



ESnet

ENERGY SCIENCES NETWORK

400GE Deployment Update

Chris Tracy

ESnet Planning & Architecture Group

Lawrence Berkeley National Laboratory

6th SIG-NGN Meeting

13th December 2021



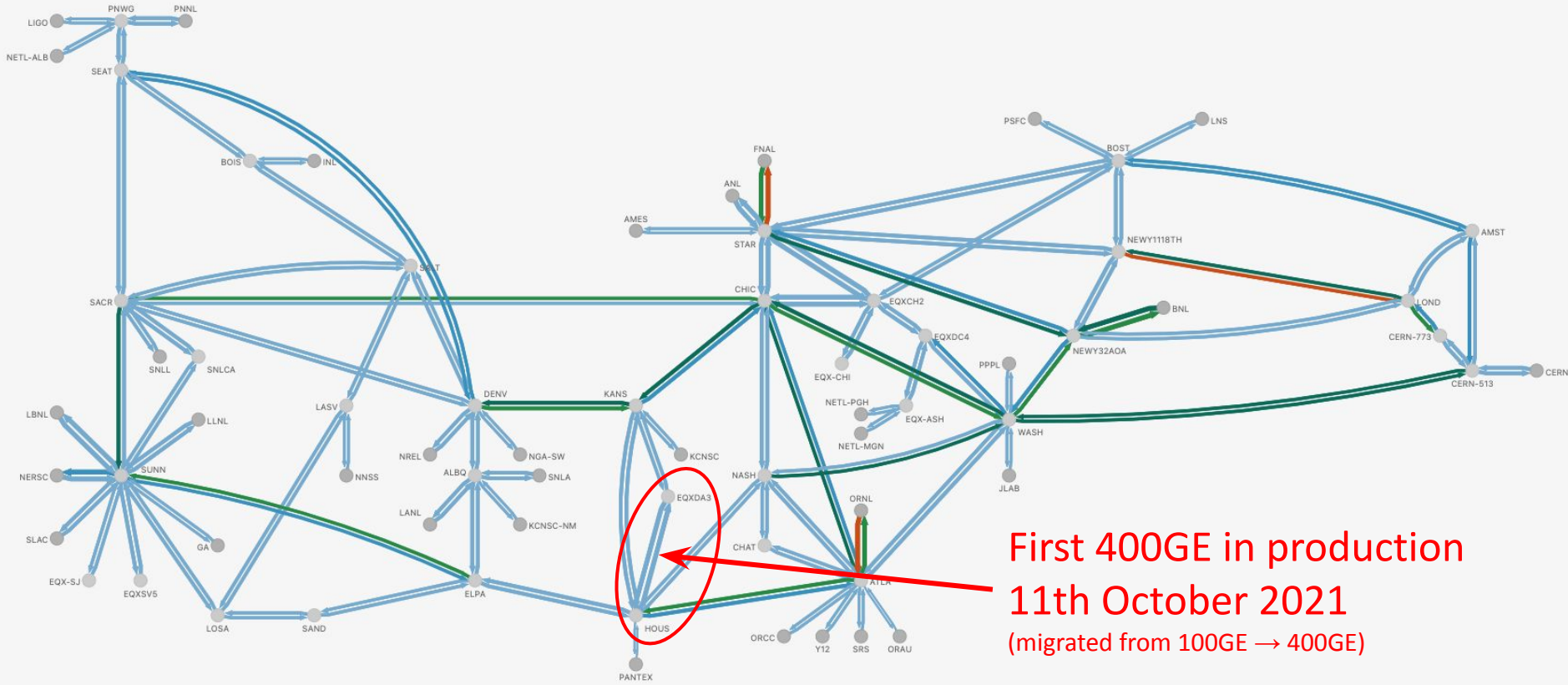
U.S. DEPARTMENT OF
ENERGY

Office of Science



ESnet6 — abstracted view (snapshot as of 9th Dec 2021)

