HPC ADVISORY COUNCIL

ANSYS FLUENT Performance Benchmark and Profiling

May 2009

AMD The future is fusion

ANSYS®

DELL

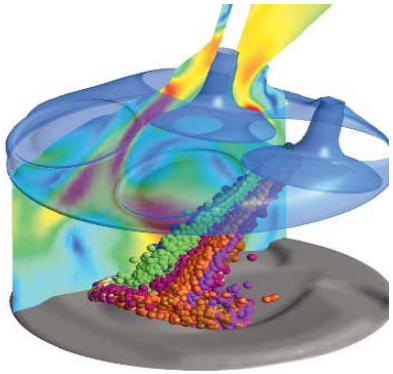
Mellanox TECHNOLOGIES

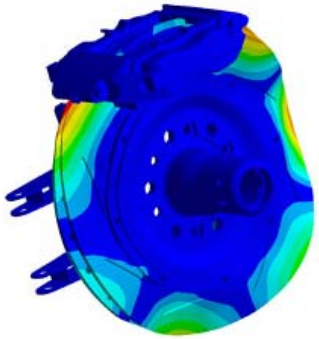
Note

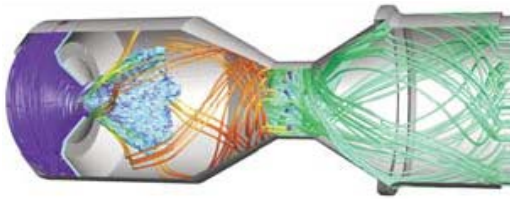
- **The following research was performed under the HPC Advisory Council activities**
  - Participating vendors: AMD, ANSYS, Dell, Mellanox
  - Compute resource - HPC Advisory Council Cluster Center
- **The participating members would like to thank ANSYS for their support and guidelines**
- **For more info please refer to**
  - www.mellanox.com, www.dell.com/hpc, www.amd.com, www.ansys.com

ANSYS FLUENT

- **Computational Fluid Dynamics (CFD) is a computational technology**
  - Enables the study of the dynamics of things that flow
    - By generating numerical solutions to a system of partial differential equations which describe fluid flow
  - Enable better understanding of qualitative and quantitative physical phenomena in the flow which is used to improve engineering design
- **CFD brings together a number of different disciplines**
  - Fluid dynamics, mathematical theory of partial differential systems, computational geometry, numerical analysis, Computer science
- **ANSYS FLUENT is a leading CFD application from ANSYS**
  - Widely used in almost every industry sector and manufactured product







Objectives

- **The presented research was done to provide best practices**
  - ANSYS FLUENT performance benchmarking
  - Interconnect performance comparisons
  - Performance enhancement of the latest FLUENT release
  - Ways to increase FLUENT productivity
  - Understanding FLUENT communication patterns

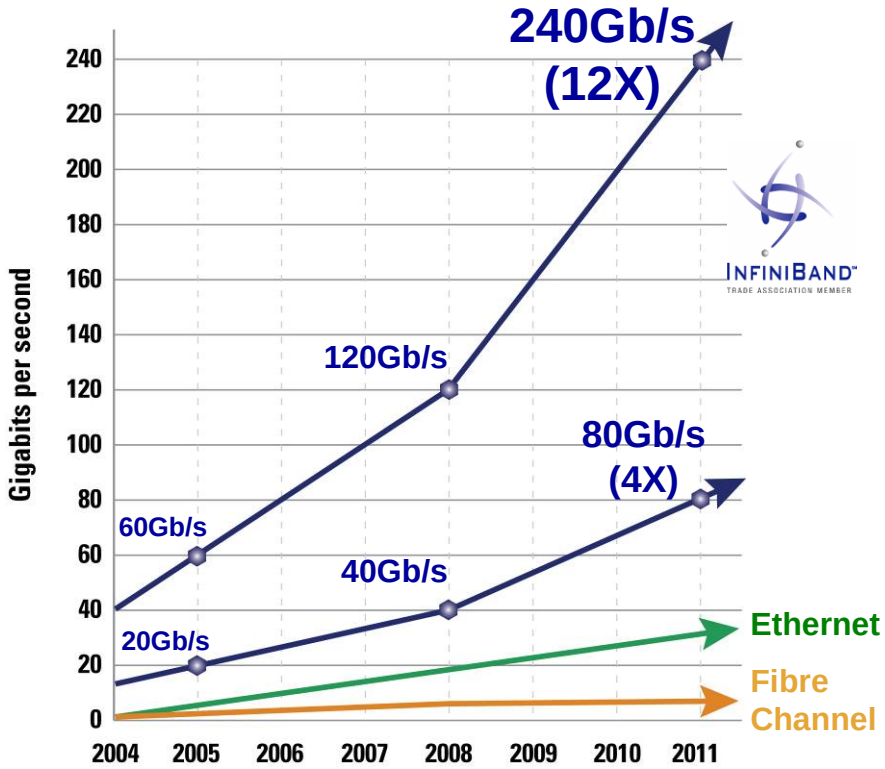
Test Cluster Configuration

- Dell™ PowerEdge™ SC 1435 24-node cluster
- Quad-Core AMD Opteron™ 2382 (“Shanghai”) CPUs
- Mellanox® InfiniBand ConnectX® 20Gb/s (DDR) HCAs
- Mellanox® InfiniBand DDR Switch
- Memory: 16GB memory, DDR2 800MHz per node
- OS: RHEL5U2, OFED 1.4 InfiniBand SW stack
- MPI: HP-MPI 2.3
- Application: FLUENT 6.3.37, FLUENT 12.0
- Benchmark Workload
  - New FLUENT Benchmark Suite

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**

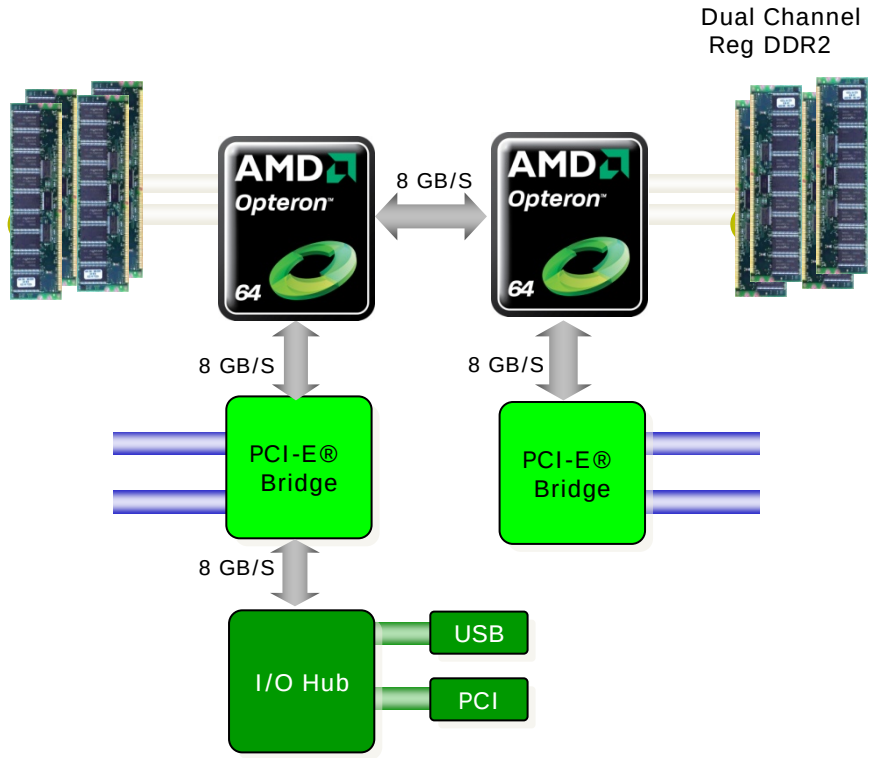
The InfiniBand Performance Gap is Increasing



