



# Abaqus 6.12-2

## Performance Benchmark and Profiling

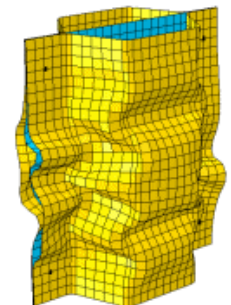
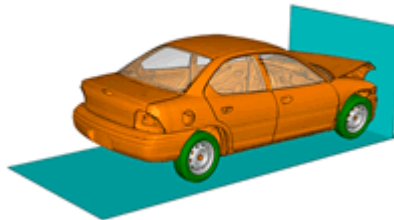
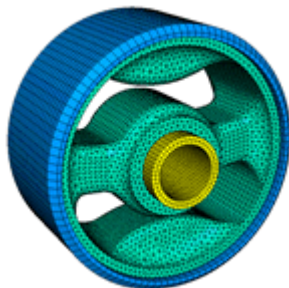
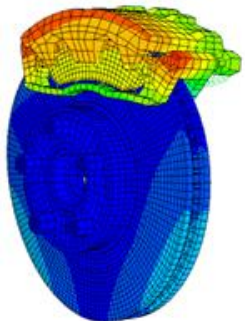
April 2013



- **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>

- **The following was done to provide best practices**
  - Abaqus performance benchmarking
  - Interconnect performance comparisons
  - MPI performance comparison
  - Understanding Abaqus communication patterns
- **The presented results will demonstrate**
  - The scalability of the compute environment to provide nearly linear application scalability
  - The capability of Abaqus to achieve scalable productivity

- **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™ R720xd 16-node (256-core) “Jupiter” cluster**
  - Dual-Socket Eight-Core Intel E5-2680 @ 2.70 GHz CPUs (Static max Perf in BIOS)
  - Memory: 64GB memory, DDR3 1600 MHz
  - OS: RHEL 6.2, OFED 1.5.3 InfiniBand SW stack
  - Hard Drives: 24x 250GB 7.2 RPM SATA 2.5” on RAID 0
- **Intel Cluster Ready certified cluster**
- **Mellanox ConnectX-3 FDR InfiniBand VPI adapters**
- **Mellanox SwitchX SX6036 InfiniBand switch**
- **Application: Abaqus 6.12-2 with Platform MPI 8.1.2 (vendor-provided)**
- **Benchmark datasets:**
  - Abaqus/Standard benchmarks: S2A – Flywheel with centrifugal load
  - Abaqus/Explicit benchmarks: E6 – Concentric Spheres

- **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



# PowerEdge R720xd

Massive flexibility for data intensive operations

- **Performance and efficiency**

- Intelligent hardware-driven systems management with extensive power management features
- Innovative tools including automation for parts replacement and lifecycle manageability
- Broad choice of networking technologies from GigE to IB
- Built in redundancy with hot plug and swappable PSU, HDDs and fans



- **Benefits**

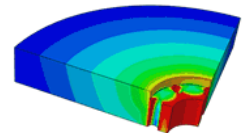
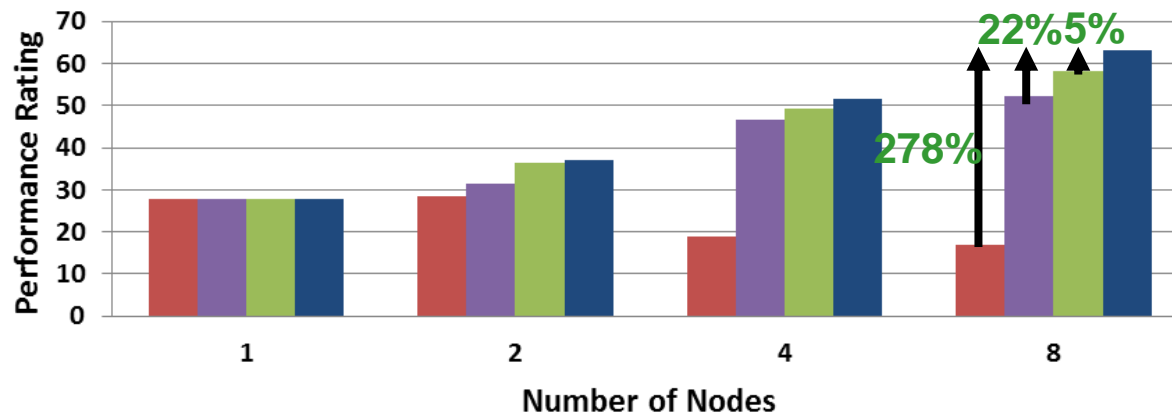
- Designed for performance workloads
  - from big data analytics, distributed storage or distributed computing where local storage is key to classic HPC and large scale hosting environments
  - High performance scale-out compute and low cost dense storage in one package

