DUNE FD DAQ Data Selection System

Georgia Karagiorgi
Columbia University

DUNE DAQ Conceptual Design Review @ CERN
December 3-4, 2018

Data Selection System (Trigger +…)

- Responsible for directing online data selection for permanent recording/offline storage from all DUNE active detectors (TPC, PDS)
- Facilitates triggering (via online data processing) and down-selection (filtering) of selected trigger records.

- Data selection decisions by this system are informed by:
  - External inputs: accelerator timing, external timing signals, external calibration signals, detector configuration, detector conditions
  - Active detector data itself: TPC and PDS raw data

Requirements and Specifications

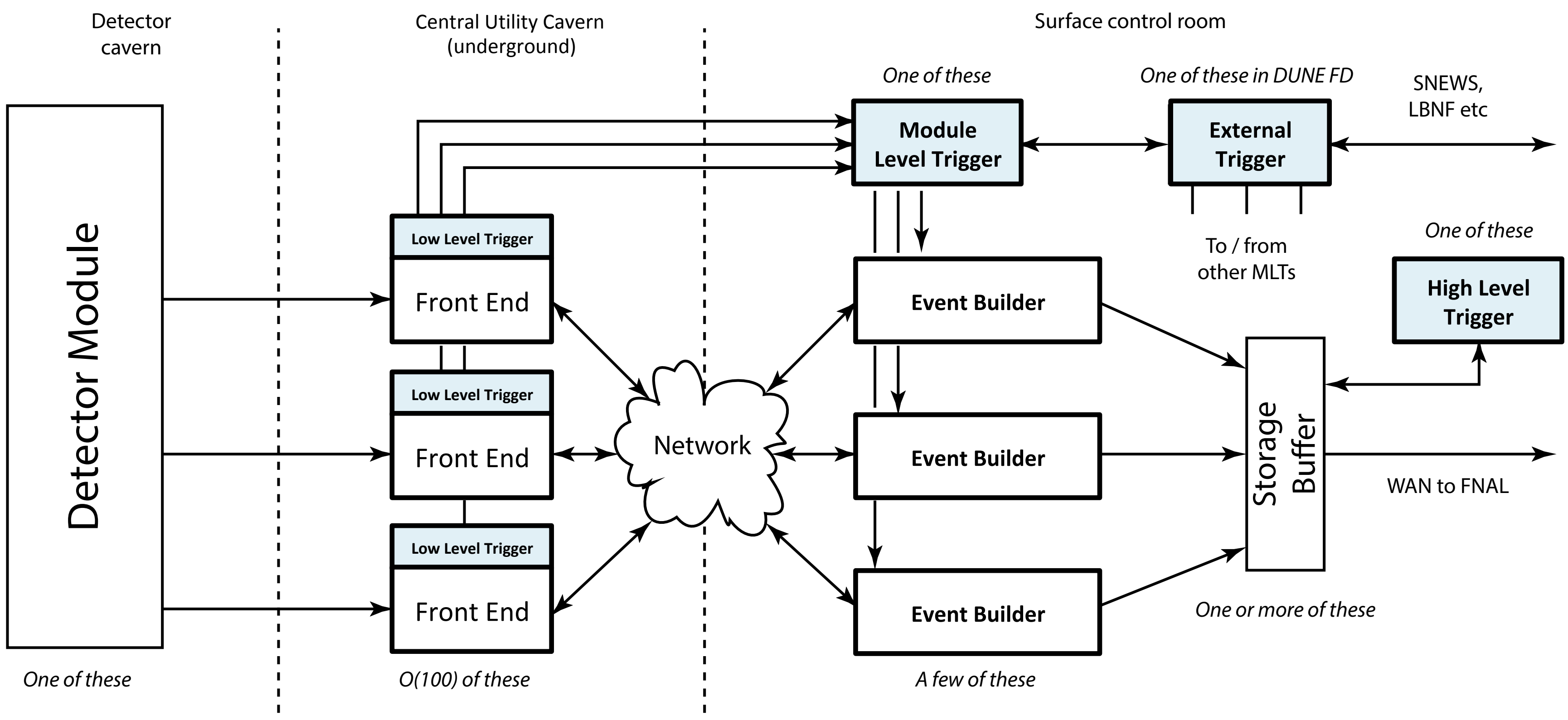
Requirement	Description	Value
Off-beam High-energy Trigger	The detector shall trigger on the visible energy* of underground physics events from decays or interactions within the active volume with high efficiency.	> 100MeV
Off-beam Low-energy Trigger	The detector shall be capable of triggering on the visible energy of single low energy neutrino interactions inside the active volume.	> 10MeV
Trigger for Beam	The detector shall trigger on the visible energy of beam interactions within the active volume with efficiency high enough that it has a sub-dominant impact on physics sensitivity.	> 100 MeV
Trigger for Calibration	The detector shall provide triggers to and trigger on calibration stimuli and tag the data from these triggers as such	
Trigger for Supernova Burst	A trigger shall be generated when a collection of signals is detected that constitute a candidate supernova burst with high galactic coverage*, while meeting offline storage requirements and overall bandwidth limitations.	
Physics Event Record	The DAQ shall merge data into a form suitable for offline analysis. Furthermore, tags shall be provided to allow the data collection conditions at the time and the livetime to be determined.	
DAQ Deadtime	The DAQ shall operate with deadtime that does not contribute significantly to overall loss of detector livetime.	

*Visible energy = deposited energy in the active volume as ionization and/or scintillation
*Galactic coverage = SBN probability-weighted efficiency, integrated over the physical extent of the Milkyway

System Conceptual Design

- The DS system design follows a **hierarchical level structure**:
  low-level processing → low-level decision
  → module-level processing → module-level decision ("trigger")
  → high-level processing → down-selection ("filter")
- Data processing is executed in **various stages of firmware and/or software**.
- Overall design **philosophy**: redundancy and scalability of processing resources, and flexibility on where the core processing payload functionality resides.
- Design is **continually evolving** → **conservative approach** for design validation for the TDR: Triggering primarily driven by TPC information; PDS information is potentially additional gain (simulations in progress).
- Design must be **directly applicable** to DP-TPC, SP-PDS, DP-PDS.

System Conceptual Design



System Conceptual Design (SP-TPC)

