Network Performance Benchmarks
Mellanox 40GE NICs, Caltech Fast Data Transfer (FDT) Application and MonALISA

NOV 2010
Azher Mughal
Caltech
Servers and Network Test Setup

**Server:**
SuperMicro 6016XT-TF  
Dual Intel X5670, 24GB DDR3 RAM  
Supporting 4 Gen2.0 x8 slots  
OS: RedHat Linux, Kernel 2.6.36

**Mellanox Driver:**
16 Receive queues + 16 Transmit queues  
Challenge: Optimize queues to CPU cores

---

24 Core Server with two Mellanox 40GE NICs

Server - 1

24 Core Server with two Mellanox 40GE NICs

Server - 2

Two NICs bonded together in Linux with round robin packet flow
Network Test: Two 40GE bonded NIC

FDT “nettest” buffer tests, 16 streams in each direction
Memory Test: Two 40GE bonded NIC

FDT reading from /dev/zero and writing to /dev/null memory tests, 16 streams in each direction
## UltraLight Kernel 2.6.36-UL3

### Requires manual IRQ shift across CPU Cores

### Improved CPU idle time and load balancing across cores
Summary

• FDT Network Tests utilizes the available PCI-Express Gen 2.0 x8 bandwidth available to 1 x 40GE NIC to a maximum of 46Gbps (IN+OUT).

• FDT Network Tests with 2 x 40GE NICs bonded together achieves a maximum of 70Gbps (IN+OUT).

• FDT Memory Tests achieves a maximum of 62Gbps (IN +OUT) when reading from memory based source /dev/zero and dropping packets to /dev/null.

• **Note:** PCI-e Gen 2.0 x8 slot offers ~24.6 Gbps real world throughput in one direction with about 7Gbps overhead due to 8b/10b encoding (20% loss per direction) and PCI-E layered architecture.

• **Note:** These transfer rates were achieved after careful IRQ steering among the CPU Cores (Dual Xeon X5670, 24 total cores).

• There is still room for more improvements in terms of Kernel, IRQ assignment and FDT.

• PCI-E 3.0 expected early 2011 will use 128b/130b encoding thus reducing loss and providing an effective double rate than PCI-E 2.0 at 1GB/s per lane or 64Gbps FD in an x8 slot.