



ESnet

ENERGY SCIENCES NETWORK

ESnet Update: with a focus on R&D

Inder Monga

Deputy of Technology, SND

CTO, ESnet

LHCONE Meeting, Berkeley

June 1, 2015



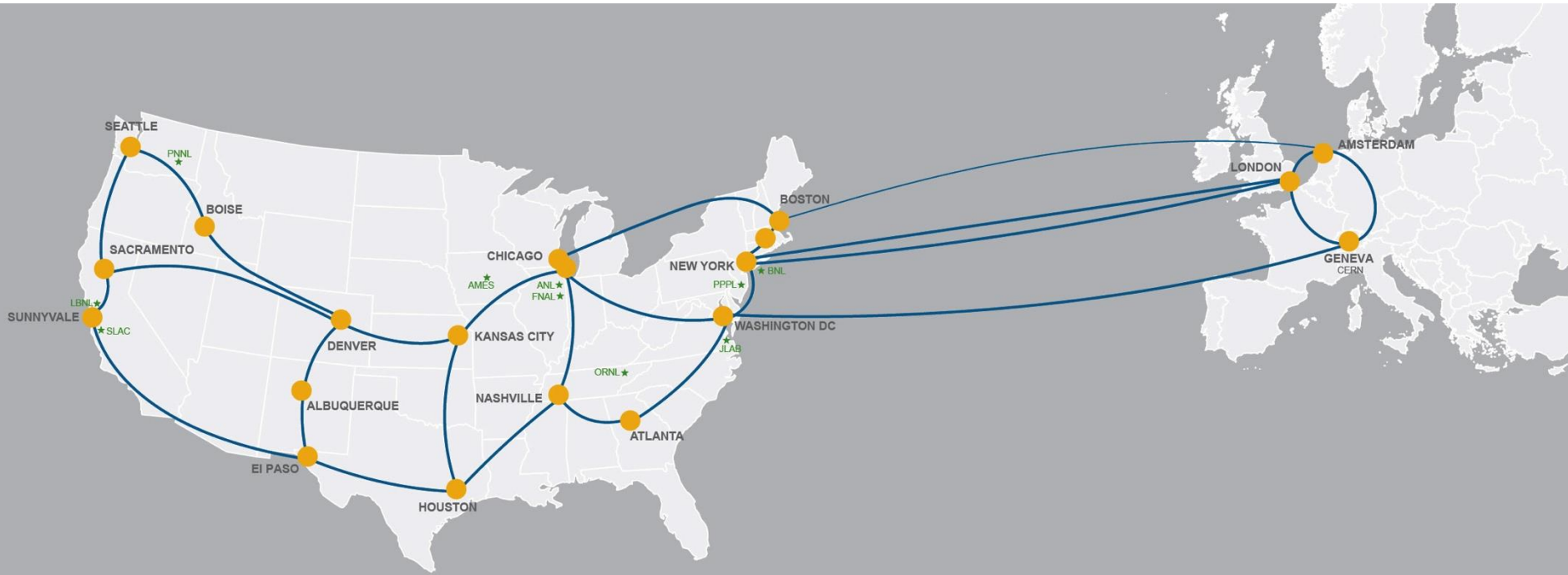
U.S. DEPARTMENT OF
ENERGY
Office of Science



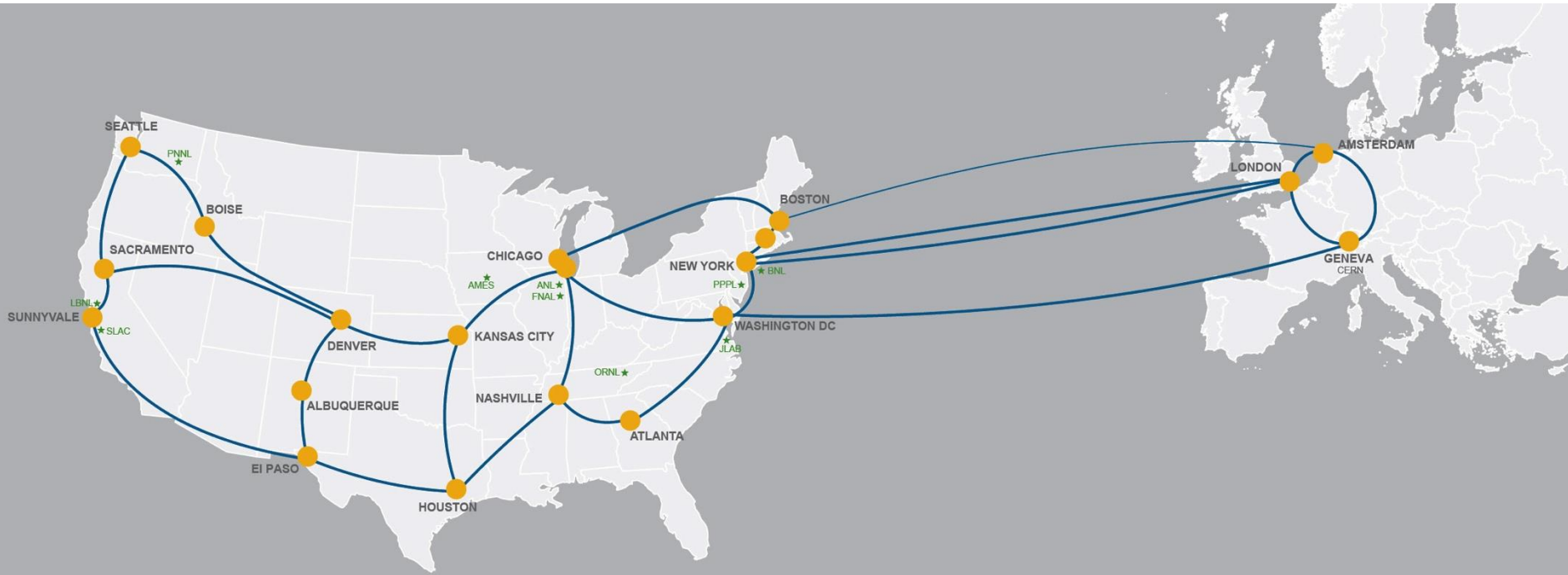
ESnet Background

1. ESnet is a special-purpose **mission network**, funded by the US Congress to support scientific goals of the Department of Energy.
2. We see networking as a means to an end: **scientific productivity**.
3. We aim to create a world in which **discovery is unconstrained by geography**.

This is not an ISP.

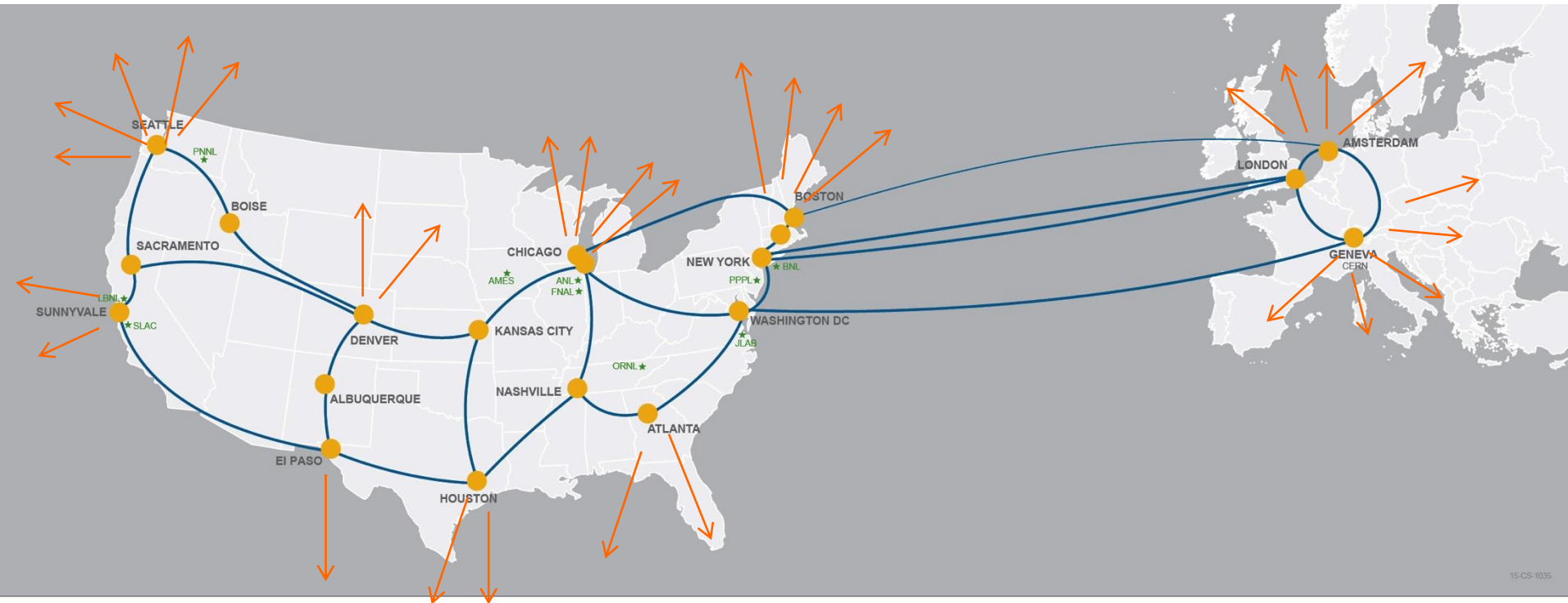


It's a DOE user facility designed to overcome the constraints of geography.



