

Rucio/SENSE Road Towards Production

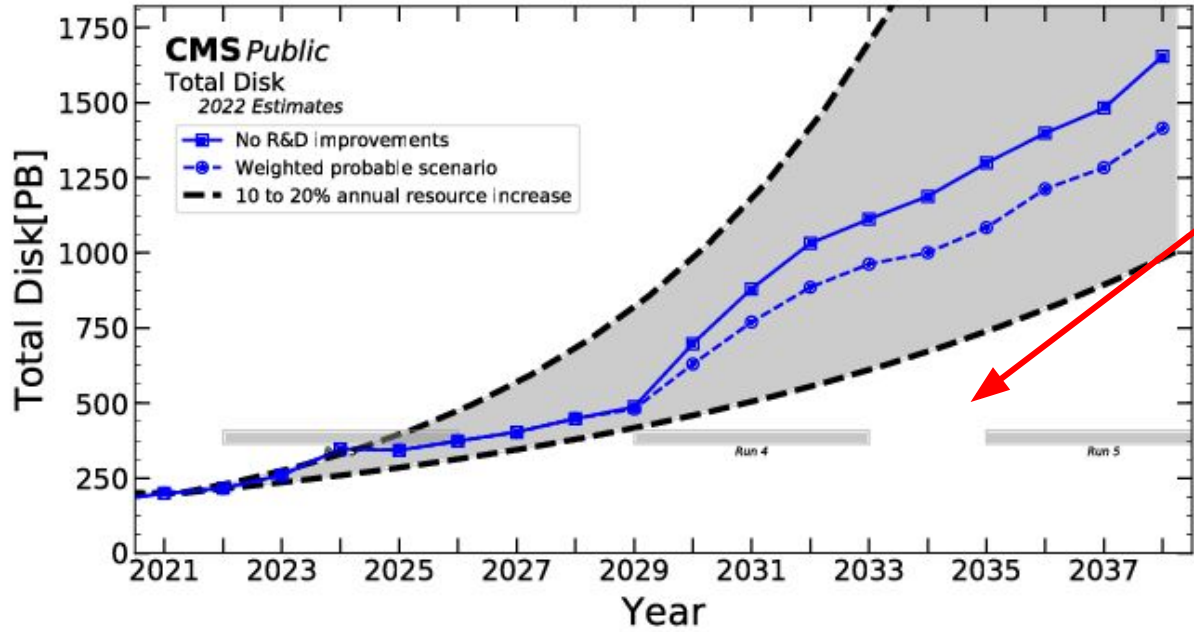
Frank Würthwein, Aashay Arora, **Diego Davila**, John Graham, Dima Mishin, Thomas Hutton, Harvey Newman, Preeti Bhat, Andres Moya, Tom Lehman, Justas Balcas, Xi Yang, Chin Guok, Oliver Gutsche, Phil Demar, Asif Shah, Dmitry Litvintsev, Garhan Attebury, Rafael Coelho, Eduardo Bach, Jessa Westclark, Robert Gardner, Fengping Hu, Lincoln Bryant, Andrew Hanushevsky



8th Rucio Workshop
Nov 5th, 2025



Motivation



2025 WLCG predictions point more into the lower band and below...

Hence: Looks like we will be tight on disk by the time of HL

Motivation (cont'd)

When you are short in disk space you have to move data around

Tight storage situation -> increased transfers



- Recently there was a big 'rebalancing' campaign for the MINIAOD
 - During a 90 day period ~15% of FTS transfer bandwidth was this activity
- "We'll pull that off tape"
- As with our usage of AAA, the more we struggle with storage capacity, the more we need to use network

Katy Ellis - WLCG/DOMA General Meeting, Oct 29, 2025:

https://indico.cern.ch/event/1550918/contributions/6753191/attachments/3159952/5615038/DC27-Coming%20to%20Network%20Rates-C+Oweek_251023.pdf

