ESnet
ENERGY SCIENCES NETWORK

BERKELEY LAB

U.S. DEPARTMENT OF ENERGY

Minimalist Approaches to Traffic Engineering using SR

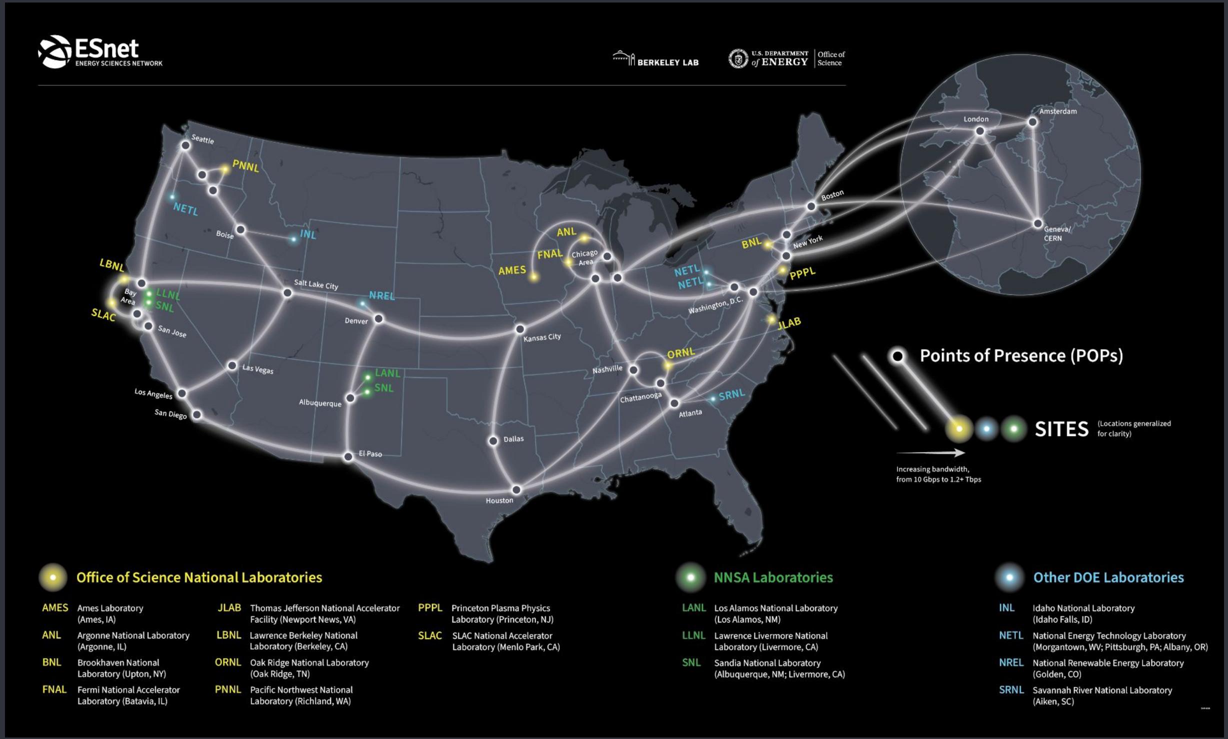
Michael Sinatra
Core Routing Network Engineer
ESnet
One of the Friendly NANOG Spokesmodels

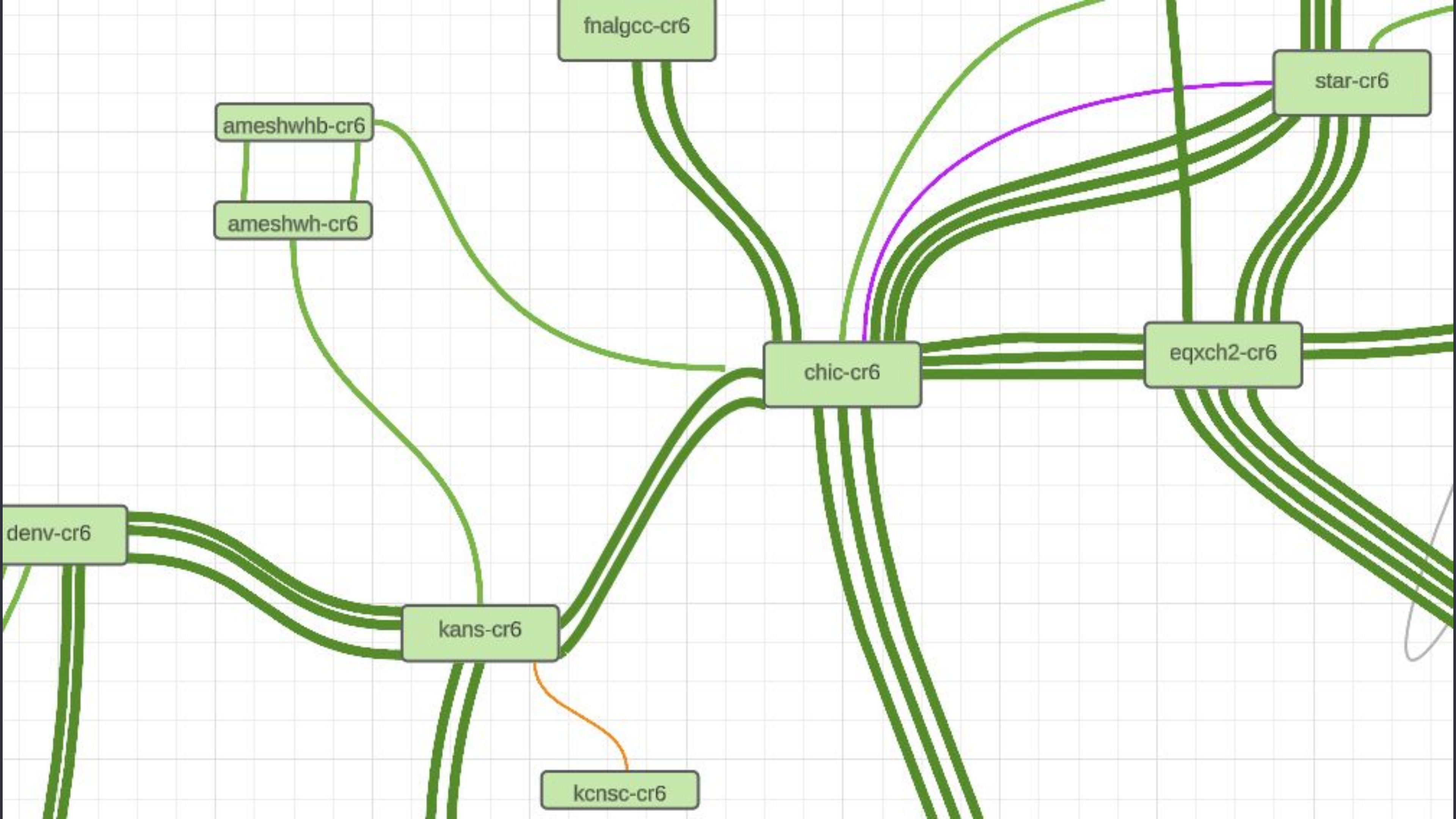
NANOG Lightning Talks
San Francisco, CA
Feb 2026

