DUNE and the fermilab VO

Andrew McNab, University of Manchester

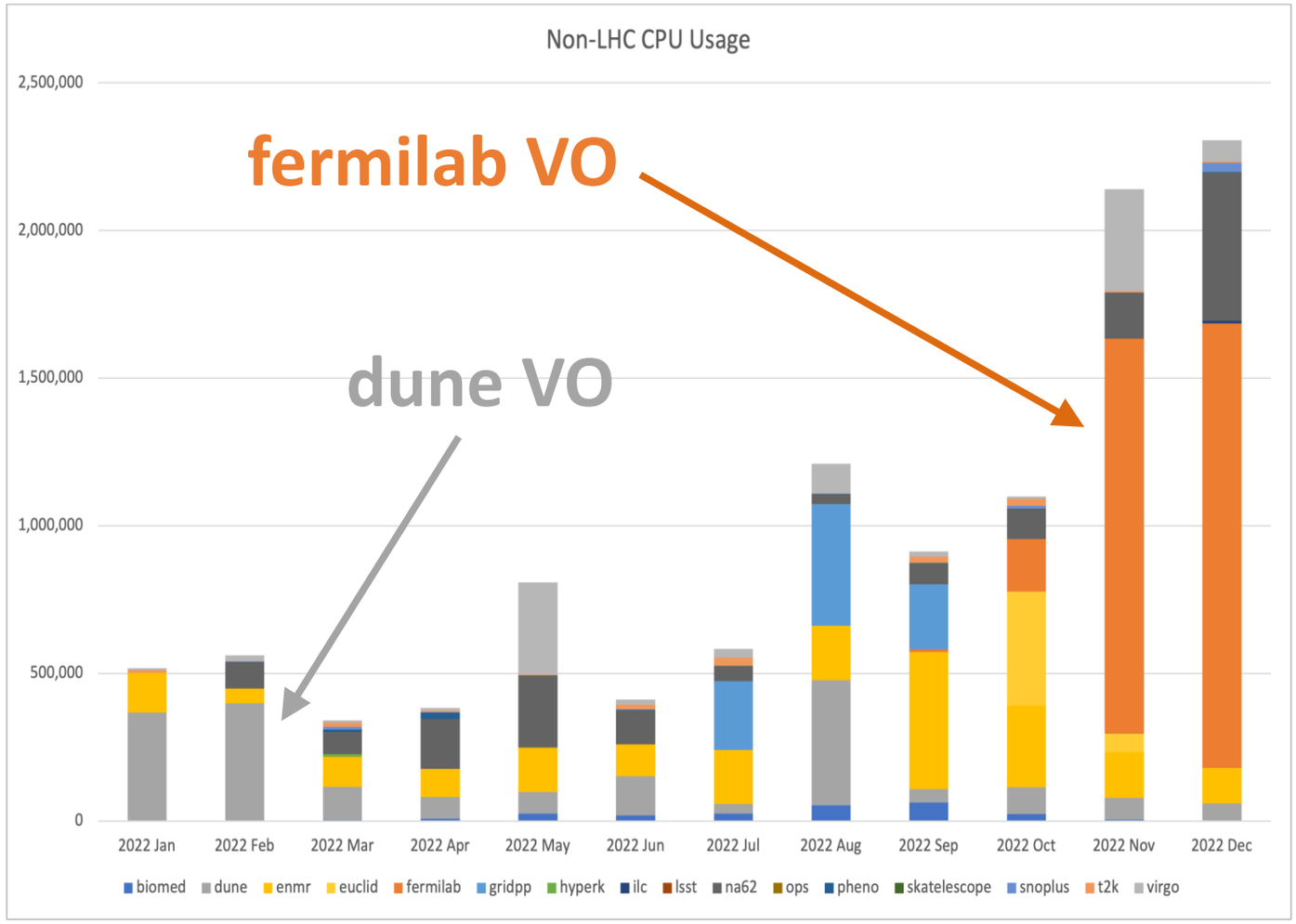
DUNE and the fermilab VO

- Two main topics
  - Where are we with DUNE?
  - What *is* the fermilab VO doing?

What *is* “fermilab” doing?

