

GDP issues

Andrej Filipčič, Rod Walker

Duplicate GUIDs

- Sometimes, two output files created on the same WN can have identical guid
- Recent incident:
 - ➔ <https://savannah.cern.ch/bugs/index.php?102508>
- UUID v4 used almost everywhere – completely random
- UUID v1 still used in some parts of the code, to be fixed by trf group
 - ➔ It has 100ms granularity, within that period the same uuid is returned on the same WN (ip)
- Not a big issue: the 2nd job is marked failed due to DQ2 registration issue, but the LFC catalogs need a manual cleanup

Double files

- In some cases, more than one version of a file is registered in the datasets, eg.:
- `dq2-list-files mc12_8TeV.159010.ParticleGenerator_e_Et7to80.simul.HITS.e1173_s1688_tid01239815_00 | grep 004077`
 - `HITS.01239815._004077.pool.root.1 F858BD8F-5C4B-4246-8DFD-95A0DA4C8603 ad:3a6cddb0 2821671`
 - `HITS.01239815._004077.pool.root.2 86A3B40A-8127-AE4F-89B2-BF82BEC949C7 ad:b964eb00 2821689`
- Several issues (Tadashi)
 - DQ2 problem (fixed in April 2013): successful dq2 registration was reported as failed to panda, so panda marked the job failed and forced a 2nd attempt. The 1st output stayed in registered. Some active datasets still affected by this bug and need manual cleaning.
 - Panda problem: not clear what is the issue (log-rotating was active at the time of registration). Possibly related to LFC registration at WNs in US.

Mintime

- Panda queue parameter mintime activated in job brokering on request by sites (FR T1)
- Jobs which are shorter (maxCpuCount) than mintime are not sent to the panda queue
- Helps the sites to optimize the resource usage and better backfill of short jobs