

Infiniband Testing Updates

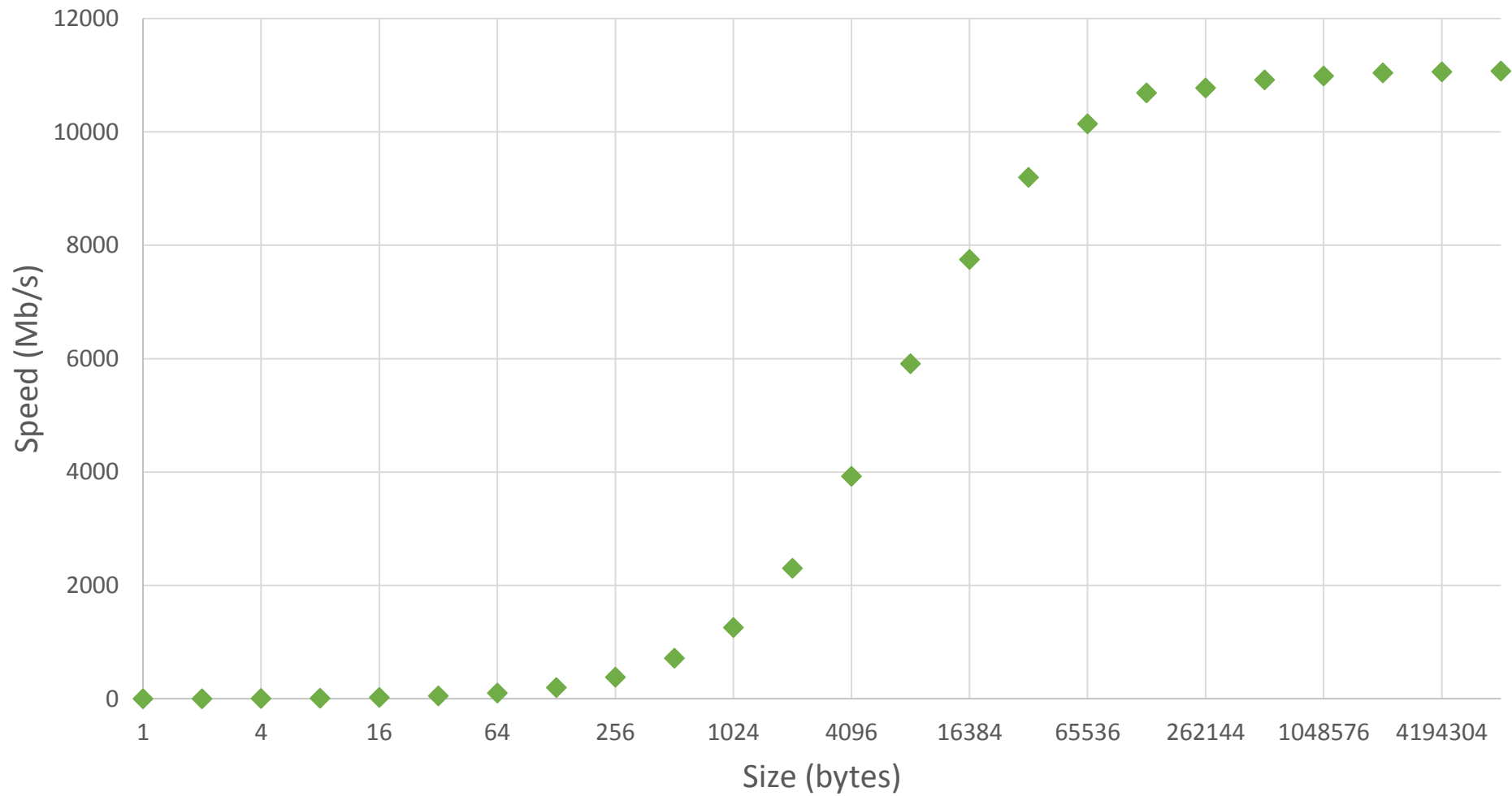
7/15/2014 – Hanna Schamis

Introduction

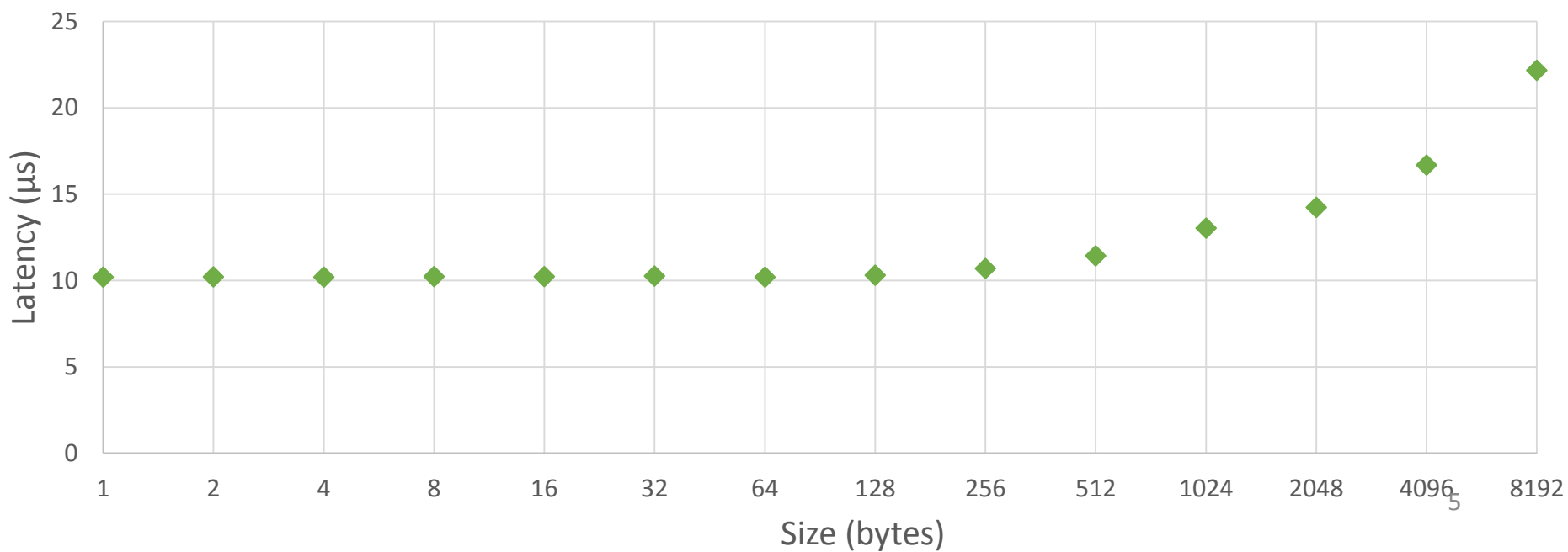
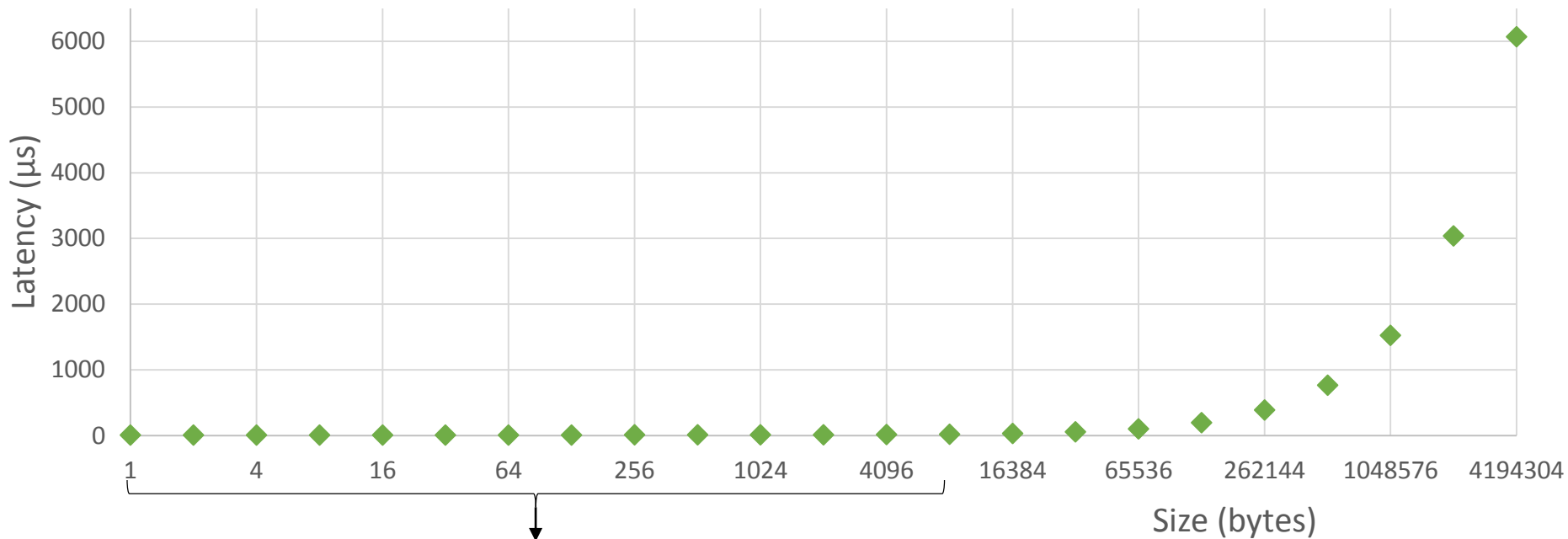
- Testing using OSU Micro-Benchmarks and Intel MPI Benchmarks
 - These are tests that use MPI through Infiniband to test the connection speed and MPI functions.
- Tested the following functions: Gather, Scatter, Broadcast, allgather, and alltoall. These the ones that I think will be the most useful for what we want to do.
 - www.cac.cornell.edu/Ranger/MPIcc/datamovement.aspx

