Tier-1 UK Mini-DC Plans

Alastair Dewhurst
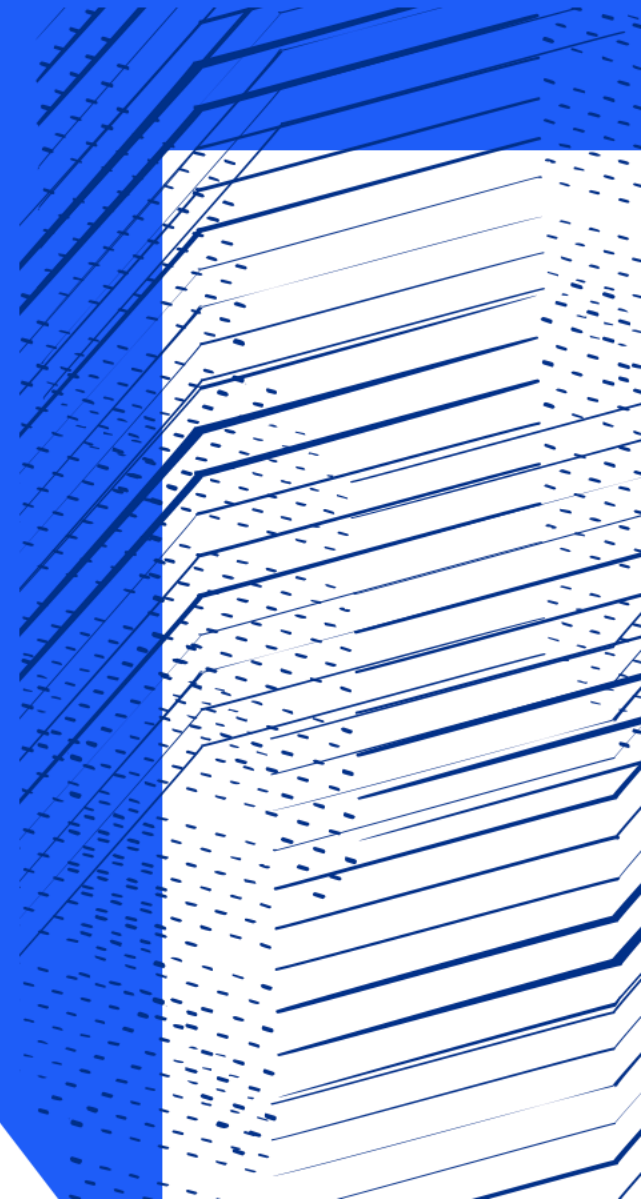
Requirements

- Table taken from Alessandra's spreadsheet.
  - DC24 level = 25% HL-LHC
- In [December there will be an OTF meeting](#) on how pledges need to include performance.
- Some tests are already reaching performance limits.
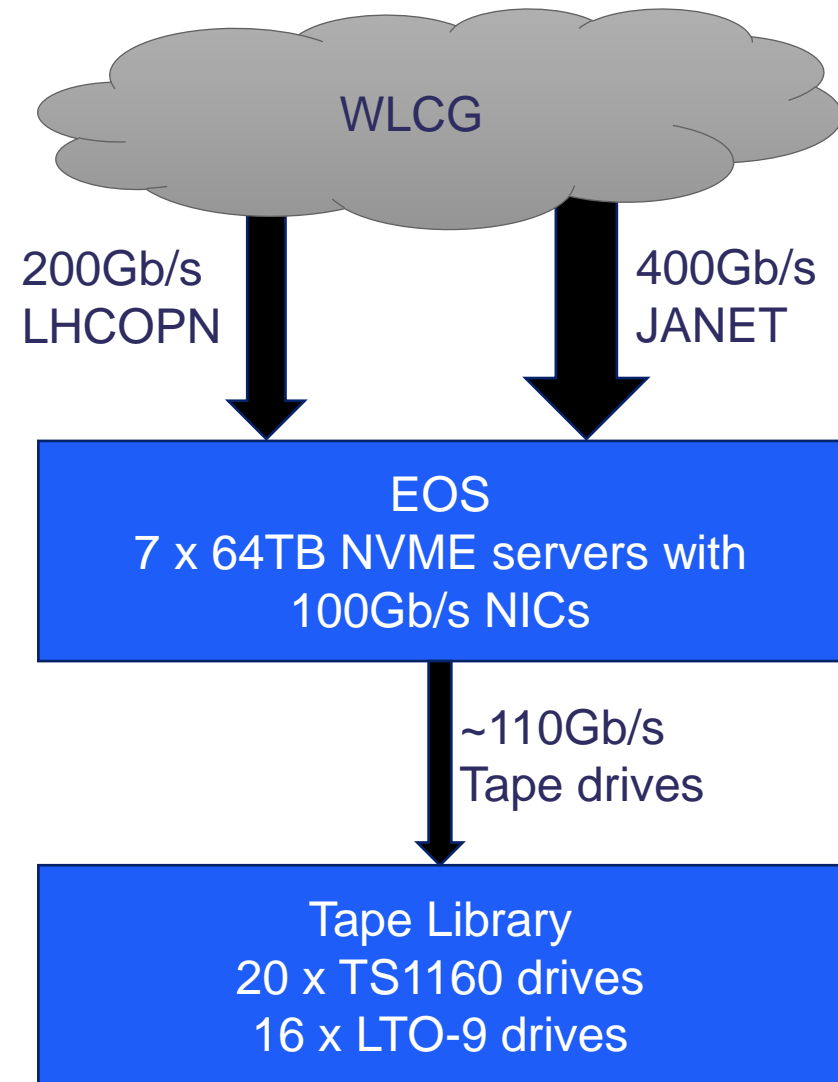 - New technology doesn't improve situation.

