U.S. DEPARTMENT of ENERGY

BERKELEY LAB

ESnet
ENERGY SCIENCES NETWORK

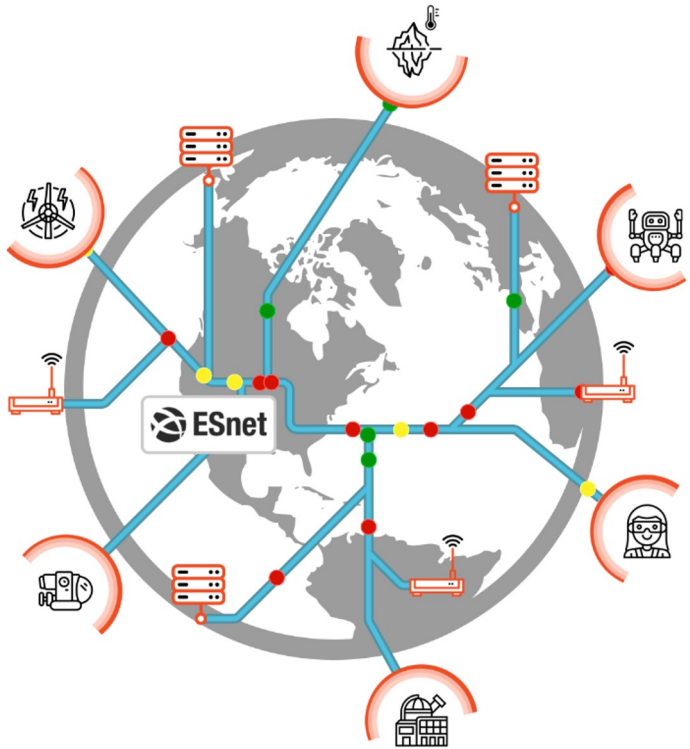
Performant Perimeter Security:
A Foundation for Data Intensive Science

Eli Dart
Network Engineer, Science Engagement
dart@es.net

ISGC 2026
Taipei, Taiwan
19 March 2026

Our World Is Data-Centric

- Many, many scientific discoveries are impossible without data analysis
- Modern science relies on computing, data, storage, all interconnected using fast networks
- ESnet even thinks of itself as a “data circulatory system”



Most People Want Semantic Reasoning

- Think about high level analysis instead of low level data sets
- This is where the value is
- This is where the discoveries are

- However, in order for this to be possible, the low level stuff has to work

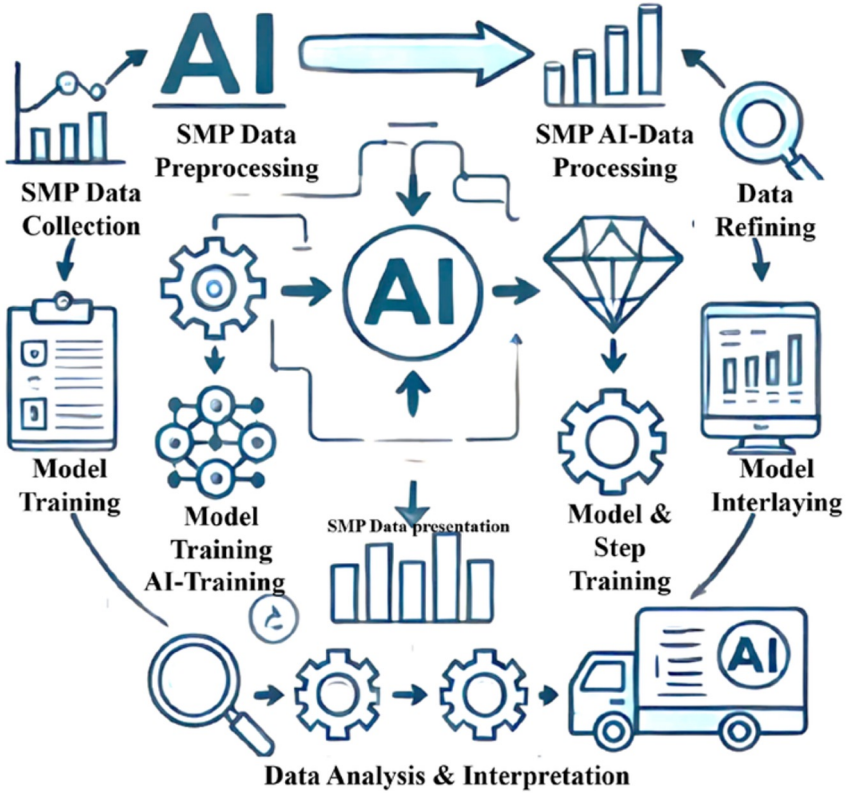
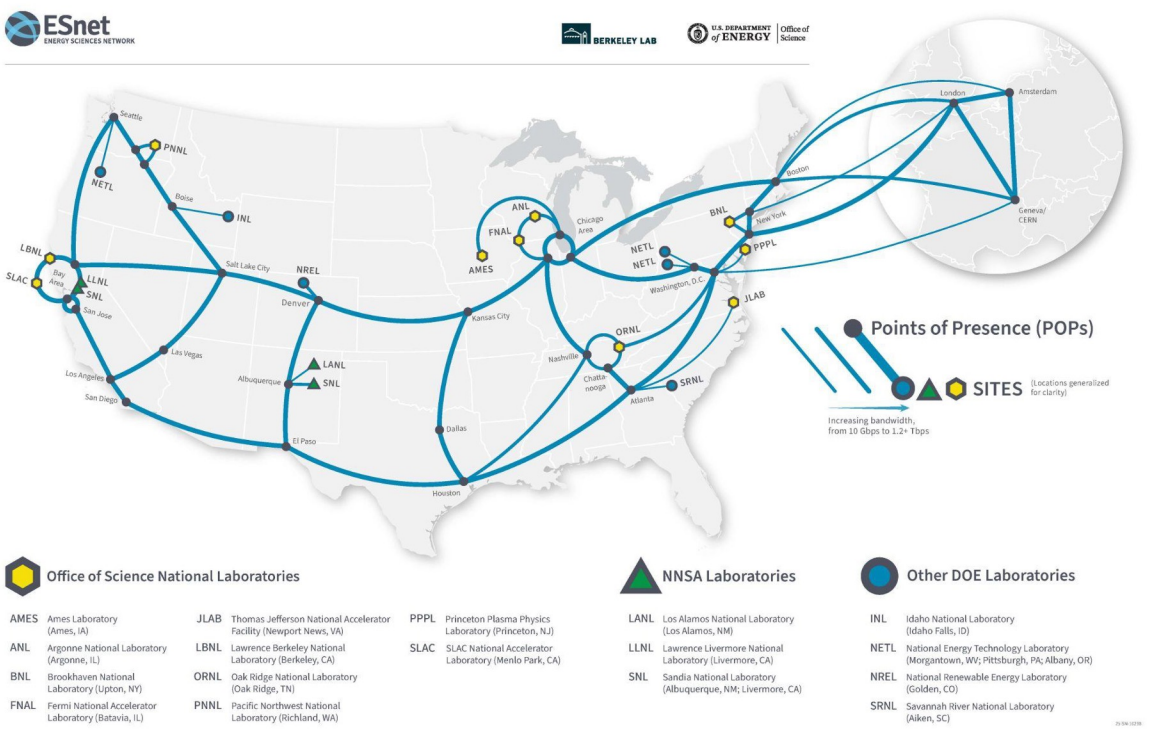


Image: Olawumi, Mattew & Oladapo, Bankole & Olugbade, Temitope & Omigbodun, Francis & Olawade, David. (2024). AI-Driven Data Analysis of Quantifying Environmental Impact and Efficiency of Shape Memory Polymers. Biomimetics. 9. 490. 10.3390/biomimetics9080490.

Data Infrastructure Under The Semantics

- ESnet supports the DOE scientific research ecosystem.
- Interconnects all US national labs and user facilities
- Provides reliable, high-performance connectivity to global research collaborations, the Cloud, and the larger Internet.



ESnet

Science Needs More Than Just The Network

- Computing and storage
  - HPC center, institute
  - cloud, personal device
- Code (workflow, analysis, …)
- **All of these things exist somewhere in physical space**
- **All of these things located in the network topology**
- It's great to be able to reason at a higher level
  - But then all the lower level stuff has to work!
- Science workflows are multi-domain in the general case
- Data and workflows cross network boundaries

What is Perimeter Security?

