



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

November 17, 2020 12:00-2:00 p.m. ET

This meeting was held virtually

Participants

Jeff Bartig, Internet2	Paul Love, NCO/NITRD
Nick Buraglio, ESnet	Chris Lowe, USDA/ARS
Rich Carlson, DOE/SC	Joe Mambretti, StarLight/MREN
Bobby Cates, NASA/Ames	Linden Mercer, NRL
James Deaton, GPN	Alex Moura, RNP
Basil Decina, NRL	Ed Moynihan, Indiana University
Dave Diller, MAX	Aruna Muppalla, NASA/GSFC
Bill Fink, NASA/GSFC	Frank Seesink, University of North Carolina
Andrew Gallo, CAAREN/GWU	Kevin Thompson, NSF
Ann Keane, NOAA	George Uhl, NASA/GSFC
Jonah Keough, PNWGP/Pacific Wave	Chris Wilkinson, Internet2
Michael Lambert, PSC/3ROX	Matt Zekauskas, Internet2
Paul Lang, NASA/GSFC	

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

I. Action Items:

- Discuss topics for the JET to brief to the LSN.
- Internet2 and ESnet updates on their respective new networks.

II. Review of the Minutes of the October meeting: One correction was received.

III. SC20 Advanced Networking for Data Intensive Science: Experiments and Demonstrations – **Joe Mambretti, Linden Mercer, Bill Fink and Paul Lang**

For reference, the slides for this talk are online at:

https://www.nitrd.gov/nitrdgroups/index.php?title=JET-Meetings-2020#November_17.2C_2020

These demonstrations were also presented and discussed at the INDIS, NRE and XNet sessions at SC20.

An overview of the large scale, data intensive science demonstrations that StarLight (SL), Goddard Space Flight Center (GSFC), the Naval Research Laboratory (NRL) and the International Center for Advanced Internet Research (iCAIR) at Northwestern University undertook for the virtual SC20. Without a physical SC this year as a focus for network demonstrations SL, GSFC, NRL and iCAIR worked with Ciena, Internet2, ESnet and CANARIE to build a demo testbed with bandwidth ranging from 100G to 400G.

- A. StarLight: The Global Research Platform (Joe Mambretti)
 - a. Architecture: “Global Science DMZ”
 - b. Services optimized for science workflows
 - c. High performance transport over WANs for large capacity data streams end-2-end
 - d. Enhanced virtualization
 - e. Highly programmable infrastructure
 - f. Specialized components (software stacks, next gen DTNs, advanced APIs, customized devices, federation, INT, P4, etc.)
 - g. Interdomain dynamic provisioning
 - h. Production resources adjacent to testbeds
 - i. Composed of multiple component Research Platforms around the world primarily interconnected at layer 2
 - j. The Global Research Platform (GRP) uses the infrastructure built by the Global Lambda Integrated Facility
 - k. StarLight hosts many specialized services and testbeds facilitating their use by the many internationals and national R&E networks connected at SL.
 - i. One of these is an NSF funded SDX - an IRNC award which allows many different pieces to be connected to its variety of services including the GRP
 - l. Several R&E organizations in the US have built interoperable SDXs. They in turn are part of the international AutoGOLE fabric
 - m. SL organized demonstrations for SC included:
 - i. DTN-as-a-Service
 - ii. ROCE over WAN
 - iii. P4 experiments and demonstrations
 - iv. Kubernetes Federation (with UCSD)
 - v. BigData Express (with FNAL)
 - vi. AutoGOLE/NSI/OpenNSA
 - vii. ROBIN (Rucio/BigData Express/SENSE) demo between SL and CERN (with FNAL and CERN)
 - viii. Composable Platform as a Service (using the Liquid infrastructure system) (with UIC)
 - ix. Petascale science using NVMe over Fabrics (NVMe-oF)
 - n. Next GRP meeting will be September 13-17, 2021, in Innsbruck, Austria
- B. Naval Research Laboratory: Resilient Distributed Processing (Linden Mercer)
 - a. The challenge:
 - i. Data to decision with global dispersion of a huge number of data sources and sinks
 - ii. Processing on supercomputers – dynamically distributed
 - iii. Faster and better decision cycles
 - b. The demo:
 - i. Resilient distributed processing of large volumes of data
 - ii. Across geographically dispersed HPC and network resources

