Data Selection

- Scope
- Overview and Status
- Interfaces
- NSF Proposed work and timeline
- Short-term tasks and priorities (inc. ProtoDUNE)
- Resources…

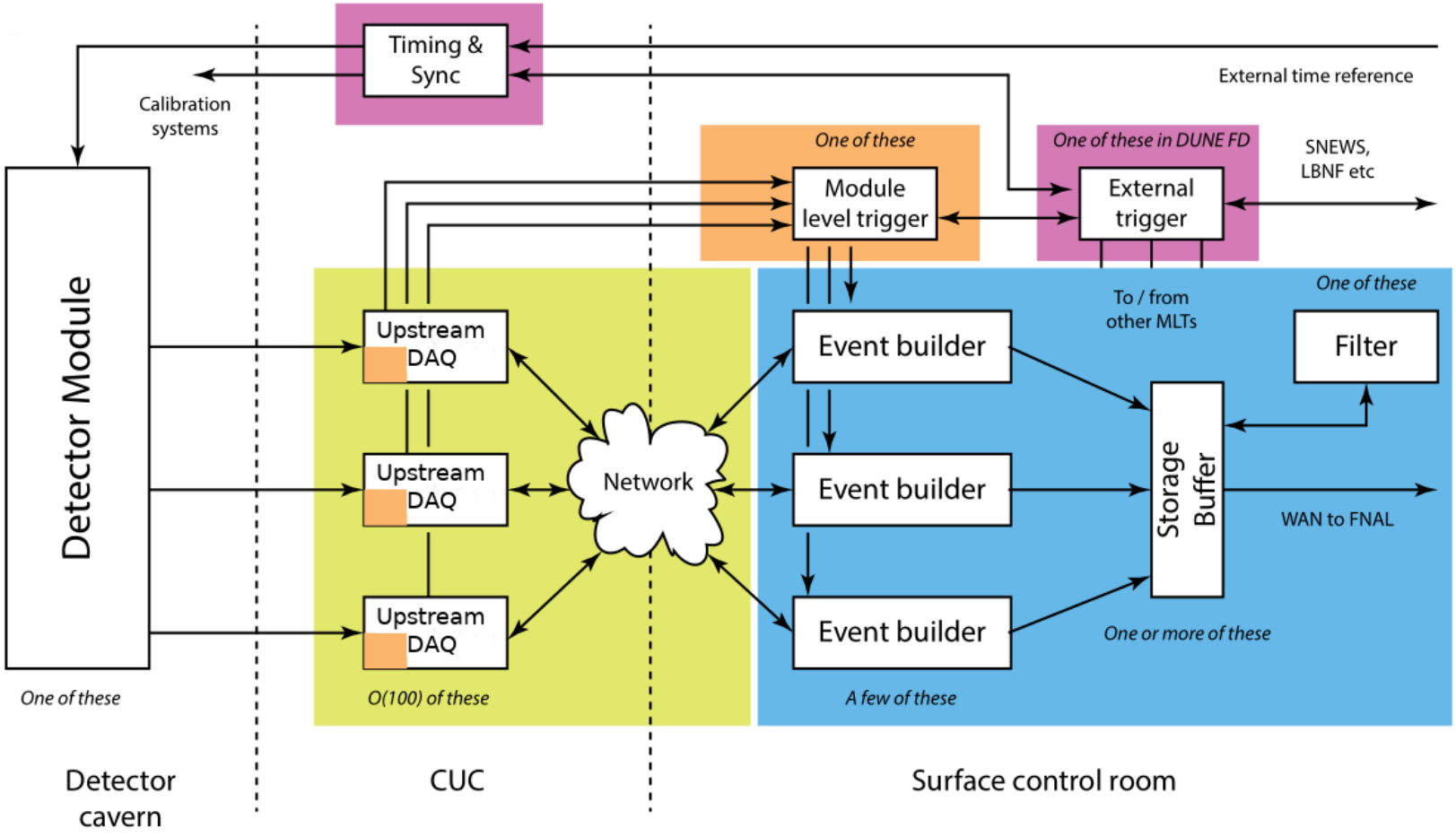
Klein, Karagiori 4 Feb 2019

Scope

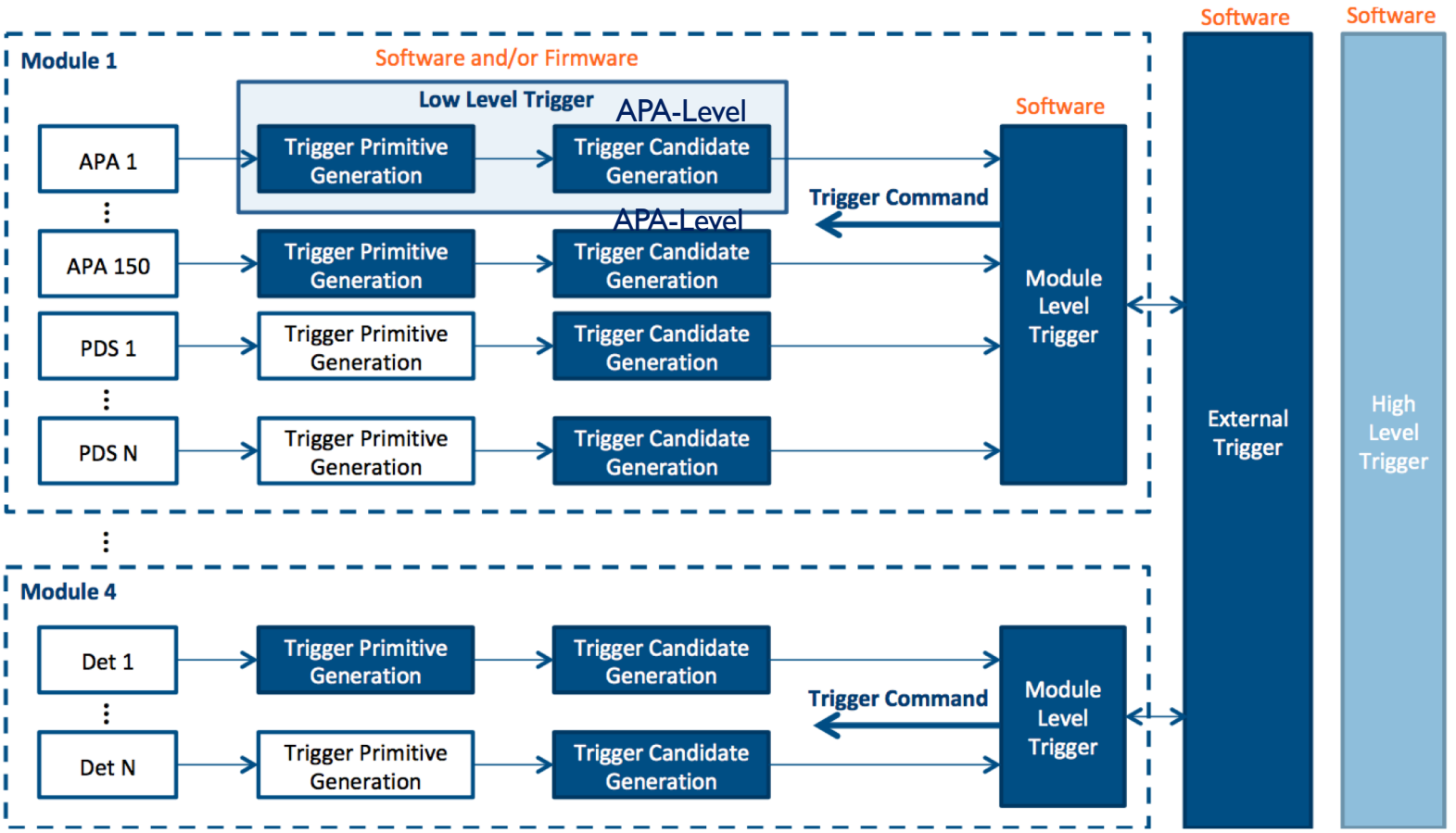
Scope: Data processing and generation of “trigger” instructions (what to be read out, and conditions)

