

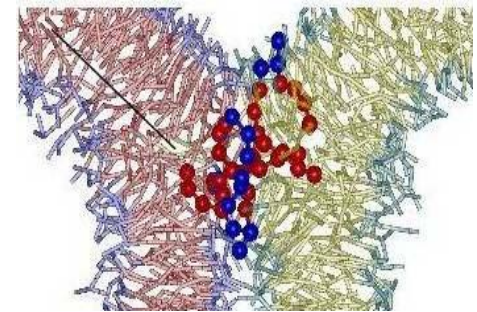
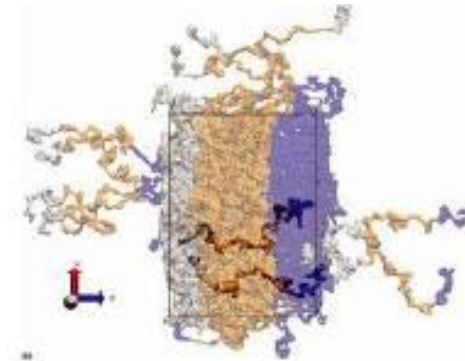
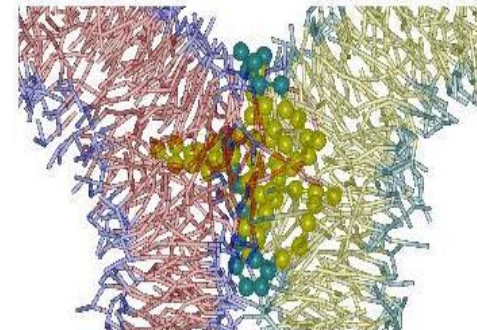
# LAMMPS Performance Benchmark and Profiling

September 2010



- **The following research was performed under the HPC Advisory Council activities**
- **For more info please refer to**
  - <http://www.dell.com>
  - <http://www.intel.com>
  - <http://www.mellanox.com>
  - <http://lammmps.sandia.gov/>

- **Large-scale Atomic/Molecular Massively Parallel Simulator**
  - Classical molecular dynamics code which can model:
    - Atomic
    - Polymeric
    - Biological
    - Metallic
    - Granular, and coarse-grained systems
- **LAMMPS runs efficiently in parallel using message-passing techniques**
  - Developed at Sandia National Laboratories
  - An open-source code, distributed under GNU Public License

