ID de Contribution: **2**   Type: **Non spécifié**

Enhancing Real-Time Astronomical Alert Processing with Scalable Kubernetes-Driven Infrastructure and Observability

mercredi 16 juillet 2025 12:00 (30 minutes)

Fink is a real-time astronomical alert broker designed to process high-throughput data streams from surveys such as ZTF and the forthcoming LSST. To meet the demands of low-latency, high-reliability data classification at scale, we have developed a production-grade Kubernetes-based deployment architecture that integrates continuous delivery (CI/CD), fine-grained observability, and resource-aware scheduling.

This contribution presents recent work on operationalizing Fink's Spark-based alert pipeline using Kubernetes-native tooling, focusing on standardizing performance monitoring and pipeline introspection. We leverage Prometheus and JMX exporters for service-level metrics, integrate sparkmeasure to capture stage-level execution details, and develop automated profiling strategies to detect bottlenecks and regressions. Our goal is to elevate Spark observability to first-class status within real-time scientific data infrastructures.

We also report on collaborations with CERN (Luca Canali), Kubeflow, and Stackable.tech to prototype next-generation workflows for distributed data science, bridging operational excellence with scientific reproducibility. These efforts aim to position Fink as a model for scalable, maintainable, and transparent alert processing in the upcoming data-intensive era of time-domain astronomy.

Author: JAMMES, Fabrice (CNRS)

Co-auteur: Dr PELOTON, Julien (CNRS-IJCLab)

Orateur: JAMMES, Fabrice (CNRS)

Classification de Session: Technology