# A Design for a Deep Underground Single-Phase Liquid Argon Time Projection Chamber for Neutrino Physics and Astrophysics

March 30, 2015

1

# Contents

1

# List of Figures

1

# List of Tables

1

# Todo list

# Chapter 1

# Data Acquisition

ch:trig

The data acquisiton (DAQ) subsystem provides the data collection in a robust fashion and optimized for the specific needs of the

long-baseline?

neutrino and underground physics of the experiment. The scope includes design, procurement, fabrication, testing, delivery and installation of a combination of custom and commercial electronics modules, (including commodity computing and networking hardware), as well as both commercial and internally developed software.

could we list out the components here?

The LBNE DAQ must collect data from (a) interactions which are associated with the beam, (b) high energy (> 100 MeV) interactions which are asynchronous such as atmospheric neutrinos, proton decay, cosmic ray muons etc., and (c) low energy interactions such as those from diffuse supernova neutrinos or a possible supernova explosion in our galaxy. For case (a) and (b), it is essential to record the maximum information about each event — the design presented allows collection of all the LArTPC data with no zero suppression whatsoever. For supernova physics, it is essential to have the maximum possible uptime, to be sure data is being collected when a supernova occurs. The sensitive time needed for a supernova is at least 10 seconds, during which the data flow will increase dramatically, and zero suppression of some form will be required. The design follows many of the principles of other neutrino and collider experiments with scope for a multi-level trigger and continuous readout to cope with events, such as atmospheric neutrinos, that are asynchronous with the beam.

An innovation of the LBNE DAQ is a modular DAQ design, as described in Section sec:daq upper 1.12, which extends the modularity of the LBNE far detector design and facilitates both staging and a possible distributed (worldwide) approach to design and procurement. It allows the different detector sections to be operated independently, and in so doing allows an added degree of robustness to data collection, in particular supernova burst detection, by giving a method for eliminating entirely even short periods when the entire detector is off. It also allows for unique design features in the different detector sections of LBNE.

The majority of this chapter describes the full conceptual design for the main data acquisition which could be used in all, or just some of the detector sections of the experiment. This is the reference design, used in the 2015 cost and schedule estimations for LBNE.

The main data acquisition is introduced in Section 1.1. It is composed of a readout for the LArTPC based on Cluster-On-Board (COB) modules (Section 1.2), a readout for the photon detectors based on SSP modules (Section 1.3), a time synchronization system (Section 1.4) and a readout and trigger generator for auxiliary signals associated with the experiment (Section 1.5). The real time data collection is performed by a toolkit called artDAQ (Section 1.6) that will implement two arms of data collection, event-building and processing; one of these is primarily dedicated to collection of full non-zero suppressed data for the most important triggered events such as beam or atmospheric neutrino candidates; the other arm receives zero-suppressed data from the whole detector for all times, it is responsible for deadtimeless triggering in software of these important events, has ring-buffers to store potential supernova events, and records low energy physics events. The run control (Section 1.7), online monitoring (Section 1.8), slow control (Section 1.9) and infrastructure for the DAQ subsystem (Section 1.11) are also described. There follows the description of the modular DAQ design approach as introduced above, in Section 1.12. This allows a heterogeneous approach to different detector sections of the DAQ and gives the ability to keep the majority of the detector running when one part needs to stop data taking for maintenance.

# 1.1 Introduction

sec:daq_intro

## 1.1.1 System overview

The DAQ subsystem will perform the primary functions of:

- Readout of raw data from the LArTPC electronics and the photon detector subsystem,

- Continuous filtering and assembly of data to be treated offline as individual events, including receiving and using the Fermilab beam-spill signal,
- Logging data to persistent storage media,
- Configuration, online calibration/checkout, and control of operations of detector electronics, including the generation and distribution of timing and control signals,
- Control of, and readout of data from, devices providing real-time information on detector, subsystem and environmental conditions, and
- Providing user/operator interfaces for these functions via a run control system.

A reference design for the DAQ subsystem is presented in this chapter. The development of this design is guided by recent experience gained in the development of relevant systems for the MINOS, NO$\nu$A [?] and MicroBooNE [?] experiments, as well as from experiments with comparable channel counts and/or experimental conditions, such as D-Zero, CDF, NA48 and ICARUS.

Need references?

Guidance and experience is also available from the design and operation of the LBNE 35t prototype

internal reference?

and the future CERN testing program.

The DAQ subsystem is to be located external to the cryostat vessel, with components in the detector halls (sharing rack space near the cryostat feedthroughs with the photon readout, and the power supplies for components inside the cryostat) and in an underground computer room for the trigger and data collection computers. A small computer room on the surface will also be required for GPS units and network connections between the fiber underground and the network to connect to Fermilab. It is not decided yet whether a larger control room or computer room on the surface is needed in addition to the one underground. The interfaces to this work package are with the front-end electronics for the LArTPC, the photon-detector subsystems and the offline computing.

To increase robustness and up-time, the modular DAQ design concept makes a comparatively loose coupling between the DAQ subsystems in each detector hall, to allow calibration or maintenance to proceed in one, while data collection continues in the others. The DAQ subsystem includes the run of optical fibers to the surface for transfer of data, GPS time synchronization and e.g., telephone. The DAQ subsystem interfaces with the Fermilab Accelerator complex (the beam-spill signal), and has a read-only interface with the cryogenics subsystem for logging of conditions.

