In-Memory Computing Platform:

Supercharge Ecommerce with In-Memory Computing

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

© 2016 GridGain Systems, Inc.

Agenda

- Introduction
- Survey Results
- In-Memory Computing
- GridGain / Apache Ignite Overview
- Use cases
- Ecommerce customers
- GridGain / Apache Ignite In-depth

© 2016 GridGain Systems, Inc.

Your Presenter

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

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

© 2016 GridGain Systems, Inc.

Look to Amazon to Understand Ecommerce Trends

- Amazon owns roughly 46% of all ecommerce
  - Best Buy 4%
  - Walmart 2.7%
- 1 billion items sold 2016 holiday season (9x order growth YOY)
- Processed > 40 million orders on Black Friday
  - 500 orders/second
  - $2.27 million revenue
  - Physically shipped 16-18 million packages
- 200K seasonal workers hired
- Cyber Monday sold 46 electronic devices/second
- Amazon is now the largest clothing retailer in the world
- US gov't statistics indicate that Ecommerce is roughly 9% of all retail US sales.

© 2016 GridGain Systems, Inc.

User Experience Drives Ecommerce Technology Choices

- Ecommerce is highly competitive
- Customer acquisition and revenue growth depend on how fast users can find what they're looking for, configure it, and buy it
- Site speed at scale is crucial
- Personalization creates unique experiences
- High-performance across multiple engines built on varied technology stacks
- High availability is a critical factor in providing consistent performance at varied user loads

© 2016 GridGain Systems, Inc.

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

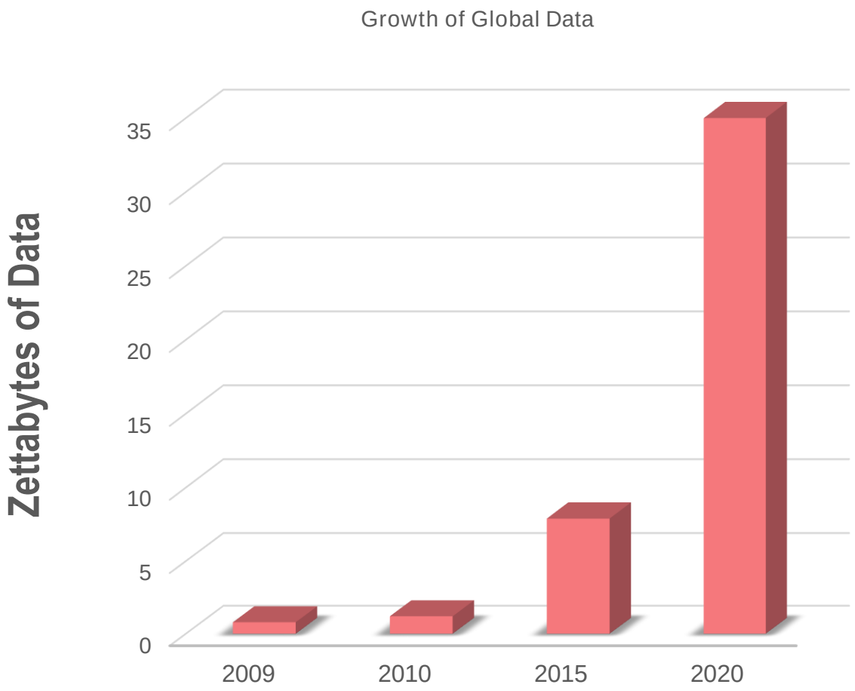
The Cost of RAM Continues to Fall

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

© 2015 GridGain Systems, Inc.

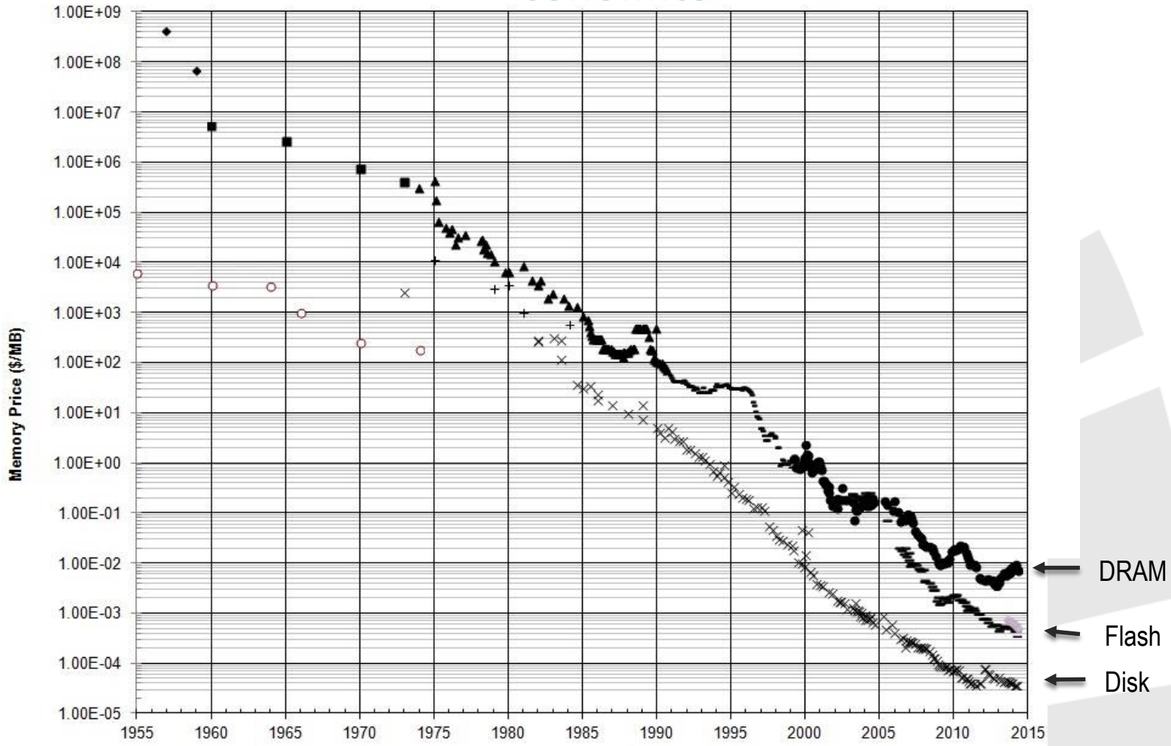
Why Now?

Data Growth and Internet Scale Driving Demand



8 zettabytes in 2015 growing to 35 in 2020

Declining DRAM Cost Driving Attractive Economics

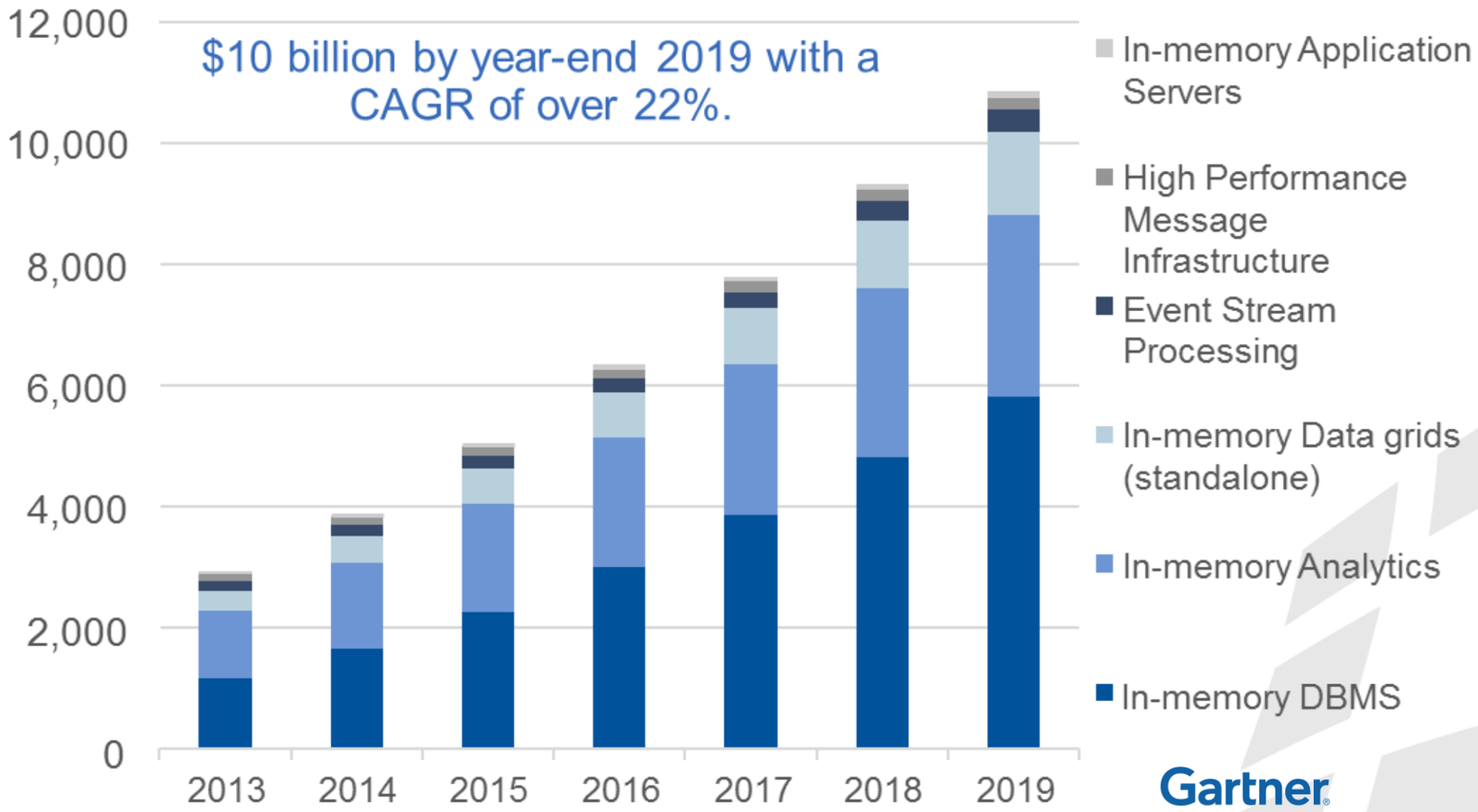


Cost drops 30% every 12 months

© 2016 GridGain Systems, Inc.

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

IMC-Enabling Application Infrastructure ($M)



© 2016 GridGain Systems, Inc.

