DUNE Computing

H Schellman and M. Kirby

12-17-2021

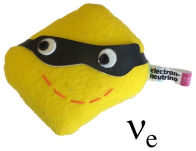
What problems are we trying to solve

- LAr TPC's have very large trigger records (200 MB for protoDUNE vs. < 10 MB for ATLAS/CMS)
- Full readout of 1 Far module is 6 GB (4 Giga 12bit-voxels)
- possible SuperNova is 20,000 times larger -> 460 TB of data
- New detector technologies
- Many subsystems in ND
- We're supposed to use 75% non-DOE computing
- But we're also supposed to run on unique HPC's with interesting IO characteristics

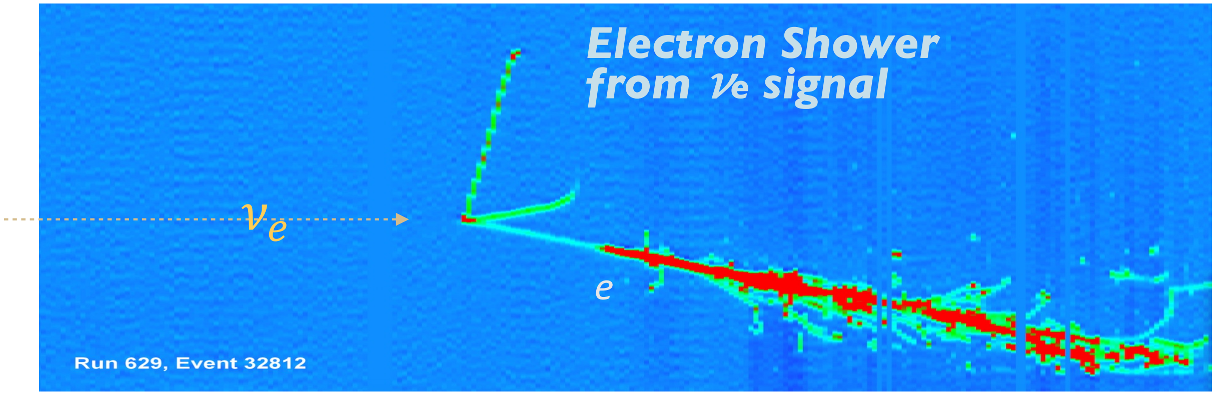
Final state – muon or electron?

e/γ separation

ArgoNeuT
FNAL
2009-10

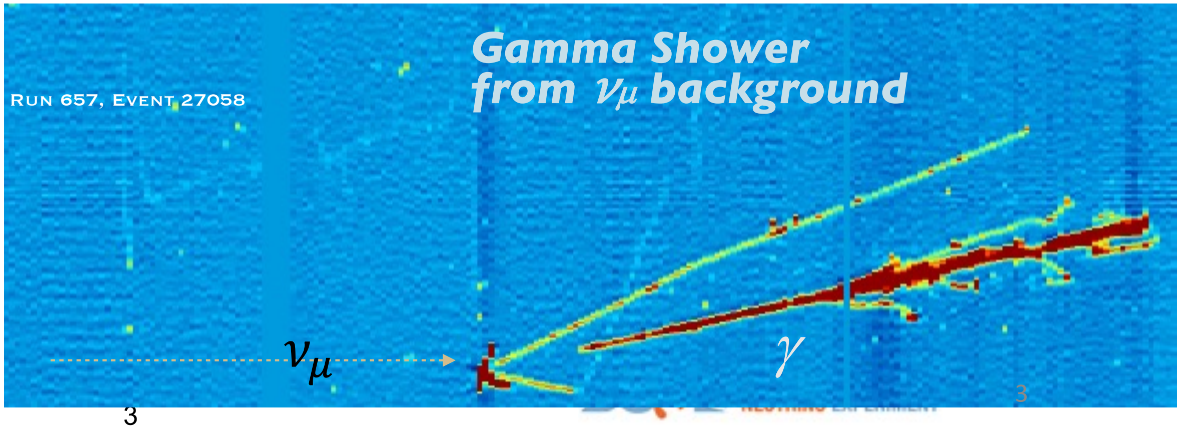


ν_e



Problem is you need to instrument ~50,000 m^3 with cm granularity and no dead material

