

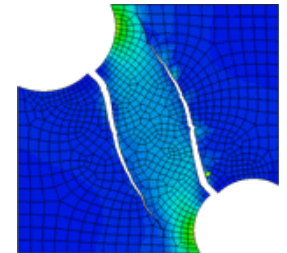
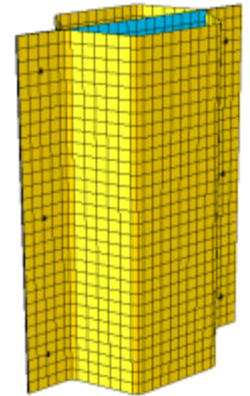
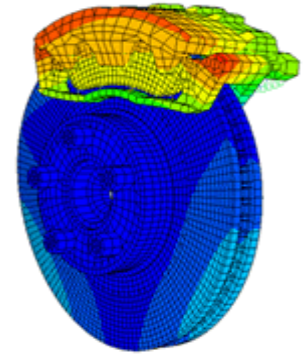
Abaqus Performance Benchmark and Profiling

December 2009



- **The following research was performed under the HPC Advisory Council activities**
 - Participating vendors: AMD, Dell, SIMULIA, Mellanox
 - Compute resource - HPC Advisory Council Cluster Center
- **The participating members would like to thank SIMULIA for their support and guidelines**
- **For more info please refer to**
 - www.mellanox.com, www.dell.com/hpc, www.amd.com
 - <http://www.simulia.com>

- **ABAQUS offers a suite of engineering design analysis software products, including tools for:**
 - Nonlinear finite element analysis (FEA)
 - Advanced linear and dynamics application problems
- **ABAQUS/Standard provides general-purpose FEA that includes a broad range of analysis capabilities**
- **ABAQUS/Explicit provides nonlinear, transient, dynamic analysis of solids and structures using explicit time integration**

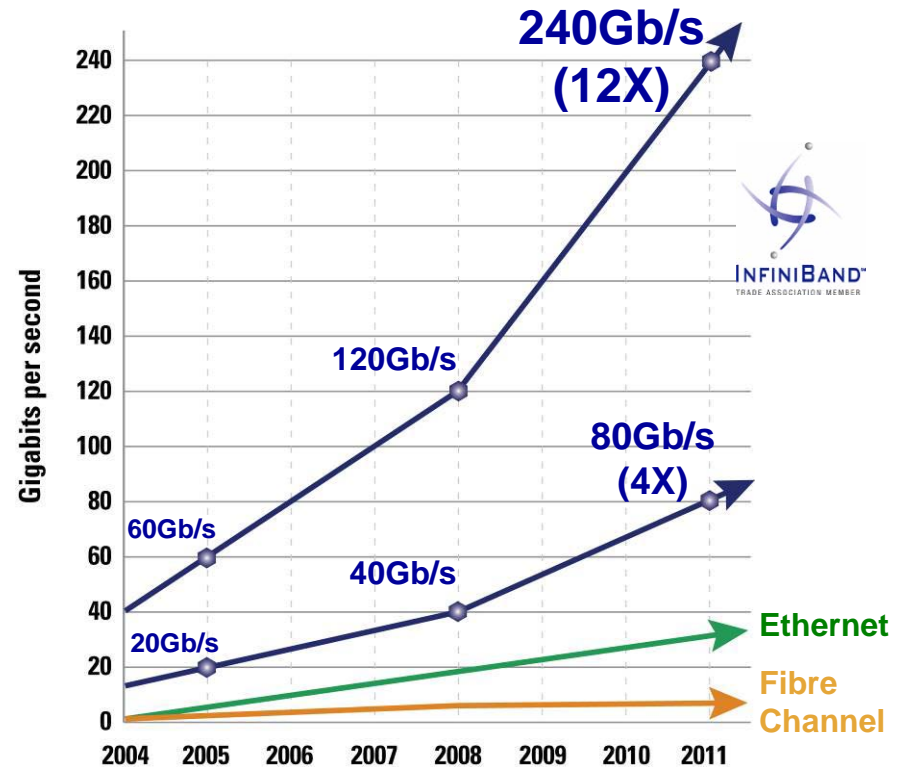


- **The presented research was done to provide best practices**
 - Abaqus performance benchmarking
 - Interconnect performance comparisons
 - Understanding Abaqus communication patterns
 - Power-efficient simulations
- **The presented results will demonstrate**
 - The scalability of the compute environment to provide good application scalability
 - Considerations for power saving through balanced system configuration

