UNIVERSITY OF BIRMINGHAM

LHCb Status

GridPP Collaboration Meeting, 2nd September, 2022

Mark Slater, Birmingham University

UNIVERSITY OF BIRMINGHAM

Overview

I will attempt to cover the following from the last 12 months of LHCb running in the UK:

- LHCb job mix and efficiencies
- LHCb data distribution and transfer rates/efficiencies
- General LHCb Status
- T1 Status
- T2 Status

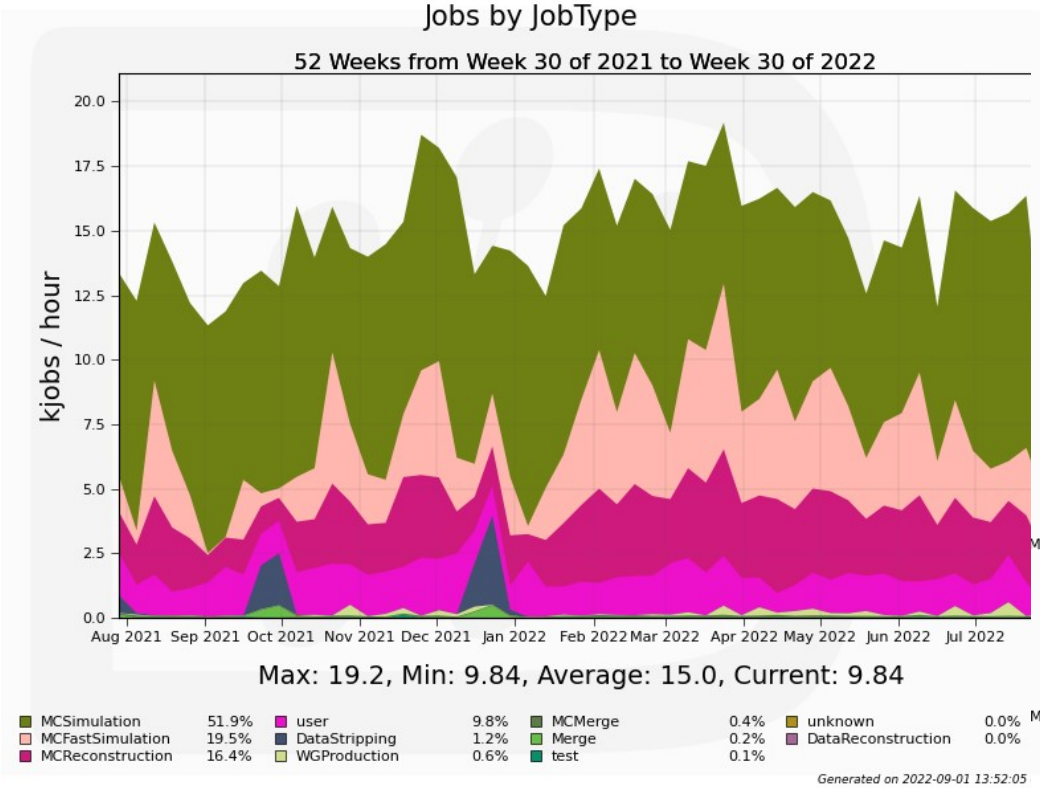
Disclaimer:

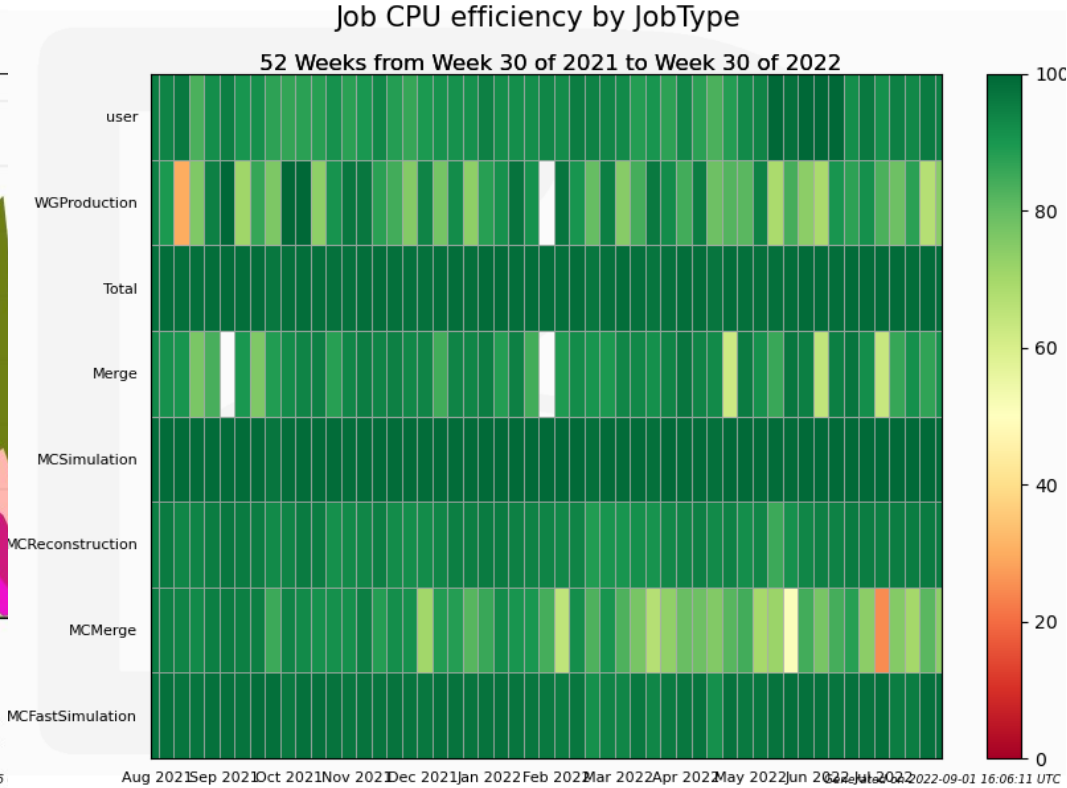
My 'usual' LHCb role is somewhat narrow and the circles involved don't completely overlap. I've tried to keep abreast of things while being the acting Tier 1 Liason but I will have undoubtedly missed/misinterpreted things – apologies!

LHCb Status - Jobs

LHCb have spent the last year running mostly running Monte Carlo jobs (not surprisingly!) with user analysis as well

Currently this is still the case as the experiment is still commissioning and preparing the sub detectors