Welcome to NANOG 95
Arlington, TX
27 – 29 October 2025
NANOG
ARLINGTON, TX
N95
2025
27 – 29 OCT
NANOG

ESnet is the DOE's data circulatory system…

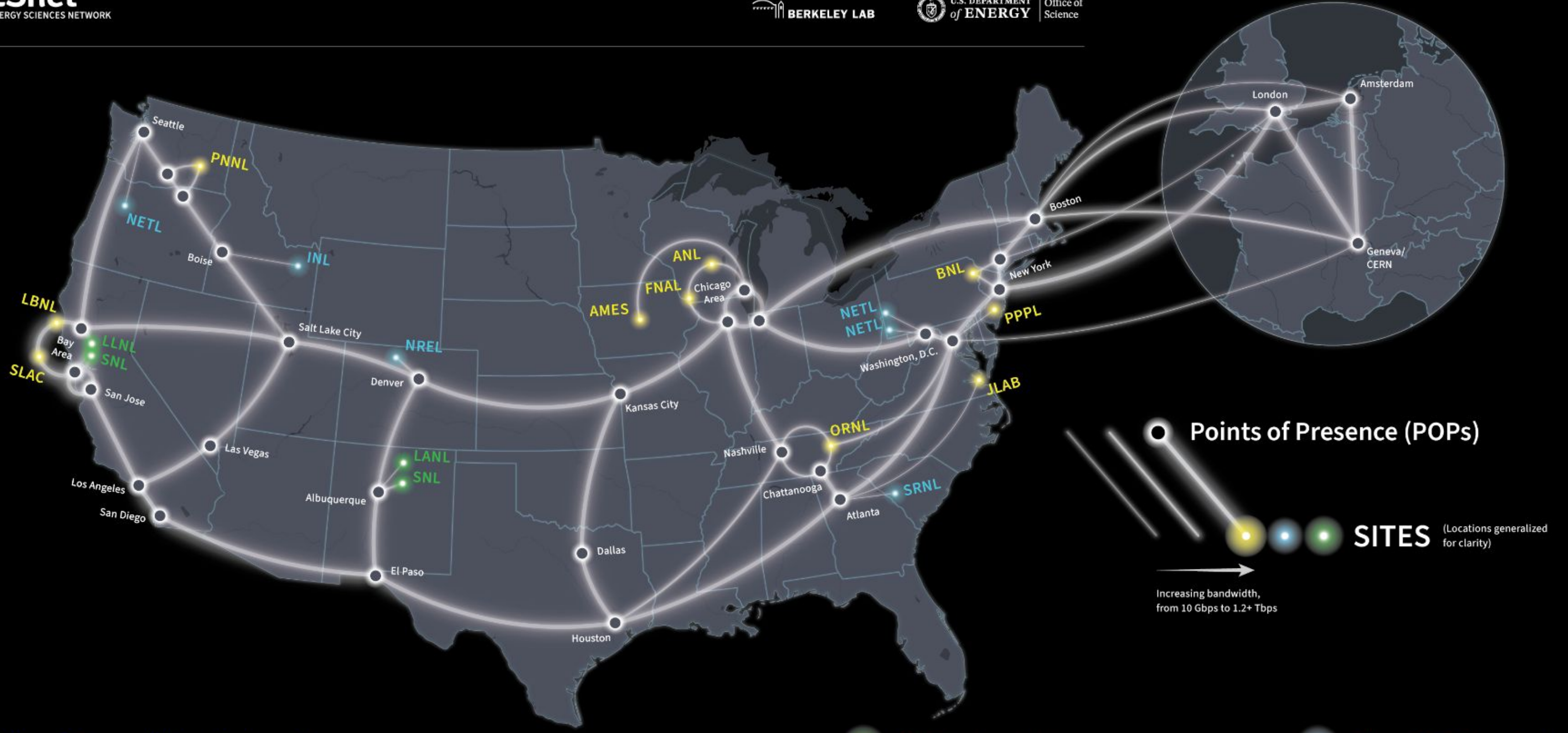
- ESnet supports the DOE scientific research ecosystem.
- Interconnects all national labs and user facilities
- Provides reliable, high-performance connectivity to global research collaborations, the Cloud, and the larger Internet.




fnalgcc-cr6
ameshwhb-cr6
ameshwh-cr6
star-cr6
chic-cr6
eqxch2-cr6
denv-cr6
kans-cr6
kcnsc-cr6

