



An Introduction to NetSage

Andy Lake, Co-PI NetSage
ESnet/Lawrence Berkeley Lab

European perfSONAR User Workshop

June 5, 2019



What is NetSage?

- Project
 - Originally funded by the NSF international program to monitor and visualize transfers across international links
 - Collaboration of Indiana University, Lawrence Berkeley National Lab, UC Davis, and University of Hawaii
- Software
 - A set of Grafana dashboards and collectors for multiple data sources
- Deployment
 - International networks dashboards (the original):
<http://portal.netsage.global>
 - Multiple other deployments. See <https://www.netsage.global>



NetSage Focus on Use Cases

- Bandwidth Dashboard: <http://portal.netsage.global>
 - How used are the links?
 - Where are congestion points?
- Flow Data Dashboards
 - What are the top sites using the my circuits?
 - What are the top sources/destinations for an organization?
 - Who's using my archive?

NetSage Data Sources

- SNMP data (Passive) - Basic bandwidth data
- perfSONAR (Active)
 - Active tests between sites
- Flow data from routers (Passive)
 - Only de-identified data collected by NetSage
- Tstat-based traffic analysis for archives (Passive)
 - TCP flow statistics: congestion window size, number of packets retransmitted, etc
 - Also de-identified before stored



NetSage Privacy

- NetSage is committed to privacy, and preemptively addressing any security or data sharing concerns
 - No PII collected
 - Remove the last octet from IP address
 - Only keep data on flows over 1M
- Data Privacy Policy
 - <http://www.netsage.global/home/netsage-privacy-policy>
- Data Flow Data Retention (De-Identification) Policy
 - <https://tinyurl.com/netsage-deid>
- Prototypes are behind a password until we're told to make it public



NetSage Science Registry

- A tool to understand science use of networks
- As part of flow collection process, we add tags
 - ASN
 - Organization
 - Science project and science discipline, if known
- Then the end of the IP address is removed
- Currently ~300 entries

NetSage International Collection Points

- SNMP and Flow in:
 - New York
 - Chicago (Starlight)
 - Miami
 - Seattle
 - Honolulu
 - Guam
- Tstat in:
 - TACC
 - NERSC
 - NOAA
 - *U Hawaii (Astronomy)**

Let's Look at some dashboards....

- SNMP Data
 - [Bandwidth](#)
 - [Bandwidth Patterns](#)
- perfSONAR Data
 - [Active Testing Maps](#)
 - [Loss Patterns](#)
- Flow Data:
 - [By Selectable](#)
 - [By Science Discipline](#)
- Tstat Data:
 - [Flow Data for Data Archives](#)