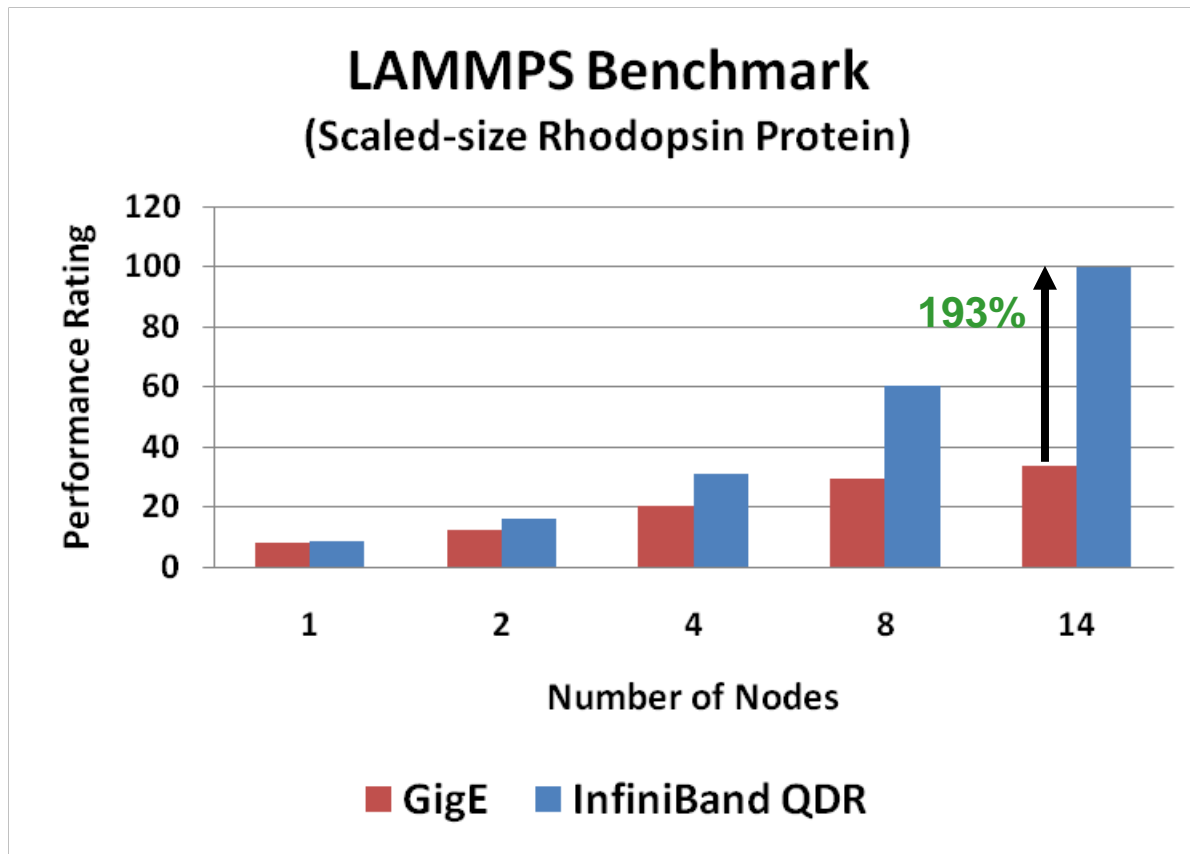


- **The following was done to provide best practices**
  - LAMMPS performance benchmarking
  - Interconnect performance comparisons
  - Understanding LAMMPS communication patterns
  - Power-efficient simulations
- **The presented results will demonstrate**
  - The scalability of the compute environment to provide nearly linear application scalability
  - The capability of LAMMPS to achieve scalable productivity
  - Considerations for power saving through balanced system configuration

- **Dell™ PowerEdge™ M610 14-node cluster**
  - Six-Core Intel X5670 @ 2.93 GHz CPUs
  - Memory: 24GB per node
  - OS: CentOS5U4, OFED 1.5.1 InfiniBand SW stack
- **Mellanox ConnectX-2 InfiniBand adapters and switches**
- **MPI: Intel MPI 4, MVAPICH2 1.5, Open MPI 1.4.2, Platform MPI 7.1**
- **Compilers Intel Compilers 11.1, GNU 4.1.2**
- **Libraries: Intel MKL 10.2.4.032, fftw-2.1.5**
- **Application: LAMMPS-30Aug10**
- **Benchmark Workload**
  - Rhodo - Rhodopsin protein in solvated lipid bilayer, CHARMM force field with a 10 Angstrom LJ cutoff

- **InfiniBand enables higher scalability**

- Up to 193% higher performance than Ethernet at 14 nodes
- Four InfiniBand connected servers deliver nearly the same performance vs 14 Ethernet connected servers

