

Incorporación de 10-Gigabit Ethernet a Myricom

En cuanto al middleware de comunicación, la inmensa mayoría está desarrollado por Myricom, y distribuido bajo la fórmula de Software Libre. Destacan las librerías a bajo nivel GM y MX, la implementación de MPI MPICH-GM y MPICH-MX y las implementaciones de Sockets de alto rendimiento Sockets-GM y Sockets-MX. Los drivers y firmware de las tarjetas están disponibles para Linux, Windows, Solaris 10, Mac OS X y FreeBSD. Para apreciaciones más específicas sobre la plataforma véase <http://www.myri.com/scs/performance/Myri10GE/>.

Los nuevos productos de Myrinet están orientados a complementar funcionalidades de las redes de 10Gb Ethernet. Ethernet tiene una gran cantidad de drivers para muchas plataformas (Windows, Linux, Mac, Solaris,...) con el máximo rendimiento, el mínimo coste y cumpliendo con los estándares Ethernet. Sobre este tipo de MAC funcionan TCP/IP y UDP/IP con un rendimiento de procesamiento entorno a los 9.8 Gb/s.

La intención de unificar los criterios de baja latencia de 10-Gigabit Ethernet y la gran variedad de soporte software que proporciona Myrinet hace que se cree MX (Myrinet Express). MX permite que la conexión de las MAC a los switch 10-Gigabit Myrinet o 10-Gigabit Ethernet se haga como si se tratase de una interfase Ethernet con una menor latencia, comunicaciones MPI y las API's de los sockets.

Todos los productos cumplen la directiva RoHS para la restricción de ciertas sustancias peligrosas en aparatos eléctricos y electrónicos. Entró en vigor el 1 de julio de 2006, pero no es una ley, es simplemente una directiva. Esta directiva restringe el uso de seis materiales peligrosos en la fabricación de varios tipos de equipos eléctricos y electrónicos.

Trjetas Myri-10G (dual-protocol)

Características de los Enlaces físicos de 10-Gigabit Ethernet:

Los puertos de Myri-10G cumplen con XAUI¹, establecido como regulación en IEEE 802.3ae, donde entre otras características, su ancho de banda tiene que ser de 10+10Gb/s en conexión full-duplex.

El cableado puede ser sobre fibra o bien sobre cobre, deben de tener una distancia máxima 200m y 15 metros, respectivamente. Evidentemente sobre fibra se consiguen mayores ratios en la velocidad de transmisión.

Características básicas de funcionamiento

¹ XAUI: estándar donde se especifican todas las características que los puertos deben cumplir para realizar conexiones a 10 Gigabit Ethernet.

Latencia de 2.3µs con MPICH-MX, ancho de banda de 1.2 GB/s (al hacer un ping) en una dirección y 2.2 GB/s en las dos direcciones (enviar y recibir). A través de los switch 10-Gigabit Ethernet se consigue el mismo ancho de banda, pero la latencia está en un rango de 2.6 a 2.8µs.

En el modo de trabajo 10-Gigabit Ethernet o la emulación de Ethernet MX-10G se consiguen ratios 9.6 Gbits/s (TCP/IP).

MAC

Pueden operar bajo dos tipos de protocolos de red 10-Gigabit Ethernet y 10-Gigabit Myrinet. Como el resto de NIC que Myrinet produce incluyen su propio procesador y soporte para el procesamiento de la información a fin de liberar a la CPU de trabajo.

Las MAC se conectan a los ordenadores a través del slot PCI-X x8² que opera a 2+2 GB/s en una conexión full-duplex, suficiente para el límite en el que trabajan los puertos de la MAC (1.25+1.25GB/s) para lograr una red a 10Gb. Las tarjetas Myri-10G son las únicas en trabajar con velocidades tan grandes sobre un soporte físico basado en cobre.

En vistas a sus funcionalidades en servidores se puede instalar físicamente sobre los mismos sin necesidad de que se incorporen ningún elemento físico (raíles) como soporte.

Los distintos tipos de tarjetas se establecen según el tipo de medio físico por el que se envía la señal: 10GBase-CX4³, 10GBase-R⁴ o con fibra bajo especificaciones de XAUI.

Las diferencias entre los modelos 10G-PCIE-8A-** y 10G-PCIE-8AL-** es el tipo de especificaciones PCI-X que se requiere.

