



2024 LArSoft Work Plan

Last updated: June 18, 2024

Introduction

This document gives the LArSoft work plan for 2024. Progress on specific items going forward are given quarterly at Steering Group meetings.

Erica and Katherine discussed priorities with each experiment in a series of meetings in September and October of 2023. The experiments detailed their plans for the next year, the requirements for LArSoft, and how the LArSoft Project Team (“the Project”) could help, as well as what the experiments might be able to contribute to LArSoft code. Some items may be raised by the Project to address general problems faced by all LArSoft users. Major observations from these discussions are listed in Appendix B since they may be of interest to other experiments and the project as a whole.

Based on those discussions, LArSoft proposes a plan of work for 2024 along with relative priorities of the various items. The Project resources responsible for executing parts of the plan are listed in cases where this is known. It should be noted, however, that the Project does not have sufficient staff at present to carry out the full program of work in this plan, or the entirety of the short term or long term programs individually. The Project will collaborate with experiments to prioritize and coordinate the work to best meet experiment needs. We also note that some work related to the plan can be performed by the experiments. Allowing experiment members to receive service credit for contributions to work plan items that benefit their experiment, but that are of a more general nature would help to fill some of this gap in effort. The more this happens, the more all experiments benefit.

1. Short term priorities for 2024

We define short term priorities as those items that should receive sufficient effort to be substantially completed in CY2024 due to the urgency or timeliness of the need, and where the Project will drive the timeline. The meaning of priorities listed is discussed in Appendix A.

1. **Thread safety, multi-threading, and High Performance Computing (HPC) for critical production workflows:**
 - **Re-architect LArSoft and experiment code to ensure thread safety.**
 - **Implement multi-threading where appropriate to improve resource utilization or to enable execution on HPC resources.**
 - **Provide support for hardware acceleration where appropriate.**
 - **Coordinate / collaborate with efforts to use LArSoft on HPC platforms.**

Priority: high

This is a multi-year effort that started in 2018 working on thread safety and multi-threading to improve resource utilization on existing grid resources, and to prepare LArSoft code to be effective in an HPC environment. A fully thread-safe production workflow and the ability to utilize GPU-as-a-Service for deep-learning inferencing within LArSoft were completed in 2021. In 2022 and 2023, the focus was on critical production workflows where the introduction of these advanced computing techniques promised improved resource utilization or increased throughput. Work completed included making production database services thread safe, and making a DUNE super-nova neutrino workflow run in multi-threaded mode.

Work in 2024 will focus on:

- replacing select algorithms with portable GPU-enabled algorithms.
- enabling the use of GPUs in production workflows more generally
- continuing efforts to ensure production workflows use multi-threading to optimize resource utilization
- facilitating use of HPC resources for LArSoft workflows (eg, via POMS)

Resources: Kyle Knoepfel, Mike Wang, Marc Paterno, Saba Sehrish, and others who were previously working on the HEPReco SciDAC

March 12 update:

- Marc Paterno recently reported on work to optimize a photon simulation module that is part of a DUNE production workflow. Initial improvements in testing showed reduced execution time by four seconds per event, and identified memory issues that adversely affect performance. The ultimate goal is to parallelize execution and accelerate execution using GPUs, should studies demonstrate a substantive gain in performance and execution time of the workflow.

June 18 update:

- A summer student, supervised by Giuseppe Cerati, is working toward incorporating Graph NNs into LArSoft via the NuSONIC client with the assistance of Mike Wang.
- Marc Paterno validated the physics results of the initial set of serial optimizations noted in the March 12 update and integrated them into LArSoft. He followed this by benchmarking the performance of various possible solutions to the memory issues previously identified. The potential solutions are well adapted to running on a GPU should eventual conversion to a GPU algorithm be indicated.

2. Provide support for a multi-experiment event display capable of dedicated and integrated displays of TPC charge collection and photo-detector systems, along with external detectors such as CRTs, spectrometers, calorimeters, etc.

Priority: high

Decide whether LArSoft will provide support for Titus (or possibly some other) experiment-agnostic framework for event displays. Titus support would entail re-architecting to make proper, extensible use of the LArSoft Geometry system, providing a library of primitives to assist in displaying LArSoft geometry and data product information and in controlling the display, and developing a system of plug-ins for experiment-specific elements. Similar work would apply to a non-Titus solution. Agreements on long-term maintenance and code ownership would be part of this work as well.

