protoDUNE Raw Data Management and SP Disk Buffer Farm

Brett Viren

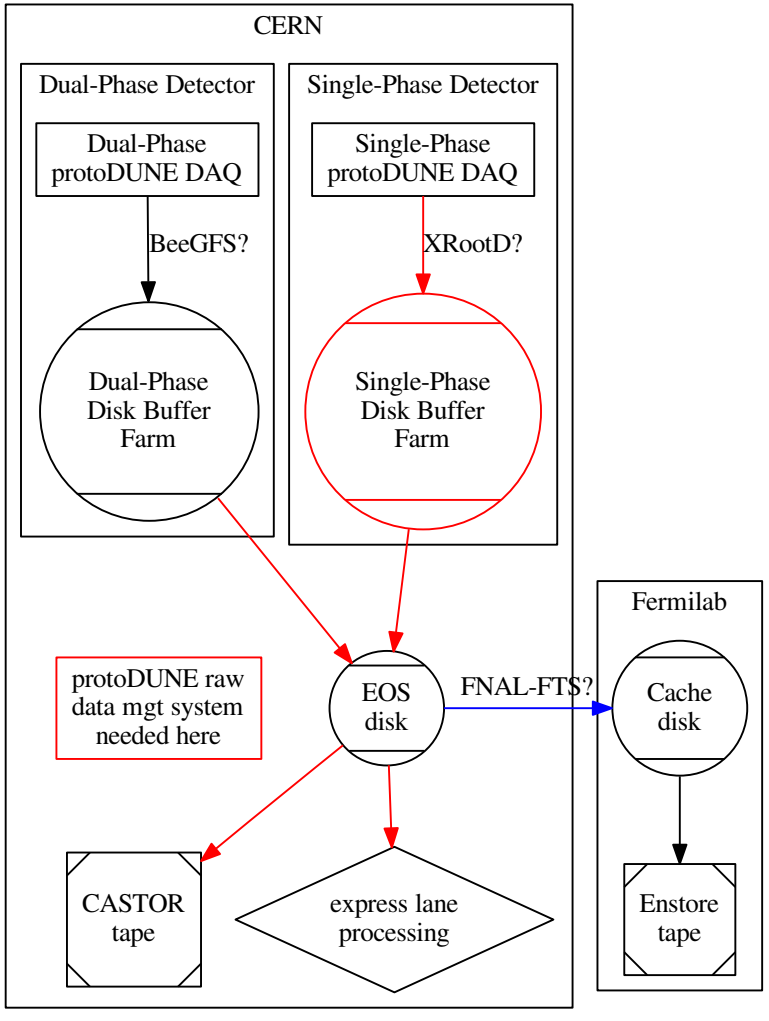
Physics Department

BROOKHAVEN
NATIONAL LABORATORY

High Level Data Flow

SP+DP functional symmetry starting at buffer farms.

- **Disk Buffer Farm** (BF)
  - isolate DAQs/EOS $\lesssim$3 days
  - exp. provides hardware and DAQ$\rightarrow$buffer transfer method
- **Raw Data Mgt Sys** (RDMS)
  - Surveyed existing systems.
  - **Likely use FNAL-FTS**
  - Setup and support effort needs understanding.
- Copy to Fermilab
  - Use FNAL-tech at `fnal.gov`
  - Little integration effort needed if use FNAL-FTS as RDMS.



High-level protoDUNE raw data flow

Online Disk Buffer Farms (DP+SP)

CERN wants BFs to isolate DAQs/EOS for $\lesssim$3 days of data.

- Dual Phase (DP) detector:
  - Nominally **2.5 PB** total data.
  - Keep all DP data on DP BF.
- Single Phase (SP) detector:
  - Uncertainty range **0.1-1.0 PB** total data.
  - SP BF is just for buffering SP data.
- Common:
  - 2 month data run, maybe ping/pong 1 week each.
  - Shared **Raw Data Management System**.
  - All raw data to CERN EOS/CASTOR, FNAL (and elsewhere).

