

IN2P3 Service Incident Report

Description

IN2P3 Tier1 site provides two types of disk storage for WLCG. The first one is Xrootd dedicated to Alice experiment and the second one is Dcache for ATLAS, CMS and LHCb experiments. Same hardware is used for the both storages systems but each of them is based on dedicated disk servers.

Each disk server is constituted of two raid systems with each of them constituted of twelve disks. A RAID6 configuration is deployed on every RAID channel (can resist at two disks failures) and the total space is provided to the system via a logical group, which merges the space of the two RAID.

Within two days (16/17 July), two disks have failed on same RAID channel. To decrease the load on disk the server was put offline just after the failure of the second disk. The reconstruction process was launched as soon as new disks have been available (24 hours). For unknown reason, the rebuilt phase failed on one disk and during the investigation to understanding the problem, two others disks have failed on same RAID channel (20 July). In this condition, the RAID array has been broken and data lost. Some efforts were done with vendor to try to retrieve some data, without success. Data were declared as lost.

Impact

110 To of data were lost on Xrootd system. Alice was able to retrieve this data from other sites. During some weeks, the capacity of storage deployed for ALICE has been decreased, because the disk server was not put online quickly to have time to analyse the incident.

Time line

July 16/17th 2018	Two disk failures on same raid6 system. Disk server has been put offline to decrease the load during the rebuild process
July 20th 2018	Rebuilding phase failed.
July 20th 2018	Two disks failed on same RAID Array.
Next week	Investigation to retrieve files without success

Analysis

The hardware where the problem has been occurred is 3 years old. No other similar problems were faced on the same generation of hardware.

Expert from labs of vendor mentioned that they already had faced a similar problem due probably to a disk who "spam" the RAID bus and disables some disk of the RAID array.

During time, we were thinking that it is possible for us to retrieve the data, but it was not the case. A better communication and a quicker escalation of the problem to hardware expert maybe could have allowed us the retrieve of the data.

At this time, ticket concerning this incident is still open, because the timeline is not yet fully understanding and as the incident is related to a hardware device that we have massively deployed on the site, we have to clarify if it is a isolate incident or not.