SC19 Network Research Exhibition: Demonstration Preliminary Abstract

"Demonstrations of 400 Gbps Disk-to-Disk WAN File Transfers using RDMA and NVMe Drives"

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

Paul Lang, NASA Goddard, paul.lang@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 multi-100 Gigabit per second (Gbps) networks being the current challenge. However it is not sufficient to simply have 100 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 high performance NVMe Server at SC19 to or from a pair of high performance NVMe servers across two national wide area 4x100G network paths, by utilizing RDMA technologies to transfer the data between the servers' NVMe drives.

Goals

1. Construct a custom built high performance NVMe server capable of sustaining near 400 Gbps disk data transfers across 4x100G 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 RDMA technology to allow offloading the CPU processing
4. Demonstrate near 400 Gbps disk-to-disk network data transfers across real world 4x100G WAN network paths
5. Determine and hopefully eradicate the bottleneck in the RDMA engine that limited network performance during SC18

Resources

- Systems with sufficient PCIe bandwidth and number of slots
- 100G NICs with RDMA capabilities
- High performance NVMe drives
- nuttcp/nuttscp network performance measurement and network data transfer tools
- 100G network switches and optics
- 4x100G WAN network paths between NASA Goddard and SC19, and between StarLight and SC19 (being arranged by Joe Mambretti and Linden Mercer together with Tom Lehman)

Involved Parties

- Tom Lehman, Mid-Atlantic Crossroads (MAX), tlehman@umd.edu
- Linden Mercer, Naval Research Laboratory (NRL), linden@cmf.nrl.navy.mil
- Joe Mambretti, StarLight, j-mambretti@northwestern.edu
- SCinet
- CenturyLink
- Fujitsu
- Marc Lyonnais, Ciena, mlyonnai@ciena.com
- Wilbur Smith, Extreme Networks, WSMITH@etremenetworks.com
- Bob Bean, Dell, Bob.Bean@dell.com
- Tom Reu, Chelsio, tomreu@chelsio.com
- Troy Leedberg, Chelsio, troy@chelsio.com

SC19

Demonstrations of 400 Gbps Disk-to-Disk WAN File Transfers using RDMA and NVMe Drives

An SC19 Collaborative Initiative Among NASA and Several Partners