



# US ATLAS Throughput Capacity challenge

Hironori Ito  
Brookhaven National Laboratory

06/06/2025



# Periodic testing of US ATLAS Sites

- Continuing to test sites
  - Identify issues as soon as possible.
  - Monitor the progress
    - Test new features and/or hardware
      - Network link
      - Jumbo frame
      - Token
      - New storage
- Testing network and storage capacity
  - Disk-to-disk test
  - Use FTS to transfer data from one storage to the other.
  - The [script](#) for this test was presented at several past meeting including [Network load generator](#).
    - The token support is not included. But, it will be coming soon.
  - Multiple monitors are checked

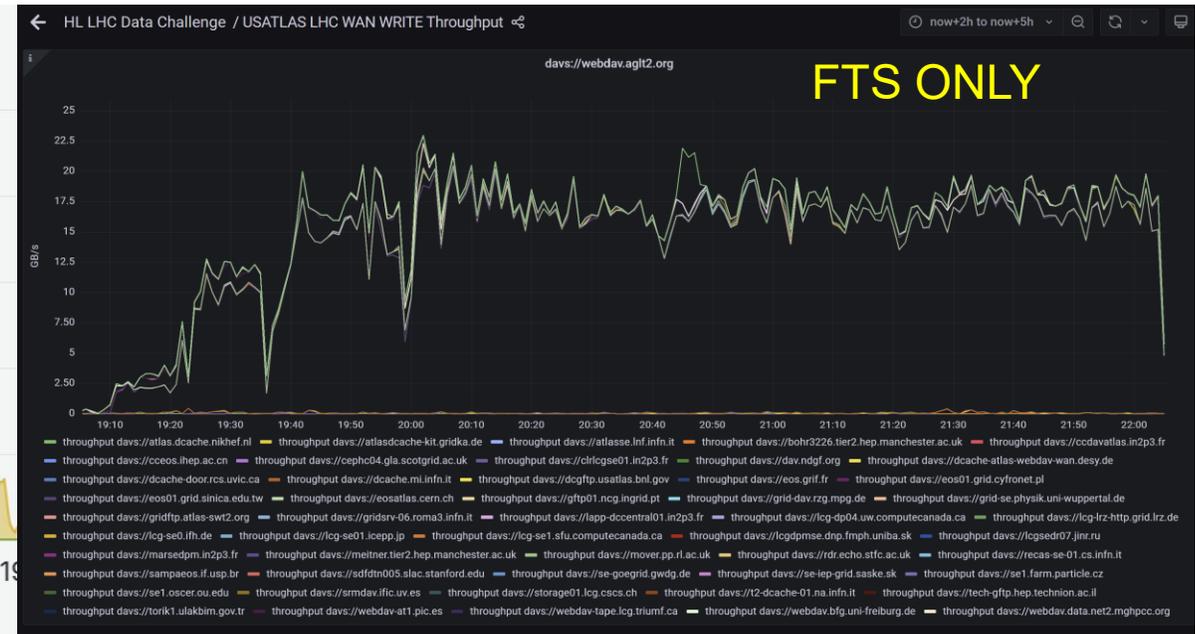
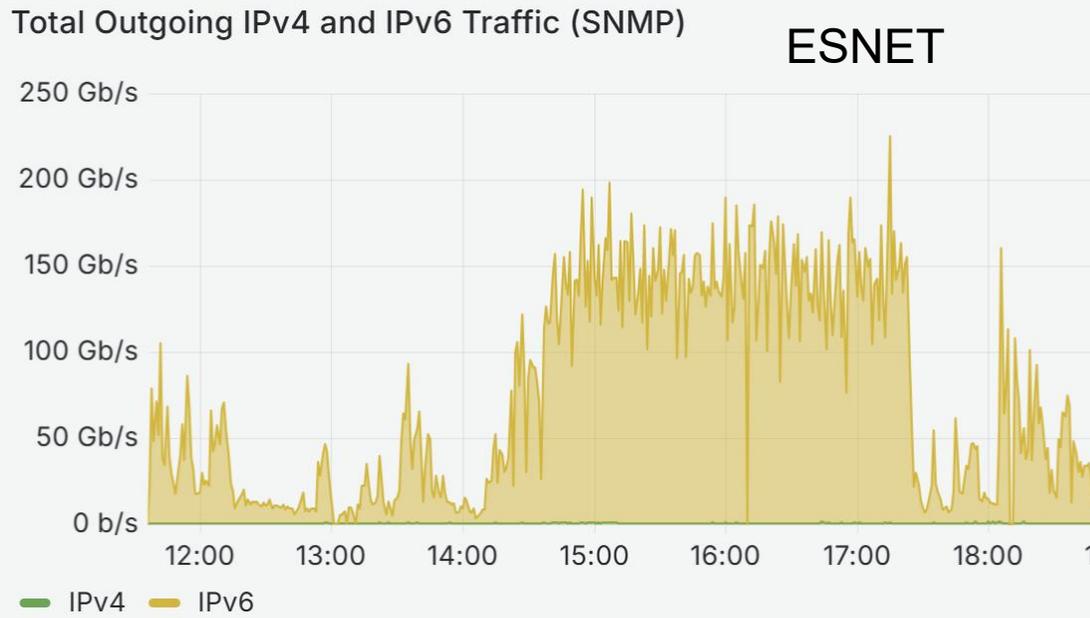
# US Site Capacity

SITE	Network Connectivity (Gbps)	Jumbo Frame
AGLT2	100 (UM) + 80 (MSU)	YES(UM), No(MSU)
BNL	800 (LHCOPN) + 800 (LHCONE)	Yes
MWT2	200 (U Chicago) + 100 (U Indiana)	Yes
NET2	400	Yes
SWT2	30 (UTA) + 100 (OU)	No (UTA), YES (OU)

NOTE: The most of these network bandwidth are well beyond HL-LHC minimum.

