Storage & transfers with tokens

WLCG DOMA Bulk Data Transfers WG

WLCG Workshop, Lancaster

November 7, 2022

WLCG DOMA Bulk Data Transfer (BDT) WG

- Coordinate, develop and integrate new technologies for data movement
  - Broader scope than Third-Party-Copy (TPC) WG
    - Already discussed within TPC WG (more accurate name)
  - Mostly projects with well defined goals
- The activities focused on topics related to the data transfers
  - ***WLCG JWT tokens for storage & transfers***
  - Network utilization visibility (packet marking)
  - Archive management (SRM replacement)

  } Data Challenge

  — GCT retirement
- Biweekly meetings first and third Wednesday starting at 16:30
  - Mailing list: wlcg-doma-tpc@cern.ch

Transfers with WLCG JWT tokens – overview

- ***Storage services (compliance, deployment / configuration)***
- Data management and transfer services
  - Both FTS and Rucio have basic support for transfers with tokens
  - Upcoming Dirac 8 basic token support and TokenManager (delegated refresh token)
  - This needs to be improved
    - New Rucio developer available with this task (WLCG Authz ideas, Rucio Workshop details)
    - Limited duration 3 year project for tokens in FTS will start at the beginning of 2023 - an additional FTS developer will be hired during this period
- User interaction with clients (gfal2, Rucio, …)
  - Storage tokens from IAM or may be only from data management services
  - Tokens should be obtained transparently without end user interaction
    - Details still needs to be discussed in WLCG AuthZ WG (ARC/HTCondor-CE Hackaton notes)

WLCG JWT storage compliance tests

- Tests with scope and group authz
  - tokens from WLCG IAM
  - storage configuration requirements
    - normal vs. protected area
  - HTTP protocol only
    - critical vs. non-critical tests
    - xroot behavior hopefuly similar
  - no combined testing with X.509
- Fresh compliance test results every day
  - critical tests OK for all supported SE implementations
    - dCache, EOS, Echo, StoRM, XRootD
      - EGI DPM GGUS migration campaign
    - include additional instance with a pull request
- New tests added as we gain experience
  - tokens with wlcg.groups sufficient w/o scope
  - 5 test added recently – total 26 tests
    - cover JWT issue#21 (+1 pending test)
- Standard protocol -> many client libraries

Statistics by Tag	Pass	Fail	Pass / Fail / Skip
critical	369	39	
not-critical	8	26	
se-cern-eos	20	6	
se-cnaf-amnesiac-storm	24	2	
se-florida-xrootd	23	3	
se-florida-xrootd-redir	23	3	
se-fnal-dcache	26	0	
se-infn-t1-xfer-storm	24	2	
se-nebraska-xrootd	20	6	
se-nebraska-xrootd-redir	18	8	
se-prague-dcache	14	12	
se-prague-xrootd	24	2	
se-prometheus-dcache	26	0	
se-ral-test-xrootd	22	4	
se-ubonn-xrootd	24	2	
se-ucsd-xrootd	23	3	
se-ucsd-xrootd-redir	22	4	
se-wisconsin-xrootd	22	4	
se-wisconsin-xrootd-redir	22	4	

