ESnet
ENERGY SCIENCES NETWORK

ESnet Status Update

Paul Wefel
Network engineer
Lawrence Berkeley National Laboratory

SA3CC - Virtual meeting

April 2022

U.S. DEPARTMENT OF ENERGY
Office of Science

BERKELEY LAB

Agenda

- **Quick overview of ESnet**
- ESnet6 Project Implementation Update
  - Backbone upgrades
  - Security black hole service
- Automation
- Monitoring
- Trans Atlantic upgrades
- Questions

ESnet

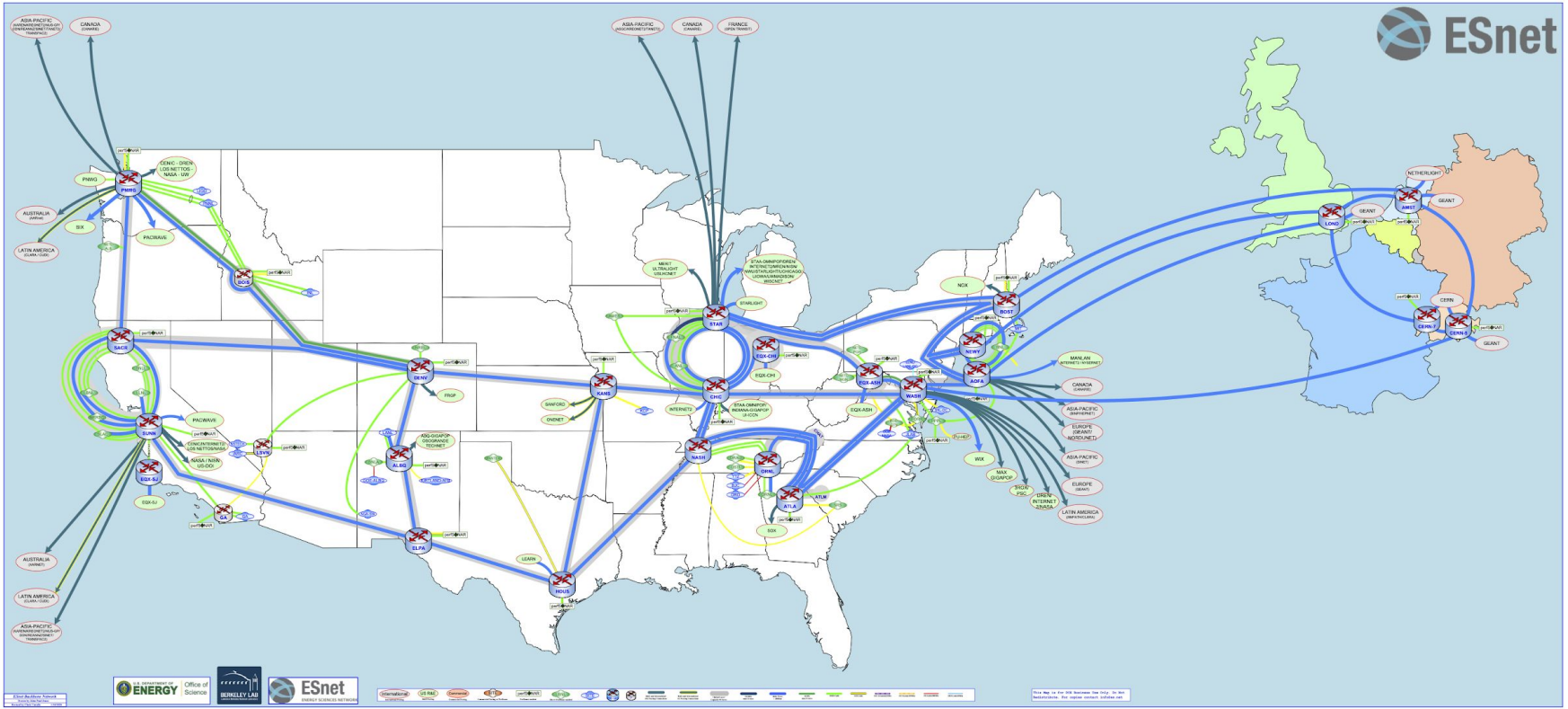
ESnet: DOE's High Performance Network (HPN) Scientific User Facility and derives its mission from Office of Science

Mission of DOE Office of Science:
...delivery of scientific discoveries and major scientific tools to transform our understanding of nature...

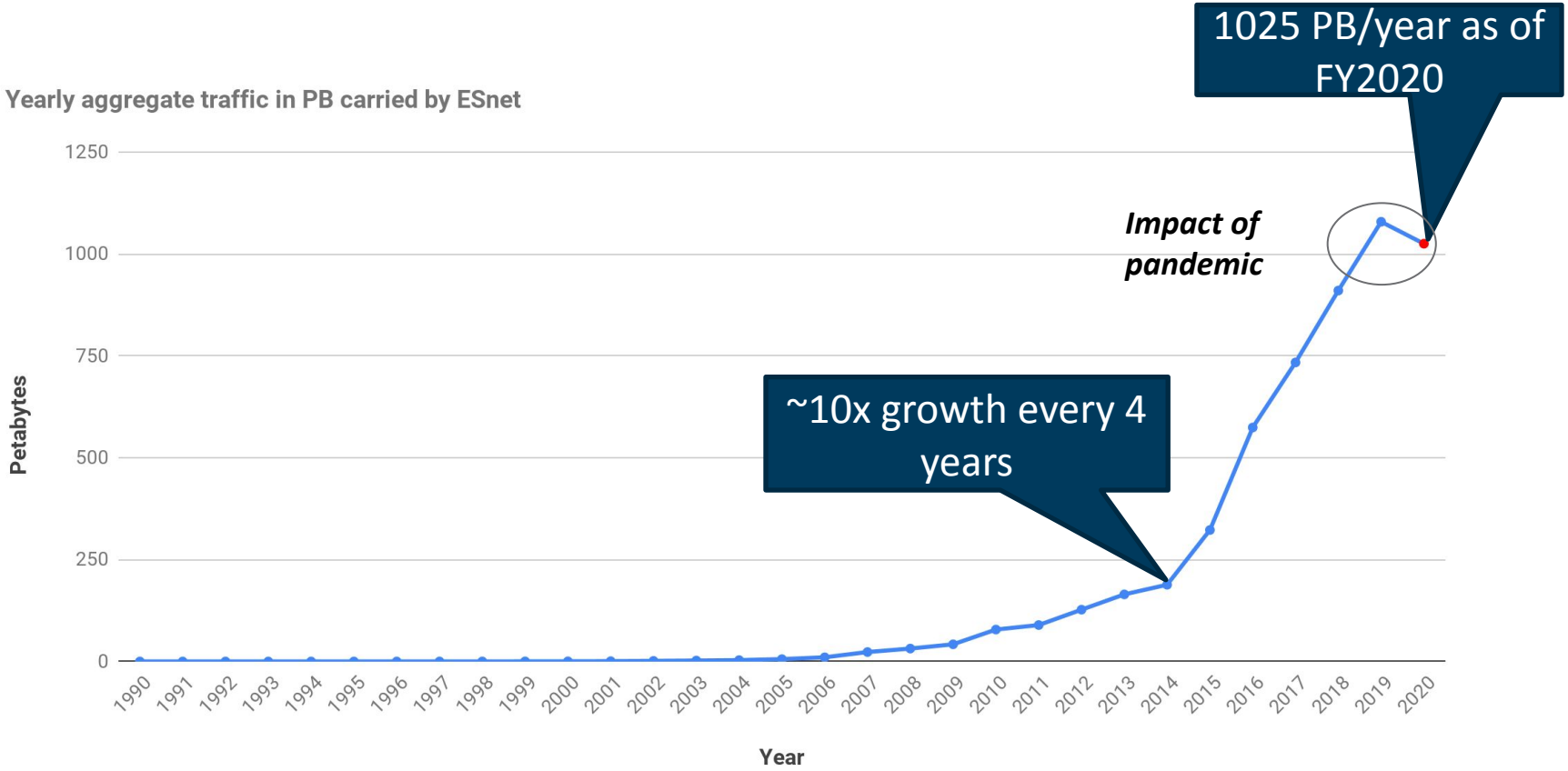
Mission of Energy Sciences Network:
Science network user facility designed to **accelerate** scientific research and discovery.

ESnet

Connects all of the DOE national labs, DOE sites, and hundreds of research and commercial networks internationally.

