

# UK mini-DC 2024: Tier 2s

Katy Ellis for **CMS**, DOMA General, 15th Jan 2025

# Why national mini-challenges?

- Many Tier 2 sites did not receive a lot of attention from the experiments during DC24
  - Of course, sites may have been on high-alert, but if rates were low then it may have been difficult to spot bottlenecks and make improvements
- In the UK, several sites have made significant changes since DC24
  - New Data Centre, new storage technology, etc.
- Tier 1s may also have changes or improvements to test
- Maintain attitude of continuous improvement
  - Prepare the system for the third data challenge

# ATLAS and CMS UK testing

- December 2024 was used for Tier 2 tests
- Tier 1 was used as a reliable source or sink, but not ‘tested’
- This talk will mention CMS site tests only
  - Results from ATLAS can be found [here](#) and will be reported at a later date
  - Equivalent notes for CMS are [here](#)
- CMS tests were performed between 9th and 13th December
- Tier 1 tests will be planned for Feb or March, possible examples:
  - Hitting DC24 rates for writes
  - What are the new limits of the system?
  - Tape testing (not in DC24 scope, except LHCb)
    - Plan to have new tape front end, and connected to LHCOPN
    - Help the experiments learn how to best test tape endpoints

# CMS Tier 2 sites

- T2\_UK\_London\_IC (Imperial College London)
  - 100Gbps link, CMS Tier 2
- T2\_UK\_SGrid\_RALPP (RAL T2 site)
  - 100Gbps link, Tier 2 for ATLAS, CMS (+many more)
- T2\_UK\_London\_Brunel
  - Requested to be write-tested up to 30Gbps and 40Gbps if possible
  - New storage technology adopted (Ceph replaced DPM)
- (T2\_UK\_SGrid\_Bristol)
  - Small site, not tested as unavailable during test week

# Test plan

Use the `dc_inject` tool to create a continuous flow of data with a specific target rate

- Monday - write tests using RAL T1 as source
  - RALPP to be switched to different source
- Tuesday - read tests
- Wednesday - more higher rate tests
- Thursday - write tests at DC24 target rate using CERN as source
- Friday - free for follow up/repeat tests

# DC24 targets

- Writing at T2s

DC24 rates and initial test setup:

Source	Dest	Mb/s	GB/s
T1_UK_RAL_Disk	T2_UK_London_IC	24341	3.042625
T1_UK_RAL_Disk	T2_UK_London_Brunel	3842	0.48025
T1_UK_RAL_Disk	T2_UK_SGrid_RALPP	6067	0.758375

# DC24 targets

- Writing at T2s

DC24 rates and initial test setup:

Source	Dest	Mb/s	GB/s
T1_UK_RAL_Disk	T2_UK_London_IC	24341	3.042625
T1_UK_RAL_Disk	T2_UK_London_Brunel	3842	0.48025
<del>T1_UK_RAL_Disk</del>	T2_UK_SGrid_RALPP	6067	0.758375

T2\_UK\_London\_IC



If these rates are easily reached then push for higher rates as requested by sites

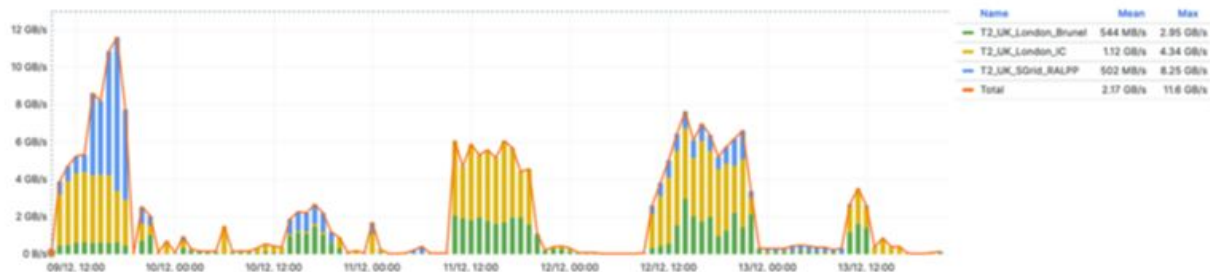
# DC24 targets

- Reading from Tier 2s

Source	Dest	Mbit/s	GB/s
T2_UK_London_IC	T1_UK_RAL_Disk	7210	0.90125
T2_UK_London_Brunel	T1_UK_RAL_Disk	744	0.093
T2_UK_SGrid_RALPP	T2_UK_London_IC	1774	0.22175

