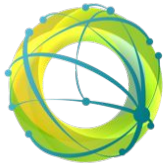
FTS & XRootD Workshop

CNRS Rhône Auvergne Delegation, Lyon – 14-18th September 2026



FTS4: State of (Testing) Affairs

Mihai Patrascoiu
on behalf of the FTS team

Content - **Steven Murray**
Slides - Mihai Patrascoiu

FTS4 goals for DC'27

- **Rucio** and **DIRAC** decide **what** to transfer
- **FTS** decides **when** to transfer

Our targets (non-exhaustive):

- Correct scheduling (deterministic, no starvation, priorities respected)
- Flow Controller (Like the Optimizer, but better)
- Double concurrency & queues of DC'24
  - 20,000 → **40,000** concurrent file transfers
  - 10M → **20M** queued requests
- Basic back-pressure
- Tape, Tokens, Tape & Tokens

> *Full details in Maria's presentation at the OTF session (Tue, 15 Sep 2026, 16:00)*

Removing the FTS4 deadlocks

The cause

- Bulk / multi-row updates in one transaction
- Denormalized counters instead of `SELECT COUNT(*)`
  - Fantastic read performance !
  - …but defeats MVCC decoupling of readers from writers

The fix

- Strict row-locking order, per table
- Ensure uniform bulk updates:
  `start transaction → identify rows → lock in predefined order → update → commit`
- Counters removed, except `SUBMITTED` queue counts
  - `ACTIVE` → `COUNT(*)`,
    all rows are in DB memory
  - `FINISHED`/`FAILED`/`CANCELED` → `COUNT(*)`
    for now, but not scalable. Maybe time series (TimescaleDB?) before DC'27

Current FTS4 installations

Each instance: **4 physical machines** (16–24 physical cores, 187–376 GiB RAM)

+ main and replica database

Instance	Role	Main database	Replica database
`fts4-atlas.cern.ch`	**Production**	130 GiB RAM	130 GiB RAM
`fts4-cms.cern.ch`	**Production**	42 GiB RAM	42 GiB RAM
`fts4-daq.cern.ch`	Test	18 GiB RAM	18 GiB RAM
`fts4-pilot.cern.ch`	Test	32 GiB RAM	32 GiB RAM

Internal tests with the CERN EOS Physics Service

Why run internal tests at CERN?

- We rely too heavily on ATLAS & CMS testing *(thanks for all the help!)*
- We would like to be more autonomous
- Strictly-controlled high-performance tests are easier when we control the storage

What we did

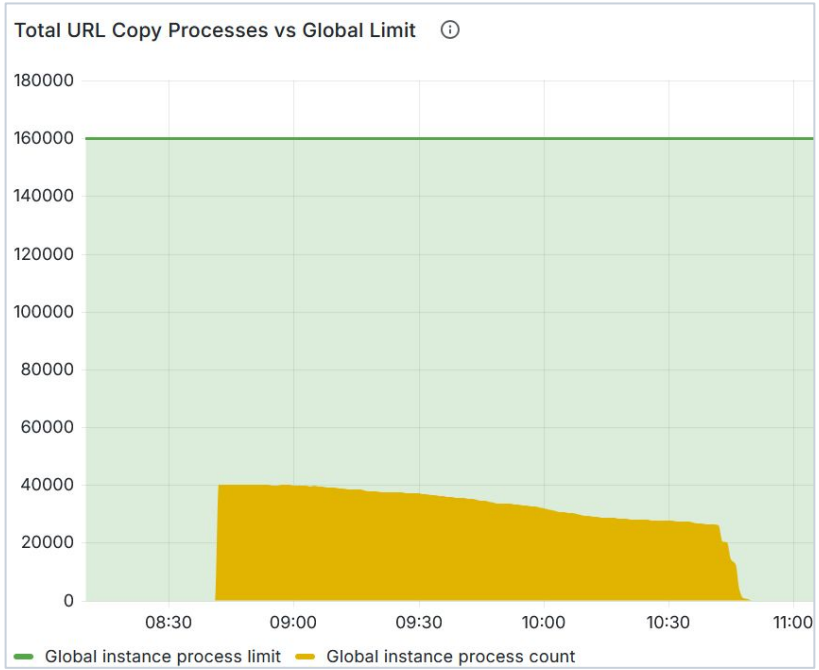
- Created **1.2 PB** on `eosp2.cern.ch`: 120,000 × 10 GB
  (thank you ALICE & EOS team for lending the storage!)
- Each test: copy the 1.2 PB to **`eospilot.cern.ch`**
  - **40,000**: scale test, all 4 `fts4-pilot` machines (errors expected)
  - **10,000**: clean transfers, 1 `fts4-pilot` machine (no errors expected)

