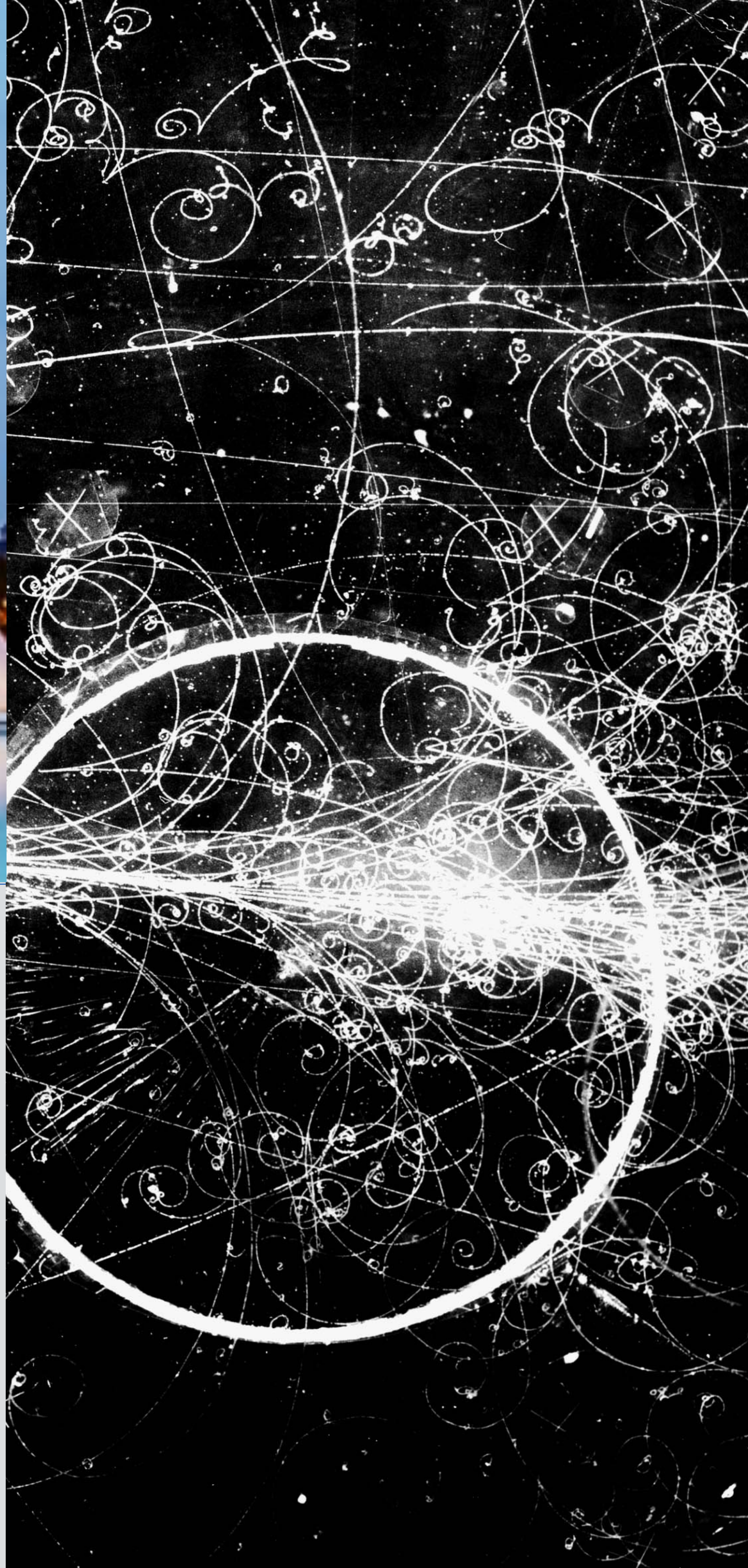


DUNE, protoDUNE
Continuous DAQ systems

ISOTDAQ 2020
University of Valencia, Valencia

January 14, 2020

Alessandro Thea
Rutherford Appleton Laboratory - PPD

UKRI Science and Technology Facilities Council

ISOTDAQ
International School of Trigger and Data Acquisition

Outline

DUNE parameters and challenges

DUNE - Physics

Origin of Matter

Could neutrinos be the reason that the universe is made of matter rather than antimatter? By exploring the phenomenon of neutrino oscillations, DUNE seeks to revolutionize our understanding of neutrinos and their role in the universe.

Unification of Forces

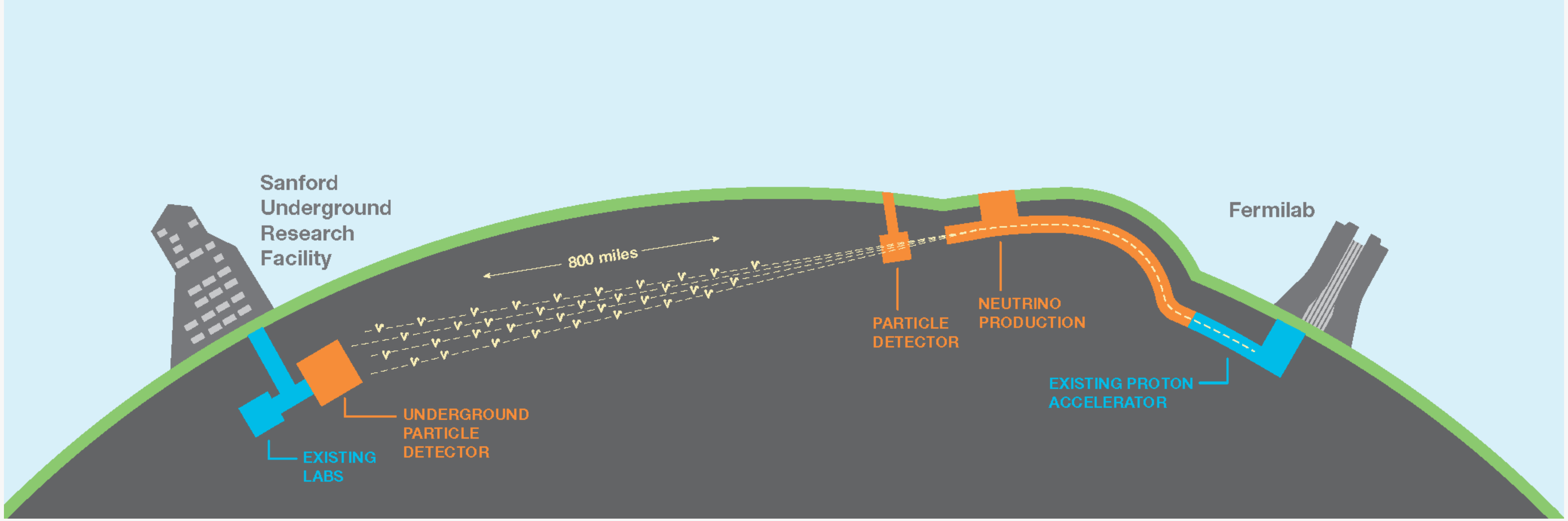
With the world's largest cryogenic particle detector located deep underground, DUNE can search for signs of proton decay. This could reveal a relation between the stability of matter and the Grand Unification of forces, moving us closer to realizing Einstein's dream.

Black Hole Formation

DUNE's observation of thousands of neutrinos from a core-collapse supernova in the Milky Way would allow us to peer inside a newly-formed neutron star and potentially witness the birth of a black hole.

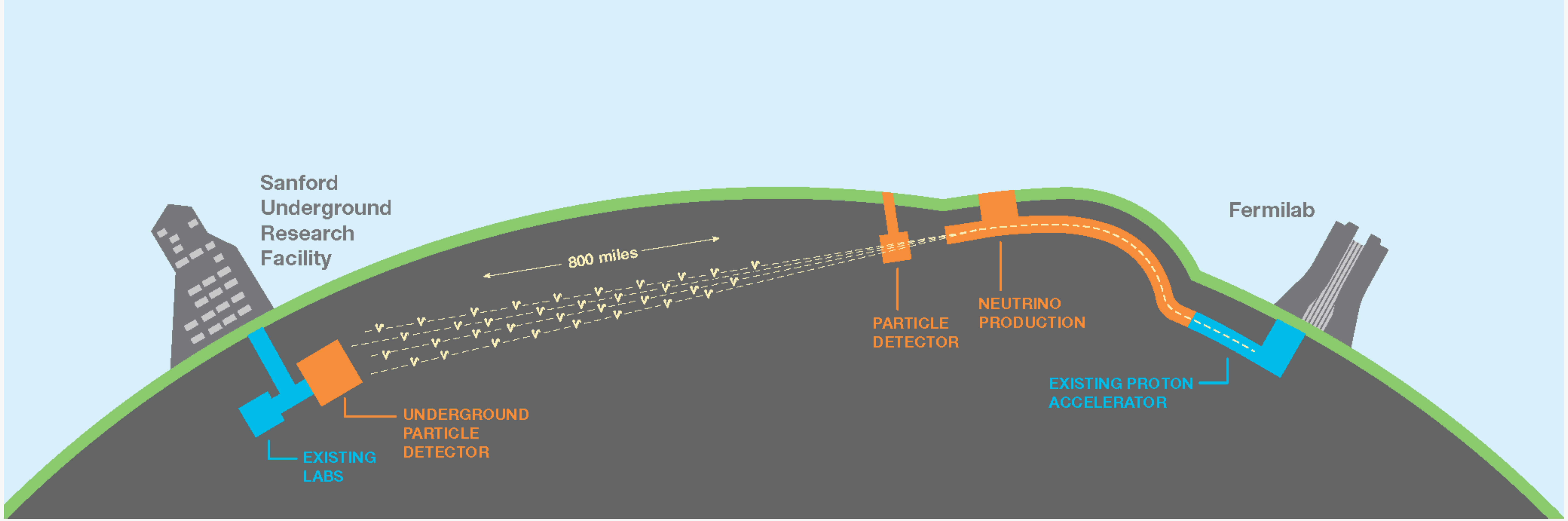
e.org

DUNE - Facility



- Accelerator generating intense neutrino beam
- Near detector measuring neutrinos close to source
- Far detector 1300 km away from source and 1.48 km underground

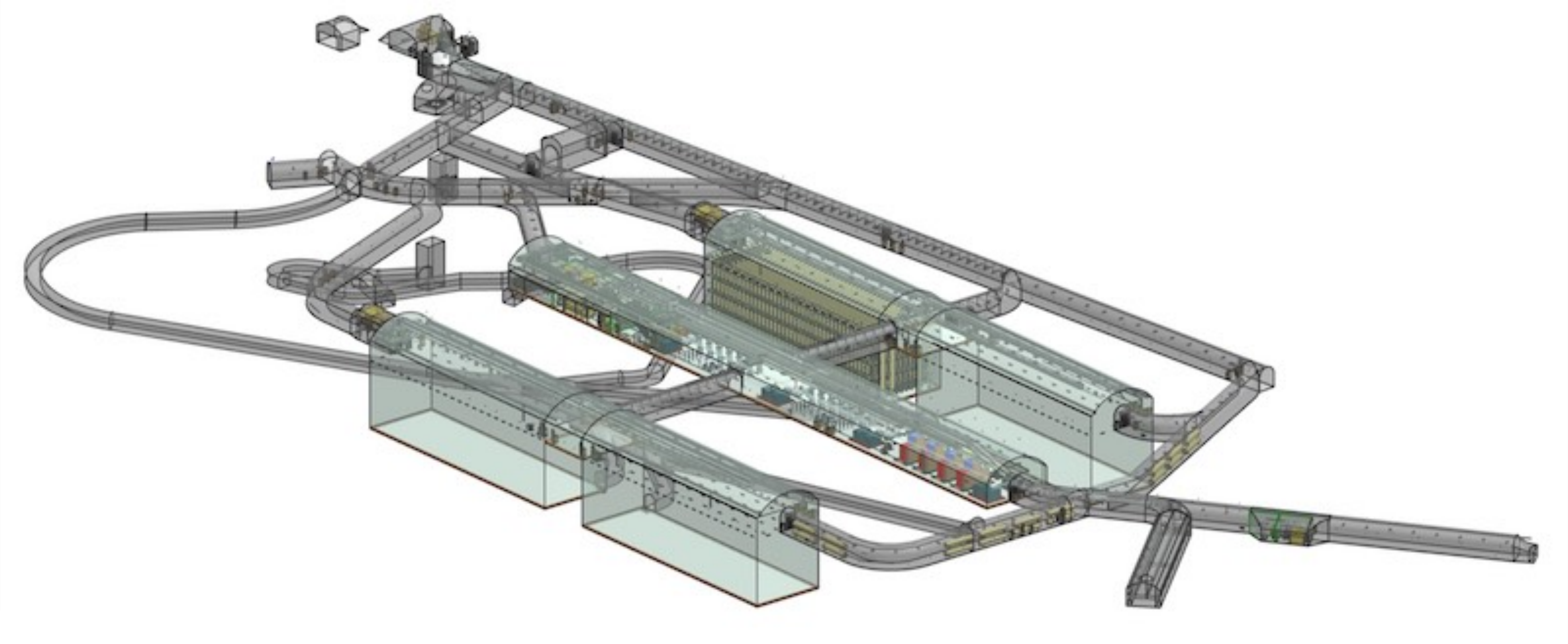
DUNE - Facility

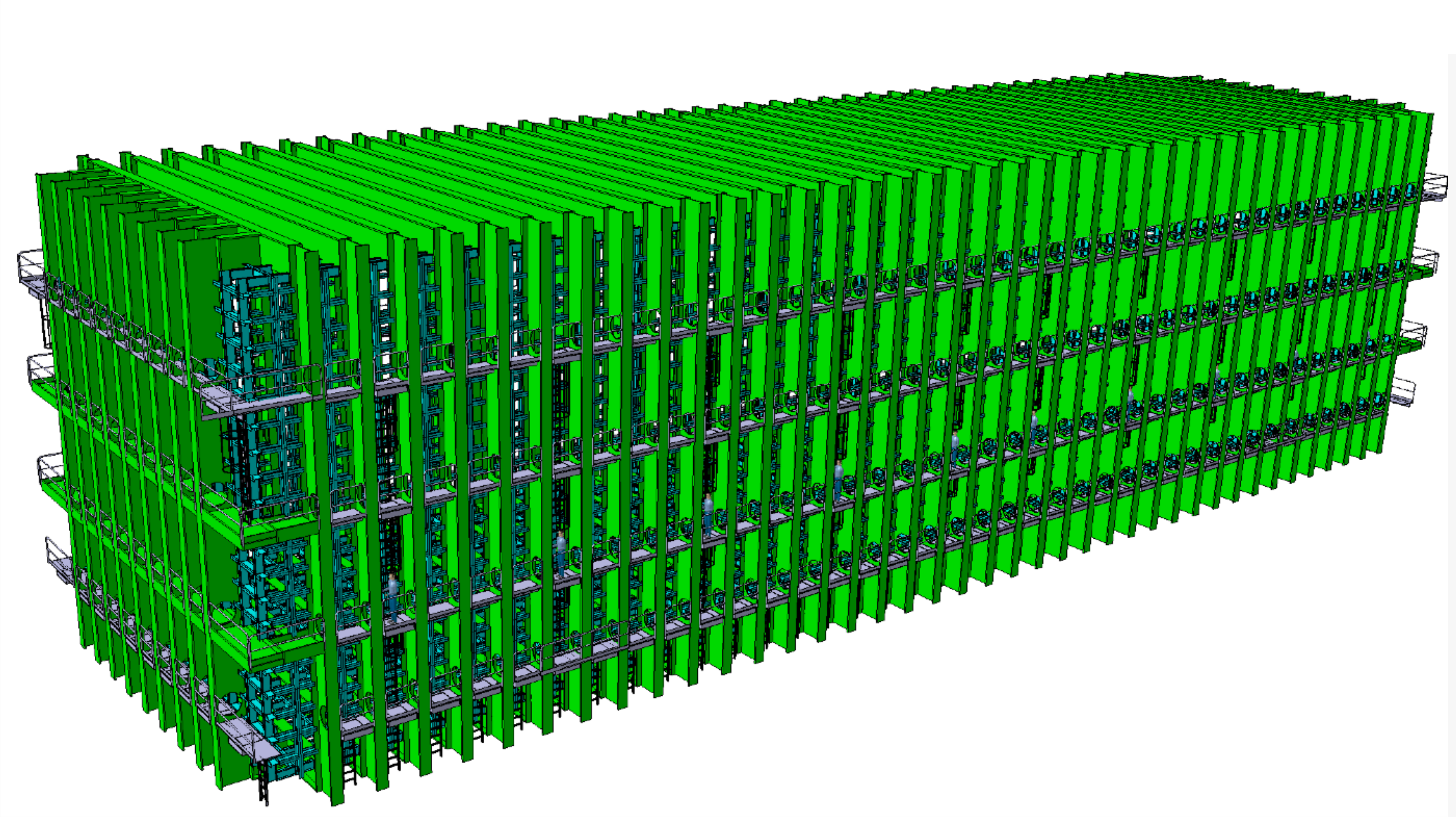


- Accelerator generating intense neutrino beam
- Near detector measuring neutrinos close to source
- Far detector 1300 km away from source and 1.48 km underground

TDAQ: no quick access and no large host lab in the vicinity!

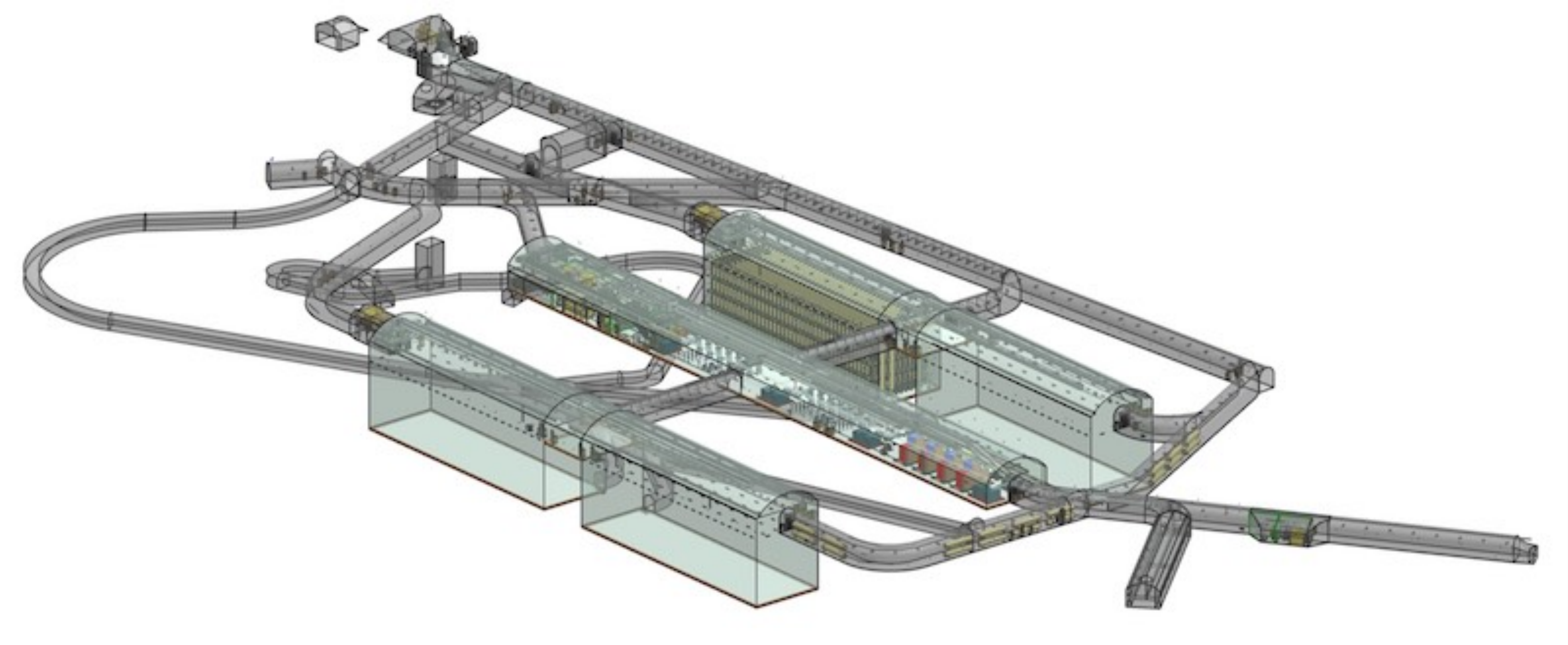
DUNE - The far detector

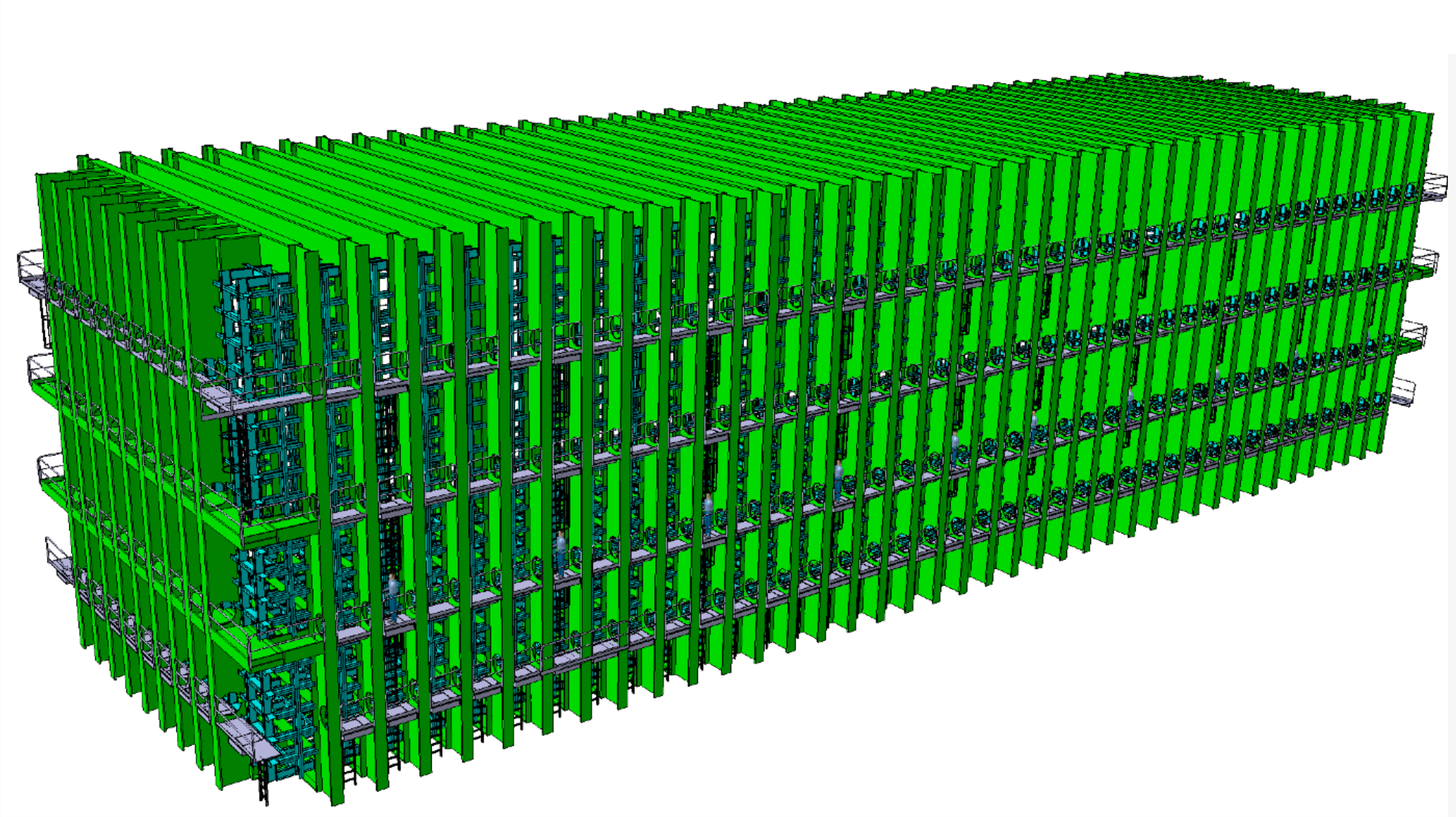




- 4 modules, each 18mx66mx17m (17 kton Lar)
- Detector:
  time projection chamber (slow) + photon detectors (fast)

DUNE - The far detector

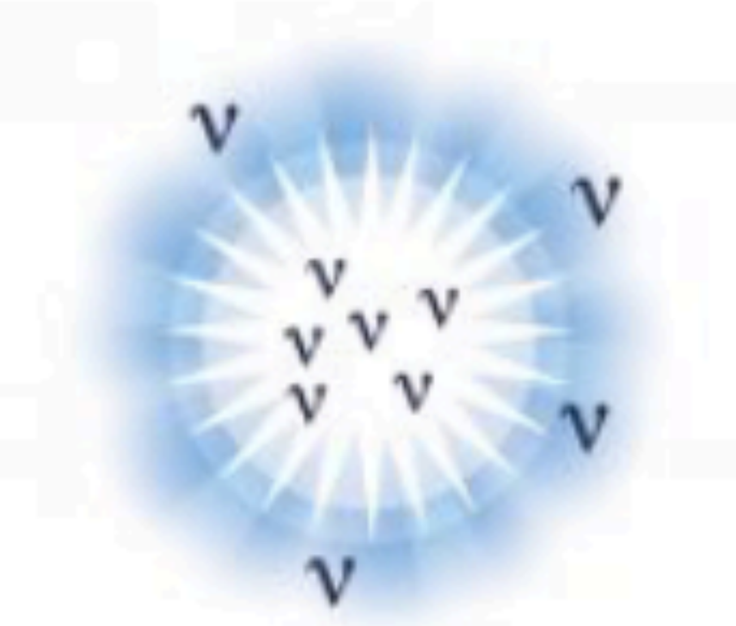




- 4 modules, each 18mx66mx17m (17 kton Lar)
- Detector:
  time projection chamber (slow) + photon detectors (fast)

TDAQ:

4 independent instances, synchronized to a common clock, supporting potentially different detector technologies



Origin of Matter

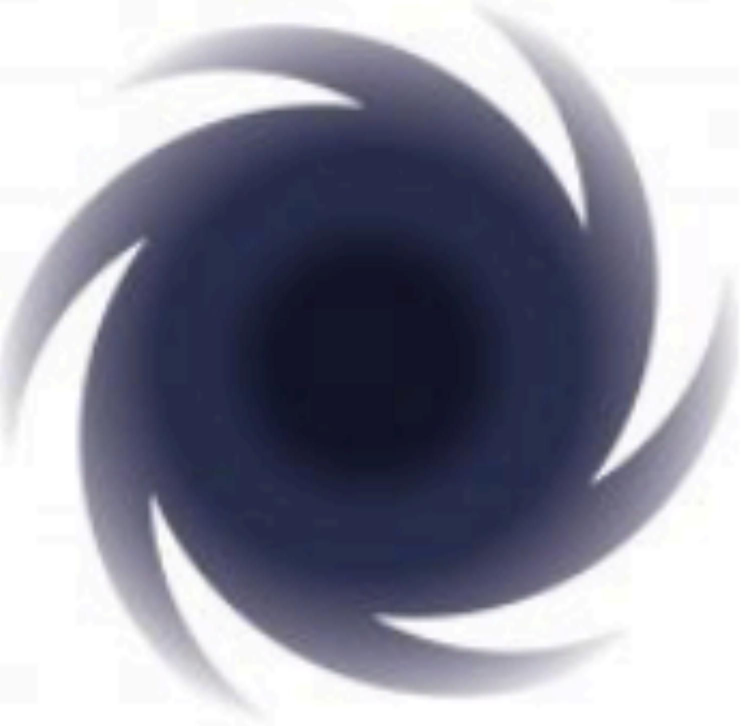
Neutrino beam -> external trigger possible

Could neutrinos be the reason that the universe is made of matter rather than antimatter? By exploring the phenomenon of neutrino oscillations, DUNE seeks to revolutionize our understanding of neutrinos and their role in the universe.

