Performance of Large Scale Implicit Crash Analysis on Multicore Processor Systems

Yih-Yih Lin, PhD

8th European LS-DYNA Users Conferences, Strasbourg—May 2011

Introduction

Main Topics

- Scalability and Performance
  - Core Placements
  - Cluster Network Interconnects
- Memory

Additional Information

- Power usage
- Rack space

Models Used in this Study

– Cylinders of 3-DOF nodes
  - Cyl1e6: 1 million nodes, 3 MDOF
  - Cyl2e6: 2 million nodes, 6 MDOF
  - Cyl4e6: 4 million nodes, 12 MDOF

Systems Used in this Study

Two HP ProLiant DL980 Servers

Architecture: 8 8-core Intel Xeon 7650 (2.3 GHz) Processors
Cache: 8 MB/processor, shared
Memory: Shared; 1TB & 0.5TB respectively

Two HP ProLiant BL2x220c Clusters

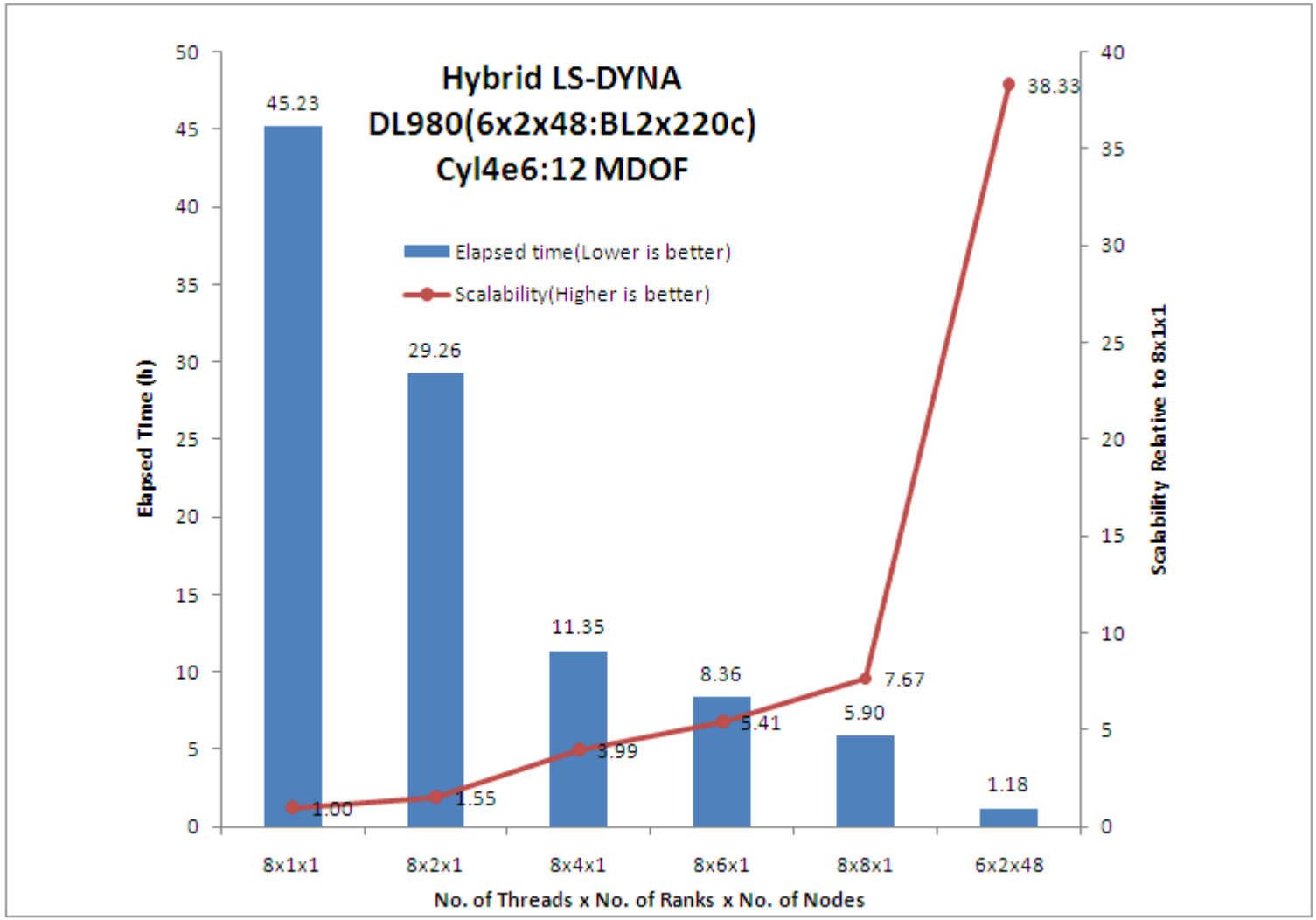
Architecture: Two 6-core Intel Xeon 5670 (2.3 GHz) Processors
Cache: 12 MB/processor, shared
Memory: Shared; 24GB/node & 48GB/node respectively
Interconnect: InfiniBand QDR and GigE

Terms and Notations

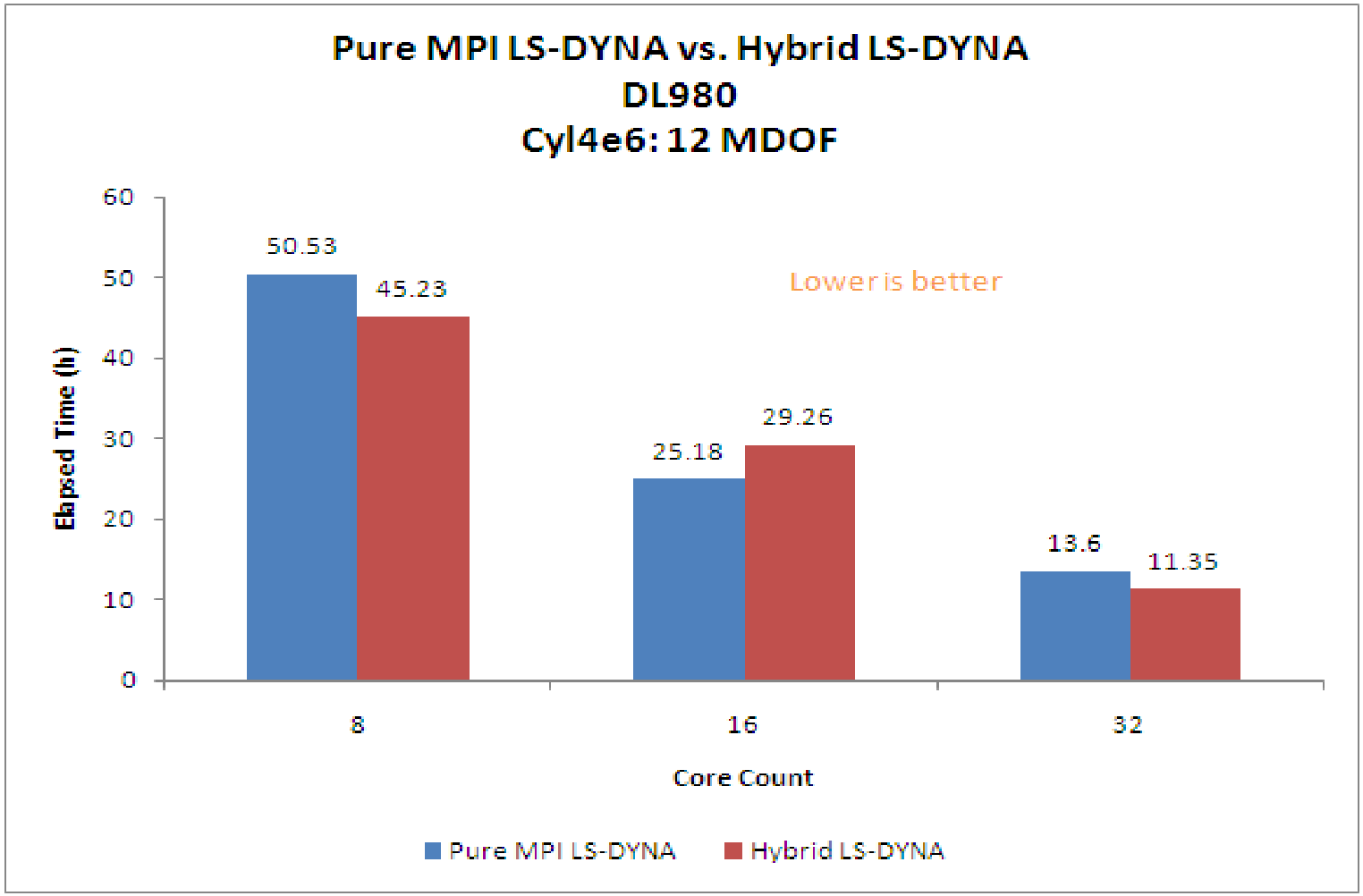
- mP/nC: A server (or node) that comprises m processors with a total of n cores
- Core count: Number of parallel processes
- Rank: Number of MPI processes
- SxRxN: A Hybrid LS-DYNA job with S SMP threads per rank, R ranks per node, and N nodes

