ESnet Update: A somewhat random collection of (hopefully) interesting topics.

Brian L Tierney

ESnet, Lawrence Berkeley National Laboratory

ScienceDMZ and other ESnet Updates

Science DMZ Summary

Consists of **three key components**, all required:

"Friction free" network path

- Highly capable network devices (wire-speed, deep queues)
- Virtual circuit connectivity option
- Security policy and enforcement specific to science workflows
- Located at or near site perimeter if possible

Dedicated, high-performance Data Transfer Nodes (DTNs)

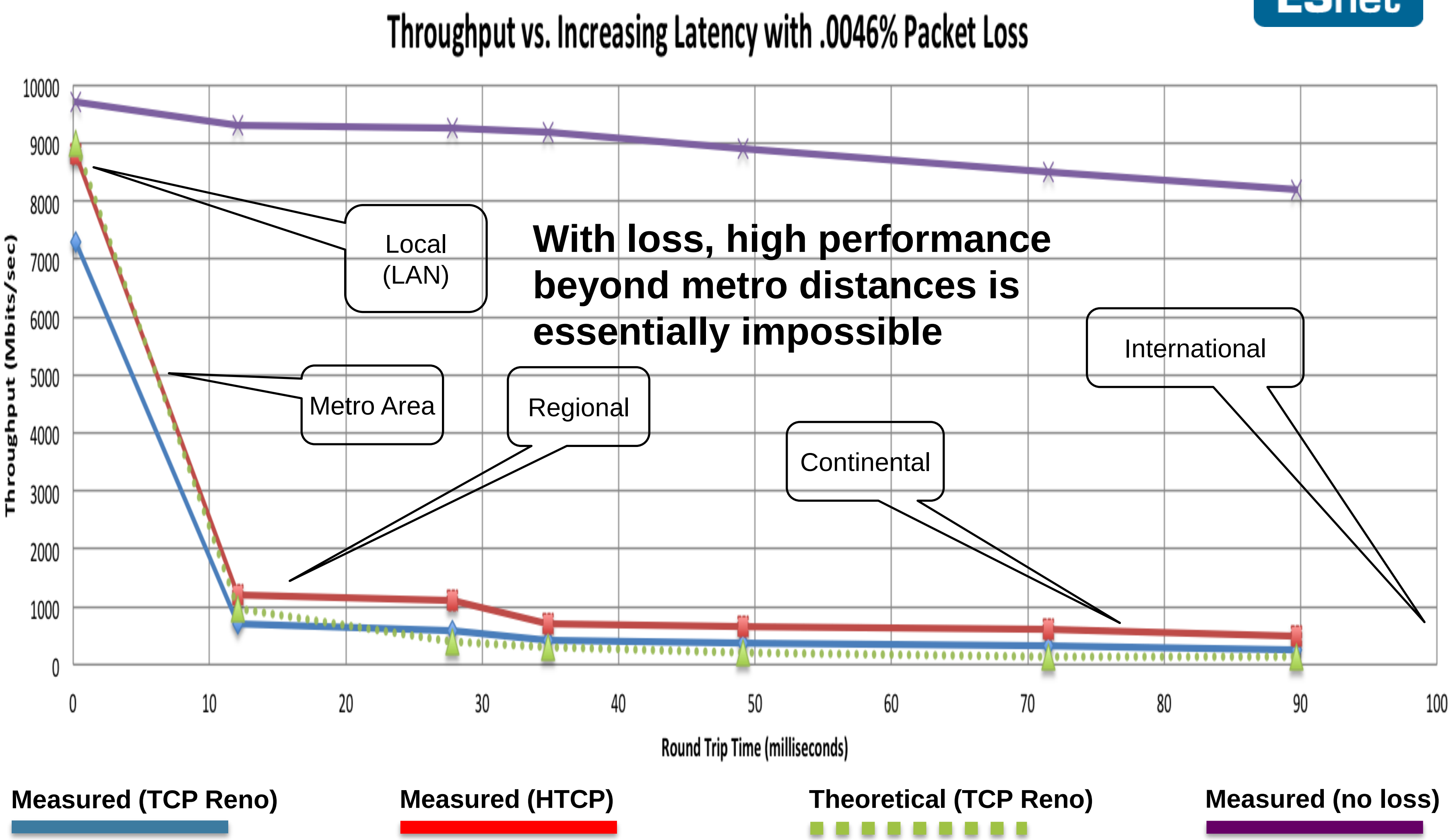
- Hardware, operating system, libraries all optimized for transfer
- Includes optimized data transfer tools such as Globus Online and GridFTP

Performance measurement/test node

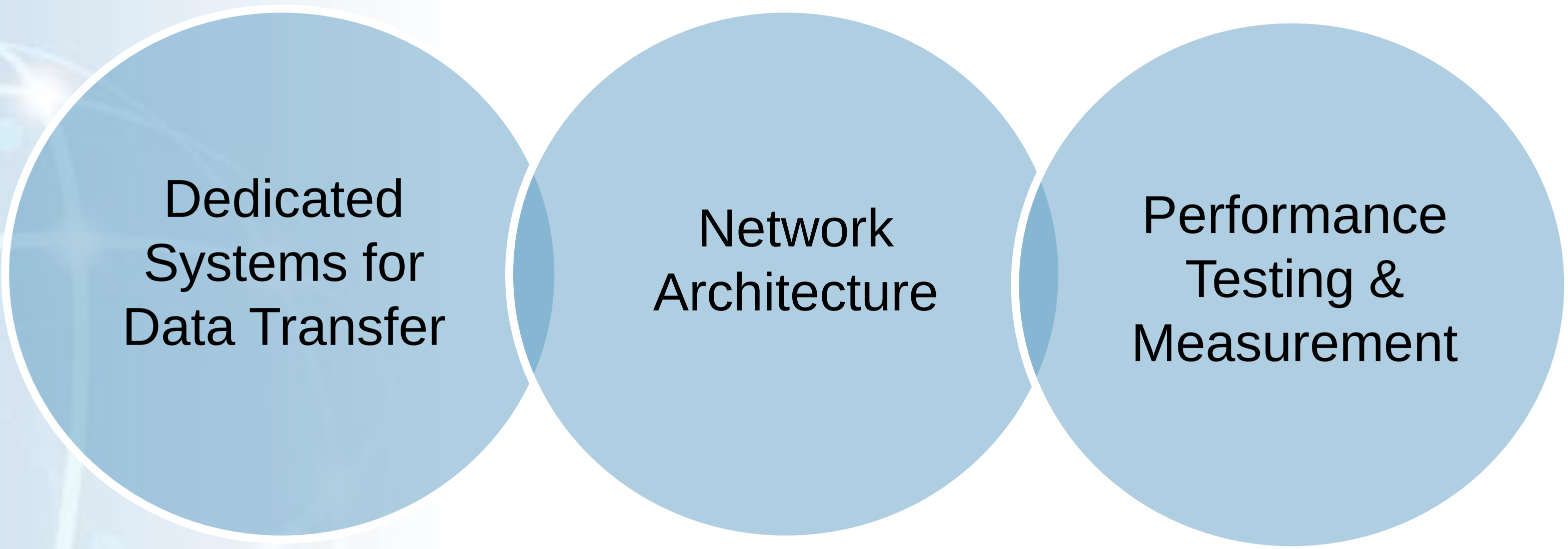
- perfSONAR

Details at http://fasterdata.es.net/science-dmz/

A small amount of packet loss makes a huge difference in TCP performance



The Data Transfer Trifecta: The Science DMZ Model



Data Transfer Node
- High performance
- Configured for data transfer
- Proper tools, such as Globus Online

Science DMZ
- Dedicated location for DTN
- Proper security
- Easy to deploy - no need to redesign the whole network
- Additional info: http://fasterdata.es.net/

perfSONAR
- Enables fault isolation
- Verify correct operation
- Widely deployed in ESnet and other networks, as well as sites and facilities

ESnet Outreach Program

Education/Consultation

- Science DMZ Architecture
- perfSONAR
- Network performance troubleshooting
- Data Transfer Nodes
    - data transfer tools

- Tutorials
- Fasterdata.es.net

- Contact: engage@es.net

my.es.net example: BNL

Brookhaven National Laboratory

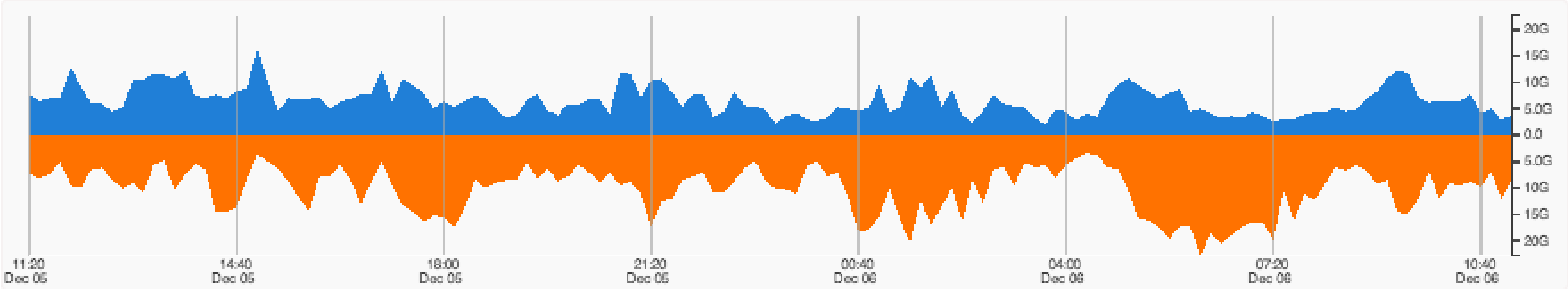
Summary | Flow details | Interfaces | Outages

Website http://www.bnl.gov/

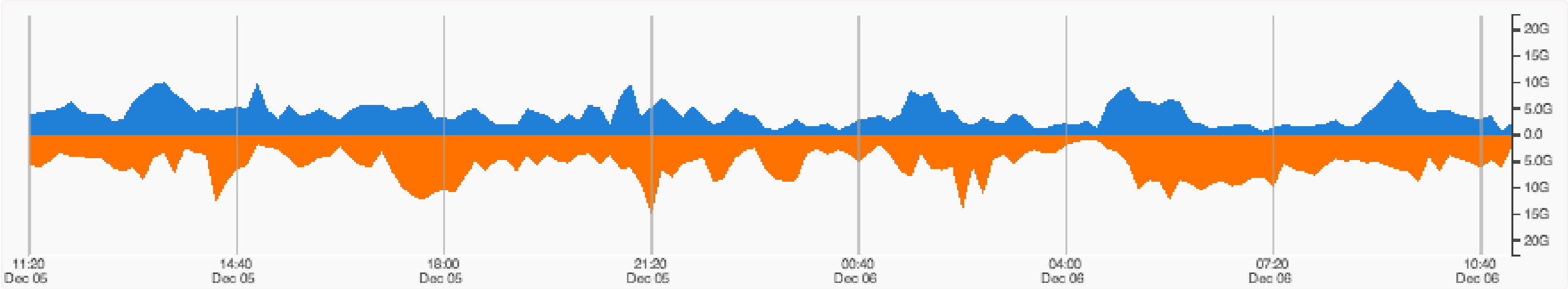
From : Thu Dec 5 11:20:16 2013 To : Fri Dec 6 11:20:16 2013