- **Hardware Capabilities**

- Flexible compute platform with dense storage capacity
  - 2S/2U server, 6 PCIe slots
- Large memory footprint (Up to 768GB / 24 DIMMs)
- High I/O performance and optional storage configurations
  - HDD options: 12 x 3.5" - or - 24 x 2.5 + 2x 2.5 HDDs in rear of server
  - Up to 26 HDDs with 2 hot plug drives in rear of server for boot or scratch

- **Benchmark input: S2A – Flywheel with centrifugal load**
  - Abaqus/Standard is more compute intensive, 1 MPI process is launched per node
- **InfiniBand enables higher scalability and system utilization**
  - Reducing the runtime by 278% compared to 1GbE
  - Enabling faster job turnaround time by up to 22% versus 10GbE, 5% vs 40GbE
  - 1GbE would not allow it to scale beyond 2 nodes

## Abaqus/Standard Performance (S2A)



*Higher is better*

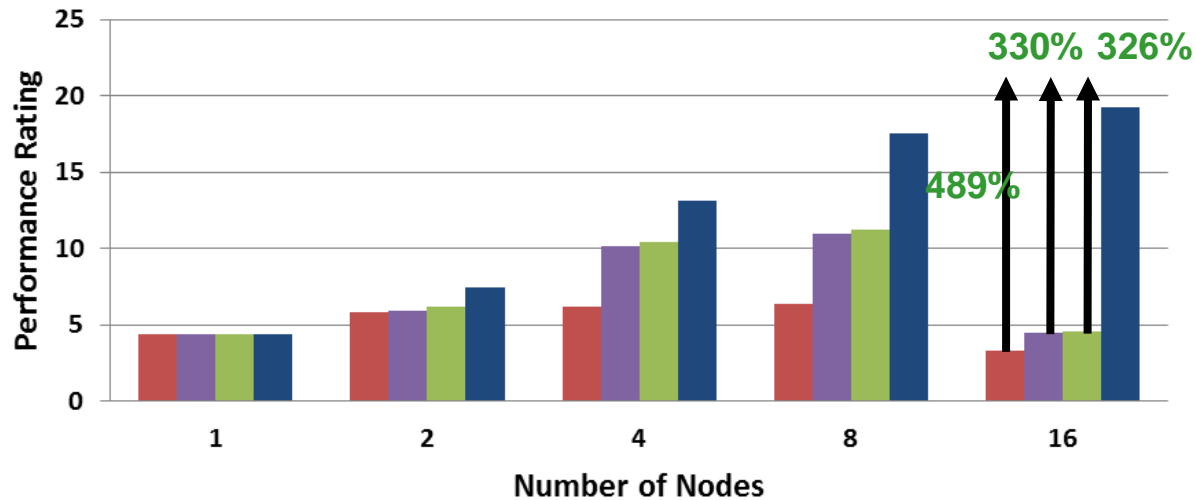
■ 1GbE ■ 10GbE ■ 40GbE ■ FDR InfiniBand

*1 MPI Process/Node*



- **Benchmark input: E6 – Concentric spheres**
  - Abaqus/Explicit is more data intensive, as more processes running per node
  - More MPI processes generate more data communications as a result
- **InfiniBand enables higher cluster productivity**
  - Reducing the runtime by 489% versus 1GbE
  - Up to 330% higher performance versus 10GbE
  - Up to 326% higher performance versus 40GbE