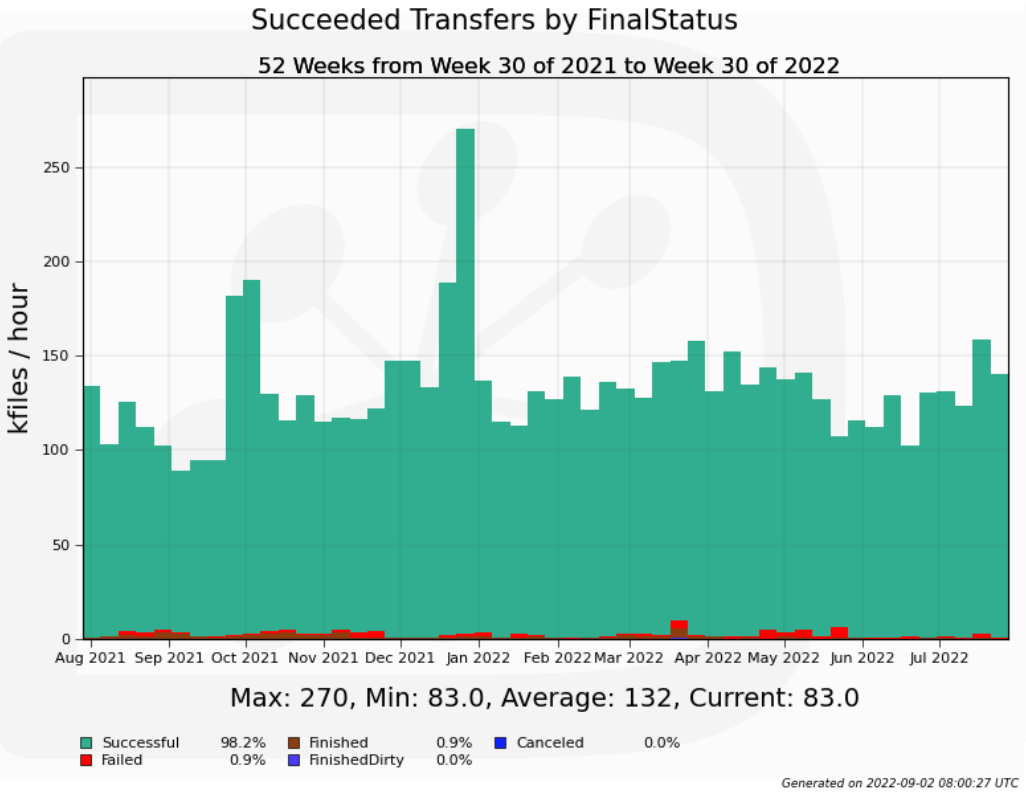


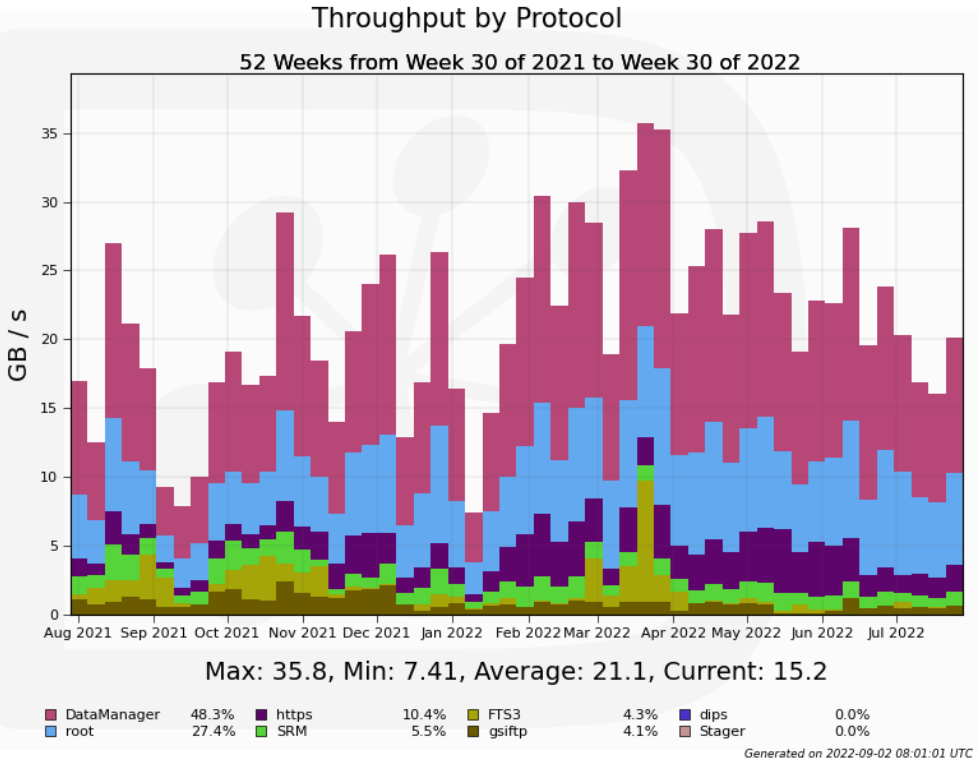


LHCb Status - Data

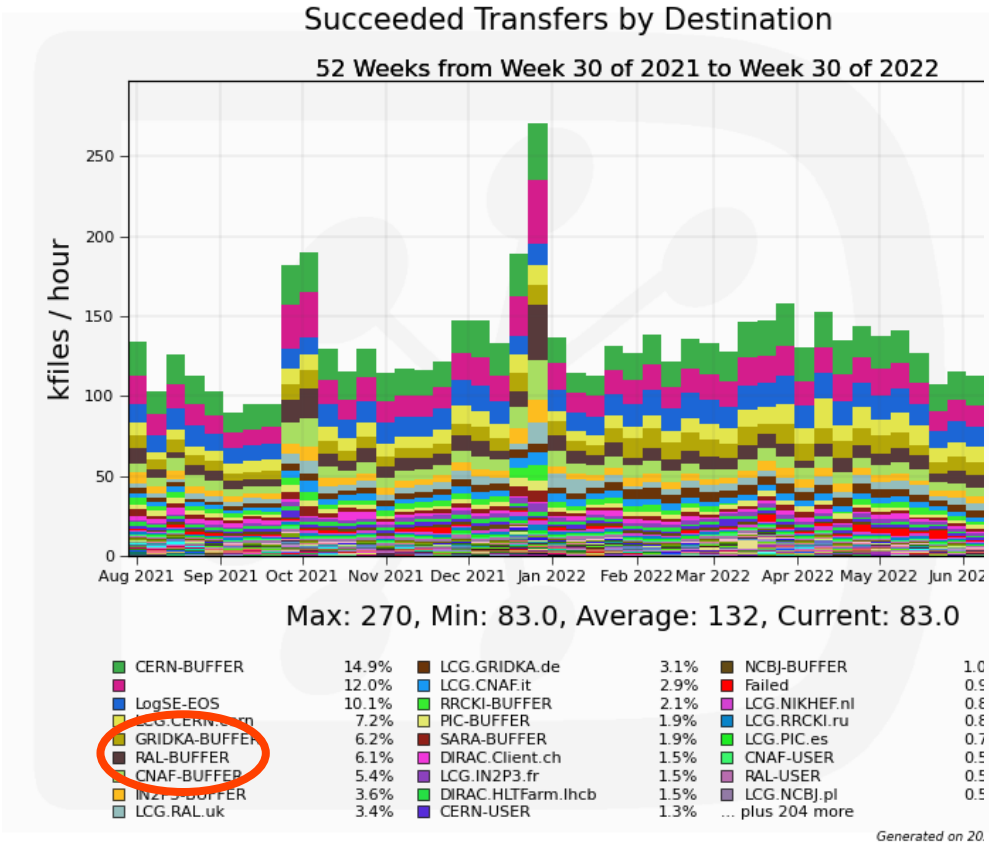
Again, data operations (which includes job up/download) has been at a relatively low level but stable over the last 12 months

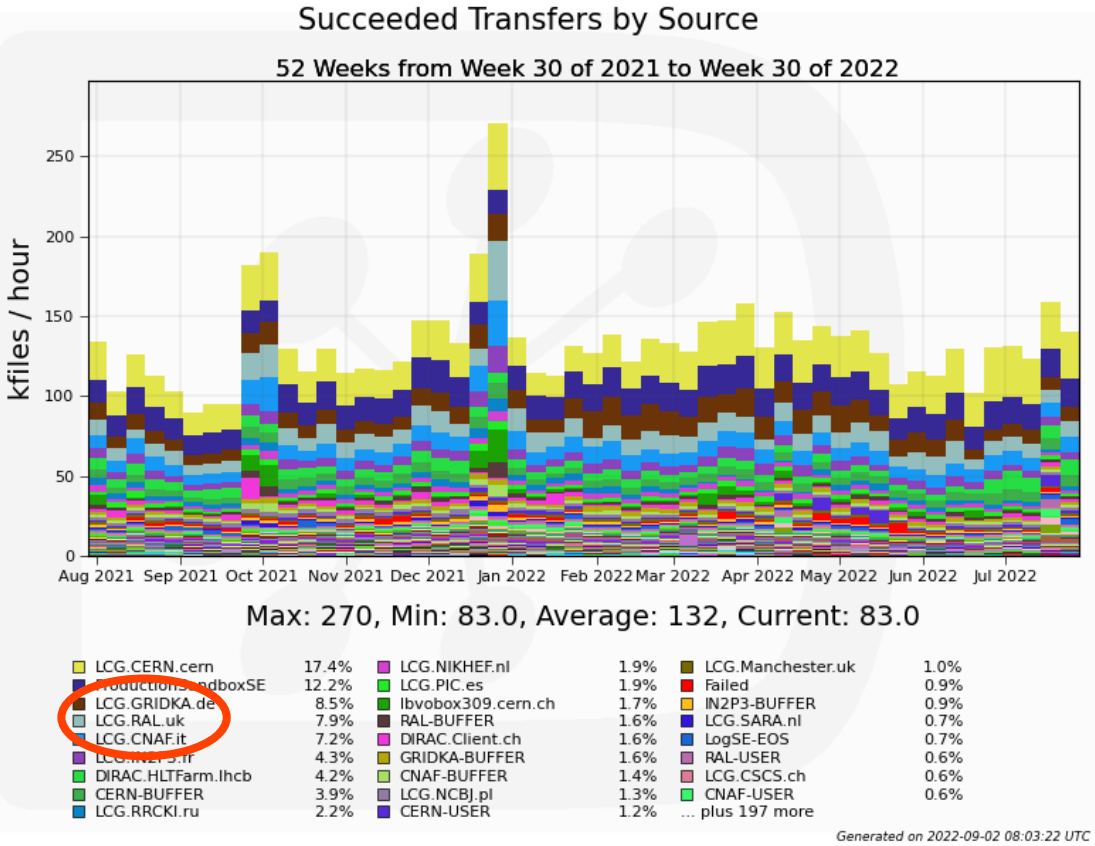
LHCb have averaged ~125000 files per hour at a rate of ~15 GB/s





LHCb Status - Data





Note: LCG.* are sites Where jobs are running

LHCb Status

Some general issues of note that have come up in the last 12 months (there are probably more but I can't remember them!)

