Apache Ignite and Apache Spark
Where Fast Data Meets the IoT

Denis Magda
GridGain Product Manager
Apache Ignite PMC

http://ignite.apache.org
#apacheignite
#denismagda
Agenda

- IoT Demands to Software
- IoT Software Stack
  - Device OS/RTOS
  - Data Collection and Enrichment
  - HTAP Platform
  - Application APIs
- Demo
IoT Demands to Software
IoT Demands to Software

- Real-time ingestion
- Real-time processing
- Time-series support
- Flexible Querying APIs
  - SQL
  - Full-text search
  - Geo-spatial
- Analytics
  - BI
  - Machine Learning
- High-availability
- Simple scalability
IoT Software Stack
IoT Software Stack

Application APIs

HTAP Platform

Data Collection and Enrichment

Device OS / Real-Time OS
Apache IoT Software Stack

- Application APIs
- HTAP Platform
- Data Collection and Enrichment
- Device OS / Real-Time OS
Device OS/RTOS
Apache MyNewt

- Open Source RTOS
  - Cortex M0-M4
  - MIPS & RISC-V
- Networking
  - Bluetooth Low Energy
  - Wi-Fi
  - TCP/IP & UPD
- Build & Package Management
- Secure bootloader and signed images
- Remote Firmware Upgrade
Data Collection and Enrichment
Data Collection and Enrichment

• Spark Streaming
  – Fault-Tolerant Streams Processing
• Data Collection
  – Sockets
  – Kafka
  – Flume
• Data Enrichment
  – Spark API
• Data Storage
  – Apache Ignite
HTAP Platform
Use Case:

Smart Metering and Utilities – delivers a comprehensive IOT platform

- **SilverSpring Requirements**
  - Migrate to in-memory processing
  - Add scalability & elasticity
  - Use open source technologies

- **SilverSpring Solution**
  - SaaS Architecture
    - Multi-Tenancy
    - Advanced Security
  - Strong compute capabilities
    - Co-located in-memory processing
  - Demonstrated best
    - On-demand elasticity & scalability
    - ANSI-99 SQL Support
    - Transactional consistency
Apache Ignite In-Memory Data Fabric

- Supports Applications of various types and languages
- Open Source – Apache 2.0
- Simple Java APIs
- 1 JAR Dependency
- High Performance & Scale
- Automatic Fault Tolerance
- Management/Monitoring
- Runs on Commodity Hardware
- Supports existing & new data sources
- No need to rip & replace
In-Memory Data Grid

- Distributed Key-Value Data Store
- Data Reliability
- High-Availability
  - Active replicas, automatic failover
- Data Consistency
  - ACID distributed transactions
In-Memory Data Grid: Partitioning

Partitioned Cache

Replicated Cache
Streaming to Ignite

- **Ignite Data Streamer**
  - Fastest Data Injection
  - Automatic data partitioning
- **Stream Receivers**
  - Custom logic execution
  - Additional data transformation
- **Stream Adapter**
  - Socket
  - Kafka
  - Flink
  - RocketMQ
  - etc.
Application APIs
Apache Ignite SQL Grid

- ANSI-99 SQL
- Geo-spatial Queries
- Full-text Search
- Always Consistent
- Fault Tolerant
- Cross-Platform
  - JDBC and ODBC drivers
  - DML (INSERT, UPDATE, etc.)
  - DDL (CREATE, DROP, etc.)
Apache Ignite SQL Grid: Queries

- Distributed Joins
- Automatic Group By, Aggregations, Sorting
- Cross-Cache Joins, Unions, etc.
- Ad-Hoc SQL Support
Apache Ignite Compute Grid

• Direct API for MapReduce
• Direct API for ForkJoin
• Zero Deployment
• State Checkpoints
• Load Balancing
• Automatic Failover
Spark Shared RDDs

- **IgniteRDD**
  - Share RDD across jobs on the host
  - Share RDD across jobs in the application
  - Share RDD globally

- **Faster SQL**
  - In-Memory Indexes
  - SQL on top of Shared RDD
Machine Learning Grid

**DSL**
Domain Specific Language

**Execution Planner**
Local “last-mile” optimizer

**Distributed Core Algebra**
Dense and sparse algebra + data preparation

**Apache Ignite**
Compute & Data Grids

Pre-Built Algorithms  ML Grid Components
Resources

• Documentation:
  – Apache Ignite: https://apacheignite.readme.io/docs
  – Apache Spark Streaming: http://spark.apache.org/streaming/
  – Ignite and Spark Integration: https://apacheignite-fs.readme.io/docs/ignite-for-spark
  – Apache MyNewt: http://mynewt.apache.org

• Demo Source Code:
  – https://github.com/dmagda/IgniteSparkIoT
ANY QUESTIONS?

Thank you for joining us. Follow the conversation.

http://ignite.apache.org

#apacheignite  #denismagda