Bulk Data Transfers WG Update (DOMA General Meeting)

August 2022

Status

- Tape REST API – details with timeline in last DOMA GENERAL
- Packet marking update planned after summer
  - Development in progress – flowd plugins (firefly, packet tagging)
    - Support for SE without native packet marking capability
  - Packet tagging demo for Supercomputing 2022
- Data transfers – today's update
  - Turning off GridFTP doors for TAPE sites (FTS & SRM+HTTPS support)
  - Transfers with tokens
    - Compliance with JWT profile
    - Progress towards DC24 with tokens
  - Missing HTTP-TPC features

GridFTP doors and TAPE sites

- BNL removed GridFTP doors, causing FTS bringonline protocol mismatch
  - The FTS protocol request list has been made configurable in 3.12.1 release
    - Currently deployed on CERN Pilot, production CERN FTS next week, BNL FTS server second week in September
    - This update still needs to be validated for production use
      - Manual tests seems to work OK
  - Technical details discussed in GGUS:158503 and FTS-1824 + FTS-1825
- Tape SE with SRM+HTTPS must keep GridFTP doors few more weeks
  - First all FTS servers must be updated to the latest version 3.12.1

Transfers with tokens

- Be ready to test tokens at scale for transfers during DC24 (WLCG TTT)
  - Including transfers that involve interaction with TAPE using new REST API
- Storage
  - Current set of compliance tests passed by StoRM, dCache and XRootD 5.5
    - EOS and Echo should be fine once they upgrade testbed to XRootD 5.5
    - Plan to include something similar for SAM/ETF tests with tokens
    - We'll add more compliance tests when we discover incompatibilities (e.g. JWT profile discussion issue#21)
  - Production SE upgrade / token configuration
    - Collected details about common storage configuration for WLCG experiments
      - Provide recommended configuration for each SE implementation (developers & sites)
    - Plan to have at least fraction of SE with token support already in Q1 2023
      - Winter shutdown may be good opprotunity to introduce tokens for SE
      - Rucio & FTS transfer tests with tokens
      - Become ready also for developing / testing jobs with tokens (WFMS)
  - Use X.509 and tokens at same time with production storages
    - Permissions (DACL vs. Capability) and namespace organization
- Rucio
  - New developer available – improve existing token support and implement new Rucio token workflows (upload & download)
  - Testbeds WLCG JWT (wlcg) vs. EGI profile (ESCAPE)
- FTS
  - Code refactoring

FTS / gfal2 HTTP-TPC "missing features"

- GridFTP protocol configuration exposed by FTS / gfal2 offered more features
  - Multi-stream transfers
    - Partially implemented (and usually broken)
      - Should we invest time to fix / implement or is ***single stream acceptable by DOMA***?
    - Could this functionality solve (or make more visible) issues with rare slow transfers
      - Reduce effort we invest in rare slow transfer investigation (GGUS:157985)
    - Technical details will be discussed in the next BDT meeting
  - TCP buffer size
  - IPv4 vs. IPv6 preference and monitoring (fixed FTS-1802)
    - Do we need an option to force IP version or instead we should rely on happy with happy-eye-ball (RFC8305)
    - Performance markers not available from StoRM HTTP-TPC
    - Improved logging including remote server IP address could be useful for debugging issues
      - Only few sites can easily analyze transfer logs (e.g. by using their own ElasticSearch instance)
      - Distributed storage sites (data lakes)
  - Transfer timeouts
    - Kill stuck transfers with zero progress or even too low speed early (DMC-1278) or let them finish
    - e.g. GridFTP killed transfers with zero progress after 5 minutes, HTTP-TPC waits for total transfer timeout (hours for big file transfer submitted by Rucio) – very long "active" FTS transfer
- We have more different HTTP-TPC implementations compared to GridFTP protocol

Same FTS configuration interface for all protocols that doesn't always support same feature set

FTS can now disable proxy delegation for HTTP-TPC DMC-1229