

NAMD Performance Benchmark and Profiling

August 2010



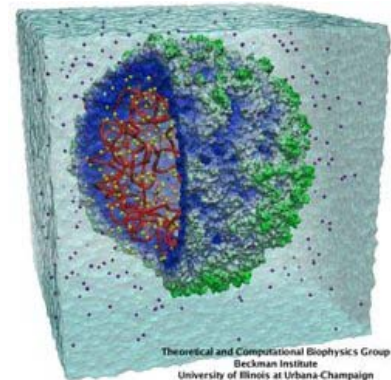
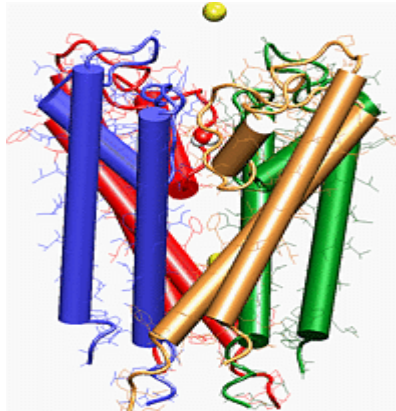
- **The following research was performed under the HPC Advisory Council activities**
 - Participating vendors: Dell, Intel, Mellanox
 - Compute resource - HPC Advisory Council Cluster Center

- **For more info please refer to**
 - <http://www.dell.com>
 - <http://www.intel.com>
 - <http://www.mellanox.com>

- A parallel molecular dynamics code that received the 2002 Gordon Bell Award
- Designed for high-performance simulation of large biomolecular systems
 - **Scales to hundreds of processors and millions of atoms**
- Developed by the joint collaboration of the Theoretical and Computational Biophysics Group (TCB) and the Parallel Programming Laboratory (PPL) at the University of Illinois at Urbana-Champaign
- NAMD is distributed free of charge with source code



Theoretical and Computational Biophysics Group
Beckman Institute
University of Illinois at Urbana-Champaign



Theoretical and Computational Biophysics Group
Beckman Institute
University of Illinois at Urbana-Champaign

- **The presented research was done to provide best practices**
 - NAMD performance benchmarking
 - Interconnect performance comparisons
 - Ways to increase NAMD productivity
 - Power-efficient simulations

- **The presented results will demonstrate**
 - The scalability of the compute environment
 - Considerations for performance optimization

