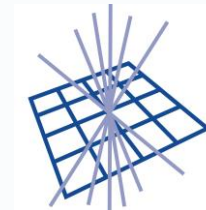


LHCb status

GRIDPP 51

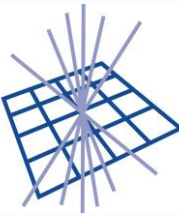
alexrg



Contents

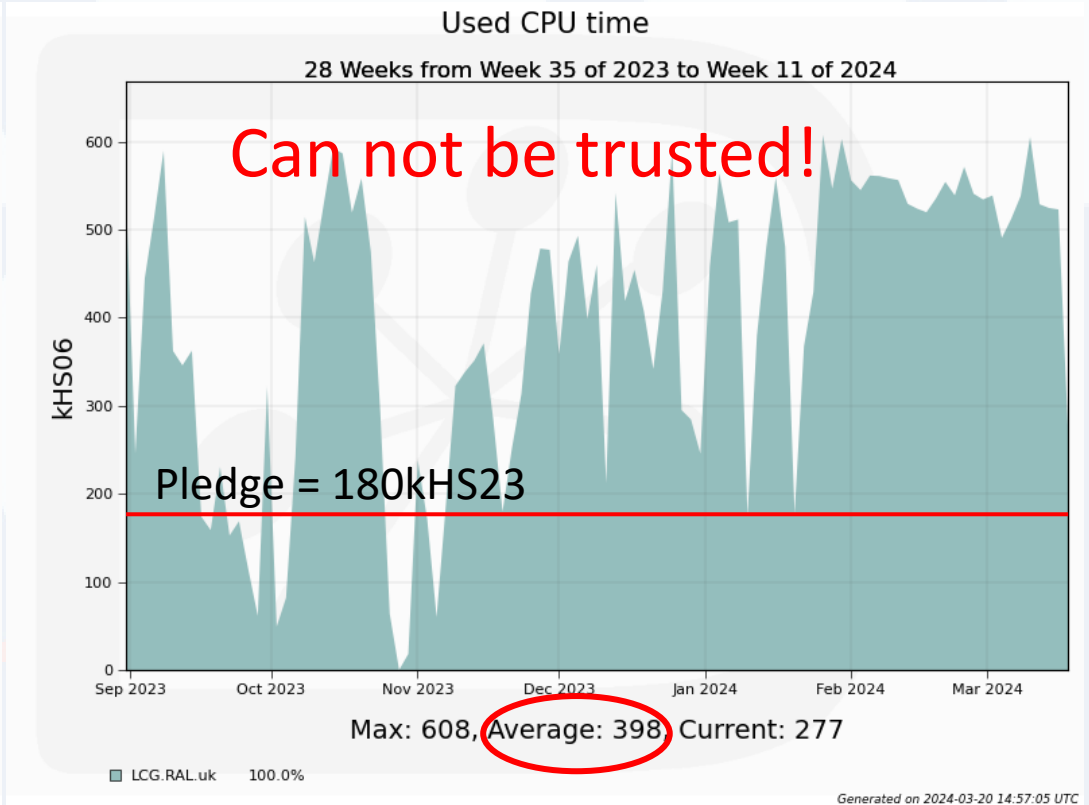
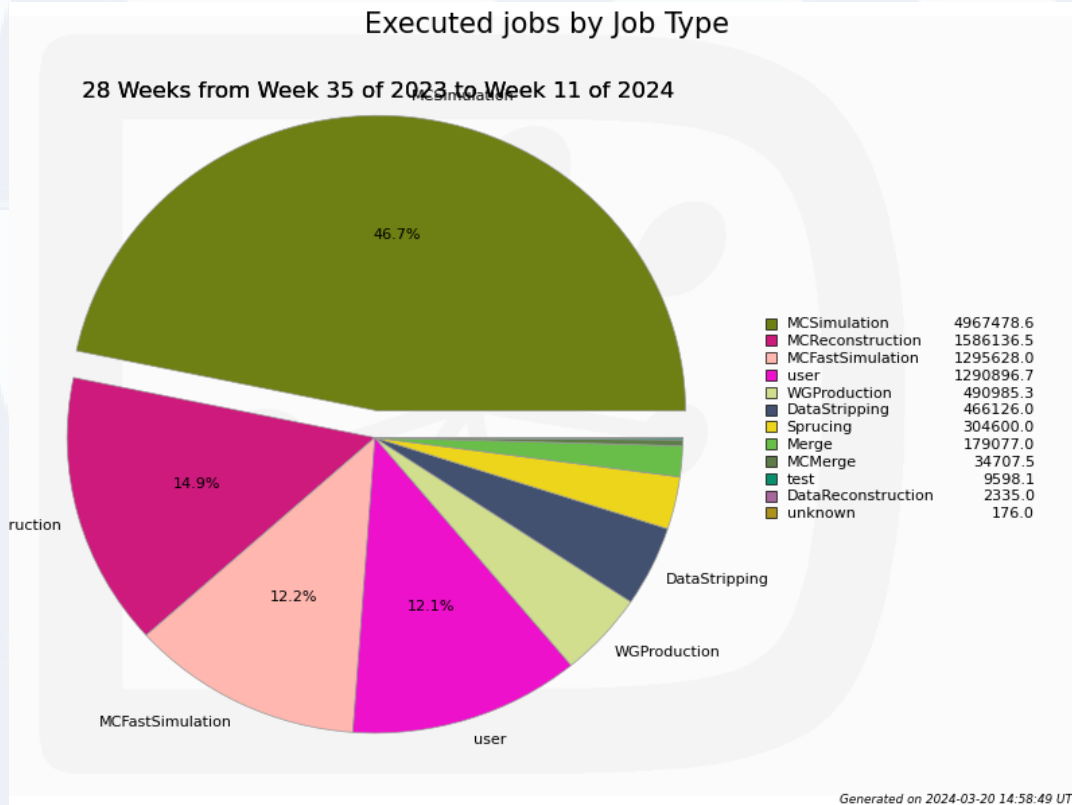
In this talk I'll try to cover the following topics:

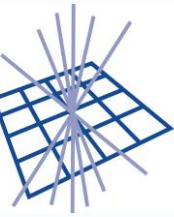
1. LHCb T1 resource usage for the last half year.
2. RAL T1 problems and solutions.
3. LHCb T2 resource usage.
4. LHCb news and plans.



LHCb Jobs at RAL T1

The majority of LHCb jobs at RAL for the last 6 months are Monte Carlo jobs, with some data processing jobs (some of them were processing 2023 data!)

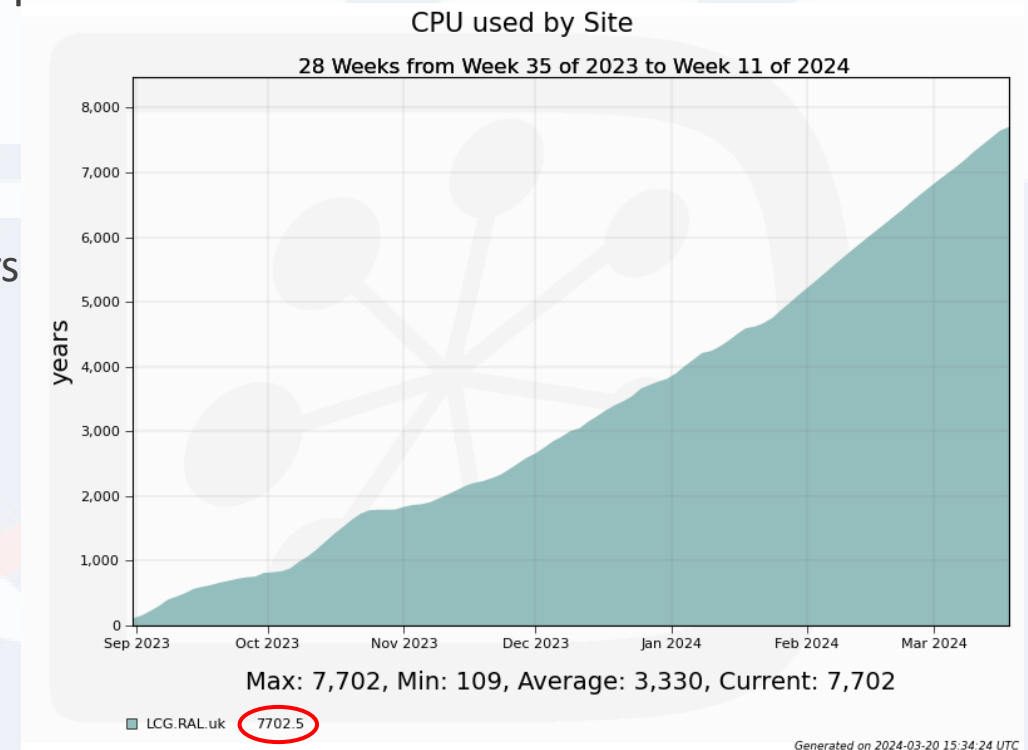


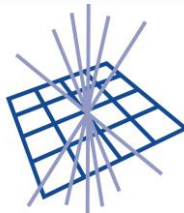


LHCb Jobs at RAL T1

Normalized CPU number looks to good to be true

- DIRAC code for calculating normalization factor is being updated
- If we use static normalization factor from RAL, result is different (should be closer to reality)
 - $7702.5 * 12.7 / 0.55 = 177857$ HS23
 - Blue = cpu years; purple = Normalization factor, green = years in the reporting period
 - Pledge is 180 kHS23
 - To be reduced to 140kHS23 next FY