- Network perimeter: boundary between zones of control
- Each organization has its zone of control
  - Hardware owned by the organization
  - Network topology under the control of the organization
  - Connection of the network to the outside world
- Cloud complicates this, but the network perimeter still exists
  - This is a necessary construct
  - ZeroTrust doesn't change it
  - On-premise assets are "inside" the perimeter, the rest of the Internet is "outside" the perimeter
- We can think of perimeter security as the policies and technologies that defend on-premise assets from network-based attacks

Science Crosses the Perimeter

- Science collaborations are inherently multi-domain
  - Labs, Universities, Cloud, etc.
  - Data must move between sites and resources
- Science mission requires that high performance science traffic moves between and across zones of control
  - Data, computing, storage, instruments, researchers
  - All are at different locations in the network topology
- Key point: **high performance data flows have a concrete location in the network topology**
  - Virtualization, cloud, workflows, etc. all hide physical performance characteristics, but that doesn't change them
  - If we want workflows to perform, we have to have a performant physical path for the data

Security and Performance In Conflict

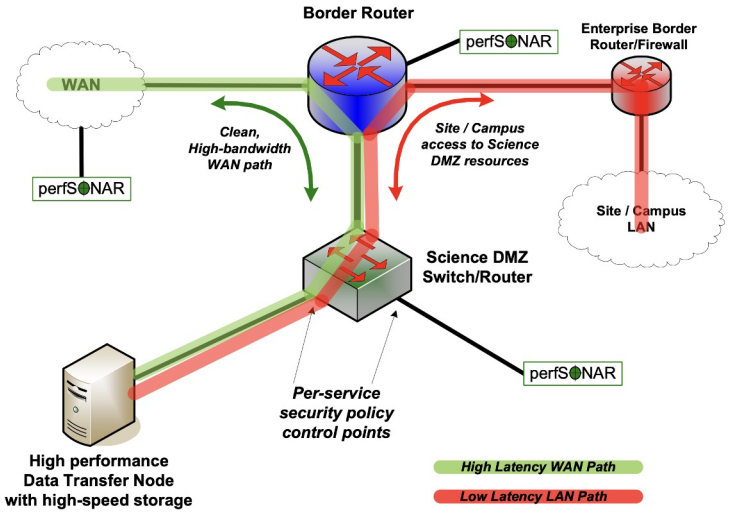
- The organization must defend its perimeter
  - Huge liability risk
  - Operational/availability risk
  - This MUST happen, so it WILL happen!
  - We can choose how it happens, but not whether it happens
- Science requires performance
  - In order to meet the science mission, we must find a way to make data flows performant
  - Because perimeter security must exist, performance must exist in the context of perimeter security
  - Find a way to run high performance workflows across the perimeter

High Performance, Safely

- This is both a technical issue and a sociological issue
- The technical solutions must be technically sound
- However, policy is critical, and there are multiple related but separate aspects to perimeter policy
  - Which hosts/systems are allowed to exchange data?
  - Which applications and protocols are permitted?
  - Which technologies are used to enforce policy?
- The answers to these questions determine whether high performance data applications can use your perimeter
  - And: can your scientists run high performance workflows at your site, using your computing, instruments, and storage?

Best Practice: Science DMZ

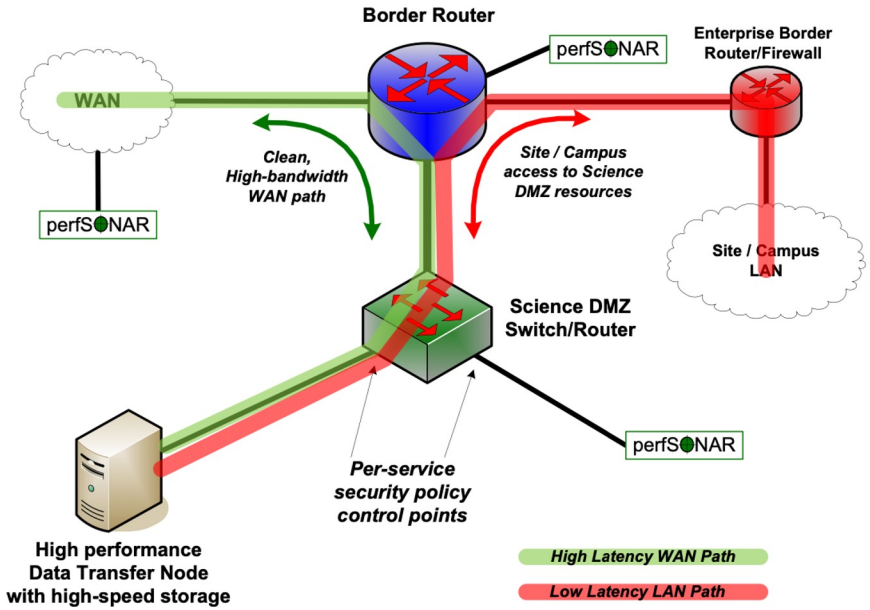
- The answer is the science dmz model.
  - Separate high performance science traffic from enterprise traffic
  - Build the science data perimeter using performant technologies
- If we can do this, we can scale up data performance to match the data scale



ESnet

Limitation Is An Enabler

- Only approved services in the Science DMZ
  - Services that can be secured with stateless firewalls
  - Complexity is the enemy
- Allows space for performant security
- Control vs. Data can be handled separately
  - Enterprise firewall is fine for control messages, auth, etc.
  - Need Science DMZ for high performance data



ESnet

Lessons Learned Over Time

- Enterprise security team will try to eliminate Science DMZ
  - “The firewall is faster now”
  - “The DMZ isn’t secure because it doesn’t have a firewall, so we will put in a firewall to make it secure”
  - Multiple variations of this
- Workflows change over time
  - You may not know about the changes in advance
  - Science need drives mission
  - Architecture must accommodate changes needed by science
- Must find a way to make performance persistent
  - Then, users can rely on it

Lessons Learned Over Time (2)

- Most users won't ask for help
  - They just suffer in silence or change to fit limitations
  - Many "know" that things can't be better
- Proving a negative is hard
  - Very difficult to find evidence that users would be doing better if they could
- Answer: build it right, and teach users to use it
- The results can be impressive

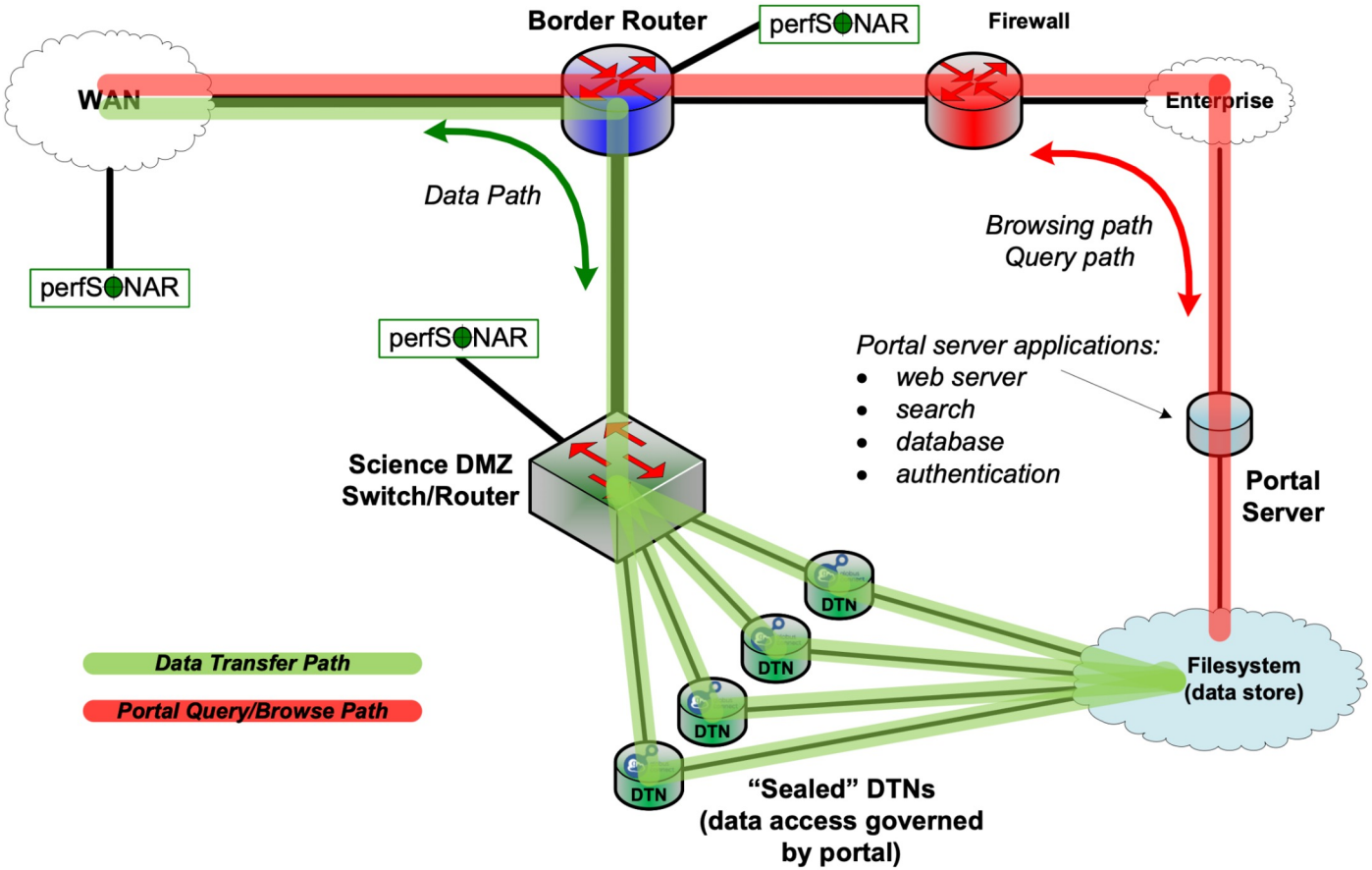
A Note About Data Locality

- I hear a lot of people say “don’t move the data”
  - “Just move the compute to the data!”
- Of course! Do that if you can. But there are often difficulties
  - Open access data, closed access computing
    - Code execution is often tightly controlled
    - Scientist often doesn’t have permission to run analysis next to the existing data store → need to move the data
  - Specific hardware required
    - Specialized systems, computing scale, other requirements
    - Move the data to the specific system
  - I don’t have to pay to run on my local system
    - More creativity if I run there
    - I have to move the data to my cluster
- Therefore, we must support data movement in the general case

What is Possible?