10GBase-CX4: 10G-PCIE-8A-C y 10G-PCIE-8AL-C

Tiene un precio de 674.1€. Su puerto trabaja a velocidades de 10+10 Gbit/s (full-duplex) bajo las especificaciones IEEE 802.3ak para 10GBase-CX4 (tipo de cable coaxial). Cuando opera en modo Ethernet el flujo de control que la tarjeta se comporta

² El PCI Express x4 / x8 / x12 estaban inicialmente destinados para operar en servidores, en concreto, el x8 puede llegar a alcanzar un ancho de banda de unos 40Gbps. Para más información:

http://www.dell.com/content/topics/global.aspx/vectors/en/2004_pcieexpress?c=us&l=en&s=corp

³ Cable coaxial especial, 4 cables en cada sentido a 3.125 Gbaud cada uno, distancia debe de ser menor a 15 m, sobre XAUI, codificación 4B5B.

⁴ 10GBASE-R (en función del tipo de fibra: SR, LR, ER). Varios medios físicos (PMD):

- Fibra MMF, 850 nm, distancia < 65 m. (“Short”)
- Fibra SMF, 1310 nm, distancia < 10 km (“Long”)
- Fibra SMF, 1550 nm, distancia < 40 km. (“Extended”)

Codificación 64B/66B. Modulación a 10.3 Gbaudios.

según las especificaciones definida por IEEE 802.3x. El puerto conecta con 10GBase-CX4 y no debe de sobrepasar los 15m. Pueden usar slots PCI Express x8 o x16.

En los host la información se envía a 2 GB/s (250 MB/s por cada hilo del cable) en conexión full-duplex.

Se incorpora el procesador Lanai-Z8E, el cual opera como mínimo a 300MHz (arquitectura RISC). Una memoria de 2MB (proporciona la información al procesador a 2,400 MB/s), EEPROM de 512KB con la configuración mínima necesaria (configuración del PCI-X, firmware entre otros) y reprogramable por el procesador. En el futuro se espera incorporar a esta memoria el driver de arranque de la tarjeta (etherboot).

Sus dimensiones físicas son alto 68.9mm, longitud 147.3mm y 22mm de anchura, su peso es de 88g. Requiere una alimentación de 3.3V (que proviene de slot del bus) con un consumo máximo de unos 8.3W. Soporta hasta un máximo de 55°C.

Cableado 10GBase-CX4

<i>Distancia</i>	<i>Código del producto</i>	<i>Precios</i>
1m, 30AWG ⁵	10G-CX4-01M	127.1 €
2m, 30AWG	10G-CX4-02M	135.6 €
3m, 30AWG	10G-CX4-03M	144.1 €
5m, 30AWG	10G-CX4-05M	161.1 €
7m, 28AWG	10G-CX4-07M	169.5 €
10m, 26AWG	10G-CX4-10M	195 €
15m, 26AWG	10G-CX4-15M	228.9 €
Dispositivo de testeo Loopback	10G-CX4-LOOPBACK	50.87 €

NOTA: Los cables que InfiniBand utiliza los mismos conectores Fujitsu CX4 que los cables 10GBase-CX4, pero los cables que InfiniBand utiliza no cumplen el estándar sobre 10GBase-CX4.

⁵ Mediad de clasificación de los cables atendiendo a la dependencia entre el diámetro y el área del conductor. Calibre 4/0 (4 ceros), al que corresponde el mayor diámetro, el número de ceros disminuye hasta alcanzar el valor 1/0. A partir de este valor el calibre del cable está asociado a un valor numérico creciente (2, 4, 6, etc). Es importante recordar que para estos calibres el diámetro del conductor se reduce cuando el valor numérico asignado aumenta.

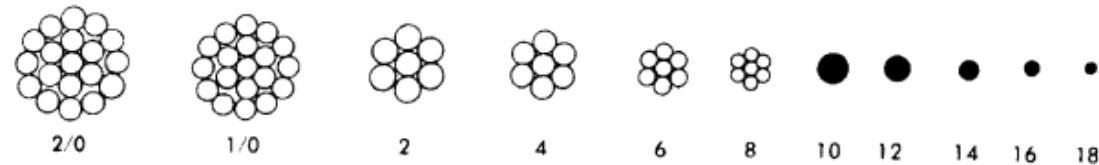


Fig. 8.1- Diámetros Relativos de Varios Calibres AWG



10GBase-R: 10G-PCIE-8A-R y 10G-PCIE-8AL-R

Ronda los 758.9 € y esto sin incluir el transceptor XFP, que se vende por separado. Puerto opera a 10+10 Gbit/s (full-duplex) sobre cada hilo de fibra la información se transmite a 10.3125 GBaud en una codificación 64b/66b. Dependiendo del transistor XFP⁶ que se encuentre ubicado la placa frontal el puerto conectará a 10GBase-SR (850nm de longitud de onda, 26-300m en fibra multimodo), a 10GBase-LR (1310nm de longitud de onda, 10km en fibra monomodo) o 10GBase-ER (1550nm de longitud de onda, más de 40 km de fibra monomodo). Puede operar con datos definidos en la capa Ethernet o Myrinet, en el primer modo las señales de control son la especificadas en IEEE 802.3x. La longitud permitida de la fibra depende del transmisor receptor de XFP y de la calidad de la fibra, pero en el modo Myrinet no debe exceder de los 200m.

