WLCG Data Challenge 2027

15th September 2026

Joint FTS and XRootD workshop, Lyon
Katy Ellis, Johannes Elmsheuser
With help from Alessandra Forti, Mario Lassnig, Christoph Wissing

WLCG
Worldwide LHC Computing Grid

Planning

AGENDA

1 Previous DC reminders

2 Progress since DC24

3 Goals for DC27

4 Planning DC27

5 FTS traffic

6 Assessing number of active transfers

Data Challenge series

RECAP

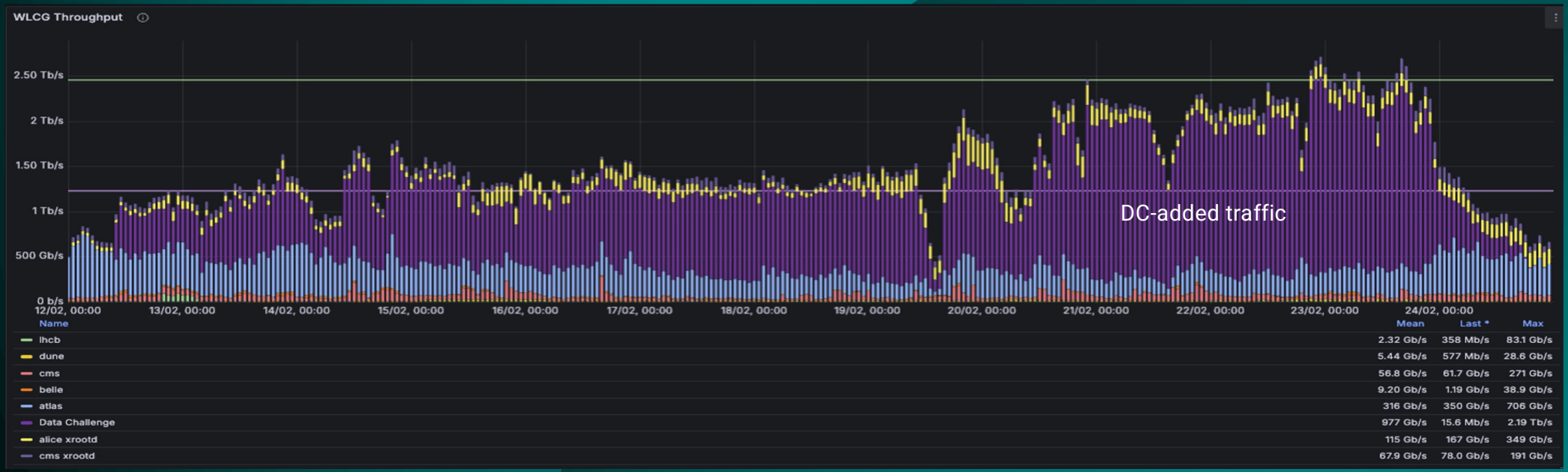
1	DC21	10% Run-4 target; learn how to do a data challenge
2	DC24	25% Run-4 scale, with minimal and flexible targets; first usage of tokens
3	DC27	50% Run-4 target (if feasible given current hardware); being discussed today
4	DC29?	Planned for the year before the start of data-taking; sites and software ready for Run-4 scale data movement

DC24: What happened?

Successful test, although hard work trying to sustain the flexible target in the second week.