To site | From site

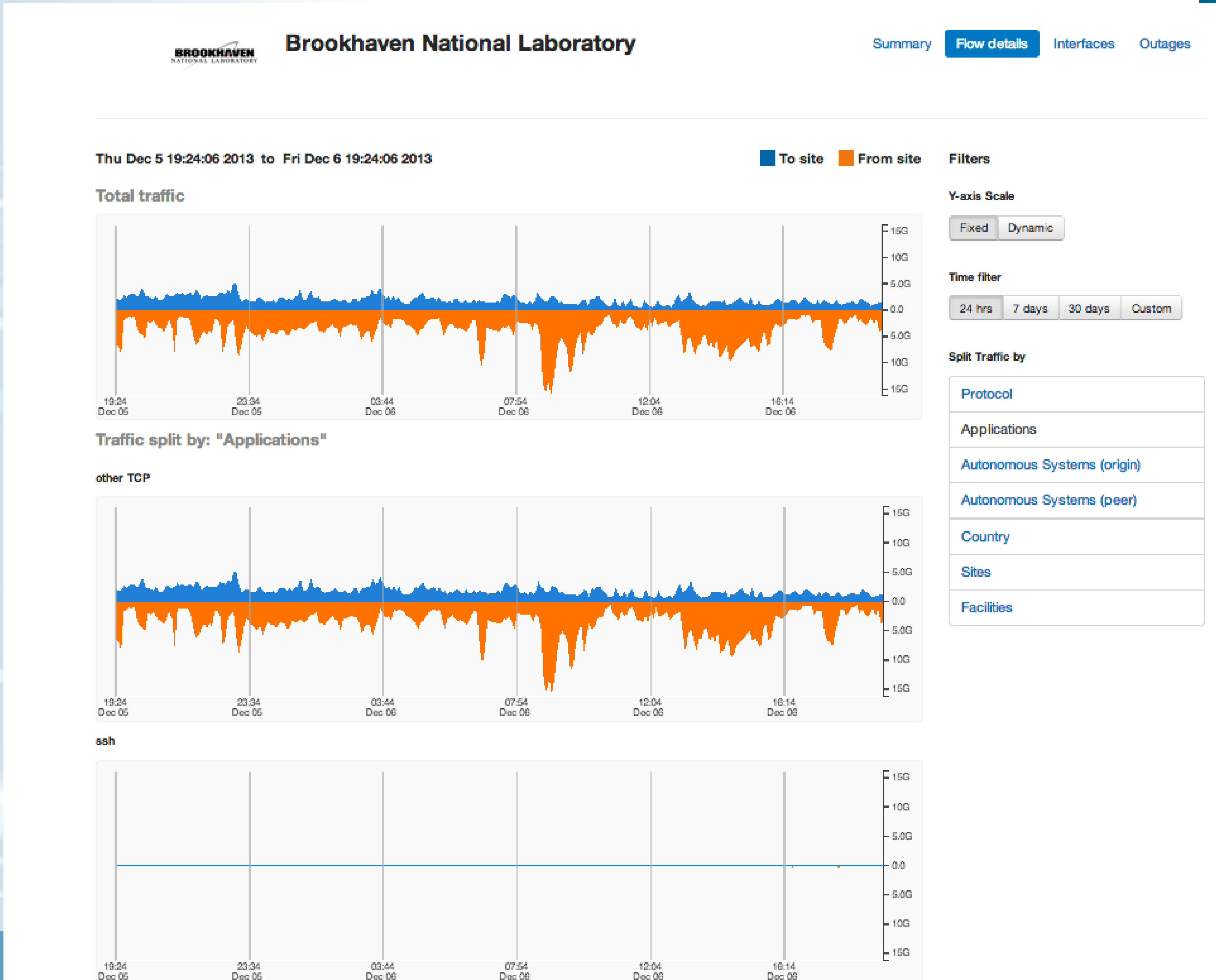
Total traffic



Total OSCARS traffic



my.es.net example: BNL



US LHC 100G sites (currently or very soon)

Tier-1 Sites:

- BNL
- FNAL

Tier-2 Sites:

- University of Chicago
- Indiana University
- Boston University
- Harvard University
- University of Nebraska
- University of Michigan
- University of Illinois
- Caltech
- Univ Florida
- UCSD
- others?

bwctl + iperf3 example

```
bwctl -T iperf3 -c nettest.lbl.gov -i1 -v
[ ID] Interval           Transfer     Bandwidth       Retransmits
[ 16]   0.00-1.00   sec   109 MBytes   912 Mbits/sec  129
[ 16]   1.00-2.00   sec   109 MBytes   912 Mbits/sec  0
[ 16]   2.00-3.00   sec  98.8 MBytes   828 Mbits/sec  0
[ 16]   3.00-4.00   sec  93.8 MBytes   786 Mbits/sec  0
[ 16]   4.00-5.00   sec  88.8 MBytes   744 Mbits/sec  24
[ 16]   5.00-6.00   sec  85.0 MBytes   712 Mbits/sec  0
Test Complete. Summary Results:
Sent:  [ 16]   0.00-10.00  sec   950 MBytes   797 Mbits/sec  161
Recvd: [ 16]   0.00-10.00  sec   955 MBytes   801 Mbits/sec
Host CPU Utilization:   98.0%
Remote CPU Utilization: 71.5%
iperf Done.
```

perfSONAR Update

World-Wide perfSONAR-PS Deployments:
695 bwctl, 734 owamp **registered** nodes as of Dec '13

perfSONAR Update

perfSONAR-PS 3.3.2 rc1 came out this week

Includes a few bug fixes, and some security enhancements

- iptables now on by default
- fail2ban host IDS
- updated versions of bwctl, iperf3, and nuttcp
- http://psps.perfsonar.net/toolkit/releasenotes/pspt-3_3_2rc3.html

Increase in perfSONAR developers

Internet2 has re-added the 1FTE that went away for a while

Indiana GR-NOC will be providing 1.5FTE starting next year

perfSONAR-PS and perfSONAR MDM will likely be combined

- will be just called ‘perfSONAR’ in the future
- more developers in Europe
  - exact role TBD

perfSONAR Roadmap

https://code.google.com/p/perfsonar-ps/wiki/RoadMap

1. perfSONAR node Cost Reduction: Support both latency testing and throughput testing on the same host
2. Extensibility and Ease of Use: Adding REST APIs for all components will make it much easier for others to extend perfSONAR.
3. Documentation overhaul
4. Additional Troubleshooting Capabilities: e,g.: ability to collect and store TCP retransmit information. Better GUIs are needed as well, but we still need to find a good GUI developer.
5. Enhanced NOC support: details TBD based on discussions with various NOCs.
6. Enhanced Release Management: utilize automated build/test systems such as OSG uses

Release Roadmap

- 3.3.2: Bug Fix release, December 2013
- 3.4: Next major release that includes deliverable #1 and part of #4: March 2014
- 3.4.1 Bug fix release: May 2014
- 3.5 Next major release that includes deliverable #2, #5, July 2014
- 3.6: Next major release that includes deliverables #4, #6, Fall 2014

Iperf3: https://code.google.com/p/iperf/

iperf3 is a new implementation from scratch, with the goal of a smaller, simpler code base, and a library version of the functionality that can be used in other programs.

Some new features in iperf3 include:

- reports the number of TCP packets that were retransmitted
- reports the average CPU utilization of the client and server (-V flag)
- support for zero copy TCP (-Z flag)
- JSON output format (-J flag)
- “omit” flag: ignore the first N seconds in the results

More at: http://fasterdata.es.net/performance-testing/network-troubleshooting-tools/iperf-and-iperf3/

