



Science & Technology  
Facilities Council

UK Research  
and Innovation

# Echo - Experiences running an erasure coded object store

Rob Appleyard, Alison Packer, Tom Byrne,  
Alastair Dewhurst

# CHEP 2018

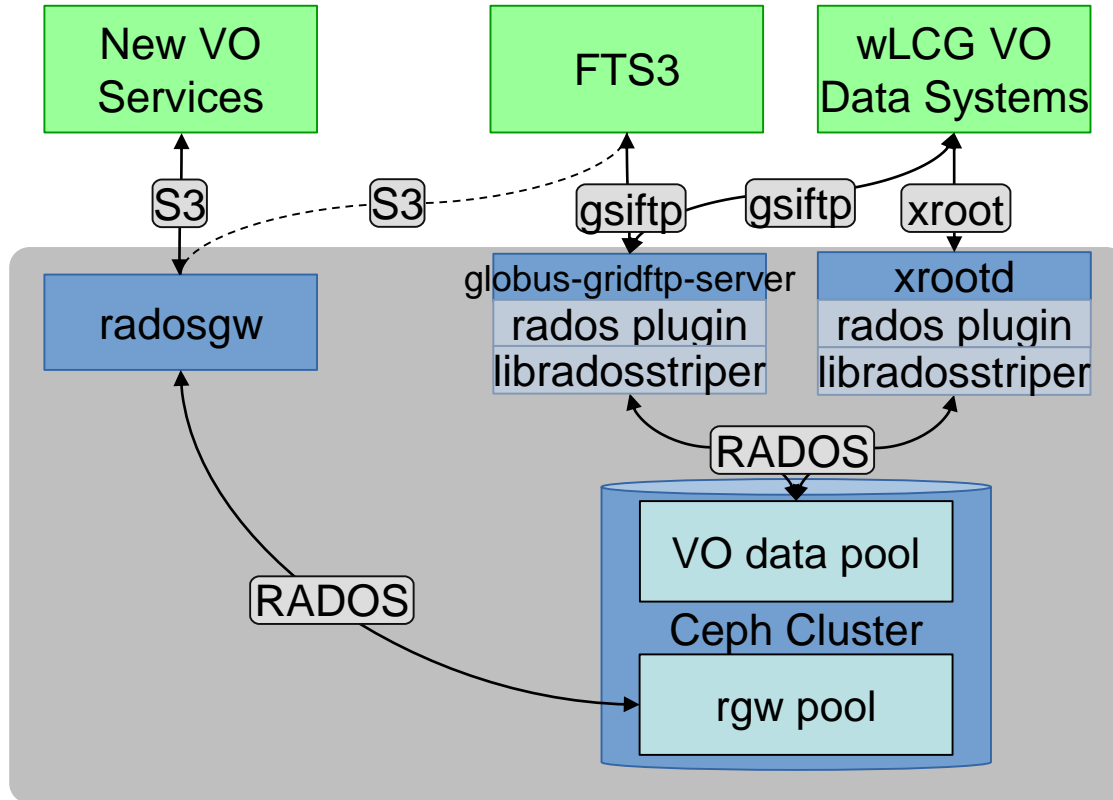
# What is Echo?

- Erasure-Coded High-throughput Object store
- Ceph backed object store
  - Not a file system
  - Access provided via GridFTP and Xrootd
- Erasure coding, not replication
- Current: ~20PB raw, ~14.5PB usable.
  - 27PB usable by end of year
- In production since Feb 2017
  - ATLAS and CMS fully migrated

