HPC ADVISORY COUNCIL

ANSYS Fluent 14.5 Performance Benchmark and Profiling

March 2013

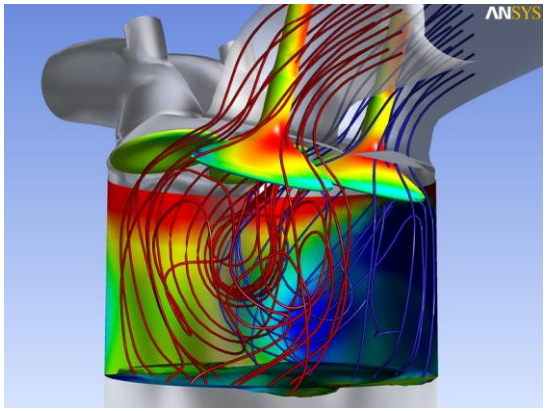
AMD ANSYS DELL Mellanox TECHNOLOGIES

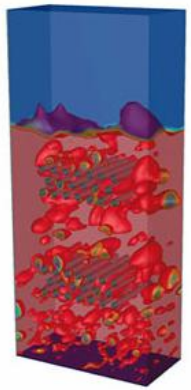
Note

- **The following research was performed under the HPC Advisory Council activities**
  - Participating vendors: AMD, Dell, Mellanox
  - Compute resource - HPC Advisory Council Cluster Center
- **For more info please refer to**
  - http:// www.amd.com
  - http:// www.dell.com/hpc
  - http://www.mellanox.com
  - http://www.ansys.com

CFD and 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 following was done to provide best practices**
  - ANSYS Fluent performance benchmarking
  - Interconnect performance comparisons
  - CPU performance
  - Understanding Fluent communication patterns
  - Ways to increase Fluent productivity
  - MPI libraries comparisons

- **The presented results will demonstrate**
  - The scalability of the compute environment
  - The capability of Fluent to achieve scalable productivity
  - Considerations for performance optimizations

Test Cluster Configuration

- **Dell™ PowerEdge™ R815 11-node (704-core) "Vesta" cluster**
  - AMD™ Opteron™ 6174 (code name "Magny-Cours") 12-cores @ 2.2 GHz CPUs
  - AMD™ Opteron™ 6276 (code name "Interlagos") 16-cores @ 2.3 GHz CPUs
  - AMD™ Opteron™ 6380 (code name "Abu Dhabi") 16-cores @ 2.5 GHz CPUs
- **4 CPU sockets per server node**
- **Mellanox ConnectX-3 VPI adapters for 40Gb/s QDR InfiniBand and 40Gb/s Ethernet**
- **Mellanox SwitchX™ 6036 36-Port InfiniBand switch**
- **Memory: 128GB memory per node DDR3 1333MHz**
- **OS: RHEL 6.2 MLNX-OFED 1.5.3 InfiniBand SW stack**
- **MPI: Intel MPI 4.0 Update 3, Open MPI 1.3.3, Platform MPI 8.2**
- **Application: ANSYS Fluent version 14.5**
- **Benchmark workload:**
  - sedan_4m (External Aerodynamics Flow Over a Passenger Sedan)
  - truck_poly_14m (External Flow Over a Truck Body with a Polyhedral Mesh. 14 million cells)
  - truck_111m (External flow case over a truck body. 111 mill cells)

Dell™ PowerEdge™ R815 11-node cluster

- **HPC Advisory Council Test-bed System**

- **New 11-node 704 core cluster - featuring Dell PowerEdge™ R815 servers**
  - Replacement system for Dell PowerEdge SC1435 (192 cores) cluster system following 2 years of rigorous benchmarking and product EOL
    - System to be redirected to explore HPC in the Cloud applications

- **Workload profiling and benchmarking**
  - Characterization for HPC and compute intense environments
  - Optimization for scale, sizing and configuration and workload performance
  - Test-bed Benchmarks
    - RFPs
    - Customers/Prospects, etc
  - ISV & Industry standard application characterization
  - Best practices & usage analysis

About Dell PowerEdge™ Platform Advantages

Best of breed technologies and partners

Combination of AMD™ Opteron™ 6300 series platform and Mellanox ConnectX®-3 InfiniBand on Dell HPC

Solutions provide the ultimate platform for speed and scale

- Dell PowerEdge R815 system delivers 4 socket performance in dense 2U form factor
- Up to 64 core/32DIMMs per server – 1344 core in 42U enclosure

Integrated stacks designed to deliver the best price/performance/watt

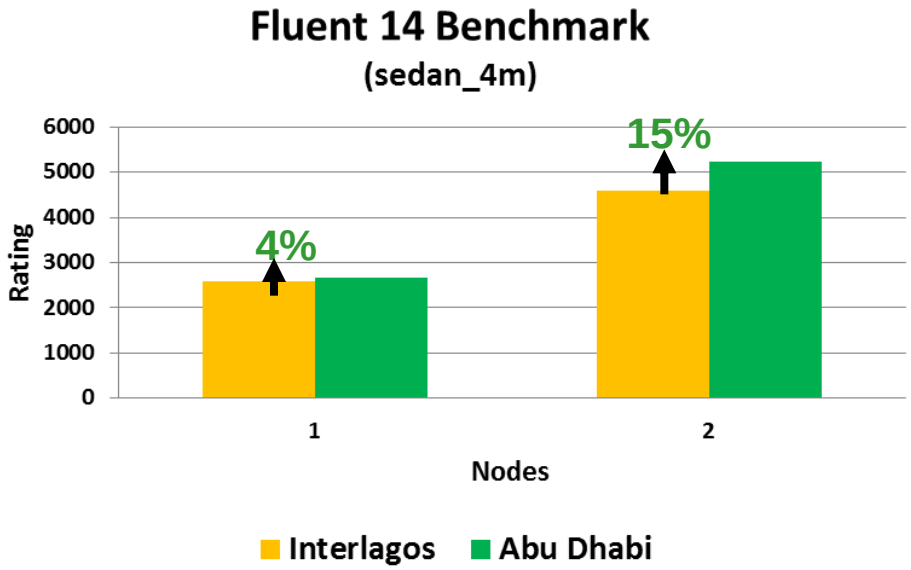
- 2x more memory and processing power in half of the space
- Energy optimized low flow fans, improved power supplies and dual SD modules

