



FTS News and Plans

Andrea Manzi
on behalf of the FTS team

Outline

- FTS team reorganization
- New Website
- FTS 3.7.8 Release
 - Monitoring via Grafana
- New FTS3 Public instance
- FTS plans for 2018
 - Scheduler enhancements
 - Scalability improvements
 - XRootD/HTTP support enhancements
 - FTS in XDC project

New FTS Website and Logo



- New Website just launched:
 - <http://fts.cern.ch/>
- More dynamic, modern and also compatible with all devices and browsers.
- It provides information regarding all FTS components, the documentation related to the project, steering meetings, release information tagged automatically
- New techblog and use case discovery to guide users selecting the best FTS components for their needs.

New FTS Website



[HOME](#)

[FEATURES](#)

[NEWS](#)

[DOCUMENTATION](#)

[USERS](#)

[CONTACT US](#)

File Transfer Service at CERN

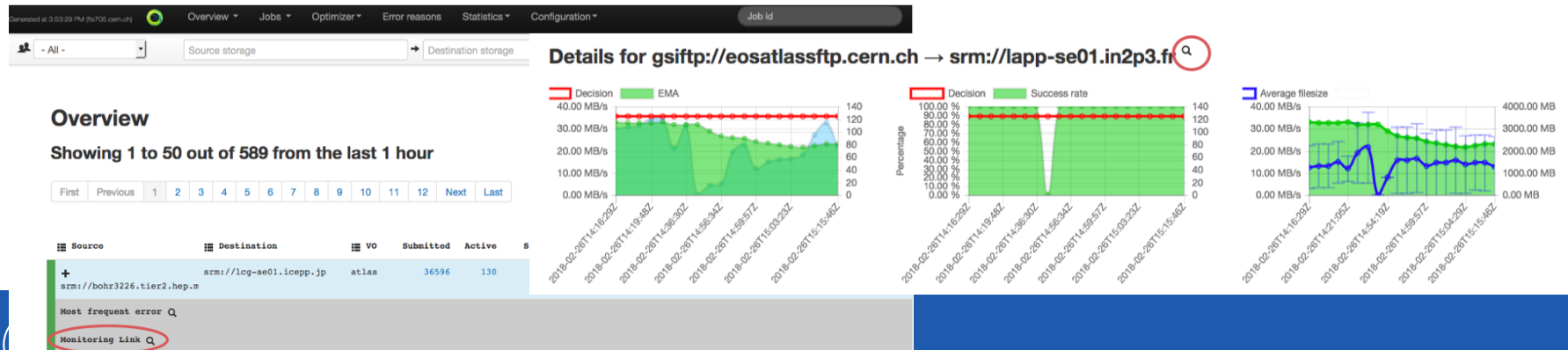
Open source software to transfer data reliably and at large scale between storage systems

[▶ GET STARTED](#)



