

LBNE Software and Computing Requirements

Brett Viren

Physics Department



November 13, 2013

Contents


Overview

Document coordinates

Document Organization

Existing Requirements Run Through

Summary and Future Work

(Note: in this talk, requirements pertaining to the use of the art software framework and related Fermilab S&C infrastructure will be marked with: )

Overview of Requirements

The requirements are:

- A list of LBNE needs pertaining to software and computing.
- Meant to focus on expected needs and problems and **not** the solutions themselves
 - no technology choices!
 - stay meta!
- Developed by the S&C group.
- Defines, to some extent, the S&C group's responsibilities.
- Needs acceptance by the larger collaboration.
- Will be used to justify high-level decisions.

Document Coordinates

DocDB Periodic releases to 8035

redmine [https://cdcvs.fnal.gov/redmine/projects/
lbne-software-and-computing/wiki](https://cdcvs.fnal.gov/redmine/projects/lbne-software-and-computing/wiki)

git/ro [http://cdcvs.fnal.gov/projects/
lbne-software-and-computing-requirements](http://cdcvs.fnal.gov/projects/lbne-software-and-computing-requirements)

live <http://www.phy.bnl.gov/~bviren/lbne/sc/requirements/>

Document Organization

Currently:

- Initial section meta-info about the document itself.
- Then, one section per category of requirements
- A sub-section per requirement
- Requirement stated as a brief assertion
- Ancillary / optional content
 - Use cases** Descriptions of any use cases that drive the choice for the requirement.
 - Justifications** Explanations as to why the requirement was chosen.
 - Accepted risks** Any known problems with the requirement which are considered acceptable.
 - Alternatives** Any alternatives that were considered with reasons for rejection.
 - Deviations** Recognized cases where deviations do not violate the intention.

Structure change?

Current structure largely invented. The LBNE project has its own:

- DocDB 3472
- Uses a **spreadsheet** instead of sectioned document.
- Defines a requirements **hierarchy**: lower-level referencing higher level
- Each requirement includes **rationale** and **verification** methods.

High-level requirements are mean to be “fluffy” and inarguable:

The software should be user friendly.

- If a low-level requirement can not derive from a high-level one we question it's usefulness (or make up a new “fluffy” one). Some existing ones are already “fluffy” but most are already lower-level.
- I'd rather stress verification after acceptance of a requirement so as not to bias a requirement to make the verification easier.
- I'm thinking to move the document to this organization (if not file format).

Thoughts on making this change?

Existing Requirements Run Through

Group Organization

Data

Database

Software

Validation

Continuous Integration

Grid

Workload Management

Geometry

Visualization

Networking

Online

Issue tracking

Conferencing

Group Organization - Technical Advisory Committee

A Software and Computing Technical Advisory Committee (S&C TAC) shall be formed. It will consist of five members. Its purpose is to advise the LBNE collaboration spokespeople on technical matters regarding collaboration software and computing. The committee is to proactively identify, examine and select technologies pertinent to LBNE software and computing. Interfaces with software and computing elements managed by the LBNE Project shall be considered in committee recommendations. Recommendations by the TAC will be reached through committee consensus.

Group Organization - Consensus

Important decisions will be made by consensus. Consensus is to be formed by discussions within the S&C group. If this can not be reached then the S&C TAC will be consulted for advice.

Data

Data Access

Access to raw or processed data by collaborators will not be limited by policy. Technical implementations must not unduly restrict access by any collaborator.

Simulated Data Production

All simulated data must be generated through official production processing methods or otherwise by agreement with the S&C group in order for the group to consider it representing collaboration results.

Distribution of full copies

Three copies of the raw data will be maintained at unique locations.

Raw data storage reliability

A host of an official copy of the raw data will employ storage technology with an expected loss of no more than 0.00XXX% per year.

Database

Unique Master Database

There must be exactly one master for any given database across the collaboration.

Slave Database Provenance

A slave database must be replicated from a master (or another slave) and must not be otherwise modified.

Master availability

All master databases must be available for replication by any collaborator.

Application-specific databases

An application may have a database which is not required to be distributed. Such databases are exempt from master/slave distribution as determined by consensus.

Database - Master Database Adoption

Master Database Adoption

Requirements and implementation for new master databases shall be discussed with the S&C group and may be subject to recommendations by the S&C TAC if needed. Discussions shall take into consideration expertise of and ongoing support by the database owner, appropriateness and long-term feasibility of the database technology and integration of the database into the collaboration-wide database replication requirements and existing implementations.

Software - 1

External Software Provenance

The group shall assure that all third party software available in source form can be compiled from pristine upstream releases and any applied patches are taken from trusted sources.

Source Code Archival

A copy of the source code used to build external and LBNE software for official releases will be archived in order to allow for reproducing results at a later time.

Reliance on Third Party Binaries

The group shall minimize the reliance on external software which is available only in binary form.

Software - 2

Supported Platforms

The S&C group shall develop a policy for supported platforms. This shall include the mechanisms by which platforms are selected for support, retired from being supported and what level of support and by what manner it will be provided.

Supported Software Versions

The S&C group determines the supported versions of software.

Software Releases

The S&C group must develop a release mechanism and enact periodic release for LBNE software and externals.

Software - 3

Software Installation from Source

The S&C group must supply an installation mechanism for LBNE software and externals which can be run in an automated fashion by collaborators with limited but a reasonable amount of expertise.

