



TPC Update



A. Forti, B. Bockelman
GDB
9th September 2020



TPC Working Group in Context

The DOMA TPC effort has been ongoing for 2 years. The aim was to grow alternative ecosystems to GridFTP for TPC and ensure they work in our multifaceted environment!

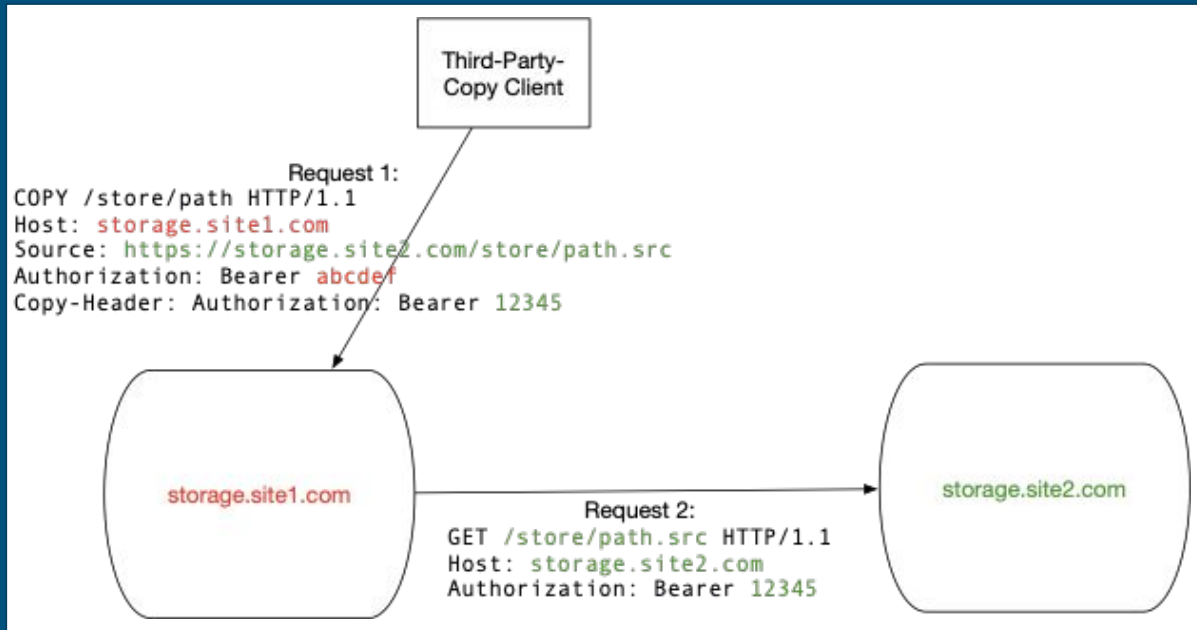
We quickly settled on two candidates protocols: xrootd and HTTP-TPC

- HTTP-TPC is a set of agreed-upon interpretations of the WebDAV protocol (built on top of HTTPS).
- TPC extensions to Xrootd were already present but required changes in the authentication for X.509 authentication and were not implemented in dCache.

Reference: Token use in HTTP-TPC

Since it's been raised here before, the way this works currently:

1. Authenticating with X.509, FTS requests a token from each side.
 - a. The token format is opaque to FTS.
2. The token from site X is included in the request to site X.
 - a. Only the issuing site needs to recognize its token.
 - b. Once a token is available, the request flow is as shown to the right.



TPC in Context

We are looking at the whole ecosystem, which also includes how authentication is done.
Beyond X.509:

- HTTP has a long history with token-based auth.
 - Current production transfers are X.509-based but, internally, use tokens (the X.509-authenticated connection receives a token that authorizes the transfer)
 - The transfer testbed is working on a pure-token-based scheme.
- After initially working on X.509 delegation, we are waiting on Xrootd-over-TLS to be used in production before enabling tokens for xrootd.
 - Work with X.509 delegation hit difficult deployment issues.

For two years we tested using smoke tests and the Rucio-based DOMA testbed:

- Now beginning to move data in production!

Putting Sites in production

ATLAS and CMS have started to cautiously enable HTTP-TPC at selected sites in production.