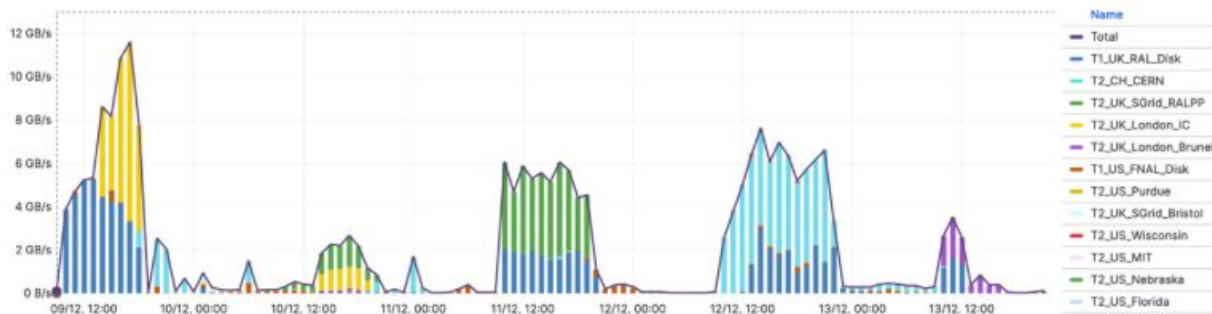


# Summary of results



FTS transfer monitoring for the test week, by destination RSE

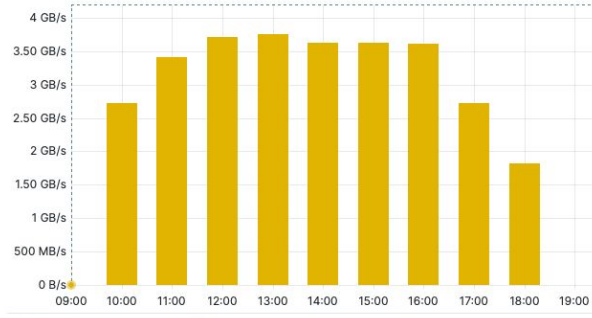
Mon      Tues      Wed      Thur      Fri



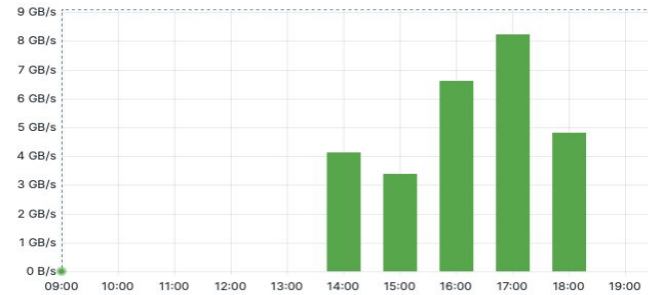
Same plot as above, but by source RSE

# Results - IC

- IC was easily able to sustain the rates requested by DC24 for read and write
- Simultaneous writes (left) and reads (right)



DC traffic only, Imperial College as destination, holding the rate above requested, with a slight drop-off when the read rate (data replicating to RALPP) was its highest.



DC reads from IC, transfers going to RALPP

- Pushed IC over their theoretical limit...

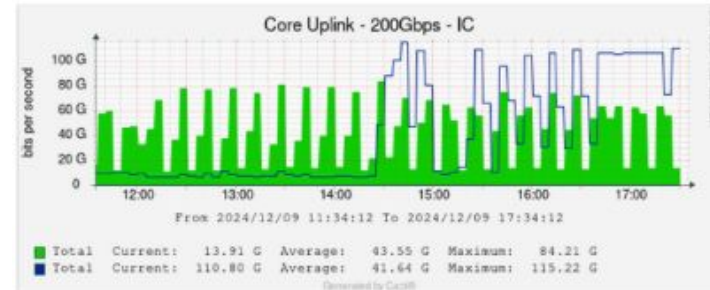


Image from Daniela. Green is writes and blue is reads

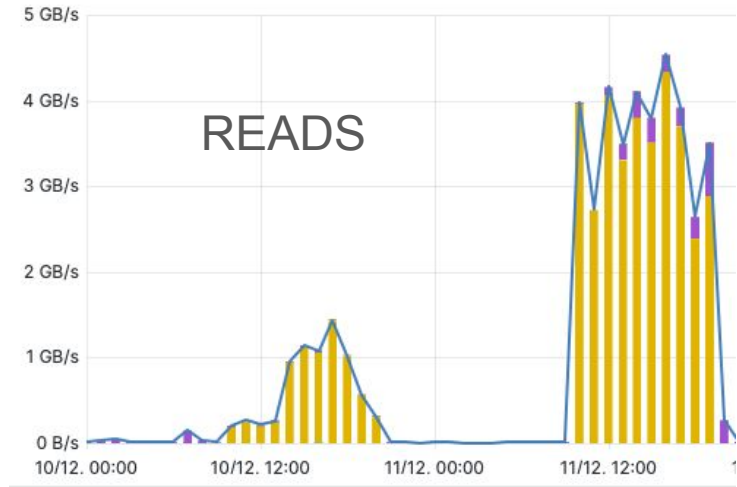
# Results - RALPP

- RALPP was also easily able to sustain the rates requested by DC24 for read and write
- Was then pushed harder

