U.S. DEPARTMENT OF ENERGY | Office of Science

ESnet
ENERGY SCIENCES NETWORK

Approaching the ESnet6 Era

ASCAC Meeting
September 25, 2020

Ben Brown
ESnet Program Manager
Advanced Scientific Computing Research

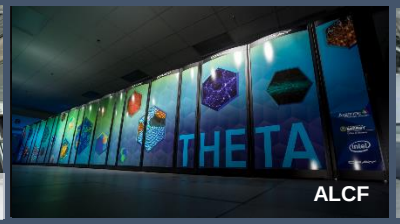
Inder Monga
Executive Director, ESnet
Lawrence Berkeley National Laboratory

The ESnet user facility: A complex system tuned for science

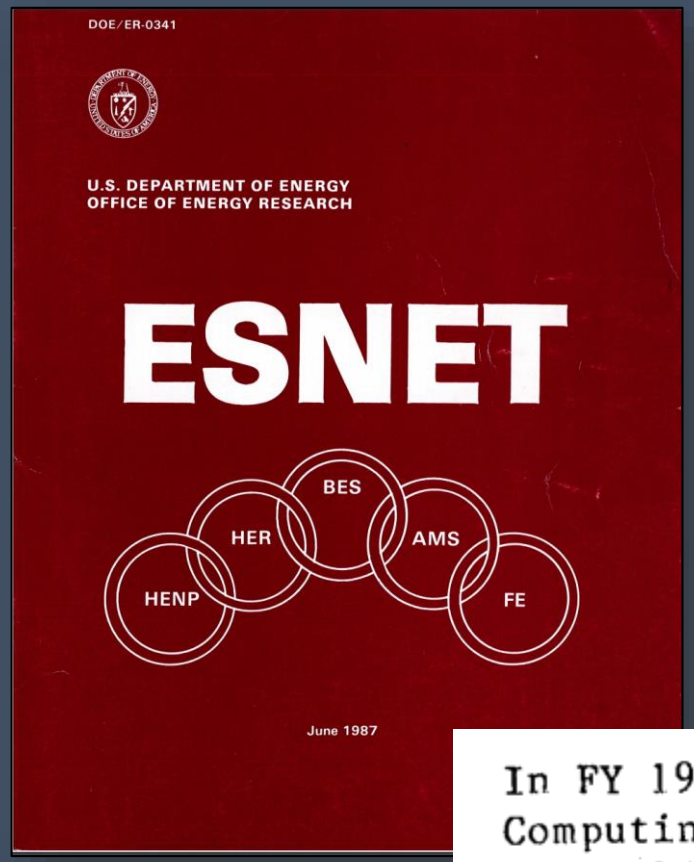
The Office of Science's world-class high performance network, providing services and innovation to enable research

ESnet
ENERGY SCIENCES NETWORK

FY 2020
27 scientific user facilities
36,000+ users



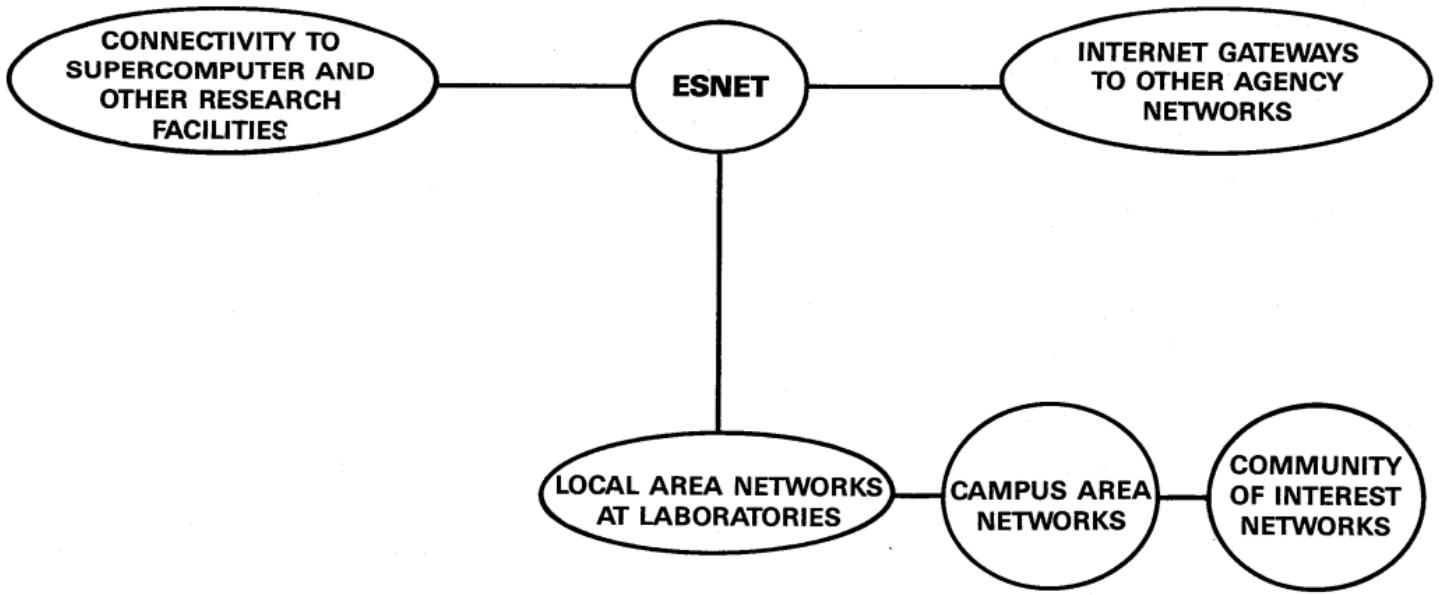
U.S. DEPARTMENT OF
ENERGY
Office of Science



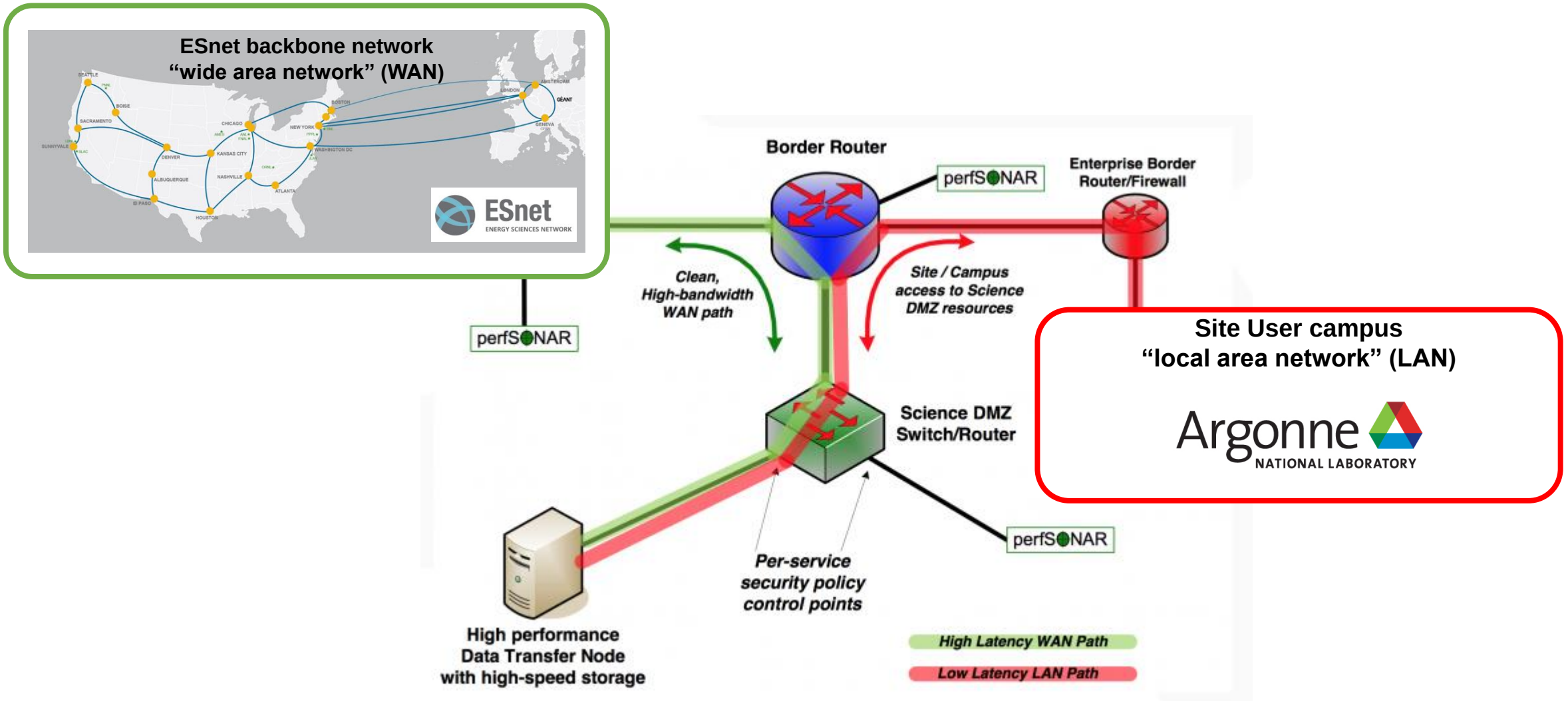
June 1987

In FY 1985, Dr. Alvin Trivelpiece, Director of ER, charged the Scientific Computing Staff (SCS) to survey the status of and requirements for computer networking throughout all ER programs. This project served as a complementary adjunct to the existing SCS charter for the provision of nationwide access to ER supercomputers. The SCS survey data demonstrated a significant need for improved computer networking to facilitate: 1.) improved access to unique ER scientific facilities, 2.) needed information dissemination among scientific collaborators throughout all ER programs, and 3.) more widespread access to existing supercomputer facilities.

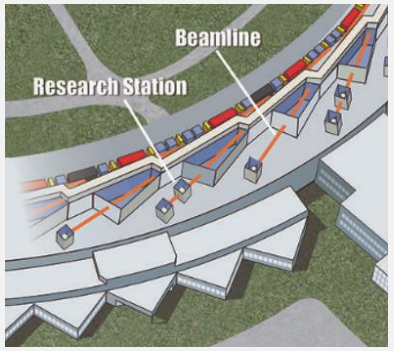
ESNET ARCHITECTURAL MODEL



A typical ESnet Site User border connection



ESnet's Site User Model



Site User

Site User

Site User

Site User

Site User

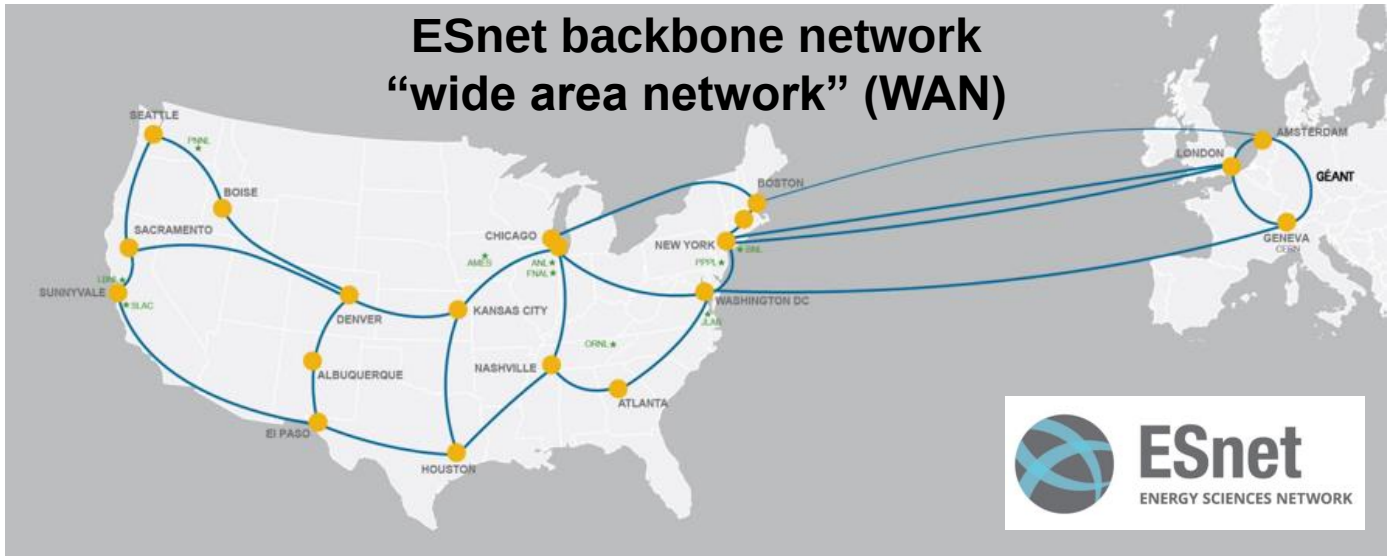
Site User

Site User

Site User

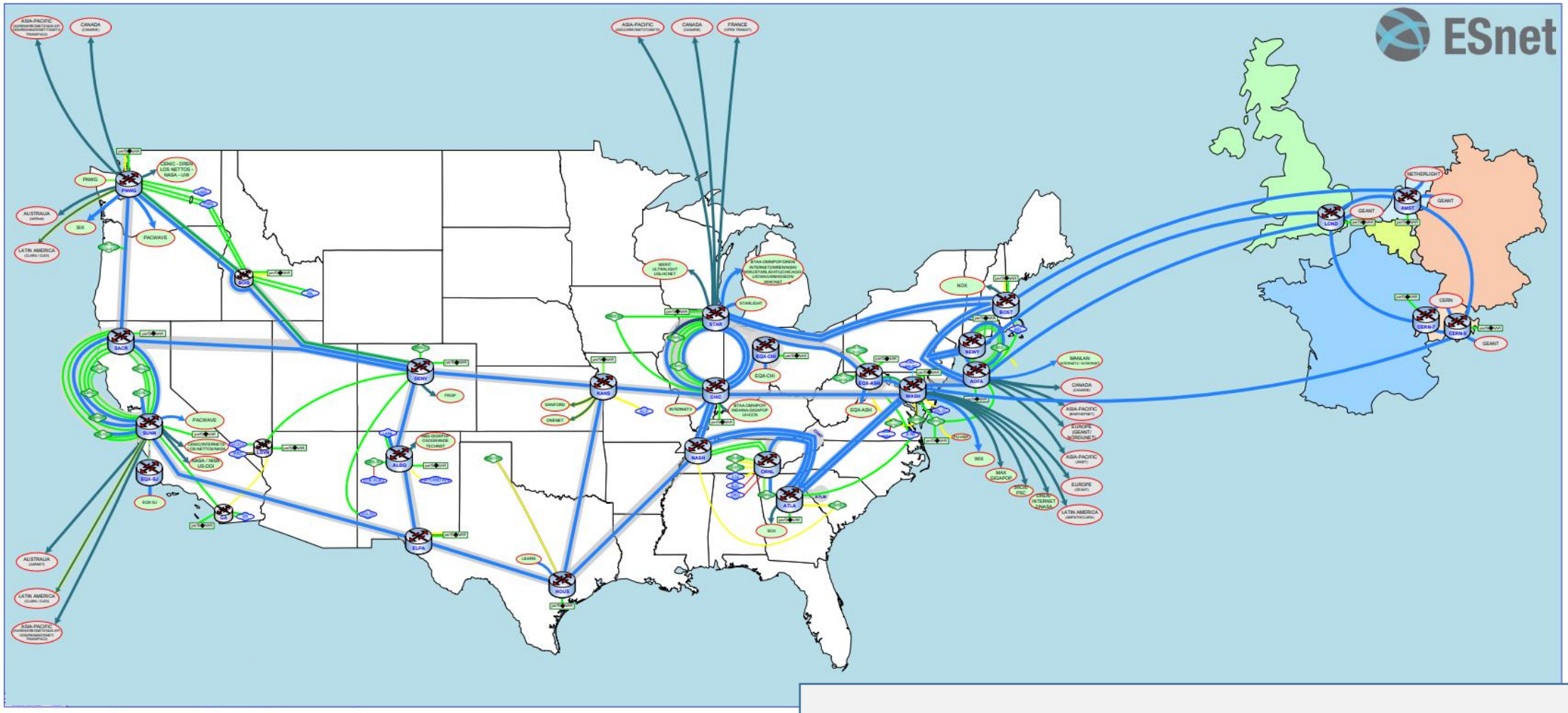
Site User

ESnet backbone network
"wide area network" (WAN)