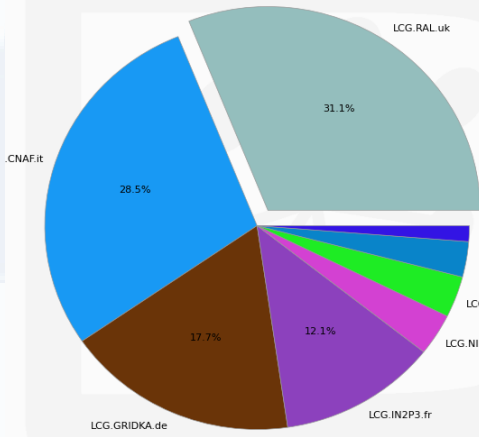




Comparison

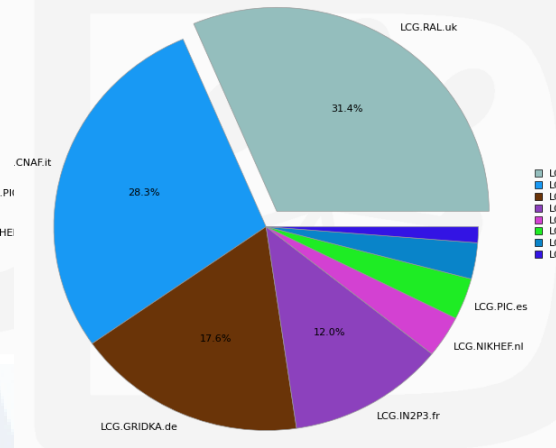
CPU time consumed by site

28 Weeks from Week 35 of 2023 to Week 11 of 2024



Wall time consumed by site

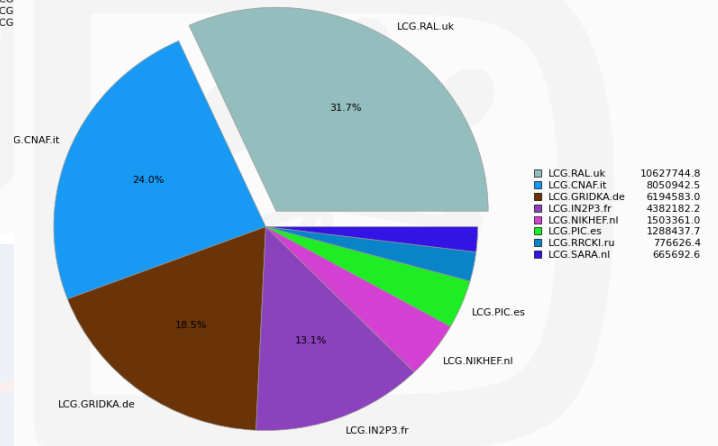
28 Weeks from Week 35 of 2023 to Week 11 of 2024



RAL has provided the biggest amount of CPU resources among all T1 sites, in terms of CPUtime (first plot), walltime (second), and number of jobs (third)

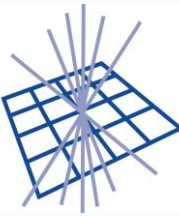
Executed jobs by Site

28 Weeks from Week 35 of 2023 to Week 11 of 2024



LCG.RAL.uk	10627744.8
LCG.CNAF.it	8050942.5
LCG.GRIDKA.de	6194583.0
LCG.IN2P3.fr	4382182.2
LCG.NIKHEF.nl	1503361.0
LCG.PIC.es	1288437.7
LCG.RRCKI.ru	776626.4
LCG.SARA.nl	665692.6

