USCMS Mini-Challenges: from DC24 to DC27

Diego Davila with contributions from Asif Shah and Garhan Attebury

FTS/XRootD Workshop 2026

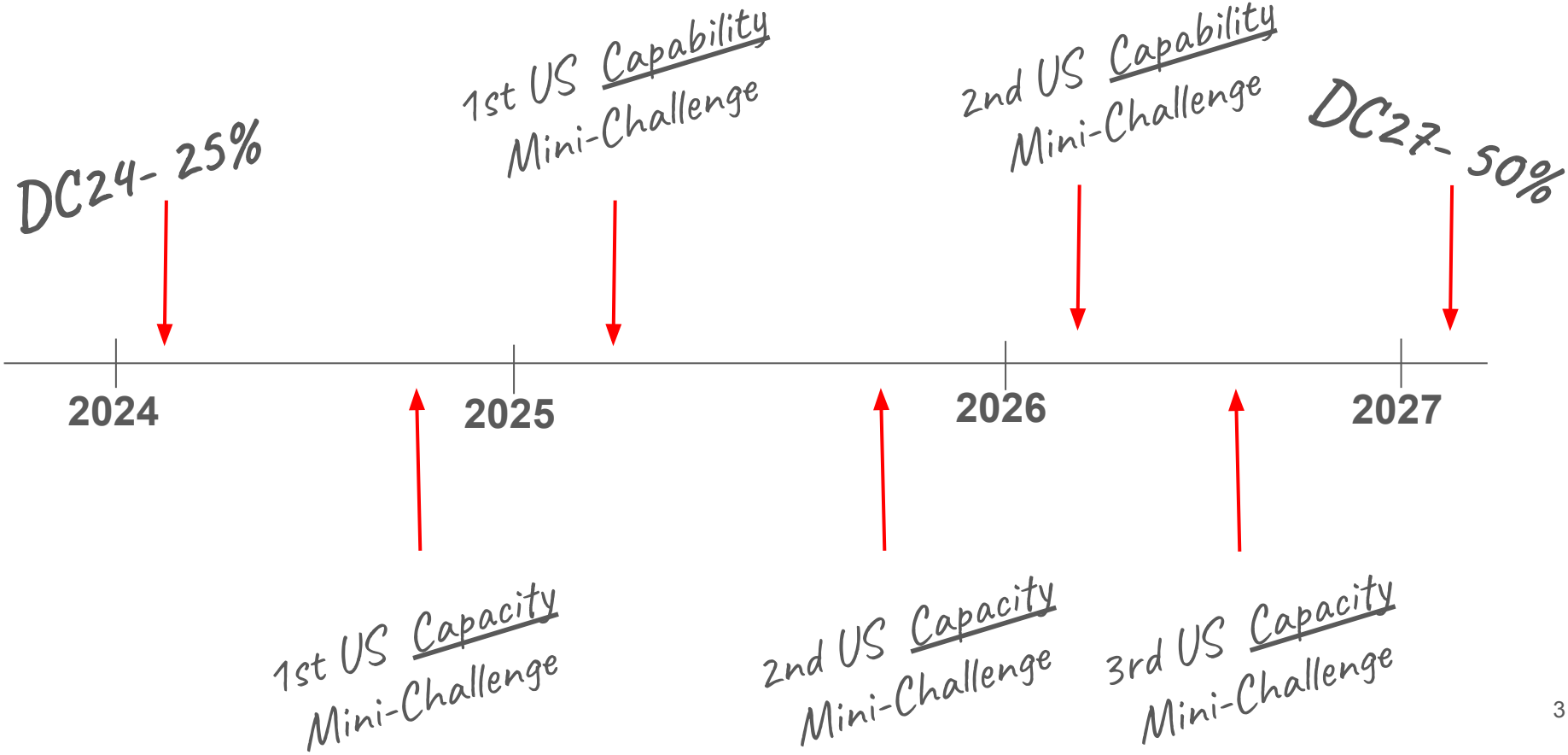
What exactly is a Mini-Challenge?

- First suggested by Mario Lassnig in 2022[1], they were described as:

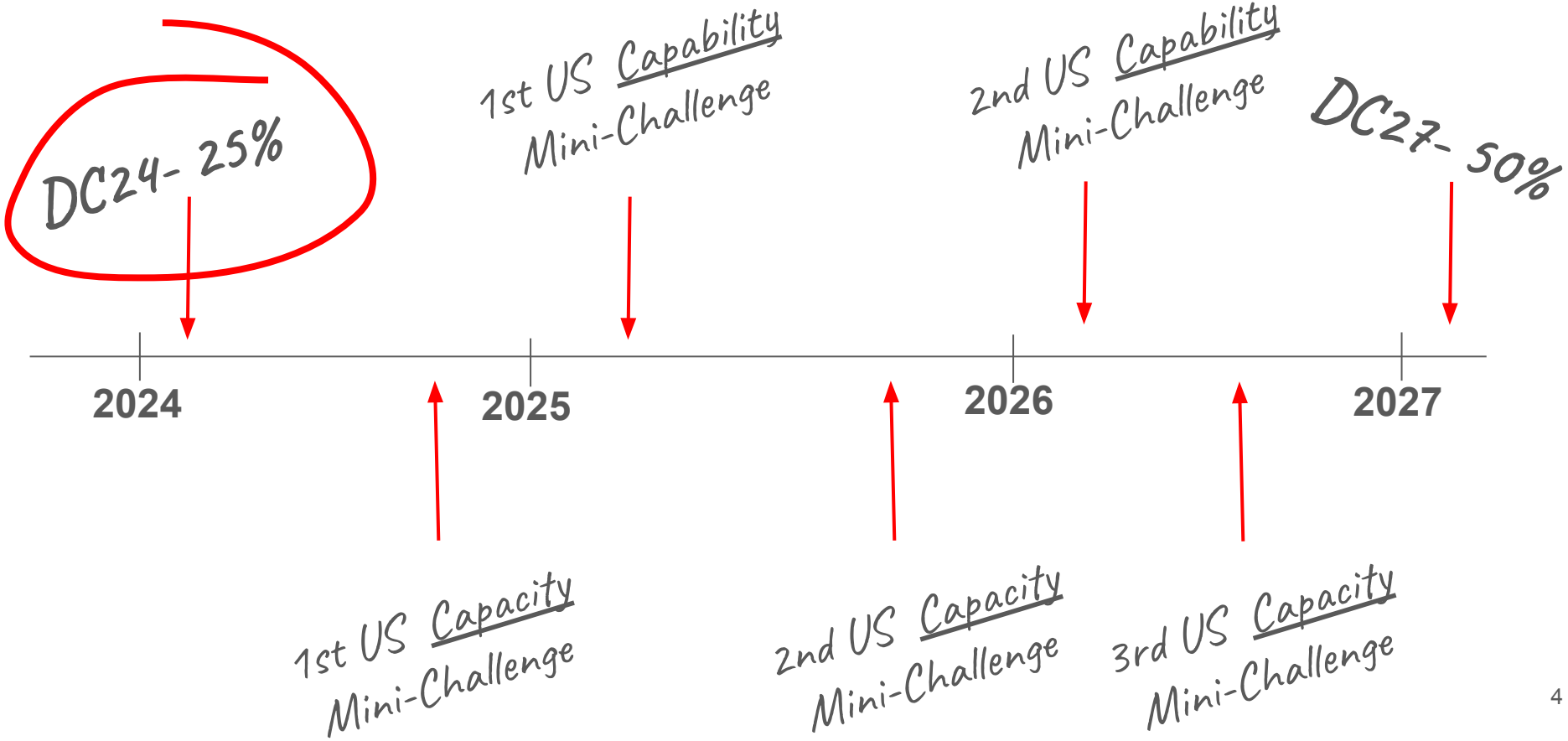
  *“a series of distributed, constrained, and isolated ramp-up challenges I**ndependently organised**, and report via WLCG/DOMA”*

