ProtoDUNE Computing Model

Maxim Potekhin (BNL)

DUNE S&C general meeting

05/16/2016

Powered by

BROOKHAVEN NATIONAL LABORATORY

DUNE

The DUNE Computing Model

- "The DUNE Computing Model" is the name given to the document which contains the computing model itself and a substantial amount of supporting materials to it:
  - DUNE Data Characteristics, for both raw and derived data, with estimates for the FD and ND
  - FD and ND DAQ and Monitoring
  - General Software and Computing Requirements
  - The Computing Model itself
  - Overview of the 35t computing
  - protoDUNE: single phase and dual phase (as an Appendix to the document, not a standalone paper)

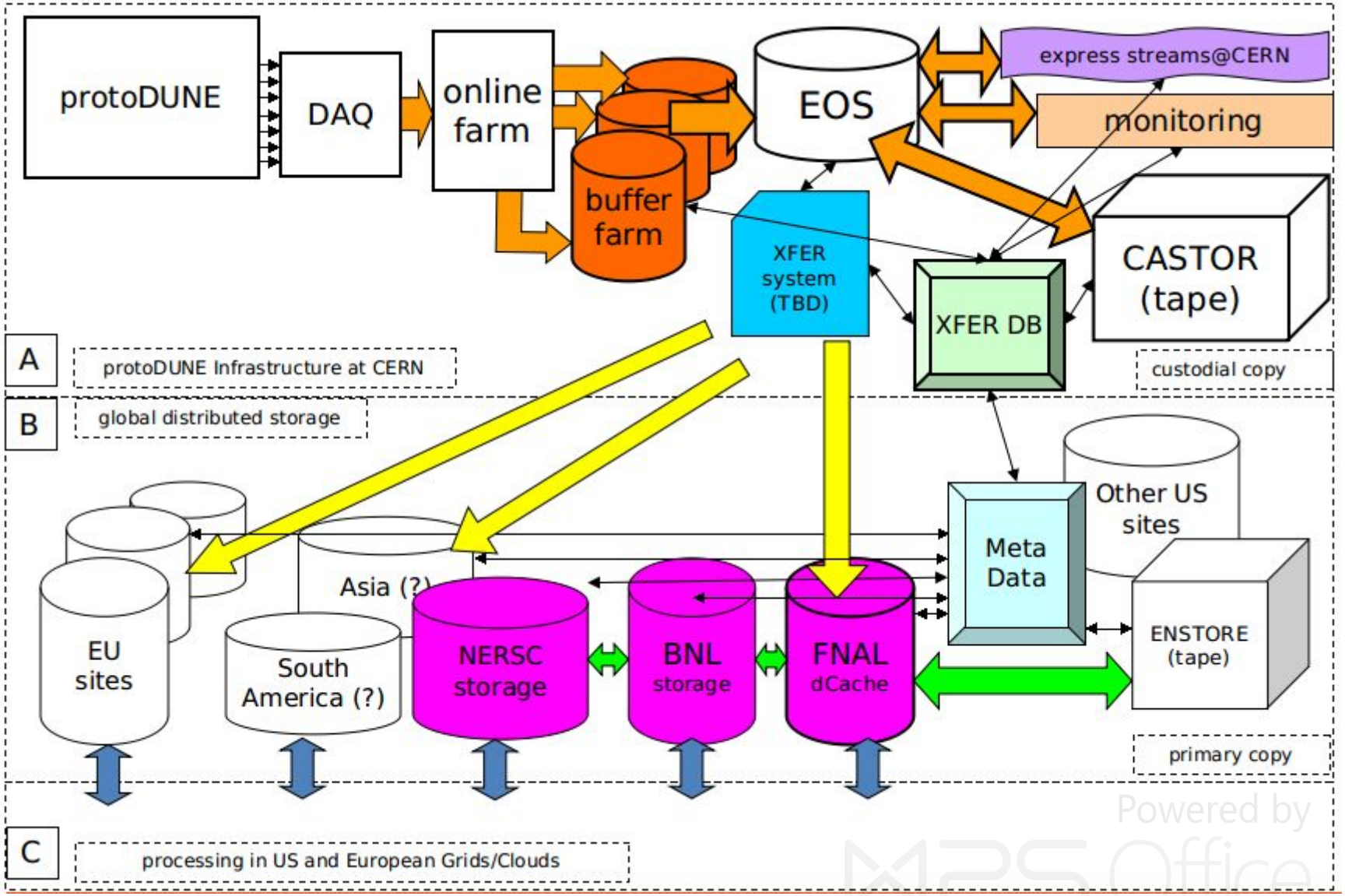
Powered by MPS Office

What's new?