perfSONAR Dashboard

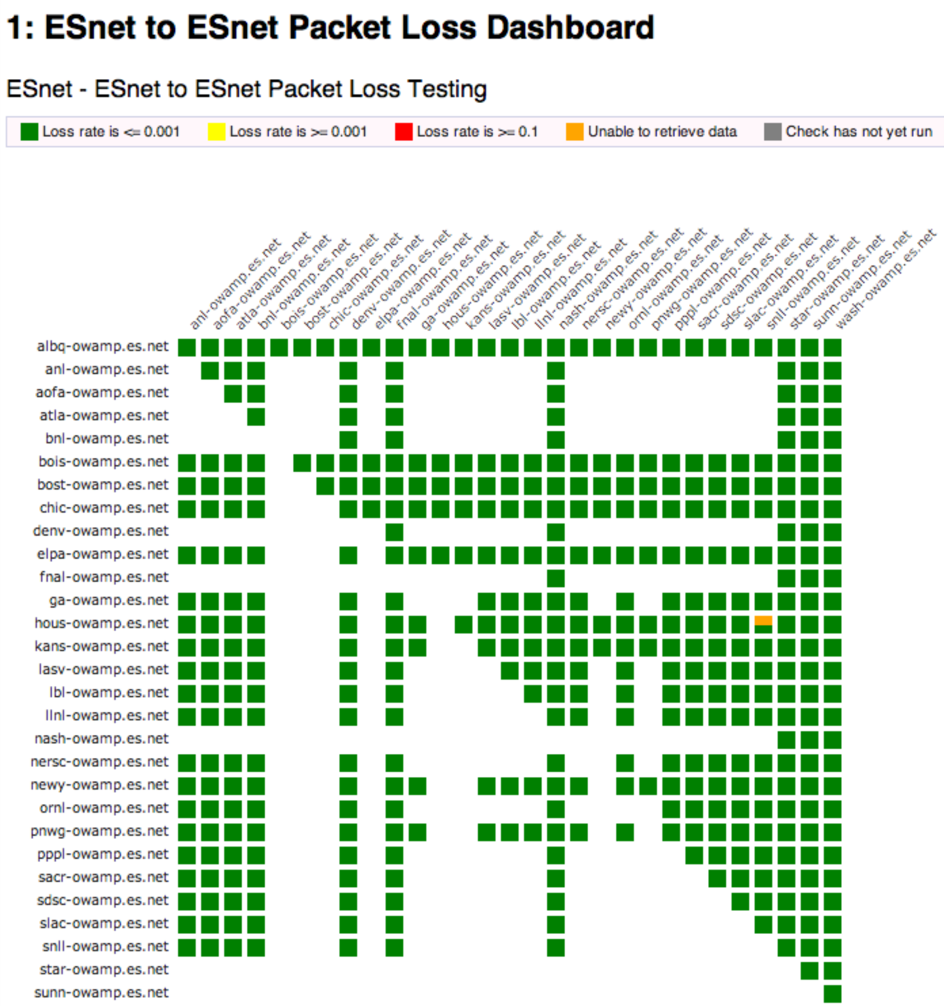
Status at-a-glance

- Packet loss
- Throughput

Current live instance at http://ps-dashboard.es.net/

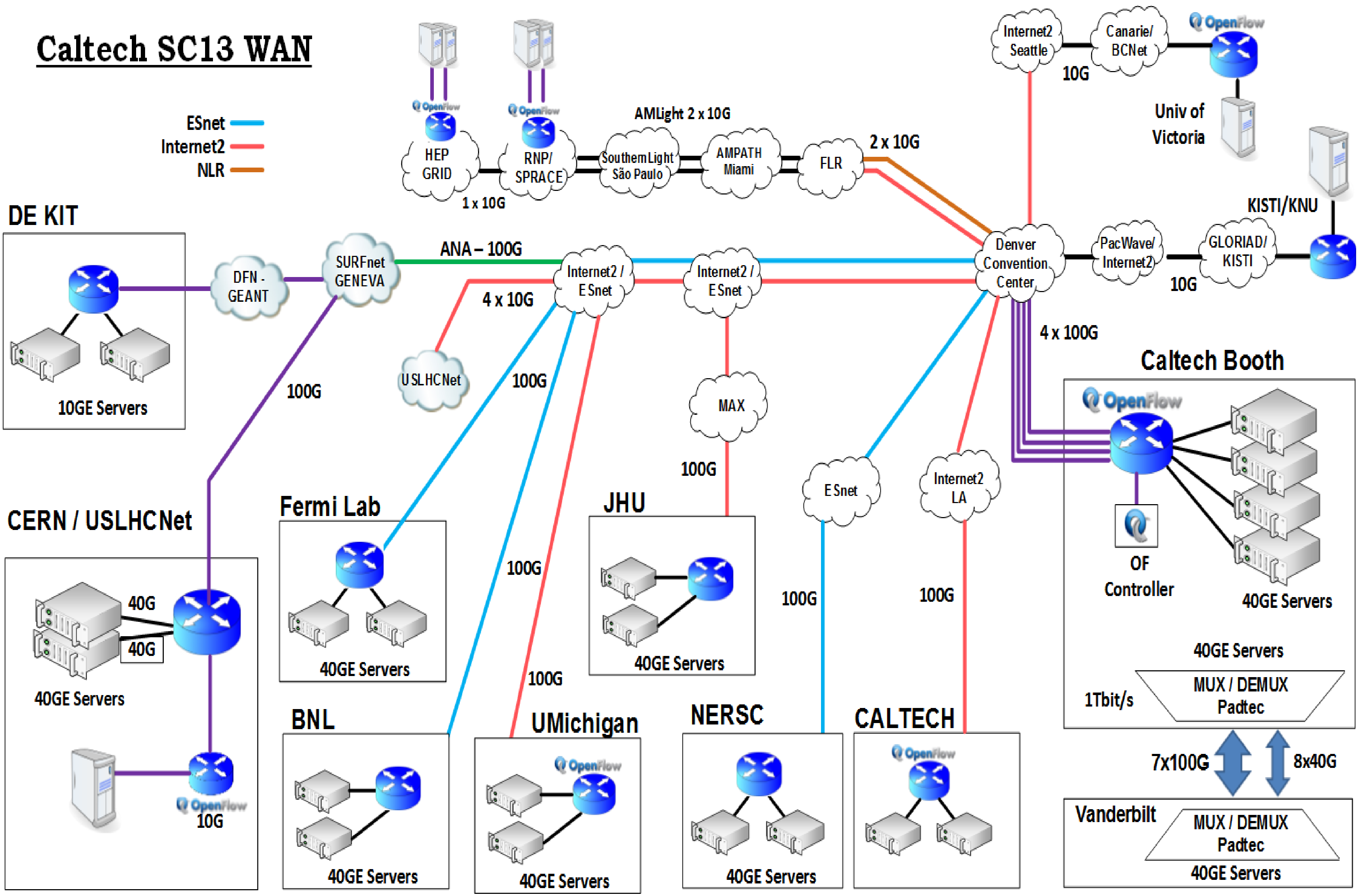
Drill-down capabilities

- Test history between hosts
- Ability to correlate with other events
- Very valuable for fault localization



(slides from Azher Mughal, Caltech)

SC13 Results

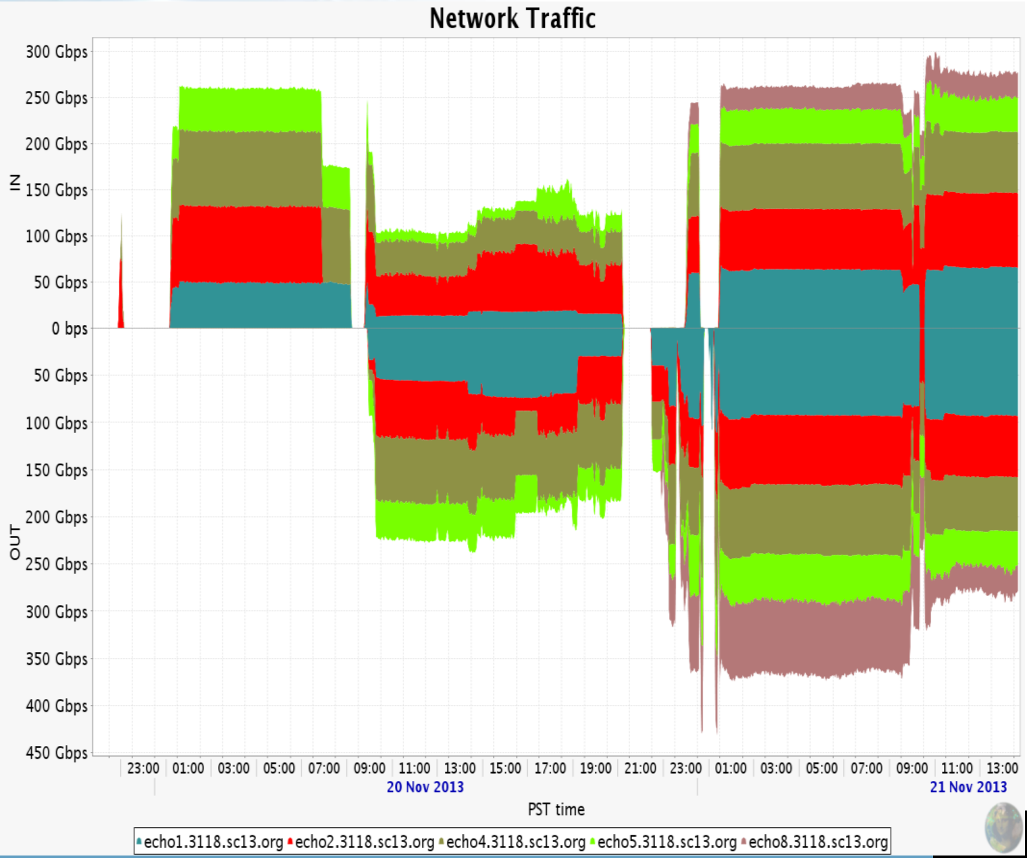


Supporting Vendors: Mellanox, Brocade, Echostreams, Intel, Cisco, Dell, Padtec, Ciena, SGI, Seagate, FusionIO, iWnetworks, Juniper, ADVA

TeraBit Demo

7x 100G links

8 x 40G links



SC13 Results

SC13 – DE-KIT
- 75Gb from Disk to Disk (couple of servers at KIT – Two servers at SC13)

SC13 BNL over ESnet:
- 80G over two pair of hosts, memory to memory

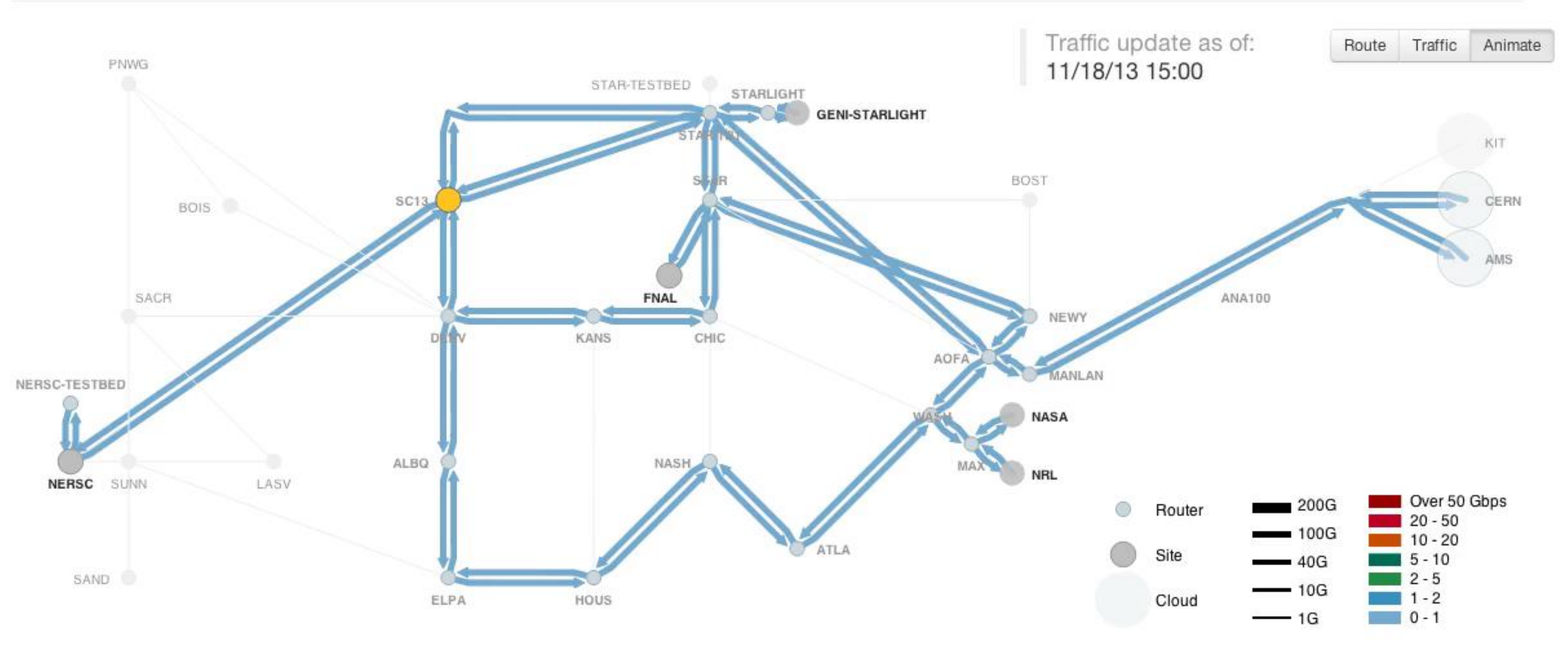
NERSC to SC13 over ESnet:
- Lots of packet loss at first, then removed the Mellanox switch from the path, and then the path was clean
- Consistent 90Gbps, reading from 2 SSD host sending to single host in the booth.

