Software and Computing
Overview

Maxim Potekhin

BNL

potekhin@bnl.gov

Intro

▸ Since the last report in August 2014, we followed the directions outlined in:

https://indico.bnl.gov/conferenceDisplay.py?confId=858

…and progress has been made.

▸ Some of the topics for today:

  ▸ Computing Model

  ▸ Software infrastructure

  ▸ Distributed computing

  ▸ Geometry

  ▸ Collaborative Tools

  ▸ 35t/DAQ

Computing Model: a Summary

▶ In early summer of 2014, we were planning to complete a draft of the Computing Model by the end of 2014, and also prepare a companion document termed "Implementation Plan" (since the review panel indicated they wanted a document of this kind). Together with the existing "Requirements" this would constitute a "Technical Design Report".

▶ Progress has been made with:

  ▸ Integrating the Requirements into the Model

  ▸ Including the LHC experience with new computing architectures as a section

  ▸ Getting a handle on the metrics e.g. data rates and volumes (there is now a spreadsheet to aggregate these data).

▶ Still work in progress and needs a lot of development, and…

▶ The extent of "reformulation" of LBNE/F became clear in late summer, making the "Implementation Plan" somewhat premature, so the proposal is to drop it for now and finish the Computing Model asap – and use it to inform LBNF.

▶ More accent will be made on the agile, distributed and international characteristics of the LBNF computing infrastructure.

Software Infrastructure

▸ "Portable Build" – with *worch* build orchestration largely complete (Brett), Ben Morgan has now achieved success with building the *art* framework software using CMake, so we are closer to the target.

▸ For better efficiency, we want to get a degree of acceptance of these more conventional and general build methods by FNAL SCD.

▸ Next major milestone is building LArSoft (ETA late 2014 to early 2015).

▸ A software workshop at FNAL is being planned in December 2014, would ideally like to co-locate with the *art* workshop around the same time.

▸ D. Adams made useful contributions and developed scripts to make it easier for the user to manage the environment, working on release management policies. David has also been raising two other important issues

  ▸ factoring the monolithic "lbnecode" into smaller packages.

  ▸ stripping UPS and rigid versioned dependency information from the source repository.

▸ Ongoing discussions of the purpose and format of unit tests of "lbnecode".

Distributed Computing

▸ OSG once again was a reliable partner. After a period of successful running, LBNE usage of distributed Grid resources (i.e. outside of FNAL) managed by the Open Science Grid has ceased due to insufficient demand.

▸ This may change due to some international groups showing interest in resource sharing (cf. Brazil).

▸ RACF helped us reconfigure the local IF cluster and set up a xrootd connection to the "xrtood door" to dCache at FNAL, which means there is now a transparent way to access the remote data – it is staged to BNL transparently to the user on demand. Tested from my account. If and when we get the global redirector in place, this configuration can be cloned to other sites with relative ease, fulfilling the shared data aspect of the Computing Model. Another work item – bidirectional access, e.g. write data to FNAL.

▸ Initial plans are being made for the potential "IF" cluster at BNL (e.g. it will be more optimal to configure disk-heavy nodes rather than CPU-heavy nodes, since effective sharing of data will be more important that sheer computing power, case in point – 35t data).

▸

Geometry

‣ **See Brett's slides**.

‣ Really complex and important work area.