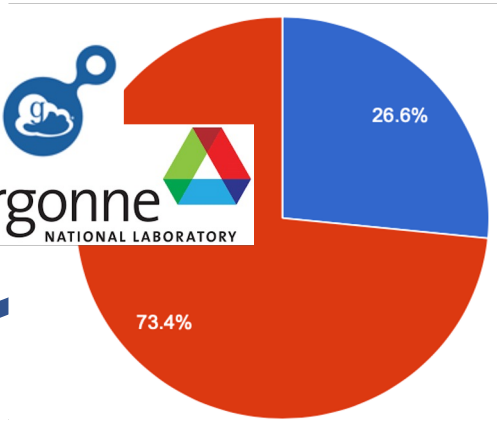
- I’m going to go through a few real-world examples
  - High performance data workflows across the perimeter
  - Multiple Science DMZ design pattern examples
- First examples: file transfer using DTNs
  - Classical use of Science DMZ
  - Transfer files or objects from one storage system to another
  - Valuable if data needed at remote location for some time
  - Many, many workflows do this
- Basic portal or cluster/HPC implementation of Science DMZ

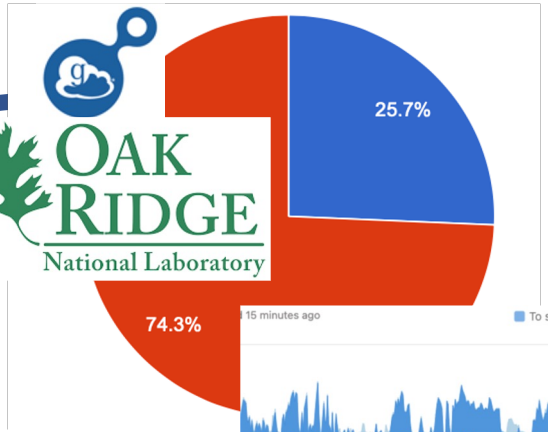
Portal Design Pattern



ESnet

Replication: Globus used to move 7.4 petabytes of data from **Livermore** to **Argonne** and **Oak Ridge** Labs





4 to 6 GB/s

1.5 GB/s

17,347,671 directories and 28,907,532 files
From February 12 to May 4, 2022

To ALCF

No	Datasets	From	Requested	Completed	Status	Directories	Files	Bytes Transferred	Faults	Rate
1	/css03_data/CMIP6/CMIP/MOHC/HadGEM3-GC31-LL/historical	LLNL	2022-03-10 13:19:03		ACTIVE (20%)	6125	6515	6138646980430	0	832 MB/s
2	/css03_data/CMIP6/CMIP/MIROC/MIROC-ES2L/historical	LLNL	2022-03-10 05:35:04		ACTIVE (79%)	37994	409095	24611252181300	12	699 MB/s
3	/css03_data/CMIP6/CMIP/MOHC/HadGEM3-GC31-LL/amip	LLNL	2022-03-10 12:12:03	2022-03-10 13:18:06	SUCCEEDED	3908	1892	3091419704055	0	780 MB/s
4	/css03_data/CMIP6/CMIP/MOHC/HadGEM3-GC31-LL/abrupt-4xCO2	LLNL	2022-03-10 11:40:03	2022-03-10 12:11:57	SUCCEEDED	1121	953	1559216858805	0	814 MB/s

To OLCF

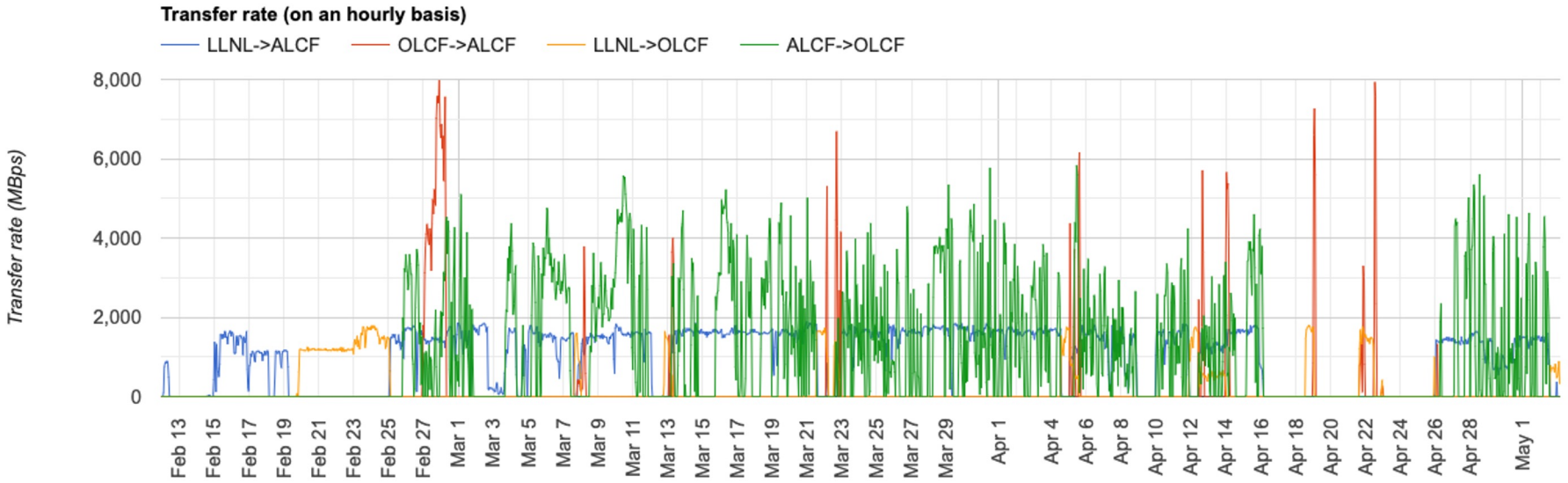
No	Datasets	From	Requested	Completed	Status	Directories	Files	Bytes Transferred	Faults	Rate
1	/css03_data/CMIP6/CMIP/MIROC/MIROC-ES2L/esm-piControl	ALCF	2022-03-10 15:14:03		ACTIVE (25%)	1236	40039	1407934487539	0	2.93 GB/s
2	/css03_data/CMIP6/CMIP/IPSL/IPSL-CM6A-LR/historical	ALCF	2022-03-09 22:02:03		ACTIVE (77%)	73193	36610	129503497305534	1	2.08 GB/s
3	/css03_data/CMIP6/CMIP/MIROC/MIROC-ES2L/esm-hist	ALCF	2022-03-10 14:51:04	2022-03-10 15:13:24	SUCCEEDED	3706	39663	2973432261868	0	2.22 GB/s
4	/css03_data/CMIP6/CMIP/MIROC/MIROC-ES2L/amip	ALCF	2022-03-10 14:47:03	2022-03-10 14:50:22	SUCCEEDED	3126	12284	446324011629	0	2.25 GB/s

Slide credit: Ian Foster, ANL

Shown: Status as of March 10, 2022

Replication: Transfer Strategy

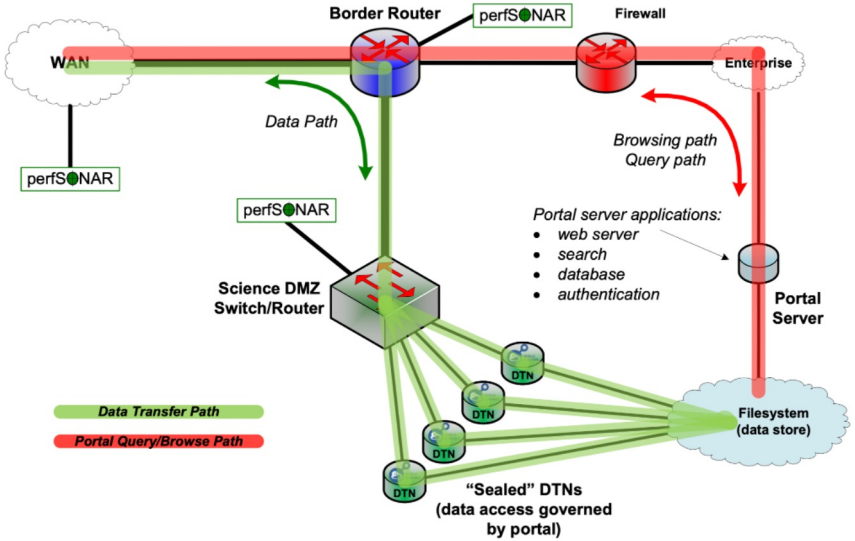
- Keep the slow part busy
  - Sync between the fast systems
  - Fail over during maintenance
- Note performance differences
  - Outbound from LLNL much slower
  - Sync between ALCF and OLCF is fast