**Abaqus/Explicit Performance  
(E6)**



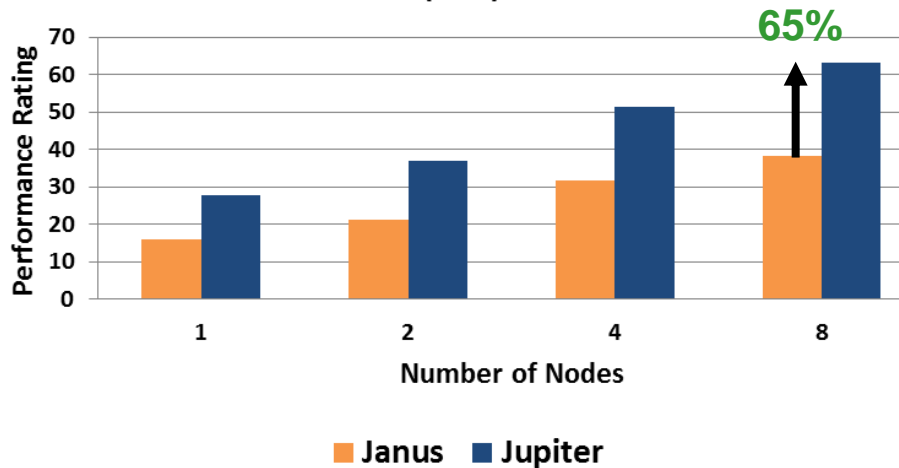
*Higher is better*

■ 1GbE ■ 10GbE ■ 40GbE ■ InfiniBand FDR

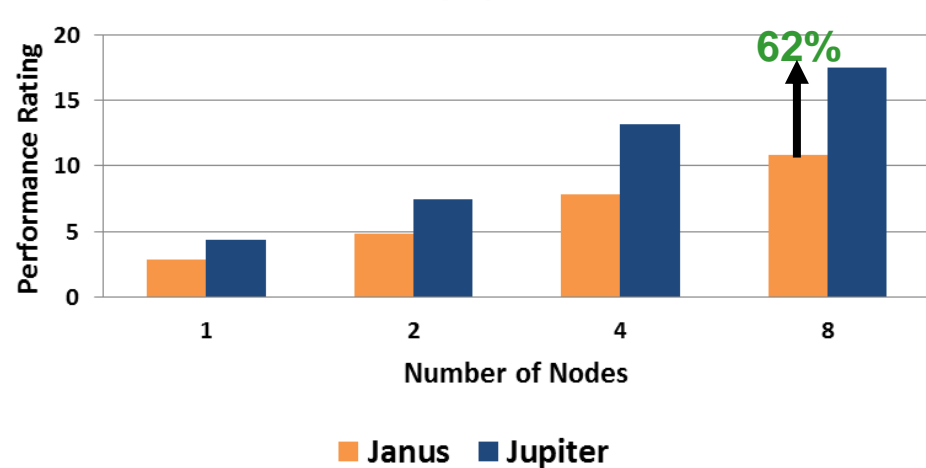
**16 MPI Processes/Node**

- **Intel E5-2680 (Sandy Bridge) cluster outperforms prior CPU generation**
  - Abaqus/Standard: Performs 65% higher than X5670 cluster at 8 nodes with S2A
  - Abaqus/Explicit: Performs 62% higher than X5670 cluster at 8 nodes with E6
- **System components used:**
  - Jupiter: 2-socket Intel E5-2680 @ 2.7GHz, 1600MHz DIMMs, FDR InfiniBand, v6.12-2
  - Janus: 2-socket Intel X5670 @ 2.93GHz, 1333MHz DIMMs, QDR InfiniBand, v6.10-3