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

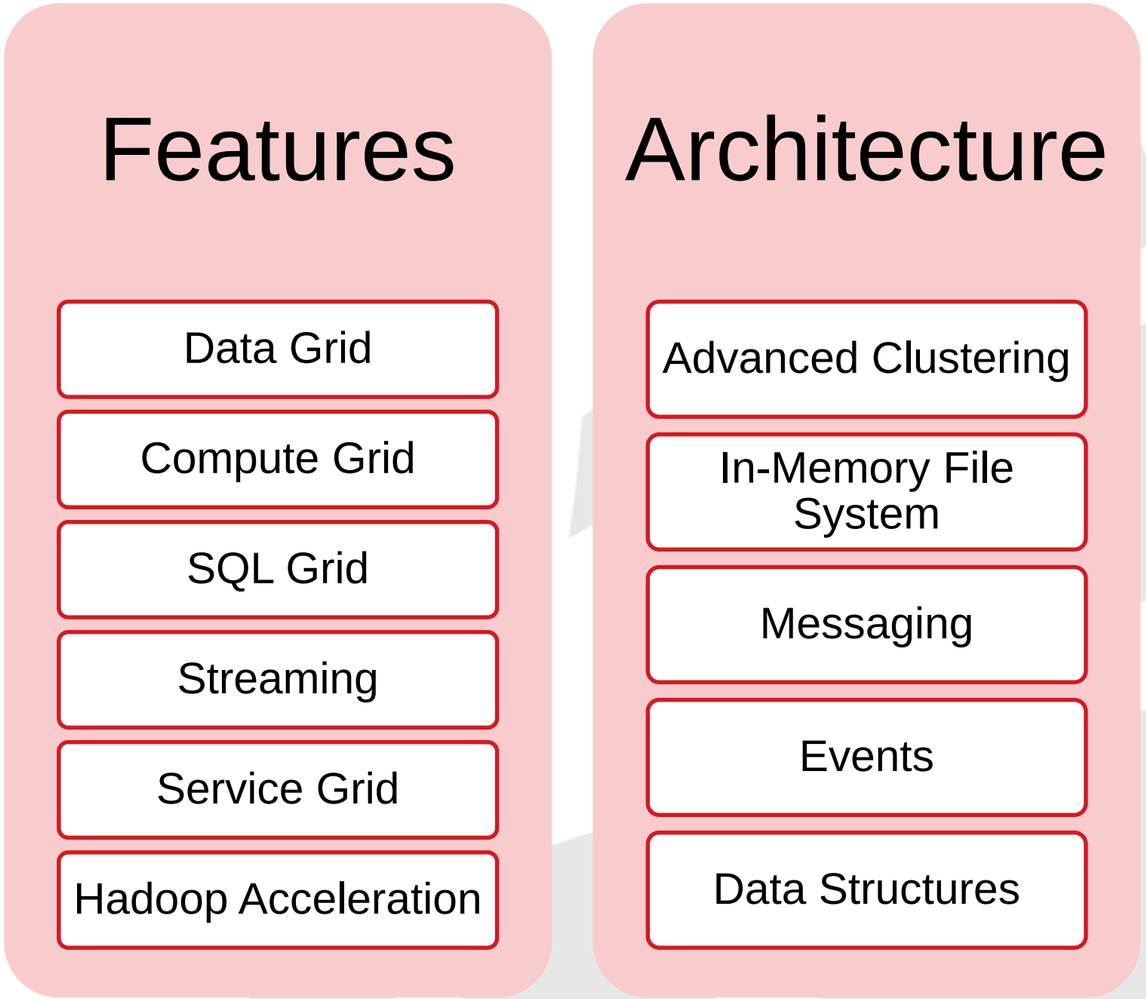
Deployable Anywhere

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

© 2015 GridGain Systems, Inc.

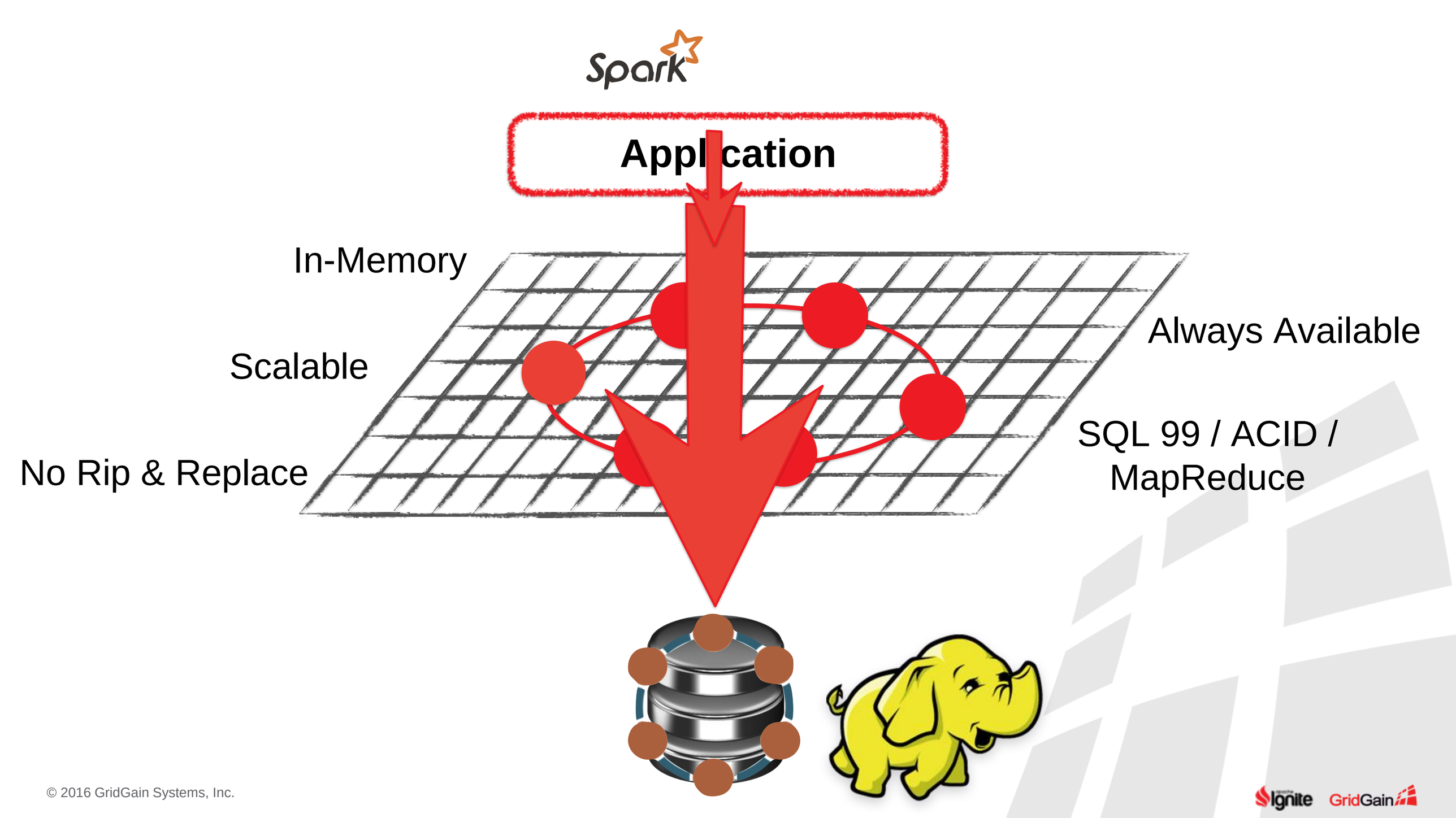
The GridGain In-Memory Computing Platform

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



© 2016 GridGain Systems, Inc.

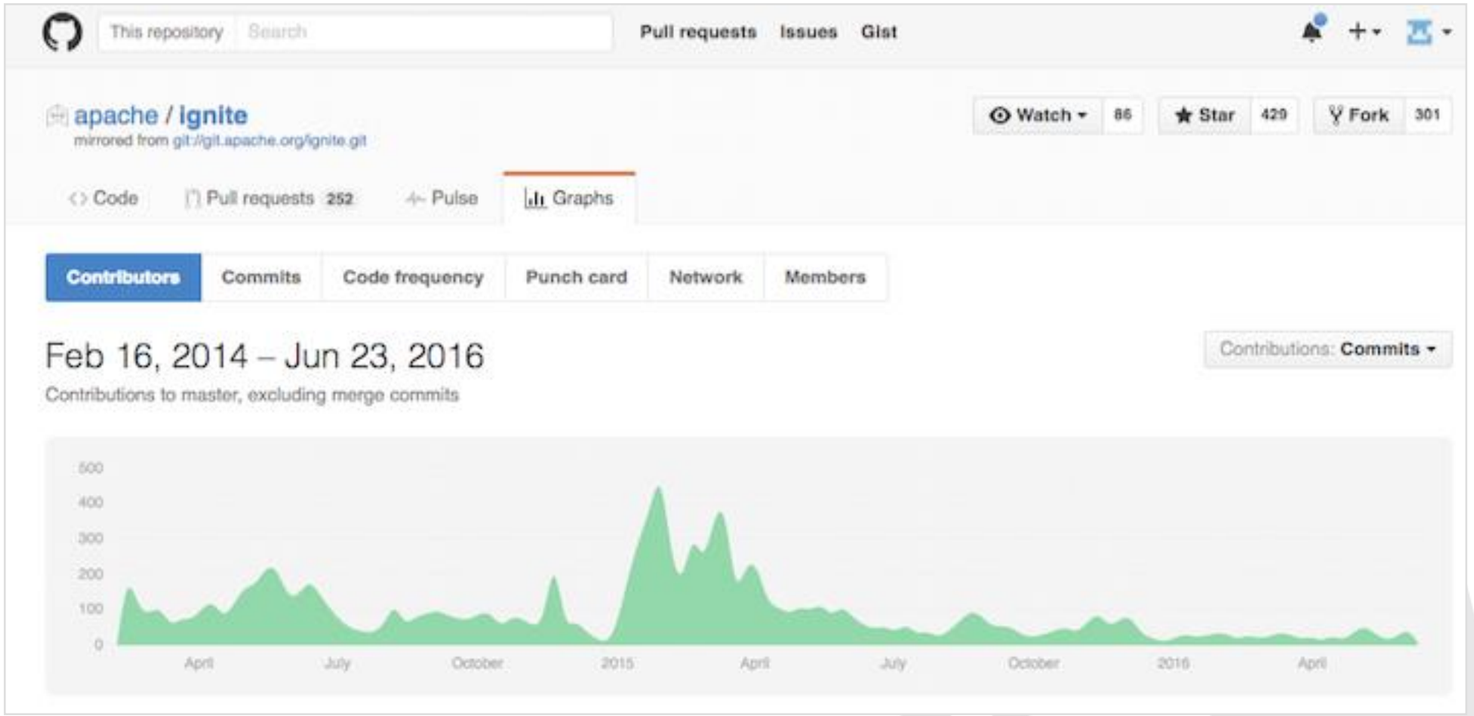
GridGain



© 2016 GridGain Systems, Inc.

