



Joint Engineering Team (JET) Meeting Minutes

National Coordination Office for Networking and Information Technology R&D (NCO/NITRD)

490 L'Enfant Plaza SW, Suite 8001, Washington, DC 20024

August 17, 2021 12:00-2:00 p.m. ET

This meeting was held virtually

Participants

Nick Buraglio, ESnet

Todd Butler, NASA/GSFC

Rich Carlson, DOE/SC

Basil Decina, NRL

Dave Diller, MAX

Bill Fink, NASA/GSFC

Andrew Gallo, GWU/CAAREN

Ann Keane, NOAA/N-Wave

Jonah Keough, PNWGP/Pacific Wave

Michael Lambert, PSC/3ROX

Paul Love, NCO/NITRD

Joe Mambretti, StarLight/MREN

Linden Mercer, NRL

Joe Metzger, ESnet

Edward Moynihan, Indiana University

Aruna Muppalla, NASA/GSFC

Kevin Thompson, NSF

Christian Todorov, Internet2

George Uhl, NASA/GSFC

Proceeding: This meeting was chaired by Rich Carlson (DOE/SC) and Kevin Thompson (NSF).

I. Action Items:

- Internet2 and ESnet updates on their respective new networks.

II. Review of the Minutes of the July 2021 meeting: The only correction was the updating of the status of one Meeting of Interest.

III. ESnet6 Update – Joe Metzger

A. Architecture summary:

- a. Hollow core with Optical Line System (OLS) and core routers.
- b. Smart edge services with service edge routers, high touch programmable data plane and compute resources.
- c. As part of the ESnet6 deployment ESnet is migrating from colocation (colo) space shared with Internet2 to its own colo spaces.

B. Summary of deployed and planned hardware:

- a. Over 15,000 miles of fiber.
- b. OLS: Infinera Flex ILS CDC; MTC chassis, a Wavelength Selectable Switch per optical degree, two Route and Select CDC add/drop muxs per add/drop node, redundant shelf processors for all nodes including at amplification sites, RAMAN amplification on segments with approximately 20db or more attenuation, and OTDRs on all metro and RAMAN segments (approximately 40%).

- c. Transponders:
 - i. Phase 1 (initial): Infinera Groove G30 with one or two CHM2T sleds. The initial deployment will only be 100G. Supports 12xQSFP28 (100GE) or 2xQSFP-DD (400GE). Can mix and match. Distance impacts bandwidth. Licensed for full bandwidth from rollout.
 - ii. Phase 2: Ciena Waveserver 5 with up to four Wave Logic 5 Extreme sleds. Supports 16xQSFP28 or 4xQSFP-DD. Can mix and match. Distance impacts bandwidth. Licensed for full bandwidth from rollout.
 - iii. To provide redundancy, both models of transponders will be deployed in all locations.
 - d. Routers – All are Nokia SR-2s:
 - i. Large sites: Two 36 connector XMA-s cards with up to 9.6 tbps with 400GE, 100GE or 10GE. Supported optics: QFSP-DD, QSFP28DD, QSFP28 and QFSP+.
 - ii. Small sites: Two 18 connector XMA-s cards licensed for 6 connectors – 600G with up to 1.2 tbps with 100GE or 10GE. Supported optics: QSFP28 and QFSP+. Each card can have its license expanded up to 2.4 tbps.
 - e. Support equipment:
 - i. Management plane router: Juniper QFX-48S.
 - ii. Out-of-Band: Opengear 7216-2-24E with LTE and serial console ports.
 - iii. perfSONAR: 1U, AMD EPYC 7302P, 64G memory, 2x480gG disk, 100GE interface.
 - iv. PDUs: For DC Alpha Technologies Smart E2, for AC ServerTech PRO2.
- C. 400G circuits
- a. Though the current backbone is Nx100G, the few ESnet5 sites with two ESnet PoPs have a single 400GE circuit between the two routers using FR4 optics.
 - b. Starting in September the Phase 2 transponder deployment will start lighting one or two 400GEs between adjacent ESnet6 routers. (400GE active optical cable will be used to connect the router port to a transponder.)
 - c. After the router deployment is complete the Phase 1 transponders will be reconfigured for 400GE circuits. When completed this will eliminate all, or nearly all, 100GE backbone circuits.
 - d. Site connections:
 - i. Several sites are interested in upgrading their connections to 400GE.
 - ii. These discussions are still in process – no firm planes nor timelines at this point.
 - iii. The first 400GE backbone needs to be completed before any 400GE site connections.
 - iv. If the timing can be made to work the second 400GE backbone in a region, say a metro ring, should be also deployed before sites upgrade.

