ESnet

ENERGY SCIENCES NETWORK

NRENs USA

Inder Monga

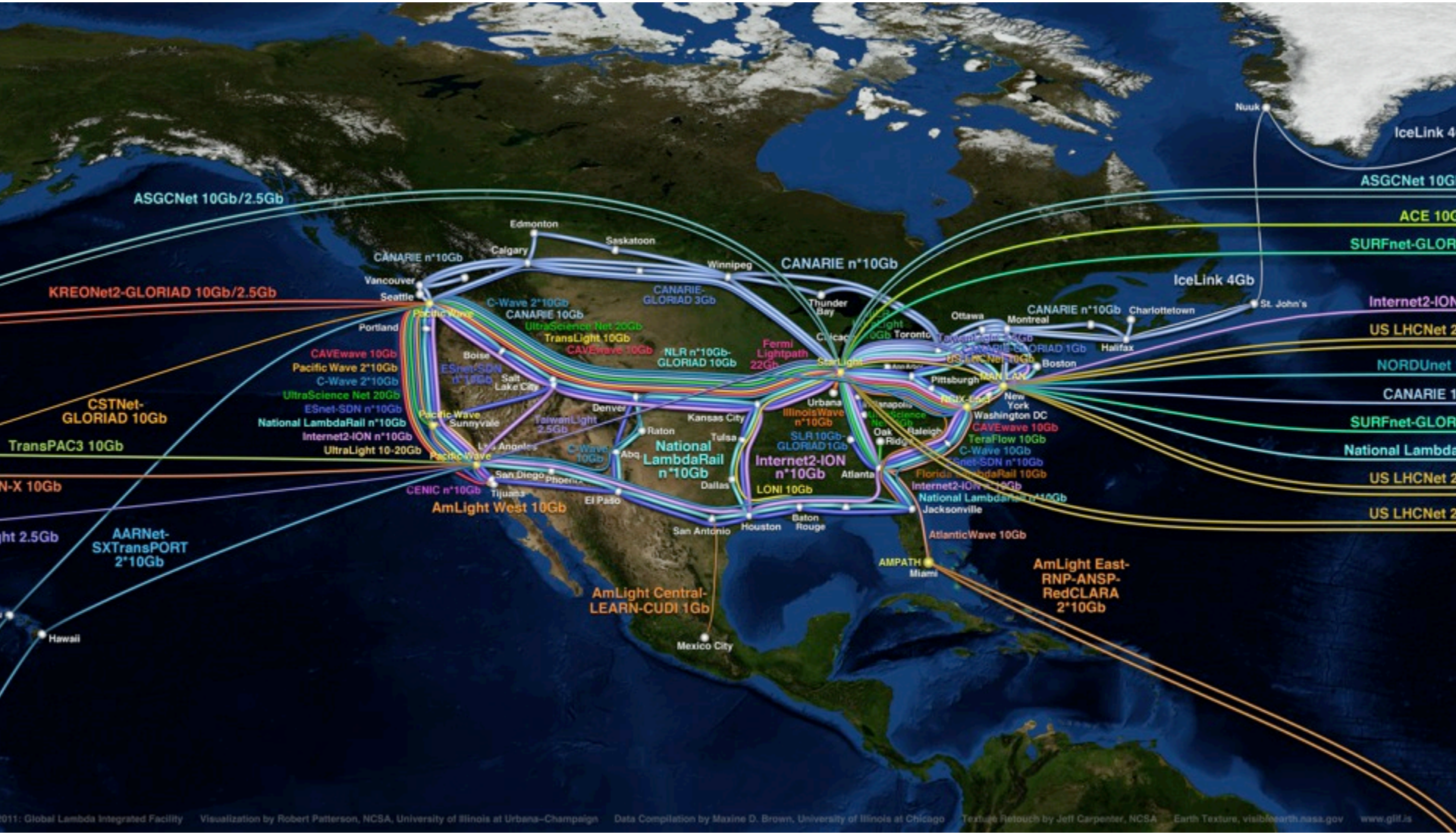
ESnet Chief Technologist

Lawrence Berkeley National Laboratory

SwitchON Workshop

Miami, Florida

January 9th, 2015


ASGCNet 10Gb/2.5Gb
KREONet2-GLORIAD 10Gb/2.5Gb
CSTNet-GLORIAD 10Gb
TransPAC3 10Gb
AARNet-SXTransPORT 2*10Gb
Hawaii
Edmonton
Calgary
Saskatoon
Winnipeg
Vancouver
Seattle
Portland
CANARIE n*10Gb
CANARIE 10Gb
CANARIE-GLORIAD 3Gb
C-Wave 2*10Gb
Pacific Wave
UltraScience Net 20Gb
TransLight 10Gb
CAVEwave 10Gb
Pacific Wave 2*10Gb
C-Wave 2*10Gb
ESnet-SDN n*10Gb
National LambdaRail n*10Gb
Internet2-ION n*10Gb
UltraLight 10-20Gb
Boise
Salt Lake City
Denver
Pacific Wave Sunnyvale
Los Angeles
San Diego
Phoenix
Tijuana
CENIC n*10Gb
AmLight West 10Gb
El Paso
Raton
Kansas City
Tulsa
Abq
C-Wave 10Gb
TaiwanLight 2.5Gb
National LambdaRail n*10Gb
Dallas
San Antonio
Houston
Baton Rouge
Mexico City
AmLight Central-LEARN-CUDI 1Gb
Thunder Bay
Fermi Lightpath 22Gb
NLR n*10Gb-GLORIAD 10Gb
StarLight
Chicago
Toronto
Urbana
IllinoisWave n*10Gb
SLR 10Gb-GLORIAD 1Gb
Internet2-ION n*10Gb
LONI 10Gb
Atlanta
Ottawa
Montreal
Pittsburgh
MAN LAN
New York
Boston
Halifax
Charlottetown
St. John's
Nuuk
IceLink 4Gb
IceLink 4
ASGCNet 10G
ACE 10G
SURFnet-GLOR
Internet2-ION
US LHCNet 2
NORDUnet
CANARIE 1
SURFnet-GLOR
National Lambda
US LHCNet 2
US LHCNet 2
Washington DC
CAVEwave 10Gb
TeraFlow 10Gb
C-Wave 10Gb
Jacksonville
AtlanticWave 10Gb
AMPATH
Miami
AmLight East-RNP-ANSP-RedCLARA 2*10Gb
2011: Global Lambda Integrated Facility
Visualization by Robert Patterson, NCSA, University of Illinois at Urbana–Champaign
Data Compilation by Maxine D. Brown, University of Illinois at Chicago
Texture Retouch by Jeff Carpenter, NCSA
Earth Texture, visibleearth.nasa.gov
www.glif.is

ESnet's Vision

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

ESnet

Fact and Figures

High-speed national network, optimized for DOE science missions:

- Connecting 40 labs, plants and facilities with >100 networks
- Managing entire path to Lab border (metro, inter-city)
- $32.6M in FY14, 42FTE
- older than commercial Internet, growing twice as fast