Storage configuration

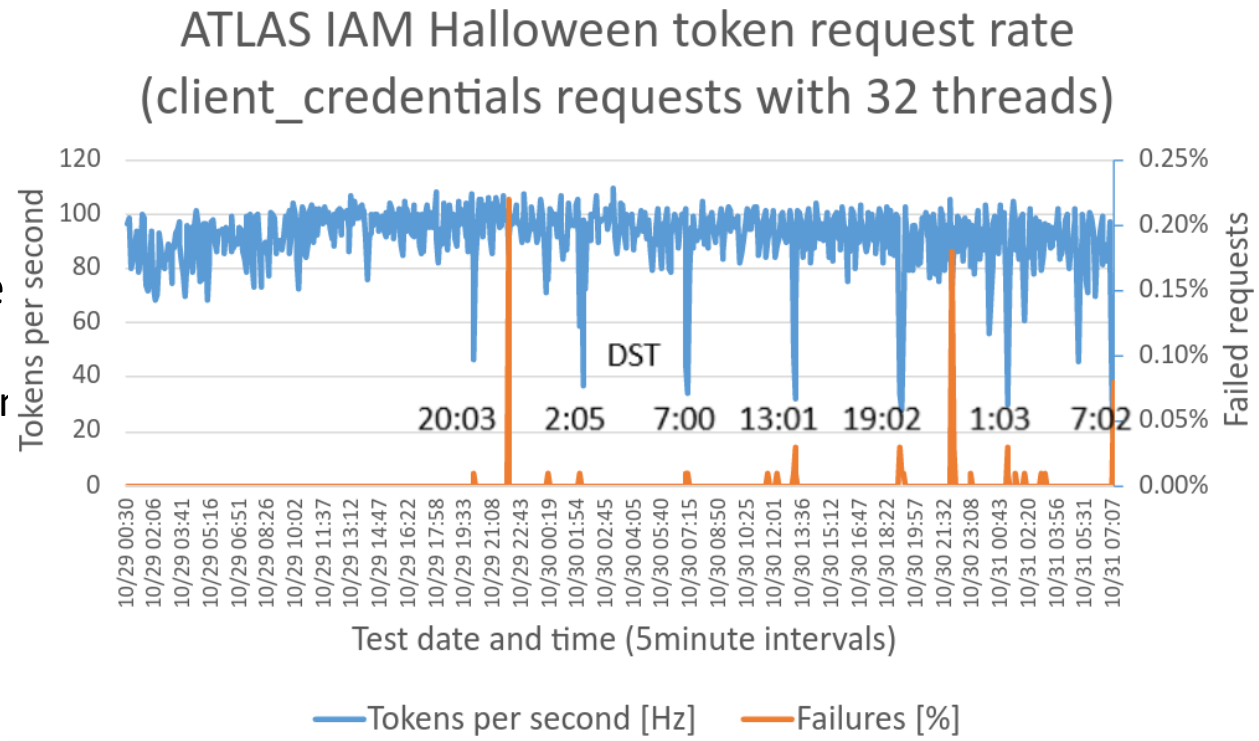
- Collected requirements from our experiments
  - Very similar mapping and access permissions
    - Usually distinguish just few VOMS roles, no read restriction for VO users
  - All experiments prefers capability based authz
    - `storage.create` significantly reduce risk of abuse for (job) tokens used to write data
- Storage must work with X.509 and tokens at the same time
  - Plan to provide simple examples for each storage implementation
    - all support SE implementations and both HTTP & xrootd protocols
  - Started discussion with developers/experts
    - tools for quick and easy deployment might be necessary
- Capability & storage namespace
  - Non-essential IAM `storage.*:/$PATH` token exchange and scope policies makes sense only when all sites provides same namespace structure
    - `storage.*:/$PATH` is unique feature of WLCG JWT profile and IAM implementation
  - Not always the case (at least for ATLAS with multiple RSEs per site)

Storage identity mapping for tokens

- Simple once we move completely to the tokens
  - With capability model access policy is defined by data management services
  - Easier for storage administrators – map whole VO to one identity
  - Can't be combined with posix access (no capability), e.g. NFS/GPFS mounts
    - Read-only access for users from corresponding VO might be OK
- Increased storage configuration complexity during transition to tokens
  - Interoperable and secure support for both authz methods – X.509 and tokens
  - Just config changes to add support for tokens (no development)
    - dCache – inheritable ACLs, for some VOs there may be a simpler configuration
    - StoRM – fine grained authorization and ACL configuration (BelleII already configured)
    - XRootD based storages (XRootD, EOS, Echo) – support for storage path mapping
- Personal storage area
  - Not yet discussed, provided only by few sites (optional for CMS)
  - Secure support for multiple token issuers
    - path mapping functionality in XRootD might make configuration easier
    - thousands of scope policies in the IAM probably never tested

Storage & tokens timeline

- Be ready to do transfers with tokens "at scale" during DC24(WLCG token timeline)
  - testbeds & few production instances close to experts already by the end of 2022
    - allow development of new SAM/ETF transfer tests with tokens
      - include also compliance tests
  - final examples for sites in January 2023
  - GGUS campaign in the spring 2023
  - storages with no token support may not be able to participate in DC24
- IAM tokens & Rucio+FTS considerations
  - naive token per-write operation don't scale
  - More pragmatic approach
    - reduce required request rate by order of magn
    - less granular tokens (storage.*:/)
    - limit security implications
      - can we avoid storage.modify(?)
      - active party tokens in HTTP-TPC push
    - may not fit requirement of all VOs