Failed pilots due to Memory usage:

Pilots failing is a common issue and is often user error. However, some specific MC productions have recently been using 4GB+ per core. The Computing team is trying to restrict them to the HLT Farm

Timeouts from hadd

There was issue found at NL-T1 where timeouts from hadd occurred. This was due to hadd attempting to open all given files at once and hitting the connection limit on some pool nodes. See https://ggus.eu/ws/ticket_info.php?ticket=153653

Reboot Campaigns

LHCb are happy if sites don't drain their workers before reboots as the system will recover. However, a warning DT in GOCDB would be good so we know what's happening when we see aborted pilots :)

Authentication and Tokens

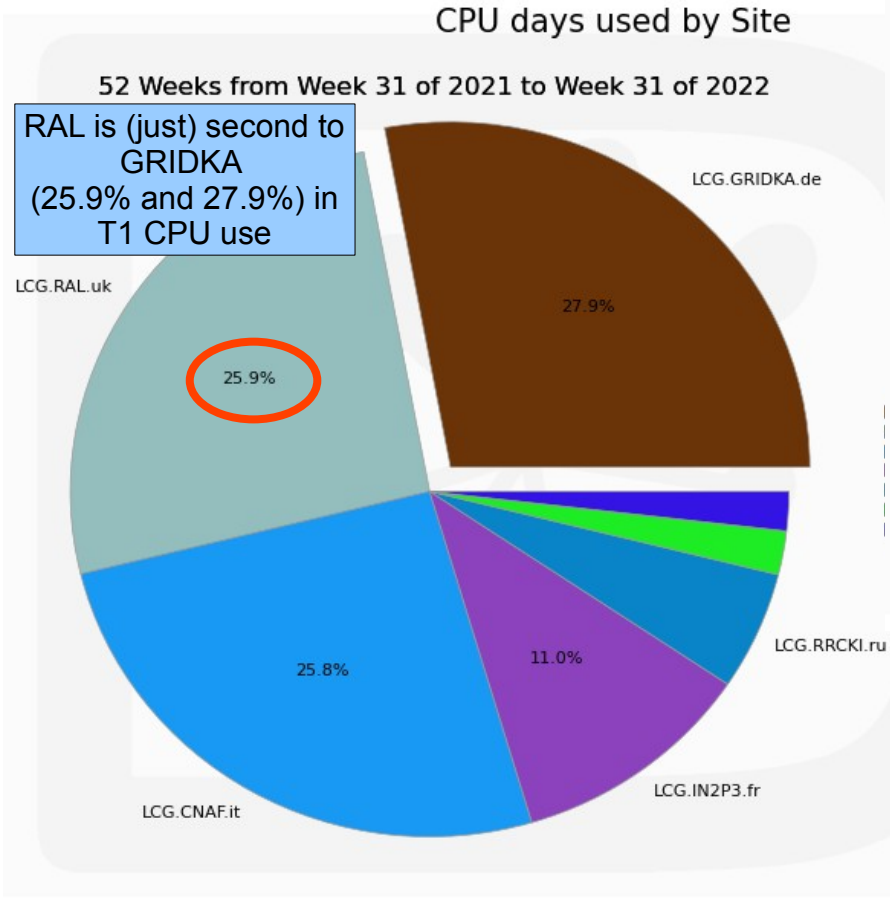
IAM Services will hopefully be ready for LHCb this Autumn though I would assume switching off the VOMS ones are still some time away

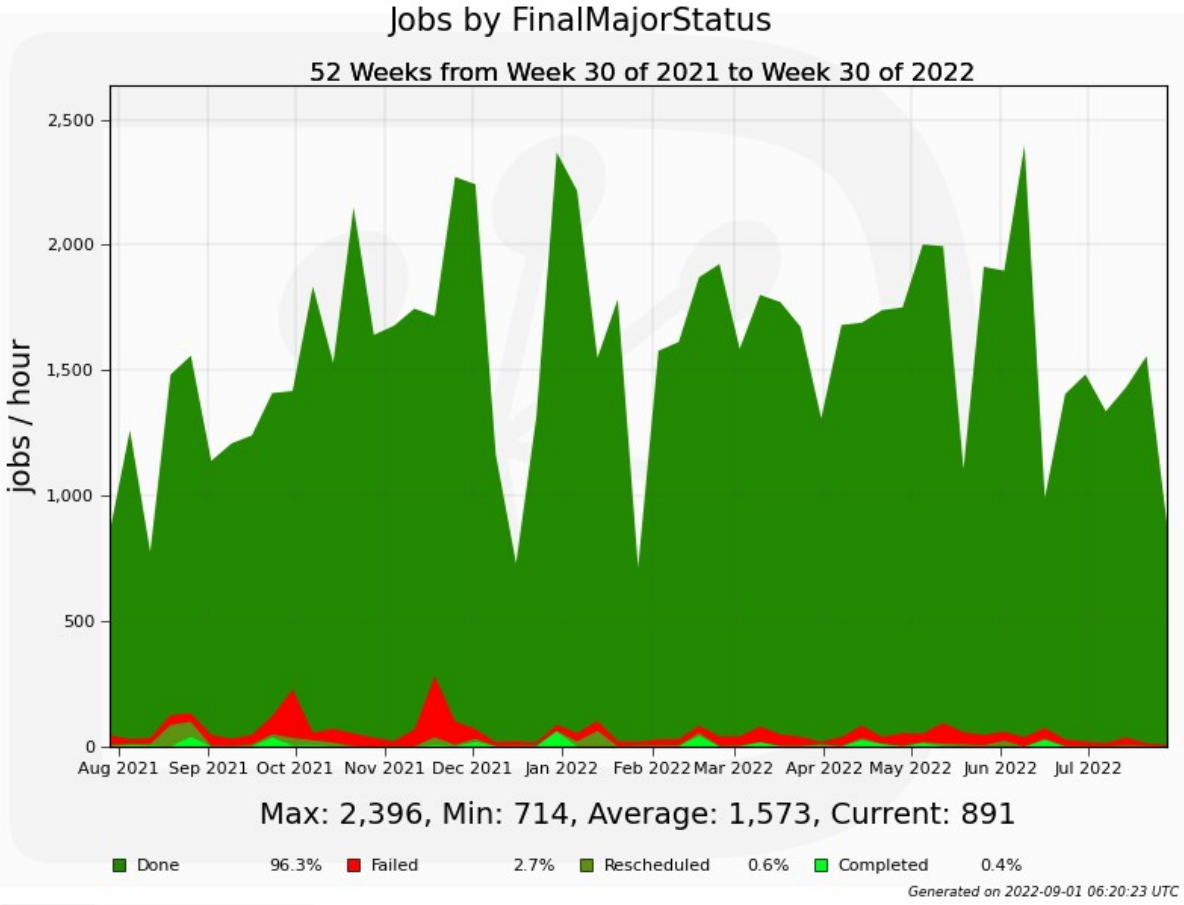
DIRAC v8 (out this Autumn I believe) introduces *experimental* support for authenticating via tokens

By the end of the year, DIRAC should hopefully be able to talk nicely with HTCondor and ARC CEs that only offer token authentication

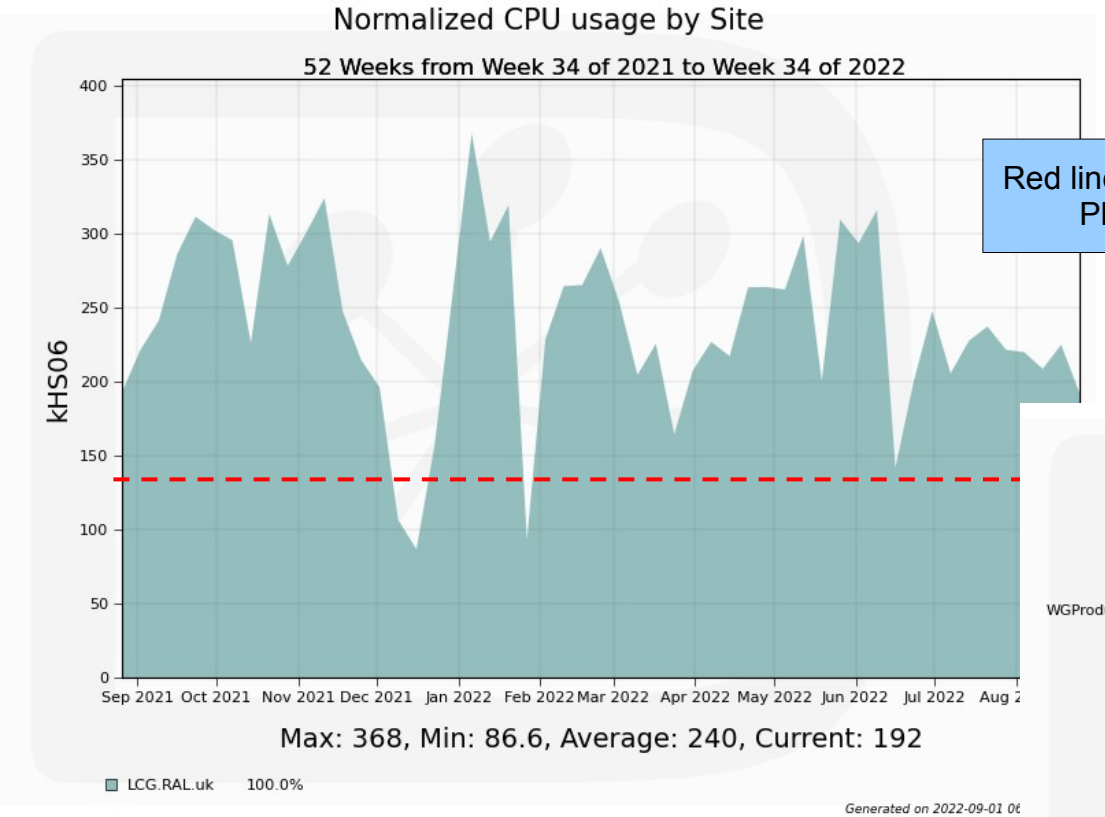
However, LHCb have done very little testing on this so there would be additional delay from the LHCb side to getting this into production

