ANSYS CFX Performance Benchmark and Profiling

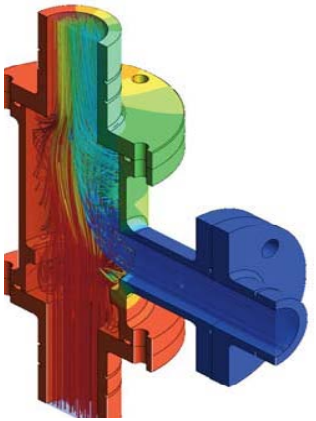
July 2009

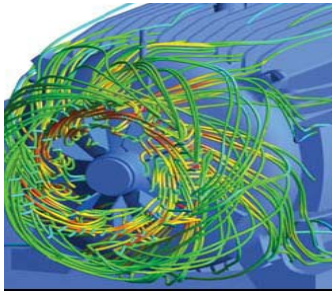
Note

- **The following research was performed under the HPC Advisory Council activities**
  - Participating vendors: AMD, ANSYS, Dell, Mellanox
  - Compute resource - HPC Advisory Council High-Performance 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 CFX

- **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 CFX is a high performance, general purpose CFD program**
  - All physical models in the ANSYS CFX solver work in parallel





Objectives

- **The presented research was done to provide best practices**
  - ANSYS CFX performance benchmarking
  - Interconnect performance comparisons
  - Ways to increase ANSYS CFX productivity
  - Understanding ANSYS CFX communication patterns
  - Power-efficient simulations

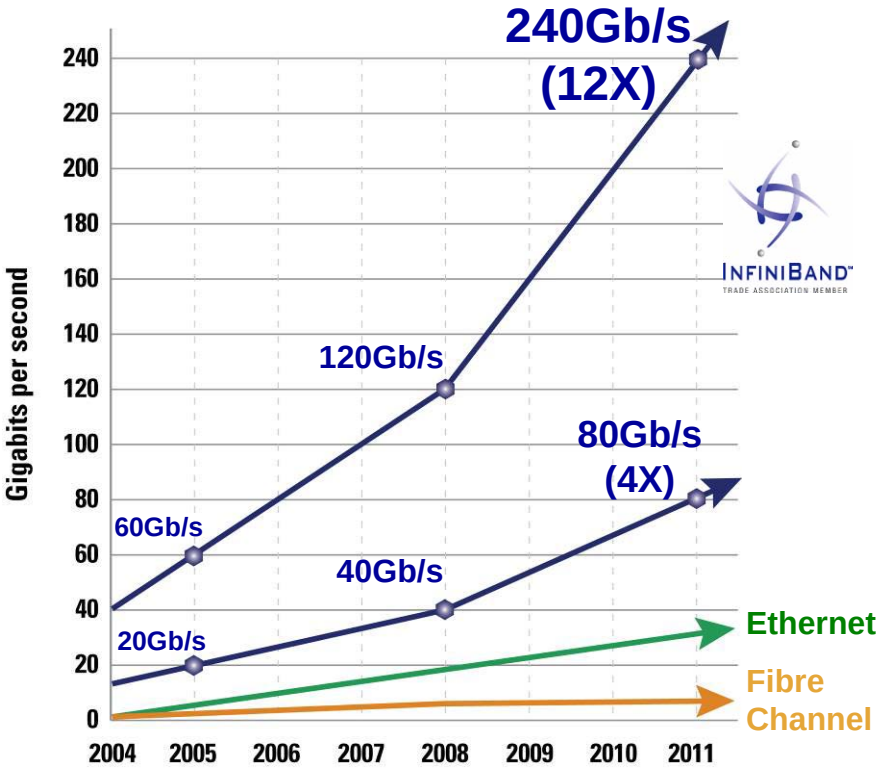
Test Cluster Configuration

- Dell™ PowerEdge™ SC 1435 20-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: ANSYS CFX 12.0
- Benchmark Workload
  - CFX Benchmark Dataset - Pump