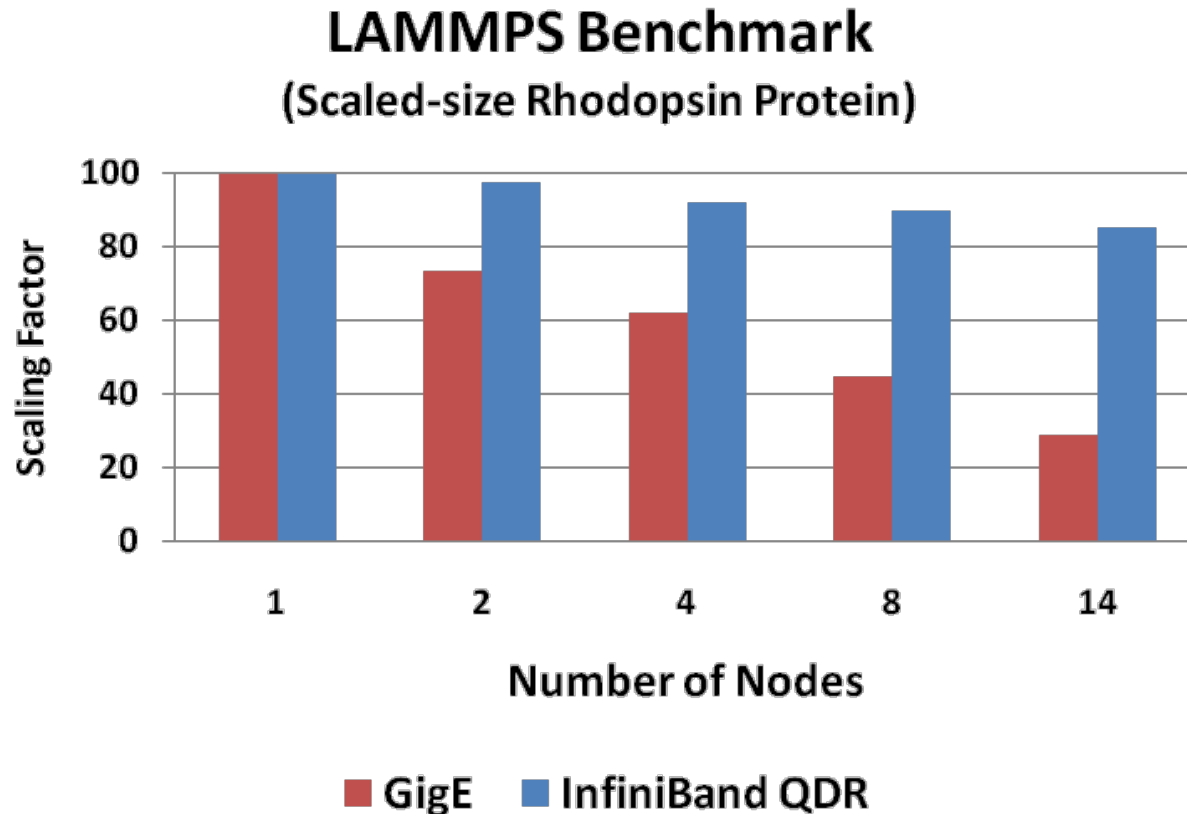


Performance Rating = 32,000 atoms x the number of cores divided by the wall-clock simulation time

*Higher is better*

**12 Cores/Node**

- InfiniBand enables higher scalability – more than 85% at 14 nodes
- Less than 30% of the system compute capability can be utilized with Ethernet at 14 nodes



