

# DAQ to Cold Electronics Protocol and Timing Distribution for the ProtoDUNE TPC

Giles Barr, Hans Berns, Hucheng Chen, Jack Fried, Dan Gastler,  
Mathew Graham, Eric Hazen, Ryan Herbst, David Newbold,  
Terri Shaw, Matthew Worcester

June 25, 2016

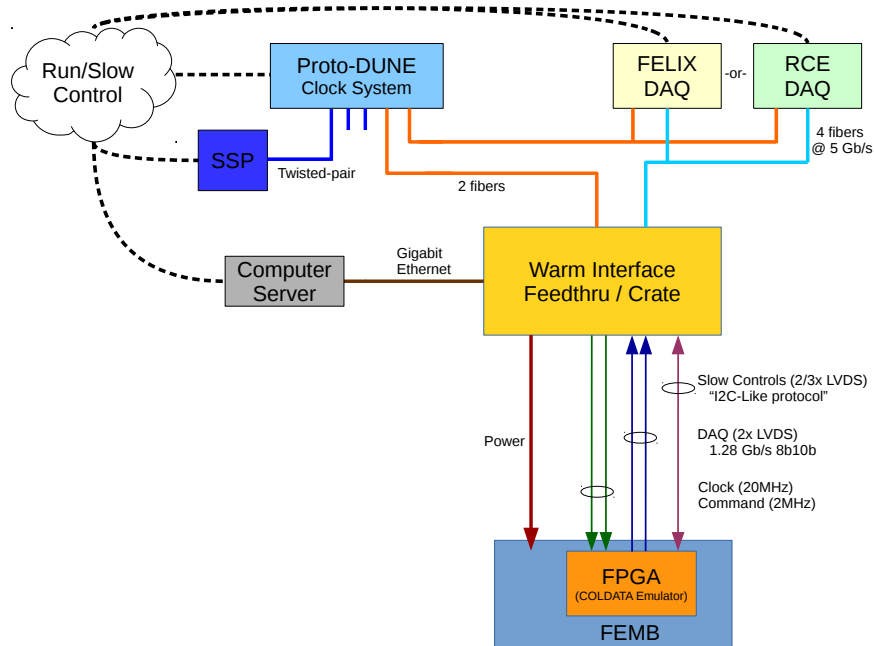


Figure 1: Overview of the proposed timing distribution system. Data flow from the cold front-end motherboards (FEMB) through the warm interface to the RCE DAQ is also shown.

## 1 Introduction

ProtoDUNE is a liquid-argon time projection chamber (LAr TPC) that is observed by six anode-plane assemblies (APAs). Each APA contains 2,560 wires which collect charge from the drift electrons produced by particles ionizing in the LAr TPC. The charge signals on the wires are processed, digitized, and read out by the TPC cold electronics.

The ProtoDUNE TPC cold electronics readout chain consists of the cold front-end motherboards, cold copper cable, warm flanges mounted on the cryostat signal feed-throughs, and warm interface electronics. The data is transmitted on optical fiber from the Warm Interface Boards (WIBs) to the Reconfigurable Computing Elements (RCEs) where it is buffered and any compression and trigger decisions are applied. The data can also be transmitted to the FELIX system, which can also apply trigger decisions. Each of the RCE or FELIX DAQ systems, cold front-end motherboards (FEMBs), and warm interface electronics require communication data protocols and timing signals that are described in this document and summarized in Figure 1.

## 2 Timing Distribution System

The ProtoDUNE timing distribution system provides a stable clock, encoded control commands and a return link for error checking and system monitoring to each readout element in the ProtoDUNE experiment. “Readout element” in this case means a card or crate which is designed to receive a link from the timing system, and may distribute it further locally within a subsystem. The timing system is inspired by those developed for the NOvA experiment, the CERN TTC system and others. One transmit (downstream) and one receive (upstream) signal path is routed to each readout element. These signal paths may be either optical fiber or copper. In the case of copper, the signal lines will be twisted pairs, using a differential signaling standard such as LVDS or RS-422/485.