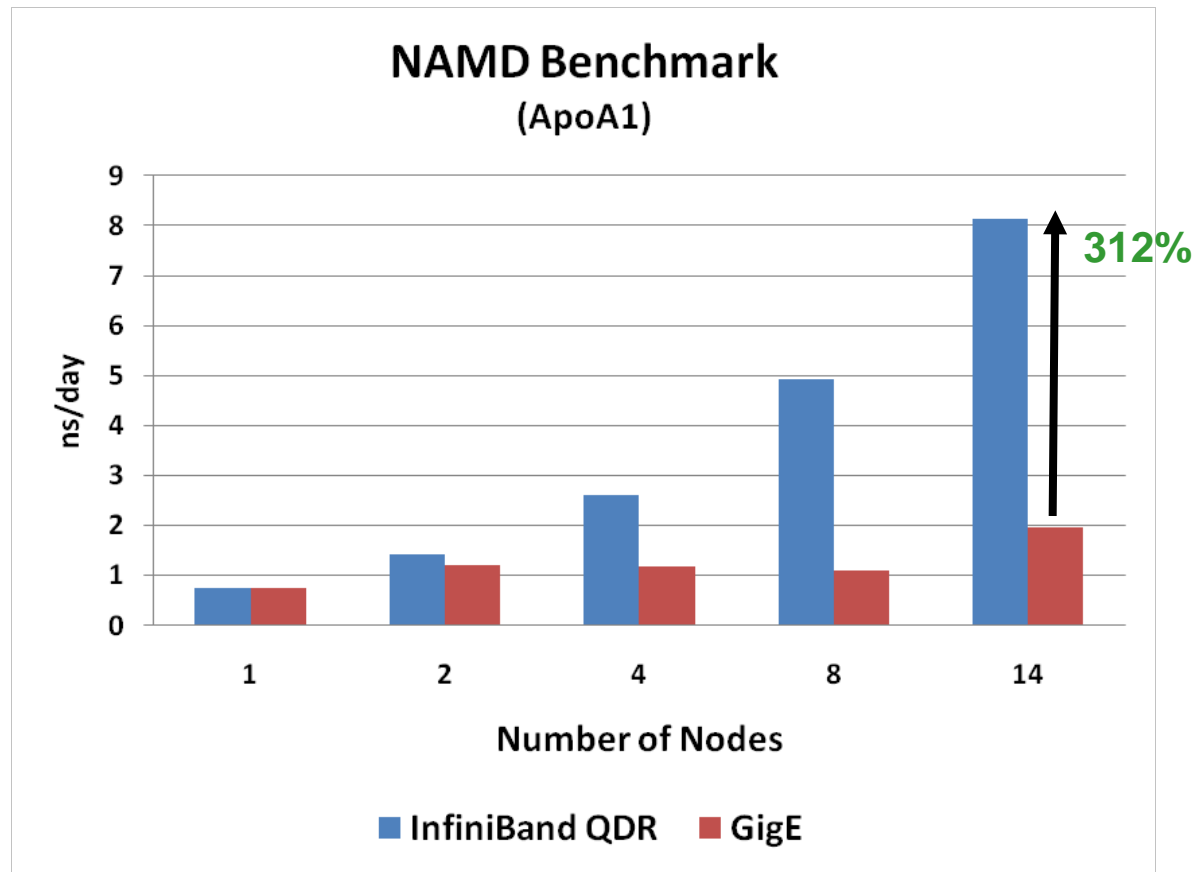
- **Dell™ PowerEdge™ M610 14-node cluster**
 - Six-Core Intel X5670 @ 2.93 GHz CPUs
 - Memory: 24GB per node
 - OS: CentOS5U4, OFED 1.5.1 InfiniBand SW stack
 - Intel Cluster Ready certified cluster
- **Mellanox ConnectX-2 InfiniBand adapters and switches**
- **MPI: MVAPICH2 1.5, Open MPI 1.4.2**
- **Application: NAMD 2.7b3, Charmm-6.2.1, fftw-2.1.3, TCL 8.3**
- **Benchmark Workload**
 - ApoA1 bloodstream lipoprotein particle model (92,224 atoms, 12A cutoff)

- **Intel® Cluster Ready systems make it practical to use a cluster to increase your simulation and modeling productivity**
 - Simplifies selection, deployment, and operation of a cluster
- **A single architecture platform supported by many OEMs, ISVs, cluster provisioning vendors, and interconnect providers**
 - Focus on your work productivity, spend less management time on the cluster
- **Select Intel Cluster Ready**
 - Where the cluster is delivered ready to run
 - Hardware and software are integrated and configured together
 - Applications are registered, validating execution on the Intel Cluster Ready architecture
 - Includes Intel® Cluster Checker tool, to verify functionality and periodically check cluster health.



- **InfiniBand enables higher scalability**

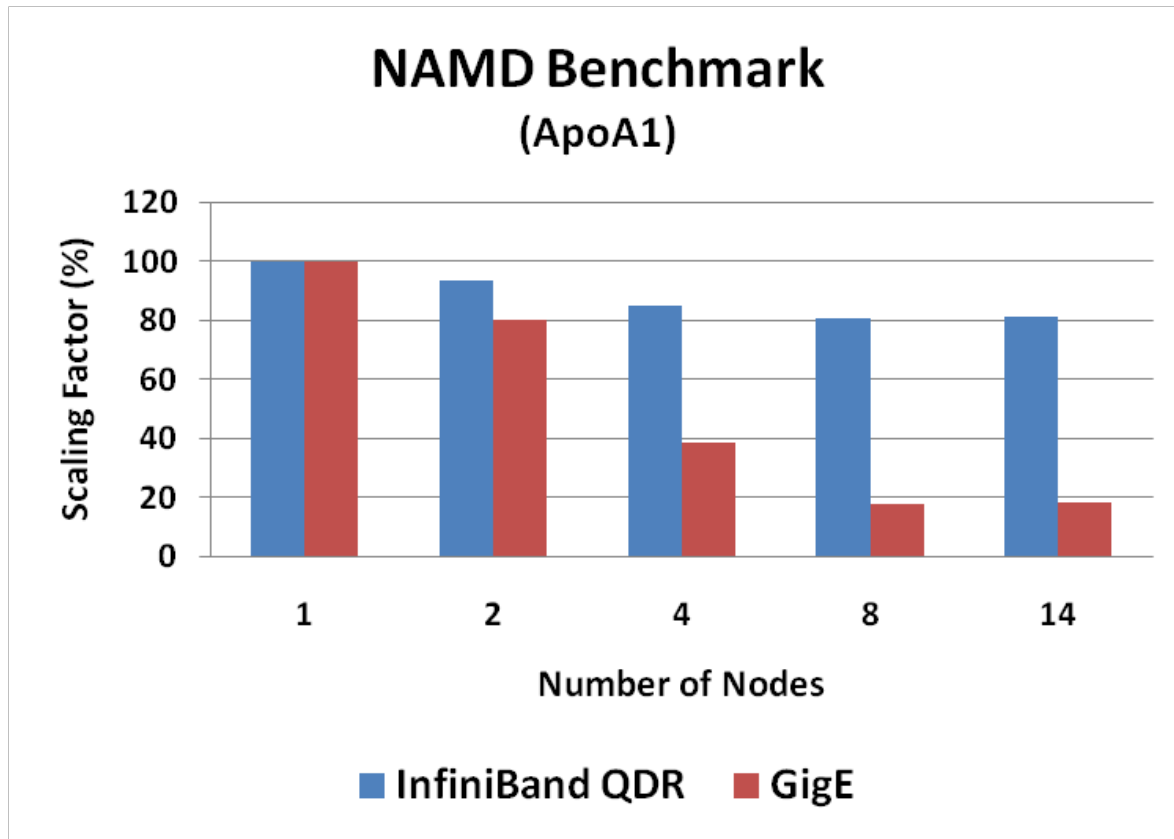
- Up to 312% higher performance than Ethernet at 14 nodes
- Four InfiniBand connected servers deliver higher performance vs 14 Ethernet connected servers