- **Dell™ PowerEdge™ SC 1435 24-node cluster**
- **Quad-Core AMD Opteron™ 2382 (“Shanghai”) CPUs**
- **Mellanox® InfiniBand ConnectX® 20Gb/s (DDR) HCAs**
- **Mellanox® InfiniBand DDR Switch**
- **Memory: 16GB memory, DDR2 800MHz per node**
- **OS: RHEL5U3, OFED 1.4.1 InfiniBand SW stack**
- **MPI: HP-MPI 2.3**
- **Application: Abaqus 6.9 EF1**
- **Benchmark Workload**
 - **Abaqus/Standard Server Benchmarks: S2A**
 - **Abaqus/Explicit Server Benchmarks: E2**

- **Industry Standard**
 - Hardware, software, cabling, management
 - Design for clustering and storage interconnect
- **Performance**
 - 40Gb/s node-to-node
 - 120Gb/s switch-to-switch
 - 1us application latency
 - Most aggressive roadmap in the industry
- **Reliable with congestion management**
- **Efficient**
 - RDMA and Transport Offload
 - Kernel bypass
 - CPU focuses on application processing
- **Scalable for Petascale computing & beyond**
- **End-to-end quality of service**
- **Virtualization acceleration**
- **I/O consolidation including storage**

The InfiniBand Performance Gap is Increasing



InfiniBand Delivers the Lowest Latency

Quad-Core AMD Opteron™ Processor

- **Performance**

- Quad-Core

- Enhanced CPU IPC
- 4x 512K L2 cache
- 6MB L3 Cache

- Direct Connect Architecture

- HyperTransport™ Technology
- Up to 24 GB/s peak per processor

- Floating Point

- 128-bit FPU per core
- 4 FLOPS/clock peak per core

- Integrated Memory Controller

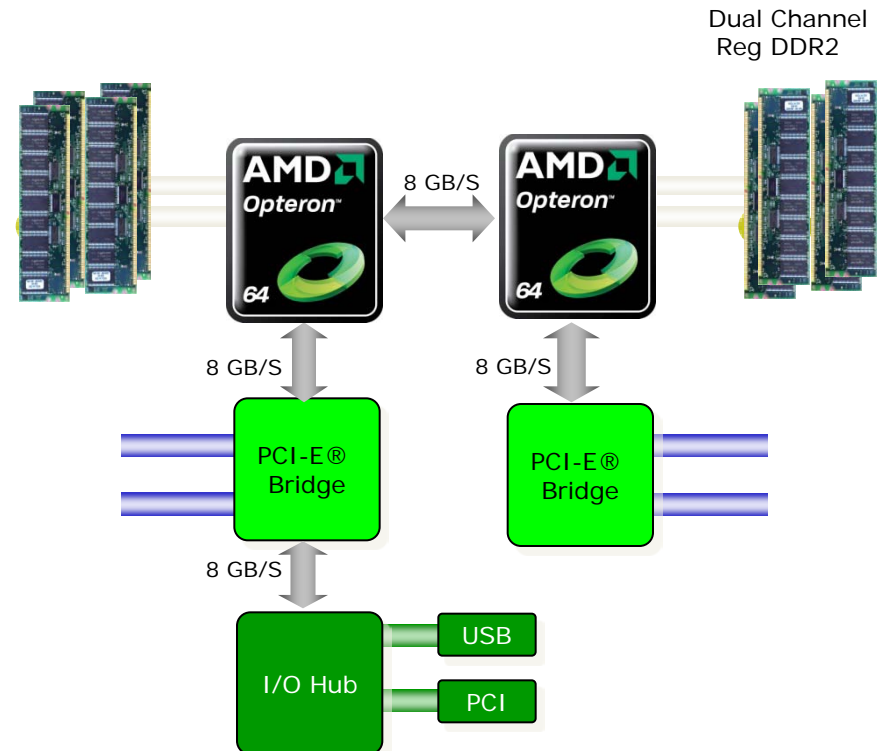
- Up to 12.8 GB/s
- DDR2-800 MHz or DDR2-667 MHz

- **Scalability**

- 48-bit Physical Addressing

- **Compatibility**

- Same power/thermal envelopes as 2nd / 3rd generation AMD Opteron™ processor



- **System Structure and Sizing Guidelines**

- 24-node cluster build with Dell PowerEdge™ SC 1435 Servers
- Servers optimized for High Performance Computing environments
- Building Block Foundations for best price/performance and performance/watt

- **Dell HPC Solutions**

- Scalable Architectures for High Performance and Productivity
- Dell's comprehensive HPC services help manage the lifecycle requirements.
- Integrated, Tested and Validated Architectures