ESnet

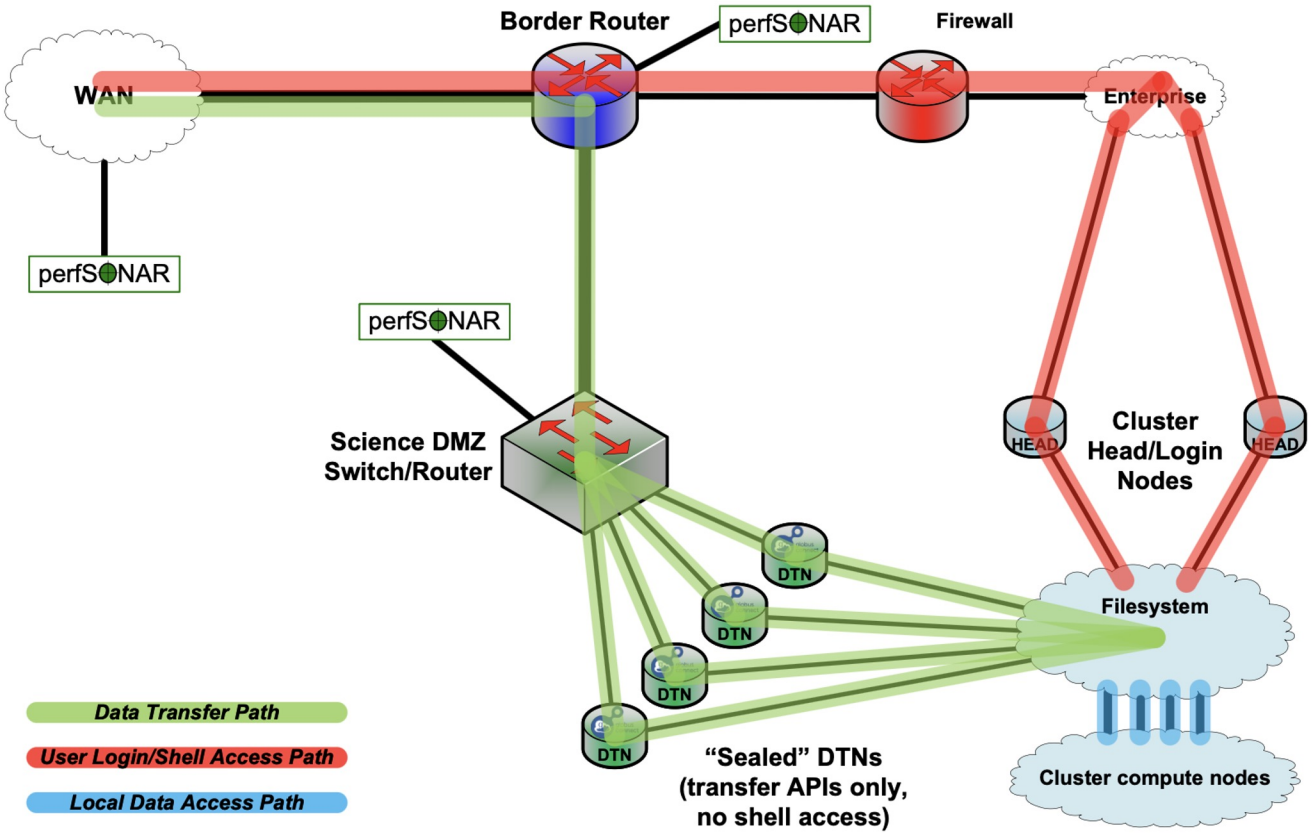
Portal Data Replication: Discussion

- Data portal infrastructure already had DTNs in place
- All the replication team had to do was use them
- If they had had to use the portal web server, the task would have been impossible
- New use for the DTNs, very valuable to have them there



ESnet

HPC/Cluster Design Pattern



ESnet

ESnet
ENERGY SCIENCES NETWORK

HOME › SITES »

National Energy Research Scientific Computing Center

Interfaces | Flow

Website http://www.nersc.gov/

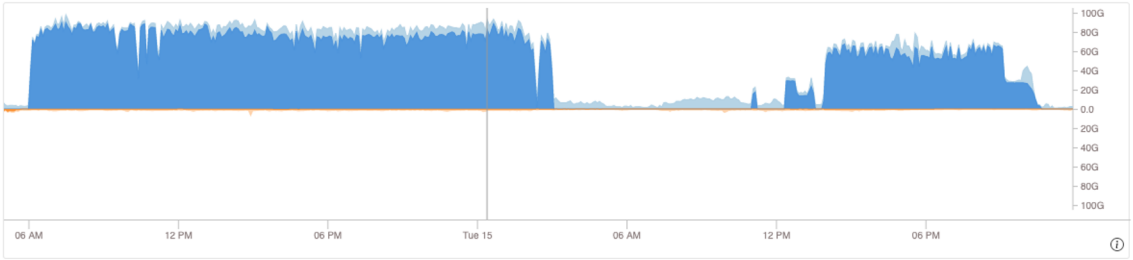
BREAKDOWN Autonomous Systems (origin)

TIME day week month custom

TOTAL SITE TRAFFIC

Tue Sep 15 12:24 AM 2020

To site 81.6Gbps | From site 2.49Gbps



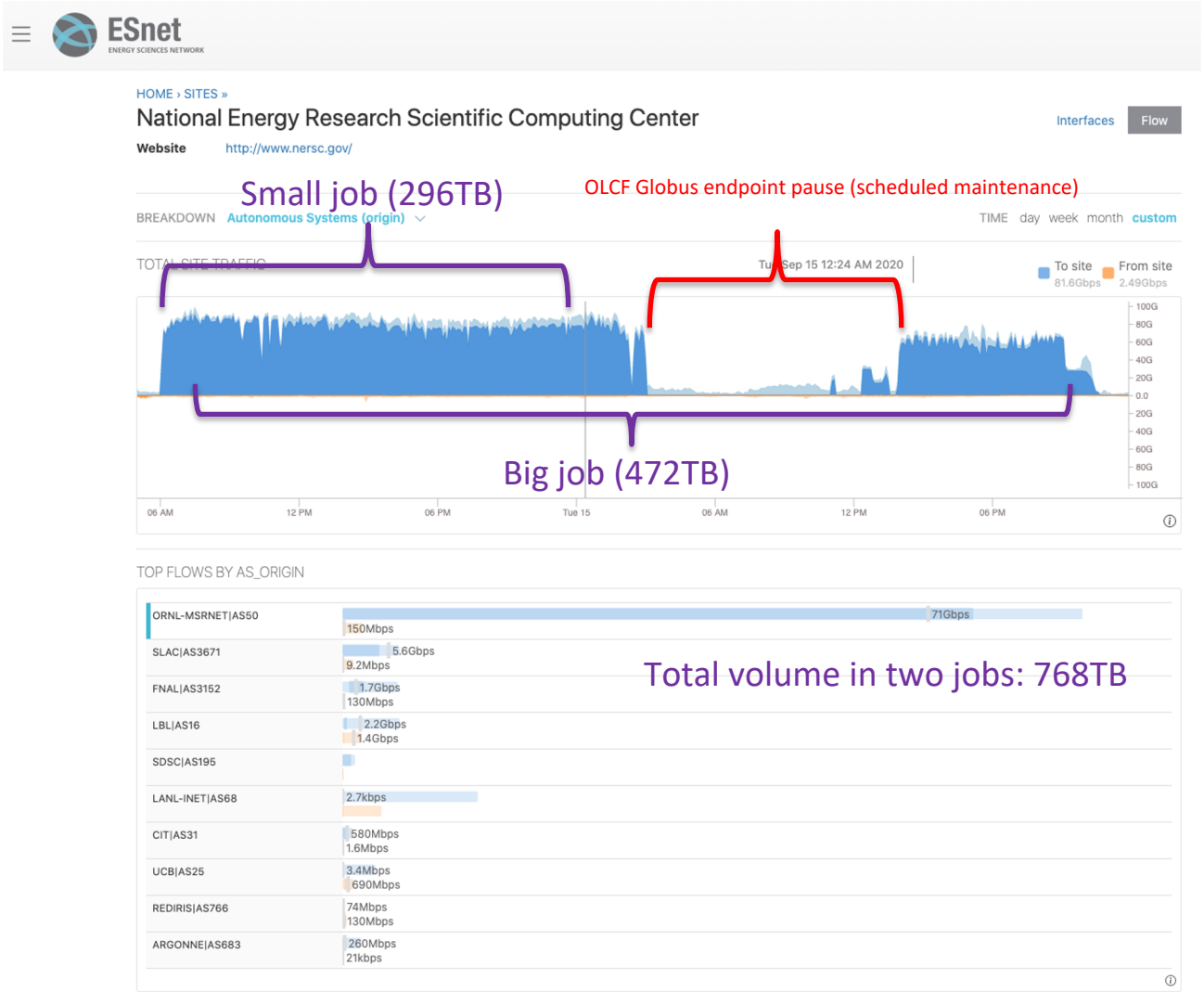
TOP FLOWS BY AS_ORIGIN

ORNL-MSRNET|AS50	71Gbps	150Mbps
SLAC|AS3671	5.6Gbps	9.2Mbps
FNAL|AS3152	1.7Gbps	130Mbps
LBL|AS16	2.2Gbps	1.4Gbps
SDSC|AS195		
LANL-INET|AS68	2.7kbps	
CIT|AS31	580Mbps	1.6Mbps
UCB|AS25	3.4Mbps	690Mbps
REDIRIS|AS766	74Mbps	130Mbps
ARGONNE|AS683	260Mbps	21kbps

