Rucio/SENSE Integration update on the path to production

Diego Davila, on behalf of the Rucio/SENSE team

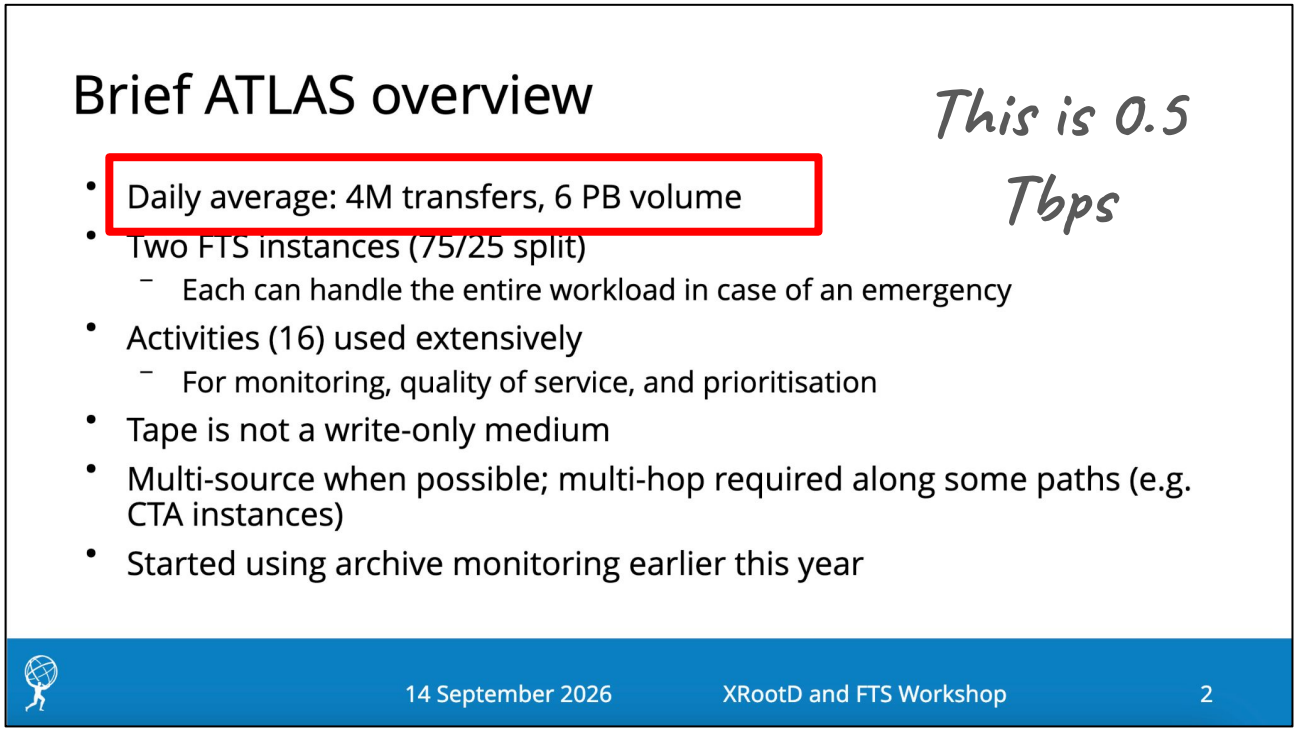
FTS/XRootD Workshop

Sep, 2026

Motivation

- SENSE is to the Network what HTCondor is to CPU *(kind of...)*
- Imagine a world without job schedulers where we shove processes into a server until it becomes slow and then we ask for a bigger server
  - Not looking at priorities nor the characteristics of the jobs (cores, mem, disk)
- That's how we use the network nowadays :(
  - We are bad bad users of the Network
  - Because we have always been spoiled
    - We've always got more bandwidth
    - HEP is the biggest player around, no need to worry about competing with other users
- It is time to become better users of the Network, starting by communicating our priorities and intentions.

More Motivation: From Dimitrios (yesterday)



https://indico.cern.ch/event/1639200/contributions/7225211/attachments/3342795/5991491/2026.09.15%20ATLAS%20Community%20Talk.pdf

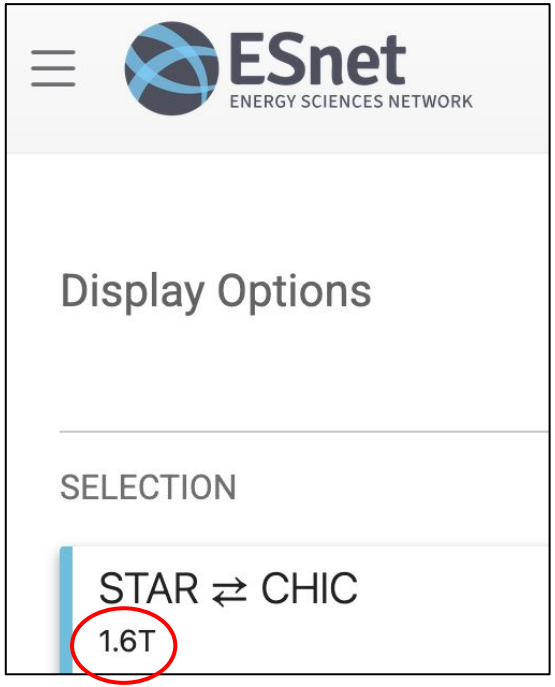
Even More Motivation

SC26 Scitags R&D Experiments

High-Fidelity Packet Marking: Conduct collaborative, real-time testing and performance impact measurements on the showfloor.

Focused on packet marking at high throughput.
We plan to have 1.6Tbps btw. two HP Gen6 servers each with 2x800Gbps (Univ of Victoria and Starlight), with CERN at 1.2 Tbps (DTN Gen5 3x400Gbps).

https://indico.cern.ch/event/1722889/



You don't need a Grid to distress the Network... just 1 big server.

... maybe it's time to worry about other users

SENSE

A Network orchestrator that can, among other things, help us scheduling the Network.

