Fabric convergence with lossless Ethernet and Fibre Channel over Ethernet (FCoE)

Technology Bulletin

BLADE NETWORK TECHNOLOGIES
The Trusted Name in Blade Server Networking

Introduction

Fibre Channel (FC) is currently the defacto block storage protocol standard for implementing Storage Area Networks (SANs) in enterprise data centers. In 2007 approximately 90% of storage devices in the data center were connected via Fibre Channel.

SANs deliver the following well-understood value propositions:

- Storage consolidation
- Centralized storage management
- High performance
- High reliability
- Rapid reconfiguration

Despite its popularity in SANs, Fibre Channel has not been accepted as a volume networking or clustering fabric. This puts Ethernet at a significant volume and therefore cost advantage to FC.

The recent introduction of 8Gb/s FC marks the 4th generation of this long standing incumbent SAN technology. However the rate of development of the Fibre Channel road map has been slower than Ethernet which has led FC to suffer a performance disadvantage to Ethernet. 10Gigabit Ethernet has been available for several years and Ethernet will move up to 40Gb/s and then 100Gb/s over the next few years.

Current Alternatives

iSCSI and SAS are block storage protocols that have both seen moderate success at meeting the demanding needs of Enterprise customers.

iSCSI is a block storage protocol that transports SCSI over TCP/IP allowing SANs to be created using Ethernet switches. iSCSI has seen success in SMB (small medium businesses), using low cost Gigabit Ethernet adapters and software iSCSI initiators. iSCSI SANs are also being created using 10Gigabit Ethernet adapters with TCP Offload Engines (TOE) to meet the high performance requirements of the enterprise.

Serial Attached SCSI (SAS) is a cost focused technology designed to replace parallel SCSI. While an excellent technology for internal and direct attached storage it is limited in performance, scalability and distance compared to Fibre Channel and has seen limited acceptance as an external storage fabric. Also like Fibre Channel SAS is only used for storage connectivity.

Network Attached Storage (NAS) appliances using NFS or CIFS protocols are a very popular file access storage solution over Ethernet networks.

Introducing FCoE

Fibre Channel over Ethernet (FCoE) transports Fibre Channel frames over an Ethernet fabric and enables data centers to increase application performance while reducing cost, power and manageability tasks by converging Storage, Networking and clustering data to a single fabric.

The key benefits of FCoE are:

- High performance storage access over lossless 10Gigabit Ethernet fabrics
- Lower capital, energy and cooling costs with less adapters and switches
- Lower cost and size of servers by reducing the number of required I/O slots
- Lower cabling costs by unifying to a single fabric type with less cables
- Lower management overhead by maintaining a single fabric
- Increase application availability by simplifying the network and server configuration
    - Easier to install
    - Easier to diagnose
    - Using common parts = faster time to fix

FCoE is an evolution of Fibre Channel that uses Fibre Channel's Network, Service and Protocol layers to carry data packets over Ethernet physical and data link layers.

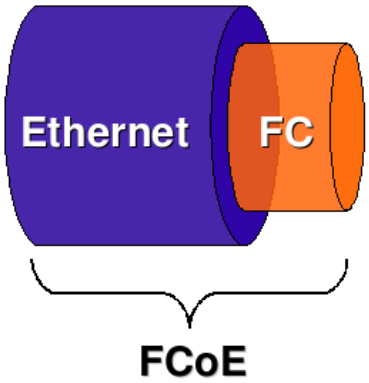

Figure 1. Conceptual view of Fibre Channel over Ethernet

Using Fibre Channel's upper layers smoothes the transition to FCoE by allowing the leverage of enterprise-proven Fibre Channel software stacks, management tools and trained administrators. Most importantly, mission critical applications do not need to change to benefit from the performance and cost benefits of FCoE.

Existing management infrastructure such as zoning policies and name server databases can be simply reused and extended. There are also a very wide range of tools and expertise to manage Ethernet networks.

FCoE enables a convergence to a single Ethernet fabric for Storage, Networking and Clustering which will provide significant cost savings in initial hardware purchase, power and manageability.