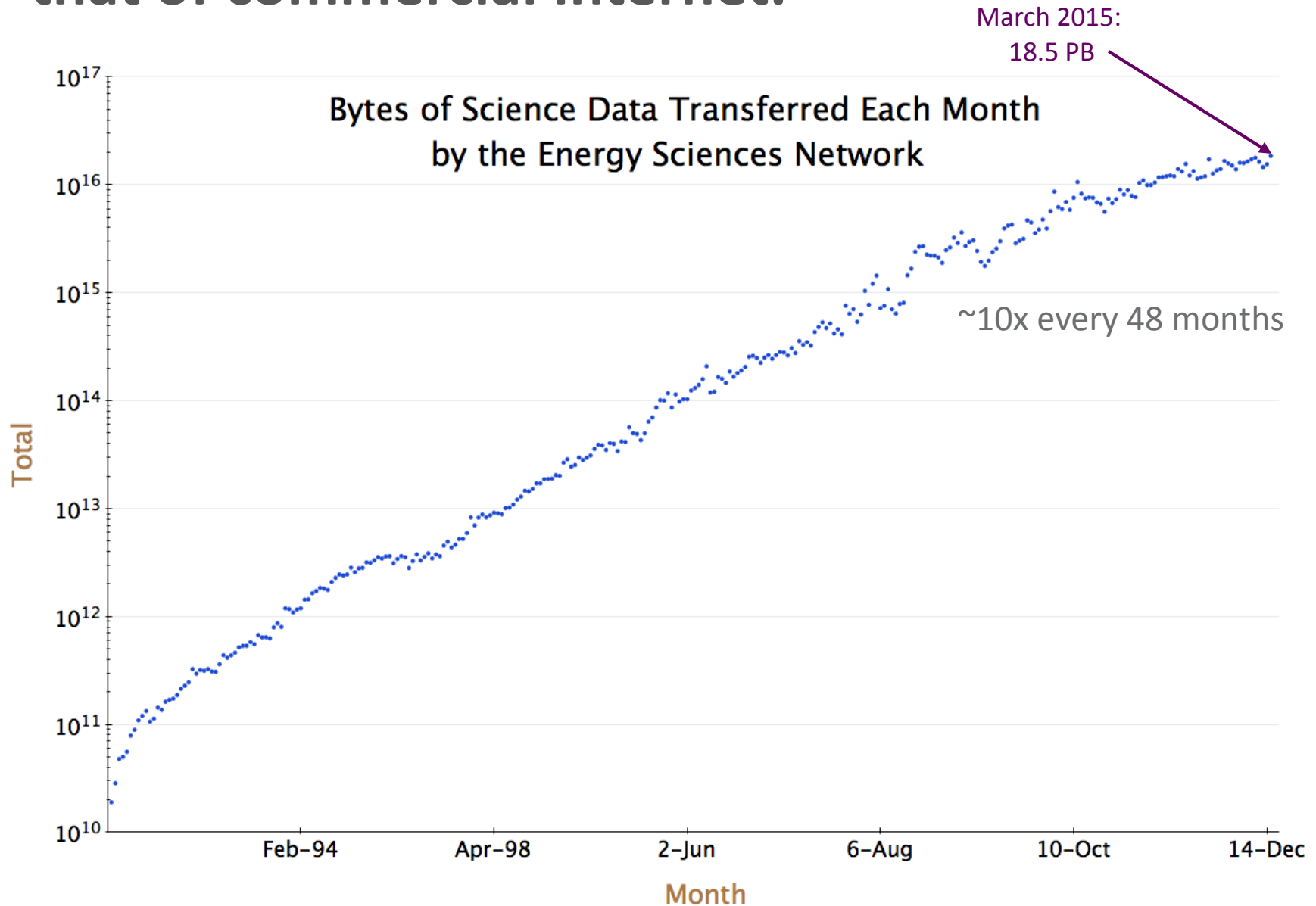
We do this by offering unique capabilities, and optimizing the facility for data acquisition, data placement, data sharing, data mobility.

80% of ESnet traffic originates or terminates outside the DOE complex.



ESnet's success is **necessary but not sufficient** for DOE, because networks share fate. SC invests nearly \$1B/year in university research: **campus networks matter.**

Exponential growth rate slowing, but still double that of commercial Internet.



ESnet Vision and Goals

(in addition to world-class operations...)

vision:

Discovery is unconstrained by geography.

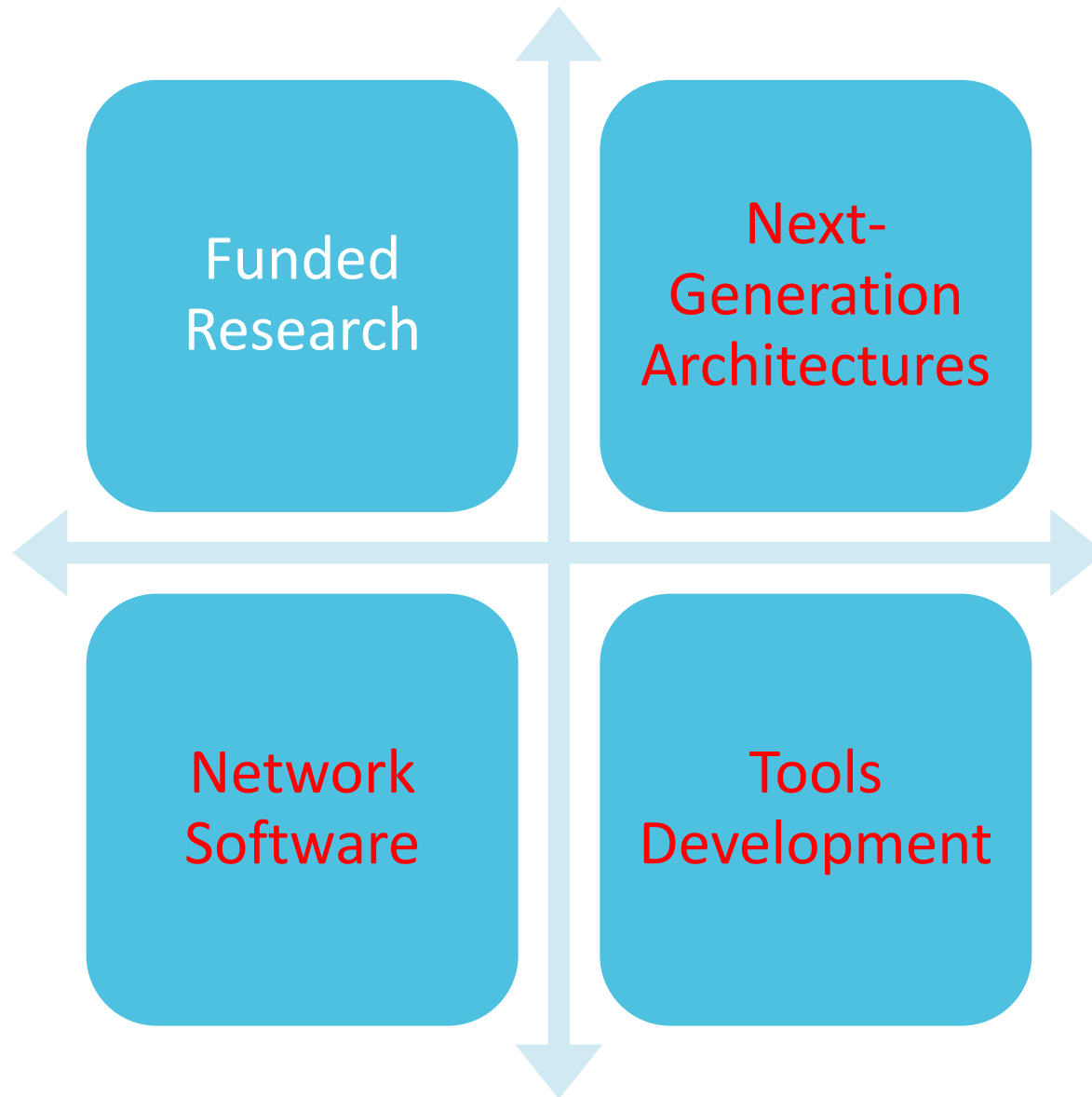
goals:

1. Improve networking practices globally.

2. Provide information and tools for optimal network use.

3. Pioneer architectures, protocols, applications.

ESnet Research and Software Development Portfolio



ESnet adopts Github

- Commitment to make it easy for the community to discover and collaborate on software
- All ESnet public repositories are on Github
 - <https://github.com/esnet>
- Documentation for the projects
 - <http://software.es.net/>

Network Software



perfSONAR

- Re-invigorated collaboration with Internet2, Indiana University and GEANT is resulting in consistent and operational software package
- perfSONAR 3.4 release full of great feature enhancements, notably, REST interface, security enhancements and new graphs
- perfSONAR 3.5 planned for June 2015
 - Highlights include:
 - GUI modernization
 - Central Management and Node Auto-configuration
 - Support for 1 inexpensive node platform (~\$100 – \$200)
 - perfSONAR Security Audit Completion
- 1300+ *public* perfSONAR hosts in 30 countries as adoption continues to rise
- New perfsonar.net website and user-manual both written from scratch

Who is running perfSONAR?

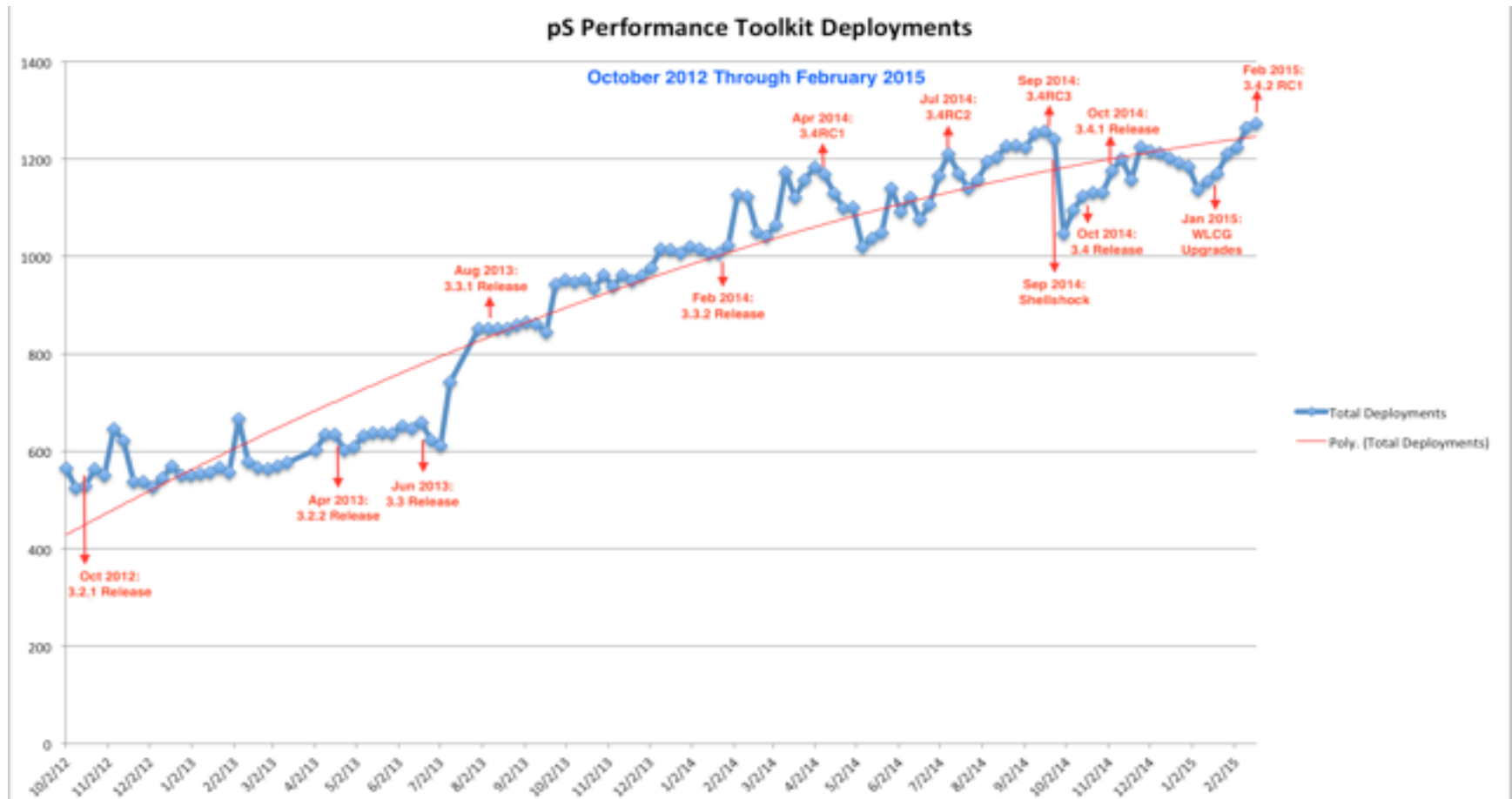
- Currently over 1300 deployments world-wide