InfiniBand Delivers the Lowest Latency

Quad-Core AMD Opteron™ Processor

- **Performance**
  - Quad-Core
    - Enhanced CPU IPC
    - 4x 512K L2 cache
    - 6MB L3 Cache
  - Direct Connect Architecture
    - HyperTransport™ Technology
    - Up to 24 GB/s peak per processor
  - Floating Point
    - 128-bit FPU per core
    - 4 FLOPS/clk peak per core
  - Integrated Memory Controller
    - Up to 12.8 GB/s
    - DDR2-800 MHz or DDR2-667 MHz
- **Scalability**
  - 48-bit Physical Addressing
- **Compatibility**
  - Same power/thermal envelopes as 2nd / 3rd generation AMD Opteron™ processor

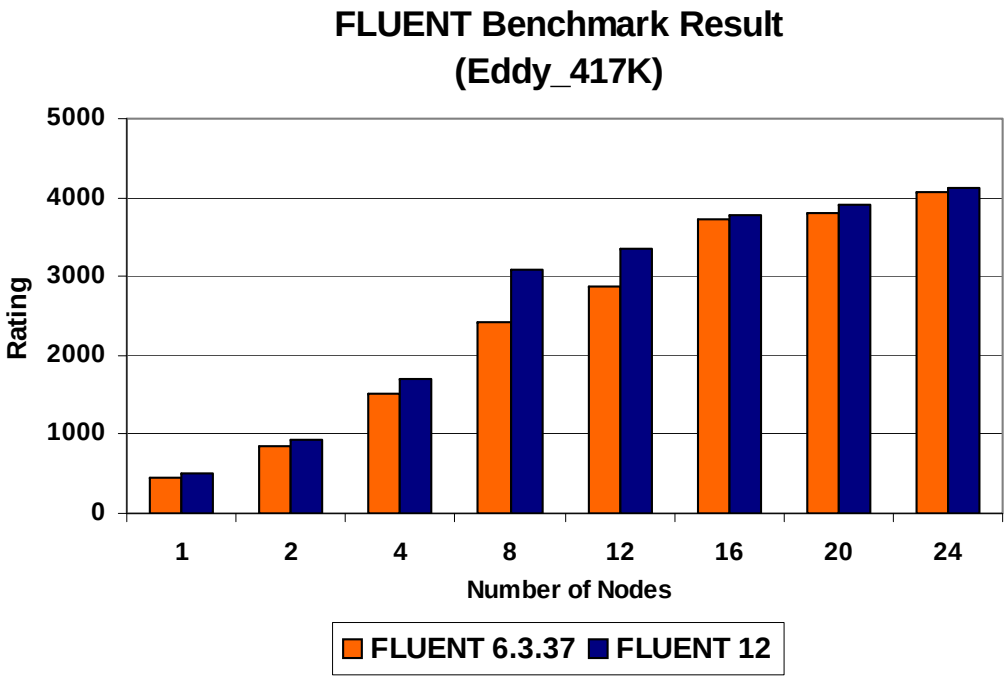


Dell PowerEdge Servers helping Simplify IT

- **System Structure and Sizing Guidelines**
  - 24-node cluster build with Dell PowerEdge™ SC 1435 Servers
  - 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

FLUENT Benchmark Results

- **Input Dataset**
  - EDDY_417K
    - Reacting Flow with Eddy Dissipation Model
- **FLUENT 12 provides better performance and scalability**
  - Utilizing InfiniBand DDR to delivers highest performance and scalability

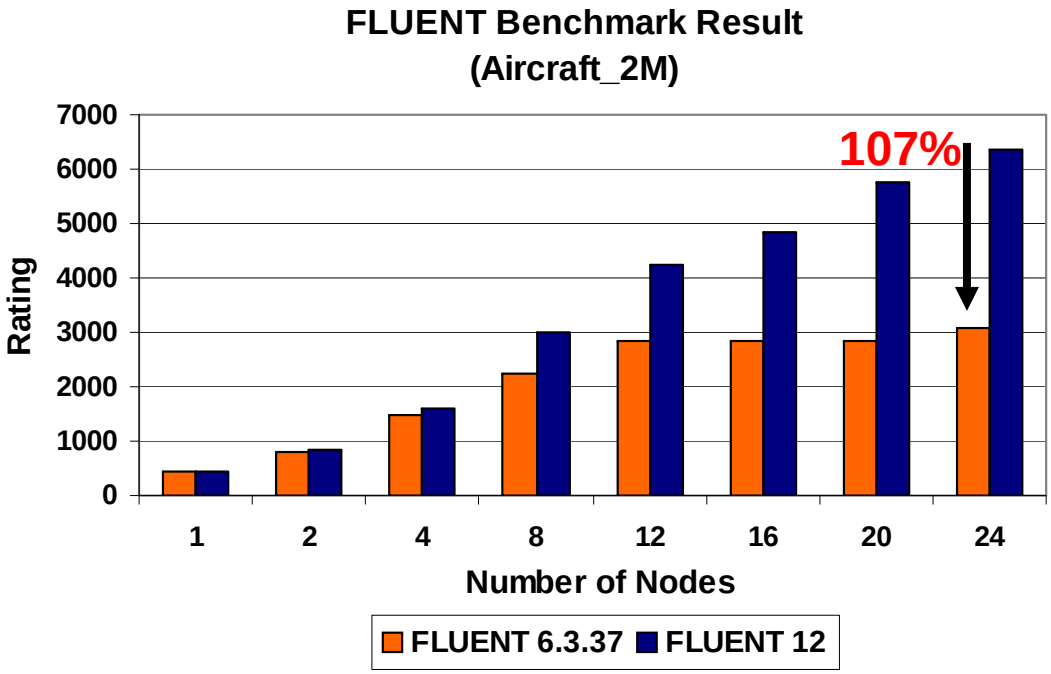


Higher is better

InfiniBand DDR

FLUENT Benchmark Results

- **Input Dataset**
  - Aircraft_2M
    - External Flow Over an Aircraft Wing
- **FLUENT 12 provides performance and scalability increase**
  - Up to 107% higher performance versus previous 6.3.37 version