What we achieved

- Improved FTS4, EOS and XRootD
- **660 Hz** ramp-up rate
- **99.563%** success rate with 10,000 concurrent transfers
- Demonstrated FTS4 can handle **40,000** concurrent transfers for DC'27

Internal test results 1: Scale test

40,000 concurrent transfers with **≥ 660 Hz** ramp-up on 24.07.2026



Lots of expected transfer failures:

- Small set of disks on `eospilot.cern.ch`
- Interested in the ramp-up rate: **660 Hz**
- Transfer failures tested FTS4 error handling

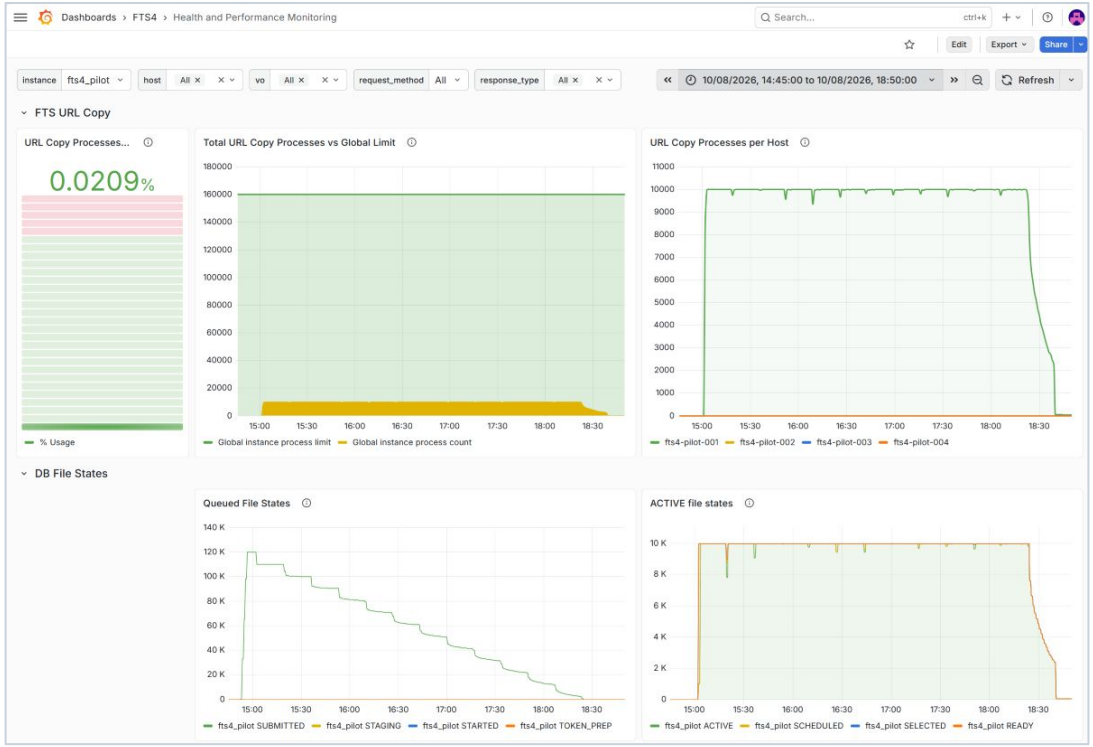
Internal test results 2: Clean transfers

- Stable performance
- 10,000 concurrent transfers

Final score: **99.563%**

```
$ ./psql_pilot_file_states.sh
Tue Aug 11 06:50:58 AM CEST 2026

 file_state | count
------------+--------
 FINISHED   | 119476
 FAILED     |    524
(2 rows)
```



Conclusion

- Database deadlocks addressed with a uniform `UPDATE` model
- FTS4 entered DC’27 testing phase

  → Large-scale tests show robustness of the system

  → Demonstrated target of sustained 10k transfers on one machine
  (remains to test a clean 40k transfers across four machines)

Our thanks to experiments and CERN Storage team for their help !

Thank you!

For everything FTS4

- FTS4: Optimising Performance & Efficiency – *Louis Regnier*
- FTS4: Scheduler and FlowController – *Nicola Pace*
- FTS4: Tape Daemon Re-Engineering for LHC Run 4 – *Joao Pedro Lopes*
- FTS4: Introducing TimescaleDB – *Mihai Patrascoiu*
- FTS status and plans (OTF Session) – *Maria Arsuaga Rios*