<http://stats.es.net/ServicesDirectory/>



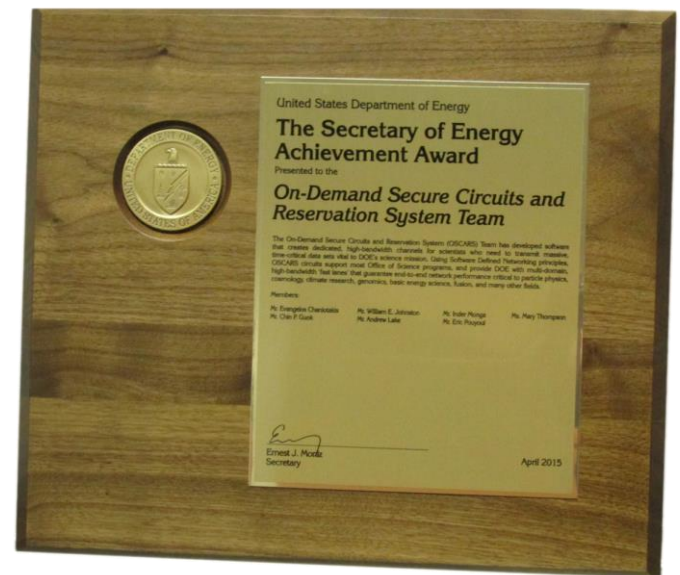
perfSONAR Growth



ESnet Public DTNs and perfSONAR hosts

- Upgraded public test DTNs at LBL, ANL, and BNL, and with a new DTN at CERN
 - Anonymous GridFTP hosts, Globus Online Endpoints
 - Capable of > 9Gbps from disk, single file read performance
 - <http://fasterdata.es.net/performance-testing/DTNs/>
- 40G perfSONAR host deployed in Boston
 - First 40G perfSONAR node on an R&E backbone network
- Beginning hardware refresh of the > 85 perfSONAR hosts we have deployed
 - Many of the current perfSONAR hosts are > 5 years old
 - New design combines bandwidth and latency testing into a single host, each on a separate NIC
 - Will reduce the number of hosts to manage by 50%

OSCARS



- OSCARS contributions to networking technology and science continue to get recognized
 - 2014 Secretary’s Honor Award from US Department of Energy
- Focus of OSCARS software development is to support LHCONE’s multi-domain virtual circuit experimentation
 - by implementing NSI
 - building community consensus
 - and helping with NSI deployments/ interoperability testing.

OSCARS/NSI Latest Updates

- NSI Bridge
 - Running in ESnet, WIX, MANLAN, RNP, UMich, CalTech
 - In process of being deployed at Internet2 AL2S
- NSI – Document Distribution Service
 - Topology translator from NM to NML up and running
 - Being used by ESnet's OSCARS
 - WIX, MANLAN, UMich and CalTech also using this software
- SURFnet/ESnet NSI Aggregator
 - Running in ESnet, SURFnet, iCAIR and RNP

Focus on Operational use of OSCARS

- New capabilities include:
 - New QoS profile to handle best-effort circuits
 - User can now specify guaranteed bandwidth or best-effort circuit
 - Policing enhancements for hard or soft QoS
 - Over-subscribed traffic dropped (Hard QoS), or marked for discard and allowed to burst (Soft QoS)
 - Circuit Protection
 - Enhanced for EEX, protection circuits are best-effort circuits

iPerf3

- Maintained actively by ESnet, and helping answer community questions
 - <https://fasterdata.es.net/performance-testing/network-troubleshooting-tools/iperf-and-iperf3/>
 - <https://github.com/esnet/iperf>
- Good uptake by the community
 - Docker build for iperf3
 - <https://github.com/nerdalert/iperf3>

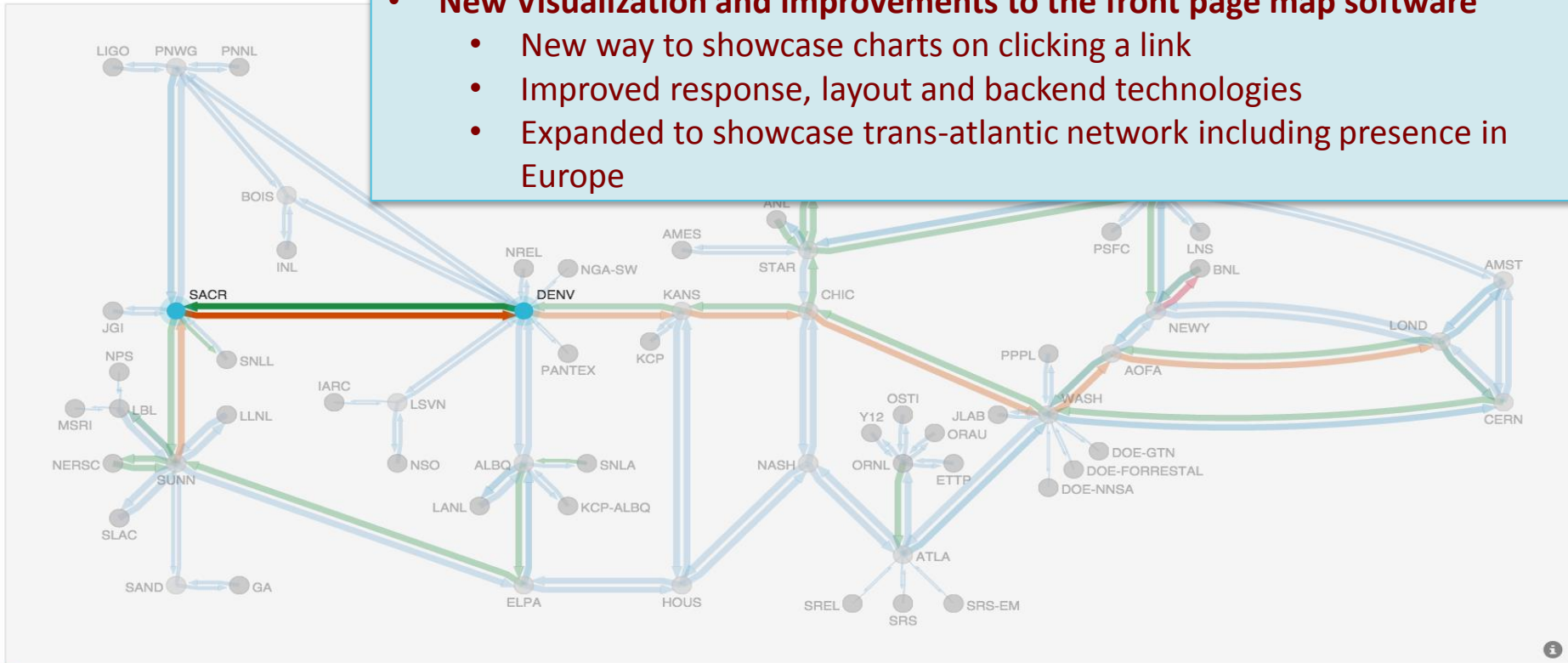
Tools Development

myESnet Portal - <https://my.es.net/>

- Vision: Transform the way we visualize network information
- Lots of software work being developed, not all of it visible in the portal
- Based on 'state of the art' in industry
 - Twitter Bootstrap
 - Node.js
 - React and Flux
- Open-Source requests
 - Code re-organized to make it modular and open-source
 - Separating ESnet specific data from the portal code
 - Plan to open-source modules by end of August



- **New Visualization and improvements to the front page map software**
 - New way to showcase charts on clicking a link
 - Improved response, layout and backend technologies
 - Expanded to showcase trans-atlantic network including presence in Europe