Tier 1 Job Statistics

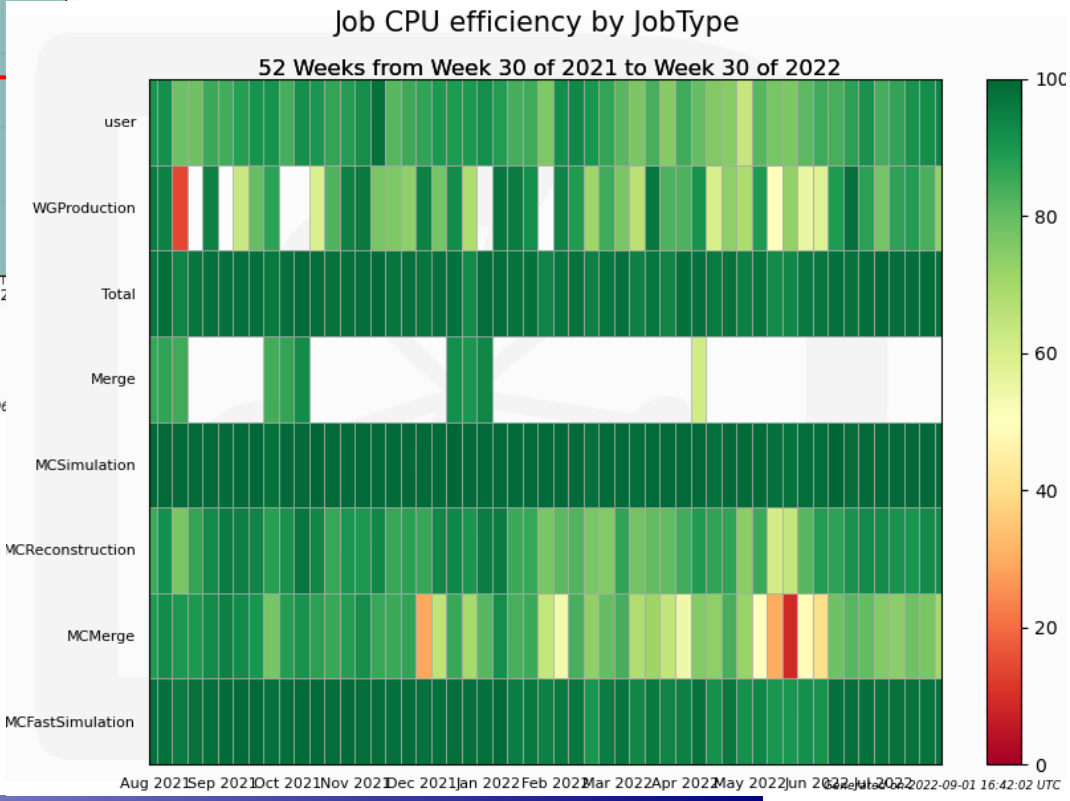




Tier 1 Job Statistics



Red line indicates Pledge

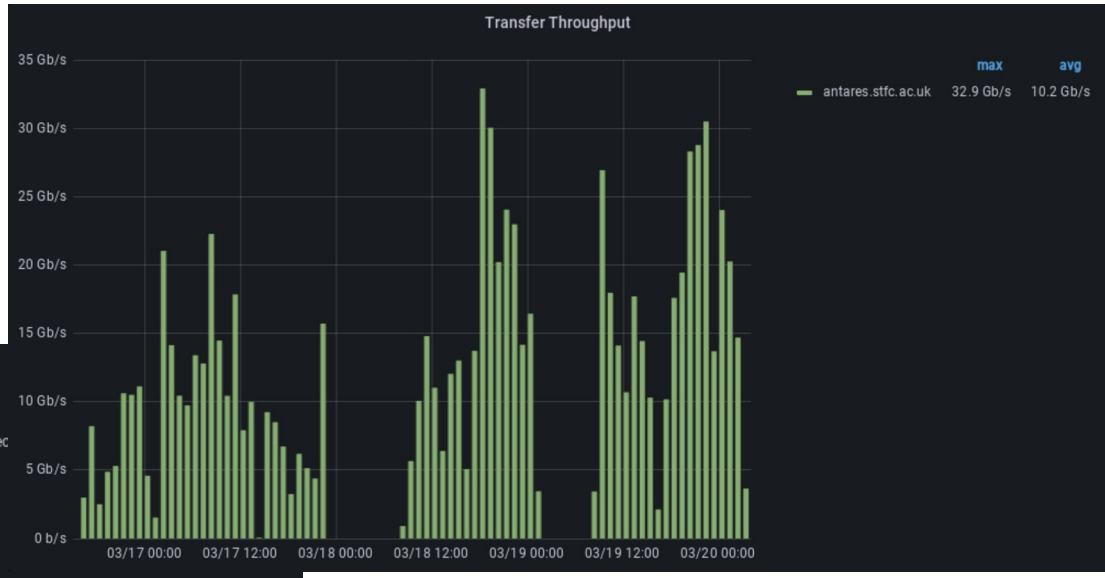


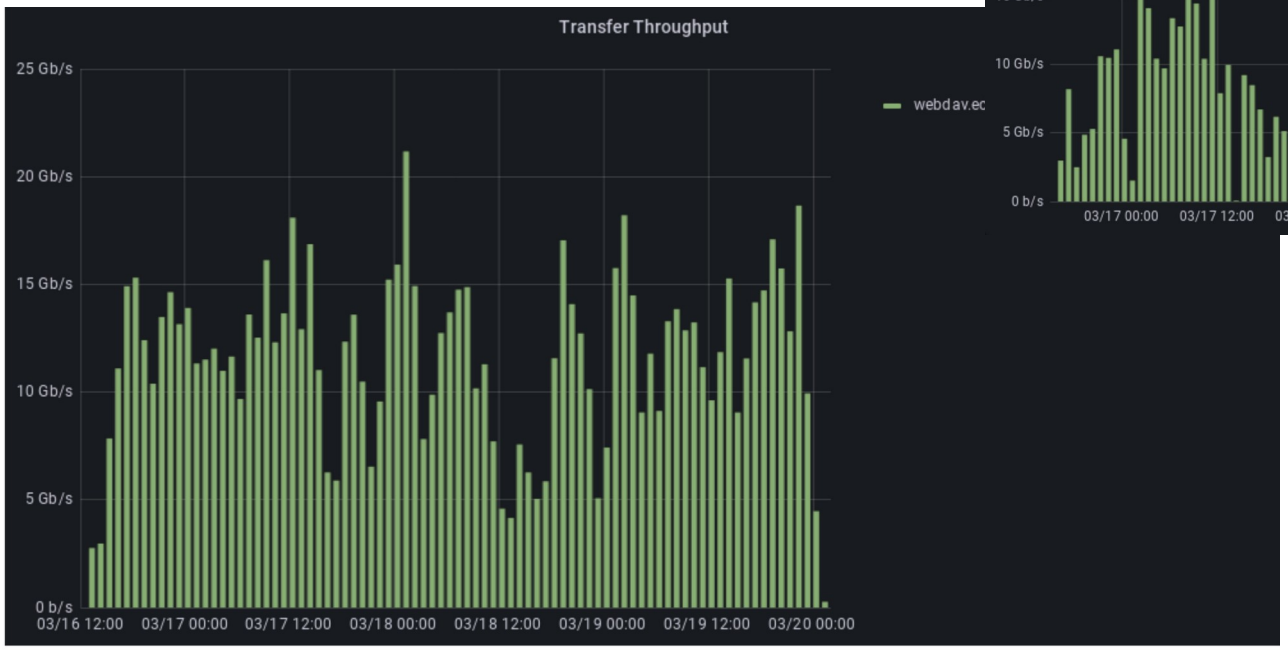
CPU efficiency very Similar to overall job efficiencies

