OSG Council Meeting
Apr 10, 2014

LBNE and OSG

Maxim Potekhin
BNL
potekhin@bnl.gov

Overview

- LBNE quick intro
- Structure and status of LBNE Software and Computing Effort
- Collaboration between LBNE and the Open Science Grid

What are the main goals of LBNE?

- Precision measurement of neutrino oscillation parameters: will shed light on charge-parity violation and matter-antimatter asymmetry

- Search for nucleon decay (Grand Unified Theories)

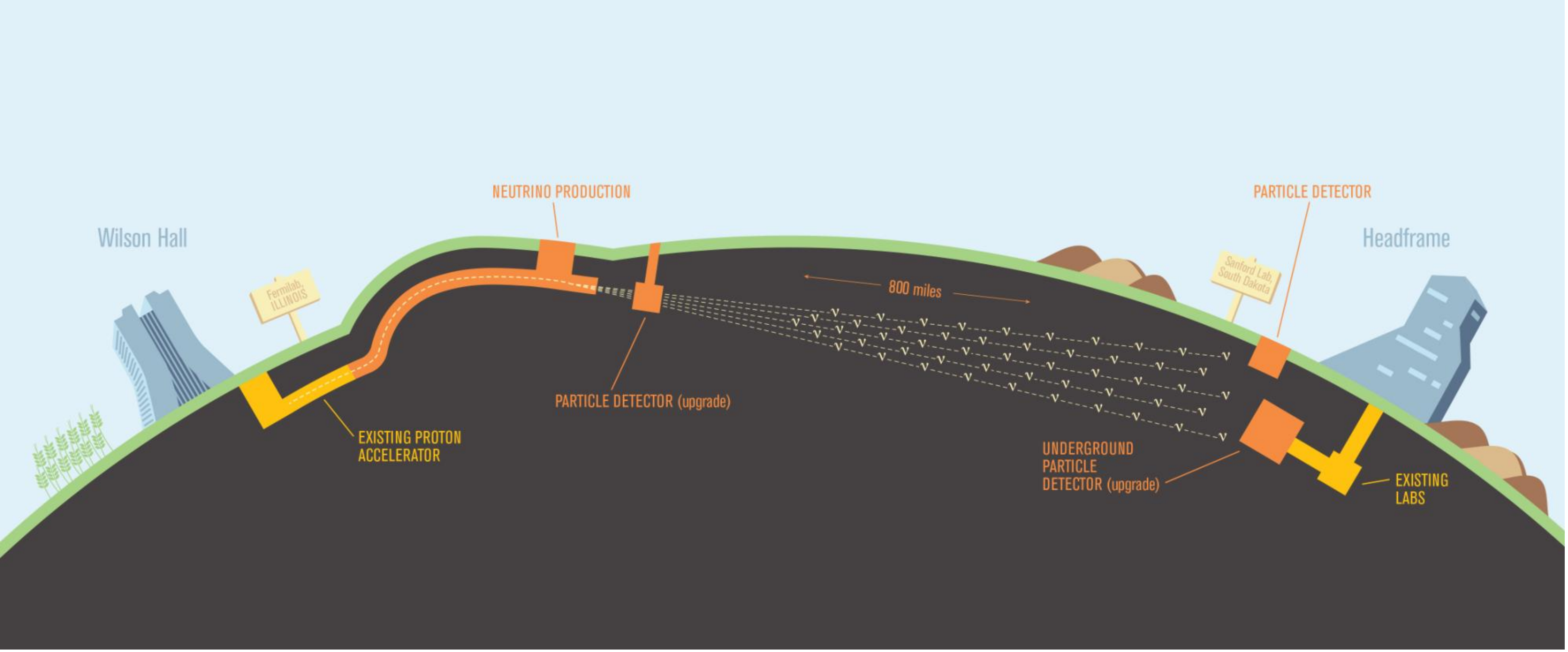
- Study of supernova bursts by detecting their neutrino signal

What components of LBNE?

To achieve its ambitious physics objectives as a world-class facility, LBNE has been conceived around three central components:

1. an intense, wide-band neutrino beam
2. a fine-grained near neutrino detector just downstream of the neutrino source
3. a massive liquid argon time-projection chamber (LArTPC) deployed as a far neutrino detector deep underground, 1,300 km downstream; this distance between the neutrino source and far detector — the baseline — is measured along the line of travel through the Earth

