

Unravel Newsletter: March 2019

Greetings from Unravel,

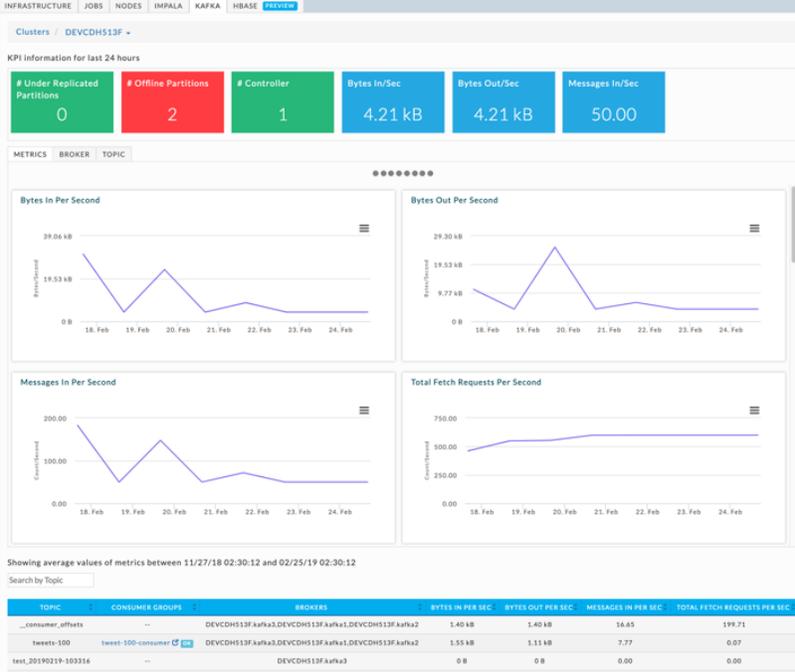
With [Pi Day](#) just around the corner, we at the Unravel Office have been talking about a Pie baking bash, hope you are planning something fun as well! Another exciting topic that's coming up a lot around here is Apache Kafka. A large number of companies are now using Kafka and running Spark Streaming Apps as part of their modern data architecture, Unravel's own design does too!

In this month's newsletter, we talk about **Unravel for Kafka**, followed by common multi-tenant cluster contention problems. We then continue the Kafka theme with community highlights on this topic. Also included are links to recent webinars and upcoming events that you might find interesting.

Unravel for Kafka

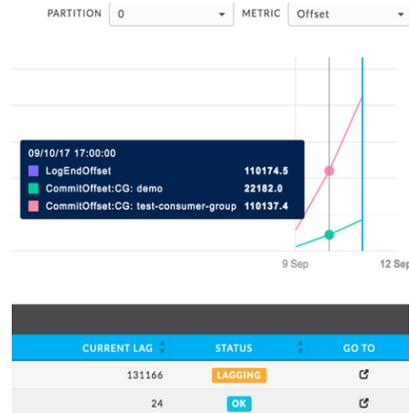
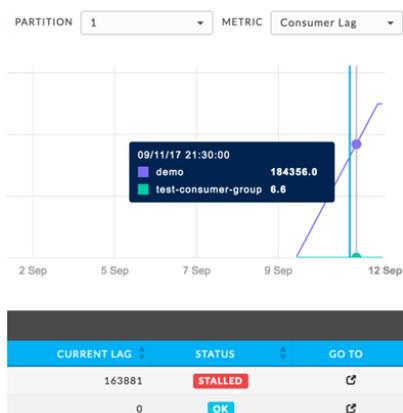
Unravel for Kafka enables DataOps teams to comprehensively monitor and manage performance, predictability, and reliability, of Kafka based streaming data applications end-to-end (i.e., across all the components such as Kafka, HBase, Spark, that make up an application).

To start with, the Unravel dashboard for Kafka shows key KPIs (e.g. total fetch requests, # of under replicated partitions, # of offline partitions etc.) for every Kafka cluster it tracks. You can drill-down further as needed into details of the metrics corresponding to "Topics", "Brokers", Consumers etc.



