DUNE Community Report

Yuyi Guo for the DUNE Collab

The 7th Rucio Community Workshop
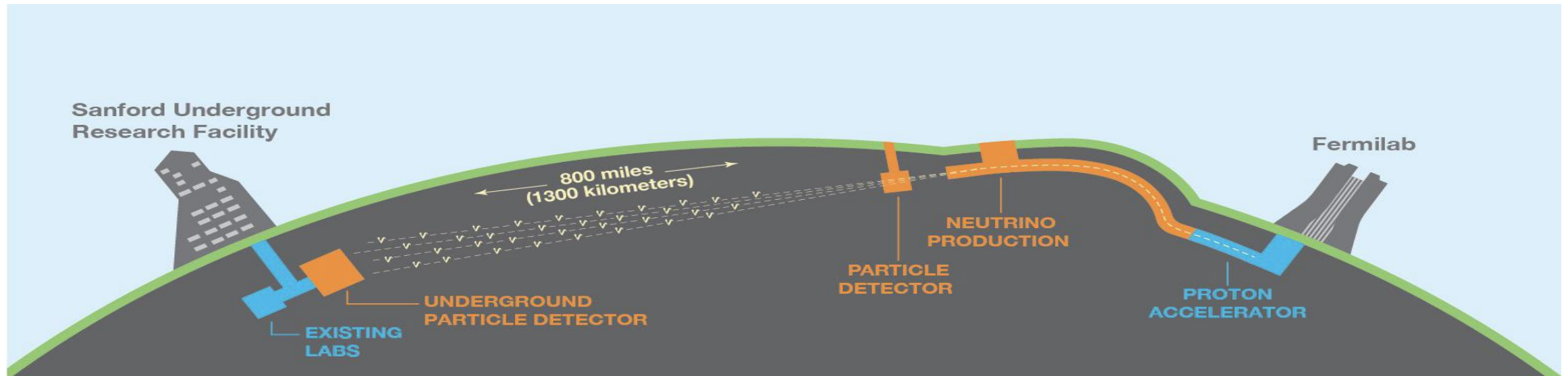
September 30, 2024

Fermilab DUNE

Outline

- DUNE Data Model

- Big Data Tests

  - Data Challenge Feb 24

  - Dress Rehearsal Apr 24

- Data Taking

  - ProtoDUNE-HD Data Taking at CERN

  - 2x2-Minerva LAr prototype testing

- Rucio Deployment

- Plan for 2024-2025

- Requests from DUNE

Fermilab DUNE

DUNE Data Model

Yuyi Guo | The 7th Rucio Community Workshop

About DUNE

- Neutrino experiment studying neutrino oscillation parameter (mass ordering, matter vs antimatter asymmetry, unitarity), proton decay, supernova neutrinos, and more.

- Four large LAr TPC (17 kT) at 4850 ft underground in Lead, SD (Homestake Mine)

- Near detector being constructed onsite at Fermilab (3 sub-detectors, two that move)

- Two prototypes at CERN - (ProtoDUNE II Horizontal Drift - ProtoDUNE II Vertical Drift)

Data Pipeline Diagram
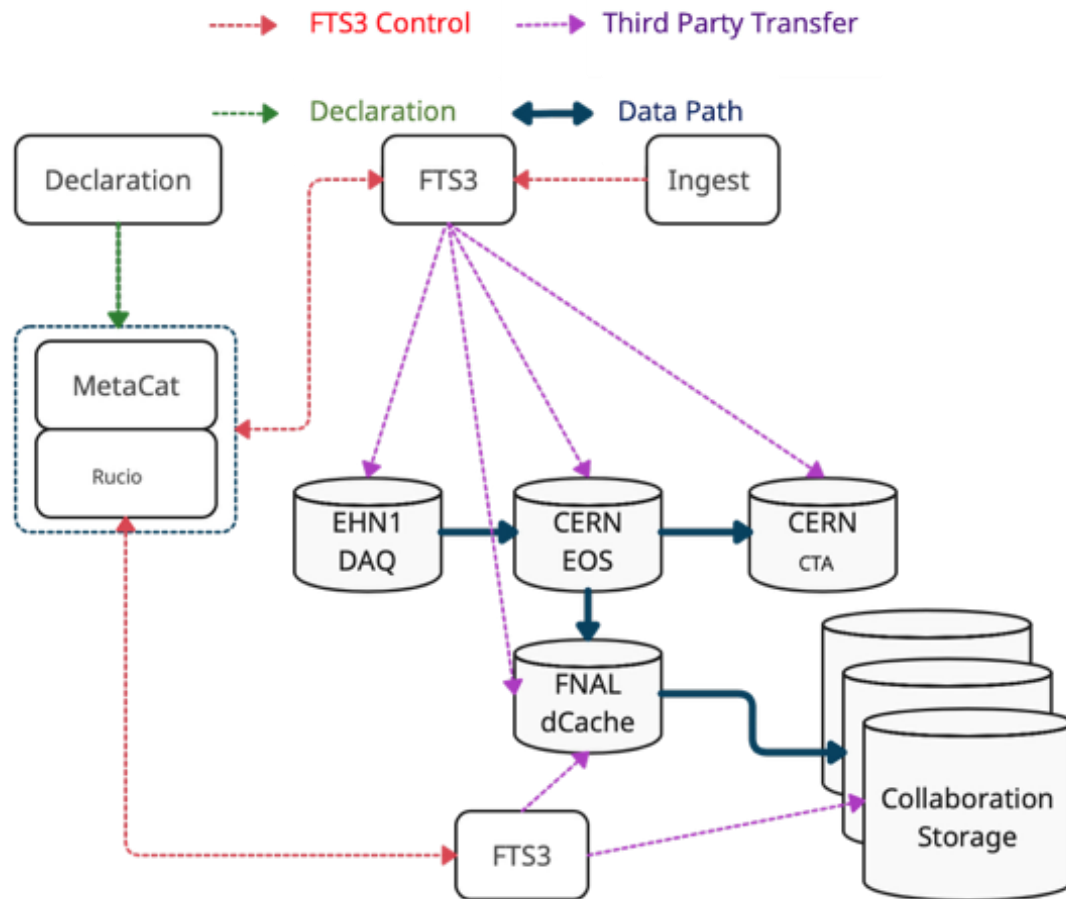
- Ingest Daemon and Declaration Daemon
  - Ingest daemon transfers files from experiment systems to dropbox
    - Can operate without connection to Fermilab
  - Declaration daemon declares data to Metacat and Rucio, copies data to the Rucio RSE and makes rules to get information to the final destinations
- 2 copies of raw data on tape, one at Fermilab Enstore and one at CERN CTA
- 1 copy of sim/reco on tape
- 2 copies of sim/reco on disk distributed across global storage elements
- Adapted for all DUNE data taking