Transfer Throughput



Transfer Throughput



# Results - RALPP

- A small adjustment to the number of possible connections allowed RALPP to fill their 100Gbps bandwidth

We can also view RALPP activity in the nice new network monitoring:



Traffic of RALPP, mostly ingress. The earlier spiky behaviour is due to the site absorbing data faster than the tool is injecting. The final hour is where the data injected is high enough to create a sustained plateau.

# Results - Brunel

- Brunel was able to hit the requested DC24 rates
- The next stage was to push for 30Gbps
- A sustained rate around 15Gbps was achieved
- However, significant failures were seen
  - FTS Optimiser started see-sawing up and down on the 'decision' over how many concurrent transfers to do (see plot, above right)
- From the site point of view the bandwidth was filled...but this included a large number of failing transfers
- The next day, the number of concurrent transfers allowed in FTS was manually limited to 300

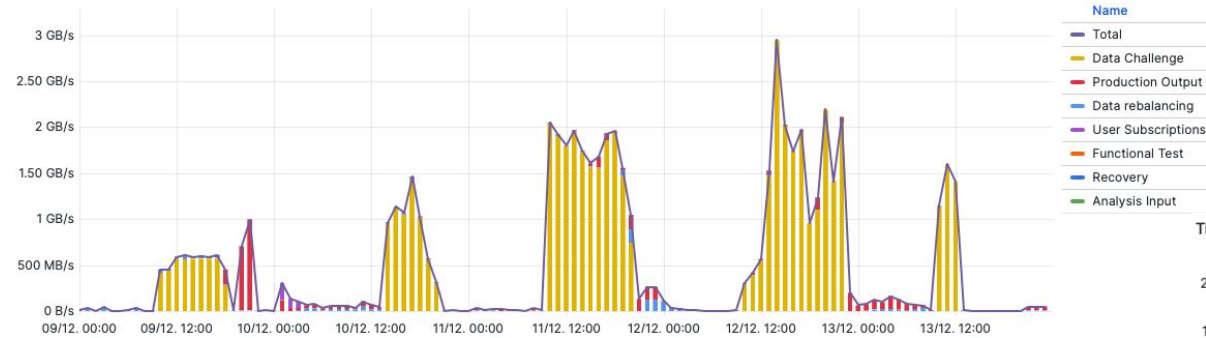


FTS failures during a short period of the test, overlaid with the 'decision' made by FTS on number of parallel transfers.

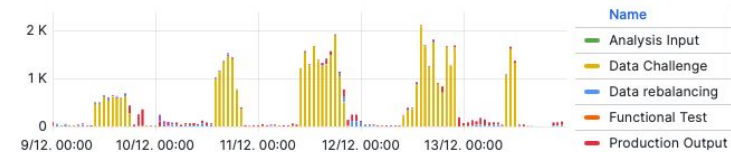
# Results - Brunel

- The limit on the number of concurrent transfers helped reduce the number of failures and increase the throughput on Thursday

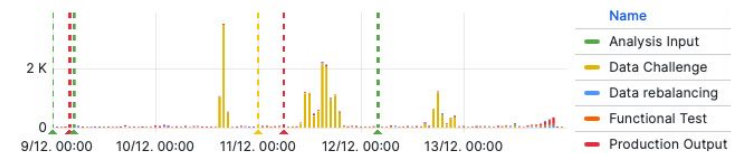
Transfer Throughput



Transfer Successes



Transfer Failures



# Results - Brunel

- On Friday, tested simultaneous ingress/egress
- Started well (green - writes at Brunel; yellow reads from Brunel)



Later, the site 'tweaked' some parts of the stack and broke things :-)

There are further plans for improvements at the site in the new year, so further tests will be welcome

# Results - Brunel



Network activity at Brunel throughout the test week (blue is writes, red is reads)



# Comments

- Few failures observed other than those mentioned at Brunel at higher load
- Each site was able to hit the DC24 target write rate with data source from CERN

# Summary

- Each of the three Tier 2 sites tested were able to sustain the read/write rates requested by CMS in DC24
  - RAL Tier 1 and CERN-EOS both used as sources
- Sites were also pushed to their limits, as requested by the sites
  - Very much an attitude of “see what it can do!”
- IC and RALPP were able to fill their available bandwidth
  - No problems seen with excessive failures or contention between ingress and egress
- Brunel still some work to do to fill their bandwidth (40Gbps)
  - No obvious issues with simultaneous ingress/egress