The SLAC RCE Platform for ProtoDUNE

K.V. Tsang[1,*], *M.* Convery[1], *M.* Graham[1], *R.* Herbst[1], *L.* Ruckman[1], and *J.* Russell[1]

[1]SLAC National Accelerator Laboratory

Abstract. The ProtoDUNE-SP is a single-phase liquid argon time projection chamber (LArTPC) prototype for the Deep Underground Neutrino Experiment (DUNE). Signals from 15,360 electronic channels are received by 60 Reconfigurable Cluster Elements (RCEs), which are processing elements designed at SLAC for a wide range of applications and are based upon the "system-on-chip" Xilinx Zynq family of FPGAs. The RCEs are housed in industry-standard ATCA shelves on a custom blade, called the Cluster on Board (COB). The RCE platform and its processing functions for the ProtoDUNE-SP will be presented.

1 Introduction

The Deep Underground Neutrino Experiment (DUNE) [1] is a leading-edge, international experiment for neutrino science and proton decay studies. DUNE will consist of two neutrino detectors placed to an intense neutrino beam: a near detector at the Fermi National Accelerator Laboratory in Batavia, Illinois, and a far detector 1300 km away from the source at the Sanford Underground Research Laboratory in Lead, South Dakota. The far detector will comprise four >10 kton liquid argon time-projection chambers (LArTPCs) sub-detectors placed 1.5 km underground. Two prototypes have been built at the CERN Neutrino Platform to demonstrate the viability of the DUNE far detector using a particle beam provided by the CERN Super Proton Synchrotron, as well as cosmic rays.

ProtoDUNE-SP [2], the second of two prototypes illustrated in Fig. 1, is a single-phase LArTPC with 0.77 kton liquid argon in a $10 \times 10 \times 10$ m^3 cryostat. It comprises two anode planes, which are separated by a central cathode plane into two drift volumes. Each anode plane consists of three adjacent 6 m $\times$ 2.3 m Anode Plane Assemblies (APAs). An APA has 2560 readout channels from three parallel wire planes (two induction signals and one for collection). Digitized wire signals of 12-bit ADCs at 2 MHz are read out via Warm Interface Boards (WIBs) installed on the top of the cryostat. Each WIB multiplexes data from four front-end broads (128 channels per board) to the data acquisition system (DAQ) via optical fibers. Each APA requires 5 WIBs, totaling 30 WIBs for ProtoDUNE-SP. The baseline DAQ system is based on the Reconfigurable Computing Elements (RCEs) [3] to read out 5 of 6 APAs, while an alternative system, Front-End LInk eXchange (FELIX) [4], is used for the remaining APA.
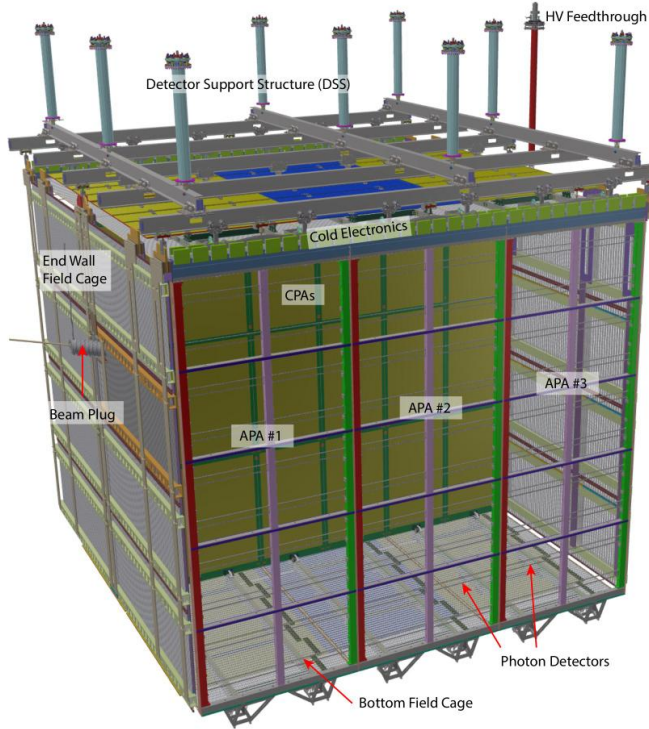
---

*e-mail: kvtsang@slac.stanford.edu

Figure 1. An overview of ProtoDUNE-SP detector.