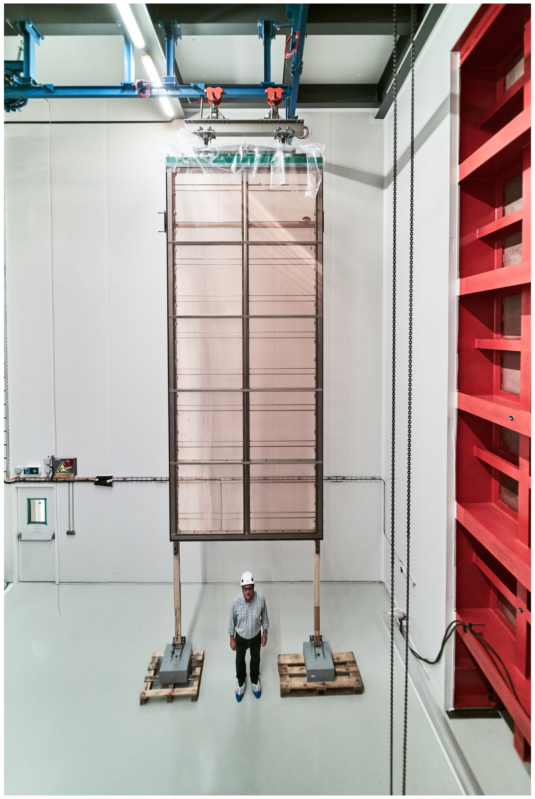
ν_μ



LAr TPC data volumes

- The first far detector module will consist of 150 **Anode Plane Assemblies (APAs)** which have 3 planes of wires with 0.5 cm spacing. Total of **2,560** wires per APA
- Each wire is read out by 12-bit ADC's every 0.5 microsecond for 3-6 msec. Total of **6-12k** samples/wire/readout.
- Around 40 MB/readout/APA uncompressed with overheads → **6 GB/module/readout**
- 15-20 MB compressed/APA → **2-3 GB/module/readout**
- Read it out ~5,000 times/day for cosmic rays/calibration → **3-4PB/year/module (compressed)**

(x 4 modules x stuff happens x decade) = ….



1 APA – 2,560 channels
150 of these per FD module

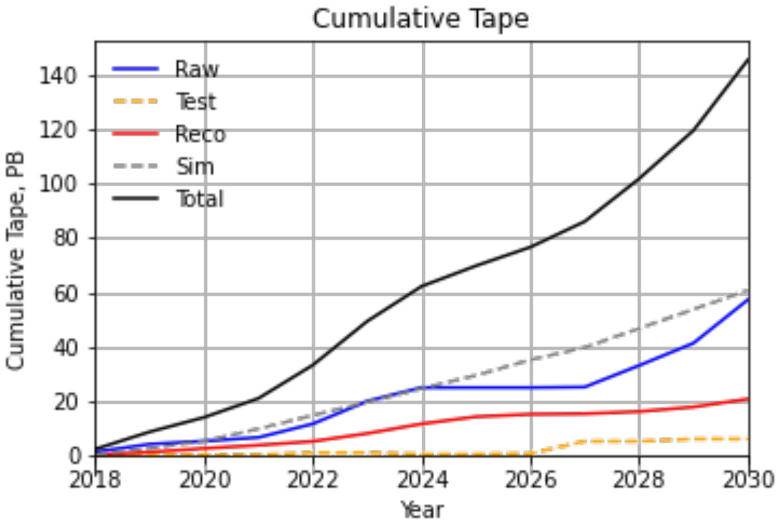
DUNE DEEP UNDERGROUND NEUTRINO EXPERIMENT

DUNE FD-Data for Supernova

Pack 150 5 ms APA readouts into a 6 GB file

Ship 20,000 time slices (x 4 modules)

This is ½ module for 100 s



CDR - Resource estimates to 2030

2 copies of raw data on tape (6 months on disk)