Wrap-up

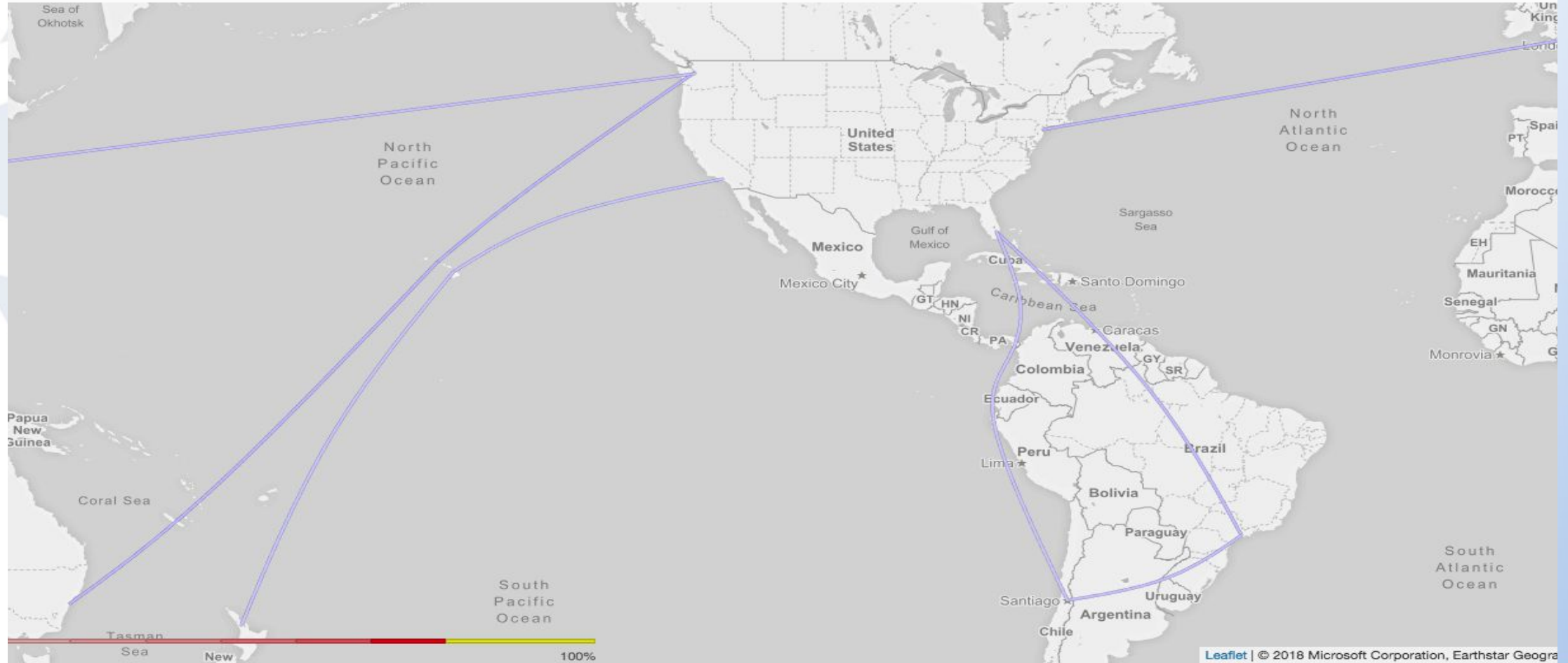
- NetSage is a real-world tool that demonstrates the power of bringing multiple metrics together to do useful things
- From a perfSONAR perspective (since this a perfSONAR workshop), NetSage is exactly the type of tool the perfSONAR project wants to enable
- perfSONAR data is useful in its own right, but the next stage of evolution is how to combine it in useful ways with other data. NetSage is a leader in this space.



