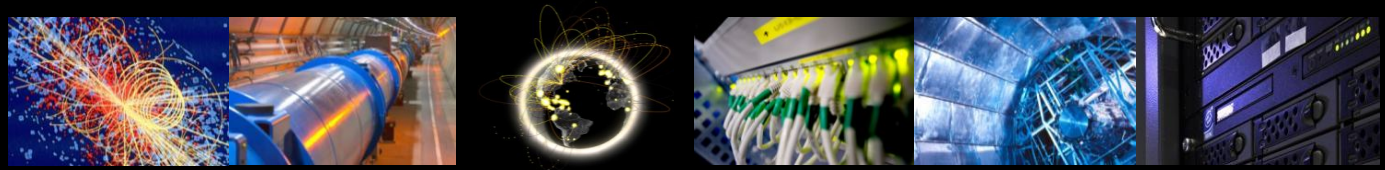
Results from US Mini-Challenges

Diego Davila / University of California San Diego, Garhan Attebury / University of Nebraska,
Shawn McKee / University of Michigan, Hiro Ito / BNL

WLCG DOMA General Meeting (https://indico.cern.ch/event/1652743/)

February 25, 2026



WLCG Worldwide LHC Computing Grid

Open Science Grid

The Last Year of Challenges

There have been a number of mini-challenges during the last half year (see WLCG DOMA Twiki)

- Earlier in 2025 tested jumbo frames and benchmarked capacity
- August 2025: mini-capacity challenge in the US
  - This was another benchmark in our progress towards DC27
- September 2025: shutdown of IPv4 on LHCOPN for BNL and FNAL
  - No issues observed. IPv4 traffic shifted to LHCONE.

We need to continue planning and executing to make progress

Mini-Challenges in 2026 So Far

USATLAS and USCMS planned and are executing 3 mini-capability challenges this month:

- IPv4 Blackout on LHCONE
  - AGLT2, NET2 for ATLAS
  - UNL, Purdue for USCMS
- SENSE/Rucio
  - MWT2, NET2 for ATLAS
  - FNAL, UCSD for USCMS
- Host-tuning Evaluation
  - AGLT2, MWT2 for ATLAS
  - FNAL, UNL, UCSD, Vanderbilt for USCMS

We will provide brief updates on each.

SENSE/Rucio Mini-Challenge

Rucio/SENSE Mini-challenge

- Our focus with this Mini-Challenge was to show the progress of each experiment towards production
- On USATLAS, the major developments are:
  - **The addition of MWT2 (UChicago) to our testbed**
    - Very Devy: Software Router (capped at 100 Gbps - needs a LAG)
    - XRootD cluster (UC AF K8s) with 3 100G nodes but little storage (using Blackhole plugin for infinite space)
    - Currently planning the deployment of dCache proxy with access to Prod (no data)
  - **Show the usage of NET2 Storage System (prod*)**
    - We use a set of 3 dCache proxies load balanced by HAproxy (NRP K8s) in front of their production storage. This was working great until it wasn't :(when the storage servers are overloaded with regular ATLAS operations, some SENSE transfers go so slow that end up timing out (under investigation; but not a SENSE problem)
  - **The new ATLAS/DMM instance**
    - Deployed (NRP K8s) and currently under testing

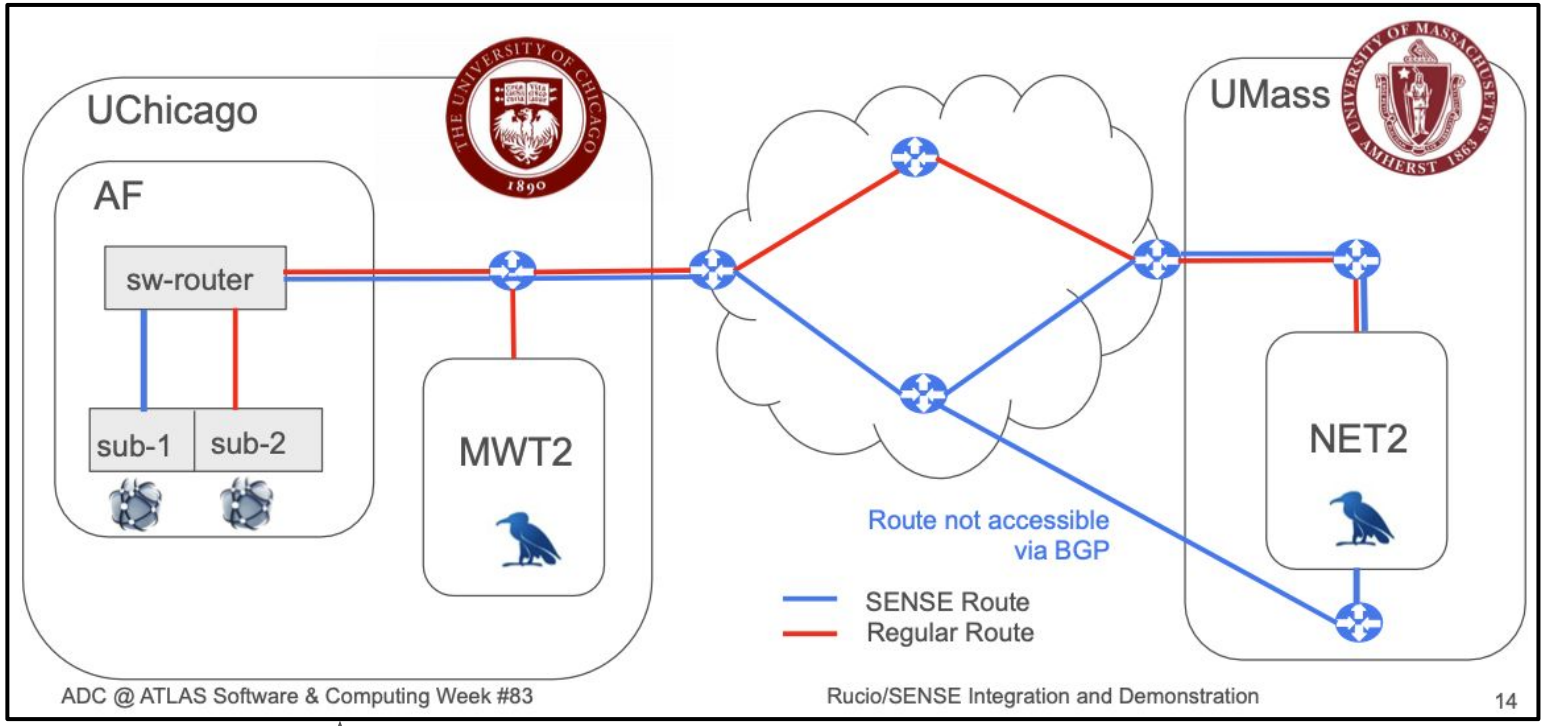
(*) In all cases we can use the storage system but do not have access to production data.