This means: making bandwidth reservations between 2 sites

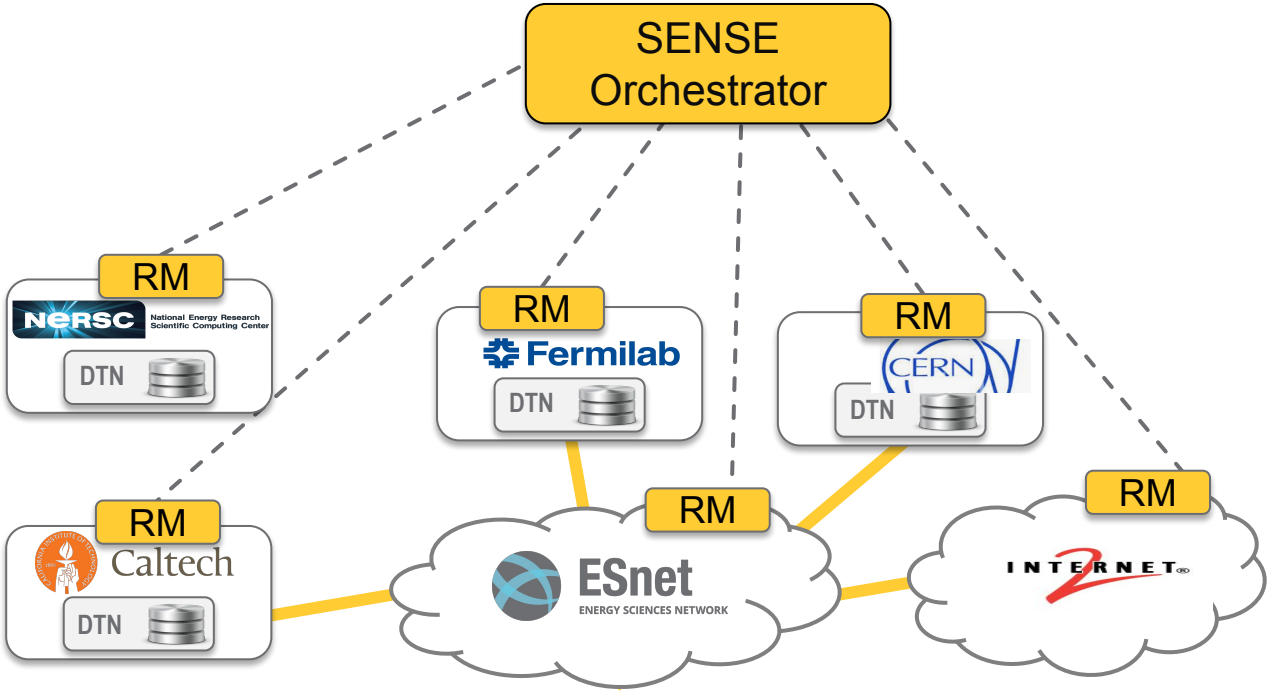
If our data management systems were able to communicate with the Network and tell:

“this transfer is more important than these others”

That would be already a great improvement

How does SENSE work?

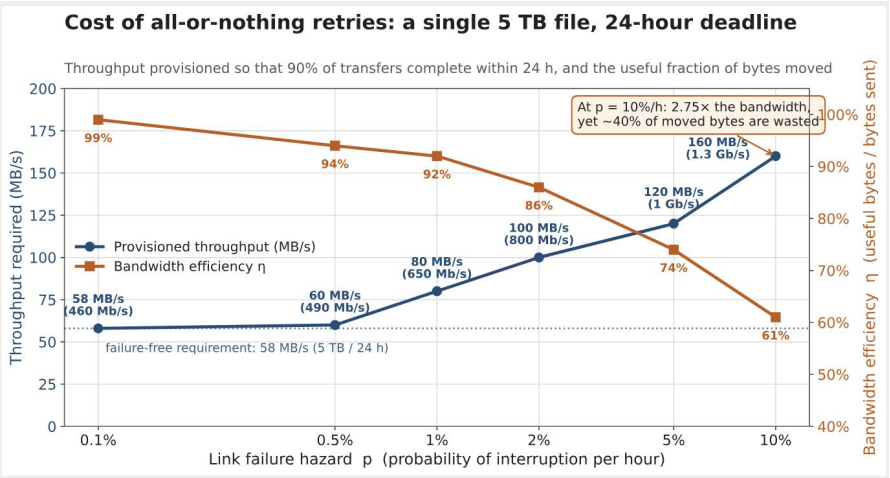
By deploying agents RM both at the Sites and in the R&D Networks (e.g. Internet2, GEANT) it can control **Routing** and **Bandwidth Allocation** between Sites (DTN to DTN)



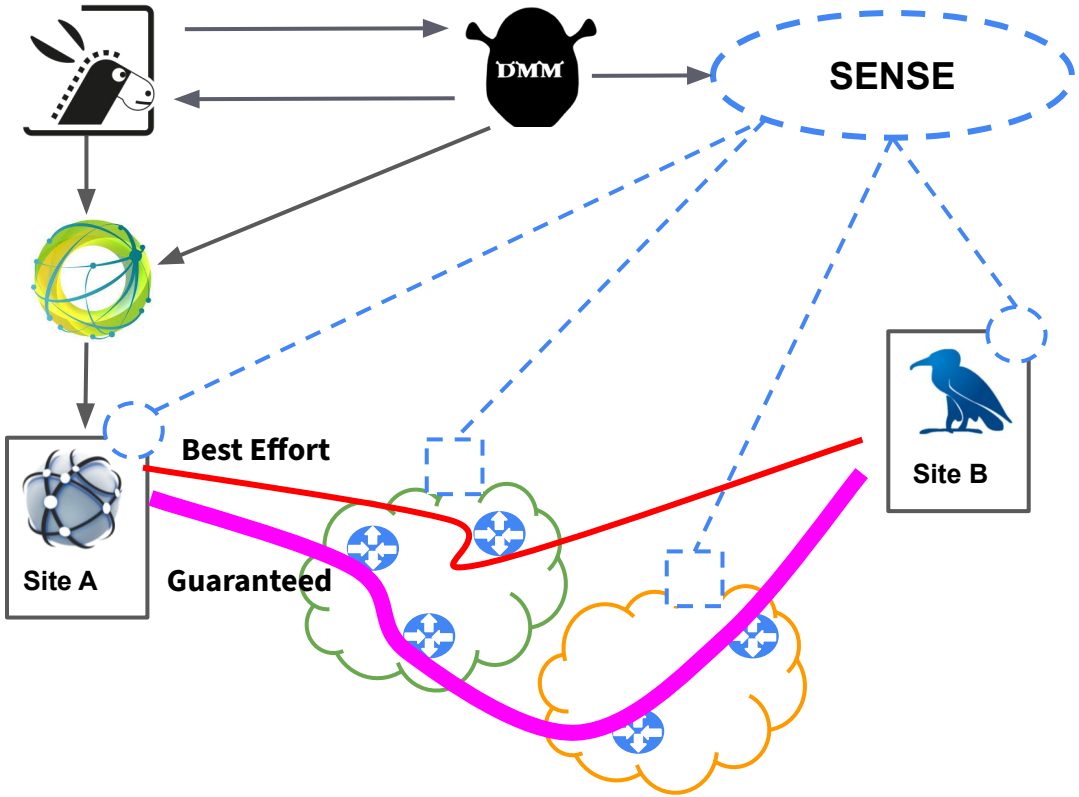
Rucio: The first SENSE use case

- SENSE can be used independently from Rucio
- We picked Rucio as a first use case because scheduling bandwidth for the largest data flows makes SENSE (pun totally intended)
- **Who doesn't want precise ETAs when it comes to moving data?**