Scalability

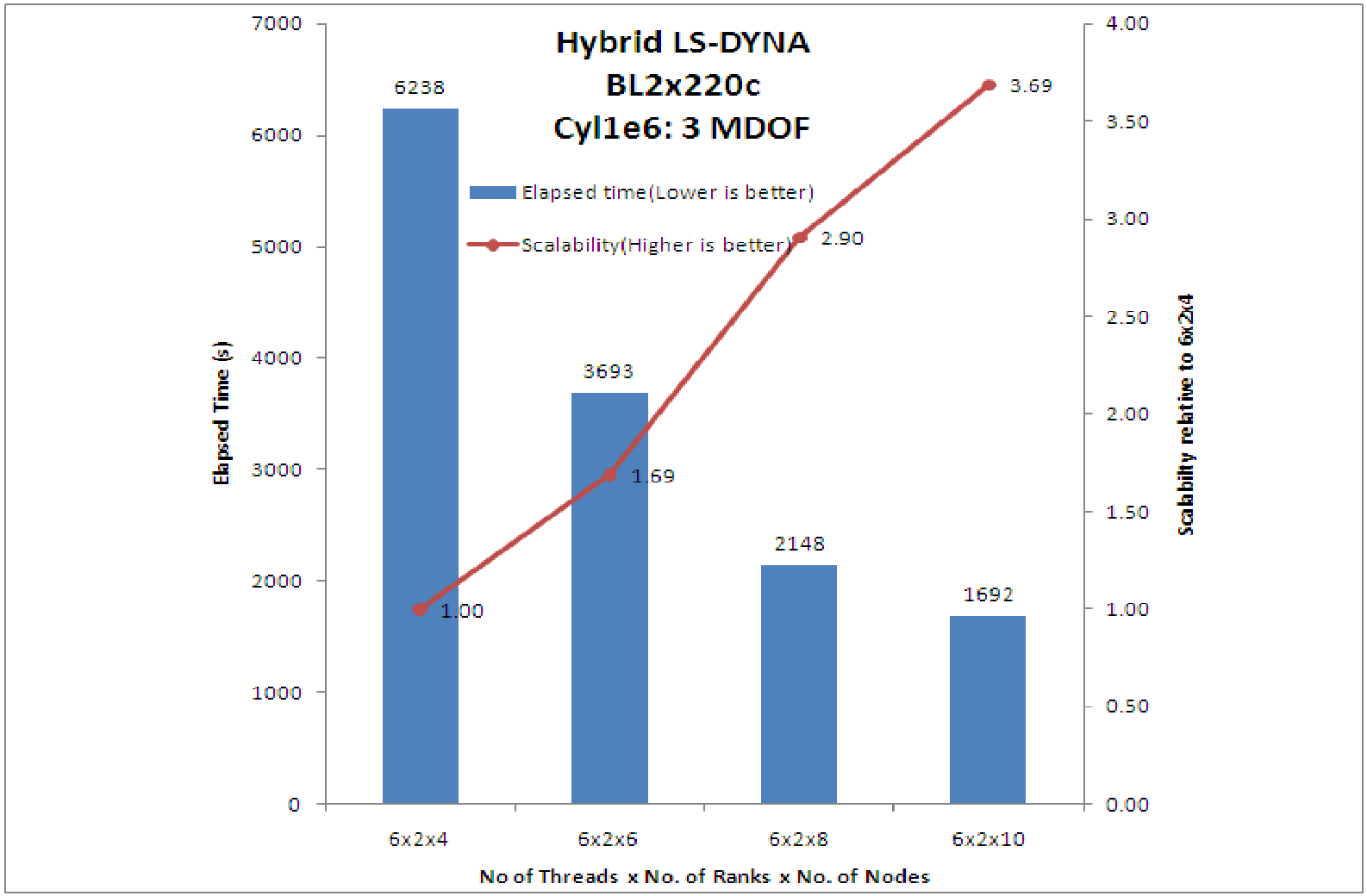
Performance of Hybrid LS-DYNA



Pure MPI LS-DYNA vs. Hybrid LS-DYNA

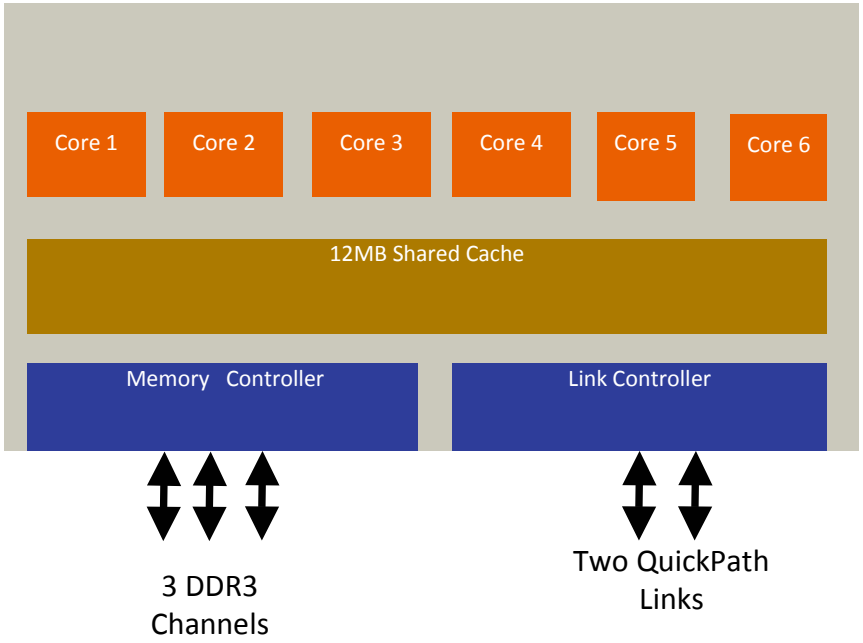