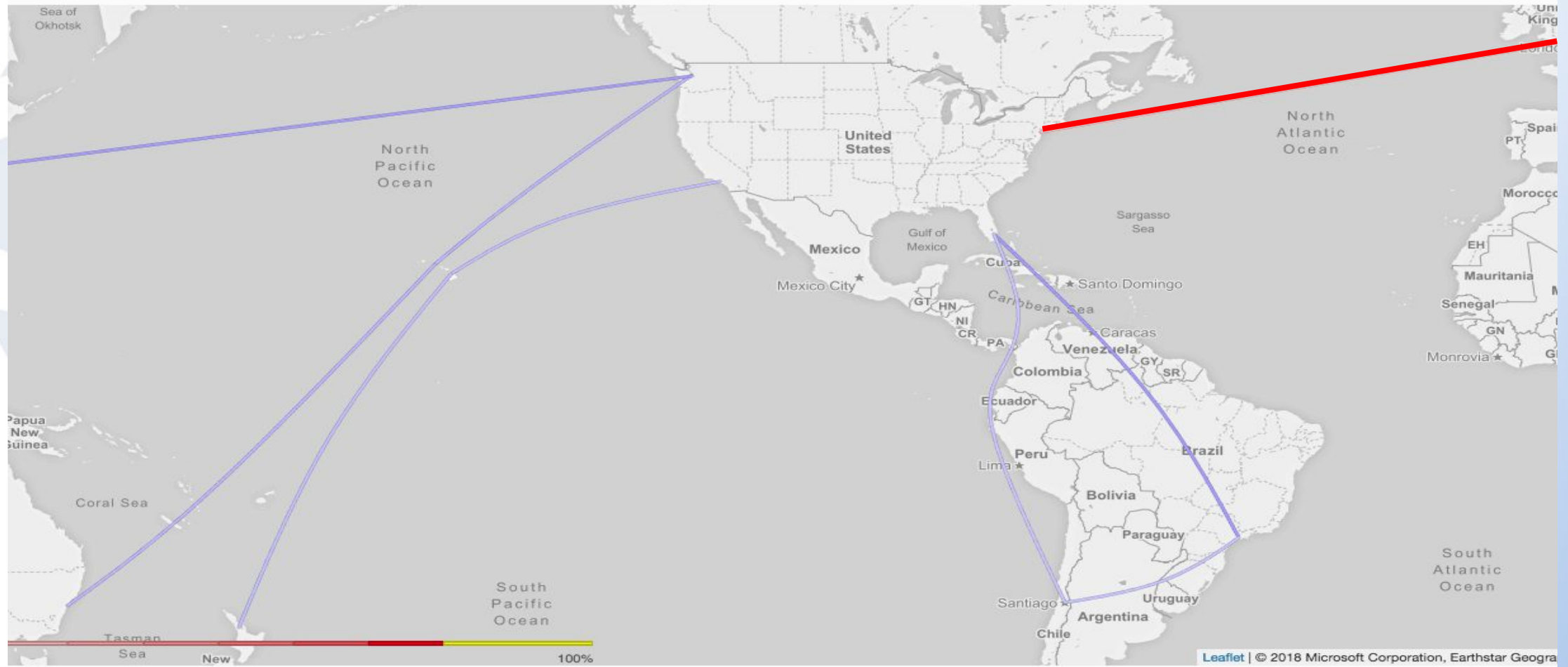
Extra Slides



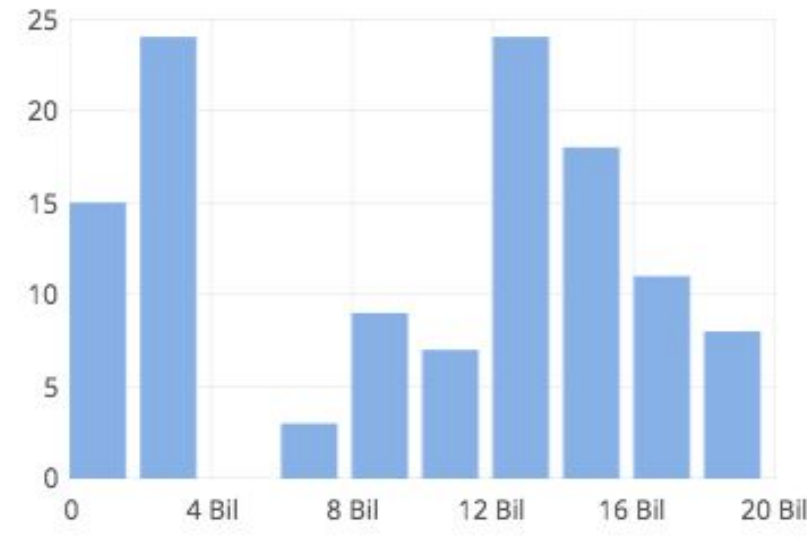
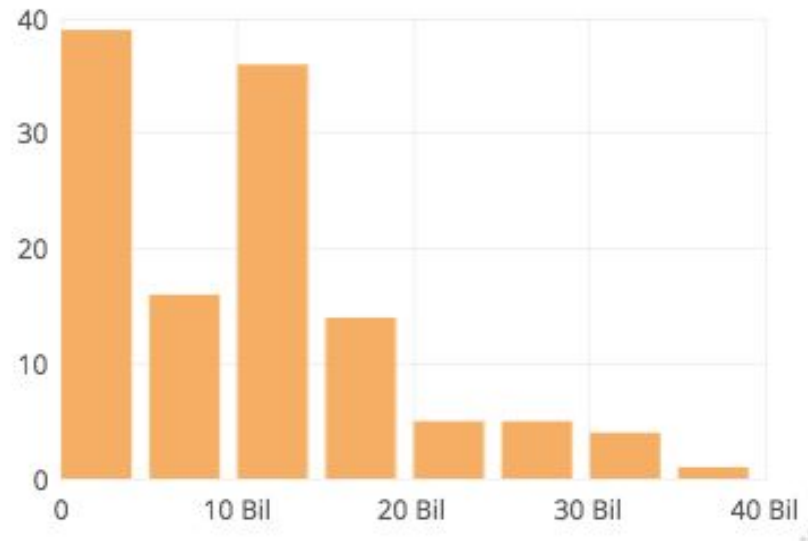
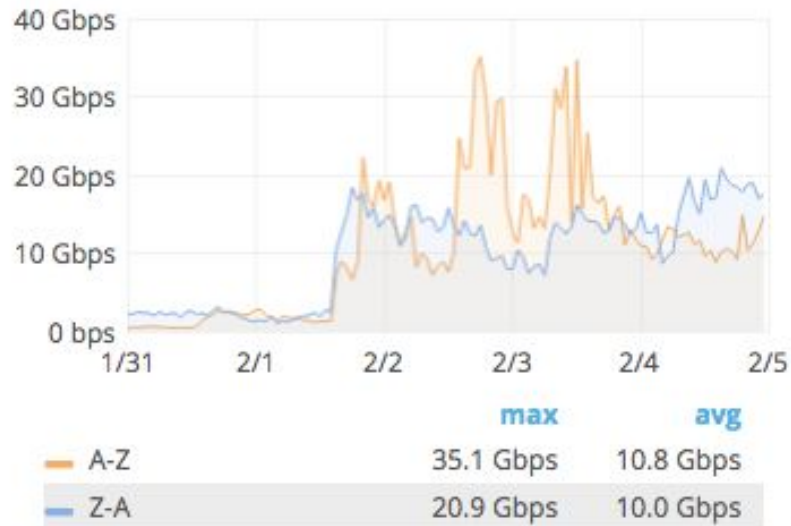
Shift of focus to PI for NEAAR - Normal day: <http://portal.netsage.global>