ESnet

Large Scale Data Transfer

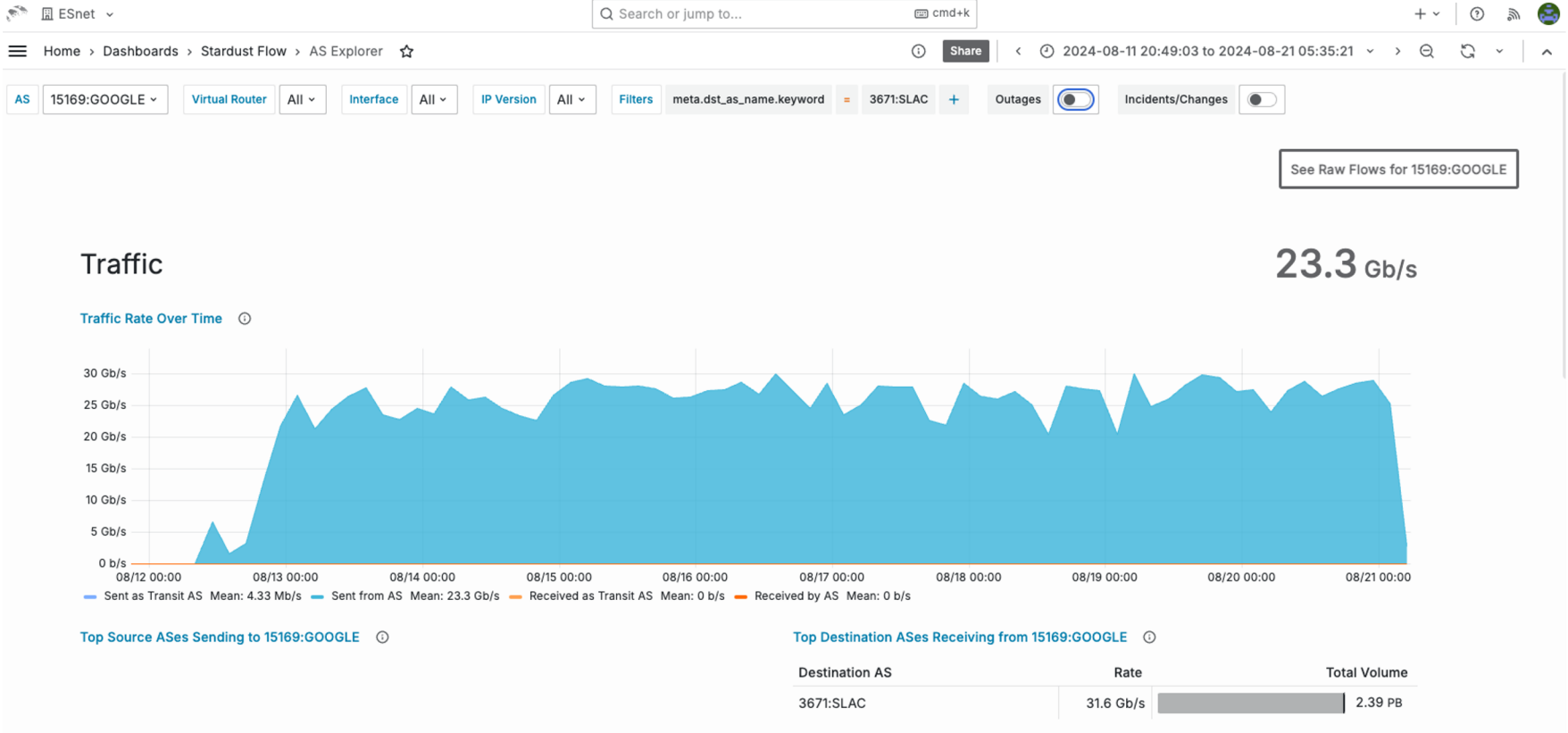
- Two Globus transfers from OLCF to NERSC were started by the same user (presumably for the same project – DESI – which is dark energy/cosmology) within 40 minutes of each other.
  - One transfer was ~350k files, ~296TB
  - Other transfer was ~1.8M files, ~472TB
  - Data transfer rate peaks over 80Gbps
  - Total transfer volume ~768TB
  - Total wall clock time ~39 hours
- Current tools (Filesystems, DTNs, Globus) handle petascale data sets easily


ESnet
ENERGY SCIENCES NETWORK
HOME › SITES »
National Energy Research Scientific Computing Center
Interfaces
Flow
Website
http://www.nersc.gov/
Small job (296TB)
OLCF Globus endpoint pause (scheduled maintenance)
BREAKDOWN
Autonomous Systems (origin)
TIME
day
week
month
custom
TOTAL SITE TRAFFIC
Tue Sep 15 12:24 AM 2020
To site
81.6Gbps
From site
2.49Gbps
100G
80G
60G
40G
20G
0.0
20G
40G
60G
80G
100G
06 AM
12 PM
06 PM
Tue 15
06 AM
12 PM
06 PM
Big job (472TB)
TOP FLOWS BY AS_ORIGIN
ORNL-MSRNET|AS50
71Gbps
150Mbps
SLAC|AS3671
5.6Gbps
9.2Mbps
FNAL|AS3152
1.7Gbps
130Mbps
LBL|AS16
2.2Gbps
1.4Gbps
SDSC|AS195
LANL-INET|AS68
2.7kbps
CIT|AS31
580Mbps
1.6Mbps
UCB|AS25
3.4Mbps
690Mbps
REDIRIS|AS766
74Mbps
130Mbps
ARGONNE|AS683
260Mbps
21kbps
Total volume in two jobs: 768TB

Key Points

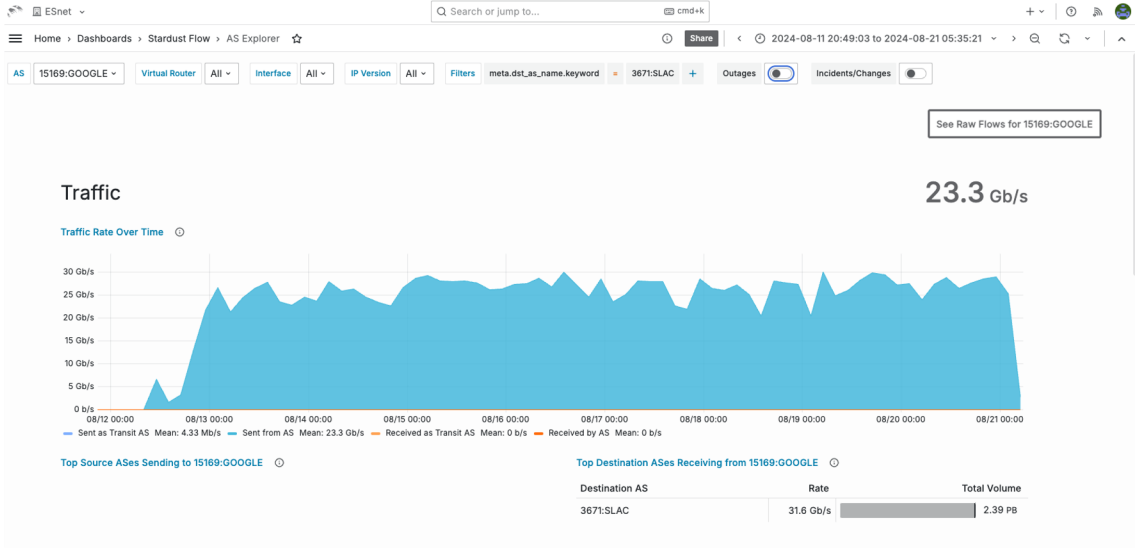
- Sophisticated tools bring multiple benefits
  - Both transfers experienced random data corruption (multiple checksum failures) which was automatically corrected by Globus (data set is ¾ of a petabyte….a lot can happen with that much data)
    - User had to expend zero effort to find and fix the corrupted files
  - OLCF endpoint paused for a workday for normal scheduled maintenance, and transfer resumed after maintenance concluded
    - OLCF staff did not have to worry about the user when planning maintenance
    - User had to expend zero effort to work around HPC facility maintenance
  - Easy to use tools with automated fault recovery reduce human effort
- Large scale data transfer is a key enabler for scientific productivity
  - Benefits of large-scale HPC allocation brought back to collaborators
- Normal, routine operations – no magic

Large Scale Replication From Cloud



Large Scale Replication From Cloud

- Rubin/LSST project at SLAC
- Large scale data replication
  - 250M files
  - Over 2.3PB
  - Approx. 8.5 days
- Six DTNs at SLAC
  - K8 environment
  - Google Cloud Storage Transfer application
- Traffic traversed three Squid proxies at SLAC



ESnet

Google to SLAC Replication - Key Points

- Petascale data transfer from cloud works!
- Small number of DTNs can achieve good performance
- Containers are viable for high performance applications, including I/O intensive

File Transfer: Discussion

- When the infrastructure is already in place, users just use it
  - Nobody has to ask permission
  - No complex config for users to do
  - No need to debug the end to end path before using it
- Users can concentrate on the workflow
- Simple protocols, high performance
  - Simple policy: Support DTN driven data transfers
  - Simple protocols: Globus, cloud-native object access, etc.
  - Safe to operate when secured with performant technologies, in keeping with Science DMZ

Storage Becoming A Limitation

- In many cases, storage I/O is now the bottleneck
- Answer: streaming
- One or both ends of the data transfer is volatile
  - Instrument
  - Memory
- Removes storage limitations from the workflow

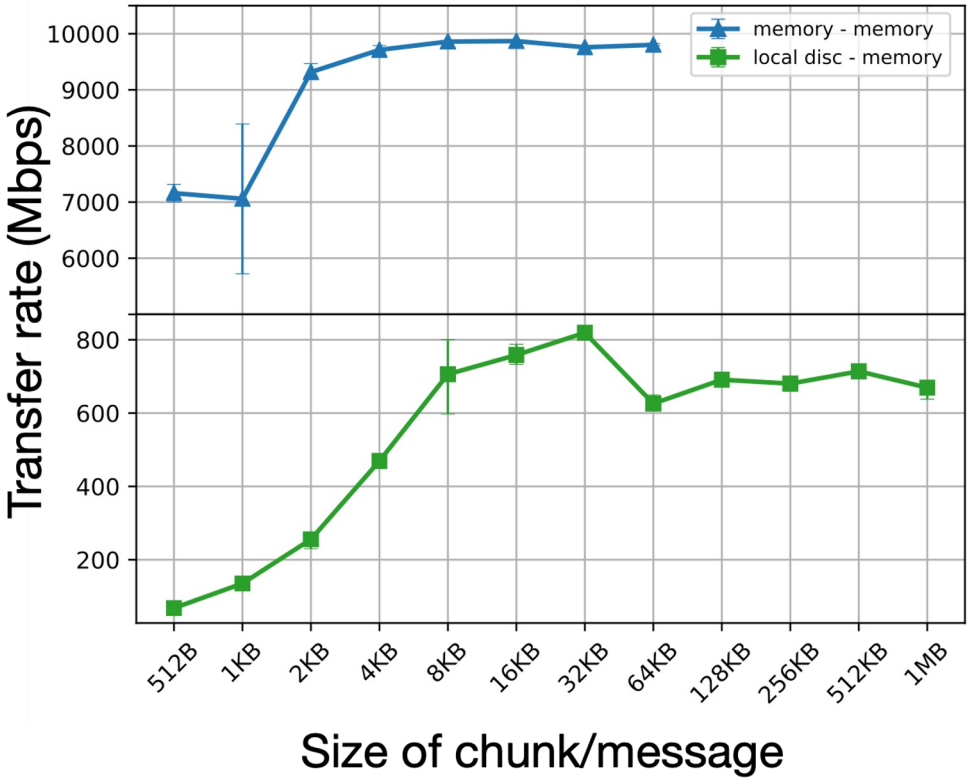


Image: Raj Kettimuthu, ANL

Stream From Site Storage To Cloud

- Workflow reads from site storage into cloud analysis job
- No local storage in the cloud environment
  - Compute job analyzes the data and then drops it
  - Need the data again? Fetch it again!
  - No cloud storage costs
- High performance perimeter required for success
  - Packet loss or other network issues reduce compute performance
  - Easy to do with Science DMZ
- Two examples follow

