FTS Community talk: LHCb

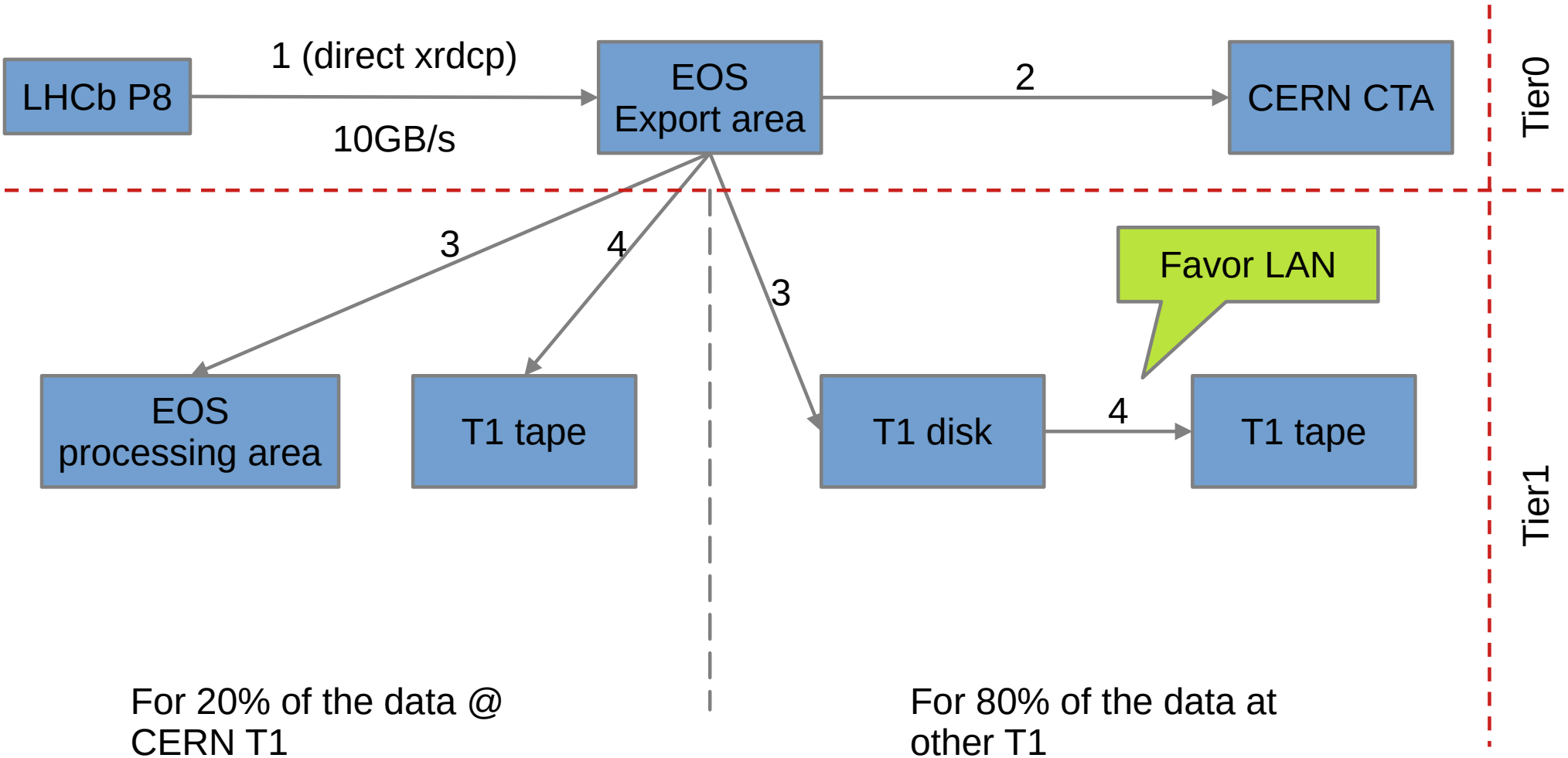
XrootD/FTS workshop 2024

Christophe Haen

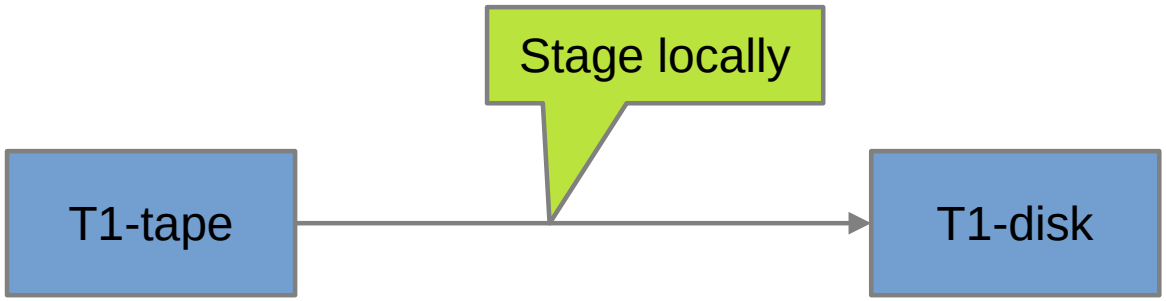
FTS at the core of LHCb DMS

- All Third Party Copy transfers of LHCb go through FTS
  - Transfers dominated by Real Data flows
  - Jobs upload data to their final destination (i.e. no rebalancing campaign)
- Everything orchestrated by DIRAC
- As usual, follow the KISS principle
