TANEJA GROUP
TECHNOLOGY ANALYSTS

T E C H N O L O G Y  B R I E F

# Fibre Channel Over Ethernet: A Unified Fabric in Sight

October 2007

THE NET NET

Fibre Channel is the heartbeat of storage in the enterprise data center, but ubiquitous Ethernet rules supreme nearly everywhere. Convergence proponents point to the advantages of a single Ethernet-based networking protocol: economy of scale, wide IT expertise, highly simplified networks. However, the facts remain that a) corporations have invested hugely in Fibre Channel and are not likely to dump that investment any time soon, b) Fibre Channel beats iSCSI (Ethernet) as the storage-friendly network *for critical applications*, and c) Fibre Channel traffic can traverse IP-based networks with the FCIP protocol, but Ethernet's packet-dropping behaviors make it unsuitable to transport Fibre Channel frames.

Does this render Ethernet and Fibre Channel convergence impossible? To the contrary, Taneja Group believes that convergence is not only possible, but probable – just not with today's available technologies and specifications. *The solution to true convergence is the ability to run Fibre Channel over lossless Layer 2 Ethernet.* That day is not here yet, but we believe that it is coming in the form of Fibre Channel over Ethernet (FCoE) and Converged Enhanced Ethernet (CEE).

Industry heavyweights with strong Fibre Channel investments are at the heart of the movement. These include Emulex, Cisco, IBM, Intel, Sun, EMC, Brocade, Nuova, and QLogic; and other companies like Woven Systems that are actively introducing related products into the market. Standards work is happening first at industry groups like ANSI and IEEE, with products expected in 2008. Emulex is one of the companies pushing hard for FCoE and CEE standards and technical development, understandably given their stated goal of pervasive server connectivity. Since FCoE will allow a direct mapping of Fibre Channel over Ethernet, it ensures pervasive server connectivity by enabling Fibre Channel to run over Ethernet with no (or minimal) changes to drivers, software stacks, and management tools.

## Challenges to a Converged Network

Consolidation is the name of the game in the enterprise data center. Servers, storage and related components are commonly consolidated, but thus far networks have resisted the siren call. It's not that the industry doesn't want it: given the ubiquitous nature of Ethernet, a converged Ethernet fabric has been a goal in the storage industry for some time. Such a fabric would be able to leverage existing Ethernet and Ethernet training and knowledge, consolidating storage network fabrics with IP-based traffic coming from blade and virtualized server farms. Instead, Ethernet and Fibre Channel co-exist uneasily in the data center. Fibre

Copyright© The TANEJA Group, Inc. 2007. All Rights Reserved
87 Elm Street, Suite 900 Hopkinton, MA 01748 Tel: 508-435-5040 Fax: 508-435-1530 www.tanejagroup.com

Channel administrators do not trust traffic to and from lossy Ethernet, even though the latter carries the majority of networking traffic.

It is not for lack of trying. The ability to consolidate fabrics sounds outstanding in theory: cost savings, leveraging Ethernet networks, and administrative simplicity to begin with. A few years ago the idea was to replace Fibre Channel with Ethernet-based block storage traffic by combining Gigabit Ethernet for speed, iSCSI for running block-level data over IP and Ethernet, and iWARP, a 10Gb Ethernet incarnation of Remote Direct Memory Access (RDMA) for high performance clustering. These protocols leveraged Ethernet and IP, and the thinking went that the enterprise would swiftly adopt them in order to use Ethernet as a unified data center fabric. In fact, when iSCSI hit the corporate scene in earnest a few years back, some industry writers proclaimed that because Ethernet training and expertise were so common, iSCSI would quickly replace Fibre Channel as the storage network protocol of choice.

They were wrong. iSCSI has become popular in the small enterprise, workgroups and remote offices but is a non-starter in the large enterprise data center. Large enterprise is heavily invested in Fibre Channel SANs and the last thing most data center managers want is to replace their expensive Fibre Channel SANs with iSCSI storage and all of its attendant management tools and supporting products. Nor are most of them willing to trust the mission critical applications that now run on Fibre Channel to IP-based networks.

