From: Jeff Templon
To: Jamie Shiers; Harry Renshall;
cc: Maurice Bouwhuis; sanden@sara.nl;
Jurriaan Saathof;
Subject: post mortem tape-system outage 25/10 in NL
Date: Friday, November 07, 2008 11:13:49

Hi Guys,

On behalf of the SARA MSS group.

JT

Post mortem unscheduled downtime tape mass storage system on 25/10 until 27/10.

The tape mass storage system is built around a CXFS clustered HSM filesystem. On Saturday 25/10 on 10:28 a non-recoverable error occurred on the CXFS filesystem. On one of the storage enclosures, which builds the CXFS filesystem, both controllers hit a non-recoverable error on which the storage enclosure lost both its paths to the SAN. This resulted in the CXFS filesystem losing a few luns, corrupting the filesystem. On 25/10 10:32 the clustered nodes lost access to the CXFS filesystem. When this happens access to the tape mass storage system is automatically disabled to the dCache front end storage.

This problem was discovered on monday morning 27/10 around 8:00. On 27/10 8:16 the srm.grid.sara.nl was placed in unscheduled maintenance in the GOCDB. After discovering the problem and investigating the cause and the scale of the problem, the choice was made to destroy the CXFS filesystem, rebuild the filesystem from scratch, and restore the latest successful backup. The backup restored was from 25/10 0:12. The restore of the filesystem was completed at around 27/10 10:53. At 27/10 10:58 the first files were flushed from dcache to the tape mass storage system again. After restoring production the system was monitored for a few hours before the unscheduled downtime was deleted in the GOCDB, at 15:31.

The time between problem and the restored backup is little more then 10 hours. The files stored in this time period are considered lost. In total there were 53 files written, 28 Atlas and 25 dteam files. The Atlas file names are communicated with Atlas via Hurng-Chun Lee, our NL-T1 representative. The dteam files are not communicated.

This concludes the post mortem.