



Contribution ID: 223

Type: Oral

Performance and Advanced Data Placement Techniques with Ceph's Distributed Storage System

Thursday 21 January 2016 15:45 (25 minutes)

The STAR online computing environment is a demanding concentrated multi-purpose compute system with the objective to obtain maximum throughput and process concurrency. Motivation for extending the STAR compute farm from a simple job processing tool for data taking, into a multipurpose resource equipped with a large storage system would lead any dedicated resources to become an extremely efficient and an attractive multi-purpose facility. To achieve this goal, our compute farm is using the Ceph distributed storage system which has proven to be an agile solution due to its successful POSIX interface and excelling its object storage in I/O concurrency. With this we have taken our cluster one step further by squeezing more performance with investigating and leveraging new technologies and key features of Ceph.

With an acquisition of a 10Gb backbone network we have ensured to eliminate the network as a limitation. With further acquisition of large fast drives (1TB SSDs) we will also show how one can customize the placement of data and make good use of the I/O performance tweaking options Ceph has to offer. Finally, we will be discussing OSD Pool mapping in the context of redundancy based on compute racks, rows, PDU's and other physical parameters. We will also present and discuss the cost comparatives of our cluster with other traditional storage systems such as NAS and SAN and the performance of using older hardware to work as one cooperative storage system. We will present our latest performance results as well as the stability, lessons learned, and overall experience with the STAR Ceph cluster and the steps taken to mitigate the problems we've come across. Furthermore we will present the tools we used to manage, maintain, and monitor the Ceph cluster with the use of tools such as the CFEngine configuration management tool and the Icinga Infrastructure monitoring system giving the STAR admins a bird's eye view of the cluster state and a centrally managed point to ensure configuration consistency. We hope our presentation will serve the community's interest for the Ceph distributed storage solution.

Authors: LAURET, Jerome (Brookhaven National Laboratory); POAT, Michael (Brookhaven National Laboratory)

Presenter: POAT, Michael (Brookhaven National Laboratory)

Session Classification: Track 1

Track Classification: Computing Technology for Physics Research