## BRIDGING THE GAP

There are existing means of transporting Fibre Channel frames over distances using IP-based networks. A prime example is Fibre Channel over IP (FCIP), a point to point wide area connection that tunnels between two Fibre Channel switches over an extended IP link. The originating Fibre Channel switch delivers Fibre Channel frames to the FCIP device, which wraps them in IP datagrams and shoots them across the IP network to the receiving tunnel gateway. There the second FCIP device strips the IP information from the frames and delivers them to the receiving Fibre Channel switch for routing. Internet Fibre Channel Protocol (iFCP) also carries Fibre Channel frames over IP networks, but unlike FCIP does not depend on switch to switch communications. This means that IT can build IP-based SANs using Fibre Channel storage devices but without requiring Fibre Channel fabric.

But if FCIP and iFCP are already available, what is the problem? Here it is in a nutshell: FCIP is acceptable for point-to-point data transmissions between Fibre Channel switches, and iFCP is adequate for IP SANs storing persistent data. *But IP-based Ethernet cannot reliably carry Fibre Channel traffic at the Quality of Service levels the production environment requires.* Why? Because whether you are using FCIP or iFCP, you must still depend on IP wrapping the Fibre Channel frames to transport them over Ethernet. And IP-based Ethernet is a lossy network that is unsuitable for transporting critical data.

Copyright© The TANEJA Group, Inc. 2007. All Rights Reserved

In a traditional LAN environment, IP-based Ethernet is subject to variable latencies called latency jitter. Left alone, severe latency jitter can even down applications. IP networks can use TCP/IP's recovery and flow control capabilities to control latency effects, but TCP/IP overhead is an unacceptable burden for Fibre Channel traffic.

Given these serious challenges, a converged fabric based on traditional Ethernet lacks a compelling value proposition. But that is changing with development around related new standards: a new specification called Fibre Channel over Ethernet (FCoE) and an advanced fabric called Converged Enhanced Ethernet (CEE).

## Meeting the Challenge with FCoE and CEE

The proposed Fibre Channel over Ethernet (FCoE) standard would enable Fibre Channel traffic to run over a single enhanced Ethernet LAN segment in the data center, and would support SAN management domains by maintaining logical Fibre Channel SANs across Ethernet. The goal of the proposed standard is to enable Fibre Channel to run over Ethernet with no performance degradation and without requiring any changes to the Fibre Channel frame.

FCoE does not stand by itself, but also requires the presence of Converged Enhanced Ethernet (CEE), an advanced network fabric built on 10Gb Ethernet (10GbE). This combination avoids using TCP/IP protocols to transport Fibre Channel frames over Ethernet. TCP/IP adds latency and overhead to frames; not a difficult issue for traffic created for IP networks, but a serious problem for Fibre Channel frames in high performance environments. Running Fibre Channel over layer 2 Ethernet would avoid TCP/IP overhead, and CEE would provide the lossless environment that the Fibre Channel SAN requires. This approach – adding features into layer 2 in order to dispense with TCP/IP baggage -- makes it a highly suitable fabric for storage and the FCoE transport mechanism.

A critical piece of bringing CEE to reality is Congestion Management (CM). The IEEE is developing CM standards to addresses latency and packet loss on Ethernet networks. Along with 10GbE, this will make CEE suitable as a Layer 2 network in short distance, high-speed networks like blade chassis backplanes and data centers.

Enhanced Ethernet and FCoE would map Fibre Channel directly into the Ethernet payload. The ability to do this is fundamental to a truly converged network environment based on Ethernet, since the technique maintains existing Fibre Channel environment, discovery and management tools. Success will enable a converged Ethernet dedicated to serving high speed, short range networks such as blade server backplanes and virtualized data centers.