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**

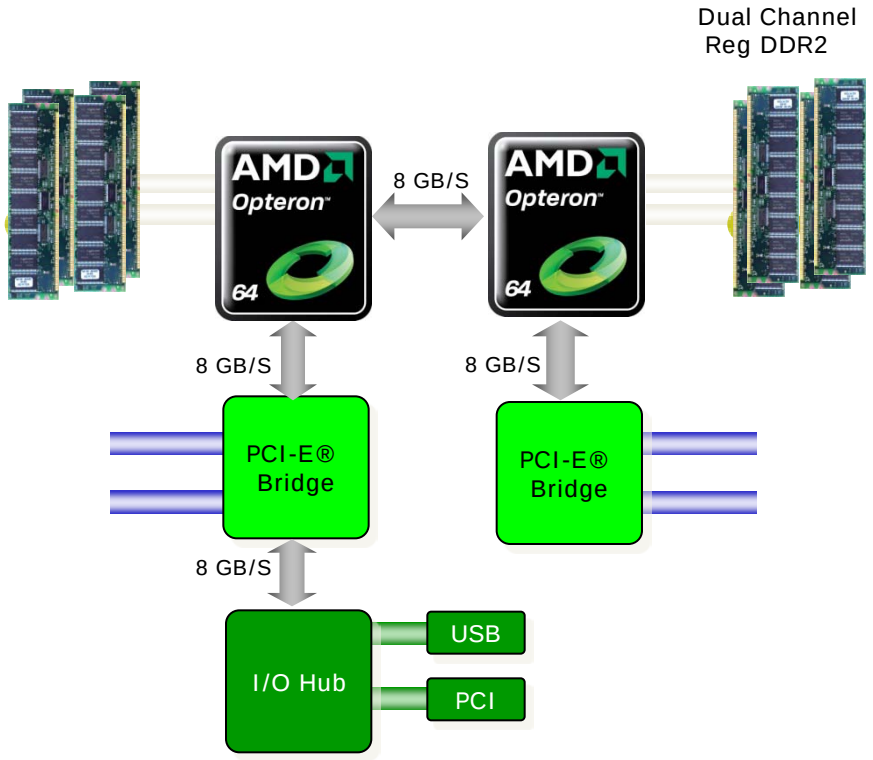
The InfiniBand Performance Gap is Increasing



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

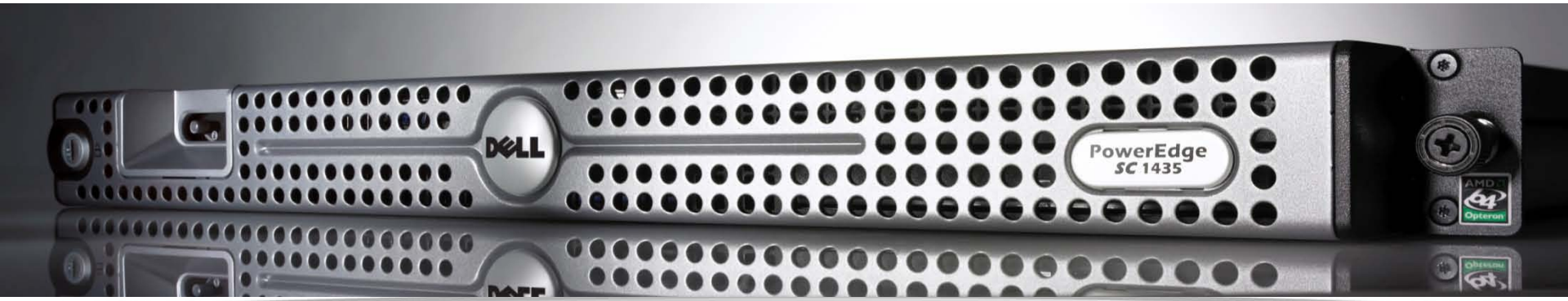


Dell PowerEdge Servers helping Simplify IT

- **System Structure and Sizing Guidelines**
  - 20-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

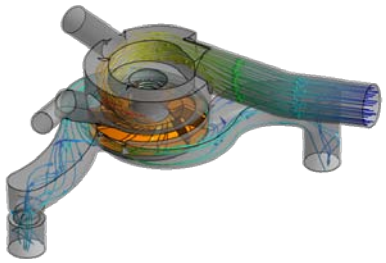
Dell PowerEdge™ Server Advantage

- **Dell™ PowerEdge™ servers incorporate AMD Opteron™ and Mellanox ConnectX InfiniBand to provide leading edge performance and reliability**
- **Building Block Foundations for best price/performance and performance/watt**
- **Investment protection and energy efficient**
- **Longer term server investment value**
- **Faster DDR2-800 memory**
- **Enhanced AMD PowerNow!**
- **Independent Dynamic Core Technology**
- **AMD CoolCore™ and Smart Fetch Technology**
- **Mellanox InfiniBand end-to-end for highest networking performance**



