HPC ADVISORY COUNCIL

RADIOSS
Performance Benchmark and Profiling

March 2013

Altair

hp

Mellanox TECHNOLOGIES

Note

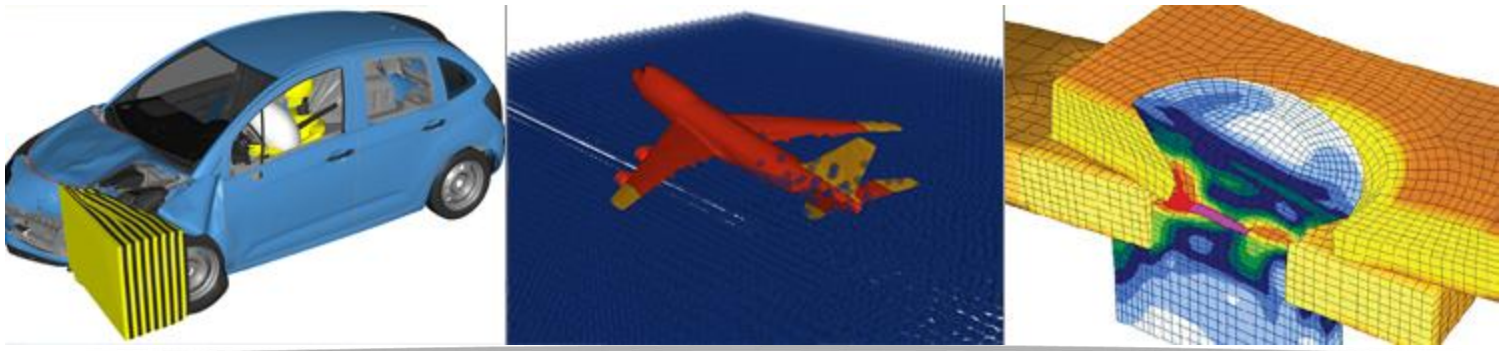
- **The following research was performed under the HPC Advisory Council activities**
  - Special thanks for: HP, Mellanox

- **For more information on the supporting vendors solutions please refer to:**
  - www.mellanox.com, http://www.hp.com/go/hpc

- **For more information on the application:**
  - http://www.altairhyperworks.com

RADIOSS by Altair

- **Altair® RADIOSS®**
  - Structural analysis solver for highly non-linear problems under dynamic loadings
  - Consists of features for:
    - multiphysics simulation and advanced materials such as composites
  - Highly differentiated for Scalability, Quality and Robustness
- **RADIOSS is used across all industry worldwide**
  - Improves crashworthiness, safety, and manufacturability of structural designs
- **RADIOSS has established itself as an industry standard**
  - for automotive crash and impact analysis for over 20 years



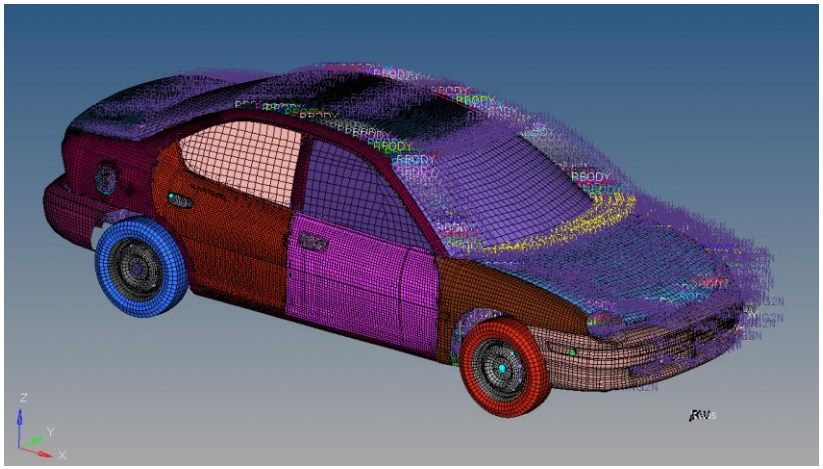
Objectives

- **The presented research was done to provide best practices**
  - RADIOSS performance benchmarking
  - Interconnect performance comparisons
  - MPI performance comparison
  - Understanding RADIOSS communication patterns