$62M ‘stimulus’ funding for 100G upgrade in 2011:

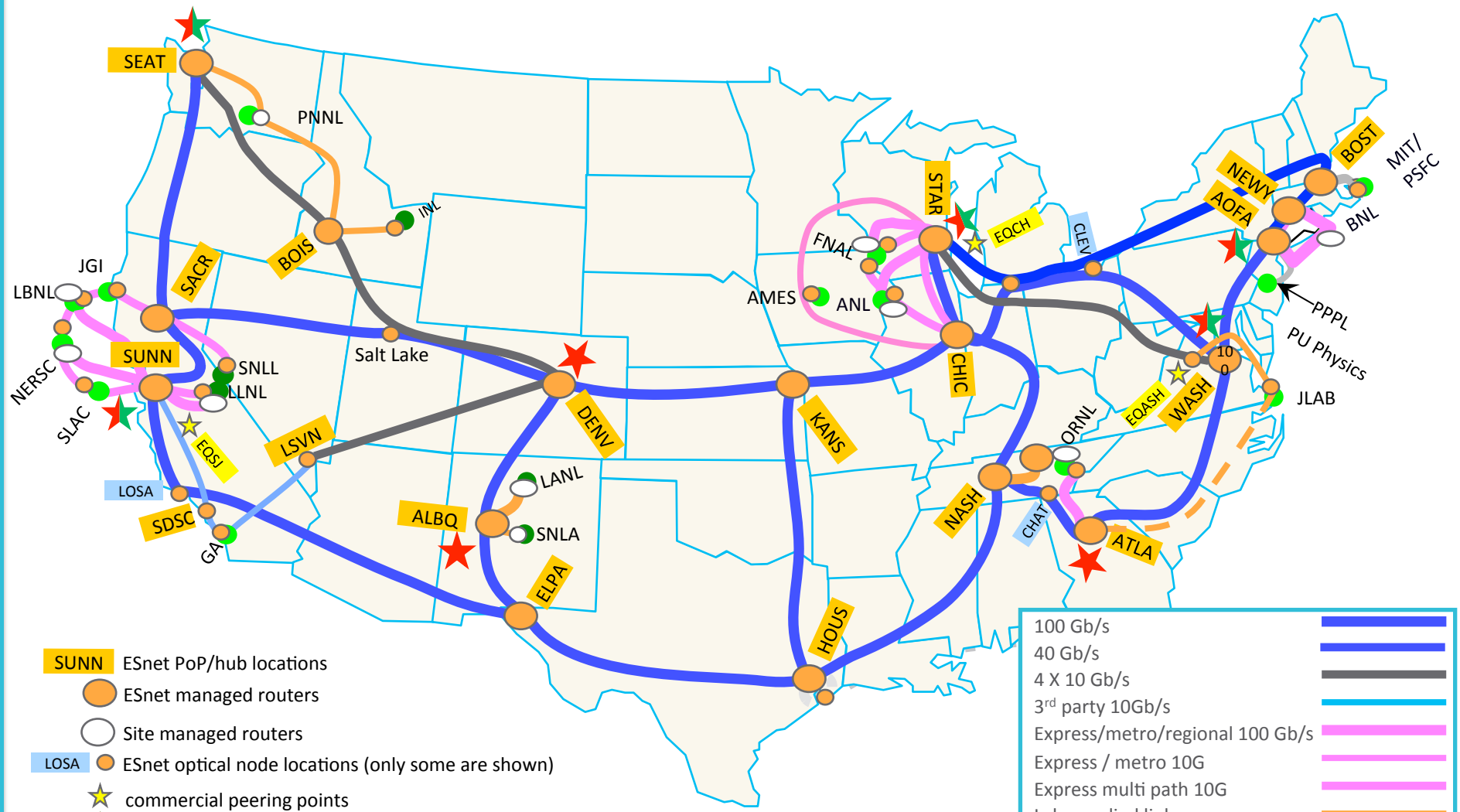
- new era of optical networking, abundant capacity
- fiber assets + access to spectrum shared with Internet2
- world’s first 100G network at continental scale

Culture of urgency:

- 4 awards in past 3 years
- R&D100 in FY13
- startup perspective

ESnet

ESnet September 2014



SUNN ESnet PoP/hub locations

ESnet managed routers

Site managed routers

LOSA ESnet optical node locations (only some are shown)

commercial peering points

R&E network peering locations – US (red) and international (green)

LBNL Major Office of Science (SC) sites

LLNL Major non-SC DOE sites

Geography is only representational

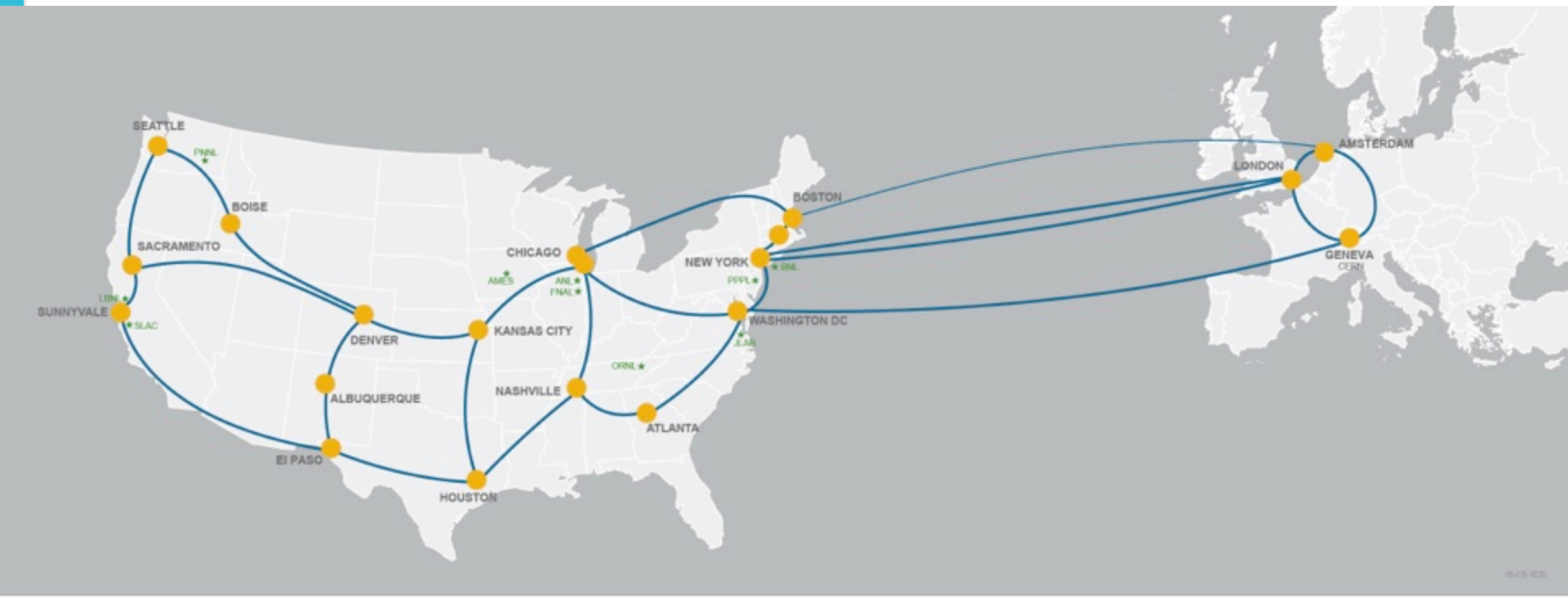
100 Gb/s	
40 Gb/s	
4 X 10 Gb/s	
3rd party 10Gb/s	
Express/metro/regional 100 Gb/s	
Express / metro 10G	
Express multi path 10G	
Lab supplied links	
Other links	
Tail circuits	