‣ It has always been on our top priority list but unfortunately its real importance was not appreciated in the physics tools group until recently, hence insufficient effort available in the past year (with Brett heroically stepping in as usual).
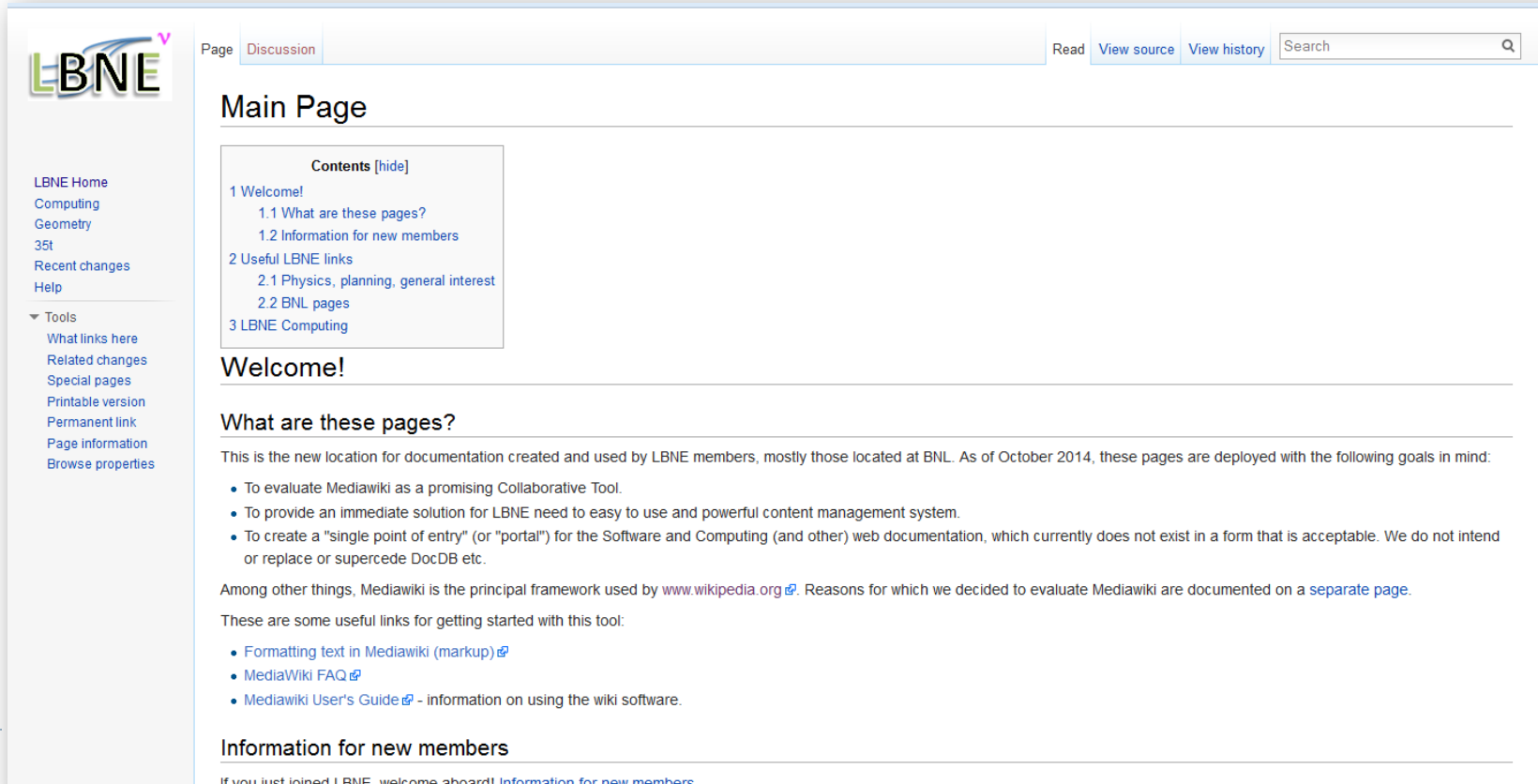
Collaborative Tools: Overview

▸ The need for quality and easy to use documentation has always been apparent to most of us. There is room for improvement (lots of it) in LBNE…

▸ In HEP, the following core functionality is typically provided by a set of tools:

  ▸ Document database (cf. CERN **CDS**, FNAL **DocDB**)

  ▸ Meeting Agenda Management (cf. CERN **Indico**, FNAL **DocDB**)

  ▸ Web Content Management Service/Framework (cf. CERN **TWiki**, FNAL **Redmine**)

▸ …also, LBNE project is using Microsoft Sharepoint.

▸ In the past year and a half, our experience with these tools was as follows:

  ▸ DocDB gets the job done, for document storage and retrieval (with some search capability).

  ▸ DocDB agenda management is cumbersome but can be made to work. Indico is a breeze compared to this part of DocDB functionality and is used to cover our Software and Computing needs (both at FNAL and BNL).

  ▸ **Indico** can still provide reference to DocDB when needed, but is more appropriate for handling more *transient material* (ad hoc slides and notes, etc). IMHO keeping all slides of temporary value in DocDB just clutters the latter and creates the illusion of having adequate documentation.

  ▸ **Redmine** - see next slide.

Collaborative Tools: Wiki

▸ A good Web Content Management System (CMS) is key to collaborative process.