Each path is encoded on a 50 MHz carrier clock, encoded using *e.g.* biphasic mark encoding, carrying a 100 MB/s serial bit stream. This encoded signal is fanned out and sent identically in parallel to all readout elements, as shown in Figure 1. In the case of an optical system, a passive optical splitter could be used to accomplish the fanout. The return signal path uses the same low-level protocol, and is intended to support a “shared” return path, so a control signal should be provided to extinguish the laser output or disable the electrical driver. The return path allows the timing electronics to verify that all subsystems are synchronized.

This system distributes a set of synchronous commands such as CONVERT, CALIBRATE, RESET synchronized to a specific clock period, and also a separate logical channel for low-speed individually addressed or broadcast commands. For example, the CONVERT command provides a synchronization check (*e.g.* every 16-bit rollover) to the 2 MHz digitize clock that is required by the cold ADC ASICs, as described in Section 5. However, the details of this high-level protocol are beyond the scope of this document.

The current subsystems that will require the system clock are the RCE DAQ (8 fibers), the Photon Detector readout electronics (SSPs, approximately 24 cables), the WIBs (either 6 or 30 fibers, depending on the implementation of the Power and Timing Card discussed in Section 4.2), and the FELIX DAQ. A reference design similar to the one on the Power and Timing Card in Figure 5 for the clock receiver logic will be a joint project of Bristol and Boston Universities. The electronics that distribute the system clock and interface to external software such as run control will be provided by Bristol University.

## 3 RCE DAQ

The data from the WIB are received by processing units called RCEs (Reconfigurable Computing Elements), which are housed in industry-standard ATCA crates on COB (cluster-on-board) motherboards that are designed at SLAC for a wide range of applications. The RCE is a SoC from the Xilinx Zynq family and contain a full Linux processor system on the chip accompanied by 1 GByte of

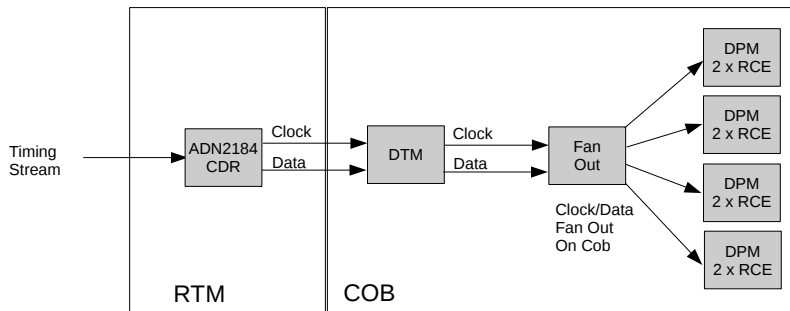


Figure 2: Timing distribution for each COB. The timing stream is the carrier clock described in Section 2.

DRAM. The primary processing function of the RCEs for ProtoDUNE are compression (or zero-suppression) and buffering of the raw data and then sending data to the backend upon the receipt of an external trigger.

The interface with the WIB is provided via the ATCA compliant rear-board, called the RTM (Rear Transition Module). This is an application-specific board that is custom made depending on the physical characteristics of the connections to-and-from the FE (e.g. fiber or electrical connections, number and types of transceivers, timing and trigger connections, etc). For ProtoDUNE, the RTM will use a set of QSFP transceivers to receive the data from the WIB, with approximately 5 Gbps per optical fiber, and an SFP+ optical interface for communication with the timing and trigger distribution system as shown in Figure 2.

A preliminary draft of the WIB-to-RCE data format can be found in Figure 3. The WIB will apply header information (marked in orange) and 8b:10b encoding to the COLDATA-formatted data from up to four FEMBs and transmit the data to the RCE. Data error checking will be done in both the RCE and WIB FPGAs; however, the WIB will simply count and flag errors in the header information and continue transmitting data to the RCE, which will handle the error response. The hardware and firmware for the RCEs, COBs, and RTMs will be provided by SLAC.