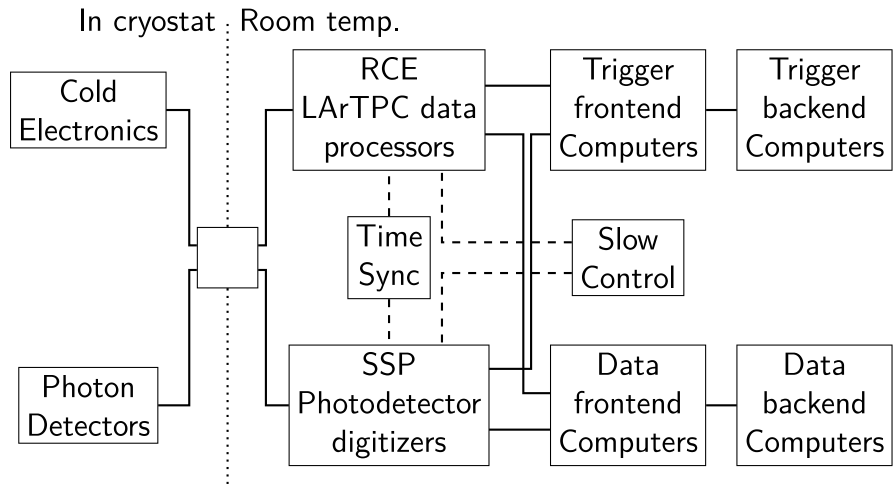


Figure 1.1: Block diagram layout of the main components of the DAQ subsystem.

The DAQ subsystem reference design described in this chapter is shown in Figure 1.1 and consists of the following components:

- logic and processing elements called “Reconfigurable cluster elements” (RCE) residing on daughter cards in ATCA crates

  What is ATCA?

  located in the detector hall to receive data from the LArTPC; these carry out data merging, buffering and compression and transmission to the local farms of commodity computers (see Section 1.2)

- custom ‘SSP photon detector digitizers’ which digitize and process the light signals (described in Section 1.3)

- two local farms of commodity computers providing two separate branches of readout, triggering, event processing and logging of the detector computers; these use the artDAQ toolkit for data acquisition systems (see Section 1.6)

- a custom timing system consisting of a master unit, situated on the surface, that locks onto a GPS clock and distributes timing signals to the data concentrator modules via slave units (see Section 1.4)

- dedicated computer nodes that host run control, online monitoring, database services and slow controls processes

The DAQ subsystem does not include power-supply hardware for the LArTPC or front-end electronics, nor does it include the cryogenics subsystem process-control and monitoring functions. The SSP readout modules for the photon detector subsystem is included in that part of the project.

### 1.1.2 Physics Considerations

The physics considerations determine the scale of the primary tasks of digitizing the LArTPC data readout, merging the photon system data, event building and online processing. In addition to rates for processes of interest, the DAQ subsystem design depends critically on the specifications for the LArTPC and front-end electronics systems, chosen to satisfy the LBNE physics requirements. The sampling rate of 2 MHz has been chosen so as to achieve the required position resolution along the ionization drift direction. Obtaining sensitivity to signals that occur independently of the LBNE beam spill, such as those from nucleon decay, atmospheric neutrinos or supernova-neutrino bursts, requires a free-running transmission of data from the LArTPC front-end electronics. In principle, this same technique can be used to collect the beam events as well, however the robustness of the beam data collection will be supplemented by transferring knowledge of the spill times from Fermilab to the far detector site using GPS.

The task of data transfer is facilitated by multiplexing and utilizing high-speed data lines (1Gbps) in front-end ASICs in the LAr, and by data lines that provide connection to data-acquisition hardware located outside the cryostat. The hardware receiving the raw TPC data then perform zero-suppression and/or data compression, as desired. A challenge in real-time is to use the information of the channels which have hits on them in a particular interaction to determine a larger set of channels, which includes the surrounding ones, that should be read out. The design, as depicted in Figure 1.1, addresses this difficulty by using two data paths. The trigger data path continuously receives the data from the channels that the online hardware determines have hits on them. This determines whenever an event occurs and designates regions of interest that are to be read out in more detail. Ring buffers in the RCEs store the non-zero suppressed data until the region-of-interest information is received. At that point they transmit non-zero-suppressed data to the second path, the data path, which, within the entire event drift window period, reads out the interesting events (all candidates of cosmic rays, beam events, atmospheric neutrinos, proton decays, etc.) for all channels within active APAs with no zero suppression. This allows the maximum available information for these, the most important LBNE events, to be recorded for offline analysis using e.g., Fourier transform deconvolution techniques.

The trigger data path handles all the zero-suppressed data and can store it for long periods, either in $\mathcal{O}$(hours)-long ring buffers or in permanent offline storage for collection of data from a supernova neutrino burst. Experience from MicroBooNE, which has a similar division of lossy-compressed and lossless-compressed data readout, will be vital in optimizing this feature of the data acquisition. During a supernova neutrino burst, the data path may also be used for data collection; the optimum way for utilizing this is still to be studied.

In addition to physics considerations, the DAQ design goals include minimizing the impact of single-point failures, maximizing the uptime and maximizing the use of commercial components. The robustness is addressed in part by some redundancy (and ability to cross-check) between the two arms of data transfer shown in Figure 1.1, and by the separation of sub-system control in different detector sections of the experiment (Section 1.12).

### 1.1.3 Event Rates and Timing

For the reference design described here, sited at the 4850L of the Sanford Laboratory, the atmospheric-muon rate is small enough — 0.1 Hz within the full LAr-FD active volume — to contribute only negligibly to the DAQ bandwidth requirement.

what does it mean to “contribute to a requirement?”

Signals associated with beam events will be localized within the LArTPC and synchronous with discrete ($\mathcal{O}(1\,s)$ rep rate) beam-spill intervals spanning approximately $10\,\mu s$ and will take $\mathcal{O}(2\,ms)$ for the ionization drifting in the LArTPC to take place. However, other physics events of interest will occur at random times, and can be dispersed throughout the TPC volume as in the case of neutrino bursts from supernovae. Other specific signatures, such as very slow-moving magnetic monopoles ($\beta < 10^{-3}$) may involve signals spanning sample times exceeding the ionization-drift time.

Cosmic-ray muons dominate the physics rate, even at the proposed 4850L site. However, this rate is negligible with respect to noise sources. The reference design proposed here would be capable of operation at shallower depths, up to about the 800L, without significantly impacting the design.

