On 21/09/2012 one of power supply of one of the enclosures of storage system assigned to LHCb (a EMC CX4-960-6) was in fault and EMC the maintenance support decided to replace it.

The replacement, a routinely intervention, started the same day (21/09/2012) at 15:00 MET.

The EMC technician did not reveal immediately another problem present on the same enclosure, more precisely on one of the cables interconnecting the JBODs.

From EMC technical support: "As shown by logs and confirmed by CE observations onsite it looks like the loop/bus 3 cable on B side of the system going from enclosure 5 to enclosure 6 had a dormant issue which was triggered during the replacement of a power supply in the enclosure 5.

This intermittent cable issue triggered the system to see the drives on the loop as malfunctioning and marked them as failed".

Hence the cable was replaced too but this operation caused a loop and consequently several connections between several disks and controllers timed out. Immediately these several disks were put off-line from the controller. As a consequence the gpfs file system stopped working.

We immediately dismounted the file system from the WNs, UIs and we stopped any operation on it.

A down was put in the GOCDB.

Then we asked for an immediate escalation.

At 19:45 MET, EMC remote support tried to reset all faulty disks one by one and then the SAS communication modules for the specific enclosure.

At 24:00 MET all faulty LUNs are now in "transition" state (i.e. the disks are in rebuild).

After EMC's advice, the LHCb file systen is kept un-mounted in order to allow a quicker and safer rebuild of all the affected LUNs.

26/09/2012: The rebuild processes end and a check on the file system is started.

15:00 MET: the file system checks ends with the message "File system is clean"

At this point, after EMC's technical support advice, an upgrade of the CX4-960-6 firmware is performed.

19:45 MET: the upgrade is completed.

20:00 MET: the file system is re-mounted on the disk-servers and tests are performed on it.

During the night, the firmware reveals some stripes of data in "non recoverable" state.

27/09/2012 at 09:00 MET: LHCb queue is re-opened on the farm.

28/09/2012 at 13:00 MET: EMC technical support starts a background verification of the raid-sets 55 and 82.

03/10/2012: the verification process ends showing that 3 stripes on LUN 55 and 2 stripes on LUN 82 have double error (this system is still with RAID5).

09/10/2012: EMC analysis recommends a consistency check of all files present on the file system to find out if there are any damaged files. To speed-up the process we asked EMC a way to find which files reside on the LUNs 55 and 82 but apparently this is not obvious.

In any case, in a month, the file system will be migrated to a newer hardware: in that occasion a complete check of all files will be performed.