ESnet
ENERGY SCIENCES NETWORK

BERKELEY LAB | U.S. DEPARTMENT of ENERGY | Office of Science



Office of Science National Laboratories

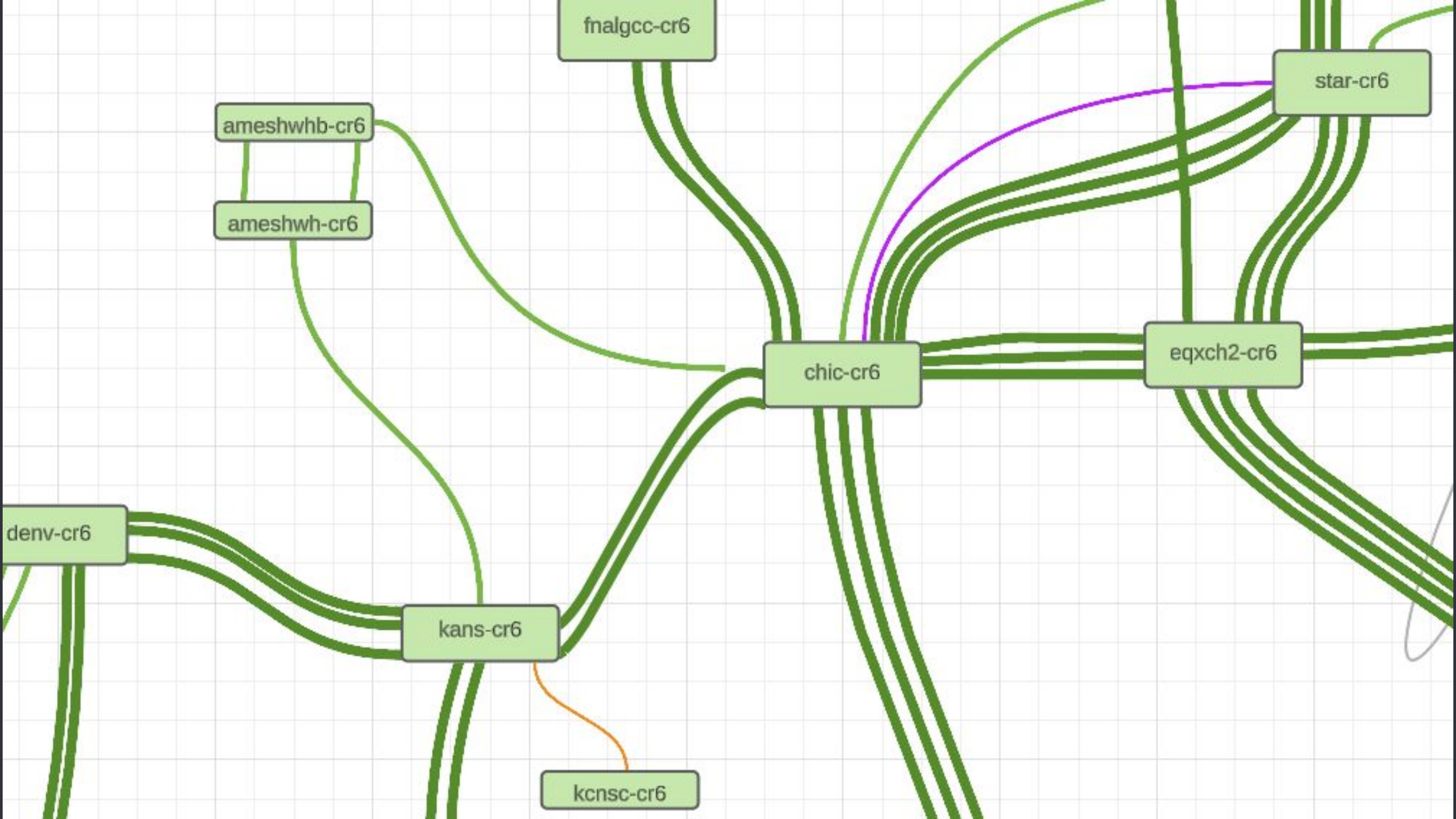
- **AMES** Ames Laboratory (Ames, IA)
- **ANL** Argonne National Laboratory (Argonne, IL)
- **BNL** Brookhaven National Laboratory (Upton, NY)
- **FNAL** Fermi National Accelerator Laboratory (Batavia, IL)
- **JLAB** Thomas Jefferson National Accelerator Facility (Newport News, VA)
- **LBNL** Lawrence Berkeley National Laboratory (Berkeley, CA)
- **ORNL** Oak Ridge National Laboratory (Oak Ridge, TN)
- **PNNL** Pacific Northwest National Laboratory (Richland, WA)
- **PPPL** Princeton Plasma Physics Laboratory (Princeton, NJ)
- **SLAC** SLAC National Accelerator Laboratory (Menlo Park, CA)

NNSA Laboratories

- **LANL** Los Alamos National Laboratory (Los Alamos, NM)
- **LLNL** Lawrence Livermore National Laboratory (Livermore, CA)
- **SNL** Sandia National Laboratory (Albuquerque, NM; Livermore, CA)

Other DOE Laboratories

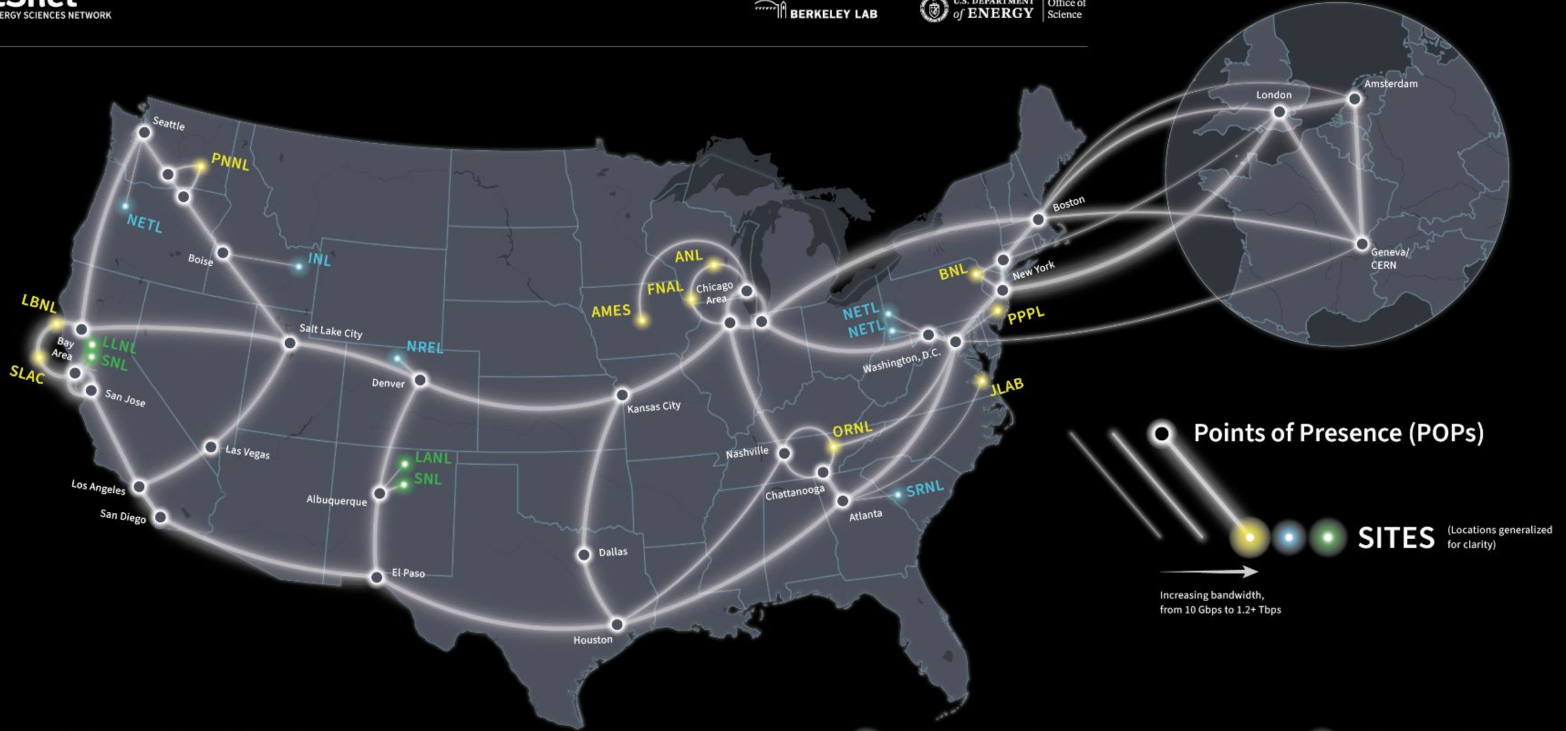
- **INL** Idaho National Laboratory (Idaho Falls, ID)
- **NETL** National Energy Technology Laboratory (Morgantown, WV; Pittsburgh, PA; Albany, OR)
- **NREL** National Renewable Energy Laboratory (Golden, CO)
- **SRNL** Savannah River National Laboratory (Aiken, SC)


fnalgcc-cr6
ameshwhb-cr6
ameshwh-cr6
star-cr6
chic-cr6
eqxch2-cr6
denv-cr6
kans-cr6
kcnsc-cr6