The FCoE protocol is being developed in the T11.3 committee of the INCITS standard body. There is broad industry agreement on the standard with key players including Blade Network Technologies, Brocade, Broadcom, Cisco, Emulex, HP, IBM, Intel, LSI, NetApp, SUN and Qlogic.

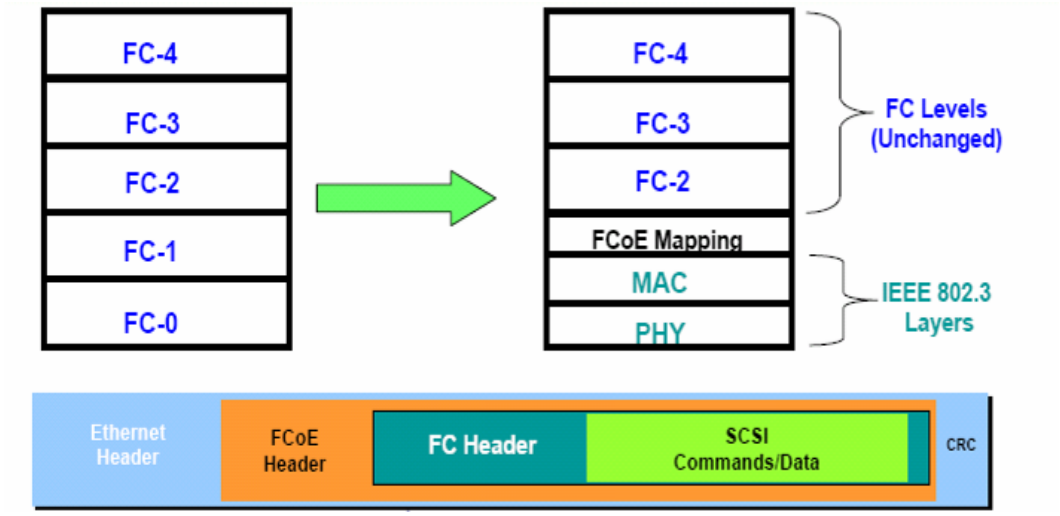

Figure 2. Fibre Channel to Ethernet Mapping

The Fibre Channel frame along with the headers and CRC fields are encapsulated in an Ethernet frame. Fibre Channel ID's are mapped to Ethernet MAC addresses which are used to route data across the switched Ethernet fabric from source to destination. By maintaining the FC header the FC frame can be easily de-encapsulated and transmitted across an FC SAN to an existing Fibre Channel device.

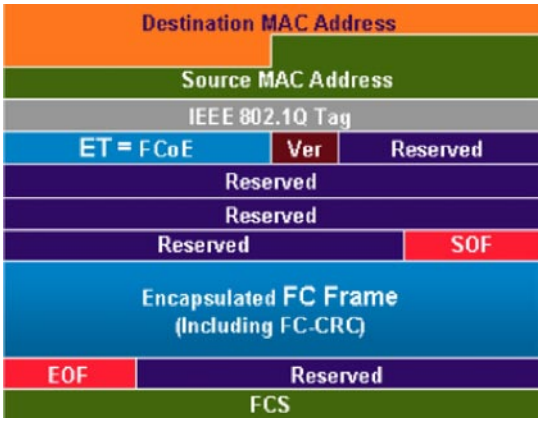

Figure 3. FCoE Frame

At 2122 bytes the typical size of FC frames is bigger than the basic Ethernet 1500 byte MTU. Using Ethernet switches that support jumbo or mini-jumbo frames (2500 bytes) will avoid fragmenting the FC frame and deliver the best performance.

Lossless Ethernet

Fibre Channel protocols assume that the underlying fabric is lossless even during network congestion, to ensure that storage traffic is delivered reliably and in a timely manner. The Fibre Channel transport protocol uses link level credit based flow control that guarantees no loss of frames under normal conditions. Ethernet by contrast is a best-effort network, and may drop packets when the network is busy, resulting in retransmissions, storage protocol time-outs, and even lost data. While FCoE SANs can be built with current Ethernet fabrics, providing the mission critical reliability and data integrity required by enterprise data centers requires enhancements to Ethernet to create fully lossless fabrics.