As described in Section **??**, the cold electronics for a single Anode Plane Assembly will consist of twenty 128-channel Front-End Readout Boards, each providing four digital inputs to the RCEs. The Front-End Boards will generate non-zero-suppressed data, i.e., a constant data rate of (1.5 bytes/sample $\times$ 32 wires/link $\times 2 \times 10^6$ samples/wire/s =) 96 MB/s.

Radioactive decay from $^{39}$Ar and $^{85}$Kr in the LAr, and to a lesser extent from detector materials (U/Th/Co/K), is estimated to provide a 65-kHz/APA rate of activity of energy above about 300 keV (0.3 MIPs) but less than $\sim$ 5 MeV, while electronics noise (assuming 10:1 S/N for 1 MIP, and a threshold of 0.3 MIPs) will contribute a relatively low rate per APA of singles. The data rate from cosmic ray muons at the 4850L is considerably lower. Table 1.1 provides a summary of these rate estimates.

Table 1.1: Per-APA estimates of rates and data sizes/rates for various processes. Mbps denotes millions of *bits* per second. Unless otherwise stated, estimated numbers of samples and data rates assume suppression of signals below 0.3 MIP. 'Inst. Data Rate' refers to the number of bits in a 2.3-ms long data block divided by this time interval, while 'Avg. Data Rate' factors in the process rate. A 12-bit ADC is assumed, and no allowance is made for data items such as time-stamp, etc.

| **Process** | **Rate (kHz/APA)** | **Samples (per APA)** | **Inst. Data Rate (Mbps)** | **Avg. Data Rate (Mbps)** |
|---|---|---|---|---|
| Generic 2.3 ms interval (not zero-suppressed) | 0.43 | $1.06 \times 10^7$ | 55,000 | 55,000 |
| Cosmic ray muons (4850L) | $6 \times 10^{-7}$ | $5 \times 10^4$ | 260 | $1 \times 10^{-4}$ |
| 10 GeV EM shower | — | $1 \times 10^6$ | 5,200 | — |
| Radioactivity: $\gamma$: U/Th | $\sim 1$ | 40 | 0.48 | 0.48 |
| $\beta$: $^{39}$Ar, $^{85}$Kr | 63 | 24 | 18 | 18 |
| Electronics noise (not common mode) | $\sim 1$ | 15 | 0.2 | 0.2 |

The conclusion from Table 1.1 is that the average data rates out of the front-end electronics system are manageable: about 20 Mbps of 'salt and pepper' per APA due to radionuclides in the Ar and LArTPC materials. Large beam- or atmospheric-neutrino interactions or showering ultra-high-energy cosmic-ray muons will result in high (Gbps-level) instantaneous rates on the scale of the maximum ionization drift period, but contribute negligibly to the average rate.

### 1.1.4 Architecture Summary

The reference design of the DAQ system is summarized in block-diagram form in Figure 1.1. Component counts are given in Table 1.2.

Table 1.2: DAQ subsystem component counts for one 10-kt module/cryostat.

| Quantity | Description |
|---|---|
| 250 | COBs (Cluster on Board) each with 8 RCEs, |
| 26 | 14-slot ATCA Shelves |
| 27 | Ethernet Switches (one 10G, 26 1G) |
| 26 | TPC Readout Compute Nodes |
| 1 | Master Timing Unit with GPS Receiver |
| 27 | Slave Timing Units |
| 11 | Event Builder Compute Nodes |
| 8 | Other computing nodes (server, run control, slow control etc) |

## 1.2 TPC Readout

The primary interface between the TPC front-end electronics (FE) and the DAQ subsystem consists of an ATCA-based system of RCEs (Reconfigurable Cluster Elements). The RCE system receives the serialized raw data for the FE, stores it with no zero suppression in ring buffers for later retrieval if a valid trigger is received, performs zero suppression on it to send to the trigger-arm. It packetizes and transmits the resulting data to the back-end farms for event building and further processing. Additionally, the RCE system transmits timing and control signals to the FE and forwards configuration data to them at start-up.

The RCE system consists of the following components: a commercial ATCA shelf (2-, 6-, or 14-slot), a Cluster-On-Board (COB) which is the "front board" in ATCA terms, and a Rear Transition Module (RTM) which is the "rear board." The COB is a custom board, developed by SLAC, which holds the processing power of the system. The COB (see Figure 1.2) consists of five bays for holding daughter boards, an onboard 10-GbE switch, and both 10- and 1-Gb ethernet connections for communications with the back-end system. Four of the daughter-board bays are for Data Processing Modules (DPM), each of which can hold up to two RCEs. The RCE is the core processing unit of the system; it is made up of a modern "system-on-chip" (SoC) — currently, the Xilinx Zynq-7045 — with multiple high-speed I/O ports (up