Site Users enjoy membership in a user group called the ESnet Site Coordinators Committee

The number of Site Users continues to grow



Today ESnet has 62 Site Users

In 2012 ESnet had ~40 Site Users

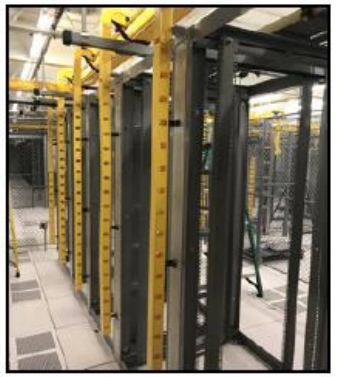
The ESnet6 Project is a greenfield build

There are 311 ESnet6 equipment "colocation" installation sites:

- **39 optical switching sites**
- **272 optical in-line amplification sites**

"Colocation" or "co-lo" is telecom-speak for a shared commercial leased space for telecom equipment

"co-lo" space



racks / power

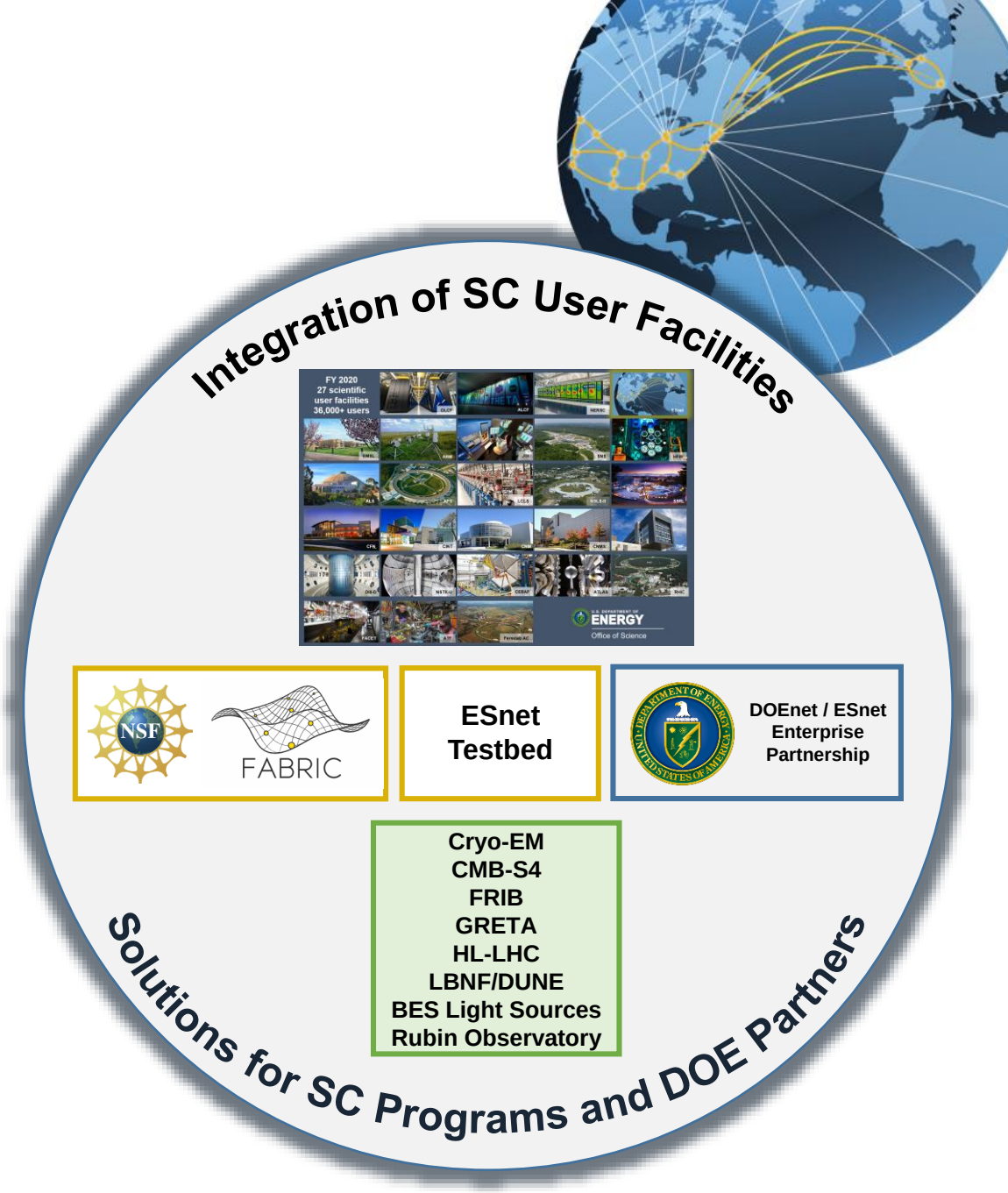
equipment

Themes of the ESnet6 Era

FOUNDATIONS
Next Generation Infrastructure & Services

INNOVATION
Testbeds and Advanced Networking R&D

CO-DESIGN
Partnerships for New Data Solutions



The ESnet6 architecture is about programmable orchestration of **network attributes and services … in a Terabit scale network**

The architecture is built on the Software Defined Networking R&D investments of the past

Music Director Marin Alsop and the Baltimore Symphony Orchestra

FOUNDATIONS

Next Generation Infrastructure & Services

U.S. DEPARTMENT OF ENERGY | Office of Science

ESnet's R&D portfolio is composed to
create strategic flexibility
generate native innovation, and
enable priority goals in the SC/DOE/national research enterprise

INNOVATION

Testbeds and Advanced Networking R&D

The GRETA (Gamma-Ray Energy Tracking Array) Instrument

To be deployed at the Facility for Rare Isotope Beams

CO-DESIGN

Partnerships for New Data Solutions

The Vera C. Rubin Observatory Project
Cerro Pachón, Chile

CO-DESIGN

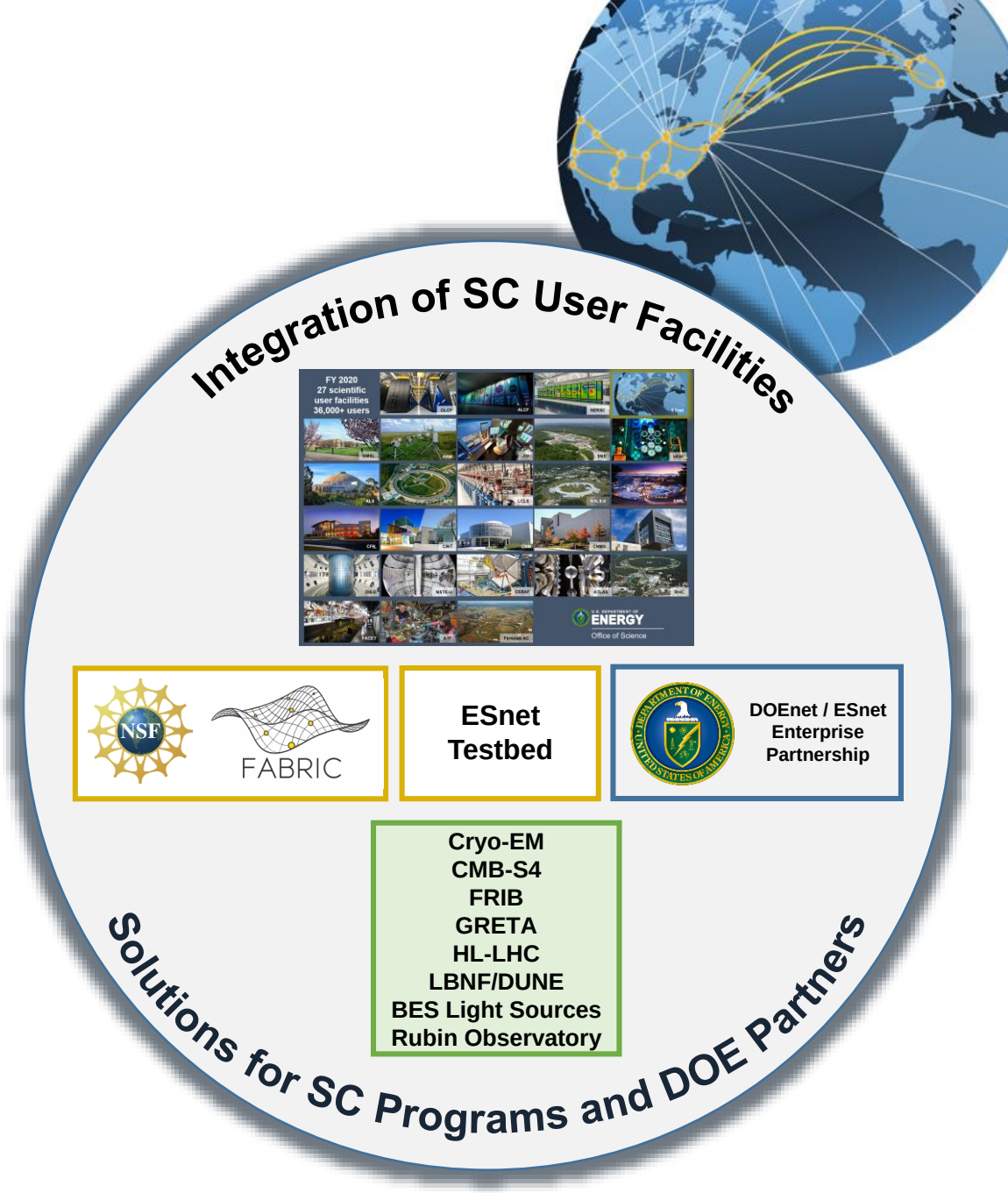
Partnerships for New Data Solutions

Themes of the ESnet6 Era

FOUNDATIONS
Next Generation Infrastructure & Services

INNOVATION
Testbeds and Advanced Networking R&D

CO-DESIGN
Partnerships for New Data Solutions



Vision: Networks accelerate scientific discovery

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

ESnet

ESnet Strategic Goals Follow the Vision

Vision:

Discovery, unconstrained by geography.

Approaches:

1. **Outstanding operations and planning.**
2. **Create and disseminate information and tools for optimal network use.**
3. **Pioneer architectures, protocols, applications.**