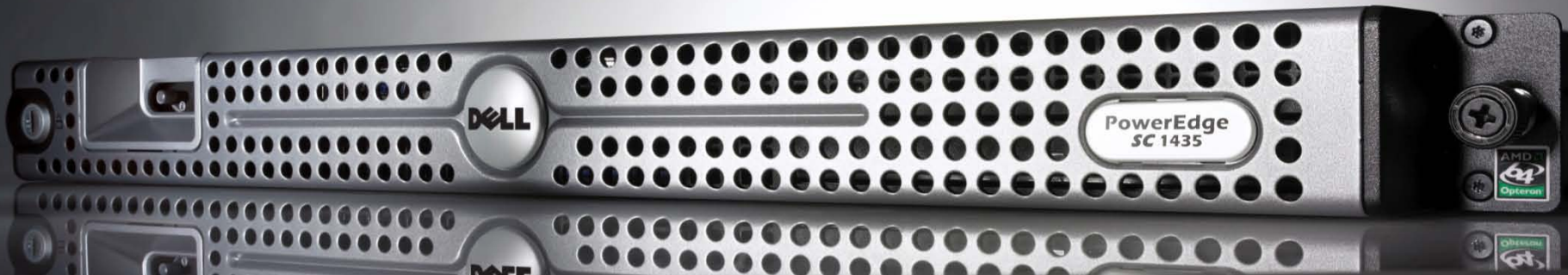
- **Workload Modeling**

- Optimized System Size, Configuration and Workloads
- Test-bed Benchmarks
- ISV Applications Characterization
- Best Practices & Usage Analysis



Dell PowerEdge™ Server Advantage

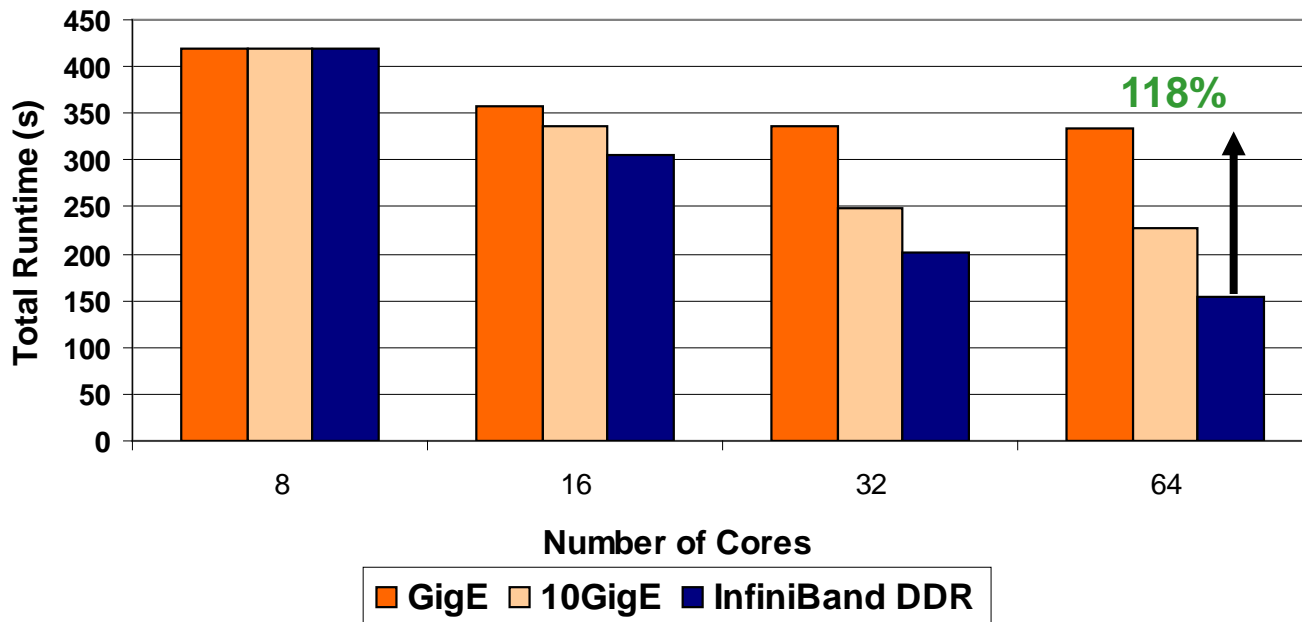
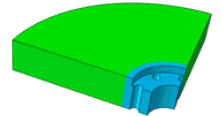
- Dell™ PowerEdge™ servers incorporate AMD Opteron™ and Mellanox ConnectX InfiniBand to provide leading edge performance and reliability
- Building Block Foundations for best price/performance and performance/watt
- Investment protection and energy efficient
- Longer term server investment value
- Faster DDR2-800 memory
- Enhanced AMD PowerNow!
- Independent Dynamic Core Technology
- AMD CoolCore™ and Smart Fetch Technology
- Mellanox InfiniBand end-to-end for highest networking performance