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



© 2016 GridGain Systems, Inc.

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** √

© 2016 GridGain Systems, Inc.

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

© 2015 GridGain Systems, Inc.

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

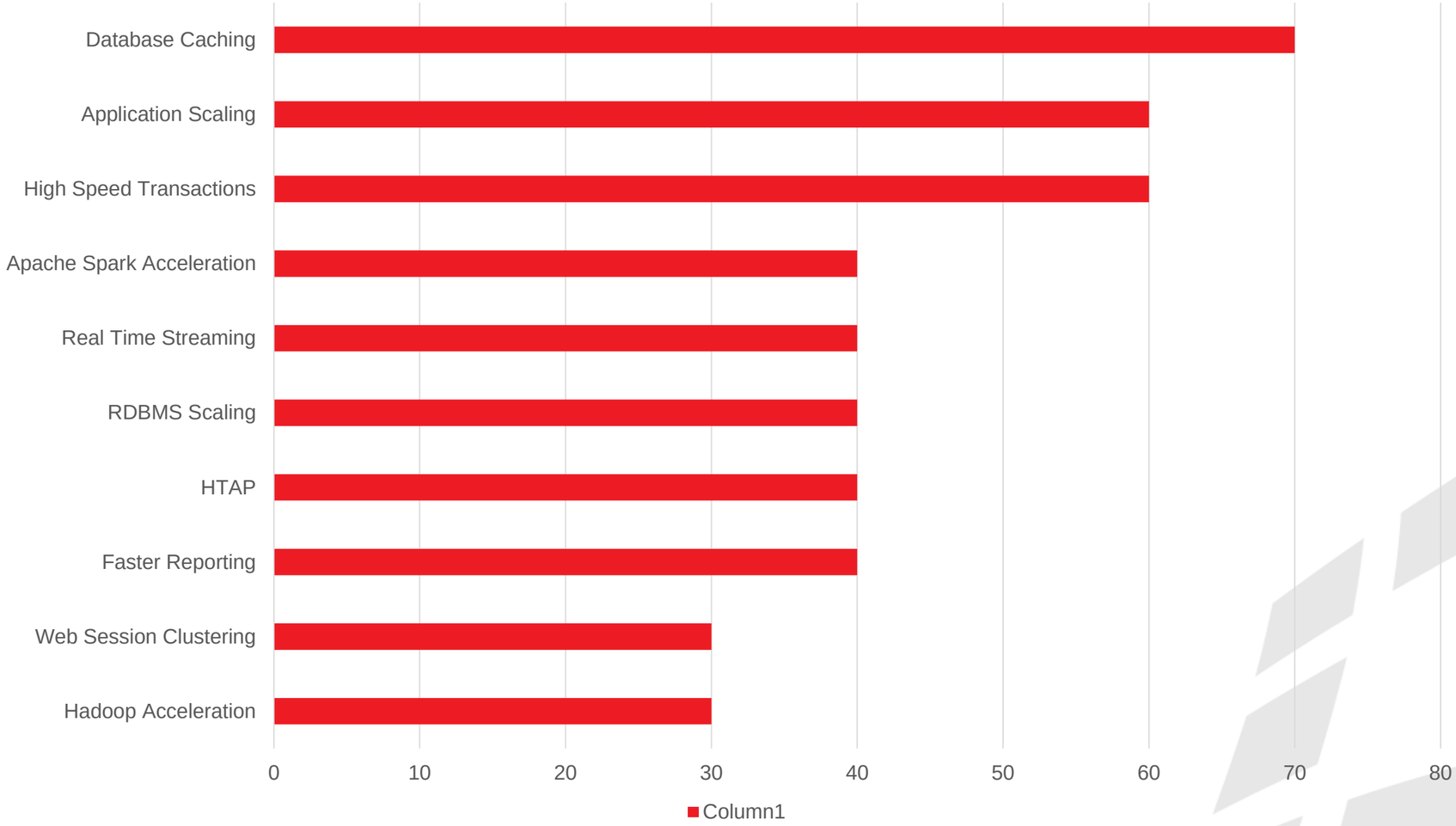
- *Real-time streaming processing*
- *Complex event processing*

Biotech

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

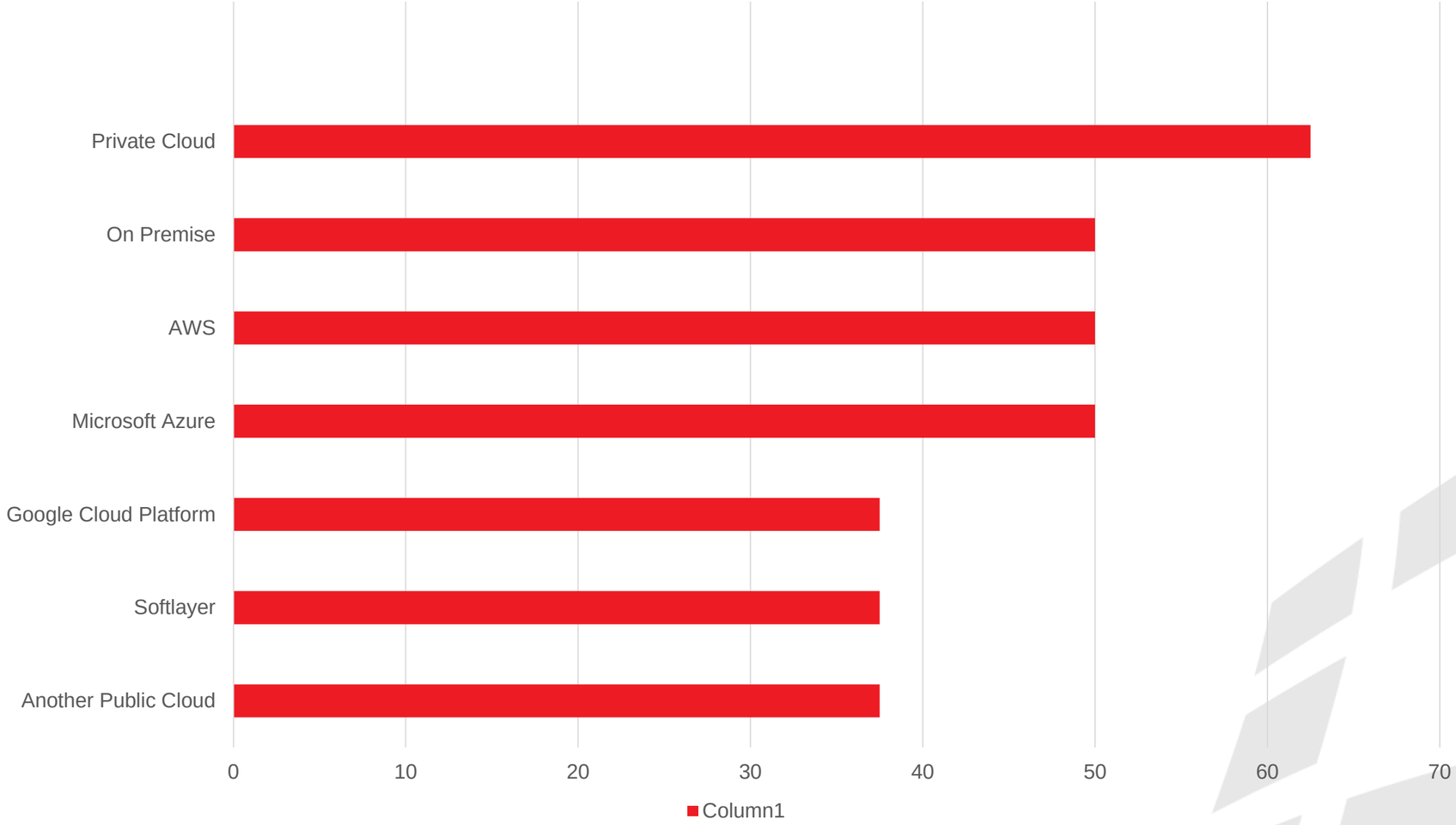
© 2016 GridGain Systems, Inc.

