

OpenFOAM Performance Benchmark and Profiling

October 2010



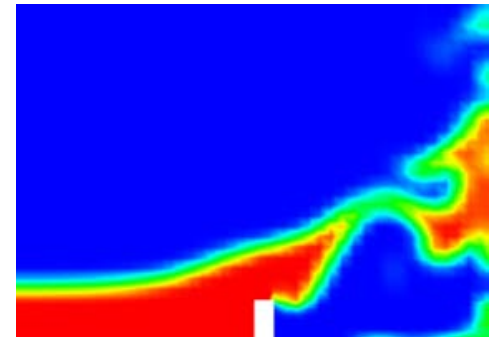
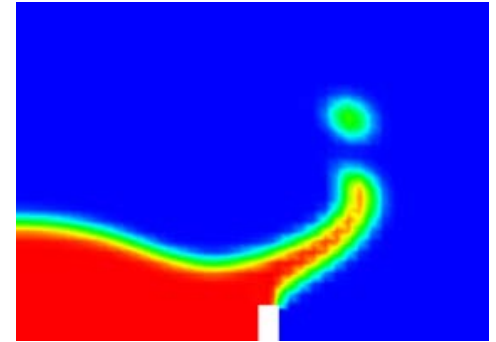
- **The following research was performed under the HPC Advisory Council activities**
 - Participating vendors: Dell, Intel, Mellanox
 - Compute resource: HPC Advisory Council Cluster Center

- **For more info please refer to**
 - <http://www.dell.com>
 - <http://www.intel.com>
 - <http://www.mellanox.com>
 - <http://www.openfoam.com/>

- **OpenFOAM® (Open Field Operation and Manipulation) CFD**

Toolbox can simulate

- Complex fluid flows involving
 - Chemical reactions
 - Turbulence
 - Heat transfer
- Solid dynamics
- Electromagnetics
- The pricing of financial options



- **OpenFOAM is Open source, produced by OpenCFD Ltd**

- **The following was done to provide best practices**
 - OpenFOAM performance benchmarking
 - Interconnect performance comparisons
 - Understanding OpenFOAM communication patterns
 - Power-efficient simulations

- **The presented results will demonstrate**
 - The scalability of the compute environment to provide nearly linear application scalability
 - The capability of OpenFOAM to achieve scalable productivity
 - Considerations for power saving through balanced system configuration