Rucio/SENSE Mini-challenge (cont'd)

- On USCMS, the major developments are:
  - **The addition of UCSD as pre-production**
    - Using our new DTNs and Production Storage System(*)
      - **For the test show here we used the Blackhole plugin given the current limitation of Ceph (more in the Host Tuning section)
    - Using it to test different SENSE supported DTN models (e.g. br vs k8s)
  - **Use of FNAL Storage System (prod*)**
    - HAproxy load balancing 2 dCache proxies with access to the Production Storage System

(*) In all cases we can use the storage system but do not have access to production data.

Rucio/SENSE (USATLAS)



I stole this slide from Rafael

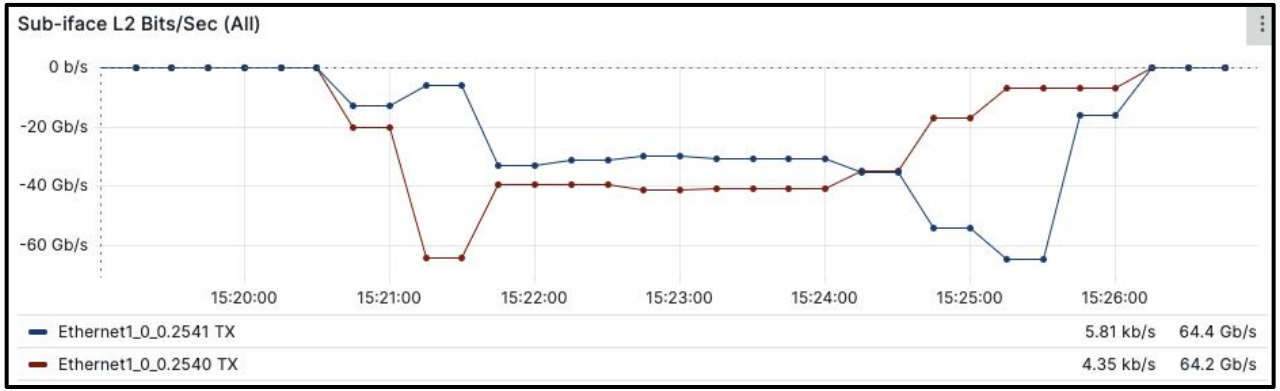
Rucio/SENSE (USATLAS) - FDT flows

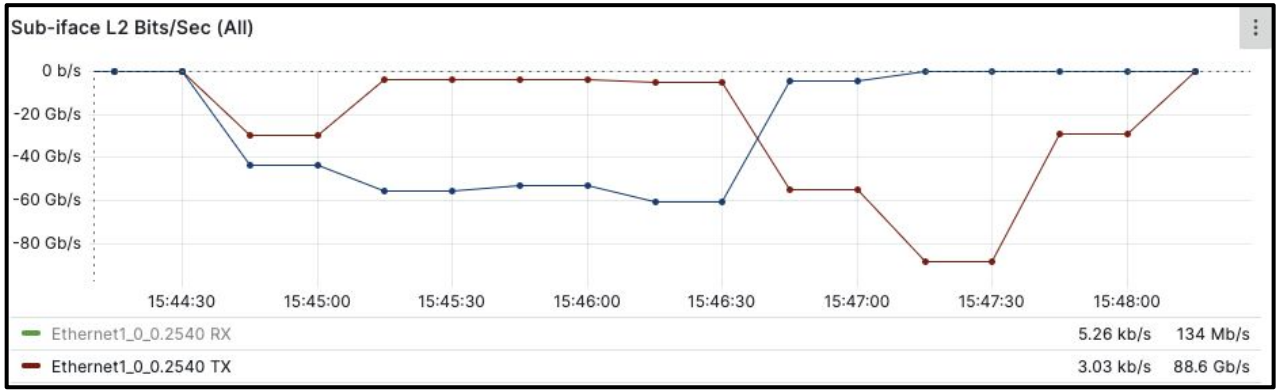
FDT: 80 Gb/s circuit, 60 Gb/s requested for high-priority transfer

ADC @ ATLAS Software & Computing Week #83 Rucio/SENSE Integration and Demonstration 15

... This one too

Rucio/SENSE (USATLAS) - FTS flows

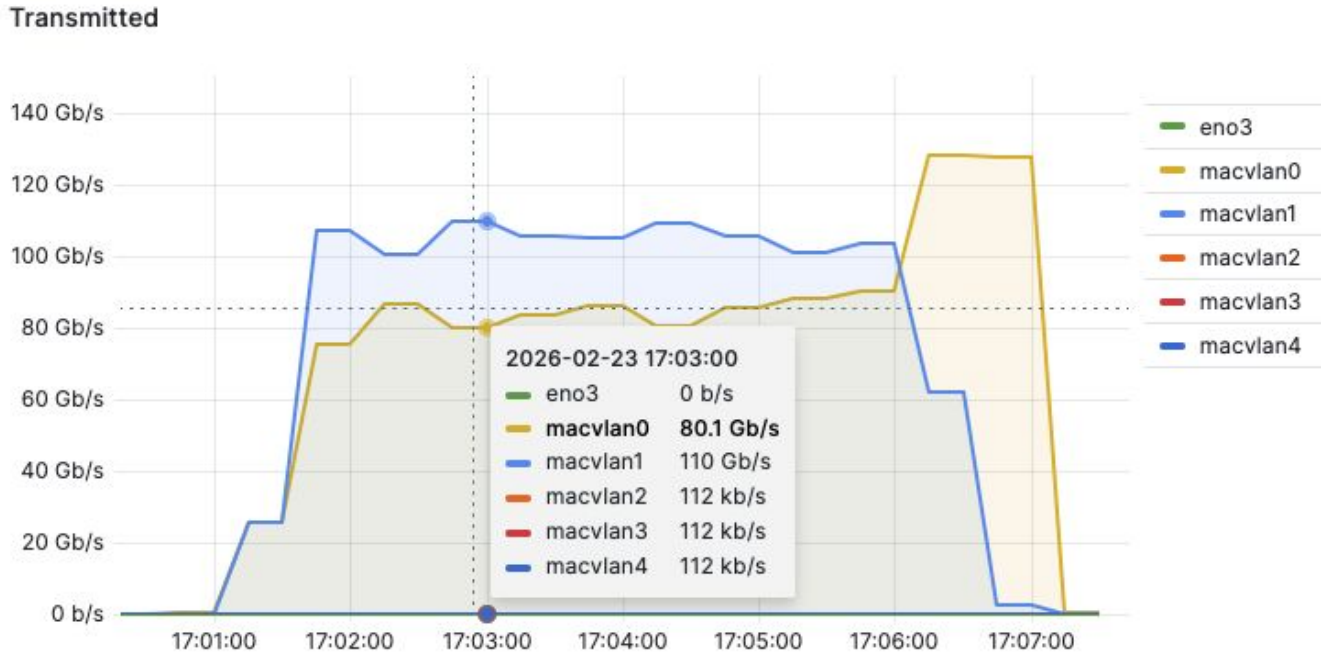




Two identical FTS flows competing for 100 Gbps available bandwidth

Regular (top) vs SENSE (bottom)

Rucio/SENSE (USCMS) FTS test FNAL -> UCSD



Two identical FTS flows managed by SENSE

Blue has priority over Yellow.