ESnet
ENERGY SCIENCES NETWORK

BERKELEY LAB

U.S. DEPARTMENT of ENERGY | Office of Science



Office of Science National Laboratories

- **AMES** Ames Laboratory (Ames, IA)
- **ANL** Argonne National Laboratory (Argonne, IL)
- **BNL** Brookhaven National Laboratory (Upton, NY)
- **FNAL** Fermi National Accelerator Laboratory (Batavia, IL)
- **JLAB** Thomas Jefferson National Accelerator Facility (Newport News, VA)
- **LBNL** Lawrence Berkeley National Laboratory (Berkeley, CA)
- **ORNL** Oak Ridge National Laboratory (Oak Ridge, TN)
- **PNNL** Pacific Northwest National Laboratory (Richland, WA)
- **PPPL** Princeton Plasma Physics Laboratory (Princeton, NJ)
- **SLAC** SLAC National Accelerator Laboratory (Menlo Park, CA)

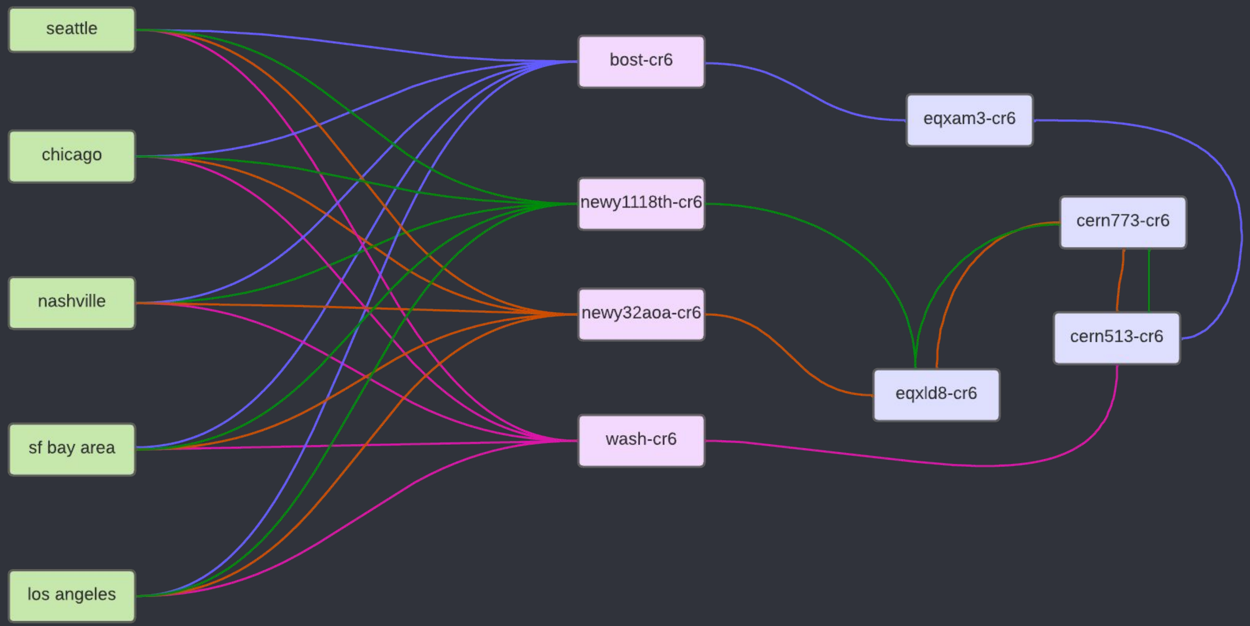
NNSA Laboratories

- **LANL** Los Alamos National Laboratory (Los Alamos, NM)
- **LLNL** Lawrence Livermore National Laboratory (Livermore, CA)
- **SNL** Sandia National Laboratory (Albuquerque, NM; Livermore, CA)

Other DOE Laboratories

- **INL** Idaho National Laboratory (Idaho Falls, ID)
- **NETL** National Energy Technology Laboratory (Morgantown, WV; Pittsburgh, PA; Albany, OR)
- **NREL** National Renewable Energy Laboratory (Golden, CO)
- **SRNL** Savannah River National Laboratory (Aiken, SC)

How does it work?



… And many more routers, where
LHCONE sites connect

Does it work?