Throughput GB/s	Echo (Disk)	Antares (Tape)
Ingress ATLAS	7	7.45
Egress ATLAS	9.5	2.7
Ingress CMS	1.64	2.51
Egress CMS	3.55	2.51
Ingress LHCb	3.96	3.96
Egress LHCb	2.4	2.4
Ingress Total	**12.6**	**13.92**
Egress Total	**15.45**	**7.61**

https://docs.google.com/spreadsheets/d/1agJoNSDISoSB2xwL3HVdJdjGo2JVkeqzb9Ue3yUck10/edit?gid=0#gid=0

Alastair Dewhurst, 15th November 2024

Antares

- Plan to replace current EOS buffer.
  - New instance will be on LHCONE and LHCOPN.

- Due to procurement moratorium, servers are still being ordered.

- Best case scenario EOS is ready for late February 2025.

WLCG

200Gb/s
LHCOPN

400Gb/s
JANET

EOS
7 x 64TB NVME servers with
100Gb/s NICs

~110Gb/s
Tape drives

Tape Library
20 x TS1160 drives
16 x LTO-9 drives

Alastair Dewhurst, 15th November 2024
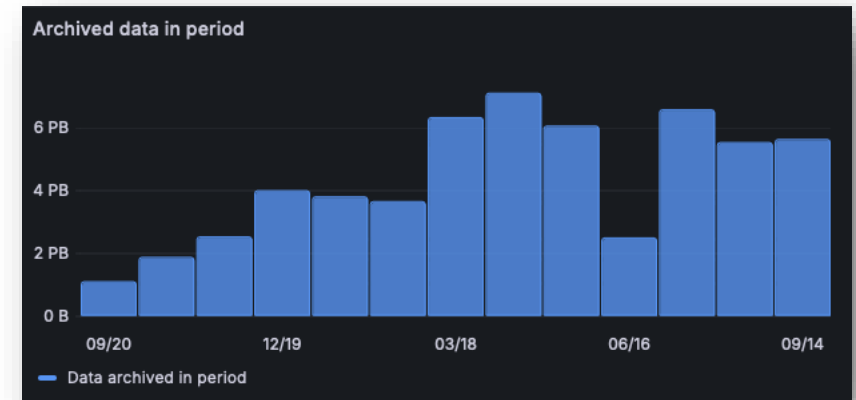
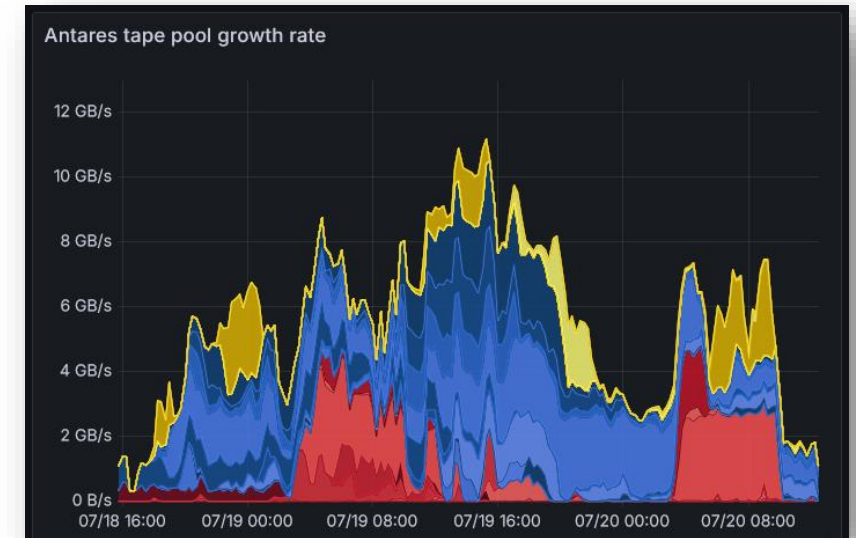GridPP
UK Computing for Particle Physics

Tape Throughput

- We have written a huge amount of data to Tape this year.