## 4 Warm Interface Electronics

The ProtoDUNE warm interface electronics are housed in a small crate attached directly to the cryostat flange. The crate contains one Power and Timing Card (PTC), up to five WIBs and a passive backplane which fans out signals and LV power from the PTC to the WIBs.

The WIB is the interface between the DAQ system (based on either the RCE or FELIX) and up to four FEMBs. It receives the system clock timing and control signals from the timing system and provides processing and fan-out to the four FEMBs. The WIB also receives the high-speed data from the

	K/D															notes
1	01	Reset Count [7:0]	K28.5													Start of Packet/Header
2	00	Reset Count [23:8]											$\approx 30\text{hz}$			
3		Convert Count											2Mhz			
4		Error-bits														
5		WIB Timestamp [15: 0]											125Mhz			
6		WIB Timestamp [31:16]														
7		Reserved											WIB			
8		Reserved														
9		Reserved											WIB-CD1			
10		ChkSm B [7:0]				ChkSm A [7:0]								COLDATA 1		
		...														
64	S8 [99:84]															
65	Reserved											WIB-CD2				
66	ChkSm B [7:0]				ChkSm A [7:0]								COLDATA 2			
	...															
120	S8 [99:84]															
121	CRC-32 [15: 0]											Trailer				
122	CRC-32 [31:16]															
123	11	K28.2				K28.1								Idle		
124		K28.2				K28.1										
125		K28.2				K28.1										

Figure 3: TPC data format. Orange boxes are header information appended by the WIB. The white boxes are the unaltered COLDATA-formatted data. When the cold electronics are not receiving the CONVERT command for the ADC digitization, idle characters will be transmitted by the WIB.

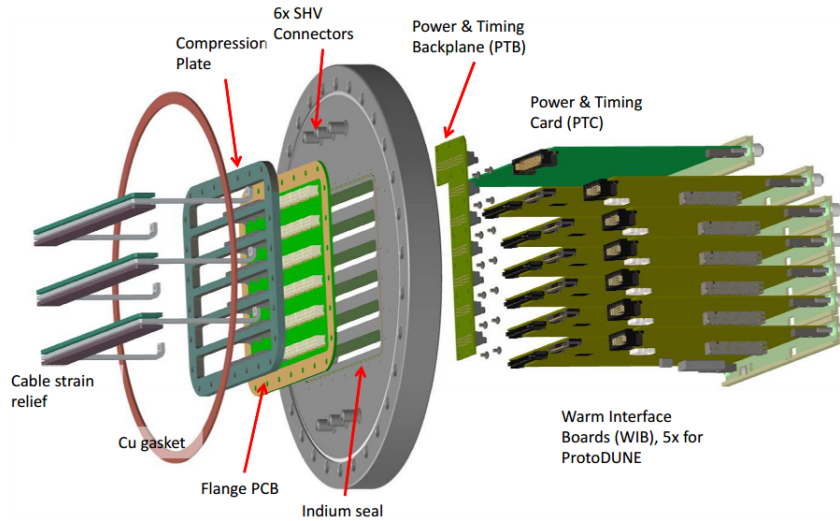


Figure 4: TPC readout electronics flange. Note that ProtoDUNE will have 5 WIB per flange, not 6 (SBND design).

four FEMBs and transmits it to the DAQ system over optical fiber. Figure 4 shows the TPC warm electronics flange (SBND design). The WIBs are attached directly to the TPC readout electronics flange on the cryostat feed-through. The flange board is a PCB with connectors to the cold signal and LV cables fitted between the compression plate on the cryostat side, and sockets for the WIB on the warm side. Cable strain relief is also provided directly on the flange, which is mounted vertically to the cryostat feed-through.

#### 4.1 Power and Timing Card

In addition to the WIB, the warm interface electronics contains a Power and Timing Card (PTC) and a Power and Timing Backplane (PTB). The PTC provides a bidirectional fiber interface to the ProtoDUNE clock system. The received data is separated into clock and data using *e.g.* an AD2814 clock/data separator IC. The clock and data streams are separately fanned-out to five WIB via the power/timing backplane. Details are shown in Figure 5. The PTC fans the clocks out to the WIB over the PTB, which is a passive backplane attached directly to the PTC and WIBs, shown in Figures 4 and 8.