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



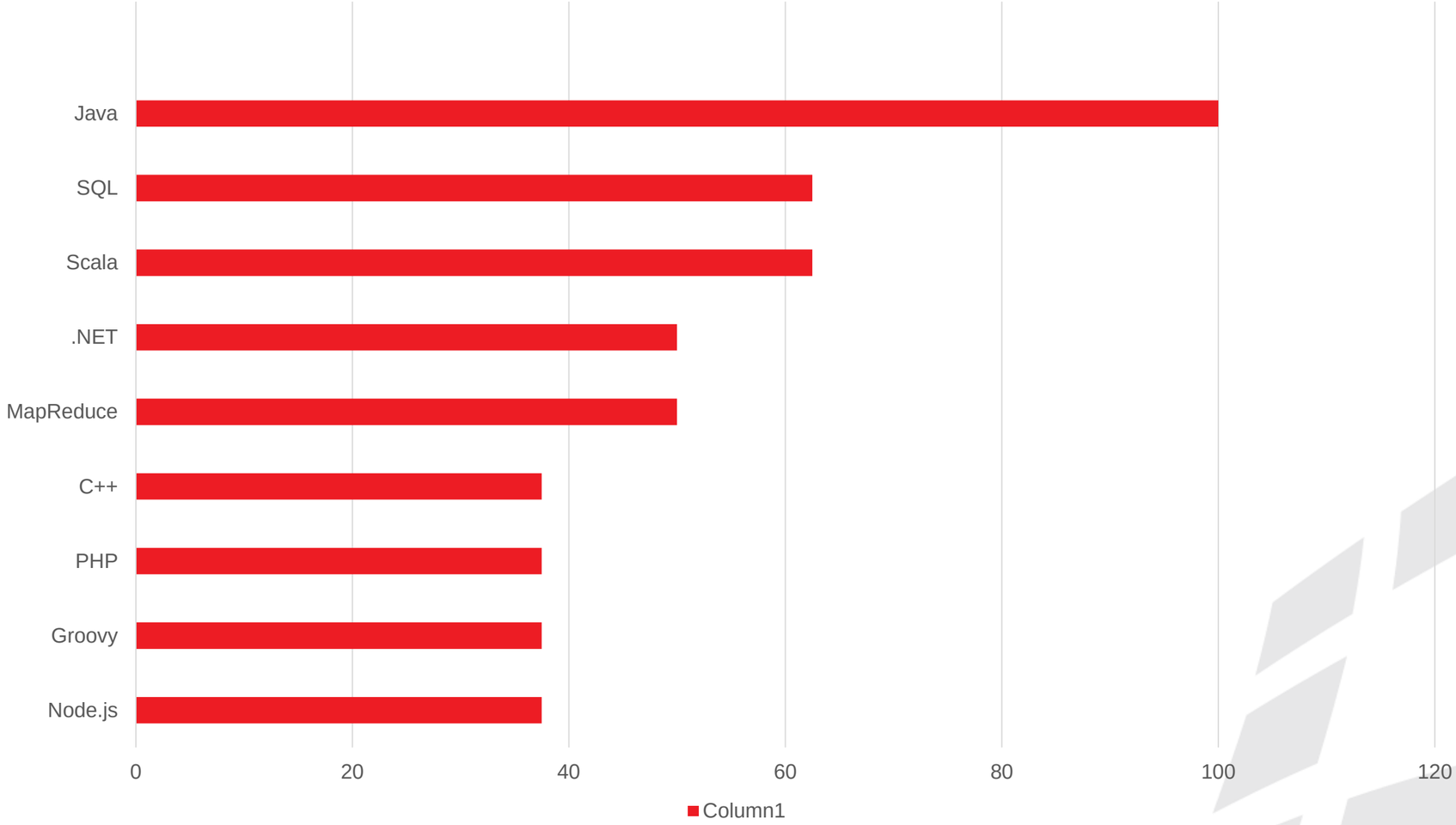
© 2016 GridGain Systems, Inc.

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



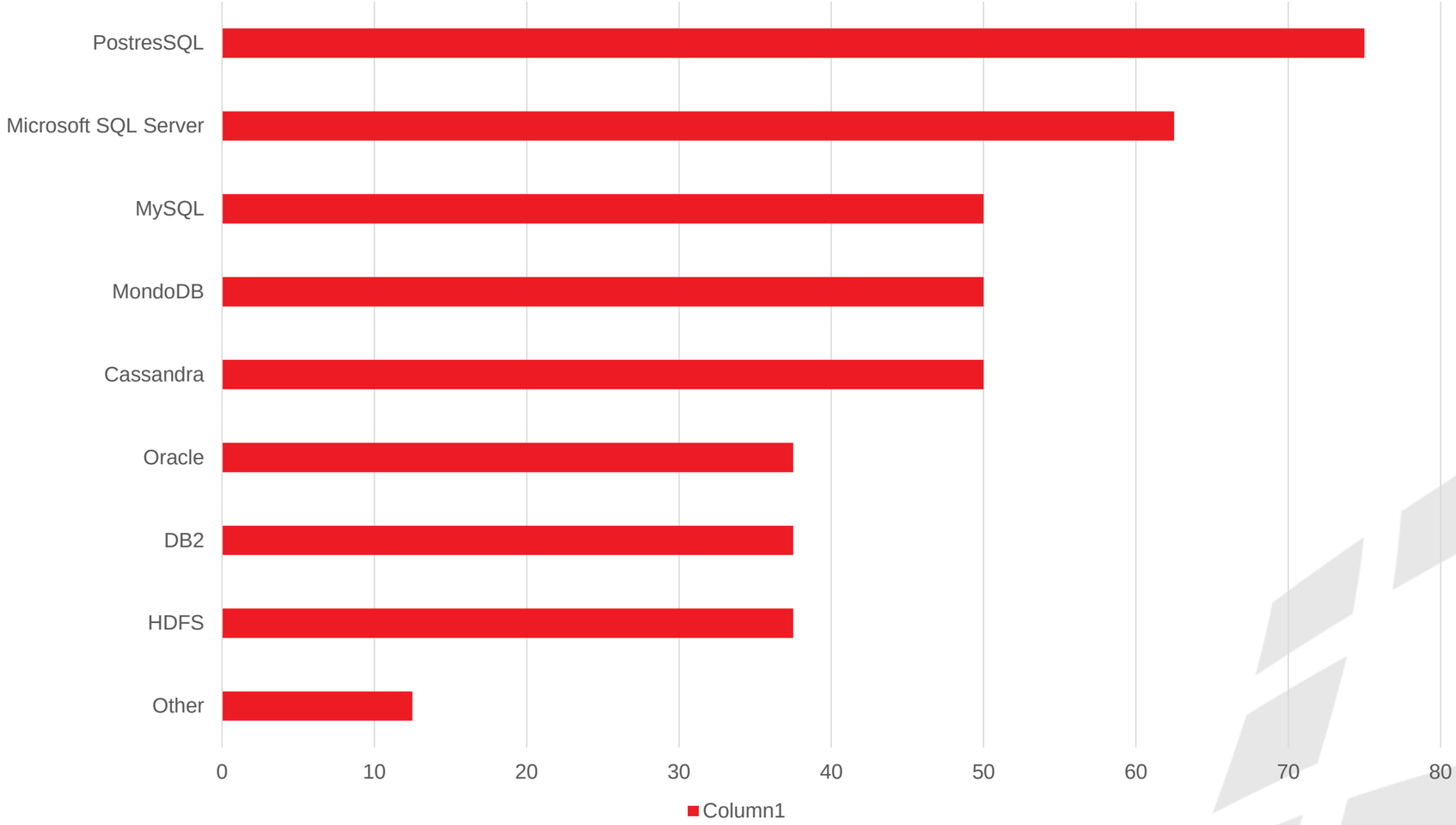
© 2016 GridGain Systems, Inc.

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



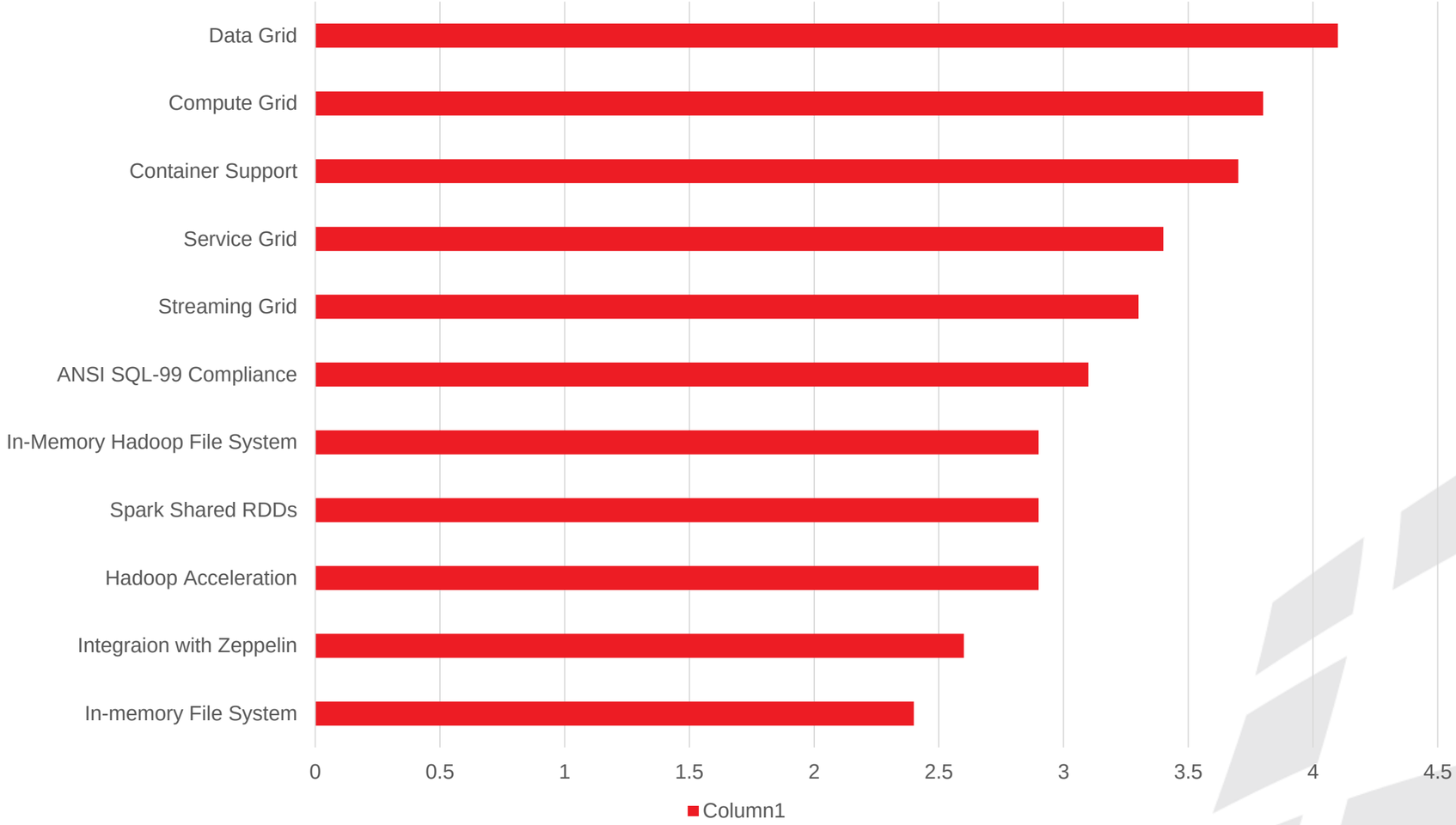
© 2016 GridGain Systems, Inc.

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



© 2016 GridGain Systems, Inc.

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



© 2016 GridGain Systems, Inc.

GridGain Ecommerce Customers

- Apple
- Newegg
- Canon
- SONY
- 24 Hour Fitness
- Avis
- InterContinental Hotels Group
- Experticity
- Cartera Commerce

© 2016 GridGain Systems, Inc.