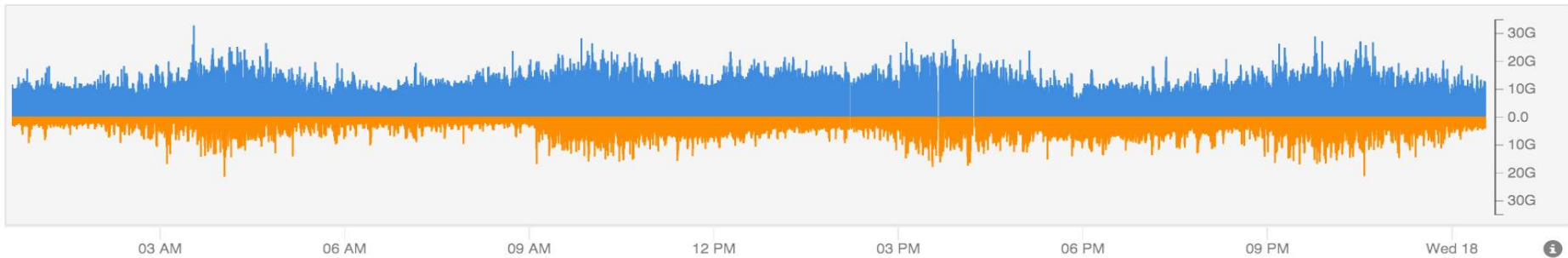
DENV ⇌ SACR
100G

Interfaces

denv-cr5 to_sacr-cr5_ip-a



■ DENV → SACR ■ SACR → DENV



Visualization of Demos at SC14

Demos

Summary

LLNL

MDTM

Caltech

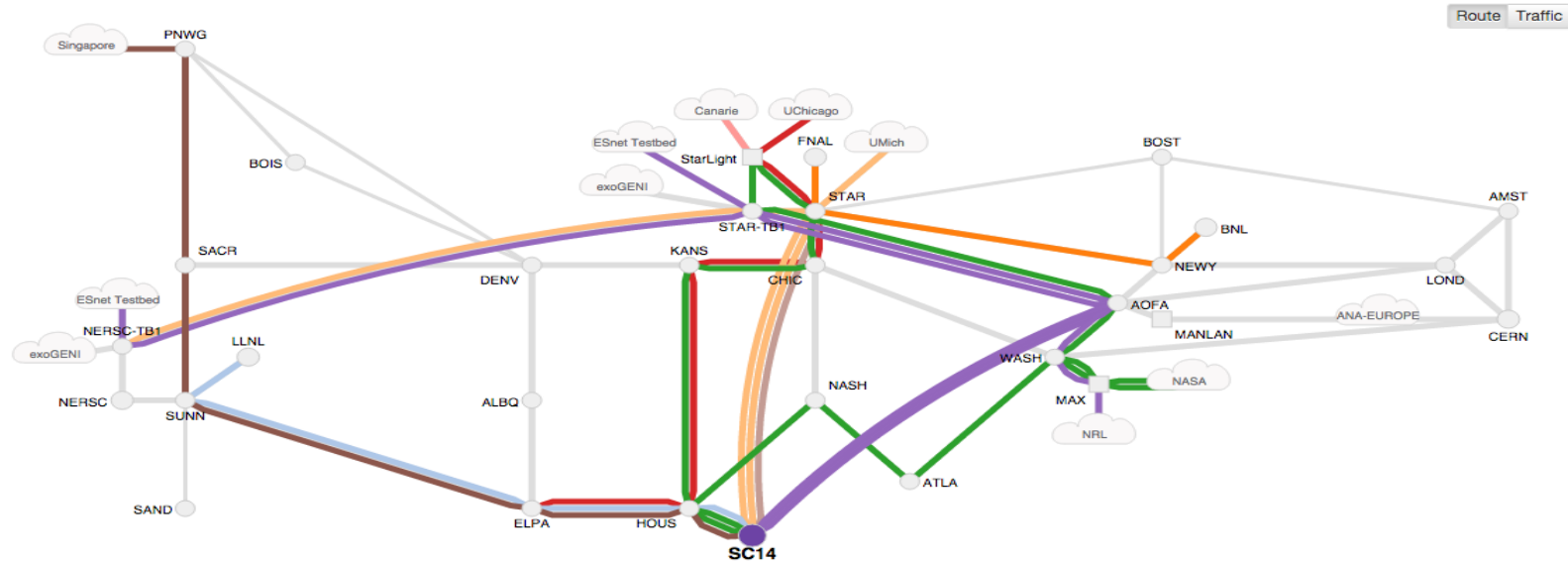
NASA

ANL UDT

ANL GridFTP

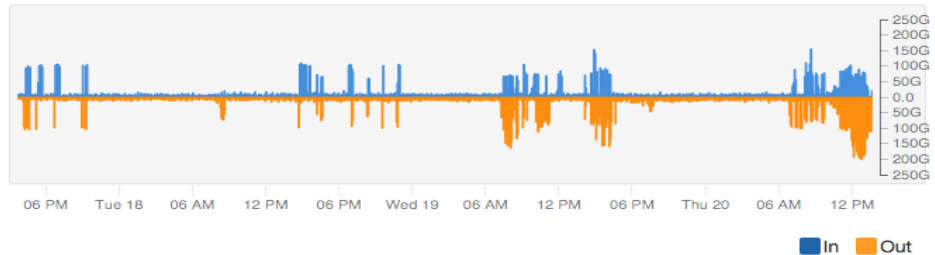
NRL

Aspera



ESnet at SC14

ESnet provided four 100 Gbps circuits to SC14 in support of various demonstrations. This page shows details about those demos. Please note that the total traffic chart in and out are relative to ESnet. All pages show traffic for the duration of SC14, starting at 5:00 PM on Monday, November 17th and concluding at 3:00 PM on Thursday, November 20th.



esmond - ESnet Monitoring Daemon

- system for collecting, storing, visualizing and analyzing large sets of time-series data
- RESTful API which allows easy access to the data which is collected
- perfSONAR project has begun using esmond to store time-series of network measurements

- Documentation
 - <http://software.es.net/esmond/>
- Software
 - <https://github.com/esnet/esmond>

ESnet Database (ESDB)

ESnet wide effort to build a common source of truth for core data sets:

- Organizations
- Contacts
- Locations
- Circuits

Provides a strong foundation for many internal and external processes.

- Web based and REST API access – enables further automation with programmatic access to ESDB data

Internal Software only – no open-source planned



ESDB Contacts

ESDB

- Locations
- Contacts
- Organizations
- Circuits

ESDB is a database for use within ESnet that has two primary roles: to be a source of truth for some kinds of data and to be a home for otherwise homeless data. In the source of truth case it will focus on providing data that needs to be consistent across other databases and systems within ESnet. In the case of data without a home it will be an approachable and robust database system that can be extended to store datasets that otherwise might end up in arbitrary text files or spreadsheets.

ESDB has three main components: the database itself, this web-based user interface and the RESTful API.

ESDB / Organizations

LBL

Lawrence Berkeley National Lab

| | |
|--------------------|--|
| Organization Type | ESnet Site |
| Website | http://www.lbl.gov/ |
| Extra Instructions | It's on top of a big hill with no parking. Take the shuttle |
| ESnet Site | SC (DOE Office of Science) great views Outage Notification Guidelines Please do no break the internet. Thank you. |
| Vendor | No |
| Peering Details | ASN 16 AS-SET: LBL-AS-SET Credentials: seekrit |
| Location | Lawrence Berkeley National Laboratory 5:11 AM, Tuesday |

Contacts

| | |
|------------------------|-------------------|
| ESnet Site Coordinator | Alice Alicedottir |
|------------------------|-------------------|

Not real-data, but shows the fields and possibilities



Modular Forms software creates flexible information gathering interface

ESDB Home ESDB Help Logged in as Jdugan Logout

ESDB / Contacts

Edit contact

General Information

CONTACT TYPE * Human

FIRST NAME * Greg

LAST NAME * Bell

TITLE The job title of this contact.

Related contacts

SUPERVISOR Select...

Communication

EMAIL ADDRESSES

TYPE * work

EMAIL * grbell@es.net

PHONE NUMBERS

TYPE * work

NUMBER *

IM ADDRESSES

Location

LOCATION Energy Sciences Network

Availability

AVAILABILITY * Availability for the contact. Typically their office hours.

BEST AVAILABILITY Optional best time to reach a contact.

✓
✗
🗑️
2 fields still required

Inline
circuit editing
and
visualization
increases data
accuracy and
operational
efficiency

ESNET-Z00110

ALBU-DENV-100GE

ESNET-Z00110



Basic

Circuit ID **ESNET-Z00110**

Description **ALBU-DENV-100GE**

Type **Equipment-Equipment**

State **In-service since Mon Oct 15 2012 00:53:00 GMT-0700 (PDT)**

Layer **Primary circuit**

Related circuits

None

Organizations

Provider **Internet2**

Customer of record **Energy Sciences Network**

Properties

Fiber type **Single-Mode Fiber**

Endpoints

ALBQ-CR5:1/1/1
Equipment Port

Device name **ALBQ-CR5**

Interface **1/1/1**

Location **ALBQ-HUB**

DENV-CR5:3/1/1
Equipment Port

Device name **DENV-CR5**

Interface **3/1/1**

Location **DENV-HUB**

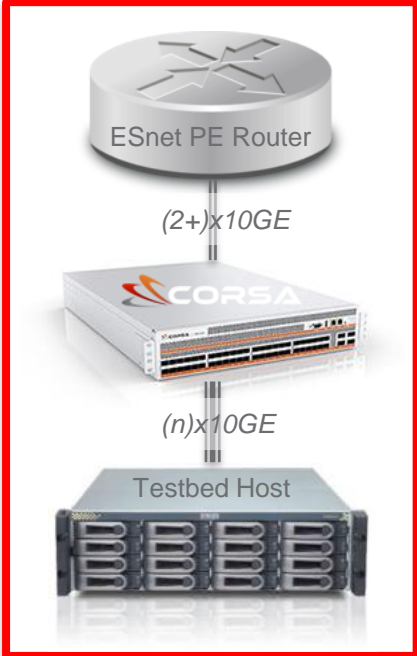
Next-Generation Architectures


investigations only


a *flexible* platform

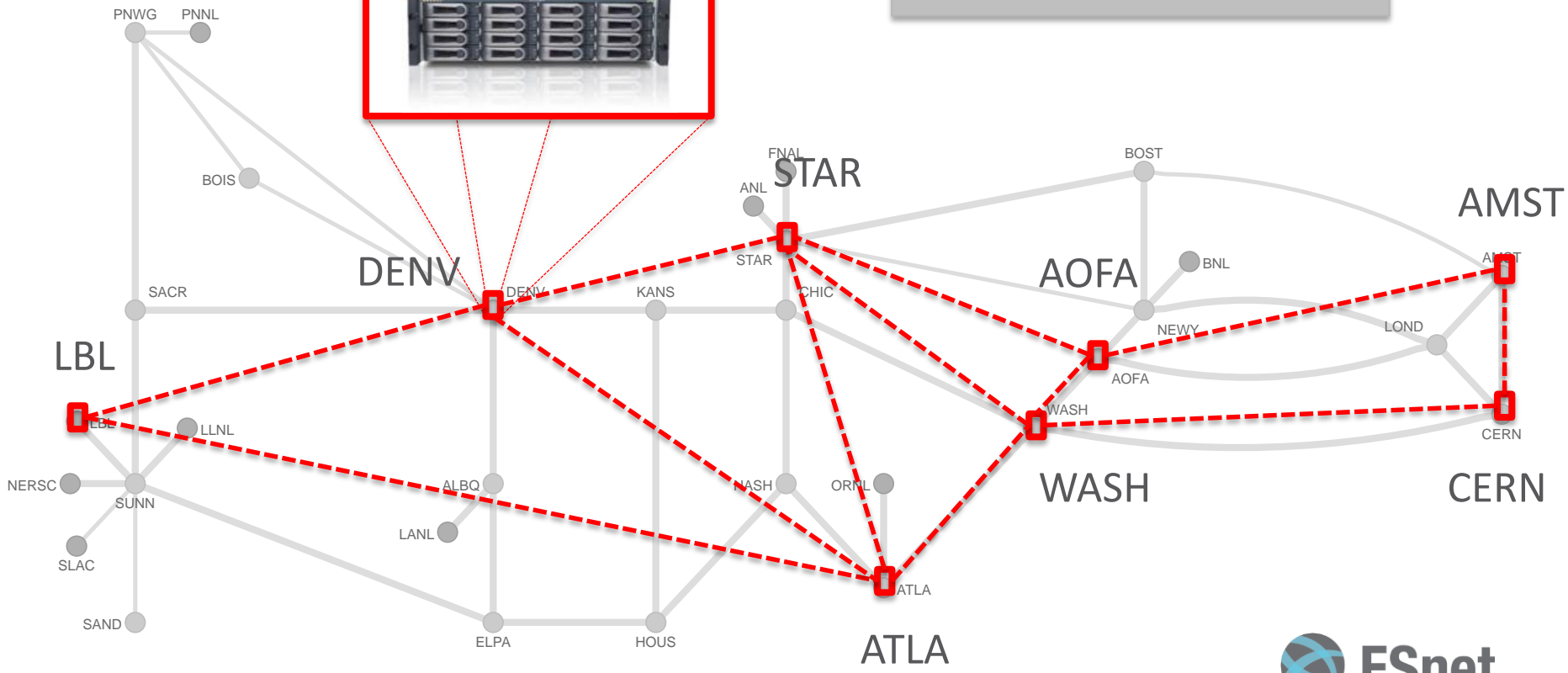
- a *flexible* white-box hardware platform to support existing ESnet services and new abstractions