From James Walder talk (Monday)
https://indico.cern.ch/event/1639200/contributions/7231113/attachments/3342186/5990239/SRCNET_FTSXROOTD_LyonSept2026.pdf



SENSE and Rucio in action



Bandwidth Guarantee != Throughput Guarantee

This is where FTS comes into play

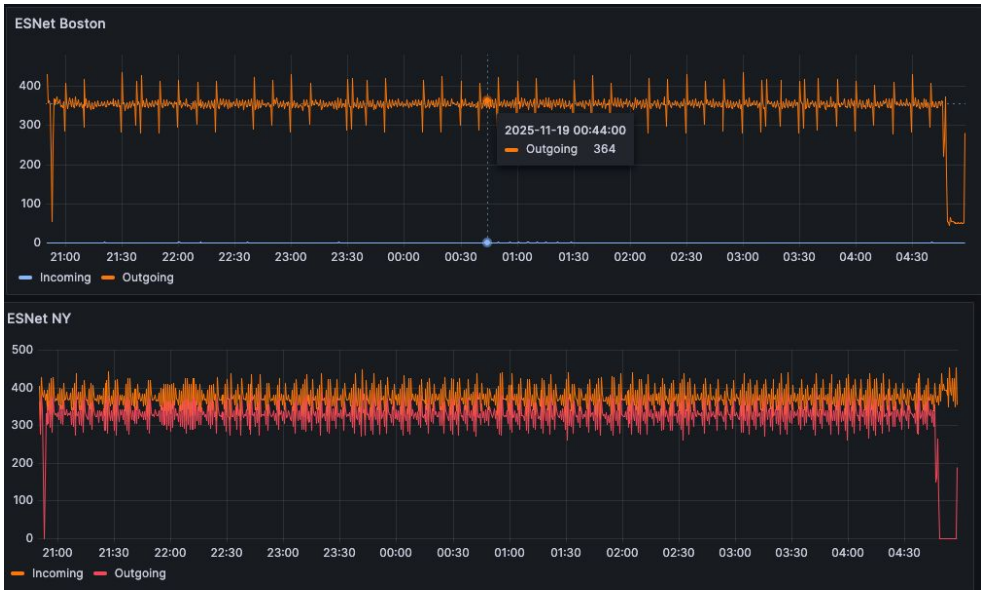
Stolen from Nicola Pace talk (Tuesday):

We know what the optimal should be

We might be able to give you a hint on what the number of active transfers should be

We need to be able to use all the allocated bandwidth ASAP

A different use case:UMass (ATLAS T2) SC25 DEMO



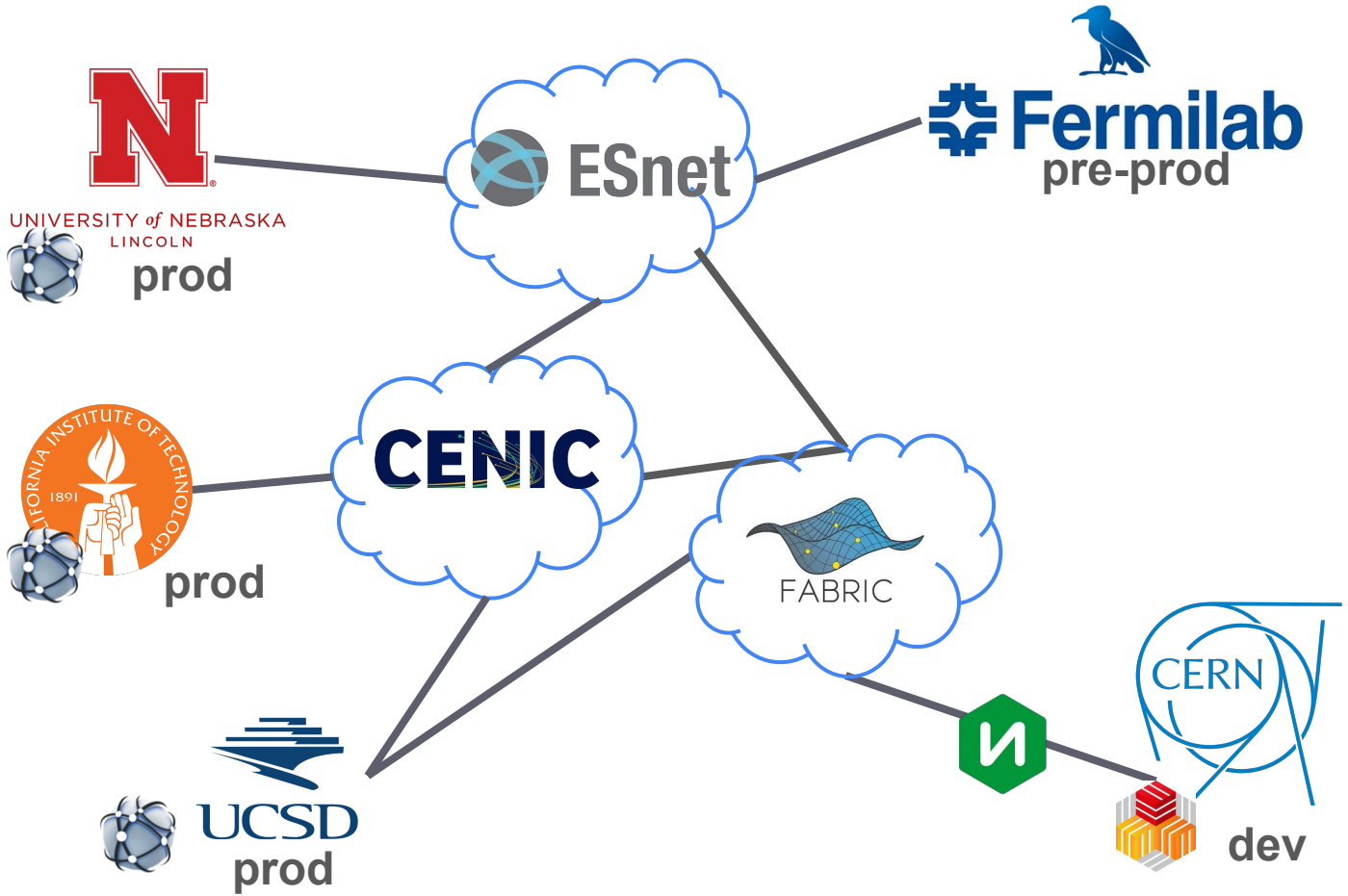
Using SENSE we were able to configure UMass to temporarily make use of their backup network path to increase their network capacity from 400 to 800 Gbps (symmetric) for the duration of the demonstration.