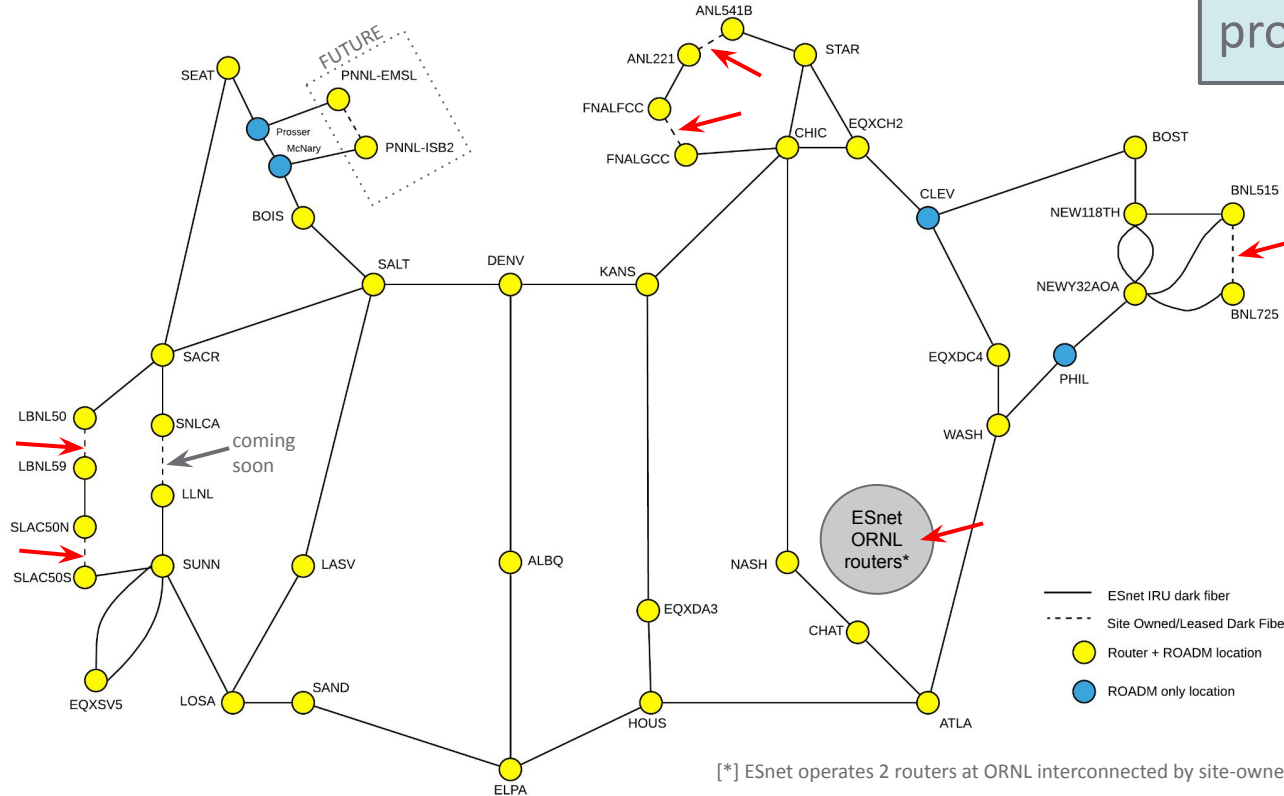
Provides connectivity to all of the DOE labs, experiment sites, and user facilities



400GBASE-FR4 deployment

(router-to-router over dark fiber)