The Max. available bandwidth is 200 Gbps

Source	Destination	VO	Submitted	Active	Staging	S.Active	Archiving	Finished	Failed	Cancel	Rate (last 1h)	Thr.
davs://cmssense6-origin-2841-1.fnal.gov	davs://redir-f012.t2.ucsd.edu	cms	–	–	–	–	–	900	–	–	100.00 %	3084.43 MiB/s
davs://cmssense6-origin-2842-1.fnal.gov	davs://redir-f018.t2.ucsd.edu	cms	–	–	–	–	–	900	–	–	100.00 %	3901.97 MiB/s
			0	0	0	0	0	1800	0	0	100.00 %	–

LHCONE IPv4 Blackout Mini-Challenge

IPv4 Blackout on LHCONE

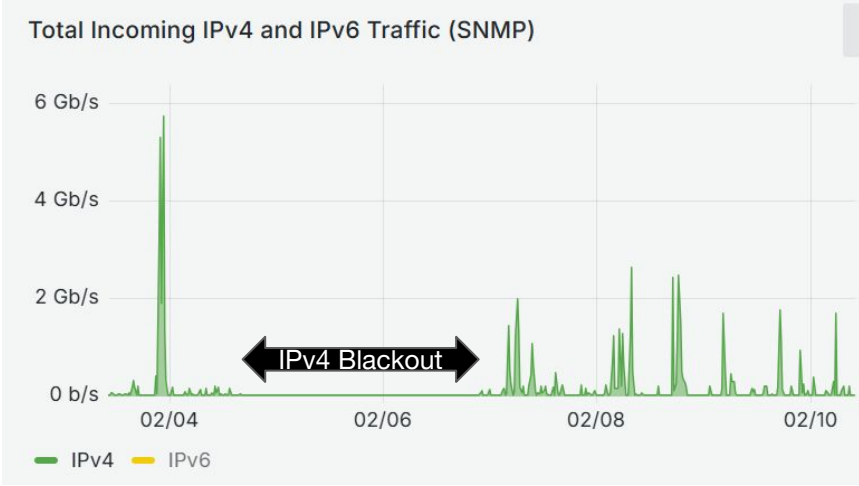
Goal: To evaluate the impact for sites from disabling IPv4 on LHCONE.
Participating sites: AGLT2+NET2 (ATLAS) and UNL+Purdue (USCMS)

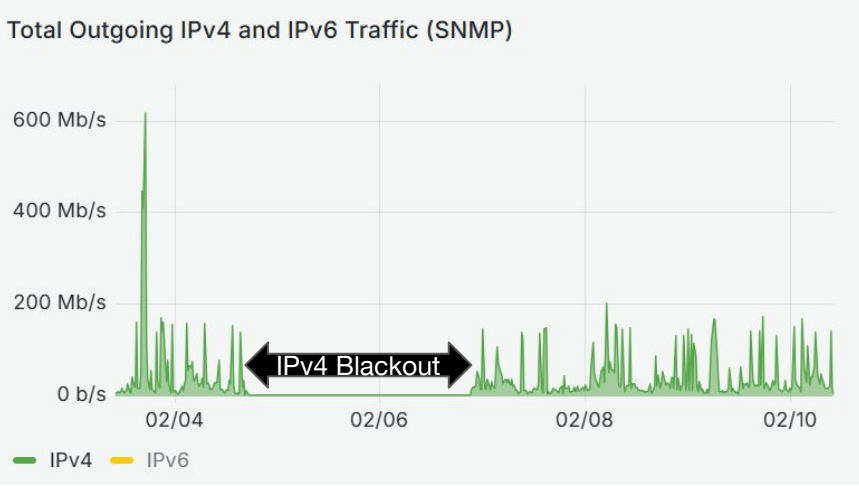
Purpose: To give our sites confidence in moving forward towards IPv6-only and reduce the use of IPv4 where possible.

Very simple: We shutdown IPv4 peering OR IPv4 network announcements into LHCONE for 2 days between February 4th-6th, 2026.

As expected, we saw **no problems**. Any IPv4 traffic that would have reached our sites via LHCONE went via R&E networks.

USATLAS: Example AGLT2 LHCONE IPv4 Traffic





USATLAS: Example AGLT2 IPv4 Routing Change

One example of shutting down IPv4 on LHCONE at AGLT2:

LHCONE IPv4 Enabled

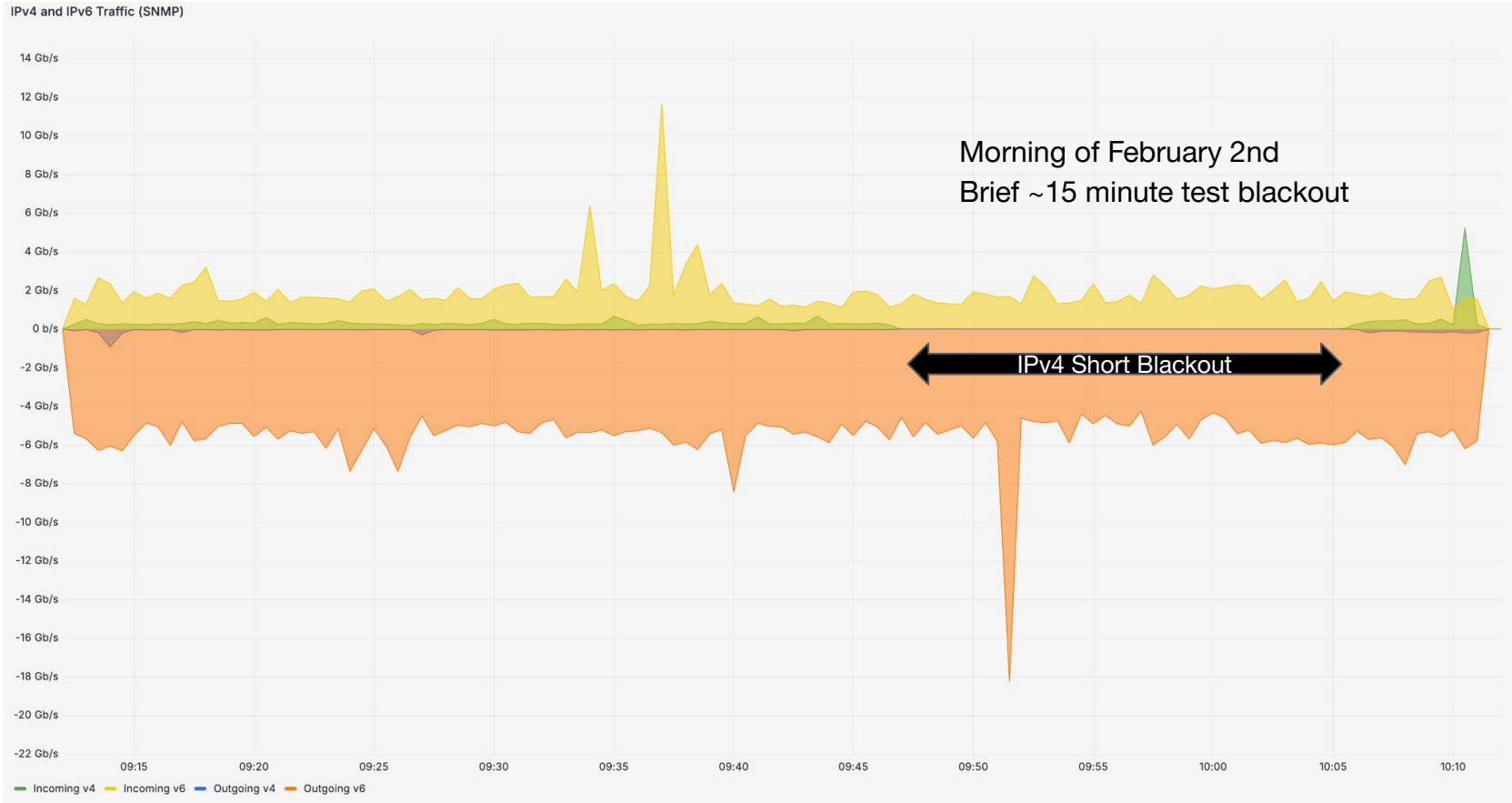
traceroute to etf-04.cern.ch (IPv4 on for LHCONE)
1 aglt2-rtr-2.aglt2.org (192.41.230.5)
2 192.41.233.162 (192.41.233.162)
3 192.41.233.150 (192.41.233.150)
4 esnet-**lhc1**-a-aglt2.es.net (198.124.80.53)
5 bost-cr6--eqxld8-bb-a.igp.es.net (134.55.57.48)
6 **eqxdc4-cr6--wash-bb-e.igp.es.net (134.55.56.18)**
7 **newy32aoa-cr6--eqxld8-bb-c.igp.es.net (134.55.57.104)**
8 **eqxld8-cr6--cern773-bb-e.igp.es.net (134.55.58.14)**
9 cern773-cr6--cern513-bb-c.igp.es.net (134.55.57.0)
10 esnet-**lhc1**-b-cern.es.net (198.124.80.69)
11 etf-04.cern.ch (128.142.196.141)

LHCONE IPv4 Disabled

traceroute to etf-04.cern.ch (IPv4 off for LHCONE)
1 aglt2-rtr-2.aglt2.org (192.41.230.5)
2 192.41.233.164 (192.41.233.164)
3 chic-cr6--aglt2_se-120.es.net
4 eqxch2-cr6--eqxdc4-bb-c.igp.es.net (134.55.57.174)
5 **eqxdc4-cr6--wash-bb-e.igp.es.net (134.55.56.18)**
6 eqxam3-cr6--cern513-bb-d.igp.es.net (134.55.57.118)
7 **newy32aoa-cr6--eqxld8-bb-c.igp.es.net (134.55.57.104)**
8 **eqxld8-cr6--cern773-bb-e.igp.es.net (134.55.58.14)**
9 esnet-**lhc1**-b-cern.es.net (198.124.80.69)
10 cern-**lhc1**-b-esnet.es.net (198.124.80.70)
11 etf-04.cern.ch (128.142.196.141)

NET2 saw the same type of behaviour and also saw no problems. Hopefully other sites are more comfortable with switching off IPv4 on LHCONE. WLCG has a **goal** of doing this before HL-LHC.

USCMS: Example UNL LHCONE IPv4 Traffic



USCMS: Example UNL IPv4 Routing Change

Example of shutting down IPv4 at UNL

LHCONE IPv4 Enabled

HOST: red-xfer1.unl.edu
1. AS7896 hcc-shor-m1428-vl20.unl.edu (129.93.239.190)
2. AS7896 hcc-shor-m1436-vl48.unl.edu (129.93.241.33)
3. **AS291 esnet-lhc1-unl.es.net (198.124.80.33)**
4. **AS??? ???**
5. **AS??? ???**
6. **AS??? ???**
7. **AS20965 eex-esnet-lhcone-rt0.lon.nl.geant.net.geant.net (62.40.126.151)**
8. **AS20965 eex-esnet-lhcone-gw-rt0.lon.nl.geant.net (62.40.126.150)**
9. AS20965 janet-lhcone-gw.lon2.uk.geant.net (62.40.126.72)
10. AS20965 janet-lhcone.lon2.uk.geant.net (62.40.126.73)
11. AS786 ae26.londpg-ban1.ja.net (146.97.35.234)
12. AS786 lhcone-londpg-imperial.ja.net (146.97.140.26)
13. AS786 sdmz-dcsl-rt2-bdr-rt2-lhcone.net.ic.ac.uk (192.195.105.181)
14. AS786 wm09.grid.hep.ph.ic.ac.uk (146.179.233.252)

LHCONE IPv4 Disabled

HOST: red-xfer1.unl.edu
1. AS7896 hcc-shor-m1428-vl20.unl.edu (129.93.239.190)
2. AS7896 hcc-shor-m1436-vl48.unl.edu (129.93.241.33)
3. AS7896 nu-wan-city-et-0-1-1-v23.unl.edu (129.93.6.69)
4. **AS??? fourhundredge-0-0-0-10.1410.core2.kans.net.internet2.edu (163.253.5.24)**
5. **AS??? fourhundredge-0-0-0-0.4079.core1.chic.net.internet2.edu (163.253.2.28)**
6. **AS??? fourhundredge-0-0-0-21.4079.core2.chic.net.internet2.edu (163.253.1.95)**
7. **AS??? fourhundredge-0-0-0-6.4079.core2.eqch.net.internet2.edu (163.253.2.75)**
8. **AS??? fourhundredge-0-0-0-0.4079.core2.clev.net.internet2.edu (163.253.2.16)**
9. **AS??? fourhundredge-0-0-0-1.4079.core1.alba.net.internet2.edu (163.253.1.20)**
10. **AS??? fourhundredge-0-0-0-0.4079.core1.bost2.net.internet2.edu (163.253.2.173)**
11. **AS??? 198.71.45.233**
12. AS20965 lag-8-0.rt0.lon2.uk.geant.net (62.40.98.106)
13. AS20965 jisc-p1b-gw.rt0.lon2.uk.geant.net (62.40.125.58)
14. AS786 ae26.londpg-ban1.ja.net (146.97.35.234)
15. AS786 imperial.ja.net (146.97.140.18)
16. AS786 sdmz-dcsl-rt2-bdr-rt2.net.ic.ac.uk (192.195.105.171)
17. AS786 wm09.grid.hep.ph.ic.ac.uk (146.179.233.252)

Again, hopefully other sites are more comfortable with switching off IPv4 on LHCONE. WLCG has a **goal** of doing this before HL-LHC.