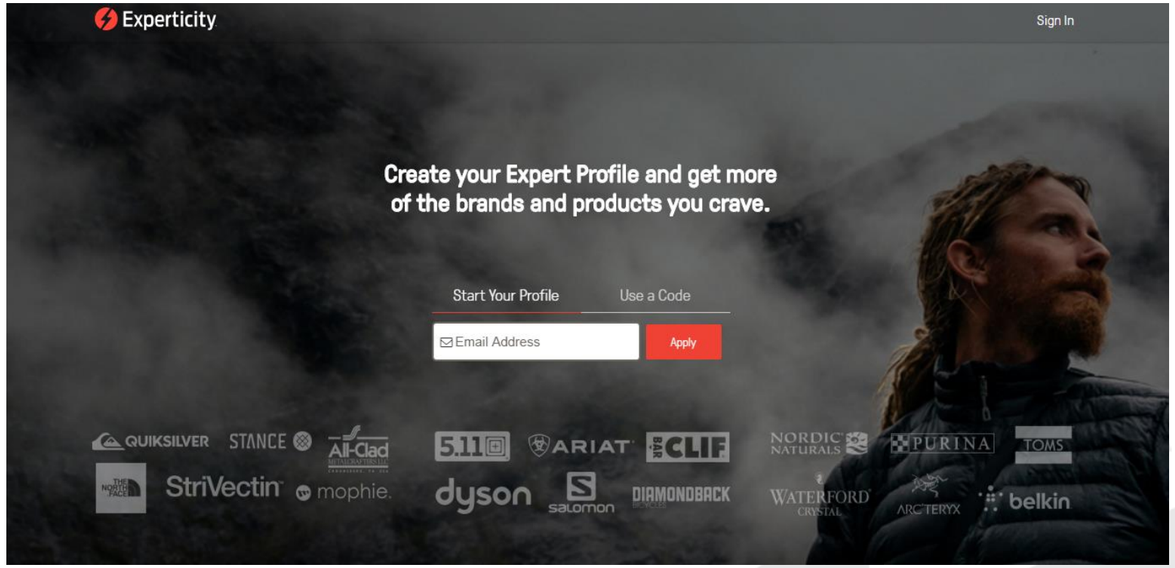
Ecommerce Use Case: Real Time Personalization and User Session Management

- React to user choices in real time
- Location and behavior based recommendations
- Store real time user session data
- Serve dynamic content with sub-millisecond latency
- Customized pricing on-the-fly

© 2016 GridGain Systems, Inc.

Case Study: Experticity

- Connects over 650 lifestyle and consumer brands (more than 4.5 million products) with brand advocates
- Brands use the network to identify experts and engage them in their sales channel
- Member profiles (interests, experience, content viewed, purchases)
- Real Time Analytics on thousands of active user sessions for personalized content



"We need it to work fast and work at scale, and GridGain does. We couldn't do what we do without GridGain, and we're confident we can scale to meet the growth we anticipate."
-Jeremy Knudsen, CTO

© 2016 GridGain Systems, Inc.

Ecommerce Use Case: High Speed Transactions

- High speed transactions create customer satisfaction and revenue growth
- Performance and scale required for entire spectrum of site functionality
- The user sees a page and a shopping cart.
- We see pages, elements (graphics, text), shopping cart, transactional elements (prices, inventory, shipping, payment)

© 2016 GridGain Systems, Inc.

Ecommerce Use Case: Real-time Analytics

- Real time analytics relies on multiple sources to enable upselling and cross-selling
  - User session analysis
  - Behavior based recommendations
  - Location-based offers
  - Trending items
  - Spot promotions
- Hybrid transaction and analytics use cases
- Where big data meets fast data
- Analysis across varied data sets and data structures
- Dynamic and customizable reporting for BI

© 2016 GridGain Systems, Inc.

Case Study: Real Time Analytics

Real-Time Reporting via home-grown analytics application

- Upgrade current In-Memory solution
  - Goal: Move to next-generation solution, while adding advanced-level ad hoc query capabilities
  - Desired Result:
    - Store 2TB of hot data in memory
    - Internal customers – run thousands of real-time ad hoc reports
  - Evaluation:
    - Competition – Terracotta & GemFire
    - GridGain demonstrated the ability to seamlessly integrate into homegrown analytics application, while adding the ad hoc query capabilities they desired

Advanced-Level SQL Support

Further offload work to the grid

Reduce trips to data warehouse

© 2014 GridGain Systems, Inc.

Ecommerce Use Case: Caching Product Catalogs

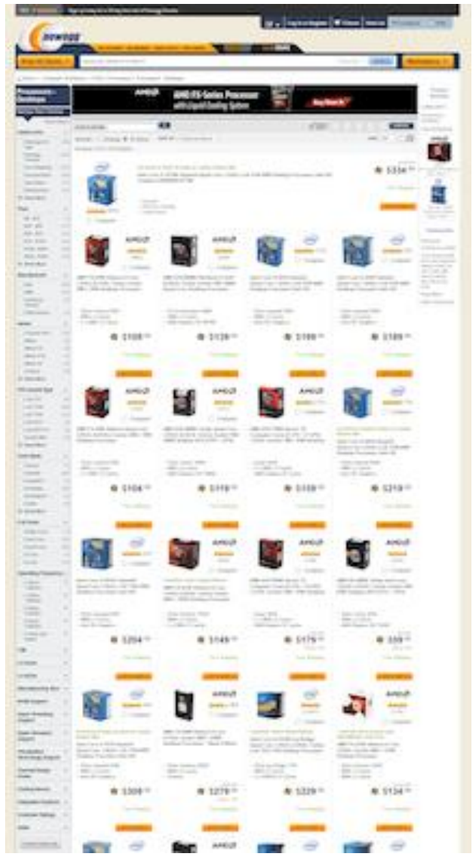
- Requirement: extreme responsiveness across multiple data stores
- Data is frequently stored and maintained in back-end RDBMS
  - Multiple images
  - Video
  - Text
- Intelligent caching across data structures, locations, back-ends
- Pre-sort and Pre-fetch based on complex rules
- Customizable expiration and eviction

© 2016 GridGain Systems, Inc.

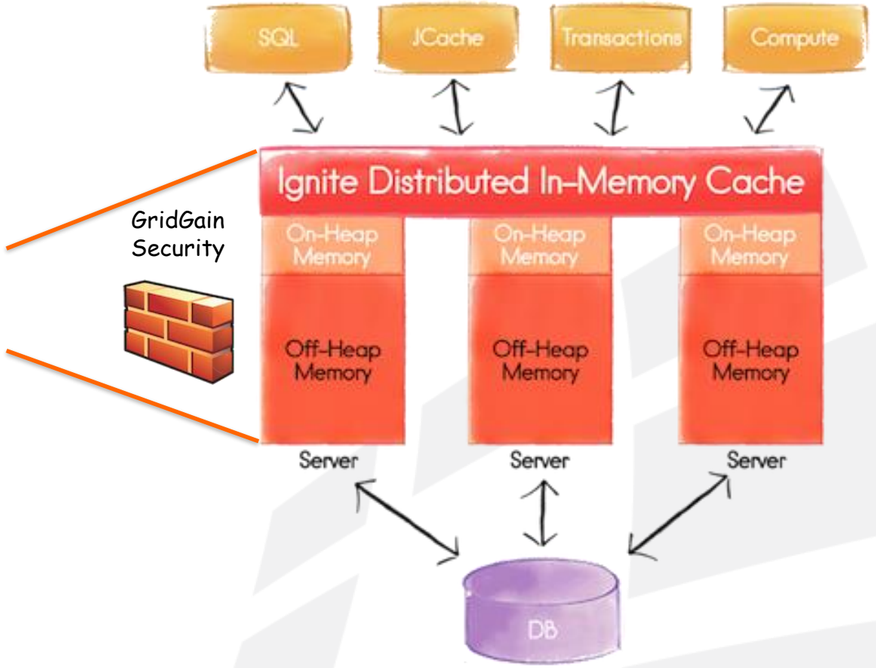
Case Study:

Online consumer electronics store – 11th largest retailer in US

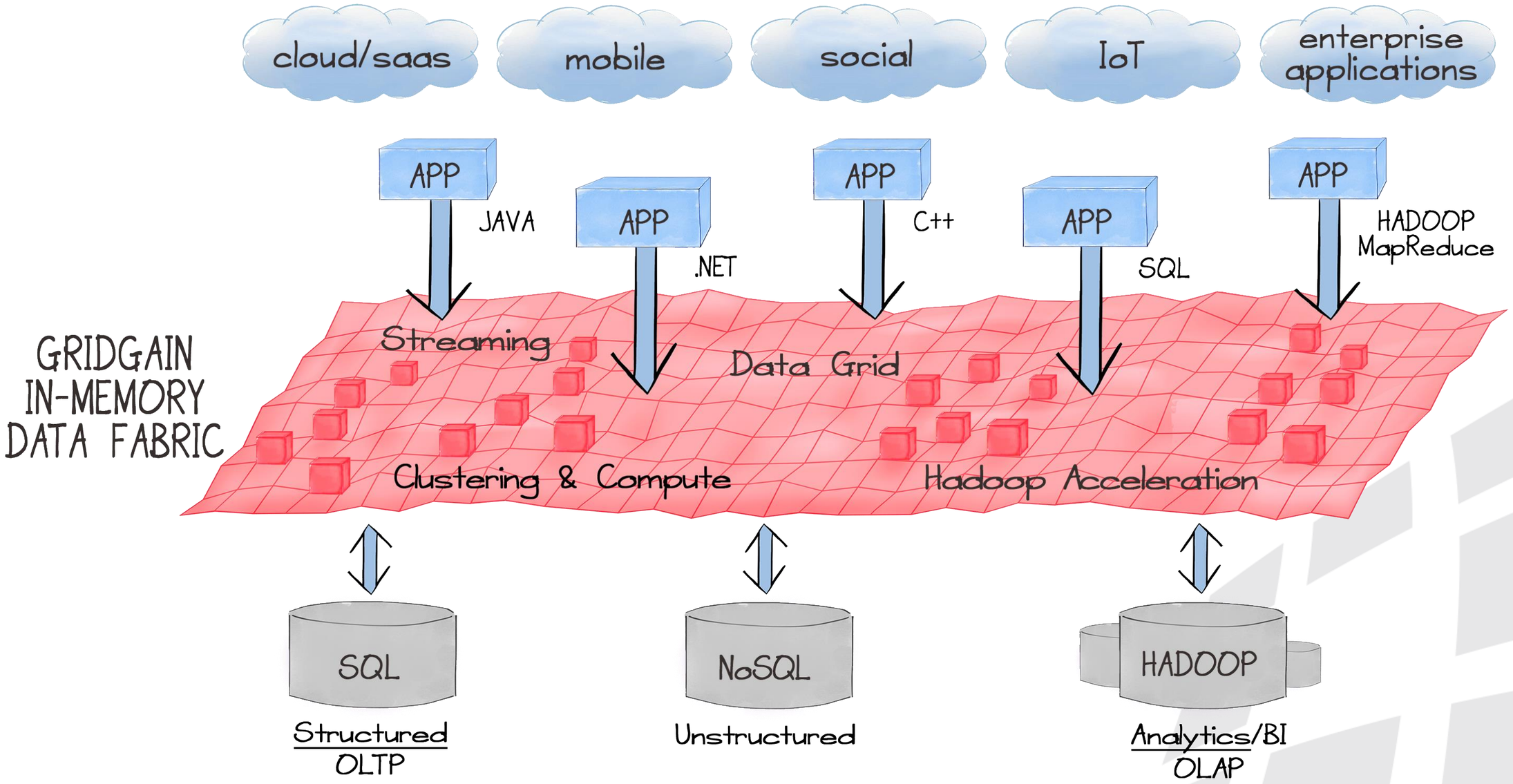
- Newegg Requirements
  - Oracle Coherence replacement
  - Smart in-memory analytics
  - Use open source technologies
- Why GridGain Won
  - Best performance
    - 2x better than incumbent
  - Demonstrated best
    - ANSI-99 SQL Support
    - Resiliency & fault tolerance
    - Transactional consistency



NewEgg Ecommerce Site



© 2015 GridGain Systems, Inc.