Libibverbs tests

Libibverbs – Bandwidth

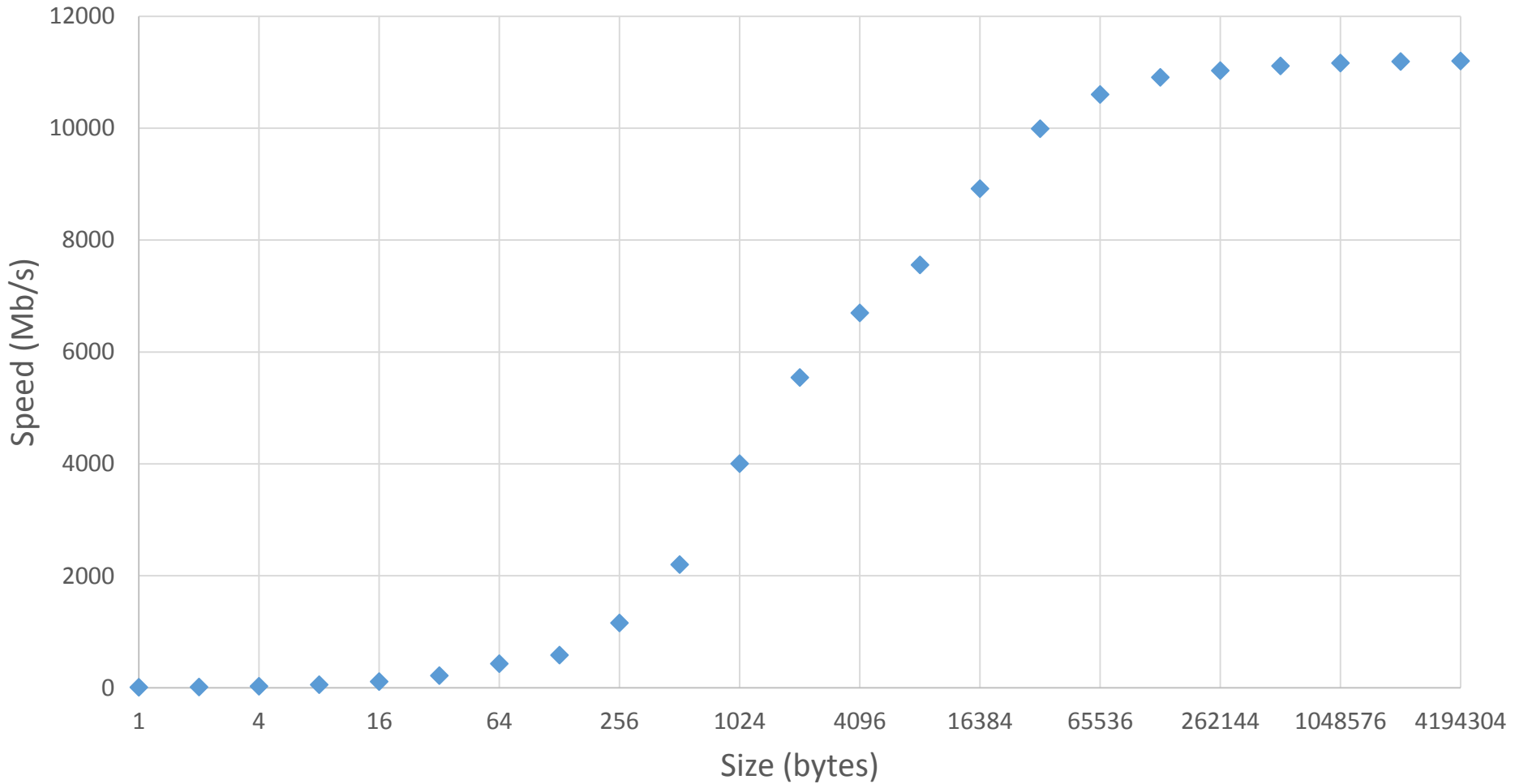


Libibverbs – Latency



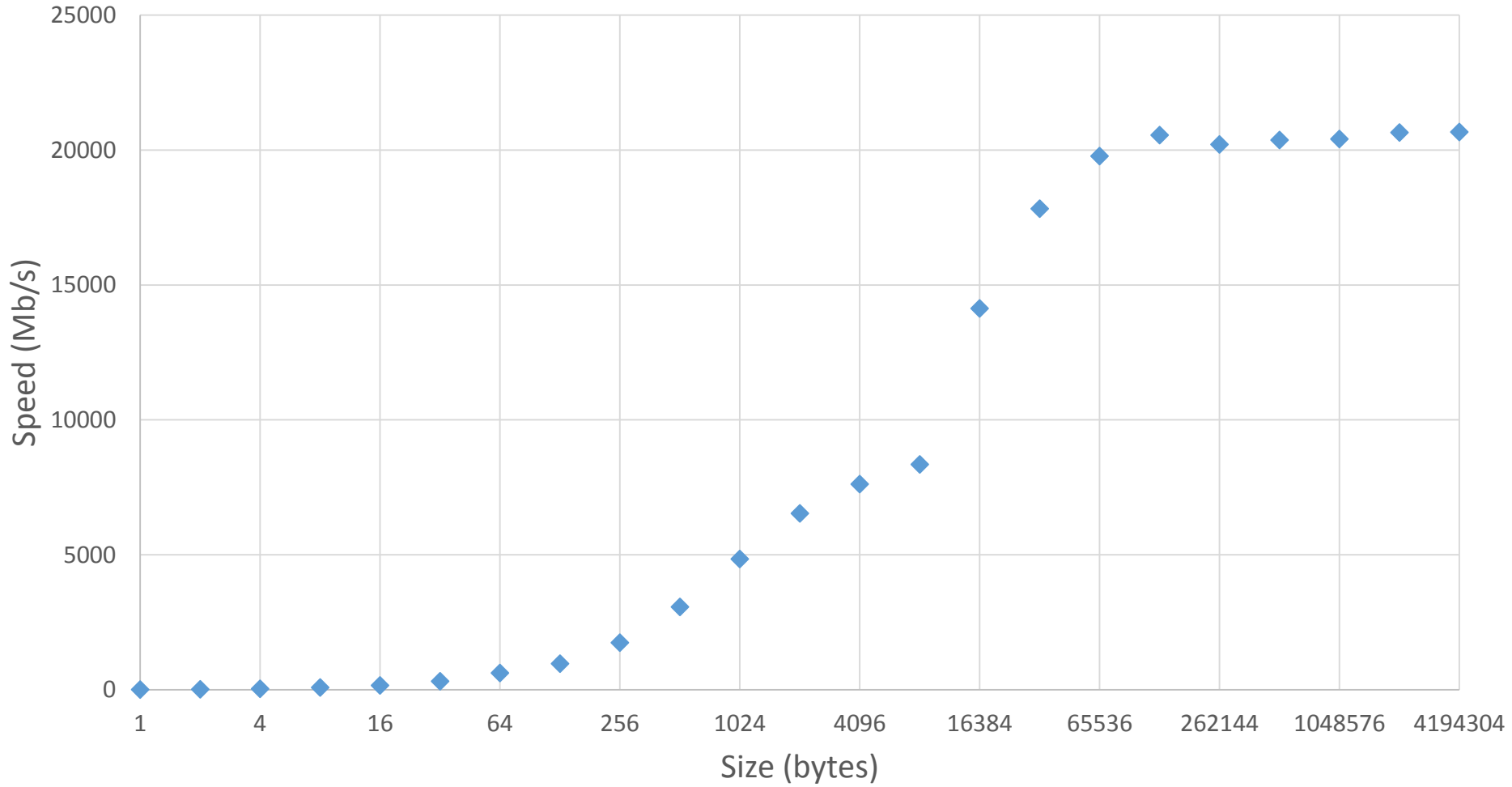
OSU Micro-Benchmarks

Unidirectional Bandwidth



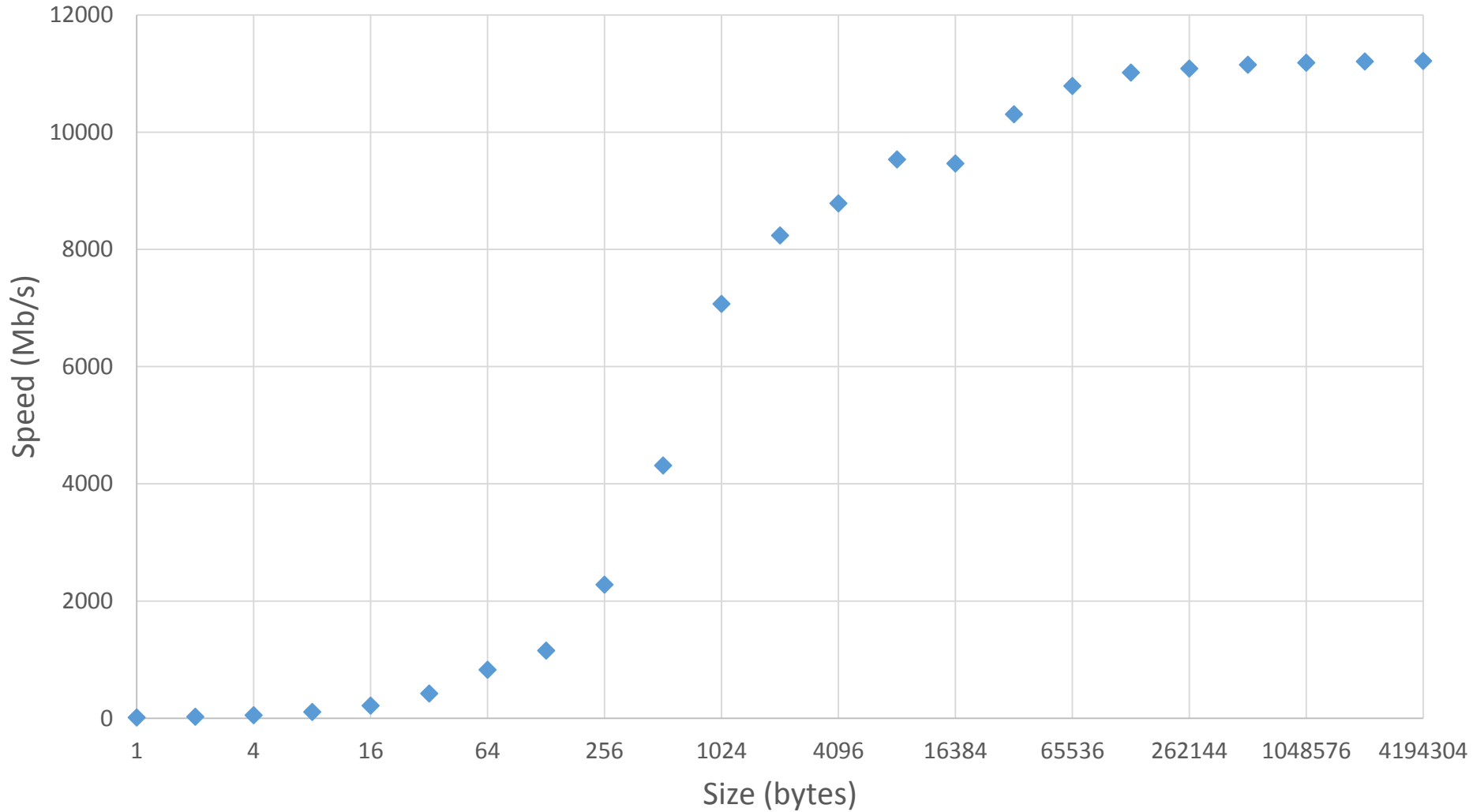
```
$ mpirun --allow-run-as-root --mca btl self,openib -np 2 -host banjo13,banjo14 ./osu_bw_7
```

Bidirectional Combined Bandwidth



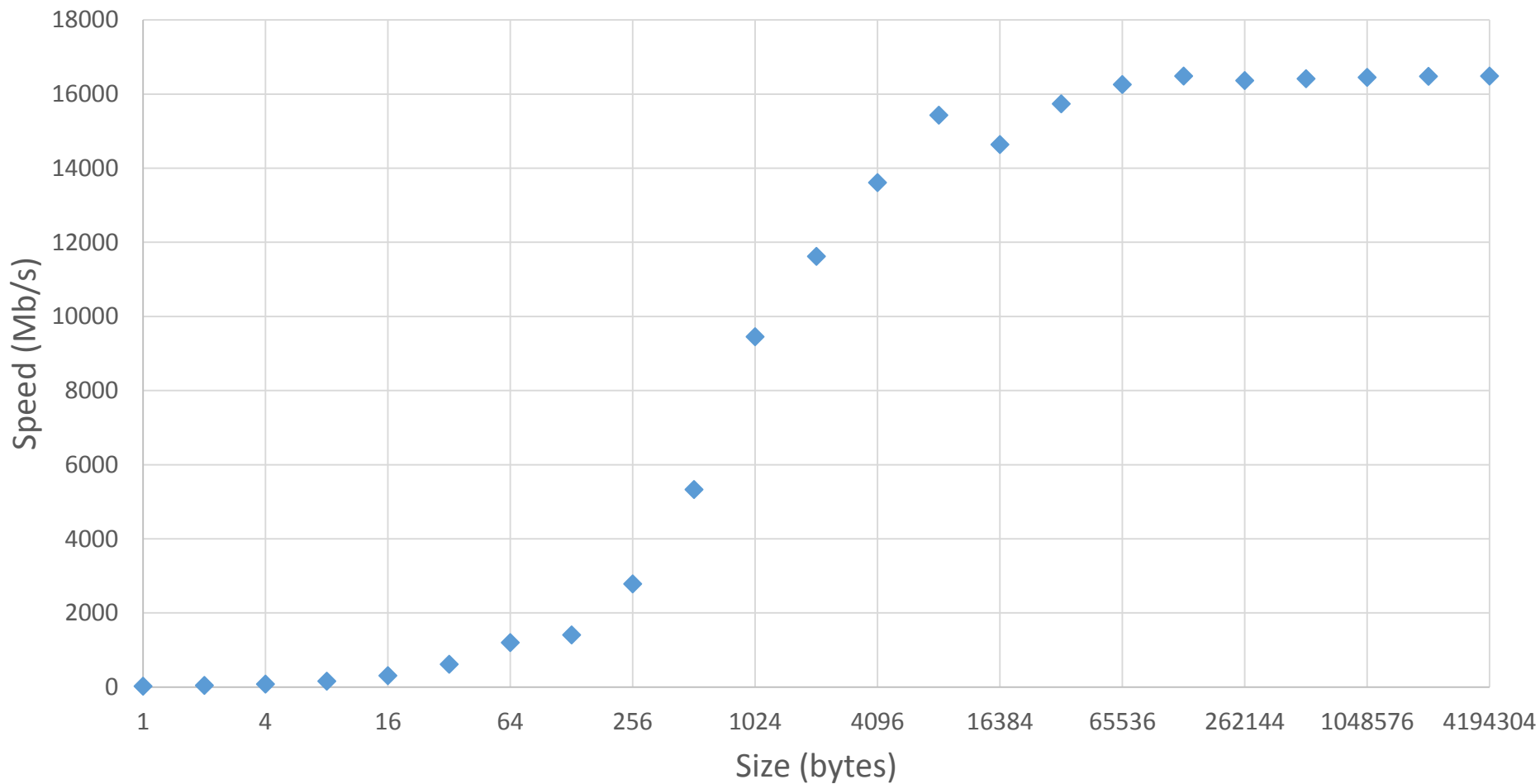
```
$ mpirun --allow-run-as-root --mca btl self,openib -np 2 -host banjo13,banjo14 ./osu_biw
```


Total Bandwidth of 2 connections between 3 nodes

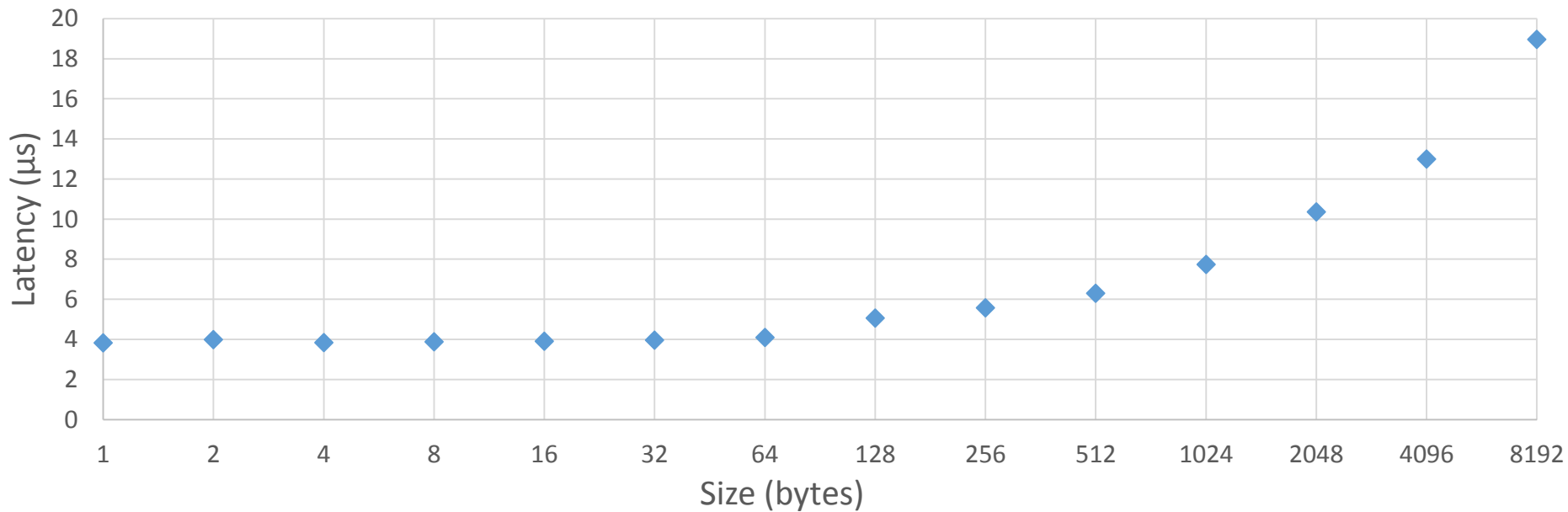
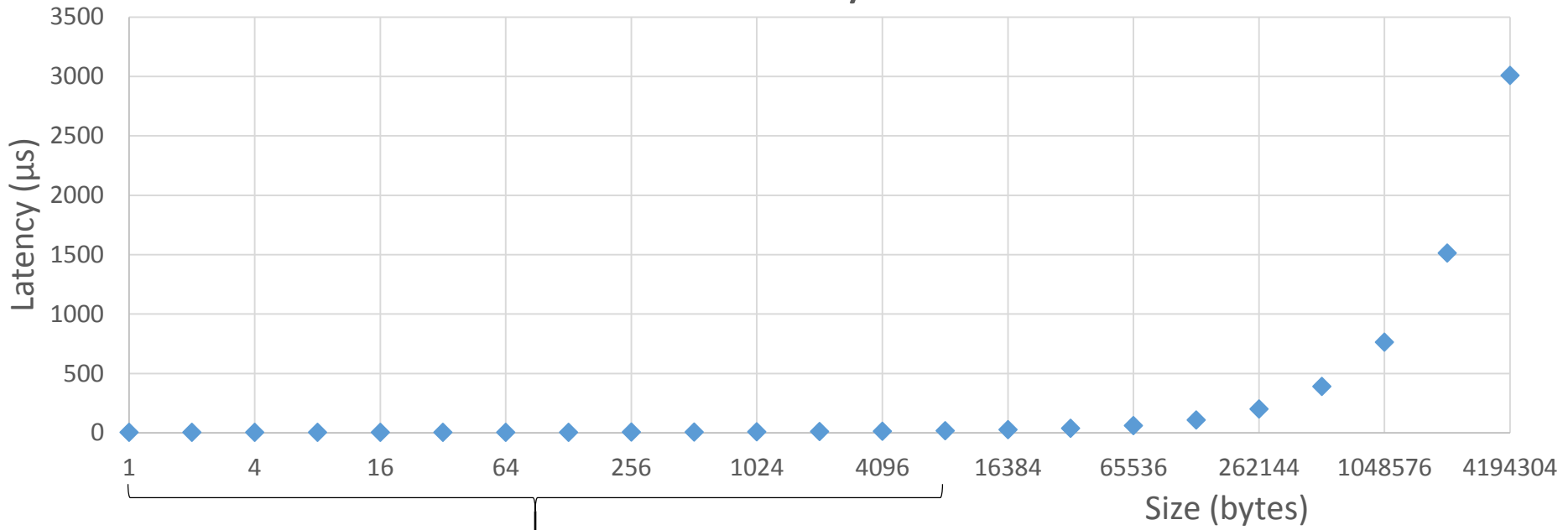