- **The presented results will demonstrate**
  - The scalability of the compute environment to provide nearly linear application scalability

Test Cluster Configuration

- **HP ProLiant SL230s Gen8 4-node "Athena" cluster**
  - Processors: Dual Eight-Core Intel Xeon E5-2680 @ 2.7 GHz
  - Memory: 32GB per node, 1600MHz DDR3 DIMMs
  - OS: RHEL 6 Update 2, OFED 1.5.3 InfiniBand SW stack
- **Mellanox ConnectX-3 VPI InfiniBand adapters**
- **Mellanox SwitchX SX6036 56Gbps InfiniBand and Ethernet Switch**
- **MPI: Intel MPI 4.1.0, Platform MPI 8.2**
- **Application: RADIOSS 12.0**
- **Benchmark Workload:**
  - NEON benchmarks: 1 million cells
    - (8ms, Double Precision)

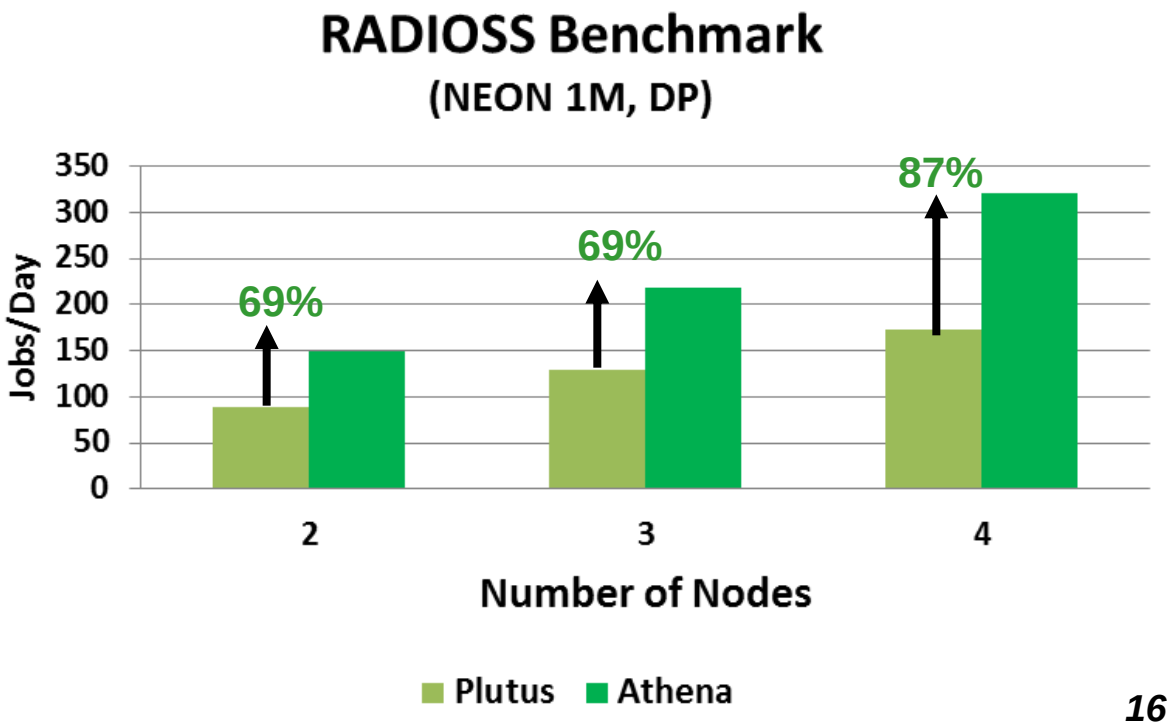


About HP ProLiant SL230s Gen8

Item	SL230 Gen8
Processor	Two Intel® Xeon® E5-2600 Series, 4/6/8 Cores,
Chipset	Intel® Sandy Bridge EP Socket-R
Memory	(512 GB), 16 sockets, DDR3 up to 1600MHz, ECC
Max Memory	512 GB
Internal Storage	Two LFF non-hot plug SAS, SATA bays or Four SFF non-hot plug SAS, SATA, SSD bays Two Hot Plug SFF Drives (Option)
Max Internal Storage	8TB
Networking	Dual port 1GbE NIC/ Single 10G NIC
I/O Slots	One PCIe Gen3 x16 LP slot 1Gb and 10Gb Ethernet, IB, and FlexFabric options
Ports	Front: (1) Management, (2) 1GbE, (1) Serial, (1) S.U.V port, (2) PCIe, and Internal Micro SD card & Active Health
Power Supplies	750, 1200W (92% or 94%), high power chassis
Integrated Management	iLO4 hardware-based power capping via SL Advanced Power Manager
Additional Features	Shared Power & Cooling and up to 8 nodes per 4U chassis, single GPU support, Fusion I/O support
Form Factor	16P/8GPUs/4U chassis