Higher is better

12 Cores/Node

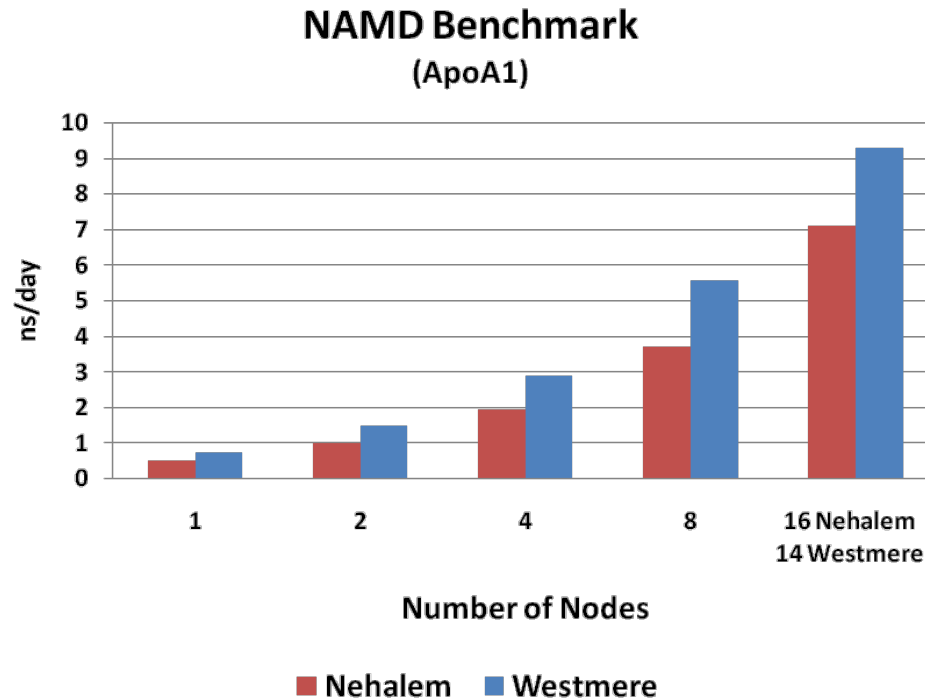
- InfiniBand enables higher scalability – more than 80% at 14 nodes
- Only 20% of the system compute capability can be utilized with Ethernet