FTS 3.7.8 release

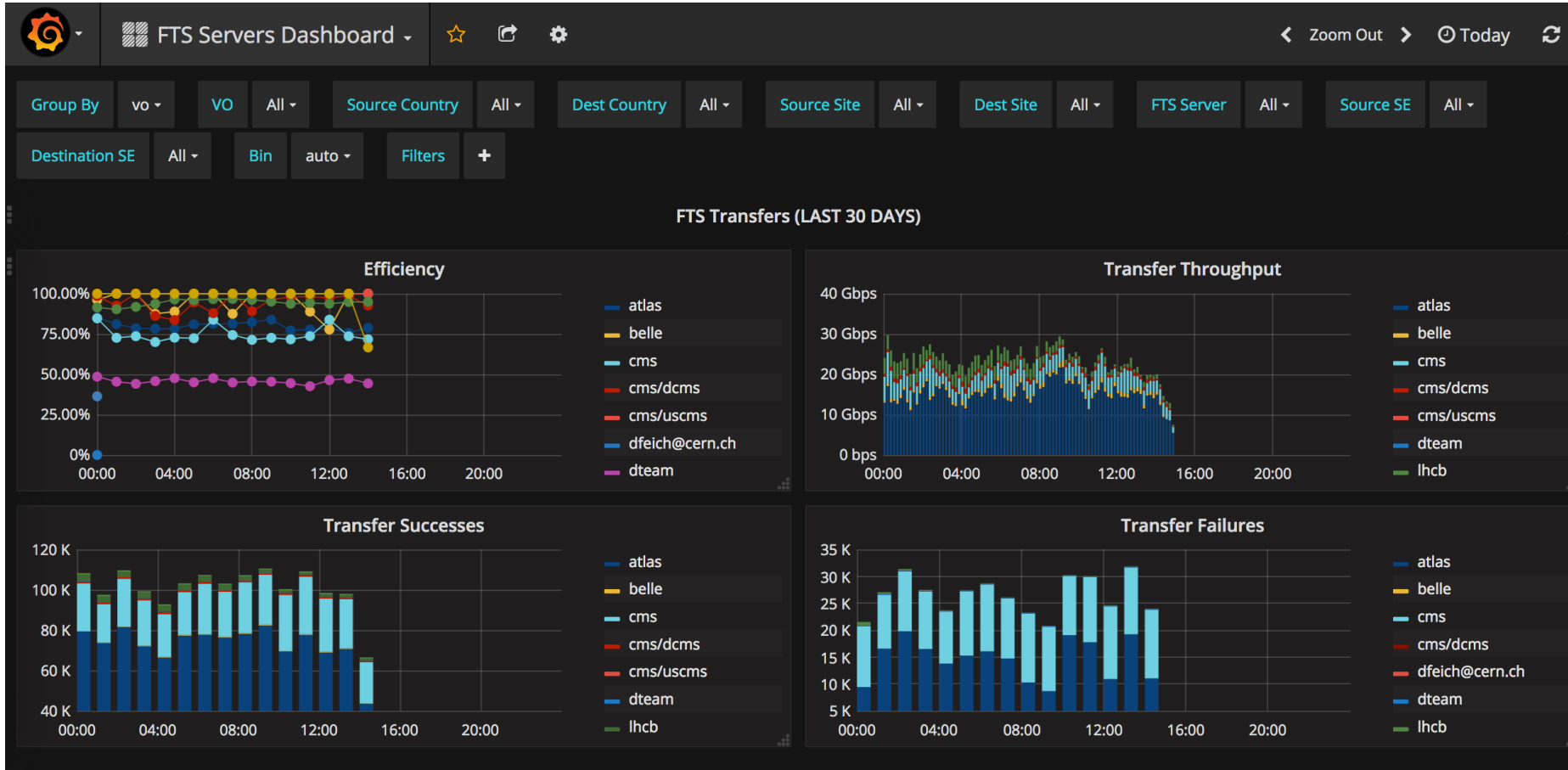
- FTS 3.7.8 released in February and installed last week at CERN prod
 - <http://cern.ch/go/Kw8g>
 - No problem reported, other sites can also upgrade.
- Small bug fixes + some enhancements
 - Ability to disable DB tables backup process (backups not used and degrading DB performances)
 - WebMon integration with the new Grafana Monitoring



New Grafana Dashboard

- <https://monit-grafana.cern.ch/dashboard/db/fts-servers-dashboard?orgId=25>
- Ability to query the last 30 days of FTS data
 - Transfer statistics
 - Data volumes
 - Queue times
 - Optimizer info
 - Staging statistics
- Reading data as JSON (using Grafana API Token)
 - https://monitdocs.web.cern.ch/monitdocs/access/monit_grafana.html

New Grafana Dashboard



FTS3 Public instance @CERN

- New **Production** cluster available since February
 - <https://fts3-public.cern.ch/fts3/ftsmon>
- The goal is to provide a more stable environment for non-LHC VOs already using the FTS Pilot instance
 - **ILC, AMS, Compass, NA62, Auger**
- *WebFTS* is also using this instance now with its VOs:
 - **belle, biomed, dream, icecube, ops, vo.access.egi.eu**
- We already contacted the VOs using the FTS Pilot to switch to the new instance for production transfers
- The Pilot instance will be dedicated to test the FTS and gfal2 RC releases

FTS Plans for 2018

- 4 areas of work
 - **Scheduler**
 - **Scalability**
 - **XRootD/HTTP support**
 - **XDC project**
- We plan to have 2 FTS **major** releases (with of course bug fix releases in between):
 - 3.8.0 Q2
 - 3.9.0 Q4
- gfal2 releases are independent from FTS
 - 2.15.x is in RC
 - 2.16.x in Q3

Scheduler Enhancements

- **Automatic gridftp session reuse (3.8.0)**
 - REST enables it automatically if the job average file size is less than a configured threshold
 - To minimize authentication overheads when moving small files
 - Already released in 3.6 but disabled cause of issues found when testing in the Pilot.
- Implement proper **links reshuffling (3.8.0)**
 - Links starvation is still possible in some cases
- Study and eventually implement the **activity shares per destination**, as requested by ATLAS (3.9.0)

Scheduler Enhancements [2]