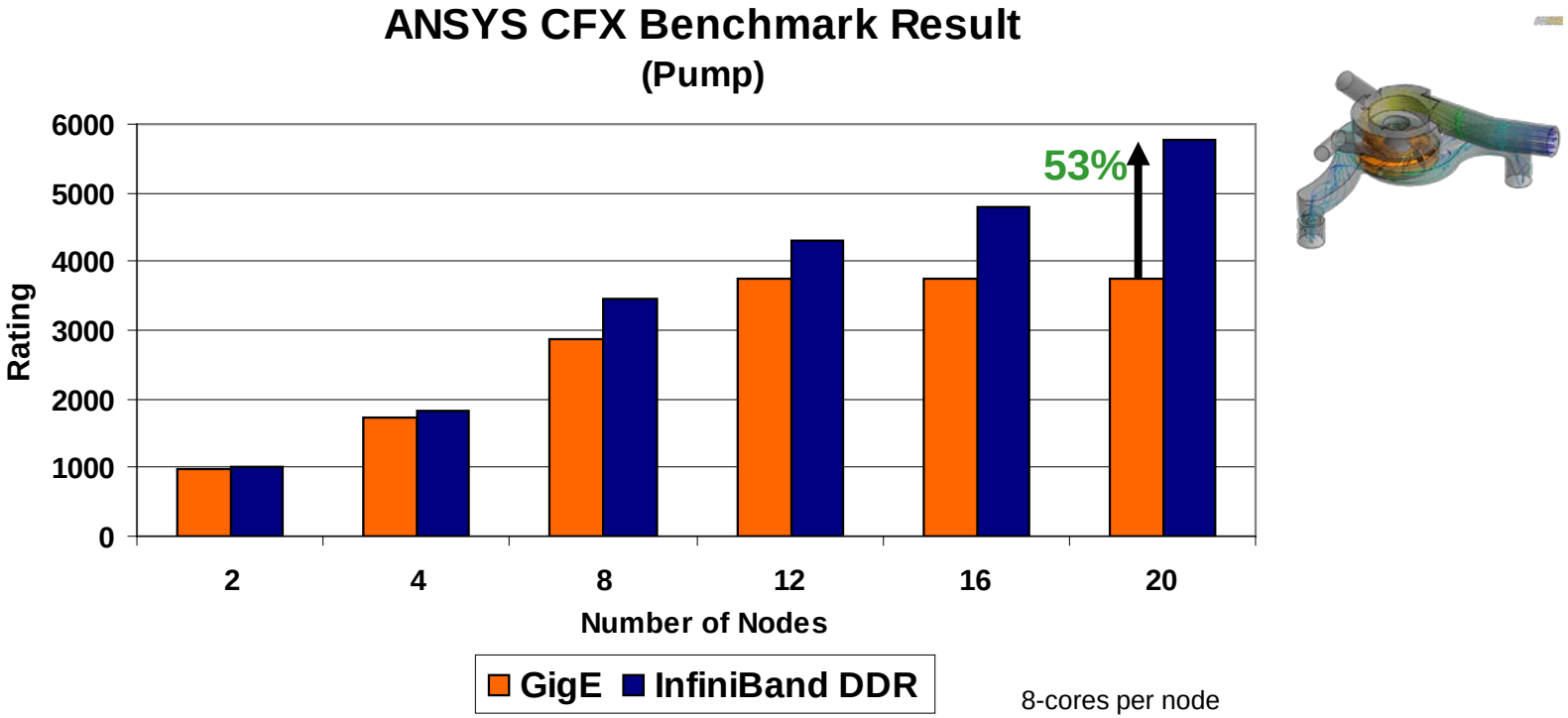
ANSYS CFX Pump Benchmark

- **Multiple frames of reference**
  - Rotating and stationary components
- **Unstructured mesh with tetrahedral, prismatic, and pyramid elements**
- **Total mesh size: approx. 600000 mesh nodes**
  - Implies following approximate mesh sizes per core
    - 2 servers using 8 cores: approx. 37500 nodes per core
    - 20 servers using 8 cores: approx. 3750 nodes per core
    - 2 servers using 4 cores: approx. 75000 nodes per core
    - 20 servers using 2 cores: approx. 15000 nodes per core



