LArSoft Work Plan for 2017 Wrapup

by: Erica Snider and Katherine Lato
Last updated 12/6/17

Background

This material is for the 12/13/17 Steering Committee meeting.

- Material from the 2/2/17 Steering Committee meeting is available at: https://indico.fnal.gov/conferenceDisplay.py?confId=13416
- 5/2/17 version is at https://indico.fnal.gov/conferenceDisplay.py?confId=14198
- 8/31/17 version is at at https://indico.fnal.gov/conferenceDisplay.py?confId=15075

Short-term priorities; happening concurrently

The order does not imply priority.

1. Re-factoring of LArG4

o Purpose: Separate GEANT4 functionality and LArSoft-specific parts into different units
- Simplifies interface between material modeling and detector simulation
- Clarifies maintenance responsibilities
- Easier to introduce or modify models of various detector effects

o Resources: GEANT4 team: Hans Wenzel and Krzysztof Genser
http://cdcvs.fnal.gov/redmine/issues/14454 - 2/2/17 - 48 done /170 total hours
3/2/17 - 90/170
3/30/17 - worked on putting changes into GEANT4, need to test those.

4/27/17 - 124/170
6/1/17 - 124/170
6/27/17 - 148/230
7/27/17 - The original plan was to incrementally replace parts of the LArG4 code with the updated Geant 4 functionality and interfaces. But it looks like it is easier to do a full replacement. This has affected the work and schedule.
8/29/17 - 163/230, Have meeting scheduled on 8/31 to discuss this.
9/26/17 - Hans has delivered a draft data product that will store energy depositions created by Geant4 that can then be used as input for ionization / scintillation modeling, electron and photon transport modeling, etc. Need feedback from experiments. Once we have agreement on this data product, we can use it as the basis for extracting the various LArTPC simulation effects into pluggable components to be used in the replacement of LArG4. Some coordination on the relevant interfaces will be needed. This work is badly behind schedule and is threatening production schedules within ProtoDUNE and ICARUS.
10/12/17 - Hans is receiving help from Bill Seligman. Hans feels the full replacement is better than the original plan to incrementally replace parts of the LArG4 code with the updated Geant 4 functionality. He knows he needs to update the ticket. But hasn't.
11/9/17 - Have discovered more work to do it the right way than expected. The right way to do it is let Geant4 deal with the physics processes. All the physics will go to nutools.

2. DONE - Track fitting / data product improvements - phase II
o Purpose: Provide data structures to store track fit information, and re-organize existing data structures to better match algorithm workflows
- Provides the additional flexibility needed to accommodate the reconstruction steps downstream of pattern recognition
- Also working to standardize the information to be produced by any track reconstruction workflow
- This is part of a longer-term effort aimed at establishing uniform policies for the output of reconstruction

o Resources: LArSoft team: Giuseppe Cerati, Gianluca Petrillo, Erica Snider
o http://cdcvs.fnal.gov/redmine/issues/14047-
2/2/17 - 70 done / 400 total hours
3/2/17 - 130/440
4/27/17 - 130/440 - working on other things (like ProtoDUNE/ICARUS integration)
6/1/17 - 130/440 - still working on other things and vacation-time.
6/27/17- 130/440 - scheduled a meeting Giuseppe, Gianluca, Erica to discuss what should happen with this.
7/27/17 - July meeting determined to return to the original scope for phase II on the tracking, instead of trying to capture parts of phase III in this earlier work.
8/2/17 - Discussion with Erica, Gianluca and Katherine where we descoped or moved all phase 3 work that had slipped into the phase 2 work plan, and some non-track work, due to lack of resources.

1. We rejected (without prejudice, we just don't have the resources in LArSoft to do this) the following tasks:
    - https://cdcvs.fnal.gov/redmine/issues/14061 - Provide an interface to access a complete reconstruction of the event
    - https://cdcvs.fnal.gov/redmine/issues/14265 - Provide an interface for access of reconstructed cluster information

2. We moved phase 3 work into a new milestone - https://cdcvs.fnal.gov/redmine/issues/17338 - Phase 3 Tracking Data Product with a couple of subtasks.

3. Left with 200 hours as part of 14047. Phase 2 focuses on making it easy to navigate information associated to tracks where the associated data products are a direct product of the pattern recognition.
8/29/17 - 110/200
9/26/17 Done with phases 1 and 2. Phase 1 changes included adding track ft information and code to enforce policies for trajectory points, etc. Phase 2 changes introduced helpers to write tracks and enforce hit and cluster association policies, and a track "proxy" (or "facade") to hide complications of finding associated hits and clusters. The proxy concept will be extended to include a number of other associated data structures in Phase 3 which is not currently scheduled. Specification of the interfaces were presented at a LArSoft Coordination meeting in September and the code was completed.

