IMPACT OF SHARP AND ADAPTIVE ROUTING ON FRONTERA

John Cazes
Director, High Performance Computing

MVAPICH User Group Meeting
August 26, 2020

Frontera

Dell 8000+ node cluster

38 Pflops

43 PB Lustre filesystem

16 NVDIMM Nodes

Each node contains:

- **4 Intel Xeon Platinum 8280M chips**
- **2x 28 core 2.2 Ghz Xeon cores**
- **384 GB DRAM**
- **2 TB NVMe RAM**
- **4 TB NVMe disk**

8008 Cascade Lake Nodes

Each node contains:

- **2 Intel Xeon Platinum 8280 chips**
- **2x 28 core 2.2 Ghz Xeon cores**
- **192 GB DRAM**
- **Mellanox HDR Infiniband**

Mellanox HDR Infiniband

90 GPU Nodes

Each node contains

- **4 NVIDIA QUADRO RTX 5000 GPUs**
- **2 Intel Xeon E5-2620 v4**
- **192 GB DRAM**

FRONTERA – UNIQUE FEATURES

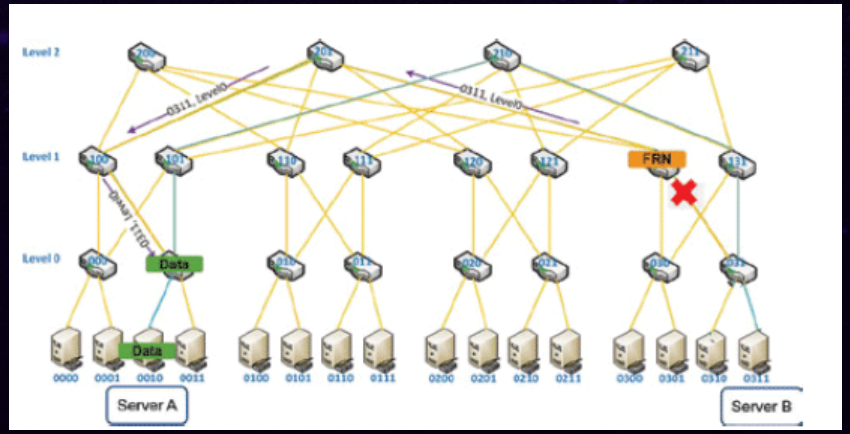
- RTX queue
  - 90 nodes
  - 4 NVIDIA Quadro 5000 RTX cards per node
  - 16 GB and 11 TFlop single-precision performance per card
  - MVAPICH2-GDR available
- NVDIMM queue
  - 16 nodes
  - 2 TB Memory per node
  - 4 local 833 GB partitions per node (3.2 TB local storage)
  - 112 Cascade Lake cores per node

INFINIBAND FABRIC TOPOLOGY

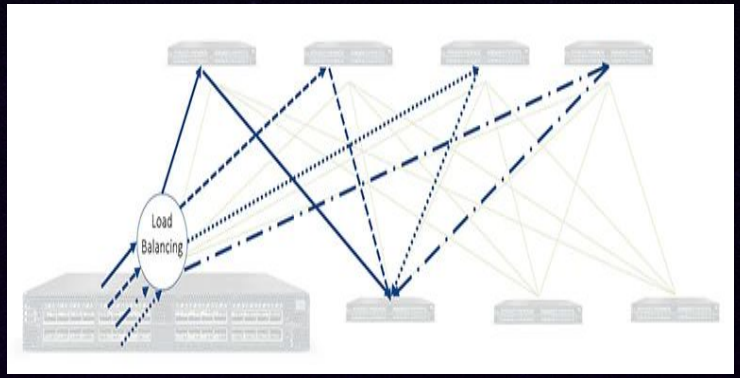
- Mellanox HDR 200Gb/s InfiniBand for interconnect
- Mellanox ConnectX-6 HDR100 InfiniBand for node
- Fat-tree topology design with 11:9 oversubscription
- Each Top of Rack switch connects:
  - 44 nodes at 100Gbps
  - 18 uplink cables at 200Gbps

ADAPTIVE ROUTING (AR)

Source : Mellanox



Traffic recovery - Recover traffic when link fails or switch reboots without intervention of Subnet Manager.



