Quantum ESPRESSO Performance Benchmark and Profiling

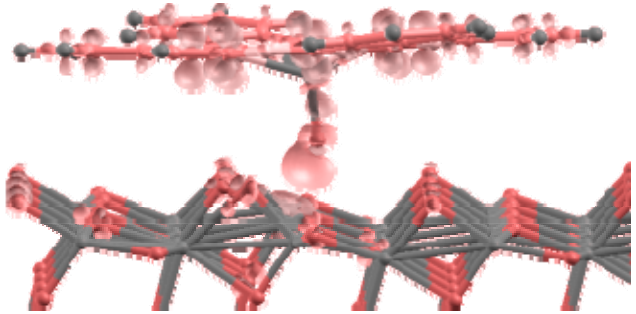
June 2010

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**
  - www.mellanox.com, www.dell.com/hpc, www.amd.com
  - http://www.quantum-espresso.org

Quantum ESPRESSO

- Quantum ESPRESSO stands for opEn Source Package for Research in Electronic Structure, Simulation, and Optimization
- It is an integrated suite of computer codes for electronic-structure calculations and materials modeling at the nanoscale
- It is based on
  - Density-functional theory
  - Plane waves
  - Pseudopotentials (both norm-conserving and ultrasoft)
- Open source under the terms of the GNU General Public License

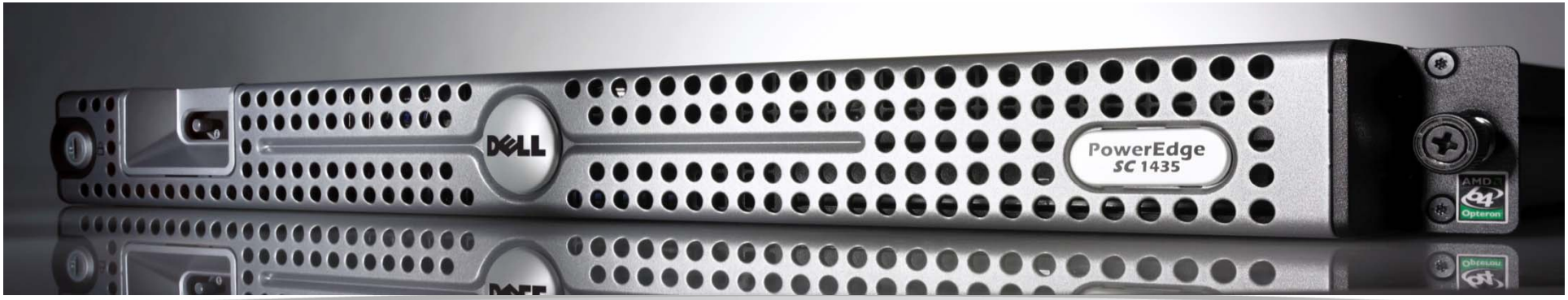


Test Cluster Configuration

- **Dell™ PowerEdge™ SC 1435 24-node cluster**
- **Quad-Core AMD Opteron™ 2382 (“Shanghai”) CPUs**
- **Mellanox® ConnectX EN 10 GigE Adapter Card**
- **Fulcrum 10 GigE Switch**
- **Memory: 16GB memory, DDR2 800MHz per node**
- **OS: RHEL5U3, OFED-1.5-RDMAoE SW stack**
- **MPI: OpenMPI-1.3.3**
- **Application: Quantum ESPRESSO 4.1.2**
- **Benchmark Workload**
  - Medium size DEISA benchmark AUSURF112

Dell PowerEdge™ Server Advantage

- **Dell™ PowerEdge™ servers incorporate AMD Opteron™ and Mellanox Mellanox ConnectX Ethernet 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**



