



**CHEP`24, Krakow, Poland, 21-25 October**

## Impact of RNTuple on Storage Resources for ATLAS Production

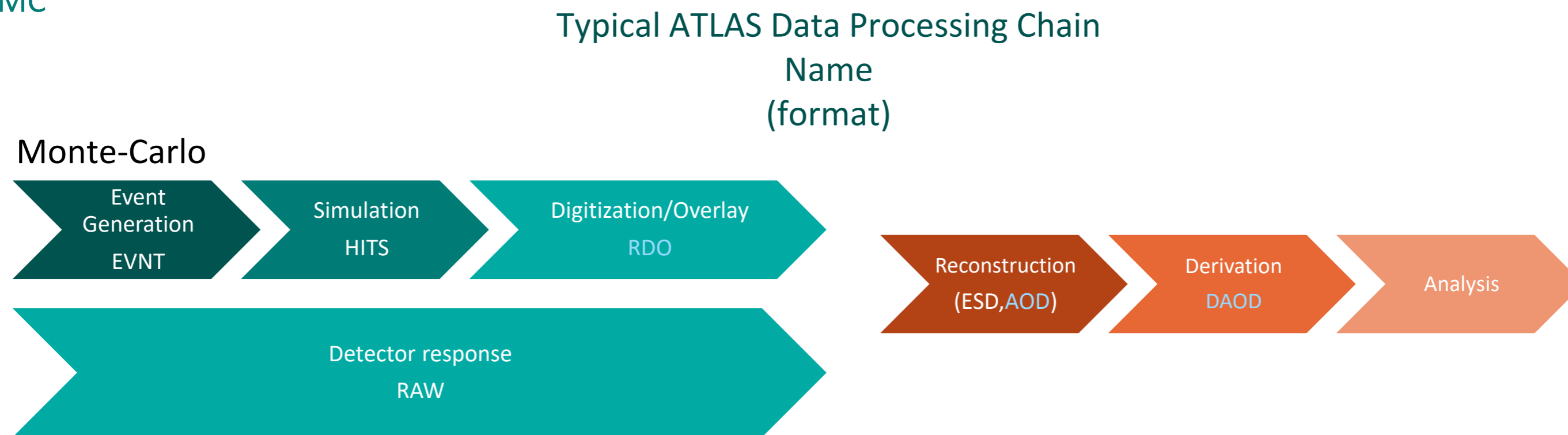
---



**T. Ovsianikova (UWashington, US) [tovsiank@uw.edu](mailto:tovsiank@uw.edu)**  
Alaettin Serhan Mete (ANL, US) [amete@anl.gov](mailto:amete@anl.gov)  
Marcin Nowak (BNL, US) [mnowak@bnl.gov](mailto:mnowak@bnl.gov)  
Peter Van Gemmeren (ANL, US) [gemmeren@anl.gov](mailto:gemmeren@anl.gov)  
on behalf of the ATLAS Computing Activity

# Motivation

- The High-Luminosity era demands optimized I/O, making RNTuple crucial for efficient data handling
- RNTuple provides significant disk space reduction and I/O CPU usage improvements via parallel, asynchronous operations and direct GPU memory transfers
- After years of testing in ATLAS, ATLAS is now able to write all production formats using RNTuple
- The ROOT team presented the detailed studies on the compression and reduction size last year
- However, these studies were limited to open-source ATLAS data (only DAOD format), and production data could not be used
- The ATLAS DAOD data2023 studies were presented on ACAT earlier this year showing around **20% disk space saving**
- The goal of this study is to measure the reduction effect using the main reconstruction and derivation production formats for both data and MC

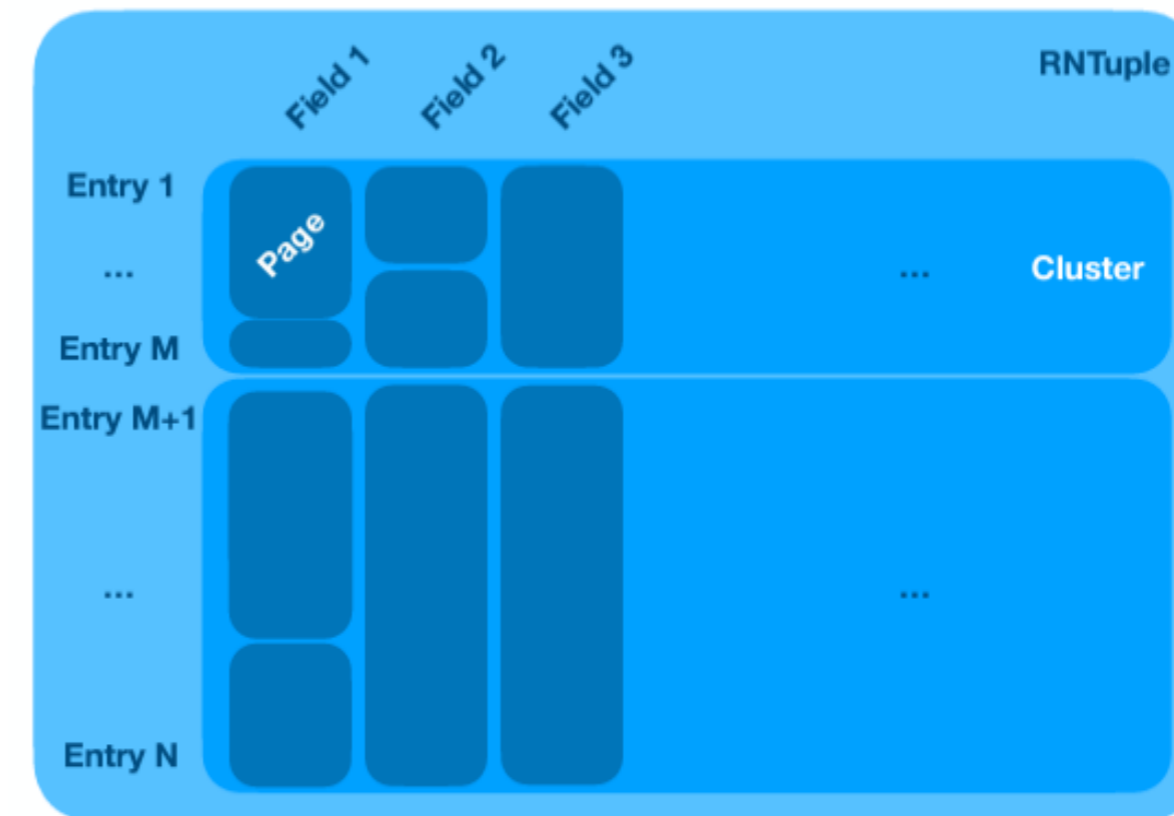
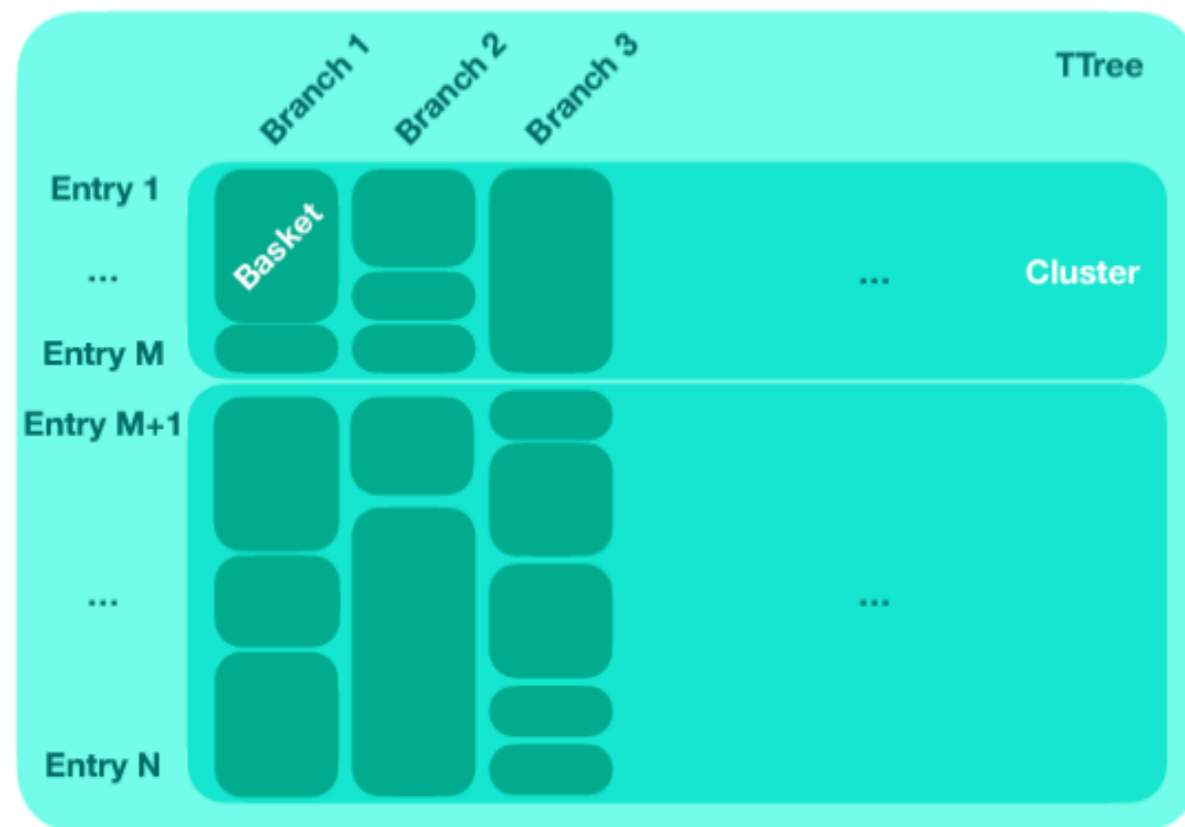