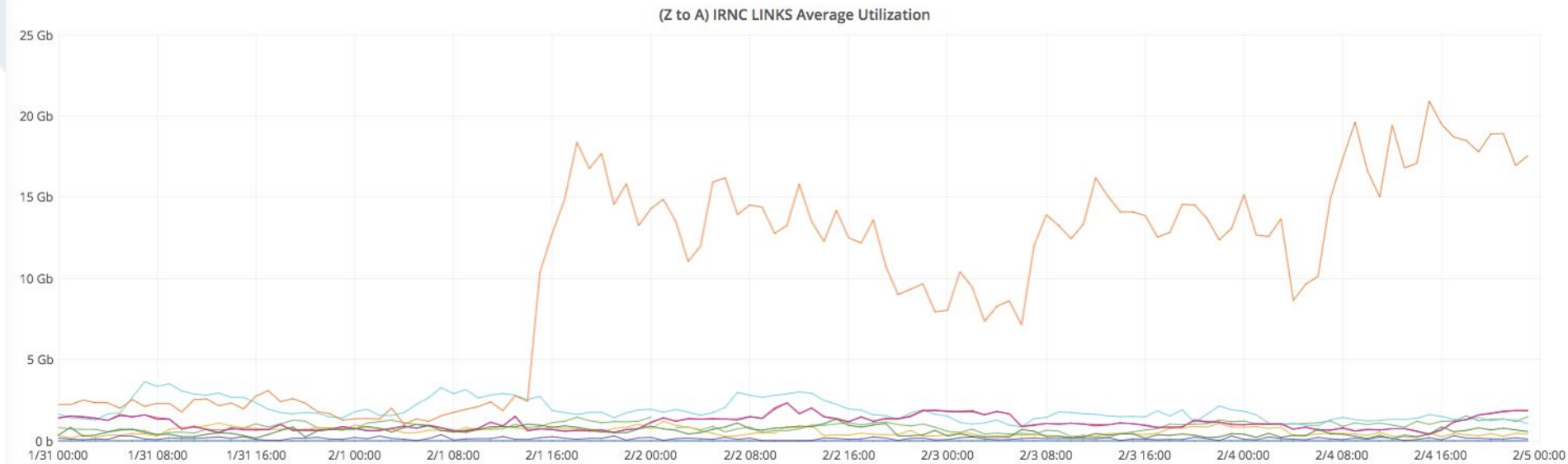
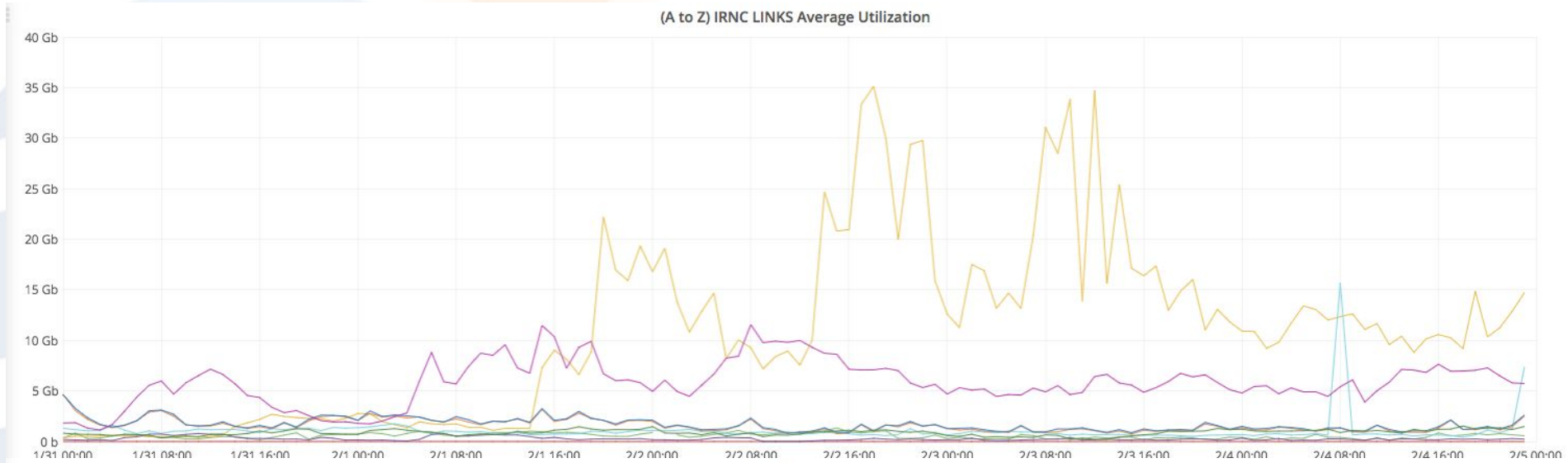
Monday Feb 5



