



Abaqus 6.10

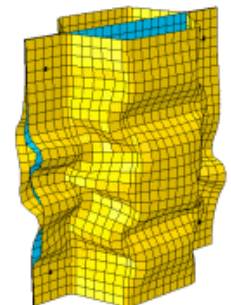
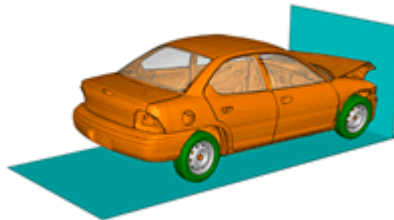
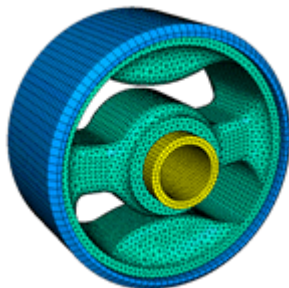
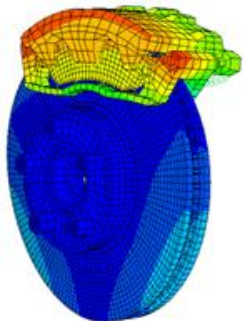
Performance Benchmark and Profiling

August 2011



- **The following research was performed under the HPC Advisory Council activities**
 - Participating vendors: Intel, Dell, Mellanox
 - Compute resource - HPC Advisory Council Cluster Center
- **The following was done to provide best practices**
 - Abaqus performance overview
 - Understanding Abaqus communication patterns
 - Ways to increase Abaqus productivity
 - MPI libraries comparisons
- **For more info please refer to**
 - <http://www.dell.com>
 - <http://www.intel.com>
 - <http://www.mellanox.com>
 - <http://www.simulia.com>

- **Abaqus Unified FEA product suite offers powerful and complete solutions for both routine and sophisticated engineering problems covering a vast spectrum of industrial applications**
- **The Abaqus analysis products listed below focus on:**
 - Nonlinear finite element analysis (FEA)
 - Advanced linear and dynamics application problems
- **Abaqus/Standard**
 - General-purpose FEA that includes broad range of analysis capabilities
- **Abaqus/Explicit**
 - Nonlinear, transient, dynamic analysis of solids and structures using explicit time integration



- **Dell™ PowerEdge™ M610 38-node (456-core) cluster**
 - Six-Core Intel X5670 @ 2.93 GHz CPUs
 - Memory: 24GB memory, DDR3 1333 MHz
 - OS: RHEL 5.5, OFED 1.5.2 InfiniBand SW stack
- **Intel Cluster Ready certified cluster**
- **Mellanox ConnectX-2 InfiniBand adapters and non-blocking switches**
- **MPI: HP-MPI 2.3**
- **Application: Abaqus 6.10-3**
- **Benchmark datasets:**
 - Abaqus/Standard benchmarks: S2A – Flywheel with centrifugal load
 - Abaqus/Standard benchmarks: S4B – Cylinder head bolt-up
 - Abaqus/Explicit benchmarks: E2 – Cell phone drop
 - Abaqus/Explicit benchmarks: E5 – Blast loaded plate

- **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



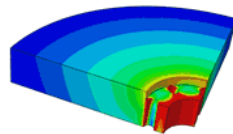
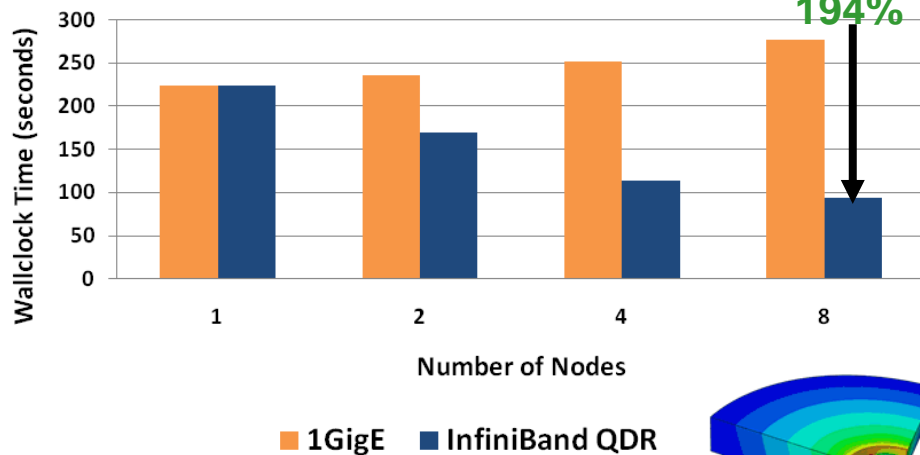
- **System Structure and Sizing Guidelines**
 - 38-node cluster build with Dell PowerEdge™ M610 blade servers
 - Servers optimized for High Performance Computing environments
 - Building Block Foundations for best price/performance and performance/watt
- **Dell HPC Solutions**
 - Scalable Architectures 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



