

Follow-up Report KIT Service Incident 09.08.2018

**Lessons learned,
Improvements applied and
Final impact in numbers**

Steinbuch Center for Computing (SCC), Scientific Data Management (SDM)



Incident specifics

- Site: KIT
- Service Area: (Online) Storage
- Date: August 9th/10th 2018
- Published [SIR](#)
- Incident description:
 - The database backend for the dCache storage element „cmssrm-kit.gridka.de“ crashed on August 9th due to disk space exhaustion.
 - Site administrators moved production to a warm stand-by database clone within the first hour of the incident.
 - However, the next day, August 10th, it was decided to move production back to the original master database, because the stand-by machine was lagging one week behind the master in terms of database updates.
 - Consequently, some work accomplished by CMS between 9th and 10th of August was given up.
 - Though the days after, CMS site contacts and GridKa administrators have recovered some data from the disks.

Lessons Learned and Improvements Made

- Fixed the original cause for the disk space exhaustion.
 - IPv6 configuration deployed on the database nodes broke, because a new site-wide strategy was being tested. → Removed legacy IPv6 configuration from the database nodes.
- Improved robustness and monitoring of the workflow
 - Because the synchronization tasks never failed, the monitoring that already existed did not raise an alarm. → Added timeouts to the synchronization tasks, which were hanging indefinitely otherwise.

Quantified Impact

- Because the dCache pools were not restarted in the aftermath of the incident, the files created between August 9th and 10th were not purged from the disks.
- We copied files from the pools that dCache could no longer identify onto temporary storage and provided access to those replicas to CMS via the VO box.
- By crawling the dCache billing log files, we generated a list mapping those replicas to logical file names as CMS understands them.
- By September 19th, when we concluded our efforts to recover any more data, we managed to restore **57538 /store/unmerged** and **299 other** files „known to Phedex“.
- A definitive number of lost files is unknown to us
 - Ordinary user files (owned by dCMS role in this case) were not cataloged. A broadcast was sent to the users that some files may be lost now.
 - Restored files were not confirmed so far.
 - A comparison between our database content and Phedex still is not finished