UMass performing 1.1Tbps at SC25.
Top: UMass/Boston => SC25 400 Gbps
Bottom: UMass/NY <=> SC25 400 + 300 Gbps

Status of the project

- Split our testbed: now ATLAS has its own Rucio/DMM instances
- Set up Automated tests
- Worked on the CERN deployment: NGINX proxy in front of EOS routed by a software router running in FABRIC
- Added support for SENSE in production at 2 new T2s
- Working with another 2 CMS T2s
- Looking for a new ATLAS T2 to join (maybe SWT2)

CMS Testbed



Our own patched Rucio

CMS DMM

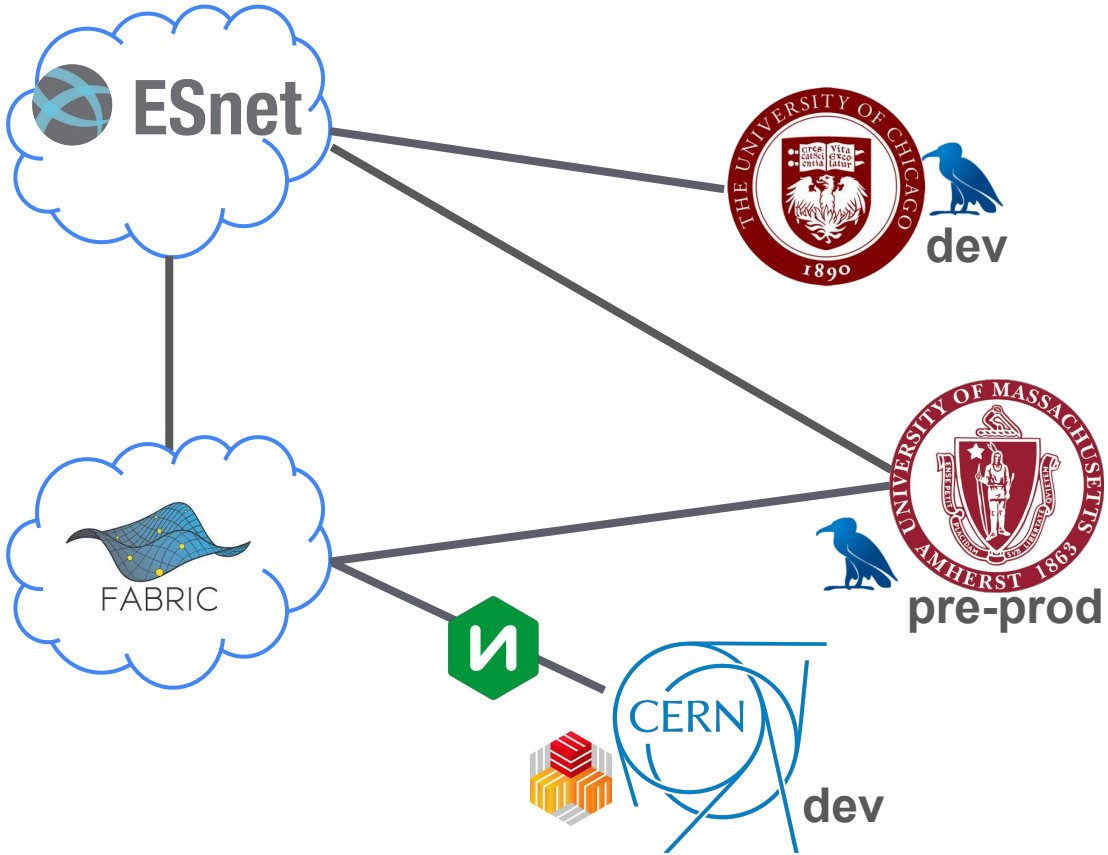
Fts(dev) @FNAL

Prod: both prod infrastructure and data

Pre-prod: just infrastructure

Dev: neither

ATLAS Testbed



Our own patched Rucio

ATLAS DMM

Fts(dev) @BNL

Prod: both prod infrastructure and data

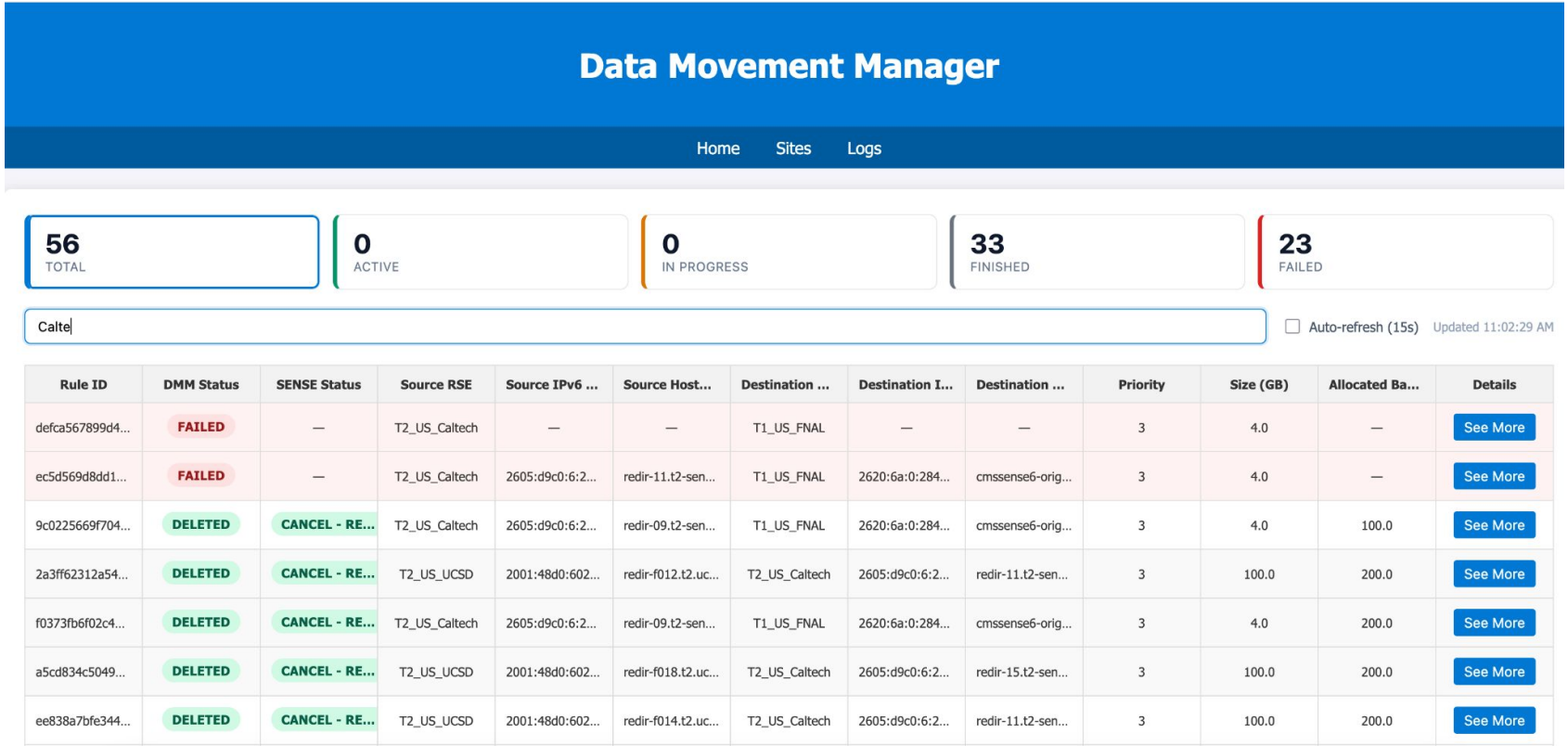
Pre-prod: just infrastructure

Dev: neither

Automated Tests

- In order to increase the stability of the overall system we started running periodic automated tests across the site in the CMS testbed
- Every 8 hrs a set of short-lived Rucio rules are injected requesting “SENSE” treatment
- These tests have help us find bugs across the whole infrastructure
- The same should be implemented in the ATLAS testbed in the future

DMM Dashboard



Rule ID	DMM Status	SENSE Status	Source RSE	Source IPv6 ...	Source Host...	Destination ...	Destination I...	Destination ...	Priority	Size (GB)	Allocated Ba...	Details
defca567899d4...	FAILED	—	T2_US_Caltech	—	—	T1_US_FNAL	—	—	3	4.0	—	See More
ec5d569d8dd1...	FAILED	—	T2_US_Caltech	2605:d9c0:6:2...	redir-11.t2-sen...	T1_US_FNAL	2620:6a:0:284...	cmssense6-orig...	3	4.0	—	See More