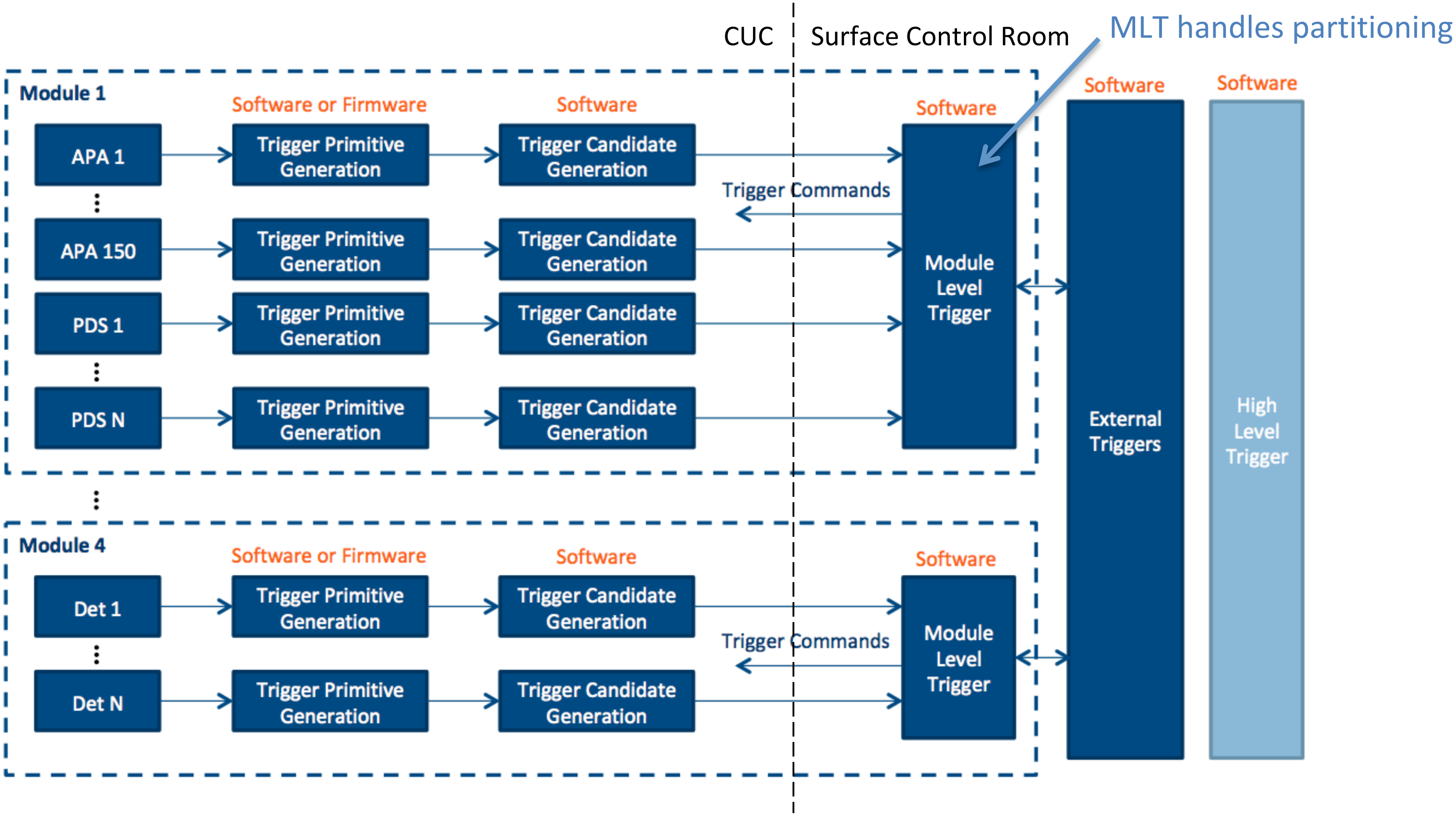


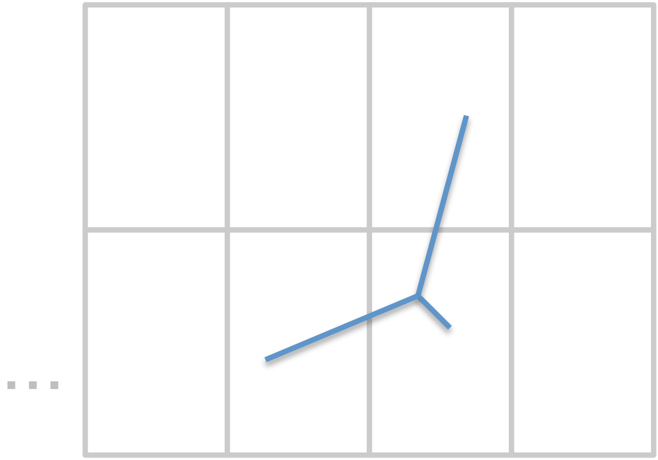
Figure 1.5: Block diagram of DUNE DAQ Data Selection sub-system, illustrating hierarchical structure of sub-system design.

System Conceptual Design (SP-TPC)

- **Trigger Primitive (TP):** defined nominally on a "per collection channel" basis ("hit on a wire"); TP summary includes channel address, time of hit, time over threshold, ADC sum, error flag (20 Bytes total)
- **Trigger Candidate (TC):** defined nominally on a "per APA" basis ("cluster of hits"); multiple definitions (e.g. low energy, high energy); TC summary includes APA address, timing, summed ADC, … (TBD)
- **Trigger Command (MLT):** defined nominally on a "per 10-kton-module" basis. The MLT is where information from TPC TC, PDS TC, and external inputs is combined to decide when and which parts of the detector are to be read out, and how (localized readout mode or supernova burst mode)
- **External Trigger (EXT):** common among the four DUNE FD modules; feeds information to MLT from other modules and vice versa (e.g. SNB Trigger Command), as well as from module-external sources: SNEWS, calibrations, accelerator,...
- **High Level Trigger (HLT):** Lives post-Event Builder; serves to **down-select** and thus limit total triggered data rate to offline (filtering, lossy data reduction, further event classification/identification); longer latency→higher sophistication.

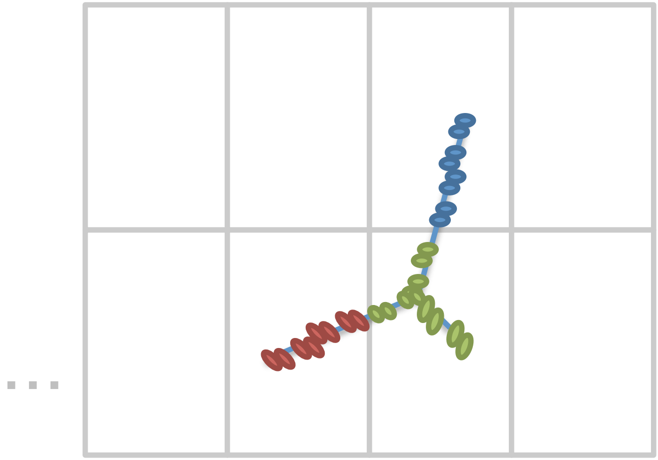
Functional Example: “localized activity”

- Localized events (in space and time): e.g. typical atmospheric neutrino interaction



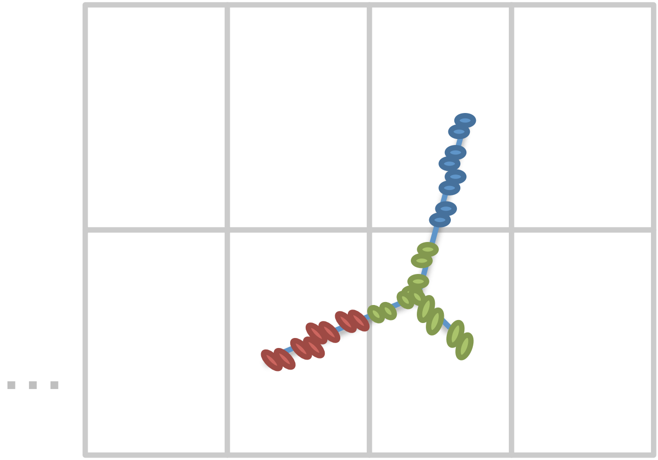
Functional Example: "localized activity"