### Abaqus/Standard Performance (S2A)



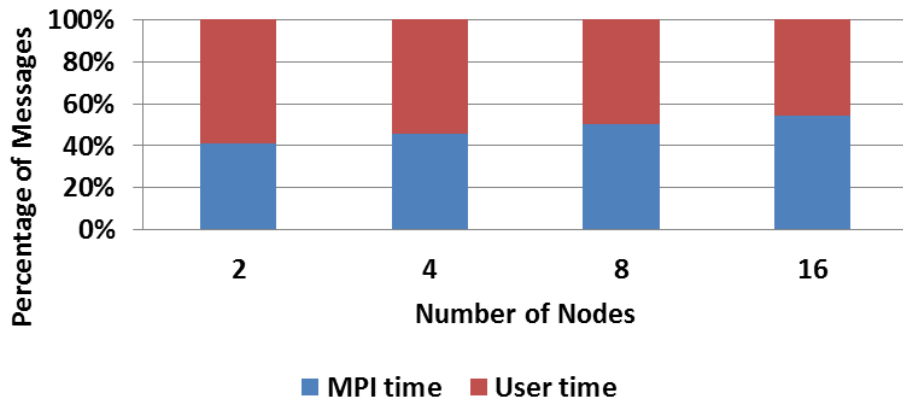
### Abaqus/Explicit Performance (E6)



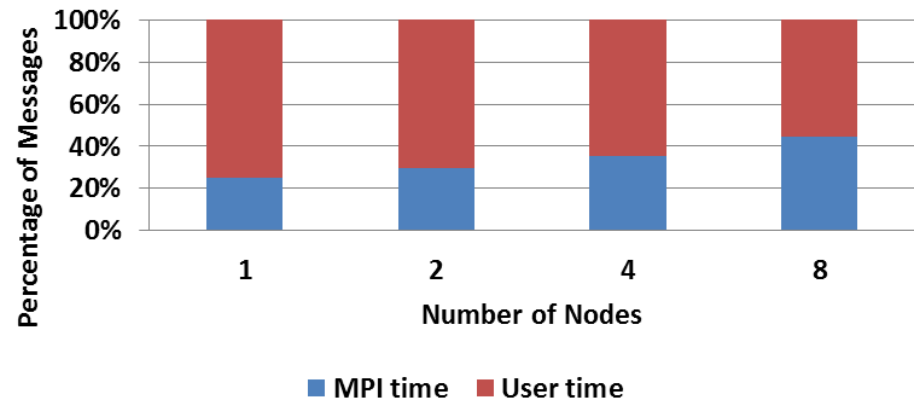
*Higher is better*

- **MPI communication percentage increases as the cluster scales**
  - The communication ratio for S4A stays flat as more nodes are running the simulation
  - Performance data shows E6 is being affected largely by network