Includes: Low-level processing hardware (CPU where trigger primitives and/or APA-level trigger candidates are formed); module-level processing hardware; ext trigger hardware; high-level processing hardware

Scope

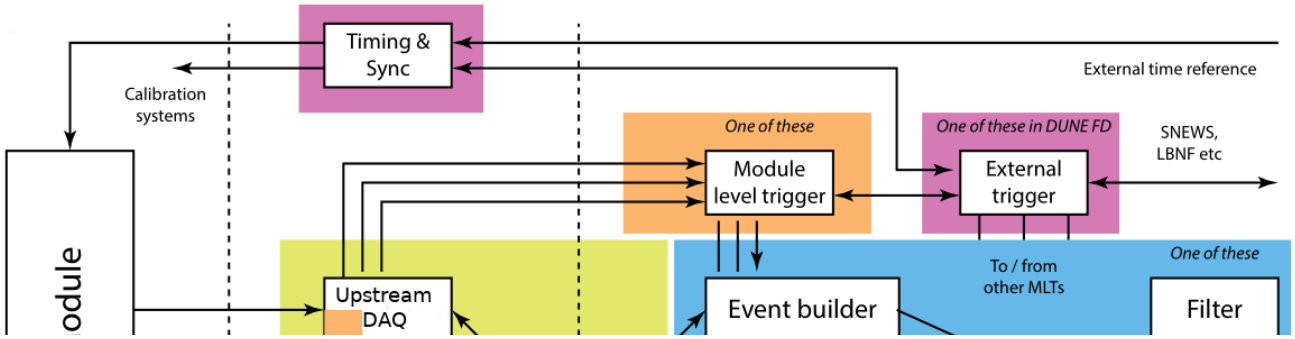

Timing & Sync
External time reference
Calibration systems
One of these
One of these in DUNE FD
SNEWS, LBNF etc
Module level trigger
External trigger
Detector Module
To / from other MLTs
One of these
Filter
Upstream DAQ
Upstream DAQ
Upstream DAQ
Network
Event builder
Event builder
Event builder
Storage Buffer
WAN to FNAL
One or more of these
One of these
O(100) of these
A few of these
Detector cavern
CUC
Surface control room

Scope


Module 1
Software and/or Firmware
Low Level Trigger
APA-Level
APA 1
Trigger Primitive Generation
Trigger Candidate Generation
Software
Trigger Command
APA-Level
APA 150
Trigger Primitive Generation
Trigger Candidate Generation
Module Level Trigger
PDS 1
Trigger Primitive Generation
Trigger Candidate Generation
PDS N
Trigger Primitive Generation
Trigger Candidate Generation
Software
External Trigger
Software
High Level Trigger
Module 4
Det 1
Trigger Primitive Generation
Trigger Candidate Generation
Trigger Command
Module Level Trigger
Det N
Trigger Primitive Generation
Trigger Candidate Generation

Scope

MLT, ELT, Timing, Calibrations…and all that….



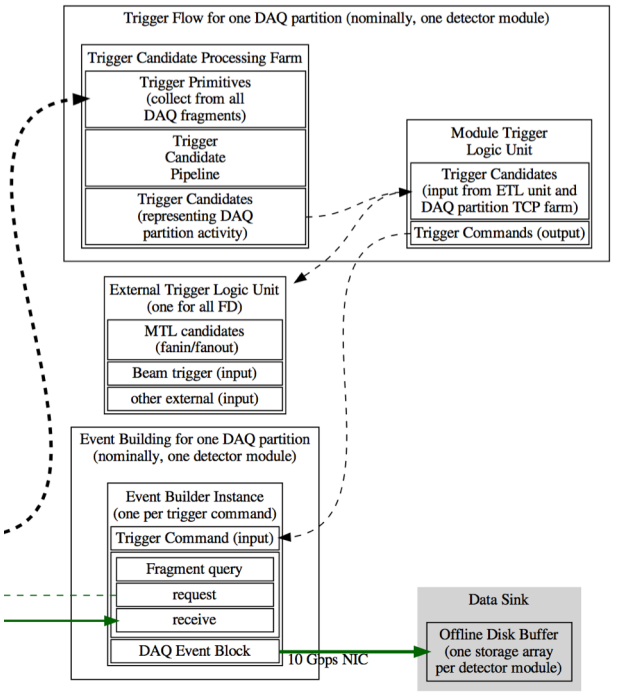
TDR model is that :

MLT functionality includes

- Decision on whether a trigger candidate or set of TCs are a valid trigger
- Issues trigger commands to data egress/orchestrator
- Accepts ELT trigger [candidates and arbitrates these and others]

ELT functionality includes

- Inter-module triggering (e.g., supernovas)
- Beam timeline handling
- Truly "external" triggering (e.g., other experiments)
- Calibration triggers



Scope

Handling Calibrations (FE pulsers, neutron bomb, laser system…)

It is agreed that:

1. When at all possible, DAQ should *drive* calibrations
2. "Firing" calibration sources should avoid pileup with physics triggers
   - In particular beam and supernova triggers
3. Trigger commands should depend on trigger type
   - E.g., readout window for laser cals depends on where and when laser fires

It is *natural* that:

- Calibration source firing is distributed by timing system

Where there is not (yet) agreement is where decision to fire a calibration source happens.
(and this raises a host of other issues)

Scope

- "MLT is GT option"
  - MLT decides on future timestamp to fire calibration source
  - Decision based on SN state and beam timeline
  - Calibration "firing" looks like a trigger command in that it is a timestamp
  - Timing system executes command [could go through ELT]
  - MLT sends trigger commands to buffers requesting data slice appropriate for source
  - Tags this time with "calibration trigger" bit
- "ELT is GT option"
  - ELT decides on future timestamp to fire calibration source
  - Decision based on MLT SN state, ELT SN state, beam timeline
  - Calibration "firing" is an ELT-to-Timing system command
  - Timing system executes command
  - ELT sends calibration timestamp to MLT
  - ELT sends other calibration info to MLT (e.g., where laser is)
  - MLT sends trigger commands to buffers
  - Tags this time with "calibration trigger" bit

Scope

The two fundamental philosophical differences here are

1. The difference between trigger ***decisions*** and trigger ***arbitration.***

These are clearly distinct processes…are they distinct hardware objects or even "blocks"?
MLT will always need to do some arbitration…e.g., PDS vs. TPC
ELT will always need to make some decisions…e.g. inter-module supernova triggers

2. Where calibration configuration information exists

MLT will need some configuration information (see interfaces slides…)
And trigger command is domain of MLT
But MLT already has a lot of interfaces (see inferfaces slides…)

And part of the problem here is that unlike "other experiments" the modules are completely independent except for supernovae (and maybe big cosmic showers).

