



Status of Monitoring of DOMA Components

Borja Garrido Bear (On behalf of the WLCG-mon TaskForce)
CERN IT Monitoring team

WLCG/HSF 2025, 08.05.2025

Index

- Why monitoring is important
- What are the DOMA components?
- How are they being monitored?
- Status
- Issues and plans
- Conclusions

Why Monitoring is Important?

- **Ensure the correct operation of the service (Operations Monitoring)**
 - “Short-term” lived, fine granularity metrics
 - Service logs
- **Visualize service performance (Accounting Monitoring)**
 - “Long-term” lived, coarse granularity (aggregated) metrics
 - Generally based on the operations monitoring metrics
- **In general anything allowing to answer questions on the service status**

**DIFFERENT
AUDIENCES!
(usually)**

What's the question WLCG wants to answer? (Goal)

- **It's actually probably more than one**
 - Number of transfers, throughput, efficiency...
 - Number of transfers over IPv6
 - Number of transfers using tokens
- **Current focus in DC27!**
 - Are we reaching the expected data rate?
 - Can we answer the question?
 - Can we trust the answer?



What are the DOMA components?



FTS
File Transfer Service

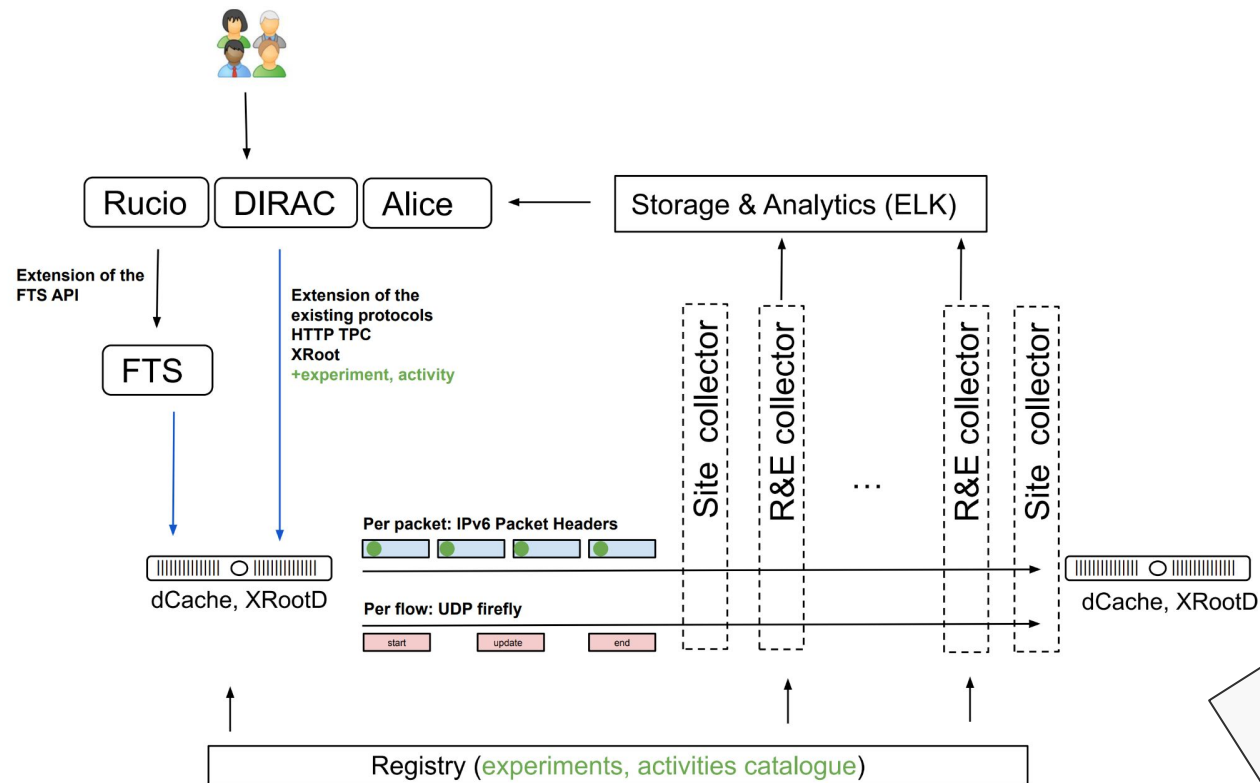


XRootD



Scitags

- Enhancement for network monitoring
- Can be used also for enriching transfer monitoring
- Will allow to contrast transfers and network monitoring at more granular levels (vo, activity)

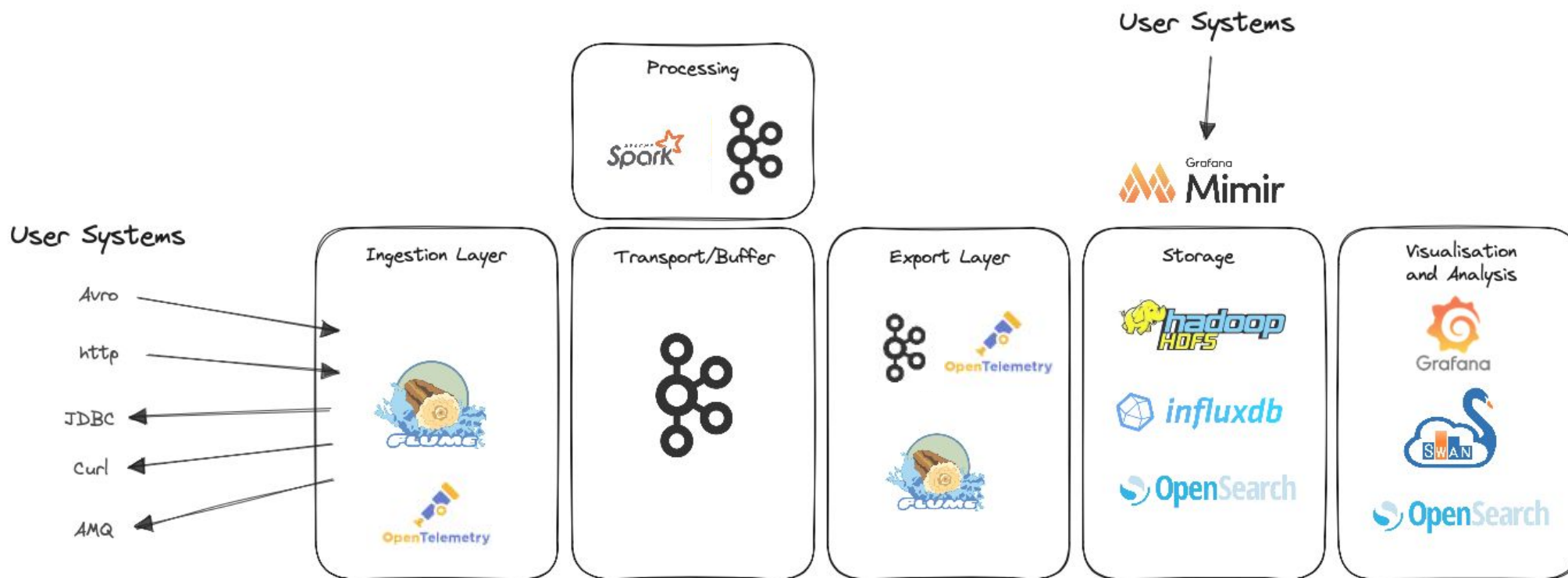


See Shawn Mc Kee latest HEPIX [slides](#)