▸ Wiki is just one type/concept of CMS.

▸ "**TWiki** " (one of the flavors of Wiki) is the central collaboration platform at CERN and by extension at many of its member institutions.

▸ "**Mediawiki**" is an equally widespread platform and is the engine behind Wikipedia.

▸ Lots of powerful functionality in both TWiki and Mediawiki. The latter successfully used at BNL.

▸ The current solution at FNAL (**Redmine**) has a different focus which is project management. This has implications for organization of material, navigation and look and feel, which I believe make it less than optimal. It also does not enjoy as wide installation base and community support when compared to the top CMS/Wikis. There are occasional performance issues.

▸ In moving towards a "more international" collaboration, it pays to use tools which are more widespread and have demonstrated scalability.

▸

Collaborative Tools: Evaluating Mediawiki

▸ Initial experience with Mediawiki is positive. Prototype pages are located at https://lbne.bnl.gov/wiki/Main_Page. Brett has set up the application and Maxim has started migrating/adding material. Work in progress. Contributions, suggestions, critique are all welcome!

Collaborative Tools:
Short Term Plans

- We plan to use Mediawiki to facilitate the development and management of S&C documentation (and any other documentation is the users agree) with the following in mind:
  - State of the art, widely available and popular software
  - Acceptable to new and/or international users
  - Appropriate "landing page" or "portal" for the S&C (and any other, as needed) documentation with the exception of the Project which is invested in Sharepoint.
  - Urgent need to have a good platform for the 35t documentation within S&C.

- The Sharepoint page is currently listed as the Software and Computing navigation page and this will change since it only serves as a link aggregator now and does not provide much value for S&C otherwise, i.e. no meaningful S&C content is managed there. It looks unattractive, the technology is unfamiliar to many and may create a PR issue with our international colleagues.

- We won't decree the use of Mediawiki for LBNE as a whole or for the S&C/PT, but will do our best to manage the material here as well as we can so people will hopefully appreciate the value it brings to LBNE.

Collaborative Tools: Sharepoint Page (deprecated)

Documentation for 35t

▸ One of motivations for the current migration to Mediawiki was to facilitate the development and management of information pertaining to the 35t (e.g. DAQ and its interface with off-line).

▸ Until recently, there was a significant amount of pretty complex information dispersed in slides, notes or communicated verbally during meetings. I found this unsatisfactory. Also missing was a compelling narrative of what is done for what reasons and how. Sometimes it is just assumed everyone knows reasons for complex decisions.

▸ One important issue is compatibility of existing frameworks with the type of data we need to handle (e.g. we are often in a situations different from the typical "collider trigger").

▸ Some existing important information:

  ▸ Run Modes

  ▸ links to recent documents

  ▸ Issues and questions that still need to be answered

▸ …is collected on the new page: https://lbne.bnl.gov/wiki/35t_Test

▸ The idea is to capture all of the info available and if needed, develop documents ourselves if they are not forthcoming from DAQ.

▸ To illustrate the current situation, let's take a look at the following slide…

35t: questions yet to be answered (Brett)

▸ what is the low level DAQ file format/encoding (ROOT? binary?)

▸ what schema does the content of DAQ files follow?

▸ where is a diagram showing all the parts of the DAQ data stream with their names (millislice, microblock, tick, "trigger", etc)?

▸ what is the start/stop criteria that defines the highest level "chunk" of data (e.g. a "trigger")?

▸ how does this "chunk" correspond to a trigger for each expected trigger criteria?

▸ what time-ordering is expected from data coming out of the DAQ, particularly between disparate sources (e.g. Wires/PDs)?

▸ what unit of data "chunk" ("event") is required and desired for offline analysis?

▸ what offline analysis decisions/limitations can we, as a collaboration, be comfortable "baking in" to this choice of unit?

▸ how are DAQ data "chunks" (triggers) numbered? How are offline data "chunks" ("events") numbered?

▸ what is the mapping from the former to the latter?

▸ what changes to art and/or larsoft are needed to accommodate the above answers?

▸

AOB

▸ Forum for Computational Excellence (Brett, Maxim, Torre) – DOE/US based initiative – meetings are being held…

▸ …and its coexistence and collaboration with the "HEP Software Foundation" (Torre) – CERN/Global initiative – a meeting at SLAC immediately preceding the LBNE Collaboration Meeting (Liz will attend)