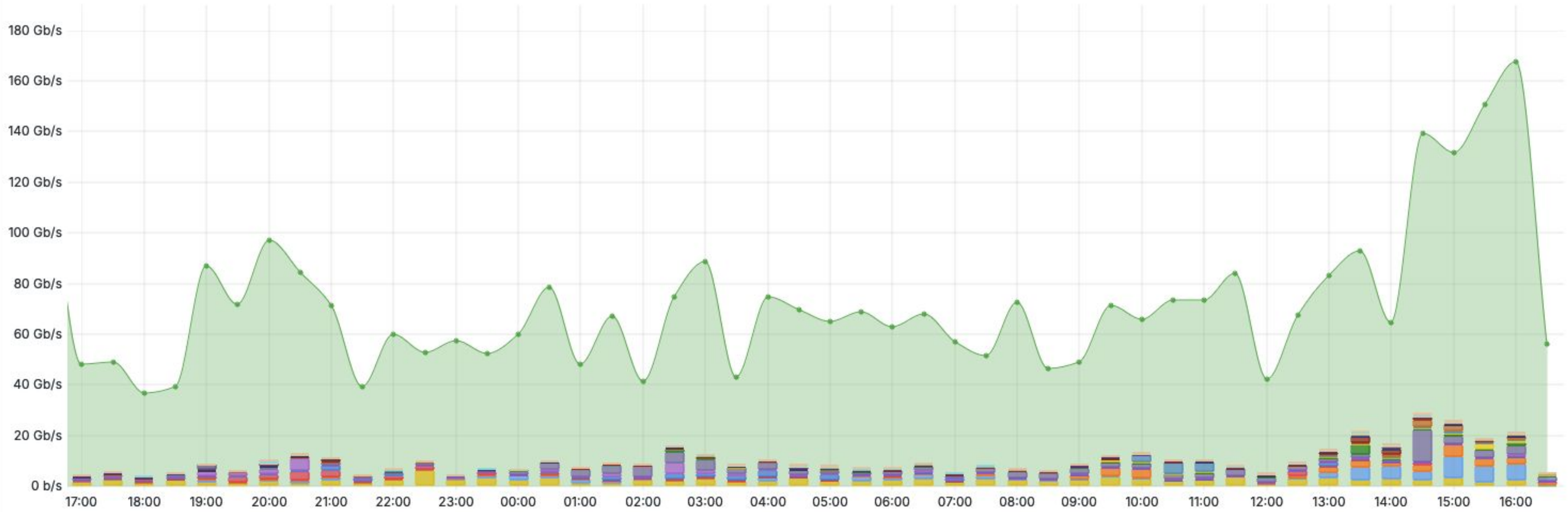
ESnet

Search or jump to... ⌘+k

Home › Dashboards › User Dashboards - NEG › TA Links - Dev2 › View panel

Share | Last 24 hours

bost-cr6 1/1/c25/1



Name	Mean	Max
bost-cr6::1/1/c25/1	71.9 Gb/s	168 Gb/s
513:CERN to 7212:VANDERBILT	2.05 Gb/s	5.74 Gb/s
0:AS0 to 0:AS0	2.01 Gb/s	12.3 Gb/s
513:CERN to 160:U-CHICAGO-AS	1.35 Gb/s	8.46 Gb/s
2200:FR-RENATER to 43:BNL-AS	670 Mb/s	4.58 Gb/s
2200:FR-RENATER to 160:U-CHICAGO-AS	536 Mb/s	3.34 Gb/s
786:JANET to 18515:UTARLINGTON	519 Mb/s	2.38 Gb/s
786:JANET to 160:U-CHICAGO-AS	507 Mb/s	4.78 Gb/s
137:ASGARR to 160:U-CHICAGO-AS	332 Mb/s	3.79 Gb/s
58069:KIT-GridKa to 3152:FNAL-AS	322 Mb/s	4.81 Gb/s

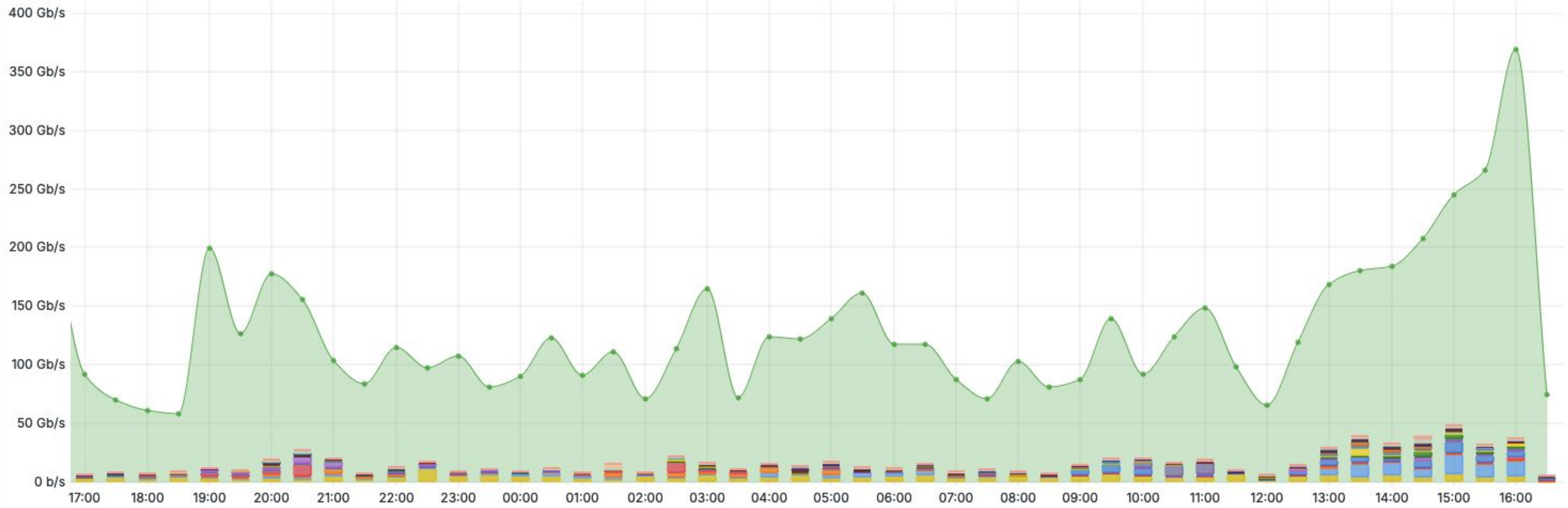
ESnet

Search or jump to... ⌘+k

Home › Dashboards › User Dashboards - NEG › TA Links - Dev2 › View panel

Share | Last 24 hours

bost-cr6 2/1/c25/1



Name	Mean	Max
bost-cr6::2/1/c25/1	128 Gb/s	369 Gb/s
513:CERN to 7212:VANDERBILT	4.10 Gb/s	10.6 Gb/s
513:CERN to 160:U-CHICAGO-AS	2.79 Gb/s	17.6 Gb/s
2200:FR-RENATER to 43:BNL-AS	1.36 Gb/s	9.70 Gb/s
786:JANET to 18515:UTARLINGTON	1.07 Gb/s	5.13 Gb/s
786:JANET to 160:U-CHICAGO-AS	1.06 Gb/s	10.6 Gb/s
2200:FR-RENATER to 160:U-CHICAGO-AS	1.05 Gb/s	6.22 Gb/s
58069:KIT-GridKa to 3152:FNAL-AS	697 Mb/s	10.3 Gb/s
137:ASGARR to 160:U-CHICAGO-AS	604 Mb/s	6.41 Gb/s
58069:KIT-GridKa to 160:U-CHICAGO-AS	519 Mb/s	4.64 Gb/s