\$ mpirun --allow-run-as-root --mca btl self,openib -np 4 -host banjo13,banjo14,banjo15 ./osu_mbw_mr

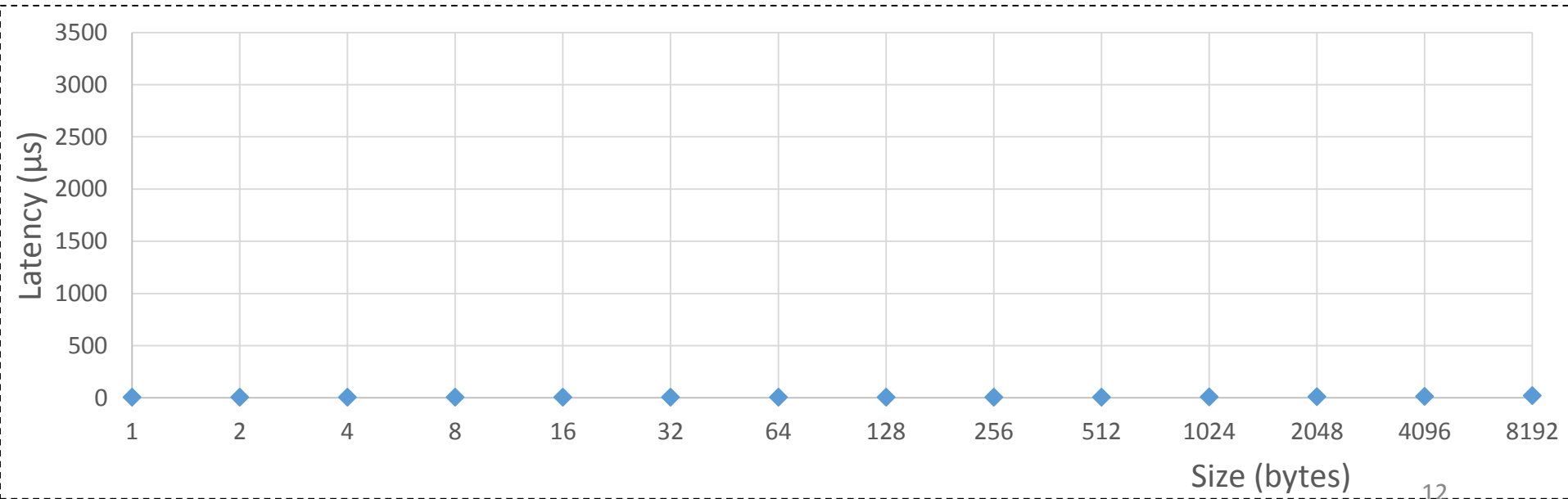
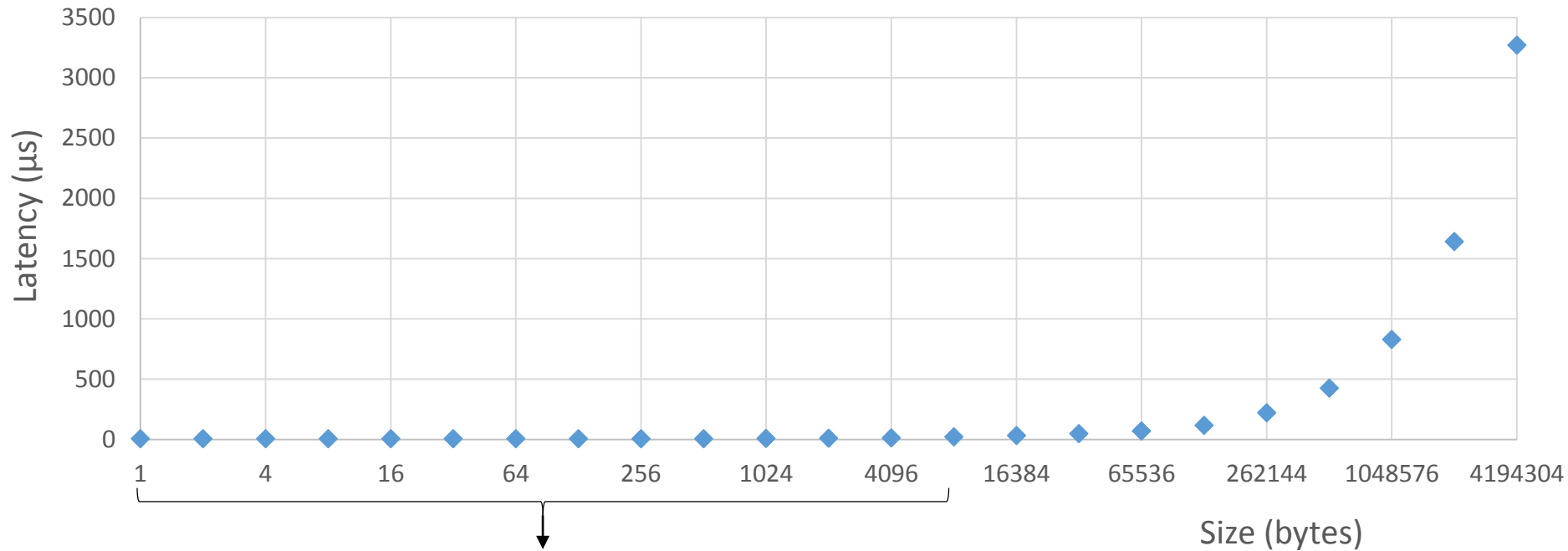
Total Bandwidth of 3 connections between 3 nodes



Latency

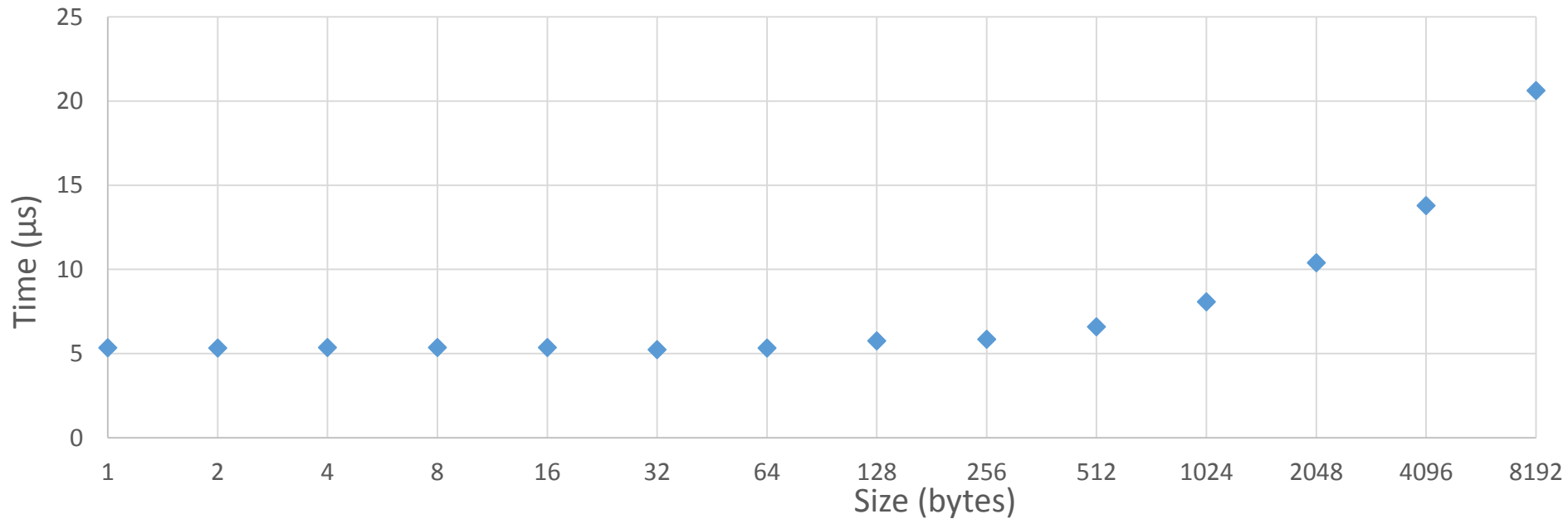
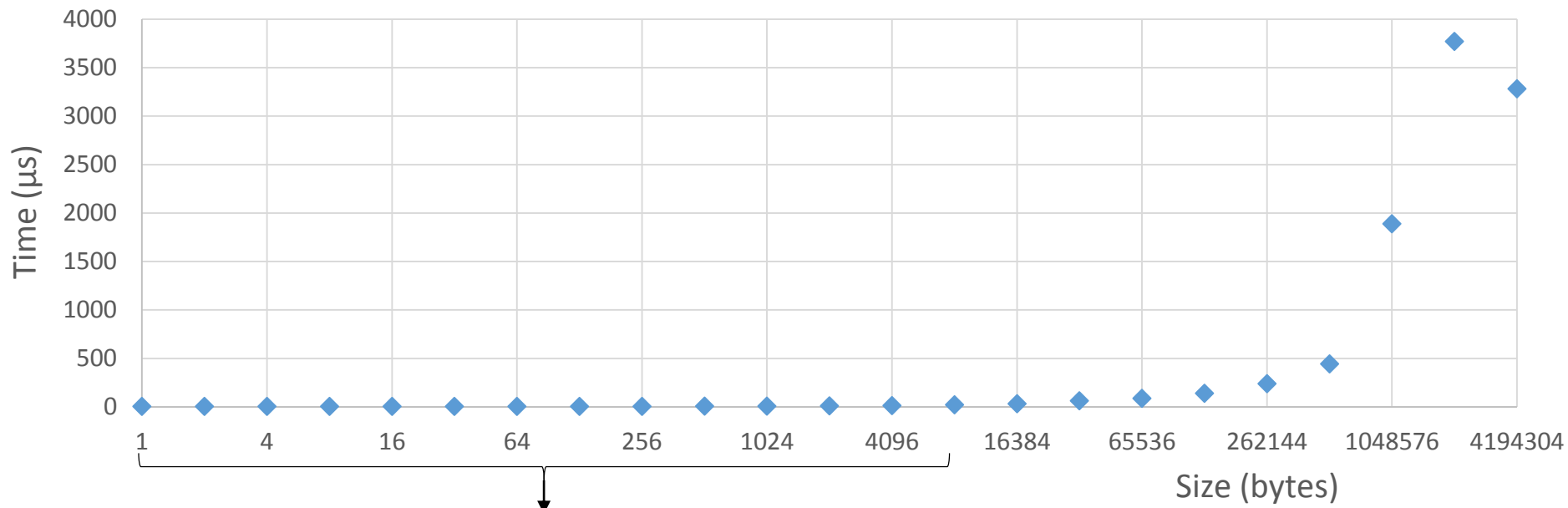


