EP-DT
Detector Technologies

High-performance storage and dataflow solutions for the data acquisition system of particle physics experiments

Adam Abed Abud
(CERN EP-DT Detector Interface)

December 7, 2022
DT Training Seminar

Outline

- Experimental challenges in neutrino detection
- The DUNE experiment
- Liquid argon TPC
- Physics constraints on the DUNE DAQ
- The DUNE Data Acquisition system
- Low-level triggering system
- High-performance storage buffer for Supernova events
- Particle identification in LAr detectors
- Conclusion

Neutrino experimental challenges

Neutrinos are *elusive* particles that interact weakly with matter

- **Large active volume** is needed to detect neutrinos
  - Interactions can happen anywhere in the detector (unlike collider experiments)
    - Good **timing resolution** is needed
- Need for a **high detector uptime**, always ready to take data:
  - We know when beam neutrinos arrive but we do not know when a galactic event happens (unlike collider experiments where the event is organized by the bunch crossing)
- **Energies and momenta range from MeV to GeV**
  - Difficult to accurately detect ν in the "same" detector
    - Good **energy resolution** is needed to distinguish between track and shower events

The DUNE experiment

Physics goals

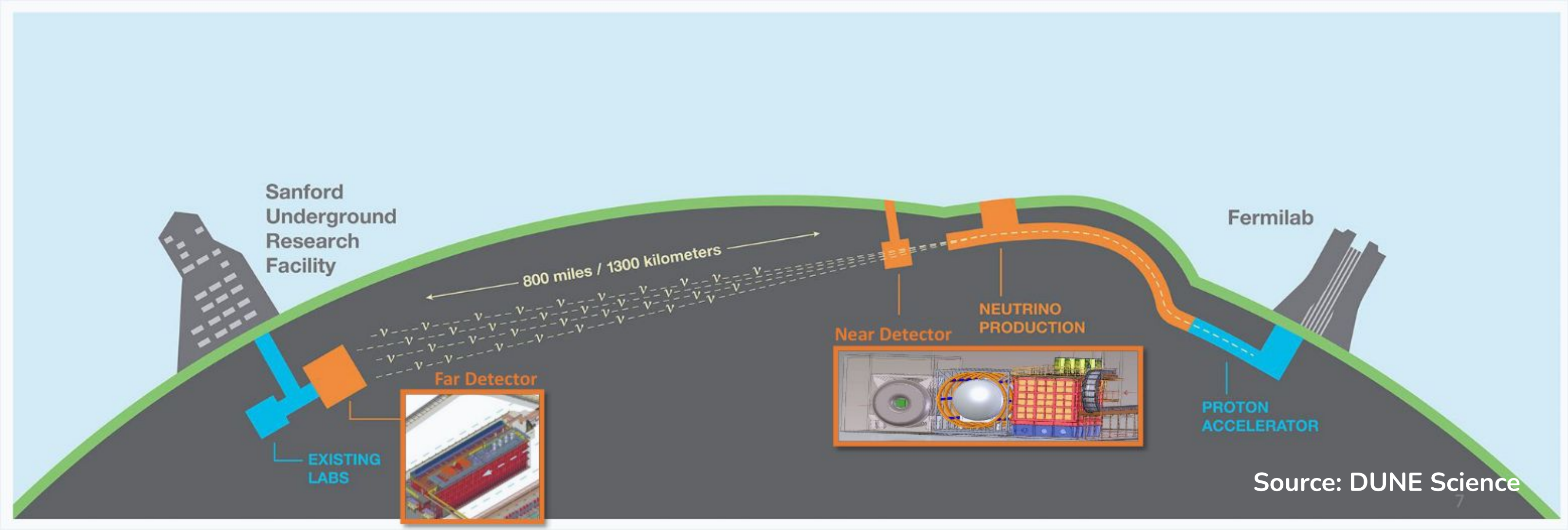
- **DUNE**: Long baseline neutrino detector located 1300 km from the source and 1.5 km underground
  - 1-6 GeV muon neutrinos/antineutrinos from a high-power proton beam from Fermilab
- **Main physics goals:**
  - Solve the hierarchy mass mystery by measuring rate of oscillation of neutrinos and antineutrinos (neutrino beam from FNAL)
  - Study neutrinos from Supernova events
  - Study solar and atmospheric neutrinos
  - Physics beyond the SM (proton decay)
- **Two detectors:**
  - **Near Detector:** characterize the beam (100s of millions of neutrino interactions)
  - **Far Detector:** 4 cryogenic LArTPC modules with a total fiducial mass of 40 kton
    - Trigger and DAQ: 4 independent instances (one for each module), synchronized to a common clock

The DUNE experiment

Physics goals

- **DUNE**: Long baseline neutrino detector located 1300 km from the source and 1.5 km underground
    - 1-6 GeV muon neutrinos/antineutrinos from a high-power proton beam from Fermilab
- **Main physics goals:**
    - Solve the hierarchy mass mystery by measuring rate of oscillation of neutrinos and antineutrinos (neutrino beam from FNAL)
    - Study neutrinos from Supernova events
    - Study solar and atmospheric neutrinos
    - Physics beyond the SM (proton decay)
- **Two detectors:**
    - **Near Detector:** characterize the beam (100s of millions of neutrino interactions)
    - **Far Detector:** 4 cryogenic LArTPC modules with a total fiducial mass of 40 kton
        - Trigger and DAQ: 4 independent instances (one for each module), synchronized to a common clock

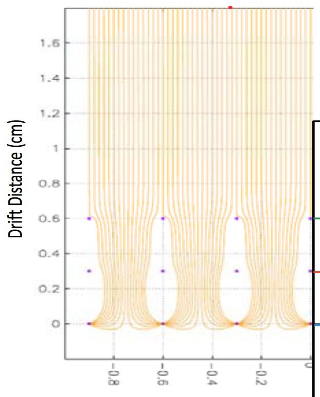
The DUNE experiment
Physics goals

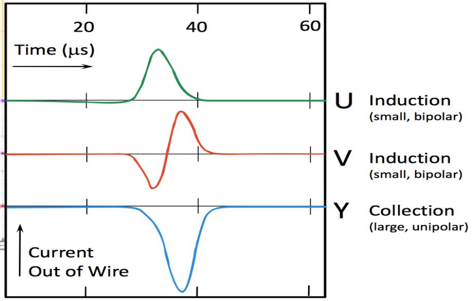


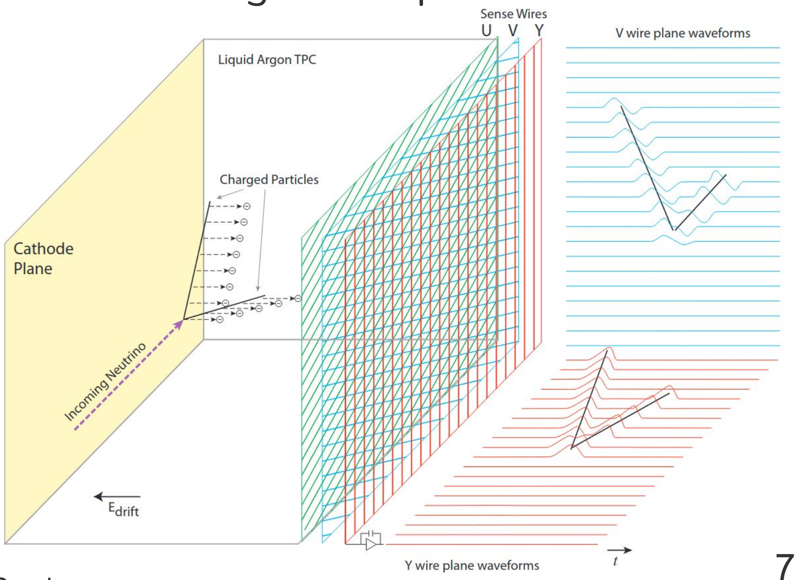
- **Two detectors:**
  - **Near Detector:** characterize the beam (100s of millions of neutrino interactions)
  - **Far Detector:** 4 cryogenic LArTPC modules (with different detector technologies) with a total fiducial mass of 40 kton and size of 15 m (w) × 14 m (h) × 62 m (l)
    - Trigger and DAQ: 4 independent instances (one for each module), synchronized to a common clock

Liquid argon time-projection-chamber (LArTPC)

- Multiple LArTPC technologies: we will focus on the horizontal drift detector technology
- Horizontal drift TPC: multiple drift volumes separated by a central cathode
  - HV at 180 kV providing an E field of 500 V/cm
  - 3 planes of wires to collect the signal (2 induction and 1 collection)
    - **Challenge**: difficult to trigger on induction planes because signal is bipolar
  - Cold electronics to minimize the noise
- Use of photon detectors to trigger the start of the event