Internally, each has a different infrastructure and deployment approach:

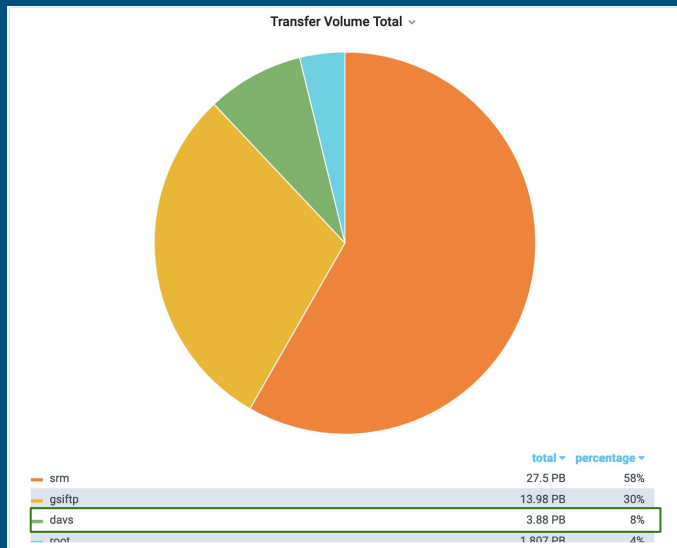
- ATLAS uses Rucio; once enabled for a destination, HTTP-TPC transfers start on all configured source sites. Rucio only avoids sites that are not configured in AGIS to serve data via HTTP-TPC.
- CMS uses PhEDEx and enables one single channel at the time. The changeover is managed by site admins.
 - CMS is in the process of changing from PhEDEx to Rucio; once the transition is done, a more aggressive rollout is expected.
 - Currently, a small software patch is required to be a HTTP-TPC destination site.

Both VOs do additional functional testing before putting a site into production with HTTP-TPC.

ATLAS

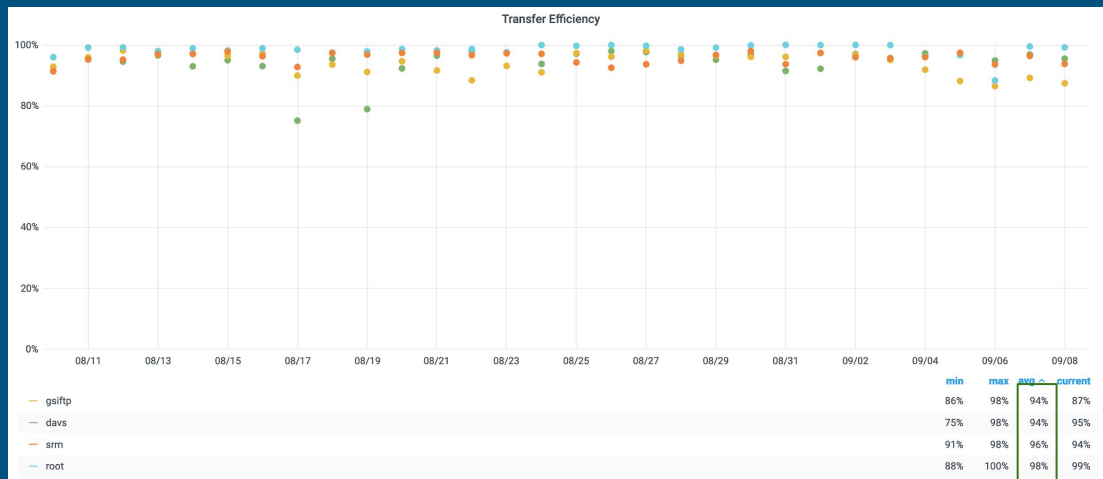
- Concentrate on HTTP-TPC
 - Configured FTS to do only pull transfers
 - Active sites copy from other sites rather than to
 - Only uses macaroons, temporarily falls back on gridsite delegation if source hasn't enabled macaroons
 - Uses other protocols if HTTP-TPC not configured at the source
- Concentrate only on dcache (>5.2.14) and DPM sites (1.14.0)
 - Dcache works out of the box, DPM sites need to upgrade (ongoing campaign).
 - Some sites will need modest configuration changes.
- 10 sites enabled as active destinations so far
 - 6 dcache, 4 DPM
 - 3 T1 (FZK, PIC, IN2P3-CC), 6 T2 (AGLT2, Prague, IF AE, DESY, Liverpool, Manchester), 1 T3 (Brunel)
 - MWT2 and CERN planned for next week, DPM sites as they upgrade

