Near Detector Data Model and Volume Estimate

Tom Junk

DUNE ND Software Integration Meeting

August 30, 2019

Fermilab DUNE

Input Data Requirements

- All three detectors read out during beam spills – shared triggers.

- Cosmic-ray data need to be collected. Cosmic rays that pass from one detector to another are especially valuable.

- Calibration data – lasers, pulsers, photon flashers, radioactive sources, etc.

- Beam-off random triggers need to be collected

- Special requests for BSM-motivated triggers ought to be honored when possible.

- Full-waveform diagnostic data subsets.

- Data need to be divisible into pieces < ~100 MB for offline processing

List is likely not complete and may contain extras

🔷 Fermilab   DUNE

Channel Counts

- ArgonCube:

  - dimensions: 7m x 5m (x3m tall).

  - Pixels are 3mm x 3mm

  - I get from this 4M TPC pixels   (cf. FD: 384,000 per SP module)

  - a small number of photon detector channels (ArcLight)

- MPD:

  - 0.7M pads in ALICE TPC chambers + hole-fillers

  - 3M calorimeter channels (E. Brianne)

  - Small number of photon detector channels + muon catcher (?)

- 3DST-S

  - Default size in the TDR Exec Summary is 2.4m x 2.4m x 2.0 m, with 1 cm cubes. Assume double-ended readout: 2x(240x240+240x200+240x200)=307,200

  - Assume 22K ECAL channels

  - TPC: T2K ND has 36,864 Micromegas pads per TPC. x4 for DUNE?

🔶 Fermilab DUNE

Data Volume from MPD

- It's the one I know best and have some MC and points of comparison for

- Monte Carlo 60-interaction spills, no noise.

  - 130,000 TPC hits. Zero-suppressed waveforms provided by electronics

  - MC event size after ROOT compression: 800 kB/event TPC + 200 kB/event in the calorimeter. Eldwan tells me 1 MB/event for the calorimeter is reasonable.

  - Wouldn't be surprised if we were off by x5 due to noise, crosstalk, and desire to sample longer sections of waveforms. Call it 5 MB/event

  - ALICE: David Rohr at the TPC workshop July 12, 2019:
    https://indico.cern.ch/event/827540/
    60 MB/trigger raw data links from detector, LHC heavy-ion collisions
    But ALICE trims and compresses it down to 5 MB/trigger

🔬 **Fermilab**   DUNE

Near Detector Issues

Each spill will contain tens of neutrino interactions
- MPD will have 60 events in the ECAL per spill plus 0.5 events in the gas
- ArgonCube will have ~tens of events per spill
- 3DST-S also similar

- MPD and ArgonCube are slow TPC's. The MPD ECAL is fast (~ns response time). 3DST-S is also quite fast, of order ns for scintillation

- Overlapping events in liquid and gas Ar ND detectors will be overlaid in space and time and have to be unentangled with faster detector components

🔶 **Fermilab** DUNE

Unknowns

- Components of the data volume

  - Some scale with intensity. Zero suppression makes most data like this

  - Some scale with channel count: Noise and calibrations

  - Intensity-scaling data scales inversely with channel sensitive volume/area

- Calibration data volume is our choice.

- Cosmic-ray rate needs to be estimated at the ND (or at least communicated to me...)

- How much full-waveform data do we need?

  - Sub-threshold noise + radiological measurement would be good to have

- Random trigger rate

- Do we need a SNB trigger?

