Streaming Analytics, Data Lakes and PI Integrators

Presented by
Matt Ziegler
Aaron Loe
Conference Theme and Keywords

Digital Transformation


Operational Intelligence, Asset Health, Sensor-based, Data IoT, Connectivity

Community, Quality, Integrators, Process, Scalability

Infrastructure, Reliability, Enterprise Agreement, Safety

High Speed, Business Impact, Operational Efficiency, CBM

Open System, Community, Partner

Enterprise, Connected Services, Big Data

PI System, Millions of Streams, Big Data, Future Data

Visualization
A journey through history

Machine Learning and IoT
Value in the data, Value in the expertise

**Constants**
Data  Expertise  Hard Work
Infrastructure  Self-Service Tools

**Trends**
Machine Learning  Big Data  IIoT

**Goals**
- Reduce scheduled repairs by up to 12%
- Reduce downtime up to 50%
- Reduce maintenance costs by up to 25%
- Reduce capital investment by 3-5%
- Eliminate 70% of breakdowns
- Cut unplanned outages by up to 50%
- Reduce unplanned outages by up to 50%
PI Integrators augment the existing set of PI capabilities like visualization, data access, and analytics with capabilities that make it easy to interact with non-OSIsoft tool sets.
Agenda

1. Recap on PI Integrators
   a. Capabilities
   b. Roadmap
   c. Success Criteria
2. Streaming
   a. Use Cases
   b. Architectures
3. Demo
4. Summary
5. Q&A
Progression

2015
Visibility / Slice and Dice / Model Training Data

2017
Prediction / Run Models / Stream Data

2019
Standardize / Enhance Models
Operations / Application Integration

Get Started

Derive More Value
<table>
<thead>
<tr>
<th>Business Intelligence &amp; Data Warehouses</th>
<th>2015-2016</th>
<th>2017</th>
<th>Future</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Available Today</strong></td>
<td>PI Integrator for Business Analytics</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>- Microsoft SQL Server, Oracle</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>- Hadoop (HDFS/HIVE)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>PI Integrator for SAP HANA</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

| **Available**                          | Cloud Platforms                      |       |        |
|                                        | - Microsoft Azure                     |       |        |
|                                        | - HANA Cloud Platform (5/2017)        |       |        |

| **Planned (1H 2017)**                  | Stream Systems                        |       |        |
|                                        | - Azure Event Hubs, IoT Hub           |       |        |
|                                        | - Apache Kafka                         |       |        |
|                                        | - SAP SDS                              |       |        |

| **Planned (1H 2017)**                  | More Platforms                         |       |        |
|                                        | - ESRI ArcGIS GeoAnalytics             |       |        |

| **Real-Time GIS**                      | Planned (1H 2017)**                    |       |        |
|                                        | PI Integrator for Esri ArcGIS          |       |        |
|                                        | - Situational Awareness                |       |        |
|                                        | - Real-Time Geoprocessing              |       |        |
|                                        | - Import ESRI features (assets)        |       |        |
| **Stream Systems**                     | Planned (1H 2017)**                    |       |        |
|                                        | Stream Systems                         |       |        |
|                                        | - Azure Event Hubs, IoT Hub            |       |        |
|                                        | - Apache Kafka                          |       |        |
|                                        | - SAP SDS                              |       |        |

| **Planned (1H 2017)**                  | Process Scale out                      |       |        |
|                                        | SSL / HTTPS                            |       |        |

| **Research**                           | All Integrators on common Framework (ESRI) |       |        |
|                                        | Node Scale Out and HA                  |       |        |

| **New Integration Patterns**           | Enable business process orchestration with PI System data – workflow, asset sync, transaction-like data, MES |       |        |

| **Research**                           | Enable partners and customers to build applications and interact programmatically using PI Integrator Framework. |       |        |
Data Scientist is the sexiest job of 21st century, but…

What's the least enjoyable part of data science?

- Building training sets: 10%
- Cleaning and organizing data: 57%
- Collecting data sets: 21%
- Mining data for patterns: 3%
- Refining algorithms: 4%
- Other: 5%
- Being Sexy: 0%

Time Series Data is Complex!

Turbine 1
- Speed
- Bearing Temp
- Oil Temp

Turbine 2
- Speed
- Bearing Temp
- Oil Temp
- Wear Factor
What do we need to approach this problem?

Operations – Up Front Work

- Collect: Collecting high-fidelity high-frequency data from a variety of sources and systems
- Enhance: Wrapping a layer of context around the data
- Calculate: Enriching the raw data by calculating KPIs, aggregations and different analysis

Debate

Process and Data Experts

GAP

1100010100 … Data Wrangling … 1001001110

Tools/System Experts

- Correlate: Finding patterns and relationships in data sets that aren’t revealed in one data set
- Apply Algorithms: Using statistics and machine learning to find insights in data across multiple variables
Soft Measurements – Ideal Use Cases for Streaming

• Predicting Soft Measurements
  – Deschutes Brewery – Fermentation Transition
  – Quality

• Enabling Data Lakes
  – Customize how data is stored & processed
Transition occurs between infrequent manual measurements

Constraints
- One manual density measurement per vessel every 8-10 hours
- Large capital expenditure not an option

Impact
- Up to 72 hours lost in production

Options
- $750k for inline density meters
- Manually predict transition in spreadsheets

Can this be predicted?