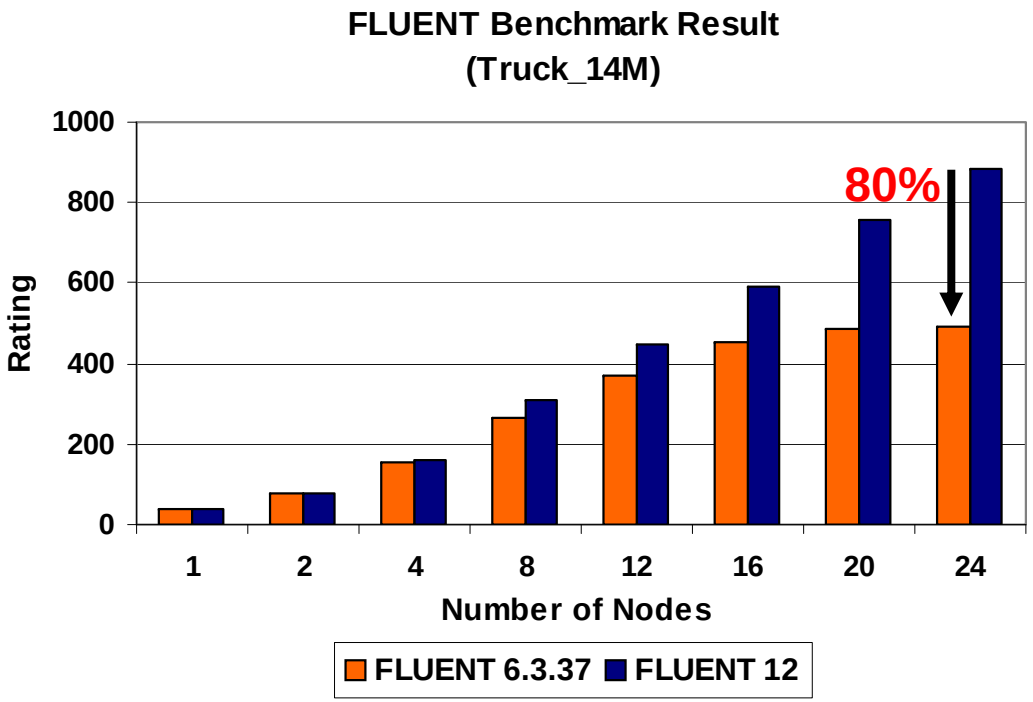


Higher is better

InfiniBand DDR

FLUENT Benchmark Results

- **Input Dataset**
  - Truck_14M
    - External Flow Over a Truck Body
- **FLUENT 12 delivers higher performance and scalability**
  - For any cluster size
  - Up to 80% higher performance versus previous 6.3.37 version

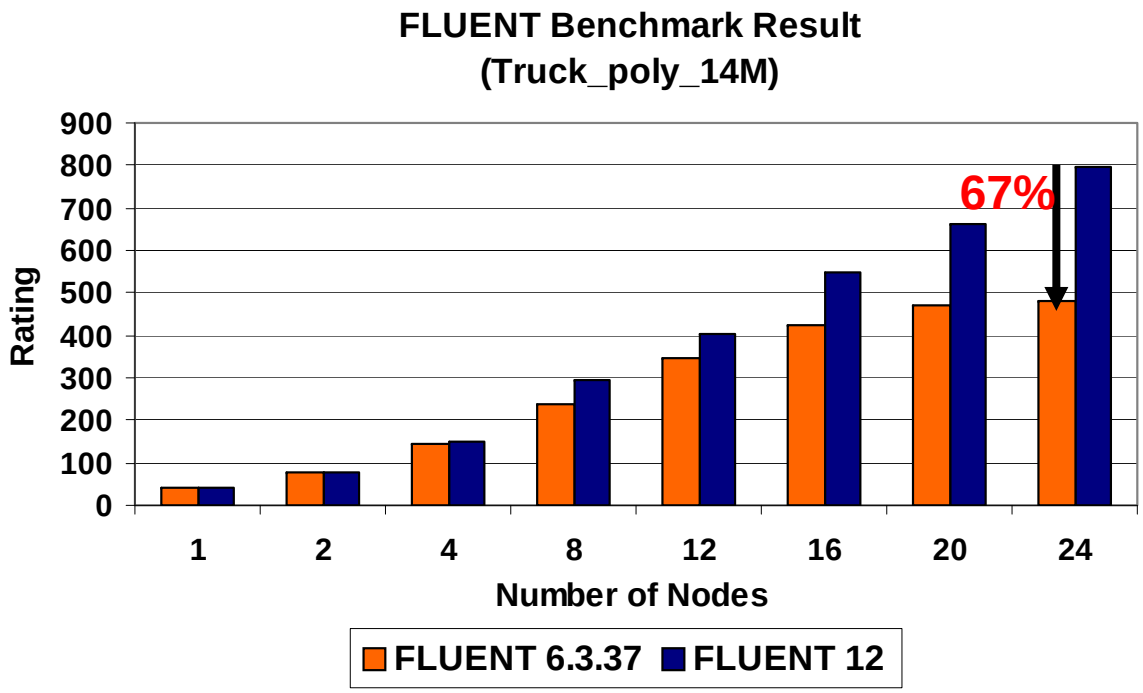


Higher is better

InfiniBand DDR

FLUENT Benchmark Results

- **Input Dataset**
  - Truck_Poly_14M
    - External Flow Over a Truck Body with a Polyhedral Mesh
- **FLUENT 12 delivers higher performance and scalability**
  - For any cluster size
  - Up to 67% higher performance versus previous 6.3.37 version