RADIOSS Benchmark – CPU Generation

- **Intel E5-2680 processors (Sandy Bridge) cluster outperforms prior CPU generation**
  - Performs 87% higher than the Plutus cluster at 4 nodes
- **System components used:**
  - Athena: 2-socket Intel E5-2680 @ 2.7GHz, 1600MHz DIMMs, FDR InfiniBand, 1HDD
  - Plutus: 2-socket Intel X5670 @ 2.93GHz, 1333MHz DIMMs, QDR InfiniBand, 1HDD

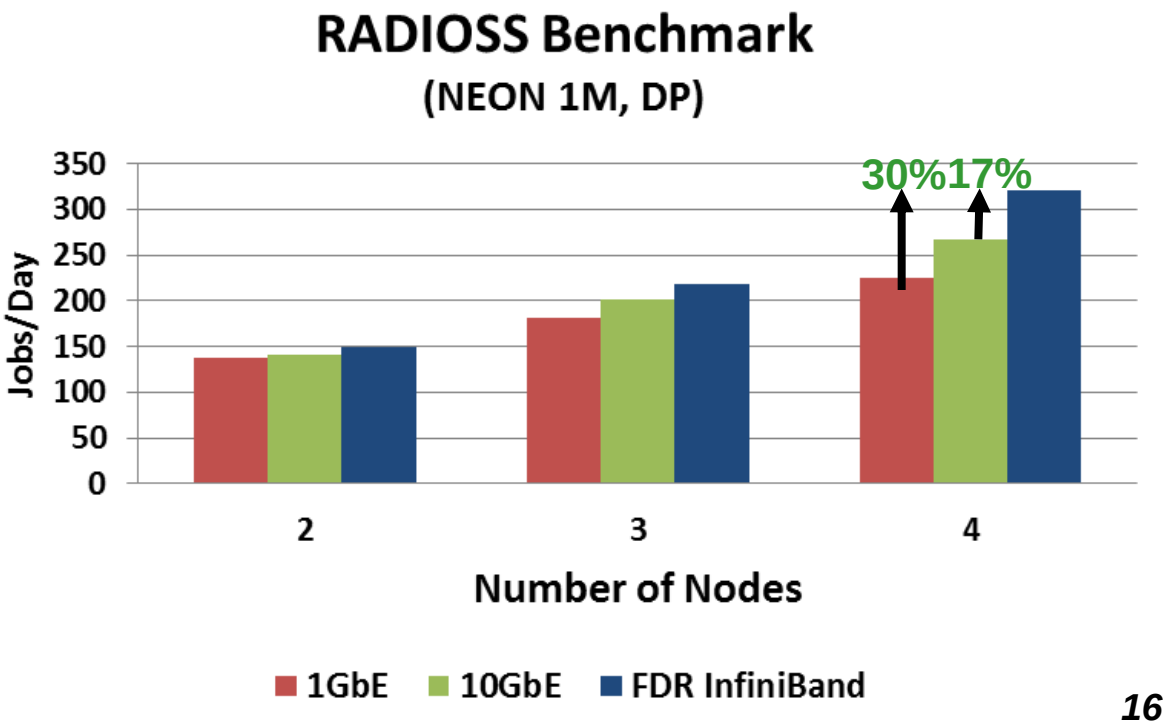


Higher is better

16 Processes/Node

RADIOSS Performance - Interconnect

- **InfiniBand FDR is the most efficient inter-node communication for RADIOSS**
  - Outperforms 1GbE by 30% at 4 nodes
  - Outperforms 10GbE by 17% at 4 nodes