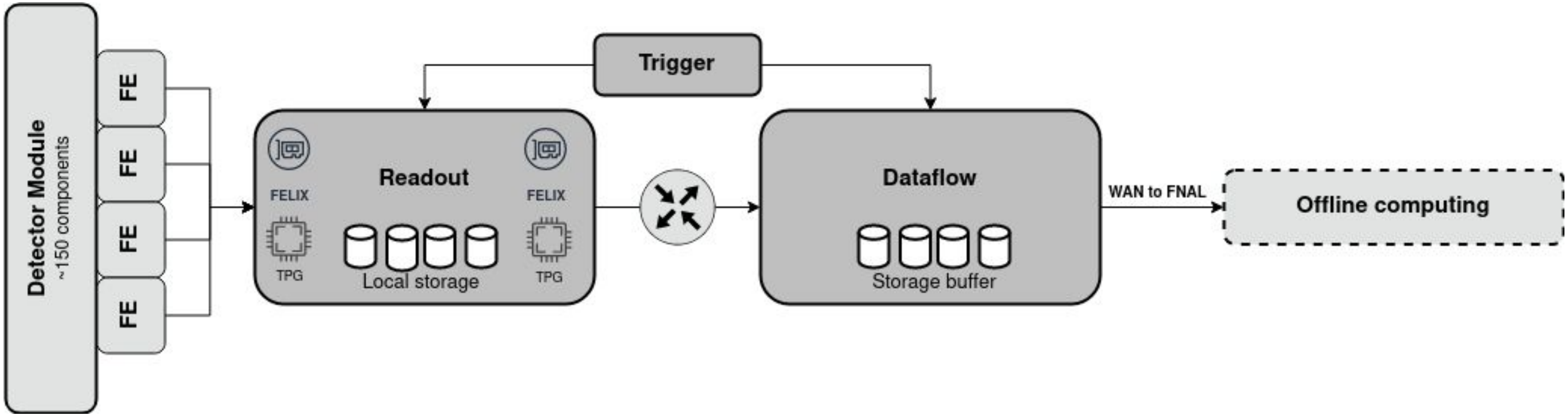






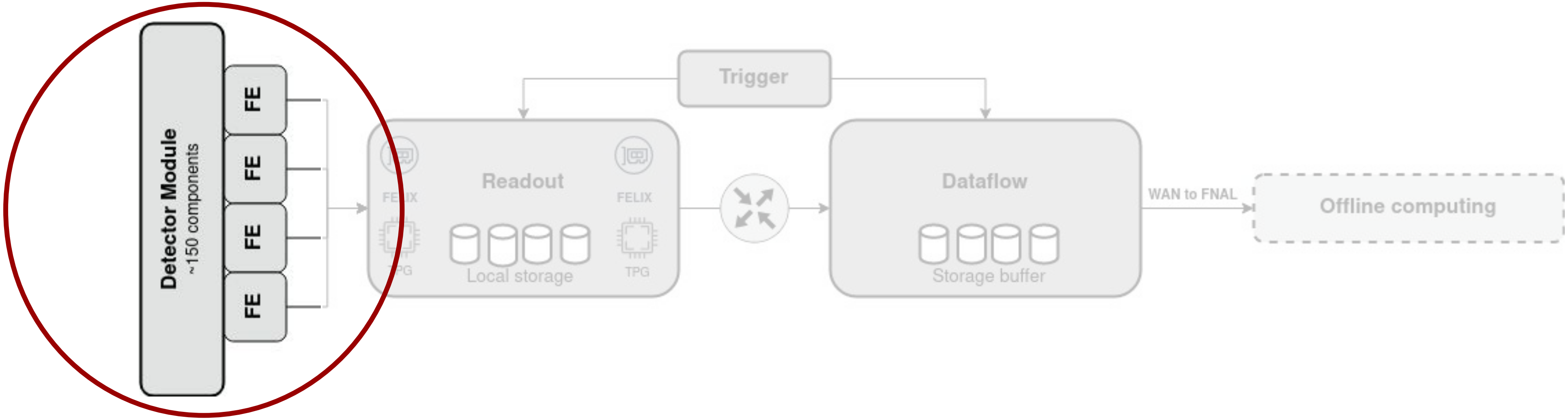
DUNE Data AcQuisition system (DAQ)

- Modular nature of the apparatus allows splitting a cryostat in ~150 identical components



DUNE Data AcQuisition system (DAQ)

- Modular nature of the apparatus allows splitting a cryostat in ~150 identical components

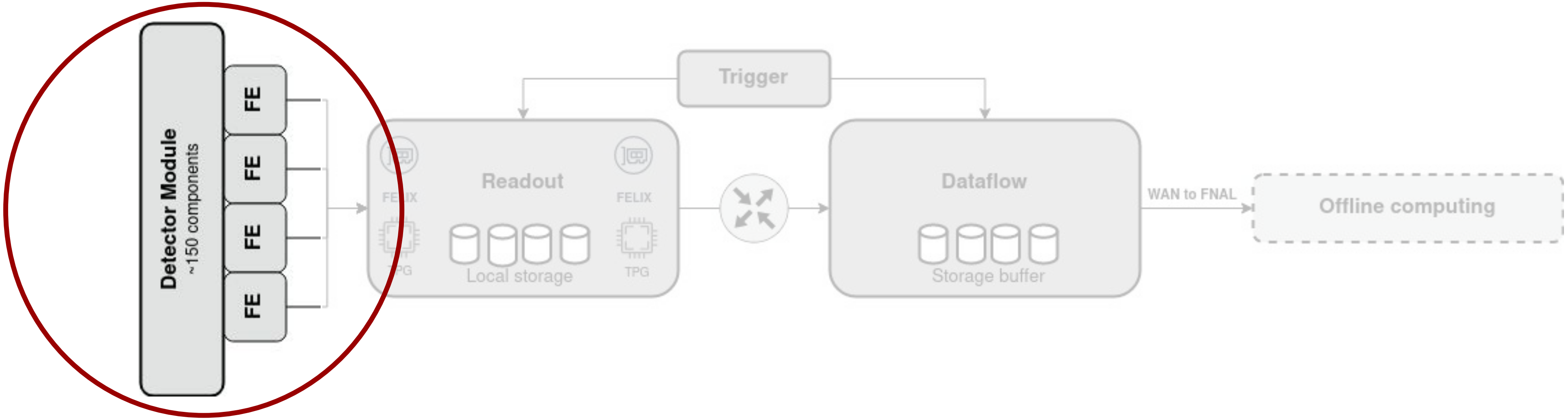


DUNE uses a continuous readout for the LArTPC

- 2 MHz sampling rate
- 384 k channels, 14 bit ADC @ 2MHz —> 9.2 Tb/s
- Adding up all the TDAQ from the four cryostats leads to ~**6 TB/s**
  - Similar rate expected for HL-LHC experiments !

DUNE Data AcQuisition system (DAQ)

- Modular nature of the apparatus allows splitting a cryostat in ~150 identical components

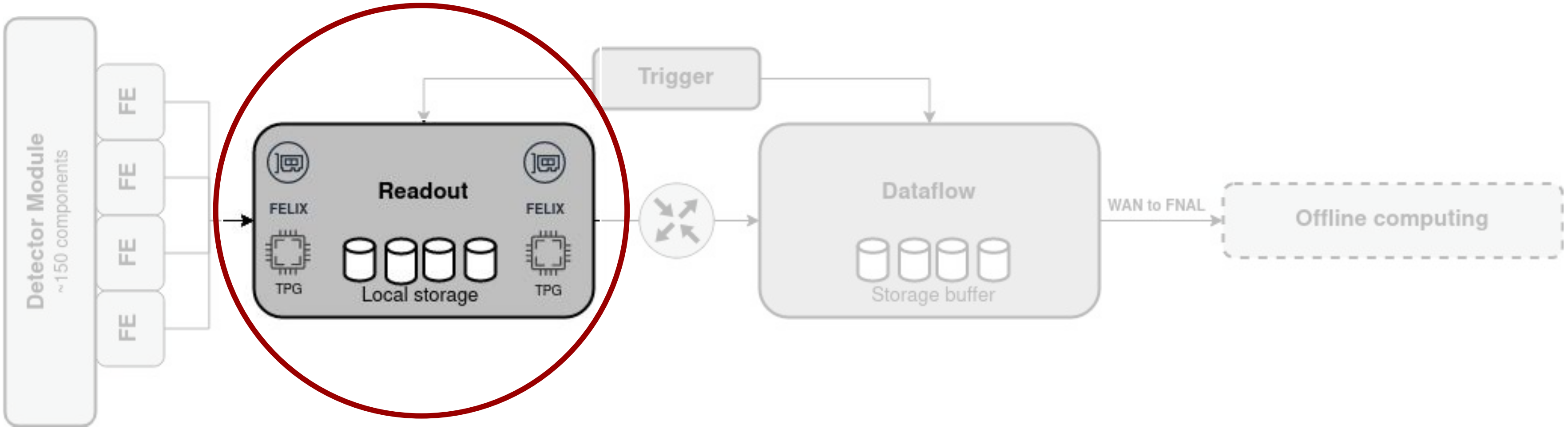


DUNE uses a continuous readout for the LArTPC

- 2 MHz sampling rate
- 384 k channels, 14 bit ADC @ 2MHz —> 9.2 Tb/s
- Adding up all the TDAQ from the four cryostats leads to ~**6 TB/s** = 1000 movies in 4k per second NETFLIX
  - Similar rate expected for HL-LHC experiments !

