HPC ADVISORY COUNCIL

NWChem Performance Benchmark and Profiling

October 2010

DELL intel Mellanox TECHNOLOGIES

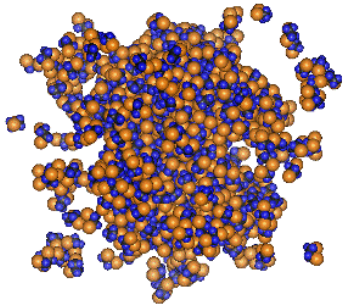
Note

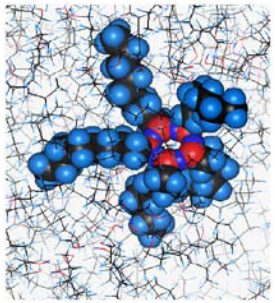
- **The following research was performed under the HPC Advisory Council activities**
  - Participating vendors: Intel, Dell, Mellanox
  - Compute resource - HPC Advisory Council Cluster Center

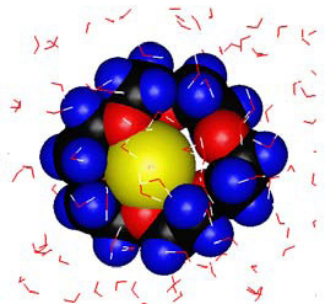
- **For more info please refer to**
  - http://www.dell.com
  - http://www.intel.com
  - http://www.mellanox.com
  - http://www.nwchem-sw.org

NWChem

- **NWChem is a computational chemistry package**
  - NWChem has been developed by the Molecular Sciences Software group of the Environmental Molecular Sciences Laboratory (EMSL) at the Pacific Northwest National Laboratory (PNNL)
- **NWChem provides many methods to compute the properties of molecular and periodic systems**
  - Using standard quantum mechanical descriptions of the electronic wave function or density
- **NWChem has the capability to perform classical molecular dynamics and free energy simulations**
  - These approaches may be combined to perform mixed quantum-mechanics and molecular-mechanics simulations







Objectives

- **The following was done to provide best practices**
  - NWChem performance benchmarking
  - Interconnect performance comparisons
  - Understanding NWChem 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 NWChem to achieve scalable productivity
  - Considerations for power saving through balanced system configuration

Test Cluster Configuration

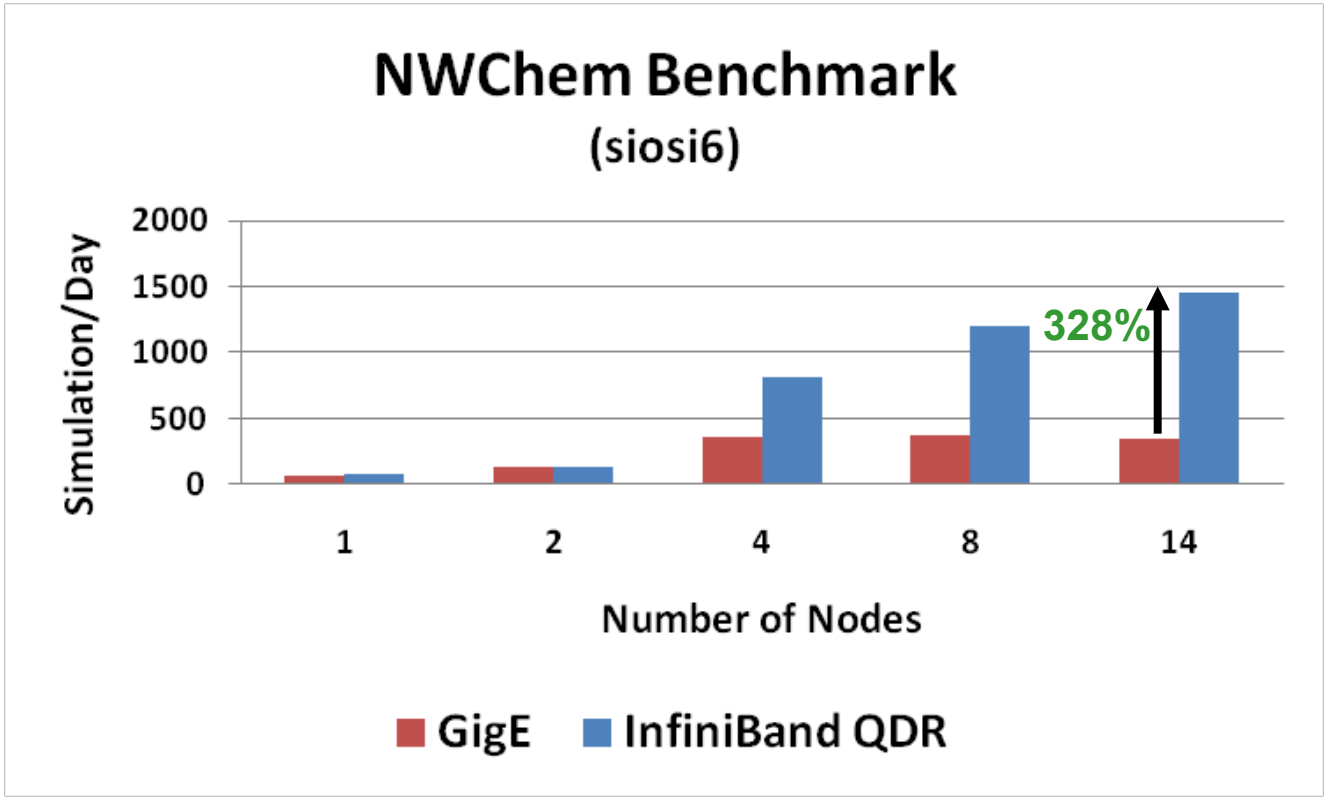
- **Dell™ PowerEdge™ M610 14-node cluster**
  - Six-Core Intel X5670 @ 2.93 GHz CPUs
  - Memory: 24GB memory, DDR3 1333 MHz
  - OS: CentOS5U4, OFED 1.5.1 InfiniBand SW stack
- **Intel Cluster Ready certified cluster**
- **Mellanox ConnectX-2 InfiniBand adapters and switches**
- **MPI: Intel MPI 4, MVAPICH2 1.5.1p1, Platform MPI 7.1**
- **Compiler: Intel Compilers 11.1 build 073**
- **Application: NWChem 6.0**
- **Benchmarks:**
  - siosi6: LDA calculations of 3 zeolite fragments (347,1687,3554) ($Si_{28}O_{67}H_{30}$)

