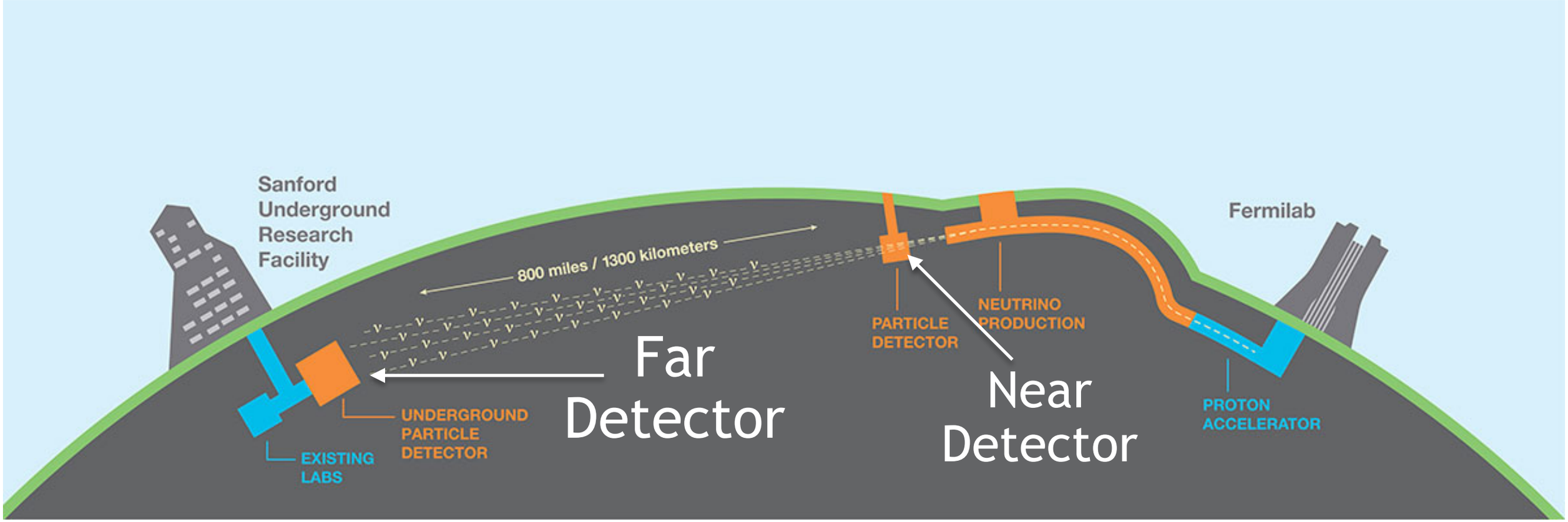
DUNE and the Fermilab Neutrino program

Andrew McNab

University of Manchester

DUNE / LBNF recap

- *"Make neutrinos at FNAL then detect some of them in South Dakota (and maybe supernovae and proton decays)"*
  - Construction underway. Start taking data later this decade
  - protoDUNE experiments running at CERN now and taking data
- DUNE is part of the **Long** Baseline Neutrino Facility (LBNF)