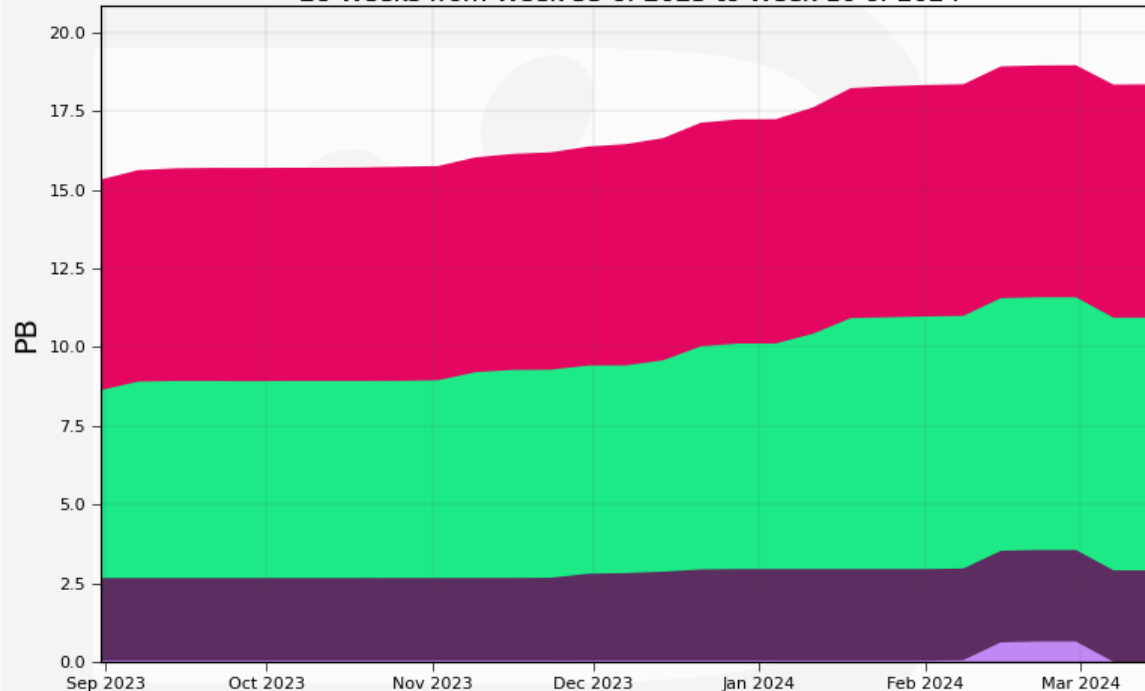
Generated on 2024-03-20 15:00:47 UTC



Tape usage

PFN space usage by StorageElement

28 Weeks from Week 35 of 2023 to Week 10 of 2024

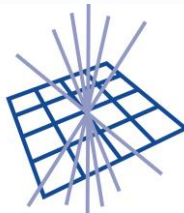


Max: 19.0, Min: 15.3, Average: 16.9, Current: 18.3

■ RAL-ARCHIVE 41.6% ■ RAL-RAW 41.4% ■ RAL-RDST 16.4% ■ RAL-DC-RAW 0.6%

Generated on 2024-03-20 15:05:16 UTC

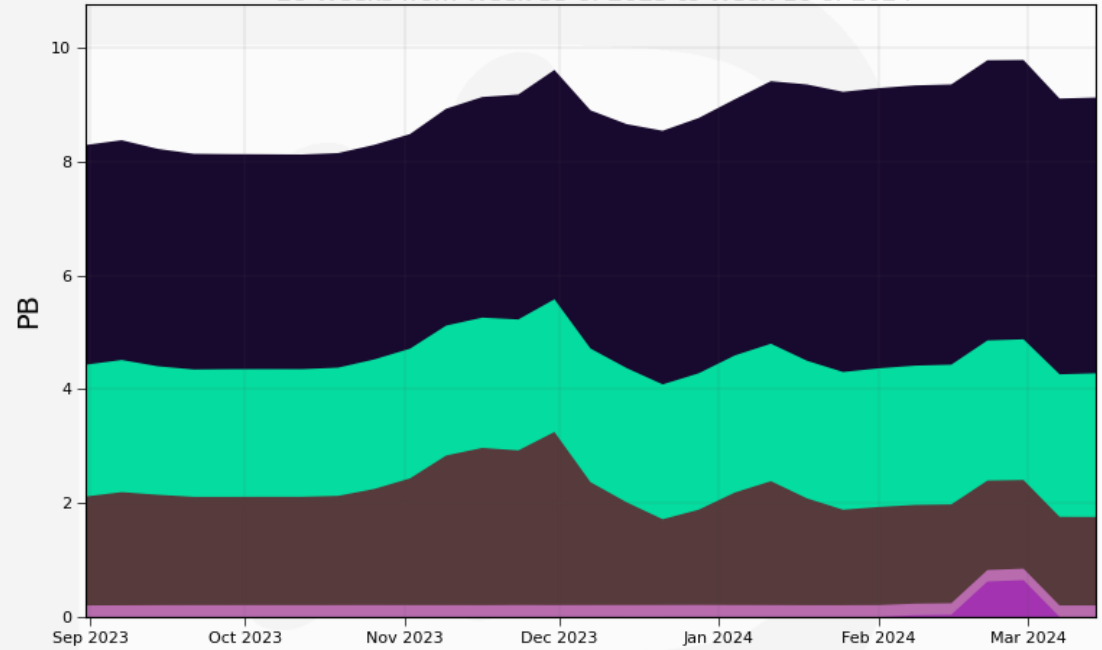
- Increased tape usage
 - 2023 data distribution
 - Evacuated 2016 data from RRCKI
- Still significantly below the pledge
- We are expecting much more data this year
- Pledge is going down next FY
 - 39PB to 33PB



Disk usage

PFN space usage by StorageElement

28 Weeks from Week 35 of 2023 to Week 10 of 2024



Max: 9.79, Min: 8.13, Average: 8.86, Current: 9.13

■ RAL-DST 48.3% ■ RAL_BUFFER 22.2% ■ RAL-DC-BUFFER 0.5%
■ RAL_MC-DST 26.6% ■ RAL-USER 2.3% ■ RAL-FAILOVER 0.0%

Generated on 2024-03-20 15:04:59 UTC

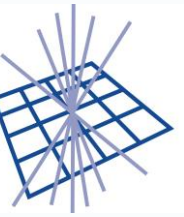
- Increased usage during reprocessing campaign
- Still below the pledge
- Pledge is going down next FY (15.7 -> 15.1PB)
- Discrepancy between WLCG and DIRAC accounting disappeared
 - Dark data has been removed

Status	Endpoint ↑	StorageEle...	% Free	Total (MB)	Guaranteed...	LastCheckTime
Active	gsiftp://gridftp.echo.stfc.ac.uk	RAL_MC-DST	41.62	14991760253.91	0	2024-03-21 11...

```

4:
  assignedendpoints:
    0: "GridFTP"
    1: "XRRootD"
  name: "LHCb-Disk"
  numberoffiles: -1
  path:
    0: "/lhcb:"
  timestamp: 1710979142
  totalsize: 1572000000000000
  usedsize: 9178022783545444
  vos:
    0: "lhcb"
  servingstate: "open"
    
```

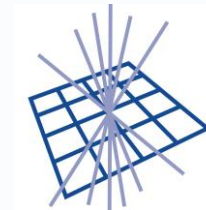
$$14991760253.91 * 1024^2 * (1 - 0.4162) = 9177336000002296$$



Problems: long term issues

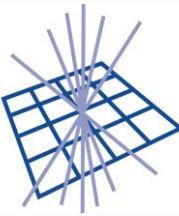
There were two long-lasting issues for LHCb:

1. Slow stat calls ([ticket](#)).
2. Vector read ([ticket](#)).



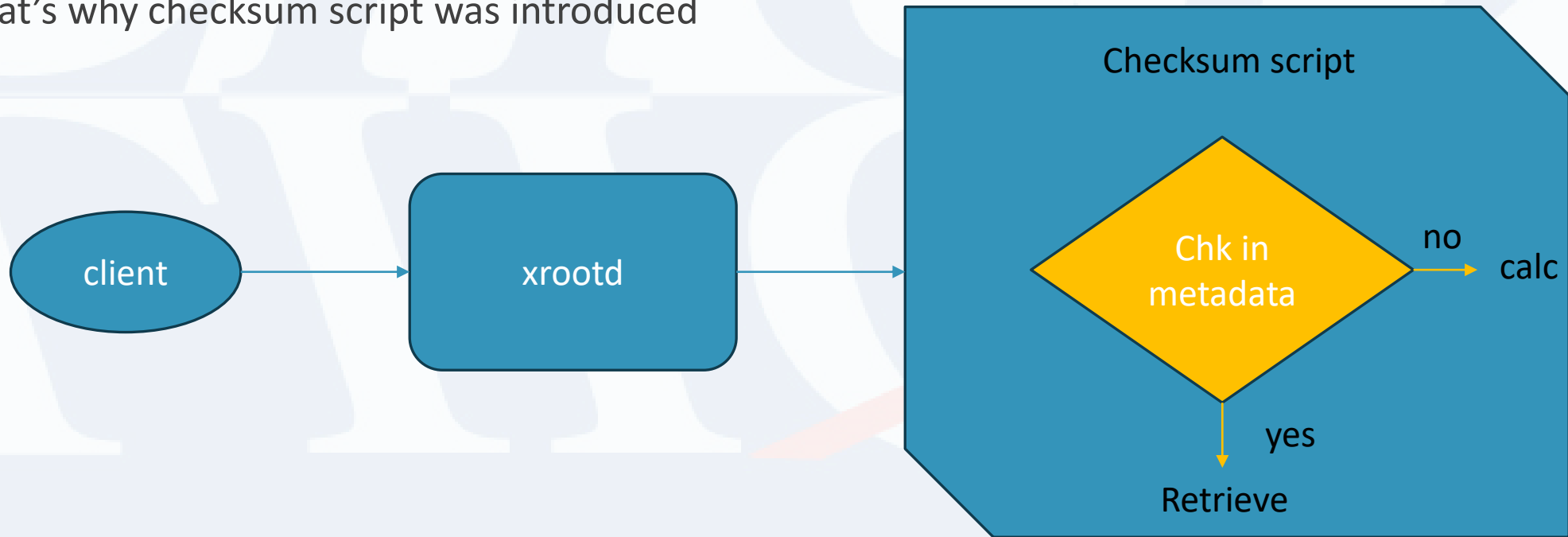
Slow “stat” calls

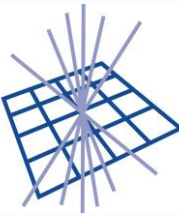
- A ticket was opened with a complaint about slow stat requests on ECHO
- The problem was spotted because some DIRAC workflows that test presence of datasets were slow
- It turned out that the problem is not with stats, but with checksum requests
 - I.e. requests to get the checksum of file that has been already written to the storage
 - Should be a simple attribute retrieval in most cases
 - DIRAC uses these requests to check file integrity (so it is stat + checksum for every file)



Slow “stat” calls: history

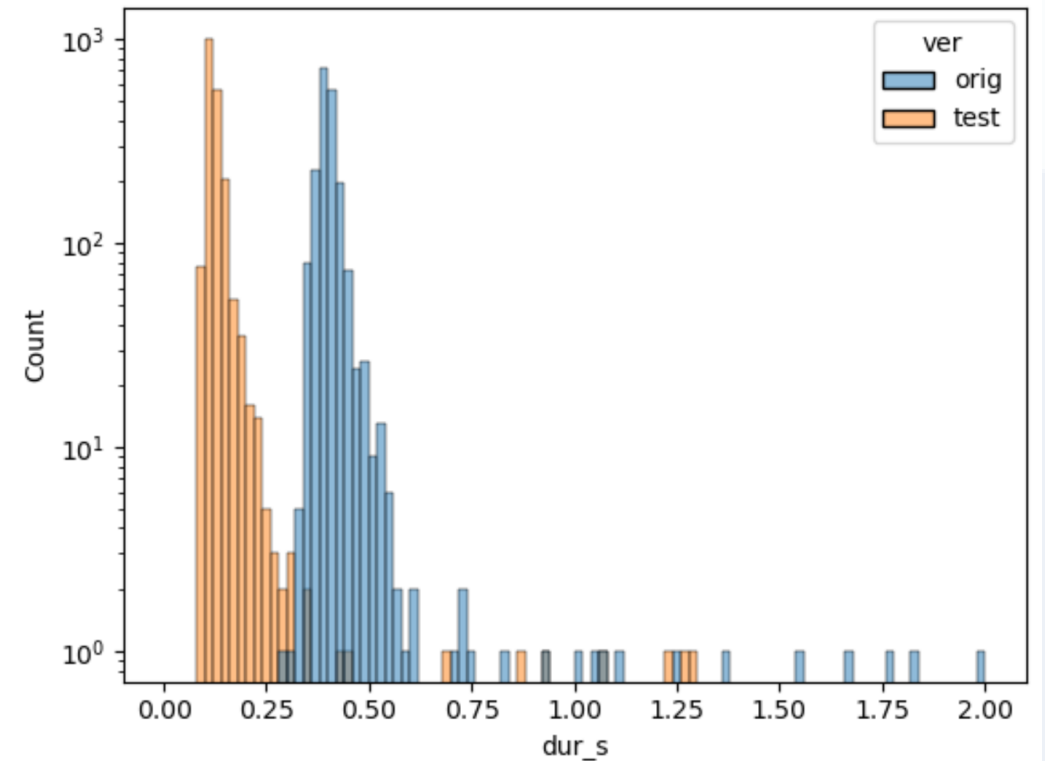
- Xrootd can handle checksum requests “automatically”, even with ofs plugins
- When ECHO had just entered production, performance of this mechanism was very poor
- That’s why checksum script was introduced