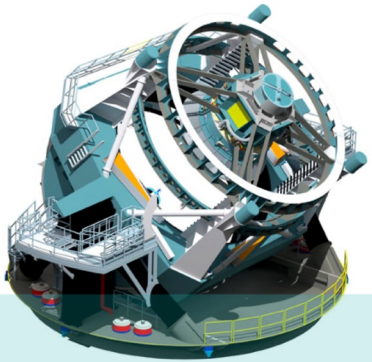
Analysis in Cloud: Vera Rubin Observatory

- Two parts to Rubin workflow
- Acquisition + alerting
  - Data comes from the telescope in Chile to SLAC in USA
    - World's largest digital camera on telescope in Chile
    - Data path traverses many networks, including ESnet
    - Prompt analysis and alert generation at SLAC
  - Multinational networking collaboration required to get the data to SLAC
  - Alerts sent to scientific community
- **User analysis**
  - **Data releases distributed from SLAC to scientists**
  - **Jupyter analysis environment for users in Google Cloud**
    - **Cloud job reads data from SLAC into Jupyter notebook**
    - **Collaboration will scale this up as data accumulates**

Rubin Data Management System includes both making data and serving data to users

Raw Data: 20TB/night

Sequential 30s images covering the entire visible sky every few days



LSST SCIENCE PIPELINES

Prompt Data Products

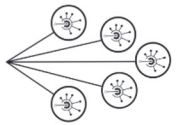
- Alerts incl. science, template and difference image cutouts — 60s — via Alert Streams
- Catalogs of detections incl. difference images, transient, variable & solar system sources — 24h — via Prompt Products
- Raw & processed visit images (PVIs), difference images — 80h — via Image Services

Data Release Data Products

Final 10yr Data Release:

- Images: 5.5 million x 3.2 Gpixels
- Catalog: 15PB, 37 billion objects

via Data Releases



Community Brokers

Rubin Data Access Centres (DACs)

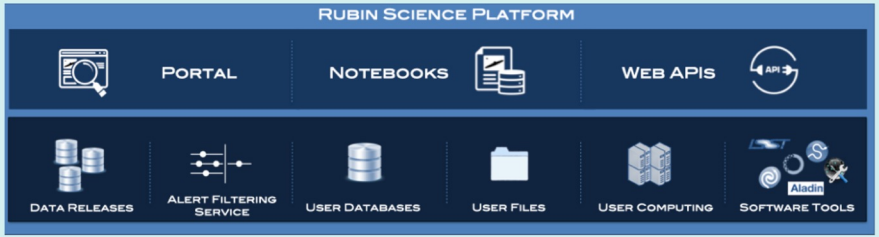
USA (USDF)
Chile (CLDF)
France (FRDF)
United Kingdom (UKDF)

Independent Data Access Centers (IDACs)

Access to proprietary data and the Science Platform require Rubin data rights

Rubin Science Platform

Provides access to LSST Data Products and services for all science users and project staff.

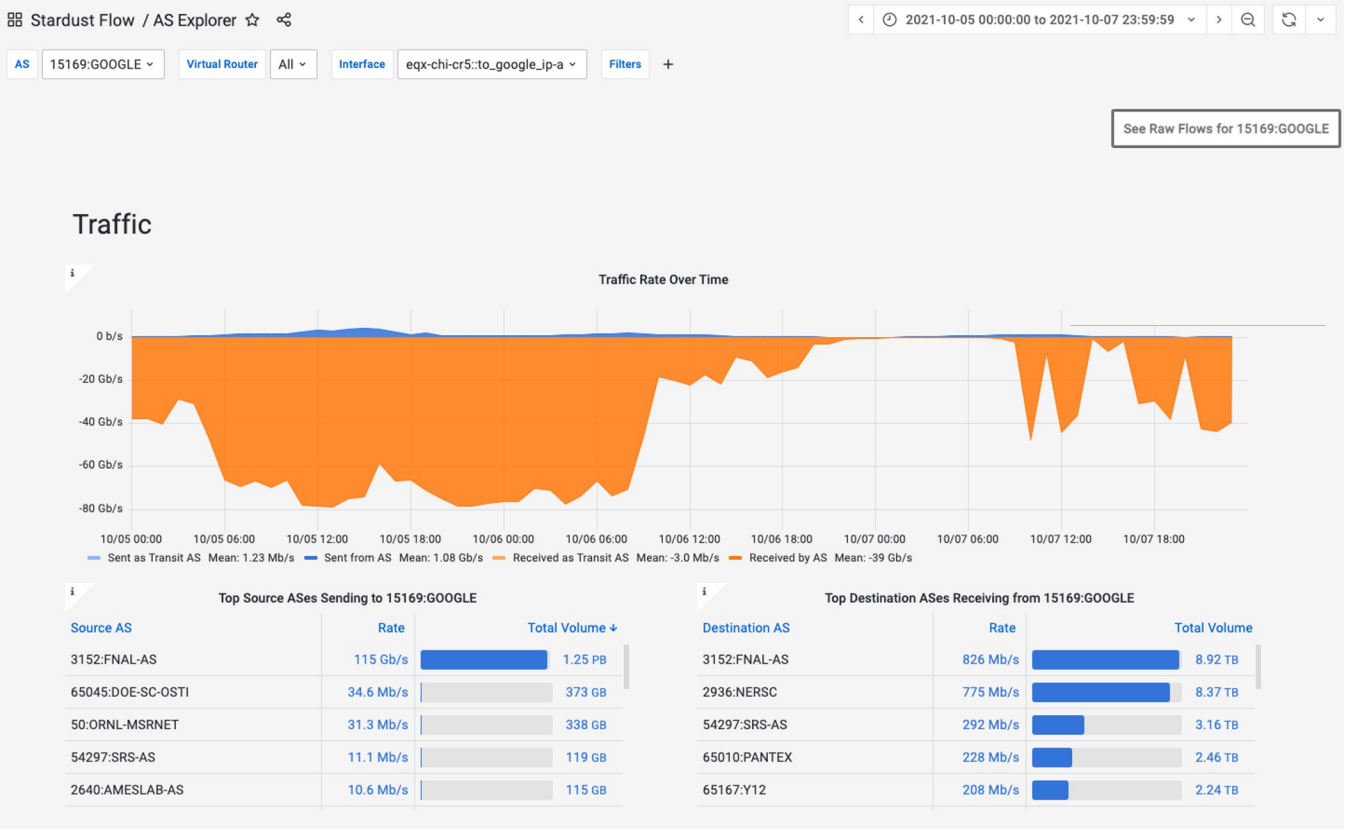


Credit: Leanne Guy

Training and Inference in Commercial Cloud

- DUNE collaboration
- ML based data reduction
- Near-real-time inferencing classifiers supporting neutrino physics workflows.
- Hybrid workflow
  - Data and CPU computing at FNAL
  - GPUs in Google
- Impossible without high performance perimeter
- Works great if you do have high performance perimeter

Data From Lab to Cloud For Inference



- ~80Gbps for hours at a time
- Over 1PB moved in ~30 hours
- No need to replicate the data first (no file transfer workflow)
- No in-cloud storage costs

