SC22 Network Research Exhibition: Demonstration Preliminary Abstract

SciStream: Mem-to-Mem Scientific Data Streaming over a Wide Area Network

Chengyi Qu, University of Missouri, cqy78@mail.missouri.edu

Abstract

Modern scientific instruments, such as detectors at synchrotron light sources, generate data at such high rates that online processing is needed for data reduction, feature detection, and experiment steering. The same high data rates also demand memory-to-memory streaming from instrument to remote high-performance computers (HPC), because local computational capacity is limited and data transmissions that engage the file system introduce unacceptable latencies. To address these issues, we developed SciStream, a middlebox-based architecture with control protocols to enable efficient and secure memory-to-memory data streaming between producers and consumers that lack direct network connectivity. SciStream operates at the transport layer to be application agnostic, supporting well-known protocols such as TCP, UDP, and QUIC. In this demonstration we will emulate a light source data acquisition workflow streaming data from the SC showfloor in Dallas to a compute node in StarLight, Chicago. We will also demonstrate a workflow that uses the FABRIC testbed to connect StarLight with the Chameleon Cloud infrastructure. The mem-to-mem data streaming over the wide area network (WAN) will be enabled by SciStream.

Overview

As scientific instruments generate data at rates on the order of tens of gigabytes per second, real-time analysis of streaming data has emerged as a solution to cope with this new paradigm. Furthermore, rapid analysis of generated data may permit real-time feedback and experiment steering. However, this often requires computational capabilities greater than those available at a single experimental facility---or may require the use of specialized computer systems. Thus, the use of remote high-performance computing (HPC) for analysis is becoming more common in large scientific streaming workflows.

We note that scientific streaming applications differ from many other streaming applications in their high throughput requirements (a single application may require >10 Gbps) and the fact that the data producers (e.g., data acquisition applications on scientific instruments, simulations on supercomputers) and consumers (e.g., data analysis applications on HPC systems) are typically in different security domains (and thus require bridging of those domains). The SciStream system establishes the necessary bridging and end-to-end authentication between source and destination, while enabling efficient memory-to-memory data streaming.

In our preliminary evaluations, however, we discovered that a naive implementation of SciStream's Data Server (S2DS) can introduce significant jitter into an end-to-end streaming pipeline.

Moreover, measurements obtained from a real testbed were different from those obtained on an emulated scenario. To get SciStream ready for production environments, we first need to evaluate it at scale. Fortunately, SCinet's NRE provides a suitable environment to test SciStream at scale thanks to its 200Gbps WAN connection between SC showfloor (Dallas, TX) and the StarLight IXP (Chicago, IL).

In this demo, we will experiment with several implementation approaches for SciStream's S2DS that may alleviate the large jitter introduced by our initial prototype, and we will evaluate their performance on SCinet facilities under different WAN setups. Furthermore, we will present an early prototype of a streaming application using the QUIC transport protocol over a SciStream-enabled infrastructure.

Goals

1. Demonstrate on-demand memory-to-memory streaming of scientific data utilizing SCinet facilities and the FABRIC testbed.
2. Show SciStream's various transport protocol support including TCP, UDP, and QUIC.
3. Evaluate SciStream on a real WAN with an emulator of scientific data streaming and multiple WAN connections.

Resources

[Any known needs which include, but are not limited to, Wide-Area connections, network technology, controllers, software tools, monitoring needs, etc.]

For our demonstration, we will use compute resources hosted at the StarLight booth in the SC showfloor. We require a 200 Gbps circuit from the StarLight booth in Dallas to the StarLight IXP in Chicago. We also plan to use the FABRIC testbed, which has several sites connected by a 100 Gbps WAN across the US (one of those sites is StarLight). In addition, at least 4 computer nodes, with a CPU of at least 32 cores and memory of 128 GB, are needed.

No need for further physical network devices. All equipment can be reserved on FABRIC or provided by StarLight. No

software tools needed. Need for a monitor at the StarLight booth.

Involved Parties

- Chengyi Qu, University of Missouri, cqy78@mail.missouri.edu
- Joaquin Chung, Argonne National Laboratory, chungmiranda@anl.gov
- Raj Kettimuthu, Argonne National Laboratory, kettimuth@mcs.anl.gov
- Joe Mambretti, StarLight, j-mambretti@northwestern.edu
- Jim Chen, StarLight, jim-chen@northwestern.edu
- Feh Yeh, StarLight, fyeh@northwestern.edu
- Se-Young Yu, StarLight, youff3@gmail.com
- Paul Ruth, FABRIC, pruth@renci.org