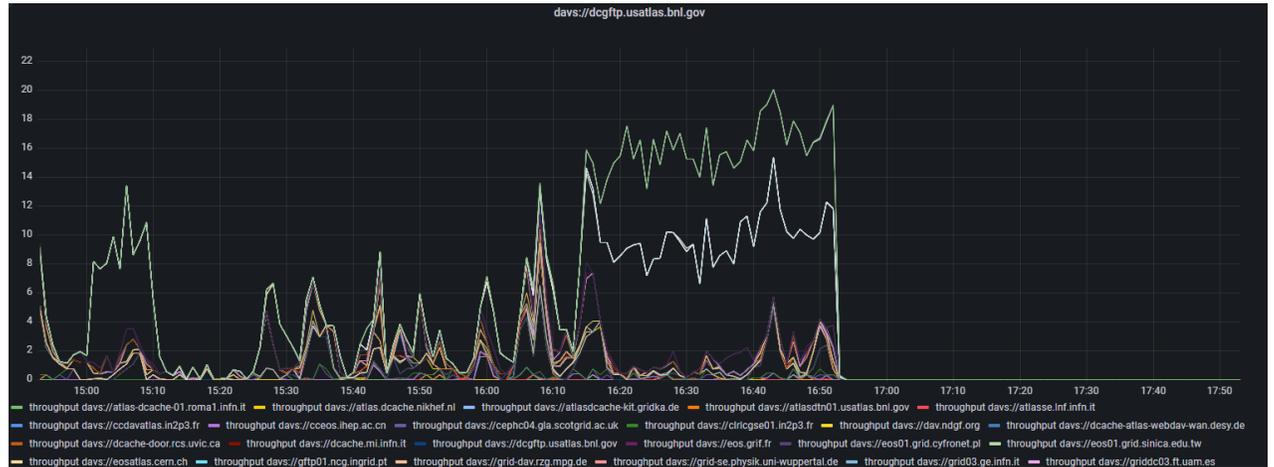
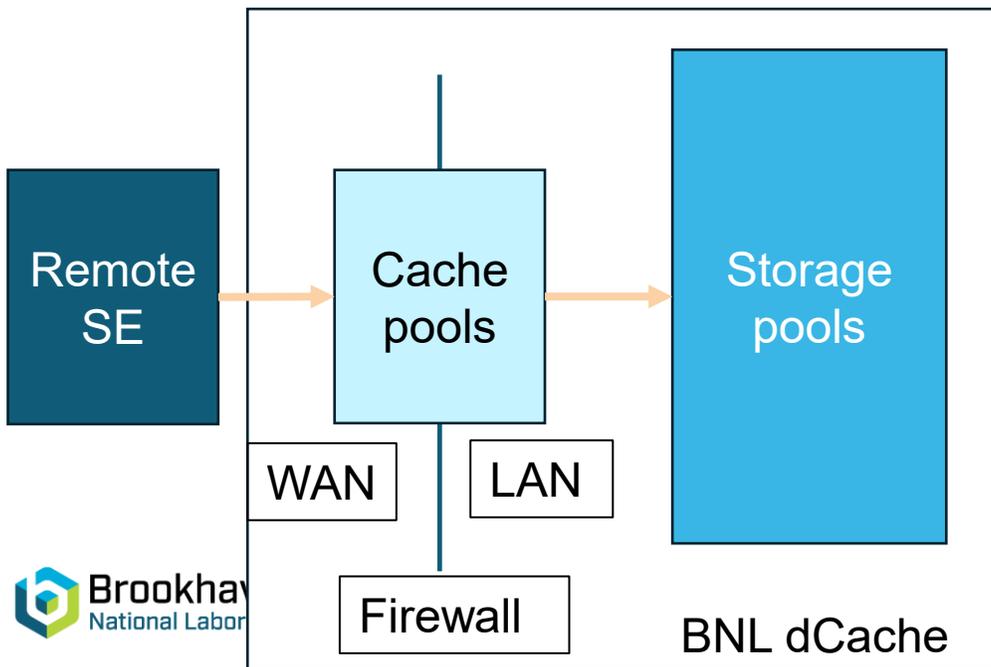
# AGLT2

- The site capacity, 180 Gbps, is achieved.
- The observed throughput of FTS / load generator is about 80-90% of the site capacity.