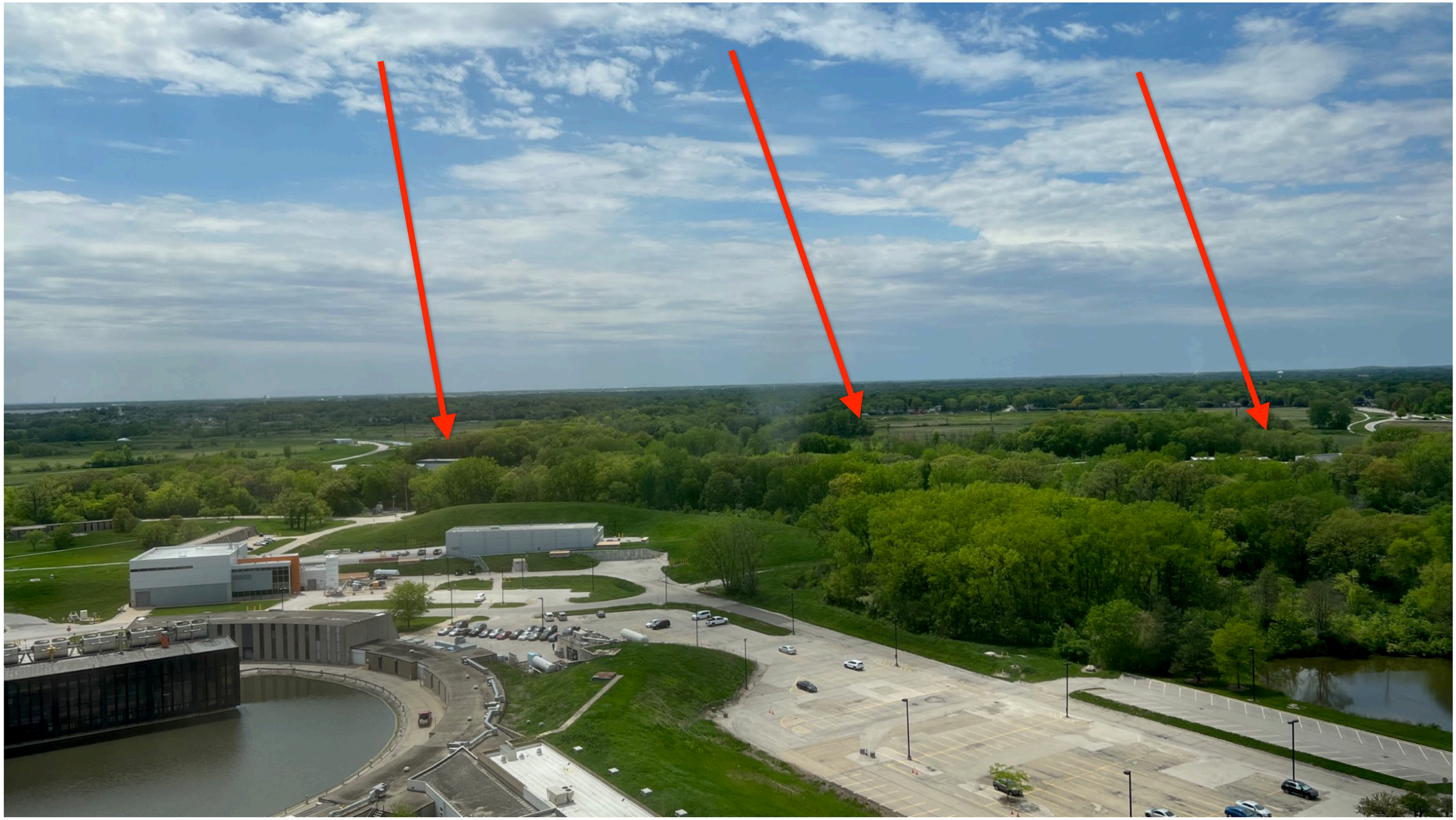


Fermilab Wilson Hall (in 1970s)

Floor 15 of