9c0225669f704...	DELETED	CANCEL - RE...	T2_US_Caltech	2605:d9c0:6:2...	redir-09.t2-sen...	T1_US_FNAL	2620:6a:0:284...	cmssense6-orig...	3	4.0	100.0	See More
2a3ff62312a54...	DELETED	CANCEL - RE...	T2_US_UCSD	2001:48d0:602...	redir-f012.t2.uc...	T2_US_Caltech	2605:d9c0:6:2...	redir-11.t2-sen...	3	100.0	200.0	See More
f0373fb6f02c4...	DELETED	CANCEL - RE...	T2_US_Caltech	2605:d9c0:6:2...	redir-09.t2-sen...	T1_US_FNAL	2620:6a:0:284...	cmssense6-orig...	3	4.0	200.0	See More
a5cd834c5049...	DELETED	CANCEL - RE...	T2_US_UCSD	2001:48d0:602...	redir-f018.t2.uc...	T2_US_Caltech	2605:d9c0:6:2...	redir-15.t2-sen...	3	100.0	200.0	See More
ee838a7bfe344...	DELETED	CANCEL - RE...	T2_US_UCSD	2001:48d0:602...	redir-f014.t2.uc...	T2_US_Caltech	2605:d9c0:6:2...	redir-11.t2-sen...	3	100.0	200.0	See More

Goal: SENSE fully in production (CMS) during DC27

With the newly increased stability on the system and 3 sites supporting SENSE in production we believe we are ready to try SENSE in the real CMS Rucio

Next steps:

1. Deploy a DMM instance controlled by the CMS-Rucio team (anytime)
2. Patch the CMS Rucio-Integration instance (by the end of the year)
3. Learn from this pre-prod environment
4. Patch the CMS Rucio-Production instance (late January 2027)

Note: The above steps are suggested by the Rucio/SENSE team. The CMS-Rucio team has not agreed on them… yet :)

Acknowledgements

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

This work would not be possible without the significant contributions of collaborators at ESNet, Caltech, FNAL, SDSC, UNL, NET2 and MWT2.

Questions?