- Localized events (in space and time): e.g. typical atmospheric neutrino interaction
  - TP's generated from a large number of wires per APA (over a few APAs)
  - Multiple TC reported to MLT (one or more per APA over a few APAs)



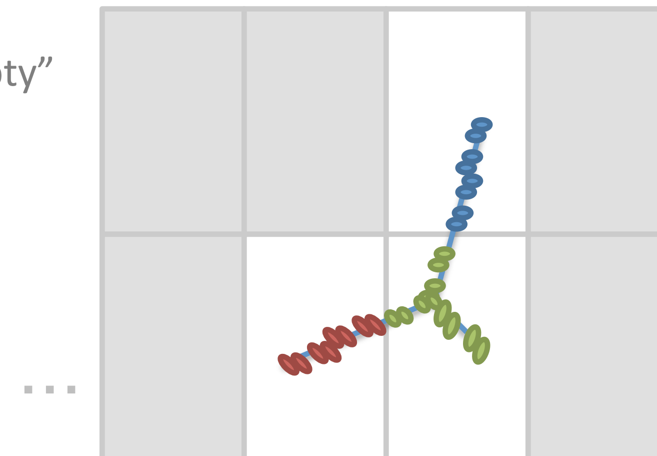
Functional Example: "localized activity"

- Localized events (in space and time): e.g. typical atmospheric neutrino interaction
  - TP's generated from a large number of wires per APA (over a few APAs)
  - Multiple TC reported to MLT (one or more per APA over a few APAs)
  - MLT is aware of detector partition and operating state of each APA and accelerator timing → decides it's a valid HE candidate event
  - MLT sends trigger command (interpreted as 5.4ms readout over entire module) to Back-end System
  - Data retrieved → Event Builder



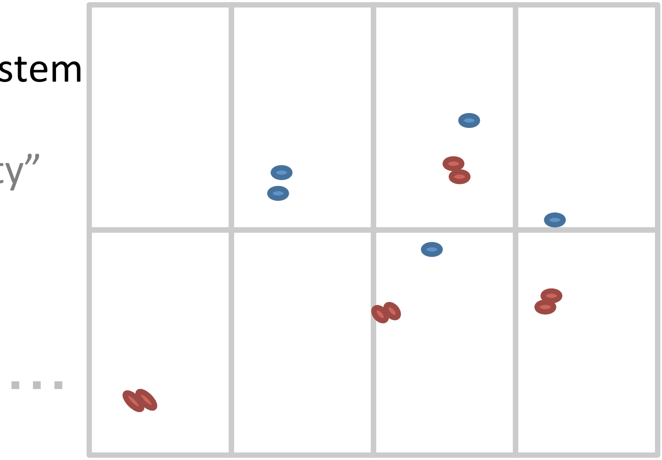
Functional Example: “localized activity”

- Localized events (in space and time): e.g. typical atmospheric neutrino interaction
  - TP’s generated from a large number of wires per APA (over a few APAs)
  - Multiple TC reported to MLT (one or more per APA over a few APAs)
  - MLT is aware of detector partition and operating state of each APA and accelerator timing → decides it’s a valid HE candidate event
  - MLT sends trigger command (interpreted as 5.4ms readout over entire module) to Back-end System
  - Data retrieved → Event Builder
  - HLT can optionally remove APAs which are “empty”



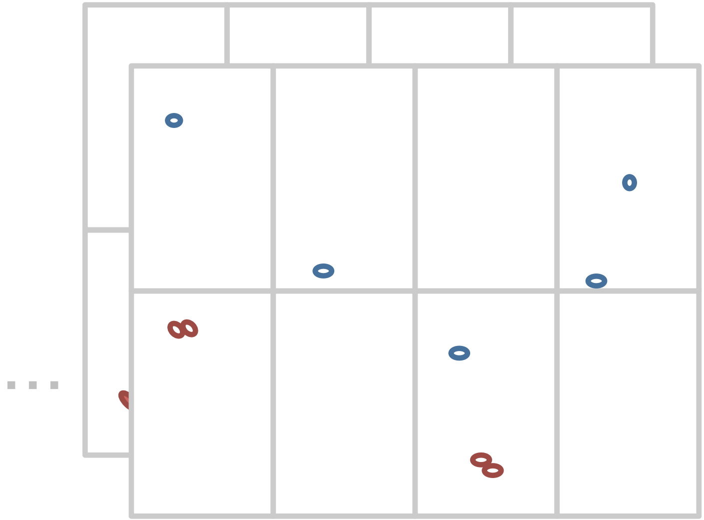
Functional Example: “extended activity”

- Extended events (in space and time): supernova burst (SNB)
  - TP’s generated from a small number of wires per APA (over multiple APAs)
  - Multiple TC reported to MLT (~one per APA over multiple APAs and over an extended period of time (longer than 1 drift))
  - MLT is aware of detector partition and operating state of each APA and accelerator timing → decides there are valid LE candidate events → decides this is a valid SNB event
  - MLT sends trigger command (interpreted as 100s readout over entire module) to Back-end System
  - Data retrieved → Event Builder
  - HLT can optionally remove APAs which are “empty”



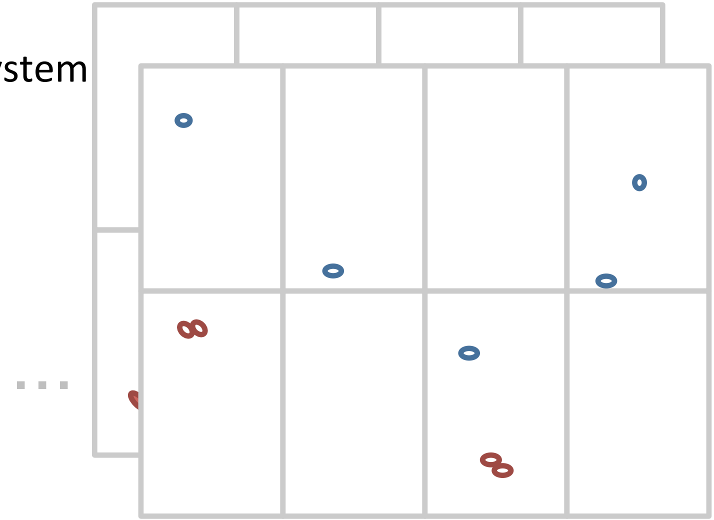
Functional Example: “extended activity”

- Extended events (in space and time): supernova burst (SNB)
  - TP’s generated from a small number of wires per APA (over multiple APAs)
  - Multiple TC reported to MLT (~one per APA over multiple APAs and over an extended period of time (longer than 1 drift))



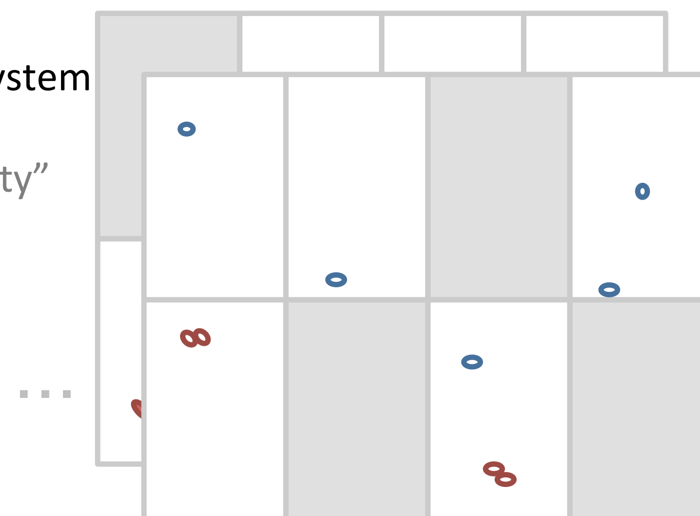
Functional Example: “extended activity”

