



Archival Storage WG DOMA Update

Oliver Keeble

DOMA

- “The WLCG DOMA project consists of several activities in the area of Data Organization, Management and Access, with a focus on the medium/long term evolution”
 - <https://twiki.cern.ch/twiki/bin/view/LCG/DomaActivities>
- A coordinating layer which tracks targeted activities, currently:
 - DOMA WGs
 - Third party copy
 - Data Access
 - Quality of Service (QoS)
 - Other activities
 - Archival storage (WLCG WG)
 - Data and networks
 - Auth* (WLCG TF)
 - ...

DOMA Nov 2018

- November general meeting concentrated on Archival storage
 - <https://indico.cern.ch/event/767209/>
 - Carousels
 - Archival Storage WG report
 - Experiment input

Today's meeting

Focuses on archive storage:

- Processing not-so-frequently-accessed-data (a.k.a. “cold data”) from high latency and low cost media (today a.k.a. TAPE) => “Tape Carousels”
- Archive storage Working Group and the ongoing activities. Technology evolution. Synergy with the work ongoing in QoS DOMA activity.
- Inputs from experiments: plans, prospects, strategies, concerns, ongoing efforts
Name one.

We should then try to understand how those initiatives will proceed coherently

Finally, a quick roundtable of other DOMA activities in the AOB

Future WG activities

- Long term evolution of experiment use of archival storage
 - Vlado's talk today
 - See experiment contributions at DOMA meeting
- Cost modelling
 - Renaud's talk today
- Should we pursue the following?
 - QoS, non-tape archive, pledges
 - Post-SRM interfaces

post-SRM interfaces

Why are we considering this?

- Two providers (CERN & RAL) are developing new systems
 - CERN has no plans to provide an SRM interface
 - CERN's EOS/CTA is using an xrootd interface
 - Tweaked to handle tape operations
 - FTS can already use this
- Is there an argument for a new standard/convention?
 - **What should the WG's role be in answering this question?**
- If we do nothing?
 - CERN and RAL will have their own interfaces

Is an agreement required?

We need a new common interface definition because...

- We use only a small part of SRM
 - A new interface could allow simpler, cheaper systems
- We could add new functions to a new interface
 - Tagging datasets
 - Expressing locality
 - Dataset operations
- A more generic QoS interface could be used
 - Candidates exist
- More systems will add new interfaces and there will be proliferation

We don't need a new common interface definition because...

- SRM is not being obsoleted
 - The number of systems needing a new interface may be very small (Just CERN & RAL in the end)
- FTS can be used and its interface will not change
 - It can absorb differences between systems
- Many smaller communities already use storage-specific commands
 - Castor “stager_get”
 - This works OK
- SRM can be extended if new operations are required