Scope

"Module Level Trigger" was "Trigger Decisions" + "Trigger Management"

- Module-Level Trigger decisions are based on Trigger Candidates
- Module-Level Trigger management is based on Trigger decisions

Trigger Primitives

TDR Version

- Channel address (32 bits)
- Time of “hit” (64 bits)
- Time-over-threshold (16 bits)
- ADC Sum (32 bits)
- Error flags (16 bits)

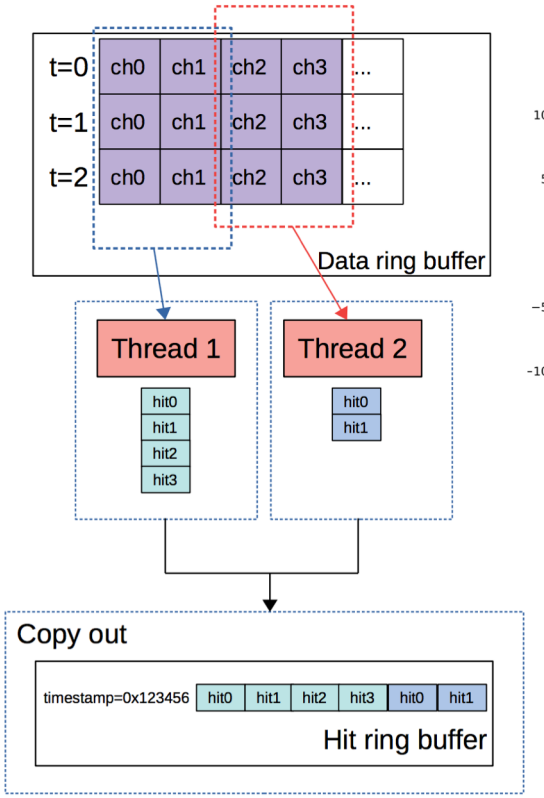
This totals 20 bytes but in principle could be smaller (e.g., time stamp)

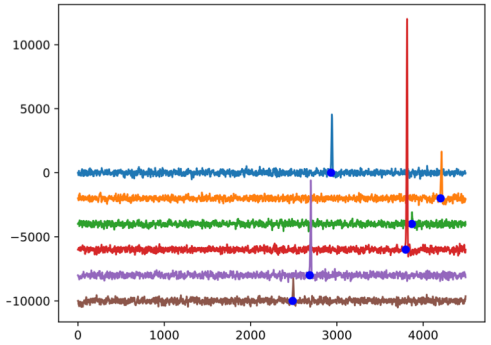
Possible that other information could be used:

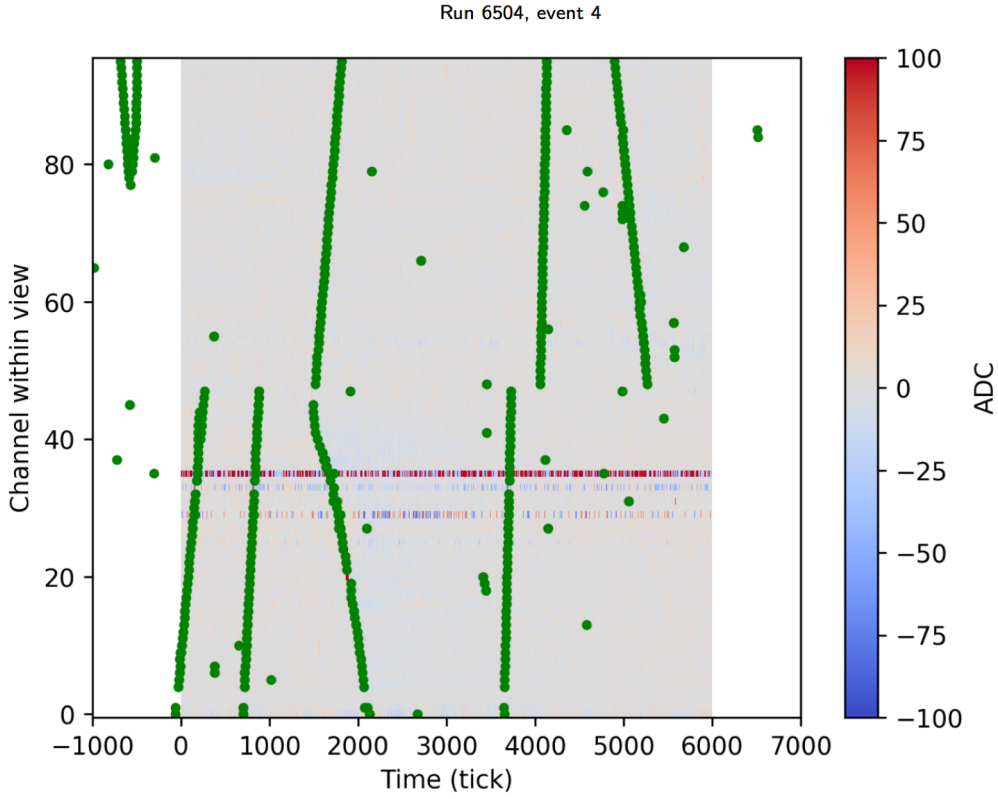
- ADC peak
- Pulse-shape information beyond ToT and sum/peak
- Nearest-neighbor information

Trigger Primitives

Rodrigues returns from The Future to tell us how TP generation is working on FELIX architecture on ProtoDUNE:



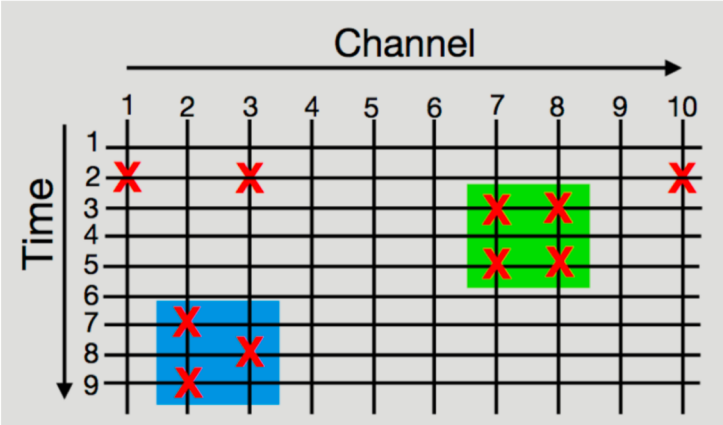




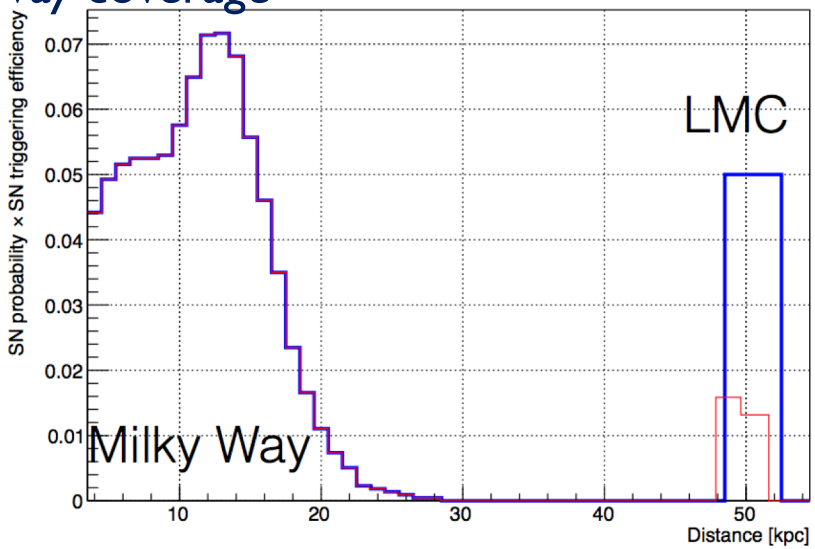
Trigger Candidates (APA-level Clustering)