- We do NOT need to do a sustained write test for Antares.

- We do not have the tape drives available to meet 14GB/s…
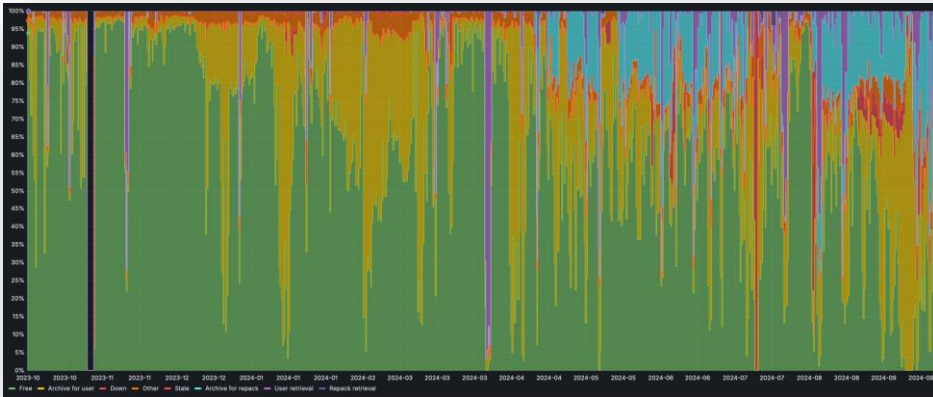
Green = Free
Yellow = Archival
Purple = Retrieval
Blue = Repack
Red/Orange = Down

