DUNE CERN prototype Computing: networking and storage requirements

Maxim Potekhin
Brett Viren
Brookhaven National Laboratory

CERN integration meeting
July 30, 2015

Executive Summary

- As was recently communicated, the Integration Team has plans to contact CERN ITD regarding initial set of computing infrastructure requirements for the test, such as network bandwidth.
- In order to maximize utilization of the beam, the nominal trigger rate (in-spill) was chosen at 200Hz, and is limited to avoid pile-up of beam particles in the liquid Argon TPC.
- Most of the data to be collected will be due not the beam particles, but to exposure to cosmic ray muons, a few dozen tracks per read-out.
- According to current estimates, the DUNE prototype will need a connection of ~2gbps bandwidth.
- Details of the run plan, including commissioning etc, are still being worked out, and so are estimates for the efficiency of reconstruction in the conditions of cosmic ray background. For this reason, only lower estimate for the total raw data volume is available at this time for the minimalistic run plan with 100% efficiency, and it's 25TB worth of "golden events".
- This is likely to increase by a factor of ~10 as the calibrations, commissioning and efficiencies are taken into account. This information will take more time and effort to be developed.

Overview

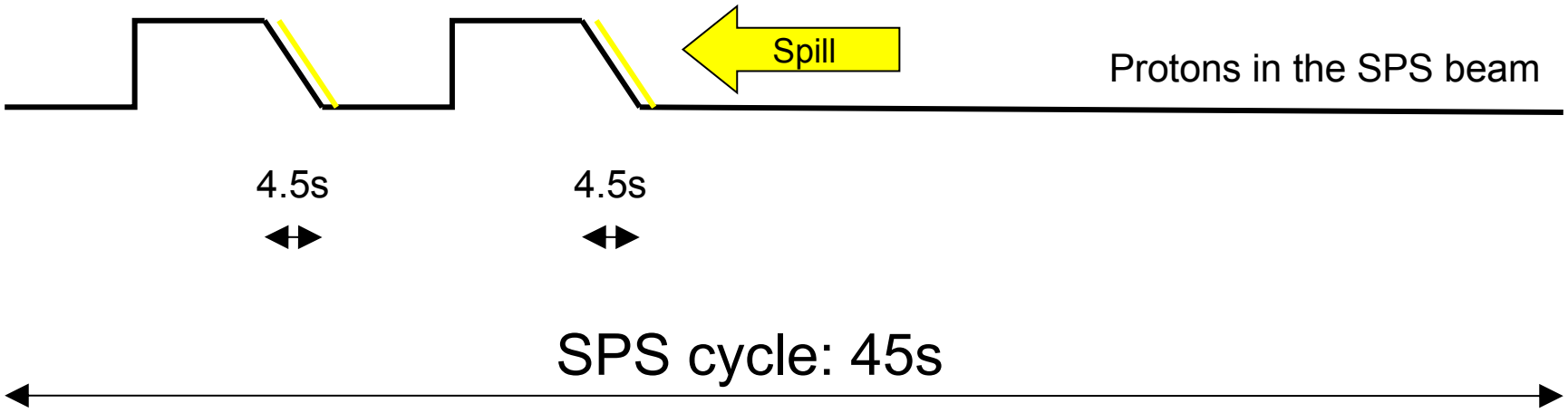
- Recording the data collected with the prototype is obviously critical to the experiment but it also constitutes one of the central points of interaction and integration between DUNE prototype and CERN.
- According to the proposal submitted to SPSC, the data collected with the DUNE prototype at CERN is to be committed to tape at CERN as the primary custodial copy.
- Data will be replicated from CERN to a few locations in the United States. There will be a full replica of the data at FNAL used for production, and partial copies at other DUNE institutions.
- Bulk of the processing will be performed at the US computing sites. An effort will be made to establish "express processing streams" for a portion of the data for QA and debugging purposes. If possible, some of these streams would run at CERN.
- Disclaimer: we are still in early stages of planning and developing relevant information, so any and all numbers quoted in this presentation should be considered as preliminary.

Cosmic ray muons

- One significant factor that defines the characteristics of LAr TPC in the context of the test beam is the combination of the relatively long electron drift time (nominally ~2.25ms) and occupancy from cosmic ray muons (cf. the top face of the detector is ~50m^2).
- As a result, it is estimated that there will be ~68 cosmic muon tracks (or track segments) in addition to the "main" triggered beam event. *This includes additional "padding"*, i.e. readout time intervals just before and just after the trigger which are necessary to fully characterize the beam event. These combined extra tracks produce significantly more ionization (and correspondingly more data) than the beam trigger events.
- The minimalistic run plan calls for a total of 5M events that pass applicable selection criteria, in a few event categories. Since the data will be dominated by signals from cosmic ray muons, it is possible to estimate of the total data volume just by looking at the data due to "cosmic muon overlay".
- Signal processing strategies such as Zero Suppression (ZS) and Region of Interest (ROI) are being considered to reduce the volume of the data to be transmitted by DAQ and committed to mass storage. This is work in progress and at this time we aim to provide a range of parameters to help planning process.
- In the following, we assume that ZS and ROOT compression are applied but no ROI. Estimates are based primarily on looking at Monte Carlo data.