- **Benchmark: dataset: S2A – Flywheel with centrifugal load**
 - Affects largely by MPI communications and network interconnect
 - 1GigE would not allow it to scale beyond 2 nodes
- **Benchmark dataset: S4B – Cylinder head bolt-up**
 - Not as network sensitive; more computationally-intensive problem
- **InfiniBand enables higher scalability and system utilization**
 - Reducing the runtime by 194% compared to 1GigE using the S2A benchmark
 - Enabling faster job turnaround time by up to 102% versus 1GigE using the S4B dataset

Abaqus/Standard Performance (S2A)

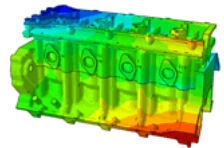
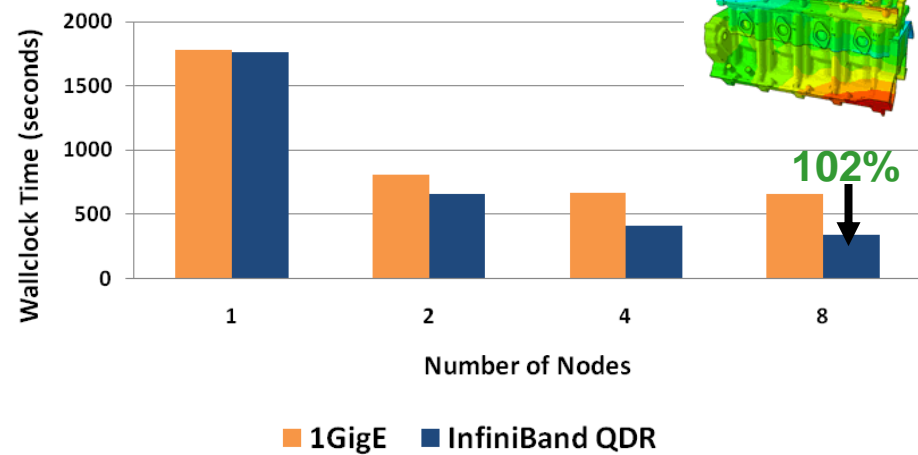
Number of MPI Calls



Lower is better

Abaqus/Standard Performance (S4B)