Performance of Hybrid LS-DYNA

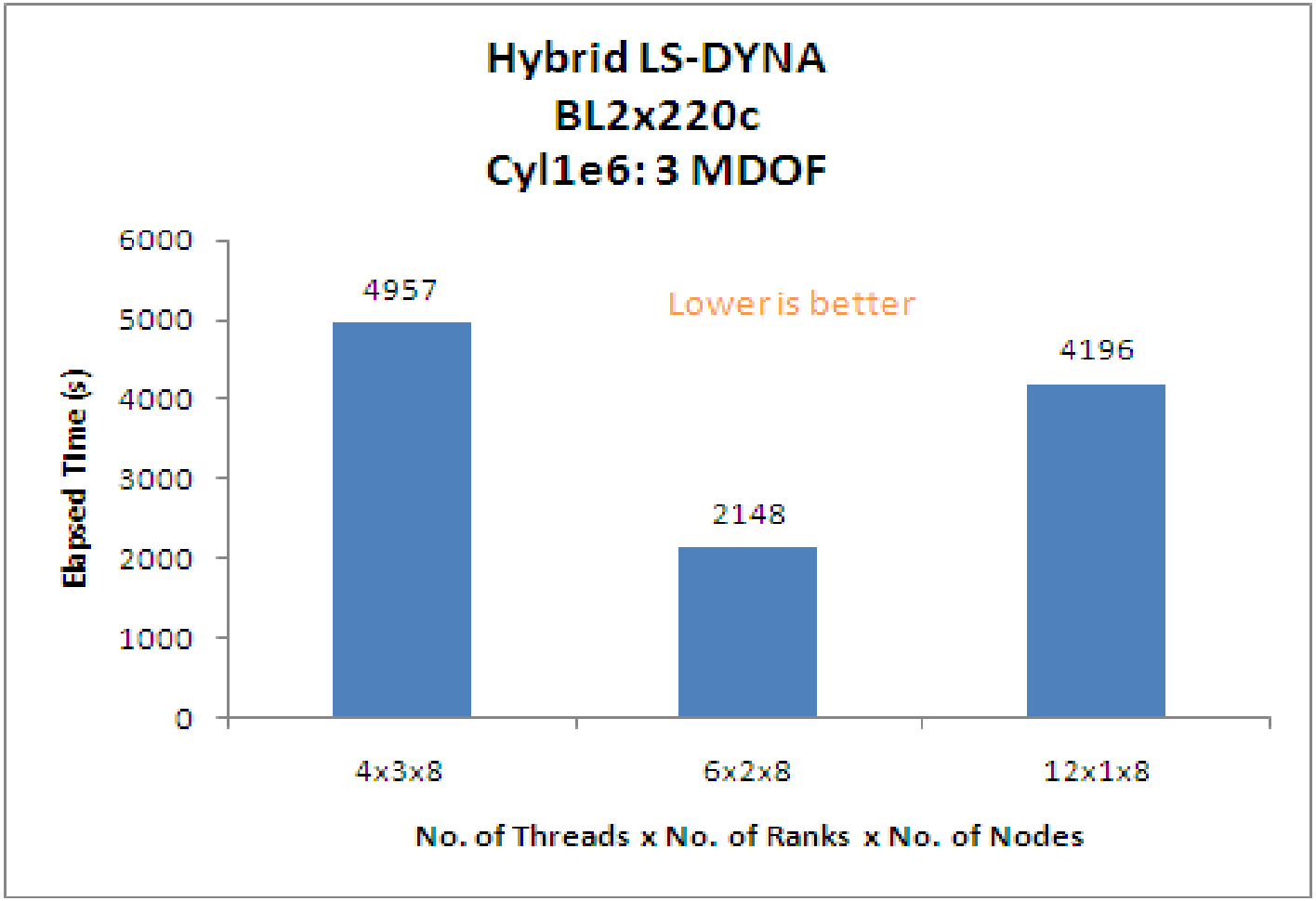


Process Placement

Simplified Intel X5670 Processor

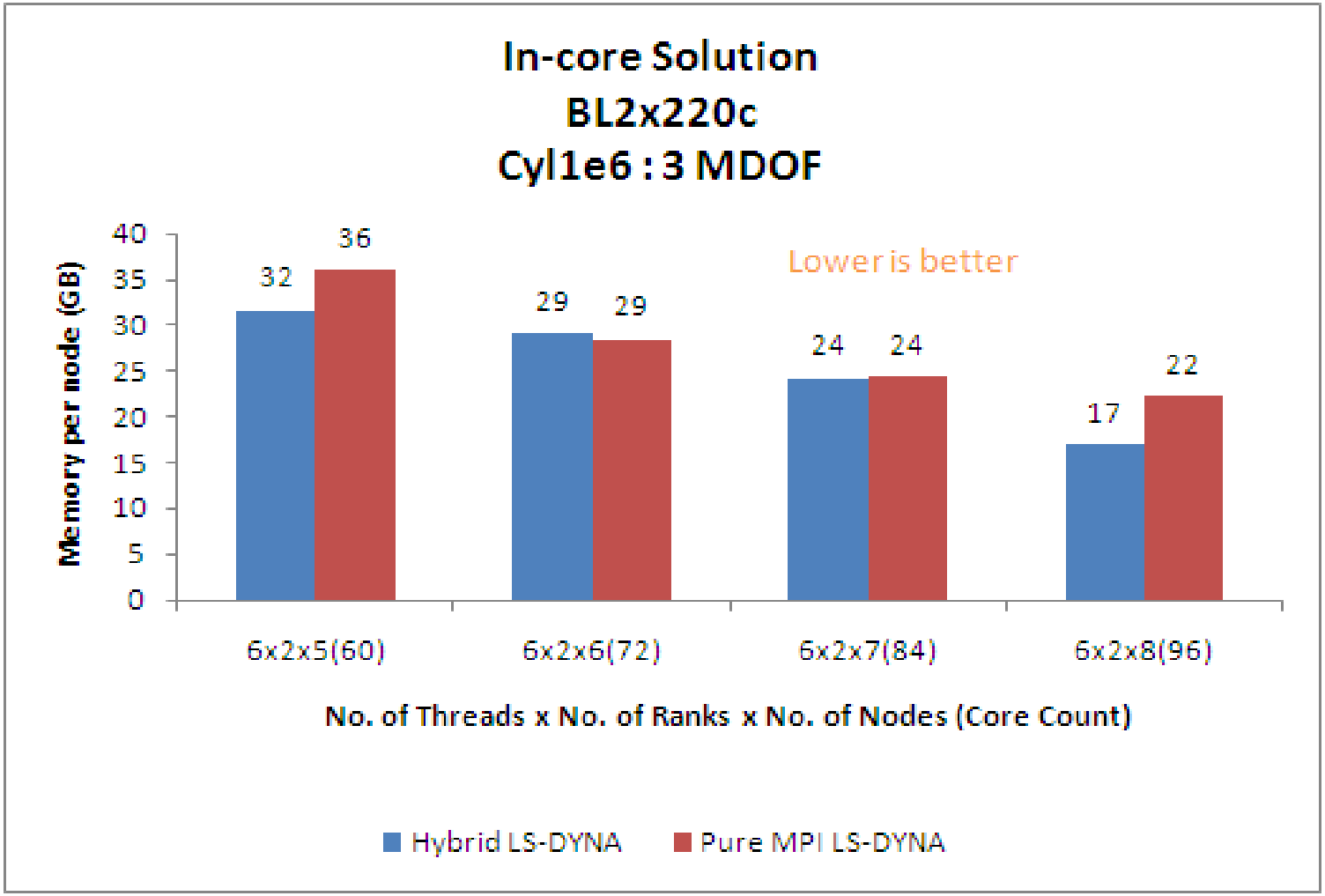


