

Building a Storage Cluster with Gluster

Tuesday, March 24, 2009 8:00 AM (20 minutes)

Gluster, a free cluster file-system scalable to several peta-bytes, is under evaluation at the RHIC/USATLAS Computing Facility. Several production SunFire x4500 (Thumper) NFS servers were dual-purposed as storage bricks and aggregated into a single parallel file-system using TCP/IP as an interconnect. Armed with a paucity of new hardware, the objective was to simultaneously allow traditional NFS client access to discreet systems as well as access to the GlusterFS global namespace without impacting production.

Gluster is elegantly designed and carries an advanced feature set including, but not limited to, automated replication across servers, server striping, fast db backend, and I/O scheduling. GlusterFS exists as a layer above existing file-systems, does not have a single-point-of-failure, supports RDMA, distributes metadata, and is entirely implemented in user space via FUSE.

We will provide a background of Gluster along with its architectural underpinnings, followed by a description of our test-bed, environmentals, and performance characteristics.

Presentation type (oral | poster)

Oral

Primary author: PETKUS, Robert (Brookhaven National Laboratory)

Co-author: SMITH, Jason (Brookhaven National Laboratory)

Presenter: PETKUS, Robert (Brookhaven National Laboratory)

Session Classification: Poster session

Track Classification: Hardware and Computing Fabrics