About Intel® Cluster Ready

- **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

NWChem Performance – Interconnect (siosi6)

- **InfiniBand enables higher scalability**
  - Up to 328% higher performance than Ethernet at 14 nodes
  - Dramatic improvement seen as communication pattern shifts starting from 4 nodes

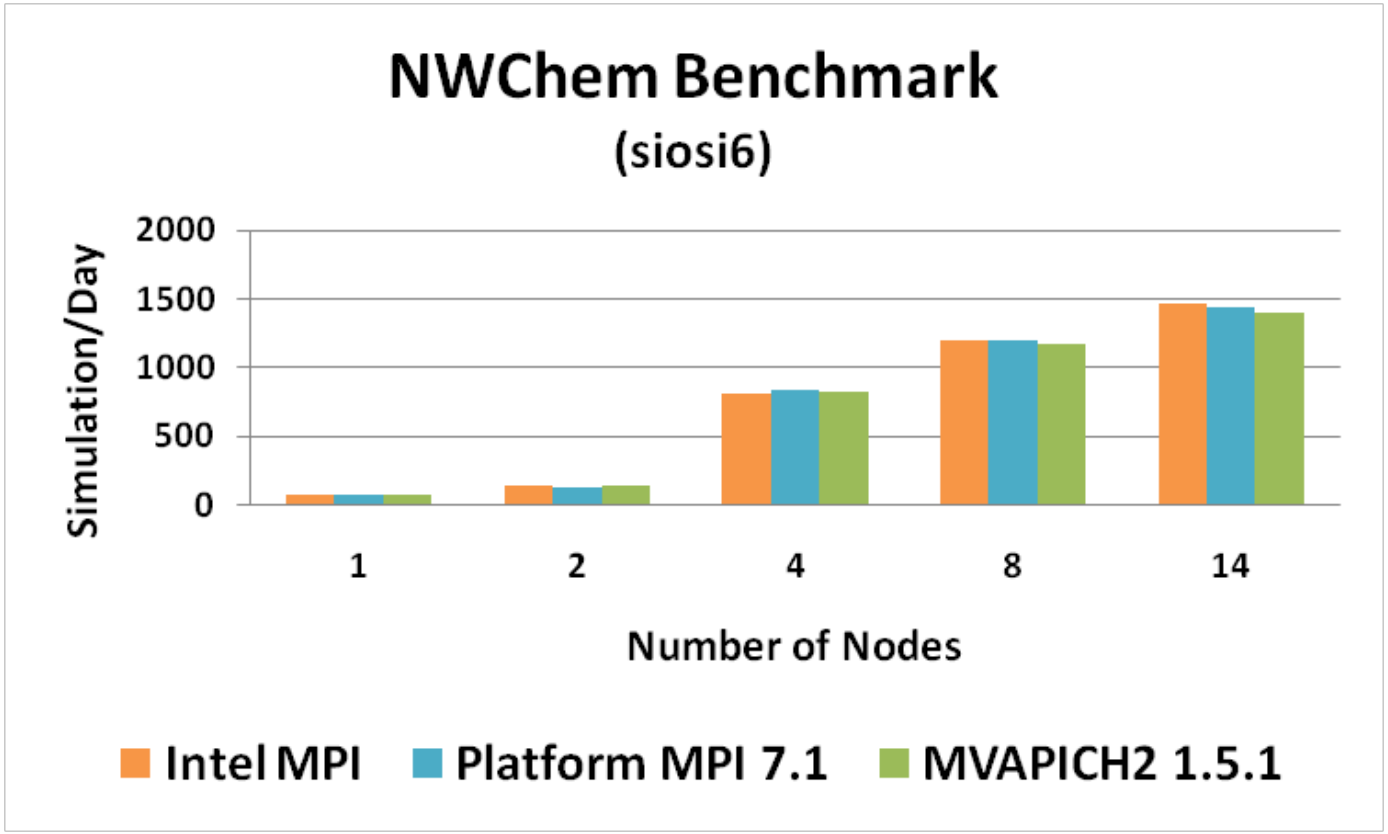


Higher is better

12 Cores/Node

NWChem Performance – MPI (siosi6)