Six pairs of
400GBASE-FR4 in
production today

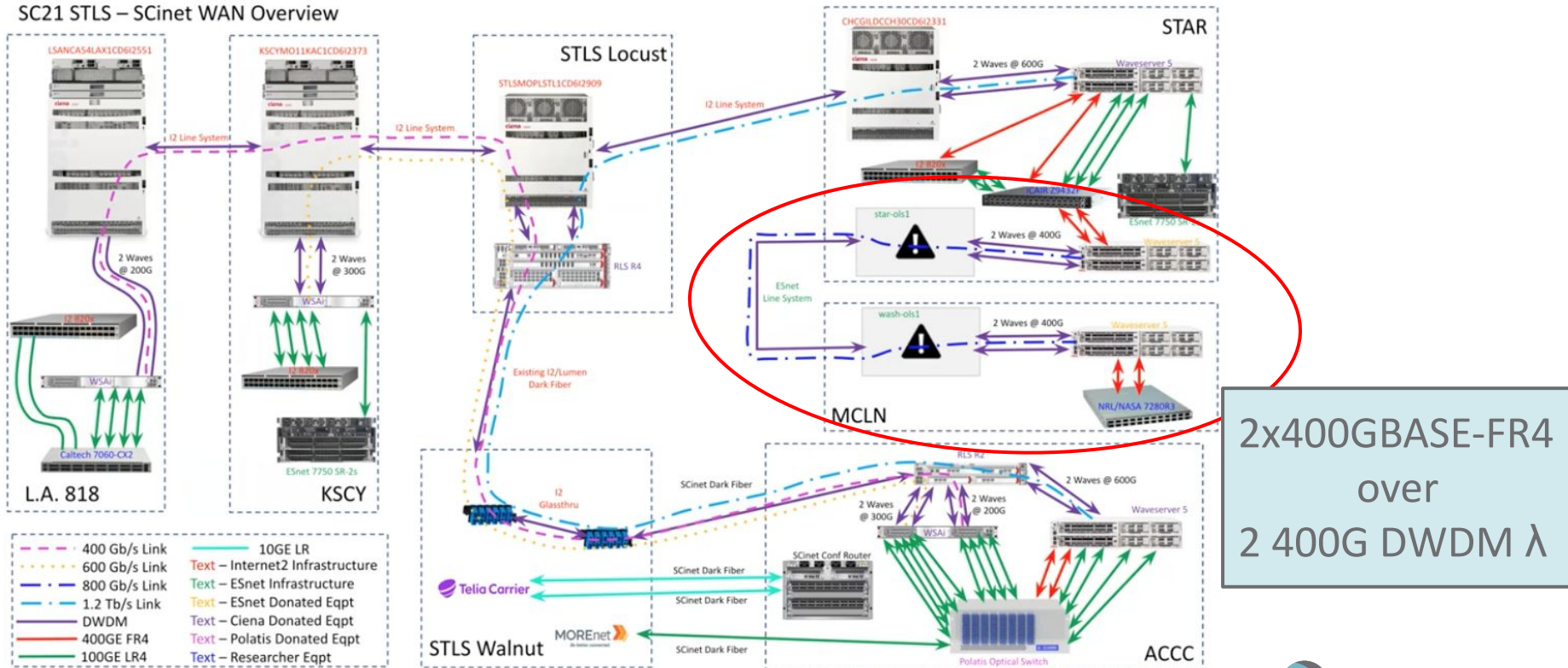


[*] ESnet operates 2 routers at ORNL interconnected by site-owned fiber which are not pictured



400GE support for SC21 (November 2021)

SC21 STLS – SCinet WAN Overview



2x400GBASE-FR4
over
2 400G DWDM λ

SC21-NRE-006 & SC21-NRE-017 demos: <https://sc21.supercomputing.org/scinet/network-research-exhibition/>

Diagram source: <https://www.youtube.com/watch?v=rH4Yvy3zGgQ> (8:56 mark) Brenna Meade - SC21 WAN Design



Outcomes and Lessons Learned

Interoperability, Reliability, and Stability:

- It was a concern given early interop reports by vendors at conferences
- ESnet opted for vendor-supported 400GBASE-FR4 optics for routers
 - No issues were expected or encountered (single-vendor solution)
- AOCs present a challenge — which vendor to purchase from?
 - ESnet opted to use AOC cables offered by the transponder vendor
 - The router is agnostic to third-party pluggables
 - Diagnostic & management capabilities are fully visible in router CLI
- First deployments of new hardware often come with some hassles
 - Some AOC link stability issues encountered, investigations ongoing

Short-Term Plans and Lab Testing

Nationwide production deployment of 2nd transponder vendor

- 84 additional 400GE backbone circuits, deployed as 400GE from day-1
- Again using transponder-vendor AOCs between transponder/router
- 95 GBaud symbol rate

High-order modulation: 2x400GE over 1x800G DWDM on short metro links

- Deploying in field, will be placed into production after acceptance testing

Lab testing: 400GE client over 2 adjacent 200G DWDM wavelengths

- Tested successfully in lab over short distances, not yet field tested
- Inverse multiplexing for long-haul 400GE transport

Later Migration Plans (for remaining 100GE clients)

The ESnet6 backbone still has over 100 100GE client circuits in production.

After the Waveserver 400G waves are deployed, we will go back and migrate groups of N x 100GE G30 circuits to 400GE in late 2022 using a similar process:

- Pre-install 400GBASE-SR8 AOC transceivers
- During maintenance window, drain traffic, disable 100GE client
- Swing electrical cross-connects to the 400GE ports
- Testing and acceptance (hardware BERT across the new 400GE path)
- Place circuit into production, monitor for stability and errors

Thank you!

Questions/comments?

Backup Slides

IGP Metric Rework

Our IGP metric algorithm is based on combination of link bandwidth and measured one-way latency.

100GE links had been using metrics in the lowest possible range

- 100G over a very short distance had a metric as low as 10

With the introduction of 400G, bandwidth scalars had to be adjusted

- Current production scheme accounts for link speeds up to 1 Tbps
- Defines scalars for speeds between 100G-1T, in increments of 100G
- 1 Tbps link over shortest distance — results in metric of 10
- 100G link over shortest distance — results in metric of 1000

New Features

ESnet's 100GE deployment did not include any of the 100GE variants that featured FEC at the Ethernet layer.

The introduction of 400GE brings FEC to the routers, as shown to the right.
(operating mode is CL119 or Clause 119)

Monitoring capabilities seem to be limited to CLI at this point.

For more info see: [Baseline Proposal for 4-lane Interleaved 100G FEC](#)

```
RS-FEC statistics:
Total Blocks          : 768844123494818
Uncorrectable Blocks  : 0
Corrected Blocks     : 99529553445
Corrected Symbols    : 99586844875
Corrected Zeros      : 94223576211
Corrected Ones       : 5376829099
Pre-FEC BER          : 2.454E-08
Post-FEC BER         : 0.000E+00

RS-FEC detailed lane information:
```

		Status	Errored Seconds		
Physical Lane	Fec Lane	AM Lock	Skew (bits)	Skew (ns)	AM Loss Lock
0	0	Locked	200	7.76	0
1	1	Locked	200	7.76	0
2	2	Locked	280	10.86	0
3	3	Locked	280	10.86	0
4	5	Locked	200	7.76	0
5	4	Locked	240	9.31	0
6	7	Locked	240	9.31	0
7	6	Locked	240	9.31	0
8	9	Locked	320	12.41	0
9	8	Locked	360	13.96	0
10	11	Locked	240	9.31	0
11	10	Locked	240	9.31	0
12	12	Locked	240	9.31	0
13	13	Locked	240	9.31	0
14	14	Locked	240	9.31	0
15	15	Locked	240	9.31	0

