Contribution ID: **588**   Type: **presentation**

Echo - Experiences running an erasure coded object store

Tuesday 10 July 2018 14:00 (15 minutes)

Since February 2017, the RAL Tier-1 has been storing production data from the LHC experiments on its new Ceph backed object store called Echo. Echo has been designed to meet the data demands of LHC Run 3 and should scale to meet the challenges of HL-LHC. Echo is already providing better overall throughput than the service it will replace (CASTOR) even with significantly less hardware deployed.

Echo relies on erasure coding rather than hardware RAID to provide data resilience. Of the publicly known Ceph clusters around the world, Echo is largest running erasure coding in production. This paper describes the erasure coding setup, its advantages over hardware RAID and our experience relying on it for data resilience.

At the start of 2017, the LHC experiments had more than 14PB of data stored on disk in Castor. Migrating this to Echo is no small challenge and will take approximately 2 years. This paper describes the different approaches taken by the experiments as well as their current usage of Echo. This paper also describes the growing usage of the S3 and Swift APIs and lessons learnt.

In the first year of operation there have been many pieces of scheduled work, including the addition of new hardware which resulted in significant data rebalancing, major software updates as well as security patching. There have also been operational problems such as a power cut and high disk failure rates. This paper describes how Echo has coped with these events and the higher level of data availability it is able to provide.

Authors: DEWHURST, Alastair (STFC-Rutherford Appleton Laboratory (GB)); VASILAKAKOS, George (STFC); BYRNE, Tom (STFC); PACKER, Alison (STFC); APPLEYARD, Rob (STFC); CANNING, Bruno (RAL)

Presenter: APPLEYARD, Rob (STFC)

Session Classification: T4 - Data handling

Track Classification: Track 4 - Data Handling