Contribution ID: **88**   Type: **Poster**

Awkward Arrays to RDataFrame and back

Tuesday, 25 October 2022 11:00 (30 minutes)

Awkward Arrays and RDataFrame provide two very different ways of performing calculations at scale. By adding the ability to zero-copy convert between them, users get the best of both. It gives users a better flexibility in mixing different packages and languages in their analysis.

In Awkward Array version 2, the ak.to_rdataframe function presents a view of an Awkward Array as an RDataFrame source. This view is generated on demand and the data is not copied. The column readers are generated based on the run-time type of the views. The readers are passed to a generated source derived from ROOT::RDF::RDataSource.

The ak.from_rdataframe function converts the selected columns as native Awkward Arrays.

We discuss the details of the implementation exploiting JIT techniques. We present examples of analysis of data stored in Awkward Arrays via a high-level interface of an RDataFrame.

We show a few examples of the column definition, applying user-defined filters written in C++, and plotting or extracting the columnar data as Awkward Arrays.

We discuss current limitations and future plans.

Significance

References

Experiment context, if any

CMS

Primary authors: OSBORNE, Ianna (Princeton University); PIVARSKI, Jim (Princeton University)

Presenter: OSBORNE, Ianna (Princeton University)

Session Classification: Poster session with coffee break

Track Classification: Track 1: Computing Technology for Physics Research