Storage Issues

Armen Vartapetian

BNL Storage Meeting; April 13, 2009

BNL Storage Configuration

- *Space Tokens*
  - **DATADISK**: **700**TB
  - **DATATAPE**: **55**TB
  - **MCDISK**: **490**TB
  - **MCTAPE**: **50**TB
  - **USERDISK**: **15**TB
  - **SCRATCHDISK**: **15**TB
  - **GROUPDISK**: **20**TB
  - **LOCALGROUPDISK**: **15**TB *(to be created)*
- *Non-Space Tokens*
  - **BNLPANDA**: **400**TB
  - **BNLDISK**: **46**TB
  - **BNLTAPE**: **33**TB
  - **USERDISK**: **110**TB

BNL Storage Status

- Usually available/free space for storage involved in production is kept at the level of at least 10%
- Additional space reserves (not on display) were available, and were added wherever needed to sustain production
- Just recently significant increase in reserve space due to installation of 30 thors (dc061-dc091), 30TB each. Some thor pools are already added to production storage. Still setup work on some of them.
- Will allow better flexibility in storage operations, and to have much more free space available any time for production storage
- All space token storage is thumpers/thors, all non-space token storage is acas (except USERDISK)
- At the moment BNLPANDA (T1D1) is still used for production
- Similar space token configuration, as in previous slide, in T2-s (+PRODDISK-DATATAPE-MCTAPE)

Next in Storage

- Move production to space token areas (will be discussed separately)
- Gradual retirement of the acas storage:
  - 270 machines, acas0128 - acas0397, 97TB, retired 2009
  - 192 machines, acas0015 - acas0127, 253TB, to be retired 2010
  - 95 machines, acas0489 - acas0583, 343TB, to be retired 2011
- General concept is clear for DATADISK, DATATAPE, MCDISK, MCTAPE storage
- Not so settled for USERDISK, SCRATCHDISK (new), GROUPDISK, LOCALGROUPDISK – still in development… (discussion in next slides)
- Cleanup of the storage is an important issue (discussion in next slides)
- Additional 1pb storage planned for 2009

Disk Cleanup

- One of the main obstacles for disk cleanup is that the records of the deleted datasets in central catalog are not accessible with regular dq2 tools (only for dq2 administrators). Though we still have replicas for some of those datasets in US, we don't have access to their records!
- Savannah ticket is submitted on this, long discussions with dq2 experts, but no resolution yet. No timescale for option/method to list those datasets.
- Additional pressure from management may help to speed up
- At the moment only possibility is to use so called bulkDeleteDatasetReplicas method, which will blindly delete those datasets (if they exist in the site). No similar method for listing.
- It is decided to go ahead with this deletion for obsolete datasets
- A notification with the cleanup list will be sent from usatlas-diskcleanup-notify listserver address to usatlas-users-l list this week
- Run the cleanup process for each storage element which may have those files, including T2 sites
- dCache cleanup is not automatically propagated to HPSS in cases when tape backup exists: → Use deletion logs to extract pfn-s, and construct a path for HPSS deletion

USERDISK → SCRATCHDISK

- Due to the confusion within users that USERDISK is for more permanent usage rather than a scratchdisk, it was ATLAS-wide decision to discontinue it and instead create a space token called SCRATCHDISK
- BNL already has created the SCRATCHDISK space token
- No problem with closing of USERDISK at BNL, since user analysis are using the non-space token USERDISK area
- All the US T2-s were also been asked to create this new space token
- We will also help T2-s to cleanup USERDISK areas for subsequent closing

USERDISK Cleanup

- As USERDISK is defined as scratchdisk, regular cleanup is needed
- We are continuing to send notifications to users who have old datasets
- As with production disk cleanup, usatlas-diskcleanup-notify listserver address is used for mass email
- We will also help T2-s to cleanup USERDISK area
- For users one of the location options, where to save the datasets they want to keep, is the LOCALGROUPDISK area
- We are in the process of creating LOCALGROUPDISK space token (details in the next slide)

LOCALGROUPDISK

- Need to create a LOCALGROUPDISK space token, for users to have a more permanent location for storage of the private analysis data
- No defined policy for LOCALGROUPDISK by ATLAS
- Several policy issues for USATLAS still to finalize:
  - Size of the storage for T1 and T2-s
  - Personal quota, methods to enforce
  - Limitation to local users, or general USATLAS access
  - Deletion/cleanup issue, who will do. Deletion tools for users.
  - No backup/replica for that data
  - …
- Nevertheless probably we can go ahead an create this space token even without having all the answers…

GROUPDISK

- Recently a dedicated meeting at CERN (April 7) to discuss policy of usage of the GROUPDISK area
- At T1 & T2-s as an analysis resource
- Start extensive utilization, set goals, collect experience, analyze and discuss
- Accounting per group per site. Have dedicated tools (timescale first half of May)
- About 15 groups
- Data from particular group how distributed? Not every T2, but distributed through clouds…
- Tools are still not optimized for this activity