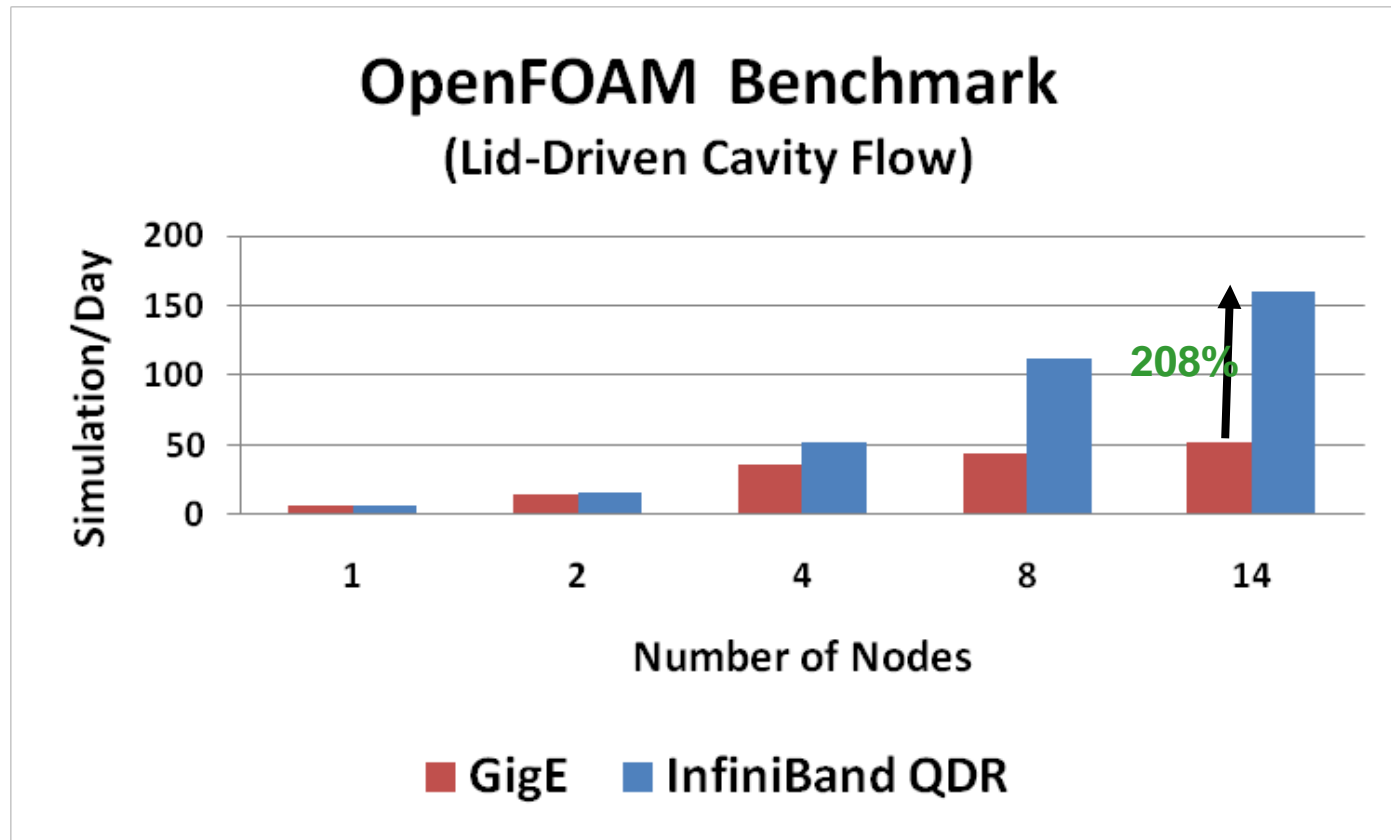
- **Dell™ PowerEdge™ M610 14-node cluster**
 - Six-Core Intel X5670 @ 2.93 GHz CPUs
 - Memory: 24GB per node
 - OS: CentOS5U4, OFED 1.5.1 InfiniBand SW stack
- **Mellanox ConnectX-2 InfiniBand adapters and switches**
- **MPI: Open MPI 1.4.1, Platform MPI 7.1**
- **Application: OpenFOAM 1.7.1**
- **Benchmark Workload**
 - Lid-driven cavity flow (icoFoam, 2D, 1 Million Cells)

