Accelerate Mobile Apps 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 Mobile Application Development

Over 1 billion smartphones

Roughly 179 billion mobile apps downloaded in 2016

Messaging, navigation, social media, readers, games, retail, banking, travel
Mobile Application Trends

Wearables
- According to IDC, 101.9 million wearable devices shipped in 2016
- Smartphone as a hub
- Wearables communicate with apps
- Enable wide range of products and services

IoT
- According to Gartner, 26 billion connected devices by 2020
- Includes app controlled smart objects
- Connected Home
Mobile Application Trends

Mobile Commerce
- Continuing to grow in popularity (dollars and users)
- Apple Pay and Google Wallet merge mobile and physical commerce
- Wearables and IoT devices
- Just beginning to scratch the surface of data collection and analysis

Motion and Location Sensing
- Know an individual’s location within 10 feet
- Provide location specific information, services, deals
- Motion sensing for security, games
- Precise indoor location sensing for personalized services, promotions and information
Mobile Application Trends

Innovative Mobile User Experience Design
- Effective display of data and content via mobile user interface
- Intuitive designs and interactive interface
- Mobile challenges of partial user attention and interruption

Application Performance Management (APM)
- Visibility into app behavior via infrastructure, network, device, app
- Statistics about device, OS, carrier
- Track user behavior and interactions
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
The In-Memory Computing Technology Market Is Big — And Growing Rapidly

IMC-Enabling Application Infrastructure ($M)

- $10 billion by year-end 2019 with a CAGR of over 22%.
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™

Features
- Data Grid
- Compute Grid
- SQL Grid
- Streaming
- Service Grid
- Hadoop Acceleration

Architecture
- Advanced Clustering
- In-Memory File System
- Messaging
- Events
- Data Structures
In-Memory
Scalable
No Rip & Replace

Application

Always Available
SQL 99 / ACID / MapReduce
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

<table>
<thead>
<tr>
<th>Features</th>
<th>Apache Ignite</th>
<th>GridGain Enterprise</th>
</tr>
</thead>
<tbody>
<tr>
<td>In-Memory Data Grid</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>In-Memory Compute Grid</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>In-Memory Service Grid</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>In-Memory Streaming</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>In-Memory Hadoop Acceleration</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Distributed In-Memory File System</td>
<td></td>
<td>✓</td>
</tr>
<tr>
<td>Advanced Clustering</td>
<td></td>
<td>✓</td>
</tr>
<tr>
<td>Distributed Messaging</td>
<td></td>
<td>✓</td>
</tr>
<tr>
<td>Distributed Events</td>
<td></td>
<td>✓</td>
</tr>
<tr>
<td>Distributed Data Structures</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Portable Binary Objects</td>
<td></td>
<td>✓</td>
</tr>
<tr>
<td>Management &amp; Monitoring GUI</td>
<td></td>
<td>✓</td>
</tr>
<tr>
<td>Enterprise-Grade Security</td>
<td></td>
<td>✓</td>
</tr>
<tr>
<td>Network Segmentation Protection</td>
<td></td>
<td>✓</td>
</tr>
<tr>
<td>Recoverable Local Store</td>
<td></td>
<td>✓</td>
</tr>
<tr>
<td>Rolling Production Updates</td>
<td></td>
<td>✓</td>
</tr>
<tr>
<td>Data Center Replication</td>
<td></td>
<td>✓</td>
</tr>
<tr>
<td>Integration with Oracle GoldenGate</td>
<td></td>
<td>✓</td>
</tr>
<tr>
<td>Basic Support (9×5)</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Enterprise Support (9×5 and 24×7)</td>
<td></td>
<td>✓</td>
</tr>
<tr>
<td>Security Updates</td>
<td></td>
<td>✓</td>
</tr>
<tr>
<td>Maintenance Releases &amp; Patches</td>
<td></td>
<td>✓</td>
</tr>
</tbody>
</table>

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

- Database Caching
- Application Scaling
- High Speed Transactions
- Real Time Streaming
- HTAP
- RDBMS Scaling
- Hadoop Acceleration
- Apache Spark Acceleration
- Faster Reporting
- Web Session Clustering
- MongoDB Acceleration
Survey Results: Where do you run GridGain and/or Apache Ignite?