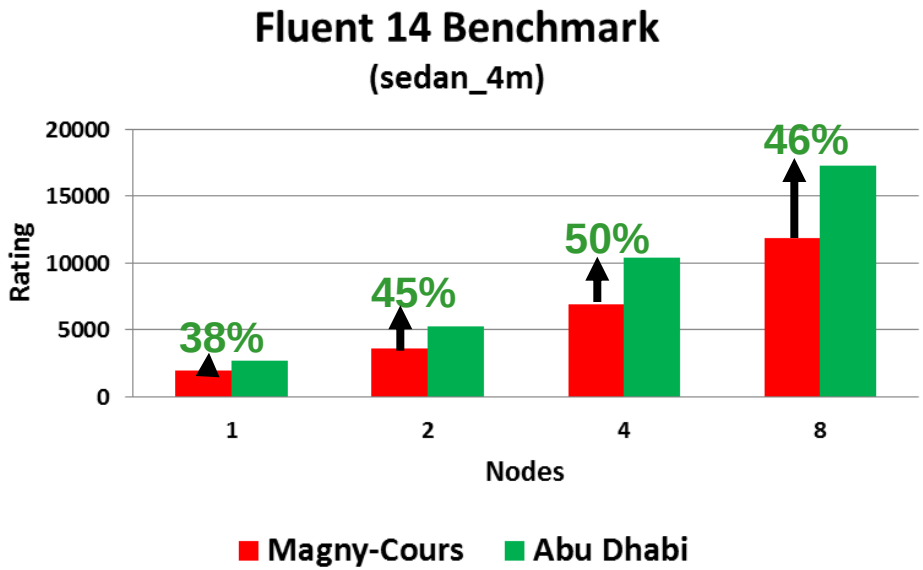
Optimized for long-term capital and operating investment protection

- Platform longevity across 3 CPU generations (AMD™ Opteron™ 6100, 6200 & 6300 series)
- System expansion, component upgrades and feature releases

Fluent Performance – Processor Generations

- **AMD Opteron 6300 series outperforms prior CPU generations**
  - Delivers up to 15% gain over the AMD Opteron 6200 “Interlagos” series
  - Delivers up to 50% gain over the AMD Opteron 6100 “Magny-Cours” series
  - Additional performance is expected if 1600MHz (instead of 1333MHz) DIMMs are used
- **The following configuration differences are noted:**
  - Mangy-Cours: Fluent 13.0, AMD Opteron 6174 @ 2.2GHz, ConnectX-2 HCA
  - Interlagos: Fluent 14.5, AMD Opteron 6276 @ 2.3GHz, ConnectX-3 HCA
  - Abu Dhabi: Fluent 14.5, AMD Opteron 6380 @ 2.5GHz, ConnectX-3 HCA





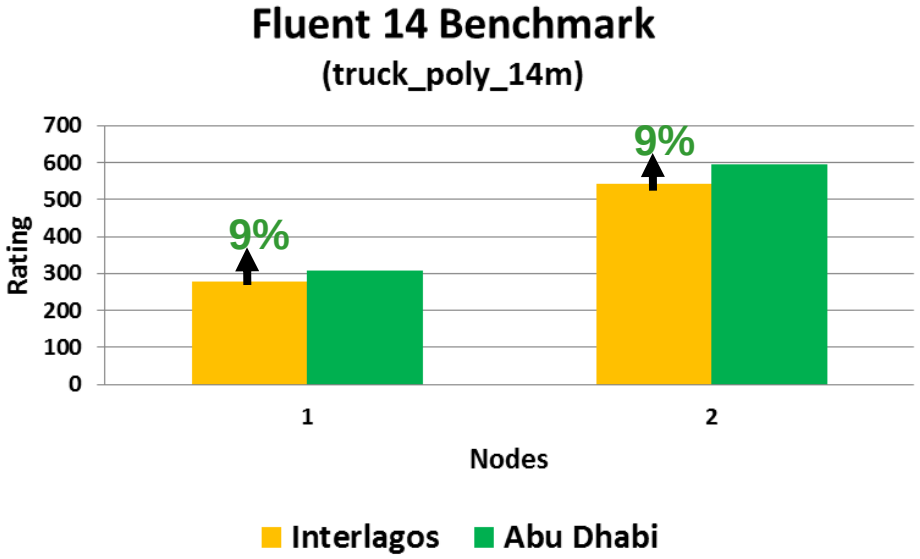
Higher is better

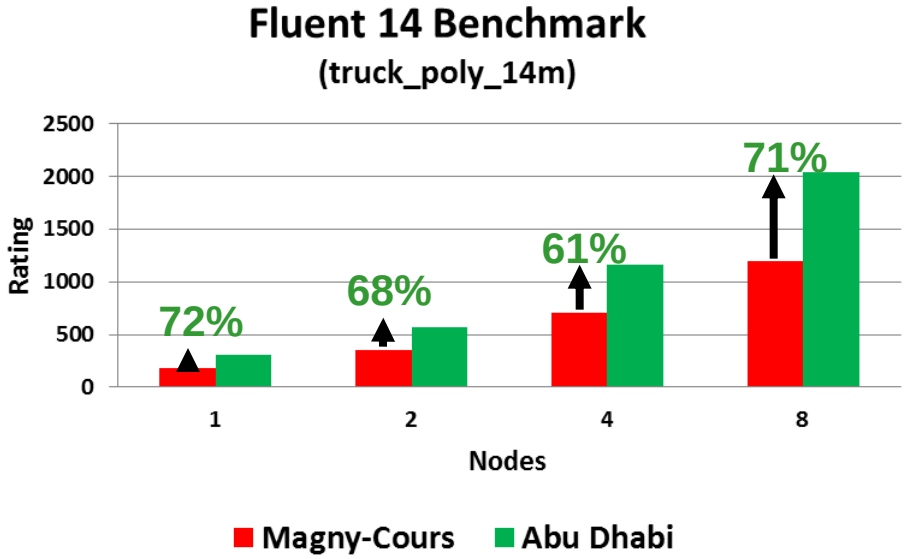
**Rating: Jobs/Day*

64 Cores/Node

Fluent Performance – Processor Generations

- **AMD Opteron 6300 series demonstrate modest gain over past generations**
  - Delivers up to 9% gain over the AMD Opteron 6200 “Interlagos” series
  - Delivers up to 72% gain over the AMD Opteron 6100 “Magny-Cours” series
  - Additional performance is expected if 1600MHz (instead of 1333MHz) DIMMs are used
- **The following configuration differences are noted:**
  - Mangy-Cours: Fluent 13.0, AMD Opteron 6174 @ 2.2GHz, ConnectX-2 HCA
  - Interlagos: Fluent 14.5, AMD Opteron 6276 @ 2.3GHz, ConnectX-3 HCA
  - Abu Dhabi: Fluent 14.5, AMD Opteron 6380 @ 2.5GHz, ConnectX-3 HCA