New standards are under development that will create a new, more capable family of Ethernet protocols. These standards, referred to collectively as Converged Enhanced Ethernet (CEE) are being developed in IEEE 802.1, 802.3, and IETF standards bodies. However, the critical technologies required to enable FCoE are available now, based on a subset of the standards that have been agreed to by the standards participants. Thus FCoE can be deployed now, with lossless characteristics.

Convergence Enhanced Ethernet (CEE) capable products will enable lossless Ethernet fabrics by using IEEE 802.1p Class Based Flow Control (CBFC), to pause traffic base on the priority levels. 802.1p allows virtual lanes to be created within an Ethernet link with each virtual lanes assigned a priority level. During periods of heavy congestion lower priority traffic can be paused, while allowing high priority and latency sensitive tasks such as data storage to continue.

IEEE 802.1Q (Virtual LAN) can be use to partition the physical Ethernet fabric to create high levels of security by isolating traffic types and to enhance Quality of Service by configuring guaranteed bandwidth and latencies per VLAN. Using VLANs and 802.1p flow control several lanes of high performance lossless FCoE can be established on a single 10 Gigabit Ethernet fabric.

Ethernet will be further enhanced by the proposed IEEE 802.1au Congestion Notification (CN), which will provide end to end flow control capabilities by allowing congestion points to notify a rate limiter when congestion is occurring, so that the traffic can be throttled back.

Comparing FCoE

iSCSI uses TCP (Transport Control Protocol), to provide the reliable in-order delivery of data over lossy/unreliable fabrics. This requires some overhead in host CPUs to process the TCP/IP software stack or dedicated TCP Offload Engines (TOE) on network adapters. FCoE does not use TCP/IP and imposes a lower CPU overhead, but does require a lossless Ethernet fabric to provide reliable and high performance storage connectivity.

	iSCSI	FCIP	iFCP	FCP	FCoE
TCP required	Yes	Yes	Yes	No	No
Loss-Less fabric required	No	No	No	Yes	Yes
CPU overhead	High	High	High	Low	Low
FC gateway overhead	High	High	High	N/A	Low
FC security model	No	Yes	Yes	Yes	Yes
FC management model	No	No	Yes	Yes	Yes
FC software stack	No	No	Yes	Yes	Yes
IP routable	Yes	Yes	Yes	No	No

Figure 4. Comparing FCoE

Like Fibre Channel, FCoE is not directly routable across IP networks, therefore for applications where data needs to transmitted across large distances to remote sites or over lossy networks, technologies such as iSCSI, FCIP (FC over IP) or iFCP (Internet FC protocol) are better choices.

FCoE maps directly into Ethernet frames allowing low cost gateways to simply strip off the Ethernet and FCoE headers and pass the FC packets on to existing FC SANs or storage devices. This protects customer investments in existing FC switches and storage devices while delivering high performance and enabling a simple transition from FC to FCoE SANs.