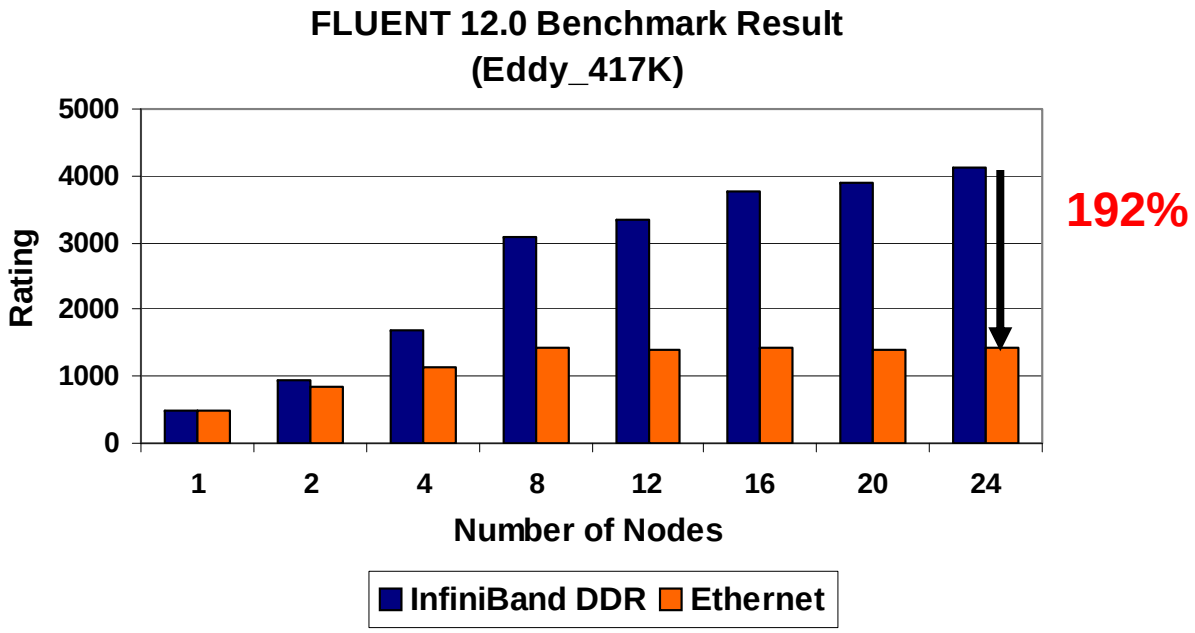


Higher is better

InfiniBand DDR

FLUENT 12 Benchmark Results - Interconnect

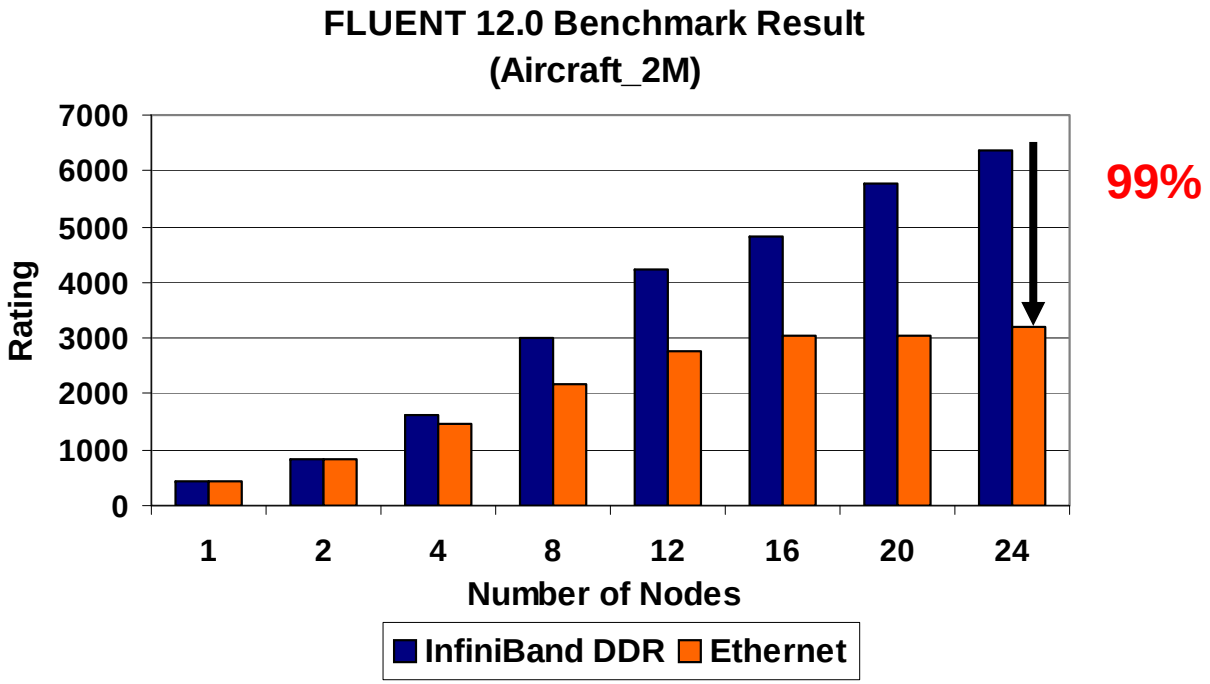
- **Input Dataset**
  - EDDY_417K (417 thousand elements)
    - Reacting Flow with Eddy Dissipation Model
- **InfiniBand DDR delivers higher performance and scalability**
  - For any cluster size
  - Up to 192% higher performance versus Ethernet (GigE)



Higher is better

FLUENT 12 Benchmark Results - Interconnect

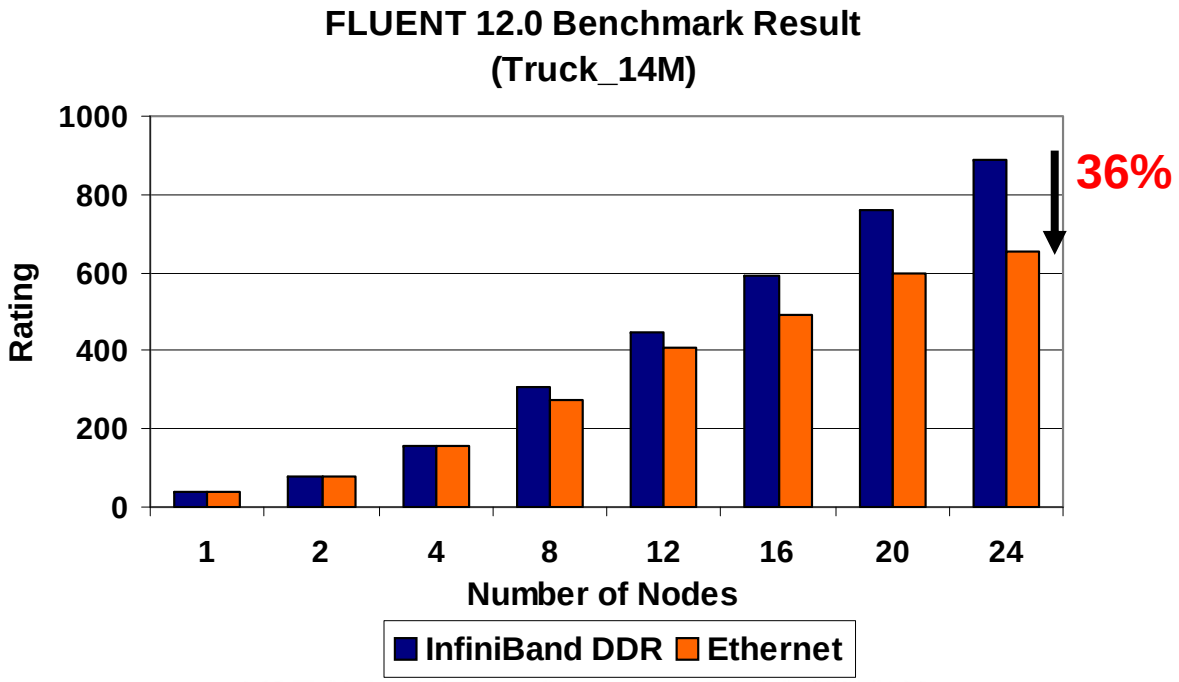
- **Input Dataset**
  - Aircraft_2M (2 million elements)
    - External Flow Over an Aircraft Wing
- **InfiniBand DDR delivers higher performance and scalability**
  - For any cluster size
  - Up to 99% higher performance versus Ethernet (GigE)



Higher is better

FLUENT 12 Benchmark Results - Interconnect

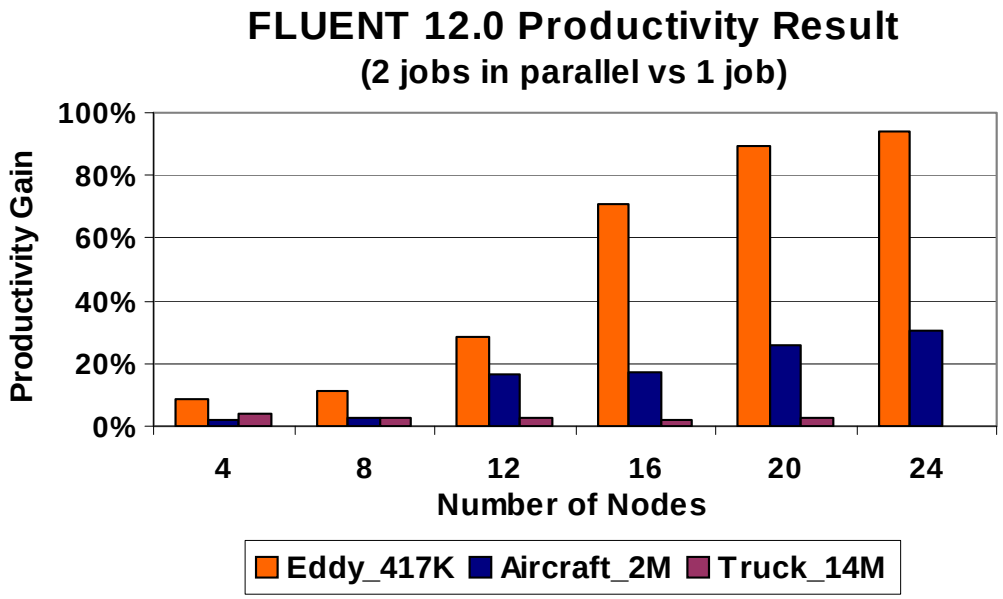
- **Input Dataset**
  - Truck_14M (14 millions elements)
    - External Flow Over a Truck Body
- **InfiniBand DDR delivers higher performance and scalability**
  - Up to 36% higher performance versus Ethernet (GigE)
- **For bigger cases (# of elements) CPU is the bottleneck for larger node count configuration**
  - More server nodes (or cores) are required for increased paternalism interconnect dependency



