DOMA-TPC testing

ESCAPE F2F meeting 2019-07-01 Amsterdam

https://indico.in2p3.fr/event/19214/

Background: we move data wholesale!

- WLCG relies heavily on **third-party copy** (TPC) to move data
  - A coordinating service (FTS) instructs one endpoint to send data, the other to receive data.
  - Data flows **directly between endpoints**: **not** relayed through the coordinating service.
- Currently TPC mostly done with (an extension of) the FTP protocol: GridFTP
  - Due to control/data channel separation, FTP support TPC out-of-the-box
  - GridFTP adds extra features, such as multi-stream support, persistent data channels, and data integrety verification.

Why work on something that's been a robust production solution for ~20 years?

- In Spring 2017, Globus announced they are **ending support** for the Globus Toolkit (GT) in 2018-01-01.

- Almost all WLCG software stacks (client & server) use GT for GridFTP support:

  dCache doesn't ☺

- There is now a community effort to support the GT software by the Grid Community Forum.

  … however, people nervous about future support effort.

- Good opportunity to evaluate alternatives.

A problem with two solutions

- Two contenders for a replacement protocol: xrootd-TPC and HTTP-TPC

    - xrootd is a proprietary protocol, championed within WLCG

    - HTTP is an industry standard, but HTTP-TPC is a proprietary extension

- Both xrootd-TPC and HTTP-TPC have multiple implementations.

- Unclear if either (or both) will provide a potential replacement to GridFTP.

- Tests are needed to make an informed decision.

DOMA-TPC

- **DOMA-TPC** is an activity within the DOMA working group.

- Charged with the **investigation** of xrootd-TPC and HTTP-TPC and to oversee the **roll-out** of these alternatives to a production environment.

- WLCG sites responded to a **call for resources** and have deployed new versions of software and allocated resources.

- Together, they form the **DOMA-TPC testbed** within which we can evaluate the different protocols and software implementation.

- It has evolved over time, but currently we have two main testbeds:

  "**Proving ground**" (32 sites, 40 storage systems)
  "**Stress testbed**" (7 sites/storage systems)

Rucio testing: our work-horse

- Rucio is software that **coordinates data placement**, using FTS to handle **data movement** (when needed)

  - Used to test **both** HTTP-TPC and xrootd-TPC

  - Storage systems registered in Rucio with a single protocols (not a limitation in Rucio, a testing choice by DOMA-TPC).

- Proving ground:

  once an hour, a single file is uploaded to an endpoint and registered in Rucio. This file is requested at another endpoint, retrying as needed.

- Stress-test:

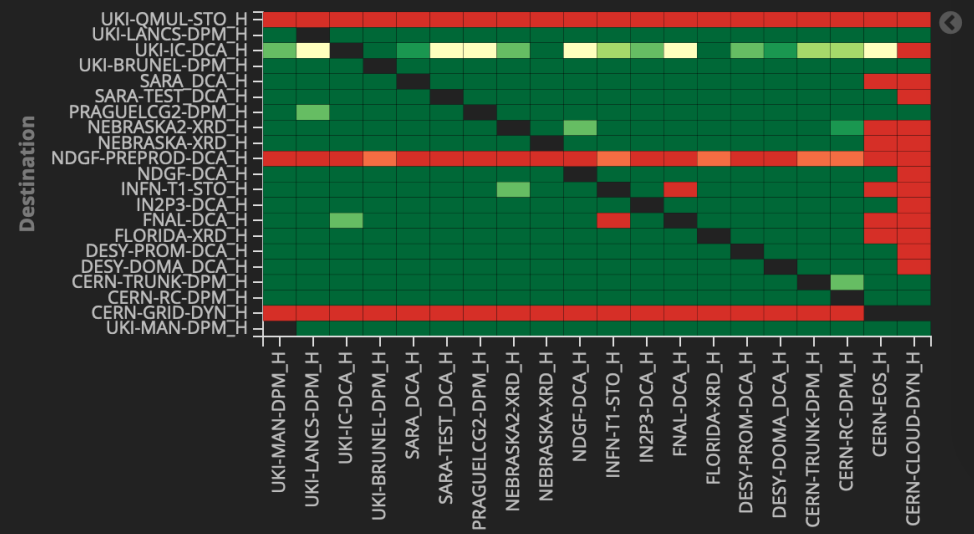 More data is registered in Rucio → more transfers and higher load.