FOUNDATION

CO-

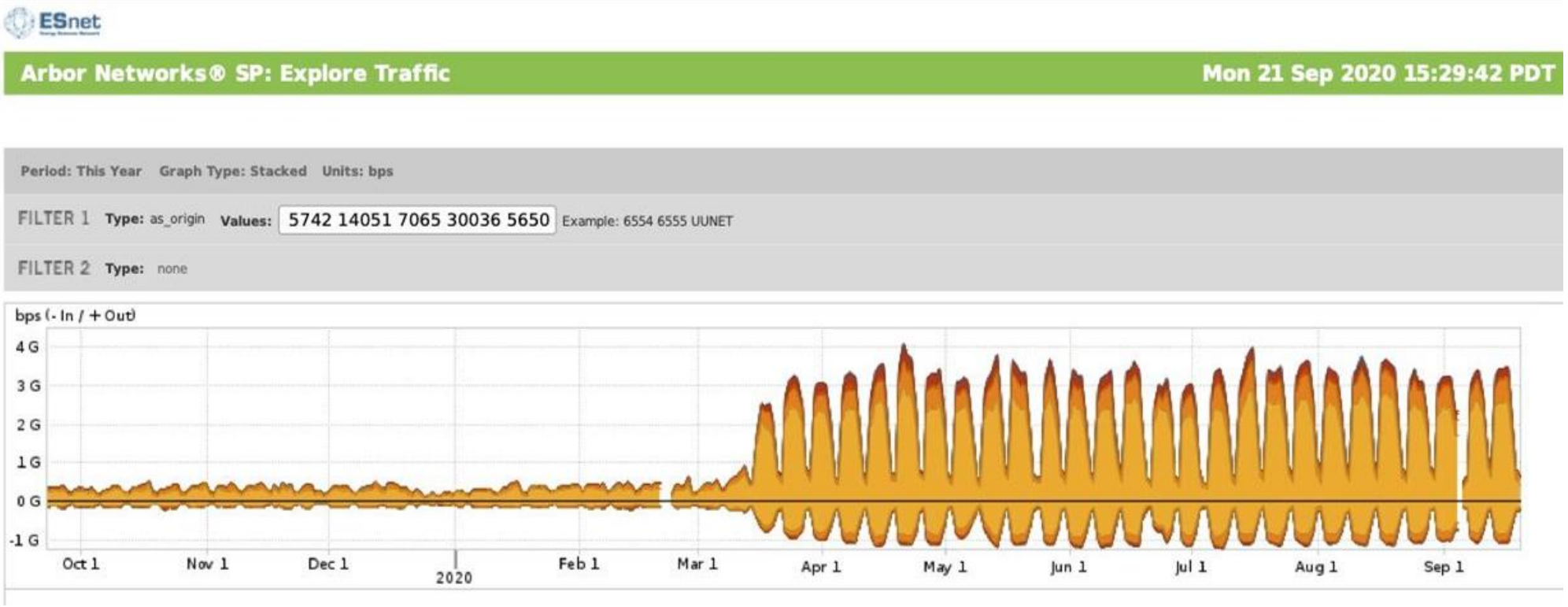
INNOVATIO

ESnet

Key takeaways

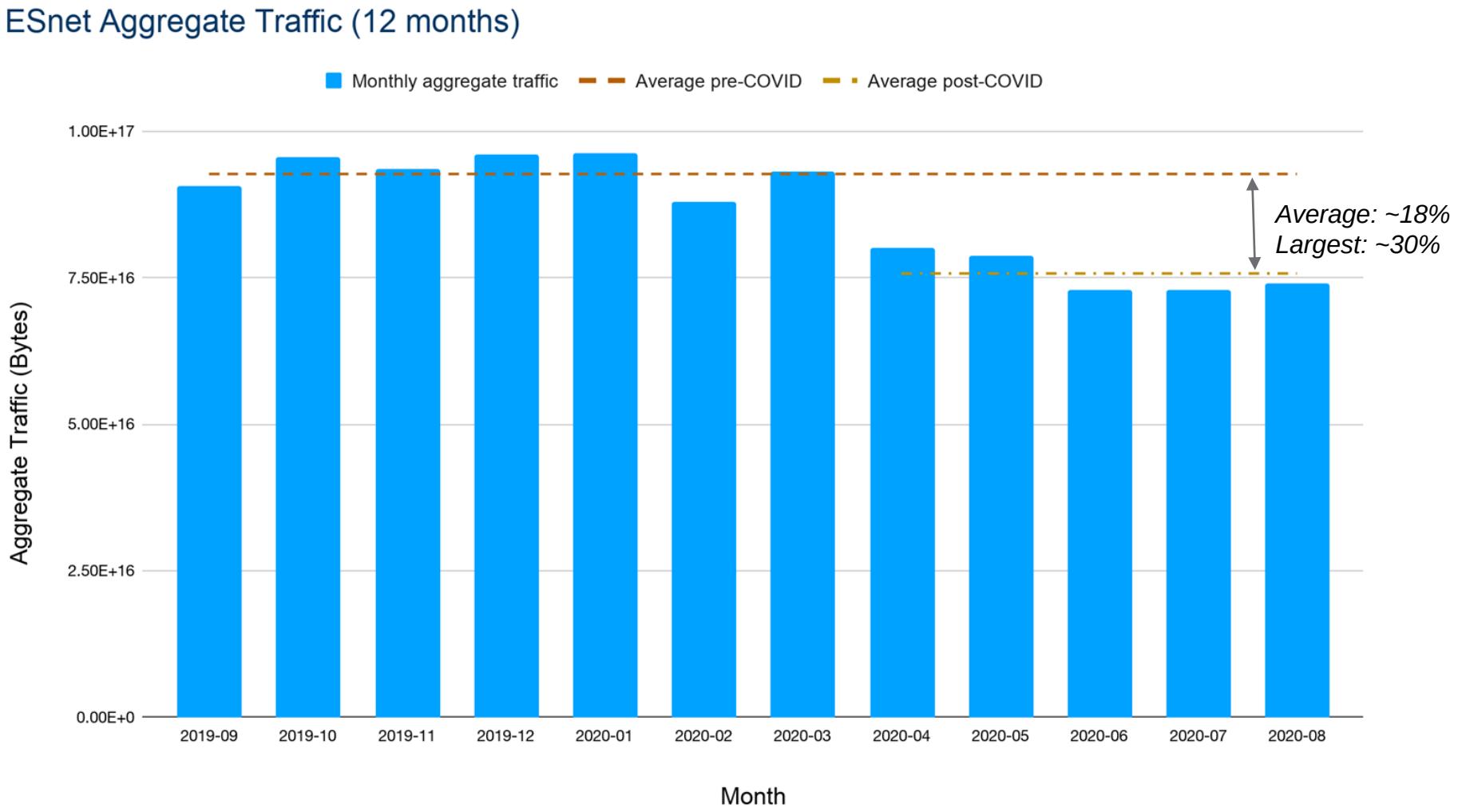
ESnet continues to execute well as a User Facility despite COVID

Operational continuity maintained for all sites in their transition to work-from-home



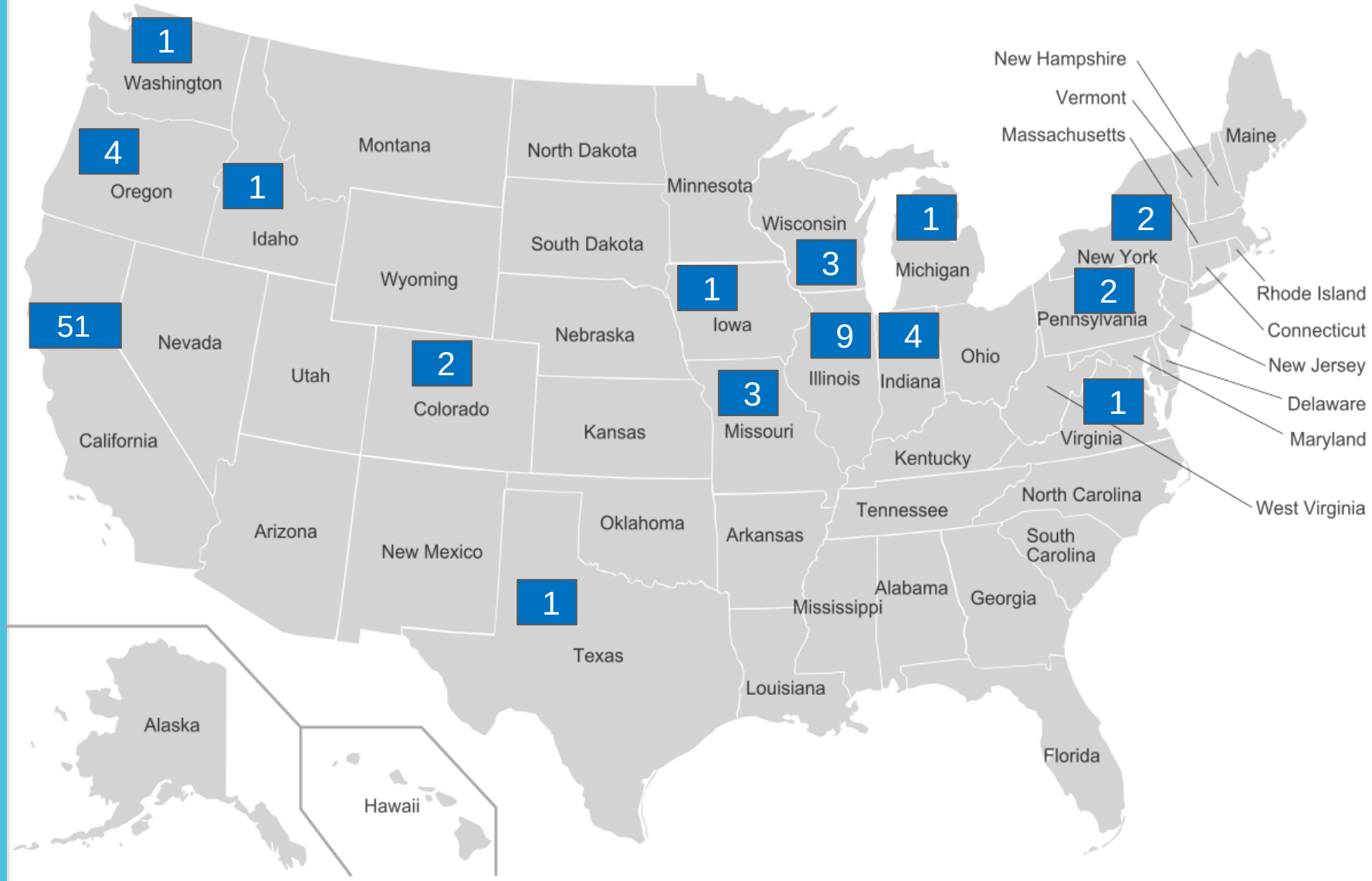
Informal best practices to configure and operate home wifi by staff from internal #homenetworks channel - shared with NLCIOs

Limited on-site personnel causing a dip in aggregate traffic over the backbone



ESnet

Hiring continues to be strong due to distributed nature of the organization

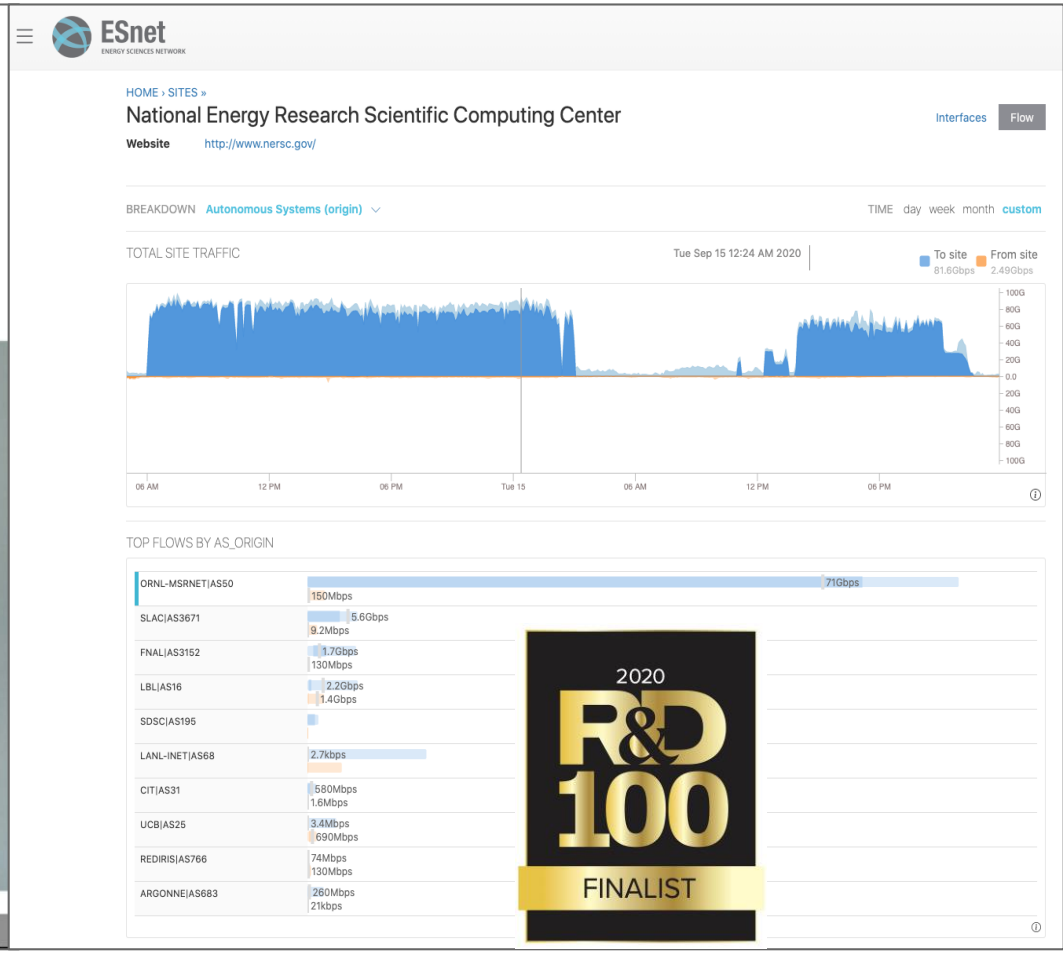


Statistics

- 25 hires in FY20
- 15 hires since March 1st
- 4 employees left, including 2 retirements
- 4 - 6 students hired and mentored every summer
- 87 total employees

ESnet

Quality of staff continues to be excellent

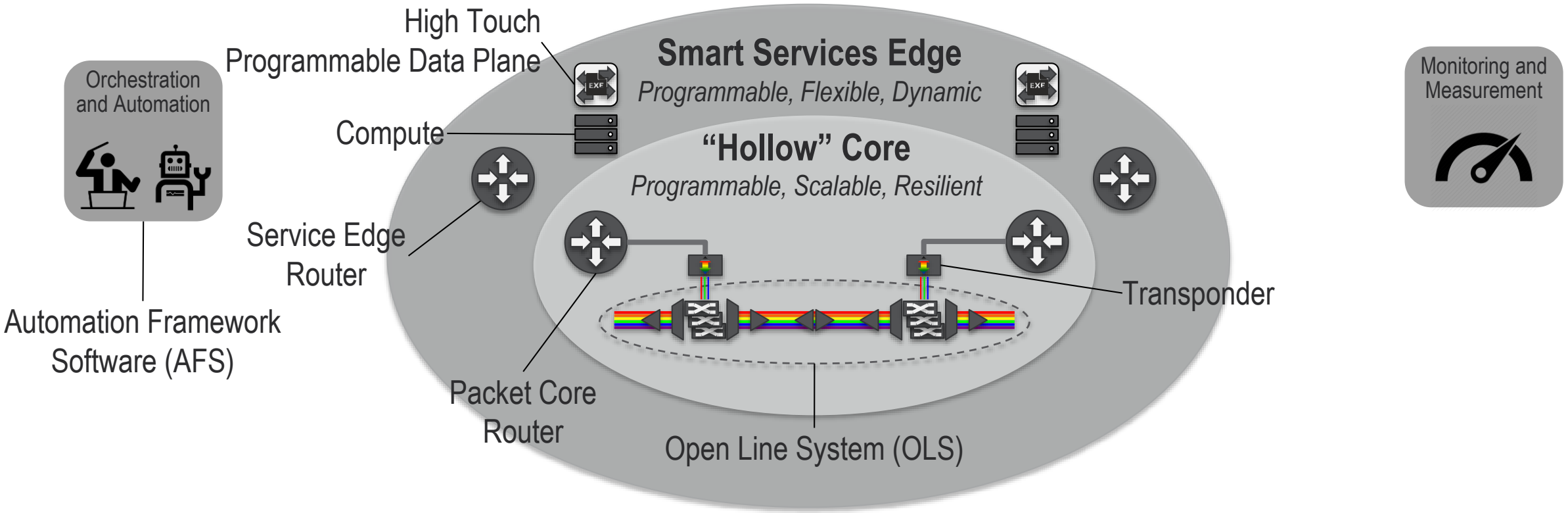


ESnet

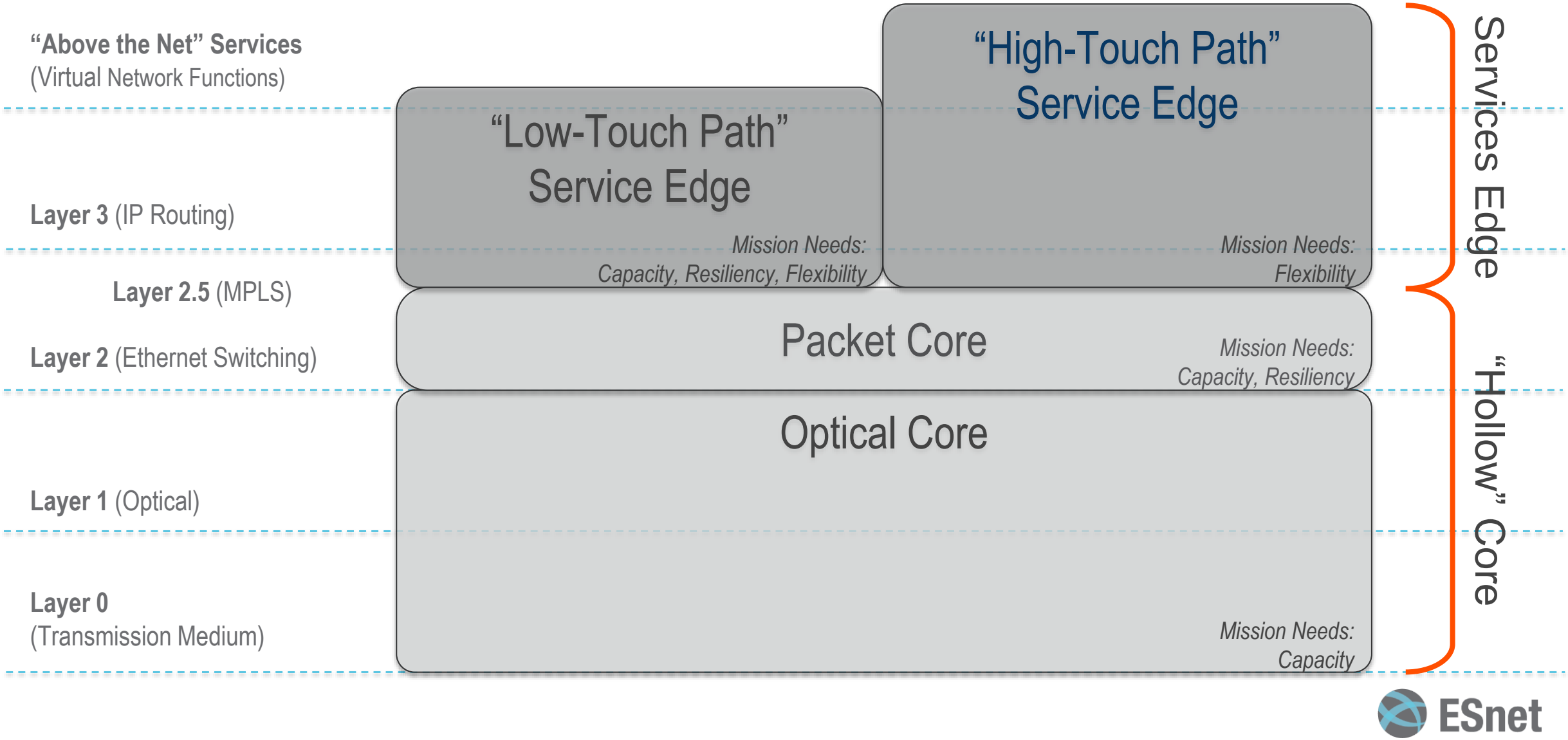
Key takeaways

ESnet6 is Foundational high-performance, flexible platform for scientific discovery and innovation