Higher is better

Enhancing FLUENT Productivity

- **Test cases**
  - Single job over the entire systems
  - 2 jobs, each runs on four cores per server
- **Running multiple jobs simultaneously improves FLUENT productivity**
  - Up to 90% more jobs per day for Eddy_417K
  - Up to 30% more jobs per day for Aircraft_2M
  - Up to 3% more jobs per day for Truck_14M
- **As bigger the # of elements, higher node count is required for increased productivity**
  - The CPU is the bottleneck for larger number of servers

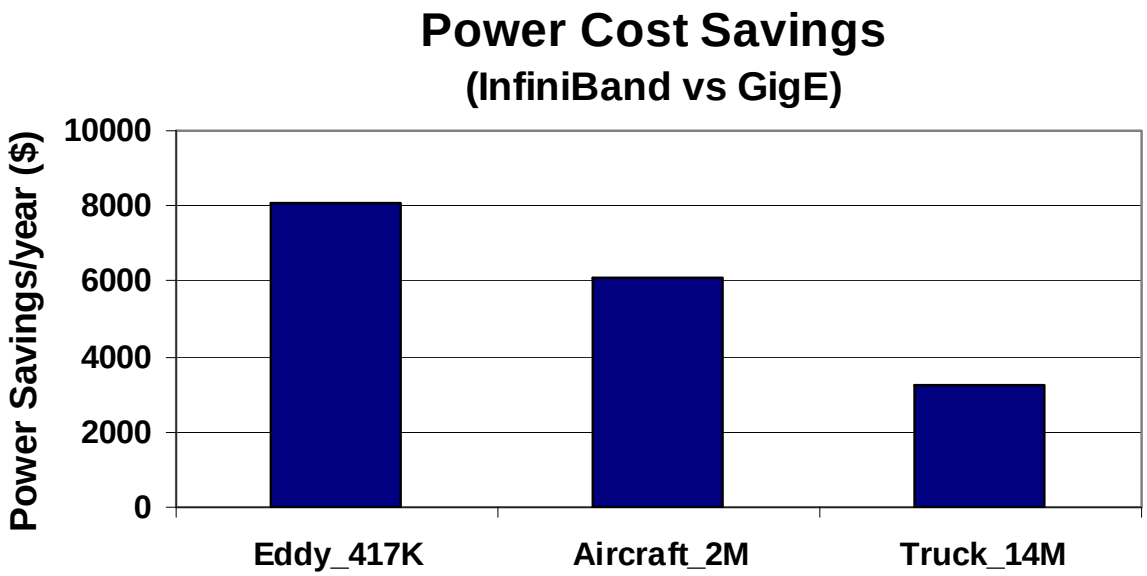


Higher is better

InfiniBand DDR

Power Cost Savings with Different Interconnect

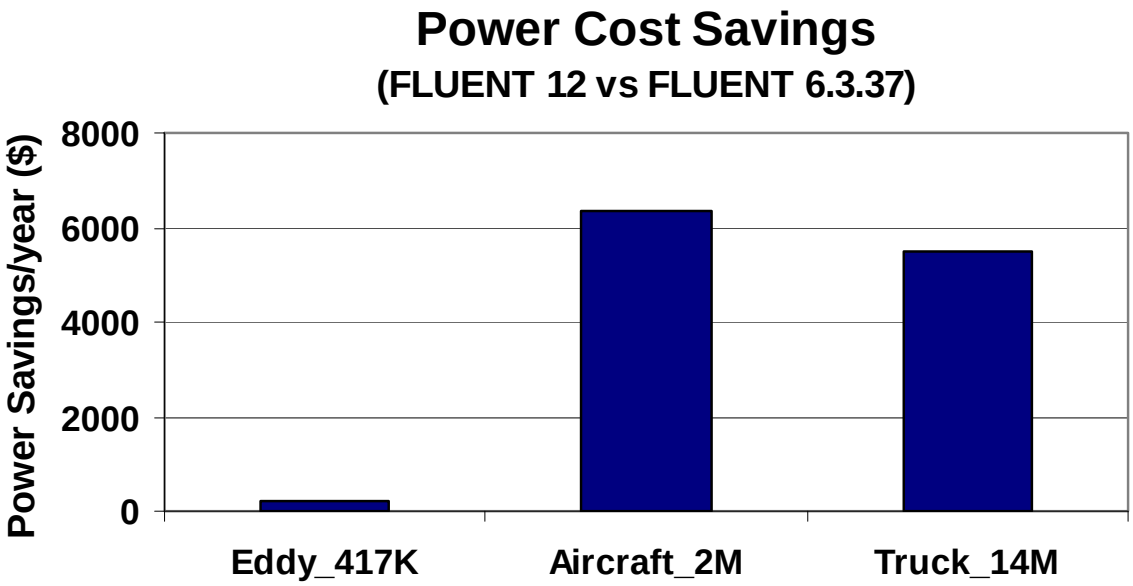
- **InfiniBand saves up to $8000 power to finish the same number of FLUENT jobs compared to GigE**
  - Yearly based for 24-node cluster
- **As cluster size increases, more power can be saved**



*$/KWh = KWh * $0.20*
For more information - http://enterprise.amd.com/Downloads/svrpwrusecompletefinal.pdf

Power Cost Savings with FLUENT upgrade

- **FLUENT 12 saves up to ~$6000 power to finish the same number of FLUENT jobs compared to FLUENT 6.3.37**
  - Yearly based for 24-node cluster
- **As cluster size increases, more power can be saved**



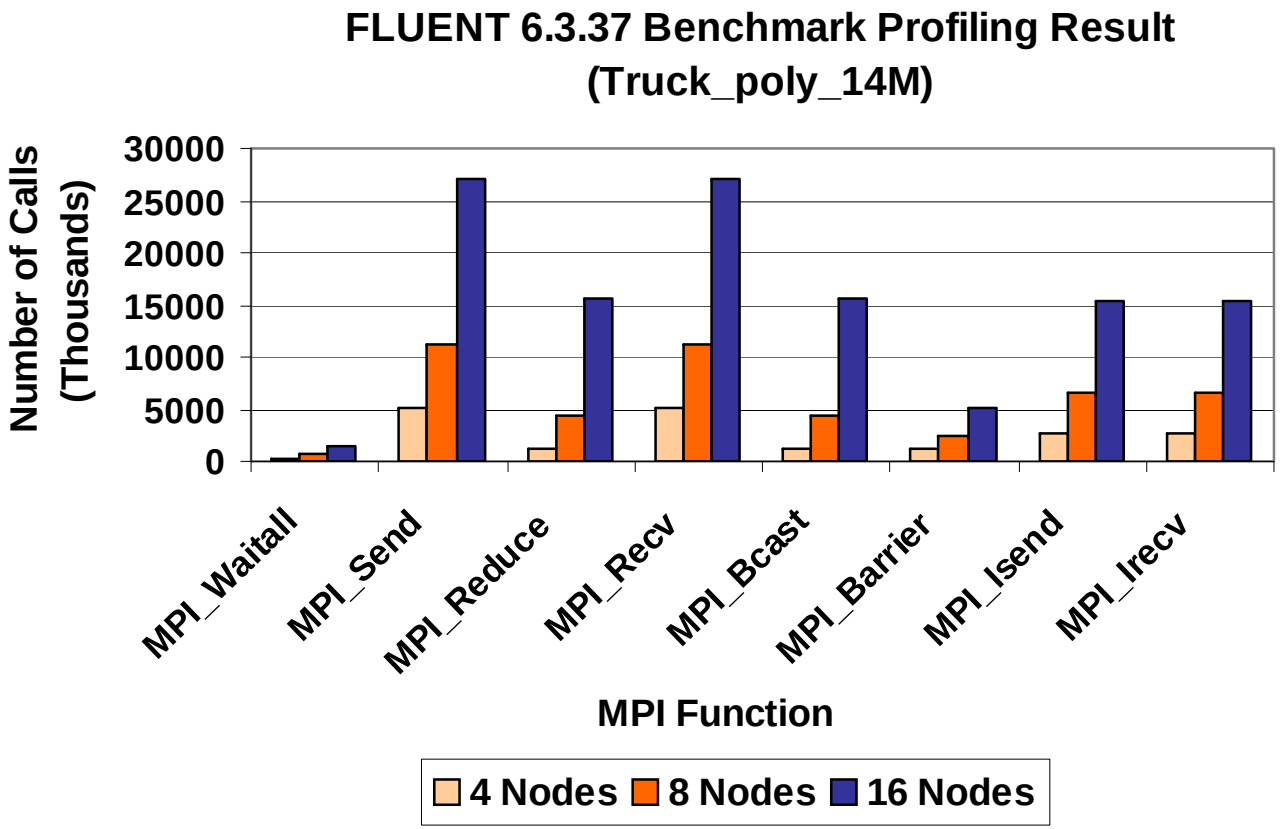
*$/KWh = KWh * $0.20*
For more information - http://enterprise.amd.com/Downloads/svrpwrusecompletefinal.pdf