- Prompted by a conversation at the RAL Tier-1 Resource Review meeting last month
- In Alastair’s talk about experiment resource usage, there was this slide
- It naturally raised the question about what the fermilab VO is actually doing
- I gave a very rough answer about the different sub-VOs
- This talk is intended to give proper answers

Non-LHC CPU Usage



Fermilab VOs

- dune VO
  - This is run in the same way as an LHC experiment
- fermilab VO
  - An umbrella VO for 44 different sub-VOs for smaller experiments
  - For GridPP, the VOs using significant UK capacity are
    - /fermilab/uboone
    - /fermilab/icarus
    - /fermilab/sbnd
    - /fermilab/gm2
  - There are others, but at UK sites they are just running a few tests, often months apart
- Always one or more user HTCondor jobs running inside OSG pilot jobs

FCRSG

- The following tables and charts are from the Fermilab Computing Requests Scrutiny Group meeting in February

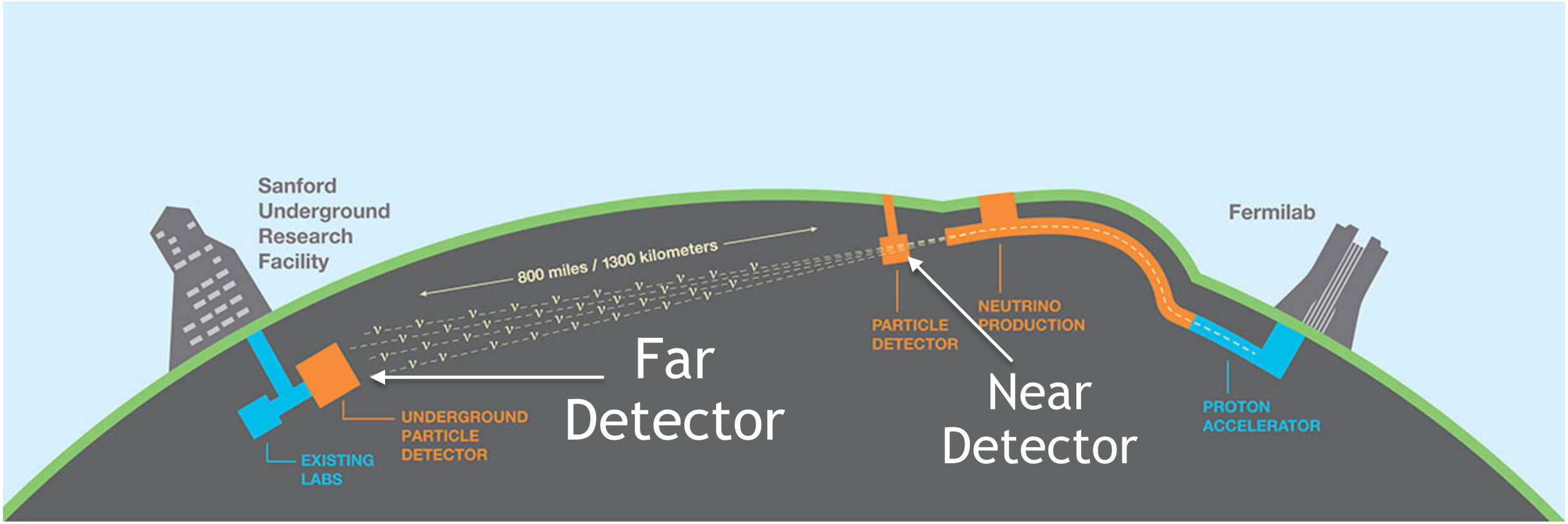
- This is an annual review of the major experiments' requirements, their plans for the future, and presentations from Fermilab about the status of the various services

- I'm only covering the experiments with significant use of GridPP

- Public slides: https://indico.fnal.gov/event/57596/

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)