TAPE & tokens

- Some implementations supports SRM with tokens – no plans to try this method
- Move away from SRM and deploy TAPE REST (next talk HTTP TAPE REST API Status)
  - Doesn't automatically means support for tokens
    - support for `storage.stage` capability
  - First focus on deployment with X.509
    - some sites would like to move to TAPE REST as soon as possible
      - e.g. RAL Antares to optimize LHCb transfers to tape
    - significant number of sites could have TAPE REST available in 2023
      - e.g. BNL plans to upgrade to dCache 8.2 already this December
  - Start with site-by-site migration already in 2023
    - CTA and dCache implementations already exists, StoRM still WIP
- No plans to use TAPE in DC24
  - More flexible timeline for tokens and TAPE transfers
  - We should identify what's missing soon
    - and run tests with tokens in 2023/4 (data management developers available)
    - long development -> testing -> deployment cycle

WebDAV Error Message Improvement Project

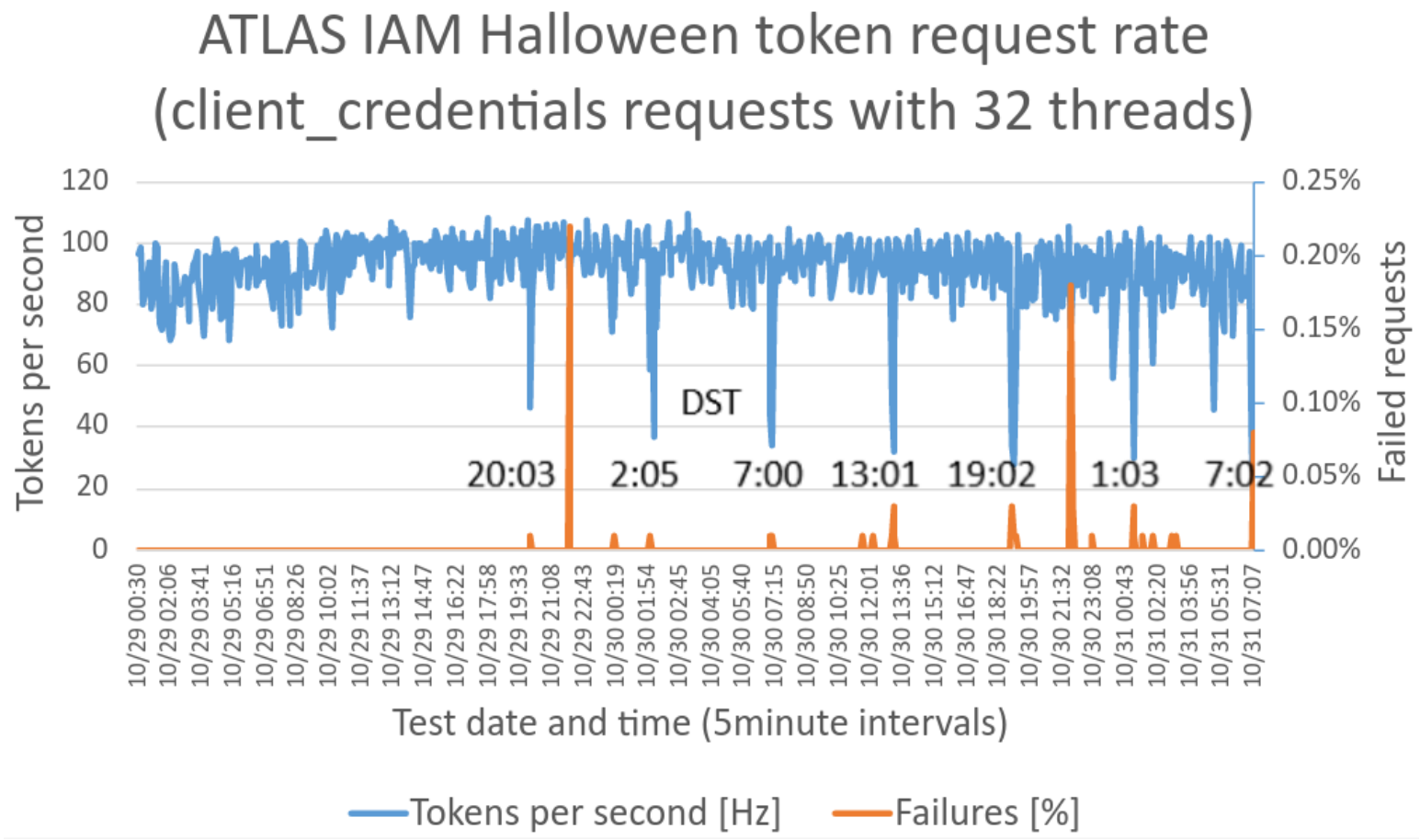
- Failed HTTP transfers don't always provides enough details
  - Too generic error messages
    - Difficult to understand error origin and what's causing transfer failure
    - Slow diagnosis, necessary to involve more people
  - Grid storages comes with several different HTTP implementations
    - GridFTP provided just by GCT or dCache
  - More complex TPC transfers with pull, push and streaming mode
- CMS came with proposal to improve HTTP error reporting
  - All experiments should collect poor HTTP error reporting in the twiki
    - Not just HTTP-TPC, but also normal two-party uploads/downloads
    - Production storages with supported sw (exclude DPM errors)
  - DOMA BDT meetings – time slot reserved for discusion with experts
    - Identify problematic component and create ticket

