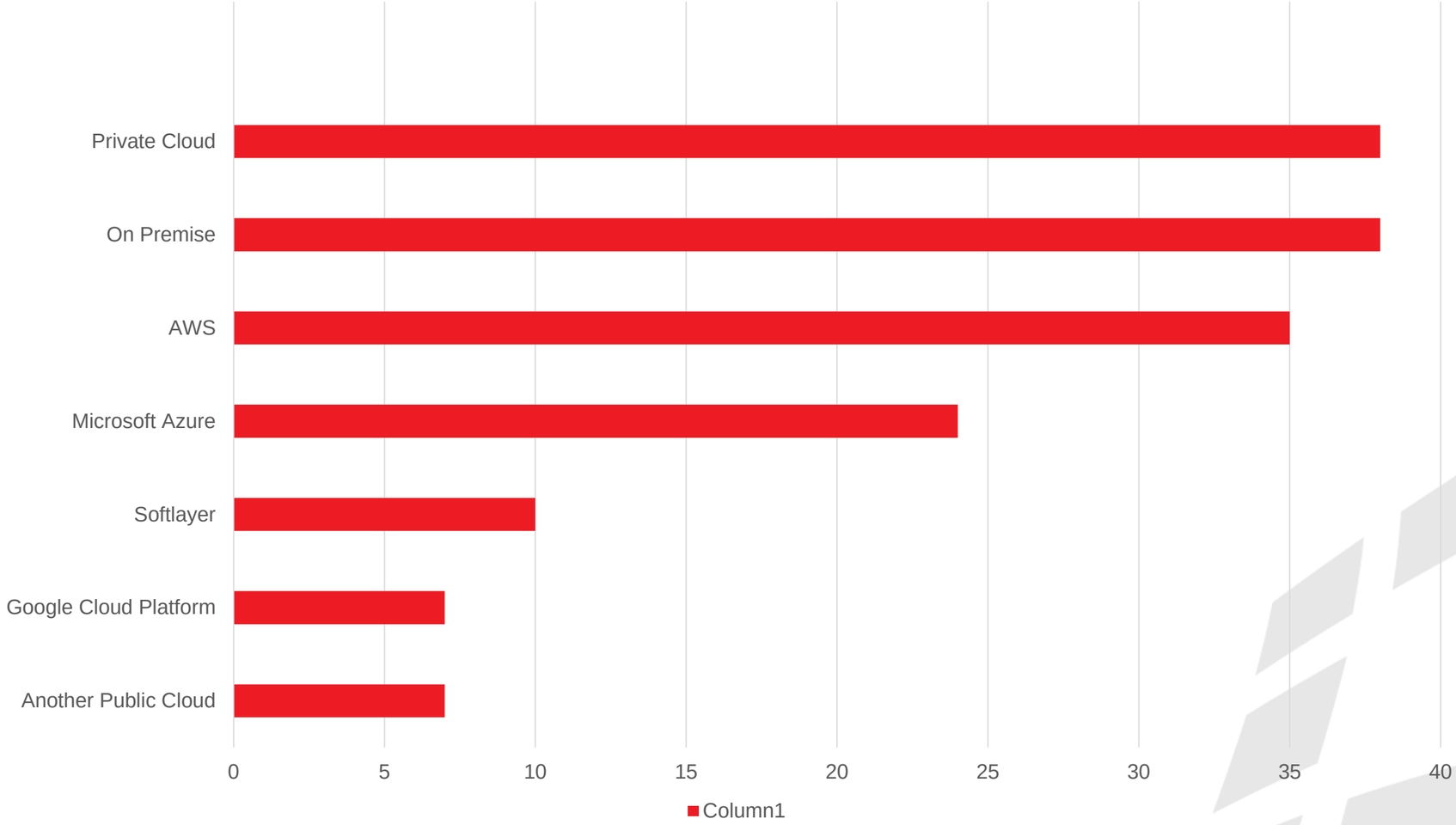
Biotech

- *High performance genome data matching*
- *Drug discovery*

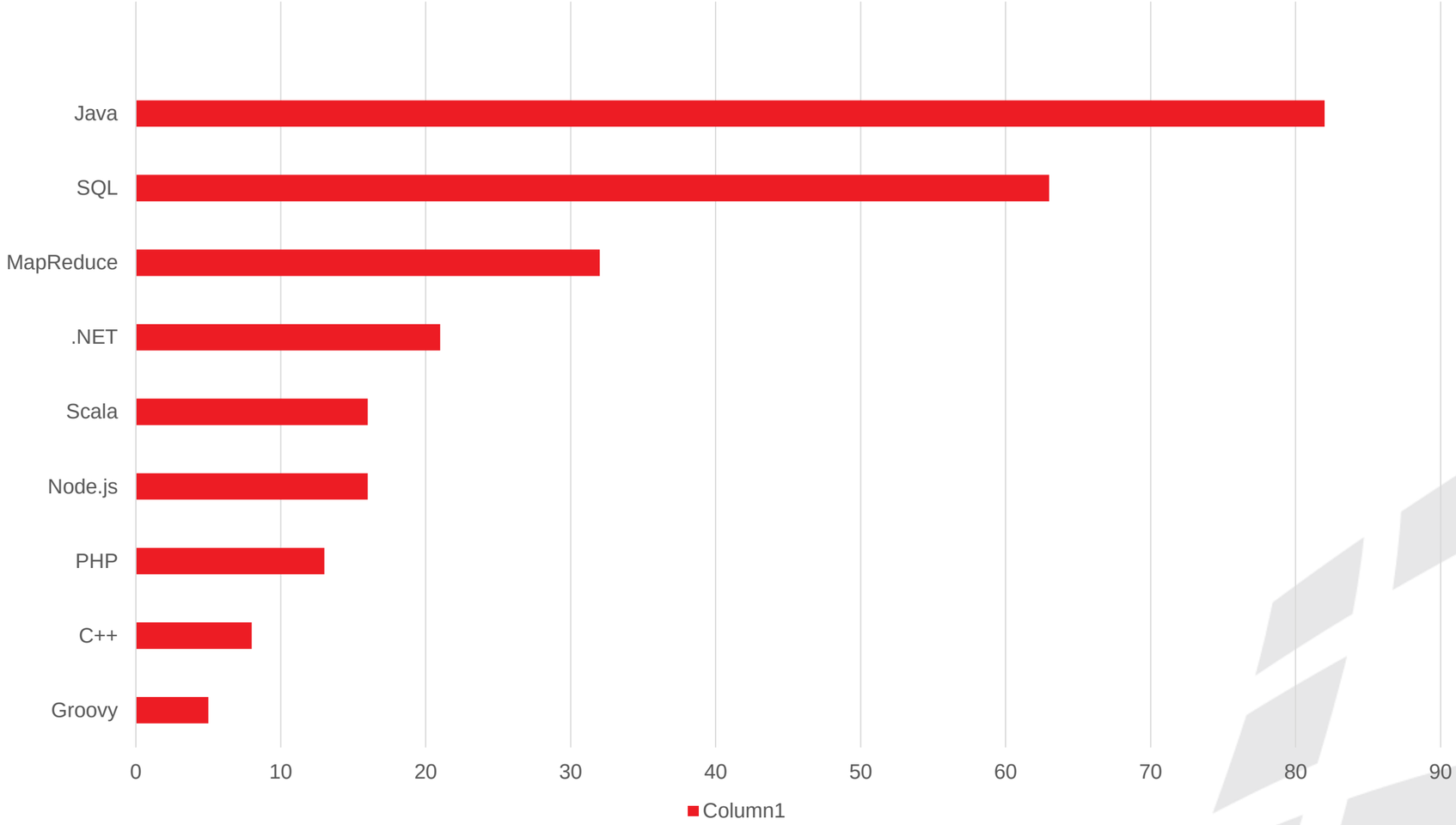