Higher is better

12 Cores/Node

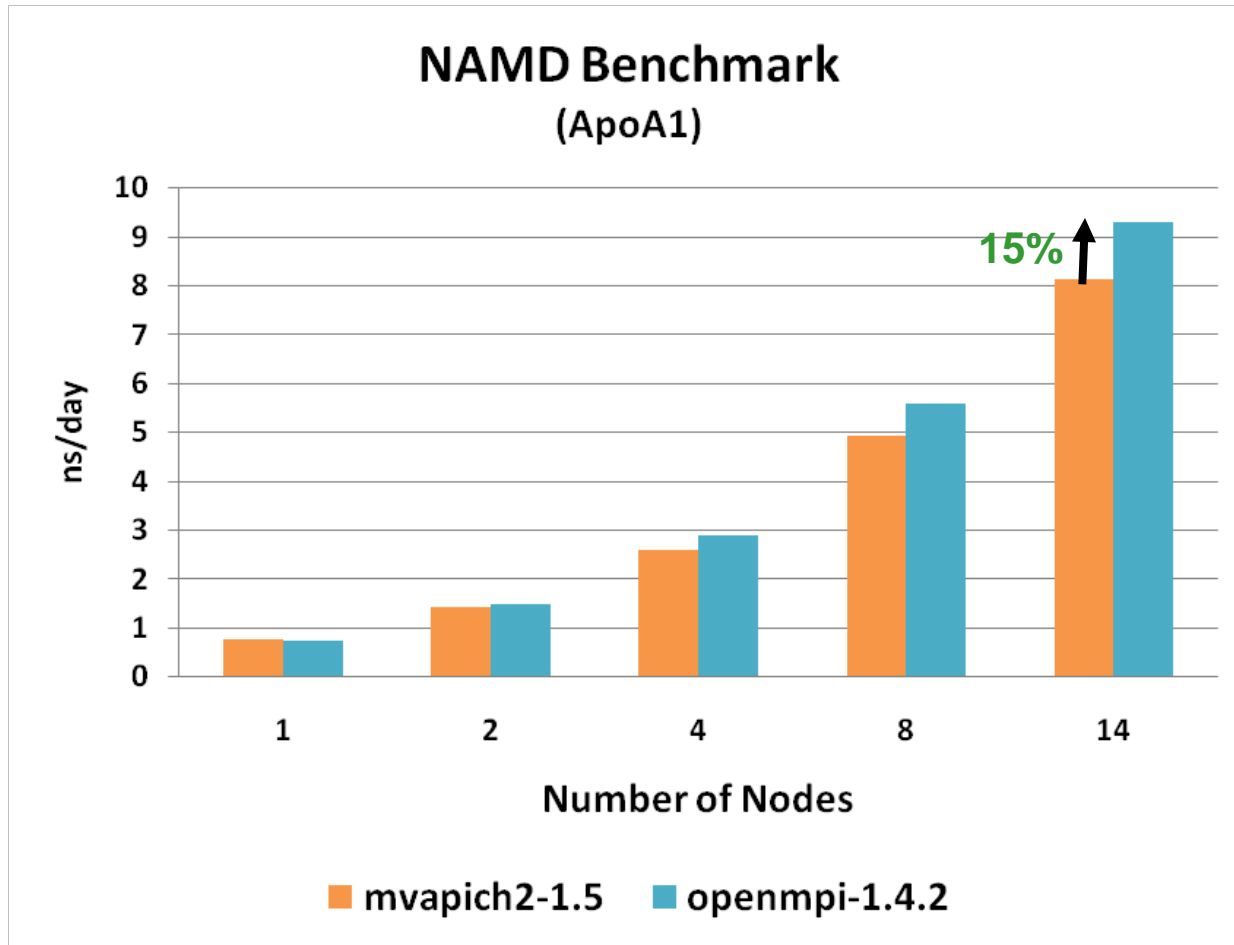
- Intel X5670 (“Westmere”) provides ~50% higher performance versus Intel X5570 (“Nehalem”)
 - Both are the same speed – 2.9GHz
 - InfiniBand QDR as the interconnect solution for both cases
 - Performance comparison per node (server) with dual sockets