**Abaqus/Standard Profiling (S2A)**  
MPI/User Time Ratio



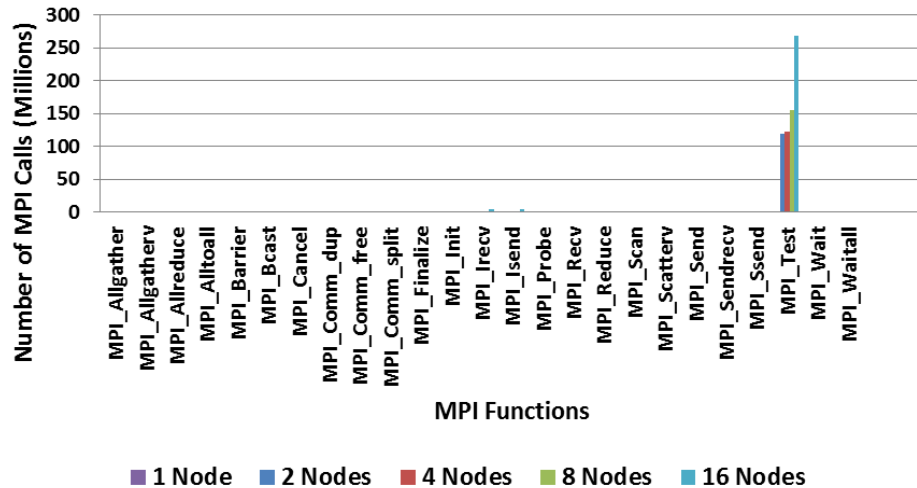
**Abaqus/Explicit Profiling (E6)**  
MPI/User Time Ratio



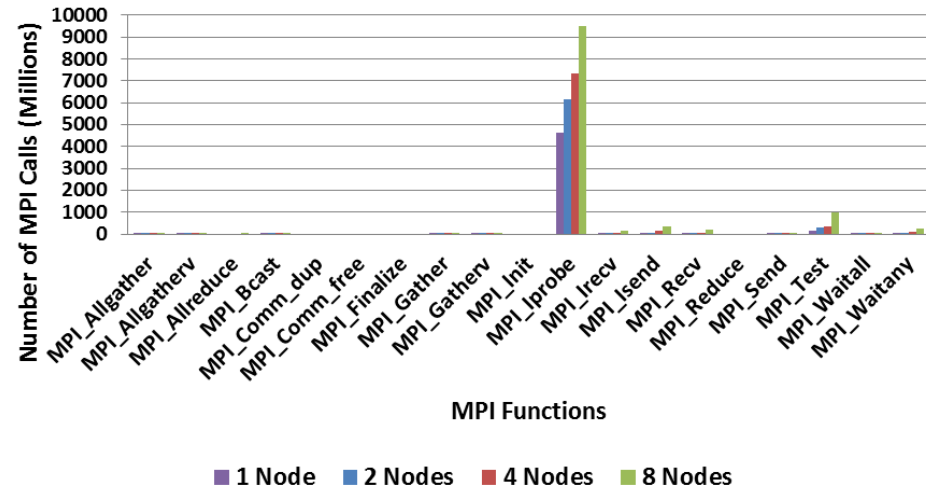
*FDR InfiniBand*

- **Abaqus/Standard uses a wide range of MPI APIs**
  - MPI\_Test dominates the MPI function calls (over 97%)
  - MPI\_Test on 8-node S4B simulation (for non-blocking communications)
- **Abaqus/Explicit uses a range of MPI calls for solving the E2 dataset**
  - MPI\_Iprobe (83%) and MPI\_Test (9%) use almost exclusively

**Abaqus/Standard Profiling (S2A)**  
Number of MPI Calls



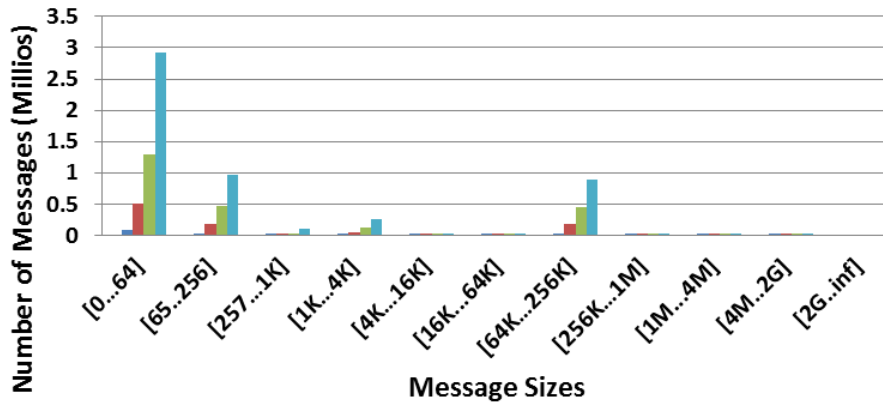
**Abaqus/Explicit Profiling (E6)**  
Number of MPI Calls