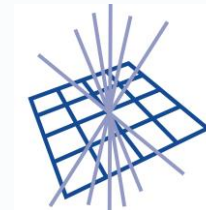




Slow “stat” calls: first attempt

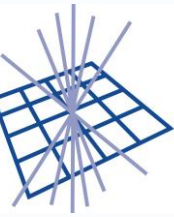
- A patch for checksum script was developed by James W to improve checksum retrieval speed
- During the non-prod tests on “empty” gateway performance was very good
- When it was deployed to production, results were not that great





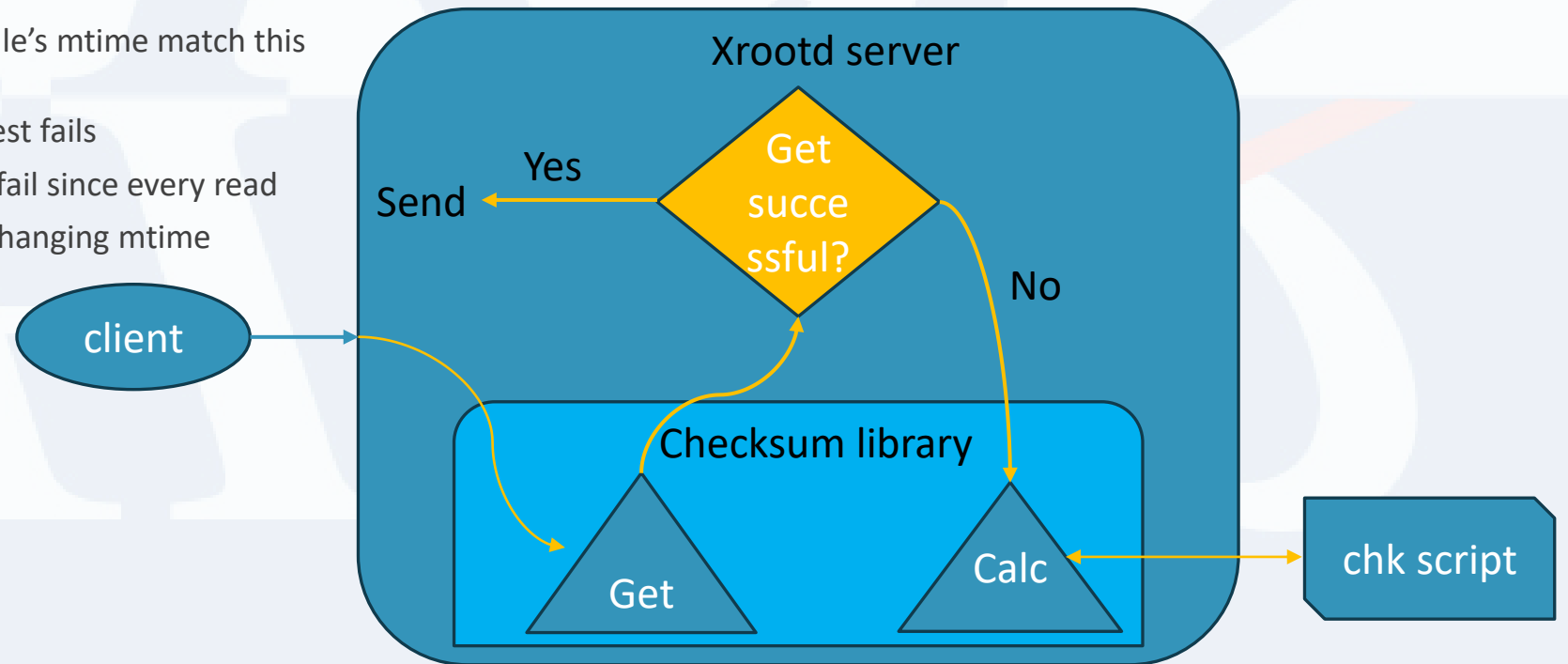
Slow “stat” calls: problem

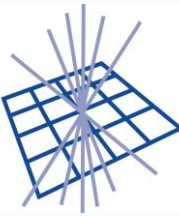
- Turned out that xrootd considers checksum script run as a time-consuming operation
 - make sense since one can not be sure that the checksum is present in the attributes
- That’s why whenever a request comes, checksum script is not executed right away but its run is scheduled
- Under heavy load this scheduling introduces significant delays
- xrootd tries to check whether checksum attribute is present or not (and retrieve it if it is)
 - Though this check is skipped when the checksum script is set up



Slow “stat” calls: attempt 2

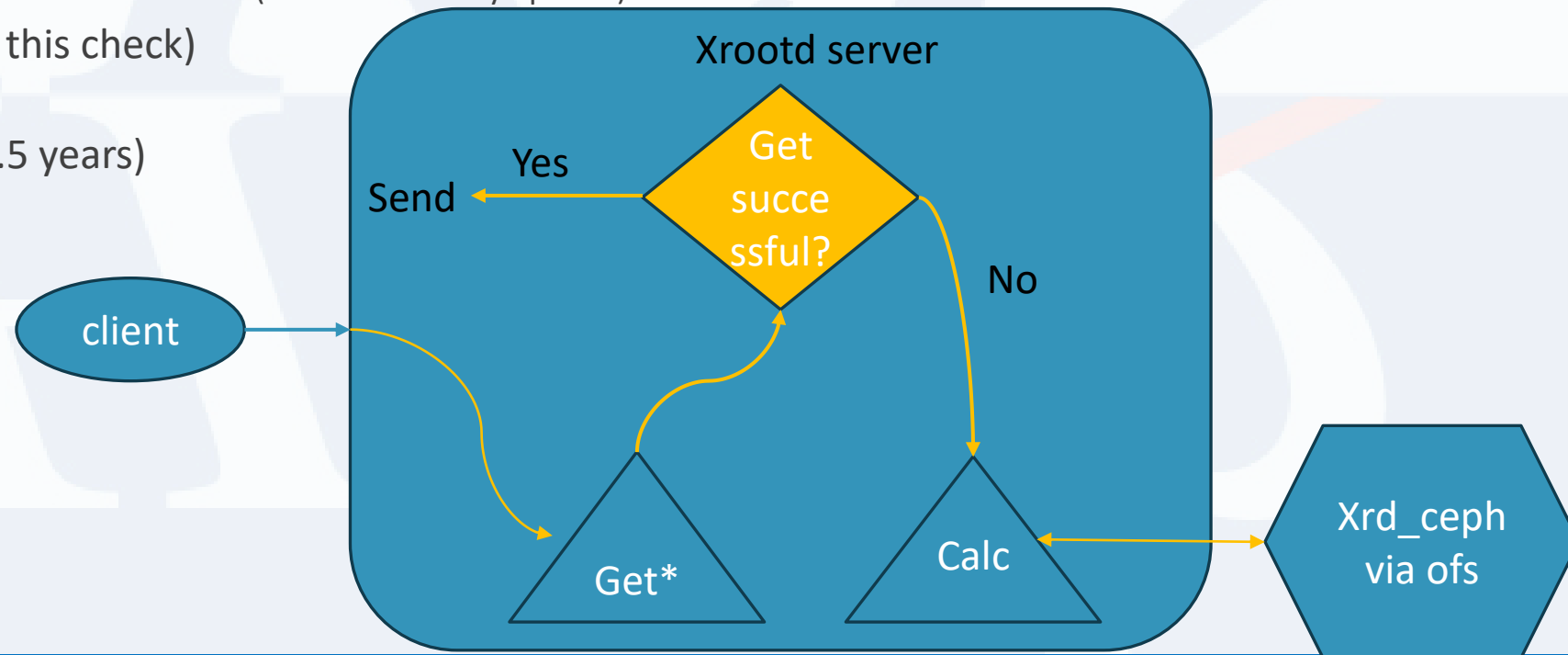
- The next attempt was to move “get or calc” logic to xrootd server itself, by implementing a checksum library
- Worked much better, but Get method had to be modified to remove **stale checksum check**
 - Xrootd writes checksum calculation time to the attribute
 - On retrieval it checks whether file’s mtime match this checksum calculation time
 - If there is a mismatch, get request fails
 - On ECHO this check will always fail since every read via libradosstriper writes lock, changing mtime