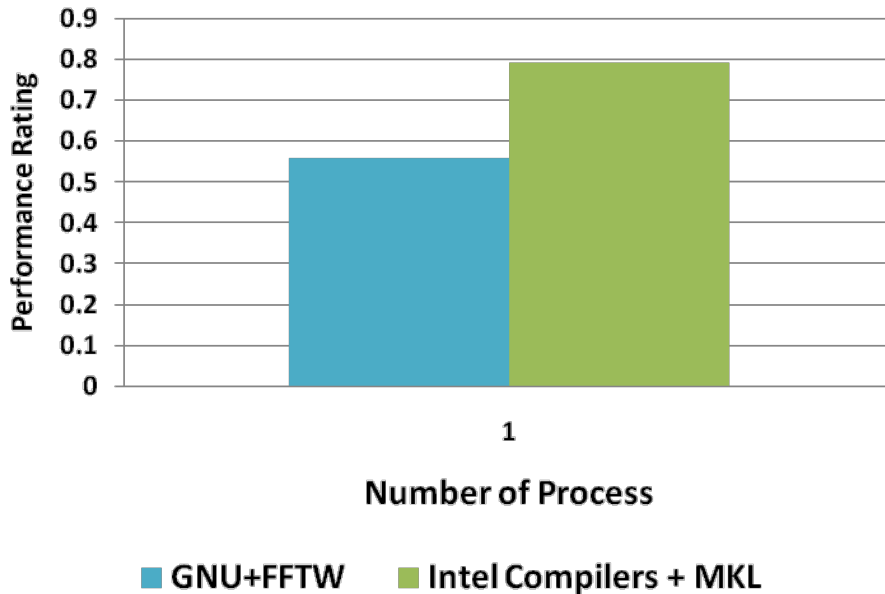
*Higher is better*

12 Cores/Node

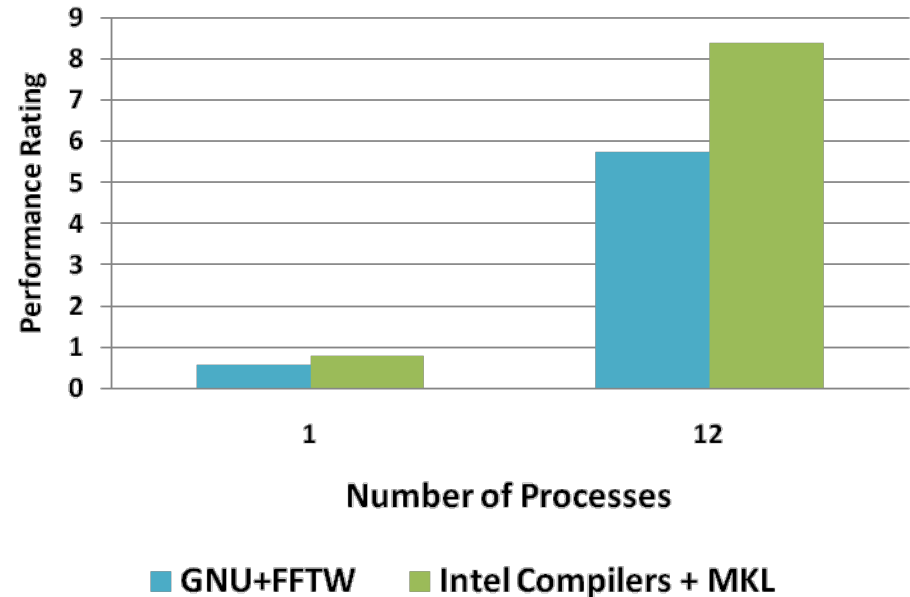
# LAMMPS Performance Result – Compilers and Libraries (Single node)

- **Intel Compilers and MKL libraries provides higher performance versus GNU compilers and FFTW libraries**
  - ~41% higher per process basis
  - ~46% higher per node basis