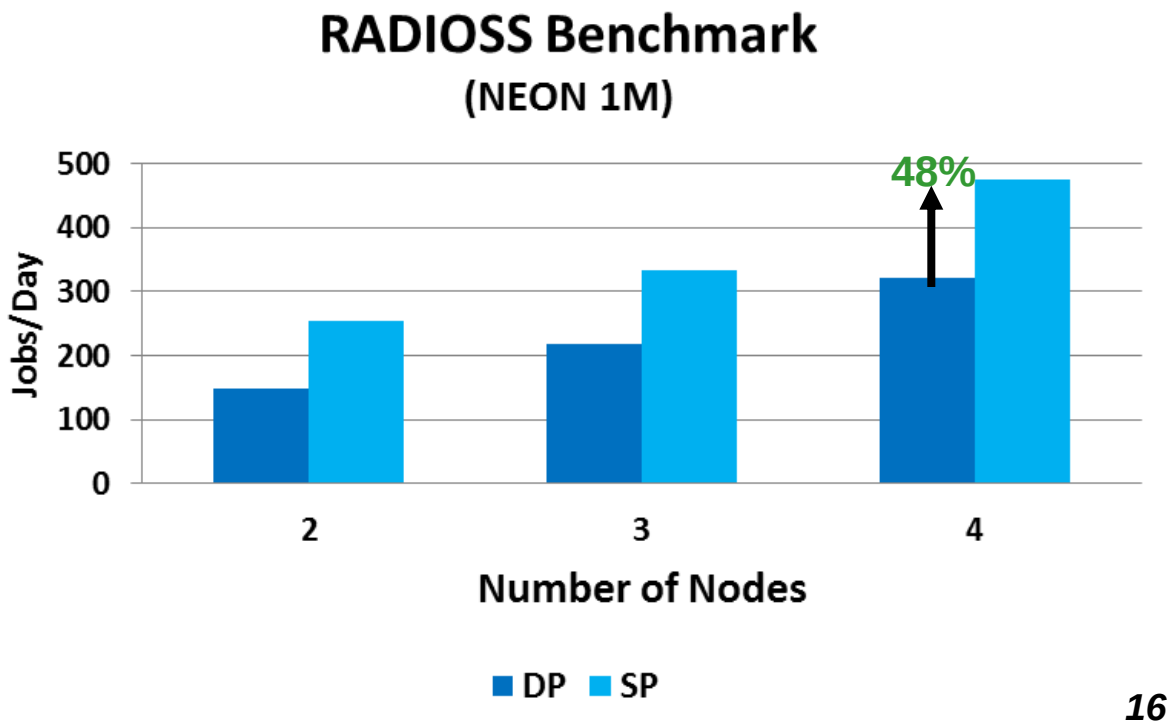


Higher is better

16 Processes/Node

RADIOSS Performance – SP vs DP

- **Running at Single Precision is faster than Double Precision**
  - Up to 48% faster at 4 nodes when running with Single Precision than Double Precision