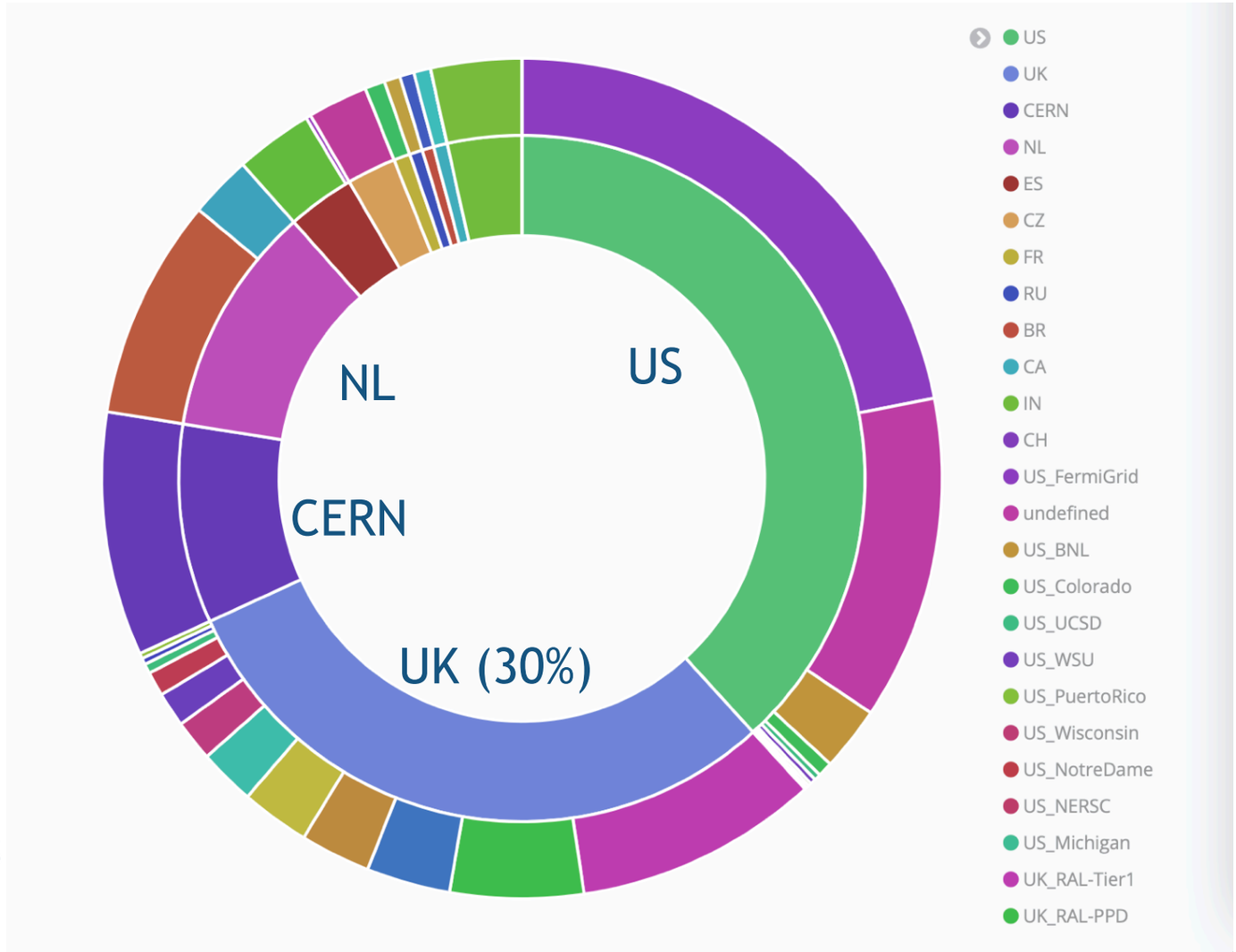
DUNE GridPP use in context

DUNE production jobs for 2022

Larger memory requirements enabled at UK sites to attract jobs

Broader set of sites being targeted

2022 Data Challenge in Q4 was testing systems with automated matching of jobs to best site



7

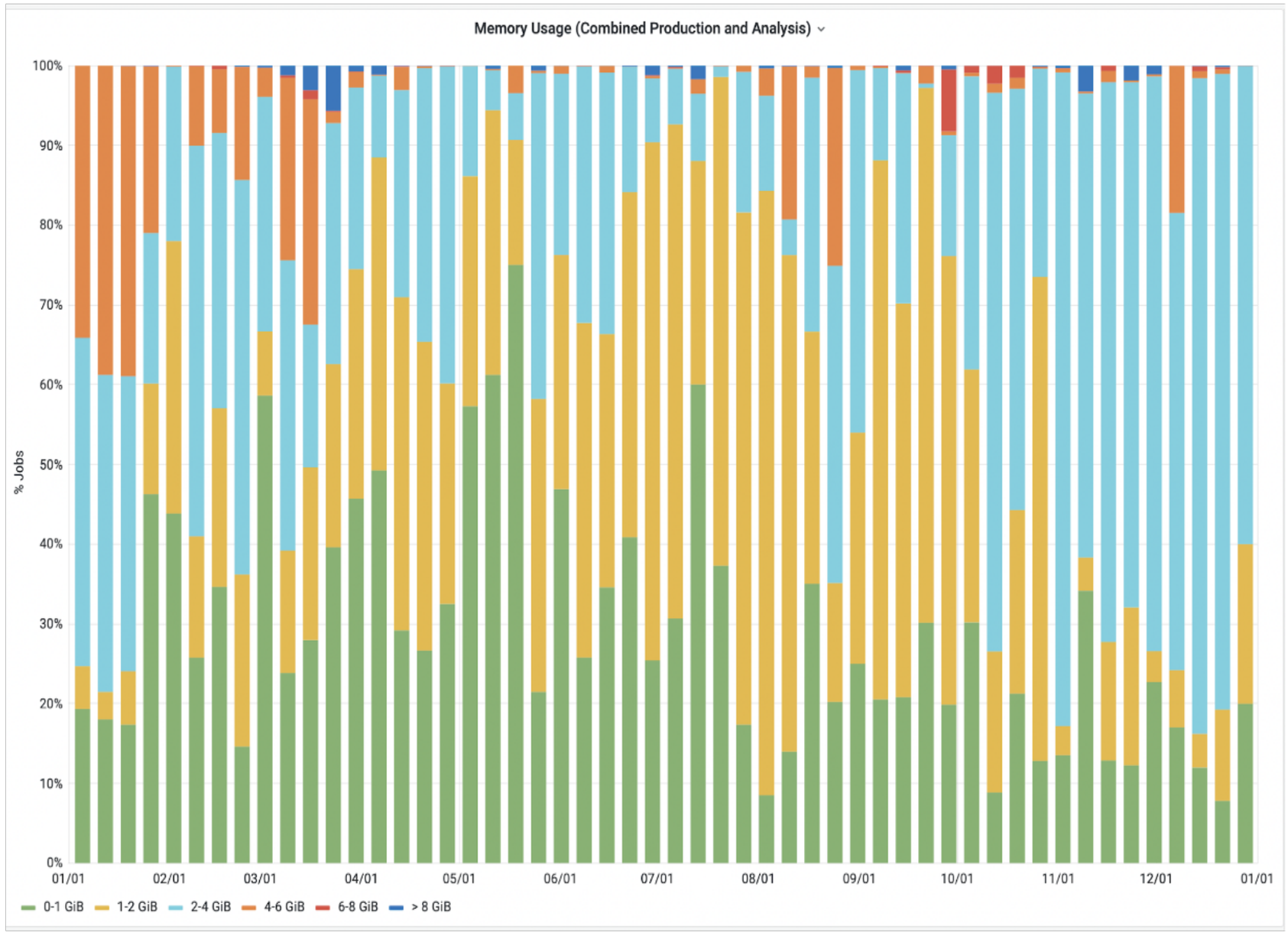
DUNE CPU requests

- Millions of 1processor,2GB wall hours per year
- Major processing goal is data from protoDUNE experiments at CERN
- Filling of the two cryostats delayed by liquid argon shortages in Europe
  - Each need 800 tonnes of LAr, ~1% of European LAr annual supply

	2019	2020	2021	2022	2023	2024	2025	2026	2027
Requested		25 M (FNAL) 36 M (Total)	29 M (FNAL) 40 M (Total)	49 M (FNAL) 98 M (Total)	56 M (FNAL) 139 M (Total)	61 M (FNAL) 153 M (Total)	70 M (FNAL) 175 M (Total)	55 M (FNAL) 138 M (Total)	33 M (FNAL) 83 M (Total)
Actual Used		29 M (FNAL) 40 M (Total)	36 M (FNAL) 56 M (Total)	38 M (FNAL) 58 M (Total)	N/A	N/A	N/A	N/A	N/A