Latency – 2 simultaneous connections



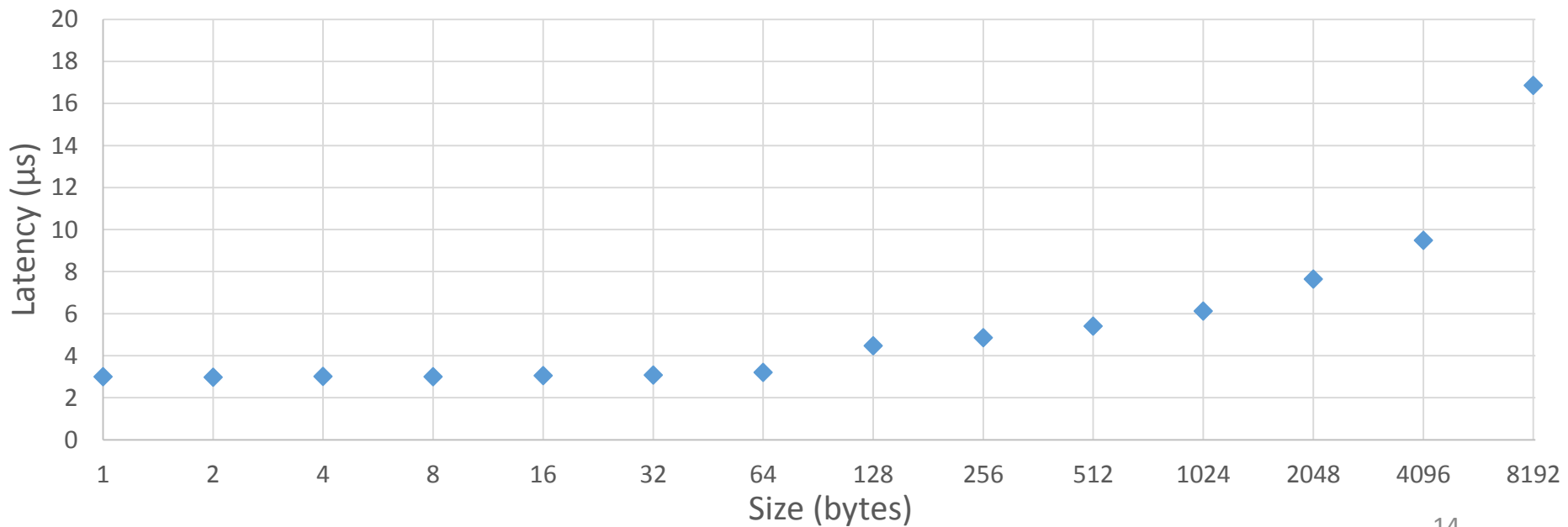
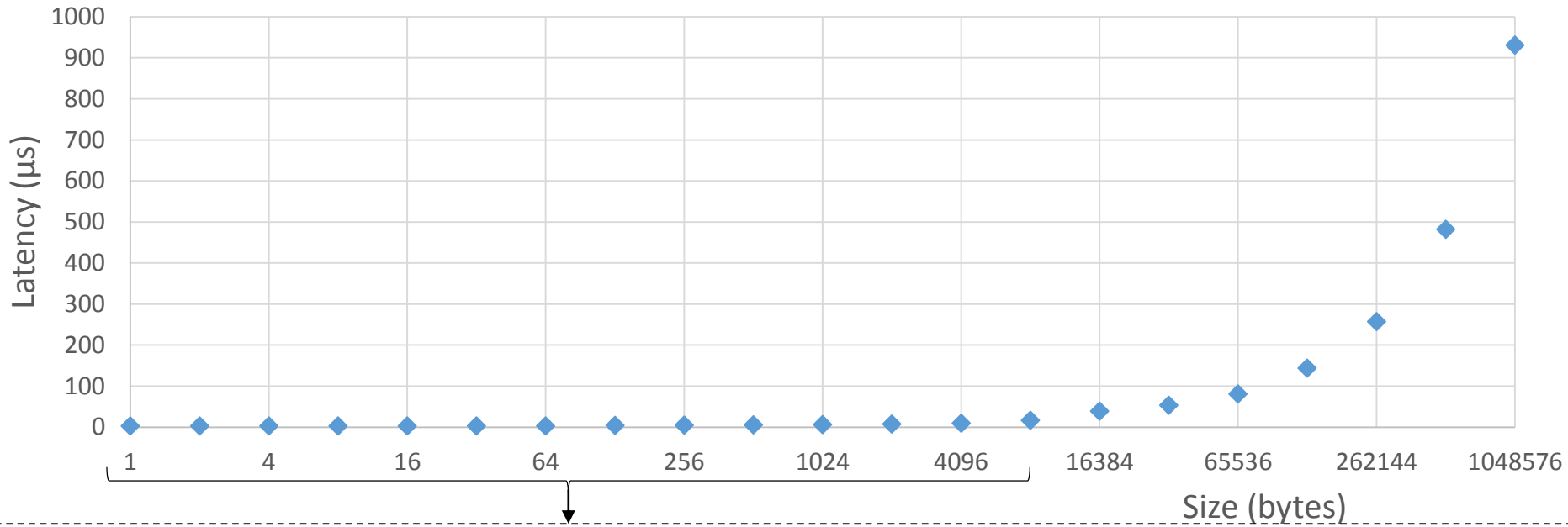
```
$ mpirun --mca btl self,openib -np 4 -host banjo13,banjo14,banjo15 ./osu_multi_lat
```

Latency – 3 simultaneous connections



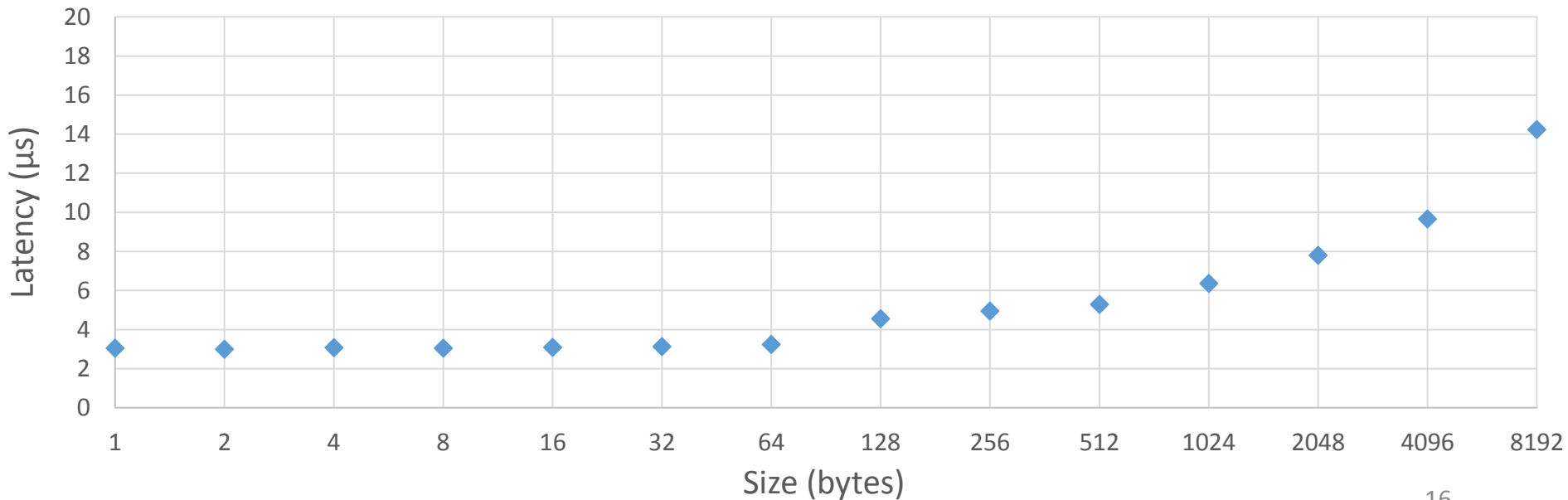
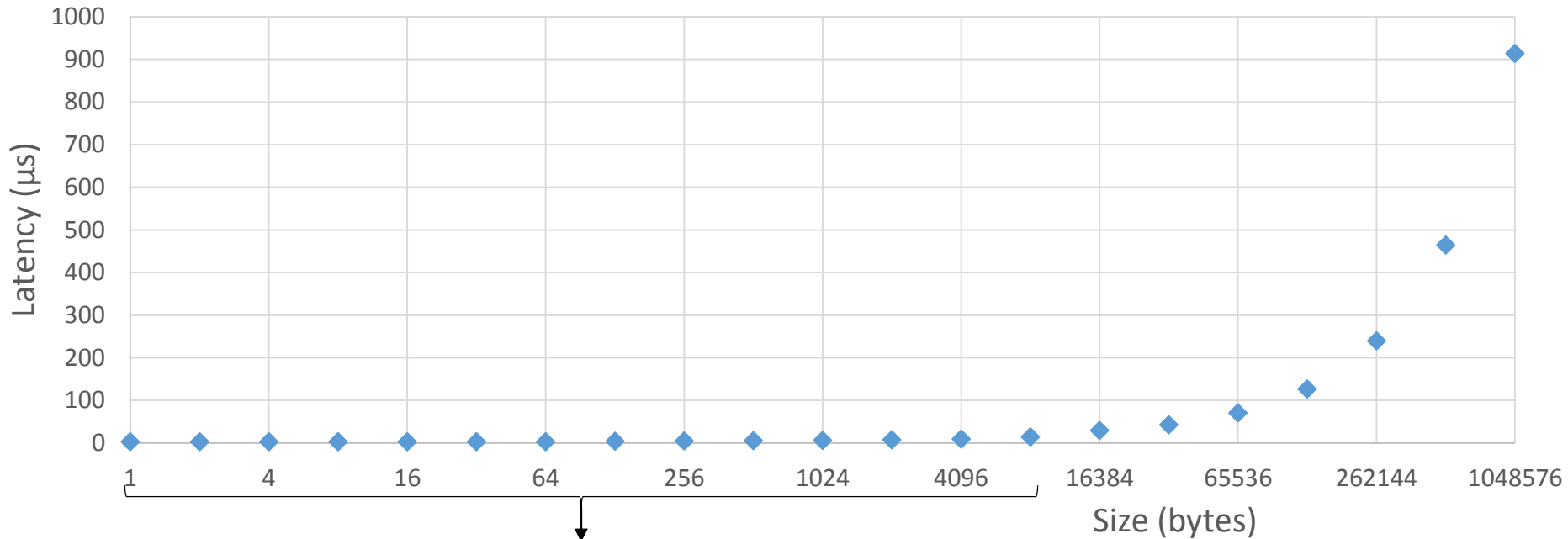
\$ mpirun --mca btl self,openib -np 6 -host banjo13,banjo14,banjo15 ./osu_multi_lat

Gather latency



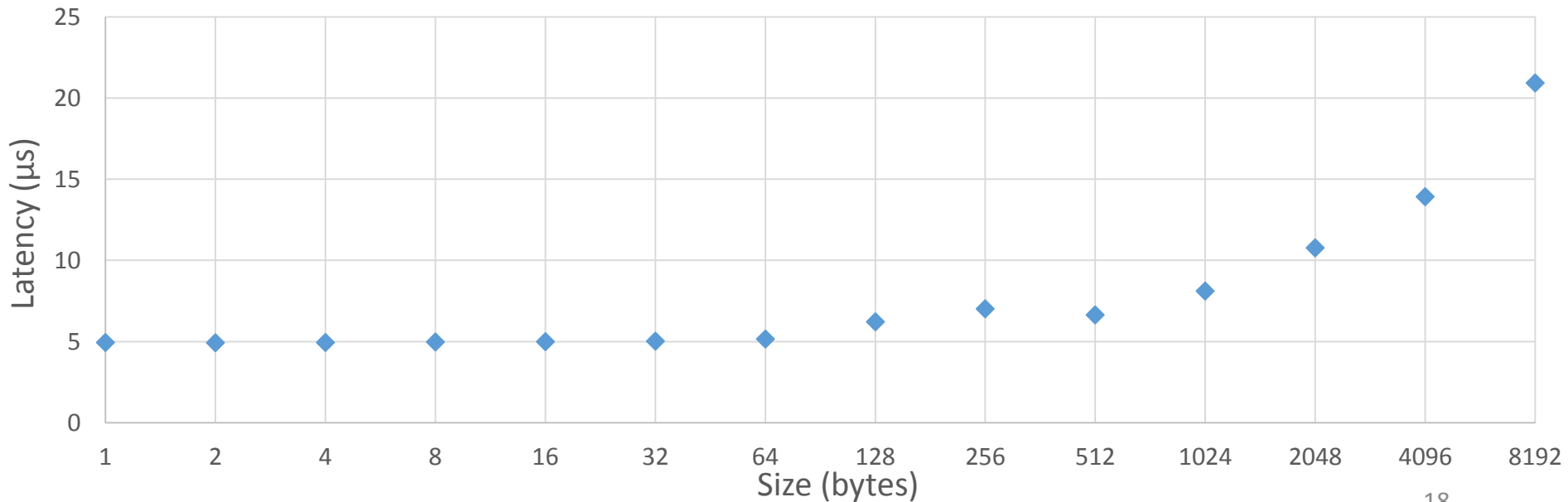
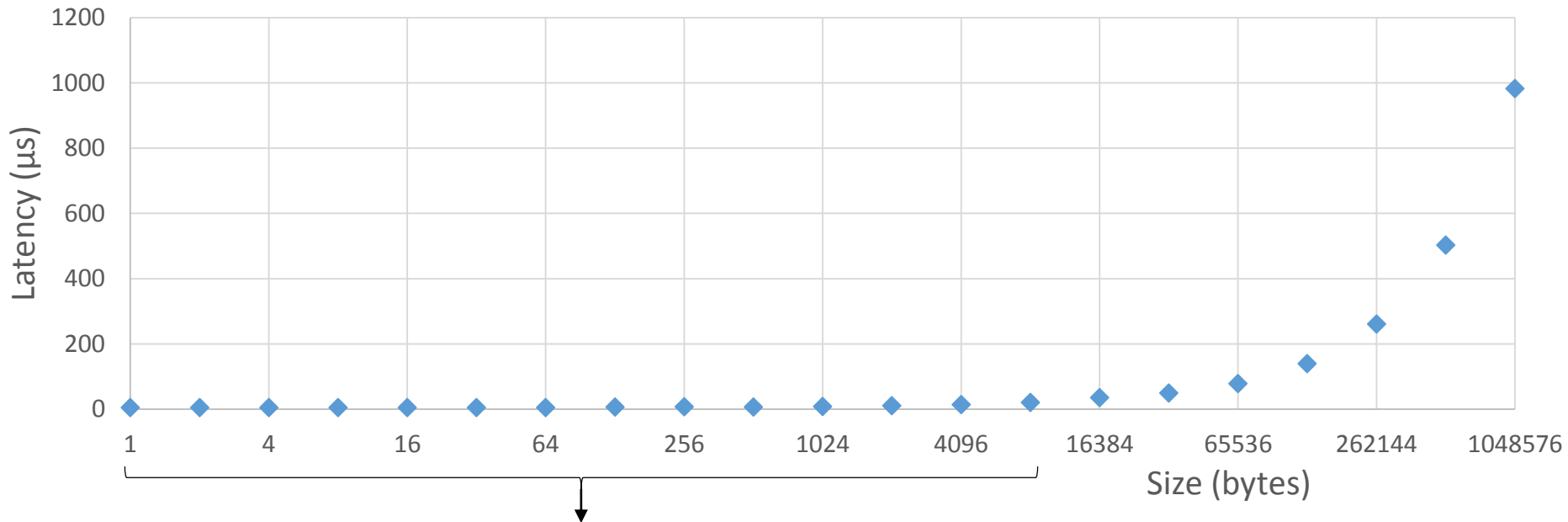
\$ mpirun --allow-run-as-root --mca btl self,openib -np 2 -host banjo13,banjo14 ./osu_gather

