# FTS Service Status at BNL

Hironori Ito

XRootD and FTS Workshop at STFC UK

Sept 10th , 2024

@BrookhavenLab

# Current Status

- FTS Agents
  - 9 hosts
    - 64 GB RAM
    - 16 CPU threads
  - RHEV Clusters
    - Scalability
      - Horizontally (more VMs)
      - Vertically (more CPUs, more RAMs, etc…)
    - Resiliency
  - RHEL 7
  - Issues
    - RHEV Ending soon (2026)
    - RHEL7 (VM support will end soon)
    - New FTS is not supported
- FTS DB
  - Physical host
  - NVME SSD 2.2 TB
  - 256GB RAM
  - MySQL 8.0
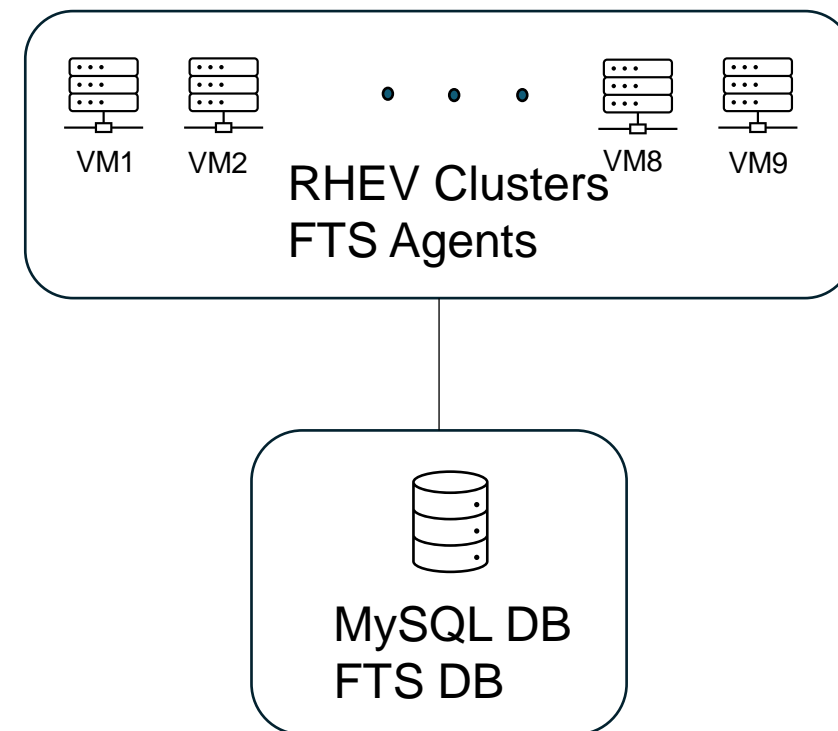
# Future Setup

- FTS Agents
  - 9 or more hosts
    - 64 GB RAM
    - 16 CPU threads
  - **Openshift**
    - <u>In progress</u>
    - Scalability
    - Resiliency
    - **Container** support
  - **RHEL 9** and beyond
    - In progress
- FTS DB
  - Physical host
  - NVME SSD 2.2 TB
  - 256GB RAM
  - MySQL 8

