RDMA Microbenchmark

The RDMA Microbenchmark was written to allow a client machine to test the maximum sustainable throughput of a DUT (which could be a BlueField-2 DPU or any other host). The client’s window size and number of threads can be varied to change the amount of load generated by the client.

This RDMA microbenchmark was created using UD QPs. To perform two tasks:

1. Generate load the test the packet processing limits of a DUT.
2. Measure response latency.

Each client/server thread is pinned to a different core and a unique queue pair (QP), which has a unique QP number.

The client knows how to connect to the server as it is provided the server’s IP address, starting QP number, and the number of server threads.

More Information on how to run the microbenchmark (with the two different modes of operation), along with all possible command line arguments is included in the repo’s README.

The microbenchmark is written with native low-level RDMA Verbs and uses the -libverbs library to compile. Please see the Makefile for more information on how to compile. UD QPs are used to enable scalability to communicate across QPs. With UD QPs, each QP on the client can send or recv messages from any QP on the server.

For any questions: please feel free to contact me at hamed@gatech.edu.