DUNE Production processing and workflow management software evaluation

Ken Herner for the DUNE Collaboration

CHEP 2019

4 Nov 2019

This document was prepared by [DUNE Collaboration] using the resources of the Fermi National Accelerator Laboratory (Fermilab), a U.S. Department of Energy, Office of Science, HEP User Facility. Fermilab is managed by Fermi Research Alliance, LLC (FRA), acting under Contract No. DE-AC

Outline

- Overview of DUNE and ProtoDUNE
- Production Group responsibilities and infrastructure
- Current and future workflow management options

DUNE and ProtoDUNE

- DUNE

- Future long-baseline neutrino experiment; near (FNAL) and far (SURF) detectors
- Far det: 4 liquid argon TPCs

- ProtoDUNE

- Two LAr TPC detectors, 1/20 size of regular DUNE far detectors
- Single-phase operational in 2018
- Dual-phase operational in 2019
- Technology demonstrators

See Heidi Schellman's plenary talk for more details

CERN and ProtoDUNE

ProtoDUNE inside EHN1 at CERN

Current Production group activities

- Serve as central processing group for ProtoDUNE data reconstruction, large-scale MC simulation, site and resource commissioning
- MC and data reco are via the LArSoft software suite (shared framework with other FNAL LAr TPC experiments; based on Art framework)
- Two major ProtoDUNE-SP processing passes of beam data so far
- ProtoDUNE-Dual-phase keepup reconstruction to begin soon

ProtoDUNE by the numbers

Raw data (SP+DP): 3329 TiB

Raw “physics” beam data (SP): 786 TiB

Most recent beam reprocessing
reconstruction output size (SP “good” runs
only): 169 TiB

Wall hours for most recent reprocessing pass
(data+new MC): 2.08 M

Current setup: Job submission

- Job submission is via POMS (see POMS poster for details)
- Resource/slot provisioning is with GlideinWMS (setup shared with other FNAL IF and muon expts.)
- DUNE software built for both SL6/7
- Copyback is generally to FNAL dCache, other sites demonstrated
- Exploring creation of a global gWMS pool similar to CMS; would allow for additional submitter resources to come online

Current Setup: Data movement

- DUNE using the FNAL SAM system for file catalog and delivery
- Data replication being handled by Rucio instance
- Most input streamed with xrootd; output usually returned via gridftp (can easily use other protocols as needed)
- Auxiliary file input (needed for MC generation) now handled via StashCache; used heavily in Spring 2019 (1.75 PB transferred)

International Contributions

DUNE already getting significant contributions from international partners
In 2019 so far, **49% of production wall hours are from outside USA**

Actively working to add more sites and countries

Future workflow management systems

- Current setup adequate for ProtoDUNE runs; full DUNE brings additional challenges
 - POMS+gWMS will evolve; will work to meet DUNE's requirements
 - DUNE has not chosen a workflow management system for future running
- Other systems being considered: PANDA, DIRAC (UK collaborators studying DIRAC now)
- How do we decide what to choose?

Future workflow management systems (2)

- How do we decide? Do it from a requirements POV (must do X,Y,Z, system must show that they can manage all job types)
- What does the future system need to do?
 - Be able to provision on variety of architectures (HTC, HPC/LCF, GPU)
 - Schedule large blocks of jobs at HPC resources when work demands it
 - Support for “pipeline” workflows
 - Jobs to data vs. data to jobs: Both?
- The full DUNE Computing Model will drive the requirements
- **Can serve as important precedent for methods of component evaluation and selection in other areas of computing**

Production Group future (6-18 month) plans

- Assist in workflow management software evaluation
- Prepare for second ProtoDUNE run in 2021-22
- Work closely with Data Management group on future file data movement protocols and Rucio integration
- Move to more distributed I/O setup as additional institutions provide storage
- Test additional workflows on HPC sites
- Actively integrate compute resources from new contributors
- Strengthen engagement with DUNE Near Detector software efforts

Timeline not to scale

Summary

- DUNE Production handling all large-scale data reconstruction, reprocessing and simulation for the collaboration
 - Plays an active role in new site integration and testing. DUNE already has an active and growing community. Exciting time to be involved!
- Current job submission and workflow system based on FNAL's POMS and resource provisioning is via GlideinWMS
- Future workflow system has not been chosen yet; multiple systems being evaluated
 - DUNE is working on developing requirements based on the computing model; no official timetable yet
- Production group planning to adopt new practices and technologies in as things evolve in preparation for ProtoDUNE-DP cosmic run and next ProtoDUNE beam run with both detectors.

BACKUP

Aside: GlideinWMS

GlideinWMS is a workload management system sitting on top of the local batch system

Extensively used by CMS, IF expts, Open Science Grid projects, IceCube, LIGO, etc.

Adding DUNE support is *trivial* if you support any of these other groups

Has long-term support

GWMS Factory

GlideinWMS Frontend

DIRAC

DIRAC in use by other experiments (e.g. LHCb)

Successful tests with integrating SAM project and DIRAC test jobs.

Further work ongoing