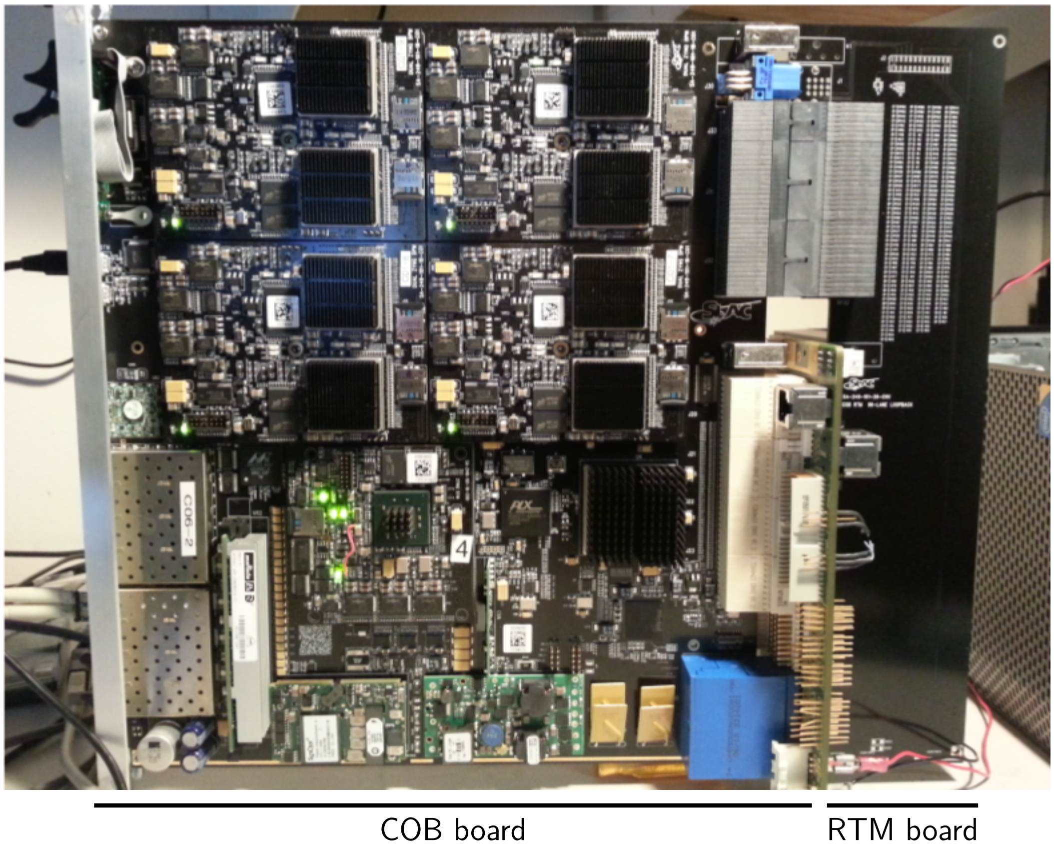


fig:daq15_cob

Figure 1.2: The COB (left) and RTM (right). The eight black square units near the upper-left of the COB modules are the RCEs.

to 10-Gbps each) and 1 GBytes of external DRAM and flash memory controllers. The other bay on the COB contains the Data Transmission Module (DTM), which is responsible for distributing timing and trigger information to and between the DPMs.

While the COB hardware is application-agnostic, the firmware and software on the RCE and the RTM boards are application-specific. The RTM provides the mechanical and electrical interfaces between the front-end (or, in our case, the flange electronics) and the back-end, as well as other external sources such as the timing or trigger systems. In the case of LBNE, fiber optic connections are used between the flange and the TPC DAQ using QSFP+ connectors. This provides the required ground isolation. These links are used for data transmission from the cryostat and also provide the synchronization clock to run the ADCs. Currently, each RTM can accommodate up to 16 QSFP+ connections.

With

Erik suggests a simple diagram to accompany this paragraph

the assumption that each cold FE board multiplexes its 128 wire channels to four outputs at 1-Gbps each, the non-zero-suppressed data for one APA can be fed into a single COB (containing eight RCEs). Each RCE would receive data from two FE boards, perform zero suppression, and send the result to the trigger front-end computers. There would also be a large DRAM ring buffer to store a non-zero-suppressed copy of the data, which would be read out on request (see Figure 1.1) to the data front-end computers.

“data front-end computers”? Why ‘data’?

Figure 1.3 shows a COB in a two-slot ATCA shelf. There are some options regarding the physical distribution of the shelves. One option is to have a smaller shelf at each flange port, each shelf collecting data from the two or four APAs accommodated by that port. Alternatively, the fibers from all APAs could be routed to a central location into a smaller number of large ATCA-shelves.


Figure 1.3: A front view of the ATCA crate with a COB in the top slot.

## 1.3 Photon Detector Readout

sec:daq_ssp

The photon detection system is digitized and readout by a custom module called the SIPM signal processor (SSP) which is described in detail in Section **??**.The module has ADC- and signal-processing hardware. The output of the SSP is Ethernet, interfaced with the same Xilinx Zynq architecture as the COB that is used for the LArTPC data. The SSP therefore is able to provide control, data and slow-control monitoring independently of each other, which facilitates a consistent interface between the hardware and software parts of the DAQ.

## 1.4 Timing System

sec:daq_time

The requirements of LBNE for timing are comparable to those of the NO$\nu$A experiment. NO$\nu$A uses a Global Positioning System (GPS) receiver on the surface, in a module called the Master Timing Unit (MTU), and a protocol that uses four twisted-pair connections (on a standard RJ45 connector) to fan out and then daisy-chain these signals to the electronics. LBNE proposes using the same concep,t with potential enhancements and modernization, for distribution of the surface-to-underground timing among the readout modules for the photon detector and also for distribution to the COBs for the LArTPC readout. This will maintain a 64-bit counter synchronized to the absolute GPS time on each of these readout modules. In addition, a separate protocol for synchronization of the cold electronics boards from the counters on the COBs will be used; this uses the same optical links as the data path described in Section sec:daq_cob 1.2. For simplicity, this

the protocol?

maintains a 64 MHz clock, but not a large counter synchronized on the cold electronics boards, the time being added when the data reach the RCE.

I'm not sure I can parse this sentence correctly; please try to rephrase it a bit.

The timing link is the only physical link other than the communications network. The system described here provides several essential functions:

- Provides a time stamp so that data can be associated with a specific accelerator extraction cycle.
- Provides a common 64-MHz clock to all front-end electronics boards.
- Synchronizes the data acquisition system so that all data from a given time period is sent to the same event builder.
- Provides dynamic synchronization to the cold electronics so that glitches in the clock signal are detected promptly.
- Enables calibration and test pulses to be generated simultaneously for multiple channels.