  - Little to no use of “fancy features”
- Reliability, performance & stability are our main requirements

Real Data distribution (most common workflow)



End of year reprocessing



~ 2024 figures

- Again, favor LAN transfers
- Valid for T0 too

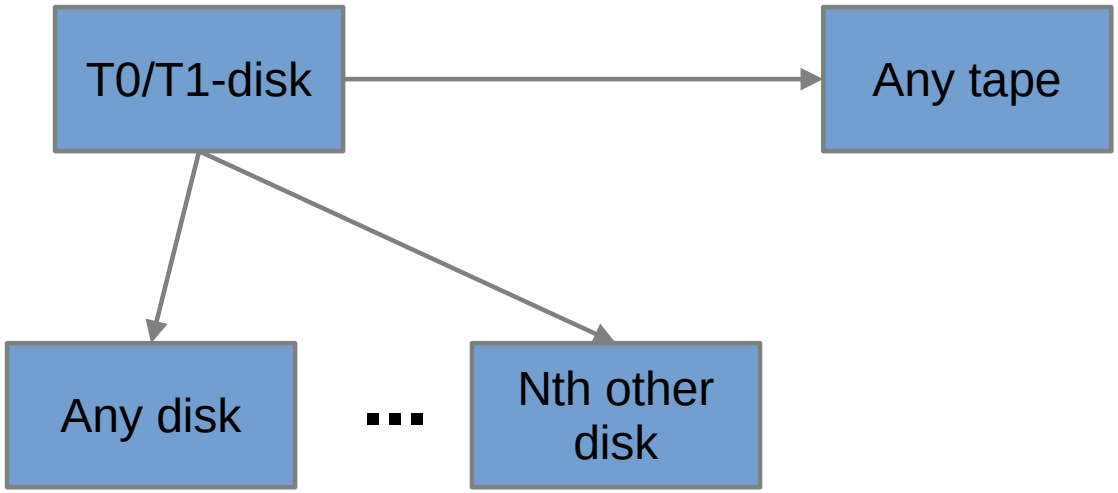
Write tests:
CERN disk → T1 disk → T1 tape

Site	Export Speed (GB/s)
CERN	11
CNAF	1.98
GRIDKA	2.49
IN2P3	1.48
NCBJ	0.74
PIC	0.49
RAL	2.48
SARA	1.34
Total T1	11