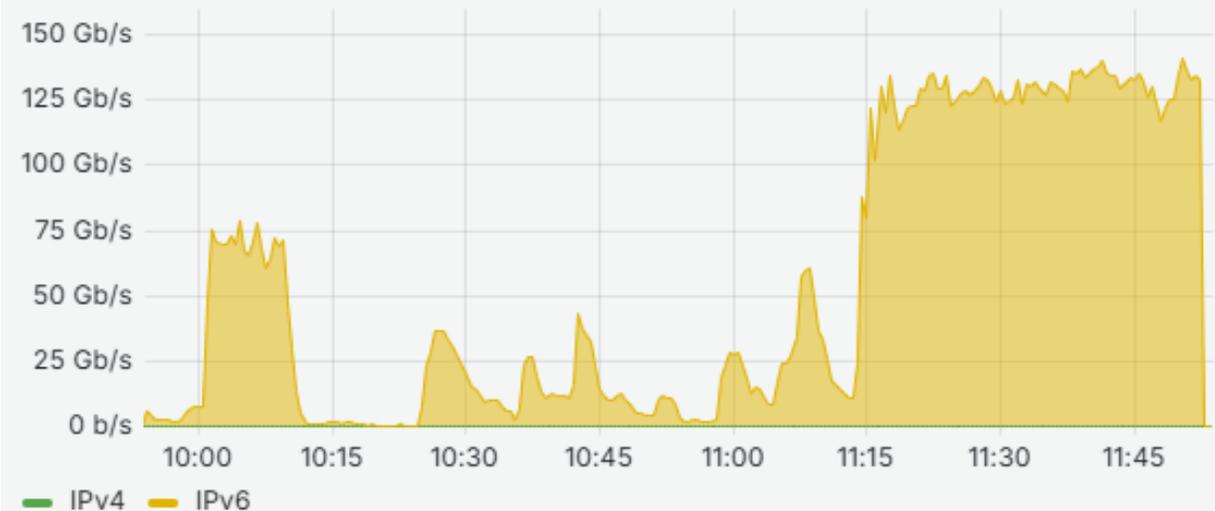


# BNL

- 125Gbps with 500 max concurrency
  - **This is concurrency limited in FTS setting.** (See the later jumbo frame section.)
- Use of the cache pools influence the behavior using WebDAV protocol in dCache.
  - It can't pass through DTN in write while it pass through in read.

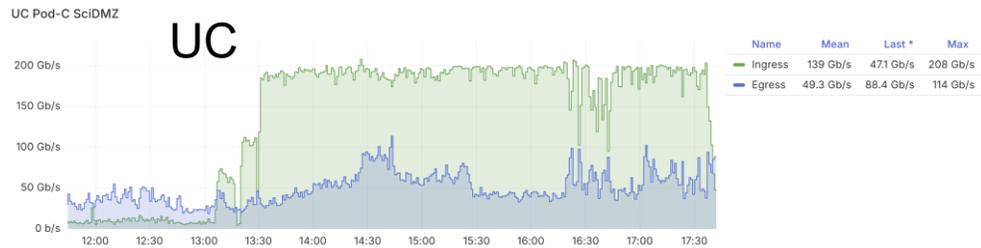
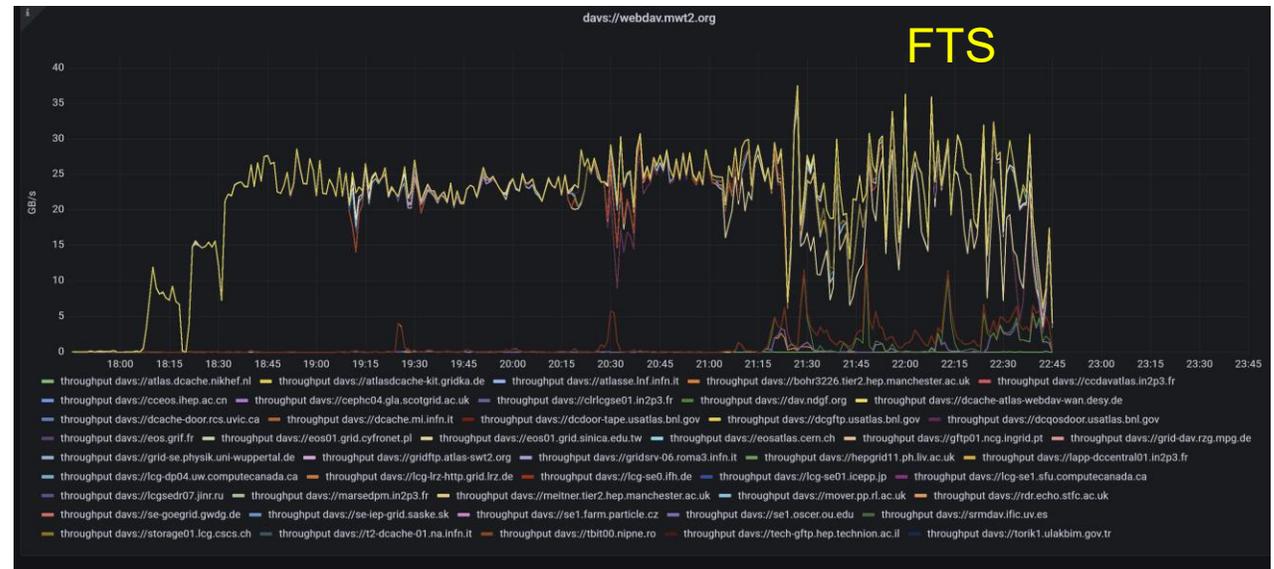


Total Outgoing IPv4 and IPv6 Traffic (SNMP)



