



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
March 21, 2023, 12:00 – 2:00p.m. ET
This meeting was held as a virtually

Participants

Hans Addleman, IU	Michael Lambert, PSC/3ROX/ACCESS
Jeff Bartig, Internet2	Paul Love, NCO/NITRD
Dale Carder, ESnet	Joe Mambretti, StarLight/MREN
Rich Carlson, retired	Ralph McEldowney, DREN
Bill Fink, NASA/GSFC	Linden Mercer, NRL
Jonah Keough, Pacific Wave/PNWGP	Edward Moynihan, Indiana University
Kevin Kranacs, NASA/EOS	Glenn Ricart, US Ignite
Padma Krishnaswamy, FCC	Adam Slagell, ESnet
Yatish Kumar, ESnet	Jim Williams, Indiana University

Proceeding: This meeting was chaired by Ralph McEldowney (DREN).

I. **Action Items:** (none pending)

II. **Review of the Minutes of the February 2023 meeting:** Received corrections were incorporated in the posted final minutes.

III. **ESnet High-Touch Telemetry: Yatish Kumar**

Minutes from the last High-Touch update in October 2021 can be found at:

<https://www.nitrd.gov/coordination-areas/lsn/jet/jet-meetings-2021/>

Minutes from the previous High-Touch update in October 2020 can be found at:

<https://www.nitrd.gov/coordination-areas/lsn/jet/jet-meetings-2020/>

- A. Since the last High-Touch (HT) update to the JET HT boxes are now being deployed with two at each ESnet node. HT is currently at about half of ESnet's 42 locations. With the completed deployment there is now six months of production experience and data with some interesting use cases.
- B. For the basic design & architecture of ESnet's FPGA telemetry collection system, HT, please see the minutes from Oct 2021. HT is interested in edge data – traffic entering and leaving ESnet. For the edge traffic HT looks at every packet's header at full packet rate. Packets are time stamped with nanosecond precision. Flow summaries are exported to Stardust. For flows of interest the packet headers are exported into a pcap file for detailed analysis. HT operates in an IPv6-only environment without Internet access.

- C. The result of the FPGA and the integrated software is a SQL database with detailed information on each flow. About 1 TB per day is generated. Data is retained using a moving window to cull to fit existing storage capacity. With the data in a SQL database, it queries across all fields of interest can be easily written.
- D. For selected flows of interest HT writes a pcap file. The specifications can include many parameters to narrow the captured traffic. While currently each HT node needs to be configured manually, plans call for fully automated, network wide capture to be added. With each ingress and egress point monitored each packet's entrance and exit from ESnet can be seen regardless of an asymmetric routing. The developed software stitches all together so that the network looks like a single node doing a tcp dump.
- E. Some use cases with the summarized flow data - the High-Touch Packet Microscope:
 - a. Data plane validation:
 - i. Looking for packets that are unexpected - an anomaly in the data plane, traffic somebody is putting into the net, etc.
 - ii. Look at what HT can do with timing – μ s accuracy without PTP
 - iii. Internet background radiation pointed at ESnet
 - b. Unexpected traffic
 - i. There should not be any Martians – traffic that has non-forwardable address – inside ESnet. There were.
 - ii. Some were RFC 1918 addresses.
 - iii. But there were also “document addresses,” addresses that aren't forwardable that are put in hardware to match the documentation of the device. These were found inside ESnet. As an example, from a box in a lab being put on the net with some issue in its configuration on an unused port. With HT these can easily be mitigated.
 - iv. A simple SQL inquiry and good network hygiene resulted.
 - v. Another example involved LHCONE and the folks at LHC trying to determine if loose uRPF could be turned on. HT revealed that there would be issues if turned on without the issues being corrected before loose uRPF was enabled. A help in this case was the importation of LHC's routing table from a BGP listener into the SQL database which allowed HT and the BGP data to be correlated.
 - c. Timing
 - i. ESnet has used NTP for a long time. NTP is good for +/- ms. By making use of its μ s timestamps on each packet captured HT can calculate the NTP offset for each HT box and therefore for each ESnet node. This calculation is aided by the large number of packets and the statistical advantage this yields. The methodology of these calculations were cross-checked by looking at the Ciena transponders and their reported delay between two end points.
 - ii. ESnet is planning on adjusting the FPGA to further improve the accuracy of its timestamps.