DUNE Data AcQuisition system (DAQ)

- Modular nature of the apparatus allows splitting a cryostat in ~150 identical components

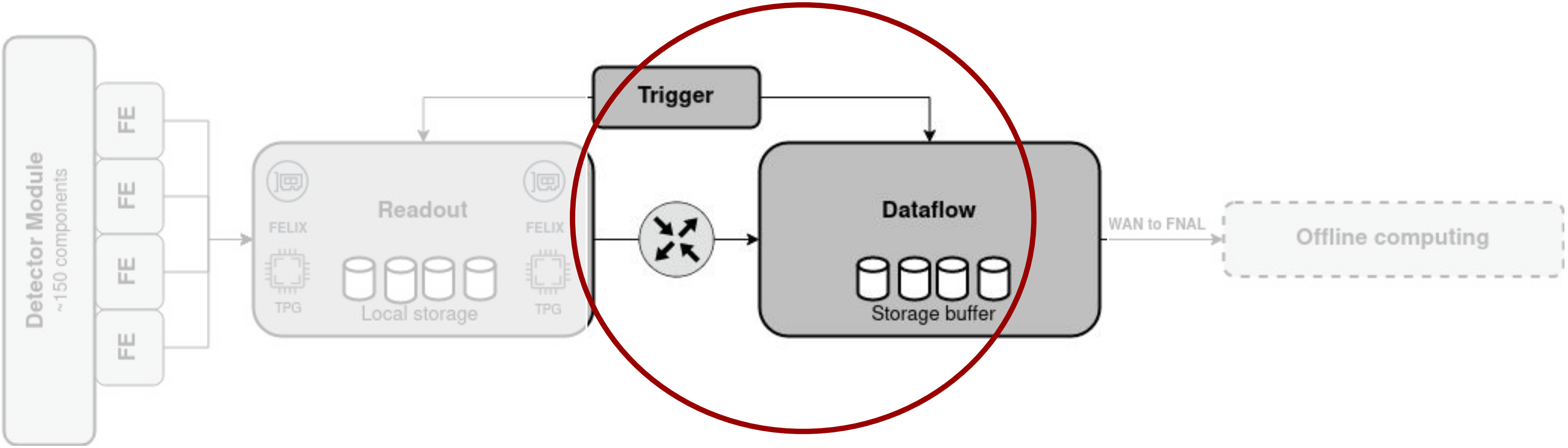


Readout system interfaces the detector front-end with the DAQ processing units

- Commercial-off-the-shelf server with multiple uses:
  - Detector interface: handle the data input from the front-end electronics of the detector
  - Low-level data selection system (*Trigger Primitive Generation*): identify time periods in which the waveforms are noise-free
  - Local storage buffer: temporary store the data while waiting for a trigger decision
- **Data throughput** for each readout unit: approximately **10 GB/s** from 10 optical links
  - Total of ~1.5 TB/s for each detector module

DUNE Data AcQuisition system (DAQ)

- Modular nature of the apparatus allows splitting a cryostat in ~150 identical components



Trigger combines a subset of readout (TPs) data into time windows of interesting signals:

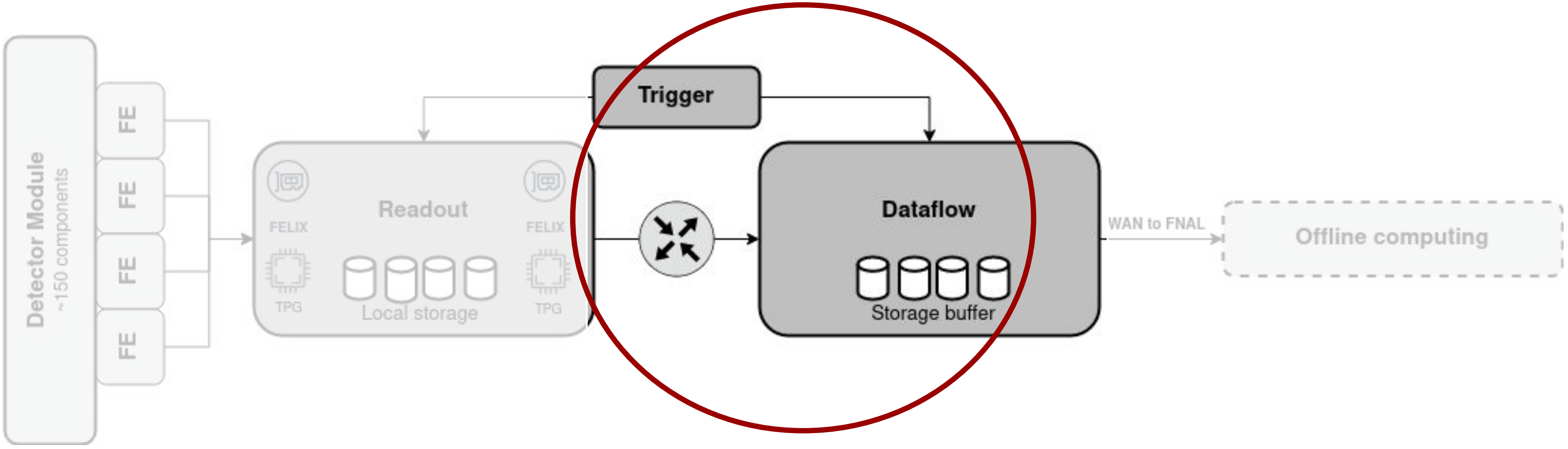
- Time "window" can vary from < 1 ms to ~100s;
- Data size ranging from few MB to ~150 TB

Dataflow moves the data fragments (identified by the trigger) from the Readout nodes to a large storage buffer

- Total storage size is 1 PB (approximately one week of data taking)

DUNE Data AcQuisition system (DAQ)

- Modular nature of the apparatus allows splitting a cryostat in ~150 identical components



Trigger combines a subset of readout (TPs) data into time windows of interesting signals:

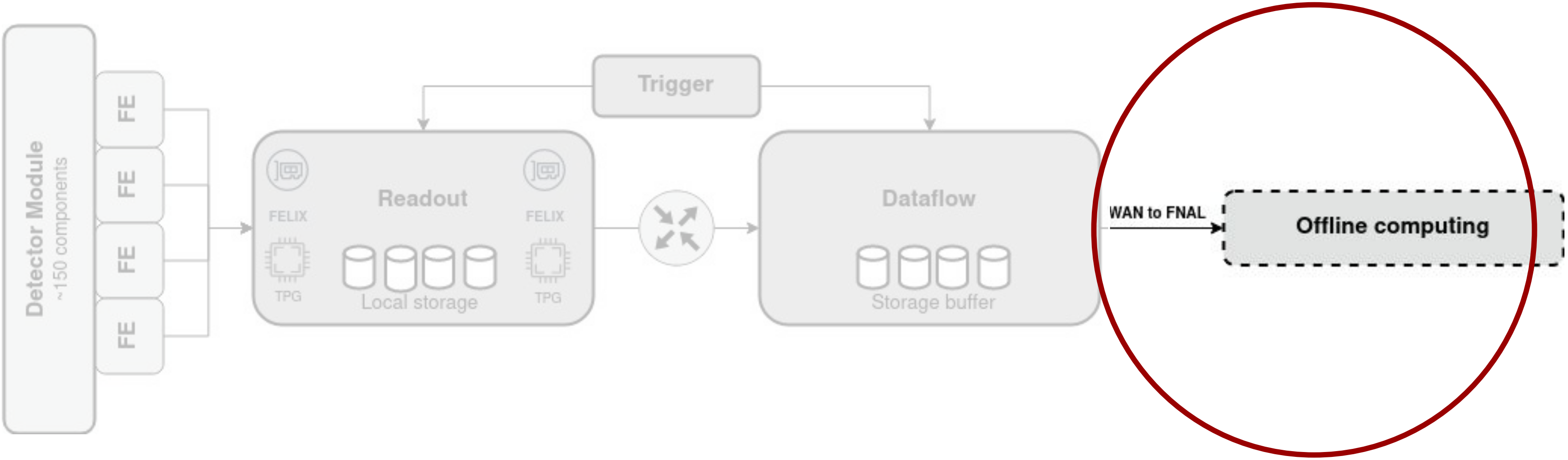
- Time "window" can vary from < 1 ms to ~100s;
- Data size ranging from few MB to ~150 TB

Dataflow moves the data fragments (identified by the trigger) from the Readout nodes to a large storage buffer

- Total storage size is 1 PB (approximately one week of data taking) = 150k movies in 4k NETFLIX

DUNE Data AcQuisition system (DAQ)

- Modular nature of the apparatus allows splitting a cryostat in ~150 identical components

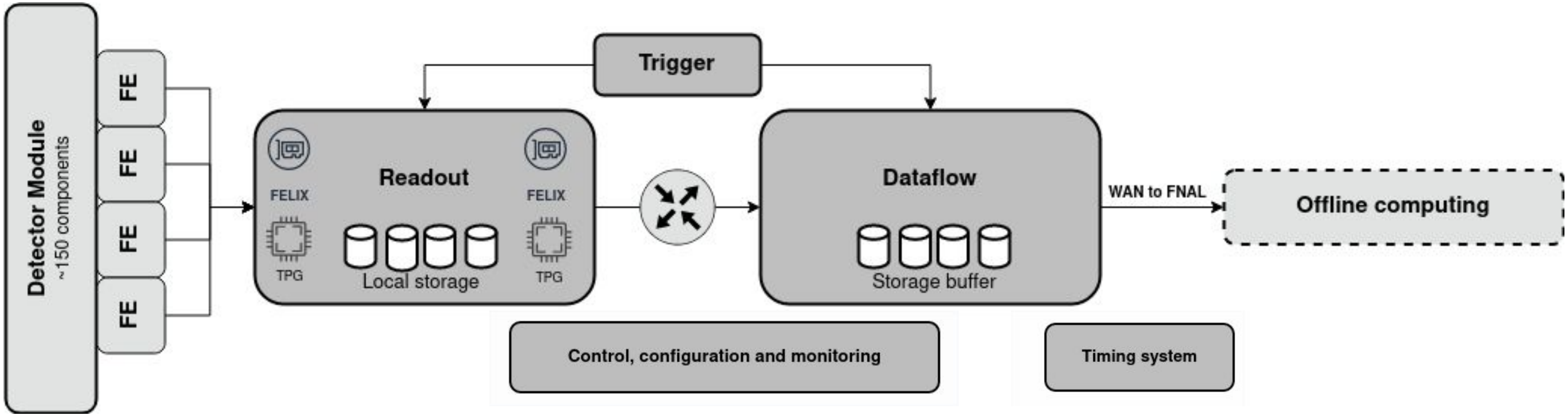


Transfer recorded data to Fermilab computing infrastructure

- Total transfer of 30 PB/year (across all detector modules)

