COMPUTER & COMPUTATIONAL SCIENCES

13th IEEE International Conference on High-Performance Interconnects
(IEEE Hot Interconnects), Palo Alto, CA, October 18-19, 2005

Panel

EtherNET vs. EtherNOT

Wu Feng
feng@lanl.gov

Research & Development in Advanced Network Technology (RADIANT)
http://www.lanl.gov/radiant
Computer & Computational Sciences Division
University of California, Los Alamos National Laboratory

NNSA

Los Alamos NATIONAL LABORATORY

HotI Background: EtherNET *and* EtherNOT

- "The Quadrics Network (QsNet): High-Performance Clustering Technology," *9th IEEE Hot Interconnects*, August 2001.
- "Initial End-to-End Performance Evaluation of 10-Gigabit Ethernet," *11th IEEE Hot Interconnects*, August 2003.
- "Performance Characterization of a 10-Gigabit Ethernet TOE," *13th IEEE Hot Interconnects*, August 2005.

Current Perceptions?

- Performance → *EtherNOT*
  - Better end-to-end latency and throughput.
- Cost → *EtherNET (GigE), EtherNOT (10GigE)*
  - Gigabit EtherNET virtually free!
  - 10-Gigabit EtherNET: $795/NIC, ~$1000/port
- Features → *EtherNOT (or maybe EtherNET)*
  - Support for one-side & collective communication
  - On-demand pinning of buffers & on-demand memory registration support
  - (Off-loaded TCP/IP)
- Compatibility & Ubiquity → *EtherNET*

Current Customer Perceptions?

- HPC Customer
  - See previous slide
- Mainstream Customer
  - Jeopardy! answer to "Quadrics and InfiniBand":
    - "What are the alpha & beta codenames for XM radio?"
- Relative to Market Share & Customer Needs
  - EtherNOT
    - The Formula One Racecars of Networking.
    - Exception: Myricom's Myrinet-2000 and now Myri-10G
      - The High-End Sports Car of Networking :-)
  - EtherNET
    - The Honda Accord of Networking

Is It Really EtherNET vs. EtherNOT?!

- Maybe 5-10 years ago …
- Today: General convergence of EtherNET & NOT
  - EtherNOT working to deal with …
    - Congestion control (see HotI'05 and OpenIB'04 & '05)
    - More generally, TCP/IP support?
  - EtherNET working to deal with …
    - General protocol offload engines (POEs): RDDP, iWARP, uDAPL
      - "Connectionless" iWARP, on-demand pinning of buffers for ESDP & iWARP.
      - PCI-X → PCI-Express
    - Routing: Virtual cut-through (SAN), store-and-forward (WAN)
- Divergence: Congestion control resides where?!

Current Observations

- Identifying Trends
  - How fast is fast? Approaching speed-of-light limits.
  - EtherNET sustained halving of prices every 12 months (or less) while enhancing feature set, e.g., virtual cut-through, better support for iSCSI, iWARP, uDAPL, and so on, while maintaining compatibility.
  - Target markets: HPC, which is a TINY market.
    - Exceptions: EtherNET, and perhaps EtherNOT's Myri-10G
  - EtherNET controller may become more EtherNOT-like, or perhaps the other way around (LANai on the motherboard?)
    - Comment: HPC will *not* drive this. Commodity market has to.