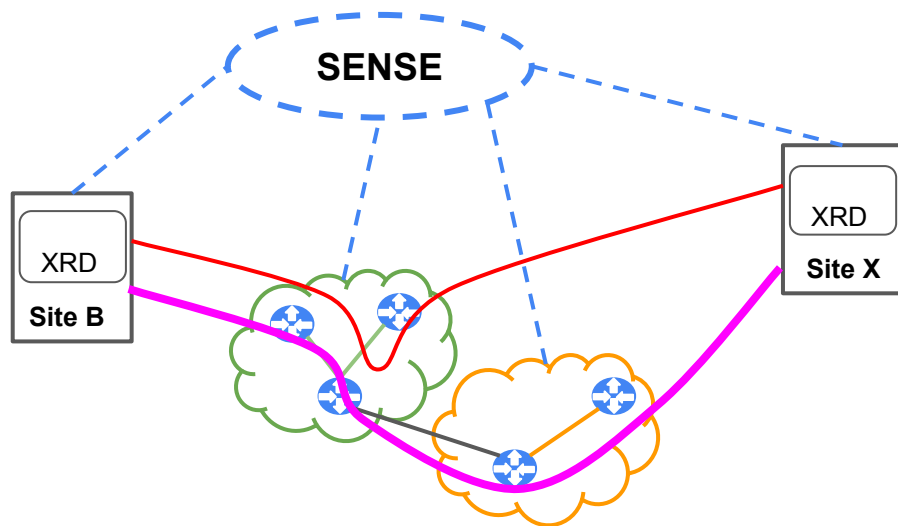
SENSE

An ESnet Network Orchestrator System that can be used to:

1. Find alternative network paths
2. Create bandwidth-guaranteed paths

In the context of CMS and ATLAS we want to use SENSE to **give preferential treatment to certain Rucio/FTS initiated data flows**