3. DONE - ProtoDUNE / ICARUS integration
o Purpose: To provide the code and interface changes needed to extend LArSoft support to ProtoDUNE and ICARUS simulation and reconstruction Resources: Gianluca Petrillo, Erica Snider and Robert Sulej
o http://cdcvs.fnal.gov/redmine/issues/15086- 2/2/17 - not estimated yet,
3/2/17 - subtasks added,
4/27/17 9/12
6/1/17 still 9/12
6/27/17 - subtasks are done

Also, ProtoDUNE work -
https://cdcvs.fnal.gov/redmine/issues/14363 + other work
- 4/27/17 112/130 (Note, estimate doesn't include Robert Sulej's work)
- 6/1/17 112/130
- 6/27/17 - 136/136 - from the LArSoft core team. Note, there are tasks in this milestone for ProtoDUNE people to work on. These are not being estimated nor tracked by LArSoft. There may be work to support their efforts on demand. This is like other work and is tracked as it comes up.
- DONE (for LArSoft)

ICARUS work -
https://cdcvs.fnal.gov/redmine/issues/16031
4/27/17 - 8/12 hours
6/1/17 - 8/12 hours
6/27/17 - 8/12 done

Total work + lots of meetings between LArSoft people & ICARUS people. Expect estimate of effort to increase.
4/27/17 - 129/154 estimated. Estimate may still increase as full scope of work is understood, but not necessarily.
6/1/17 - in-depth conversation in person to further understanding and ran tests with Tracy Usher
6/27/17 - don't believe they are waiting for anything from the LArSoft core team at this point.
7/6/17 - requested input on https://cdcvs.fnal.gov/redmine/issues/15086 by end of July.
8/2/17 - No reply, so marked it resolved.

4. DEFERRED TO 2018 SPACK - new build system for *art* and LArSoft
o Purpose:
- To migrate to a standard set of build tools that have broad community support

- To address portability and configurability issues raised by experiments with the current build system
- To allow continuing use of Mac OSX as a LArSoft development platform within the context of Apple's System Integrity Protection system

o Resources: Jim Amundson, Patrick Gartung, Lynn Garren
o http://cdcvs.fnal.gov/redmine/issues/15313 - 2/2/17 - no estimate of hours, nor by 3/30, nor by 4/27, promised by early May, still no estimate by 6/1/17
6/27/17 Not estimated, running late.
8/30/17 Not estimated, running late, but assigned to Chris Green, expect a plan in early September.
9/27/17 - not estimated.
10/11/17 - Due to other work priorities, this effort has not yet been planned.

5. DONE - Documentation improvements such as a training page, updates to wiki pages, new LArSoft notes
o Purpose: Keep the LArSoft Collaboration aware of changes in tools and process. Highlight information produced within LArSoft Collaboration.
o Resources: Katherine Lato
o Various redmine issues throughout the year. 40 -120 hours
o As an example: http://cdcvs.fnal.gov/redmine/issues/14691 - LArSoft notes for January
3/2/17 - 9 hours done, extra hour to update redmine pages related.
3/2/17 - drafted LArSoft notes for March. https://cdcvs.fnal.gov/redmine/issues/15716
4/3/17 - https://cdcvs.fnal.gov/redmine/issues/15716 - 9 hours total done
4/3/17 - https://cdcvs.fnal.gov/redmine/issues/15816 - New document on Geometry (6/23 hours)
4/27/17 - 17/35 hours (new estimate, involving Erica & Gianluca)
6/1/17 - 29/40
6/27/17 - 29/40 - Tom Junk is reviewing material.
8/24/17 - 40/40 - done
https://cdcvs.fnal.gov/redmine/projects/larsoft/wiki/Geometry_Package
6/22/17 - new LArSoft notes on workshop with videos
6/27/17 - 12 out of 21 hours done
8/3/17 - 21 out of 21 hours done https://cdcvs.fnal.gov/redmine/issues/16933
8/15/17 - new LArSoft notes on the Geometry document
8/31/17 - 3 out of 5 hours done
9/6/17 - 5 out of 5 hours done

6. DONE Packaging NuWro. Note, a helper like GENIEhelper might require a fair amount of work and is probably not within the effort LArSoft has available.
o Purpose: To package NuWro as a UPS product that is distributed with the LArSoft suite.
- A more ambitious level of support would be to provide a direct interface to NuWro within LArSoft, much like GENIEHelper does for GENIE, but that work isn't being estimated here.

o Resources: LArSoft code management: Lynn Garren
https://cdcvs.fnal.gov/redmine/issues/15448
o Note: the original request also included integrating NuWro into GENIEhelper or a similar new class. While evaluating the costs and benefits of that part of the request, we intend to address this accessibility issue.
3/2/17 - 8 / 8 - done (no further updates on this item)

7. DONE - CI improvements

o Purpose: Migrate the LArSoft CI system to the new system and produce a second tier of CI tests.