Survey Results: What uses were you considering for in-memory computing



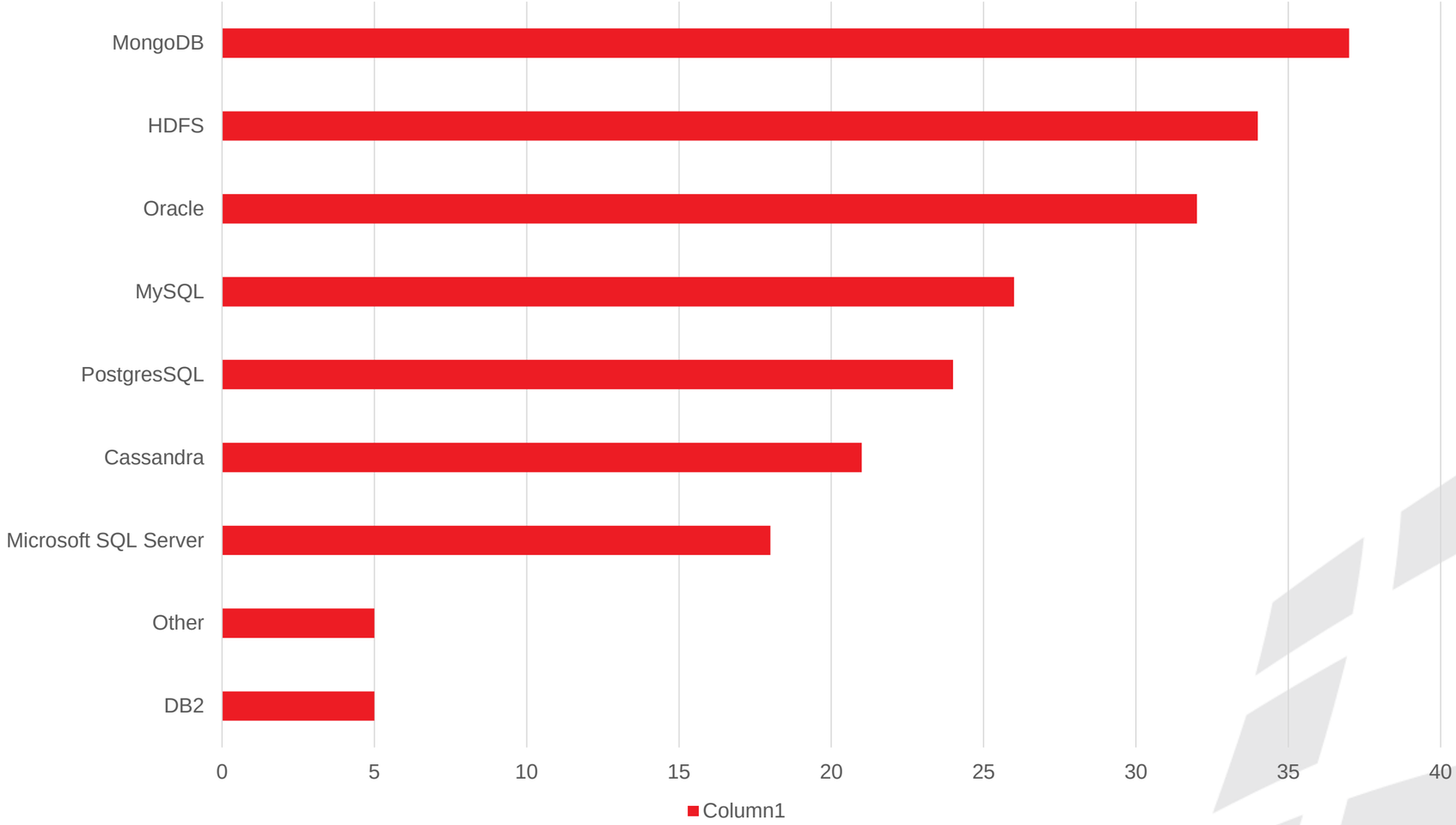
Survey Results: Where do you run GridGain and/or Apache Ignite?