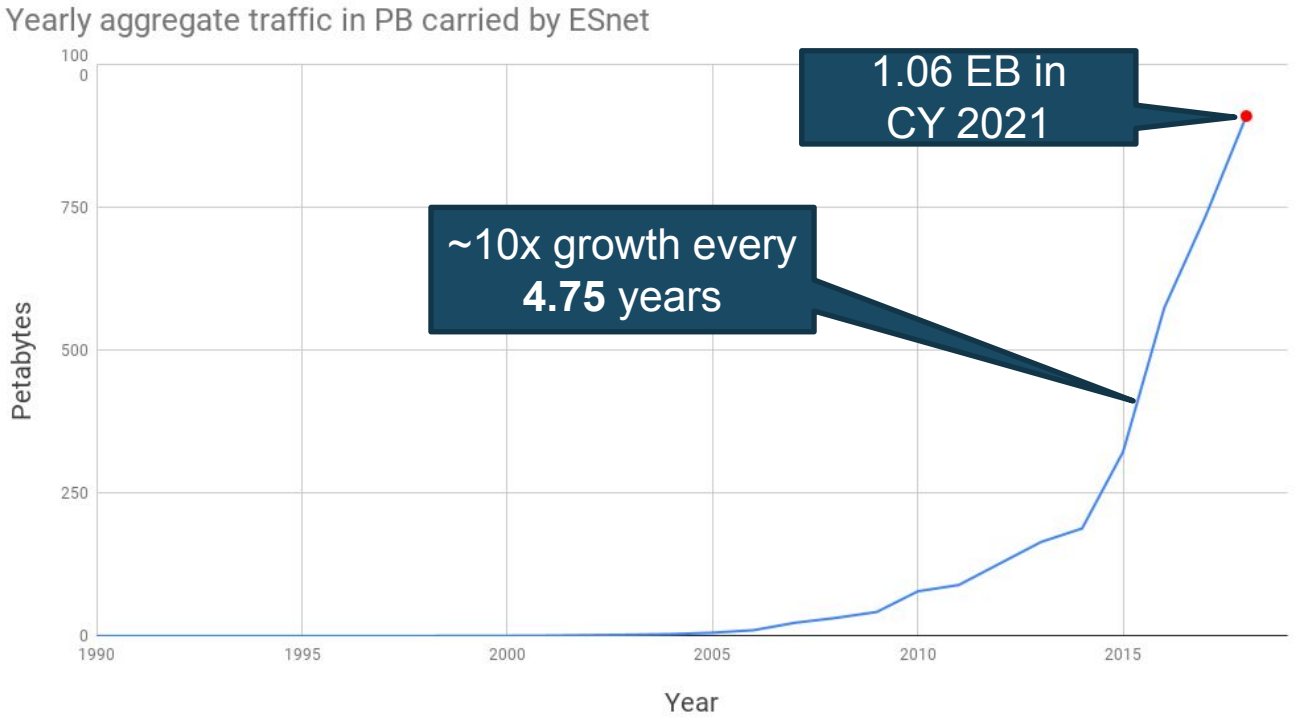


An Exabyte Network Today



Science Applications Take Full Advantage of Well Engineered Networks

Exponential traffic growth over past 28 years



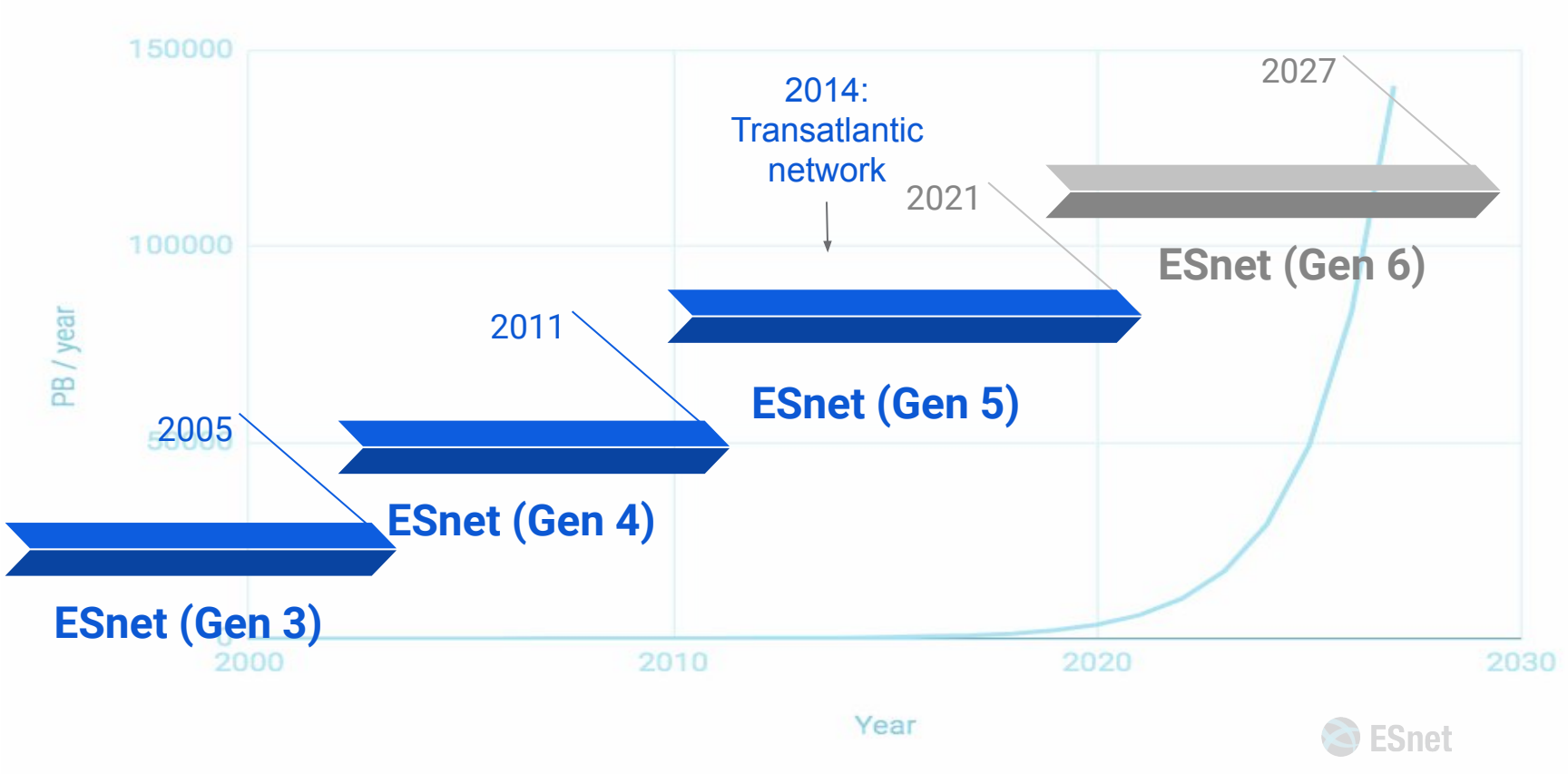
ESnet

Agenda

- Quick overview of ESnet
- **ESnet6 Project Implementation Update**
  - Backbone upgrades
  - Security black hole service
- Automation
- Monitoring
- Trans Atlantic upgrades
- Questions