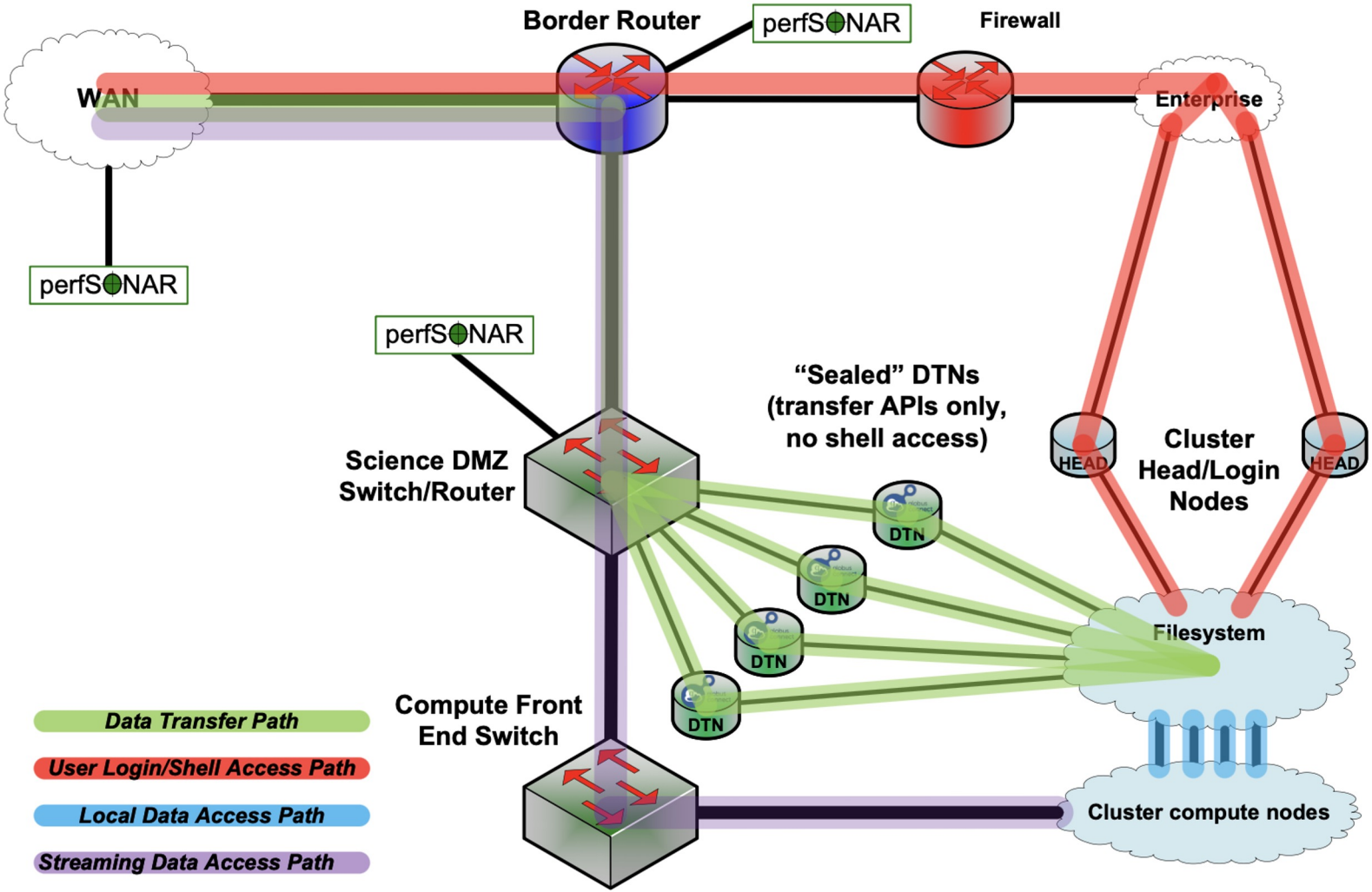
ESnet

Site to Cloud Streaming: Discussion

- Data storage and servers are on infrastructure owned by the community/collaboration
  - Storage systems with DTNs
  - POSIX or object store, doesn't matter
- Analysis/compute clients in the cloud fetch the data directly
  - No copying to cloud storage first
  - Just read the data in from remote storage, analyze it
- Site perimeter design: DTNs for outbound data
  - Simple security policy (address and port filters)
  - Safe to operate with performant security technologies
- Science workflow requires high performance
  - Scaling enterprise firewalls to support this is difficult and expensive
  - No added benefit from enterprise firewalls (just address/port)

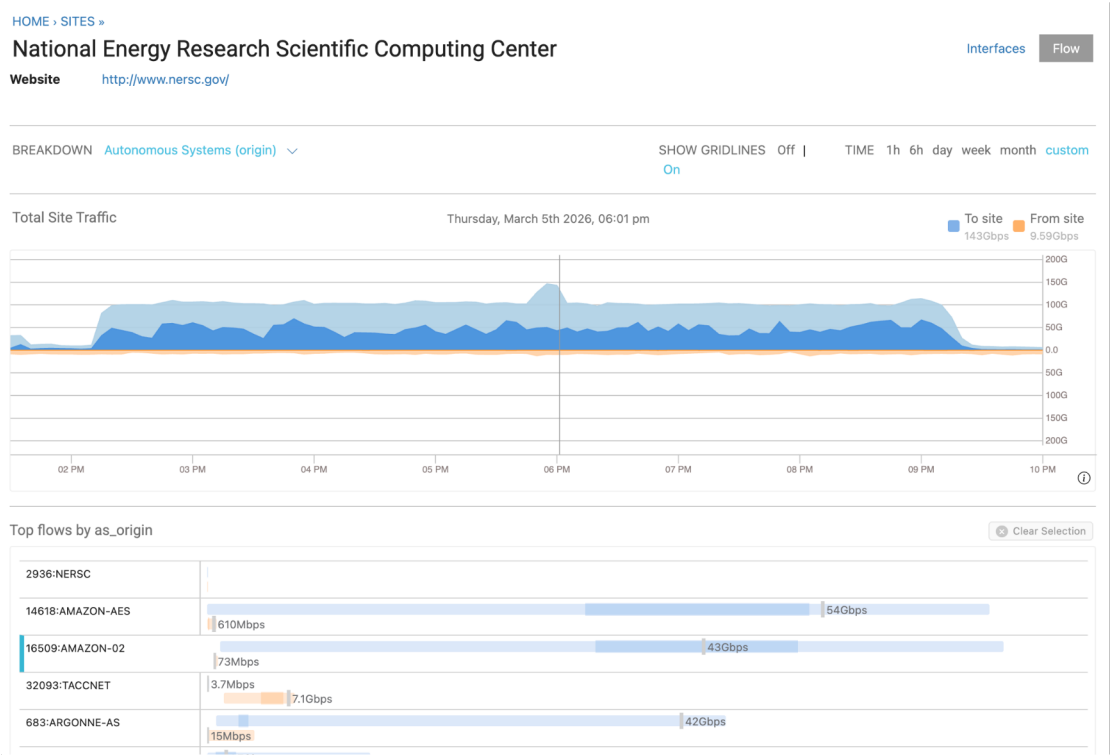
Stream From Cloud To HPC

- These workflows go in the other direction
- Data comes from the cloud, across the perimeter, and into a running analysis job on a supercomputer
- Can be many many nodes on both sides
- Bandwidth can get quite large
- Not limited by HPC storage I/O
  - No need to wait for data replication
  - Workflow can go faster than the DTN cluster


Border Router
perfSONAR
Firewall
WAN
Enterprise
perfSONAR
perfSONAR
"Sealed" DTNs
(transfer APIs only,
no shell access)
Science DMZ
Switch/Router
HEAD
Cluster
Head/Login
Nodes
HEAD
DTN
DTN
DTN
DTN
Filesystem
Compute Front
End Switch
Cluster compute nodes
Data Transfer Path
User Login/Shell Access Path
Local Data Access Path
Streaming Data Access Path

Direct Transfer to Compute: Amazon to NERSC

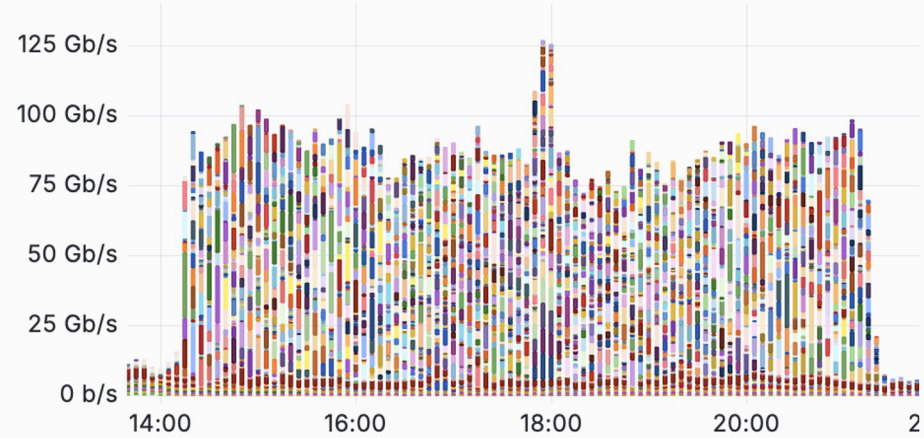
- Data set in commercial cloud
- Weather forecasting code
- No replication before job launch
- Direct ingest by compute job takes pressure off the HPC facility storage system



ESnet

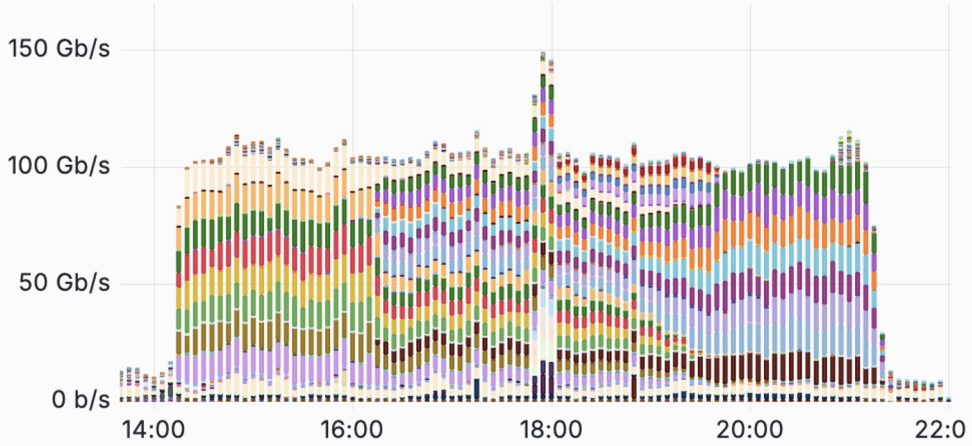
Many cloud servers to a few compute nodes

Traffic Rate Over Time of Top 1000 Source IPs By Volume Wh...



Name	Max
16.182.33.153	23.5 Gb
52.217.197.241	21.0 Gb
3.5.27.127	20.8 Gb
16.15.217.51	19.6 Gb

Traffic Rate Over Time of Top 1000 Destination IPs By Volume...