Turn off IPv4 for LHCONE Persistently for Sites?

In presenting the IPv4 blackout to ATLAS during software & computing week, the question was raised “Why did you turn it back on?”

- Main answer: because that was the plan

However, it is a good point. We are moving towards IPv6-only (single stack again) and the testing showed no issues.

Plan for AGLT2 is to announce a shutdown for ~2 weeks from now (March 10), at which point we turn off IPv4 announcements into LHCONE. (R&E IPv4 still supported). Other sites may wish to do the same thing.

We will have to watch longer term for issues…

Host Tuning Mini-Challenge

Host-Tuning Evaluation

Quick summary: we want to test ESnet Fasterdata recommendations on our production storage hosts.

Goal: Improve the throughput and/or reduce the load per transfer, decreasing time for transfers & improving net use.

As part of the upcoming perfSONAR upgrade campaign, we created a fasterdata-tuning.sh (bash) script to implement the recommendations from the Fasterdata web pages.

The WLCG mini-capability challenge for host-tuning was done Feb 2-3 and involved USATLAS and USCMS sites.

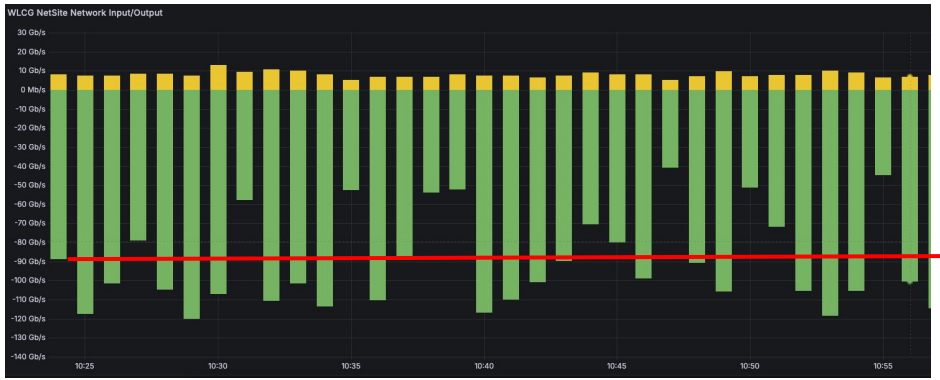
USCMS: Data Transfer Testing Details

- Participant sites: FNAL, UNL, Vanderbilt and UCSD
  - FNAL: Only a subset of 30 (out of 150) dCache servers are used in this test.
- FNAL is always used as Source.
- We carried out 2 types of tests: **Iperf3** (multi-threaded) and **FTS** transfers
- The iperf3 tests were done between single nodes and using 32 streams
- The set of files used in the FTS were selected in a way that they target all the 30 nodes, in FNAL, evenly
- Both iperf3 and FTS test were executed before and after tuning the systems on all sites.

Full report here:
https://docs.google.com/document/d/1nmjQaqjSLzZngKvEARxfbQM7PCeL8riOhtooxlZ49AQ/edit?usp=sharing

USCMS: UNL plots

60 Gbps - untuned



90 Gbps - tuned

USCMS Results

Destination site	iperf3		FTS	
	untuned	tuned	untuned	tuned
UNL	65	97	60	90
Vanderbilt	24	22	70	65
UCSD	21	95	55	55

Table 1. Summary of Results (Units are Gbps)

Summary of the summary

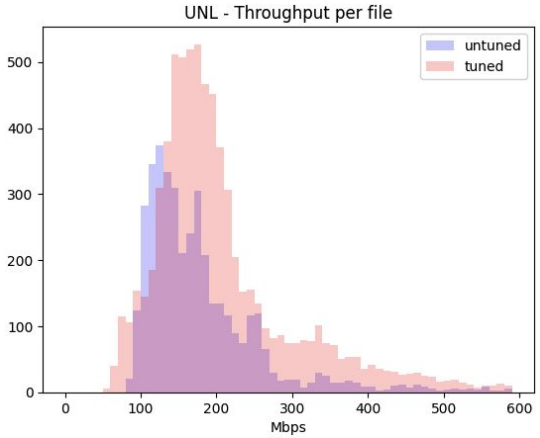
Better / Better

Worse / Worse?

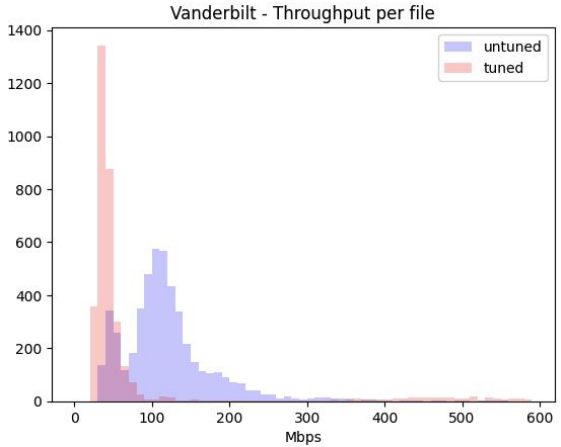
Better / Same

Quick Conclusion: Overall the tuning seems to improve at best and do little harm at worst

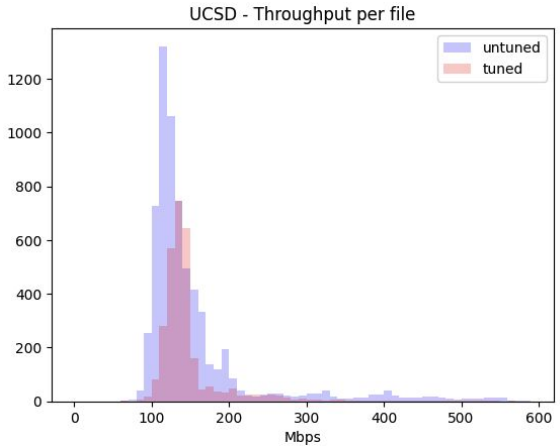
USCMS: Throughput per file (input data: FTS records)



Better



Definitely Worse



Same

... also maybe something to look at

USCMS: About Storage, the UCSD case

- At UCSD we have 400 Gbps connectivity to ESnet
- We used 7 new DTNs (all 100G connected) exclusively for this test
- **Iperf3 went from 20 to 90 (untuned vs tuned) in a single DTN**
  - If only UCSD is tuned the result is still 20 Gbps (i.e. no improvement)
- **FTS stayed the same at 55 Gbps**
- During the SENSE (see Rucio/SENSE slide above) we managed to write at 200 Gbps to UCSD
  - In that test we got rid of CEPH and wrote to /dev/null using the Blackhole plugin
- It is obvious that our bottleneck is at the Storage System level
  - Upgrades coming soon…
- No matter how well tuned our DTNs are, if the Storage is slow the overall System will be slow

USATLAS: Data Transfer Testing Details

The tests used the production data set and their associated files are larger than 4GB on average.