ESnet

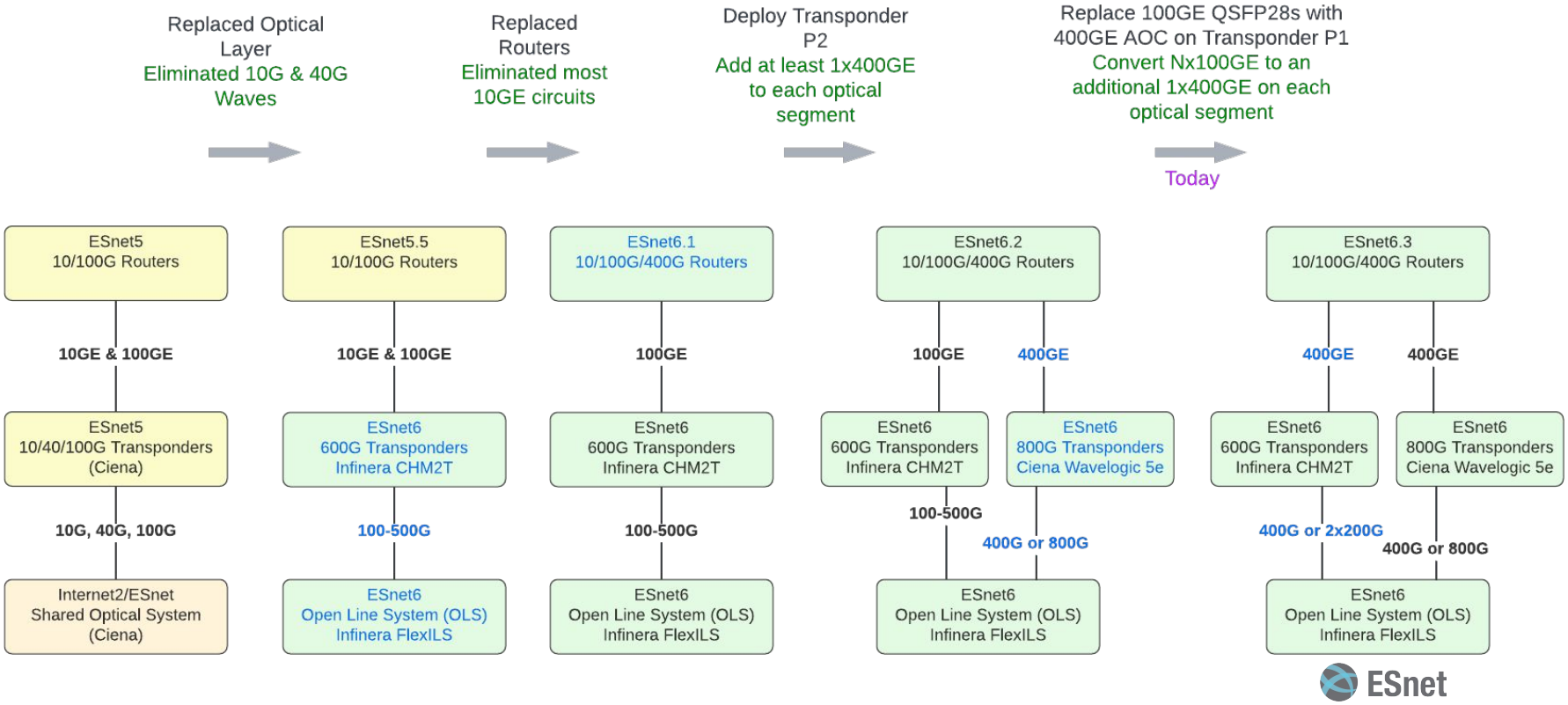
The evolution of the ESnet network



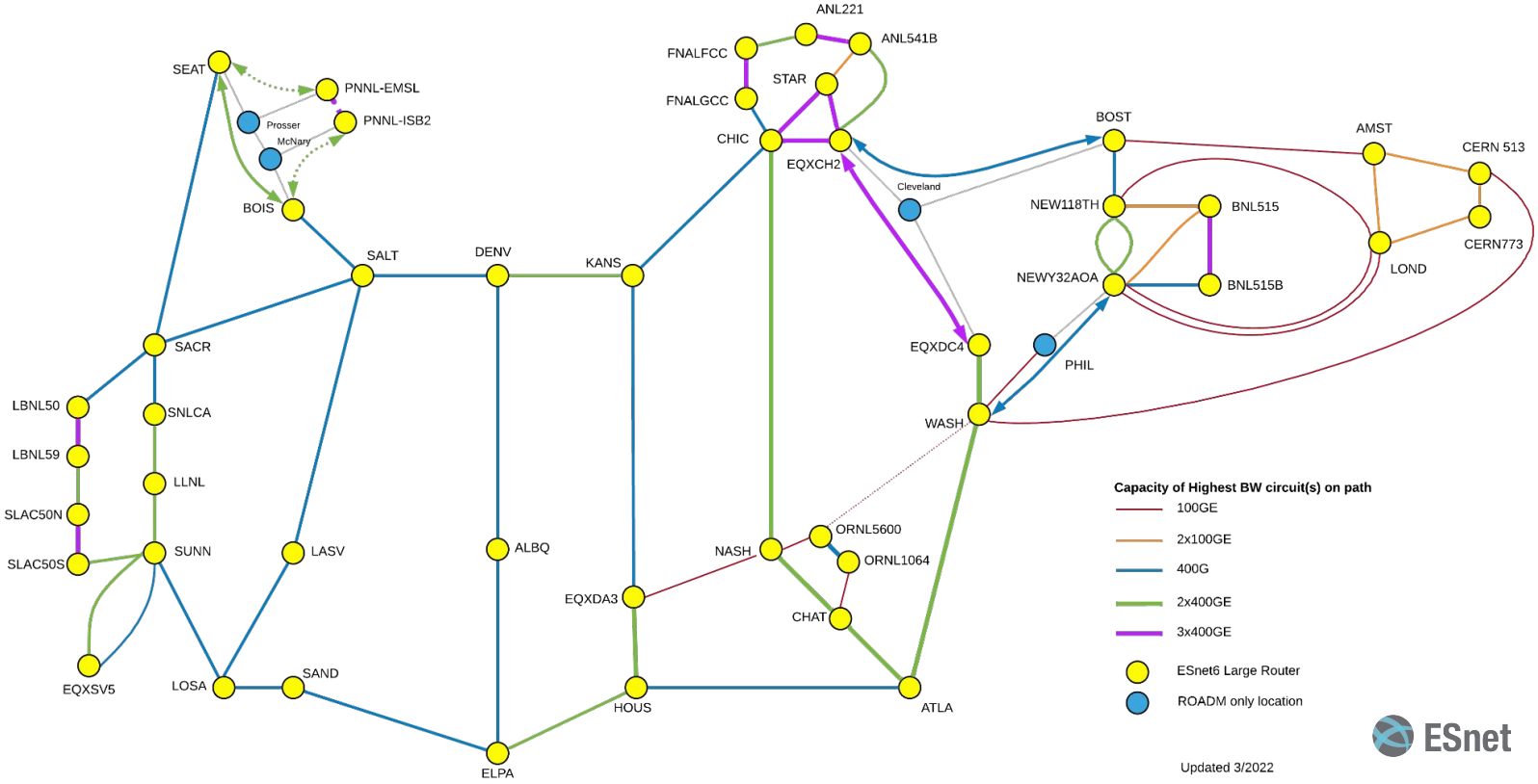
Major Backbone Changes in last Year

- Optical
  - Deployed 74+ Ciena 400G circuits
  - Upgraded 3 of the Infinera CHM2T circuits to 400G.
  - We now have both Infinera G30 and Ciena WL5E waves on every segment on the optical network.
    - This significantly reduces the risks when doing software upgrades to one of the platforms, or when updating the software automation stack for one of the platforms!
- Routers
  - Completed deploying and transitioning 42 ESnet6 routers into service
  - Retired more than 3/4 of the ESnet5 routers, including all Juniper MX960s, MX480s.
  - Rolled routed traffic off of almost all core 100G links to 400G links.

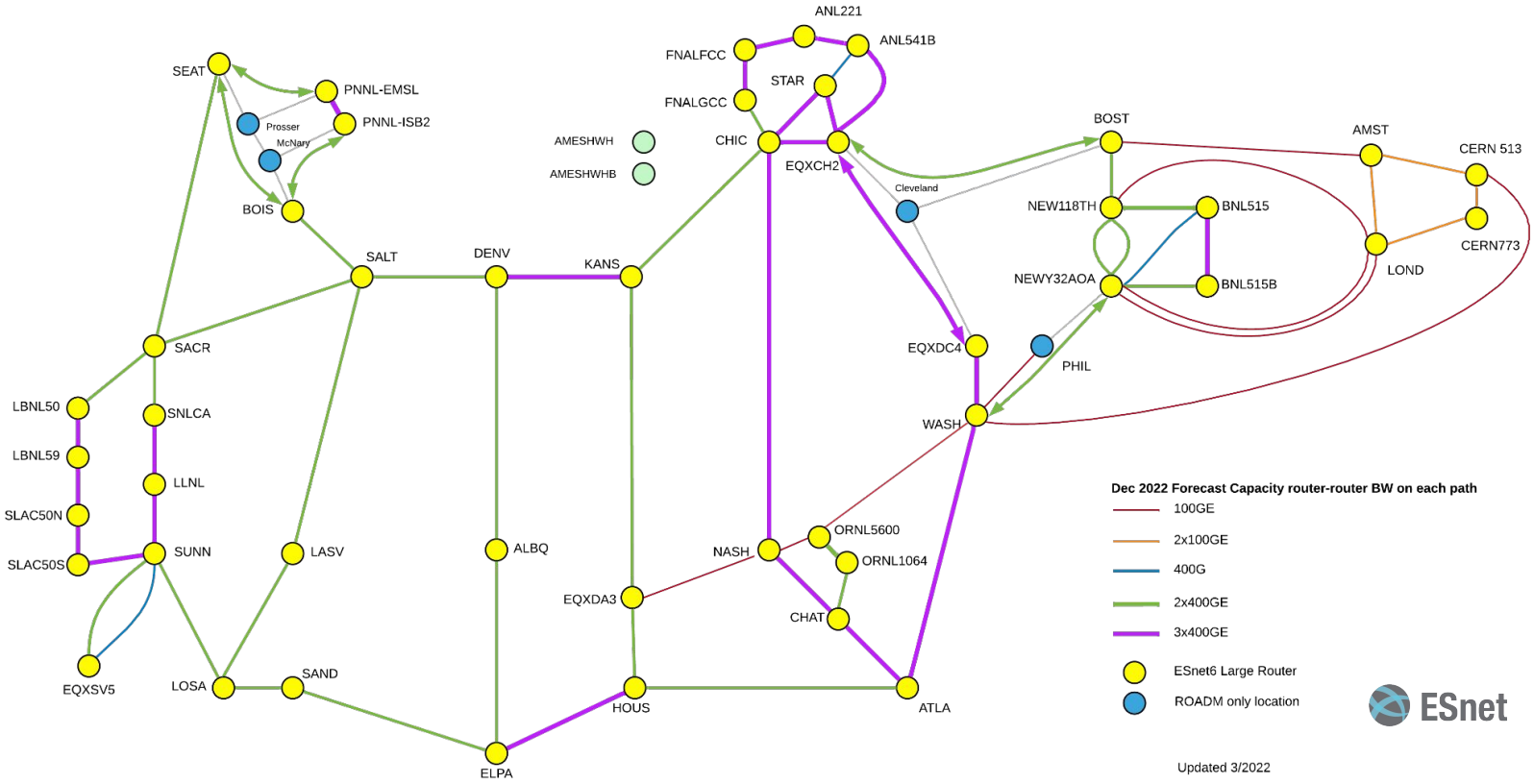
ESnet Backbone Transitions



March 2022 Routed Network Primary Capacity



December 2022 Target Backbone Capacity



WAN Black Hole Routing Service

A new ESnet6 Security Service replacing an ad hoc manual process used previously.

It discards traffic destined **to** IPv4 or IPv6 destinations at the ingress edges of the ESnet network.

