TAPE REST status (SRM retirement)

GDB, September 11th 2024

P.Vokac on behalf of WLCG experiments

TAPE file transfers

- No direct access to files stored on the TAPE
  - Staging files from TAPE to the disk buffer
    - Special protocol / protocol extensions to trigger data staging
      - Client ask storage to "bringonline" files from tape to disk buffer
      - May be transparent (e.g. dCache), but this comes with certain limits
  - File uploads transparent for clients
    - Again client transfer file just to the disk buffer in front of tape system
    - Data transferred asynchronously to the tape
- Different "bringonline" methods for staging files from tape
  - SRM(v2) with support for heterogenous storage systems
    - Complex protocol designed two decades ago
    - Implemented by CASTOR, dCache, StoRM
    - Some SE implementation and FTS/gfal rely on GCT libraries
      - GCT WLCG retirement plan, OSG done
    - Only dCache implementation supports access with tokens
  - xroot protocol extension implemented by CTA
  - TAPE REST

Design

- Modern, simple, minimalistic and uniform way to manage tape transfers
  - Designed in 2022 by WLCG tape system and dmc developers (CTA, dCache, StoRM, FTS)
  - Manage disk residency of tape-stored files and observe progress
  - Support bulk operations for efficient handling large number of files
  - HTTP REST with auth by both X.509 and tokens
    - no dependency on special and complex libraries
- TAPE REST API v1 specification
  - Finalized in May 2022
  - TAPE REST endpoints
    - stage (submit, progress, cancel, delete)
    - release
    - archiveinfo
- BDT TAPE REST twiki documentation

a lot of freedom in the specification to allow different implementation optimize tape access

Implementation

- CTA
  - Preliminary specification implemented already in March 2022 (EOS Workshop 2022)
  - Allowed FTS/gfal2 to start development/testing with CERN test CTA instance
- dCache
  - Build on top of more generic Bulk Service v2 (dCache Workshop 2022)
    - Included in TAPE REST testbed in July 2022
  - Available in golden dCache release 8.2 (September 2022)
  - Important Ops features (stale request cleanup) available only in dCache 9.2
    - Sites with specific configurations may need 9.2.23
- StoRM
  - Available since the end of 2023
  - Introduction of TAPE REST was part of bigger project and replacing GCT
- FTS/gfal
  - Gfal2 2.21 brings support for TAPE REST API (DMC-1301)
  - FTS support since 3.12.2
  - Transfer and stress tests done in July 2022