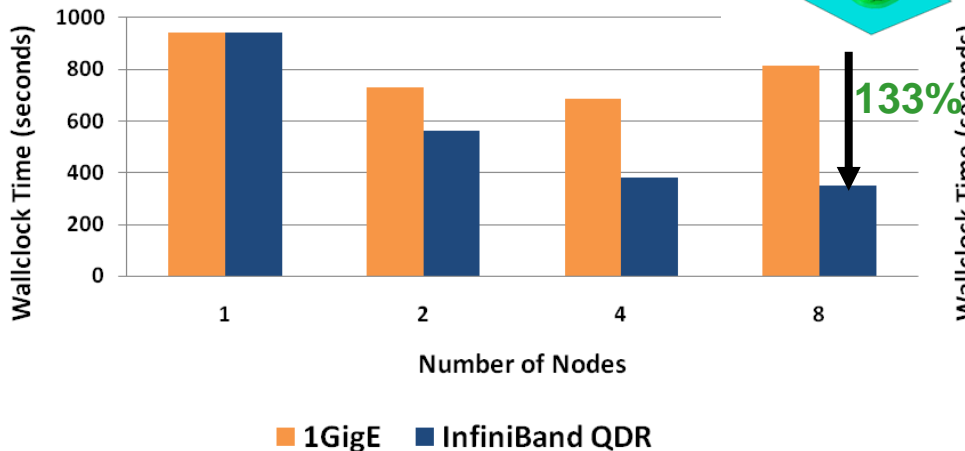
Number of MPI Calls



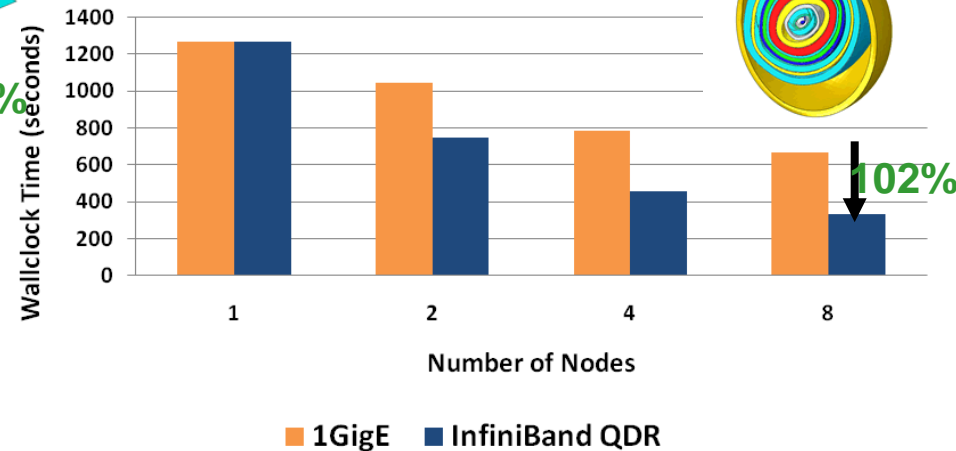
InfiniBand QDR

- **Benchmark dataset: E2 – Cell phone drop**
 - 1GigE would not allow work done beyond 4 nodes
- **Benchmark dataset: E6 – Concentric spheres**
 - Less affected by network latency; more computationally sensitive dataset
- **InfiniBand enables higher cluster productivity**
 - Reducing the runtime by 133% for the E2 dataset
 - Up to 102% higher performance versus 1GigE using the E6 dataset