Antares Commissioning

Early in the year we had the commissioning of Antares and the subsequent Tape Data Challenge. Setup for LHCb was:

- LHCb provide ~ 2 days worth of data
- Transfer from EOS → CTA (T0)
- Transfer EOS → T1-disk → T1-tape
- Remove from T1-disk





Avg. Speed of 1.41/1.28 (disk/tape)
With a max of 2.65/4.11
Expected speed was 2.96

Antares Commissioning

Overall, the Tape challenge went well and LHCb were happy with it overall with it. Conclusions were:

- Antares suffered 2 outages during the challenge which lowers the overall TAPE efficiency (not disk)
- Discounting those, the efficiency was very good for a new service
- ECHO efficiency was affected by some issues with webdav. More gateways are planned to alleviate these problems (GGUS:156277)
- Overall throughput was a factor 2 below target but reasons for this are understood

XRoot/WebDAV access

One significant issue that LHCb had with Antares was the initial lack of WebDAV support

Initially, Third Party Copy (TPC) to Antares was only available via xroot but all other sites (for LHCb at least) can use SRM with gsiftp and https

This would mean direct TPC from an external site → Antares would not be possible

With advice from CERN, RAL enabled WebDAV access with some known and accepted drawbacks

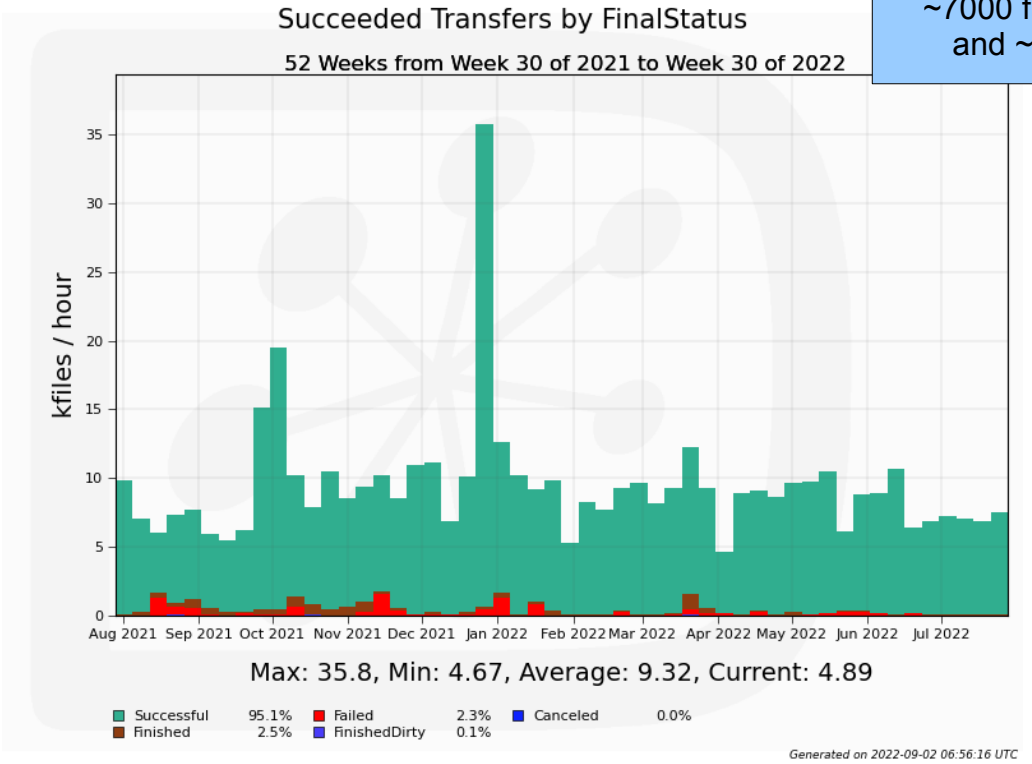
Note that other VOs can do a 'multihop' strategy but DIRAC is unable to register all intermediate steps without significant development

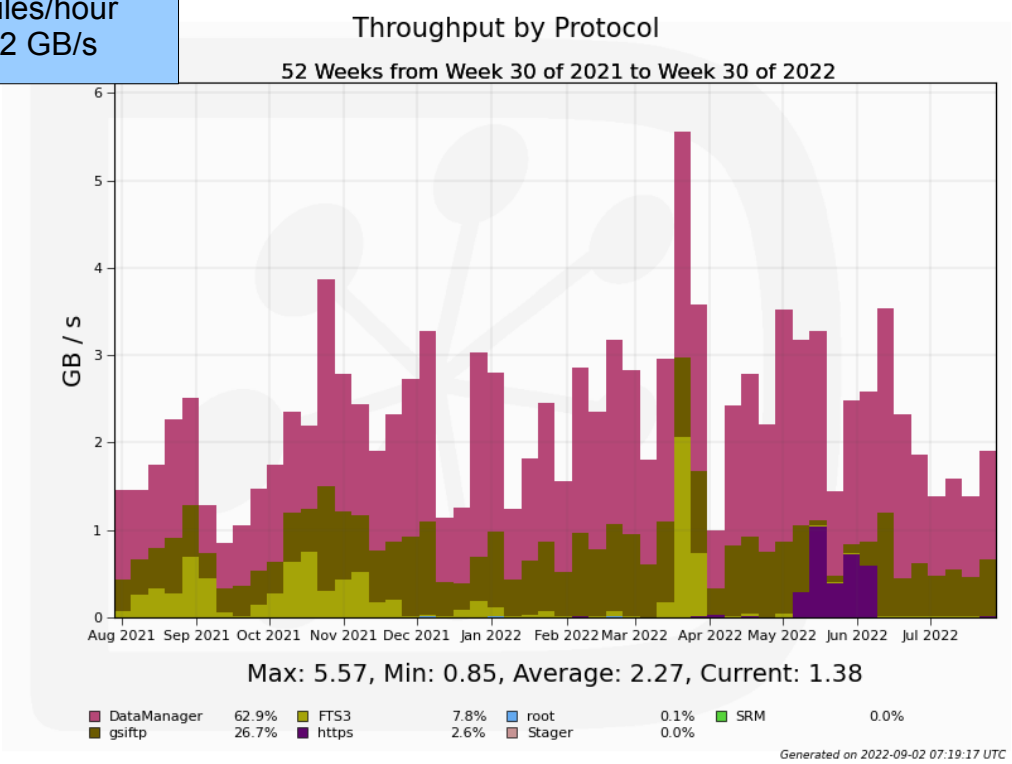
WLCG HTTP Rest tape will be permanent solution to this

Data Operations (Writing)

Writing operations for RAL have been stable with no long term outages or significant problems