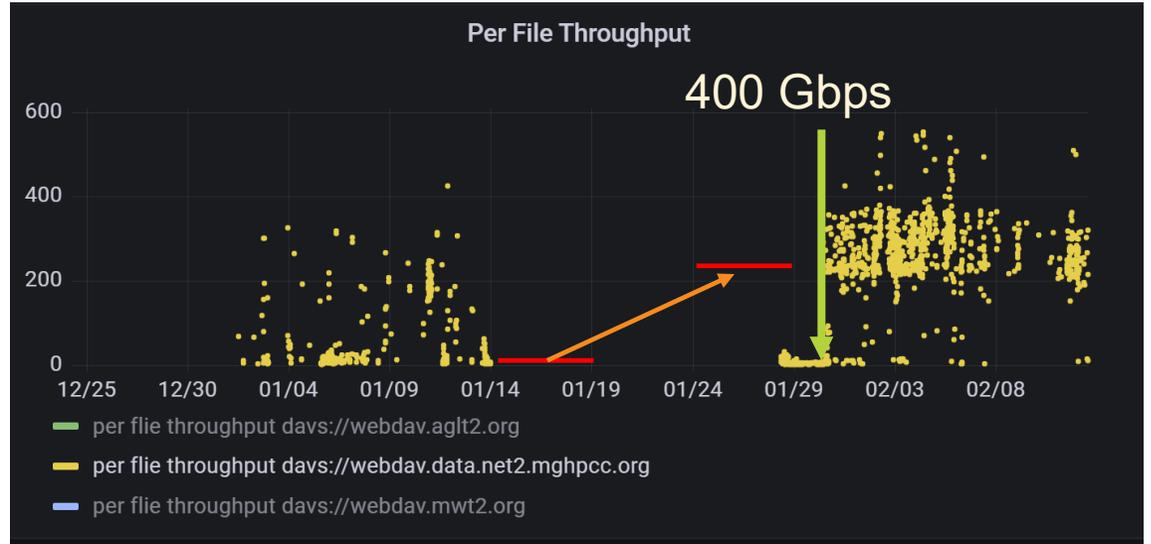
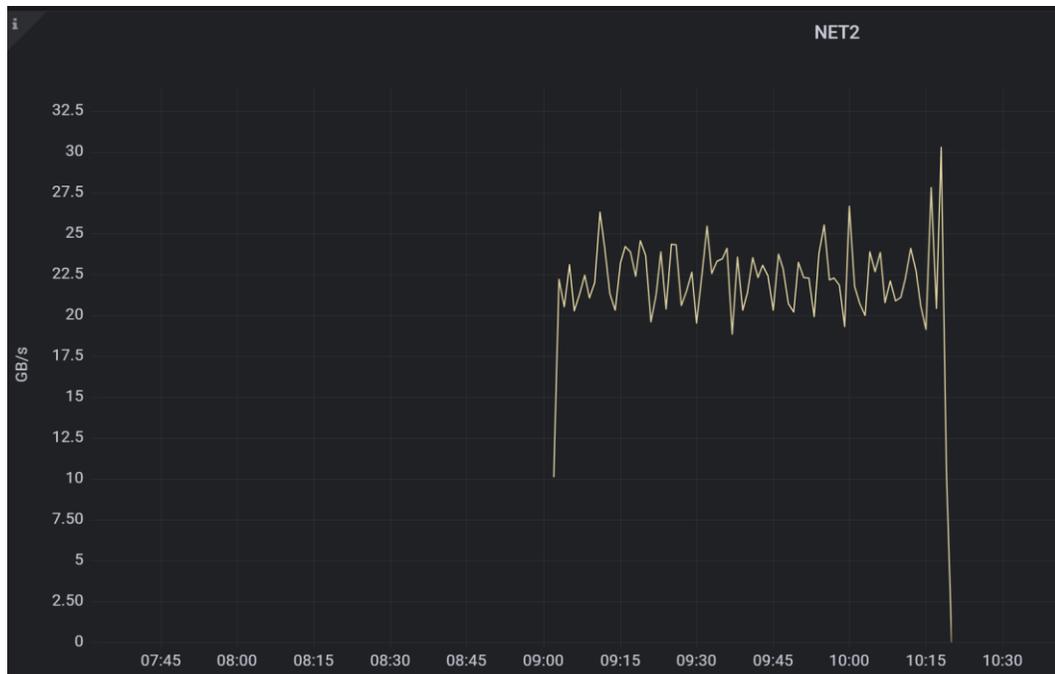
# MWT2

- Reached the site capacity (at UC)
- A bit tricky to reach max throughput at both sites with asymmetric network/storage capabilities.