FLUENT Productivity Results Summary

- **FLUENT 12 has tremendous performance improvement over version 6.3.37**
  - Optimizations made for higher performance and scalability
  - Optimizations included for AMD Opteron™ processor technology
    - AMD contributed to the development and the QA stages of Fluent 12
    - Performance results on AMD technology established baseline performance
- **InfiniBand enables higher performance and scalability than Ethernet**
  - Performance advantage extends as cluster size increases
- **Efficient job placement can increase FLUENT productivity significantly**
- **Interconnect comparison shows**
  - InfiniBand delivers superior performance in every cluster size
  - Low latency InfiniBand enables unparalleled scalability
- **InfiniBand enables up to $8000/year power savings compared to GigE**
- **FLUENT 12 reduces yearly power consumption by up to $6000 compared to FLUENT 6.3.37**

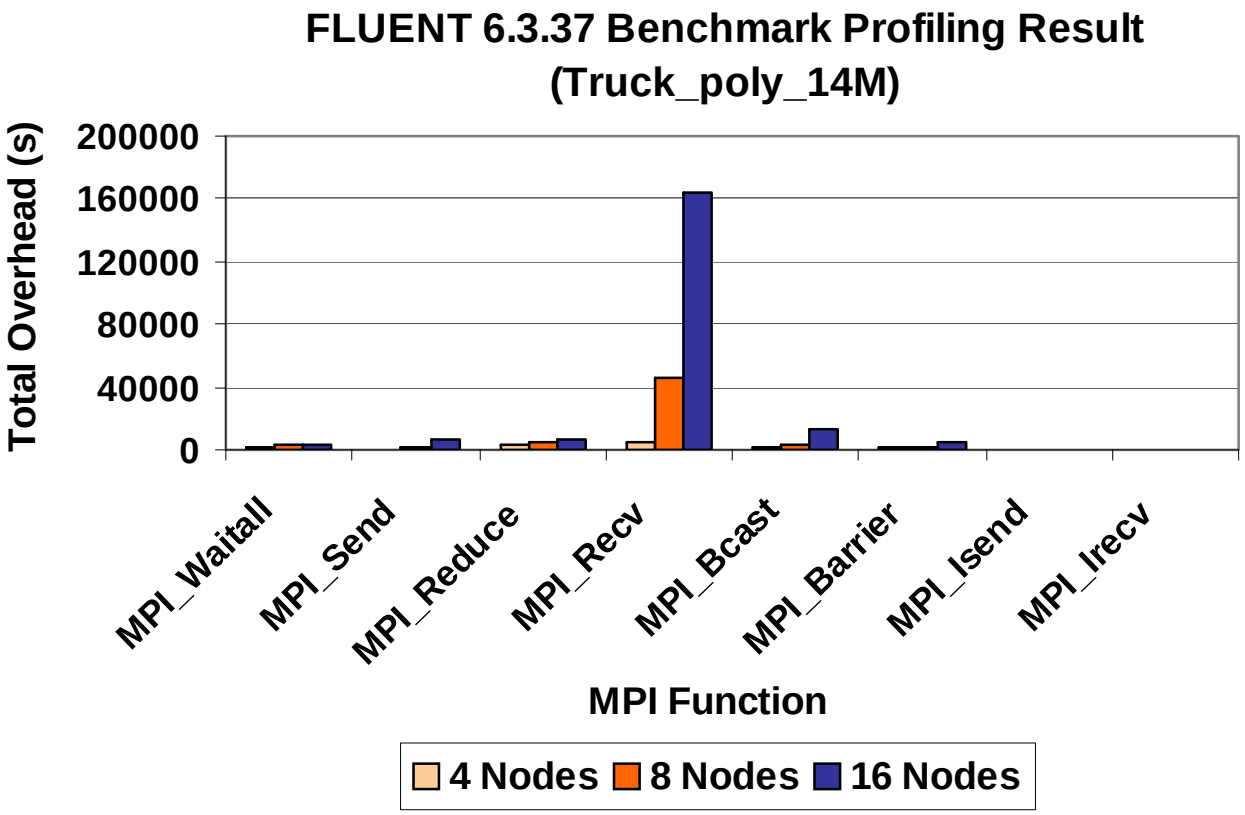
FLUENT 6.3.37 Profiling – MPI Functions

- **Mostly used MPI functions**
  - MPI_Send, MPI_Recv, MPI_Reduce, and MPI_Bcast