- The new system has significantly improved reporting and email messaging that will enhance usability and allow the system to send emails with error reports directly to developers
- Support for the current version in use by LArSoft is ending

o Resources: Vito Di Benedetto and experiment contacts for the CI tests
o http://cdcvs.fnal.gov/redmine/issues/15125 - done
o http://cdcvs.fnal.gov/redmine/issues/15124 - in progress
3/2/17 - Have related tasks assigned. Anna tracking details.
4/3/17 - done with 15125
4/26/17 - ⅗ out of 15124 related task, one done, but may be another one.
6/1/17 - no estimate on second subtask in CI system. (Vito tracking.)
6/27/17 - Two subtasks in the CI system are being tracked for this item. Are at 8 hours out of 25 for both of them. (One is mostly done, the second hasn't been started.)
7/27/17 - 10/25 hours are done.
8/17/17 - work was required to get a more stable grid support. This work was not done in the LAr CI, but in the general CI infrastructure. Now the grid support is mature enough to start to implement the functionality.
8/29/17 - this remains not finished, 10/25 hours done.
9/26/17 - this remains not finished, 10/25 hours done
10/11/17 -Added the possibility to build the common base of LArSoft modules and store it as tarball on dCache, then use this tarball to build/test the experiment code in independent jobs. This way the load of the test is spread among different build nodes and the feedback about the experiment code status is much faster (The CI build time is decreased from the current ~1 hour, to about 10-15 minutes). Finished 15124, 25/25 hours are done.

Longer-term priorities

A. Concurrency -
   - Purpose:
     - To use multi-threading to address memory usage issues and provide flexibility in resource utilization
     - To introduce vectorization to those components of the code where speed improvements could be obtained, and to make use of currently unutilized resources
   - Resources: LArSoft team and (proposed) experiment effort
     - LArSoft and users need to be educated about concurrency, and the coding rules that apply to candidate code. - 7/6/17 update -- There was a session on concurrency at the 6/20/17 LArSoft Workshop. Slides and video are available at: http://larsoft.org/larsoft-workshop-on-tools-and-technologies/
     - Multi-threading will depend on support within art. LArSoft may be asked to provide realistic scenarios to test their changes.
     - MicroBooNE has already contributed effort to this project
     - 10/11/17 update - Vectorizing optimizations to identify and vectorize appropriate targets across LArSoft to take advantage of resources within our existing computing infrastructure that has not been utilized. Will be carried out by Giulherme Lima, -

https://cdcvs.fnal.gov/redmine/issues/17920 Will be talking about this in 2018 plans as well as optimization.

B. Architectural changes to define algorithm interfaces for layered algorithms.
   - Purpose: Provide a well-defined set of interfaces for layering algorithms within a framework that allows for run-time configurability below the level of art modules, e.g., the art "tool" currently under development
     - The primary target for this change are the algorithms used to generate, process and deconvolve raw signals from the detector
     - A major thrust is to provide a set of tools that will allow algorithms that are currently detector dependent to be generalized and incorporated into the core LArSoft code
   - Resources: The LArSoft team and (proposed) experiment effort.
     - DUNE has already contributed effort to a similar project
   - Note: new architectures do not need to be based around the art "tool". This is primarily about agreeing what certain interfaces look like.
   - 6/1/17 - There are efforts within DUNE and MicroBooNE to implement this strategy. See David Adams presentation on tool interfaces -- https://indico.fnal.gov/conferenceDisplay.py?confId=14491 that occurred at the LArSoft Coordination Meeting. There is a MicroBooNE one as well. LArSoft needs to get involved to ensure there is an accord across the project on interface definitions.
   - 11/15/17 - We'll examine this further in 2018.

C. TPC topology / locality in the Geometry service.
   - Purpose: To provide a mechanism to efficiently determine the TPCs adjacent to a given TPC, as would be needed to follow tracks and showers that cross TPC boundaries
   - Resources: The LArSoft team and (proposed) experiment effort
   - http://cdcvs.fnal.gov/redmine/issues/9818 -- requested by DUNE
   - 11/15/17 - asking DUNE for priority call on this.