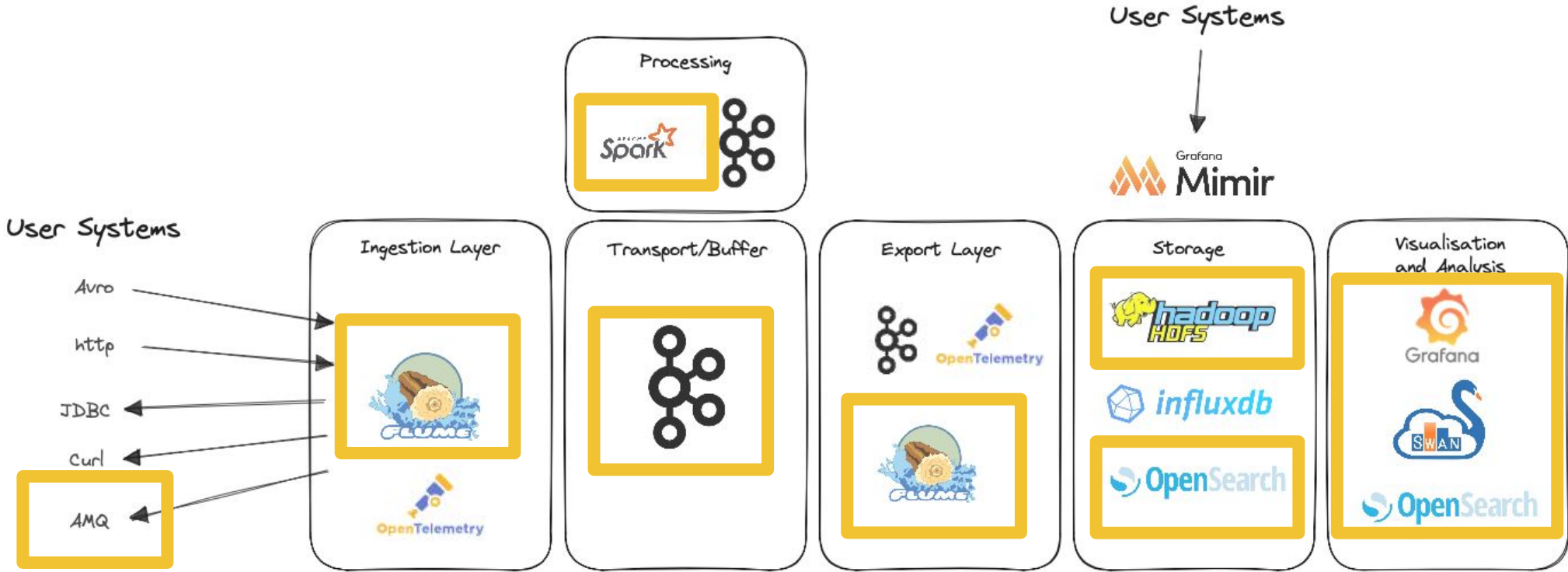
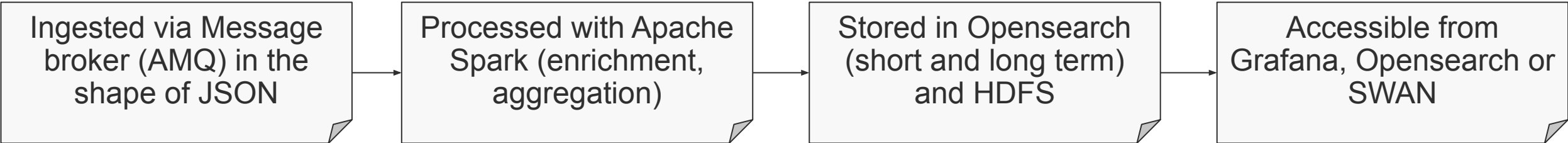
How are they being monitored?

The MONIT infrastructure

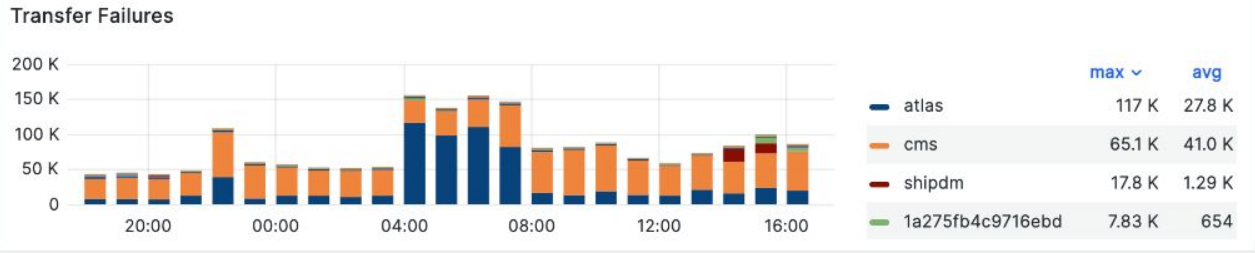
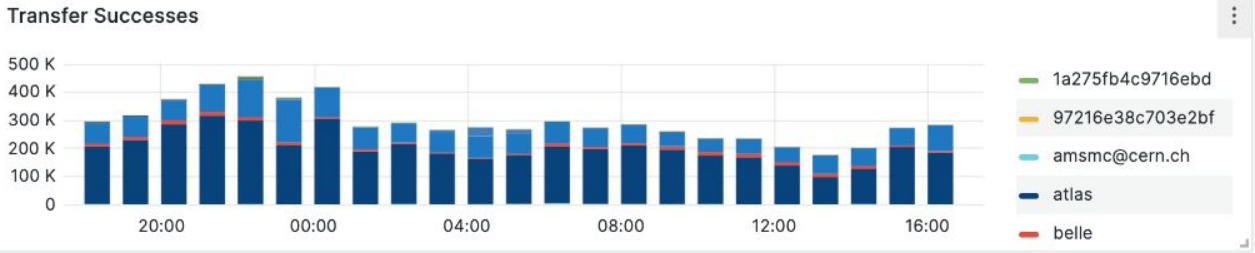
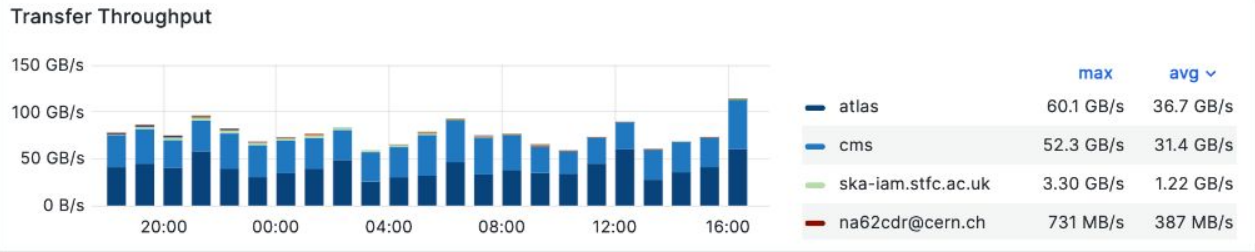
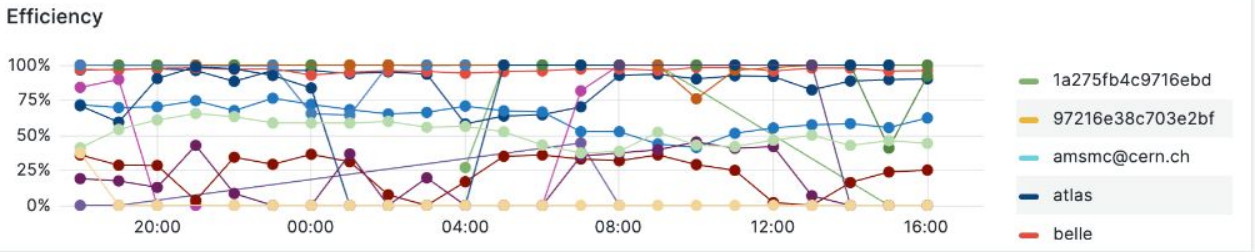
The aim of the monitoring team is to provide an infrastructure allowing the usage of tools and service to ease the load of monitoring CERN data centres (host and services) and WLCG experiments