- First version of the Computing Model completed in December 2015, with caveats
  - the dual-phase work was not represented in much detail, in either the main computing model or the prototype sections
  - it reflected our earlier intention to take most or all data in NP04 with zero suppression, a decision that since has been revisited
- Since this version was released, a few developments took place:
  - protoDUNE meeting at CERN in early 2016
  - a new strategy has emerged where NP04 will aim to take a large fraction of their data without Zero Suppression
  - a follow-up meeting of BNL and FNAL teams which resulted in a draft design document for the data management system
  - the protoDUNE "interface group" is being (has been?) set up, with the mandate to coordinate protoDUNE computing among NP02/04 and CERN/FNAL. Convener: Amir Farbin
- As a result of these developments, the protoDUNE portion of the Computing Model has been updated.

Powered by MPS Office

Raw data flow in protoDUNE (conceptual diagram)



Powered by MPS Office

Data handling in NP02 and NP04

- NP02 has plans for substantial computing resources (both storage and CPU) to be deployed "in the pit".
- Processing is to take place at the same rate as data is produced by the experiment, with calibrations "embedded" into the process.
- The goal is to reconstruct ~70 cosmic ray tracks in real time, while collecting some metrics about the detector performance.
- In the process described above, event data resulting from reconstruction is not to be saved to permanent storage (or at all).
- A select sample of data will be preserved for more detailed offline reconstruction (presumably according to a run plan and operators' decisions). XFER to EOS completely asynchronous!
- According to plans, NP04 is not going to possess substantial real-time processing capability, outside of monitoring to be implemented as a part of DAQ.
- In NP04, the emphasis is on high-bandwidth, reliable data buffering and transmission to CERN EOS and beyond - to FNAL and other sites
  - different bandwidth utilization compared to NP02
- We believe this approach will allow to reconstruct a portion of data in near-time for QA and other purposes with good scalability.
- File size of 5GB is the nominal choice in NP02 and NP04.

Powered by MRS Office

Data characteristics estimates for NP04

- A range of scenarios is being considered since the parameters of the beam, triggering capabilities, detailed run plan and other factors are still being investigated
- Brett has compiled a table containing the "low", "high" and "Goldilocks" scenario, the latter getting its name from the fact that it is technologically quite feasible.

Parameter	Low	High	Medium
trigger (Hz)	10	100	10
readout window (ms)	2	5	2
number of APAs read out	1	6	3
instantaneous rate (GB/s)	0.08	11.5	0.23
average rate (GB/s)	0.02	3.3	0.07
instantnaneous HDD writes	1.5	230	4.6
average HDD writes	0.4	66	1.3
disk space for 3 days (TB)	5.7	850	17
total data volume (TB)	76.8	1,150	230