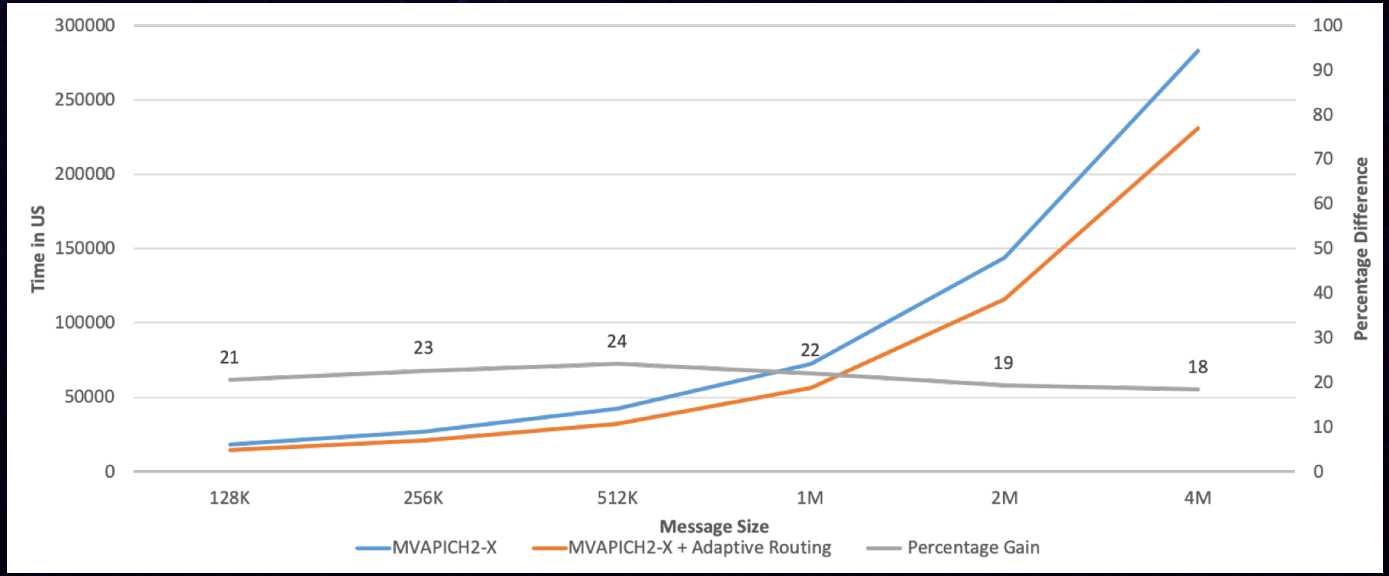
- **Load Balance** : Allows the switch to select the outgoing port based on port's load.

Source : Mellanox

ADAPTIVE ROUTING ON FRONTERA

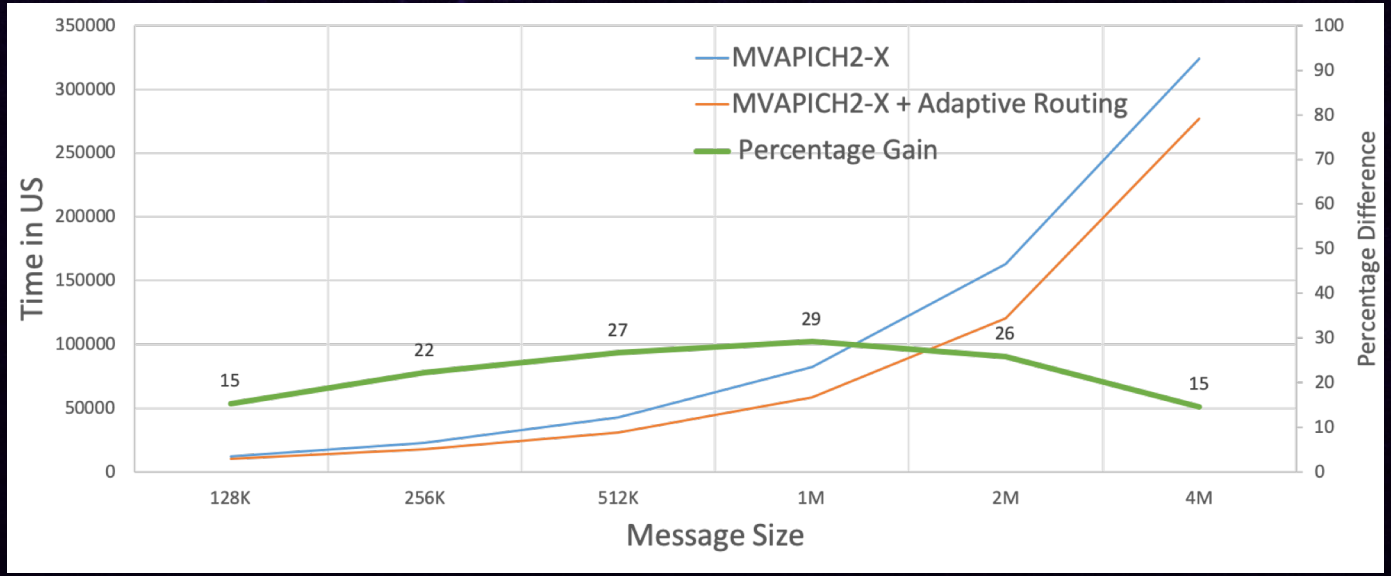
- Began experimenting with AR last year
- Admins enabled and disabled AR many times over the past year
- ~ 3 months ago AR enabled
- Admins discovered a performance degradation problem
- Mellanox and admins correlated performance degradation to AR
- AR disabled last month
- Admins and Mellanox will investigate a fix on Sep. 1