# NET2

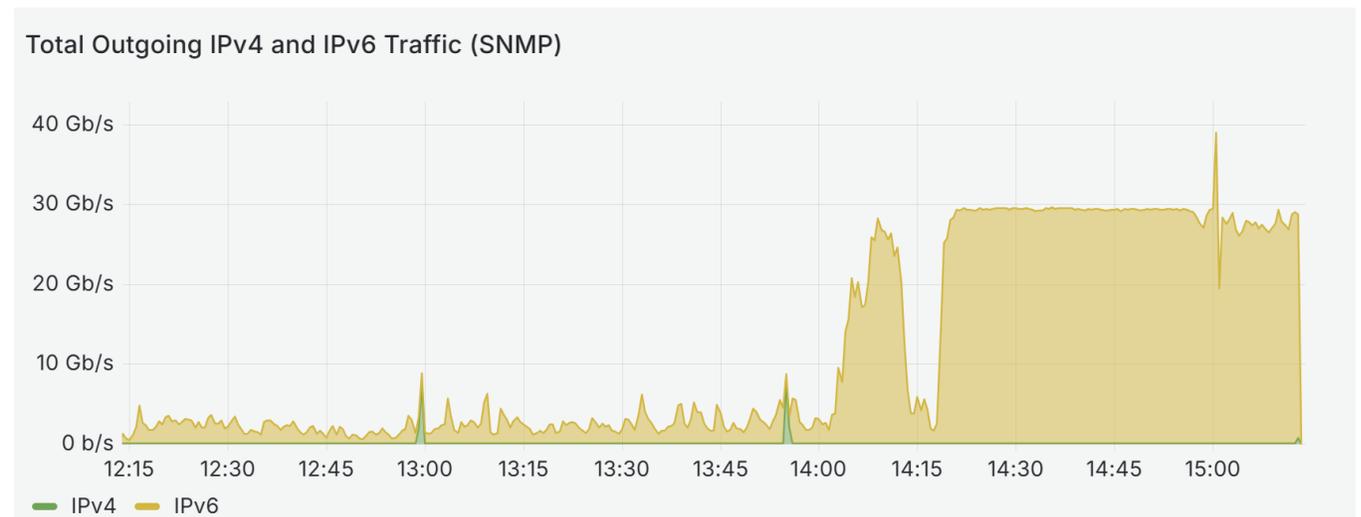
- NET2 has achieved ~180 Gbps



- The monitoring see the big effect of the upgraded network bandwidth.
- The throughput per file transfer improved tremendously.

# SWT2 (UTA)

- SWT2 (UTA) has achieved 30 Gbps.
- This is still the limit of the network at the site.
  - The flatness of the plot indicates that it is indeed the network limit.
- Discussion with the network engineer is under way to increase the bandwidth. (Needs both bandwidth and DTN capacity increases)

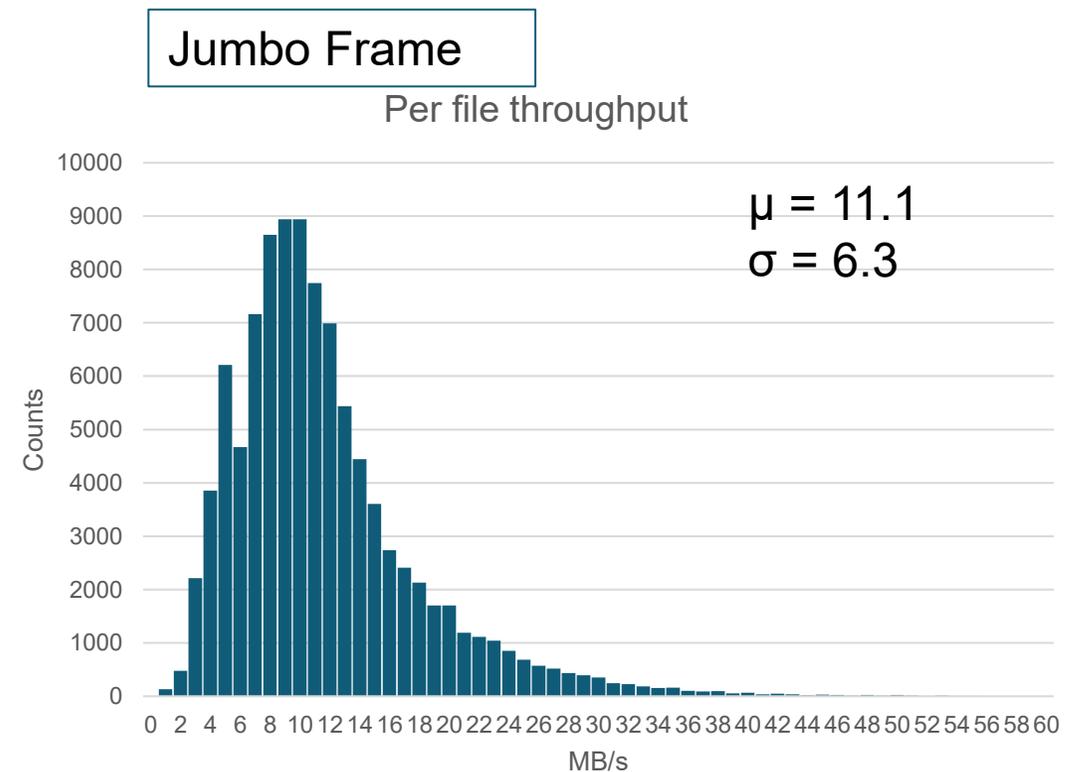
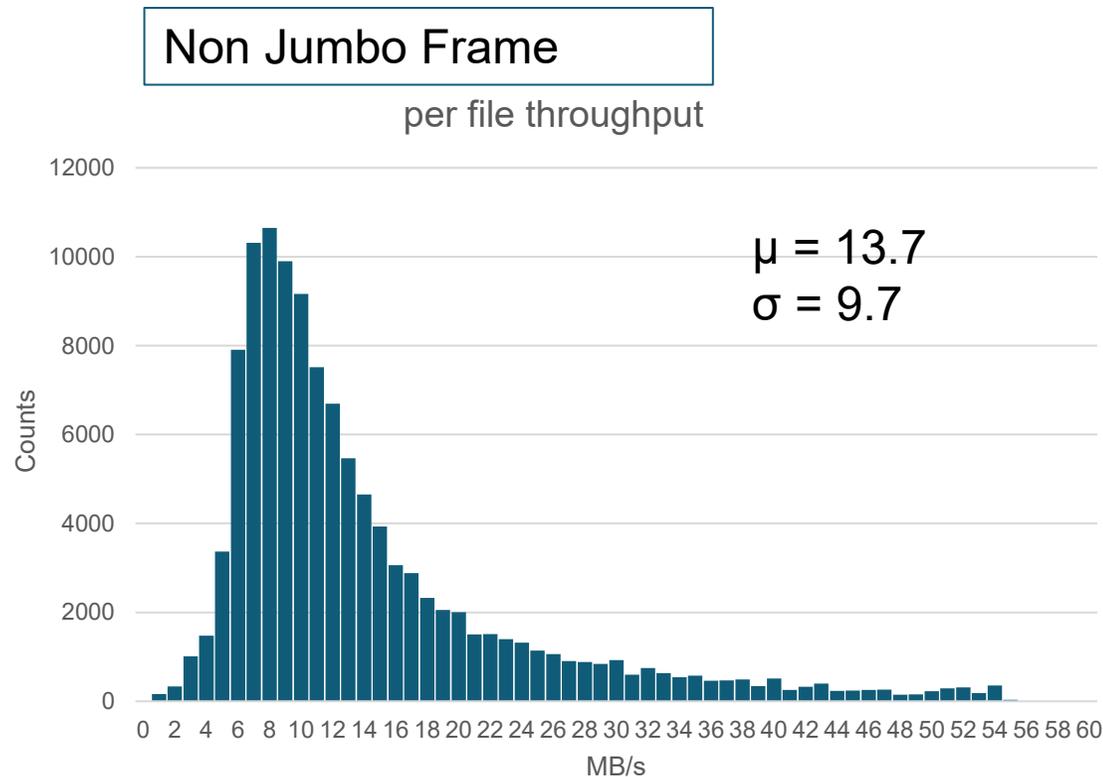