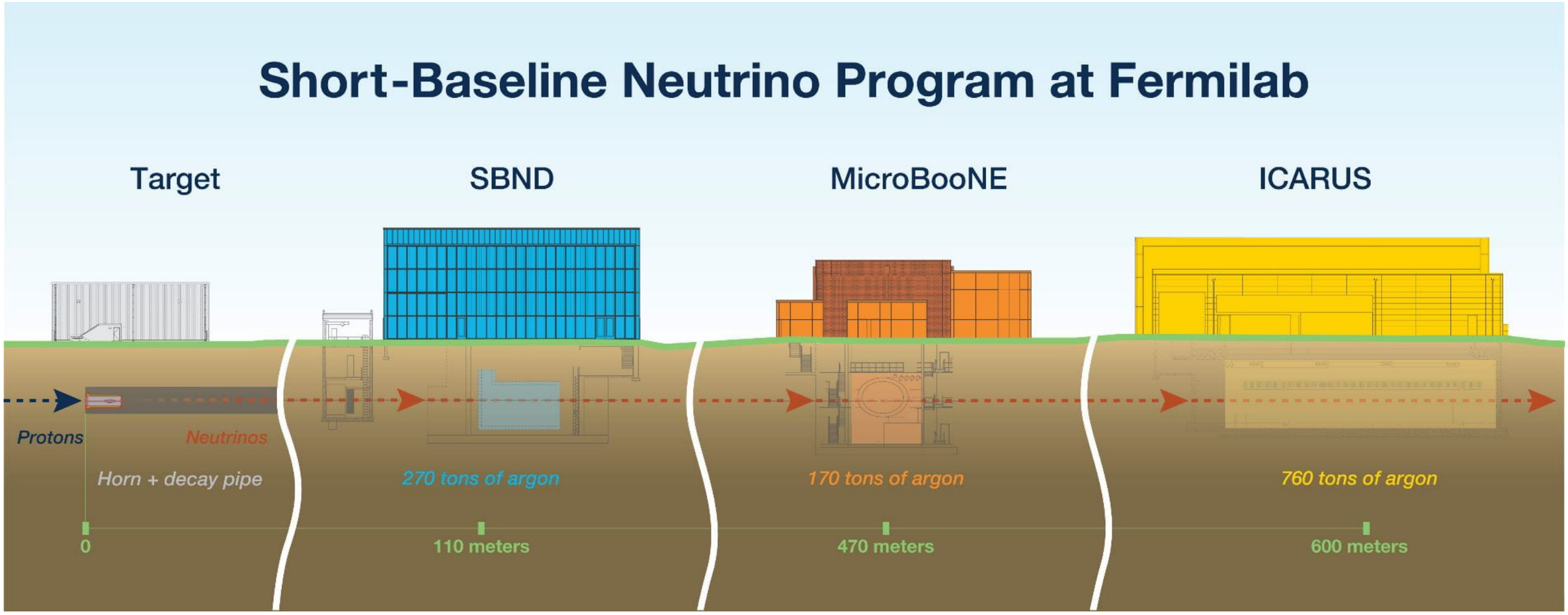
DUNE memory profile

- Significant amounts of 2-4GB,1processor jobs during DUNE data challenge at the end of the year
- DC was to prepare for protoDUNE data processing which needs ~4GB single processor jobs
- Typically several jobs inside 8processor pilots with 16GB to 32GB



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



Short baseline CPU requests

- Millions of 1processor,2GB wall hours per year

MicroBooNE (Fermilab)

	2019	2020	2021	2022	2023	2024	2025	2026	2027
Requested	20	54	43	30	30	25	20	10	
Actual Used	29.1	40.6	30.0	26.4	N/A	N/A	N/A	N/A	N/A

MicroBooNE (Off site)

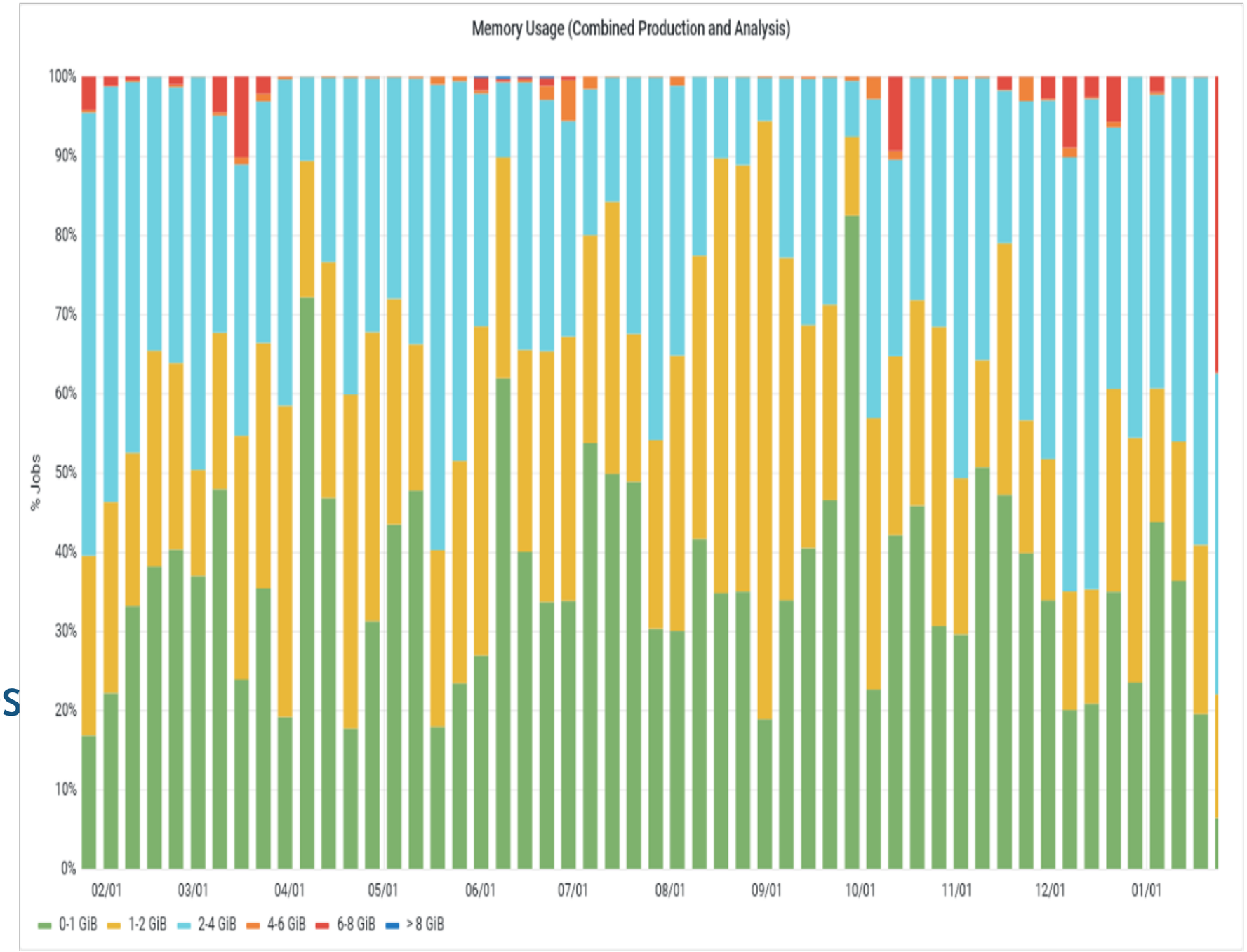
	2019	2020	2021	2022	2023	2024	2025	2026	2027
Requested	0	0	0	0					
Actual Used (OSG)	3.6	6.3	6.8	10.7	N/A	N/A	N/A	N/A	N/A