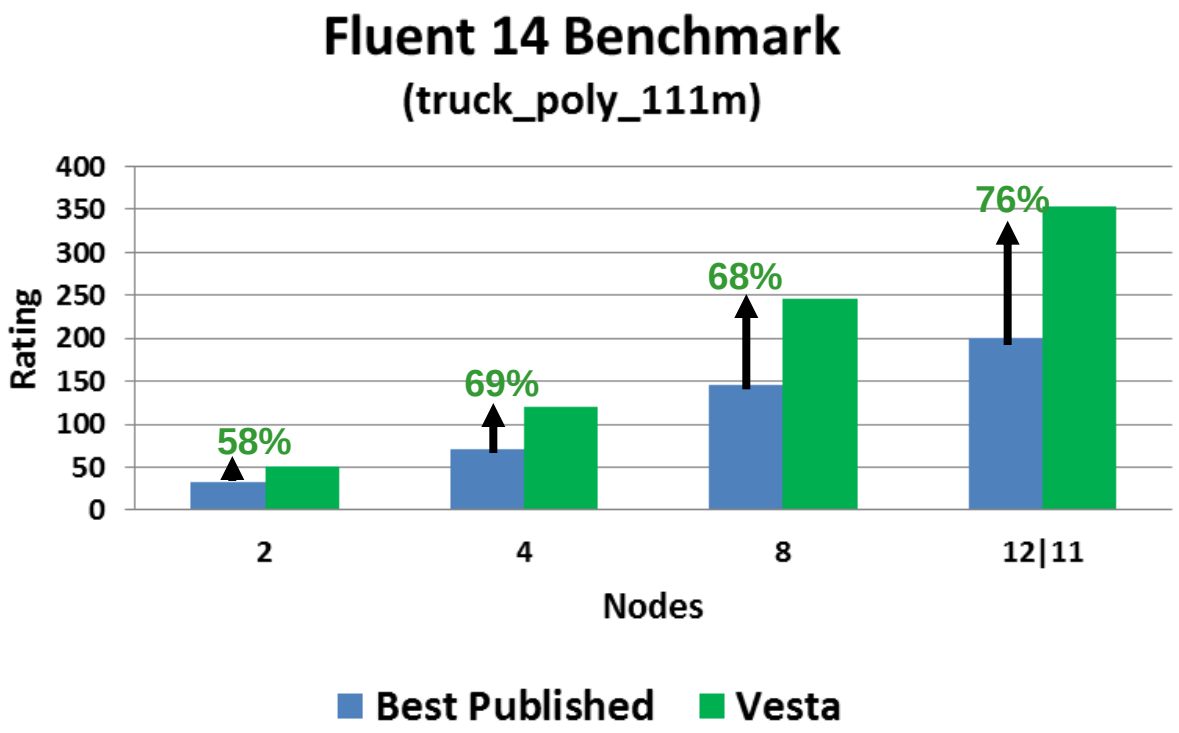


Higher is better

64 Cores/Node

Fluent Performance – Best Published

- **The 4P AMD "Abu Dhabi" cluster delivers the best performance**
  - 4-socket PowerEdge R815 Vesta cluster delivers compute density and performance
  - Up to 76% higher performance compared to best published results
- **Published results for Fluent 14:**
  - http://www.ansys.com/Support/Platform+Support/Benchmarks+Overview/ANSYS+Fluent+Benchmarks/

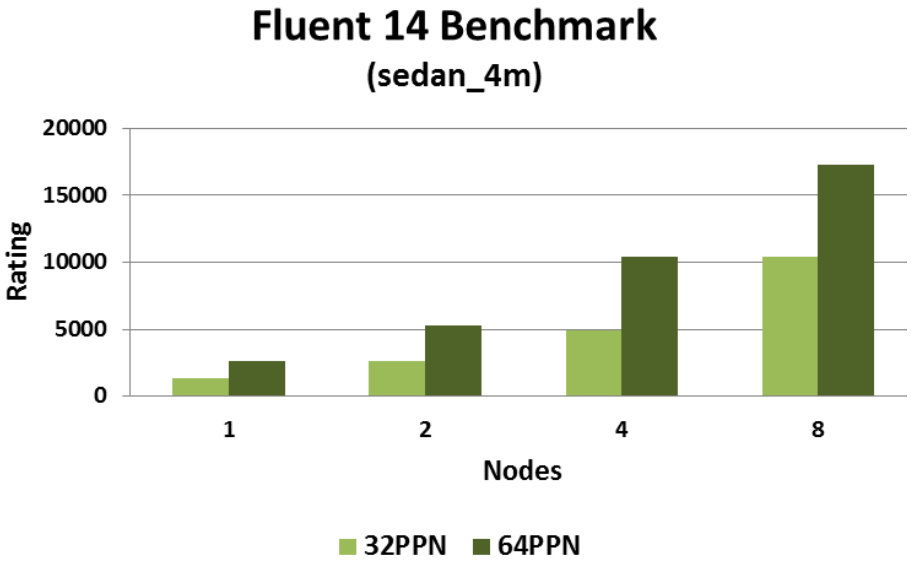


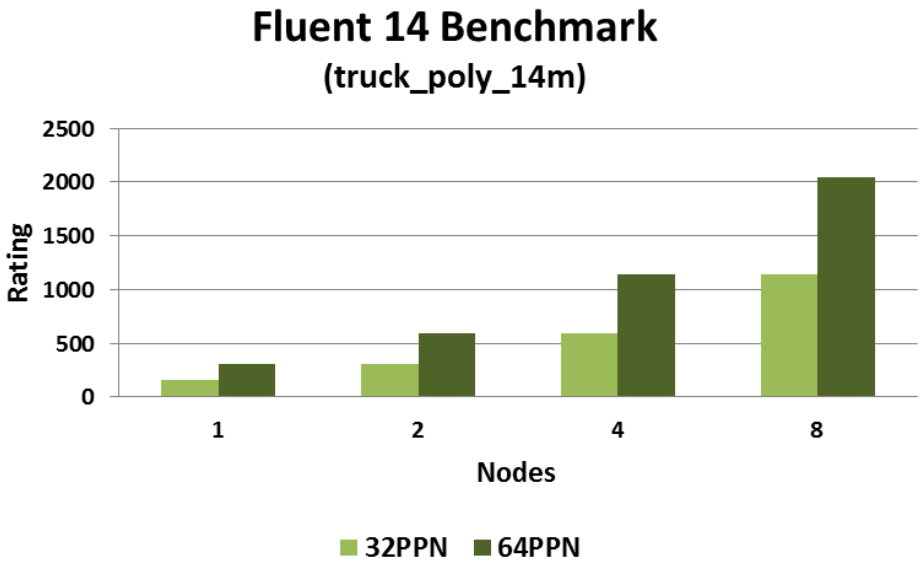
Higher is better

64 Cores/Node

Fluent Performance – Processes Per Node

- **Running with 64PPN yields doubles the system utilization than with 32PPN**
  - 4P servers (64 cores/node) delivers almost twice the performance over a 2P server