Abaqus/Standard Benchmark Results

- **Input Dataset: S2A**
 - Flywheel with centrifugal load, direct solver
- **InfiniBand provides higher utilization, performance and scalability**
 - Up to 118% higher performance versus GigE and 49% higher than 10GigE

**Abaqus/Standard Server Benchmark
(S2A)**



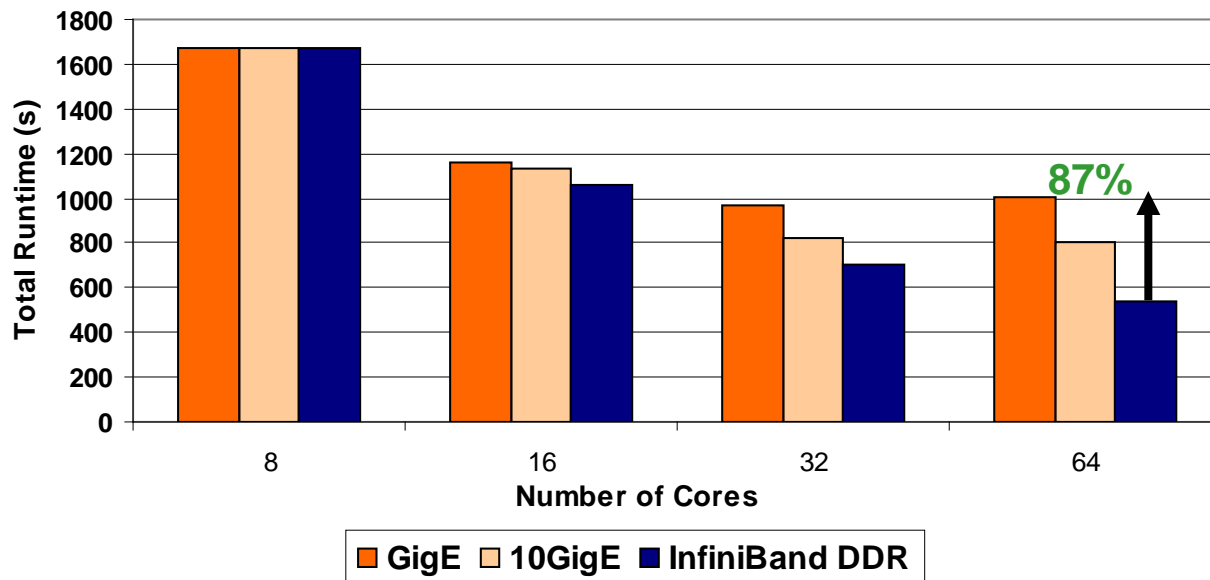
Lower is better

8-cores per node

Abaqus/Explicit Benchmark Results

- **Input Dataset: E2**
 - Cell phone impacting a fixed rigid floor
- **InfiniBand provides higher utilization, performance and scalability**
 - Up to 87% higher performance versus GigE and 50% higher than 10GigE

