

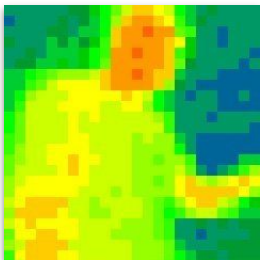


Real-Time Stock Processing With Apache NiFi, Apache Flink and Apache Kafka

Timothy Spann, Principal DataFlow Field Engineer @ **CLOUDERA**
Pierre Villard, Senior Product Manager @ **CLOUDERA**



Who?



Tim Spann

@PaasDev // Blog: www.datainmotion.dev

Principal DataFlow Field Engineer. Princeton Future of Data Meetup.

<https://github.com/tspannhw/EverythingApacheNiFi>

<https://community.cloudera.com/t5/Community-Articles/Real-Time-Stock-Processing-With-Apache-NiFi-and-Apache-Kafka/ta-p/249221>



Pierre Villard

Twitter & Github - @pvillard31 // Blog: www.pierrevillard.com

Committer and PMC member for Apache NiFi (in the community since 2015)

Senior Product Manager at Cloudera for products around Apache NiFi, NiFi Registry and MiNiFi

Previously at Google & Hortonworks



What?

This talk is about ingesting real-time data from many sources and build a dashboard on top it to track in real time what our stocks are.

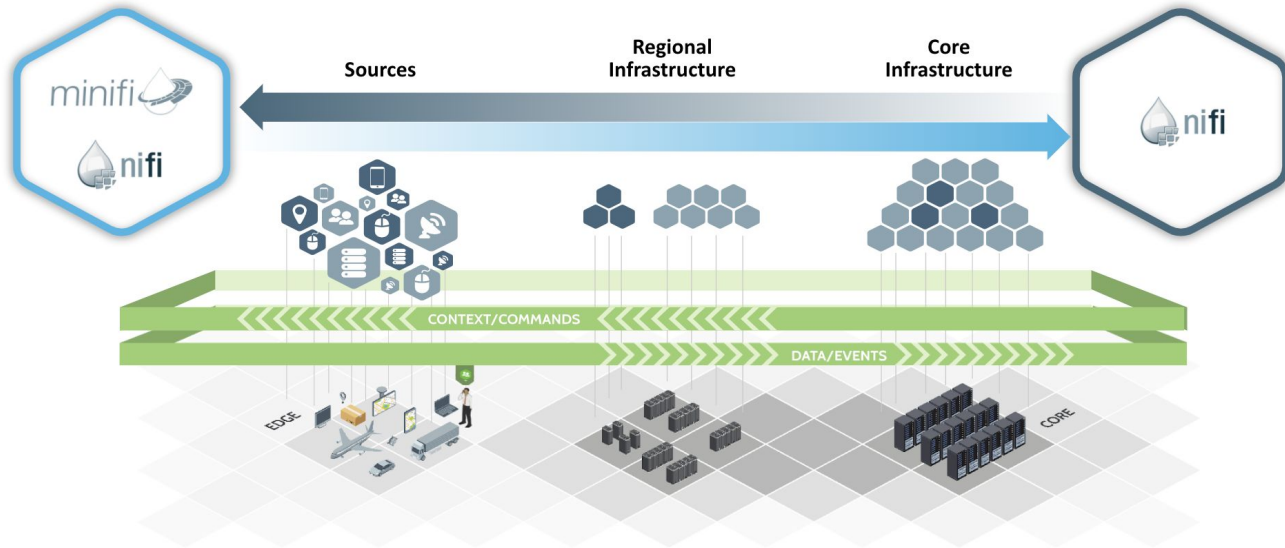
This use case is a good example to show the combination of some of the best Apache solutions for streaming applications.