## ATLAS I/O with TTree:

- Each branch is stored in baskets, which are compressed separately. After the first flush (around 500 events), baskets are re-optimized to improve compression

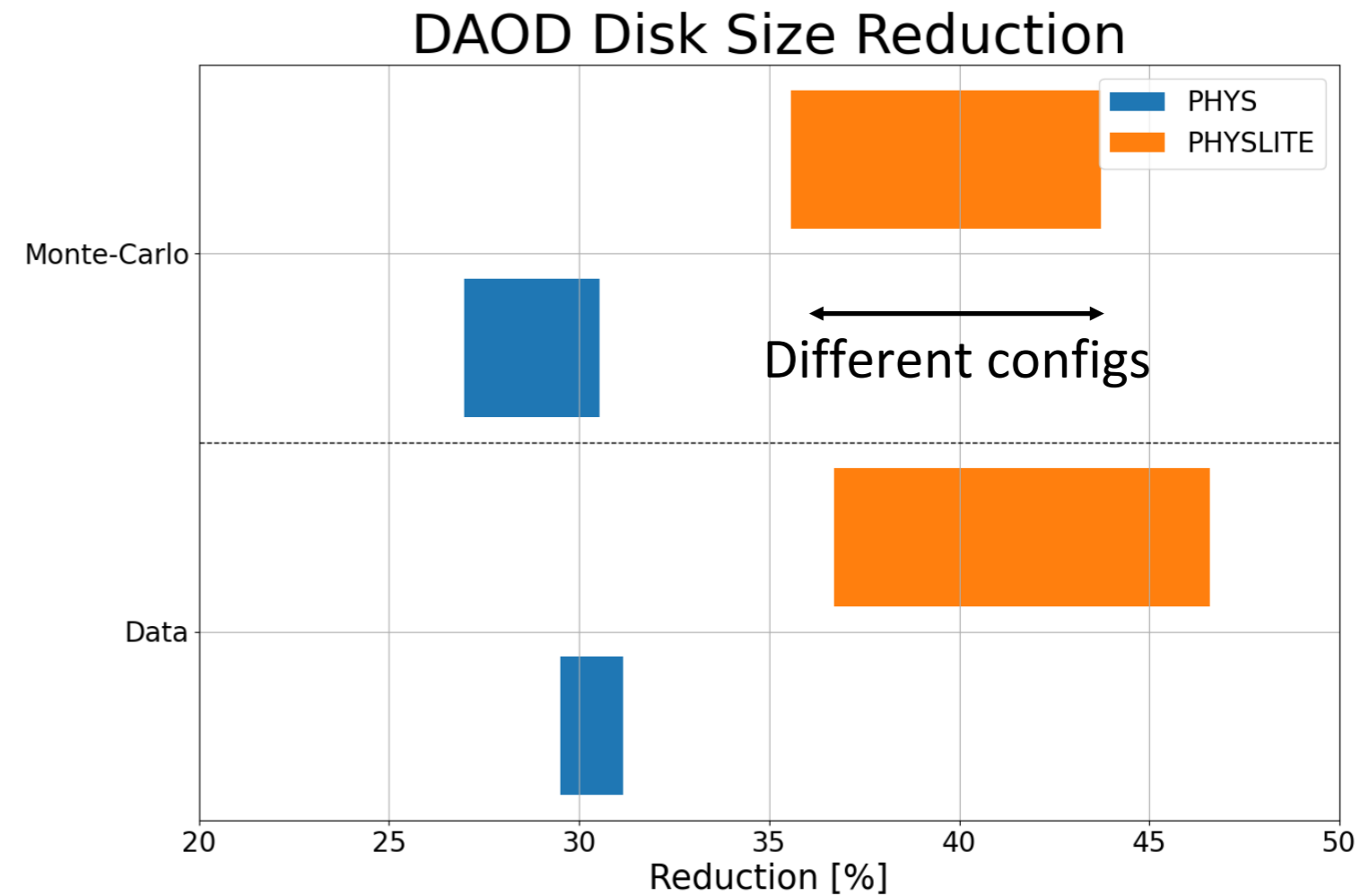
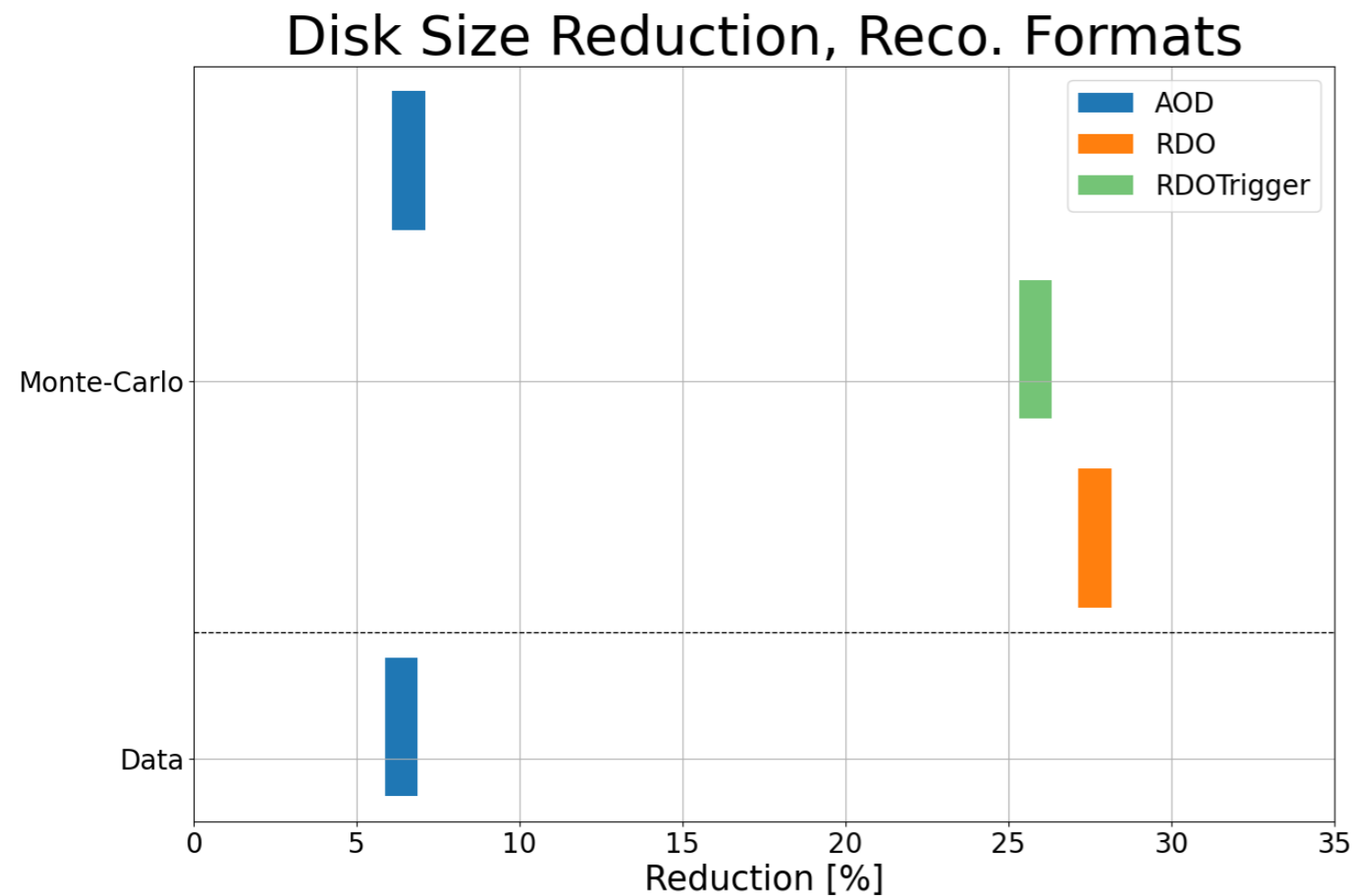
## ATLAS I/O with RNTuple:

- Fields are stored in columns as pages, usually page sizes smaller than optimized TTree baskets
- Very recently addressed by a new, adaptive algorithm to adjust page sizes
- More details can be found in the Marcin Nowak talk [“Adoption of ROOT RNTuple for the next main event data storage technology in the ATLAS production framework Athena”](#)



# RNTuple vs TTree size reduction

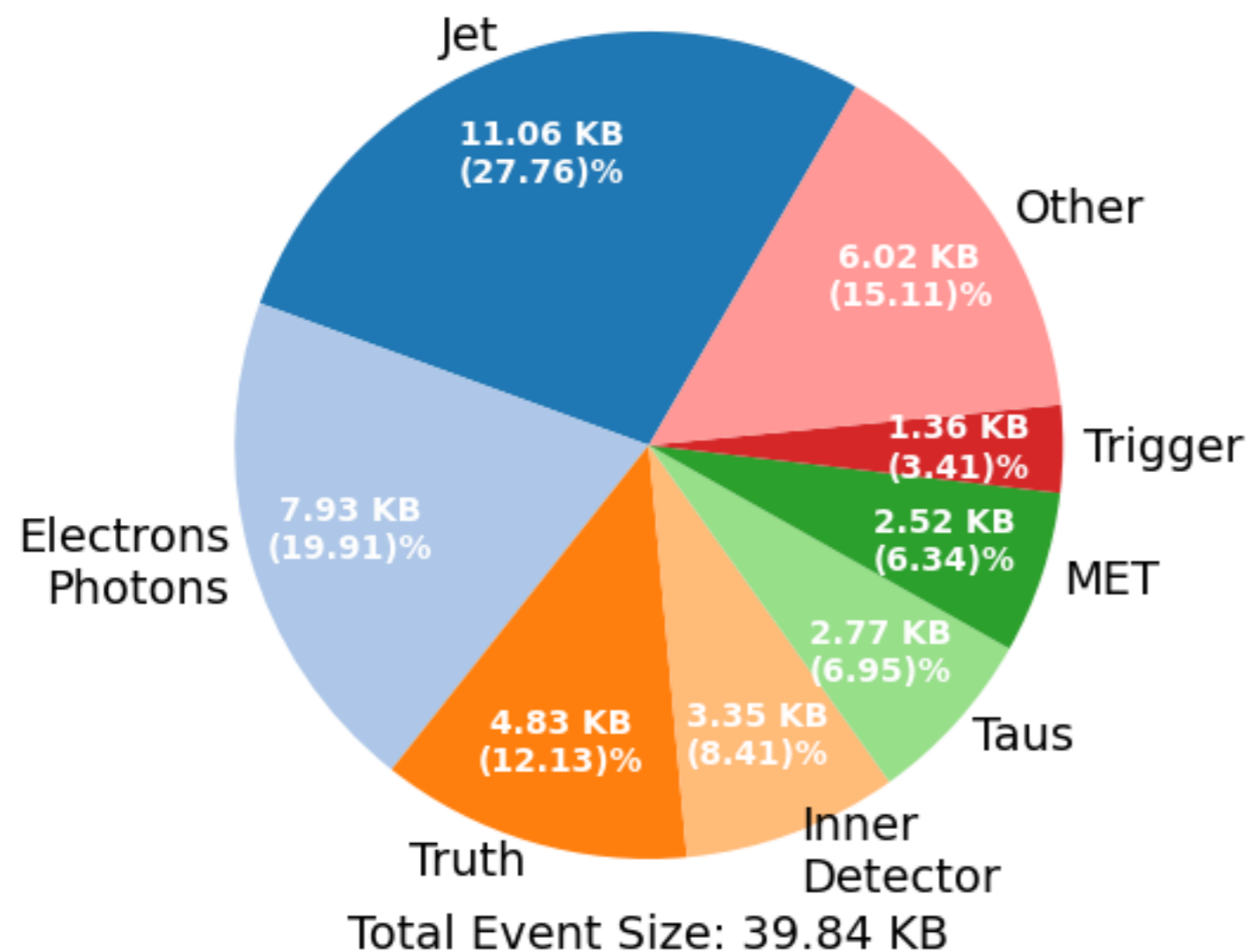
- RNTuple vs. TTree output size was studied for reconstruction and derivation ATLAS formats
- The ROOT head of master version as of 01 October 2024 was used
- Compression was studied using the zstd compression algorithm (standard for ATLAS)
- **Reconstruction:** 10k events sample Data and MC for the common formats (single sample with data23 pile-up profile for each format)
- **Derivation:** 100k events samples with various configurations for both Data and MC (different pile-up profiles and conditions)