**LAMMPS Benchmark**  
(Scaled-size Rhodopsin Protein)



**LAMMPS Benchmark**  
(Scaled-size Rhodopsin Protein)

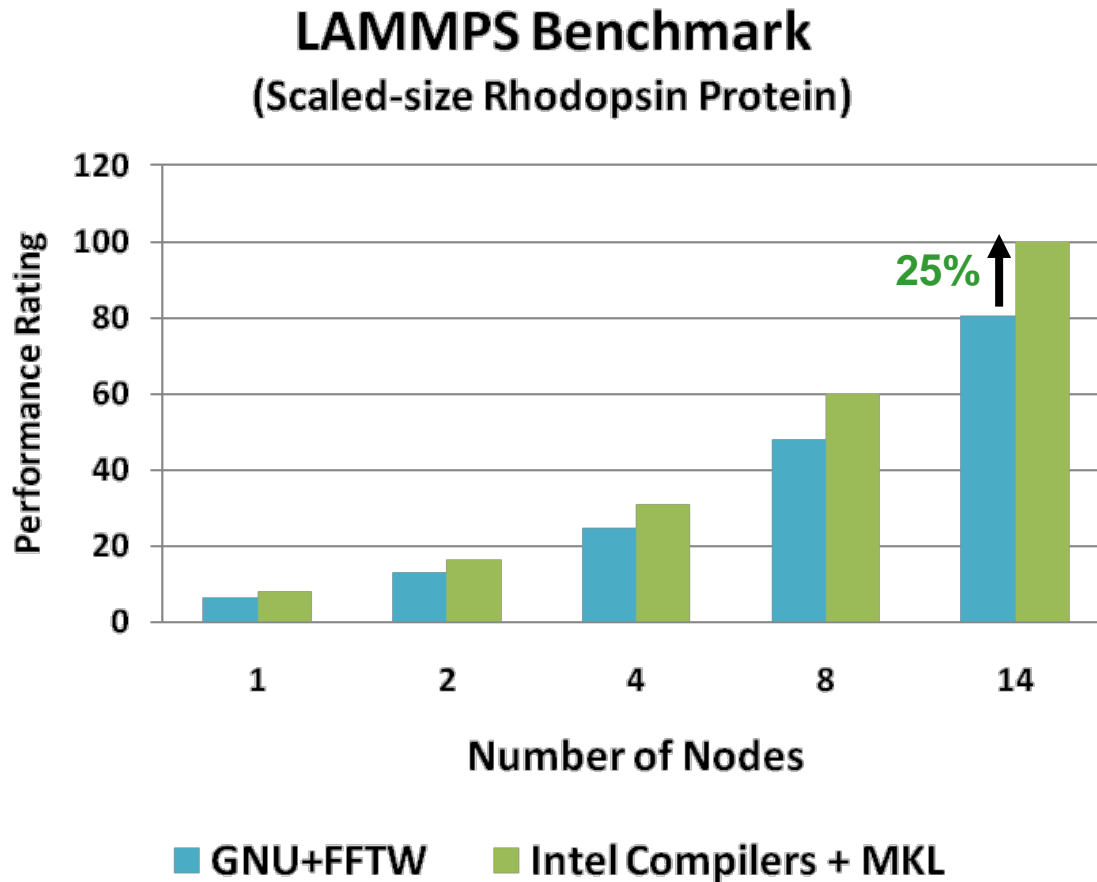


**Higher is better**

**12 Cores/Node**

# LAMMPS Performance Result – Compilers and Libraries (Multi-node)

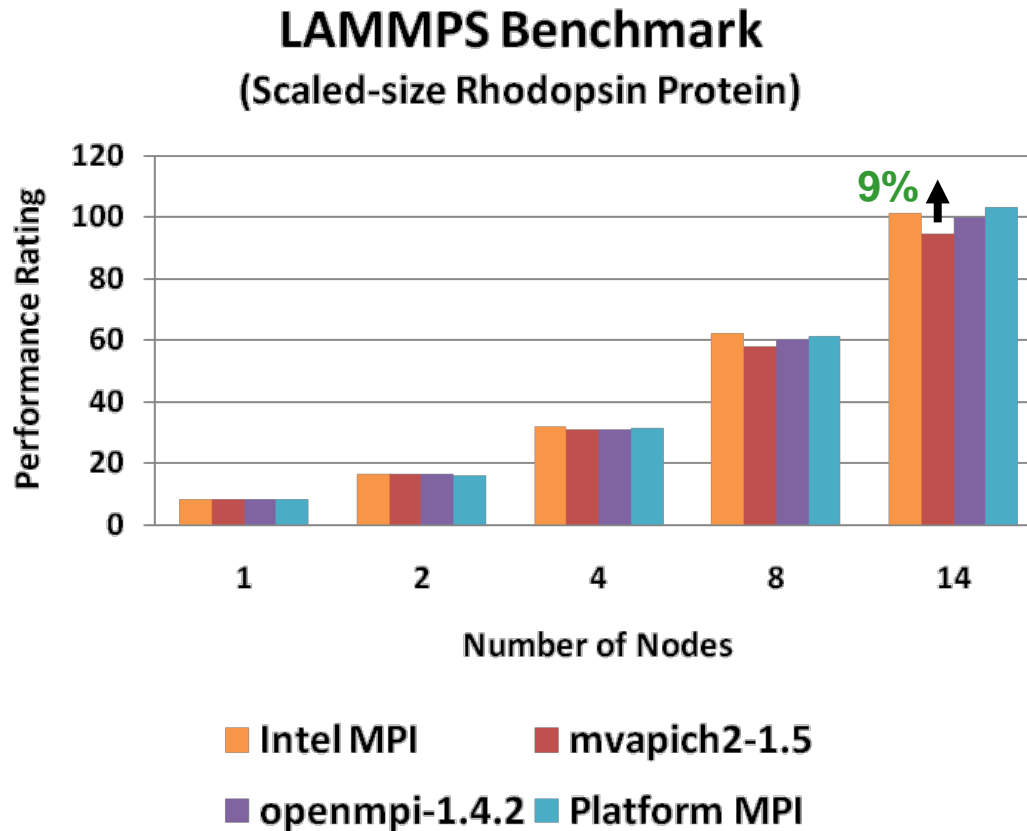
- Intel Compilers enables 25% higher performance than GNU Compilers and FFTW library