✜ Fermilab DUNE

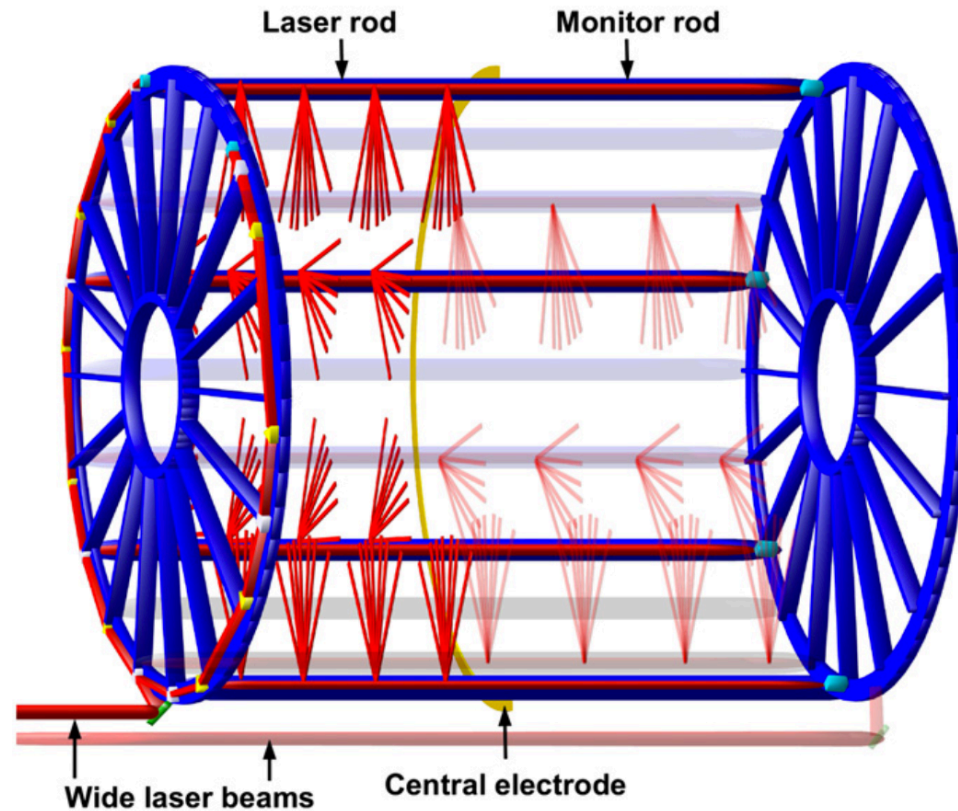# Example: ALICE Laser Calib System

Fig. 45. Schematic 3D view of the TPC and the laser system. Two wide pulsed laser beams enter horizontally at the bottom of the TPC and are guided around the two end-caps by mirrors, prisms and beam splitters before entering the TPC. Bundles of micromirrors in the hollow laser rods intersect the beams and generate a large number of thin rays in the TPC drift volume. The undeflected part of the beams continue through the monitor rods to cameras at the far end. All elements are fixed mechanically, except for the remote controllable entrance mirrors at the bottom.
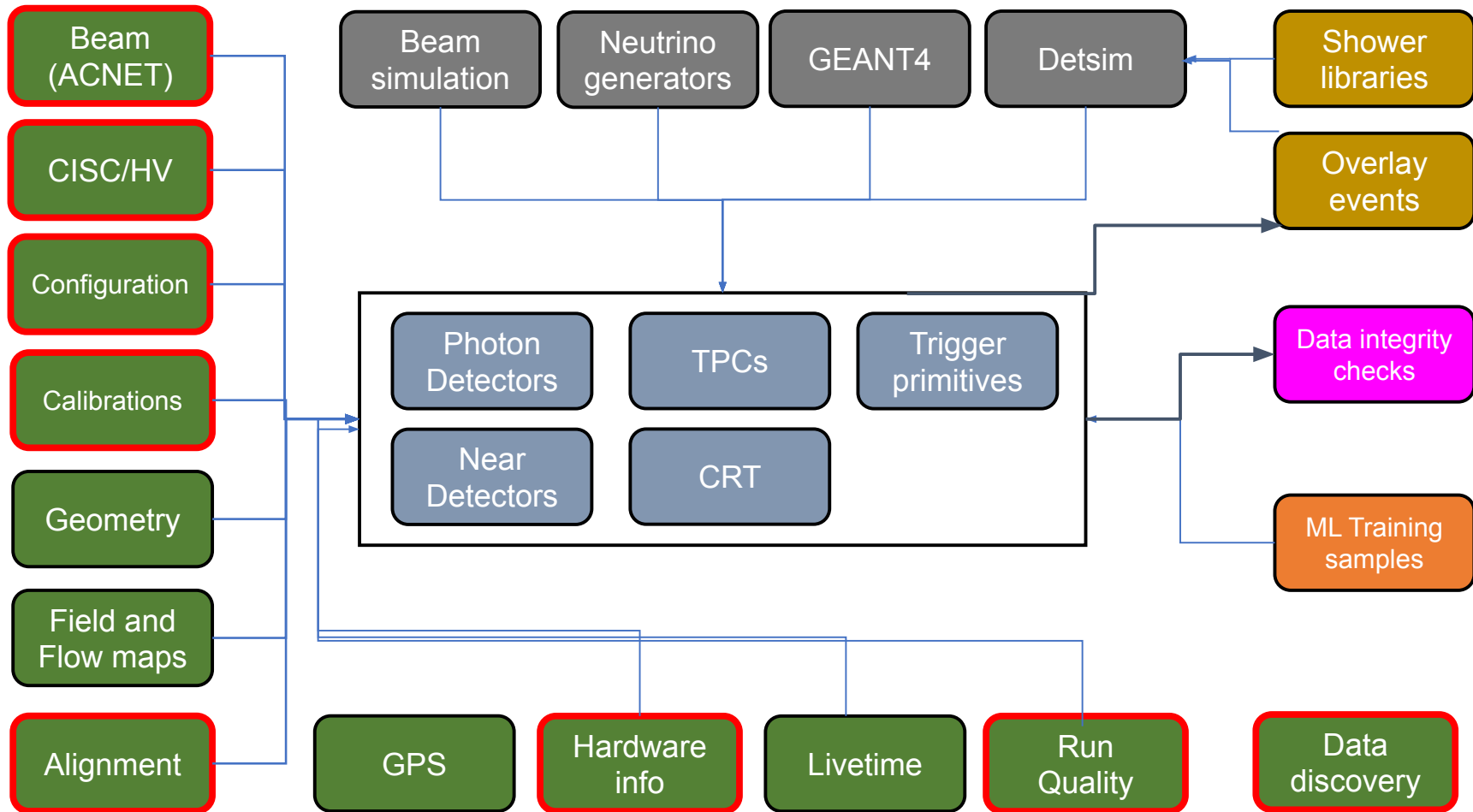
🔵 **Fermilab** DUNE

Data Volumes

- Generous spill data volume estimate: 100 MB/spill (20x CMS event size).

