

# Mini-Data Challenge Program of Work

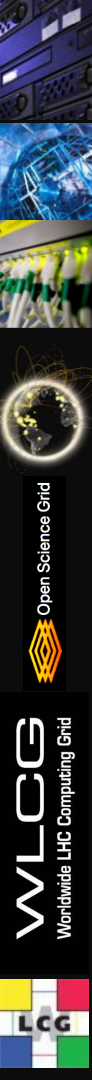
## *Advancing R&D and Operations for HL-LHC Readiness*

Diego Davila / **UCSD**, Shawn McKee / **UMich**

IRIS-HEP Retreat (<https://indico.cern.ch/event/1576135/>)

October 9th, 2025





# Section 1. Data Challenge and Mini-challenges Overview

# Data Challenge Overview: Introduction

The WLCG Data Challenges are a ~biennial series of four increasingly-complex exercises which started in 2021 and are aimed at demonstrating readiness of our infrastructure and middleware wrt Data Transferring at the HL-LHC scale.

Basically: Baby steps to make sure that our SEs, Network and Middleware (Rucio/FTS) will be ready for HL-LHC Rates

# Data Challenge Overview: Rough Numbers

From DC original document[1]: Two main **use-cases** driving the numbers:

1. Export of RAW data from CERN =>T1s
2. Data Reprocessing T1s  $\leftrightarrow$ T2s

Two models depending on whether we can do both uses-cases together or not

## **minimal model:**

- Export and Reprocessing are done separately
- Requirement of 4.8 Tbps

## **flexible model:**

- Export and Reprocessing can be done at the same time
- Requirement of 9.6 Tbps



# Data Challenge Overview: Original vs Current Plan

**Original Plan**[1] targets fraction of HL-LHC planned rates:

- 2021 (10%)
- 2023 (30%)
- 2025 (60%)
- 2027 (100%)
- Run 4 scheduled to start in 2028

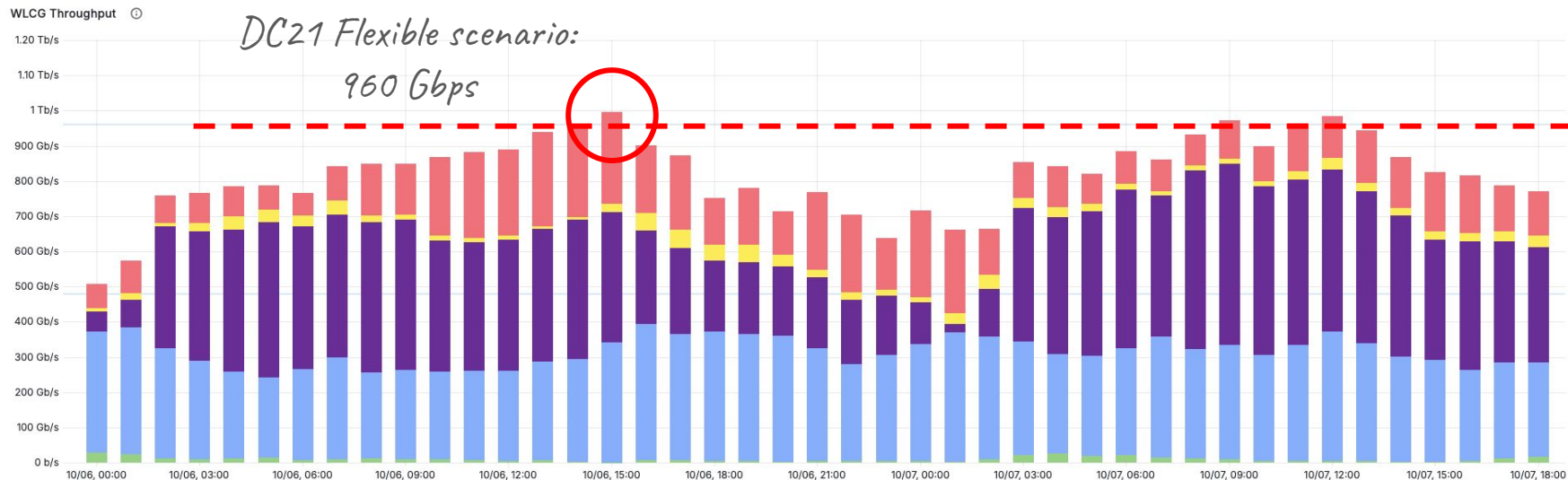
**Current Plan:**

- 2021 (10%)
- 2024 (25%)
- **2027 (50%)? (An open question is what should the HL-LHC rate be?!)**
- 2029 (100%)?
- Run 4 is scheduled to start in 2030



# Data Challenge Overview: DC21

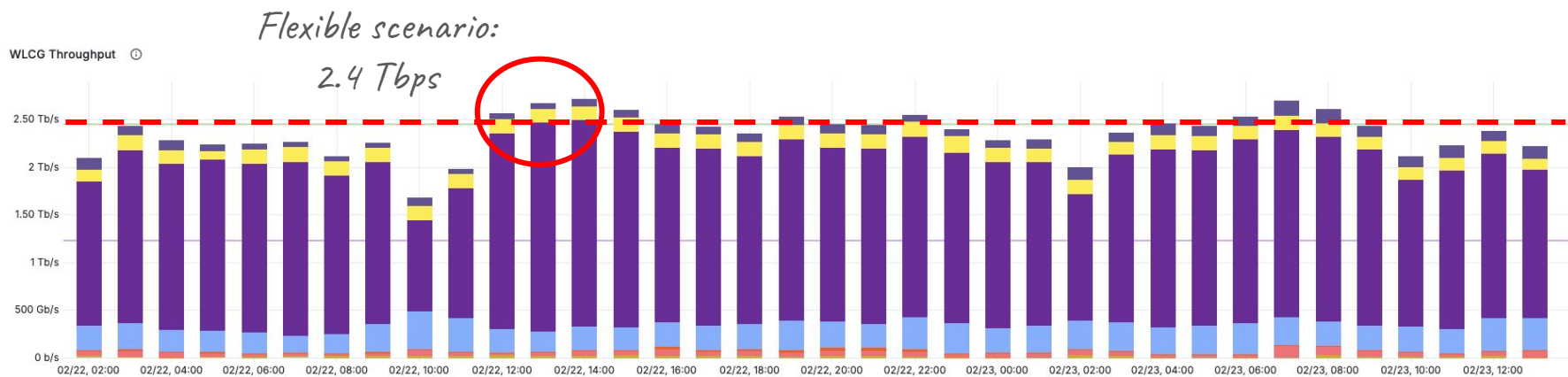
Flexible scenario end goal: 9.6Tbps, DC21: 10%



<https://monit-grafana.cern.ch/d/ZqU5ugjMz/fts-status-board?from=1633327200000&orgId=20&to=1633903200000&var-activity=Data+Challenge&var-bin=1h&var-dst=country=All&var-dst=experiment=site=All&var-dst=rse=All&var-dst=site=All&var-dst=tier=All&var-fts=server=All&var-group=by=vo&var-protocol=All&var-src=country=All&var-src=experiment=site=All&var-src=rse=All&var-src=site=All&var-src=tier=All&var-staging=All&var-vo=All&viewPanel=111>

# Data Challenge Overview: DC24

Flexible scenario end goal: 9.6Tbps, DC24: 25%



<https://monit-grafana.cern.ch/d/d3543f53-950b-4a60-b353-16611bf7f5f7/dc-2024-draft?orgId=20&from=1707692400000&to=1708815599000>

# Data Challenge Overview: Capabilities

The Data Challenge is **not only about Scale Testing** but it also has been used as a platform to showcase/test-at-scale/commission new technologies.