Scatter Latency



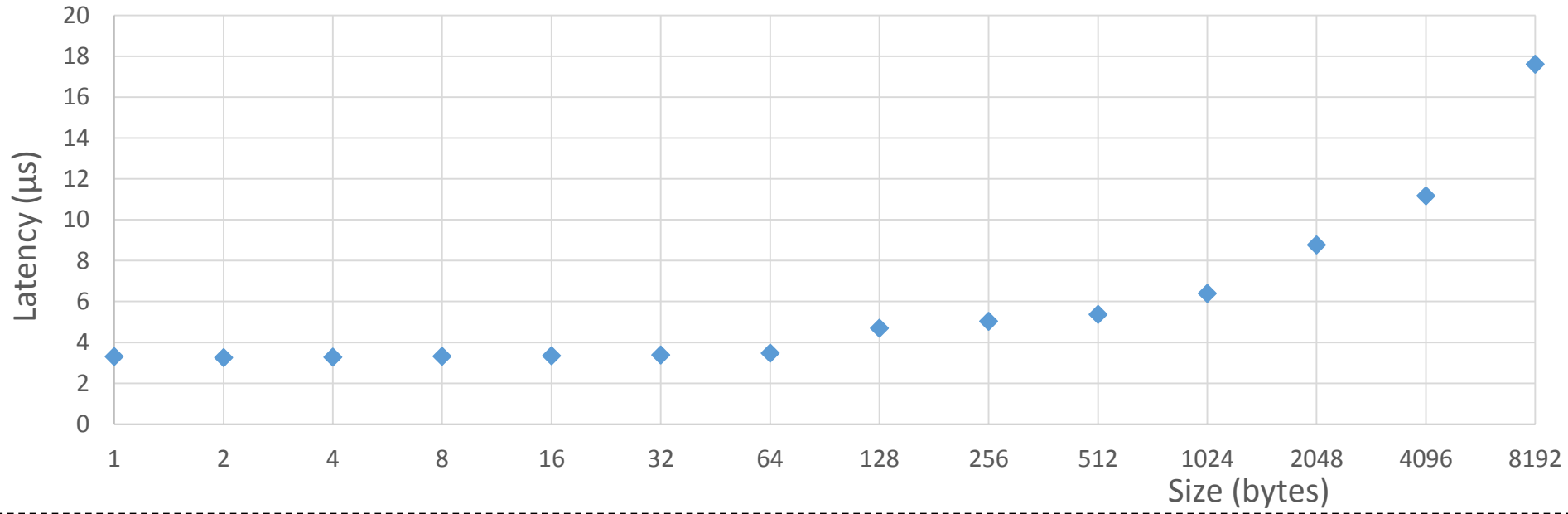
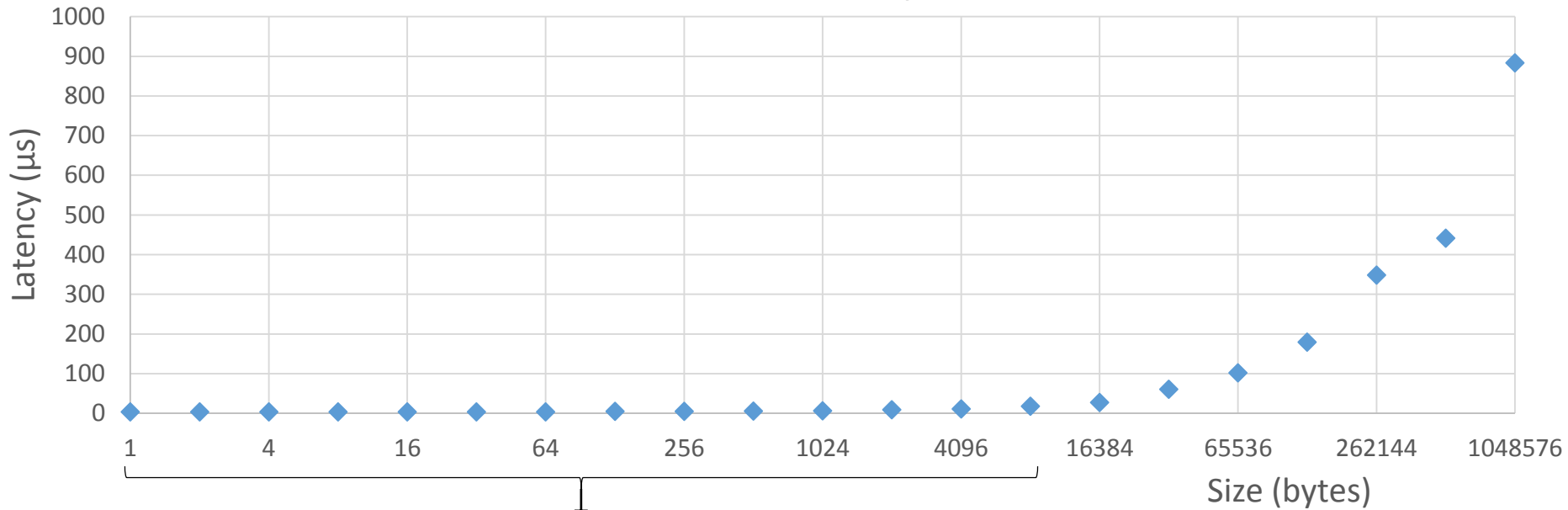
\$ mpirun --allow-run-as-root -np 2 --mca btl self,openib -host banjolo13,banjolo14 ./osu_scatter

Allgather latency

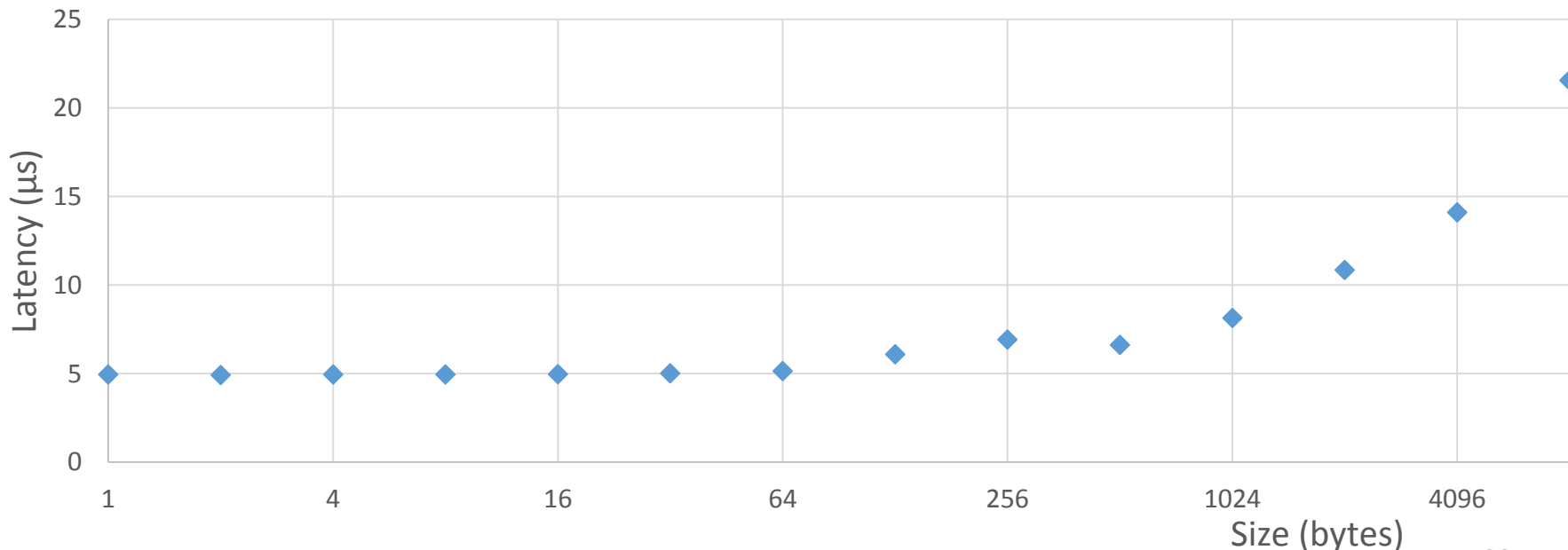
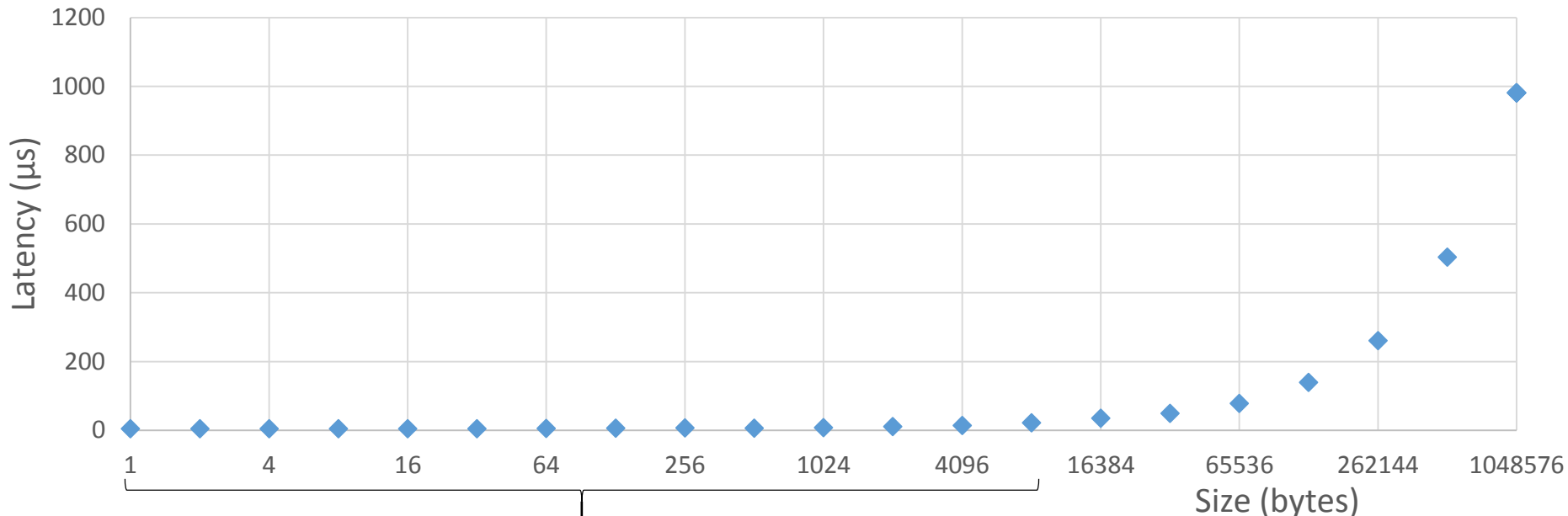


\$ mpirun --allow-run-as-root -np 2 --mca btl self,openib -host banjo13,banjo14 ./osu_allgather

