Contribution ID: **533**   Type: **Oral**

Evaluation of ZFS as an efficient WLCG storage backend

Monday, 10 October 2016 12:00 (15 minutes)

ZFS is a combination of file system, logical volume manager, and software raid system developed by SUN Microsystems for the Solaris OS. ZFS simplifies the administration of disk storage and on Solaris it has been well regarded for its high performance, reliability, and stability for many years. It is used successfully for enterprise storage administration around the globe, but so far on such systems ZFS was mainly used to provide storage, like for users home directories, through NFS and similar network related protocols. Since ZFS became available in a stable version on Linux recently, here we will present the usage and benefits of ZFS as backend for WLCG storage servers based on Linux and its advantages over current WLCG storage practices using hardware raid systems.

We tested ZFS in comparison to hardware raid configurations on WLCG DPM storage servers used to provide data storage to the LHC experiments. Tests investigated the performance as well as reliability and behavior in different failure scenarios, such as simulating failures of single disks and whole storage devices. The test results comparing ZFS to other file systems based on a hardware raid vdev will be presented, as well as recommendations for a ZFS based storage setup for a WLCG data storage server based on our test results. Among others, we tested the performance under different vdev and redundancy configurations, behaviour in failure situations, and redundancy rebuild behaviour. We will also report on the importance of ZFS'own unique features and their benefits for WLCG storage. For example, initial tests using ZFS'built in compression on sample data containing ROOT files indicated a reduction in space of 4% without any negative impact on the performance. We will report on space reduction and how the compression performance scales to 1PB of LHC experiment data. Scaled to the whole LHC experiments'data amount, that could provide a significant amount of additional storage at no extra costs to the sites. Since more sites provide data storage also to other non-LHC experiments, be able to use compression could be of even greater benefit to the overall disk capacity provided by a site.
After very promising first results on using ZFS on Linux at one of the NGI UK distributed Tier-2 ScotGrid sites together with the much easier administration and better reliability compared to hardware raid systems, we switched the whole storage on this site to ZFS and will report about the longer term experience of using it, too.

All ZFS tests are based on a Linux system (SL6) with the latest stable ZFS-on-Linux version instead of using a traditional Solaris based system. To make the test results transferable to other WLCG sites, typical storage servers were used as client machines managing 36 disks of different capacity, used before in hardware raid configurations based on typical hardware raid controllers.

Tertiary Keyword (Optional)

Secondary Keyword (Optional)

Primary Keyword (Mandatory)

Storage systems

Primary author: EBERT, Marcus (University of Edinburgh (GB))

Co-author: WASHBROOK, Andrew John (University of Edinburgh (GB))

Presenter: EBERT, Marcus (University of Edinburgh (GB))

Session Classification: Track 4: Data Handling

Track Classification: Track 4: Data Handling