For example:

- The commissioning of HTTP-TPC as transfer protocol happened during DC21
- Scitokens were thoroughly tested and adopted by many sites during DC24
- Other technologies like SENSE and NOTED were also tested at lower scale during DC24



Open Science Grid



WLCG  
Worldwide LHC Computing Grid



# MiNi-Data Challenges: What they are?

Data Challenge = baby steps towards HL-LHC

Mini-Challenges = **baby steps between Data Challenges.**

Mini-data challenges are coordinated across IRIS-HEP and WLCG to incrementally prototype, test, and validate infrastructure/software innovations in preparation for HL-LHC.

Each mini-challenge targets a specific aspect of data movement, monitoring, or orchestration, building towards robust, scalable operations by 2030.

- Make sure that things (Sites, Network, Middleware) are as ready as possible for the next DC.
- Normally (but not necessarily) done within the same country and experiment
- Removes the need to synchronize across experiments, time zones, etc
- Separates Capacity (load tests) from Capabilities (new technology)



# MiNi-Data Challenges: Previous exercises

2024				
Period	Location	Experiment	Type	Description
Nov-Dec	UK	ATLAS	Capacity	T2s DC24 rates against other UK sites
Dec	UK	ATLAS	Capacity	T2 sites with 100Gb/s against other UK sites and CERN
Dec	UK	CMS	Capacity	T2 sites, 100Gb/s for Imperial College and RALPP, and 30Gb/s for Brunel
Nov-Dec	US	ATLAS	Capacity	T1 and T2s
Nov-Dec	US	CMS	Capacity	T2s DC24 rates against FNAL
2025				
Jan	CERN/FNAL	CMS	Capability	Jumbo frame evaluation on short and long RTT
Feb	CERN/BNL	ATLAS	Capability	Jumbo frame evaluation on long RTT
Feb	US	CMS	Capability	Jumbo Frames, Scitags and SENSE
March	UK	N/A	N/A	Tier1 disk
Apr	CERN/BNL	ATLAS	Capability	Jumbo frame evaluation on long RTT
N/A	CZ/CERN	N/A	Capacity	T2 upstream connection upgrade from 200Gbps to 400Gbps
N/A	CZ	N/A	Capacity	CZ Geant peering upgrade 200Gbps to 400Gbps (two paths: Viena and Frankfurt)
Aug	US	CMS	Capacity	T2s and FNAL at DC27 rates
Aug	US	ATLAS	Capacity	T2s and BNL at max capacity
Sep	CERN/US	CMS/ATLAS	Capability	LHCOPN to IPv6 only for BNL and FNAL

Source: <https://twiki.cern.ch/twiki/bin/view/LCG/DomaMiniChallenges>

# MiNi-Data Challenges: Previous exercises

2024				
Period	Location	Experiment	Type	Description
Nov-Dec	UK	ATLAS	Capacity	T2s DC24 rates against other UK sites
Dec	UK	ATLAS	Capacity	T2 sites with 100Gb/s against other UK sites and CERN
Dec	UK	CMS	Capacity	T2 sites, 100Gb/s for Imperial College and RALPP, and 30Gb/s for Brunel
Nov-Dec	US	ATLAS	Capacity	T1 and T2s
Nov-Dec	US	CMS	Capacity	T2s DC24 rates against FNAL
2025				
Jan	CERN/FNAL	CMS	Capability	Jumbo frame evaluation on short and long RTT
Feb	CERN/BNL	ATLAS	Capability	Jumbo frame evaluation on long RTT
Feb	US	CMS	Capability	Jumbo Frames, Scitags and SENSE
March	UK	N/A	N/A	Tier1 disk
Apr	CERN/BNL	ATLAS	Capability	Jumbo frame evaluation on long RTT
N/A	CZ/CERN	N/A	Capacity	T2 upstream connection upgrade from 200Gbps to 400Gbps
N/A	CZ	N/A	Capacity	CZ Geant peering upgrade 200Gbps to 400Gbps (two paths: Viena and Frankfurt)
Aug	US	CMS	Capacity	T2s and FNAL at DC27 rates
Aug	US	ATLAS	Capacity	T2s and BNL at max capacity
Sep	CERN/US	CMS/ATLAS	Capability	LHCOPN to IPv6 only for BNL and FNAL

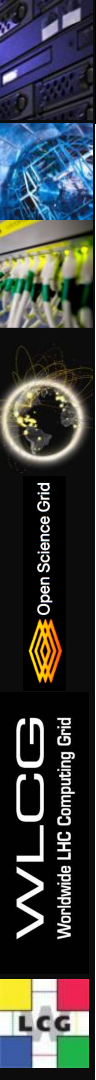
*Lead by  
IRIS-HEP  
effort*

Source: <https://twiki.cern.ch/twiki/bin/view/LCG/DomaMiniChallenges>

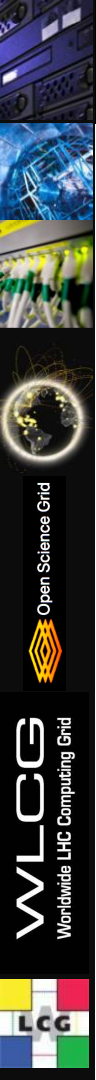


# IRIS-HEP Role

1. Leads both Capacity and Capabilities Mini-challenges in the US
  - a. See evolving plans in [Capability Planning folder](#).
2. Serves as liaison between CMS and US for Data Challenge activities
3. Communicates plans and activities to:
  - a. WLCG (DOMA General Meeting)
  - b. ESnet (ESnet-HEP Blueprint Meeting) and more generally LHCOPN/LHCONE
  - c. CMS/ATLAS



## Section 2. Summary of Latests US Mini-challenges



## Section 2.1 Latest US **Capabilities** Mini-challenges

# Latest Capabilities Mini-challenge

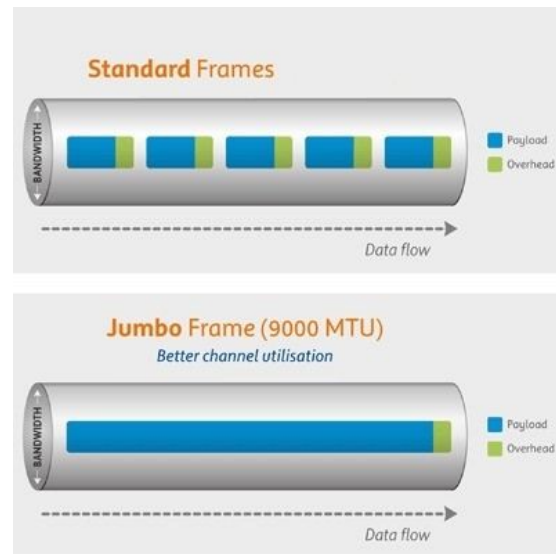
- The idea behind a “Capability” test is to:
  - try new technologies, configurations and or tunings
  - that have shown (in test beds) to enhance or TPC transfers, either by increasing throughput, visibility manageability, etc
  - as close to production as possible
- The latest Capability Mini-challenge was carried out during February 2025.
- The following Capabilities were tested:
  - Jumbo Frames
  - Rucio/SENSE
  - Scitags

# Latest Capabilities Mini-challenge: Jumbo Frames

Jumbo Frames was by far the most heavily tested capability.