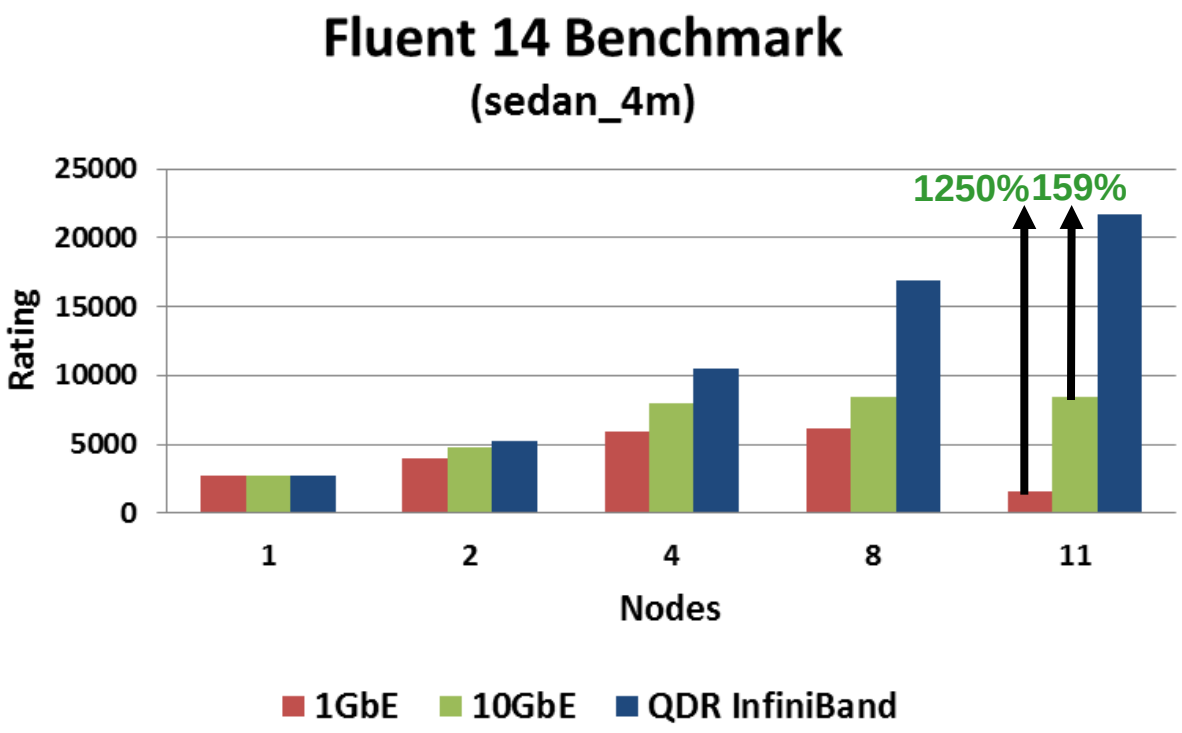




Higher is better

Fluent Performance – Interconnects

- **InfiniBand shows continuous gain as the cluster scales for sedan_4m**
  - Up to 159% higher productivity compared to 10GbE at 11 nodes (704 cores)
  - Over 12 times higher productivity compared to 1GbE at 11 nodes (704 cores)
  - Ethernet does not scale; performance declines form 4 nodes and beyond

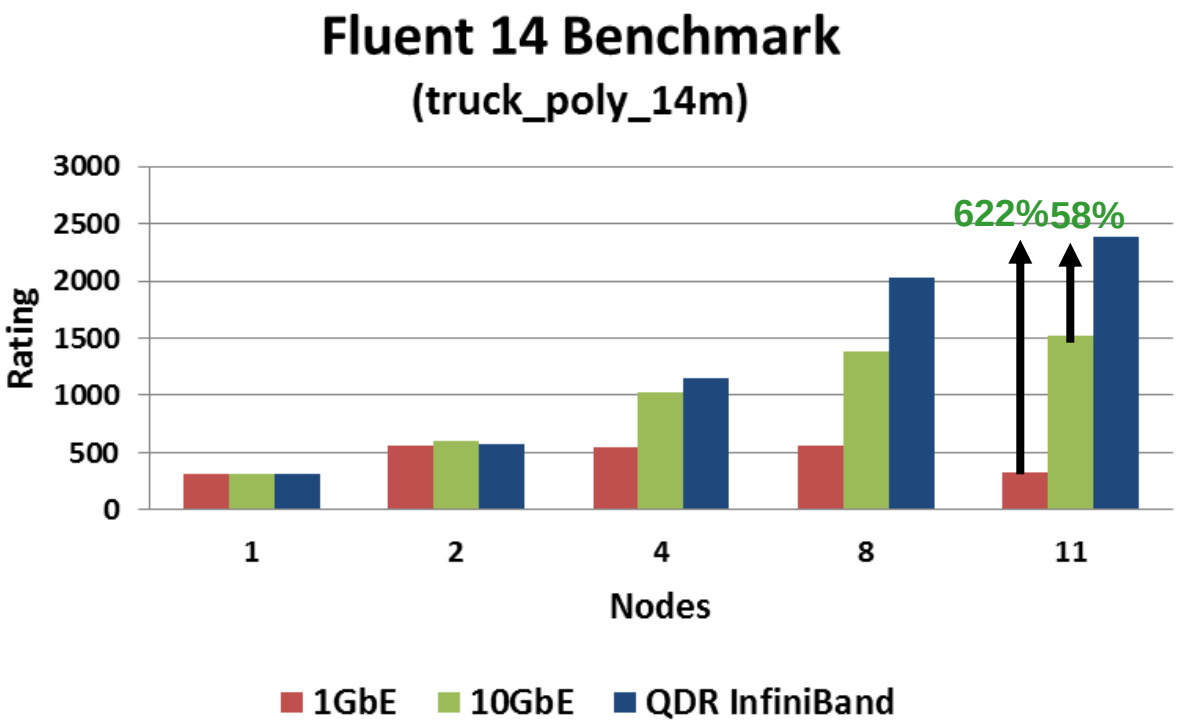


Higher is better

64 Cores/Node

Fluent Performance – Interconnects

- **InfiniBand shows continuous gain as the cluster scales for sedan_4m**
  - Up to 58% higher productivity compared to 10GbE at 11 nodes (704 cores)
  - Over 6 times higher productivity compared to 1GbE at 11 nodes (704 cores)
  - Ethernet solutions shows performance decline from 8 nodes and beyond