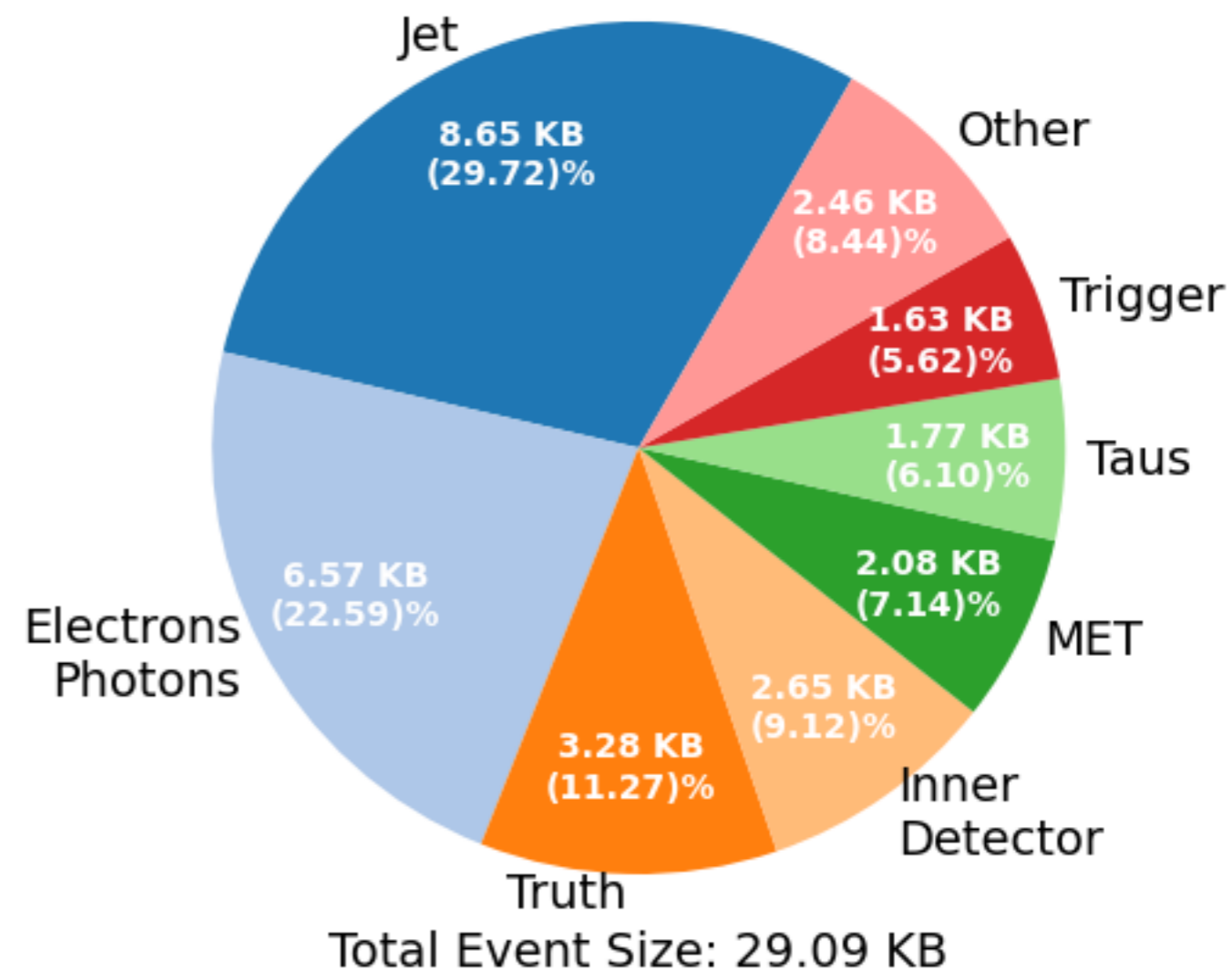
## DAOD per events size reduction

- Reduction seen for most of the domains with few exceptions for some branches
- Should be resolved before HL-LHC:
  - As example: the latest trigger optimization reduces the size of the trigger domain with factor 4 for both RNTuple and TTree