Throughput average ~7000 files/hour and ~2 GB/s

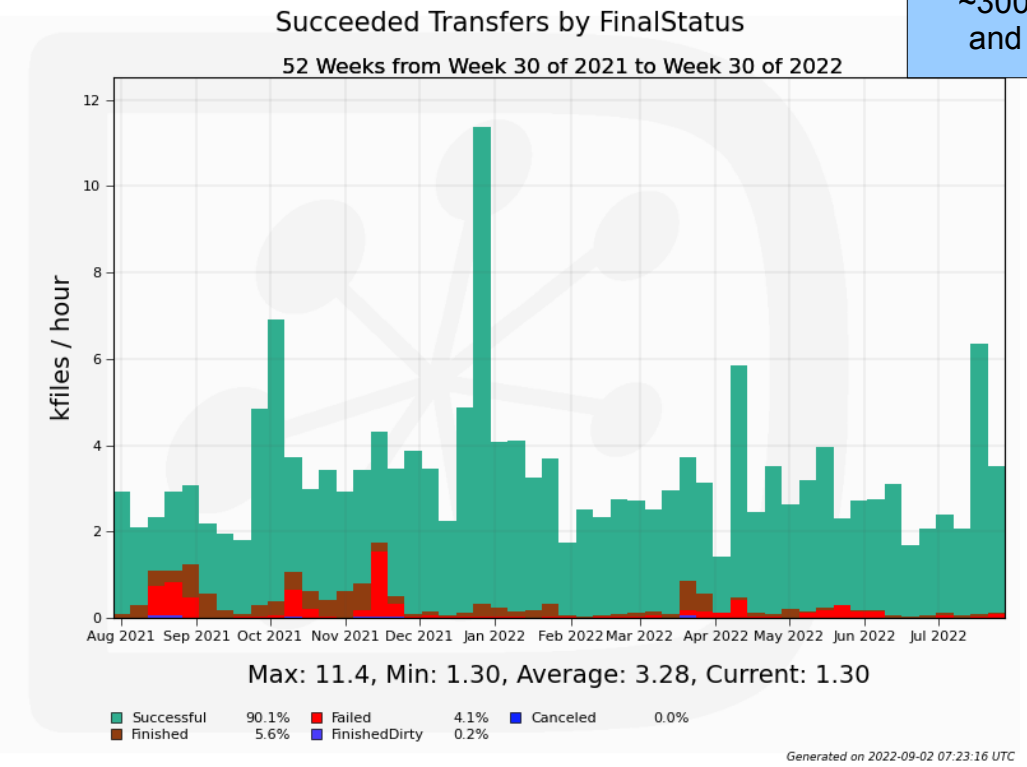


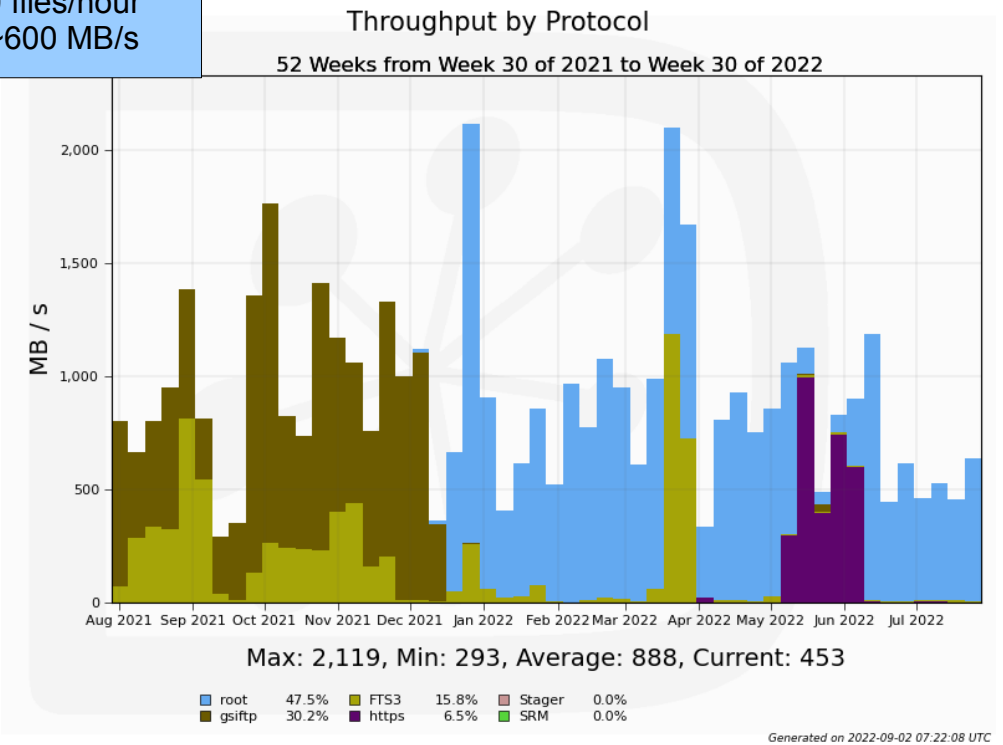


Data Operations (Reading)

Reading operations were a bit more unstable last Autumn with WebDAV but switching to root fixed this

Throughput average ~3000 files/hour and ~600 MB/s

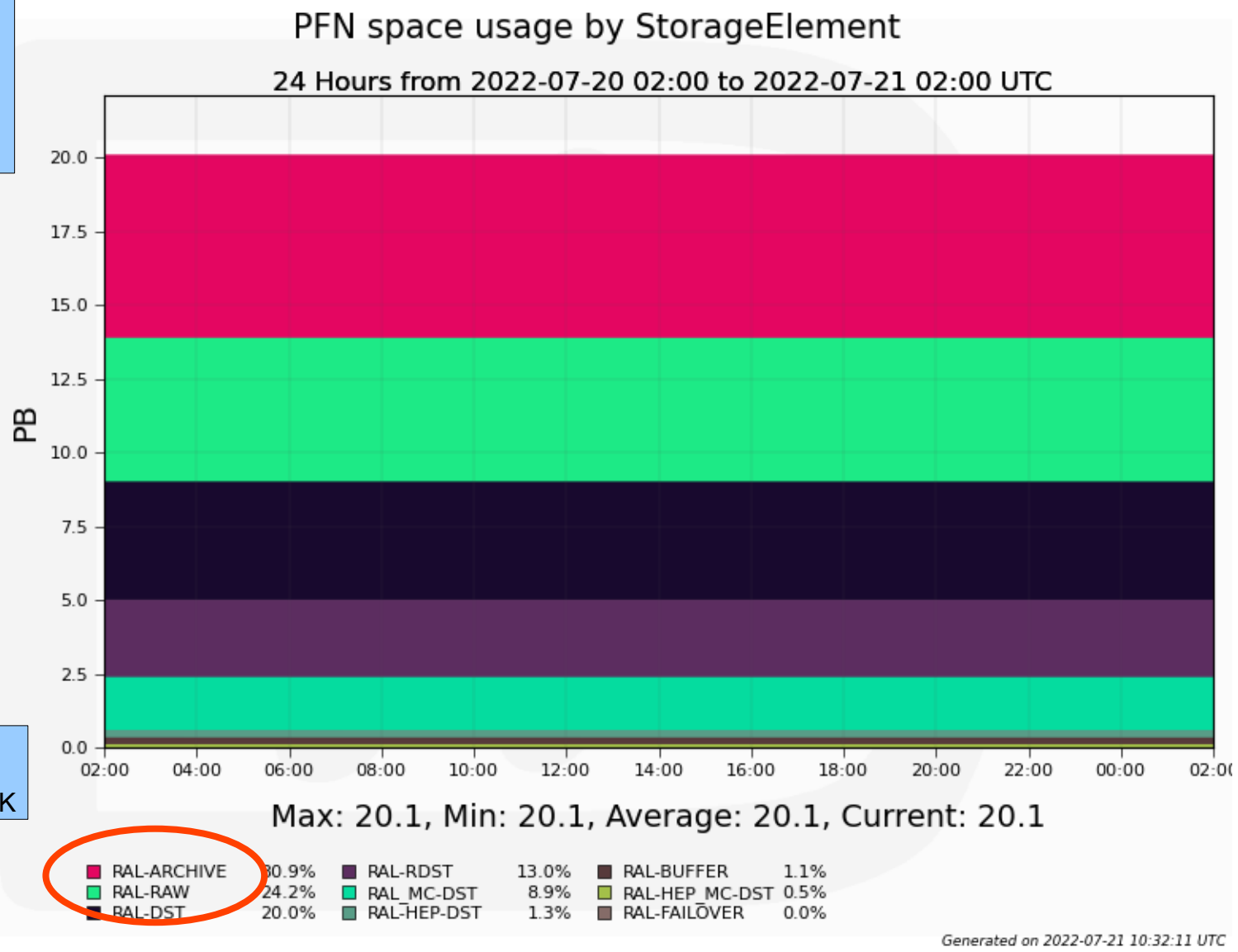




UNIVERSITY OF BIRMINGHAM

Current Storage

Usage/pledge
TAPE = 12PB/32.8PB
DISK = 8PB/12.4PB



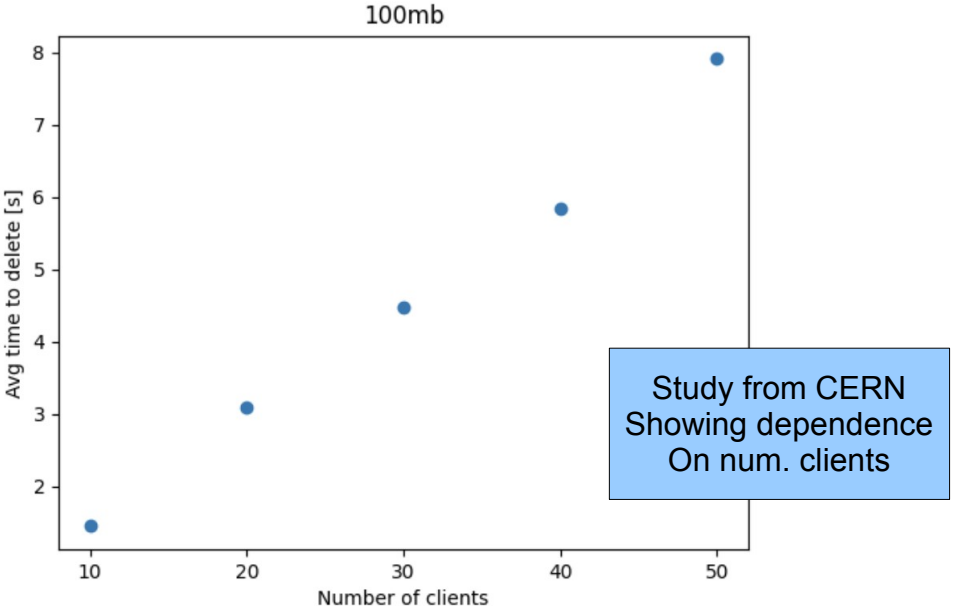
RAL-RAW and RAL-ARCHIVE are TAPE, all others DISK

Long Term Problems

For some time now, LHCb has seen problems with ECHO deletions timing out or being very slow.