Deployment

DC24 relevant experts bussy

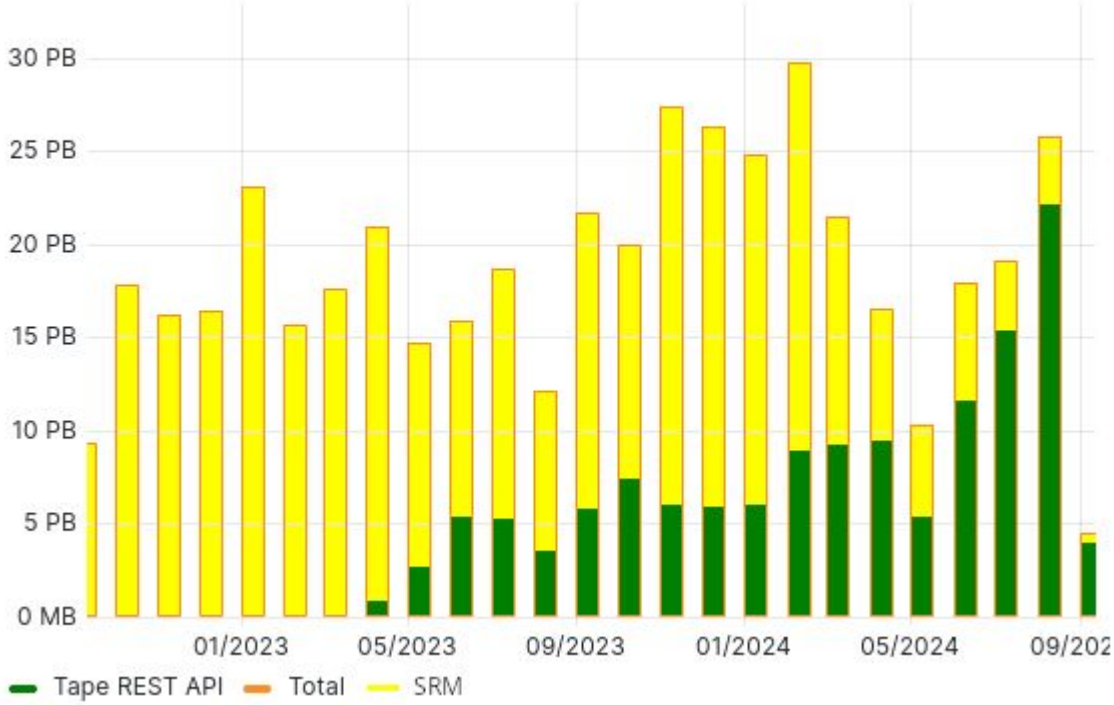
- Crystal ball predictions from 2022
  - Second half of 2022 will be for testbed instances and software verification.
  - Leading-edge T1s will have production instances in 2023 available for tests.
  - ***Site-by-site migration*** during 2023; ask all sites to finish by winter shutdown 2023/2024?
- March/**April 2023** first site with TAPE REST in production – FZK **dCache**
  - few updates to get everything stable and include also “DATATAPE” (GGUS:160869)
- **CTA** @ CERN with TAPE REST available since **March 2023**
  - mid of April 2023 ATLAS started to use TAPE REST for all CERN production RSEs
  - CMS and LHCb followed in next months
- CTA @ RAL in production since May 2023
- First **StoRM** TAPE REST used by LHCb since **beginning of 2024**

Operation

- Issues found by early adopters were quickly resolved
- Several problems with dCache namespace prefix took longer to resolve
  - non-default `webdav.root` was not handled properly by TAPE REST
    - dCache with this configuration did not work at all – fixed in 9.2.14
    - no cleanup of disk buffer – fixed in 9.2.23
  - took time to discover these issues
    - T1s with more complex configuration adopted TAPE REST later
    - tricky to push sites to TAPE REST upgrade when it is not 100% clear their configuration can't cause troubles
- Configuration tuning, limits and potential issues
  - Staging queue limits in FTS
    - Missing limit lead to problem with StoRM when queue reached 50k+ transfers (HTTP 500 response from REST)
  - Storage implementation protects TAPE REST from excessive request size and number of active requests
    - Compatible numbers must be configured in FTS
    - Default dCache values compatible with 1 FTS server (experiments usually rely on more FTS instances)
  - FTS storage and link limits applied to scheme://fqdn
    - Different hostname (e.g. DNS alias) for TAPE doors necessary to allow different DISK vs. TAPE active transfer limit
- TAPE disk buffer management (FTS-2043)
  - Specification leave optimal decision on SE implementation
  - FTS default pin lifetime (7 days) may be too long for small buffer
  - CTA can wipe staged files from buffer after 1 day regardless of suggested pin lifetime, StoRM in 3 days, dCache respect pin lifetime
- GOCDB new service type "`wlcg.webdav.tape`" and OSG "`WebDAV.tape`"
  - Allows to distinguish downtime for TAPE vs. DISK with shared same doors / hostname
- WLCG SRR (Storage Resource Reporting) is the only way to get space occupancy without SRM