Read tests
T1 tape → T1 disk

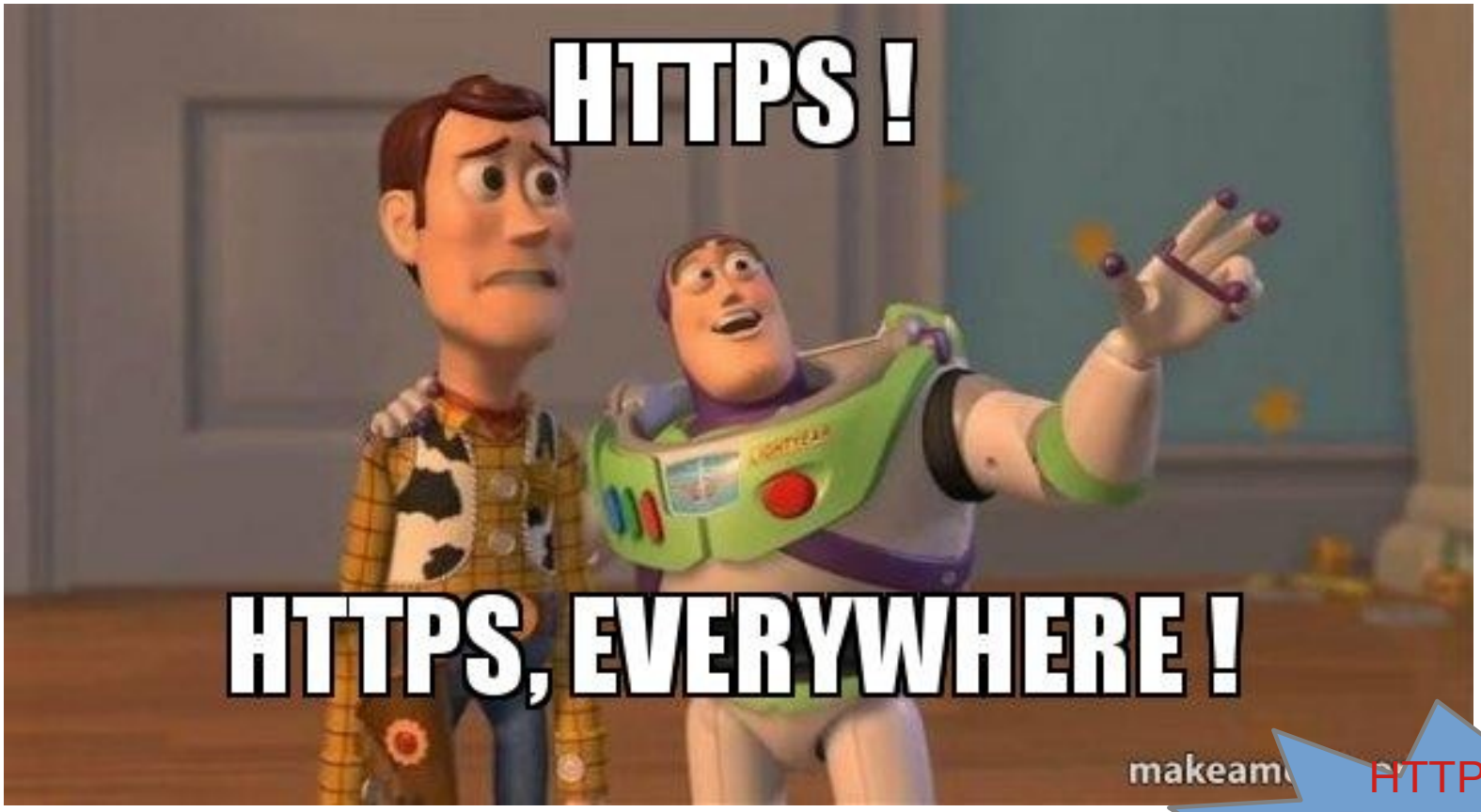
Site	Stage Speed (GB/s)
CERN	2.92
CNAF	1.56
GRIDKA	1.31
IN2P3	1.17
NCBJ	0.69
PIC	0.31
RAL	1.6
SARA	0.62
Total T1	7.27