ADAPTIVE ROUTING - ALLGATHER BENCHMARK



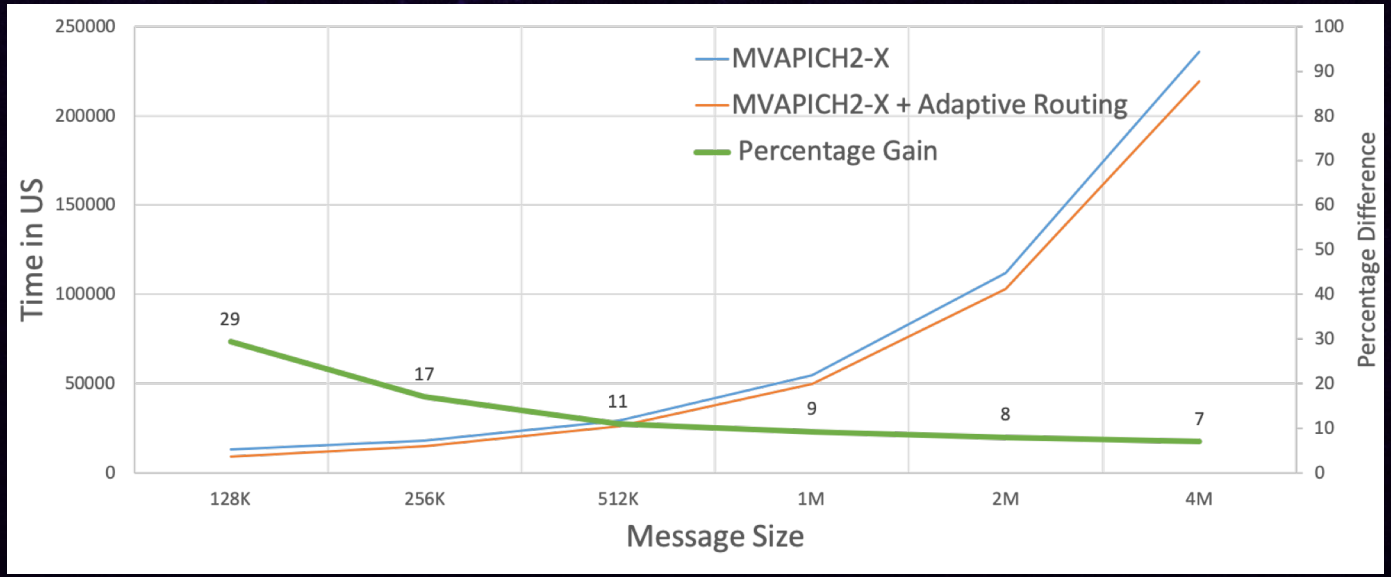
Nodes = 512 **Processes per node = 1**

ADAPTIVE ROUTING - ALLTOALL BENCHMARK



Nodes = 512 **Processes per node = 1**

ADAPTIVE ROUTING - ALLTOALLV BENCHMARK



Nodes = 256 **Processes per node = 1**

Source : Mellanox

ADAPTIVE ROUTING ON FRONTERA

- Initial benchmarks show impressive performance
- Continue benchmarking with applications
- Will enable by default after fix is validated
- Greatly intrigued by our initial results

Offloads collective operations from the CPU to the switch network

- Data aggregation and reduction while packets traverse network.
- Significant reduction in the amount of data transfer through the network, and therefore improves latency of MPI Collective operations
- Assist CPU for doing computation rather spending cycles in process communication.