TTree: Container Sizes per Domain  
MC23 PHYS

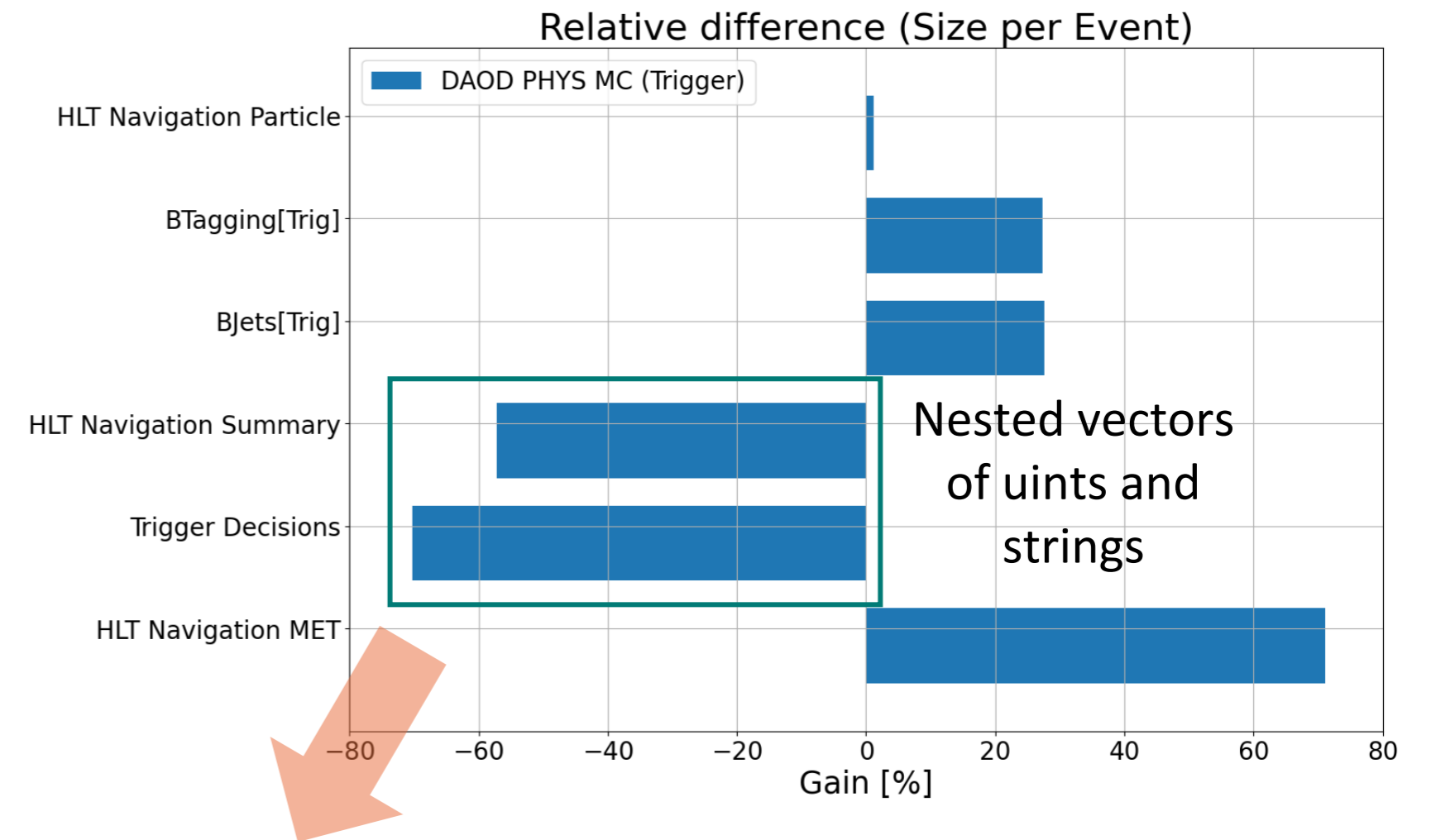
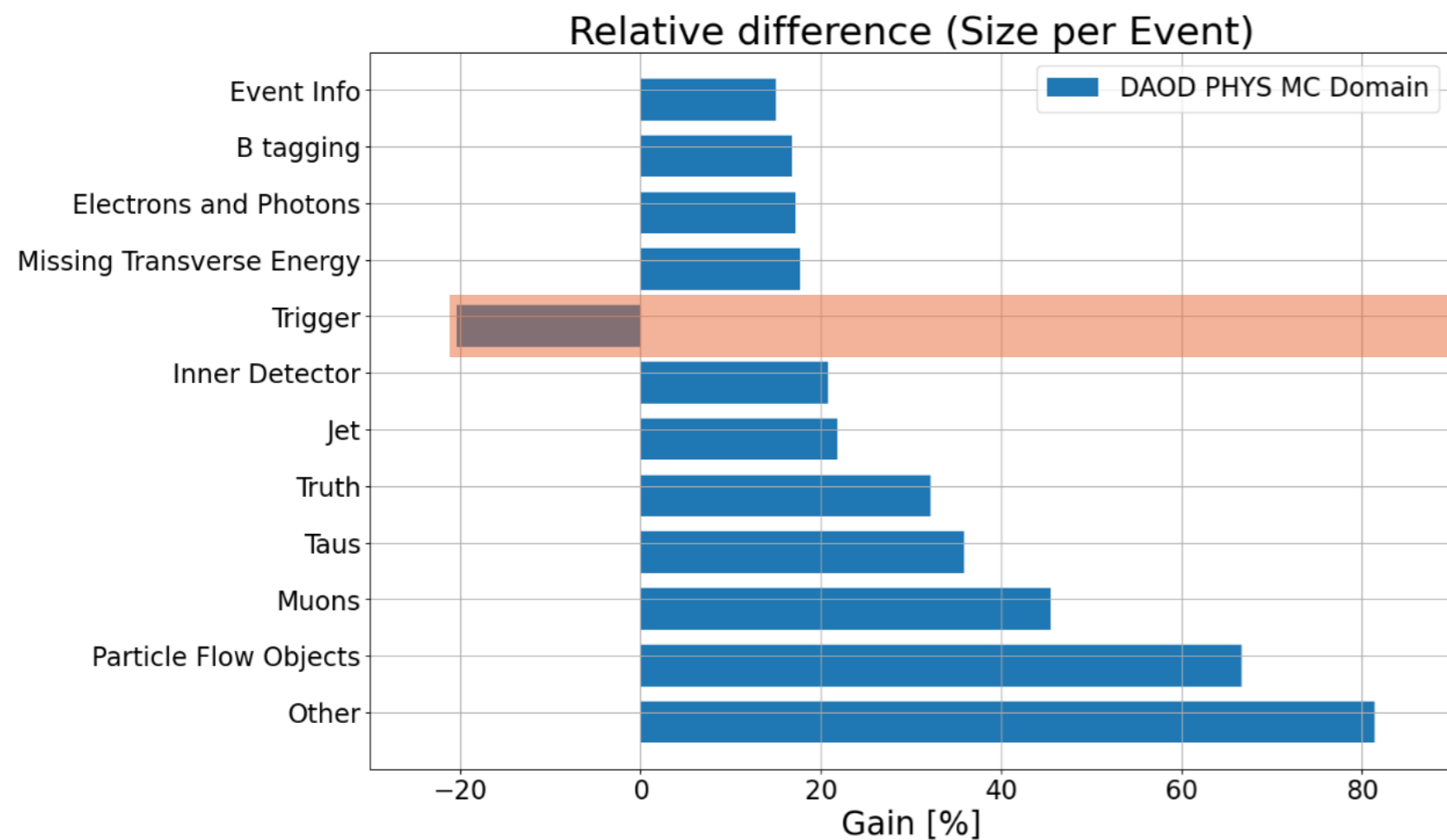


RNTuple: Container Sizes per Domain  
Job: MC23 PHYS



# DAOD per events size reduction

- Reduction seen for most of the domains with few exceptions for some branches
- Should be resolved before HL-LHC:
  - As example: the latest trigger optimization reduces the size of the trigger domain with factor 4 for both RNTuple and TTree