Performance comparison

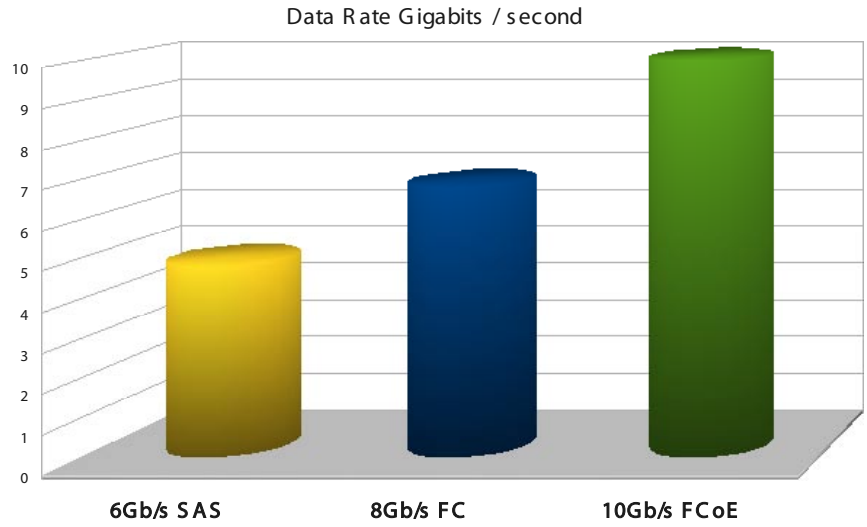

Figure 5. Date Rates of SAS, Fibre Channel and FCoE

The latest generations of SAS and Fibre Channel (6Gb/s SAS and 8.5Gb/s FC), both use 8b/10b encoding and send 10 bits for every 8 bits of data. The true theoretical maximum data rates are appropriately 4.8 Gb/s for SAS and 6.8Gb/s for FC. 10 Gigabit Ethernet uses a 64b/66b encoding scheme and has a maximum data rate of 10 Gigabits/s therefore FCoE can provide significantly higher data rate performance.

The road map for Ethernet shows an evolution to 40Gb/s and to 100Gb/s over the next few years and the standards bodies are well on the way to final specification for both speeds. The rate of performance increase in Ethernet is significantly faster than SAS or native Fibre Channel and will widen the performance advantage for FCoE.

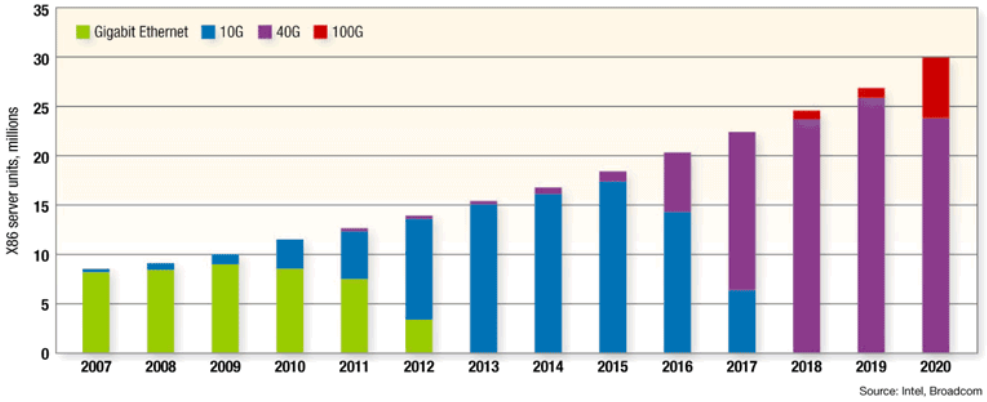

Figure 6. Proposed industry road map for Ethernet

Fabric Convergence

In today's data center servers are often configured with at least 4 adapters, typical 2 or more Ethernet NICs plus a pair of dedicated Fibre Channel HBAs. As the number of CPU cores per server and usage of Virtual Machines increase the demand for more I/O, the number of adapters per server is growing quickly.

Loss-less Ethernet and FCoE makes it possible to converge enterprise storage, networking, management and clustering data onto a single fabric that is simple to manage, high performance and cost effective fabric.