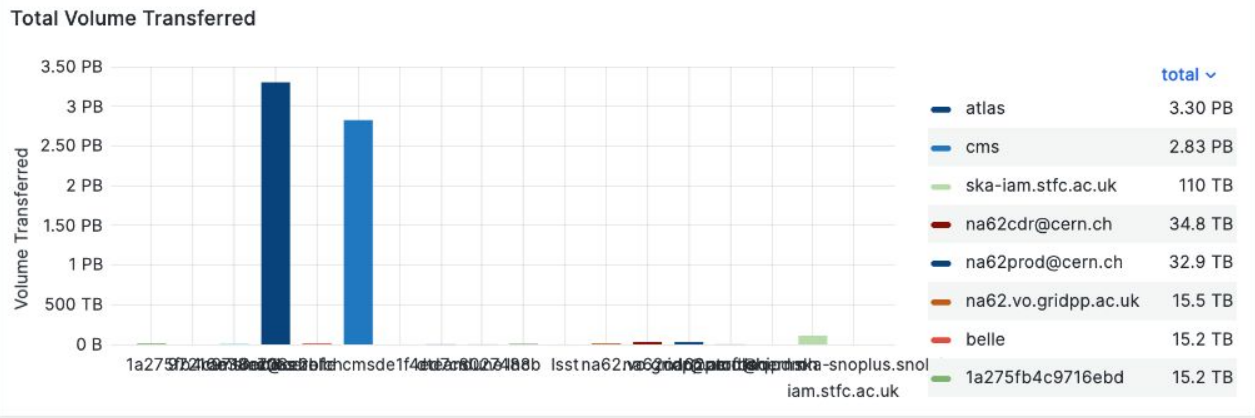
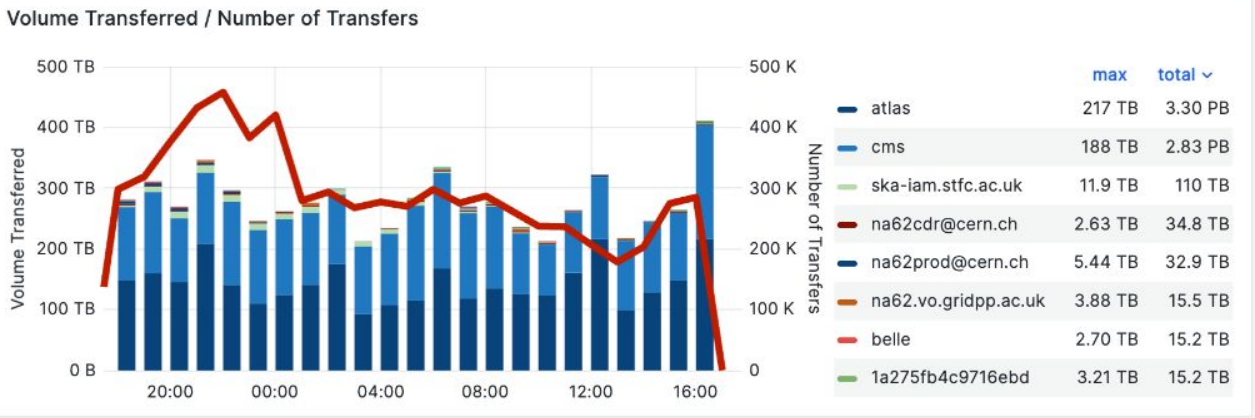
Same set of tools



