

DUNE Frameworks Workshop Summary

Mathew Muether – Wichita State University

Dune Computing Meeting

July 8, 2024

Workshop Summary

- Met June 27-June 28
- Goal to "finalize" initial requirements with devs and plan next steps to move to initial design.
- 27 registered participants participated in Person and Online
- Following introductory talks on computing and past process, focused on general discussion with dev on organizing process to formalize requirement and milestones.
- Remained of WS focused on discussion around data products, persistency, checkpointing, error-handling, and backwards compatibility.
- Live notes doc:
<https://docs.google.com/document/d/1S0xfWP8AH83n-UJ0vQfsjmDIvCbU-sDdScafK14o52E/edit#heading=h.d7stj43yjfe7>

Outcomes

- The developers will be requirement custodians with DUNE as requirement owners.
- Requirements will be managed using Jama Connect.
- Requirement lifecycle will follow the flow-chart on p. 13 of <https://indico.fnal.gov/event/65009/contributions/294777/attachments/179211/244599/dune-framework-workshop-intro-2024-06-27.pdf> (with potential adjustments such as adding an “Assigned” state).
- Collaborative development model will be as described on p. 14 of <https://indico.fnal.gov/event/65009/contributions/294777/attachments/179211/244599/dune-framework-workshop-intro-2024-06-27.pdf>.
 - Code will be developed collaboratively in GitHub.
 - Need to decide which repo to use, which core software is connected.
- Deliverable milestones will be formed based on approved requirements (“baselines”) to be met.
- Not everything has to be runnable on a CPU.
- Stakeholder group will persist for 2-3 years.
- docdb will be used to communicate design ideas between the developers and the collaborators. (High level diagrammatic content preferred. Having examples that compare art vs. the new framework would be helpful.)

Action Items

- Create Slack channel for developer/DUNE interaction
- Determine how to increase/coordinate efforts in evaluating RNtuple.
- Clarify if ML training needs to be able to run inside a framework job.
- Determine whether DUNE intends to use software (Wire-Cell etc) as an in-between, non-framework process.
- Figure out the politics of where to put code, related to ownership etc
- Figure out the transition from the TF to the stakeholders group
 - Chris M. will lead the formation of the DUNE stakeholders group. Work with spokes and Gary to ensure that it counts as a Service Task. Required by October LBNC meeting, goal August.
- DUNE needs to consider future ML needs — the FW is a natural way to overcome issues of reproducibility and provenance for training datasets, and has requirements to work with HPC and GPU. But DUNE's needs are not well enough understood to make progress on this discussion now

Closeout and Planning

- Discussion items for next time:
 - Resource requirements (e.g. don't schedule this particularly memory-intensive task at the same time as other things)
 - Adjacency requirements/time ordering
 - Provenance
 - Reproducibility
- Future Workshop: General agreement that it would be very useful but will be a challenge to schedule. likely late August pre-FNAL closure or late September after collab meeting.
- Overall timeline planning
 - Requirements finished early September
 - DUNE stakeholders and devs iterate as needed
 - First pass of design doc by October LBNC meeting
 - Commission external review of design doc (with a carefully written charge so the reviewers don't try to rewrite all the requirements). Likely late October/early November
 - External review complete before DUNE-US DOE review late 2024/early 2025