UNIVERSITY OF MASSACHUSETTS
NET2

NRP
NATIONAL RESEARCH
PLATFORM

UNIVERSITY of NEBRASKA
LINCOLN

XRootD

Cres. Cat Sci entia
Vita Exco latur

ESnet

iris
hep

Fermilab

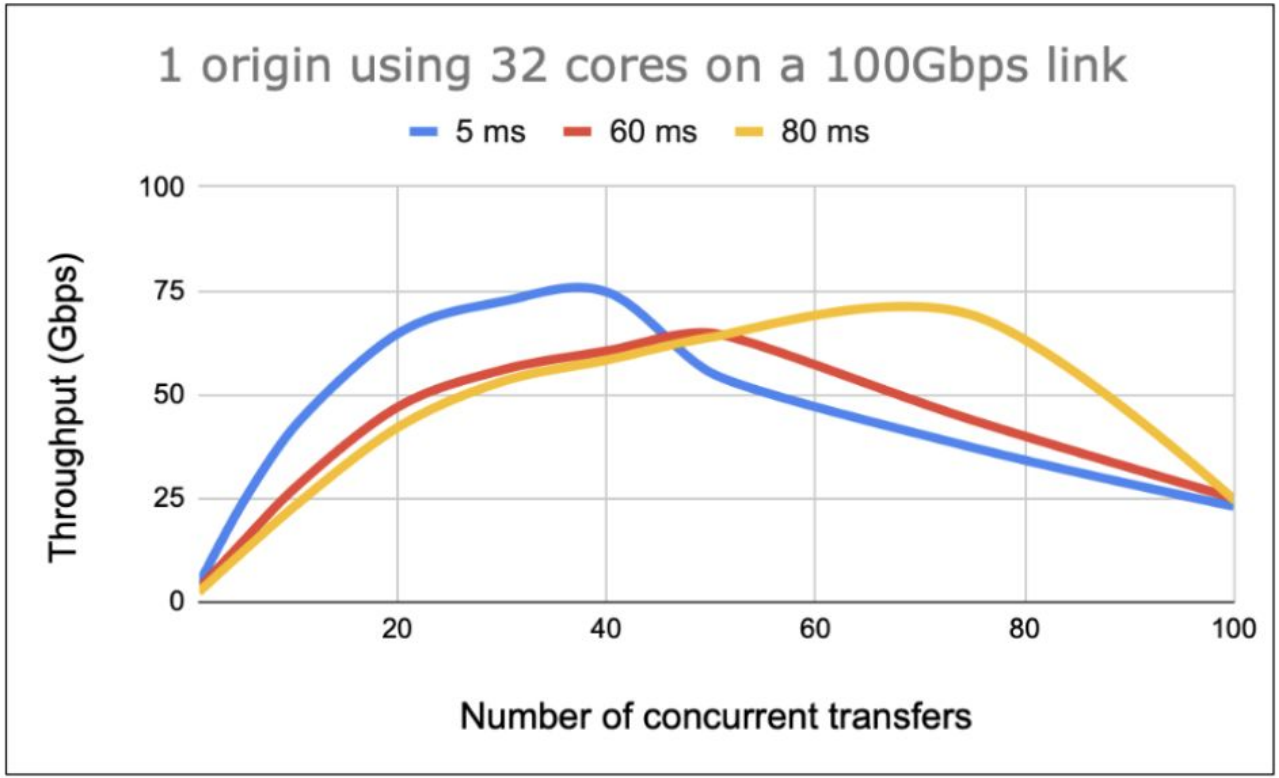
CALIFORNIA INSTITUTE OF TECHNOLOGY
1891

UCSD

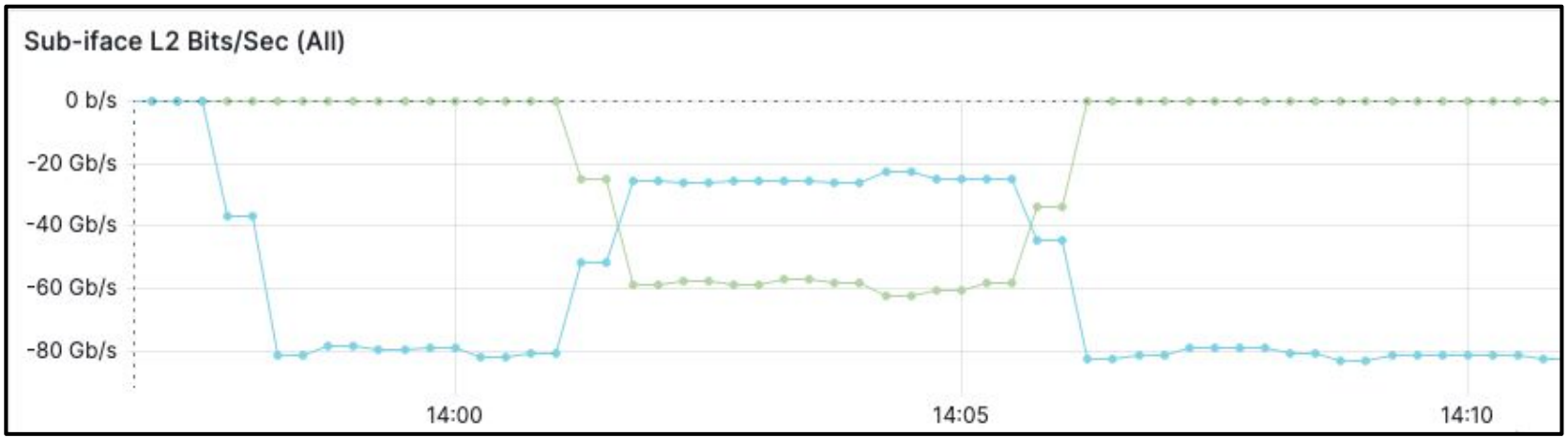
CMS

Backup Slides

Throughput, Active Transfers and RTT

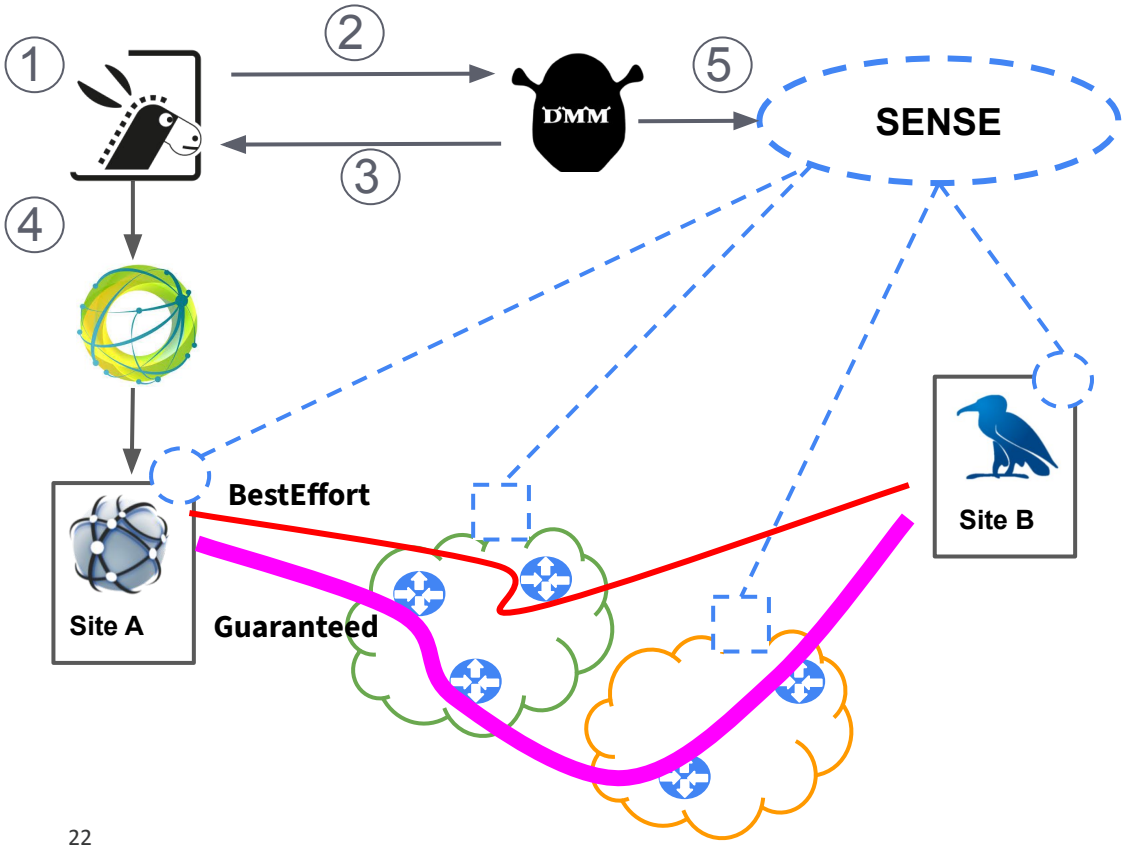


Story of Success: UChicago



During the WLGC 2026 Capability Mini-challenge we were able to demonstrate a 100Gbps SENSE managed data flow between UChicago and UMass

SENSE and Rucio in action



1. A user creates a Rucio rule with the −sense flag.

2. Rucio contacts DMM to get a new pair of endpoints

3. DMM pick a free endpoint on each site and gives them back to Rucio. These endpoints point to the **same sites** but on **different subnets**

4. Rucio replaces the old endpoints with the new ones and continues as usual i.e. sends the list of files to transfer to FTS