NetSage for NEAAR link Feb 1-5



Average over all the IRNC Circuits



Flow Data



Feb 1, 2018 09:02:49 to Feb 5, 2018 09:02:49



Sensors NEAAR New York sFlow

Source Scope meta.src_organization.keyword

Destination Scope meta.dst_organization.keyword

Flow Dashboards

Top Senders

By Volume

Source	Total	Packets	Biggest	# Flows
Fermi National Accelerator Laboratory (Fermilab)	676.0 TB	444.6 Bil	18.9 GB	491,296
JANET Jisc Services Limited	296.8 TB	196.1 Bil	23.0 GB	234,376
Brookhaven National Laboratory	170.4 TB	112.3 Bil	9.6 GB	161,704
ACONET AConet Backbone	132.3 TB	91.8 Bil	7.5 GB	92,029
Massachusetts Institute of Technology	86.4 TB	56.8 Bil	7.3 GB	51,228
CERN	77.8 TB	51.6 Bil	15.4 GB	63,697
ESnet	64.9 TB	16.3 Bil	28.3 GB	12,571

By Rate

Source	Peak	Average	# Flows
ESnet	9 Gbps	2 Gbps	12,571
Queen Mary and Westfield College, University of London	7 Gbps	205 Mbps	12,264
GEANT The GEANT IP Service	6 Gbps	5 Gbps	1,283
MCNC	6 Gbps	5 Gbps	293
DFN Verein zur Foerderung eines Deutschen Forschungsnetzes e.V.	5 Gbps	733 Mbps	3,010
IN2P3 IN2P3 Autonomous System	5 Gbps	265 Mbps	6,051
Florida International University	5 Gbps	3 Gbps	95

Top Destinations

By Volume

Destination	Total ▼	Packets	Biggest	# Flows
Fermi National Accelerator Laboratory (Fermilab)	190.7 TB	130.3 Bil	9.6 GB	153,624
JANET Jisc Services Limited	171.9 TB	113.7 Bil	18.9 GB	145,632
IN2P3 IN2P3 Autonomous System	139.7 TB	91.8 Bil	13.3 GB	88,799
KIT-GRIDKA	103.2 TB	67.8 Bil	13.9 GB	56,083
SLAC National Accelerator Laboratory	96.9 TB	63.8 Bil	7.7 GB	74,994
Boston University	90.6 TB	59.8 Bil	11.3 GB	83,593
DESY-HAMBURG	86.7 TB	57.0 Bil	8.2 GB	48,132
			10.0	