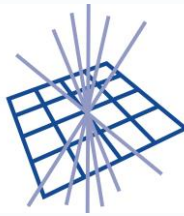




Slow “stat” calls: solution

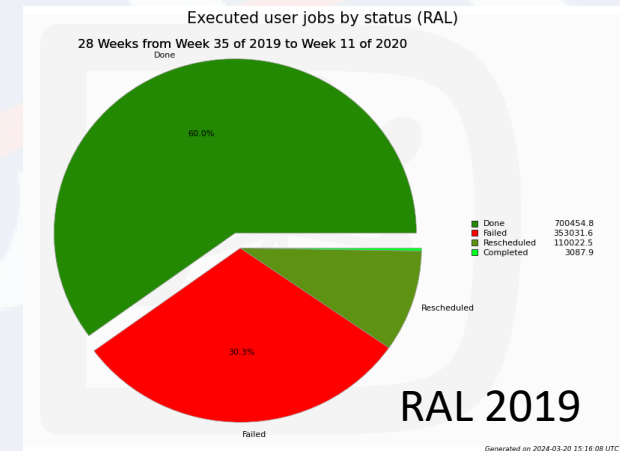
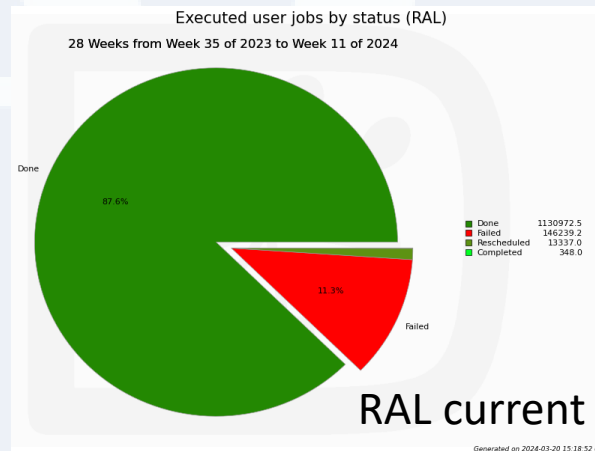
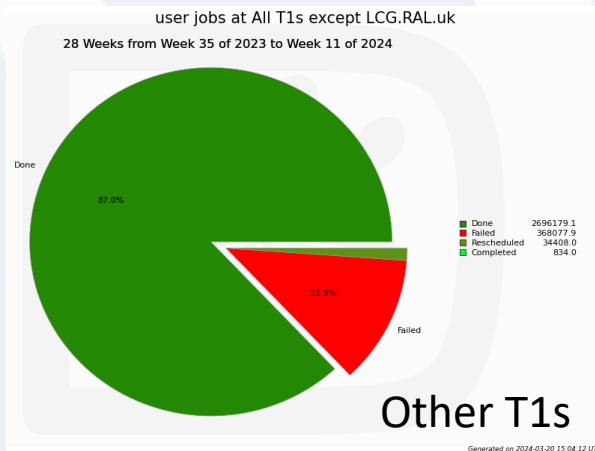
- After a discussion with xrootd devs it was found out (thanks to Jyothish) that Calc method is already implemented in xrootd
- Poor performance of the checksums in the early days of ECHO was caused by the stale checksum check
 - I.e. all checksums were re-calculated from scratch (at least that’s my opinion)
- Patching xrootd (removing this check) solves the problem
- Ticket was closed (after ~1.5 years)

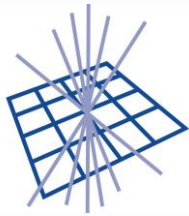




Vector read issue

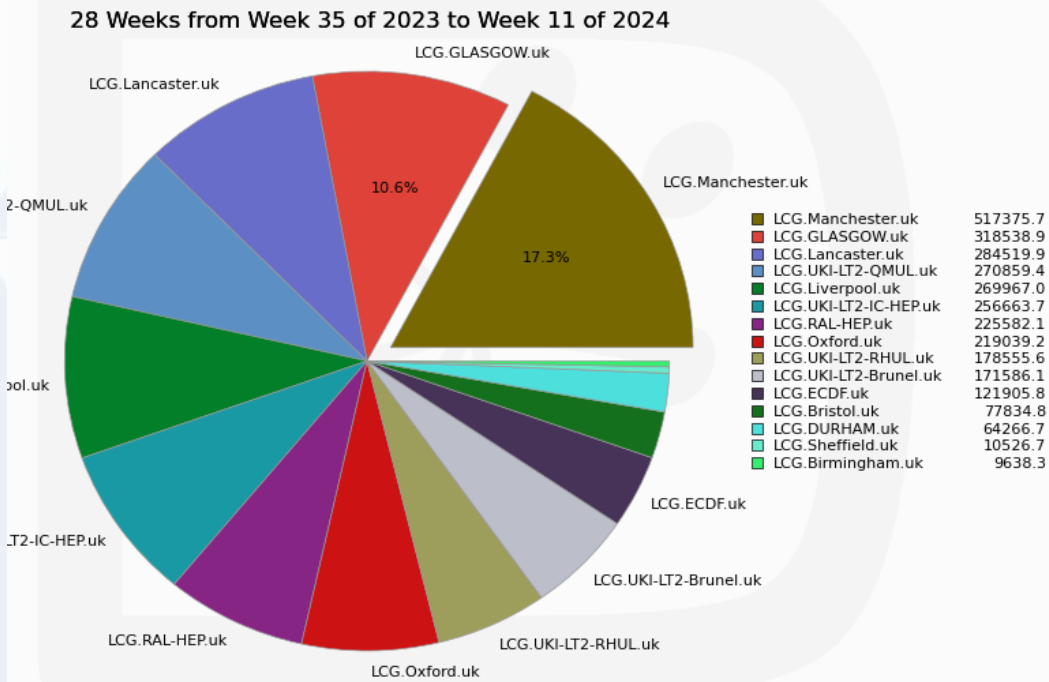
- Vector read operations on ECHO were slow
- It was improved by patching xrootd-ceph
- There is still some room for improvement: we can achieve better performance by turning prefetch off on local WN gateways
 - This change is scheduled
- However, even now the performance is much better than before