2 RCE Platform Architecture

The RCE platform was developed at SLAC as a generic DAQ platform that packages the most common requirements of a data acquisition system. It has been adopted by various experiments and applications - LSST, Heavy Photon Search, ATLAS Muon and nEXO (Baseline).

The RCE platform consists of two boards residing in an industry standard ATCA shelf: Cluster On Board (COB) and Rear Transition Module (RTM). As shown in Fig. 2, a COB is a collection of five daughter boards, 4 Data Processing Modules (DPMs) and 1 Data Transport Module (DTM). Each DPM consists of 2 system-on-chip (SOC) clusters, while the DTM consists of 1 SOC cluster. The DPM SOC is a Xilinx 7045 FPGA, a dual-core ARM A9 processor and 1 GBbyte DDR3 memory. The DTM is a Zynq 7030. The DPMs are programmed for digital signal processing, while the DTM provides the switch configuration and serves as a timing distribution path for the DPMs and to/from the backplane. The internal 10-Gbps Ethernet switch connects the 8 RCEs and supports a full mesh ATCA backplane. It allows low latency communication between multiple COBs populated in an ATCA crate. There is a 10-Gbps port out of a COB to read out the processed data. The hardware design of a COB is common to all applications, while user firmware and software are implemented for different experiments.

A RTM is a small board plugged into the back of an ATCA crate. It is designed to be application specific to support digital or optical connections as required by an experiment. For ProtoDUNE-SP, there are 4 QSFP transceivers in a RTM to receive data from 4 WIBs. Each RCE is routed to two of the QSFP's 4 lanes that originated from two front-end boards
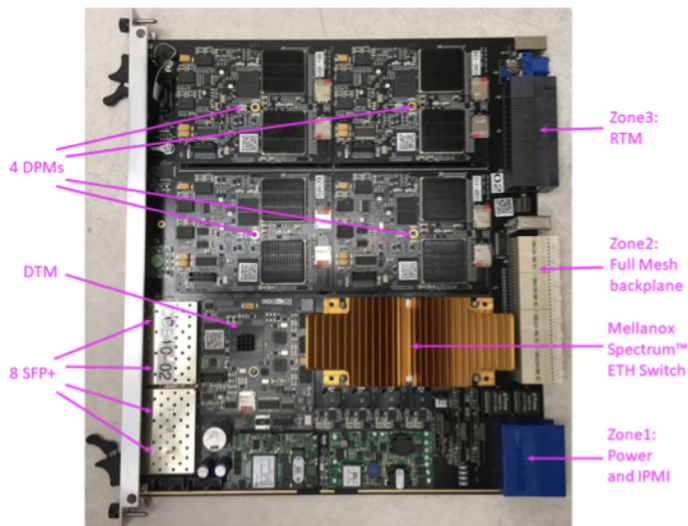
Figure 2. An illustration of a COB. Four DPMs and one DTM daughter boards are installed to the COB. Connectors of the RTM, full mesh backplane and power supply are shown in the right hand side.

(2 x 128 channels). In addition to data receiver, the RTM also takes the timing and trigger messages from external source.

3 ProtoDUNE-SP RCE Data Processing

Data coming from the WIB are organized into frames, where boundaries using a predefined 8B10B characters. Data integrity checks are performed by comparing the computed CRC checksum of a data frame to the checksum provided in the tail of the frame. ADC values are then extracted from the received frame, reorganized into channels, and then delivered to the next processing state in time order. The extraction process is written in a self-contained module in firmware that allows changes in data format as the experiment evolves. Trigger and timestamp information received at RTM are routed to the DTM and DPMs. Trigger messages are loaded to the direct memory access (DMA) of the RCE. These messages are used later to select data to be sent downstream to the DAQ system. A block diagram of ProtoDUNE-SP dataflow is illustrated in Fig. 3. The RCE is also capable of emulating noise data to allow firmware / software development and detector commissioning without the present of WIBs.