- Private Cloud
- On Premise
- Amazon Web Services (AWS)
- Microsoft Azure
- Softlayer
- Google Cloud Platform
- Another Public Cloud

© 2017 GridGain Systems, Inc.
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?

- MongoDB: 35
- HDFS: 34
- Oracle: 31
- MySQL: 27
- PostgreSQL: 22
- Cassandra: 20
- Microsoft SQL Server: 18
- Other: 5
- DB2: 5
Survey Results: How important are each of the following product features to your organization?

- Data Grid
- Compute Grid
- In-memory File System
- Service Grid
- ANSI SQL-99 Compliance
- Streaming Grid
- Support for Mesos/YARN/Docker
- Spark Shared RDDs
- Hadoop Acceleration
- In-Memory Hadoop MapReduce
- Integration with Zeppelin
Case Study:

- **Background:**
  - Intelligentpipe is a big data software company serving the global telecommunications industry by developing solutions for mobile operators to improve their business and operational processes

- **The Challenge**
  - Collect and analyze massive amounts of mobile user traffic data in real time
  - Tens of millions of users
  - Consumption of network resources
  - Type of network traffic (voice or data)
Case Study: GridGain Systems, Inc.

• GridGain Professional Edition used to build a high performance low latency analysis platform
• “GridGain ensures responsiveness regardless of how much information we need to search through.” Sakari Paloviita, CTO, Intelligentpipe
• Collect and analyze multiple terabytes per day
Case Study: GridGain Systems, Inc.

- Real-time analytics provides fast insight
- Easy integration with existing systems due to GridGain’s Unified API and ANSI SQL-99 support
- Linear scaling across deployed server to keep up seamlessly as the business grows

We’ll want to use technology GridGain offers so we can focus on our core business ourselves.”
- Jari Kuusela, Director of Product Management
Mobile Apps Use Case: High Speed Transactions

- High speed transactions create customer satisfaction, increase user base and revenue
- Performance and scale required for entire spectrum of app/infrastructure functionality
- The user sees content (product catalog, reviews, etc) and a shopping cart.
- Developers see pages, elements (graphics, text), shopping cart, transactional elements (prices, inventory, shipping, payment)
GridGain Provides High Speed Transactions

- In-memory is roughly 1,000x faster than disk
- Distributed compute and data create additional speed gains
- ACID compliant transactions
- ANSI SQL-99 compatibility for interoperability to other systems like inventory management, shipping, and analytics like fraud detection
Case Study:

• Background:
  – Cyber Dust is a platform for text messages: “A safer place to text.”
  • Untraceable
  • Encrypted
  • Disappearing
  • Screenshot blocking
  – Available for Android and iOS
  – Mark Cuban funded

• The Challenge
  – To build a real-time, reliable and highly available server infrastructure to support a mobile messaging platform
  – More than 500K users
  – Millions of messages a day
  – Avoid writing messages to disk
Case Study:

- GridGain Professional Edition used to build a messaging platform
- Runs completely on Amazon EC2
- All user account data, configurations, and messages held in memory
- Messages deleted without a trace because they were never written to disk
- Extensive use of Unified API
Case Study:

- “Blast” feature performance: capable of broadcasting disappearing messages to all of a user’s contacts
- Real-world performance of 300,000 messages sent and disappeared in 30 seconds

I was pleasantly surprised by the GridGain solution and performance.  
-Igor Shpitalnik, CTO

I keep learning about additional capabilities GridGain offers. It’s what I expected and more.  
-Igor Shpitalnik, CTO
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™

Features
- Data Grid
- Compute Grid
- SQL Grid
- Streaming
- Service Grid
- Hadoop Acceleration

Architecture
- Advanced Clustering
- In-Memory File System
- Messaging
- Events
- Data Structures
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
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
**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

© 2017 GridGain Systems, Inc.
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
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