LBNE – the concept

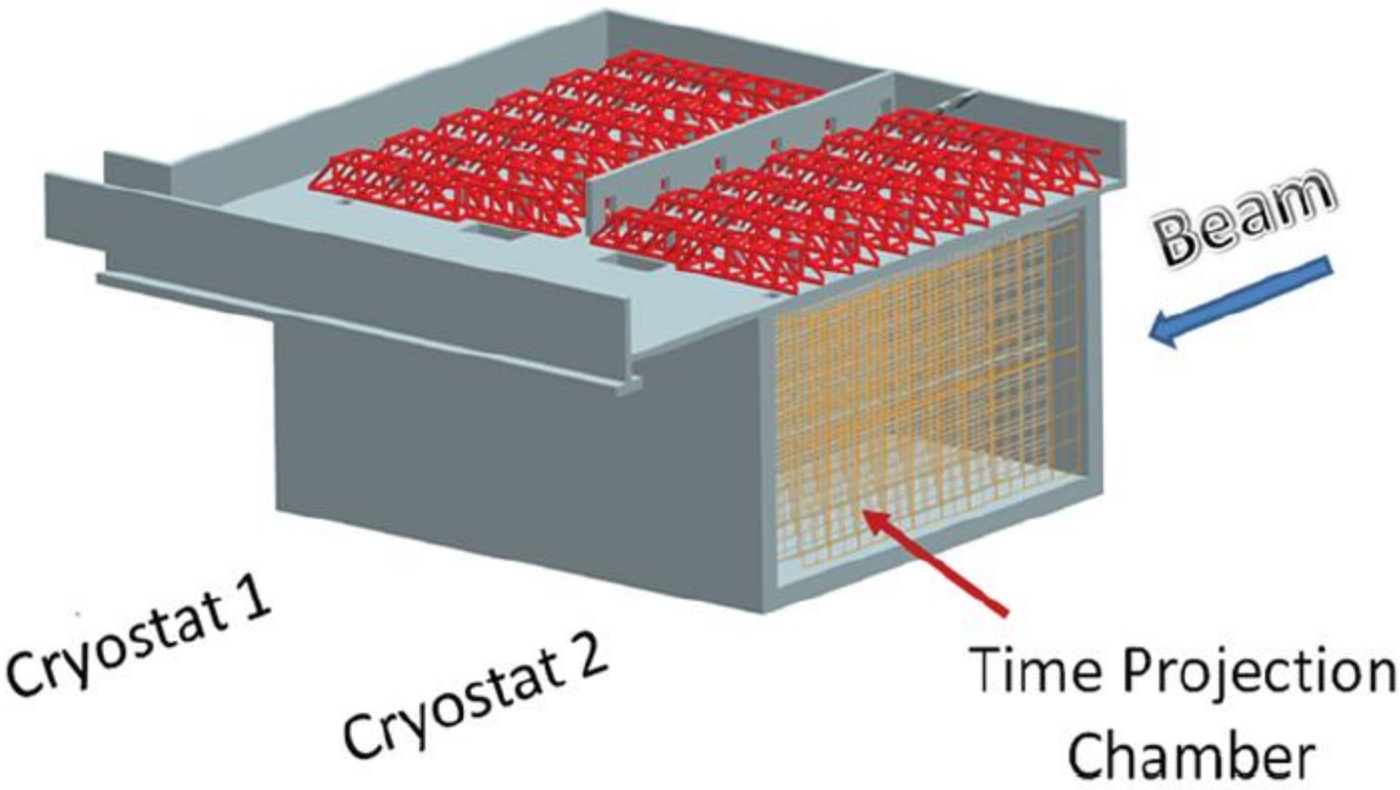


LBNE Detectors

The liquid argon TPC far detector technology combines fine-grained tracking with total absorption calorimetry. Installed 4,850 ft underground to minimize backgrounds, this detector will be a powerful tool for long-baseline neutrino oscillation physics and underground physics such as proton decay, supernova neutrinos and atmospheric neutrinos. The far detector design is scalable and flexible, allowing for a phased approach, with an initial fiducial mass of at least 10 kt and a final configuration of at least 34 kt.

A high-precision near detector is planned as a separate facility allowing maximal flexibility in phasing and deployment.

LBNE: Liquid Argon TPC



LBNE vs a HEP Collider Experiment

HEP Collider Experiment: HD Flash Photography. Events are built based on a trigger decision.

LBNE: HD Video. The detector is live and streaming data most of the time. There is some triggering mechanisms that can be used (and self-triggering) but there are physics objectives that require constant digitization – during the spill and out of spill as well.

Until recently, two possibilities were considered – surface placement of the detector vs the deep placement option. This obviously has dramatic effect on background conditions and data rates. Currently the preference is given to the deep underground configuration.

LBNE readout and data rates

Part of front-end electronics will be placed inside the Lar volume. Total readout channel count ~$2*10^5$ – of course there is multiplexing, zero suppression and compression of the data.

Data is dominated by background by a very large margin.