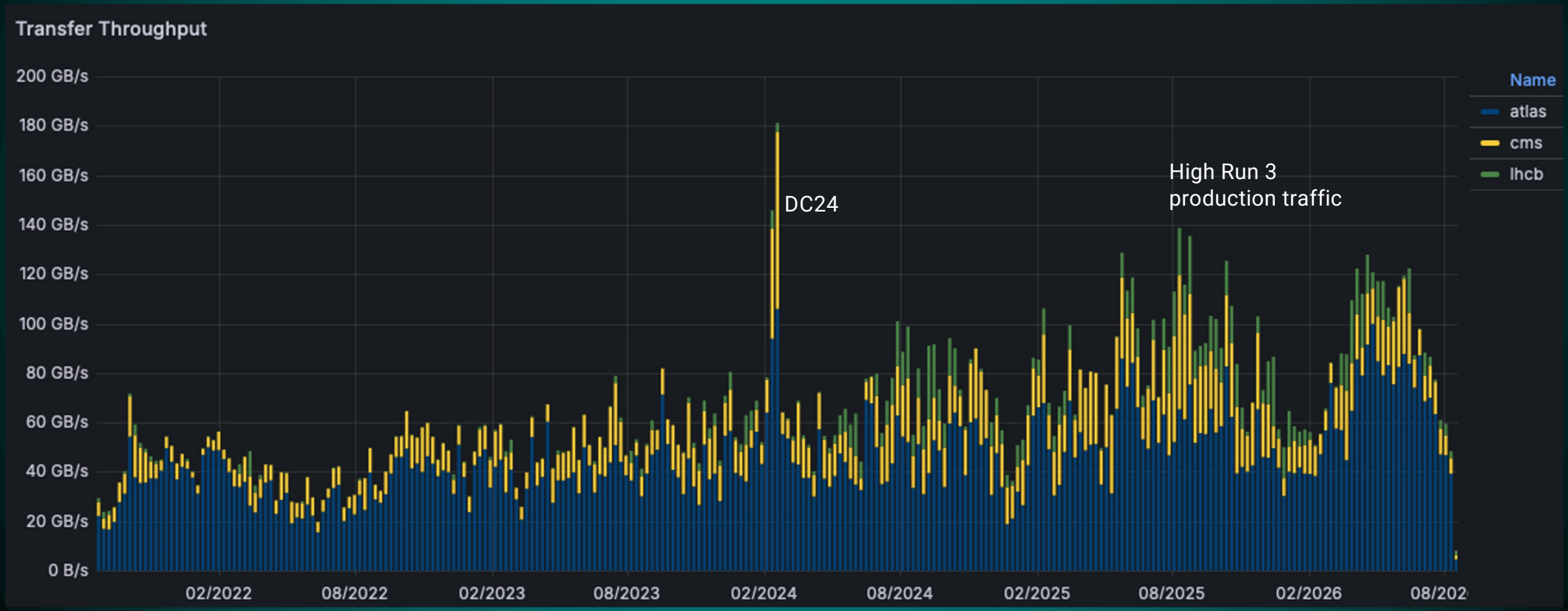
Several Tier 1 storage systems (disk only - ATLAS and CMS) under extreme pressure.

Highlighted bottlenecks in FTS (File Transfer Service) software → FTS4



Note of current FTS traffic

The LHC in Run 3 surpassed expectations - leading to high transfer rates



1 week binning

Looking forward to FTS4

- FTS4 is in test phase
  - ATLAS and CMS
    - https://fts4-atlas.cern.ch:8449/fts/webui/#/
    - https://fts4-cms.cern.ch:8449/fts/webui/#/
- Has been tested in Production by ATLAS and CMS
  - Some issues found! (Currently not in production)
- No longer a bottleneck..?

Improving our estimates for Run-4

Previous targets were defined in 2020 - we can do better now!

- CMS and ATLAS have been working hard to estimate better what rates will be needed
  - A bit lower than before?
  - We want to maintain the ambition of a challenge (but not ask the sites for more than we need)
- LHCb and ALICE already have a good handle on their Run-4 rates, as they should be similar to Run-3

- Although the flexible target may not be double what it was in DC24, we will still face huge challenges in DC27!
- As always, we do not encourage sites to invest in hardware or network capacity to ‘pass’ these tests, but it is a good opportunity to optimise and learn from what we have

Defining DC27

1 Target rates at 50% of Run-4

Taking into account the new and improved Run-4 estimates

2 Access only via tokens

Can every transfer be done without X509 certificates? Even tape?

