



## 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

October 20, 2020 12:00-2:00 p.m. ET

This meeting was held virtually

### Participants

Shawn Armstrong, University of Alaska  
Lotfi Benmohamed, NIST  
Joe Breen, UTEN/University of Utah  
Nick Buraglio, ESnet  
Rob Buttler, USDA/ARS  
Rich Carlson, DOE/SC  
Bobby Cates, NASA/Ames  
James Deaton, GPN  
Basil Decina, NRL  
Dave Diller, MAX  
Bill Fink, NASA/GSFC  
Dale Finkelson, Internet2  
Ann Keane, NOAA  
Jonah Keough, PNWGP/Pacific Wave  
Kevin Kranacs, NASA/GSFC – EOS

Padma Krishnaswamy, FCC  
Yatish Kumar, ESnet  
Michael Lambert, PSC/3ROX  
Paul Love, NCO/NITRD  
Joe Mambretti, StarLight/MREN  
Linden Mercer, NRL  
Alex Moura, RNP  
Aruna Muppalla, NASA/GSFC  
Karl Newell, Internet2  
Glenn Ricart, US Ignite  
Kevin Thompson, NSF  
George Uhl, NASA/GSFC  
Chris Wilkinson, Internet2  
Matt Zekauskas, Internet2

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

### I. Action Items:

- ESnet update on its operational network security use of Rapid7.
- Internet2 and ESnet updates on their respective new networks.

II. **Review of the Minutes** of the September meeting: No corrections were received.

### III. US Ignite: Project OVERCOME - Glenn Ricart

For reference: <https://www.us-ignite.org/program/overcome/>

Project OVERCOM is an NSF award to US Ignite to find creative ways to connect the unconnected using creative ways to connect them such as TV white space, CBRS, and self-install fiber kits. The award is for five pilot communities with a one-year timeline. The award is looking for social impact. Final proposals are due January 8, 2021. See the web site above for details.

### IV. ESnet6 High Touch Services II – Yatish Kumar

- A. As part of ESnet's greenfield ESnet 6 it is instrumenting the new network with FPGA based hardware capable multiple use cases for packet editing, telemetry and other uses

to be determined. These will permit dynamic traffic engineering and other steps to increase the efficiency of ESnet 6 and enhance the user experience.

- B. The high touch hardware (as in the sense of touching of packets) will provide:
  - a. High precision, real time visibility into the network
    - i. Process every packet of interest in real time
    - ii. Ns precision timing
    - iii. Software defined functionality
  - b. This contrasts with “low touch” services with fixed functions which have been optimized for high speed and low cost
  - c. Enabling technologies
    - i. Software defined networking
    - ii. Programmable network data plane hardware with accurate timestamps
    - iii. High speed packet processing libraries such as DPDK
- C. High touch architecture
  - a. The mirror service in ESnet 6’s router replicates the first portion of each packet entering the router for selected flows and sends to the FPGA in the high touch (HT) server. This adds meta data such as a time stamp and returns to the router into a L2VPN. The L2VPN traffic then exits the router into a Mellanox 100G capable card in the HT server. The traffic can also be sent via the router’s management plane to remote HT servers to load balance for additional capacity.
    - i. The FPGA builds a new, outer header for routing the telemetry packet to the right process. The packet’s VLAN field is used to carry the generated Flow ID which maps to the queue to process the packet. As ESnet 6 management is IPv6 only there is also IPv6 header and finally a UDP header. Following are the original IP & TCP headers, then the HT high precision time stamp, flow counters plus other items from the original packet.
  - b. The Mellanox NIC separates each flow into a queue it’s maintaining on the HT server’s PCIE card where DPDK is running - one queue per available core on the card. The flows (up to 8k) are mapped to a particular queue based on a 5-tuple. The server is also running Kafka which segregates each flow at full line rate with each flow a Kafka Topic. Backend services in the HT server then take one or more Topic(s) and process for such things as TCP rate, packet loss, etc. The output from the backend services is a much lower rate - histograms, time series, etc. These products are also stored off-line for long term analysis.
    - i. Currently if a single flow is at full line rate Kafka is unable to keep up.
- D. Current FPGA hardware:
  - a. ESnet was looking for:
    - i. 100G bps port speed with a road map to higher speeds
    - ii. Timing & performance guarantees
    - iii. Easy programming (P4 style)
    - iv. Established vendor support