**Openshift**
FTS Agents

VM1  VM2  ...  VM8  VM9

MySQL DB
FTS DB
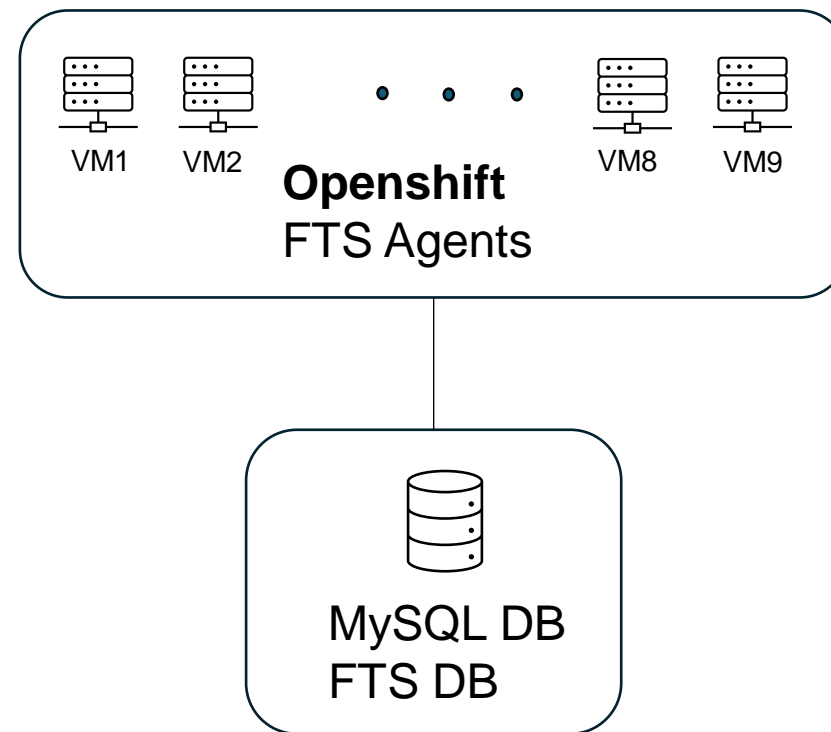
# ATLAS & Belle II VOs

- BNL FTS serves primarily ATLAS and Belle II
- ATLAS
  - North American sites
- Belle II
  - All transfers except RAW data SEs
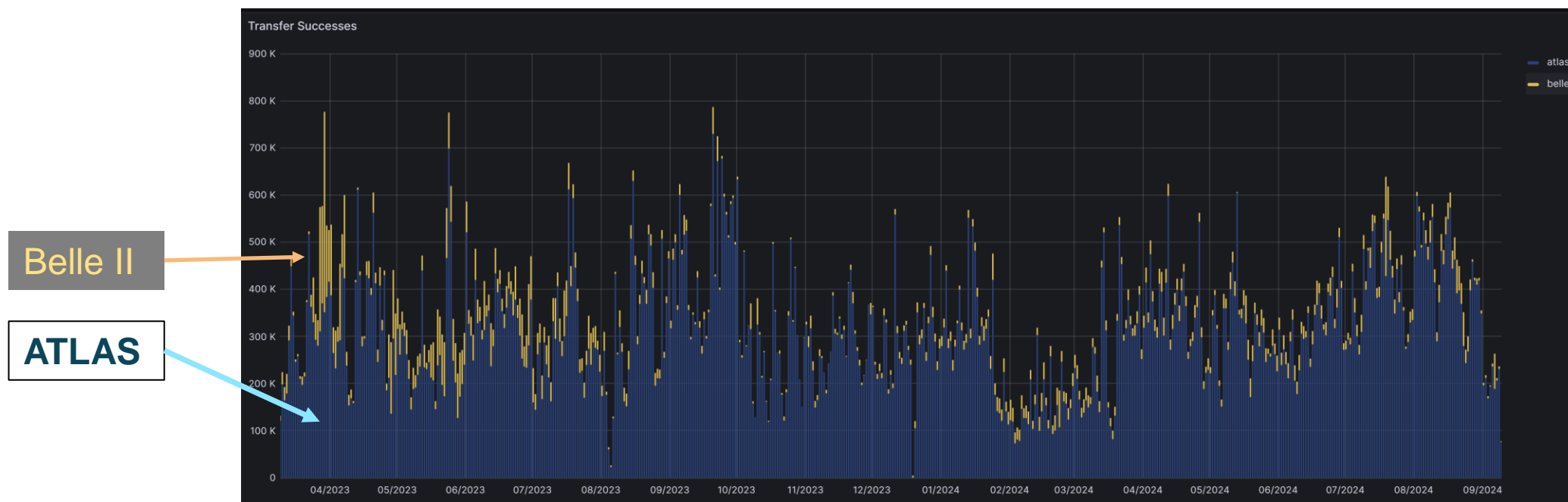
# Mini Data Challenges in US