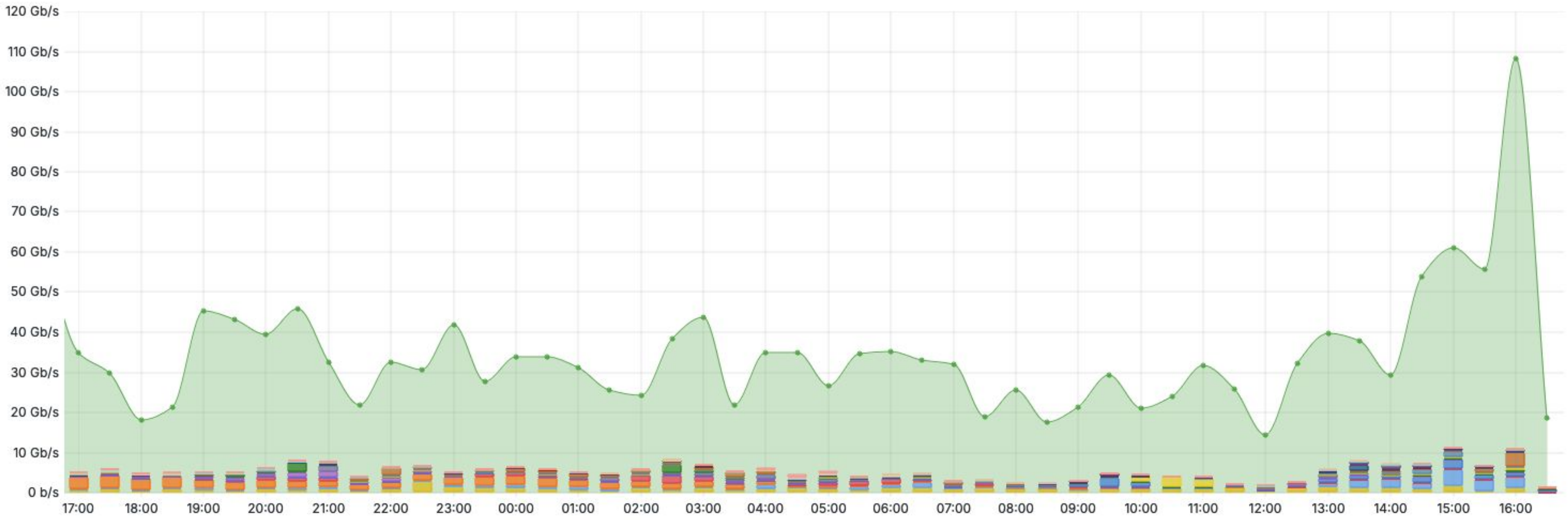
ESnet

Search or jump to... ⌘+k

Home › Dashboards › User Dashboards - NEG › TA Links - Dev2 › View panel

Share | Last 24 hours

bost-cr6 2/1/c6/1

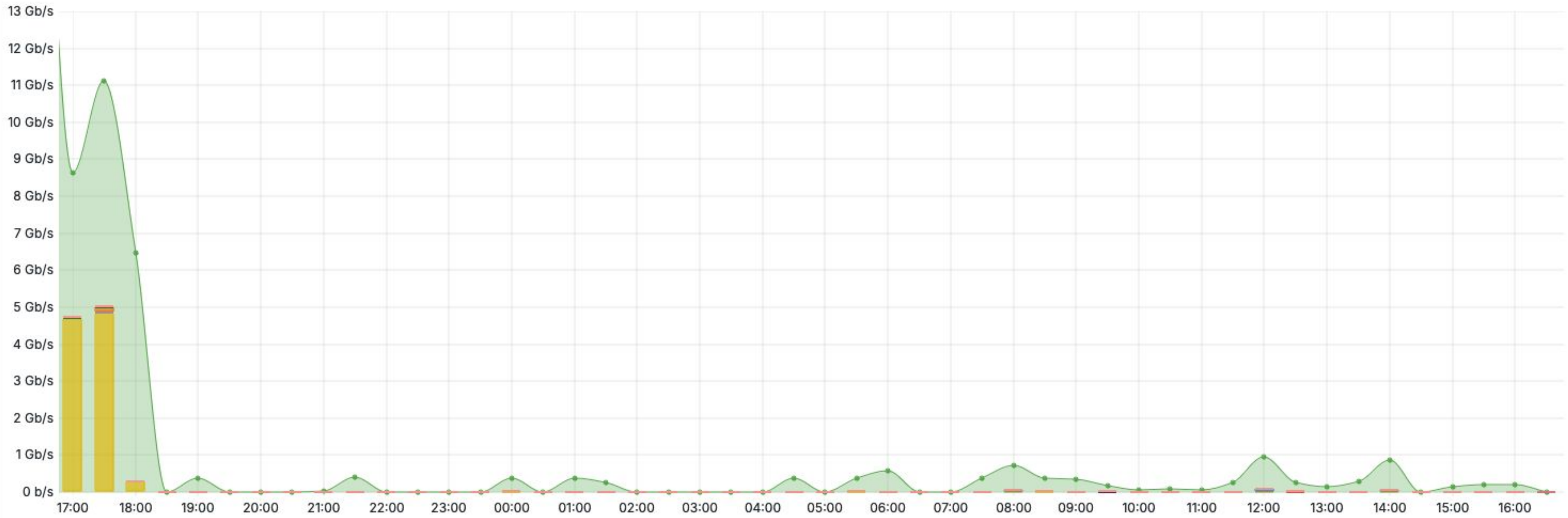


Name	Mean	Max
bost-cr6::2/1/c6/1	34.2 Gb/s	108 Gb/s
1754:DESY-HAMBURG to 7896:NU-AS	1.01 Gb/s	3.12 Gb/s
513:CERN to 7212:VANDERBILT	899 Mb/s	2.62 Gb/s
513:CERN to 160:U-CHICAGO-AS	602 Mb/s	4.13 Gb/s
0:AS0 to 0:AS0	490 Mb/s	3.42 Gb/s
1754:DESY-HAMBURG to 3:MIT-GATEWAYS	400 Mb/s	2.05 Gb/s
2200:FR-RENATER to 43:BNL-AS	297 Mb/s	2.34 Gb/s
786:JANET to 18515:UTARLINGTON	250 Mb/s	1.65 Gb/s
2200:FR-RENATER to 160:U-CHICAGO-AS	233 Mb/s	1.23 Gb/s
786:JANET to 160:U-CHICAGO-AS	228 Mb/s	2.46 Gb/s

ESnet   Search or jump to...   ⌘+k

Home › Dashboards › User Dashboards - NEG › TA Links - Dev2 › View panel   Share   Last 24 hours

newy32aoa-cr6 eqxam3-bb-a



Name	Mean	Max
newy32aoa-cr6::eqxam3-bb-a	1.11 Gb/s	19.6 Gb/s
513:CERN to 5723:JHU	220 Mb/s	4.82 Gb/s
207592:GOENET to 43:BNL-AS	8.13 Mb/s	84.6 Mb/s
291:ESNET-EAST to 291:ESNET-EAST	6.79 Mb/s	193 Mb/s
513:CERN to 2715:AS2715	1.79 Mb/s	36.4 Mb/s
291:ESNET-EAST to 10508:UARK-FAYETTEVILLE	995 kb/s	11.2 Mb/s
65292:ESNET to 65291:ESNET	730 kb/s	1.74 Mb/s
2603:NORDUNET to 3671:SLAC	659 kb/s	21.8 Mb/s
293:ESNET to 65291:ESNET	424 kb/s	939 kb/s
2200:FR-RENATER to 43:BNL-AS	189 kb/s	9.25 Mb/s