1 copy of "test" data stored for 6 months

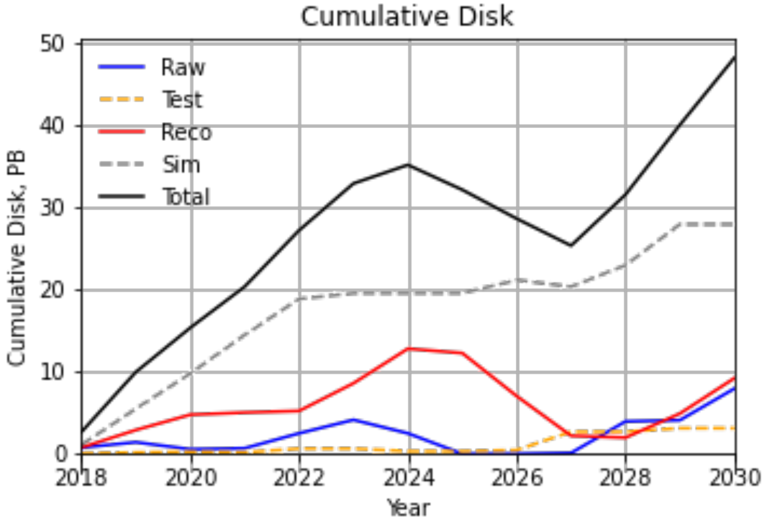
1 copy of reco/sim on tape

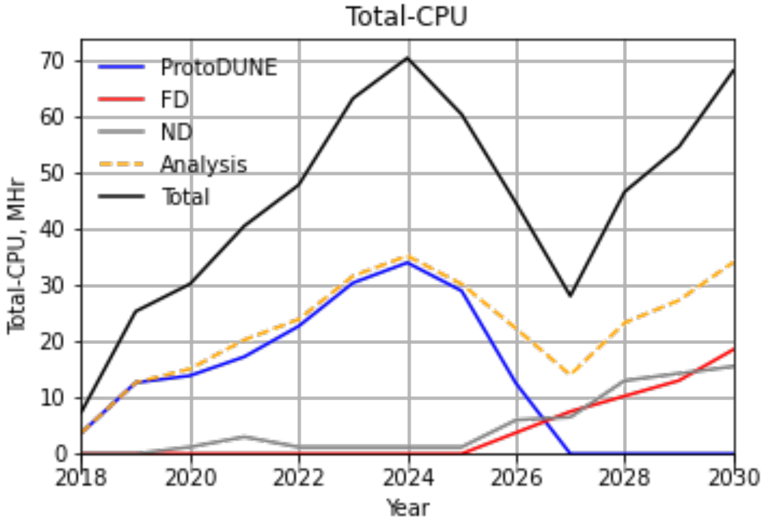
Currently assume 1 reco pass over all data and 1 sim pass/year

Assume reco/sim resident on disk for 2 years

Assume 2 disk copies of reco and sim

impose shorter lifetimes on tests and intermediate sim steps.





Longer term projections

VD assumed to be similar to HD, raw data may be larger due to longer drift.

2 copies of raw data on tape (6 months on disk)

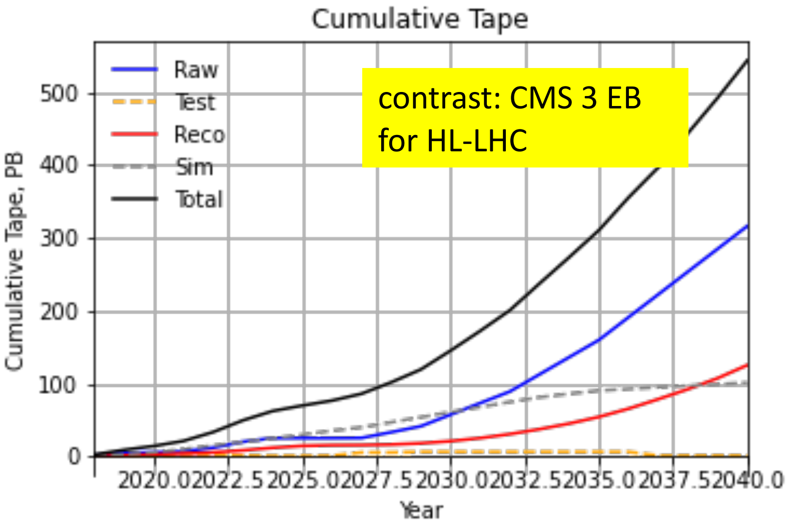
1 copy of "test" data stored for 6 months