ICARUS/SBND (Fermilab)

	2019	2020	2021	2022	2023	2024	2025	2026	2027
Requested	13.2	16.5	19.5	27	**29**	**53±15**	**48±15**	**58±15**	**74±15**
Actual Used	4.46	7.7	11.4	15	N/A	N/A	N/A	N/A	N/A

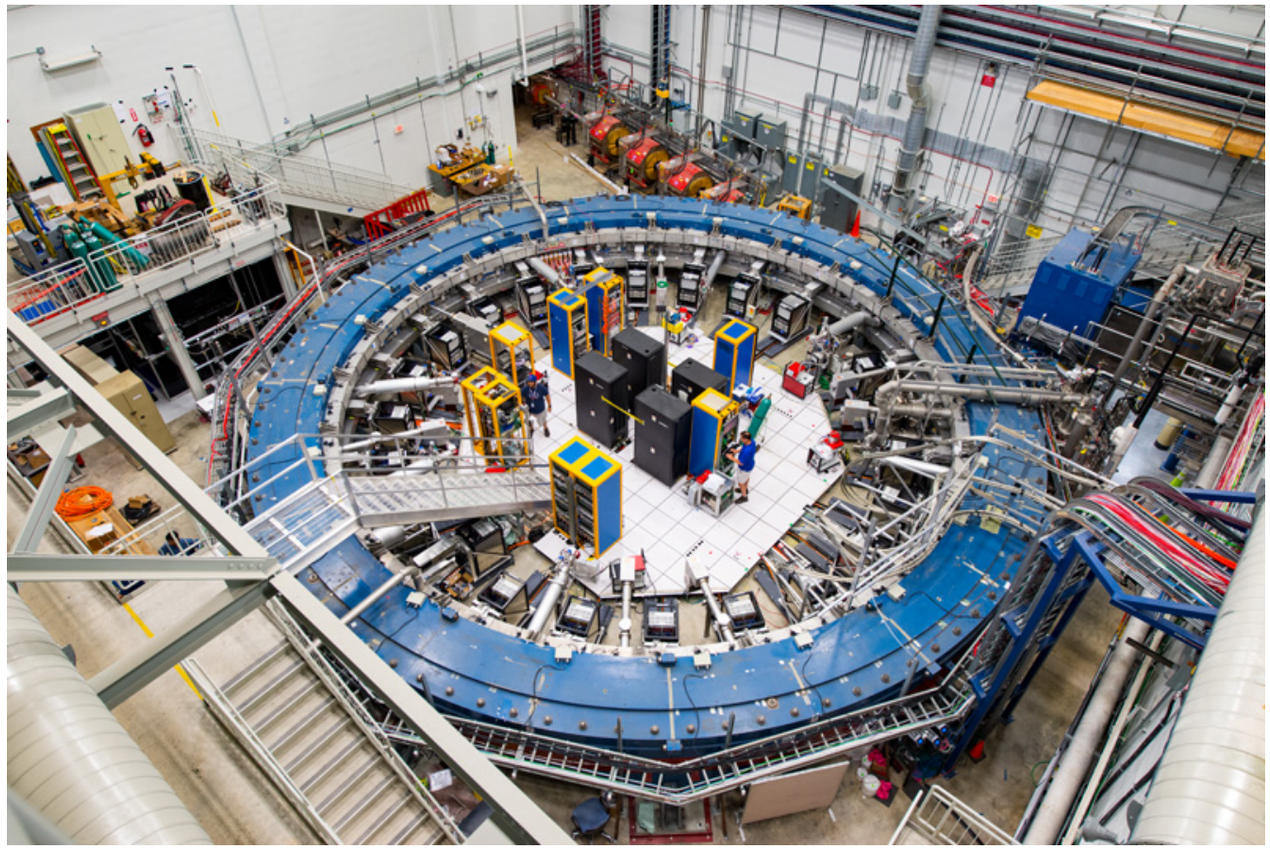
MicroBooNE memory profile

- A bit better than DUNE as not as many 2-4GB jobs
- Very similar workload framework to DUNE
- Typically 8cpu pilots with 16GB