- **All MPI implementations tested demonstrate relatively good performance**

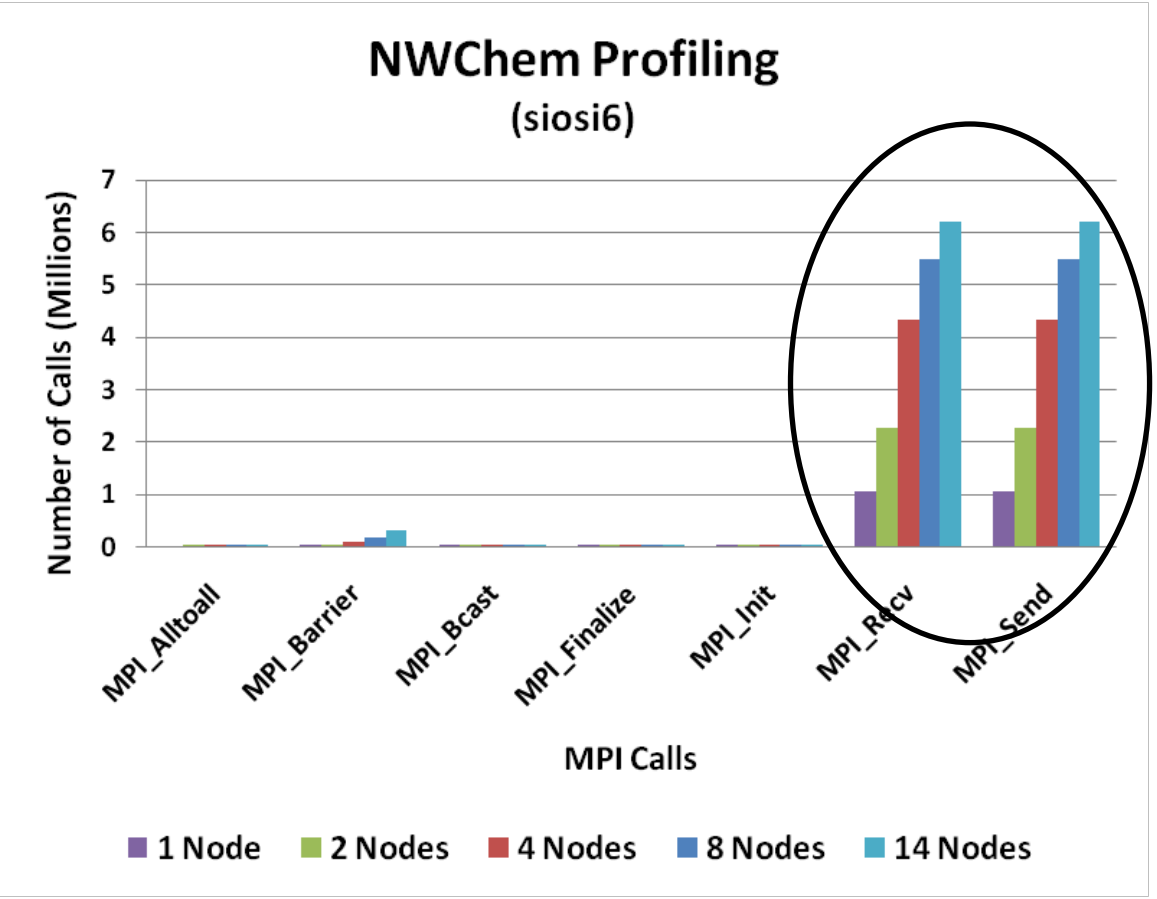