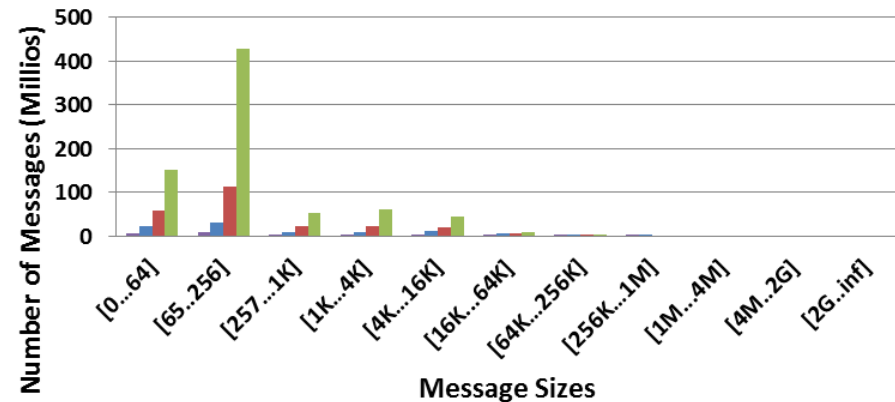
- **Abaqus/Standard uses small and medium MPI message sizes**
  - Most message sizes are between 0B to 64B, and 65B to 256B
  - Some medium size concentration in 64KB to 256KB
- **Abaqus/Explicit has the highest concentration in small message sizes**
  - Highest around 65B to 256B