- Extended events (in space and time): supernova burst (SNB)
  - TP’s generated from a small number of wires per APA (over multiple APAs)
  - Multiple TC reported to MLT (~one per APA over multiple APAs and over an extended period of time (longer than 1 drift))
  - MLT is aware of detector partition and operating state of each APA and accelerator timing → decides there are valid LE candidate events → decides this is a valid SNB event
  - MLT sends trigger command (interpreted as 100s readout over entire module) to Back-end System
  - Data retrieved → Event Builder



Functional Example: "extended activity"

- Extended events (in space and time): supernova burst (SNB)
  - TP's generated from a small number of wires per APA (over multiple APAs)
  - Multiple TC reported to MLT (~one per APA over multiple APAs and over an extended period of time (longer than 1 drift))
  - MLT is aware of detector partition and operating state of each APA and accelerator timing → decides there are valid LE candidate events → decides this is a valid SNB event
  - MLT sends trigger command (interpreted as 100s readout over entire module) to Back-end System
  - Data retrieved → Event Builder
  - HLT can optionally remove APAs which are "empty"



Interfaces

- Timing System (global timing system interface in place)
- Front-end (Trigger Primitive Generation)
  - Algorithms are part of Data Selection
  - Implementation (where algorithms live) in Front-end
- Back-end System
  - Configuration
  - Communication
  - Trigger Command (format)
- Calibration Systems
  - Via External Trigger

Technical Constraints and Key Challenges

- Sufficient data reduction needed to meet bandwidth constraints for data transmission to above-ground and to off-site
  - Scalability of processing resources, spare bandwidth (at cost)
- Power and space constraints underground in CUC and above ground
  - Development of power-efficient and scalable solutions, with demonstrated power usage performance
- Multi-component system and trigger latency
  - All hardware components synchronized and time-stamped to the same clock
  - Sufficient buffering (up to 10 seconds) built into front-end for SNB trigger
- Radiological backgrounds
  - Overdesign for higher MLT trigger rate, and implementation of more sophisticated (and slower) data selection at the HLT level
  - Variable thresholds and prescales allow throttling of rates
- Uncertainty on anticipated noise levels and features
  - Overdesign for higher MLT trigger rate, and implementation of more sophisticated (and slower) data selection at the HLT level
  - 30 PB/year cap is moderated by staged construction (first module has full 30 PB/year capacity)

System Design Development

- Now-2020 (EDR):
  - Algorithm development informed by simulations (ongoing)
  - Algorithm development informed by protoDUNE data (ongoing)
  - Implementation and testing on various COTS hardware (CPU, GPU) (ongoing)
  - Implementation and testing on custom FPGA boards (planned)

- 2020-2022 (PRR)
  - Full Vertical Slice demonstrators (planned)

Ongoing Validation and Simulation

- Algorithm validation on simulations and protoDUNE data:
  - **Trigger Primitives:** Single-wire baseline subtraction, filtering, and hit finding algorithms:
    - Demonstrated on simulation and protoDUNE data to perform well
    - Remains to be tested against coherent noise protoDUNE data, although coherent noise removal algorithms (offline) have been developed by MicroBooNE
  - **Trigger Candidates:** Generation for high-energy (atmospheric and cosmic) events:
    - Demonstrated on simulation to perform well, with low fake trigger rate from radiologicals
    - Remains to be tested for latency (target is less than 2.25ms drift time); demonstrated only for radiologicals
  - **Trigger Candidates:** Generation for low-energy (SN) events:
    - Demonstrated on simulation to perform well
    - Remains to be tested against a wide variety of SNB models

Ongoing Validation and Simulation

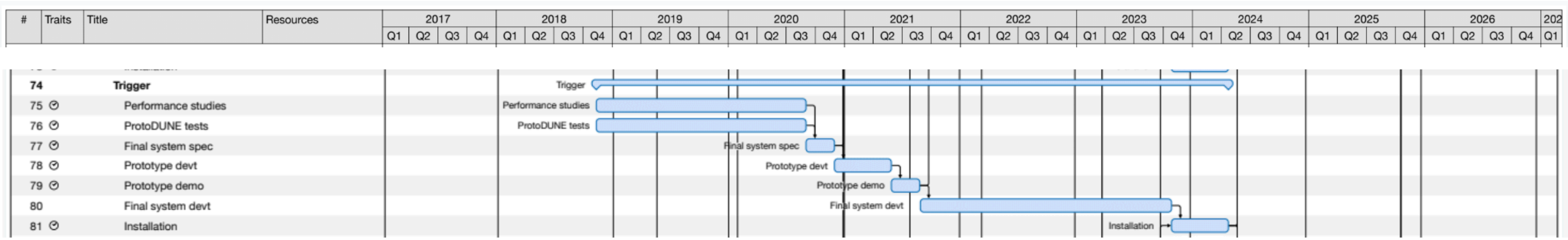
- Algorithm validation on simulations and protoDUNE data (cont'd):
  - **Trigger Command:** Generation for SNB's:
    - Demonstrated on simulation to perform well. High galactic coverage with target fake rate (1/month) achieved. Reasonable LMC coverage in addition.
  - **High Level Trigger:** Event classification performance:
    - Demonstrated on simulation to perform well
    - Remains to be tested for robustness against alternate noise/radiological models

Ongoing Validation and Simulation

- Implementation on FPGA, CPU, GPU:
  - Given sufficient resources, in principle all (FPGA, CPU, GPU) implementations can keep up with detector raw data rates.
  - Trigger Primitive generation demonstrated performance with 4 CPU cores/APA (Xeon Gold 6140 system) for <2.25ms
  - High Level Trigger filtering demonstrated performance with 1 GPU/APA (NVIDIA GeForce TITAN X 12GB) for 20-30ms

Schedule

- Expanded schedule available here: [link]



Development Plan

- Strategy (TDR):
  - Use ProtoDUNE as a design demonstration platform.
  - Use Vertical Slice Teststands for further development and testing of individual DAQ sub-systems as well as for key aspects of the overall DAQ.
  - Use FD MonteCarlo simulations and studies to augment demonstrations at ProtoDUNE and testing at Vertical Slice Teststands.
  - Cross-check against developments and measurements from other ongoing LArTPC experiments.