SC13 to FNAL over ESnet:
- Lots of packet loss; TCP max around 5Gbps, but UDP could do 15G per flow.
- Used 'tc' to pace TCP, and then at least single stream TCP behaved well up to 8G. But using multiple streams was still a problem. This seems to indicate something in the path with too small buffers, but we never figured out what.

SC13 – Pasadena Internet2:
- 80G read from the disks and write on the servers (disk to memory transfer). Link was lossy the other way.

SC13 – CERN over ESnet:
- About 75Gb memory to memory. Disks about 40Gb

SC13 Traffic Animation: https://my.es.net/demos/sc13



Other Technical Topics

Speed Mismatch Issues

More and more often we are seeing problems sending from a faster host to a slower host

- This can look like a network problem (lots of TCP retransmits)
- The network is rarely the bottleneck anymore for many sites

This may be true for:

- 10G to 1G host
- 10G host to a 2G circuit
- 40G to 10G host
- Fast host to slower host

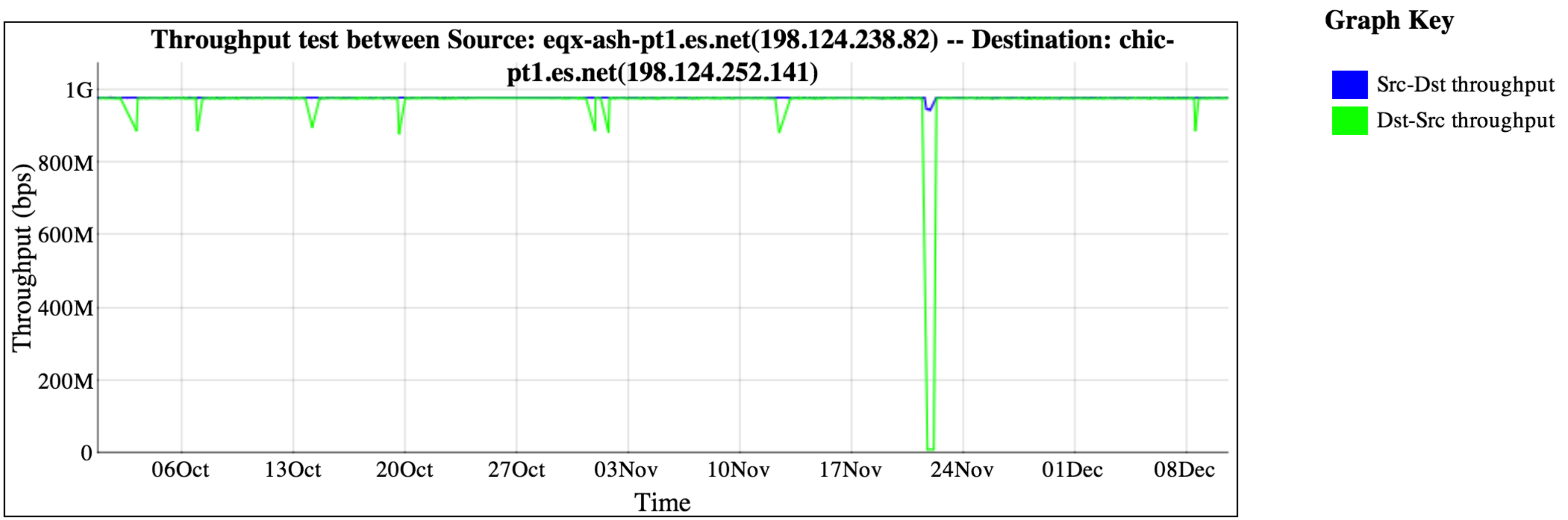
Pacing at the application level does not help

The linux ‘tc’ command does help

- But only up to speeds of 8Gbps
- And perhaps this will just mask problems with under-buffered switches?

http://fasterdata.es.net/host-tuning/packet-pacing/

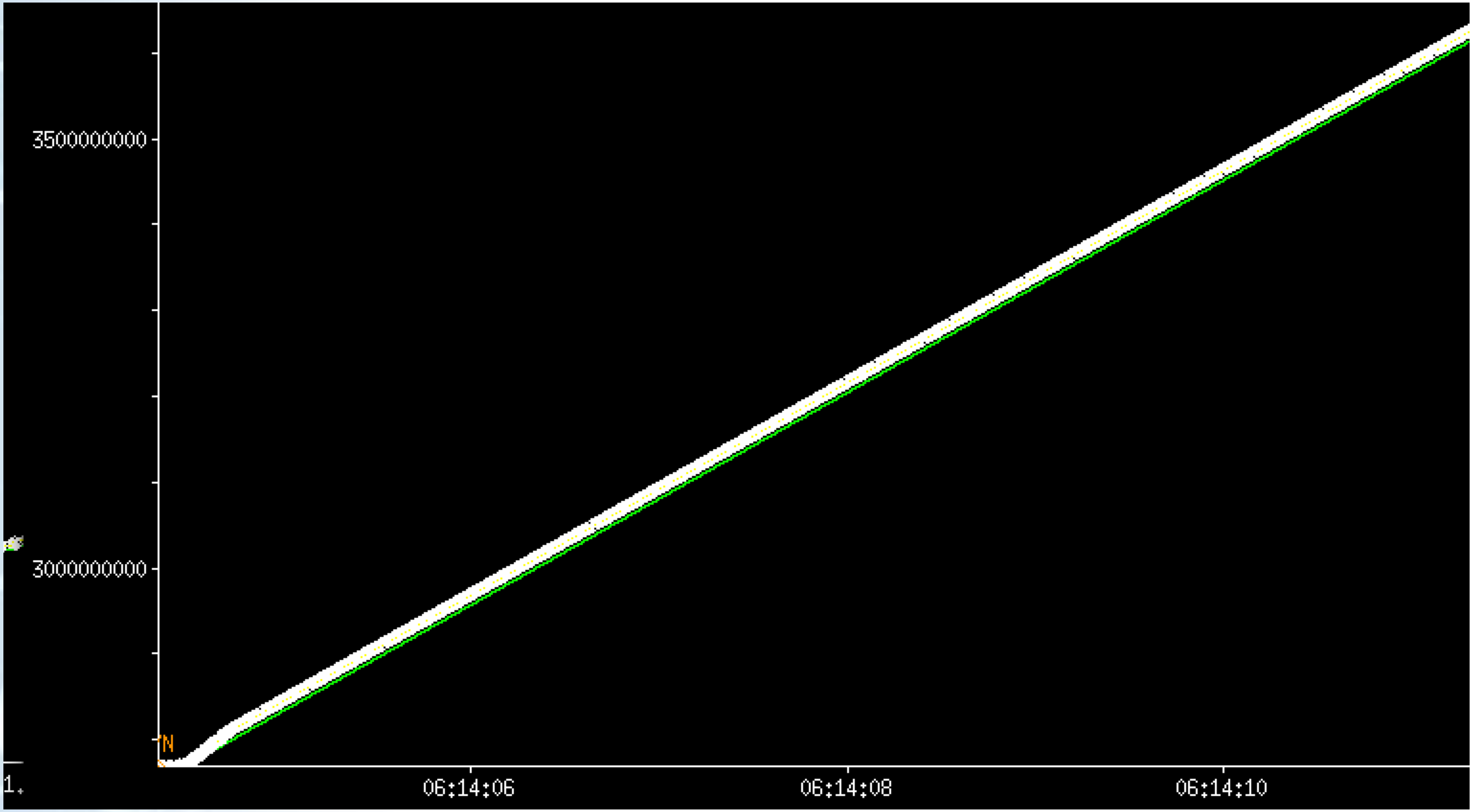
But 10G to 1G can work just fine too….



<- 1 month

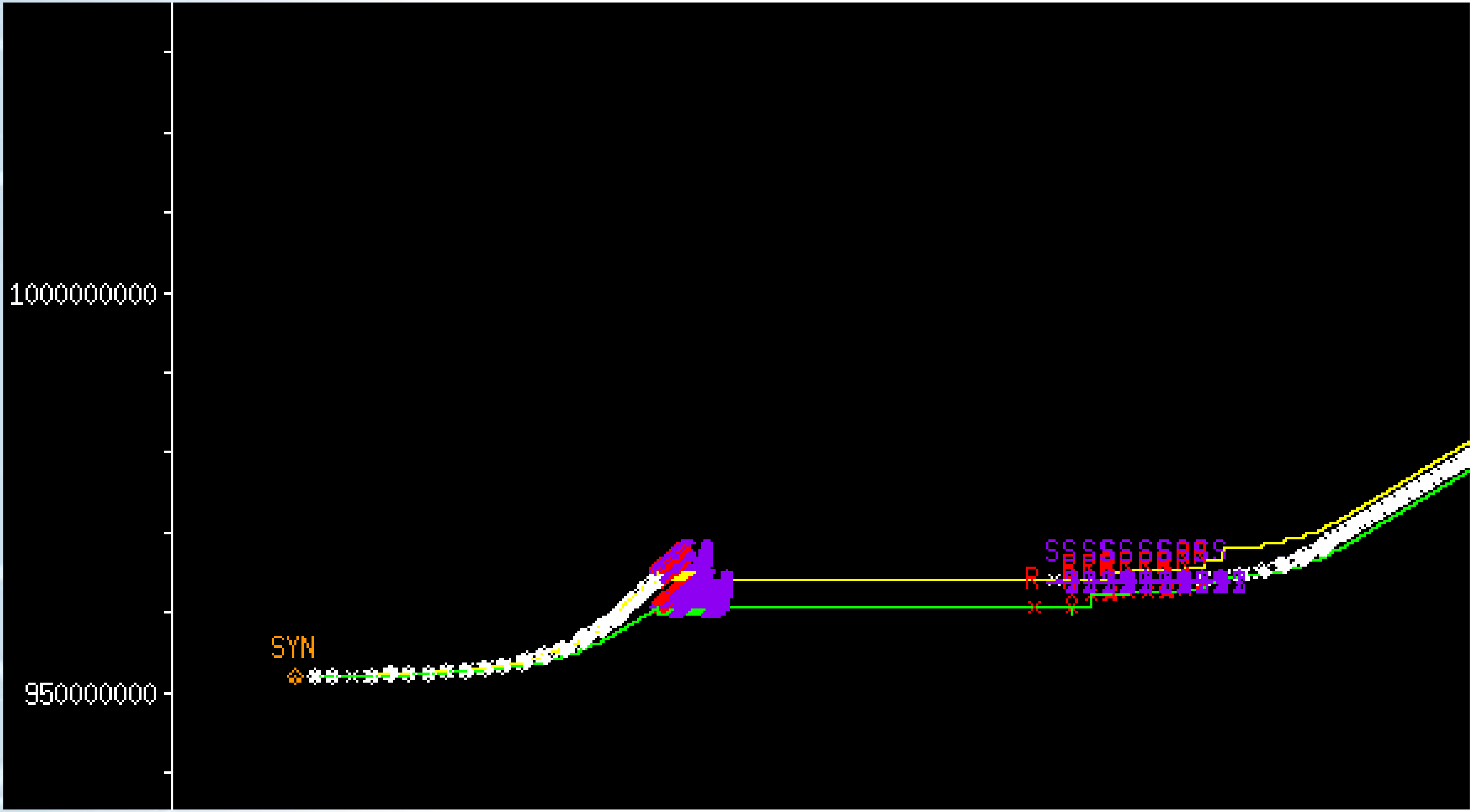
1 month ->

Compare tcpdumps:
kans-pt1.es.net (10G) to eqx-chi-pt1.es.net (1G)