Fermilab DUNE

Big Data Tests

DUNE in WLCG Data Challenge 24 1/5

- Three Tests were planned

  - Test 1 - "FD" Raw Data

    - Simulate the archival of 25% of the raw data rate from the Far Detector

    - translates to 1 GB/s from BNL (stand in for SURF) to FNAL

    - replicate that "FD" raw data to archival storage facilities around the world

    - replicate the "FD" raw data to disk storage elements around the world for prompt access from compute elements

  - Test 2 - "FD" Raw Data keep up processing

    - Maintain continuous processing workload at distributed sites commensurate with 25% "FD" raw data rate (1 GB/s)

      - Utilize compute resources across sites in Europe and North America

      - Match the locality of jobs with locality of data at nearby RSEs

  - Test 3 - SuperNova Raw Data rapid transfer & processing

    - 3.5 GB/s SURF (BNL) to FNAL to NERSC

    - Unfortunately, test3 was not done due to network congestion from other tests.

🔷 Fermilab DUNE
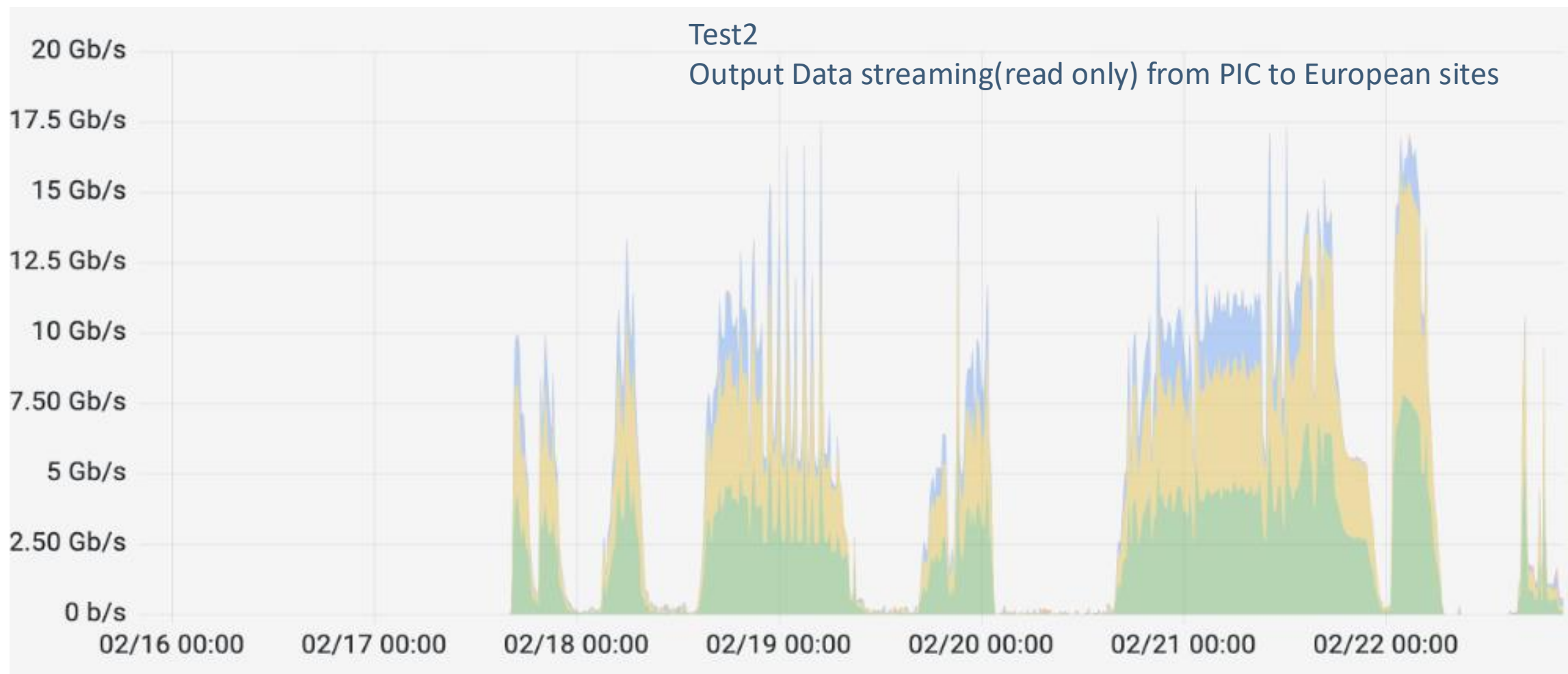
DUNE in WLCG Data Challenge 24 2/5