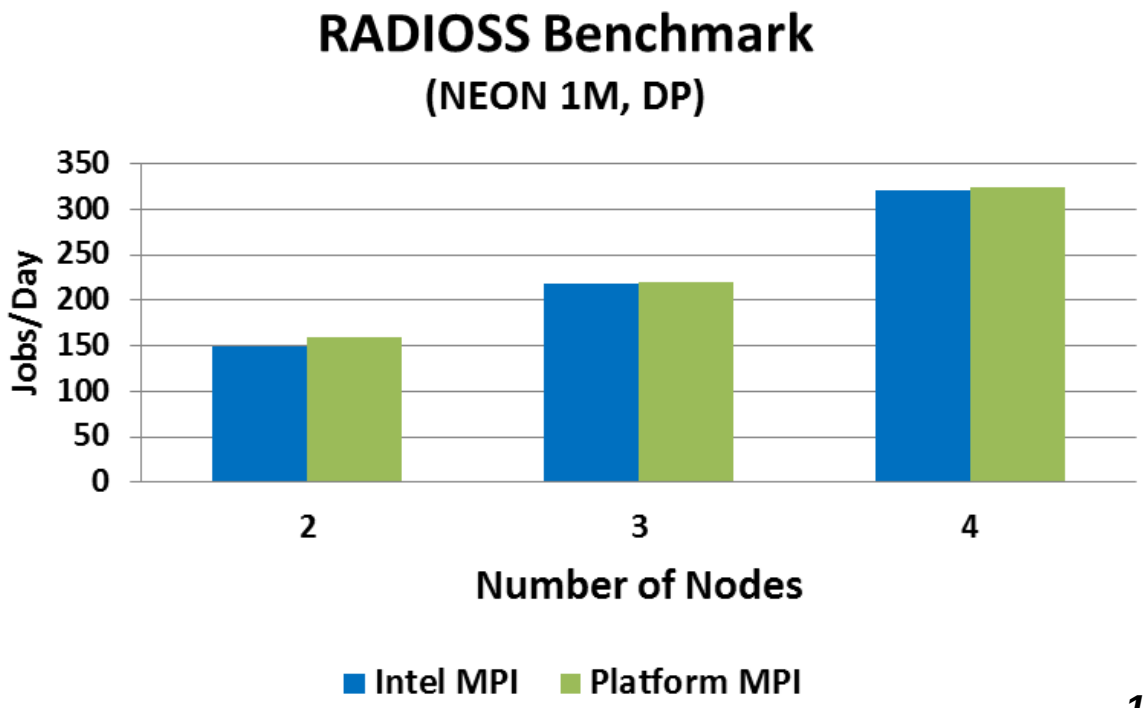


Higher is better

16 Processes/Node

RADIOSS Performance – MPI

- **Both Platform MPI and Intel MPI performs similarly in performance**
  - Platform MPI has processor binding enabled using the "-cpu_bind" flag
  - Intel MPI flags used: "-genv I_MPI_FABRICS_LIST ofa –genv I_MPI_FALLBACK 0 -genv I_MPI_ADJUST_BCAST 1 -genv I_MPI_ADJUST_REDUCE 2"