- GlideInWMS/ HTCondor manages user jobs inside the pilots



Muon g-2 (gm2)

- Not a neutrino experiment
- Storage ring for muons, allowing observation of their precession
- Is the muon magnetic moment consistent with Standard Model predictions?
- Started data taking in 2018 and expected to finish data taking in 2023
- So far there is currently a 4 sigma discrepancy



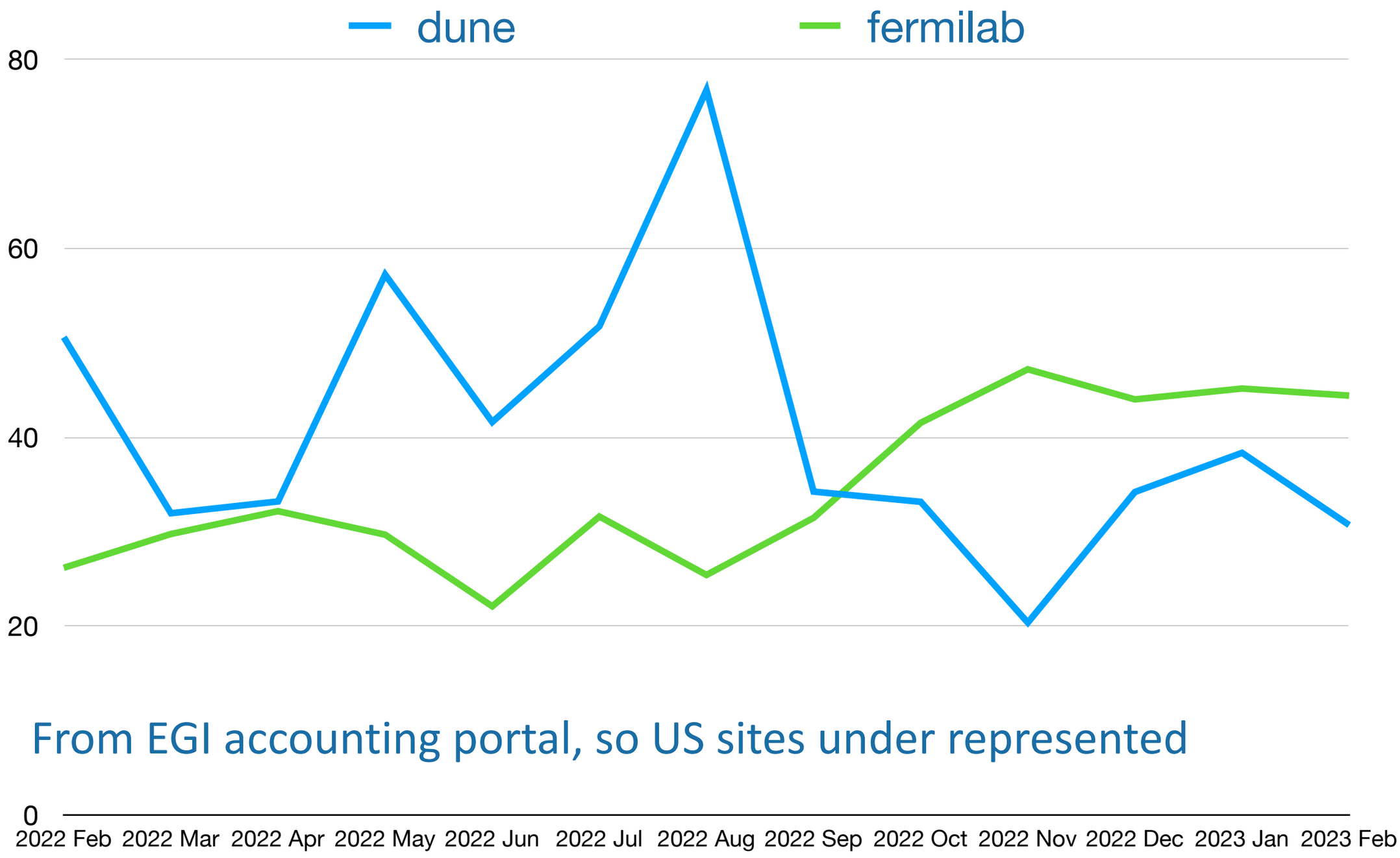
Muon g-2 (gm2)

