SC22 Network Research Exhibition: Demonstration Preliminary Abstract

Transfers above 100gbit/s using EScp

Charles Shiflett, ESnet, cshiflett@es.net

Abstract

EScp, a high performance transfer tool with an interface similar to SCP, is being developed with the goal of providing a way of transferring data at the line rate of the network interface.

Recently we have seen two significant changes in how we approach high performance transfers, one is that physical CPU sockets consist of multiple chiplets which are bonded together. The 2nd change is that 400gbit WAN networks are available and thus the desire to transfer at 400gbit/s.

This demo will show a transfer on a 400 gbit/s WAN link, how we arranged the system to maximize performance, what we changed to support these new architectures, and what steps are needed to support data transfers at 400gbit/s.

Goals

While we would like to show 400gbit/s transfers, there are a number of issues that will likely prevent that from occurring:

1. It's still hard to pack enough disks in a system to support a 400gbit/s (46 GiB/s) transfer.
2. The NUMA domain w/ chiplet based CPU Sockets means you have a complicated memory model and your Ethernet card is no longer directly connected to all the cores in a socket.
3. There is a limit on the core count available to support these transfers.
4. Since this demo is showing transfers using standard API's / tooling on Linux, it inherits all the limitations on that platform.

Resources

This demo will use the ESnet testbed infrastructure to demo the transfer. While this demo is not intended to require anything special beyond what the testbed provides, the expectation is that:

- 400 gbit/s WAN connection will be available.
- The testbed system(s) will have an AMD chiplet based processor configured in a way to optimize performance.
- The network backend will consist of multiple (>=2) 200gbit ports, ideally w/ those ports connected to different chiplets and/or sockets.
- The system will be optimized to support high speed transfers (BIOS performance settings, NUMA domain set to expose AMD chiplets, network stack optimized for WAN transfers).

Involved Parties

This demo was put together with the help of people working at ESnet.