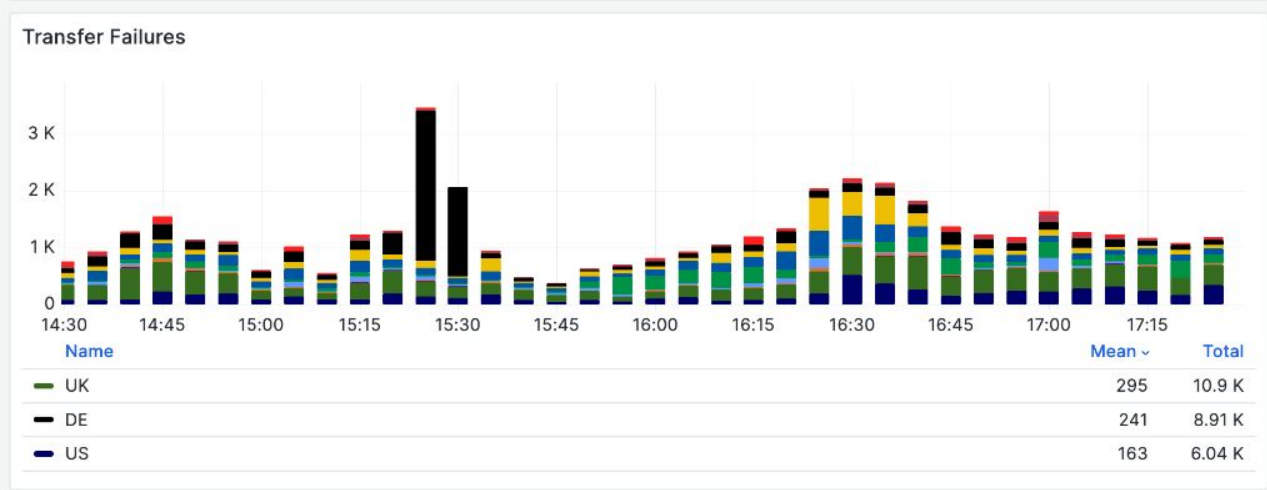
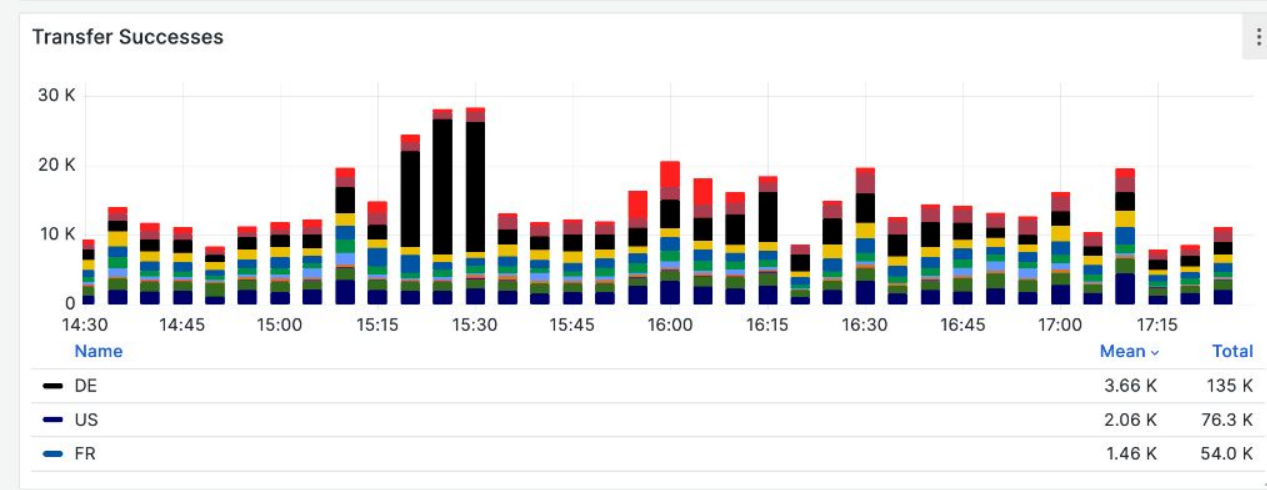
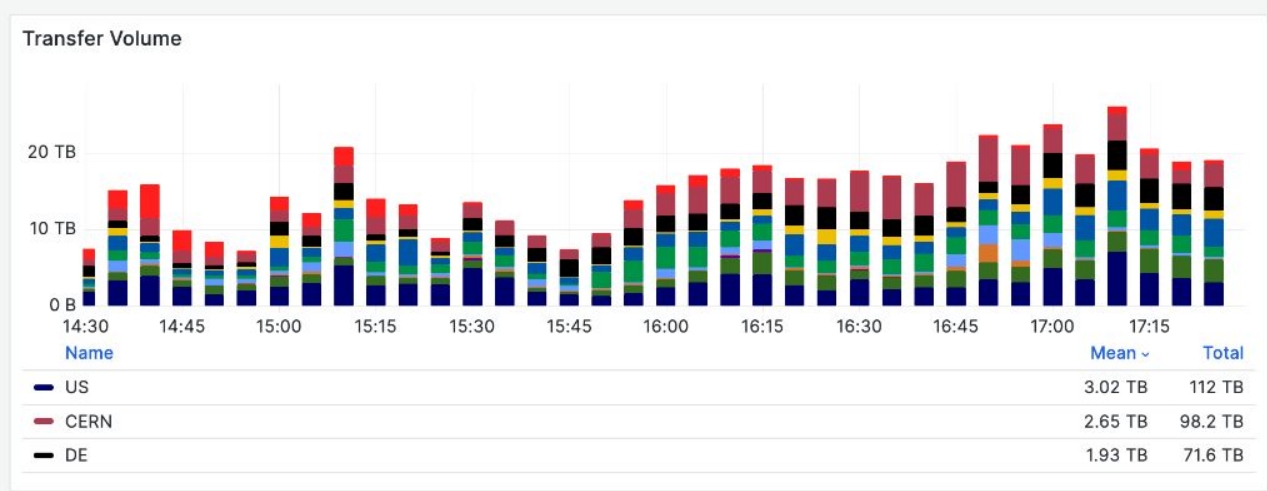
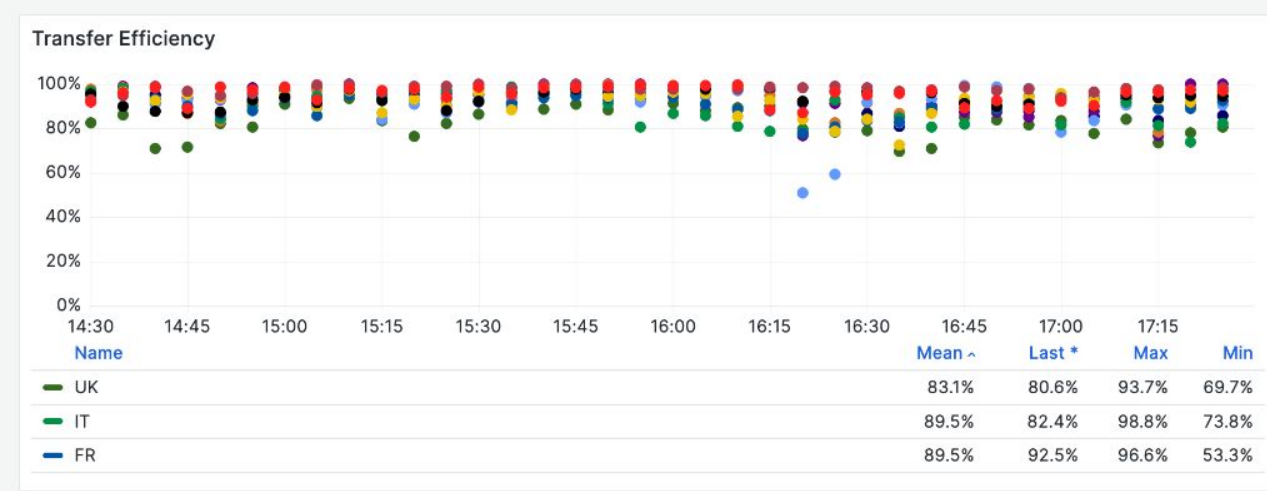
Transfer plots



Volume Statistics



Transfer Plots



Transfer Throughput

Transfer Volume



CMS Rucio (WIP)