March 12 update:

- No update

June 18 update:

- No update

3. SPACK - migrate to a new, community-supported build system for art and LArSoft

Priority: high

This is a continuation of a multi-phased project to migrate LArSoft to a Spack-based packaging and build system, which has broad HEP community and industry support. Via this migration, we will address a number of long-standing portability and configurability concerns raised by the experiments that are inherent to the current system. The overall strategy centers on minimizing disruption to the experiments by providing backward compatible changes until the final migration. The first major milestone, preparing LArSoft for the transition to a system with a Spack-based back end, was completed in Q4 of 2021. The final phase will end support for UPS. An aggressive education campaign is part of this project.

Details on Spack at Fermilab can be found in a [presentation by Marc Mengel](#) at the LArSoft Coordination meeting. Further information on the build system can be found in the following presentations by Chris Green: [Feb 23, 2020](#), [Nov 2, 2020](#). Update on [September 2023](#). and an overall plan – [Spack strategy.docx](#) with instructions at <https://fifewiki.fnal.gov/wiki/Spack>

Completion of the transition to Spack is expected by Q2 2024.

Resources: Patrick Gartung, Chris Green, Kyle Knoepfel

March 12 update:

- Kyle Knoepfel demonstrated a model for a development environment under Spack. Though a number of details remain to be worked out, the model showed a promising, practical approach for creating a work environment that avoided many of the problems that plagued earlier approaches.
- CSAID has committed to technical solutions for experiments that need to use either UPS, the legacy packaging and runtime environment tool that is being phased out this summer, or SL7 to build and run their code, as that operating system reaches end of life this June..

June 18 update:

- Kyle gave an update at LArSoft Coordination Meeting on May 28th on developing under Spack - <https://indico.fnal.gov/event/64878/contributions/292297/attachments/178392/243150/developing-with-spack-2024-05-28.pdf>.
- As of mid-May, Marc Paterno is now leading the final stages of the Spack migration project, and is drafting a plan and timeline for completion.

4. Pixel detectors within LArSoft

- **Re-architect geometry systems to allow integration of multiple readout schemes**
- **Support other development work needed for pixel detectors in LArSoft**

Priority: high

Refactor the LArSoft geometry system to separate volume geometry hierarchies from readout geometries and integrate wire-readout and pixel-readout geometries. Work with experiments to ensure any other adaptations to data products, simulation or reconstruction code needed to support pixels are made. Work in 2023 completed the refactoring and validation of the geometry system. Continuing work in 2024 will include integration and validation of a pixel readout geometry aligned with DUNE ND, and support for pixel simulation and reconstruction integration.

Details on the project and the progress of work can be found in [this presentation at the LArSoft Coordination Meeting](#).

<https://cdcv.s.fnal.gov/redmine/issues/19328>

Deployment is expected early in 2024.

Resources: Erica Snider, Kyle Knoepfel

March 12 update:

- Work has focused on bringing the new geometry system into production. Documentation for the system was completed in February, which cleared the way for experiments to begin validating the system in preparation to sign-off on deployment. A draft pixel readout geometry already exists, and will be integrated into the geometry soon after deployment.

June 18 update:

- We are working with DUNE collaborators to resolve an issue that arose during DUNE FD validation. Once completed, we intend to deploy the new system and begin integration of a pixel readout geometry that can be used in the simulation.

5. Neutrino event generator refactoring

Priority: high

Refactor the interface to the GENIE neutrino event generator such that the GENIE version can be selected through the runtime environment configuration. One side effect will be to provide a template for how to interface other event generators so as to simplify integration and maximize flexibility. The work will commence after the final phase of the Spack migration is completed.

Resources: Robert Hatcher, Steven Gardiner, Liang Liu

March 12 update:

- Experts met in late February to discuss the final steps needed to complete this task. A new postdoc, Liang Liu, has joined the group and will be performing the bulk of development work.

June 18 update:

- Working to understand how Spack affects the eventual solution, since that transition is imminent and will change the handling of dependencies.

6. LArSoft infrastructure updates

There are several items that came up in discussion with the experiments about their plans for 2024. These include, for example:

- Sampling frequencies vary across TPCs in protoDUNE, while LArSoft supports only a single value.
- Support for non-planar cathode geometries to facilitate tracking across non-planar cathodes.
- Support for TPC-dependent drift velocities and electron lifetimes.

March 12 update:

- No update

June 18 update:

- No update

7. Review and update / augment LArSoft documentation in key areas as requested by experiments.

Priority: high

Continuing project to provide improvements covering areas noted by experiments, and new developments in LArSoft capabilities and the surrounding ecosystem. Initiatives will include:

- Maintain and enhance an up-to-date introductory slide deck on LArSoft - <https://indico.fnal.gov/event/49621/#1-see-attached-source-material>, ensuring alignment with other sections of LArSoft.org
- Update Geometry documentation on larsoft.org and GitHub wiki.
- Verify links in GitHub wiki material and correct ones that are now invalid. (This happens over time.)
- Continuously review, enhance and update LArSoft documentation as needed.