3 Baseline and flexible targets?

DC24 defined a minimal model, and a flexible model which was ~double the minimal. No doubling of the flexible this time

4 Usage of tape

Reads and writes with tape should be tested. These storages are using much of the same network as disk.

5 Test profile

Fast ramp-up in first week to baseline target or above using disk transfers, T0/T1/T2. Schedule still to be decided

6 Other?

Possibility for some targets around IPv6, using SDNs, etc. Defined by experts in that area. Small file testing?

DC27 scheduling

When to hold the challenge / how long?

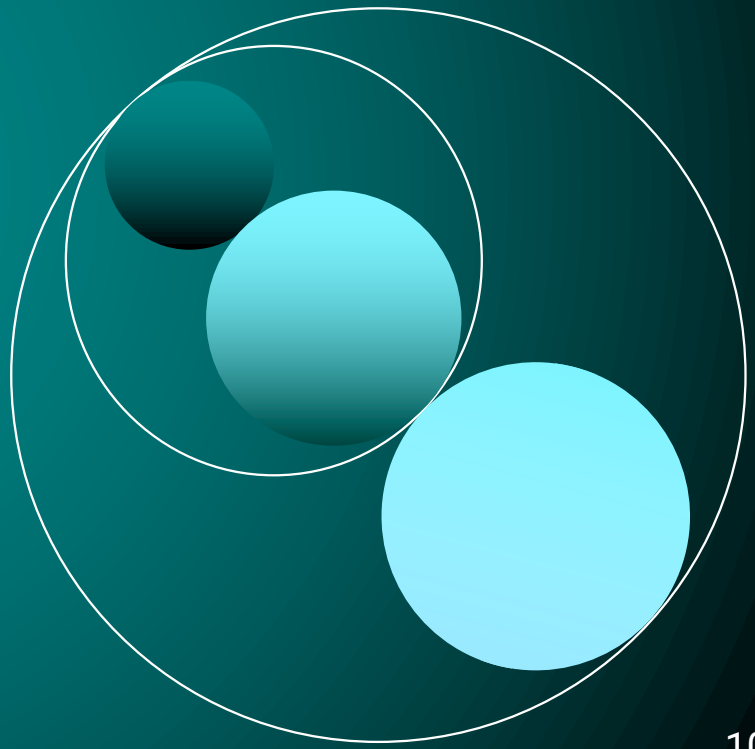
- DC27 is agreed for 22nd Feb - 5th March
- It will be a packed schedule!
- Details to be agreed during the WLCG Workshop (Bologna, Nov)

Tape questions

What, when, how much?

- Read tests
  - Easy to execute (Rucio) with disk as destination
  - Recall range of datasets
- Write tests
  - More clean-up for the sites
  - How to do a write test that is easy to clean up but is realistic in write pattern (e.g. tape families)
- Simultaneous read and write
  - Reports that some sites struggle

- **During the main DC period (simultaneous with disk transfers)?**
- What rates can sites handle even in theory?
  - As always, we do not want to encourage sites to invest in hardware too early



Mini-DCs

Continue to run pre-challenge tests between now and DC27

Mini-DCs are tests targeting specific parts of the system, with the intention of improving components

- Capacity tests
  - Example: improve throughput at sites in the UK
- Capability tests
  - Example: try new technology

+

Test FTS4, tokens at scale

Julien's tape emulation test

Confirmed participants?

FTS users

- Summing the cumulative rate at multi-VO sites could show higher demand than supply on the network and/or storage I/O
- Early involvement will help with rate calculations
  - Experiments have been entering their target rates per Tier 1 site into a big spreadsheet - thanks!

Challenge planning

STEP 1

Calculate rates

Talk from ATLAS on rate estimation at DOMA-General meeting (1st April)
Talk from CMS on same topic from Oct 2025

STEP 2

Add per-site rates together for each VO…can the sites and networks handle it?

Now!

STEP 3

Decide what to do if the summed rates are too high. How can we try to share the bandwidth between VOs?