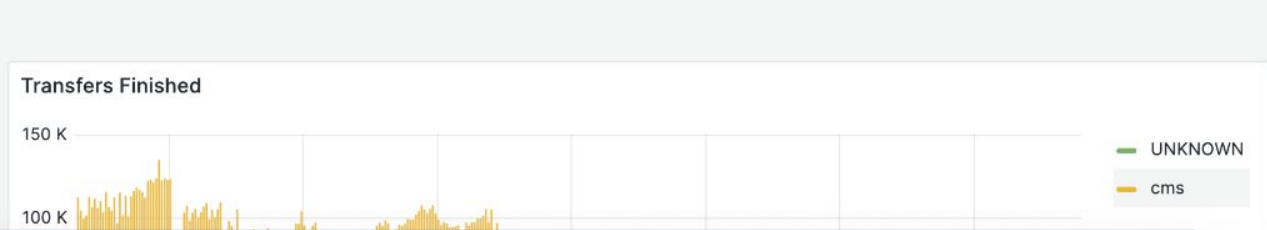
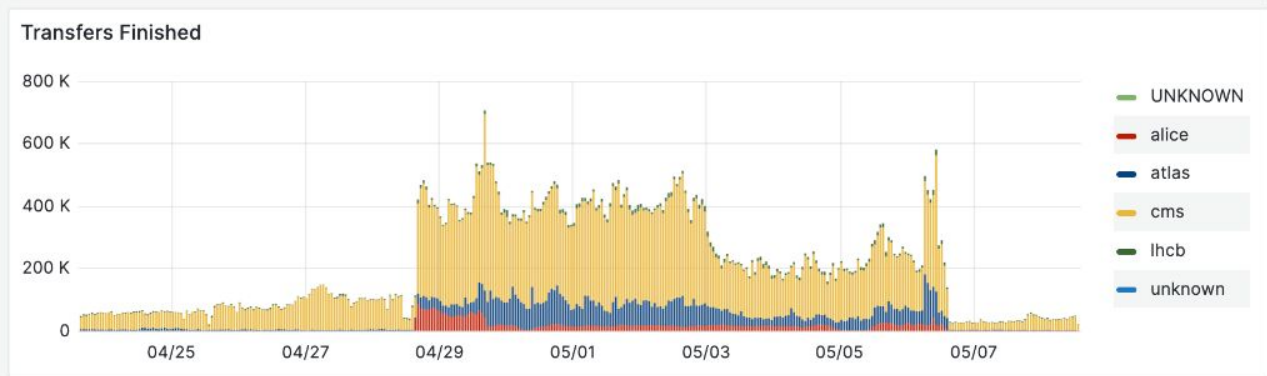
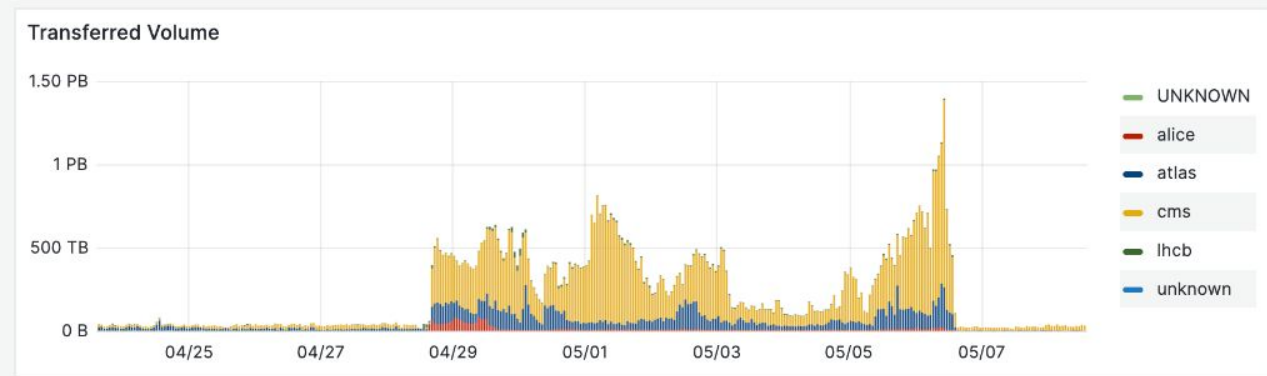
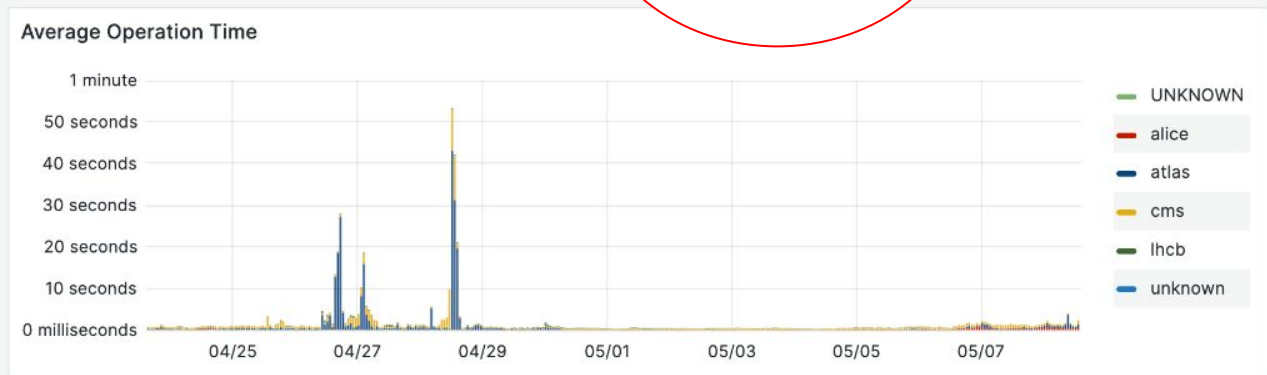
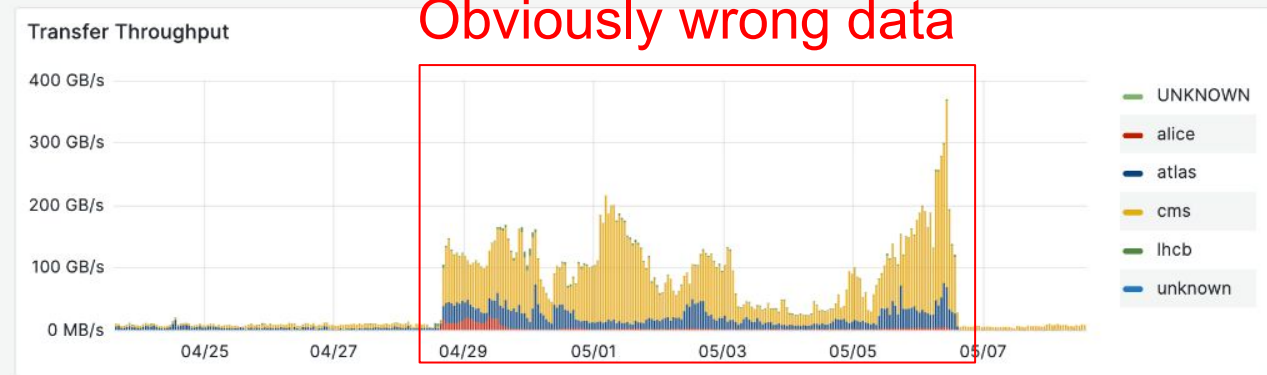
- **CMS is using Rucio and they are interested on monitoring things similar to ATLAS**
- **ATLAS DDM required a set of jobs in Apache Spark to work**
 - These are developed/run by MONIT
 - Currently these have some ATLAS specificities
- **On the MONIT side it's on our plans to provide an equivalence for CMS**

Group By: vo Bin 1h VO All Source Country All Dest Country All Source Tier All Dest Tier All Source Site All Dest Site All

Remote Access: true IP Version: All Filters +

Not validated data!