1 copy of reco/sim on tape

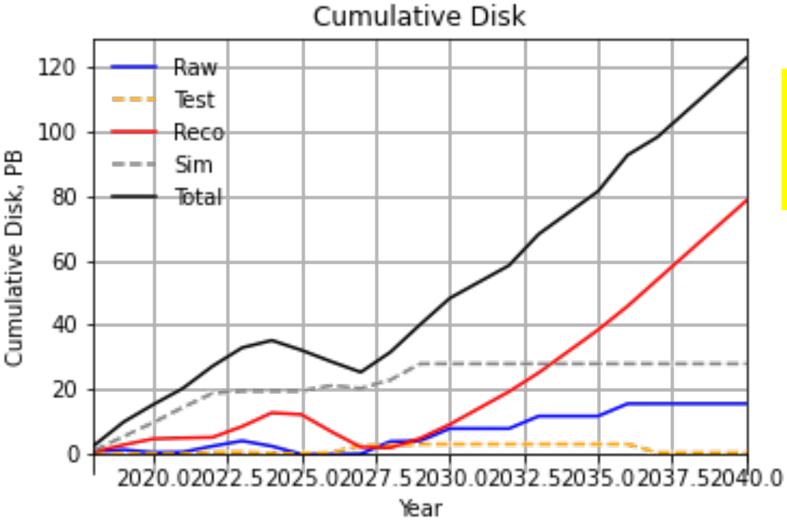
Currently assume 1 reco **pass over all data** and 1 sim pass/year

