

ESnet Network Update Network Engineering Group

ESCC
February 2014

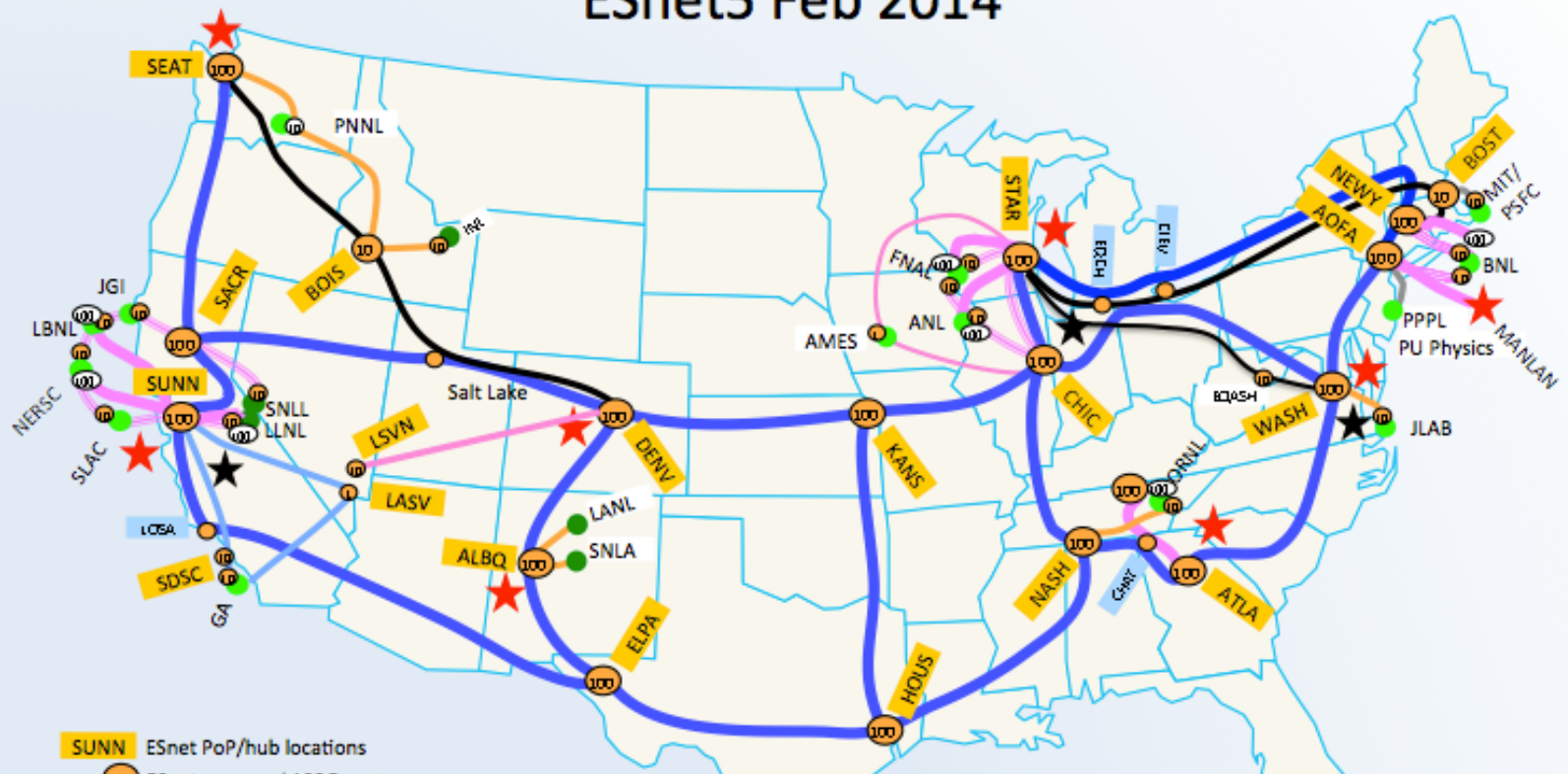


Outline



- Maps
- Topology Changes
- Service Changes
- Operational & Miscellaneous

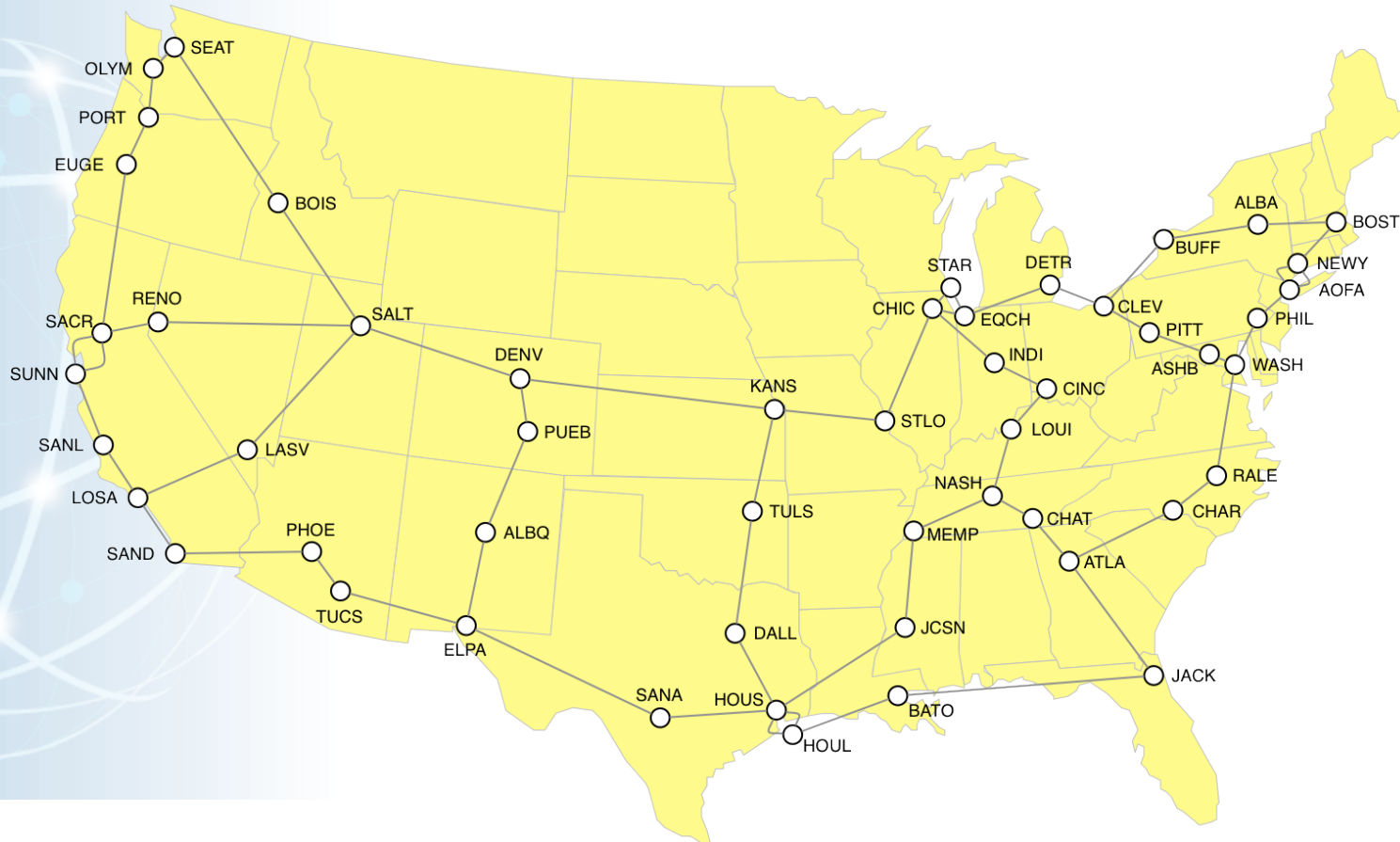
ESnet5 Feb 2014



- SUNN** ESnet PoP/hub locations
- ESnet managed 100G routers
- ESnet managed 10G router
- Site managed routers
- LOSA ESnet optical node locations (only some are shown)
- ESnet optical transport nodes (only some are shown)
- commercial peering points
- R&E network peering locations
- LBNL Major Office of Science (SC) sites
- LLNL Major non-SC DOE sites

- Routed IP 100 Gb/s
- Routed IP 4 X 10 Gb/s
- 3rd party 10Gb/s
- Express / metro 100 Gb/s
- Express / metro 10G
- Express multi path 10G
- Lab supplied links
- Other links
- Tail circuits

ESnet Optical Footprint: Add/Drops



Topology Changes



- Recent or in Progress
 - **Current 100G Connectivity**
 - **Las Vegas**
 - **Equinix Ashburn & DC Ring**
 - Chicago - Added optical nodes at ANL and FERMI
 - LIMAN - Changing dark fiber vendors for 1 side of the ring
 - Ames Lab 1G to 10G upgrade
- Future
 - **EEX**
 - Optical node at LLNL
 - 100G wave from Chicago to New York City
 - Swapping low density (1-port) for high density (2-port) 100G cards at SUNN, STAR & AOFA

ESnet 100G Connections



- Sites

- ANL
- BNL*
- FERMI*
- LLNL*
- NERSC
- LBNL
- ORNL

* Not currently used for production

- Peers

- Internet2
- MAX* (Temporary)
- MANLAN
- OMNIPOP
- PACWAVE*/CENIC
- STARLIGHT* (Ciena)
- WIX
-