## Converged Networks and the Standards Bodies

As of August 2007 a projected product timeline of 2009 moved up, largely thanks to the unexpectedly timely agreement between Cisco and Brocade over conflicting standards. The two companies hammered

Copyright© The TANEJA Group, Inc. 2007. All Rights Reserved

out a compromise proposal and submitted it to ANSI T11 workgroup, which promptly accepted it. The lack of an agreement was a major roadblock for a FCoE standard, and without a compromise between the rival networking companies it was impractical to move forward on the chip development that is the next step towards developing FCoE.

**SUMMARY OF KEY ELEMENTS IN CONVERGED ETHERNET**

- No changes to Fibre Channel frame or to Fibre Channel SAN services
- Advanced Ethernet: no dropped frames, robust congestion management, ability to prioritize Fibre Channel traffic over non-Fibre Channel traffic, frame latency comparable to Fibre Channel
- FCoE transparency to endpoints, including gateways and forwarders between Ethernet and Fibre Channel
- No or minimal changes to OS drivers, software stacks and management tools
- MAC addressing dispenses with IP overhead
- Native Fibre Channel and FCoE devices have seamless connectivity

## CEE: ADVANCING ETHERNET

FCoE members also support 10Gb Converged Enhanced Ethernet, and are actively engaged in standards development through IEEE. These standards are driven by 10GbE development, which is in turn driven by storage needs. Network-only 10GbE does not have the same requirements around data qualification and interoperability that storage does, and an IP-based network like traditional Ethernet suits network-only needs. However, allowing Fibre Channel and Ethernet storage traffic to run over a converged fabric requires significant work in order to make 10GbE Ethernet a lossless, native network layer.

Ethernet already has some ability to isolate traffic using Virtual Local Area Network (VLAN) protocol, and service differentiation with the Quality of Service (QoS) protocol. However, Ethernet is still prone to network congestion, latency, and frame dropping – an unacceptable state of affairs for a converged network solution. The Ethernet 'pause' frame was designed for flow control but is rarely implemented today because it pauses the entire traffic on the link.

IEEE and related bodies have stepped into the gap, developing standards for an enhanced Ethernet capable of converging Fibre Channel traffic. The four major standards for CEE include congestion management (CM), per priority PAUSE, CM discovery and capability exchange, and priority processing/packet scheduling.

- **Congestion management** operates at the link level to control unicast network traffic. IEEE 802.1au is actively pursuing this function.

- **Per priority PAUSE** works at the congestion point to support differentiated services and to minimize or eliminate packet drops. Standard Ethernet has a PAUSE command but cannot distinguish between multiple data flows occurring over the same path, so a PAUSE stops not only the congested flow but all others sharing the path. Per priority PAUSE enables the

Copyright© The TANEJA Group, Inc. 2007. All Rights Reserved

network to practice selective pausing by assigning different priority classes. Storage traffic will continue to flow through the path without being impeded by a PAUSE command, allowing storage traffic and data traffic to share converged links.

- **CM discovery and capability exchange protocol** defines interoperability between CM-compliant bridges and endpoints.

- **Priority processing and packet scheduling** observes priority processing for high priority traffic in endpoints and switches.

A related standard from an IETF workgroup is likely to include an Ethernet link layer routing protocol, which provides shortest-path frame routing in multi-hop IEEE 802.1-compliant Ethernet networks.

## ENHANCED ETHERNET AND FCOE

The ANSI T11 workgroup is shepherding the FCoE storage protocol to run over enhanced 10GbE Ethernet. FCoE encapsulates complete Fibre Channel frames onto Ethernet frames, observing a strict one-to-one encapsulation. This means that FCoE does not segment Fibre Channel frames or pack multiple Fibre Channel frames into a single frame. FCoE frames can encapsulate extended and optional FC headers as well as virtual fabrics.

To be successful, FCoE must keep to traditional Ethernet standards wherever possible. For example, autonomous fabric support is a necessity for true convergence. Large enterprises never adopted iSCSI into the data center partially because it meant replacing familiar Fibre Channel services. FCoE must support native Fibre Channel services like name server, domain controller and zone server.