Wilson Hall (May 2022)

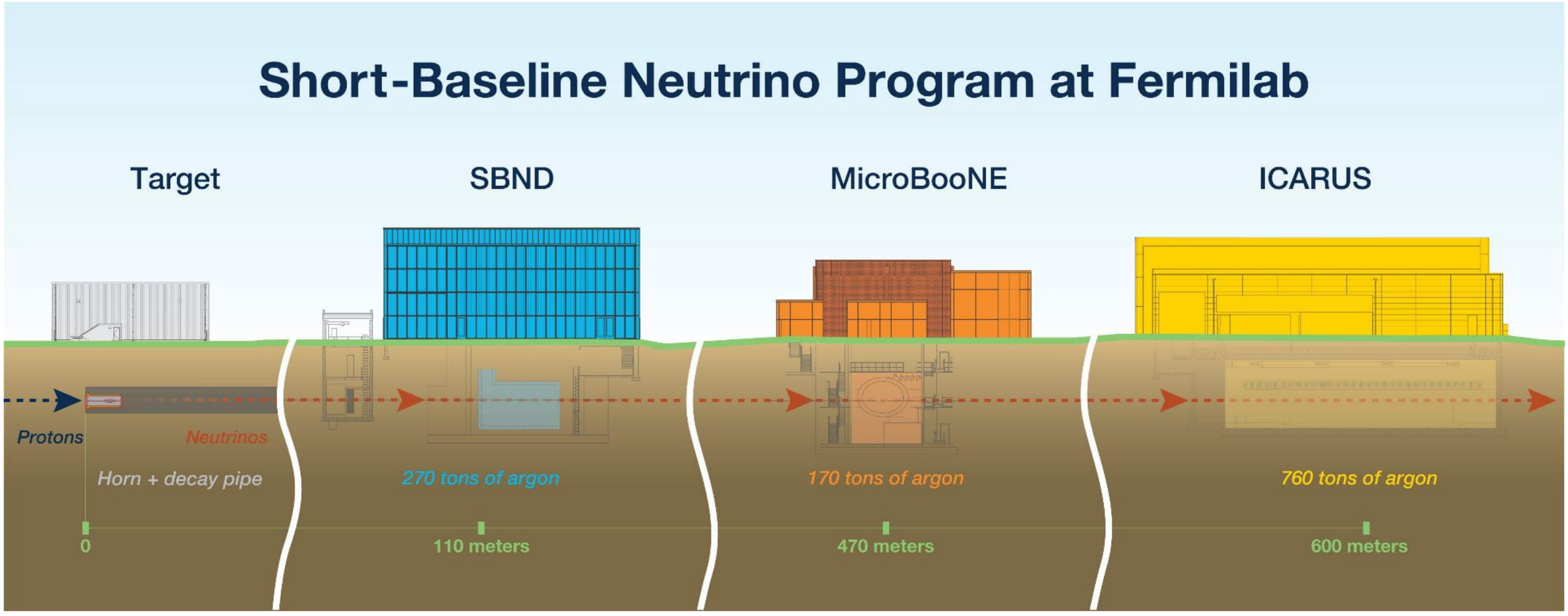
South Dakota is that-a-way!