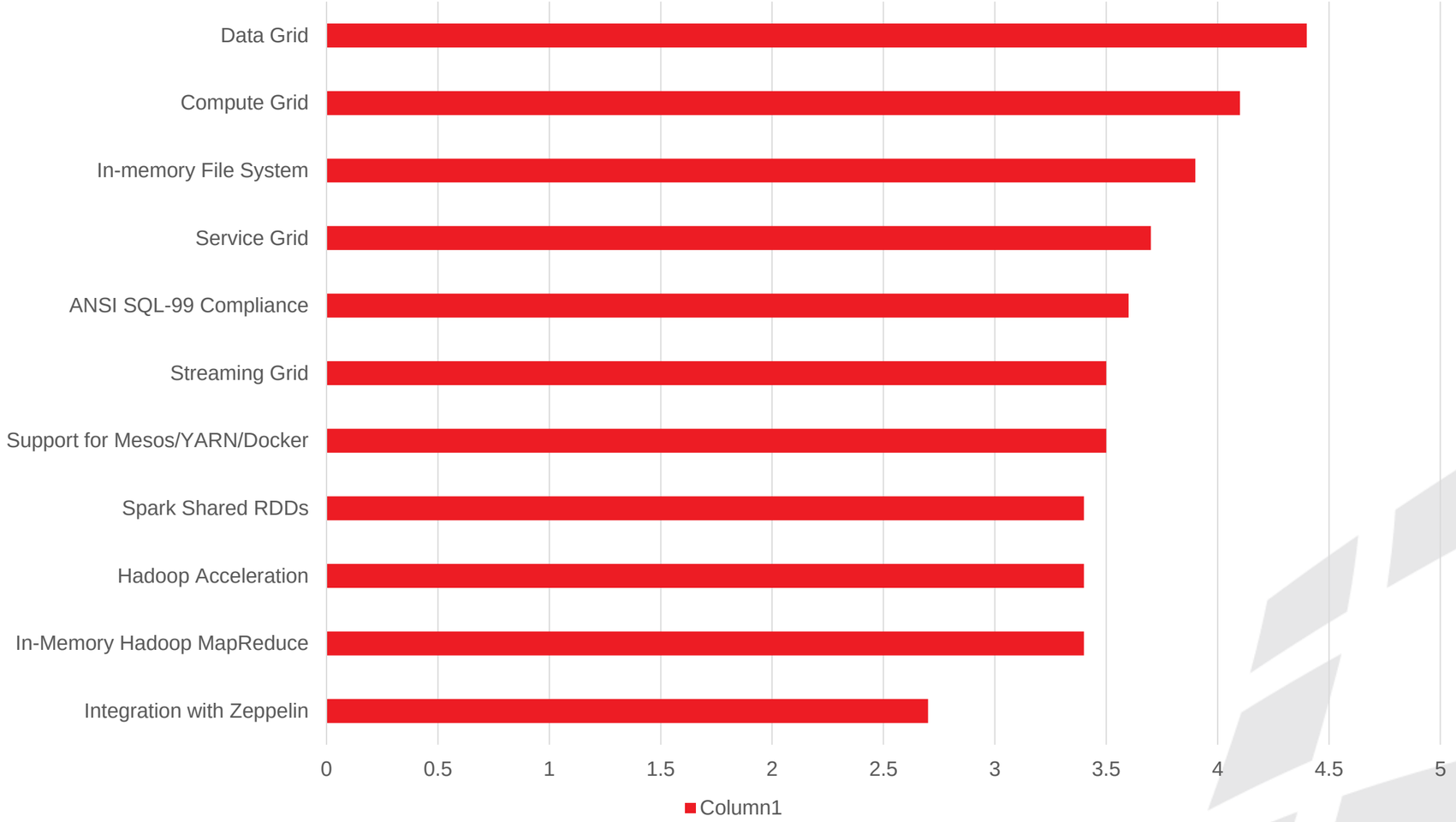
Survey Results: Which of the following protocols do you use to access your data?



Survey Results: Which data stores are you/would you likely use with GridGain/Apache Ignite?



Survey Results: How important are each of the following product features to your organization?



GridGain Software and SaaS Customers

- Workday
- Advent Software
- Thingworx
- Microsoft
- Siemens
- IBM
- Misys
- Optym
- Silver Spring Networks
- The Glue

Case Study:

- Financial services software
  - Retail and corporate banking
  - Lending
  - Treasure
  - Capital markets
  - Investment management
  - Enterprise risk
- More than 2,000 customers in 130 countries
- Used by 48 of the world's 50 largest banks

- The Challenge: Eliminating Data Processing Bottlenecks
  - Huge amounts of trade and accounting data
  - Customers need
    - High-speed transactions
    - Real-time reporting
  - New Java-based IT stack with data lake support
  - Global regulatory compliance

Case Study:

- Commodity servers (256GB RAM)
- Data stored in memory
  - Transactions
  - Market data
- Parallel processing across cluster
  - Calculation heavy reporting for regulatory compliance

Case Study:

FusionFabric.cloud

- Integrates trading systems with cloud-based components
  - OTC derivatives
  - Exchange traded derivatives
  - Inflation
  - Fixed income
  - FX/MM
  - Hybrids
  - Developing additional modules

"With GridGain, we have achieved real-time processing of massive amounts of trade and transaction data, eliminating bottlenecks and enabling us to offer next-generation financial services to our customers."
-Felix Grevy, Director of Product Management for FusionFabric.cloud at Misys

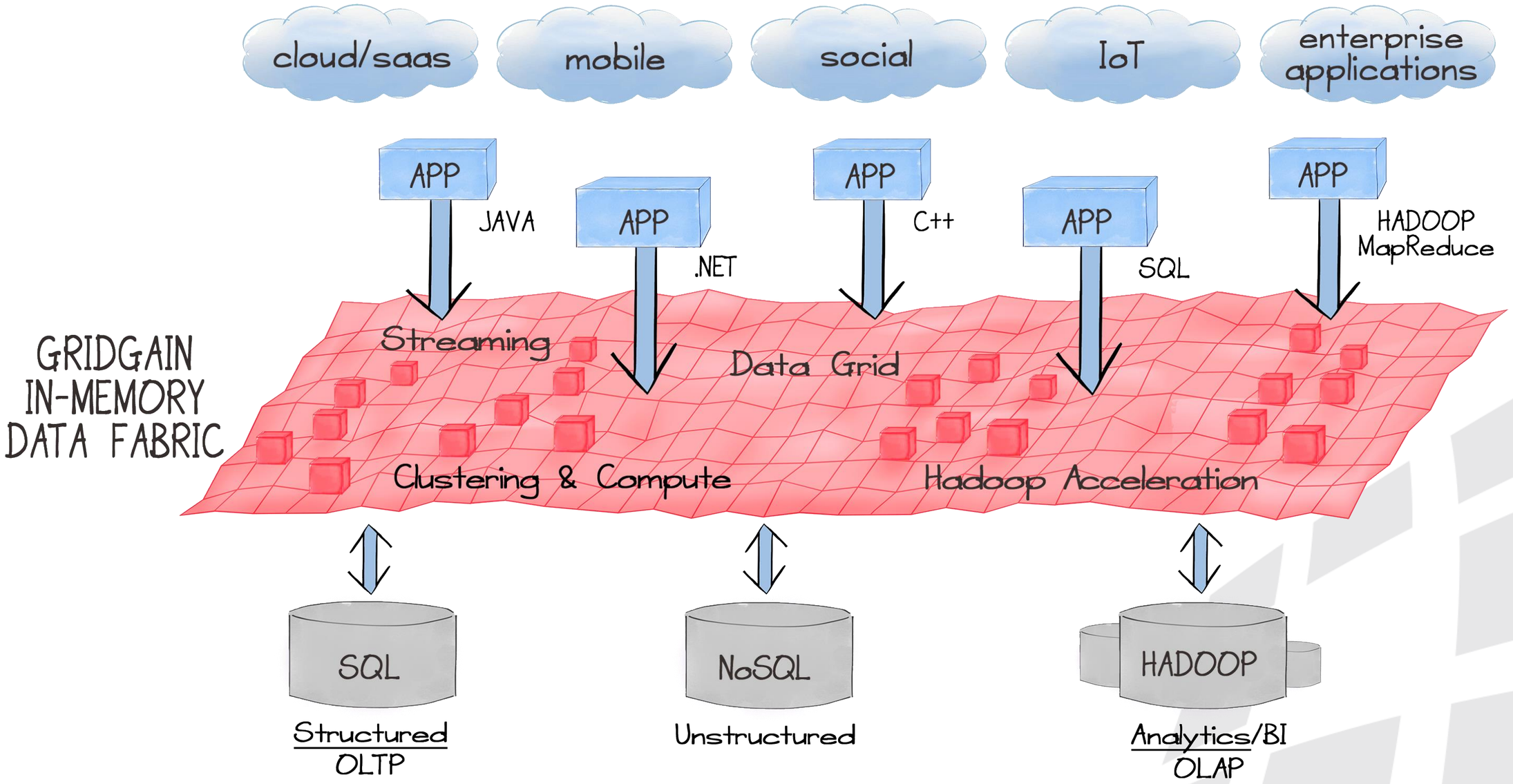
Software and SaaS Use Case: Search

- Customer sells over 20,000 information products worldwide
- Needed a faster search across products
- Search across over 1 billion properties
- Currently 5M+ information assets
- Anticipate growing to 50M+ information assets
- Search platform forecasted to generate over $275 million in revenue
- “GridGain is the only platform capable of providing the performance and scale we need.”

Software/SaaS Use Case: IoT Management and Analytics

- Requirements
  - Performance
  - Scalability
  - High Availability
  - Streaming
  - APIs for analytics