# Jumbo Frame

- MTU 1500 → 9000
  - To improve the data transfer efficiency.
    - Better performance in per-file transfer rate?
- Within US (T1 and T2s), (from BNL) RTT is in the range of 8 ms to 50 ms, which is relatively small.
  - The effect of the use of jumbo frame is relatively small.
- Across the trans-Atlantic or more distance, the effect should/could be visible.
  - RTT is 86.5ms to CERN from BNL.
- Issue.
  - In IPv4, all hosts in the same subnets must use jumbo frame.
    - IPv6 has dynamic MTU discovery

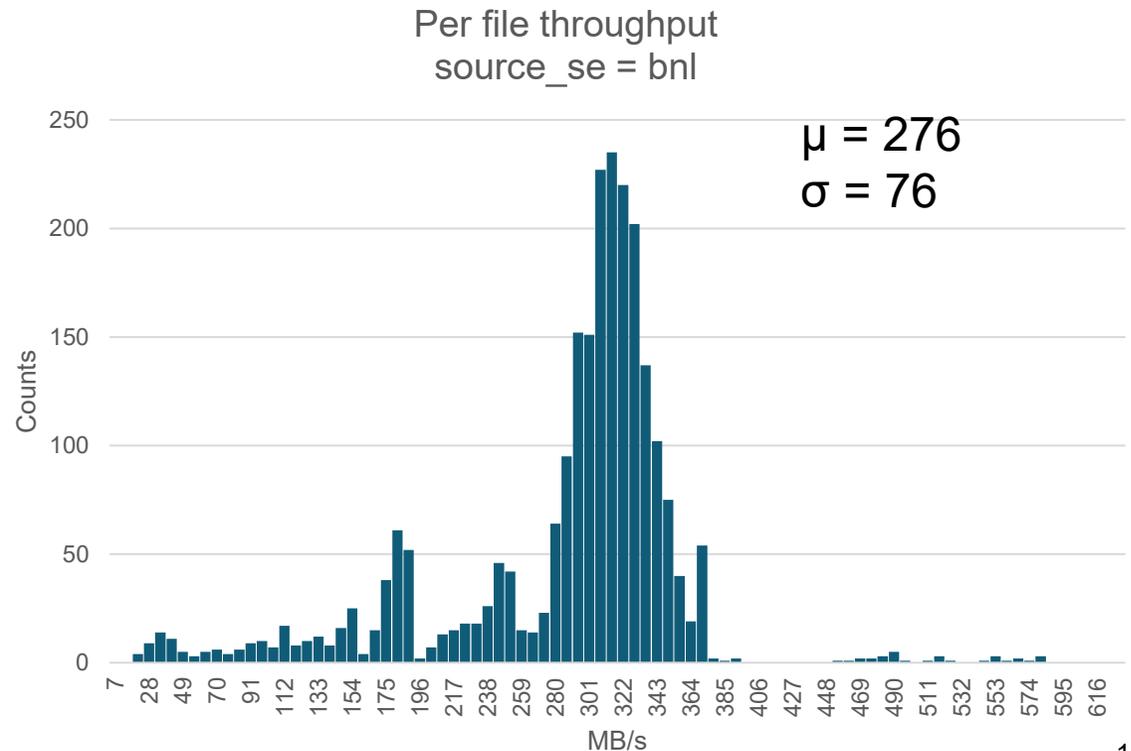
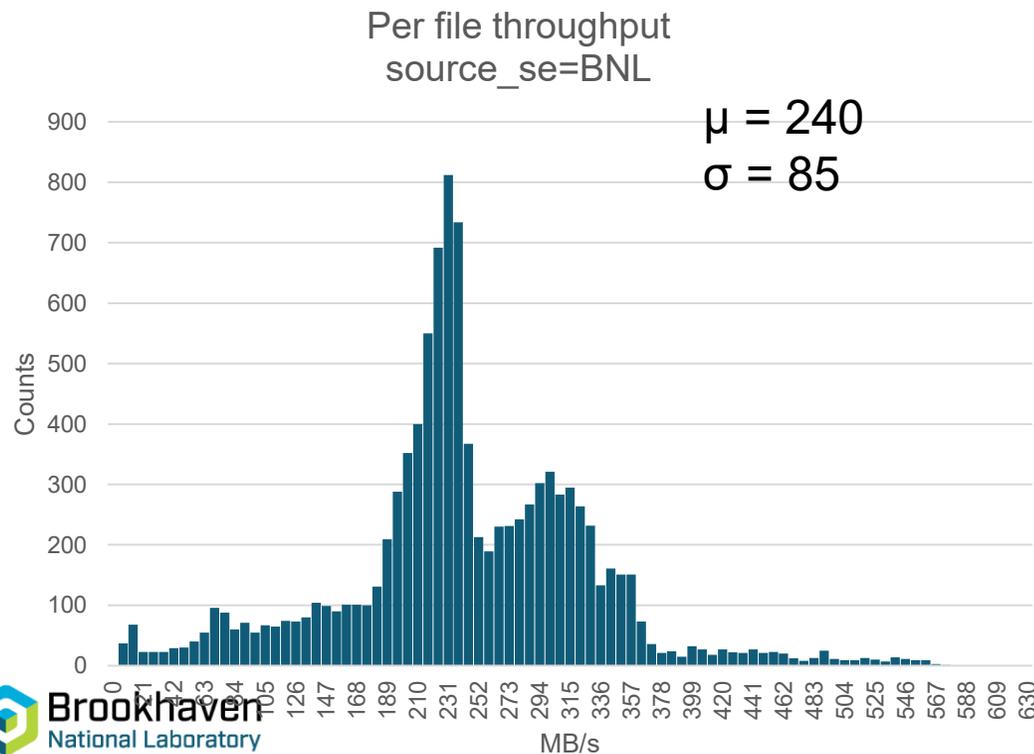
# Jumbo Frame at BNL