Higher is better

12 Cores/Node

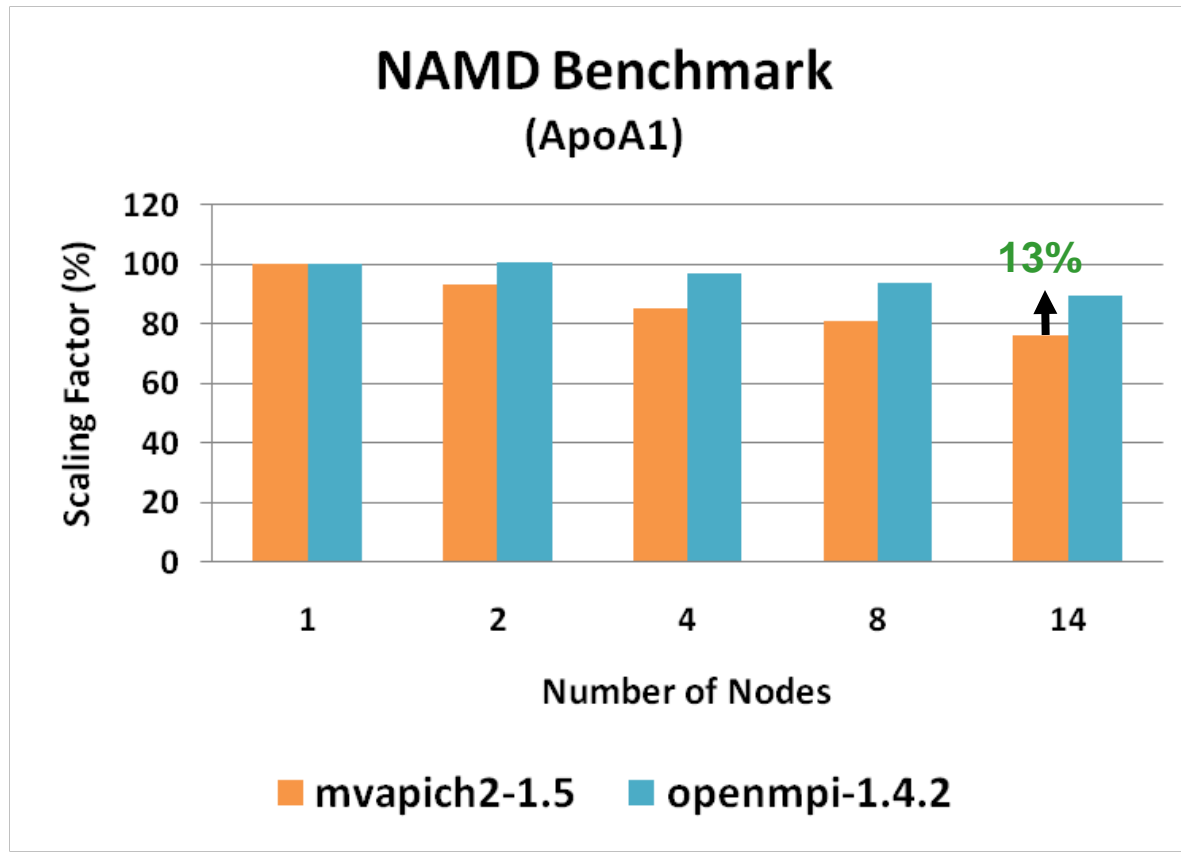
- **Open MPI shows higher performance over MVAPUCH2**
 - Up to 15% gain at 14 nodes



Higher is better

12 Cores/Node

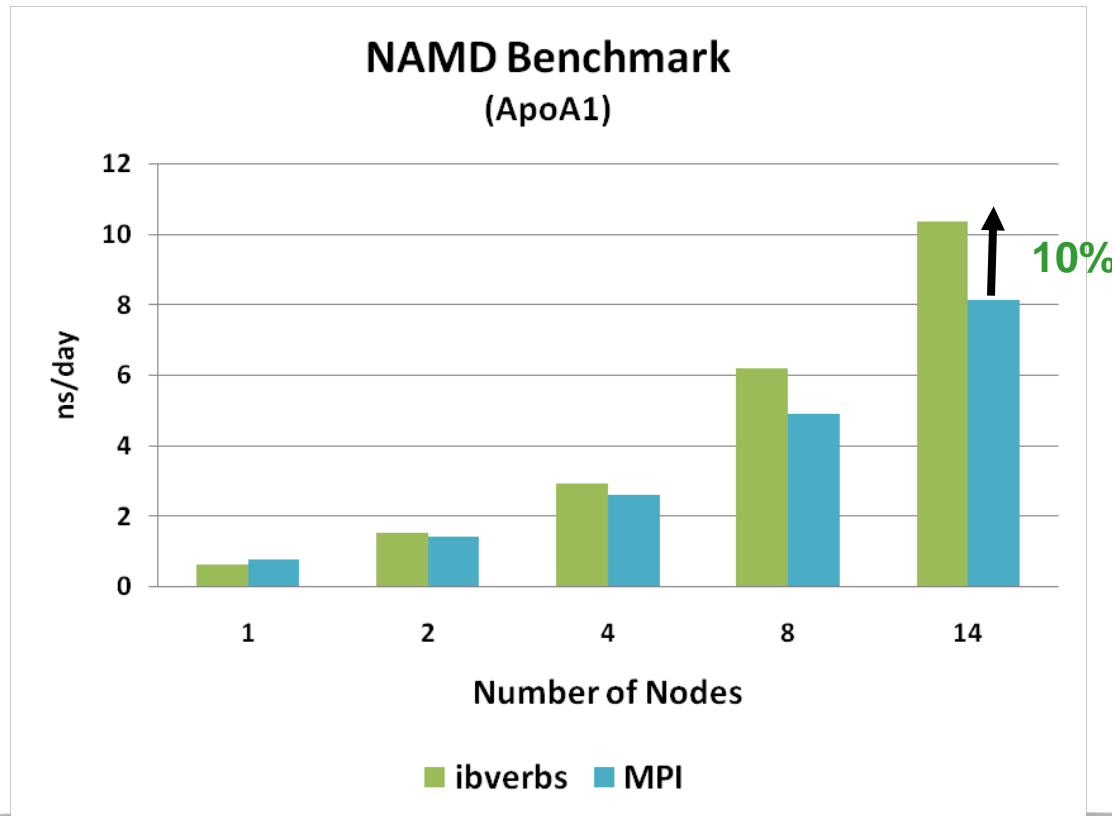
- **Open MPI demonstrates better scalability**
 - Up to 13% better at 14 nodes