3.1 Compression

The reorganized data is passed to a compression module. This module and the compression algorithm were developed using Vivado HLS design flow, written in C/C++, and then synthesized for deployment on the FPGA. The compression is channel based and fully dynamic using Arithmetic Probability Encoding (APE) to achieve the maximum stable compression. The values compressed are not the absolute ADC values, but the differences of successive ADCs. Doing so avoids tracking pedestal drifts, essentially serving as a poor man's filter, removing low frequency noise which shifts the baseline. In order to reduce the size of the histograms, which form the basis of the compression tables and could be as large as 13 bits
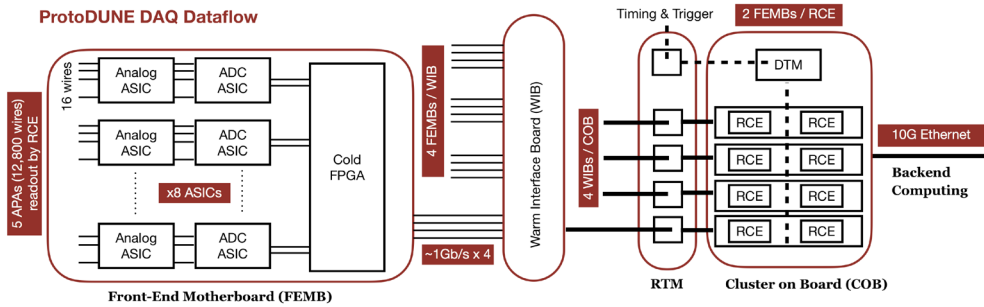
Figure 3. A block diagram of ProtoDUNE-SP dataflow

(8192) bins, only differences within ±15 are encoded. Differences outside this range are labeled with an exception symbol and their exact values are stored in a look-aside list. This is effective because the TPC data is mostly pedestals and, even for signals, the slew is limited by the electronics shaping time. To limit the overhead, the decoding tables are custom encoded, saving a factor of 2, and only the maximum number of bits needed to encode the largest look-aside value is used, usually reducing a 13-bit difference down to 3-5 bits. Typically there are only a handful of these exception values.

On a macroscopic scale, the 128 channels of 1024 differences streams can be compressed in parallel groups. Keeping up in realtime (512 µs for 1024 samples @ 2 MHz) determines the required parallelism with the number of FPGA clock cycles needed to compress 1 symbol being the major determining factor. At 1 clock per symbol, 64K symbols could be compressed with 2 parallel engines each handling 64 channels. Overheads and practical considerations reduce this available 512 µs, limiting it to less than the 64K symbols, thus increasing the needed number of parallel engines.

Achieving 1 cycle per FPGA was a major challenge. Textbook examples of the APE kernel do 1 bit at a time. For this implementation, there is a minimum of 12 ($\log_2$ of 1024 + 2 guard bits) FPGA cycles per symbol, consuming too many resources for the number of parallel compression engines this implies. The solution required both the discovery of how to compress a symbol as a unit and the correct instruction mix to fit within the 8 ns FPGA clock cycle.

Once 1 clock cycle thru-put was achieved, the amount of parallelism was set. The result is 4 compression engines each serially processing 32 channels × 1024 samples. The compression stage thus took roughly 256 µs of the available 512 µs. The time to serialize the 4 compressed output streams into a single output stream is variable, being dependent on the size of the compressed data. However, even the worst case scenario fits comfortably in the remaining 256 µs.