- **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



- **InfiniBand enables higher scalability**

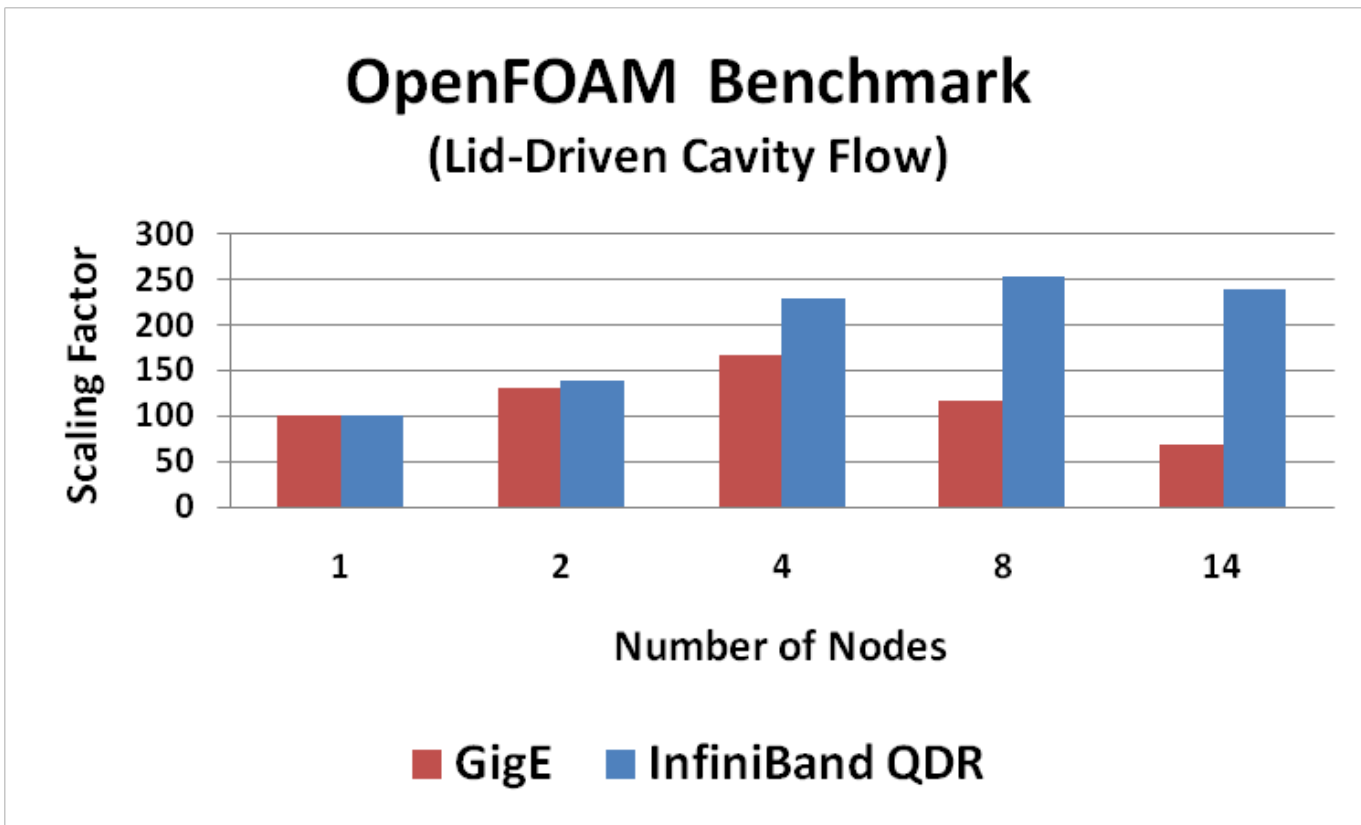
- Up to 208% higher performance than Ethernet at 14 nodes
- 4 InfiniBand connected servers deliver nearly the same performance vs 14 Ethernet connected servers



Higher is better

12 Cores/Node

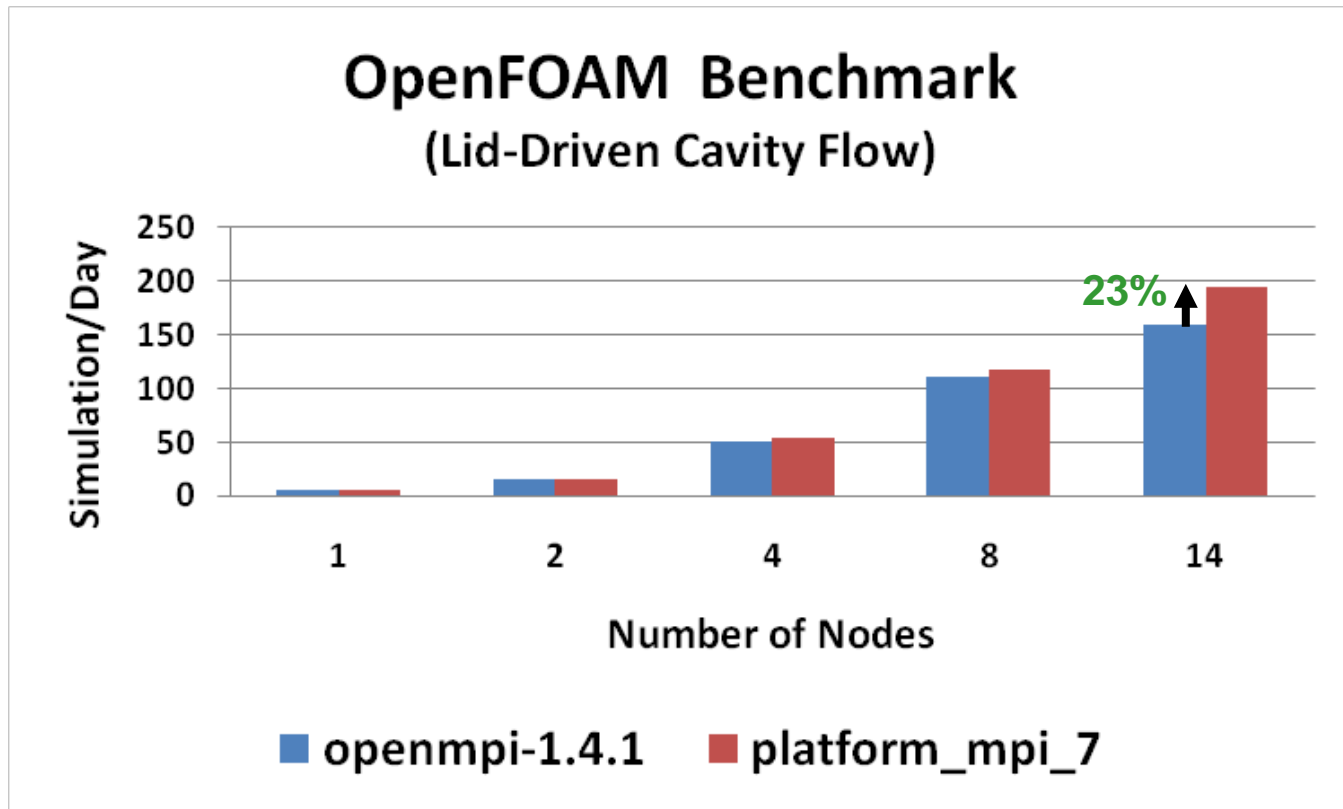
- **InfiniBand enables higher scalability**
 - Performs better as the number of nodes increase
 - Continues to scale as Ethernet scalability drops after 4 servers.