By placing the detector at the deep underground, it became possible to cut the overall Far Detector data rates from a few 100 MB/s down to ~100MB/s (rough estimates).

Data collection

LBNE will collect information from beam monitors, the Near Detector (which is only in the initial planning stage), and the Far Detector – Liquid Argon TPC.

The Far Detector data will be buffered at the South Dakota location and transmitted to the primary storage facility at FNAL over the network – and after QA the data is placed in storage and becomes the responsibility of LBNE offline processing team; the data can then be deleted from the buffer.

LBNE Software and Computing Effort in 2012

LBNE started with leveraging software and experience from other Intensity Frontier experiments and support that exists at FNAL.

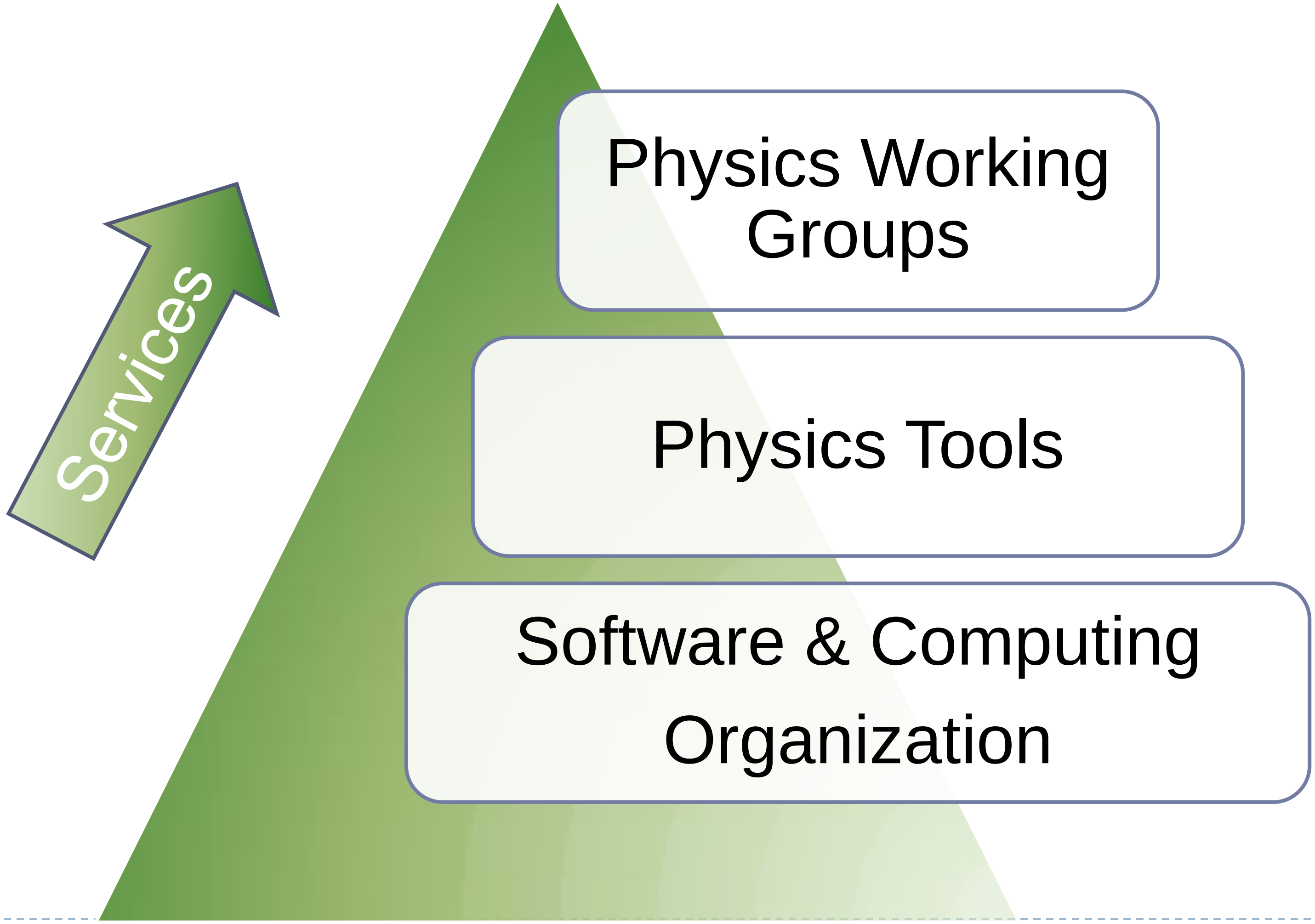
The Physics Working Groups were formed to cover the essential science deliverables needed by the LBNE project, e.g. beam simulations (target R&D), TPC (tracking and pattern recognition), Cosmogenic backgrounds etc.