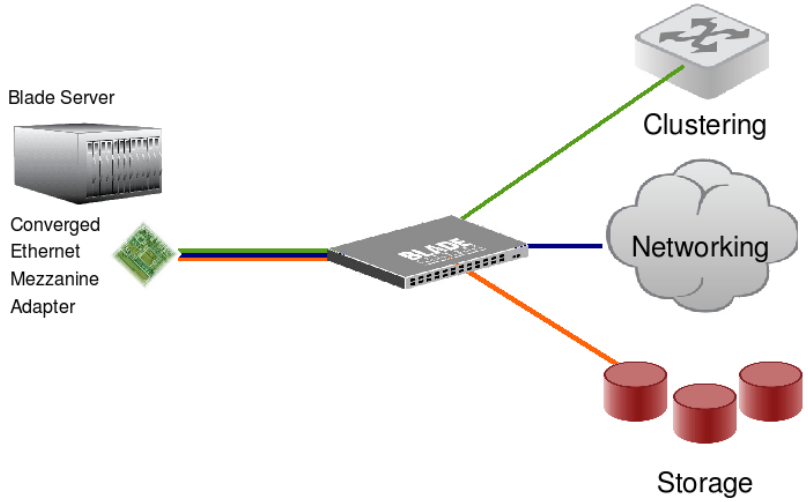

Figure 7. Converging Storage and Networking to a single fabric

Using a single converged fabric reduces equipment and management costs by eliminating the need to use multiple adapters for specific data types such as FC for storage or InfiniBand for clustering. Reducing the number of adapters required allows smaller and low cost servers to be used for additional savings.

Converging to a single Ethernet fabric for Blade servers enhances the key benefits such as lower cost, better power / cooling efficient and simpler management that make Blade servers the fastest growing segment in the market today.

Cost saving

Significant cost saving can be expected by converging storage and networking onto Ethernet fabrics. Not only are energy, management and maintenance costs reduced, but the initial cost of Ethernet adapters, cables and switches is lower than Fibre Channel due to the higher volumes and ultra competitive market forces.

While 10Gigabit Ethernet and FCoE enabled adapters will provide the highest performance FCoE connectivity, the FCoE protocol can be run on existing Ethernet ports allowing customers to connect existing systems to an FCoE SAN with no additional cost. The open source project Open-FC.org is creating FCoE driver and software solutions for the Linux operating system

Power Saving

Converging to a single lossless Ethernet fabric not only reduces equipment and management cost, it also significantly reduces the power consumption across the data center by reducing the number of adapters and switches required.

Based on studies by thegreengrid.org the total power used in data centers is >3X the power used directly by the IT equipment. Therefore every Watt saved by converging to an Ethernet fabric actually saves 3 Watts. In large data centers the cost saving and reduced impact on global warming can be significant.

See "Guidelines for Energy Efficient Data Centers" for more details available from www.thegreengrid.org

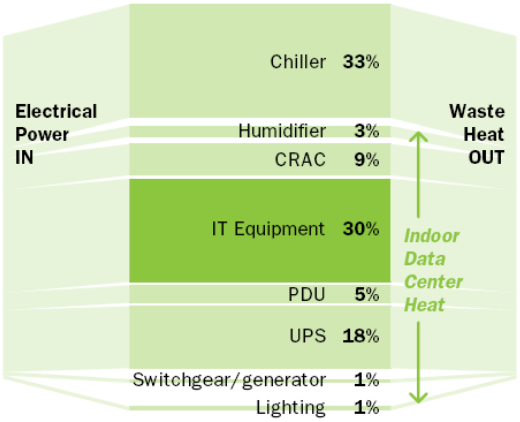

Figure 8. Data Center energy usage

Rolling FCoE out in the data center

The 1st FCoE products will be gateways that connect FCoE capable initiators (server side) to existing Fibre Channel SANs and enable customers to consolidate server interconnectivity to a single Ethernet fabric. These are expected to be available before the end of 2008.

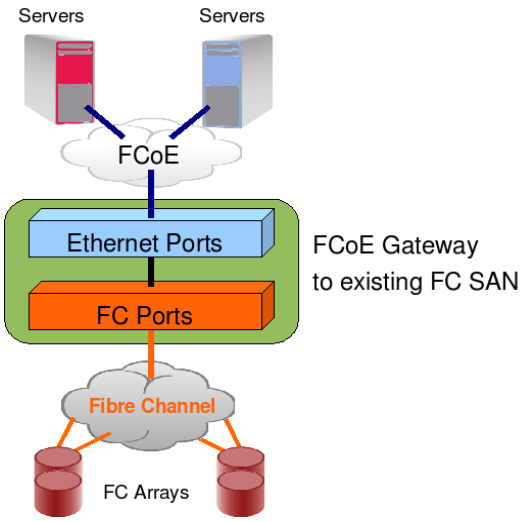

Figure 9. FCoE to FC Gateway

To the other nodes in the FC SAN the FCoE initiators appear to be directly connected and can be managed and maintained with the same tools.