- We have categorized into 2 types: ***Capacity*** (load test) and ***Capability*** where new technologies/techniques are showcased in production
- The cadence of these tests in the US has been: 1 of each per year

Timeline


DC24- 25%
1st US Capability Mini-Challenge
2nd US Capability Mini-Challenge
DC27- 50%
2024
2025
2026
2027
1st US Capacity Mini-Challenge
2nd US Capacity Mini-Challenge
3rd US Capacity Mini-Challenge

Timeline


DC24- 25%
1st US Capability Mini-Challenge
2nd US Capability Mini-Challenge
DC27- 50%
2024
2025
2026
2027
1st US Capacity Mini-Challenge
2nd US Capacity Mini-Challenge
3rd US Capacity Mini-Challenge

DC24 USCMS Targets

DC24 target was **25%** of the HL-LHC expected rates.

- FNAL: 400 Gbps
- T2s : 10 - 100 Gbps

DC24 USCMS Overview

DC24 target was **25%** of the HL-LHC expected rates.

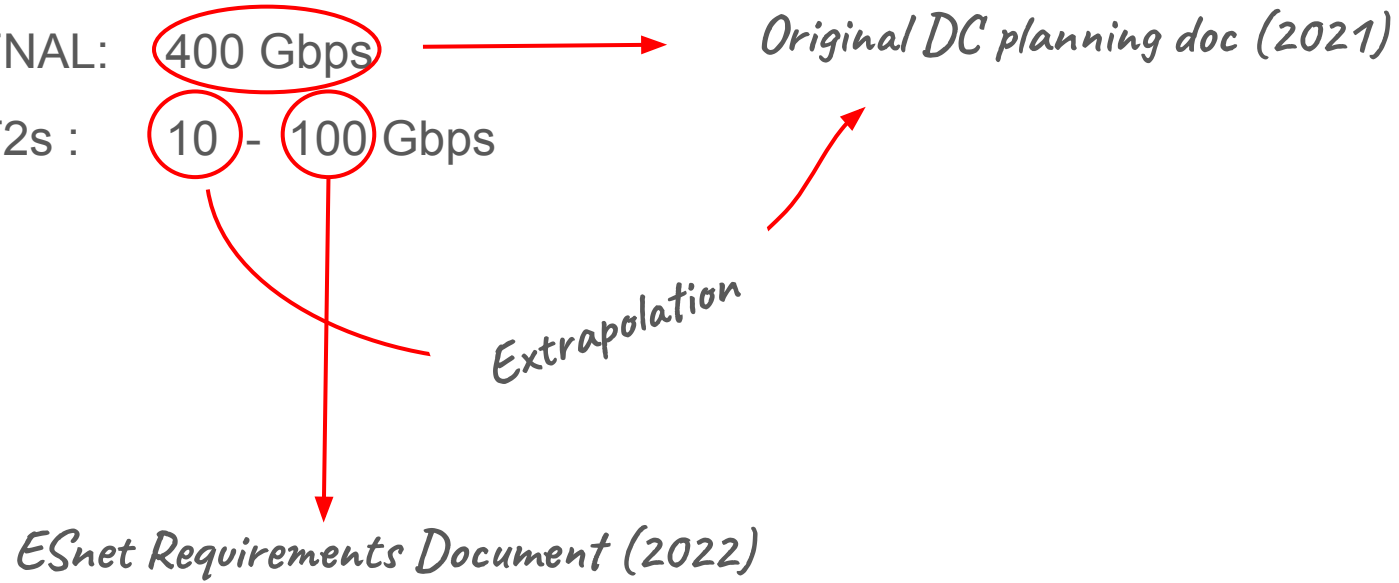
- FNAL: 400 Gbps → Original DC planning doc (2021)
- T2s : 10 - 100 Gbps

Extrapolation

DC24 USCMS Overview

DC24 target was **25%** of the HL-LHC expected rates.

- FNAL: 400 Gbps
- T2s : 10 - 100 Gbps



DC24 USCMS Overview

DC24 target was **25%** of the HL-LHC expected rates.

- FN
- T2

Original DC planning doc (circa 2021)

So far we have run with the 400 and 100 targets (T1 and T2s)
Updated numbers suggest that these targets are higher than what we need

ESnet Requirements Document (circa 2022)

DC24 USCMS Results

Writes

Site	DC24
FNAL	210
Caltech	73
Florida	90
MIT	50
Nebraska	70
Purdue	59
UCSD	28
Vanderbilt	39
Wisconsin	61

+

Reads

Site	DC24
FNAL	300
Caltech	68
Florida	24
MIT	34
Nebraska	49
Purdue	21
UCSD	33
Vanderbilt	25
Wisconsin	23

