

GROMACS Performance Benchmark and Profiling

September 2011

- **The following research was performed under the HPC Advisory Council activities**

- Special thanks for: HP, Mellanox



- **For more information on the supporting vendors solutions please refer to:**

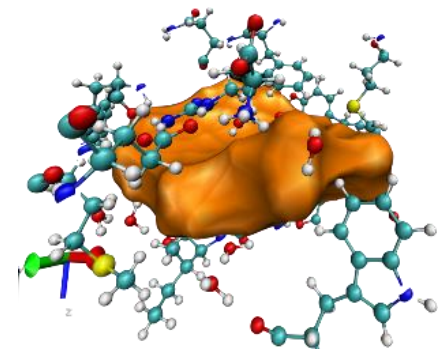
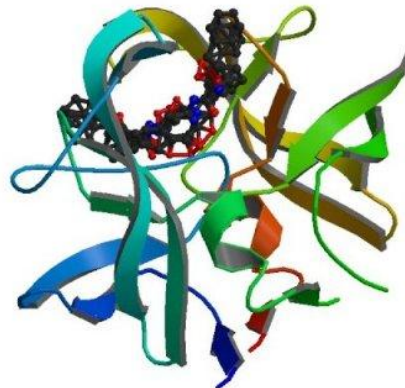
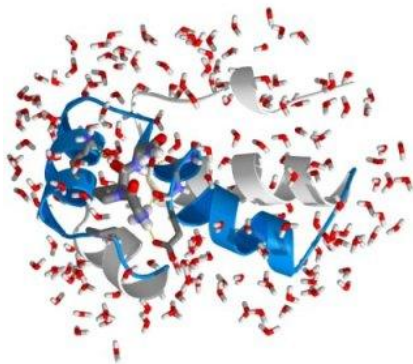
- www.mellanox.com, <http://www.hp.com/go/hpc>

- **For more information on the application:**

- <http://www.gromacs.org>



- **GROMACS (GRoningen MACHine for Chemical Simulations)**
 - A molecular dynamics simulation package
 - Primarily designed for biochemical molecules like proteins, lipids and nucleic acids
 - A lot of algorithmic optimizations have been introduced in the code
 - Extremely fast at calculating the nonbonded interactions
 - Ongoing development to extend GROMACS with interfaces both to Quantum Chemistry and Bioinformatics/databases
 - An open source software released under the GPL



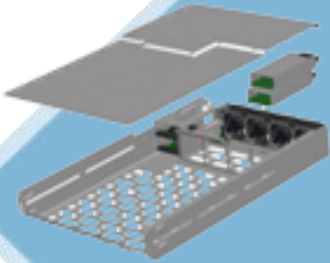
- **The presented research was done to provide best practices**
 - GROMACS performance benchmarking
 - Interconnect performance comparisons
 - MPI performance comparison
 - Understanding GROMACS communication patterns

- **The presented results will demonstrate**
 - The scalability of the compute environment to provide nearly linear application scalability

- **HP ProLiant SL2x170z G6 16-node cluster**
 - Six-Core Intel X5670 @ 2.93 GHz CPUs
 - Memory: 24GB per node
 - OS: CentOS5U5, OFED 1.5.3 InfiniBand SW stack
- **Mellanox ConnectX-2 adapters and switches**
- **Fulcrum based 10GigE switch**
- **MPI: Platform MPI 8.1.1, Intel MPI 4.0.3, Mvapich2-1.6, Open MPI 1.5.3**
- **Application: GROMACS 4.5.4**
- **Benchmark Workload**
 - MD case with 20000 steps

About HP ProLiant SL6000 Scalable System

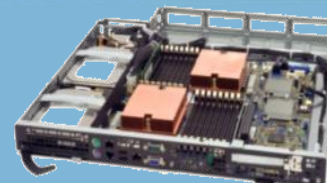
- **Solution-optimized for extreme scale out**



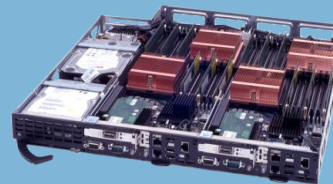
ProLiant z6000 chassis
Shared infrastructure
– fans, chassis, power