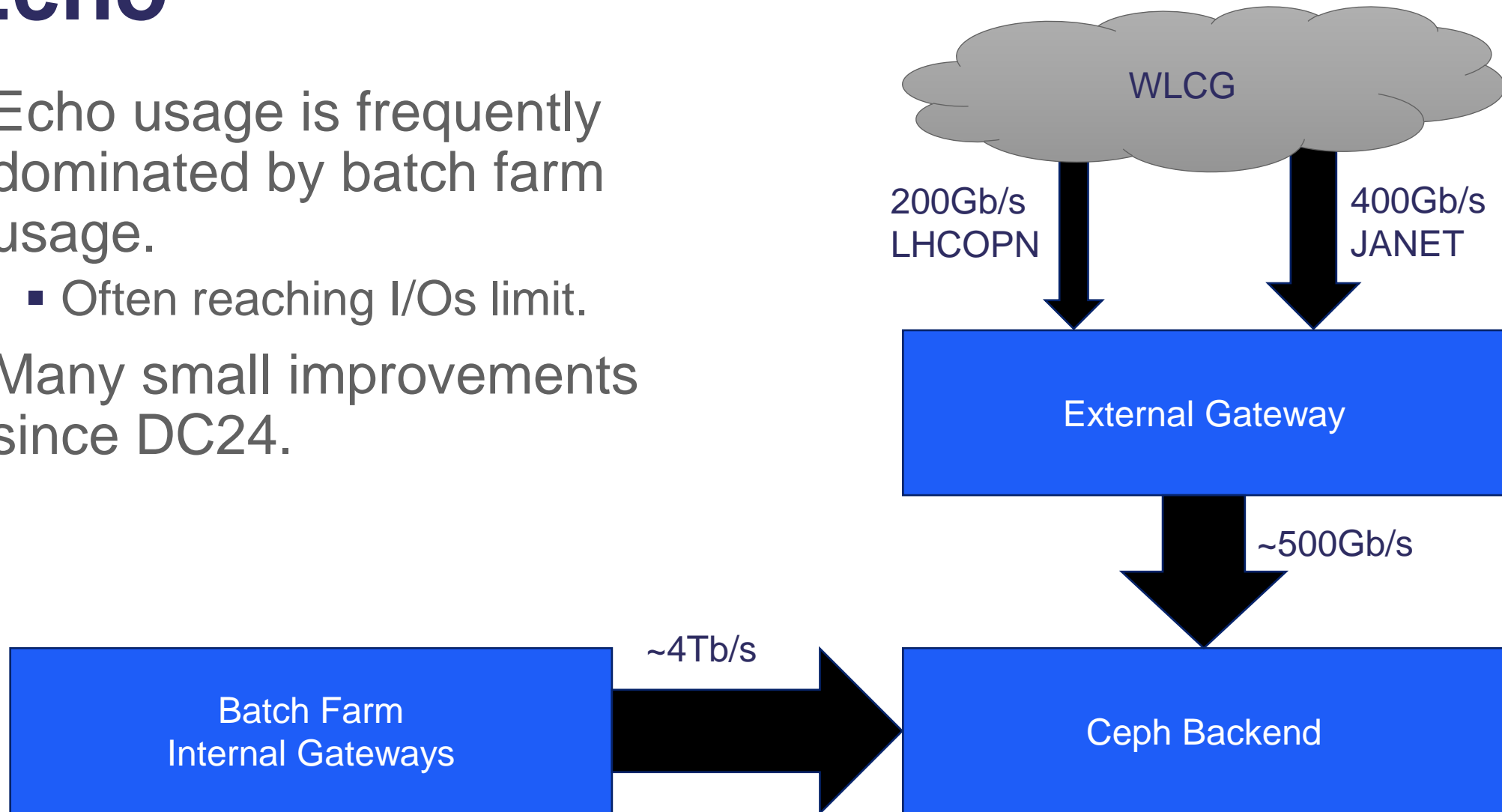Alastair Dewhurst, 15th November 2024

Tape tests

- We want to be able to test the connectivity of the new EOS instance:
  - Several short (~100TB) write test from various sites (e.g. CERN, other Tier-1s).
    - Goal is to comfortably exceed what we can do with Tape drives.
- We want to run a sustained read test (~1PB)
  - Requests need to be from multiple VOs / Tape Families.
    - Goal is to see what kind of sustained efficiency we can get from reading.

Echo

- Echo usage is frequently dominated by batch farm usage.
  - Often reaching I/Os limit.
- Many small improvements since DC24.

WLCG

200Gb/s
LHCOPN

400Gb/s
JANET

External Gateway

~500Gb/s

Batch Farm
Internal Gateways

~4Tb/s

Ceph Backend

Alastair Dewhurst, 15th November 2024

Echo test

- We want to do sustained read and write tests.

- Slow deletes caused a problem in the previous test.

- While we want to test deletes we have not made any major improvements to them.
  - We should ensure that we are not relying on deletes for writing test.

- Read test can be done when doing write test to Tier-2s?

Need to monitor WN usage as it can limit external throughput.
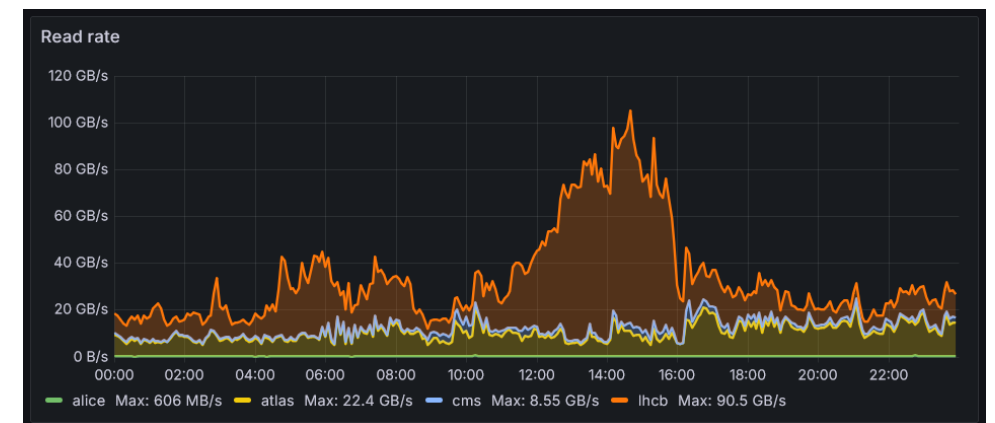
100Gb/s Gateway testing

- We have a single new gateway server with 100Gb/s NIC.
- We want to identify the bottlenecks in this particular node.
  - CPU, memory, network etc
- Several short tests where we try and saturate the machine and then tweaking performance.
  - reads, writes, deletes
- We have the ability to test Jumbo frames on this server.

FTS Balancing

- Echo is a shared service.
  - Allows individual VOs to use much more than their fairshare when they have a big use case.
- Design choice means it is difficult to limit individual VO throughput.
- FTS is the main tool to balance between VOs (number of channels):
  - Goal is to get writes to be balanced:
    - 50% ATLAS
    - 30% LHCb
    - 15% CMS
    - 5% ALICE

Alastair Dewhurst, 15th November 2024

Network

- During DC24 the network was disrupted.
- We are in the process of moving Echo re-directors to new network.
  - Final service being moved off old network means we can fully provision LHCOPN.
- We could consider doing a LHCONE failover test.
  - Disconnect LHCOPN links during a challenge and see it failover to LHCONE.
  - MUST be done with agreement from JISC and CERN!