How Komatsu is Improving Mining Efficiencies using IoT and Machine Learning

September 2018
Shawn Terry
Analytics Solutions Architect
@globalcoder1

www.mining.komatsu
@KomatsuMining
AGENDA

• An Introduction to Komatsu
• What Motivates Us
• The Komatsu Mining Data Analytics Journey
• Architecture
• Lessons Learned
• Questions?
COMPANY OVERVIEW

Smart Solution Centers
• Around the clock monitoring
• Custom Data Solutions
• Productivity analysis
• Training
• And more…

Provider of:
Equipment
Technology
Service solutions

Industries served:
Mining
Construction
Industrial
Forestry

Headquarters:
Tokyo, Japan

200 distributors
serving customers in
160 countries
BIG DATA for BIG MACHINES

4800XPC – Electric Rope Shovel

- 3.65 million pounds
- Nearly 75 feet high
- Move 270,000lbs of every 33s
- More than 1250 sensors reporting at frequencies up to 10Hz
- 50,000 unique measurement points every minute
CUSTOMER FOCUS

Optimizing mineral extraction for the least impact

- **Safety**
  - Autonomy, increased reliability

- **Environment**
  - Less waste, improved efficiency
  - Minimal footprint

- **Productivity**
  - Continuous improvement
  - Bringing lean practices to mining

- **Maintenance**
  - Less wear and tear on the machines
  - Reducing unplanned downtime
  - Reducing maintenance time
MINING MARKET CHALLENGES

Mines are driven by cost per ton

Today’s challenges:
- Increasing social and regulatory issues
- Difficult mining conditions – deeper ore deposits
- Declining commodity prices
- Inventory management
- Aging workforce
MOVING TO A HOLISTIC SOLUTION

From machine focus to **Systems Focus** to help our customers solve their toughest challenges:

- Smart equipment with built-in sensors
- Automated Monitoring of machine health
- Proactive Recommendations for Maintenance and Operations
- Understanding dependencies between machines and optimizing the entire process
- Delivering the right information to the right people at the right time to make decisions that lower $/ton.
MANAGING THE DATA

This machine:
1,250 sensor points
550 alarms & events

Each day

All machines:
17 billion time series
data points with 3xPeak
6 million Alarms & events
150k analytics executed
ANALYTICS PLATFORM OVERVIEW
OUR ANALYTICS JOURNEY

2008-2014

Gen 1
Historian Based
- Limited analytical capabilities
- Data processing and storage costs
- Limited machine learning capabilities
- Data silos

2014-2015

Gen 2
Out-of-the-Box Proprietary Solution
- Limited flexibility/interoperability
- Scaling to data growth was a challenge
- Data processing and storage costs
- Challenges around custom use case

2016-Now

Gen 3
Cloud based Open Architecture & Platform
- Cloudera on Microsoft Azure
- Processing, analytics and machine learning for IoT data
- Processing 200,000 data points/second
- Rapid prototyping

Future State

Gen 4
Edge Analytics and Improved Orchestration

Future Opportunities
- Edge Analytics
- Optimize Data Handling
- Improved Dev/Ops
- Leverage Cloud Services (separate compute/storage)
- Scale for Broader Growth

Increasing value and business impacts
CORE FUNCTIONS OF PLATFORM

INGEST

STORE

CALCULATE

VISUALIZE

REPORT
ORIGINAL PLATFORM (2008)

- First connected machine
- Near Real-Time Trending and Alerts in a browser
- New visibility into equipment health and operations
- New insights and actions never before possible
- Everything fresh and cool
REACHED THE LIMIT -- (~2013)

- Max sustained ingest ~60,000 points per second per environment
- ~150 pieces of connected equipment
- Complex and costly to scale out storage
- Operationally a nightmare to support
- Performance Sucked
- Unable to bring on new workloads / limiting the business
THE TEAM OBJECTIVES

• Get out of the old platform / cloud
• Ensure new platform can grow with business
• No new budget for first year
• Fixed deadline (Feb1 - 2017)
• Same core team (platform team: 5 + 3 offshore)
• Lift and shift what we can
• Address the new feature wish list
DATA DEMOCRATIZATION

