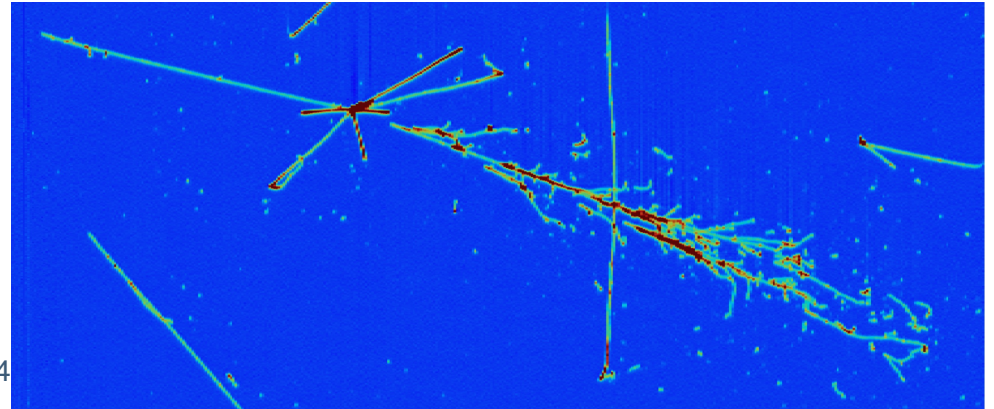


Heidi Schellman
Oregon State University for
the DUNE collaboration

DUNE offline computing

- 30 PB of raw data per year
- ~1500 scientists, ~700 have active computing accounts
- A few million very large trigger records/year (4-8 GB per “event”)
 - $\sim 5 \times 5 \times 1 \text{ mm}^3$ voxels over a volume of $4 \times 10^4 \text{ m}^3$
 - TPC data for the first far detector module are 450 (5000x2560x14 bit) 2-D arrays. Second module is slightly larger.
- Currently setting up to run prototypes at CERN - data rates of 2-3 GB/s



