Six-port 4X QDR InfiniBand Router

Crossbow R400-6 by Obsidian

Operation

Crossbow R400-6 is a six-port InfiniBand router, supporting native routing between multiple InfiniBand subnets at full line rates and very low port-to-port latencies.

Single subnets comprising just switches and HCAs have a single subnet manager and thus a single LID space administrator, a simple topology and limited mechanisms for fault or control isolation. Crossbow enables a single InfiniBand fabric to span many subnets, providing many operational advantages:

1. The ability to scale beyond a single LID space – a concern for very large InfiniBand installations.
2. The potential to support many different subnet topologies (fat-tree, toroid, mesh) in a single fabric without resorting to dropping out of InfiniBand to provide connectivity between them.
3. Faster subnet manager operation (since the work is distributed across multiple subnet managers) – again facilitating scaling.
4. Controlled inter-organizational InfiniBand connectivity without sharing management sovereignty or dropping out of InfiniBand fabric (particularly relevant when combined with Longbow wide area InfiniBand connectivity).
5. Fault isolation – very large systems need not be dependent on a single subnet's connectivity.

Each of R400-6's six InfiniBand ports can belong to different subnets; the R400-6 providing InfiniBand specification-compliant GID-based routing between them.

Sample Applications

Segmenting very large HPC fabrics

Very large clusters may use multiple Crossbow R400-6s for very high bisection bandwidth linkages between neighboring LAN subnets, allowing for parallel subnet manager operation. Four R400-6 devices fill a 1RU shelf, providing a very high 48Gbytes/second/RU throughput density, for 120W and sub 250ns latency penalties.

Dividing a compute cluster across multiple subnets overcomes the LID-space limitation imposed by the InfiniBand standard, allowing much larger systems to be deployed without the need to resort to specification changes that would break backward compatibility with existing hardware and software.

Global shared scratch storage

It is common for supercomputer clusters to use the same InfiniBand fabric used for IPC as a fast SAN fabric to accelerate check pointing or large-scale dataset I/O. In a site or campus with multiple supercomputers, each machine is typically associated with its own high-speed storage fabric, and so sharing data sets between the machines is problematic and cannot be done within a single InfiniBand subnet. Existing subnet managers are not able to cope with compound topologies, comprising multiple domains routed for dissimilar inter-switch connection topologies.

Using Crossbows to span them, multiple separate HPC installations (perhaps with very different internal subnet topologies) may share one or more InfiniBand SAN fabrics (perhaps using Longbow range-extenders to cover distance within a site) as high-performance global scratch. Data may now be freely shared between machines, and storage may be physically consolidated into a single building – perhaps with different security or environmental operating conditions.

Inter-site InfiniBand connectivity

Obsidian's companion Longbow C400 devices provide 40Gbits/s connectivity across as much as 80km of fiber, seamlessly connecting remote locations with a single InfiniBand subnet, since they operate as two-port InfiniBand switches. It may be undesirable for each end of such a connection to share administrative subnet control, due to performance, stability and especially political considerations.

Combining Longbow C400s with Crossbow R400s provides demarcation points between multiple subnets hosted at each geographical site. Obsidian's unique ¼ RU form-factor enables a QDR router, two QDR range-extenders and a CWDM module (to map all 80Gbits/s of InfiniBand WAN traffic onto a single fiber pair) to occupy a single RU space and a modest sub 100W power footprint.

Multiple sites thus enjoy full InfiniBand connectivity – including efficient data RDMA transfers – without compromising site administration sovereignty, with the security risks that may impose.

Subnet Manager Support - BGFC

Crossbows enable multi-subnet InfiniBand fabrics, however existing open source and commercial subnet managers do not support multiple subnets.

Obsidian has developed a next-generation subnet manager, BGFC, which is designed to support multiple subnets out of the box. BGFC is a complete re-imagining and re-implementation, not a derivative from existing software.

In addition to multi-subnet capability, BGFC also brings lockless clustered implementation, graph-theory-based guaranteed deadlock free forwarding table calculation, topology persistence through write-once databases and a very powerful Python based topology description environment.

BGFC is required in order to deploy Crossbow technology.

Specifications

Chassis

Mounting (Desktop)	Fitted with rubber feet
Mounting (Quad Pack)	19" rack-mount shelf, front mounting no-rail system
Physical	1.5"x4.25"x13"
System power	Max 45W / Typical 30W
Input power	80-264 VAC at 47-63 Hz (universal)
Environmental	10-45 degrees C (32-113 degrees F) ambient
Airflow	Pressurised, rear to front flow, filtered intake option
External Ports	1 AC input, 1 USB (RS232), GbE (management), 6 QSFP+ IB QDR 4X
Acoustics	Intelligent fan speed control for quiet operation and long bearing life (dual internal redundant)

Management

Ethernet	Full duplex 10/100/1000 Base-T Ethernet with auto MDI-X
Protocol Support	IPv4, IPv6, HTTP/HTTPS, SNTP, DNS, ZeroConf and DHCP
GUI	Web based interactive management
User Management	Single user account
HTTP	SSL v2/v3 and TLS with HTTP digest challenge/response password exchange
Configuration	Through Web GUI or via a text configuration file
Firmware	Web upgradeable. Primary/Secondary high-availability FLASH storage with scrubbing.

InfiniBand Interface

Connector	IB QSFP+ 4X, QDR at 32Gbits/second (payload)
Node Type	Six-port switch or router
Physical Layer	InfiniBand Architecture v1.2.1
Subnet Manager Agent	Integrated InfiniBand Architecture v1.2.1
GID routing entries	Up to 65536 (Unicast), Up to 32768 (Multicast)
Port-to-port latency	<250 ns (small packet store-and-forward)
Packet routing rate	Up to 80 million packets per second per port
VLs	4 data, 1 management

Crossbow R400 Series Ordering Code

R400-6	Six-port 4X QDR InfiniBand router

Contacts

For sales or channel partner inquiries:

sales@obsidianstrategics.com