As always, we do not ask sites or networks to invest in hardware in order to meet data challenge targets, unless it fits well with their existing plans

Then define the schedule…

Schedule and tools

The schedule will be discussed during the WLCG workshop (Bologna)

- Likely to start simple and build in complexity
- Tape read and write tests
- Simultaneous read/write tests
- Small file / metadata tests
  - FTS and IAM do not care about throughput volume - only the number of files
  - Including different file sizes is realistic
- Probably various network technology tests
  - SDNs etc.

ATLAS and CMS: Rucio injection tool - new and improved version!

\+

XRootD small read load-gen

XRootD in DC27

VO jobs read data via XRootD

- LHCb showed that they want to do internal reads
  - Does not count towards the official DC activities - only WAN
  - Still puts pressure on the storages
- CMS also stream data from remote via AAA
  - Currently there is no monitoring to show this (see Panos' talk)
  - Past monitoring shows AAA streaming doing similar rates to FTS transfers at times
- New tools in progress to replicate XRootD reads
  - See Rahul's talk - a load-test tool for CMS remote reads
  - Alex R writing something similar for LHCb

Attempt to sum requested DC27 rates for all VOs

Provisional!!!

Note: e.g. CMS Computing week next week
ATLAS beginning of October

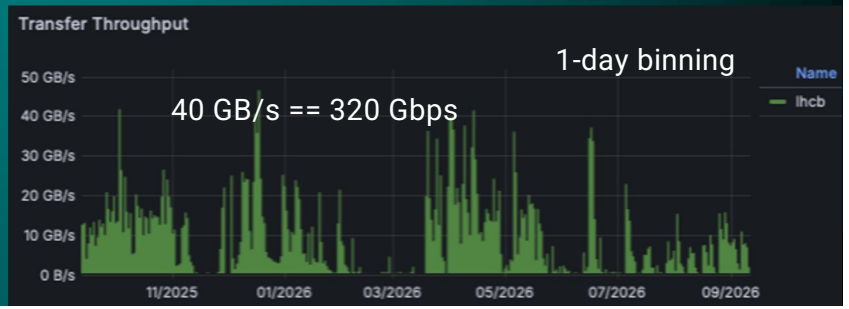
Attempt to sum requested DC27 rates for all VOs

- CMS is proposing to target 1 Tbps baseline sustained, and 1.5 Tbps flexible (stretch) goal
- ATLAS is looking at 1.4 Tbps as baseline and 2.1Tbps as flexible target (still ruminating though)

CMS
Minimal 0.5 Tbps (DC24) → 1 Tbps (DC27)
Flexible 1 Tbps → 1.5 Tbps

ATLAS
Minimal 0.7 Tbps (DC24) → 1.4 Tbps (DC27)
Flexible 1.4 Tbps → 2.1 Tbps

LHCb
Demonstrating Run-3 rates, with increased control and complexity compared with DC24



Attempt to sum requested DC27 rates for all VOs

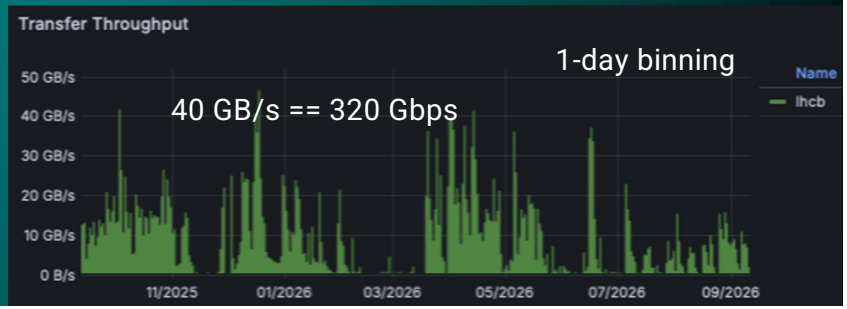
- CMS is proposing to target 1 Tbps baseline sustained, and 1.5 Tbps flexible (stretch) goal
- ATLAS is looking at 1.4 Tbps as baseline and 2.1Tbps as flexible target (still ruminating though)

CMS
Minimal 0.5 Tbps (DC24) → 1 Tbps (DC27)
Flexible 1 Tbps → 1.5 Tbps

1.5 + 2.1 + 0.32 + other VOs >= **4 Tbps** (including Production traffic)

ATLAS
Minimal 0.7 Tbps (DC24) → 1.4 Tbps (DC27)
Flexible 1.4 Tbps → 2.1 Tbps

LHCb
Demonstrating Run-3 rates, with increased control and complexity compared with DC24



What do FTS actually want to know?

