SC25 Network Research Exhibition: Demonstration Abstract

400 Gbps E2E WAN Services: Architecture, Technology and Control Systems

Joe Mambretti, Jim Chen, Fei Yeh
International Center for Advanced Internet Research -
Northwestern University, j-mambretti, jim-chen, fyeh@northwestern.edu,
Linden Mercer, Naval Research Laboratory, linden.mercer.ctr@nrl.navy.mil,
Bill Fink NASA Goddard Space Flight Center, bill@sci.gsfc.nasa.gov

Abstract

Data production among science research collaborations continues to accelerate, a long term trend, in part, propelled by large scale science facilities, including high luminosity research instruments and next generation HPC computational science centers. Consequently, the networking community is preparing for high performance WAN services, including 400 Gbps, 800 Gbps and 1 Tbps WAN and LAN services. In this progression, 400 Gbps E2E WAN services, the focus of these NRE demonstrations, comprise a key building block. Currently, the requirements and implications of 400 Gbps WAN services are being explored at scale, including 400 Gbps E2E over thousands of miles across multiple domains. Recently, these techniques were used to stream 800 Gbps E2E WAN services over 4,000 at the OFC conference in San Francisco, supported by an OFCnet path between the venue and the StarLight Exchange in Chicago. The 400 Gbps NRE demonstrations described here will showcase 400 Gbps E2E WAN services from the StarLight International/National Communications Exchange Facility in Chicago to the SC25 venue, between StarLight and the multi-agency Joint Big Data Testbed (JBDT) Facility in McLean, Virginia, and between the JBDT Facility and the SC25 venue.

Goals

With its research partners including the SCinet WAN group and the StarLight Consortium, the International Center for Advanced Internet Research (iCAIR) at Northwestern University is designing, implementing and demonstrating an E2E 400 Gbps WAN services among several sites. Multiple issues related to 400 Gbps transport are being investigated and resolved to enable to transition to 400 Gbps WAN services.

1. At the ends of all current 400 E2E paths, 400 Gbps Gen5 Data Transfer Nodes (NREs) have been implemented, each with a 400 Gbps smart NIC.
2. These DTNs are connected at 400 Gbps to switches with deep buffers.
3. These switches are interconnected to high performance WAN optical transport switches, which support high performance WAN services on open line systems.
4. These demonstrations will leverage iCAIR's experimental research into the optimal design, configuration, components, and integration technologies for 400 Gbps (DTNs), including techniques for kernel bypass using zero-copy for memory and disk copy to avoid bottlenecks in data transfers over 400 Gbps WANs, NVMe over fabric, and optimal affinity bindings for NUMA architecture for optimal resource utilization
5. The demonstrations will also showcase software middleware for reliable high-speed network data transfer to orchestrate infrastructure resources for optimal high performance transfers.
6. In addition, measurement techniques for real-time monitoring, benchmarking and evaluation at 400 Gbps LAN and WAN will be shown.
7. Other technologies being investigated are high performance optical transport switches, 400 Gbps ZR/ZR+ transceivers (for distances of over 1,000 miles) and 400 Gbps breakout cables.

Impacts

1) This series of research projects is directed at developing resources to meet the demands of increasingly large capacity WAN transport for large scale data intensive science. This NRE is focused on WAN infrastructure that supports a core WAN transport building block – 400 Gbps E2E infrastructure and services.
2) SCinet supports capabilities that enable innovative 400 Gbps WAN techniques to be

demonstrated at a national and international scale.

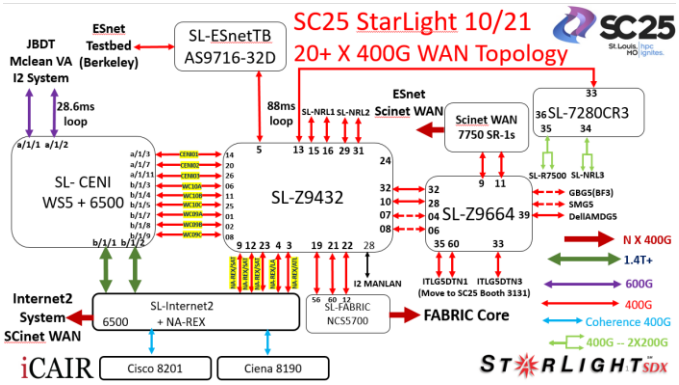
3) Multiple next steps include migration to Gen6 servers, 800 G smart NICs, integrating new protocols, integrating ZR/ZR+ coherent optics, developing additional monitoring and analytics using AI techniques etc.

Resources

Required resources from SCinet WAN are 1.2 Tbps E2E WAN services from the StarLight International/National Communications Exchange Facility in Chicago to the SC25 venue, between StarLight and the JBDT Facility in McLean, Virginia, and between the JBDT Facility and the SC25 venue. Another site utilized will be NERSC, which has a 400 Gbps ESnet testbed connection to the StarLight Facility.

Involved Parties

- Joe Mambretti, iCAIR, j-mambretti@northwestern.edu
- Jim Chen, iCAIR, jim-chen@northwestern.edu
- Fei Yeh, iCAIR,fyeh@northwestern.edu
- Rod Wilson, Ciena, rwilson@ciena.com
- Marc Lyonnais, Ciena, mlyannai@ciena.com
- Scott Kohlert, Ciena, skohlert@ciena.com
- Gaurauvdeep Shami, Ciena, gshami@ciena.com
- Linden Mercer, NRL, linden.mercer.ctr@nrl.navy.mil
- Basil Decina, NRL, basil.decina@nrl.navy.mil
- Bill Fink NASA Goddard Space Flight Center, bill@sci.gsfc.nasa.gov,
- StarLight International/National Communications Exchange Facility and Consortium
- Metropolitan Research and Education Network (MREN)
- SCinet


SC25 StarLight 10/21
20+ X 400G WAN Topology
JBDT
Mclean VA
I2 System
ESnet
Testbed
(Berkeley)
SL-ESnetTB
AS9716-32D
28.6ms
loop
88ms
loop
ESnet
Scinet WAN
Scinet WAN
7750 SR-1s
SL-7280CR3
SL-R7500
SL-NRL3
SL-NRL1
SL-NRL2
SL- CENI
WS5 + 6500
SL-Z9432
SL-Z9664
GBGS(BF3)
SMG5
DellAMDG5
N X 400G
1.4T+
600G
400G
Coherence 400G
400G -- 2X200G
ITLGSDTN1
(Move to SC25 Booth 3131)
ITLGSDTN3
I2 MANLAN
Internet2
System
SCinet WAN
SL-Internet2
6500
+ NA-REX
SL-FABRIC
NCS5700
FABRIC Core
iCAIR
Cisco 8201
Ciena 8190
STARLIGHT SDX