ESnet SDN Testbed



 Planned SDN Testbed node locations

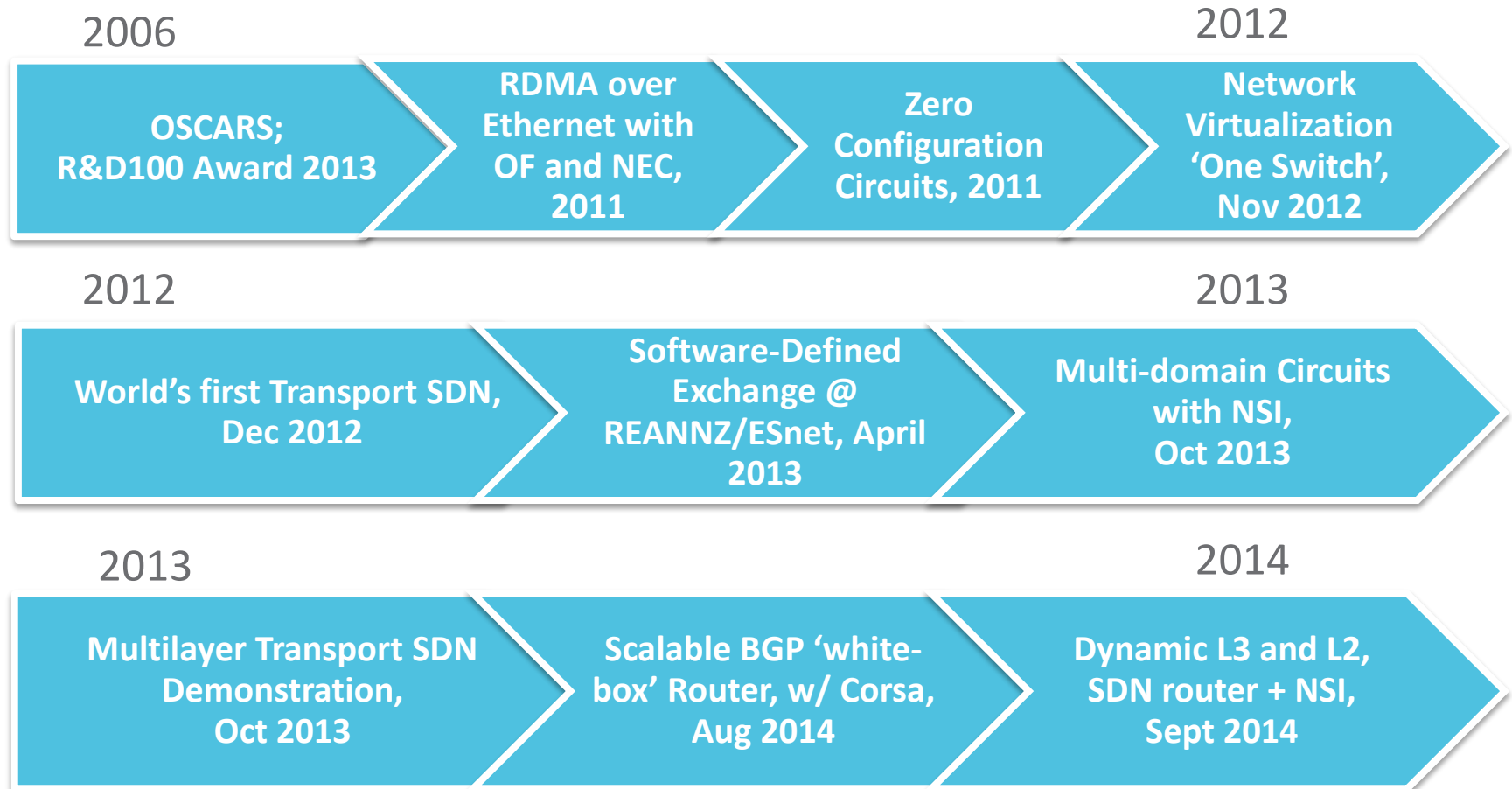
 Planned SDN Testbed connectivity overlay (using OSCARS circuits)



a *flexible, programmable* platform

- a *flexible* hardware platform to support existing ESnet services and new abstractions
- an *environment* to support ‘the right’ abstractions for science applications

ESnet SDN investigations, in a nutshell:



ESnet Operating System

- Software platform with capabilities to run multiple science applications
- Integrating the capabilities from above point-demonstrations, into a working platform running on flexible hardware infrastructure
- Concept still under development, presentations and paper by the end of the year



Single App

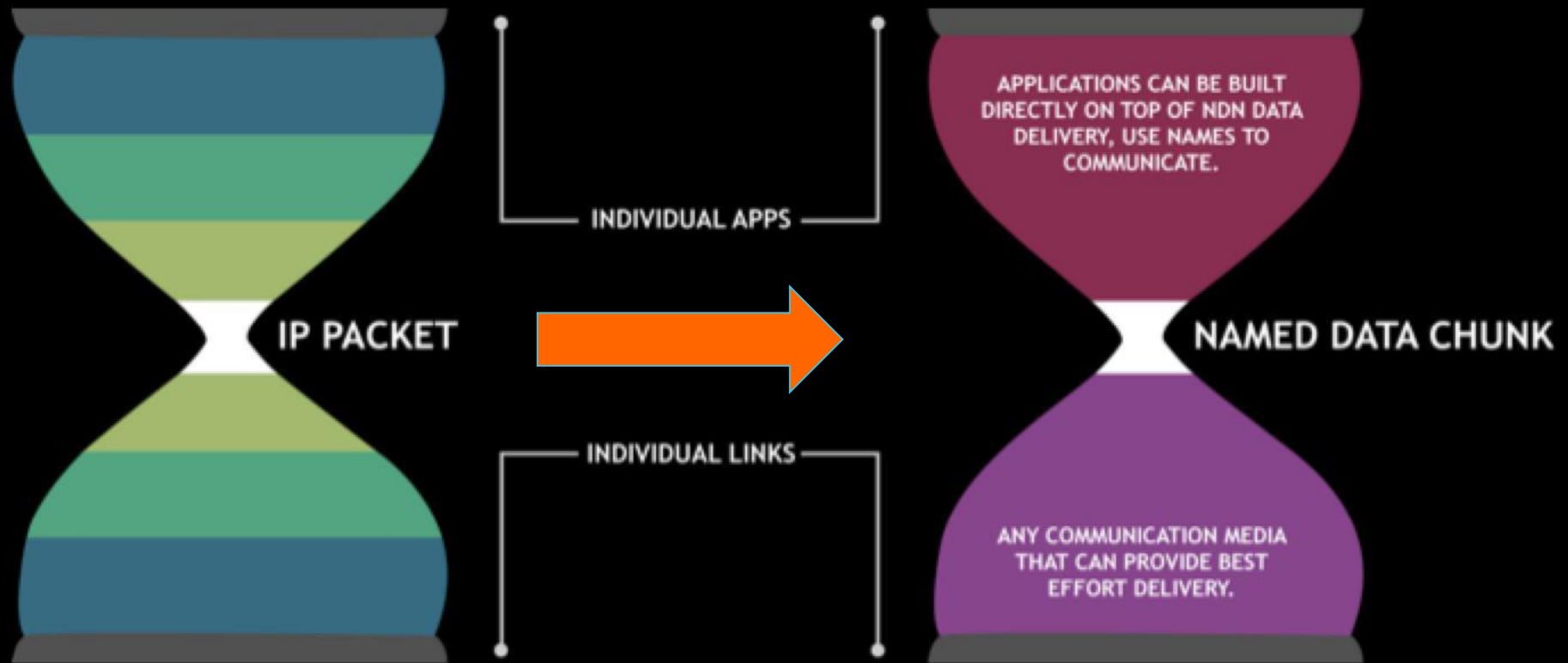


Multiple Apps

a *flexible, programmable, content* platform

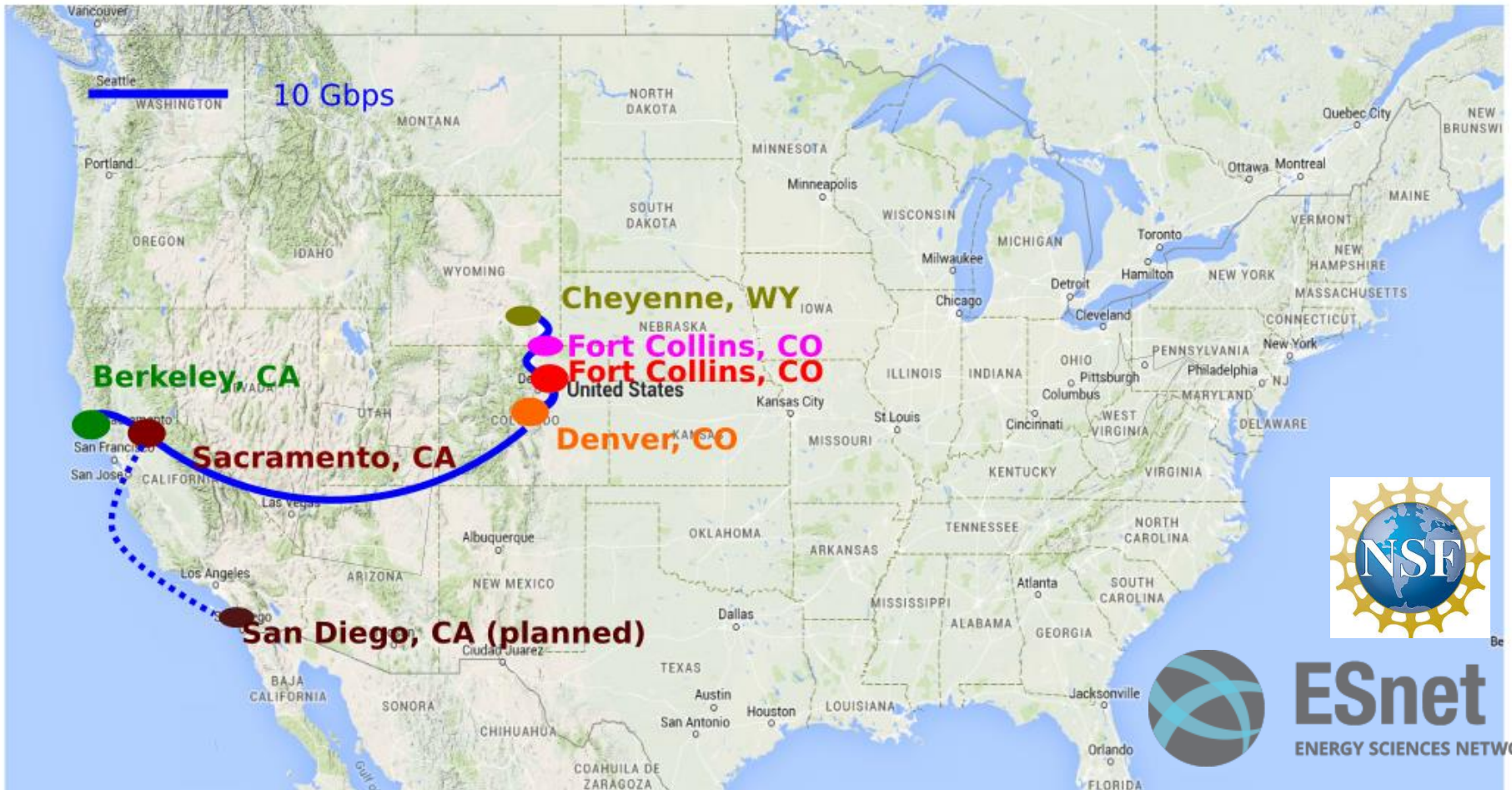
- a *flexible* hardware platform to support existing ESnet services and new abstractions
- an *environment* to support ‘the right’ abstractions for science applications
- a *content* platform to abstract the placement and retrieval of content

Named-Data Networking (NDN)



- The new idea: **Name the data, not the hosts!**
- ..so you can ask the network directly for the data!
- New abstractions for science applications looking for data