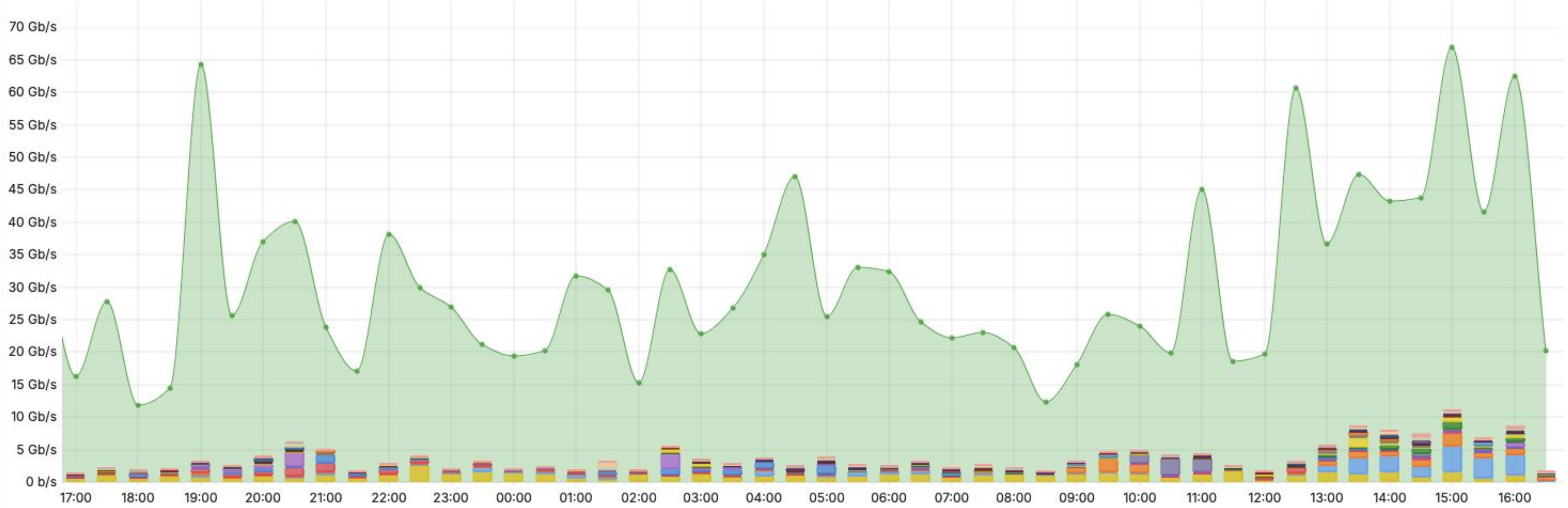
ESnet

Search or jump to... ⌘+k

Home › Dashboards › User Dashboards - NEG › TA Links - Dev2 › View panel

Share | Last 24 hours

newy1118th-cr6 1/1/c5/1



Name	Mean	Max
newy1118th-cr6::1/1/c5/1	30.6 Gb/s	67.0 Gb/s
513:CERN to 7212:VANDERBILT	954 Mb/s	2.47 Gb/s
513:CERN to 160:U-CHICAGO-AS	594 Mb/s	3.86 Gb/s
2200:FR-RENATER to 43:BNL-AS	285 Mb/s	2.11 Gb/s
786:JANET to 18515:UTARLINGTON	248 Mb/s	1.35 Gb/s
786:JANET to 160:U-CHICAGO-AS	245 Mb/s	2.47 Gb/s
2200:FR-RENATER to 160:U-CHICAGO-AS	232 Mb/s	1.35 Gb/s
58069:KIT-GridKa to 3152:FNAL-AS	165 Mb/s	2.73 Gb/s
137:ASGARR to 160:U-CHICAGO-AS	153 Mb/s	1.48 Gb/s
58069:KIT-GridKa to 160:U-CHICAGO-AS	108 Mb/s	1.22 Gb/s

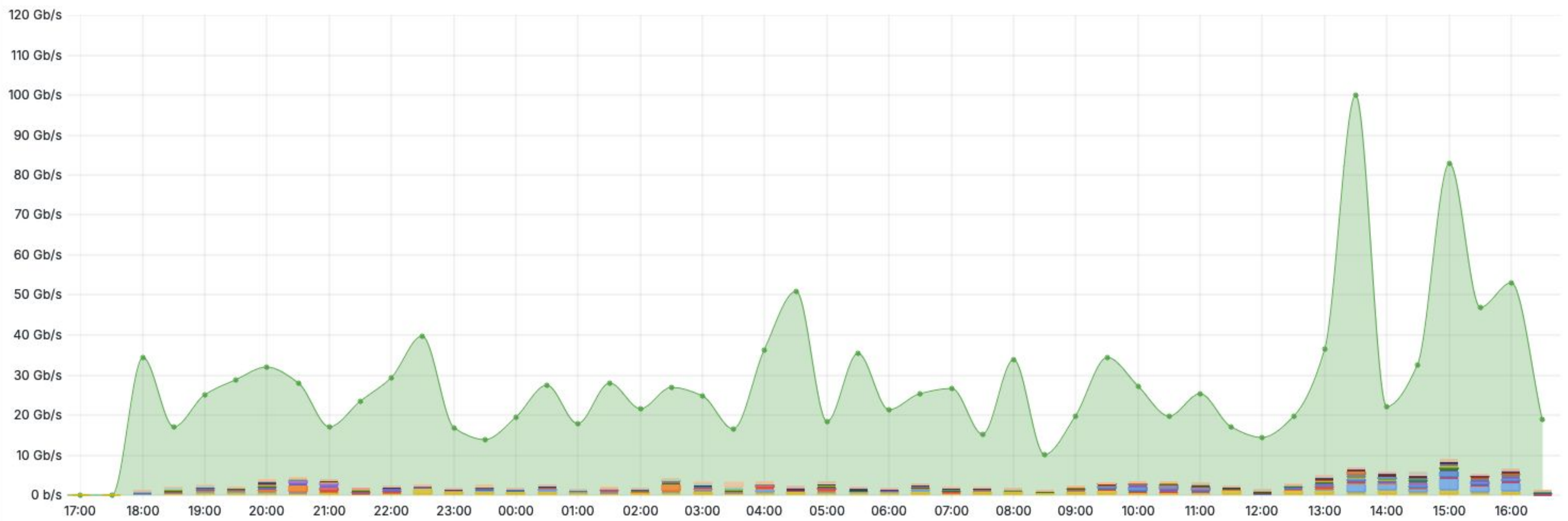
ESnet

Search or jump to... ⌘+k

Home › Dashboards › User Dashboards - NEG › TA Links - Dev2 › View panel

Share | Last 24 hours

wash-cr6 1/1/c5/1



Name	Mean	Max
wash-cr6::1/1/c5/1	27.2 Gb/s	100 Gb/s
513:CERN to 7212:VANDERBILT	656 Mb/s	1.70 Gb/s
513:CERN to 160:U-CHICAGO-AS	427 Mb/s	3.12 Gb/s
2200:FR-RENATER to 43:BNL-AS	229 Mb/s	1.66 Gb/s
786:JANET to 160:U-CHICAGO-AS	181 Mb/s	1.86 Gb/s
2200:FR-RENATER to 160:U-CHICAGO-AS	175 Mb/s	1.07 Gb/s
786:JANET to 18515:UTARLINGTON	174 Mb/s	1.03 Gb/s
680:DFN to 44:S1-DOMAIN	142 Mb/s	432 Mb/s
137:ASGARR to 160:U-CHICAGO-AS	107 Mb/s	1.08 Gb/s
58069:KIT-GridKa to 3152:FNAL-AS	103 Mb/s	1.61 Gb/s