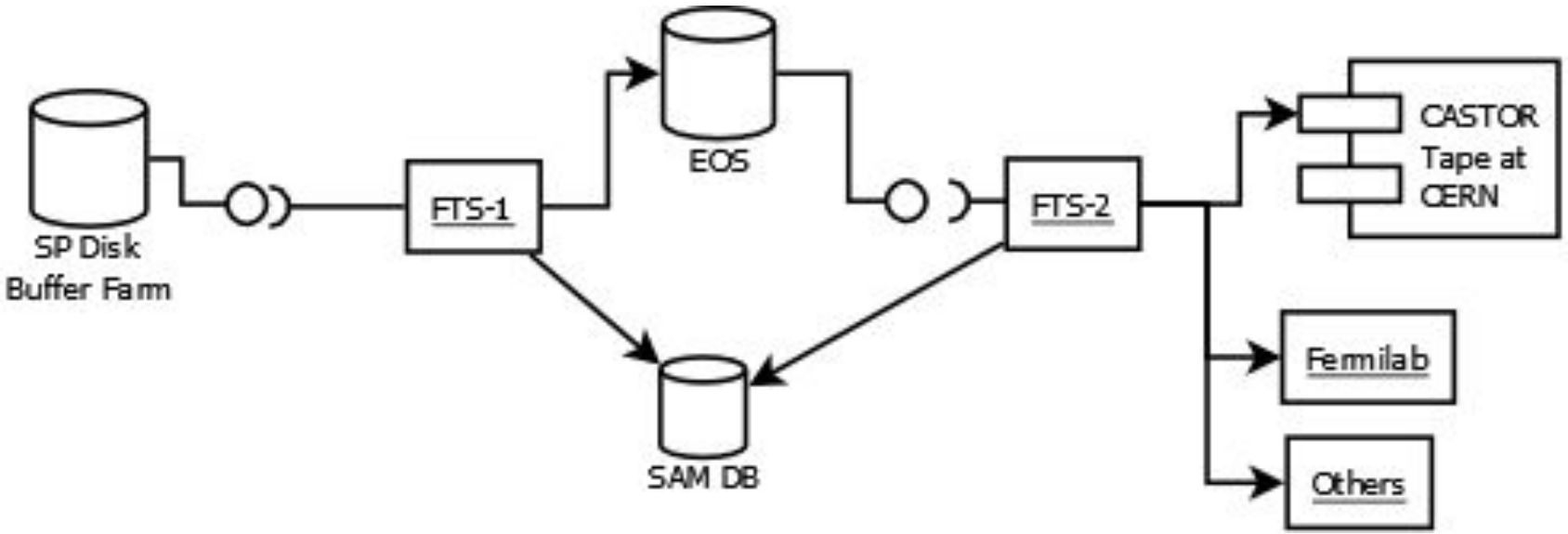
Powered by WPS Office

protoDUNE Raw Data Management System (RDMS)

- The high-level (technology-agnostic) requirements were summarized in DUNE DocDB **1209**
- They formed a basis with a series of discussions with FNAL data experts
- The F-FTS and SAM in combination appear to satisfy the requirements
- Originally, two-stage FTS-based system was envisioned for moving the data:
  - FTS-1 reading data from a dropbox (the buffer attached to DAQ) and moving the data saved in the buffer by the DAQ, to EOS
  - FTS-2 creating a tape replica at CERN and also moving data to FNAL (dCache, Enstore) and other sites such as BNL
  - SAM database keeping track of the state of every unit of data and outcome of various operations on it
  - Same architecture for Single (NP04) and Dual Phase (NP02) protoDUNEs
  - See next slide
- Feedback from NP02 indicated that at this stage the team wants to reduce dependencies on components not native to their DAQ and local (in the pit) data processing system in order to facilitate the latter
  - decision has been made to defer the choice of technology for moving data to EOS to a later date (i.e. no FTS-1 in current approach)
  - FTS-2 is still seen as the right tool to perform data transfer to mass storage downstream from EOS
- Reference by NP02 to "CERN standard scripts". Any layer of automation on top of it?