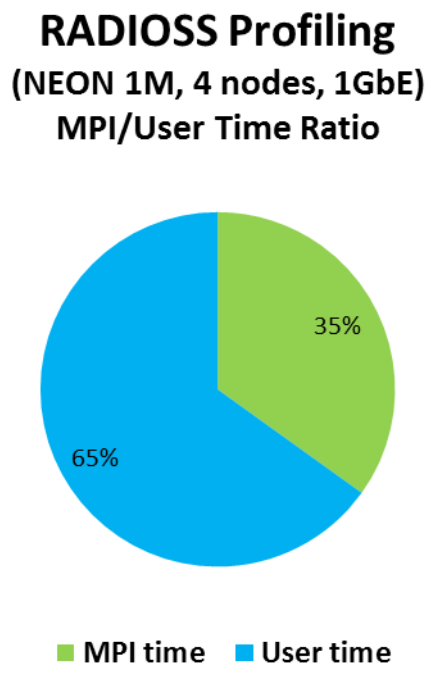


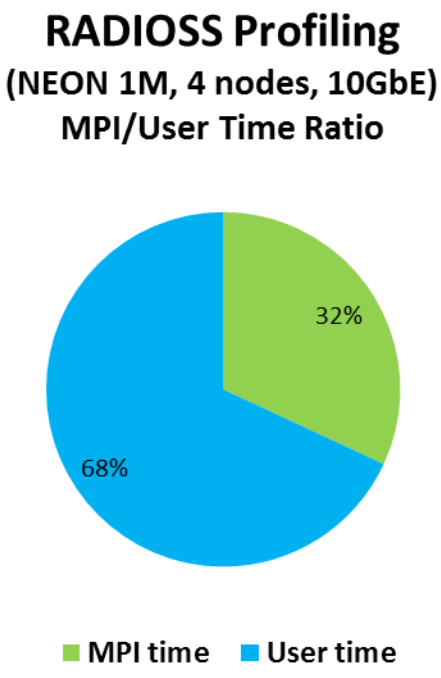
Higher is better

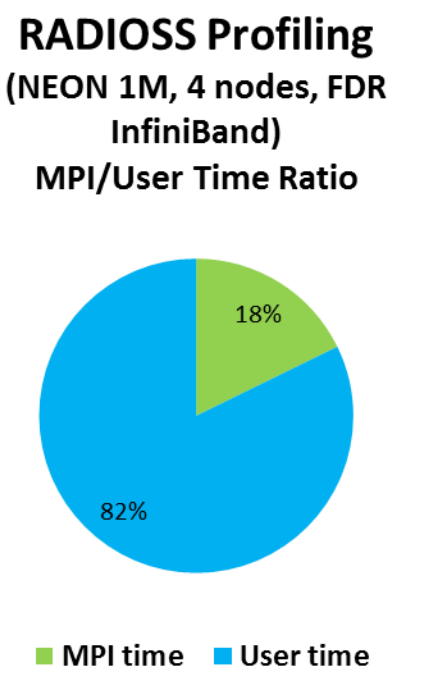
16 Processes/Node

RADIOSS Profiling – MPI Time Ratio

- **FDR InfiniBand reduces the MPI communication time**
  - FDR InfiniBand consumes about 18% of total runtime at 4 nodes
  - Ethernet solutions consume from 32% to 35% at 4 nodes