- ESnet Testbed*

- MANLAN
- NERSC
- ESnet Backbone



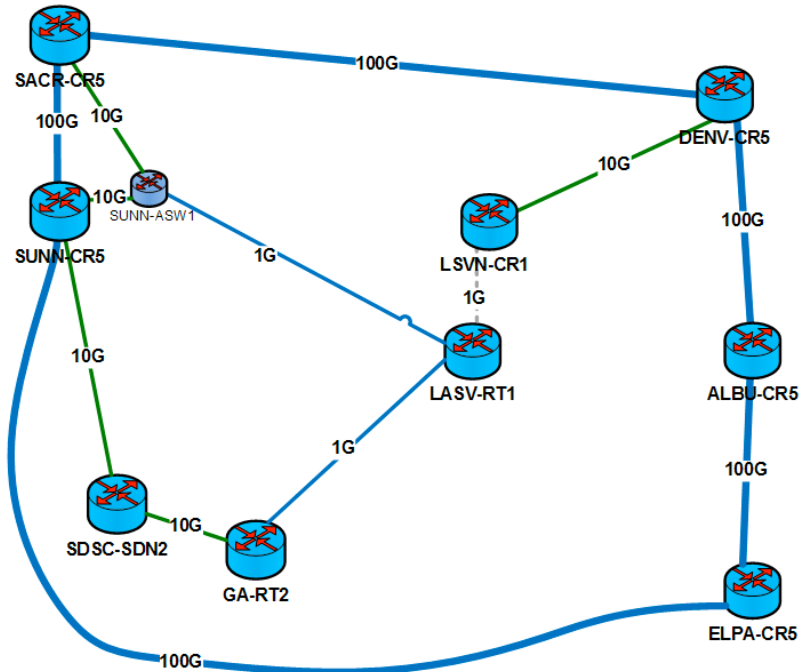
Moving Las Vegas Hub

- Deploying a new hub in Las Vegas to:
 - Improve services
 - Reduce costs
- Current Status
 - New hub with a 10G link to DENV brought up last week
- Next Steps
 - Turn up connection between Las Vegas hubs
 - Migrate tail circuits to sites from old hub to new hub
 - Decommission 1G from Las Vegas to Sunnyvale
 - Decommission old hub