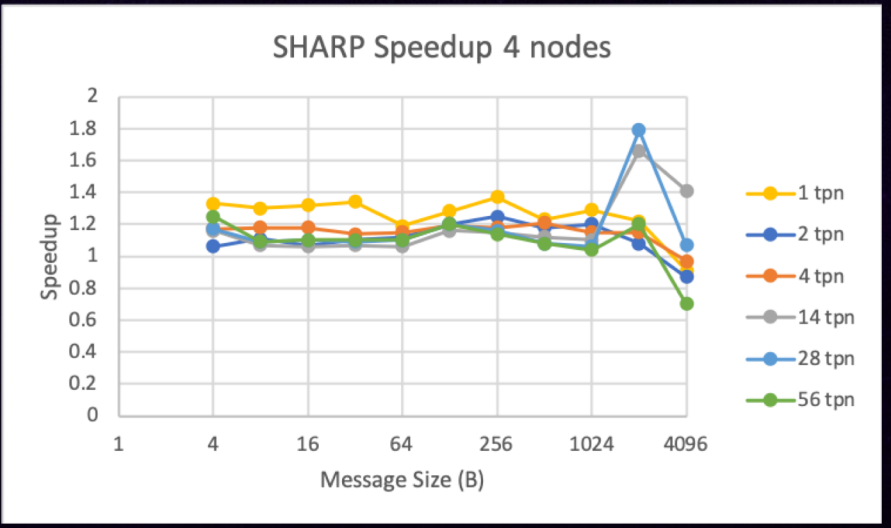
SHARP BENCHMARKS

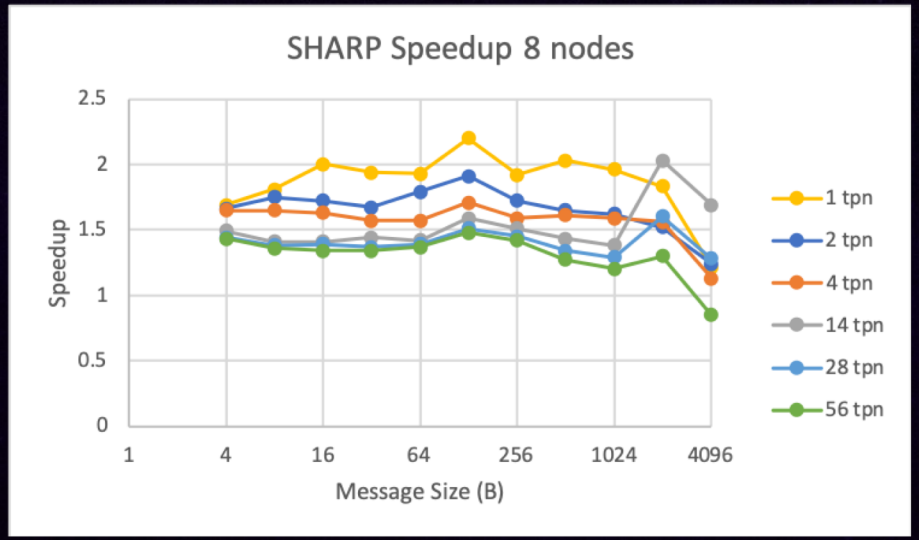
- SHARP available on Frontera since last maintenance
  - Previously, SHARP was not enabled
- Ran OSU micro benchmarks using MVAPICH2-X 2.3 pre-release
  - osu_allreduce
  - osu_barrier
- Working with MVAPICH2 group to validate Frontera SHARP configuration
  - Issues with task count per node
    - Task counts of 1, 2, and 4 tasks per node cause hangs for allreduce **– FIXED!**
  - Failed at 1024 nodes – **FIXED!**
- No application results using SHARP yet
  - Could be Frontera SHARP configuration – using the default – Not tuned

SHARP BENCHMARKS

- Timings from 9 runs spread across 3 different sets of nodes
- Enable SHARP
  - MV2_ENABLE_SHARP=1 -- Turns SHARP on
  - SHARP_COLL_LOG_LEVEL=3 – Enables Logging to determine executable using SHARP
- osu_allreduce
  - Time for allreduce
  - Up to 4k message size – not implemented for larger message size
  - Speedup = (time without SHARP)/(time with SHARP)
- osu_barrier
  - Time for barrier
  - Speedup = (time without SHARP)/(time with SHARP)