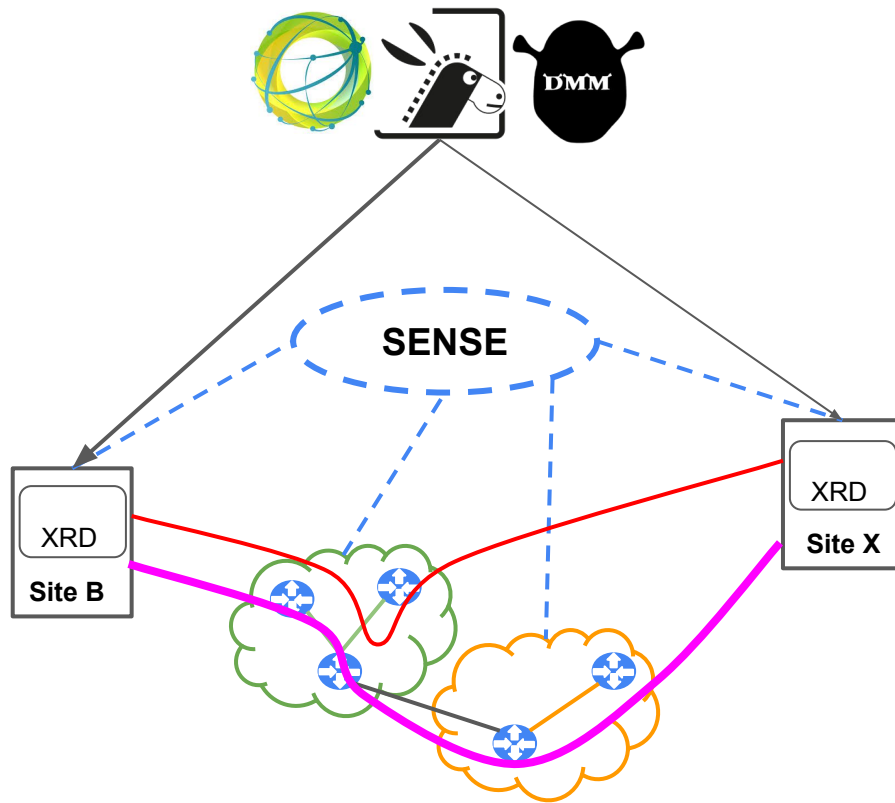
(*) <https://sense.es.net/>



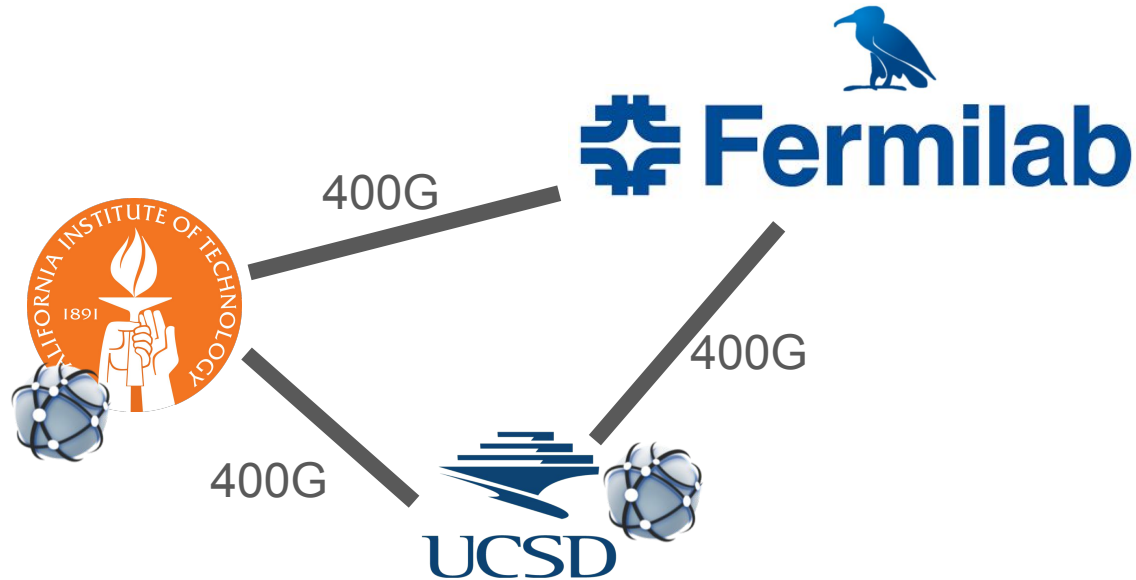
Objective of Rucio/SENSE

Enable Rucio to use SENSE to create **bandwidth-guaranteed** paths for its more important and large data flows

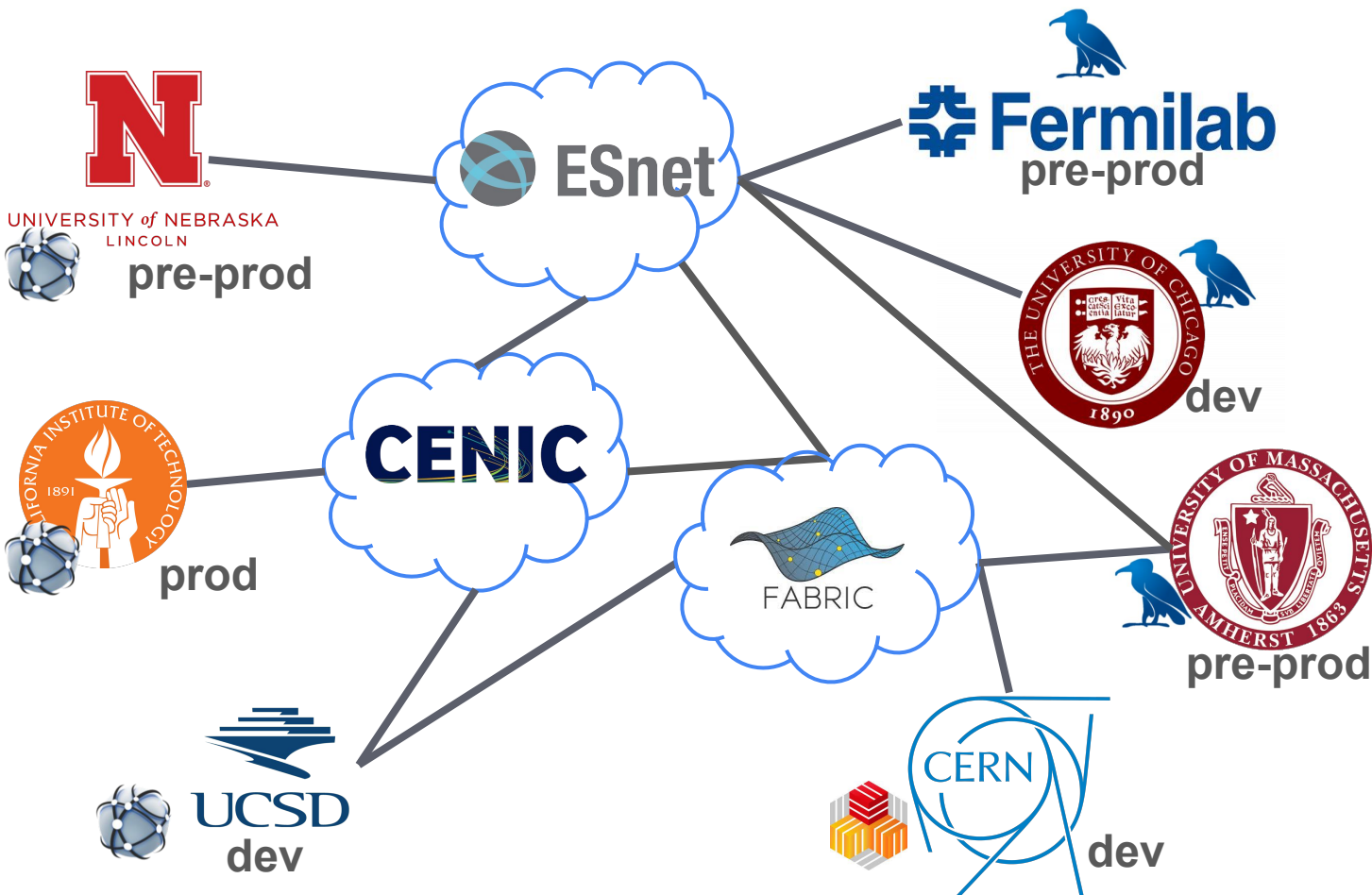
Bandwidth Guarantee =
Predictability &
Resilience to Traffic variability