Higher is better

12 Cores/Node

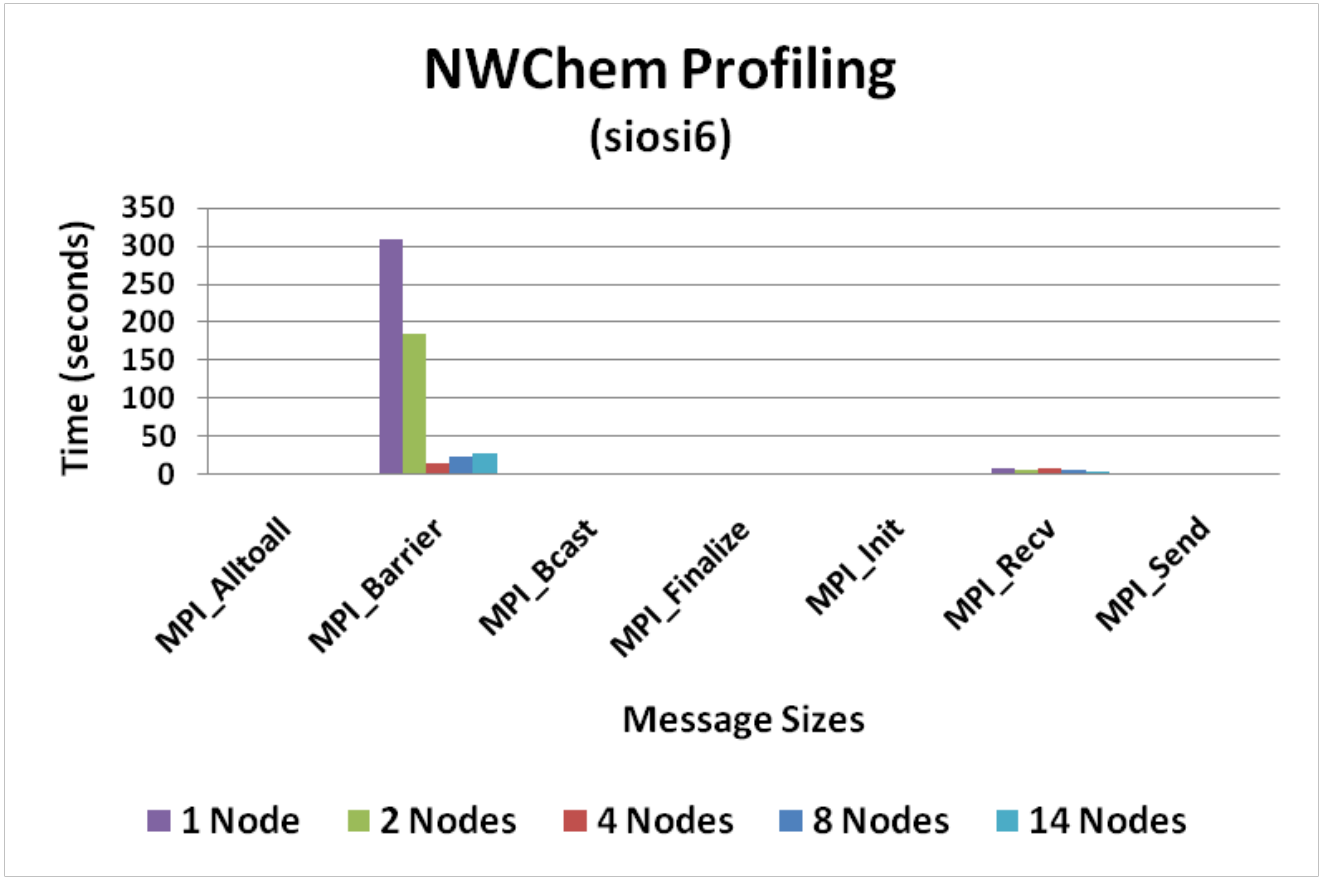
NWChem Profiling – # of MPI Calls (siosi6)