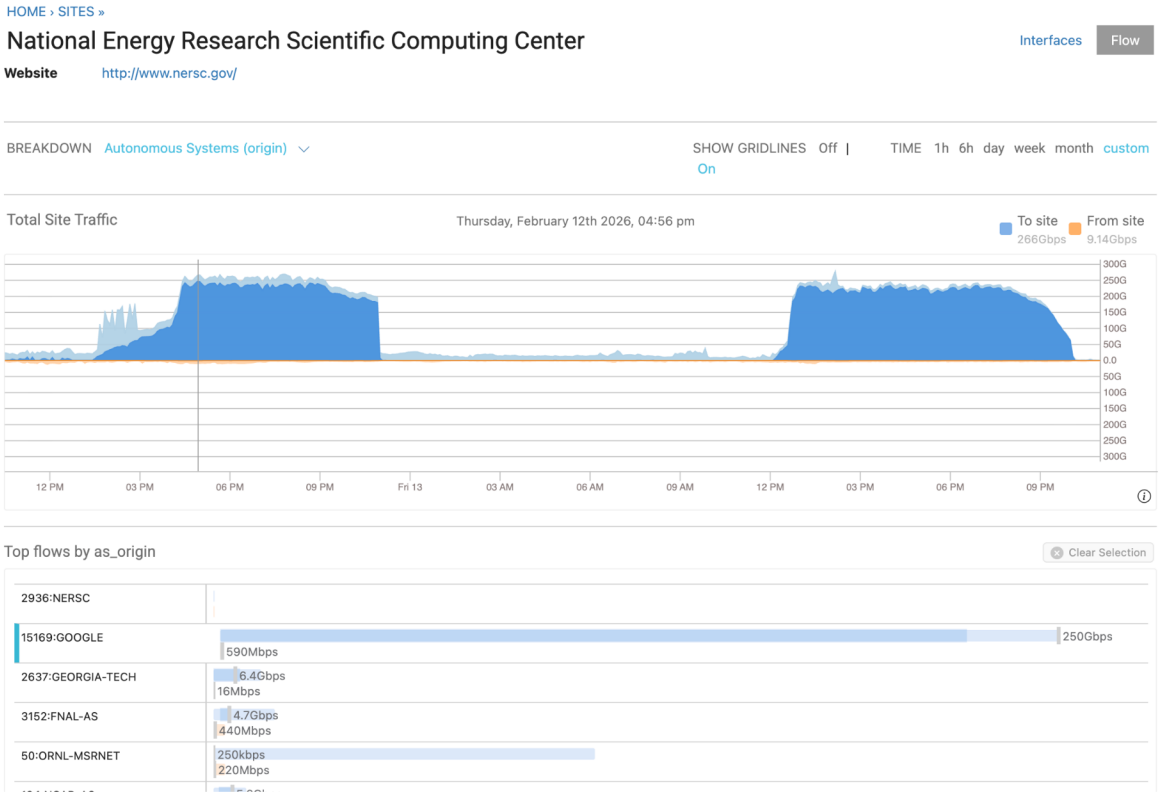


Name	Max
128.55.75.239	14.9 Gb/s
128.55.77.32	14.3 Gb/s
128.55.73.78	13.8 Gb/s
128.55.77.24	13.7 Gb/s

ESnet

Direct Transfer to Compute: Google to NERSC

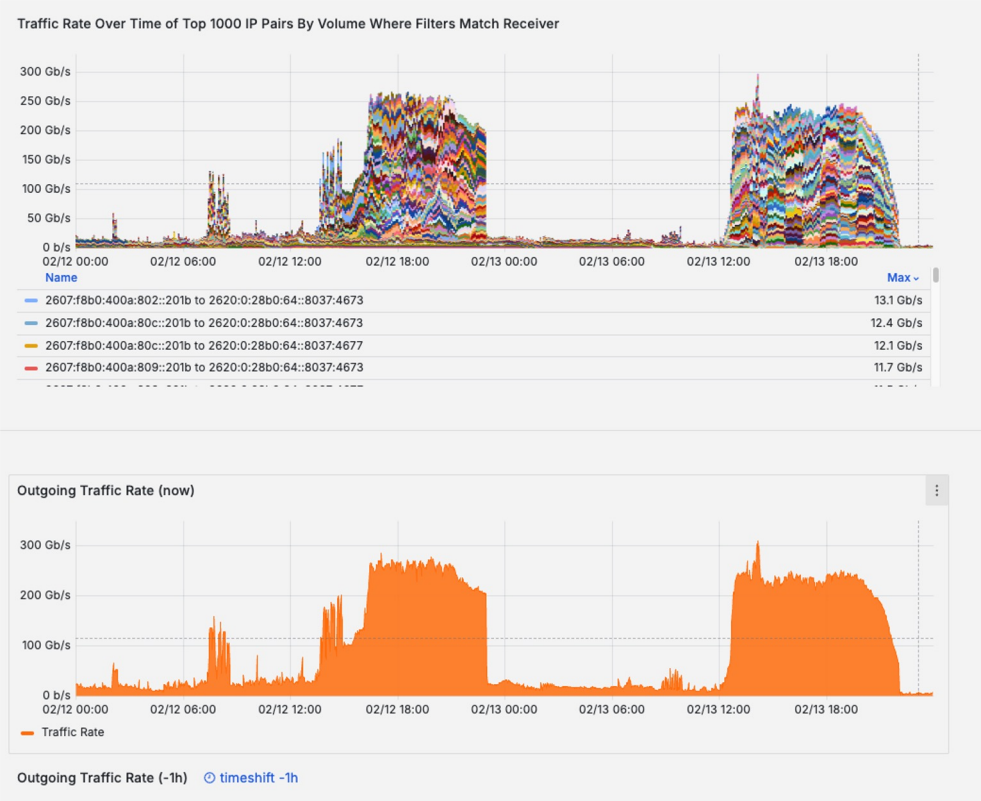
- Large (over 5PB) public data set in commercial cloud
  - ECMWF ERA5
  - AI-ready
- ML-accelerated earth system model application
- Direct data ingest into running code



ESnet

Flow statistics: massively parallel

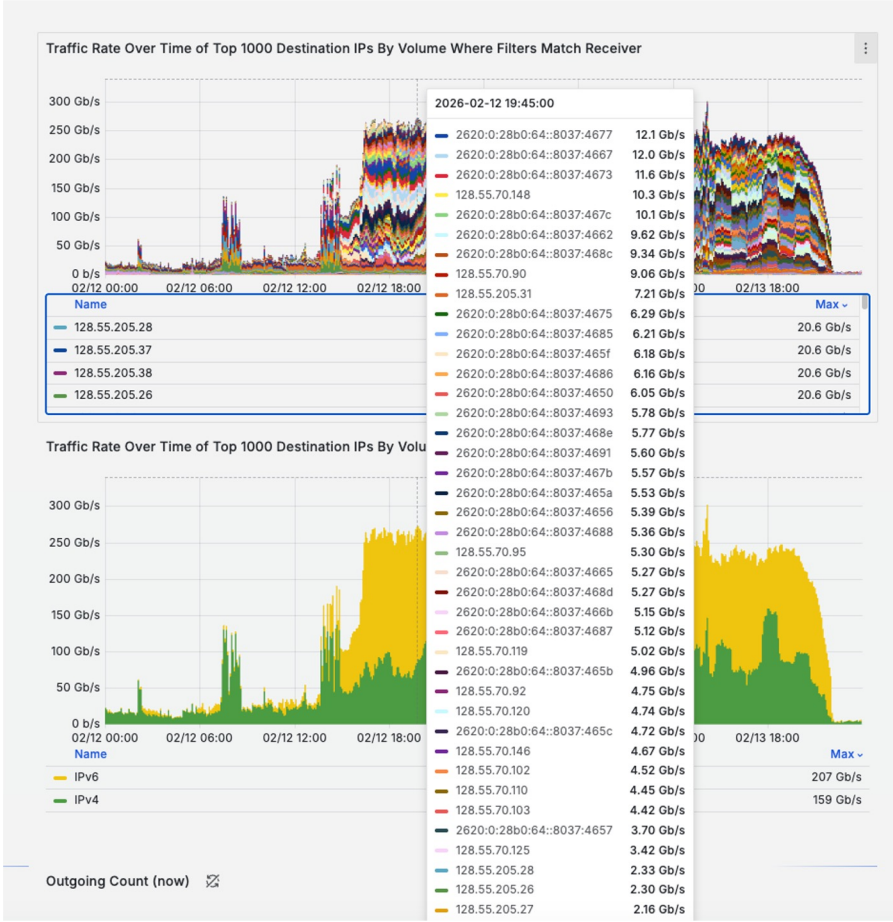
- Host breakdown:
  - Many, many parallel streams in these transfers
- Many cloud hosts delivering over 10Gbps
- User was able to just launch the job and move forward.
  - No special setup
  - No debug before use



ESnet

IPv6 Works!

- A lot of this transfer occurred over IPv6
- It was already set up and configured - no need for user to care
- IPv6 is real now, even in commercial cloud environments



ESnet

Evolution Of Data-Intensive Science

- Science DMZ enabled high performance file transfer
  - DTNs
  - Performant tools
- Now, we are running into the limitations of storage
  - File transfer before job launch increases friction
  - Storage layer is slow and expensive
- New workflows drive ever higher performance across network boundaries
- But this only works if the perimeter is performant

In conclusion – ESnet's vision:

Scientific progress will be **completely unconstrained** by the physical location of instruments, people, computational resources, or data.

U.S. DEPARTMENT of ENERGY

BERKELEY LAB

ESnet
ENERGY SCIENCES NETWORK

Thanks!

Eli Dart

dart@es.net

https://my.es.net/

https://www.es.net/

https://fasterdata.es.net/

U.S. DEPARTMENT of ENERGY

BERKELEY LAB

ESnet
ENERGY SCIENCES NETWORK

Thanks!

Eli Dart

dart@es.net

https://my.es.net/

https://www.es.net/

https://fasterdata.es.net/