

# GridPP DUW2024 Update

Daniela, Simon & Janusz

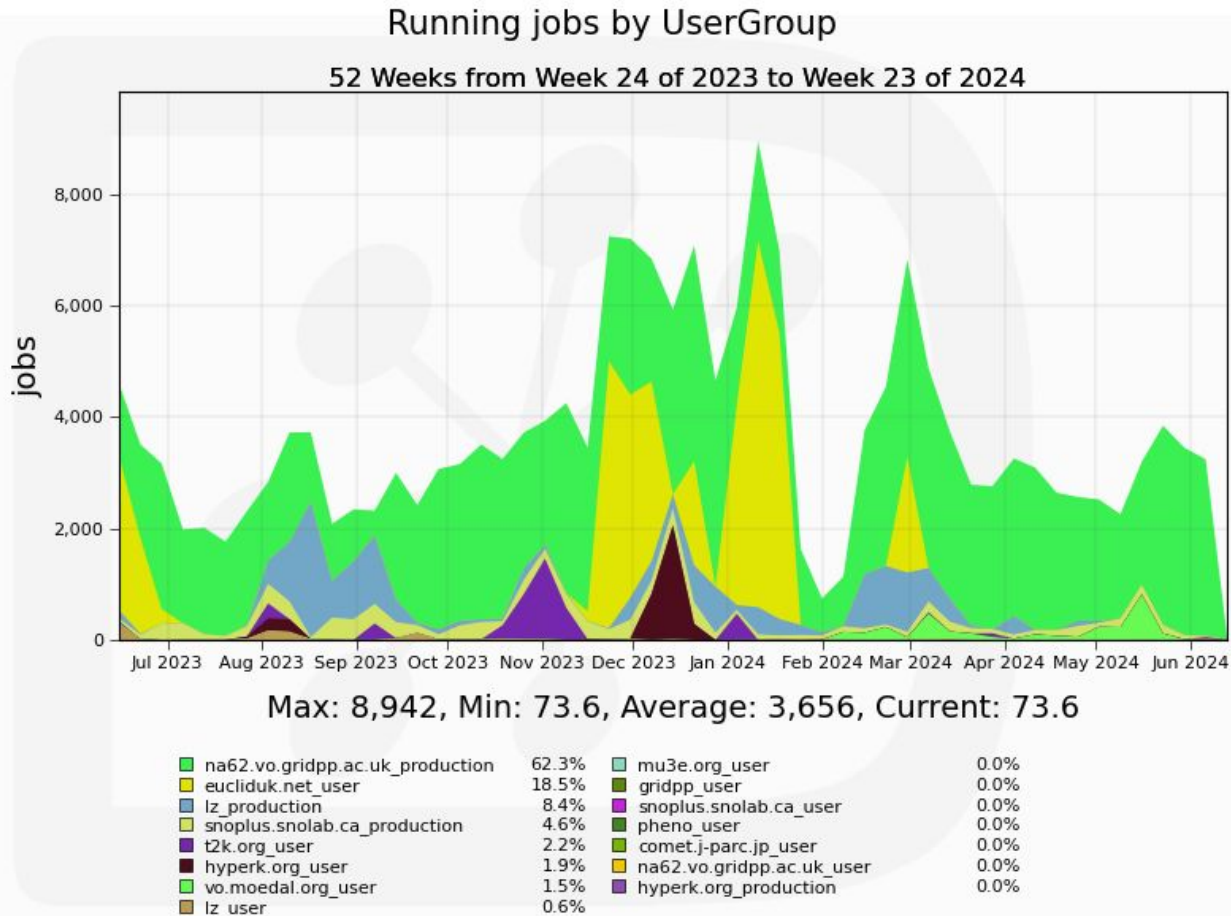
# GridPP DIRAC Overview

- The GridPP DIRAC instance supports small to medium size communities with UK participation:
  - Neutrinos: T2K, HyperK, SoLid, SNOPLUS
  - Dark Matter: LZ
  - Astronomy: Euclid (simulation workflows)
  - Phenogrid
  - Muon decays: COMET, Mu3e
  - Smaller CERN based experiments:
    - NA62 (rare Kaon decays)
    - Moedal (magnetic monopoles)



# Obligatory usage plot June 15<sup>th</sup> 2023 to June 15<sup>th</sup> 2024

No change since the last meeting.



# Developments since the last workshop

- Like most WLCG sites we've been focussing on the upgrade to EL9, so we mainly focussed on **keeping DIRAC running**:
  - The DIRAC servers were upgraded to Rocky 9 without too much trouble
  - We are now using tokens for pilots (almost) everywhere: The one exception is RAL-LCG2.
  - We maintain our own IAM server, but we don't like it:
    - It's not high-availability, unlike our three geographically distributed voms servers
    - It was clear at the Indigo-IAM workshop that things aren't set in stone yet
    - We'd like to do more work with the developers, but it's not clear we'll have the time
  - We deployed the TransformationSystem in production, but so far it's only undergone cursory tests by T2K

# Data management

- We recently had three VOs (T2K, Hyper-K and Moedal) use our DIRAC server for data transfers (using our FTS as the back end):
  - These transfers typically concern a large number of files (500k to 4 mio)

We see timeouts in the file catalogue queries, but no indication that our file catalogue is overloaded, i.e. the dreaded:

```
$ dirac-dms-find-lfns Path=/t2k.org --SE=CA-TRIUMF-T2K1-disk
```

Can not access File Catalog: FileCatalog error ( 1604 : Failed to perform findFilesByMetadata from any catalog)

- It seems very difficult to verify that a transfer is complete:
  - Maybe we just don't know how ?
  - E.g. how do I hunt down checksum errors ?

# Goals for the workshop

- IAM:
  - Testing the “get users from IAM” module without breaking my production server.
  - Develop a bit of a strategy on how to handle “multi-VO” IAM servers on DIRAC (I will have to loop in Simon remotely for that).
- Large scale file transfers:
  - Is there any tuning we can do ?
  - Do we need to implement batching/paging in the dirac-dms-\* commands ? (And if so, can I get Janusz to do it ?)
  - How do I verify transfers ?
- Improve the hackathon:
  - My personal pet peeve: Make the hackathon results more meaningful by making sure the targeted resources actually accept jobs/data:  
<https://github.com/DIRACGrid/DIRAC/issues/7658>
  - Make sure the “non-cvmfs pilots do not run at RAL-LCG2” issue is fixed  
(<https://github.com/DIRACGrid/DIRAC/issues/7657>)
- Get Janusz’ DIRACX into a state where he can focus on development.

# Speaking of DIRAC-X

- As you might have noticed, the GridPP goals for this workshop are all rather acute:
  - We try to keep up with DIRAC-X, but only have limited resources.
  - We need to ensure our current service doesn't degrade while waiting for the future.