ESnet6 architecture focuses on ability to deliver terabit-scale performance with programmability and custom science services

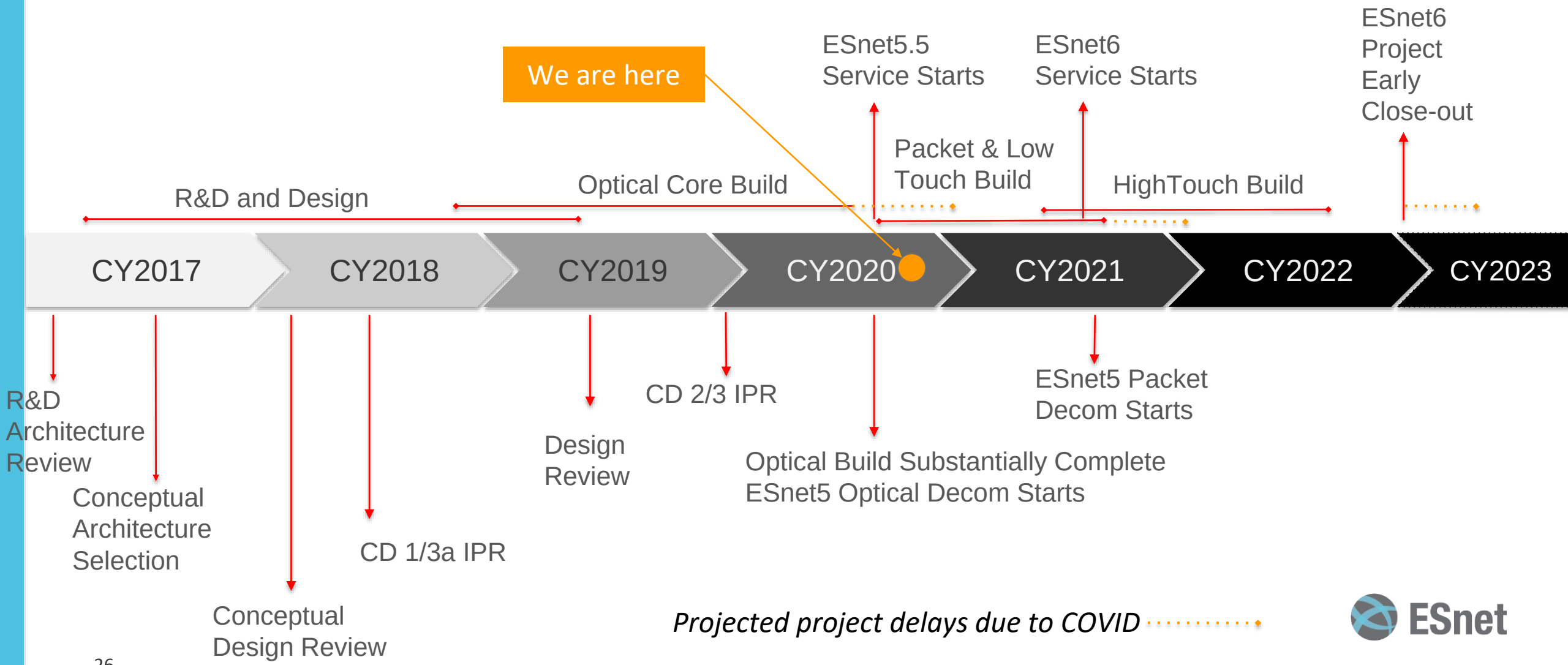


ESnet6 Architecture Components (Layered View)

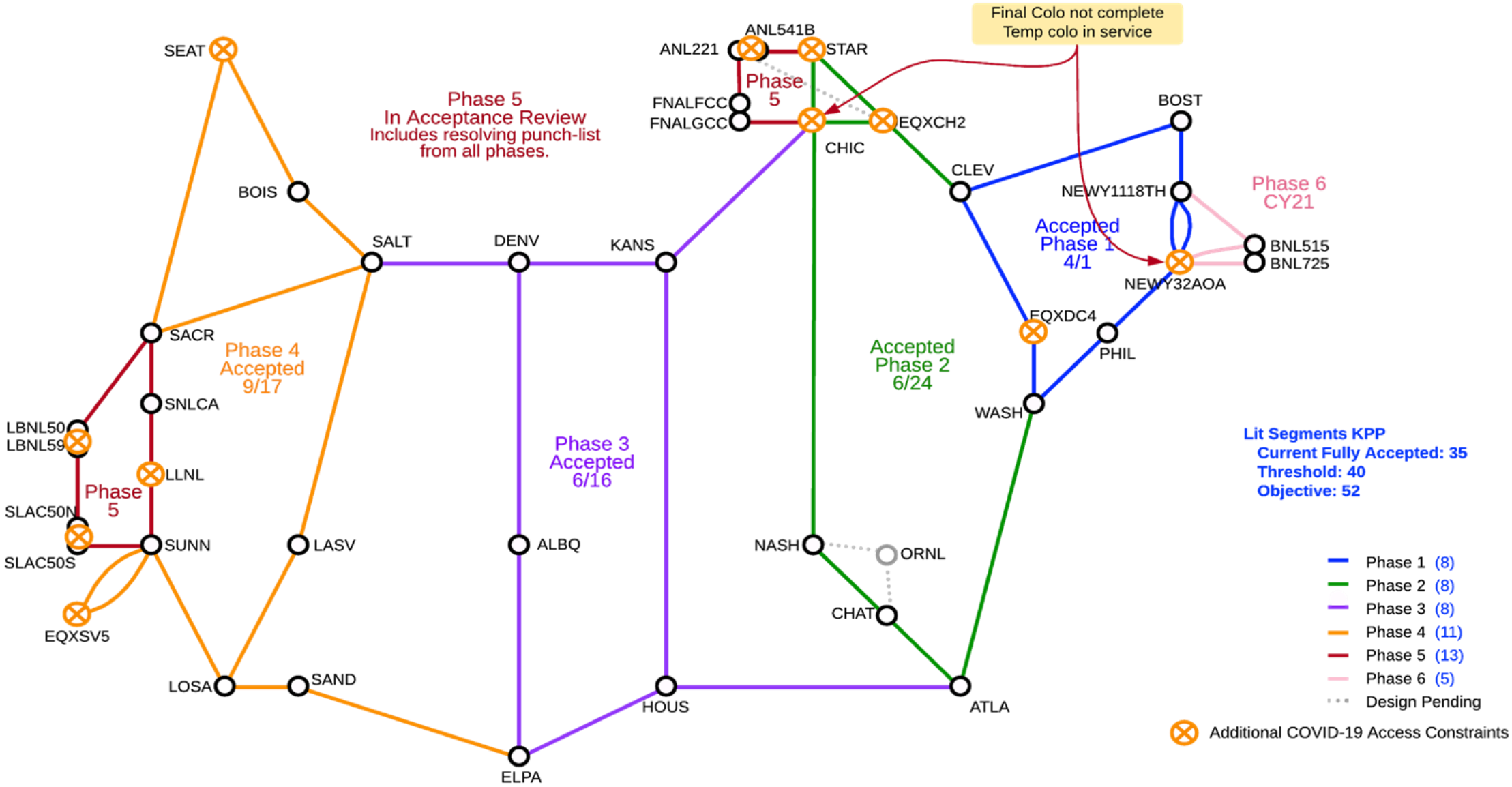


ESnet

ESnet6 Project Implementation Timeline

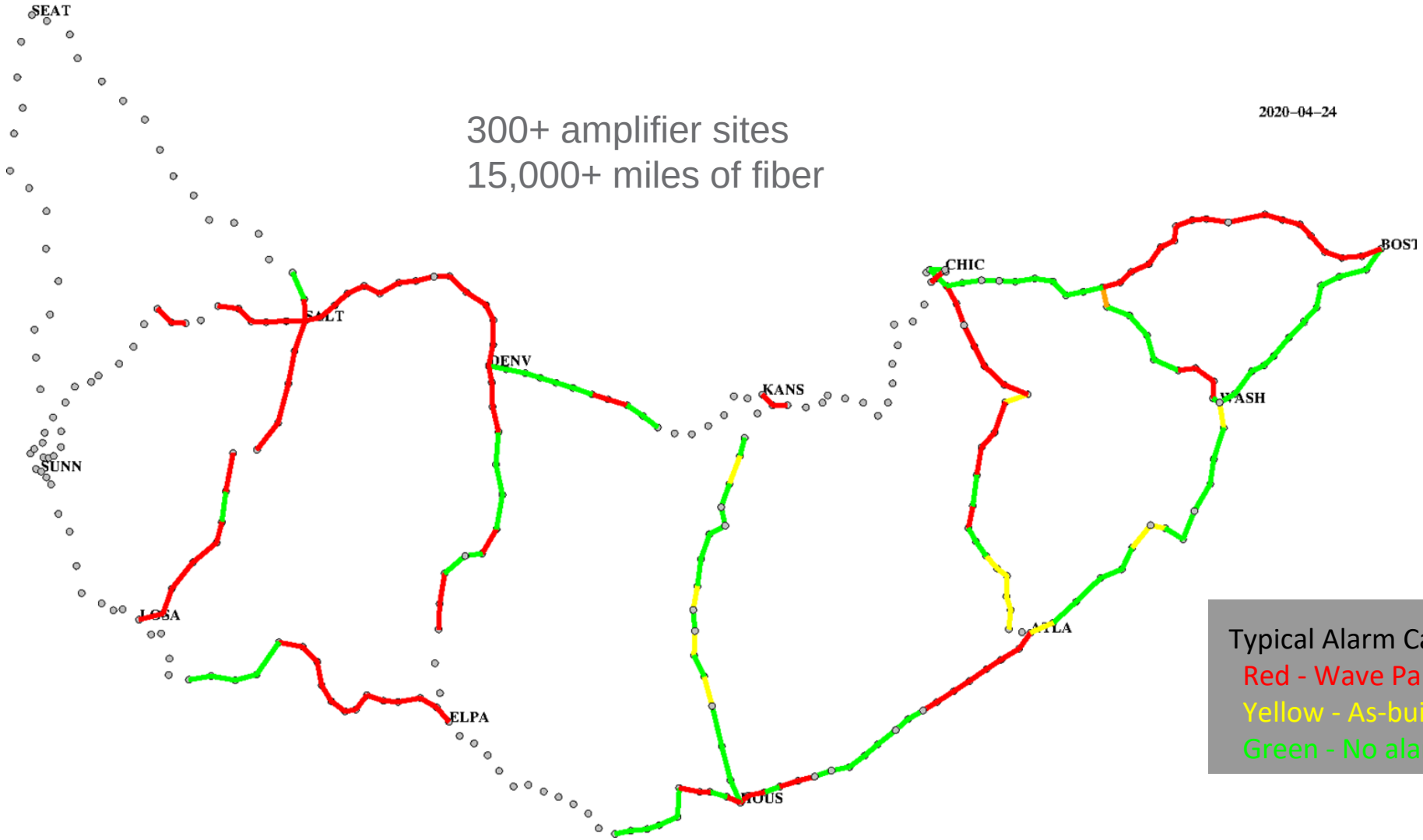


ESnet6 OLS Status
As Of September 22 2020



Apr 28 - Aug 19th: Building the Optical Substrate

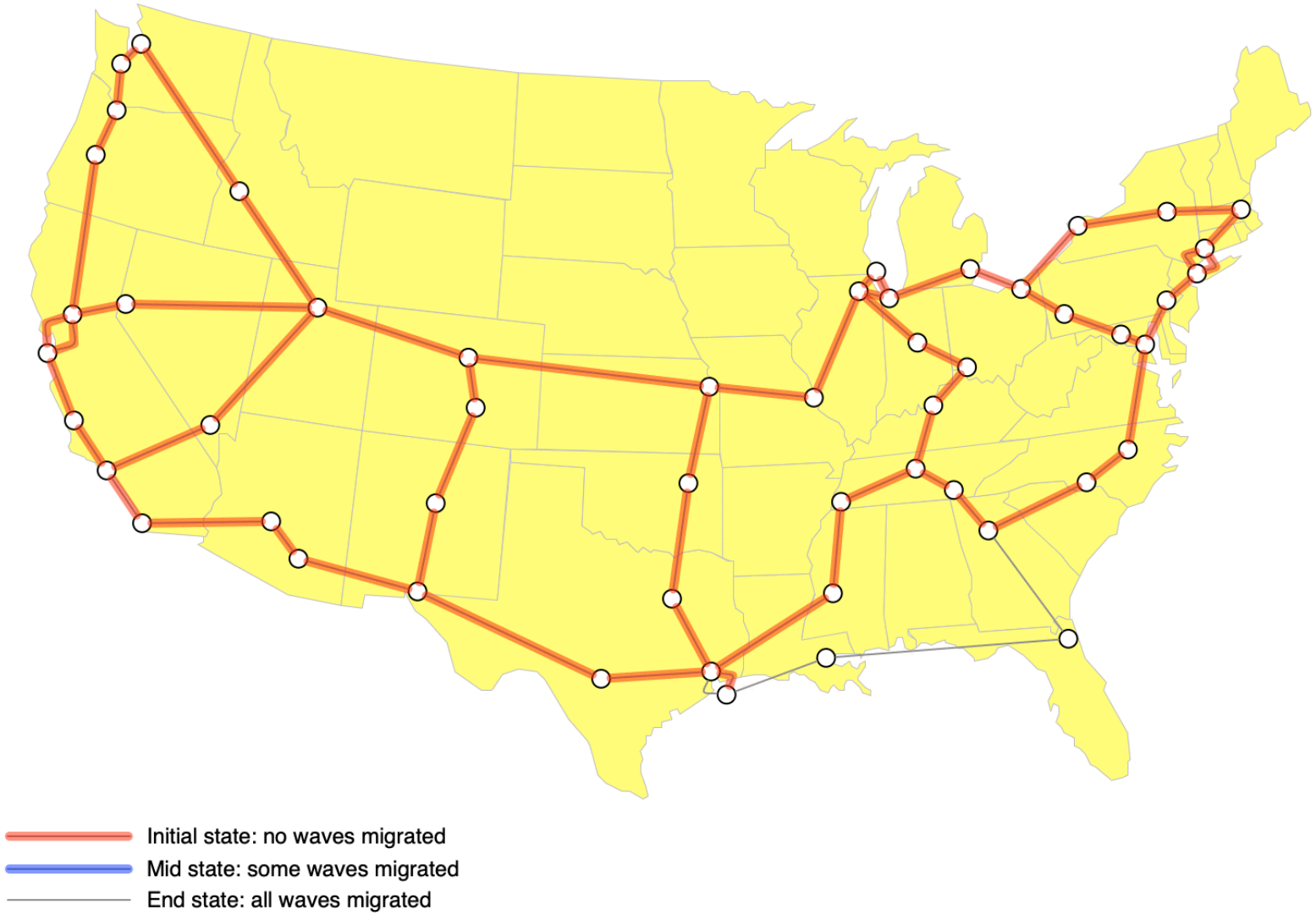
300+ amplifier sites
15,000+ miles of fiber