Files are read from and write to the AGLT2/MWT2 production storage using BNL FTS.

First day was running data transfers both directions between AGLT2 and MWT2 in the “as built” configuration

Second day was repeating the previous day test but with all the storage hosts running in “tuned” configuration.

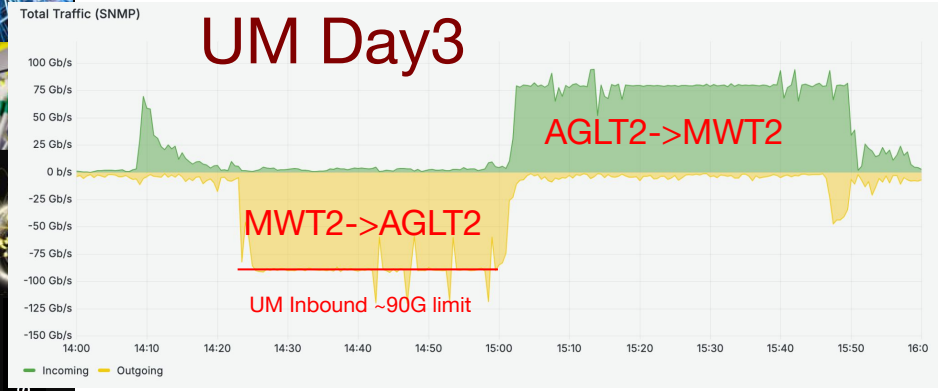
USATLAS: Per-file Rates for Feb 2-3 Testing

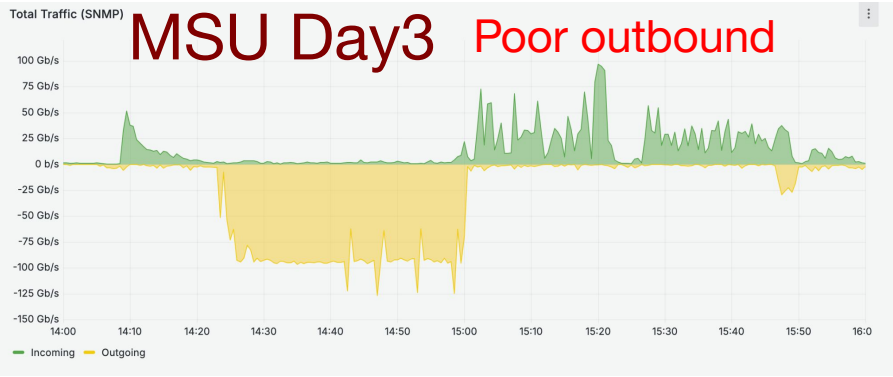
The top row shows the MWT2->AGLT2 file rates.

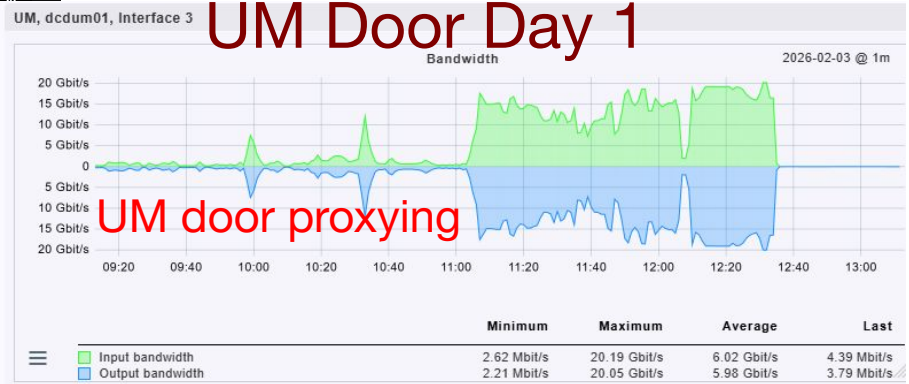
The 2nd row shows AGLT2-MWT2 file rates.

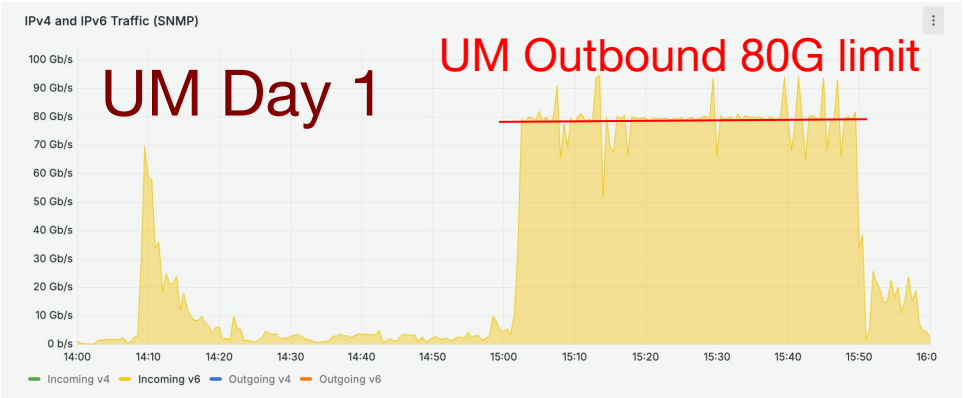
Unclear on the tuning impact…

USATLAS: Some Plots Showing Issues









USCMS Outcomes

- We saw significant benefit in the iperf3 test i.e. network usage but only when DTNs have large (100G vs 25G) NICs
- UCSD needs to address the low performance of their Storage System to match the rest of the infrastructure
- Vanderbilt might benefit from larger NICs in their DTNs also they should take a look at the seemingly grouping observed at the throughput-per-file histogram
- We need to understand if the drop in overall performance seen for UNL and Vanderbilt (this test vs last capacity Mini-Challenge) is due to the fact that we were using only a subset of nodes at FNAL

USATLAS Outcomes

- We are still working on interpretation.
- We set up to gather a significant amount of monitoring: CheckMK plots, Ganglia plots, ESnet Stardust and various logs
- Problems with conflicting production (merge) jobs at MWT2 made tests on 2nd day fail
- Re-ran two days later (only tuned config) which showed best transfers we have seen (MWT2->AGLT2)
  - Other direction exposed MSU outbound issue being investigated.