HTTP-TPC (COPY) protocol updates

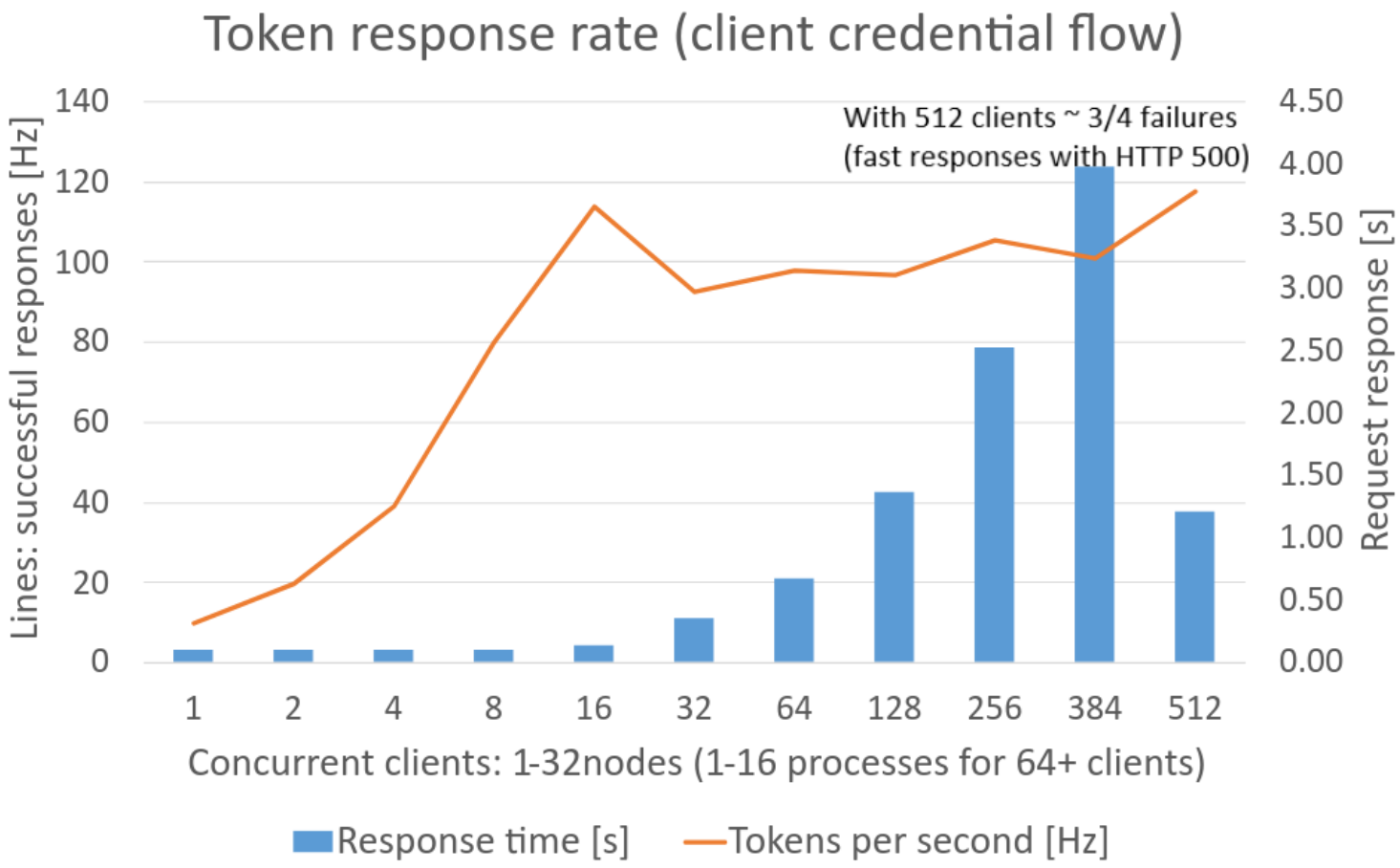
- There are more HTTP protocol implementations compred to proprietary GridFTP
- With year+ prod experince we see original specification might need updates
  - Clarify requirements in the existing technical specification
  - Come to conclusion which GridFTP features should (not) be implemented
    - e.g. Multistream, TCP buffer size, IPv4 vs. IPv6 preference, …
    - Same FTS configuration interface for all protocols
  - Improve operational experience / better transfer traceability and error reporting
- Process to propose HTTP-TPC improvements documented in twiki
  - Collect all information at one place
    - Icluding `TransferHeader` used by clients (pass HTTP headers to the passive party)
  - Discuss in BDT meeting / via associated mailing list
  - Get agreement from involved parties and set timeline
    - Protocol updates – HTTP-TPC draft
    - Supported SE – active vs. passive TPC party
    - Client changes (FTS, gfal2, Rucio, Dirac, …)
- Make everything backward compatible
  - Protocol versioning not defined