- iii. Rapidly changing resources to meet demand and availability
 - iv. Leverage additional capacity (beyond 2x100G SL<>NRL) as available
 - v. Use of improved RDMA long distance performance for terabyte bulk transfers with a goal of 98% link utilization. (This comes about due to performance issues at SC19 with 1MB RDMA transfers over both InfiniBand and Ethernet on distances of 12k km. NRL worked with Mellanox who provided a correction in their generally available firmware in August.)
 - vi. The performance needs to be very deterministic. To that end:
 - 1. Fast (sub second) fault detection and location using active probes
 - 2. Sensing with remote probes and control to obtain lossless WAN data flow even with congestion at remote location. (Buffering can be tolerated, drops cannot.)
 - 3. Policy driven, dynamic network control
 - 4. Use of Kafka messages
 - vii. Traffic flow security equivalent to legacy TDM
 - 1. NRL is working with IEEE and IETF to update security standard as network data encryption moves from data encryption to network encryption. IEEE MACsec 802.1 AEAD and IETF IPsecME WG TFS
<https://datatracker.ietf.org/doc/draft-ietf-ipsecme-iptfs/>
<https://1.ieee802.org/security/802-1aedk>
 - 2. Satisfies full range of network TRANSEC requirements
- C. Goddard Space Flight Center: Disk-to-disk WAN file transfers using NVMe-oF/TCP (Bill Fink and Paul Lang)
- a. Continuation of GSFC's decade long project to do high speed, disk-to-disk (D2D) transfers over long distances.
 - b. This year's hardware was a loaner Supermicro chassis with two AMD Epic processors each with 16 cores and 10 x16 gen4 PCIE slots. The slots to be populated with five dual 200G port NICs and five PCIE adaptors each able to hold four NVMe SSD drives.
 - c. With the pandemic and ESnet being in the middle of deploying ESnet 6 the anticipated 400G circuits were not available. The network to be used for testing was the CENI testbed with 200G GSFC<>SL. Either as a loopback to GSFC or making use of some older GSFC hardware already in SL. From SL a loop was also available of SL<>NYC (32 Avenue of the Americas)<>Montreal<>Ottawa<>SL with segment speeds ranging from 100G to 300G. With the loop the time from GSFC to SL, around the loop and then back to GSFC was 80ms.
 - d. This year the transport was again NVMe-oF with TCP (NVMe-oF/TCP). In the past NVMe-oF used IWARP for transport. This had some issues. Last year the Linux kernel was starting to support the use of TCP for transport and this has had more success. At SC19 GSFC was able to read at 200Gbps over the wide area. There are still some issues with writing over the wide area which GSFC is investigating.
 - e. For quite a while the bottleneck was processors – all the CPUs would saturate. With NVMe-oF/TCP that was no longer the case. Rather the bottleneck was the

total bandwidth of the server. The new chassis, etc. should get past that bottleneck. It will be installed in GSFC and WAN testing will start over a 100G path to SL.

- i. On the horizon are new NVMe Gen4 M.2 drives that use the new Phison E18 controller. These are claimed to be able to do read and writes at 7GB.

IV. Operational network security roundtable

No updates were received.

V. Network roundtable

- A. CAAREN (Andrew Gallo): Nothing to report.
- B. ESnet (Nick Buraglio): Very deep implementing ESnet 6.
 - a. The optical plant is functionally done.
 - b. Preparing for the packet network installation. Scheduled to start 1Q2021. The procurements are underway.
 - c. ESnet's staff is growing including several new network engineers to handle the additional workload.
- C. International Networks – Indiana University (Ed Moynihan):
 - a. NEA³AR (née NEAAR): One of the new trans-Atlantic 100G circuits is due in a few days, the second will follow.
 - b. TransPAC:
 - i. The GOREX relocation started early last month and should be completed.
 - ii. There are discussions with Pacific Wave, Internet2 and SingAREN on a potential 100G circuit Guam<>Singapore.
- D. Internet2 (Chris Wilkinson): Interenet2 (I2) is working through both the optical and packet services for its NGI.
 - a. For the optical deployment Lumen is its prime contractor.
 - i. Most of Kansas, Missouri, Illinois and Indiana are completed along with northern Ohio and New York from Albany west. The built Flexgrid system extends beyond these areas.
 - ii. Work is currently happening from the mid-Atlantic to the Northeast. It's anticipated that Charlotte, NC, to Boston, MA, will be done this week.
 - iii. The west coast will be next followed by the southwest and south.
 - iv. 1Q2021 is the target for the new backbone to be up.
 - b. Packet services:
 - i. I2 is working with its packet install partner, GDT, to stage and preconfigure the hardware. GDT will be providing four install teams.
 - ii. Hardware staging and colo preparation are in progress. A preliminary target for service transition is May-June 2021.
- E. NASA GSFC (Bill Fink/George Uhl): No update today.
- F. NOAA/N-Wave (Ann Keane): No update.
- G. NRL (Linden Mercer): Nothing beyond what was reported above.
- H. Pacific Wave (Jonah Keough):
 - a. No major updates.