Low Energy (SP): Supernova, solar interactions

Lasorak approach so far uses only hit clusters (no charge, ToT)



Strictly speaking this satisfies our requirements for Milky Way coverage



But how dumb is it that we don't cover LMC?

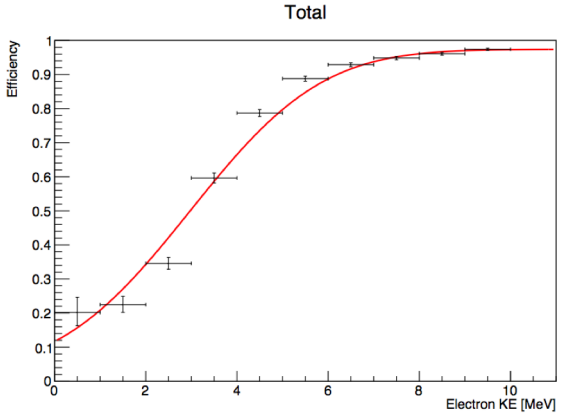
And some people fantasize about solar νs

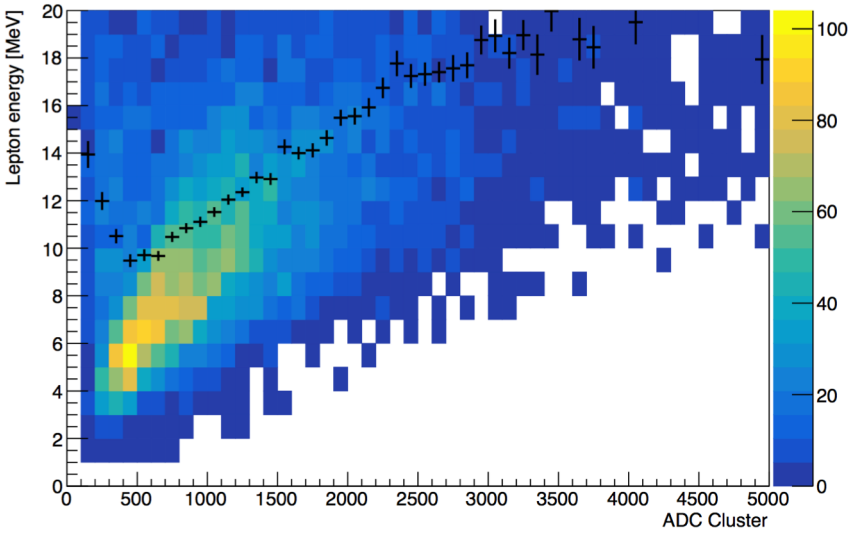
Trigger Candidates (APA-level Clustering)

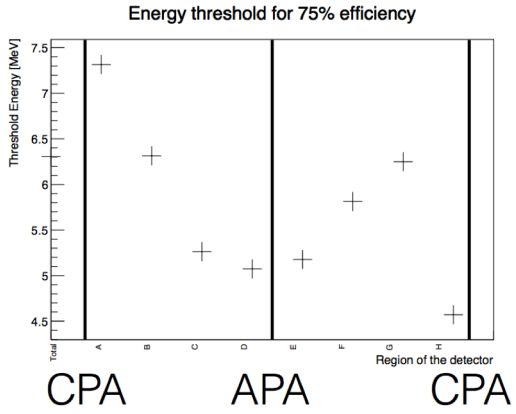
Low Energy (SP): Supernova, solar interactions

Using SADC in trigger candidates gen at low energy may help

Going to SADC cut at low E may require PDS for t_0





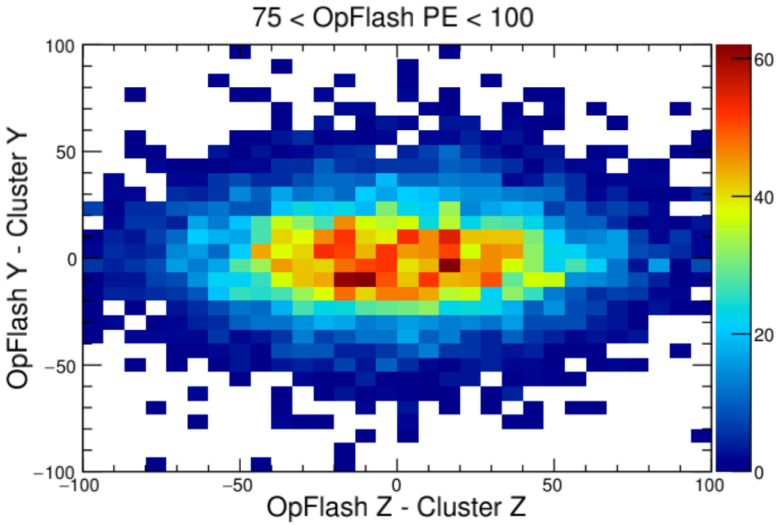


Will depend on achieved electron lifetime at DUNE

Trigger Candidates (APA-level Clustering)

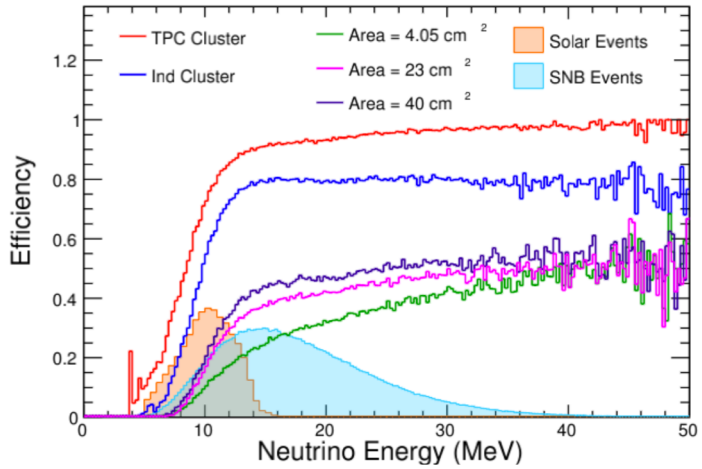
Low Energy (SP): Supernova, solar interactions

More "extreme" version using PDS is Pershey:



Requires induction information for TPC yz "recon"

Requires PDS "recon" for match



	4.05 cm^2	23 cm^2	40 cm^2
Interacting rate	0.49 mHz	0.49 mHz	0.49 mHz
3+ hits in collection plane	0.37 mHz	0.37 mHz	0.37 mHz
Rate with y-coordinate reco'd	0.29 mHz	0.29 mHz	0.29 mHz
Rate with OpFlash reco'd	0.064 mHz	0.105 mHz	0.126 mHz
Rate with NHit-Charge cut	0.062 mHz	0.101 mHz	0.121 mHz

This looks very encouraging! But…

Trigger Candidates (APA-level Clustering)

PDS Polemics:

xArapuca design is creative, cool, beautiful!