By Rate

Destination	Peak ▼	Average	# Flows
Ubuntunet Alliance For Research and Education Networking	9 Gbps	5 Gbps	128
ESnet	8 Gbps	3 Gbps	6,249
MCNC	8 Gbps	7 Gbps	374
TKY-AS	8 Gbps	4 Gbps	229
GEANT The GEANT IP Service	7 Gbps	5 Gbps	1,094
JANET Jisc Services Limited	6 Gbps	84 Mbps	145,632
IN2P3 IN2P3 Autonomous System	5 Gbps	75 Mbps	88,799

Top Pairs by Total Data Transferred

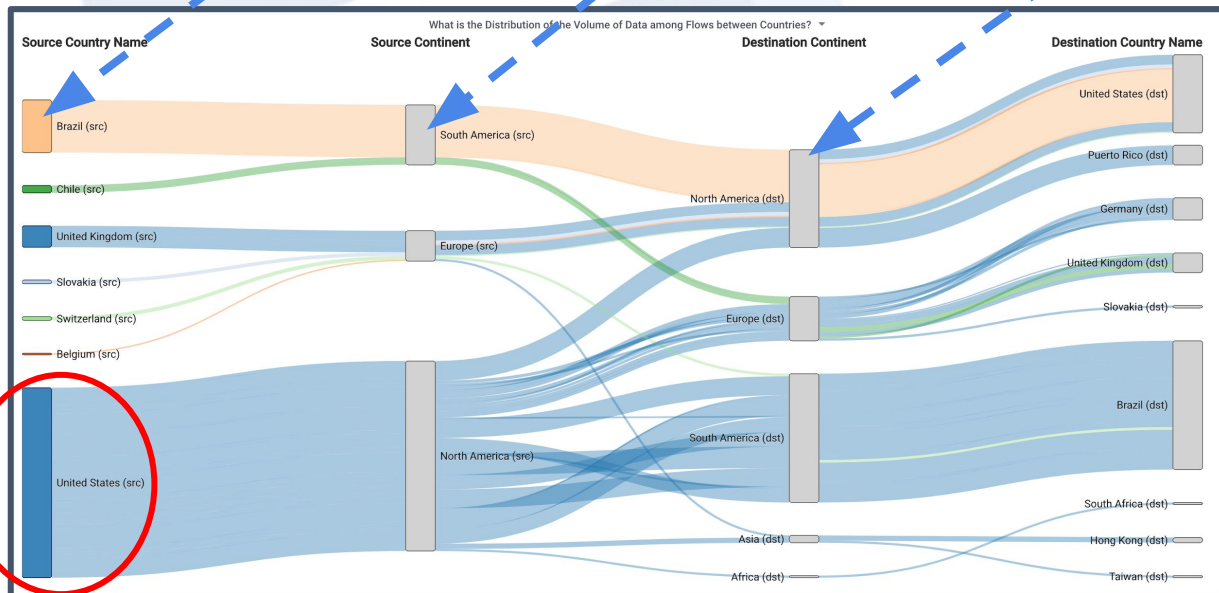
Source	Destination	Total ▾	Packets	Biggest	# Flows
Fermi National Accelerator Laboratory (Fermilab)	IN2P3 IN2P3 Autonomous System	136 TB	89.1 Bil	12.8 GB	86,618
ACONET AConet Backbone	Fermi National Accelerator Laboratory (Fermilab)	132 TB	91.8 Bil	7.5 GB	92,029
Brookhaven National Laboratory	JANET Jisc Services Limited	87 TB	57.4 Bil	9.6 GB	87,949
Fermi National Accelerator Laboratory (Fermilab)	KIT-GRIDKA	86 TB	56.8 Bil	13.9 GB	44,928
JANET Jisc Services Limited	Massachusetts Institute of Technology	70 TB	46.1 Bil	4.7 GB	43,369
JANET Jisc Services Limited	SLAC National Accelerator Laboratory	67 TB	44.4 Bil	7.0 GB	52,774
Fermi National Accelerator Laboratory (Fermilab)	NCBJ	65 TB	42.6 Bil	11.3 GB	41,405
Fermi National Accelerator Laboratory (Fermilab)	JANET Jisc Services Limited	65 TB	42.9 Bil	18.9 GB	44,042
JANET Jisc Services Limited	Boston University	62 TB	40.6 Bil	11.3 GB	57,011
Massachusetts Institute of Technology	DESY-HAMBURG	61 TB	40.0 Bil	7.3 GB	35,229

