AN END-TO-END OPEN SOURCE ARCHITECTURE FOR THE INTERNET OF THINGS (IOT)

Dave Shuman, Industry Lead for IoT, Cloudera
James Kirkland, Chief Architect for IoT, Red Hat
CHALLENGE: PULLING TOGETHER THE BUILDING BLOCKS FOR IoT

1. Data endpoints
2. Field connectivity
3. Data acquisition & control
4. Data pre-processing
5. Edge intelligence & analytics
6. Back-end connectivity
7. Device Management
8. Information processing & integration
9. Data Management & Advanced Analytics
10. BI & Machine Learning
11. Enterprise Applications & Services

Enterprise IT & Business Applications
Data Management | Advanced Analytics
IoT Integration Platform
IoT / M2M Communication
Multi-service Gateway
KEY CHARACTERISTICS FOR AN END-TO-END IoT ARCHITECTURE

- **Modular Architecture**
  Integrated yet independent components that you can utilize to make best of your existing investments

- **Open & Interoperable**
  Based on open source and open standards to ensure interoperability and accelerate innovation

- **End-to-End Security**
  Enterprise-grade data security and compliance – from the edge all the way through to storage and access

- **Deployment Flexibility**
  Deploy anywhere you want it – on the edge, on any of the leading cloud platforms or in your data center
Device Management & Connectivity
Securely connect, authenticate and manage disparate connected devices that speak different protocols

Intelligent Edge Processing & Analytics
Apply analytics at the edge with machine learning and business rules to enable local, low-latency decision making

Advanced Analytics & Machine Learning
Centralize IoT data processing, analytics and machine learning to enable deep business insights and actionable intelligence

Business & Application Integration
Enable integration with enterprise and business applications to bridge the gap between OT and IT and reduce complexity

End-to-End Security & Compliance
Tools to enable end-to-end data security, compliance, authorization and authentication
Why Open Source for IoT?

- Use open standards
- Take advantage of community innovation and advances
- Wider integration with platforms and devices
- Accelerate time to market to deliver solutions
- Maintain control of your IoT technology selection and licensing
- Iterate more quickly with less risk
- Lay the foundation for future system and requirements scaling
OPEN SOURCE DEVELOPMENT DRIVES RAPID INNOVATION

64% think open source is very or extremely important in their deployment of IoT

“We believe the best way to support this complex environment is to base our commercial IoT platform, the Bosch IoT Suite, on open source components and open standards. These projects establish a horizontal open technology for IoT and provide the technical breeding grounds for successful business ecosystems.”

- Dr. Stefan Ferber, VP of Engineering, Bosch Software Innovations

Source: IDC, Global IoT Decision Maker Survey, August 2016
ECLIPSE IoT OVERVIEW

- 2.2M lines of code
- 26 projects
- 210+ developers
- 132K monthly visitors

Sponsors:
- Bosch
- Samsung
- Eurotech
- Red Hat
- GE
- Google
- Siemens
- Cloudera
- Ericsson
- SAP
BASED ON OPEN STANDARDS
OPEN END-TO-END IoT ARCHITECTURE
Integrating IoT operating technology, data management, analytics, and applications

- Modular, secure, end-to-end architecture
- Streaming analytics and machine learning
- Open, interoperable on hybrid cloud
- Modern application agility and integration
CONNECTED "THINGS"

**Sensors, Actuators, Data Sources**

**Edge Processing & Analytics**
- Device connectivity
- Data transformation
- Intelligent routing
- Business logic
- Edge analytics & real-time decisions

**IoT Gateways**

**Data Integration, Routing, Device Command/Control**

**Advanced Analytics & Machine Learning**

**Application Development, Deployment, Integration**

**IoT Integration Hub**

**Data Management & Analytics Platform**

**Enterprise Applications**
DATA INTEGRATION, ROUTING, DEVICE COMMAND/CONTROL

Sensors, Actuators, Data Sources

Edge Processing & Analytics

Data Integration, Routing, Device Command/Control

Advanced Analytics & Machine Learning

Application Development, Deployment, Integration

CONNECTED "THINGS"