- VO plans and rates…but more important than throughput is number of files
  - How many active transfers?
  - How many file transfers launched per second?
- Also, how many cancelled transfer jobs?

What do FTS actually want to know?

Some back-of-the-envelope calculations:

	Avg for DC27	Flex model	Flex model	Per-file rate	Time-in-flight	Flex model
VO	File size (GB)	FTS rate (GB/s)	New file rate (Hz)	Avg (MB/s)	Avg (s)	Files in flight
ALICE	10					
ATLAS	4	263				
CMS	4	156				
LHCb	4					
DUNE						
JUNO						
Belle II						
SKA	1000?					

= 2.1 Tbps

Many values in this table are not calculated by the experiments - this analysis purely for illustration!

What do FTS actually want to know?

Some back-of-the-envelope calculations:

	Avg for DC27	Flex model	Flex model	Per-file rate	Time-in-flight	
VO	File size (GB)	FTS rate (GB/s)	New file rate (Hz)	Avg (MB/s)	Avg (s)	Files in flight
ALICE	10					
ATLAS	4	263	66			
CMS	4	156				
LHCb	4					
DUNE						
JUNO						
Belle II						
SKA	1000?					

Many values in this table are not calculated by the experiments - this analysis purely for illustration!

What do FTS actually want to know?

Some back-of-the-envelope calculations:

	Avg for DC27	Flex model	Flex model	Per-file rate	Time-in-flight	
VO	File size (GB)	FTS rate (GB/s)	New file rate (Hz)	Avg (MB/s)	Avg (s)	Files in flight
ALICE	10					
ATLAS	1	263	263			
CMS	4	156				
LHCb	4					
DUNE						
JUNO						
Belle II						
SKA	1000?					

Many values in this table are not calculated by the experiments - this analysis purely for illustration!

What do FTS actually want to know?

Some back-of-the-envelope calculations:

	Avg for DC27	Flex model	Flex model	Per-file rate	Time-in-flight	
VO	File size (GB)	FTS rate (GB/s)	New file rate (Hz)	Avg (MB/s)	Avg (s)	Files in flight
ALICE	10					
ATLAS	0.25	263	1050			
CMS	4	156				
LHCb	4					
DUNE						
JUNO						
Belle II						
SKA	1000?					

Many values in this table are not calculated by the experiments - this analysis purely for illustration!

What do FTS actually want to know?

Some back-of-the-envelope calculations:

	Avg for DC27	Flex model	Per-file rate	Time-in-flight	
VO	File size (GB)	FTS rate (GB/s)	Avg (MB/s)	Avg (s)	Files in flight
ALICE	10				
ATLAS	4	263			
CMS	4	156	???		
LHCb	4				
DUNE					
JUNO					
Belle II					
SKA	1000?				

