Measuring InfiniBand Network Performance: Bandwidth and Latency

John Henry Mejia

March 2, 2024

1 Introduction

InfiniBand is an extensively used network communication standard, particularly favored for its high throughput and low latency in high-performance computing environments. To comprehend its performance characteristics, an experiment was conducted to measure the bandwidth and latency by transmitting messages of varying sizes and analyzing the response times.

2 Methodology

A program was developed in C to send messages between two nodes within an InfiniBand network. The message sizes were incremented in 16MB intervals, starting from 16MB up to 2.1GB. Each distinct size was sent ten times to allow for around 2000 data points for a statistically significant analysis.

These transmissions were orchestrated using a SLURM job scheduler to enable the processes on two nodes to communicate through the InfiniBand network. We measured the time elapsed to send and receive each message (loopback time). By halving that loopback time, we were able to find the average one-way communication time.

3 Results

Time needed to transmit a message highly correlates with the message size:

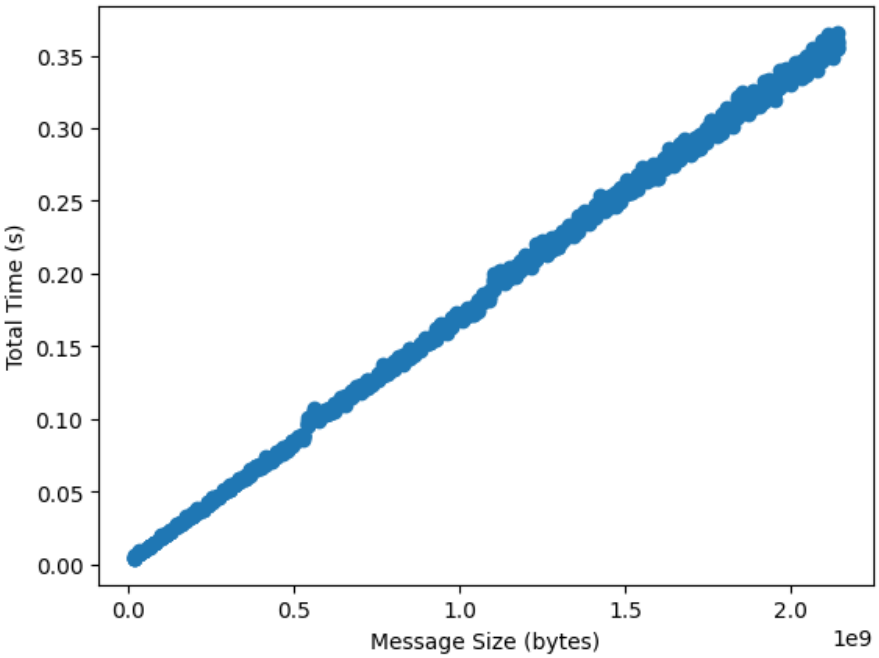

Figure 1: Graph of Time to Send Message Versus Message Size

The slope of the trendline in the graph represents the network bandwidth, and the y-intercept accounts for the latency. These parameters were estimated as follows:

- Bandwidth: 5939141600 bytes/s (5.94 GB/s)
- Latency: 0.00231 s (2.31ms)
- R^2 value: 0.9989

The R^2 value, being close to 1, indicates that the size of a message is a good predictor of the time to send it via the network.

4 Discussion

The observed bandwidth aligns particularly well with the theoretical capabilities of InfiniBand technology, which offers a maximum bandwidth of 56 Gbits/s (7000 MB/s). Taking into account the conversion between bits and bytes, and overhead from encoding and other factors, the measured bandwidth demonstrates significant utilization of the network's potential.

However, in terms of latency, the observed value of 2.31 ms is above the sub-millisecond latency typically associated with InfiniBand networks. InfiniBand networks are known for their sub-millisecond latency and this is surprising given the program. This discrepancy may be attributed to the additional latency introduced by the network configuration, software overhead, or experimental setup. Nonetheless, the latency is within a reasonable range for applications demanding high performance.

5 Conclusion

The experiment successfully quantified the network performance of InfiniBand, demonstrating its potent bandwidth and minimal latency capabilities. The findings align well with the standard's specifications, corroborating the technology's high reputation in the realm of efficient data networking.

6 Materials

The following links provide access to the original source code, the SLURM job configuration, the raw data, and the scripts used to calculate bandwidth and latency from the collected data.

- C Program for Message Transmission: `https://github.com/jhmejia/hppc-hw3/blob/main/pingpong.c`
- Bandwidth and Latency Calculation Script: `https://github.com/jhmejia/hppc-hw3/blob/main/pingpong.ipynb`
- SLURM job script: `https://github.com/jhmejia/hppc-hw3/blob/main/pingpong.slurm`
- Raw Data: `https://github.com/jhmejia/hppc-hw3/blob/main/pingpong_raw_data.csv`