- **The most used MPI functions for this dataset are**
  - MPI_Send and MPI_Recv
- **This dataset drives more communications between processes**
  - MPI communications accelerates starting at 4 nodes



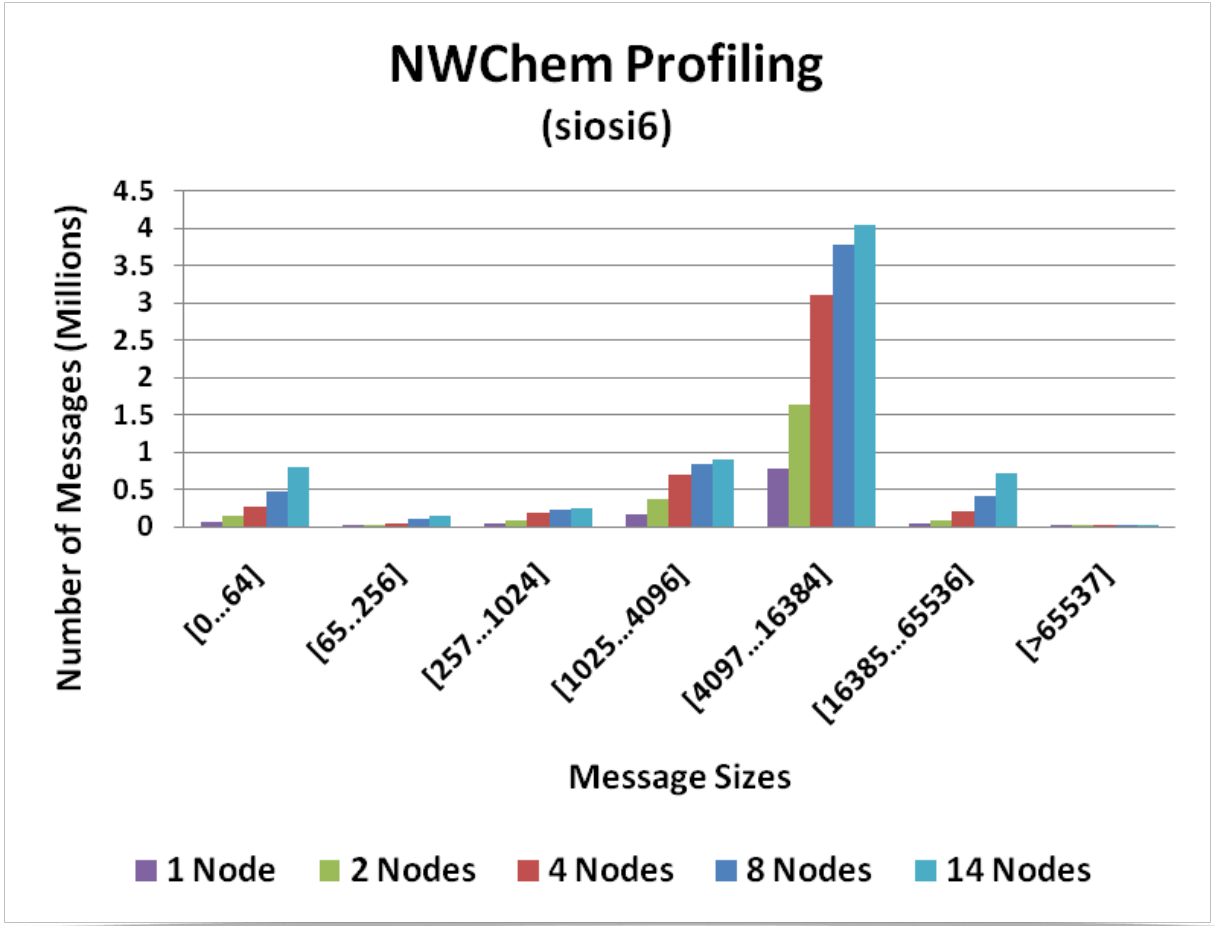
NWChem Profiling – Time of MPI Calls (siosi6)

- **Majority of time is spent on MPI_Barrier**
  - Sudden drop for synchronization time at 4 nodes