- b. Pacific Wave has completed the deployment of its layer 3 upgrade.
- c. It is investigating possible synergies with Internet2 on their shared west coast backbone.
- I. RNP (Alex Moura): No update.
- J. 3ROX (Michael Lambert): No updates for 3ROX or XSEDE. For PSC the 100G circuit between the PSC machine room and 3ROX is now in place. This is implemented between a pair of Ekinops using an alien wave over the exiting 15454 infrastructure. The upgrade is in support of PSC's newest supercomputer, Bridges-2. PSC is looking for a December turnup with a friendly user period beginning sometime in January.
- K. USDA – ARS (Chris Lowe): No update this month.

VI. Exchange Points Round Table

- A. PNWGP (Jonah Keough): No updates.
- B. StarLight (Joe Mambretti):
 - a. StarLight (SL) currently supports around 25 research and computer science testbeds. SL is moving to make these persistent . The motivation for this is experience with the virtual SC.
 - b. SL is participating in NSF's exploration of Japan-US joint research in networking. There may be an opportunity for additional testbeds to support this joint research.
 - c. SL's FABRIC node is being assembled. SL anticipates its arrival in early January.
 - d. For both FABRIC and ESnet 6 SL is exploring a new antenna that will allow for very precise time stamps for the High Touch hardware being developed by ESnet. FABRIC is developing its own tools for the ESnet developed hardware. SL will make the high precision time stamps a service available to others.
- C. MAN LAN and WIX (Chris Wilkinson): No update this month.
- D. MAX (Dave Diller): No updates. MAX is also looking forward to the arrival of its FABRIC node.
- E. Albuquerque (Nick Buraglio): This new IXP is about ready to accept customers. It has nodes in the three main locations in Albuquerque. The remaining piece needed is fiber into one node (the other two already have). This IXP is being done as a community service and has either a small or no connection fee. For more information contact Nick at:

[Nick Buraglio <buraglio@es.net>](mailto:nick.buraglio@es.net)

VIII. LSN update – Rich Carlson

The LSN welcomes suggestions of topics on which its teams – JET, MAGIC and BRD - can update it. A discussion of possible topics will be on the agenda for the JET's next meeting. If anybody has a possible topic, please bring it up then during the January meeting. You can also let Paul Love, Kevin Thompson and Rich Carlson know ahead of time:

[E Paul Love - CTR <Love@nitrd.gov>](mailto:E_Paul_Love - CTR <Love@nitrd.gov>)

[Kevin Thompson <kthompson@nsf.gov>](mailto:Kevin.Thompson <kthompson@nsf.gov>)

[Richard Carlson <Richard.carlson@science.doe.gov>](mailto:Richard.Carlson <Richard.carlson@science.doe.gov>)

Meetings of Interest 2020

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

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
Feb 1-5	APAN 51 , virtual
Feb 8-10	NANOG 81 , virtual
Feb 9-11	The Quilt Winter Member Meeting , virtual
Feb 17-18	HIC , virtual
Mar 8-12	IETF 110 , virtual
Apr 11-14	ARIN 47 , Tampa, FL
Apr 18-21	Internet2 Global Summit , virtual

Next JET meetings

Note: It is anticipated that many of the JET's meetings in CY2021 will be virtual due to COVID-19 guidelines.

Dec 15	12-2 p.m. ET <i>n.b.: Will be held only if needed</i>
Jan 19, 2021	12-2 p.m. ET
Feb 16, 2021	12-2 p.m. ET