ATLAS (2)



Last 30 days: 3.8PB, 8% of total transfers

Last 30 days: efficiency in line with other protocols
Performance comparison tests are also planned

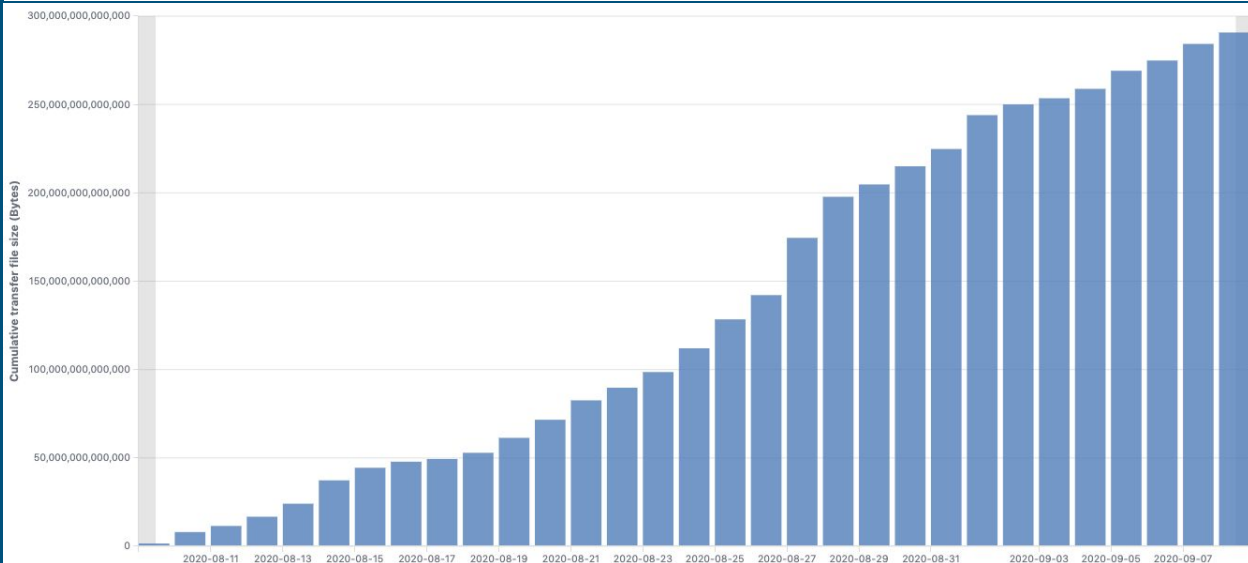


CMS

- Currently only deploying HTTP-TPC.
- Using 2 sites as main destination SEs: UCSD and Nebraska
- Enabling only “reads” from other sites (until PhEDEx gets retired)
- Using macaroons for auth for the COPY (falls back to X.509)
- 11 sites in Production (dCache, XRootD)
 - T1s: FNAL, KIT and JINR
 - T2s: DESY, MIT, Florida, Caltech, Wisconsin, Nebraska, Purdue, UCSD
- 7 sites in progress (dCache, XRootD, DPM)
 - CSCS, PIC, TIFR, IN2P3, CNAF, UCL, London_Brunel
- Main issues found when enabling a site:
 - Wrong rule(s) in the Trivial file Catalog for LFN to PFN translation
 - Lack of PhEDEx Debug transfers from the site even when enabled
 - Setting up the macaroons configuration

CMS (2)

Cumulative transferred bytes using WebDAVS within the last 30 days



Within the last 30 days:

- Total transferred: 18.3PB
- Transferred with WebDAV: 0.3PB (1.6%)
- Total transferred by UCSD and Nebraska: 0.82PB
- Transferred by UCSD and Nebraska via WebDAV: 0.22PB (26.8%)