Typical Alarm Causes
- Red - Wave Partially Provisioned/Functioning
- Yellow - As-built span loss != design span loss
- Green - No alarms on services on this fiber

ESnet

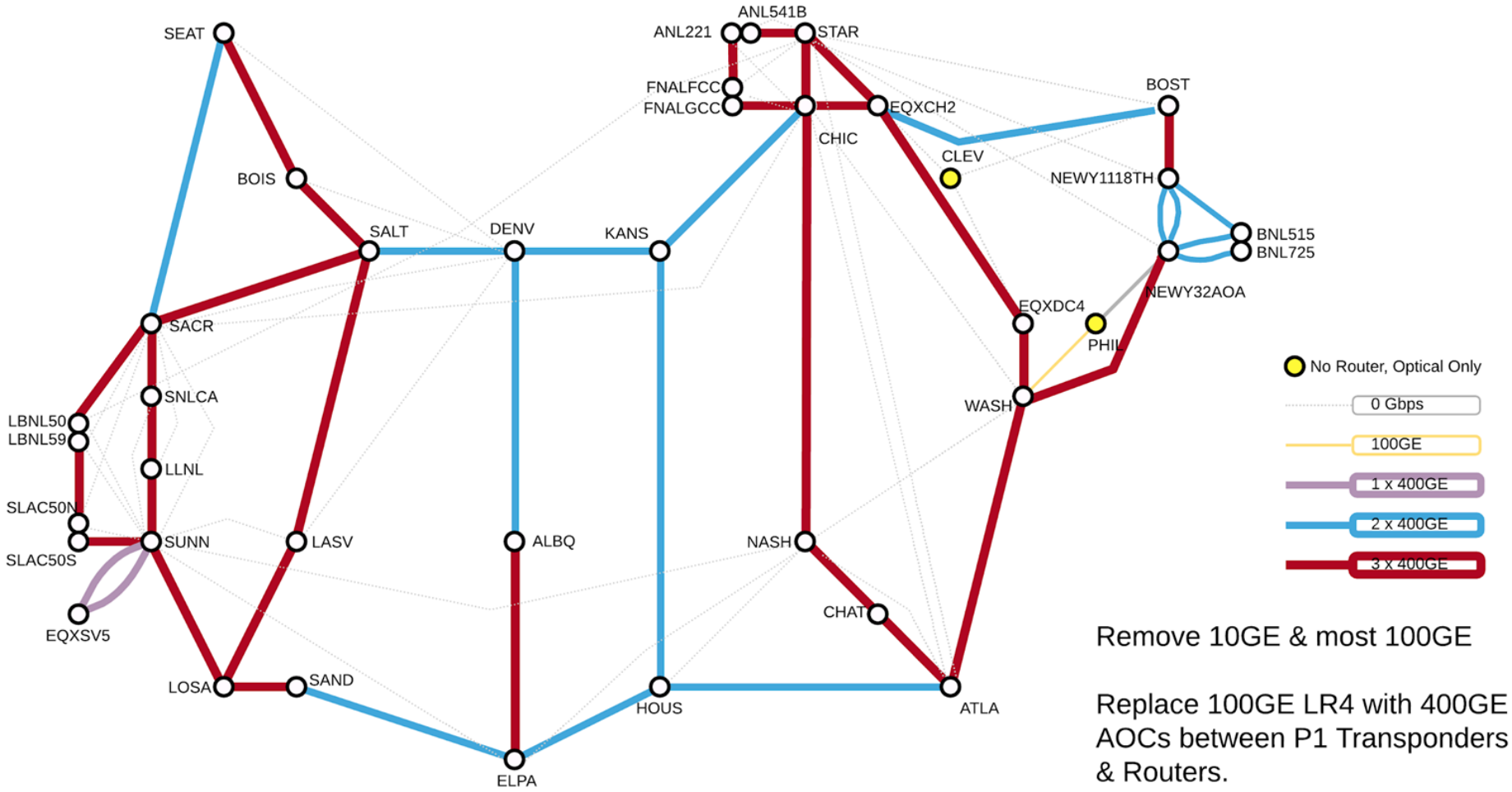
Tearing down ESnet5 Optical happened in parallel



- 75%+ of all production traffic on new optical network
- Rest to be transitioned by Nov. ‘20

ESnet

ESnet 6.3 FY22
High-Touch Deployed
Reconfigured Phase 1 Transponders to 400G Circuits
Concept Diagram, Capacity Plan Not Final!



Remove 10GE & most 100GE

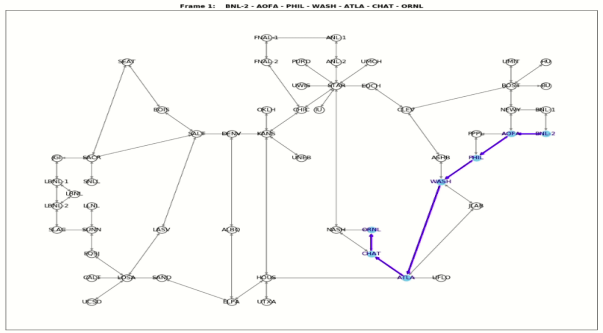
Replace 100GE LR4 with 400GE AOCs between P1 Transponders & Routers.

Reconfigure P1 Transponders to 1x400G wave where possible, otherwise 2x200G waves

Aug 13, 2020 11:33 AM

ESnet

Key takeaways

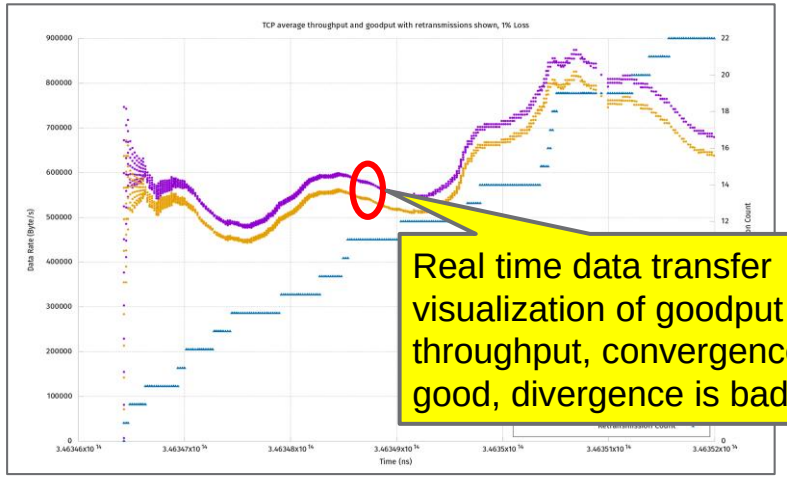


Strong focus on Innovation within the organization with a growing software and prototyping team that matches operational expertise

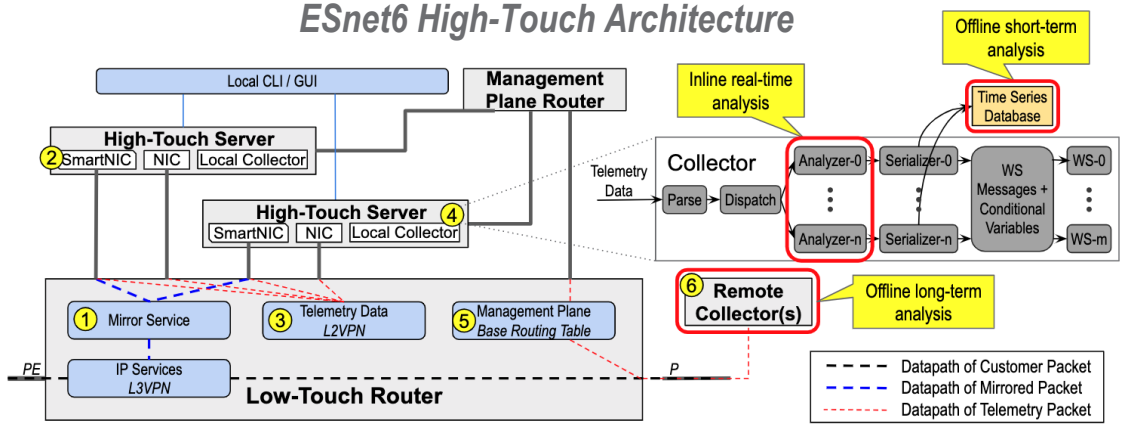
ESnet6 High Touch: Integrating (Edge) Compute, Storage, and SmartNICs with Networking

- **Motivation and Objective**
  Design, develop, and deploy a highly flexible and programmable architecture that integrates (edge) compute, storage, and SmartNICs with networking within the ESnet Wide Area Network (WAN)
- **Significance and Impact**
  - Create new services for science and high-speed networking, currently not available commercially

*Example of an **PacketScope** - an application leveraging (edge) compute, storage, SmartNIC, and networking*



ESnet6 High-Touch Architecture



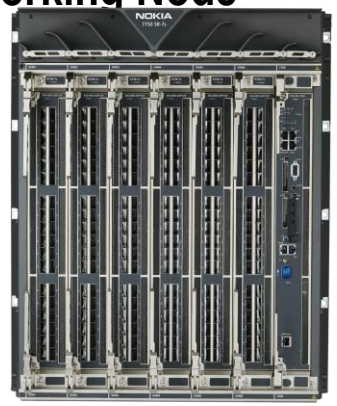
1. Mirror Service - Allows selective flows in the dataplane to be duplicated and sent to the SmartNIC for processing.
2. SmartNIC - Appends meta-data and repackages packet for transmission to Collector code.
3. Telemetry Data L2VPN - Provides option to connect SmartNIC and Collector and bypass PCIe bus if needed.
4. Collector - Performs (limited) in-line real-time analysis as well as inserts telemetry data into database for offline local (short-term 1-2 hr) analysis.
5. Management Plane Base Routing Table - Provides connectivity to remote collector where aggregated telemetry data is sent for offline global analysis.
6. Remote Collector - Stores aggregated telemetry data for long-term global analysis.

ESnet6 Integrated Compute, Storage, SmartNIC, and Networking Node

Compute and Storage

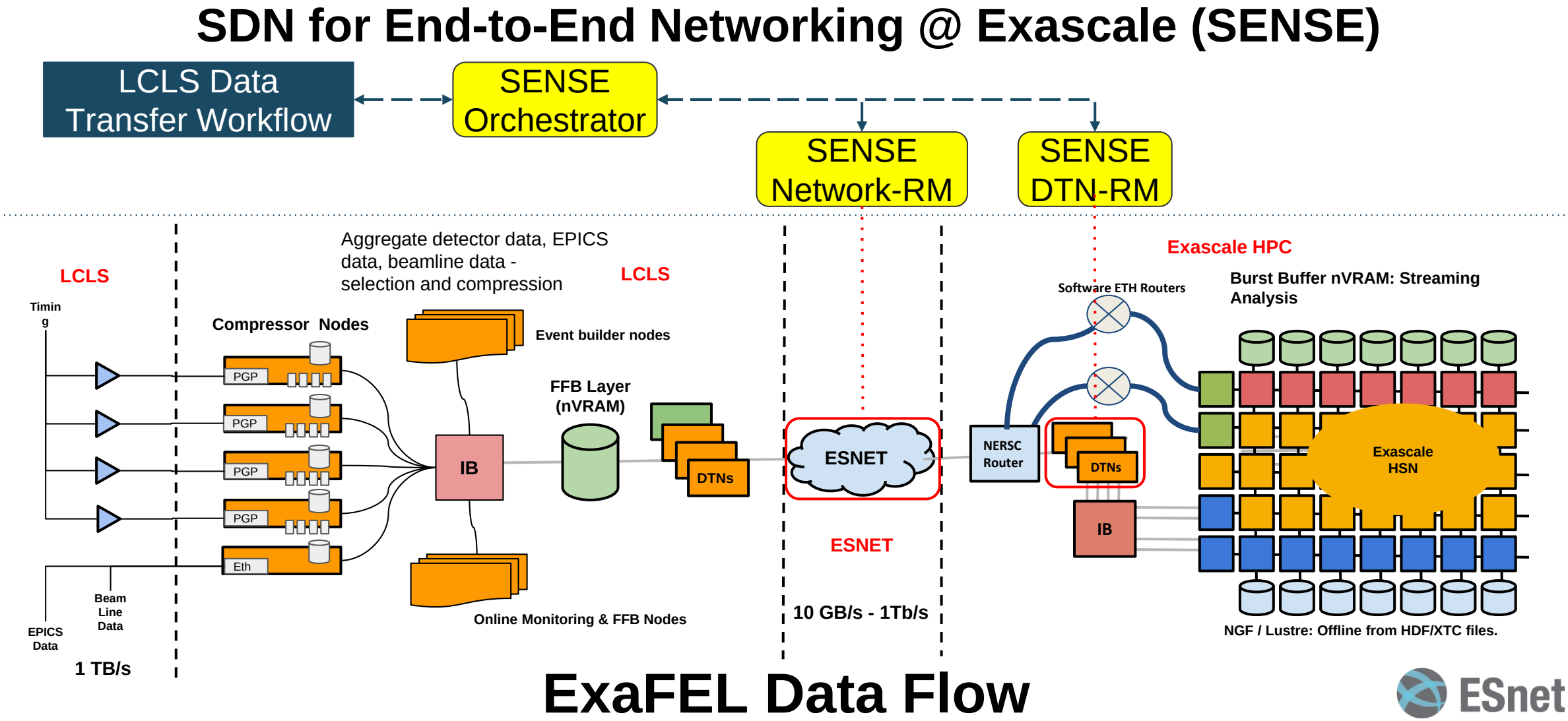
Built in Berkeley

SmartNIC (FPGA)