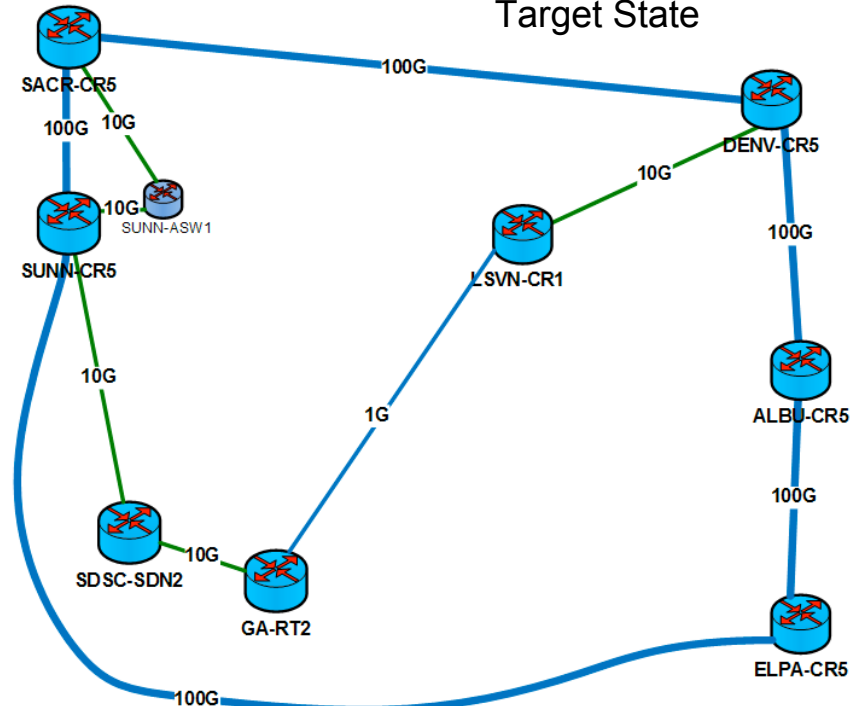
Las Vegas Hub Move



Current State



Target State

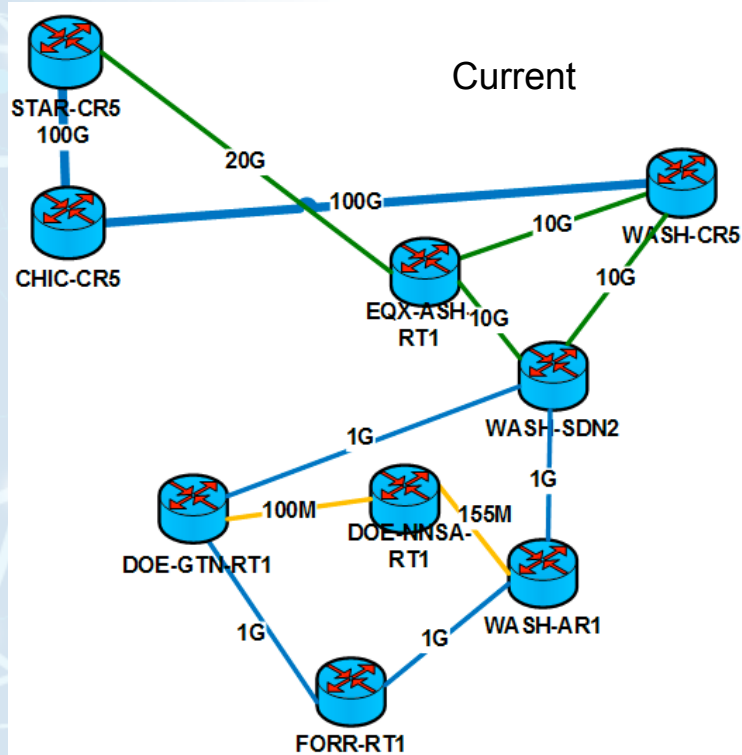




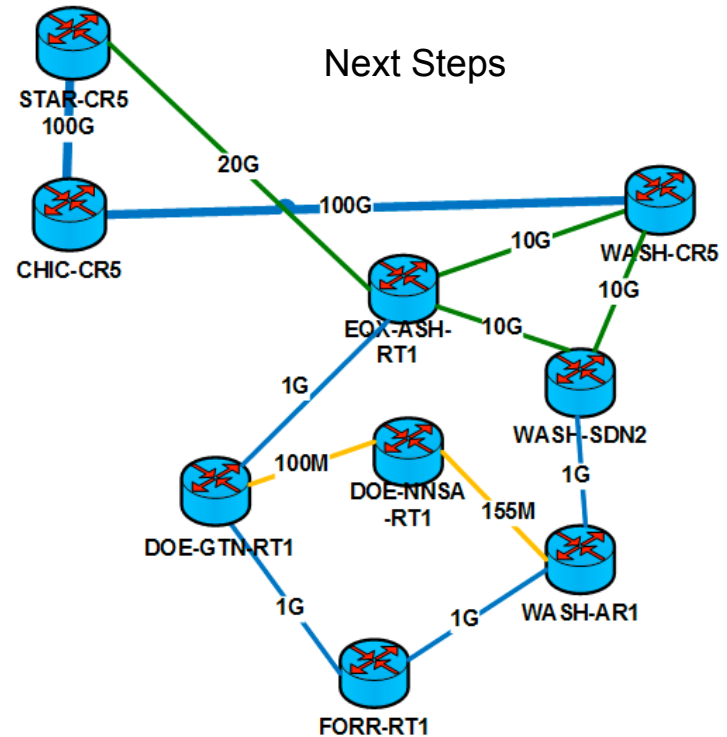
Washington DC Changes

- Re-arranging Circuits in DC to:
 - Improve redundancy/reliability
 - Reduce Costs
 - Simplify Operations
- Current Status
 - Diverse Egress from eqx-ash-rt1 done.
 - Orders placed for rehomeing DOE-GTN-RT1 circuit
- Next Steps
 - Finish transition to a ring with diverse egress points
 - Collapse routers at Germantown
 - Collapse routers at WASH hub