# Key Design Considerations

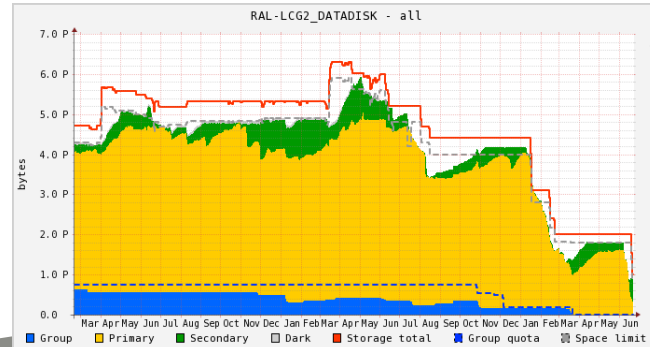
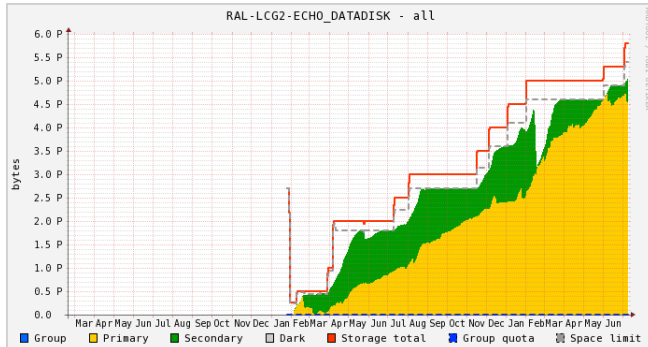
- Price
- Scalability
- Whole-cluster throughput
- Support for HEP protocols

# Access Protocols



# Experiment Migration: ATLAS

- ATLAS
  - 2 'sites' within RAL. Increase quota on Echo, decrease on CASTOR.
  - Separate sites let us go faster or slower as needed



# Experiment Migration: CMS

- 2 PhEDEx 'sites' at RAL
- Jobs got submitted to just one endpoint
  - Site can use Trivial File Catalogue to direct data requests to correct storage endpoint.
- Plan:
  - Use XRootD redirection + multiple TFCs to allow transparent data access to Echo and Castor.
  - External and Internal access would also be directed to different gateways.
  - Successfully confused CMS shifters.
    - Migration happened really fast

# Experiment Migration: LHCb and ALICE

- LHCb
  - For each DIRAC endpoint we will sync data between Castor and Echo and then switch the SPath.
  - Have migrated 500TB+ to Echo.
    - ~4PB still on Castor.
- ALICE
  - Still need to finish testing XRootD configuration
    - Many fixes in XRootD 4.8
  - Only ~400TB space in total.



# Erasure Coding

- Echo uses an 8/3 EC scheme
  - 8 data shards, 3 parity shards, each shard on a different storage node
  - 73% raw storage capacity is usable.
  - Largest known (to us) EC Ceph cluster
- Why use EC?
  - Price
  - Throughput
  - Fault tolerance
- Disadvantages?
  - Sparse reads of the data are inefficient (see Alastair's talk later\*)
  - Complexity (sync between 11 machines)
  - Smaller deployed base

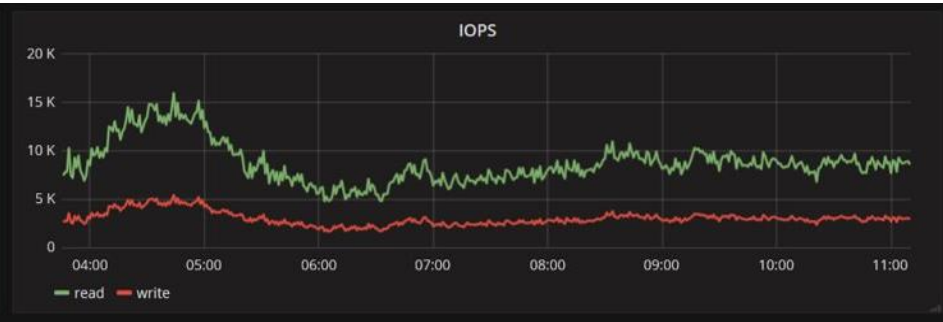
\* <https://indico.cern.ch/event/587955/contributions/2936908/>

# Why 8+3?

- The erasure coding scheme for our main data store is 8+3:
  - $m=8, k=3$
  - 8 data stripes
  - 3 parity stripes
- Both  $m$  and  $k$  are tuneable
  - Ratio and sum are both relevant
  - High total number of stripes – smaller parity stripes
    - ...but also increased computational complexity of synchronisation operations
  - Low  $m/k$  ratio: Greater data security
    - ...but also lower ratio of raw:usable storage
- 8+3 is the final compromise
  - Synchronisation is still computationally nontrivial