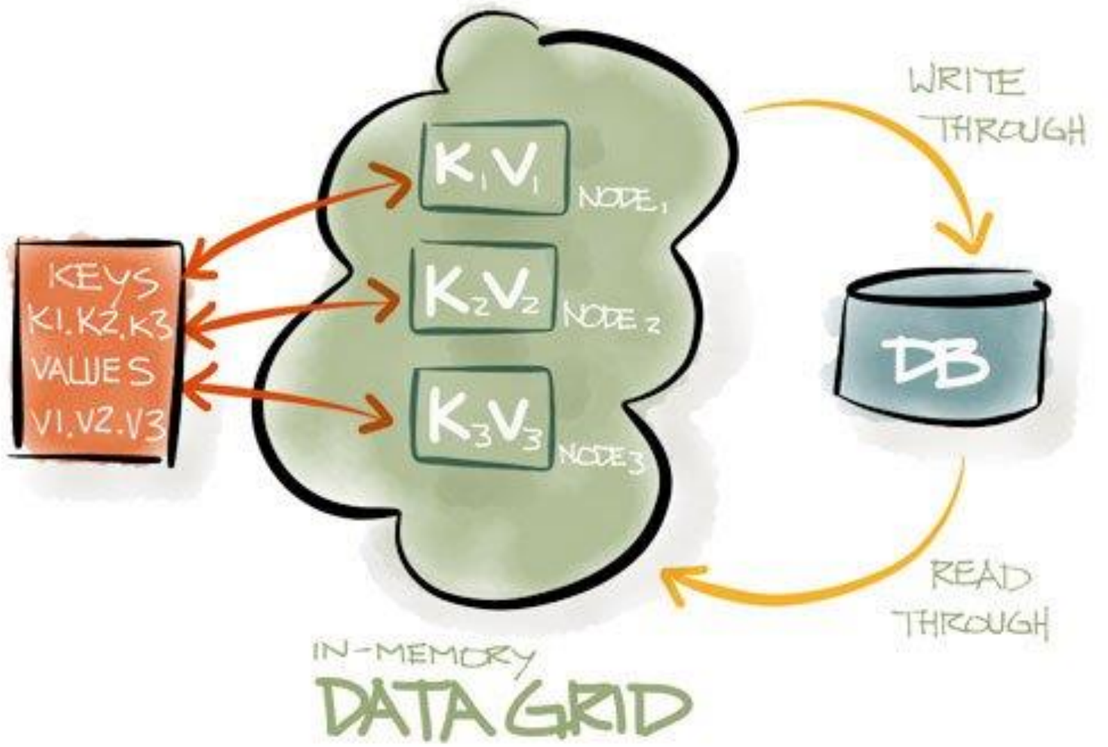
- Silver Spring Networks
  - Smart meter solution
  - Performance
  - Scalability
  - High Availability
- ThingWorx
  - Enterprise IoT platform
  - Performance
  - Scalability
  - Integrations
  - Analytics



In-Memory Data Grid

- *Inserted between the application and data layers. Moves disk-based data from RDBMS, NoSQL or Hadoop databases into RAM*

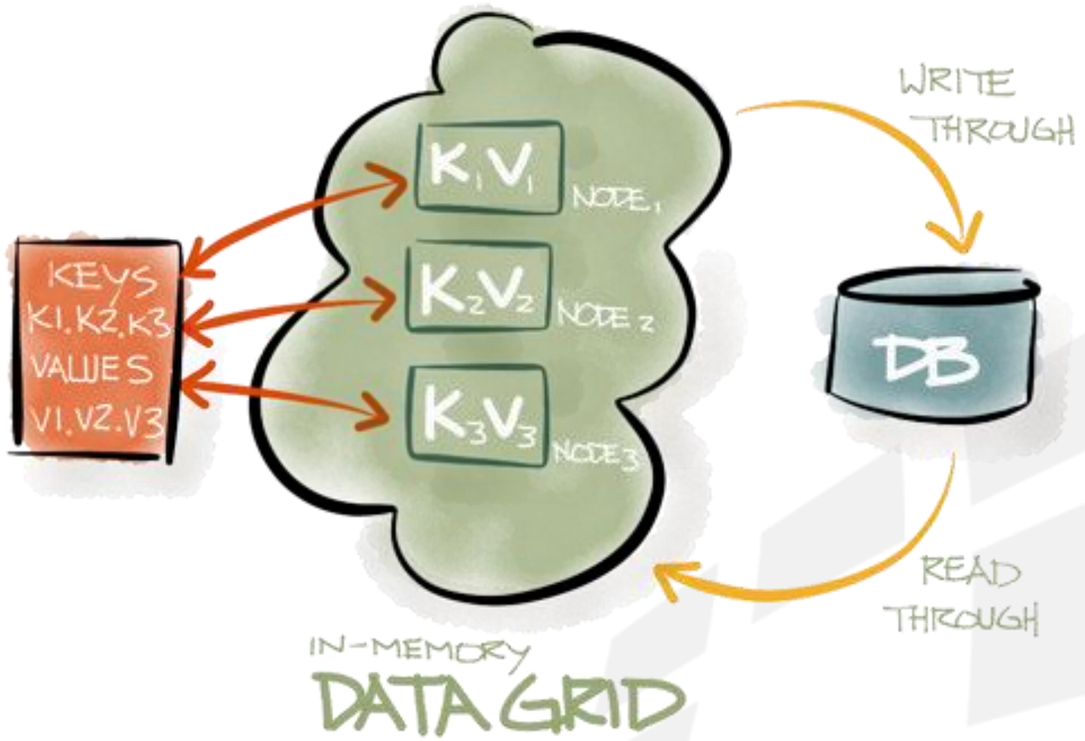
- ***Features:***
  - *Distributed In-Memory Key-Value Store*
  - *Replicated and Partitioned Data Caches*
  - *Lightning Fast Performance*
  - *Elastic Scalability*
  - *Distributed In-Memory Transactions (ACID)*
  - *Distributed In-Memory Queue and Other Data Structures*
  - *Web Session Clustering*
  - *Hibernate L2 Cache Integration*
  - *On-Heap and Off-Heap Storage*
  - *Distributed SQL Queries with Distributed Joins*