Along these same lines, although FCoE requires advanced Ethernet like CEE, it does not require a dedicated fabric. FCoE must be able to operate over standard Ethernet switches carrying other Ethernet traffic, although the switches must be capable of operating in a lossless mode and supporting Fibre Channel's larger frame sizes. As FCoE and CEE specifications mature, CEE-enabled switches will become part of advanced Ethernet product development. These switches will carry all Ethernet traffic, not just FCoE.

Cost-effective gateways also figure into the equation. FCoE will require high bandwidth gateways interconnecting FCoE and Fibre Channel, and the better the ROI the faster the time to adoption. Minimizing time spent in the gateway is key, with approaches like cut-through behaviors, simple FC encapsulation and no TCP. FCoE services databases should be simply mapped to Fibre Channel databases. Routing must also maintain standard Ethernet learning techniques and routing algorithms. And some products are already drawing close to the standard. For example, Woven has dispensed with Layer 3 routing with a 10 GbE modular switching platform containing congestion management technology.

Another vendor that is presently working in this field is Finisar, who has already

Copyright© The TANEJA Group, Inc. 2007. All Rights Reserved

produced CEE network testing products for Emulex.

## Putting It All Together: Converged Ethernet and FCoE

What does a converged 10Gb Ethernet with FCoE look like? Keep in mind that CEE is not a dedicated fabric, but will be capable of running all Ethernet traffic over a lossless layer 2 network. Network components will include enhanced switches, some with integrated FCoE-Fibre Channel forwarding functions and converged network adapter cards.

Enhanced Ethernet Switches will contain congestion management and traffic isolation features based on IEEE's 802.1au Congestion Management and related standards. The standard manages network congestion which will improve Ethernet latency and packet loss. In turn this enables truly isolated traffic lanes in a lossless Ethernet environment, which enables Fibre Channel to be directly mapped to Ethernet without using processing-intensive IP for the transport protocol.

Converged Network Adapter Cards use drivers to present both Fibre Channel functionality and networking/clustering devices to the host server OS. This method preserves Fibre Channel hardware, software, and existing operations. The host OS sees the FC drivers as HBAs to be managed as Fibre Channel end-points, preserving traditional Fibre Channel features such as World Wide Names.

The FCoE-FC forwarding function uses an Enhanced Ethernet switch to send FCoE frames from multiple FCoE end-points to native Fibre Channel devices, and uses a traditional Fibre Channel switch to connect to the Fibre Channel SAN. Unlike the more demanding iSCSI to Fibre Channel gateway, the FCoE to FC conversion is a simple, stateless function that simply adds or strips Ethernet and FCoE headers to or from Fibre Channel frames.

### FCOE BENEFITS

Making converged Ethernet a reality requires tremendous amounts of ongoing work and R&D investment at the front-end, and enterprise investment at the back-end. Is it worth it? Taneja Group believes that it is for some very compelling reasons:

**FCoE Benefit #1. Unified fabric in the data center.** FCoE and CEE together can create a unified data center fabric that meets the reliability, latency and performance requirements for storage and broader data-center connectivity. Converged 10GbE speeds up performance for storage and applications, supports multiple protocols and traffic types, and preserves Fibre Channel frames and services; all while offering Ethernet's ease of management and a simplified interface.

**FCoE Benefit #2: Common interface for Fibre Channel and IP traffic.** Using the same Ethernet interface for both storage and IP traffic leverages common Ethernet knowledge. Meanwhile, FCoE retains familiar Fibre Channel software stacks, OS drivers, or management for Fibre Channel administrators.

Copyright© The TANEJA Group, Inc. 2007. All Rights Reserved

**FCoE Benefit #3: Cost and Power savings.** FCoE leverages Ethernet for the data center by shrinking the number of networking components needed for multiple fabrics and simplifying the management process. In turn this eases the management burden on Fibre Channel and Ethernet administrators alike, freeing them up to pursue high-value IT projects.