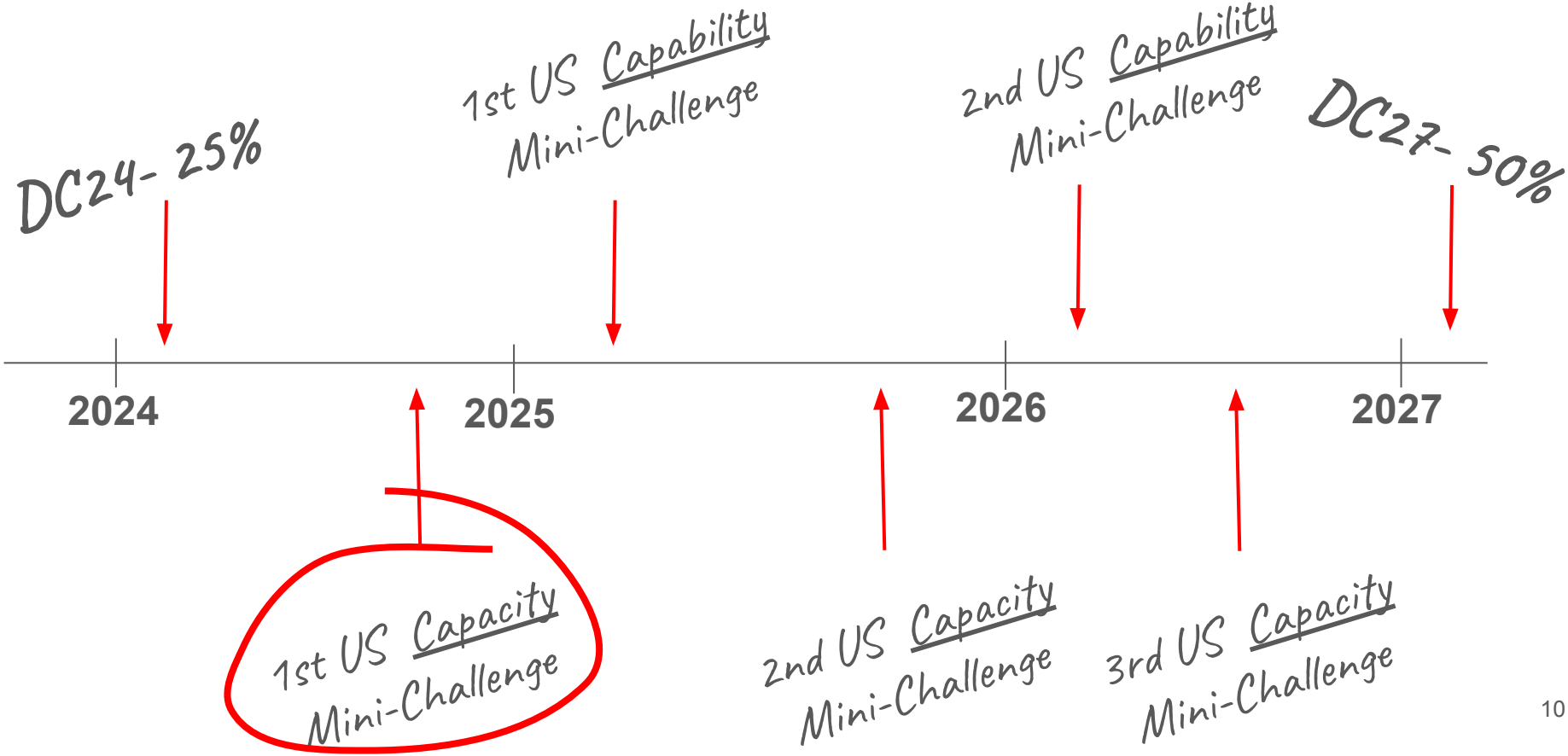
=

Writes + Reads

Site	W+R	Target
FNAL	510	128%
Caltech	141	141%
Florida	113	113%
MIT	84	84%
Nebraska	119	119%
Purdue	80	80%
UCSD	61	61%
Vanderbilt	64	64%
Wisconsin	84	84%

N.B. During DC24 Reads and Writes were done simultaneously. During the Mini-Challenges we carried out these 2 separately so the results should be consider separately

Timeline


DC24- 25%
1st US Capability Mini-Challenge
2nd US Capability Mini-Challenge
DC27- 50%
2024
2025
2026
2027
1st US Capacity Mini-Challenge
2nd US Capacity Mini-Challenge
3rd US Capacity Mini-Challenge

2024 - 1st USCMS **Capacity** Mini-Challenge

Note: All USCMS Capacity Mini-Challenges presented in the following follow the same strategy:

- Writes and Reads are performed separately
- Every T2 is tested separately using FNAL as the other side
- A group of T2s is used to test FNAL

About this Mini-Challenge:

- FNAL, Caltech and UCSD were not tested for different reasons
- The T2 Target was the same as in DC24: 100 Gbps
- Secondary objective: deploying/fixing the WLGC Site Network Monitoring
  - All but 1 site did it :)
- Miscellaneous Outcomes
  - Various fixes to the monitoring deployments were done
  - Found routing issues in Purdue, Nebraska and Wisconsin
  - FTS limits were increased in a number of sites

DC24 + 1st MC Results

Writes

Site	DC24	MC1
FNAL	210	N/A
Caltech	73	N/A
Florida	90	135
MIT	50	28
Nebraska	70	90
Purdue	59	60
UCSD	28	N/A
Vanderbilt	39	80
Wisconsin	61	60

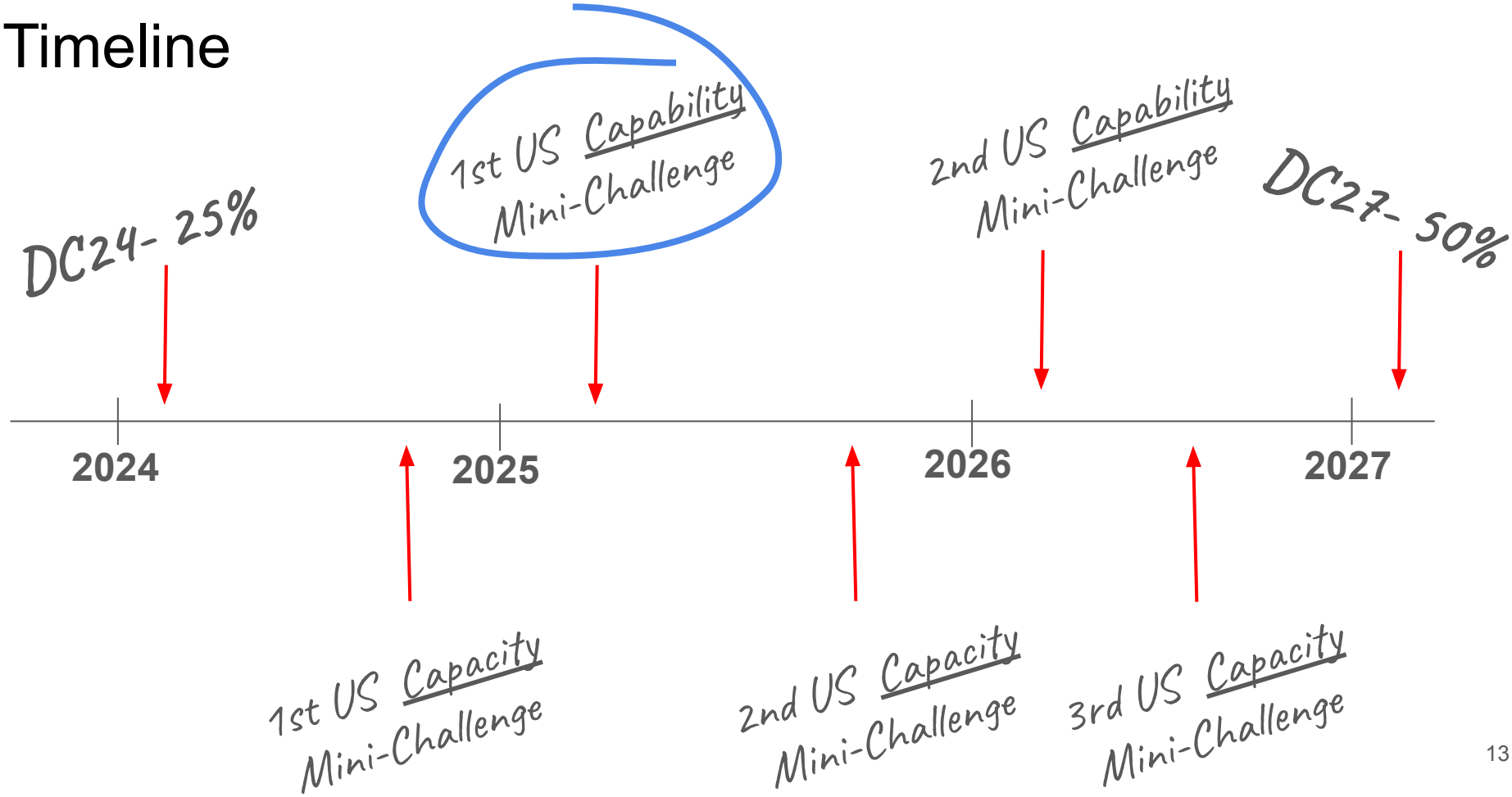
Due to write failures and lack of FTS pressure