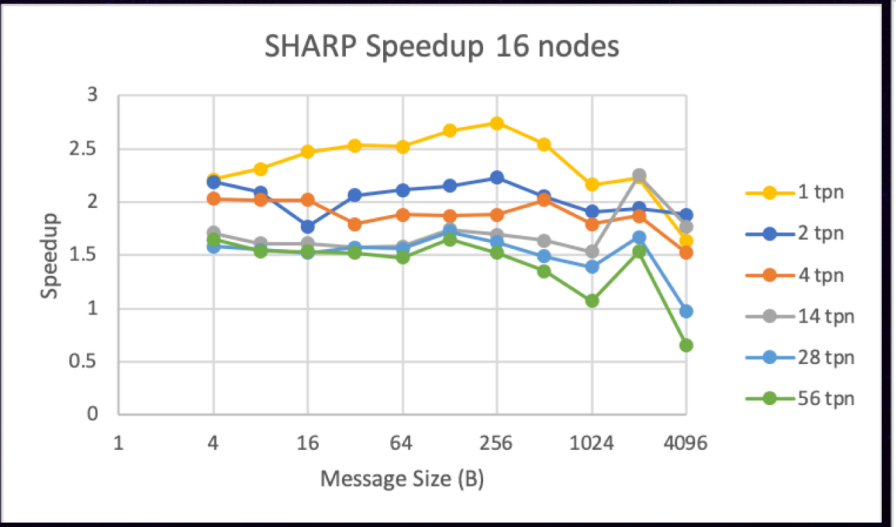
SHARP ALLREDUCE

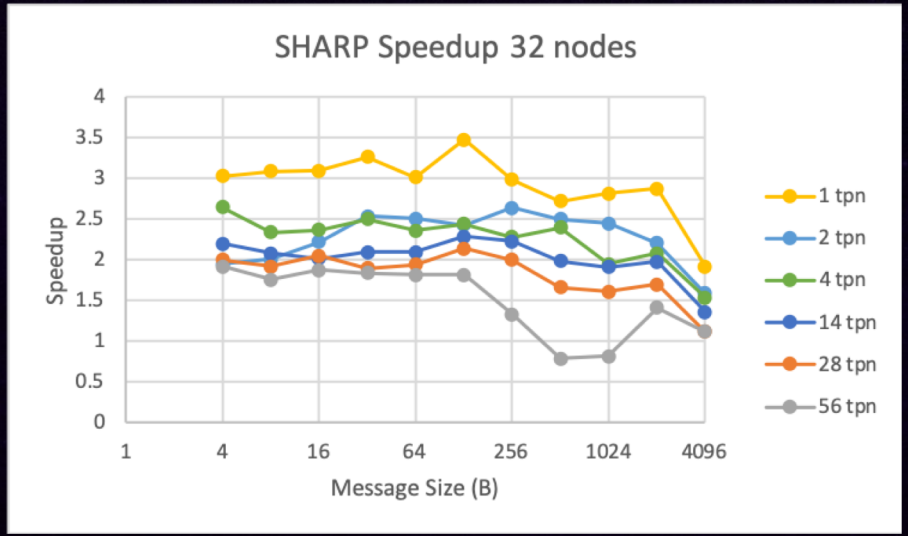




- Small improvement for small messages
- Consistent improvement except for 56 tasks per node at 4K message size