### 1.4.1 Beam Triggers and Time Stamps

A similar GPS-based system will be used at Fermilab to record the times of the accelerator spills that generate the 10-$\mu$s bursts of neutrinos. There should be

should or must?

fig:daq15_main

sufficient buffering in the trigger path of the readout system shown on Figure 1.1 so that all data can be buffered while the time stamp of the spill is sent to the far detector over the internet.

Since the Fermilab accelerator complex is timed off the 60-Hz electric line frequency from the local power company, and the time between accelerator cycles is a variable number of 60-Hz periods, it may be possible to predict the time of the spills and avoid the need for such a long ring buffer. Experience on MINOS and NO$\nu$A have shown that the variability in the 60 Hz is too great to accurately predict the arrival time of a burst of neutrinos in the future for those scintillation detectors, however LBNE has a considerably longer readout window (several drift times) — long enough that it could work.

A 64-bit time stamp has been chosen for the 64-MHz clock so that all detector and accelerator data will have unique times throughout a 20-year run of the experiment. This will allow correlation with non-accelerator physics events such as supernovae as well as such things as equipment failures.

### 1.4.2 Front-End Clocks

The LBNE detector has very low noise front-end amplifiers. The APA system is connected to 7-meter-long wires so that noise pick up from the clock is a significant concern. One way to nearly eliminate this noise is to select a clock frequency that is well outside the bandwidth of the front-end electronics. This has proven to be very effective in several other experiments. A frequency of 64 MHz has been chosen for this design. The internal capacitance of the SiPMs in the photon system limits the useful frequency range to below about 30 MHz. Thus, a double-pole (6 db/octave) filter can be placed in front of the digitizers. This reduces any noise from the clock system by 12 db. The SiPM's have large internal gain, so 12 db coupled with careful cable design should be adequate to eliminate any possible clock noise.

The LArTPC front-end amplifiers digitize at 2 MHz. The Nyqvist frequency is 1 MHz, so a 4-MHz single-pole filter is quite adequate. Any noise form the clock system is then suppressed by 48 db, which should be adequate for this system.

### 1.4.3 Time stamping and synchronization

The transfer of counter synchronization from the master timing unit (containing the GPS receiver) to the 64-bit counters in the readout modules is achieved using the four twisted-pair lines. Three of the four lines send signals from the master to the readout modules and are used for (a) the continuous 64-MHz clock, (b) a serial command line and (c) a SYNC line. The fourth line, operating in the opposite direction, is (d) a return-SYNC. The synchronization procedure occurs in a similar manner to the NO$\nu$A experiment; The master loads a predetermined time in the near future over the serial command line to a time-load register in all the readout modules. A specific time $T$ (described in the next paragraph) before the GPS time reaches the predetermined time, the master sends a SYNC to all the modules. The SYNC causes each front end to load its time-stamp register from the time-load register. The time-stamp register then increments on each 64-MHz cycle. This register is 64 bits wide and it is appended to each data packet.

To compensate for the delay in the cables, the SYNC pulse is delayed by a preprogrammed number of 64 MHz steps at the receiving end of each fan-out and daisy chain step. The time $T$ is chosen so that the loading in each readout module reflects the correct GPS time. As in the NO$\nu$A timing system, the cable delays are determined automatically with a special delay-calibration sequence in which outgoing SYNC pulses are returned immediately over the return-SYNC line to the preceding fanout which measures twice the delay. As the detector is nearly a mile underground, the four signals (a)-(d) will be converted to optical for the journey down the shaft. The delay on this link will be large, and will be measured and corrected by the same procedure as described above.

The data rates at the far detector are low enough that a software trigger can be used instead of a dedicated hardware trigger. This system operates by sending data to a special trigger farm. This requires that all the data come from the same physical time period. The time stamp system easily provides this synchronization. Each front end has a “data enable” bit that must be set before any data is recorded. At initialization, this bit is turned off. When the synch signal arrives to load the time stamp register, it also sets the “data enable” bit. Since this occurs at the same physical time for the entire detector, it provides the necessary synchronization. The data acquisition software need only monitor the “data enable” to know when data taking has started. It can read the time load register to know the time that data taking started.

### 1.4.4 Dynamic Synchronization and Executes

:dyn-sync-exec

For the LArTPC readout, 64-bit registers exist in each of the COBs. For simplicity, only the 64-MHz clock is transmitted upstream and fanned-out to the cold electronics, not the absolute synchronization to GPS of any counters. To insert time markers on this clock line, phase encoding (or equivalently, modification of the mark-space ratio) is used. The first use of this technique is to detect and correct for possible glitches on the clock line that could put the data stream off by one or more bits resulting in lost data. This is achieved by phase-encoding the start of a 16-channel conversion cycle on the clock line. If the front end was not internally at the start of a conversion cycle, it would reset its internal clock and send out an error message. The warm electronics would also check the phase-encoding to spot failures in the clock distribution system. The second use of the phase encoding is to send an ‘execute’ pulse to the cold electronics which can be used to cause calibration pulses to apear all at the same time. These techniques are used successfully on the MINOS experiment.

## 1.5 Readout of Auxiliary Signals

sec:daq_penn

In addition to the main readout of the LArTPC and photon detector systems, a certain number of auxiliary signals may need to be read out. A trigger module, designed for the 35t test, will be used to perform trigger counter logic and supply external calibration pulsers with triggers at predetermined times. This module is connected to the time-synchronization network as described above

the prev section? Please use Section 1.4.4 or whatever the right label is

. It is read out using a Xilinx Zynq FPGA using the same techniques as for the RCE system. At the CERN single-phase test, this module will also be used to record beam counter information and to digitize the warning of extraction from the SPS accelerator. The module will be retained for the far detector operation for calibration triggers and to readout auxiliary information into the data stream such as e.g., the phase of the 60-Hz line voltage for correlated noise studies.