- Mini Data Challenges
  - US ATLAS
    - Prior to DC24, there were series of data challenges in US
      - Identify the current capabilities and deficiencies to prepare US sites for HL-LHC
      - Testing upto the full network and/or storage capabilities
        - Higher than DC target
      - Plan to do periodic tests to check any issues and monitor progress
      - Do tests on demand by sites
      - Using the load generator script
  - Joint mini challenges with US CMS
    - NOTE: CMS has their own FTS
    - Identify the multi VO issues on network
      - Any chock points?
- BNL FTS has been also used for Belle II mini data challenge
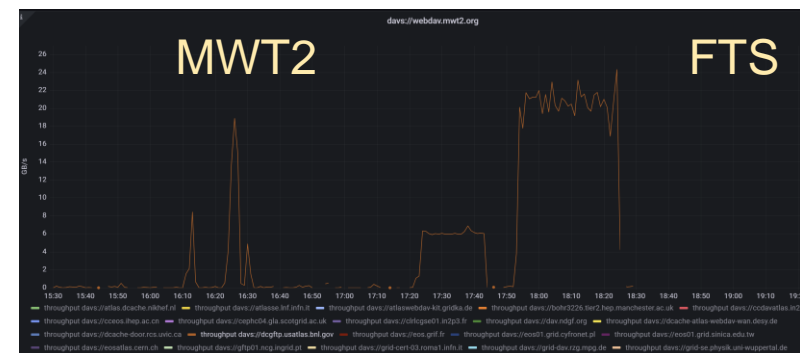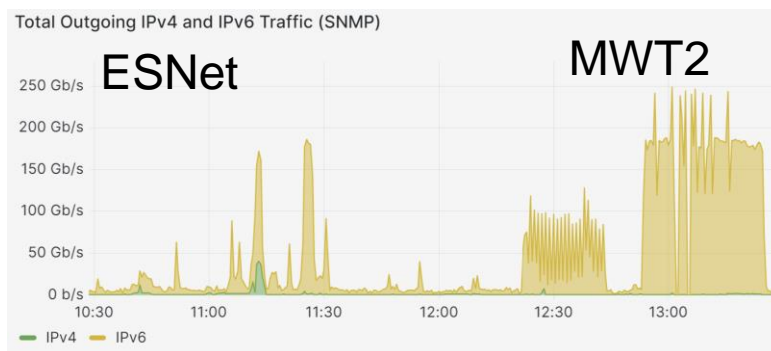  - To check the site performance

# USATLAS T2 Mini Tests

- AGLT2
  - Destinations
  - 200 Gbps
- MWT2
  - Destinations
  - 200 Gbps
- BNL
  - Source
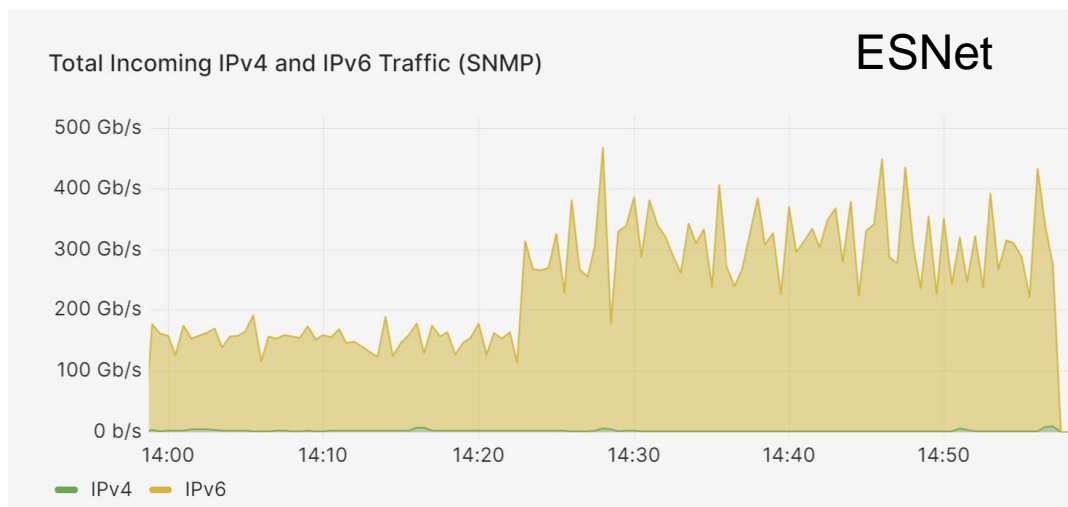  - 2x400 Gbps (at the time of this test. It is now 4x400)