Single-Phase Online Disk Buffer Farm R&D

- The need for a buffer farm only recently understood.
- Size, performance, features not yet well understood
  - Driven by data size (see next).
  - In all likely scenarios, sustained **parallel writes** to buffer farm.
- BBL intends to do some R&D:
  - Will present letter of intent to pDCIG[1].
  - Look at using **XRootD** for DAQ→BF transfer.
  - Understand **scaling** and **requirements on DAQ**.
  - Culminate in **design, cost, plan** to be given to pDCIG

[1] "protoDUNE Computing Interface Group": representation from CERN, FNAL, DUNE Software & Computing, both protoDUNE detectors

Single-Phase Data Rate/Volume Scenarios

	low	high	Goldilocks
trigger (Hz)	10	100	10
readout (ms)	2	5	2
#APA read	1	6	3
inst. rate (GB/s)	0.08	11.5	0.23
avg. rate (GB/s)	0.02	3.3	0.07
inst. $\|\|$ HDD writes	1.5	230	4.6
avg. $\|\|$ HDD writes	0.4	66	1.3
3 days disk (TB)	5.7	850	17
total volume (TB)	76.8	1150	230

- Common assumptions: spill time = 4.8s, cycle=16.8s, 2560 ch/APA, digitize at 2MHz$\times$12bit, 2x compression, 50MB/s HDD.
- Uncertainties: (+)trigger efficiency, (+)calibration data, (-)faster disks
  - $\rightarrow$ Already added Goldilocks$\times$2 for dedicated cosmic muon triggers
  - $\rightarrow$ Probably 10Hz is too slow.

High Level Requirements for RDMS

- Transfer raw data files between different storage elements at same and between different sites.
- Record metadata about file status and outcome of file operations.
- Operate at CERN and FNAL with support for initial set up and ongoing operations.
- Provide monitoring of overall system health, alerts on error and support debugging of problems.
- Provide triggers to perform file operations (copy, delete) based on configurable rules.
- Support "express lane" processing at CERN, FNAL and other institutions.

Discussion points for F-FTS

Already lots of great, positive answers in email on questions about F-FTS in the areas of:

- Complexity
- Environment
- Network
- Compatibility
- Monitoring
- Spooling

A few more open ended points of discussion$\longrightarrow$

Discussion: Support

Agreements probably need to be worked out by others, so let's try to think about what makes sense from a technical point of view.

- Who provisions FTS servers (HW & SW) to run at CERN?
- Who does OS selection, install and maintenance?
- How about issue tracking/resolution?
  Eg: find, understand and fix:
  - Mismatchs between file catalog and reality.
  - Transfer failures.

Discussion: Flexibility

How does F-FTS react to changes in participants?

- What is involved in adding a new source/sink.
- What if we want to change "trigger logic":
  - Eg: never delete files from DP BF but delete a file from SP BF as soon as copy to EOS succeeds)
  - How can FTS help support "express lane" processing.

Discussion: Wider Use

How can we make use of F-FTS to solve other problems?

- Examples:
  - Run "express lane" processing at other institutions with fat pipes (eg, FNAL, BNL, EU/UK labs)
  - Institution XXX wants copy of all Kaon data from SP.
  - Transfer files of data set ZZZ to my Ubuntu workstation at BNL.
- What requirements must a site satisfy to participate, integrate or use F-FTS?
  - Eg, say Brazil wants to host a partial mirror of data through F-FTS.

Possible Next Steps After Today

Not necessarily in order and we should discuss if/how order matters.

- Set up some F-FTS components at BNL.
  - Gain deeper understanding of F-FTS and confirm no show stoppers.
  - Set up mock transfers ("BNL as CERN") to FNAL.
  - Experience later useful when FNAL→BNL transfers are done.
  - Maybe need to balance usefulness/coverage vs. setup effort?
- Get sandboxes at CERN and repeat the above.
- Communication (here I think order matters)
  1. Invite involvement from DP (probably Jacques Marteau?)
  2. Submit letter of intent to pDCIG group will pursue F-FTS for protoDUNE.
  3. Submit proposal for F-FTS based RDMS with design, costs, support model, plans, timeline, etc.