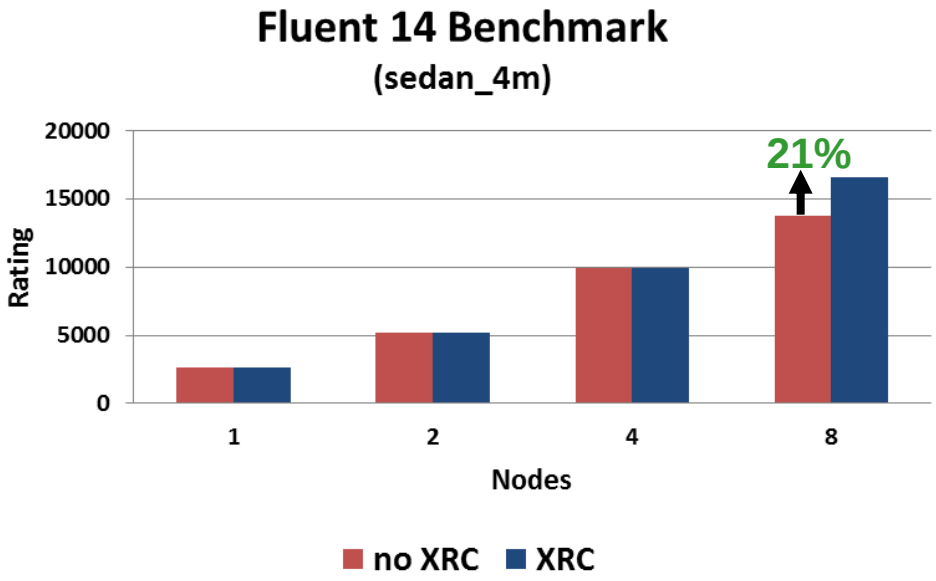


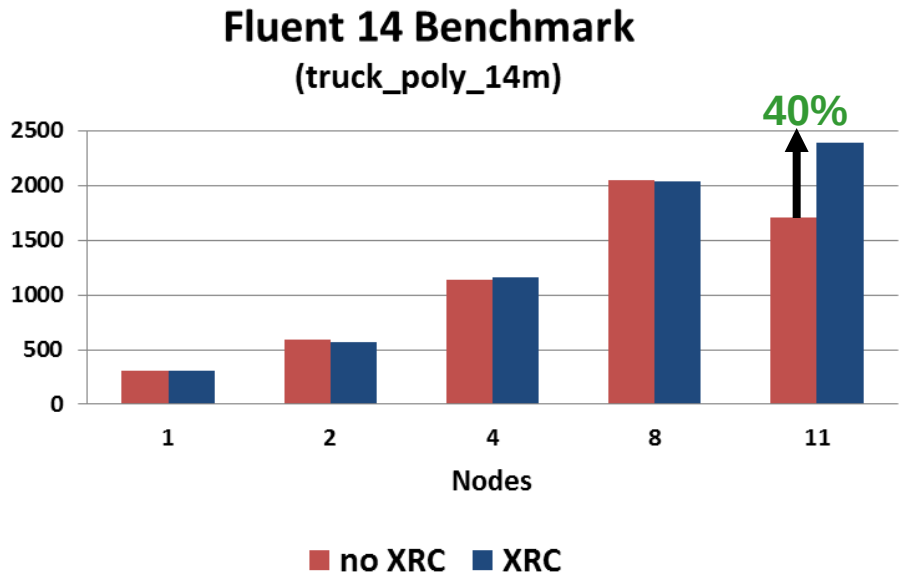
Higher is better

64 Cores/Node

Fluent Performance – XRC

- **Running XRC with InfiniBand scalability at higher CPU core counts**
  - Up to 40% better performance for 11 nodes (or 704 cores) for truck_poly_11m
  - Up to 21% better performance for 8 nodes (or 512 cores) for sedan_4m
- **To Enable XRC in ANSYS Fluent by modifying the mpirun.fl file:**
  - Include "-mca btl_openib_receive_queues X,9216,256,128,32:X,65536,256,128,32" for openmpi
  - Include the "-xrc" flag in the my_protocol_flags for pcmpi
  - Need to build Open MPI separately to run as the vendor-supplied not built with XRC support