- BNL has enabled Jumbo frame in the 1<sup>st</sup> week of Feb 2025.
- Look into the transfer of the production data transfers to see if the effect is visible.
- Looking at FTS records of the file transfer with their file size larger than 1GB.



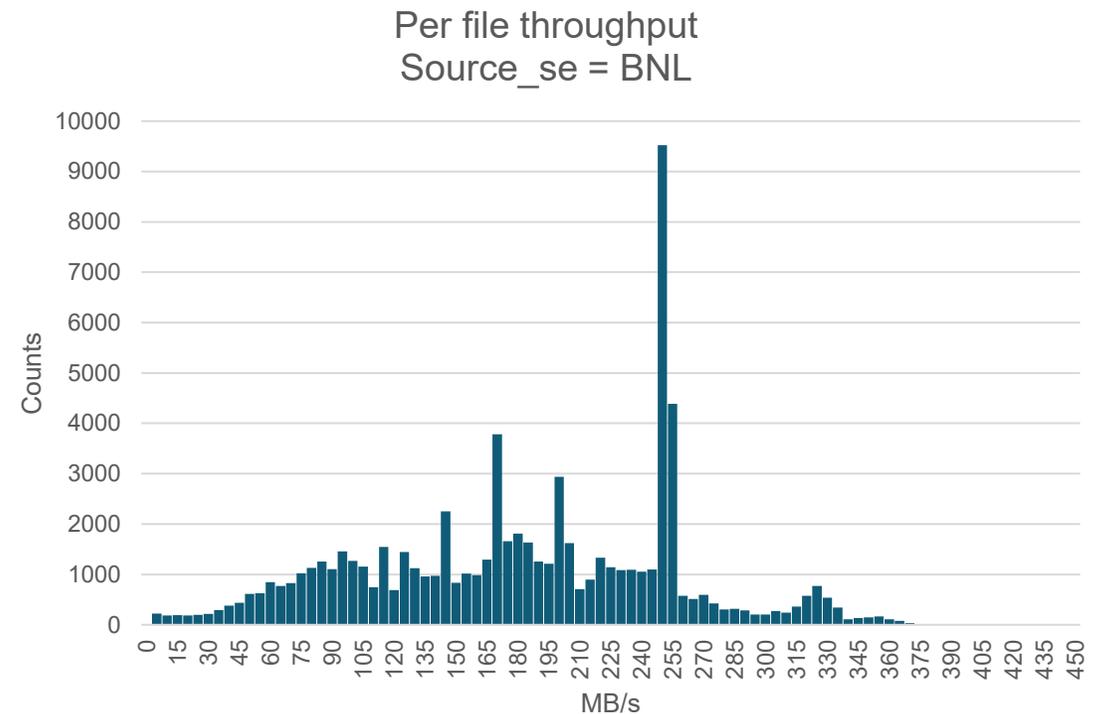
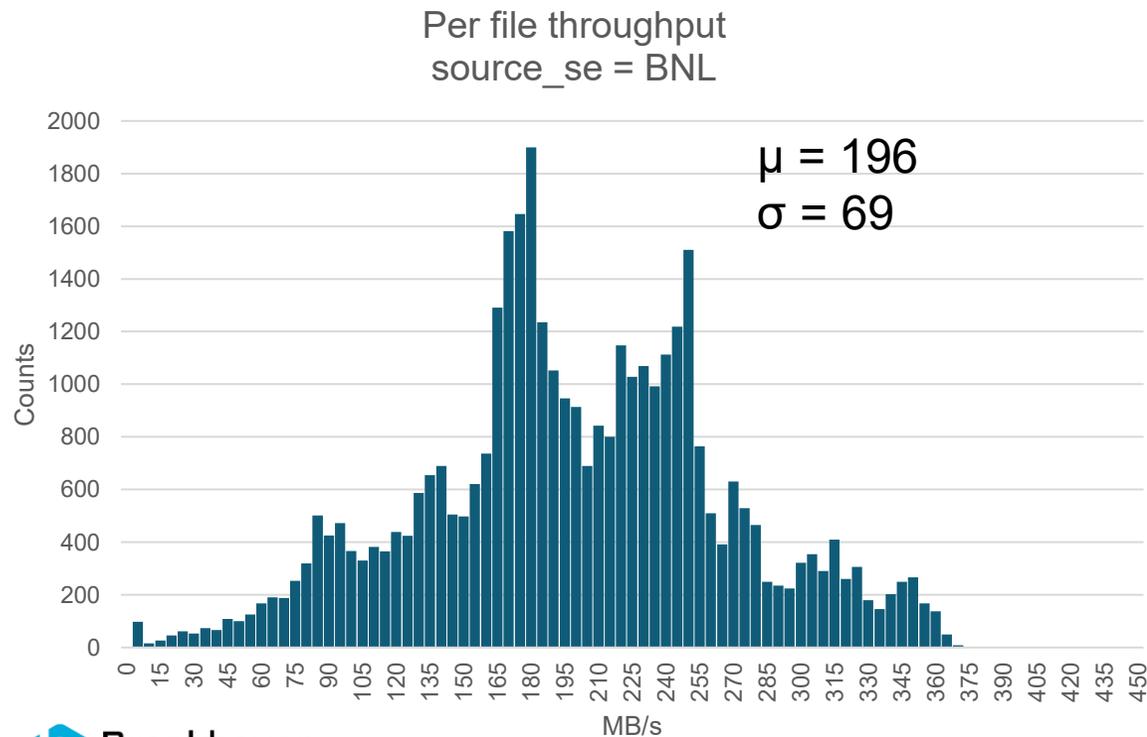
# Jumbo Frame at NET2