This was not just isolated to LHCb but it hit LHCb hardest due to deletions usually being requested in bulk

Mean	21.084608
Std Dev	33.432989
Min	0.183855
25.00%	9.902463
50.00%	17.754771
75.00%	26.446587
Max	888.245457



This causes problems as the slow deletes 'bung up' the threads in DIRAC and make it very difficult to clean productions

After a lot of work by many people, looks to be due to Xroot proxies serialising the requests

Config change being tested at the moment – many thanks to all involved!

Long Term Problems

The two other long term tickets:

Vector Reads:

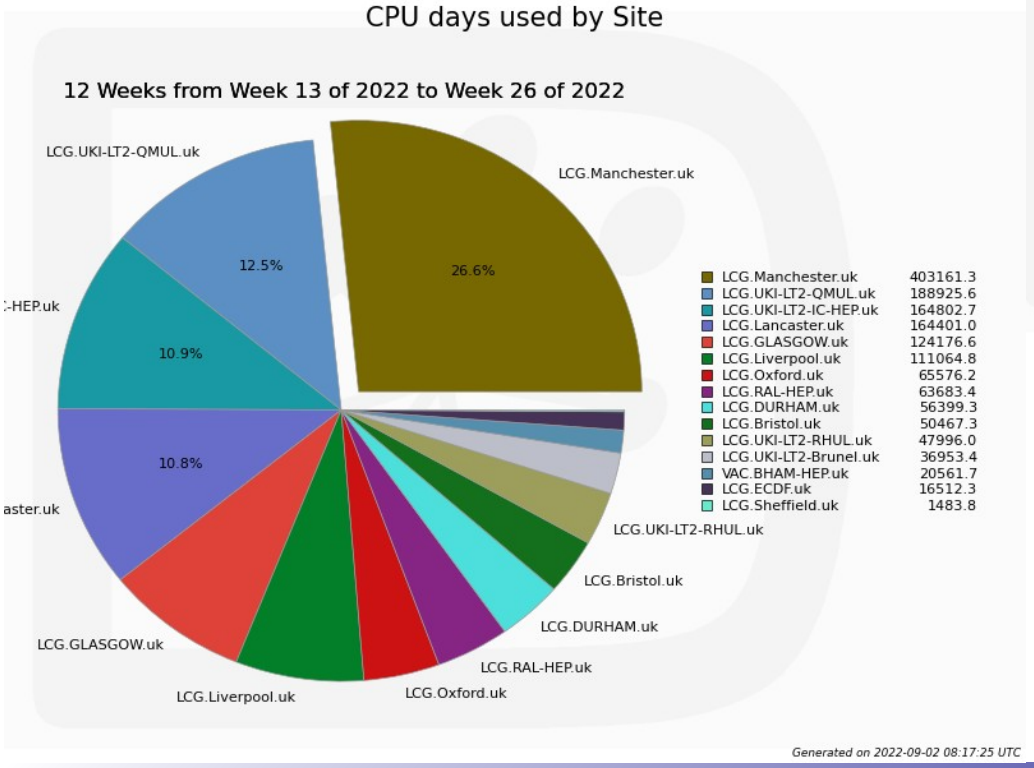
- Jobs streaming files using vector reads were seen to fail at a high rate
- This will become a significant way of working going forward so is important to LHCb
- Great work from Rob Currie indicated issues in XRoot not dealing with disconnects
- A possible cause of these disconnects has recently been found by Jyothish

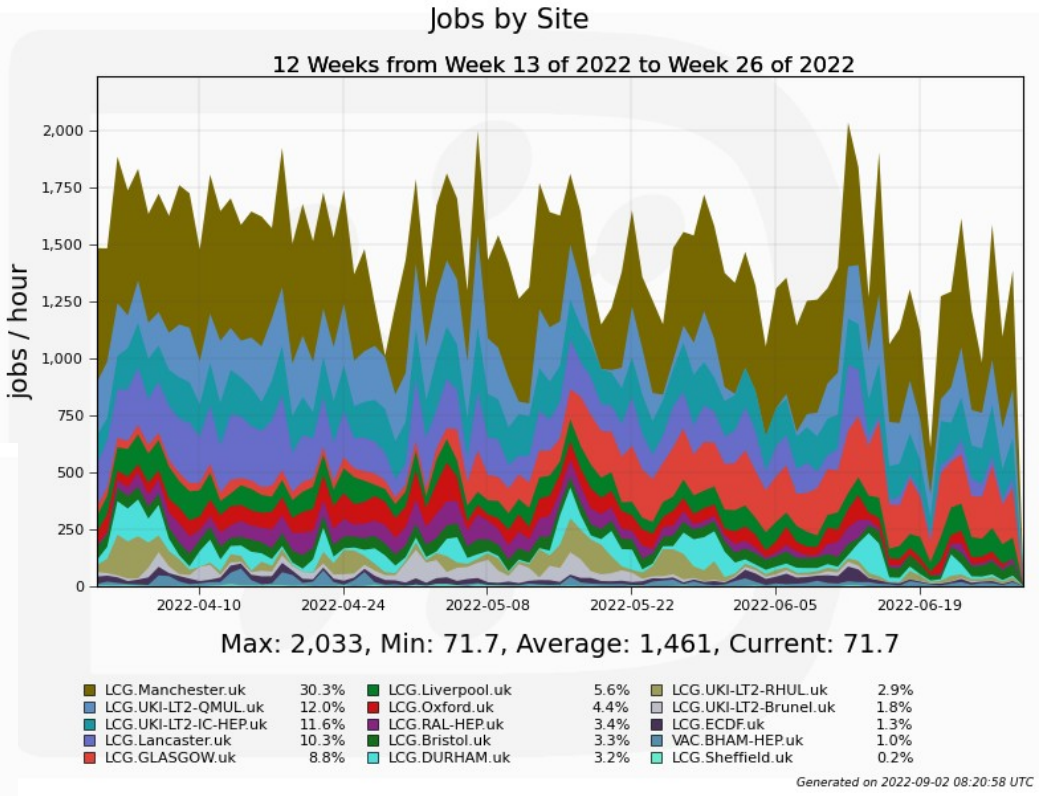
Slow performance of WebDAV:

- General WebDAV performance has caused issues over the last 6+ months
- This was the reason for a switch to root protocol for reads
- New gateway hardware and other updates will hopefully alleviate this problem

The last couple of months have seen a lot of progress – many thanks to Alasdair, Jyothish, Rob et al. for all their hard work!