- Good to an order of magnitude.

- Assume 2E7 spills/year, get 2 PB/year.

- "small" compared to FD – 30 PB/year

- Offline data tiers – reconstruction, analysis, several reprocessings will be needed. NOvA: x7. Andrew Norman predicted ProtoDUNE would need 2x raw for offline copies.

Far Detector Data Model

- See
  https://indico.fnal.gov/event/21160/other-view?view=standard

- Little discussion about ND, mostly talking about FD

- Some details:

  - Isolated events (cosmics, beam, atmospherics, nucleon decay...)
    are not too hard – 3.3 GB/event/module. Ideally we'd like to stream
    the raw digits on input to batch jobs.

  - SNB events are about 170 TB each. Definite need to stream and
    stitch.

  - Calibration samples are turning out to be a big load. Laser
    calibrations can run many seconds and pulse the laser very
    frequently.

FD Data External Dependencies

Heidi Schellman

Fermilab DUNE

Row-Wise vs. Column-wise data recording

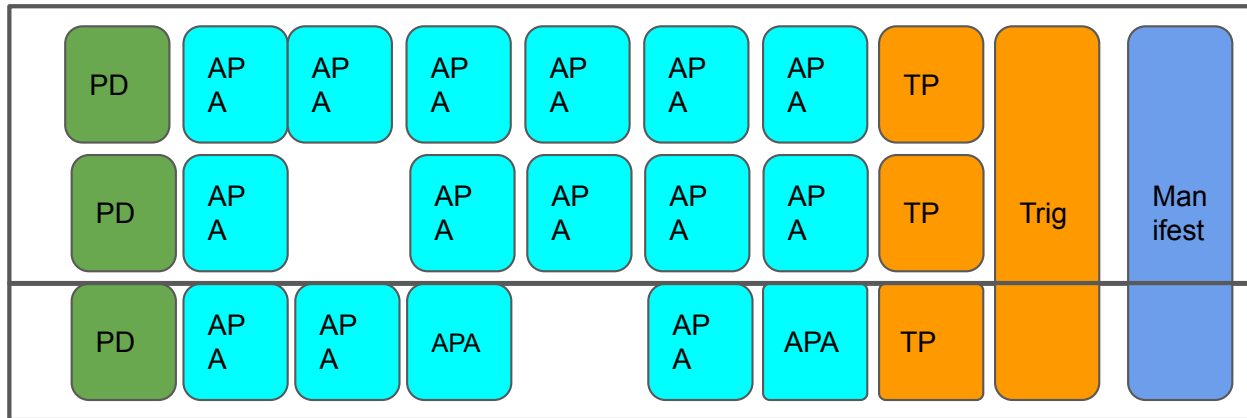Brett and Georgia's slides describe this

Or

Data need to be rearranged (channel by tick) $\rightarrow$ (tick by channel) already

🔷 Fermilab DUNE

Far Detector Data Aggregation

Localized readout aggregate (cosmics/beam)
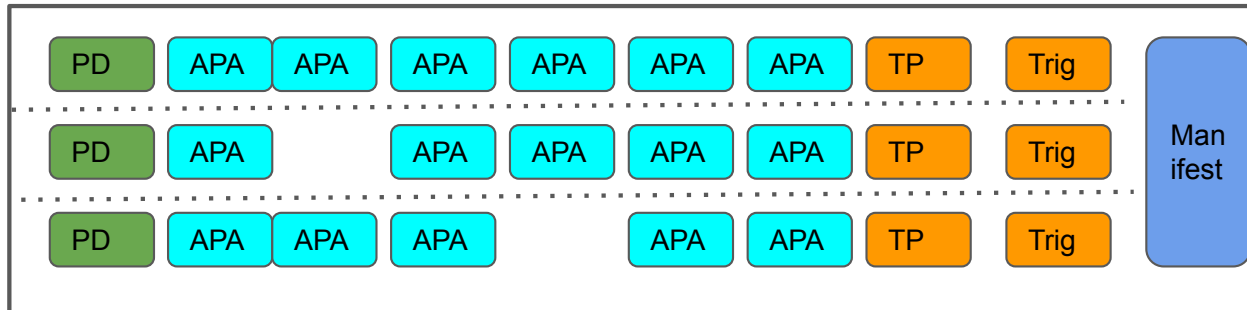
Extended (SNB) readout aggregate

Black outline box represents a file. A file must be less than 10 GB. An aggregate of files can be larger than 10GB, and needs a "manifest" i.e. enough info to prescribe what is included in the aggregate. For localized readouts the minimum number of rows is 1, and rows should be complete event records.

Heidi Schellman

🟦 **Fermilab** DUℕE