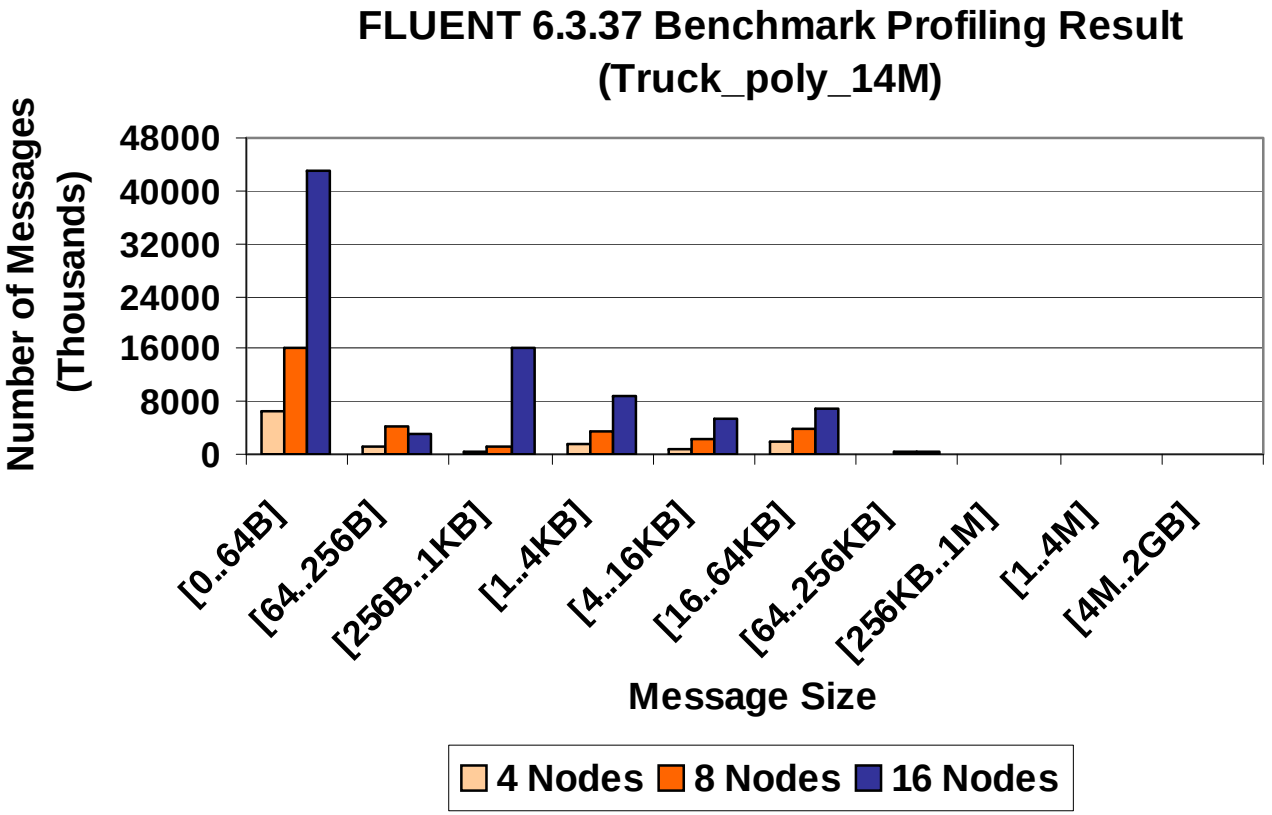
FLUENT 6.3.37 Profiling – Timing

- **MPI_Recv shows the highest communication overhead**



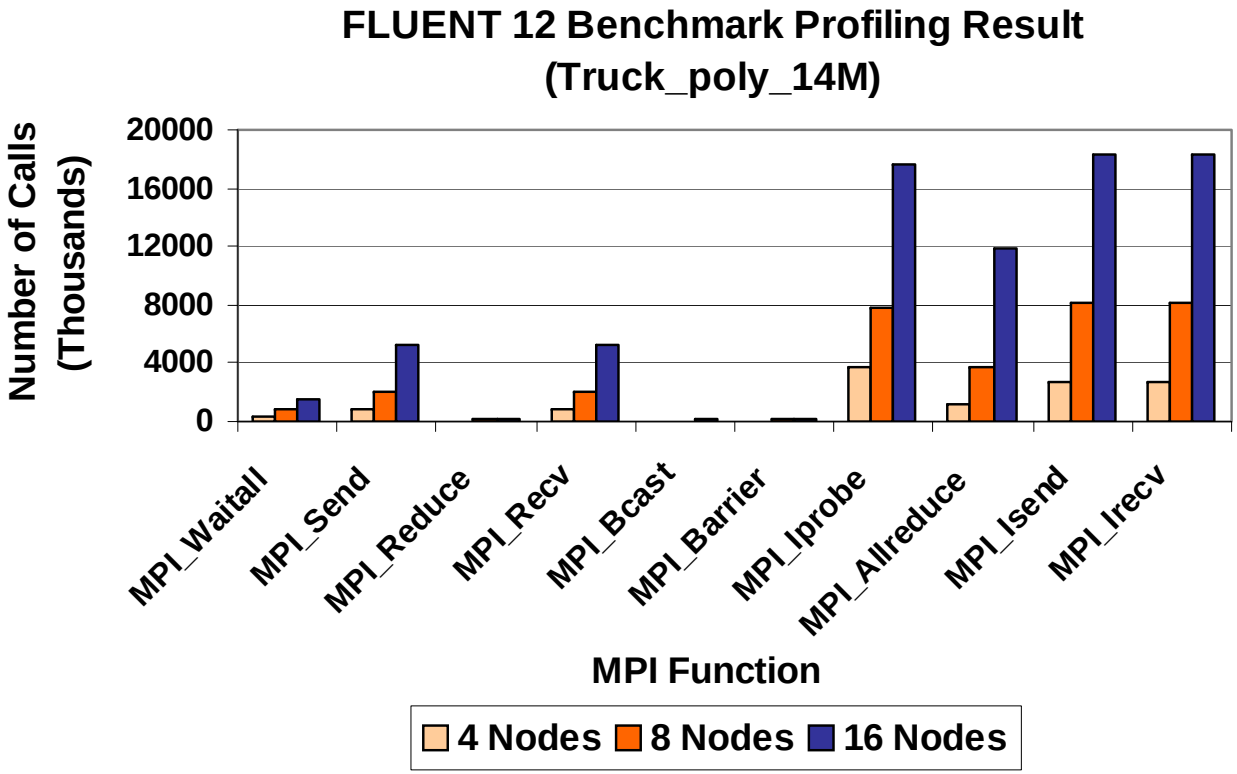
FLUENT 6.3.37 Profiling – Message Transferred

- Most data related MPI messages are within 256B-1KB in size
- Typical MPI synchronization messages are lower than 64B in size
- Number of messages increases with cluster size



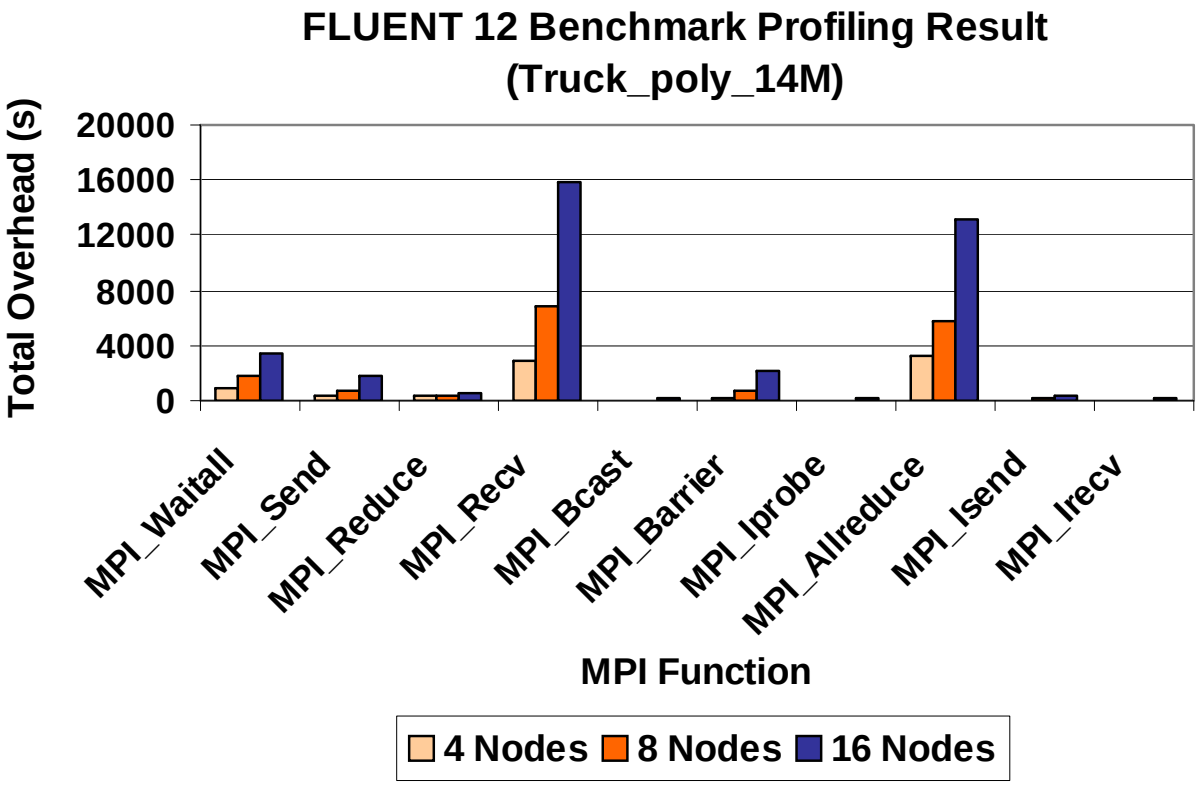
FLUENT 12 Profiling – MPI Functions

- **Mostly used MPI functions**
  - MPI_Iprobe, MPI_Allreduce, MPI_Isend, and MPI_Irecv