The hypothesis was that by increasing the size of the Ethernet frames from 1.5 to 9k (6x) we could reduce the overall overhead to  $\frac{1}{6}$ .

- fewer packets to process: less CPU usage
- fewer bytes transmitted: less network traffic



# Latest Capabilities Mini-challenge: Jumbo Frames

Numerous tests between different pairs of sites were done with the objective of answering the following questions:

1. Would throughput increase?
2. Would RTT influence the above
3. Would it make things worse?
4. Would CPU usage decrease?

The following pairs were tested:

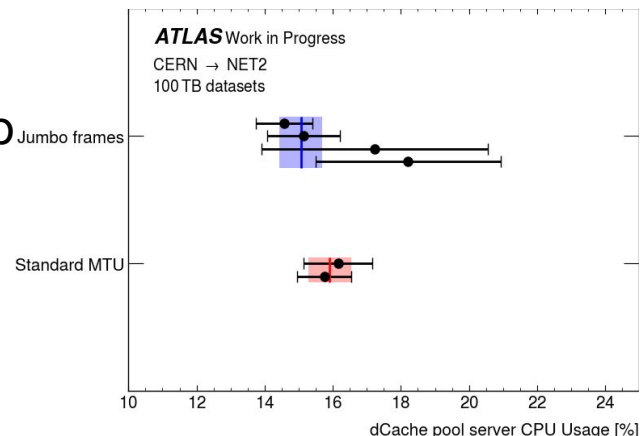
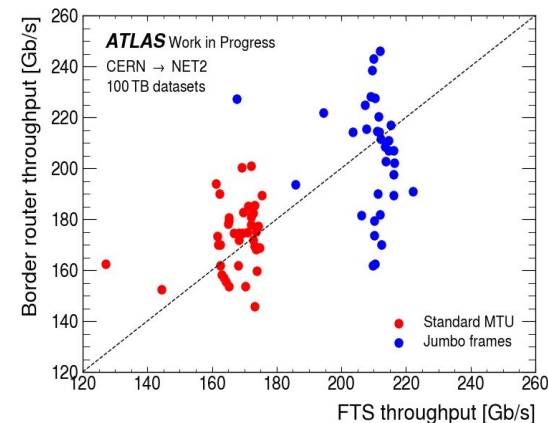
CERN-Lead[3]: CERN-P5/CERN, CERN/BNL, CERN/NET2, CERN/FNAL

USCMS-Lead[4]: FNAL/Florida, FNAL/UNL, FNAL/UCSD

# Latest Capabilities Mini-challenge: Jumbo Frames

## Key Takeaways of enabling Jumbo Frames:

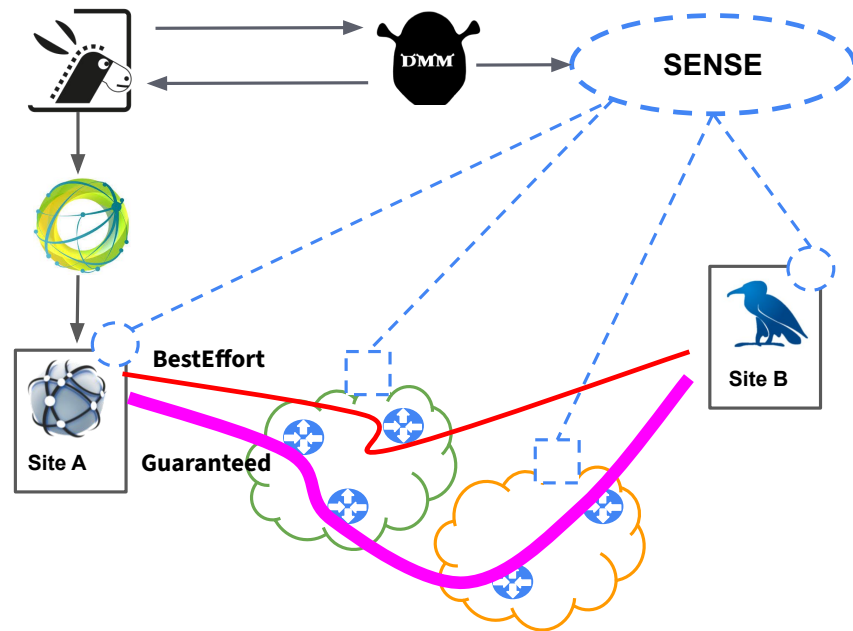
1. Improves throughput at high RTT
  - a. CERN to FNAL: ~20% improvement[3]
  - b. CERN to NET2: ~15% improvement[6]
  - c. CERN to BNL: no improvement apparently due to dCache internals (needs further investigation). More info here[5]
2. Not significant difference at low RTT
  - a. USCMS tests didn't show a significant difference[3]
  - b. CERN-P5 to CERN slightly worse (RTT: 1ms)[4]
3. CPU usage is lower on average but more variability is observed
  - a. See NET2 plots on the right[6]
4. Make sure to set **net.ipv4.tcp\_mtu\_probing=1**



# Latest Capabilities Mini-challenge: Rucio/SENSE

By enabling Rucio to use SENSE, we can:

- Use alternative network paths to transfer data between a given pair of sites
- Create bandwidth-guaranteed paths over the default routes for high-prio data transfers
- The above works at the Rucio rule granularity, allowing you to isolate and manage individual rules differently, according to their priority

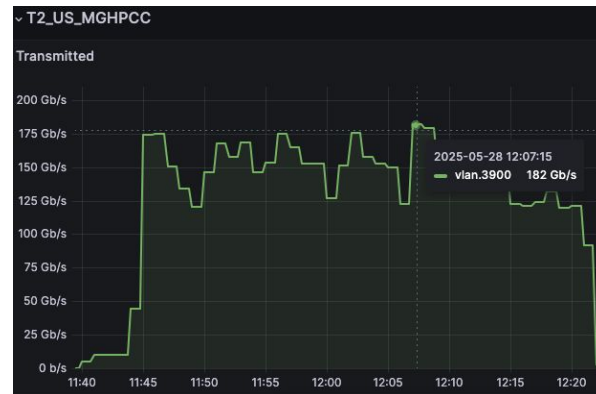
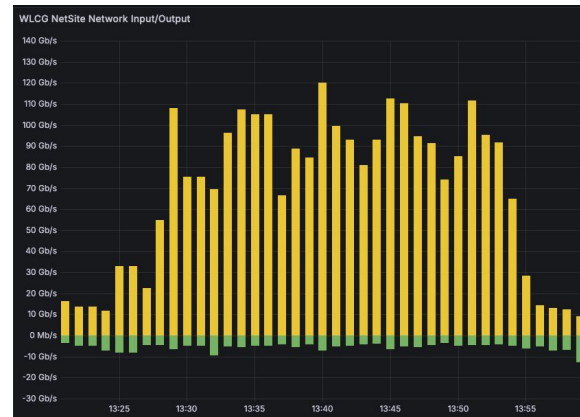


# Latest Capabilities Mini-challenge: Rucio/SENSE

This mini-challenge was presented in HTC25[8].