Reads

Site	DC24	MC1
FNAL	300	N/A
Caltech	68	N/A
Florida	24	100
MIT	34	65
Nebraska	49	80
Purdue	21	100
UCSD	33	N/A
Vanderbilt	25	90
Wisconsin	23	100

- Most of the T2s either improved or stayed the same.
- 2/3 of the tests (8 of 12) reached at least the 80% of the goal

Timeline


DC24- 25%
1st US Capability Mini-Challenge
2nd US Capability Mini-Challenge
DC27- 50%
2024
2025
2026
2027
1st US Capacity Mini-Challenge
2nd US Capacity Mini-Challenge
3rd US Capacity Mini-Challenge

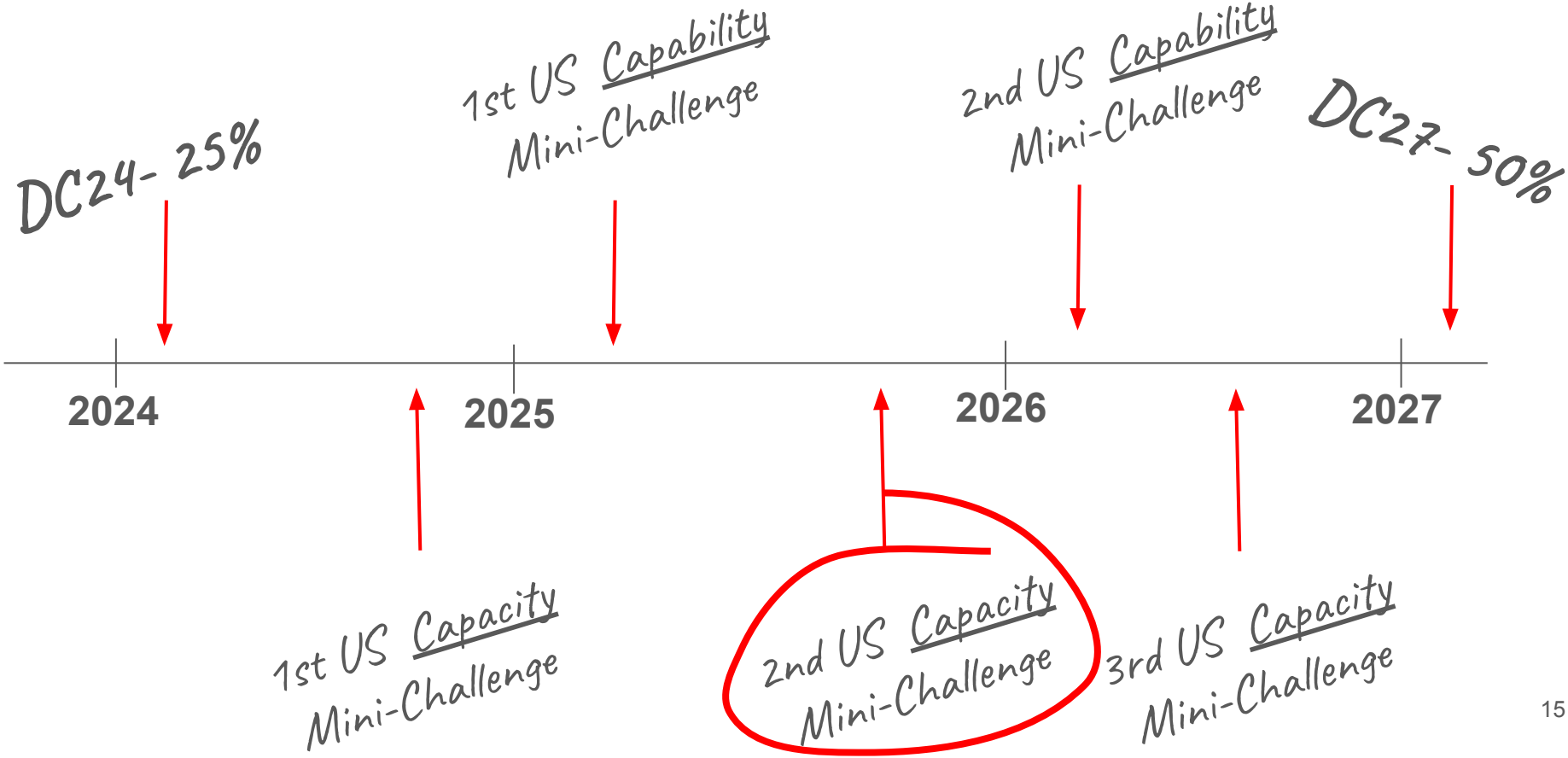
2025 - 1st USCMS **Capability** Mini-Challenge

Jumbo Frames. Tested between CERN and FNAL (high RTT) by our CERN folks[3] and between FNAL and the T2s (by Asif Shah[4]). The former showed a 20-25% improvement. The latter show neither improvement nor degradation.

Scitags. We managed to get 3 more T2s (6 total) configured to support scitags and carried out a simple validation. A known amount of files were transferred and tagged as “Data Challenge”, then the relevant scitags were queried and compared.