ESnet

100G testbed

SF Bay Area — SUNN — Chicago — STAR — New York — AOFA

ESnet expands to Europe (December 2014)



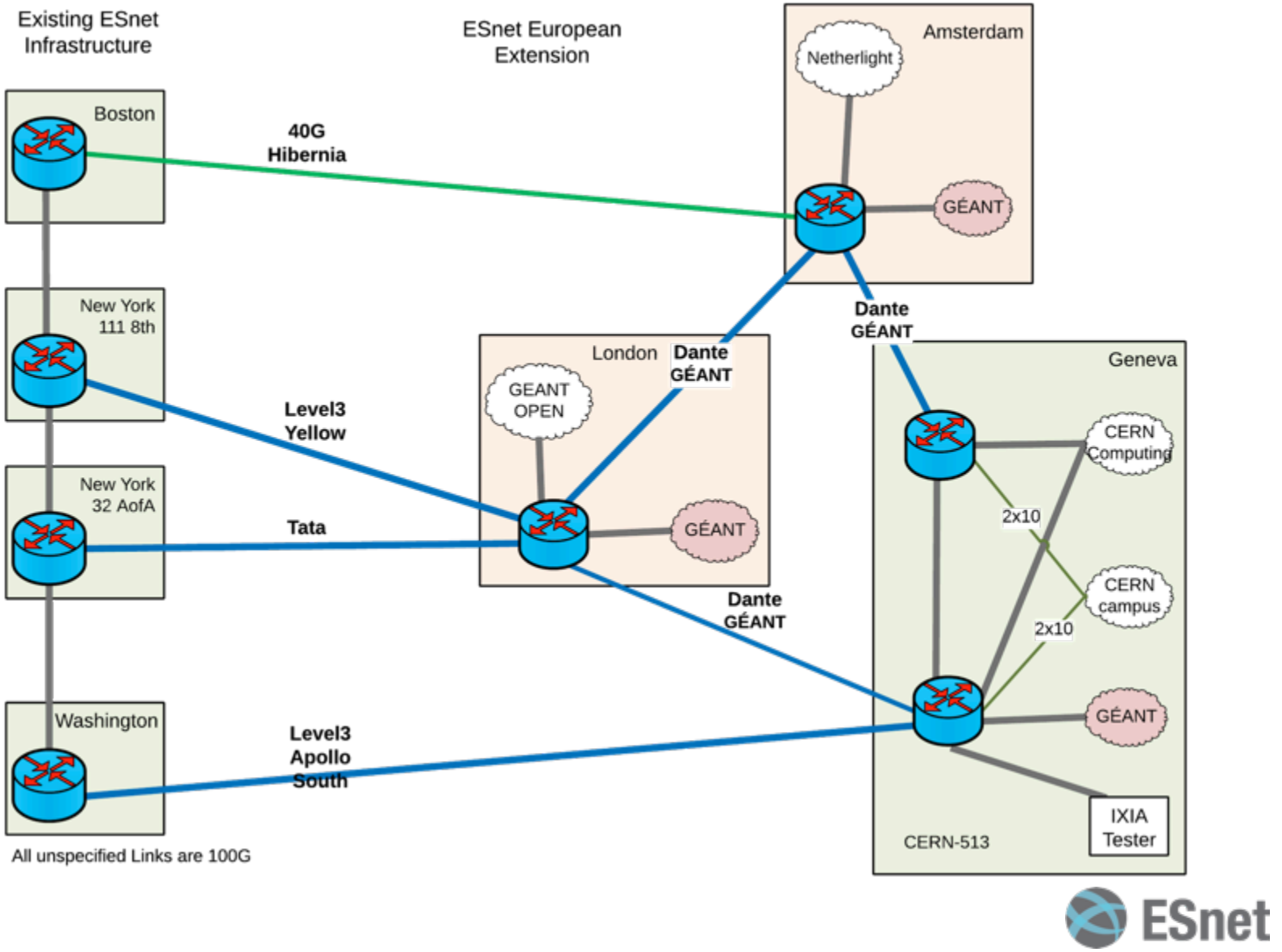
ESnet
ENERGY SCIENCES NETWORK

★ Department of Energy Office of Science National Labs

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)