Test 2 – "FD Raw Data"
Raw processing
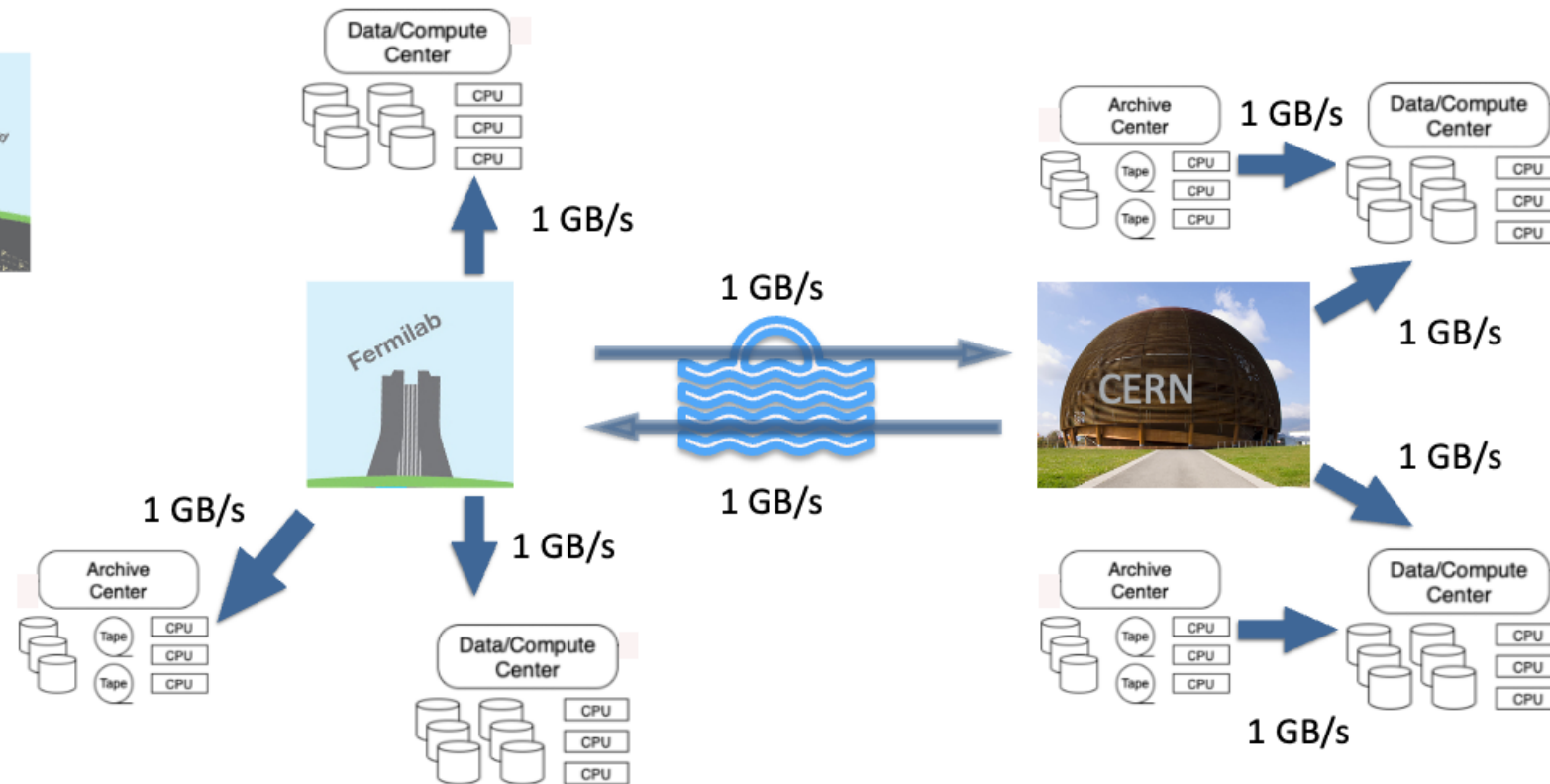
- Maintained continuous processing workload at distributed sites commensurate with 25% "FD" raw data rate
  - utilized compute elements across the WLCG and OSG