GridGain

Data Grid: RDBMS Integration

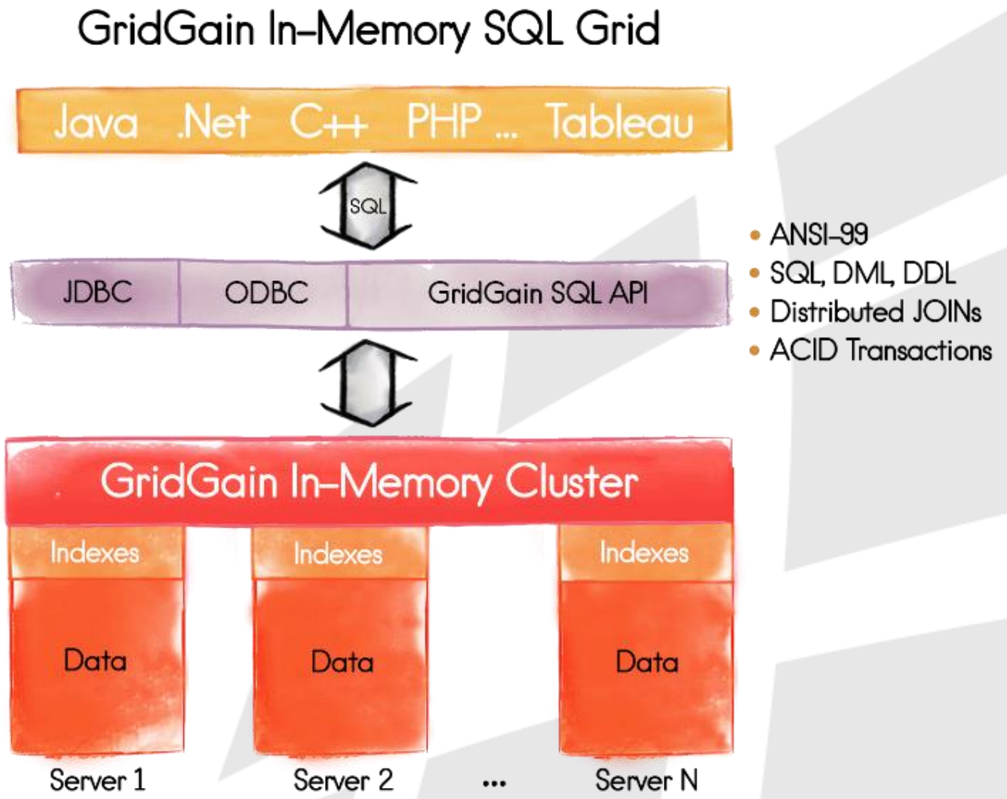
- Read-through & Write-through
- Support for Write-behind
- Configurable eviction policies
- DB schema mapping wizard:
  - Generates all the XML configuration and Java POJOs



In-Memory SQL Grid

- *Horizontally scalable, fault tolerant, ANSI SQL-99 compliant, and fully supports all SQL and DML commands*

- ***Features:***
  - *Supports SQL and DML commands including SELECT, UPDATE, INSERT, MERGE and DELETE Queries*
  - *Distributed SQL*
  - *Geospatial Support*
  - *SQL Communications Through the GridGain ODBC or JDBC APIs Without Custom Coding*
  - *ANSI SQL-99 Compliance*