SENSE. A 100 Gbps SENSE controlled data flow was showcased between Caltech (prod) and Nebraska (pre-prod)

Timeline


DC24- 25%
1st US Capability Mini-Challenge
2nd US Capability Mini-Challenge
DC27- 50%
2024
2025
2026
2027
1st US Capacity Mini-Challenge
2nd US Capacity Mini-Challenge
3rd US Capacity Mini-Challenge

2025 - 2nd USCMS **Capacity** Mini-Challenge

- In addition to testing every USCMS site we carried out a joint test with USATLAS looking for possible collisions.
- T2 targets were picked by the admins between 100 and 200 Gbps
- FNAL target was set at 800 Gbps following DC27 expected rate
- Miscellaneous Outcomes
  - Florida’s bad results were later linked to a skewed time server
  - Routing issues at Nebraska due to SENSE misconfiguration
  - dCache’s max number of active transfers was reached at FNAL. It was increased afterwards

DC24 + 1st + 2nd MC Results

Writes

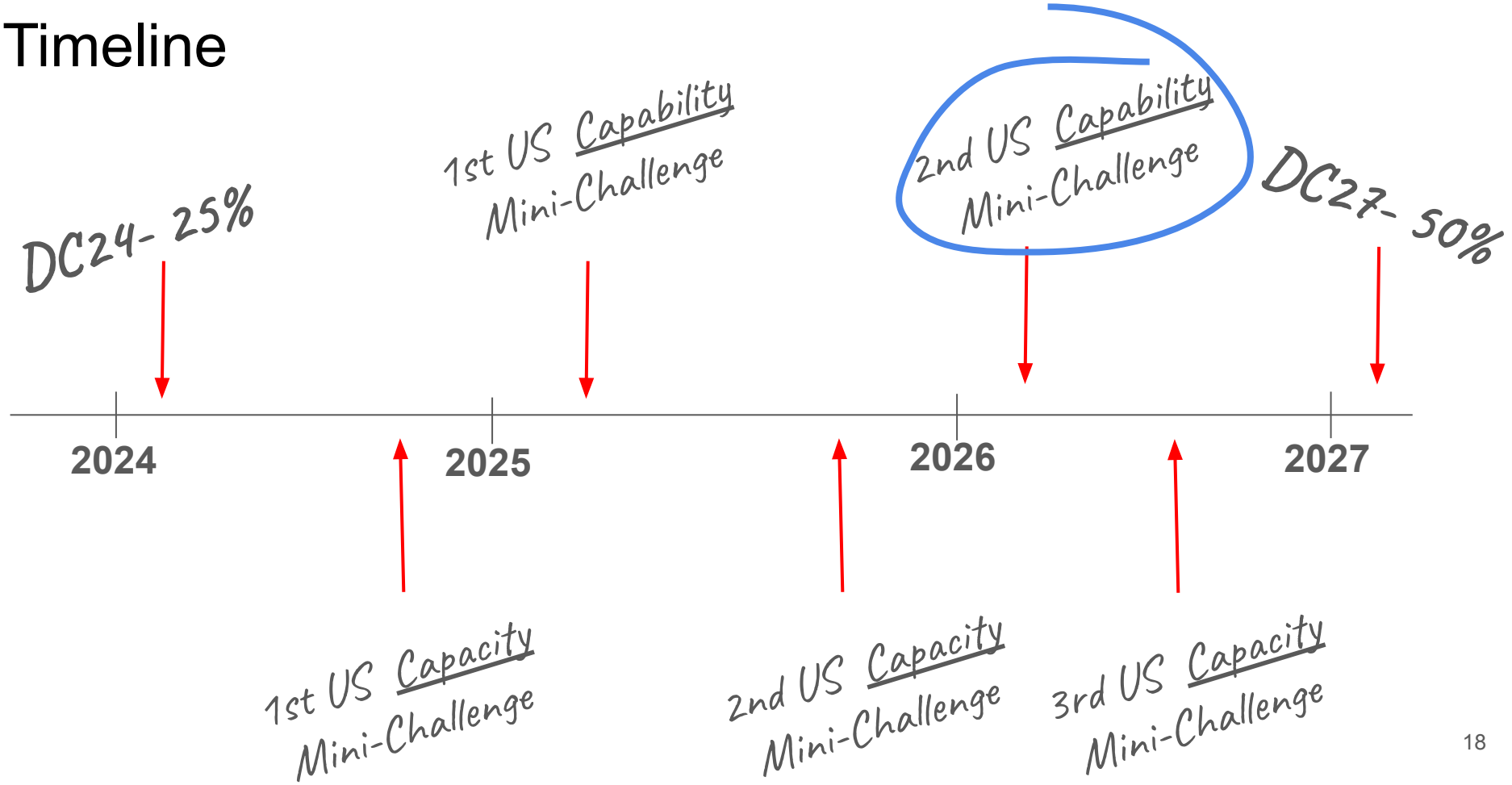
Site	DC24	MC1	MC2
FNAL	210	N/A	636
Caltech	73	N/A	70
Florida	90	135	35
MIT	50	28	47
Nebraska	70	90	115
Purdue	59	60	70
UCSD	28	N/A	50
Vanderbilt	39	80	50
Wisconsin	61	60	83

Reads

Site	DC24	MC1	MC2
FNAL	300	N/A	270
Caltech	68	N/A	85
Florida	24	100	50
MIT	34	65	80
Nebraska	49	80	65
Purdue	21	100	180
UCSD	33	N/A	45
Vanderbilt	25	90	120
Wisconsin	23	100	65

- Issues at Florida, Vanderbilt and Wisconsin are notable
- The low result of FNAL is due to lack of pressure from the T2s
- DC27 levels seem very far for most of the T2s

Timeline

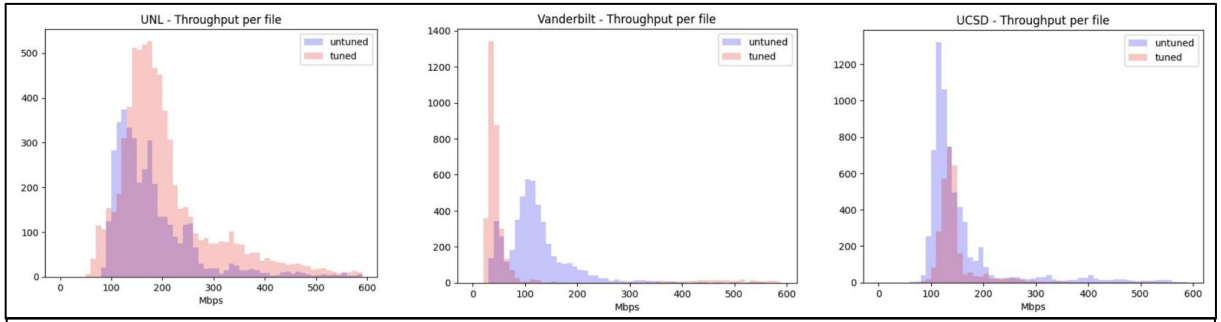

DC24- 25%
1st US Capability Mini-Challenge
2nd US Capability Mini-Challenge
DC27- 50%
2024
2025
2026
2027
1st US Capacity Mini-Challenge
2nd US Capacity Mini-Challenge
3rd US Capacity Mini-Challenge

2026 - 2nd USCMS **Capability** Mini-Challenge

LHCONE IPv4 Blackout. IPv4 was disabled on the LHCONE for 2 days on the participat sites. No problems were spotted, IPv4 traffic was routed through a different R&D Networks (e.g. Internet2)