ProLiant SL160z G6 ProLiant SL165z G7
Large memory
-memory-cache apps



ProLiant SL170z G6
Large storage
-Web search and database apps



ProLiant SL2x170z G6
Highly dense
- HPC compute and
web front-end apps

Save on cost and energy -- per node, rack and data center

Mix and match configurations

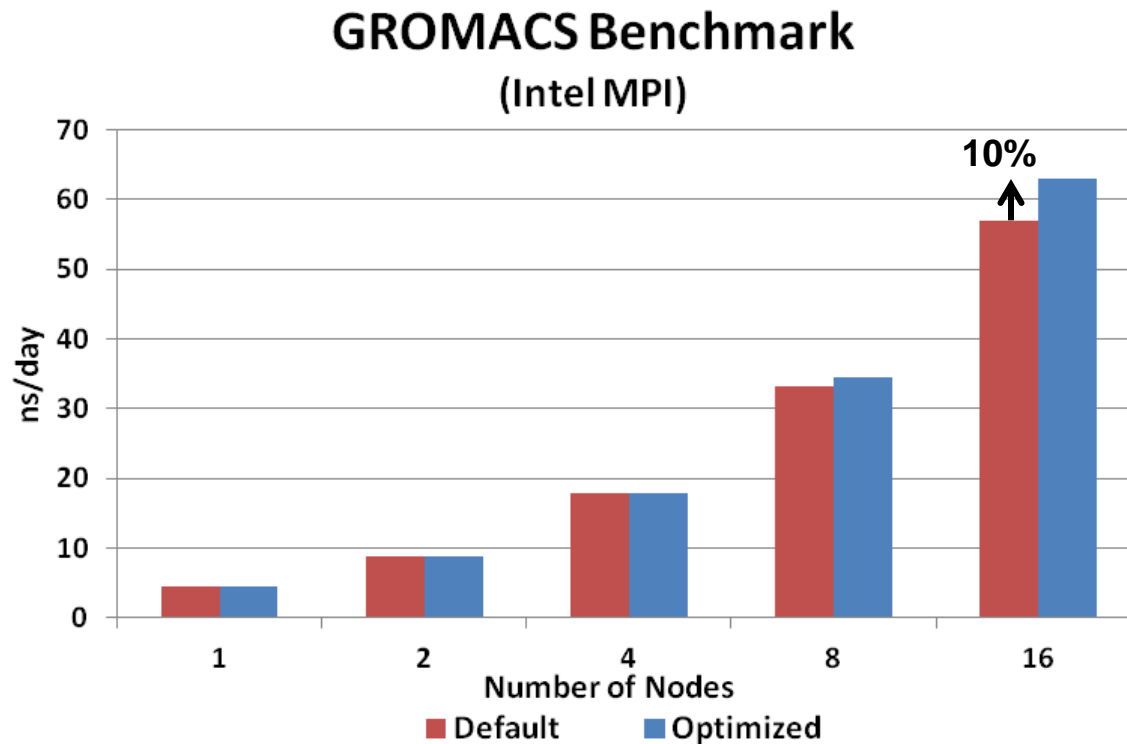
Deploy with confidence

#1
Power
Efficiency*

* SPECpower_ssj2008
www.spec.org
17 June 2010, 13:28

GROMACS Benchmark Result - Tuning

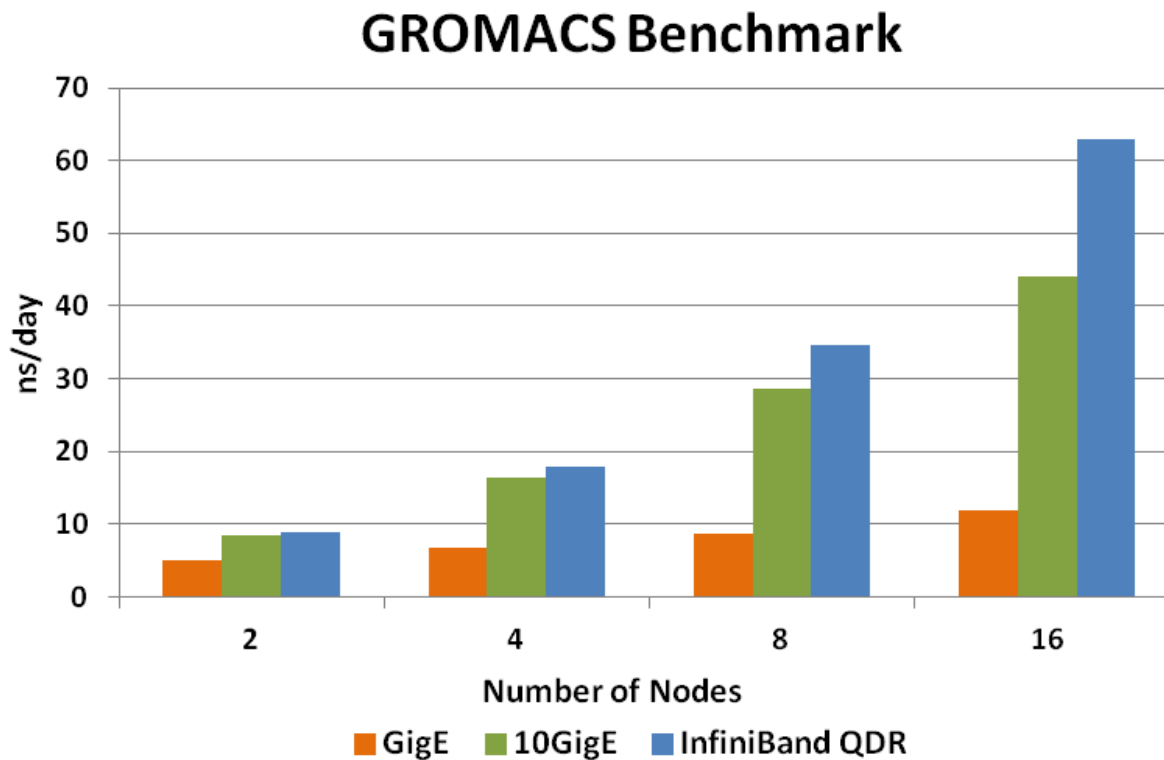
- Intel MPI tuning improves application performance by 10%
 - As cluster size increases, MPI tuning can provide higher advantage