Higher is better

12 Cores/Node

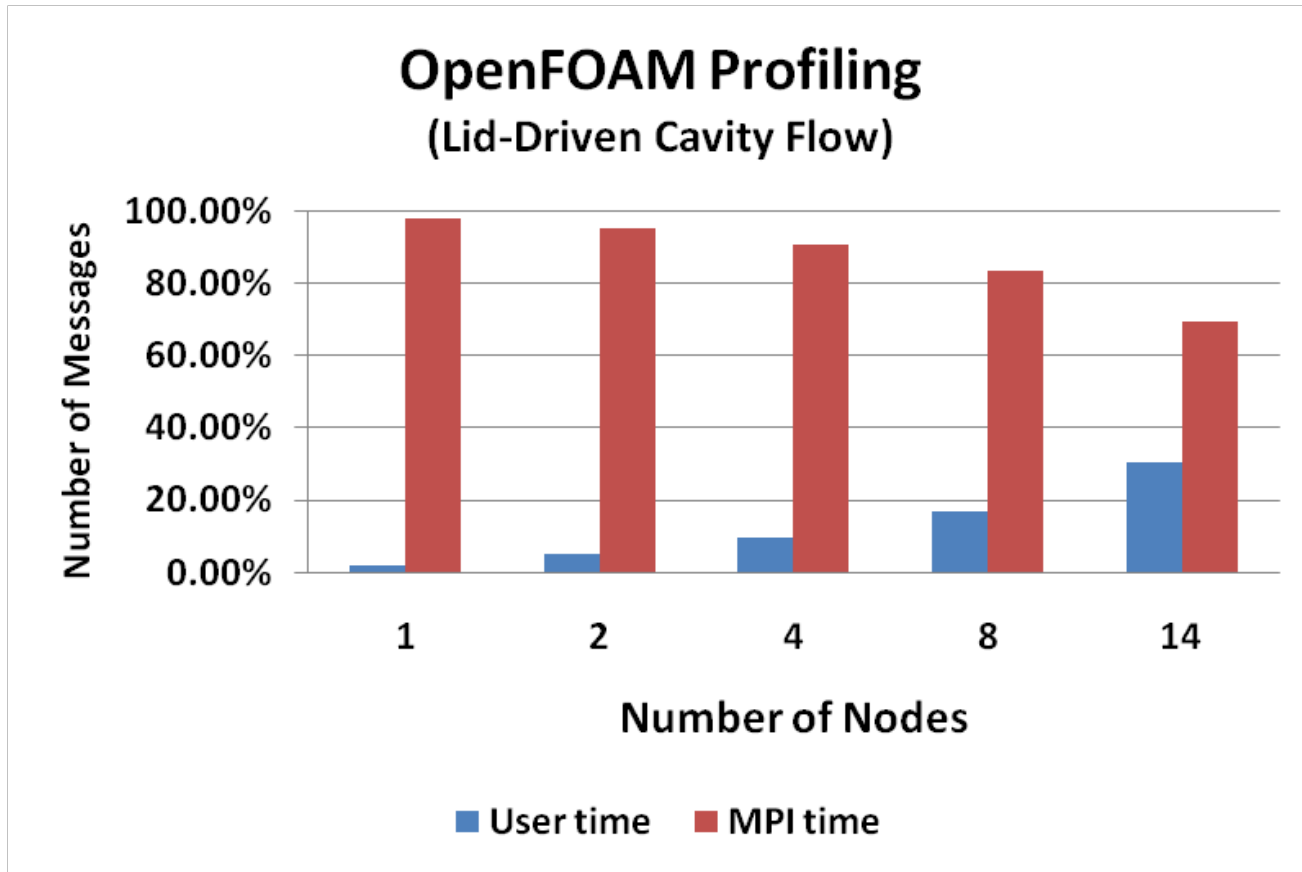
- **Platform MPI demonstrates better performance over Open MPI**
 - 23% gain compared to Open MPI at 14-node
 - Shows dramatic difference especially on larger number of nodes



