The future of data protocols

Nordugrid conference
Cedric Serfon

Introduction

- Overview of the protocol zoo
- Getting rid of SRM and gridFTP

Disclaimer : talk heavily biased toward ATLAS

SRM : one interface to rule them all

- Not a protocol per se
- Storage Resource Management (SRM) is a pure HEP product (initiated by LBNL)
  - “SRM services are middleware components whose function is to provide dynamic space allocation and file management in spaces for shared storage components on the Grid. SRM services implement SRM interface.” [1]
- Main features :
  - Space Management and reservation (Space Tokens) : Way to allocate space orthogonal to the namespace. Never used it as it was expected
  - Retention Policy (now handled by VO specific data management layer)
  - Access Latency (On-line, near-line)
  - Provide translation between Storage URL (SURL, storage agnostic) and Transport URL (TURL)

SRM

- SRM has been the standard interface to Storage Element for WLCG (except Alice) for the last decade
- But :
  - Performance overheads for experiments
  - Work on developers to maintain
  - Restricts sites technology choices (e.g. Ceph)
- Recommendation from Storage/Data “TEGs”* :
  - “It is recognized that SRM scope is mainly restricted to the HEP community, and therefore a replacement should be considered for the long term if possible.” [2]

→ Plan to get rid of SRM

* Yet another WLCG working group (2012)

gridFTP

- Supported by Globus
- GridFTP is an extension of the File Transfer Protocol (FTP) for grid computing
- Underlying protocol for every 3rd party copy transfers
- But :
  - “We are announcing that, starting in January 2018, the Globus team at the University of Chicago will no longer support the open source Globus Toolkit, except for its use with the Globus cloud service by Globus subscribers.” [3]
  - The free lunch is over, now you’ll need $$$

→ Need (also) to get rid of gridFTP

xrootd

- Yet another HEP product, developed by SLAC
  - Protocol specification
- Primary (only) protocol for Alice
- Some nice features like redirection, native support in root

→ One candidate to replace SRM/gridFTP

WebDAV

- Web Distributed Authoring and Versioning (WebDAV) is an extension of the Hypertext Transfer Protocol (HTTP)
  - https://tools.ietf.org/html/rfc4918
  - Widely used
  - Supported by multiple clients
- WebDAV supported by almost all storage technologies used in HEP (dCache, DPM, StoRM…)

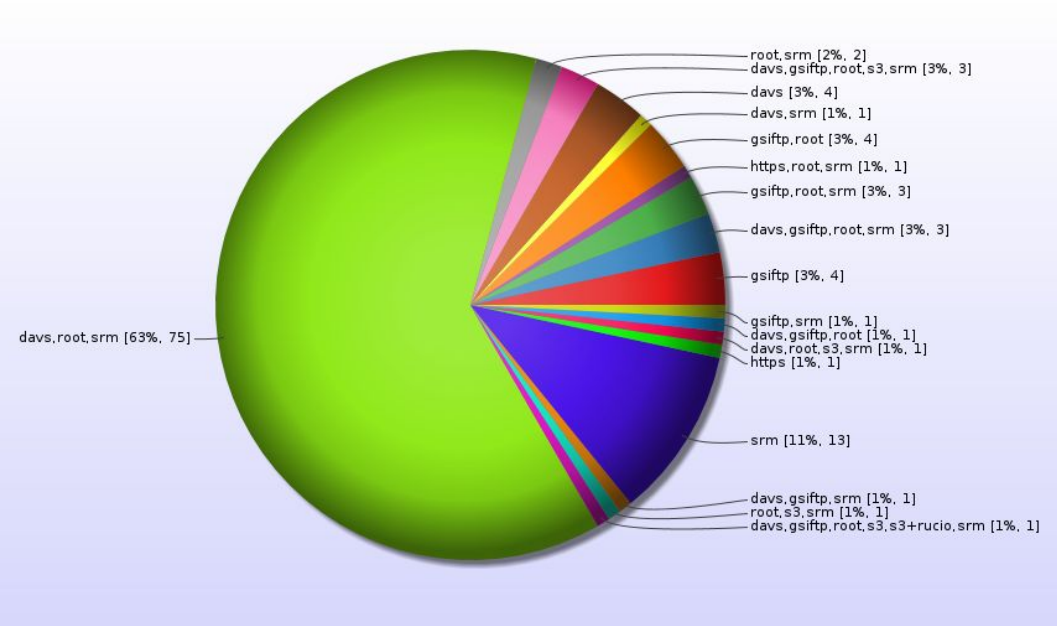
→ One other good candidate to replace SRM/gridFTP

Other protocols - aka the protocol zoo

- dCache related :
  - dcap : dCache access protocol
  - gsidcap : Same as the previous one + authentication
- rfio :
  - Deprecated
- S3, Swift
  - Cloud solutions
- You name it

The protocol zoo

Example of ATLAS (120 sites) :



- Most sites now provide now multiples protocols
  - The majority provides both WebDAV and xrood, in particular the biggest ones
  - Still some work to reach 100% deployment

Going away from SRM and gridFTP

- Need to check that protocols that can replace SRM/gridFTP, provide the same features :
  - Transfer 3rd Party (FTS)
  - Copy in/out
  - Deletion
  - Space reporting
  - Tape recall functionalities

Read/write/delete

- Read/write/delete supported by both WebDAV and xrootd
- Read/write with WebDAV/xrootd already in production on some sites
  - For WebDAV we already use some nice features like multi-source streaming
- Deletion through WebDAV enabled on more than 60% of ATLAS sites and through xrootd on some big sites (e.g. CERN) for more than a year
  - No big issues, but sometimes some “instabilities” maybe linked to the fact that we never requested the same QoS for these protocols (no critical SAM tests)

Space reporting

- One of the feature of SRM is the possibility to report in real time the space allocated/used by Space Tokens
  - Useful functionality used to monitor the consistency of the Storage wrt catalog (Dark Data)
- One solution proposed :
  - Publish a json format with information on space available/used, etc.
  - But 2 implementations with different format : One from ATLAS and one from WLCG. Work ongoing to uniformize it

TAPE staging in an SRMless world

- SRM provides tools to interact with TAPE (e.g. bring online)
- A successor of SRM should be able to provide the same functionalities.
  - New TAPE software, e.g. CTA (CERN Tape Archive) are already providing the possibility to use xrootd
  - What about the other Storage provides like dCache ?
- In the end, this functionality is just for Tier1s, i.e. ~10 sites
  - It should prevent us to move away from SRM for the 90% other sites

3rd party copy

- No standard for 3rd party copy (TPC)
  - For gridFTP, done via globus-url-copy
  - For WebDAV, done via davix-cp
  - For xrootd, done via xrdcp
- Test are currently being conducted
  - Fine with DPM, dCache so far
  - Some bugs find for the other storage flavours are being addressed
  - Plan is to have validation of TPC on all storage flavours by the end of the year
- For sites providing incompatible protocols, no TPC, but possibility to stream through the FTS server
  - Should be avoided as much as possible

New protocols functionalities

- Data Lakes (or Fjords) has become the new buzzword in the last months
- Allow to hide the complexity of the storages systems
- Protocols might evolve to handle new semantics like Quality of Service (see Vincent's talk)

Conclusion

- Are we ready to move away from SRM/gridFTP ?
  - For sites without TAPE system, almost (still needs a bit of work on the 3rd party copy). Timescale a few months
  - For sites with TAPE endpoints, more long term plans

References

[1] Introduction to SRM

[2] Data Storage TEG report

[3] Globus announcement end support gridFTP