Example use cases include:

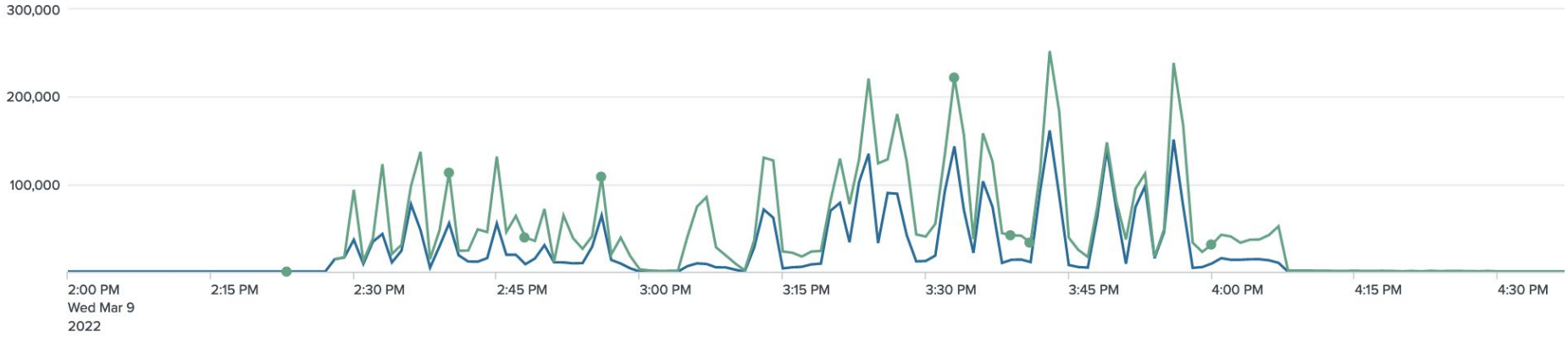
- Mitigate DDOS attacks aimed at a single ESnet customer host
- Stop communication destined to botnet command & control nodes

This is a simple service that we are using to develop the internal frameworks and processes necessary to develop and support more advanced security services

ESnet

WAN BHR First Production Use Case:
The day **before** the planned internal release…

A site report a DDoS attack and asked us to mitigate it by blocking all traffic to that host. Projected packet rates based on sampled flow data peaking at ~250k/min with unique IPs peaking at ~150k/min



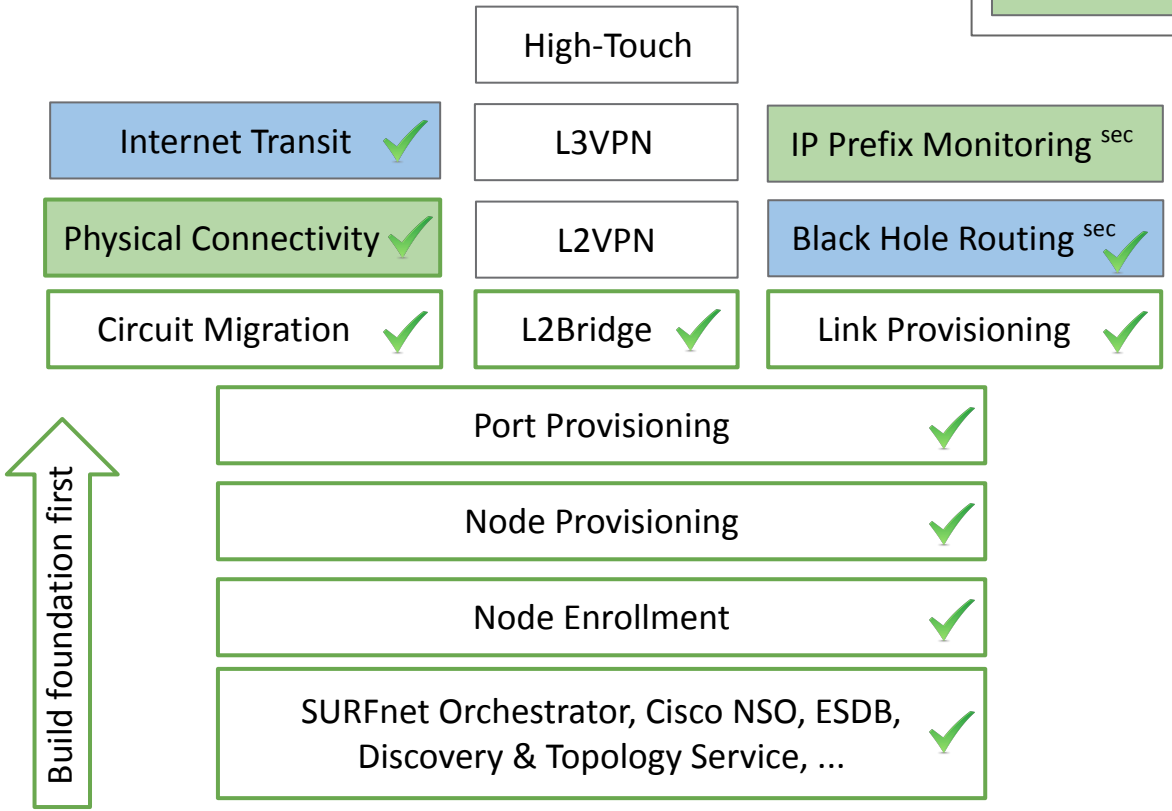
ESnet

Agenda

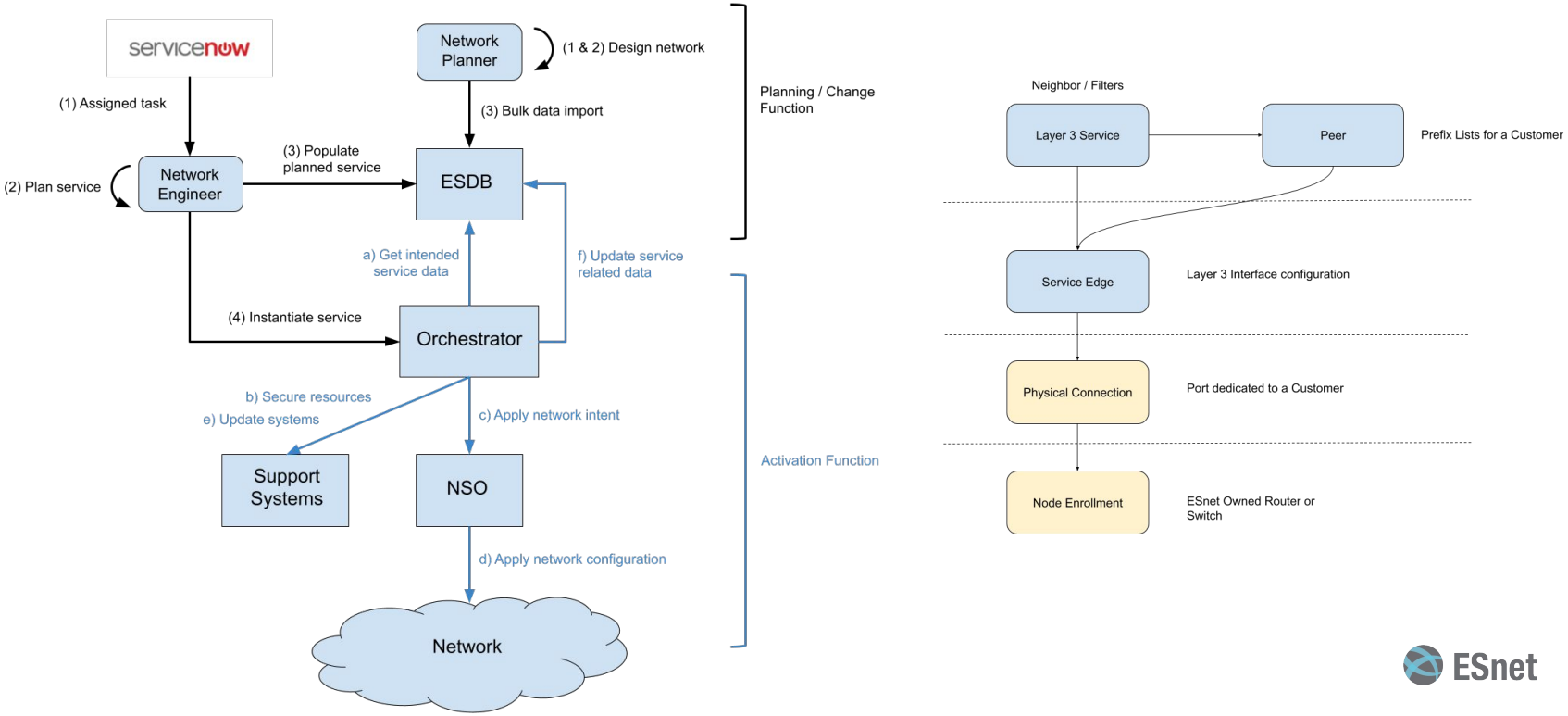
- Quick overview of ESnet
- ESnet6 Project Implementation Update
  - Backbone upgrades
  - Security black hole service
- **Automation**
- Monitoring
- Trans Atlantic upgrades
- Questions

ESnet

Delivering ESnet6 Automation



Most of our router provisioning activities are now using the automation stack to deploy services



ESnet

Automation by the numbers

- Migrated Service Edges & BGP Peers on 42 routers to orchestrated and automated services.
- Currently managing 1882+ service subscriptions in the orchestrator!
- Most network service turn ups and transitions are being done using automated tools at this point!
- It's not all rainbows and ponies yet, but we are heading in the right direction.
- Snowflake configs are still a challenge.