Washington DC Changes



+

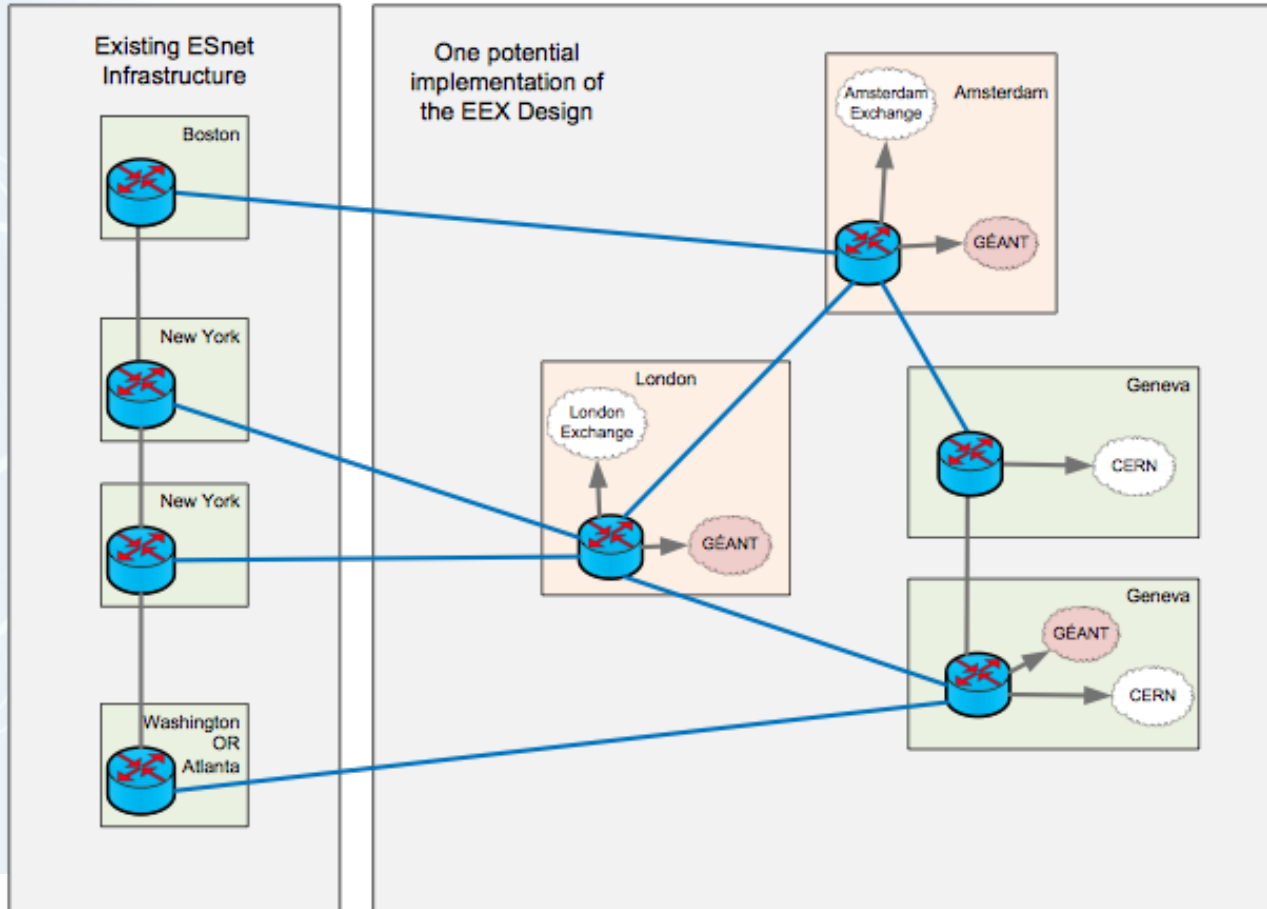




EEX - ESnet Extension to Europe

- Drivers
 - Improve quality & quantity of transit to Europe for all ESnet supported programs
 - Support LHC services to Europe
- Architecture
 - Hubs
 - CERN
 - London
 - Amsterdam
 - Links
 - 100G Ring between European Hubs
 - 4 trans-atlantic circuits between 4 US hubs and 3 EU hubs
 - 40G & 100G circuits, no 10G bottlenecks
- AUP
 - We will be expanding the ESnet AUP if EEX is approved/funded by LHC
- Timeline
 - US LHC project approval in mid March
 - Goal is completion before end of Fiscal Year

EEX



Service Changes



- **OSCARS**
 - Feature Updates: NSI Roadmap, Protection, hard policing and 0 BW reservations.
 - QOS Model
- **LHCONE LHC layer 3 overlay network**
 - Three peers added (since July 13)
 - LHCONE ANA-100 Integration testing begins March 2014
 - CANET, DANTE, ESnet, I2, NORDUNET, SURFNET
- **Service Now ticket system**
 - We are adding a bit more structure and rigor to our internal processes.
 - This shouldn't be customer visible. If it is, let us know.

OSCARS Timelines



- Implementation of OSCARS NSI Bridge to support of NSI CS v2.0 (r99) (4Q13)
- Deployment of NSI CS v2.0 (r99) in production infrastructure; ESnet*, MANLAN (4Q13 – current)
- Update of OSCARS NSI Bridge to support NSI CS v2.0 (ratified) (1Q14)
- OSCARS enhancements to support ESnet service extensions
 - Soft/hard enforcement of reservation bandwidth (4Q13)
 - “Zero” bandwidth best effort VCs (1Q14)
 - VC protection/restoration (1Q14)
 - OpenFlow (including multi-layer transport SDN) (3Q13)

*NB: Topology limited to A-GOLE demo STPs

OSCARS enhancements to support ESnet service extensions*



On-demand Secure Circuits and Advance Reservation System
 The OSCARS service is provided by ESnet, the Energy Sciences Network.
 OSCARS software developed by ESnet, in conjunction with Internet2 and ISI East.

February 19, 2014 13:00

Reservations | Reservation Details | Create Reservation | User Institutions | Authorizations | Authorization Details

Reservation creation form

Required inputs are bordered in green. The source and destination can be used to bring up a calendar widget. The reservation time slot defaults to now, and now + 15 minutes, respectively, if you leave the dates and times empty.