As FCoE becomes more popular native FCoE devices will appear on the market. These will provide the additional benefit of connecting directly to cost effective Ethernet switches and bypassing the albeit low overhead of the FCoE to FC gateways for even better performance.

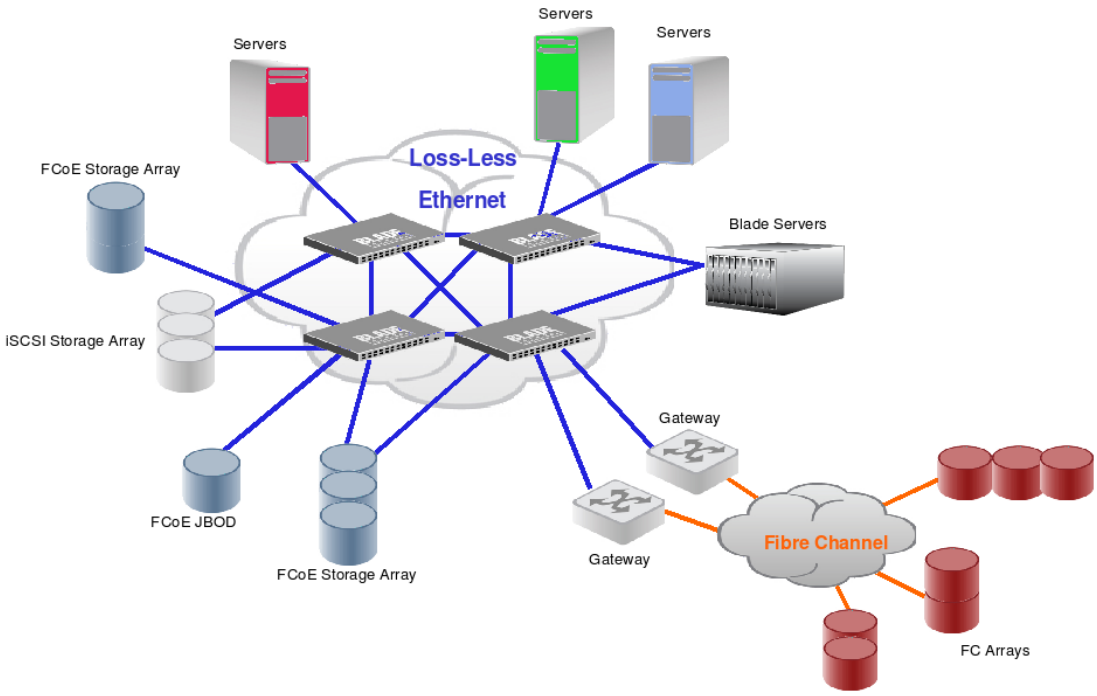

Figure 10. FCoE SAN

As shown in the diagram above not only can native Fibre Channel devices be connected via gateways, but iSCSI devices can also be seamless attached to the same lossless Ethernet fabric.

FCoE in SMB (Small Medium Businesses)

The price / performance values that attracted SMB customers to iSCSI very much applies to FCoE. Without the overhead of TCP/IP, FCoE will be able to offer even higher value to this segment as lossless Ethernet switches become available.

Summary

Fibre Channel over Ethernet (FCoE) is a rapidly evolving standard that transports storage data over reliable Ethernet networks to enable the convergence of storage, networking and clustering onto a cost effective fabric.

FCoE has wide support from Ethernet and Fibre Channel industry leaders and enables the following key benefits:

- Increased performance
- Greater data center density
- Lower capital cost
- Lower energy and cooling costs for a lower CO^2 footprint
- Lower manageability overheads
- Higher application availability with simplified server and network configurations