



**Storage Virtualization**

- iSCSI & Fibre Channel
- Fail-Safe Data Protection
- Go Green Thin Provisioning



**Go Virtual!  
Click here.**

## Interviews

### Mellanox on InfiniBand and FCoE

posted on 10 May 2008 18:08

#### InfiniBand leads the way

B&F talked with Thad Omura, VP for product marketing at Mellanox and asked him a set of questions about InfiniBand, Fibre Channel over Ethernet (FCoE) and datacenter Ethernet. Here's what he said:-

#### Blocks and Files: How would Mellanox position InfiniBand and Ethernet? Will they co-exist and, if so, how?



**Thad Omura:** InfiniBand (IB) will always be ahead of Ethernet in terms of performance, price performance, latency and features. Ethernet is familiar to data center users and is evolving, albeit slowly, to try and do the things InfiniBand has been doing for years (fabric reliability to carry storage traffic, congestion management, etc.) IB will co-exist with Ethernet, provide the leading (in terms of performance, price, power, etc.) connectivity for servers and storage as a convergence fabric, and gateway out to Ethernet LAN services. There will also be data centers that will deploy only Ethernet where capabilities, features and performance are not critical and desire to leverage FCoE and other convergence technologies as they evolve.

Mellanox's Virtual Protocol Interconnect (VPI) recently launched in our ConnectX server and storage adapters provides the ability for any networking, clustering, storage, and management protocol to seamlessly operate over any converged network (InfiniBand, Ethernet, Data Center Ethernet) with a consolidated software stack and the ability to auto-sense what network is connected. With this capability, we are well positioned to support any convergence fabric with the same adapter device.

#### Blocks and Files: What is Mellanox' view of Data Center-class (loss-less) Ethernet? How will this compare and contrast to InfiniBand? (Blade Network Technologies has just [announced](#) its loss-less RackSwitch)



**Thad Omura:** The goal of all of these efforts in the industry is to bring the capabilities that have existed for years on InfiniBand to Ethernet fabrics. We are involved in the standardization committees within IEEE (Data Center Bridging Group (DCB)) and with our history with InfiniBand, will take a leadership position on these features as well. There are three elements to loss-less Ethernet - Class-based flow control (CBFC) or per-priority pause (PPP), hardware-based congestion management, and how to assign different types of network traffic to different classes of service. Most vendors today have implemented PPP or CBFC and are touting loss-less Ethernet.

We believe it will take several years for the multiple initiatives to become standard, and in between, vendors will try to deploy pre-standard/proprietary solutions, or partial solutions (e.g., PPP/CBFC only). In the end, IT managers prefer standards-based infrastructure and believe there will not be mass deployment until this is established. At the same time, InfiniBand will continue to evolve, add capabilities and performance and continue to be ahead of Ethernet.

#### Blocks and Files: How will InfiniBand withstand the inexorable onrush of what Cisco calls Ethernet economics?

**Thad Omura:** InfiniBand offers a significant price and performance advantage over Cisco's Ethernet solutions. In fact, Cisco continues to demand higher GigE, let alone 10GigE prices vs.

InfiniBand. In addition, there is a continuous battle between server vendors and Cisco - Cisco is driving to tie down deployments of the server vendors to their networking solutions (pre-standard, management, etc.) and server vendors prefer to have an alternative that is open. This is where InfiniBand shines as a viable alternative convergence fabric that is here today.

The fact of the matter is multi-core CPUs are scaling so fast in performance that InfiniBand is really the only I/O solution that is scaling fast enough to get the most out of these very fast computing and storage platforms. So when looking from a total system platform perspective, InfiniBand will continue to maximize ROI as there is always a need to execute faster and more efficiently.

#### Blocks and Files: What is Mellanox' view of FCoE? Will FC SAN users be likely to move to a unified server/storage interconnected for SAN (iSCSI, FC) and NAS based on Ethernet, which they use for LAN traffic, rather than adding

### **InfiniBand to their existing Ethernet/FC mix?**

**Thad Omura:** Mellanox sees FCoE as a convergence fabric for those data centers looking to remain Ethernet-based and continue to leverage their investments in FC storage equipment. FCoE's value is to consolidate a FC and Ethernet adapter into a single adapter on the server side and remove the FC switch that enables servers to connect to the SAN (saving power, cost, cabling and complexity). Initial deployments will happen through gateways to FC storage and over time, FC storage will migrate to FCoE adapters too.