Development Plan

- Plans for further Data Selection development:
- TP validation in Software (ongoing)
- TC validation in Software (ongoing)
- HLT validation in Software (ongoing)

- **ProtoDUNE CPU Demonstration:** with addition of TP and TC generation
- **ProtoDUNE FPGA Demonstration:** with addition of TP and TC generation
- **ProtoDUNE DAQ Demonstrator:** Includes integration with calibrations, timing system; utilizes real data

- **TP validation in Firmware:** beyond protoDUNE, first in evaluation FPGA boards, then prototype co-processing modules (DAQ kits)
- **Test-bench Demonstration(s):** System prototype, Felix based (DAQ kit), utilizing fake data; testing up to MLT of data selection system.
- **Scalability:** "Horizontal Slice" demonstration starting at Trigger Candidate stage

Summary

- Data Selection System is currently in Conceptual Design stage
- Well-defined requirements, design philosophy
- Working on further substantiating design and ultimately demonstrating feasibility
- Strong institutional interest in US (leading design effort) and UK; working on securing resources for completing design (US)
- Goal is to have a technical design by Q4 2020

Backup Slides

Trigger Primitive definitions

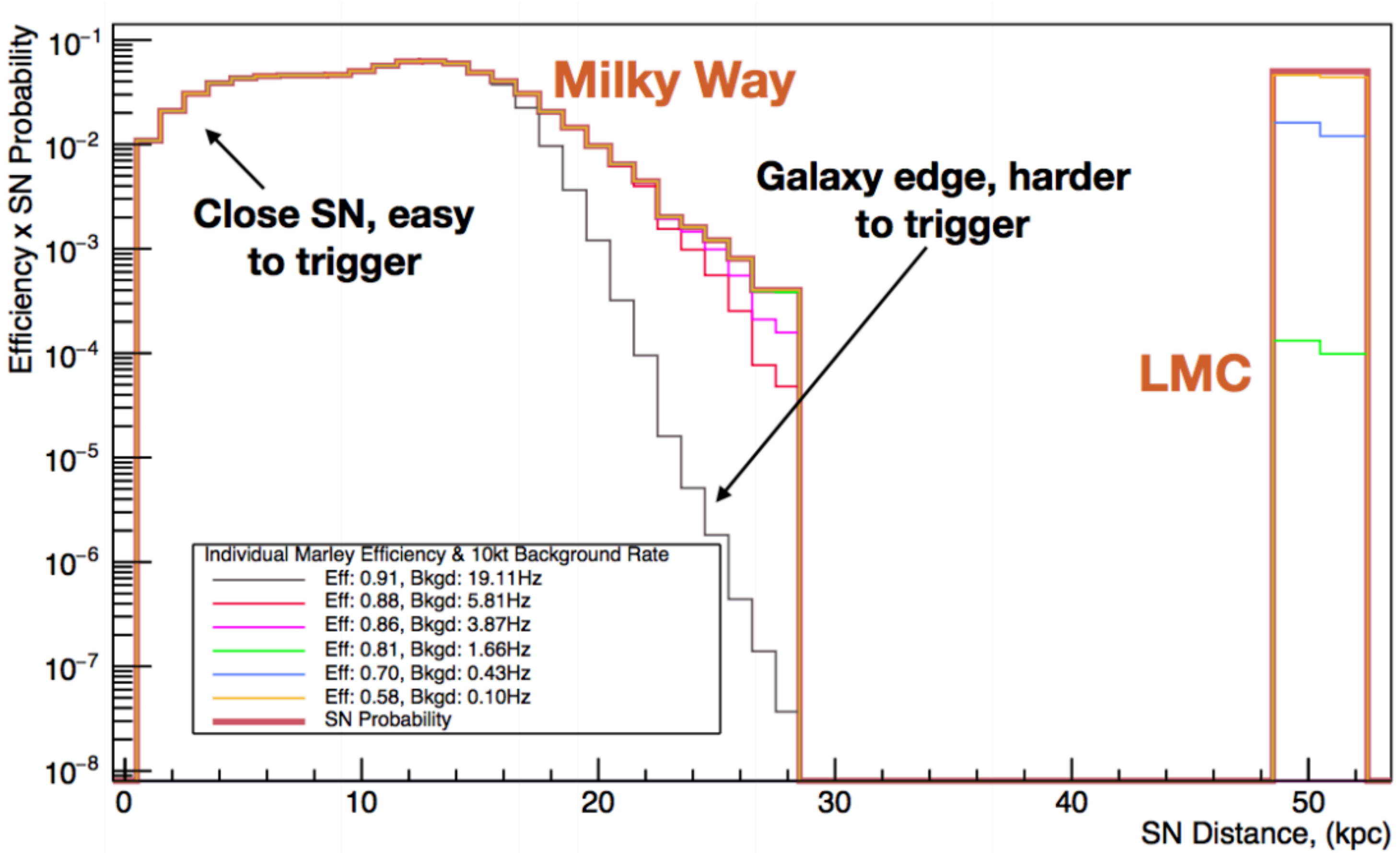
- Currently in TDR:

  - channel address $32bit$
  - time of hit $64bit$
  - time over threshold $16bit$
  - ADC sum $32bit$
  - error flag $16bit$

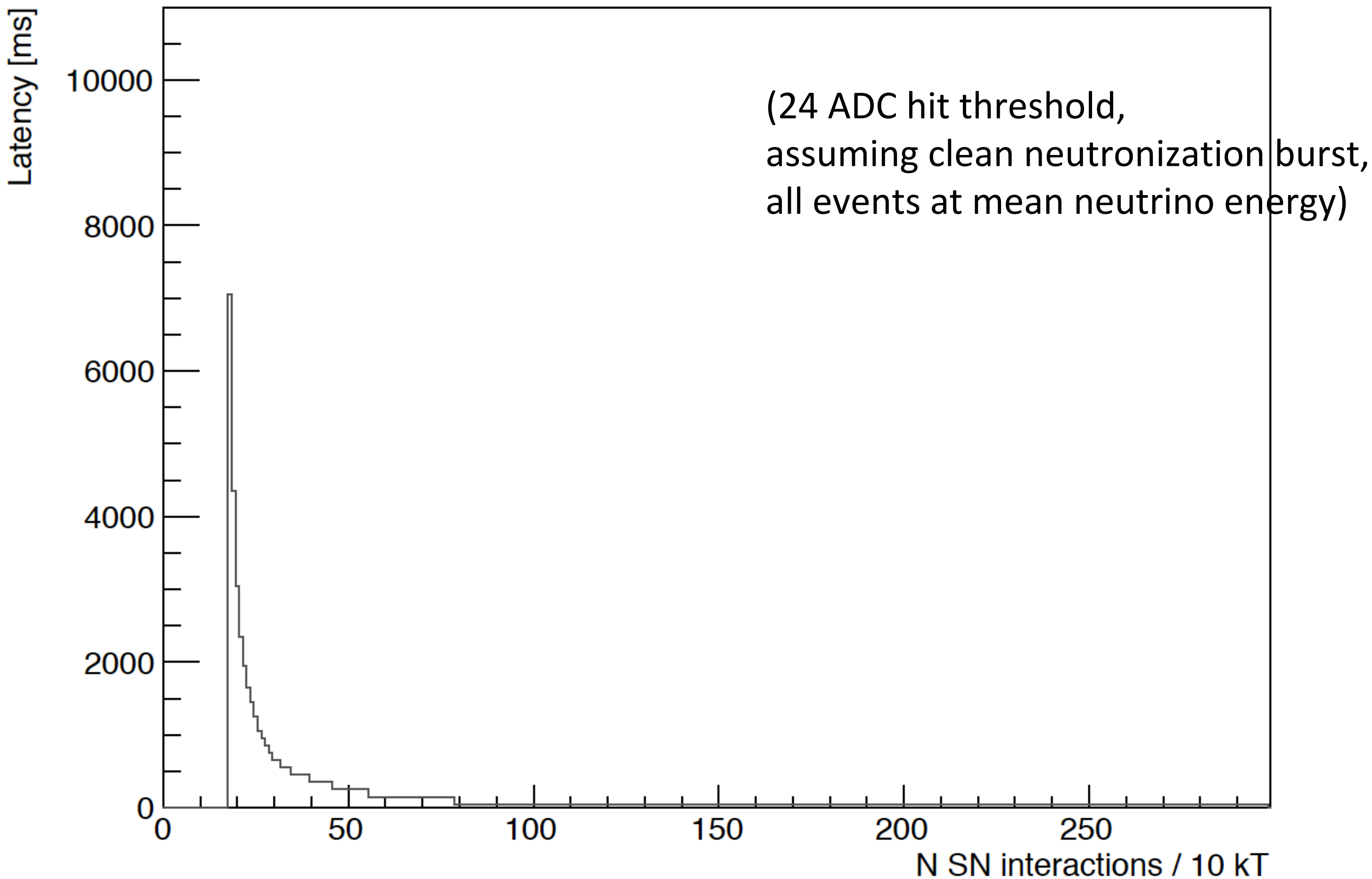
  corresponding to a total of 20 Bytes.