Initially, most of the work (with some exceptions) was done inside the FNAL perimeter and while thought was given to a uniform and "portable" infrastructure and software management, there was little effort available to pursue that.

The Three Layer Cake (concept)

- The entirety of the LBNE Software effort is complex
- We need to control this complexity by identifying principal segments of our software and computing activity and establishing relationships among these segments/silos
- We use the concept of Service Organization which is successfully applied in industry
- …and so we arrived at the concept of the Three Layer Cake:
  - Physics Working Groups
  - Physics Tools
  - Software and Computing Organization

Three Layer Cake (hierarchy)



S&C Technology Areas – Overview

- Software Frameworks
- Distributed Code Management
- Build, Testing, Validation and Distribution
- Simulation Tools
- Event Display
- Geometry Model and Description
- Databases
- Messaging
- Data Storage, Access and Management
- Grid Tools and Distributed Computing
- Workload and Workflow Management
- Networks
- Information Services and Web Frameworks
- Security
- Collaborative Tools

LBNE Software and Computing Effort in 2013

The Collaboration rapidly grew and so did its needs in the area of software and computing. Many institutions joined, both from the US and abroad.

What was missing?

- reliable build procedures for most components, on supported platforms anywhere
- true Grid capability at the transparency and scale typical of a modern HEP experiment
- distributed data management
- unified and version-controlled geometry description
- … a lot of other stuff.

LBNE Software and Computing Effort in early 2014

- Proper organizational structure has been put in place.
- A document outlining our technology roadmap was delivered to DOE in September 2013 and the briefing was successful. Question of OSG participation was raised
- A lot of progress has been made in the software infrastructure area, where we are now close to having a system for building software which is documented, reliable and truly portable.
- We have had an initial success in running a demo on the Open Science Grid resources with realistic Monte Carlo payloads
- Right now we are working around the clock to prepare for the major DOE review in mid-May, and this includes both documentation and the technology demonstration…
- …and the 35t prototype detector will come on line later in the year, posing its own set of challenges

LBNE Software and Computing: some numbers

- The 35t prototype will generate about 50 to 100TB of data in a few weeks, which will need to be processed and reprocessed multiple times; the format and other characteristics of these data will differ form the eventual solutions we'll have to choose
- The full detector (ETA 2024) will end up producing data *almost* on the scale comparable to ATLAS, in the PB range (TBD)
- Current usage in each segment of the Physics Tools Group ranges from a few dozen batch slots to a few hundred, and from a few TB of storage to a few hundred – at FNAL
- Our assessment of future needs of LBNE indicate that we'll need more resources in future (as early as this summer) and we need to find ways to scale out – sometimes we need resources on short notice and for example yesterday we used a few thousand CPU hours on a single simulation that was needed right away.

Why is LBNE interested in working with the Open Science Grid?

- OSG has proven instrumental in the success of major scientific efforts of the past decades, and notably RHIC and LHC experiments
- The software stack maintained and packaged by the Open Science Grid forms the foundation of the infrastructure for most of HEP and IF research in the US – and is also used outside of HEP
- OSG has developed scalable ways to federate globally distributed resources of organizations of different sizes and capabilities
- There has been a focused effort to maintain robust, state-of-the-art practices in software build and validation in the Open Science Grid, which gives us confidence in the quality of the software stack
- A computing site which is not Grid-aware can be put on the Grid rather quickly by using the OSG software stack and maybe a little bit of help from OSG personnel

Why is LBNE interested in working with the Open Science Grid? (cont'd)

- OSG has moved into promising technology areas e.g. CVMFS, which is well aligned with the plans and strategy of LBNE
- OSG has substantial expertise in tools for distributed data access, which will be absolutely crucial for LBNE
- Combination of distributed software, distributed data and distributed computing resources will allow LBNE to meet its principal challenges, such as:
  - Federation of a variety of resources managed by the Collaboration members – there are more than 80 participating institutions in LBNE right now and some can make meaningful infrastructure contributions
  - We need to provide a software environment to the LBNE researchers which allows them to be optimally productive in both production and analysis scenarios

Summary

- Many of the challenges and planned deliverables of LBNE Software and Computing Effort can be met in an efficient manner if we collaborate with the Open Science Grid Consortium
- We are deferring some technology choices (such as the Workload Management System) until a later point in time, and the tools supplied by OSG for submission and management of payload on the Grid will play an important role in immediate future
- OSG experts (thanks Tanya and Marko!) are already helping us on a very technical level and this cooperation has already proven successful.
- Statements to this effect have been included in our documentation package (the current development version) for the DOE review.