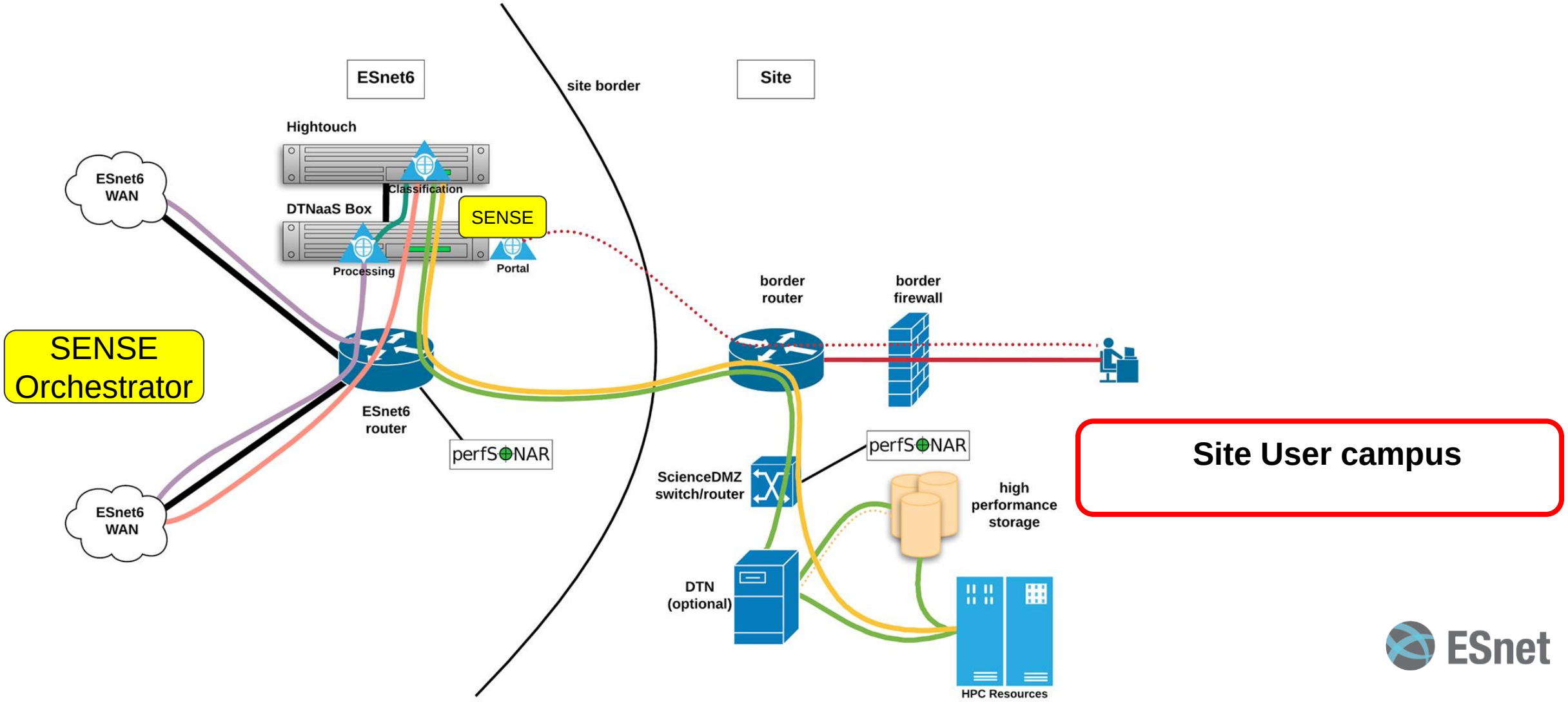


Network Router

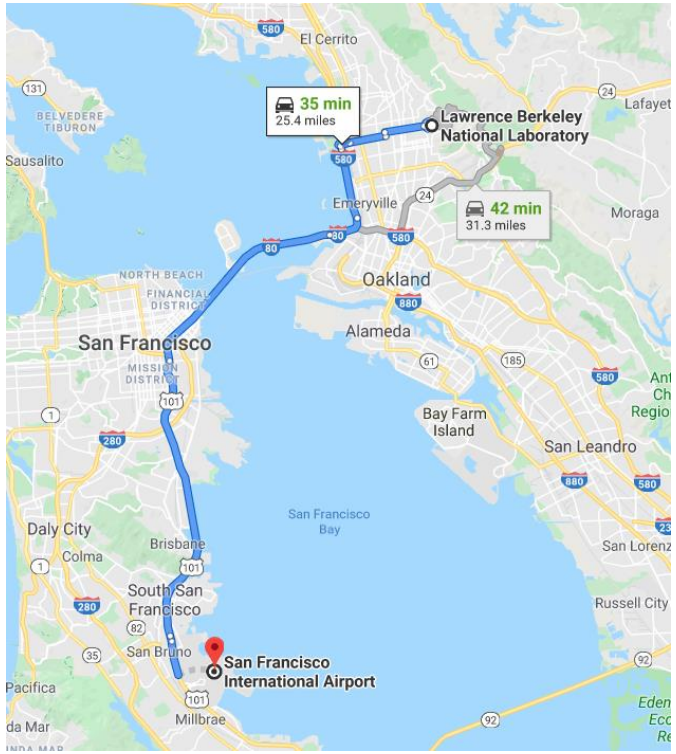
SENSE Automation: ASCR research funded project component of ExaFEL and deployed internationally for HL-LHC experimentation

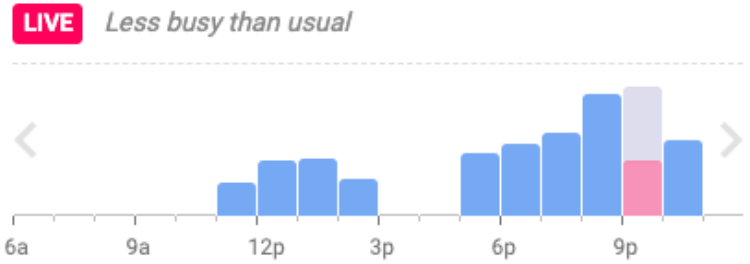


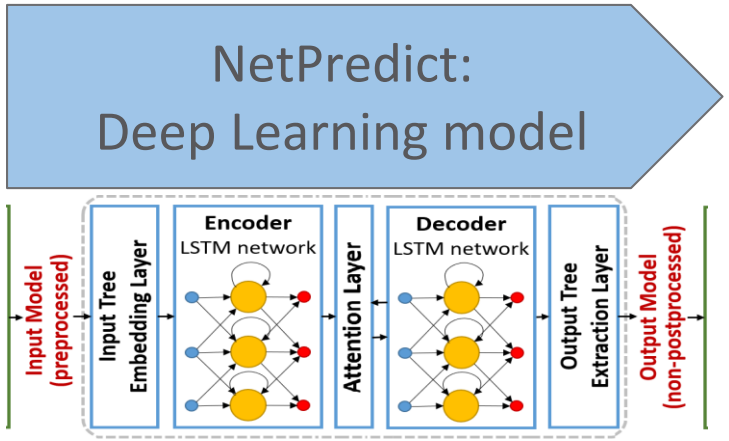
Intelligent Edge helps create new services: Data Transfer Nodes as a Service



Creating Google Maps for networks







Planning your next transfer?

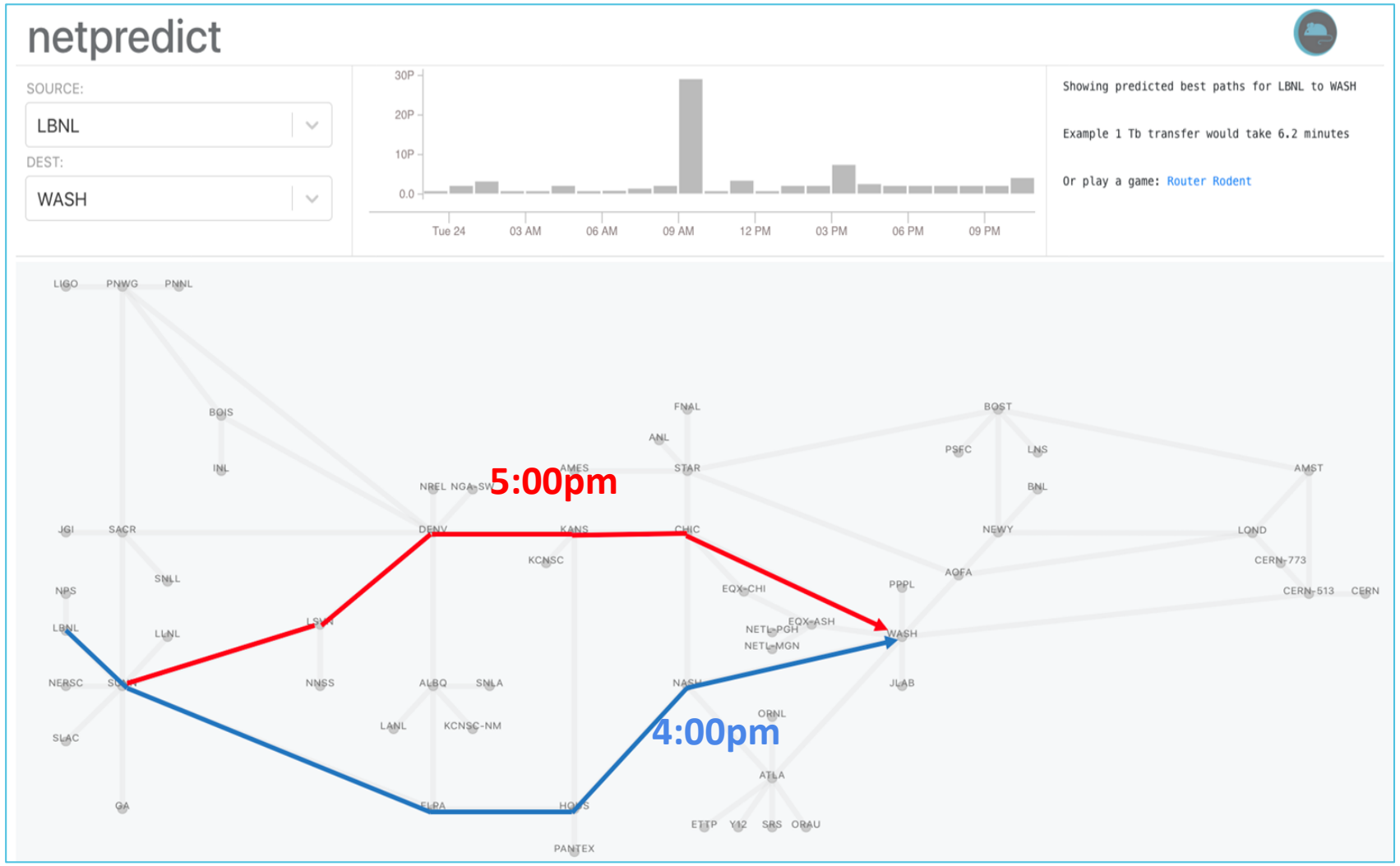
Real-time Data

- PerfSonar (Loss, Throughput)
- Traffic: SNMP data
- Flow behavior: Netflow log

Mariam Kiran
DOE Early Career Researcher

ESnet

NetPredict



Deployed on Google Cloud Platform

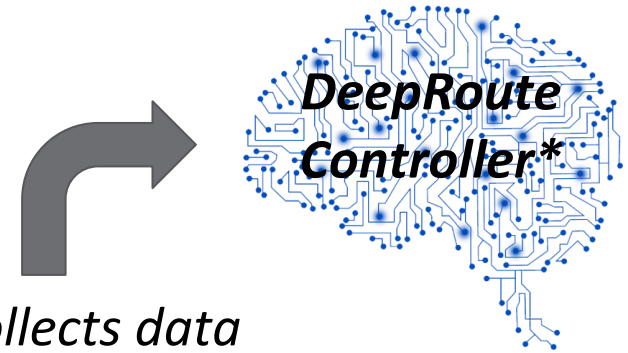
- Different models can run at the same time to compute least congested paths
- Estimates transfer completion time

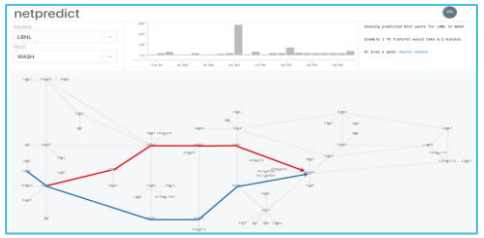
Trust dashboard

- Real-time ML performance
- Build engineer's confidence in predictions

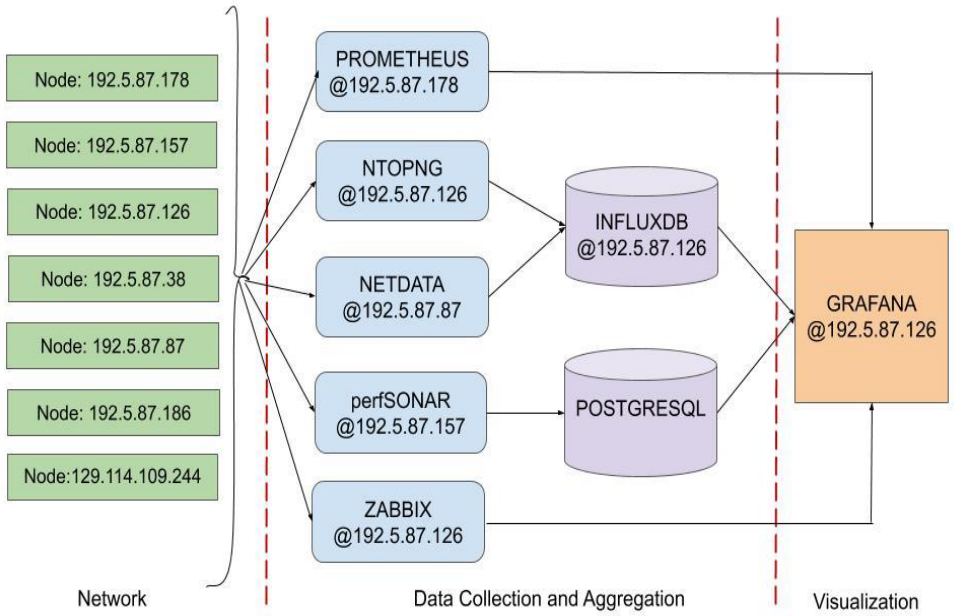
ESnet

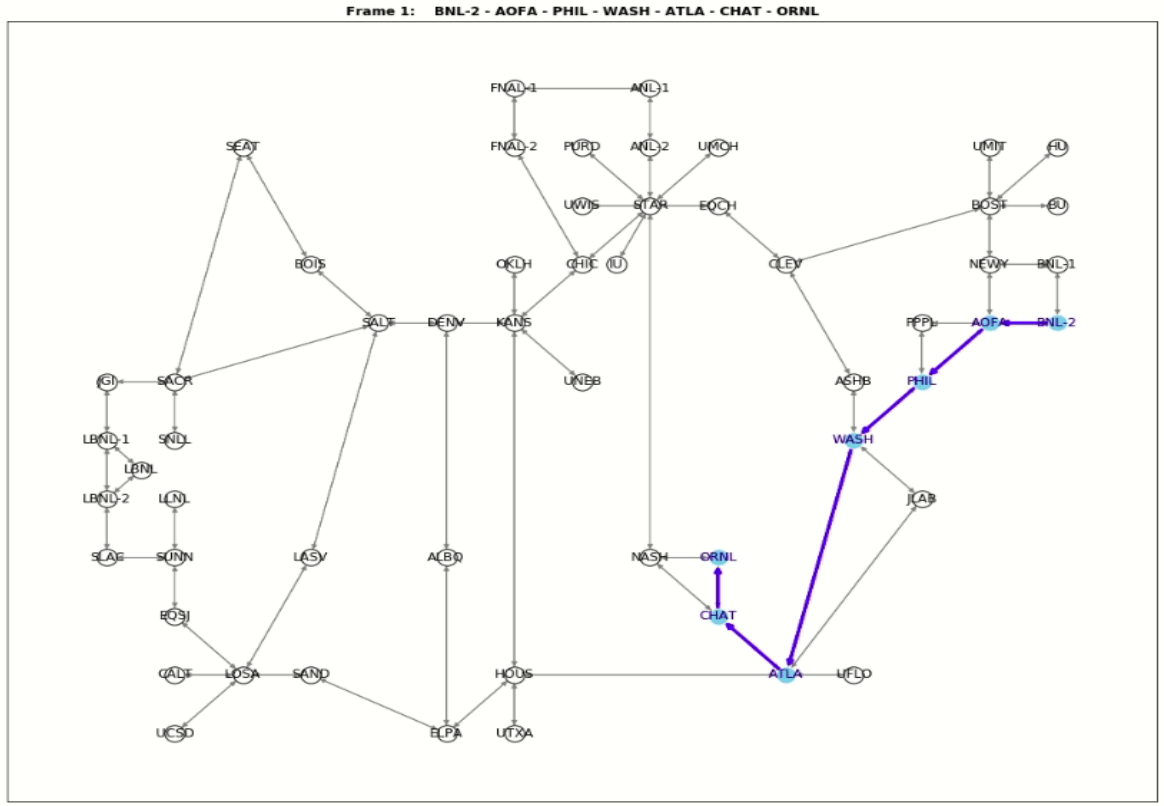
Simulating Intelligent Traffic Engineering