Test beam in a subsection of one of the 450 readout planes

Shared and Unique Tools and Challenges

Shared

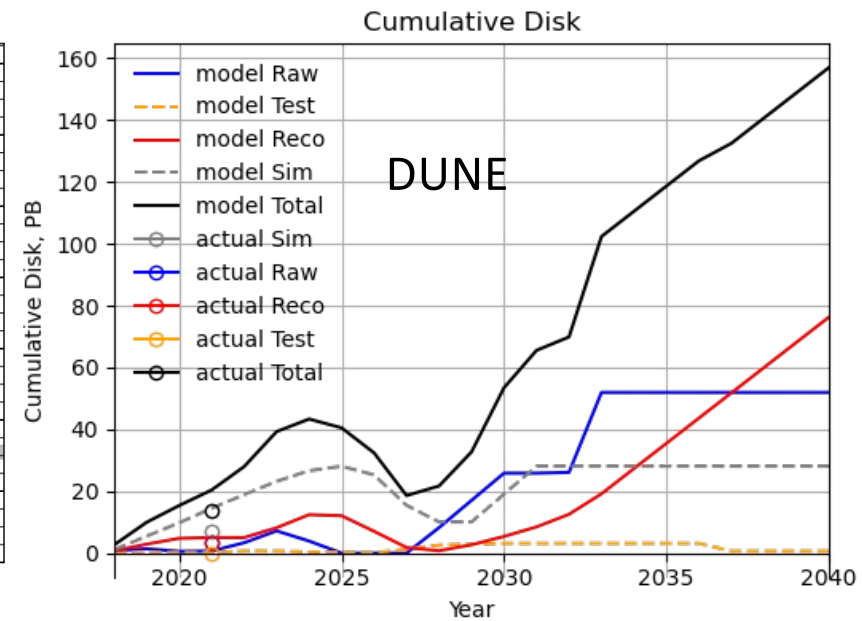
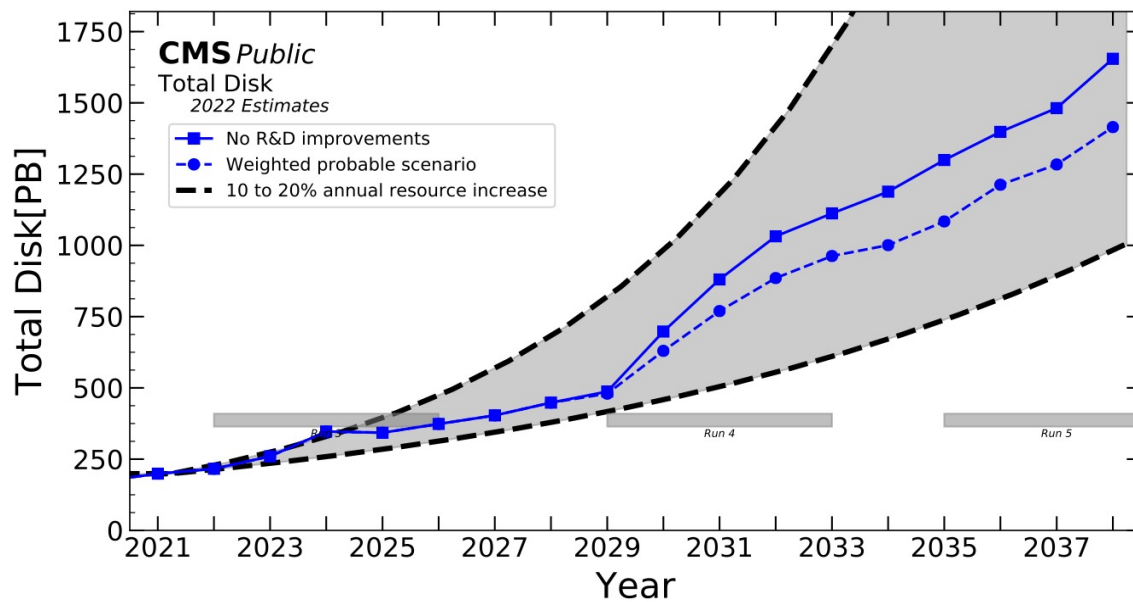
- OSG/WLCG
- CVMFS
- ROOT
- Rucio
- GEANT4
- LArSoft (joint with many others)
- Neutrino event generators
- People – lots of them
- Many training materials

Unique

- Data are produced 1300 km away almost a mile down in a mine with limited space and utilities
- Memory management with 4 GB “events”
- Supernova → ½ PB in 100 seconds
- Calibration of 4x10,000 m³ of flowing charged liquid

Disk compared to CMS

DUNE is ~10% of an HL-LHC expt.



Things I worry about

- Training and documentation – draft text at https://docs.google.com/presentation/d/1KSaGbU1l3DeJ4tC0nKJx3Lt76wwG_AQczEXCIES-0n0/edit?usp=sharing
 - How do we get people the information they need, when they need it?
 - How do we match 1500 people up with the expert help they need?
 - How do we share training and documentation across experiments? Our docs are often not googleable anymore.
 - How do we retain knowledge for decades when our experts have to move on?
- Easy access to resources
 - Getting and renewing computer accounts is getting harder and harder
 - Divergence in authentication regimes makes this even harder
 - What is the tradeoff between flexible access for many (and the need to chase down rogue summer students) vs. tightly controlled applications

Issues for the draft

- Overall, **already in great shape!** Text needed for training...
- DUNE has successfully built out a global computing system using OSG/WLCG/HSF tools created for the LHC like Rucio, cvmfs ... Joint projects and common tools work and should be emphasized. We should **encourage multi-experiment funding mechanisms for international common efforts** and help the smaller experiments join in.
- **Supernova neutrino predictions are important** for large detectors (IceCube, HyperK, DUNE, NOvA, T2K, Juno). They have a different computational footprint than normal event generators.
- Authentication and authorization. Access to heterogeneous resources is hard even if your code can run there due to differences site by site. We need to keep encouraging **robust, secure but standard access methods** to resources.
- And a final one. **Public data access is not just a problem for the Cosmic Frontier.** All US based experiments could and should use common portals for results distribution that have long-term support.