Assume reco/sim resident on disk for 2 years

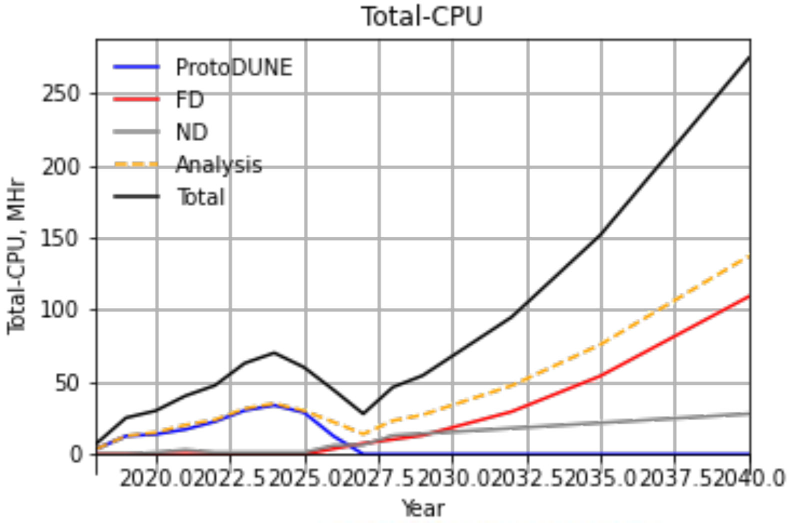
Assume 2 disk copies of reco and sim



contrast: CMS 3 EB for HL-LHC

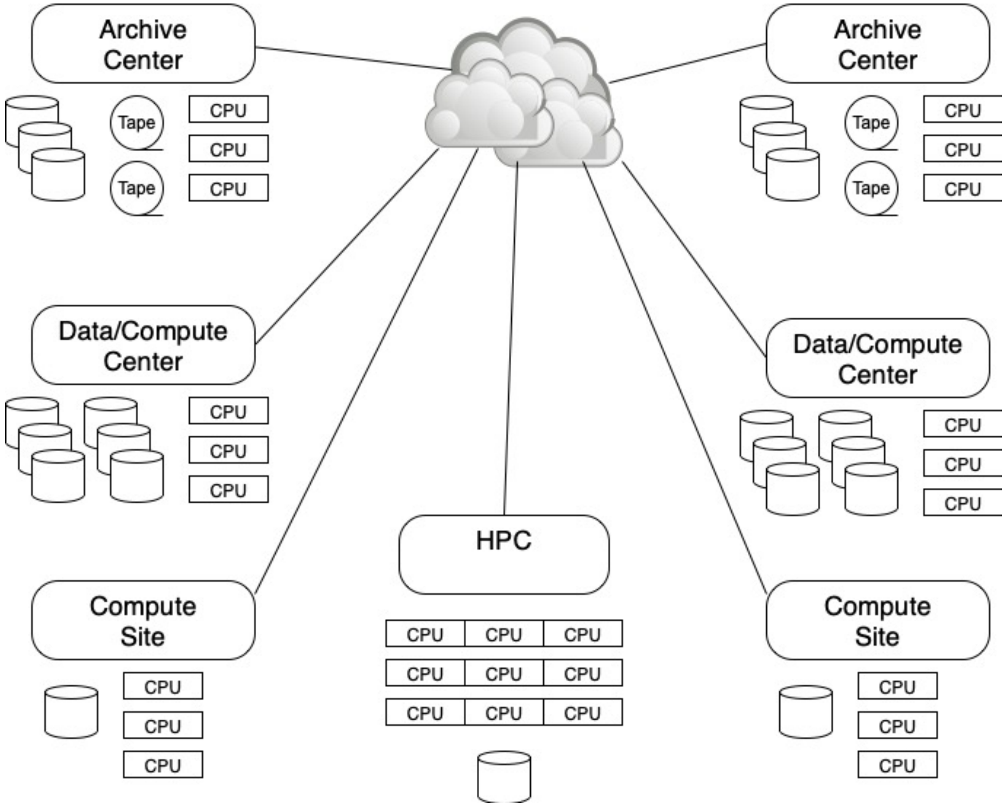


contrast: CMS 150 PB->2.2 EB in HL-LHC



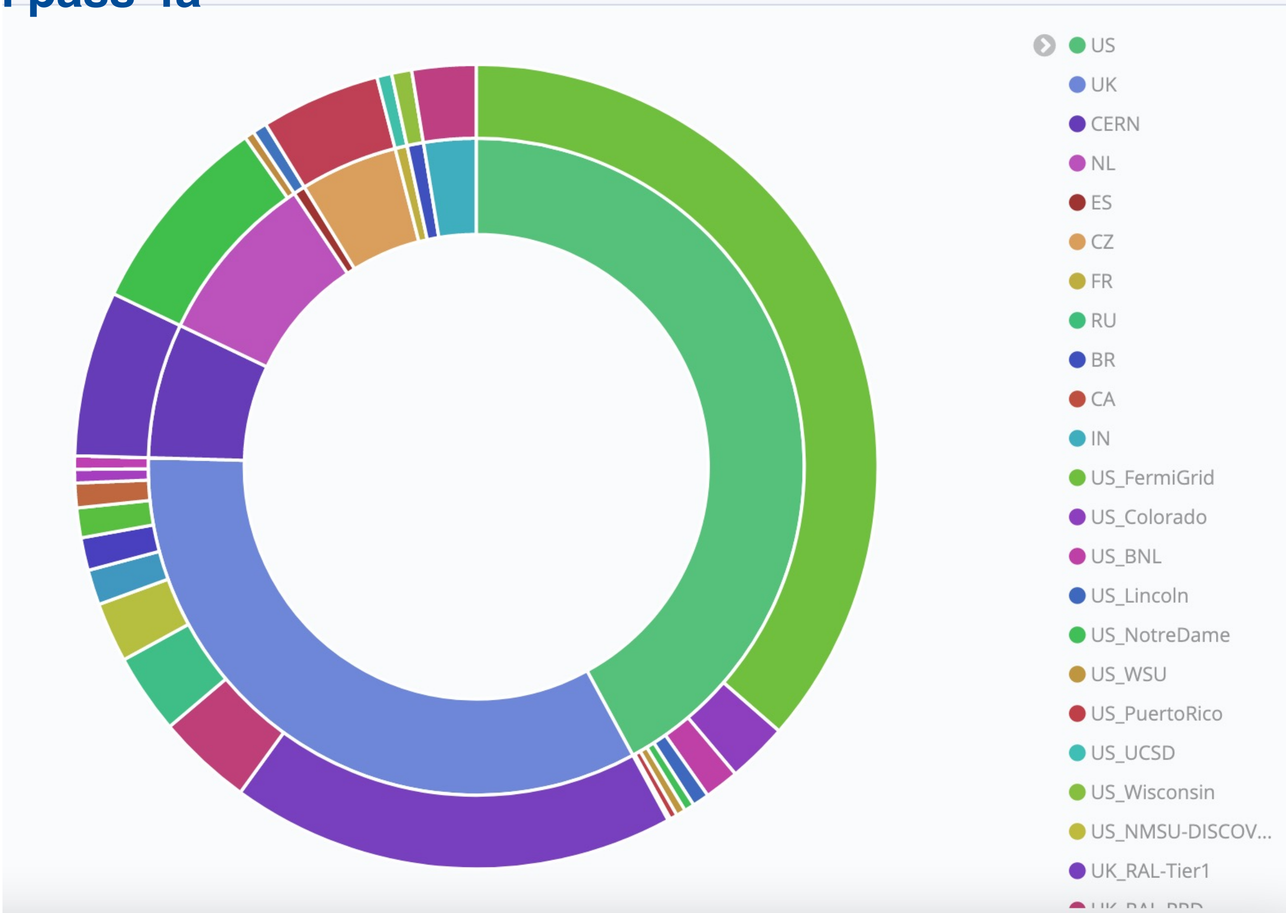
Distributed computing model

- Less “tiered” than current WLCG model → **DOMA**
- Collaborating institutions (or groups of institutions) provide significant disk resources (~1PB chunks)
- **Rucio** places multiple copies of datasets
- **We likely can use common tools:**