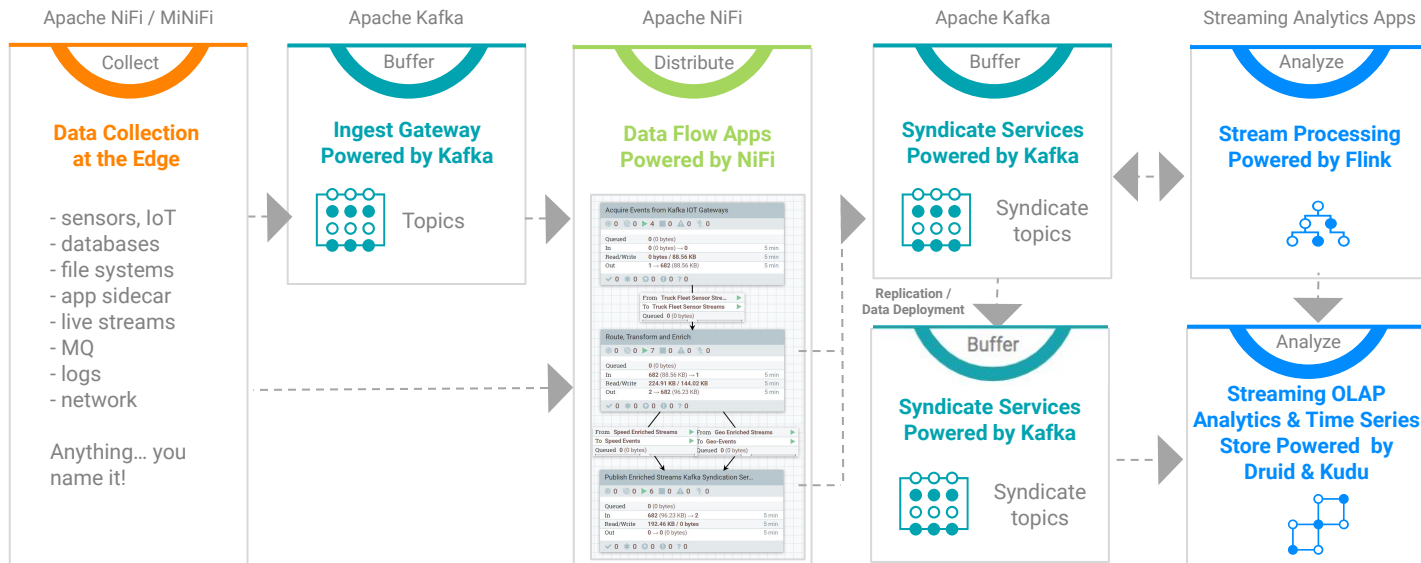
NiFi, Kafka and Flink in a few numbers

- Apache NiFi (version 1.12.x) - created and open sourced by the NSA - initial release in 2006
350+ contributors, 1200+ people in Slack, 3.1M+ docker pulls
Many sub-projects: NiFi, MiNiFi Java, MiNiFi C++, NiFi Registry, etc
- Apache Kafka (version 2.6.x) - created and open sourced by LinkedIn - initial release in 2011
700+ contributors
- Apache Flink (version 1.11.x) - initial release in 2011
750+ contributors, 2nd top repository by number of commits, top active project on mailing lists

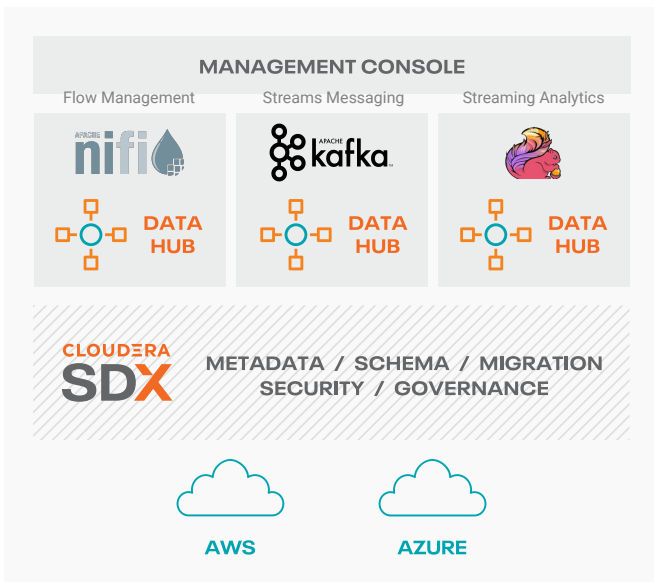
What is NiFi used for?