Goal

• Let regional teams interact with the data
• Let regional teams develop analytics that can run quickly, reliably, and responsibly
• Let regional teams create reports and dashboards specific for their customers
• Share knowledge and best practices globally

Challenges

• Multi-tenant security
• Analytics development and testing processes and change management (BYOA)
• Isolating analytics development from backend changes
• Maintaining consistent SLAs
A NEW BEGINNING -- (2016)

- One unified cloud in Azure with Cloudera at the core.
- Microservices protect from over coupling
- Easily scalable storage but not all bottlenecks removed
- From “Go” to Production in 6 months
- Operations – Still challenging, but sustainable
- Happy Users = Rapid Growth and Adoption
- Still need to rethink our Analytics patterns
ANALYTICS PLATFORM STACK

VISUALIZATION & REPORTING

DATA INGEST / PIPELINES

ANALYTICS

DATA STORAGE AND MANAGEMENT
Enabling Next Gen Analytics -- (2018+)

- Further Decoupling and Removing Bottlenecks
- Stream Based Analytics
- Containerization
- Shifting Analytics to Edge
- Simplifying Architecture
- Improving DEV/OPS
- Improving DR and HA
CHALLENGES AT THE EDGE

• Getting data out of the mine site (customer friendly/extreme environments)
• Data Quality / Reliability
• Monitoring edge communications
  • Is it down or off / planned or unplanned?
• Managing time and time zones
• Every machine is unique and lots more to enable
• Explosive data volumes (Catch Up / Bad Sensors)
• Accurate and timely responses – Right Data to Right Person at Right Time
• Maintaining the Edge Footprint
Lessons Learned

Do

• Start with a clear problem to solve
  • Limit scope – get quick wins
• Expect to experiment, fail, and iterate quickly
• Engage experts
• Decouple as much as possible
  • Expect every piece to be replaceable
• Ensure your sponsors understand iterative/agile development processes
• Learn to automate dev/ops
• Instrument and Measure Performance
Lessons Learned

Don’t

• Expect cloud to solve all of your problems
  • Though it does creates amazing new possibilities

• Expect Cloud infrastructure to be “Always On”
  • Things will fail in unexpected ways – expect and plan for it.

• Expect technology selection to be a one and done decision
  • There will be something better tomorrow.

• Be afraid or unable to change course if something doesn’t work

• Be afraid of open source

• Underestimate the impacts of parallelization
  • Horizontal scaling creates different bottlenecks – you are only as strong as your weakest link.
Even Whole Data Centers Can Fail…

Microsoft Blames "Severe Weather" for Azure Cloud Outage

Azure outage: Microsoft working to restore key services after US regional disruption
COMPLEXITY

IT’S NOT A PROBLEM
If no one can see it
DATA VALUE

**Analytic solutions**

**What has happened**

<table>
<thead>
<tr>
<th>Summary</th>
<th>%</th>
<th>Total</th>
<th>7-Day Average</th>
<th>%Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Start Time</td>
<td>0%</td>
<td>100</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total Cycles</td>
<td>0%</td>
<td>100</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total Watts</td>
<td>0%</td>
<td>100</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Non-Productive Time</td>
<td>0%</td>
<td>100</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cycle Time</td>
<td>0%</td>
<td>100</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Idle</td>
<td>0%</td>
<td>100</td>
<td></td>
<td></td>
</tr>
<tr>
<td>No-Cycle Time</td>
<td>0%</td>
<td>100</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Average Cycle Stats</td>
<td>%</td>
<td>Average</td>
<td>Best Dev</td>
<td>%Difference</td>
</tr>
</tbody>
</table>
Summary

Komatsu is leveraging the power of IOT, big data and the cloud to help our mining customers get the lowest $/ton in safe, sustainable ways.

Thanks to our incredible team, our partners in Cloudera, Microsoft, and others, we are transforming the way we do business, creating new opportunities to serve our customers and now have the road open for future growth.

Through democratization of data analytics and BI we are enabling our Smart Services teams to partner with our customers to solve their toughest challenges.
QUESTIONS?