ANSYS CFX Benchmark Results

- **Input Dataset: Pump benchmark**
- **InfiniBand provides higher utilization, performance and scalability**
  - Up to 53% higher performance versus GigE with 20 nodes configuration
  - Continue to scale while Ethernet max up system utilization at 12 nodes
  - 12 nodes with InfiniBand provide better performance versus any cluster size with GigE



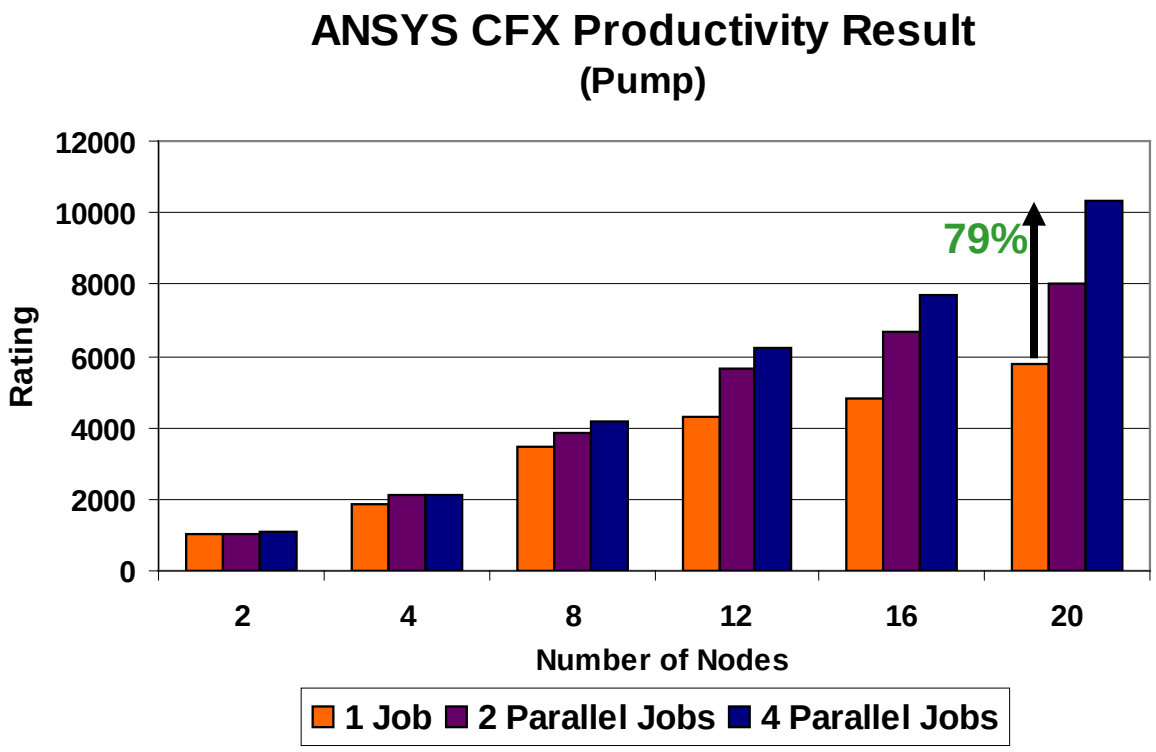
8-cores per node

Rating=24*3600 / CFD Solver wall clock time

Higher is better

Enhancing ANSYS CFX Productivity

- **Test cases**
  - Single job, run on eight cores per server
  - 2 simultaneous jobs, each runs on four cores per server
  - 4 simultaneous jobs, each runs on two cores per server
- **Running multiple jobs simultaneously can significantly improve CFX productivity**
  - Pump benchmark shows up to 79% more jobs per day