- b. ESnet is currently prototyping Xilinx FPGAs – Alveo U280
  - i. 2x100G port
  - ii. 8GB HBM2 memory (3.2 Tbps I/O bandwidth)
  - iii. 32GB DDR4 memory
  - iv. 1.2M logic cells
  - v. Around \$6k
- c. FPGA notes
  - i. A TCAM is used to generate the Flow ID critical to the handling of the generated telemetry packets. Currently a new flow will have its first few packets missed while the x86 sorts out and sticks the new Flow ID into the TCAM. ESnet is working towards a hash approach which will result in no missed packets.
  - ii. The P4 programmable packet distribution chain can process on the order of 150Mpps. As the telemetry packets move through the chain various flow and packet statistics counters are incremented. Additionally programmable, stateful NPUs can edit and forward the telemetry packets.
  - iii. Just before the telemetry packet exits the chain it can be sent to the HBM2 memory where QoS, shaping, flow control, etc. are done.
  - iv. HT current uses about 10% of the FPGA resources.
- d. Three models for others using FPGAs
  - i. Statistics (easy)
    - 1. Install copy of ESnet’s telemetry solution
    - 2. Zero FPGA development
    - 3. Customize, Splunk, Kentik, Grafana, etc.
  - ii. Packet editing (intermediate)
    - 1. Program the embedded NPUs
    - 2. Zero FPGA development (FPGA bit file provided)
    - 3. Packet editing no different than in an SDN switch
    - 4. Retains statistic services as above.
  - iii. Custom drivers and applications (advanced)
    - 1. Reconfigure the FPGA using P4 and Verilog
    - 2. User defined hardware – arbitrary & stateful L2-L7 solutions.
- E. Status: The above is what ESnet has or expects to have over the course of the next year.
  - a. Question: How is the collected data exposed?  
Operational data from ESnet’s production network is not available externally. It’s used internally by ESnet - e.g. its planning group. They have access to the 24x7 stats, flows with events such as packet loss, etc. The HT data is also used to enrich other stats such as SNMP.
  - b. High Touch timeline
    - i. 2019 Q3: Service design
    - ii. 2020 Q1: Design validation
    - iii. 2020 Q4: Design refinement
    - iv. 2021 Q1: Pre-pilot (gear to be placed at StarLight, Sunnyvale and Berkeley)

- v. 2022 Q1: Pre-deployment
  - vi. 2022 Q3: Deployment.
- F. ESnet 6 baseline
- a. Each ESnet 6 hub will have at least a pair of HT servers. Each will be hefty - good number of cores, etc.
  - b. The capability of sending traffic to remote HT servers can, beside just load balancing, be used to send traffic to a HT server with a specialty configuration or service software.
- G. High Touch use cases
- a. TCP rate and retransmission tracking
    - i. Monitor TCP rate on a per packet basis allows peaks and abnormal rates to be found in the shortest amount of time
    - ii. A tool for network operations and engineering
    - iii. Able to catch retransmissions as they happen – determine if the cause is within ESnet, the source network or the destinations network
    - iv. These all yield themselves to easy visual detection.
  - b. TCP congestion control identification
    - i. Detect flows unable to utilize the available bandwidth
    - ii. Detect TCP flows taking more than their fair share
    - iii. Misconfigured flow detection permits ESnet to:
      - 1. Tune the configuration of DTNs
      - 2. Automatically notify ESnet sites suboptimal configurations
      - 3. Inform fair share of the network
    - iv. Visually easy to determine BBR vs. Cubic; they each have a distinctive signature
    - v. The raw telemetry data (1:1 with the original packet) collapses well into a histogram which needs to be sampled at a much lower rate
    - vi. ESnet is exploring using machine learning on the aggregated congestion control histograms for determination.
  - c. ESnet's security group is investigating using ML to apply a microscope to the telemetry

## V. Discussion of the JET's tasking on tools to help with inter-domain issues - Joe Breen, all

Work is beginning to pick up as many participants' institutions CARES Act projects are completing.