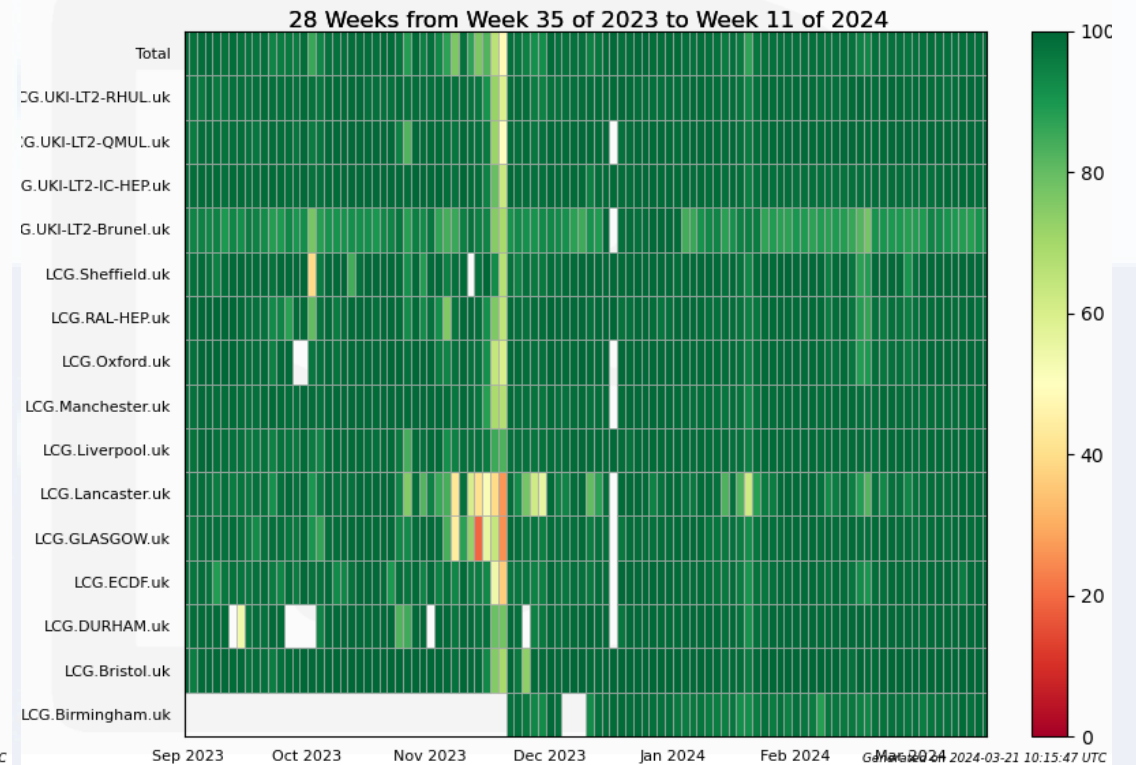
Tier-2 statistics

CPU days used by Site



Generated on 2024-03-21 10:19:13 UTC

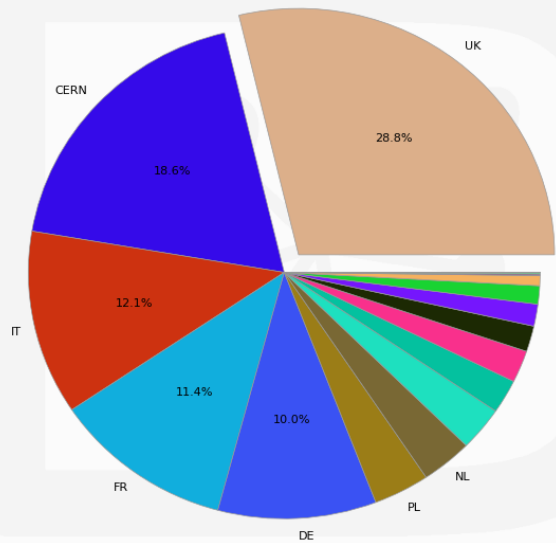
Job CPU efficiency by Site



Tier-2 statistics

Total Number of Jobs by Country

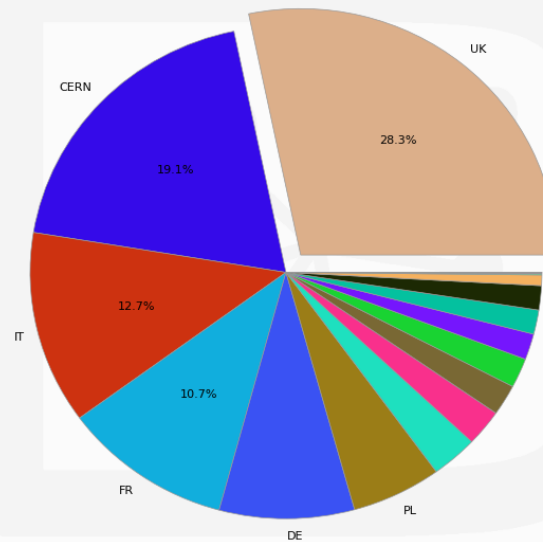
28 Weeks from Week 35 of 2023 to Week 11 of 2024



UK provides more computing resources than CERN!

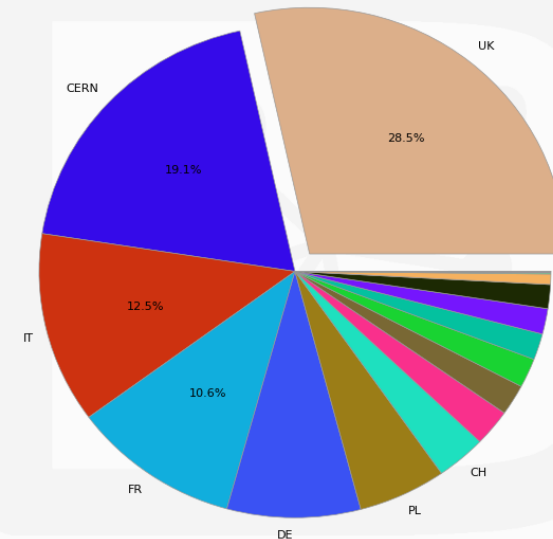
CPU days used by Country

28 Weeks from Week 35 of 2023 to Week 11 of 2024



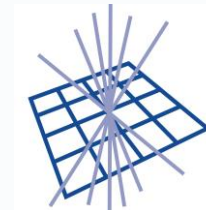
Wall time days used by Country

28 Weeks from Week 35 of 2023 to Week 11 of 2024



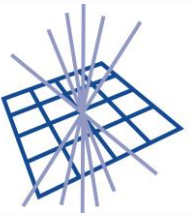
UK	6029974.5
CERN	4045343.9
IT	2657817.8
FR	2251004.6
DE	1788071.1
PL	1178031.9
CH	666920.3
RU	508926.3
NL	425842.9
BR	407234.7
ES	363675.7
US	354685.2
CN	333276.4
RO	144261.1
AU	20842.3
IL	12344.8
CR	2654.9
ANY	155.6
MULTIPLE	76.9

Generated on 2024-03-21 10:07:39 UTC



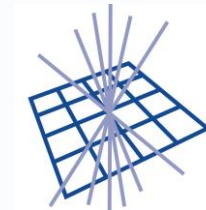
Tier-2D consistency

- Consistency checks were done for the following Tier-2Ds (hopefully, all UK Tier-2Ds, unless I've forgotten some site):
 - Liverpool
 - QMUL
 - Imperial
 - Manchester
 - RALPP
 - Glasgow
 - RAL-T1
- Lost files restored (where possible), some dark data removed



Plans

- 2024 is expected to be the first “normal” data taking year
 - We have ambitious plans to take a lot of data! So resource usage increase expected
- Tokens
 - Tokens were tested during DC24, results are not very good
 - Tokens were not used for RAL
 - Because of the xrootd peculiarities, should be fixed now
 - In general, we are not ready to use tokens for storage in production
 - Condor CEs are set up to use tokens, experimenting with some arc CEs as well (e.g. at RALPP)
- We have new [ETF test suite](#), and it includes storage tests (finally!)
 - Still in preprod, compare to prod [here](#)
 - Token tests to be added
 - Migration expected soon



New Tier-1 sites

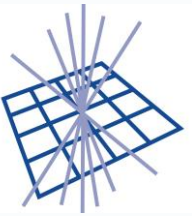
- Two new LHCb sites were/to be promoted to Tier-1
 - NCBJ in Swierk, Poland
 - Officially Tier-1 now
 - T1 only for LHCb, T2 for CMS
 - Should receive 7.30% of RAW data
 - IHEP in Beijing, China
 - Proto Tier-1, to be promoted soon
 - Should receive 7.45% of RAW data



Narodowe Centrum Badań Jądrowych
National Centre for Nuclear Research
ŚWIERK

JRC collaboration partner





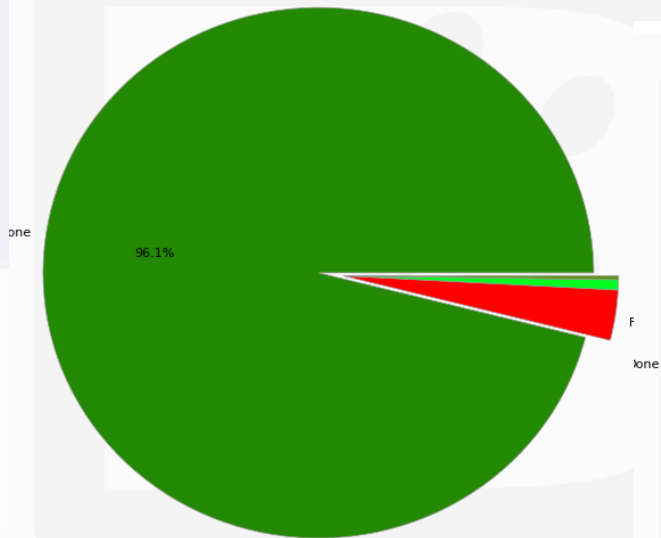
Summary

- UK provided a lot of resources to LHCb during the last 6 months
 - Around 30%, more than CERN!
- Relatively smooth operations
- Long lasting issues are almost completely resolved

RAL T1 Job success rate (backup)

