Table of Contents

Archival Site Survey Conclusions

Conclusions from the survey.

Main Message

- Submit recalls as far in advance as possible
    - ♦ Keep the queue as full as possible

Campaign Planning

- Group recall requests by creation time or tape family if possible
- Inform the site with as much warning as possible about recall plans
    - ♦ Allows synchronisation with local activities such as repack
- Understand how priority requests are handled
    - ♦ Submitting priority requests will degrade throughput
    - ♦ Withholding recall submissions to keep latency down will degrade throughput
- Synchronise data use with recalls to avoid purge/recall loops
- The client should delete a staged file from the disk buffer once the workflow requiring the retrieval has completed.
- Do not wait for the last byte to be recalled before beginning processing

Client Behaviour

- Consider queue size to be unlimited
    - ♦ Exceptions
        - ◊ FNAL, PIC - 15k per VO
        - ◊ KIT - 2k per pool
        - ◊ UNIKHEF-SARA - 1k (?)
- Back off on a combination of SRM_INTERNAL_ERROR (request status) and SRM_FILE_BUSY (file status) (Castor).
- Back off when the number of files in SRM_REQUEST_QUEUED approaches the server-configured limit (dCache).
- Use bulk requests
    - ♦ Best bulk recall size unknown. 1k is the reference, some sites want more, some want fewer.
- Interaction rates under 10Hz typically acceptable
- Run with no timeouts, or at least 48hrs
- Ignore disk buffer occupancy
    - ♦ Exception: CNAF

Writing strategy

Recommendations on writing strategy are difficult, both in defining a good strategy and understanding how it should be implemented. In principle, one could select particular pools or resources for particular types of data or based on the probability of future deletes. In some cases, any locality gained will be destroyed by repacking campaigns.

Discussions are underway between Atlas, FTS and dCache to understand how datasets can be tagged thus allowing the storage system to promote locality.

This topic: HEPTape > Survey_Conclusions

Topic revision: r5 - 2018-10-10 - OliverKeeble