DUNE DAQ readout R&D at ProtoDUNE Single-Phase

Roland Sipos - for the DUNE Collaboration
CERN

24th International Conference on Computing in High Energy & Nuclear Physics
4th November 2019
Outline of the talk

- DUNE and its single-phase prototype
  - DAQ differences
- Integration objectives and results
  - FELIX readout
  - Self-triggering
- Summary and outlook
DUNE - Deep Underground Neutrino Experiment

- Future experiment
doing to take data in 2025
- Varied physics program
  - Neutrino beam from Fermilab
  - Explosion of supernovas
  - Atmospheric & solar neutrinos
  - Proton decay
- Underground: shielded environment

DUNE “Far Detector”
  - 4 super-module in 4 caverns excavated in a former gold mine
  - Liquid Argon Time Projection Chamber detectors (LAr-TPC)
  - Cryostat internal dimensions:
    14m x 14m x 62m = 17,000 ton LAr
  - Liquid argon @ 87K (-186 °C)
ProtoDUNE Single-Phase

- Demonstrate design, construction, and operation of the single-phase technology DUNE TPCs
- External cryostat dimensions: 10m x 10m x 10m 750 ton of LAr
- Charged particle beam from SPS

- Ionisation tracks are collected by the wires of the Anode Plane Assemblies (APAs)
- 6 APAs in ProtoDUNE-SP (4% of the 150 APAs of a DUNE supermodule)
- Scintillation light collected by Photon Detectors installed on APA frame
- Successful beam run in Q4 2018
ProtoDUNE-SP DAQ

TPC readout
- Continuous digitization
  - 464 Bytes @ 2MHz
- 15,360 channels
- 55 GByte/s
- Large buffers O(GBs)

Photon Detectors
- Continuous digitization
  - @ 150 MHz
- 240 channels
- Self triggering

DAQ farm
- On-detector electronics connected to DAQ via ~700 optical fibers
- ~20 high performance servers for dataflow, monitoring and control
- 700TB on-site storage
- Maximum 20 Gb/s data rate towards EOS

Timing and trigger
- Phase-aligned master clock to all components
- Aggregate trigger inputs from CRT/PD/BI
- External trigger due to high rate of cosmic flux
ProtoDUNE-SP DAQ

External global trigger:
- Aggregated inputs
- ~30Hz trigger rate
- 3ms extraction windows
- Event size: ~60MB (compressed)

The talk is focused on this!
- Implications due to trigger
- TPC readout with FELIX

William Panduro Vazquez - FELIX: commissioning the new detector interface for the ATLAS trigger and readout system
FELIX in ProtoDUNE

1 FELIX card handles a single APA (10x ~1GB/s links)

Modest modifications on FPGA gateware for lower memory I/O rate:
- Increasing DMA payload size
- Aggregating detector data frames

Data routing software is customized:
- Scatter-gather: collects pointers for detector data fragments
- Single copy pipeline: serialization to user buffer
- Data published on Infiniband over Ethernet

BoardReader process
- Subscribes to single link & buffers data
- Extracts data fragments for triggers
- Reorders data (AVX2 & 512)
- Hardware accelerated compression
  - Intel® QuickAssist (QAT)
DUNE DAQ
DUNE DAQ

Real-time processed datastream!

Self-triggered system!
“Hit” finding

Real-time processing and selection of interesting data regions for trigger decision

HitFinder thread
- Extract collection channel
- Do pedestal, filter, hit-finding (FIR filter)
- Implementation with specific AVX2 registers and instructions

Operational in ProtoDUNE: With full self-triggering chain!
DUNE DAQ

High density in favor of routing feature!
OnHost BoardReader

Merged FELIX data routing software, with data selection (trigger-matching) algorithm

- Eliminated 100Gb P2P connection
- Reduced space and cost requirement

Covered:
- Extensive server evaluation
  - PCIe riser configurations,
  - BIOS settings
- Optimized for memory throughput
  - Performance profiling
- Heavy NUMA balancing between processing threads and allocated memory
- Interrupt moderation of 10Gb NIC

Operational in ProtoDUNE: Provides expected performance characteristics! (Ex. compression)
HitFinding in FELIX FPGA

FPGA gateware R&D for HitFinding

- Bitwise and byte operations are CPU heavy
- FPGA HitFinder implementation is ready
- Integration with FELIX is ongoing
- Outgoing hits have dedicated virtual links

Ongoing work, in collaboration with DUNE DAQ and FELIX developers.
DUNE DAQ

Supernova burst (SNB) buffer is ~100 seconds!
Supernova Burst trigger

A supernova could produce thousands of neutrino events within several seconds!

But recording the data is tricky:
- **Long time** for trigger decision
- Physics event is distributed over time
- Most critical data: avoid any potential losses

Requirements:
- 10s transient buffer (**15TB** for one detector module)
- On trigger: 100s **continuously persisted data stream** (**150TB**)  

Possible solutions:
- Dedicated ultra-fast distributed NAS?
  - How to **dimension a network + storage** to cope with this I/O rate?
- **Bring storage close to readout?**
  - **NVMe SSDs**?
- ???
SNB buffer prototype

Evaluation of RDIMM + NVDIMM (Non-Volatile Memory Module) solutions

Hardware:
- C628 chipset + Cascade Lake Xeon® Platinum L SKU processor
- 192GB RAM + 6TB Intel® Optane™ DC NVDIMMs

Software:
- Persistent Memory Development Kit (PMDK - libpmem)
- pmem in AppDirect (persistent) mode, mounted as dax filesystem in interleaved configuration
- For each APA link:
  - Fixed 100GB pmem pool
  - On SNB trigger:
    - Persist 10s buffer
    - Persist incoming 1GB/s data for 100s!

Preliminary, SNB trigger functional prototype in ProtoDUNE! (Half APA)
Outlook

- Continue hit-finding optimization
- Integrate fake SNB trigger to ProtoDUNE-SP
  - In order to validate the storage and data transmission flow
- Aim for 2 FELIX cards in the same server
  - Testing FELIX driver with NUMA support
- Continue incorporating emerging technologies for data processing, storage, and compression
Conclusions

- We applied the FELIX system to a new experiment
- Substantial progress on DUNE DAQ prototypes and their integration to PDSP
- DUNE DAQ benefits from advancing data center and server technologies
- There are a lot of further possibilities to explore

Also, thanks to ATLAS FELIX developers for their support