ESnet

Active Subscriptions

Advanced search: Search on resource types

id	Description	Status	In Sync	Abbr.	Tag	Start date	End date	Notes
id	description...							notes...
528e03b0	Node ornl5600-cr6 Production	active	☑	ESNET	NodeEnrollment	25-10-2021		
07a82574	Node ornl1064-cr6 Production	active	☑	ESNET	NodeEnrollment	25-10-2021		
db827489	Node lbnl59qa-cr6a Production	active	☑	ESNET	NodeEnrollment	25-1-2022		
95f51dd7	Node frib-cr6a Production	active	☑	ESNET	NodeEnrollment	26-3-2022		
ee2aa427	Node chat-cr6 Production	active	☑	ESNET	NodeEnrollment	29-9-2021		

1 2 3 4 5 ... 13 Rows per page: 5

Initial, Provisioning and Terminated Subscriptions

Advanced search: Search on resource types

id	Description	Status	Customer	Abbr.	Product	Tag
id	description...	☑				
55e94d5f	Node eqxld8-mpr1 Provisioned: Ready for Backbone Link	provisioning	ESNET::Energy Sciences Network	ESNET	Enroll Management Router	NESMPR
49acc9e5	Service Edge DOENET-WAPA-PX - Base - sand-cr6:lag-20 - doenet-wapa-px_se-509	terminated	DOENET-WAPA-PX::DOENET WAPA Phoenix Office	DOENET-WAPA-PX	Service Edge	ServiceEdge
d7874b3d	sacr-mpr1 to sacr-cr6	terminated	ESNET::Energy Sciences Network	ESNET	Management Link	MgmtLink
afa56f63	Service Edge DOENET-WAPA-LW - Base - denv-cr6:lag-20 - doenet-wapa-lw_se-508	terminated	DOENET-WAPA-LW::DOENET WAPA Lakewood Office	DOENET-WAPA-LW	Service Edge	ServiceEdge
5864f0c0	sand-mpr1 to sand-cr6	terminated	ESNET::Energy Sciences Network	ESNET	Management Link	MgmtLink

1 2 3 4 5 ... 77 Rows per page: 5

Agenda

- Quick overview of ESnet
- ESnet6 Project Implementation Update
  - Backbone upgrades
  - Small router deployments
  - Security black hole service
- Automation
- **Monitoring**
- Trans Atlantic upgrades
- Questions

ESnet

Stardust

Network Measurement and Analysis for ESnet

Extensible / Open Architecture

NSF NetSAGE project derived

Leverage Open Source components where we can, and innovate where it makes a difference.

Multiple access methods

Dashboards, Indexed APIs and “Raw”

Today, we are focused on users creating and sharing visual dashboards.

In the future, we expect direct programmatic access will become increasingly common for ML work and external collaboration.

Multi Datasource

Extensible and Open

Traffic Accounting, Link and Resource Use, Performance Testing Results, Others not yet invented.

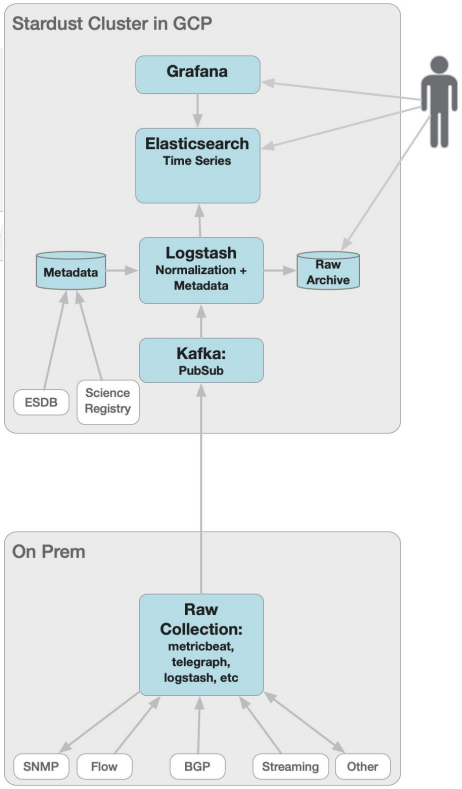
Each has its own set of measurements to which we add a common core set of metadata.

Flexible aggregation

Time Frame and Relationships

The metadata we add to measurements is used to summarize data to tell stories, and having common metadata allows us to use multiple data sets in a story.

- How are researchers moving science data and how has that changed over the last 3 years?
- What just caused that huge spike in traffic on the links to europe in the last 15 minutes and is that likely impacting data transfers?



ESnet

Current Scale

Ingest rate

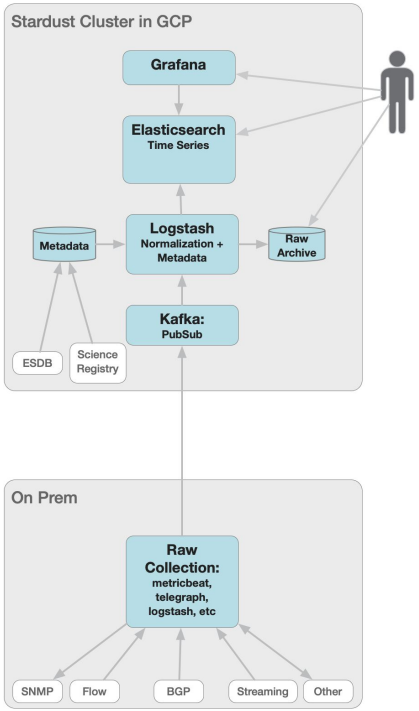
- Average: 10,000 eps (design target = 7000 eps)
- Peak: 20,000 eps (design target = 12,000 eps)

Data retention

- Raw data = 30 days
- Aggregates (90 days, 180 days, 3650 days)

Replication factor: 1

Storage: 147TB



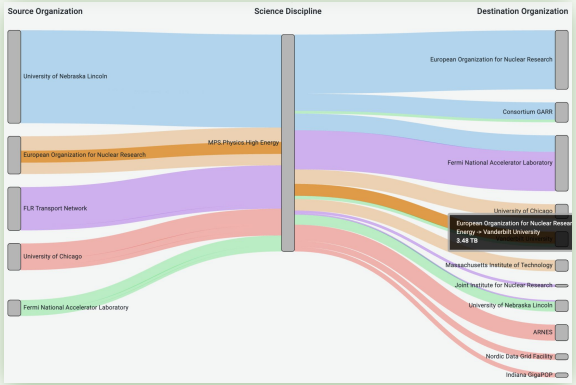
ESnet

Stardust Dashboards

- Combination of curated and custom dashboards used internally by ESnet staff to visualize and contextualize measurements

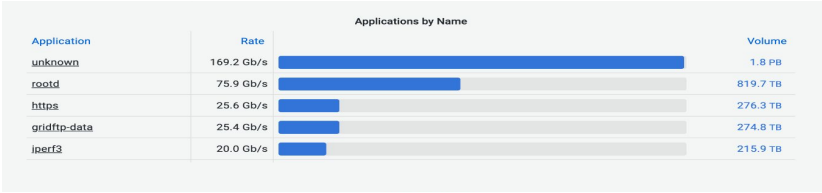
Not This

This

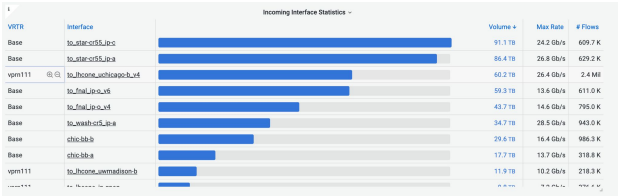


Stardust Flow Dashboards

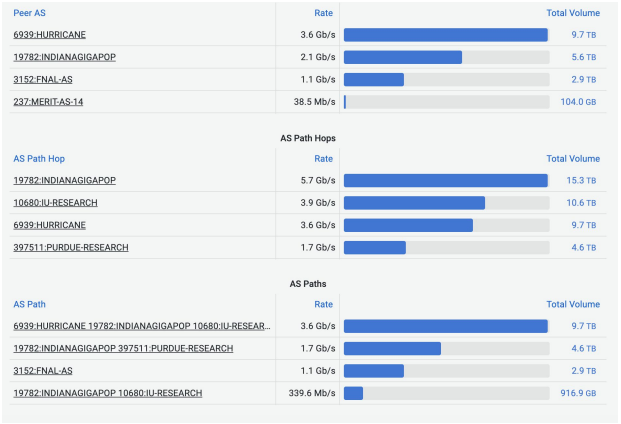
Breakdown by Application:



Breakdown by Interface:



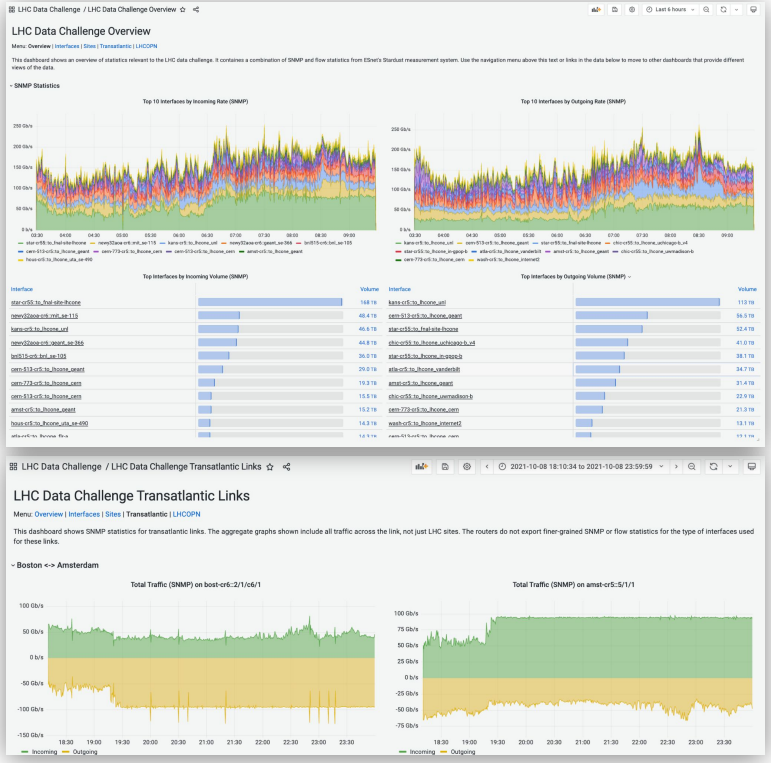
Breakdown by Peer and AS Path:



ESnet

LHC Data Challenge

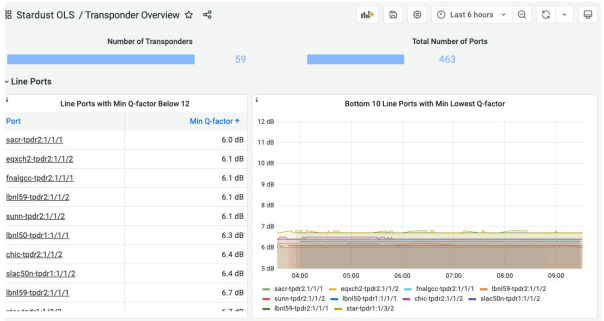
- First of 4 “data challenges” conducted from Oct 4-7, 2021 as readiness test for LHC high luminosity era
- With BNL (ATLAS) and FNAL (CMS) as US Tier 1 sites, ESnet a critical part of network path
- Specific ask was to provide a view of network metrics relevant to LHC sites
- Dashboards here



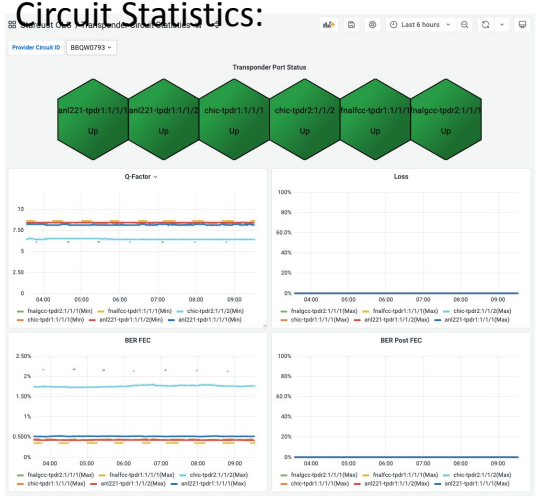
ESnet

Stardust Transponder Dashboards

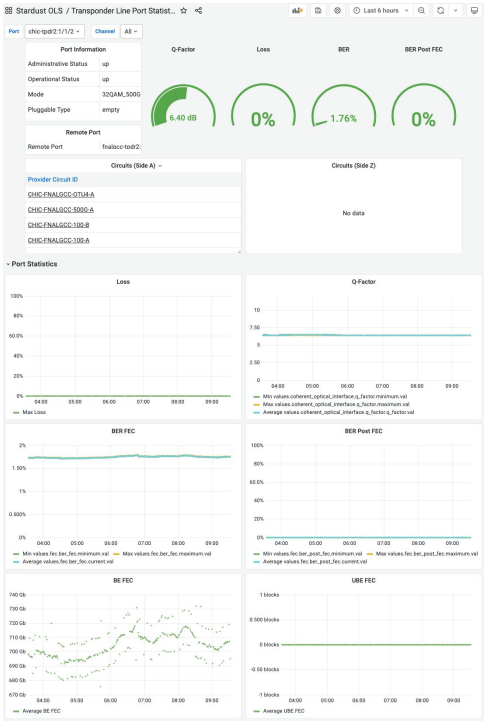
Metrics Across Network:



Circuit Statistics:

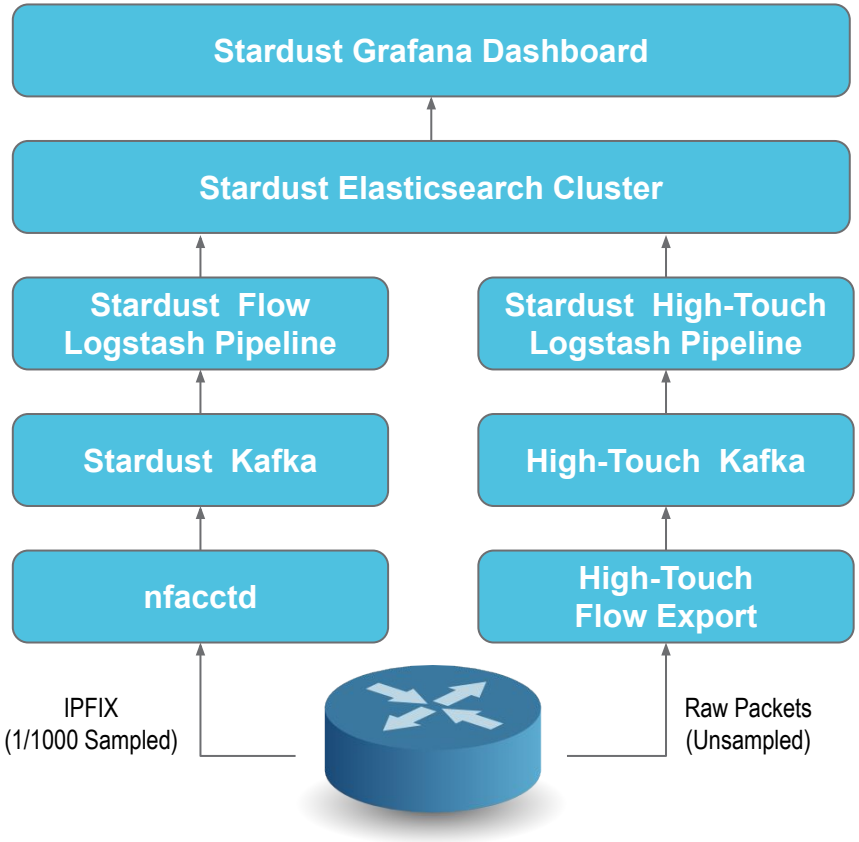


Port Statistics:



ESnet

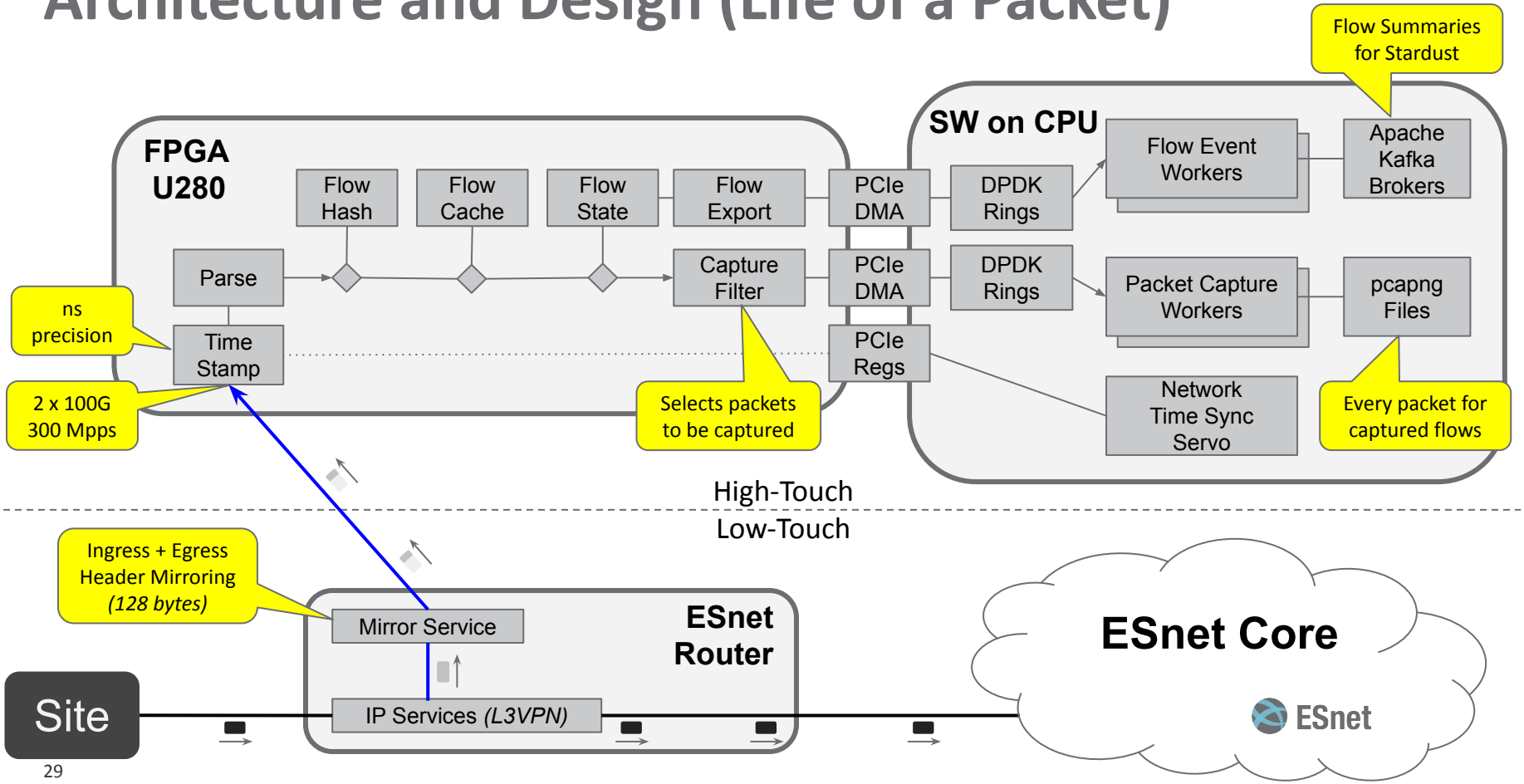
High-Touch + Stardust Prototype Architecture