Streaming Reference Architecture



Where?



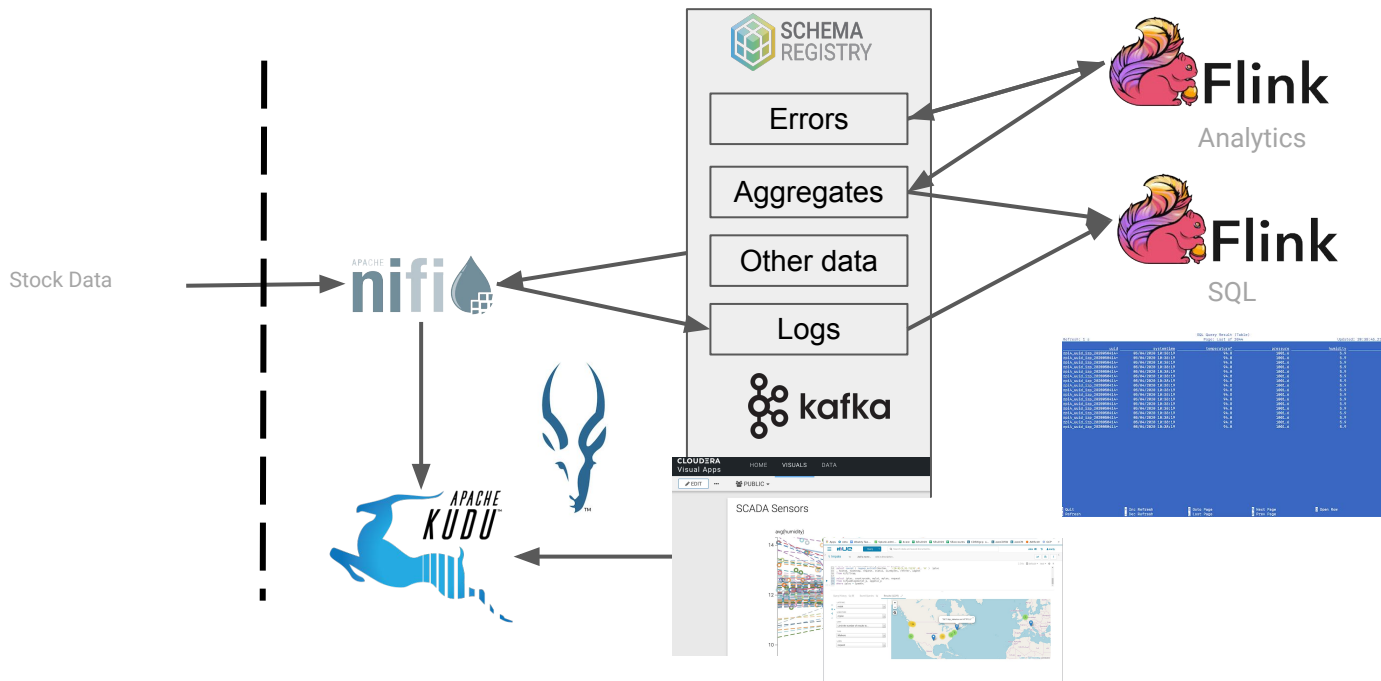
Web service hosted and managed by Cloudera

Hosted in the your cloud environment, but managed by the CDP Management Console

Shared Data Experience (SDX) technologies form a secure and governed data lake backed by object storage (S3, ADLS, GCS)

CDP services are optimized for the elastic compute & 'always-on' storage services provided by any cloud provider

How? This use case architecture





Demonstration

Let's see all of this in action...



Thanks! Questions?

Timothy Spann, Principal DataFlow Field Engineer @ **CLOUDERA**

Pierre Villard, Senior Product Manager @ **CLOUDERA**