Note:

- No effort is available to work on Event Display at this time. It is, however, a nagging issue.
- Need someone to lead an analysis on deep learning and integrating Convolutional Neural Networks (CNNs) before initiating this project. LArSoft should be available to consult, but isn't in a position to lead this effort.
  - 11/15/17 update: We should assess whether the degree to which the community believes this integration is needed. We can make arguments for why we *should* want close integration with CNN inference steps with LArSoft -- namely, in order to construct a full reconstruction workflow where CNNs are interleaved with conventional algorithms that refine CNN results.
    - The work by ProtoDUNE is already partly integrated, so there is evidence that at least part of the community supports this idea. Then it would be a matter of assessing requirements and priorities
    -
- 8/29/17 update - Paraview work has been experimental, not supported by SCD. (They don't support any event display.) Someone worked over the summer on a Paraview driver in Geant, much faster, almost done. Last summer's work laid the groundwork. But won't get additional effort for that. Need someone from an experiment to take it on.

- 11/15/17 update - may want to use Qt to rewrite Event Display. Note this needs a Qt build on osx for development purposes with LArSOft providing a framework (Qt distribution, basic interface to art run control, interface to geometry).

Work began in 2016, continuing into 2017

Here are the LArSoft assigned projects actively being worked on:

1. http://cdcvs.fnal.gov/redmine/issues/14454 - Refactoring LArG4 - covered in item 1 above
2. DONE http://cdcvs.fnal.gov/redmine/issues/14047 - Data product review ("phase II") - covered in item 2 above
3. LArSoft Part is DONE http://cdcvs.fnal.gov/redmine/issues/14363 - Support detectors with drift direction different than x axis (ProtoDUNE)
    - 6/1/17 update - 112
    - 8/29/17 update - 136 hours spent by Gianluca, more tasks assigned (to non-LArSoft core team members) for this work.
4. DONE http://cdcvs.fnal.gov/redmine/issues/11994 - Enable use of Kalman Filter as final fitter
    - As of 3/30/17 What is left is to provide the documentation needed.

5. DONE http://cdcvs.fnal.gov/redmine/issues/12237 - Usability: add facilities for better/easier use of associations

6. DONE http://cdcvs.fnal.gov/redmine/issues/14048- Provide a uniform recommendation for physic vector data structures in LArSoft data products - part of item 2 above

Other work:

- User support
- Continue to have code analysis
- Continuous architecture improvement
- Yearly workshop. Concurrency - multi-threading and vectorization options

Other Topics

Other topics under consideration to work on include the list of accepted, but not assigned, redmine issues. These issues originated from a combination of problems identified by the experiments and those identified by the core LArSoft team.

1. https://cdcvs.fnal.gov/redmine/issues/13711 Concurrency: Explore multi-threading and vectorization options - assigned to Erica on 2/20/17 - part of long term priorities. Resolved, off this list.
2. https://cdcvs.fnal.gov/redmine/issues/9818 Add knowledge of TPC topology in GeometryCore - Assigned to Gianluca, and made a milestone, so it's off this list.
3. https://cdcvs.fnal.gov/redmine/issues/12778 LArSoft needs an error handling policy - Assigned to Katherine, starting work on this. Will propose as a 2018 priority.
4. https://cdcvs.fnal.gov/redmine/issues/13563 Obsolete LArSoft code - staying in accepted, but not assigned state due to resource lack, will be on 2018 list.

5. https://cdcvs.fnal.gov/redmine/issues/12785 review typedef cryo_tpc_view_hitmap - will be on 2018 list
6. https://cdcvs.fnal.gov/redmine/issues/12602 Add a compact representation for the plane ID, TPC ID and cryostat ID - will be on 2018 list
7. https://cdcvs.fnal.gov/redmine/issues/11872 - Establish a standard way to represent particle identification results will be on 2018 list
8. https://cdcvs.fnal.gov/redmine/issues/11871 - Decay products in pandora do not start from the same vertex, while Projection Matching Algorithm's do will be on 2018 list
9. https://cdcvs.fnal.gov/redmine/issues/11066 - Load geometry from ROOT file instead of GDML file will be on 2018 list
10. https://cdcvs.fnal.gov/redmine/issues/10937 - Remove references to AuxDetGeo objects from Geometry service and related classes - assigned to Saba, off the list

Topic for discussion

Infrastructure that benefits a lot of people could be the target for collaboration effort. This may be a big thing that needs lots of people, or it could be smaller but the core project can't develop it, so experiments need to contribute. Discuss how experiments might contribute on things that have broad application or impact across the community, but that are not typically or easily pursued by any single experiment.