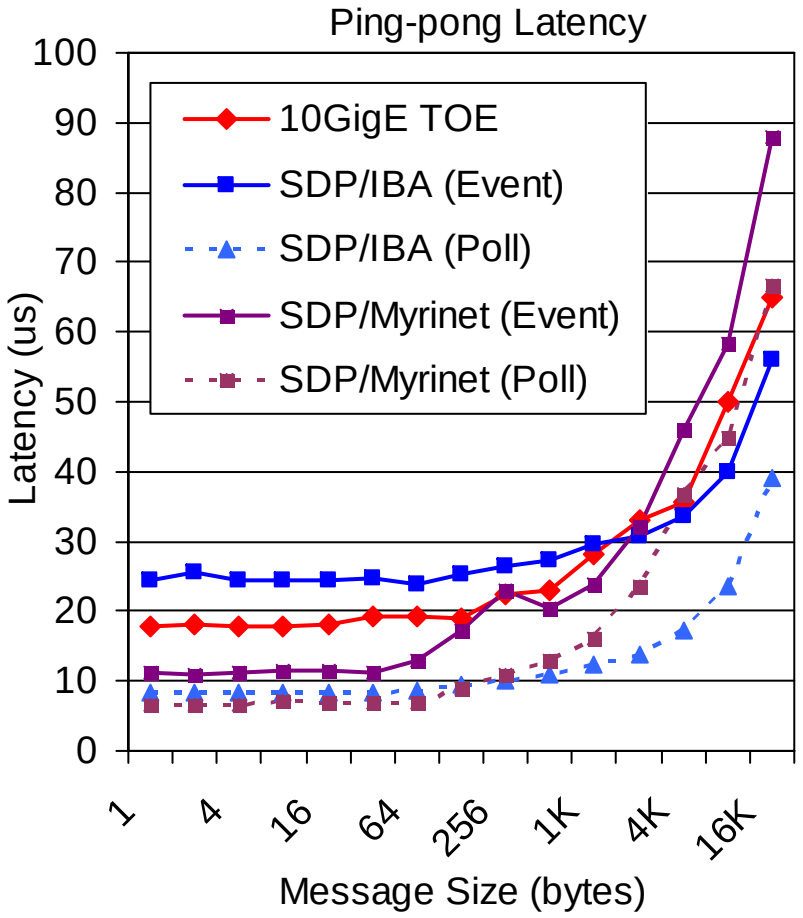
Ethernet Trends

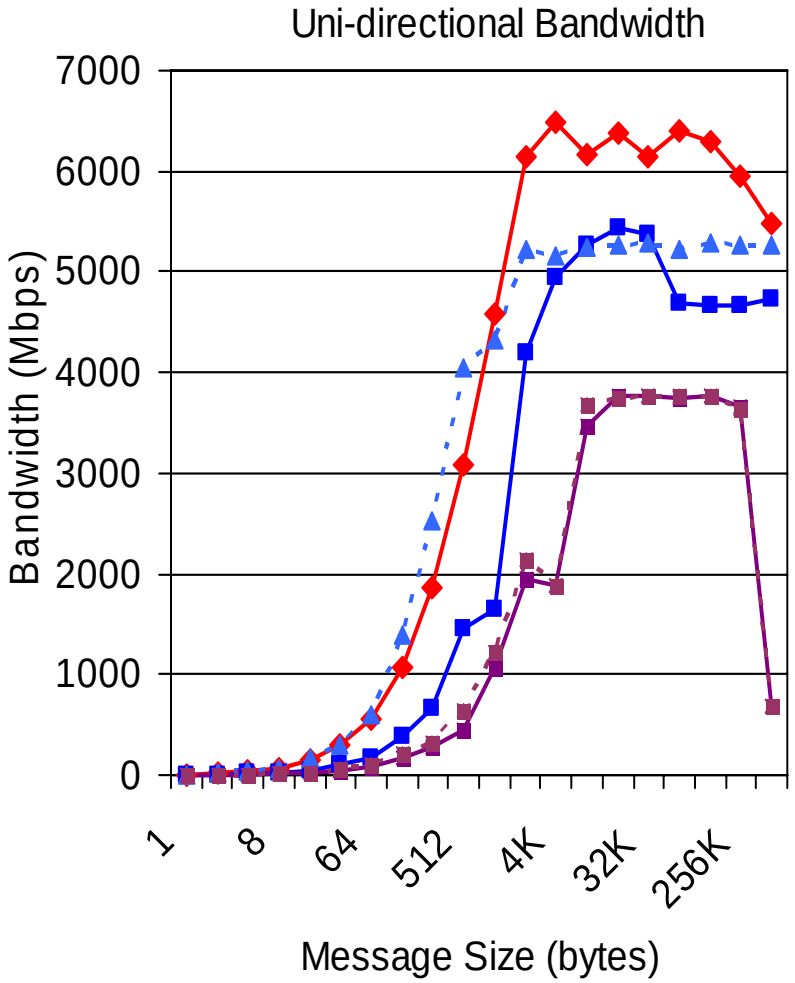
- Top500 Supercomputer List
  - 2001: 0% of Top500 is Gigabit Ethernet
  - 2005: 42.4% of Top500 is Gigabit Ethernet
    70.6% of Top500 is EtherNET+Myrinet
    - And now that Myricom has Myri-10G …
- Ethernet Everywhere
  - Suspect that 100% of Top500 has an Ethernet network.
- Price/Performance of EtherNET
  - Reasonable performance at low cost and complexity.

- Disturbing trend in reporting performance
- Virtually all BW (e.g., 9+ Gb/s / stream) and end-to-end (e.g., 2 us) latency numbers are produced in isolation. These numbers are generally not achieved without exhaustive tuning and with an application.