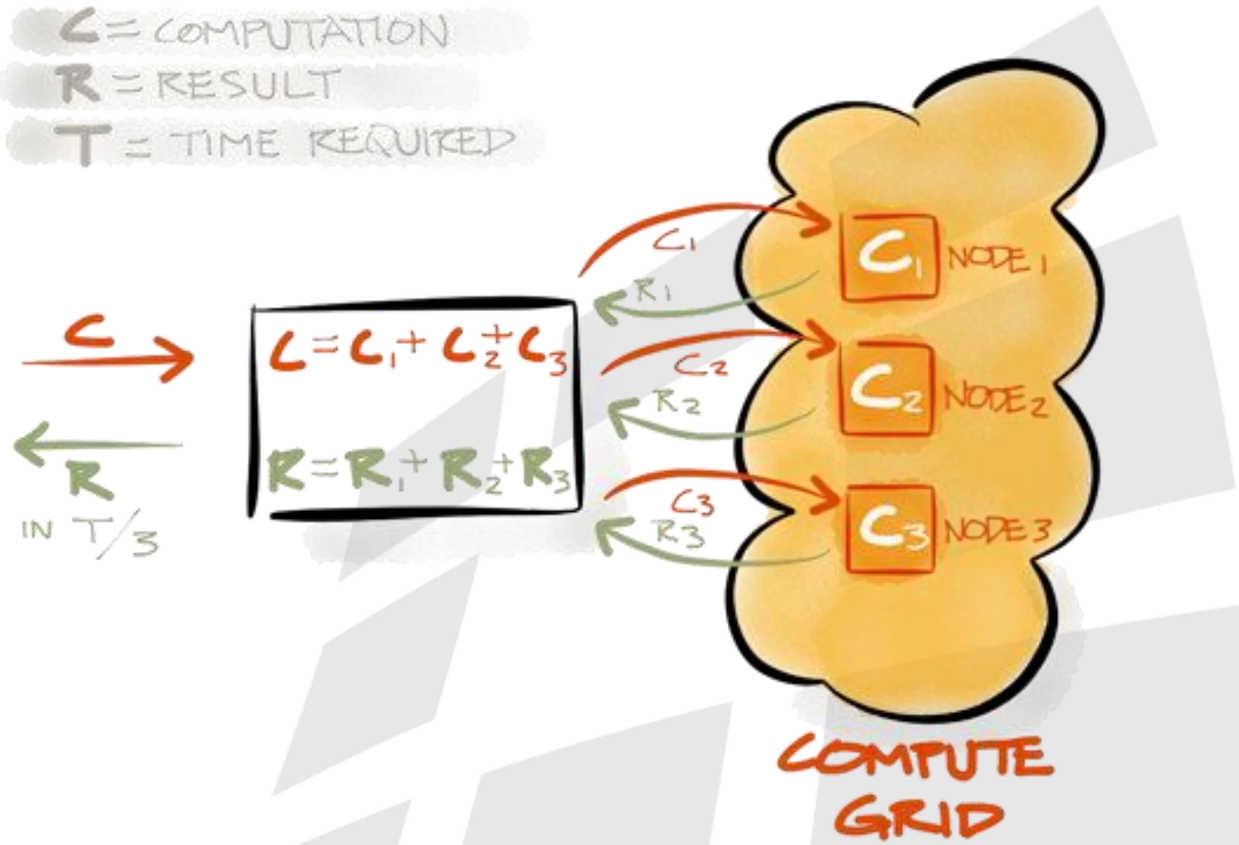


GridGain

In-Memory Compute Grid

- *Enables parallel processing of CPU or otherwise resource intensive tasks*

- ***Features:***
  - *Dynamic Clustering*
  - *Direct API for Fork-Join & MapReduce Processing*
  - *Distributed Closure Execution*
  - *Adaptive Load Balancing*
  - *Automatic Fault Tolerance*
  - *Linear Scalability*
  - *Custom Scheduling*
  - *State Checkpoints for Long Running Jobs*
  - *Pluggable SPI Design*

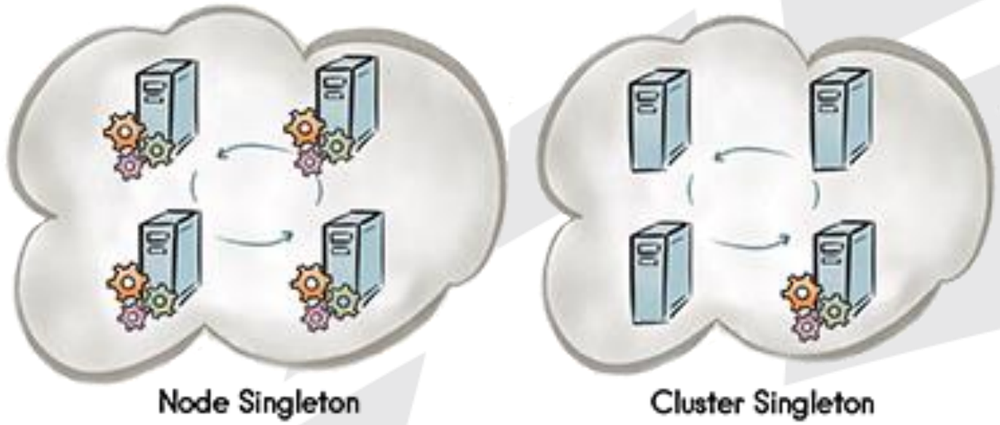


GridGain

In-Memory Service Grid

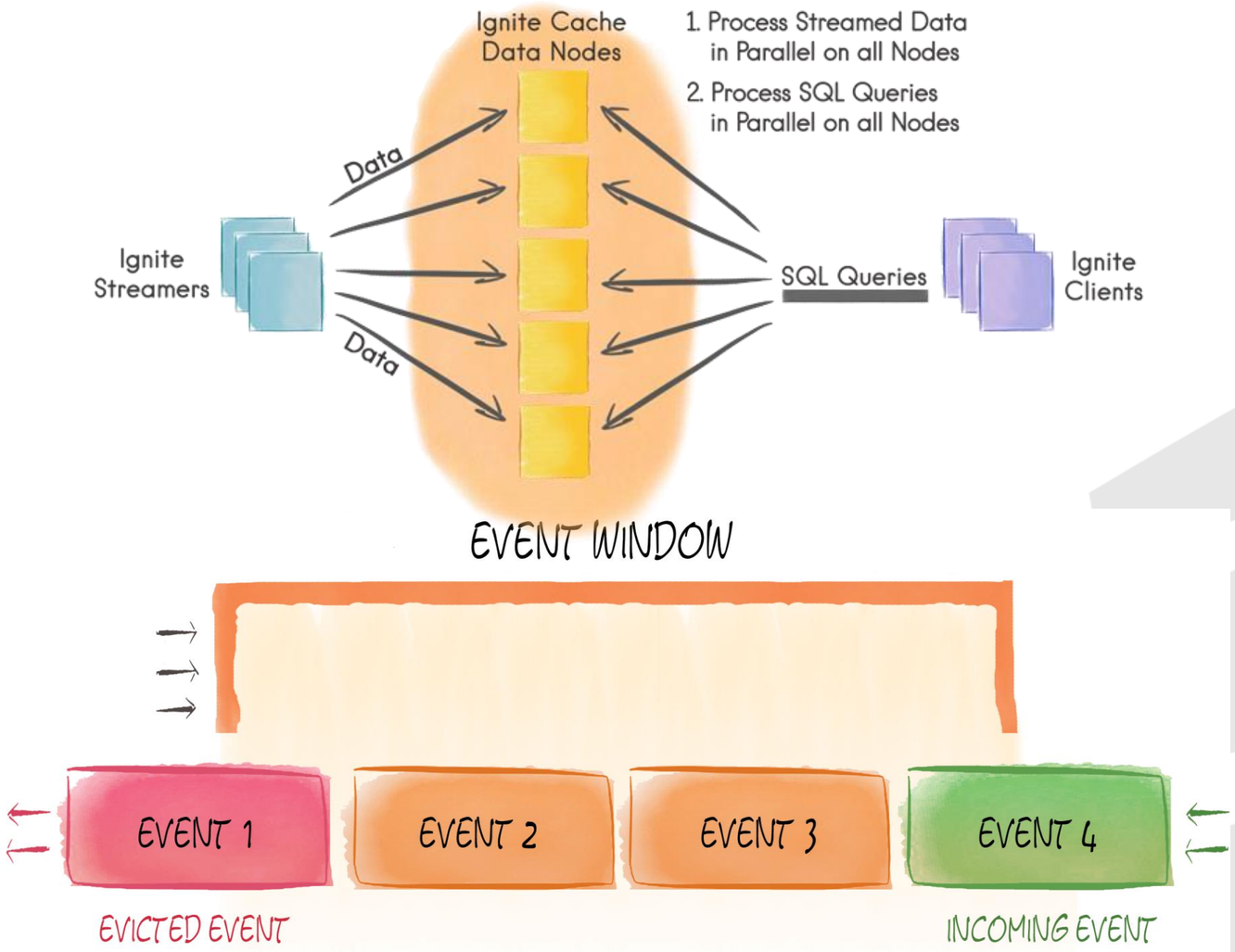
- *Provides control over how many instances of your service should be deployed on each cluster node and guarantees continuous availability of all deployed services in case of node failures*