  - matched the locality of jobs with locality of data at nearby RSEs
- **Job submission used token authentication/authorization, X509 proxies was for Data transfers**
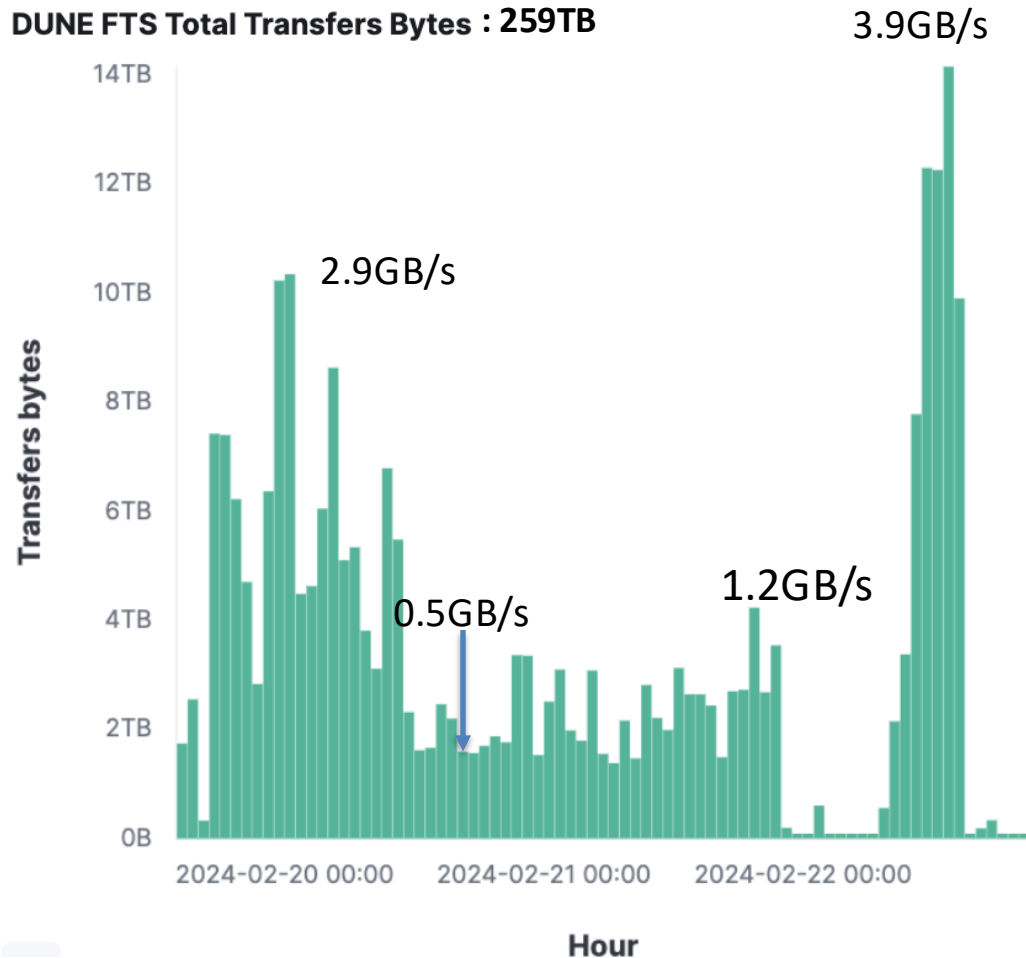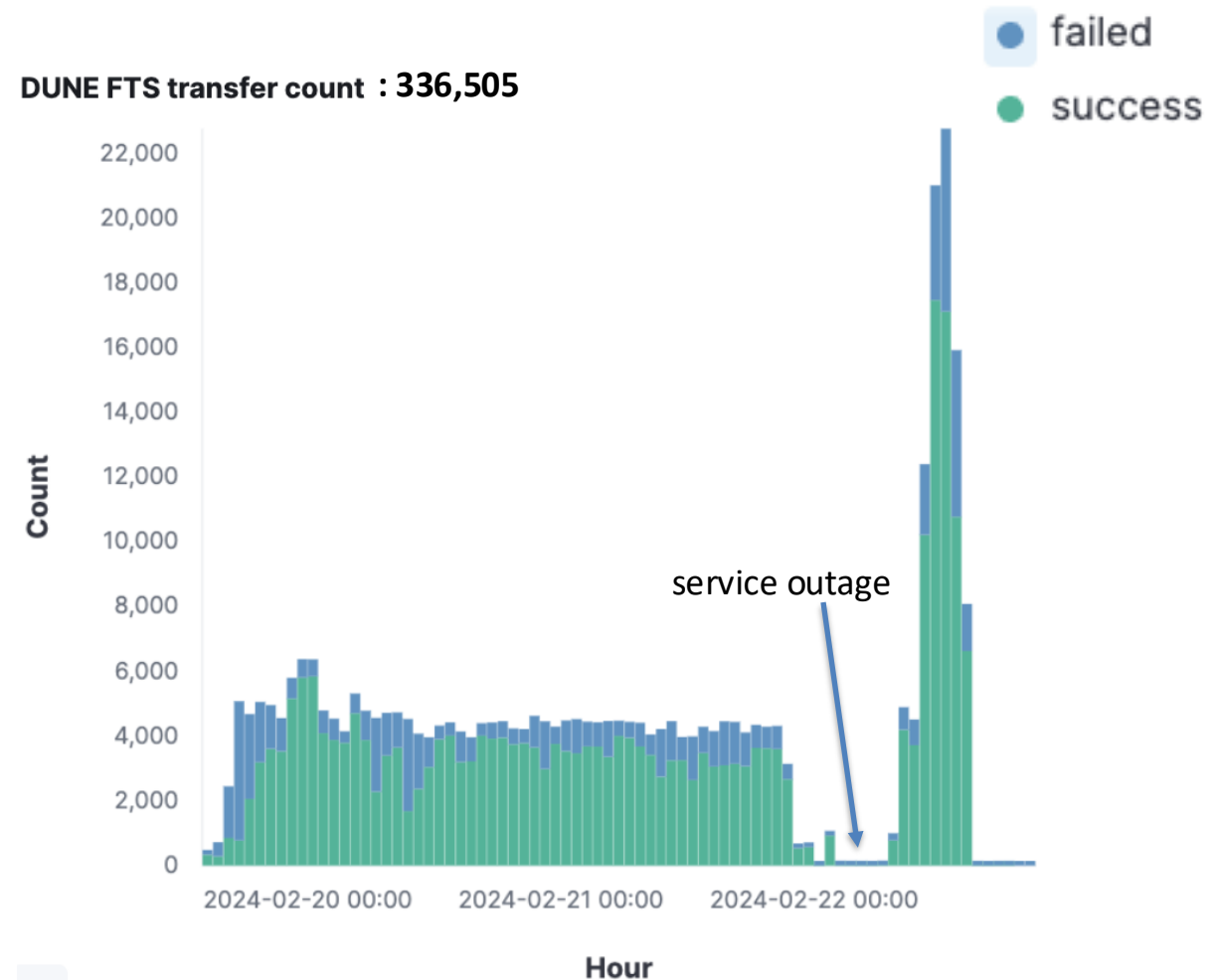
DUNE in WLCG Data Challenge 24 3/5