USCMS: Initial lessons learned

- The tuning script focus on making the systems better at using the network, i.e. increases the TCP buffers, changes the congestion algorithm, etc.
- We see 3 different cases on the 3 T2s:

Site	DTNs capacity	Storage System Performance(*)	Effect of tuning
UNL	100G	Good	Improves
Vanderbilt	25G	Regular	Worsens
UCSD	100G	Bad	No Difference

Conclusion: Tuning helps if the DTNs have room-to-grow AND the Storage System is not the bottleneck. It might hurt if the DTNs are the bottleneck

(*) Based on write rates observed in the previous capacity Mini-Challenge (see backup slides)

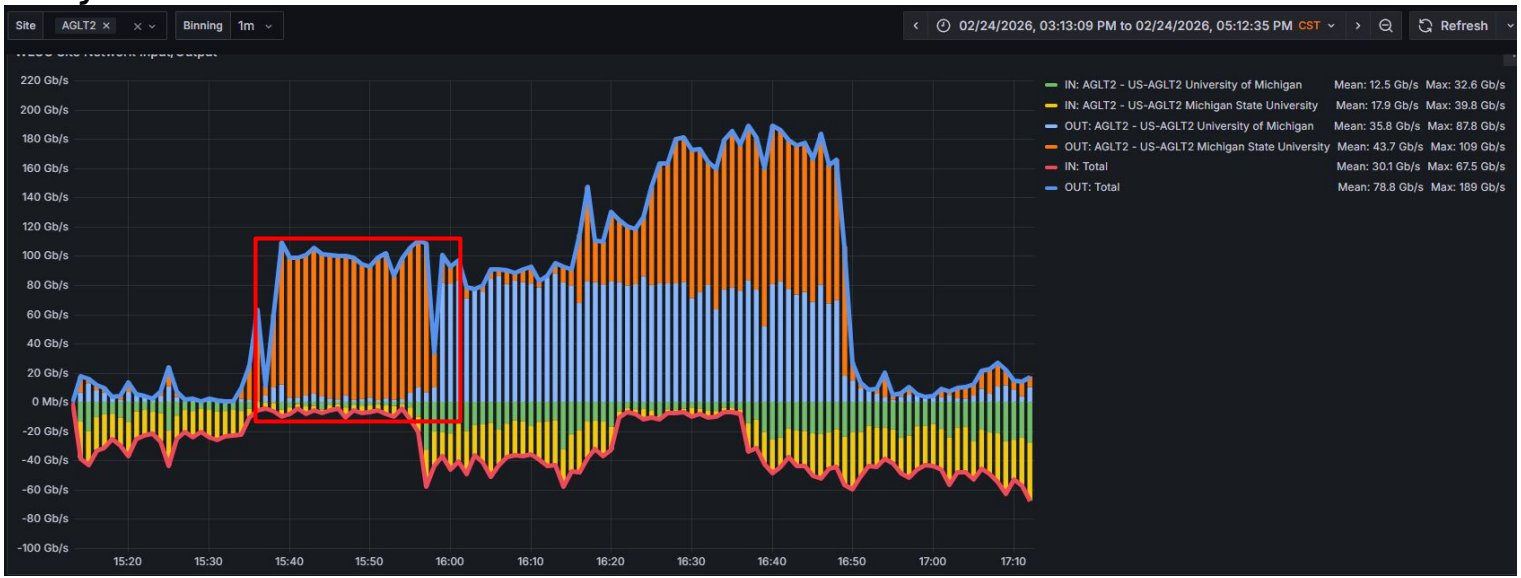
USATLAS: Initial lessons learned

For USATLAS, we had a few quick lessons learned as part of setting up and monitoring these tests:

- AGLT2 had a **misconfigured dCache** which was proxying read transfers via our (VM) doors and overloading MSU VMs
- **Low-level disk configs need fixing** for some hosts at UM/MSU
- We identified **a possible outbound issue** at MSU which is reducing transfer rates from AGLT2_MSU to other sites (next slide)
- UM still has a “known” **output bottleneck** at 80 Gbps, but was was a surprise is than inbound traffic can bypass that, reaching 90 Gbps
- MWT2 was the target of a large amount of merge jobs which hit on the second day of testing, resulting in 95% failures for our added testing transfers, exposing a need to protect from overload.

Checking Outbound at AGLT2_MSU

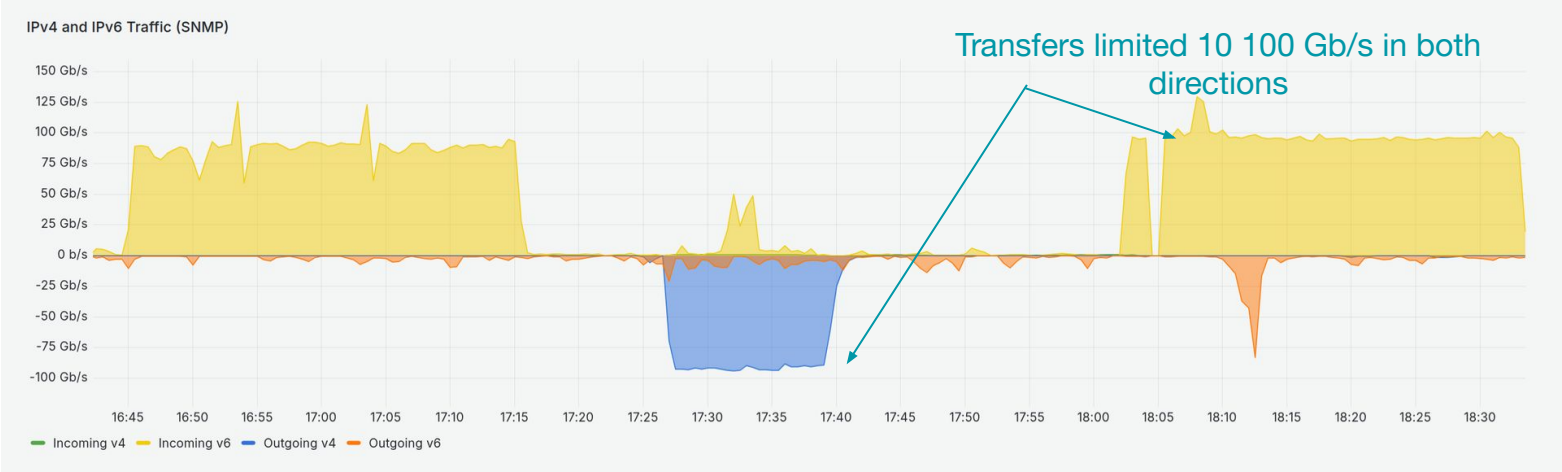
Yesterday we ran some tests to check AGLT2_MSU's outbound ability.
We identified a set of large files for AGLT2_MSU and AGLT2_UM, so we can separately test each site. Results below



MSU has 35% of the storage but higher network bandwidth. This shows the importance of balancing requests based upon achieved throughput…

Capacity challenge between NET2 and PRG

Quick micro-challenge last week: Both NET2 and PRG are Tier-2 centers connected to LHCONE at 400 Gb/s
Previous capacity tests with NET2 indicate that **site can transfer at 400 Gb/s** inside the US (eg., to BNL)
This test transfers high-throughput across transatlantic link.



Based on discussion with ESnet, limitation at 100 Gb/s seems to indicate a problem with (lack of) load-balancing issue within ESnet.
The traffic got pinned to one shortest path, not the best path(s).
Work ongoing to improve the situation before tests resume.