Unification of Forces

Very local, rare signature

With the world's largest cryogenic particle detector located deep underground, DUNE can search for signs of proton decay. This could reveal a relation between the stability of matter and the Grand Unification of forces, moving us closer to realizing Einstein's dream.



Black Hole Formation

Very distributed, rare signature

DUNE's observation of thousands of neutrinos from a core-collapse supernova in the Milky Way would allow us to peer inside a newly-formed neutron star and potentially witness the birth of a black hole.

DUNE - Signatures

Origin of Matter

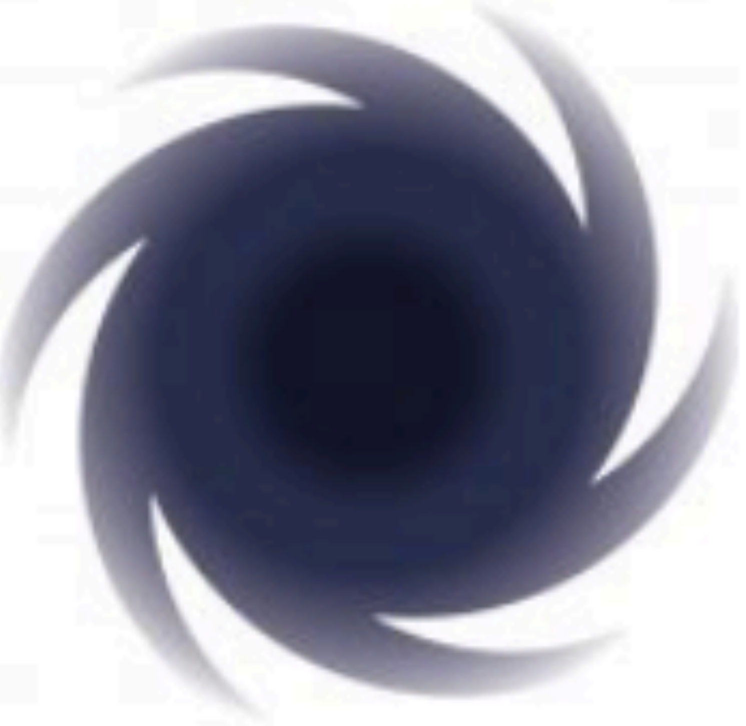
Neutrino beam -> external trigger possible

Could neutrinos be the reason that the universe is made of matter rather than antimatter? By exploring the phenomenon of neutrino oscillations, DUNE seeks to revolutionize our understanding of neutrinos and their role in the universe.

Unification of Forces

Very local, rare signature

With the world's largest cryogenic particle detector located deep underground, DUNE can search for signs of proton decay. This could reveal a relation between the stability of matter and the Grand Unification of forces, moving us closer to realizing Einstein's dream.



Black Hole Formation

Very distributed, rare signature

DUNE's observation of thousands of neutrinos from a core-collapse supernova in the Milky Way would allow us to peer inside a newly-formed neutron star and potentially witness the birth of a black hole.

TDAQ: active at "all" times!

A short digression on triggering…

Nomenclature

Globally triggered

- An “external device” decides what data is interesting
- There is a coherent event ID throughout the readout
- Front-end data are organized into fragments associated to the event ID

Locally triggered

- A local trigger element fires when data has to be readout (e.g. signal above threshold)
- The trigger is relative to individual or groups of channels, not to the full front-end
- The readout can process incoming data to create fragments corresponding to the trigger
- There is no concept of a global event ID at the readout level

Continuous readout

- The front-end sends data to the readout at a fixed rate, irrespective of the data content
- Data rate and data size (if there is no zero suppression) are constant in input
- There is no indication for the readout on how to group front-end data into fragments corresponding to a physics event

Use cases for different readouts

Colliders

- Normally use global trigger: if something interesting has been seen somewhere, take all the data corresponding to trigger bunch crossing.

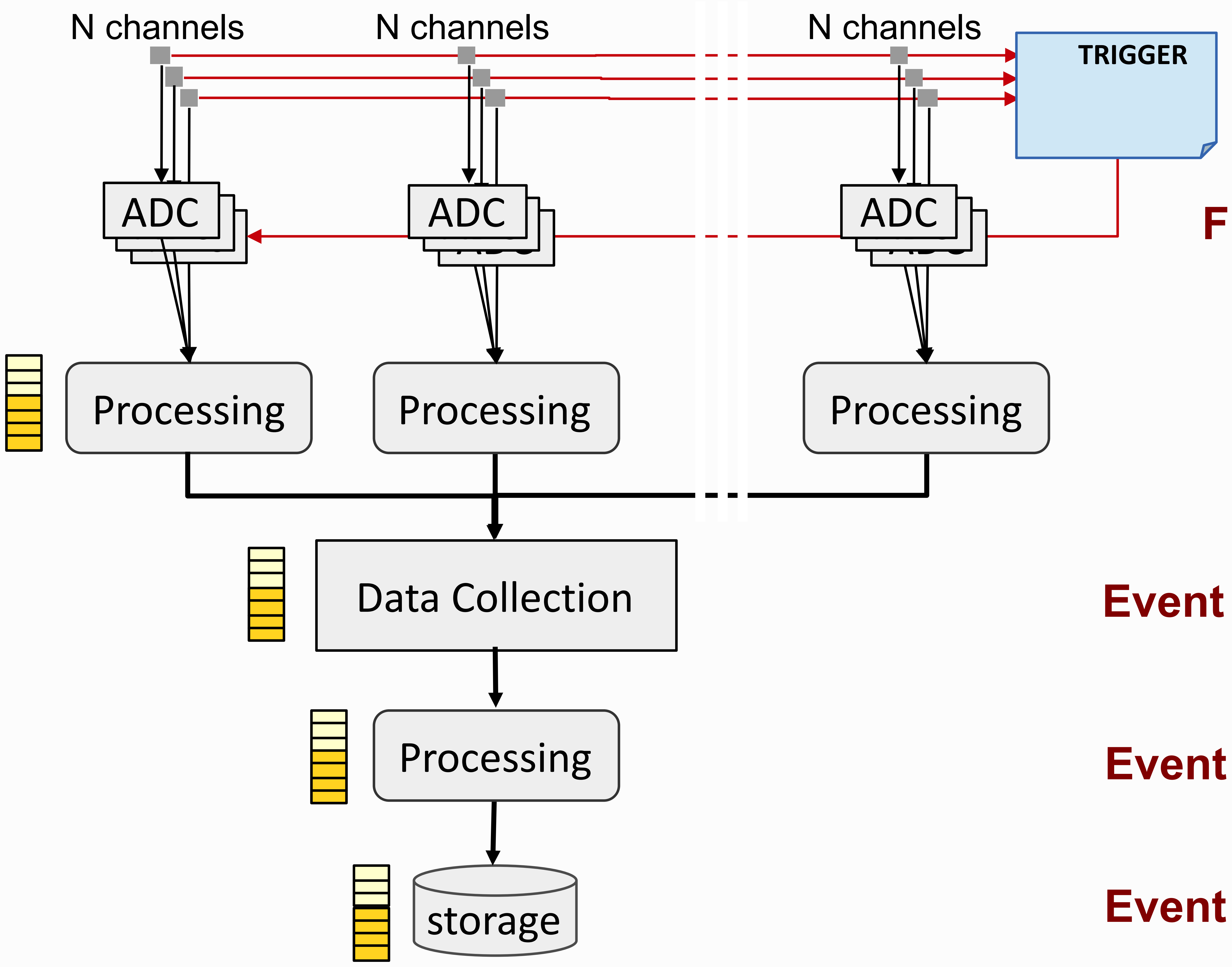
Large distributed telescopes

- Often use local trigger: readout data for the portions of the detector that have seen something.

Very slow detectors

- Sometimes use continuous readout: sample the analogue signals at a fixed rate and let the downstream DAQ decide any interesting signal is present.

Globally triggered



Front-End | data digitization, data buffering

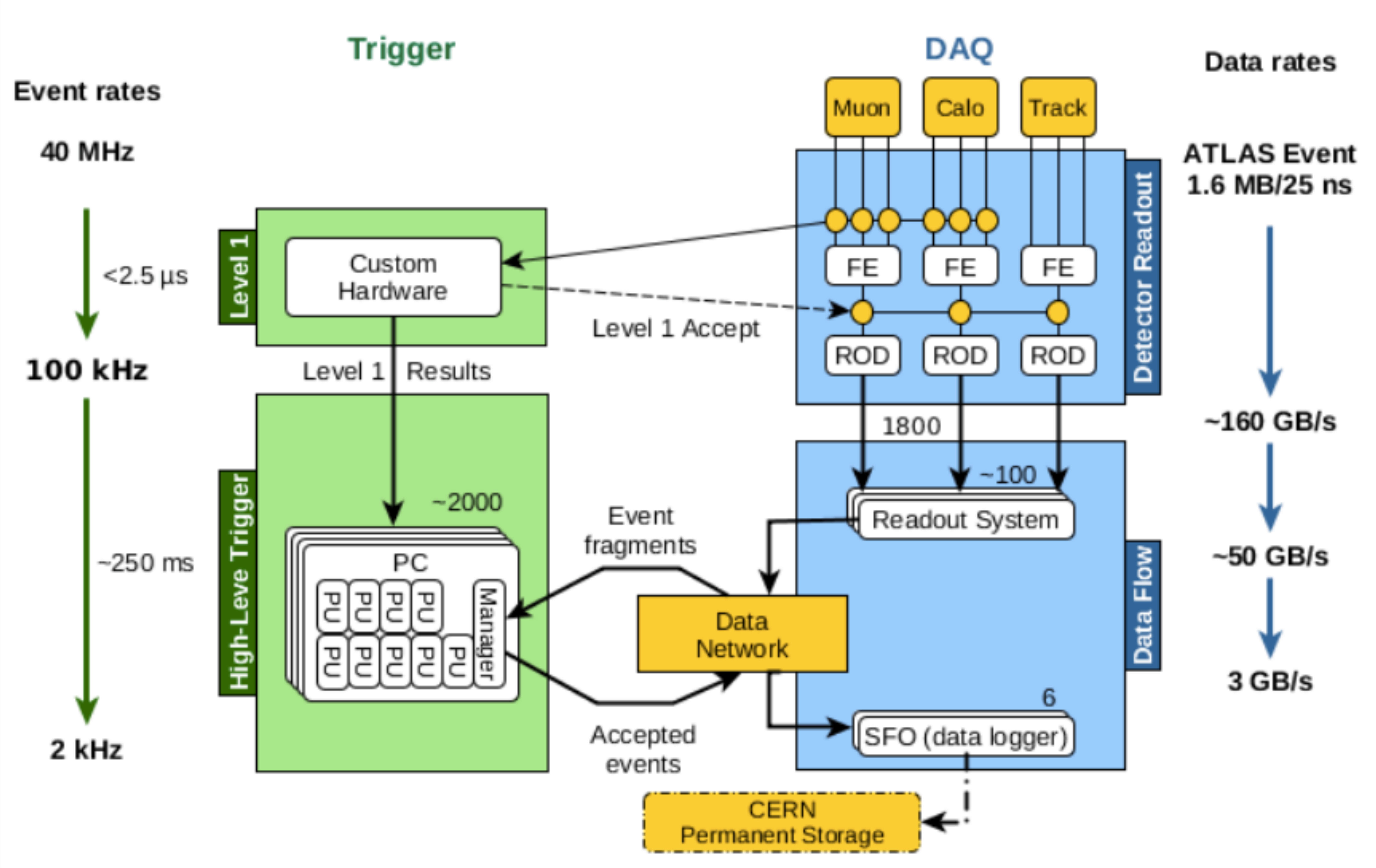
Readout | data extraction, data formatting, data buffering

Event Building | event assembly, event buffering

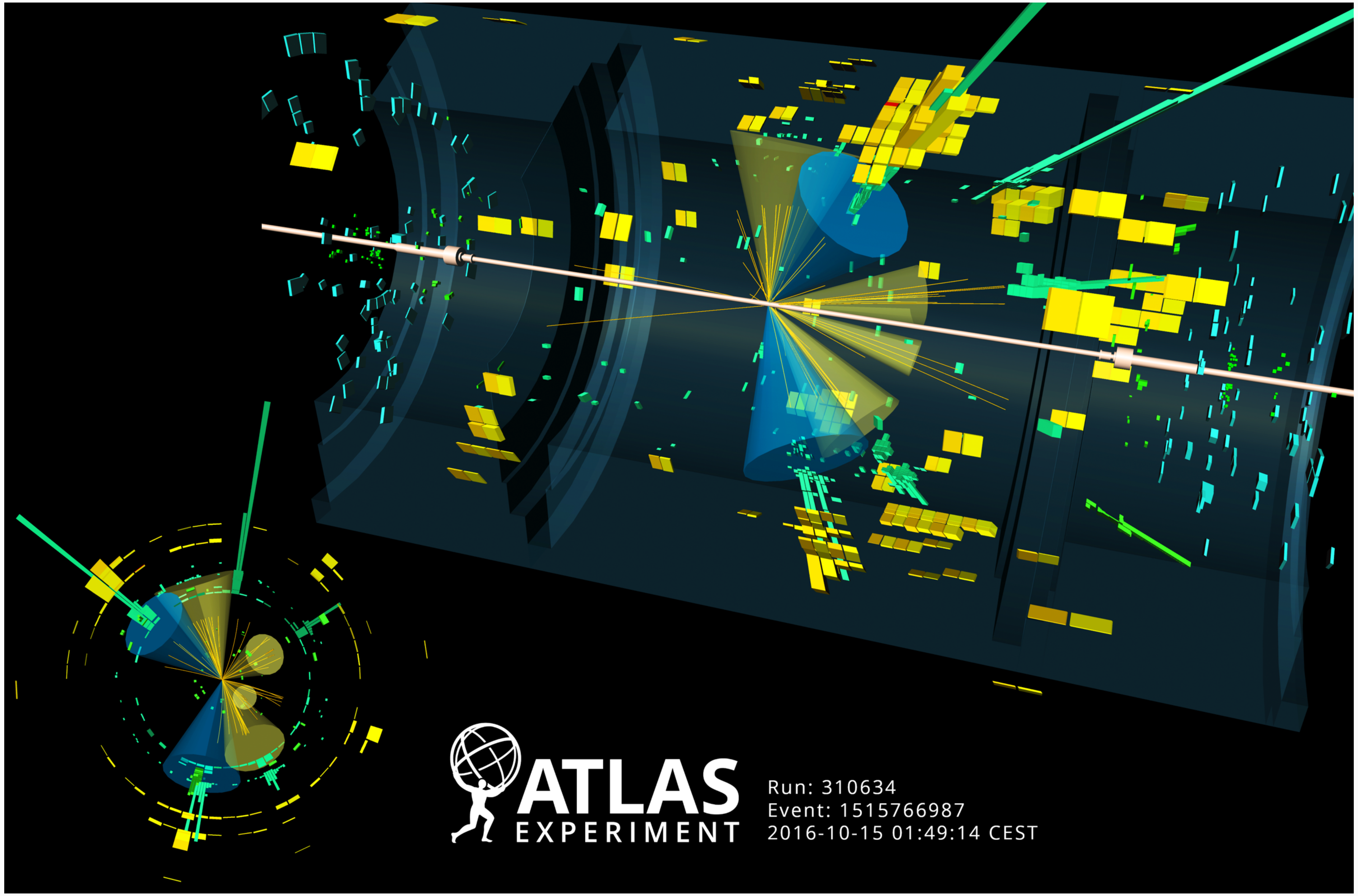
Event Filtering | event rejection, event buffering

Event Logging | file storage, file buffering

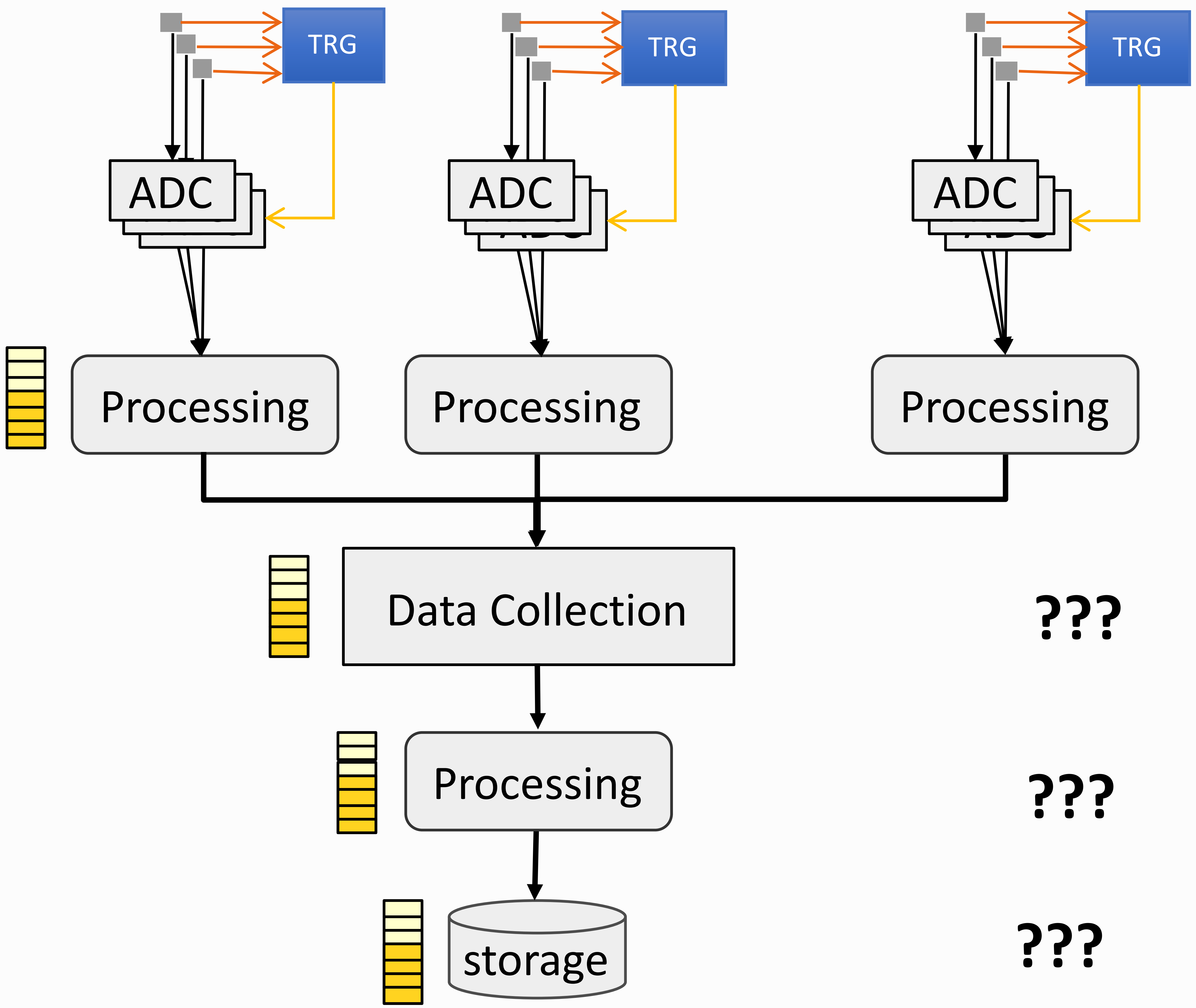
Globally triggered: ATLAS@LHC



An event


ATLAS
EXPERIMENT
Run: 310634
Event: 1515766987
2016-10-15 01:49:14 CEST

Locally triggered



Front-End — data digitization, data buffering

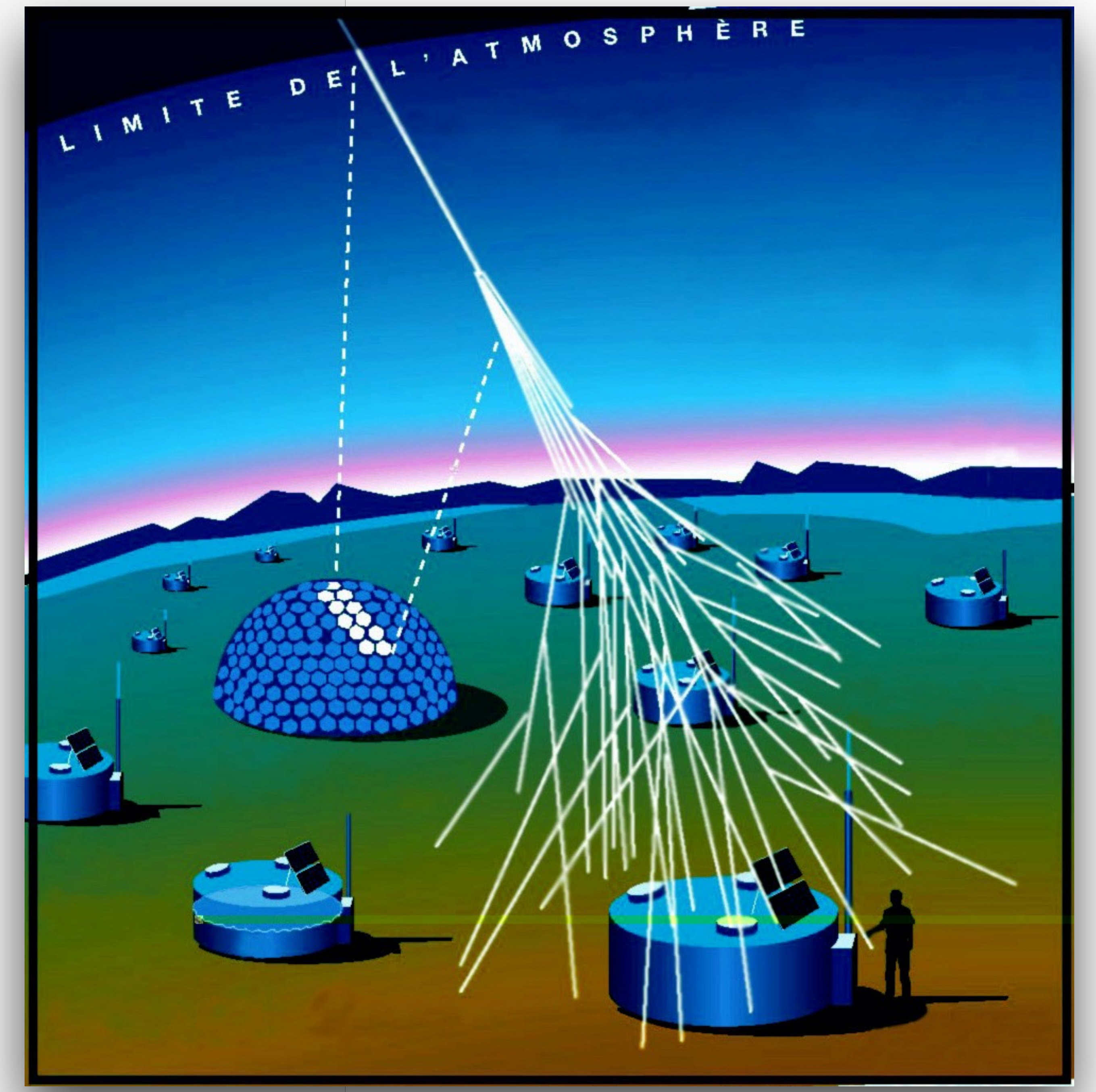
Readout — data extraction, data formatting, data buffering