DUNE Data AcQuisition system (DAQ)

- Modular nature of the apparatus allows splitting a cryostat in ~150 components



More systems are included in the DAQ:

- **Run Control**: steer the data acquisition in coherent way, propage commands to all the applications
- **Configuration**: description and configuration of all the applications running in the DAQ
- **Operational Monitoring** and **Data Quality Monitoring**: check the quality of the data and the operational metrics
- **Timing system**: provide synchronization to a common clock

Physics constraints on the DUNE DAQ

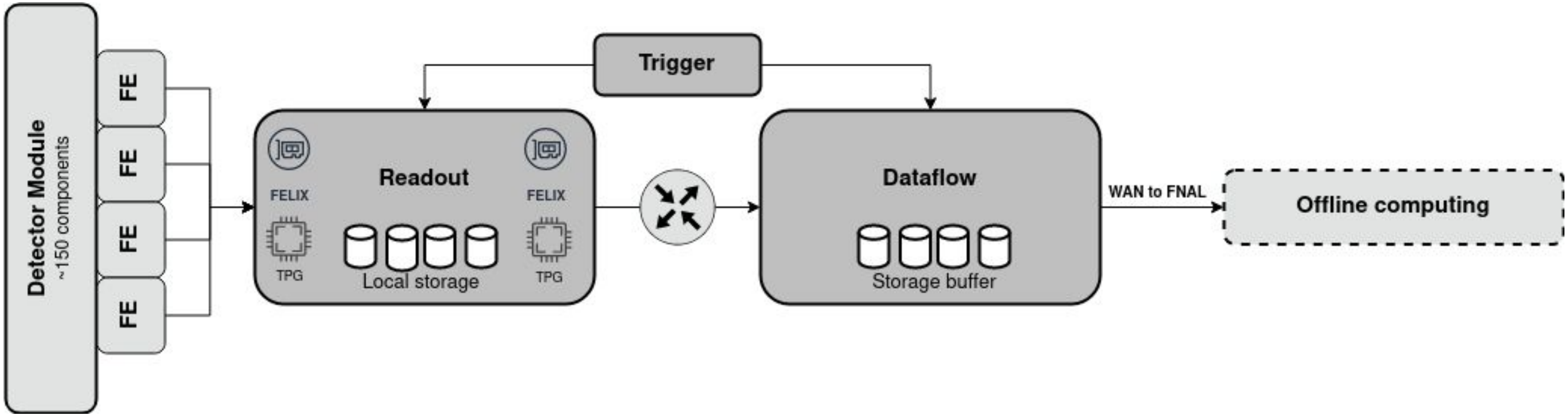
The physics goals of the DUNE experiment heavily drive the DAQ design

- Wide physics program results in the study of many different types of events
  - Support data taking over a wide energy spectrum
    - DAQ must support a very wide range of readout windows
    - Trigger system will need both a self triggering mechanism for the many low-energy deposits as well as a triggering system for the high energy (>100 MeV) interactions
  - **Storage system and buffering becomes crucial to support all data taking operations**
- Detector located 1.5 km underground
  - No quick access and no large host lab in the vicinity !
    - DAQ must be fully configurable, controllable and operable from remote locations
    - Quality monitoring and operational monitoring become essential
    - Automated error detection and recovery capabilities


PR-538
C.M.U. 40/10t
DE SERIE 10886
ANNEE 2015
Focus of this seminar

Source: CDS

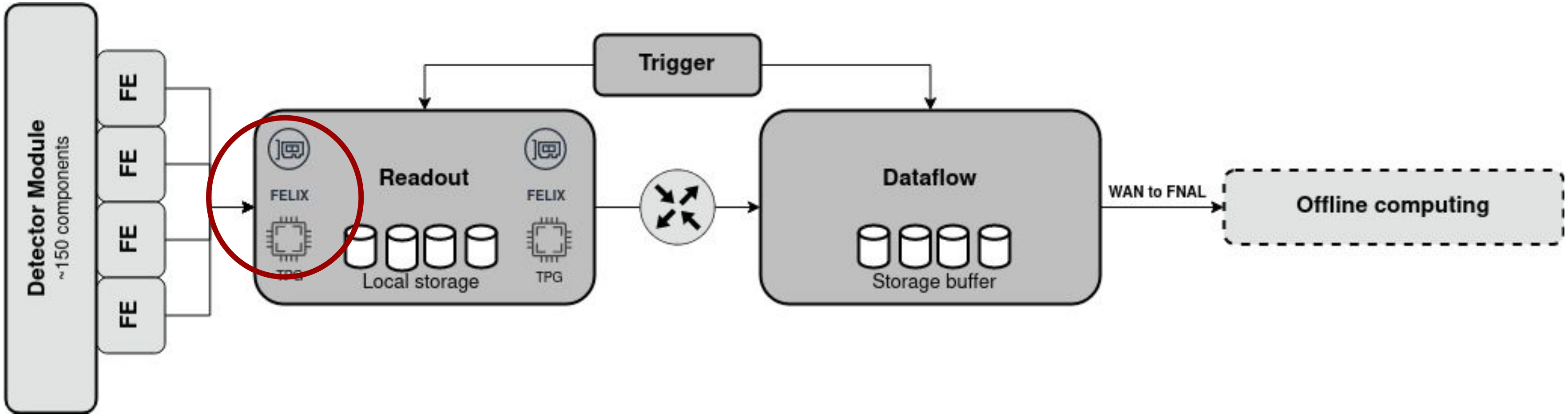
Focus of this seminar



In this seminar, we will focus on three components of the system:

- Real-time processing and streaming of interesting data regions for trigger decisions
- Local storage buffer for Supernova events
- Particle identification in liquid argon detectors

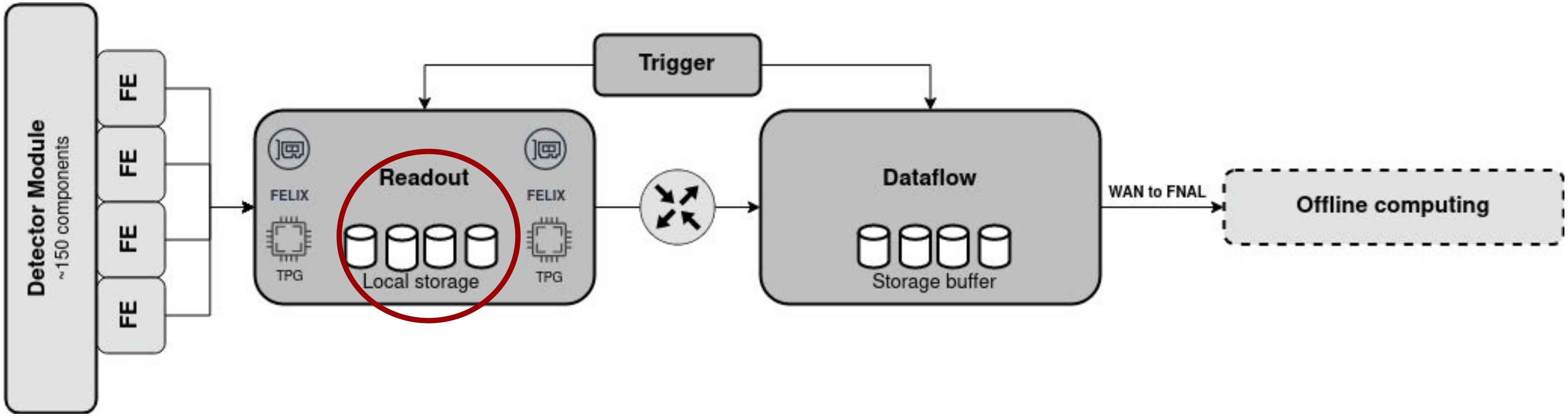
Focus of this seminar



In this seminar, we will focus on three components of the system:

- Real-time processing and streaming of interesting data regions for trigger decisions
- Local storage buffer for Supernova events
- Particle identification in liquid argon detectors

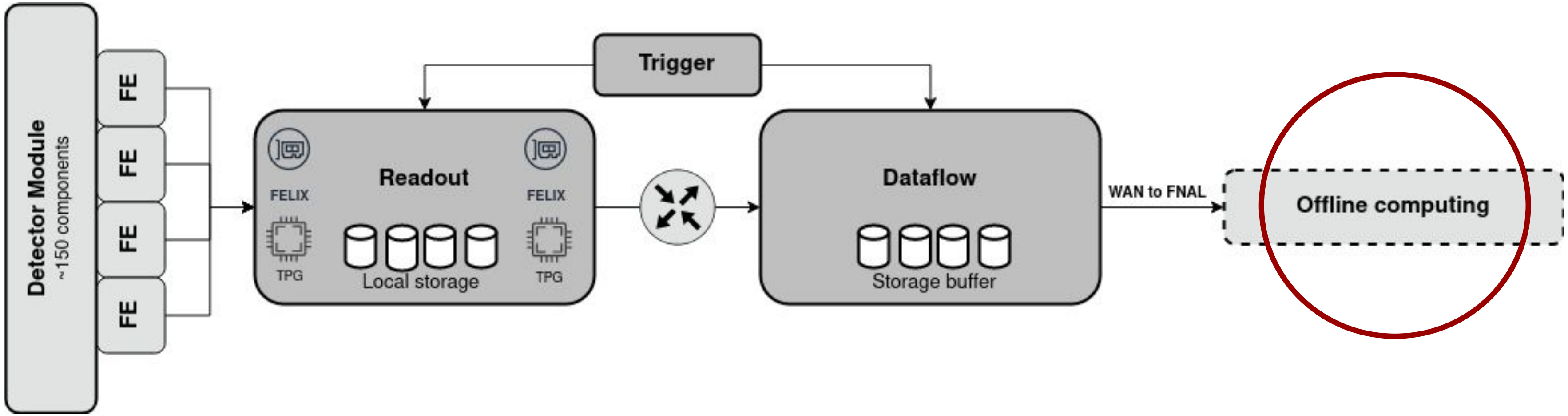
Focus of this seminar



In this seminar, we will focus on three components of the system:

- Real-time processing and streaming of interesting data regions for trigger decisions
- Local storage buffer for Supernova events
- Particle identification in liquid argon detectors

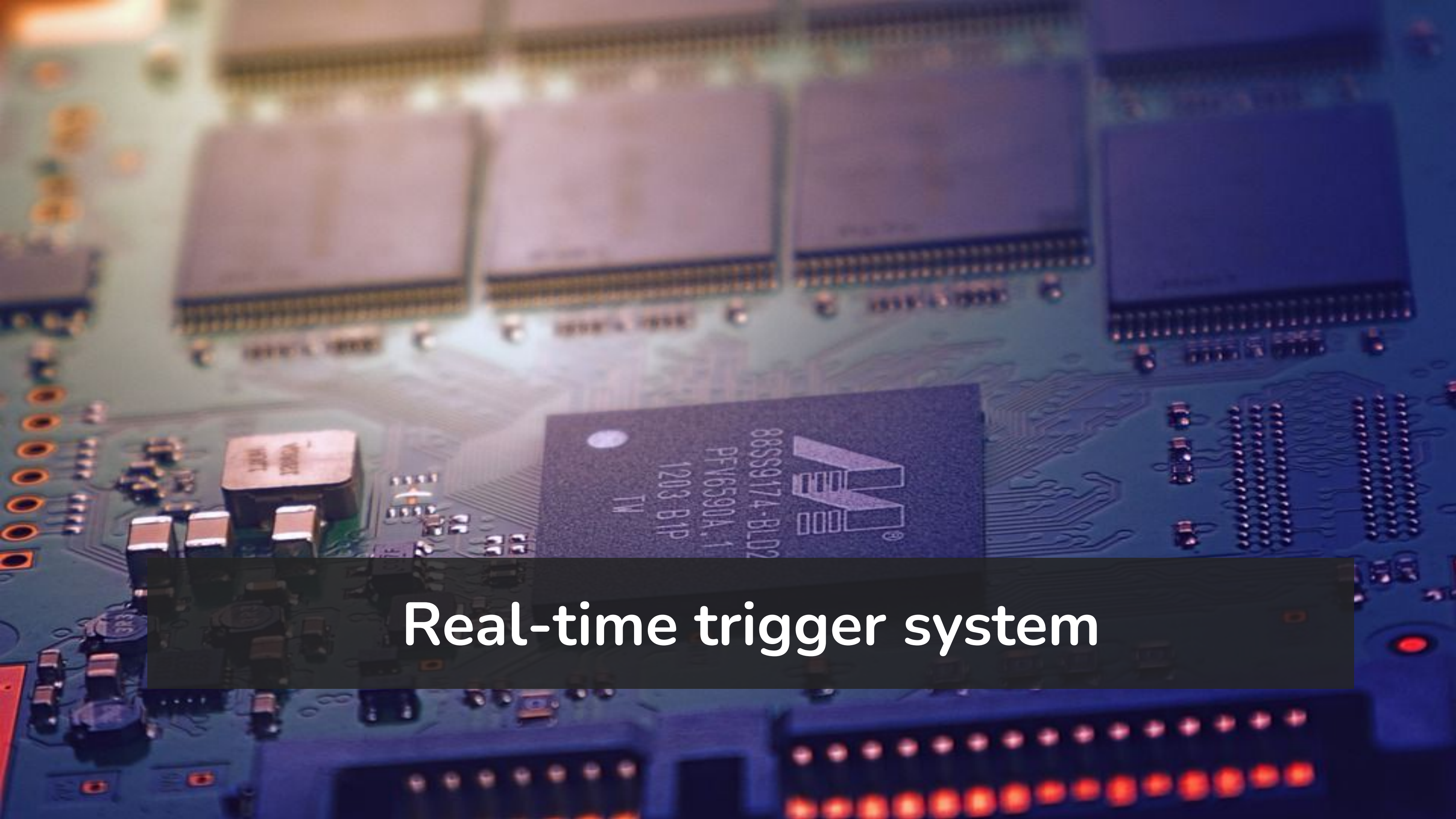
Focus of this seminar



In this seminar, we will focus on three components of the system:

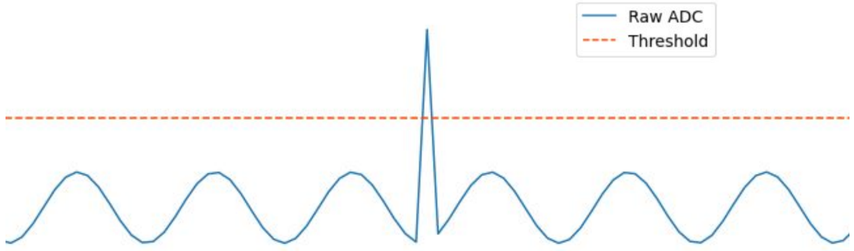
- Real-time processing and streaming of interesting data regions for trigger decisions
- Local storage buffer for Supernova events
- Particle identification in liquid argon detectors

Real-time trigger system



Real-time trigger system - Motivation

- A Real-time triggering system is a crucial for the success of the DUNE physics program
- Hit finding is the process of identifying a signal (Trigger Primitive) above a threshold
  - Objective is to trigger on signal and discard noise data

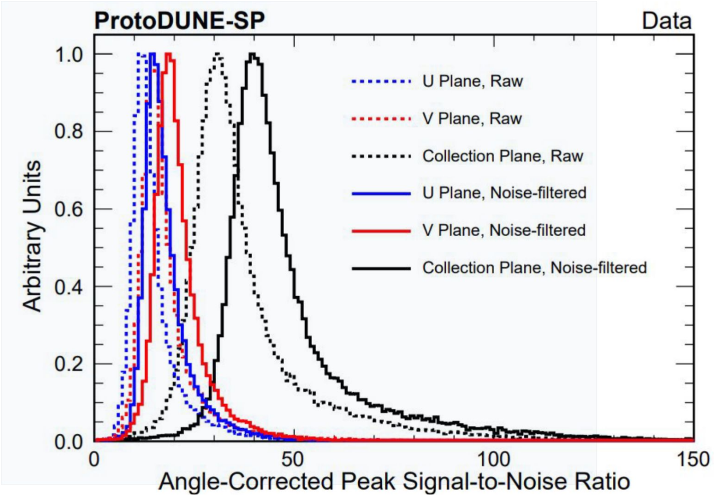


- **Advantages**
  - Improved clustering/triggering efficiencies for events with low hit multiplicities
  - Improve signal-to-noise ratio for low energy regions
  - Trigger based on Regions-of-Interest (3-view of the hit information becomes possible)
    - Select only regions of the LArTPC where an activity has been detected

Real-time trigger system - Challenges

- Identifying a signal across all detector planes is challenging
  - Noise RMS is not constant across all three detector planes
  - Lower hit finding efficiency because signal to noise ratio is lower across different detector planes

- **Requirements** for a **real-time triggering** algorithm for the DUNE **DAQ**
  - **Accuracy**: maximize the signal to noise ratio, in real-time, for any type of signal
  - **Universality**: all the waveforms from the detector front-end electronics must be treated equally
  - **Frugality**: the algorithm needs to process signals at the detector sampling rate of 2 MHz



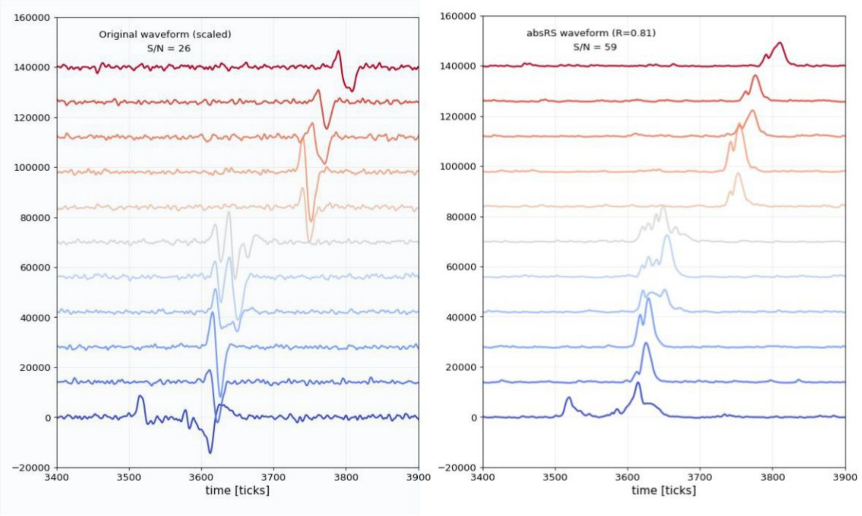
Signal is 40x higher than noise for the collection plane

Absolute Running Sum Algorithm

- Absolute Running Sum (AbsRS) integrates the incoming signal using the **absolute value** to compensate for negative pulses, a scale factor **s** and a weighting parameter **R**

$$y_n = R \cdot y_{n-1} + \frac{|x_n|}{s}$$

- Simulation studies show that signal-to-noise improve by a factor of 2 (**universality**)
- The algorithm works on all planes of the LAr TPC detector with satisfactory S/N ratio (**accuracy**)