  - **But need our own contribution system**
  - **And may have different requirements for dataset definition and tracking**



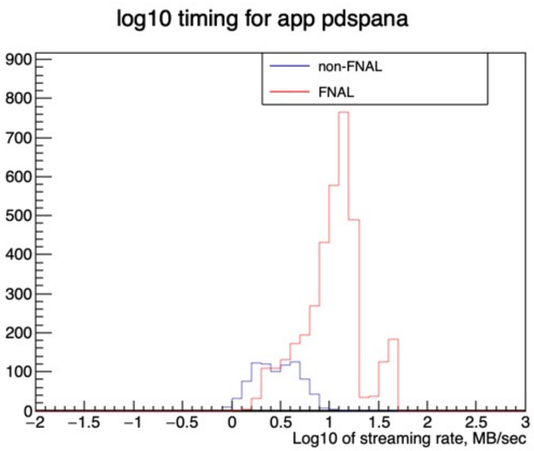
Where we are now: Production pass 4a

- This shows August 2021 contributions to production
- Production team now led by
  - Ken Herner and Elisabetta Pennacchio
  - Team from OSU (US), UNICAMP and UNIFESP (BR), York U. (CA) and Cambridge (UK)
- Each MC pass generates ~1.5 PB of detsim and 1.3 PB of reconstructed events.
- Actual reco data is only ~300 TB