Short Baseline neutrino line



Short Baseline neutrino line

- MicroBooNE (uboone) has finished data taking but is still doing lots of processing
- ICARUS is starting up and SBND coming soon



DUNE Computing organisation

- There are weekly computing ops meetings at 2:30pm on Mondays
  - Anyone here is welcome to join
  - Ask me for details. Also a dune-computing-sites@fnal.gov list.
- DUNE has a Computing Consortium, composed of DUNE members
  - Covers computing as in grid etc plus some shared software framework, databases etc
  - UK computing contribution comes from DUNE UK Construction Project with dev effort funded at Edinburgh, Manchester, RAL-PPD
  - UK dev effort focussed on Rucio / data management (Ed) and Workflow (Man + PPD)
- If you see problems with DUNE jobs or storage use, please either use the ops meeting or contact us
- The DUNE VO also exists in GGUS with its own support unit

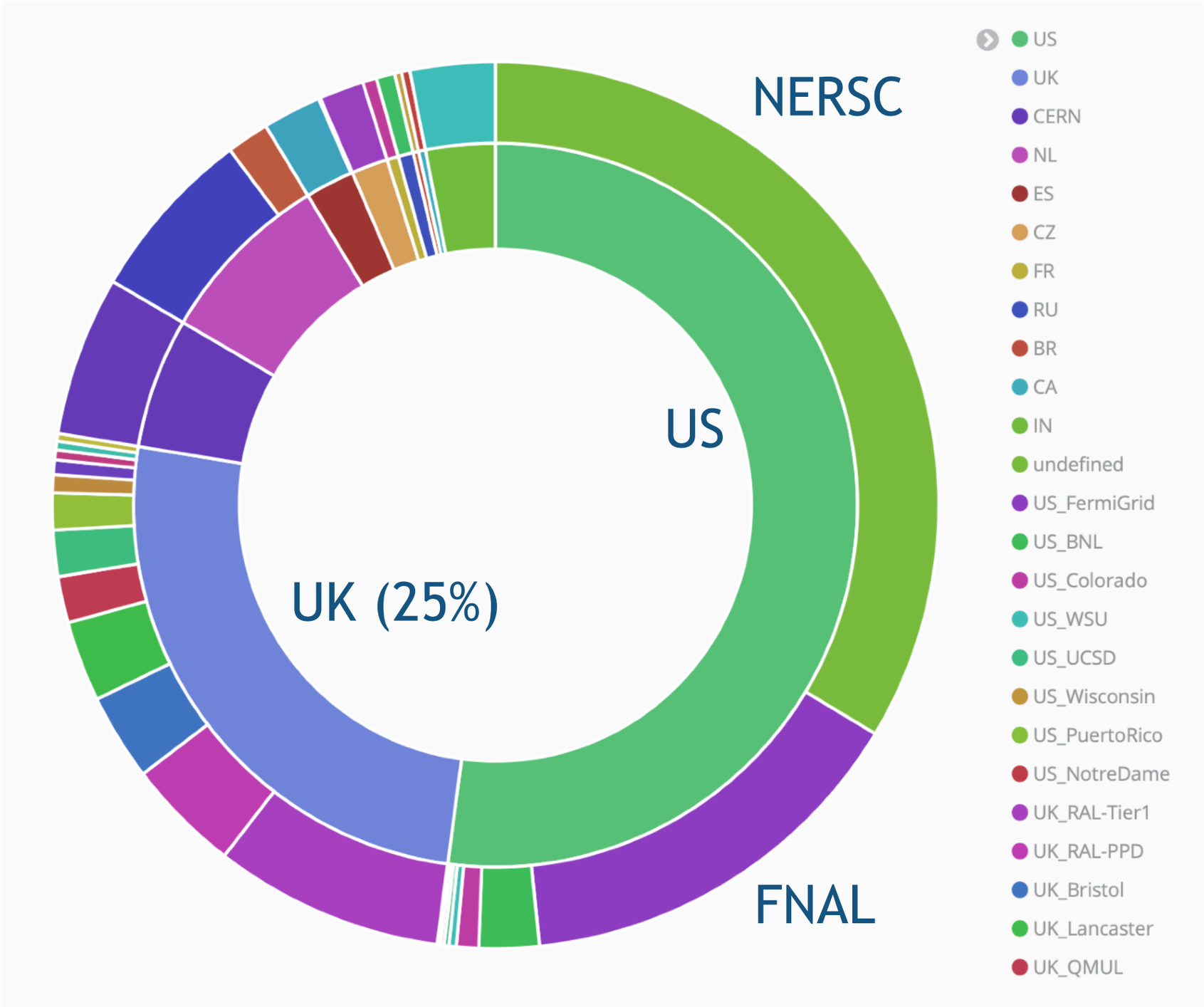
DUNE use of GridPP sites

- Pilot jobs submitted by OSG pilot factories at UCSD or CERN
  - HTCondor-CE and ARC are ok
  - DUNE exclusively uses dune VO; MicroBooNE, SBND, ICARUS use fermilab VO
- Pool of user jobs managed by glideInWMS/HTCondor at Fermilab
  - HTCondor inside the pilot jobs does the “masonry” of matching multiple user jobs into the space of processors, memory and time obtained by the pilot job
- Single processor user jobs often need 4GB or sometimes more
  - Site won’t get many matches if pilot job slots are less than “4GB” (4096MiB to be safe, not 4000)
- For storage, GridFTP, xroot, WebDAV are supported but xroot has better support for streaming in the applications used by DUNE (lar/art)