NDN for Climate Testbed with Colorado State University



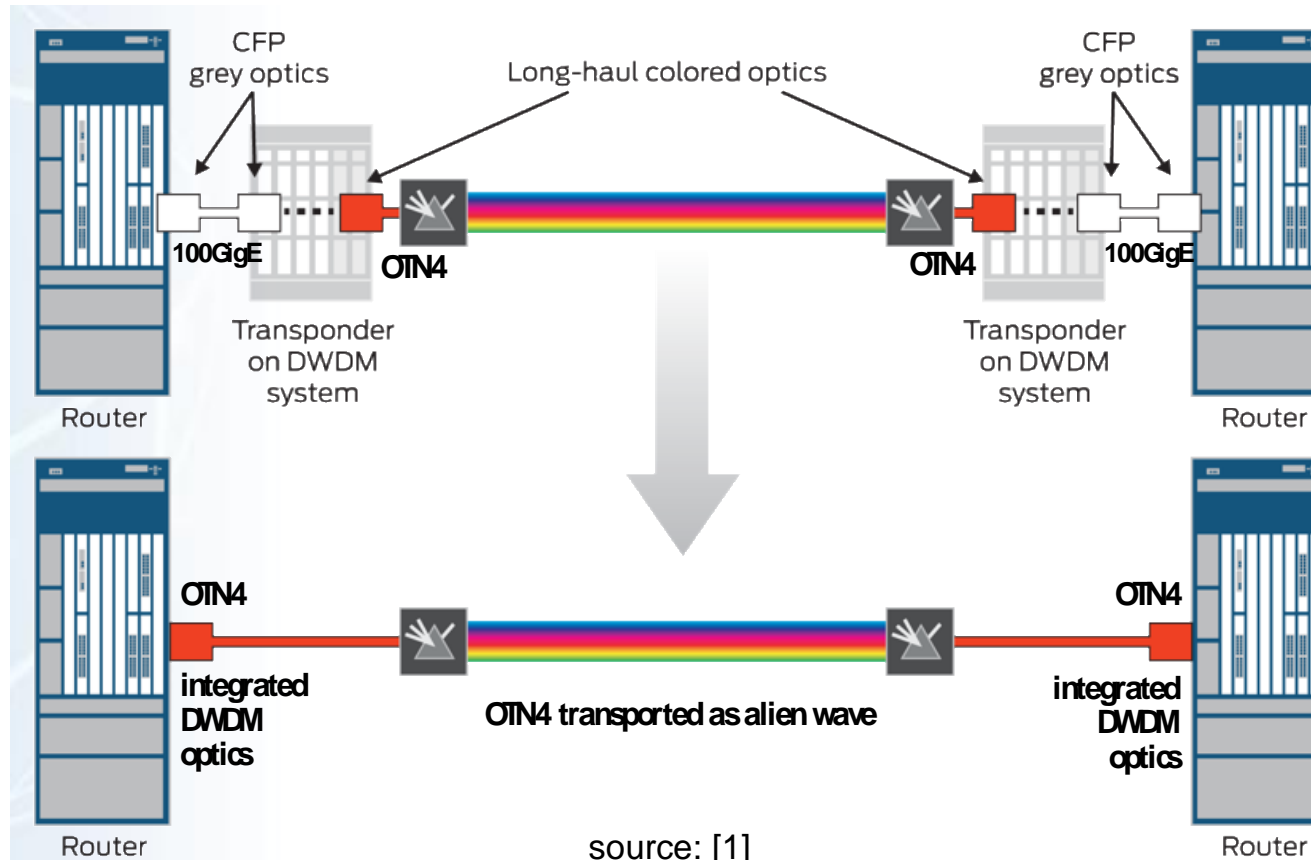
ESnet
ENERGY SCIENCES NETWORK



a *converged, flexible, programmable, content* platform

- a *flexible* hardware platform to support existing ESnet services and new abstractions
- an *environment* to support ‘the right’ abstractions for science applications
- a *content* platform to abstract the placement and retrieval of content
- a *converged* platform to optimize the cost of the moving bits end-to-end

Packet-Optical convergence

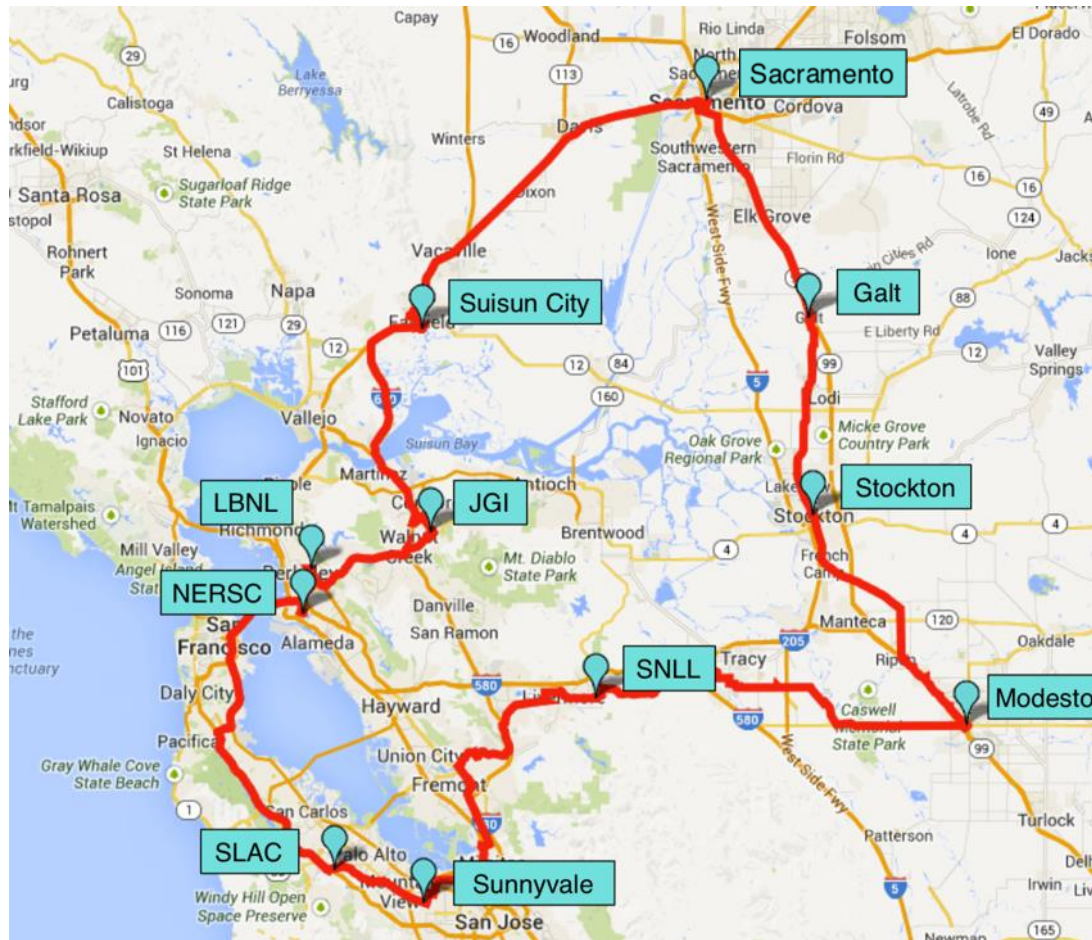


From Chris Tracy's OFC presentation

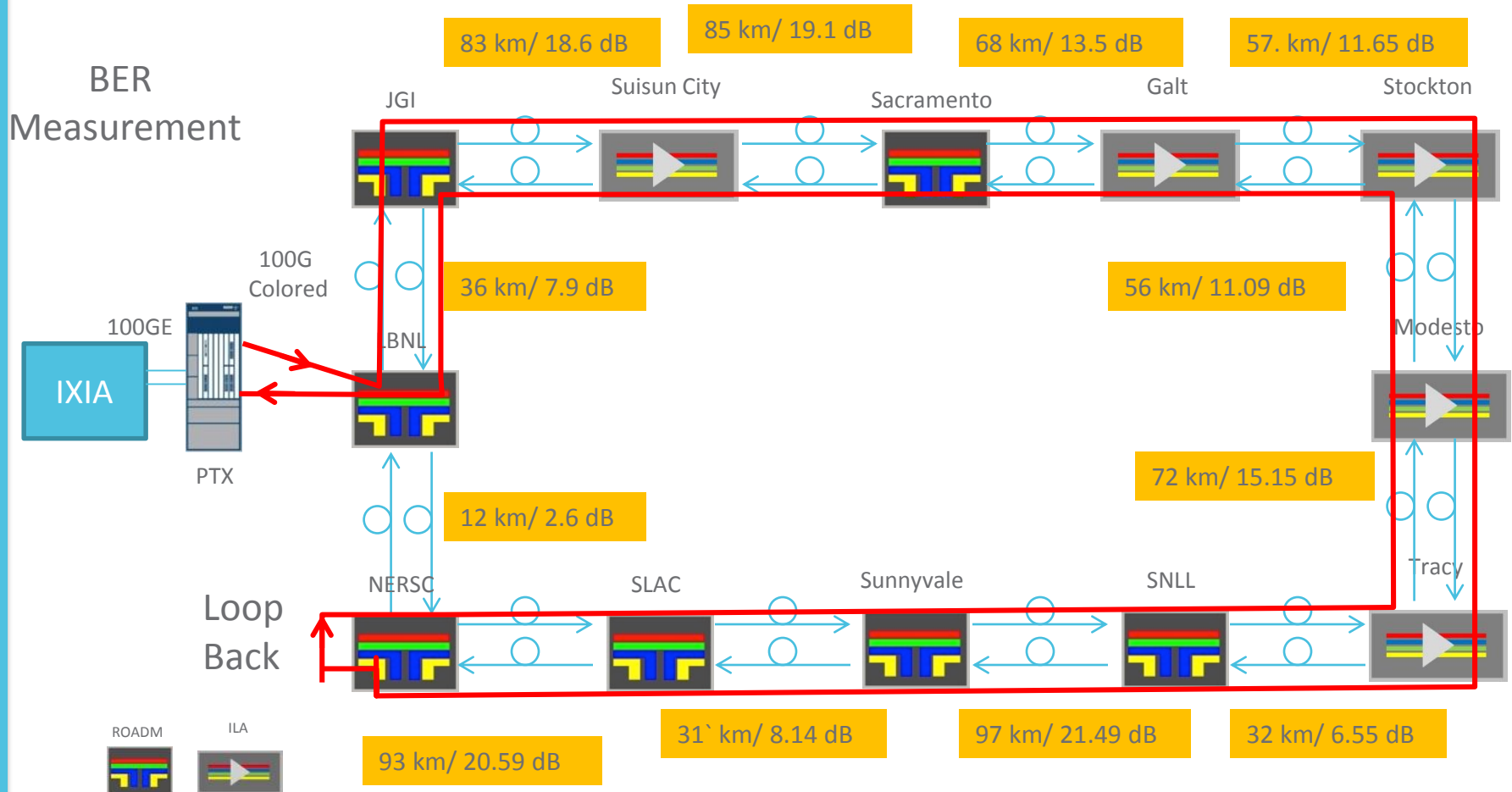
JUNIPER
NETWORKS

ESnet

Field Trial of ONT4 Router Interfaces



Testing on live, 1400 km Bay Area Loop



Field Test Results

- Demonstrated:
 - Error-free transmission of integrated DWDM optics in the router over entire distance
 - over mix of fiber types (1400km was LEAF + SMF-28), and - in a multi-vendor environment
- Bit-error Rates (BER) of production wavelengths were not affected
- Operational integration of integrated DWDM router optics with third-party DWDM system
 - Optically, the transport system treated like any other wavelength
 - Fully automated provisioning/teardown via optical control plane
 - Fully automated power balancing/equalization, gain adjustments, etc.
- More work to be done to incorporate this architecture into the programmable, content, platform envisioned

a *operationally supportable, converged, flexible, programmable, content* platform

- a *flexible* hardware platform to support existing ESnet services and new abstractions
- an *environment* to support ‘the right’ abstractions for science applications
- a *content* platform to abstract the placement and retrieval of content
- a *converged* platform to optimize the cost of the moving bits end-to-end
- an *operationally supportable* platform with abstractions suitable for automation and analytics

Operationalizing SDN: Calling attention to Manageability, Implementation and Security

OPERATIONALIZATION OF SOFTWARE/DEFINED NETWORKS (SDN) PROGRAM REVIEW

Dec 16-17, 2013



Conducted in Arlington, Virginia by
National Science Foundation
Department of Energy
National Coordination Office

<https://www.ora.gov/sdnpr2013/default.htm>

Roadmap to Operational SDN

- July 14 – 16th @ Berkeley

Thank you!