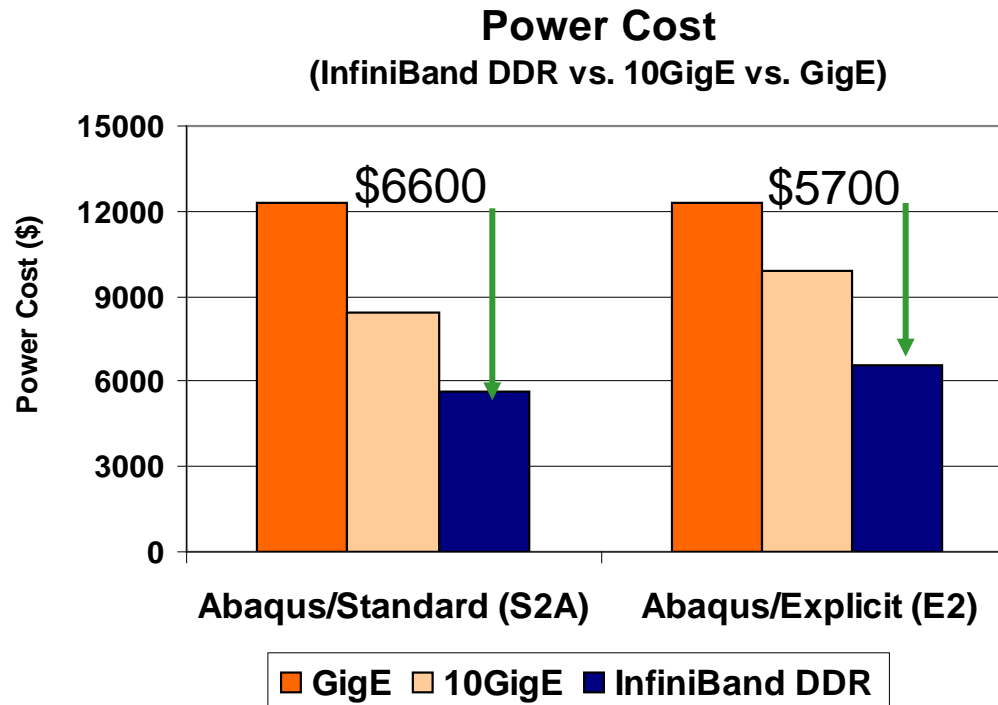
**Abaqus/Explicit Server Benchmark
(E2)**



Lower is better

8-cores per node

- **Dell economical integration of AMD CPUs and Mellanox InfiniBand**
 - Saves power up to \$6600 to achieve same number of application jobs over GigE
 - Up to \$3800 to achieve same number of application jobs with 10GigE
 - Yearly based for 24-node cluster
- **As cluster size increases, more power can be saved**