# EC and throughput

- 100+Gb/s cluster throughput...
- Mix of internal and external traffic



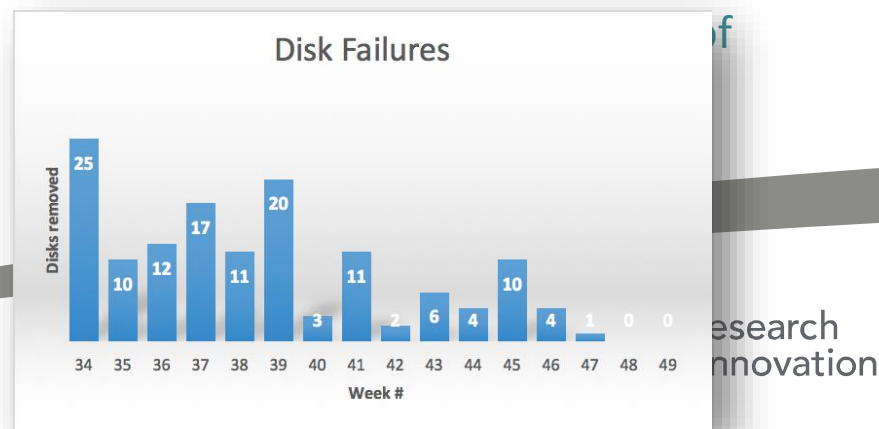
# EC and data resilience

- No data losses *directly* due to hardware failure.
- ‘EC Backfill’: Major bug encountered during late 2017
  - In some cases, when a disk failed Ceph would identify the wrong disk.
  - When this first happened we removed 5 disks before understanding the problem.
    - We can tolerate any 3 losses...
    - By the time we put the healthy disks back enough changes had happened that not all the data was recoverable.
    - 23k ATLAS files lost.

# Operational Problems:

## Disk read errors

- Echo has proved to be very sensitive to disk read errors
  - Read errors are passed to Ceph by OS
  - Highest error state ('HEALTH\_ERROR') if found during scrubbing
    - Easy fix
    - Appropriate behaviour for replicated pools, less so for EC
- Also 'bad' batch of disks
  - High temperature at back rack caused failures.
  - 10 disk servers had more than 5 disk failures in 2 months.



# Aside: Why treat a read error so seriously?

- Ceph isn't designed for things to change on disk
  - It's good at dealing with 'hard' failures
    - Lose entire disk
    - Lose entire server
    - Lose entire rack
- For replicated pools (most common configuration)
  - Data changing on disk is dangerous
    - All is fine if we have 2 matching copies left (they are probably correct)
    - If we don't, placement group\* is inconsistent – may need human intervention - alarm
- Not such a problem in EC land
  - We can easily figure out which version is correct because of the 3 parity shards
  - Ideally, Ceph should understand this – devs are aware.