??? **Event Building** — event assembly, event buffering

??? **Event Filtering** — event rejection, event buffering

??? **Event Logging** — file storage, file buffering

Locally triggered: Auger observatory

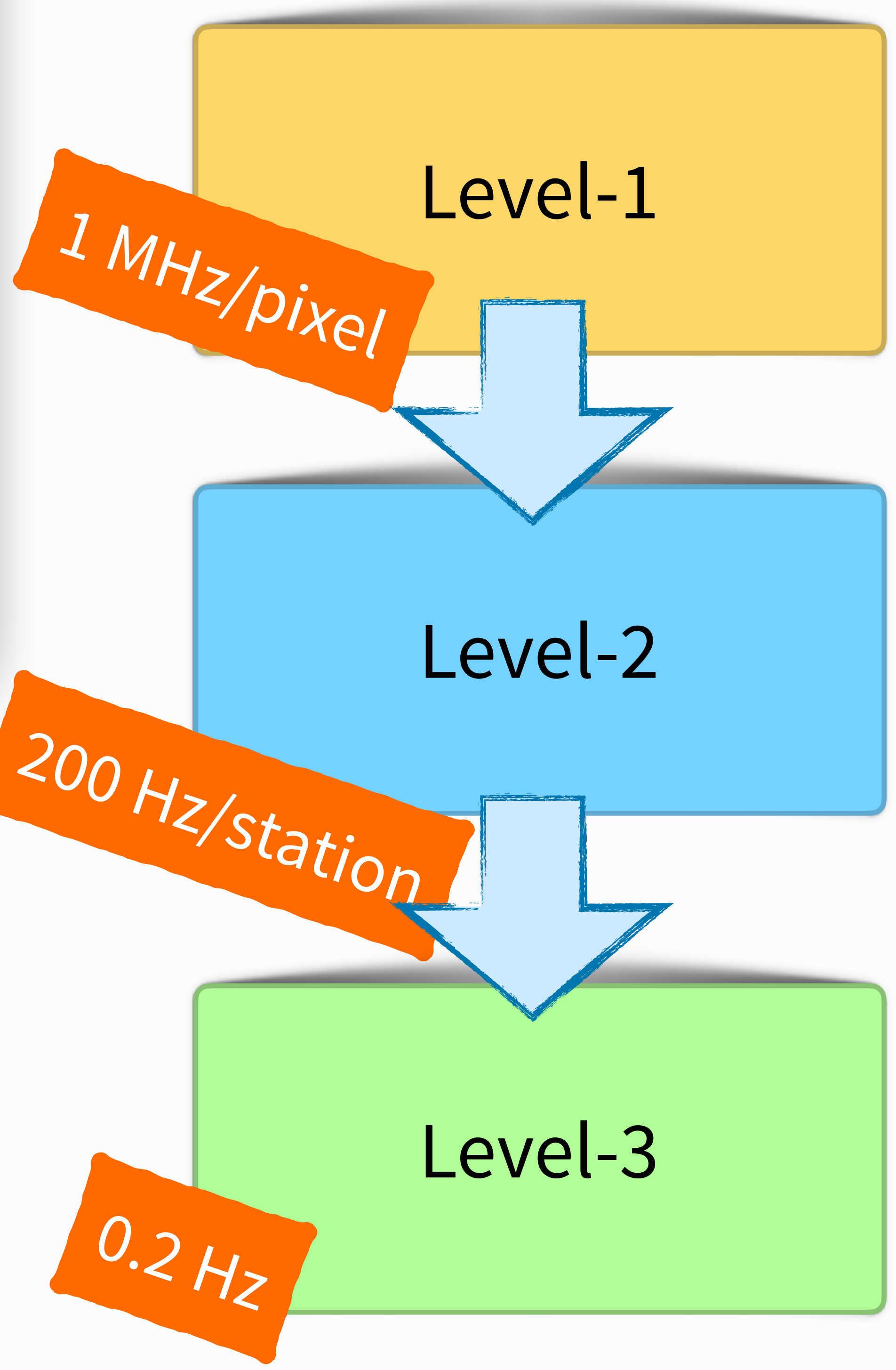


Surface Detector: array of ~1600 water Cherenkov stations over 3000 km^2 on ground, to identify secondary particles
Florescent Detector: 4 UV telescopes measure the shower Energy longitudinally

Detect air showers generated by cosmic rays above 10^{17} eV

- Expected rate < 1/km2/century,
- 2 large area fluorescence detectors.

3-level trigger installed on each detector

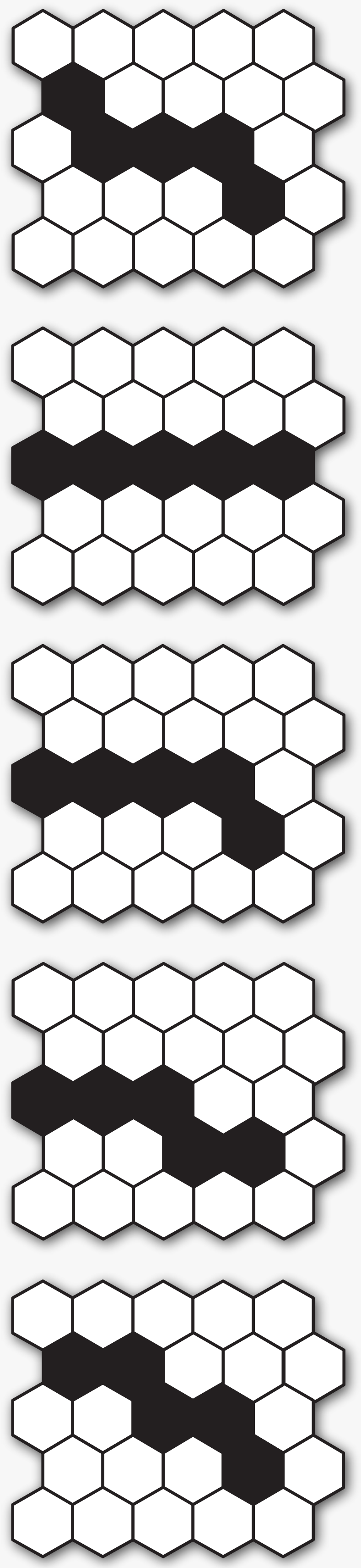


L1: (local) select active pixels
- ADC counts > threshold
- ADC digitises every 100 ns (time resolution)
- ADC values stored for 100 μs in local buffers
- Synchronised via GPS clock signal

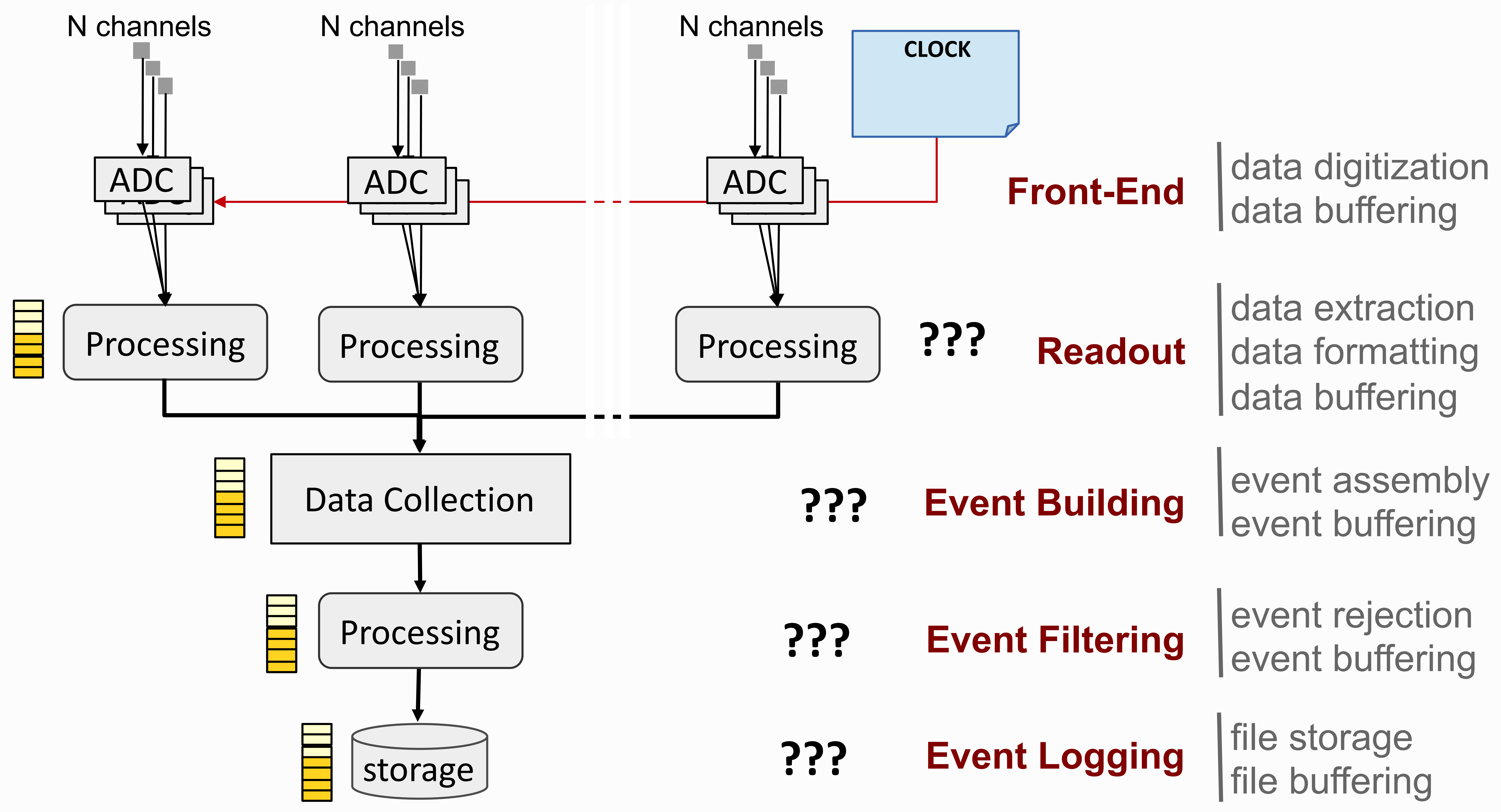
L2: (local) identifies track segments
- Geometrical criteria with programmable pattern recognition algorithms

L3: (central) 3-D correlation between L2 triggers

Example of L2 patterns



Continuous

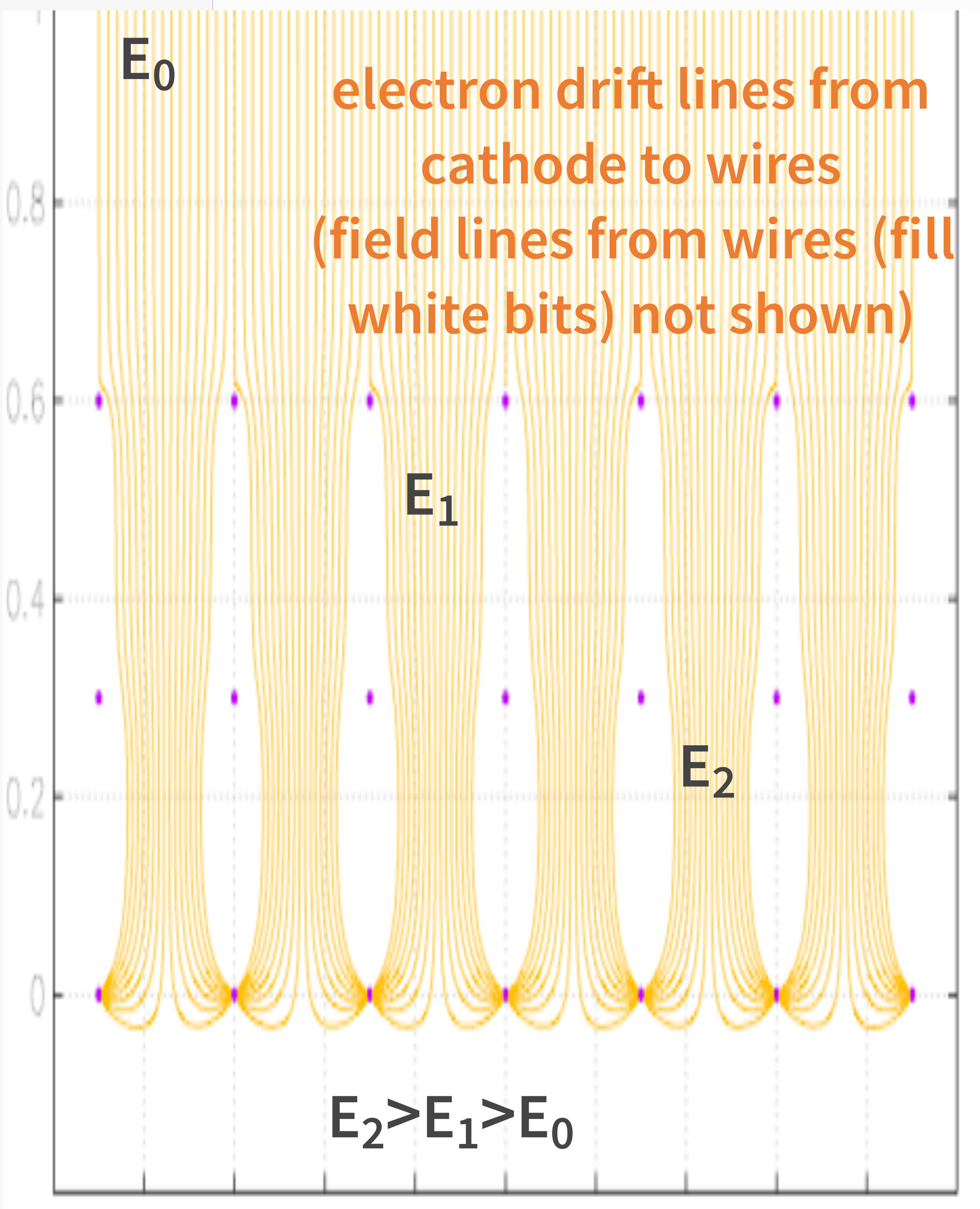


N channels
N channels
N channels
CLOCK
ADC
ADC
ADC
Processing
Processing
Processing
???
Data Collection
???
Processing
???
storage
???
Front-End
data digitization
data buffering
Readout
data extraction
data formatting
data buffering
Event Building
event assembly
event buffering
Event Filtering
event rejection
event buffering
Event Logging
file storage
file buffering

Not events, but rather a movie

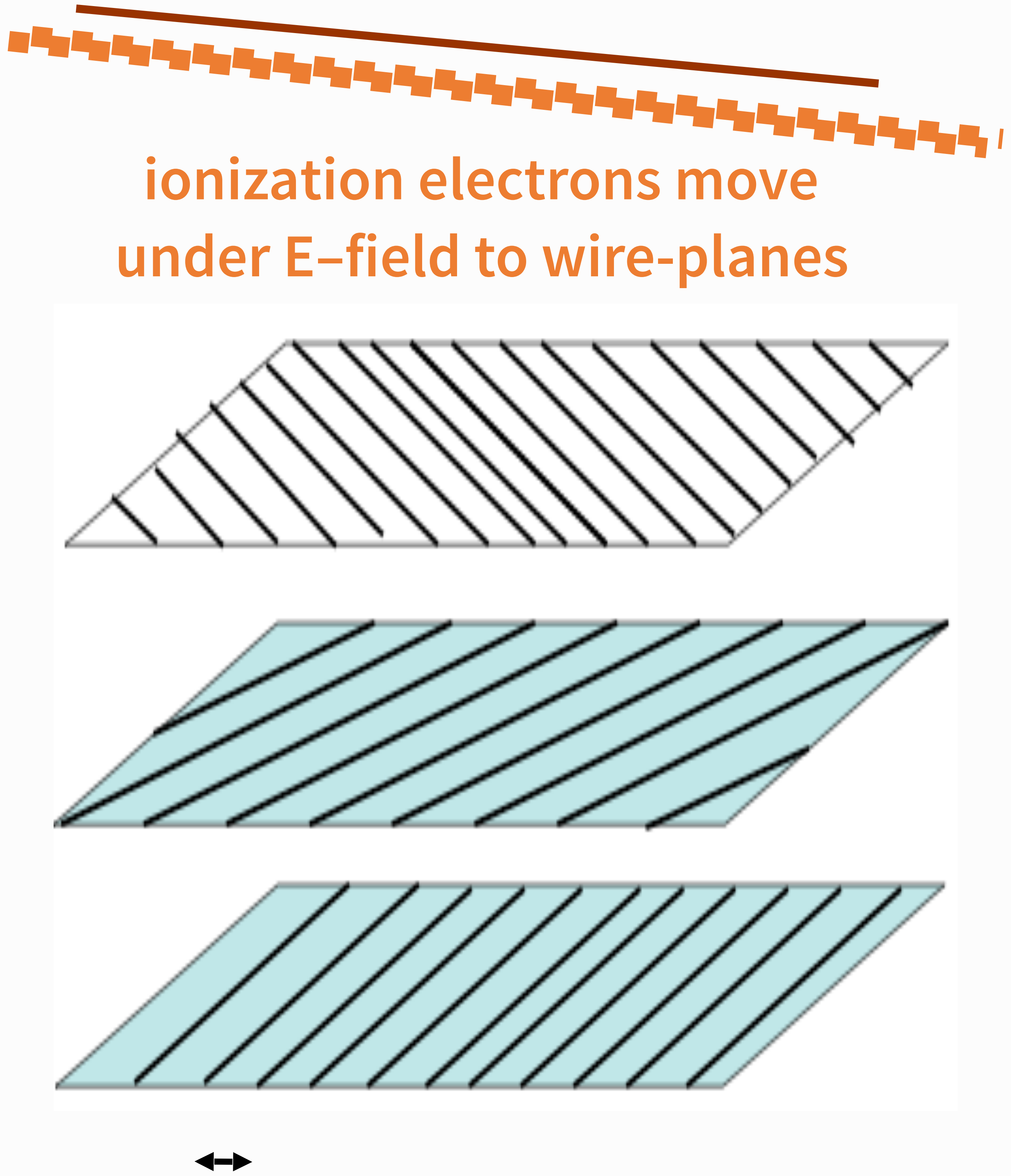
. . .end of digression, back to DUNE

TPC working principle



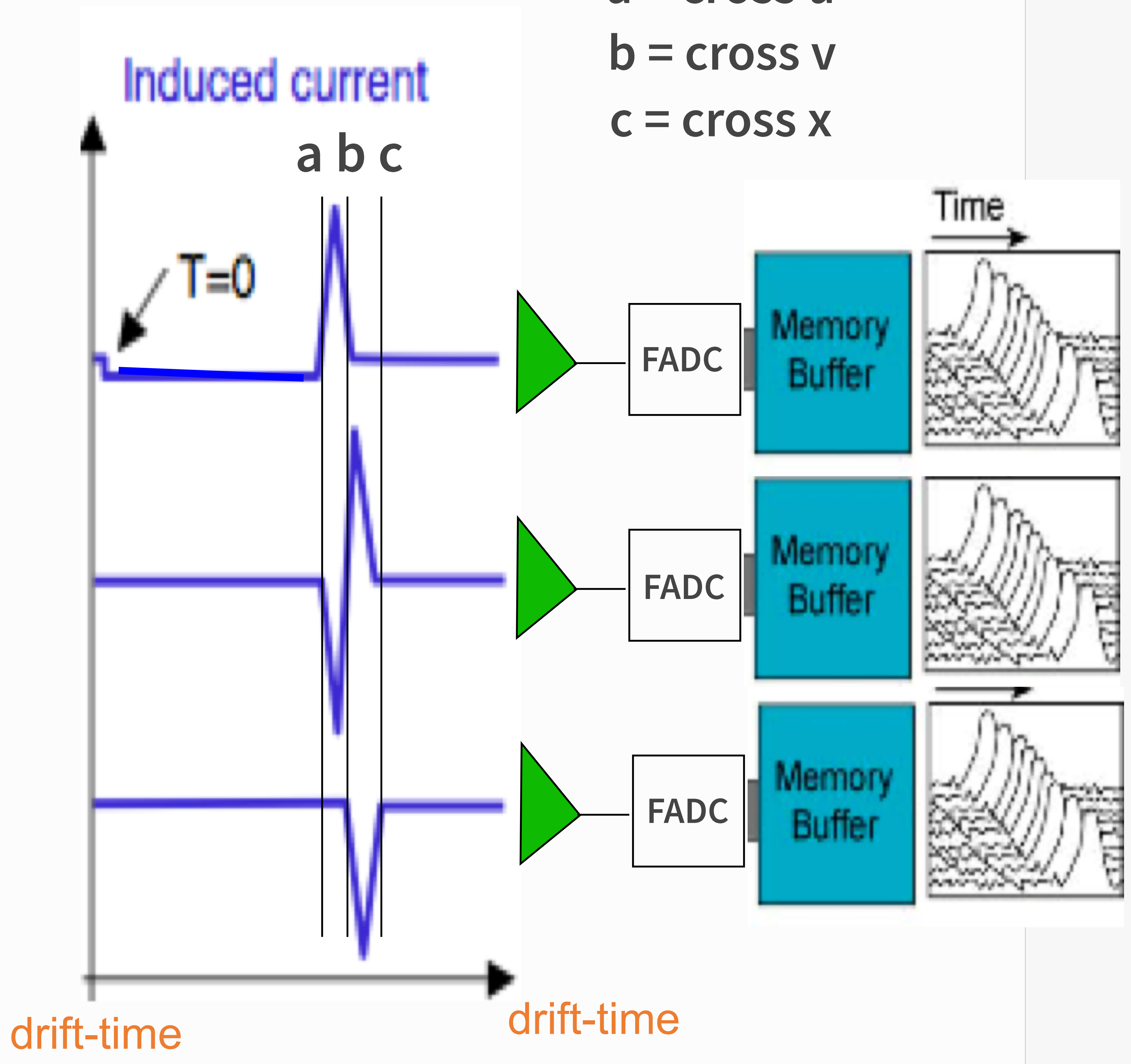
Bunneman, Cranshaw, Harvey,
Can. J. Res. 27 (1949) 191

charged particle ionizes argon



p
p~d~3-5 mm

a = cross u
b = cross v
c = cross x



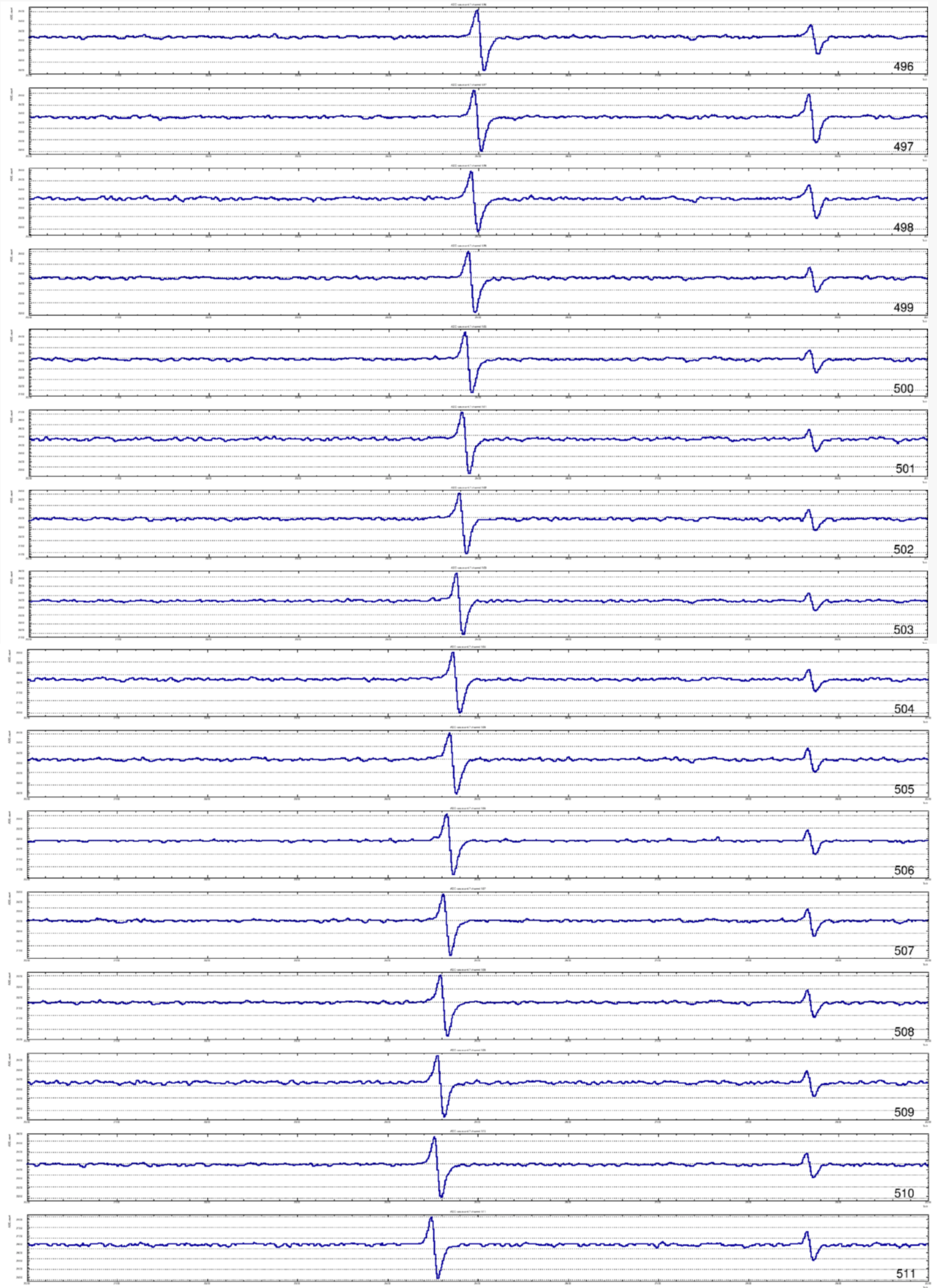
Gatti, Padovini, Quartapelle, Greenlaw, Radeka
IEEE Trans. NS-26 (2) (1979) 2910]