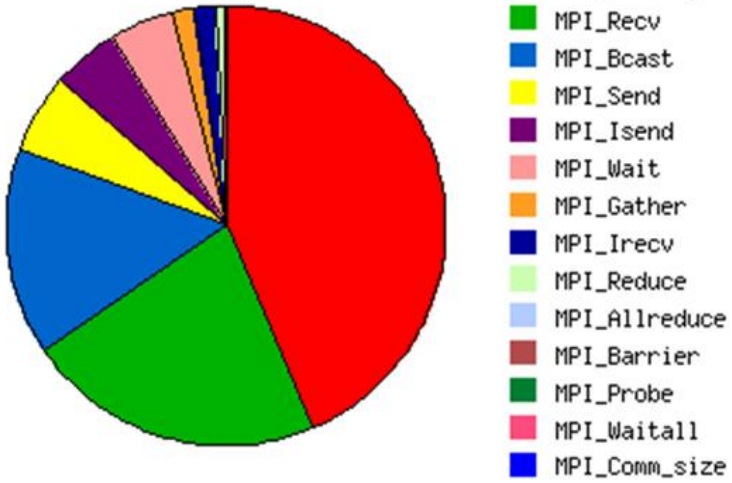


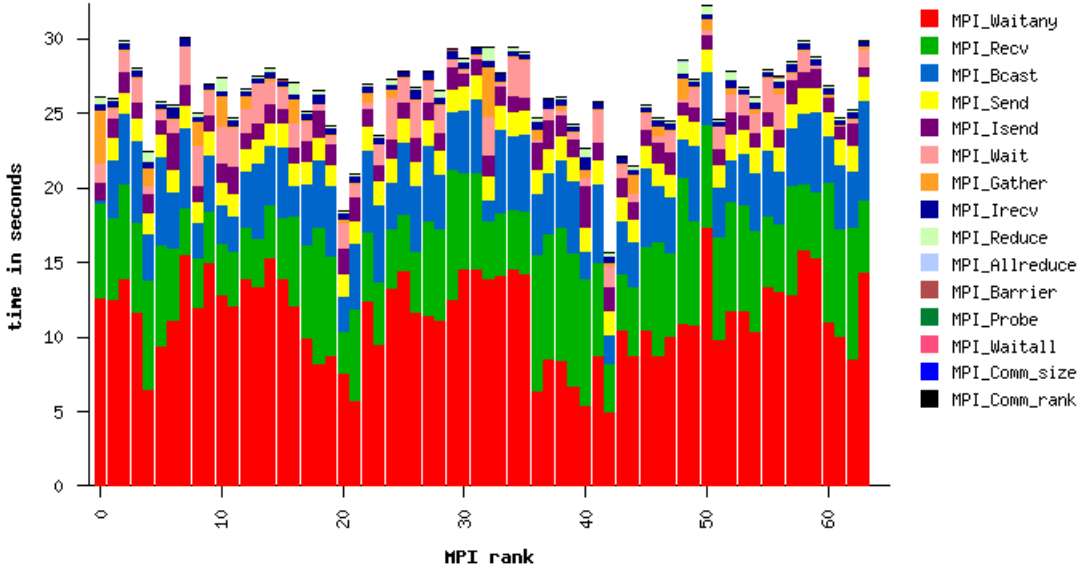




RADIOSS Profiling – Time Spent in MPI

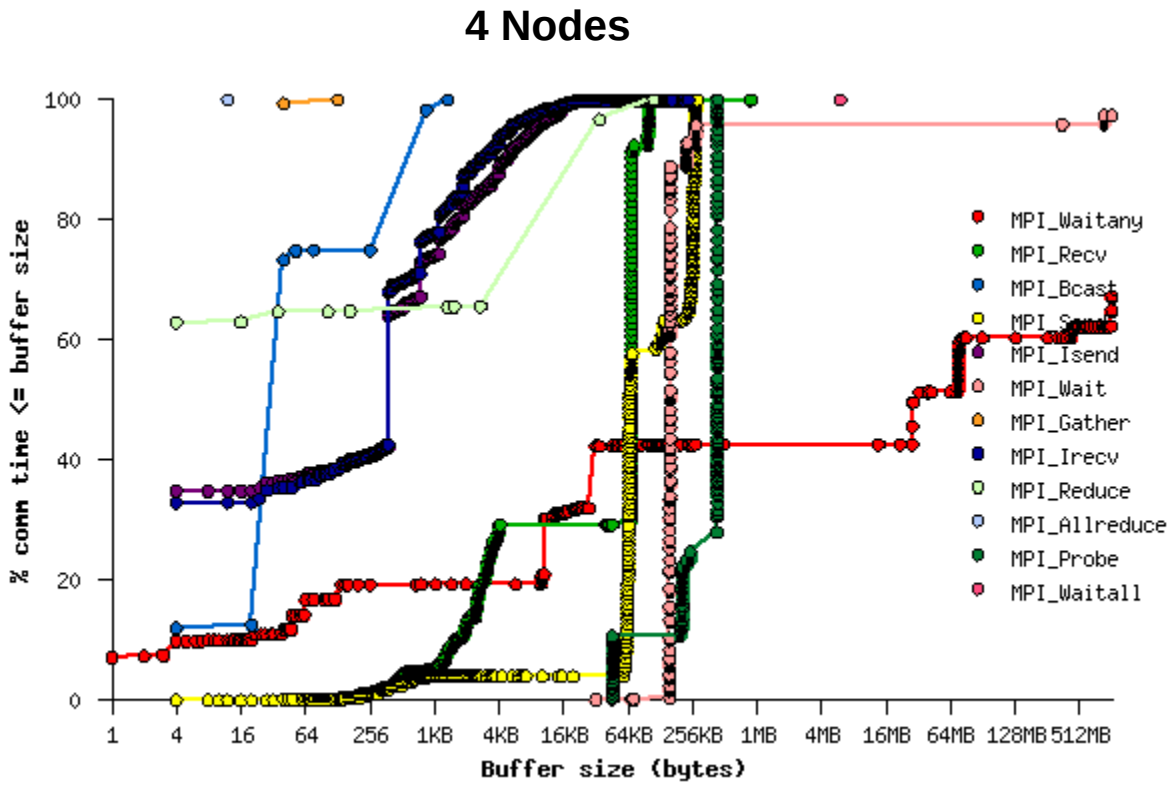
- **RADIOSS: More time spent on non-blocking MPI communications**
  - Time spent on MPI_Waitany takes the most time for waiting on non-blocking comms





RADIOSS Profiling – Message Sizes

- **RADIOSS shows distribution of small to midrange message sizes**
  - Small messages peak in the range from 64KB to 256KB



RADIOSS Summary

- **HP ProLiant Gen8 servers delivers better performance than its predecessor**
  - ProLiant Gen8 equipped with Intel E5 series processes and InfiniBand FDR
  - Up to 87% higher performance than ProLiant G7 when compared at 4 nodes
- **InfiniBand FDR is the most efficient inter-node communication for RADIOSS**
  - Outperforms 1GbE by 30% at 4 nodes
  - Outperforms 10GbE by 17% at 4 nodes
- **RADIOSS Profiling**
  - InfiniBand FDR reduces communication time; provides more time for computation
    - InfiniBand FDR consumes 18% of total time, versus 32-35% for Ethernet solutions
  - MPI:
    - MPI time is spent mostly on non-blocking communications

Thank You

HPC Advisory Council