



Contribution ID: 19

Type: Oral

## Seamless transition from TTree to RNTuple analysis with RDataFrame

*Monday, March 11, 2024 5:30 PM (20 minutes)*

As the High-Luminosity LHC era is approaching, the work on the next-generation ROOT I/O subsystem, embodied by the RNTuple, is advancing fast with demonstrated implementations of the LHC experiments' data models and clear performance improvements over the TTree. Part of the RNTuple development is to guarantee no change in the RDataFrame analysis flow despite the change in the underlying data format.

In this talk, we present integration of RNTuple and RDataFrame. The engine can process RNTuple datasets on a local machine, sequentially with one core or using implicit multithreading with multiple cores. Furthermore, RNTuple processing is also introduced in the distributed RDataFrame layer and benchmarked using SWAN, a web-based platform, to transparently offload analysis tasks to the CERN HTCondor pools. The new workflow is demonstrated using existing RDataFrame analyses on one or multiple nodes with no change in the API. One notable example is the t-tbar Analysis Grand Challenge benchmark, which is also used as a blueprint to showcase differences in performance of (distributed) execution with the two data formats.

### Significance

LHC experiments are already involved in the process of testing and validating the next-generation ROOT I/O. ROOT will progressively fade out support for writing new datasets with TTree, so RNTuple will have a clear impact on future HEP computing workflows at many levels, from infrastructures to final analyses. This presentation demonstrates how the ROOT efforts go in the direction of making the transition as effortless as possible for the HEP users, while aligning with the experiments' expected computing challenges.

### References

CHEP 2023 <https://indico.jlab.org/event/459/contributions/11582/>

ACAT 2022 <https://indico.cern.ch/event/1106990/contributions/4998129/>

### Experiment context, if any

**Author:** CZURYLO, Marta (CERN)

**Co-authors:** FALKO, Andrii; PIPARO, Danilo (CERN); TEJEDOR SAAVEDRA, Enric (CERN); GUIRAUD, Enrico (Princeton University, CERN); BLOMER, Jakob (CERN); CANAL, Philippe (Fermi National Accelerator Lab. (US)); Dr PADULANO, Vincenzo Eduardo (CERN)

**Presenter:** CZURYLO, Marta (CERN)

**Session Classification:** Track 1: Computing Technology for Physics Research

**Track Classification:** Track 1: Computing Technology for Physics Research