Slide from S. Pordes

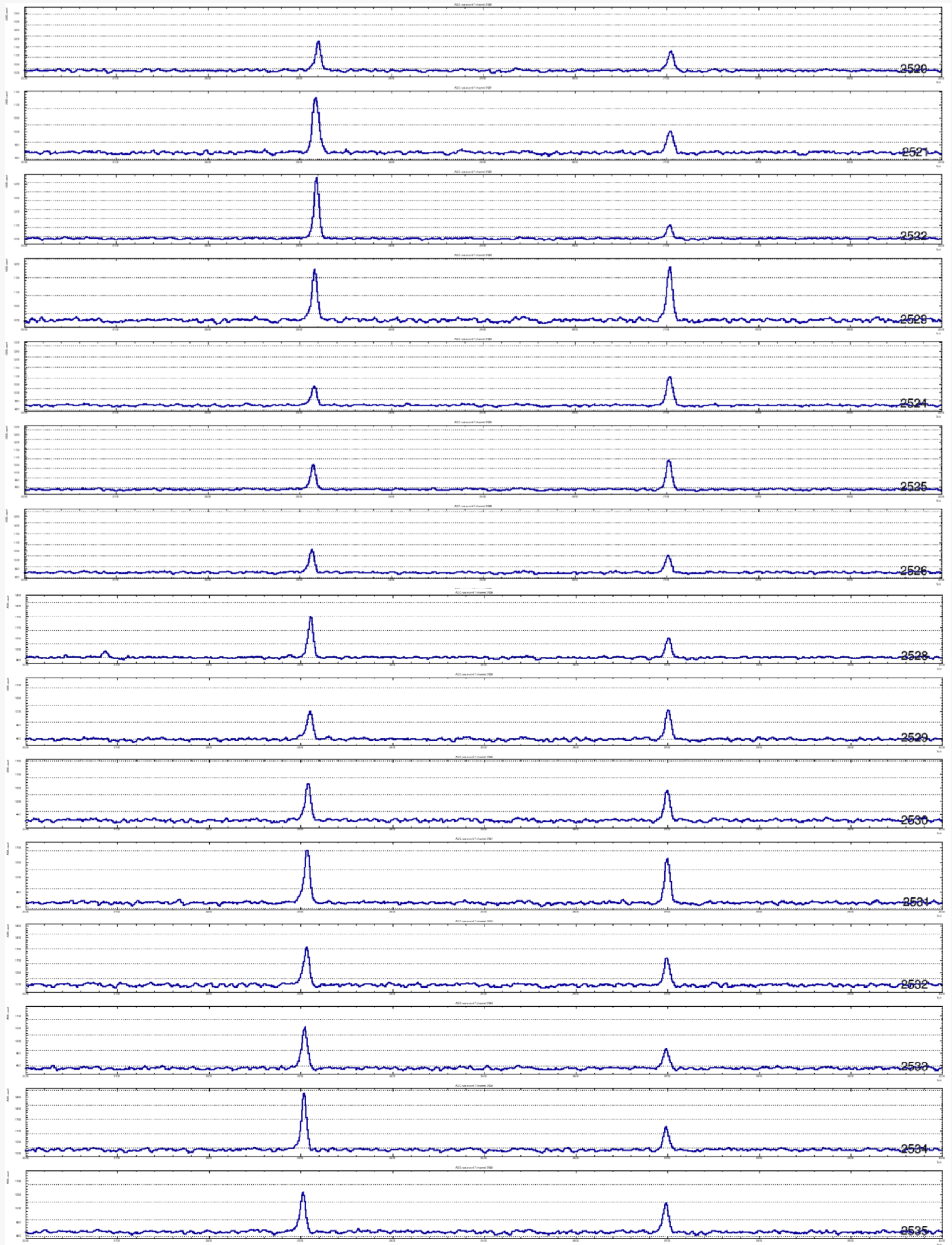
What does the raw signals look like (16 sequential wires)

Slide from S. Pordes

induction plane signals



collection plane signals



DUNE Front-End readout - 1 Single-Phase Module

DUNE mixes **continuous** readout (TPC) and **locally triggered** readout (Photon detectors)

- TPC sampling rate = 2 MHz
- Photon detectors sampling rate = 150 MHz (but data only when there is a signal)
- TPC wires (single phase technology) = 384000 per cryostat
- 384 k channels (12 bit ADC) @ 2 MHz = 9.2 Tb/s (dominates data size)
- Adding all up the TDAQ has to sustain a readout of **~5 TB/s**
  - Sounds very much like HL-LHC...

DUNE post-readout data selection

The post-readout system to combine data snippets into time windows of interesting detector regions

- In DUNE the "window" can be anything from **few ms to ~100s** for the supernova core collapse
- The data corresponding to a trigger can have a size ranging from **<< 1 GB to ~150 TB**!

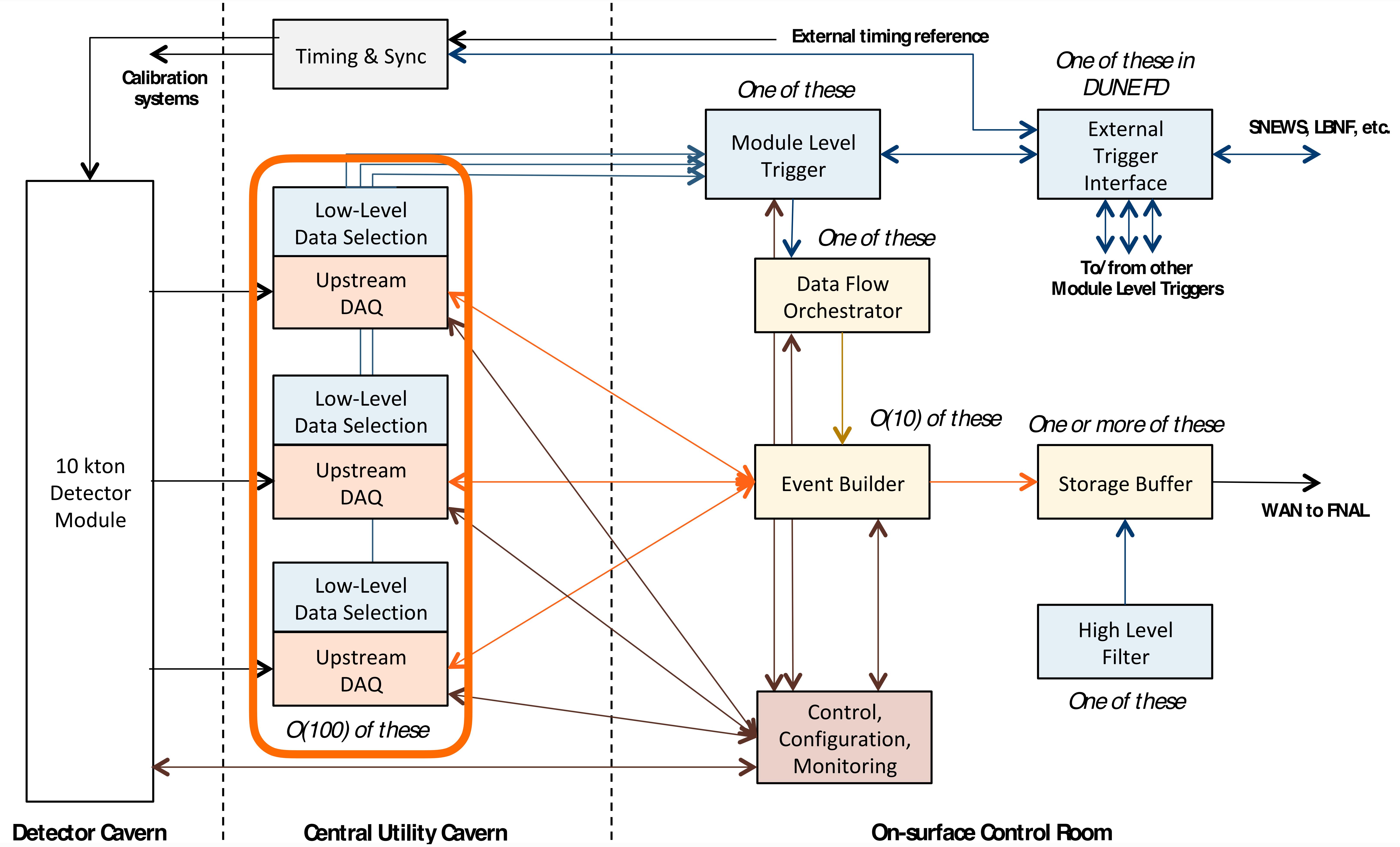
The rate of events varies widely from **few Hz to <<1/month**

The data selection needs to accumulate a view on the activity of the detector over several seconds to identify some signatures (SNB)

- The readout needs to have **very large** buffers to accommodate for the decision **latency**

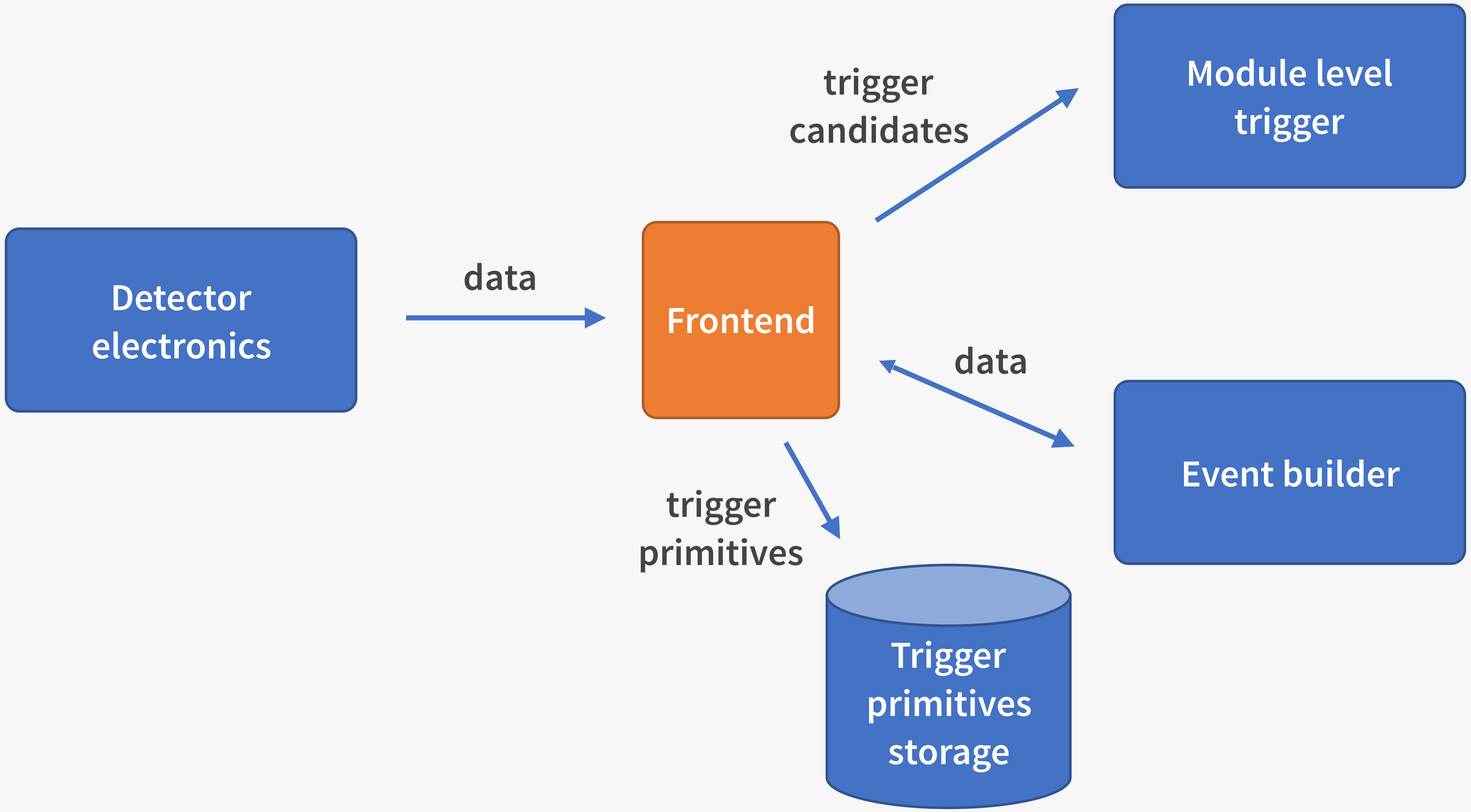
Baseline design of the DUNE data flow

DUNE data flow

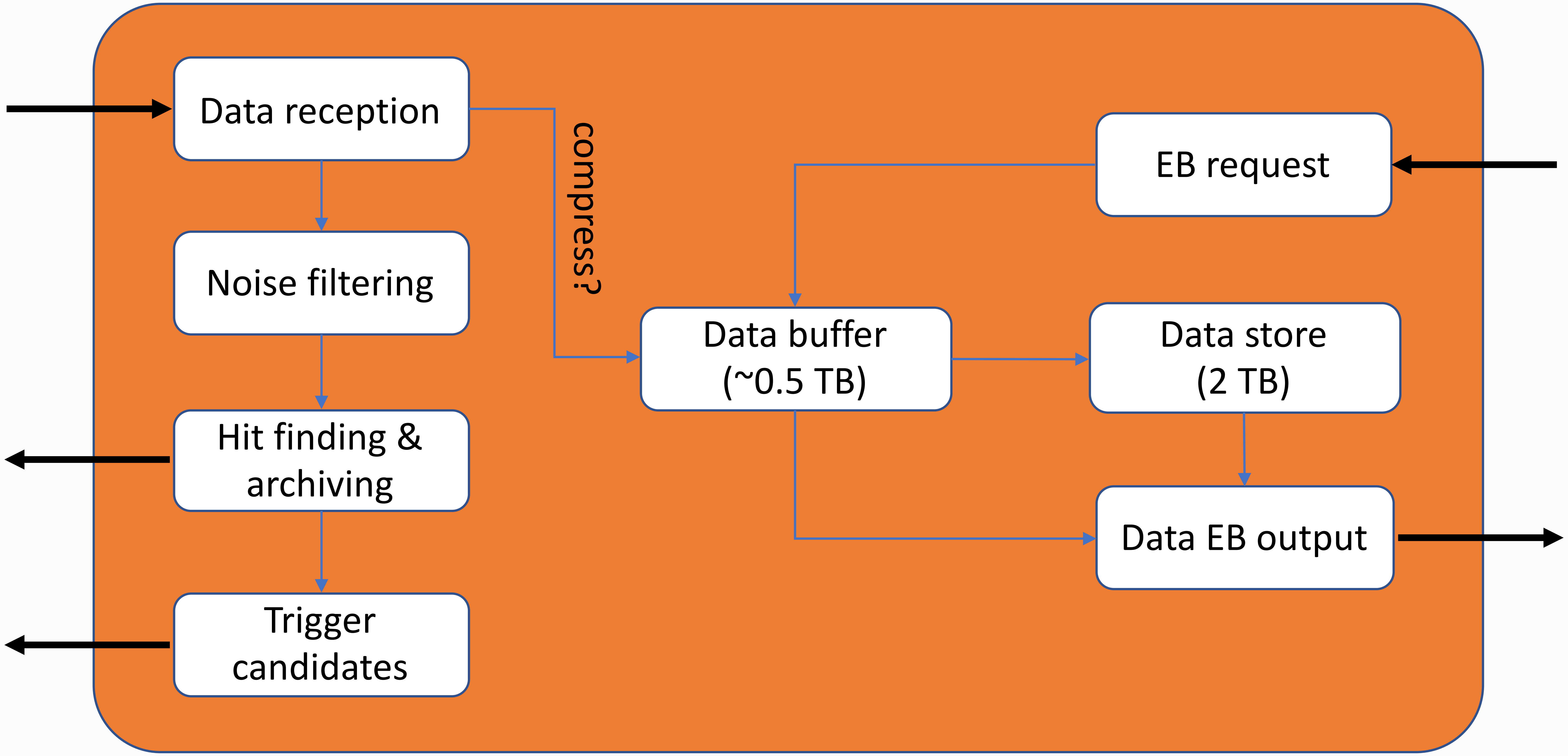


Front end - data flow interfaces

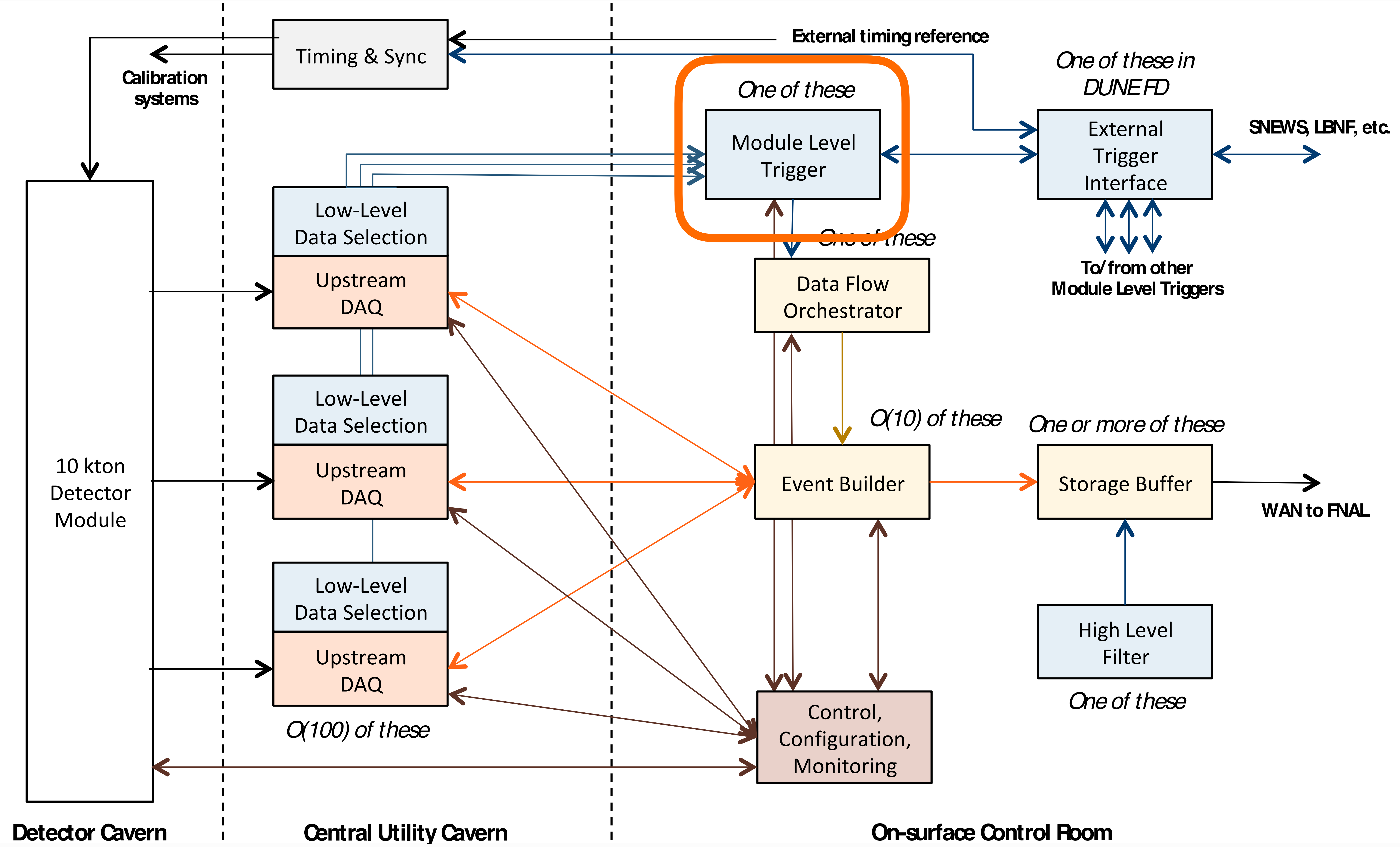
- Many similar modular elements (~100 / module)