- Small time windows after the network upgrade (400Gbps) when it was not using jumbo frame and BNL enabled Jumbo frame.
- Looking at FTS records of the production data transfers from BNL with their file size larger than 1GB.



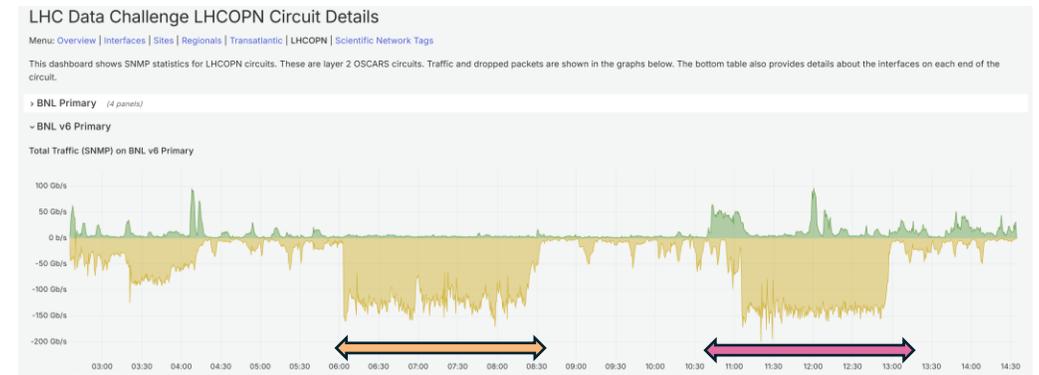
# Jumbo Frame at MWT2

- MWT2 has been using Jumbo frame for long time.
- Looking at FTS records for the production data transfers from BNL with their file size larger than 1GB.



# Further testing for BNL from CERN

- Series of the test from CERN to BNL with varying concurrency limit and on/off jumbo frame
- Max concurrency has bigger effect than jumbo frame.
- Effect of the jumbo frame seem to be visible.
  - ~10%?
- The total throughput at ~275 Gbps is achieved with large concurrency.



800 Max

3000 Max



# Token

- Token is coming for the data transfer.
- ATLAS and Belle II are planning to use long-lived access tokens in RUCIO transfer
  - FTS won't be dealing with requesting new access token.
    - Avoid generating many access tokens
    - --unmanaged-token option in FTS API
- The load generator needs to be modified to use long-lived token option.
- Testing transaction rate issue.
  - The use of token was causing issue during the last DC test in FTS.
    - New version 3.14.x fix the issue
    - FTS 3.14.0 has been deployed at BNL
  - Long-lived token will/should avoid FTS token generation issue.
  - Does storage have any rate limitation caused by token access?
    - Testing by transferring small files at high rate.

# Discussion

- Next round of tests.
- Proposed US ATLAS tests.
  - Capacity test in the week of June 23<sup>rd</sup> using the token (and X509 for comparison)
  - (small) token transaction rate test in the week of June 9<sup>th</sup> or as soon as the load generator is updated to use tokens.
- Any other?
  - Other Non-US ATLAS sites.
  - Together with (US)CMS?