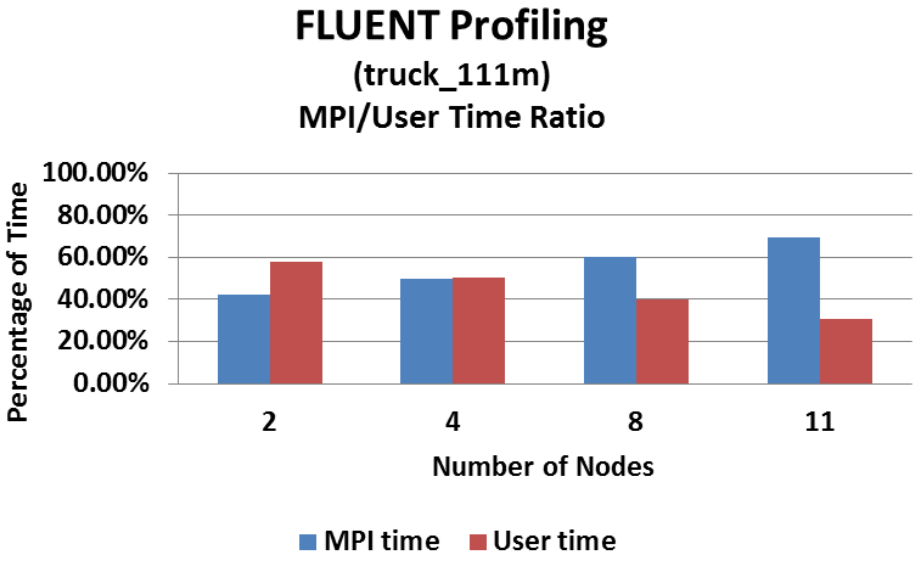


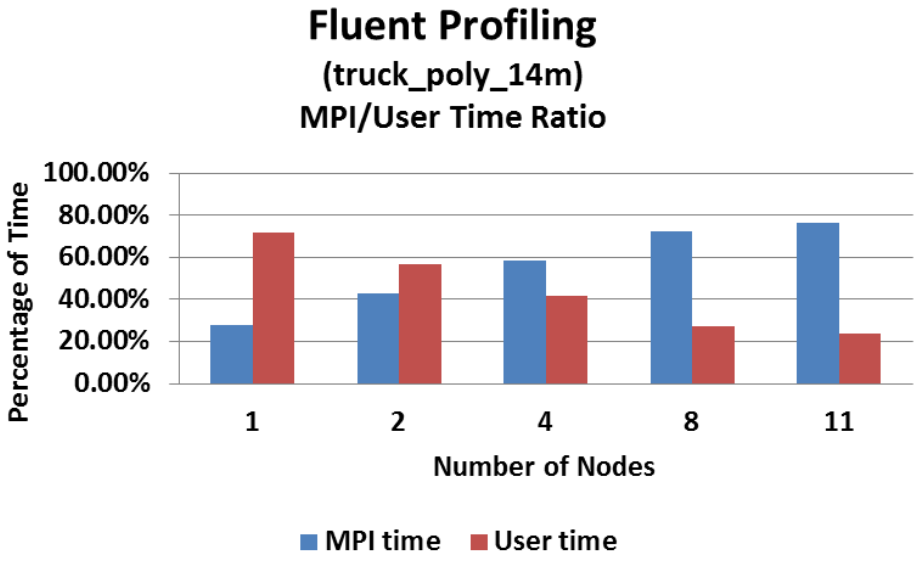
Higher is better

64 Cores/Node

Fluent Profiling – MPI/User Time Ratio

- **Gradual increase in communications time as the cluster scales**
  - truck_poly_14m spends more time on MPI communications than truck_111m
  - Network infrastructure like InfiniBand allows Fluent to run at scale