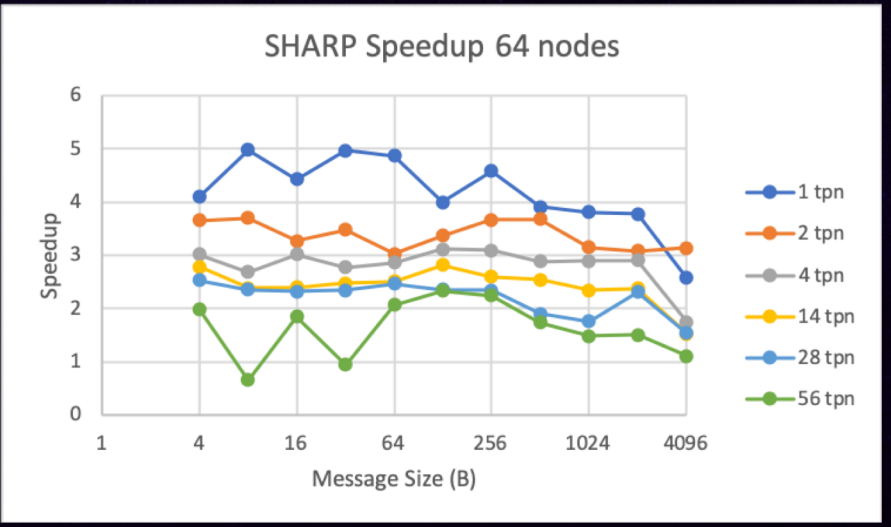
SHARP ALLREDUCE

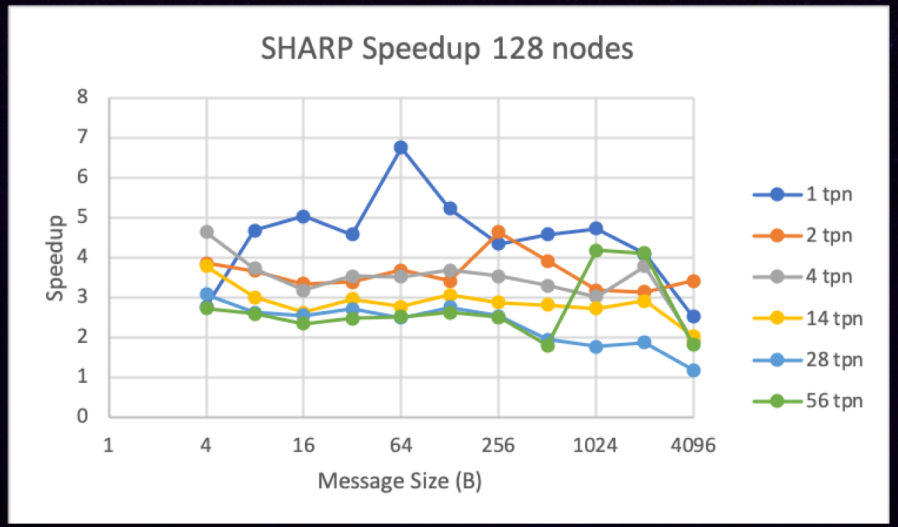




- Much better speedup with higher node counts
- Consistent improvement except for 56 tasks per node