Effects of Varying Number of SMP Threads

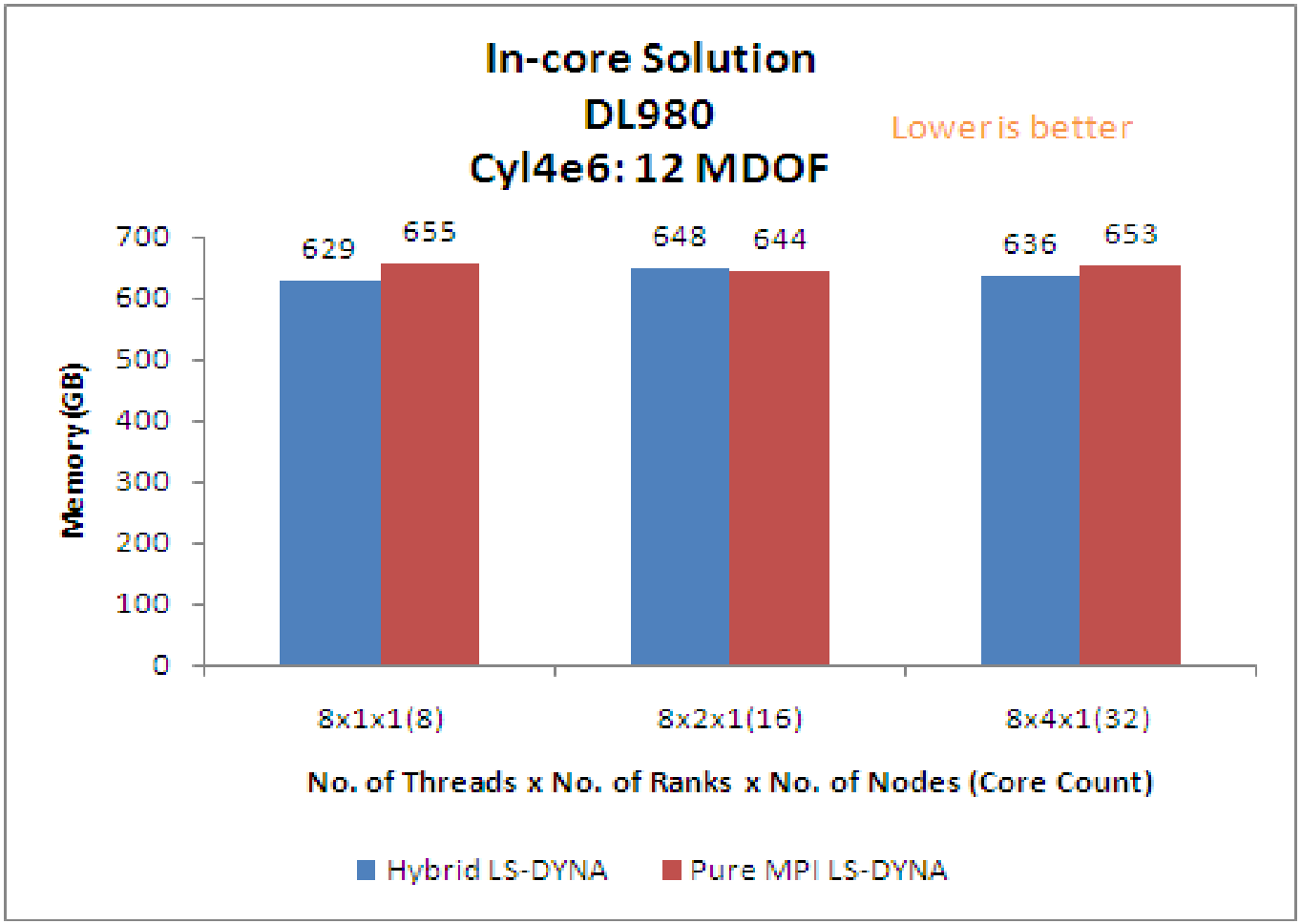


Memory Requirement

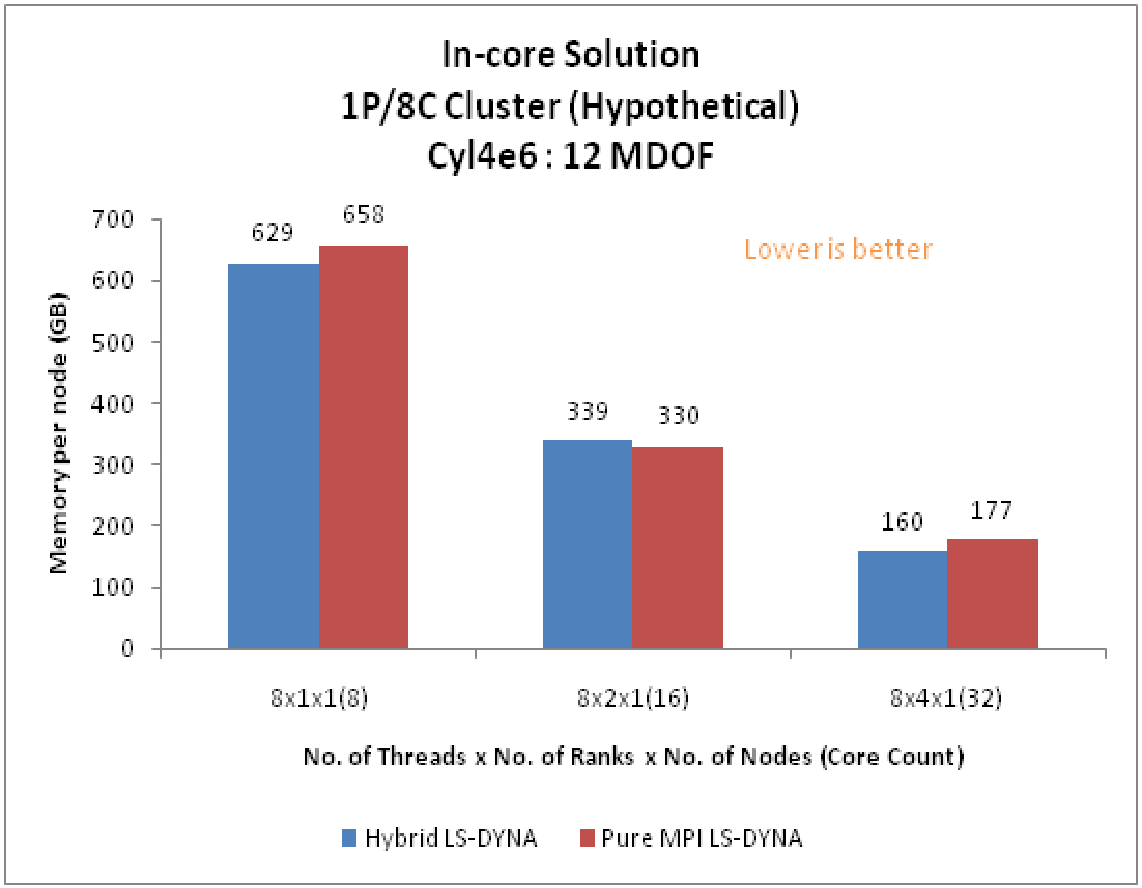
Minimal Per-node Memories for In-core Solutions