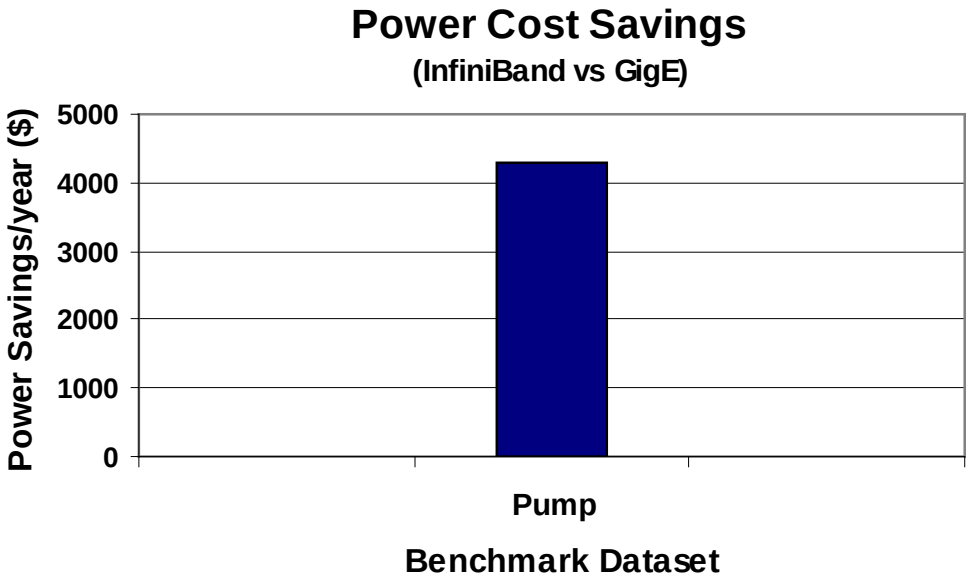


Higher is better

InfiniBand DDR

Power Cost Savings with Different Interconnect

- **Dell economical integration of AMD CPUs and Mellanox InfiniBand saves up to $4000 in power**
  - Versus Gigabit Ethernet as interconnect
  - Yearly based for 20-node cluster
- **As cluster size increases, more power can be saved**



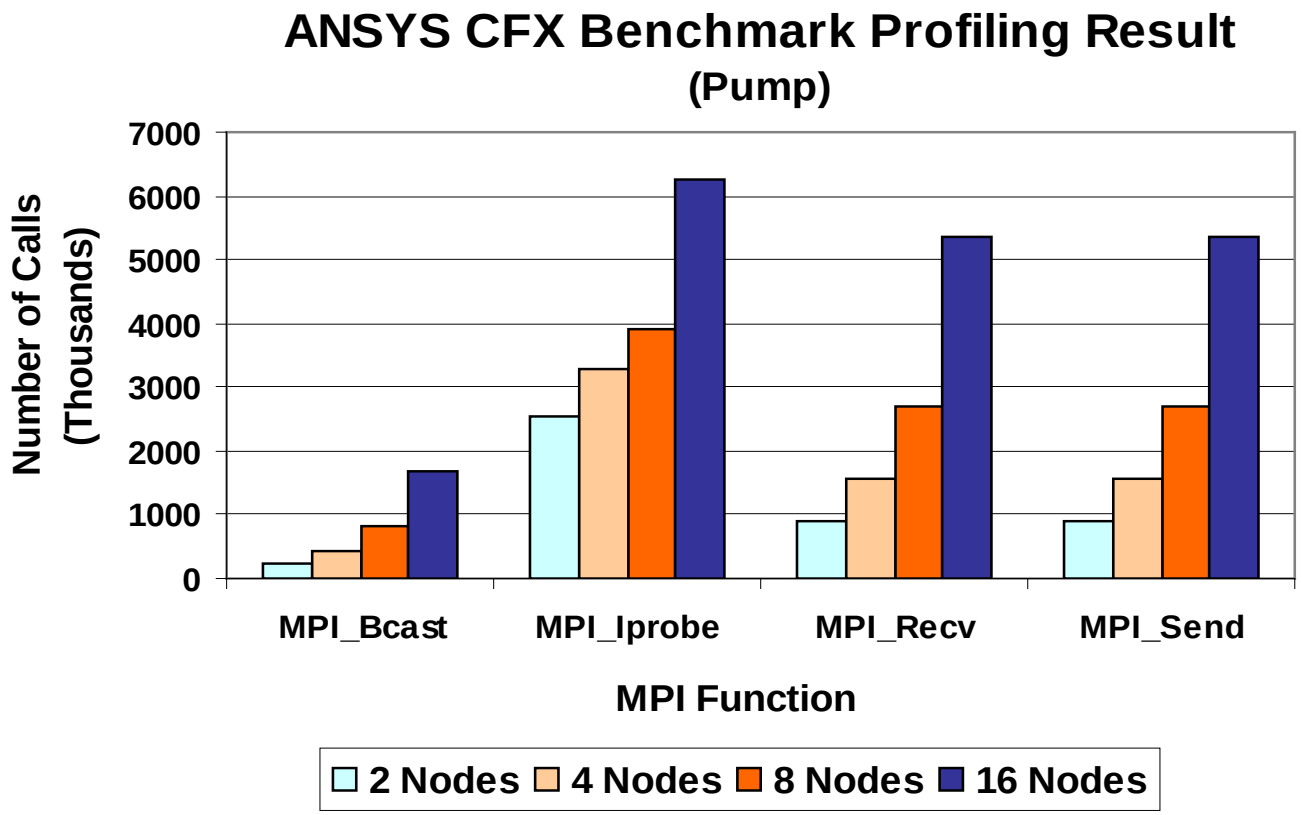
*$/KWh = KWh * $0.20*
For more information - http://enterprise.amd.com/Downloads/svrpwrusecompletefinal.pdf