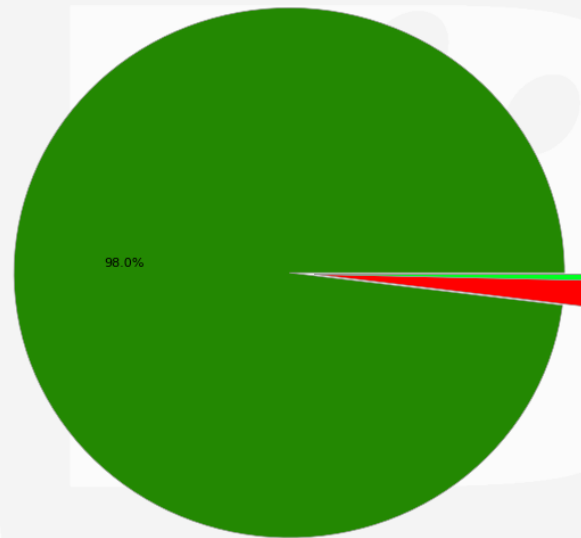
RAL jobs by Status

28 Weeks from Week 35 of 2023 to Week 11 of 2024



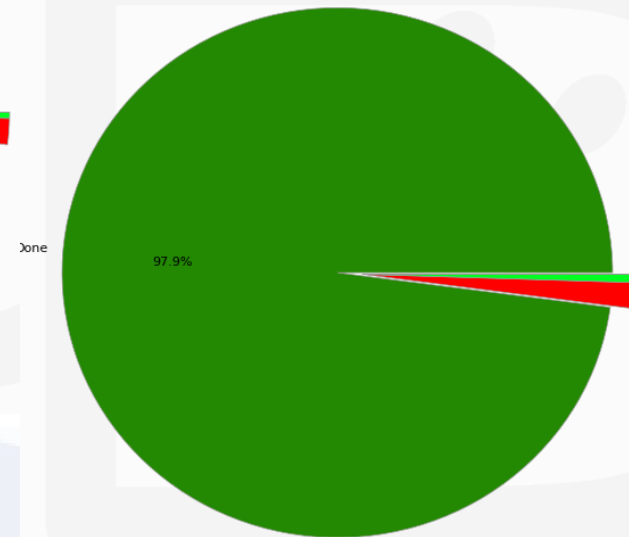
RAL CPU time consumed by status

28 Weeks from Week 35 of 2023 to Week 11 of 2024



RAL wall time consumed by status

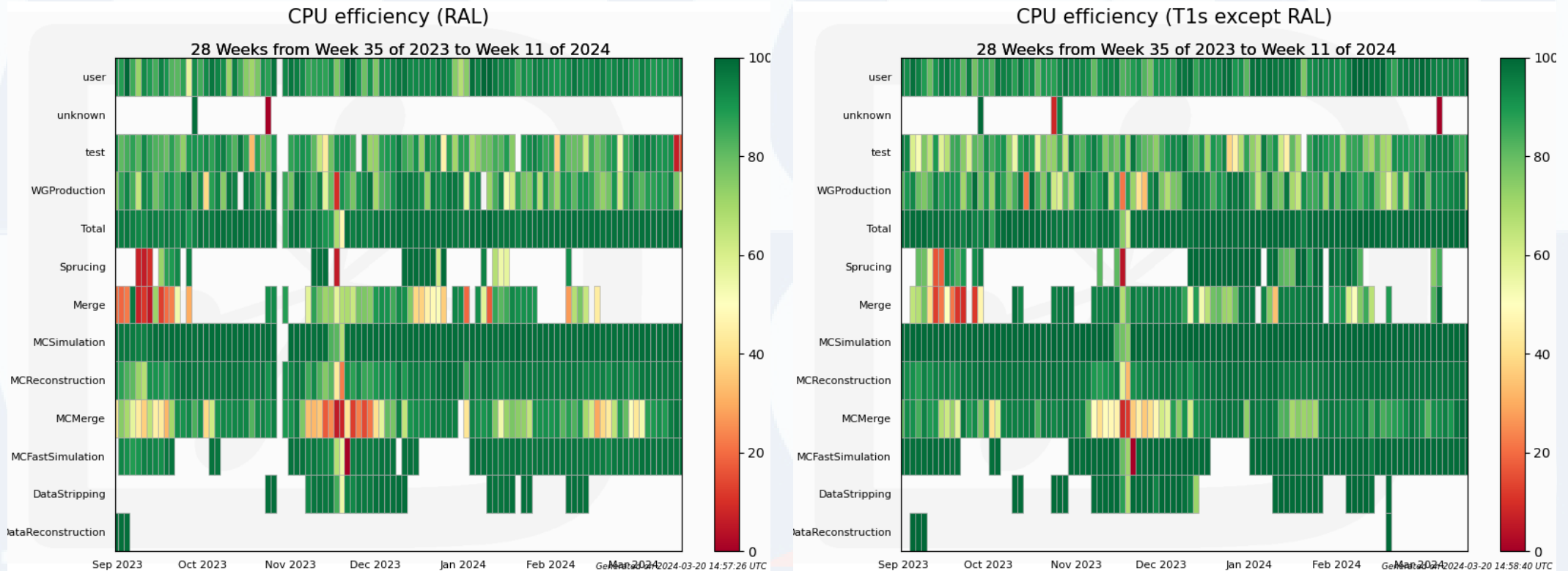
28 Weeks from Week 35 of 2023 to Week 11 of 2024



Done	2850880.4
Failed	46029.3
Completed	14577.1
Rescheduled	138.8

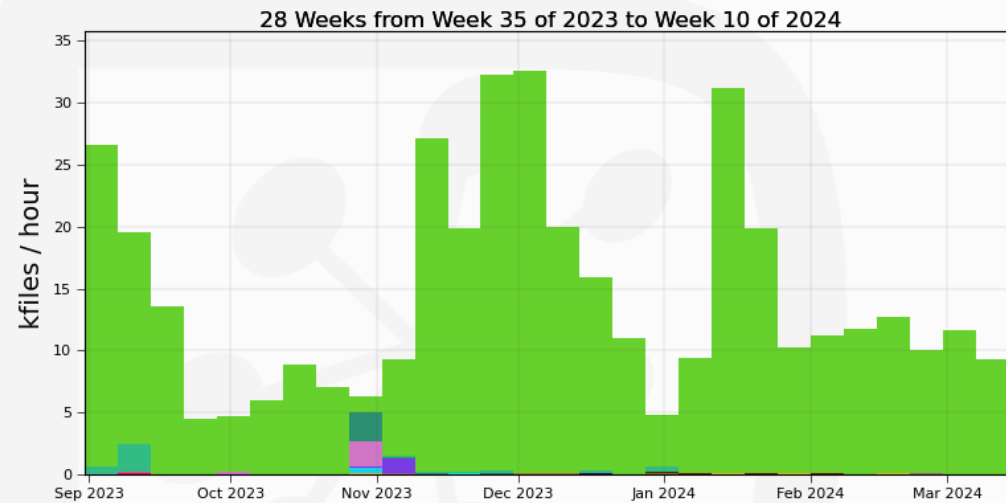
Generated on 2024-03-20 14:59:36 UTC

RAL T1 CPU efficiency (backup)



Transfers to ECHO (backup)

Transfers to ECHO

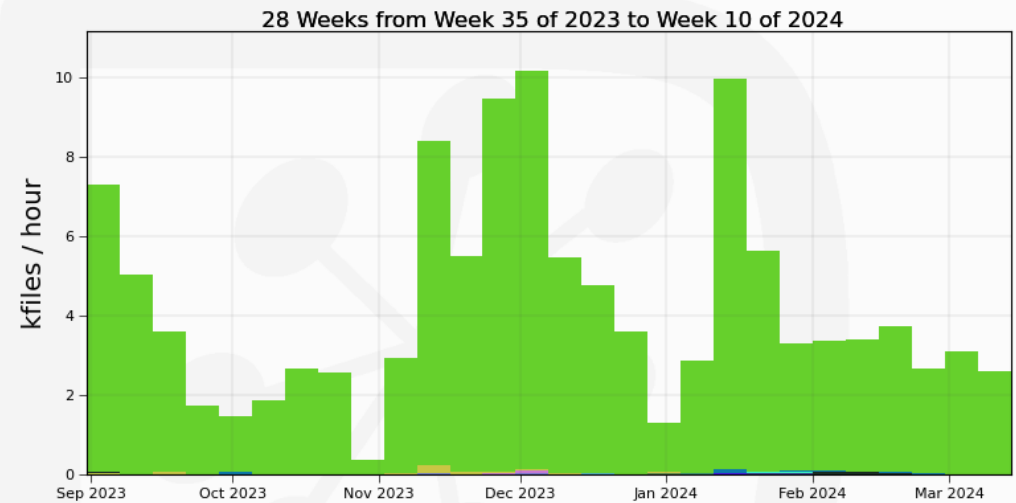


Max: 32.5, Min: 4.49, Average: 14.3, Current: 5.98

Succeeded	96.9%	DIRAC.Client.ch -> RAL-USER	0.0%
LCG.RAL.uk -> RAL-BUFFER	1.0%	CNAF_MC-DST -> RAL_MC-DST	0.0%
RRCKI-RDST -> RAL-BUFFER	0.6%	DIRAC.Client.cn -> RAL-USER	0.0%
lbvobox310.cern.ch -> RAL-BUFFER	0.5%	DIRAC.Client.uk -> RAL-USER	0.0%
CERN-RAW -> RAL-BUFFER	0.3%	LCG.Glasgow.uk -> RAL-BUFFER	0.0%
RRCKI-RAW -> RAL-BUFFER	0.1%	LCG.RAL.uk -> RAL_MC-DST	0.0%
LCG.Beijing.cn -> RAL-BUFFER	0.1%	GRIDKA-DST -> RAL-DST	0.0%
LCG.RAL-HEP.uk -> RAL-BUFFER	0.0%	LCG.GRIDKA.de -> RAL-BUFFER	0.0%
LCG.RAL.uk -> RAL-USER	0.0%	... plus 261 more	