Source: data selection slides from Klaudia Wawrowska

Trigger Primitive Generation using AbsRS algorithm

Decode format from Front-End

- Decode ADC data using specific detector format
- Expand 14-bit ADCs to 16-bits

Find pedestal median and variance

- Pedestal subtraction is used to bring the waveform baseline to zero
- Use of a frugal streaming method [paper]**:** use one entry at at time and update the median if N entries are higher than a fixed threshold
- Variance given by interquartile range (IQR): difference between the 75th and the 25th percentiles

Apply AbsRS

- Apply the absolute running sum on the incoming ADC entry using information
- Resource intensive task. Performance impact: ~10% CPU utilization

RS pedestal subtraction

- Apply a pedestal subtraction on the RS value

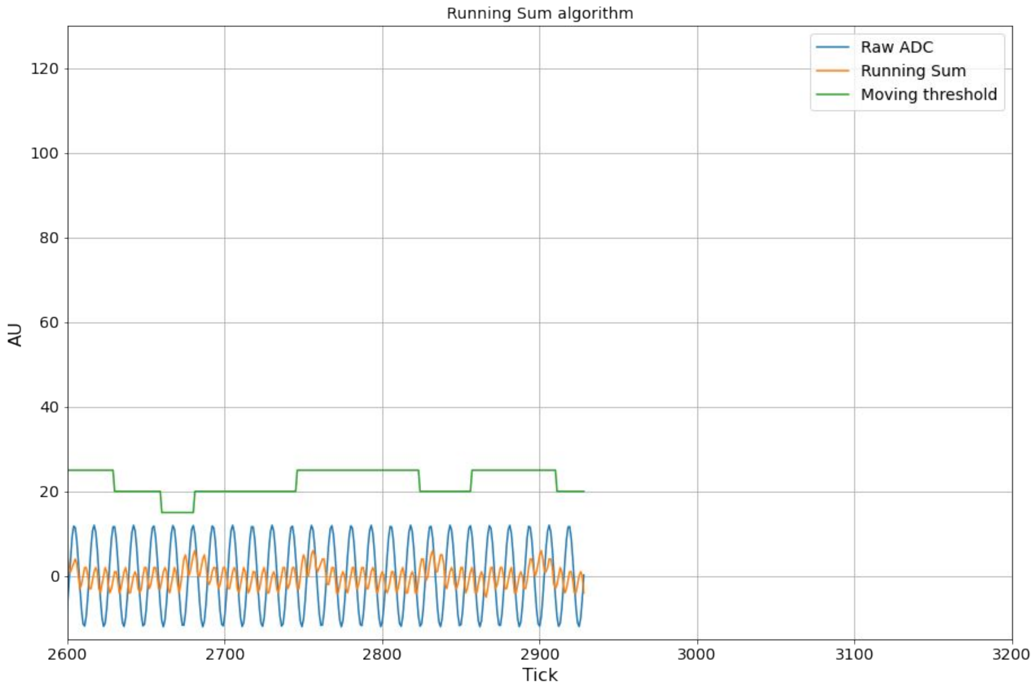
Evaluate TPs

- A TP is identified if the incoming signal is greater than a threshold
- Threshold is given by the IQR. Threshold “adapts” to the incoming data due to the frugal streaming

Parallelize using AVX2-based instructions: extension to x86 instructions in order to perform with a single instruction multiple data

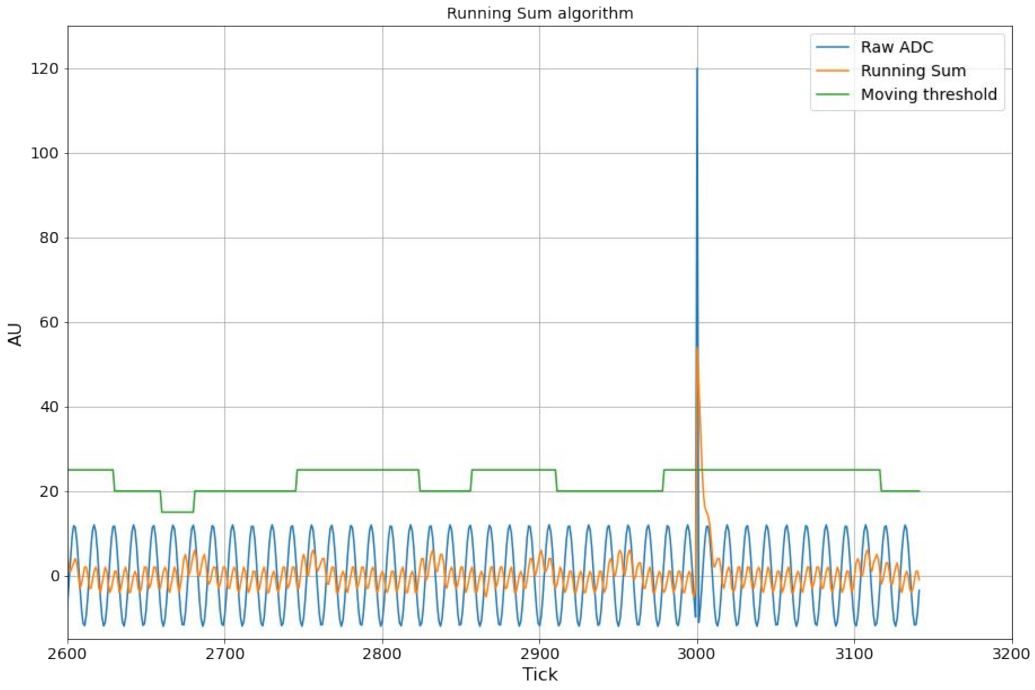
Response of the Absolute Running Sum to noise

- Simulation of noise data and investigation of the response of the Absolute Running Sum algorithm



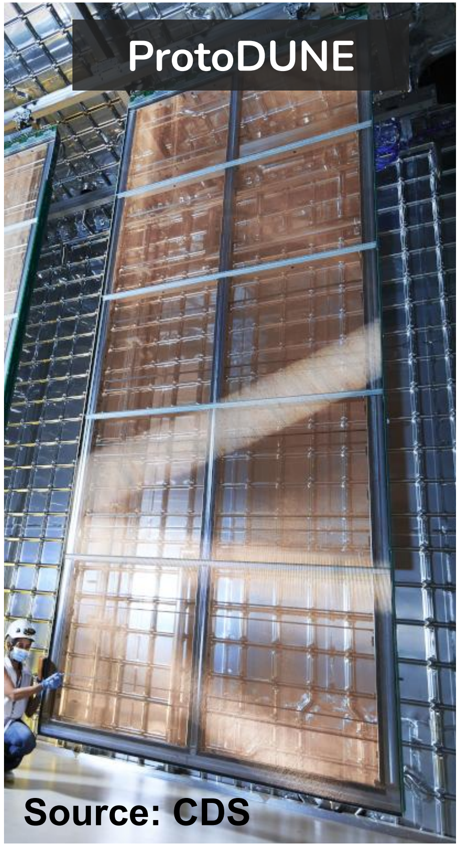
Response of the Absolute Running Sum to noise

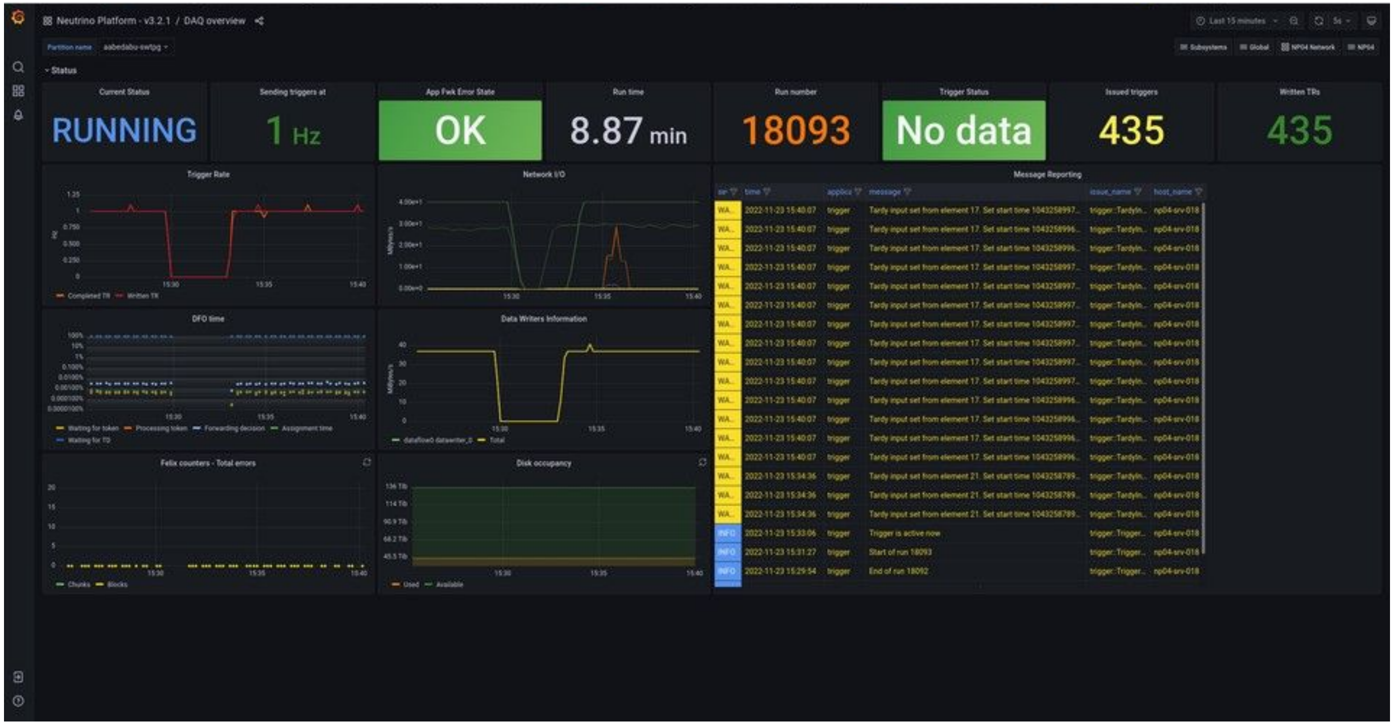
- Simulation of noise data and investigation of the response of the Absolute Running Sum algorithm