Final data/MC distribution



- Jobs upload their output on 1 disk
- Replicate from that disk with FTS
  - 1 archive to any tape
  - N replicas to any disk (T0,T1,T2)
  - Initially, N=2 (can be reduced depending on popularity)

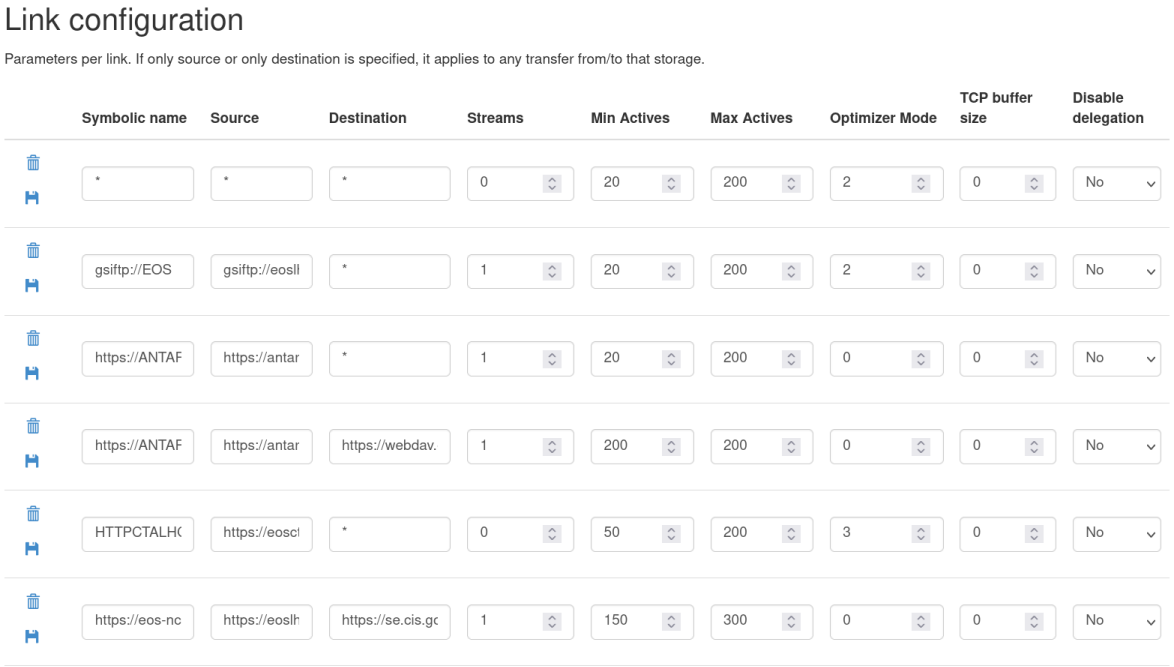
Protocols



HTTPs
Tape Rest
API
everywhere

Configuration

- Default except for T0 ↔️ T1 links (empirical outcome of Data Challenge)
- Useful for protecting storages when doing some tests



Link configuration

Parameters per link. If only source or only destination is specified, it applies to any transfer from/to that storage.

	Symbolic name	Source	Destination	Streams	Min Actives	Max Actives	Optimizer Mode	TCP buffer size	Disable delegation
	*	*	*	0	20	200	2	0	No
	gsiftp://EOS	gsiftp://eosll	*	1	20	200	2	0	No
	https://ANTAF	https://antar	*	1	20	200	0	0	No
	https://ANTAF	https://antar	https://webdav.	1	200	200	0	0	No
	HTTPCTALH(	https://eosc!	*	0	50	200	3	0	No
	https://eos-nc	https://eoslh	https://se.cis.gc	1	150	300	0	0	No

~~Misslanous,misclanous,misclenaous~~, OTHER !!