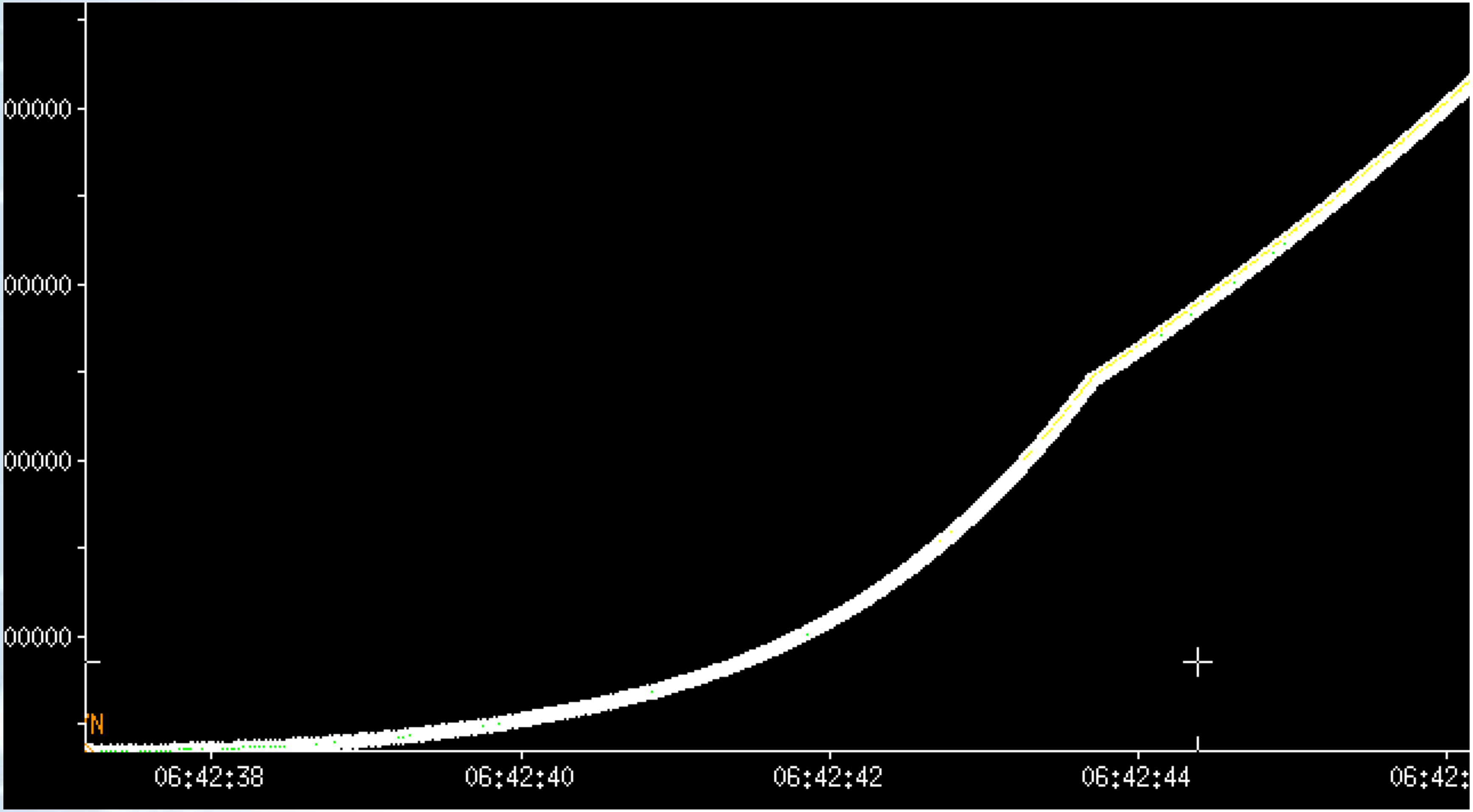
Compare tcpdumps

kans-pt1.es.net (10G) to uct2-net4.uchicago.edu (1G)



Compare tcpdumps

kans-pt1.es.net (10G) to uct2-net4.uchicago.edu (1G) with pacing



40G Lessons Learned

Tuning for 40G is not just 4x Tuning for 10G

Some of the conventional wisdom for 10G Networking is not true at 40Gbps

- e.g.: Parallel streams more likely to hurt than help

UDP needs to be tuned too

“Sandy Bridge” Architectures require extra tuning as well

Lots of details at http://fasterdata.es.net/science-dmz/DTN/tuning/

Sample results: TCP Single vs Parallel Streams

```
1 stream: iperf3 -c 192.168.102.9
[ ID] Interval           Transfer     Bandwidth       Retransmits
[  4]   0.00-1.00   sec  3.19 GBytes  27.4 Gbits/sec    0
[  4]   1.00-2.00   sec  3.35 GBytes  28.8 Gbits/sec    0
[  4]   2.00-3.00   sec  3.35 GBytes  28.8 Gbits/sec    0
[  4]   3.00-4.00   sec  3.35 GBytes  28.8 Gbits/sec    0
[  4]   4.00-5.00   sec  3.35 GBytes  28.8 Gbits/sec    0

2 streams: iperf3 -c 192.168.102.9 -P2
[ ID] Interval           Transfer     Bandwidth       Retransmits
[  4]   0.00-1.00   sec  1.37 GBytes  11.8 Gbits/sec    7
[  6]   0.00-1.00   sec  1.38 GBytes  11.8 Gbits/sec   11
[SUM]   0.00-1.00   sec  2.75 GBytes  23.6 Gbits/sec   18
……
- - - - - - - - - - - - - - - - - - - - - - - - -
[  4]   8.00-9.00   sec  1.43 GBytes  12.3 Gbits/sec    8
[  6]   8.00-9.00   sec  1.42 GBytes  12.2 Gbits/sec    7
[SUM]   8.00-9.00   sec  2.85 GBytes  24.5 Gbits/sec   15
- - - - - - - - - - - - - - - - - - - - - - - - -
[  4]   9.00-10.00  sec  1.43 GBytes  12.3 Gbits/sec    4
[  6]   9.00-10.00  sec  1.43 GBytes  12.3 Gbits/sec    6
[SUM]   9.00-10.00  sec  2.86 GBytes  24.6 Gbits/sec   10
- - - - - - - - - - - - - - - - - - - - - - - - -
[ ID] Interval           Transfer     Bandwidth       Retransmits
[  4]   0.00-10.00  sec  13.8 GBytes  11.9 Gbits/sec   78             sender
[  4]   0.00-10.00  sec  13.8 GBytes  11.9 Gbits/sec                  receiver
[  6]   0.00-10.00  sec  13.8 GBytes  11.9 Gbits/sec   95             sender
[  6]   0.00-10.00  sec  13.8 GBytes  11.9 Gbits/sec                  receiver
[SUM]   0.00-10.00  sec  27.6 GBytes  23.7 Gbits/sec  173             sender
[SUM]   0.00-10.00  sec  27.6 GBytes  23.7 Gbits/sec                  receiver
```

iperf3: https://code.google.com/p/iperf/

Sample results: TCP On Intel "Sandy Bridge" Motherboards

30% Improvement using the right core!

```
nuttcp -i 192.168.2.32
 2435.5625 MB /   1.00 sec = 20429.9371 Mbps     0 retrans
 2445.1875 MB /   1.00 sec = 20511.4323 Mbps     0 retrans
 2443.8750 MB /   1.00 sec = 20501.2424 Mbps     0 retrans
 2447.4375 MB /   1.00 sec = 20531.1276 Mbps     0 retrans
 2449.1250 MB /   1.00 sec = 20544.7085 Mbps     0 retrans

nuttcp -i1 -xc 2/2 192.168.2.32
 3634.8750 MB /   1.00 sec = 30491.2671 Mbps     0 retrans
 3723.8125 MB /   1.00 sec = 31237.6346 Mbps     0 retrans
 3724.7500 MB /   1.00 sec = 31245.5301 Mbps     0 retrans
 3721.7500 MB /   1.00 sec = 31219.8335 Mbps     0 retrans
 3723.7500 MB /   1.00 sec = 31237.6413 Mbps     0 retrans
```

nuttcp: http://lcp.nrl.navy.mil/nuttcp/beta/nuttcp-7.2.1.c

Sample results: TCP On Intel “Sandy Bridge” Motherboards: Fast host to Slower Host

Intel(R) Xeon(R) CPU 2.90GHz to 2.00GHz

```
nuttcp -i1 192.168.2.31
  410.7500 MB /   1.00 sec = 3445.5139 Mbps      0 retrans
  339.5625 MB /   1.00 sec = 2848.4966 Mbps      0 retrans
  354.5625 MB /   1.00 sec = 2974.2888 Mbps    350 retrans
  326.3125 MB /   1.00 sec = 2737.3022 Mbps      0 retrans
  377.7500 MB /   1.00 sec = 3168.8220 Mbps    179 retrans

nuttcp -r -i1 192.168.2.31  (reverse direction)
 2091.0625 MB /   1.00 sec = 17540.8230 Mbps      0 retrans
 2106.7500 MB /   1.00 sec = 17672.0814 Mbps      0 retrans
 2103.6250 MB /   1.00 sec = 17647.0326 Mbps      0 retrans
 2086.7500 MB /   1.00 sec = 17504.7702 Mbps      0 retrans
```

(slides from Mike O'Conner, ESnet)

LHCONE

LHCONE Collaborating NSPs and Compute Centers

- CANET(6509)
  - BCNET(271)
  - UTORONTO(239)
  - UVIC(16462)
  - MCGILL(15318)
  - TRIUMF(36391)
  - UALBERTA(3359)
- ESNET(293)
  - FNAL(3152)
  - BNL(43)
  - SLAC(3671)
- I2(11537)
  - UIUC(38)
  - UNL(7896)
  - MIT(3)
  - AGLT2(229)
  - **MICH-Z(230)**
  - UOC(160)
  - CSUNET(2153)
  - ULTRALIGHT(32361)
  - **VANDERBILT(39590)**
- INDIAN(19782)
  - IUPUI(10680)

North America

Europe

- CERN-LHC1(20641)
  - CERN-WIGNER(61339)
  - CERN(513)
- DFN(680)
  - KIT(34878)
  - DESY(1754)
- GEANT(20965)
  - ROEDUNET(2614)
  - ASGARR(137)
  - ARNES-NET(2107)
  - **CZECH-ACAD-SCI(2852)**
- LHC1-RENATER(2091)
  - IN2P3(789)
  - CEA-SACLAY(777)
- NORDUNET(2603)
  - NDGF(39590)

Migration to 100GE Substrate

LHCONE is deployed primarily over shared infrastructure.