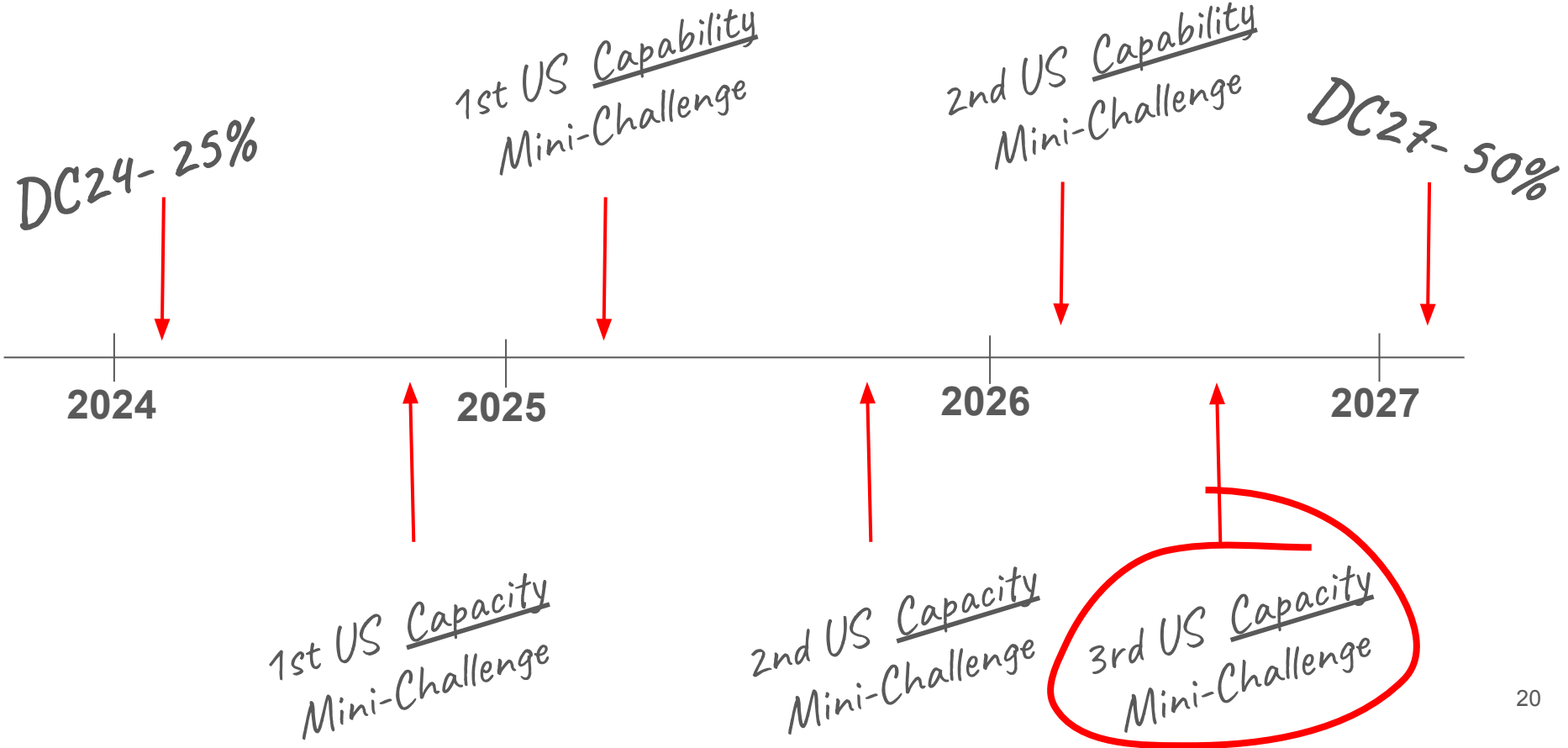
SENSE ("yes" again). Showed progress on the USATLAS between NET2 (prod) and MWT2 (dev) and on USCMS between FNAL (pre-prod) and UCSD (pre-prod)

Host Tuning. Based on the Fasterdata web site a script was written to tune hosts. The results were unclear



Throughput per file before (blue) and after tuning (red) on 3 different sites

Timeline


DC24- 25%
1st US Capability Mini-Challenge
2nd US Capability Mini-Challenge
DC27- 50%
2024
2025
2026
2027
1st US Capacity Mini-Challenge
2nd US Capacity Mini-Challenge
3rd US Capacity Mini-Challenge

2026 - 3rd USCMS **Capacity** Mini-Challenge

- Carried out in July this year
- Targets: 100-200 Gbps for the T2s and 800 Gbps for FNAL
- Using direct submission to FTS this time
  - This allowed us to modify the FTS limits on the fly
- Jumbo Frames and IPv4 vs IPv6 analysis done in parallel
- Miscellaneous Outcomes:
  - Various fixes to the WLCG Site Network Monitoring deployments
  - Caught some traffic not using Jumbo Frames: UCSD and SPRACE
  - Noticed some traffic fragmentation when writing to FNAL (under investigation)
  - Found some IPv4 traffic when reading from FNAL (under investigation)

DC24 + 1st + 2nd + 3rd MC Results

Writes

Site	DC24	MC1	MC2	MC3
FNAL	210	N/A	636	700
Caltech	73	N/A	70	88
Florida	90	135	35	180
MIT	50	28	47	60
Nebraska	70	90	115	87
Purdue	59	60	70	265
UCSD	28	N/A	50	*140
Vanderbilt	39	80	50	42
Wisconsin	61	60	83	168