IoT GATEWAYS

IoT INTEGRATION HUB

Telemetry Data

• Device management, security, and access control
• Data aggregation
• Event processing
• Integration services

DATA MANAGEMENT & ANALYTICS PLATFORM
ADVANCED ANALYTICS AND MACHINE LEARNING

CONNECTED “THINGS”

- Sensors, Actuators, Data Sources
- Edge Processing & Analytics
- Data Integration, Routing, Device Command/Control
- Advanced Analytics & Machine Learning
- Application Development, Deployment, Integration

IoT GATEWAYS

- Telemetry Data

IoT INTEGRATION HUB

- Telemetry Data
- Machine Learning Model
- DATA MANAGEMENT & ANALYTICS PLATFORM
- cloudera (CDH)

- Data ingest
- Stream / batch processing
- Secure data storage
- Machine learning and real-time analytics

ENTERPRISE APPLICATIONS

ADVANCED ANALYTICS AND MACHINE LEARNING

Machine Learning Model
Sensors, Actuators, Data Sources

CONNECTED “THINGS”

IoT GATEWAYS

Edge Processing & Analytics

Data Integration, Routing, Device Command/Control

Advanced Analytics & Machine Learning

Application Development, Deployment, Integration

APPLICATION DEVELOPMENT, DEPLOYMENT, INTEGRATION

DATA MANAGEMENT & ANALYTICS PLATFORM

- Application lifecycle management
- Integration services
- Self-service provisioning
- Hybrid cloud portability via containers

APPLICATION DEVELOPMENT, DEPLOYMENT, INTEGRATION

DATA MANAGEMENT & ANALYTICS PLATFORM

- Application lifecycle management
- Integration services
- Self-service provisioning
- Hybrid cloud portability via containers
END-TO-END DATA MANAGEMENT & ANALYTICS FOR IoT

Data Sources

Data Collection/Aggregation

Data Ingest

Data Storage & Processing

Advanced Analytics & Machine Learning

IoT Gateway

IoT Integration Hub

IoT Data Management

Connected Machines/ Data Sources

Enterprise IT & Business Applications

Visualization / BI

Advanced Analytics

Machine Learning

Data Streams

Machine Learning Models & Control

Enterprise Data Hub

OpenShift

Amazon Web Services

Microsoft

Google Cloud Platform
END-TO-END ANALYTICS
Data flow to derive deep business insights and actionable intelligence

Protocol Translation  Intelligent Filtering  Aggregation  Routing

CONNECTED “THINGS”  IoT GATEWAYS  IoT INTEGRATION HUB  ENTERPRISE APPLICATIONS

Telemetry Data  Management  Application Integration  Application Data

Deep data analysis & insights

DATA MANAGEMENT & ANALYTICS PLATFORM

Data Ingest  Real-Time Processing  Data Storage
Machine Learning  Real-Time Analytics  Data Security
END-TO-END ANALYTICS
Data flow to derive deep business insights and actionable intelligence
Industry 4.0 Demo - Business challenges

The scenario

- A manufacturing company is incurring **unplanned machine downtime** leading to increased maintenance and costs.
- Lack of a **real-time view on asset condition** makes it difficult to predict downtime and therefore optimize factory productivity.
- Lack of **business and operations visibility** due to inability to correlate real-time with historic data results.

The goal

- Optimize production factory capacity by scheduling **planned maintenance windows**.
- Predict failures on connected machines to **reduce downtimes & increase margins**.
INDUSTRY 4.0 DEMO ARCHITECTURE

- **IoT Gateways**
  - Device connectivity
  - Data transformation
  - Intelligent routing
  - Business logic
  - Edge analytics & real-time decisions

- **IoT Integration Hub**
  - Device management, security, and access control
  - Data aggregation
  - Event processing
  - Integration services

- **Data Mgmt, Analytics & ML**
  - Data ingest
  - Stream / batch processing
  - Secure data storage
  - Machine learning and real-time analytics