D. Deployment status:

- a. OLS
 - i. Substantially completed and in service for over a year.
 - ii. The Long Island MAN OLS, supporting BNL, will be completed in September (for a variety of reasons the routers needed to be deployed before the OLS).
 - iii. Part of ESnet6 buildout was getting fiber into PNNL – in process with estimated completion by March 2022.
 - iv. The Chicago area has lacked diversity on a few segments in the metro area. ESnet has found a vendor with the new fiber expected to be in service this fall.
- b. Transponders
 - i. Phase 1: Infinera Grove G30s have been in service for over a year supporting approximately 115 100GE circuits.
 - ii. Phase 2: Waveservers supporting 84 400GE circuits deploying this fall.
- c. Router deployment has had delays impacting transition schedules
 - i. Factors include delayed deliveries, subcontractor issues and COVID.
 - ii. Impacts:
 - 1. Focus from now through November is transitioning off ESnet5 routers in spaces shared with Internet2 as leases will terminate in December.
 - 2. Small router installs will be pushed out until after November.
 - 3. European installed will be pushed out until after the first of the year.
 - iii. These delays have given time to significantly increase the amount of automation used to configure and manage the installations and service transitions.
 - iv. Current estimated domestic completion dates: Large domestic – Nov 2021, small domestic and PNNL – April 2022.
 - v. European routers – June 2022 but this is still fluid as ESnet has decided to relocate its London and Amsterdam PoPs from space shared with GÉANT into its own colo space in the same facility.
- d. High Touch
 - i. Significant supply chain constraints.
 - ii. ESnet is deploying two early prototype High Touch servers from the lab into field locations. This will allow it to meet a Project Threshold Key Performance Parameter in a low risk and timely manner.
 - iii. The schedule for High Touch deployments at the large router nodes is unknown – perhaps half a year delay.
- e. Greenfield deployment (first 40 routers) was done using six teams of one to three engineers each. During a transition there are multiple maintenance events but only the final one, service migration, is service impacting for most.

- f. Brownfield deployment is done using fewer teams based on feedback from the Greenfield sites. Typically this is one 2+ day event with several short service impacting windows.
- E. Capacity targets for Q1CY22 (not final capacity)
 - a. Provisioning targets between adjacent router/optical nodes:
 - i. 1Tb (2x400GE plus 2x100GE):
 - 1. Chicago<>Nashville<>Atlanta<>Washington, DC<>Chicago.
 - 2. Chicago metro ring.
 - ii. 600G (1x400GE plus 2x100GE): Remaining ESnet domestic optical footprint.
 - iii. 2x100GE: London<>Amsterdam<>CERN.
 - iv. 4x100GE: Transatlantic circuits (plus shared circuits with NEA³R and GÉANT).
 - b. Deviations from first quarter projected capacity:
 - i. More 400GE than planned:
 - 1. Short metro links came in at 2x400GE vs. 1x400GE.
 - 2. Most nodes have an odd number of degrees and the transponders come with two modems per card. The extra modems are being turned up where the extra bandwidth is beneficial.
 - ii. Less 100GE than expected between adjacent routers:
 - 1. Current 100GE includes many express routes that do not connect adjacent routers.
 - 2. ESnet6 has four additional routers vs ESnet5 (Chattanooga, San Diego, Los Angeles and Salt Lake City).
 - 3. The existing express routes will be kept as backup until the Phase 1 transponders are converted to 400GE (to be done after all Ciena 400GE transponders are deployed).
- F. Questions/discussion:
 - a. Why both Infinera and Ciena? The goal was to have an OLS. By selecting two vendors ESnet anticipates future transponder purchases will remain competitive and permit its obtaining the best available at any point in time. It also looks to the future with the possibility of colored optics in the routers. Perhaps initially over shorter distances in a metro area.
 - b. RAMAN vs 20db – How is the performance over distance looking? ESnet believes it'll be able to run dozens of waves at 400GE with the transponders they have. 800GE with transponder they expect to be able to get. The sole concern is one segment of fiber which is over 130km. They are investigating putting an amplifier hut in the middle.