The compression module uses 50k LookUp Tables (LUTs) and 500 18k Block RAMs (BRAMs) to encode a serial stream of 128 channels. One RCE services 2 such streams which are handled by 2 identical firmware modules. This means one RCE handles 256 channels of 12-bit ADCs at 2MHz. The LUT usage is nearly evenly split between the front-end acquisition and back-end encoding phases, while the BRAM usage is dominated by the double buffering of the differences. In operation, data from protoDUNE-SP has an overall compression factor of 4, limited by the entropy of the data, not the compression method. Storing the decompression tables and exception list adds 5-7% to the data volume.

3.2 Buffering and Selection

Data processed in the RCE is stored in a latency queue for later readout when a trigger is received. Since the FPGA has limited memory, the data must be stored in the ARM's processor memory for DMA. Compressed data are buffered continuously in a 1024-frames packet regardless of any trigger. Each packet is timestamped to provide its time range to be used when selecting for event inclusion. Newly arrived packets are pushed to a latency queue to cover the pre-trigger readout period. When a trigger message is received, an event is allocated with a preset readout window. All packets in the latency queue are checked, and then added to the event if any portion of the data is contained in the readout window. The unused packets in the memory are released. Once the latency queue is drained, subsequently received packets are added directly to the event, if and only if they fall within the readout window. Since packets arrive in time order, the first packet that is outside the window indicates the completion of the event. The completed event is then sent downstream to the DAQ system, while the out-of-window packet is stored to the latency queue for the next available trigger. The nominal readout window for ProtoDUNE-SP is 3 ms with 0.25 ms before trigger. It is sufficient to cover TPC signal with a maximum drift time of 2.25 ms at 500 V/cm electric field (corresponding to 180 kV between cathode and anode planes). The readout window is adjustable up to 50 ms to accommodate lower electric field in operation.

3.3 DAQ Software

The backend of the ProtoDUNE-SP is provided by artDAQ, a generic DAQ toolkit for high-energy physics experiments. The artDAQ consists of three major processes on multiple servers: board reader, event builder and data logger. A broad reader is an interface for data transfer between hardware and DAQ software. Each RCE is served by one independent instance of broad reader. At run start, the board readers read configuration parameters from database/files, and then feed into the RCEs for initialization. These parameters include a toggle for emulation or WIB data, the size of readout window and the duration of pre-trigger, as discussed in the previous sections. The board readers continuously fetch completed events from the RCE, and send to the event builders as artDAQ fragments. The data transmission between the RCE platform and the board reader process are carried via the COB's 10-Gbps Ethernet port using a firmware based protocol, the reliable SLAC streaming interface (RSSI) over UDP. A high performance multi-threaded C++/Python library called Rogue is provided to interface between the RCE platform and the board reader process. The maximum output bandwidth for RCE with RSSI protocol is 9.6 Gbps per COB, which is approaching the 10-Gbps Ethernet port limit. At the end of dataflow, the event builders assemble all fragments from the same trigger timestamp, and log to storage using the data loggers. The full DAQ chain has been tested to run steadily at the baseline trigger rate of 25 Hz in nominal ProtoDUNE-SP operation. A stress test has also demonstrated the ProtoDUNE-SP DAQ system is sustainable up to 50 Hz, end-to-end from RCE to storage.

4 Experience on Vivado HLS

Vivado HLS's augmenting of C++ to support FPGA features and its design flow are well thought out. The more time-consuming phases are only done when the code is in very good shape, so the initial and more often done stages turn-around quickly. However, in the course of developing the code, there were issues. A few of the more major ones were:

- Error messages could be obscure and sometimes not related to the problem.

- A change in one piece of the code could break another previously and seemingly unrelated working piece of code.

- The generated output code was sometimes incorrect, causing failures in the CoSimulation or full-up simulation stages. These were the most serious and time-consuming to track down and work around.

- A newer version of the Vivado HLS development package could substantially change resource usage or even break a previously working piece of code.

Some of these problems are likely related to the fact that Vivado HLS does not produce native FPGA modules (like .dcp files) directly, but rather, writes VHDL or Verilog. This means that errors discovered when compiling/building the VHDL or Verilog code can be hard to trace back to the source code.

