
CERN

FTS and DMC News and Plans

Andrea Manzi
on behalf of the FTS team

Outline

- 2018 Review
  - FTS numbers
  - Software releases
  - Achievements
- FTS and DMC team
- FTS and DMC plans for 2019

FTS numbers

- 7 WLCG instances deployed
  - BNL, CERN (3), FNAL, RAL, MIT
- 10 non-WLCG instances
  - CERN (DAQ, Public), RAL, KEK(2), Imperial, PIC, PNNL, MWT2, CESNET (WebFTS + RCAuth prototype)
- ~20 Virtual Organizations
  - ATLAS, CMS, LHCb, Mice, Xenon, Snoplus, AMS, NA62, Compass, ILC, Magic, Belle, Gridpp, Dune, LZ, Solidexperiment.org, SKA, Ligo, Icecube, Elixir, Opera
- 830 PB and 1.1 Billion files transferred
  - (numbers only from FTS instances centrally monitored)
  - ~60% by CERN FTSs (~30% is via IPV6)

FTS/DMC releases 2018 overview

- FTS 3.8 (Oct) + 4 patch releases
  - FTS nagios probes for C7
- Gfal2 2.16 (Sept) + 6 patch releases
  - Gfal2 bindings for python3 released to EPEL7 (Nov)
    - Needed packaging boost-python3 in EPEL7
- Davix 0.7 (Oct) + 3 patch releases
- Srm-ifce 1.24.4 (July)
- CGSI-GSOAP 1.3.11 (June)

2018 Achievements

- **EOS-CTA integration**
  - New Tape Solution @CERN
    - SRM-less
  - Staging via Xrootd implemented in gfal2-xrootd plugin
  - Validated by ATLAS

- **Xrootd and HTTP TPC enhancements**
  - Support for Bearer tokens (Macaroons/Scitokens)
    - Many contributions from B. Bockelman integrated and released (Thanks!)
  - Support for X509 Delegation for Xrootd TPC

2018 Achievements[2]

- **New FTS long term monitoring Dashboard**
  - New ES cluster to store aggregated data and make them available for 5 years via a dedicated dashboard, including staging metrics

- **Scalability improvements (ongoing)**
  - Improve the DB schema and indexes
  - Study DB partitioning
  - See Presentation from Eddie

2018 Achievements[3]

- **Automatic Session Reuse**
  - FTS Server automatic enables session reuse based on the Job parameters (number of files, size of the files, etc)
  - N.B. Issues discovered in session reuse implementation (fixed in 3.8.3). Disabled in production for now
- **Cloud Support**
  - Support for Google Cloud implemented in davix/gfal2
    - ATLAS Data Ocean project
  - Support for S3 Multipart upload in davix
- **Participation to the XDC EU project**
  - OpenID connect integration
  - Support for Storage QoS started

XRootD / HTTP support

- Many enhancements and bug fixes mainly driven by DOMA TPC and CTA
  - https://twiki.cern.ch/twiki/bin/view/LCG/ThirdPartyCopy
- HTTP
  - Bearer Tokens Support
  - HTTP 3pc mode **(Push vs Pull)** selection via config or query parameters (Integration in Rucio?)
- XRootD
  - X509 delegation for TPC
  - Clean destination on transfer failure
  - Many fixes on checksum support
  - Bringonline for CTA

Grafana Dashboards

- https://monit-grafana.cern.ch/dashboard/db/fts-servers-dashboard?orgId=25
  - last 30 days of FTS data
- https://monit-grafana.cern.ch/d/000000913/fts-servers-dashboard-yearly?orgId=25
  - Last 5 years (starting from June 2018)
- Reading data as JSON (using Grafana API Token)
  - https://monitdocs.web.cern.ch/monitdocs/access/monit_grafana.html

EU Project XDC

- 2 years software development project started in Feb ‘18
- http://www.extreme-datacloud.eu/
  - *‘Developing scalable technologies for federating storage resources and managing data in highly distributed computing environments*’
- Funded FTS activities
  - Integration with OIDC (OpenID Connect)
  - CDMI protocol integration to support QoS transitions

XDC: OIDC in FTS [1]

- FTS Auth/Authz historically done only with X509 proxy certificates and VOMS groups/roles
- 2 types of OIDC integrations implemented
  - Directly accepting access tokens from users via CLI/REST API
  - Redirect WebFTS users to a provider in order to acquire a token and using it via the FTS REST API
- Tokens are used both to authenticate to FTS and to the storages
  - dCache and StoRM are supporting OIDC

XDC: OIDC in FTS[2]

- FTS-REST component has been modified in order to accept an access token and refresh it when needed
  - Access tokens are verified via introspect endpoint of the provider or via offline validation
  - A refresh token related to the access token is acquired and saved to the FTS DB
  - A daemon refreshes the access tokens that are about to expire through the provider token endpoint by using the refresh tokens (needed for transfers staying long on queue)
- FTS Server can now use access tokens for transfers
  - Access tokens are retrieved from the DB and set to gfal2 API as BEARER credentials