- Single FTS instance at CERN
  - Proved to be very reliable and sufficient
  - No configuration discrepancy
  - If lasting downtime: just recreate another cluster
- Plan to use FTS for pinning files on tape for jobs (i.e. tape → tape cache)
- Activity: currently used only for DC, but interested in adding more
- Priority: not now, but if needed, we can
- Tape metadata: interested in adding them (I just need a few spare cycles)

Tokens ?

- Ran DC24 with it for about half the size
- We all know the outcome, no need to further discuss it
- Support ~ is in DIRAC
- File specific tokens

Shopping list

Easier way to query/set the configuration ?

- Visual and edition
- Programatically and web
- YAML ? (happy to help)

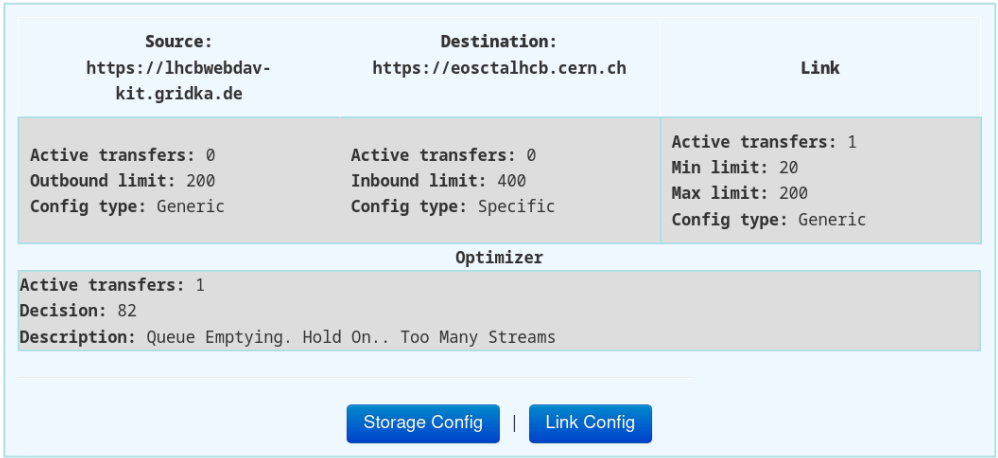
Archiving limit ?

Per file search ?

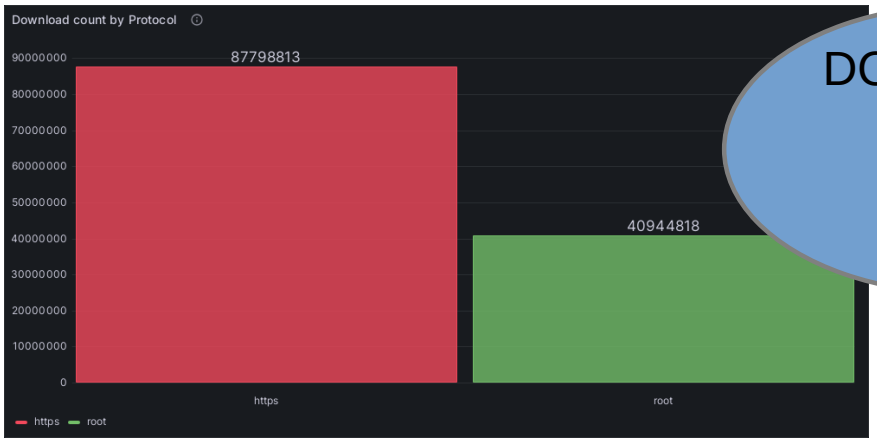
Multihop intermediate files removal ?

