

Unravel Newsletter: June 2019

Greetings from Unravel,

This month's newsletter focusses on Streaming Data Applications. We'll talk about the need for Application Performance Management (APM) for such applications and how Unravel can help. We'll continue with community highlights on this topic and include some links to recent conference recordings and upcoming events that you may find interesting.

[This short 2 min video](#) shows how Downer Group in Australia is transforming the train industry by harnessing the power of Big Data. They get 30,000 signals from each train every 10 minutes (from IOT sensors and cameras) that tell them about the health of the train. A cloud-based Stream Analytics platform then helps automate inspections, optimize operations and introduce predictive maintenance, saving time and money. Soon, Sydney Trains' drivers and guards will be alerted to technical problems in less than a second. This is their TrainDNA solution!

The above is just one example of Streaming Data Applications. These days, Streaming Data Analytics is prevalent across many industries and enables a wide variety of use cases. These applications process continuously flowing data generated from sources like sensors, web activity, monitoring data, incoming logs, Internet of Things (IoT) applications etc. and their use cases range from real-time online recommendations to instant fraud detection to real-time security intelligence operations to ad-auction platforms powered by real-time data and many more.

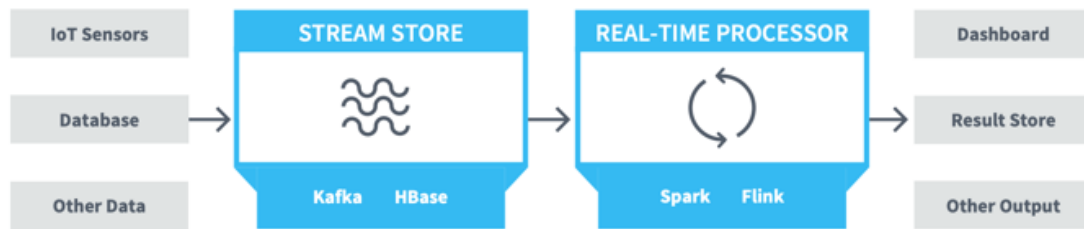
Spark Streaming has become a popular platform to implement such applications. Also, while streaming data can come from many different sources e.g. Kafka, Flume, HDFS, Databases etc., Kafka has risen as one of the more popular ones.

[Unravel Performance Management for Streaming Data Applications](#)

Anatomy of a Streaming Application

Several distributed components together make up a streaming data application. There can be various sources of this data that make into a

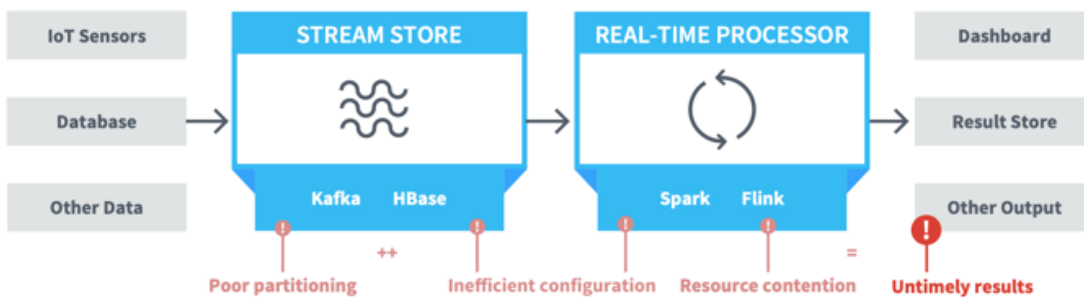
Stream store like Kafka, Flume, etc. and can be processed using Spark Streaming for example. The processed results are then displayed on live dashboards, saved to filesystems, databases etc.



Challenges

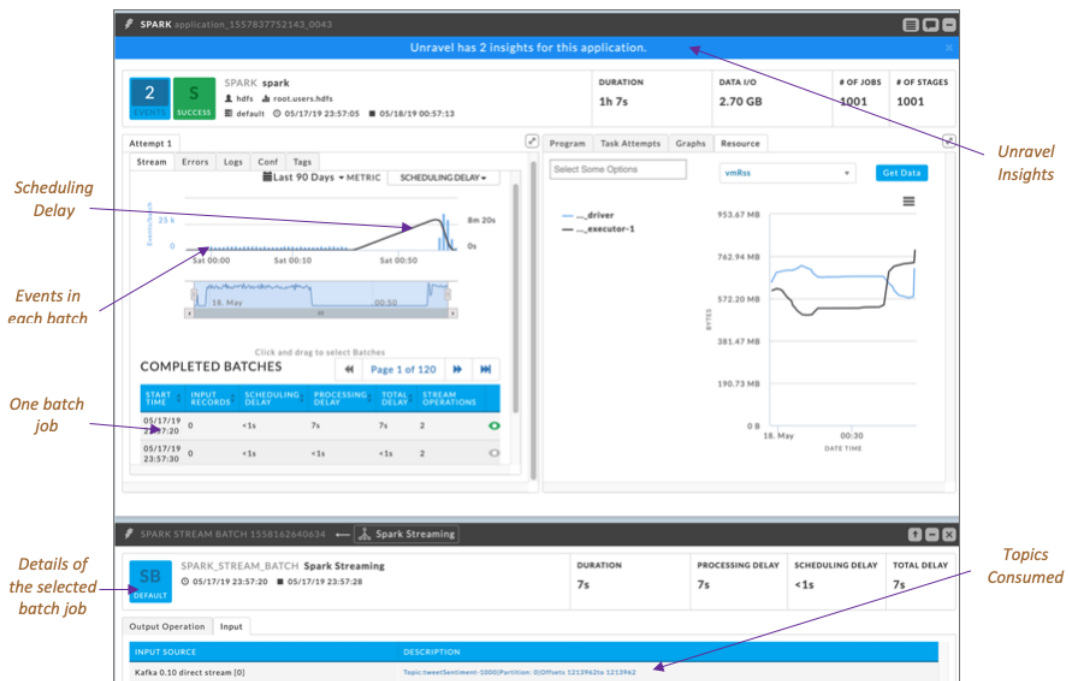
Troubleshooting problems in such distributed applications is not easy. In the absence of a single pane of glass to monitor and manage all the components together (e.g., Kafka, Spark, etc.), it is almost impossible to track down causes of failures and slowdowns in a reasonable amount of time.