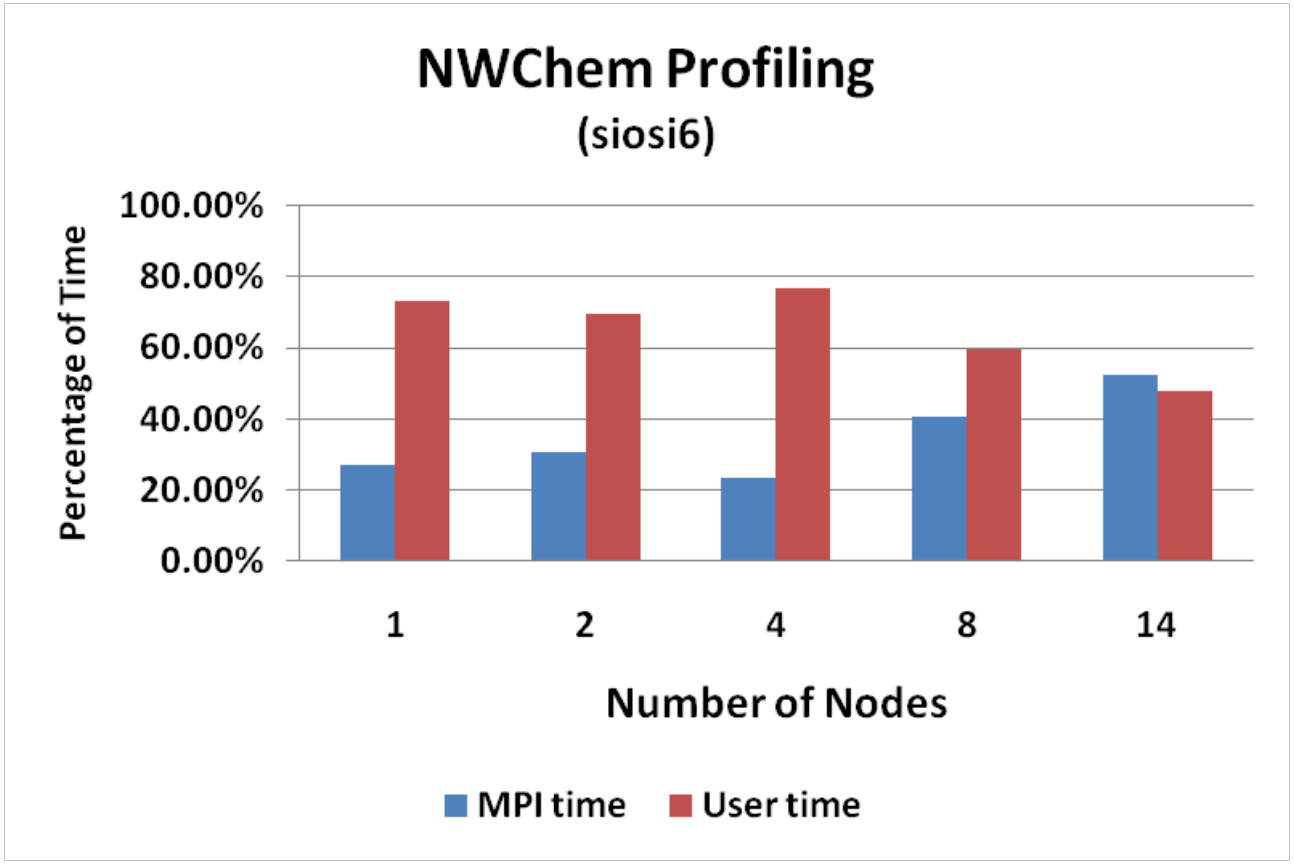
NWChem Profiling – # of Message Sizes (siosi6)

- **Majority of messages are small and medium messages**
  - Messages between 4K and 16K are mostly used for any node sizes
- **Number of messages increases proportionally with the number of nodes**



NWChem Profiling - MPI/user Time Ratio (siosi6)

- **Percentage of MPI Communication increases as cluster scales**



12 Cores/Node

Summary

- **Interconnects effect to NWChem performance**
  - InfiniBand enables higher performance/scalability
  - Up to 328% higher performance than Ethernet at 14 nodes
- **MPI_Send, MPI_Recv and MPI_Barrier mostly used MPI calls**
- **Majority of communication time is spent on MPI_Barrier**
- **Message sizes**
  - Siosi6 has the most MPI messages between 4K and 16KB

Thank You

HPC Advisory Council