KIT Service Incident Report

Description

During a routine replacement of broken hard disks a bug in the storage controllers was triggered causing both of them to reboot subsequently. Because of this, 3 GPFS clusters and the Oracle instance running the FTS DB lost contact to the disks.

Impact

Many dCache pools – ATLAS: 36 x disk-only, 2 x write-tape – died because their associated mounted file systems were gone. The FTS and LFC services hosted by GridKa were down since the Oracle database backend is located in the same storage system. The ATLAS conditions DB was down, too.

Time line of the incident

27.11.2012 16:50	Crash of the disk clusters.
27.11.2012 17:38	Disks and dCache pools back online.
27.11.2012 18:15	Restart of FTS, LFC services and conditions DB.
27.11.2012 18:45	Oracle DB found to be (partially) corrupted. The FTS service could not be used for regular production.
28.11.2012 08:30	Begin recovery of corrupted DB blocks.
28.11.2012 12:25	DB recovery finished. Restart of FTS successful.

Analysis

The firmware for the storage controllers was identified to be faulty. Once a fixed firmware is released, we will install it on our systems asap. This should avoid a similar issue in the future. Although the redundancy between the two storage controllers failed, the actual data integrity was left unharmed. Therefore, we were able to restart the dCache pools as soon as the file systems were usable again.

The Oracle DBs got corrupted due to the sudden loss of connection to the disks.

We are awaiting a fix from the storage vendor. It is unclear when the fix, which may also include HW changes, is available.

Follow up actions

-   Install updated storage firmware once available.

Summary

The problem will either be solved by hardware replacement and/or by firmware update as soon as a downtime for this storage system permits. Until then maintenance on the storage system, e.g. disk replacement, will take place only when all depending services have been taken offline before. The vendor has not been conclusive on the prospective date of possible replacement of the storage hardware but has guaranteed a timely solution.