For example, consider the case of untimely results - how does one figure out whether this is caused by a resource contention issue, or a configuration issue, or poor partitioning and what is the best way to fix and prevent these issues across distributed systems like Kafka, HBase, Spark, Flink, etc.? Essentially, APM for Streaming Data Applications is a must.



How Unravel helps

Unravel enables monitoring and managing of streaming data applications end-to-end, meaning across all the components of the stack such as Kafka, HBase, Spark and others that make up such an application. To start with, Unravel brings all performance data into one complete and correlated view and then in addition, provides out of the box intelligence to derive useful insights and recommendations. Below is a sample screenshot of how Unravel shows you all the relevant information about your Spark Streaming Application.



Spark Streaming Application in Unravel

[Sandip Honde](#), Senior Staff Engineer at Unravel and [Hari Nair](#), Senior Software Engineer at Unravel recently published a blog illustrating how to leverage Unravel’s capabilities to debug issues with a Spark Streaming application that consumes data from a Kafka topic: [Monitoring Data Pipelines for Spark Streaming in a Kafka Cluster](#).

Also check out the [Spark – Streaming](#) section in the Unravel User Guide to help you get started with monitoring Spark Streaming Applications running in your cluster.

New Product Documentation Portal

Notice a new look? We are excited to share the launch of our [new Product Documentation Portal](#)!

Conference Session Recordings

- Check out the videos and slides from the talk by Shivrath Babu, Unravel CTO at the Kafka Summit, London - [Using Machine Learning to Understand Kafka Runtime Behavior](#)
- Watch the talk from Eric Chu, Unravel VP of Data Insights at Spark AI Conference, San Francisco – [Use Machine Learning to Get the Most out of Your Big Data Clusters](#)

- Watch the talk from Shivnath Babu, Unravel CTO at Spark AI Conference, San Francisco - [An AI-Powered Chatbot to Simplify Apache Spark Performance Management](#)
-

Upcoming Events

- Join us at the [Big Data/AI Conference and Expo, Toronto](#), in Toronto, Canada for a session on June 12 at 2:45 pm where Muji Qadri, Unravel Solutions Engineer will present Taming BigData DevOps with Automation. Join us at the Unravel booth at the Expo as well.
 - Join us at the Unravel Booth at [AWS Summit, New York](#), July 11.
-

Community Highlights

Keeping with the Streaming Applications theme, here are

- Some well-regarded books to develop in-depth knowledge in Streaming Applications:
 - [Streaming Data: Understanding the real-time pipeline](#): Streaming Data introduces the concepts and requirements of streaming and real-time data systems. The book is an idea-rich tutorial that teaches you to think about how to efficiently interact with fast-flowing data.
 - [Streaming Systems: The What, Where, When, and How of Large-Scale Data Processing](#): This practical guide helps data engineers, data scientists, and developers learn how to work with streaming data in a conceptual and platform-agnostic way.
 - [Pro Spark Streaming: The Zen of Real-Time Analytics Using Apache Spark](#): This book will help you learn the right cutting-edge skills and knowledge to leverage Spark Streaming to implement a wide array of real-time, streaming applications. It walks you through end-to-end real-time application development using real-world applications, data, and code.
 - See how Netflix uses Kafka and Spark Streaming for various business use cases - [Spark and Spark Streaming at Netflix](#).
-

Resources

- [Learn more](#) about Unravel.
- [Online Product Demo](#)
- [Unravel Partners](#)
- [Unravel Product Releases and Documentation](#)
- [Unravel Datasheet](#)
- [More Unravel News](#)



[Contact Us](#). [Sign Up for 30-day Trial](#).

2019 Unravel. All Rights Reserved. 2 Palo Alto Sq, Suite 120, Palo Alto, CA 94306