ESnet

EEX Architecture



ESnet

Internet2 Network – Layer 1 Services

Internet2 Network – Layer 2 Services

Internet2 Network – Layer 3 Services

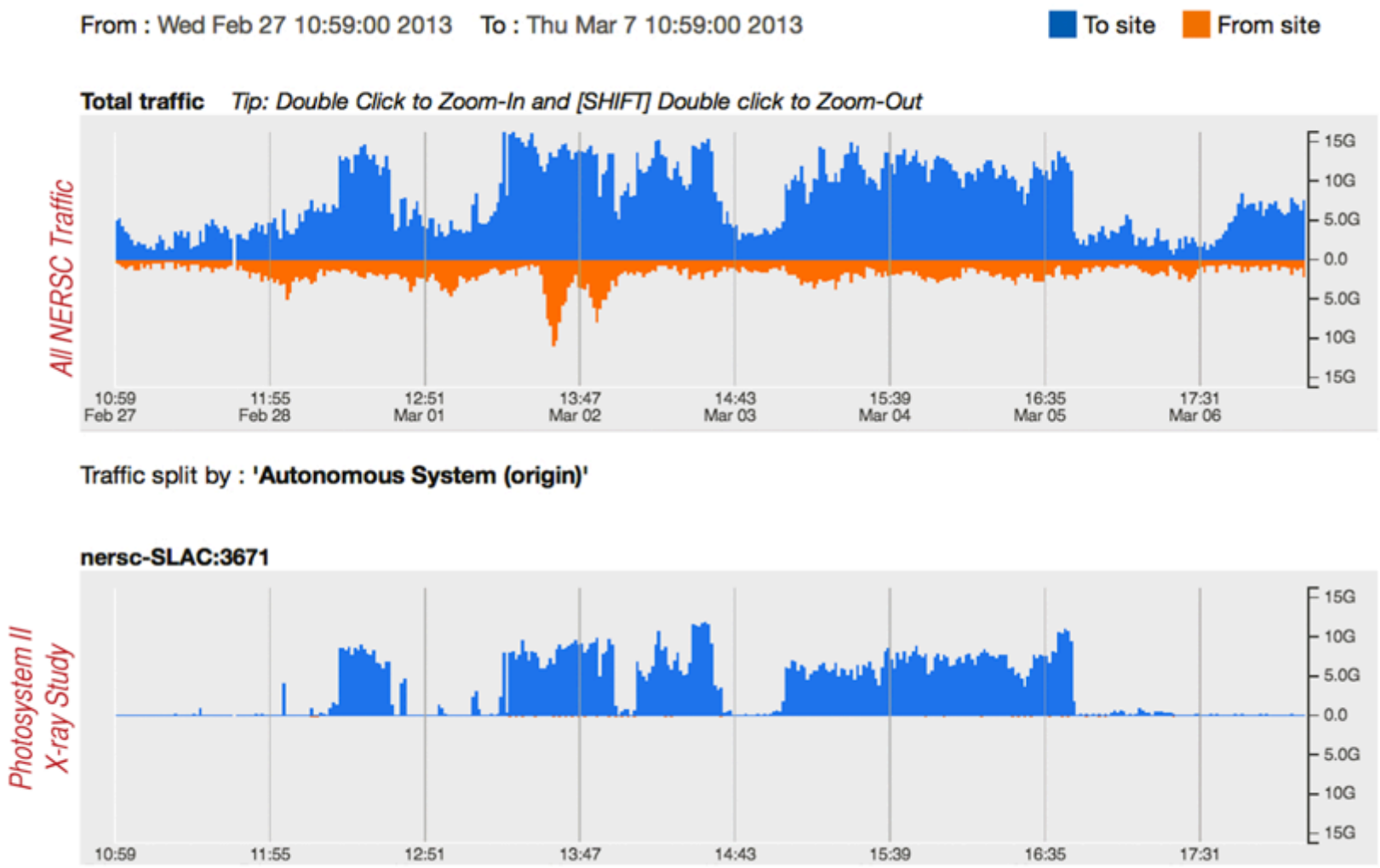
Science R&E Trends

First Emerging Trend in US: Super Facilities, Coupled by Networks

Experimental facilities are being transformed by new detectors, advanced mathematics, robotics, automation, advanced networks.

ESnet

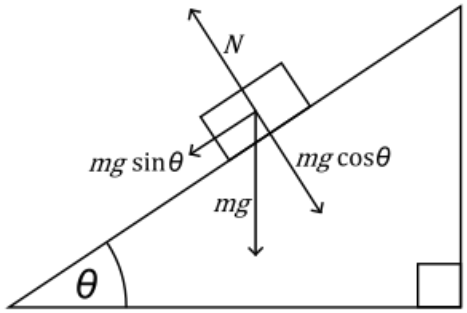
Data flow from single FEL detector *triples* network utilization for major HPC center