The PTC also receives the low voltage power for the entire cold electronics connected to that flange, approximately 200W at 12V for a fully-loaded flange (5 WIB + 20 FEMB). The LV power then fanned out on the PTB to each WIB, which provides the DC/DC conversion and fans the LV power out to each of the cold FEMBs supplied by that WIB, as shown in Figure 6. The majority of the 200W drawn by a full flange is dissipated in the LAr by the cold FEMB.

ProtoDUNE has 6 cryostat signal feed-throughs with warm interface elec-

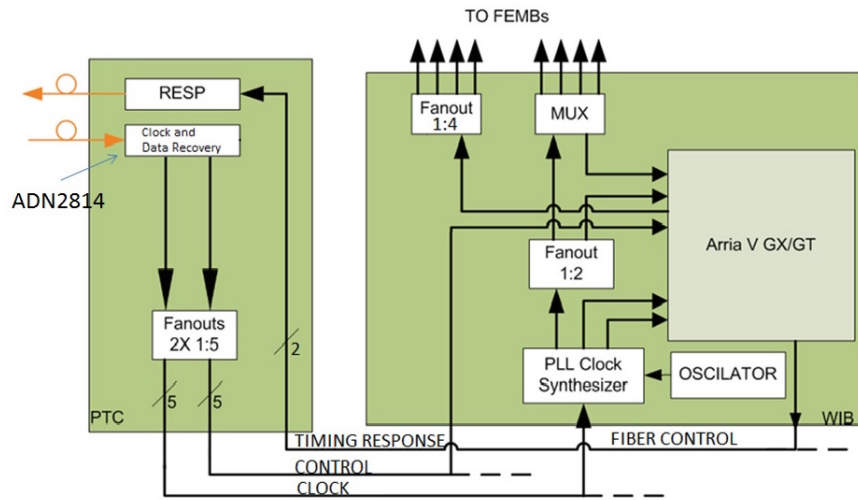


Figure 5: Power and Timing Card (PTC) and timing distribution to the WIB and FEMBs.

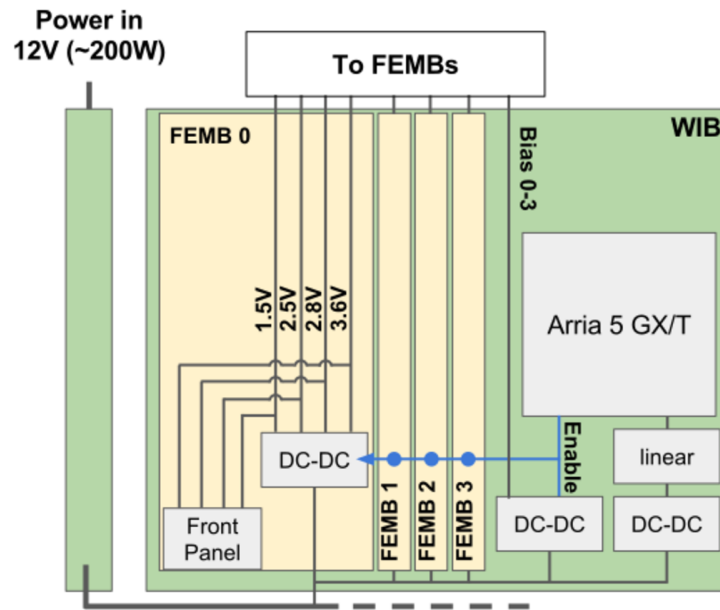


Figure 6: Power and Timing Card (PTC) and LV power distribution to the WIB and FEMBs. 200W is for a fully-loaded crate with the majority of the power dissipated by the 20 cold FEMBs in the LAr.