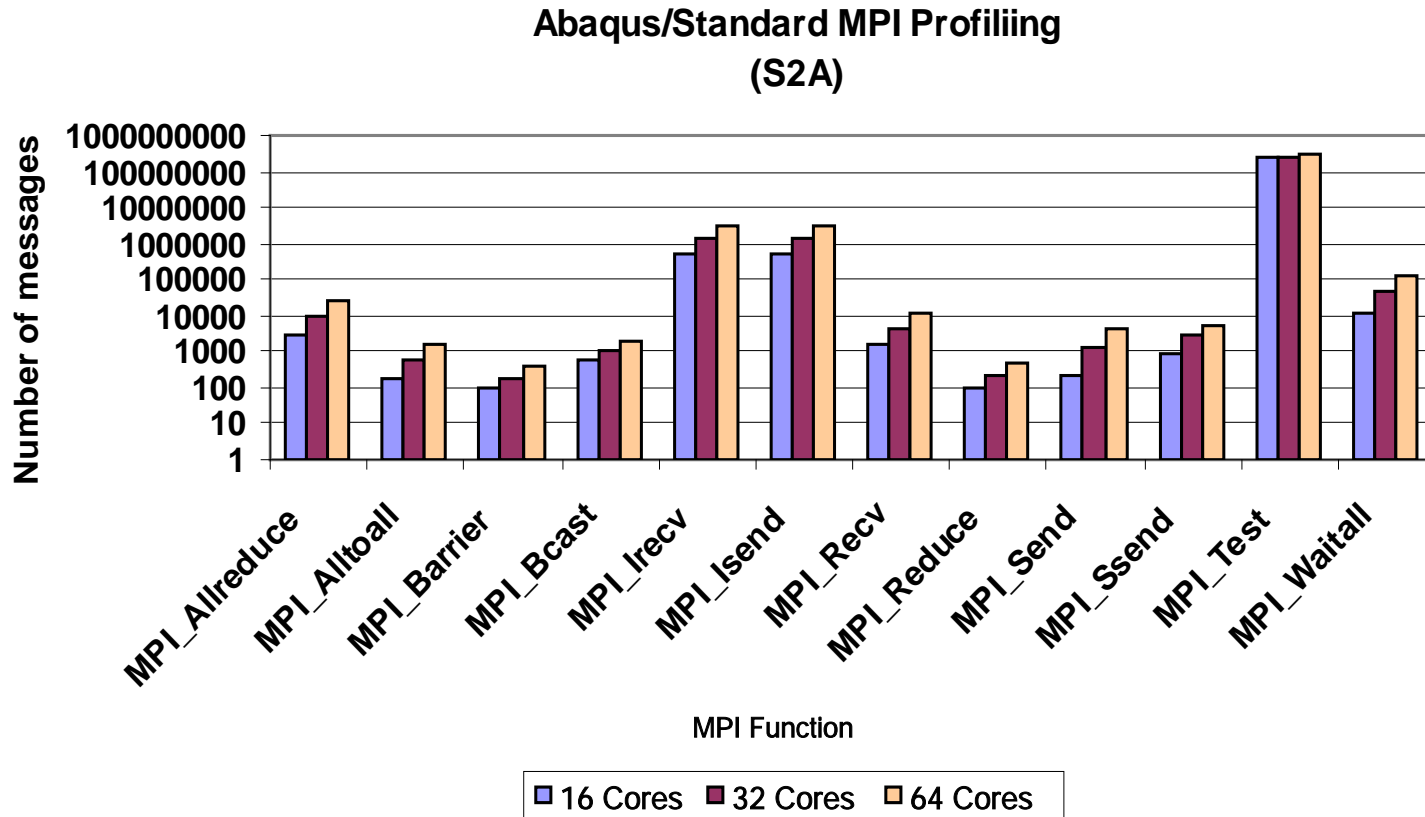
$\$/KWh = KWh * \0.20

For more information - <http://enterprise.amd.com/Downloads/svrpwrusecompletefinal.pdf>

- **Interconnect comparison shows**
 - InfiniBand delivers superior performance in every cluster size
 - Performance advantage extends as cluster size increases
- **InfiniBand enables power saving**
 - Up to \$6600/year power savings versus GigE
 - Up to \$3800/year power savings versus 10GigE
- **Dell™ PowerEdge™ server blades provides**
 - Linear scalability (maximum scalability) and balanced system
 - By integrating InfiniBand interconnect and AMD processors
 - Maximum return on investment through efficiency and utilization

