RNTuple API Review: ATLAS Framework Open Items/Close Out

Alaettin Serhan Mete[1], Marcin Nowak[2], Peter Van Gemmeren[1]
[1]*Argonne National Laboratory,* [2]*Brookhaven National Laboratory*

Weekly CCE-SOP Meeting
20 November 2024

Argonne NATIONAL LABORATORY

Current Status of RNTuple in ATLAS/Athena

- **Athena has a fully functional prototype for some time now**
  - Over 80 MRs with the first one dating as early as January 2023
  - Already integrated into CI/CD (DAOD) and larger-scale release tests (more)

CI

Result	Test name	Optional or required
✔	CITest_DerivationRun3Data_Train_RNTuple-test	Required
✔	CITest_DerivationRun3MC_Train_RNTuple-test	Required

ART

test_data23_daod_phys_physlite_diff.sh	0 1 2 3 4 5	succeeded
test_data23_daod_phys_physlite_diff_metadata.sh	0 1 2 3	succeeded
test_data23_legacy_sharedwriter_rntuple_phys_physlite.sh	0	succeeded
test_data23_rawtoall_rntuple_esd_aod.sh	0	succeeded
test_mc23_daod_phys_physlite_diff.sh	0 1 2 3 4 5	succeeded
test_mc23_daod_phys_physlite_diff_metadata.sh	0 1 2 3	succeeded
test_mc23_legacy_sharedwriter_rntuple_phys_physlite.sh	0	succeeded
test_mc23_overlay_trigger_reconstruction_rntuple.sh	0	succeeded

- **In a nutshell, RNTuple is a functional storage backend in Athena**
  - All official data products we support in TTree are also supported in RNTuple at this point
    - Not all workflows are fully supported yet, e.g., AthenaMP + SharedWriter w/ parallel compression (more later)
    - We’re currently looking into optimizing various aspects of our RNTuple data already!
  - Still, ATLAS has plenty of work to do before deploying RNTuple in production
    - Certain aspects go above and beyond the core framework support, e.g., analysis etc.

API Review: General Remarks

- **ATLAS thanks CCE for facilitating this review and the ROOT team's collaboration**
  - Most, if not all, core ATLAS requirements are addressed before/during the API review
    - E.g., model updating (late attributes), emplacing new values (generating default objects), etc.
- **So far, naturally, we've mostly concentrated on the functionality**
  - That's not to say we haven't looked at a number of optimization aspects
    - Split vs non-split encoding of fields etc. (already discussed in previous meetings, e.g., see here)
- **Over the course of the API review, we made a number of improvements**
  - Using the proper API, e.g., `RNTupleReader` → `RPageSource` → `RNTupleReader`
  - Adopting API changes that result from the review, as well as internal developments
- **TL;DR Current RNTuple API provides sufficient functionality for ATLAS**
- **Having said all these, ATLAS still has significant work ahead**
  - Optimizing RNTuple usage in various different modes and for data different formats
    - Serial, multi-process (MP), multi-thread (MT), and even hybrid (MT/MP) Athena jobs
    - HITS, RDOs, ESDs, AODs, and DAODs all have different characteristics
  - Elephant in the room: Adoption in physics analysis (beyond our scope here)

Open Items on the API and Functionality

- **There are a still few to-be-addressed functional points:**
  - **These were all previously discussed and agreed to be addressed in the next version**
  - **On-the-fly fast merging of RNTuples**:
    - Current merging works in an `hadd` scenario (merging existing/finalized files)
      - A few shortcomings there too, e.g., no way to pass `RNTupleMergeOptions` through `TFileMerger` (e.g., select union)
    - However, it is **missing incremental merging of open/in-memory RNTuples**
      - This is used in AthenaMP + SharedWriter (w/ parallel compression) derivation production jobs producing TTrees
      - N clients produce their own TTrees and a server merges client data (via TFileMerger) and resets the objects
        - A functionality akin to parallelMergeServer+parallelMergeClient setup (via networking)
  - **Custom indexing (a la `TTree::BuildIndex("...")`)**:
    - `auto view = reader->GetView<...>("foo");`
    - `auto result = view(idx);` where `idx` is a custom field value instead of `[0...N]`
  - **Friend RNTuples (a la friend TTrees)**
    - This is a functionality we've been investigating for Run 4, see Custom Event Sample Augmentation

Open Items on the API and Functionality (cont'd)