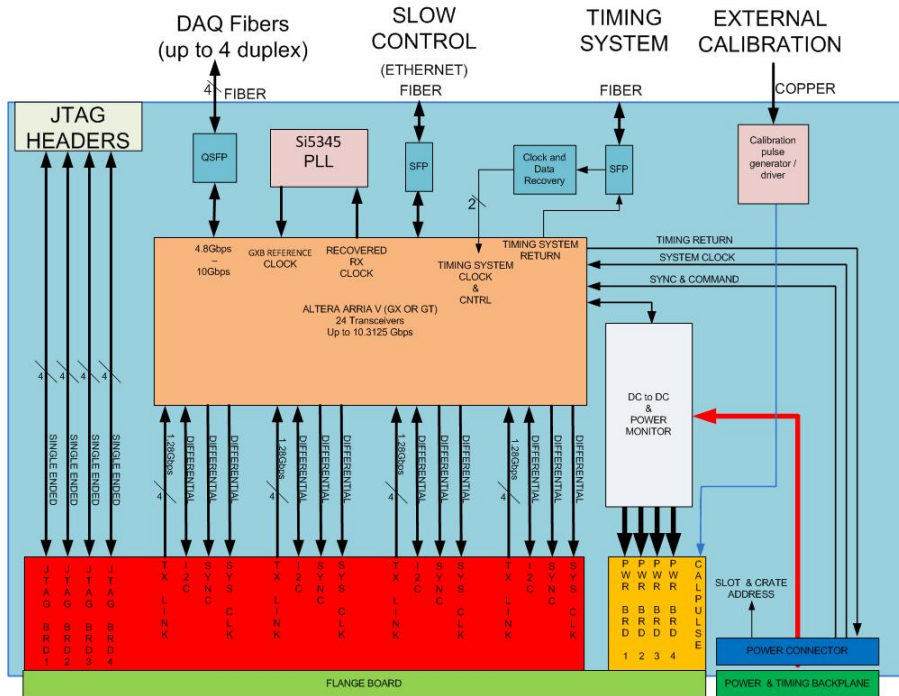


Figure 7: Warm interface board (WIB). Note that front panel inputs will include a LEMO connector and alternate inputs for LV power.

tronics. If the system clock is received by the PTC, it will require 6 fiber links. The PTC will be provided by UC Davis. All of the hardware for the warm electronics crate, including the PTB, will be provided by BNL.

## 4.2 Warm Interface Board

The WIB receives the separated clock and data from the timing system from the backplane as shown in Figure 7. Each WIB in ProtoDUNE will contain a unique IP address for its UDP slow control interface. The IP address for the WIB is derived from a crate and slot address: the crate address is generated on the PTC board via dipswitches and the slot address is generated by the PTB slot, numbered from one to five, shown in Figure 8. Note that the WIB also has connectors to receive LV power on its front panel, bypassing the PTC.

The WIB is also capable of receiving the encoded system clock over bi-directional optical fibers on the front panel; in this scenario, the warm electronics would require 30 fiber links to the clock system. The clock will be divided by two using either the FPGA or an on-board clock synthesizer chip to provide the 50 MHz clock required by the cold electronics. The clean clock and processed control stream is fanned out to up to four FEMBs.



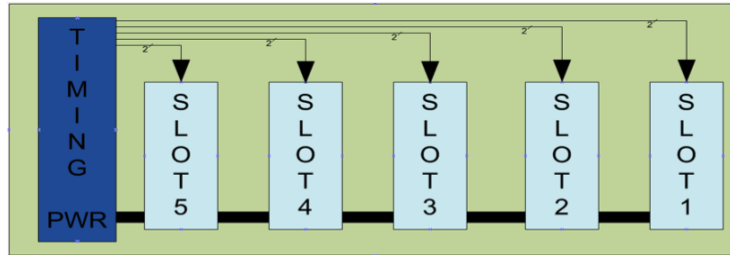


Figure 8: Power and Timing Backplane

The FPGA on the WIB is an Altera Arria V GT variant, which requires a 125 MHz clock for its state machine which is provided by crystal oscillators onboard the WIB. The GT variant of the Arria V transceivers can drive the high-speed data to the DAQ up to 10.3125 Gbps per link, which is capable of transmitting all data from two FEMB ( $2 \times 5$  Gbps) on each link. However, the current design is to use a QSPF socket on the WIB, and deliver 5 Gbps on four optical fibers (1 fiber per FEMB) to two RCEs. The FPGA also has Gbps Ethernet transceiver I/O, also using the 125 MHz clock, which provides real-time digital data readout to slow control.