*Higher is better*

*12 Cores/Node*

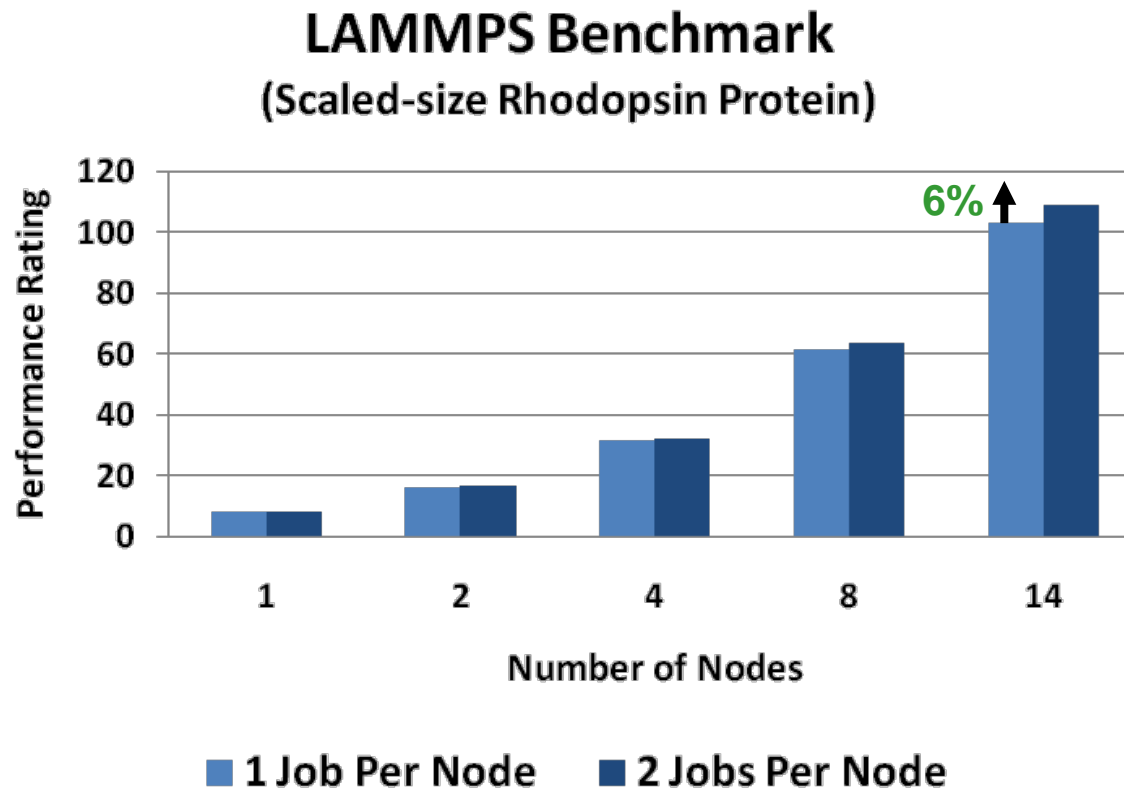
- **Intel MPI and Platform MPI demonstrate better performance**
  - Used with Intel compilers and MKL libraries together to deliver the highest performance
  - 9% gain compared to Open MPI at 14-node



*Higher is better*

**12 Cores/Node**

- **Running 2 jobs concurrently can provide slightly higher productivity**
  - 2 jobs setup: 6 cores/node per job; 1 job setup: 12 cores/node
  - Up to 6% higher performance when 2 jobs running concurrently