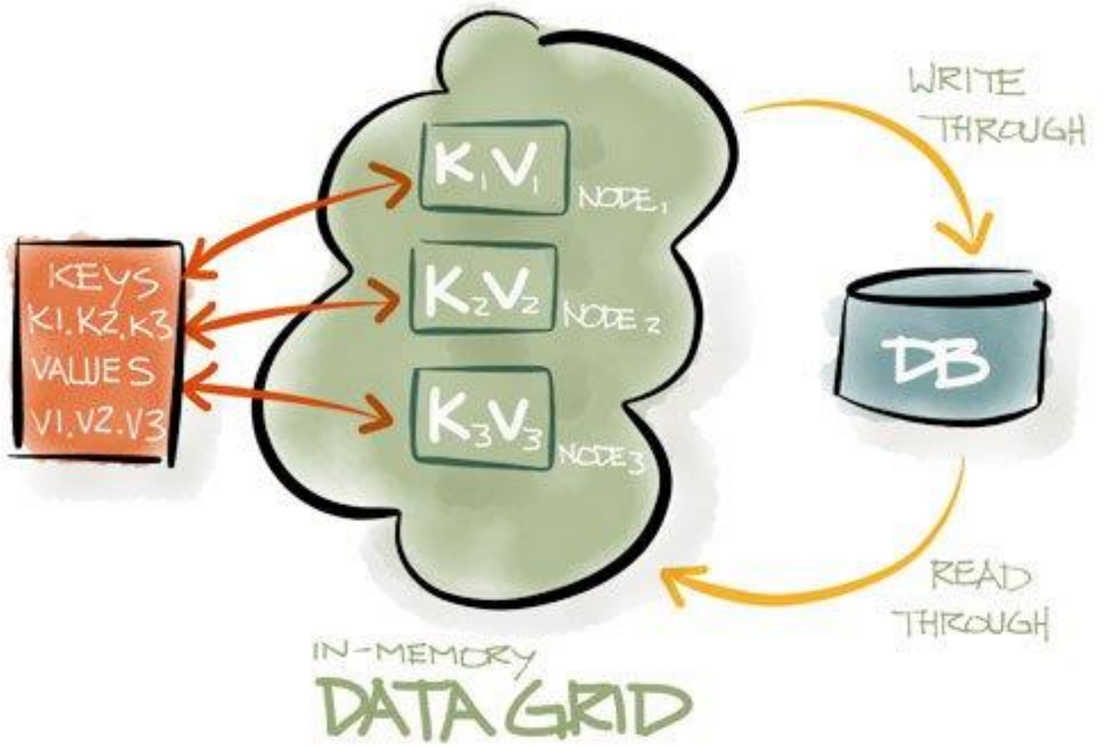


© 2016 GridGain Systems, Inc.

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*

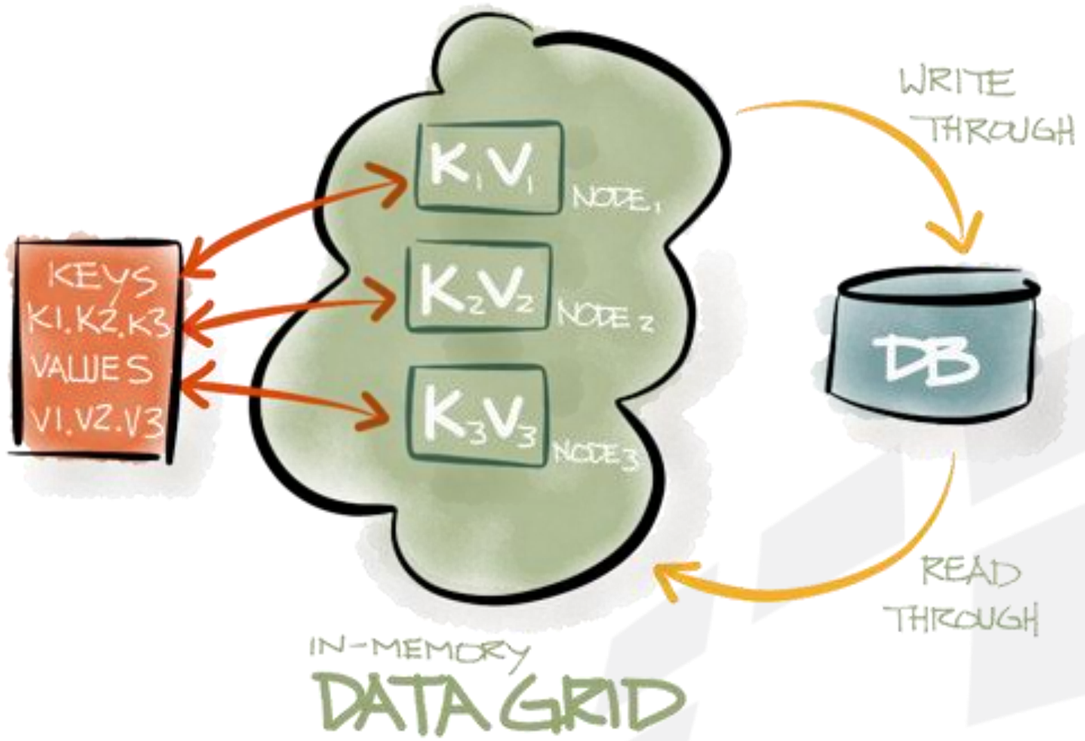


© 2016 GridGain Systems, Inc.