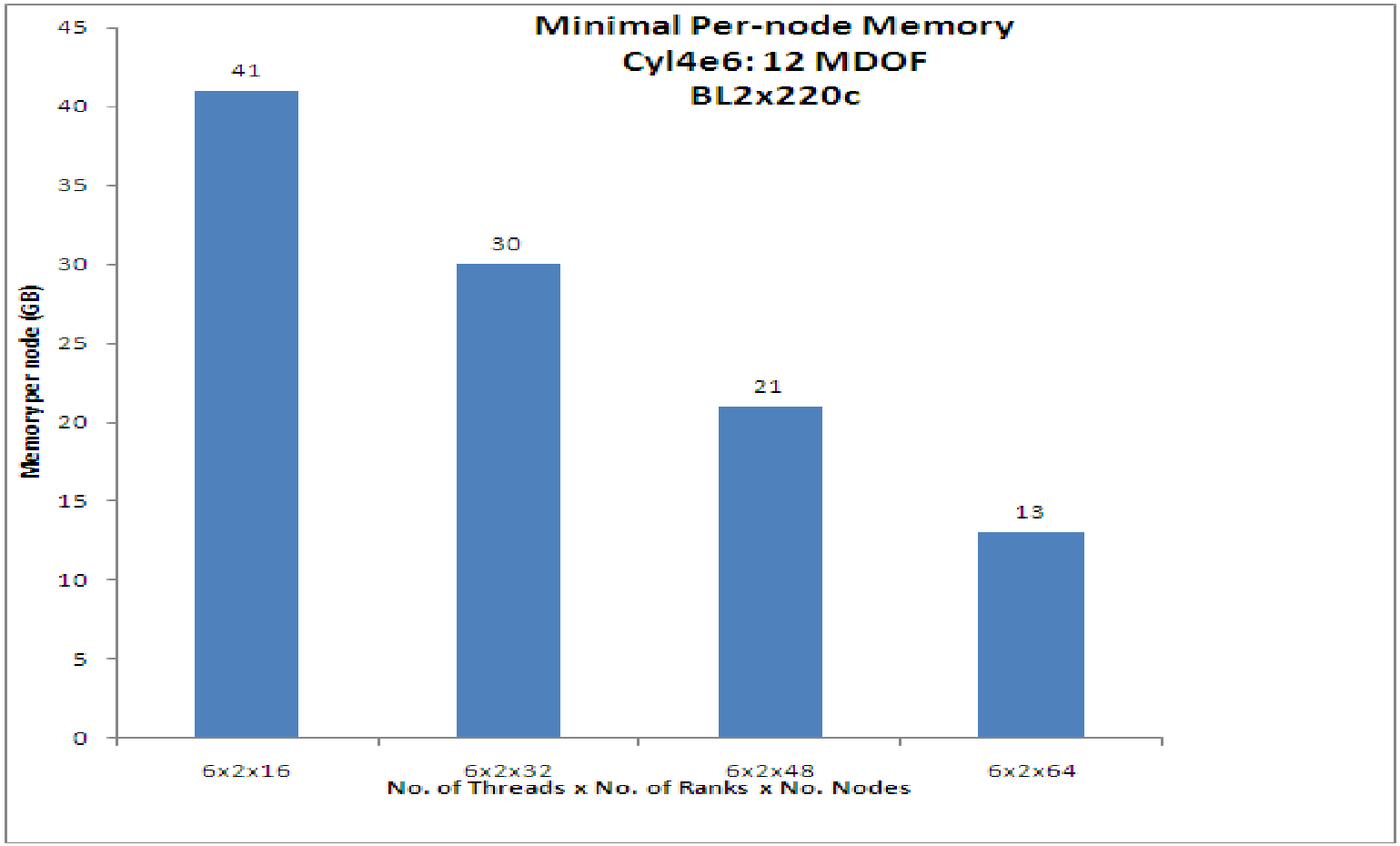
Minimal Per-node Memories for In-core Solutions



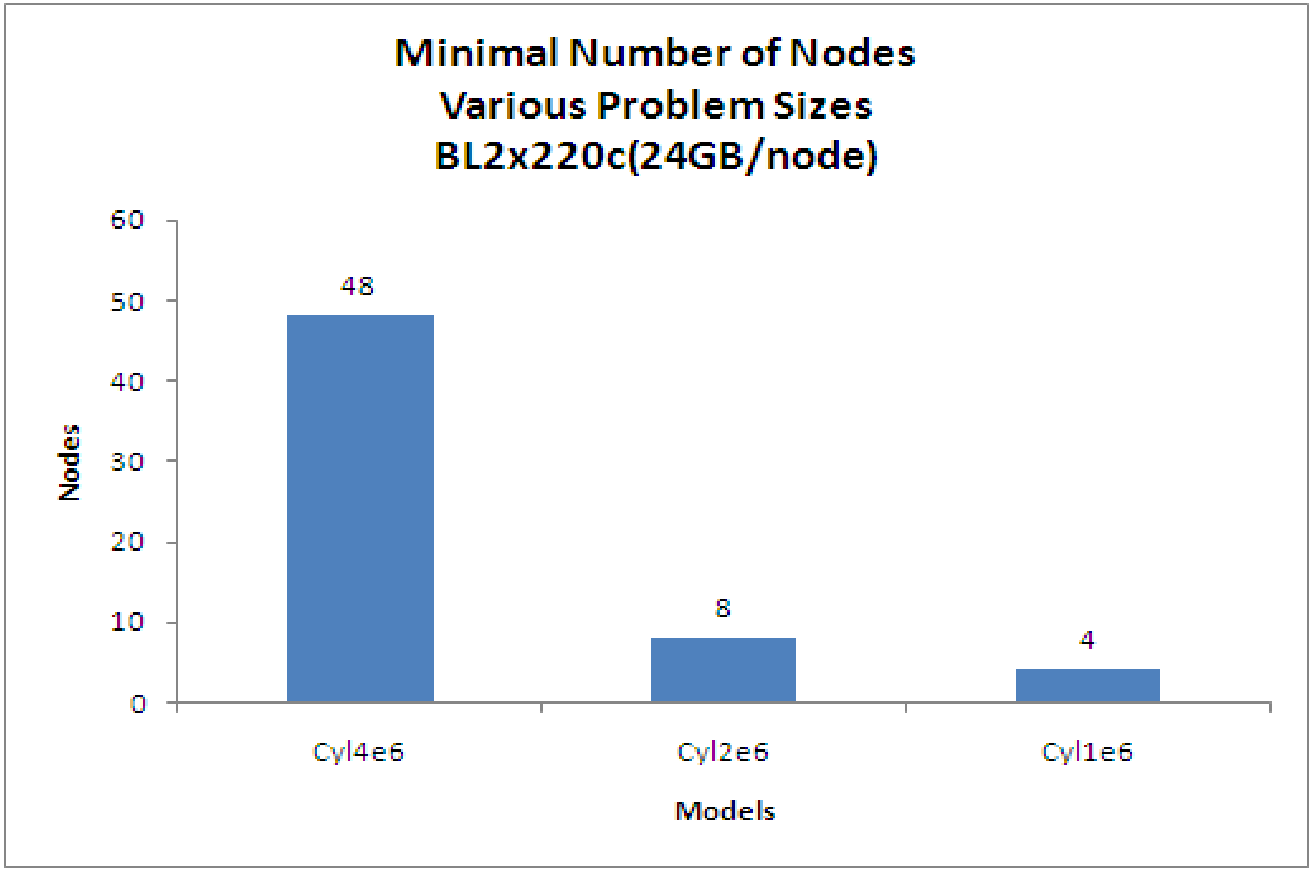
Minimal Per-node Memories for In-core Solutions