AGLT2 ~ 150Gbps
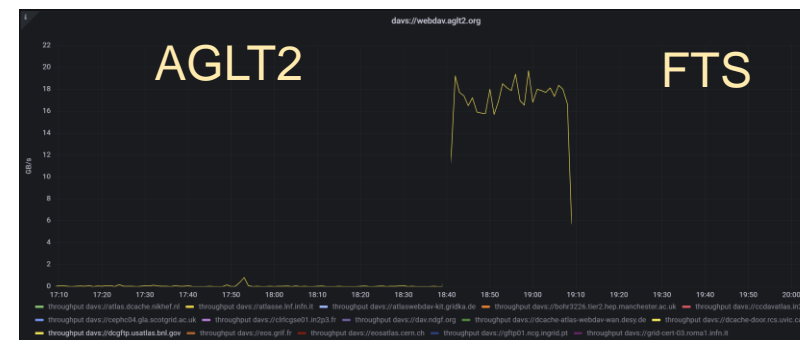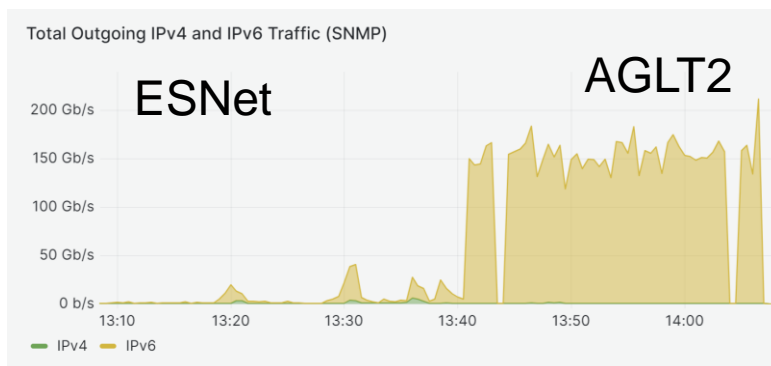


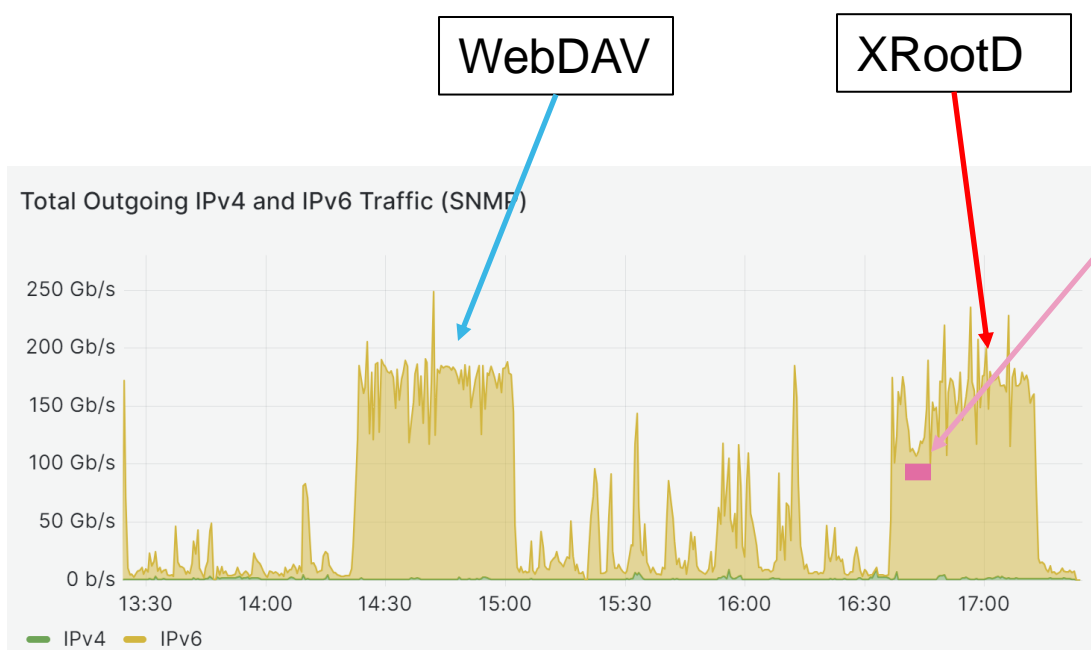MWT2 ~ 180Gbps

# US ATLAS T1 Mini Test

- BNL as source
  - BNL had 2x400 Gbps at that time of this test. (Now, it is 4x400 Gbps.)
- AGLT2 and MWT2 as destinations

# WebDAV vs XRootD

- XRootD is not primarily used for storage-to-storage WAN transfers in ATLAS and Belle II.
- ATLAS uses XCache to cache the data at a site.
  - XCache can trigger large data transfers.

