SC22 Network Research Exhibition: Demonstration Preliminary Abstract

"Demonstrations of 400 Gbps Disk-to-Disk WAN File Transfers using **NVMe-oF/TCP**"

Bill Fink, NASA Goddard, bill.fink@nasa.gov

Abstract

NASA requires the processing and exchange of ever increasing vast amounts of scientific data, so NASA networks must scale up to ever increasing speeds, with 400 Gigabit per second (Gbps) networks being the current challenge. However it is not sufficient to simply have 400 Gbps network pipes, since normal data transfer rates would not even fill a 10 Gbps pipe. The NASA Goddard High End Computer Networking (HECN) team will demonstrate systems and techniques to achieve near 400G line-rate disk-to-disk data transfers between a pair of high performance NVMe Servers for SC22 across Scinet national wide area 2x400G network paths, by utilizing NVMe-oF/TCP technologies to transfer the data between the servers' PCIe Gen4 NVMe drives.

Goals

1. Construct a custom built high performance NVMe server capable of sustaining near 400 Gbps disk data transfers across 400G network paths
2. Tune and optimize the system configuration, including disk, network, PCIe, CPU and memory subsystems, to eliminate potential performance bottlenecks
3. Since previous demos identified CPU resources being the primary blockade to surpass the 200 Gbps performance level, investigate using NVMe-oF/TCP technology to allow offloading the CPU processing
4. Demonstrate near 400 Gbps disk-to-disk network data transfers across real world 400G WAN network paths
5. Determine and hopefully eradicate the bottlenecks that limited network performance during SC21

Resources

- Systems with sufficient PCIe Gen4 bandwidth and number of slots
- 200G NICs using NVMe-oF/TCP technology
- High performance PCIe Gen4 NVMe drives
- nuttcp/nuttscp network performance measurement and network data transfer tools
- 400G network switches and optics
- 2x400G SCinet WAN network paths forming a triangle between NASA Goddard NVMe servers in McLean, VA, StarLight in Chicago, IL, and SC22 in Dallas, TX (being arranged by SCinet WAN team in coordination with Joint Big Data Testbed (JBDT) high performance network R&D partners Joe Mambretti / Jim Chen (StarLight) and Linden Mercer (NRL))

Involved Parties

- Dave Diller, Mid-Atlantic Crossroads (MAX), dave@maxgigapop.net
- Linden Mercer, Naval Research Laboratory (NRL), linden.mercer.ctr@nrl.navy.mil
- Joe Mambretti & Jim Chen, StarLight, j-mambretti@northwestern.edu, jim-chen@northwestern.edu
- SCinet SC & WAN Team
- Esnet
- Internet2
- Ciena
- Wilbur Smith, Arista Networks, wilbur.smith@arista.com
- James Wynia, Dell, James_Wynia@DELL.com
- Mellanox