Rucio DOMA - Heatmap (Root)
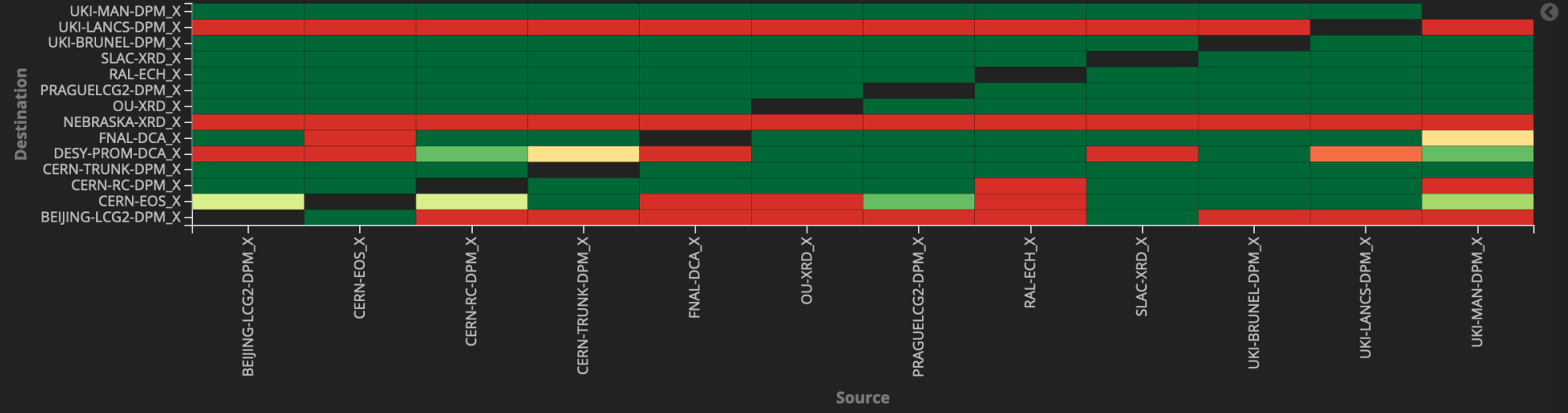
Rucio DOMA - Heatmap (DAVS) 1

Rucio DOMA - Heatmap (DAVS) 2

Rucio DOMA - Heatmap (DAVS) 1
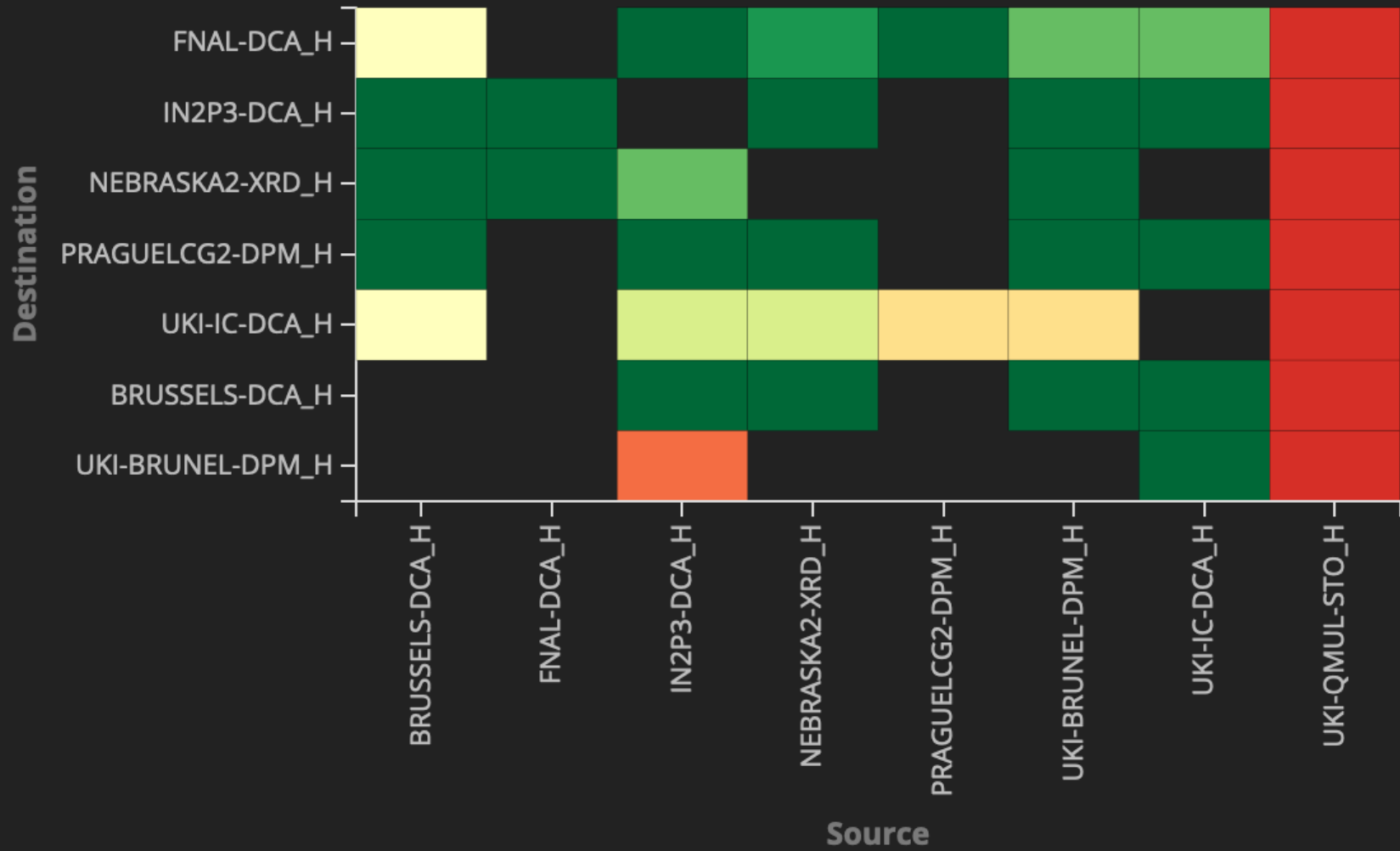
HTTP smoke tests:

- Originally written to overcome some **limitations** with the Rucio/FTS testing.

- A "simple" script, just runs different **curl commands**.

  Limited requirements, relatively easy to deploy

- Two groups of test: **direct tests** and **TPC tests**:
  - Direct tests simulate an HTTP-TPC transfer (test each interface),
  - TPC tests test if HTTP-TPC transfers work.

- Two **modes**: "overview" and "stop-on-failure"

- Includes some **work-rounds** for problems/bugs in server software where a solution is not yet available.

HTTP smoke tests: example

Target: https://dcache-se-doma.desy.de:2880/dteam/smoke-test-discordia.desy.de-4117

DIRECT TRANSFER TESTS

Checking 131.169.191.70 (1 of 2)

Uploading to target with X.509 authn: **SUCCESS**
Downloading from target with X.509 authn: **SUCCESS**
Obtaining ADLER32 checksum via RFC 3230 HEAD request with X.509 authn: **SUCCESS**
Deleting target with X.509 authn: **SUCCESS**
Request DOWNLOAD,UPLOAD,DELETE,LIST macaroon from target: **SUCCESS**
Uploading to target with macaroon authz: **SUCCESS**
Downloading from target with macaroon authz: **SUCCESS**
Obtaining ADLER32 checksum via RFC 3230 HEAD request with macaroon authz: **SUCCESS**
Deleting target with macaroon authz: **SUCCESS**

Checking 2001:638:700:10bf::1:46 (2 of 2)

Uploading to target with X.509 authn: **SUCCESS**
Downloading from target with X.509 authn: **SUCCESS**
Obtaining ADLER32 checksum via RFC 3230 HEAD request with X.509 authn: **SUCCESS**
Deleting target with X.509 authn: **SUCCESS**
Request DOWNLOAD,UPLOAD,DELETE,LIST macaroon from target: **SUCCESS**
Uploading to target with macaroon authz: **SUCCESS**
Downloading from target with macaroon authz: **SUCCESS**
Obtaining ADLER32 checksum via RFC 3230 HEAD request with macaroon authz: **SUCCESS**
Deleting target with macaroon authz: **SUCCESS**

THIRD PARTY PULL TESTS

Initiating an unauthenticated HTTP PULL, authn with X.509 to target: **SUCCESS**
Deleting target with X.509: **SUCCESS**
Initiating an unauthenticated HTTP PULL, authz with macaroon to target: **SUCCESS**
Deleting target with macaroon: **SUCCESS**
Requesting (from prometheus) DOWNLOAD macaroon for a private file: **SUCCESS** (doesn't count)
Initiating a macaroon authz HTTP PULL, authn with X.509 to target: **SUCCESS**
Deleting target with X.509: **SUCCESS**
Initiating a macaroon authz HTTP PULL, authz with macaroon to target: **SUCCESS**
Deleting target with macaroon: **SUCCESS**

THIRD PARTY PUSH TESTS

Third party push target: https://prometheus.desy.de:2443/VOs/dteam/smoke-test-push-discordia.desy.de-4117

Requesting (from prometheus) UPLOAD,DELETE macaroon to third party push target: **SUCCESS** (doesn't count)
Uploading target, authn with X.509: **SUCCESS**
Initiating a macaroon authz HTTP PUSH, authn with X.509 to target: **SUCCESS**
Deleting file pushed to third party, with X.509: **SUCCESS**
Initiating a macaroon authz HTTP PUSH, authz with macaroon to target: **SUCCESS**
Deleting file pushed to third party, with X.509: **SUCCESS**
Deleting target with X.509: **SUCCESS**

Of 32 tests: 32 successful (100%) Work-arounds: (none)

Daily smoke tests

- Smoke tests proved popular, now done daily with reports sent to mailing list.

  Initially done manually, now fully automated.