Higher is better

12 Cores/Node

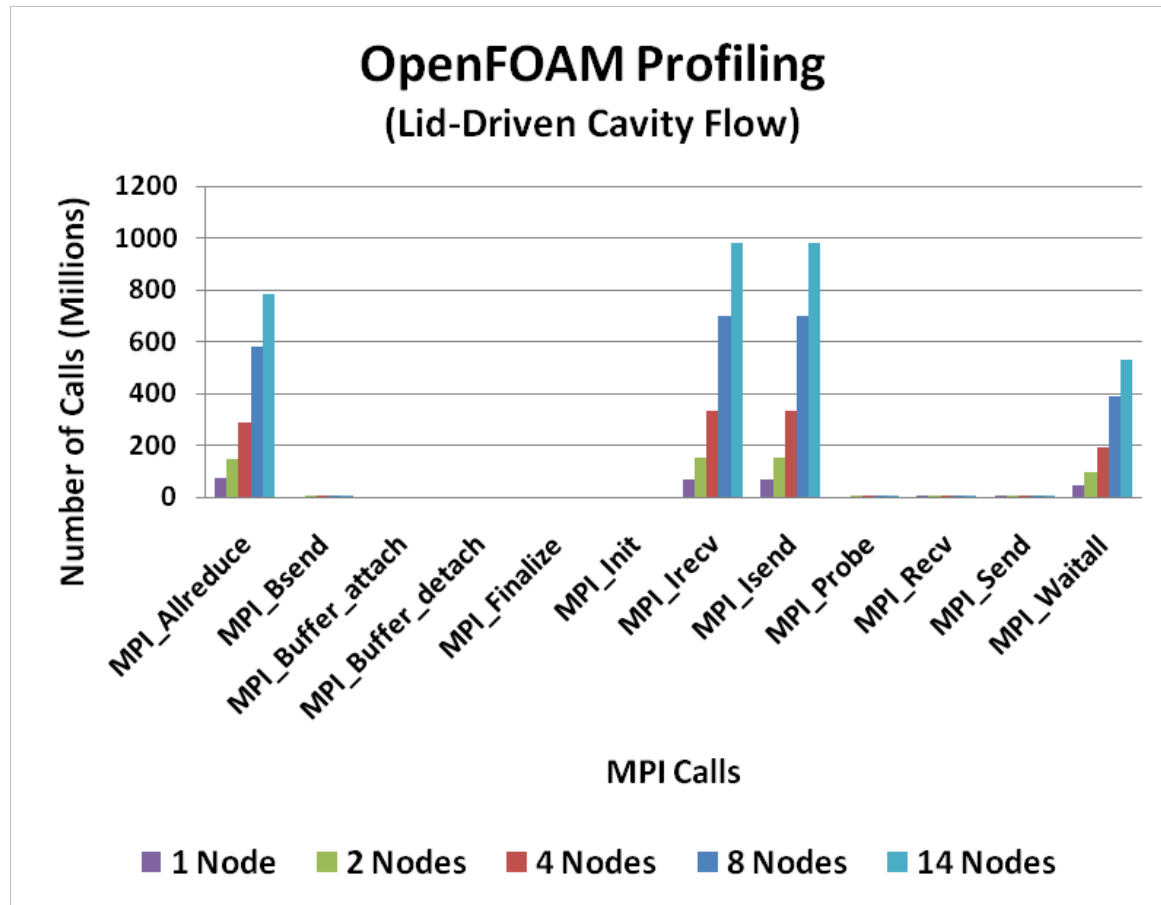
- **The ratio of MPI communication time to user time**
 - decreases as the number of nodes increases