COMPUTER & COMPUTATIONAL SCIENCES

EtherNET vs. EtherNOT: Latency/Bandwidth (circa July 2005)



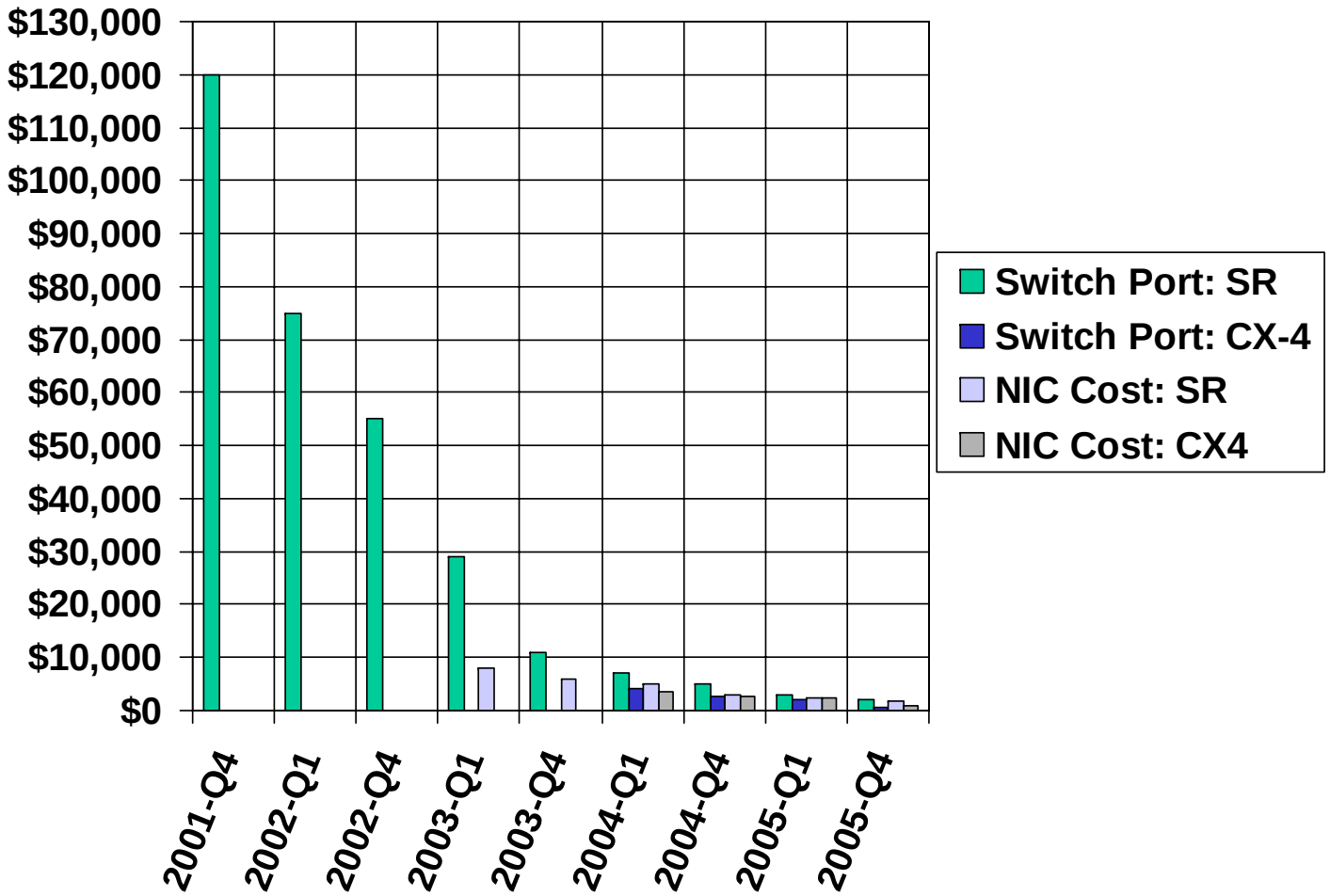


Recent SDP/Myrinet (Poll): 4.5 μs

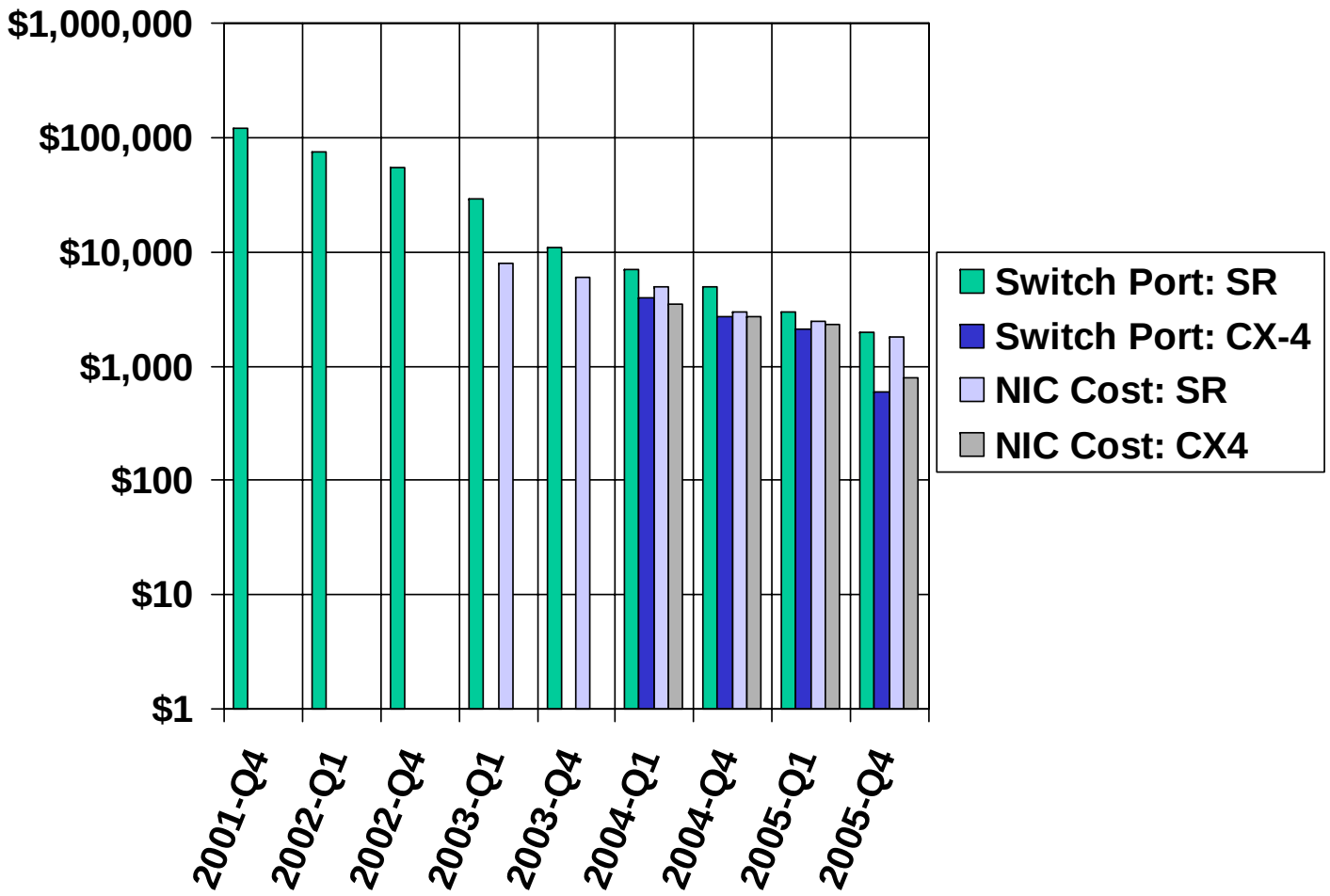
EtherNET vs. EtherNOT Paper

- P. Balaji, W. Feng, Q. Gao, R. Noronha, W. Yu, and D. Panda, "Head-to-TOE Evaluation of High-Performance Sockets over Protocol Offload Engines," *7th IEEE International Conference on Cluster Computing (IEEE Cluster)*, Boston, MA, September 2005.

Exponential Drop in Pricing



Exponential Drop in Pricing



Concluding Remarks

- I don't know what the next programming language will look like, but it will be called *Fortran* …
- I don't know what the next transport protocol will look like, but it will be called *TCP/IP* …
- I don't know what the next interconnect technology will look like, but it will be called *Ethernet* …
  - 80% of all networks are Ethernet.
  - 99.9% of all Internet traffic is over Ethernet
- What is the point that I am trying to make? Inevitably, the above ubiquitous technologies adopt the best features of non-ubiquitous solutions. Latest for 10-Gigabit Ethernet:
  - Network processors on adapters. TOE. RDDP/iWARP/iSCSI.
  - Next, virtual cut-through routing for 10-Gigabit Ethernet in system-area networks.
    - Problem: Eliminates the advantage of ubiquitous deployment (for now).

Concluding Remarks

- Performance
- Cost
- Features
- Compatibility & Ubiquity

- Differentiator:
  - Target Markets

- A final pet peeve …