Front end - functional blocks

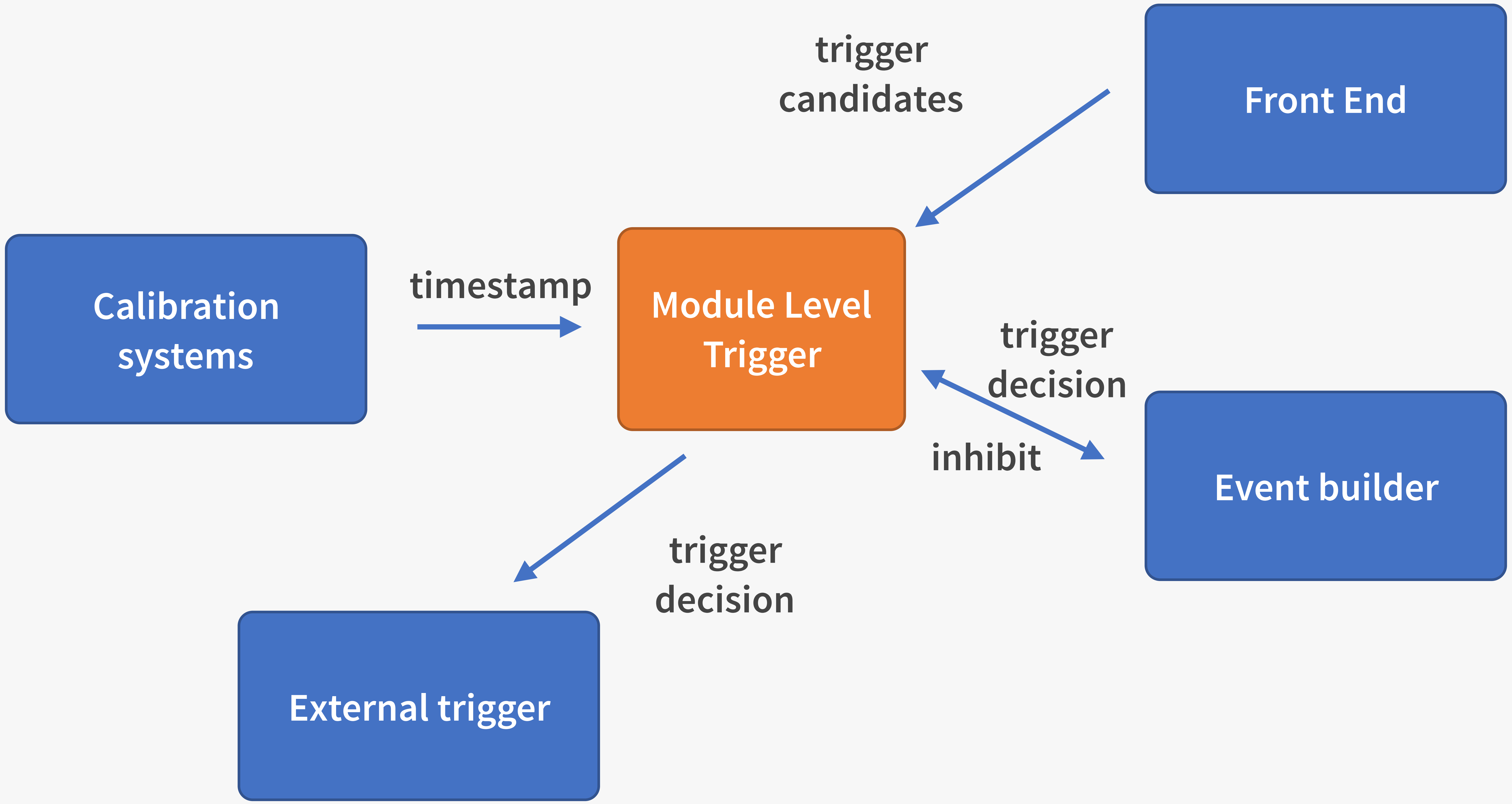


DUNE data flow

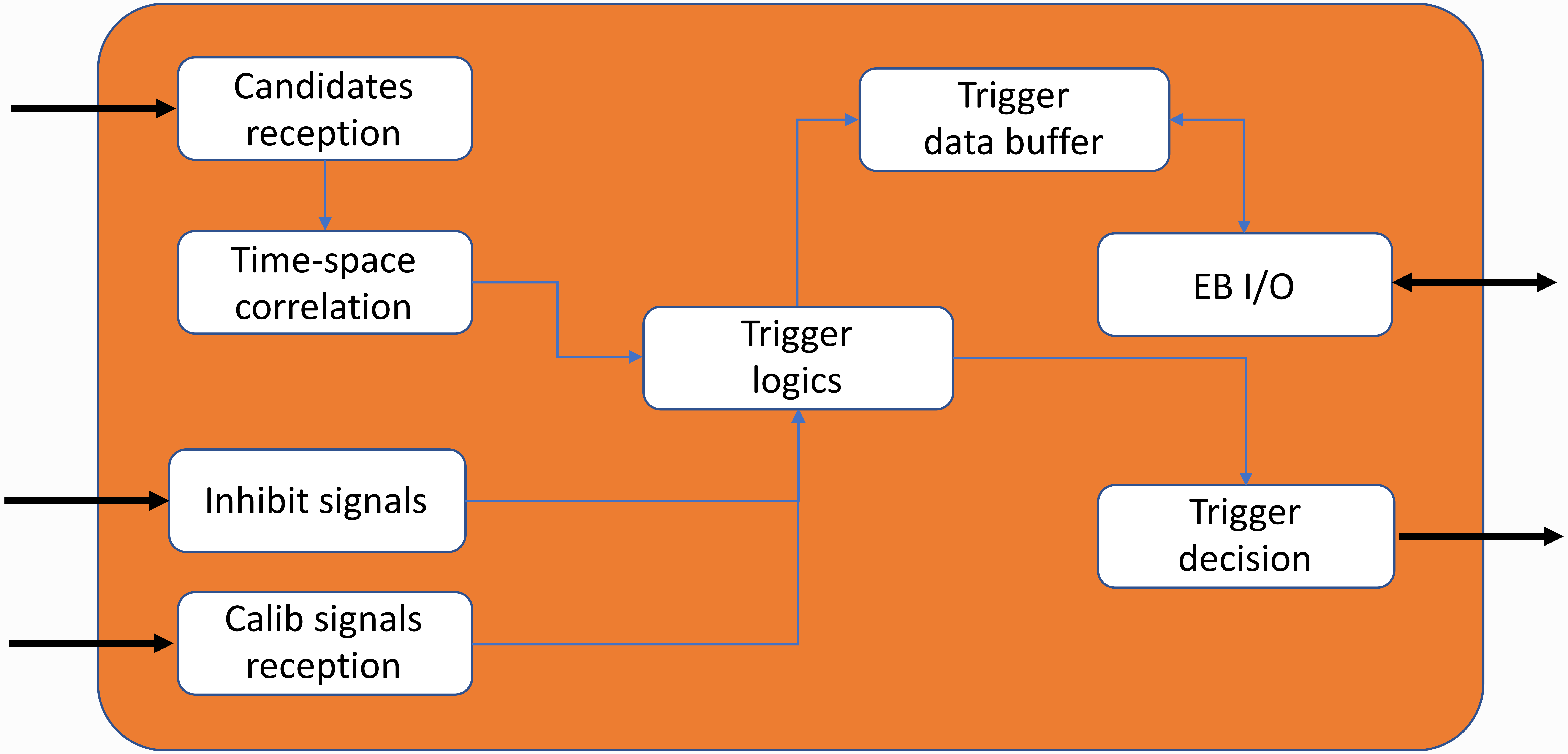


Module Level Trigger - data flow interfaces

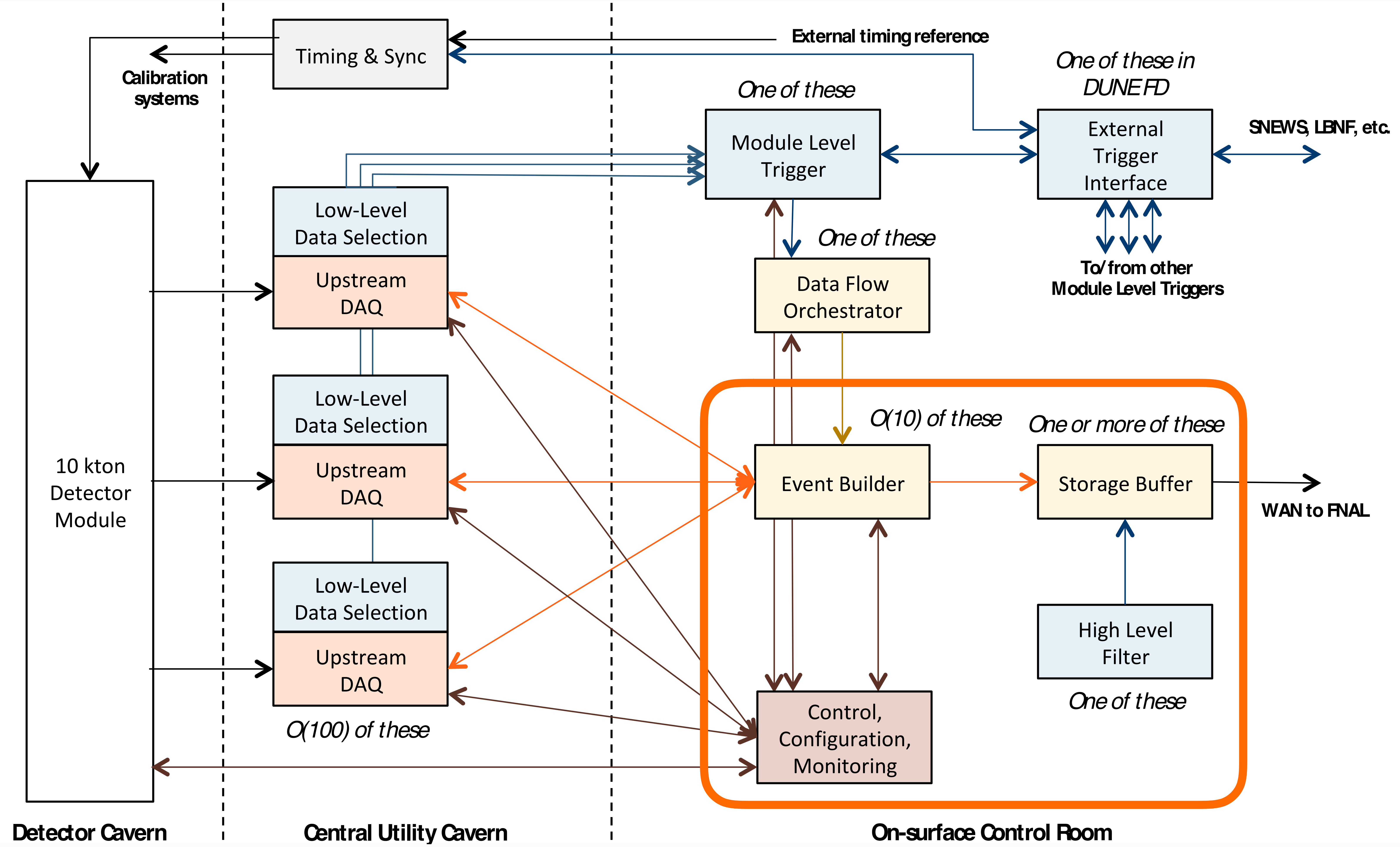
- 1 per module



Module Level Trigger - functional blocks

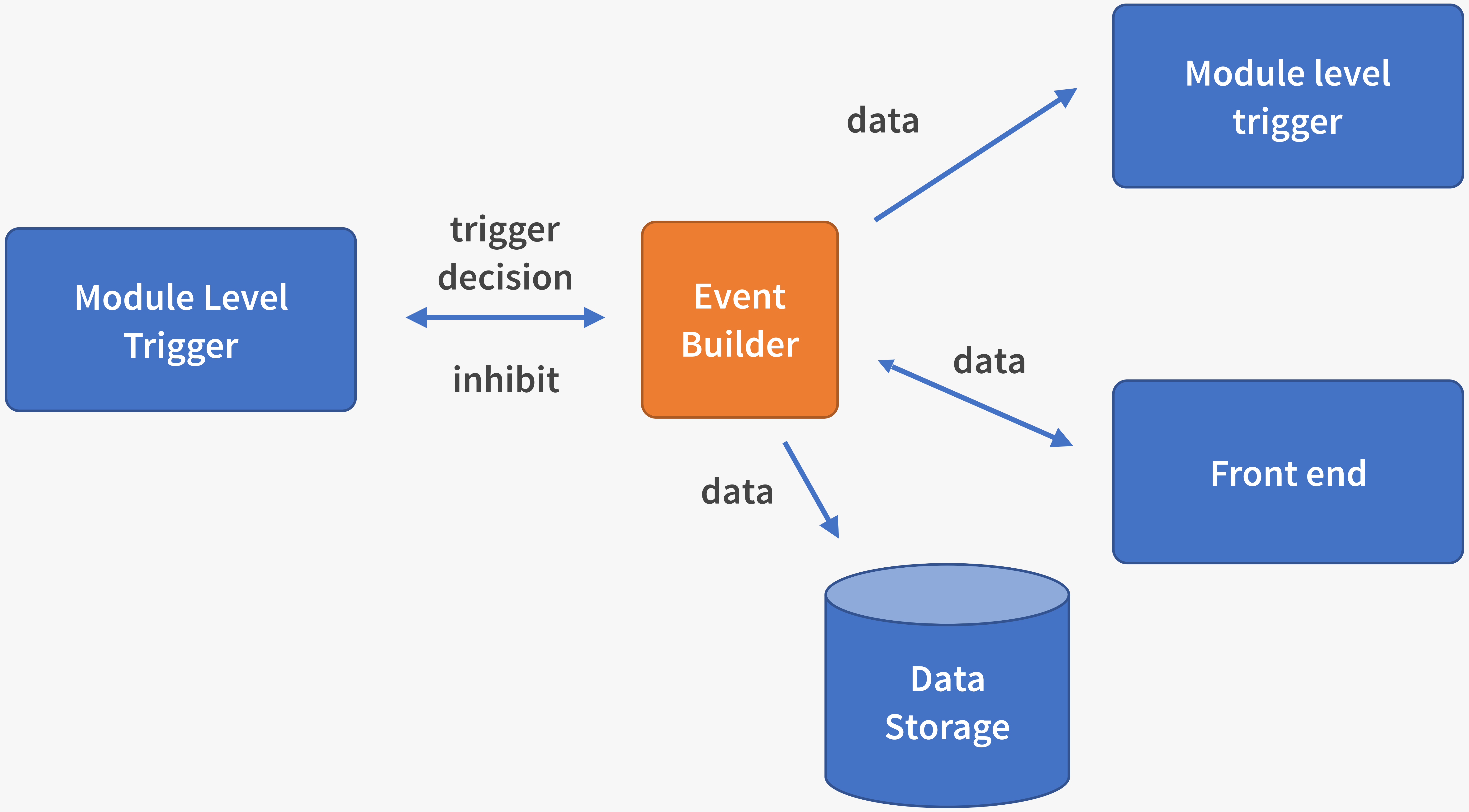


DUNE data flow

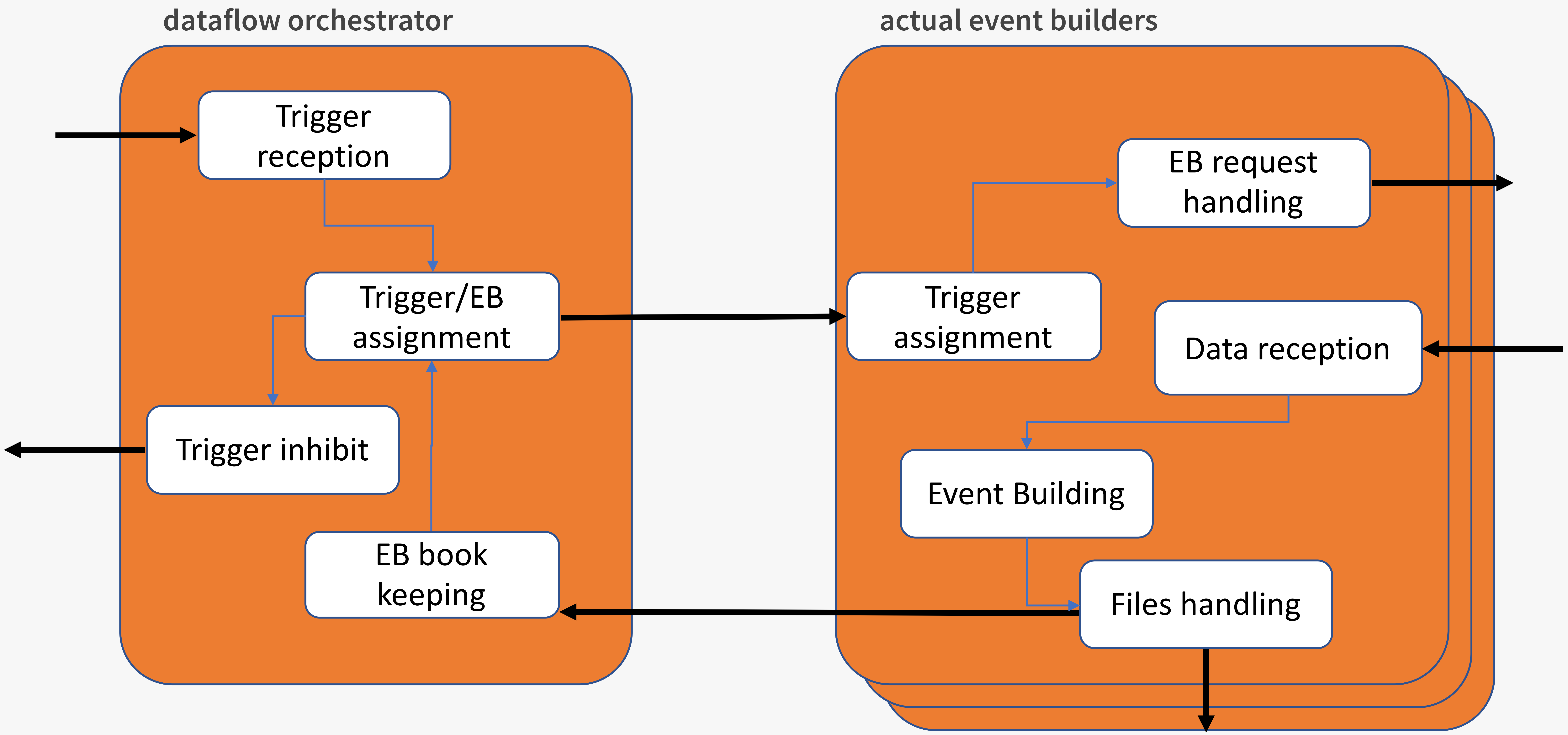


Event builder - data flow interfaces

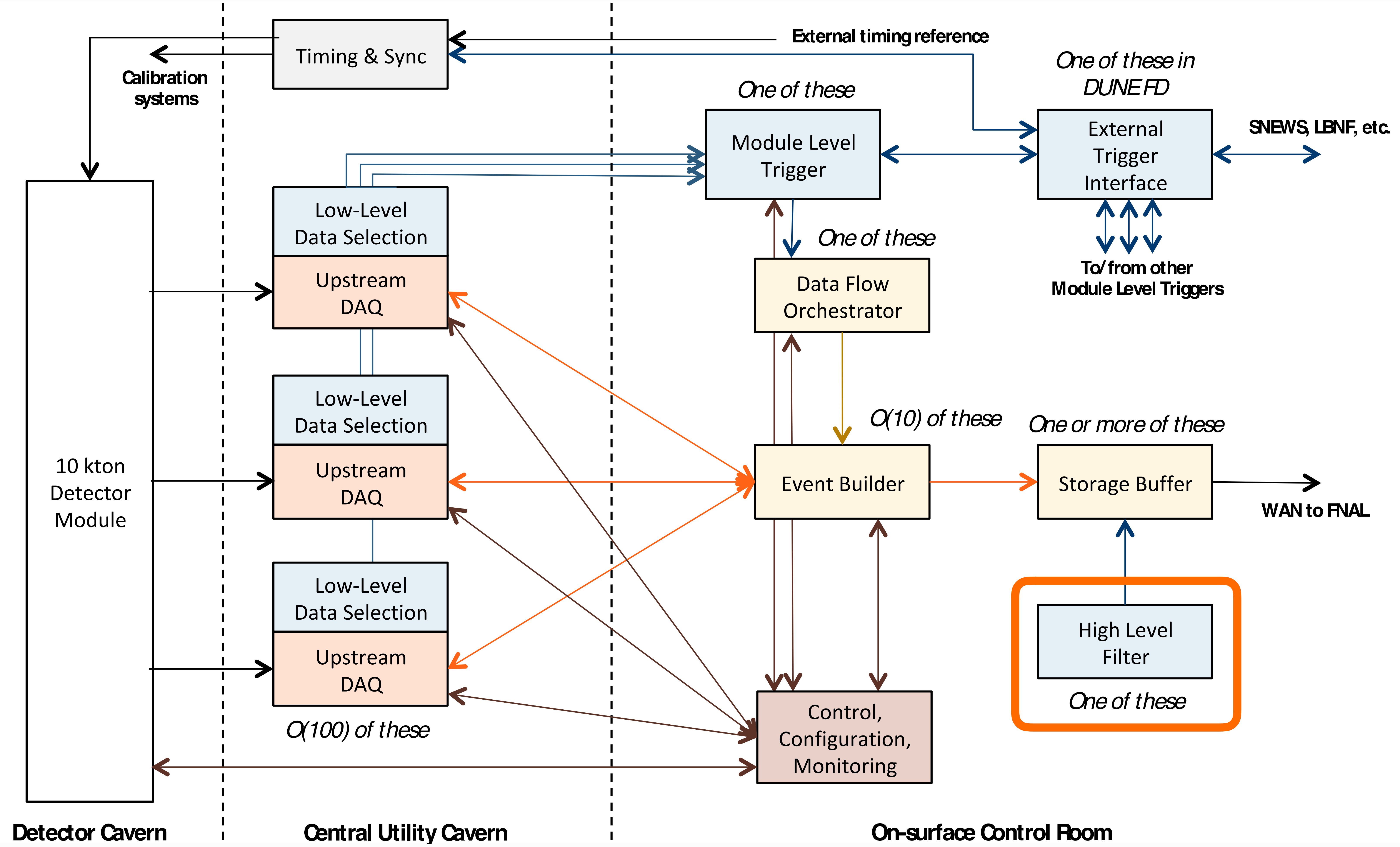
- One orchestrator, few similar building elements (~10 / module)



Event builder - functional blocks

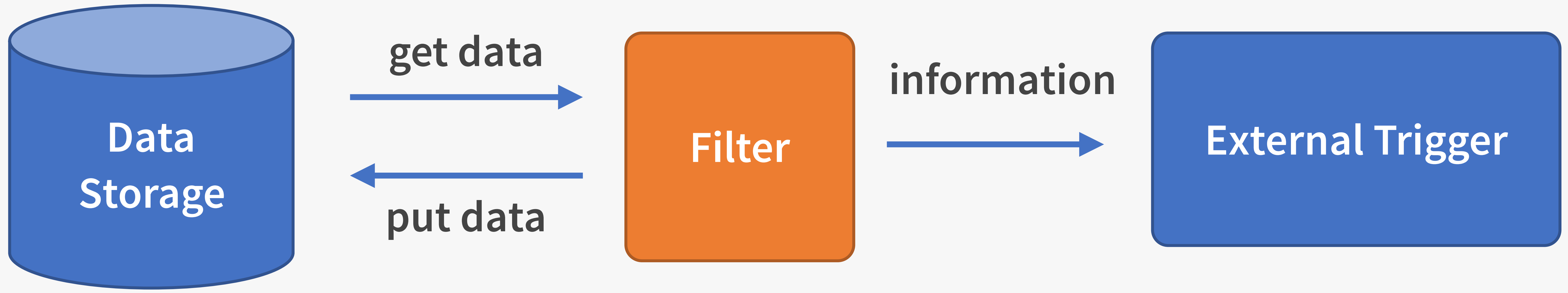


DUNE data flow

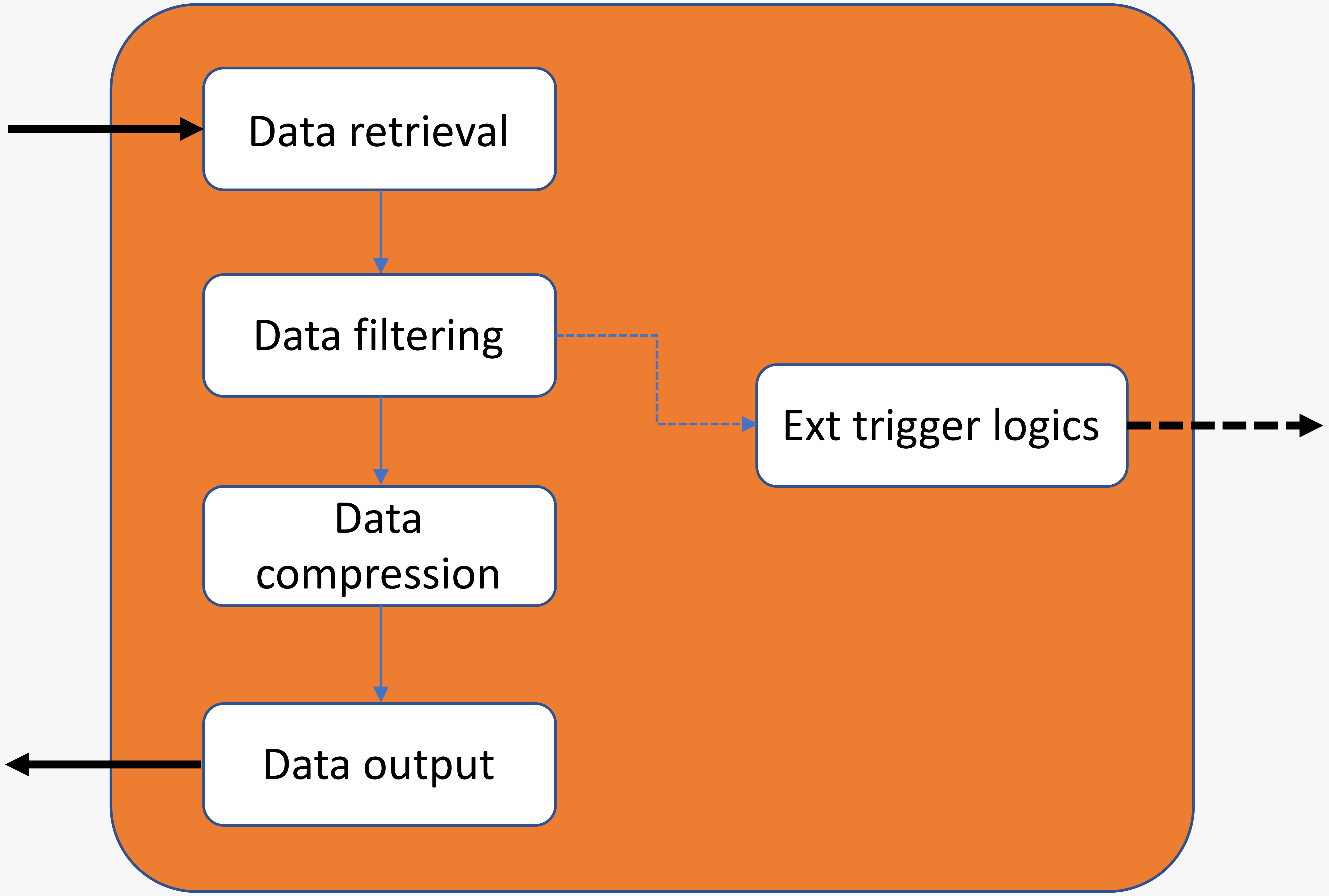


Filter - data flow interfaces

- Several similar elements (~10 / module)