Efficiency issues

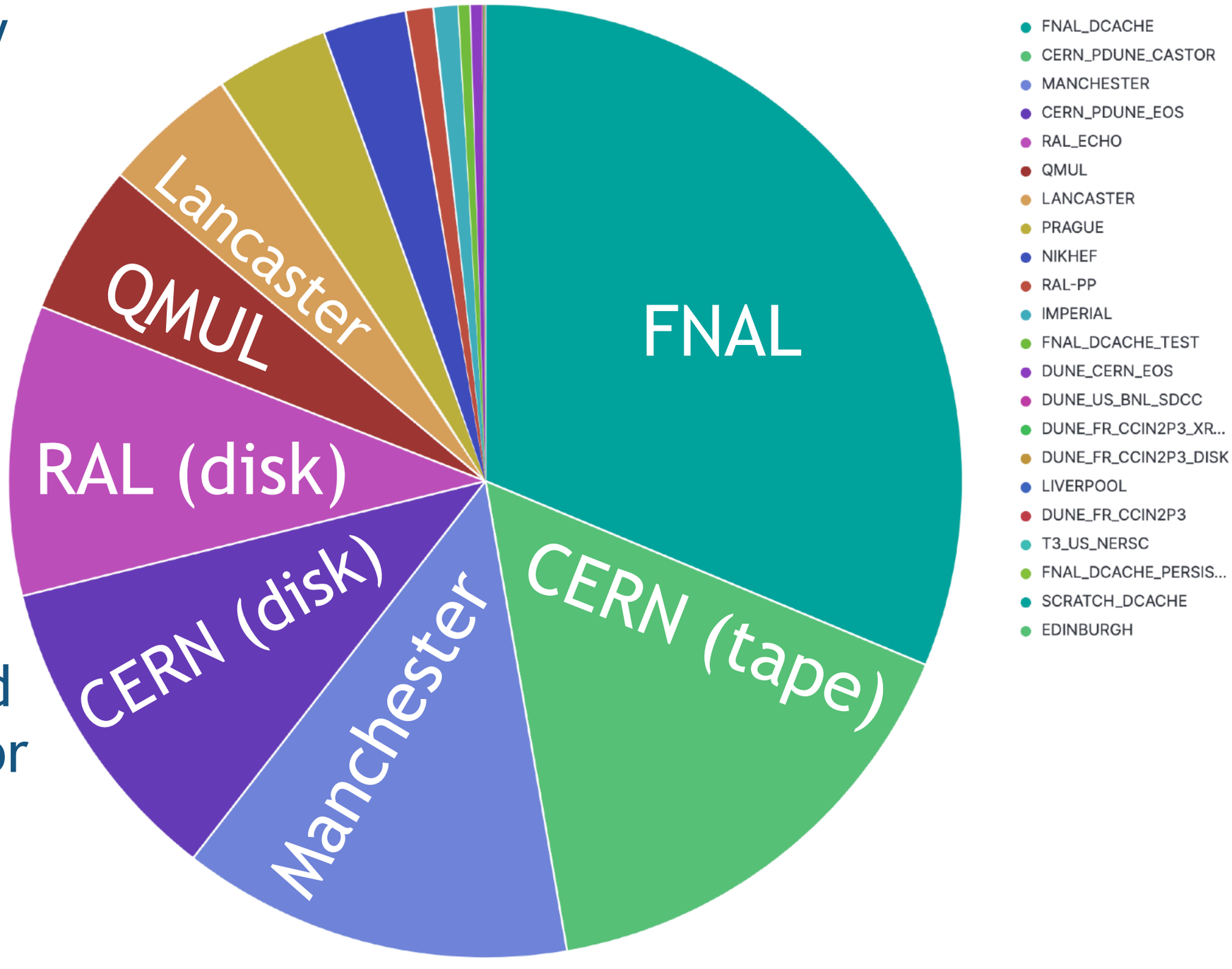
- In multiprocessor pilot job slots, to get enough memory, jobs may end up occupying multiple processors
  - This is part of the way glideInWMS/HTCondor works
- DUNE has relied on SAM for data management and assignment of files to jobs, and POMS for creating the jobs
  - Job starts and asks SAM for the next unprocessed file to work on
- This framework has a limited ability to assign jobs to sites where the input data is nearby
  - So can easily find yourself streaming data from Fermilab or CERN, with corresponding loss of efficiency
  - Work is underway in the new Workflow System to address this, and to base job matching on locations of replicas of unprocessed files
- The other neutrino experiments still use the old SAM etc framework