- There are no differences in observed maximum throughput between WebDAV and XRootd
- Dip in the throughput seen during XRootD testing was caused by the small number of failed transfers
  - XRootD door is not well monitored
  - Small number of failures have a big impact to the maximum throughput due to FTS optimizer regardless of the protocol.

WebDAV

XRootD



Total Outgoing IPv4 and IPv6 Traffic (SNMP)

250 Gb/s
200 Gb/s
150 Gb/s
100 Gb/s
50 Gb/s
0 b/s

13:30  14:00  14:30  15:00  15:30  16:00  16:30  17:00

— IPv4  — IPv6
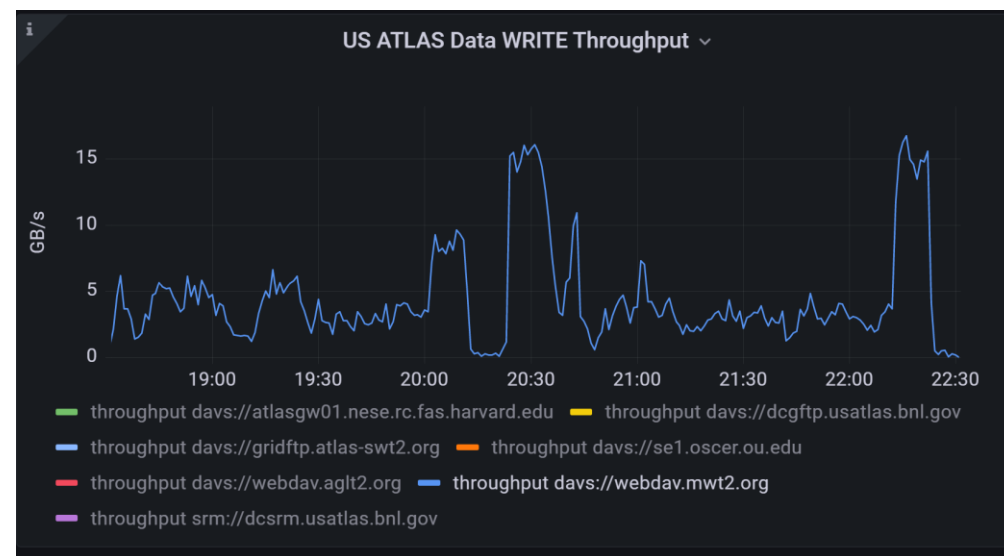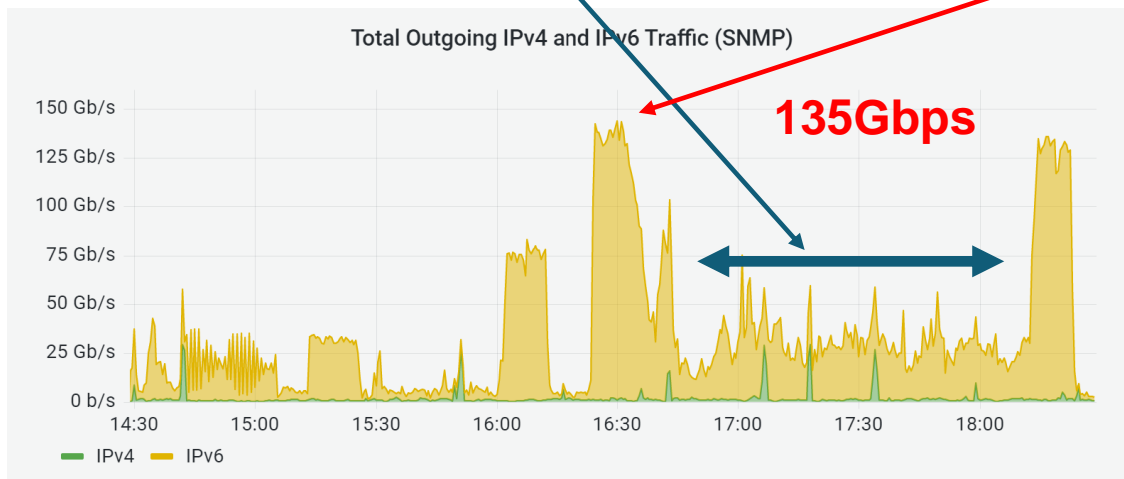
**Brookhaven**
National Laboratory