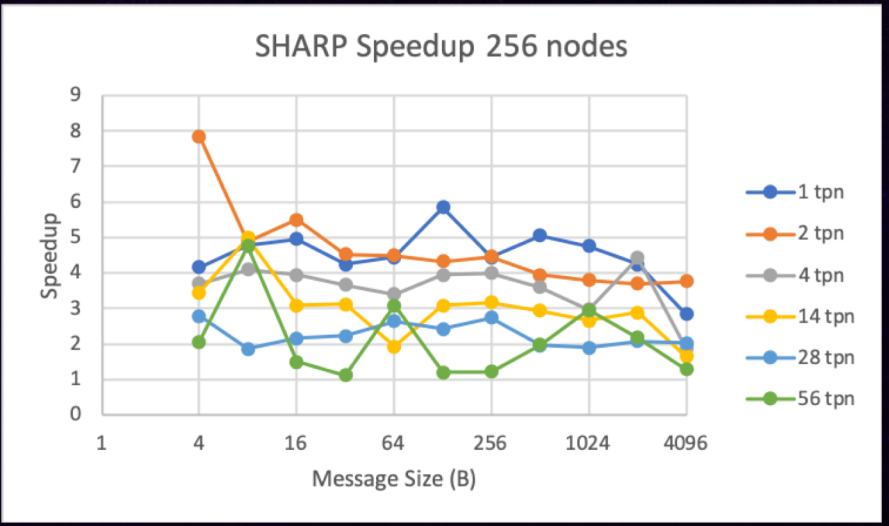
SHARP ALLREDUCE

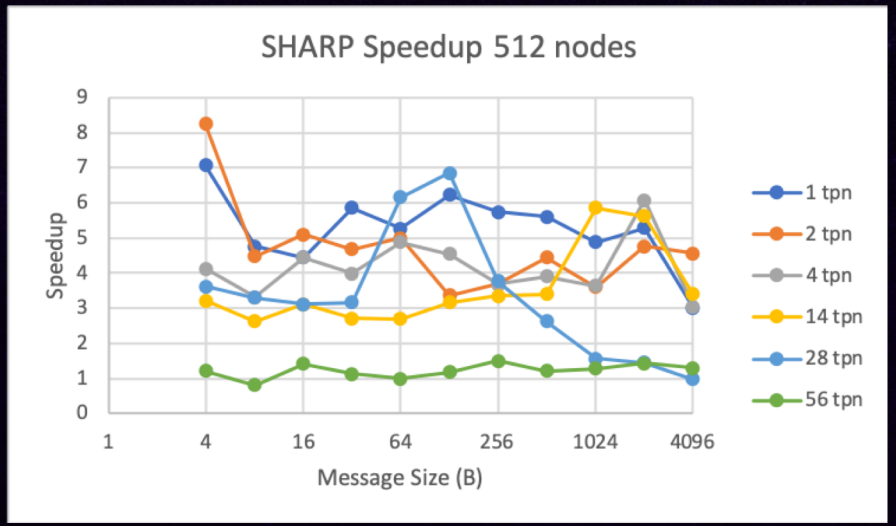




- 2x increase in performance for up to 28 tasks per node
- Less consistent using 56 tasks per node