5. Behind the curtain DMM requests the creation of the special path

The patch: `/lib/rucio/transfertool/fts3.py`

```
+    # SENSE modifications
+    use_sense = config_get_bool('dmm', 'use_sense', False, None)
+    dmm_url = config_get('dmm', 'url', False, None)
+
+    dmm_response = {}
+
+    if use_sense and dmm_url:
+        for file in files:
+            # Get rule ID
+            try:
+                rule_id = file['metadata']['rule_id']
+                logging.debug(f"Trying to change job endpoints for {rule_id}")
+                if rule_id not in dmm_response.keys():
+                    logging.debug("Rule ID not in cache, getting from DMM")
+                    response = requests.get(dmm_url + '/query/' + rule_id)
+                    if response.status_code == 200:
+                        logging.debug(f"Got response 200 from DMM: {response.json()}")
+                        dmm_response[rule_id] = response.json()
+                    else:
+                        raise Exception(f"Could not get SENSE addresses for {rule_id}")
+
+                logging.info(f"job endpoints changed for {rule_id} with sense hosts")
+
+                if dmm_response[rule_id]:
+                    logging.debug("Rule ID in cache, changing job endpoints")
+                    # replacement
+                    src_url = file['sources'][0]
+                    src_hostname = src_url.split("/")[2]
+                    src_sense_url = src_url.replace(src_hostname, dmm_response[rule_id]['source'], 1)
+                    file['sources'][0] = src_sense_url
+
+                    dst_url = file['destinations'][0]
+                    dst_hostname = dst_url.split("/")[2]
+                    dst_sense_url = dst_url.replace(dst_hostname, dmm_response[rule_id]['destination'], 1)
+                    file['destinations'][0] = dst_sense_url
+                else:
+                    raise Exception("Illegal response from DMM")
+
+            except Exception as e:
+                logging.error(f"Error getting SENSE addresses: {e}, continuing as normal")
```

Single HTTP GET request with rule_id as query param

Replace endpoint with SENSE controlled endpoint.

In a nutshell:

```
if not sense:
  continue
else:
  for transfer in transfers_to_be_submitted_to_fts:
    new_endpoints = query_dmm(file.rule_id)
    file.replace(curr_endpoint, new_endpoint)
...
transfer.submit_to_fts()
```

Network Control - Requirements

From the point of view of a site and in order to achieve the above, SENSE requires access to the following network capabilities:

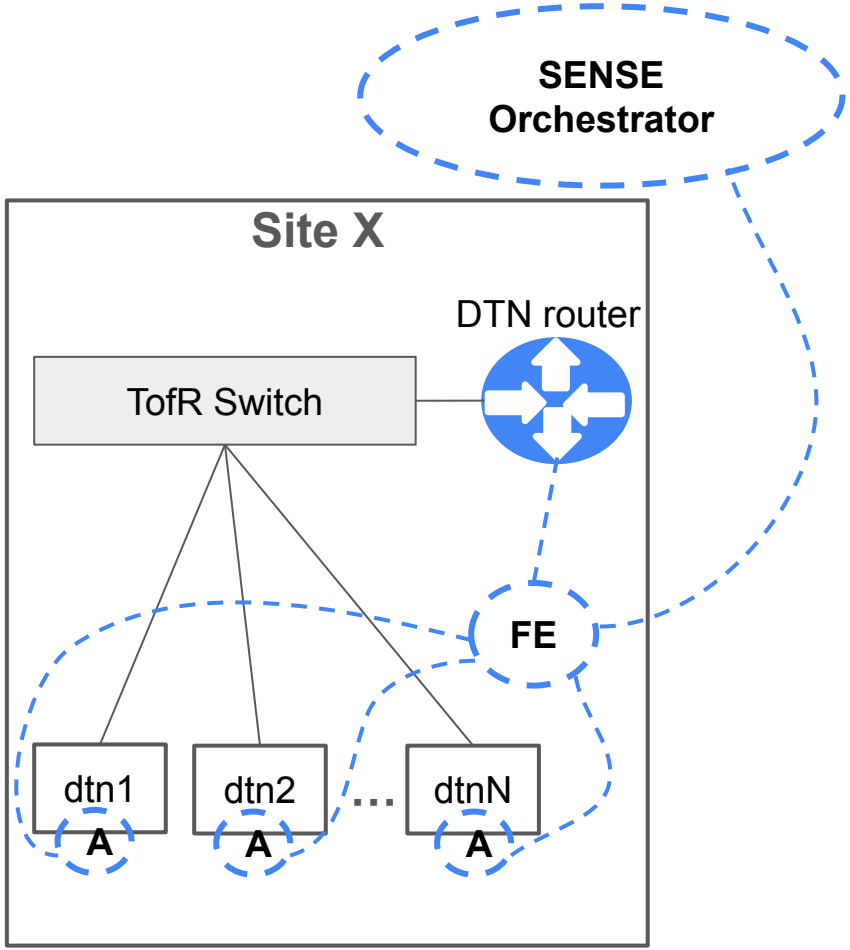
- Modify the routing of the DTNs data flows
  - Manage VLANS to create Point-2-Point L2 network paths
  - Manage BGP announcements to route desired traffic over the above L2
- Configure QoS policies on the DTNs
  - Used to provide bandwidth allocations
  - Done using the Linux Traffic Controller

These requirements are **restricted to a defined set of subnets and VLANS** which are used only for SENSE-managed data flows.

Network Control - SiteRM

At the site level, SENSE requires the deployment of 2 main Site Resource Managers (SiteRM):

1. SiteRM Frontend.
   a. Receives commands from the Orchestrator
   b. Applies changes in the DTN router
   c. Sends commands to the Agents
2. SiteRM Agent.
   a. Runs in the DTNs
   b. Execute commands on the DTN
   c. Enforces QoS policies



More info about SiteRM: https://sdn-sense.github.io/#siterm-resource-manager

Software-Router

The SENSE system is based on traffic engineering IPv6/64 subnets for the purpose of providing end-to-end priority data transfer service. SENSE is typically allowed to apply a **limited set of configurations** on the DTN router to effect these features at the site level.

When the above network access is not possible, SENSE provides an option for a software based traffic engineering approach .This solution is based on the following 3 software suites:

1. Free Range Routing (FRR)
2. Vector Packet Processing (VPP)
3. Data Plane Development Kit (DPDK)

Software-Router

Generally speaking, network routing is composed of 2 elements: control-plane and the data-plane. The former is in charge of the logic and the later of the execution.