Production circuit

Source:

Destination:

Path (series of hops):

Bandwidth (Mbps):

Description:

Start date:

Start time:

End date:

End time:

Policing:

Protection:

Apply QoS:

Use layer 2 parameters Use layer 3 parameters <--> Same VLAN on source and destination

Source VLAN:

Source VLAN type:

Destination VLAN type:

WBUI

- Soft: Oversubscribed VC traffic is reclassified into lower priority
- Hard: Oversubscribed VC traffic is discarded

API (OptionalConstraints)
 Category: "policing"
 StringValue: "soft" (default) / "hard"

Protection

WBUI

- None: No secondary/backup path
- Loose Secondary Path: Backup path is active and will use IGP metrics for path-finding (and automatic re-route) with VC traffic classified as best-effort

API (OptionalConstraints)
 Category: "protection"
 StringValue: "none" (default) / "loose-secondary-path"

Apply QoS

WBUI

- Yes: Enforce VC bandwidth guarantee
- No: Do not enforce VC bandwidth guarantee, but classify traffic as best-effort

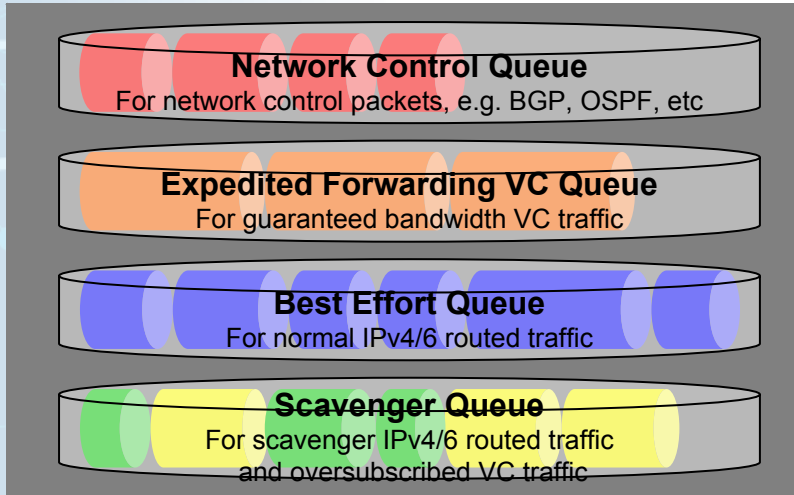
API (OptionalConstraints)
 Category: "apply-qos"
 StringValue: "true" (default) / "false"

