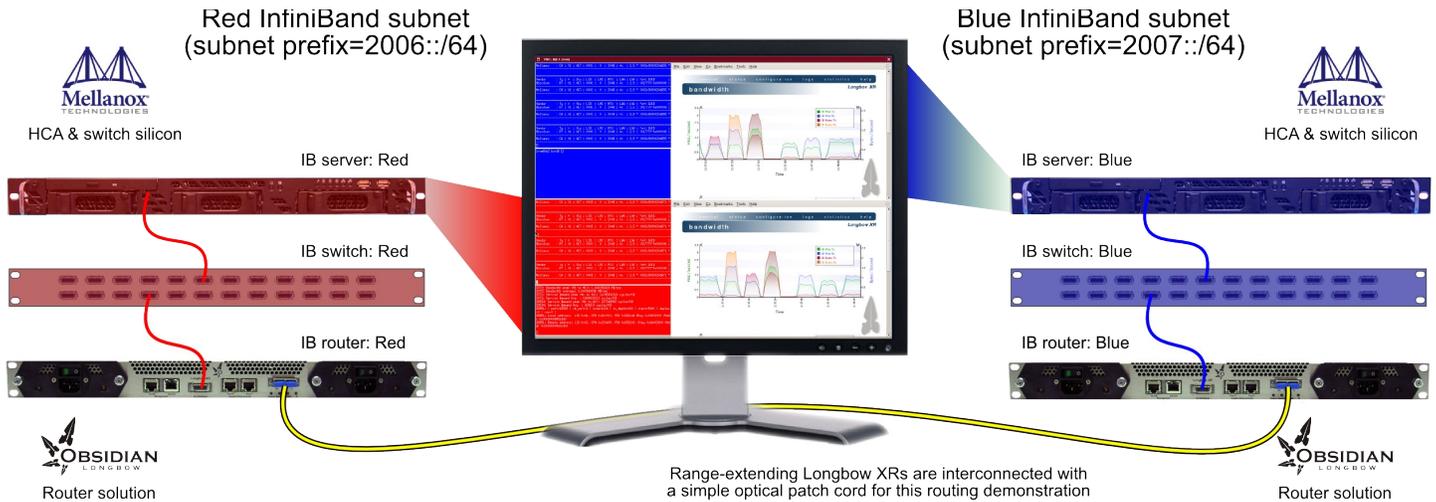


# XNet: Native InfiniBand Routing Hardware

Obsidian's InfiniBand range extension technology highlighted a need for subnet isolation through native InfiniBand routing. Longbow XR demonstrates its full bandwidth InfiniBand routing capabilities when deployed with Mellanox HCAs and switches by streaming data between two servers running in their own subnets.



## InfiniBand routing Background

To date, Obsidian's Longbow InfiniBand range extension technology has linked remote InfiniBand facilities by appearing as an InfiniBand switch to each side. Since each subnet hosts exactly *one* active subnet manager, the subnet manager's domain spans the range extended link. In some circumstances this may not be optimal on the grounds of scalability, performance or administrative access concerns.

Obsidian is announcing that it has equipped Longbow XR with both switch and now *router mode*. Router mode supports users wishing to preserve local subnet topology and management arrangements while exchanging data with other sites using native InfiniBand protocols.

## The Demonstration – a World First

Obsidian is pleased to announce that it has teamed with Mellanox Technologies to provide this XNet InfiniBand routing exhibit. Mellanox silicon powers the PCI-Express HCAs and switches that are used to connect a pair of servers through Longbow XR units. Each server is running its own subnet manager – OpenFabric's OpenSM. The two subnets are designated **Red** and **Blue**. The servers each drive three windows in their half of the display:

- The subnet manager windows indicate that two subnets are indeed present in the fabric, and that each subnet manager discovers exactly one HCA, one switch and one router.
- The traffic generator windows running the OpenFabrics ibperftest diagnostic reflect sustained full-duplex transfers between the two servers in separate subnets.
- The Longbow XR GUI windows chart the sustained bandwidth performance across the inter-subnet link.

The optical connection between the Longbow XRs could be an OC-192c SONET/SDH, ATM, 10GbE WAN or simply a dark fiber (Longbows defaulting to SONET mode in this case) – the link itself does not explicitly manifest as an InfiniBand connection in either subnet.

## Performance

Longbow XR's architecture features an entirely hardware-based datapath implementation. InfiniBand routers are required to provide high speed GID to LID address resolution; Longbow XRs include large CAMs for this function, supporting up to 64k routes. Sustained full-duplex throughput in switch or router mode matches that of an SDR 4x InfiniBand link – 1Gbyte/s. Additionally, Longbow XR is capable of maintaining this throughput across many thousands of kilometers.

The optical link between Longbows incurs an unavoidable flight-time delay of 5  $\mu$ s/km. Longbow XRs show a port-to-port latency of 1.5 $\mu$ s (switch mode) and 1.7  $\mu$ s (router mode) with small packets (store-and-forward).

## Subnet Manager support

While the InfiniBand v1.2 specifications adequately define router hardware, the details of subnet management interaction are not fully addressed. OpenSM, the OpenFabrics open source InfiniBand subnet manager, has been adjusted to recognize routers and assist in the initialization of the fabric and Longbow XR routing tables. Obsidian is working with Mellanox and others to complete the specification and reflect the changes in OpenSM.

*It is important to recognize that while this XNet exhibit demonstrates that the commercially available Obsidian Longbow XR readily supports high performance InfiniBand routing, a complete end-to-end solution depends on the process of InfiniBand specification completion and subsequent software stack support.*

## Applications

Longbow XR's router mode further supports its primary applications – remote InfiniBand visualization, storage, clustering, bulk data transport and latency-sensitive messaging – through superior network scalability, performance and administrative control.



[www.openfabrics.org](http://www.openfabrics.org)



[sc06.supercomputing.org/  
conference/scinet\\_xnet.php](http://sc06.supercomputing.org/conference/scinet_xnet.php)



[www.obsidianresearch.com](http://www.obsidianresearch.com)



[www.mellanox.com](http://www.mellanox.com)