### A. Prototype/pilot status:

- a. The University of Michigan pilot: now able to get data from across its backbone. The data from DTN sites such as the one for LP2 will soon be reachable.
- b. NCSA's pilot rejoined last month after being on hold due to campus COVID projects. Soon able to again pull in its telemetry). The project is anticipating expanded work at NCSA.
- c. Discussions continue with Clemson on a possible pilot.
- d. Other pilots are being reimaged after COVID related project load put them on hold.

- e. The capability developed by Andy Lake to traverse both Internet2 and ESnet will be demonstrated either at next month's JET or in early 2021.
- B. One additional comment on the draft Letter of Intent to Share is pending.
- C. Background on efforts lead by Eric Boyd, Joe Breen, James Deaton, Dan Doyle, Dale Finkelson 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 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\\_PNVpeT42nAxa3bW4QostMxcCHTXkWSPbZOpFwE/edit#gid=0](https://docs.google.com/spreadsheets/d/1pMW_PNVpeT42nAxa3bW4QostMxcCHTXkWSPbZOpFwE/edit#gid=0)  
The spreadsheet also has an embedded link to measurement templates for campus, regional and national networks setting out what data is desired. See: <https://drive.google.com/drive/folders/1l-LRyri16u4AvBeY6NlvyYYaINRpiByA>
  - 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. While NASA polices 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)

## VI. Operational network security roundtable

No updates were received.

## VII. Network roundtable

- A. ESnet (Nick Buraglio):
  - a. Packet procurement is well underway. ESnet is working out the details of its rolling update of packet gear. It is also identifying those locations that need a capacity boost for packet services before ESnet 6 rolls out. They will become ESnet 5.5 locations.
    - i. ESnet 6 phased roll out: OLS, ESnet 5.5 (identified in DNS as CR 5.5 vs CR 5 ), then ESnet 6.
  - b. As part of the router RFP ESnet has obtained licenses to run all of ESnet 6 in simulation with the possible exception of the High Touch services.
  - c. During the past month ESnet had a blip in their optical system due to a pushed-out software update. The silver lining was that ESnet was able to recover by way

- of the management network and its out-of-band (OoB) access via the cellular network. A couple of minor issues were encountered in the OoB access. These didn't prevent bringing the optical layer back up and have now been addressed.
- d. ESnet 6 routers will be deployed in 2021 – perhaps early.
  - B. Great Plains Network (James Deaton): Great Plains is excited about all going on around it. It's working to be ready to best leverage the new capabilities when they arrive. It's also working on some projects with its members
  - C. Internet2 (Chris Wilkinson):
    - a. Packet services: Internet2 (I2) is planning the installations of its new Cisco 8200 boxes. They'll start arriving December/January.
    - b. Part of the agreement with Cisco includes some community benefits. If you are interested in those, contact Chris: [Chris Wilkinson <cwilkinson@internet2.edu>](mailto:cwilkinson@internet2.edu)
    - c. I2 is revising its planning from the optical layer for the packet deployment – colocation planning, etc. It expects to deploy the 8200s in step with the Ciena Waveserver 5s in a pseudo greenfield environment with 400G coast to coast. Dependent on the timing of software releases they are looking to a Q2 migration from the current MX platform to the new gear. I2 is finding substantial COVID related delays in fiber delivery but still looking for Q2.
    - d. Lots of software work going on in I2 for the network.
  - D. NASA GSFC (Bill Fink/George Uhl): No update today.
  - E. NOAA/N-Wave (Ann Keane): No update.
  - F. NRL (Linden Mercer): NRL continues to develop plans for the SC20 related demos.
  - G. Pacific Wave (Jonah Keough):
    - a. Pacific Wave (PW) has finished deploying its new MX10k platform in Los Angeles, CA, Sunnyvale, CA, and Seattle, WA. All customers have been migrated onto them.
    - b. PW is doing some SC prep for demos by PW partners around AutoGOLE.
  - H. 3ROX (Michael Lambert): No updates for PSC, 3ROX or XSEDE.
  - I. University of Alaska (Shawn Armstrong): No network updates.