**FCoE Benefit #4: Inherits Ethernet technology advancements.** FCoE running on 10GE will allow Fibre Channel to leverage high network speeds. And since CEE is a lossless environment, traditional Ethernet's challenges will not impact Fibre Channel traffic.

## Taneja Group Opinion

Taneja Group is positive about iSCSI SANs in the right environment. Small business, small enterprise, and non-mission critical applications certainly benefit from the strong economies and management simplicity that many iSCSI SANs provide. But we disagree with some prominent iSCSI supporters over an important point: that increased Ethernet performance thanks to 10GbE will make iSCSI adequate to replace Fibre Channel throughout the data center. In this view, FCoE is simply unnecessary because Fibre Channel is unnecessary over CEE.

Taneja Group strongly disagrees. Not to mention existing heavy investment in Fibre Channel, storing mission-critical data is not primarily about speed but about reliability and data integrity. Network-based speeds are not inconsequential, and development continues on faster Fibre Channel. But Fibre Channel's strengths rise from its ability to speedily process and effectively transport data based on its highly efficient four-layer architecture. Fast Ethernet speeds combined with FCoE will enhance this ability, not replace it.

We believe that FCoE and CEE will begin to create a unified fabric for the enterprise data center. Existing fabric types are not going anywhere yet: Ethernet serves normal traffic perfectly well, Fibre Channel is widely used for storage, and InfiniBand is ideal for clustering traffic. If there is a universally converged network fabric coming down the pike, it will be years before it establishes network hegemony. Nevertheless, some vendor-level FCoE products already exist and FCoE and CEE can be productized for IT consumers by the end of 2008 given the recent agreement in T11. But real products that can be put into production are unlikely until 2009.

We grant that Emulex, QLogic, Brocade, EMC, IBM, HP, HDS, Sun and others have a lot invested in Fibre Channel and enjoy excellent margins on their FC products. Each of them has also recognized the role of iSCSI and has released products in the market, but each company also knows that iSCSI cannot – and from a financial perspective, must not -- replace FC entirely. So it is little surprise that they are all supporting FCoE and CEE efforts.

We believe their logic and motivation are appropriate for technical and financial reasons. This is because in spite of all of iSCSI's excellent qualities, it still rides on

Copyright© The TANEJA Group, Inc. 2007. All Rights Reserved

TCP/IP and a lossy layer 2 Ethernet. This requires iSCSI to inevitably make up for the deficiencies of the lower layers. The fact that many of the new iSCSI-based products today are easy to use and provision has more to do with the way they are architected than the fact that they are iSCSI-based. They are good products and will continue to have a place in the IT shops, but their deficiencies relative to FC will remain. That is the reason why FCoE/CEE makes sense to us: it allows the use of the 10GE fabric and yet maintains tight control on congestion, QoS and in-order flow control.

We are bullish on this standard. Some commentators have stated that FCoE/CEE is merely being driven by vendor greed. We believe that the commentators are quite simply wrong, and that this standard holds genuine appeal to IT. Clearly it appeals to the vendors as well. Emulex has taken an aggressive stance towards promoting the effort and driving the development of the standard, and has a lot riding on it. Even QLogic, who seems to have a leg in both camps, will ultimately benefit from this as will all players with strong FC products.

But it is not only the vendors who will benefit. In our view the new standards will ultimately benefit the end-user as well. And in the end, it is the end-user who truly matters.

*NOTICE: The information and product recommendations made by the TANEJA GROUP are based upon public information and sources and may also include personal opinions both of the TANEJA GROUP and others, all of which we believe to be accurate and reliable. However, as market conditions change and not within our control, the information and recommendations are made without warranty of any kind. All product names used and mentioned herein are the trademarks of their respective owners. The TANEJA GROUP, Inc. assumes no responsibility or liability for any damages whatsoever (including incidental, consequential or otherwise), caused by your use of, or reliance upon, the information and recommendations presented herein, nor for any inadvertent errors which may appear in this document.*

Copyright© The TANEJA Group, Inc. 2007. All Rights Reserved