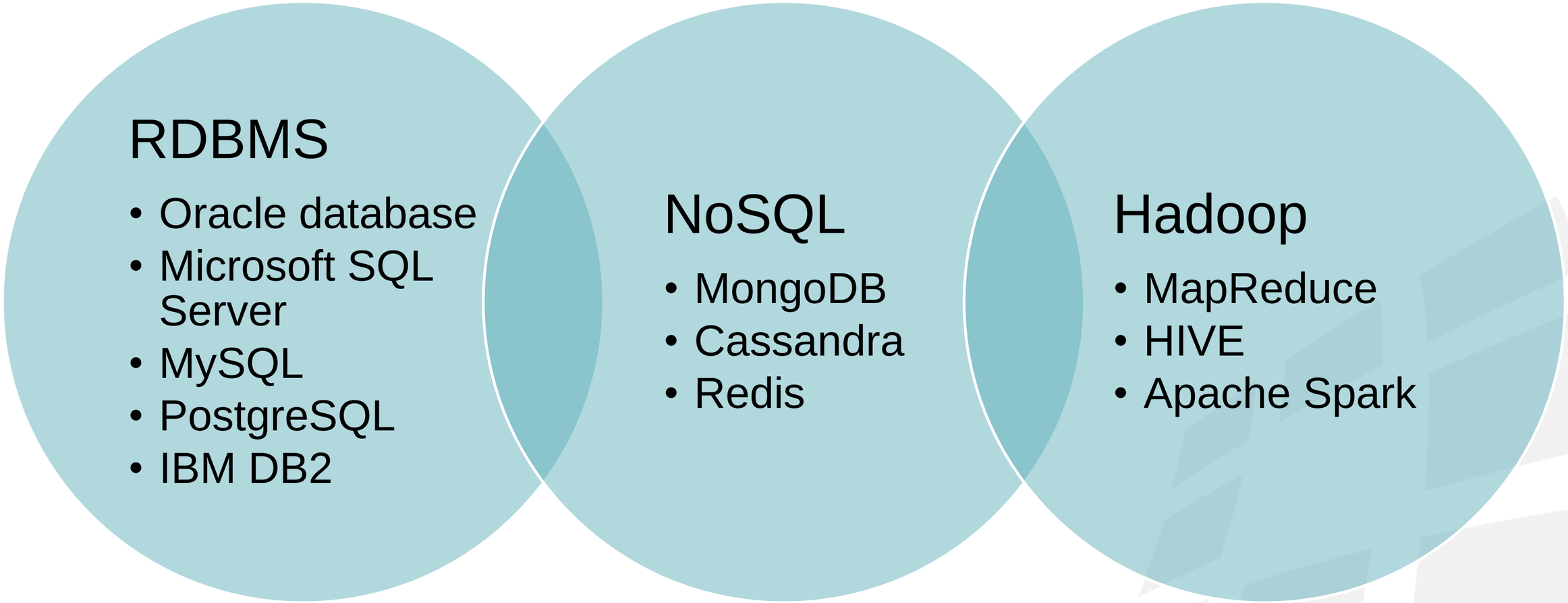
GridGain

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



Support for All Major Databases Including

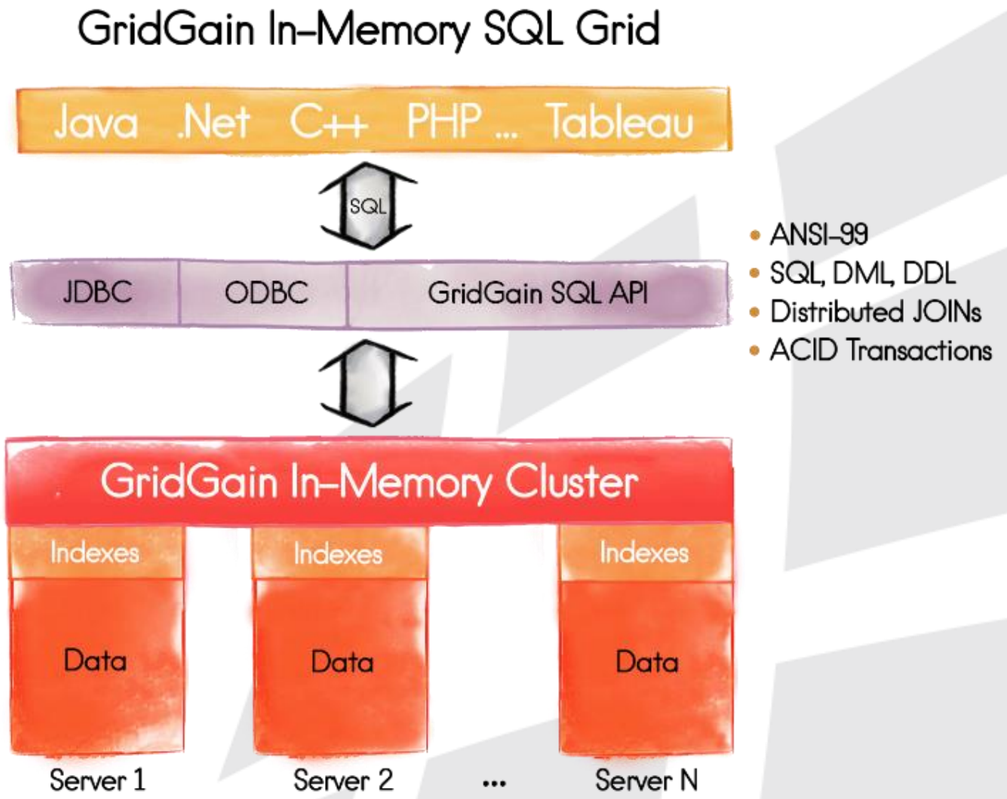


© 2015 GridGain Systems, Inc.

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*



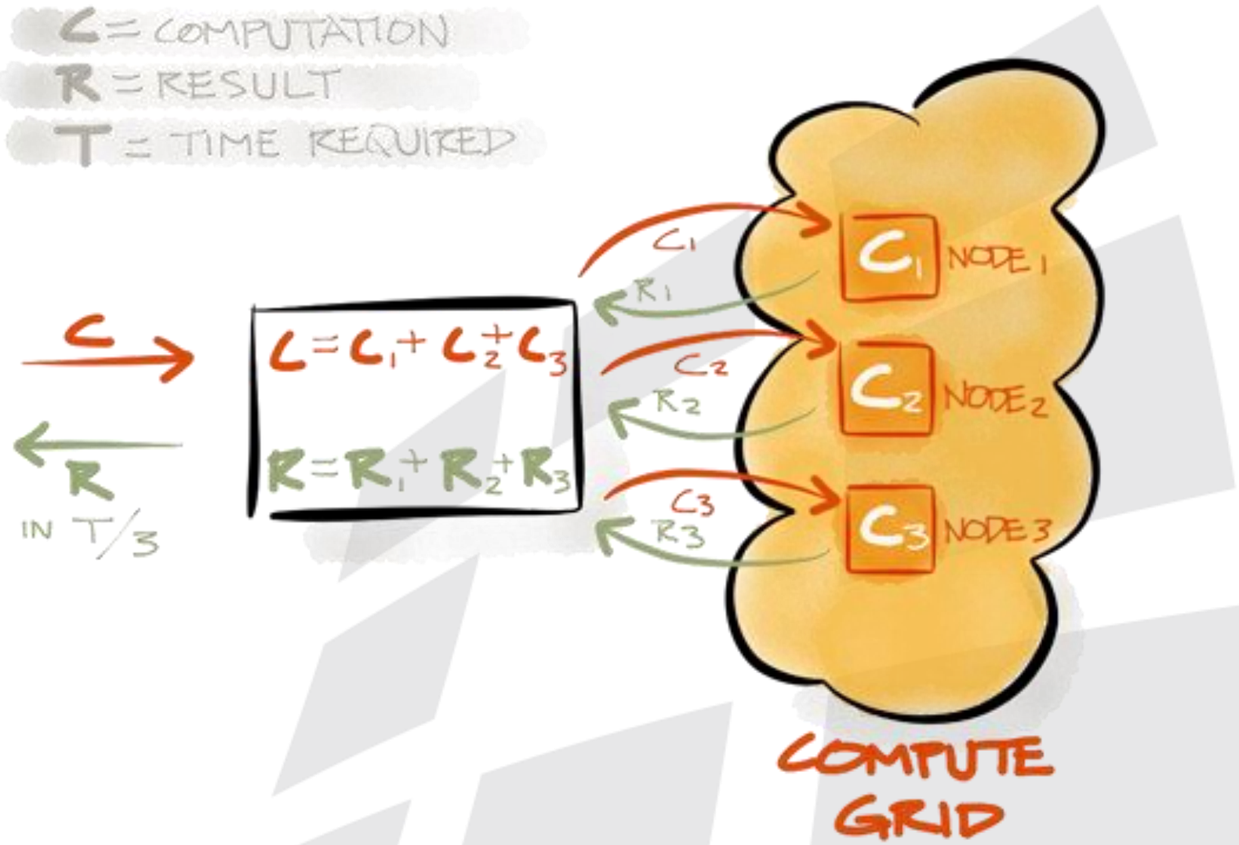
© 2016 GridGain Systems, Inc.

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*



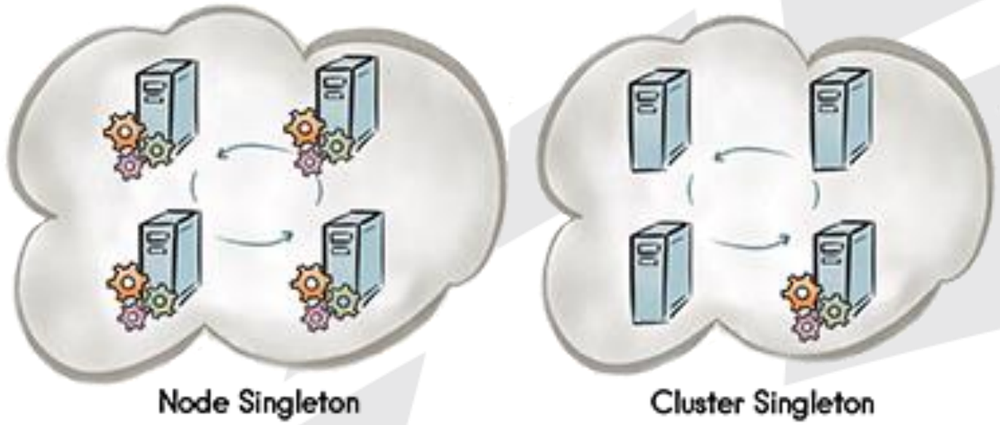
© 2016 GridGain Systems, Inc.

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*

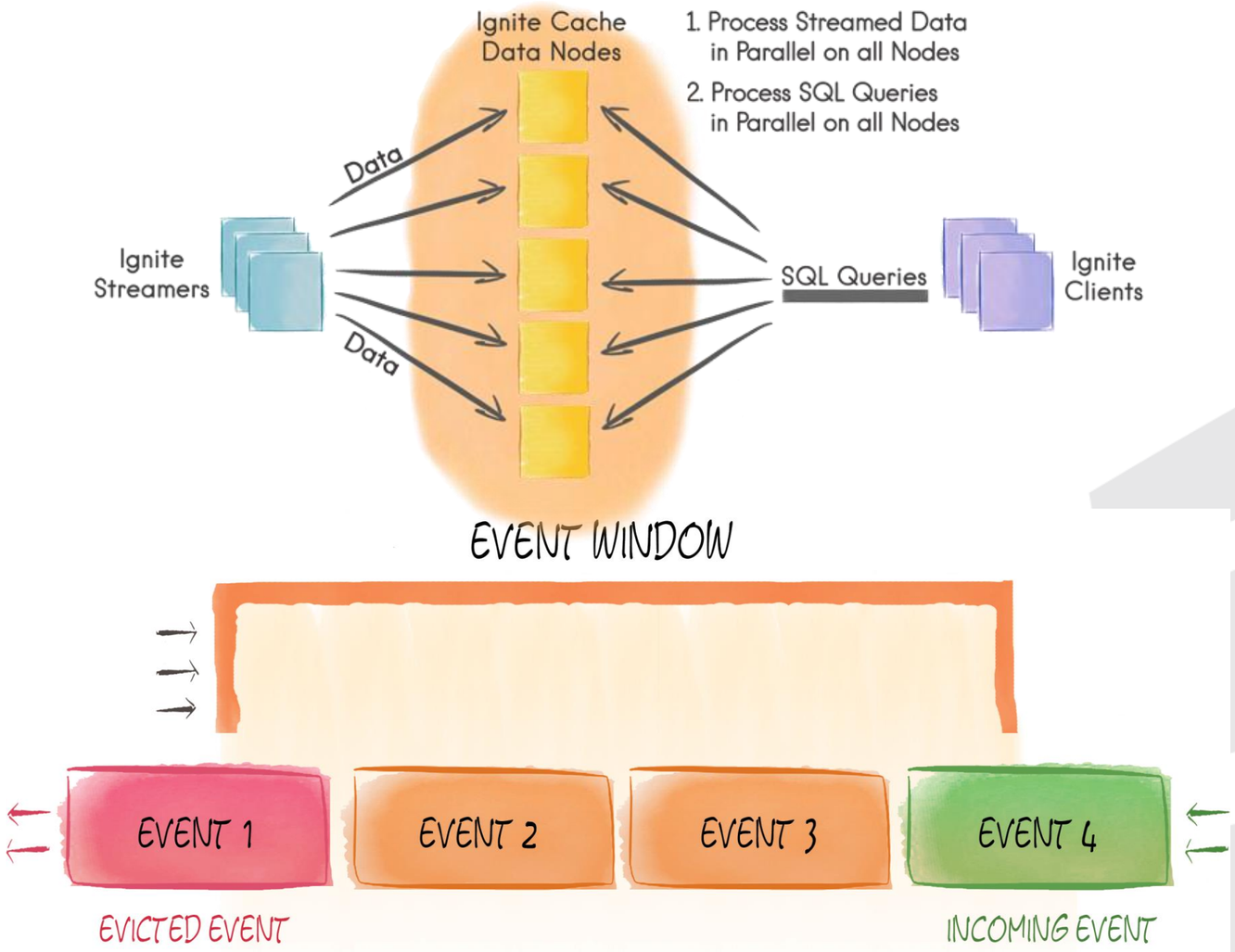


© 2016 GridGain Systems, Inc.