Filter - functional blocks



Predesign prototyping studies

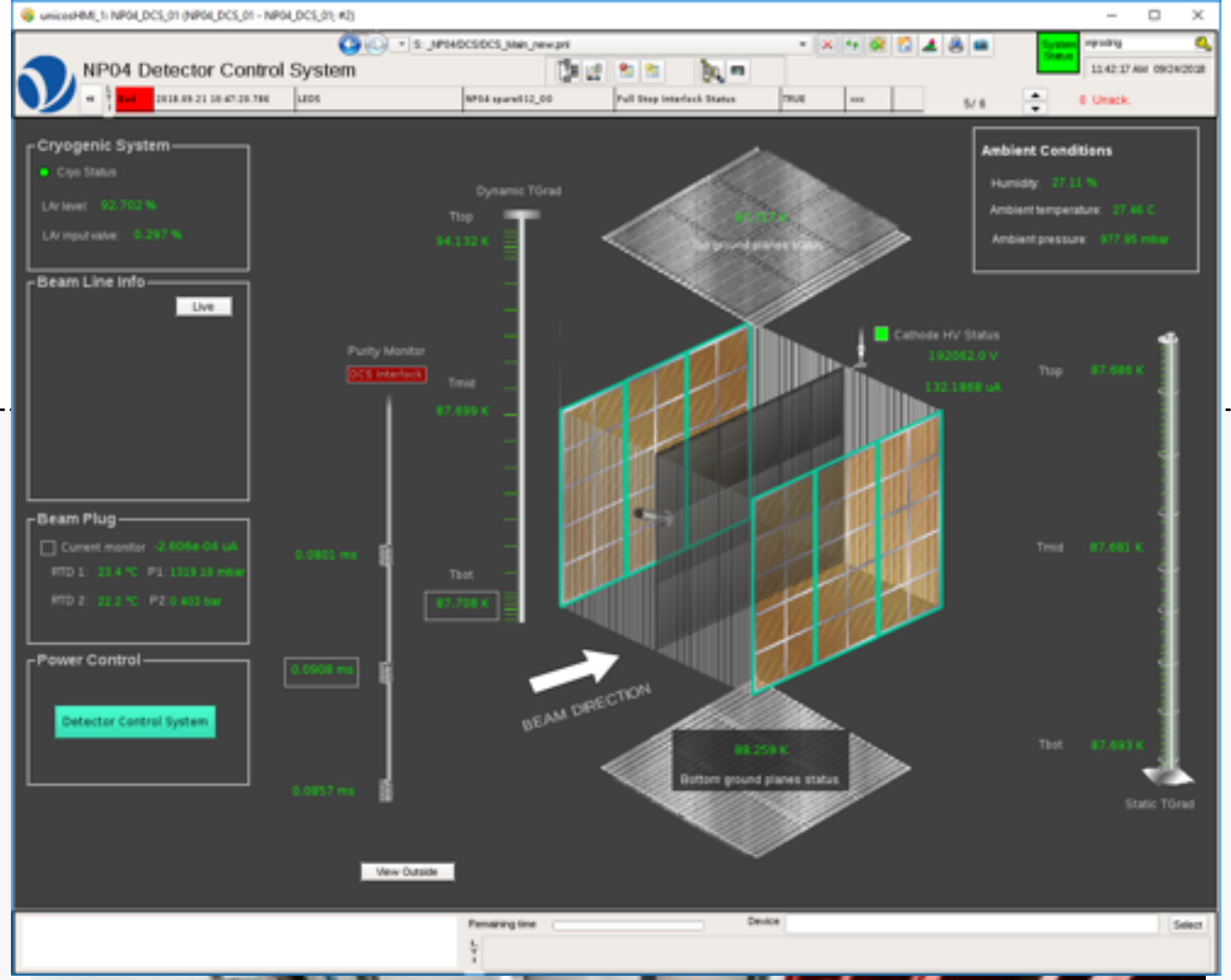
ProtoDUNE Single Phase

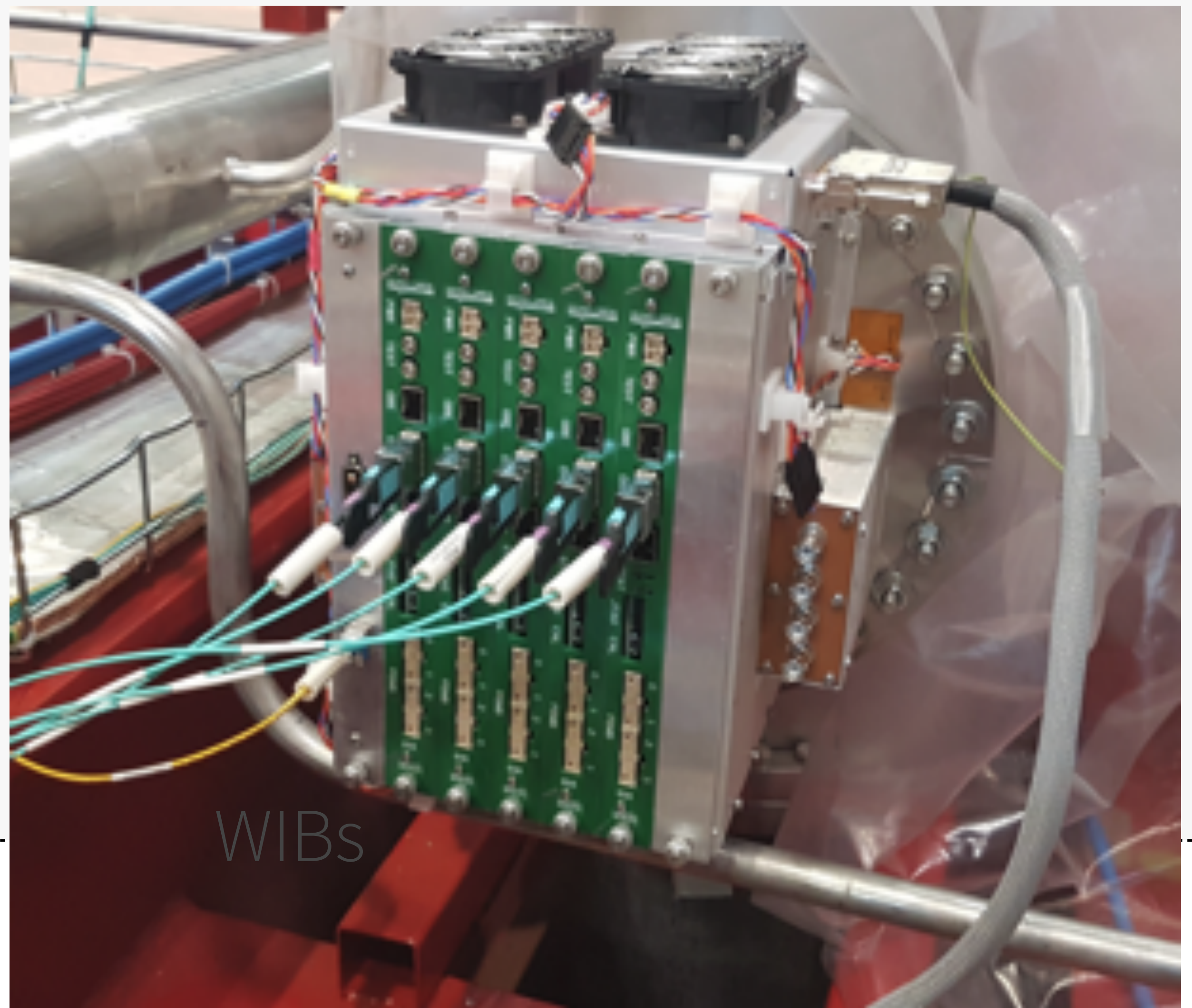
Largest monolithic single phase LArTPC detector and test beam built to date:

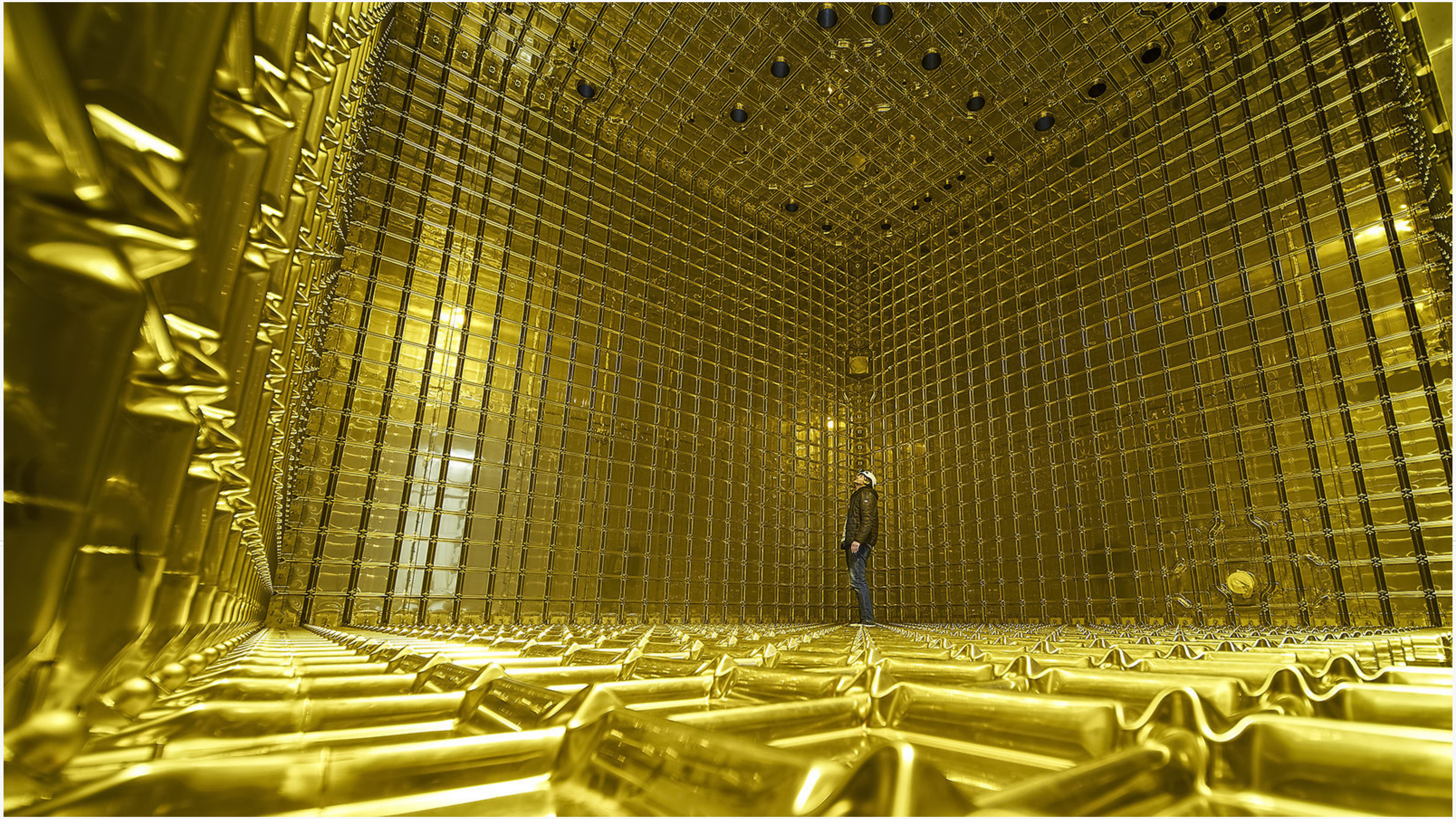
- Goal is to **validate detector design, construction and data acquisition** solutions for DUNE’s Single Phase Far Detectors
  - 10x10x10 **LArTPC**
  - 800 tonnes of LAr
  - Located **on surface** → **external trigger needed**
- Extreme schedule:
  - Project launch: Q1 2016
  - Data taking with beam: **Q4 2018**
- DAQ approach:
  use ready-to-use solutions
  - minimise development time

Data flow and volume

- **LArTPC** → ionisation tracks are collected by the wires of the **Anode Plane Assemblies (APAs)**
- **Cold electronics** in the detector **digitise** signals recorded by wires at **2 MHz**
- **Warm interface boards (WIBs)** then group the resulting channels into **frames**, each of which consists of a single **500 ns time slice of the grouped channels (128 or 256)**
- Output via **optical links** to DAQ:
  - 2x 9.6 Gb/s or 4x 4.8 Gb/s supported, depending on readout solution
  - Continuous **timestamped data frame** streams
- Each APA (2560 channels) is read out by 5x WIBs for a **total payload of about 74 Gb/s**



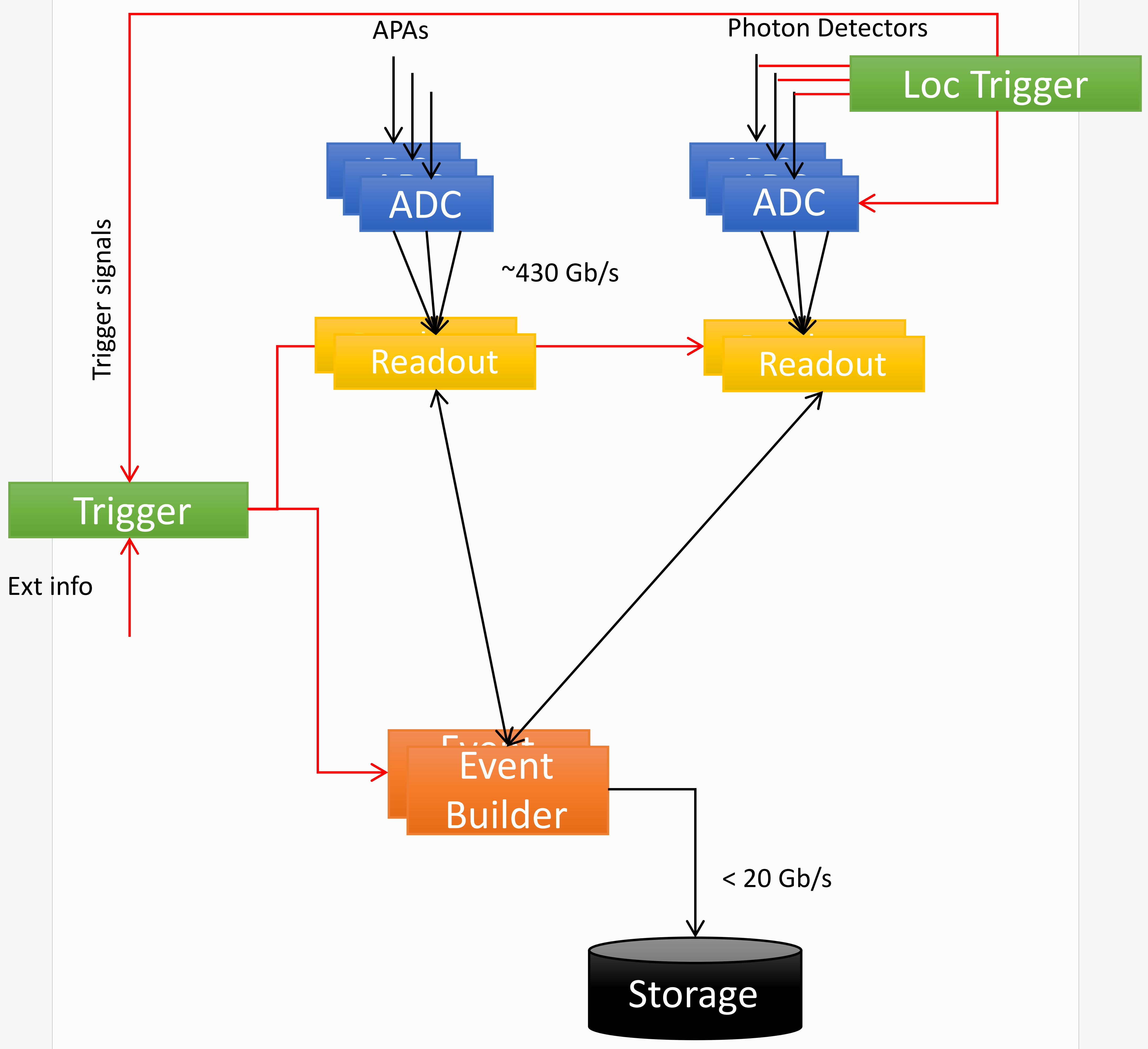




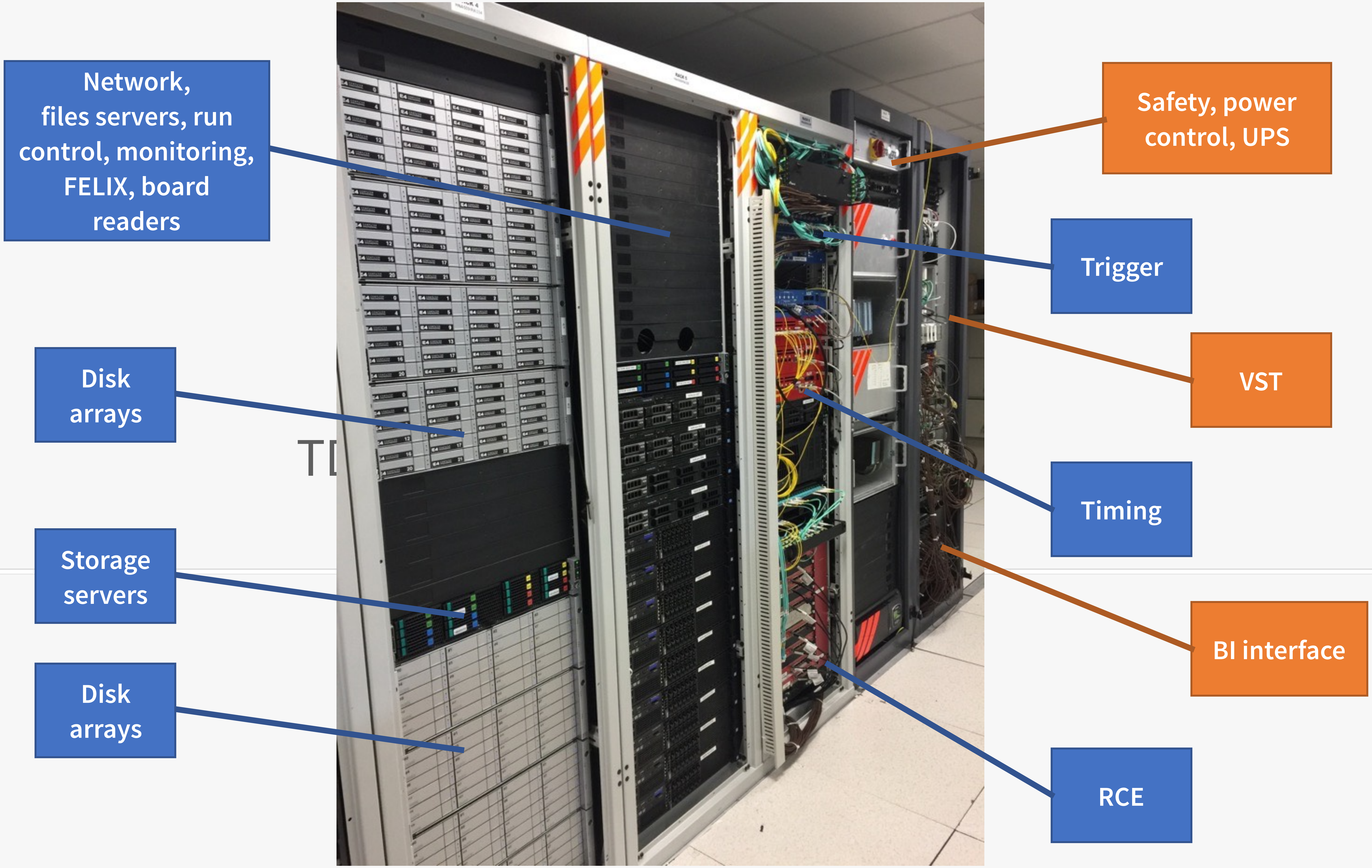
By M. Brice, CERN


20MVL

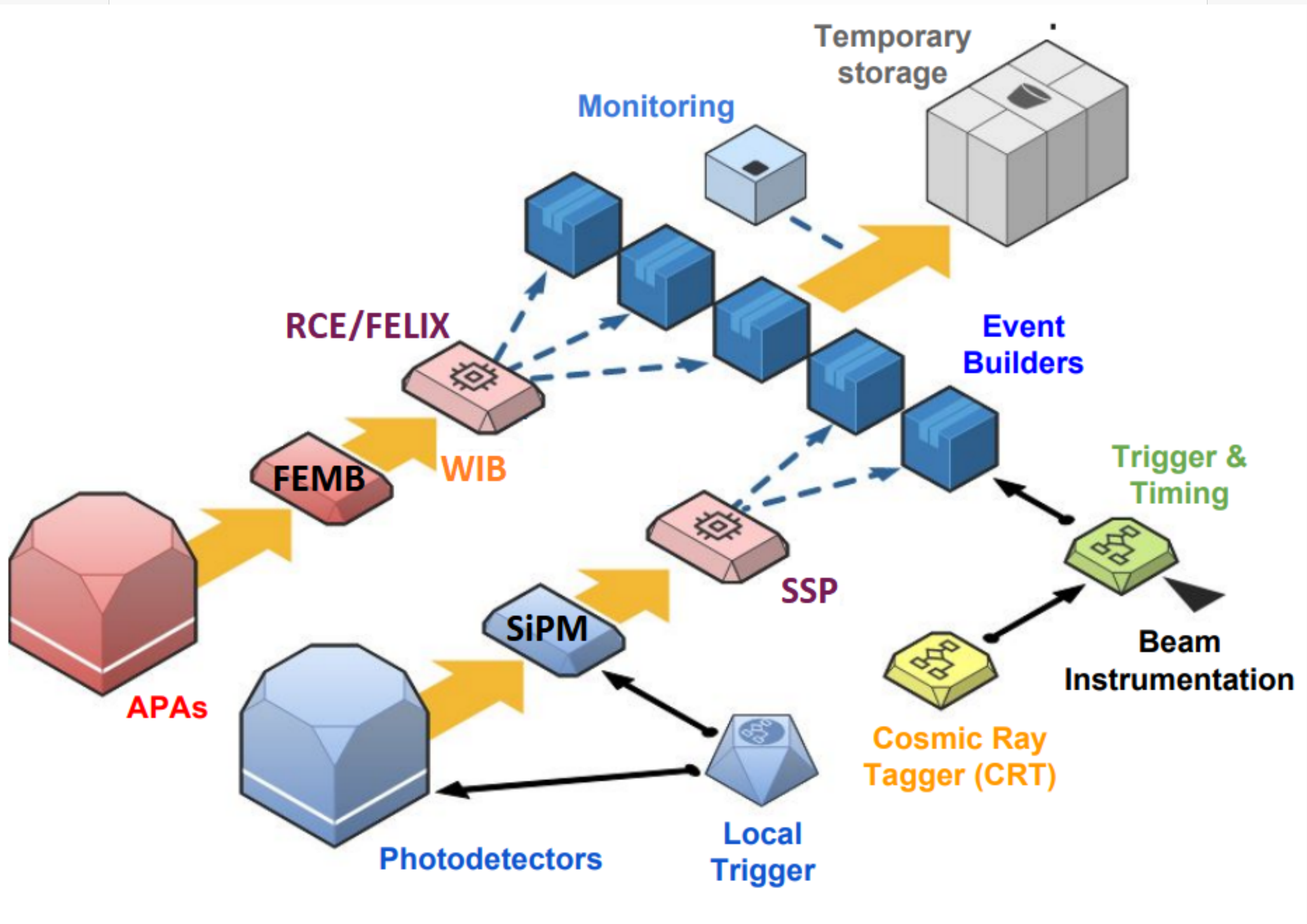
ProtoDUNE SP TDAQ



- Readout with large buffers to allow exploiting the spill structure
- Trigger logics implemented in a custom board
  - Inputs from beam instrumentation, muon tagger, photon detectors
- Data compression to reduce storage and network needs
- An event is a 3 ms window of all data contained in the readout corresponding to the timestamp of a trigger


Network, files servers, run control, monitoring, FELIX, board readers
Disk arrays
Storage servers
Disk arrays
Safety, power control, UPS
Trigger
VST
Timing
BI interface
RCE