If the applications demand the highest throughput to storage, we see data centers will go to InfiniBand for storage connectivity. If storage I/O convergence on Ethernet is the highest requirement (as opposed to best price/performance, scalability, etc.) then FCoE will eventually become a viable option once standard products are set in the market (multiple years).

With VPI, Mellanox provides the leading InfiniBand, Ethernet, Data Center Ethernet, FCoIB and FCoE adapter solution - so no matter what the IT manager is looking to deploy, we'll support it with the same device. This future proofs data center IT infrastructure.

### **Blocks and Files: The Mellanox white paper claims that InfiniBand is set to become the interconnect standard for SANs and mentions wide-scale deployments are coming. Some people might find this quite hard to believe. What evidence is there to justify this claim please?**

**Thad Omura:** The adoption of InfiniBand continues to increase on both the server and storage side. We are seeing an increasing number of storage vendor provide native IB connectivity, and several tier-1 vendors support solutions through gateways and believe this will lead to native IB products as well. Certainly, in the HPC market, InfiniBand storage is widely deployed as a storage interconnect in the form of clustered or scalable file systems. As trends in HPC spill over into enterprise deployments, we believe the same value of InfiniBand for clustering servers is happening for storage clustering.

### **Blocks and Files: Is there an InfiniBand roadmap? If so what is it? When will an 80Gbit/s InfiniBand arrive?**

**Thad Omura:** Absolutely. 120Gb/s (12X, 10Gb/s per lane) is already defined and Mellanox announced this will be supported on switch to switch links on the next switch silicon device coming to market (InfiniScale IV) with systems expected the second half of this year. The InfiniBand Trade Association is working on performance roadmaps that scale into the 100s of Gigabits per second. Again, this is to address the I/O demands of multi-core CPUs, systems running a large number of virtual machines that want to share the same I/O connection for power and cost concerns, and the need to access massive data files and storage.

### **Blocks and Files: Does Mellanox think that data center SAN and virtual server users want to virtualise network connections? Is InfiniBand a better basis for doing this than Ethernet?**

**Thad Omura:** Yes, there is a big move to virtualize network connections and have multiple virtual machines in a single server utilize the same interface for the I/O services it requirements. What makes InfiniBand ideal for this is the mechanisms defined to allow virtual machines to set up direct I/O connections with the IB adapter, effectively offloading the hypervisor for this compute intensive task. Even more, for each of the different I/O connections that are established between VM and the adapter, you can assign QoS, bandwidth restrictions, etc. depending on the type of I/O is required. You can then map these I/O channels into the virtual lanes supported by the InfiniBand fabric and maintain end-to-end delivery of I/O services.

The I/O channels available in Mellanox ConnectX adapter products enable deployment of multiple virtual NICs and virtual HBAs over the same 20 or 40 Gb/s InfiniBand wire. Today, there are multiple deployments of virtualized I/O over InfiniBand using VMware ESX server where up to 32 GigE NICs and 32 FC HBAs can be dynamically (on-the-fly) enabled over a single InfiniBand adapter, while maintaining seamless end-to-end Ethernet and FC experience.

Mellanox's VPI architecture allows the ability to bring these I/O services on both InfiniBand (deployable today and with higher performance and capabilities), Ethernet and Data Center Ethernet. Ethernet and Data Center Ethernet currently lack the fabric reliability and virtual lane capabilities of InfiniBand in addition to other features so you wouldn't have end to end I/O services, so it is yet to be determined if the DCB efforts will truly enable Ethernet to provide what IB can do today.

### **Blocks and Files: How does InfiniBand cope across longer distances than inter-rack data center links?**

**Thad Omura:** InfiniBand has both copper and fiber solutions that span all of the different ranges of interconnect required. For very short rack distances, there are passive copper cables. For longer data center connections, there are active fiber cables that are cost-effective. 10GigE and Data Center Ethernet share copper and fiber solutions used by InfiniBand and therefore are subject to the same distance capabilities. High speed copper and fiber solutions are more proven over InfiniBand that has enjoyed millions of ports of deployments in the field (nearly 10 fold than 10GigE ports). For campus connectivity and WAN, there are several vendors providing solutions including Obsidian, Bay Microsystems and NET. So, for any distance connection, there are solutions in the market today being deployed for InfiniBand.

[Chris Mellor, editor.]

---

tags: [InfiniBand](#) [FCoE](#)

---