Bcast latency



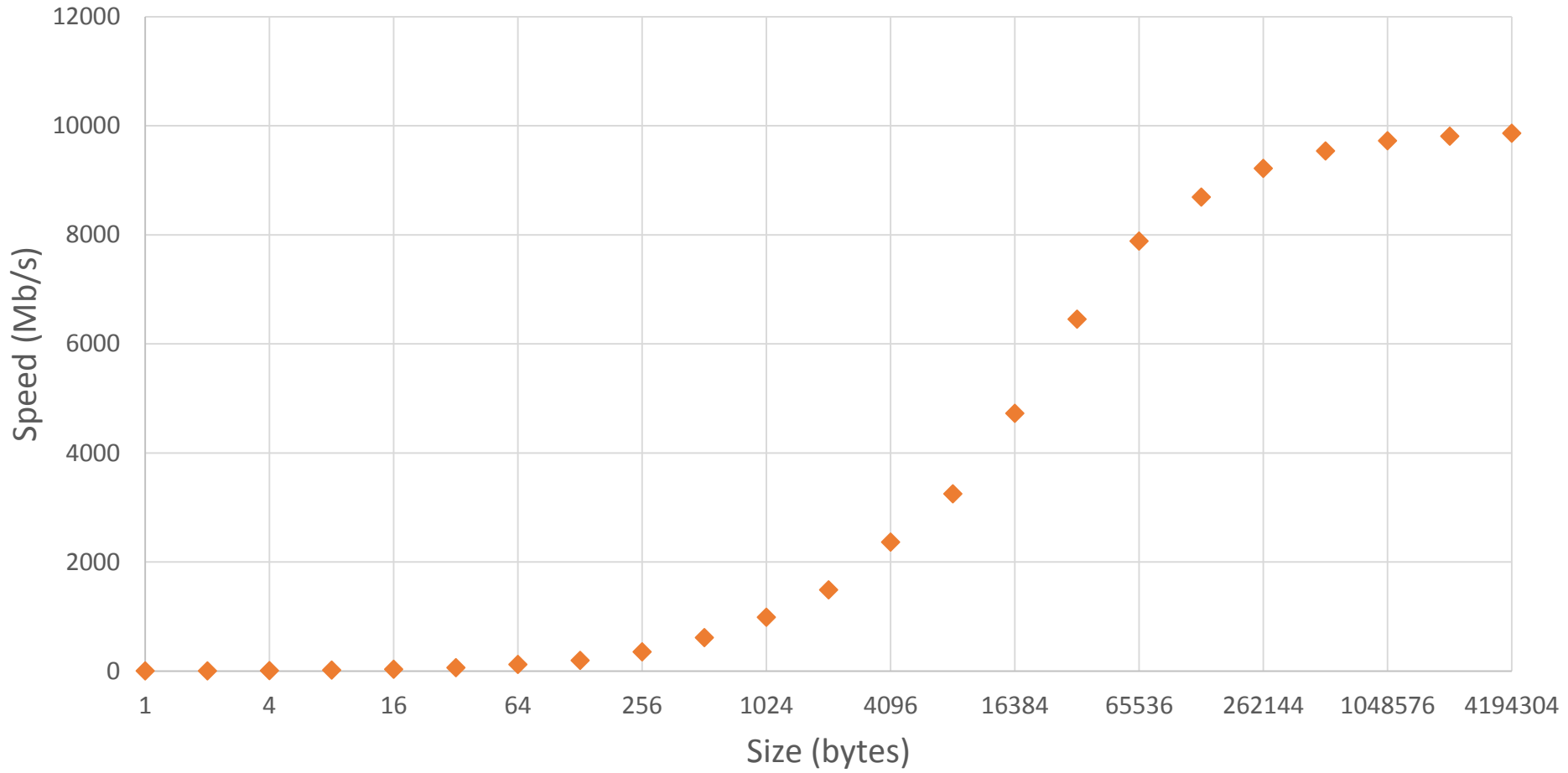
Alltoall latency



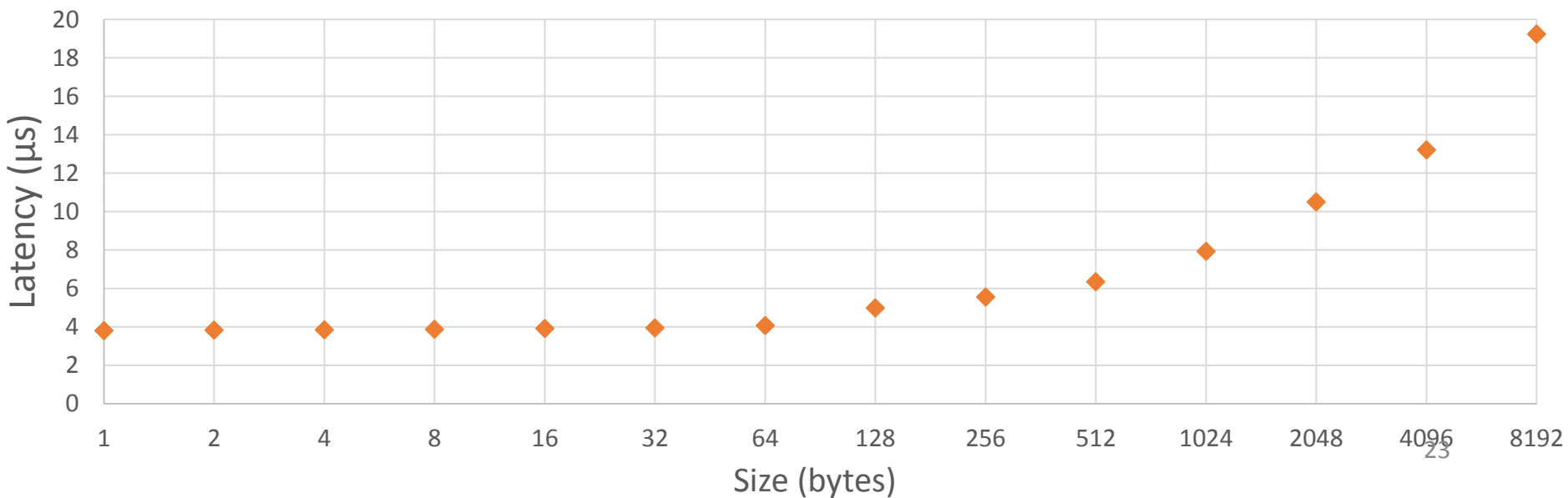
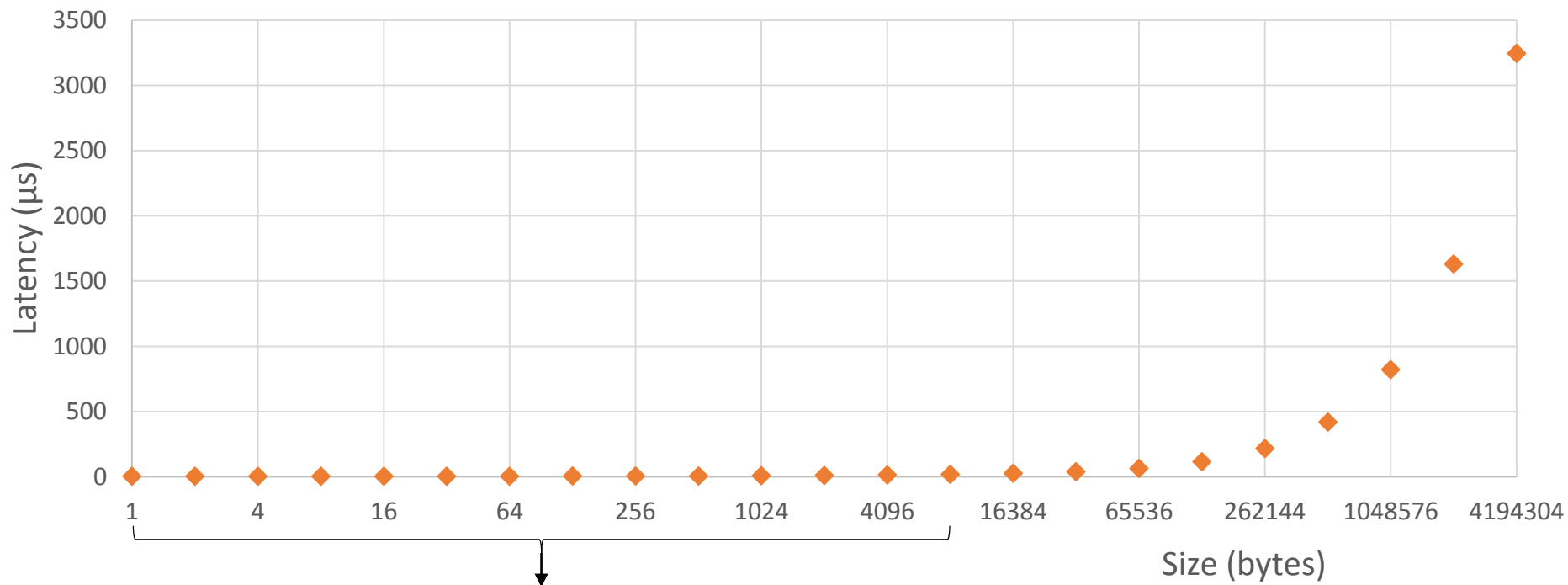
```
$ mpirun --allow-run-as-root -np 2 --mca btl self,openib -host banjo13,banjo14 ./osu_alltoall
```

Intel MPI Benchmarks

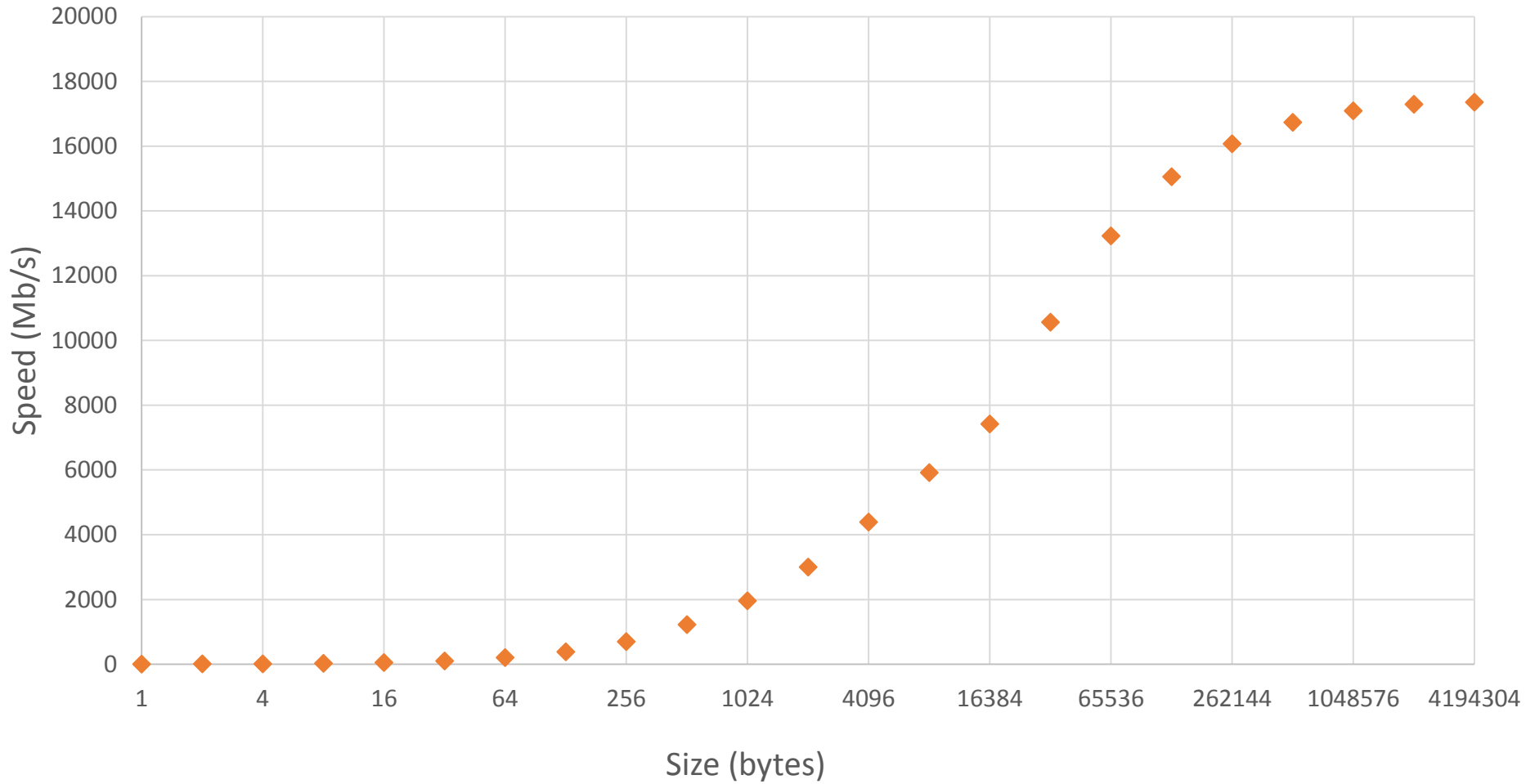
Pingpong – Bandwidth



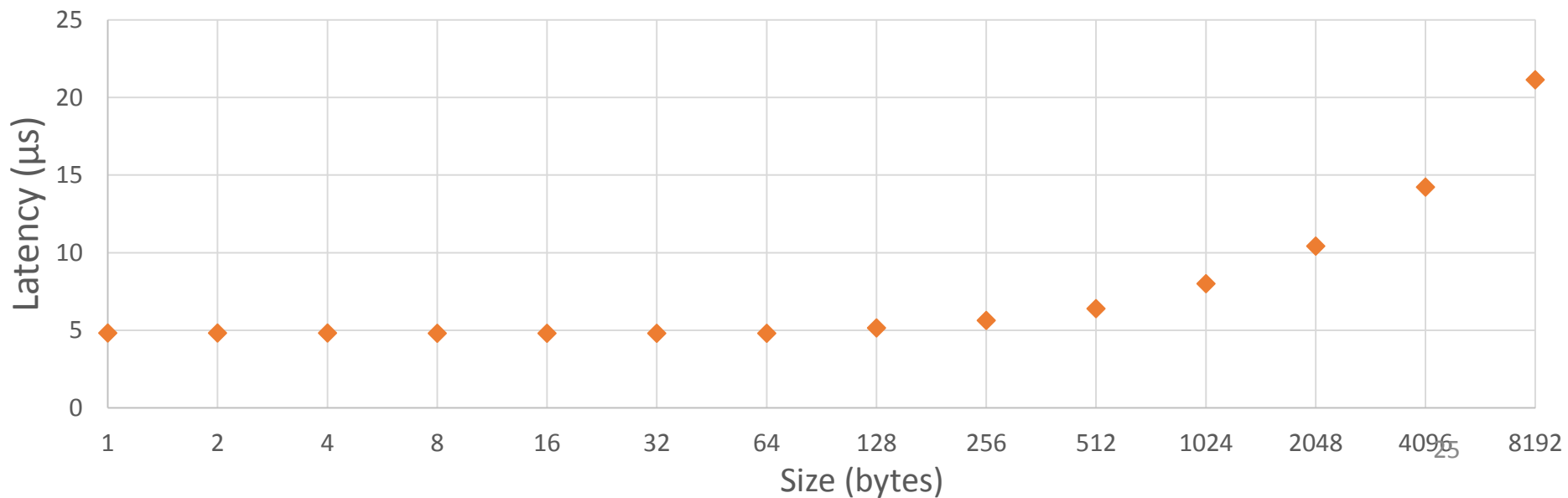
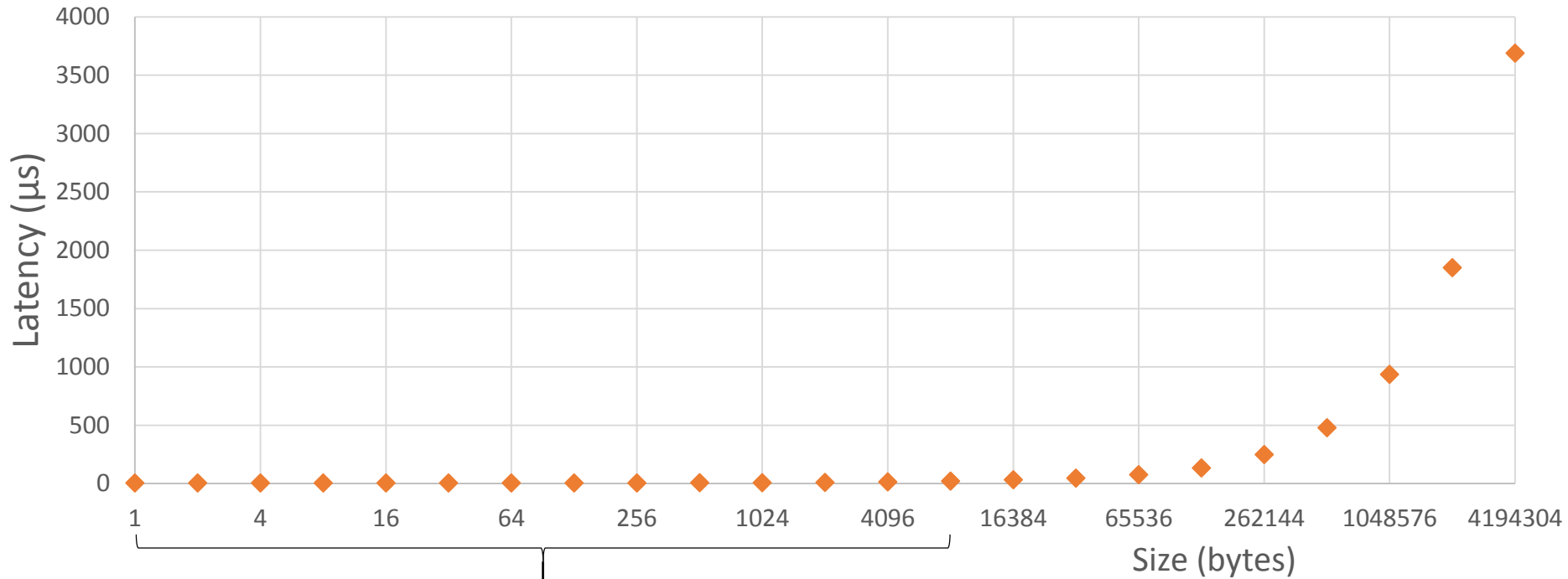
Pingpong – Latency