## 1.6 Event Building and Triggering

sec:daq_artDAQ

Subsequent to the data collection and processing by the RCE system (for the LArTPC), the SSP (for the photon detectors) and the auxiliary readout module, the data is passed to the two local data processing farms, as described in Section 1.1.2 and shown in Figure 1.1, to perform the triggering and event building. Event data will be staged locally before being transmitted in quasi real time for archival to persistent storage (nominally the primary store will be at Fermilab, and the responsibility of the DAQ will end when it reaches this store; it will then be copied to further locations).

### 1.6.1 artDAQ

The data acquisition software in both of the local processing farms of commodity computers will be based on artDAQ, which is a toolkit for building DAQ systems. It has been developed at Fermilab, written in C++11 and is in use in several other experiments. Within artDAQ, core DAQ functions are provided by the toolkit, and experimenters develop the modules that perform the functions specific to the experiment. For the 35t detector, LBNE collaborators have developed modules that configure and read out the RCE and SSP boards that are connected to the LArTPC and photon system detectors, respectively. Other members of the experiment are developing reconstruction and filtering software modules that will analyze the data as it is acquired. The artDAQ-based DAQ system for the 35t detector has been successfully used to acquire data in electronics and detector-integration tests, and artDAQ has been the default choice for the DAQ framework for the full LBNE detector for some time.

ArtDAQ uses the concept of *fragments*. These are used in different ways in the triggering processing farm and in the data processing farm. In the triggering processing farm, in order to implement a completely deadtimeless trigger, the data are read in a continuous stream with no interruption. This is done by reading blocks of data, called millislices, corresponding to fixed windows of time immediately adjacent to each other. The boundaries between one millislice and the next are at precisely the same time across all the subcomponents of the readout, as determined by the 64-bit time counters, which are synchronized as described in Section 1.4. The triggering front-end computers (Figure 1.1) receive zero-suppressed data from one particular part of the readout of the detector and parcel it up into blocks corresponding to the millislice interval. These are the fragments in artDAQ terminology. The data from the beginning of the time interval of one fragment is copied into the end of the preceding fragment to provide sufficient overlap for correct treatment of neutrino events close to the boundary of the millislices.

ArtDAQ designates a back-end trigger farm node for each millislice in the sequence to allow parallelization of the triggering and event-building processing. One fragment from each front-end is sent to each back-end node in sequence so that the back-end node receives all the data from the whole detector for one millislice period. The number of back-ends operating in parallel can be extended if the processing required is large. The trigger farm back-end then runs triggering algorithms on the data to recognize periods of time with clusters of hits corresponding to neutrino, cosmic ray and other interactions in the detector. The times and locations within the detector corresponding to interactions are communicated to a centralized trigger master.

In addition, the trigger back-end nodes store all the zero-suppressed hit data in data files locally, in order to provide a historical buffer of hits for supernova studies. As storage costs decrease, it may become possible to store this data indefinitely, but even with current disk storage capability, the history of many hours of data may be buffered in case a supernova is detected by some other group elsewhere in the world. The trigger master will also provide a self-trigger mode for supernova detection by receiving a summary of the number of hits above a threshold from the back-ends for each millislice. If an increase in interactions, integrated over $O(100\,\mathrm{ms})$ or $O(1\,\mathrm{s})$ indicates a possible supernova, the data in the data storage for that period will be retained indefinitely.

The second local data processing farm operates in a similar way to the first, except the object of this farm is to collect *all* the data, rather than just the zero-suppressed data, during the full drift time for the interesting interactions (beam or atmospheric neutrinos, proton decay candidates, cosmic muons, etc.). The data is stored in ring buffers in the RCEs while the trigger processing farm is working on the data. The trigger master will communicate trigger information (including a mask indicating the 'regions of interest', i.e. which APAs), which determines the time interval and which APAs to readout as a complete event. The *fragment* in the data processing farm corresponds to one trigger. The fragments are merged in exactly the same way as in the trigger farm, i.e., all the fragments corresponding to one event are sent to a particular data farm back-end for merging. Further trigger processing is possible at this stage. The events are written to disk file and transferred to the host laboratory for data archival.

In thinking about the data acquisition system for a reconfigured long-baseline neutrino experiment, it is recognized that using a toolkit such as artDAQ has advantages in that it allows experimenters to leverage existing infrastructure and focus on developing only the components that are necessary for their experiment. In addition, there are advantages to using the same event analysis framework online and offline (artDAQ uses art). This allows experimenters to develop and substantially test their software in an offline environment and only include the full DAQ infrastructure for final testing.

# 1.7 Run Control

The LBNE run-control system will be the primary control and monitoring interface to all data acquisition detector systems for use by detector operators and other users. This run-control system will provide several key functions that will make the collections of specific data acquisition and monitoring systems (referred to as "components") appear as a single, integrated system with a common control, monitoring, alert and information display interface. If the modular DAQ design concept described in Section 1.12 is adopted, the run control will independently control each detector section, with a separate run-state while an overall user interface will give an overview of the state of all the detector sections. This common view of the "system of systems" will ease the training burden on detector shift operators, allowing for rapid response in the event of system error conditions.

The design is modeled on the successful IceCube Neutrino Observatory experiment-control system, known as IceCube Live[**?**]. The overall design presented here is simplified, since the geographical and networking constraints for LBNE are considerably more straightforward than for the extremely remote South Pole site.

The run-control system has a modular design, which is shown in Figure 1.4. It has a dedicated Run Control server, which handles the control messages and communication with the components, the GUI and command line front-end control interfaces, and the database server to manage recording of reported information in an organized form.

## 1.7.1 Control and Monitoring of LBNE Components

The run-control system provides a control and monitoring interface that can be integrated into each component. This allows operators to monitor and change the state (for example: Running, Stopped, or Error) of each component. This component control will be available by both command line interface and a web GUI.This centralized control of all components that make up DAQ systems will support complicated dependencies between components, for example a requirement that the DAQ be in a running mode before starting a calibration system.