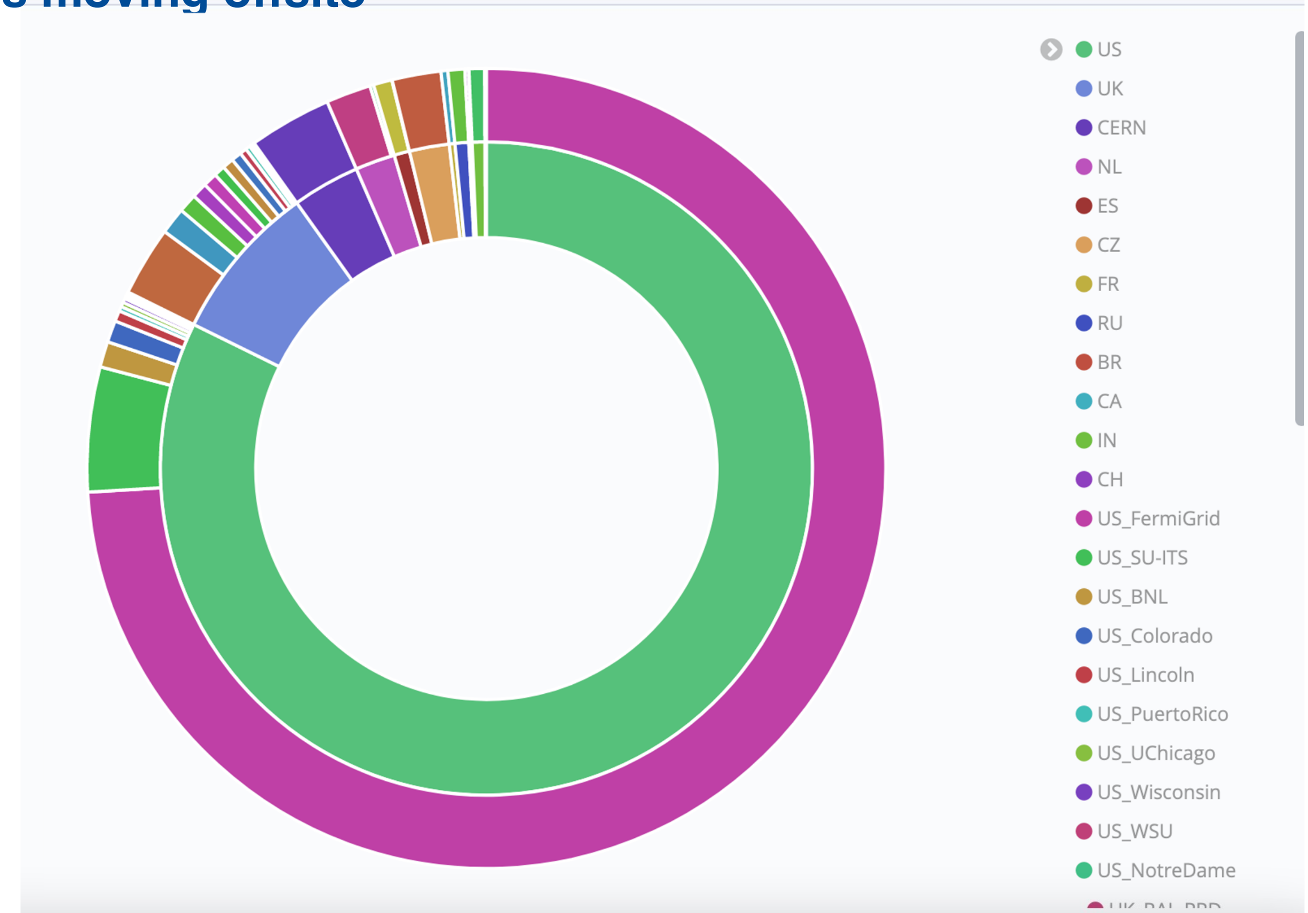


Where we are now: Analysis is moving offsite

- This shows August 2021 contributions to analysis
- US sites are contributing as are many sites worldwide