Its objectives were:

1. Show a Rucio/SENSE workflow as close to production as possible
  - a. We used Caltech and Nebraska
  - b. The test used both Site's prod-File System
  - c. Caltech used prod-DTNs
  - d. The data flow (top image) uses prod-network
  - e. The Rucio and FTS instances used are not-prod
2. Show a Rucio/SENSE high-throughput / high RTT workflow
  - a. In the bottom image we can see a Rucio/SENSE workflow between NET2 and SDSC reaching 180 Gbps at 70ms RTT in the SENSE testbed



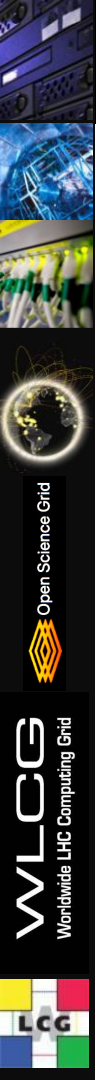
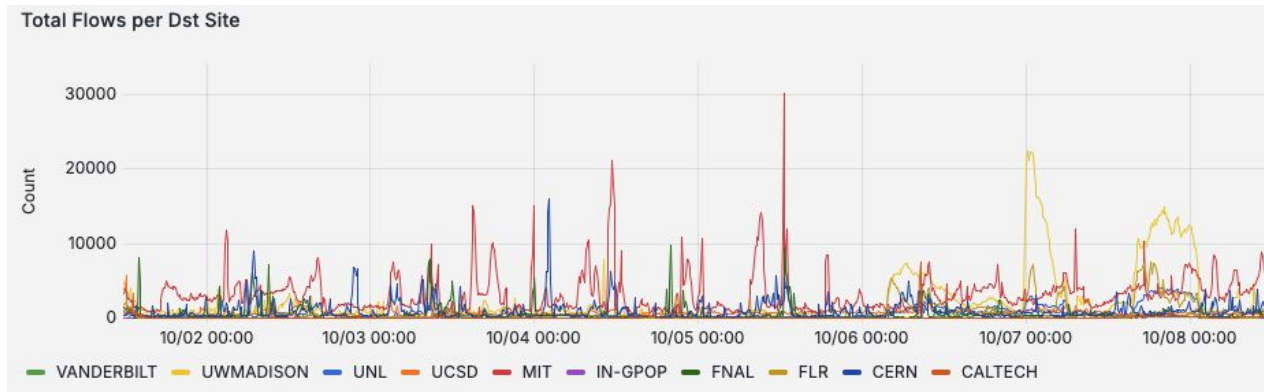
# Latest Capabilities Mini-challenge: Scitags

Fireflies are one type of scitags[9] that uses udp packets to **tag data flows**. These fireflies are collected by R&E Network providers and **capture information like the experiment and the activity** associated with the data flows among other data. The 2 objectives of this mini-challenge were:

## 1. Promote the support of scitags amongst the US CMS T2s

- During Feb 2025 we increase the number of T2 sites supporting scitags from 3 to 6.

**Note:** At the moment all USCMS T2 support scitags.



# Latest Capabilities Mini-challenge: Scitags

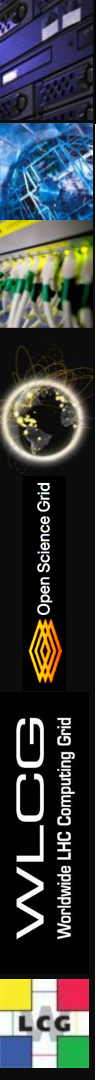
## 2. Validate the data produced (more info here[10])

We created a Rucio transfer of a single dataset between Wisconsin and UCSD. Then pulled the corresponding scitags data (script in background slides) and compare. Key Takeaways:

- Every file transfer produces 4 fireflies, 2 from each host involved
- Src and Dst inside the scitag do NOT represent the direction of the data flow but the origin of the firefly (src) and the other side of the connection (dst).
- 'Usage.received' and 'usage.sent' are larger than file size, because overhead is taken into account
- The sum of 'usage.received' and 'usage.sent' do not match, not sure why
- The flag: `--scitag=65` is parsed different in
  - `gfal-copy` (experiment\_id: 3, activity\_id:1) VS
  - `FTS` (experiment\_id: 1, activity\_id:12)

	syslog.hostname	state	esdb.src.org.short_name	esdb.dst.org.short_name	usage.sent	usage.received	netlink.rtt
0	g24n10.hep.wisc.edu	end	UWMADISON	UCSD	757792127	1442	80.542
1	xrd-2.t2.ucsd.edu	end	UCSD	UWMADISON	5993	757787322	82.212
2	xrd-2.t2.ucsd.edu	start	UCSD	UWMADISON	5614	1320	80.191
3	g24n10.hep.wisc.edu	start	UWMADISON	UCSD	6034	1418	80.081

*Example of the 4 scitags produced by a single file transfer*



## Section 2.2 Latest US **Capacity** Mini-challenges

The following is a summary of this[12] presentation in the last  
Open Technical Forum

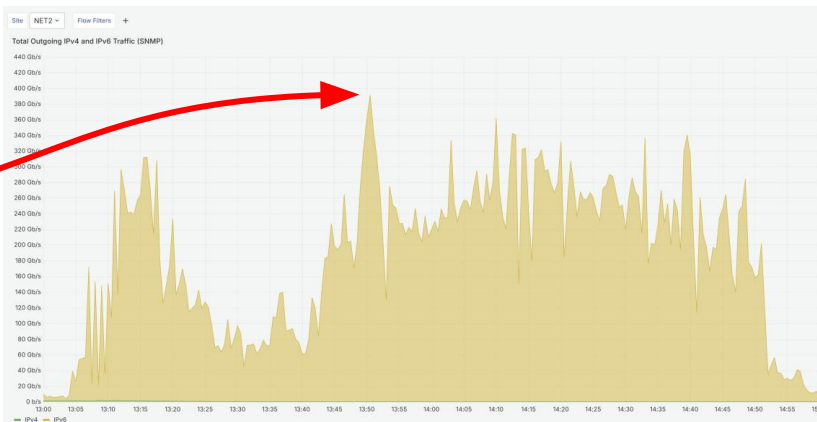
# Latest Capacity Mini-challenge

- Overall idea was to use the last two week's in August 2025 for testing.
- The first week would allow each experiment (USCMS/USATLAS) to “benchmark” each site individually
- The second would focus on **joint testing** trying to identify any bottlenecks.
- There were some slight differences on the first's week benchmark approach between USCMS and USATLAS
  - Tool used: [dc inject](#) vs Hiro's load-test tool
  - Targets: based on DC27 vs Max Available
- In general the idea was to use the T1 to benchmark the individual T2s and then a set (or all) T2s to benchmark the T1