**Abaqus/Standard Profiling (S2A)**  
MPI Message Sizes



1 Node 2 Nodes 4 Nodes 8 Nodes 16 Nodes

**Abaqus/Explicit Profiling (E6)**  
MPI Message Sizes

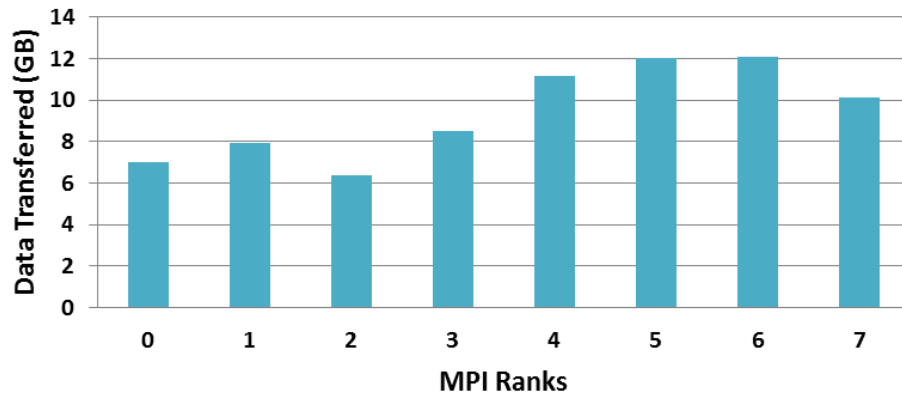


1 Node 2 Nodes 4 Nodes 8 Nodes

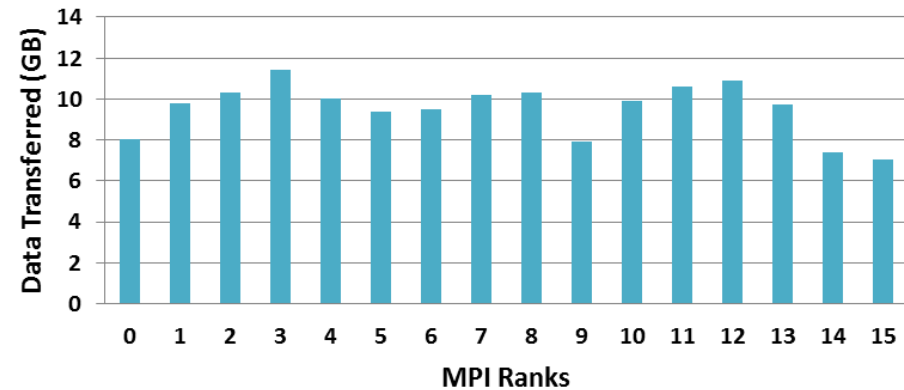


- **Abaqus/Standard uses a hybrid of MPI and threading**
- **Substantial data transfers between the MPI processes**
  - Growing data communications as more nodes involved in simulation

**Abaqus/Standard Profiling**  
(S2A, 8-node)  
Data Transferred by Ranks

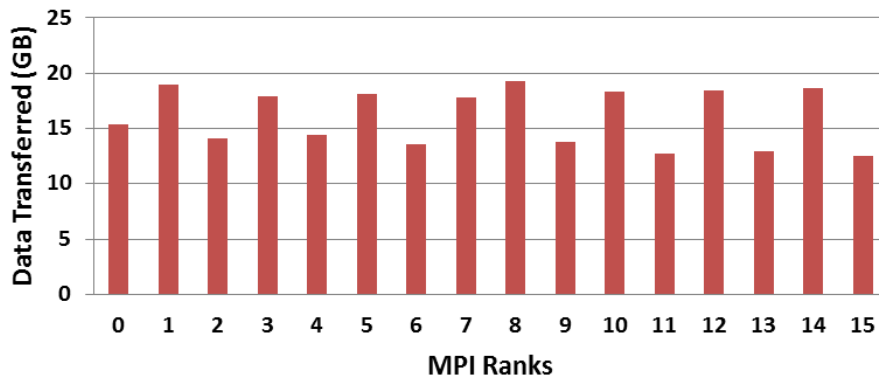


**Abaqus/Standard Profiling**  
(S2A, 16-node)  
Data Transferred by Ranks

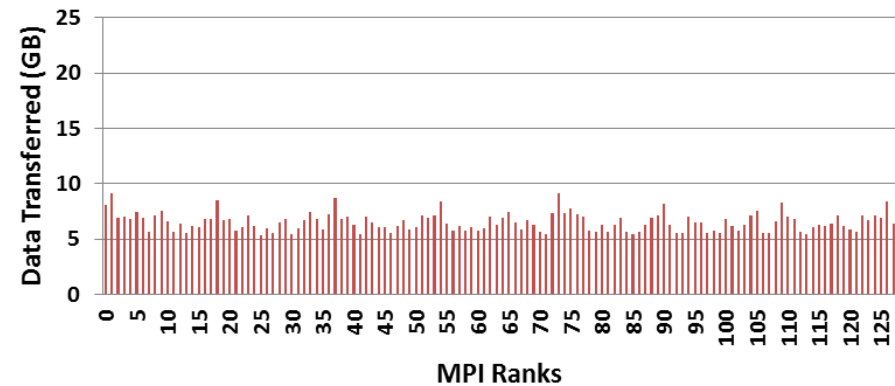


- **Abaqus/Explicit uses pure MPI for job partitioning and execution**
- **Abaqus/Explicit spreads out data transfers as more MPI processes in the job**
  - From 8-18GB per process on a single node, down to 6GB per process in 8-node job