- d. Internet background radiation
 - i. Looking at the same database used for the timing work, a SQL query was done to see who is sending a single syn probe and what responded with a syn-ack. The top source was Censys so not an unexpected source. It aimed at a packet at ~72k DOE desktops with a response from [a comprehensive] 1%. Other sources of these type of probes are Google Cloud Platform, China Telecom, etc. as many places are doing their own version of Censys. Now these are identified, and responses can be discussed & designed.
 - ii. How effective is this in the IPv6 world? Not very – only a few organizations are attempting.
 - iii. Besides the big warehouses probing the database can also be queried for the bottom end. Who is sending a single probe packet to a single destination? What are they up to?
 - iv. These sorts of queries are done with the flow data, not the pcap file with full packet header capture so the database is always there, not only there when something becomes of interest.

F. Questions:

- a. In IPv6 do you assume that the Protocol Header is immediately following the IP Header, or can it keep parsing headers until it comes to the protocol?
Answer: Right now, it's designed to skip past options. If you're doing source routing, it will do one. It could do more if needed. There aren't any instances on the internet so there's no requirement. Would be more work, but no limitation on what could be parsed in the FPGA. Similarly, MPLS isn't parsed at this point as there isn't a need. MPLS could be added if needed.

G. Implementations details:

- a. The goal was to make network-wide pcap files easy to use. Something a network engineer can pull into Wireshark (or whatever tool) as if all of ESnet was a single node. A lot of software to make this work. Once the filters are in place each packet of interest, regardless of where it enters & leaves ESnet will have its header captured into a pcap file for analysis.
- b. pcap files are automatically translated into a columnar format for fast data analysis.
- c. If there are pcap files and flow records, they easily fit together as they share the same identifier (a 5-tuple) and timestamp. This allows the combination of the detailed pcap file and the much broader flow records.
- d. The common technology and common "data lake" HT provides can be used to answer many different questions by the many different groups in ESnet.

H. Questions:

- a. Given two network flows, can HT show the time evolution of each? Are the two flows different or not?
Answer: Yes. Just about to do exactly this for some DTN traffic. HT is able to plot each packet in the right place.

- b. Are 8k packets good for a network?
Answer: HT should also be able to answer the question which revolves around the jumbo frame blocking a port. 1500-byte packets that are behind it can't move out even though there isn't any congestion, just head of queue blocking. The suspicion is that this results in the small packets coming in 8k micro-bursts affecting the network's traffic shape. Is this detrimental to the network?
Unknown, but first off is the need to see if it's happening.
- c. This was followed by some discussion about if jumbo frames are still good or not.

IV. JET's tasking on tools to help with inter-domain problem resolution

- Remains on hiatus while Joe Breen is engaged with other matters.

V. Operational Security Round Table:

- Jeff Bartig reported that Internet2's turning on Route Origin Validation on 16 March was "remarkably uneventful".

VI. Network roundtable

- A. DREN (Ralph McEldowmey):
 - a. DREN's main focus continues to be the transition from DREN III to DREN 4. It has tested 160 with 120 transitioned. The deadline for completing the transition remains mid-June.
 - b. In the last month all internet connections, internet transit and external peering have moved onto DREN 4.
- B. ESnet (Dale Carder):
 - a. ESnet is performing acceptance testing of the first 400G trans-Atlantic link. Two more are to be delivered this year. This circuit is on a cable ESnet is already using.
 - b. In concert with the new trans-Atlantic links ESnet is piloting traffic engineering techniques using segment routing.
 - c. ESnet is working to bring up connectivity in Los Angeles, CA, where ESnet 5 did not have a presence. ESnet 6 has add/drop there on the optical line system and a 400G capable router. ESnet will be contacting folks in the Los Angeles area for connectivity.
- C. Indiana University/International Networks: (Ed Moynihan):
 - a. No operational update - all links are stable.
 - b. Indiana University/International Networks (IU/IN) was at SupercomputingAsia earlier in the month and at APAN last week. It will be at LHC's meeting in a few weeks.
- D. Internet2 (Jeff Bartig):
 - a. James Deaton has come onboard as Ingernet2's Vice President for Network Services.
 - b. Intenet2 is working to add capacity to it I2PX commercial peering service and its cloud providers.