*NB: Service extensions only implemented within ESnet OSCARS

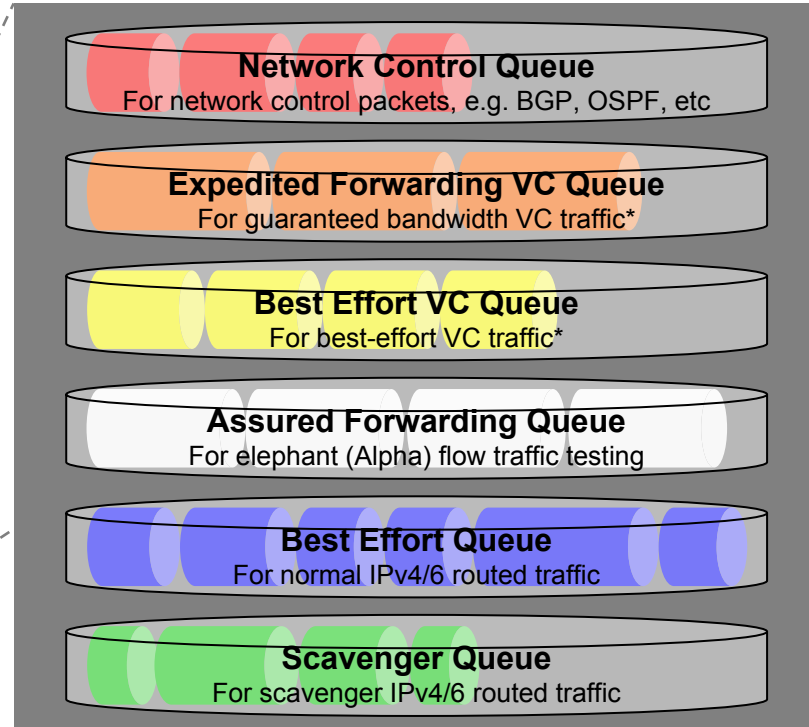
ESnet QoS changes



Before

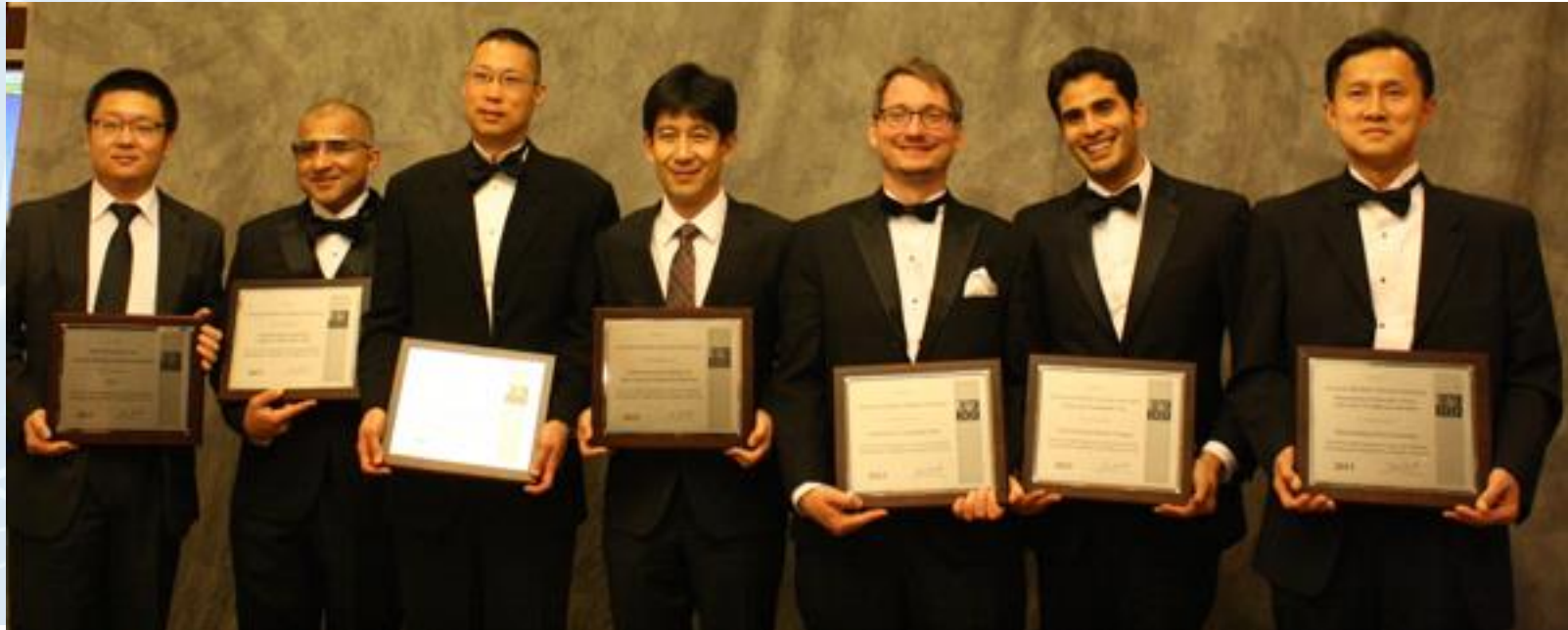


Now



*NB: High and low priority traffic classification within queue dictates which packets should be dropped first (e.g. oversubscribed VC traffic)

Chin Guok (3rd from left) accepts the R&D Magazine 100 Award for OSCARS in Orlando, FL on Nov. 7 2013.



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)



CERN-LHC1(20641)
CERN-WIGNER(61339)
CERN(513)
DFN(680)
KIT(34878)
DESY(1754)