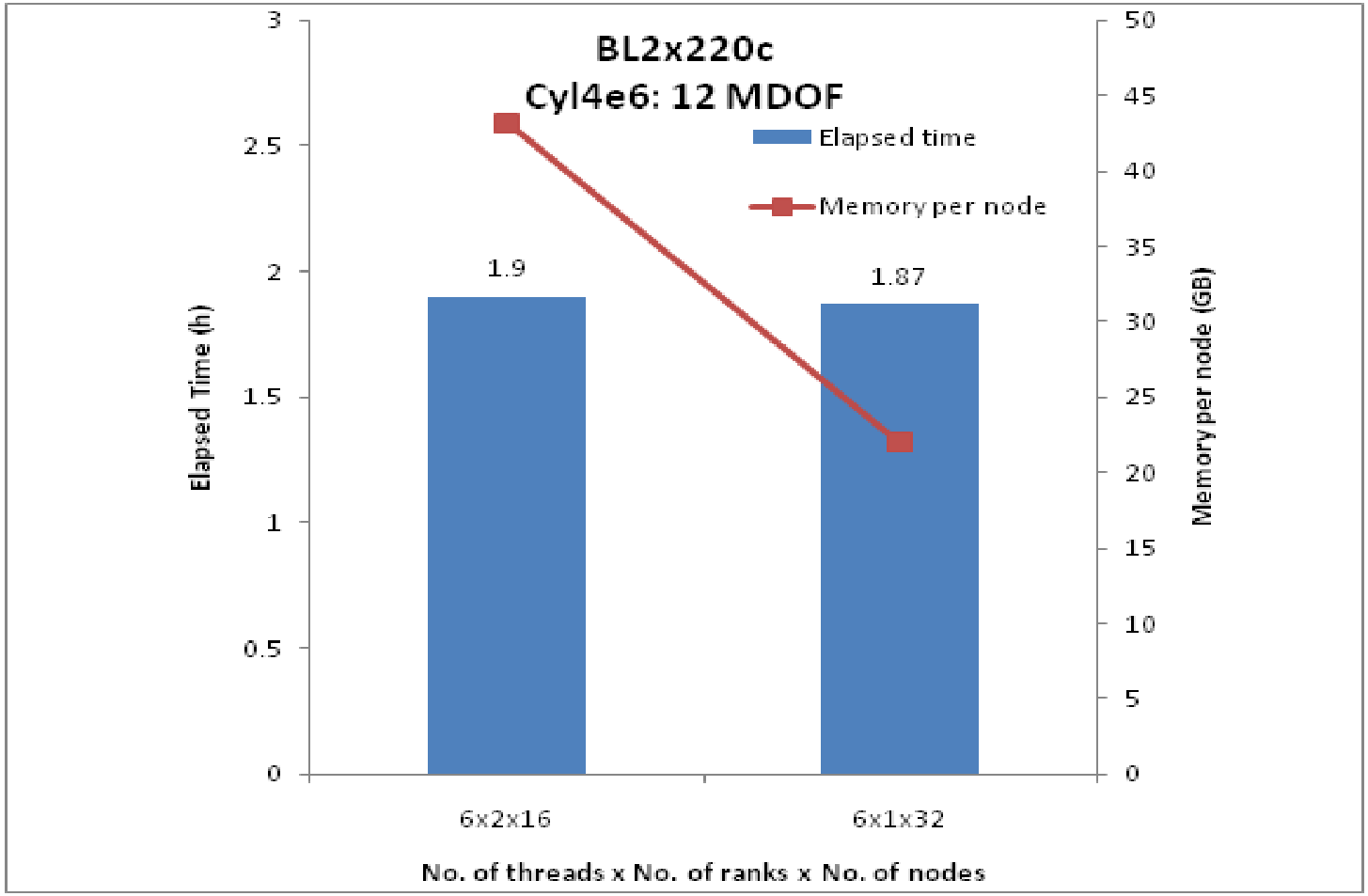
Minimal Per-node Memories for In-core Solutions with Various Number of Nodes



Minimal Number of Nodes for In-core Solutions with Various Problem Sizes



Elapsed Time and Per-node Memory Requirement: Full vs. Half Loaded, Same Core Count at192



Cluster Network Interconnects

Gigabit Ethernet

– Latency : 43 µs

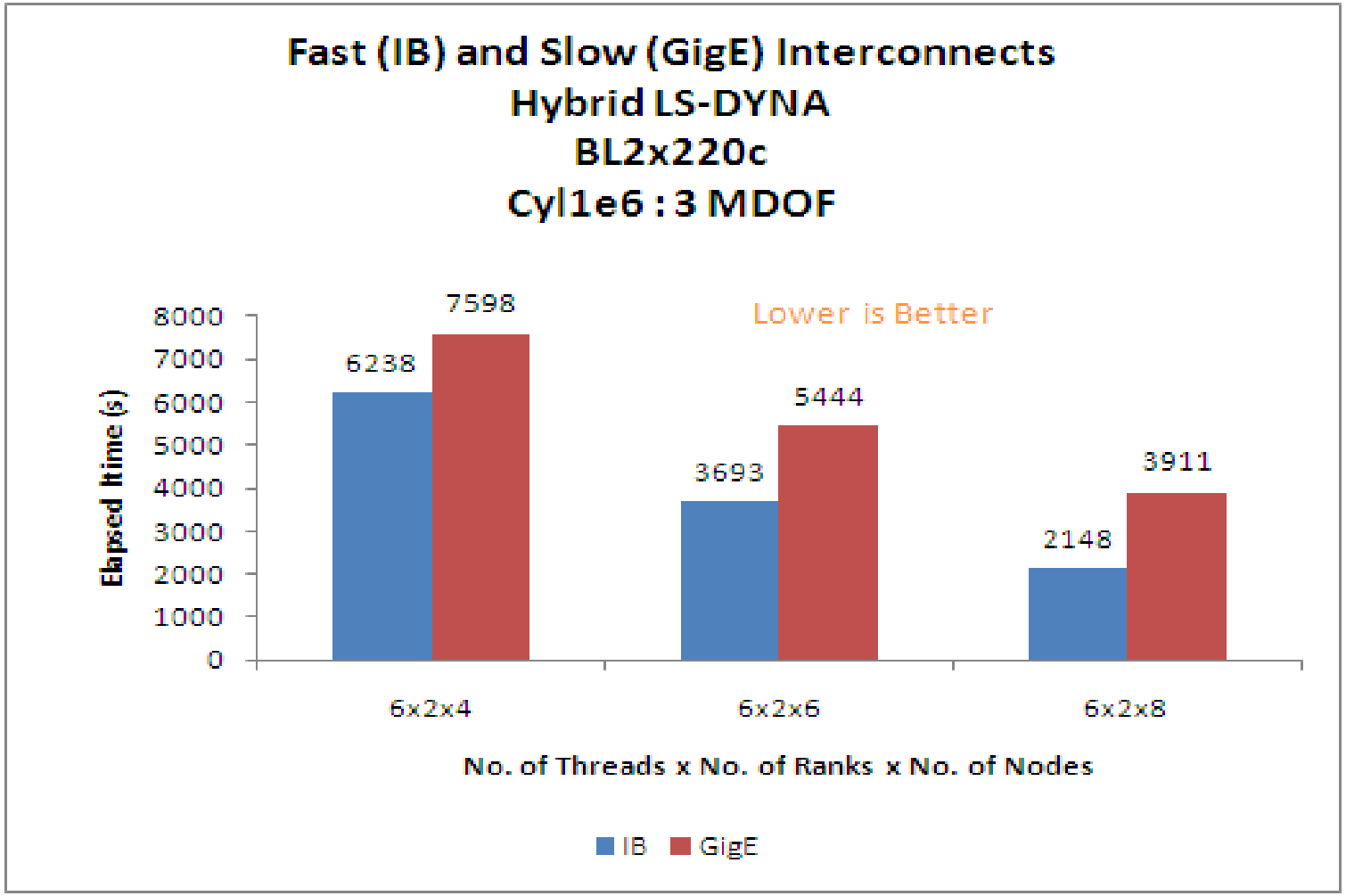
– Bandwidth: 128 MB/s

InfiniBand (QDR)