Software Binaries

The group must provide packages in a form ready to run (ie, executable binaries, libraries, properly placed Python modules or scripts) and a mechanism to install them in a user friendly manner.

Long term maintenance

The S&C group shall produce and enforce a policy on long term software maintenance which addresses the eventual loss of software developers from the collaboration.

Software - 4

Version Control System

All LBNE software source code shall be maintained in a version control system.

Coding Standards

The S&C group shall develop and document a set of coding standards to be followed by all experiment software.

Requirements for a process for initiating new LBNE software packages

The S&C group shall develop a process by which developers of new LBNE software packages should follow in order to integrate the package into the greater code base.

Validation - 1

Valid compilation

The source code for experiment and externals must compile.

Valid running

Newly compiled binary programs must continue to run as they did without error.

Reproducible Results

Results from prior runs of the same program must be reproducible.

Benchmark Validation Results

When software modifications lead to novel results the change must be investigated and either the software regression must be removed or the novel result become the target for future reproduction.

Validation - 2

Validation Platform Coverage

The software must be valid on all supported platforms.

Unit test harness

The group shall select a suitable test harness to form the basis of unit tests. Working with issue tracking, continuous integration shall be considered.

Unit Testing Coverage

Every experiment software package must provide tests of the functionality it provides (unit tests).

Validation - 3

Unit Test Integration

Unit tests must be exercised by the Continuous Integration System.

Issue resolution

When an issue is reported and fix the fix shall be tested with a unit test.

Continuous Integration

Continuous Integration System Existence

The S&C group shall set up and maintain a continuous integration (CI) system which promptly tests recent commits to LBNE software for proper compilation and Physics based validation.

Continuous Integration Reporting

A centralized part of the CI system shall collect results from all build elements and present and report (via email at least) the success and failure of the elements.

Continuous Integration Platform Coverage

The group shall assure that all supported platforms have representation in the CI system and that additional platforms managed in an unofficial manner may participate.


Grid

CURRENTLY EMPTY

Listed notes:

- Focus only on OpenScience Grid?
- What requirements on the software stack?
- Security and certificates?
- We have LBNE OSG VO already, what responsibilities does that entail to member institutions?
- Grid node memory or other resource issues.

A fluffly here: “LBNE shall exploit Grid resources”?

(may have connections with: )

Workload Management

CURRENTLY EMPTY

Listed notes:

- Resource management and brokerage procedures for user/group driven jobs and for production processing (who gets what where)
- Mechanism for analysis groups requesting/driving large scale production runs
- Monitoring system needs: correct execution, usage, debugging failure, bottlenecks due to sub-optimal submission pattern

(may have connections with: )

Geometry

Geometry management(?)

The S&C group shall develop a central geometry management system from which specific applications may derive their geometry information.

Geometry scope

The geometry management system must provide geometry information for near, far and prototype detectors, for beamlines, for overburden and other geometry information used for collaboration results.

Geometry Versions and Provenance

The geometry management system must support the storage, tracking and updating of different versions of any set of geometry information.

Visualization

Visualization tools(?)

The S&C group shall provide applications and supporting tools to visualize geometry, event data and simulated "truth" information.

Visualization Accessibility

The S&C group shall assure that visualization methods adequate for both end-users who may lack programming expertise (for example for hand scanning studies) as well as for programmers requiring fast feedback during algorithm development.

Networking

CURRENTLY EMPTY

Listed notes:


- Far site bandwidth.

Online

CURRENTLY EMPTY

Listed notes:

- provenance of data
- schema version changes

(due to use of artdaq: )

Issue Tracking

Issue tracker existence

Experiment software projects will use an issue tracker to record the status of issues.

Issue reporting

Any significant report made by individuals in some forum outside of the project's issue tracker will be added to the issue tracker by a member of the project.

Issue reproducibility

Package maintainers shall produce a unit test that reproduces the issue by failing and that succeeds after the issue is resolved.

Conferencing - 1

Basic functionality for conferencing

Collaboration conferencing systems must support audio, video and desktop sharing with any client software must be cross-platform (ie, Linux, Mac OS X, MS Windows).

Collaboration conferencing systems must support audio, video and desktop sharing with any client software must be cross-platform (ie, Linux, Mac OS X, MS Windows).

Conferencing connectivity.

Collaboration conference systems must be accessible both via Internet connected computers and telephone.

Cost model

The cost model must include no direct immediate cost to participants.

Conferencing - 2

Conference scheduling

Both scheduled and ad-hoc conferences must be supported.

Conference size

The system must be able to handle 100 participants spread over as many as 10 simultaneous conferences. (numbers are wild guesses)

Conference chairing

Conferences chairpersons must have remote mute capability to suppress sources of echo.

Summary and future work

Summary:

- An LBNE S&C requirements document is taking shape.
- Received good, initial input from S&C members.
- Most categories identified, many with some initial requirements.
 - I guess we have 30-40% of needed coverage.

Near-future work:

- Plan to emulate LBNE Project requirements structures
 - incorporate hierarchy, justification and verification
- Gap filling, substantially more coverage is needed!
 - Focus on just requirements, not specific solutions or technologies.
 - Propose brief assertions, not vague wondering paragraph.
- When S&C group is comfortable: present requirements to the collaboration for review with some acceptance criteria.

Longer term:

- Revisit and revise as needed.