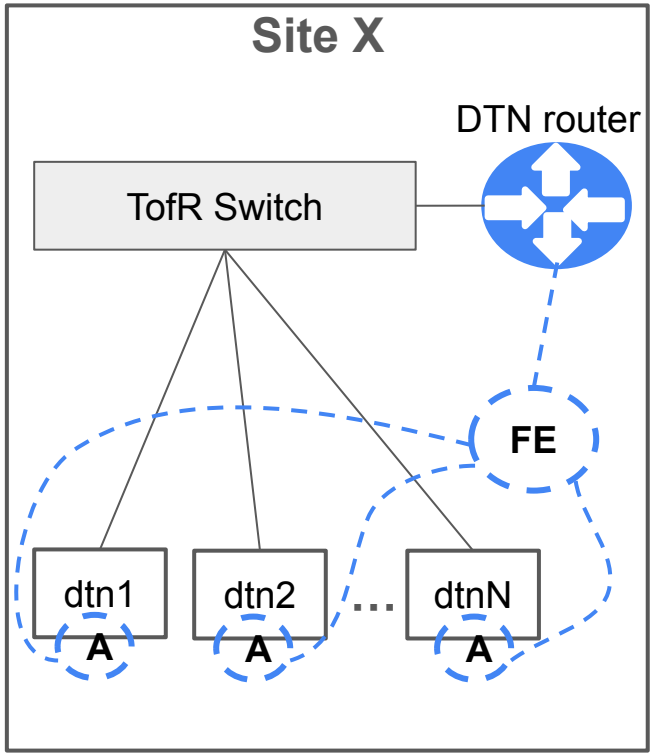
FRR implements the control-plane logic. When used alone, it offloads the data-plane logic to the Linux Kernel which process packets one-by-one.

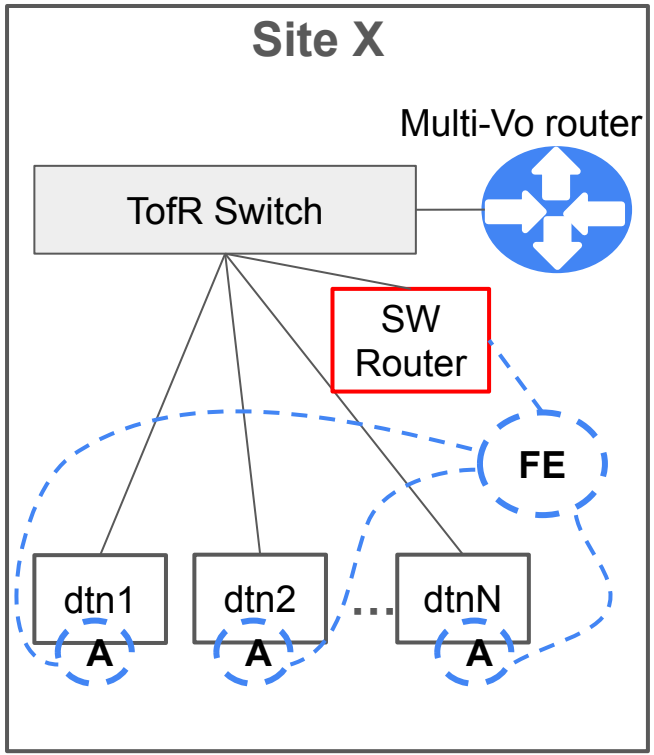
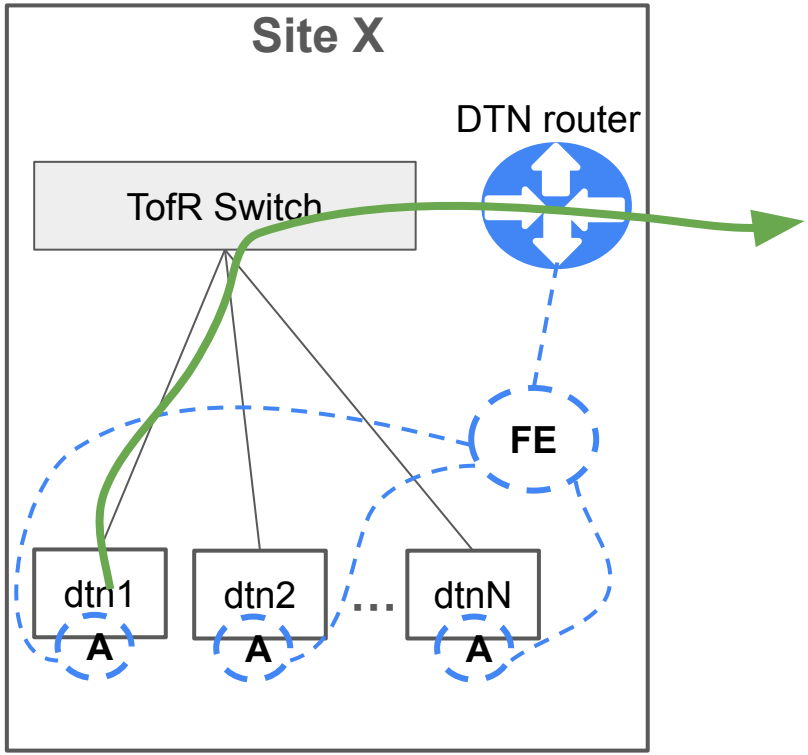
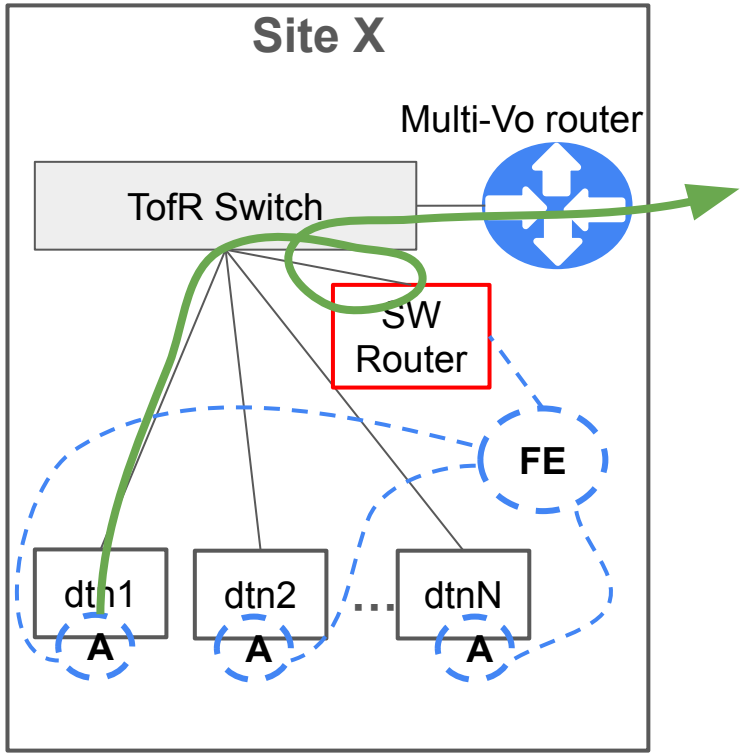
When dealing with high-throughput the above becomes a bottleneck. This is where **VPP/DPDK** comes into play.

VPP process packets in vectors which makes the process faster. Moreover it can be integrated with **DPDK** allowing for directly control of the NIC, bypassing the Linux kernel network stack hence eliminating the interrupt overhead

More information here: https://sdn-sense.github.io/optional-install/software-based-router-dpdk/

Regular Router vs Software-Router





Routing Difference