Obviously wrong data



What do we need to answer the question?

- **This specific question is more aimed for aggregated (accounting) data**
 - Put the service in perspective over a long time interval to see how DC performed
 - Based on operations data for these flows (Needed!)
- **We need to distinguish activity, vo and site for the transfers**
- **Big difference between FTS and XRootD monitoring**
 - FTS data was integrated and **consumed** by the FTS users
 - The fact it's used by experts in a daily basis makes it easier to trust/curate it
 - XRootD Monitoring (access via xrootd protocol) of the servers and dCache (XRootD door) is **not actively being consumed** by experts

Status

Roadmap up to now



DC2021

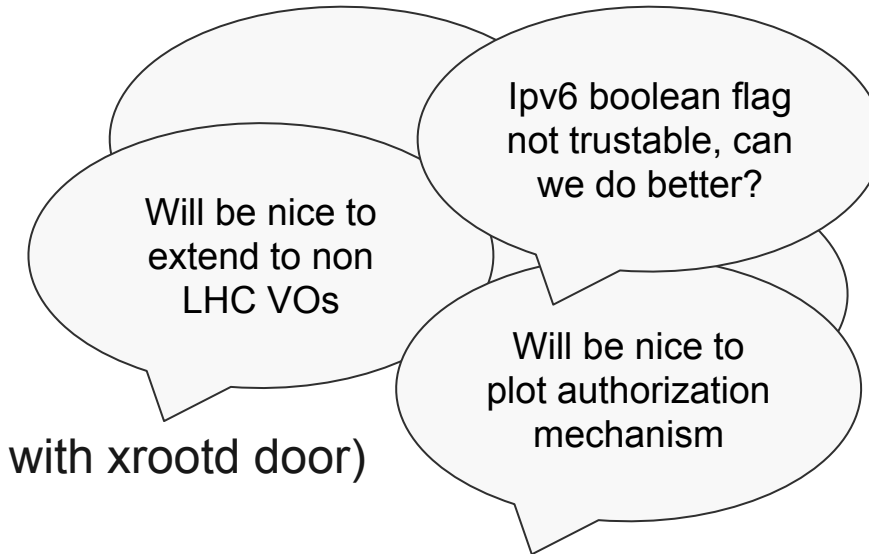
- We need network monitoring
- We need XRootD monitoring (CMS, Alice)
- Dashboards could be improved ([DC21 Dashboard](#))

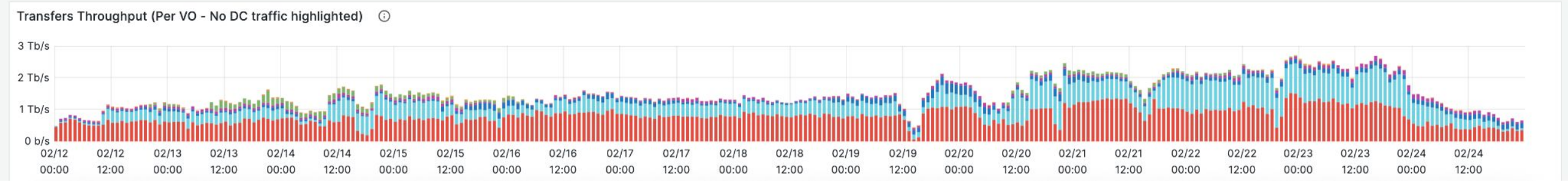
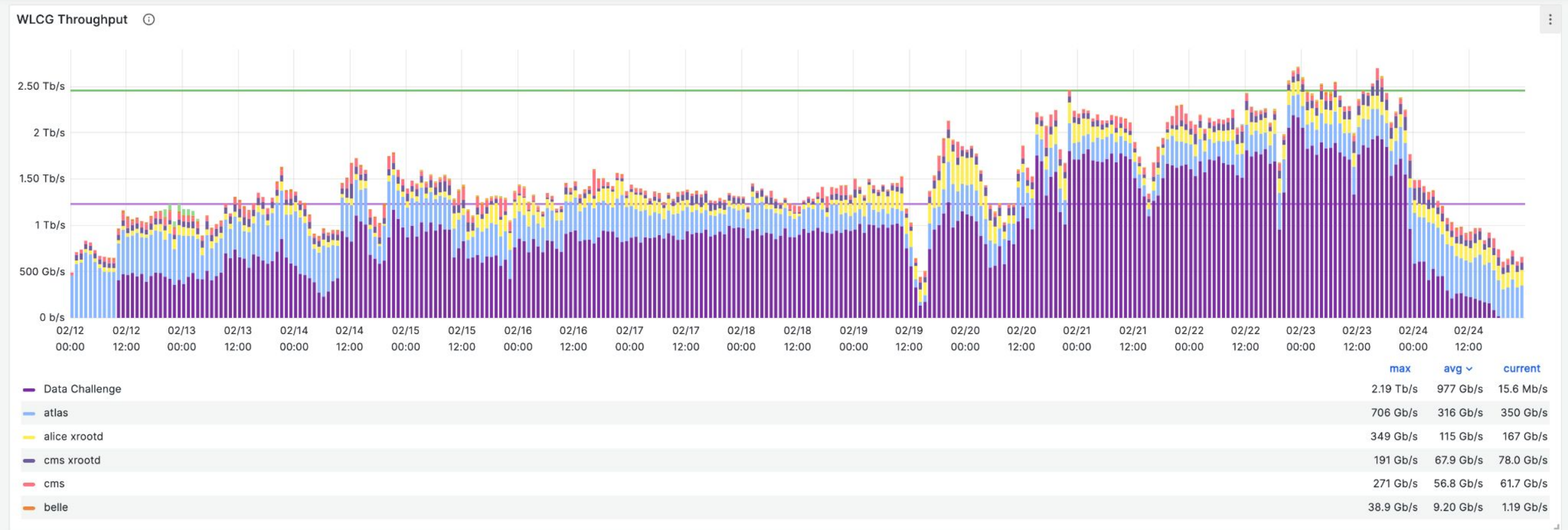
Taskforce working points

- Provided network monitoring ([Dashboard](#))
- Work started on providing XRootD monitoring
 - Big push to have CERN and FNAL before DC24
- Dashboard was improved ([DC24 Dashboard](#))
 - It's draft to this moment!