1 2 3 4 5 6 7 8 9

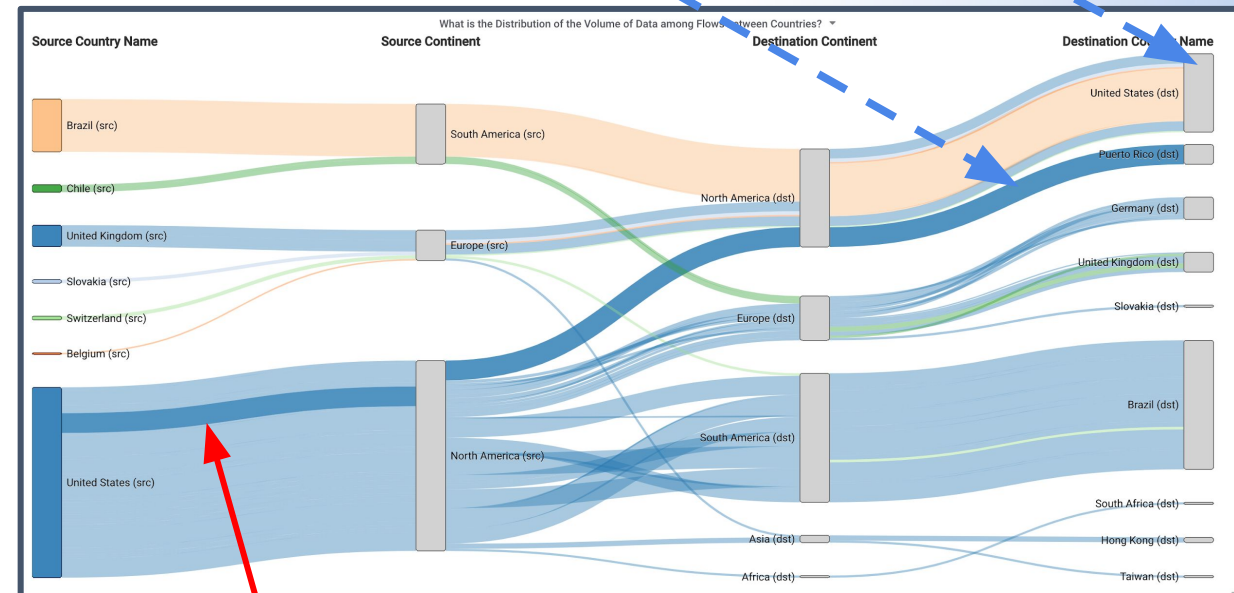
What is the Distribution of the Volume of Data among Flows between Countries?

5 dimensions of data shown in 1 chart

(Source Country, Source Continent, Destination Continent, Volume of Data, Destination Country)



At a glance we can see that United States sends a lot of data which is distributed to 5 continents, which further distributed to 8 different countries