ProtoDUNE SP TDAQ



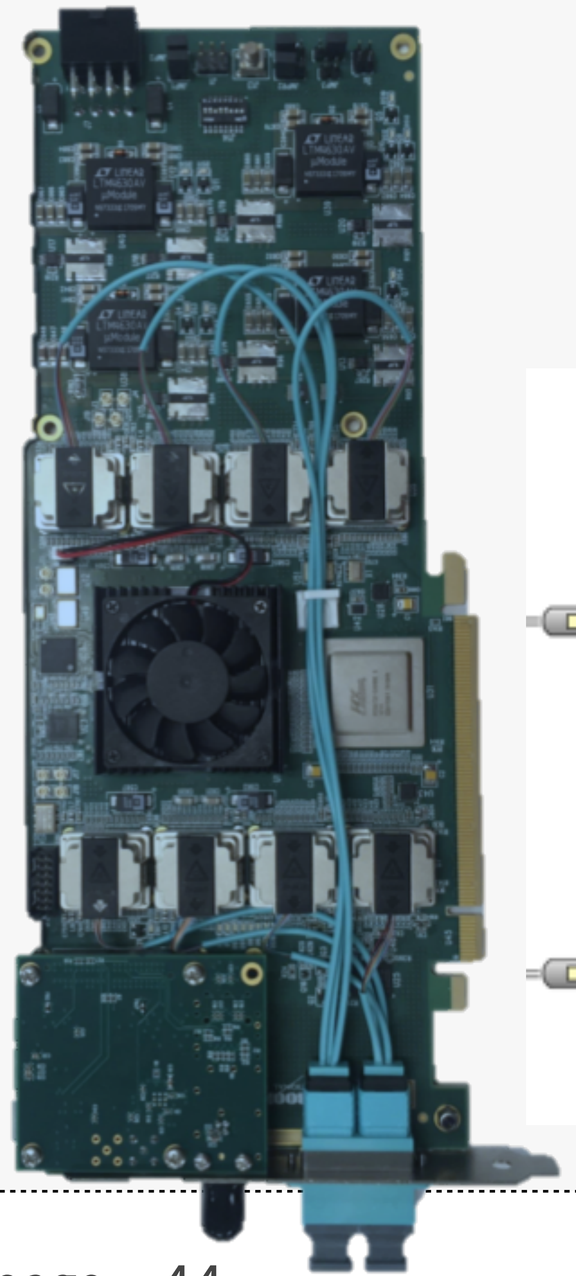
Not a “little” DUNE TDAQ but an excellent test bench for TDAQ technologies for DUNE

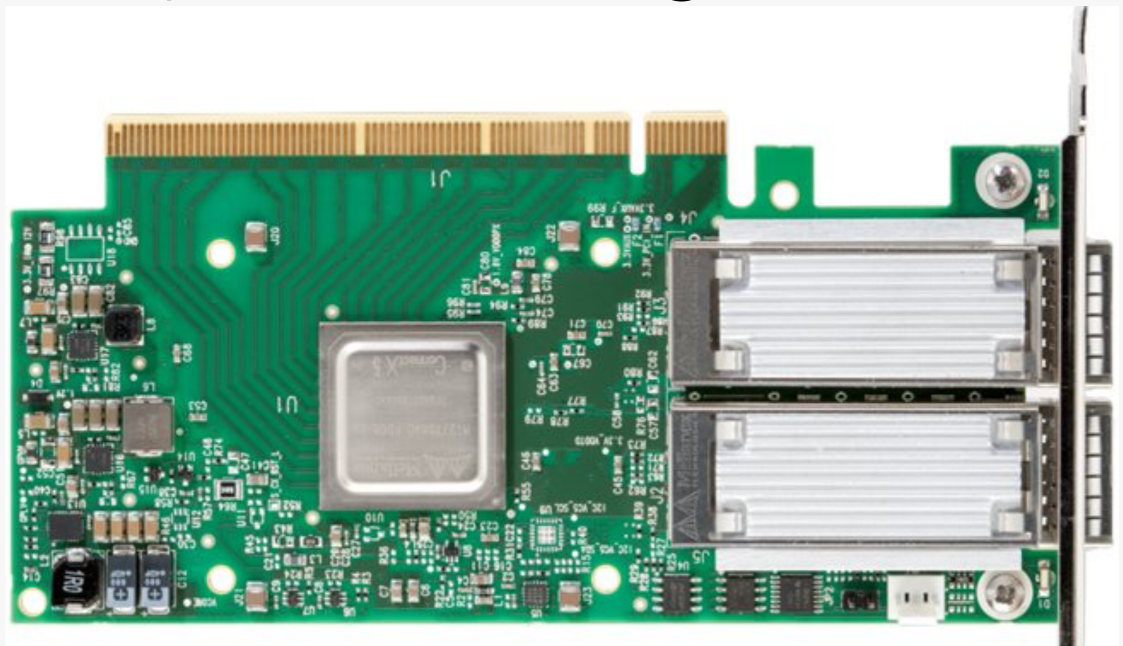
- Timing system prototype
  - Global time via White Rabbit
- Readout system prototypes
  - Buffering in FPGA or server
- Compression scenarios
  - On FPGA or SW
- Hit finding
  - SW or FW
- Buffering
  - RAM + SSD or
    new approaches such as Intel Optane Memory+QLC 3D NAND storage

An example: FELIX based readout

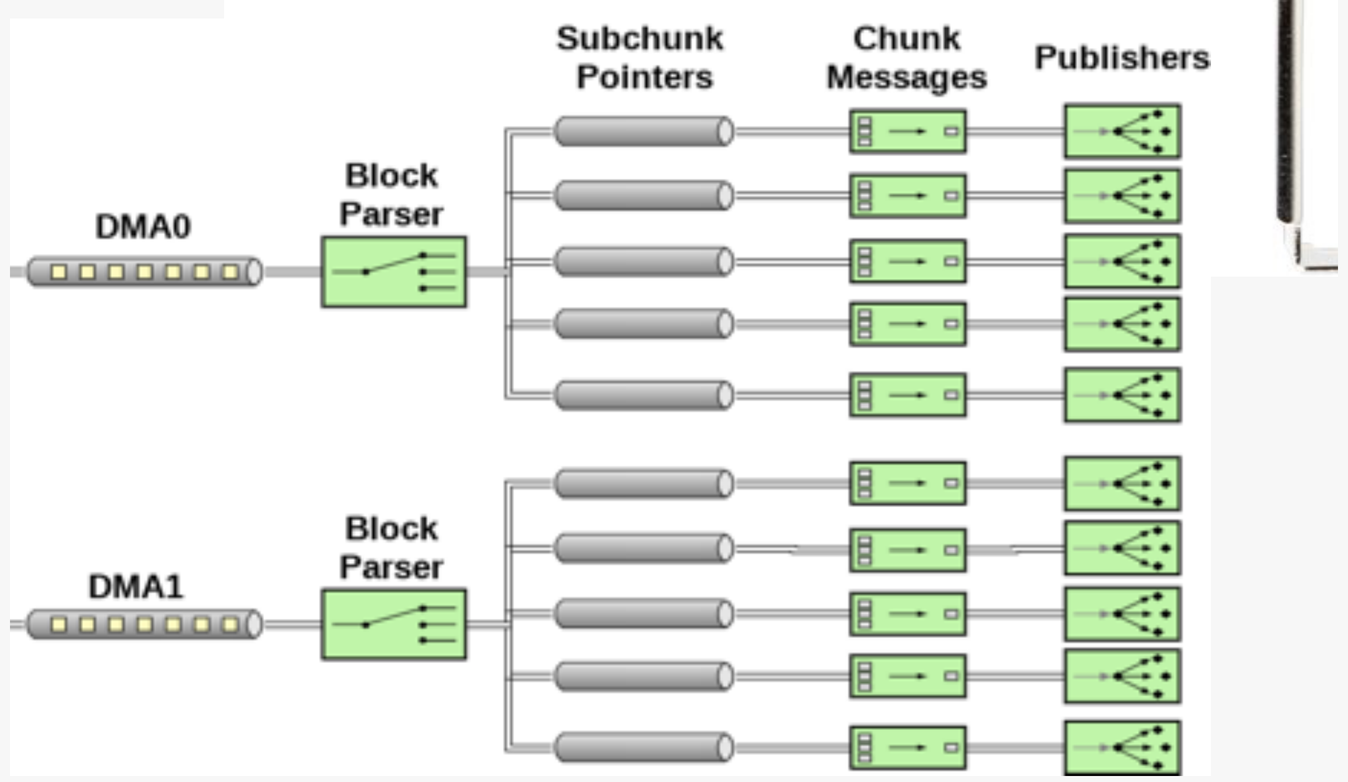
- P-to-p link throughput ✓
- 10 links -> host memory over 1 FELIX card ✓
  - Need to switch to PCIe4 or greater to do 20 links
- HW aided data compression using Intel QAT technology ✓

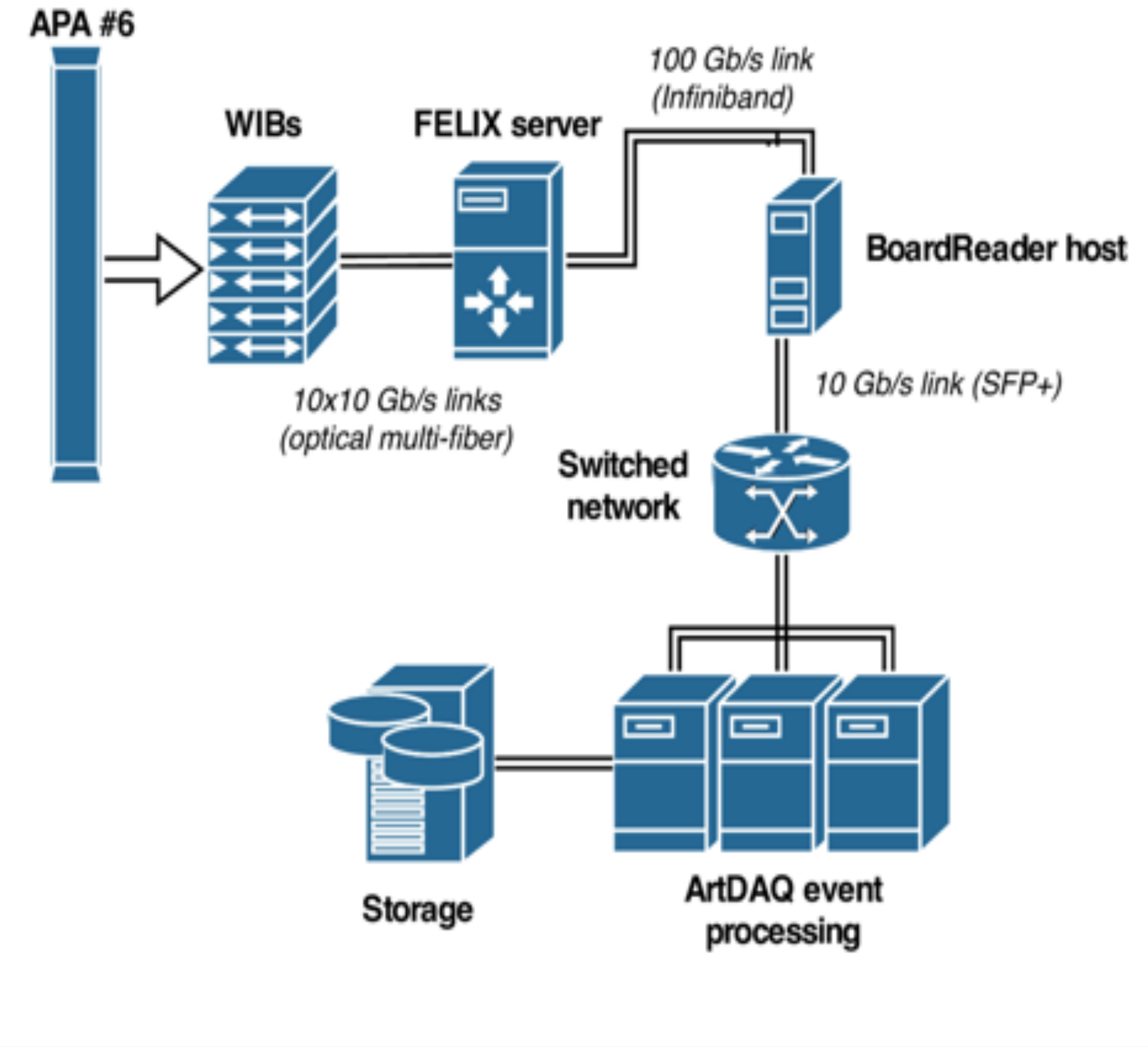
FLX-712





Mellanox ConnectX-5





Full I/O over InfiniBand for ProtoDUNE ✓

- Need much less network I/O for DUNE
- Need longer and high throughput storage (but new technologies go in the right direction)
  - Joint R&D with ATLAS, CMS, DUNE and Intel on DAQ DB

SW triggering from FELIX readout

- Get the complete stream of raw data
- Reformat WIB frames to
  - Expand 12 bit ADCs into 16 bits
  - Reorder wires in order to select only collection plane
- Identify each time a wire has a "hit"
- Combine information of hits in order to form track candidates
- Implement a sw based trigger logics

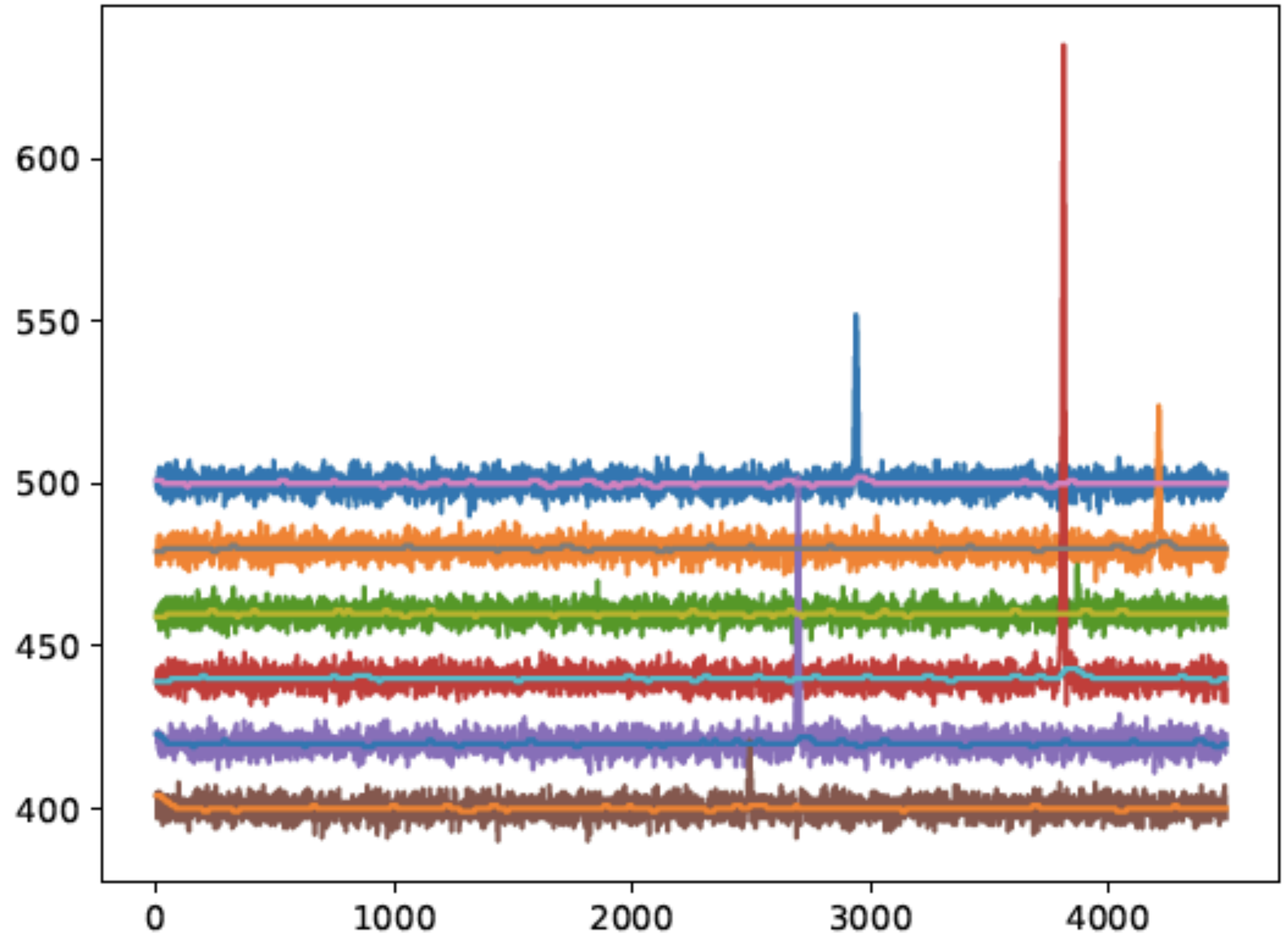
This work is ongoing now! (next few slides from P. Rodrigues)

Data Reordering

<table>
<tr><td>31</td><td>30</td><td>29</td><td>28</td><td>27</td><td>26</td><td>25</td><td>24</td><td>23</td><td>22</td><td>21</td><td>20</td><td>19</td><td>18</td><td>17</td><td>16</td><td>15</td><td>14</td><td>13</td><td>12</td><td>11</td><td>10</td><td>9</td><td>8</td><td>7</td><td>6</td><td>5</td><td>4</td><td>3</td><td>2</td><td>1</td><td>0</td></tr>
<tr><td colspan="8">0x00</td><td colspan="8">0x00</td><td colspan="8">0x00</td><td colspan="8">SOF</td></tr>
<tr><td colspan="8">Reserved (8)</td><td colspan="3">SlotNo</td><td colspan="5">CrateNo</td><td colspan="3">FiberNo</td><td colspan="5">Version = 0x1</td><td colspan="8">0x0</td></tr>
<tr><td colspan="16">WIB Errors</td><td colspan="14">Reserved (14)</td><td>OOS</td><td>MM</td></tr>
<tr><td colspan="32">Timestamp [31:0]</td></tr>
<tr><td>Z</td><td colspan="15">Timestamp [62:48] or WIB counter [3]</td><td colspan="16">Timestamp [47:32]</td></tr>
<tr><td colspan="8">ChkSm B [7:0]</td><td colspan="8">ChkSm A [7:0]</td><td colspan="8">Reserved (8)</td><td colspan="4">Stream 2 ERR</td><td colspan="4">Stream 1 ERR</td></tr>
<tr><td colspan="16">COLDDATA Convert Count</td><td colspan="8">ChkSm B [15:8]</td><td colspan="8">ChkSm A [15:8]</td></tr>
<tr><td colspan="16">Reserved</td><td colspan="16">Error Register</td></tr>
<tr><td colspan="4">HDR8</td><td colspan="4">HDR6</td><td colspan="4">HDR7</td><td colspan="4">HDR5</td><td colspan="4">HDR4</td><td colspan="4">HDR2</td><td colspan="4">HDR3</td><td colspan="4">HDR1</td></tr>
<tr><td colspan="4">ADC2 CH2[3:0]</td><td colspan="4">ADC2 CH1[11:8]</td><td colspan="4">ADC1 CH2[3:0]</td><td colspan="4">ADC1 CH1[11:8]</td><td colspan="8">ADC2 CH1[7:0]</td><td colspan="8">ADC1 CH1[7:0]</td></tr>
<tr><td colspan="8">ADC2 CH3[7:0]</td><td colspan="8">ADC1 CH3[7:0]</td><td colspan="8">ADC2 CH2[11:4]</td><td colspan="8">ADC1 CH2[11:4]</td></tr>
<tr><td colspan="8">ADC2 CH4[11:4]</td><td colspan="8">ADC1 CH4[11:4]</td><td colspan="4">ADC2 CH4[3:0]</td><td colspan="4">ADC2 CH3[11:8]</td><td colspan="4">ADC1 CH4[3:0]</td><td colspan="4">ADC1 CH3[11:8]</td></tr>
<tr><td colspan="4">ADC2 CH6[3:0]</td><td colspan="4">ADC2 CH5[11:8]</td><td colspan="4">ADC1 CH6[3:0]</td><td colspan="4">ADC1 CH5[11:8]</td><td colspan="8">ADC2 CH5[7:0]</td><td colspan="8">ADC1 CH5[7:0]</td></tr>
<tr><td colspan="8">ADC2 CH7[7:0]</td><td colspan="8">ADC1 CH7[7:0]</td><td colspan="8">ADC2 CH6[11:4]</td><td colspan="8">ADC1 CH6[11:4]</td></tr>
<tr><td colspan="8">ADC2 CH8[11:4]</td><td colspan="8">ADC1 CH8[11:4]</td><td colspan="4">ADC2 CH8[3:0]</td><td colspan="4">ADC2 CH7[11:8]</td><td colspan="4">ADC1 CH8[3:0]</td><td colspan="4">ADC1 CH7[11:8]</td></tr>
</table>

- 1 WIB frame = 464 B => 256 ADCs + headers @ 2 MHz
- Unpack collection channels with AVX2 code
- Spoiler: this appears to be the biggest CPU consumer

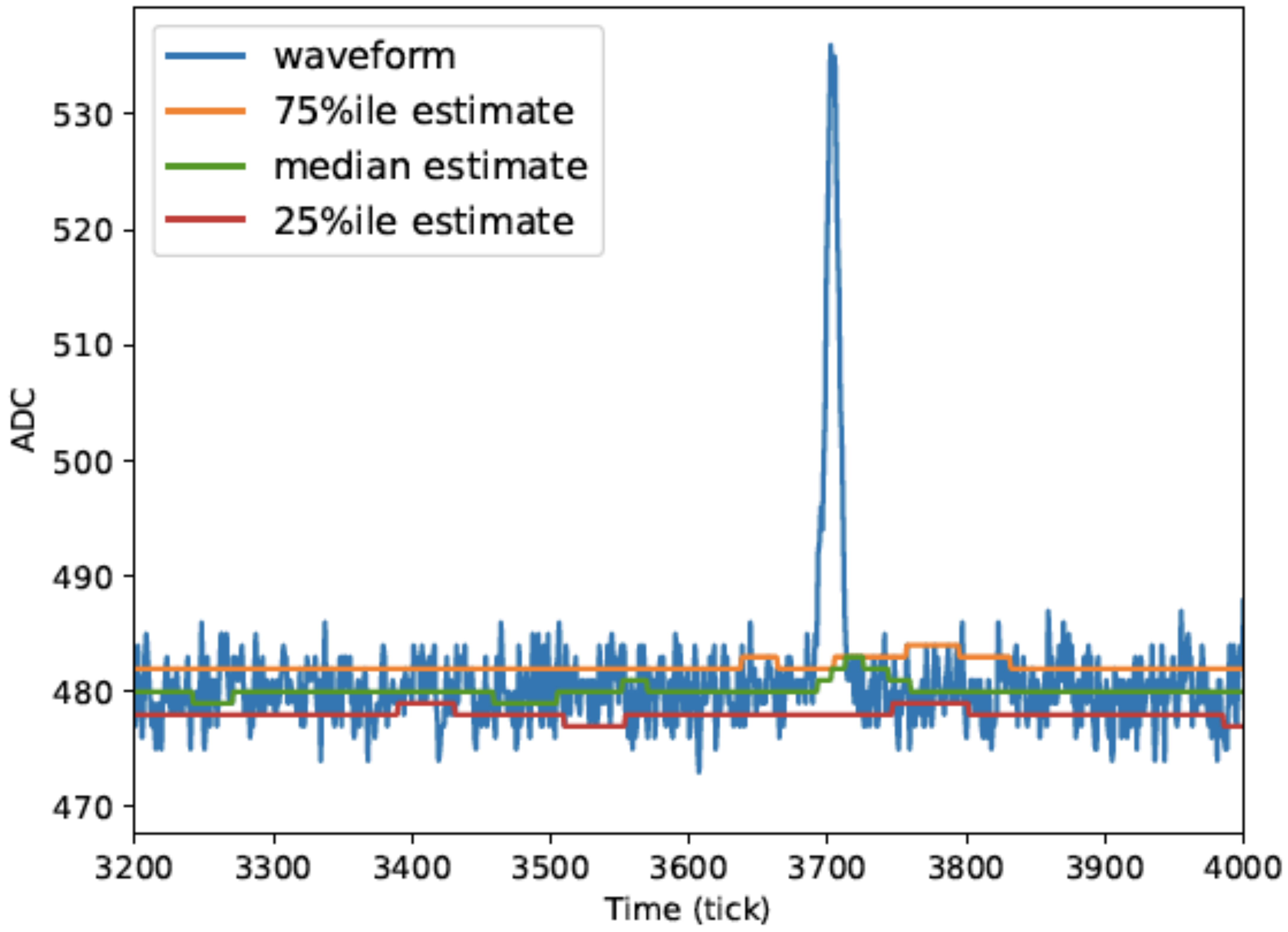
Pedestal Finding



1. Start with an accumulator=0, an estimate of the median, and read the next sample
2. If sample > median, increase accumulator by 1
3. If sample < median, decrease accumulator by 1
4. If accumulator $= X$, increase median by 1, reset accumulator to 0
5. If accumulator $= -X$, decrease median by 1, reset accumulator to 0