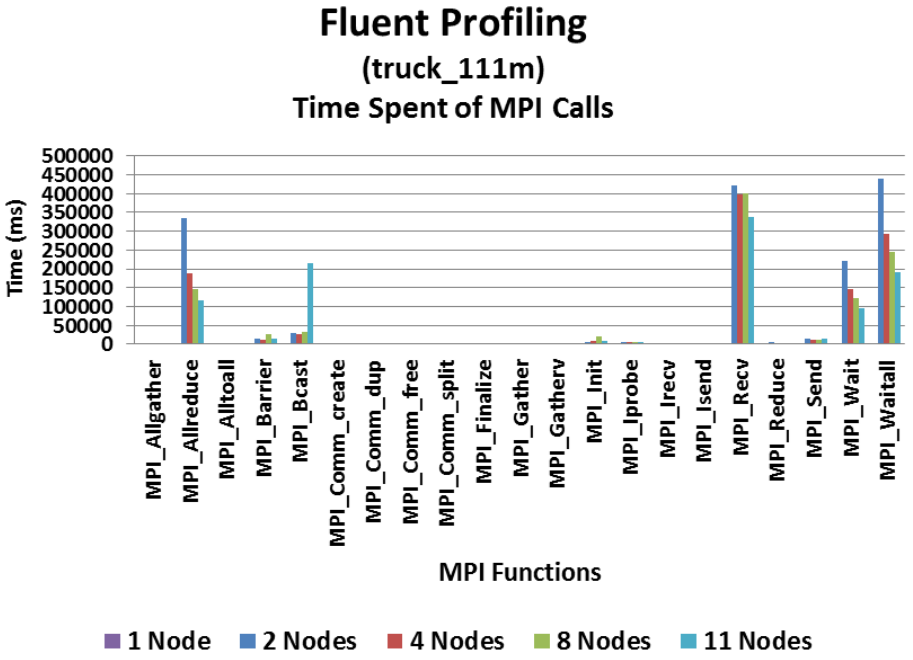


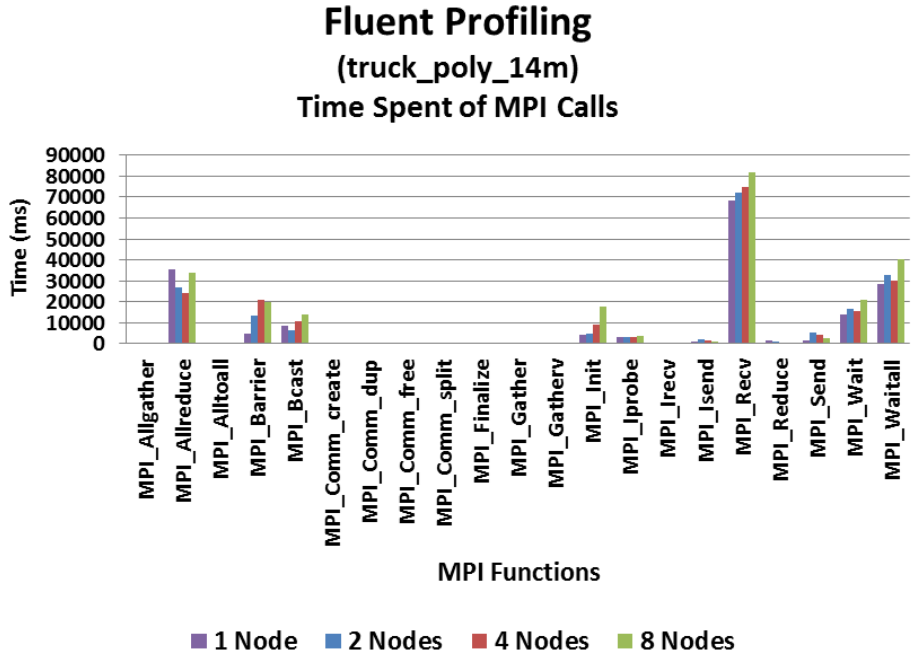
QDR InfiniBand

64 Cores/Node

Fluent Profiling – Time Spent of MPI Calls

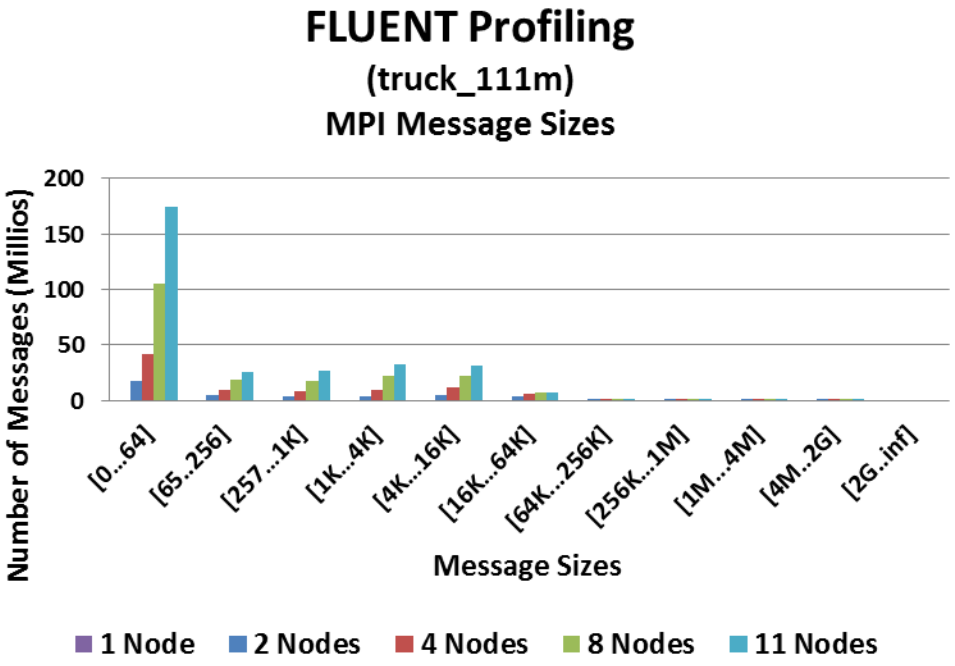
- **MPI_Recv is the largest time consumer for truck_poly_14m & truck_111m**
  - Occupies 34% of all MPI time for 8 node in both truck_poly_14m and truck_111m
- **More time spent on data MPI communication than MPI synchronization**
  - Both truck_poly_14m and truck_111m show large time in MPI_Recv

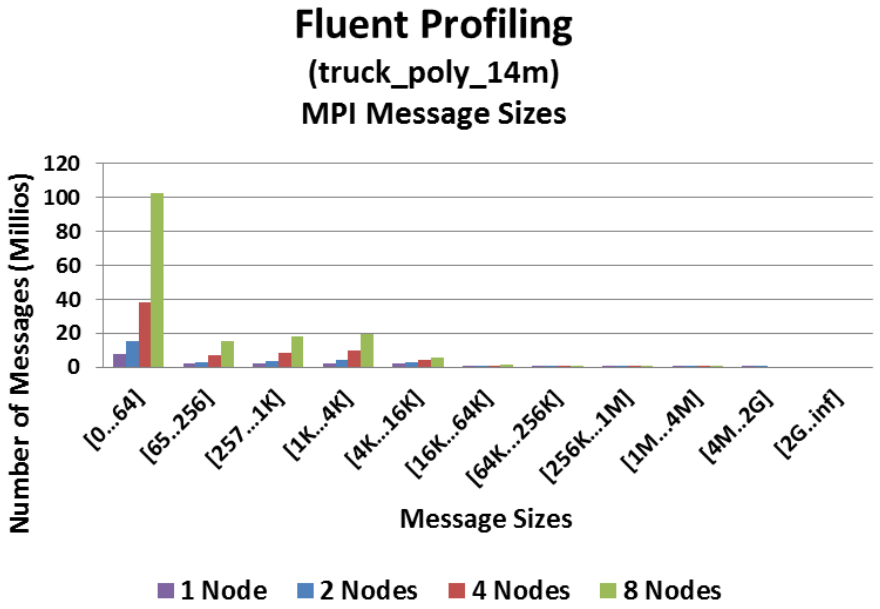




Fluent Profiling – MPI Message Sizes

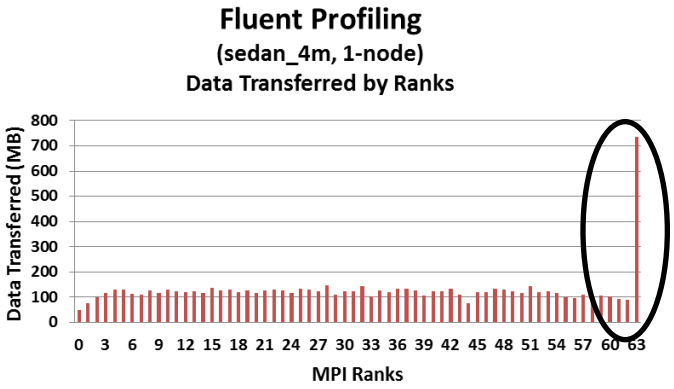
- **MPI message sizes are concentrated in range of small message sizes**
  - Majority are in the range of 0B and 64B
  - Small messages are typical used for synchronization, implies Fluent is latency sensitive
- **Large message sizes do exist but at a smaller percentage**
  - Larger messages (65B to 4MB) responsible for data transfers between the MPI ranks
  - Implies that Fluent does data movement which requires good network throughput