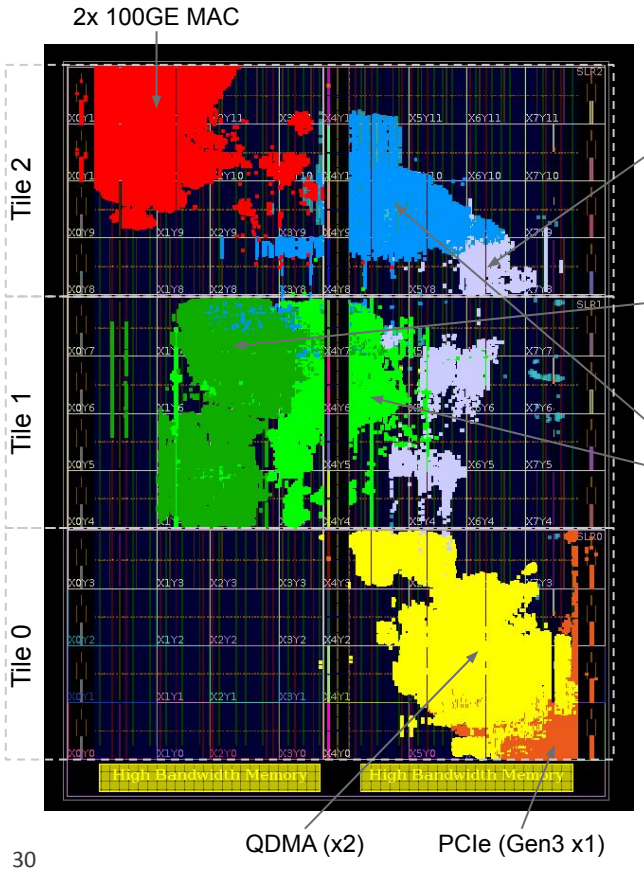
- The router sends IPFIX flow information to nfacctd (left column) and packet information to High Touch Exporter (right column).
- These follow parallel paths through a Kafka message bus, a logstash pipeline and are stored in separate indices of the same Elasticsearch cluster
- The Grafana dashboard displays results of both
- Stardust Flexibility of cloud vs VM:
  - The High Touch components from exporter to Logstash all run in local VMs before exporting to an Elasticsearch cluster in the cloud
  - The nfacctd component is a local VM but remaining Stardust flow components are cloud-based

ESnet

Architecture and Design (Life of a Packet)



High-Touch Integrated FPGA Logic Blocks



Xilinx Open NIC Shell (open source)

- Provides pin mappings, CMAC + PCIe/DMA interfaces
- ESnet was a pre-release user and provided user feedback

Xilinx SDNet (P4 program -> logic)

- Packet parsing, table lookups, packet filtering, packet edits
- Compiles a user-provided P4 program into FPGA logic

ESnet Custom Logic

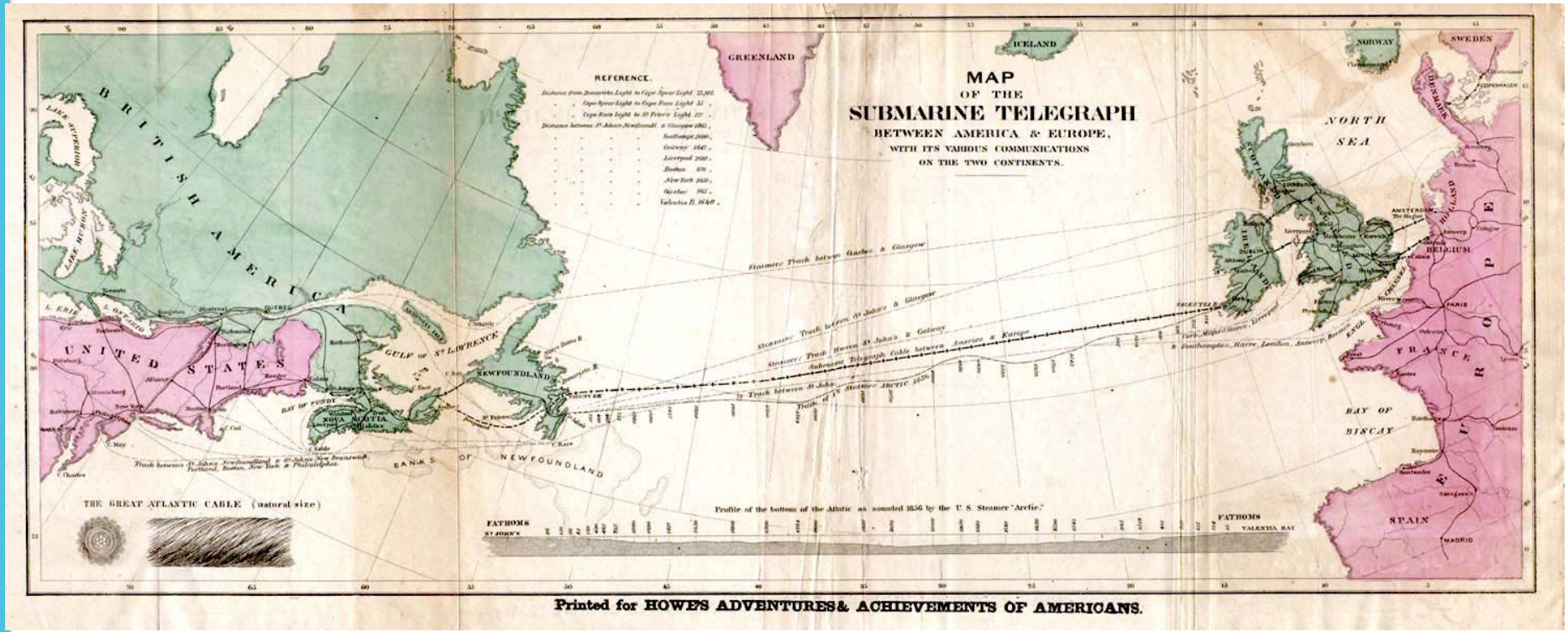
- Processes 100% of the packet headers on the wire
- Per-Flow state tracking block (new function for P4 program)
  - Unsampled packet/byte counts
  - Packet size histograms
- PCIe register interfaces
- (Room for more stuff!)

ESnet

Agenda

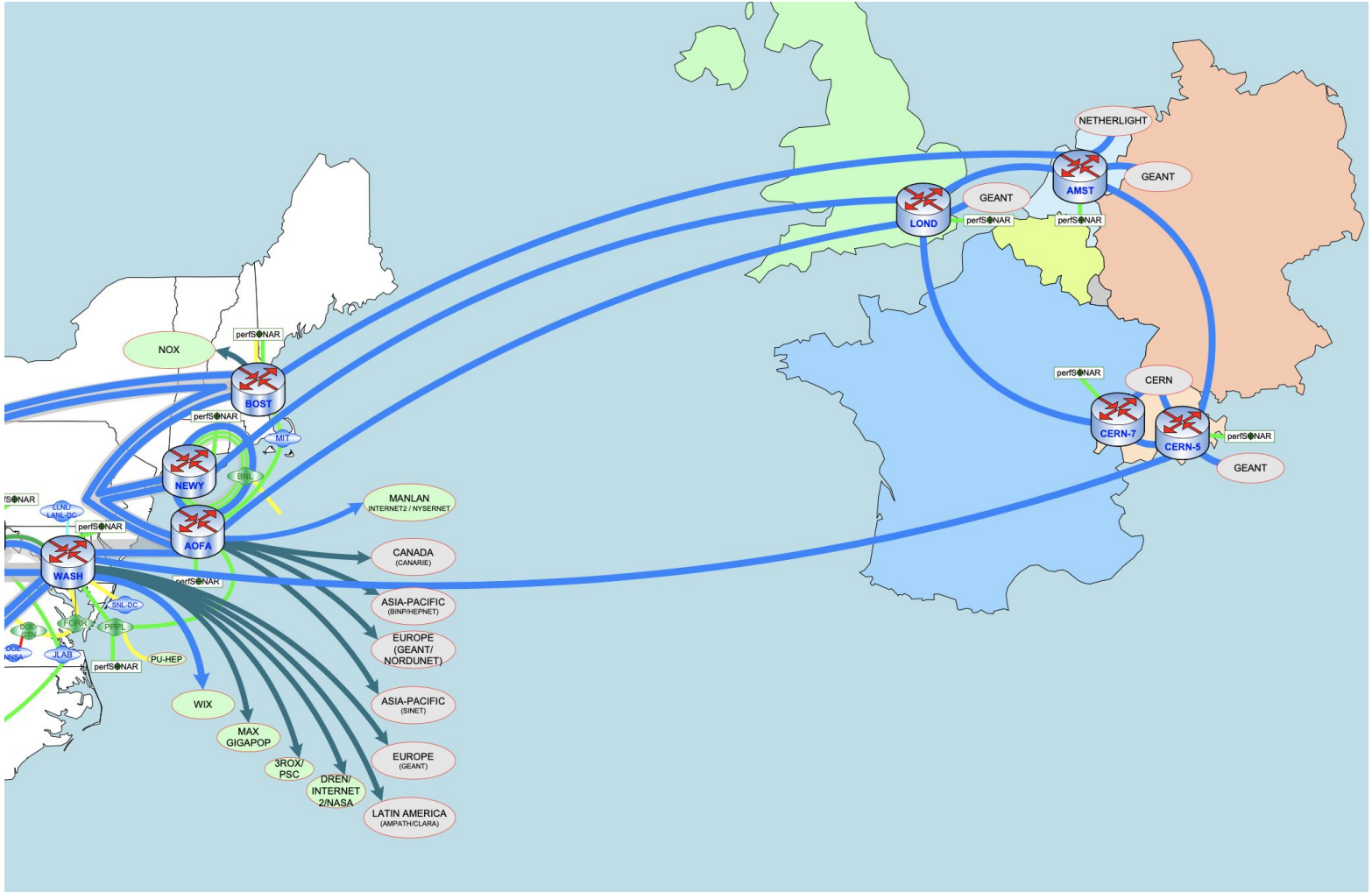
- Quick overview of ESnet
- ESnet6 Project Implementation Update
  - Backbone upgrades
  - Small router deployments
  - Security black hole service
- Automation
- Monitoring
- **Trans Atlantic upgrades**
- Questions

ESnet



Printed for HOWE'S ADVENTURES & ACHIEVEMENTS OF AMERICANS.