- Expected to be dominated by Ar39 radiological backgrounds (~100Hz per channel for 0 threshold)

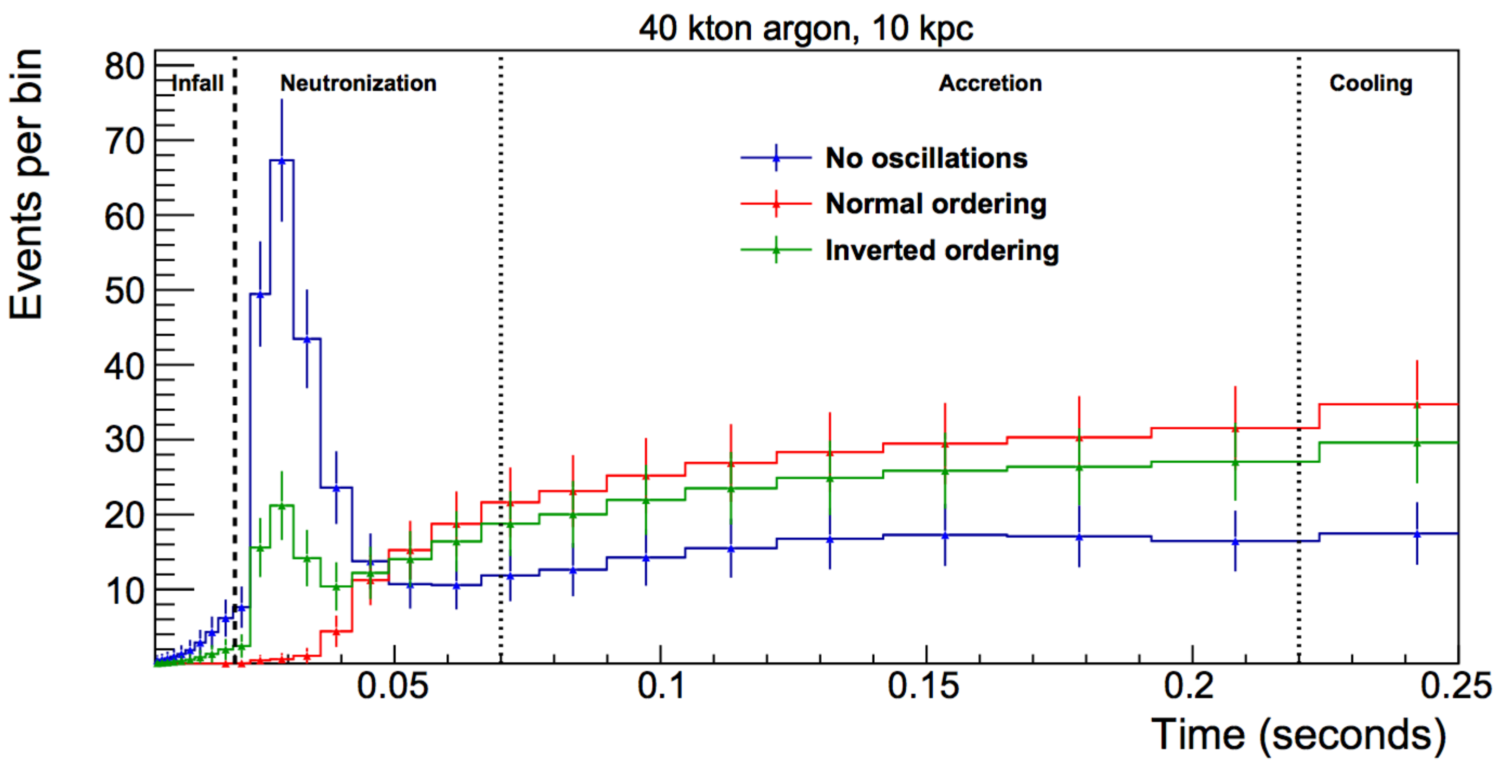
Supernova Burst Trigger Efficiency



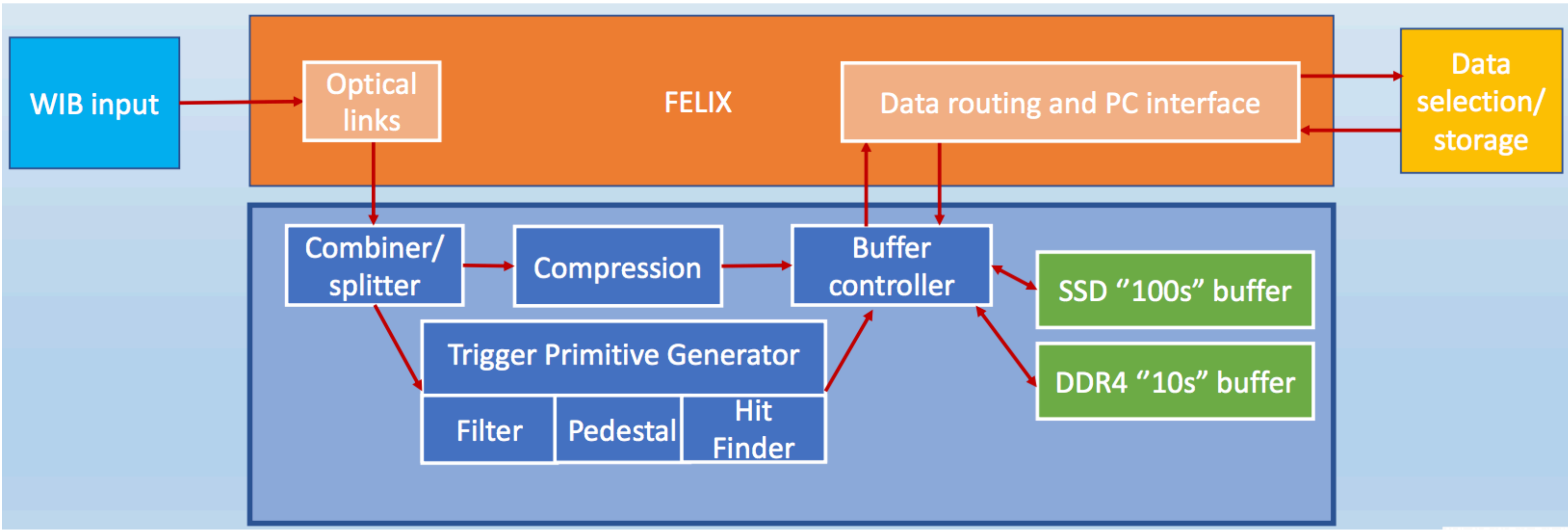
Supernova Burst Trigger Efficiency



SNB Profile (for just one SN model!)