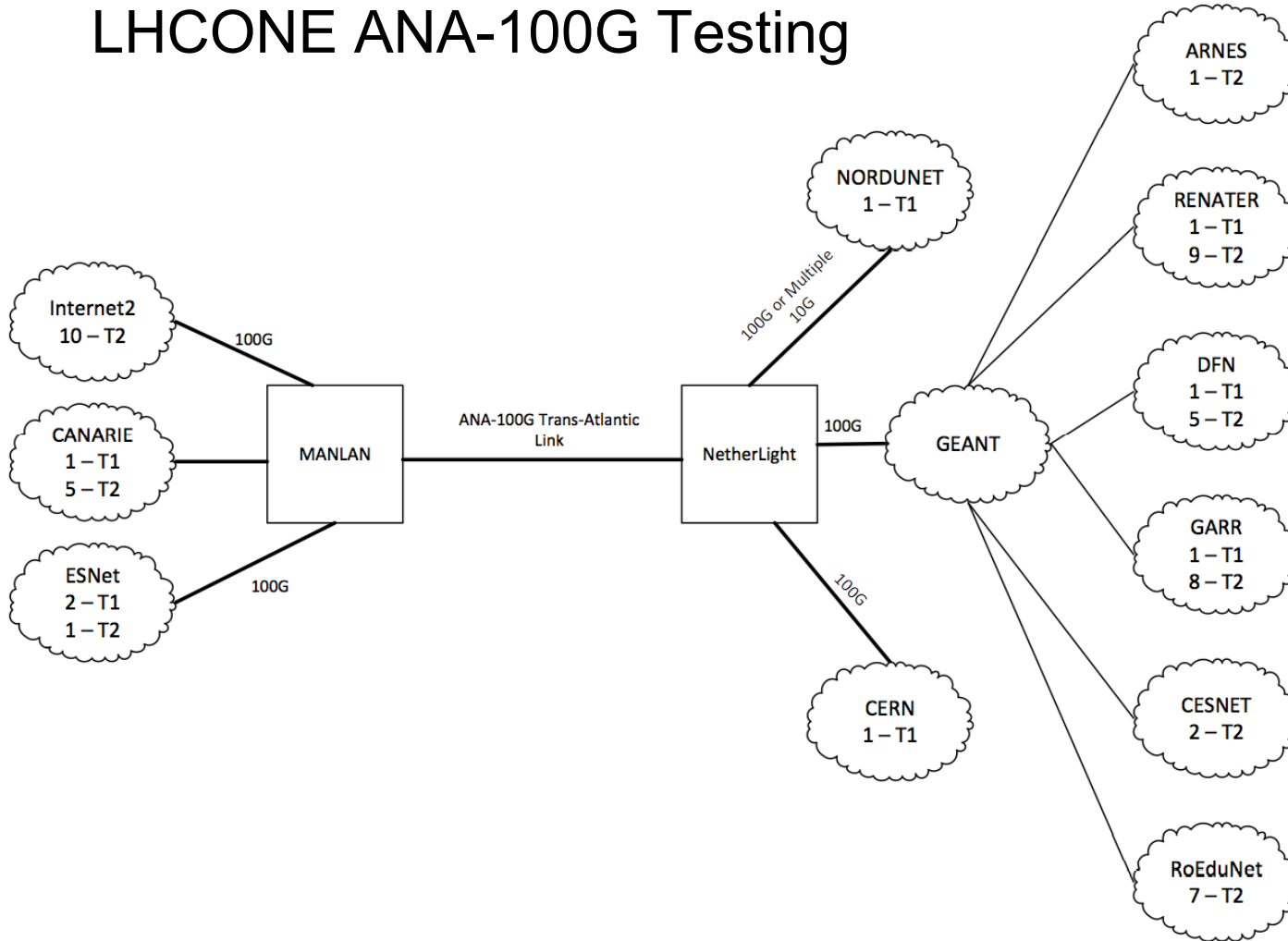
GEANT(20965)
ROEDUNET(2614)
ASGARR(137)
ARNES-NET(2107)
CZECH-ACAD-SCI(2852)

LHC1-RENATER(2091)
IN2P3(789)
CEA-SACLAY(777)

NORDUNET(2603)
NDGF(39590)

Bold - New Peers

LHCONE ANA-100G Testing



Other Topics



- [NTP Monlist Incident](#)
- [Junos In-Service Software Update](#)
- Research Efforts
 - 400G Ciena trials at SC
 - HNTES (next presentation)
 - LHCONE over ANA-100
- 2014 IPV6 Mandate
 - Requires native IPv6 connectivity for the organization. In other words, systems that are intended to connect to the Internet (including workstations) need to be able to access IPv6-only resources by the end of this fiscal year.
 - Understanding is that most labs are not currently covered by the mandate, but it is a good idea to know what would be required, and have a road map to get there.
 - Most labs will not want to be in a position of having to do an emergency crash program for deploying IPv6. But that need may develop.

NTP Incident



- Standard UDP reflection/amplification attack.
- Only an issue if ntp daemon is allowing administrative queries.
 - Admin queries not needed to provide time service (although they are useful for troubleshooting).
 - Time queries not good for amplification--only admin queries.
- Response from some corners of the R&E security community has been disappointing.
 - “Turn off or block NTP.”
 - NTP is a useful protocol for security, too!
 - Not everyone bothered to understand the actual protocol before trying to fashion a remedy. (E.g. NTP is not a client-server protocol, but was being treated as such, with bad results.)
- Standard UDP reflection/amplification attack. Translation: We really need to get everyone to implement BCP 38 unless we enjoy playing whack-a-mole on the Internet.

Juniper In-Service Software Upgrades



- We needed to upgrade all of our Juniper routers to address security concerns, and to support new OSCARS features late last Calendar year
- After extensive lab testing, we developed a detailed process and tried it.
- It didn't go well
 - ~ 20 % of our routers didn't support it due to hardware limitations including
 - Single RE
 - MX-FPC~
 - **~ 50% of the routers which experienced a total of about 7 seconds of packet loss distributed across several sub-second and 2-3 second outages**
 - ~ 20 % Failed
 - BGP Peers reset
 - ~ 10% Failed badly with various random behaviour
 - Processes core dumping & hanging
 - One router required manual intervention to physically pull out a corrupted RE.
 - Multiple Juniper tickets opened...