Rapid Development of Open Real-Time Drilling Analytics System

Dingzhou Cao, Chad Loesel, Sanjay Paranji

10/09/2017
Outline

- Motivations
- Streaming Analytics Development Tools
  - Open Source or Vendor Product
- Open Real-Time Drilling (RTD) Analytics System
  - Analytics Modules
- Conclusions and Challenges
Motivations

- Advanced Analytics is the trend in oil and gas
  - Machine learning / deep learning is the trend
  - RTD has already been proven to be a vital tool
  - RTD + Advanced Analytics to gain significant competitive advantages

- Quick delivery RTD to fill gaps in operation.
  - Drilling engineering team has immediate need for the RTD system
  - Quick delivery what they need to get their buy-in

- Proof of Concept (POC) in a fast track
Real Time Streaming Architecture

1. Data source
2. Data Collection
3. Messaging System
4. Real-time Processing
5. In-memory Storage/caching
6. Live UI
### Streaming Analytics Development Tools (no complete list!)

<table>
<thead>
<tr>
<th>Product</th>
<th>Framework</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Open source</strong></td>
<td><strong>Closed source</strong></td>
</tr>
<tr>
<td>Apache Flink</td>
<td>Event Hubs</td>
</tr>
<tr>
<td>STORM</td>
<td>Cloud Pub/Sub</td>
</tr>
<tr>
<td>HERON</td>
<td>Amazon Kinesis</td>
</tr>
<tr>
<td>Apache Flink</td>
<td>EsperTech</td>
</tr>
<tr>
<td>Spark Streaming</td>
<td></td>
</tr>
<tr>
<td>samza</td>
<td></td>
</tr>
<tr>
<td>kafka</td>
<td></td>
</tr>
</tbody>
</table>
Real Time Streaming Architecture cont.

configure multiple tools to make the pipeline/streaming process works

Data source → Data Collection → Messaging System → Real-time Processing → In-memory Storage/caching → Live UI

Open Source

WITSML Connector

Stream Base

TIBCO® StreamBase®

Single Tool (complete processing pipeline)
Stream Analytics Development Tools

**Open Source Frameworks**
(Kafka, Spark Streaming / Apache Flink / Apache Storm et al.)
- Fault-Tolerance and scalability – running on nodes/clusters
- Easy implementation / deployment of machine learning / deep learning algorithms
  - Organic integration
- Lots of developer communities
- Too many tools available
  - More than 20 open source streaming tools available,
  - New tools emerges and evolves in a fast pace
  - Need to configure multiple tools to make the pipeline/streaming process works

**Vender Product**
(StreamBase)
- Visual IDE
- Single Tool (complete processing pipeline)
- Single node server, fault-tolerance and scalability issue
- Issue of implementation / deployment of machine learning / deep learning algorithms
  - Rely on TERR, or adaptors
- Lack of developer communities (almost none)
Streaming Analytics Development Tools cont.

Comparison

<table>
<thead>
<tr>
<th>TECHNOLOGY</th>
<th>UP-FRONT MATERIALS</th>
<th>UP-FRONT PERSONNEL</th>
<th>LONG-TERM MATERIALS</th>
<th>LONG-TERM PERSONNEL</th>
<th>TIME-TO-VALUE</th>
</tr>
</thead>
<tbody>
<tr>
<td>StreamBase</td>
<td>High</td>
<td>Low</td>
<td>Low</td>
<td>Low</td>
<td>Short (3+ months)</td>
</tr>
<tr>
<td>Open Source</td>
<td>Low</td>
<td>High</td>
<td>Low</td>
<td>Medium-high</td>
<td>Long (12+ months)</td>
</tr>
</tbody>
</table>

- **Up-front Materials** – Any initial costs associated with software licensing or subscription, along with project hardware and other resources
- **Up-front Personnel** – Initial team required for the development cycle, expressed over time
- **Long-term Materials** – Ongoing costs associated with software subscription or maintenance
- **Long-term Personnel** – Ongoing team required to enhance and support the project
- **Time-to-Value** – The time from the initiation of the project to the delivery of functional business value

Data source: TIBCO StreamBase and the TIBCO Accelerator for Apache Spark

**After POC, we explore the alternative way of RTD with open source tools**
- Kafka + Spark Streaming + OpenTSDB
Open Real-Time Drilling (RTD) Analytics System

Architecture

Data layer

Analytics layer

UI layer

Static data

Streaming data

RTD Analytics Modules
- Activity Recognition
- SDGS
- RDGS
- DKPI
- T&D model
- Hydraulic model
- Optimal Wellbore Trajectory Control

Service Company

WITSML Data Store

WITSML Connector

MS SQL Server

APC

Tibco Live Datamart

Live UI in web browser

Storage

EDMPL Oracle database

TIBCO StreamBase

MAPR
OPEN means…

- Data layer is interchangeable
  - Current with WISTML, we get 1 second data with latency ~10-15 seconds
  - We are currently seeking alternative method other than WISTML to reduce the latency to ~2-3 seconds for 1 second data

- Analytics layer
  - Easy for Data Scientist to build model (machine learning based model or physics based model), plug in and run
  - APC AAET has a strong advanced analytics team

- UI layer
  - UI can be customized to according to analytics modules
RTD Analytics System - modules

- Four modules are in production
  - Rule-based Drilling Activity Recognition
  - Sliding Drilling Guidance System (SDGS)
  - Drilling Key Performance Indicator (DKPI)
  - Torque and Drag (T&D)

- One is ready for deployment
  - Rotational Drilling Guidance System (RDGS)

- Two are under development
  - Hydraulic Model
  - Wellbore Trajectory Correction Control

- Majority of them are physics based model
  - Business unit has an immediate need these physics based model
  - More machine learning based model will be developed and added to system
Analytics Module - Drilling Activity Recognition

- Real-time categorize the drilling activity every second (trip-in, trip-out, rotation drilling, sliding drilling et al.)
  - Enable the real-time monitor of drilling operations
  - Foundation for the remaining analytics modules
Automated directional calculations including motor yield and build rate needed to land as surveys are posted.
Analytics Module – T&D

- Provides real-time information on downhole friction and hole issues as well as automatic monitoring of casing running operations.

  - It helps make necessary adjustments to reduce the chance of stuck pipe event and damaged equipment and ensure casing integrity.
Analytics Module – RDGS

Learning the best drilling patterns from analog wells using machine-learning algorithm, generate the best drilling roadmap (best composite well) to guide the drilling of a new well.
Analytics Module – Wellbore Trajectory Correction

- Real time monitor the actual well path; propose a well path back to plan well path once the actual well is deviated from plan well path
- Real-time Determination of the “Best Yet-to-Drill Path”
Conclusions and Challenges

Conclusions

- Within in three months
  - Build the RTD Analytics System (POC) from scratch within a small team
  - Deliver what drilling engineering team urgent need and get their buy-in
- RTD Analytics System is growing as more modules added

Challenges

- As a real time system, getting all data models need in the right time/manner is very demanding
- Smooth UI/UX is very important
- Seeking alternative streaming data connection other than WITSML to reduce latency
- Seeking alternative open source streaming development tools to make the system more OPEN
  - Easy implementation / deployment of machine learning / deep learning
Thank you!