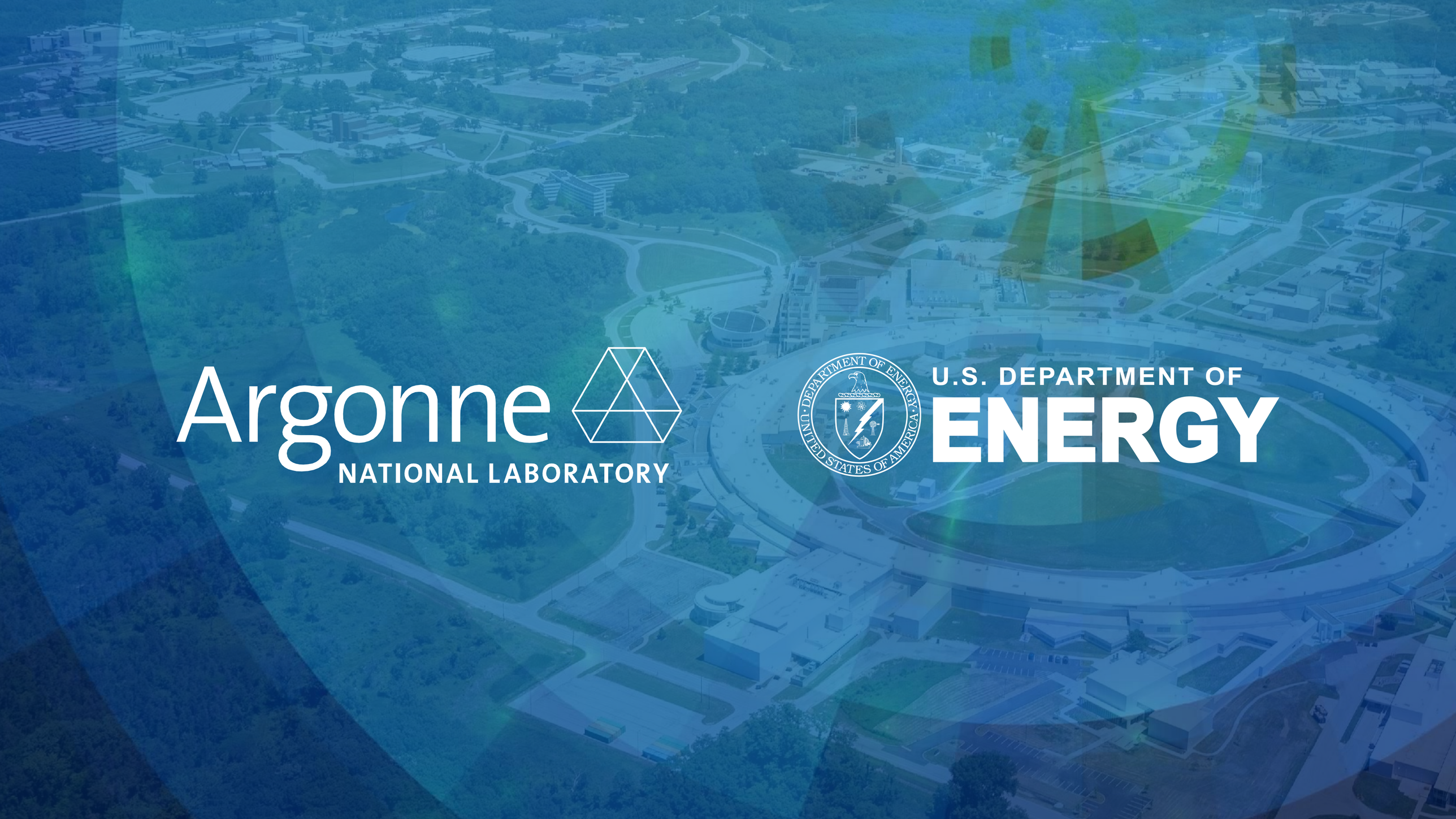
- **There are also other less critical points:**
  - We internally cache `RNTupleView` objects for performance
    - Currently no way to create an empty view, we need to use dynamic allocation
      - `viewMap[fieldID] = std::make_unique<RNTupleView<void>>(reader->GetView<void>(fieldName, nullptr));`
      - Perhaps this can be improved on both ROOT and Athena sides
  - Streamlined/simplified API with more configurability:
    - Configure split/unsplit encoding in an easier way (now requires looping over all subfields by hand)

```
for (auto& subfield : *field) {
    if(subfield.GetTypeName() == "float") {
      subfield.SetColumnRepresentatives({{ROOT::Experimental::EColumnType::kReal32}});
    }
    else if(subfield.GetTypeName() == "std::int32_t") {
      subfield.SetColumnRepresentatives({{ROOT::Experimental::EColumnType::kInt32}});
    } // Check/set each type's column representatives by hand
}
```

    - Page size configuration per (sub)field (this can be done globally)

Conclusions and Outlook

- **Current RNTuple API provides sufficient functionality for ATLAS**
  - Here, we cover the framework-level requirements, not analysis use-cases
  - However, there are still desired missing functionality (see previous slide for examples)
- **Ongoing effort in measuring the performance/ performing optimizations**
  - Athena has many execution modes: serial, MT, MP, and MT/MP, with different signatures
  - Moreover, different formats have different characteristics, e.g., ESD vs DAOD etc.
- **Next steps for ATLAS involve:**
  - Addressing the missing functionality, especially for derivation production (fast merge)
  - Profiling the performance in all aspects and optimizing the workflows
    - E.g., reproduce all DAOD PHYSLITE OpenData in RNTuple v1.0 for 1-to-1 comparisons w/ TTree
  - Move the existing prototype in the framework to a production-ready state
  - Help with the RNTuple adoption on the analysis side
- **ATLAS looks forward to continued collaboration with the CCE/ROOT teams!**


Argonne
NATIONAL LABORATORY
U.S. DEPARTMENT OF
ENERGY