ANSYS CFX Benchmark Results Summary

- **Interconnect comparison shows**
  - InfiniBand delivers superior performance in every cluster size
  - 12 nodes with InfiniBand provide higher productivity versus 20 nodes with GigE, or any node size with GigE
  - Performance advantage extends as cluster size increases
- **Efficient job placement**
  - Can increase CFX productivity significantly
  - Running 4 jobs concurrently can enhance productivity by up to 79% on 20 node cluster
  - Productivity advantage increases as cluster size grows
- **Power saving**
  - InfiniBand enables up to $4000/year power savings versus

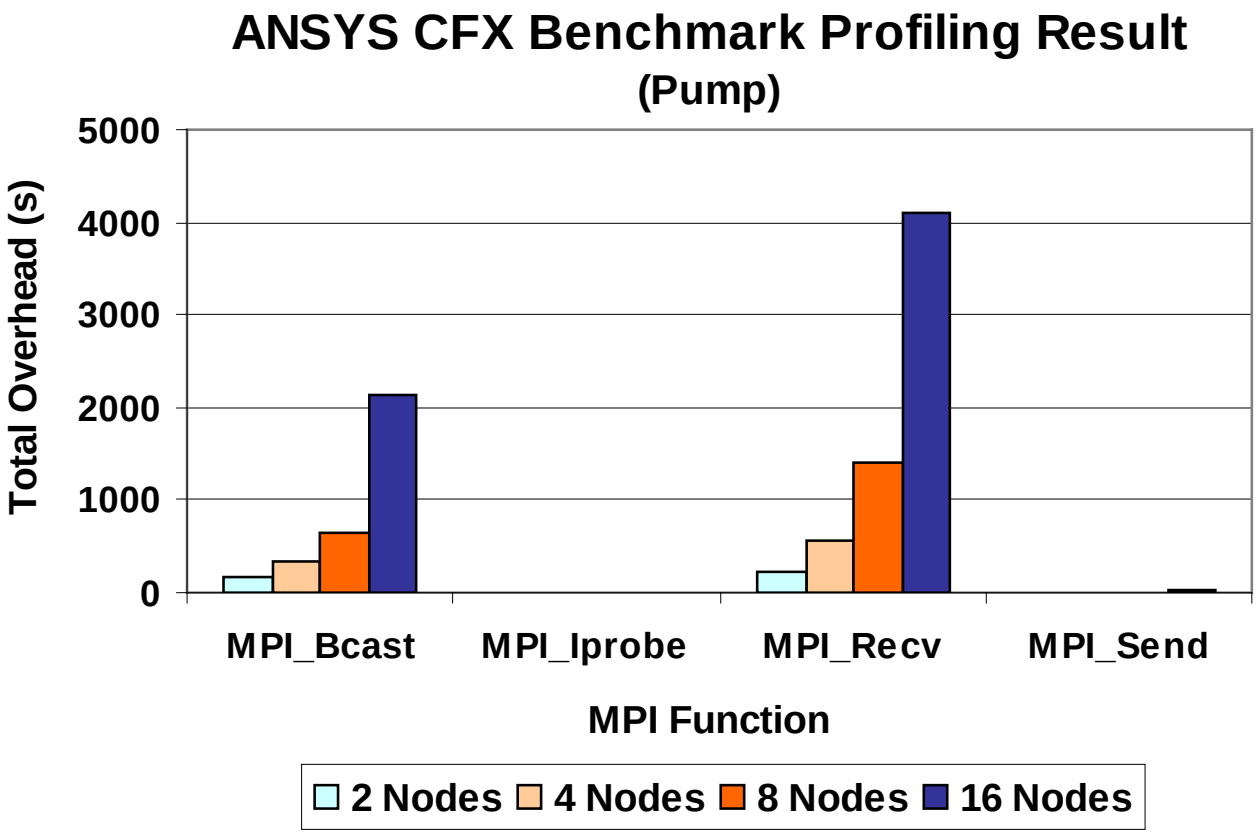
ANSYS CFX Profiling – MPI Functions

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



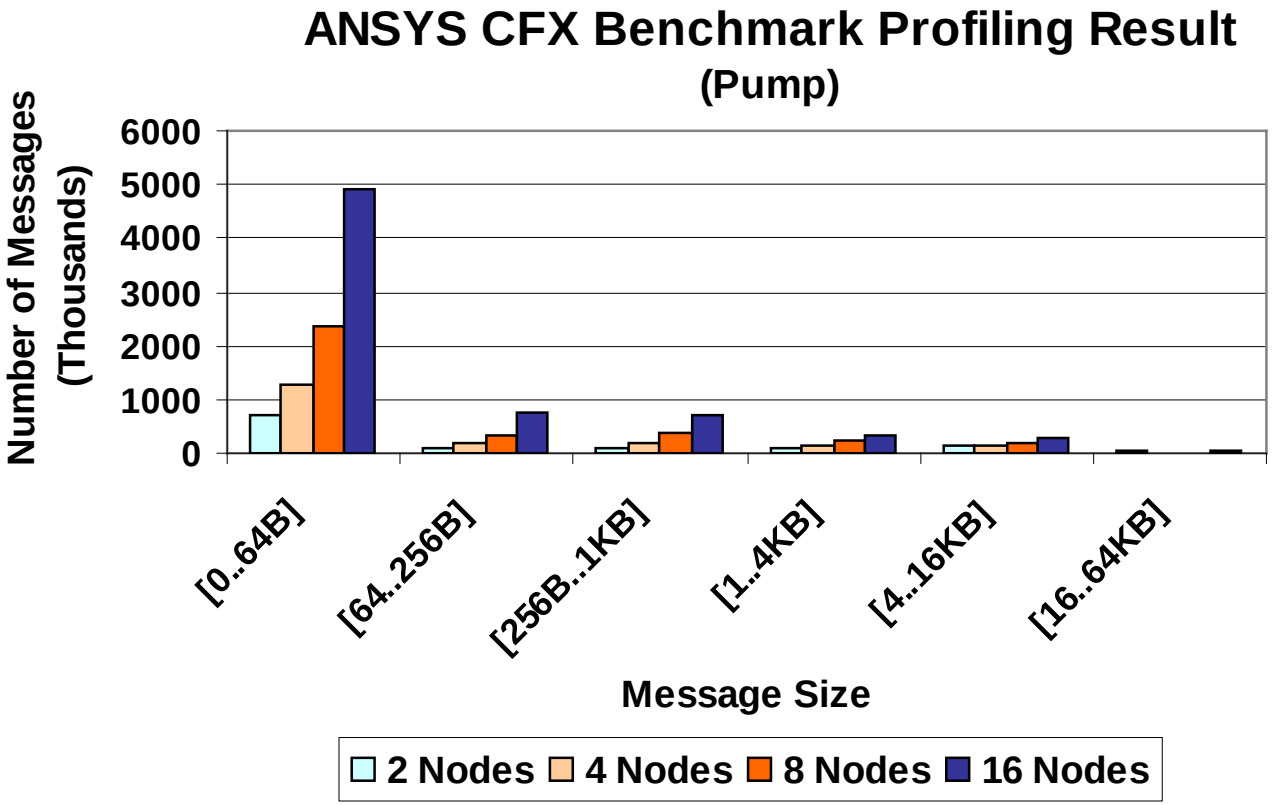
ANSYS CFX Profiling – Timing

- MPI_Recv and MPI_Bcast show the highest communication overhead



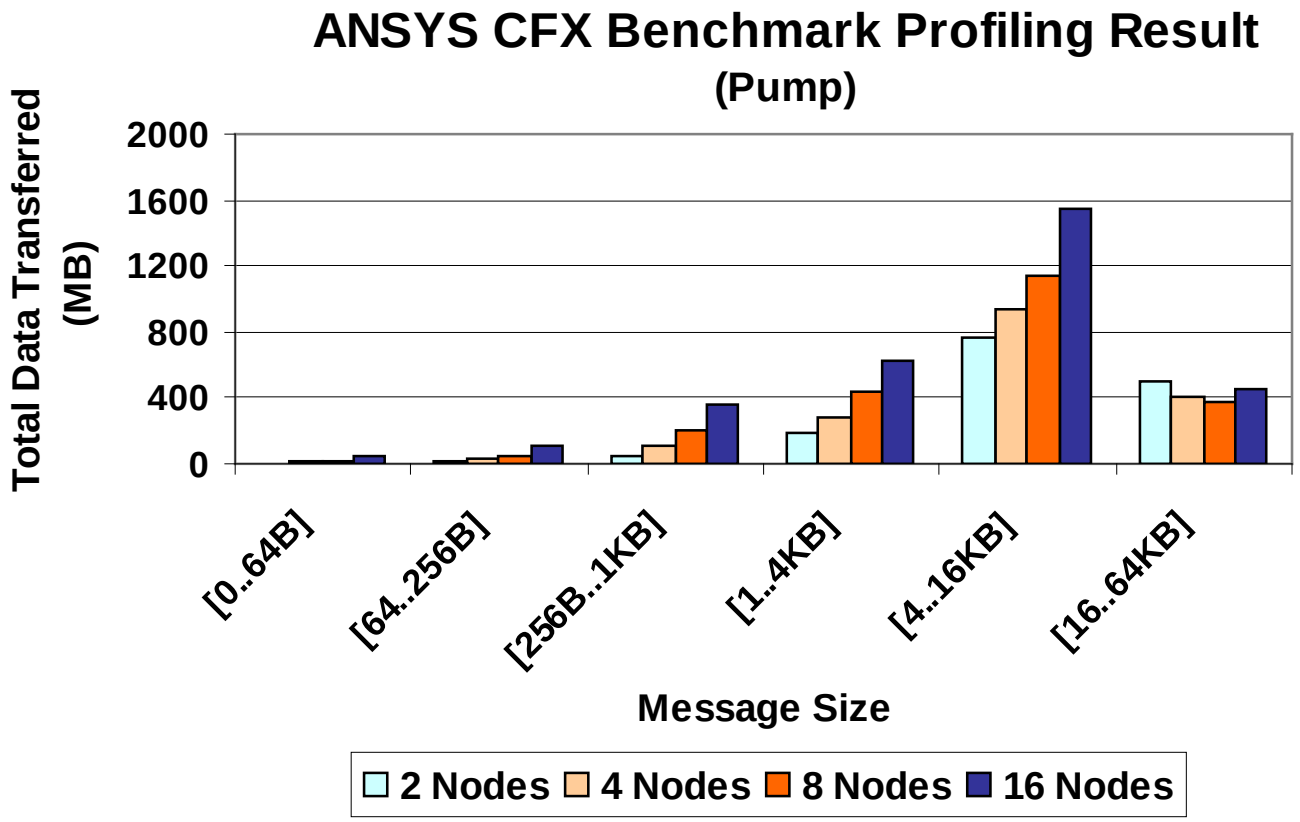
ANSYS CFX Profiling – Message Transferred

- Typical MPI synchronization messages are lower than 64B in size
- Number of messages increases with cluster size



ANSYS CFX Profiling – Message Transferred

- Most data related MPI messages are within 256B-64KB in size
- Total data transferred increases with cluster size



ANSYS CFX Profiling Summary

- **ANSYS CFX 12.0 were profiled to identify their communication patterns**
- **Frequent used message sizes**
  - 256B-64KB messages for data related communications
  - <64B for synchronizations
  - Number of messages increases with cluster size
- **Interconnects effect to ANSYS CFX performance**
  - Both interconnect latency (MPI_Bcast) and throughput (MPI_Recv) highly influence CFX performance
  - Further optimization can be made to take bigger advantage of high-speed networks

Thank You

HPC Advisory Council