Feedback

- Very happy
  - At least, as happy as one can be doing DMS….
- Very stable and reliable
- The “explain” is very useful
- Grafana dashboard top notch
- **FTS is absolutely paramount to LHCb computing**
  - This is a subliminal message to the IT dpt management…

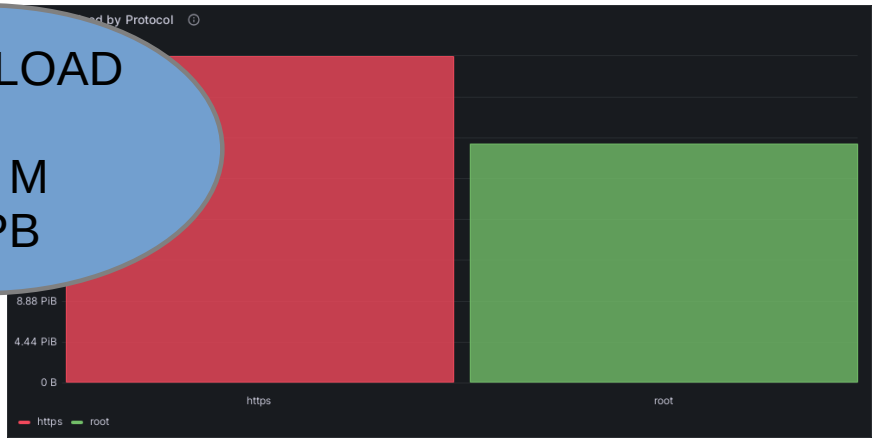


FTS is just the tip of the iceberg



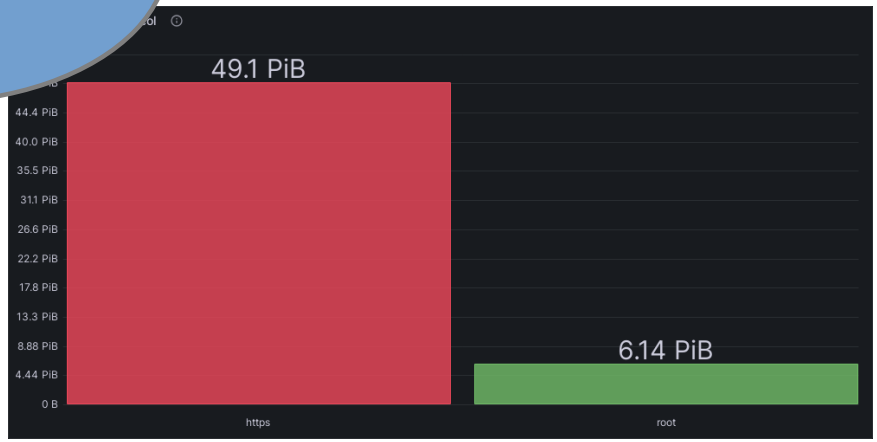
DOWNLOAD

130 M
51PB

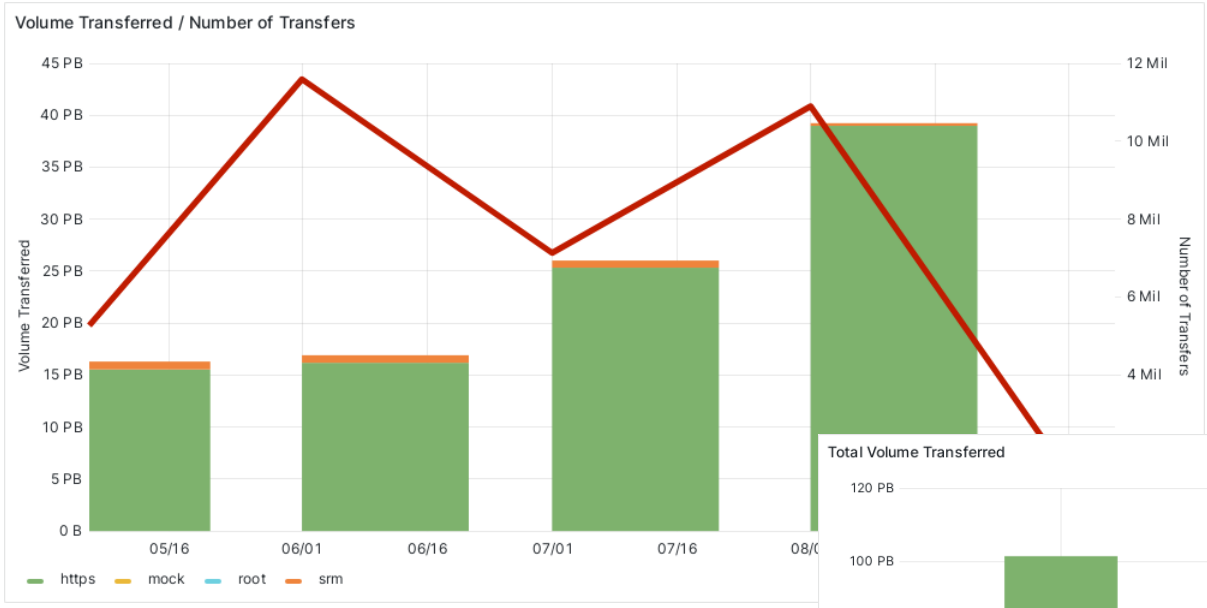


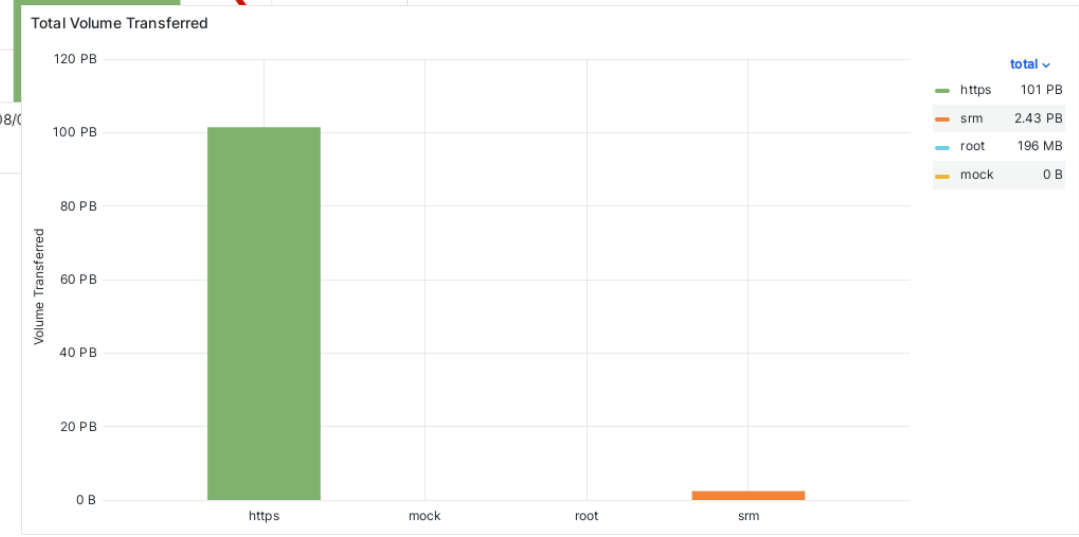
UPLOAD

140 M
56PB



FTS is just the tip of the iceberg





32 M
100PB

FTS is just the tip of the iceberg

GFAL2 for LHCb

- Absolutely critical for LHCb/DIRAC
- Every single file operation performed via Gfal2 python binding
  - Removal
  - Stats
  - Checksums
  - Direct download/upload
  - TPC
  - Few directory listing
- True for all protocols DIRAC supports (SRM,GSIFTP, HTTPS, XROOT)

What really matters for something like Gfal2

- All the operations I mentioned before
- Protocols: https and xroot would suffice
- Shared between DIRAC/FTS/RUCIO
- Canonical way of reporting a problem to a site
- Python friendly
  - Can you please build wheels ?
- fsspec ?

Conclusions

- FTS and gfal2 is what underpins all of the LHCb Data Management
- FTS has been operating flawlessly
  - Performance, stability, monitoring are our most desired features
- A common library like gfal2 is paramount
  - Happy to contribute
- Side note: xrootd streaming read is MASSIVELY used in LHCb
  - We nuked sites
  - Worth taking into account for future DC...

Finally…

Big thank you to the FTS team (present and past)