DC2024

- We should validate XRootD monitoring (servers and dcache with xrootd door)
- We should extend XRootD monitoring to more sites
- Most of the side questions answered by FTS
 - Some are still working points for XRootD (see plans later)
- Dashboards could be improved





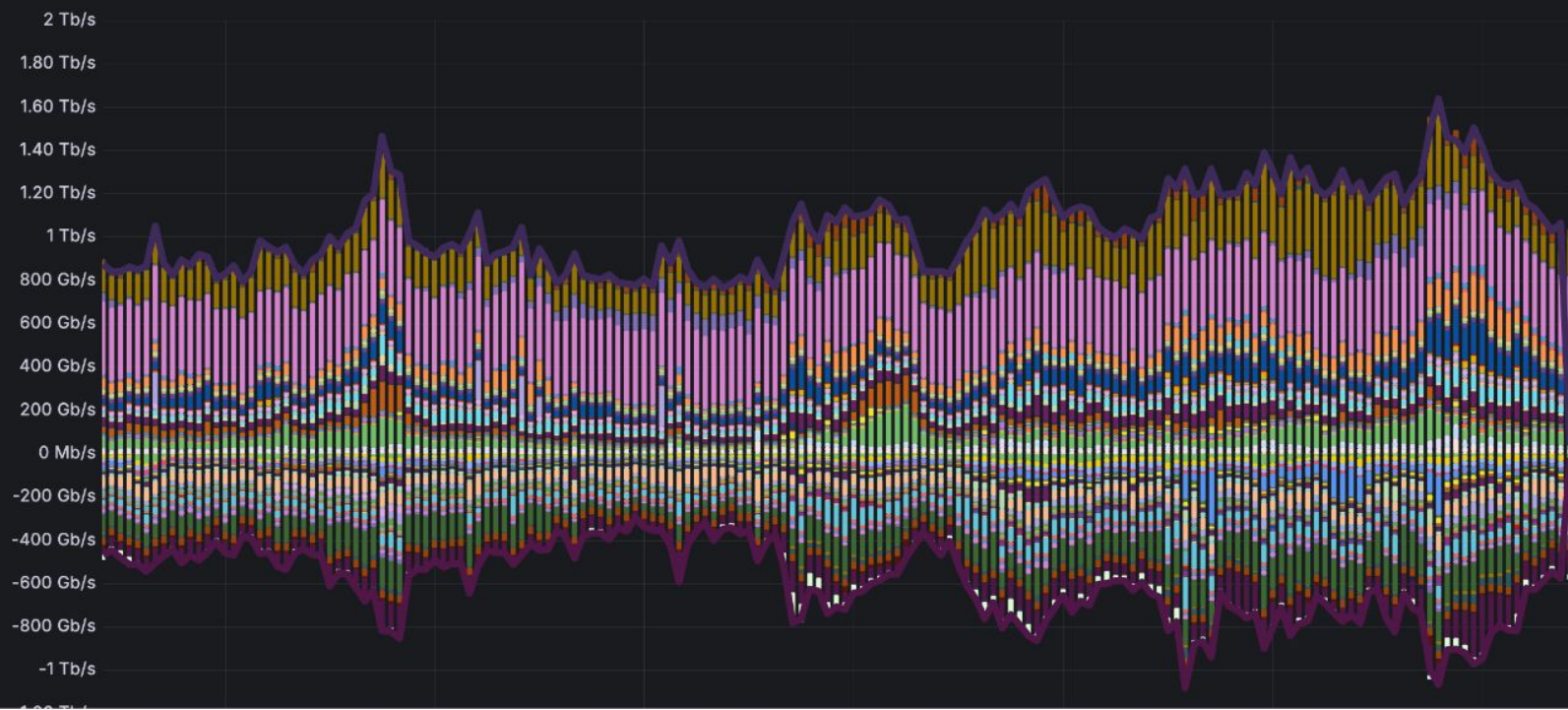
Number of s...

48

Number of sites that answer to request in time



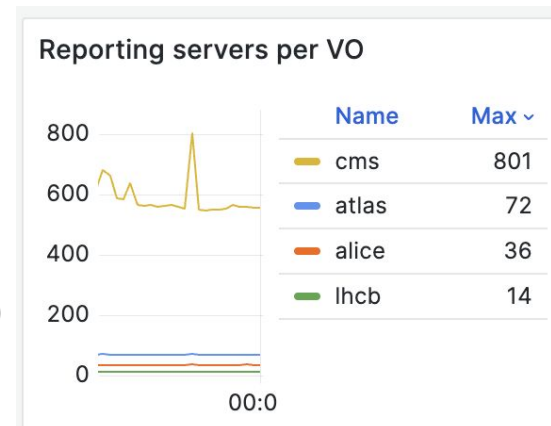
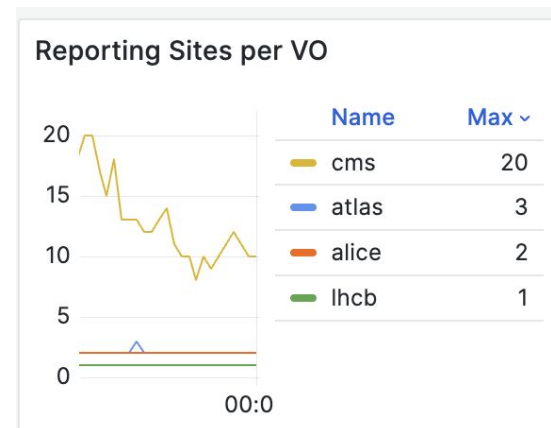
WLCG Site Network Input/Output



IN: pragueicg2 - CZ-pragueicg	Mean: 14.2 Gb/s	Max: 56.0 Gb/s
IN: pic - ES-PIC	Mean: 16.1 Gb/s	Max: 38.8 Gb/s
IN: ifae - ES-IFAE	Mean: 16.1 Gb/s	Max: 38.8 Gb/s
IN: WEIZMANN-LCG2 - IL-WEIZMANN	Mean: 322 Mb/s	Max: 881 Mb/s
IN: Vanderbilt - US-Vanderbilt	Mean: 7.22 Gb/s	Max: 21.9 Gb/s
IN: USCMS-FNAL-WC1 - US-CMS	Mean: 20.2 Gb/s	Max: 244 Gb/s
IN: UKI-SOUTHGRID-RALPP - UKI-SOUTHGRID-RALPP	Mean: 2.69 Gb/s	Max: 14.8 Gb/s
IN: UKI-SCOTGRID-GLASGOW - UKI-SCOTGRID-GLASGOW	Mean: 3.88 Gb/s	Max: 8.48 Gb/s
IN: UKI-NORTHGRID-MAN-HEP - UKI-NORTHGRID-MAN-HEP	Mean: 4.74 Gb/s	Max: 16.7 Gb/s
IN: UKI-NORTHGRID-LANCS-HEP - UKI-NORTHGRID-LANCS-HEP	Mean: 6.93 Gb/s	Max: 26.9 Gb/s
IN: UFlorida-HPC - US-UFlorida-HPC	Mean: 2.29 Gb/s	Max: 21.8 Gb/s
IN: UCSDT2 - UCSD	Mean: 6.78 Gb/s	Max: 22.4 Gb/s
IN: UAM-LCG2 - ES-UAM	Mean: 111 Mb/s	Max: 4.77 Gb/s
IN: TRIUMF-LCG2 - CA-TRIUMF	Mean: 6.81 Gb/s	Max: 6.81 Gb/s
IN: TOKYO-LCG2 - JP-University-Tokyo	Mean: 9.79 Gb/s	Max: 66.9 Gb/s
IN: SPRACE - BR-SPRACE	Mean: 960 Mb/s	Max: 8.12 Gb/s
IN: RWTH-Aachen - DE-RWTH-Aachen	Mean: 12.7 Gb/s	Max: 84.4 Gb/s
IN: RO-13-ISS - RO-13-ISS	Mean: 167 Mb/s	Max: 167 Mb/s
IN: RO-07-NIPNE - RO-NIPNE_1	Mean: 5.26 Gb/s	Max: 15.3 Gb/s
IN: RAL-LCG2 - UK-RAL	Mean: 48.4 Gb/s	Max: 110 Gb/s
IN: Purdue - US-Purdue	Mean: 549 b/s	Max: 549 b/s
IN: Nebraska - US-Nebraska	Mean: 714 Gb/s	Max: 43.8 Gb/s