But what does PDS actually add at the trigger level?

1. t0 if light yield is good enough at low energies---potentially real added value
2. Improved energy resolution if PDS resolution ~ TPC resolution

Trigger Candidates (APA-level Clustering)

PDS Polemics:

xArapuca design is creative, cool, beautiful!

But what does PDS actually add at the trigger level?

1. t0 if light yield is good enough at low energies---potentially real added value
2. Improved energy resolution if PDS resolution ~ TPC resolution
   - If this is true, please call Eric James and tell him we are building the wrong detector

Trigger Candidates (APA-level Clustering)

PDS Polemics:

xArapuca design is creative, cool, beautiful!

But what does PDS actually add at the trigger level?

1. t0 if light yield is good enough at low energies---potentially real added value
2. Improved energy resolution if PDS resolution ~ TPC resolution
   - If this is true, please call Eric James and tell him we are building the wrong detector
   - Please let me be on that call

Trigger Candidates (APA-level Clustering)

PDS Polemics:

xArapuca design is creative, cool, beautiful!

But what does PDS actually add at the trigger level?

1. t0 if light yield is good enough at low energies---potentially real added value
2. Improved energy resolution if PDS resolution ~ TPC resolution
   - If this is true, please call Eric James and tell him we are building the wrong detector
   - Please let me be on that call
3. Redundant (coincident) information that event is not noise
4. Possibly PSD information

Are these worth creating a system dependency at the trigger level?

1 could help if electron lifetime is really bad (but then we have other issues and pDUNE OK)

2 better be wrong

3 could help if noise is really bad (but then we have other issues again…)

4 has not been explored by anyone and we don't expect αs to be an issue

Trigger Candidates (APA-level Clustering)

PDS Polemics:

If we are going to use PDS it should be in HLT (or maybe Module Level)

AND/OR

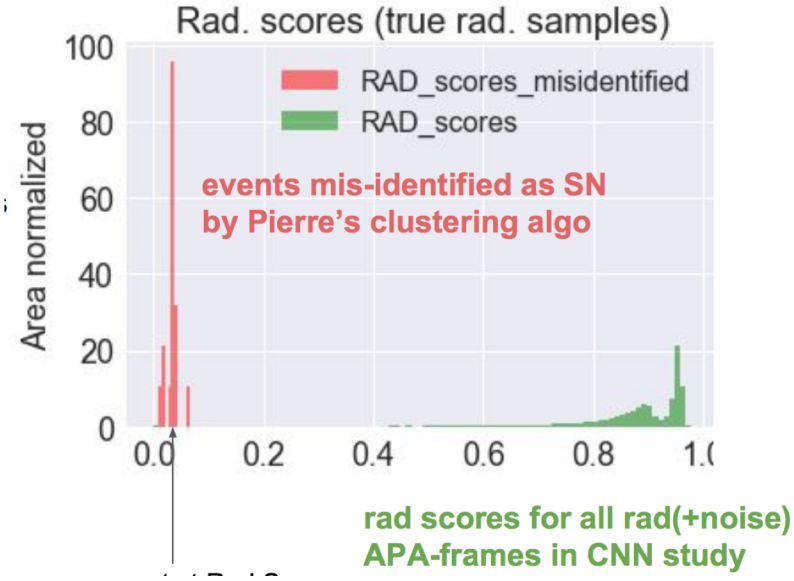
Have entirely distinct PDS-only trigger

Trigger Candidates (APA-level Clustering)

Low Energy (SP): Supernova, solar interactions

Machine Learning (CNN) approaches (Jwa, Karagiorgi)

RAD score cut	RAD frame efficiency	SN frame efficiency	n-nbar frame efficiency	atmo. nu frame efficiency	p-decay frame efficiency	cosmic frame efficiency
<0.1	0.73% (99.44% rejection)	89.18%	99.98%	92.24%	99.29%	92.57%
<0.01	0.14% (99.82% rejection)	83.27%	99.98%	91.01%	99.18%	92.46%
<0.001	0.033% (99.969% rejection)	77.11%	99.98%	89.76%	99.04%	92.24%
<0.0001	0.011% (99.989% rejection)	69.74%	99.97%	88.39%	98.74%	91.71%

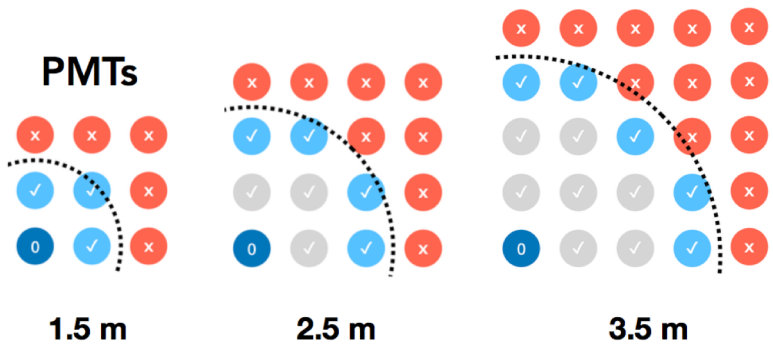


Does a great job of tagging mis-ID events for current Low E clustering approach

Trigger Candidates (APA-level Clustering)

Low Energy (DP): Supernova, solar interactions

Trigger based so far exclusively on PDS (Gallego Ros)



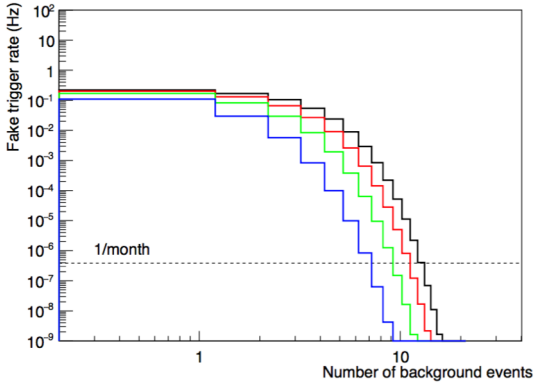
Parameters

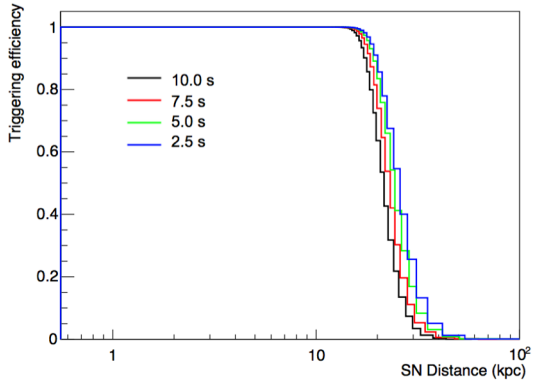
< 250 ns of cluster time window
> 48 hits per cluster
< 2.5 m between PMTs
5 s of burst time window