Current Status: Testbed Last Year



Current Status: Testbed Today



Our own
patched
Rucio



Fts(dev)
@FNAL

Prod: both prod
infrastructure and data

Pre-prod: just
infrastructure

Dev: neither

Latest Tests - FNAL(pre-prod) to UCSD (dev)



Source: FNAL

- Load Balanced dCache proxies
- Backed by Prod-Storage
- Fake data

Destination: UCSD

- XRootD cluster
- NVMe RAID

Tool: Rucio/FTS

<https://autogole-grafana.nrp-nautilus.io/d/000000048/throughput-by-interface?orgId=1&from=1757539495586&to=1757540834942>

Latest Tests - MGHPCC (prod) to UCSD (dev)



Source:

UMass/MGHPCC

- Prod-Storage
- dCache redirect mode

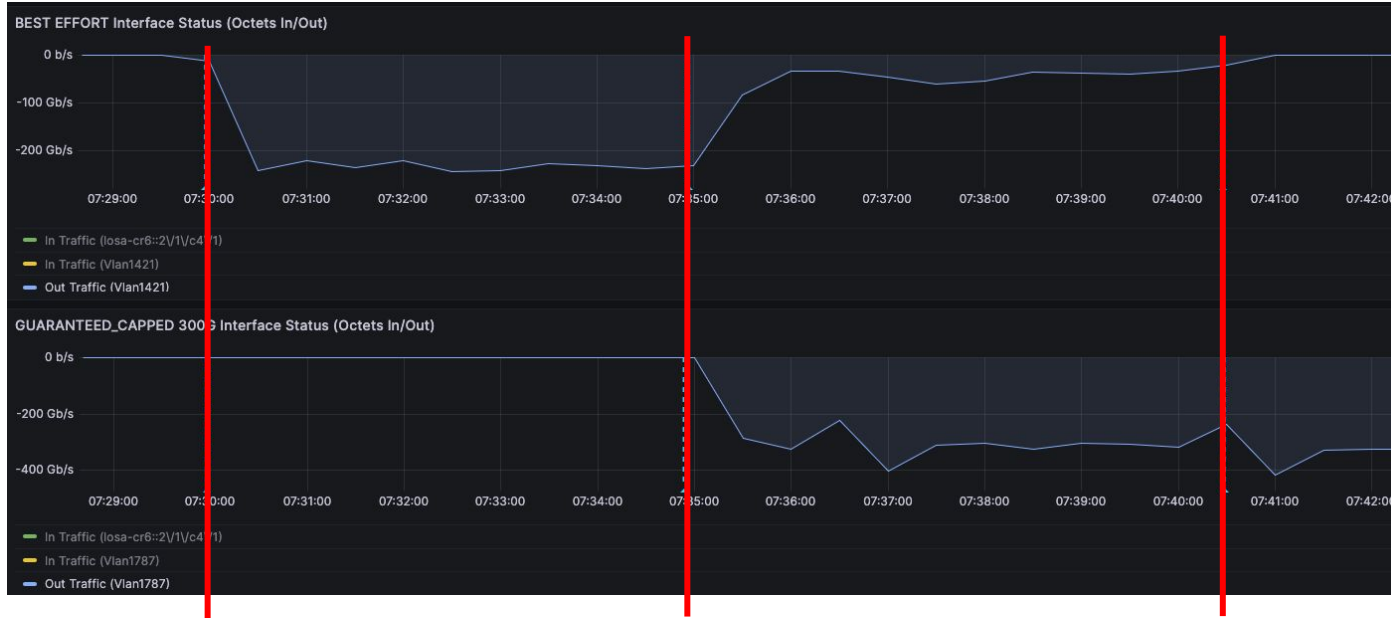
Destination: UCSD

- 2 x Gen5 nodes

Tool: FDT

<https://autogole-grafana.nrp-nautilus.io/d/000000048/throughput-by-interface?orgId=1&from=1761858185036&to=1761861178119>

Best Effort vs BW-Guaranteed at high throughput



<https://autogole-grafana.nrp-nautilus.io/d/bc1f1c44-b3f2-4007-8e63-6ec1d02f0659/oscars-qos-testing?orgId=1&from=1761748120847&to=1761749226767>

1. A Best Effort (BE) flow is injected and takes over the available bandwidth

2. High-Prio (HP) flow is injected. BE flow is capped

3. Once BE finishes HP takes over the full bandwidth

What's new from Last Year?

- Caltech's full support in production
- FNAL and Nebraska: pre-production
- 2 new ATLAS Sites in the testbed: NET2 (UMass) and MWT2 (UChicago)
- XRootD improved support for SENSE:
https://indico.cern.ch/event/1542060/contributions/6710219/attachments/3155479/5604636/xrootd_improved_support_for_sense.pdf
- 2 deployments using the Software-Router approach: CERN and MWT2:
https://docs.google.com/presentation/d/1O73_G-lthl6kLxbT5lCT7s4J1vZCGzD3yUJFhWl5hXs/edit?usp=sharing

Roadmap to production (CMS)

Last year **during the 7th Rucio Workshop** we agreed upon this plan with Eric V.:

1. Get 2 sites to support SENSE in production
2. Transfer knowledge into CMS Ops to deploy and operate (*)DMM
3. Patch the Rucio-Integration instance to talk to DMM
 - a. Similar infrastructure that prod but slightly different namespace
 - b. Used to run periodic tests
4. Perform SENSE-supported data transfers via the Integration testbed
5. If all the above goes well consider patching Rucio-prod

(*) DMM - Data Movement Manager. Interface between Rucio and SENSE

https://www.epj-conferences.org/articles/epjconf/pdf/2025/22/epjconf_chep2025_01208.pdf

Roadmap to production (CMS) cont'd

Unfortunately we've been stuck in step 1:

1. Get 2 sites to support SENSE in production

Currently we have 4 good contenders:



We have access to a subset of their prod-FS via a balanced set of dCache proxies



UNIVERSITY of NEBRASKA
LINCOLN

Dealing with unexpected behavior of multi-subnet setup



READY!



Stuck in a network upgrade (99% completed now) for what it feels forever

Roadmap to production (ATLAS)

ATLAS joined the Rucio/SENSE effort significantly later than CMS, but thanks to the efforts made by NET2(UMass) and MWT2(Uchicago) we currently have deployments for these 2 in the SENSE testbed.

At NET2 we even got access to their prod-storage (infrastructure, not data) via dCache proxy and the plan is to go further and expose their DTNs directly to SENSE

In principle, by having a second ATLAS site supporting SENSE in prod, we could follow the same plan as with CMS but that conversation has just started.

Next Steps

- SC25 Demo (current focus)
 - Start a “background traffic” flow of 800 Gbps from NET2 (UMass)
 - Add a 200 Gbps SENSE managed data flow on top of the above
- Get a second CMS site to fully support SENSE in prod
 - Patch CMS Rucio-Integration instance
- Add a Rucio/DMM ATLAS instance in our testbed
- Agree on a “Roadmap to production” with ATLAS
- Engage with CERN/EOS
- Continue engaging with more Sites (both CMS and ATLAS)

Question to the Audience:
How would you use this technology?



Acknowledgements

This ongoing work is partially supported by the US National Science Foundation (NSF) Grants OAC-1836650, PHY-2323298, PHY-2121686 and OAC-2112167.

Finally, this work would not be possible without the significant contributions of collaborators at ESNet, Caltech, FNAL and SDSC.

Thanks!

Questions?

Answers (prev question)?

*If you are shy, send me
an email :)*

—————→ didavila@ucsd.edu