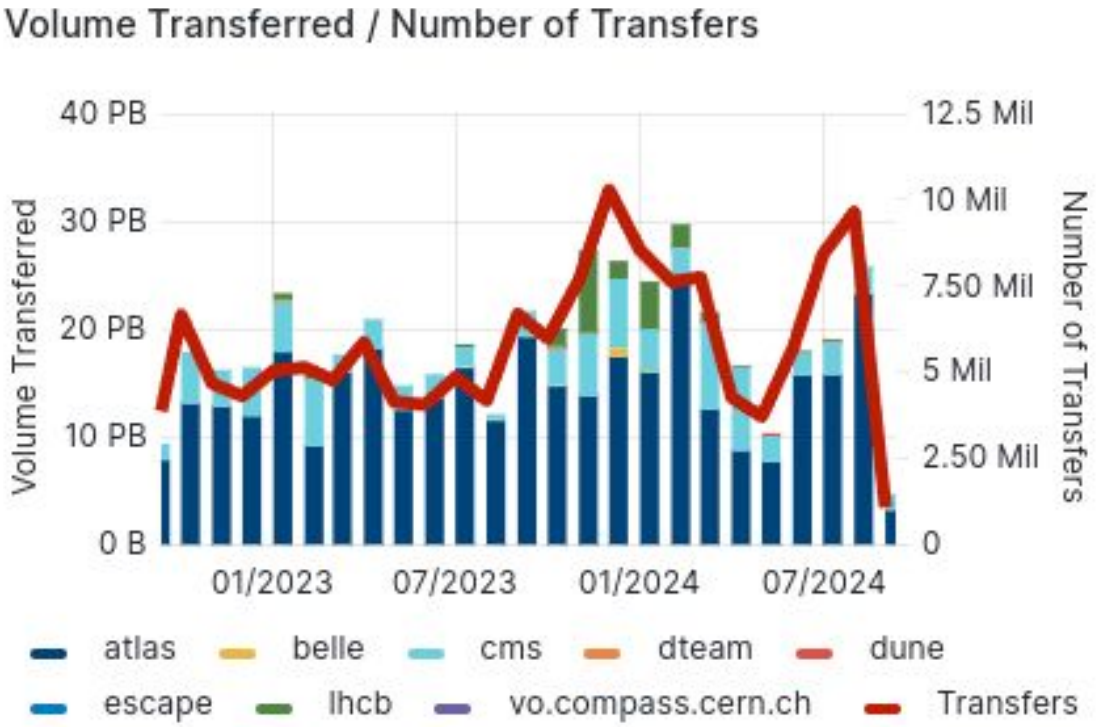
SRM vs. TAPE REST volume

- TAPE REST production transfers since April 2023
- T1 sites asked before summer 2023 to plan upgrades required to support TAPE REST
- Push for TAPE REST after DC24
- TAPE recalls comes mostly from ATLAS (187PB), CMS (51PB), LHCb (19PB), BelleII (1PB)



SRM vs. TAPE REST volume

- TAPE REST production transfers since April 2023
- T1 sites asked before summer 2023 to plan upgrades required to support TAPE REST
- Push for TAPE REST after DC24
- TAPE recalls comes mostly from ATLAS (187PB), CMS (51PB), LHCb (19PB), BelleII (1PB)



WLCG deployment status

Experiment	Production	Testing	No progress	Deadline	Comment
ATLAS	**9** (2*CTA, 6*dCache, 1*StoRM)	**2** (dCache)	**0**	end of 09/2024	RU Tape not considered
ALICE	-	-	**-**	-	Rely exclusively on xroot protocol
CMS	**6** (2*CTA, 3*dCache, 1*SToRM)	**2** (dCache)	**0**	asap	
LHCb	**8** (2*CTA, 5*dCache, 1*Storm)	**1** (CTA)	**0**	asap	RU Tape not considered
BelleII	**1** (dCache)	**2** (dCache)	**2** (dCache, StoRM)	none	“Pushed by sites” that would like to drop SRM
DUNE	**?**				Share T1 with REST enabled

Conclusion – SRM retirement

- LHC experiments are very close to fully retire SRM
  - End of September still seems realistic for ATLAS
    - Sites that moved to TAPE REST can stop providing SRM
      - Already done at some T1s
  - Future DMC tools may drop SRM from our point of view
    - Already discussed in FTS Workshop 2024
    - SRM is not considered to be part of future FTS & DMC tools
  - One more item that we can mark fulfilled on WLCG Globus Retirement list
- EGI is also working on SRM replacement
  - Ideally by the end of 2025
  - Plan to use HTTP & Check-In for transfers
- All users / experiments should plan how to **replace SRM by 2026**