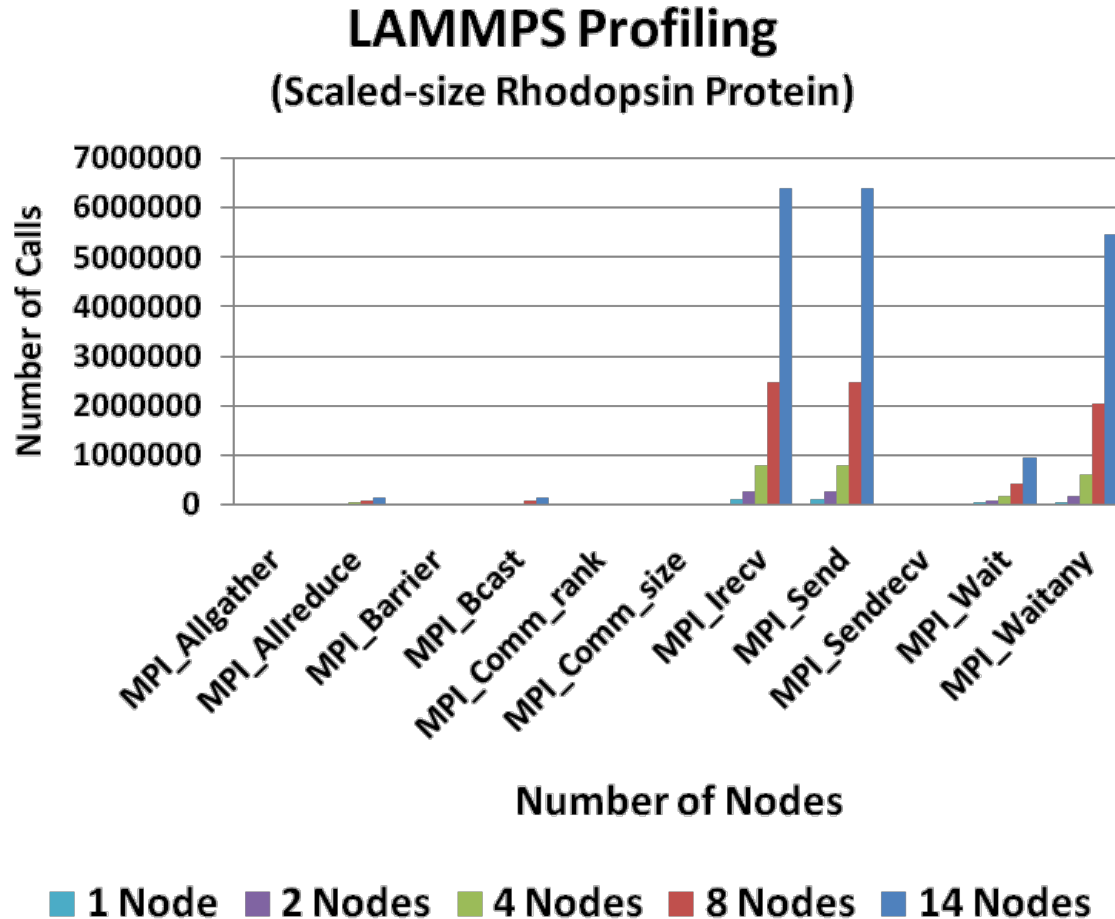


*Higher is better*

*12 Cores/Node*

# LAMMPS Profiling Result – # of MPI Calls

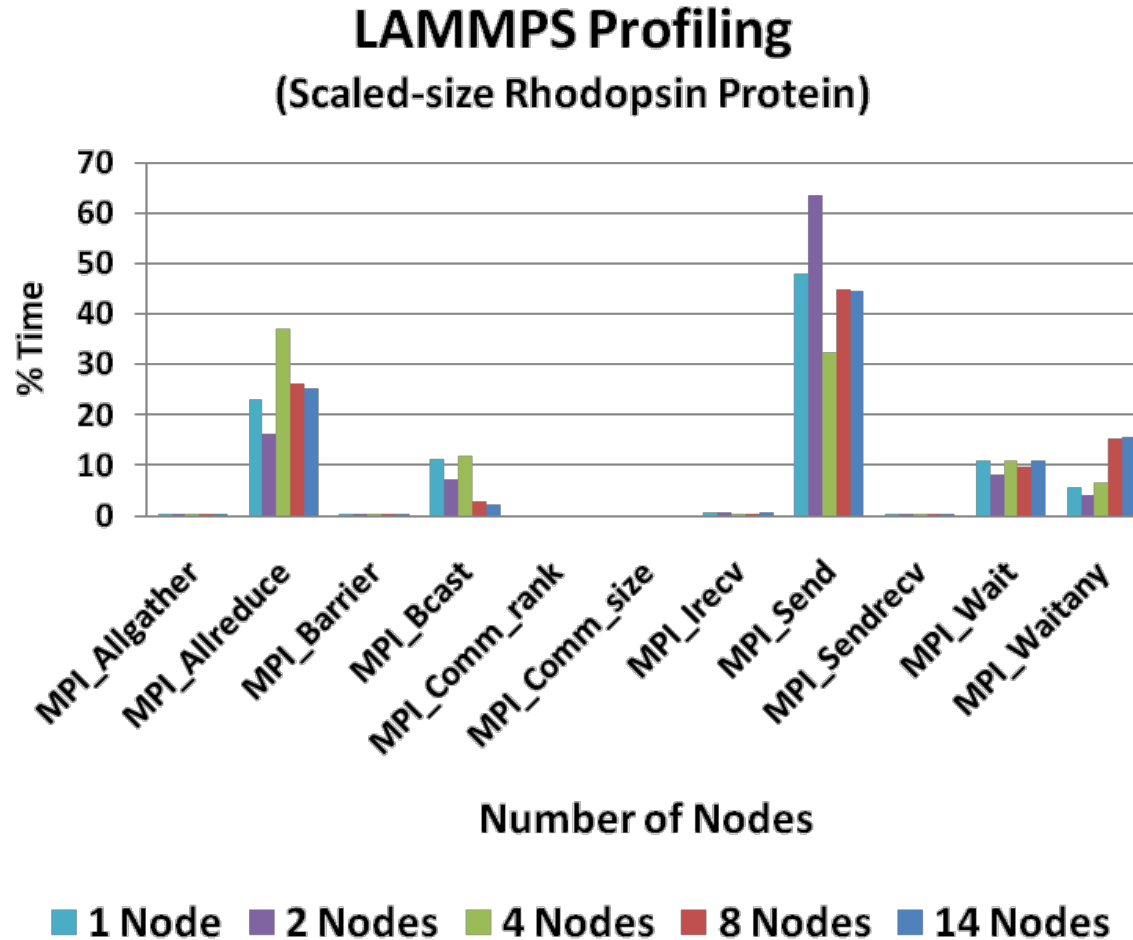
- MPI\_Send, MPI\_Allreduce and MPI\_Waitany the mostly used calls