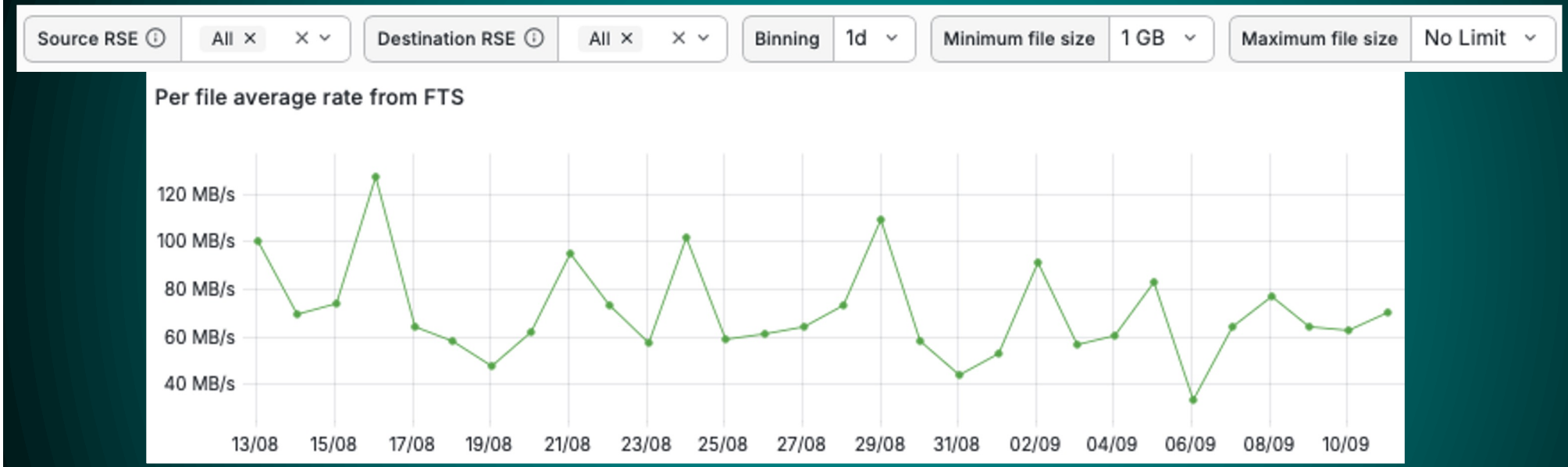
Many values in this table are not calculated by the experiments - this analysis purely for illustration!

What is ‘per-file rate’?

- i.e. the transfer rate in MB/s of an individual file
- Maybe FTS have some (new) monitoring to tell us this?
- For now, I try to use some (unverified) CMS monitoring…

What is ‘per-file rate’?

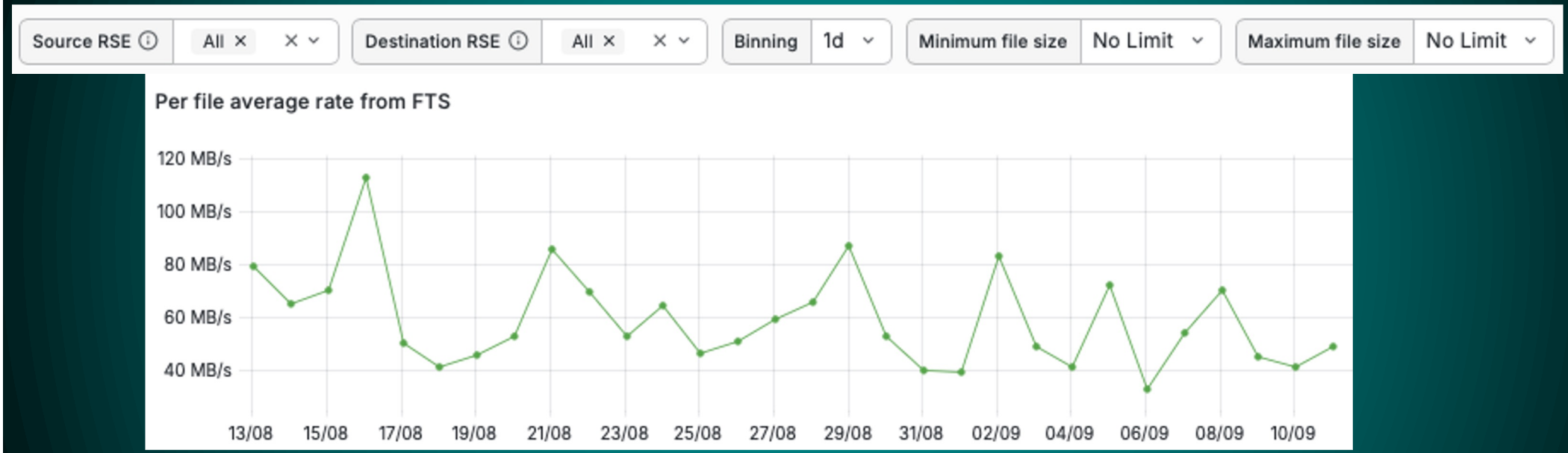
- For now, I try to use some (unverified) CMS monitoring…