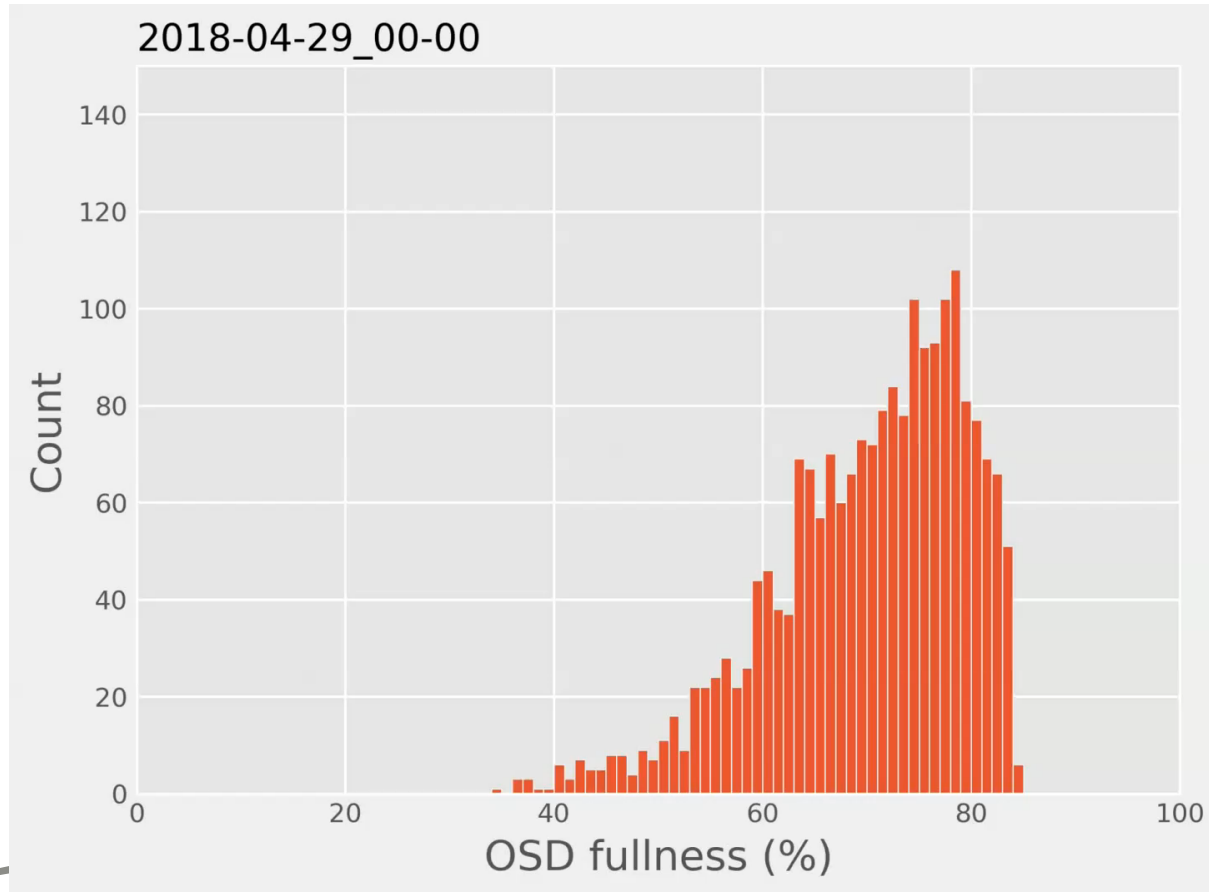
\*A 'placement group' is a set of 11 disks within Echo

# Operational Problems: Data distribution

- Large variations in the amount of data\* that Ceph's placement algorithm chooses to put in each disk (OSD)
  - No protection against overfilling
- One disk fills, cluster enters error state, problems start
- Also disrupts new hardware deployment
- Regular rebalancing operations

\*Actually, number of placement groups

# Data Distribution





# Conclusion

- Echo is in production at scale for ATLAS and CMS
  - LHCb and ALICE still to come.
  - Problems overcome in the first year.
- Confident about the future
  - New hardware coming on stream
  - Ceph devs have taken note of disk management problems
  - Working with new users with S3/Swift.

# Previous talks

- HEPiX Spring 2015 - Ceph storage at RAL
  - <https://indico.cern.ch/event/346931/contributions/817835/>
- HEPiX Fall 2015 – Ceph object storage at RAL
  - <https://indico.cern.ch/event/384358/contributions/909231/>
- HEPiX Spring 2016 – Why so Sirius? Ceph backed storage at the RAL Tier-1
  - <https://indico.cern.ch/event/466991/contributions/2136880/>
- CHEP 2016 – The deployment of a large scale object store at the RAL Tier-1
  - <https://indico.cern.ch/event/505613/contributions/2230932/>
  - <http://iopscience.iop.org/article/10.1088/1742-6596/898/6/062051/pdf>
- ISGC 2018 - Experiences of hard disk management in a erasure coded 10 petabyte-scale Ceph cluster
  - <http://indico4.twgrid.org/indico/event/4/session/15/contribution/66>
- Cephalocon APAC 2018 – Erasure coding big data stability
  - <https://youtu.be/cX8LIFZDfqU>, <https://ceph.com/cephalocon/>
- Regular presentations at HEP Ceph meetings:
  - <https://indico.cern.ch/category/4866/>

# Any Questions?



Image from Hubble Space Telescope, owned by NASA, marked for non-commercial reuse.  
[https://www.nasa.gov/mission\\_pages/hubble/news/arp\\_194.html](https://www.nasa.gov/mission_pages/hubble/news/arp_194.html)