Source: <https://monit-kibana.cern.ch/kibana/goto/fdf879b76459f058796abff4e1437e31>

Technology Deployment Status

- As with any significant new feature, typically the “latest greatest” release captures the learned operational experience (aka, bugfixes!)
- A note on versions for HTTP-TPC:
 - dCache: HTTP-TPC works with 5.2.* releases better if $\geq 5.2.15$
 - DPM: 1.14.0 with macaroons enabled
 - WLCG Ops coordination opened 45 GGUS tickets, deadline end of the year. WLCG Ops is currently only checking for the minimum version, not that the configuration is working for a given VO.
 - StoRM: 1.11.18
 - XRootD: 4.12.4. Several sites rely on the HTTP-TPC implementation shipped with the XRootD software (recall: XRootD implements multiple transport protocols!).
- The DOMA-TPC working group smoke test is a great way to validate functionality.

Technology Deployment Status

Notable omissions from the prior slide:

- EOS: Uses XRootD *but*, given the sheer size of deploys, has been finding scale bugs not observed at other sites. We believe **known issues are fixed in XRootD 4.12.4**.
- Echo: xrootd-TPC works; tweaking configuration of XRootD to get HTTP-TPC working.
- CTA: Current versions work only with xrootd-TPC. For ATLAS and CMS, this is OK: all transfers to/from CTA are within CERN to EOS. LHCb wants to do direct transfers from CTA to T1s; at least CNAF will require HTTP-TPC.
 - Approach currently under investigation.

Technology Deployment Status

NOTE that, in the OSG:

- The current release series (OSG 3.5) is the last one to support GridFTP.
- **The next release series (OSG 3.6; expected Q1 2021) will not support GridFTP.**
 - The two series will be supported in parallel for some time.

Consider the timing of Run 3: the HTTP-TPC transition is a **LS2 activity**, not LS3!

XRootD 5.0

The XRootD 5.x series adds many new features - the most exciting of which is the support for xrootd-over-TLS (xrootd protocol version 4.0).

- This provides support for integrity checking and encryption using the well-understood TLS protocol.
- Critically, encryption permits **use of the bearer tokens for authorization**.

We strongly encourage the rollout of new clients that understand the protocol.

- Integrators are working to validate the server.
- One current blocker bug (fixed, waiting on 5.0.2) prevents HTTP-TPC from working with FTS.

Token Transition

Now that alternate transfer protocols are beginning to hit production, R&D is focusing on the transition away from X.509 client authentication:

- As of the WLCG Hackathon in January 2020, we can do “X.509-free” transfers through FTS with several of the storage implementations.
- We are working on a similar pattern as before:
 - Adapt the smoke-test framework to use tokens; establish a baseline of sites that work at least once every morning.
 - Increase sites in the [rucio-based tests](#) to continuously test at a low-level.
 - Use the Rucio tests to gain operational comfort until we are ready for production.
- Leveraging the WLCG IAM instance to issue transfer authorizations.

Replacing SRM at T1s

- Outside of operating tape archives, any site can today transition away from SRM.
- We still have no agreed-upon, interoperable protocol for managing archives over the WAN.
 - Although we are relatively close!! CTA has shown a tape archive can be managed completely through the xrootd protocol.
- In the meantime, we are looking at:
 - Splitting up the task in FTS: SRM to stage to disk, HTTP-TPC to transfer.
 - Adding non-X.509 auth methods to SRM. dCache already has a working prototype for this!

Next step

Add more sites to production HTTP-TPC

More tests!

- Performance tests comparison between protocols in production (soon)

- Increase OIDC testbed (soonish)

- Add SRM tape endpoints to the tests (longer term)

- Networking tests (long term)

Second token-based [AuthN/Z Hackaton](#)

Summary

- At long last, real production transfers are being performed by a non-GridFTP protocol!
 - In February 2020, there was ~zero production activity.
 - This is currently below 10% of volume globally but growing quickly. Over 50% of the volume at some sites!
- All sites should be preparing a production deployment.
 - Current priority is HTTP-TPC.
 - This is a LS2 activity!
- XRootD 5.0 brings important capabilities to TPC and transfers in general - integrators should be targeting this version.