

AFS/OSD: massive production experience and work in progress

Wednesday 27 May 2009 08:30 (30 minutes)

AFS/OSD was presented two years ago on the HEPiX Spring 2007 in Hamburg. At this time the results of the R&D project sponsored by CERN and ENEA were presented which showed that performance of AFS/OSD scales linearly with the number of OSDs used. For RZG, however, the goal was to replace MR-AFS which for the last 12 years had offered HSM-features to AFS. Therefore the concept of “archival OSDs” and “wipeable OSDs” has been introduced. Files in OSDs automatically get copies on archival OSDs with the double purpose to protect the files against loss of a disk system and to allow wiping of files in order to free disk space. In 2007 these features have been added to AFS/OSD and at the begin of 2008 MR-AFS has been replaced by AFS/OSD in place (without moving data on tapes). In Garching TSM-HSM is used as an underlying HSM system, but any filesystem based HSM system can be used. In cooperation with DESY also an interface to dCache/Chimera has been developed. The AFS cell at RZG today contains nearly 300 TB 80% of which are stored in OSDs. Felix Frank from DESY added the policy support which allows to specify for a volume or for a directory which files should go into OSDs, should be striped over multiple OSDs, or get copies in multiple OSDs. Policies can be based on file name patterns or size.

Additional work is in progress to use cluster filesystems such as Lustre or GPFS for AFS. This technique offers fastest access to data stored in AFS inside the cluster while world wide access with normal AFS speed is still possible. It also adds HSM features Lustre.

Author: REUTER, Hartmut (Rechenzentrum Garching)

Presenter: REUTER, Hartmut (Rechenzentrum Garching)

Session Classification: Storage and Filesystems I

Track Classification: Storage