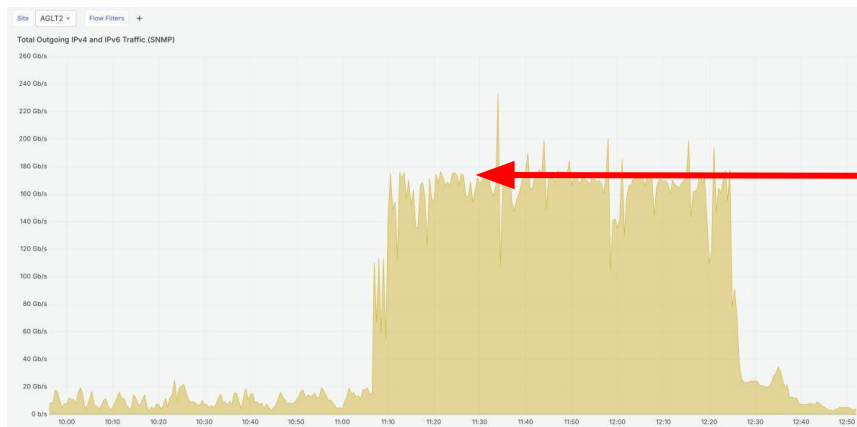


# Latest Capacity Mini-challenge: USATLAS Highlights

*BNL => NET2 Max:  
380 Gbps*



*BNL => AGLT2  
Sustained at: 180 Gbps*



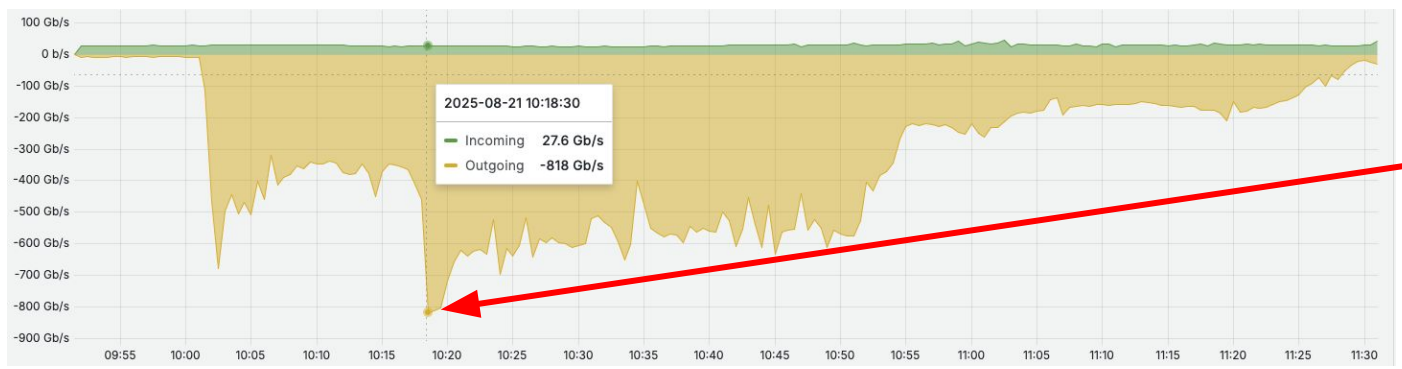
# Latest Capacity Mini-challenge: USATLAS Results

## USATLAS Capacity Tests Results: Fall 2024 vs Summer 2025

Site	Fall 2024	Summer 2025	2025/2024
AGLT2	150 (180)	180 (180)	<b>120%</b>
MWT2	200 (200)	250 (300)	<b>125%</b>
NET2	NP (10)	380 (400)	-
SWT2	30 (60)	55 (60)	<b>183%</b>
BNL	200 (1600)	200 (1600)	<b>100%</b>

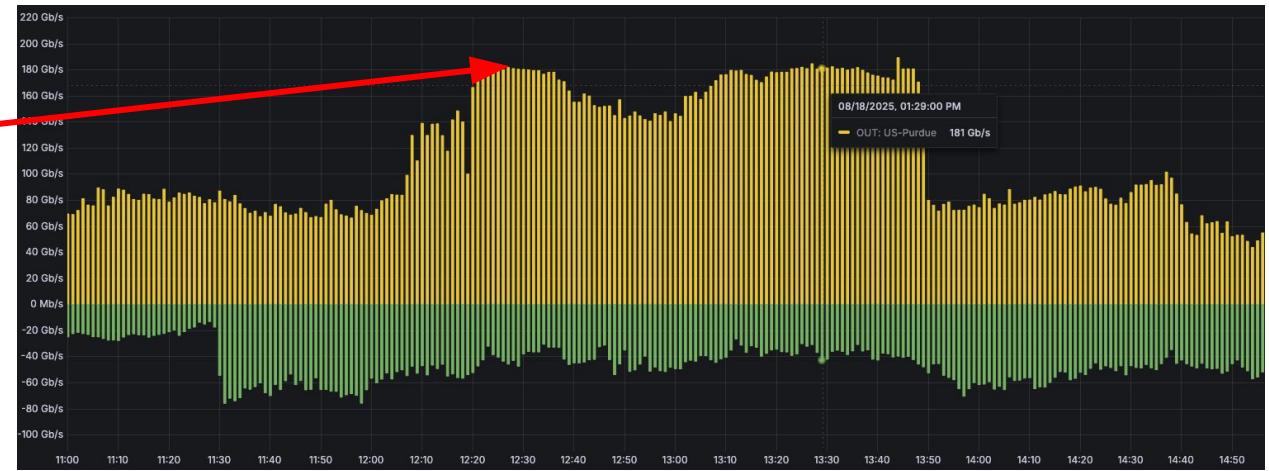
- The numbers in parenthesis are the physical wirespeed possible for each site
- The numbers in the table represent the **max peak write rate** to each site
- SWT2 Fall results only include only UTA, Summer results include both OU and UTA
- NET2 wasn't able to participate in the Fall test.

# Latest Capacity Mini-challenge: USCMS Highlights



*T2s => FNAL  
Max: 818 Gbps*

*Purdue => FNAL  
Avg 180 Gbps*



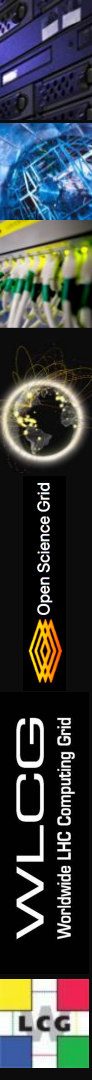
# Latest Capacity Mini-challenge: USCMS Results

## USCMS Capacity Tests Results Summer 2025

		Reads		Writes	
Site	Target	Max	Avg	Max	Avg
FNAL	800	350	270	818	636
Caltech	100	90	85	96	70
Florida	200	52	50	45	35
MIT	100	85	80	62	47
Nebraska	200	109	65	145	115
Purdue	150	189	180	120	70
UCSD	100	47	45	48	50
Vanderbilt	200	170	120	60	50
Wisconsin	200	120	65	86	83

**Max** is the maximum value observed either in the [WLGC](#) or the [Stardust](#) dashboard.

**Avg** is the maximum value sustained for at least 10 minutes



## USCMS Capacity Tests Results: Fall 2024 vs Summer 2025

Site	Writes Avg			Reads Avg		
	Fall 2024	Summer 2025	2025/2024	Fall 2024	Summer 2025	2025/2024
Florida	135	35	26%	100	50	50%
MIT	28	47	168%	65	80	123%
Nebraska	90	115	128%	80	65	81%
Purdue	60	70	117%	100	180	180%
Vanderbilt	80	50	63%	90	120	133%
Wisconsin	60	83	138%	100	65	65%

- The numbers in the above table represent the **max rate sustained for at least 10 min**
- The sites that did not participate in the Fall mini-challenge have been excluded
- Full USCMS report can be found here[13]

# USCMS/USATLAS Joint Test

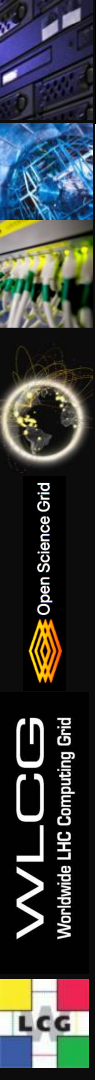
The plan for the second week was to jointly stress the network by generating the maximum traffic possible:

- **USATLAS:** BNL + NET2 => MWT2 + AGLT2 + SWT2
- **USCMS:** FNAL => T2\_US\_\*

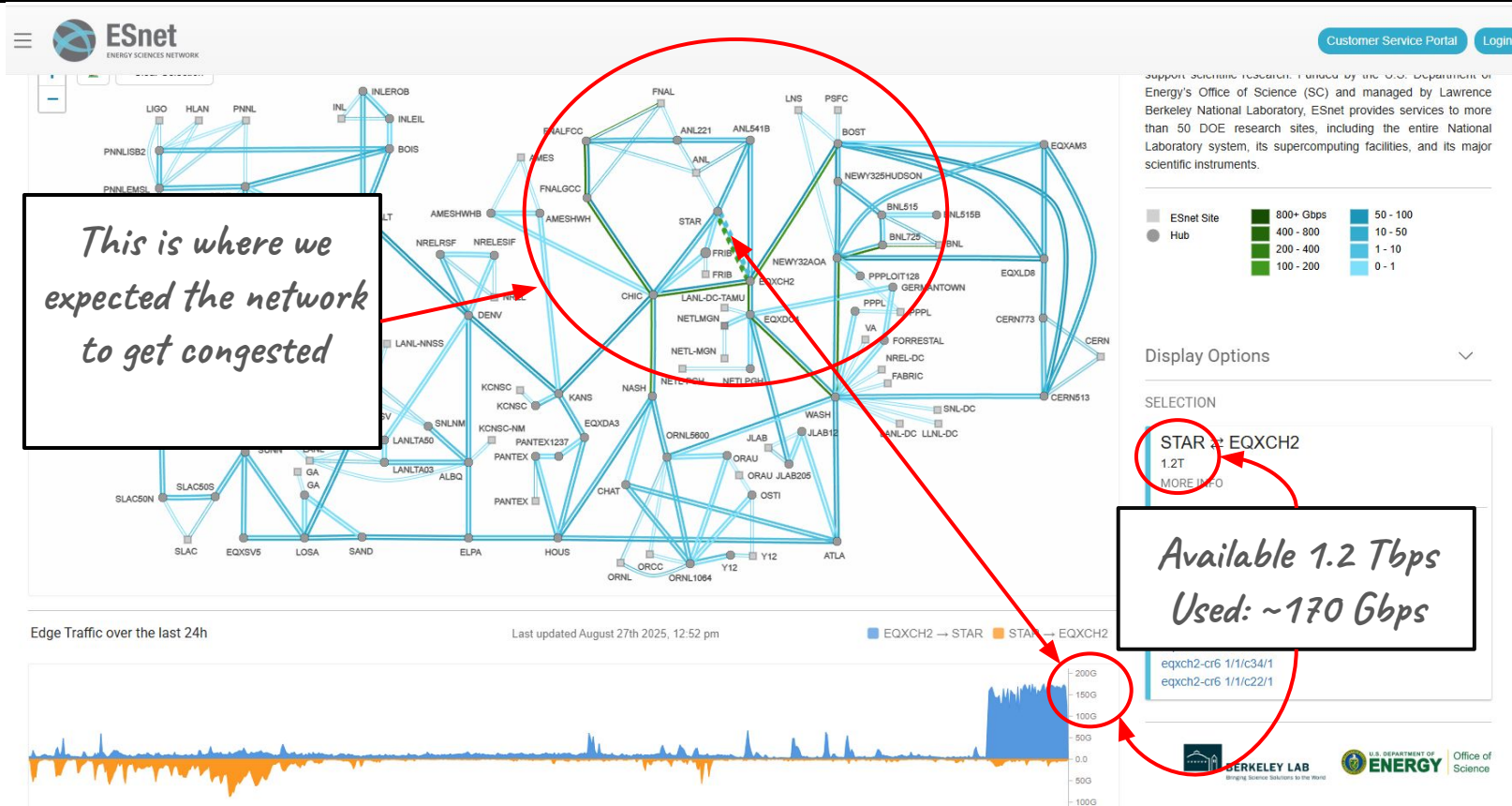
We wanted to identify places in the network where both experiments overlap. Using traceroute between the T1s and their respective T2s we found 2 ESnet routers shared by the experiments

*ESnet router	Sites passing by on their way to/from their respective T1
chic-cr6	CMS: caltech, nebraska, florida, ucsd, mit, vanderbilt ATLAS: mwt2, ou
eqxch2-cr6	CMS: mit ATLAS: mwt2, aglt2_msu, aglt2_um, mwt2_uc, ou

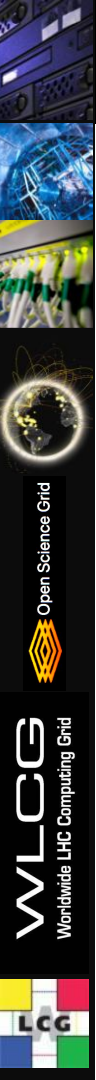
(\*) These locations can be monitored in the [my.es.net](http://my.es.net) web page



# ESnet Monitoring from August 27, 2025



<https://my.es.net/>



## Section 3. Issues and Lessons Learnt

# Issues

## General Issues

- Need consensus on the HL-LHC expected rates for the T2s
  - During DC24 some US T2 got  $O(10G)$  while they expected  $O(100G)$
  - From the US T2s PoV it's clear that 400 Gbps is the HL-LHC target
- Many Capabilities that we want to test but very little volunteers
  - The bottom-up approach only works when there's a sizable group behind a capability
- Need more people (apart from Diego and Hiro) that can run capacity tests
  - Basically: people with write permissions in all sites

# Issues (cont'd)

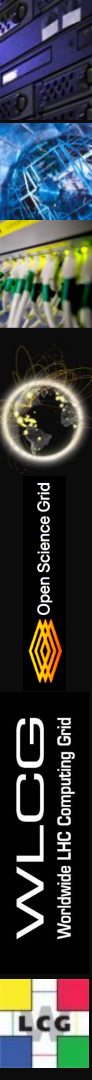
## Technical Issues

- FTS scaling (hopefully will be improved in FTSv4)
  - We loaded cms-fts3 with our first T2 test
- FTS Limits: hard to define and easy to overlook
  - For a single pair of sites e.g. FNAL  $\leftrightarrow$  UCSD there are 8 different limits to check
- FNAL's FTS needs more servers (in the works)
- We need Rucio quotas shielded from Rucio use
  - Increased quotas for the test ended up being used by Rucio
  - USCMS will try to use the 't2\_local\_users' account for this (not trivial to add people to it)
- Broken/Missing Monitoring
  - Our tests can only be as good as our monitoring

# Lessons Learnt

## About Capabilities

- Jumbo frames
  - can improve throughput at high RTT and are harmless otherwise
  - **net.ipv4.tcp\_mtu\_probing=1** to avoid issues with sites using regular frames
- Scitags (Fireflies)
  - are well supported in XRootD ( $\geq 5.6.2$ ), recently supported by dCache (11.0)[12]
  - easy to fetch ( $< 50$  lines of python), tricky to interpret
- SENSE
  - can give Rucio the ability of finding alternative paths and providing bandwidth guarantees for priority data flows
  - sites have started to support SENSE in production



# Lessons Learnt

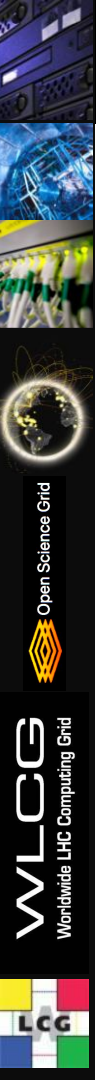
## About Capacity

- Most sites improved their throughput capacity from Fall 2024 to Summer 2025 (last 2 capacity tests)
- Some sites broke between mini-challenges
  - Do we need nano-challenges to make sure sites are at least as good as the previous test?
- The joint test didn't show any signs of network congestion
- dCache (proxy-mode) scalability needs to be studied
  - Issues mentioned both in the Jumbo Frames study and the USATLAS capacity tests
  - We have seen signs of it in NET2 within the SENSE testbed