SHARP ALLREDUCE

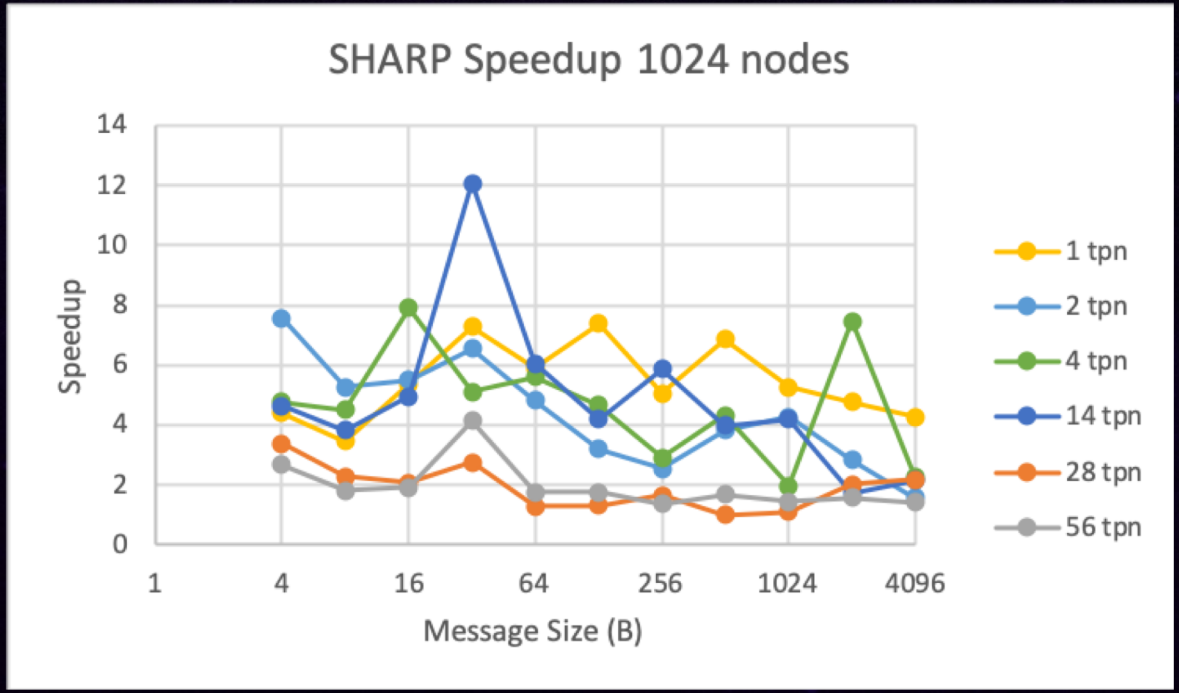




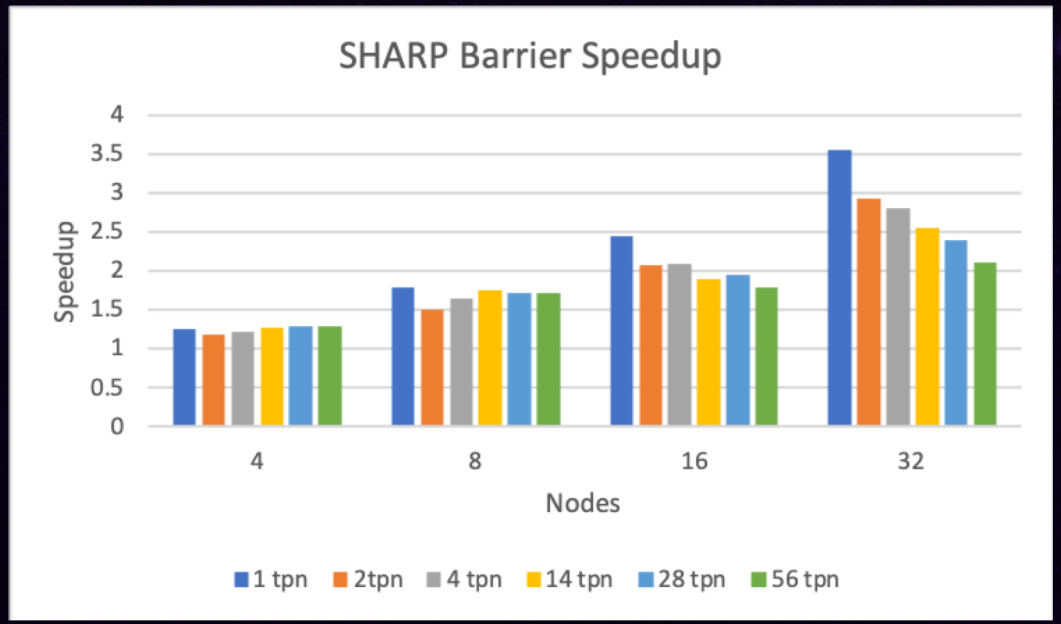
- Impressive performance using up to 28 tasks per node
- Less consistent using 56 tasks per node

SHARP ALLREDUCE

- Impressive performance using up to 14 tasks per node
- Less consistent using 28 and 56 tasks per node

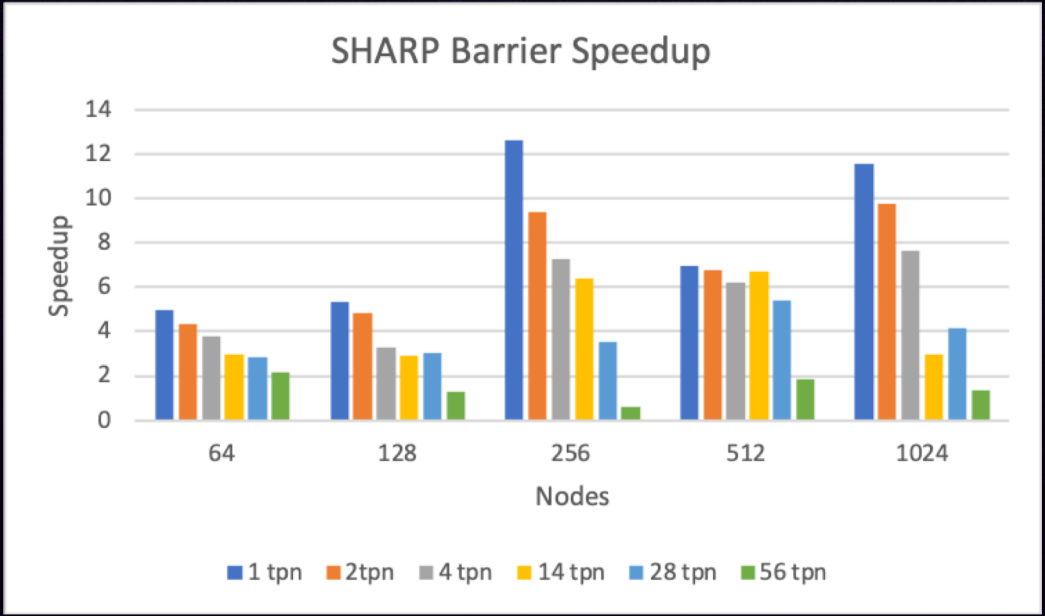


SHARP BARRIER



- Performance improves with node count
- Consistently better – should always enable

SHARP BARRIER



- Performance improves with node count up to 256
- Performance using 56 tasks not as good at high node count but > 1
- One case showed worse performance for SHARP – one anomalous measurement of 9

SHARP ON FRONTERA- IN PROGRESS

- Limited to Allreduce and Barrier operations in beta release
  - Assume other collectives are coming
- Performance variability increases at full task count per node
  - Not sure if this is Frontera specific or a more general characteristic
- Applications don't show SHARP getting triggered
  - OSU and TACC investigating
  - Tested with WRF and MILC

SUMMARY

- Resume AR testing once it's available
- Enable AR by default via module settings
- Continue to work with MVAPICH2 group to resolve SHARP issues
- Add SHARP instructions to MPI training for Frontera

TACC greatly appreciates the effort put in by the MVAPICH2 team to continuously improve performance and stability for large scale systems