Second Emerging Trend in US: Science DMZ

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



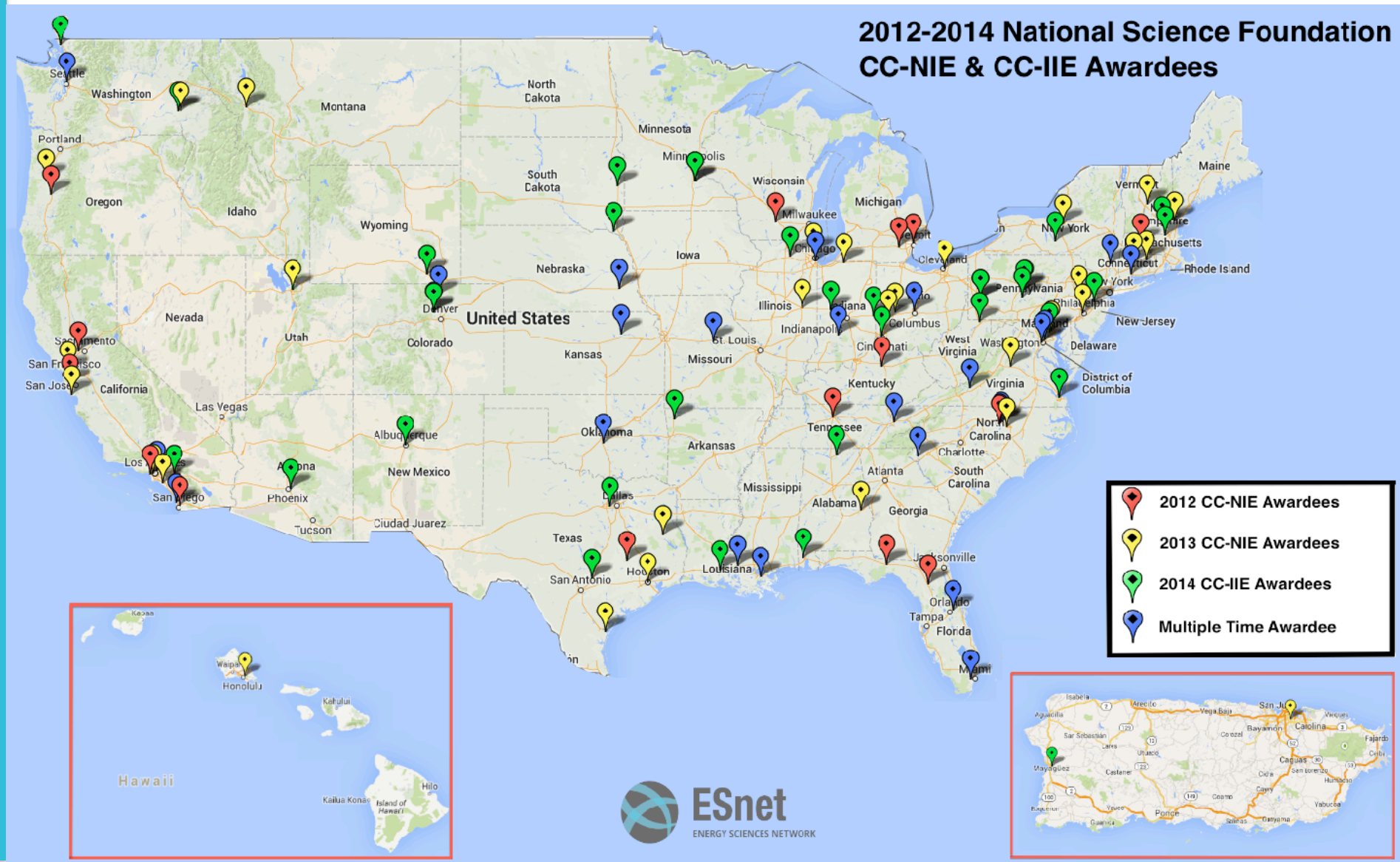
globus

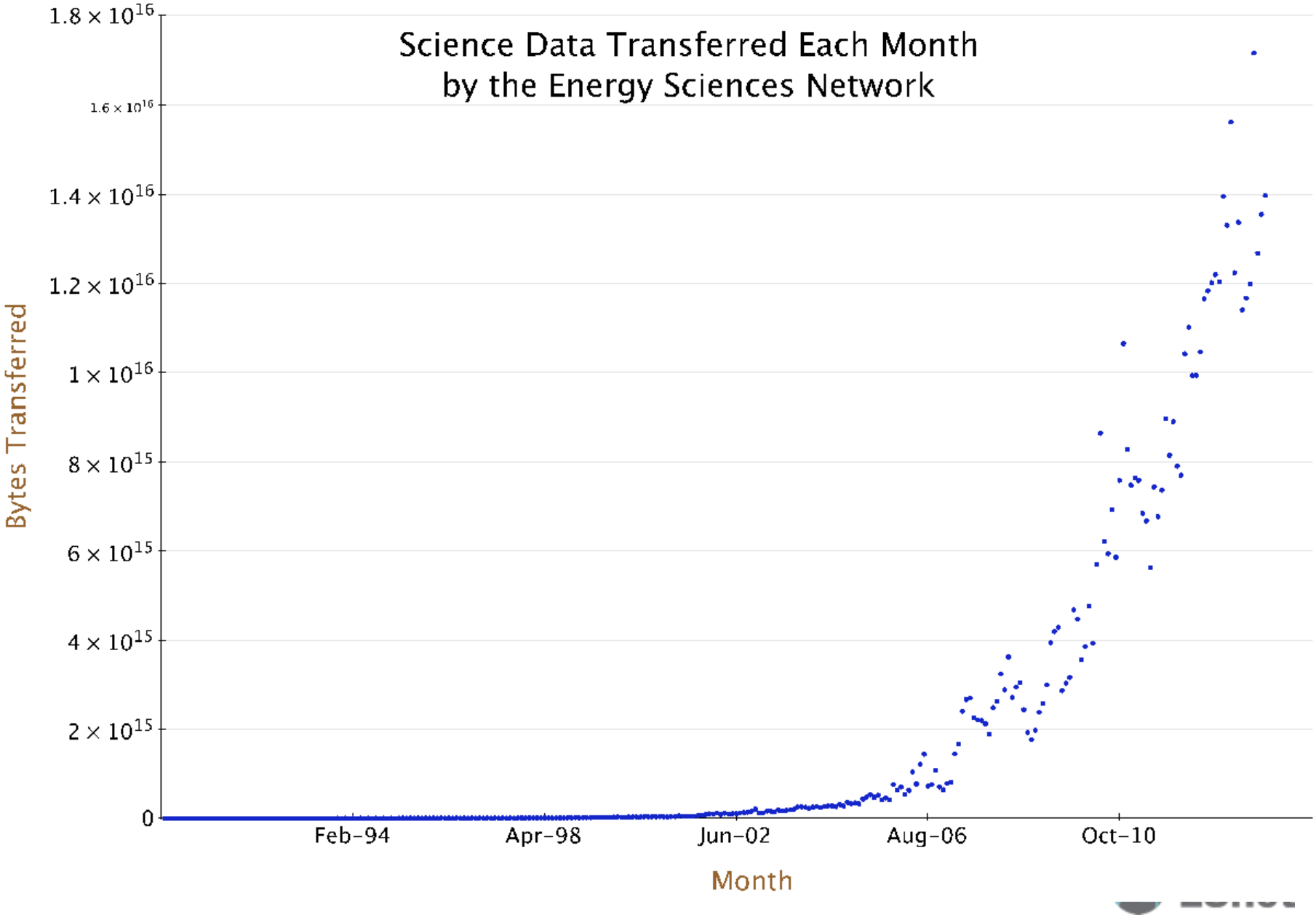
perfSONAR

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

ESnet

Map of NSF Grants to Promote Science DMZ




Science Data Transferred Each Month
by the Energy Sciences Network
Bytes Transferred
Month
1.8 × 10^16
1.6 × 10^16
1.4 × 10^16
1.2 × 10^16
1 × 10^16
8 × 10^15
6 × 10^15
4 × 10^15
2 × 10^15
0
Feb-94
Apr-98
Jun-02
Aug-06
Oct-10