Last 30 days, averaged per day over all CMS FTS transfers
NOTE: minimum file size = 1 GB
Average per file transfer rate is approx. 70 MB/s ??

What is ‘per-file rate’?

- For now, I try to use some (unverified) CMS monitoring…

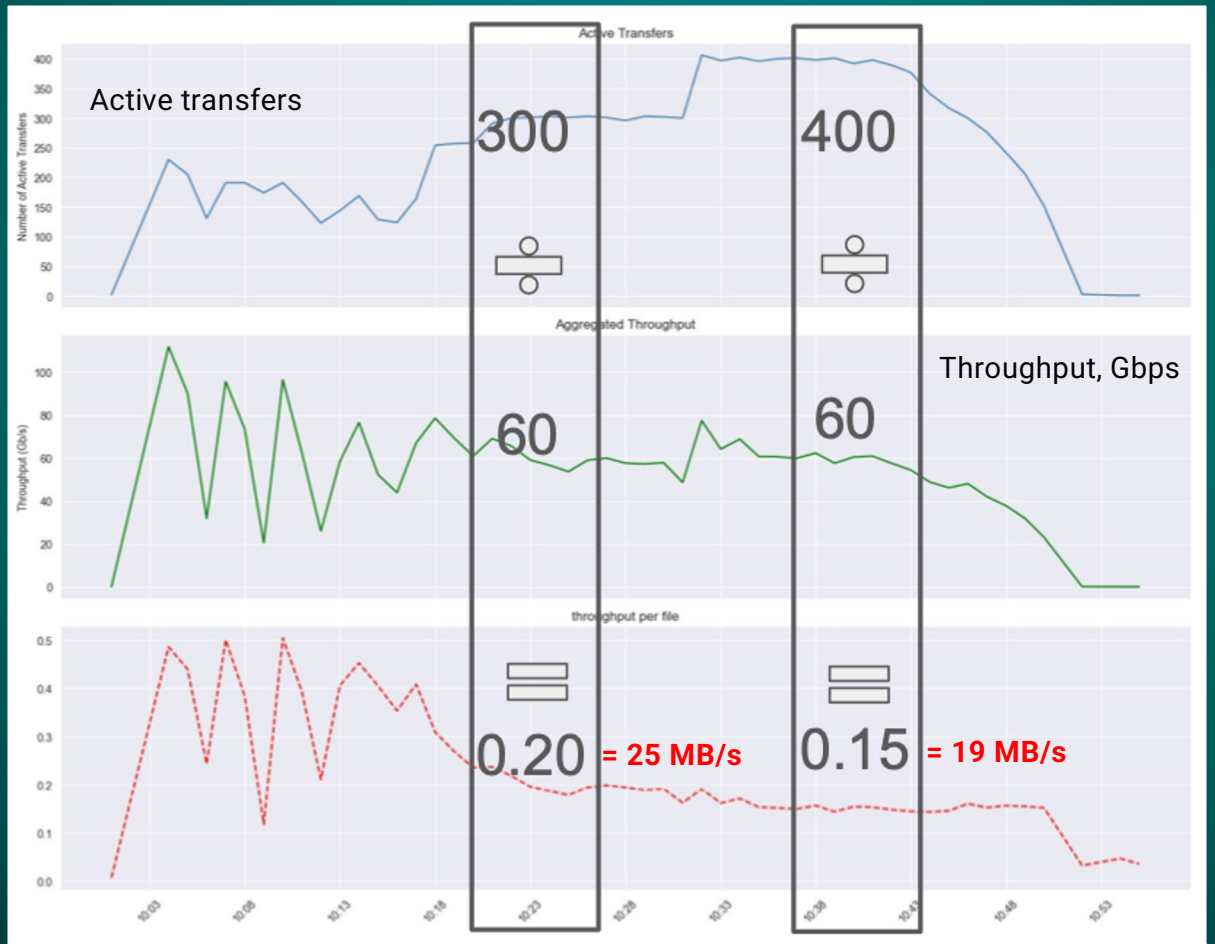


Remove the minimum file size, rate drops slightly - less than I feared
Average per file transfer rate now approx. 55 MB/s ??

HOWEVER, many things may affect this number, as you can see even averaging over a day shows quite some variability. I expect the per-file rate to DECREASE as load and number of parallel transfer increases

What is ‘per-file rate’?

- I can also cross-check with results from Diego (USCMS mini-DC)



Diego Davila, DOMA meeting
2nd Sept 2026

What do FTS actually want to know?

Some back-of-the-envelope calculations:

	Avg for DC27	Flex model	Per-file rate	Time-in-flight	
VO	File size (GB)	FTS rate (GB/s)	Avg (MB/s)	Avg (s)	Files in flight
ALICE	10				
ATLAS	4	263	25	160	10500
CMS	4	156	25	160	6250
LHCb	4				
DUNE					
JUNO					
Belle II					
SKA	1000?				

= target rate (MB/s) divided by per-file rate

Many values in this table are not calculated by the experiments - this analysis purely for illustration!

What do FTS actually want to know?

Some back-of-the-envelope calculations:

	Avg for DC27	Flex model	Per-file rate	Time-in-flight	
VO	File size (GB)	FTS rate (GB/s)	Avg (MB/s)	Avg (s)	Files in flight
ALICE	10				
ATLAS	4	263	10	400	26250
CMS	4	156	10	400	15625
LHCb	4				
DUNE					
JUNO					
Belle II					
SKA	1000?				

What if?

= target rate (MB/s) divided by per-file rate

Many values in this table are not calculated by the experiments - this analysis purely for illustration!

What do FTS actually want to know?