Test2
Output Data streaming(read only) from PIC to European sites

🔷 **Fermilab** DUNE

DUNE in WLCG Data Challenge 24 4/5

**Test 2**

DUNE in WLCG Data Challenge 24 5/5

- Lessons Learned/Conclusions
  - Initial Rucio mis-configuration
    - Wrong priority in RSEs lead to many TPC transfers using xrootd instead of davs. (many timeouts at RAL)
  - Increased Rucio server resources and optimized the settings
    - Rucio/Metacat scale to 16k simultaneous jobs
    - Increased Rucio server from 1 , 2, 4 , eventually 16 servers during the tests.
    - Increased Rucio DB connection to 500.
  - a bug in Rucio client found and fixed now
- Identified services that need to be hardened.
- DUNE used MC production reconstruction (reco2) as part of the DC24 activity.
- DC24 extremely helpful as a stress test for our processing system.

‡ Fermilab   DUNE

ProtoDUNE Dress Rehearsal 1/3

- The April 2024 dress rehearsal ran the DAQ for three days

  - April 18: 6 TB   (run 25077)

  - April 22: 34 TB (run 25101) 4 hrs

  - April 24: 27 TB (run 25136) 3.5 hr

- Data transfer rates from EHN1 to EOSPUBLIC approximately 7 TB/hour at expected beam rates.

  - Expect to need only 5.6TB/hour

- Keep-up processing

  - At past ran up to 20k simultaneous jobs in the JustIN and the DUNE global pool.

  - The rehearsal ran about 4k of HD-Protodune keep-up processing due to the high memory requirement of these jobs.
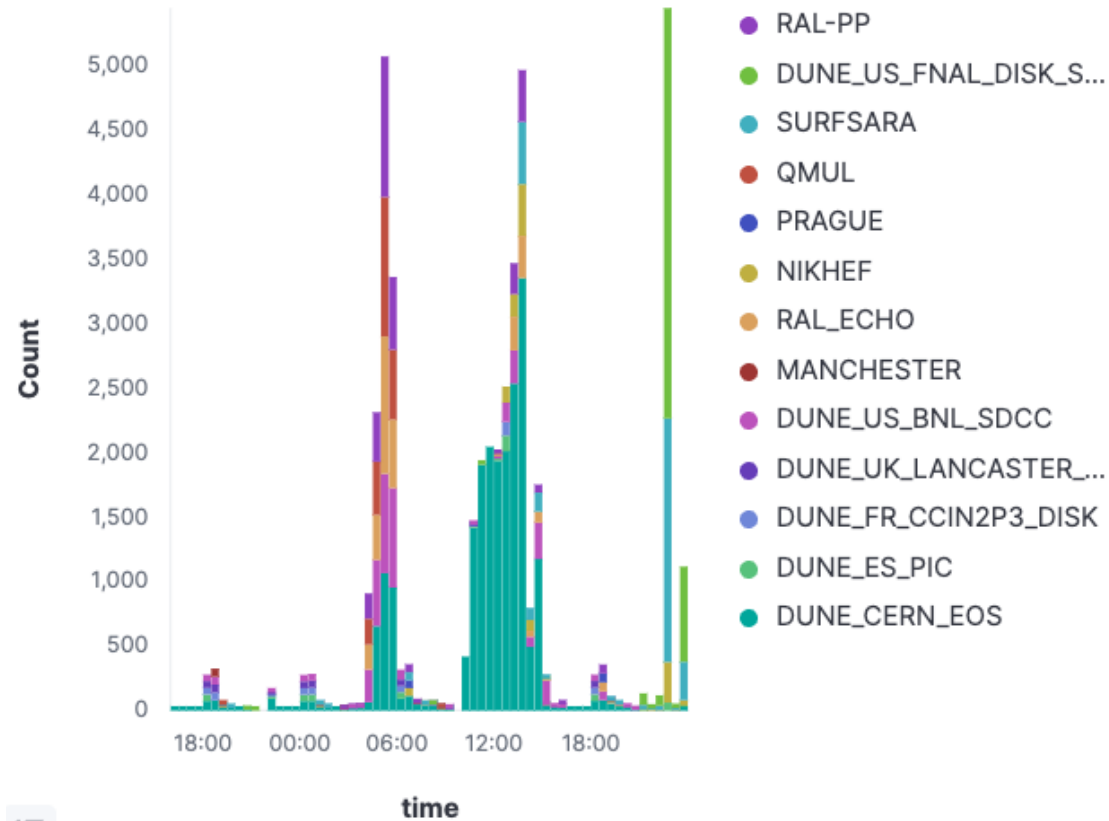
‡ Fermilab DUNE

