Interconnect Panel

HPC User Forum
Norfolk, Virginia
April 2008

Josh Simons
Distinguished Engineer
Sun Microsystems
The Two Winners

• Ethernet
  > 10 GbE acceptance will grow as prices fall
  > iWARP with hardware RDMA support will extend Ethernet reach into HPC

• InfiniBand
  > Finally – “The MPI of Interconnects”
  > And, more important, starting to get traction with enterprise customers (e.g. Financial Services desire for high msg rates)
  > Aggressive bandwidth roadmap, excellent latencies ensure IB longevity
Density Attributed to Connector Choice

- Sun's Datacenter Switch 3456
  - 1152 physical CSFP connectors
  - Yields 3456 4x DDR ports in single chassis (W48” x H62” x D28.5”)
  - Connects to CSFP enabled leaf switch (NEM) in each blade shelf

Not drawn to scale
IBTA 12X Standard Expected May/June 08

- Sun utilized 12x CSFP on the Magnum 3456 switch with straight thru and splitter copper cables
  - Both passive and active designs, supporting lengths from 3M to 16M
  - Cable reports S/N, length, & equalization parameters to switch
  - Superior mechanical design to CX4 with 3:1 cable reduction

- IBTA defining a new 12x Multi-Source Agreement (MSA)
  - Compact Small Form-Factor Pluggable (CSFP) Transceiver
  - Dual-Channel CSFP has the same mechanical dimensions as the industry standard SFP transceiver
  - Compatible with the standard SFP cage
Sun Multithreaded 10GbE

- Dual 10 GbE on-board
- Multithreaded to support CMT (24 TX; 16 RX DMA channels)
- On-chip version: UltraSPARC T2 NIU
- Plug-in card available
- Crossbow network stack
Some Sun Interconnect Work

• System-on-a-Chip (SOC)
  > On-silicon interconnect support (e.g. UltraSPARC T2 NIU)
• Possible semantics beyond current IB / RDMA
  > MPI acceleration
  > Support for new programming models (e.g. PGAS)
• UNIC – Ultraperformance Nanophotonic Intrachip Communications