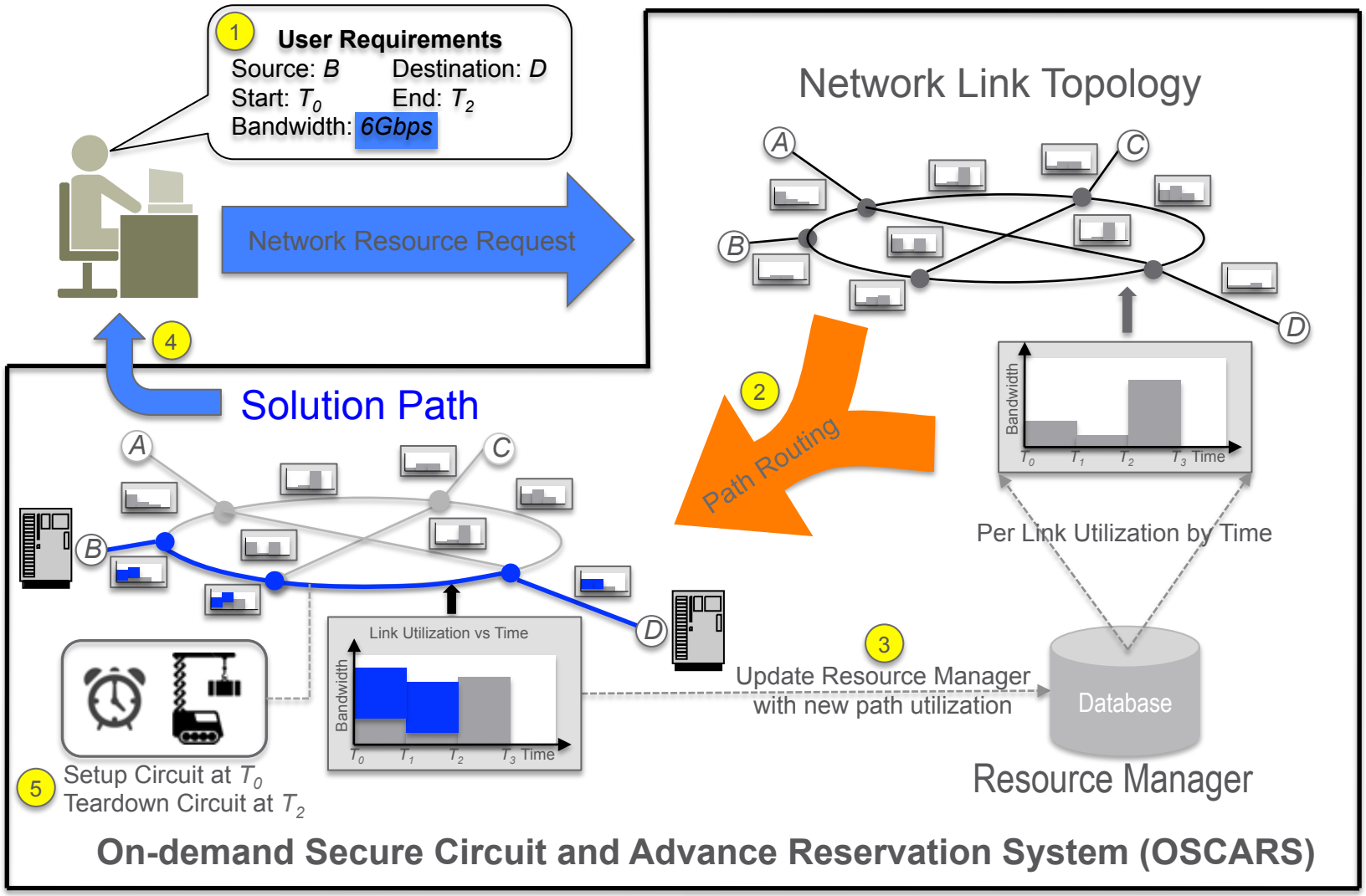
Current APIs

OSCARS Introduction

1. Multi-domain, Dynamic Circuit Service within ESnet
   - **Enforces QoS and Guaranteed Bandwidth**
   - Implements Inter-Domain Control Protocol (IDCP) and Network Services Interface (NSI) multi-domain protocols
   - GUI-based user-interface, and roles/responsibilities

2. Open Source project
   - 40+ R&E networks have OSCARS deployed in some form (production, prototype, tests)
   - In many cases, have to test their own NRM functions (ex. OESS)

ESnet


1
User Requirements
Source: B
Destination: D
Start: T_0
End: T_2
Bandwidth: 6Gbps
Network Resource Request
Network Link Topology
A
B
C
D
2
Path Routing
Bandwidth
T_0 T_1 T_2 T_3 Time
Per Link Utilization by Time
4
Solution Path
A
B
C
D
Link Utilization vs Time
Bandwidth
T_0 T_1 T_2 T_3 Time
3
Update Resource Manager
with new path utilization
Database
Resource Manager
5
Setup Circuit at T_0
Teardown Circuit at T_2
On-demand Secure Circuit and Advance Reservation System (OSCARS)

OSCARS Information

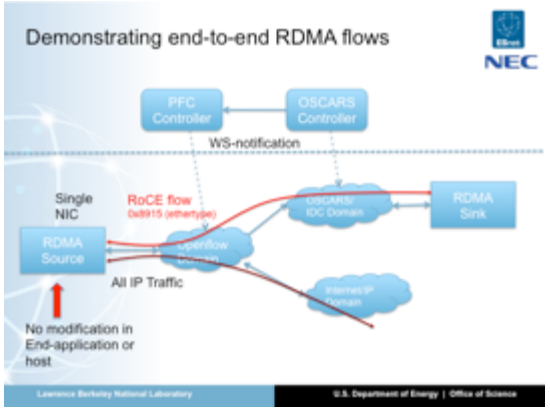
- OSCARS v0.6 Project (including NSI Bridge)
  - http://code.google.com/p/oscars-idc/
- OSCARS Mailing Lists
  - **oscars-announce** (https://listserv.es.net/mailman/listinfo/oscars-announce)
    - This list will be used to announce general information such as upcoming releases and roadmaps.
  - **oscars-dev** (https://listserv.es.net/mailman/listinfo/oscars-dev)
    - This list will be used to discuss code development issues, track bugs, and report code commits.
  - **oscars-users** (https://listserv.es.net/mailman/listinfo/oscars-users)
    - This list is for the user community to share deployment and usage experiences, as well as submitting enhancement requests and reporting bugs.
    - This mailing list is in the process of being renamed to **oscars-support**

ESnet

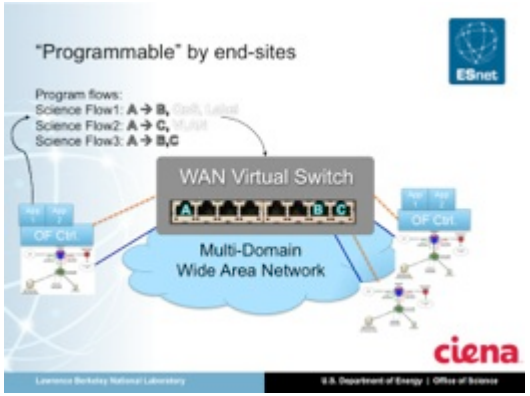
SDN Testbeds

Last three years of SDN investigations

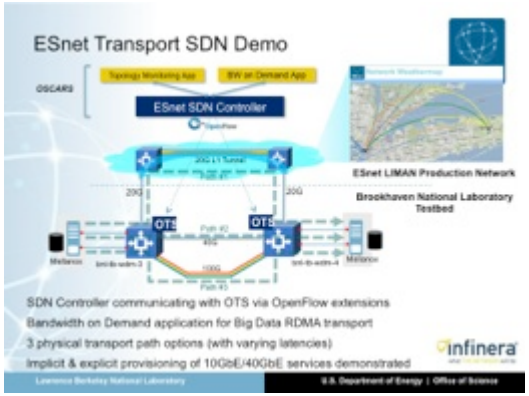
ONS 2011, SC 2011

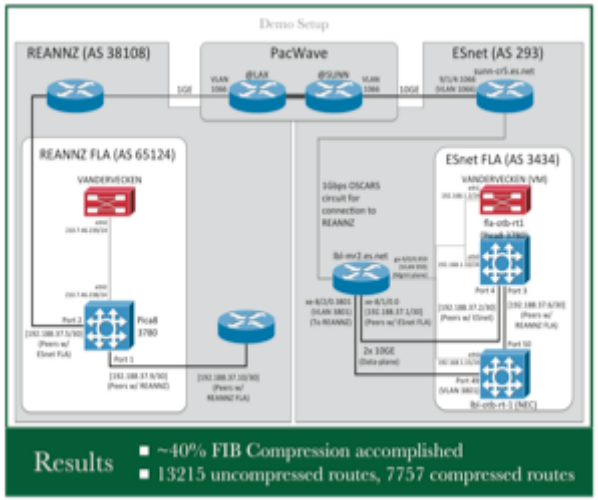


SC 2012



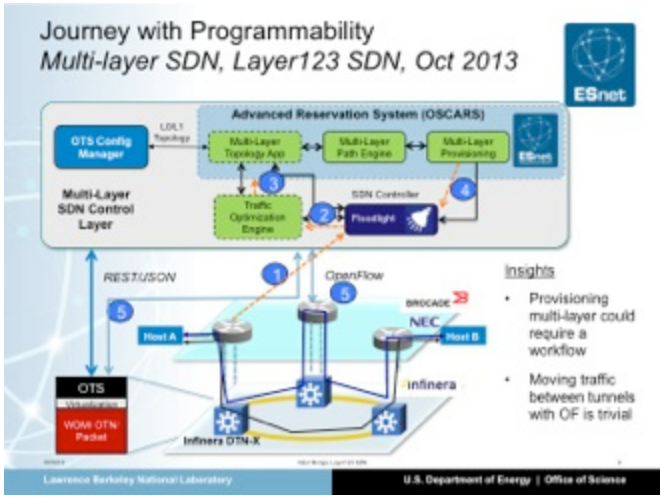
December 2012





Treehouse w/REANNZ, Google

ONS March 2013



Layer 123 SDN, Nov 2013

SDN Workshop
'Operationalizing SDN'
December 2013

ESnet

Building a scalable SDN WAN Testbed

1. Prototype for a potential ESnet6 design

2. Multi-domain federation for the science complex

3. International SDN collaboration

ESnet

Building a scalable SDN WAN Testbed

- Requirements
  - Openflow 1.3 support [Cannot do Guaranteed Bandwidth or QoS with OF 1.0]
  - L2/L3 transport with QoS, MPLS
  - Support for Federation
  - Support internet-scale routing and forwarding table
  - Diagnostic and Troubleshooting capabilities