IV. Discussion of the JET's tasking on tools to help with inter-domain issues – Joe Breen (via email)

This is a community project to collect shared data from all who are willing to share.

- A. Prototype/pilot: The various pilots continue to progress. Work continues to get basic measurement data from different universities and RONS.

- a. Red Clara is now sending data to the project.
 - b. After a meeting with MERIT they will be adding the data from interfaces from ten routers
 - c. Work with LEARN is expected to start after a meeting with them in two weeks
 - d. Pacific Wave engineers are investigating making other interfaces available
 - e. Additional discussions with the University of Hawaii's Institute for Astronomy.
- B. Background on efforts lead by Eric Boyd, Joe Breen, James Deaton, Dan Doyle, and Karl Newell:
- a. The project gets basic SNMP metrics from groups around the country that are willing to share for trouble shooting and research. Metrics include link utilization, discards and errors. These are collected hop by hop as the path crosses multiple domains.
 - b. Several prototypes are going along with the drafting of a basic letter of intent for those wishing to participate.
 - c. Tools: Telegraf container as an option for local collection. Nearly ready for production use.
 - d. Tracking sheet of networks willing to share data. Please update your network's entry. See:
https://docs.google.com/spreadsheets/d/1pMW_PNVpeT42nAxa3bW4QostMxcHTXkWSPbZOpIFwE/edit#gid=0
 Templates for campus, regional and national networks setting out what data is desired can be found at:
 Campus template: (for a Science DMZ or research segment)
https://docs.google.com/spreadsheets/d/1v7iFw8_YoMpa3wigwcmlZgy0QsTi1bHb4Qk1cV6qfAM/edit#gid=1161461998
 Regional template:
<https://docs.google.com/spreadsheets/d/1ElqYiLTLn-Q07doDzHb5vtUCUosFLNbNSgiumm145d4/edit#gid=0>
 National backbone template
https://docs.google.com/spreadsheets/d/14CQq67LjJ_hlnrpiL8WpTbHmQSW112zzvKPBp6fx8Gw/edit#gid=0
 - e. The Internet2 Performance Working Group Community Measurement, Metrics, and Telemetry project holds meetings on the second Tuesday for those participating or interested. If you are interested, please contact Joe:
[Joe Breen <Joe.Breen@utah.edu>](mailto:Joe.Breen@utah.edu)
 - f. General information about this project can be found at:
<https://spaces.at.internet2.edu/display/PerformanceWG/Internet2+Community+Measurement%2C+Metrics+and+Telemetry+Project>
 - g. While NASA policies preclude EOS from sharing this data, EOS has an internal perfSONAR (pS) mesh. They are happy to open their firewalls to permit pS testing by prior arrangement. Contact George at:
["Uhl, George D." <george.d.uhl@nasa.gov>](mailto:George.D.Uhl@nasa.gov)

V. **Operational network security roundtable** No updates were received.

VI. **Network roundtable**

- A. CAAREN (Andrew Gallo): No update today.
- B. ESnet (Nick Buraglio): Work on the OMB mandate for IPv6 only networks continues with DOE. Several from ESnet have been working in different groups within DOE on policy, the implementation plan, and to build and document the pilot programs. The needed policy and the implementation plan are nearly complete. The group has developed three implementation plans. Two have been documented and will be submitted. The third will be documented if needed.