- **From the WLCG perspective we have all information needed***
 - Although we thought the same and only discovered few issues during the DC24
 - **I.e:** Accounted transferred volume for failed transfers was using full file size
 - Way to prevent this will be to have experts looking at the data used for the DCs! (Mini Challenges?)
- **There will be some evolution towards DC27**
 - See [Mihai's talk](#) for more information

- **Monitoring based on the shoveler, Collector flow**
 - Using XRootD monitoring stream to shape a “transfer” document
- **Currently few sites are (partially?) integrated**
 - We decided not to ask many sites to join before validating properly the data
- **Data has been validated at high level**
 - “More or less” seems to match (fine for accounting data, bad for operations)
 - We need to do a proper validation to ensure it can be trusted
 - And experts need to look to it from time to time!
- **Will require some extra work to add extra streams information**
 - Sci-tags: New stream provides activity information (Required)
 - Tokens: New stream, provides token issuer information
 - Token issuer can be “easily” mapped to a VO (Required)
 - It also will allow to know if the transfer was scheduled with tokens (nice to have)



dCache with XRootD door



- **Monitoring based on script integrating monitoring data available in Kafka**
 - Sites need to run the full monitoring stack (Kakfa + script to forward WLCG data)
- **Currently few sites are integrated**
 - A bit hard to tell based on the Monitoring but mainly FNAL
- **Data has not been validated**
 - As far as I personally know
- **Missing activity information**
 - We will need to follow with dCache developers how the integration with scitags is going to happen



Issues and plans

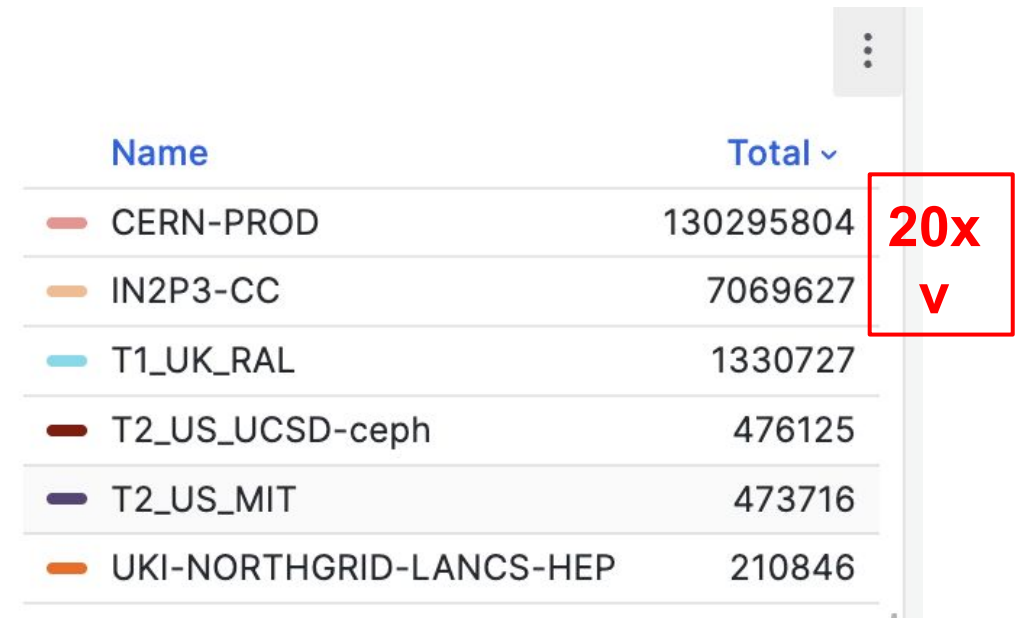
After one year of running new flow...

- **Three wrong initial assumptions**

- Sites are monitoring XRootD (using the streams)
 - Is there anyone using XRootD streams seriously?
 - We focused on aggregated data
 - As we only had ALICE Monalisa data for comparison
- Let's follow what we did for FTS
 - FTS operations data is of interest for FTS team
 - Data is validated and aggregated can be compared to it
- Integration is only for WLCG high level view
 - Interest from sites to use this for operations monitoring
 - Is the current deployment a good match for this?

- **Flow scalability issues**

- CERN numbers are big!
 - We have been working on improving this
 - Still some things that make it hard: [#1873](#)
- This made validation not possible at all
- Will it hold all the sites?



Name	Total
CERN-PROD	130295804
IN2P3-CC	7069627
T1_UK_RAL	1330727
T2_US_UCSD-ceph	476125
T2_US_MIT	473716
UKI-NORTHGRID-LANCS-HEP	210846

Plans

- **Low hanging fruits (2025 Q3-Q4)**
 - Provide shoveler under WLCG repository
 - Implement missing streams support in Collector: token and sci-tag
- **Adding failure monitoring to XRootD**
 - On XRootD roadmap [#2352](#)
 - Important for operations monitoring, nice to have for the DC27
- **Validation/first curation of data (2025 - 2026)**
 - Tried to do validation at high level
 - As we could only compare to ALICE Monalisa
 - New plan (start from small controlled environment) and focus on getting good operations data
 - **This will require CMS sites involvement (main audience)**
- **Campaign to get more sites in (Q1-Q2 2026)**
 - IMO, should not happen before we are happy with the current data
 - Realistic Target: All T1 and T2?
- **DC Dashboard improvements (< 2027)**
 - Experts should ask for them a bit in advance (Please)
- **Data Challenges (2027)**

Conclusions

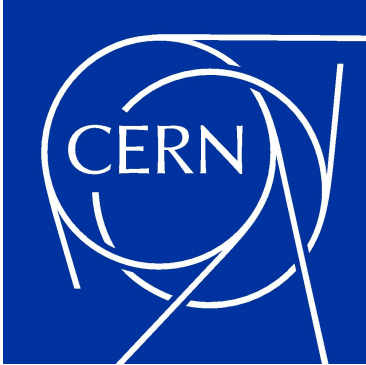
- **Experts involvement is required**
 - Little value on having data flowing if no one validates it
 - We can build “any” dashboard for we might not know exactly what’s required
- **We should focus on having good operations data**
 - Will caught the attention of the experts and helm them monitoring their system
 - If operations data is good, accounting data based on it should also be
- **Running the shoveler comes with extra work**
 - Depending how big the site is might require to be scaled
 - If you want to run it properly it requires also to monitor/alarm it
- **We should not wait until the last minute**
 - Already twice there were big pushes on the day/s before and during the Data Challenges
 - We should try to think in advance what will be needed and plan things properly

Thank you !

Q & A

Contact: wlcgmon-tf@cern.ch





home.cern