ESnet

Corsa Product Introduction

DP6440
4 x 100G CFP2 Ports
24x 10G SFP+ Ports

DP6430
2 x 100G CFP2 Ports
24x 10G SFP+ Ports

100G

10G

DP6420
48 x 10G SFP+ Ports
4 x 40G QSFP+ Ports

DP6410
24 x 10G SFP+ Ports

Open vSwitch + intel

OpenFlow 1.3 / 1.4 / 1.5+

- Fully programmable OpenFlow data plane with multiple pipelines
- Embedded OVS based OpenFlow 1.3+ / ovsdb SDN controller interface
- Ultra fast flow setup and control plane interfaces
- Full line rate, OpenFlow ports with multiple flow tables on all ports.

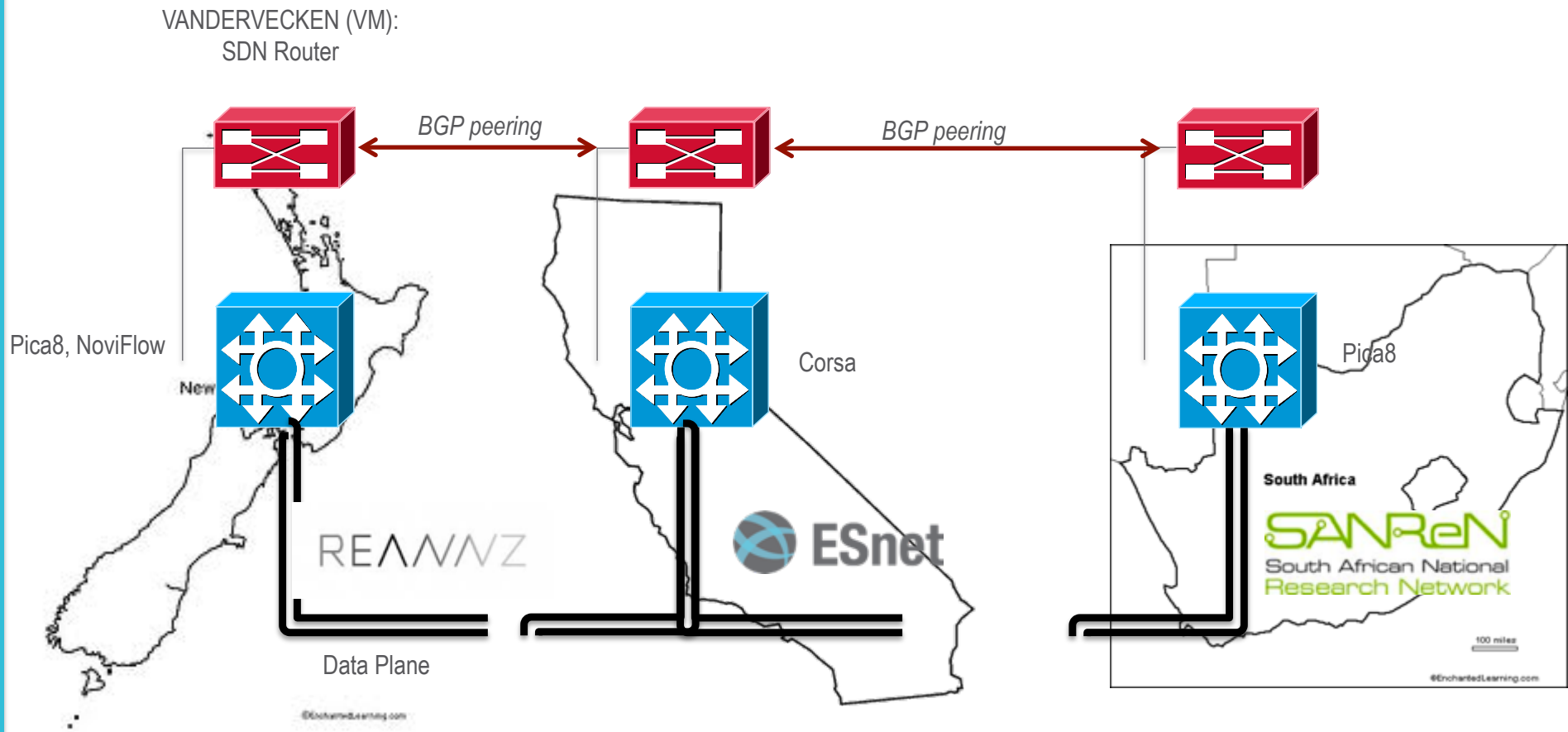
ESnet

Fills in 5 Key Features Missing From current whitebox OpenFlow Switches

1. Multiple match/action tables
2. Millions of flow entries
3. Large scale packet buffers
4. 100-Gigabit ports with full OpenFlow 1.3
5. Extremely fast flow modifications per second 10k-1M entries per second.
   (as opposed to 100 flow mods / sec)

ESnet

Routing with SDN
ESnet SDX, Multi-continent SDN BGP peering



With help from -

SDN BGP Gateway Pipeline: Under the Hood

- A flexible structure with millions of flow entries through multiple tables
- Performs at-scale SDN BGP gateway function to apply policies, and move data quickly and reliably.