GridGain

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

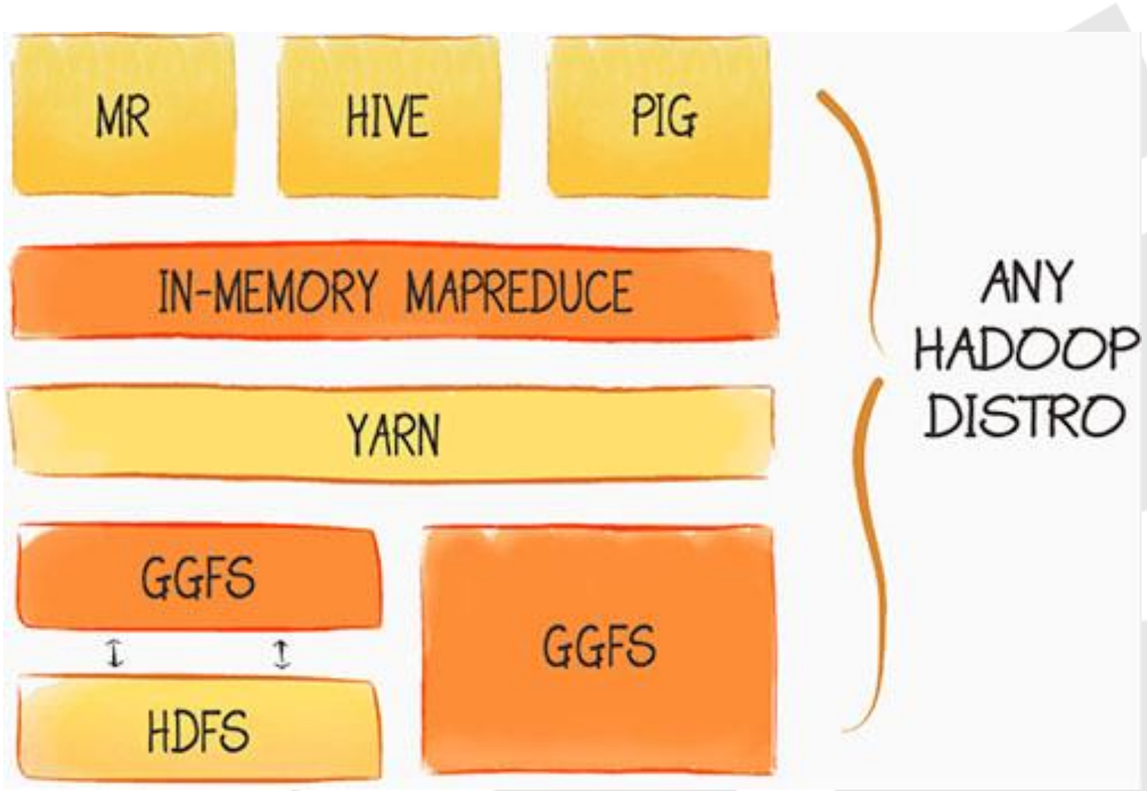


© 2016 GridGain Systems, Inc.

In-Memory Hadoop Acceleration

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

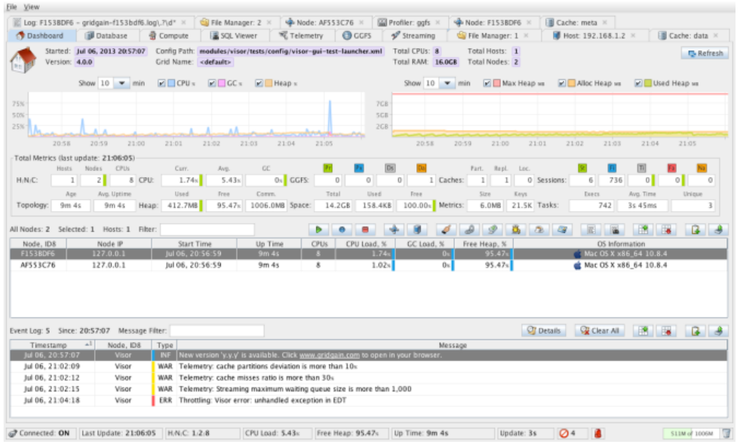
- ***Features:***
  - *Use existing MapReduce / Pig / Hive*
  - *10x Faster Performance*
  - *In-Memory MapReduce*
  - *Highly Optimized In-Memory Processing*
  - *Standalone File System*
  - *Optional Caching Layer for HDFS*
  - *Read-Through and Write-Through with HDFS*

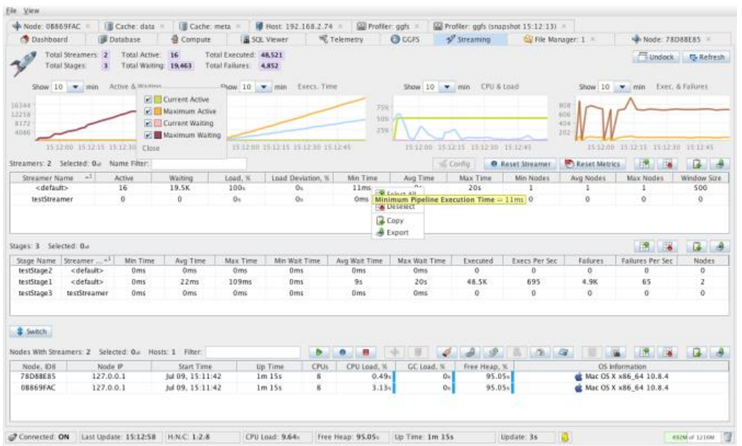


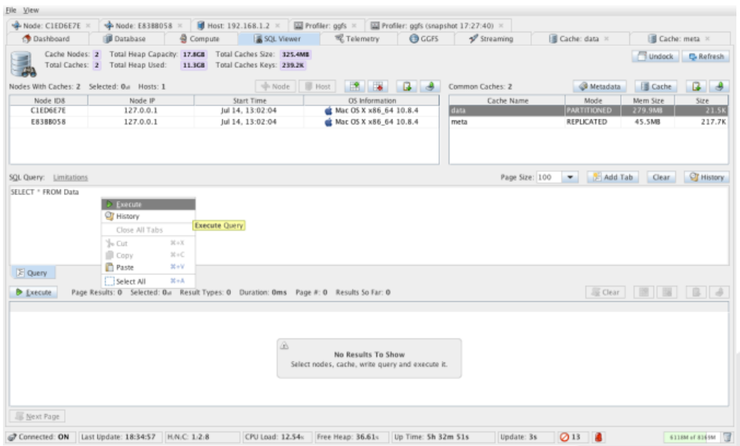
© 2016 GridGain Systems, Inc.

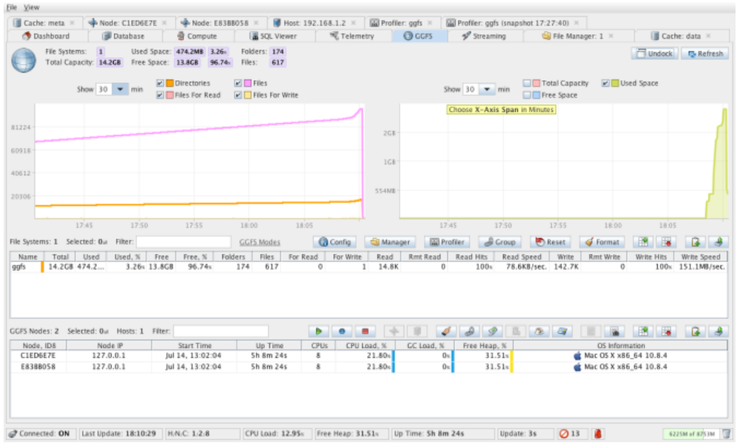
GridGain

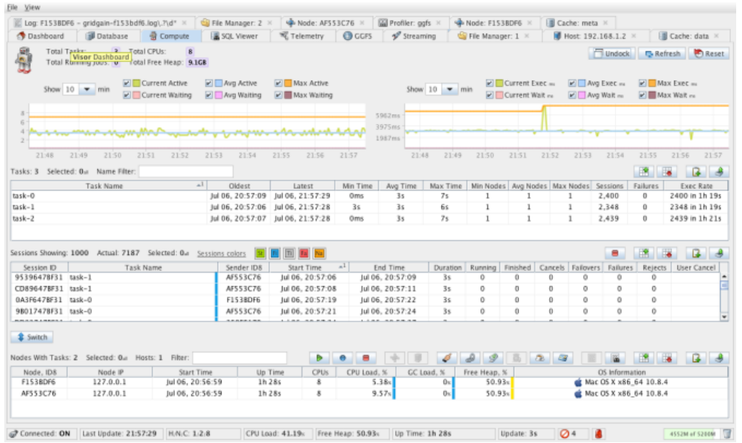
Visor: Monitoring & Mgmt for DevOps

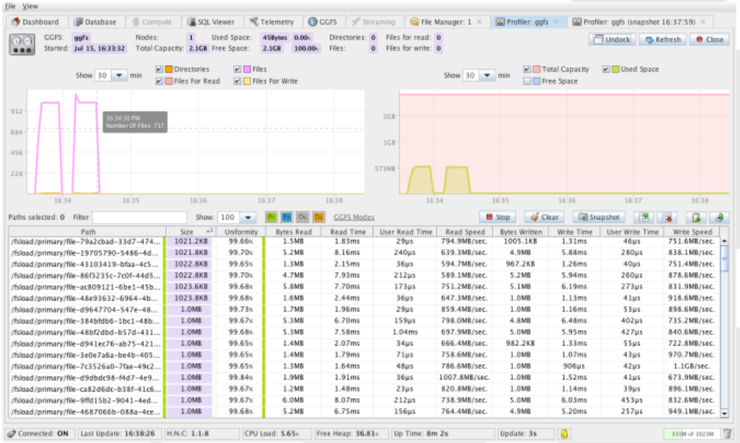












© 2016 GridGain Systems, Inc.

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

© 2016 GridGain Systems, Inc.