

Coffea Framework: Current Status, Recent Updates, and Community Impact

Monday, 7 September 2026 09:30 (15 minutes)

The Coffea (Columnar Object Framework for Effective Analysis) framework continues to evolve as a cornerstone tool for high-energy physics data analysis, providing physicists with efficient, scalable solutions for processing complex event data. This talk presents the current status of Coffea, highlighting significant recent developments and their impact on the HEP analysis community.

A major milestone has been transitioning from Dask-Awkward to Awkward Array's new virtual arrays feature as the default backend. This shift improves Coffea's user-friendliness, increases performance, and simplifies the execution model. The transition requires minimal to no code modifications for existing analyses, providing a seamless migration path with substantial improvements.

Recent enhancements include advanced workflow features such as checkpointing for robust, resumable analyses, improved branch preloading and caching for network-efficient data access, and workflow tracing to identify required data branches. These optimizations benefit both interactive and batch processing scenarios, allowing physicists to focus on physics rather than data management details.

This talk will also present usage and community feedback demonstrating Coffea's growing adoption across CMS and other HEP experiments. Looking toward the High-Luminosity LHC era, Coffea's architecture positions it as a key enabling technology for handling unprecedented data volumes while maintaining intuitive, user-friendly interfaces for data analysis at scale.

Primary author: KROMMYDAS, Iason (Rice University)

Presenter: KROMMYDAS, Iason (Rice University)

Session Classification: Talks