DOE had a data call on the current state of affairs on IPv6 with the anticipated results: a very mixed back of capabilities and knowledge of IPv6. The biggest concern was the needed resources to implement the mandate.

Questions:

- a. Q: How much of the experimental infrastructure supports IPv6?

A: A very broad question and the answer depends by what you mean as experimental. For the experiments' compute it depends on the software. The specialty software that controls the equipment it's very much a mixed bag. If it relies on the base systems OS it usually just works. If it's custom written software it may have hard coded IPv4 items. Most likely these packages were written with no thought of anything besides IPv4.

Software for telescopes, light sources, etc. is quite often imbedded and barely does IPv4. Ancillary systems can mostly do IPv6. The management systems quite often can do IPv6.

The planning teams have from the start focused on the solved problems – enterprise networks. Concurrently they are documenting the specialty items and working out how to deal with them in the later stages of the migration. Most likely a big portion of what stays IPv4, or dual stacked, will be the specialty systems.

ESnet also has to think about collaboration software, especially where the collaboration is with entities outside the US government who may have little or no incentive to use IPv6. (US universities are a good example as most received large IPv4 allocations.) This will be exacerbate performance issues with any high throughput flows as passing it through NAT64s will drastically the performance. These systems will also need to be in the 20% remaining IPv4 or dual stacked.

- C. Internet2 (Christian Todorov): Internet2 (I2) is making good progress with its NGI rollout.
 - a. The 400G optical infrastructure is up. To date it is outperforming the design specifications.
 - b. The service transitions from the old routed platform to the new has begun.
 - c. I2 is using an automated process for this. Some tweaks to the scripts were discovered at the initial couple of sites but now going well.

D. International Networks – Indiana University (Ed Moynihan):

- a. All circuits are stable.
- b. The Guam<>Singapore 100G contract is still being competed.
- c. International Networks – Indiana University (IN/IU) held a BOF session at TNC earlier this year and at the recent APAN meeting on community-wide routing issues on trans-oceanic flows. GNA-G and APAN are hosting a formal Working Group on this issue.

If you are interested in joining the WG contact Ed:

[Edward Moynihan <edmoyn@iu.edu>](mailto:edmoyn@iu.edu)

E. NASA GSFC (Bill Fink and George Uhl):

- a. No update for today.
- b. Nick already mentioned this in the IPv6 discussion but in the science community you have all kinds of partners who you have no control over their use of IPv6 and their capabilities to use it. We raised this issue with meeting within NASA at ESDIS. We have science and mission support and this is an issue we're working through internally. Some of the missions have been in operation for a long time and we're trying to work out the scope of the mandate on them.

Nick: ESnet found it very useful to find out/confirm what systems, especially old systems, it had. ESnet didn't have the expectation these old systems would even be considered in the first round or two. Some really old systems barely do IPv4. ESnet has had discussions with OMB. OMB wasn't very receptive to any sort of exemption. OMB take: that's what the 20% is for. The ESnet6 management network, after two years of work, still has 2 to 5% of its components not capable of IPv6-only. Despite the additional upfront costs and efforts, when a network gets to IPv6-only it is simpler to be just doing the one protocol. What OMB is asking for is doable, the difficulty is the timeline and the un-funded nature of the mandate. Servers and desktops are easy to do. It's the other systems – VoIP, sensors, door controllers, etc. that are the difficult part of the mandate. The IoT will become the majority of items to be concerned with.

ESnet had extensive discussions on exactly what an IPv6 node. Its conclusion was that a node is anything running an IP stack that can be managed discretely. A VM system is. A container, managed by a container management system, is not (you can't ssh into a container).

F. NOAA (Ann Keene): N-Wave is finishing work with vendors to complete a first generation architecture for services in Alaska. Initially it is a 1 GE ring from Seattle to Anchorage to Fairbanks and back to Seattle. There is a spur from Anchorage to Juneau to support NOAA's fisheries group's two Juneau sites. The existing service in support of NOAA's satellite group near Fairbanks and at Utqiagvik (Barrow) will be migrated to the new infrastructure in the fall.

