SC26 Network Research Exhibition: Preliminary Abstract

Resilient, Performant Networks and Distributed Processing

Basil Decina, Linden Mercer, Dardo Kleiner, Naval Research Laboratory,
Joe Mambretti, Jim Chen, International Center for Advanced Internet Research - Northwestern University

Abstract

This demonstration will build on our previous NRE demonstrations with significant advancements. We aim to show intelligent, dynamic arrangement and re-arrangement of widely distributed processing of large volumes of data across a set of compute and network resources organized in response to resource availability and changing application demands. We also aim to explore performance limitations and enablers for high volume bulk data transfers. Our network will be assembled using a number of switches and multiple SCinet Tbps connections. We plan to show rapid automated deployment and redeployment with prioritized optimization, real-time monitoring and QOS management of application data flows with very different network demands. Technologies we intend to leverage include SDN, RDMA, RoCE, NVMe, GPU acceleration and others.

Goals

Similar to previous efforts but planned SC25 focus is on prioritized dynamic network monitoring and control, and Tbps enabled workflows.

1. Network deployment, monitoring, reporting, and redeployment.
2. Tbps RDMA performance over global distance for timely Terabyte bulk data transfers (building on previous demonstrations of 10-20 second Tbyte NVMe to NVMe transfers on N by 400G network).
3. Prioritized dynamic shifting of processing and network resources from one location/path/system to another (in response to demand and availability).

Impacts

The expected outcomes of this NRE are confirmation of the stated goals in a complex, large scale environment that exceeds what our organization could expect outside of this collaborative environment. The leverage here is substantial not only in terms of the scale of the infrastructure but even more so in the benefit of working together with the SCinet team and other NRE collaborators. Accomplishments here translate into confident early transition of technologies into operational environments.

Resources

For this set of demonstration goals, we would like to have use of multiple Tbps WAN connections from VA/DC to StarLight in Chicago and on to SC26 and the StarLight booth. We will leverage in place expanded CENI and anticipated FABRIC Tbps connections between StarLight and JBDT in McLean, VA. Initial JBDT consortium network submission NRI103.

Here are some details:

- Primary locations: NRL, JBDT in McLean, VA, StarLight at NWU in Chicago and of course SC26 at the McCormick in Chicago.
- Bandwidth to support multiple Tbps application flows.
- Multiple paths between locations to allow more concurrent activities and network control/restoration test/demos (multiple - time shared - with full restoration option).
- Multiple Tbps systems established at StarLight and McLean (and possibly Berkeley), by iCAIR, NRL, NASA. We will be maturing the substantial aggregation of 400G systems and extending this work to include new early access Gen 6 systems with Gen 6 NVMe storage and 800 Gbps network cards we hope to obtain late in August.
- Engaging with vendors to augment and expand.

Involved Parties

- StarLight, Joe Mambretti, Jim Chen
- Mid-Atlantic Crossroads (MAX), Dave Diller
- CENI
- SCinet
- ESnet Testbed
- Naval Research Laboratory (NRL)
- Other vendors and collaborators

JBDT SC26 WAN (NW iCAIR, NRL)

- Primary locations:
  JBDT at MAX in McLean, VA;
  StarLight at NW, Chicago, IL
  SC26 in Chicago, IL
  ESnet TB in Berkeley, CA
  CENIC DTN in LA, CA
  UCSD in SD, CA
- Bandwidth to support multi-Tbps application flows
- Multiple paths between locations to allow more concurrent activities and network control/restoration test/demos (multiple – time shared – with full restoration options)
- 800G Gen5 data transfer systems established at StarLight, Mclean, Berkeley by iCAIR, NRL, NASA (adding new Gen5 NVMe systems)
- Negotiating Gen6 early access systems
- Expecting 800G switches
- Engaging with multiple vendors to augment and expand