Near To Mid-Term Plans

The **host-tuning** will need multiple iterations before DC27

- If the things we tuned are NOT the bottleneck, no improvement
- Once we finish current evaluations, will schedule v2 in about 3 months
- Tuning clearly helps network/iperf3, but storage needs further study

In general, continued mini-challenges for both **capacity** and **capability**.

Not detailed in this talk but very important: getting more advocates to drive capability challenges for WLCG:

- File sizes (increase x10 to understand cost-benefit)
- Broader deployment of Network optimizations (Jumbo frames, pacing, new protocols)
- SDN (SENSE/Rucio), assuming testing shows benefits

Running more capacity challenges in regions we have yet to benchmark.

Conclusions

Concern was voiced that some IPv4 only sites may have their non-LHCONE uplinks overloaded if major/many sites cease IPv4 connectivity with LHCONE. Not observed during blackout test but perhaps worth future consideration.

We are as yet unsure of the benefit from host-tuning (as defined by our current script version) and will need further testing to demonstrate benefits.

Summary & Plans

We have successfully tested current USATLAS and USCMS capacity during our Feb 2026 testing, identifying some issues for further work.

We (IRIS-HEP/OSG-LHC/US-LHC) need to further **clarify** and **document** existing plans, **mini-challenges** and goals for the next year and onwards to DC27.

We have an **opportunity** to leverage mini-capability challenge results and findings to improve our infrastructure, to drive technology deployment, to show value and to demonstrate capabilities at scale.

Questions or Discussion?

Acknowledgements

We would like to thank the **WLCG**, **IRIS-HEP**, **HEPiX**, **perfSONAR** and **OSG** organizations for their work on the topics presented.

In addition we want to explicitly acknowledge the support of the **National Science Foundation** which supported this work via:

- IRIS-HEP: NSF OAC-1836650 and PHY-2323298

Background Material

Relevant resources we know about:

WLCG DOMA wiki page: https://twiki.cern.ch/twiki/bin/view/LCG/DomaMiniChallenges (includes link to Google Calendar to track activities)

ESnet Stardust page
https://dashboard.stardust.es.net/dashboards/f/ddhq1z9u9ca9sc/lhc-data-challenge

Presentations

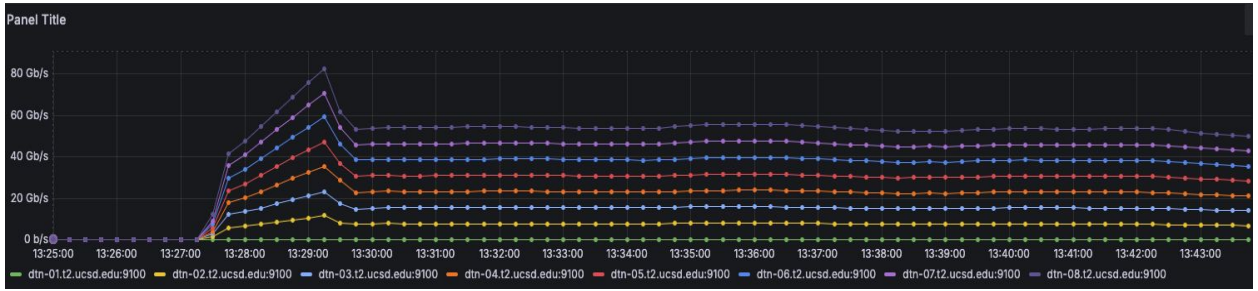
- Network Planning for Sites #77 S&C (2024)
- Results form US Mini-Challenges (Jan 2025)
- US Mini Data Challenge Plans (Feb 2025)
- Updates from perfSONAR, SciTags and the RNTWG HEPiX IPv6 WG (2026)

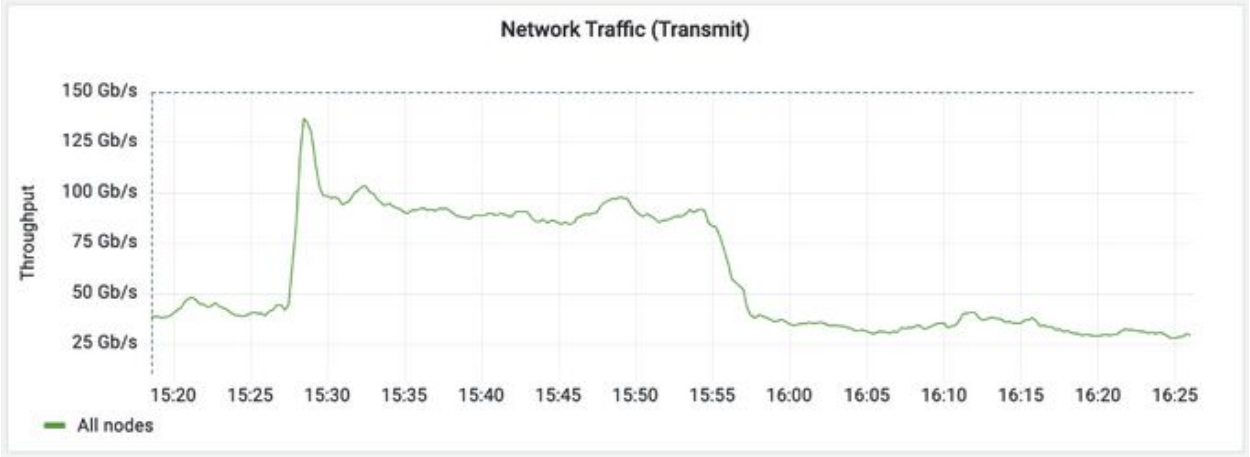
Some Google Docs

- WLCG/DOMA Data Challenge 2024: Final Report
- 2025-08 US Capacity Mini-challenge
- Host-tuning notes

Backup Slides

FNAL => UCSD 2 different PoV (background traffic)





No background traffic at UCSD. New pre-prod DTNs were used exclusively for these tests

~30 Gbps of background traffic at FNAL

These plots are always tricky to read because of the background traffic. SENSE can help with this :)

DC24 Links

Official DC24 report
https://zenodo.org/records/11402618

DC24 Network Activities and Results:
https://docs.google.com/presentation/d/1s0VvbXEpj1PN9umFT8wgsHsHmG9EYucymbalKNrvuKQ/edit#slide=id.p1

Katy Ellis LHCONE/LHCOPN DC24 presentation:
https://docs.google.com/presentation/d/1Tm3pCMkfHj5KHTW3PXbgS7mdHf72Ir27qr1JgMbrnRg/edit#slide=id.g1ea89411ecb_0_4

Next Steps Towards DC26:
https://docs.google.com/presentation/d/1mMx6QaihWJWpbVEQgxNjZXRT5_s4SkBTXu0SpELtuvI/edit#slide=id.gd170caf633_1_0

DC24 ATLAS Retrospective:
https://docs.google.com/presentation/d/1Lh_D57BvWn13AFClhhuczm-j-tKV-yMez_oD4yYUtBo/edit#slide=id.gd170caf633_1_0

USCMS Last capacity Mini-Challenge Results

		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