Trigger Primitives in FPGA Co-processor



Trigger Primitive Rates: Simulation

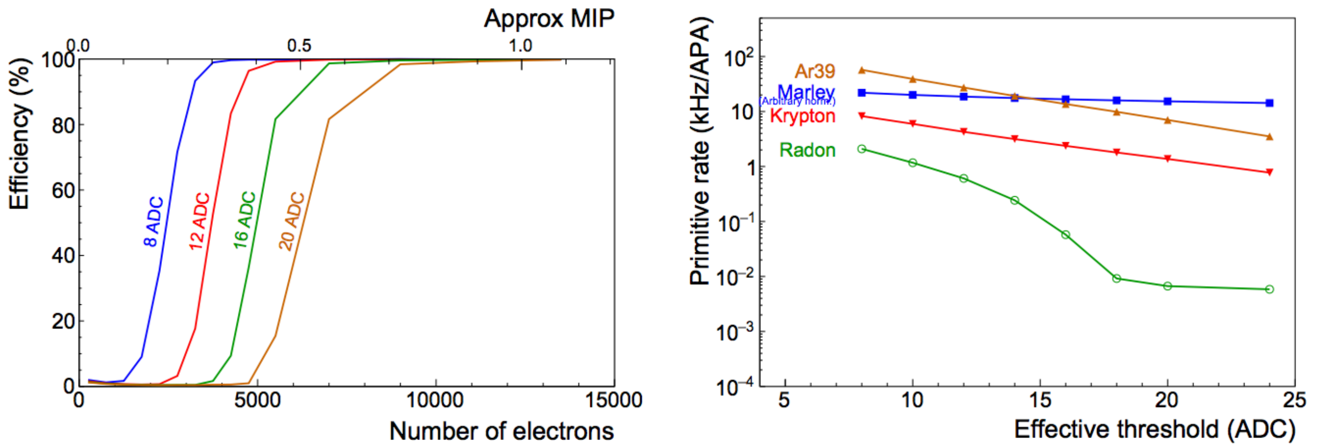


Figure 5: Efficiency and hit rate, default noise. The rate of hits due to noise (ie, not matched with any true energy deposition) is too low to show up on this plot.

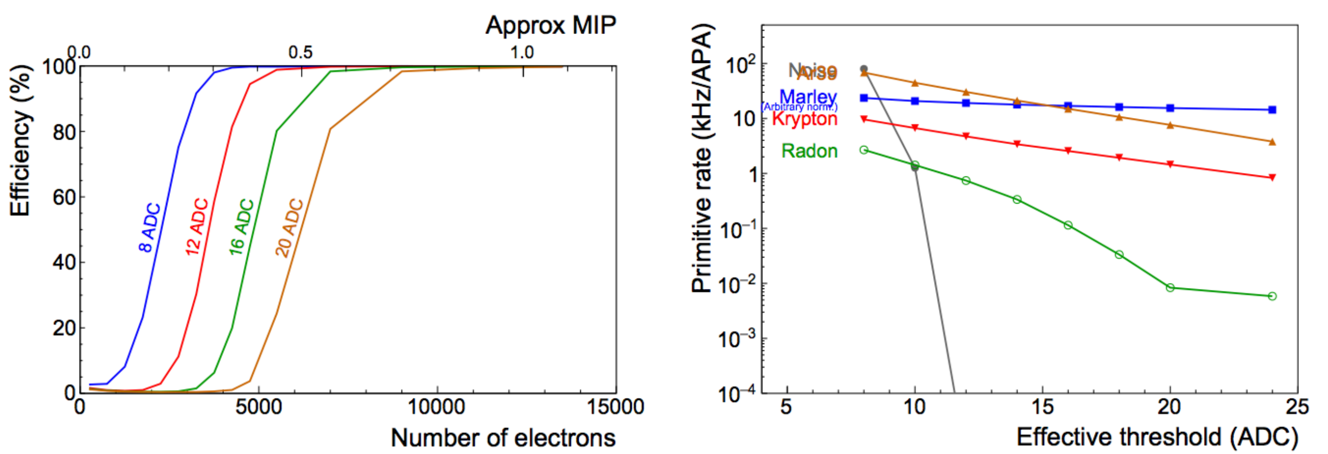


Figure 6: Efficiency and hit rate, with default noise increased by 50%.

Trigger Primitive Generation in CPU

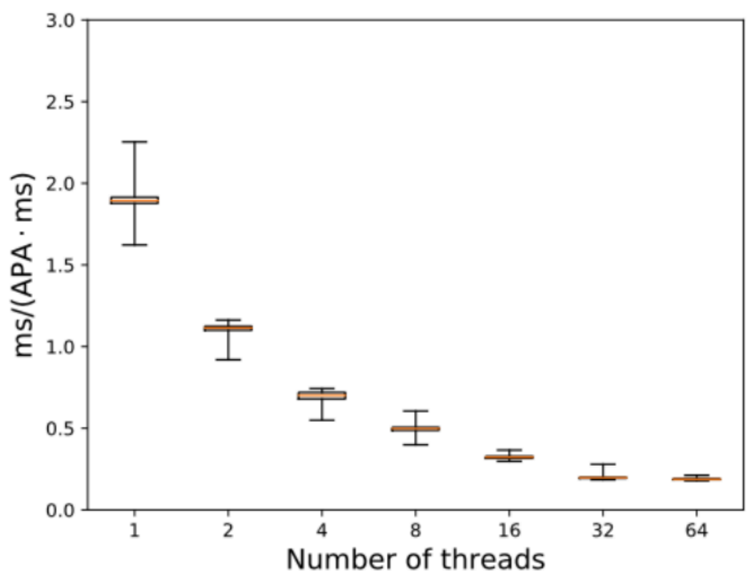


Figure 7: Time for Trigger Primitive generation in milliseconds of processing over APA milliseconds, as a function of the number of threads. The test was done on an existing Xeon Gold 6140 system. The boxes indicate the median and interquartile range of times for various iterations of the test, and the whiskers indicate the minimum and maximum after 1000 repeats of thet test. By 4 threads (cores), the algorithm can keep up with detector data rate.