- Validation of the algorithms for **Source Replica Selection**
 - Already implemented but never intensively tested (and documented)
 - **orderly(default)** : no rank, try replicas in the order provided
 - **queue or auto**: rank based on the number of pending files in the queue.
 - **throughput**: rank based on the total throughput rate between a source destination pair in the last 1 hour
 - **file-throughput**: same as above but ranking for file throughput
 - **pending-data**: rank based on the total pending data in the queue between a source destination pair.
 - **waiting-time**: rank based on the waiting time for the incoming job in the queue
 - **waiting-time-with-error**: using the failure rate info, calculate the amount of data that will be resent. Rank based on the waiting time plus the time for resending failed data
 - **duration**: rank based on the waiting time with error plus the time required to transfer the file

Scalability Improvements

- **DB Scalability improvements**
 - FTS overall performances degrade when the number of queued transfers on the DB $> \sim 1\text{M}$
 - Study and eventually implement MySQL Partitioning for t_job and t_file tables (3.9.0)
- **Maximize parallel transfers** per FTS node
 - Try to reduce as much as possible the memory footprint for fts-url-copy
- **fts3-atlas** dedicated instance at CERN
 - ATLAS -> 70% of the total files transferred
 - Ready! <https://fts3-atlas.cern.ch:8449>

XRootD / HTTP support

- ATLAS and ILC are testing **XRootD 3rd party copy** (3pc) transfers via FTS
 - Plan to follow up any issue reported and implement enhancements
 - i.e. stuck transfers are not cancelled at the moment (as we do in gridftp)
- XRootD 3pc supported by DPM, EOS, Castor (available for dCache via XRootD Proxy)
 - DPM does not support query for checksums via XRootD
- Integration of the **EOS-CTA** project (3.9.0)
 - Staging via XRootD

XRootD / HTTP support [2]

- New gfal2 2.15 (in the Pilot) introduces changes on the **HTTP 3pc process**
 - No need to explicitly use `davs+3rd://` notation
 - *Pull mode => Push mode => Streamed mode*
 - <http://fts-docs-devel.web.cern.ch/fts-docs-devel/docs/3rdpartycopy.html> (preview docs)
- 3pc supported by dCache, DPM and XRootD (via XrdHTTP which will be used also by EOS)
- Clouds:
 - HTTP 3pc transfers to S3 implemented
 - S3 gCloud 3pc via signed URLs under implementation in Davix

Token Authentication

- Brian Bockelman has implemented a first support for Token Authentication (**SciTokens** and **Macaroons**) in gfal2 and FTS (thanks Brian!)
- This will allow HTTP transfers without x509
- DPM (Dav 0.19.0) and dCache (since 3.2.x) already support Macaroons. XRootD (XrdHTTP) support for SciTokens is also there
- WLCG HTTP Transfer group
 - <https://groups.google.com/forum/?hl=en#!forum/wlcg-http-transfer>
- WLCG HTTP Ecosystem Demonstrator talks at upcoming WLCG workshop and CHEP.

FTS in XDC Project



- 2 years EU software development project started in Feb
 - <http://www.extreme-datacloud.eu>
- “Develop scalable technologies for federating storage resources and managing data in highly distributed computing environments”
- CERN IT is participating with FTS, EOS and Dynafed
- Main activities for FTS in 2018
 - AAI
 - **OpenID Connect** authentication for FTS
 - Exploitation of network information (**Perfsonar**)
 - Storage **QoS** exploitation
 - By integrating CDMI interface

FTS discussion at ATLAS Site Jamboree

- Presentation from **Ilija**
 - <https://indico.cern.ch/event/692124/contributions/2899983/attachments/1611875/2559846/go>
 - Proposal to give scheduling priority to links based on **Throughput**
 - We would like to work together to evaluate the performances at other FTS instances and links:
 - possible deployment change at the BNL FTS
 - Automatic Session reuse and better DB scalability may already improve the situation
- FTS Configuration ??
 - Site managers do not know the different conf options available
 - We should aim at **Zero** conf as much as possible.

FTS discussion at ATLAS Site Jamboree

- Knowledge of the network usage on a link **external** to FTS transfers
 - To check with the monitoring team if it's available
- Usage of a **single** FTS instance (with a fallback)
 - Not feasible at the moment, scalability issues to be solved first
 - Each FTS instance “aware” of the others? We need FTS4☺
- Staging file **priority** ?
 - Need to pass this information via SRM, not available AFAIK
 - To evaluate in EOS/CTA (staging via xrootd)
- Ability to group SE endpoints, **FTS topology awareness**.
 - Grouping of SE has been removed since FTS 3.7.0 as it was not used for scheduling purposes.
 - With the increase protocol zoo (each protocol+endpoint is a different SE in FTS), it would be useful to implement the scheduler topology aware. To plan for next year

Questions?