Resources: SciSoft team

March 12 update:

- Efforts were focused on documenting the refactored geometry system in full, as noted above. The resulting documentation addressed a number of shortcomings in the documentation for the legacy system, so is generally an improvement over what was there before.
- Several efforts involving [Machine Learning with LArSoft](#) have been documented at larsoft.org. Specifically:
 - [NuGraph2 GNN event reconstruction](#)
 - [CNN-based event classification and particle ID using CVN](#)

June 18 update:

- Updated dependency diagram on https://larsoft.github.io/LArSoftWiki/LArSoft_repositories_packages_and_dependencies and in our presentation deck for LArSoft.
- Added information on [Waveform region-of-interest finding for supernova event triggering](#) to larsoft.org under the Machine Learning area. Also added a wiki page on [Using the 1DCNN raw waveform algorithm](#) to the github documentation for LArSoft.
- Created temporary instructions on using SPACK in LArSoft at: https://larsoft.github.io/LArSoftWiki/LArSoft_with_SPACK

8. Provide consultative support to experiments for transitioning to the refactored LARG4 detector simulation framework, or developing simulations for new detectors

Priority: high

Provide consultative support for efforts within the experiments to migrate to the new simulation framework. Advise and provide guidance in refactoring experiment code to be compliant with the workflows and interfaces within that framework.

Resources: Hans Wenzel for LArG4 / artg4tk, SciSoft team for geometry, other infrastructure

March 12 update:

- There have been discussions related to this, but nothing of note to report.

June 18 update:

- No update.

2. Long term and continuing priorities for 2024

We define long term priorities as those that we do not anticipate completing within CY2024 due to the nature of the work involved. Unless explicitly indicated, a project defined as a long-term priority should not be construed as implying it is “low priority”. A “continuing priority” is a set of tasks that by construction represent a continual, recurring stream of work that may or may not span more than a calendar year.

- 1. Provide technical expertise, advice and leadership in matters related to LArSoft and software development**
- 2. Develop and execute support plans as needed in areas where assistance from the Project has previously been requested by experiments, or advised by the Project.**

Priority: per request

- a. Align Wire-cell reconstruction and simulation with LArSoft workflows and interfaces (eg, factoring electron drift from anode response simulations via drifted electrons)
- b. Extend the NuWro direct integration model to include the GiBUU event generator framework and the NEUT event generator.
 - i. June 2021: Work proceeding within Genie Collab to provide event library solution to integrating external generators. This requires stand-alone running of the generators to produce the libraries, so is not the direct integration sought. Discussing with Genie team about the direct alternative
- c. Investigate need for magnetic field map within LAr volume. Design and develop the interface for such a service. -- Make this a named project. Assign architect, etc.
- d. Coordinate between GArSoft and LArSoft as needed to allow GARTPC design and development
 - i. Ensure sharing of data products and geometry (and with other detectors)
 - ii. Ensure sharing of any services, should that be needed (eg, for magnetic field)
 - iii. Possible introduction of non-uniform magnetic field to LAr volume
- e. SBND data reduction strategies.
- f. Participate in the discussion of how the DUNE data streams (SN + possibly others/all) will be handled within *art* / LArSoft. May involve moving away from root-based I/O.

- i. Sept update: In light of Fermilab Frameworks Workshop held in June 2023 and on-going work to develop a prototype framework that can meet DUNE framework requirements, we consider that this item is now out of scope and being addressed by the larger framework effort within CSAID.
 - g. Architecture work to support a common framework for data preparation
 - h. Education and assistance with optimization of critical production workflows, particularly for the SBN experiments.
 - i. Provide a solution suitable for running TextFileGen at production-scale
- 3. Architecture work to facilitate integration of machine learning into LArSoft algorithms**
Priority: per request
Possible targets include standard modules to provide translation and downsampling for image-based algorithms. A work plan item can be opened once a set of deliverables is defined.
- 4. Support for integration of community-supported event displays within art and gallery based LArSoft jobs.**
Priority: per request
- 5. Work on user support, bug fixes, emergency feature requests, software builds, releases.**
Priority: per request
Resources: SciSoft team
- 6. Other topics under consideration to work on include the list of accepted, but not assigned, redmine issues.** These can be found under [accepted redmine issues](#).
Priority: per request

Other considerations on work planning

In response to input received from the experiments, the Project will pursue the following topics as time allows:

- Improving informal channels of communication, which have suffered during the Covid era. To facilitate more efficient communication, community members are encouraged to engage directly with SciSoft team members before and after opening issue tickets. In addition, SciSoft team members understand that greater visibility at experiment software meetings would be helpful.
- Seeking information on recent LArG4 migration experiences, then using this information to enhance the available documentation.
- Seeking information on recent experiences using LArSoft on HPC resources, then using this information to create a common resource for others interested in using HPC.
- Developing a plan to ensure LArSoft is compatible with reading and writing data in HDF5 format, as well as other data formats.