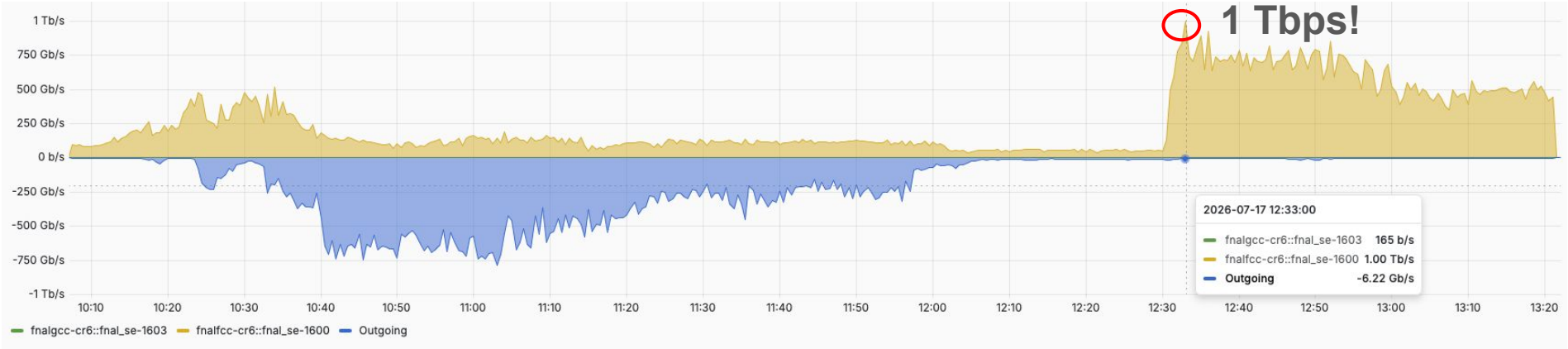
Reads

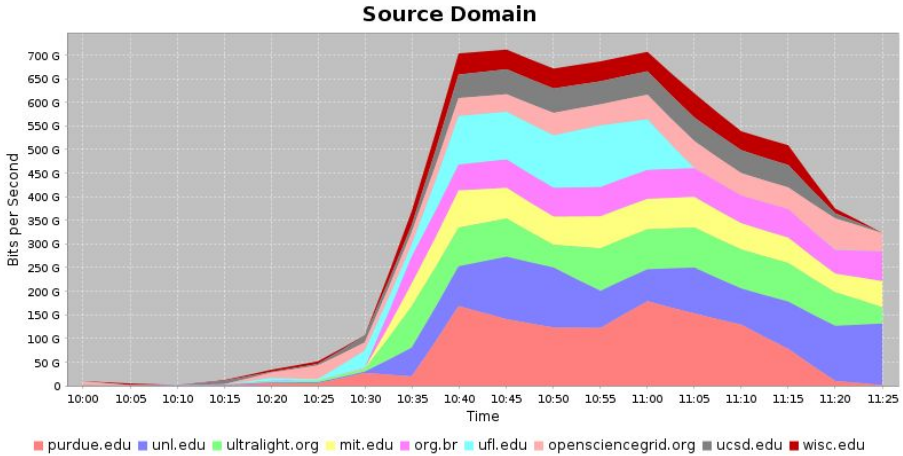
Site	DC24	MC1	MC2	MC3
FNAL	300	N/A	270	750
Caltech	68	N/A	85	80
Florida	24	100	50	125
MIT	34	65	80	82
Nebraska	49	80	65	175
Purdue	21	100	180	305
UCSD	33	N/A	45	52
Vanderbilt	25	90	120	56
Wisconsin	23	100	65	190

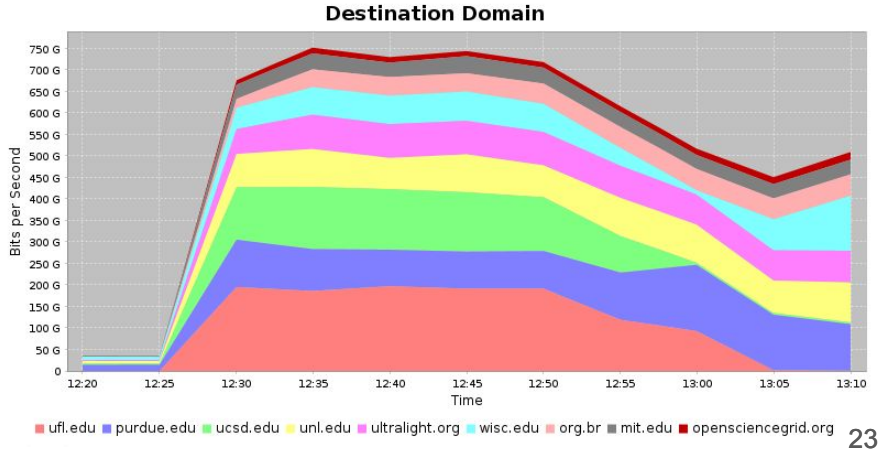
- Lots of improvements but still most T2s are far from the 200 Gbps goal
- Significant disparity across T2 capacities

* UCSD using blackhole plugin (writing to /dev/null)