The sync/command signal protocol is currently under discussion, but one likely implementation is to send a 25-bit serial word at 50 Mb/s every 500 ns, with the word encoding one or more of the four commands described in the COLDATA ASIC specifications:

1. CONVERT – sent every 500ns
2. CALIBRATE – generate a calibration pulse using previously-programmed parameters
3. SYNC – check/reset timestamp counter on COLDATA or cold FPGA
4. COLDATA.RESET – reset all ADCs and the COLDATA or cold FPGA

The protocol for sending these commands to the COLDATA or cold FPGA is documented in detail in the COLDATA ASIC specification. The WIB FPGA will translate the external control protocol described in Section 2 to the one used by the COLDATA, so there is reasonable flexibility in the external protocol which can be accommodated.

The 2 MHz CONVERT clock is distributed by the WIB on up to four differential links, one to each FEMB, by a zero-delay commercial fan-out chip. The configure and control signals to the FEMBs are transmitted over two or three differential I2C links. The 12 links between the WIB and each FEMB, which are currently differential signals transmitted over twin-axial cables, except where noted, are:

- $4 \times 1.28$  Gbps high-speed data

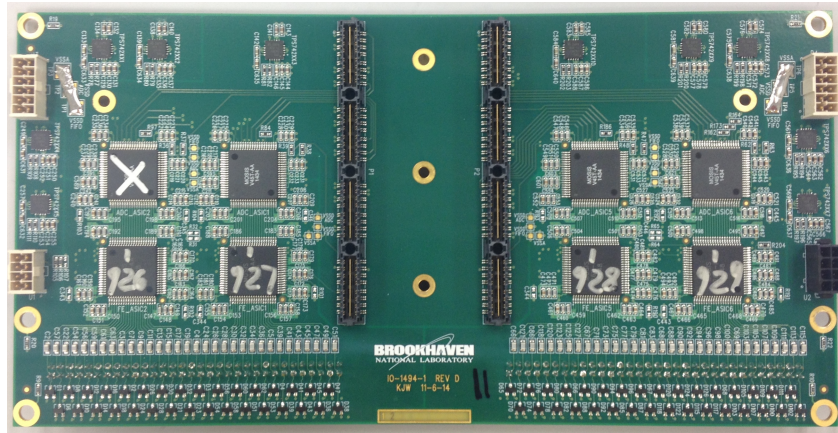


Figure 9: 35-ton analog motherboard.

- One 50 MHz system clock
- One 2 MHz CONVERT clock
- 2/3 I2C control and configure (3 for COLDATA)
- 4 single-ended JTAG programming for the FPGA (not used for COLDATA)

Note that when the cold FPGA is replaced with the COLDATA ASIC, there is a small change in signal assignments required, with one pair of JTAG signals reassigned to the “I2C-Like” COLDATA slow control interface, which uses three LVDS pairs.

The WIB is a joint project between Brookhaven National Lab and Boston University, with BNL providing the hardware and layout development and Boston providing feedback into the layout design and programming for the Arria V FPGA on the WIB. UC Davis will also take a role in providing feedback to the WIB design.

## 5 Front-end Motherboard

The ProtoDUNE FEMB consists of an analog motherboard, shown in Fig. 9, and an FPGA mezzanine. The analog motherboard is assembled with 8 16-channel front-end (FE) ASICs that provide filtering, gain, and shaping to the signals coming from the APA wires. The processed signals coming from the FE ASICs are digitized at 2 MHz on the analog motherboard by 8 16-channel ADC ASICs, providing 128 channels per FEMB of digitized waveform read-out. The cold FPGA configures and controls the FE and ADC ASICs, and packages and sends the digital data to the cryostat flange over four 1.28 Gbps links. In the

DUNE far detector, the cold FPGA will be replaced by two COLDATA ASICs. Note that the COLDATA interface currently requires 9 differential links.