Higher is better

192 cores/16Nodes

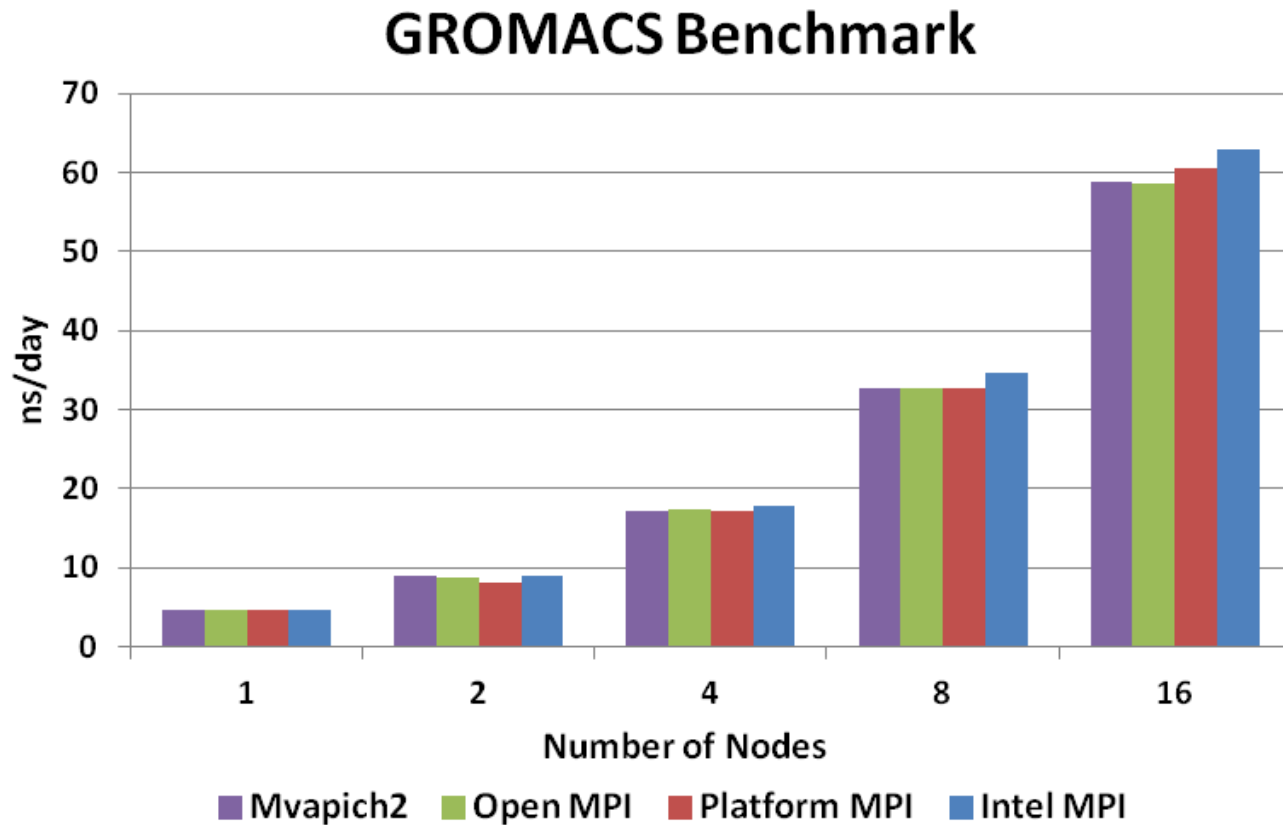
- **InfiniBand QDR enables higher application performance**
 - Up to 43% higher than 10GigE and 430% higher than GigE
 - As cluster size increases, bigger difference is expected