DUNE prod jobs: 2022 Q1/Q2

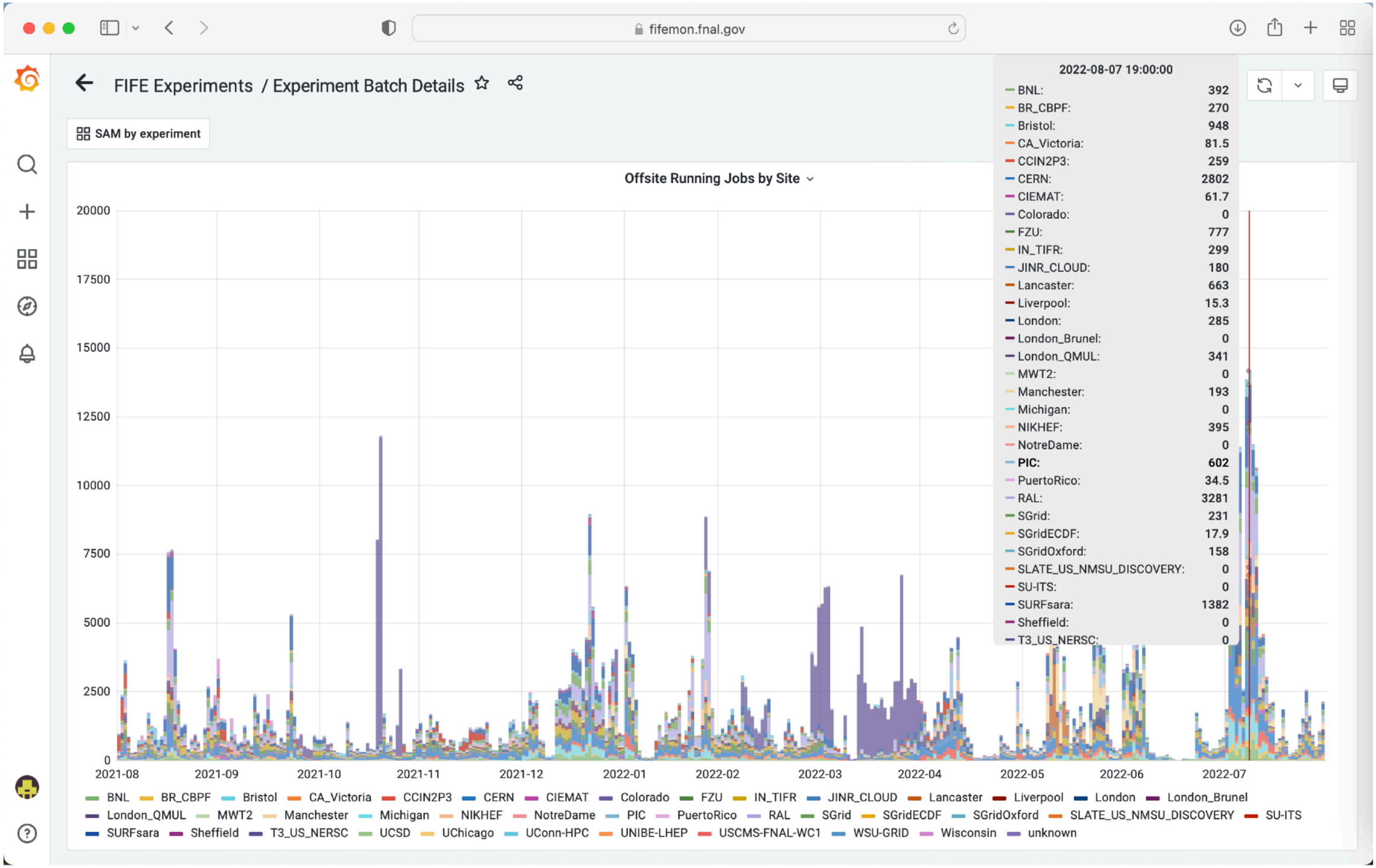


DUNE files in Rucio (July 2022)

