Contribution ID: **327**   Type: **Poster**

Evolution of the Ceph Based Storage Systems at the RACF

Tuesday 11 October 2016 16:30 (15 minutes)

Ceph based storage solutions and especially object storage systems based on it are now well recognized and widely used across the HEP/NP community. Both object storage and block storage layers of Ceph are now supporting production ready services for HEP/NP experiments at many research organizations across the globe, including CERN and Brookhaven National Laboratory (BNL), and even the Ceph file system (CephFS) storage layer is now used for that purpose at the RHIC and ATLAS Computing Facility (RACF) at BNL for more than a year. This contribution gives a detailed status report and the foreseen evolution path for the 1 PB scale (by usable capacity, taking into account the internal data redundancy overhead) Ceph based storage system provided with Amazon S3 complaint RADOS gateways, OpenStack Swift to Ceph RADOS API interfaces, and dCache/xRootD over CephFS gateways that is operated in RACF since 2013. The system is currently consisting of two Ceph clusters deployed on top of a heterogeneous set of RAJD arrays altogether containing more than 3.8k 7.2krpm HDDs (one cluster with iSCSI / 10 GbE storage interconnect and another one - with 4 Gb/s Fibre Channel storage interconnect) each provided with an independent IPoIB / 4X FDR Infiniband based fabrics for handling the internal storage traffic. The plans are being made to further increase the scale of this installation up to 5.0k 7.2krpm HDDs and 2 PB of usable capacity before the end of 2016. We also report the performance and stability characteristics observed with our Ceph based storage systems over the last 3 years, and lessons learnt from this experience. The prospects of tighter integration of the Ceph based storage systems with the BNL ATLAS dCache storage infrastructure and the work being done to achieve it are discussed as well.

Tertiary Keyword (Optional)

Distributed data handling

Secondary Keyword (Optional)

Storage systems

Primary Keyword (Mandatory)

Object stores

Authors: ZAYTSEV, Alexandr (Brookhaven National Laboratory (US)); ITO, Hironori (Brookhaven National Laboratory (US))

Co-authors: RAO, Tejas (Brookhaven National Laboratory); WONG, Tony (Brookhaven National Laboratory); ZHAO, Xin (Brookhaven National Laboratory (US))

Presenter: ZAYTSEV, Alexandr (Brookhaven National Laboratory (US))

Session Classification: Posters A / Break

Track Classification: Track 4: Data Handling