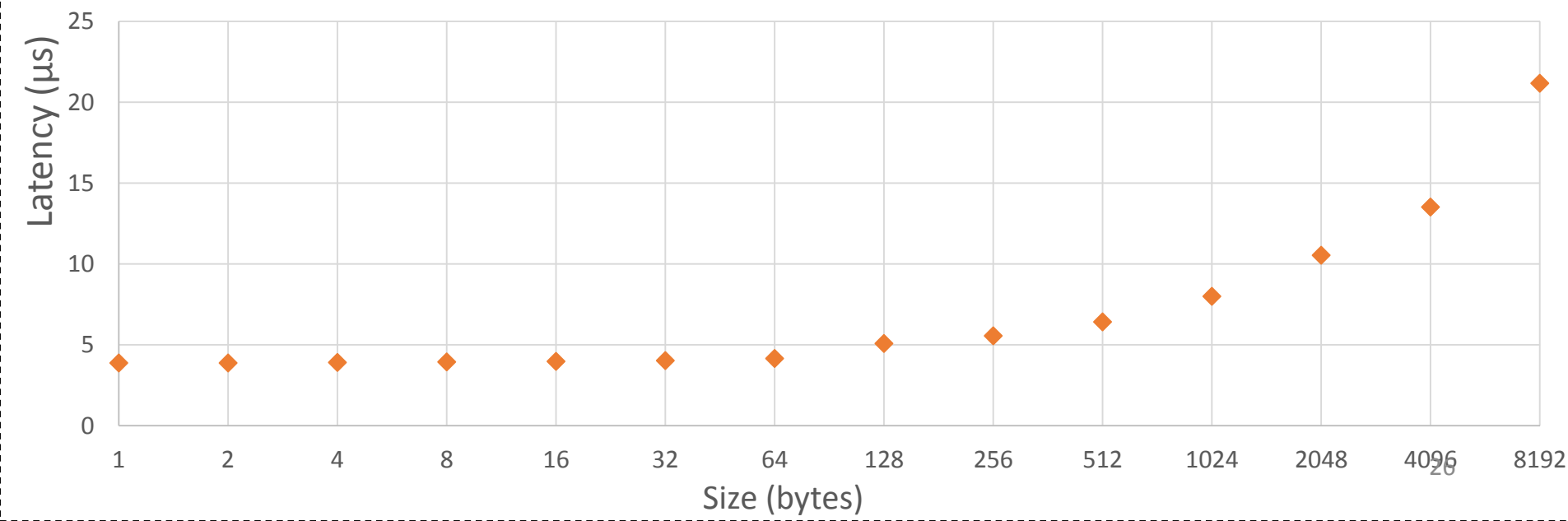
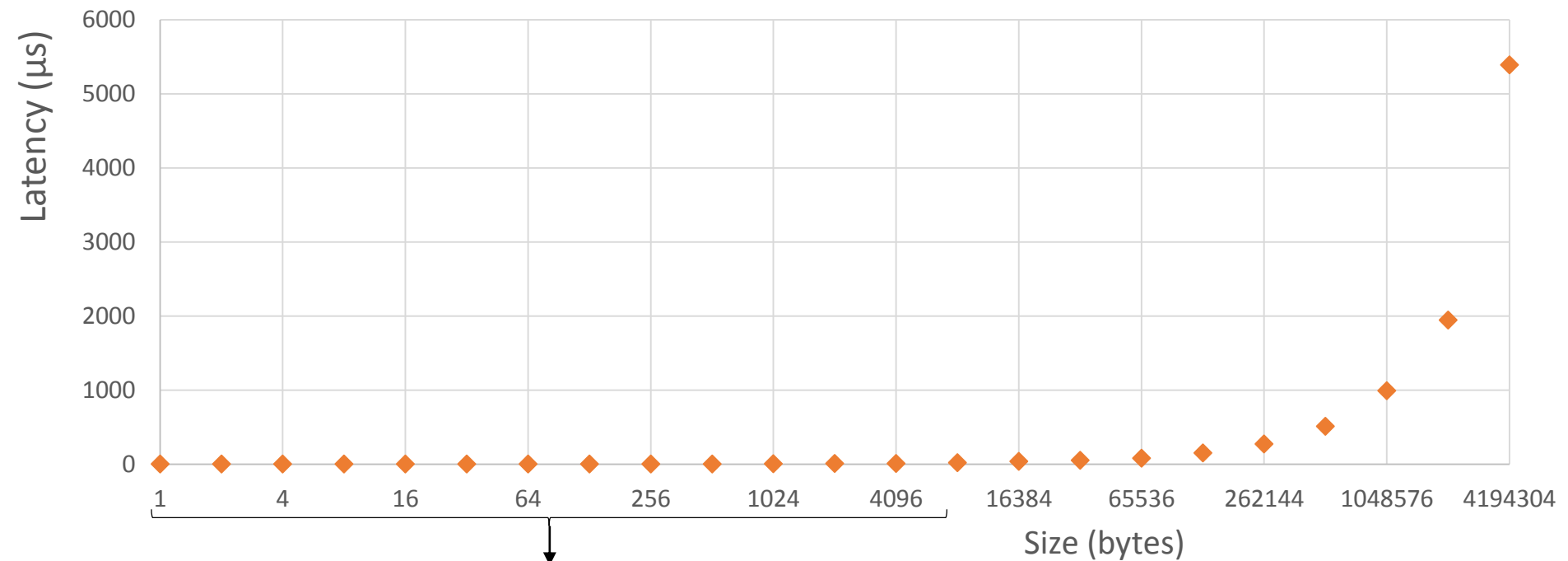
Sendrecv test – Bidirectional Bandwidth



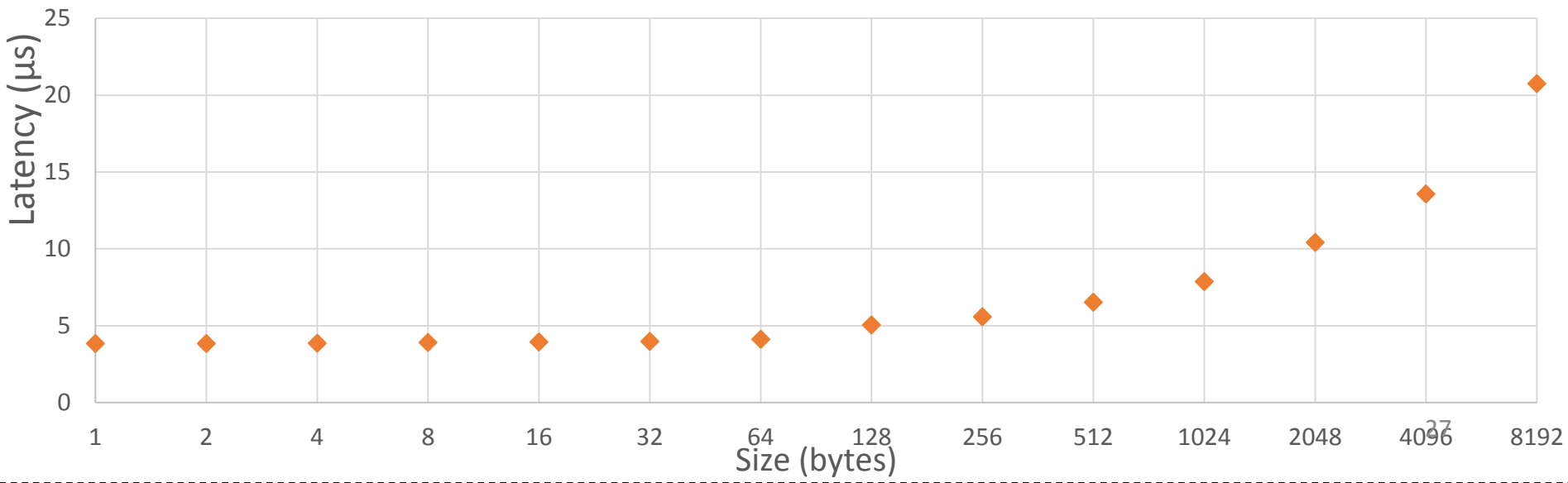
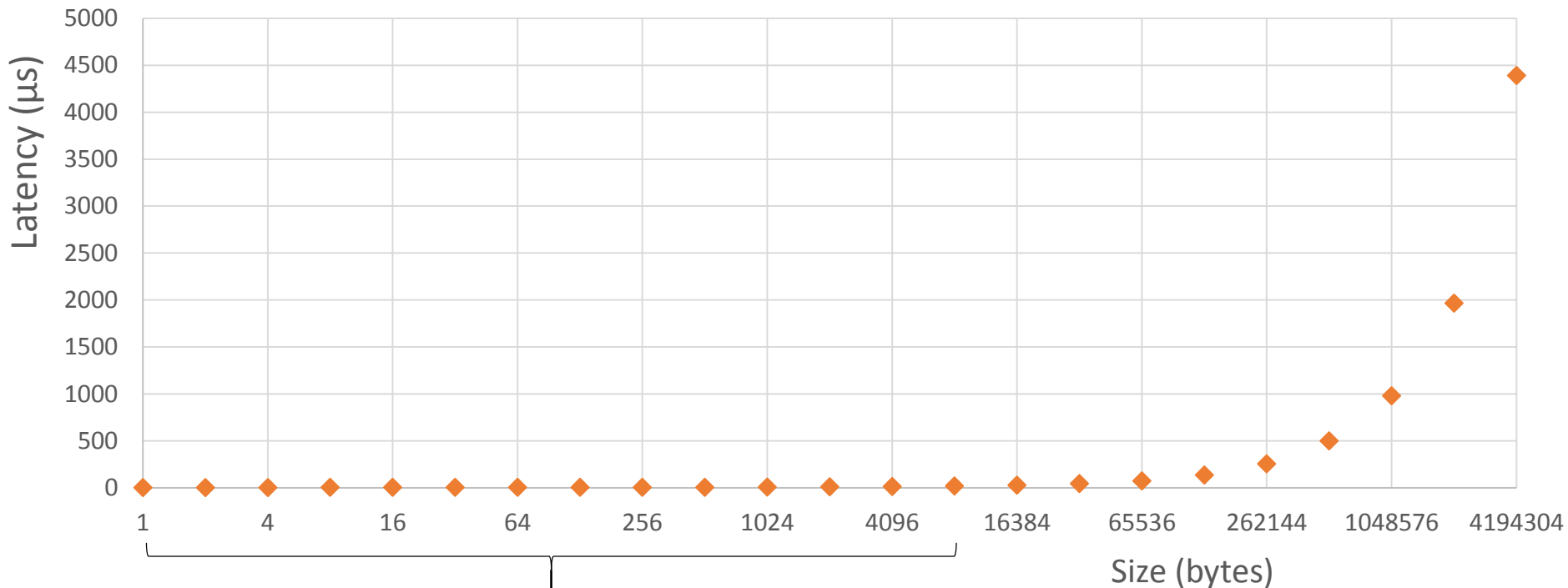
Sendrecv test – latency



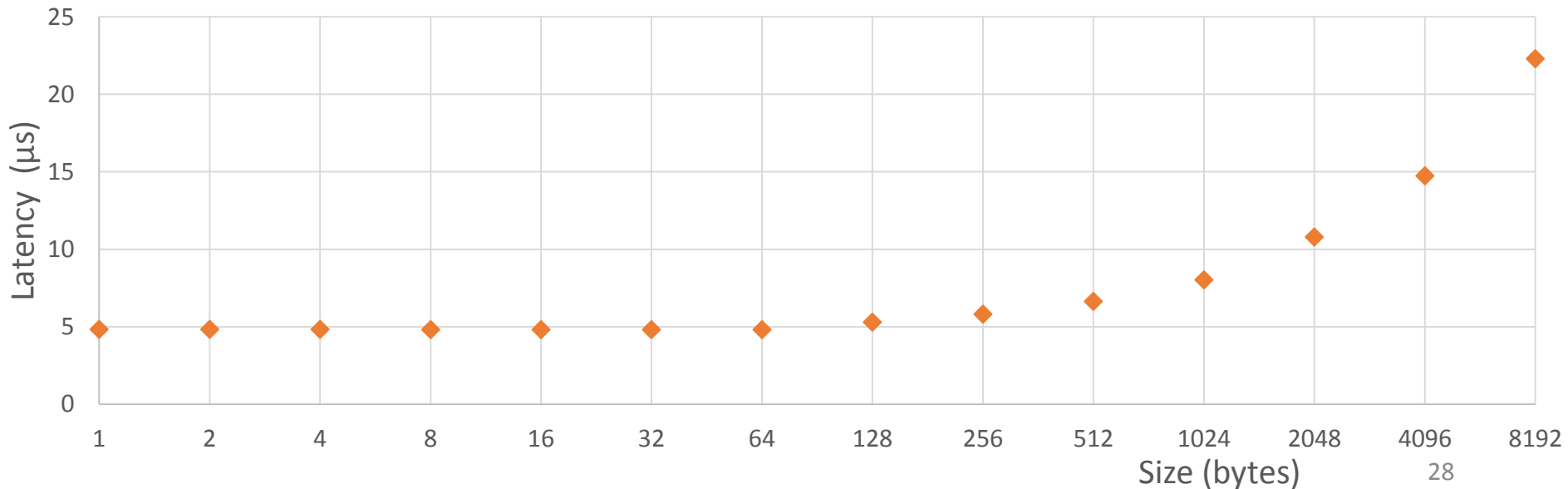
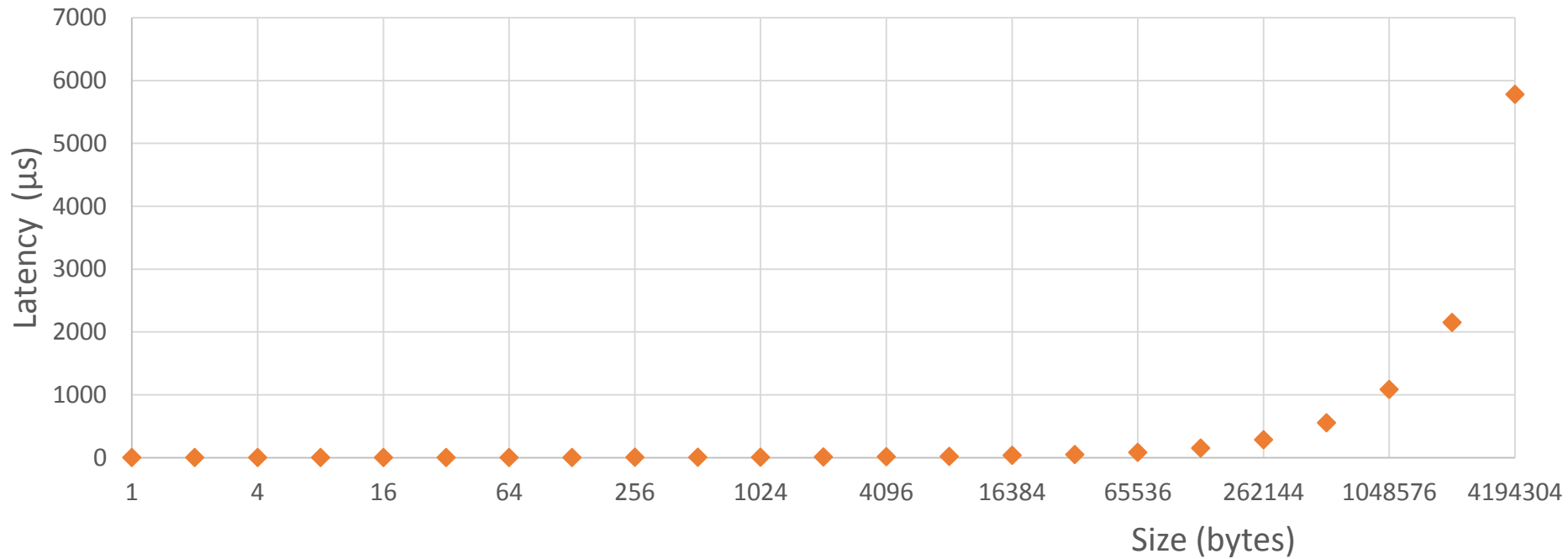
Gather latency



