Abaqus Performance Benchmark and Profiling

December 2009
• The following research was performed under the HPC Advisory Council activities
  – Participating vendors: Intel, SIMULIA, Dell, 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
• 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
Objectives

• 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
Test Cluster Configuration

- Dell™ PowerEdge™ M610 24-node cluster
- Quad-Core Intel X5570 @ 2.93 GHz CPUs
- Intel Cluster Ready certified cluster
- Mellanox ConnectX MCQH29-XCC 4X QDR InfiniBand mezzanine card
- Mellanox M3601Q 32-Port Quad Data Rate (QDR-40Gb) InfiniBand Switch
- Memory: 24GB memory 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: S4B
  - Abaqus/Explicit Server Benchmarks: E5
Mellanox InfiniBand Solutions

- **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
Delivering Intelligent Performance
Next Generation Intel® Microarchitecture

Bandwidth Intensive
• Intel® QuickPath Technology
• Integrated Memory Controller

Threaded Applications
• 45nm quad-core Intel® Xeon® Processors
• Intel® Hyper-threading Technology

Performance on Demand
• Intel® Turbo Boost Technology
• Intel® Intelligent Power Technology

Performance That Adapts to The Software Environment
Dell PowerEdge Servers helping Simplify IT

- **System Structure and Sizing Guidelines**
  - 24-node cluster build with Dell PowerEdge™ M610 blades server
  - 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
Abaqus/Standard Benchmark Results

- **Input Dataset: S4B**
  - Cylinder head bolt-up
- **InfiniBand provides higher utilization, performance and scalability**
  - Up to 51% higher performance versus GigE

![Abaqus/Standard Server Benchmark (S4B)](image)

*Lower is better*
Abaqus/Explicit Benchmark Results

• Input Dataset: E5
  – Blast loaded plate

• InfiniBand provides higher utilization, performance and scalability
  – Up to 57% higher performance versus GigE

Lower is better

Abaqus/Explicit Server Benchmark (E5)

<table>
<thead>
<tr>
<th>Number of Cores</th>
<th>Total Runtime (s)</th>
<th>8-cores per node</th>
</tr>
</thead>
<tbody>
<tr>
<td>8</td>
<td>GigE: 480, InfiniBand QDR: 380</td>
<td>57%</td>
</tr>
<tr>
<td>16</td>
<td>GigE: 320, InfiniBand QDR: 220</td>
<td></td>
</tr>
<tr>
<td>32</td>
<td>GigE: 200, InfiniBand QDR: 110</td>
<td></td>
</tr>
<tr>
<td>64</td>
<td>GigE: 100, InfiniBand QDR: 55</td>
<td></td>
</tr>
</tbody>
</table>
Power Cost Savings with Different Interconnect

- InfiniBand saves up to $6400 power to finish the same number of Abaqus jobs compared to GigE
  - Yearly based for 24-node cluster
- As cluster size increases, more power can be saved

Abaqus Power Cost
InfiniBand QDR vs. GigE
(24 Nodes)

<table>
<thead>
<tr>
<th>Power Cost ($)</th>
<th>Abaqus/Standard (S4B)</th>
<th>Abaqus/Explicit (E5)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>GigE</td>
<td>InfiniBand</td>
</tr>
<tr>
<td>$5750</td>
<td>$6400</td>
<td></td>
</tr>
</tbody>
</table>

$/KWh = KWh * $0.20  
Abaqus Benchmark Results Summary

- Interconnect comparison shows
  - InfiniBand delivers superior performance in every cluster size
  - Performance advantage extends as cluster size increases
- InfiniBand enables power saving
  - Up to $6400/year power savings versus GigE
• **Mostly used MPI functions**
  – MPI_Test, MPI_Isend, MPI_Irecv, MPI_Waitall, and MPI_Allgather/Allgatherv
• MPI_Allgatherv and MPI_Allgather show highest communication overhead
- Majority messages are small and medium messages
- Number of messages increases with cluster size
• Most data related MPI messages are within 65B-256B
• Total data transferred increases with cluster size
Abaqus Profiling Summary

- Abaqus/Explicit was profiled to identify its communication patterns
- Frequent used message sizes
  - Abaqus/Explicit has large number of both small and medium messages
  - Number of messages increases with cluster size
- Interconnects effect to Abaqus performance
  - Both Interconnect latency (MPI_Allgather/Allgatherv) and bandwidth
    (MPI_Isend/Irecv) are important to Abaqus/Explicit performance
- Balanced system – CPU, memory, Interconnect that match each other capabilities, is essential for providing application efficiency
Productive Systems = Balanced System

- **Balanced system enables highest productivity**
  - Interconnect performance to match CPU capabilities
  - CPU capabilities to drive the interconnect capability
  - Memory bandwidth to match CPU performance

- **Applications scalability relies on balanced configuration**
  - “Bottleneck free”
  - Each system components can reach it’s highest capability

- **Dell M610 system integrates balanced components**
  - Intel “Nehalem” CPUs and Mellanox InfiniBand QDR
  - Latency to memory and Interconnect latency at the same magnitude of order
  - Provide the leading productivity and power/performance system for Desmond simulations