

PIC Service Incident Report

Description

In October 2015 we started to use T10KD technology for ATLAS at PIC. After two months writing on tape with this technology, on December 17th while checking a migration process, we detected read errors in one T10KD cartridge.

We never ever (or rarely) observed such problems, and since the technology was pretty new for us, we decided to check all the tapes written with this technology (79 cartridges) in order to check if other T10KD cartridges could be affected. On this same day, we reverted back ATLAS to use/write to T10KC, as we were not sure about the impact of the incident. After checking all T10KD cartridges, we verified that 8 cartridges were affected.

We looked for the drives which wrote the affected cartridges, and we realized that all of them were written by a single drive (DMOD 4721) which was replaced on December 10th due to a write error in one cartridge.

In February 2016 we migrated these 8 cartridges and all files could be migrated except 114 files. A new ORACLE private firmware (4.10.106 – currently in production) with an enhancement called “read accumulation code” was installed in one of our T10KD drives.

Using this new firmware, it was possible to read 93 out of the 114 problematic files. Only 21 files (13 DataTape and 8 MC) located in 3 cartridges were not recovered. Therefore we sent these 3 cartridges to ORACLE Data Recovery for analysis and data recovery.

Finally, by the end of April 2016, all files were completely recovered, and the DataTapes were injected back to PIC file system.

Impact

21 files were not available during this incident. These 21 files were successfully recovered by Oracle Data Recovery service.

Timeline of the incident

04.12.2015	Write error detected in T10000D drive with S/N 579004004721. Set offline this drive in the application and reported to Oracle (SR opened).
10.12.2015	Drive with S/N 579004004721 was replaced by Oracle.
17.12.2015	First read error (FSC 37F6) in a T10000D cartridge (reported to ORACLE). Writes to T10000D were stopped. We started to read all T10000D cartridges written (79 cartridges).
23.12.2015	New ORACLE firmware (4.10.106) was installed in one T10KD drive.
28.01.2016	We finished to read all T10000D cartridges: 8 cartridges affected. We start to migrate data from the affected cartridges to T10000C technology.
09.02.2016	We informed ATLAS about the T10KD issue and the ongoing investigations.
11.02.2016	After several retries the migration of 8 cartridges is finished: 21 unrecovered files located in 3 cartridges.
12.02.2016	21 files were reported to ATLAS as unavailable.
15.02.2016	We reported in detail about the T10KD issue during a WLCG operation meeting [1]

18.02.2016	Sent 3 affected tapes to Oracle data recovery for analysis.
14.03.2016	Received 1 cartridge from Oracle data recovery with 1 file recovered.
04.04.2016	Update from Oracle (detailed in "Analysis" section).
27.04.2016	Received last 2 cartridges from Oracle data recovery with 20 files recovered.
28.04.2016	Reported to ATLAS experiment: all files have been recovered and injected back to PIC file system [2], waiting for ATLAS to re-inject them in their File Catalog.
03.05.2016	ATLAS started to write again on T10000D technology.
18.05.2016	ATLAS re-registered the recovered files into their File Catalog on RUCIO.

[1] <https://twiki.cern.ch/twiki/bin/view/LCG/WLCGDailyMeetingsWeek160215>

[2] <https://its.cern.ch/jira/browse/ATLDDMOPS-5253>

Analysis

Oracle experts reported that the problem was related to one single drive (DMOD 4721) which was writing data off track:

*The problem that this drive has caused can be called "writing off track". It's not that the data hasn't been written or that the data written was different from what's supposed to go on tape. The data has been written "incompatible" (off track) in a sense that other drives couldn't read the data back (depending on the build-in tolerances). **The particular error with this drive made the read-back check to succeed and hence the problem went unnoticed**, the drive itself could read without a problem. Unfortunately other drives couldn't always adjust their read heads to the "off track" data and therefore failed with a FSC 37F6.*

Follow-up actions

We asked ORACLE support if there is a periodic check of the tape drives to pass or if there is a way to detect drives "writing off track" in order to prevent these read errors to happen. ORACLE support stated that they have not seen this issue anywhere else and it is to be considered as an exceptional case.

As a follow-up action we will increase the check frequency of the random access to tape files in order to detect bad tapes.

Summary

The detection and recovery procedure of the T10KD issue affecting ATLAS took around 4 months. It was operationally costly, because it was necessary to take several dumps when read errors were detected to investigate the problem. It required to read all tapes written with this technology in order to understand the impact of the incident, and to migrate all data contained in the 8 affected tapes after performing several retries. Eventually 21 files resulted unrecoverable.

After informing ATLAS about it, they recommended us to recover at least the DataTape files because they were unique copies. Therefore we sent the related cartridges to the US for analysis and data recovery. Based on the analysis carried out by the ORACLE Engineering team, we've received confirmation that we can put back all our T10KD drives into production, and that the problem was related to one single drive writing data off track.

Once the file recovery was completed, we re-injected just the DataTape files to PIC file system, as requested by ATLAS. After being informed, ATLAS DDM team re-registered those files in their File Catalog on RUCIO.