Absolute Running Sum in Prevessin

- The algorithm is running within the DUNE-DAQ software

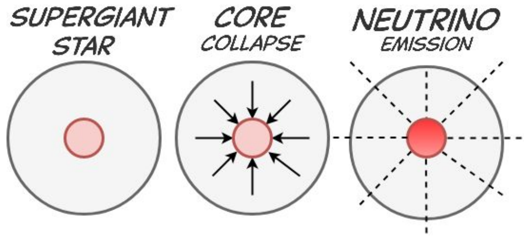




High-performance storage system

Storage buffer for Supernovae

Physics use-case



- DUNE objective: store Supernova Neutrino Burst (SNB) events
  - Detection of **rare**, **low energy** and **distributed** signatures

- **DAQ requirements:**
  - Sustain **high data rates**: 10 GB/s for each DAQ readout unit
  - Transient buffer of ~10 seconds to temporary store the data until a trigger signal is issued
  - On trigger: persist ~100 seconds (i.e. 150 TB per detector module)
    - **High data volumes:** minimum capacity of 2 TB per readout unit

- **Solution**:
  - Use commercial-off-the-shelf (COTS) storage devices to meet the performance demands of the DUNE experiment

Persistent memory devices

Device with DRAM-like properties!

- Technology candidate to store the Supernova data: **persistent memory devices (PMEMs)**
  - **Non-volatile memory technology**: memory devices capable of permanently storing data
  - Higher storage capacity (up to 9 TB per server) and higher latency (1 μs) compared to DRAM

- **DAQ application** [paper]
  - Implementation of the supernova storage buffer with persistent memory devices could be a promising solution
    - Better than DRAM devices or single solid-state drives
  - EP-DT-DI **DAQlab** has a cluster of 6 servers with PMEMs !

- Benchmarking results:
  - Disk performance: write throughput of ~10 GiB/s
    - Target range needed for the DUNE local storage system

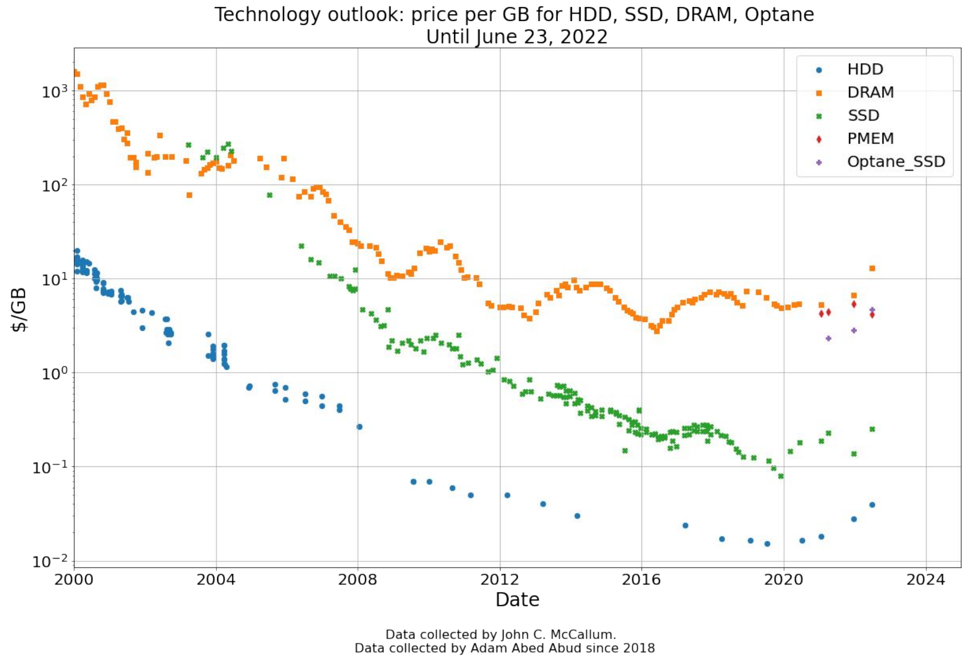
Integration with ProtoDUNE

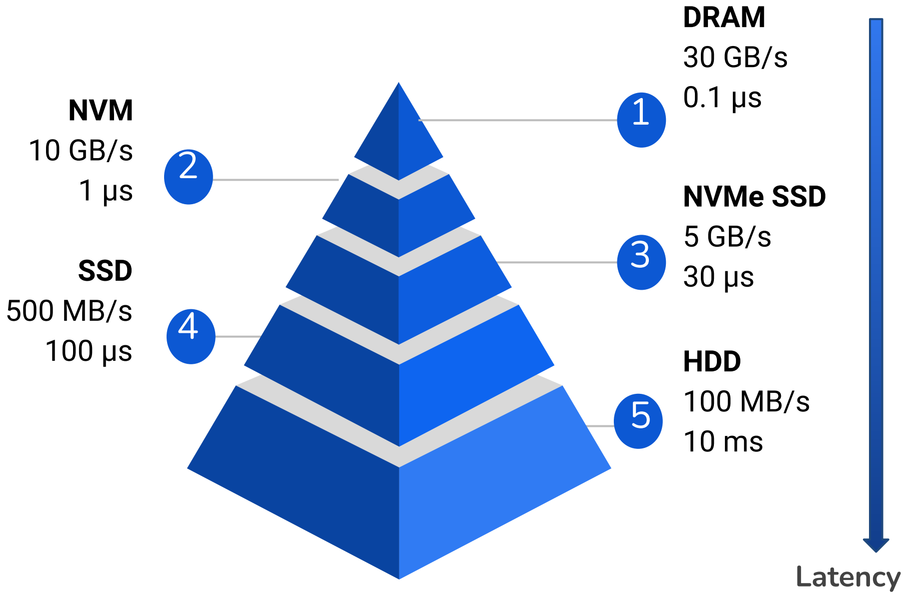
- **Testing procedure** :
  - Emulate the data workload from the Readout system and write into PMEMs ✔
  - Use the DUNE Readout system in emulation mode: ✔
    - Data generated from the electronics and not by the Front-End electronics
  - Full integration with a single ProtoDUNE readout unit ✔

- Each detector optical link is processed independently and assigned to a thread
  - Setup **CPU affinity** of the executing thread to the right NUMA node
- **Software optimizations** needed to achieve the target throughput:
  - **Direct I/O writing**: indicates that data is copied directly to the device, bypassing kernel buffer and the operating system's page cache

Evolution of storage technologies

- **Bandwidth of storage technologies is evolving fast**. Write throughput for NVMe SSD devices
  - In **2018**: ~1.5 GB/s
  - In **2022**: ~5.0 GB/s
- Use COTS technologies as much as possible in DAQ systems

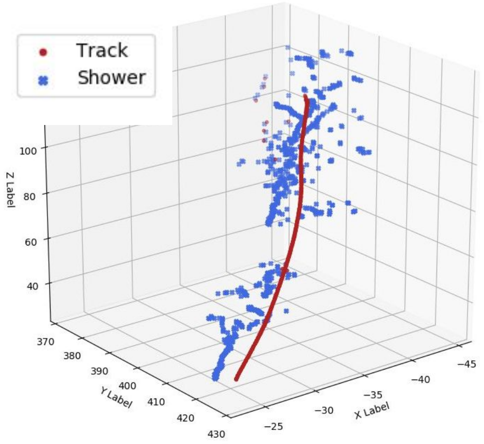




Particle identification in a LAr detector

Source: CDS

Track vs Shower identification
Introduction



- Particles produced from the neutrino interaction with argon have to be properly identified (**particle identification**) and their energies need to be measured (**reconstruction**)

- **Challenges:**
  - Correctly identify particle showers initiated from photons or electrons
    - Many DUNE beam events have this signature (e.g. decay of π^0)
    - Correctly identify track events (muons, pinos) from shower events (electrons, photons)
  - Energy reconstruction of an event involves correctly identifying all hits produced in an interaction
    - Improve energy resolution of the event

- **Solution:** use state-of-the-art tools for the classification task between track and shower particles

Deep Learning method

- **Problem statement:** Binary classification task (track vs shower) from sparse data
- Approach based on **Submanifolds Sparse Convolutional Networks** (arXiv:1706.01307, Facebook Research)
    - Convolutional neural networks applied on sparse data
        - Suitable to (Proto)DUNE use-case: large detector and few interactions
    - From paper: "Testing results show that they perform on par with state-of-the-art methods (sparse convolutional neural networks) but **require substantially less computation**"
    - Similar approach adopted in MicroBooNE (arxiv:2012.08513)

- Use **ProtoDUNE** as testing platform for the algorithm development and testing
    - Use of both MC data and from the 2018 data taking session

Training and validation

- Input features: identified in the hits 9 discriminating features between track and shower particles
  - Angle between neighbouring hits;
  - Dot product between neighbouring hits
  - Charge deposition of single hits
  - Number of neighbouring hits within a certain distance. R = [3 cm, 10 cm, 30 cm]
  - Total charge over a certain distance. R = [3 cm, 10 cm, 30 cm]