All participating NSPs have either upgraded, or are in the process of upgrading their core networks to 100GE circuits.

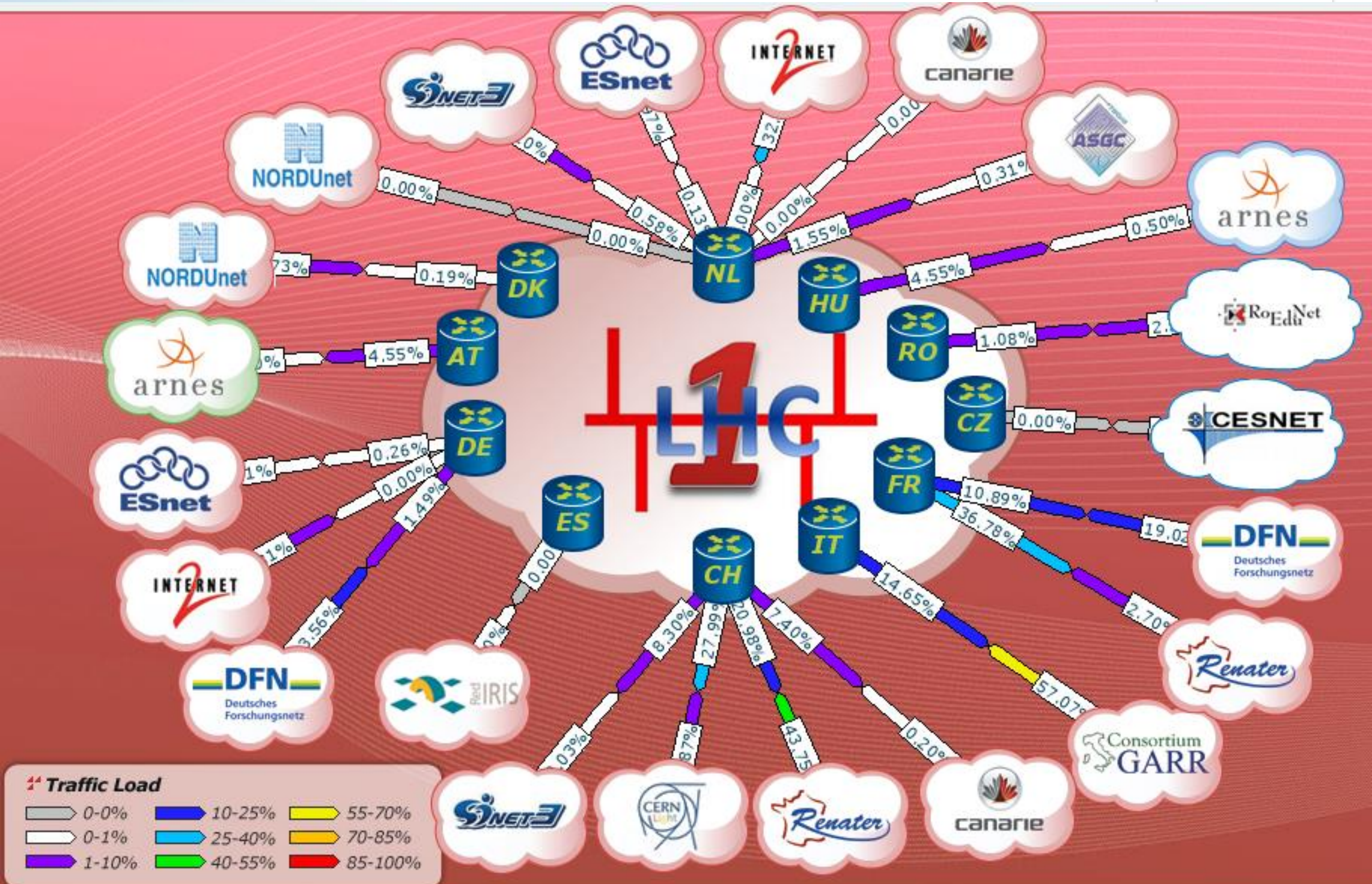
Single 10GE circuits that transport LHCONE along with other traffic are likely to experience periods of saturation.

NSPs should identify their remaining non-aggregated10GE segments and plan to eliminate them, this includes inter domain connections used for BGP peering.

A 100G network substrate will be essential for deploying various kinds of virtualized networks to address the needs of the growing number of distributed scientific collaborations world-wide.

LHCONE in Europe: GEANT:

(from Mian Usman's talk at last week's LHCONE meeting: http://indico.cern.ch/conferenceDisplay.py?confId=269840



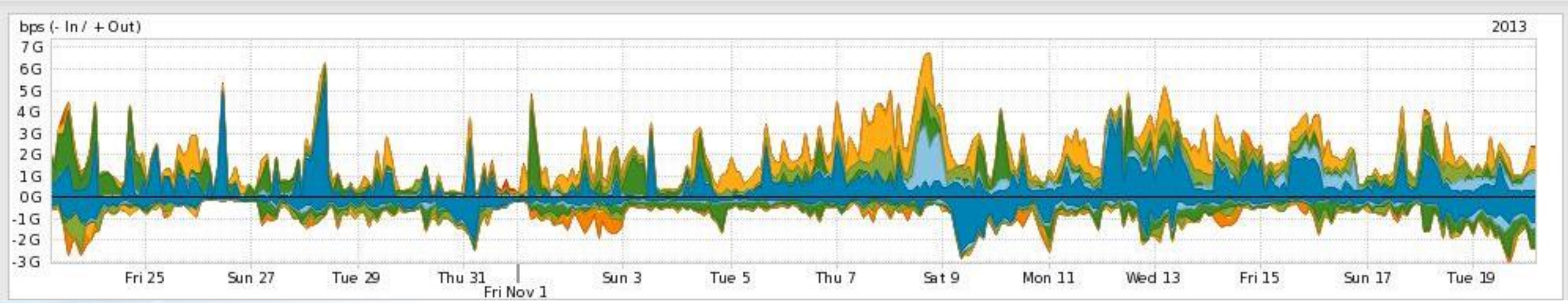
LHCONE and ESNet

Configured as L3VPN VRF and connects following sites:

- BNL (Atlas T1) FNAL (CMS T1) SLAC (ATLAS T2)

ESNet LHCONE VRF peers with Internet2, GEANT, NORDUNET, CERNLight and CANARIE LHCONE VRF

ESNET is present at StarLight, MANLAN, WIX and PNWG



LHC perfSONAR MPs

Tier	Type	Hostname	IP address
RAL	Latency:	perfsonar-ps02.gridpp.rl.ac.uk	130.246.179.197
	Bandwidth:	perfsonar-ps01.gridpp.rl.ac.uk	130.246.179.196
CC-IN2P3	Latency:	ccperfsonar2-lhcopn.in2p3.fr	193.48.99.78
	Bandwidth:	ccperfsonar-lhcopn.in2p3.fr	193.48.99.79
CERN	Latency:	perfsonar-ps2.cern.ch	128.142.223.237
	Bandwidth:	perfsonar-ps.cern.ch	128.142.223.236
TRIUMF	Latency:	ps-latency.lhcopn-mon.triumf.ca	206.12.9.71
	Bandwidth:	ps-bandwidth.lhcopn-mon.triumf.ca	206.12.9.70
SARA	Latency:	ps.lhcopn-ps.sara.nl	145.100.17.9
	Bandwidth:	ps.lhcopn-ps.sara.nl	145.100.17.9
ASGC	Latency:	lhc-latency.twgrid.org	117.103.105.188
	Bandwidth:	lhc-bandwidth.twgrid.org	117.103.105.187
BNL	Latency:	lhcperfmon.bnl.gov	192.12.15.26
	Bandwidth:	lhcmon.bnl.gov	192.12.15.23
CNAF	Latency:	perfsonar-ps.cnaf.infn.it	131.154.254.11
	Bandwidth:	perfsonar-ow.cnaf.infn.it	131.154.254.12
NDGF	Latency:	perfsonar-ps.ndgf.org	109.105.124.86
	Bandwidth:	perfsonar-ps2.ndgf.org	109.105.124.88
PIC	Latency:	psl01.pic.es	193.109.172.188
	Bandwidth:	psb01.pic.es	193.109.172.187
FNAL	Latency:	psonar2.fnal.gov	131.225.205.141
	Bandwidth:	psonar1.fnal.gov	131.225.205.139
KIT	Latency:	perfsonar2-de-kit.gridka.de	192.108.47.12
	Bandwidth:	perfsonar-de-kit.gridka.de	192.108.47.6

LHCONE Reachable

None of these MPs are located within NRENs.

Note: ESnet is considering deployment of dedicated LHCONE perfSONAR Infrastructure at STARLIGHT, MANLAN, and WIX.

LHCONE Summary

- LHCONE is critical to get to EU sites
  - with new Atlas compute model that is less hierarchal, this is even more important
- More 100G sites are coming online soon
- Need to redesign your site architecture to deal with this?
  - do you have a Science DMZ?
- Email engage@es.net if you want help

Questions?

Thanks!

Brian Tierney, bltierney@es.net or engage@es.net

http://www.es.net/

http://fasterdata.es.net/

Extra Slides

100Gbps Networks
Network Engineering Perspective

Eli Dart, Joe Metzger

100 Gbps transatlantic science trials workshop at SC13

Denver, CO

November 18, 2013

Experience With 100G Equipment

ESnet experiences

- Advanced Networking Initiative
- ESnet5 production 100G network
- Helping other people debug their stuff

Important takeaways

- R&E requirements are outside the design spec for most gear
  - Results in platform limitations – sometimes can't be fixed
  - You need to be able to identify those limitations before you buy
- R&E requirements are outside the test scenarios for most vendors
  - Bugs show up when R&E workload is applied
  - You need to be able to troubleshoot those scenarios

Platform Limitations

We have seen significant limitations in 100G equipment from all vendors with a major presence in R&E

- 100G single flow not supported
  - Channelized forwarding plane
  - Unexplained limitations
  - Sometimes the senior sales engineers don’t know!
- Non-determinism in the forwarding plane
  - Performance depends on features used (i.e. config-dependent)
  - Packet loss that doesn’t show up in counters anywhere

If you can’t find it, nobody will tell you about it

- Vendors don’t know or won’t say
- Watch how you write your procurements

Second-generation equipment has proven to be much better

Vendors have been responsive in rolling new code to fix problems

They Don't Test For This Stuff

Most sales engineers and support engineers don't have access to 100G test equipment

- It's expensive
- Setup of scenarios is time-consuming

R&E traffic profile is different than their standard model

- IMIX (Internet Mix) traffic is normal test profile
  - Aggregate web browsers, email, YouTube, Netflix, etc.
  - Large flow count, low per-flow bandwidth
  - This is to be expected – that's where the market is
- R&E shops are the ones that get the testing done for R&E profile
  - SCinet provides huge value
  - But, in the end, it's up to us

New Technology, New Bugs

Bugs happen.