Precio de los transceptores XFP para 10GBase-R

<i>Distancia</i>	<i>Código del producto</i>	<i>Precios</i>
10GBase-SR	10G-XFP-SR	423.9 €
10GBase-LR	10G-XFP-LR	763.1 €

⁶ Los transceptores XFP de 3Com son dispositivos hot-swappable estándar de la industria que se conectan a una ranura 10-Gigabit Ethernet, uniéndola con una red basada en fibra óptica o cobre. Véase la página:

http://www.3com.com/prod/es_LA_AMER/prodlist.jsp?tab=cat&cat=209632&subcat=229567

Cables de fibra multimodo(50/125), conexión duplex para 10GBase-SR

<i>Distancia</i>	<i>Código del producto</i>	<i>Precios</i>
1m	M3F-CB-1M	59.35€
3m	M3F-CB-3M	63.59 €
5m	M3F-CB-5M	67.83 €
10m	M3F-CB-10M	76.31 €
25m	M3F-CB-25M	127.1 €
50m	M3F-CB-50M	169.5 €

El resto de características es igual a la 10GBase-CX4, procesador, memoria... exceptuando que su longitud y peso es mayor 165.1 mm y 91g respectivamente (sin incluir el XFP) y su consumo máximo es de 10W.



XAUI sobre fibra: 10G-PCIE-8A-Q y 10G-PCIE-8AL-Q

Con un precio de 928.5 €. Los puertos cumplen XAUI sobre la fibra a través de la interconexión con MTP/MPO (pequeño conector de fibra capaz de conectar hasta 24 fibras) que permite ratios de 10+10 Gbit/s, full-duplex. El puerto es compatible con conectores XAUI o 10GBase-CX4, para la fibra tiene un “conversor hardware” externo, pero internamente funciona un POP4⁷ (4 dispositivos MTP/MPO). Los conectores de la fibra que se usen en el puerto deben de cumplir las especificaciones del estándar MTP/MPO fibra multimodo (50/125 hilos) y no superar los 200m.

⁷ <http://www.popoptics.org/>
http://www.popoptics.org/POP4_Specification.pdf

Físicamente sus dimensiones son similares a las 10GBase-CX4, pero su peso es superior (94g), pero con un menor consumo 6.3W.

Precio de cables de fibra bajo XAUI (sobre el cuarteto de puertos)

<i>Distancia</i>	<i>Código del producto</i>	<i>Precios</i>
1m	M3Q-CB-1M	127.1 €
3m	M3Q-CB-3M	148.3 €
5m	M3Q-CB-5M	169.5 €
10m	M3Q-CB-10M	211.9 €
25m	M3Q-CB-25M	339.1 €
50m	M3Q-CB-50M	466.3 €
100m	M3Q-CB-100M	635.9 €
150m	M3Q-CB-150M	847.9 €
200m	M3Q-CB-200M	1102 €



Protocolo iSCSI para realización de test sobre las distintas plataformas: El protocolo iSCSI utiliza TCP/IP para sus transferencias de datos. Al contrario que otros protocolos de red diseñados para almacenamiento, como por ejemplo Fibre Channel (que es la base de la mayor parte de las SANs), solamente requiere un simple y sencillo interfaz Ethernet (o cualquier otra red compatible TCP/IP) para funcionar. Esto permite una solución de almacenamiento centralizada de bajo coste sin la necesidad de realizar inversiones costosas ni sufrir las habituales incompatibilidades asociadas a las soluciones Fibre Channel storage area networks.

Los críticos de iSCSI argumentan que este protocolo tiene un peor rendimiento que el Fibre Channel ya que se ve afectado por la sobrecarga que generan las transmisiones TCP/IP (cabeceras de paquetes, por ejemplo). Sin embargo las pruebas que se han realizado muestran un excelente rendimiento de las soluciones iSCSI SANs, cuando se utilizan enlaces Gigabit Ethernet