XDC: QoS in FTS

- Use FTS to steer Storage QoS (e.g. multi-replica, low latency, etc.)
- New QoS daemon prototyped to include the current bringonline daemon functionalities + implement QoS transitions via CDMI
- Extension of Gfal2 HTTP plugin to support CDMI implemented
- First “simple” use case covered, requesting and monitoring a QoS transition
  - Supported by dCache

FTS/DMC Plans for 2019

- 2 FTS major releases
  - 3.9.x in March/April
  - 3.10.x by the end of year (including XDC contributions)
- 2 gfal2 major releases
  - 2.17.x in March/April
  - 2.18.x by the end of the year (including XDC contributions)
- Main Activities (in order of priority)
  - Scalability/Scheduler improvements (see Eddie's slides)
    - From last year delayed due to missing effort
  - Tape Migration Monitoring for CTA and other CTA related tasks
  - FTS-REST migration to new framework
  - More enhancements on HTTP and XRootD support
  - Complete XDC tasks

FTS and DMC Team

Andrea
(80%)

Project lead

Gfal2

Service Manager @CERN

Maria)
(30%)

FTS Rest and Monitoring

Eddie
(100%)

FTS Server, Bring-online

Aris
(100%)*

XDC

*till April

New Fellow to join the team in Q2/Q3

Oliver
(10%)

XDC

Georgios
(10%)

Davix

Migrations to Tape [1]

- FTS, in case of a Tape Storage, is now unaware of file migrations to tape:
  - Transfers to a tape storage are considered completed when the file is on the disk buffer
  - Clients need an extra step in order to validate that the file is on Tape (i.e. checking the “m” bit on Castor)
- Plan to implement migrations to tape monitoring
  - Transfers in Final state only when files are stored on Tape
- This will also help implementing other mechanisms, like back-pressure on number of files/size of migrated data
  - FTS will stop scheduling new transfers if files under migration are over a certain threshold

Migrations to Tape [2]

- We plan to extend the QoS daemon (Disk->Tape is a QoS Transaction), under implementation in the context of XDC, to implement a first version of the Migration to Tape monitoring this year
  - Targeting first CTA where disk buffer size is limited by design
  - Extension to SRM to be planned afterwards
- We plan to involve the experiments in the design phase
  - Many details to discuss
  - N.B. CMS has already expressed high interested in this topic as a requirement to move to Rucio

Staging + transfer with different protocols

- When running Staging + transfer jobs, possible protocol mismatch between the source and the destination
  - Staging with XrootD url and transfer to a Srm/Gridftp url destination or viceversa
- Plan to automatically adapt the source protocol to match the destination protocol when performing the transfer
- Need to (re)introduce the concept of StorageGroup to know the endpoints associated to a storage
  - i.e. The Xrootd endpoint can be different from the Gridftp gateways endpoint

FTS-Rest migration to new web framework

- Pylons, the framework we built FTS-REST upon, is quite obsolete (in maintenance-only mode now) and only available in Python2
- We have planned to move the implementation to a new Web Framework: Pyramid
  - https://trypyramid.com/
  - The project developed to replace Pylons
- Quite some effort to spend as LOT of things changed
  - First prototype version with core functionality to be implemented this year
  - Python 3 since the beginning

XRootD / HTTP support in gfal2/davix

- Enhancements/fixes planned for HTTP
  - TPC transfer cancel in Davix/Gfal2
  - Add support for Checksum algorithm hinting in HTTP COPY
  - Support for different tokens for source/dest in gfal-copy
  - IPV4/6 monitoring (needs changes Storage side)
- Enhancements/fixes planned for XrootD
  - Fix transfer cancel in gfal2
  - IPV4/6 monitoring
  - Transfer canceling based on performance markers (as in gridftp)
  - Integration of Xrootd 5 new methods for staging

XDC plans - OIDC

- Complete extension to OIDC tokens of all FTS-REST operations now requiring X509
  - E.g. User banning
- Integration of Token translation service (to be confirmed)
  - Present a token – get an X509 certificate
  - Needed for EOS in XDC, but of course for all the other storages which do not support OIDC yet
    - Needed also to use other protocols than HTTP
- Follow closely the output of the WLCG Authz WG

XDC plans - QoS

- Full integration of QoS logic
  - Transfer/Transition logic
    - Use existing multi-hop logic to serialise transfer-then-QoS-transition
  - Completer QoS daemon
- Validate integration of all QoS methods in gfal2
  - Still need a working test endpoint
- Definition of FTS QoS interface for Rucio/Orchestrator

Other Tasks

- Best replica selection algorithm: reshuffle the order chosen at submission time by failing transfers more than 1 hour on the queue ?
- gfal2-util migration to python3 (prototype available)
- WebMon upgrade to latest Django version and Python3
- Network Topology/SDN
  - We are part of the Hepix NVF WG, but no activity are planned this year
  - Noted Project together with ATLAS

Other Tasks[1]

- Davix migration to libcurl
  - Move from libneon to libcurl so as to drop the hard dependency on OpenSSL, which is going away on MacOS. Necessary to continue offering davix on MacOS, requested by EP-SFT

- CentOS8 support?
  - The New OS will appear during the year
  - SL6 EOL in 2020

Questions?