Some back-of-the-envelope calculations:

	Avg for DC27	Flex model	Per-file rate	Time-in-flight	
VO	File size (GB)	FTS rate (GB/s)	Avg (MB/s)	Avg (s)	Files in flight
ALICE	10				
ATLAS	4	263	6	667	43750!!
CMS	4	156	10	400	15625
LHCb	4				
DUNE					
JUNO					
Belle II					
SKA	1000?				

What if?

= target rate (MB/s) divided by per-file rate

Many values in this table are not calculated by the experiments - this analysis purely for illustration!

What do FTS actually want to know? Tape

- FTS: Careful counting of tokens required
- VOs: Rates to tape determined and average file size
- DC27: Tape a likely limitation
  - Gathering information from sites about tape capabilities
  - Actual summed rate across tape systems will be the limiting factor

Summary

- DC27 is planned for end of Feb/beginning of March 2027
- A 2-week test of our data transfer system used by WLCG and other experiments
  - Storage (disk and tape), network, middleware (FTS, XRootD, Rucio, IAM)
  - Experiments: ATLAS, CMS, LHCb, ALICE, DUNE, Belle II, Juno, SKA, others?
  - CERN, T1s, many T2s
- ATLAS and CMS expect a big jump in data rates as Run-4 ramps up; ALICE + LHCb as Run-3
- Other experiments may not have comparable throughput expectations but making sure they can operate in the environment is still a useful exercise
- Pre-challenge tests to take place between now and DC27 - using FTS4
- Number of ‘active transfers’ in FTS strongly dependent on individual file transfer speeds?
  - This itself could be dependent on many things:
    - Source and destination
    - File size
    - Number of concurrent transfers per link
    - Other?
- More details on VO rates at the next DOMA-General(s) and discussion of schedule at WLCG