A1 fx Vanderbilt --> CERN path

	A	B	C	D	E
1	**Vanderbilt --> CERN path**	**avg traffic (gb/s)**	**weight**	**traffic %age**	**weight %age**
2	1 - AM3 -- BOST	0.899	4	8.394808105	8.510638298
3	2 - LD8 -- NEWY1118th	0.954	4	8.908394808	8.510638298
4	3 - LD8 -- BOST	4.1	18	38.28555421	38.29787234
5	4 - CERN -- WASH	0.656	3	6.125688673	6.382978723
6	5 - LD8 -- NEWY32AOA	2.05	9	19.1427771	19.14893617
7	6 - CERN -- BOST	2.05	9	19.1427771	19.14893617
8					
9	Totals	10.709	47	100	100
10					

```
nash-ps:~> traceroute perfsonar-bw.cern.ch
traceroute to perfsonar-bw.cern.ch (128.142.208.132), 30 hops max, 60 byte packets
 1  nashcr6-nash-ps-tp.es.net (198.125.56.28)  0.218 ms  0.230 ms  0.220 ms
 2  chat-cr6--atla-bb-b.igp.es.net (134.55.57.231)  109.186 ms chic-cr6--eqxch2-bb-c.igp.es.net (134.55.5
6.121)  130.659 ms chat-cr6--atla-bb-b.igp.es.net (134.55.57.231)  116.266 ms
 3  atla-cr6--wash-bb-h.igp.es.net (134.55.56.90)  112.928 ms  108.929 ms eqxch2-cr6--bost-bb-b.igp.es.ne
t (134.55.58.27)  121.097 ms
 4  bost-cr6--eqxam3-bb-a.igp.es.net (134.55.57.69)  113.804 ms  113.801 ms wash-cr6--newy32aoa-bb-f.igp.
es.net (134.55.56.215)  116.283 ms
 5  newy32aoa-cr6--eqxld8-bb-b.igp.es.net (134.55.56.246)  109.188 ms eqxam3-cr6--cern513-bb-d.igp.es.net
 (134.55.57.118)  114.602 ms newy32aoa-cr6--eqxld8-bb-b.igp.es.net (134.55.56.246)  109.164 ms
 6  eqxld8-cr6--cern773-bb-c.igp.es.net (134.55.56.178)  125.305 ms  125.287 ms  125.279 ms
 7  cixp-esnet-cern513cr6.cern.ch (192.65.184.214)  132.339 ms cixp-esnet-cern773cr6.cern.ch (192.65.184.
194)  130.702 ms  130.701 ms
 8  e773-e-rjup1-2-te5.cern.ch (192.65.184.193)  108.542 ms cixp-esnet-cern513cr6.cern.ch (192.65.184.214
)  130.613 ms e513-e-rjuxm-v10-pe4.cern.ch (192.65.184.81)  115.267 ms
 9  e513-e-rjup1-1-ne0.cern.ch (192.65.184.189)  115.948 ms e513-e-rjup1-1-te22.cern.ch (192.65.184.213)
 113.806 ms g513-e-rjuxm-10-fi1.cern.ch (192.65.196.149)  125.001 ms
10  e513-e-rjuxm-v10-pe4.cern.ch (192.65.184.81)  109.807 ms g513-e-fpa78-1-fe1.cern.ch (192.65.184.61)
121.317 ms e513-e-rjuxm-v10-pe4.cern.ch (192.65.184.81)  109.817 ms
11  g513-e-fpa78-1-fe1.cern.ch (192.65.184.61)  113.916 ms b513-b-rjuxl-1-pg4.cern.ch (192.65.196.65)  11
8.507 ms l513-b-rjupl-3-ht3.cern.ch (192.65.196.226)  124.452 ms
12  l513-b-rjupl-3-ht3.cern.ch (192.65.196.226)  131.447 ms g513-e-rjuxm-10-fi1.cern.ch (192.65.196.149)
 131.833 ms l513-b-rjupl-1-ht1.cern.ch (192.65.196.202)  115.631 ms
13  b513-b-rjuxl-1-pg4.cern.ch (192.65.196.65)  118.362 ms l513-b-rjuxl-2-ht2.cern.ch (192.65.196.210)  1
19.219 ms b513-b-rjuxl-1-pg4.cern.ch (192.65.196.65)  118.378 ms
14  l513-b-rjupl-3-ht3.cern.ch (192.65.196.226)  131.508 ms 185.249.56.70 (185.249.56.70)  119.893 ms psb
01-gva.cern.ch (128.142.208.132)  112.899 ms
```

Other lessons learned

What else did I learn?

- You can incorporate as many or as few components of SR and still do useful things. I.e. **you can run SR-TE without a controller.**
- SR fits really well in situations where TE is only needed in one part of the network.
- There's even an advantage to doing "SR-Shortest Path" even if you don't need TE!
- Moving to SR from traditional MPLS is dirt-simple (there can be gotchas, but if you can "turn it on" in your entire network, you can reap benefits quickly and easily).

…a few principles…

- TE is something like a band-aid, or more like stitches
  - These are very useful tools!
  - You don’t need a body-cast if you only have a cut!
- TE can (should?) be used sparingly, where needed
- Shortest-path routing is your friend 90-99% of the time…
- Do you even need TE?
  - “Throwing bandwidth at the problem” *may still work*, in some circumstances
  - Having a good algorithm for setting IGP costs, and using automation (or using built-in algorithms) may also work

...a few principles...

In the early 2000, Thomas Telkamp was managing the worldwide GLOBAL CROSSING backbone from Amsterdam, the Netherlands. This was one of the first RSVP-TE deployment and likely the biggest at that time. I had the chance to work directly with him and learned the three following concepts through the experience.

1. the always-on RSVP-TE full-mesh model is way too complex because it creates continuous pain for no gain as, in most of the cases, the network is just fine routing along the IGP shortest-path. A tactical TE approach is more appealing. Remember the analogy of the raincoat. One should only need to wear the raincoat when it actually rains.
2. ECMP is key to IP. A traffic engineering approach for IP should natively support ECMP
3. for real networks, routing convergence has more impact on SLA than bandwidth optimization

- Filsfils, et al, *Segment Routing, Part I*

...a few principles...

- "...use IGP shortest path routing whenever possible and then deploy some SR-TE policies for traffic that needs to be forwarded via a different path."
  - Dmytro Shypovalov, Eircom, APNIC Blog
- I'd like to create a new saying to go along with nonsensical ones like "cattle not pets" and "inside baseball" (I am a lifelong MLB fan, and I still don't get "inside baseball," BTW), and this *specifically refers to traffic engineering*: "**band-aids not body-casts.**"

- You're welcome.

ESnet

...a few principles...

- “traffic engineering” != “you have to use RSVP”
- if you need bandwidth guarantees, use QoS (ESnet has been doing this for 16+ years)
- “SR-TE” != “you have to use a controller and all of the SR bells & whistles”
- “**band-aids not body-casts.**”

ESnet

End of Spoken Presentation (the rest you get to read yourself!)

Resources

- (Refresher) SR tutorial (geared toward R&E but should be applicable): https://www.youtube.com/watch?v=QW7_vq1MWOY

- A much longer version of this presentation, which includes config! https://www.youtube.com/watch?v=mwVlPIu9OEs

ESnet