Results

DE = 15 % & BGR = 0.24 Hz
95% of the events of the burst considered
Min bkg cluster multiplicity (for FT<1/month): 10

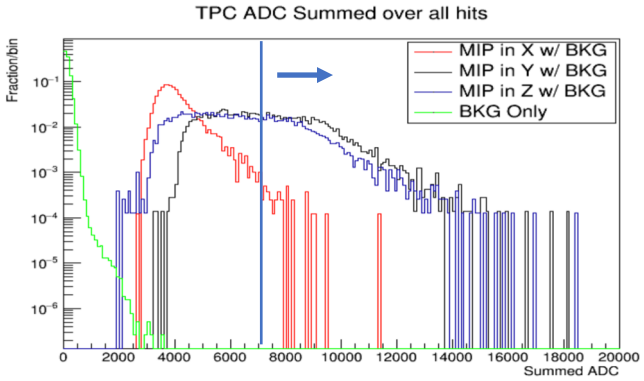
TE for a SNB at a 10 kpc distance: 100%
TE for a SNB at a 15 kpc distance: 99%
TE for a SNB at a 20 kpc distance: 83%

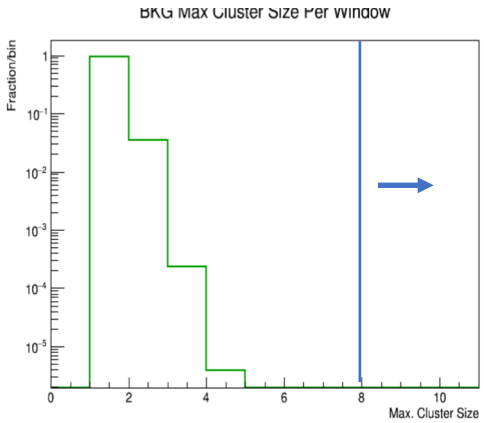


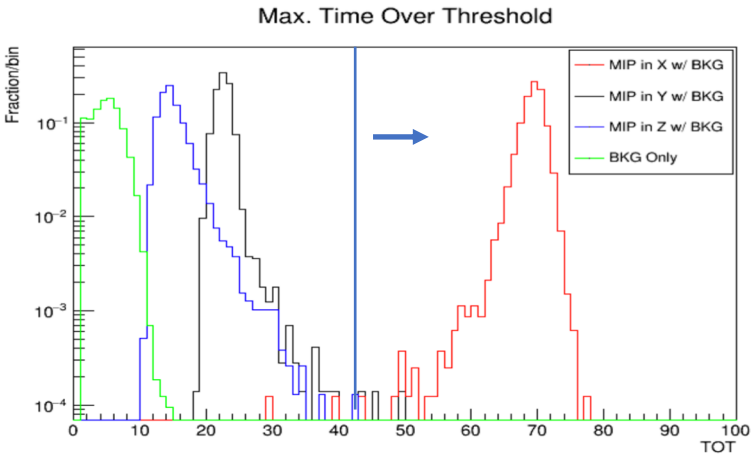


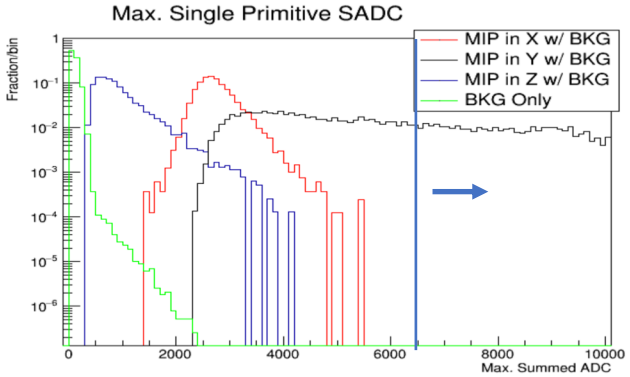
Trigger Candidates (APA-level Clustering)

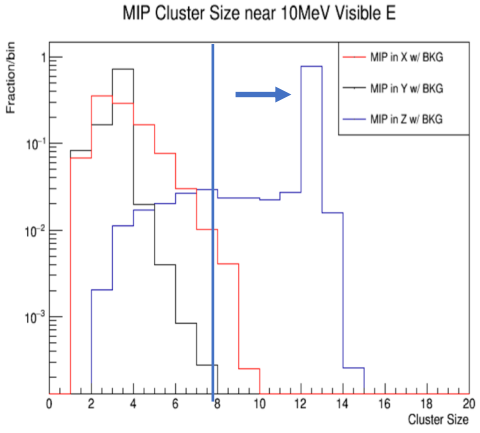
High Energy: Beam, atmospherics (Last, Rivera)