ProtoDUNE Dress Rehearsal 2/3

[rucio] Total Transfers vs Time

Total 185.4TB
Transferred Bytes

🎲 **Fermilab**   DU(VE)
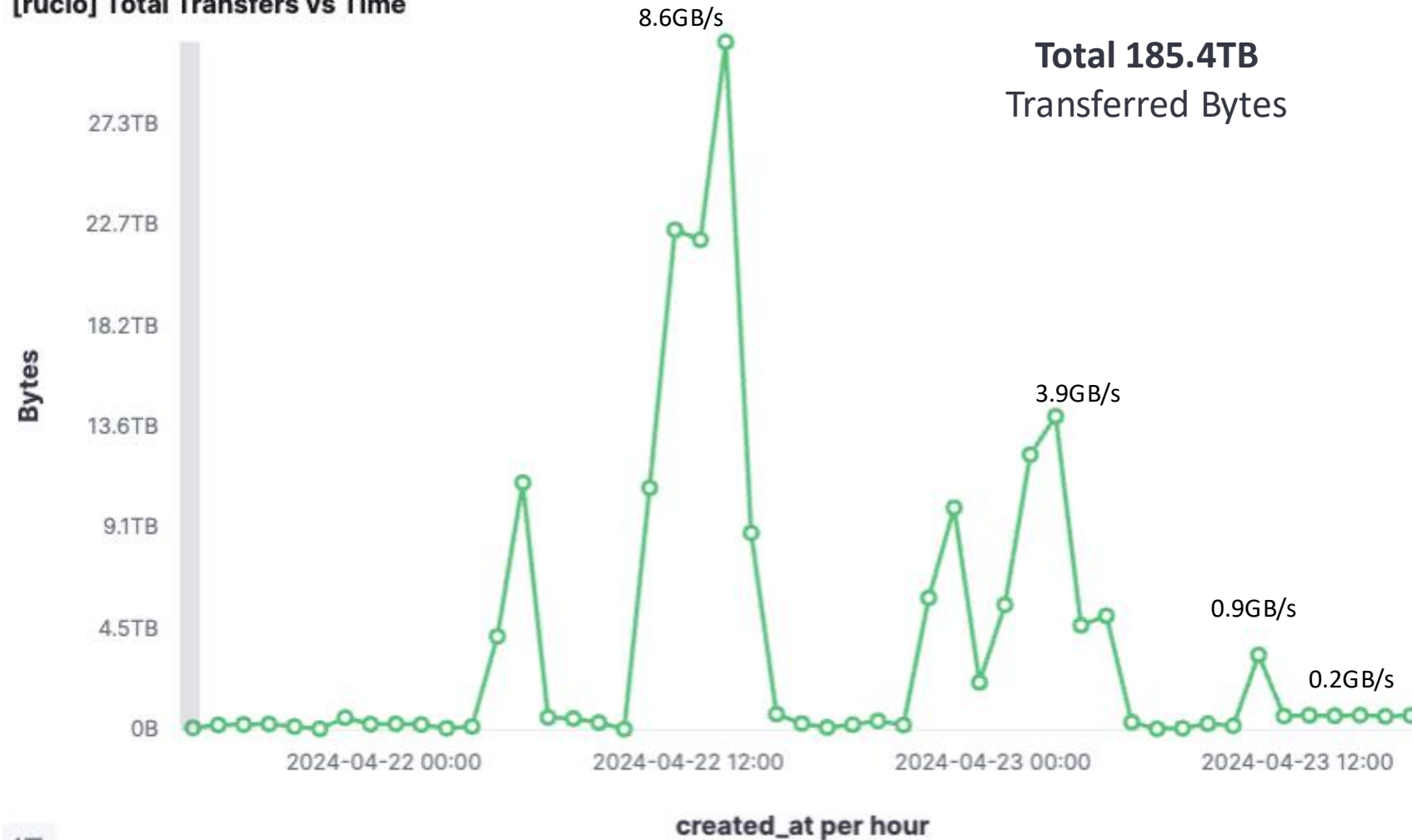
ProtoDUNE Dress Rehearsal 3/3
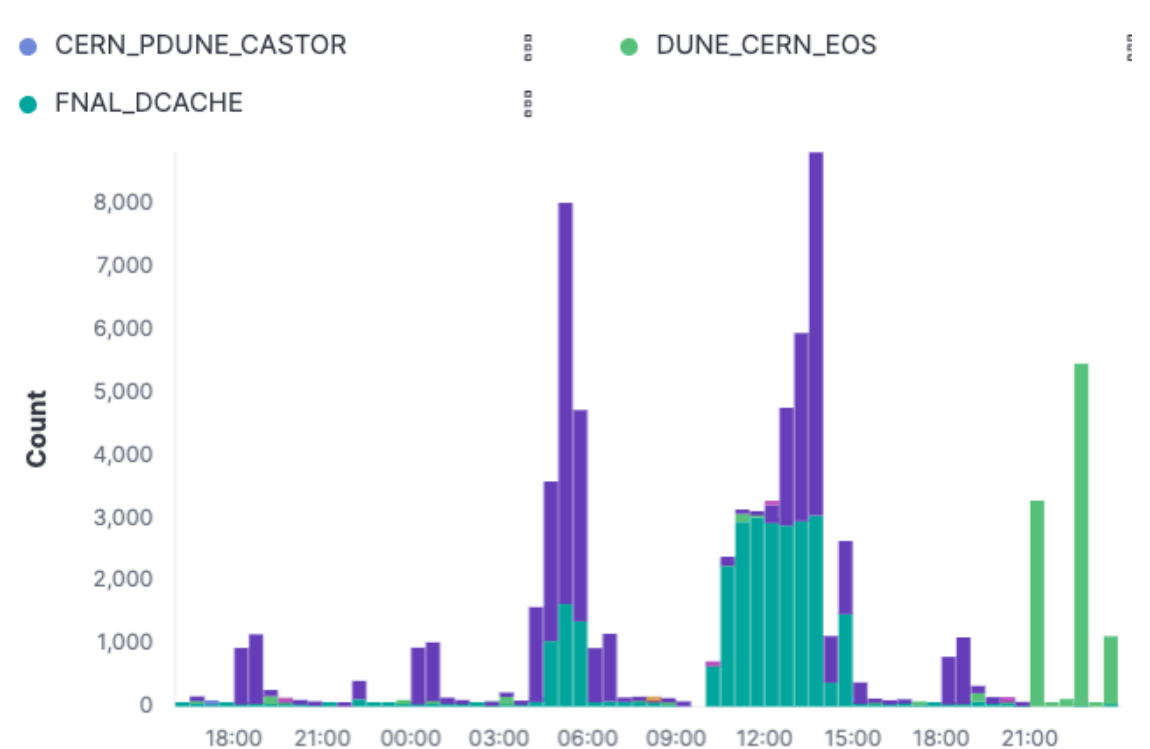
April 22, Run 25101