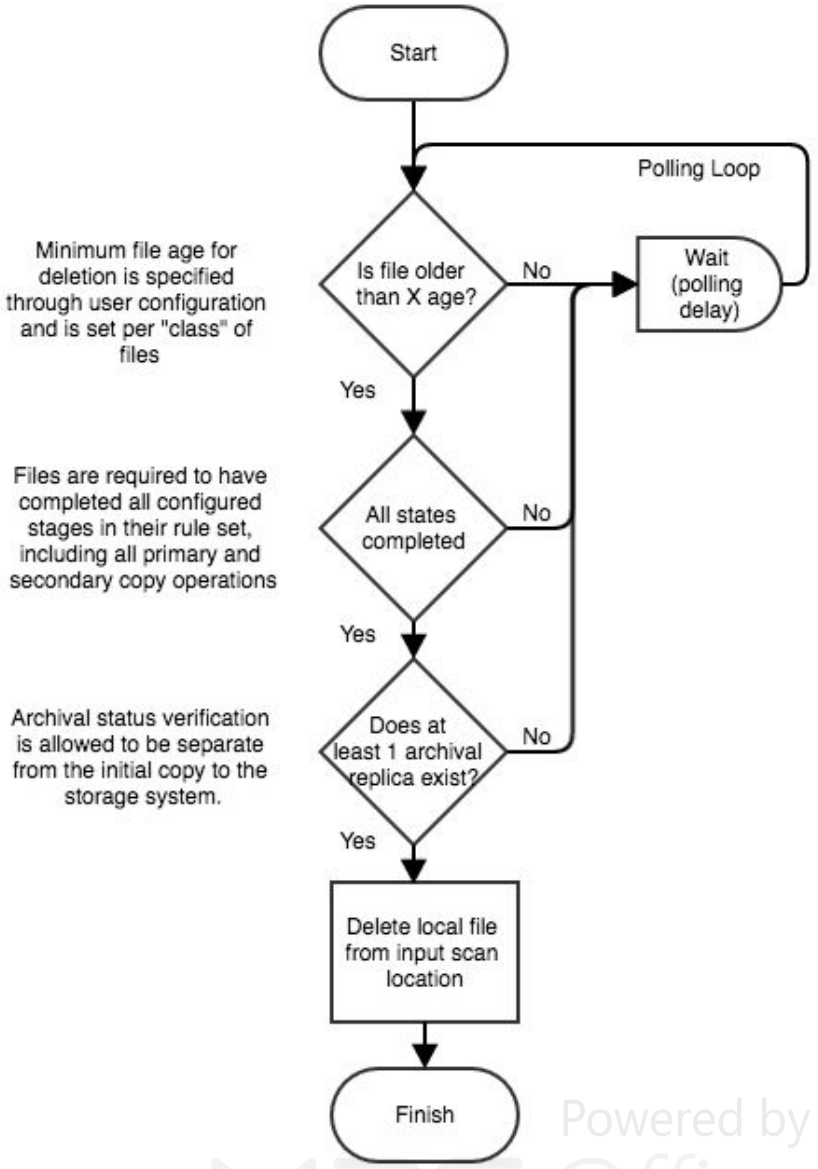
Powered by WPS Office

F-FTS instances



Powered by
WPS Office

Management of file lifetime and deletion in F-FTS



Powered by WPS Office

Current design of the RDMS

- The document "Design of the Data Management System for the protoDUNE Experiment" has been developed based on the original draft and then reworked, filed as DUNE DocDB **1212**
- Revision control in github: https://github.com/DUNE/protodune-raw-data-mgt
- Feedback from FNAL and NP02/04 received which motivated changes such as a provision for a separate data transfer mechanism from NP02 to EOS, not tied to F-FTS

Powered by MPS Office

Updates to the Computing Model

- The document "Design of the Data Management System for the protoDUNE Experiment" (DocDB **1212)** has been used to provide online computing portion of the protoDUNE appendix.
- Close to a complete rewrite of previously available material.
- A few new quantitative data characteristics included (see the table in this presentation).
- Offline part for protoDUNE (quite small) was carried over from the 2015 version as no progress has been made in developing that.
- Revision control in github: https://github.com/DUNE/dune-computing-model

Powered by MPS Office

Known unknowns

- NP02
  - To an uninitiated reader, it may appear that there is no sufficient detail on the design of online data handling and computational power deployment in NP02 - the goals seem quite ambitious and the scale of real-time computing considerable.
  - Interface to CERN storage should be better documented (in addition to "standard scripts"
  - IMHO needs more development and documentation - NP02 team is the best group of people to do that - the document can be edited using the repository in github.
- Data placement: nominally all of the NP02 data will need to be transported to FNAL, stored and eventually processed. Question - is that optimal? What other Grid sites can be involved in this?
- Offline computing for protoDUNE in general: at some point we'll need to conisder tools for task and job management, perhaps beyond jobsub and closer to a full-fledged workflow control system.

Powered by MPS Office