- Data integrity (traffic forwarded, but with altered data payload)
- Packet loss
- Interface wedge
- Optics flaps

Monitoring systems are indispensable

Finding and fixing issues is sometimes hard

- Rough guess – difficulty exponent is degrees of freedom
  - Vendors/platforms, administrative domains, time zones

Takeaway – don't skimp on test gear (at least maintain your perfSONAR boxes)

Design For Easy Debug

International circuits often have special circumstances

- Undersea cables
- Multiple administrative domains for one circuit

These things can make debugging harder than for terrestrial circuits

TCP loss impact and other issues are more damaging

It must be easy to run tests on international circuits

- Regular monitoring with perfSONAR
- As-needed testing for debugging specific issues

Workflow Decomposition

Many people still think in terms of one program running inside one system image on one computer

Workflows that process tens of terabytes of data must work differently

What does your workflow look like?

- What produces the data?
- Where is the storage?
- What does the analysis? (What storage goes with analysis?)
- Where can data be reduced?
- What can be automated?

Different components have different requirements

Proper decomposition can have significant benefits

Component Reuse

Many people understand about software reuse

Not many people understand workflow component reuse

Do you really want to re-invent the wheel?

- High-speed data transfer (Globus)
- Integration of virtualized components (OpenStack)
- Volume rendering, feature detection, FFT, CFD, …

Many scientists/experiments think they are a unique snowflake

- In some ways they are
- However, there is a set of tasks common to many workflows

Find your commonalities and exploit them – we can't scale otherwise

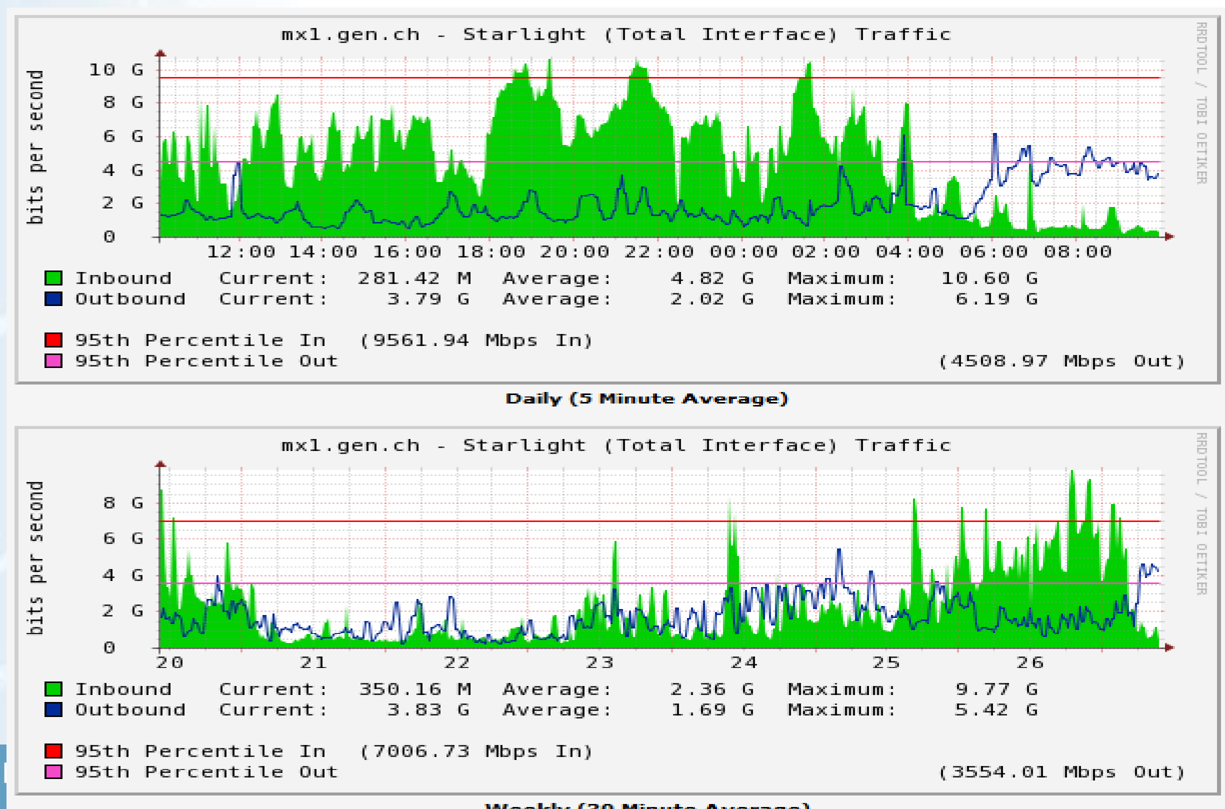
Trans-Atlantic Links

NORDUNET ~ 1.5Gbps peaks

NREN funded LHCONE dedicated links ~12Gbps peaks

ACE and GEANT funded shared links ~8Gbps

USLHCNET LHCONE dedicated link ~



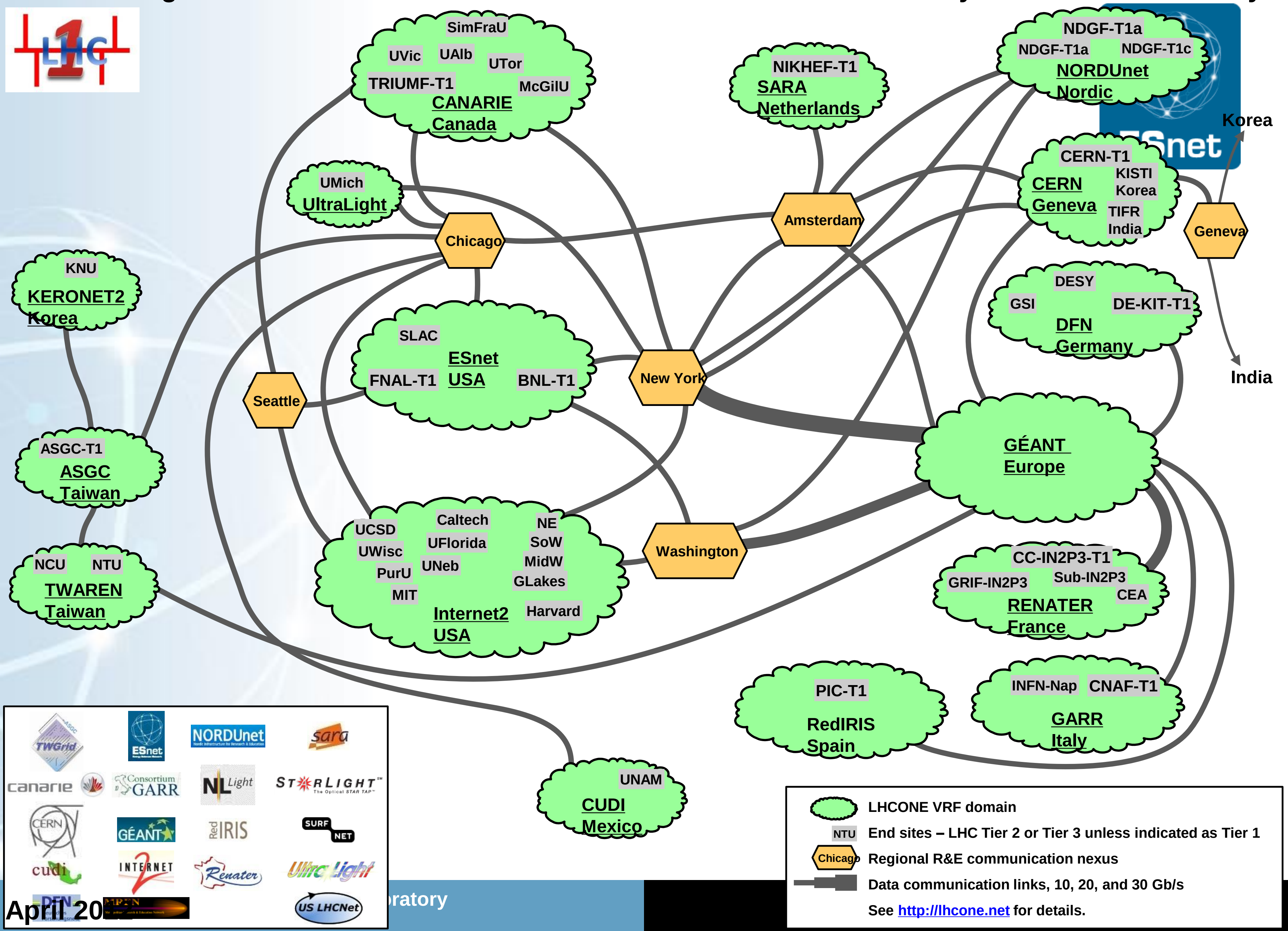
LHCONE in Open Exchanges

	MANLAN	StarLight	WIX	NetherLight	CERNLight
GEANT	★	★	★	★	★
NORDUnet	★			★	
Internet2	★	★	★		
ESnet	★	★	★		
CANARIE	★	★			
ASGC		★		★	

The LHC's Open Network Environment – LHCONE

- LHCONE provides a private, managed infrastructure designed for LHC Tier 2 traffic (and likely other large-data science projects in the future)
- The approach is an overlay network whose architecture is
  - A collection of routed "clouds" using address spaces restricted to subnets that are used by LHC systems
    - The clouds are mostly local to a network domain (e.g. one for each involved domain – ESnet, GEANT ("fronts" for the NRENs), Internet2 (fronts for the US universities), etc.
  - The clouds (VRFs) are interconnected by point-to-point circuits provided by various entities (mostly the domains involved)
- In this way the LHC traffic will use circuits designated by the network engineers
  - To ensure continued good performance for the LHC and to ensure that other traffic is not impacted – this is critical because apart from the LHCOPN, the R&E networks are funded for the benefit of the entire R&E community, not just the LHC

LHCONE: A global infrastructure for the LHC Tier1 data center – Tier 2 analysis center connectivity



April 20

oratory

The LHC's Open Network Environment – LHCONE

LHCONE could be set up relatively "quickly" because

- The VRF technology is a standard capability in most core routers, and
- there is capacity in the R&E community that can be made available for use by the LHC collaboration that cannot be made available for general R&E traffic

LHCONE is essentially built as a collection of private overlay networks (like VPNs) that are interconnected by managed links to form a global infrastructure where Tier 2 traffic will get good service and not interfere with general traffic