**REMOTE MAINTENANCE & SUPPORT**

**MACHINE LEARNING & ADVANCED ANALYTICS**

**ON-SITE DATA ANALYTICS**
EXAMPLE USE CASE: INDUSTRY 4.0 DEMO ARCHITECTURE: PROJECTS

Connected “Things”

IOT Gateways

MQTT Telemetry Data

Management

IOT Integration Hub

App Integration

Enterprise Applications

JDBC

App Data

Telemetry Data

Management

Machine Learning Model

Cloudera’s Distribution including Hadoop (CDH)

Spark

Kudu

kafka

ActiveMQ

Apache Camel

Kapua

Openscoring.io

Apache Camel

OpenShift Origin

Future integration with Drools

EXAMPLE USE CASE: INDUSTRY 4.0 DEMO ARCHITECTURE: PROJECTS

Connected “Things”

IOT Gateways

MQTT Telemetry Data

Management

IOT Integration Hub

App Integration

Enterprise Applications

JDBC

App Data

Telemetry Data

Management

Machine Learning Model

Cloudera’s Distribution including Hadoop (CDH)

Spark

Kudu

kafka

ActiveMQ

Apache Camel

Kapua

Openscoring.io

Apache Camel

OpenShift Origin

Future integration with Drools

EXAMPLE USE CASE: INDUSTRY 4.0 DEMO ARCHITECTURE: PROJECTS

Connected “Things”

IOT Gateways

MQTT Telemetry Data

Management

IOT Integration Hub

App Integration

Enterprise Applications

JDBC

App Data

Telemetry Data

Management

Machine Learning Model

Cloudera’s Distribution including Hadoop (CDH)

Spark

Kudu

kafka

ActiveMQ

Apache Camel

Kapua

Openscoring.io

Apache Camel

OpenShift Origin

Future integration with Drools

EXAMPLE USE CASE: INDUSTRY 4.0 DEMO ARCHITECTURE: PROJECTS

Connected “Things”

IOT Gateways

MQTT Telemetry Data

Management

IOT Integration Hub

App Integration

Enterprise Applications

JDBC

App Data

Telemetry Data

Management

Machine Learning Model

Cloudera’s Distribution including Hadoop (CDH)

Spark

Kudu

kafka

ActiveMQ

Apache Camel

Kapua

Openscoring.io

Apache Camel

OpenShift Origin

Future integration with Drools
**Device connectivity**
Open standards – MQTT, AMQP, OPC-UA, CoAP, HTTP(s)

**Data management & analytics**
Based on Apache open source ecosystem libraries for machine learning & advanced analytics

**No vendor lock-in**
No rigid architectures or proprietary formats & components

**Flexible deployment**
Any of the leading cloud providers or your data center or hybrid cloud

**Open application interfaces**
Enterprise visibility | real-time anomaly detection | future-proof

**Community innovation**
Collaboration driven by some of the leading enterprises in the IoT space
VALUE PROPOSITION

**Open and interoperable**
Future-proof open source architecture | open standards | deployment flexibility

**End-to-End Analytics**
Analytics at the edge | advanced analytics & machine learning | ML model execution at the edge

**Control your data**
Privacy | security | regulatory

**Modular**
Avoid lock-in | capitalize on existing investments

**Reduce risk and complexity**
Simplify development, deployment, and integration tasks | save costs

**End-to-end security**
Security across devices, access, authentication and applications as well as data in motion and at rest
ENTERPRISE GRADE PRODUCTS FOR END-TO-END ARCHITECTURE

- Sensors, Actuators, Data Sources
- Edge Processing & Analytics
- Data Integration, Routing, Device Command/Control
- Advanced Analytics & Machine Learning
- Application Development, Deployment, Integration

CONNECTED “THINGS”

IoT GATEWAYS

IoT INTEGRATION HUB

Telemetry Data

APPLICATIONS

Enterprise Data Hub

Telemetry Data

Application Data
TO LEARN MORE

Download the Demo

Cloudera Booth # 225
THANK YOU
A NEW BEGINNING