BACKUP

ATLAS IAM Halloween test (source data)

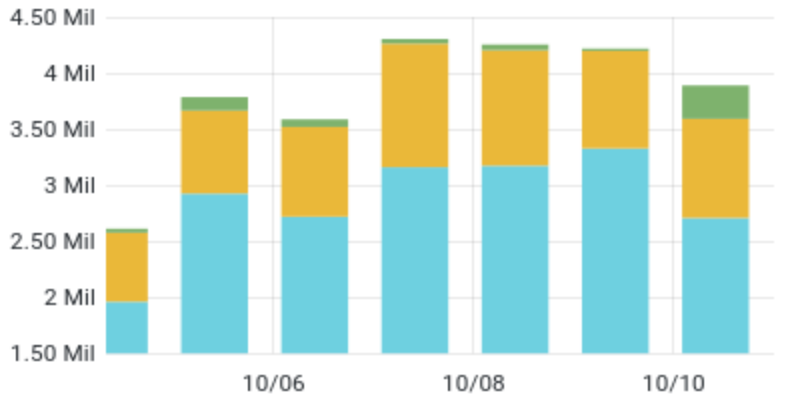


IAM access tokens response rate & time



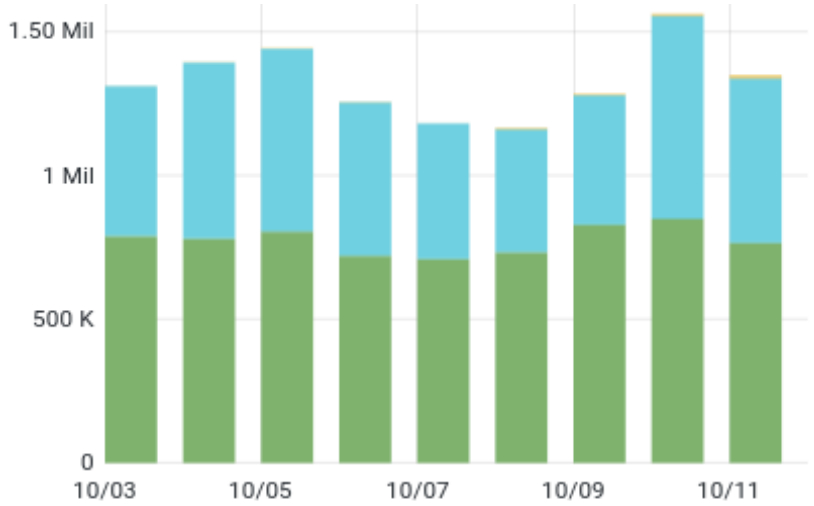
DC21 – daily FTS transfers & ATLAS stageout

Daily FTS during DC21 by experiment



	max	avg	current	total
atlas	3.33 Mil	2.85 Mil	2.71 Mil	20.0 Mil
cms	1.10 Mil	863 K	884 K	6.04 Mil
lhcb	297 K	90.2 K	297 K	632 K

Daily ATLAS jobs successfull stageout during DC21



	avg	total
Production Upload	700 K	7.00 Mil
Analysis Upload	494 K	4.94 Mil
CLI Upload	2.69 K	26.9 K