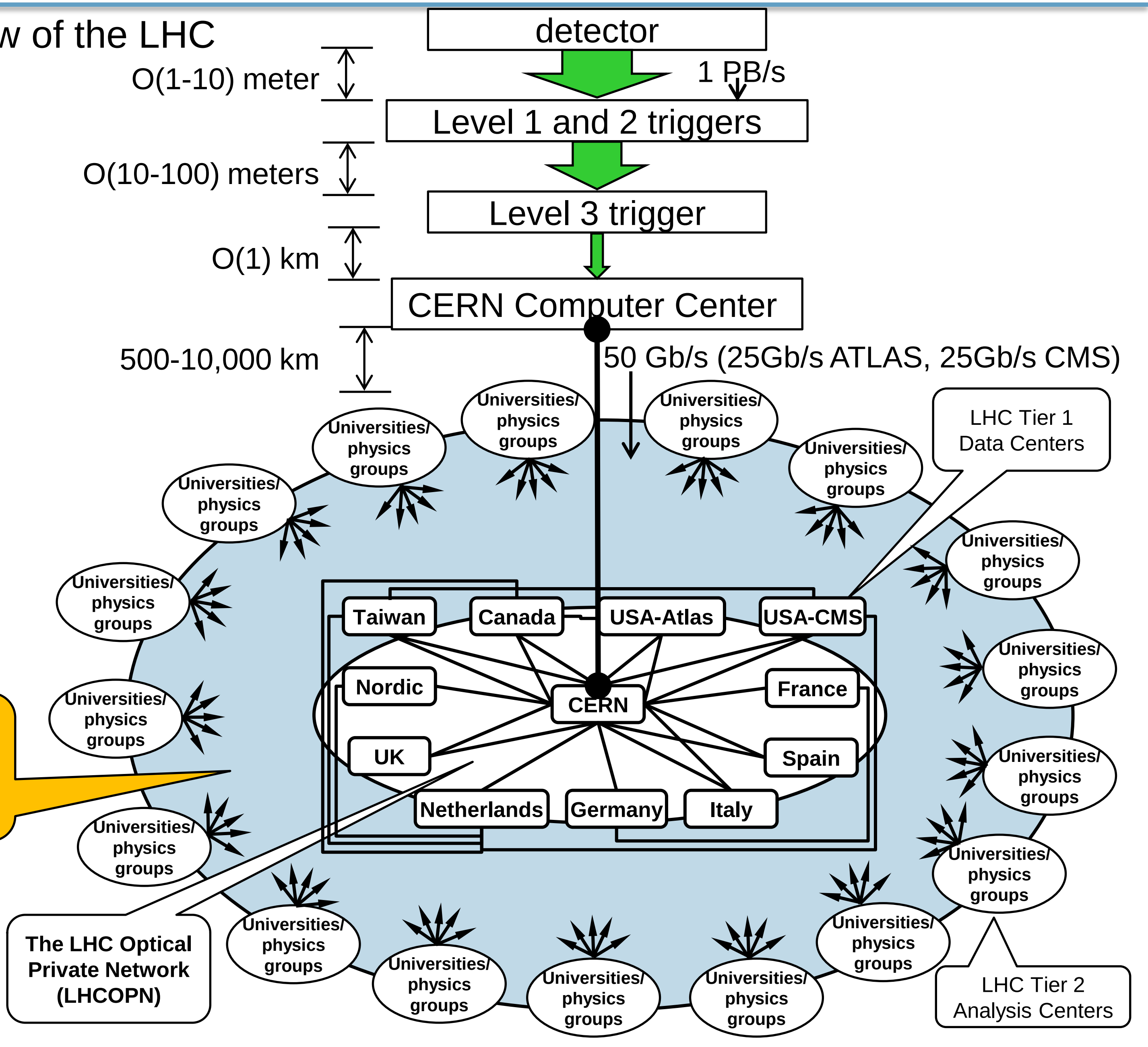
From the point of view of the end sites, they see a LHC-specific environment where they can reach all other LHC sites with good performance

See LHCONE.net

LHCONE is one part of the network infrastructure that supports the LHC

A Network Centric View of the LHC

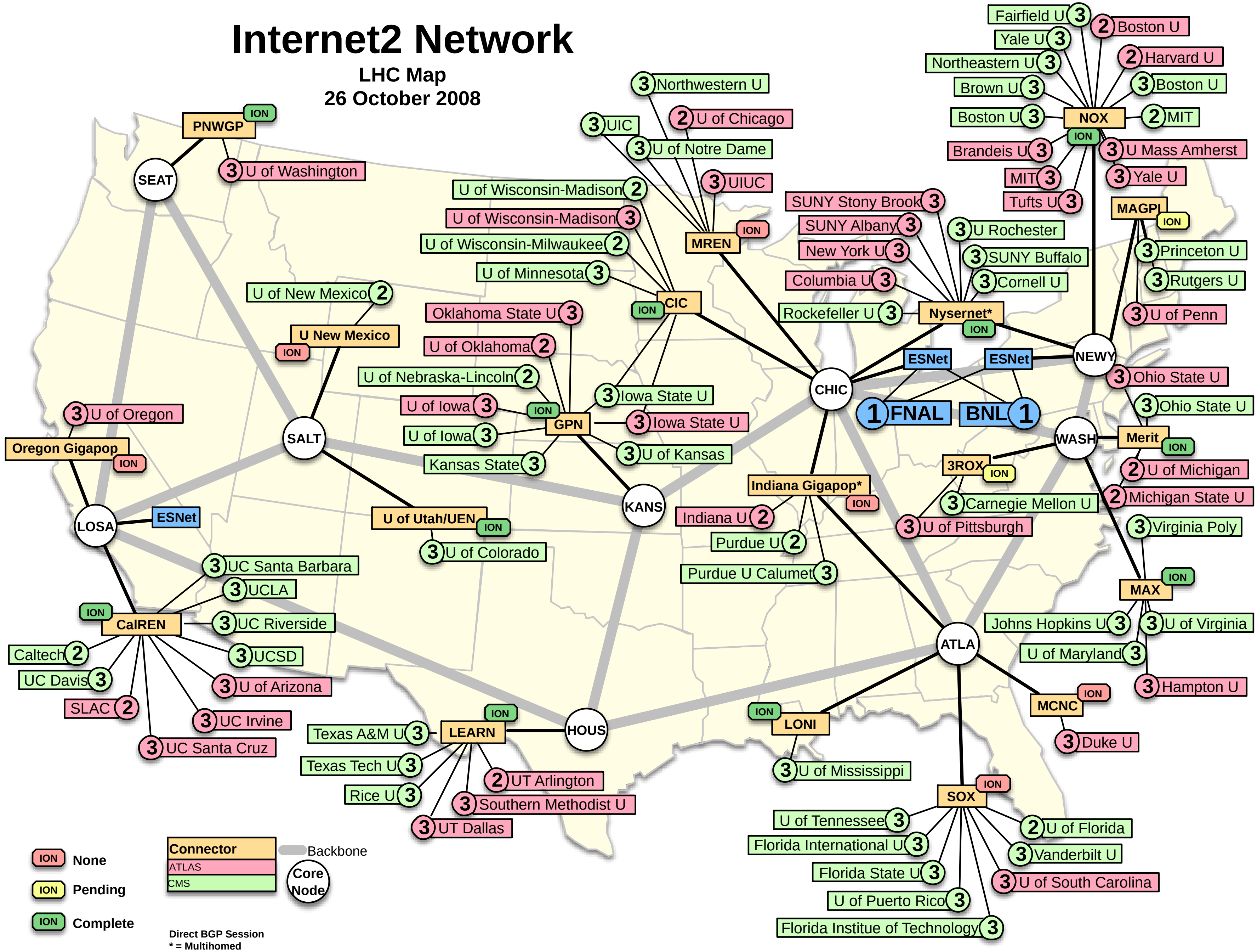
CERN →T1	miles	kms
France	350	565
Italy	570	920
UK	625	1000
Netherlands	625	1000
Germany	700	1185
Spain	850	1400
Nordic	1300	2100
USA – New York	3900	6300
USA - Chicago	4400	7100
Canada – BC	5200	8400
Taiwan	6100	9850



The LHC Open Network Environment (LHCONE)

This [arrow symbol] is intended to indicate that the physics groups now get their data wherever it is most readily available

LHC Site Map



Internet2 Network
LHC Map
26 October 2008
PNWGP
ION
SEAT
3 U of Washington
U of New Mexico 2
U New Mexico
ION
3 U of Oregon
Oregon Gigapop
ION
SALT
LOSA
ESNet
U of Utah/UEN
ION
3 U of Colorado
3 UC Santa Barbara
3 UCLA
ION
CalREN
3 UC Riverside
Caltech 2
UC Davis 3
3 UCSD
3 U of Arizona
SLAC 2
3 UC Irvine
3 UC Santa Cruz
U of Wisconsin-Madison 2
U of Wisconsin-Madison 3
U of Wisconsin-Milwaukee 2
U of Minnesota 3
Oklahoma State U 3
U of Oklahoma 2
U of Nebraska-Lincoln 2
U of Iowa 3
U of Iowa 3
Kansas State 3
ION
GPN
3 Iowa State U
3 Iowa State U
3 U of Kansas
KANS
HOUS
ION
LEARN
Texas A&M U 3
Texas Tech U 3
Rice U 3
3 UT Dallas
2 UT Arlington
3 Southern Methodist U
3 Northwestern U
3 UIC
2 U of Chicago
3 U of Notre Dame
3 UIUC
ION
MREN
ION
CIC
CHIC
1 FNAL
BNL 1
ESNet
ESNet
Indiana Gigapop*
ION
Indiana U 2
Purdue U 2
Purdue U Calumet 3
ION
LONI
3 U of Mississippi
ATLA
ION
MCNC
3 Duke U
ION
SOX
U of Tennessee 3
Florida International U 3
Florida State U 3
U of Puerto Rico 3
Florida Institue of Technology 3
2 U of Florida
3 Vanderbilt U
3 U of South Carolina
Fairfield U 3
Yale U 3
Northeastern U 3
Brown U 3
Boston U 3
Brandeis U 3
MIT 3
Tufts U 3
2 Boston U
2 Harvard U
3 Boston U
NOX
2 MIT
ION
3 U Mass Amherst
3 Yale U
SUNY Stony Brook 3
SUNY Albany 3
New York U 3
Columbia U 3
Rockefeller U 3
3 U Rochester
3 SUNY Buffalo
3 Cornell U
Nysernet*
ION
MAGPI
ION
3 Princeton U
3 Rutgers U
3 U of Penn
NEWY
3 Ohio State U
3 Ohio State U
WASH
Merit
ION
3ROX
ION
2 U of Michigan
2 Michigan State U
3 Carnegie Mellon U
3 U of Pittsburgh
3 Virginia Poly
ION
MAX
Johns Hopkins U 3
3 U of Virginia
U of Maryland 3
3 Hampton U
ION None
ION Pending
ION Complete
Connector
ATLAS
CMS
Backbone
Core Node
Direct BGP Session
* = Multihomed