Higher is better

12 Cores/Node

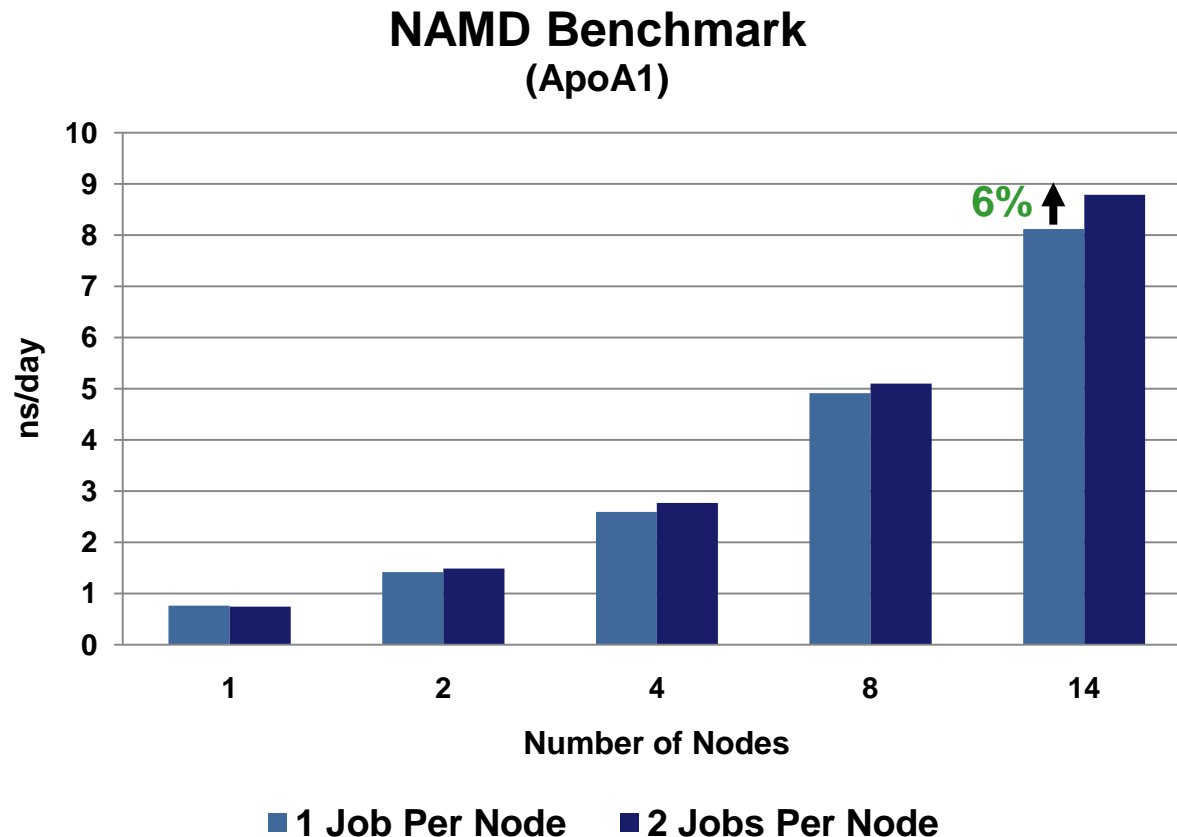
- NAMD can use MPI or the native InfiniBand interface (verbs) as the interface to the interconnect (InfiniBand)
- IB verbs version provides better productivity versus MPI
 - 10% improvement over 14 nodes
 - IB verbs provides a lower level interface to the interconnect versus MPI – lower overhead