- ***Features:***
  - *Automatically Deploy Multiple Instances of a Service*
  - *Automatically Deploy a Service as Singleton*
  - *Automatically Deploy Services on Node Start-Up*
  - *Load Balanced and Fault Tolerant Deployment*
  - *Un-Deploy Any of the Deployed Services*
  - *Get Service Deployment Topology Information*
  - *Access Remotely Deployed Service via Service Proxy*



In-Memory Streaming and CEP

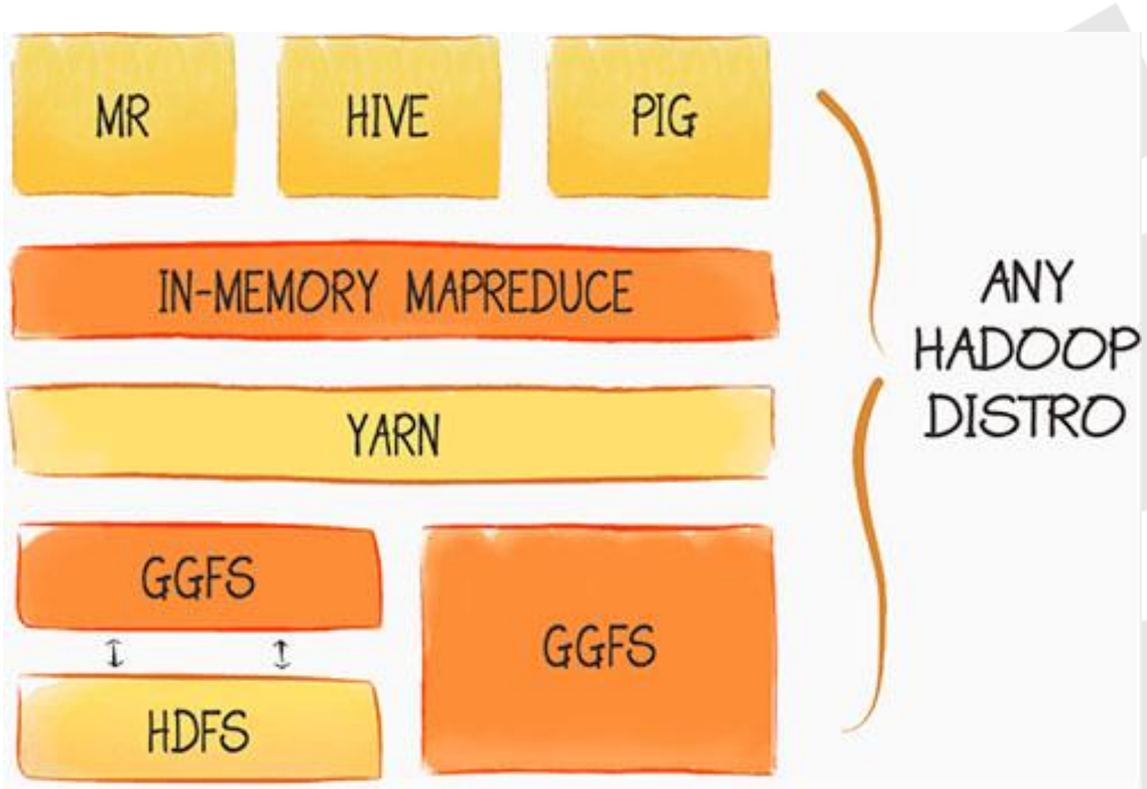
- Streaming Data Never Ends
- Sliding Windows for CEP/Continuous Query
- Customizable Event Workflow
- Branching Pipelines
- Pluggable Routing
- Real Time Analysis
- Data Indexing
- Distributed Streamer Queries



In-Memory Hadoop Acceleration

- *Provides easy to use extensions to disk-based HDFS and traditional MapReduce, delivering up to 10x faster performance*

- ***Features:***
  - *Use existing MapReduce / Pig / Hive*
  - *10x Faster Performance*
  - *In-Memory MapReduce*
  - *Highly Optimized In-Memory Processing*
  - *Standalone File System*
  - *Optional Caching Layer for HDFS*
  - *Read-Through and Write-Through with HDFS*



GridGain

ANY QUESTIONS?

Thank you for joining us. Follow the conversation.

www.gridgain.com
www.gridgain.com/resources/blog

@gridgain
#gridgain #inmemorycomputing
@msarrel
matt.sarrel@gridgain.com