# Effect of slow transfers



Slower Production Transfers from various sites prevent the throughput test transfers due to the max concurrency setting for MWT2 SE (400 at the time)

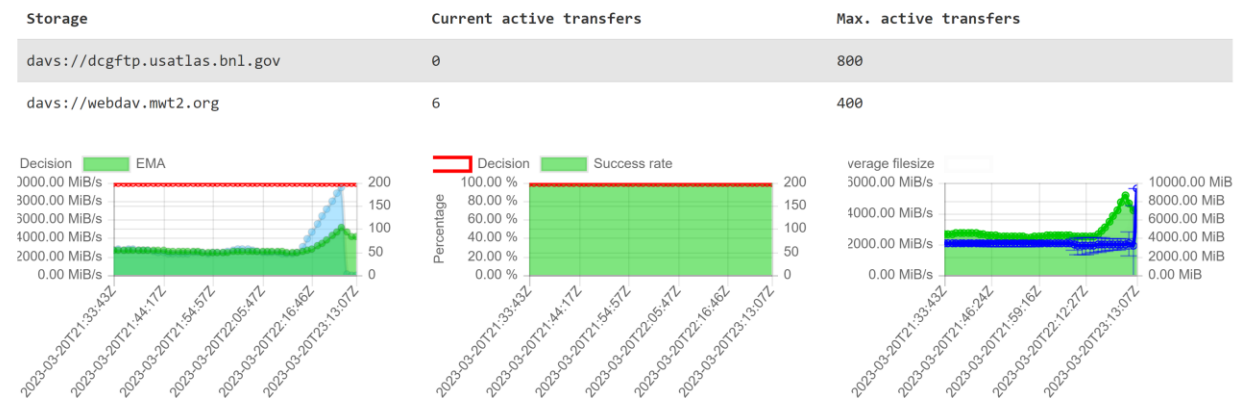Max 135Gbps is achieved once the slow transfers are done.

**US ATLAS Data WRITE Throughput**

- throughput davs://atlasgw01.nese.rc.fas.harvard.edu
- throughput davs://dcgftp.usatlas.bnl.gov
- throughput davs://gridftp.atlas-swt2.org
- throughput davs://se1.oscer.ou.edu
- throughput davs://webdav.aglt2.org
- throughput davs://webdav.mwt2.org
- throughput srm://dcsrm.usatlas.bnl.gov

ESNet Monitor

**Total Outgoing IPv4 and IPv6 Traffic (SNMP)**

**135Gbps**

IPv4   IPv6

Transfers were successful

Details for davs://dcgftp.usatlas.bnl.gov → davs://webdav.mwt2.org

| Storage | Current active transfers | Max. active transfers |
|---|---|---|
| davs://dcgftp.usatlas.bnl.gov | 0 | 800 |
| davs://webdav.mwt2.org | 6 | 400 |