## 1.7.2 Display of Component Monitoring and Alerts

During normal operation, all LBNE components will report system monitoring and health values to the run control and will be available for graphical display. These values can include items such as overall component health, trigger rates, memory

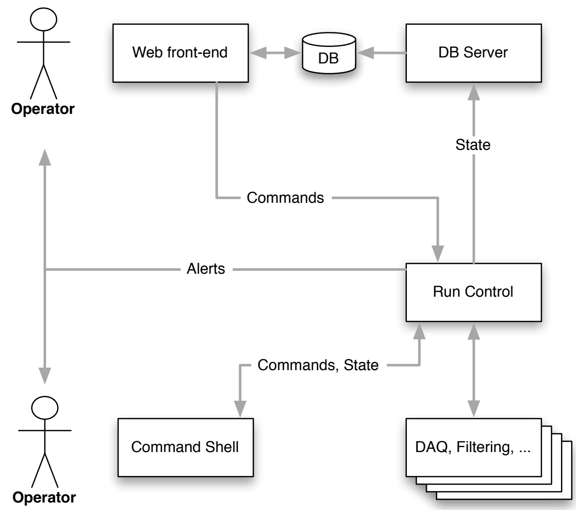


Figure 1.4: LBNE run control system design. fig:expcont

buffer status or rack temperatures. These values will be archived in a database and be presented in a detailed GUI. This interface can be customized to the intended audience, with shift operator pages presenting just high-level status information, and expert views that show detailed, per-channel information for use in detector debugging.

Historical views of component monitoring information will also be available via the user interface. This enables exploration of historical reported information and correlation analysis between detector components. Detector components are not required to be under LBNE run-control-state control, but can simply report useful information. This type of reporting-only integration could potentially be useful for offline systems, and would allow the creation of a more complete historical view of detector history, with information from data collection displayed side-by-side with information from offline data analysis.

The LBNE run control can also generate alerts to detector operators and/or component experts if reported monitoring information is outside expected ranges or a system component is in an error state. These alerts can take any form (email, SMS, etc.) and will be a unified mechanism for alerting operators/experts of operational issues.

### 1.7.3 Management of Per-Run Information

Run control is responsible for assigning a unique number to each segment of data collected by the data acquisition system (commonly referred to as a run number), as well as collecting all user settings selected for a run. This information will include identifiers for the configurations used, components selected for operation during a run, and any provided operator comments. This information will be stored in a database, and made available for collaboration-wide use in online and offline analysis processes.

### 1.7.4 Run-Control Technology

The run control uses several standard network and web technologies. These include:

- XML-RPC - A remote procedure-calling library using XML-formatted messages to issue commands to remote processes, allowing the run-control server to remotely control components

- JSON - An open standard for message formatting that allows arbitrary data to be packaged in a human-readable format using name-value pairs for transmission. Used by the run-control components to format and send arbitrarily complex monitoring records.

- ZeroMQ - A high-performance message transport system that allows a large number of clients to report information to the run-control server.

- DJANGO - A graphical web content framework that can easily query and display information from database records. The run-control server stores recorded monitoring information in database tables, which are available for graphical presentation within the web-base user interface.

While these technology choices are selected for their robust designs, the modular design of the LBNE run-control system allows for flexibility in utilizing other technologies as replacements.

## 1.8 Online Monitoring

sec:daq_om

Rapid access to the data during data collection will be provided by the artDAQ framework and used for making online monitoring histograms and event display. The online monitoring histograms will be used to display the performance of the detector including e.g., noise rates, dead channels and the online drift-velocity measurement. A framework based on the ATLAS LHC experiment online monitoring will be used to display the histograms and associated historical data. Since the data are presented in the art framework online, the development of new histograms can be easily made offline first.

# 1.9 Slow Control Systems

## 1.9.1 Monitoring

The slow-controls system will allow monitoring of data both within and outside normal data taking. The central server is based on the same infrastructure as the run control described in Section 1.7. It will accept individual ZeroMQ messages from subsystem components containing measurement readings and will store them in a database on a server at the far detector. This database will be replicated to a server at Fermilab which will allow access via the standard offline interface being developed at Argonne National Lab, based on the NO$\nu$A database system.

The slow control will generate status and historical plots of the monitored values on demand from the run control DJANGO system so that shift-takers and experts can access the information easily through the same framework as the run control. Alarms will be used to notify shift-crew and experts of values that are beyond their normal range.

Several example data collection routines which send the zeroMQ message to the central control will be provided, e.g., for the Wiener power supplies, the rack protection units, the SSP monitoring functions, the GPS time synchronization and the ATCA crate managers. Storage of performance parameters derived from inside artDAQ, or from the data analysis in the trigger or the online monitoring, will also be reported by zeroMQ messages to the slow-control infrastructure. Collaborators who supply nonstandard equipment will be able to provide their own software to send the zeroMQ messages. A centralized program will harvest read-only data about the cryogenics and about the beam at Fermilab which will also be available for status and historical display. All this data will be replicated to Fermilab and accessible offline through the LBNE offline database framework as described above.

pls indicate section

### 1.9.2 Slow Control

The control functions of the LBNE slow control will allow experts to power off/on crates and racks remotely when it is necessary to restart the experiment after a crash, and no DAQ expert is present underground, e.g., during off hours. The rack protection system will automatically shut down electronics if a smoke, over-temperature or absence of cooling condition is detected (independently of the slow controls). The slow controls will provide the interface to set up the rack protection thresholds and to send reset commands after a trip.

The slow control will have no control over the cryogenics system or the beamline at Fermilab; it will only provide a display of harvested monitoring data.

## 1.10

**What?**

## Interface with Offline Computing

ec:daq_offline

This section needs some cleanup

As described in Section sec:daq_slowcontrol 1.9, the interface

from ??

to the slow-control monitoring data will be handled via replication of the online slow-control database tables to Fermilab where they will be accessible via the standard database-access tools developed for LBNE. This method will also be used to give access to the run-logging database table that provides information about each run. The values of parameters used to configure

to configure the run?

at the start of each run will also be stored in database tables and made available in this way. [The full

set of files, or the full file?

configuration files for each run will be backed up, and a separate job will harvest the relevant ones

relevant to what? maybe say the meaningful ones?