Higher is better

192 cores/16Nodes

- Intel MPI with tuning delivers higher GROMACS performance

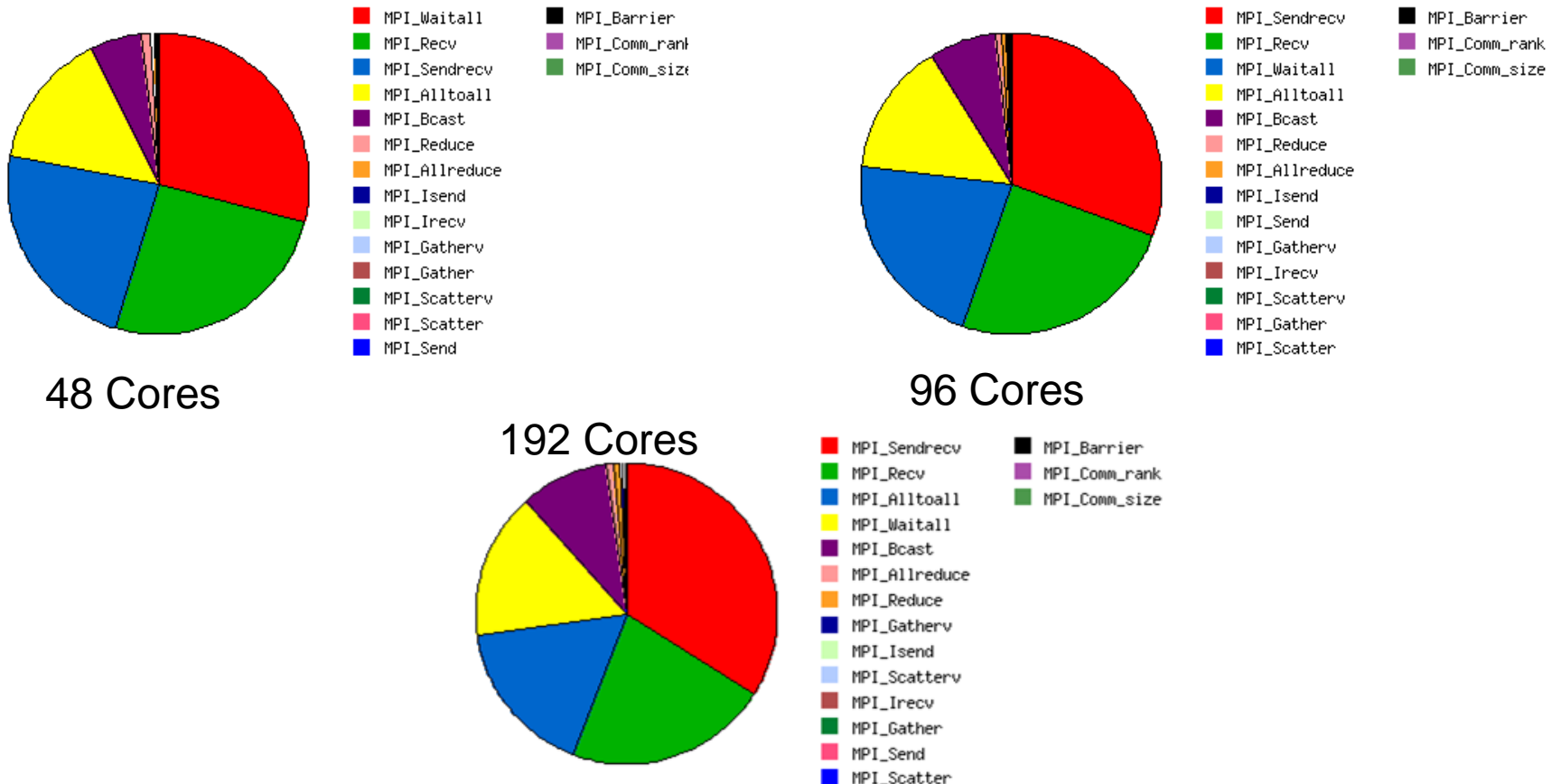


Higher is better

192 cores/16Nodes

- **Mostly used MPI functions**

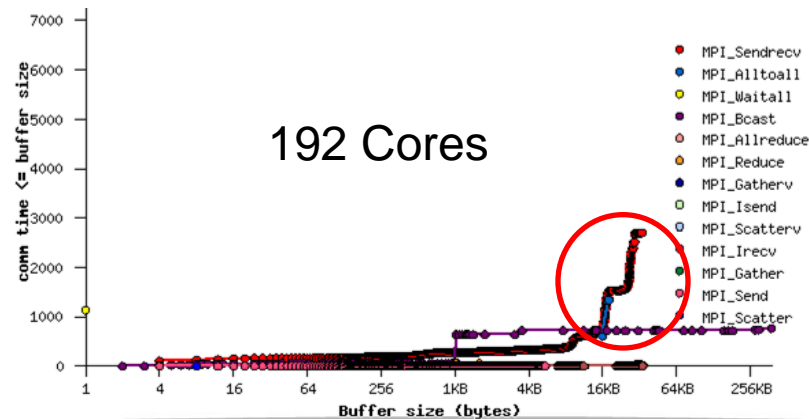
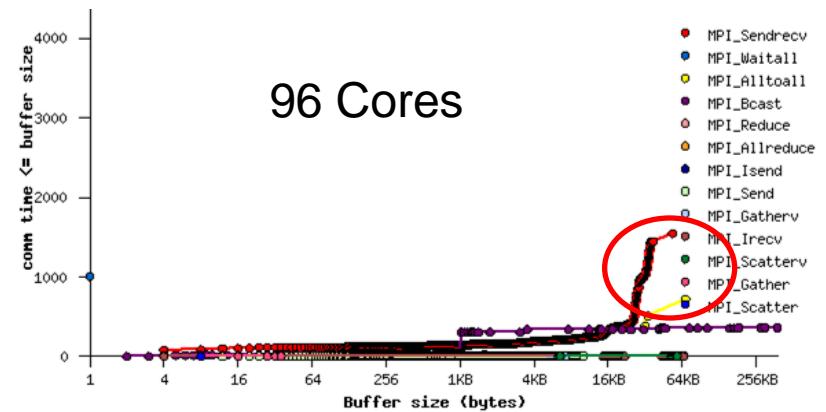
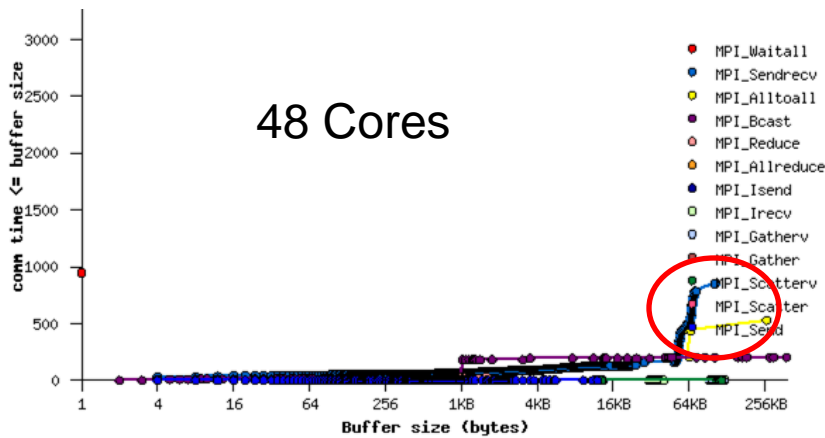
- MPI_Sendrecv, MPI_Recv, MPI_Alltoall, and MPI_Waitall
- Collective overhead increases faster than point-to-point functions
 - MPI_Alltoall and MPI_Bcast



GROMACS Profiling – Message Size

- **As cluster size grows, message size becomes smaller**

- 48 Cores: MPI_Alltoall and MPI_Sendrecv message are typical larger than 48KB
- 96 Cores: MPI_Alltoall and MPI_Sendrecv message are typical between 27KB and 48KB
- 192 Cores: MPI_Alltoall and MPI_Sendrecv message are typical smaller than 27KB



- **GROMACS Profiling**
 - Frequent used message sizes
 - Fairly large message are used at small scale
 - As cluster size increases, messages becomes smaller
 - Frequent used MPI functions
 - Both point-t-point and collectives are heavily used by GROMACS
 - Collectives becomes more critical as cluster size scales
 - Interconnects effect to GROMACS performance
 - Interconnect bandwidth and latency highly influence GROMACS performance
 - Latency becomes more important as cluster size increases
- **GROMACS Performance benchmark**
 - Performance test approved that balanced system (CPU, memory, and Interconnect) 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