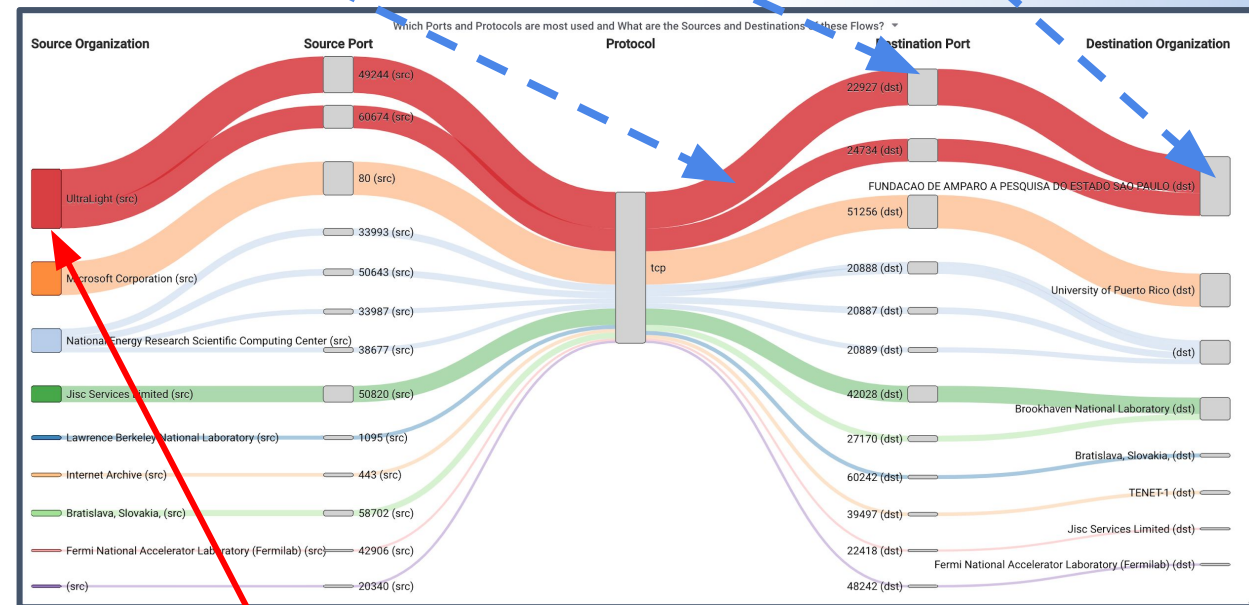
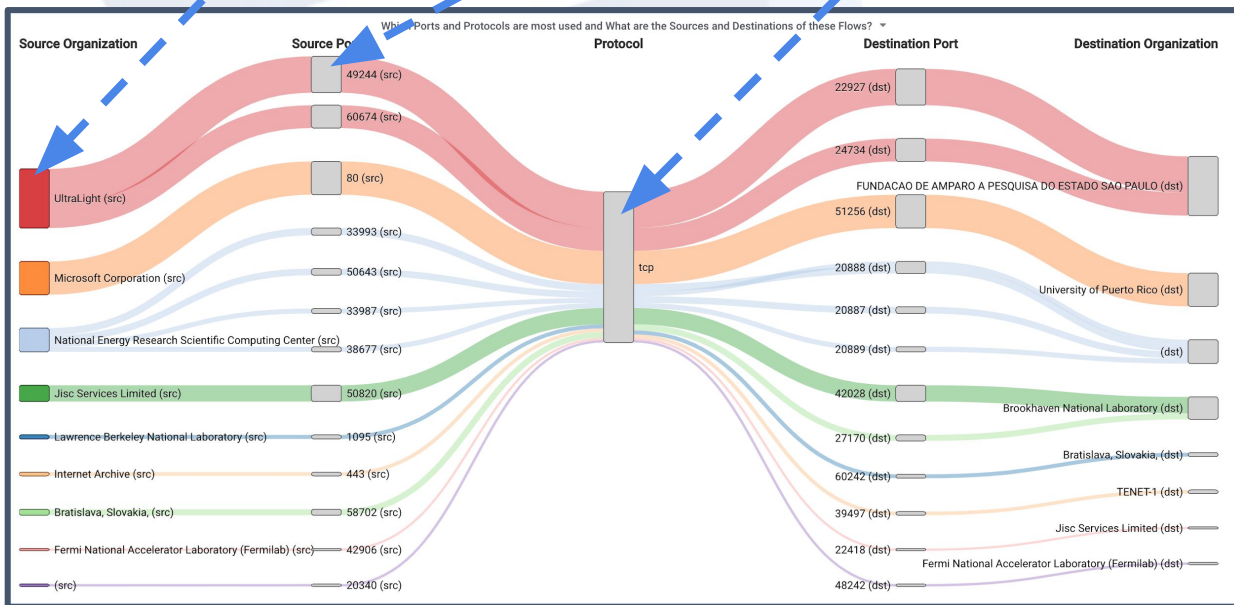
Mouseover to show distribution of flow from Source to Destination



Which Ports and Protocols are most used and What are the Sources and Destinations of these Flows?

6 dimensions of data shown in 1 chart

(Source Organization, Source Port, Protocol, Destination Port, Volume of Data, Destination Organization)



Identify which ports and protocols have the largest flow

Mouseover to highlight full flow path of data