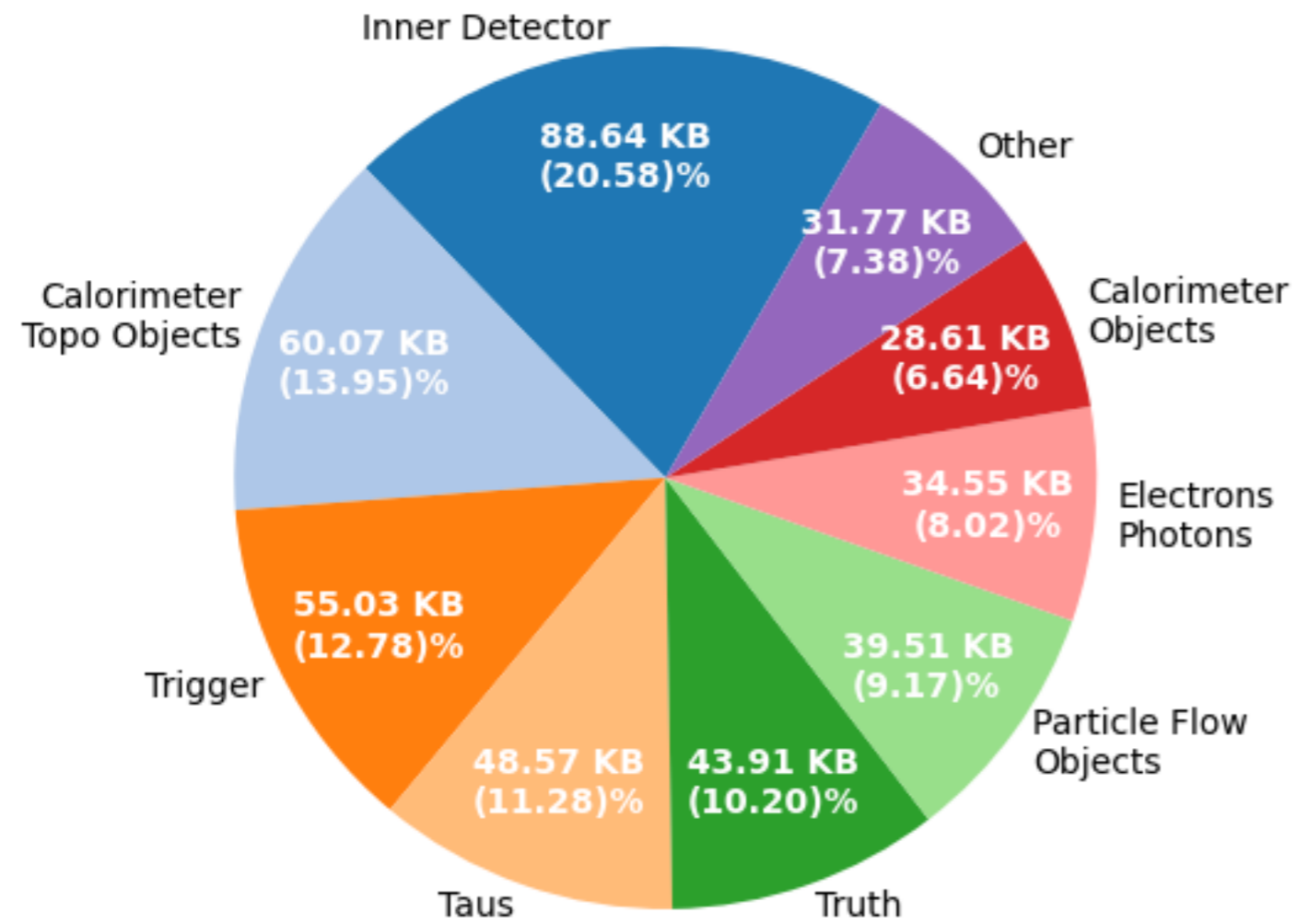


mostly understood would be resolved soon

# AOD per events size reduction

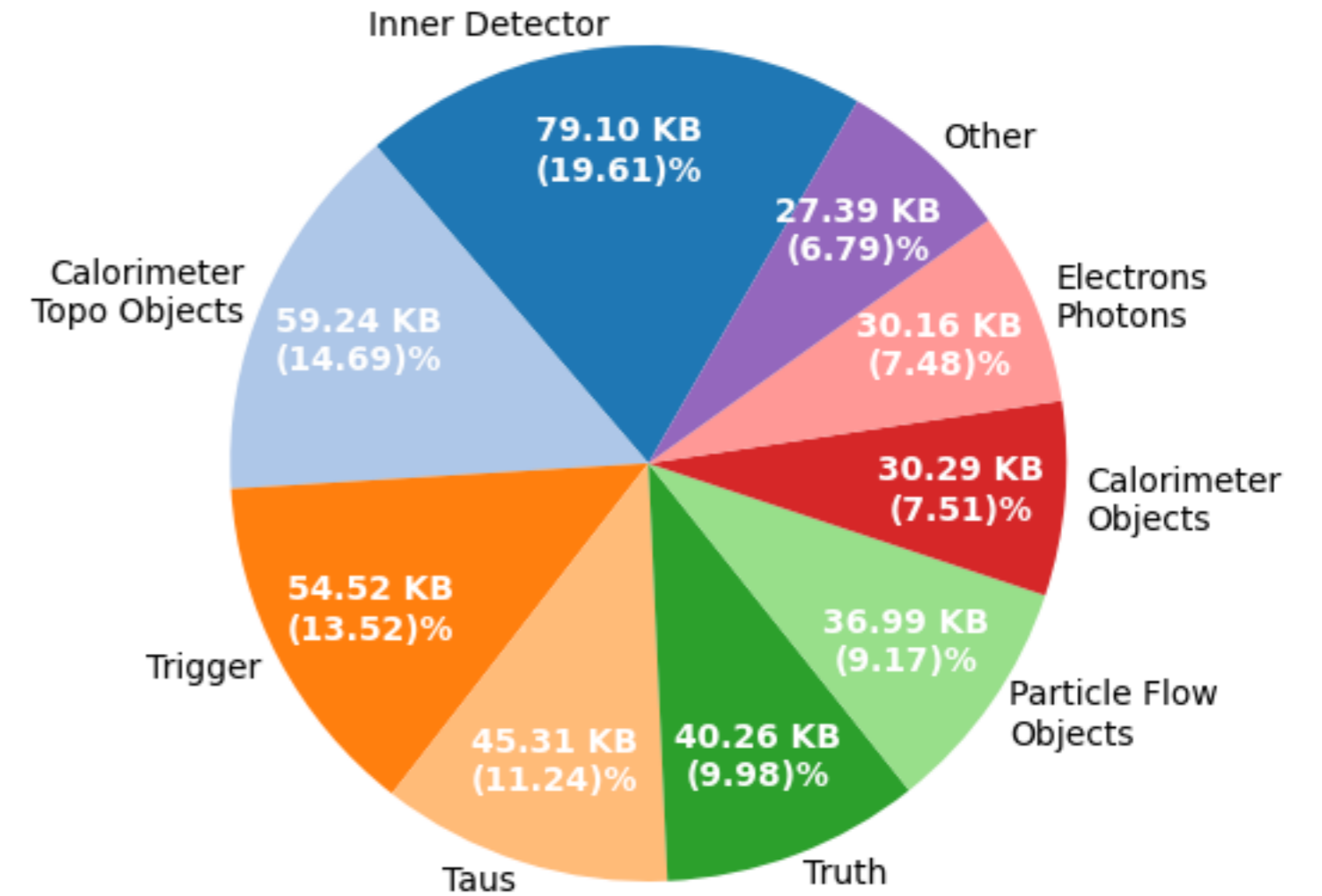
- In each domain there are some branches require more storage for RNTuple
- Custom page sizes for larger branches could improve performance