## VII. Exchange Points Round Table

- A. MAN LAN and WIX (Dale Finkelson):
  - a. MAN LAN: NEA3R has added a second 100G to Europe. It's being brought up this week. Internet2 is working to finish up the NSI implementation.
  - b. WIX: No changes.
- B. MAX (Dave Diller): Several internal changes: moving out of a pop, out of one data center and into another, and helping NOAA relocate TICAP equipment due to the data center changes.
- C. StarLight (Joe Mambretti): The demos for the virtual SC20 are coming together. They'll include:
  - a. Linden Mercer from NRL and Bill Fink and Paul Lang from GSFC.
  - b. AutoGOLE.
  - c. Pacific Research Platform and Global Research Platform.

- d. There will be a 400G circuit from StarLight (SL) to Ottawa, ON, with a couple of 100G circuits to the East Coast and a number of 100G circuits to elsewhere around the globe.
- e. The results will be presented at the XNet forum, the INDIS forum, and a forum being organized with SL's research partners on the 19<sup>th</sup>.
- D. Ames (Bobby Cates): No updates.
- E. PNWGP (Jonah Keough): No updates.

#### **VIII. Determination of date and time of the JET's meeting in November 2020**

After discussion on possible conflicts with the SC20 agenda none on the call found they had a conflict. Nor was any email received regarding conflicts with the JET's normal day and time (third Tuesday, 12-2 PM eastern). With that it was agreed that the JET will hold its November meeting at the usual day and time (November 17, 12-2PM eastern).

#### **IX. Update on the LSN's annual planning retreat as pertains to the JET – Rich Carlson**

The LSN accepted the proposed tasking for the JET for 2021. The LSN would also like to start receiving more detailed briefings from the JET and the LSN's three other teams. The LSN will look at possible tasks for the teams to work on together with a report back to the LSN.

For reference the JET's tasking for CY2021 is:

- A. Ongoing JET tasks:
  - a. Assist in the planning of technology and application demonstrations of SDN & Big Data at SC21
  - b. Technology tracking: perfSONAR, SDN/SDX/SDI, Science DMZs, network automation & orchestration, and segment routing
  - c. Hold two meetings collocated with R&E networking community conferences:
    - i. Internet2 Global Summit (18-21 April)
    - ii. SC21 (November)
  - d. Continue to schedule meeting round tables of updates on members' networks, operational network security, exchange points and meetings of interest to the community
  - e. Continue coordinating the development of tools to monitor cross-domain workflows and automate the detection of transport issues. Additionally facilitate the sharing of measurement data between networks - anonymized as needed.
- B. Potential JET Workshop: TBD

#### **Meetings of Interest 2020**

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

- Oct 14-15 & 23      [ARIN 46](#), in person cancelled, moved to a virtual meeting
- Oct 19-21            [Fórum RNP 2020](#), virtual meeting
- Oct 19-21            [NANOG 80](#), in person cancelled, moved to a virtual meeting
- Oct 22-23            [ESCC](#), virtual meeting
- Nov 2                 [TechEXtra: perfSONAR Day](#), virtual meeting
- Nov 9-10            [SC20 Tutorials](#), virtual meeting

Nov 11-13 [SC20 Workshops](#), virtual meeting  
Nov 12 [TechEXtra: Route Origin Validation in the Cloud](#), virtual  
Nov 14-20 [IETF 109](#), in person cancelled, moved to a virtual meeting  
Nov 16-19 [SC20](#), in person cancelled, moved to a virtual meeting  
Dec 8 [TechEXtra: Continuous Integration & Delivery](#), virtual  
Jan 17-20, 2021 [PTC'21](#), in person cancelled, moved to a virtual meeting  
Jan, dates TBA [HIC](#), virtual  
Feb 1-5 [APAN 51](#), virtual  
Feb 8-10 [NANOG 81](#), virtual  
Feb 9-11 [The Quilt Winter Member Meeting](#), virtual

### **Next JET meetings**

*Note: It is anticipated that JET meetings into CY2021 will be virtual due to COVID-19 guidelines.*

Nov 17 12-2 p.m. ET *Confirmed during the October meeting*  
Dec 15 12-2 p.m. ET *n.b.: Will be held only if needed*  
Jan 19, 2021 12-2 p.m. ET