Both the 50 MHz system clock and the 2 MHz sync/command are received on the FEMB in the FPGA. From the 50 MHz system clock, a PLL in the cold FPGA generates the 200 MHz clock that is required for the operation of the ADC ASICs. The FPGA then distributes the 200 MHz clock to each of the 8 ADC ASICs on the analog motherboard. The 2 MHz sync/command is also distributed by the FPGA to the ADC ASICs on the analog motherboard. The rising edge of the 2 MHz clock is the command to each ASIC to digitize. The FPGA is an Altera Cyclone IV, which requires a 100 MHz clock for its state machine and transmitters, which is provided by a commercial crystal oscillator onboard the mezzanine.

The analog motherboard, including the FE and ADC ASICs, and the FPGA mezzanine will be provided by BNL. The COLDATA ASIC and mezzanine are a separate development provided by Fermilab.

## 6 Software Interfaces to the DAQ

The following is a discussion that is still in progress at an early stage.

**Interfaces:** The software interaction with the readout hardware can be classified as three separate functional parts, even if they could be implemented in a more coherent whole:

1. Slow control — Allowed to read status information such as temperatures, voltages, error counters on link status, read configuration etc. at any point (irrespective of whether a run is in progress or not). May need to initiate control functions, but none have been identified in discussions so far.
2. Config control — Write configuration registers, usually at start of run, but also during a run in the case of calibration settings that cycle through a known pattern.
3. Run control — Commands needed to step through the run-start and run-stop state changes to get the system running.

The WIB will communicate using UDP over the 1Gbit Ethernet interface. It can run a server-model (receive command, issue response). It may do sanity checks and report errors for badly sequenced or formed commands. The timing system will probably employ a similar interface with a similar software program. The RCE will accept commands to a process running on its ARM processors.

**Data links** The two ends of the WIB to RCE data links should start up shortly after power on, using default settings without needing to receive commands from the experiment software. Until they receive the start command and CONVERT clock, the FEMB and WIB will transmit idle characters. Commands can be sent to the WIB as data is being sent and will affect future data in a few ticks in the future (i.e. no requirement to change instantaneously). Likewise for the RCE, it will start up in a reasonable default data collection mode

and accept configuration changes while the data is being processed. In the case of many registers, changing during a run is not sensible, and the software will operate by stopping the run, changing, waiting a while and then restarting. The ones which are envisioned to change in the run are concerning collection of calibration pulses of different settings, in which case the calibration triggers will be stopped during the change, and then restarted, but not the rest of the data collection (supernova uptime maximisation). There will be a mode where the WIB can be instructed to send a test data pattern rather than the real data from the FEMB (also possible with no FEMB connected).

**Slow control monitoring of hardware** Each piece of hardware (WIB, RCE, etc.) maintains a set of counters that allow the performance to be monitored (count link errors, give compression ratios etc.) This data is accessed for the WIB via a slow control command over the 1Gbit/s interface and similarly for the RCE. To be decided: whether/when these counters are reset (e.g. once per second, never (software can subtract successive records to get rate)). Error flags are also recorded in the data stream, but note that some of these are lost from blocks that are removed by triggering. These data stream error flags are essential for use offline to determine if there is an error present in that piece of data.

**Run state machine** The mechanism to start/stop the runs will likely follow closely the scheme used in the 35t, which follows the state model implemented in artDAQ (init-ready-running-ready-init; error). However artDAQ can be modified to permit the customisation of this set of states if we need it. In the 35t, the control of the change of states, including starting and stopping processes in the DAQ farm, and clearing up after an error is performed in a program called DAQiFACE, which is implemented as a python script. The plan is to adapt this for the 35t rather than rewriting it if possible.