Innovation

Is Worthless

unless it solves real problems for real people.

Backup

ESnet Vision and Goals

(in addition to world-class operations...)

vision:

Discovery is unconstrained by geography.

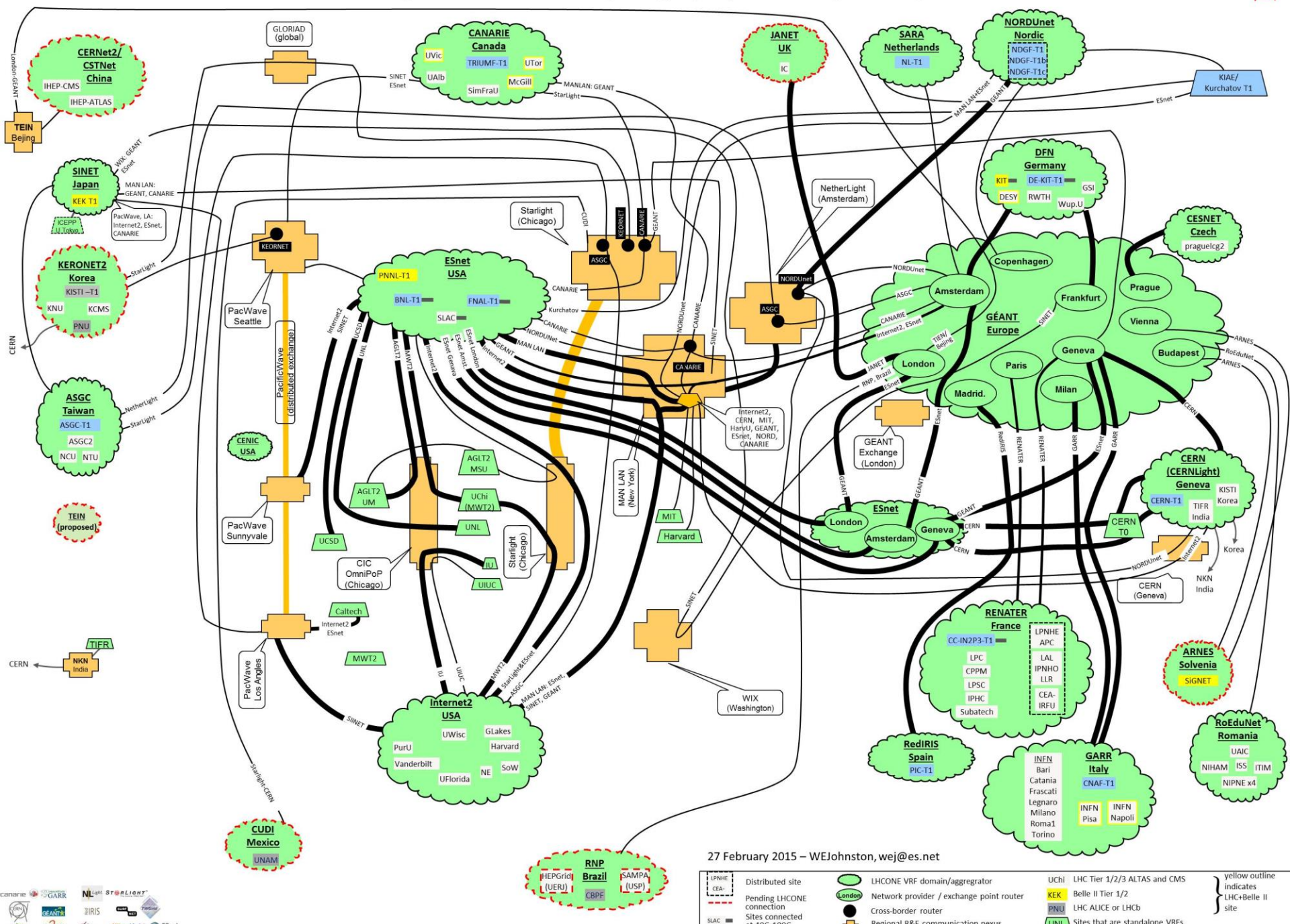
goals:

1. Improve networking practices globally.

2. Provide information and tools for optimal network use.

3. Pioneer architectures, protocols, applications.

LHCONE: A global infrastructure for the High Energy Physics (LHC and Belle II) data management



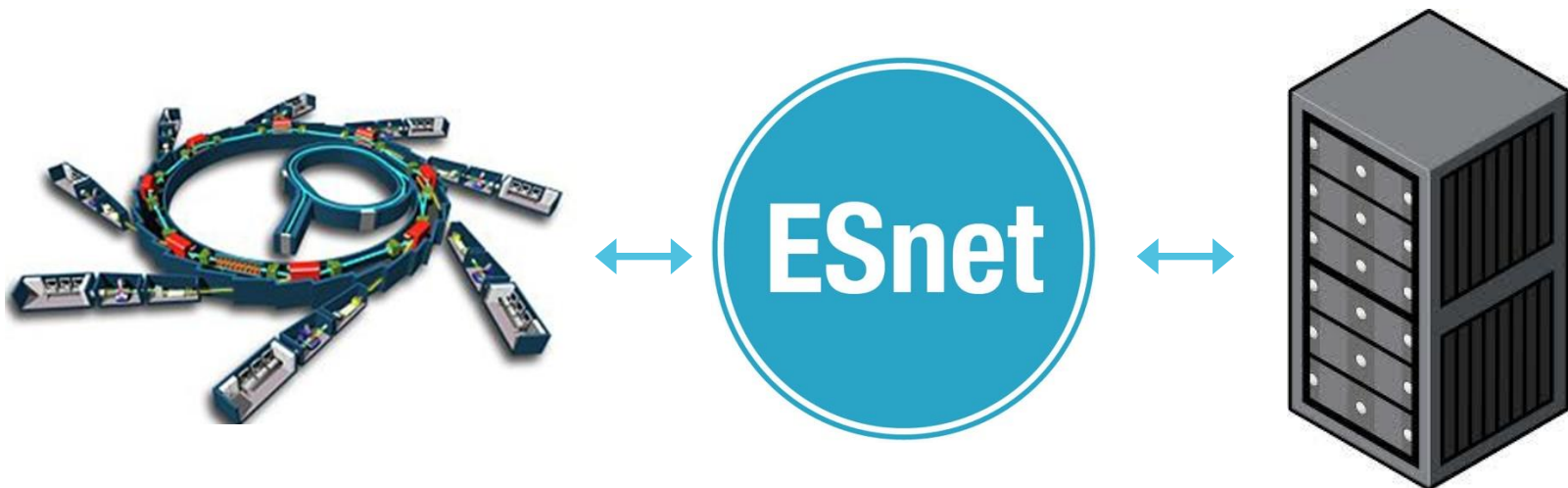
27 February 2015 – WEJohnston, wej@es.net

| | | |
|-----------------------------|---|--|
| Distributed site | LHCONE VRF domain/aggregator | LHC Tier 1/2/3 ALTAS and CMS |
| Pending LHCONE connection | Network provider / exchange point router | Belle II Tier 1/2 |
| Sites connected at 40G-100G | Cross-border router | LHC ALICE or LHCb |
| Broadcast VLAN | Regional R&E communication nexus w/ switch providing VLAN connections | Sites that are standalone VRFs, |
| | | Uchi LHC Tier 1/2/3 ALTAS and CMS |
| | | KEK Belle II Tier 1/2 |
| | | PNU LHC ALICE or LHCb |
| | | UNL Sites that are standalone VRFs, |
| | | RENATER France: CC-IN2P3-T1, LPNHE, APC, LAL, IPNHO, LLR, LPSIC, IPHC, Subatech |
| | | GARR Italy: INFN Bari, Catania, Frascati, Legnaro, Milano, Roma1, Torino; INFN Pisa, INFN Napoli |
| | | RedIRIS Spain: PIC-T1 |
| | | RoEduNet Romania: UAIC, NIHAM, ISS, ITIM, NIPNE x4 |
| | | ARNES Slovenia: SIGNET |
| | | CERN (CERNLight) Geneva: KISTI Korea, TIFR India, NKN India |
| | | DFN Germany: KIT, DE-KIT-T1, GSI, DESY, RWTH, Wup.U |
| | | NORDUnet Nordic: NDFG-T1, NDFG-T1b, NDFG-T1c |
| | | JANET UK: IC |
| | | SARA Netherlands: NL-T1 |
| | | CANARIE Canada: UVic, TRIUMF-T1, Utor, UAib, SimFraU, McGill |
| | | Internet2 USA: PurU, UWisc, Glakes, Harvard, Vanderbilt, NE, SoW, UFlorida |
| | | Esnet USA: PNNL-T1, BNL-T1, FNAL-T1, SLAC, Starlight (Chicago) |
| | | Esnet Europe: Amsterdam, Frankfurt, Prague, Vienna, Budapest, Paris, Geneva, Milan, Madrid, London |
| | | ASGC Taiwan: ASGC-T1, ASGC2, NCU, NTU |
| | | KERNET2 Korea: KISTI-T1, KNU, KCMS, PNU |
| | | SINET Japan: KEK T1 |
| | | CERNET2/ CSTNet China: IHEP-CMS, IHEP-ATLAS |
| | | RNP Brazil: IHEPGrid (UERJ), SAMPA (USP), CBPE |
| | | CUDI Mexico: UNAM |
| | | TEIN (proposed) |
| | | TEIN Beijing |
| | | KEONET |
| | | KEONET2 |
| | | KEONET3 |
| | | KEONET4 |
| | | KEONET5 |
| | | KEONET6 |
| | | KEONET7 |
| | | KEONET8 |
| | | KEONET9 |
| | | KEONET10 |
| | | KEONET11 |
| | | KEONET12 |
| | | KEONET13 |
| | | KEONET14 |
| | | KEONET15 |
| | | KEONET16 |
| | | KEONET17 |
| | | KEONET18 |
| | | KEONET19 |
| | | KEONET20 |
| | | KEONET21 |
| | | KEONET22 |
| | | KEONET23 |
| | | KEONET24 |
| | | KEONET25 |
| | | KEONET26 |
| | | KEONET27 |
| | | KEONET28 |
| | | KEONET29 |
| | | KEONET30 |
| | | KEONET31 |
| | | KEONET32 |
| | | KEONET33 |
| | | KEONET34 |
| | | KEONET35 |
| | | KEONET36 |
| | | KEONET37 |
| | | KEONET38 |
| | | KEONET39 |
| | | KEONET40 |
| | | KEONET41 |
| | | KEONET42 |
| | | KEONET43 |
| | | KEONET44 |
| | | KEONET45 |
| | | KEONET46 |
| | | KEONET47 |
| | | KEONET48 |
| | | KEONET49 |
| | | KEONET50 |

Also see <http://lhcone.net> for details.