Soft Measurement
Automate and Optimize

**Lab Measurements**
- Viable Cell Count (VCC)
- Sugar Concentration
- API Concentration
- ...

**Online Measurements**
- Agitation Speed
- pH
- **Spectroscopy**
- ...

**Goals:** Optimize Cell Production, Eliminate Coming to the Lab

**Technique:** Use indirect measurements and machine learning to predict outcomes of lab measurements
Exercise Caution – Think PI System First

• Condition Based Maintenance / Detecting Failures
  – PI is already really good at this

• Predicting Failures
  – Contextualized Data Required
  – Not Enough Failures for predictions
  – Better suited for batch analytics
PI Integrator for Business Analytics View Types

- Asset View - Tables
- Event View - Tables
- Streaming View (In/Out) – Streams
Tables vs Streams

- Business Intelligence
- Human readable
- Batch / Bulk Process
- Normalized data
- Regularly scheduled
- Large data, few messages

- In-line (Streaming) Analytics
- Computer readable
- Specific Data / Targeted Process
- Raw or “Packages” of data
- Triggered
- Small data, many messages
Integration Patterns

Tables

Streams

Other Patterns

Files

Databases

Files

Queues

Messaging

External Analytics Engines

Metadata

Programming

On-Demand

Workflow & Transactions
Current Architecture with Streaming Analytics

A lot of stuff
Deschutes Brewery Operationalizing Architecture

DATA SOURCES ➔ PREPARE ➔ INGEST ➔ ANALYZE ➔ PUBLISH ➔ CONSUME

PI System ➔ PI Integrator for Microsoft Azure ➔ SQL Data Warehouse ➔ Machine Learning ➔ SQL Data Warehouse ➔ Cortana

Predictions as Future Data (to PI 2015)

Azure Data Factory (Orchestration)

On Premises ➔ On Cloud

SQL Data Warehouse ➔ Power BI ➔ Web/LOB Dashboards
2017 Simplified Architecture and Integrated Workflow

Focus on your computations with fully integrated services for Cortana and SAP HANA

DATA SOURCES ➤ PREPARE ➤ ANALYZE

- PI System
- PI Integrator for Microsoft Azure
- Machine Learning
- Web Services
- Predictions as Future Data

PI Integrator for Business Analytics
PI Integrator for SAP HANA
PI Integrator for Microsoft Azure
San Leandro Tech Campus View Glass System

View Glass Tint Changes with Outdoor Brightness

Finding Outliers – Health Score
- Transition Time Relative to Peers
- Drift in Transition Time
<table>
<thead>
<tr>
<th>Name</th>
<th>Run Status</th>
<th>Type</th>
<th>Run Mode</th>
<th>Start Time</th>
<th>End Time</th>
<th>Last Run Time</th>
</tr>
</thead>
</table>

No Rows To Show
<table>
<thead>
<tr>
<th>Name</th>
<th>Run Status</th>
<th>Type</th>
<th>Run Mode</th>
<th>Start Time</th>
<th>End Time</th>
<th>Last Run Time</th>
</tr>
</thead>
<tbody>
<tr>
<td>View Glass Tinting</td>
<td>Streaming</td>
<td>StreamingOut</td>
<td>Stream</td>
<td>8h</td>
<td>*</td>
<td>12/31/99 3:59 PM</td>
</tr>
</tbody>
</table>

### Run Status
- **View Name**: View Glass Tinting
- **PI AF Database**: Facilities - 9000 Alvarez
- **Publish Target**: Kafka
- **View Type**: StreamingOut
- **Run Mode**: Stream
- **Start Time**: 12/31/99 3:59 PM
- **End Time**: *
- **Last Run Time**: 12/31/99 3:59 PM
- **Sample Frequency**: 1 minute
Wrap Up

• Constants – Data Infrastructure, Expertise, Self-Service Tools

• Soft Measurements are Ideal for Streaming

• PI Integrator for Business Analytics 2017
  – Data in **Context** for streaming analytics
Have an idea how to improve our products?

OSIsoft wants to hear from you!

https://feedback.osisoft.com/
Contact Information

Matt Ziegler
mziegler@osisoft.com
Product Manager
OSIsoft, LLC

Aaron Loe
aloe@osisoft.com
Sr. Software Developer
OSIsoft, LLC
Questions

Please wait for the **microphone** before asking your questions

State your name & company

Please remember to...

Complete the Online Survey for this session

Download the Conference App for OSIsoft Users Conference 2017

- View the latest agenda and create your own
- Meet and connect with other attendees

Thank You

감사합니다

谢谢

Merci

Danke

Gracias

ありがとう

Спасибо

Obrigado