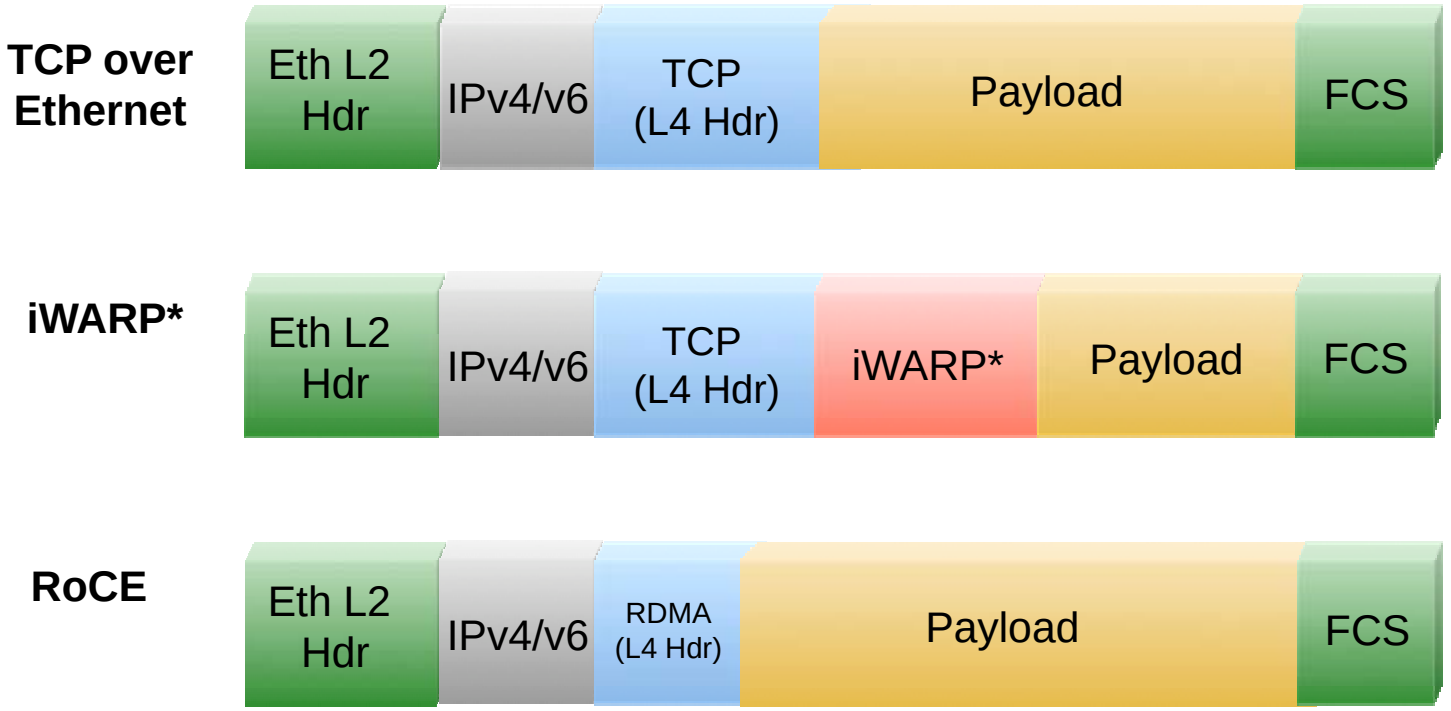
RDMA for Ethernet - RoCE

- **RoCE is an efficient RDMA solution for Ethernet**
  - Reliable/unreliable, connected/datagram
  - Unicast and multicast
- **Provide scalable and flexible RDMA solution for data centers**
- **Operates on Data Center Bridging (DCB) networks**
  - Efficient, light-weight transport over Layer 2 (L2) Ethernet
  - Ensure L2 interoperability on existing Ethernet infrastructure
  - Takes advantage of virtual links (per-priority pause)

RoCE Benefits

- **Real low latency Ethernet**
- **Kernel bypass, SEND/RCV, atomic operations**
- **UDP, multicast**
- **High clustering efficiency**
- **Higher scaling than TCP based or iWARP based**
- **Support for 10 and 40G Ethernet**

Packet Formats (TCP/IP, iWARP, RoCE)

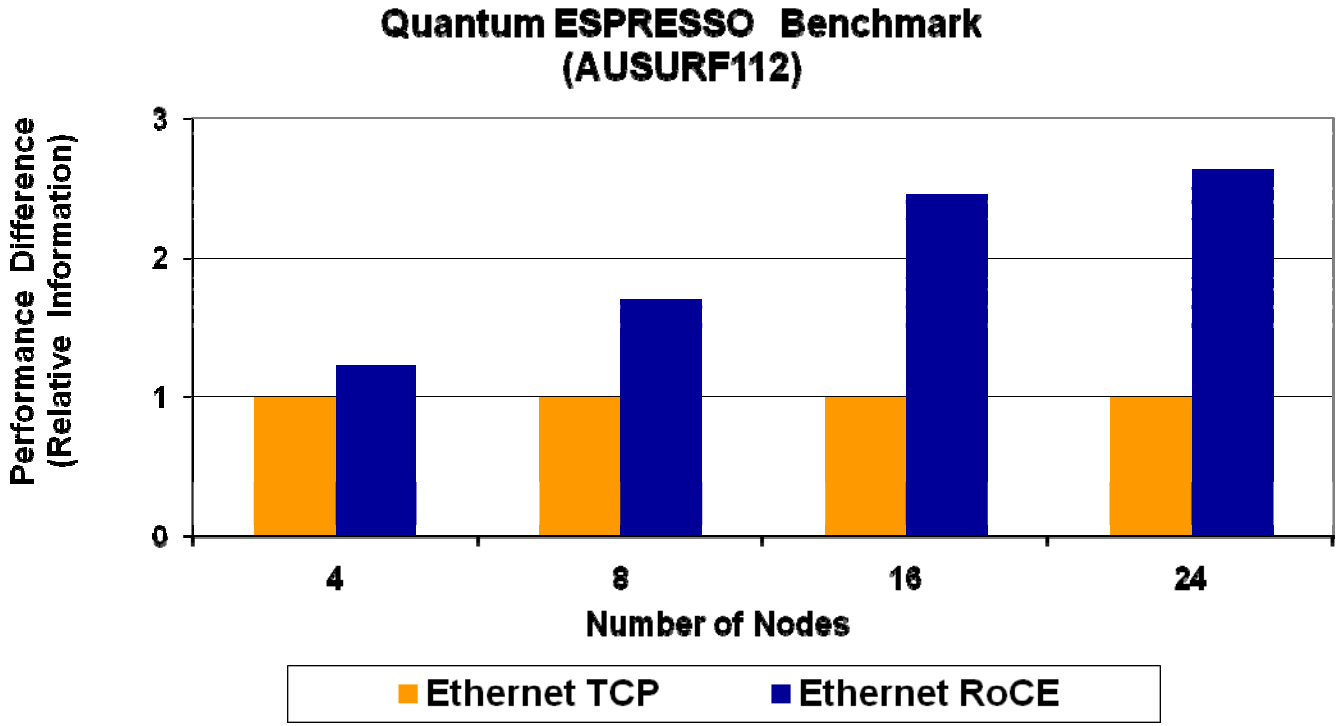


- **Layer 2 RoCE**
  - Ethernet (MAC,VLAN ID) based forwarding
  - Ethernet Pause based flow control

iWARP* = RDMAP/RDDP/MPA

Quantum ESPRESSO Benchmark Results

- **RoCE enables better application performance and scalability**
  - Up to 164% higher performance versus TCP over 10GigE
- **Application performance over RoCE scales as cluster size increases**



Higher is better

8-cores per node

Thank You

HPC Advisory Council