ESnet


NETHERLIGHT
GEANT
GEANT
AMST
LOND
GEANT
perfSONAR
perfSONAR
perfSONAR
CERN
CERN-7
CERN-5
perfSONAR
GEANT
NOX
perfSONAR
BOST
perfSONAR
MIT
NEWY
BNL
AOFA
WASH
perfSONAR
perfSONAR
LNL
LANL-DC
perfSONAR
SNL-DC
FCHR
PPPL
DOE GTN
NNSA
JLAB
PU-HEP
perfSONAR
MANLAN
INTERNET2 / NYSERNET
CANADA
(CANARIE)
ASIA-PACIFIC
(BINP/HEPNET)
EUROPE
(GEANT/
NORDUNET)
ASIA-PACIFIC
(SINET)
EUROPE
(GEANT)
LATIN AMERICA
(AMPATH/CLARA)
WIX
MAX
GIGAPOP
3ROX/
PSC
DREN/
INTERNET
2/NASA

ESnet's first 400G Transatlantic additions

- Spectrum on Amitié cable system
- Expected initial provisioning
  - 400G Boston-London
  - 400G Boston-Bordeaux
  - late 2022 / early 2023
- Internet2/CANARIE also on this cable
- Terrestrial builds & upgrades will be required
- Spectrum plans for additional TA cables are in process

ESnet

Questions...

ESnet

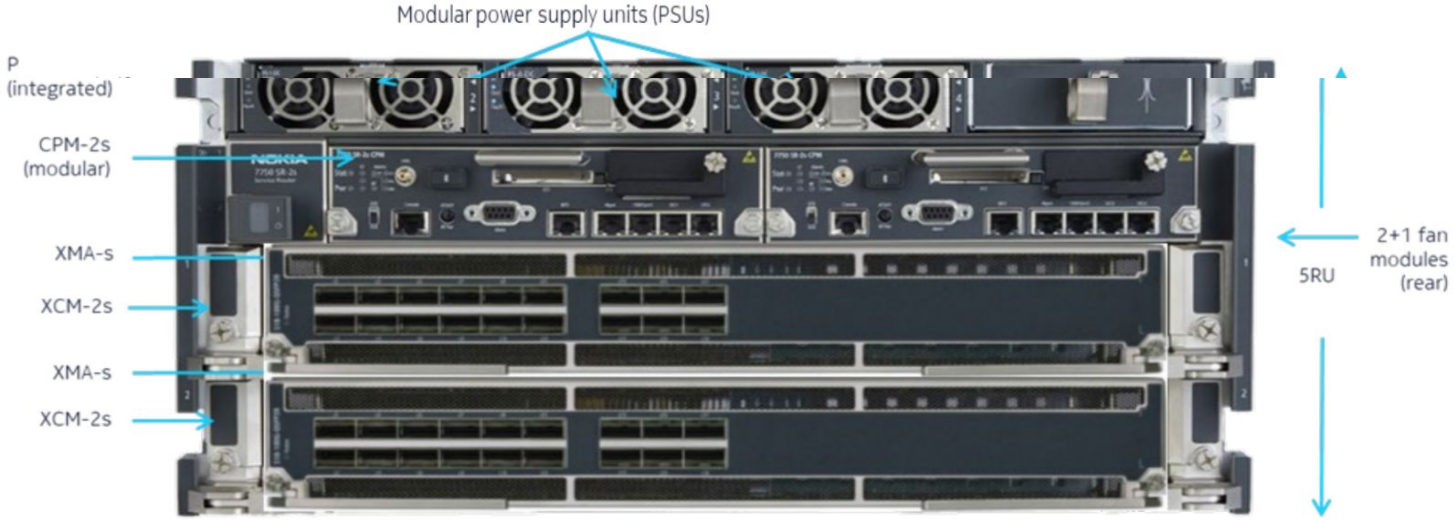
Nokia SR-2s Large Router Configuration

- Two 36-connector XMA-s cards licensed for the full 4.8 tbps
- Max supported per Router:
  - 24 x 400GE
  - 96 x 100GE
  - Other combinations of 10GE, 100GE & 400GE totaling 9.6 tbps
- QSFP-DD connectors support:
  - QSFP+: 4x10GE or 1x 40GE
  - QSFP28: 10x10GE or 1x100GE
  - QSFP28-DD: 2x100GE
  - QSFP56-DD: 1x400GE

ESnet

36

Nokia SR-2s Small Router Configuration

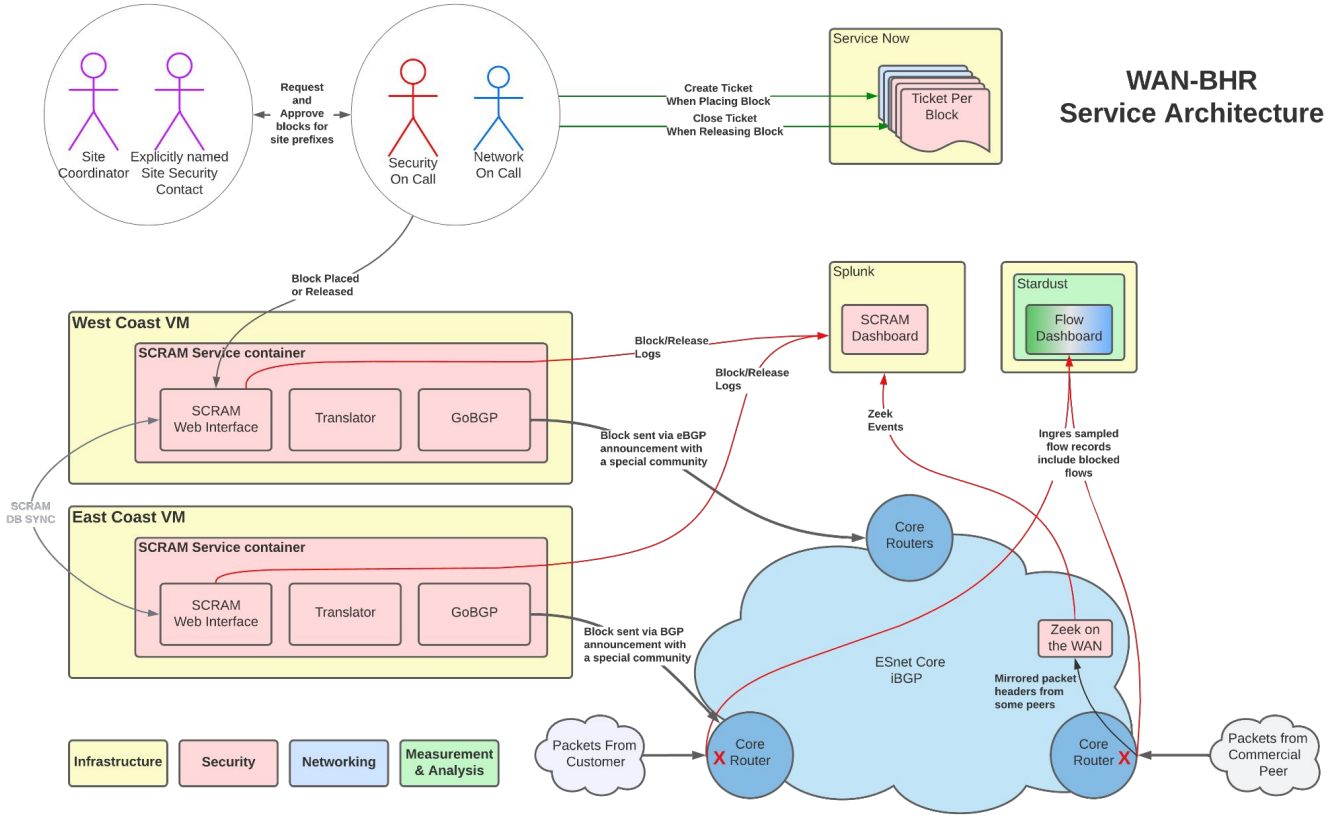


- Two 18-connector XMA-s cards licensed for 6 connectors & 600 gbps
- Max supported per Router:
  - 12 x 100GE
  - 120 x 10GE
  - Other combinations of 10GE and 100GE totaling 1.2 tbps

- QSFP28 connectors support:
  - QSFP+: 4x10GE or 1x 40GE
  - QSFP28: 10x10GE or 1x100GE
- License upgrade-able to 2.4 tbps per card

ESnet

WAN Black Hole Routing - Service Architecture



ESnet