Assumptions about the beam

- 45s SPS cycle
- Two spills per cycle, each of 4.5 duration
- Debunching
- Low intensity of the beam to minimize pile-up (i.e. <=200Hz during the spill)



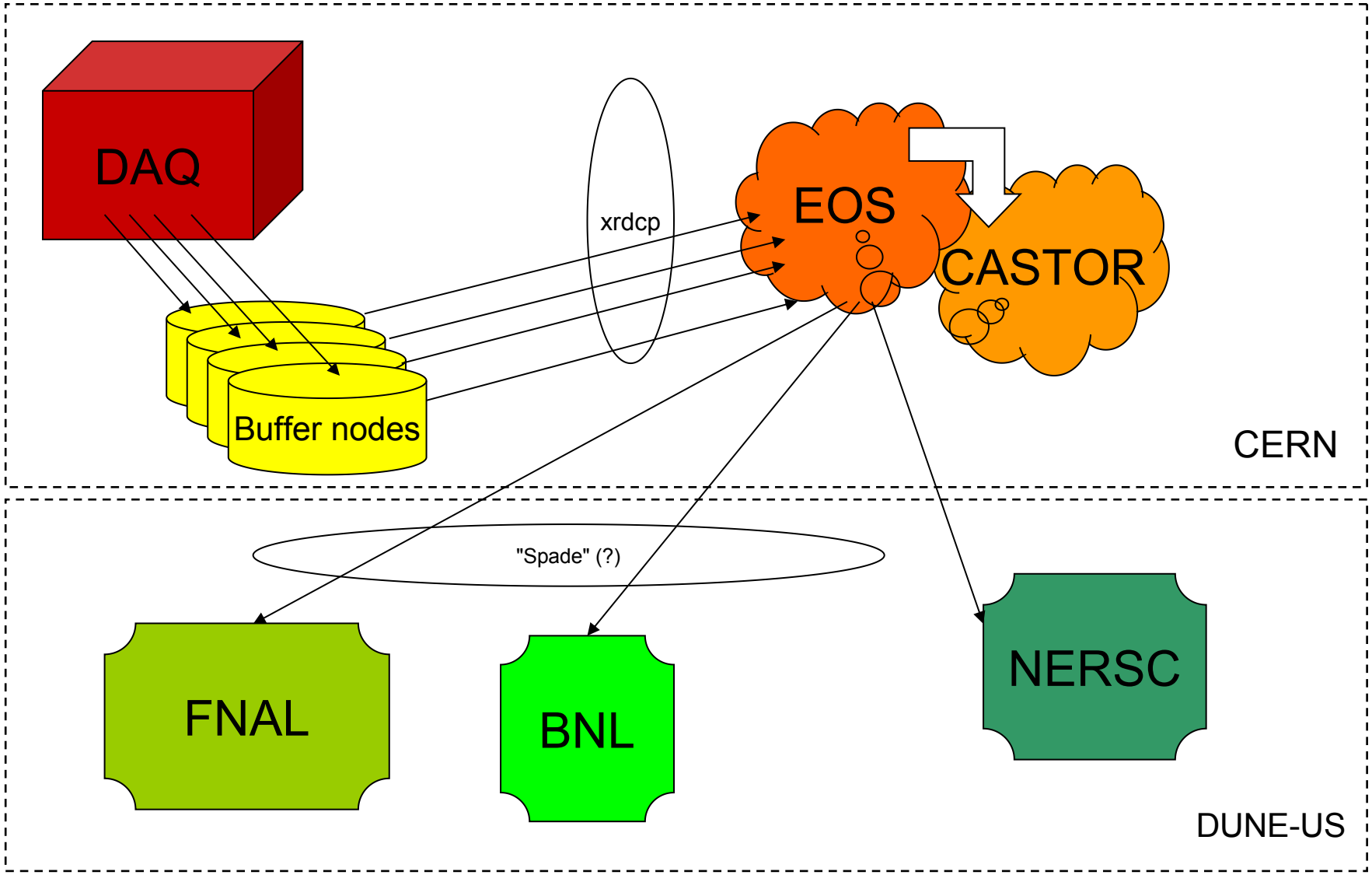
Estimating the Data Rates

- At this point we estimate the volume of zero-suppressed TPC data due to a "nominal" 4GeV muon track as 0.1MB (with ROOT compression). This number obviously depends on data design, formats and algorithms and currently is still dominated by less than perfect packaging of the data, so its dependence on energy is minimal.
- Data from the photon detector subsystem is of substantially smaller volume so it won't be considered here.
- When running at 200Hz trigger rate, we arrive to:
  - ✓ Instantaneous DAQ rate of ~1GB/s (will need buffering nodes in DAQ).
  - ✓ Sustained average rate of 200MB/s.
  - ✓ Required 2gbps connectivity out of the DAQ room.
  - ✓ EOS buffer - assuming 24hr hold time, 20TB disk space.
  - ✓ Tape storage capacity - TBD O(100TB).

Capturing the Data

- LHC experience - both ATLAS and CMS transmit data from their respective DAQ to EOS, with subsequent creation of a custodial copy in CASTOR.
- Buffering in DAQ: several buffer nodes may be needed to absorb the instantenous rate coming out of DAQ.
- DAQ to EOS: multiple connections via xrdcp.
- We are considering the possibility of running limited "express analysis streams" at CERN for data QA purposes. This could be done with the EOS buffer, before the data is committed to CASTOR.
- There are multiple tools under consideration to transport data *from* CERN to data centers in the US and elsewhere, one possibility is to leverage existing expertise with the system called "Spade" (IceCube, Daya Bay) which provides ease of installation, redundancy and monitoring.
- Procedures will need to be put in place to restore data from tape (stored at CERN) in case replicas are corrupted.

Moving the data – initial thoughts…



Other items to be considered

- Commissioning of the detector will take place in ~2 years from now, so to arrive to it with a functional data handling system at the scale being considered we must start planning and prototyping rather soon.
- We would like to obtain guidance from CERN Central Services (and ATLAS and CMS) regarding the EOS and CASTOR interface, since no expertise in that exists in DUNE at this point.
- Ideally, we would like to have a "sandbox" system at CERN with which to practice and prototype the software for the run (e.g. EOS and CASTOR interactions and data transmission to remote sites). The sooner we get onto the learning curve, the better.