# Lessons Learnt

## About running these tests (TODO List):

- Make sure Monitoring for the participant sites is up, running and validated
- Involve Data Management Operators early in the process
- Get everyone in the loop:
  - WLGC-DOMA, Experiments, DM Ops, FTS devs, Sites, USCMS/USATLAS Management
- Make sure to dial up the correct FTS limits



## Section 4. What Next?

# Next Mini-challenge

## DOMA Milestones and Deliverables – Year 8

ID	Description	Date
D4.9	60% of production transfers executed using tokens	Jan 2026/ Y3Q2
D4.X	(Mini-DC) Run a round of “capability tests” with at least 2 new capabilities	Jan 2026/ Y3Q2

- According to Brian’s presentation[15] from yesterday. The next Mini-challenge will happen early next year. Stay tuned!
- We need to select the new capabilities to try out and find the volunteers

# Help needed

Current list of capabilities we are interested in testing:

[https://drive.google.com/drive/folders/1Af7hWa0Zm30EuqsV1PbekSjb--gXAsVG?usp=drive\\_link](https://drive.google.com/drive/folders/1Af7hWa0Zm30EuqsV1PbekSjb--gXAsVG?usp=drive_link)

Cloud Storage Testing

*ongoing*

Datafile Size Testing

*help needed*

Network Monitoring Validity Testing

*very needed*

TCP Congestion Protocol Testing

*done (missing info)*

Traffic Visibility via SciTags Testing

*help needed*

Network Flow Monitoring Testing

*help needed*

Batch System Resource Consumption Testing

*ongoing?*

Netbird VPN Testing

*PofC done*

IPv6-Only Capability Testing

*help needed*

WireGuard Deployment and Use Testing

*included here*

~~Recent Linux Improvements that impact TCP Throughput.pdf~~

SENSE/Rucio Testing

*ongoing*

~~Copy of Capability Template~~

Host Network Tuning Optimization Testing

*highly desirable*

Tape System Metadata Scaling

*help needed*

If any of these caught your eye or want to test something along these lines please get in touch

# NANO-Challenge... really?

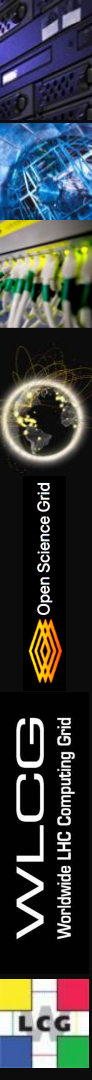
- There are a couple of USCMS sites that have expressed their wish for a re-test after getting a lower than expected result in the previous Mini-Challenge.
- If time permits we will run these before the end of the year

# What else?

... well that depends on the discussion next :)

Please jump here:

<https://docs.google.com/document/d/1OD8pE-I-nnjlhmTIFx7tQ9NLumxAWIYvzPNzazK3nZOE/edit?usp=sharing>



# References

[1] Original DC planning document:

<https://zenodo.org/records/5532452#.YXqMTC0RqqA>

[2] Mini-Challenges twiki:

<https://twiki.cern.ch/twiki/bin/view/LCG/DomaMiniChallenges>

[3] Jumbo Frames - CERN

<https://indico.cern.ch/event/1511535/contributions/6374829/attachments/3018021/5323446/JumboFramesTests.pdf>

[4] Jumbo Frames - USCMS

[https://docs.google.com/presentation/d/13Ya-4rQvWFjHQkWxwACRoY3sbNzHH\\_IF6Zh3\\_ixd76A/edit?slide=id.g3332a174d21\\_0\\_74#slide=id.g3332a174d21\\_0\\_74](https://docs.google.com/presentation/d/13Ya-4rQvWFjHQkWxwACRoY3sbNzHH_IF6Zh3_ixd76A/edit?slide=id.g3332a174d21_0_74#slide=id.g3332a174d21_0_74)

[5] Jumbo Frames- BNL

<https://docs.google.com/document/d/1P2pNG0rSW5MZhOFDaiDwjijwtnLZsEaCulCBQdYdHOM/edit?tab=t.0#heading=h.q8dgotne59tj>

[6] Jumbo Frames - NET2

[https://docs.google.com/presentation/d/1oqb3\\_BdCF6mO2OCn9PVXDqUQ1loyAvbrV1oBu9Or8bo/edit#slide=id.p](https://docs.google.com/presentation/d/1oqb3_BdCF6mO2OCn9PVXDqUQ1loyAvbrV1oBu9Or8bo/edit#slide=id.p)

[7] Rucio/SENSE

[https://www.epj-conferences.org/articles/epjconf/pdf/2024/05/epjconf\\_chep2024\\_01009.pdf](https://www.epj-conferences.org/articles/epjconf/pdf/2024/05/epjconf_chep2024_01009.pdf)

# References (cont'd)

[8] Capability Challenge: SENSE

[https://agenda.hep.wisc.edu/event/2297/contributions/34309/attachments/10375/13341/SENSE\\_Rucio%20-%20HTC25.pdf](https://agenda.hep.wisc.edu/event/2297/contributions/34309/attachments/10375/13341/SENSE_Rucio%20-%20HTC25.pdf)

[9] Scitags

<https://scitags.org/>

[10] Scitags validation report

[https://docs.google.com/document/d/14es\\_UeJX7reOt2B7PfrftqRmL4A1L1i68\\_2LWaMD02A/edit?usp=sharing](https://docs.google.com/document/d/14es_UeJX7reOt2B7PfrftqRmL4A1L1i68_2LWaMD02A/edit?usp=sharing)

[11] RNTWG/Scitags presentation at LHCOPN/LHCONE meeting #55, October 9th

[https://indico.cern.ch/event/1534556/contributions/6653197/attachments/3150644/5594871/Scitags%20OPN%20meeting%20\(fall%202025\).pdf](https://indico.cern.ch/event/1534556/contributions/6653197/attachments/3150644/5594871/Scitags%20OPN%20meeting%20(fall%202025).pdf)

[12] Results of Mini Data Challenges in the U.S. (Summer Capacity mini-challenge)

<https://indico.cern.ch/event/1562124/contributions/6680945/attachments/3133245/5558850/Results%20of%20Mini%20Data%20Challenges%20in%20the%20U.S..pdf>

[13] USCMS Summer 2025 Capacity Mini-challenge

[https://docs.google.com/document/d/1-P185vz3qQyU70fiBq0\\_JcKWjRZjzXd164XibrZguE/edit?usp=sharing](https://docs.google.com/document/d/1-P185vz3qQyU70fiBq0_JcKWjRZjzXd164XibrZguE/edit?usp=sharing)

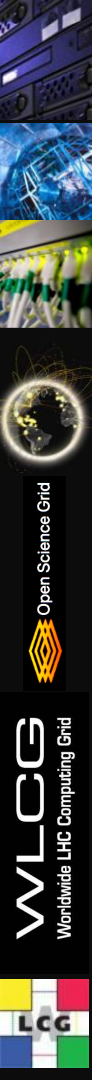
[14] dCache 11.0 Release notes

<https://www.dcache.org/old/downloads/1.9/release-notes-11.0.shtml>

# References (cont'd)

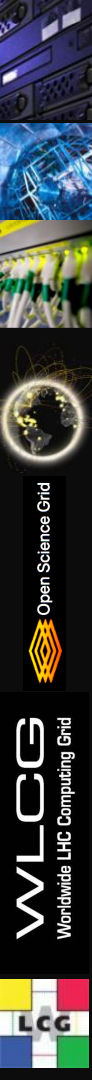
[15] Brian's presentation (DOMA deliverables)