(some are 'plumbing parameters' used for setting registers used only for lab debugging of the module)].

The default output format for artDAQ is readable by the art framework directly. As described in Section 1.12, the data will be stored briefly at each detector section and then a merging job will form complete 'runs' corresponding to all the data in the entire detector for a given running period. This will be archived to Fermilab, recorded in the run control database and accessible directly by the offline reconstruction and analysis programs.

## 1.11 DAQ Infrastructure

### 1.11.1 Wide Area Network

As in the case of MINOS and NO$\nu$A, it is expected that event data can be transmitted over the network to Fermilab. Although rates for events of interest are comparable, data throughput for the LBNE LArTPC is expected to be at least an order of magnitude higher. A bundle of single-mode fibers (the latest estimate is a 96-fiber bundle) should be sufficient for transmitting the data from all four 10-kt modules to the surface, for transmitting the GPS synchronization underground (independently for each of the four 10-kt modules) and to provide expansion capacity.

### 1.11.2 Online Data Storage

To protect against significant periods of absent network connectivity, it is desired to store a significant amount of the data emerging from the DAQ to local storage. A local data storage facility of $\sim 100\,\mathrm{TB}$ is expected to be more than adequate for five days worth of detector data, even without prescaling cosmic-ray muon events. This storage is provided by a RAID disk array on one of the computing nodes in the trigger farm.

### 1.11.3 Power and Cooling

Power and cooling requirements for the DAQ system described here are fairly modest. COB modules operate at around 100 W each and the commodity computers for the data collection and triggering operate at around 200 W each. More detailled power measurements will be provided by the 35t test to refine these numbers and add the other electronics components into the estimate.

You don't give any totals here; might be useful

## 1.12 Modular DAQ Design

sec:daq_upper

The modular DAQ design concept described here allows the detector sections of the experiment to have semi-independent and autonomous data taking — essentially each detector section has sufficient components (timing system, readout of detectors, data storage and database storage) to run on its own, and also to maintain an independent run-control state. The modular DAQ design allows the detector modules to be monitored and controlled as a single big detector, for simplicity in operation. It also allows triggers in one detector section to initiate data collection in the other detector sections. This strategy addresses the problem of ensuring that the possibility of a supernova occurring while data is not being collected is very close to zero, as long as power is present underground. It leverages the inherent redundancy, without increasing cost, in the modular nature of the experiment. It also has benefits in staging and allowing different detector sections to implement different readout features.

The essential difference between the modular design and a single data acquisition for the entire far detector is that if one detector section needs maintenance, or if the data collection stops due to a communication problem or other reason, the unaffected detector sections will continue data collection uninterrupted. If the complete detector is not in data taking mode, this will be indicated on the operations status and alarms screen, and the operator will be able restart the detector section and add it back into the overall run. The run-log database will indicate the periods when any compartment is missing from the overall data collection. Given the rarity and scientific value of a supernova event, it is deemed essential to do everything possible to ensure continuous operation of at least part of the detector at all times.

This sentence was a mouthful... does the previous sentence capture the right idea?

The modular DAQ design concept will improve overall uptime of the detector.

Each detector section runs as a separate data acquisition system with all the real-time functions capable of being run independently. Each compartment has its own run-start and run-stop mechanisms (with a standardized interface for initiating these commands and obtaining state status information); its own timing system (required to be synchronized to UTC to within 10 $\mu$s, although it is likely with GPS that a more stringent limit will be met by the individual detector sections); its own data storage and data catalogue (with a standard format for data and for indicating the period of collection in each file). Each detector section will also have its own independent storage for slow-control measurements.

A centralized 'manager' for triggers across the entire detector will be implemented. This places a number of requirements on the readout of the individual detector sections, although none of these is as arduous as

what?

collecting beam-spill data on its own.

This comes out of the blue... what's collecting it, why is it so arduous and why is it used as a comparison here?

The individual detector sections are able to report when they detect an activity that warrants triggering of the adjacent (or all) detector sections; this could be detection of e.g., a cosmic muon, atmospheric or beam neutrino. These reports are likely to occur only once every minute, and a maximum acceptable reporting rate will be about one every five seconds. The reports should arrive at the central manager within 150 ms of the event occurring.

The readout in the detector sections should be capable of storing the complete set of signals for at least 200 ms. The central manager will report back to the detector sections within this time any requests from adjacent detector sections to collect data within a certain period. The central manager will also distribute the spill timing from the beam to the detector sections. It is likely that the communication between the detector sections and the central manager will be implemented via a dedicated Ethernet network.To establish a connection setup to the central management service the individual detector section will subscribe to it; if the central management is not present or stops, the detector section can run independently (although, in this case, it will not receive trigger requests from the other detector sections). It is possible that an alternative way for a detector section to obtain the spill-timing information independently of the central manager may be available.

this sentence is pretty vague

The

something about 'modular'?

run control will consist of a dedicated run control for each detector section that permits run-starts, run-stops, monitoring, logging, run-mode selection and configuration. The run control functionality can vary between detector sections, but will conform to a standard overall scheme to allow control and monitoring of the entire experiment via a single interface. A separate section of the run-control will give the overall status of the detector sectionsȦll detector sections are expected to be taking data in production configurations; anything else will be regarded as a fault situation.

The ability for the run-control to treat the detector sections as autonomous and to provide an overall status are the required features of the real-time portion of the DAQ.

The status portion runs offline, shortly after the data collection has been completed, operating as a batch system on a farm of computers. The individual detector sections will have data files spanning different periods of time; this operation combines the data into a single sequence of time-ordered files for archival and production use offline. The beamline status and monitoring data will be transferred from Fermilab and incorporated into the files as if this data came from another detector section.