2022 LArSoft Work Plan

Last updated: December 16, 2021

Introduction

This document gives the LArSoft work plan for 2022. 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 October of 2021. The experiments detailed their plans for the next year, the implied 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 LArSoft, not a particular experiment, 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 2022 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 2022

We define short term priorities as those items that should receive sufficient effort to be substantially completed in CY2022 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. The focus will remain on critical production workflows where the introduction of these advanced computing techniques will provide substantial added value through improved or expanded resource utilization, or gains in execution speed. A number of such targets have been identified.

   Resources: Kyle Knoepfel, Saba Sehrish, Mike Wang, in collaboration with HEPReco SciDAC

2. **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 and the cetbuildtools-based build system. 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 Feb. 23rd LArSoft Coordination. Further information on the build system can be found in the following

presentations by Chris Green: [Feb 23, 2020 LArSoft Coordination Meeting](#), [Nov 2, 2020 LArSoft Coordination Meeting](#).
https://cdcvs.fnal.gov/redmine/projects/spack-planning/issues
Resources: Patrick Gartung, Chris Green

3. **Pixel detectors within LArSoft**
   - **Re-architect geometry systems to improve integration of multiple readout schemes**
   - **Identify and implement common low-level and 3D data products needed to represent data from LArTPCs with pixel readouts**
   - **Support development of pixel-based anode simulations**

Priority: high

Work with experiments to develop requirements and design solutions, and to carry out the work necessary to adapt reconstruction and simulation codes.
https://cdcvs.fnal.gov/redmine/issues/19328
Resources: Erica Snider, Kyle Knoepfel

4. **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

5. **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. Examples include:
   - Foundational concepts, principles and best practices, e.g., reinforcing framework / algorithm separation.
   - Migrating simulations for existing detectors, and developing simulations for new detectors within the refactored LArG4 framework
   - Guidance on writing more efficient and maintainable code
   - Maintain and enhance an up-to-date introductory slide deck on LArSoft - https://indico.fnal.gov/event/49621/#1-see-attached-source-material
   - Update and add pages to the concepts section of LArSoft.org to be inline with the updated slide deck.

Continuously review and update this documentation as needed.
Resources: SciSoft team

6. **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

7. **Migrate the LArSoft wiki to a platform that is searchable via common search engines**
   Priority: high
   The migration is intended to solve wiki access problems caused by a previous move of Fermilab-hosted web pages behind the Fermilab SSO pursuant to a new DOE requirement.
   Resources: SciSoft team

2. Long term and continuing priorities for 2022

We define long term priorities as those that we do not anticipate completing within CY2022 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.

    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 October of 2021

(Topics in common across multiple experiments)

1. **DUNE** - Met with Heidi Schellman, Tingjun Yang, Michael Kirby. Tom Junk, Paul Laycock (two meetings).
   a. Being flexible about framework. DUNE has a framework task force that has started to discuss requirements for the DUNE event processing framework. While LArSoft is in principle framework independent, LArSoft should be informed of a decision to deviate from an art type framework with as much lead time as possible.
   b. Using HDF5 for raw data, and being able to write data to multiple formats.
      i. DUNE has committed that all the raw data from here on out will be written in HDF5. This fits well with existing machine learning formats.
   c. Shipping separate detector elements or "time slices" to different threads within a single readout frame. Consequently, thread safety and generally supporting access to HPC is still important.
   d. Being able to use GPU-as-a-service is a great idea.
      i. Updating TensorFlow was also a tremendous help.
      ii. Machine Learning is becoming more popular, LArSoft is making progress there.

2. **ICARUS** - Met with Tracy Usher
   a. Titus was built before there was any connection to LArSoft. Can't view reconstructed objects. Can't use services. At some point ICARUS has to either abandon Titus, or find people in ICARUS dedicated to fixing the problem correctly.
   b. Early thoughts of a large-scale hand-scanning effort for physics (as an alternative to automated reconstruction) have been abandoned.
   c. Improved machine learning integration

        i.   In the ICARUS reconstruction scheme, there are lots of noise hits, so try to filter them out. Machine learning can be used to investigate the space points. Work is ongoing to evaluate how effective this approach is. Hoping they can be competitive to the standard Pandora.

        ii.   Led by Kazu. Talk to Francois about this also.

   d. Multi-threading, particularly as part of signal processing phase

        i.   Jobs there consume 6 GB.

        ii.   Running on GPU for this might be interesting, which could make of GPUaaS .

   e. Not being able to search redmine pages easily is a huge problem.

3. **MicroBooNE** - Met with Herbert Greenlee.
   a. Concerned about Spack transition being disruptive.
   b. Wants LArsoft to do something about the root io thing for the PID. Herb will characterize the problem in current system, provide example program. Then we can seek help from root team. [This has since been solved, so the issue is moot.]
   c. Possibly interested in running Geant4 in multi-threaded mode
      i. Geant4 is thread safe and multi-threaded already. So believe we just need to configure it to run that way. Maybe ask whether it would be useful to figure out how to enable this for MicroBooNE G4 jobs

4. **SBND** - Met with Andrzej Szelc.
   a. Lessons learned about using new LArG4 from SBND. Want continued support with new LArG4.
   b. Overlays. Not a LArSoft problem per se. Need to have enough overlay data to generate the Marlo Carlo. Because of data throughput, not clear that SBND can do it the way MicroBooNE did. In SBND case, the limited data throughput capacity of DAQ means that not much can be devoted to beam-off or no-bias during data taking. Can we generate enough Monte Carlo for what we need?
   c. Andrzej will give Patrick (?) a call about their experience about running things on Theta. It's easier now than it was a year and a half ago.
      i. Would be useful to create some documentation as a community resource for this
      ii. Would like to hear about the running of data from SBND. Guiseppi et al might come up with something useful for this based on their experience.
   d. Theta workflow seems to be a good solution to running Cosmic jobs.
   e. Want something that allows TextFileGen to scale on the grid.
      i. Direct integration would be better, but isn't easy.
   f. Want thread safety work to continue.

5. **SBN Data/Infrastructure** - Met with Joseph Zennamo, Wesley Ketchum.
   a. **LArSoft communication**. With people working remotely, there aren't the hallway conversations, or the lunch-time discussion that can help enhance the communication about an issue. (Note, the regular Coordination Meetings, Offline Leads meetings and Steering Group meetings are providing opportunities for people to comment on issues or concerns with LArSoft.) LArSoft encourages discussion with a SciSoft team members before opening a ticket. It helps to have the dialogue and context going into crafting the ticket. Can get a solution more quickly this way.

       i.   Having a presence at experiments' meetings would provide another opportunity to interact with the developers. Something to consider.

b. In the migration to LArG4 they bumped into bugs with what is in there. Surprised since ProtoDune was already using it. Challenge has been some of the issues within the workflows didn't match up to the use cases they needed. Also, some confusion over whether an issue was an art issue or LArG4.

c. Kyle's presentation on optimization: needed in advance of active opt periods, roughly Nov to Feb during development and testing, and March to June during production phase

d. When the Spack transition disruption happens, we should make it very disruptive so everyone knows what is disrupting them. Suggested this would lead to less confusion overall. So managing disruption of Spack transition is important.

e. Being able to run on HPC is important

f. Integrating DL workflows

g. Data volumes are an issue. The experiments will likely need to find ways to reduce event sizes.