**Abaqus/Explicit Profiling**  
(E6, 1-node)  
Data Transferred by Ranks

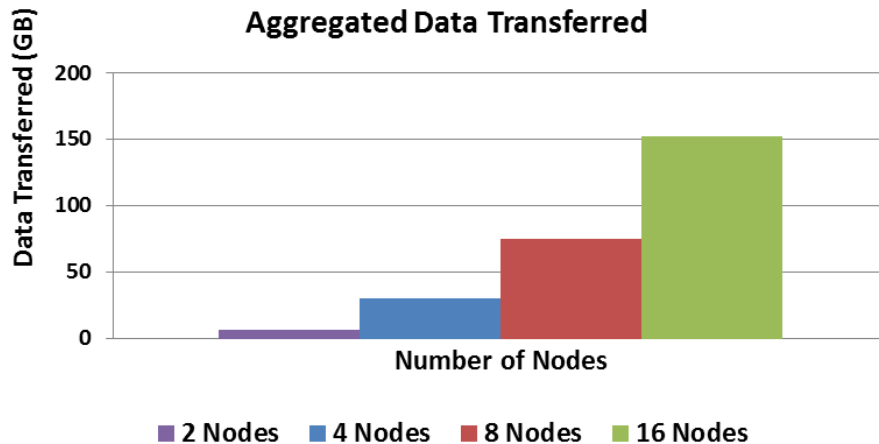


**Abaqus/Explicit Profiling**  
(E6, 8-node)  
Data Transferred by Ranks

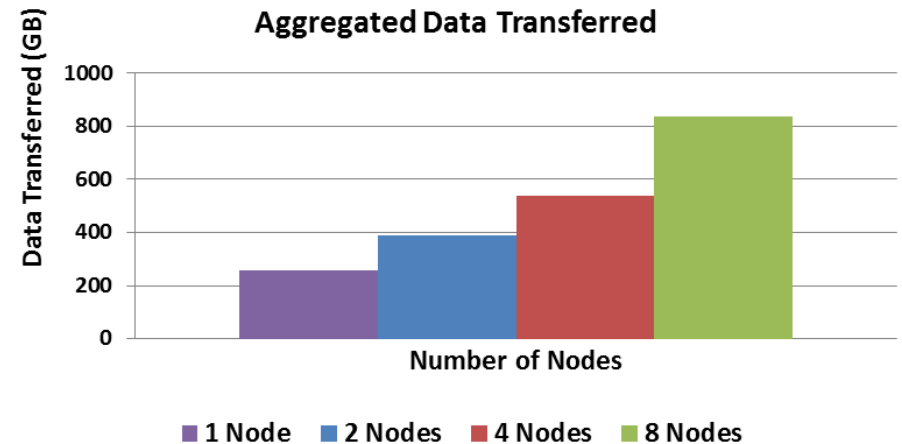


- **Aggregated data transfer refers to:**
  - Total amount of data being transferred in the network between all MPI ranks collectively
- **Substantially larger data transfers take place in Abaqus/Explicit**
  - 16 process per node in Abaqus/Explicit would take part in MPI communication
  - Only 1 process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

### Abaqus/Standard Profiling (S2A)



### Abaqus/Explicit Profiling (E6)



- **Abaqus 6.12 provides scalable, high-quality realistic simulation solutions**
- **Abaqus/Standard**
  - Uses hybrid of MPI and threading; 1 process/node responsible for communications
  - MPI\_Test is the most dominant MPI function call
- **Abaqus/Explicit**
  - Uses pure MPI for job execution and communications
  - Creates significantly more communications compared to Abaqus/Standard
- **InfiniBand enables Abaqus to achieve fastest runtime and highest cluster scalability**
  - Abaqus/Standard: IB provides faster runtime by 278% vs 1GbE, 22% vs 10GbE
  - Abaqus/Explicit: InfiniBand performs better by 489% versus 1GbE, 330% vs 10GbE
  - Ethernet would not allow scale, ended up wasting valuable system resources
  - 1GbE would not allow it to scale beyond 2 nodes
- **Intel E5-2600 Series (Sandy Bridge) cluster outperforms prior generation**
  - Performs 62-65% higher than X5670 cluster at 8 nodes

# 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