3rd Capacity Mini-Challenge - **FNAL Test**

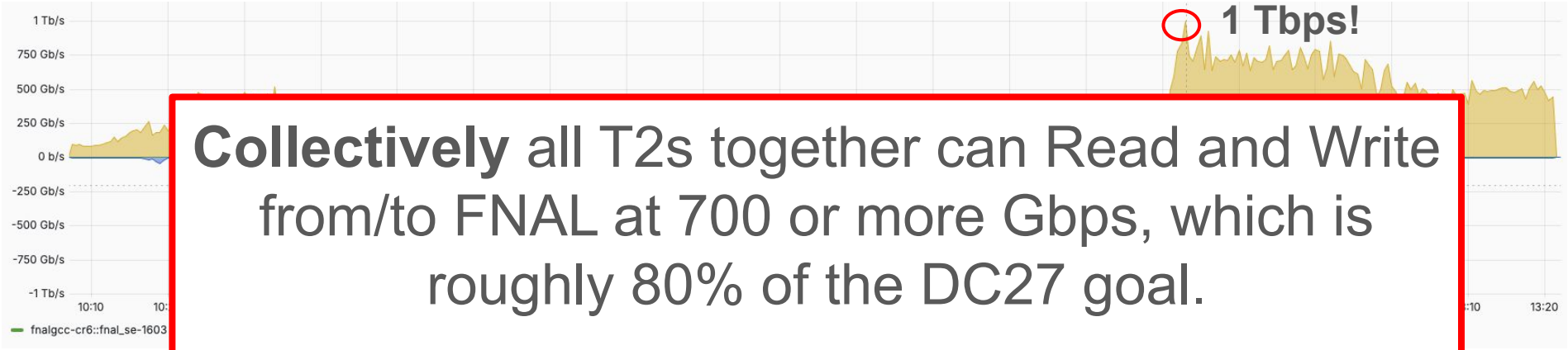




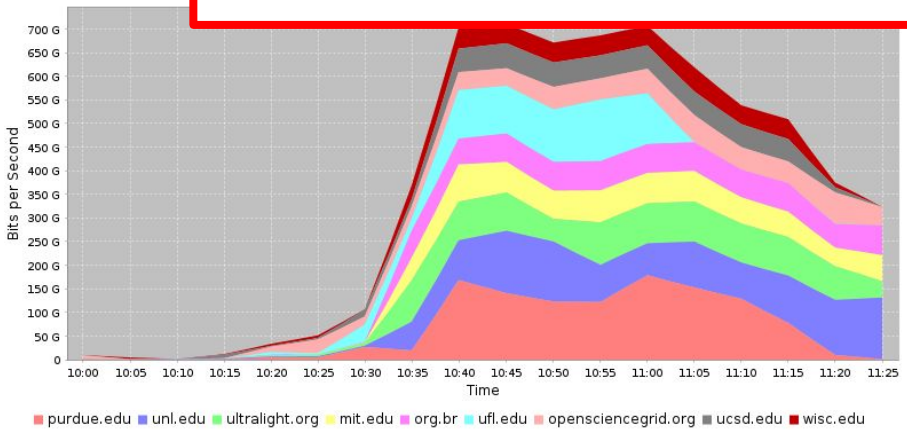


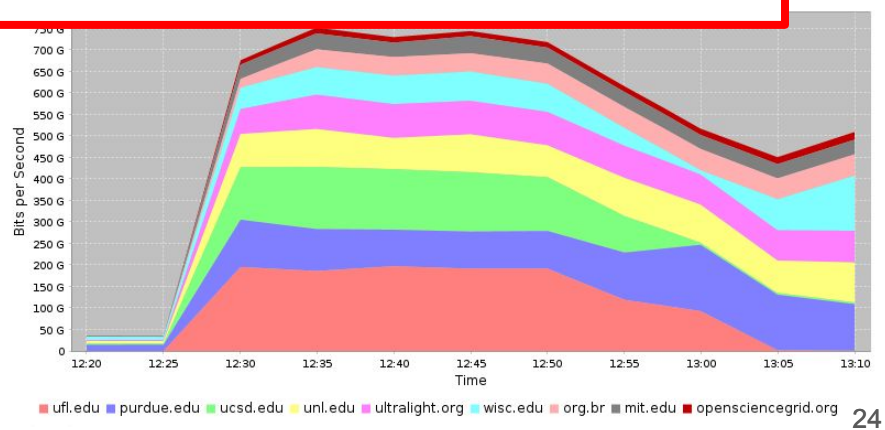
Note: Opensciencegrid representing “Vanderbilt” and org.br representing “SPRACE”

3rd Capacity Mini-Challenge - **FNAL Test**



Collectively all T2s together can Read and Write from/to FNAL at 700 or more Gbps, which is roughly 80% of the DC27 goal.





Note: Opensciencegrid representing “Vanderbilt” and org.br representing “SPRACE”

Summary

- After DC24 we started running Mini-Challenges twice a year
- Capacity Mini-Challenges help sites identify shortcomes
- Capability Mini-Challenges help us test new technology in a production environment
- 5 Mini-Challenges have been carried out so far
- We have performed some of these test together with USATLAS
- The 200 Gbps is still far for most of the T2s
- Huge performance disparity across T2s
- Collectively all T2s can reach 80% of FNAL’s DC27 goal

Links to talks/reports

- DC24 CMS Overview presentation: https://indico.cern.ch/event/1340661/contributions/5832693/attachments/2822185/4929423/DC24%20results%20for%20O&C%20week-2.pdf
- DC24 USCMS Overview report: https://docs.google.com/document/d/1m05fvGPO4rdmYQaE-7LZg0R2-WJL9ImeQCP1cMuSOeU/edit?usp=sharing
- 1st US (CMS and ATLAS) Capacity Mini-Challenge presentation in WLCG DOMA: https://indico.cern.ch/event/1495675/
- 1st USCMS Capacity Mini-Challenge report: https://docs.google.com/document/d/1rtfhxfsvHYfqc5xQXsFQot-Vu_Ek83WGJMEfKCiVs1I/edit?usp=sharing
- 1st USCMS Capability Mini-Challenge report: https://docs.google.com/document/d/1mTDX-GwgaEHRyGfpOtRIjz_SY86oS7VS_HGhXj7_Ti0/edit?usp=sharing

Links to talks/reports

- 1st Capacity and Capability Mini-Challenges presentation: https://indico.cern.ch/event/1479019/contributions/6300516/attachments/3033757/5356906/Results%20from%20US%20Mini-Challenges.pdf
- 2nd US (CMS and ATLAS) Capacity Mini-Challenge presentation in WLCG OTF https://indico.cern.ch/event/1562124/contributions/6680945/attachments/3133245/5558850/Results%20of%20Mini%20Data%20Challenges%20in%20the%20U.S..pdf
- 2nd US (CMS and ATLAS) Capability Mini-Challenge presentation in WLGC DOMA: https://indico.cern.ch/event/1652743/contributions/6954332/attachments/3226796/5751333/Results%20from%20US%20MIni-Challenges.pdf
- 3rd USCMS Capacity Mini-Challenge presentation in WLCG DOMA:
- https://indico.cern.ch/event/1722889/contributions/7246986/attachments/3336281/5978010/2026%20Summer%20USCMS%20Mini-Challenge.pdf

References

[1]https://indico.cern.ch/event/1146558/contributions/4908525/attachments/2534611/4361807/The%20path%20to%20the%20next%20WLCG%20Data%20Challenge.pdf

[2]https://escholarship.org/uc/item/3jj0h54n

[3]https://indico.cern.ch/event/1511535/contributions/6374829/attachments/3018021/5323446/JumboFramesTests.pdf

[4]https://docs.google.com/document/d/1a57YtNJ7hSUV1BIelud-48l0ET6LulsFeV0LnjEvaQk/edit?tab=t.0#heading=h.u5yp6wnvquqo

Acknowledgements

This work was supported by the US National Science Foundation (NSF) Grants: PHY-2323298 and PHY-2121686

Special thanks to Asif Shah (FNAL) and Garhan Attebury (UNL) for their significant contributions and to our of collaborators at ESNet.

Questions?

Backup slides

MIT in 1st Capacity Mini-Challenge

2024-12-12T18:44:02Z	460	201	2568	100.00%	1008.81 MiB/s	1008.81 MiB/s	10	Good link efficiency. Increment
2024-12-12T18:43:24Z	450	200	2570	100.00%	944.44 MiB/s	944.44 MiB/s	10	Good link efficiency. Increment
2024-12-12T18:42:46Z	440	195	2581	100.00%	901.02 MiB/s	901.02 MiB/s	10	Good link efficiency. Increment
2024-12-12T18:42:07Z	430	193	2583	100.00%	837.09 MiB/s	837.09 MiB/s	10	Good link efficiency. Increment
2024-12-12T18:41:29Z	420	196	2568	100.00%	765.63 MiB/s	765.63 MiB/s	10	Good link efficiency. Increment
2024-12-12T18:40:49Z	410	202	2562	100.00%	719.01 MiB/s	719.01 MiB/s	10	Good link efficiency. Increment

Figure 10. After reaching 200 active transfers, FTS is not pushing further even though the limits are way higher

USCMS Fall-2024 mini-challenge - Outcomes

- Foud an issue at Wisconsin (under investigation)
  - Reads go on a 200 Gbps path whilst writes go over a 100 Gbps one.
- Found a bottleneck at Nebraska. (upcoming upgrade should take care)
  - They are technically connected via 2x100G but logically traffic into their storage system is always assigned to a single link which caps Writes to 100G
- Found a problem with the go implementation of the WLCG Monitoring
  - It's been fixed
- Found missing checksums at Florida
  - This was fixed the same day it was found
- Increased FTS limits to 300+ from the 200 default for most T2s
- Switched: Florida and Wisconsin to use FNAL's FTS instead of CERN's
- Found an asymmetry within ESnet for Purdue: reads/writes over different interfaces.
  - This issue hasn't been solved yet.