Summary of DUNE Data Model W/S @ BNL 2019-08-14/17

Brett Viren

Physics Department

BROOKHAVEN
NATIONAL LABORATORY

Goals and Major Topics

Goals Produce a TDR-like document covering DUNE Data Model.

Major topics:

DAQ design, raw data format, interface to Offline.

Event Model metadata, framework interface, simulations.

Databases scope, interfaces, approach for conditions DB, use cases.

Offline processing current processing, future requirements.

DAQ

- How to split a raw "event" (6 GB - 120 TB) across files.
- Details types of info in DAQ output data.
- Features DAQ may provide to optimize offline read-in.
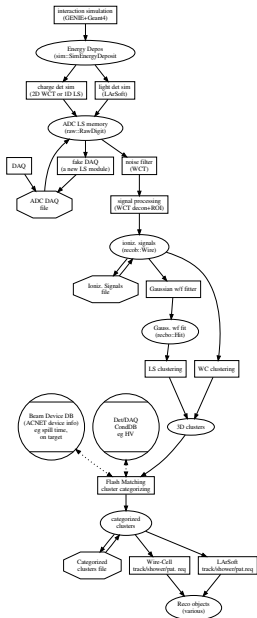- Expected content of DAQ database fed to offline Conditions DB.

Event Model

- Sketch of a file "metadata" aka "data discovery" system
  - Aka SAM's replacement
  - Connect file provenance with Conditions
- Data tiers starting with "raw" then "signals" (10^4 reduction after detector response deconvolution).

Databases

- Proposal for general-purpose R&D of noSQL-style solutions (Maxim)
- Conditions: current ProtoDUNE and BelleII systems described.
- Interest in using BelleII CondDB solution. Fermilab has system with very similar design (UConDB). Comparison study, DUNE requirements and formal(ish) review, tbd.
- DUNE can not (in general) have 1-to-1 mapping between "events" and files. Solution has possible synergies with ATLAS Event Service, et al. Smart/efficient association between file metadata and conditions needed.

Offline Processing

- Overview of current resource contributions. FNAL $\approx$ 25-50% CPU, CERN $\approx$25-33%, BNL #5 at 5-10%.
- Initial "big DAQ", full but incomplete processing chain
- Geometry (gegede source generating GDML for consumption)

interaction simulation
(GENIE+Geant4)

Energy Depos
(sim::SimEnergyDeposit)

change det sim
(2D WCT or 1D LS)

light det sim
(LArSoft)

ADC LS memory
(raw::RawDigit)

DAQ

fake DAQ
(a new LS module)

noise filter
(WCT)

ADC DAQ
file

signal processing
(WCT deconv+ROI)

ioniz. signals
(recob::Wire)

Ioniz. Signals
file

Gaussian w/f fitter

Gauss. w/f fit
(recbo::Hit)

LS clustering

WC clustering

Beam Device DB
(ACNET device info)
eg spill time,
on target

Det/DAQ
CondDB
eg HV

3D clusters

Flash Matching
cluster categorizing

categorized
clusters

Categorized
clusters file

Wire-Cell
track/shower/pat. rec.

LArSoft
track/shower/pat.req.

Reco objects
(various)

Next steps

- Continuing to capture goals, ideas, conclusions, requirements, interfaces, action items in a set of documents started at the w/s.
- 9-11 Sep 2019, DUNE Computing Model workshop at FNAL.
  - Major topics: data management, workload and workflow management, resource board, offline workflow design, computing centers.
  - https://indico.fnal.gov/event/21231/