## TTree: Container Sizes per Domain AOD MC23



Total Event Size: 430.66 KB

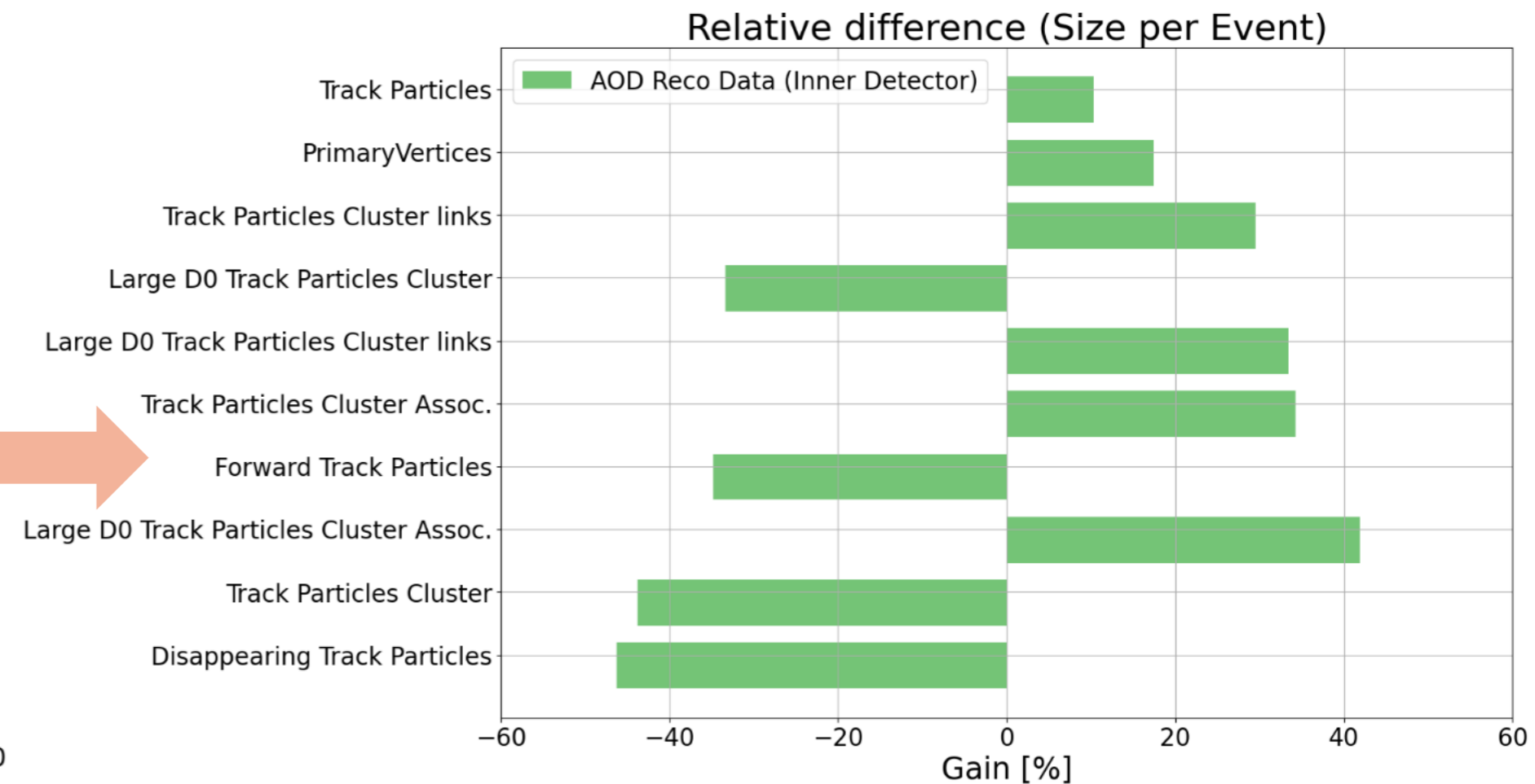
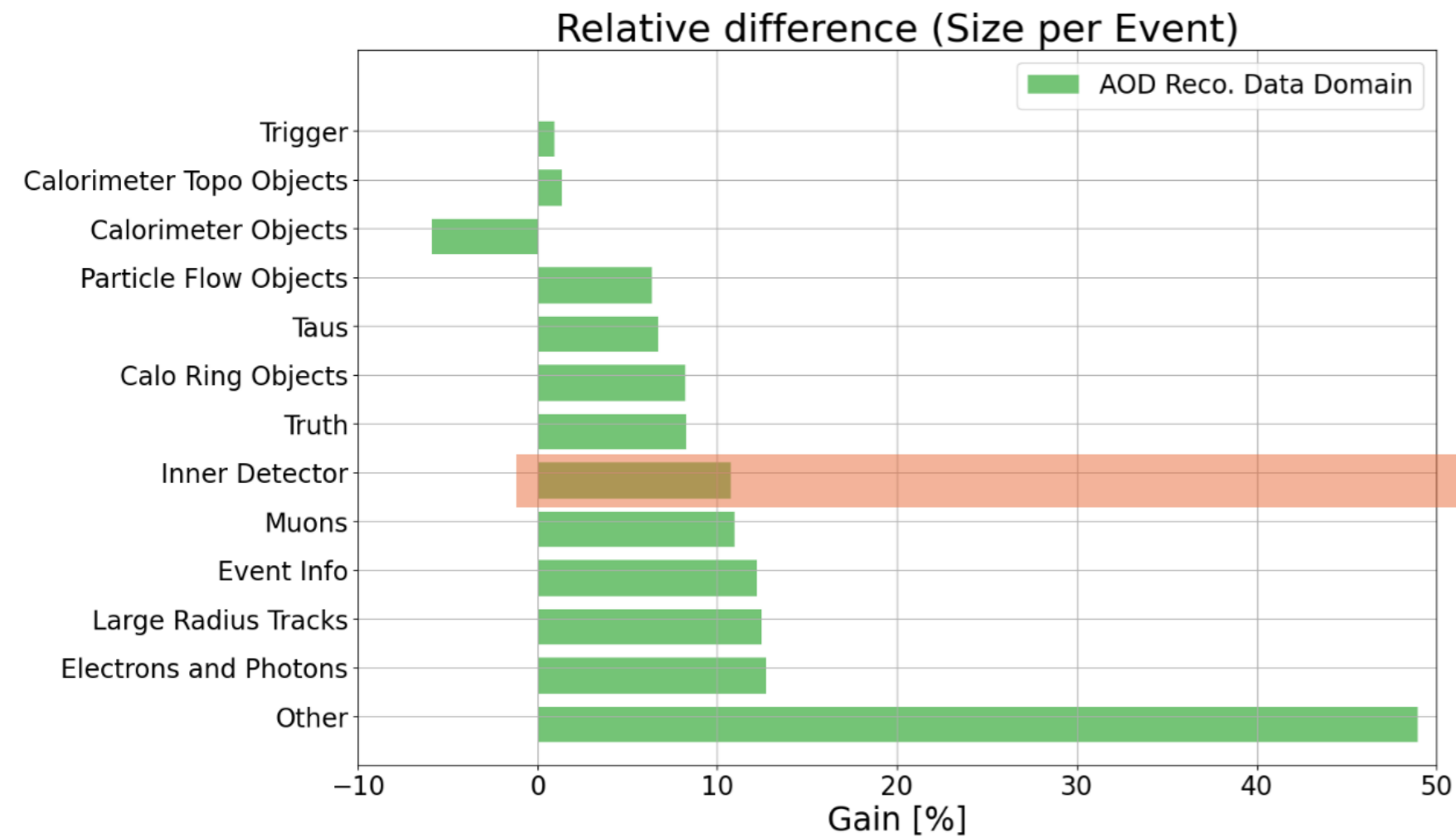
## RNTuple: Container Sizes per Domain AOD MC23