<https://indico.cern.ch/event/1576135/contributions/6729160/attachments/3150545/5594672/DOMA-Area-Overview.pdf>

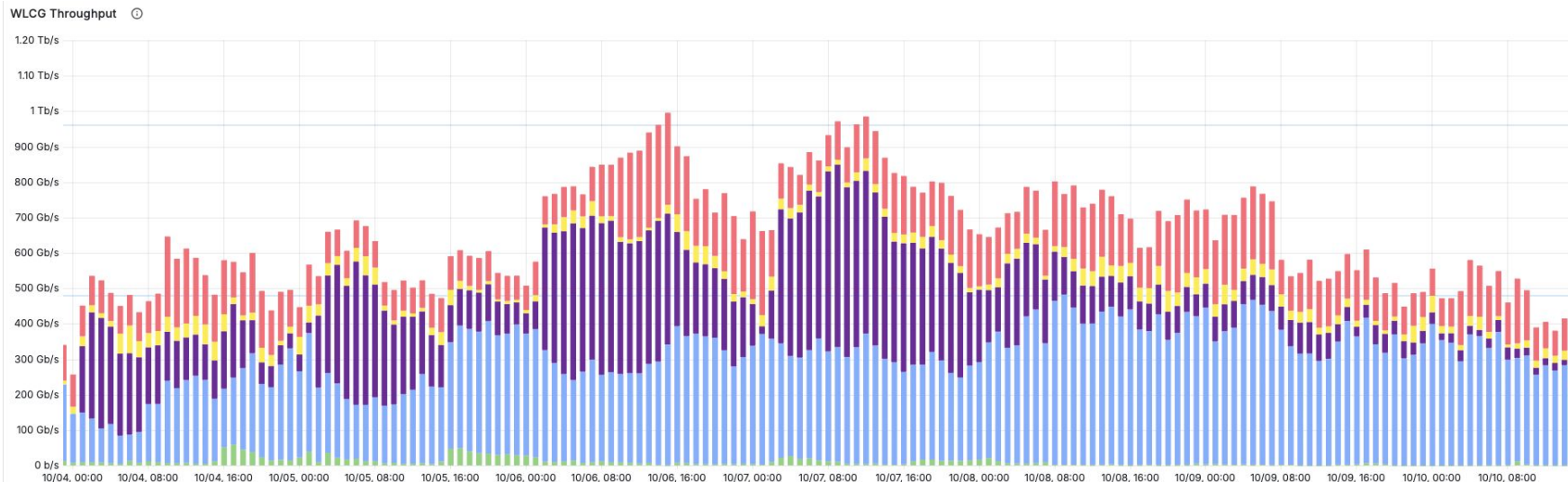


# Background Slides

Background Slides start here



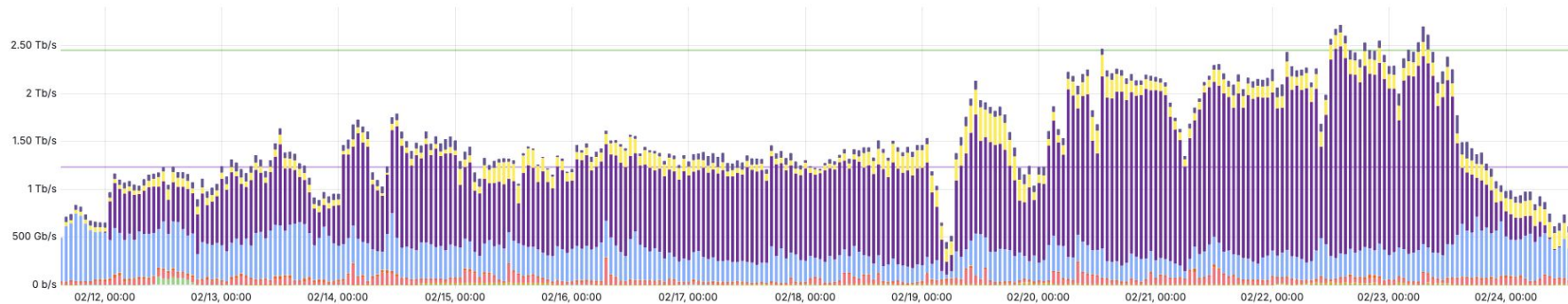
# DC21 Full plot



[https://monit-grafana.cern.ch/d/ZqU5ugjMz/fts-status-board?from=1633327200000&orgId=20&to=1633903200000&var-activity=Data+Challenge&var-bin=1h&var-dst\\_country=All&var-dst\\_experiment\\_site=All&var-dst\\_rse=All&var-dst\\_site=All&var-dst\\_tier=All&var-fts\\_server=All&var-group\\_by=vo&var-protocol=All&var-src\\_country=All&var-src\\_experiment\\_site=All&var-src\\_rse=All&var-src\\_site=All&var-src\\_tier=All&var-staging=All&var-vo=All&viewPanel=111](https://monit-grafana.cern.ch/d/ZqU5ugjMz/fts-status-board?from=1633327200000&orgId=20&to=1633903200000&var-activity=Data+Challenge&var-bin=1h&var-dst_country=All&var-dst_experiment_site=All&var-dst_rse=All&var-dst_site=All&var-dst_tier=All&var-fts_server=All&var-group_by=vo&var-protocol=All&var-src_country=All&var-src_experiment_site=All&var-src_rse=All&var-src_site=All&var-src_tier=All&var-staging=All&var-vo=All&viewPanel=111)

# DC24 Full plot

WLCG Throughput 



<https://monit-grafana.cern.ch/d/d3543f53-950b-4a60-b353-16611bf7f5f7/dc-2024-draft?orgId=20&from=1707692400000&to=1708815599000>

# Scitags: example code to pull scitag data

```
import requests
import json

headers = {'Content-Type': 'application/json'}
query='meta.esdb.dst.org.short_name:UCSD AND meta.esdb.src.org.short_name: UWMADISON AND meta.activity_name:"Data Challenge"'
#March 5th, 2:30pm (pacific) in milliseconds
start=1741213813000
fields=['meta.activity_name', 'meta.application', 'meta.syslog.hostname','meta.src_ip',
'meta.state', 'meta.esdb.dst.org.short_name', 'meta.esdb.src.org.short_name']
fields = fields +['meta.ip_version', 'meta.dst_ip', 'values.netlink.rtt', 'values.usage.sent', 'values.usage.received', 'duration', 'start', 'end']
data={
    "size":10000,
    "query":{
    "bool":{
    "filter":[
        {"range":{"start":{"gte":start, "lte":(start+3600000)}},
        {"query_string":{"analyze_wildcard":"true", "query":query}}
    ]
    }
    },
    "_source":fields
}

data_string = json.dumps(data)
response = requests.post('https://e1.gc1.prod.stardust.es.net:9200/stardust_firefly/_search', headers=headers, data=data_string)

d = response.json()
records_list=[]
total_received=0
for record in d['hits']['hits']:
    total_received+=record['_source']['values']['usage']['received']

print(str(total_received/(10**9))+ " GB")
```