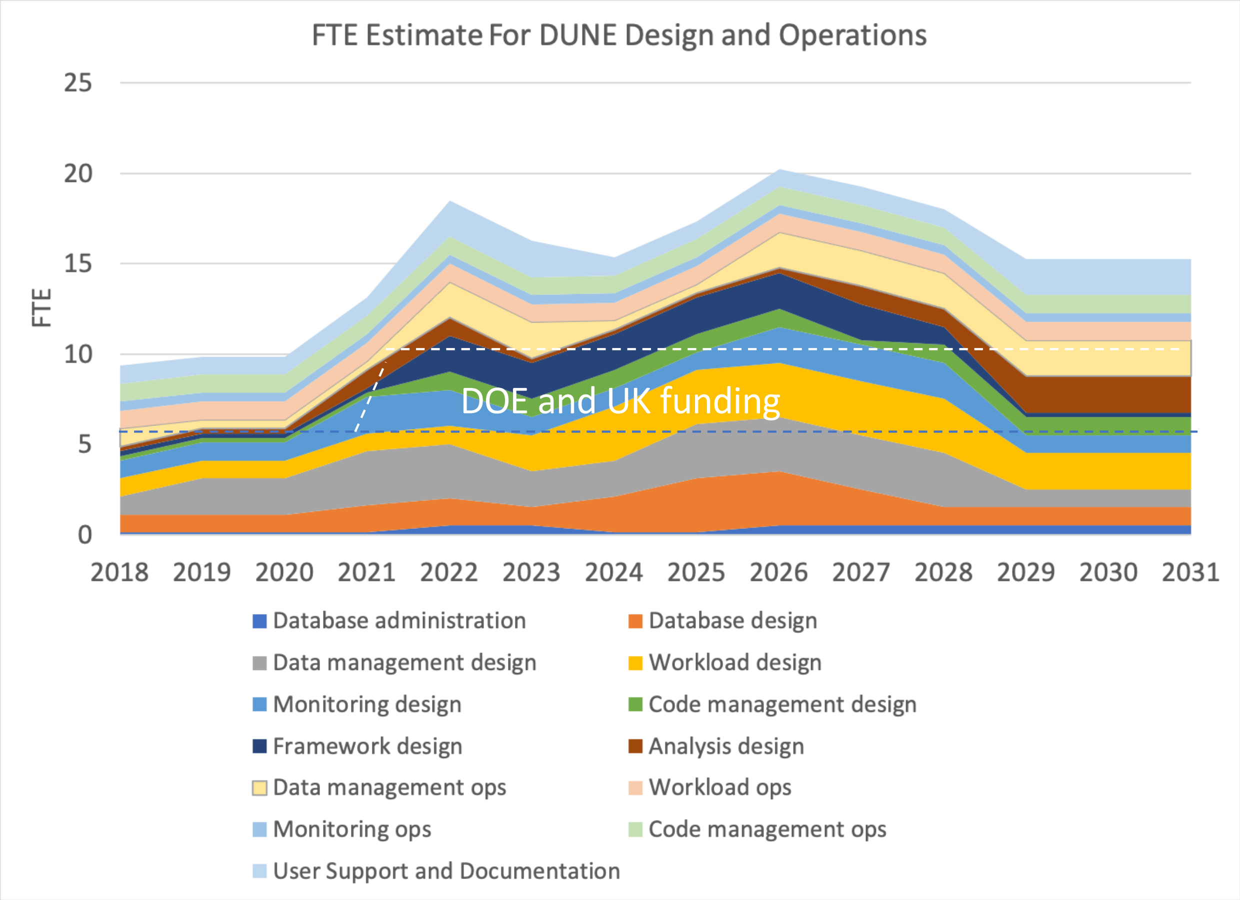


Fast jobs run slower far away



FTE estimate. Does not include shared facility (storage etc.) costs

- Some effort (mainly operations – pastels at top) can be trained collaboration physicists.
- Rest requires experts
- Until recently had 5 FTE experts (FNAL + collab), all in-kind contributions except UK DUNE funded personnel.
- DOE grant has added 4+1 postdocs and more lab FTE
- UK has added 1.5 FTE
- Expert need is greatest for ProtoDUNE 2 and pre-operations in 2024-2028. 5-10 FTE > 50% US


FTE Estimate For DUNE Design and Operations
FTE
DOE and UK funding
Database administration
Database design
Data management design
Workload design
Monitoring design
Code management design
Framework design
Analysis design
Data management ops
Workload ops
Monitoring ops
Code management ops
User Support and Documentation