Higher is better

12 Cores/Node

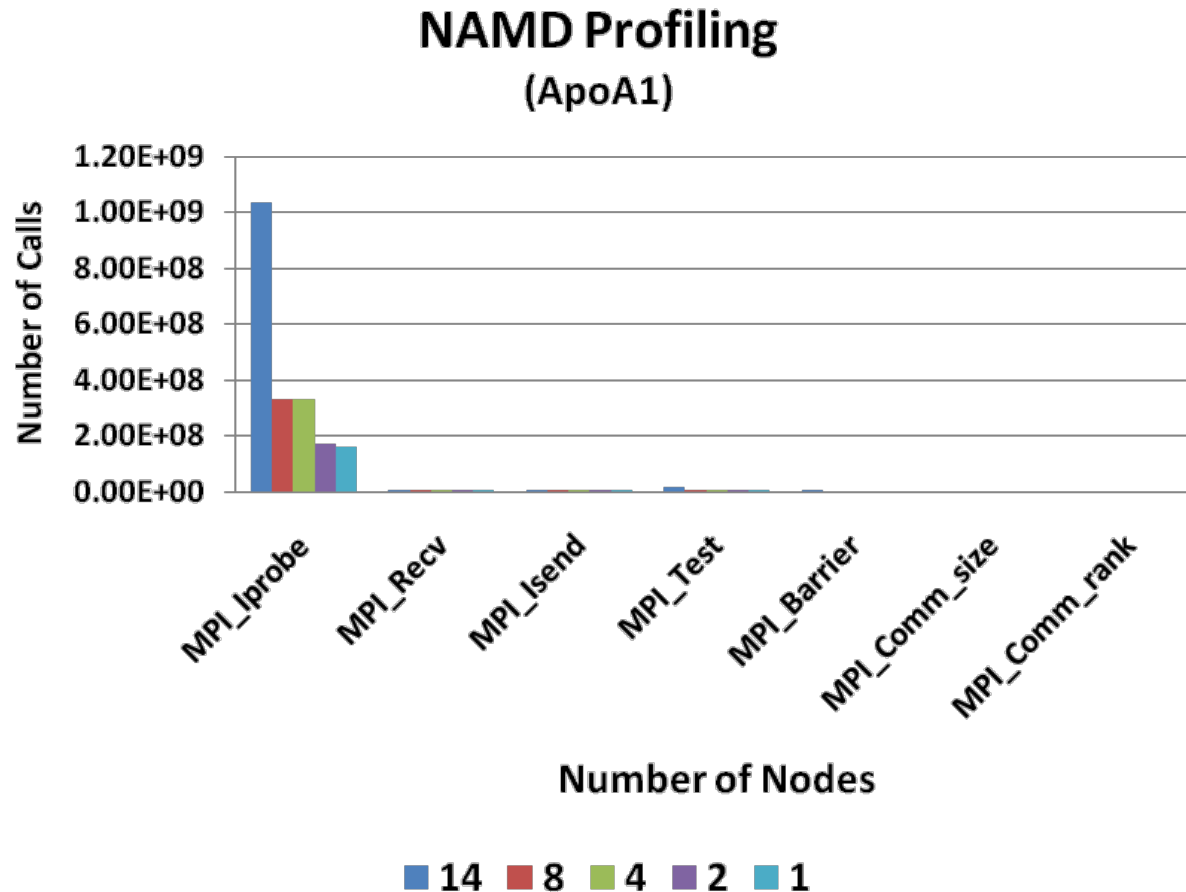
- **Running 2 jobs concurrently can provide slightly higher productivity**
 - 2 jobs setup: 6 cores/node per job; 1 job setup: 12 cores/node
 - Up to 6% higher performance when 2 jobs running concurrently