Data Taking

ProtoDUNE-HD Data Taking at CERN 1/2

- It was from June 19, 2024 to September 16, 2024

- Data volume

  - 2943 TB of Beam TPC Data

  - 727 TB of Cosmics TPC Data

  - 872 TB of Trigger Primitives

- Some raw data files moved to BNL, PIC, NIKHEF, FNAL, PRAGUE, RAL-PP, SURFSARA due to space shortage at CERN

🍀 Fermilab   DUNE

ProtoDUNE-HD Data Taking at CERN 2/2

[rucio] Total Transfers vs Time

Total transferred bytes 7.7PB
Total transfers 2,545,842
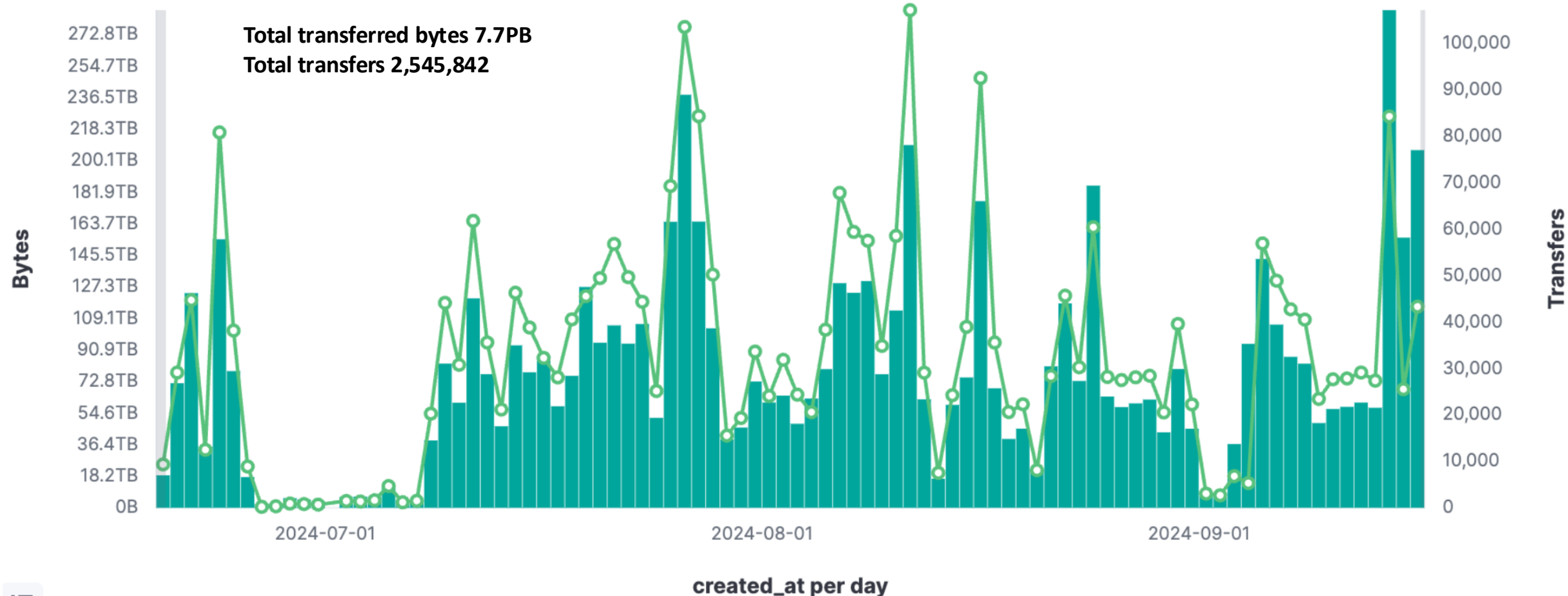
2x2-Minerva LAr prototype testing

- Data pipeline developed for ProtoDUNE was replicated at Fermilab for 2x2-Minerva LAr prototype

- Approx. 1 week of good data taken with neutrino beam just before July shutdown

- 1.2TB of Minerva chambers data

- 13.2TB of LAr Light readout

- 0.631TB of LAr charge readout.