Unravel Kafka Dashboard (main)

The intelligence provided by Unravel can easily help identify bottlenecks in the Kafka cluster that affect the overall throughput and efficiency of a streaming application. E.g., Unravel provides insights into a Kafka cluster’s activities through the auto-detection of the status of relevant Consumer Groups (OK/Lagging/Stalled) and other KPIs.



Examples of “Stalled” and “Lagging” Consumer Groups (name = “demo”) showing in the Unravel UI

Find more details about this topic (no pun intended) in the [Unravel Solution Guide, Kafka](#). Also, hear our CTO, Shivnath Babu talk about [Using Machine Learning to Simplify Kafka Operations](#).

Don’t forget to check out the [Kafka Insights](#) and [Spark Application Manager – Spark Streaming](#) sections in the Unravel User Guide to help you get started!

Fighting for Precious Resources on a Multi-Tenant big Data Cluster



In his new blog, Unravel CEO Kunal Agarwal discusses some of the most common multi-tenant cluster contention problems seen in the field:

- Oversubscribed clusters – too many apps or jobs to run, just not enough resources
- Bad container sizing – too big or too small
- Poor queue management – sizes of queues are inappropriate
- Resource hogging users or apps – bad apples in the cluster

and how Unravel helps solve each of these. [Read the blog.](#)

New Website

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

Webinars

- Unravel CTO Shivnath Babu explains how taking an [#AI](#)-driven approach can simplify [#Spark](#) performance management challenges. Watch the full video from the [#ApacheSpark](#) meetup [here](#).
- Check out this Bloor Group and Eckerson Group webinar with Unravel - [How to Build Reliable Data Pipelines Using AI and DataOps to learn about the role of DataOps](#) in supporting modern data apps, how AI is needed to manage and monitor complex data pipelines and environments, and how modern perf. management software can reduce the risk of running complex data apps.
- Learn why the cloud is forcing an evolution of thinking and investment in big data programs: [Planets Align – The Irresistible](#)

[Forces Pulling Big Data to the Cloud](#). It also discusses the role of AI, automation, and market forces in the shift towards big data in the cloud.

- [Doing DevOps for Big Data? What you need to know for AIOps](#). Explore with us the promise AIOps has to create hyper-efficiency within DevOps teams as they struggle with the diversity, complexity, and rate of change across the entire stack.

Upcoming Events

- Join us at the Unravel Booth on the Expo floor at [DataWorks Summit, Barcelona](#), March 18-21.
- Join us at the Unravel Booth at [AWS Summit, Santa Clara](#), March 27.
- Join us at the [#StrataConf](#) in SF on March 28 at 11:50 am PT to learn how to easily root cause Big Data Stack App failures with an automated technique that uses deep learning techniques. Details [here](#).
- Join us at the [#SparkAISummit](#) in SF for a session on April 25 at 11:50 AM PT where our CTO, Shivnath Babu, will discuss how organizations can leverage [#AI](#) chatbots to enhance their [#Spark](#) performance mgmt. Details [here](#).

Community Highlights

Keeping with the Kafka theme for this time

- Some well-regarded books to develop in-depth knowledge in Kafka:
 - [Kafka: The Definitive Guide](#): Engineers from Confluent and LinkedIn, responsible for developing Kafka explain how to deploy production Kafka clusters, write reliable event-driven microservices, and build scalable stream-processing applications.
 - [Apache Kafka 1.0 Cookbook](#): For developers and admins who are looking for practical solutions to problems encountered while operating, managing or monitoring Kafka.
- An interesting tutorial featuring Kafka:
<https://tech.marksblogg.com/minimalist-guide-tutorial-flume.html>. This talks about feeding NGINX logs into Kafka and using Flume to write data from Kafka to HDFS.

Resources

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

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

© 2019 Unravel - All Rights Reserved

2 Palo Alto Square, Suite 120, Palo Alto CA 94306

[Privacy Policy](#) | [Manage your preferences](#) | [Unsubscribe](#)