CMS Storage Use Cases

Storage Classes WG Meeting 28.07.06

Artem Trunov

CC-IN2P3

artem.trunov@cern.ch

CMS storage

- We don’t yet have a full story to tell
- T0 workflow is being developed
- Reconstruction workflow is being developed
- Dataset placement policies under development
- Dealing with user files, physics workgroup files under development too.
- But you should not expect anything extraordinary from these items

Short list of use cases for T1

- Transfer raw from CERN for archiving.
- Staging raw for reprocessing, archiving reconstructed data (custodial)
- Transfer reconstructed data from another T1 (second copy)
- Keeping entire AOD on disk for analysis
- Transfer of MC data from T2, archiving, making available for analysis.

- Require disk-only pool for intermediate results of production (files are then merged to get a reasonable file size and stored in tape pool)
- 85% of disk space – for analysis, the rest – for reprocessing and MC. (from Computing TDR)
- Data Tiers: RAW ~ 1.5MB/ev, RECO ~ .25 MB/ev, AOD ~ .05MB/ev

CMS file transfer

- Done with Phedex, CMS own tool, running on CMS “VO BOX”
- Transfer is driven by the target site.
- Source site generates TURLs and passes them to the target site’s Phedex.
- SRM is mandatory
- Moving to FTS-only transfers.
- Some sites prefer to operate Phedex in such mode that it uses “direct” access to site’s storage, i.e. doing **ls –l /pnfs/etc//,** staging a file or checking whether it’s been migrated to tape
- New srm standard will help to eliminate this, but CMS is a bit conservative, and will deploy new features carefully.
- With SRM it’s not even required to have a Phedex on each site, could manage transfers remotely.
- *CMS is flexible with what and how sites deploy Phedex and storage*

Transfer and use of storage classes

- So far used only equivalent of Tape1Disk0 and Tape0Disk1
  - The former is only for testing, use in production is not foreseen too.
- Use of Tape1Disk1 will be evaluated
  - Need more sophisticated DM tools
  - Due to Phedex flexible configuration, CMS site admins will be free to use it or not.

“Processing” vs. “transfer” pools

- Some sites use it, some not
- Not a have a common policy, site CMS contacts use their wisdom and talk to local admins to setup storage appropriately.
- However we will be moving toward having standard setup and recommendations.
- Also considering “analysis” buffer and “reconstruction” buffer (to avoid mixing raw and the rest), but may end up with only one common buffer.

Staging, pinning, purging…

- Not used in analysis workflow.
- Limited use in transfer (with Castor)
- Plans – will use prestaging for sure in reconstruction workflow.
- For analysis – can be used to enforce dataset placement policy (i.e. making sure that new aod dataset is on disk, and old not). Not yet clear, whether prestaging will be used or not. May get away with dynamic staging.

CMS namespace

- Have a common LFN naming convention
  - /store
  - /store/unmerged – maps to disk-only production buffer
  - /store/raw – to facilitate archiving raw separately
- Such LFN convention allows to use trivial file catalog, i.e. SURL is composed of storage end point and an LFN
  - **$SEproto://$SEname:$SEport/$SEVOpath**/store/etc/
- Single SRM endpoint is not yet enforced, but gearing toward it
  - Some sites even use different types of storage together.