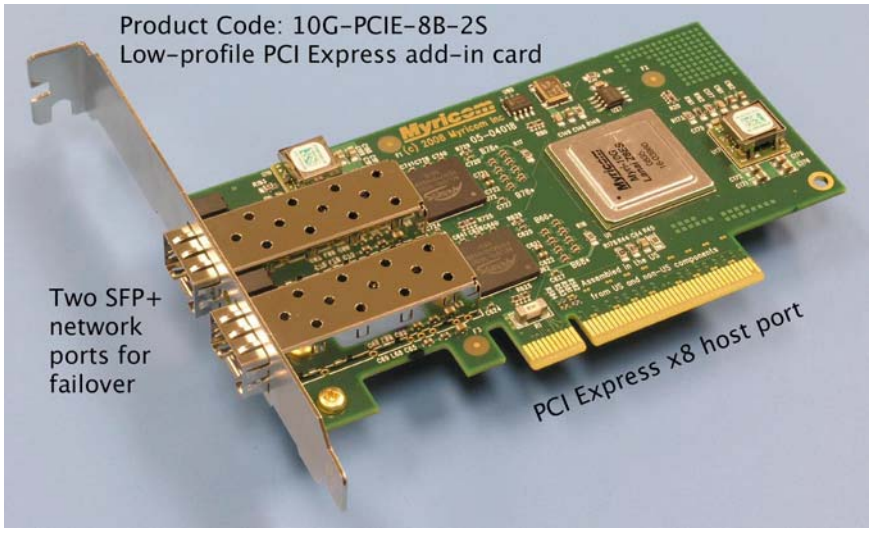
Myri-10G NICs and Software

10-Gigabit Ethernet with a Supercomputing Heritage

Myricom, the company that pioneered high-performance cluster interconnect, launched its Myri-10G family of products in 2005. Myri-10G is not only a new generation of Myrinet, the very successful High Performance Computing (HPC) "specialty network" that Myricom created and has been shipping since 1994, but is also a performance-optimized implementation of mainstream 10-Gigabit Ethernet, both NICs and switches.

Myricom labeled Myri-10G as "a **convergence** that leverages 10-Gigabit Ethernet into the HPC world, and HPC techniques into the Ethernet world." An insightful pundit amusingly called Myri-10G "10-Gigabit Ethernet with special sauce." Marketing phrases and sound bites aside, Myri-10G products immediately won customers, some for conventional 10-Gigabit Ethernet networking applications, some for HPC applications, and some for special applications such as IPTV video streaming. Why? Myri-10G Network Interface Cards deliver uncompromising, wire-speed performance – a Myricom focus and tradition – at an affordable price, and are fully interoperable with the 10-Gigabit Ethernet products of other companies. Myri-10G switches scale economically to thousands of ports, a capability not offered by the traditional Ethernet-switch vendors.

Myri-10G Network Interface Cards (NICs) are unique in the 10-Gigabit Ethernet NIC market in being internally programmable. For conventional TCP/IP and UDP/IP operation, the processors and firmware in the NIC are used for highly effective stateless offloads, resulting in **wire-speed throughput with low host-CPU utilization.**



Myricom supplies Ethernet drivers with associated firmware for **Linux**, **Windows** (WHQL certified), **Solaris**, **Mac OS X**, **FreeBSD**, and **VMware ESX**.

Myri-10G NICs are available as standard low-profile PCI Express add-in cards with a choice of 10-Gigabit Ethernet ports: 10GBase-CX4, 10GBase-SR and 10GBase-LR (either XFP or SFP+ transceivers), SFP+ Direct Attach, or QSFP for either copper or fiber cables. These NICs are also available in special form factors for blades, including the IBM BladeCenter H. Although the network ports adhere to 10-Gigabit Ethernet Physical (PHY, layer-1) standards, Myri-10G NICs are capable with different firmware of supporting either Ethernet or Myrinet network protocols at the Data Link level (layer 2).

© 2010, Myricom, Inc.

Myri-10G NICs with one port or with two ports for failover connect to hosts through PCI Express x8, a 2+2 GigaByte/s full-duplex I/O fabric that is fast enough, even after PCI Express protocol overhead, to keep up with a 1.25+1.25 GigaByte/s network port. NICs with two ports for performance connect to hosts through "Gen2" (5 GT/s) PCI Express x8 to provide sufficient bandwidth for two wire-speed 10-Gigabit ports. Myri-10G NICs were the first wire-speed 10-Gigabit Ethernet NICs, and they are also great 10-Gigabit Myrinet NICs.

Myri-10G Performance and Software

The excellent **netperf TCP/IP benchmarks** below are with Linux (2.6.18) between servers with two Intel quad-core 2.93GHz Xeon X5570s:

Netperf Test	MTU	BW	TX_CPU %	RX_CPU %
TCP_STREAM	9000	9910.33	4.52	2.84
TCP_SENDFILE	9000	9910.32	2.71	2.82
TCP_STREAM	1500	9477.10	4.62	5.57
TCP_SENDFILE	1500	9452.54	2.56	5.63

This throughput performance and low host-CPU utilization is achieved without troublesome stateful offloads such as are used by "TCP Offload Engines" (TOEs), which necessarily "break the protocol stack" (see http://www.linux-foundation.org/en/Net:TOE). Myri-10 Ethernet software, which includes both driver and NIC firmware, implements zero-copy on the send side with all operating systems, and, depending on the OS, uses a variety of stateless offloads in the driver and NIC firmware, including:

- IP and TCP checksum offload, send and receive
- Interrupt Coalescing
- TSO (TCP Segmentation Offload, also known as Large Send Offload)
- RSS (Receive-Side Scaling)
- LRO (Large Receive Offload)
- Multicast filtering

Myricom also supplies optional, firmware-accelerated software distributions for low-latency, low-host-CPU-load, kernel-bypass operation of Myri-10G NICs over either 10-Gigabit Ethernet or 10-Gigabit Myrinet networks, most notably:

- MX (Myrinet Express) for MPI and Sockets communication, and for cluster file systems. MPI and Sockets middleware is included in the MX software distribution, and alternative middleware that operates directly over MX is available from other sources.

- Video Pump™ for UDP streaming of IPTV video at ultra-low host-CPU utilization.

- DBL (Datagram Bypass Layer) for low-latency UDP/IP communication. DBL is used principally for high-speed financial trading.

- Sniffer10G™ for wire-speed 10-Gigabit Ethernet packet capture at high packet rates.

Myricom® and Myrinet® are registered trademarks of Myricom, Inc.

© 2010, Myricom, Inc.