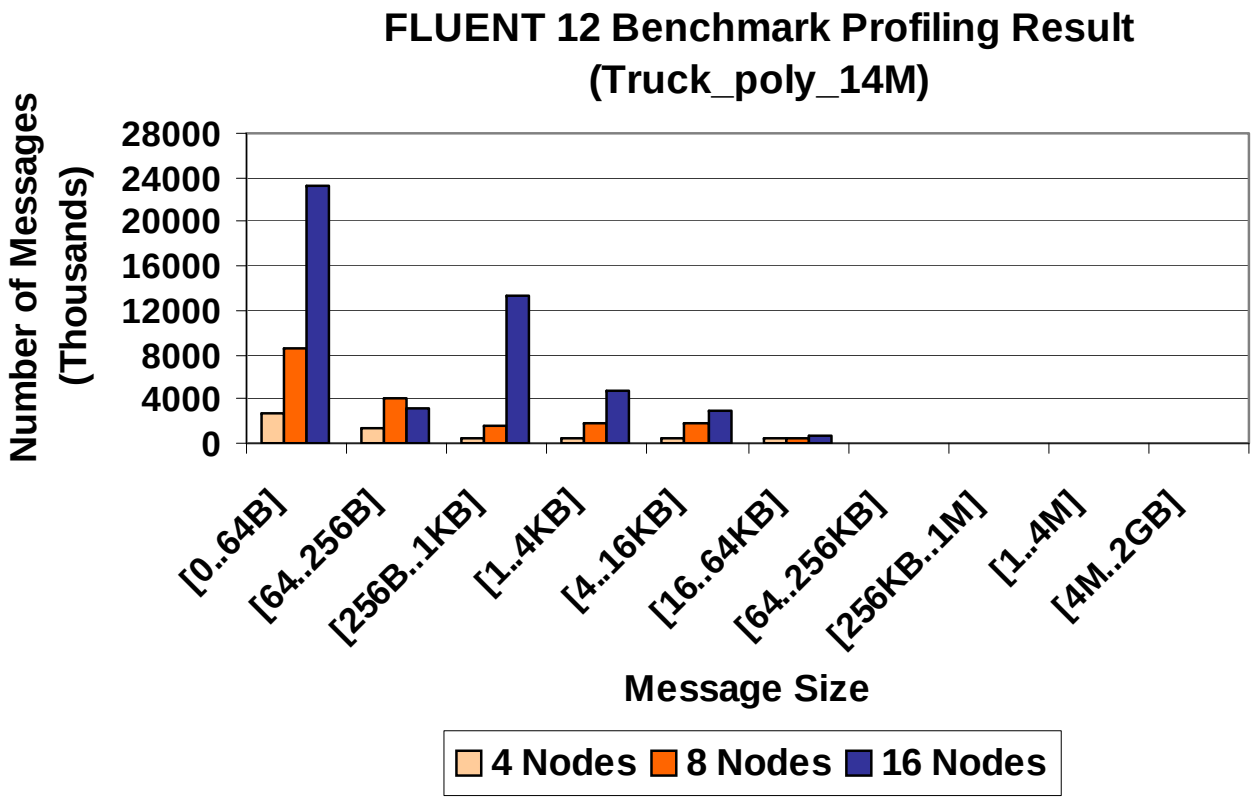
FLUENT 12 Profiling – Timing

- **MPI_Recv and MPI_Allreduce show highest communication overhead**



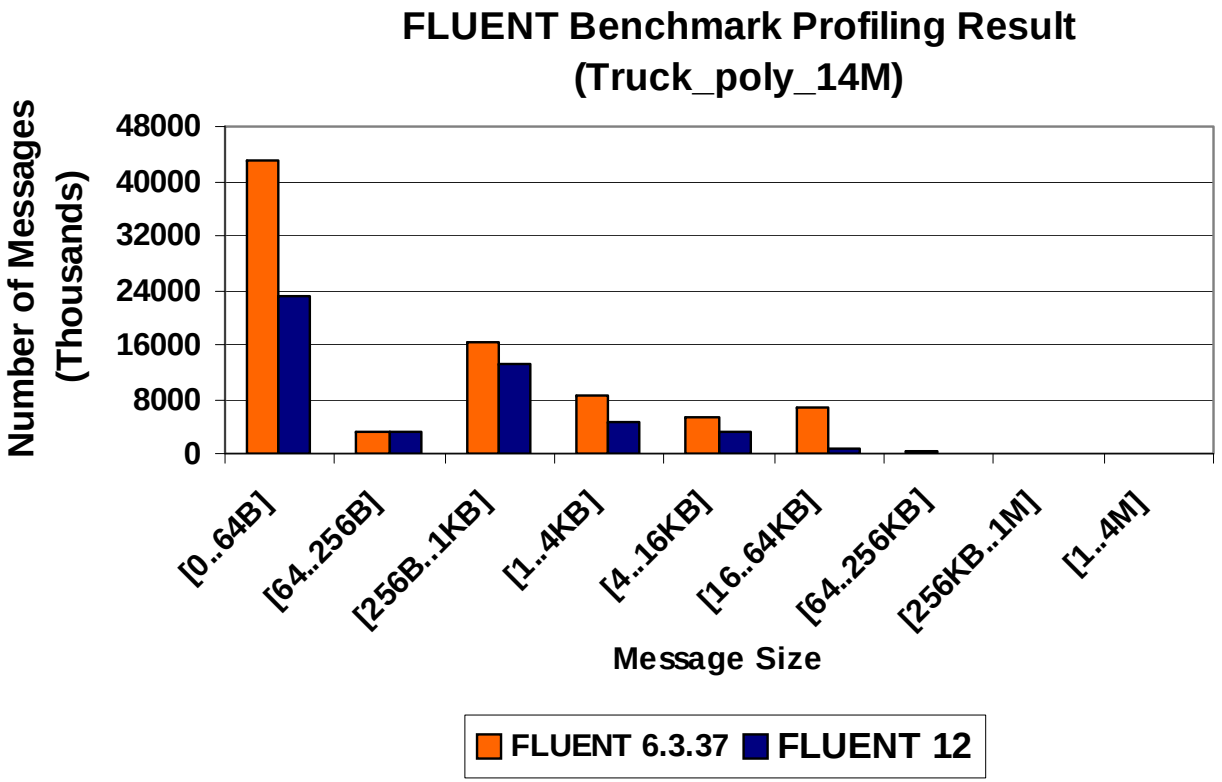
FLUENT 12 Profiling – Message Transferred

- Most data related MPI messages are within 256B-1KB in size
- Typical MPI synchronization messages are lower than 64B in size
- Number of messages increases with cluster size



FLUENT Profiling - Message Transferred

- **FLUENT 12 reduces total number of messages**
- **Further optimization can be made to take bigger advantage of high-speed and low latency interconnects**



FLUENT Profiling Summary

- **FLUENT 12 and FLUENT 6.3.37 were profiled to identify their communication patterns**
- **Frequent used message sizes**
  - 256-1KB messages for data related communications
  - <64B for synchronizations
  - Number of messages increases with cluster size
- **MPI Functions**
  - FLUENT 12 introduced MPI collective functions
    - MPI_Allreduce help improves the communication efficiency
- **Interconnects effect to FLUENT performance**
  - Both interconnect latency (MPI_Allreduce) and throughput (MPI_Recv) highly influence FLUENT performance
  - Further optimization can be made to take bigger advantage of high-speed networks

Thank You

HPC Advisory Council