- Trained the model on **40k events (> 10M hits)**

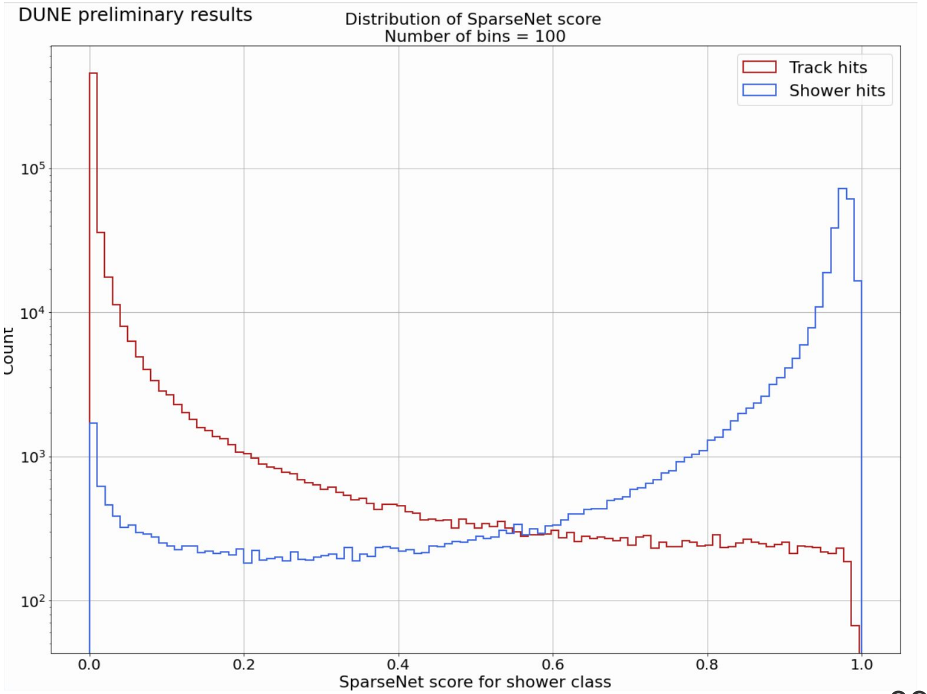
	Purity [%]	**Efficiency [%]**
Track	96.3	97.3
Shower	90.8	87.7

Inference on ProtoDUNE data

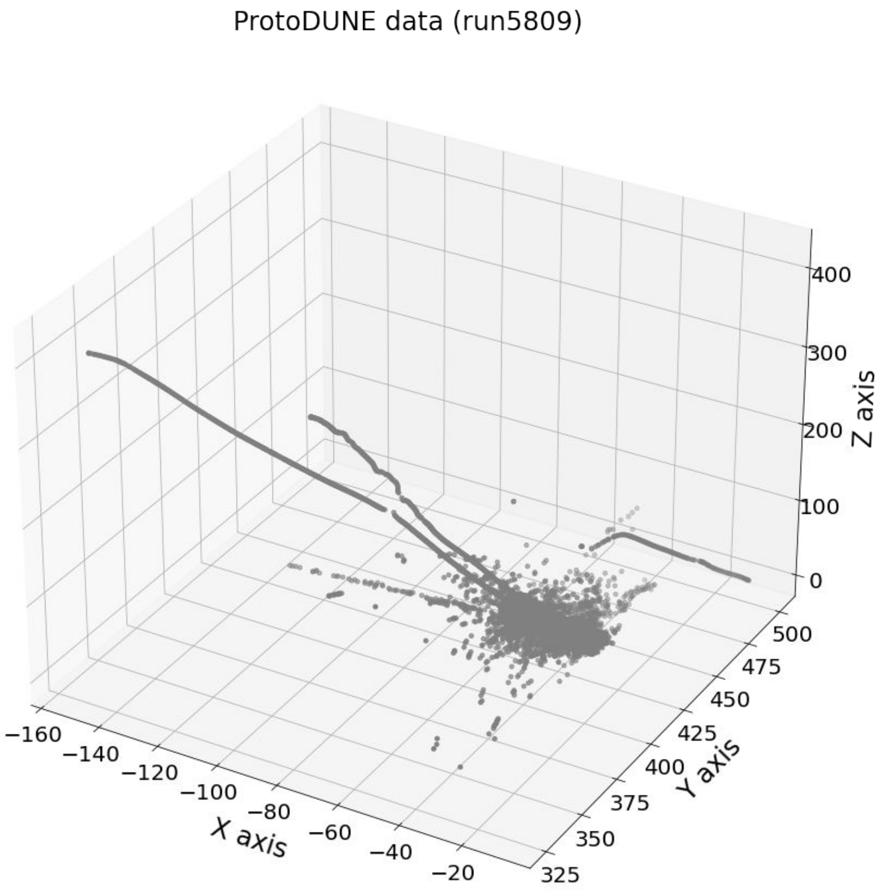
- The network output is a value between 0 (track) and 1 (shower)
- Hit level analysis: for each hit in the dataset, investigate the SparseNet output for predicting both track (red) and shower (blue) particles

- Inference results on the ProtoDUNE data show purities and efficiencies above 90%
  - Most of the track and shower hits have been correctly identified
  - Outperforming currently adopted algorithm in ProtoDUNE analysis !

	Purity [%]	Efficiency [%]
Track	97.7	97.8
Shower	95.6	95.3

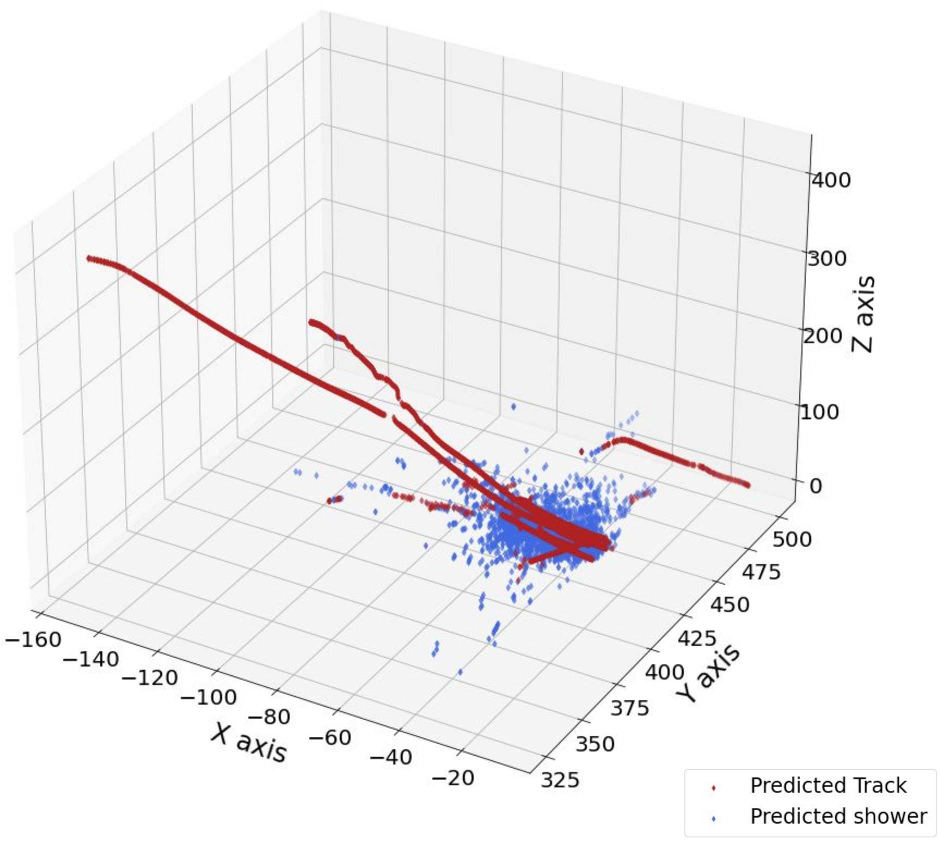


Inference in real life



Inference in real life

ProtoDUNE data (run5809)



Conclusion

- Neutrinos are elusive particles that still hide many mysteries
- Wide physics program of the DUNE experiment drives the design of its Data Acquisition system
- Different strategies have been investigated to support the physics program of the experiment
  - Real-time triggering mechanisms for low energy deposits across the detector
    - System tested and deployed on a section of the ProtoDUNE detector
  - Local high performance storage technologies for recording Supernova events
    - Tested and integrated modern storage technologies into the DAQ system with successful results
  - Use of particle identification methods to improve energy reconstruction of neutrino events
    - Successfully investigated innovative Deep Learning method specifically designed for large and sparse detector like the DUNE experiment

CERN

home.cern

Neutrinos: what is unknown?

- What is the neutrino mass ordering?
    - We do not know the mass of the three observed neutrinos but only mass differences
    - We do not know which neutrino is the lightest or the heaviest
- Is there CP violation in the neutrino sector? Could it be responsible for the matter/antimatter asymmetry in the universe?
    - We do not know if neutrinos and antineutrinos behave alike
- Do additional neutrinos exists?
    - (Inconclusive) evidence suggest the existence of a “right-handed” sterile neutrino
- Can we detect neutrinos from Supernova Burst events?

Other interesting storage solutions

- Take advantage of PCIe 4 storage adapters
  - Connect multiple SSD drives together: up to 4 x PCIe 4.0 devices
- A single PCIe 4 NVMe SSDs provides a sequential writing throughput of ~4.5 GB/s

- Connecting 4 drives with RAID-0* provides the target throughput needed for the DUNE local storage
  - Interesting solution for the DUNE DAQ local storage buffer
- Testing with the whole ProtoDUNE apparatus
  - Successfully tested the local storage system with a single readout unit ✅

RAID0 or disk striping = process of dividing data into several data blocks (or stripes) and distributing them across multiple storage devices.