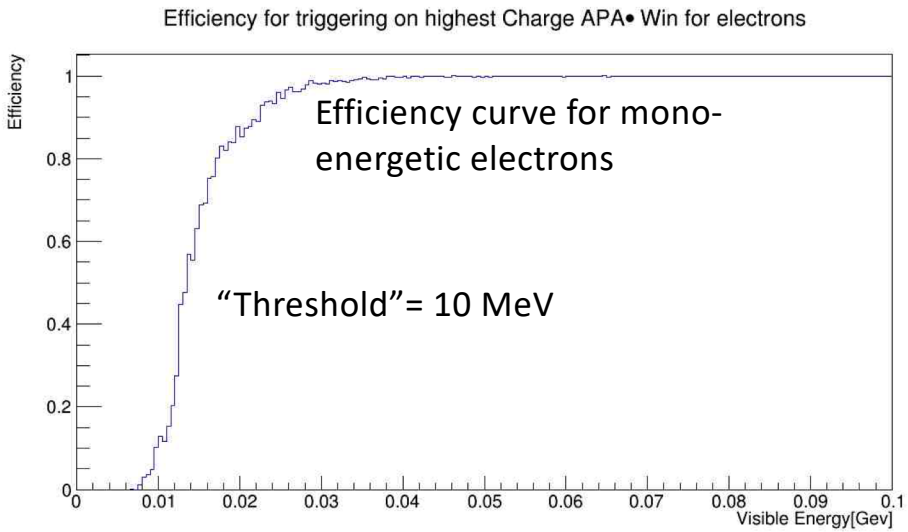




- Adjacency > 8
- TOTADC > 7000 cts
- MAX SADC > 6500 cts
- ToT > 45

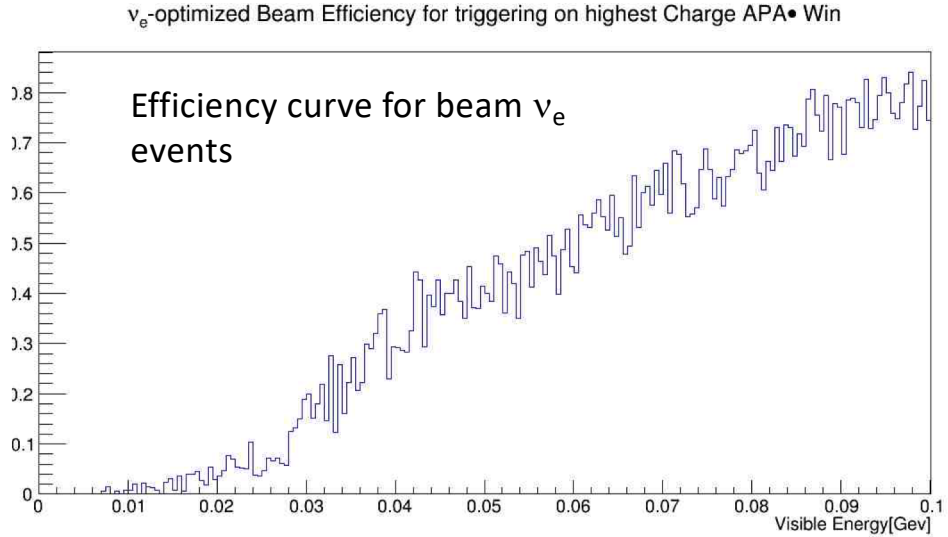
Trigger Candidates (APA-level Clustering)

- Beam interaction triggering (Last, Rivera)



D. Last, D. Rivera

NC events lead to loss of efficiency



D. Last, D. Rivera

Integral efficiency for beam:

- nu_mu with NC: 97.950%
- nu_mu without NC: 99.992%
- nu_e with NC: 97.958%
- nu_e without NC: 99.991%

[Logical] Interfaces: Trigger Primitives

Input Side

- FELIX waveforms
- TP configuration (e.g., filter function(s), thresholds,…) from Run Control/DB
- Detector configuration (noisy/ignored channels, gains,…) from Run Control/DB

Output Side

- Trigger Candidate processor---network, backplane, RAM...?
- TP streamer (to disk)
- Monitor processor (poss. Including ^{39}Ar cal)

[Logical] Interfaces: Trigger Candidates (TPC)

"APA Level"

Input Side

- TP processors---network, backplane, RAM…?
- TC configuration (cluster sizes, cuts,…) from Run Control/DB
- Detector configuration (dead channels,…) from Run Control/DB
- Clock/timing system ? (for timestamp checks on TCs…?)
- Backpressure throttler? (only for testing we hope!)

Output Side

- Module Level trigger processor (network prob.)
- TC streamer? (to disk)
- Monitor processor

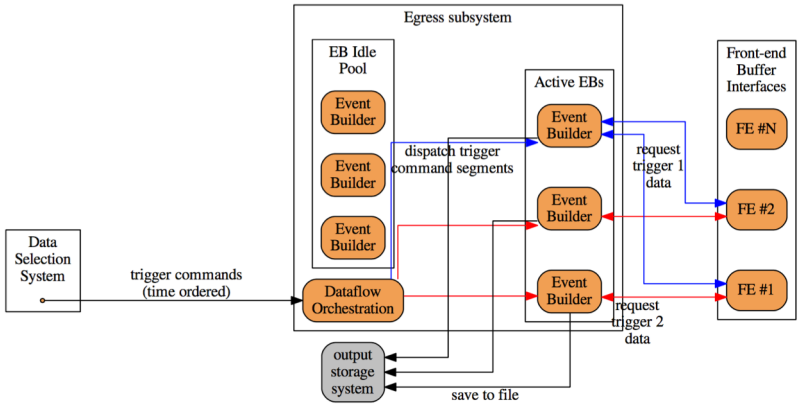
[Logical] Interfaces: Module Level Trigger

Input Side

- [75?] TC processors---network prob
- MLT config (e.g., SN burst criteria, trigger mask…) from Run Control/DB
- Partition "server"
- Trigger-generating calibrations (e.g., radioactive source)
- Calibration configuration (e.g., where is laser now?)
- External Trigger processor (network prob.)
- PDS System (PDS trigger primitives)
- Clock/timing system ? (for timestamp checks on MLTs?)
- Detector configuration (offline APAs, noisy APAs,…) from Run Control/DB
- Backpressure throttler? (only for testing we hope!)

Output Side

- Data "Orchestrator"
- Out-of-band Data Orchestrator (for SNs)
- External Trigger
- Calibration sources
- Monitoring



[Logical] Interfaces: External Level Trigger

Input Side

- [4+/-] MLT processors
- Truly external sources (SNEWs, THEIA, Super-LZ,…)
- Partition "server"
- Timing system
- Beam timeline
- Detector configuration (offline APAs, noisy APAs,…) from Run Control/DB
- Backpressure throttler? (only for testing we hope!)

Output Side

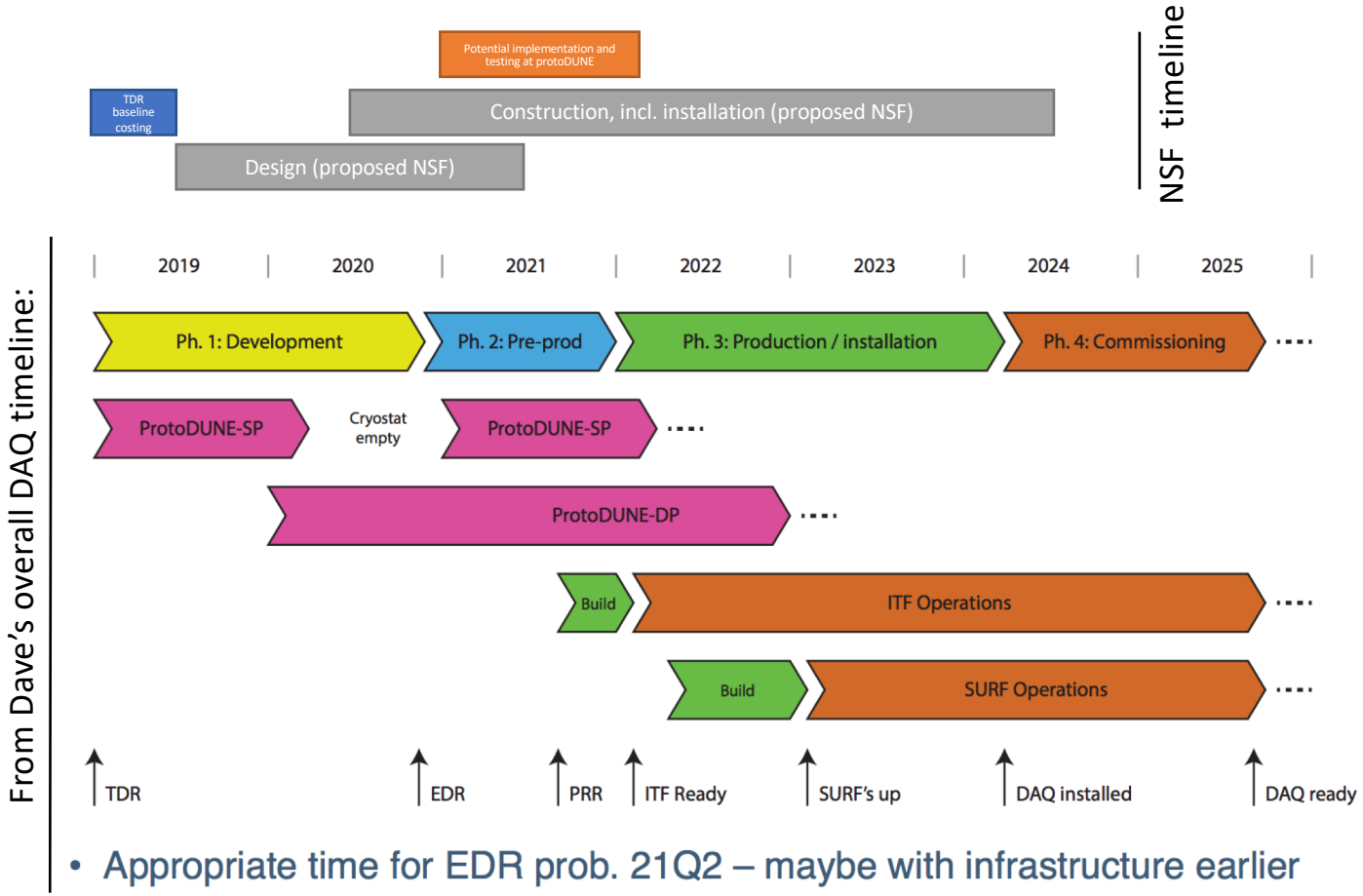
- [4+/-] MLTs
- Truly external sources (SNEWs, THEIA, Super-LZ…)
- Timing system
- Monitoring