Others problems may be more inherent in the global nature of FPGA code. In a serial processor, the code is developed and checked in a modular fashion. While the HLS development supports modular development, the fact that the realization of this code on the FPGA is very dependent on the entire mix of code, means that one realization may work, but another, due to, for example, routing or data dependency conflicts, may fail.

5 Summary

The RCE platform has been fully integrated to the ProtoDUNE-SP DAQ system. Five out of six APAs, a total of 12800 wire channels, are read out using the RCE system. At the time of CHEP2018 in mid-July, tests on individual APA have been carried, including measuring the wire noises in a cold box at ~150 K (Fig.4). Data compression has been deployed since September 2018. An overall compression factor of 4 has achieved to meet the ProtoDUNE-SP DAQ requirement. The dataflow has been tested to be sustainable beyond the baseline trigger rate of 25 Hz for nominal data taking conditions. Beam data has been successfully taken from September to mid-November 2018 at CERN. The performance of the RCE platform exceeds the baseline requirements described in the technical design report [2].

References

[1] DUNE Collaboration, *The DUNE Far Detector Interim Design Report Volume 1: Physics, Technology and Strategies* (2018), `arXiv:1807.10334`

[2] B. Abi et al., *The Single-Phase ProtoDUNE Technical Design Report* (2017), `arXiv:1706.07081`

[3] R. Herbst et al., *IEEE Nuclear Science Symposium and Medical Imaging Conference (NSS/MIC)* (Seattle, WA, 2014) pp. 1-4 *Design of the SLAC RCE Platform: A general purpose ATCA based data acquisition system*, in *2014 IEEE Nuclear Science Sympo-sium and Medical Imaging Conference, NSS/MIC 2014* (IEEE, 2016), pp. 1–4, `ISBN9781479960972`

[4] A. Borga et al., *FELIX based readout of the Single-Phase ProtoDUNE detector* (2018), `arXiv:1806.09194`
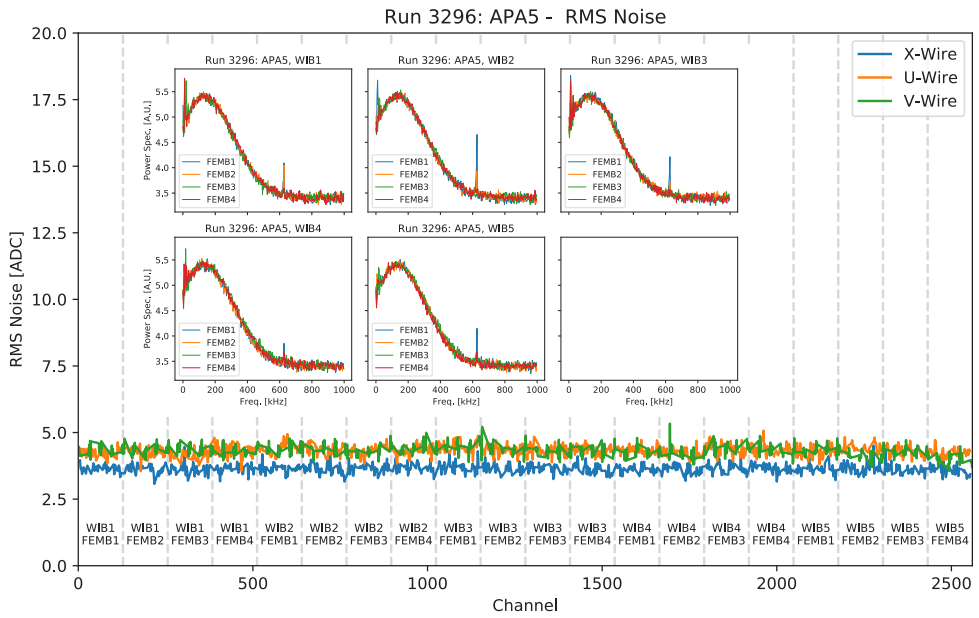
Figure 4. An example of TPC wire noise measurement of an APA. The source of excessive noise at 600 kHz was identified due to the camera system.