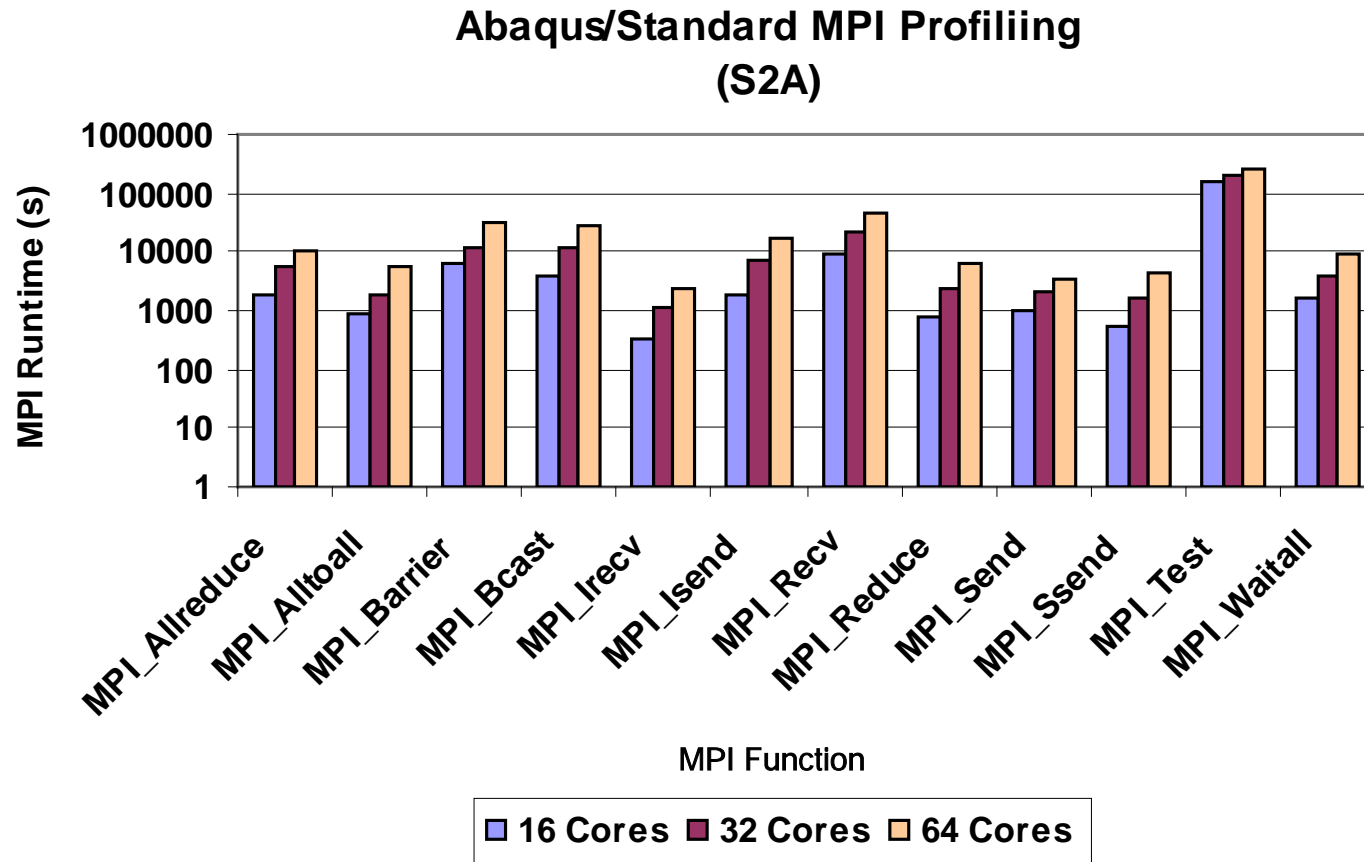
- **Mostly used MPI functions**

- MPI_Test, MPI_Isend, MPI_Irecv, MPI_Waitall, and MPI_Allreduce
- Except MPI_Test, number of other MPI functions increases with cluster size



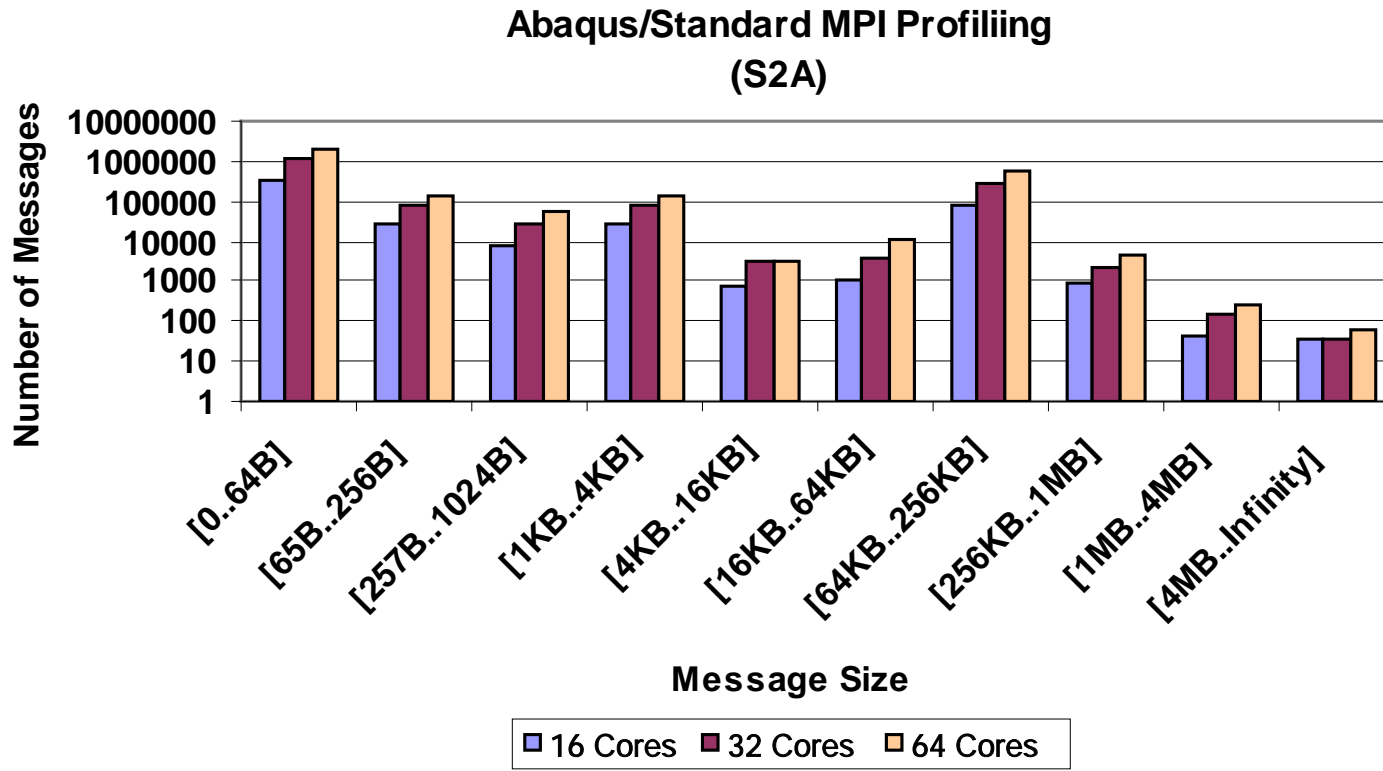
Abaqus/Standard MPI Profiling – Timing

- **MPI_Test, MPI_Recv, and MPI_Barrier/Bcast** show the highest communication overhead