Generated on 2024-03-20 15:04:48 UTC

Transfers from ECHO



Max: 10.1, Min: 0.36, Average: 4.16, Current: 1.71

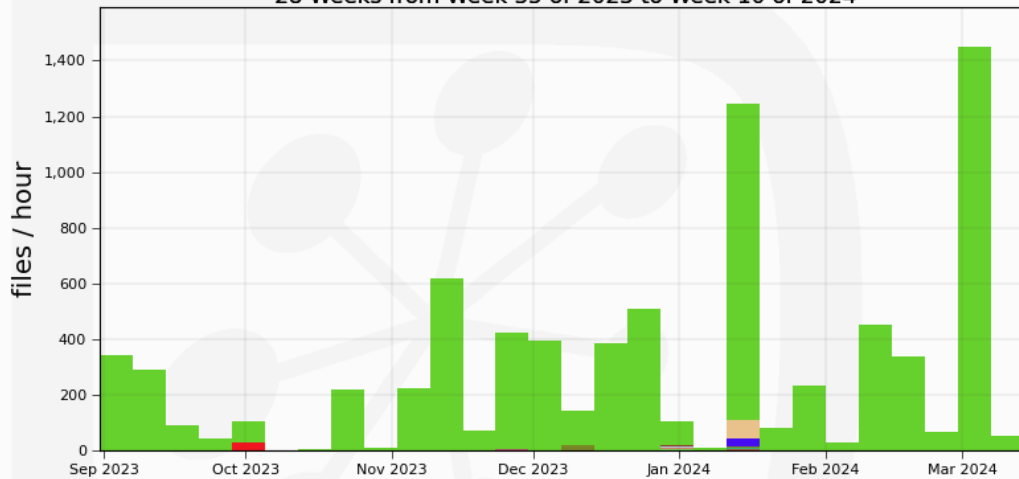
Succeeded	98.8%	RAL-USER -> LCG.Beijing.cn	0.0%
RAL-BUFFER -> LCG.RAL.uk	0.4%	RAL-USER -> LCG.RAL.uk	0.0%
RAL-USER -> DIRAC.Client.ch	0.3%	RAL-USER -> LCG.CERN.cern	0.0%
RAL-USER -> DIRAC.Client.cn	0.1%	RAL-DST -> CERN-ARCHIVE	0.0%
RAL-USER -> DIRAC.Client.de	0.1%	RAL-DST -> DIRAC.Client.ch	0.0%
RAL-DST -> CNAF-ARCHIVE	0.1%	RAL_MC-DST -> CERN-ARCHIVE	0.0%
RAL-USER -> DIRAC.Client.uk	0.0%	RAL-BUFFER -> LCG.GRIDKA.de	0.0%
RAL_MC-DST -> IHEP_MC-DST	0.0%	RAL-USER -> DIRAC.Client.es	0.0%
RAL-BUFFER -> RAL-RAW	0.0%	... plus 190 more	

Generated on 2024-03-20 15:04:57 UTC

Transfers to Antares (backup)

Transfers to Antares

28 Weeks from Week 35 of 2023 to Week 10 of 2024



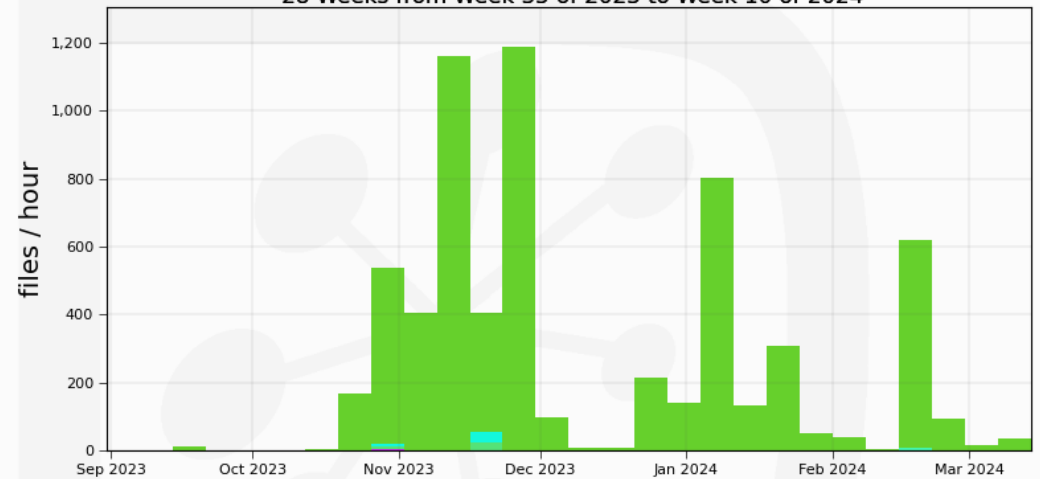
Max: 1,447, Min: 0.73, Average: 275, Current: 49.9

Succeeded	97.6%	LCG.RAL.uk -> RAL-ARCHIVE	0.0%
CERN-DAQ-EXPORT -> RAL-RAW	0.8%	LCG.GRIDKA.de -> RAL-ARCHIVE	0.0%
RAL-BUFFER -> RAL-RAW	0.4%	CNAF-DST -> RAL-ARCHIVE	0.0%
lbvobox309.cern.ch -> RAL-ARCHIVE	0.4%	LCG.IN2P3.fr -> RAL-ARCHIVE	0.0%
GRIDKA-DST -> RAL-ARCHIVE	0.3%	PIC-DST -> RAL-ARCHIVE	0.0%
RAL-BUFFER -> RAL-RDST	0.1%	LCG.NIKHEF.nl -> RAL-ARCHIVE	0.0%
LCG.CERN.cern -> RAL-ARCHIVE	0.1%	LCG.CNAF.it -> RAL-ARCHIVE	0.0%
RAL-DST -> RAL-ARCHIVE	0.1%	LCG.NCBJ.pl -> RAL-ARCHIVE	0.0%
CERN-DST-EOS -> RAL-ARCHIVE	0.0%	... plus 43 more	

Generated on 2024-03-20 15:05:08 UTC

Transfers from Antares

28 Weeks from Week 35 of 2023 to Week 10 of 2024



Max: 1,186, Average: 225, Current: 72.6

Succeeded	98.7%	RAL-RAW -> DIRAC.Client.ch	0.0%
RAL-RAW -> RAL-BUFFER	0.6%	RAL-ARCHIVE -> DIRAC.Client.de	0.0%
RAL-RDST -> RAL-BUFFER	0.5%	RAL-ARCHIVE -> DIRAC.Client.cn	0.0%
RAL-DC-RAW -> RAL-DC-BUFFER	0.1%	RAL-RAW -> DIRAC.Client.it	0.0%
RAL-ARCHIVE -> CERN_MC-DST-EOS	0.1%	RAL-RDST -> LCG.GRIDKA.de	0.0%
RAL-ARCHIVE -> DIRAC.Client.ch	0.1%	RAL-ARCHIVE -> DIRAC.Client.it	0.0%
RAL-RAW -> CNAF-BUFFER	0.0%	RAL-ARCHIVE -> DIRAC.Client.edu	0.0%
RAL-RAW -> LCG.RAL.uk	0.0%	RAL-RAW -> DIRAC.LocalProdTest.local	0.0%
RAL-RDST -> LCG.RAL.uk	0.0%	... plus 18 more	

Generated on 2024-03-20 15:05:16 UTC