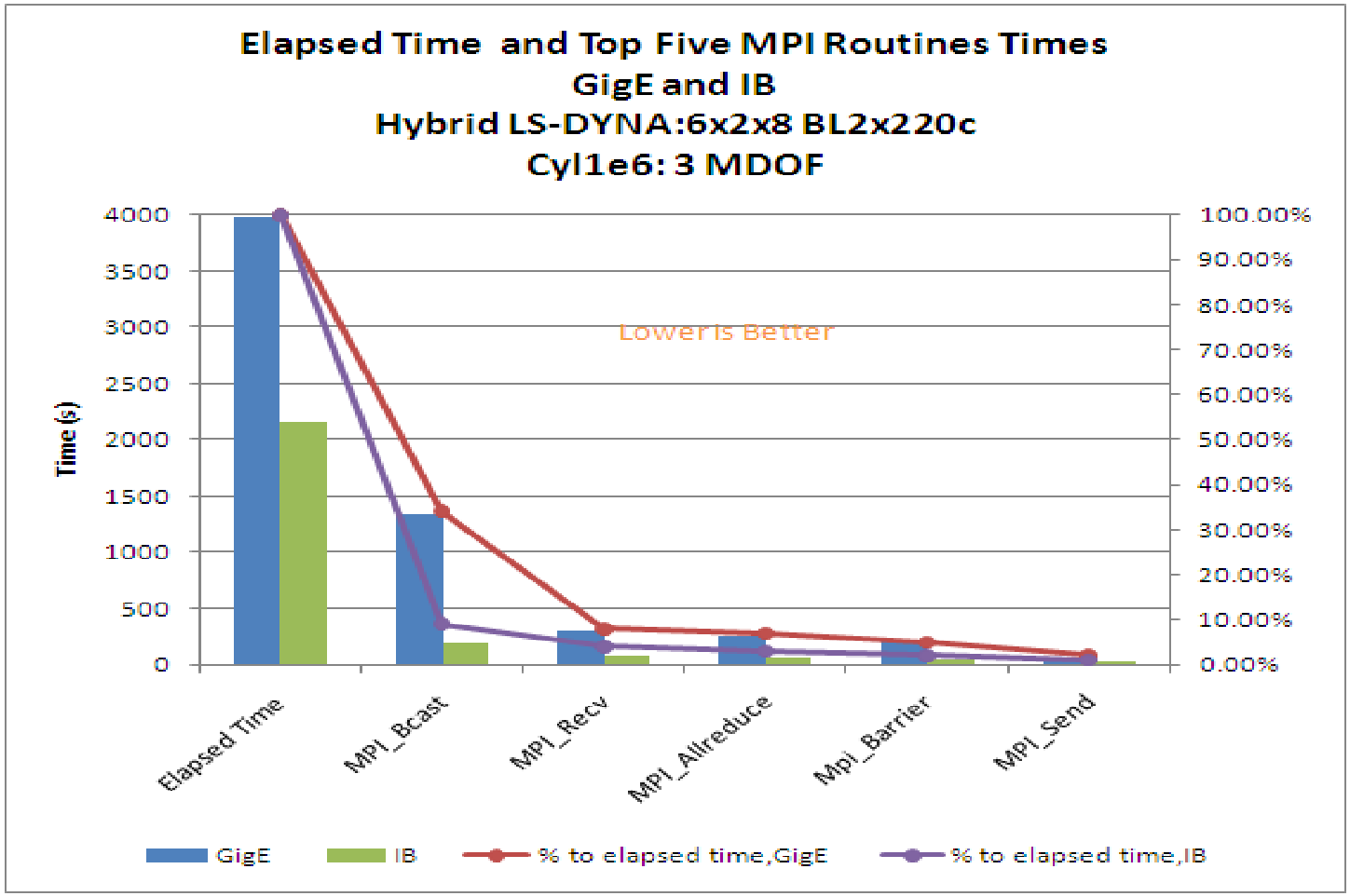
– Latency: < 2 µs

– Bandwidth: 3 GB/s

IB versus GigE Performance

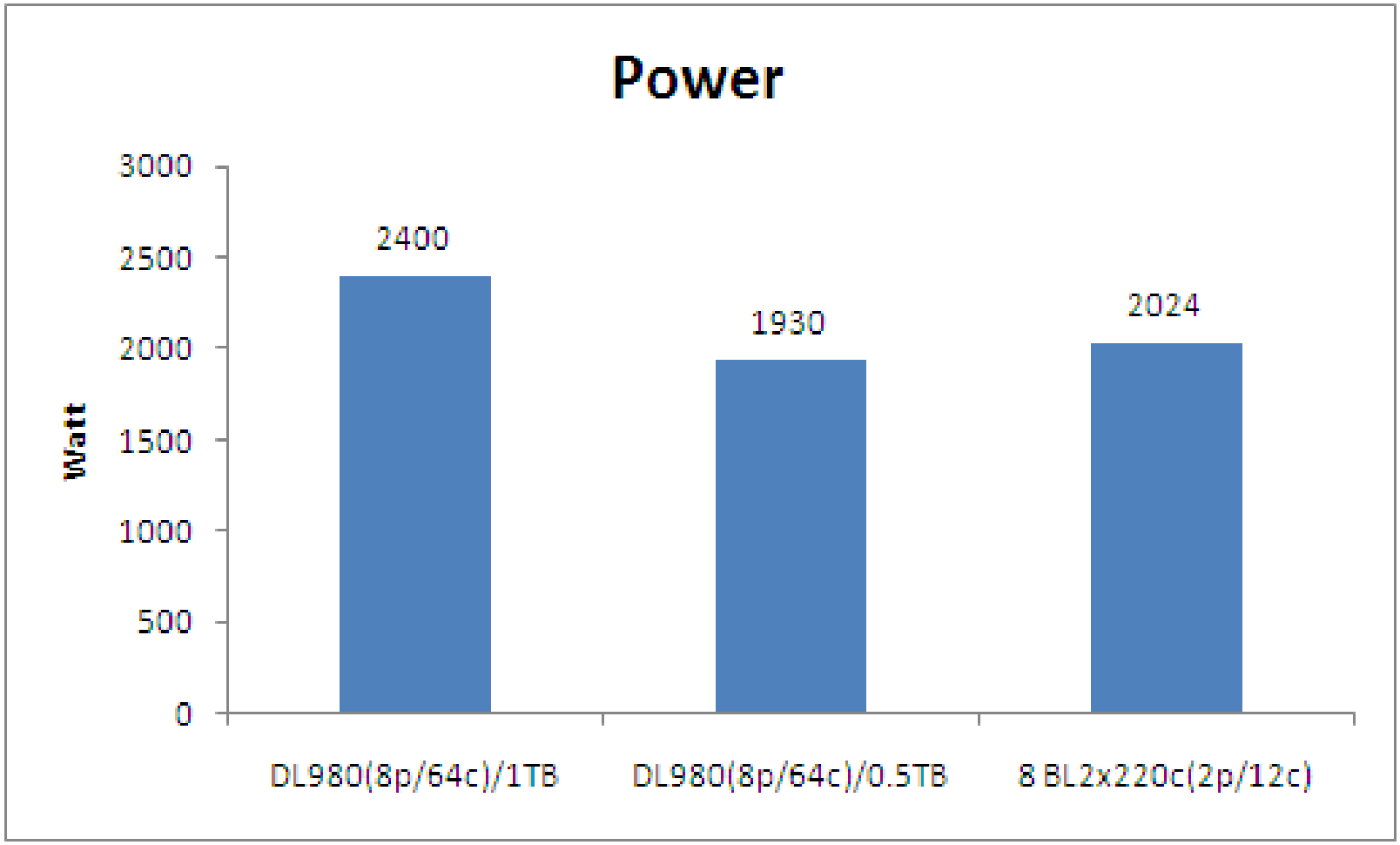


Top Five MPI Routines Times vs. Elapsed Time

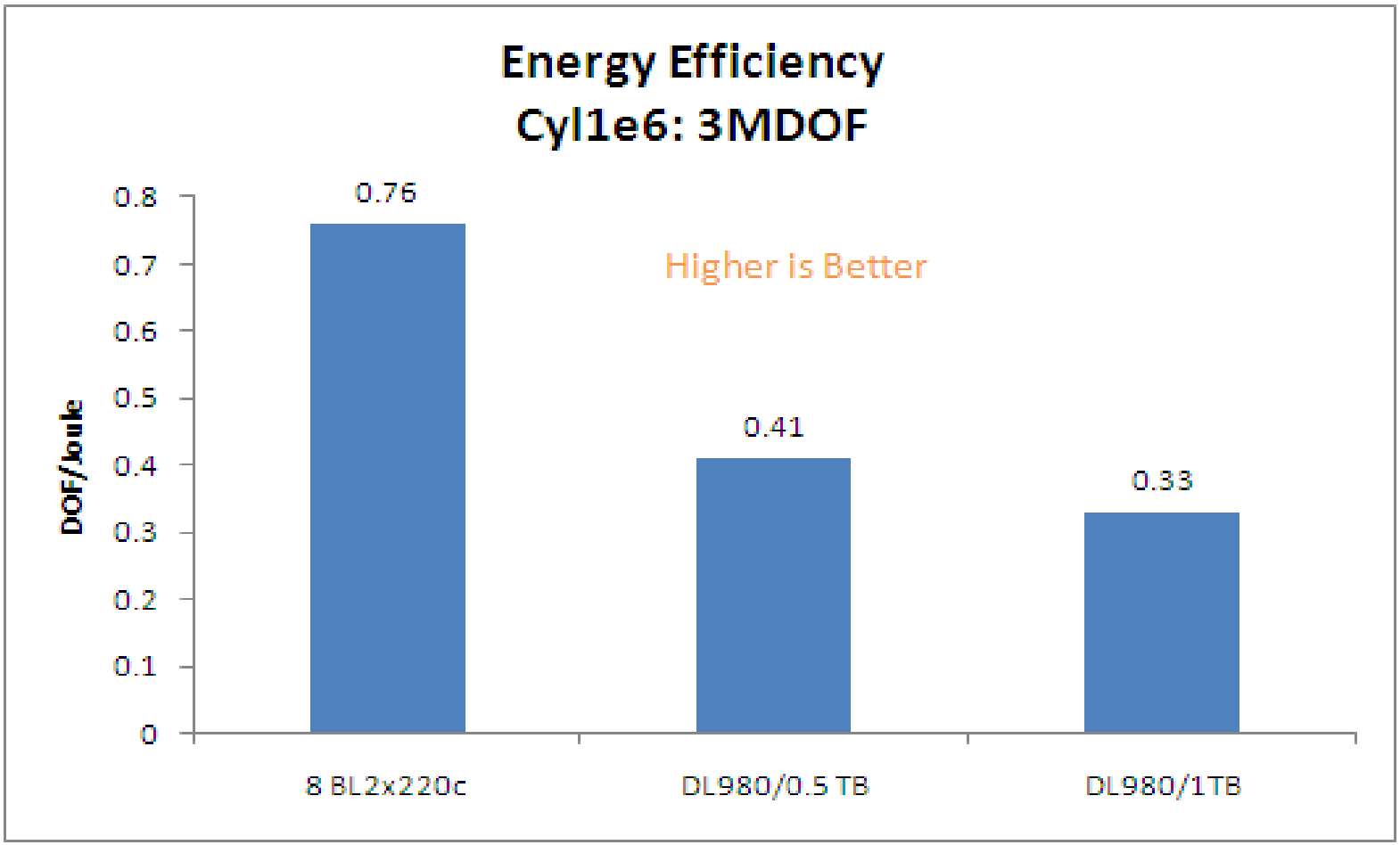


Power and Other Costs

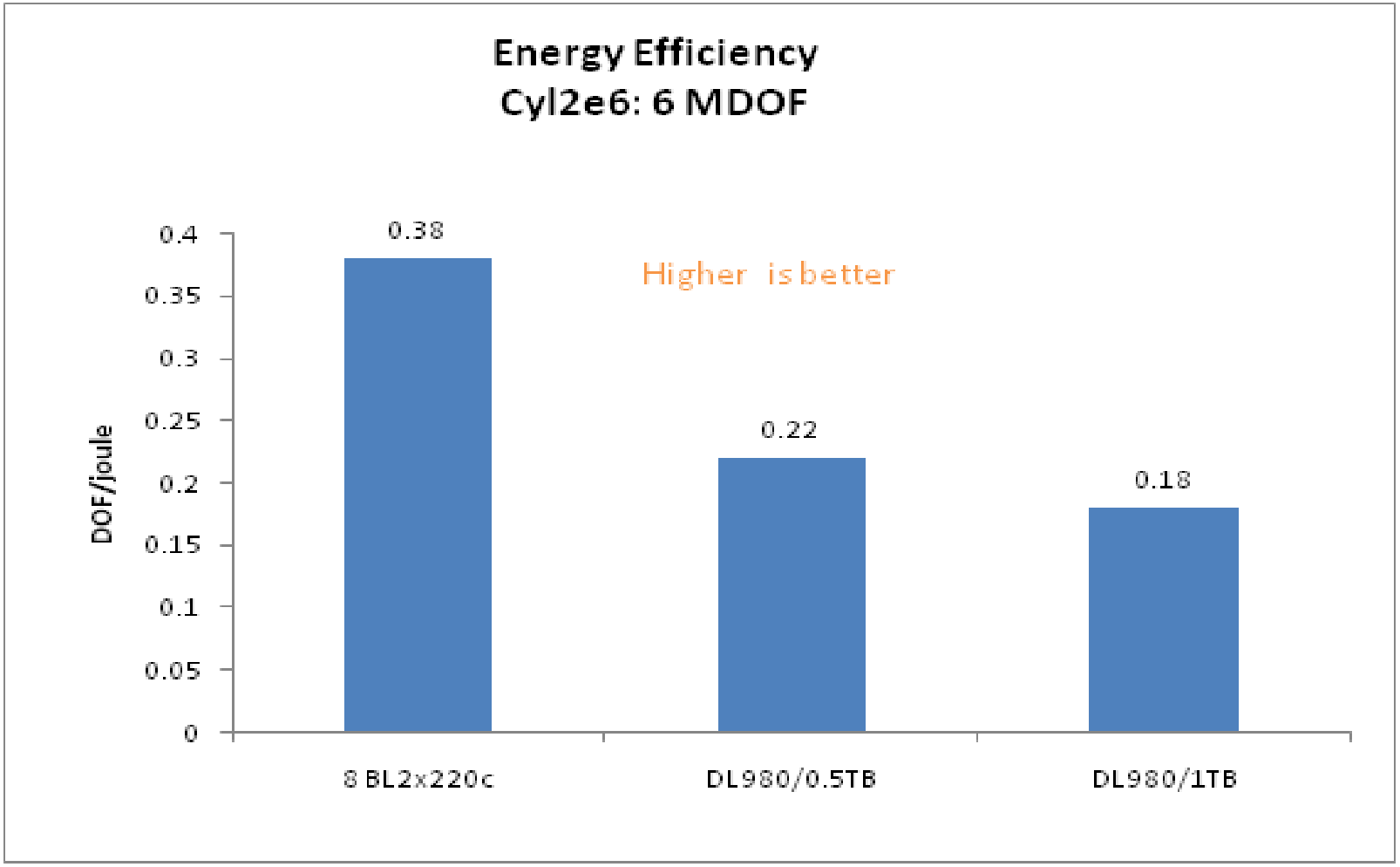
Power Usages



Energy Efficiency: 3 MDOF



Energy Efficiency: 6 MDOF



Rack Space

	Vertical Space of Standard Rack (cm)	No. of Units	Total Vertical Space (cm)
DL980 Rack	35.36	8	283
C7000 Server Enclosure	18.17	10	182
Ratio between the above two total vertical spaces			1.6

Conclusion

– Both Hybrid and pure MPI LS-DYNAs scale well for implicit analysis.

– The optimal number of threads per rank for Hybrid LS-DYNA is the number of cores that share the same cache in a processor.

– Hybrid LS-DYNA in general requires less memory for in-core solutions than Pure MPI LS-DYNA.

– Tow ways to meet memory requirement for Hybrid LS-DYNA on a cluster:

  • Decreasing the number of SMP threads per rank
  • Increasing the number of nodes

– Speed of interconnect is important. InfiniBand is recommended.

– A large shared-memory system, containing more memory per core than a smaller shared-memory system, requires less total number of processors to avoid out-of-core solutions.

– But the latter is more energy efficient than the former.