- Recent innovations:

  – Testing each interface (IP address) registered in DNS,

  – Integration with GOCDB: don't test endpoints in downtime.

  – Point system: endpoints have between 0 points (bad) and 20 points (good).

Example daily smoke tests

```
DOMA-TPC smoke test 2019-06-28T13:03+0200

SOUND ENDPOINTS

SCORE   ENDPOINT      SOFTWARE        WORK-AROUNDS
12      AGLT2         dCache                          [in 02:06]
12      BEIJING-TEST  DPM          [S]                [in 01:04]
12      BNL           dCache                          [in 00:39]
12      DESY-DOMA     dCache                          [in 00:31]
12      DESY-PROM     dCache                          [in 00:23]
12      FLORIDA       xrootd/Lustre                   [in 01:27]
12      KIT           dCache                          [in 00:31]
12      LRZ-LMU       dCache                          [in 00:40]
12      NEBRASKA2     xrootd/HDFS  [L]                [in 01:17]
12      TRIUMF-PPS    dCache                          [in 01:00]
12      UKI-MAN       DPM          [S]                [in 00:36]
11      SARA          dCache                          [in 02:39]
10      IN2P3         dCache                          [in 00:29]
10      NEBRASKA      xrootd/HDFS                     [in 00:53]
10      PRAGUELCG2    DPM                             [in 00:29]
10      SARA-test     dCache                          [in 00:33]
10      UKI-LANCS     DPM          [S]                [in 00:48]
9       FNAL          dCache                          [in 01:05]
9       NDGF-PREPROD  dCache                          [in 00:27]
8       CERN-TRUNK    DPM                             [in 00:42]
8       INFN-T1       StoRM                           [in 00:59]
6       PIC-PPS       dCache                          [in 17:00]
4       UKI-BRUNEL    DPM          [S]                [in 00:52]
3       BRUSSELS      dCache                          [in 01:50]
2       PIC-PROD      dCache                          [in 02:06]
1       FLORIDA2      xrootd/Lustre [L]               [in 01:06]
1       UNI-BONN      xrootd/CephFS [L]               [in 00:30]
```

```
PROBLEMATIC ENDPOINTS

SCORE   ENDPOINT        SOFTWARE      SUMMARY
10      NDGF            dCache        Of 32 tests: 1 skipped (3%), 31 attempted (96%):
30 successful (96% of tests run, 93% of all tests), 1 failed (3% of tests run, 3% of
all tests) Work-arounds: (none)     [in 10:35]
4       TRIUMF-DYNAFED  DynaFed       Of 23 tests: 18 skipped (78%), 5 attempted
(21%): 0 successful, 5 failed (100% of tests run, 21% of all tests) Work-arounds:
(none)                                          [in 06:31]
0       CA-IAAS         DynaFed       Of 23 tests: 15 skipped (65%), 8 attempted
(34%): 1 successful (12% of tests run, 4% of all tests), 7 failed (87% of tests run,
30% of all tests) Work-arounds: (none)     [in 00:12]
0       CALTECH         xrootd/HDFS   Of 33 tests: 25 successful (75%), 8 failed (24%)
Work-arounds: L
                                            [in 01:30]
0       CERN            EOS           Of 32 tests: 25 skipped (78%), 7 attempted
(21%): 0 successful, 7 failed (100% of tests run, 21% of all tests) Work-arounds:
(none)                                          [in 00:05]
0       CERN-RC         DPM           Of 32 tests: 28 successful (87%), 4 failed (12%)
Work-arounds: (none)
                                            [in 00:36]
0       UKI-GLASGOW     DPM           Of 23 tests: 10 skipped (43%), 13 attempted
(56%): 8 successful (61% of tests run, 34% of all tests), 5 failed (38% of tests run,
21% of all tests) Work-arounds: F       [in 00:27]
0       UKI-QMUL        StoRM         Of 32 tests: 9 skipped (28%), 23 attempted
(71%): 18 successful (78% of tests run, 56% of all tests), 5 failed (21% of tests
run, 15% of all tests) Work-arounds: (none)  [in 00:17]
0       UKI-RAL         DynaFed/ECHO  Of 23 tests: 13 skipped (56%), 10 attempted
(43%): 5 successful (50% of tests run, 21% of all tests), 5 failed (50% of tests run,
21% of all tests) Work-arounds: (none)  [in 00:15]

SKIPPED ENDPOINTS

SCORE   ENDPOINT   SOFTWARE   WHY
4       UKI-IC     dCache     GOCDB Downtime: Move to Slough part 2
```

DOMA-TPC future testing

- Move towards production testing:

    ATLAS and CMS are involved, ATLAS taking the lead.

- Xrootd smoke tests under development.

- Evaluating both HTTP-TPC and xrootd-TPC against GridFTP, in terms of performance.

Thanks for listening