Abaqus/Explicit Performance (E2)
Number of MPI Calls



Abaqus/Explicit Performance (E6)
Number of MPI Calls

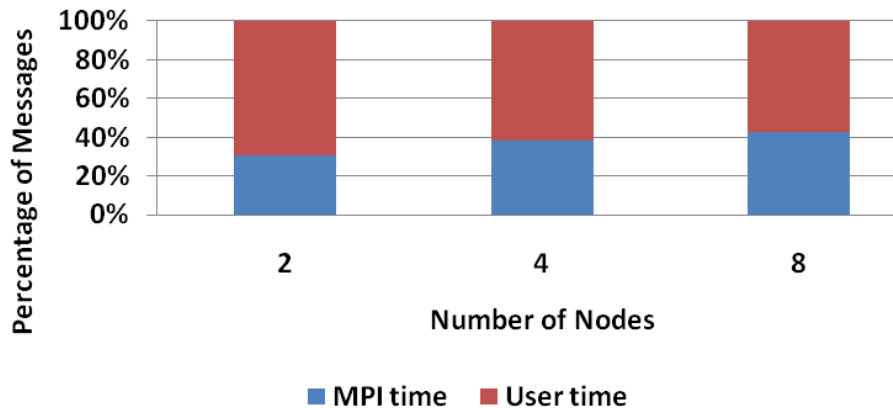


Lower is better

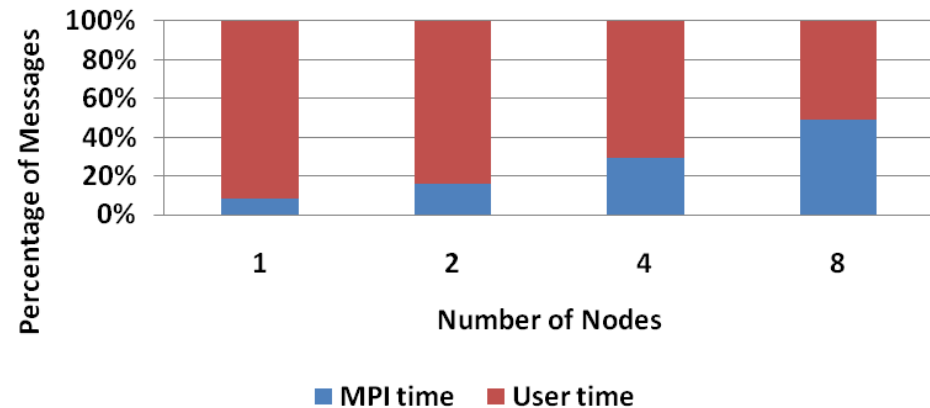
InfiniBand QDR

- **MPI communication percentage increases as the cluster scales**
 - The E2 benchmark dataset has the communication percentage increase at a faster pace
 - The communication ratio for S4B stays flat as more nodes are running the simulation
 - Performance data shows E2 is being affected largely by network, more so than S4B

**Abaqus/Standard Profiling
(S4B)**
MPI/User Time Ratio



**Abaqus/Explicit Profiling
(E2)**
MPI/User Time Ratio

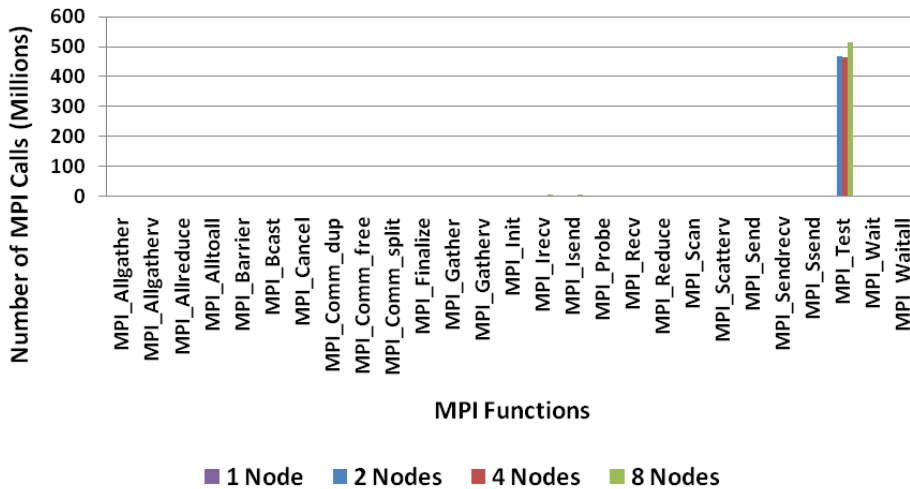


InfiniBand QDR

- **Abaqus uses a wide range of MPI APIs**
- **MPI_Test dominates the MPI function calls for the Abaqus/Standard**
 - Over 96% of the MPI function calls are for MPI_Test on 8-node S4B simulation
- **Abaqus/Explicit uses a range of MPI calls for solving the E2 dataset**
 - MPI_Iprobe (44%) for a 8-node job
 - MPI_Test (16%)
 - MPI_Isend (15%)
 - MPI_Irecv (8%)