Decision  EMA

Decision  Success rate

verage filesize

Brookhaven National Laboratory

9

# Some other fun stuff

- Some other clients for GFAL and FTS
  - Written purely in Ruby and Crystal
    - webdav-ls.rb, webdav-rm.rb, ftsDelegate.rb, ftsSubmit.rb, etc…
    - Load Generator for mini data challenge
      - NetworkLoadGen.rb
      - Submitting FTS requests periodically with the list of files with the data size corresponds to desired data volume periodically.
      - Token not supported yet.
  - They are found at
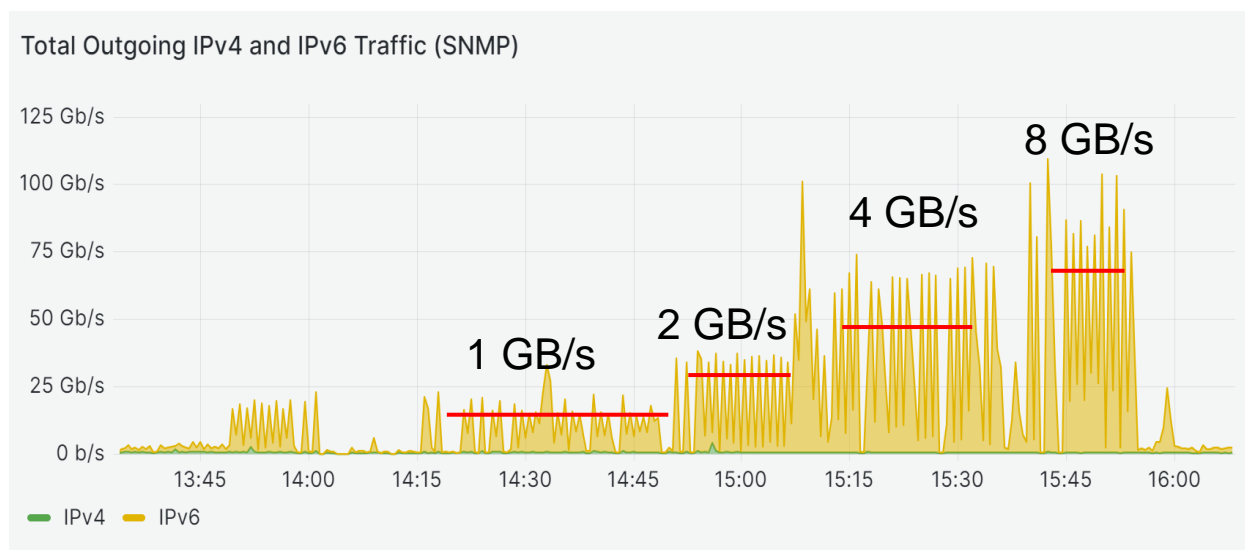    - https://bnlbox.sdcc.bnl.gov/index.php/s/XGs6LJEGNzf69zK

# Load Generator

Load generator can generate consistent network load at the desired rate.
Network and BNL FTS monitor shows the same throughput

BNL FTS Monitor
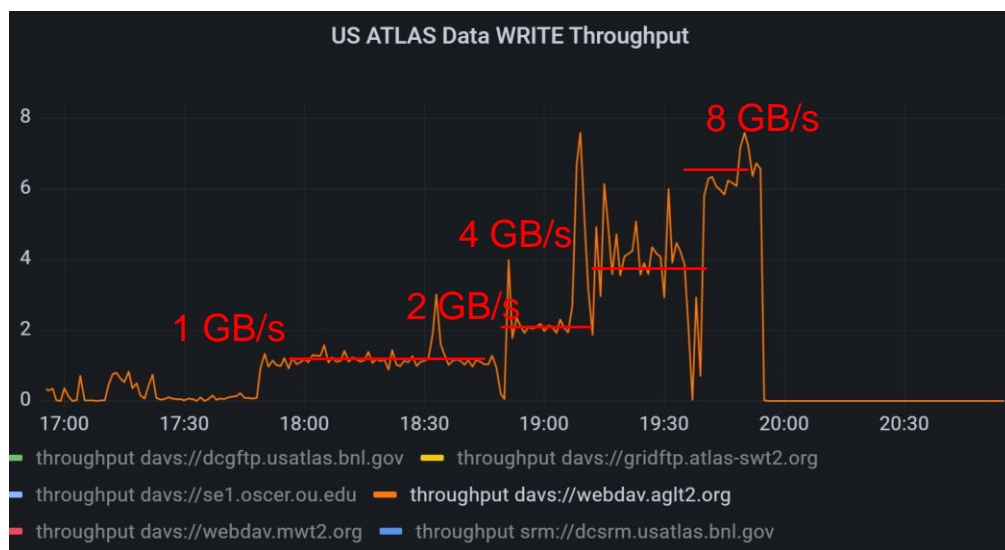
Network Monitor

# Conclusion

- BNL FTS has been in production successfully for two experiments; ATLAS and Belle II
  - It helps to have more than one FTS
    - ATLAS:  CERN and BNL
    - Belle II:  KEK and BNL
    - As a backup
    - For testing
    - Split loads and functions

- BNL FTS has been used for data challenge
  - US mini data challenge will continue