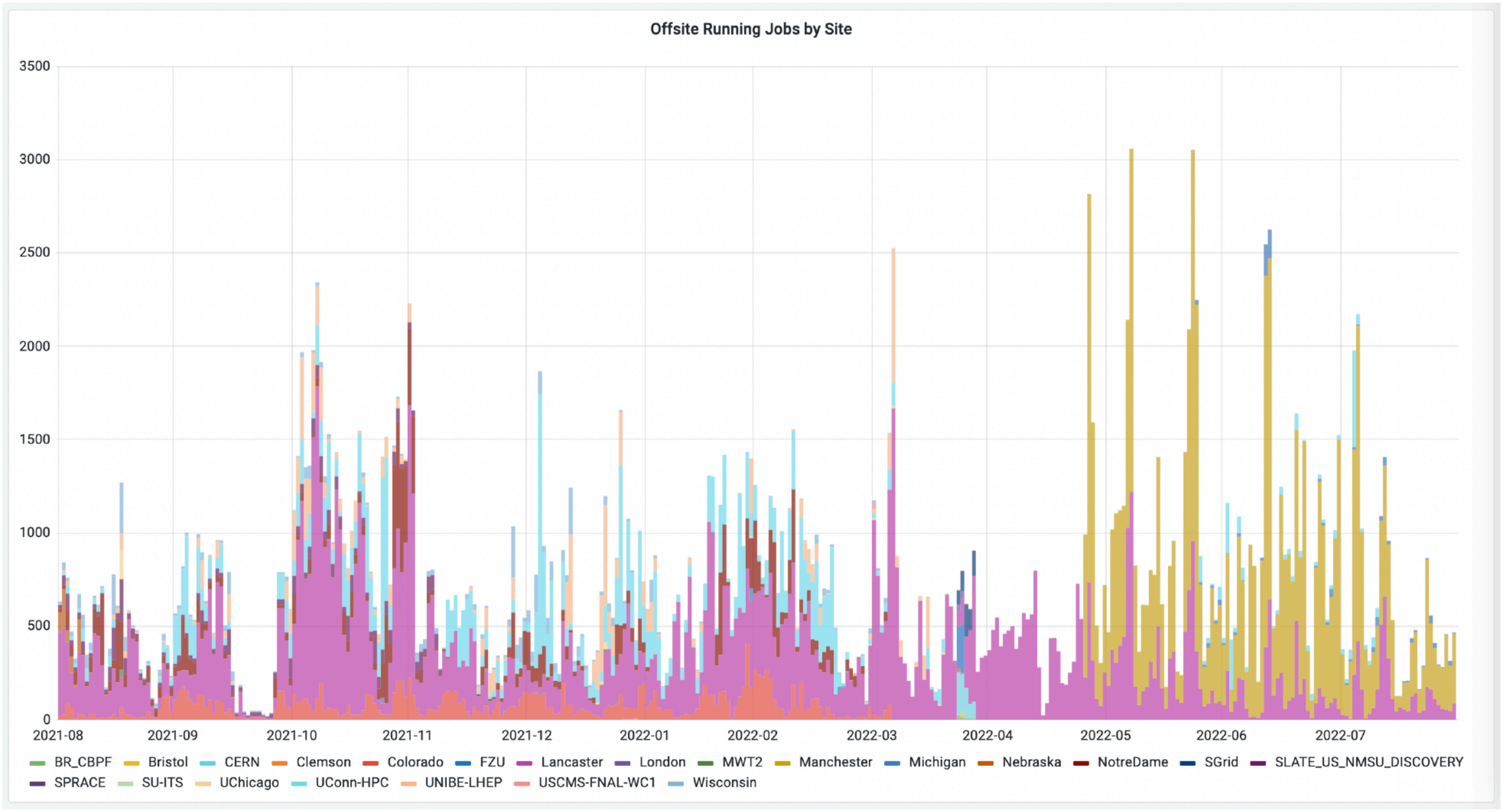
- Storage used by DUNE files known to Rucio in early July
  - So NOT just the amount of space provided by the sites!!!
- Some space had been cleared for the Data Challenge, esp at CERN



DUNE during the year



MicroBooNE during the year



Looking forward

- DUNE is working on causes of job inefficiency
  - Better matching of jobs to storage in particular
- MicroBooNE (uboone) is already using significant resources and wants to increase
  - And then SBND which has a lot of UK physicists, and maybe ICARUS too
- DUNE is already using the IRIS RSAP procedure to get a resource allocation
  - The Fermilab neutrino experiments in the UK are also planning to make a formal request
- IRIS support for the neutrino sector, fulfilled by GridPP using the IRIS capital money, further broadens the pool of physicists aware of / using / happy with ☺ GridPP beyond the LHC

2022 DUNE Data Challenge

- Split into phases
- Phase 1 earlier this summer demonstrated funnelling data from the protoDUNE DAQ machines at CERN to Fermilab, registering them in Rucio and in DUNE's metadata catalogue.
- Phase 2a is about to start and will test Rucio and the new workflow system in processing data stored at Fermilab, CERN, several UK sites, CC-IN2P3 and some other European sites.
- Phase 2b will be a repeat of 2a once some planned changes to the Fermilab job submission system are available, and anything arising from Phase 2a.
- UK sites should not see unusually large amounts of DUNE jobs during this testing, but we will flag up anything important through the usual channels.

DUNE and WLCG/GridPP

Package your work up in a standard way.

(Chicago O'Hare, May 2022)

DUNE and WLCG/GridPP

Package your work up in a standard way.

Lots more delivery mechanisms become available!

(Chicago O'Hare, May 2022)

Summary

- GridPP is making a significant and visible contribution to the CPU and storage used by DUNE
  - Underpinned by STFC funding for hardware via IRIS
- The DUNE UK Construction Project effort is working on making DUNE's exploitation of resources outside Fermilab more efficient
- We are already seeing growing demand from MicroBooNE and expect more Fermilab neutrino experiments to follow
  - The hope is that this benefits from DUNE's experience with the UK sites and IRIS resource allocation
- Getting this to work is good for everyone …