- Tracking progress on DUNE and related community-based data processing framework developments.

Appendix A: Definition of priorities

The priorities listed above have the following meaning. The impacts noted are those negotiated with the experiments.

- High: the most urgent projects that have high impact to the community, and that should be fully staffed at all times until completion
- Medium: a project that has a significant impact to the community, and that should be staffed sufficiently to ensure completion within the calendar year
- Low: a project that does not have immediate impact to the community, could be rolled into the next planning period without imposing undue burdens, and should be staffed after ensuring all other obligations are met.
- Per request: The priority will be set in consultation with the requesting experiment or ticket author.

Appendix B: Major observations from one-on-one meetings with each experiment in September and October of 2023

(Topics in common across multiple experiments)

1. Common

- Event display that is useful in current environment.
- HPC.
- Faster processing.
- Event generators

2. DUNE - Met with Mike Kirby, Tom Junk, Tingjun Yang, Laura Paulucci

1. Continue with weekly releases.
2. Continue to focus on multi-threading.
3. Be engaged with LArSoft in discussion about framework and managing data flow within events.
4. Features needed for FD, eg, diff sampling frequencies per TPC (DetectorClock service). Top and bottom TPCs use different electronics (VD only? Need to clarify...)
5. Get DUNE (and other) WCT plugins into LArSoft repository, like [larpandoracontent](#).
 - a. This will be handled outside the work plan
6. Can LArSoft / art make using accelerators for WC easy?
7. Eager writing for framework instead of per-event. (Ensure that LArSoft supports the data processing model needed. This is a framework issue.)

3. ICARUS - Met with Tracy Usher and Daniele Gibin

- Event display

- LArSoft display is not useful for hand scanning, which is the current need. Titus does not work on reconstructed objects
 - Need to ensure that geometry re-factoring does not break the interface to Titus. This is C++ to python interface in gallery. Lives in icarusalg repository
 - Filtering particles that do not traverse the cryostat prior to simulation
 - This work is being handled outside the work plan.
 - Non-planar cathode geometries
 - Major problem seems to be cm scale displacements over half-meter-scale areas. Causing reco issues due to time-to-distance conversion, which scale to max drift distance. Cannot simulate either.
 - Did not discuss the need to represent finer details of cathode structure, eg, support elements.
 - CR event catalog, or infrastructure for managing CR data and simulated event data streams for overlays.
 - Need discussion with MicroBooNE to understand what code is in LArSoft vs uboonecode. Do not believe there will be a large component that falls to SciSoft.
- 4. **MicroBooNE** - Met with Herbert Greenlee.
 - a. Want assistance with respect to anything relevant to the future of MCC9 and any migrations that might be needed that stem from that. (eg, container support/education)
 - b. Event generators might be useful, but MicroBooNE is not driving requirements.
 - c. Light model. Semi-analytic ported to MCC10
- 5. **SBND** - Met with Andrzej Szalc.
 1. HPC working, esp for simulations (ie, low input data requirements)
 - a. Making POMs run for HPC campaigns would be useful
 - b. Or anything that would make workflows run faster
 - c. Any simulation job would be a reasonable prototype
 - d. They need info on using FermilCloud HPC allocation
 2. Generator integration (so completing the current Genie decoupling work item and extending to others)
 - a. Related: long term item (2): “provide a solution suitable for running TextFileGen at production scale” if direct integration is not feasible
 3. Event display
 4. Anything that allows us to process quicker (parallelization, better overlay strategy – uses DBs now??).
- 6. **SBN Data/Infrastructure** - Met with Steven Gardiner, Giuseppi Cerati
 1. Memory consumption in simulation
 - a. Currently driven by photon sim. Not sure if large N photons, or visibility service. Uses library lookup (ie, not semi-analytic)
 - b. Might address initially as LArSoft issue rather than work plan item
 2. CPU concerns in simulation
 - a. Drift simulation takes majority of time
 - b. Might be a target for GPU
 3. Overall optimization of simulation workflows needed (cf the above)
 4. Event display,
 5. Integration of ML workflows also important, but SBN has no effort to pursue

6. Generator refactoring. Steven Gardiner + PD might be available to do the work