Higher is better

12 Cores/Node

- **MPI_Iprobe has the most number of calls**
 - Number of calls increases dramatically with cluster size

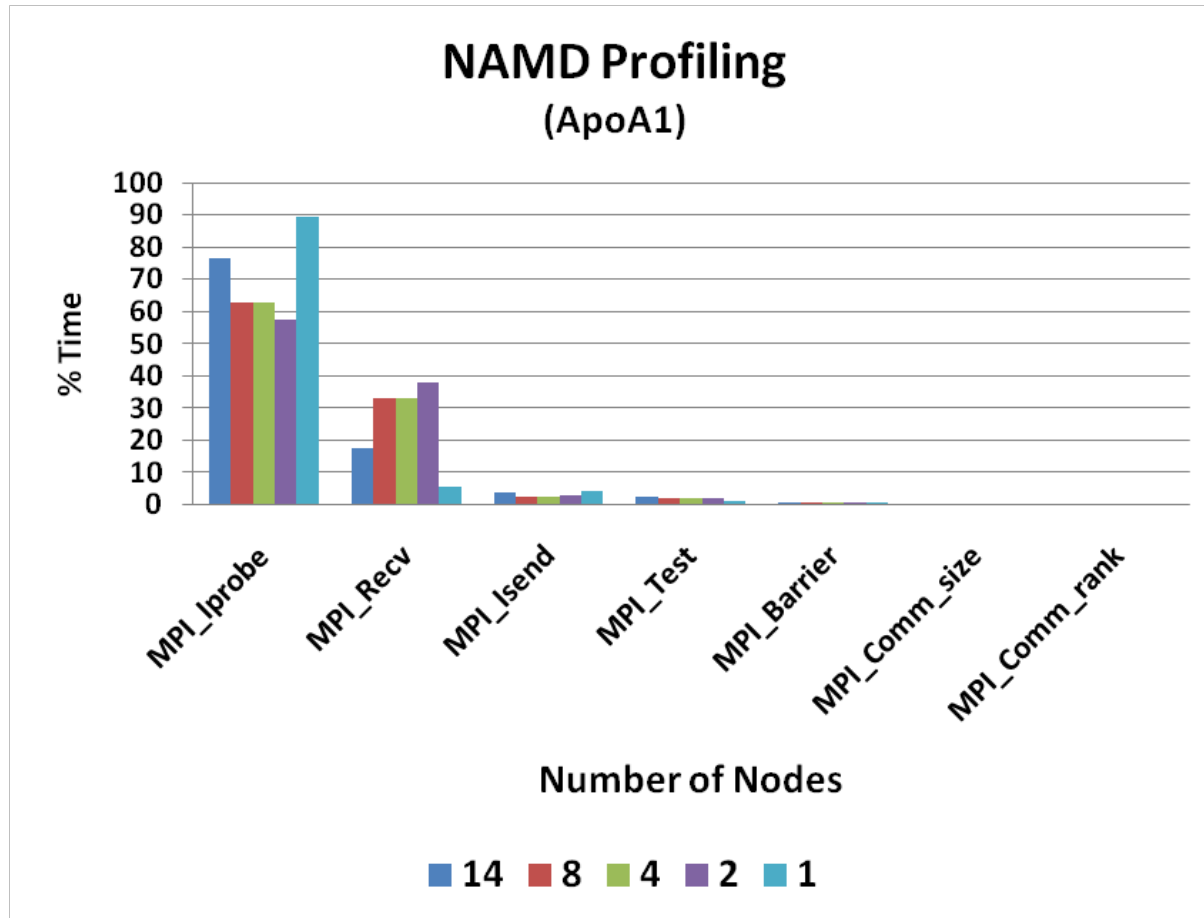


Higher is better

12 Cores/Node

NAMD Profiling Result – % Time of MPI Calls

- **Majority of communication time is spent on MPI_Iprobe**
 - Percentage time are relatively consistent as number of nodes increases



Higher is better

12 Cores/Node

- **Interconnects effect to NAMD performance**
 - InfiniBand enables higher performance/scalability
 - For same productivity (number of jobs per day) InfiniBand requires 20% of the energy needed when using Ethernet
- **Intel Westmere delivers 50% higher performance per node versus Nehalem**
 - Aligned with the increase in the number of cores per socket (6 versus 4)
- **Using lower level interfaces to the network, can increase performance in 10%**
- **MPI libraries and running multiple jobs per day, can increase NAMD productivity from 6% to 13%**

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