G. Pacific Wave (Jonah Keough):

- a. Pacific Wave (PW) is continuing its work on performance measurement and monitoring through its collaboration with UC San Diego.
- b. For its west coast optical systems PW is upgrading to Ciena Wave Logic 5s for both its Ciena and Cisco optical line systems.

H. PSC/3ROC/XSEDE (Michael Lambert): No updates this month.

VII. Exchange Points Round Table

A. PNWGP (Jonah Keough): No update this month.

B. StarLight (Joe Mambretti):

a. SC21 planning is going forward:

- i. StarLight (SL) is working with the SCinet WAN team to provision a national testbed to support demos at SC21.
- ii. In the DC area SL is working with NRL and GSFC on a rack of hardware in the McLean, VA, exchange. From there SL is working with SCinet on having two 400G circuits provisioned to SL.
- iii. At SL there will be a variety of switches to handle the multiple 400G and 100G connections.
- iv. Also with SCinet's WAN team work is going forward to provision 2x400G and 4x100G circuits from SL to SC (in St Louis, MO).
- v. Discussions are ongoing with the WAN team on bringing in circuits to SC from the west coast.
- vi. Some of the anticipated demos are:
 1. A 500G demo with the Open Science Grid
 2. A NOTED AI demo
 3. Demos from both the Pacific Research Platform and the National Research Platform on software defined provisioning.

b. The FABRIC node at SL is complete. Initial tests are related to the needed VLANs on it.

c. Work on the Data Mover Challenge for SupercomputingAsia is pretty much completed. SL will be submitting an application for the competition side of the Challenge.

d. SL is very involved in the development of next month's Global Research Platform workshop. The final agenda should be out next week.

C. MAX (Dave Diller):

a. MAX has had delays in getting the POs out for its upgrade. The only equipment received is for the Out-of-Band (OoB).

b. For OoB MAX has identified a single box solution for OoB from Opengear which combines serial and Ethernet ports in the one box. It's very early in the box's development, documentation, etc. so there are challenges. They are very close to being able to operate in an IPv6-only environment.

c. MAX is adding commodity capacity in preparation for the return of students.

D. WIX/ML (Christian Todorov): Internet2 (I2) has transitioned these to the NGI infrastructure. I2 is considering some additional features and functionality for the future.

Meetings of Interest 2020

Note: Meetings cancelled since the August JET have been removed from this list. Those moved to a virtual format have been updated.

Sep 20-24	The 2nd Global Research Platform (2GRP) Workshop , virtual meeting n.b.: 20-22 N&S America and Europe; 22-24 N&S America, Asia Pacific and SE Asia
Sep 24	Americas' Research Platform Working Group , virtual meeting
Sep 28-30	The Quilt Fall Members Meeting , virtual meeting
Oct 19-20 & 28	ARIN 48 , virtual meeting
Nov 1-3	NANOG 83 , Minneapolis, MN, hybrid meeting
Nov 4	ARIN 48 , Minneapolis, MN, hybrid meeting
Nov 6-12	IETF 112 , in person cancelled, moved to a virtual meeting
Nov 14-19	SC21 , St. Louis, MO, Hybrid meeting
Jan 16-19, 2022	PTC'22 , Honolulu, HI
Feb, date TBA	APAN53
Feb 14-16	NANOG 84 , Austin, TX
Mar 1-3	SupercomputingAsia 2022 , Singapore
Mar 19-25	IETF 113 , Bangkok, Thailand

Next JET meetings

Note: It is anticipated that JET meetings will remain virtual for the foreseeable future. The possible exception will be the November meeting if SC21 remains a hybrid conference.

Sep 21, 2021 12-2 p.m. ET

Oct 19, 2021 12-2 p.m. ET

Nov 16, 2021 8:30-10 a.m. CT,

n.b. This meeting is concurrent with SC21. It will be held in room 127 of the America's Center Convention Complex, St. Louis, MO. Remote access is expected to be available.