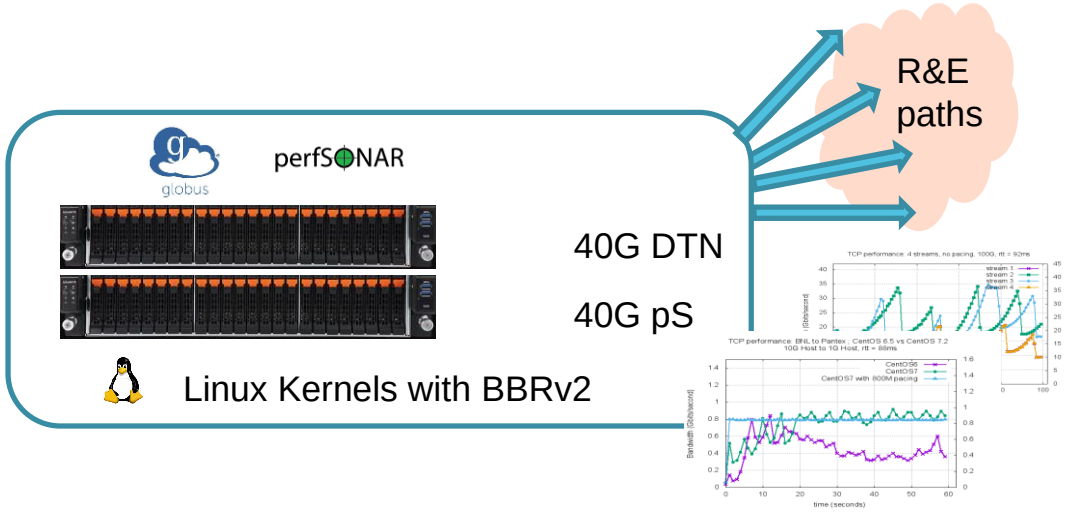
Network Dashboard collects data and feeds to controller





*DeepRoute: M Kiran, B Mohammed, N Krishnaswamy, “Herding Elephant and Mice Flows with Reinforcement Learning”, 2nd IFIP International Conference on Machine Learning for Networking (MLN’2019)

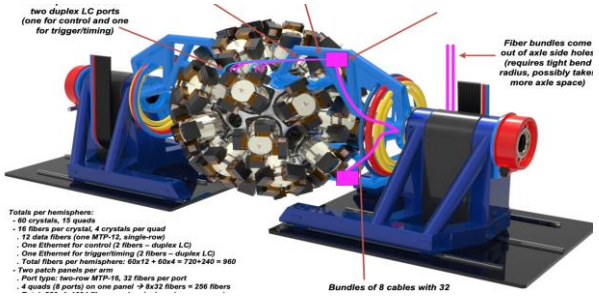
BBRv2 Evaluation (Collaboration)



- Explore BBRv2 developments on ESnet
  - Follow-up on 2016 BBR eval, and renewed discussions with Matt Matthis
  - Understand behavior on R&E networks and share results with community
  - Anticipate future small-buffer, high-BDP environments and wider adoption
- Leverage 40G DTN/perfSONAR deployments at Boston and El Paso
  - New dev workflow for installing and maintaining BBRv2 patches
- Project status: Just Started

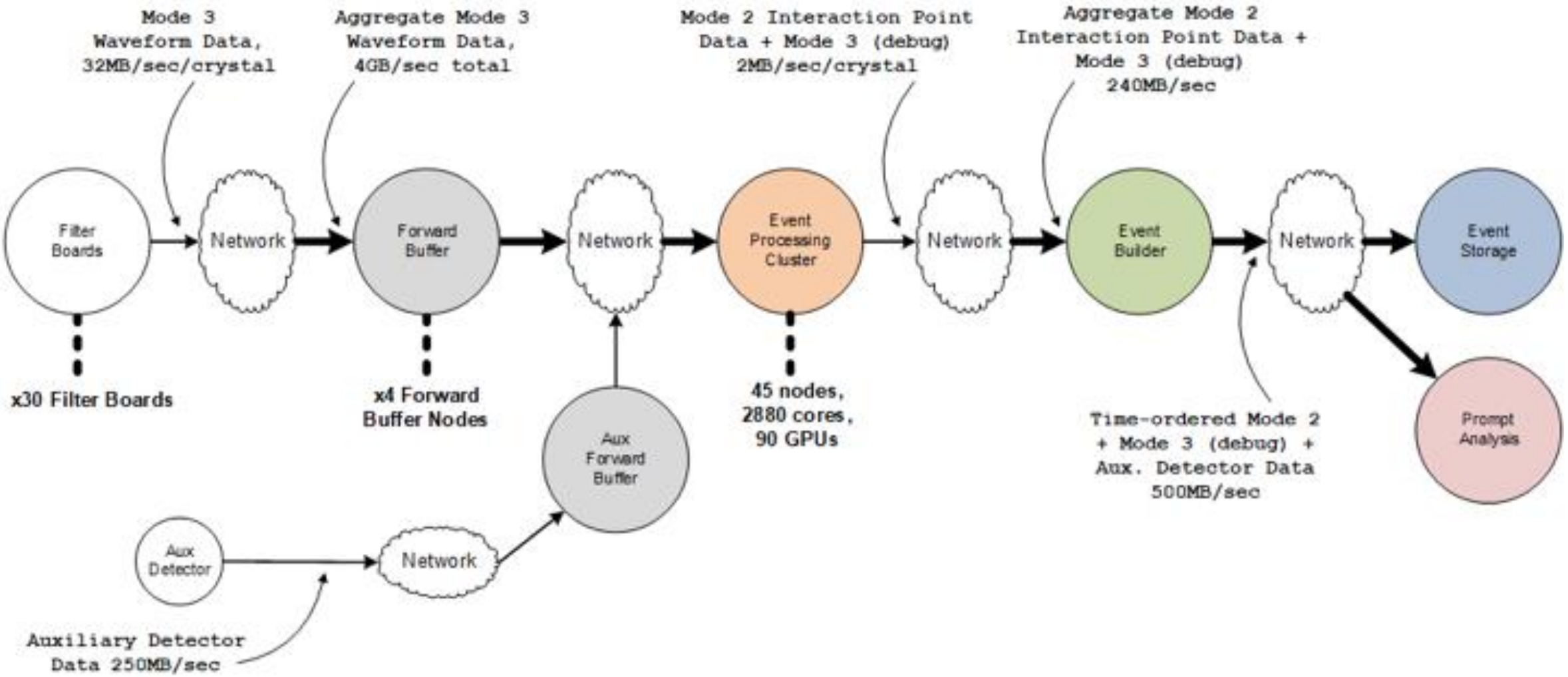
https://fasterdata.es.net/science-dmz/DTN/reference-implementation/

Key takeaways



Networking Codesign gaining momentum in new instrument and facility designs

ESnet responsible for Network and Fiber Design and Forward Buffer design of GRETA Data Pipeline

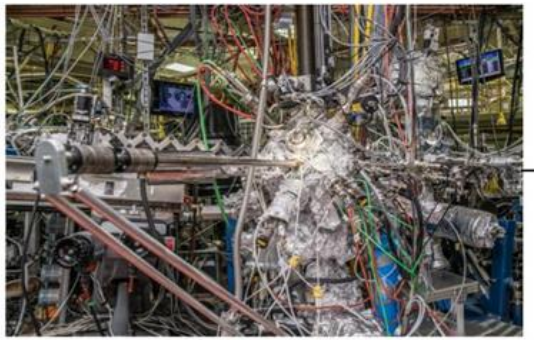


- Custom cable assembly, 32 fibers → two each MTP-12, 4 each duplex LC
. One 2-row 32-fiber MTP-16 port (one cable) serves two crystals
. Two cables per quad

ESnet

Berkeley Extensible Processor (BXP) LDRD: innovative approach to combining edge computing and networking for science instruments

Integrated Edge Compute / Layer 1.5 Services



Typical Large Science Instrument

A/D

DSP

Event Filter (more DSP)

DTN

DTN

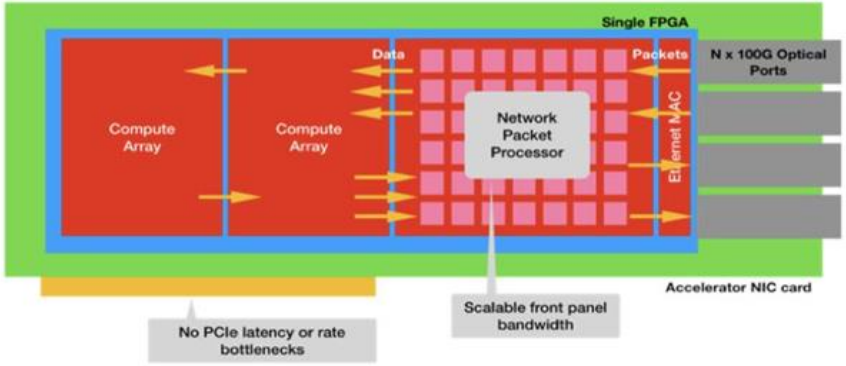
HPC Compute Storage

Replace with FPGA/ Compute

Emerging Acceleration and Networking Hardware

Same Platform
Different Use Case !

Direct to Network
No CPU / DTN / Disk / Router
Rate Limitations



LUDICROUS Speeds

Campus / Metro / WAN

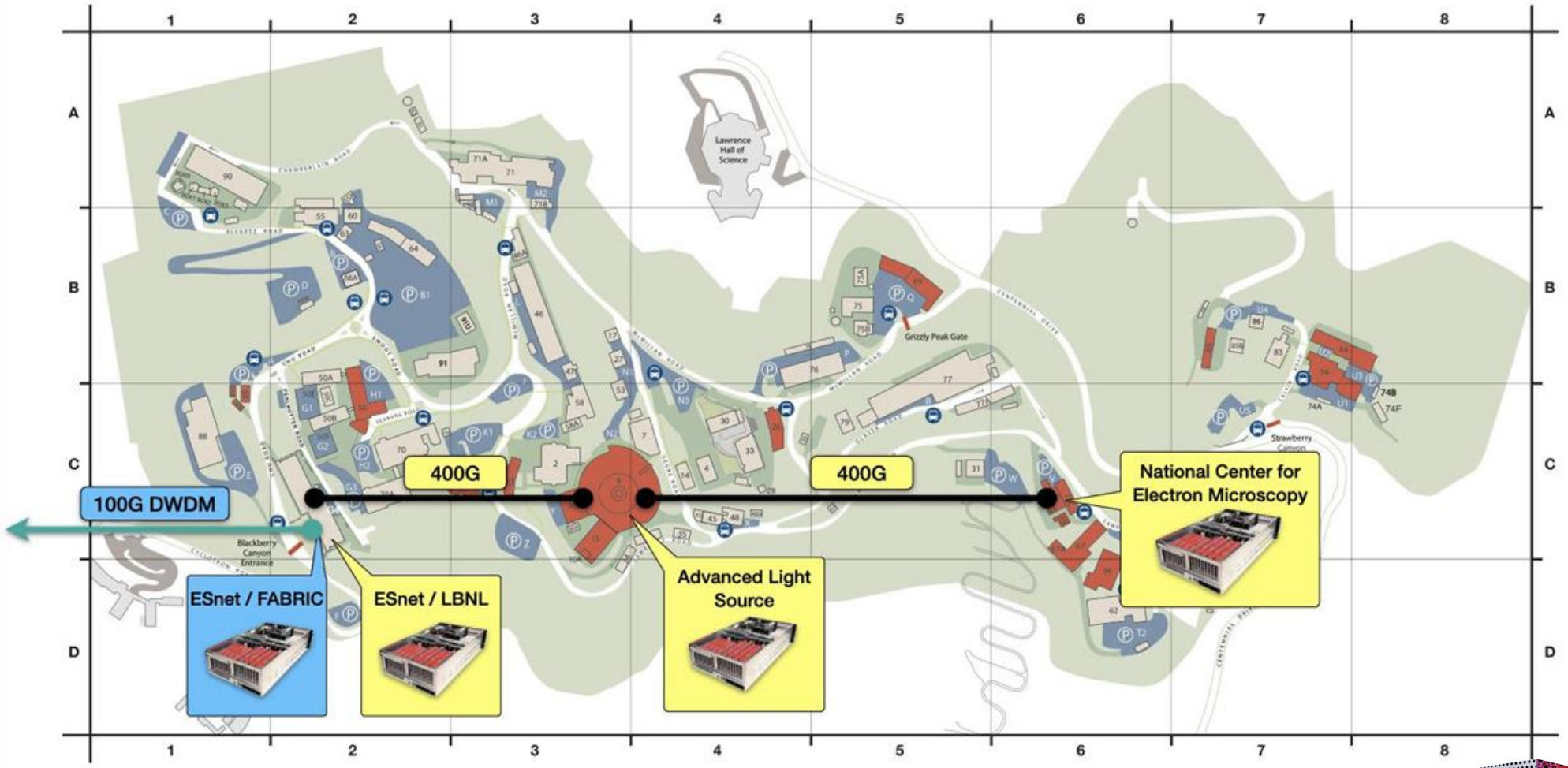
DWDM - No Routers
ROADM / Regen

ESnet

CRD

ESnet

Significant deployment planned with ALS and NCEM



8 Card Server
Dual CPU
8 x PCIe-16 slots

80,000 DSP Cores
1.6 Tbps Ethernet I/O

Composable Stack.
Units can be moved or combined for different use cases.