**Trigger master unit** The ProtoDUNE DAQ will use the concept of a hardware trigger, which is different from what was done in the 35t. This may be implemented as a part of the timing master unit by the timing group, but it can be viewed as a separate entity from a functional point of view. Its job is to (a) define the in-spill and out-spill periods based on signals received from the accelerator, (b) accept triggers from *e.g.* scintillator counters etc. in the beam line (which may be combined in a separate unit such as the PTB or in ‘a NIM crate’, and allow logic to veto triggers where the particles pile up too closely within a time window), (c) generate calibration triggers, (d) apply a software-enable-trigger (SET) condition to only allow triggers when a software enable is present and (e) to pass triggers to the timing system, where they will be distributed over the timing network. At the start of a run, the run control will first put all the hardware into a running state and will finally assert the SET condition to start the run. Similarly, at the end of a run, the first action will be to deassert the SET condition. The SET condition will also be deasserted to implement a data throttle when a backpressure condition exists (see paragraph below).

**Backpressure.** The precise way backpressure will be applied is under discussion. The following four methods summarise the discussion so far for ap-

plying backpressure due to exhaustion of buffer space in the RCEs. Further backpressure conditions may exist in the downstream software, which will need further careful thought.

1. The trigger master is supplied with information about how much of the data has been sent downstream by the RCEs, combined with the knowledge in the trigger master of how many triggers it has sent, this allows the trigger master to know when buffer space is being limited, and it can assert the SET signal.
2. Communication paths within the COB are used to OR together a signal from each RCE that indicates lack of availability of space for more events. These can be ORed together across the eight COBs by the trigger master to generate the SET.
3. Allow the RCEs to run data collection from the start of the spill until they fill up with data, at which point they stop until the start of the next spill. Under the assumption of fixed compression ratio, the trigger master will determine when this limit is reached and can stop sending more triggers. This works because the total memory capacity of an RCE divided by the inter-spill time is a bit less than the data rate out from the RCE.

The difficulty is that fluctuation in the compression ratio will mean that the size of the data storage is not precisely predictable. So for the fourth option, the RCEs could be allowed to collect a few more events beyond the capacity calculated from fixed compression in order to try to get a bit more data.

**Scripting** It is requested that in addition to having a GUI to control the slow control, configuration control and run control as described above, that a scripting interface is available, which allows a sequence of actions on the experiment to be performed. This could include *e.g.* running a self-test of a communication path (load a default configuration, override a few parameters to enable the test, run, restore config to default). Another example would be a power-on sequence if setup after a power cut requires items to be powered in a specific order. This is a difficult thing to implement but will be very useful.

**Partitioning** It will be necessary to allow a production run and a test run to be made independently and concurrently, so that debugging of one part of the experiment can take place while another part is collecting data. The software implementation of this is still under discussion and is complex. However the required functionality must be included in the hardware and firmware from the beginning to allow this to be possible. Since there is only one timing system, this means that the regular 2MHz digitisation commands will be common to all partitions. Triggers however will not be common, a trigger command that is distinct for each partition will be necessary, and firmware that receives the triggers (*e.g.* in the RCE) will need a configuration register to indicate which partition it is running in, so it can determine which triggers to respond to. It is recommended that a minimum of four partitions can be implemented, *i.e.* that the partition identifier is 2-bits long at least.

## 7 Conclusion

The ProtoDUNE timing system distributes one encoded 50 MHz clock, carrying a 100 Mbps serial bit stream, which contains all timing information needed for each of the receiving subsystems: RCE, FELIX, SSP and WIB. A return path allows the timing electronics to verify that all subsystems are synchronized.

The high-speed data links between the RCE and WIB are specified to be 4×5 Gbps optical fibers from one WIB to two RCEs. The data format is based on the COLDATA specifications, and will not be altered in the WIB; the WIB will only apply a header, including error checking and timestamp, and 8b:10b encoding before transmitting to the RCE.

The warm electronics will receive the encoded clock from the timing system via the PTC. The PTC will separate the clock and encoded control data and fan out the 50 MHz system clock and sync/control signals to the WIBs on the PTB. The WIB will fan-out the 50 MHz system clock and the 2 MHz CONVERT clock to up to four FEMBs. An alternate path for the system clock is provided on the front panel of each WIB.

A conceptual design of the software interface to the RCE DAQ and cold electronics provides an outline for slow, configuration, and run control.