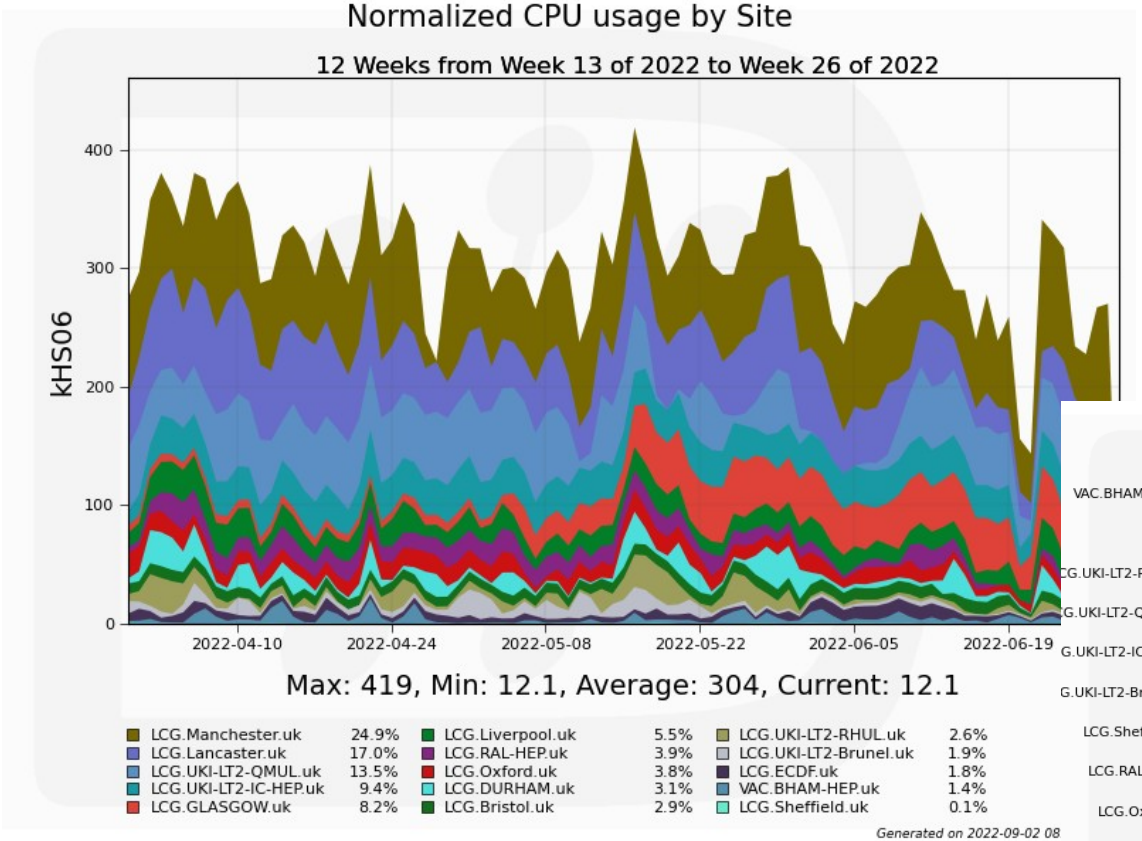
Tier 2 Job Statistics

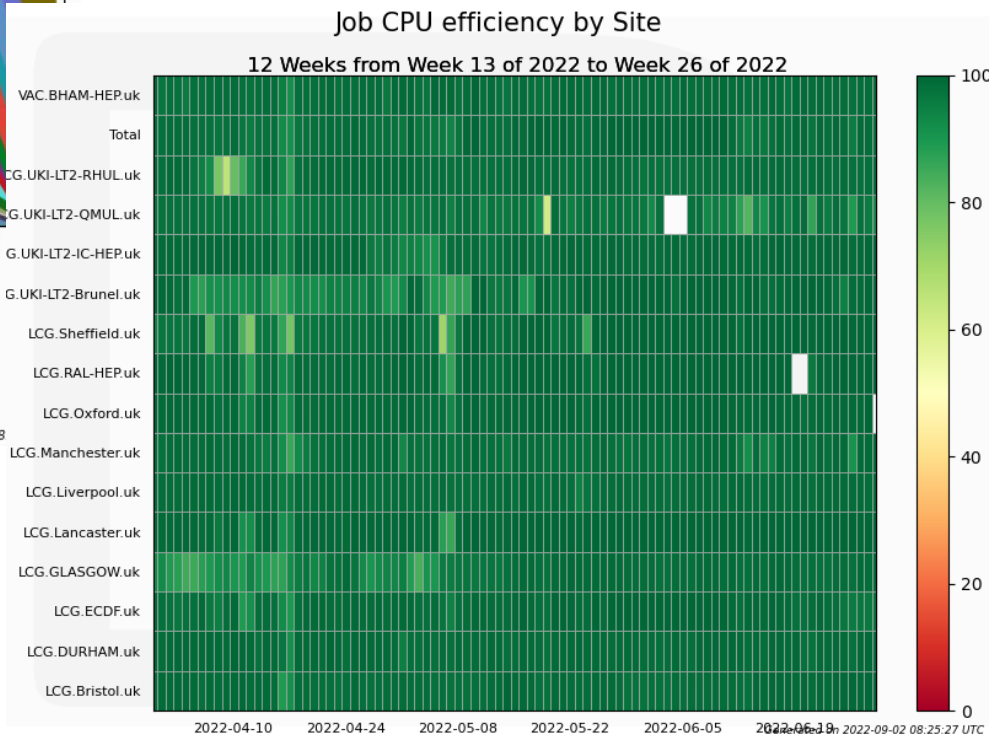




UNIVERSITY OF BIRMINGHAM

Tier 2 Job Statistics

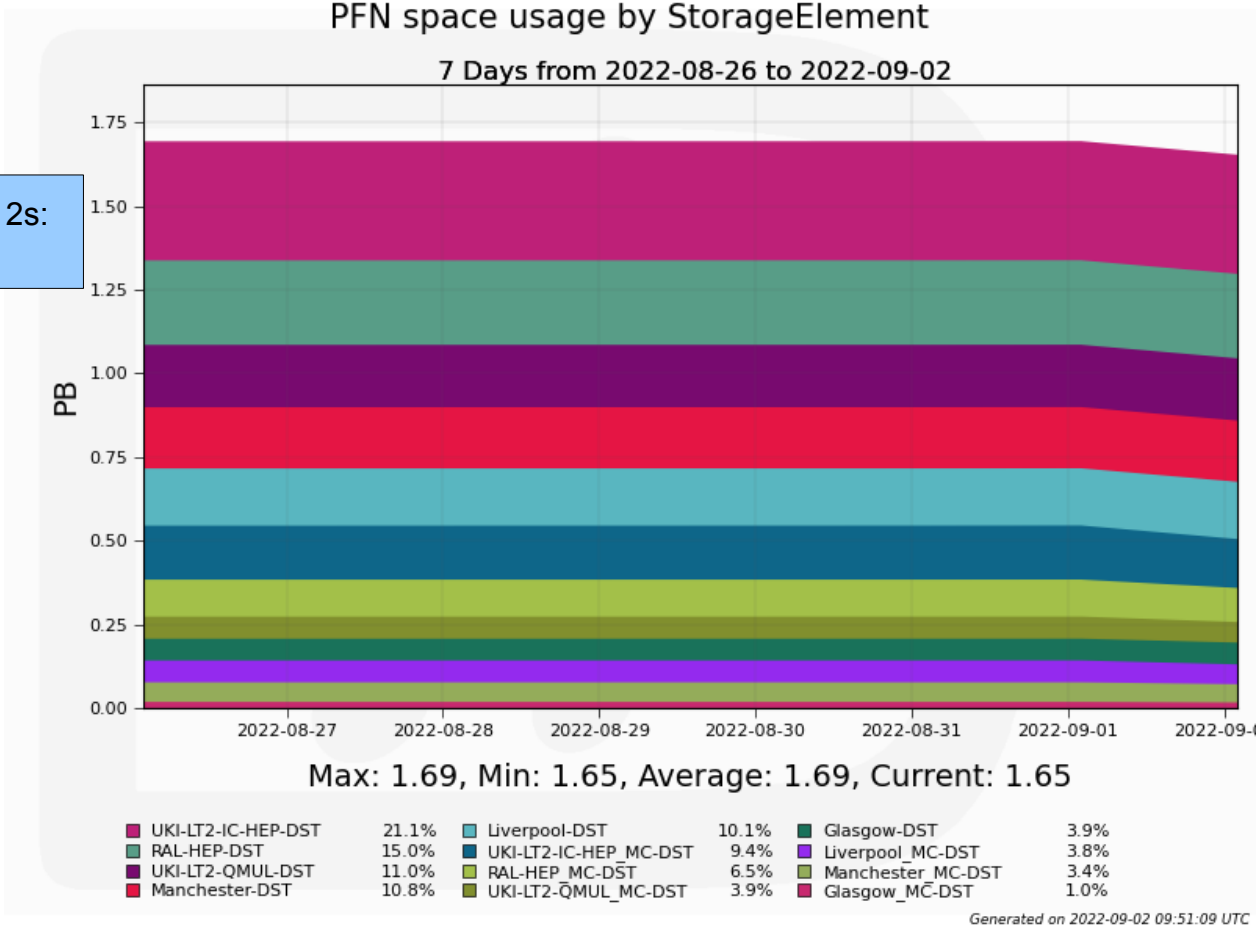




Tier 2 Data Operations

Tier 2 Storage is not being used a lot at the moment but will be when we start data processing

Usage across Tier 2s: ~1.65PB



Ticket Summary

Quick overview of the ticket situation which is pretty good!

SITE	Num. Of Tickets	Open Tickets
RAL-LCG2	39	2
UKI-LT2-QMUL	13	
UKI-SCOTGRID-DURHAM	13	
UKI-LT2-IC-HEP	8	
UKI-SCOTGRID-GLASGOW	7	2
UKI-SOUTHGRID-RALPP	6	
UKI-SCOTGRID-ECDF	5	1
UKI-LT2-RHUL	4	
UKI-LT2-Brunel	3	
UKI-NORTHGRID-SHEF-HEP	3	
UKI-NORTHGRID-LANCS-HEP	2	
UKI-SOUTHGRID-BRIS-HEP	2	
UKI-NORTHGRID-LIV-HEP	1	
UKI-NORTHGRID-MAN-HEP	1	
TOTAL	107	5

Typical problems are caused by:

- Broken CVMFS on some nodes
- Large memory usage
- Transfer errors

Summary

Overall, LHCb has run smoothly in the last 12 months in the UK

- We have been able to use all pledged resources (and beyond)
- General smooth operations
- Problems dealt with quickly
- Significant progress made on the few long term tickets at RAL

Again, Many Thanks to all those who have helped keep things going!