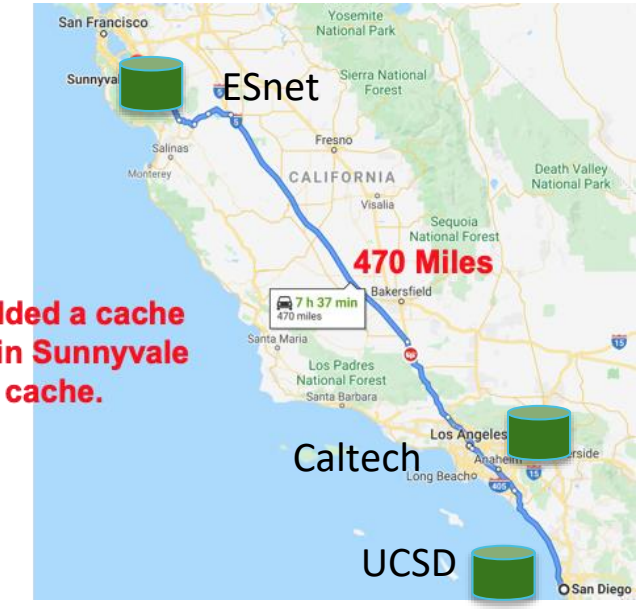
48 FPGA Cards
480,000 DSP Cores Total
9.6 Tbits / sec I/O

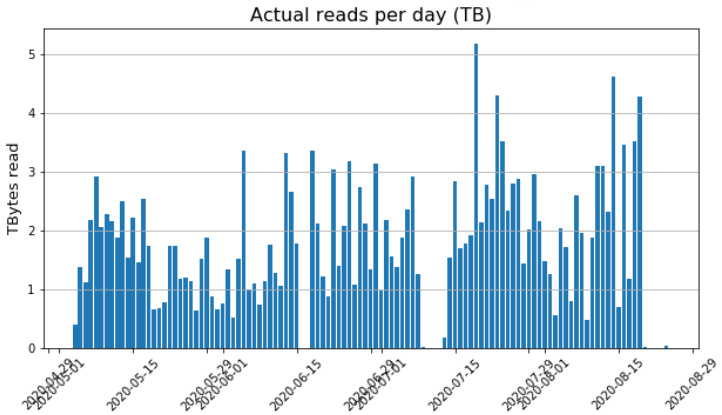
CRD

ESnet

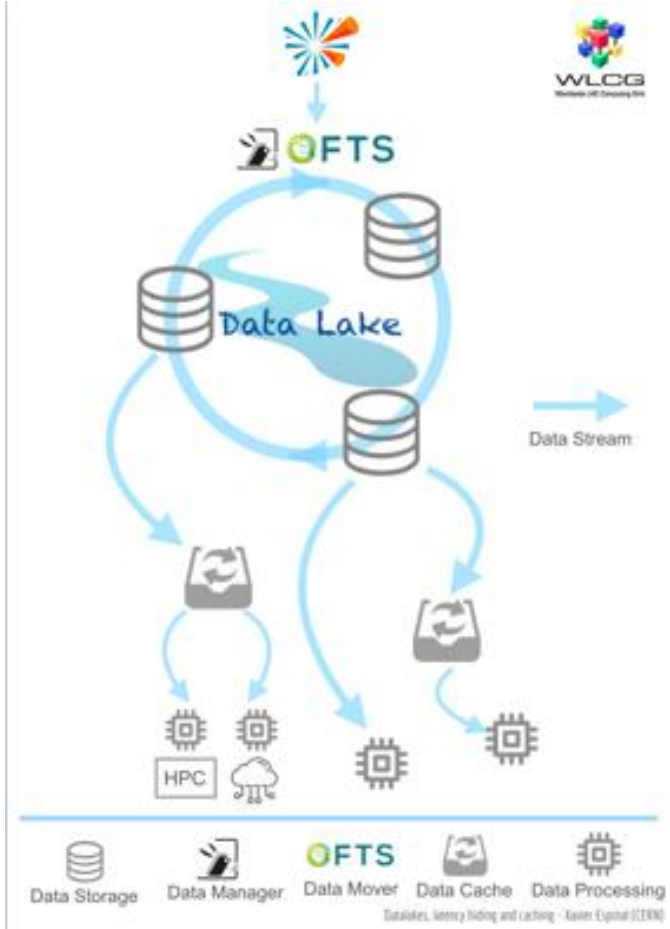
Caching collaboration with HEP CMS and ATLAS

In early May, we added a cache at the ESNet POP in Sunnyvale to the SoCal cache.





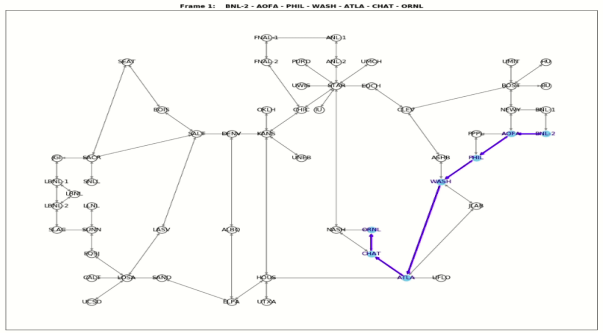
About 1.5 TB/3500 accesses per day avg.

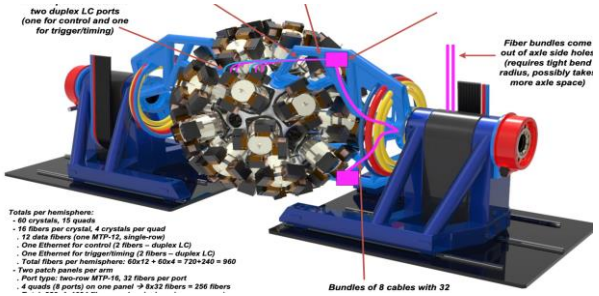


Thanks to Frank Wurthwein, Rob Gardner, Alex Sim, Wei Yang, Harvey Newman, others OSG and SLATE projects

ESnet

Key takeaways





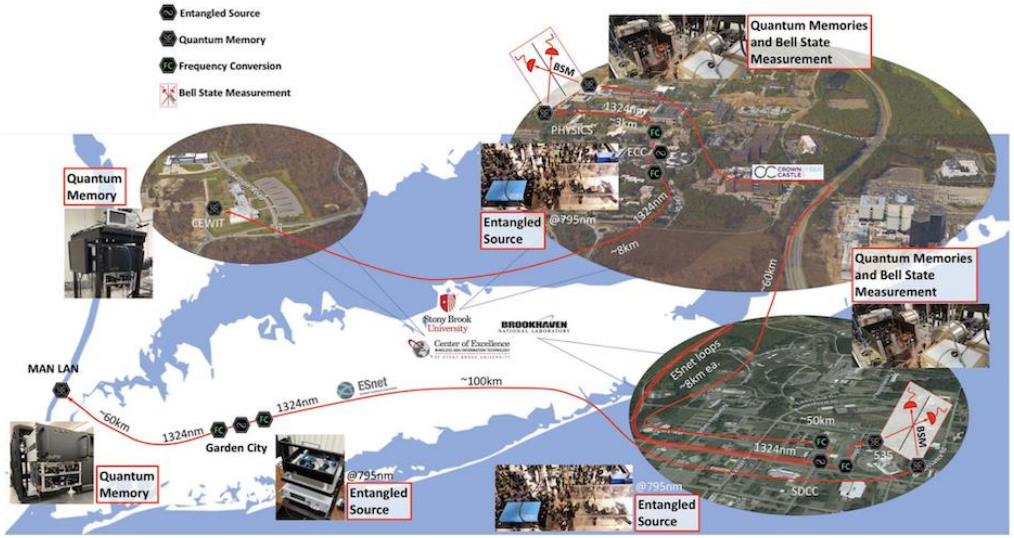
ESnet continues to execute well as a User Facility despite COVID

ESnet6 is Foundational high-performance, flexible platform for scientific discovery and innovation

Strong focus on Innovation within the organization with a growing software and prototyping team that matches operational expertise

Networking Codesign gaining momentum in new instrument and facility designs

Epilogue: ESnet supporting DOE / national strategy on Quantum and 5G



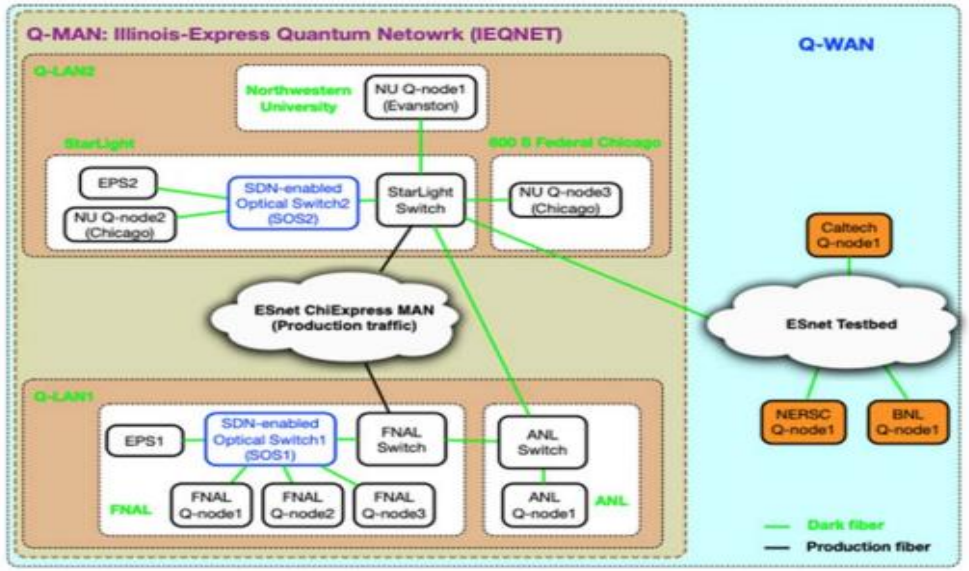
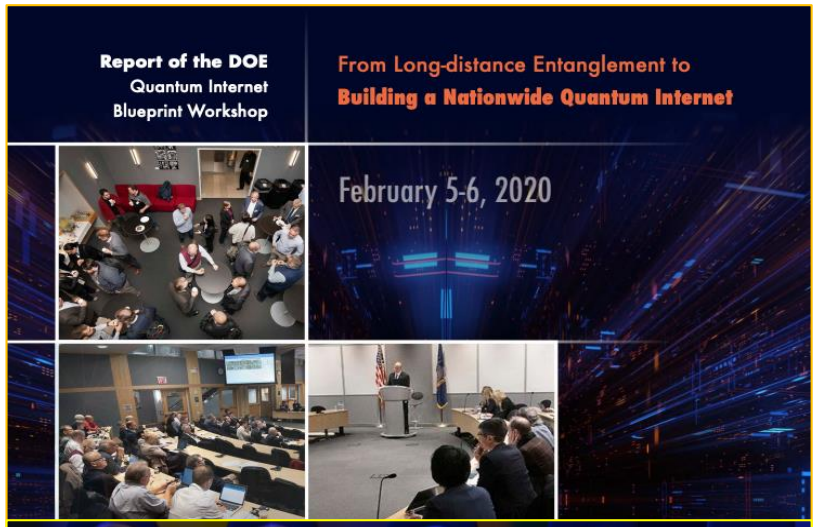


Figure 3:2. IEQNET Topology



ESnet

NSF FABRIC is being built on ESnet6 architecture

- A National-Scale Programmable Experimental Network Infrastructure
- Revolutionize Internet architectures with @scale integration of in-network compute, storage, accelerators with high-speed optical

This work is funded by NSF grant: CNS-1935966

FABRIC

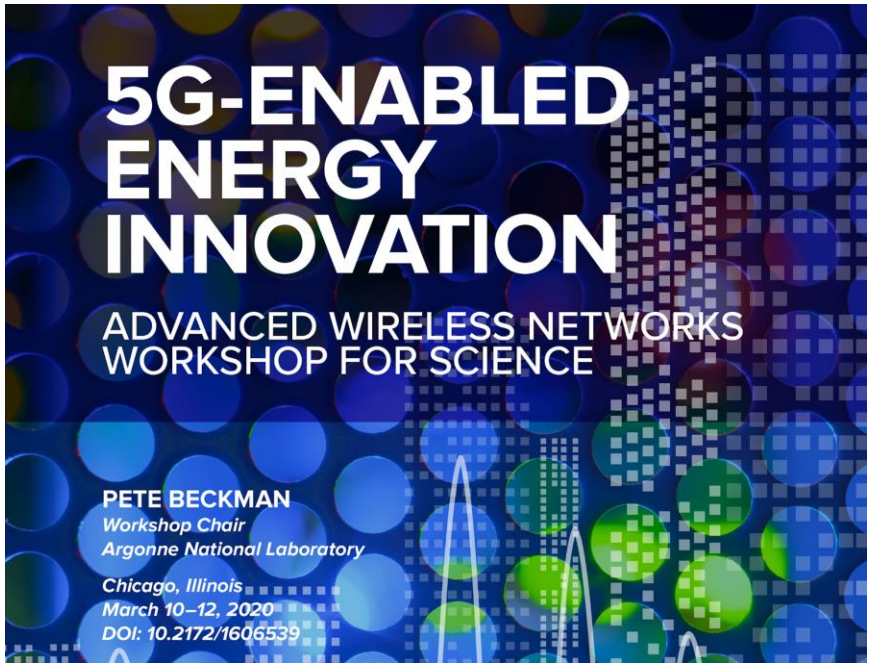
Questions...

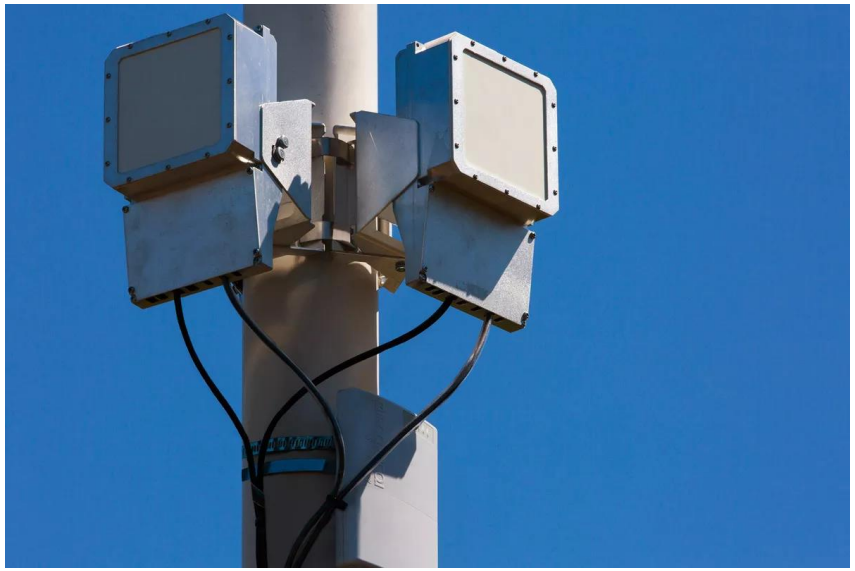
ESnet

Backup Slides

ESnet

Epilogue: ESnet supporting DOE / national strategy on Quantum and 5G/6G





ESnet