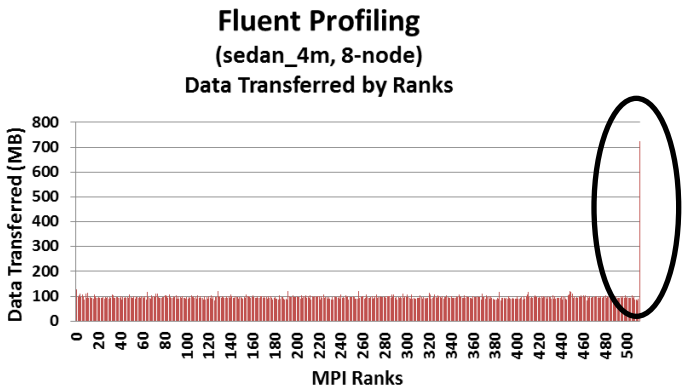


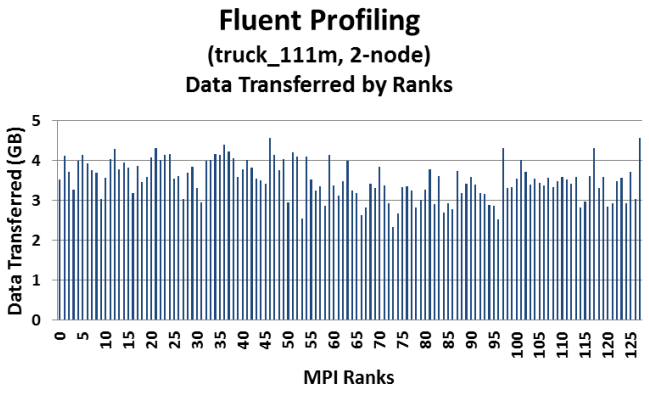


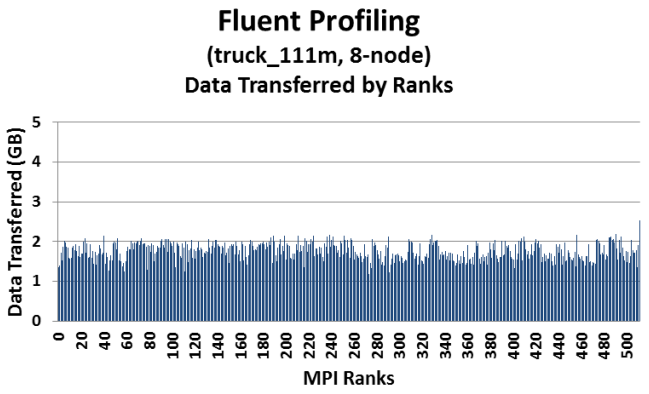
Fluent Profiling – Data Transfer Per Process

- **Data transferred to each MPI rank is generally the same except for the last**
  - Around 500MB per MPI rank for truck_poly_14m; around 2GB to 4GB for truck_111m
  - The last MPI rank has a significantly higher data rate than the rest for sedan_4m
  - Relatively



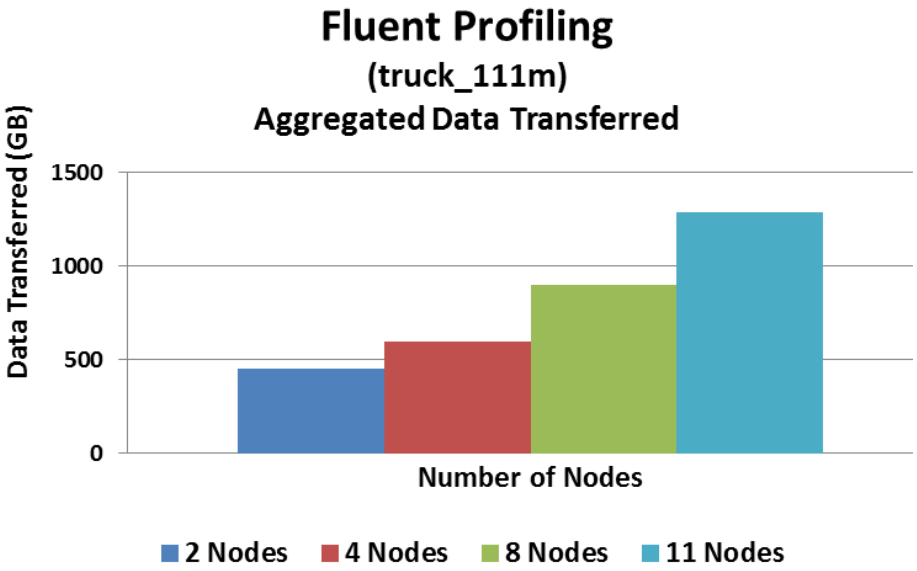


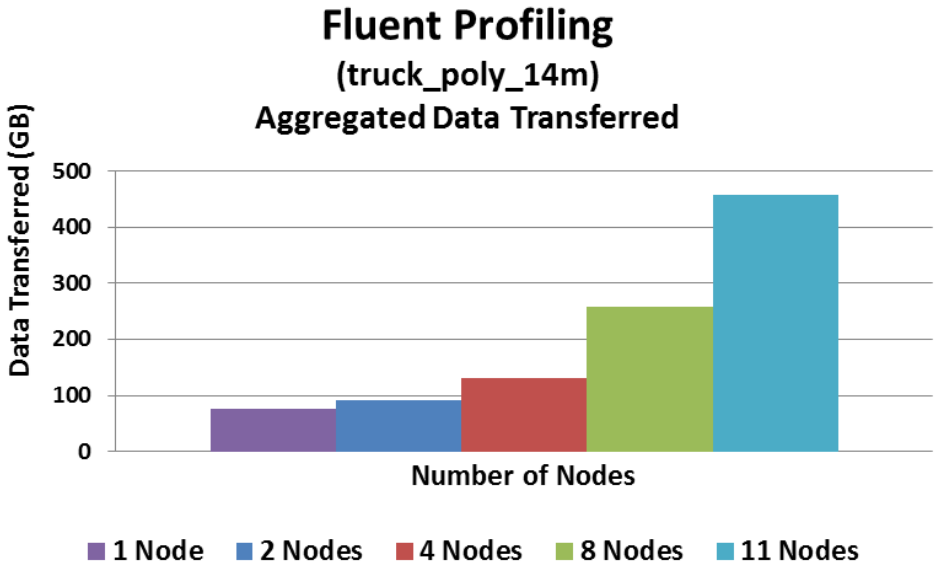




Fluent Profiling – Aggregated Data Transfer

- **Aggregated data transfer refers to:**
  - Total amount of data being transferred in the network between all MPI ranks collectively
- **The total data transfer steadily increases as the cluster scales**
  - As a compute node being added, more data communications will happen
- **Substantially more communications occur for larger dataset**
  - The rate of the increase is consistent between the 2 datasets





QDR InfiniBand

Summary

- **Fluent is a leading CFD application from ANSYS**
- **CPU**
  - AMD Opteron 6380 “Abu Dhabi” CPUs provides higher system utilization over prior generations
    - Delivers up to 72% higher performance versus AMD Opteron 6172 “Magny-Cours” CPUs
    - Delivers up to 15% higher performance versus AMD Opteron 6276 “Interlagos” CPUs
  - Running with 4P servers (64 cores/node) delivers twice the performance over a 2P server
  - 4P platform provides up to 76% higher performance compared to best published ANSYS benchmark results
- **Networking**
  - QDR InfiniBand allows Fluent to scale as it provides low latency and high throughput:
    - Runs up to 159% faster compared to 10GbE at 11 nodes (704 cores)
    - Runs up to 12 times faster compared to 1GbE at 11 nodes (704 cores)
  - Ethernet solutions do not scale; performance declines from 4 nodes and beyond
  - Running XRC with InfiniBand delivers 40% higher performance at 11 nodes (or 704 cores)
- **Data transfer on the network**
  - Significantly more data being transferred for the larger dataset
  - Tends to increase steadily as cluster scales
- **MPI**
  - Shows Fluent uses a range of MPI API for communications and synchronizations

Thank You

HPC Advisory Council