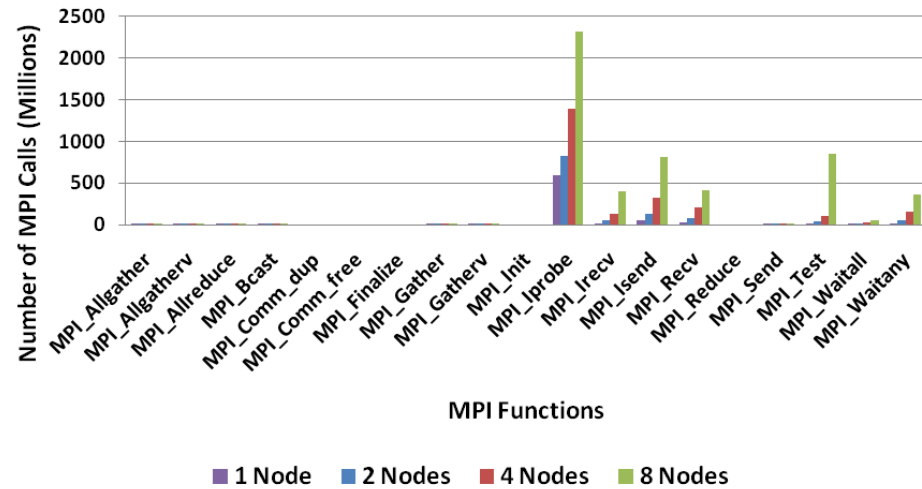
Abaqus/Standard Profiling (S4B)

Number of MPI Calls



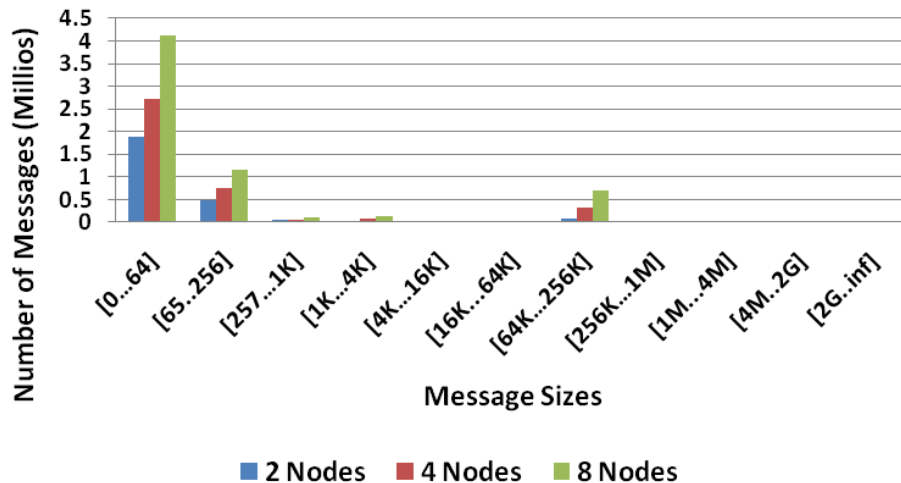
Abaqus/Explicit Profiling (E2)

Number of MPI Calls

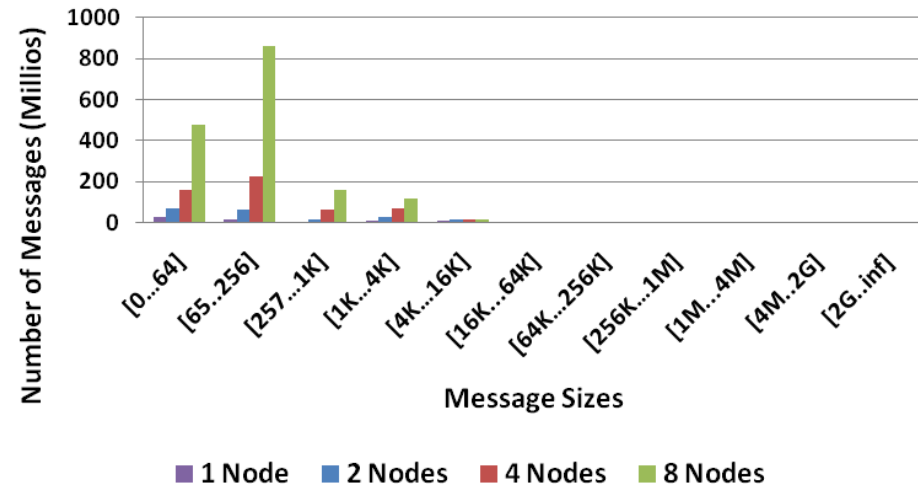


- **Abaqus/Standard uses small and medium MPI message sizes**
 - Most message sizes are between 0B to 64B, and 65B to 256B
 - Some medium size concentration in 64KB to 256KB
- **Abaqus/Explicit has the highest concentration in small message sizes**
 - Highest around 65B to 256B