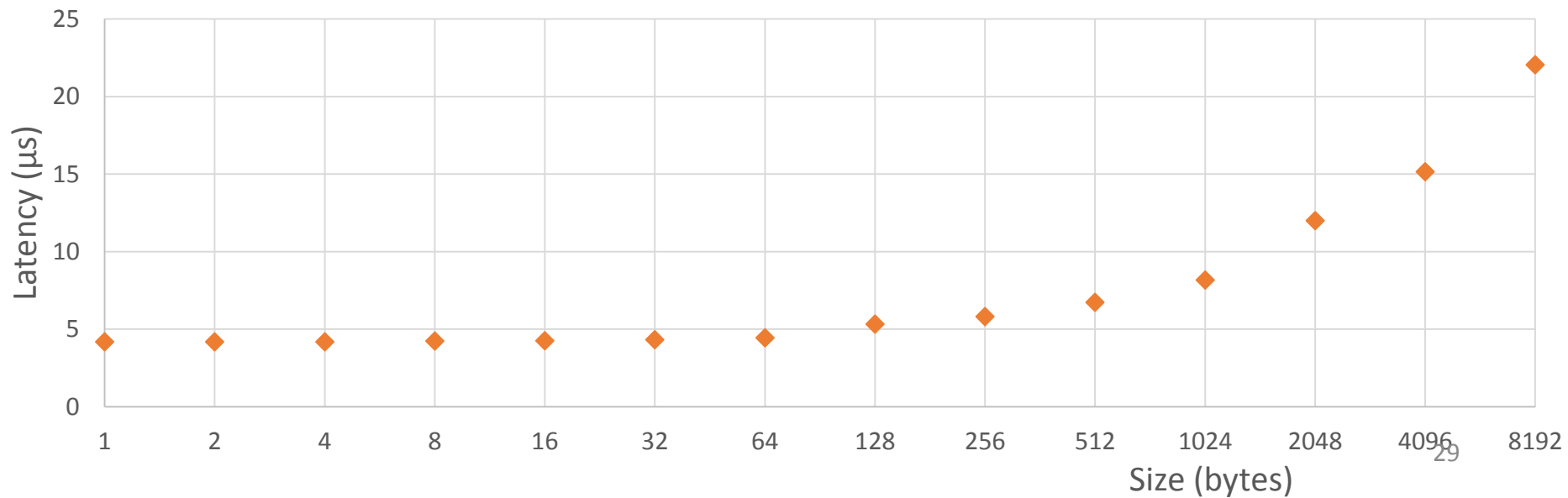
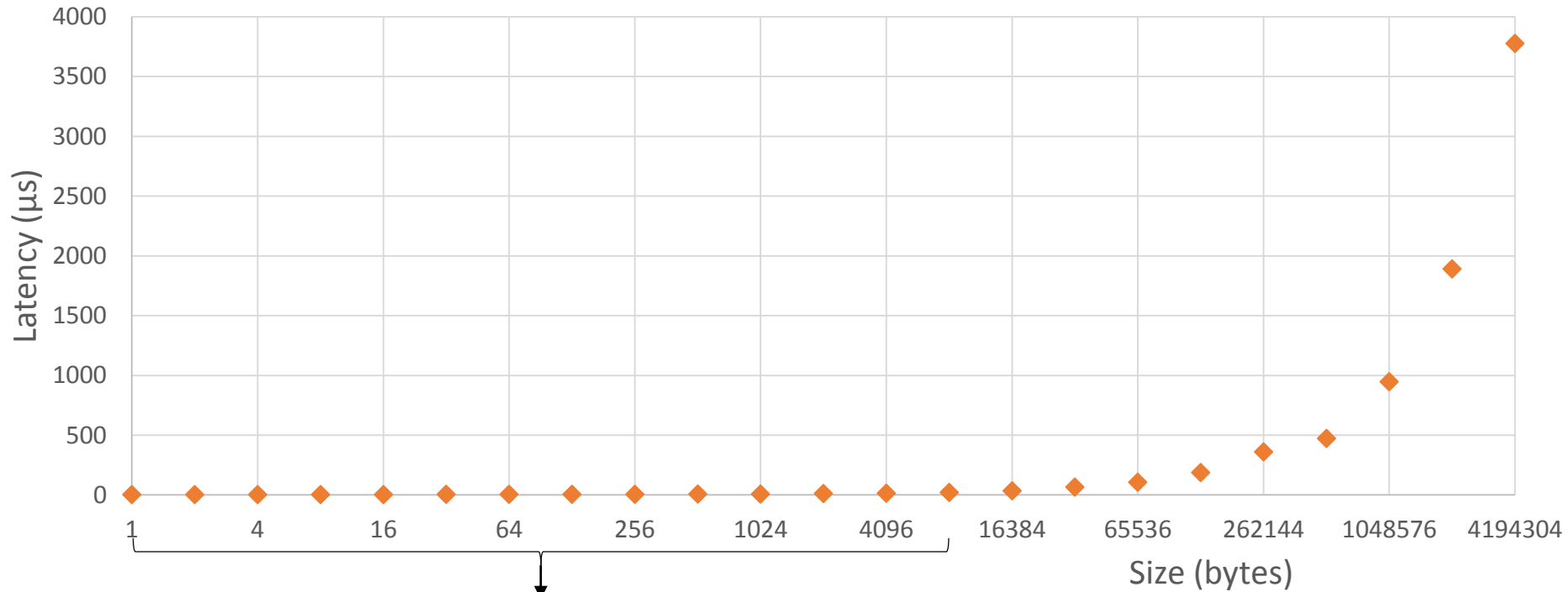
Scatter latency



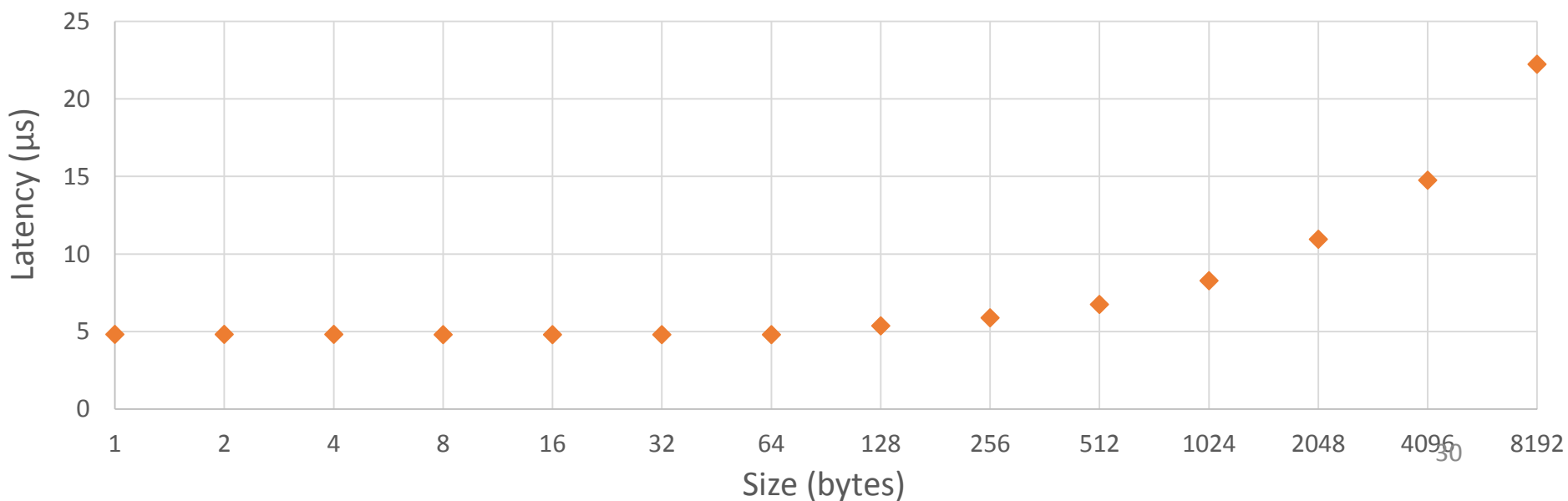
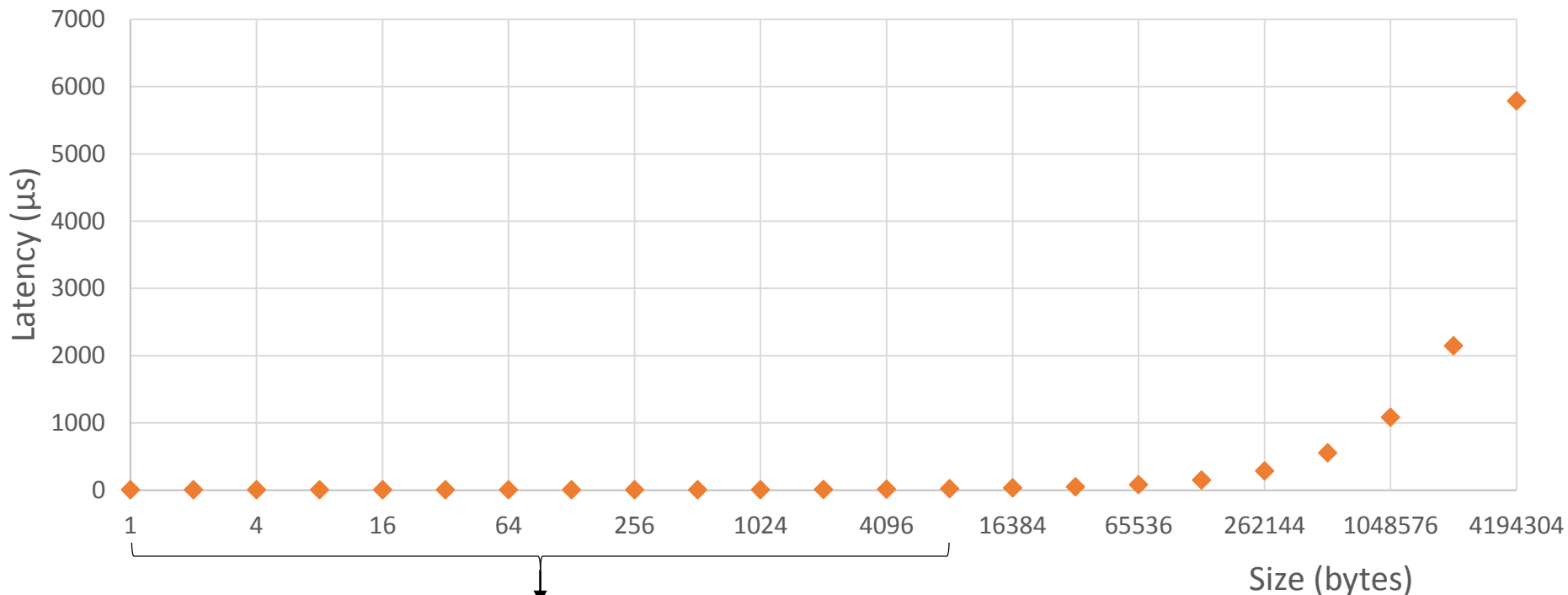
Allgather latency



Bcast latency



Alltoall Latency



Summary of Results:

Comparison to libibverbs:

- For the unidirectional latency (pingpong test), the benchmarks show significantly lower latencies than libibverbs:

Libibverbs: lower bound $\sim 10 \mu\text{s}$.

upper bound $\sim 6 \text{ ms}$.

OMB / IMB: lower bound $\sim 4 \mu\text{s}$.

upper bound $\sim 3 \text{ ms}$

Comparison between OMB and IMB:

- The OSU benchmarks show a slightly higher bandwidth than the Intel benchmarks

OMB: $\sim 10.9 \text{ Gb/s}$. IMB: $\sim 9.5 \text{ Gb/s}$.

- Similar latencies for MPI functions (gather, scatter, allgather, bcast, alltoall).