- **The data pipeline can be used by all the DUNE tests and data taking.**

🐝 Fermilab   DUNE

Conclusions of Data Taking

- Handled more than 2x our initial estimate of data volume

- Saw 100Gbit/s between CERN and Fermilab for max rate, during keep-up

- Saw peak 26TB/hr to CERN CTA. (well over design parameters)

- **Rucio was doing its jobs.**
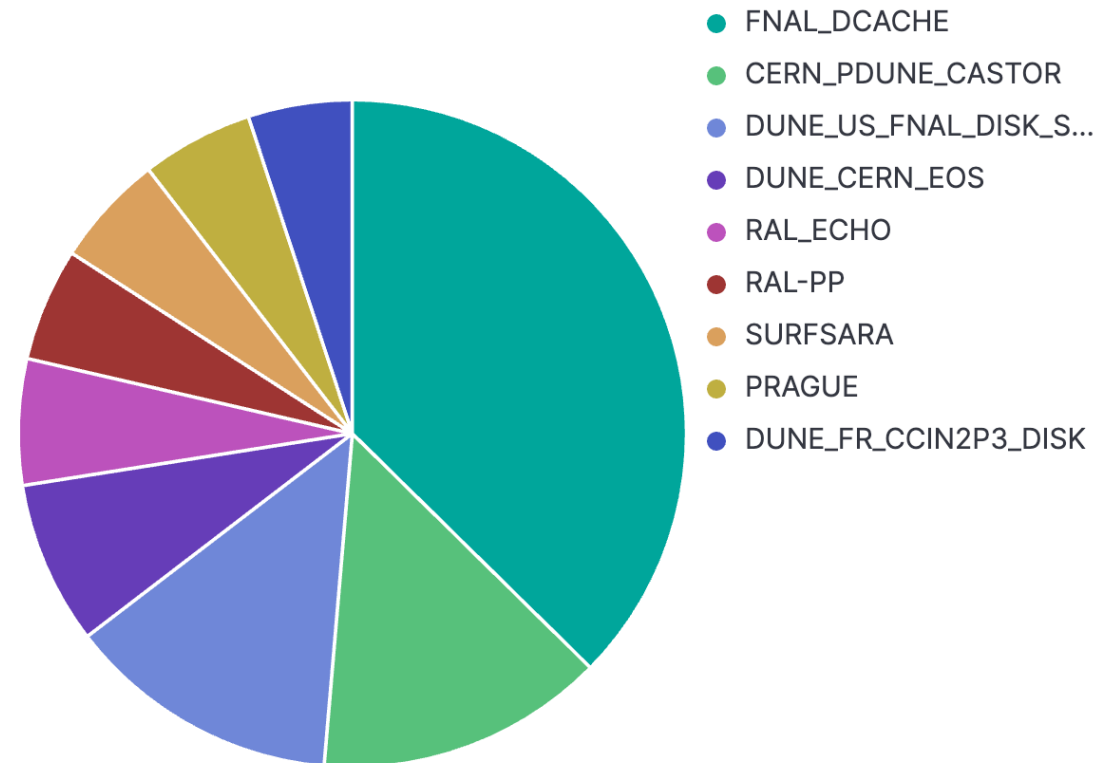
DUNE Rucio Deployment

- Running version 34 of Rucio

- Running on the Fermilab OKD cluster

  - Simple spin up of additional pods for server and daemons

  - Currently running 16 server pods and 2 reaper pods to handle data challenge loads

- Since the last Rucio Workshop, transition has been completed from a custom deployment framework to one based on use of Kustomize

  - Improved flexibility of configuration and speed of version upgrades

  - Greatly reduced chance for errors in configuration

- Utilizing a custom policy package for DUNE

  - Implementation functionally bound to current operational status of MetaCat

🟂 Fermilab DUNE

Rucio RSEs & Replicas

[rucio] Replicas per RSE

RSE	Replicas	Total bytes
FNAL_DCACHE	6,726,418	18.84PB
CERN_PDUNE_CASTOR	2,515,923	11.05PB
DUNE_US_FNAL_DISK_STAGE	2,384,065	1.53PB
DUNE_CERN_EOS	1,423,822	3.28PB
RAL_ECHO	1,099,744	909.77TB
RAL-PP	987,253	559.67TB
SURFSARA	979,045	614.43TB
PRAGUE	974,697	812.03TB
DUNE_FR_CCIN2P3_DISK	907,688	1.06PB
DUNE_ES_PIC	831,922	1.41PB
DUNE_US_BNL_SDCC	746,401	743.93TB
NIKHEF	737,781	1.1PB
QMUL	607,917	295.42TB
DUNE_UK_MANCHESTER_CEPH	360,466	476.62TB
DUNE_UK_LANCASTER_CEPH	304,090	461.59TB

[rucio] Replicas pie

- FNAL_DCACHE
- CERN_PDUNE_CASTOR
- DUNE_US_FNAL_DISK_S...
- DUNE_CERN_EOS
- RAL_ECHO
- RAL-PP
- SURFSARA
- PRAGUE
- DUNE_FR_CCIN2P3_DISK

DUNE has 24 RSEs in eight countries

Plan for 2024-2025

- Adapting to CILogon-issued Token is DUNE's top priority

  - Fermilab will keep x509 cert for one more year.

  - CILogon token for all DUNE RSEs.

  - CILogon token in Rucio

  - CILogon token in FTS3

- Complete migration SAM to Rucio

- Consistency checking between Metacat and Rucio

- Consistency checking between Rucio and storages.

Requests from DUNE

- CILogon-issued tokens working with Rucio

- Enhanced QoS support

  - Support Rucio rule for tape backed disk w/o data transfer

- Completion of VO-specific Policy Package testing

- Rucio Globus integration for HPC

Acknowledgements

I would like to extend my heartfelt thanks to my colleagues Steve Timm, Douglas Benjamin, Brandon White, Dennis Lee, Wenlong Yuan and James Perry for their invaluable contributions to the work presented in this talk, as well as for providing essential information and slides.

Thanks to Rucio team for their quick response on several issues during data challenges.

Fermilab   DUNE

Thank You!

Questions?