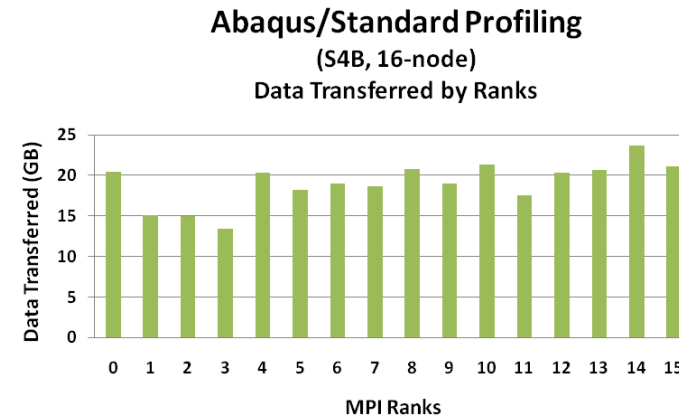
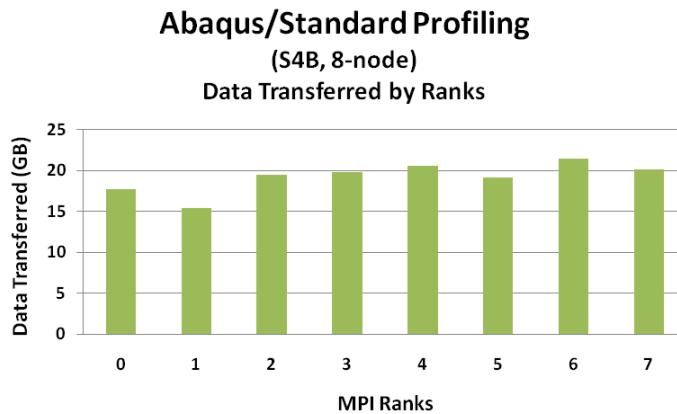
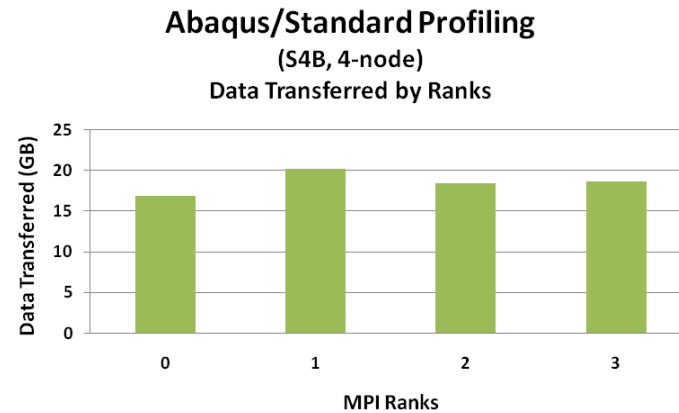
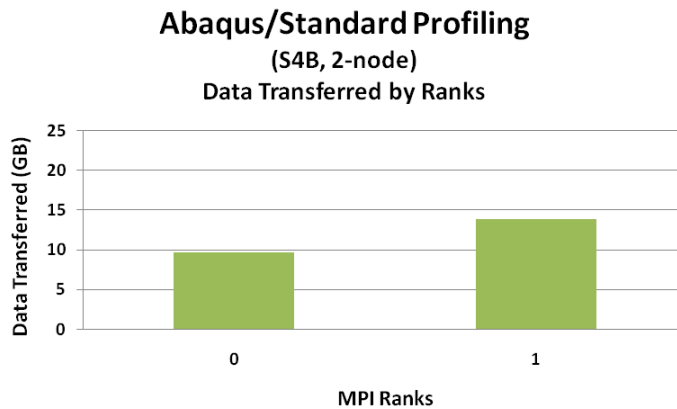
Abaqus/Standard Profiling
(S4B)
MPI Message Sizes