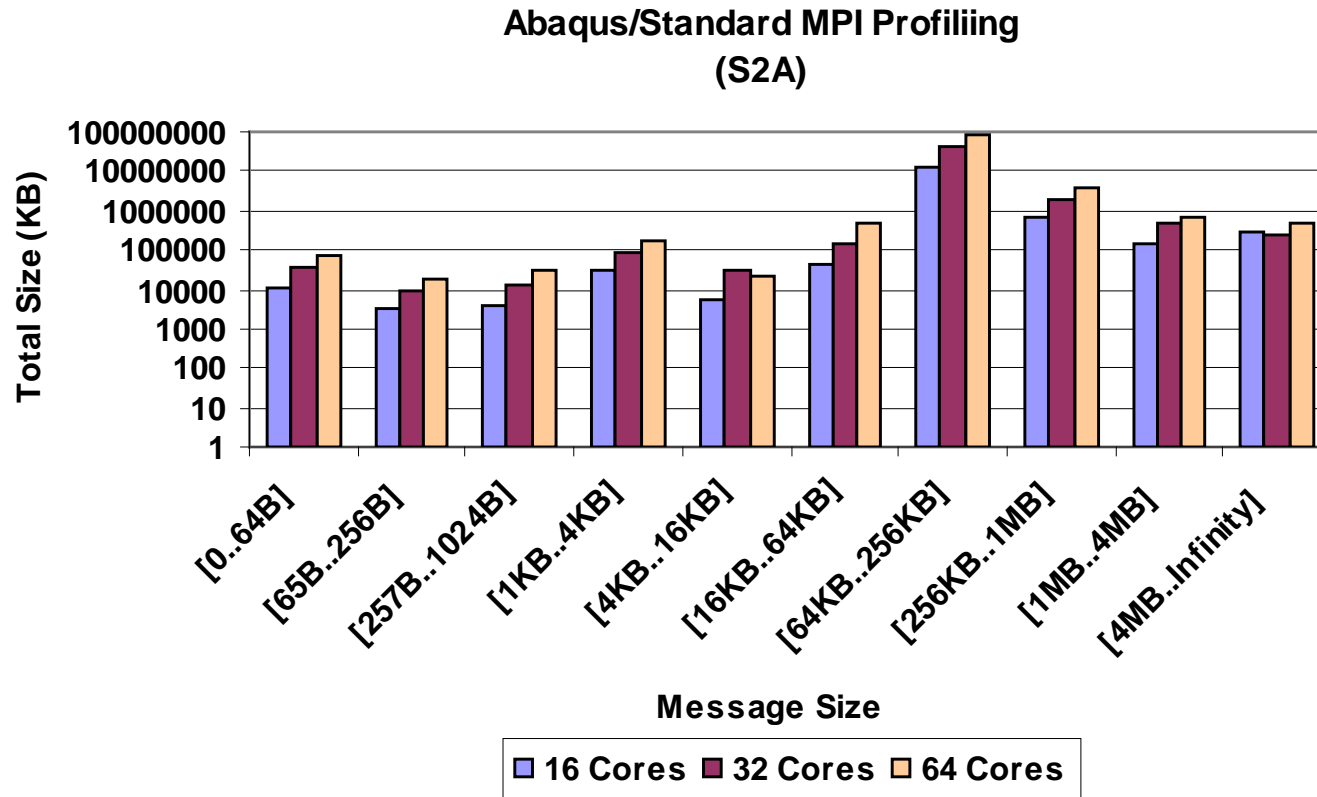
Abaqus/Standard MPI Profiling – Messages

- Majority messages are small and large messages
- Number of messages increases with cluster size



Abaqus/Standard MPI Profiling – Messages

- Most data related MPI messages are within 64KB-256KB
- Total data transferred increases with cluster size



- **Abaqus/Standard was profiled to identify its communication patterns**
- **Frequent used message sizes**
 - Abaqus/Standard has large number of both small and large messages
 - Number of messages increases with cluster size
- **Interconnects effect to Abaqus performance**
 - Both Interconnect latency (MPI_Barrier/Bcast) and bandwidth (MPI_Recv) are important to Abaqus/Standard performance
- **Balanced system – CPU, memory, Interconnect that match each other capabilities, is essential for providing application efficiency**

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