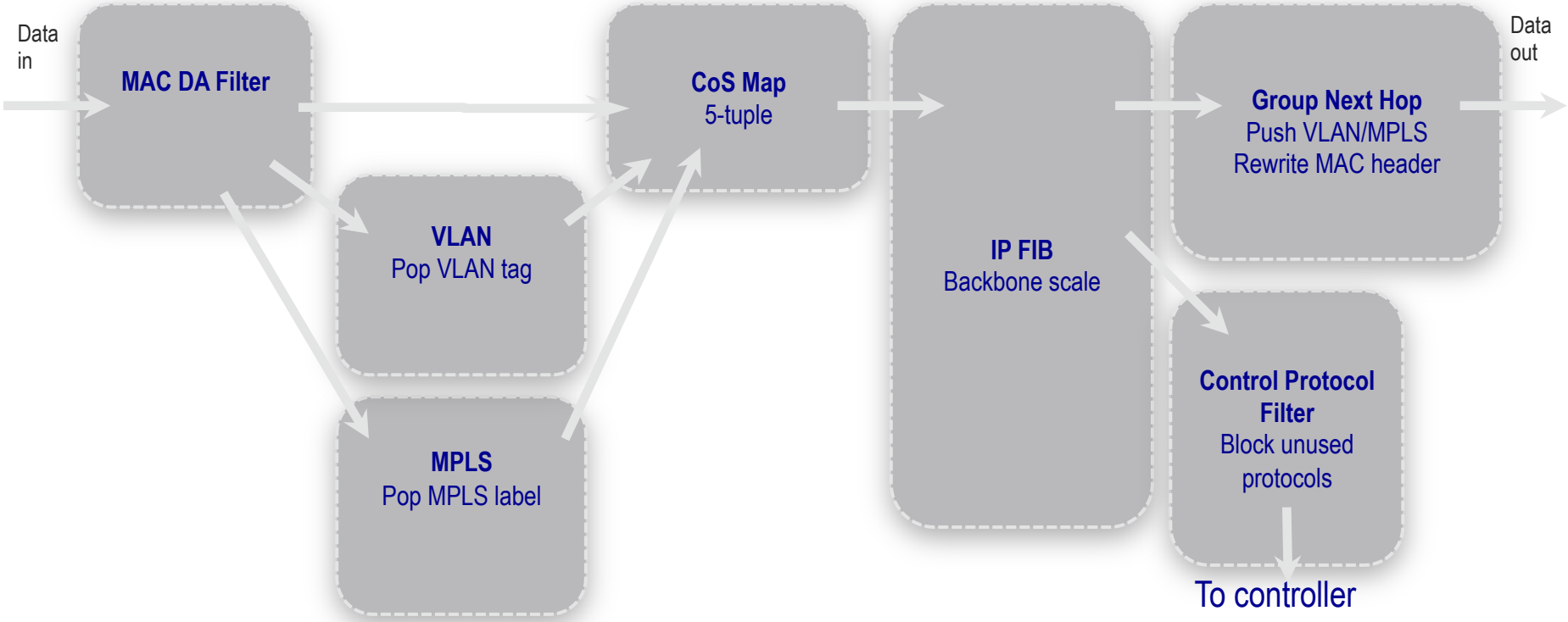


Table definitions shown are for this demo.
Same hardware platform can assume any pipeline structure.

CORSA

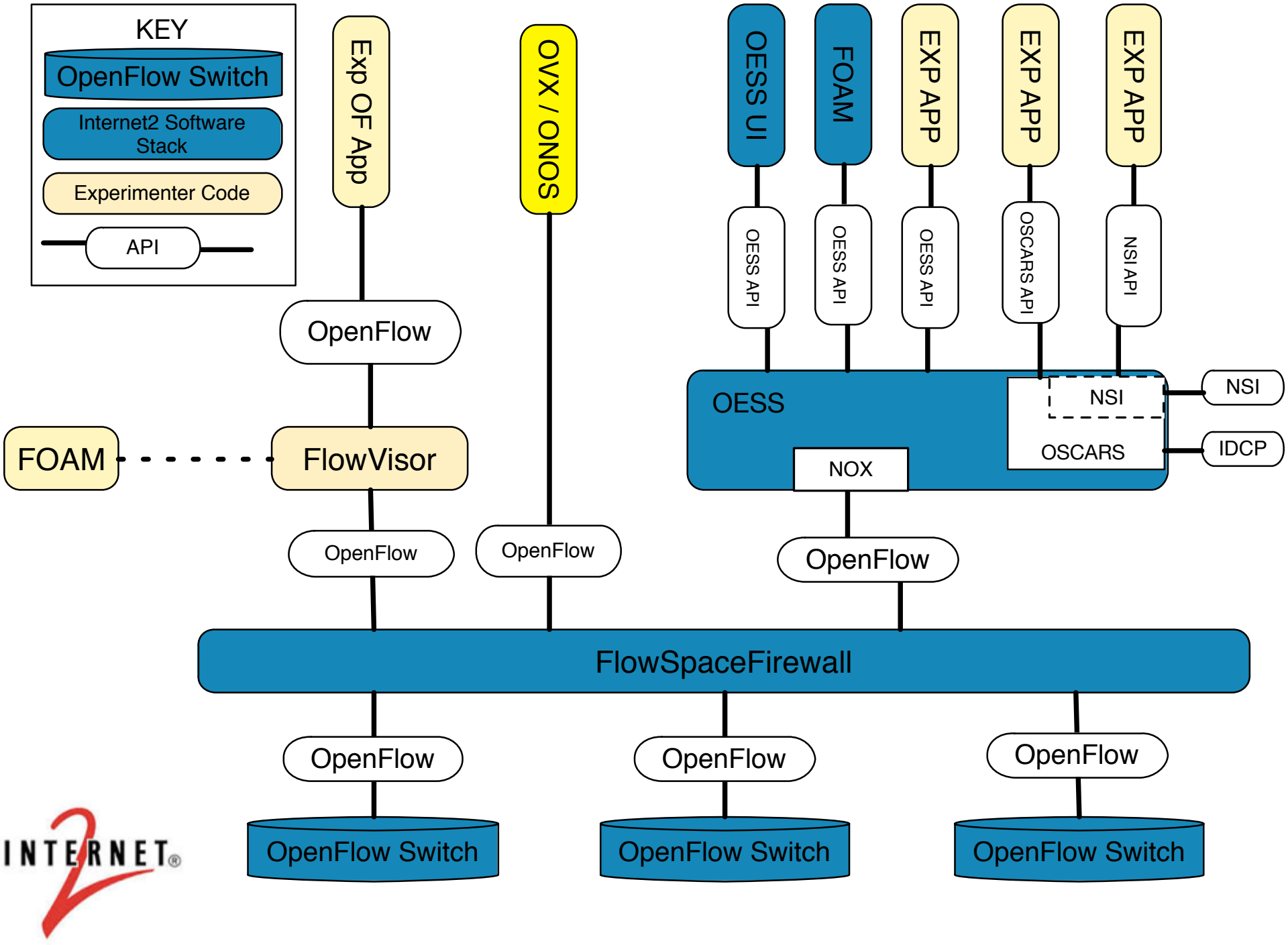
ESnet

Locations for new SDN testbed
under discussion/planning

ESnet
ENERGY SCIENCES NETWORK

★ Office of Science National Labs

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	Stanford Linear Accelerator Center (Menlo Park, CA)


KEY
OpenFlow Switch
Internet2 Software Stack
Experimenter Code
API
Exp OF App
OVX / ONOS
OESS UI
FOAM
EXP APP
EXP APP
EXP APP
OESS API
OESS API
OESS API
OSCARS API
NSI API
OpenFlow
OESS
NSI
OSCARS
NSI
IDCP
FOAM
FlowVisor
NOX
OpenFlow
OpenFlow
OpenFlow
FlowSpaceFirewall
OpenFlow
OpenFlow
OpenFlow
OpenFlow Switch
OpenFlow Switch
OpenFlow Switch
INTERNET2

Thank you!
Questions? imonga@es.net