Planning Overview

- Requirements are well understood (from DAQ). Must be vetted with physics groups.
  - Uptime requirement needs some thinking about failsafes and redundancy
- What is needed is the development of a full vertical slice (including ProtoDUNE), employing online software framework and accepted communication protocols. For that:
  - Hardware for all trigger levels must be specified.
  - Online software framework must be developed.
  - Priority scheme for "Trigger Manager"
  - Interface between DS and CCM must be specified.
  - Accelerator signals (number, types, latencies) and calibration signals must be specified.

Also, development will continue to explore non-CPU implementations (GPU, FPGA) of algorithms.

Timeline for DS Development and Demonstrations



- Appropriate time for EDR prob. 21Q2 – maybe with infrastructure earlier

Scope of (NSF proposed) Design Work

- Construction of demonstrators at multiple US institutions for the low-level trigger (APA-level) (year 1).
- Merging of teststands for module-level trigger demonstration (year 2): "vertical slice test".
- "Horizontal slice test" for module-level trigger and high level trigger demonstration at HPC center (year 2).
- Along the way, development of:
  - online software for data selection, including low-, module-, ext., and high-level trigger
  - data-processing algorithms (trigger primitives, low-and high-level trigger)
  - data-selection logic for module-level and ext trigger
  - data-processing algorithms at post-event-builder high-level trigger

For low-level trigger, different implementations (CPU, GPU, FPGA, will be tested and compared for performance, power, cost, etc)

Proposed Construction Work:

- Construction of DS system including CPU servers for low-level data processing; CPU servers for module-level trigger; node for EXT trigger; C/GPU servers for high-level trigger
- Production and testing at ProtoDUNE
- Production and installation at far site

US Collaborating institutions:
Columbia, Penn, Iowa State, UMD, Duke, UCDavis, SDSMT, Colorado State

ProtoDUNE deployments

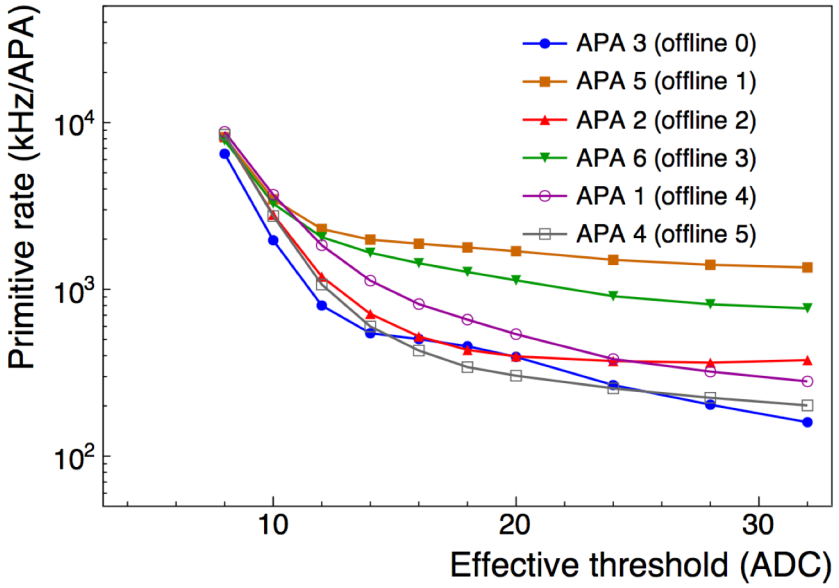
- Design work is proposed to be carried out in the US
- Production would include production testing at ProtoDUNE, ideally in mid 2020 or later.

- It may be possible to test some levels of online data selection at protoDUNE sooner (e.g. summer 2019) if sufficient (critical mass) people power available. Interest from Penn, Columbia, possibly others.

Short-term Tasks and Priorities

Trigger Primitives

- Expand TP algorithm on ProtoDUNE SP FELIX to full APA
- Compare "cleaned" TPs on ProtoDUNE to simulation
- Plot cleaned TP rate vs. time to demonstrate stability
- Send requirement to CE for WIB channel remapping
- Full-stream test (has this been done?) and throughput
- Make a decision on what to do about ^{39}Ar already
- Make a decision on firmware vs. software TP gen



Short-term Tasks and Priorities

Trigger Candidates

- Define software framework (from TP to MLT to ELT…!)
- Check background trigger rate on MC for high-energy algorithms with reasonable statistics
- Test HE TCs on ProtoDUNE data offline (SP+DP)
- Implement HE TC algorithms on local ProtoDUNE servers to test throughput
- Compare HE TCs to simulation
- Implement low E clustering on ProtoDUNE servers---compare to simulation
- Develop low E/high E (or better) tags for TCs
- Investigate inclusion of more information for Low E TPs
- Develop PDS TC algorithm (decide if this lives in MLT or not)

Short-term Tasks and Priorities

Module Level trigger

- Define software framework (from TP to MLT to ELT…!)
- Tests of simple self-triggering at ProtoDUNE (simple clustering, trigger management)
- Continue discussions with calibration WG
- Investigate using more information for SN burst definition to cover LMC
  - Can we develop a minimum bias criterion?
- Develop PDS MLT algorithm (decide if this lives in MLT or not)
- Compare ProtoDUNE PDS simulation to data

Short-term Tasks and Priorities

Simulation

- FIXME Radiological background model FIXME
- Generate much more radio statistics
- Data Selection simulation (in particular clustering algorithms but also TPs)
- MIP-only differential trigger efficiency curves

Summary

- We have "working" TP and TC algorithms
- Software framework for data selection needs to be defined
  - From TP to TC to MLT to DO
- Need further development and testing at ProtoDUNE
- Local test stands for data selection development will get built when NSF proposal(s) funded
- Short-term priorities include
  - TP streaming at ProtoDUNE
  - (When will a decision be made on FPGA vs CPU?)
  - Simple clustering/trigger management tests at ProtoDUNE