- Millions of 1processor,2GB wall hours per year, for Fermilab (top) and offsite (bottom)
- Data taking stops this year and requirements stop in 2025

	2019	2020	2021	2022	2023	2024	2025	2026	2027
Requested	21 all platforms	33 all platforms	45 all platforms	60 all platforms	15 analysis/ sim + 20 production = 35 FermiGrid	15 analysis/ sim FermiGrid	10 analysis/ sim FermiGrid		
Actual Used	29	39	25	22	N/A	N/A	N/A	N/A	N/A

	2019	2020	2021	2022	2023	2024	2025	2026	2027
Requested	21 all platforms	33 all platforms	45 all platforms	60 all platforms	3 production + analysis	3 analysis/ sim	2 analysis		
Actual Used	??	3.3	4.5	2.7	N/A	N/A	N/A	N/A	N/A

Unsatisfactory job efficiency



From EGI accounting portal, so US sites under represented

Tackling the efficiency deficit

- Low efficiency due to two major issues
  - Jobs within pilots needing larger memory/cpu ratio than pilots get
    - Four 4GB/1processor jobs in one 16GB/8processor pilot wastes 4 processors
  - Jobs at grid sites reading input data from Fermilab
    - Not a big issue for reconstruction jobs unless site bandwidth gets maxed out, but significant for analysis jobs
- Work underway on several fronts to tackle these two problems
  - Improvements in application software used by the liquid Argon based experiments
  - New DUNE workflow management system justIN with better matching
  - “Can we fix it?”

Slide from GridPP1

Can we fix it?

Yes we can!

(Subject to available resources …)

Andrew McNab - Manchester HEP - 24 May 2001

THE UNIVERSITY of MANCHESTER

Efficiency: memory ratio

- DUNE and the other neutrino experiments use LArSoft framework to process data
  - So opportunities for improvements to help several experiments
- Recent LArSoft multithreading/processor workshop
  - Explicitly driven by desire to get back down to 2GB/processor
  - Share memory usage between multiple threads running in parallel
  - (Also interest in exploiting many-processor HPC architectures)
- justIN workflow system development by DUNE includes better job to site matching
  - justIN knows geometry of pilots submitted to each site
  - And processor,memory requirements of each type of user job
  - Targets suitable sites already; in future, will rank sites' suitability for a particular type of user job

Efficiency: remote data

- DUNE's justIN workflow system uses "just in time" matching of jobs to job slots, mainly based on replica info from Rucio
  - Finds highest priority, unprocessed input file, nearest to the job slot
  - Provides the job script for that processing campaign and allocates file(s) to the job
- Demonstrated in recent DUNE data challenge with input files at several GridPP sites, CC-IN2P3, CERN, Prague, NIKHEF, BNL, FNAL, …
  - Increased fraction of files processed at same site as the job by 60%
  - … while still using all job slots that became available
  - Could be more aggressive about rejecting unsuitable site/storage combinations
  - And more campaigns, Monte Carlo, user jobs in the mix will help
    - Less temptation to settle for poor matches

WLCG tokens

- Fermilab is strongly influenced by US Open Science Grid timescales for dropping X.509 support as soon as possible
- Pilots are submitted by OSG Pilot Factories at UCSD and CERN
- Aim to drop use of VOMS proxies and rely on WLCG tokens only
- Fermilab users are now doing manual job submission with WLCG tokens
  - Fermilab dCache user scratch areas are now accessible by token
- DUNE plans to participate in WLCG DC24
  - With particular interest in testing storages which have added token support
- DUNE's justIN workflow system entirely relies on tokens for user authorization
  - But can still use DUNE VOMS proxies to access storage on users' behalf - so decoupled from OSG urgency to drop X.509

Summary

- DUNE is ramping up computing requirements
  - Currently driven by the two protoDUNE experiments at CERN
- MicroBooNE and gm2 will scale down
  - With computing requirements shifting to ICARUS and then SBND
  - Then DUNE data taking will start
- The neutrino experiments are aware of and concerned about job efficiencies
- Work underway to improve
  - Software frameworks to make better use of memory/processors
  - Job matching to match memory requirements to job slots
  - Input file locations to job slots to reduce streaming inefficiency
- Tokens are transitioning from nice to have / will be used / required

Fermilab Wilson Hall (in 1970s)

Floor 15 of Wilson Hall (May 2022)

South Dakota is that-a-way!

Short Baseline neutrino line