Data Management for DUNE and FIFE

Robert Illingworth
Fermilab Scientific Computing Division
Rucio Community Workshop
2 March 2018

Definitions

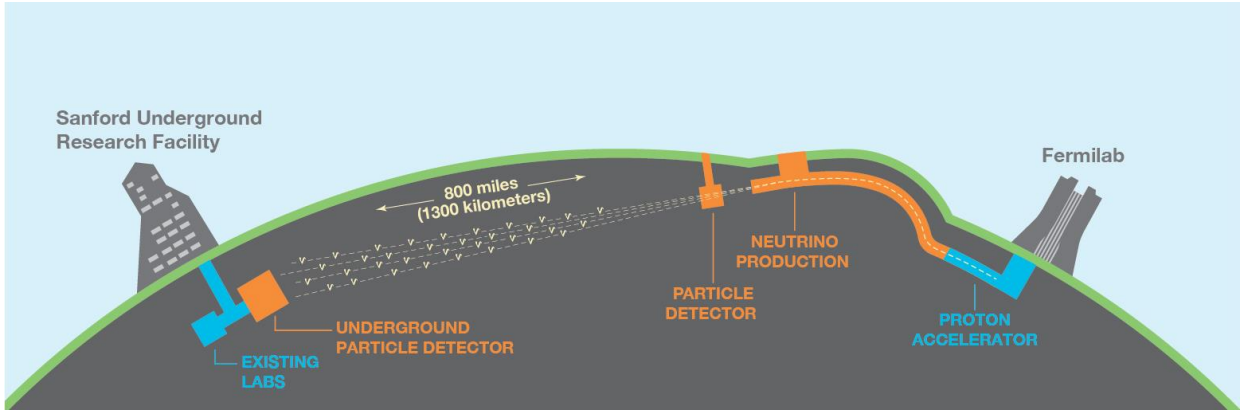
DUNE – Deep Underground Neutrino Experiment

http://www.dunescience.org/

Send a beam of neutrinos from Fermilab to South Dakota starting in ~2026

40kt Liquid Argon TPC Far Detector

Smaller Near Detector at FNAL (tracker/calorimeter)



Definitions

FIFE – Fabric for Frontier Experiments

- Project to provide common computing services and tools for multiple Fermilab experiments. FIFE is modular so experiments can take what they need, and new tools from outside communities can be incorporated as they develop.

- FIFE experiments are relatively small by recent HEP standards (10s to few 100s of scientists)
  - Usually very limited local computing expertise
- Data is up to ~5 PB/yr scale

DUNE and FIFE current DM status

- DUNE and multiple FIFE experiments are currently using the Fermilab SAM data management system
  - SAM dates back to Tevatron Run II
  - Monolithic metadata catalogue, replica catalogue
    - Rich metadata system with very flexible queries and dataset definitions
  - Operations workload is low: <0.5 FTE support for a dozen experiments (with varying levels of activity)
    - FNAL SCD runs the central services and provides consultation; we don't run experiment operations
    - Getting multiple experiments to use a common system has been a huge gain in reducing the support load

DUNE and FIFE current DM usage

- Data workflows are mostly archival and cache-based
  - Archival data uploaded to FNAL tape system
  - Data access done through cache-based access (dCache)
  - Some limited use of replication, but most data access is to the central FNAL copy

- This year the test beam ProtoDUNE will archive data at FNAL + CERN
  - Test beam experiment so the requirements are something of a moving target
    - But you can't really handle multiple PB this way…

Common requirements

- Everyone seems to want to store rich physics metadata
  - We see a lot of abuses
    - file names that track the processing ancestry of the file
    - Text fields in the database contorted to store all sorts of things
  - Linking metadata/datasets to experiment conditions and quality databases is also popular
  - Need APIs that facilitate this

- Interoperability with experiment frameworks
  - Many, but not all, use a common processing framework
  - Need a system that can interface with different workload management systems and frameworks

Common requirements

- Data lifecycle management
  - This isn't something that most of our experimenters think about
  - Currently they get little assistance
    - The result is that the tape system probably contains a lot of junk, but usually nobody is sure and it's hard to get people to take responsibility for deleting it
    - This isn't viable in the long term
  - We need to force this as a consideration right from the beginning

DUNE requirements

- Large international collaboration
  - Far detector data will start 1,475 m underground in South Dakota
  - Has to be distributed to FNAL and other archive sites
- Data rates are still somewhat conjectural
  - Depends on matters like choice of readout technology
  - The detector is very large, but mostly empty of activity
    - Typical readout ~1-2 MB
  - Beam events are rare: $< 10^5$/yr
  - Cosmic rays: ~10^3/hr
  - Supernova readout is very large (over ~10s period) but rare
  - DAQ system is working to 30 PB/yr raw data

FIFE experiment requirements

- FIFE experiments are smaller than DUNE in data rates
  - Scalability is less of a concern
  - Limited or no ability to run services at most sites
- Main requirements are
  - Flexibility: handle varying workflows
  - Easy integration: experiments have limited expertise to get their DAQ & processing workflows adapted to the system. They need something easy to plug into
  - Low experiment workload: again, limited expertise – e.g. data replication will have to be automatic or at least policy based; manual management isn't going to work.
  - Low support requirements: we need to support multiple experiments with a small team

Unusual (for HEP) workflows

- Some of the workflows depart slightly from common HEP practices
  - Near and Far detectors: several experiments have detectors at both ends of a long beamline; potentially requiring processing correlated datasets
    - In practice this has not been much of an issue
  - One experiments incorporated data recorded by another
    - Minerva effectively incorporated the Minos detector
    - Access was much simpler because both used the same DM system
  - Rather than traditional track fitting then data classification/reduction, some of the analysis is effectively doing hit level image recognition
    - This workload could potentially benefit from object store style access

Plans

- The SAM data management basic design dates back 20 years
  - Scalability is a concern
  - Continuing use would likely require major architectural changes
- We're currently evaluating options for the future
  - A suitable community wide (or beyond) system would potentially be a major benefit
  - Existing experiments tend to be very conservative and will only agree to change if a new system is very similar to the old one
  - Upcoming experiments tend to have vaguely defined data models and so can be pushed to adapt
  - Rucio seems to hit many of the requirements
    - We're setting up an evaluation system

Final thoughts

- While writing this I found I was much less concerned about the technical details of data management than I was about the “soft” stuff
  - Ease of integration
  - Usable with limited expertise and effort by experiments
  - Simplicity of operations support

- I don’t think the Fermilab intensity frontier program would have achieved what it has without the many of the experiments (especially the larger ones in terms of data volumes) making use of a common DM system

- I think there would be real benefits if there was an even more widely used common system