Trigger Primitive Rates at ProtoDUNE

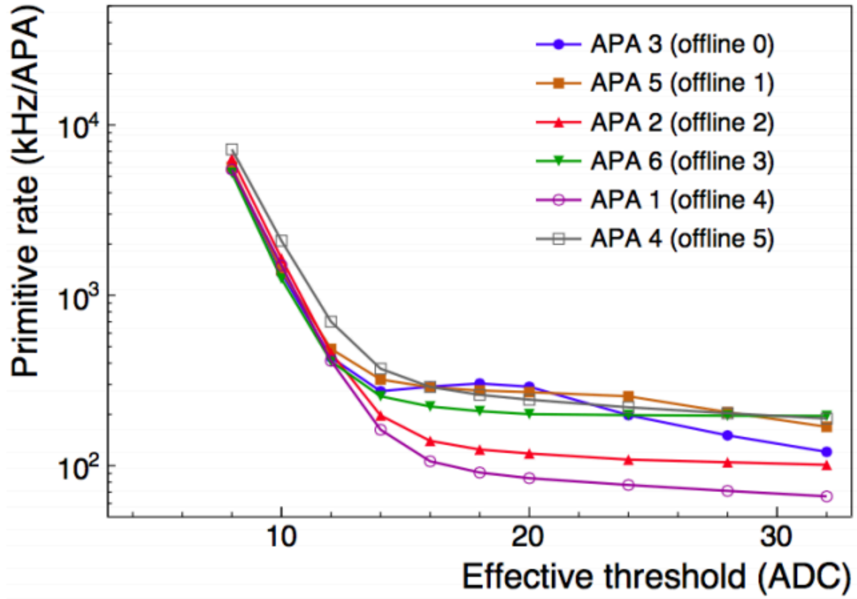
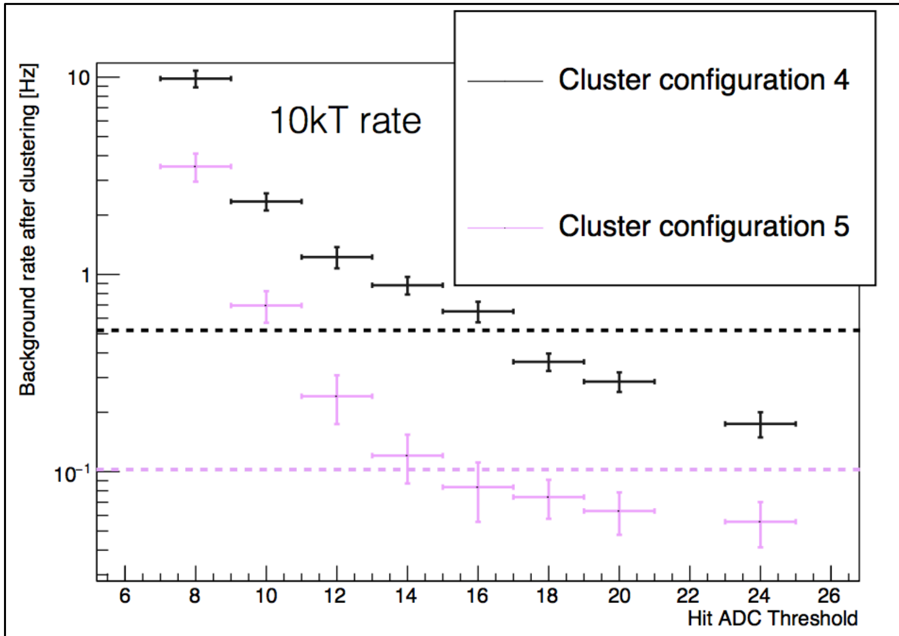
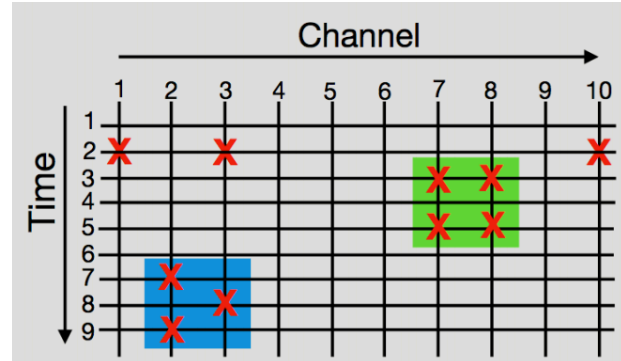
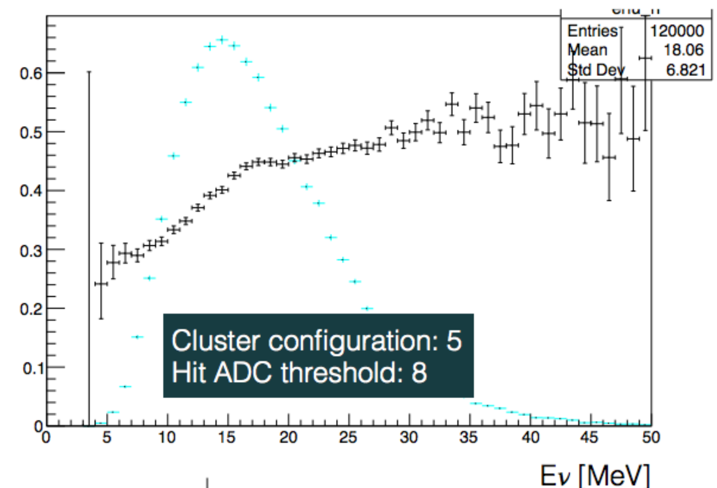


Figure 8: Trigger Primitive rate from early ProtoDUNE data. The "floor" on the rate above about 16 ADC counts is due in part to the high rate of cosmics in ProtoDUNE. In addition, known "noisy" channels have not been removed from these curves.

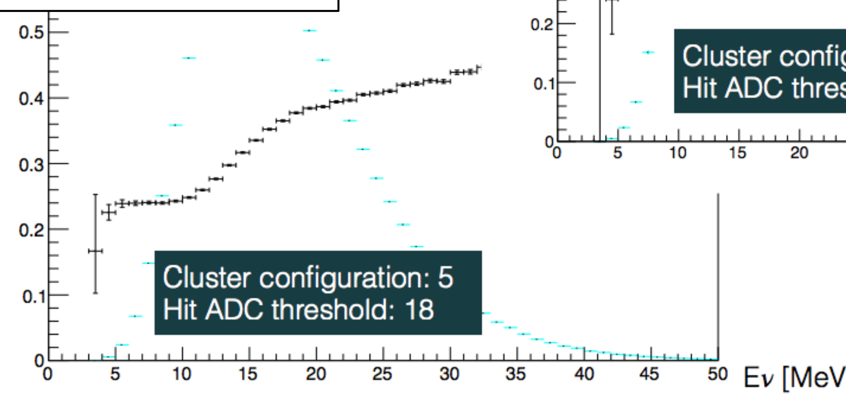
Trigger Candidate Simulations







<0.2ms to cluster 1 APA



High Level Trigger: Simulations

- High Level Trigger classification efficiency on a “per APA-frame” basis:

  (APA-frames containing true event vertex)

“empty” score cut	“empty” eff.	SN eff.	n-nbar osc. eff.	atmo. eff.	p-decay eff.	cosmic eff.
0.1	0.73%	88.18%	99.98%	92.24%	99.29%	92.57%
0.01	0.14%	83.27%	99.98%	91.01%	99.18%	92.46%
0.001	0.033%	77.11%	99.98%	89.76%	99.04%	92.24%
0.0001	0.011%	69.74%	99.97%	88.39%	98.74%	91.71%

[CNN vgg16 network, 20-30ms classification time per APA-frame on single GPU.]