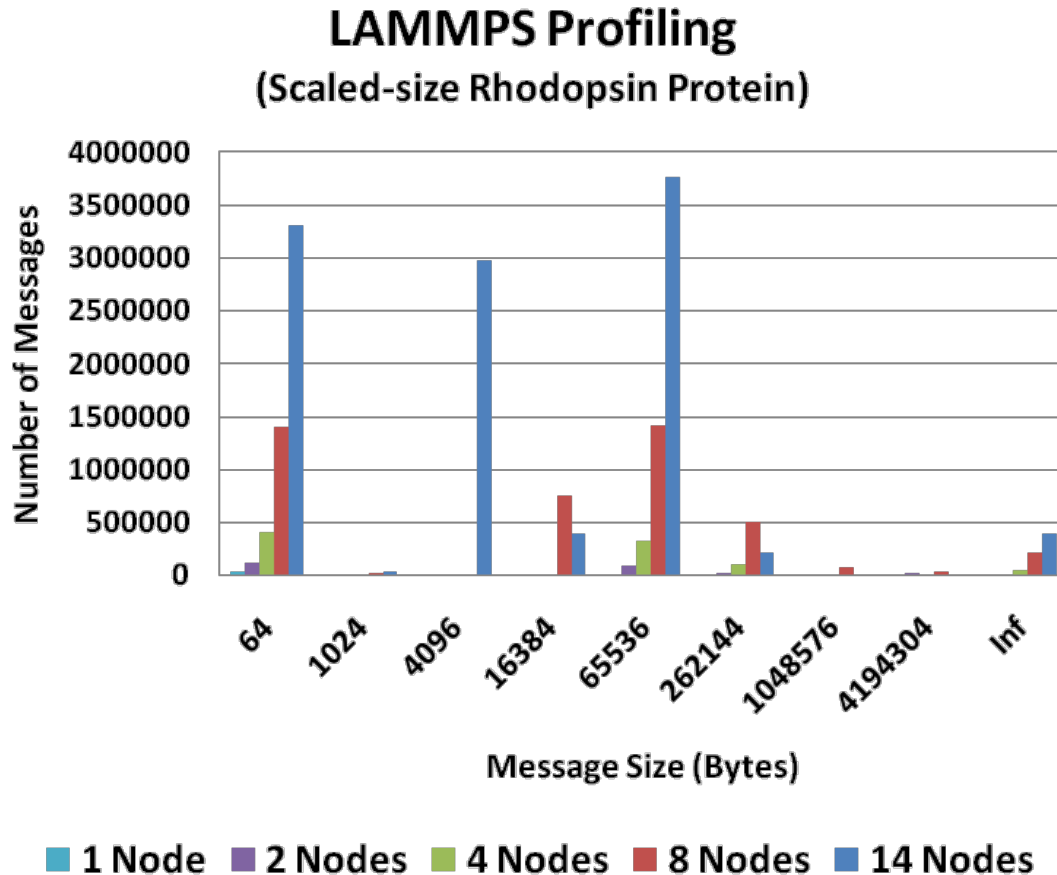
*Higher is better*

*12 Cores/Node*

- **Majority of communication time is spent on MPI\_Send and MPI\_Allreduce**
  - Percentage time is relatively consistent as number of nodes increases



- **Majority of messages are small and medium messages**
  - Messages around 64B and 64KB are mostly used
- **Number of messages increases dramatically with the number of nodes**



- **Interconnects effect to LAMMPS performance**
  - InfiniBand enables higher performance/scalability
  - Ethernet provides only 30% scalability at 14 nodes
- **Intel Compilers and MKL Libraries can increase single-node performance by 46%**
- **Running multiple jobs per day, can increase LAMMPS productivity by 6%**
- **MPI\_Send, MPI\_Allreduce and MPI\_Waitany mostly used MPI calls**
- **Majority of communication time is spent on MPI\_Send and MPI\_Allreduce**
- **Message with sizes around 64B and 64KB mostly used**

# Thank You

## HPC Advisory Council



All trademarks are property of their respective owners. All information is provided "As-Is" without any kind of warranty. The HPC Advisory Council makes no representation to the accuracy and completeness of the information contained herein. HPC Advisory Council Mellanox undertakes no duty and assumes no obligation to update or correct any information presented herein