Total Event Size: 403.24 KB

# AOD per events size reduction

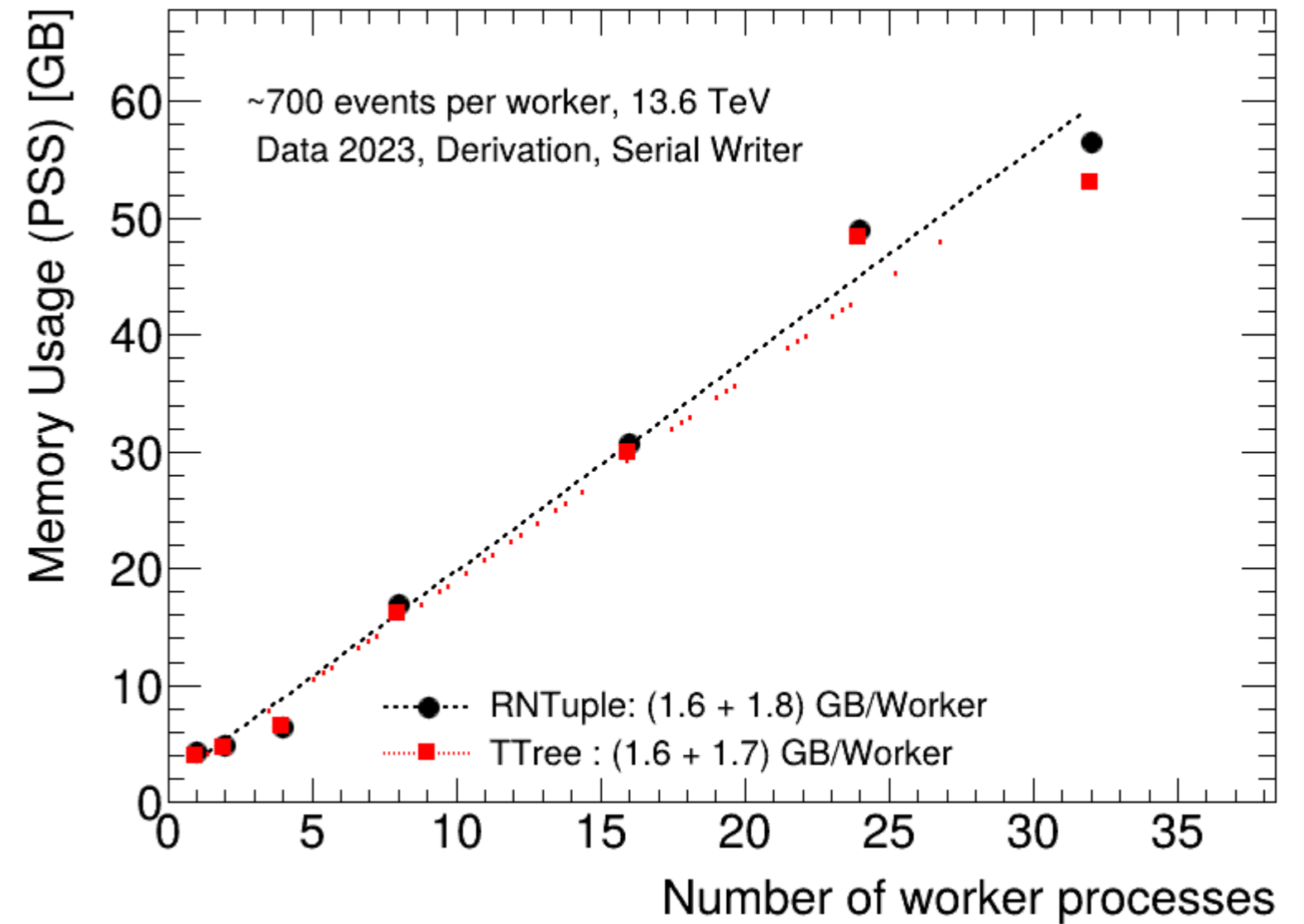
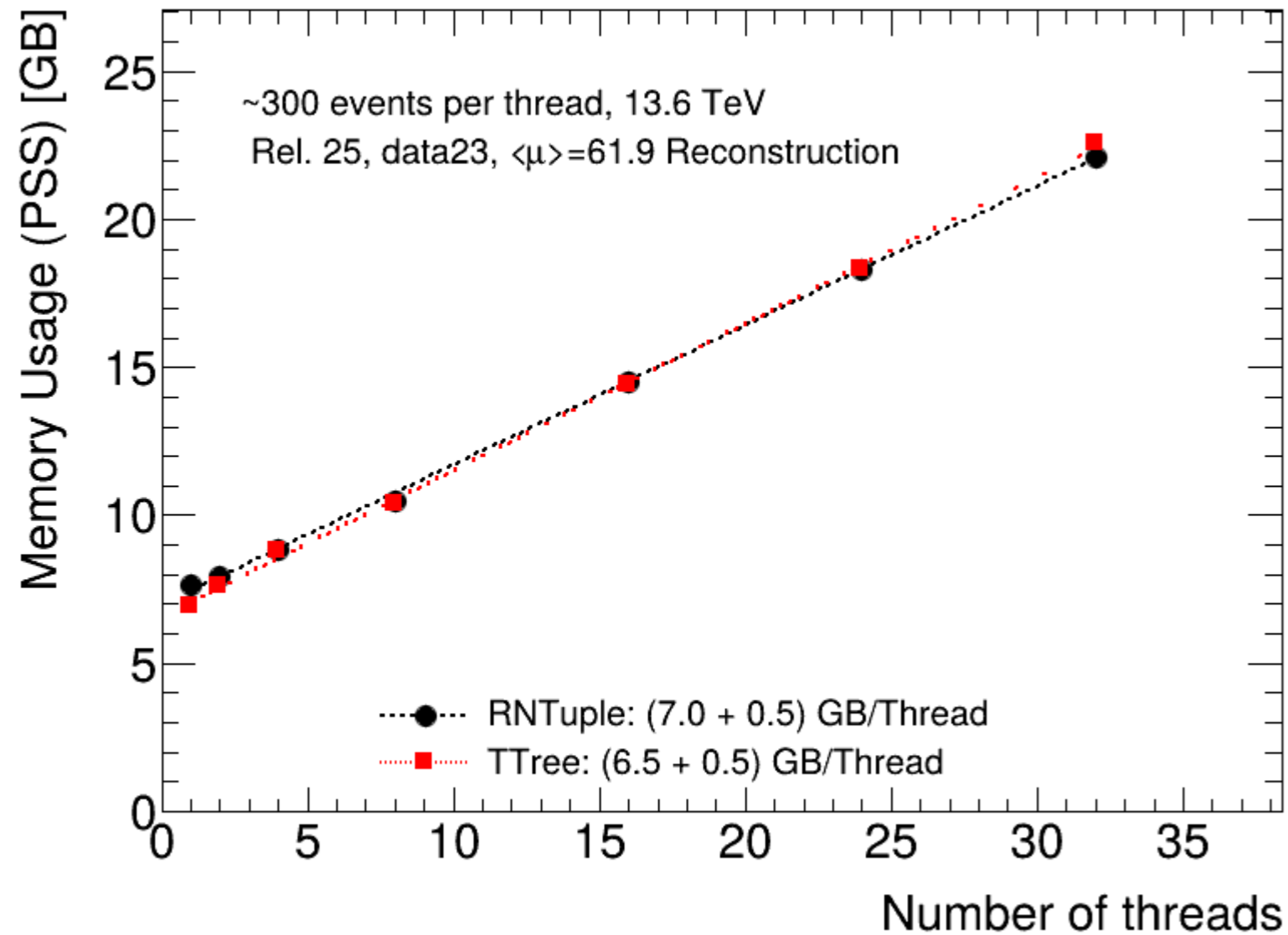
- In each domain there are some branches require more storage for RNTuple
- Custom page sizes for larger branches could improve performance



Most of the listed here branches/fields are vectors of or vectors of links to the complex objects that contains set of various C++ containers and data types



A small increase in memory usage was observed especially in derivation jobs



- Benchmarking in multi-processor mode shows an increase of approximately 100 MB per worker
- To be investigated...
- Since ATLAS has strict GRID memory limits, reducing memory usage is essential

## Conclusion

---

- RNTuple prototype is available for all ATLAS production formats
- RNTuple prototype offers significant improvements in disk space and I/O performance, but memory usage requires optimization
- While most containers shrink, some areas, such as nested vector<int>, show size increases, suggesting the need for further refinements
- Significant progress has been made on trigger data in the latest developments, but further improvements would be beneficial

**Thank you for your attention!**