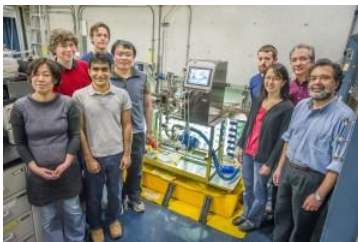
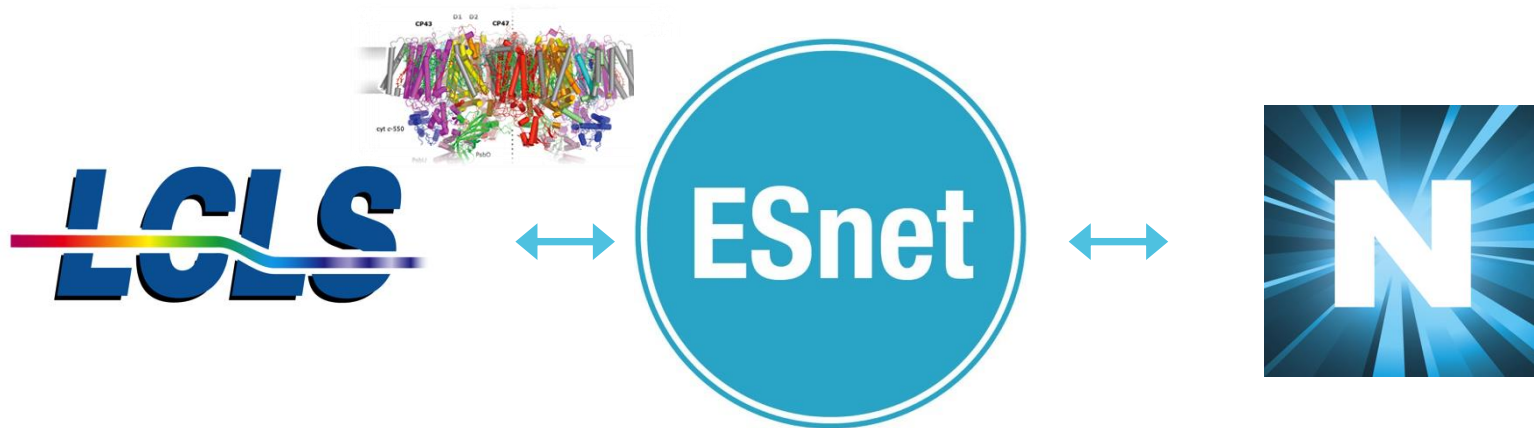
This architecture (instruments and computational resources coupled by networks) now spreading outside HEP: 'super-facilities.'

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



Super-facility example #1:

Researchers from Berkeley Lab and SLAC conducted protein crystallography experiments at LCLS to investigate photoexcited states of PSII, with near-real-time computational analysis at NERSC.



“Taking snapshots of photosynthetic water oxidation using femtosecond X-ray diffraction and spectroscopy,” *Nature Communications* 5, 4371 (9 July 2014)

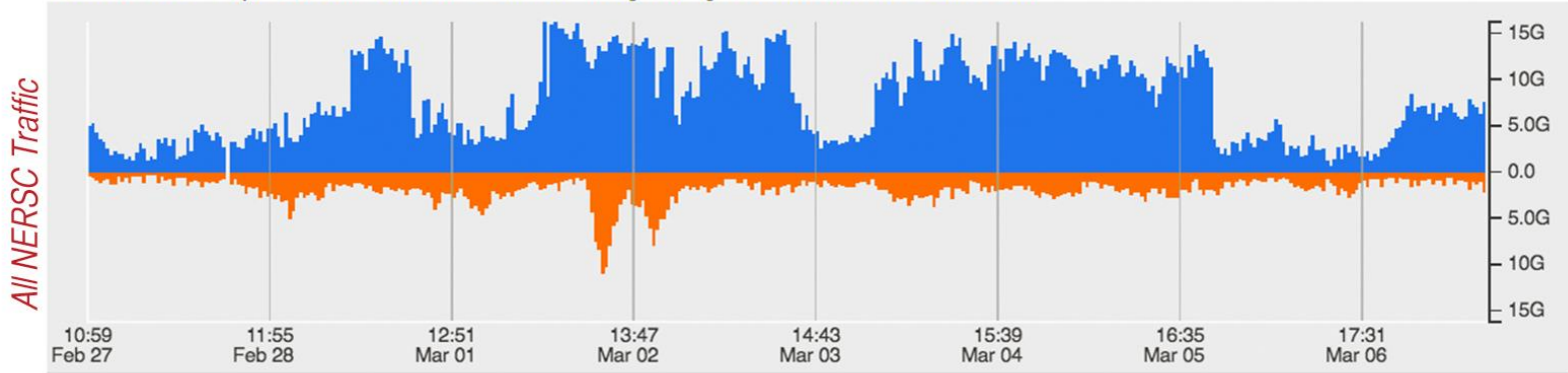


Data flow from single LCLS detector *tripled* network utilization for major HPC center.

From : Wed Feb 27 10:59:00 2013 To : Thu Mar 7 10:59:00 2013

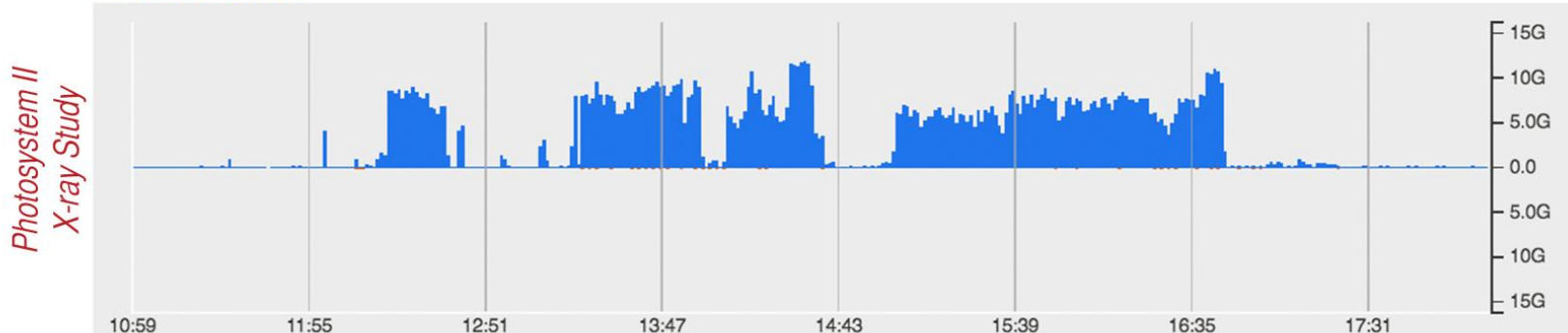
■ To site ■ From site

Total traffic Tip: Double Click to Zoom-In and [SHIFT] Double click to Zoom-Out



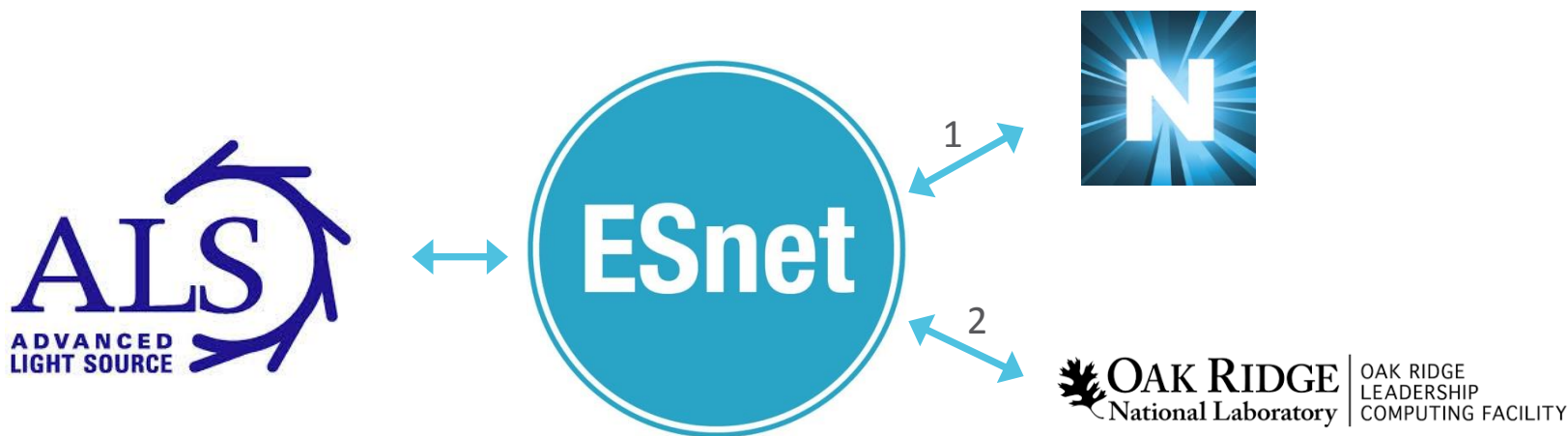
Traffic split by : 'Autonomous System (origin)'

nersc-SLAC:3671



Super-facility example #2:

Real-time analysis of 'slot-die' technique for printing organic photovoltaics, using ALS + NERSC (SPOT Suite for reduction, remeshing, analysis) + OLCF (HipGISAXS running on Titan w/ 8000 GPUs).



<http://www.es.net/news-and-publications/esnet-news/2015/esnet-paves-way-for-hpc-superfacility-real-time-beamline-experiments/>

Results presented at March 2015 meeting of American Physical Society by Alex Hexemer.

Additional DOE contributions: **GLOBUS** (ANL), **CAMERA** (Berkeley Lab)

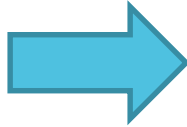


Super-facility-on-demand demo at NSF GENI conference



SPADE instance - server at Argonne

ESnet,
Internet2

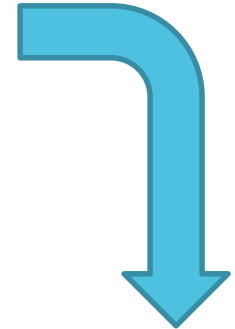


ExoGENI SPADE VM @
Starlight, Chicago

ESnet



ExoGENI SPADE VM @
Oakland, California



Compute Cluster
NERSC, LBL



Data from ALS
beamline

- fictional - but realistic - workflow
- **dedicated systems for data transfer and network circuits** created *programmatically*
- future vision: application declares *intention* for super-facility, network responds
- “Science DMZ as a service”

In collaboration with Ilya Baldine and his team from RENC1 and Craig Tull and his team from CRD. <http://portal.nersc.gov/project/als/sc14/>