Abaqus/Explicit Profiling
(E2)
MPI Message Sizes

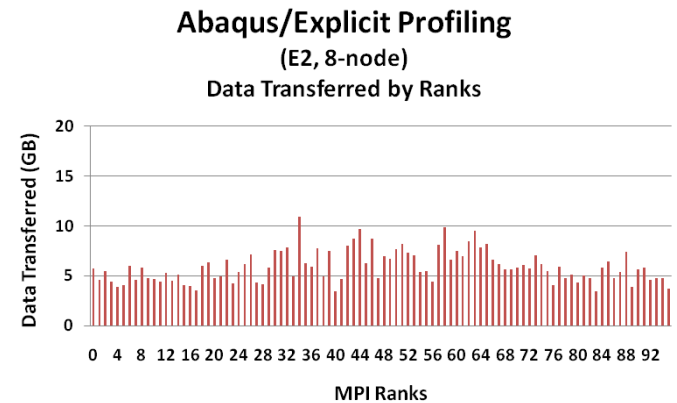
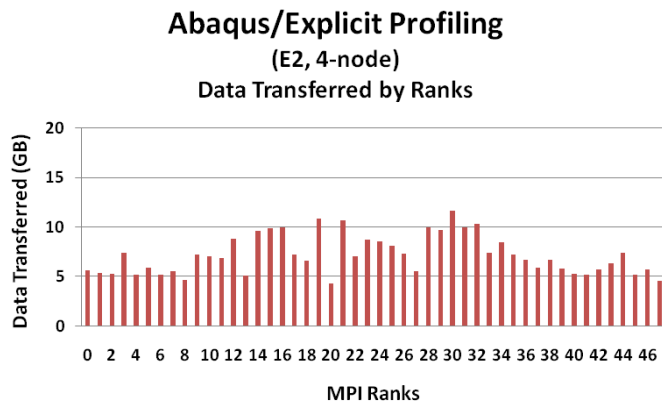
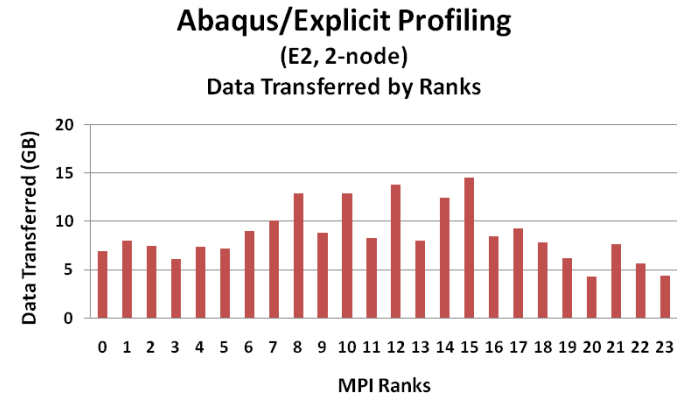
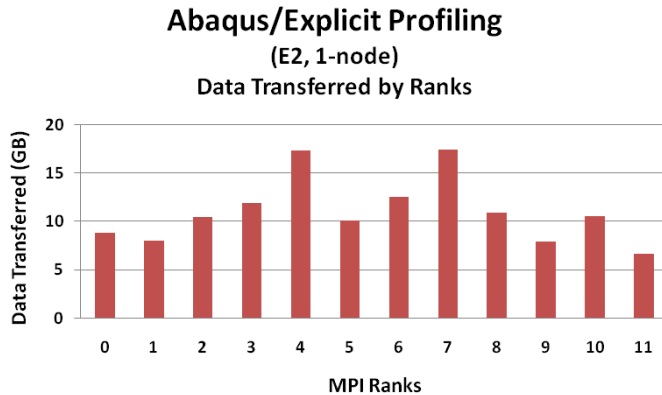


- **Abaqus/Standard uses hybrid of MPI and threading for computation**
- **It shows substantial data transfers between the MPI processes**
 - Growing data communications from around 10GB per process in an 2-node simulation, to 15-20GB for a 8-node simulation