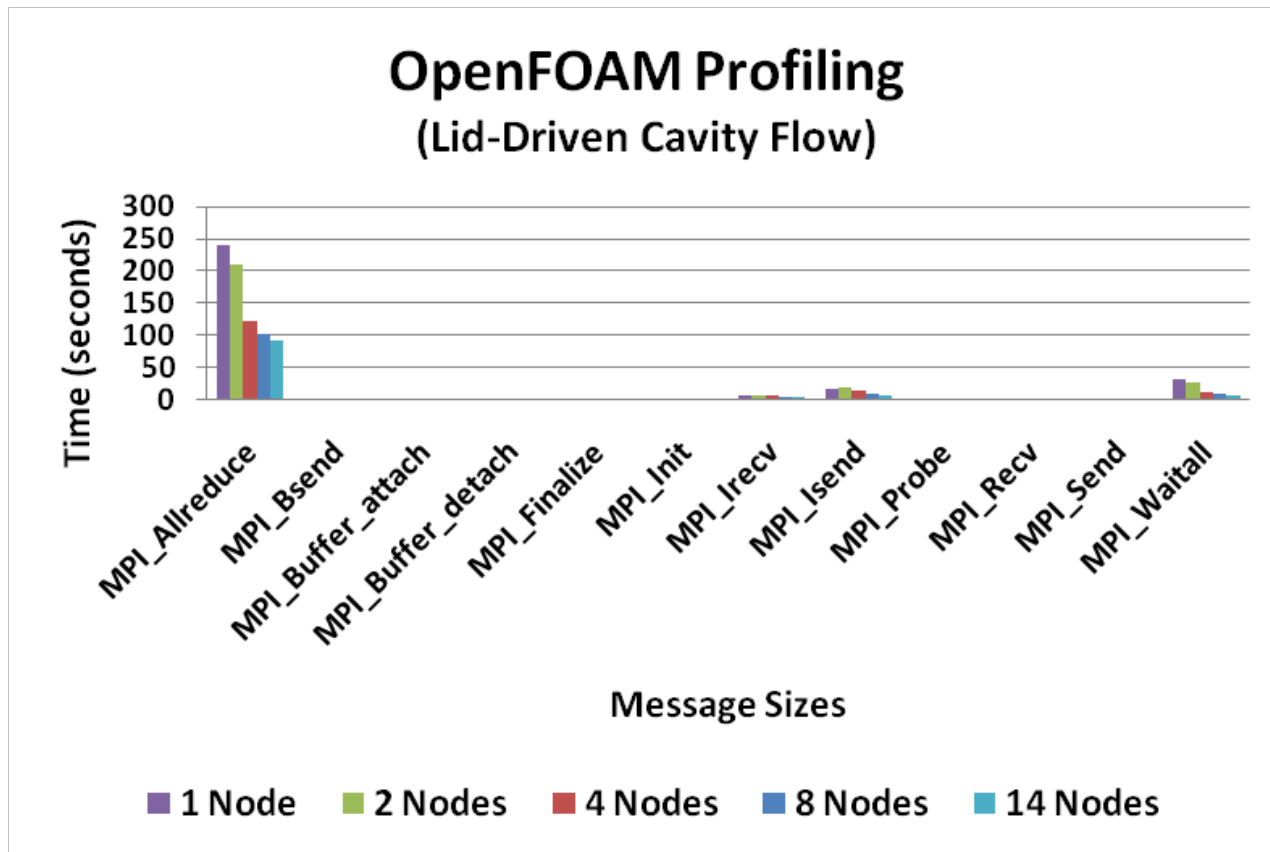
Higher is better

12 Cores/Node

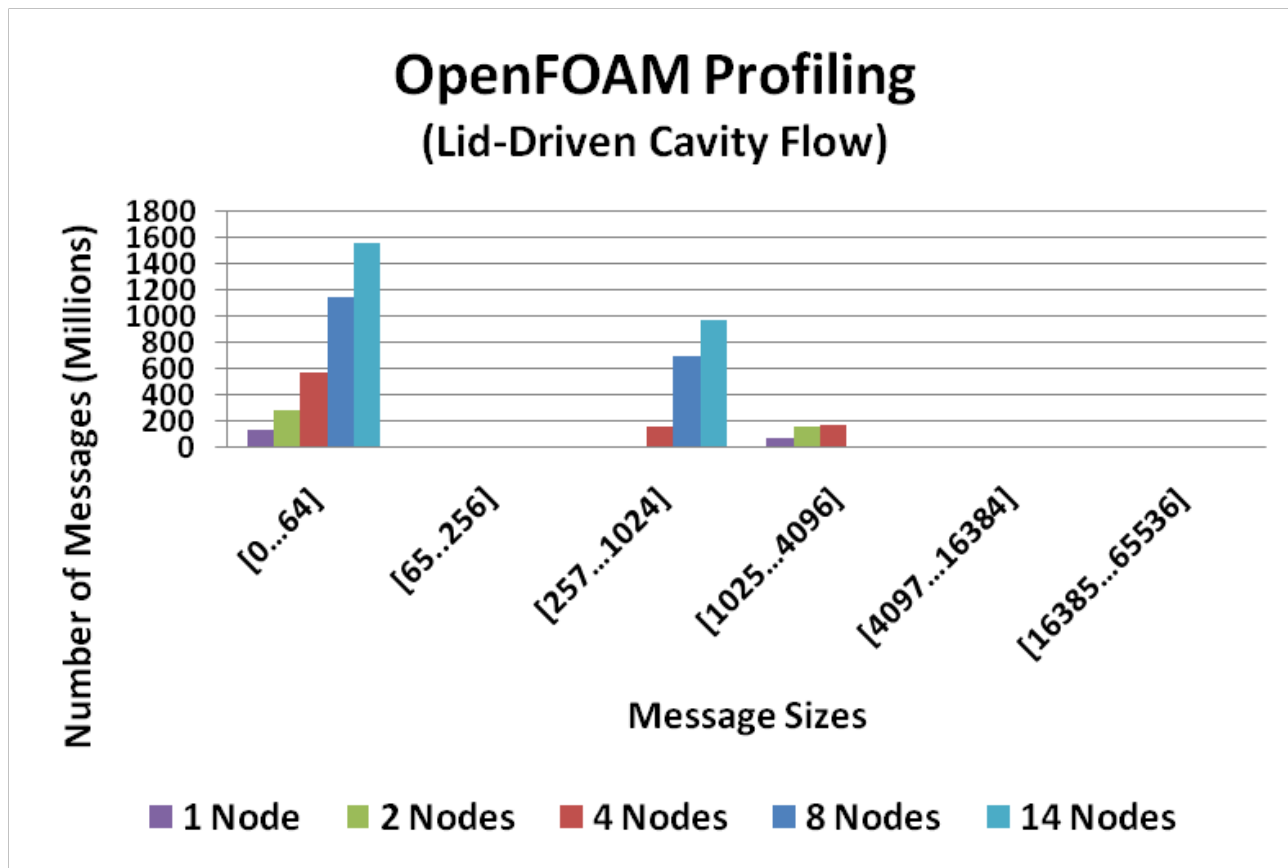
- **MPI_Allreduce, MPI_Irecv, MPI_Isend and MPI_Waitall** are most used calls
 - Percentage time is relatively consistent as number of nodes increases



- **Majority of MPI communication time is spent on MPI_Allreduce**
 - Time is relatively consistent as number of nodes increases



- **Majority of messages are small messages**
 - Messages below 64B and 256 to 1KB are mostly used
- **Number of messages increases dramatically with the number of nodes**



- **Interconnects effect to OpenFOAM performance**
 - InfiniBand enables higher performance/scalability
- **Running in parallel mode enables better system utilization**
- **Platform MPI performs better on larger node count**
- **The ratio of MPI communication over user time decreases as the cluster scales**
- **MPI_Allreduce, MPI_Irecv, MPI_Isend and MPI_Waitall are most used calls**
- **Majority of communication time is spent on MPI_Allreduce**
- **Message with sizes below 64B and between 256B to 1KB mostly used**

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