- E. NRL (Linden Mercer):
 - a. Starting to plan for SC23.
 - b. It was noted that the request for required circuits has been distributed by SCinet for next November in Denver.
- F. Pacific Wave: (Jonah Keough): A couple of interesting things:
 - a. Pacific Wave (PW) participated in a MANRS workshop sponsored by APNIC and the APAN meeting.
 - b. PW supported demos at OFC earlier in the month.
 - c. PW is dealing with its colocation facility in Tokyo being rebuilt while in use.
 - d. PW has announced the general availability of its route servers to any participant at its peering points. Each location has a physical server. This is a complementary service to direct peering – PW is encouraging all to make use of it. Currently some participants are using, and some are testing. It was discussed at APAN. A couple of networks there were particularly interested.
 - e. Discussion with DREN on needed paperwork for DREN 4 to connect.
- G. 3ROX (Michael Lambert): No updates today.
- H. US Ignite (Glenn Ricart):
 - a. US Ignite (USI) is working with Columbus, OH, and Albuquerque, NM, on metropolitan area fiber upgrades either done by the city or in partnership with somebody else.
 - i. If there is interest by any JET participant in making use of dark fiber in either region now would be the perfect time to make your requirements know.
 - ii. If any are in those regions and would be able to help with the project, please contact Glenn.

VII. Exchange Points Round Table

- A. StarLight (Joe Mambretti):
 - a. The installation of the DREN 4 node at StarLight (SL) has been completed.
 - b. With its international partners SL organized a Global Research Platform workshop at SupercomputingAsia (SCA) in Singapore.
 - c. There was also a planning meeting during SCA on the Data Mover Challenge. This runs during the year, concluding for SCA in 2024.
 - d. The annual Optical Fiber Conference this year built a network testbed, OFCnet, for demonstrations. SL streamed 400G to OFC in San Diego where the flow was looped back to SL giving a path length of 9,300 km. SL saw 380 gbps over the path.
 - e. SL also demonstrated some of the work it's doing with ANL and Fermilab on quantum networking on co-propagating the quantum signal over an existing, classical network's fiber. Testing progressed over initially a 10G and then a 100G classical link. Both worked. SL went booth to booth with entangled photons, then outside the convention center to Aero Drive, which is slightly longer distance. The observed fidelity was from 91% to 95%.
- B. PNWGP (Jonah Keough): No update today.

VIII. Concluding discussion

Interest was expressed on an in depth look at OFCnet at a future JET. Suggested contacts were Chris Tracy, ESnet, and Scott Kohlert, CIENA.

Meetings of Interest 2022-2023

Note: Meetings whose format has changed have been updated.

Mar 25-31	IETF 116 , Yokohama, Japan
Apr 16-19	ARIN 51 , Tampa, FL
Apr 20	HPC Summit , virtual
May 8-11	Internet2 Community Exchange , Atlanta, GA
Jun 5-9	TNC23 , Tirana, Albania
Jun 12-14	NANOG 88 , Seattle, WA
Jul 22-28	IETF 117 , San Francisco, CA
Aug 21-25	APAN56 , Colombo, Sri Lanka
Sep 18-21	Internet2 Technology Exchange , Minneapolis, MN
Sep 25-28	The Quilt Fall Meeting , Columbus, OH
Oct 8-9	GRP workshop at IEEE eScience , Limassol, Cyprus
Oct 16-18	NANOG 89 , San Diego, CA
Oct 19-20	ARIN 52 , San Diego, CA
Nov 4-10	IETF 118 , Prague, Czech Republic
Nov 12-17	SC23 , Denver, CO
Dec 12-14	AINTEC , Hanoi, Vietnam

Next JET meetings

Note: It is anticipated that most JET meetings will remain virtual for the foreseeable future

Apr 18, 2023	12-2 p.m. ET
May 9, 2023	11:45AM - 1:00PM ET, hybrid
	Note: This meeting is collocated with Internet2's Community Exchange in the Westin Peachtree Plaza, 210 Peachtree St NW, Atlanta, GA 30303
Jun 20, 2023	12-2 p.m. ET