Abaqus/Explicit Profiling – MPI Data Transfer

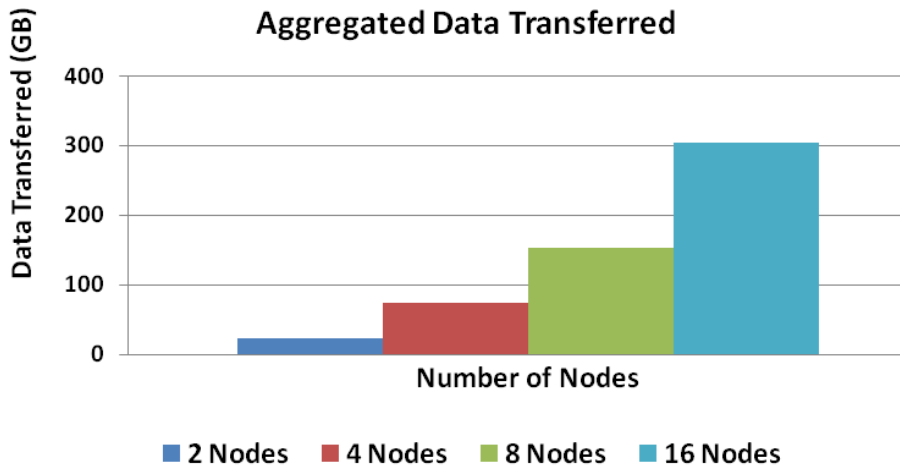
- **Abaqus/Explicit uses pure MPI for job partitioning and execution**
- **Abaqus/Explicit spreads out data transfers as more MPI processes in the job**
 - From 8-18GB per process on a single node, down to 4-10GB per process in 8-node job



- **Aggregated data transfer refers to:**
 - Total amount of data being transferred in the network between all MPI ranks collectively
- **Substantially larger data transfer takes place in Abaqus/Explicit**
 - 12 process per node in Abaqus/Explicit would take part in MPI communication
 - Only 1 process per node in the MPI hybrid of Abaqus/Standard, which reduces communications needed to take place

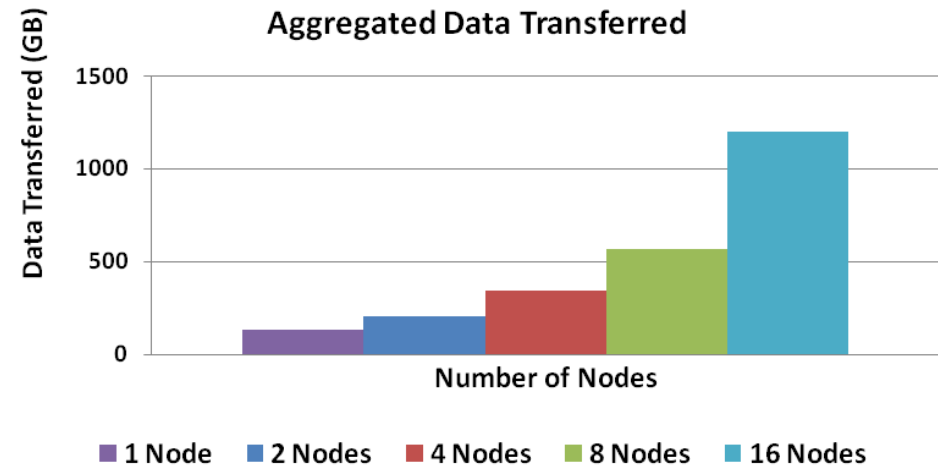
Abaqus/Standard Profiling
(S4B)

Aggregated Data Transferred



Abaqus/Explicit Profiling
(E2)

Aggregated Data Transferred



InfiniBand QDR

- **InfiniBand allows Abaqus to run at the most efficient rate**
 - InfiniBand enables Abaqus to achieve fastest runtime and highest cluster scalability
 - Ethernet would not allow scale, ended up wasting valuable system resources
- **Abaqus/Standard**
 - Uses hybrid of MPI and threading
 - Threading for computation
 - Only 1 process per node is responsible for communications
 - MPI_Test is the most dominant MPI function call
- **Abaqus/Explicit**
 - Uses pure MPI for job partitioning and execution
 - Significantly more communications is taken place compared to Abaqus/Standard
 - MPI_lprobe is the most used MPI function

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