Slide from P. Rodrigues

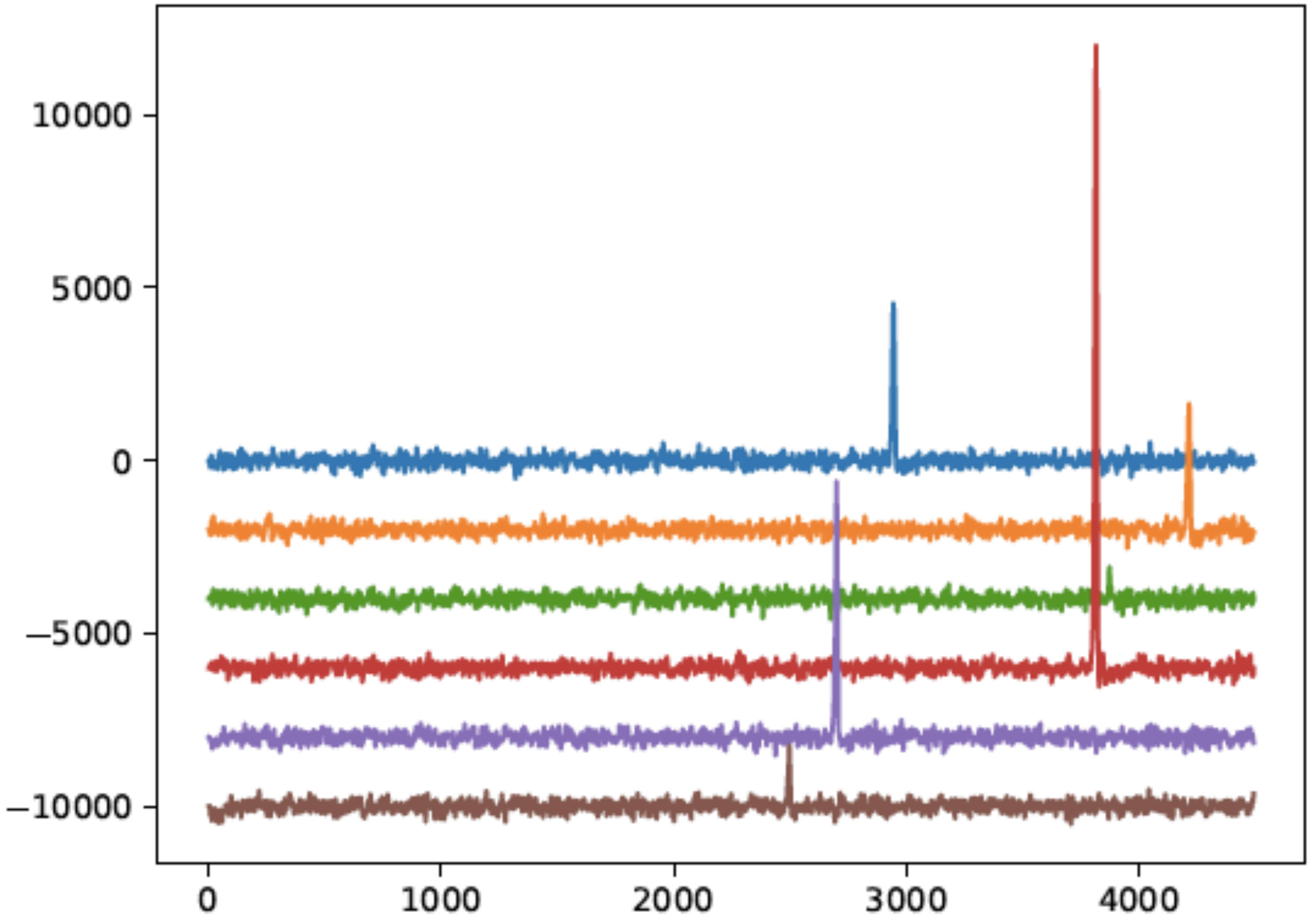
RMS estimate



- If next sample above median, use for frugal streaming of 75%ile
- If next sample below median, use for frugal streaming of 25%ile
- Call the difference "σ"

Slide from P. Rodrigues

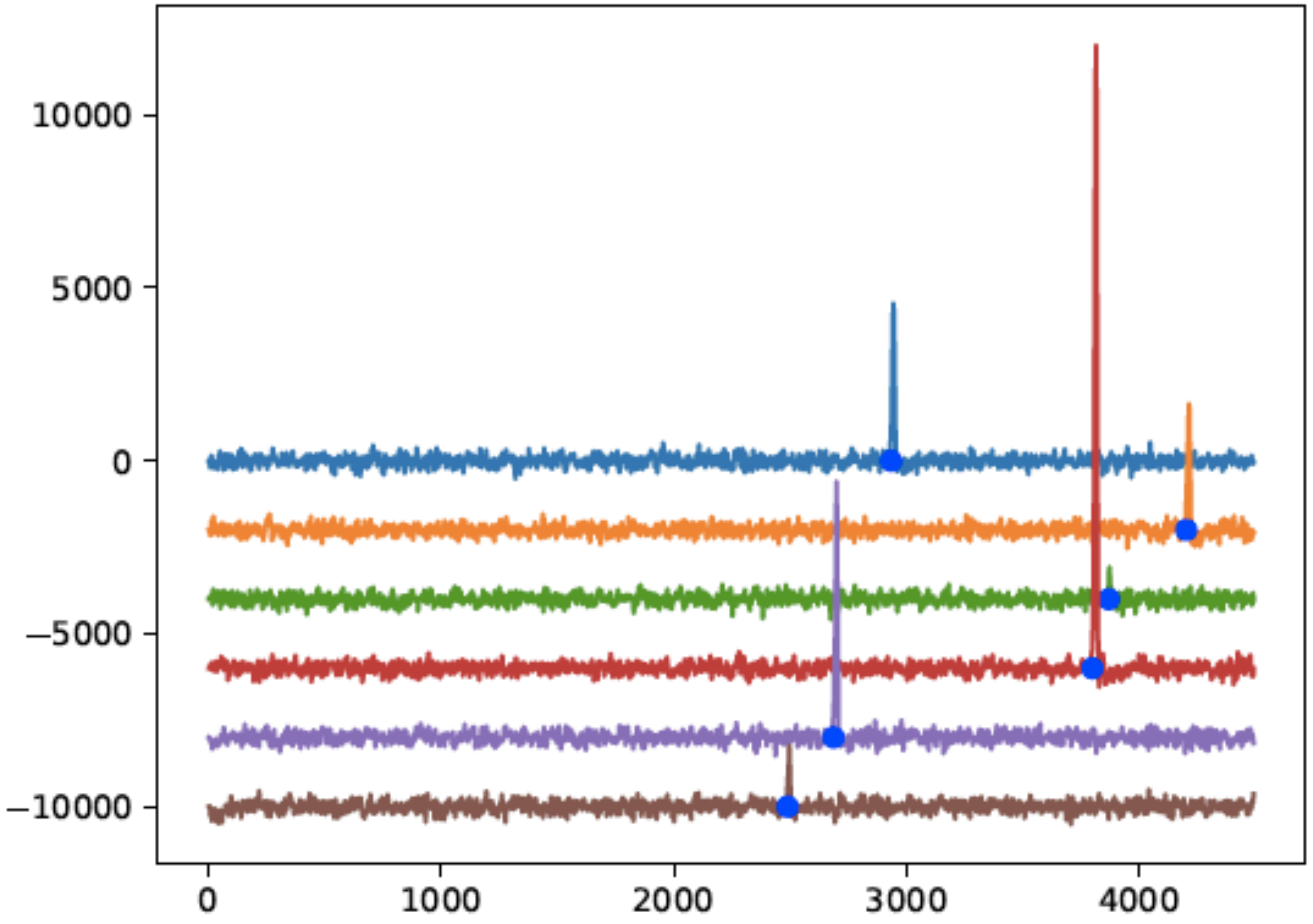
Noise Filtering



- I used a simple FIR lowpass filter
- Hardcoded filter size (7 taps), unrolled inner loop

Slide from P. Rodrigues

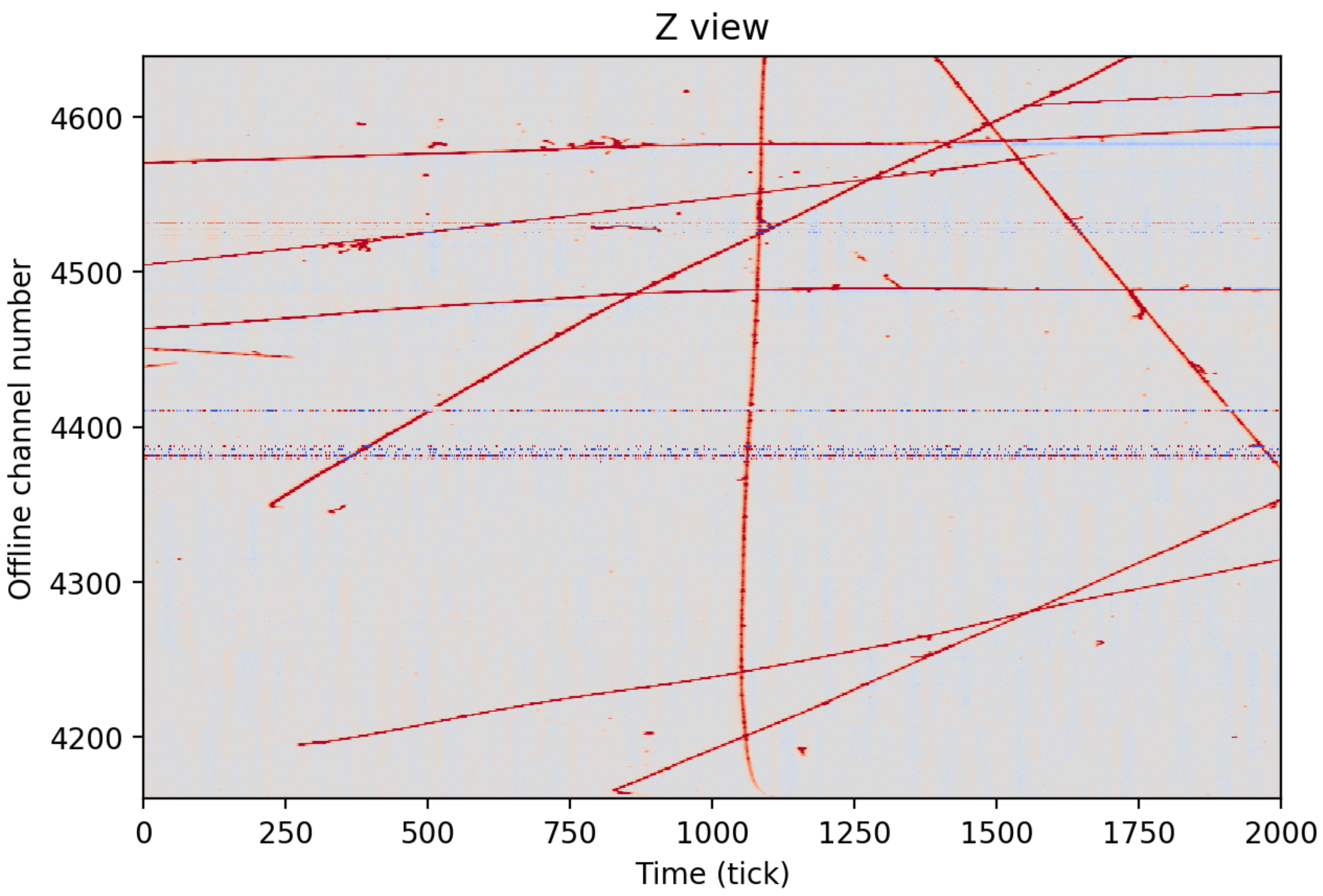
Hit Finding



- Algorithm: first sample over 5σ starts a hit. Integrate time and charge until fall below threshold again

Slide from P. Rodrigues

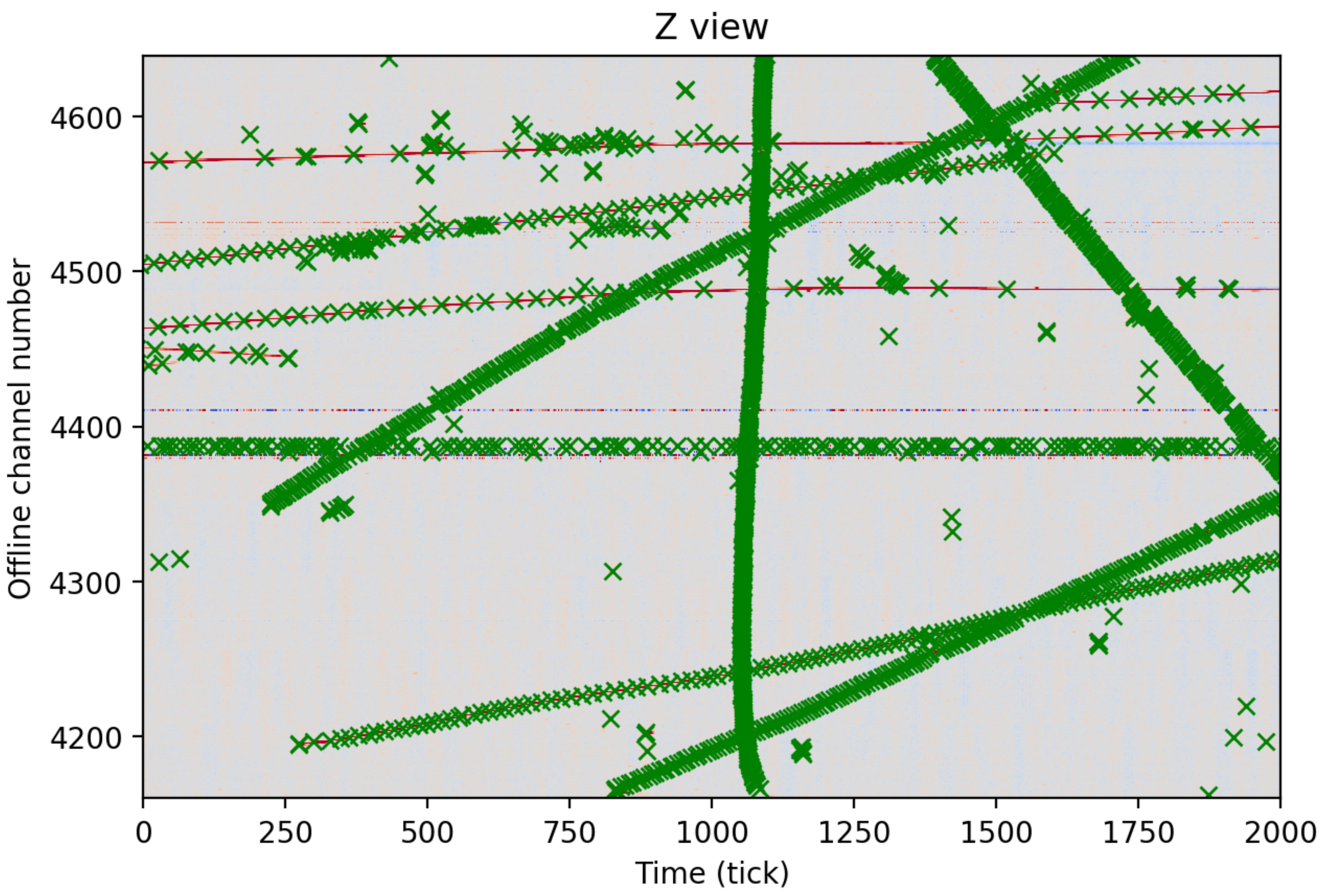
Hit finding at work



- Hits found continuously; read out from buffer in response to trigger

Slide from P. Rodrigues

Hit finding at work



- Hits found continuously; read out from buffer in response to trigger

Slide from P. Rodrigues

Clustering and Module Level Trigger Prototype

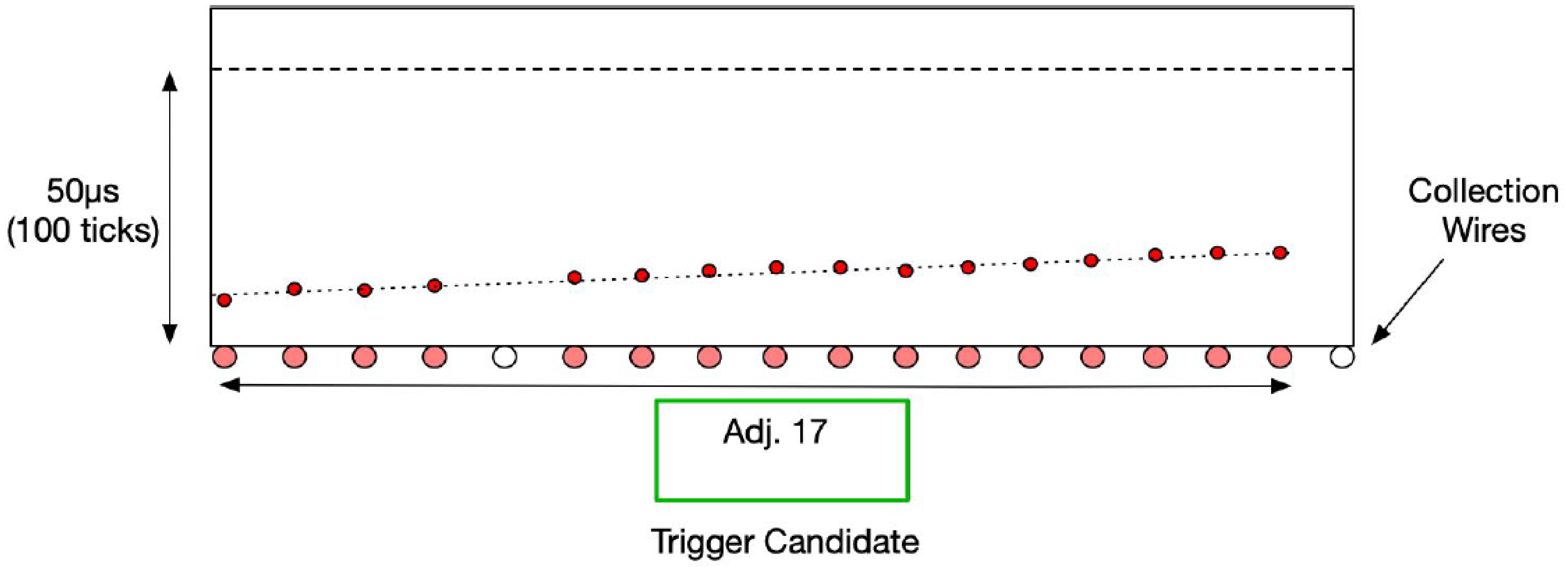
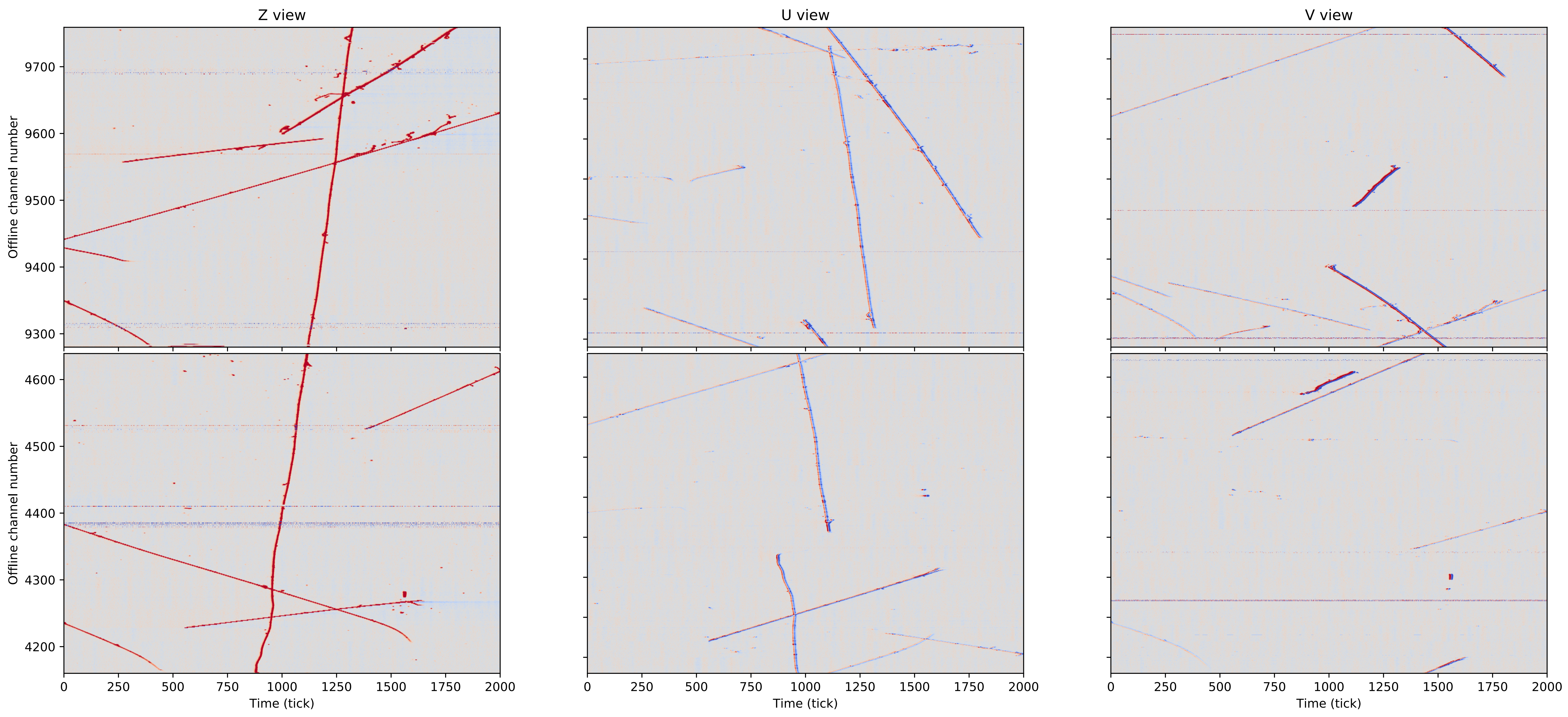


Figure: Jon Sensenig

- Take 50 μs windows, find groups of hits contiguous in channel (with up to 4-wire gap)
- Generate TC if channel range of largest group is 100 or more
- Stitch together TCs from individual APAs
- Make a sliding window of one drift time, look for consistent slope (Δ(time)/Δ(channel)) between TCs
- Trigger if stitched TCs result in 350 hits or more in each APA

Slide from P. Rodrigues

SW-triggered Horizontal Muon



- Read out 1000 ticks before and after the trigger time

Slide from P. Rodrigues

SW triggering next steps

Carry out data reformatting in FPGA

- Measure benefits

Move the complete hit finding into FPGA

- Measure benefits and assess any drawbacks

DUNE DAQ Control & Monitoring

What is special about DUNE?

Complexity and size of a large collider experiment

- Uptime <<30%

Uptime requirement of a “rare events” experiment (gravitational waves, supernovae detectors, double beta neutrino decays, etc

- Uptime “100%”

Accessibility of installation quite limited

The combination of those three doesn’t work well together. . .

The whole system must be conceived and setup in a different way

- Redundancy and fail over mechanisms
- Automated anomaly detection and recovery
- Remote monitoring and control

The control and monitoring system will have a predominant role for the success of the DUNE TDAQ (i.e. the experiment)

Few guidelines

Components must be as loosely coupled as possible

- Allow for tolerance to and recovery of local problems without affecting the data taking

Single points of failure (module level trigger, EB orchestrator) must have a running backup in standby mode

Running conditions must be as stable as possible

- Forget about stopping runs and regular full reconfigurations of the system

Assessment of data quality must be immediate and continuous

- Automated correction for bad data

System administration (computers, networks, storage) and repair must happen on an active system,

- i.e. be staggered and non intrusive

All tools need to be thought from the start for remote operators

- Heavily rely on web, but still ensure security and safety

Design of the Control and Monitoring for DUNE

Do we know how to do it?

- Not really…
- Ideas are mainly on paper and need to be tried out
- System and software engineering skills are essential to get this right

Design of the Control and Monitoring for DUNE

Do we know how to do it?

- Not really…
- Ideas are mainly on paper and need to be tried out
- System and software engineering skills are essential to get this right

How would you go about designing and implementing such a system?

Summary and Outlook

DUNE is a new giant experiment scheduled to start taking data in 2025

The TDAQ system is being designed now

- Challenging readout performance
- Very challenging operational requirements

Predesign prototyping allows us to identify suitable technologies and validate ideas

- Advancing well on the main data flow path
- Still embryonic stage for the control and configuration

If ISOTDAQ awoke your desire of becoming a TDAQ expert

- DUNE is surely an experiment where you will be able to challenge your skills!